

EpCAM differentially regulates individual and collective migration of human carcinoma cells

Azam Aslemarz

Department of Biology

McGill University, Montreal

October 2021

*A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of
Doctor of Philosophy*

© Azam Aslemarz 2021

Table of Contents

Acknowledgments.....	1
Abstract.....	4
Résumé	6
Contributions	8
Chapter I.....	9
Introduction	10
EpCAM molecular structure and modifications.....	12
Glycosylation.....	14
Proteolytic cleavage.....	15
Phosphorylation.....	16
Interacting partners	16
Cytoplasmic partners	16
Membrane proteins.....	17
Expression patterns and subcellular localization.....	19
EpCAM biological functions	21
EpCAM as a homophilic cell-cell adhesion molecule.....	21
EpCAM as a signaling molecule.....	24
Nuclear signaling by EpCAM	24
PKC inhibition by EpCAM	26
Actomyosin contractility in cell adhesion and cell migration.....	29
Myosin.....	29
Cell-cell adhesion	30
Myosin contractility and cell-cell adhesion	31
Cell- matrix adhesion	32
Myosin contractility and cell-matrix adhesion.....	33
EpCAM as a pro-adhesive and pro-migratory signaling molecule via myosin modulation	34
EpCAM role in cell adhesion in other cell systems	36
Other embryonic models	36
Differentiated cell models.....	39
EpCAM in cell-ECM adhesion.....	42
EpCAM position in cancer invasion.....	43

Cancer invasion	44
EpCAM and EMT	46
EpCAM and circulating tumor cells (CTCs).....	47
EpCAM role in cell migration	49
MCF7 cells	51
Thesis objectives	51
References	52
Figure legends.....	69
Figures.....	72
CHAPTER II	79
Abstract.....	80
Introduction	81
Results.....	84
Material and methods	104
References	112
Figure legends.....	120
Supplementary figure legends	125
Chapter III.....	141
Abstract.....	142
Introduction	142
Results.....	144
Discussion.....	147
Material and methods	150
Acknowledgments.....	152
References	152
Figure legends.....	155
Supplementary figure legends	157
CHAPTER IV	167
Discussion.....	168
Chapter II- Summary of findings	168
Chapter II- Future perspectives	170
Chapter III-Summary of findings	175
Chapter III-Future perspectives	176

References177

Acknowledgments

I would like to first thank my PhD supervisor, Francois Fagotto, mainly for providing me with a safe space to grow scientifically. I sincerely say that the constant trust he had in me, and the fact that he never questioned my potential were truly the reasons that kept me going through this path. I also thank him for making me learn that before anything science is about enjoying it and for making me learn to do what I'm doing the most correct way possible. Beyond work, I thank him and his family for being always caring and supportive.

I then thank Frieder Schock and Claire Brown, my committee members for their helpful feedback, and my co-supervisor Paul Lasko for his great help with many things, especially during the process of thesis submission.

I would then like to thank Sirous Zeinali, the CEO of the biotechnology company in Iran that I worked for before PhD, who supported my decision to pursue PhD and very kindly helped my visa application process to enter Canada. I thank my friends back in Tehran, Saeedeh, Maryam (Sharafi), Nogol, and especially Forouzan for their incredible friendship and their support during my preparation for IELTS and my departure from Iran.

I thank Leily and Pedram for welcoming me at my arrival in Montreal and for helping me with almost everything to start my life there.

I thank Mina, for treating me like an old friend shortly after we met in Montreal and for all the nice memories we created together in such a short time before I moved to Montpellier. I also thank her for the uncountable times that since then she has patiently listened to me talking about every aspect of my life and provided me with great advice.

I thank David for his everyday company that made us go beyond being only lab mates; I enjoyed our short random breaks at any time of the day during the past years and will miss all the chats and the jokes. I'm also happy to have met him as a reasonable person and a great scientist.

I thank all my lab mates, past and present, from Montreal, Leily, Laura (Canty), Katerina, Henry, and Eleyine; from Montpellier, Lina, Guillaume, Laura (Chaptal), Marie (Carta), Rim, Marie (Fagotto), Elise and Laetitia for their conversations and the nice working atmosphere they created in the lab. I'm also grateful for meeting all the incredible colleagues (scientists and staff) at CRBM and I specifically thank the staff of MRI for their help with microscopy and image analysis. I also thank Estelle, our social assistant at work, for her advice with many paper works.

I then would like to thank Nour, for her stunning friendship and for all the quality time we spent together. I also thank Charbel, Maryam (Mojallal), and Lorene for their friendship and for being always there for me during all these years. And Antonella, Himanshu, Toufic, Jesus, Igor, Saeid, and Khaled for their friendship and all the fun moments that I had with them.

And finally, I thank my family for their eternal love and care during every step of my life. I thank my father for his great and constant respect, trust, and love for me. I mention my beloved mother and wish she was still with us to see the end of this great chapter of my life. I then thank my stepmother, that I have always called mom, for her beautiful heart and being the most patient and strong person, I know in my life. I thank my sisters, Ati, Afagh, and Pani (Effat) for sending me their unconditional love and care continuously. I thank my brothers Soroush, Ahmad, Nasir, and my loveliest Amir for being always proud of me and for their endless love and support. I thank my sisters/brothers-in-law and my nieces and nephews for all the greetings and love they have sent me during the past years.

I thank anyone else that has not been mentioned here and has been in any possible way part of this journey.

Abstract

Epithelial Cell Adhesion Molecule, or EpCAM, is a cell surface glycoprotein highly expressed in most human carcinomas, and therefore being used as a major biomarker for cancers of epithelial origin. The strong correlation between EpCAM levels and malignancy suggests that EpCAM may be functionally implicated in metastasis, but current experimental evidence is scarce and contradictory. EpCAM was long considered to be a Ca^{2+} - independent homophilic cell-cell adhesion molecule, however, recent data tend to argue against this function, and rather point to signaling functions. It has been shown in particular to influence cell adhesion and cell migration indirectly, through its ability to downregulate myosin activity. The goal of this thesis was to provide an in-depth characterization of the impact of EpCAM on myosin regulation and on single and collective migratory capacities of MCF7 cells, a human cell line used as a model for pre-metastatic, high EpCAM, breast cancer cells. Spheroids formed by aggregation of MCF7 cells were used to mimic the three-dimensional organization of a solid tumor. We found that EpCAM levels indeed had major effects on the migration and adhesion of MCF7 cells. Interestingly, while single-cell migration was inhibited upon EpCAM depletion, collective migration was on the contrary strongly stimulated. Moreover, EpCAM depletion increased the coherence of spheroids, blocking the occasional detachment of single cells, observed for wild-type MCF7 cells. Thus, high level of EpCAM in carcinoma cells seems to have a dual role, repressing collective migration, while promoting detachment and migration of individual cells. Biophysical measurements together with immunofluorescence of actin and myosin showed that EpCAM knockdown increased cell cortical tension, as well as tension exerted on the matrix. Tension was also higher at cell-cell contacts but was compensated by a strong E-cadherin reinforcement reaction, resulting in effective higher cell-cell adhesion. In addition to EpCAM, most human epithelial tissues, normal and

cancerous, including MCF7 cells, also express Trop2, the only close relative of EpCAM, which was therefore included in this study. We found that Trop2 also acts as a repressor of myosin activity, however, its depletion appeared to have the exact opposite impact of EpCAM on the cellular behavior, for both single cell and collective migration. These opposite phenotypes result from a somewhat subtle difference in the balance of tensions, which can be explained by partially distinct subcellular distributions of the two proteins. Our results provide a coherent picture of the impact of EpCAM and Trop2 expression on the single and collective behavior of carcinoma cells and reveal how these two closely related genes fulfill both similar and antagonistic morphogenetic functions in epithelial tissues.

Résumé

La molécule d'adhésion cellulaire épithéliale (EpCAM, pour Epithelial Cell Adhesion Molecule) est une glycoprotéine membranaire fortement exprimée dans la majorité des carcinomes humains. Elle est utilisée comme biomarqueur majeur des cancers d'origine épithéliale. La forte corrélation entre le niveau d'expression d'EpCAM et la malignité suggère que l'EpCAM peut être impliquée dans les métastases, mais les preuves expérimentales actuelles sont rares et contradictoires. L'EpCAM a longtemps été considérée comme une molécule d'adhésion cellule-cellule indépendante du Ca^{2+} . Des données récentes suggèrent une opposition à cette fonction et indiquent plutôt un rôle de signalisation. Il a notamment été démontré qu'EpCAM influence indirectement l'adhésion et la migration cellulaires, par sa capacité à réguler négativement l'activité de la myosine. L'objectif de cette thèse était de caractériser l'impact de l'EpCAM sur la régulation de la myosine et sur les capacités migratoires uniques et collectives des cellules MCF7 : une lignée cellulaire humaine utilisée comme modèle pour le cancer du sein pré-métastatique à haut niveau d'expression d'EpCAM. Des sphéroïdes formés par agrégation de cellules MCF7 ont été utilisés pour imiter l'organisation tridimensionnelle d'une tumeur solide. Nous avons constaté que les niveaux d'EpCAM avaient des effets majeurs sur la migration et l'adhésion des cellules MCF7. Curieusement, alors que la migration cellulaire unique était inhibée lors de l'épuisement d'EpCAM, la migration collective était fortement stimulée. De plus, la déplétion d'EpCAM a augmenté la cohérence des sphéroïdes, bloquant le détachement occasionnel de cellules individuelles, observé dans des cellules MCF7 de type sauvage. Ainsi, un niveau élevé d'EpCAM dans les cellules cancéreuses semble avoir un double rôle : la répression de la migration collective et la favorisation du détachement et de la migration des cellules uniques. Des mesures biophysiques et l'immunofluorescence de l'actine et de la myosine ont montré que le knockdown d'EpCAM

augmentait les tension corticaux cellulaires, ainsi que la tension exercée sur la matrice. La tension était également plus élevée aux contacts cellule-cellule, mais était compensée par une forte réaction de renforcement de l'E-cadhérine. Ceci entraînait une adhésion cellule-cellule plus forte. En plus d'EpCAM, la plupart des tissus épithéliaux humains normaux et cancéreux, dont les cellules MCF7, expriment également Trop2 : le seul proche parent d'EpCAM, qui a été inclus dans cette étude. Nous avons constaté que Trop2 agit également comme répresseur de l'activité de la myosine. Cependant, sa depletion semble avoir l'effet inverse de celui de la depletion d'EpCAM sur la migration cellulaire unique et collective. Ces phénotypes opposés résultent d'une différence quelque peu subtile dans l'équilibre des tensions, pouvant s'expliquer par des distributions subcellulaires partiellement distinctes des deux protéines. Nos résultats fournissent une image cohérente de l'impact de l'expression d'EpCAM et de Trop2 sur le comportement unique et collectif des cellules de carcinome, et révèlent comment ces deux gènes liés ont des fonctions morphogénétiques à la fois similaires et antagonistes dans les tissus épithéliaux.

Contributions

Chapter I

I reviewed EpCAM literature and wrote this chapter with the guidance of my thesis supervisor, Francois Fagotto.

Chapter II

This chapter is currently in preparation to be submitted to a journal for publication. All experiments were designed and carried out under the guidance of Francois Fagotto. I conducted and analyzed all the spheroid experiments, as well as cell migration assays. I performed and analyzed the geometry-based force inference experiments. Characterization of the phenotypes of individual cells and small groups of cells was carried out by Marie Fagotto and Francois Fagotto. Artur Ruppel and I carried out micropatterning and traction force microscopy experiments under the supervision of Martial Balland. I performed the mixing/sorting experiments, and the analysis was done by the contribution of Marie Fagotto using a custom-made ImageJ plug-in developed by Volker Backer.

Chapter III

This chapter is currently in preparation to be submitted to a journal for publication. I carried out all the experiments under the guidance of Francois Fagotto. Image quantification was done by Francois Fagotto.

Chapter IV

I wrote this chapter with the guidance of Francois Fagotto.

Chapter I

Literature review and thesis objectives

Introduction

Epithelial cell adhesion molecule, EpCAM (also known as DIAR5, EGP-2, EGP314, EGP40, ESA, HNPCC8, KS1/4, KSA, M4S1, MIC18, MK-1, TACSTD1, TROP1) was originally discovered as a tumor antigen of human colorectal carcinoma (Herlyn *et al.*, 1979; Gires *et al.*, 2020). Due to its overexpression in most human carcinomas (Went *et al.*, 2004; Rao *et al.*, 2005), it has been long used as an important cancer biomarker for diagnostic and therapeutic purposes (Baeuerle & Gires, 2007; Keller, Werner & Pantel, 2019). It is a single-pass cell surface glycoprotein and was initially suggested to be a Ca^{2+} - independent homophilic cell-cell adhesion molecule (Balzar *et al.*, 1999; Litvinov *et al.*, 1994). However, firm experimental support for this function is still lacking (Gaber *et al.*, 2018). Moreover, while EpCAM is highly expressed in malignant tumors and its overexpression has been correlated with aggressiveness and poor prognosis, it is considered to be downregulated in invasive cells having undergone epithelial-mesenchymal transition (EMT) (Gires *et al.*, 2020). Therefore, the effect of high EpCAM expression on the behavior of cancer cells is still not fully understood, and whether EpCAM promotes or suppresses cancer invasion remains unclear. In vitro studies have revealed a signaling function that stimulates proliferation, although the actual role of this activity in tumorigenesis remains unclear (see below). In zebrafish and *Xenopus* embryos, EpCAM is expressed from the very early stages, and its experimentally induced loss decreases both cell-cell adhesion and cell motility (Slanchev *et al.*, 2009; Maghzal *et al.*, 2010, 2013). Interestingly, in the *Xenopus* model, our team showed that the positive impact of EpCAM on adhesion and migration is accomplished by acting as a signaling molecule. Both effects result from the ability of EpCAM to moderate contractility of the actomyosin cytoskeleton. Specifically, EpCAM acts as a direct inhibitor of the subclass of the so-called “novel” protein kinase C (nPKC). Inhibition of nPKC impacted a

downstream PKD-Raf-Erk pathway that inhibits non-muscle myosin II (NM II) activity (Maghzal *et al.*, 2013). In embryonic tissues, this inhibition of myosin in turn stimulates protrusive activity and migration, as well as cadherin cell-cell adhesion, consistent with the known principles (Fagotto & Aslemarz, 2020). Adhesion and migration are key parameters involved in cancer invasion (Friedl & Wolf, 2003; Wolf & Friedl, 2006), and thus these newly discovered functions of EpCAM may be highly relevant to understand its role in cancer development. While there have been over the past years scattered reports regarding the effect of EpCAM on human cancer cells in vitro, the studies were rather superficial and led to conflicting conclusions (reviewed in Fagotto & Aslemarz, 2020). This is perhaps not surprising considering the rather complex interplay between actomyosin contractility, and adhesive and migratory properties (Vicente-Manzanares *et al.*, 2009; Collinet & Lecuit, 2013; Pandya, Orgaz & Sanz-Moreno, 2017). The work of this thesis, therefore, has aimed to better define the impact of EpCAM on cell contractility and adhesive and migratory behavior of cancer cells.

EpCAM is only found in vertebrates and does not belong to any known family of membrane proteins (see below). Its sequence is highly conserved throughout all vertebrates. It forms a single gene family in fish and amphibians. Although it is duplicated in tetraploid/allotetraploid species such as zebrafish and *Xenopus laevis*, the sequences of their two EpCAMs are nearly identical. However, a second EpCAM gene called Trop2 exists in the amniote genome. Trop2 (also called as TACSTD2, EGP1; GP50; M1S1; EGP-1; TROP2; GA7331; GA733-1) is known to have appeared by retrotransposition of the EpCAM gene via an mRNA intermediate (Linnenbach *et al.*, 1993). EpCAM and Trop2 are the most common names used for these two genes, and we will keep these names in this thesis. EpCAM and Trop2 share 48% identity and 78% similarity in human. Like human EpCAM (hEpCAM), human Trop2 (hTrop2) is overexpressed in most carcinomas

and its high levels have been correlated with decreased survival of carcinoma patients (Cubas *et al.*, 2009). EpCAM and Trop2 are co-expressed in all human epithelia, except for the intestine, which only expresses EpCAM. The high similarity between hEpCAM and hTrop2 sequences and with *Xenopus* EpCAM (xEpCAM) makes their involvement in PKC-Erk-myosin pathway very likely. This is also supported by the fact that hEpCAM depletion in human cells resulted in elevated actomyosin contractility (Maghzal *et al.*, 2013; Salomon *et al.*, 2017; Barth, Honesty & Riedelkruse, 2018) and in the previous study of the team, hTrop2 was shown to directly bind to PKC. Therefore, Trop2 was included in the work of this thesis as well, however, EpCAM remains the focus of the work.

EpCAM molecular structure and modifications

EpCAM does not structurally belong to any of the four major families of the cell adhesion molecules (CAMs): cadherins, selectins, integrins, and immunoglobulin-like CAMs (Elangbam *et al.*, 1997). While we will focus here on hEpCAM, the EpCAM protein sequence is highly conserved, and the structural features presented here can be extended to other mammals and beyond. The human EpCAM gene (*hEpCAM*) is characterized by 9 coding exons that encode for a protein of 314 amino acids (291 without the signal peptide) with 35kDa molecular weight, which gets glycosylated, yielding a glycoprotein of approximately 40kDa. It is composed of an extracellular domain of 242 residues (residues 24 to 264), a single transmembrane domain (TM) of 23 amino acids (265 to 288), and a particularly short cytoplasmic domain of 26 residues (289 to 314) (Schnell, Cirulli & Giepmans, 2013). In hTrop2, these three regions correspond to respectively residues 27 to 274, 275 to 297, and 298 to 323 (Cubas *et al.*, 2009).

The overall structure of the hEpCAM extracellular domain was solved by Chong and Speicher (Chong & Speicher, 2001). It is composed of three structural domains, starting with an

N-terminal domain with a unique cysteine disulfide bonding pattern, followed by a second cysteine-rich domain, and finally a cysteine-poor C-terminal domain without similarity with any other animal protein (Chong & Speicher, 2001; Pavšič *et al.*, 2014). The second domain is the only one with some structural similarity, mostly based on an evolutionarily conserved disulfide bonding pattern, which is related to a thyroglobulin-type A1 domain (Linnenbach *et al.*, 1989; Molina, 1996; Chong & Speicher, 2001). The three domains are arranged in a triangular fashion where each domain contacts the other two, resulting in a rather compact extracellular domain conformation (Pavšič *et al.*, 2014). Based on the high sequence similarity, hTrop2 extracellular domain is likely to have an identical conformation (Pavšič *et al.*, 2015), which can be readily extended to all EpCAMs throughout the vertebrate class.

The hEpCAM extracellular domain crystallizes as a dimer, and it is well established that hEpCAM is expressed at the plasma membrane as a cis-homodimer (Pavšič *et al.*, 2014; Gaber *et al.*, 2018). Cis-dimerization mainly depends on the large contact interface between the extracellular domains, specifically the thyroglobulin-like and the cysteine poor domains (Pavšič *et al.*, 2014). Homodimerization of hTrop2, although not shown, is expected based on the very similar sequences between EpCAM and Trop2 (Pavšič *et al.*, 2015). Whether EpCAM and Trop2 may form heterodimers has not yet been addressed.

The TM domain of hEpCAM is prominently rich in valine and poor in leucine, which is unusual as leucine is generally the most abundant residue found in TM domains (Mbaye *et al.*, 2019). Furthermore, many of its residues are strongly evolutionarily conserved in both EpCAM and Trop2, and throughout vertebrates (F. Fagotto, unpublished), consistent with its proposed function for their interactions with other membrane components (Ladwein *et al.*, 2005; Nakatsukasa, 2010). Based on molecular dynamics modeling, it has been suggested that the TM

domain may also contribute to cis dimerization (Pavšič *et al.*, 2014). However, key residues of the proposed TM motif happen to be among the few that are not conserved, even in mammals, questioning the accuracy of this prediction. Another TM dimerization motif was proposed for hTrop2 (Pavšič *et al.*, 2015), which again is poorly conserved.

As mentioned above the cytoplasmic C-terminal tail is short, however, it holds distinct parts including a highly conserved juxtamembrane motif that is rich in basic residues and a distal motif that varies among vertebrates but includes three to four glutamic residues positioned at conserved intervals. hTrop2 cytoplasmic domain carries similar features (Pavšič *et al.*, 2015; Fagotto & Aslemarz, 2020). The cytoplasmic domain of both proteins has been shown to be involved in signaling pathways that I will discuss in other sections.

The amino acid alignment of the two proteins is shown in Figure 1 of this chapter and their domain structure in Figure 2.

Glycosylation

The extracellular domain of hEpCAM contains three N-linked glycosylation sites at Asn74, Asn111, and Asn198 residues that are positioned in its second and third motifs (Pavšič *et al.*, 2014). While the second site is the only site that is conserved in all vertebrates, the first site is found only in hEpCAM and the third site is found in most but not all vertebrates, questioning the physiological importance of the two latter sites, although, for example, it has been shown that glycosylation of Asn198 was important for the stability of hEpCAM and its cell surface expression when ectopically expressed in HEK293 cells (Munz *et al.*, 2008). Moreover, it is reported that EpCAM is differentially hyperglycosylated in head and neck carcinoma cells (Pauli *et al.*, 2003). However, this observation has not been reported in other carcinoma cells/cancer types. Four N-

linked glycosylation sites are predicted to exist in hTrop2 extracellular domain (Cubas *et al.*, 2009).

Proteolytic cleavage

Another feature of the EpCAM extracellular domain is holding proteolytic cleavage sites (Pavšič *et al.*, 2014). In addition to the initial cleavage of the signal peptide during translation, a second cleavage occurs between Arg80 and Arg81, located in the thyroglobulin-like domain. The small cleaved fragment (6kDa) remains attached to the rest of the protein via disulfide bonding (Thampoe *et al.*, 1988; Schnell *et al.*, 2013). This cleavage is mediated by a membrane-anchored protease called matriptase, which is widely expressed in epithelial tissues (List, Bugge & Szabo, 2006). hTrop2 has been also shown to be cleaved by this protease (Wu *et al.*, 2020) and the cleavage site is conserved in vertebrates, suggesting that this proteolysis is important for EpCAM/Trop2 function. Interestingly, the matriptase cleavage is not compatible with cis-homodimerization of EpCAM (Pavšič *et al.*, 2014), thus dimerization is possibly protective of the proteolysis. Reciprocally, cleavage of the monomeric EpCAM could prevent dimerization, resulting in degradation of the cleaved EpCAM monomer (Wu *et al.*, 2017, 2020). However, information on the actual rate of the cleavage and its impact on the EpCAM stability is still missing.

Recent studies have detected multiple additional cleavage sites in both EpCAM extracellular and transmembrane domains, eventually leading to the release of the short cytoplasmic tail, thus defining a case of regulated intramembrane proteolysis (RIP) (Maetzel *et al.*, 2009). This process has been detected in various carcinoma cell lines (Maetzel *et al.*, 2009, Schnell *et al.*, 2013; Pavšič *et al.*, 2014). EpCAM RIP is thought to be of functional importance, as the cleaved intracellular domain is involved in nuclear signaling (Maetzel *et al.*, 2009). The

detail of this signaling activity and the extent of RIP contribution to EpCAM biological functions will be further discussed in a separate section. hTrop2 has also been shown to undergo RIP, similarly releasing a soluble fragment acting in nuclear signaling (Stoyanova *et al.*, 2012).

Phosphorylation

The cytoplasmic domain of hEpCAM contains a serine (Ser289) and a tyrosine (Tyr297) (Schnell *et al.*, 2013), conserved in hTrop2 (Ser303, Tyr306), which has also an additional Ser322 (Cubas *et al.*, 2009; Pavšič *et al.*, 2015). There are no reports of hEpCAM phosphorylation. The Ser303 and 322 of hTrop2 can be experimentally phosphorylated by PKC (Basu *et al.*, 1995; Mori *et al.*, 2019), and testing the effect of Ser322 phospho-mimic and phospho-defective mutants suggests effects on interactions with claudins and cell motility (Mori *et al.*, 2019). However, the poor conservation of this residue in Trop2 of various species casts some doubts about the physiological significance of its potential phosphorylation. Although Tyr297 (Tyr306 in hTrop2) is a highly conserved residue in all vertebrates, it is unlikely to be a phosphorylation site, based on the near absence of hits in phospho-proteomic databases (PhosphositePlus) (Fagotto & Aslemarz, 2020). A more likely role of this tyrosine is as part of the pseudosubstrate inhibitory motif, as predicted by molecular modeling and supported by biochemical binding assays (Maghzal *et al.*, 2013).

Interacting partners

Cytoplasmic partners

EpCAM has been shown to interact with different cytoplasmic partners through which it has been associated with different biological functions. Balzar *et al.*, described that the cytoplasmic domain of EpCAM contains two α -actinin binding sites at positions 289 to 296 and 304 to 314

residues. Evidence included in vitro binding to α -actinin of peptides of different parts of the cytoplasmic tail, as well as coprecipitation experiments (Balzar *et al.*, 1998). Although this interaction was never confirmed since then, I will include it below as I discuss the question of EpCAM homophilic adhesion. In addition to α -actinin, the cytoplasmic domain has been also shown to participate in the role of EpCAM as a signaling molecule through interacting with cytoplasmic/nuclear complexes or membrane-associated proteins. The functional significance of these interactions will be discussed in separate sections.

Membrane proteins

Claudins

EpCAM has been shown to interact with tight junction proteins claudin-7 and claudin-1 (Ladwein *et al.*, 2005; Wu *et al.*, 2013) and these interactions are conserved in Trop2 (Nakatsukasa, 2010). The interaction has been suggested to involve a motif in the transmembrane domain (AXXXG) found in both EpCAM and Trop2 (Nubel *et al.*, 2009, McDougall *et al.*, 2015). Claudins are a group of 27 tight junction membrane proteins. Their classical function is in building the seal of the junction, forming permselective barriers, and maintaining membrane polarity. They have also been suggested to serve other biological roles such as contributing to cellular signaling, cell adhesion, and cell migration. Consistently, they have been shown to not only localize to the apical tight junctions but also the basolateral membrane (Hagen, 2017). Interestingly, EpCAM was found to colocalize with the basolateral claudin-7, and its depletion results in a decrease in both total levels of claudin-7 and of its basolateral pool, the remaining protein is restricted and/or displaced to the apical tight junctions (Barth, Honesty & Riedel-kruse, 2018). These effects seemed associated with changes in signaling activity such as levels of phospho-Erk and phospho-myosin and collective migratory behavior of MDCK cells (Barth, Honesty & Riedel-kruse, 2018).

Conversely, claudin-7 depletion decreases EpCAM levels (Tanaka *et al.*, 2015), indicating that they mutually stabilize each other along the basolateral domain. Such mutual relationship is still mysterious, since Barth *et al.*, showed that both total levels and basolateral localization of claudin-7 could be rescued by an EpCAM variant mutated in the above-mentioned TM motif, and which indeed did not bind claudin-7, suggesting that the direct interaction of the two proteins was not required for their crosstalk.

Tetraspanins-enriched microdomains

In fact, EpCAM and claudin-7 also depend on each other to associate with tetraspanins-enriched microdomains (TEMs), where they form a complex with CD44v6 and tetraspanin CO-029, an association that is thought to potentially support tumor progression (Kuhn *et al.*, 2007; Nubel *et al.*, 2009). TEMs are known as signaling platforms composed of membrane lipids and tetraspanin proteins that recruit and associate with a variety of transmembrane and cytoplasmic proteins through which they regulate or alter the function of the recruited proteins and influence different processes such as cell signaling, cell adhesion, and cell migration (Hemler, 2005). Therefore, the direct interaction between EpCAM and claudin-7 may be required for placing/localizing EpCAM in TEMs. It is thus possible that EpCAM function in signaling activity and cellular behavior may depend on its location in these special membrane domains and the cooperation with other interacting components. This view is supported by the observation that the EpCAM TM mutant was not able to rescue the effects of EpCAM loss on phospho-Erk/phospho-myosin levels and cellular migratory phenotype in MDCK cells (Barth, Honesty & Riedel-kruse, 2018). Together, while the effects of EpCAM loss on tight junction dynamics, formation, and stabilization remains obvious in several studies (Wu *et al.*, 2013; Tanaka *et al.*, 2015; Barth, Honesty & Riedel-kruse, 2018), whether these effects are through its direct interaction with

claudins remain unclear. A possible explanation for the mechanism of EpCAM impact on tight junctions will be discussed in another section.

Expression patterns and subcellular localization

Detailed studies of EpCAM expression patterns in different species aren't completely available. However, according to current data EpCAM is known to be generally expressed during early embryonic stages that will be downregulated in most adult tissues and remains restricted to epithelia. In adult human most epithelial tissues are EpCAM positive except stratified squamous epithelium and terminally differentiated epithelial cells, for instance, keratinocytes and hepatocytes have been shown to be EpCAM negative (Trzpis *et al.*, 2007). The level of expression has been reported to vary from very low in gastric epithelium to probably the highest in the colon. As mentioned earlier EpCAM levels are known to increase in the majority of tumors that originate from epithelium including adenocarcinomas and squamous cell carcinomas (Gires *et al.*, 2020).

The murine homolog of hEpCAM is shown to be expressed in non-epithelial cells including thymocytes, T-cells, and antigen-presenting cells (Nelson *et al.*, 1996), which are interestingly examples of EpCAM expression in non-epithelial cells in adult tissues. Spatiotemporal study of EpCAM expression patterns in early stages of gastrulating mice embryos has shown the expression of EpCAM in endodermal cells and its repression in mesodermal cells (Sarrach *et al.*, 2018).

Similarly, EpCAM expression in zebrafish is shown to be restricted to the superficial ectoderm layer of early gastrula embryos that give rise to the epidermis in late gastrula embryos (Slanchev *et al.*, 2009). Interestingly, EpCAM expression during *Xenopus laevis* embryonic development is shown to be ubiquitous in all early gastrulating layers, including ectoderm, mesoderm, and endoderm. Moreover, in late stages, EpCAM was detected in tissues that would

generate future non-epithelial cells like notochord and posterior mesoderm (Maghzal *et al.*, 2010). However, in situ hybridization data shows the epithelial restricted expression of EpCAM in adult *Xenopus* tissues according to Xenbase.

Early embryonic expression pattern of Trop2 has not been explored but at later stages, it has been shown to be restricted to epithelia, however with a different spatiotemporal pattern compared to EpCAM (Mcdougall *et al.*, 2015). In human adult epithelial tissues, EpCAM and Trop2 levels do not seem to always correlate, for instance in the intestine epithelium that EpCAM is highly expressed, Trop2 levels are very low or in esophagus and skin, epithelium low levels of EpCAM are concomitant with high expression of Trop2 (Balzar *et al.*, 1999; GEPIA database). A comparison of the gene expression profile of EpCAM and Trop2 in normal and malignant tissues is shown in Figure 3 of this chapter.

At the subcellular level, EpCAM is generally known to localize at the basolateral membrane of cells in normal epithelial tissues; this has been proposed based on immunohistochemical analysis of EpCAM in normal tissues of the colon in human and rat (Schiechl and Dohr, 1987; Xie *et al.*, 2005). This pattern has been suggested to be different in malignant tissues; while in normal human colon tissue EpCAM is restricted to the basolateral membranes, in the corresponding malignant tissue it is distributed over the whole surface (Xie *et al.*, 2005). In addition, the cytoplasmic tail of EpCAM has been detected in the cytoplasm and the nucleus in cancer tissues, likely as a result of its release through RIP (Ralhan *et al.*, 2010). Recently, EpCAM has been also found to localize in Rab35/EHD1 endosomal compartments of fast recycling pathway (Gaston *et al.*, 2021) through which it has been suggested to be involved in RhoA-dependent rearrangement of actomyosin cytoskeleton required for front-rear polarization of intestinal individual cells.

hTrop2 has been found mostly membrane-localized in normal epithelium, with a mixed membrane and cytoplasmic localization in cancer tissues (Stepan *et al.*, 2011). EpCAM and Trop2 have been both shown to be lost from the plasma membrane due to mutation where they have been associated with congenital tufting enteropathy (CTE) and Gelatinous Drop-like corneal Dystrophy (GDLD) respectively (Sivagnanam *et al.*, 2008; Mcdougall *et al.*, 2015).

It should be mentioned that systematic analyses of EpCAM/Trop2 subcellular localization in different normal and cancer tissues are still missing and these propositions are mainly based on a limited number of immunohistochemical studies.

EpCAM biological functions

There are multiple reports on the biological functions of EpCAM, however, they can be viewed as two main categories, EpCAM acting as a cell-cell adhesion molecule and/or as a signaling molecule.

EpCAM as a homophilic cell-cell adhesion molecule

The first functional study defined EpCAM as a Ca^{2+} - independent homophilic cell-cell adhesion molecule (Litvinov *et al.*, 1994). This conclusion was based on the observation that EpCAM expression in murine fibroblastic L-cells, which lack endogenous cell-cell adhesion molecules, resulted in cell aggregation, and directed homotypic sorting of cells ectopically expressing EpCAM when mixed with the parental cells. However, the kinetics of EpCAM-mediated aggregation of L-cells appeared to be slower than aggregation by E-cadherin, the classical cell-cell adhesion molecule that was used as a standard control in these experiments. Moreover, only L-cells expressing a high level of EpCAM could form tight colonies in Matrigel, and much less efficiently than E-cadherin-expressing cells, suggesting EpCAM was a weak

adhesion molecule (Litvinov *et al.*, 1994). A subsequent study by the same researchers showed that expression of EpCAM in L-cells already stably expressing E-cadherin affected the cell-cell adhesion strength negatively, a puzzling observation that was not consistent with the initial view of EpCAM being a cell-cell adhesion molecule (Litvinov *et al.*, 1997). The authors interpreted these results by suggesting a crosstalk between the two adhesion systems: Cadherin and EpCAM would respectively mediate strong and weaker adhesion, which, in combination, could tune the global strength of cell-cell adhesion in epithelial cells. At the molecular level, the authors proposed that EpCAM expression disturbed the association of E-cadherin with the actin cytoskeleton. Indeed, they observed a decrease in the total levels of α -catenin and its detergent-insoluble fraction, the latter parameter being an indication of a potential decrease in interaction between the E-cadherin- β -catenin- α -catenin complex and the actin cytoskeleton (Litvinov *et al.*, 1997). Besides the detection of a detergent-insoluble fraction, additional evidence supporting an interaction of EpCAM with the cytoskeleton included the observation that both an intact actin cytoskeleton and EpCAM cytoplasmic domain were required for EpCAM localization to the cell-cell boundaries and cell aggregation. Furthermore, α -actinin accumulated at cell-cell contacts upon EpCAM expression and coprecipitated with full-length EpCAM, but not with a mutant lacking the cytoplasmic domain. These latter results led the authors to suggest that α -actinin was responsible to bridge EpCAM to the actin cytoskeleton and that sharing α -actinin could be the limiting factor for the two adhesion systems (Balzar *et al.*, 1998, 1999). However, this model is questionable, in particular, because only 5% to 20% of EpCAM was found in the detergent-insoluble fractions, moreover, its cytoplasmic domain was not required for its detergent insolubility.

Subsequent studies failed to support the role of EpCAM as a genuine adhesion molecule: Analysis of EpCAM-mediated adhesion structures in L-cells by immunogold labeling showed that,

while EpCAM expression seemed to move the two adjacent cell membranes into close proximity, no distinguishable junctions were formed. The same observation was made for endogenous EpCAM in differentiated epithelial cells and epithelial tissues (Balzar et al., 1999). A second electron microscopy study did detect EpCAM-positive finger-like contacts but found no trace of EpCAM clustering (Balzar et al., 2001). These structures were proposed to be EpCAM trans-octamers formed from its cis tetramers as in this study, using biochemical techniques, EpCAM was shown to have a multimeric state. However, multimerization did not require cell-cell adhesion, since it was found in both single and monolayer cultures of EpCAM expressing cells (Balzar et al., 2001). An additional biochemical study showed that full-length EpCAM molecules were found as high-affinity dimers and low-affinity tetramers in solution and colon carcinoma cells, but the extracellular domain was found to be monomeric. The authors proposed the stronger associated dimers to exist laterally and the weaker associated tetramers to account for trans interaction of the molecule in homophilic cell-cell adhesion (Trebak *et al.*, 2001). Moreover, in a recent crystal structure study, the extracellular domain of EpCAM was demonstrated to form cis-dimers and was predicted to interact with another EpCAM cis-dimer on a neighboring cell surface thus forming trans-tetramers (Pavšič *et al.*, 2014). Although in these biochemical studies oligomerization and *cis* dimerization of EpCAM or its extracellular domain has been shown, its capability to mediate *trans* interaction remains to be demonstrated. A recent study that attempted to verify this interaction in fact failed to obtain any evidence supporting the existence of trans-interactions, thus casting strong doubts on the assumed activity of EpCAM as a homophilic adhesion molecule (Gaber *et al.*, 2018). It is however possible that such trans interactions may simply be difficult to capture for technical reasons (weak affinity, specific conditions lost in vitro experiments, etc.). Another possibility would be that EpCAM forms heterotypic interactions with

some yet unknown partners. Yet a different interpretation could be that the adhesive effects attributed to EpCAM may in fact be the indirect result of its signaling activity. Indeed, as presented in the next section, EpCAM impinges on a pathway that modulates myosin activity, which in turn impacts adhesion (Maghzal *et al.*, 2010, 2013). A summary of the different models that help explain the function of EpCAM as a pro-adhesive molecule is shown in Figure 4 of this chapter.

EpCAM as a signaling molecule

Various studies on EpCAM have indicated that EpCAM has signaling activities. These activities impact two main biological processes: Gene regulation and proliferation on one side, and the above-mentioned regulation of myosin on the other side. Currently, the best characterized functions for these two aspects are stimulation of the Wnt- β -catenin signaling (Maetzel *et al.*, 2009), and inhibition of the PKC-PKD-Erk-myosin pathway (Maghzal *et al.*, 2013). In Figure 5 of this chapter diagrams of these two pathways are depicted.

Nuclear signaling by EpCAM

EpCAM has been shown to upregulate c-Myc and cyclin D1, key components that stimulate cell proliferation (Münz *et al.*, 2004; Chaves-Pérez *et al.*, 2013). Investigations on the underlying mechanisms showed that this signaling activity was due to regulated intramembrane proteolysis (RIP) (Maetzel *et al.*, 2009). Indeed, the released short intracellular C-terminus peptide can bind to adaptor four-and-a-half LIM domain protein 2 (FHL2), forming a complex with β -catenin. Translocation of this complex to the nucleus and association with transcription factors of TCF/Lef-1 family stimulates expression of targets of the canonical β -catenin/TFC pathway, such as c-myc, resulting in enhanced cell proliferation in colon carcinoma cell lines. Injecting HEK293 cells expressing intracellular domain of EpCAM to mice induced tumor formation, however, larger tumors were formed when the injected cells were expressing full-length EpCAM (Maetzel *et al.*,

2009). Back to the RIP reaction, the process was shown to involve cleavage and shedding of its extracellular domain by ADAM10/17 and BACE1 proteases and subsequent cleavage of its intracellular domain by γ -secretase.

In addition to this direct signaling function of the intracellular domain, the cleaved extracellular domain was also shown to be active: It appeared to act as a soluble signal that induced additional cleavages of the cell surface EpCAM molecules (Maetzel *et al.*, 2009). Recent work has widened the range of activity of this shedded extracellular domain: it was also reported to activate epidermal growth factor receptor (EGFR), leading to increased activation of the extracellular signal-regulated kinase1/2 (Erk1/2) pathway that promoted cell proliferation and migration (Liang *et al.*, 2018). A similar activation of EGFR-dependent cell proliferation via Erk1/2 and Akt was also reported in head and neck squamous cell carcinomas (Pan *et al.*, 2018). In this latter case, however, it suppressed EGF-mediated EMT, due to repression of EMT driving transcription factors such as Snail and Zeb1 (Pan *et al.*, 2018).

Generally, RIP not only serves as a mechanism for membrane-to-nucleus signaling, but it also can result in the degradation of its membrane protein substrates (Kuhnle *et al.*, 2019). As was mentioned earlier, the EpCAM intracellular fragment that acts as the signal destined to the nucleus is the result of the last cleavage in the RIP process through the activity of γ -secretase. A recent study examined in detail the fate of the various EpCAM fragments produced by RIP (Pan *et al.*, 2018). They found that these fragments were efficiently degraded by the proteasome, significantly faster than the rate of γ -secretase dependent cleavage of EpCAM, which is a rather slow process. The authors concluded that signaling through EpCAM RIP was unlikely under normal conditions, since levels of intact intracellular fragments will be kept very low, but may only occur in specific situations of massive cleavage (Huang *et al.*, 2019). Therefore, whether EpCAM RIP is a general

function of EpCAM, under which conditions, and to which extent it may significantly contribute to the assumed role of EpCAM in cell proliferation remains to be elucidated. As mentioned earlier hTrop2 can also undergo RIP and get involved in β -catenin signaling, causing epithelial hyperplasia and stem cell self-renewal (Stoyanova *et al.*, 2012).

EpCAM has also been found to be able to regulate the Wnt pathway through a different mechanism. During liver development in zebrafish, EpCAM was found to directly bind to Kremen1, resulting in disruption of its interaction with Dickkopf2 (DKK2) and the Wnt receptor Lipoprotein-receptor-related protein 6 (Lrp6). This Kremen1-DKK2-Lrp6 complex normally drives the removal of Lrp6 from the cell surface, thus downregulating the Wnt pathway. EpCAM-Kremen1 interaction resulted in stabilization of Lrp6 and cooperative activation of Wnt2bb signaling for hepatic development (Lu *et al.*, 2013).

PKC inhibition by EpCAM

Inhibition of Protein Kinase C (PKC) is another signaling activity of EpCAM. This function was discovered in studies on the morphogenetic behavior of *Xenopus* embryonic tissues (Maghzal *et al.*, 2010, 2013). It turns out that the EpCAM cytoplasmic tail has a short sequence very similar to the pseudosubstrates contained in PKCs. Thus, EpCAM is able to bind directly and inhibit PKCs, a rather unique mode of signal regulation by a cell surface protein. PKC is a family of serine/threonine kinases that are involved in multiple signaling pathways. Structurally, they contain a C-terminal catalytic kinase domain and an N-terminal regulatory region constituted of different domains and motifs controlling PKC localization and activity. The family consists of three sub-families, a) conventional or classical, b) novel and c) atypical PKCs. This classification is based on the differences in their N-terminal regulatory domain, thus their mode of activation.

All PKCs are initially localized in the cytoplasm in an inactive and autoinhibited state. The first step of activation requires phosphorylation events on the kinase domain, called the priming or maturation step. PKCs remain autoinhibited and they will only be active once this autoinhibition is relieved, which is coupled to their recruitment to membranes, in particular (but not exclusively) the plasma membrane. The autoinhibited state is mediated by an intramolecular interaction between the catalytic domain of the molecule and a motif in its regulatory domain that resembles PKCs consensus substrate sequence, but lacks serine/threonine residues for phosphorylation, thus called pseudosubstrate domain. Various mechanisms cause the release of this interaction, by changes in global PKC conformation that pull away the pseudosubstrate. Classical PKCs are characterized by an activation that requires binding of N-terminal sequences to both diacylglycerol (DAG) and calcium. Interaction with DAG inserted in the inner leaflet of the plasma membrane drive recruitment to the membrane, and “unfolding” pulling the regulatory and catalytic domains apart. Novel PKCs are also activated by DAG, but without calcium requirement. Activation of atypical PKCs does not involve DAG nor calcium, but interaction with other proteins (Newton, 2001; Lipp and Reither, 2011; Kang *et al.*, 2012).

Interestingly, the EpCAM cytoplasmic domain turned out to contain a motif that is strikingly similar to the PKC pseudosubstrate domain (as well as actual consensus substrates), more specifically to substrates/pseudosubstrates of the novel PKCs. Our team demonstrated that this motif was necessary and sufficient for direct binding and specific inhibition of nPKCs (PKC δ and η). EpCAM showed an only weak effect on classical PKC β . This conclusion was based on in vitro GST pull-down and surface plasmon resonance analysis (Maghzal *et al.*, 2013), validated in cell extracts by immunoprecipitation and by functional experiments including the use of specific chemical activators and inhibitors. In particular, embryonic and cellular phenotypes caused by

EpCAM LOF and GOF were effectively rescued or mimicked by respectively chemical inhibitors or activators of nPKCs (Maghzal *et al.*, 2010, 2013).

It should be noted here that PKCs have a widespread expression and they are known to have a variety of cellular substrates with the common feature of being rich in basic amino acids before and/or after the target serine/threonine residues. The consensus phosphorylation sequences are slightly different between PKC isoforms and these differences have been shown to provide PKC isoforms with distinct optimal substrate sequences (Nishikawa *et al.*, 1997). However, the substrate specificity of different PKCs is thought to occur through a complex process beyond the classical enzyme-substrate specificity mechanisms involving their requirements for activation and their spatial regulation via local signaling domains (Lipp and Reither, 2011; Kang *et al.*, 2012). Therefore, individual PKCs are known to participate in distinct cell processes, being central to coordinating spatial signal transduction in cells (Rosse *et al.*, 2010).

Maghzal *et al.*, showed that by inhibiting nPKCs in *Xenopus* embryos, EpCAM blocked the PKC-PKD-Raf-Erk-MLCK (myosin light chain kinase) pathway. MLCK is a major kinase responsible for phosphorylation of the myosin light regulatory chain (MLC), which directly leads to activation of non-muscle myosin II. Through inhibiting this pathway, EpCAM was found to moderate actomyosin cytoskeleton contractility of *Xenopus* embryonic cells. Through this activity EpCAM appeared to have a dual effect on cellular properties: it stimulates protrusive activity and cell motility, and at the same time it favors cell-cell adhesion. Note that myosin regulation by EpCAM was also observed in Caco-2 human intestinal culture cells (Maghzal *et al.*, 2013), which increased stress on the intestinal epithelium, affecting its integrity (Salomon *et al.*, 2017). The discovery of this negative regulation of contractility positioned EpCAM as an important regulator of tissue morphogenetic properties. Note also that the cytoplasmic tail of hTrop2 can also bind

nPKCs (Maghzal *et al.*, 2013), thus this inhibitory function appears to be conserved for the EpCAM/Trop2 family, and throughout vertebrates. I will discuss below the potential consequences on human tumor progression and metastasis. One should mention that this mode of PKC inhibition is probably not unique to EpCAM/Trop2, as our team identified a series of unrelated cell surface proteins harboring a similar juxtamembrane pseudosubstrates motif and directly demonstrated for selected candidates their capacity to bind PKCs (Maghzal *et al.*, 2013).

Finally, a recent study has proposed a different mechanism through which EpCAM may control myosin contractility via interaction with RhoA in a recycling endosomal compartment in Caco-2 cells (Gaston *et al.*, 2021). The conditions of this study were somewhat peculiar, as these epithelial cells were cultured as single cells on a matrix, which is not their physiological situation. Further investigations will be needed to evaluate the relevance of this proposed mechanism.

Actomyosin contractility in cell adhesion and cell migration

In order to address the morphogenetic function of EpCAM, it is important to consider the multiple aspects through which changes in myosin activity may impact cell adhesion and cell migration. I will briefly summarize here basic concepts on myosin, adhesion, and migration, and present the principles of their interrelation.

Myosin

Non- muscle myosin II (NM II) belongs to the actin-binding motor proteins superfamily that among its other functions is its ability to exert tension on the actin filaments making the actomyosin cytoskeleton contractile. This function of NM II is initiated by obtaining a competent conformation to bind to the actin filaments upon phosphorylation of its light chain domain (MLC) followed by producing energy through ATPase activity of its heavy chain (MHC). The

phosphorylation of MLC, thus the activity of NM II, is mediated by different kinases mainly myosin light chain kinase, MLCK, and Rho-associated protein kinase, ROCK. MLC is dephosphorylated through the action of MLC phosphatase, MLCP, making myosin activity reversible (Vicente-Manzanares *et al.*, 2009).

Cell-cell adhesion

Cell adhesion is a major property of animal cells that allows them to assemble into tissues. Cell-cell adhesive structures have complex functions, as besides mediating physical adhesion, allowing the formation of tissues and their homeostasis, they are the source of inputs for signaling pathways, impacting various other cellular processes such as cell proliferation and cell migration. Classical cadherins are the major family of cell-cell adhesion molecules, responsible for most cell-cell adhesion in animal cells. The formation of a cell-cell adhesion relies upon actin polymerization mediated protrusive activity to bring two cells into contact with each other. However, expansion and stability of the contacts require further steps; cadherins anchor to the actin cytoskeleton indirectly through cytoplasmic partner proteins including α -catenin and β -catenin which results in recruiting more cadherin molecules to the sites of contacts forming clusters of cadherin-based adhesion complexes. Subsequently, the interaction between the adhesion complexes and the actin cytoskeleton results in remodeling of the linked cytoskeleton from branched state to linear state that also becomes contractile through the activity of myosin which is a critical step in the maturation of cell-cell adhesion by supporting the contacts to resist stress and remain stable (Harris & Tepass, 2010; Collinet & Lecuit, 2013).

Myosin contractility and cell-cell adhesion

It is well known that the activity of myosin is tightly regulated in cells as controlled levels of actomyosin contractility is highly important in the proper driving of tension-required processes such as cell adhesion. For instance, it has been shown that the elevated uncontrolled contractility of the actin network can disrupt adhesive structures (Zandy, Playford and Pendergast, 2007; Maghzal *et al.*, 2013).

More clearly, the importance of myosin activity regulation in cell-cell adhesion can be better explained if adhesion is considered between two single cells that are completely isolated. As explained in detail in our recent review, in this state cells adopt a spherical shape resisting their cytoplasmic hydrostatic pressure through their cell cortex (Fagotto & Aslemarz, 2020) which is a dynamic and contractile network of actin filaments bound to the inner side of the plasma membrane (Maître & Heisenberg, 2011; Chugh & Paluch, 2018). Cortex function is mainly controlling cell shape, which is defined by the mechanics of the cortex that is reflected as the contractility of the actomyosin network at the cell cortex, called cortical tension or cortical contractility, and the physical interaction of the cell with its environment (Chalut & Paluch, 2016). Cortical tension acts as a barrier to the expansion of the nascent adhesion formed between the two cells, as a decrease in cortical tension at the cell-cell interface has been shown to control the expansion of an adhesion (Maître *et al.*, 2012, 2015). Cell adhesion molecules (CAMs) subsequently recruit regulators of the actomyosin cytoskeleton to downregulate the cortical contractility to be able to expand the adhesion. Decreasing the cortical tension leading to the expansion of the cell-cell contact continues until eventually there is equilibrium between the tension at the free surface of the cells and the tension at the cell-cell contact with the latter being required for the maturation and stability of the adhesion as explained earlier in this section. Therefore, while cortical contractility acts

antagonistically to the adhesion during early cell-cell adhesion expansion, it is an essential part of the mature and stable contacts, showing the importance of its regulation throughout the cell-cell contact establishment process; for instance, if myosin contractility is very low it can be expected that adhesion structures may form but not be capable of resisting stress, on the other hand, if it is too high it can lead to very stable adhesive structures or on the contrary it can disrupt the adhesion complexes (Chalut & Paluch, 2016; Chugh & Paluch, 2018; Fagotto & Aslemarz, 2020).

Cell- matrix adhesion

Interaction with the extracellular matrix (ECM) serves as not only anchorage to the extracellular environment that facilitates attachment and cell migration but also as a signaling platform that enables cells to sense the biochemical and physical properties of their surroundings and respond to them. Integrin heterodimers are the prominent mediator of this adhesion by being receptors for matrix, constituting proteins such as collagen as their ligands. Engagement of integrins with their ligands initiates small nascent adhesions that will subsequently cluster into larger structures by recruiting other integrin molecules and partner proteins to the sites of the adhesions forming multiprotein cell-matrix adhesion structures called the adhesion complex. Like in cell-cell adhesion, linkage to the actomyosin cytoskeleton through cytoplasmic adaptor proteins (here, such as talin and vinculin) is essential in the maturation of the adhesion complexes (Chastney, Conway & Ivaska, 2021).

Based on their size and stage of their maturation, the cell-matrix adhesion complexes form a continuum of different types with the most mature and largest of them called focal adhesions. Integrins involvement in cell-matrix adhesions are dynamically intra- and extracellularly regulated (Lock, Wehrle-Haller & Strömblad, 2008; Chastney, Conway & Ivaska, 2021).

Myosin contractility and cell-matrix adhesion in cell migration

Cell-matrix adhesion and actomyosin cytoskeleton rearrangement (controlled by the activity of Rho family of GTPases) are interlinked processes and both together impact cellular migratory behavior. At the leading edge of a migrating cell, cellular protrusions are formed as a result of actin polymerization. This step is mainly mediated by the activity of Cdc42/Rac1 GTPases. Nascent cell-matrix adhesions form close to the cellular protrusions, stabilizing them and connecting the ECM to the actin cytoskeleton. This also leads to further Cdc42/Rac1 signaling that reinforces actin polymerization. Nascent adhesions disassemble or grow into larger/mature adhesions due to the contractility of the actomyosin cytoskeleton that is also coupled with exerting traction forces on the ECM. This leads to RhoA signaling and sustaining the cellular tension. On the other hand, disassembly of adhesions at the cell rear is driven by high myosin contraction (mediated by the activity of RhoA GTPase) that lets the cell body retract and the migrating cell move forward (Parsons, Horwitz & Schwartz, 2010; Goicoechea, Awadia & Garcia-Mata, 2014). Together, the adhesion assembly, maturation, and disassembly are tightly linked to actin polymerization and myosin contraction which themselves are regulated by the activity of Rho GTPases. The crosstalk of all, in turn, regulate cell migration. In an individual cell, for instance, migration can directly depend on the strength of its adhesion to ECM; in other words, efficient/fast migration occurs when the adhesions are not too weak or too strong (Schwartz & Horwitz, 2006). This relation has been further modeled by Gupton and Waterman-Storer that the rate of cell-matrix adhesion assembly and disassembly depends on myosin contractility (Gupton and Waterman-Storer, 2006). One could expect that balance between contractile forces and adhesion changes when each of adhesion strength or cellular contractility conditions are suboptimal resulting in different cellular migratory behavior (Schwartz & Horwitz, 2006; Fagotto & Aslemar, 2020).

The same principles of migration and the interplay between adhesion and actomyosin contractility apply to the migration of cells as groups (collective migration), however with more complex regulations. Collective migration includes adhesion to both the ECM and the neighboring cells. Differential distribution of actomyosin contractility between the front/leader cells and the follower cells and its lower levels at the cell-cell contact allow coherent/coordinated collective migration (Pandya, Orgaz & Sanz-Moreno, 2017).

Together, regulation of myosin contractility during cell-cell adhesion or cell migration is an important aspect of proper cell functioning and its misregulation has been linked to pathological conditions including cancer (Chalut & Paluch, 2016; Chugh & Paluch, 2018).

EpCAM as a pro-adhesive and pro-migratory signaling molecule via myosin modulation

As mentioned above EpCAM appeared to act as a regulator of myosin contractility through its signaling activity and by inhibiting nPKCs. This novel function of EpCAM was found in studies on morphogenetic behavior of *Xenopus* embryonic tissues using EpCAM LOF and GOF approaches (Maghzal *et al.*, 2010, 2013). Expectedly, the outcome of this function was well characterized to impact adhesive and migratory behavior of the gastrulating tissues. In detail, the *Xenopus* orthologue of EpCAM was identified in a gain of function screen for gene products that could induce mixing of ectoderm and mesoderm during gastrulation. EpCAM overexpression also stimulated cell migration inside tissues. EpCAM depletion using antisense morpholino oligonucleotides perturbed ectoderm epiboly during gastrulation, a morphogenetic process that requires cell intercalation movements. All these effects were consistent with EpCAM favoring “intercellular migration”, i.e., migration of cells relative to each other. This effect was very interesting since intercellular migration is a crucial mechanism involved in morphogenesis as well

as during other processes of collective migration, including during cancer invasion. At the same time, however, both gain and loss-of-function data showed that EpCAM positively regulated cadherin levels and cell-cell adhesion. The effect of EpCAM depletion was particularly striking after gastrulation, as cells rounded up, cadherin was internalized and degraded, and tissues were disaggregated. At the molecular level, as mentioned above, signaling downstream of EpCAM turned out to involve inhibition of nPKCs and, as a key downstream target, myosin activity; the short cytoplasmic tail of EpCAM was shown to contain a PKC pseudosubstrate domain, which bound and specifically inhibited members of the nPKCs sub-family, through which it blocked the PKC-PKD-Raf-Erk-MLCK pathway, resulting in lower MLC phosphorylation.

Negative regulation of cortical actomyosin contractility accounted for all EpCAM-related phenotypes; it explained why gain/loss of EpCAM expression increased or decreased, respectively, the ability of cells to emit protrusions and to move, thus accounting for epiboly and tissue mixing phenotypes; it also explained the positive effect on cell adhesion; high cortex tension was found to be the cause toward cell rounding and thus against cell adhesion. This was strikingly demonstrated by the fact that treatment with the myosin inhibitor Blebbistatin fully rescued all phenotypes in EpCAM depleted embryos, including cell adhesion, cadherin levels, tissue integrity, and overall normal embryonic development.

So far, experiments in *Xenopus* embryos indicated that at least in these tissues, the major effect of inhibition of myosin activity by EpCAM was probably to moderate myosin contractility and ‘soften’ the cells, thus favoring both adhesion and motility. Whether EpCAM preferentially targeted a particular pool of myosin, or whether it simply insured the right general level of myosin activity that was sufficient to maintain adhesion without producing a tense/stiff cortex was not investigated.

EpCAM role in cell adhesion in other cell systems

I will discuss here the still open question of the effect of EpCAM on cell-cell adhesion. As mentioned above, the status of EpCAM as a bona fide cell adhesion molecule remains unsettled, and the current consensus is rather against such activity. However, there is clear evidence that EpCAM has a positive impact on cell-cell adhesiveness and tissue integrity by favoring cadherin-mediated cell-cell adhesion (Slanchev *et al.*, 2009; Maghzal *et al.*, 2013; Salomon *et al.*, 2017), and/or by influencing tight junction formation and dynamics (Lei *et al.*, 2012; Wu *et al.*, 2013; Salomon *et al.*, 2017). Although the molecular basis of these positive impacts has not been fully explored, the capacity of EpCAM to moderate actomyosin contractility via nPKC inhibition seems to explain most of these observed effects on adhesion and tissue organization, both in other embryonic models and in adult tissues, that I will discuss them in the following parts of this section. Additionally, studies on EpCAM association with tight junctions and its role in cell-ECM adhesion will be included in this section as separate parts.

Other embryonic models

In addition to the studies in *Xenopus* embryos, EpCAM has been shown to be important in ensuring normal epidermal development during zebrafish gastrulation and epithelial integrity at the later stages using EpCAM mutants (Slanchev *et al.*, 2009). Like the EpCAM LOF phenotype in *Xenopus* embryos, epiboly morphogenetic movements were delayed in gastrulating EpCAM mutant zebrafish embryos. Epiboly is a predominant morphogenetic movement during gastrulation that includes expansion and thinning of the ectoderm tissue to cover the other embryonic tissues. In zebrafish, it involves cell stretching of the superficial layer of ectoderm tissue called the enveloping layer (EVL), and the rearrangement of the deep layer cells through radial intercalation.

Although EpCAM is exclusively expressed in EVL cells, the mutant embryos showed moderate defects in EVL and deep layer epiboly both. The impaired epiboly in EVL was shown to be layer autonomous whereas the deep cell layer defects were concluded to be a consequence of the EVL malfunction. At the cellular level, the EpCAM mutant EVL cells showed reduced basal protrusive activity that was suggested to be important as part of the underlying mechanisms of EVL epiboly. Additionally, E-cadherin, α and β catenin levels were dropped in these cells and the tight junction complexes showed increased levels and were basally extended, that collectively led the authors to propose EpCAM to be important in controlling cell-cell adhesiveness by promoting basolateral membrane localization of cadherin-catenin complex and apical positioning of tight junctions that would allow normal thinning of EVL and epiboly process (Slanchev *et al.*, 2009). However, re-expression of E-cadherin or depletion of tight junction components did not rescue the impaired epiboly, suggesting the effects to be more complex.

Combined and partial depletion of E-cadherin using antisense morpholinos enhanced the EVL epiboly defects and it resulted in the disaggregation of this tissue in the case of complete depletion of E-cadherin. These results and the assumed putative adhesive function of EpCAM directed the authors to conclude that EpCAM functions as a partner of E-cadherin to control cell-cell adhesiveness. However, it should be noted here that generally, the *Xenopus* phenotypes upon EpCAM LOF appeared stronger than in zebrafish EpCAM mutants; this could most likely be due to the existence of another EpCAM gene in zebrafish called pan-epithelial glycoprotein that was not included in the study. Therefore, the lack of more severe effects on epiboly and presence of E-cadherin even though decreased, could have been due to redundant effects of the second EpCAM gene and thus subtler effects on cadherins, as in *Xenopus* studies the complete depletion of EpCAM was against the stability of cadherins and the tissue integrity per se without combined

partial or complete knockdown of cadherins, that showed the stability of cell-cell adhesions to be downstream EpCAM/myosin signaling rather than having a redundancy relation between EpCAM and cadherin.

This is also suggested based on the high similarity of the EpCAM LOF phenotypes in both species at the tissue and cellular levels including delayed epiboly and reduced protrusive activity of cells that could be a result of high cortical actomyosin tension leading to less motility of cells, as well as decreased cadherin levels that again could be because of instability of adhesive structures under high myosin generated tensions. Moreover, in EpCAM mutant zebrafish embryos constriction of the marginal EVL cells, that require myosin activity, happened at a different position and closer to the animal pole and led to extrusion of the vegetal pole that does not occur in normal situations. These observations could also be interpreted as a result of high myosin activity. In addition, cell protrusive activity requires rearrangement of the cortical actin cytoskeleton, and its reduction in EpCAM mutants could again have been a result of high myosin activity. However, cortical tension and myosin levels of the cells/tissues were not investigated in the zebrafish study, thus these assumptions would require further experiments, perhaps similar to experiments that have been done in *Xenopus* embryos, for instance, blebbistatin treatment of the EpCAM mutant gastrulating embryos can be a very direct way to come up with more clear conclusions in comparing both studies.

Nevertheless, in both cases, the epiboly process eventually recovered, and gastrulation was completed. However, post-gastrula embryos lacking EpCAM showed defects in the organization of the epidermis, suggesting that EpCAM is required for epithelial integrity (Slanchev *et al.*, 2009; Maghzal *et al.*, 2013).

In mice embryo models, there have been different EpCAM knockout reports including embryonic lethal due to placenta defects (Nagao *et al.*, 2009) or normal embryonic development but followed by death shortly after birth (Guerra *et al.*, 2012; Lei *et al.*, 2012). The cause for the latter case has been reported to be intestinal defects. It should be mentioned here that the normal development of EpCAM knockout embryos might be because of redundancy and/or compensation by Trop2, the second gene of EpCAM family in amniotes, which has not been included in these studies. Knowing that the intestine expresses a high level of EpCAM and a very low level of Trop2 could explain the intestinal defects and the normal phenotype of the other tissues. Consistently, Trop2 knockout mice have been reported to be viable and not showing any obvious developmental or physiological phenotype, although mice are more susceptible to developing cancer (Wang *et al.*, 2011).

Differentiated cell models

Since EpCAM is linked to pathological conditions, several studies have attempted to study its role in the relevant contexts using differentiated cell models. As mentioned earlier EpCAM is overexpressed in most human carcinomas, thus different cancer cell lines have been used in EpCAM studies. In addition to its link to cancer, EpCAM has been found to be associated with congenital tufting enteropathy (CTE), a rare intestinal disorder that is characterized by villus atrophy and crypt hyperplasia (tufting) of the small intestine leading to diarrhea. In this case, the pathological condition results from EpCAM loss due to mutations (Sivagnanam *et al.*, 2008) that is suggested to be related to the lack of its role in ensuring epithelial development and integrity (Sivagnanam *et al.*, 2008; Salomon *et al.*, 2017). Accordingly, I have summarized studies on cell adhesion in both cases below.

As for the cancer cells, it should be noted that most EpCAM studies in cancer cell models have been focused on its effects on cell growth and invasion (Martowicz *et al.*, 2012, 2013), however, blocking EpCAM activity using an antibody against its extracellular domain was shown to have a negative effect on aggregating behavior of ovary and breast carcinoma cells (Litvinov *et al.*, 1994). This is probably in line with the general assumption of EpCAM mediating cell-cell adhesion, however, as explained earlier this function of EpCAM is still questionable. One possible explanation could be again the negative regulation of myosin by EpCAM found in *Xenopus* model; although this effect was through EpCAM intracellular domain, its extracellular domain was still suggested to be important in inducing the cellular phenotypes; an EpCAM mutant lacking the extracellular domain, while able to inhibit PKC activity, stimulate cell motility and tissue mixing, and rescue epiboly, was unable to prevent the loss of cell-cell adhesion and tissue integrity after gastrulation (Maghzal *et al.*, 2010, 2013). Nevertheless, the role of EpCAM in cell-cell adhesion in cancer cells has not been adequately explored and it needs further studies.

Regarding CTE, separate studies of EpCAM or tight junction protein claudin-7 knockout mice have reported intestinal defects that share similar features with the CTE phenotype (Ding, 2012; Guerra *et al.*, 2012; Lei *et al.*, 2012). A link between the two molecules was found in a study that used colonic cell lines, T84 and Caco-2, showing that EpCAM depletion results in decreased levels of claudins, between which claudin-7 was the most affected (Wu *et al.*, 2013). As already discussed earlier, this study showed that EpCAM and claudin-7 interact directly, and they localize to the lateral cell-cell intercellular sites distinctly from the apical tight junctions. They reported that in addition to the effects on the levels, the absence of EpCAM led to the redistribution of the remaining claudin-7 from the lateral sites to the apical sites, suggesting that the association between EpCAM and claudin-7 is required for proper subcellular localization and stabilization of

claudin-7 that could be important in intestinal epithelium integrity and function (Wu *et al.*, 2013). However, in another study, their direct interaction was shown to not be required the claudin-7 basolateral distribution (Barth, Honesty & Riedel-kruse, 2018). Furthermore, although, Lei et al., showed the impaired barrier function of EpCAM knock out mice to be concomitant with alterations in apical tight junctions (Lei *et al.*, 2012), another EpCAM knockout mice study showed the effect to be related to the dysregulation of E-cadherin and β -catenin localization and function (Guerra *et al.*, 2012). Moreover, while claudin-7 deficient mice showed defects in intestinal architecture, at the cellular levels the phenotype did not involve disruption of apical tight junctions proposing that the participation of claudin-7 in the regulation of epithelial integrity was not related to effects on tight junctions (Ding et al., 2012). Therefore, the role of EpCAM and/or its claudin-7 association in the cellular mechanisms causing intestinal epithelium defects found in CTE and the molecular mechanism of their effects await further investigations.

Recently Salomon et al., focused on the cellular consequences of EpCAM loss in the intestinal cell line, Caco-2, and CTE patients derived tissues (Salomon *et al.*, 2017) showing that depletion of EpCAM caused E-cadherin adhesion defects at bicellular lateral membranes and affected tight junction components that led to an unusual tight junction positioning at tricellular contacts in epithelial monolayers, therefore impacting both adherence and tight junctions. Interestingly, consistent with EpCAM effects on cell adhesion through regulation of actomyosin contractility in *Xenopus* embryos, increased myosin contractility at tricellular junctions was suggested to account for the cell-cell adhesion defects and apical domain displacement in this study. The authors concluded that the absence of EpCAM could lead to loss of actomyosin network homeostasis, resulting in epithelial dysplasia (Salomon *et al.*, 2017).

Although the molecular link between EpCAM depletion and increased myosin contractility was not investigated in this study, the identified molecular pathway of EpCAM function in *Xenopus* embryos can be readily applied to the LOF phenotypes here in differentiated tissues and in the case of CTE. This is also suggested based on our data on Caco-2 cells that were used in parallel with *Xenopus* embryos as they also showed upregulation in PKC-Erk-Myosin pathway upon EpCAM depletion (Maghzal *et al.*, 2013). It should be added here that the study in MDCK cells showed that EpCAM depletion similarly increased Erk-Myosin activity, however, they suggested the interaction between EpCAM and claudin-7 to be required for this effect, although not necessary for claudin-7 basolateral localization (Barth & Kim, 2018). Together, the role of EpCAM in cell and molecular mechanisms of CTE seems to be complex and requires additional studies.

EpCAM in cell-ECM adhesion

Besides the effect on cell-cell adhesion, EpCAM has been reported to impact the adhesion to the ECM as well. This role is suggested for EpCAM based on two different observations including its association with integrins, the ECM proteins receptors which their interaction mediates cell-ECM adhesion or the consequences of its knockdown/knockout on cell-ECM adhesion and cellular behaviors. Examples of the former are EpCAM association with tetraspanin complexes that contain $\alpha 3$ or $\alpha 3\beta 1$ integrins (Schmidt *et al.*, 2004; Claas *et al.*, 2005) or its association with $\beta 1$ integrin revealed by the co-immunoprecipitation analysis in tumor cell lines. This study suggested EpCAM and $\beta 1$ integrin association plays a role in cell adhesion and cell migration by impacting on FAK/ERK signaling pathway (Yang *et al.*, 2020). hTrop2 has been also reported to associate with $\alpha 5\beta 1$ integrins through which it modulated their localization from focal

adhesions to the leading edges and promoted prostate cancer cell migration (Trerotola *et al.*, 2013). Further characterization of EpCAM/Trop2 association with integrins is still required.

As for the latter observations, a recent study has provided detail on EpCAM role in cellular behavior on collagen ECM. This study has shown that in intestinal Caco-2 cells, EpCAM was required for stress fiber maturation process and subsequently proper localization of the focal adhesions during single-cell spreading on the ECM. Interestingly the underlying mechanism was found to be again via EpCAM contractility modulating function through which it influenced actomyosin cytoskeleton rearrangement, therefore spreading and front-rear polarization of the cells. The authors showed that by controlling active RhoA remodeling through Rab35/EHD1 fast recycling pathway in the cell cortex, EpCAM ensured the required actomyosin cytoskeleton rearrangements for stress fiber formation and cell shape changes during cell spreading that were absent upon EpCAM knockdown resulting in a nonpolarized cell phenotype (Gaston *et al.*, 2021).

It should be stated that despite all the data on EpCAM having a pro-/or anti-adhesive effect, there are reports on its dispensability in cell-cell and cell-ECM adhesion (Tsaktanis *et al.*, 2015; Hsu *et al.*, 2016). However, most available data are supporting the pro-adhesive function (Fagotto and Aslemarz, 2020) which based on its unclear state as a direct cell-cell adhesion molecule and the implications of its indirect role in controlling cell-cell and cell-ECM adhesion through regulation of actomyosin contractility, it can be concluded that EpCAM adhesive function is most likely mediated by its signaling activity and modulation of myosin contractility.

EpCAM position in cancer invasion

The fact that EpCAM is highly expressed in most human carcinomas and has been correlated with tumor progression has led to the obvious question of its actual role in tumor

invasion and metastasis. The question remains unresolved, despite multiple studies on its potential link with the epithelial to mesenchymal transition (EMT), its expression in circulating tumor cells (CTCs), and its potential to promote cell motility/migration. I will summarize the current knowledge in this section after providing a brief general background on cancer progression/invasion.

Cancer invasion

The formation of metastases from cells disseminating from the original tumor represents the endpoint of cancer, as it generally leads to death (Chambers et al., 2002). To reach this stage, important steps must be undergone by tumor cells, which must first detach from the primary tumor, pass through interstitial tissue, enter into the lymphatic or blood circulation system, which will spread them far away from their origin. They must then re-exit circulation, colonize secondary distant sites, and re-form secondary tumors (Chambers et al., 2002). This cycle can continue, as cells can further disseminate from secondary tumors for another round of metastasis. At the cellular and molecular level, metastasis results from a complex set of alterations in multiple processes, including cell adhesion and cell migration (Bogenrieder & Herlyn, 2003). Such changes are particularly dramatic in carcinomas, which are tumors that originate from epithelial tissues. In the classical view of metastasis in carcinomas, the migratory/invasive phenotype is characterized by the phenomenon of EMT (Thiery & Sleeman, 2006). EMT is a switch in cell differentiation, in which cells lose their epithelial characteristics and gain mesenchymal properties. The process appears mediated by non-transcriptional and transcriptional changes, and it involves the participation of various signaling pathways, in particular, signaling by the family of transforming growth factor- β (TGF- β) (Lamouille et al., 2014). Note that while metastasis is a pathological case where EMT is “harmful” for the organism, it is also known to be important in normal conditions,

and is implicated in multiple key steps of embryonic development, such as gastrulation or neural crest migration, (Thiery & Sleeman, 2006).

Over the past years, the view of EMT in cancer has been redefined. It is not viewed anymore as a binary process, but more as a continuum, with intermediate phenotypes, known as partial or hybrid EMT states. Cells displaying such intermediate states have been identified, and there is increasing evidence that they are associated with increased invasion and migration capacity (Cayrefourcq *et al.*, 2015; Pastushenko & Blanpain, 2019). In addition, partial EMT has been linked to metastasis by computational modeling (Puram *et al.*, 2017). Note in this context that in carcinomas, invasive cells must be able to form again tumors with epithelial characteristics (including strong cell-cell adhesion). While in the past this was thought to occur through the mirror process of mesenchymal to epithelial transition (MET), the existence of invasive cells that have preserved part of their epithelial nature has opened the possibility that the formation of secondary tumors may require a subtler reversal than classical MET. It has been even proposed that, at least in some cases, groups of cells may escape the primary tumor without undergoing EMT and migrate collectively as a compact group of epithelial cells (Friedl *et al.*, 2012).

It is important to emphasize the differences between motility, migration, and invasion. Motility is the ensemble of intrinsic dynamic properties that allow a cell to move, either on a matrix substrate or among other cells. It involves for instance the capacity to emit protrusions, such as lamellipodia or blebs, as well as that of modulating cell-matrix and/or cell-cell adhesion in order to pull on adhesive contacts, but also remove existing contacts and establish new ones. Migration defined the actual capacity of cells (or groups of cells in the case of so-called collective migration) to change their position, driven by cell motility. Invasion is defined as the ability of cells to pass through tissue barriers (including interstitial tissue, mostly made of extracellular matrix, but also

the endothelial layer of blood vessels). Therefore, motility and migration are prerequisites for invasion, along with other requirements, such as degradation and remodeling of the ECM by proteases. Cancer cells use different types of cells migration, driven by different motility mechanisms. Individual cell migration can be of the classical “mesenchymal” type or ameboid. Cells can also migrate collectively as cell sheets, strands, or cell clusters. All these types are thought to be used during the invasion. It is even suggested that metastatic cells can switch between different modes of motility while invading other tissues, adopting the mode most adapted to a particular environment. Note that despite this plasticity, it is probable that a specific mode may be predominant for a specific type of cancer. While this remains a very open issue, it has been proposed that collective migration is believed may be the major mode in carcinoma invasion. (Friedl & Wolf, 2003; Wolf & Friedl, 2006; Friedl & Gilmour, 2009).

EpCAM and EMT

The first links between EpCAM expression and EMT were the immunoreactivity of colon cancer cell lines in vitro, and the fact that once introduced in mice, the resulting large secondary tumors were EpCAM positive, but not small metastases. This study suggested that EpCAM was transiently downregulated as a result of EMT during migration to the sites of metastasis (Jojović et al., 1998). EpCAM downregulation during EMT has been since then corroborated by in vitro studies, where EMT was artificially induced, as well as by the characterization of circulating tumor cells (CTCs) (van der Gun *et al.*, 2010; Gires *et al.*, 2020). For instance, extracellular signal-regulated kinase (ERK), a key regulator of EMT, has been shown to repress EpCAM transcription by directly binding to its promoter and indirectly by activation of Zeb1, one of the EMT key transcription factors. The negative regulation of EpCAM expression through activated ERK was shown to be associated with increased migration and invasion. Interestingly, this study also showed

that an increased level of EpCAM resulted in decreased ERK activity, suggesting a double-negative feedback loop (Sankpal *et al.*, 2017). Consistent with these data another study has reported that EpCAM affected ERK activity and EMT negatively through EGFR. This effect relies on the above-mentioned activity of the cleaved extracellular domain of EpCAM to act via EGFR, while competing with epidermal growth factor (EGF) (Pan *et al.*, 2018). In addition, EpCAM expression has been found to be directly repressed by Zeb1, both during Zebrafish gastrulation using Zeb1 gain and loss of function approaches, as well as in breast and pancreatic cancer cell lines (Vannier *et al.*, 2013). Based on these data, EpCAM would be a functional antagonist of EMT, and at the same time one of the targets that gets repressed during EMT. Yet, this model may not be generally applicable, as it is contradicted by other observations, which support a positive role of EpCAM in promoting EMT. For instance, EpCAM was shown to associate with breast cancer invasion through modulation of the JNK/AP-1 signaling pathway (Sankpal *et al.*, 2011). EpCAM has also been shown to be required for EMT induction by TGF- β in breast cancer cell line, MCF7, as its expression levels increased upon TGF- β treatment and its knockdown inhibited TGF- β mediated EMT (Gao *et al.*, 2014). Moreover, EpCAM has been reported to positively regulate EMT through AKT/mTOR pathway in nasopharyngeal carcinoma cells (Wang *et al.*, 2018). Therefore, both roles of suppressing and promoting EMT have been documented for EpCAM which seems to be the same in the case of Trop2 (Remšík *et al.*, 2018; Zhao *et al.*, 2019).

EpCAM and circulating tumor cells (CTCs)

An important approach to understanding metastasis under more physiological conditions is the characterization of circulating tumor cells, CTCs. Although EpCAM has been frequently used as the base of CTCs capturing approaches, other methods that are independent of EpCAM have reported the isolation of both EpCAM positive and negative CTCs (Königsberg *et al.*, 2011). These

observations warned about the commonly accepted view that EpCAM is a general marker to detect CTCs. They rather suggested the existence of different, EpCAM positive and EpCAM negative populations, which is after all expected considering the heterogeneity of the primary tumors (Raimondi et al., 2015). Moreover, it has been reported that CTCs dynamically change their epithelial and mesenchymal composition and contain subpopulations of cells with epithelial, epithelial/mesenchymal, and mesenchymal characteristics (Yu *et al.*, 2013). Consistently, EMT can happen after cancer cells have already entered into the bloodstream, which has been for instance shown to be induced through platelet-derived TGF- β (Labelle et al., 2011). Although mesenchymal CTCs have been shown to be enriched in the blood circulation of cancer patients, and are considered to associate with disease progression, the functional contribution of different phenotypes of CTCs to invasion and metastases formation has not been often explored. Recent studies that have been attempting to investigate the functional relevance of different CTCs, are highlighting the importance of CTCs with intermediate phenotypes. Liu et al. have reported that CTCs of metastatic breast cancer with an epithelial and a limited mesenchymal phenotype had the highest capacity of metastases formation (Liu *et al.*, 2019), suggesting that a complete mesenchymal phenotype may not be always relevant to cancer progression. Moreover, CTCs are found both as individuals and as clusters in a cancer patient blood, and it turns out that CTC clusters have an intermediate EMT state, and that they have an increased metastasis capacity compared to individual CTCs (Aceto *et al.*, 2014; Jolly, Mani & Levine, 2018).

Although EpCAM is the common biomarker used in capturing CTCs from the peripheral blood, it has been shown that Trop2 can be potentially used as another cell surface marker to detect CTCs especially in EpCAM-independent enrichment of CTCs (Schneck *et al.*, 2015; Chen *et al.*, 2019).

EpCAM role in cell migration

Besides the studies on EpCAM expression levels being high or lost during different steps of cancer progression, its capacity in promoting or suppressing cell migration per se, as a prerequisite of cancer invasion, has been included in research studies on EpCAM which would potentially give rise to more direct information in completing the current knowledge on EpCAM role in cancer invasion. An extensive compilation of these studies has been done in a recent review by us (Fagotto & Aslemarz, 2020); both pro-migratory and anti-migratory effects have been recorded for EpCAM. There are only a few in vivo studies on EpCAMs role in cell motility that have been done in normal cell systems, including *Xenopus* or Zebrafish embryos or adult mice, all of which have shown that EpCAM is required, for example, for morphogenetic movements such as epiboly and thus cell motility (Slanchev *et al.*, 2009; Maghzal *et al.*, 2010, 2013) that were explained in detail earlier, or for migration of differentiated cells (Gaiser *et al.*, 2012). The latter investigated the migratory role of EpCAM in conditional knockdown mice showing that it was required for migration of Langerhans cells from skin epidermis to dermis. Likewise, Trop2 has been shown to promote cell motility during embryonic development and cancer progression (Mcdougall *et al.*, 2015).

Considering the complexity and plasticity of tumor cells' migratory behaviors during cancer invasion, it seems challenging to mimic physiologic conditions in experimental studies and this may partly explain the ambiguity of the reported pro-/or anti-migratory effects of EpCAM. As discussed in our review, most in vitro studies on EpCAM migratory function in cancer cells rely on two main approaches including scratch and trans-well assays. Each of these assays has its

limitations. The latter which is based on using two chambers separated by a porous filter can be used as a tool to run migration assay when the separating filter is just a non-coated porous membrane or invasion assay by having an ECM coating, for example, Matrigel on top of the separating filter. In the first case, cells are not provided with ECM, which would be a condition poorly relevant to physiological environments. Additionally, in both cases, the assays serve only as enumerative readouts that do not allow further characterization including live imaging or immunofluorescence staining. As for the scratch assay, while the technique allows for the above-mentioned further investigations, it still is limited by some drawbacks such as variations in the scratch area, damaged ECM, or damaged cells caused by mechanical creation of the scratch (Kramer *et al.*, 2013). These issues could be addressed by using alternative methods reviewed by Kramer *et al.*, among which we find the spheroid migration assay to be of interest for cancer studies since the three-dimensional (3D) multicellular spheroid formed from tumor cells can mimic a primary tumor-like tissue and the assay can be done as a two-dimensional (2D) migration assay by laying the spheroids on an ECM or 3D as an invasion assay by embedding them in the ECM which allows not only live imaging and other observations but could be a good model that let studying moving out or detachment of individual or small groups of cells from the main mass of cells, the spheroid's core, similar to *in vivo* conditions during cancer progression (Kramer *et al.*, 2013). In addition, the assay can be combined with using ECM proteins such as collagen as a soluble coating or polymerized solid ECM providing the cells with a soft ECM that is more relevant to the real cases, a combination that would not be possible to have in the two common methods mentioned above. To our knowledge spheroid migration assays are missing from EpCAM studies.

Together, it seems that more in vivo studies and more careful in vitro experimental designs are still required to provide a clearer picture of EpCAMs role in cell migration and cancer invasion.

MCF7 cells

These cells originate from the sites of metastases of a patient with invasive breast ductal carcinoma and are one of the most studied breast cancer cell lines (Soule *et al.*, 1973; Lee, Oesterreich & Davidson, 2015). They retain epithelial characteristics (Soule *et al.*, 1973; Dai *et al.*, 2017), thus can be used as a model for pre-metastasis tumors. In addition, they express the highest level of EpCAM among human breast cancer cell lines (Martovicz A. *et al.*, 2012).

Thesis objectives

Main objective

The work of this thesis is based on observations from the previous work of our laboratory on EpCAM function as a myosin activity regulator in cell adhesion/cell motility processes in the normal context of gastrulating *Xenopus laevis* embryos. The main objective of this thesis was thus to investigate these effects in a pathologically relevant context, human carcinoma cells.

Chapter II objective

The objective of this chapter was to elucidate the adhesive and migratory behavior of human breast carcinoma cells by modifying levels of EpCAM. This objective also included designing in vitro assays that are more relevant to the physiological conditions mainly by choosing cell spheroids to study the collective migration of cells and providing them with soft collagen I ECM. The data presented in this chapter show that EpCAM while maintaining its function as a negative regulator of myosin contractility in carcinoma cells, the cellular phenotypes resulting

from EpCAM depletion were interestingly different than what was observed in *Xenopus* embryonic tissues.

Chapter III objective

The initial objective of this chapter was to investigate the subcellular sites of EpCAM and nPKC interaction in order to know whether the interaction and subsequently the inhibition of myosin activity was distributed globally over the cell membrane/cortex, or it was located at some particular sites/subdomains. However, the chosen approach to answer this question, proximity ligation assay (PLA), turned out rather unspecific. Although not used for the initial purpose, the issue was addressed to point out the limitation of the method.

Chapter IV objective

The objective of this chapter was to discuss the major findings of this thesis and mention the perspective objectives that could be addressed in the future.

References

Aceto, N. *et al.* (2014) ‘Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis’, *Cell*, 158(5), pp. 1110–1122. doi: 10.1016/j.cell.2014.07.013.

Baeuerle, P. A. and Gires, O. (2007) ‘EpCAM (CD326) finding its role in cancer’, *British Journal of Cancer*, 96(3), pp. 417–423. doi: 10.1038/sj.bjc.6603494.

Balzar M, Winter MJ, de Boer CJ, L. S. (1999) ‘The biology of the 17 – 1A antigen (Ep-CAM)’, pp. 699–712.

Balzar, M. *et al.* (1998) ‘Cytoplasmic Tail Regulates the Intercellular Adhesion Function of the Epithelial Cell Adhesion Molecule’, *Molecular and Cellular Biology*, 18(8), pp. 4833–4843. doi:

10.1128/mcb.18.8.4833.

Balzar, M. *et al.* (1999) 'The structural analysis of adhesions mediated by Ep-CAM', *Experimental Cell Research*, 246(1), pp. 108–121. doi: 10.1006/excr.1998.4263.

Balzar, M. *et al.* (2001) 'Epidermal Growth Factor-Like Repeats Mediate Lateral and Reciprocal Interactions of Ep-CAM Molecules in Homophilic Adhesions', *Molecular and Cellular Biology*, 21(7), pp. 2570–2580. doi: 10.1128/mcb.21.7.2570-2580.2001.

Barth, A. I. M., , Honesty, K. and Riedel-kruse, I. H. (2018) 'Regulation of epithelial migration by epithelial cell adhesion molecule requires its Claudin-7 interaction domain', pp. 1–24.

Basu, A., Goldenberg, D. M. and Stein, R. (1995) 'The epithelial/carcinoma antigen EGP-1, recognized by monoclonal antibody RS7–3G11, is phosphorylated on serine 303', *International Journal of Cancer*, 62(4), pp. 472–479. doi: 10.1002/ijc.2910620419.

Bogenrieder, T. and Herlyn, M. (2003) 'Axis of evil: Molecular mechanisms of cancer metastasis', *Oncogene*, 22(43), pp. 6524–6536. doi: 10.1038/sj.onc.1206757.

Cayrefourcq, L. *et al.* (2015) 'Establishment and characterization of a cell line from human Circulating colon cancer cells', *Cancer Research*, 75(5), pp. 892–901. doi: 10.1158/0008-5472.CAN-14-2613.

Chalut, K. J. and Paluch, E. K. (2016) 'The Actin Cortex: A Bridge between Cell Shape and Function', *Developmental Cell*, 38(6), pp. 571–573. doi: 10.1016/j.devcel.2016.09.011.

Chambers, A. F., Groom, A. C. and MacDonald, I. C. (2002) 'Dissemination and growth of cancer cells in metastatic sites', *Nature Reviews Cancer*, 2(8), pp. 563–572. doi: 10.1038/nrc865.

Chastney, M. R., Conway, J. R. W. and Ivaska, J. (2021) 'Integrin adhesion complexes', *Current*

Biology, 31(10), pp. R536–R542. doi: 10.1016/j.cub.2021.01.038.

Chaves-Pérez, A. *et al.* (2013) ‘EpCAM regulates cell cycle progression via control of cyclin D1 expression’, *Oncogene*, 32(5), pp. 641–650. doi: 10.1038/onc.2012.75.

Chen, Q. *et al.* (2019) ‘Epithelial membrane protein 2: a novel biomarker for circulating tumor cell recovery in breast cancer’, *Clinical and Translational Oncology*, 21(4), pp. 433–442. doi: 10.1007/s12094-018-1941-1.

Chong, J. M. and Speicher, D. W. (2001) ‘Determination of Disulfide Bond Assignments and N-Glycosylation Sites of the Human Gastrointestinal Carcinoma Antigen GA733-2 (CO17-1A, EGP, KS1-4, KSA, and Ep-CAM)’, *Journal of Biological Chemistry*, 276(8), pp. 5804–5813. doi: 10.1074/jbc.M008839200.

Chugh, P. and Paluch, E. K. (2018) ‘The actin cortex at a glance’, *Journal of Cell Science*, 131(14), pp. 1–9. doi: 10.1242/jcs.186254.

Claas, C. *et al.* (2005) ‘The tetraspanin D6.1A and its molecular partners on rat carcinoma cells’, *Biochemical Journal*, 389(1), pp. 99–110. doi: 10.1042/BJ20041287.

Collinet, C. and Lecuit, T. (2013) *Stability and dynamics of cell-cell junctions*. 1st edn, *Progress in Molecular Biology and Translational Science*. 1st edn. Elsevier Inc. doi: 10.1016/B978-0-12-394311-8.00002-9.

Cubas, R. *et al.* (2009) ‘Trop2: A possible therapeutic target for late stage epithelial carcinomas’, *Biochimica et Biophysica Acta - Reviews on Cancer*, 1796(2), pp. 309–314. doi: 10.1016/j.bbcan.2009.08.001.

Dai, X. *et al.* (2017) ‘J o u r n a l o f C a n c e r Breast Cancer Cell Line Classification and Its

Relevance with Breast Tumor Subtyping’, 8. doi: 10.7150/jca.18457.

Ding, L. (2012) ‘Inflammation and Disruption of the Mucosal Architecture in Claudin-7–Deficient Mice’, pp. 305–315. doi: 10.1053/j.gastro.2011.10.025.

‘Elangbam et al., 1997-Cell adhesion molecules--update. Vet Pathol..pdf’ (no date).

Fagotto, F. and Aslemar, A. (2020) ‘EpCAM cellular functions in adhesion and migration, and potential impact on invasion: A critical review’, *Biochimica et Biophysica Acta - Reviews on Cancer*, 1874(2), p. 188436. doi: 10.1016/j.bbcan.2020.188436.

Friedl, P. *et al.* (2012) ‘Classifying collective cancer cell invasion’, *Nature Cell Biology*, 14(8), pp. 777–783. doi: 10.1038/ncb2548.

Friedl, P. and Gilmour, D. (2009) ‘Collective cell migration in morphogenesis, regeneration and cancer’, *Nature Reviews Molecular Cell Biology*, 10(7), pp. 445–457. doi: 10.1038/nrm2720.

Friedl, P. and Wolf, K. (2003) ‘Tumour-cell invasion and migration: Diversity and escape mechanisms’, *Nature Reviews Cancer*, 3(5), pp. 362–374. doi: 10.1038/nrc1075.

Gaber, A. *et al.* (2018) ‘EpCAM homo-oligomerization is not the basis for its role in cell-cell adhesion’, *Scientific Reports*, 8(1), pp. 1–16. doi: 10.1038/s41598-018-31482-7.

Gaiser, M. R. *et al.* (2012) ‘Cancer-associated epithelial cell adhesion molecule (EpCAM; CD326) enables epidermal Langerhans cell motility and migration in vivo’, *Proceedings of the National Academy of Sciences of the United States of America*, 109(15), pp. 889–897. doi: 10.1073/pnas.1117674109.

Gao, J. *et al.* (2014) ‘Epithelial-to-Mesenchymal Transition Induced by TGF- β 1 Is Mediated by API-Dependent EpCAM Expression in MCF-7 Cells’, (November 2013). doi: 10.1002/jcp.24802.

Gaston, C. *et al.* (2021) ‘EpCAM promotes endosomal modulation of the cortical RhoA zone for epithelial organization’, *Nature Communications*, 12(1), pp. 1–22. doi: 10.1038/s41467-021-22482-9.

Gires, O. *et al.* (2020) ‘Expression and function of epithelial cell adhesion molecule EpCAM: where are we after 40 years?’, *Cancer and Metastasis Reviews*, 39(3), pp. 969–987. doi: 10.1007/s10555-020-09898-3.

Goicoechea, S. M., Awadia, S. and Garcia-Mata, R. (2014) ‘I’m coming to GEF you: Regulation of RhoGEFs during cell migration’, *Cell Adhesion and Migration*, 8(6), pp. 535–549. doi: 10.4161/cam.28721.

Guerra, E. *et al.* (2012) ‘mTrop1/Epcam Knockout Mice Develop Congenital Tufting Enteropathy through Dysregulation of Intestinal E-cadherin/ β -catenin’, *PLoS ONE*, 7(11). doi: 10.1371/journal.pone.0049302.

van der Gun, B. T. F. *et al.* (2010) ‘EpCAM in carcinogenesis: The good, the bad or the ugly’, *Carcinogenesis*, 31(11), pp. 1913–1921. doi: 10.1093/carcin/bgq187.

Gupton, S. L. and Waterman-Storer, C. M. (2006) ‘Spatiotemporal Feedback between Actomyosin and Focal-Adhesion Systems Optimizes Rapid Cell Migration’, *Cell*, 125(7), pp. 1361–1374. doi: 10.1016/j.cell.2006.05.029.

Hagen, S. J. (2017) ‘Non-canonical functions of claudin proteins: Beyond the regulation of cell-cell adhesions’, *Tissue Barriers*, 5(2), pp. 1–14. doi: 10.1080/21688370.2017.1327839.

Harris, T. J. C. and Tepass, U. (2010) ‘Adherens junctions: From molecules to morphogenesis’, *Nature Reviews Molecular Cell Biology*, 11(7), pp. 502–514. doi: 10.1038/nrm2927.

Hemler, M. E. (2005) 'TETRASPANIN FUNCTIONS AND ASSOCIATED MICRODOMAINS', 6, pp. 801–811. doi: 10.1038/nrm1736.

Herlyn, M. *et al.* (1979) 'Colorectal carcinoma-specific antigen: Detection by means of monoclonal antibodies', *Proceedings of the National Academy of Sciences of the United States of America*, 76(3), pp. 1438–1442. doi: 10.1073/pnas.76.3.1438.

Hsu, Y. T. *et al.* (2016) 'EpCAM-regulated transcription exerts influences on nanomechanical properties of endometrial cancer cells that promote epithelial-to-mesenchymal transition', *Cancer Research*, 76(21), pp. 6171–6182. doi: 10.1158/0008-5472.CAN-16-0752.

Huang, Y. *et al.* (2019) 'Membrane-associated epithelial cell adhesion molecule is slowly cleaved by -secretase prior to efficient proteasomal degradation of its intracellular domain', *Journal of Biological Chemistry*, 294(9), pp. 3051–3064. doi: 10.1074/jbc.RA118.005874.

Jojović M, Adam E, Zangemeister-Wittke U, S. U. (1998) 'Epithelial glycoprotein-2 expression is subject to regulatory processes in epithelial ± mesenchymal transitions during metastases : an investigation of human cancers transplanted into severe combined immunode ® cient mice', 729.

Jolly, M. K., Mani, S. A. and Levine, H. (2018) 'Hybrid epithelial/mesenchymal phenotype(s): The “fittest” for metastasis?', *Biochimica et Biophysica Acta - Reviews on Cancer*, 1870(2), pp. 151–157. doi: 10.1016/j.bbcan.2018.07.001.

Kang, J. H. *et al.* (2012) 'Protein kinase C (PKC) isozyme-specific substrates and their design', *Biotechnology Advances*, 30(6), pp. 1662–1672. doi: 10.1016/j.biotechadv.2012.07.004.

Keller, L., Werner, S. and Pantel, K. (2019) 'Biology and clinical relevance of EpCAM', *Cell Stress*, 3(6), pp. 165–180. doi: 10.15698/cst2019.06.188.

Königsberg, R. *et al.* (2011) ‘Detection of EpCAM positive and negative circulating tumor cells in metastatic breast cancer patients’, *Acta Oncologica*, 50(5), pp. 700–710. doi: 10.3109/0284186X.2010.549151.

Kramer, N. *et al.* (2013) ‘In vitro cell migration and invasion assays’, *Mutation Research - Reviews in Mutation Research*, 752(1), pp. 10–24. doi: 10.1016/j.mrrev.2012.08.001.

Ku, N., Dederer, V. and Lemberg, M. K. (2019) ‘Intramembrane proteolysis at a glance : from signalling to protein degradation’. doi: 10.1242/jcs.217745.

Kuhn, S. *et al.* (2007) ‘A Complex of EpCAM , Claudin-7 , CD44 Variant Isoforms , and Tetraspanins Promotes Colorectal Cancer Progression’, 5(June), pp. 553–568. doi: 10.1158/1541-7786.MCR-06-0384.

Labelle, M., Begum, S. and Hynes, R. O. (2011) ‘Article Direct Signaling between Platelets and Cancer Cells Induces an Epithelial-Mesenchymal-Like Transition and Promotes Metastasis’, *Cancer Cell*, 20(5), pp. 576–590. doi: 10.1016/j.ccr.2011.09.009.

Ladwein, M. *et al.* (2005) ‘The cell – cell adhesion molecule EpCAM interacts directly with the tight junction protein claudin-7’, 309, pp. 345–357. doi: 10.1016/j.yexcr.2005.06.013.

Lamouille, S., Xu, J. and Derynck, R. (2014) ‘Molecular mechanisms of epithelial-mesenchymal transition’, *Nature Reviews Molecular Cell Biology*, 15(3), pp. 178–196. doi: 10.1038/nrm3758.

Lee, A. V, Oesterreich, S. and Davidson, N. E. (2015) ‘MCF-7 cells--changing the course of breast cancer research and care for 45 years.’, *Journal of the National Cancer Institute*, 107(7). doi: 10.1093/jnci/djv073.

Lei, Z. *et al.* (2012) ‘EpCAM contributes to formation of functional tight junction in the intestinal

epithelium by recruiting claudin proteins’, *Developmental Biology*, 371(2), pp. 136–145. doi: 10.1016/j.ydbio.2012.07.005.

Liang, K. H. *et al.* (2018) ‘Extracellular domain of EpCAM enhances tumor progression through EGFR signaling in colon cancer cells’, *Cancer Letters*, 433(March), pp. 165–175. doi: 10.1016/j.canlet.2018.06.040.

Linnenbach, A. J. *et al.* (1989) ‘Sequence investigation of the major gastrointestinal tumor-associated antigen gene family, GA733’, *Proceedings of the National Academy of Sciences of the United States of America*, 86(1), pp. 27–31. doi: 10.1073/pnas.86.1.27.

Linnenbach, A. J. *et al.* (1993) ‘Retroposition in a family of carcinoma-associated antigen genes’, *Molecular and Cellular Biology*, 13(3), pp. 1507–1515. doi: 10.1128/mcb.13.3.1507-1515.1993.

Lipp, P. and Reither, G. (2011) ‘Protein kinase C: The “Masters” of calcium and lipid’, *Cold Spring Harbor Perspectives in Biology*, 3(7), pp. 1–17. doi: 10.1101/cshperspect.a004556.

List, K., Bugge, T. H. and Szabo, R. (2006) ‘Matriptase: Potent Proteolysis on the Cell Surface’, *Molecular Medicine*, 12(1–3), pp. 1–7. doi: 10.2119/2006-00022.list.

Litvinov, S. V. *et al.* (1994) ‘Ep-CAM: A human epithelial antigen is a homophilic cell-cell adhesion molecule’, *Journal of Cell Biology*, 125(2), pp. 437–446. doi: 10.1083/jcb.125.2.437.

Litvinov, S. V. *et al.* (1997) ‘Epithelial cell adhesion molecule (Ep-CAM) modulates cell-cell interactions mediated by classic cadherins’, *Journal of Cell Biology*, 139(5), pp. 1337–1348. doi: 10.1083/jcb.139.5.1337.

Litvinov, S. V *et al.* (2016) ‘Evidence for a Role of the Epithelial Glycoprotein 40 (Ep-CAM) in Epithelial Cell-Cell Adhesion Evidence for a Role of the Epithelial Glycoprotein 40 (Ep-CAM)

in Epithelial Cell-Cell Adhesion’, 9061(February). doi: 10.3109/15419069409004452.

Liu, X. *et al.* (2019) ‘Epithelial-type systemic breast carcinoma cells with a restricted mesenchymal transition are a major source of metastasis’, (June).

Lock, J. G., Wehrle-Haller, B. and Strömblad, S. (2008) ‘Cell-matrix adhesion complexes: Master control machinery of cell migration’, *Seminars in Cancer Biology*, 18(1), pp. 65–76. doi: 10.1016/j.semcancer.2007.10.001.

Lu, H. *et al.* (2013) ‘EpCAM Is an Endoderm-Specific Wnt Derepressor that Licenses Hepatic Development’, *Developmental Cell*, 24(5), pp. 543–553. doi: 10.1016/j.devcel.2013.01.021.

Maetzel, D. *et al.* (2009) ‘Nuclear signalling by tumour-associated antigen EpCAM’, *Nature Cell Biology*, 11(2), pp. 162–171. doi: 10.1038/ncb1824.

Maghzal, N. *et al.* (2010) ‘The tumor-associated EpCAM regulates morphogenetic movements through intracellular signaling’, *Journal of Cell Biology*, 191(3), pp. 645–659. doi: 10.1083/jcb.201004074.

Maghzal, N. *et al.* (2013) ‘EpCAM Controls Actomyosin Contractility and Cell Adhesion by Direct Inhibition of PKC’, *Developmental Cell*, 27(3), pp. 263–277. doi: 10.1016/j.devcel.2013.10.003.

Maître, J. L. *et al.* (2012) ‘Adhesion functions in cell sorting by mechanically coupling the cortices of adhering cells’, *Science*, 338(6104), pp. 253–256. doi: 10.1126/science.1225399.

Maître, J. L. *et al.* (2015) ‘Pulsatile cell-autonomous contractility drives compaction in the mouse embryo’, *Nature Cell Biology*, 17(7), pp. 849–855. doi: 10.1038/ncb3185.

Maître, J. L. and Heisenberg, C. P. (2011) ‘The role of adhesion energy in controlling cell-cell

contacts’, *Current Opinion in Cell Biology*, 23(5), pp. 508–514. doi: 10.1016/j.ceb.2011.07.004.

Martowicz, A. *et al.* (2012) ‘Phenotype-dependent effects of EpCAM expression on growth and invasion of human breast cancer cell lines Phenotype-dependent effects of EpCAM expression on growth and invasion of human breast cancer cell lines’, *BMC Cancer*, 12(1), p. 1. doi: 10.1186/1471-2407-12-501.

Martowicz, A. *et al.* (2013) ‘EpCAM overexpression prolongs proliferative capacity of primary human breast epithelial cells and supports hyperplastic growth EpCAM overexpression prolongs proliferative capacity of primary human breast epithelial cells and supports hyperplastic growth’.

Mbaye, M. N. *et al.* (2019) ‘A comprehensive computational study of amino acid interactions in membrane proteins’, *Scientific Reports*, 9(1), pp. 1–14. doi: 10.1038/s41598-019-48541-2.

Mcdougall, A. R. A. *et al.* (2015) ‘Trop2: From development to disease’, *Developmental Dynamics*, 244(2), pp. 99–109. doi: 10.1002/dvdy.24242.

Molina, F. (1996) ‘Characterization of the type-1 repeat from thyroglobulin, a cysteine-rich module found in proteins from different families’, *European Journal of Biochemistry*, 240(1), pp. 125–133. doi: 10.1111/j.1432-1033.1996.0125h.x.

Mori, Y. *et al.* (2019) ‘Trophoblast cell surface antigen 2 (Trop-2) phosphorylation by protein kinase C / (PKC/) enhances cell motility’, *Journal of Biological Chemistry*, 294(30), pp. 11513–11524. doi: 10.1074/jbc.RA119.008084.

Munz, M. (2008) ‘Glycosylation is crucial for stability of tumour and cancer stem cell antigen EpCAM’, pp. 5195–5201.

Münz, M. *et al.* (2004) ‘The carcinoma-associated antigen EpCAM upregulates c-myc and induces

cell proliferation’, *Oncogene*, 23(34), pp. 5748–5758. doi: 10.1038/sj.onc.1207610.

Nagao, K. *et al.* (2009) ‘Abnormal placental development and early embryonic lethality in EpCAM-Null mice’, *PLoS ONE*, 4(12). doi: 10.1371/journal.pone.0008543.

Nakatsukasa, M. (2010) ‘Tumor-Associated Calcium Signal Transducer 2 Is Required for the Proper Subcellular Localization of Implications in the Pathogenesis of Gelatinous Drop-Like Corneal Dystrophy’, 177(3), pp. 1344–1355. doi: 10.2353/ajpath.2010.100149.

Nelson, A. J. *et al.* (1996) ‘The murine homolog of human Ep-CAM, a homotypic adhesion molecule, is expressed by thymocytes and thymic epithelial cells’, *European Journal of Immunology*, 26(2), pp. 401–408. doi: 10.1002/eji.1830260220.

Newton, A. C. (2001) ‘Protein kinase C: Structural and spatial regulation by phosphorylation, cofactors, and macromolecular interactions’, *Chemical Reviews*, 101(8), pp. 2353–2364. doi: 10.1021/cr0002801.

Nishikawa, K. *et al.* (1997) ‘Determination of the specific substrate sequence motifs of protein kinase C isozymes’, *Journal of Biological Chemistry*, 272(2), pp. 952–960. doi: 10.1074/jbc.272.2.952.

Pan, M. *et al.* (2018) *EpCAM ectodomain EpEX is a ligand of EGFR that counteracts EGF-mediated epithelial-mesenchymal transition through modulation of phospho-ERK1/2 in head and neck cancers*, *PLoS Biology*. doi: 10.1371/journal.pbio.2006624.

Pandya, P., Orgaz, J. L. and Sanz-Moreno, V. (2017) ‘Actomyosin contractility and collective migration: may the force be with you’, *Current Opinion in Cell Biology*, 48, pp. 87–96. doi: 10.1016/j.ceb.2017.06.006.

Parsons, J. T., Horwitz, A. R. and Schwartz, M. A. (2010) ‘Cell adhesion: Integrating cytoskeletal dynamics and cellular tension’, *Nature Reviews Molecular Cell Biology*, 11(9), pp. 633–643. doi: 10.1038/nrm2957.

Pastushenko, I. and Blanpain, C. (2019) ‘EMT Transition States during Tumor Progression and Metastasis’, *Trends in Cell Biology*, 29(3), pp. 212–226. doi: 10.1016/j.tcb.2018.12.001.

Pauli, C. *et al.* (2003) ‘Tumor-specific glycosylation of the carcinoma-associated epithelial cell adhesion molecule EpCAM in head and neck carcinomas’, *Cancer Letters*, 193(1), pp. 25–32. doi: 10.1016/S0304-3835(03)00003-X.

Pavšič, M. *et al.* (2014) ‘Crystal structure and its bearing towards an understanding of key biological functions of EpCAM’, *Nature Communications*, 5. doi: 10.1038/ncomms5764.

Pavšič, M. *et al.* (2015) ‘The cytosolic tail of the tumor marker protein Trop2 - A structural switch triggered by phosphorylation’, *Scientific Reports*, 5(November 2014), pp. 1–14. doi: 10.1038/srep10324.

Puram, S. V *et al.* (2017) ‘Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Article Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer’, *Cell*, 171(7), pp. 1611.e1-1611.e24. doi: 10.1016/j.cell.2017.10.044.

Raimondi, C., Nicolazzo, C. and Gradilone, A. (2015) ‘Circulating tumor cells isolation: The “post-EpCAM era”’, *Chinese Journal of Cancer Research*, 27(5), pp. 461–470. doi: 10.3978/j.issn.1000-9604.2015.06.02.

Ralhan, R. *et al.* (2010) ‘Nuclear and cytoplasmic accumulation of Ep-ICD is frequently detected

in human epithelial cancers’, *PLoS ONE*, 5(11). doi: 10.1371/journal.pone.0014130.

Rao, C. G. *et al.* (2005) ‘Expression of epithelial cell adhesion molecule in carcinoma cells present in blood and primary and metastatic tumors’, *International Journal of Oncology*, 27(1), pp. 49–57. doi: 10.3892/ijo.27.1.49.

Remšik, J. *et al.* (2018) ‘Trop-2 plasticity is controlled by epithelial-to-mesenchymal transition’, *Carcinogenesis*, 39(11), pp. 1411–1418. doi: 10.1093/carcin/bgy095.

Rosse, C. *et al.* (2010) ‘PKC and the control of localized signal dynamics’, *Nature Reviews Molecular Cell Biology*, 11(2), pp. 103–112. doi: 10.1038/nrm2847.

Salomon, J. *et al.* (2017) ‘Contractile forces at tricellular contacts modulate epithelial organization and monolayer integrity’, *Nature Communications*, 8. doi: 10.1038/ncomms13998.

Sankpal, N. V. *et al.* (2011) ‘Activator protein 1 (AP-1) contributes to EpCAM-dependent breast cancer invasion’, *Breast Cancer Research*, 13(6), pp. 1–13. doi: 10.1186/bcr3070.

Sankpal, N. V. *et al.* (2017) ‘A double-negative feedback loop between EpCAM and ERK contributes to the regulation of epithelial-mesenchymal transition in cancer’, *Oncogene*, 36(26), pp. 3706–3717. doi: 10.1038/onc.2016.504.

Sarrach, S. *et al.* (2018) ‘Spatiotemporal patterning of EpCAM is important for murine embryonic endo-And mesodermal differentiation’, *Scientific Reports*, 8(1), pp. 1–18. doi: 10.1038/s41598-018-20131-8.

Schiechl, H. and Dohr, G. (1987) ‘Immunohistochemical studies of the distribution of a basolateral-membrane protein in intestinal epithelial cells (GZ1-Ag) in rats using monoclonal antibodies’, *Histochemistry*, 87(5), pp. 491–498. doi: 10.1007/BF00496823.

- Schmidt, D. S. *et al.* (2004) ‘CD44 variant isoforms associate with tetraspanins and EpCAM’, *Experimental Cell Research*, 297(2), pp. 329–347. doi: 10.1016/j.yexcr.2004.02.023.
- Schneck, H. *et al.* (2015) ‘EpCAM-independent enrichment of circulating tumor cells in metastatic breast cancer’, *PLoS ONE*, 10(12), pp. 1–23. doi: 10.1371/journal.pone.0144535.
- Schnell, U., Cirulli, V. and Giepmans, B. N. G. (2013) ‘EpCAM: Structure and function in health and disease’, *Biochimica et Biophysica Acta - Biomembranes*, 1828(8), pp. 1989–2001. doi: 10.1016/j.bbamem.2013.04.018.
- Schnell, U., Kuipers, J. and Giepmans, B. N. G. (2013) ‘EpCAM proteolysis: New fragments with distinct functions?’, *Bioscience Reports*, 33(2), pp. 321–332. doi: 10.1042/BSR20120128.
- Schwartz, M. A. and Horwitz, A. R. (2006) ‘Integrating Adhesion, Protrusion, and Contraction during Cell Migration’, *Cell*, 125(7), pp. 1223–1225. doi: 10.1016/j.cell.2006.06.015.
- Sivagnanam, M. *et al.* (2008) ‘Identification of EpCAM as the Gene for Congenital Tufting Enteropathy’, *Gastroenterology*, 135(2), pp. 429–437. doi: 10.1053/j.gastro.2008.05.036.
- Slanchev, K. *et al.* (2009) ‘The epithelial cell adhesion molecule EpCAM is required for epithelial morphogenesis and integrity during zebrafish epiboly and skin development’, *PLoS Genetics*, 5(7). doi: 10.1371/journal.pgen.1000563.
- Soule, H. D. *et al.* (1973) ‘A human cell line from a pleural effusion derived from a breast carcinoma.’, *Journal of the National Cancer Institute*, 51(5), pp. 1409–1416. doi: 10.1093/jnci/51.5.1409.
- Stepan, L. P. *et al.* (2011) ‘Expression of Trop2 cell surface glycoprotein in normal and tumor tissues: Potential implications as a cancer therapeutic target’, *Journal of Histochemistry and*

Cytochemistry, 59(7), pp. 701–710. doi: 10.1369/0022155411410430.

Stoyanova, T. *et al.* (2012) ‘Regulated proteolysis of Trop2 drives epithelial hyperplasia and stem cell self-renewal via β -catenin signaling’, *Genes and Development*, 26(20), pp. 2271–2285. doi: 10.1101/gad.196451.112.

Tanaka, H. *et al.* (2015) ‘Intestinal deletion of claudin-7 enhances paracellular organic solute flux and initiates colonic inflammation in mice’, *Gut*, 64(10), pp. 1529–1538. doi: 10.1136/gutjnl-2014-308419.

Thampoe, I. J., Ng, J. S. C. and Lloyd, K. O. (1988) ‘Biochemical analysis of a human epithelial surface antigen: Differential cell expression and processing’, *Archives of Biochemistry and Biophysics*, 267(1), pp. 342–352. doi: 10.1016/0003-9861(88)90040-9.

Thiery, J. P. and Sleeman, J. P. (2006) ‘Complex networks orchestrate epithelial – mesenchymal transitions’, 7(February), pp. 131–142. doi: 10.1038/nrm1835.

Tobias Nubel, Preobraschenski, J. *et al.* (2009) ‘Claudin-7 Regulates EpCAM-Mediated Functions in Tumor Progression’, 7(March), pp. 285–300. doi: 10.1158/1541-7786.MCR-08-0200.

Trebak, M. *et al.* (2001) ‘Oligomeric state of the colon carcinoma-associated glycoprotein GA733-2 (Ep-CAM/EGP40) and its role in GA733-mediated homotypic cell-cell adhesion’, *Journal of Biological Chemistry*, 276(3), pp. 2299–2309. doi: 10.1074/jbc.M004770200.

Trerotola, M. *et al.* (2013) ‘Trop-2 promotes prostate cancer metastasis by modulating β 1 integrin functions’, *Cancer Research*, 73(10), pp. 3155–3167. doi: 10.1158/0008-5472.CAN-12-3266.

Trzpis, M. *et al.* (2007) ‘Epithelial cell adhesion molecule: More than a carcinoma marker and adhesion molecule’, *American Journal of Pathology*, 171(2), pp. 386–395. doi:

10.2353/ajpath.2007.070152.

Tsaktanis, T. *et al.* (2015) 'Cleavage and cell adhesion properties of human epithelial cell adhesion molecule (HEPCAM)', *Journal of Biological Chemistry*, 290(40), pp. 24574–24591. doi: 10.1074/jbc.M115.662700.

Vannier, C. *et al.* (2013) 'Zeb1 regulates E-cadherin and Epcam (epithelial cell adhesion molecule) expression to control cell behavior in early zebrafish development', *Journal of Biological Chemistry*, 288(26), pp. 18643–18659. doi: 10.1074/jbc.M113.467787.

Vicente-Manzanares, M. *et al.* (2009) 'Non-muscle myosin II takes centre stage in cell adhesion and migration', *Nature Reviews Molecular Cell Biology*, 10(11), pp. 778–790. doi: 10.1038/nrm2786.

Wang, J. *et al.* (2011) 'Loss of Trop2 promotes carcinogenesis and features of epithelial to mesenchymal transition in squamous cell carcinoma', *Molecular Cancer Research*, 9(12), pp. 1686–1695. doi: 10.1158/1541-7786.MCR-11-0241.

Wang, M. H. *et al.* (2018) 'Epithelial cell adhesion molecule overexpression regulates epithelial-mesenchymal transition, stemness and metastasis of nasopharyngeal carcinoma cells via the PTEN/AKT/mTOR pathway', *Cell Death and Disease*, 9(1). doi: 10.1038/s41419-017-0013-8.

Went, P. T. *et al.* (2004) 'Frequent EpCam Protein Expression in Human Carcinomas', *Human Pathology*, 35(1), pp. 122–128. doi: 10.1016/j.humpath.2003.08.026.

Wolf, K. and Friedl, P. (2006) 'Molecular mechanisms of cancer cell invasion and plasticity.', *The British journal of dermatology*, 154 Suppl, pp. 11–15. doi: 10.1111/j.1365-2133.2006.07231.x.

Wu, C. J. *et al.* (2013) 'Epithelial cell adhesion molecule (EpCAM) regulates claudin dynamics

and tight junctions’, *Journal of Biological Chemistry*, 288(17), pp. 12253–12268. doi: 10.1074/jbc.M113.457499.

Wu, C. J. *et al.* (2017) ‘Matriptase-mediated cleavage of EpCAM destabilizes claudins and dysregulates intestinal epithelial homeostasis’, *Journal of Clinical Investigation*, 127(2), pp. 623–634. doi: 10.1172/JCI88428.

Wu, C. J. *et al.* (2020) ‘Matriptase Cleaves EpCAM and TROP2 in Keratinocytes, Destabilizing Both Proteins and Associated Claudins’, *Cells*, 9(4). doi: 10.3390/cells9041027.

Xie, X. *et al.* (2005) ‘Expression pattern of epithelial cell adhesion molecule on normal and malignant colon tissues’, *World Journal of Gastroenterology*, 11(3), pp. 344–347. doi: 10.3748/wjg.v11.i3.344.

Yang, J. *et al.* (2020) ‘EpCAM associates with integrin and regulates cell adhesion in cancer cells’, *Biochemical and Biophysical Research Communications*, 522(4), pp. 903–909. doi: 10.1016/j.bbrc.2019.11.152.

Yu, M. *et al.* (2013) ‘Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition’, *Science*, 339(6119), pp. 580–584. doi: 10.1126/science.1228522.

Zandy, N. L., Playford, M. and Pendergast, A. M. (2007) ‘Abl tyrosine kinases regulate cell- cell adhesion through Rho GTPases’, *Proceedings of the National Academy of Sciences of the United States of America*, 104(45), pp. 17686–17691. doi: 10.1073/pnas.0703077104.

Zhao, W. *et al.* (2019) ‘The role and molecular mechanism of Trop2 induced epithelial-mesenchymal transition through mediated β -catenin in gastric cancer’, *Cancer Medicine*, 8(3), pp. 1135–1147. doi: 10.1002/cam4.1934.

Figure legends

Figure 1. Alignment of human EpCAM and Trop2 protein sequences. A. Alignment of amino acid sequences of human EpCAM and Trop2 was done using the Uniprot alignment tool, showing that the two proteins share 47.692% identity (marked by *). Their signal peptides residues are highlighted in orange and their transmembrane domains sequence in yellow. The juxtamembrane motif of the cytoplasmic domain containing the basic residues that interact with PKC is highlighted in pink.

Figure 2. Domain structure of human EpCAM and Trop2. The structural domain of hEpCAM and hTrop2 were depicted using Domain Graph (DOG) software. The signal peptides are shown as SP and in orange. The thyroglobulin-type A1 motif of their extracellular domain is shown as TY and in blue. The transmembrane domains are shown as TM and in yellow and the PKC binding motif of their cytoplasmic domains are shown in pink.

Figure 3. The gene expression profile of EpCAM and Trop2 across all tumor samples and paired normal tissues. A. The data were obtained from the gene expression profiling interactive analysis (GEPIA) database and depicted in excel that shows overexpression of both genes in most cancers. Their profiles in breast cancer are particularly shown by an oval. **B.** The table shows the extension of the tumor names abbreviations used in the graphs of section A.

Figure 4. The function of EpCAM as a cell-cell adhesion molecule. Different models of EpCAM function as a pro-adhesive molecule are shown. **1.** The very early model shows that EpCAM molecules interact in trans with each other. This model suggested EpCAM to be a weak cell-cell adhesion molecule. Its novel function as a negative regulator of cortical contractility can still favor EpCAM homophilic binding. **2.** The negative regulation of cortical contractility

indirectly stimulates cell-cell adhesion by favoring cadherin-cadherin trans interactions. **3.** Cortical contractility moderation by EpCAM possibly can result in cell-cell adhesion even in the absence of cadherin-cadherin adhesions and through non-specific interaction between cell surface components such as hyaluronans. **4.** And finally EpCAM might have heterotypic interacting partners that have remained unknown. Nevertheless, downregulation of cortical contractility by EpCAM can stimulate the cell-cell adhesion in this case as well. In all diagrams, EpCAM is shown as a homodimer (more detail in the text).

Figure 5. The function of EpCAM (and Trop2) as signaling molecules. Diagrams are showing the involvement of EpCAM in nuclear signaling and regulation of cortical myosin contractility. **1.** EpCAM undergoes regulated intramembrane proteolysis (RIP). This results in cleavage and release of EpCAM cytoplasmic C-terminus that binds to FHL2 then forming a complex with β -catenin. Translocation of this complex to the nucleus and association with transcription factors of TCF/Lef-1 family stimulates expression of targets genes involved in the cell proliferation process. Trop2 proteolysis through RIP has been shown thus same activity has been proposed for Trop2 as well. **2.** EpCAM interacts with nPKC directly and through the pathway of PKD, Raf, Erk, MLCK indirectly and negatively regulates cortical myosin contractility that can, in turn, facilitate cadherin-cadherin adhesion. The same impact could apply to the adhesion to the extracellular matrix (not illustrated), thus cell migration. Direct binding of Trop2 to nPKC has been shown, thus the same signaling function is likely by Trop2 as well.

Figure 6. Table of animal / cellular phenotypes involving EpCAM (and Trop2) functions. This table summarizes some of the animal and cellular phenotypes observed upon knockdown (KD), overexpression (OE) or knockout (KO) of EpCAM or Trop2. An extensive version of the

impact of these modifications on cell-cell adhesion and cell migration is provided in our recent review (mentioned in the text).

Figures

Figure 1.

```
SP|P16422|EPCAM_HUMAN -----MAPPQVLAFGLLLAATATFAAAQEECVCENYKLAVNCFVNNNRQCQCTSVGA 53
SP|P09758|TACD2_HUMAN MARGPGLAPPPLRLPLLLLVLAAVTGHTAAQDNCTCPTNKMTVCSPDGPGGRCQCRA LGS 60
                        ** * : **: **: .:***:*. * . *: * . . . :*** **:

SP|P16422|EPCAM_HUMAN QNTVICSKLAACKLVMKAEMNGSKLGRRA--KPEGALQNN DGLYDPDCDESGLFKAKQCN 111
SP|P09758|TACD2_HUMAN GMAVDCSTLT SKLLLKARMSAPKNARTLVRPSEHALVDNDGLYDPDCDPEGRFKARQCN 120
                        :* **.*: **: **:*. * . * * * :***** * * ** :***:

SP|P16422|EPCAM_HUMAN GTSMCWCVNTAGVRRTDK-DTEITCSE RVRTYWI IELKHKAREKPYDSKSLRTALQKEI 170
SP|P09758|TACD2_HUMAN QTSVCWCVNSVGVRRTDKGDL SLCDELVRTHHILIDL RHRPTAGAFNHSDLD AELRRLF 180
                        **:*****:***** * .: *. * ***: *:***: : : . * : * : :

SP|P16422|EPCAM_HUMAN TTRYQLDPKFITSILYENNVITIDL VQNSSQKTQNDVDIADVAYYFEKDVKGESLFHSK- 229
SP|P09758|TACD2_HUMAN RERYRLHPKFVAAVHYEQPTIQIELRQNTS QKAAGVDVIGDAAYYFERDIKESLFQGRG 240
                        **:*.***: : : **: . * * : * ***: **: .****.*.*****:*****: :

SP|P16422|EPCAM_HUMAN KMDLTVNGEQLDLDPGQTLIYYVDEKAPEFSMQGLKAGVIAVIVVVVIAVVAGIVVLV I S 289
SP|P09758|TACD2_HUMAN GLDLRVRGEPLQ--VERTLIYYLDEI PPKFSMKRLTAGLIAVIVVVVVALVAGMAVLVI T 298
                        .** *.** *: :*****.* *:***: *.**:* *****:***:*.****:

SP|P16422|EPCAM_HUMAN RKKRMAKYEKAEIKEMGEMHRELNA 314
SP|P09758|TACD2_HUMAN NRRKSGKYKKVEIKELGELRKEPSL 323
                        .: : .**:*.*****:***:*. *
```

Figure 2.

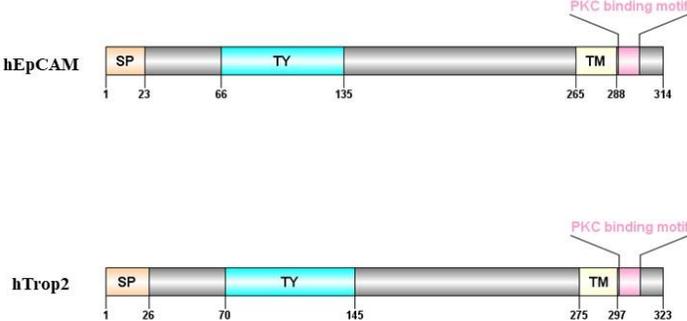
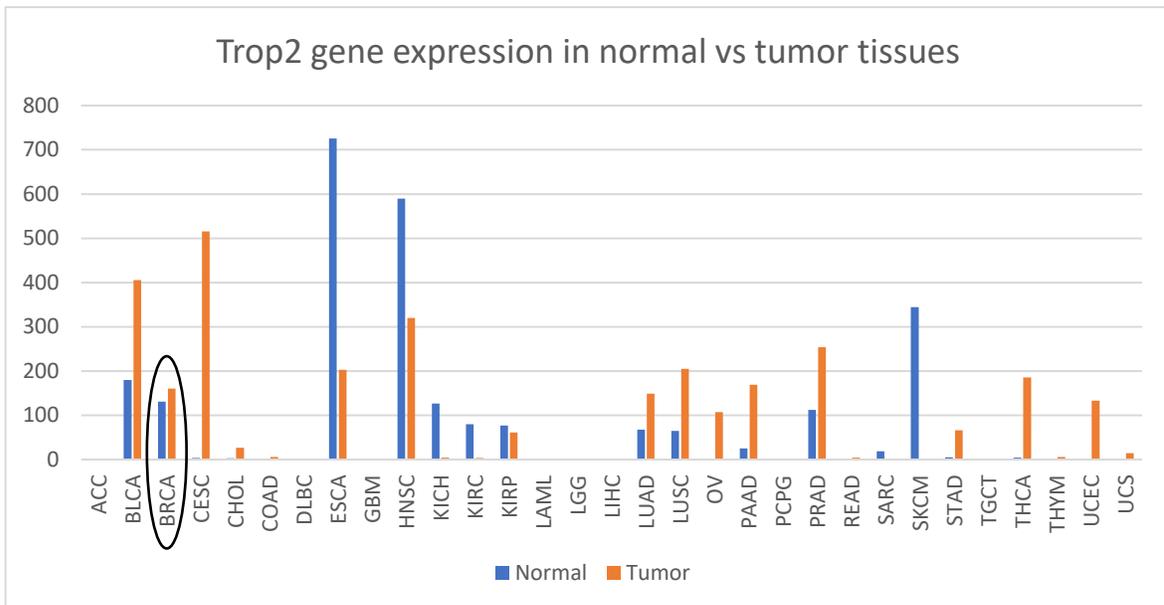
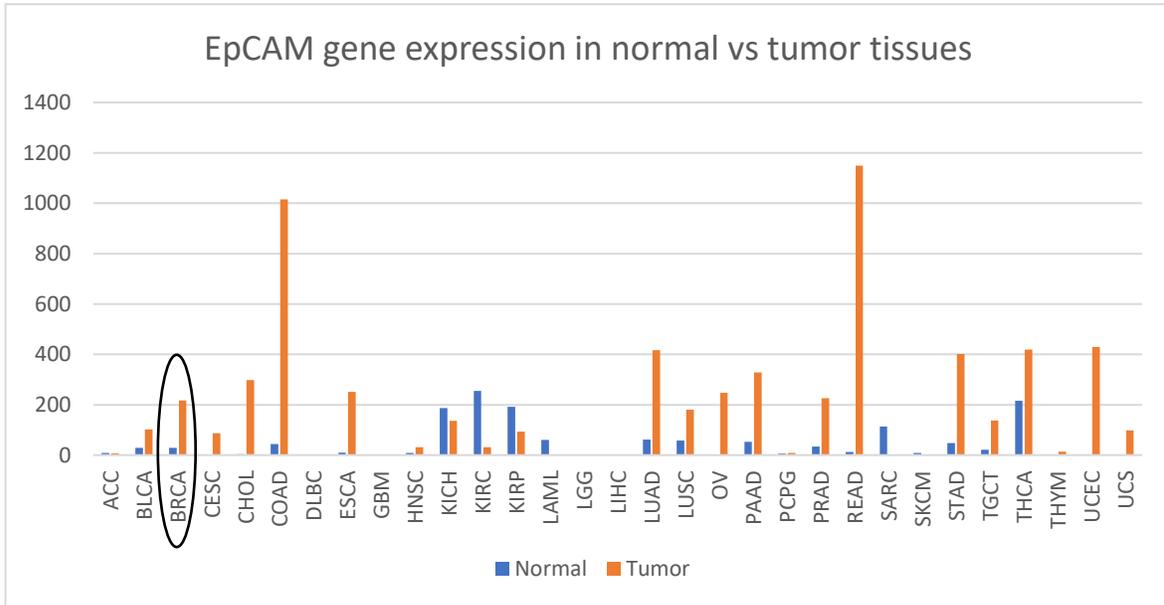


Figure 3.

A.



B.

Tumor name abbreviation	Tumor name extension
ACC	Adrenocortical carcinoma
BLCA	Bladder Urothelial Carcinoma
BRCA	Breast invasive carcinoma
CESC	Cervical squamous cell carcinoma and endocervical adenocarcinoma
CHOL	Cholangio carcinoma
COAD	Colon adenocarcinoma
DLBC	Lymphoid Neoplasm Diffuse Large B-cell Lymphoma
ESCA	Esophageal carcinoma
GBM	Glioblastoma multiforme
HNSC	Head and Neck squamous cell carcinoma
KICH	Kidney Chromophobe
KIRC	Kidney renal clear cell carcinoma
KIRP	Kidney renal papillary cell carcinoma
LAML	Acute Myeloid Leukemia
LGG	Brain Lower Grade Glioma
LIHC	Liver hepatocellular carcinoma
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
OV	Ovarian serous cystadenocarcinoma
PAAD	Pancreatic adenocarcinoma
PCPG	Pheochromocytoma and Paraganglioma
PRAD	Prostate adenocarcinoma
READ	Rectum adenocarcinoma
SARC	Sarcoma
SKCM	Skin Cutaneous Melanoma
STAD	Stomach adenocarcinoma
TGCT	Testicular Germ Cell Tumors
THCA	Thyroid carcinoma
THYM	Thymoma
UCEC	Uterine Corpus Endometrial Carcinoma
UCS	Uterine Carcinosarcoma

Figure 4.

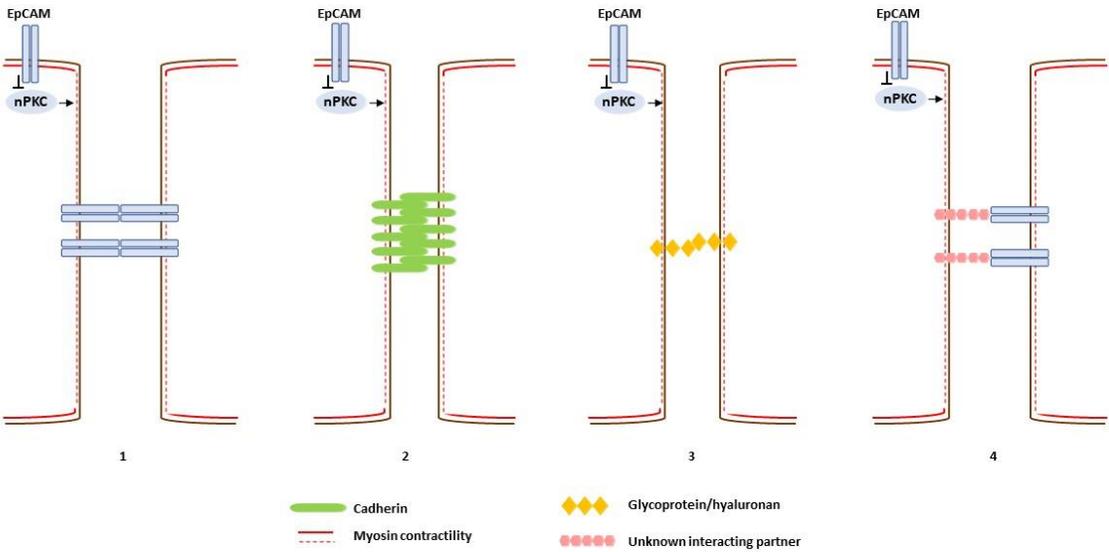


Figure 5.

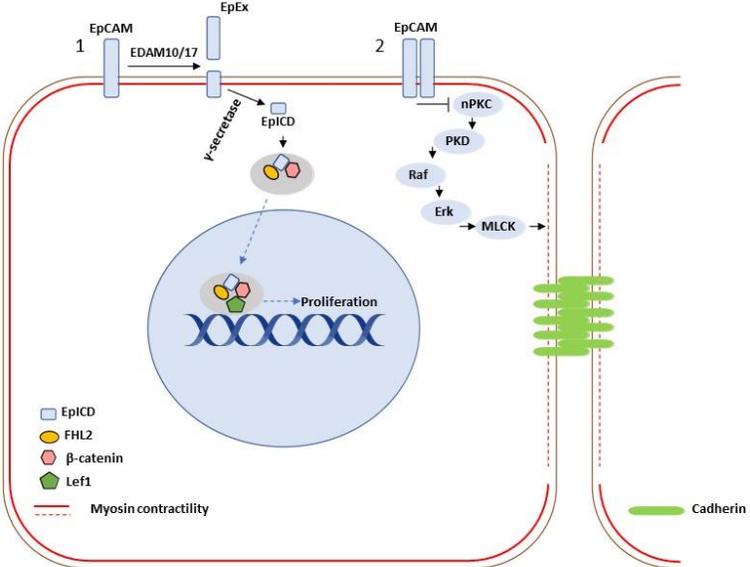


Figure 6.

Animal/ cellular model	Approach	Phenotype	Reference
L-cells/L153S	Ectopic expression	Intercellular adhesion and aggregation in suspension	(Litvinov et al., 1994)
Xenopus gastrula	EpCAM KD	Impaired motility/ loss of adhesion and tissue integrity	(Maghzal et al., 2013)
Xenopus gastrula	EpCAM OE	Induced motility	(Maghzal et al., 2010)
Zebrafish gastrula	EpCAM KD	Impaired motility/decreased E-cadherin/loss of skin integrity	(Slanchev et al., 2009)
Zebrafish lateral line	EpCAM KD	Impaired neuromast deposition	(Villablanca et al., 2006)
Mice	EpCAM KO	Embryonic lethal (placental defects)	(Nagao et al., 2009)
Mice	EpCAM KO	Death after birth (intestinal defects)	(Guerra et al., 2012)
MDCK cells	EpCAM KD	Increased monolayer migration	(Barth et al., 2018)
Caco-2 cells	EpCAM KD	Cadherin defects/ tight junction displacement/ epithelial dysplasia	(Salomon et al., 2017)
T84 and caco-2 cells	EpCAM KD	Decreased levels and displacement of claudin-7	(Wu et al., 2013)
Caco-2 cells	EpCAM KD	Loss of front-rear polarity in single cells	(Gaston et al., 2021)
Mice with EpCAM-deficient skin Langerhans cells	Conditional KO	Decreased migration from epidermis to lymph nodes	(Gaiser et al., 2012)
MCF7 cells	EpCAM OE/KD	Increased/decreased monolayer migration	(Gao et al., 2015)
Mice	Trop2 KO	Viable mice/ susceptible to developing cancer	(Wang et al., 2011)
Rat fetal lung fibroblasts	Trop2 KD	Decreased migration/proliferation	(McDougall et al., 2013)
Rat fetal lung fibroblasts	Trop2 OE	Increased migration/proliferation	(McDougall et al., 2013)

CHAPTER II

EpCAM differentially regulates individual and collective migration of human carcinoma cells

**Azam Alsemarz^{1,2}, Marie-Fagotto Kaufmann¹, Artur Ruppel³, Martial Balland³, Paul
Lasko² and François Fagotto^{1,2}**

¹CRBM, University of Montpellier and CNRS, Montpellier 34293, France
and ²Dept. of Biology, McGill University, Montreal, QC, Canada H3A1B1.

³LIPHY, UMR5588, University of Grenoble, 38400 Grenoble, France

Correspondence to François Fagotto, email: francois.fagotto@crbm.cnrs.fr

Abstract

EpCAM is well known as a carcinoma-associated cell surface protein that is commonly used as an epithelial cancer biomarker. Despite its clear link to cancer development, whether it plays an actual role in cancer metastasis remains unclear. It is known, however, to downregulate myosin contractility, a key parameter involved in cell adhesion and migration. We have examined here the potential morphogenetic impact of the high EpCAM expression that characterizes epithelial breast cancer cells, using spheroids of MCF7 cells as a model of the pre-metastatic tumor. We found that EpCAM depletion repressed single-cell migration but strongly stimulated cohesive collective migration. Using a combination of cell biological and biophysical approaches, we show that EpCAM depletion, while globally increases cell contractility, induces a strong reinforcement of cadherin cell-cell adhesion, which accounts for increased tissue cohesion and explains the switch toward collective migration. Interestingly, EpCAM negative and positive cells tend to sort in mixed spheroids.

MCF7 cells also express Trop2, a close relative of EpCAM. Intriguingly, Trop2, similar to EpCAM, also contributes to moderating cell contractility, yet its depletion leads to the exact opposite global phenotype, stimulating single-cell migration while decreasing tissue cohesiveness and collective migration. Comparison with EpCAM depletion indicates that, while both proteins display largely overlapping cellular activities, they differentially modulate distinct myosin functions, EpCAM in cell-cell adhesion, Trop2 in cell-matrix adhesion. These results highlight the importance of a fine balance of distribution of myosin contractility downstream of EpCAM (and Trop2) in controlling tumor morphogenesis, in particular in favor of different modes of migration.

Introduction

EpCAM, also called TACSTD1, is a major cell surface protein highly expressed in most human carcinomas (Went et al., 2004; Rao et al., 2005). It is used since many years as an important cancer biomarker for diagnostic and therapeutic purposes (Baeuerle & Gires, 2007; Keller, Werner & Pantel, 2019). EpCAM was first considered to act as a homotypic cell adhesion molecule (Balzar et al., 1999; Litvinov et al., 1994), but firm experimental support for this function is lacking and has been altogether questioned by recent studies (Gaber et al., 2018). In the meantime, signaling activities have been identified, proposing EpCAM as a regulator of proliferation (Maetzel *et al.*, 2009; Chaves-Pérez *et al.*, 2013) as well as of physical cell properties, the latter due to its ability to control myosin contractility (Maghzal *et al.*, 2010, 2013; Salomon *et al.*, 2017; Barth, Honesty & Riedel-kruse, 2018; Gaston *et al.*, 2021). Myosin regulation offers an obvious potential link between EpCAM and cancer metastasis which remains to be established.

The status of EpCAM in cancer invasion is unclear: Various in vitro studies have led to conflicting results, concluding that EpCAM was either pro-adhesive or anti-adhesive, pro-migratory or anti-migratory (reviewed in Fagotto & Aslemarz, 2020). Furthermore, EpCAM is known to be downregulated upon epithelial-mesenchymal transition (EMT) (Vannier *et al.*, 2013; Sankpal *et al.*, 2017; Pan *et al.*, 2018), yet it has also been proposed to be itself required for EMT (Sankpal *et al.*, 2011; Gao *et al.*, 2014, 2015; Wang *et al.*, 2018). Note more generally, that the classical view of that posed EMT as a necessary and key step in cancer invasiveness is currently a matter of debate. There is emerging evidence that, at least in some cases, the most invasive cells may have only undergone limited EMT, and still display many aspects of the epithelial cells from which they derive (Ebright *et al.*, 2020). Along these lines, circulating tumor cell lines with high

metastatic potential have been established, which happen to be EpCAM positive (Koch *et al.*, 2020).

In addition to EpCAM, all epithelia also express its closely relative Trop2 (except for the intestine epithelium, which only expresses EpCAM). EpCAM and Trop2 (also called TACSTD2) constitute a gene family on their own, Trop2 being the result of EpCAM gene duplication due to retrotransposition, that occurred in reptilians, at the root of amniote evolution. Human EpCAM and Trop2 show very high sequence similarity between themselves, and with the EpCAM of lower vertebrates. Trop2 has so far shown identical characteristics to EpCAM in virtually all features studied so far, including the structure and biochemical interactions (Fagotto & Aslemar, 2020; Lenárt *et al.*, 2020). Trop2 has been also linked to cancer, but, again, its actual role remains to be established (Lenárt *et al.*, 2020).

EpCAM morphogenetic function is best understood in non-cancer systems, specifically in embryonic tissues. In both zebrafish and *Xenopus* early embryos, gain and loss-of-function experiments showed that EpCAM expression positively impacted both cell-cell adhesion and intercellular migration, i.e. on the ability of cells to rearrange by moving relative to neighboring cells (Slanchev *et al.*, 2009; Maghzal *et al.*, 2010, 2013). Thus, EpCAM is first required during gastrulation for the morphogenetic movement of ectoderm epiboly (Slanchev *et al.*, 2009; Maghzal *et al.*, 2010), then later in development, it plays an essential role in controlling tissue cohesion, as its loss leads to a severe loss of tissue integrity (Slanchev *et al.*, 2009; Maghzal *et al.*, 2013). At the molecular level, our team demonstrated that EpCAM directly inhibited the novel class of PKC kinases (nPKC), and through this inhibition, it repressed non-muscle myosin II (NMII) activation by the nPKC-PKD-Raf-Erk-MLCK pathway (Maghzal *et al.*, 2013). This role in the downregulation of myosin activity accounted for all the EpCAM gain and loss-of-function

embryonic phenotypes. Importantly, the loss of cadherin-mediated adhesion observed upon EpCAM depletion was shown to be a secondary result of exacerbated actomyosin contractility (Maghzal et al., 2013). Interestingly, loss of EpCAM was also shown to lead to deregulated contractility and disruption of epithelial integrity in a rare human intestinal disorder called congenital tufting enteropathy (CTE) (Sivagnanam et al., 2008). Here, again, EpCAM was shown to act as a myosin negative regulator, its loss leading to increased tension, causing defects in particular at tricellular junctions (Salomon *et al.*, 2017). Note that a recent study on the migration of single intestinal Caco2 cells proposed an additional/alternative role of EpCAM in spatially organizing RhoA and myosin activity via fast recycling endosomes (Gaston *et al.*, 2021).

Since the ability of EpCAM to regulate myosin appears to be conserved and considering that EpCAM expression is highest in carcinomas, it is fair to predict that this function should be relevant for cancer metastasis. One must emphasize, however, that, due to the multiplicity and complexity of myosin functions, one can hardly predict *a priori* how its up- or downregulation will impact adhesion and migration in various cell types and under different conditions. While contractility of the cell actomyosin cortex tends to antagonize cell adhesion, the system is also able to respond to increased contractility by reinforcing adhesion (Leckband & de Rooij, 2014). In turn, both cortical contractility and adhesion have complex effects on cell motility: Softening the cell cortex is required for emission of cell protrusions, yet the cortex is also responsible for force transmission, which is also essential for migration, in particular for collective migration (Mayor & Etienne-Manneville, 2016). Adhesion to a support, whether matrix or other cells, is also a basic requirement for movement, yet strong/stable adhesion is bound to slow down and eventually stall migrating cells (Charras and Sahai, 2014).

In an attempt to clarify the role of EpCAM in a cancer context, we chose the breast epithelial MCF7 cells as a model of a pre-metastatic state (Soule *et al.*, 1973; Dai *et al.*, 2017). We evaluated the impact of EpCAM levels on the ability of MCF7 cells to migrate as single cells as well as a group, and we systematically characterized the effect on myosin contractility and cell adhesion, using a combination of cellular and biophysical approaches. We also included in the study Trop2. Trop2 is very similar to EpCAM, and they have been shown to be at least partly redundant (Wu *et al.*, 2020), and Trop2 expression could partially rescue CTE defects caused by EpCAM loss (Nakato *et al.*, 2020). Yet, the conservation of the two genes throughout amniotes argues for distinct functions. We thus considered it important to carry their systematic characterization side by side in our experimental system, aiming at defining their relative contribution to cell and tissue properties.

Results

EpCAM KD increases the collective migration of MCF7 spheroids and their cohesiveness

In order to investigate the role of EpCAM in collective migration, we used an assay where MCF7 formed spheroids in suspension as an *in vitro* model that mimics a primary solid tumor (Kramer *et al.*, 2013) that were then placed on a soft matrix of collagen I as a physiologically relevant ECM (Insua-Rodríguez and Oskarsson, 2016). Under these conditions, spheroids actively spread, a process that we typically imaged over 24 to 48hrs. We chose this 2D model of migration over spheroids fully embedded in collagen, because, in the latter setting, migration also depends on degradation of the extracellular matrix, a limiting and confounding parameter that we wanted to avoid. EpCAM was depleted by transfection of siRNA one day before starting to form the spheroids, which were laid on collagen two days later. By that stage, EpCAM depletion was close to complete and remained the same until the end of the assay (Fig.S1).

Spheroids from cells transfected with control siRNA (siCtrl) spread efficiently, typically increasing their area by 3 to 4 folds (Fig.1A,E). Note that the center of the spheroid tended to remain rather compact, while the periphery showed highly irregular contours, with cells or cell groups frequently sticking out of the cell mass (Fig.1A,H,J). Non-transfected spheroids behaved identically (not shown).

EpCAM KD resulted in a strong, highly reproducible increase in spreading (Fig.1B and E). After 24hrs, spheroids had the shape of a flat pancake, only about 2-3 cells thick (Fig.1H'). Interestingly, these EpCAM KD spheroids expanded as a very cohesive sheet: While numerous protrusions emanated from the edge of the cell mass (Fig.1J), this edge was overall strikingly smooth (Fig.1H,I,J). The distinctive morphologies of control and EpCAM KD spheroids, indicative of differences in tissue cohesiveness, were quantitatively expressed as “solidity” (Fig.1F). We verified that the increased migration/spreading of siEpCAM spheroids was not due to a higher proliferation rate by running the spheroid assay in the presence of Mitomycin C (MMC), an anti-mitotic agent (Fig.S2). MMC totally blocked the cell cycle, as validated by thymidine analog EdU incorporation (Fig.S2A). We concluded that the increased spreading upon EpCAM KD was purely due to changes in morphogenetic properties.

Based on the similarities between EpCAM and Trop2, and their assumed partial redundancy, we expected that Trop2 KD would also cause increased spreading and that the double KD might potentiate the EpCAM KD phenotype. Quite unexpectedly, however, Trop2 KD gave a completely opposite phenotype: Spheroid spreading was on the contrary strongly decreased compared to siCtrl (Fig.1C,E). In addition, solidity also appeared dramatically decreased, indicating that the tissue had become less cohesive (Fig.1C,F). Double EpCAM/Trop2 KD (dKD)

showed a phenotype more similar to single EpCAM KD, although slightly less severe, both in terms of spreading and cohesiveness (Fig.1D,E,F).

EpCAM KD-induced collective migration depends on nPKC and myosin

As the EpCAM loss-of-function phenotypes observed for embryonic tissues were due to overactivation of the nPKC-myosin pathway (Maghzal *et al.*, 2013), we asked if there was a similar dependency for the effect on MCF7 spheroids. We thus tested spreading of both control and EpCAM-depleted spheroids in the presence of calphostin, as a specific inhibitor of nPKC, or blebbistatin, the well-established inhibitor of myosin II activity. We observed a significantly reduced spreading with both inhibitors. Spreading of siEpCAM spheroids was reduced close to the levels of controls while spreading of the control spheroids themselves was further inhibited. These results clearly showed that, under these conditions, MCF7 collective migration was positively regulated by nPKC and myosin activity, downstream of EpCAM depletion (Fig.S3).

EpCAM KD decreases migration of individual MCF7 cells

Spheroid spreading involves two types of cell motility: The ability to migrate on the collagen matrix, and the capacity to intercalate within the cell mass. One obvious possible cause for the spheroid migration/spreading phenotypes could be a difference in the migratory property of the individual cells. We analyzed the migration of single cells, also plated on a layer of collagen I gel. siCtrl MCF7 cells showed heterogeneous morphologies, ranging from rather round to well spread (example Fig.2A, quantification of the area in Fig.2G). They were generally not very motile, with an average motility of $0.5\mu\text{m}/\text{min}$ (Fig.2A,E). siEpCAM cells spread significantly more than controls, adopting a characteristic radial shape (Fig.2B,G). Their migration was strongly decreased (Fig.2B,E,F). Thus, EpCAM depletion, while stimulating collective spreading, had the

opposite effect on single-cell migration. Conversely, Trop2-depleted cells were rounder and taller (Fig.2C,G). Surprisingly, despite this round shape, they showed to be capable of fast migration (Fig.2C,E,F). dKD cells were similar to siEpCAM cells, both in terms of extended morphology and low migration (Fig.2D-G).

Hence, the comparison of single cells and spheroids yielded two seemingly contradictory conclusions: In terms of morphology, EpCAM KD had the same effect at the cell scale and that of the global scale of cell masses, i.e. strong flattening and expansion. Conversely, Trop2 KD decreased matrix contact spreading, also at both scales. On the other hand, the effect of these depletions on migration was exactly the opposite for single-cell migration and collective spreading, EpCAM KD decreasing the former and stimulating the latter. Again, Trop2 KD had the inverse impact.

EpCAM depletion prevents detachment of individual cells

Considering the high cohesiveness of collective migration of EpCAM KD cells, opposed to the strong inhibition of their single-cell migration, we went back to have a closer look at the spheroid spreading time lapses. Indeed, while we had observed occasional detachment of individual cells in wild-type spheroids (Fig.3A), this did not seem to occur in EpCAM KD spheroids (Fig.3B). Such events, even if rare, are highly relevant for cancer invasion. In fact, we observed that 25 out of 84 control spheroids (from 13 experiments) had at least one isolated cell that detached and migrated away from over the 24hrs of imaging, and 7 additional spheroids even showed detachment of small groups of cells (Fig.3E). In contrast, the analysis of 85 EpCAM KD spheroids revealed only three cases of detachment (Fig.3E). The chances of observing detachments were much higher with Trop2-depleted spheroids, here mostly small groups of cells rather than single cells (Fig.3E, a pair of cells in the example of Fig.3C). Detachments were fairly rare in dKD

(less than 6%). These results suggested that EpCAM and Trop2 play a role in setting the balance between tissue cohesion and mode of migration, either single (and dispersive) or collective. High EpCAM levels in breast cancer MCF7 cells seemed to favor single-cell migration and dispersion, while Trop2 has an antagonistic function. This conclusion was further supported by the observation that experimental EpCAM overexpression in the same MCF7 cells reduced collective migration (Fig.3F).

Common and distinct effects of EpCAM and Trop2 depletions on pMLC and E-cadherin levels

In order to better grasp the effect of these depletions, we set to characterize spheroids and isolated cells in terms of adhesive and cytoskeletal properties. Considering the known activity of EpCAM, myosin activation was a prime suspect to analyze.

In whole-mount immunofluorescence (IF) of spheroids, phosphorylated myosin light chain (pMLC) was detected both along cell-cell contacts and along the free edge of the spheroids, the latter signal being generally most prominent (Fig.4A). Importantly, total pMLC levels were significantly increased both by siEpCAM and by siTrop2. This result confirmed that both EpCAM and Trop2 have indeed a negative regulatory activity on myosin II in MCF7 cells. Double KD did not further increase pMLC levels (Fig.4A,C).

E-cadherin levels are a good indicator of cell-cell adhesion, with strong correlation between both parameters (Winklbauer, 2015). We observed that both EpCAM KD and Trop2 KD, significantly increased E-cadherin intensity. Note, however, that, unlike for pMLC, the increase was milder in the case of Trop2 KD (Fig.4B,D). We concluded that loss of EpCAM stimulated both myosin contractility and E-cadherin-based cell-cell adhesion, a situation typical of adhesive

contacts being reinforced in response to increased tension. Loss of Trop2 also led to the same phenotype, although in this latter case, reinforcement of adhesion seemed to be somewhat less effective.

Beyond total levels, the relative subcellular distribution of myosin, in particular between cell-cell contacts and the free cell cortex, is an important parameter influencing the tensile and adhesive properties. Such analysis would be extremely complex in whole spheroids. We thus used as simpler model small groups of cells, ranging from doublets to groups of 5-6 cells. Since cells had a wide range of complex morphologies, including multiple large protrusions, the comparison of signal intensities over the whole cell surface and/or whole cell contacts was not very meaningful. Myosin and cadherins tend indeed to concentrate at particular regions, which vary with each cell configuration, and which correspond to sites where the most tension is exerted. Thus, a better readout was to locate and measure “peak” intensity. We thus determined peak intensity for pMLC at the cortex bordering free cell edges, as well as peaks of pMLC and E-cadherin at cell-cell contacts. This analysis yielded several important results. Firstly, EpCAM KD increased all three parameters, confirming that myosin activity was indeed globally upregulated by EpCAM depletion, and that contacts responded by recruiting more E-cadherin, as observed in whole spheroids (Fig.5,A,B,C). Similarly, peak pMLC and E-cadherin were also elevated upon Trop2 KD, again validating the results in spheroids (Fig.5,A,B,C). Beyond these global changes, a comparison of cortical pMLC, contact pMLC and E-cadherin provided a more nuanced view of the phenotypes. Indeed, in the case of EpCAM KD, the ratio between contact and cortical pMLC was not significantly changed compared to siCtrl (Fig.5C). In the case of Trop2 KD, however, pMLC was clearly more intense at cell-cell contacts than the free cortex (Fig.5C). This is a key observation, which indicates that EpCAM and Trop2 differentially impact these two myosin pools.

In particular, the relative tension at cell-cell contacts compared to the free cortex is directly implicated in setting cell-cell “adhesiveness” (Winklbauer, 2015; Fagotto, 2020). The higher contact/free cortex ratio in Trop2 KD is thus fully consistent with a lower coherence. This hypothesis was corroborated by the pMLC to E-cadherin ratio, which was also highest for siTrop2, while it was only slightly and non-significantly increased for siEpCAM (Fig.5D). All these samples were also stained with phalloidin, and we also quantified F-actin (Fig.5E). Global trends caused by either EpCAM or Trop2 depletion were similar to pMLC. Double depletion gave more complex results, which were difficult to interpret, considering the complexity of the effect of single depletions. One may note that while the pMLC levels in dKD spheroids were generally similar to siEpCAM, levels at the free cortex tended to be even higher (similar relative median, 1.8 versus 1.65, but higher relative average, 3.3 versus 2.4), a trend consistent with both EpCAM and Trop2 contributing to moderate contractility of the free cortex.

Differential effect of EpCAM KD and Trop2 KD on focal adhesions in individual cells

Another key parameter in cell migration is adhesion to the matrix. We thus examined cell-matrix adhesive structures using two markers, vinculin, and paxillin, specifically paxillin phosphorylated at residue Tyr118 (pPax). Both paxillin and vinculin are core cytoplasmic components participating in the complex linkage between integrins and the actin cytoskeleton. Vinculin is a well-established marker of mature/stable focal adhesions, recruited in a mechanosensitive manner. Paxillin is a more general marker of focal adhesion, but its phosphorylation is considered to be the signature of more nascent and/or dynamic adhesive structures (Tsubouchi *et al.*, 2002; Deakin & Turner, 2008; Parsons, Horwitz & Schwartz, 2010). Thus, the comparison of both markers should provide a read-out for the state of focal adhesions. Because of the high complexity of the patterns obtained with spheroids, we restricted this analysis

to single cells, where focal adhesion could be clearly identified. This had also the advantage to look at intrinsic matrix-adhesion in the absence of cell-cell interactions (Collins & Nelson, 2015). We measured levels of pPax and vinculin accumulating in focal adhesions, and we expressed the results both as total levels per cell (Fig.6B) and normalized to the cell surface (Fig.6C). One had indeed to keep in mind the strong differences in spreading caused by EpCAM or Trop2 depletion.

EpCAM depletion led to a significant increase in total pPax, and only a modest, non-significant increase for vinculin (Fig.6B). For both markers, the relative levels normalized to cell surface remained similar compared to siCtrl (Fig.6C), and the relative vinculin/pPax ratio was unchanged (Fig.6B,C). For Trop2 depletion, we also observed a small, non-significant increase in total vinculin signal, similar to EpCAM KD (Fig.6B). However, since cells had a much smaller contact area when normalized to the surface, vinculin was here strongly increased (Fig.6C). Furthermore, pPax was low in these cells, and the vinculin/pPax ratio climbed to twice that of controls (Fig.6B,C). In double KD, total vinculin was significantly higher (Fig.6B). Comparison with a single KD suggested that both depletions acted in an additive manner for this marker. The increase in vinculin was significant even when normalized to the cell surface (Fig.6C), despite the relatively large surface of these cells (Fig.1F). The vinculin/pPax ratio was also significantly elevated (Fig.6B,C). Altogether, this quantification of FA markers showed that spreading of EpCAM cells was only accompanied by mild changes in FA markers, their average intensities per surface area remaining similar to controls. The situation was different for Trop2 depletion: While this condition did not lead to any significant changes in total levels, both vinculin density, as well as the vinculin/pPax ratio, were strongly increased in these cells, which is consistent with Trop2 depletion stimulating reinforcement of matrix adhesion. These observations were interesting as they drew a parallel with EpCAM depletion, which seemed to preferentially trigger reinforcement

of cadherin-based cell-cell adhesion. Two striking results could be extracted from the double depletion phenotype: A strong vinculin recruitment indicative of an additive contribution of both depletions, and a high vinculin/pPax ratio, which was similar to Trop2 KD. These results seemed to indicate that matrix adhesion may be less sensitive to EpCAM KD than to Trop2 KD.

Traction and cell-cell forces are differentially elevated upon EpCAM and Trop2 depletion

Our IF results on pMLC, E-cadherin, and focal adhesion components were consistent with EpCAM and Trop2 depletions leading both to a global upregulation of cortical myosin contractility while having a different impact on cell-cell and cell-matrix adhesive contacts, EpCAM KD causing a stronger reinforcement of cadherin adhesions, Trop2 KD on matrix adhesions. This prompted us to measure the actual forces exerted on these different structures. We opted for an approach that used traction force microscopy (TFM) to measure forces exerted on cell doublets adhering to H micropatterns. This method had unique advantages for our purpose: The H patterns constrained the cell doublet to adopt a stable configuration, which provided the possibility to directly measure traction exerted on the matrix and, at the same time, to indirectly calculate the force at the cell-cell contact, simply based on the counterbalance of forces (Fig.7A)(Tseng *et al.*, 2012). A bonus from this technique was that the H pattern, by imposing a fixed geometry and a limited surface for the contacts to the matrix, at least partly solved a major challenge, i.e. the wide differences in cell morphology of cells on non-constrained collagen substrate, observed among wild type MCF7 cells, and further exacerbated by EpCAM and Trop2 depletions.

We found that EpCAM depletion caused a significant increase in both traction forces on the matrix, reflected as total contractile energy (Fig.7C,D), and in cell-cell forces, reflected as the Fcc magnitude (Fig.7E). These results directly demonstrated the increased contractility in EpCAM KD cells, consistent with the pMLC IF data. We also measured the length of cell-cell contacts,

which is a readout of adhesiveness (see below). siEpCAM doublets showed a significantly broader contact (Fig.7B,G). This was again consistent with the increased E-cadherin levels, observed both in small groups of cells and in spheroids, thus confirming that these cells respond to high tension by cadherin recruitment and reinforcement of cell-cell adhesion.

Trop2 KD also increased both total contractile energy and cell-cell forces (Fig.7C,D). However, contact length was only marginally increased (Fig.7B,G), supporting, together with the IF data, the notion that while both EpCAM and Trop2 depletions upregulate myosin-dependent cortical tension, in Trop2 KD cells the response through adhesion reinforcement was less efficient. As expected, double depletion also increased contractile energy, yielding the highest cell-cell force and highest cell-cell to total force ratio (Fig.7E,F).

Note that the collagen-coated acrylamide used as substrate in TFM has different properties than the fibrillar collagen substrate. We thus verified the collective migration of spheroids on this acrylamide-based substrate. We found essentially the same phenotypes as on fibrillar collagen gel: EpCAM depletion strongly increased spheroid spreading, while maintaining high tissue coherence. Interestingly, control spheroids appeared even less coherent, with a higher frequency of detachment of single cells or small groups of cells (Fig.7H).

EpCAM KD modulates cortical tension and adhesiveness in suspended cell doublets

While our TFM data provided a simultaneous quantitative measurement of cell-matrix and cell-cell adhesion, both types of adhesions are likely to influence each other. We thus also wanted to evaluate changes in cortical tension and cell-cell adhesiveness in the absence of cell-matrix adhesion. This is important both for a better dissection of the various parameters, but also because cells of a spheroid that are not in immediate contact with the matrix, will be obviously only exposed

to cell-cell contacts, and to free cortical tension for those located at the surface. Relative cortical tension and adhesiveness can be accurately measured simply based on the geometry of free cell doublets laid on a non-adherent surface (Winklbauer, 2015; Canty *et al.*, 2017; Fagotto, 2020). Such doublets adopt a typical configuration, where the cell contact expands until contact tension (T) is precisely balanced by the cortical tension at free edges (Ct) (Fig.8A). T is a global tension that results from the sum of the cortical tensions of each cell along the contact and the negative so-called adhesive tension, produced by the released energy due to cadherin trans interactions (Winklbauer, 2015). Note that besides adhesive tension, one major effect of cell-cell adhesion is to signal to the cell cortex and decrease its contractility. The resulting difference between cortical contractility along the contact compared to the free edge is a major force driving cell adhesion (Winklbauer, 2015). Thus, a high T/Ct corresponds to a poorly adhesive contact, a low T/Ct to a highly adhesive contact. A convenient way to express this relationship is the “adhesiveness” (Parent, Barua & Winklbauer, 2017), an absolute number between 0 and 1, which is inversely related to the T/Ct ratio. This balance of forces can be directly deduced from the geometry of the membranes at the contact vertex, from which one can extract important information: Given a doublet composed of cells A and B (Fig.8A’), measuring the angles between the three tangents to the membranes meeting at the vertex allows to calculate a relative value for the three tensions, Ct_A , Ct_B and T_{AB} (Canty *et al.*, 2017; Fagotto, 2020; Kashkooli *et al.*, 2021). In addition to the comparison between T and Ct , which gives the degree of adhesiveness, this geometry also provides a readout for relative Cts : A perfectly symmetric doublet indicates that both cells have an identical cortical tension $Ct_A = Ct_B$ (Fig.8A). If the doublet is composed of two cells with different cortical tensions ($Ct_A < Ct_B$), the softer cell (cell A) will tend to engulf the stiffer one (cell B), thus yielding

a curved cell-cell interface (Fig.8A'). Adhesiveness can also be directly calculated from the angle formed by the two free membranes (Parent, Barua & Winklbauer, 2017).

We adapted the protocol, previously used for embryonic cells (Rohani *et al.*, 2014; Canty *et al.*, 2017; Kashkooli *et al.*, 2021), to mildly dissociate MCF7 monolayers to single cells, which we mixed at low density to favor the formation of doublets (Fig.8B,C). We systematically produced doublets formed of two cells from the same condition (homotypic doublets), as well as doublets formed of a wild-type cell A and a manipulated cell B (heterotypic doublets).

Heterotypic doublets made of control cell A and an EpCAM KD cell B appeared asymmetric, cell A systematically engulfing wild-type cell B (Fig.8C). Consistently, the ratio between C_{tA} and C_{tB} was significantly lower than that of homotypic doublets, confirming that EpCAM KD increased cortical tension (Fig.8D). Comparing homotypic doublets, we found that EpCAM KD doublets had a less acute contact angle, corresponding to a strong, highly significant increase in adhesiveness (Fig.8B,E). These measurements demonstrated under conditions where cells were isolated from any other factor, EpCAM depletion led to both higher cortical tension and higher adhesiveness, the latter further confirming the occurrence of cadherin adhesion reinforcement.

Analysis of Trop2 KD and dKD cells were fully consistent with our other assays: The asymmetry of heterotypic doublets showed that cortical tension was increased. Although not significantly different from single depletions, dKD tended again to show the highest tension (lowest C_{tA}/C_{tB} , Fig.8C,D). As for adhesiveness measured from homotypic doublets, Trop2 KD caused only a mild, non-significant increase, while dKD increased it to a level similar to single EpCAM KD (Fig.8B,E).

Calculation of adhesiveness for heterotypic doublets (Fig.8E) provided another important, although puzzling, result: When confronting two populations with differences in adhesiveness/contact tension, the simplest model predicts that heterotypic adhesiveness should be of intermediate value (discussed in Canty et al, 2017 and Fagotto2020). However, we obtained, in all three heterotypic cases, i.e. siCtrl versus siEpCAM, siTrop2, or dKD, values that were significantly lower than those of control cells. The inescapable conclusion is that contact tension is highest at heterotypic contacts. One potential explanation is that adhesion reinforcement in response to increased cortical tension is only effective when tension is exerted symmetrically from both cells, while otherwise, the purely tensile effect predominates, resulting in weaker adhesion. As further addressed below, this situation of “high heterotypic interfacial tension”, or HIT, fulfills the unique condition to generate sorting of the two populations (Canty *et al.*, 2017; Fagotto, 2020).

Differential EpCAM and Trop2 localization

In an attempt to understand the cause for the observed differences in cellular behavior for EpCAM and Trop2 KD conditions, we compared the subcellular localization of the two proteins by immunofluorescence (Fig.9). We focused on small groups of cells, which offer a simple model to analyze various cell contacts. Both EpCAM and Trop2 were found all along the plasma membrane, i.e. at free cell edges, on the ventral side in contact with the collagen matrix, and at cell-cell contacts. In general, the EpCAM signal appeared more homogenous (Fig.9A), while Trop2 tended to be more punctate (Fig.9B). Furthermore, numerous intracellular Trop2 spots were found (Fig.9B, concave arrowheads), while they were much rarer for EpCAM. These spots appeared to be a mixture of organelles positive for Rab5, Rab7, or LAMP1, typical markers of different steps along the endosomal-lysosomal pathway (data not shown). In contrast to what was reported for isolated Caco2 intestinal cells (Gaston *et al.*, 2021), Rab35-positive fast recycling

endosomes were scarce in groups of MCF7 cells, and neither EpCAM nor Trop2 showed significant colocalization (data not shown).

The potentially most relevant difference in the localization of the two proteins was the relatively lower signal for Trop2 at cell-cell contacts compared to EpCAM (Fig.9A,B). We thus quantified both signals in four distinct subcellular sites: cell-cell contacts, free edges, and the bottom and the top of the cells. Relative average intensities were normalized to the signal at free edges (Fig.9B,C). We found that the relative Trop2 to EpCAM ratio was indeed significantly lower for the cell-cell contact site compared to the edges (Fig.9D). In addition, Trop2 tended to be also slightly more abundant at the ventral side. These overlapping yet distinct distributions are consistent with the functional data from KD experiments, indicating that, besides their common function in downregulation of contractility at the free edge, EpCAM may act more on a myosin pool engaged in controlling cell-cell contacts, while Trop2 would be more active toward myosin involved at the cell-matrix interface.

EpCAM positive and negative cells sort from mixed spheroids

Both differences in cortical tension/adhesiveness and migration are conditions that may lead to differential cell distribution and even cell sorting (Fagotto, Winklbauer & Rohani, 2014; Fagotto, 2020). Though differential adhesion/tension per se is not necessarily sufficient for sorting, HIT is most effective at generating this phenomenon (Canty *et al.*, 2017; Fagotto, 2020). Differences in migration (both intercellular motility and migration on the matrix) are also likely to impact the relative positioning of cell populations. Our characterization of wild-type and EpCAM depleted MCF7 cells (Figs.1 and 8) showed that such conditions were fulfilled in this system. Therefore, I determined the topography of spheroids made of mixed populations of cells expressing different EpCAM levels. Note that such settings are highly relevant for reputedly

heterogeneous tumors. For this purpose, we exploited the observation that, when using low amounts of siEpCAM RNA, we obtained a very heterogeneous depletion, yielding a mosaic population. Mixed spheroids were left to adhere and spread on the collagen matrix. EpCAM-positive cells were then detected by immunofluorescence (Fig.10). We used a custom-made ImageJ plugin to locate all nuclei and allocate positions to nuclei of EpCAM-positive cells. Most EpCAM-positive cells tended to be found in clusters, surrounded by cells with low or no EpCAM signal (Fig.10A). Focusing on those cells still expressing high EpCAM levels, our quantification of 9 spheroids showed that the majority of EpCAM-positive cells were significantly concentrated in the core of the spheroids (Fig.10B and C), consistent with these cells being left behind as EpCAM-depleted cells more effectively underwent collective migration. Note, however, that a few isolated or small groups of EpCAM-positive cells were found at the most periphery (Fig.3D).

Differences in polarization may account for differences in cell migration

Our migration data showed that the degree of cell spreading inversely correlated with single-cell migratory speed: EpCAM KD cells were flat and were almost immobile, while Trop2 KD cells were very compact and showed the fastest migration (Fig.2). The failure of siEpCAM cells to migrate could be explained by the fact that they emitted protrusions in all directions (Figs.2B,6A). siTrop2 cells, on the contrary, were rounded (Figs.2C,6A), and, based on vinculin staining, did seem to exert strong tension on focal adhesions (Figs.6A,B), which at first glance did not seem favorable conditions for migration. These cells did emit protrusions, however (Figs.2C,6A). We thus hypothesized that Trop2 KD cells, despite their compact morphology, may tend to be more polarized than siEpCAM KD cells. We evaluated the general geometry of the cell-matrix contacts of cells plated on collagen gel and immunolabelled for p-paxillin and vinculin (Fig.S4). Cells displayed a wide range of configurations, which we compiled into five categories

reflecting both degrees of spreading and symmetry/asymmetry, as indicated in Fig.S4F, with examples provided in Fig.S4A-E. siCtrl cells displayed a wide range of spreading, but the majority of them had focal adhesions distributed in asymmetric patterns (Fig.S4C). siEpCAM cell, on the other hand, did not only spread quite extensively, but the focal adhesions were, in the large majority of cases, arranged in strikingly symmetric patterns, either forming a triangle or a simple polygon (Fig.S4A) or distributed radially for cells emitting a large number of circumferential protrusions (Fig.S4B). On the contrary, siTrop2 cells were typically compact, with focal adhesion markers typically arranged as a ring (Fig.S4D). However, a closer look at the focal adhesion showed that there was very often an imbalance, one side showing clear enrichment compared to the other (Fig.S4D, filled and concave arrowheads). Furthermore, siTrop2 cells often displayed a dominant protrusion, thus adopting a unipolar configuration (Fig.S4E). These geometries were in good agreement with the migratory behavior of single cells: In particular, the fact that siTrop2 round cells tended to show a break in symmetry in their matrix contacts accounted for their relatively high migration capacity.

Discussion

Our study sheds light on the impact of EpCAM expression/repression on the morphogenetic properties of MCF7 spheroids, used as an in vitro proxy for a solid 3D breast cancer tissue. Our data revealed strong effects on spreading/collective migration and tissue cohesion/adhesiveness. The results are interesting in the context of tumors and metastasis: The relative low cohesiveness of control, EpCAM-positive tissue could correspond to the situation in a primary tumor, where high EpCAM expression would contribute to the escape of individual cancer cells or small groups of cells. On the contrary, the high cohesiveness of EpCAM KD spheroids drastically lowers the chances for the cells to detach and leave the “primary sites”. At the same time, however, EpCAM

KD spheroids show a strikingly higher capacity to spread as a whole population, reaching almost the maximal possible extension (close to a monolayer).

The cohesive migration of EpCAM KD cells is unexpected and rather counter-intuitive when one considers the fact that EpCAM depletion causes a clear increase in actomyosin contractility. High cortical contractility is generally associated with higher tissue stiffness, lower adhesion, lower intercalation, globally decreased motility and even loss of tissue coherence and integrity. These were precisely the type of phenotypes that have been observed by us and others, in early embryonic tissues (Slanchev *et al.*, 2009; Maghzal *et al.*, 2010, 2013) as well as in the intestine, where pathological lack of EpCAM causes strong damages of the epithelium (Salomon *et al.*, 2017).

In the case of breast cancer MCF7 cells, we have shown here through a variety of assays, that EpCAM depletion leads indeed an upregulation of myosin activity and cortical contractility, and the calphostin and blebbistatin rescue experiments have confirmed that the increased collective migration does result from derepression of the “canonical” nPKC-myosin pathway (Maghzal *et al.*, 2013). How does one then reconcile the strikingly opposite tissue phenotypes when compared to embryonic tissues? A key element of an explanation is the ability of MCF7 cell groups to react to increased tension by recruiting additional cadherin at cell contacts. Note that while the standard model of balance between cortical tension and contact tension (Brodland, 2002; Maître & Heisenberg, 2013) is often interpreted as a simple antagonistic relationship between contractility and adhesion, this reductionist view does not take into account the ability of the system to adapt cadherin to tension, although there is ample evidence for this process (Charras & Yap, 2018). This is precisely what appears to happen in EpCAM-depleted MCF7 cells.

In the embryo model, the loss of tissue integrity upon EpCAM depletion is most likely due to an extreme deregulation of contractility, which cadherin adhesions were simply not able to cope with. We do not know the precise reason for these different outcomes, and *Xenopus* embryonic cells and tissues and mammalian culture cells are much too different to undertake meaningful quantitative comparisons, but obvious potential causes are plentiful, including differences in any aspect of the pathway, from EpCAM levels and turnover, basal nPKC activity, to intrinsic cortical properties, quantitative and/or qualitative differences in mechanosensitive reactions coupling myosin and cadherin recruitment, among many others.

In the case of breast cancer MCF7 spheroids, the system appears to be set in such a way that, upon EpCAM depletion, the tissue acquires properties, increased cohesiveness, high traction on the substrate, and intercalation within the tissue, which are precisely those that together would be predicted to favor collective migration. As for wild-type spheroids, this should not be viewed as a tissue impaired in collective migration, as it still spreads at a quite good pace. It is however slightly less tensile and less coherent, which decreases its ability to spread as a homogenous tensile unit. The fact that cells can escape and migrate away can be explained both by this decreased cohesiveness of the tissue and by the ability of single cells to migrate on the matrix. The fact that this ability was largely lost for single EpCAM KD cells could be explained by the combination of high spreading and increased cortical tension: With radially oriented protrusions, cells pull symmetrically on the matrix, in all directions, which obviously results in their immobilization. In addition, moving actively as a collective group is readily accounted for by the reason that cells at the edge of the group will unavoidably be polarized, simply due to the asymmetry built between a free edge in the front and cell-cell contacts on the rear (and the sides). Such asymmetry is a well-established feature of collective migration (Mayor and Etienne-Manneville, 2016).

A major unexpected result of this study is the diametrically opposite phenotypes observed upon depletion of Trop2. This represents the first evidence for a functional difference between EpCAM and Trop2. So far, each feature that had been characterized for EpCAM had been also found in Trop2 (Fagotto & Aslemar, 2020). As both proteins are co-expressed in all epithelia, except for the intestine, it had been assumed that they were redundant. Such redundancy explained the lack of embryonic phenotype in the mouse of EpCAM and Trop2 single KOs (Wang *et al.*, 2011; Guerra *et al.*, 2012; Lei *et al.*, 2012), as opposed to the strong phenotypes observed in *Xenopus* embryos, which have only EpCAM. Consistently, EpCAM KO in mice mimicked human CTE, thus affecting precisely the intestine, which expresses only EpCAM (Guerra *et al.*, 2012; Lei *et al.*, 2012). It is thus likely that in normal development and healthy adult tissues Trop2 compensates for the loss of EpCAM, as shown by the partial rescue of the CTE phenotype by Trop2 (Nakato *et al.*, 2020). We have confirmed here that at the global cellular level, Trop2 KD also has similar effects as EpCAM on actomyosin contractility (increased pMLC levels, increased traction and cell-cell forces, increased cortical tension), and also induced cadherin reinforcement. The key to the observed phenotypic differences has to be found in subtle differences in relative levels and subcellular distribution of tension (Figs 9 and 5.), as well as the lower reinforcement of E-cadherin (Fig.4D). This latter most likely relates to the markedly paucity of Trop2 at cell-cell contacts. The underlying mechanism remains to be investigated. We know that Trop2 is not necessarily excluded from contacts, as when overexpressed, it distributed quite homogeneously all along the plasma membrane, very much like EpCAM (data not shown). We thus hypothesize that these differences reflect more subtle preferred locations, for example in association with membrane microdomains (Kuhn *et al.*, 2007; Nubel *et al.*, 2009). Note that while Trop2 clearly does not compensate for EpCAM in our acute loss-of-function experiments, it may well do so on

a longer time scale. It remains that these experiments have revealed that these two molecules have clear non-redundant roles, providing the first evidence of a functional specialization of these two twin genes. Note also that EpCAM regulation seems to dominate the properties of these breast cancer cells since the double KD yields similar single and collective migration phenotypes as EpCAM KD. More generally, this study presents a fascinating example that highlights the importance of a precise tensile balance between free cortex, cell-cell contacts, and matrix adhesions, and how changing this balance can lead to such dramatically different tissue properties.

Lastly, the observed sorting in mixed spheroids is another interesting finding, also directly relevant to cancer development. At first, sight, sorting under these conditions was unexpected. Indeed, we had previously demonstrated that cell sorting from a mixed population cannot be driven purely by differences in cortical contractility or cell-cell adhesion (Canty *et al.*, 2017). We and others had shown the required condition for efficient sorting was a higher tension at contacts between the two cell types (Canty *et al.*, 2017; Fagotto, 2020). The juxtaposition of cells with different degrees of contractility should produce contacts with an intermediate tension, thus not a HIT situation (Canty *et al.*, 2017). The key observation that explained sorting here came from the analysis of heterotypic doublets, which showed a contact tension unexpectedly higher (thus adhesiveness lower) than that of the corresponding homotypic doublets (Fig.8D). While this case was not predicted by the original differential interfacial hypothesis (Brodland, 2002), it is consistent with reinforcing reactions contributing to set tension/adhesion at a cell contact. We saw that EpCAM depletion induced such reinforcement, yielding at the same time in higher Ct (inferred from heterotypic doublets) AND higher homotypic adhesiveness. In the case of heterotypic doublets, however, it makes perfect sense that reinforcement may fail, as the softer cell (here wild-type cells) were unable to respond sufficiently to the traction from the more contractile cell. This

is to our knowledge the first explicit model of sorting based on this principle. Note that mixed spheroids were quite heterogeneous, with a large range of EpCAM expression. The fact that high EpCAM cells were clustered but not fully segregated by smooth interfaces is explained by considering the broad range of tensile/adhesive properties within these cell masses. Considering this heterogeneity, which we purposely produced to mimic realistic situations, the simple fact of observing sorting is in itself most remarkable. This study proposes a conceptual framework that could be widely used to account for potential similar unmixing phenomena in actual tumors.

Material and methods

Cell culture. MCF7 cells, originally acquired from ATCC, were provided by the SIRIC Montpellier center. They were grown in Dulbecco's modified Eagle's medium (DMEM) 4.5 g/l glucose supplemented with 10% fetal bovine serum (FBS), 1% antibiotic-antimycotic, and 1% non-essential amino acids (all Gibco, Thermo Fischer Scientific) for a maximum of 10 passages. 0.05% Trypsin-EDTA (Gibco, Thermo Fischer Scientific) was used for dissociation. Cells were cultured at 37°C and 5% CO₂. For all experiments except TFM, cells were plated on top of a thin gel layer of fibrillar collagen type-I (Corning rat tail collagen-I, #354236) prepared from a ~ 3 mg/ml solution. EpCAM and Trop2 KD were carried out using commercial siRNAs from Santa Cruz Biotechnology, #sc-43032 and #sc-72392, respectively. A mix of both siRNAs was used for the double KD of both proteins, and a non-targeting siRNA (Santa Cruz Biotechnology, #sc-37007) was used as a negative control.

Antibodies, reagents, and solutions. The following antibodies were used. Unless stated otherwise, dilutions are for immunofluorescence. Mouse monoclonal antibody directed against EpCAM (323/A3 #sc-73491, Sant Cruz Biotechnology, 1:1200). Rabbit monoclonal antibody

directed against Trop2 (#MA5-29593, Invitrogen, 1:800). Rabbit monoclonal antibody directed against E-cadherin (24E10, #3195, Cell Signaling Technology, 1:400). Mouse monoclonal antibody directed against phospho-myosin light chain (Ser19) (#3675, Cell Signaling Technology, 1:200). Mouse monoclonal antibody against vinculin (#MA1103, Boster Bio, 1:100). Alexa Fluor 647-conjugated phalloidin (#A22287, Invitrogen, 1:200). Hoechst 33342 (#H3570, Invitrogen/Thermo Fischer Scientific, 1:2000). Alexa Fluor 488-conjugated anti-mouse (#A-21202) and anti-rabbit (#A-21206) Alexa Fluor 594-conjugated anti-mouse (#A-11005) and anti-rabbit (#A-21207) were all from Invitrogen/Thermo Fischer Scientific and diluted 1:200.

Fixation buffer: PFA 3.7%, was prepared by diluting PFA 32% aqueous solution (#15714, Electron Microscopy Sciences) in PHEM buffer, pH 7.0; 50ml of 2X PHEM buffer was made by diluting 1.814g PIPES, 0.65g HEPES, 0.38g EGTA and 0.099g MgSO₄ in distilled water. The pH was adjusted using 10M KOH. Methylcellulose solution used for spheroid formation: for 100 ml of the solution, 1.2g methylcellulose (#M0262-100G) was autoclaved in a bottle containing a magnetic bar. Then the powder was dissolved using 50 ml of culture medium (without additives) preheated at 60°C on a stirrer for 1 hour. 50 ml of culture medium containing 2X additives were then added and the solution was stirred at 4°C overnight. This was used to produce a 20% methylcellulose solution in culture medium, which was filtered using disposable syringe filters (#146561, pore size 0.45 µm, Clear Line). Mitomycin C used in proliferation-prevented migration assays was from Sigma (#M4287-2MG). EdU click reaction kit (#C10337, Click-iT EdU Cell Proliferation Kit for Imaging, Alexa Fluor 488 dye, Invitrogen, Thermo Fisher Scientific). EdU was fluorescently labeled in fixed cells using this kit that allowed copper-catalyzed covalent reaction between an azide that was coupled to Alexa Fluor 488 and an alkyne found in ethynyl moiety of EdU. Dissociation buffer: for 50 ml buffer, 4.4 ml of NaCl 1M solution, 50 µl of KCl

1M solution, and 0.04 grams of NaHCO₃ powder were diluted/dissolved in 41.2 ml of distilled water. The pH of the buffer was adjusted using 1 drop of NaOH 5N solution. Tris-buffered saline (TBS 1X) contained 100mM Tris-HCl pH 7.2 and 150 mM NaCl as final concentration.

Collagen gel preparation. The dishes were cleaned with a plasma cleaner machine for 5 min before coating with collagen solution. This step created a hydrophilic surface that let the collagen solution spread evenly making a thin layer of the collagen gel. It also helped the gel to stick to the glass surface better. After cleaning, dishes were kept for 5 min at 4 °C, and then the prepared collagen solution was added to them. This solution was prepared using the manufacture recipe. The collagen stock solution is in 0.02 N acetic acid. To obtain a gel state its pH must be brought to alkalinity using NaOH solution. To prepare collagen solution with alkaline pH, the final required volume for the 2mg/ml concentration of the collagen solution was defined, then based on this and the concentration of the stock solution, the amounts of collagen, 10X PBS, NaOH 1N, and culture medium were calculated. Using aseptic techniques all the reagents except collagen were mixed in a tube that was placed on ice. These reagents were mixed well and then the calculated volume of collagen was added to the rest and mixed to obtain an even solution. A moderate pink color was usually indicative of the right pH of the solution. This solution was kept on ice until used (maximum 1 hour). 10 µl of it were used to cover the surface of each glass-bottom culture dish (Cellvis, #D35-20-1.5-N) and then incubated at 37 °C for 30 min for gelling. At end of 30 min incubation, 1.5 ml of complete culture medium was added to each dish to be subsequently used for cells/spheroids plating. Glass-bottom dishes were replaced with coverslips (Diameter: 12 mm) for single-cell/groups of cells immunostaining, following the same procedure.

siRNA transfection. Cells were seeded at 0.2×10^6 density one day before siRNA transfection. The transfection mixture was prepared by diluting 40 picomoles of each siRNA into

200 μ l of Opti-MEM reduced serum media (Gibco, Thermo Fischer Scientific) as solution A and 4 μ l of transfection reagent (Invitrogen, Lipofectamine RNAiMAX, Thermo Fischer Scientific) into 200 μ l of Opti-MEM reduced serum media as solution B. Both solutions were kept for 5 min at room temperature (RT) and then solution A was added to solution B, the new mix was shortly vortexed and left for 15 min at RT. Within this time, cells' medium was replaced with 1.6 milliliters of fresh medium, and then the transfection mix was added to them drop by drop to ensure an even distribution of the whole dish surface. In the case of double KD, 40 picomoles of each siRNA thus a total of 80 picomoles and 8 μ l of the transfection reagent were used. Very efficient KD was obtained after 48 to 72 hours.

Spheroid forming. The siRNA transfected cells were dissociated by trypsin and counted using the trypan-blue solution and an automated cell counter (Invitrogen, Countess). A cell suspension of 400 cells per 200 μ l in complete culture medium containing 20% methylcellulose was prepared from the dissociated cells. 200 μ l of the cell suspension were added to each well of round-bottom 96-well plates (Greiner Bio-One Cellstar, #650185) and the plate was kept in the incubator for 48 hours. Well-formed cell spheroids of the roughly same size were obtained.

Spheroid migration assay. The formed spheroids were transferred to glass-bottom dishes coated with collagen gel using a 200 μ l pipette; the pipette tips were cut at about 0.5 centimeters from their tip, each spheroid was taken from each well by going to the very bottom of the well and the medium was poured off as much as possible without losing the spheroid from the tip. It then was transferred to the dish with collagen gel. Spheroids adhered to the gel within 30 min. They were imaged with an Olympus IX83 inverted widefield video-microscope controlled by Metamorph software using a 10x (0.3NA) or 20x (0.45NA) objective. Time-lapse images were acquired every 30 min for 24hrs.

Single-cell migration assay. Cells were dissociated and counted as above, 0.5×10^5 cells were plated on top of collage gel and left to adhere for about 3 hours before acquisition of time lapses. Imaging conditions were as for spheroids, with a 20x (45NA) objective, every 5 to 7 min for 6 hours.

Immunostaining. Single cells and cell spheroids were fixed in 3.7% paraformaldehyde (PFA), preheated to 37 °C . The fixation duration was 10 and 30 min for single cells/groups of cells and spheroids, respectively. The samples were then permeabilized with 1% Triton 100X in PBS for 10 min (single cells/groups of cells) or 30 min (spheroids). This step was followed by rinsing the samples with TBS 1X, pH 7.2 for 5 min, twice, then a 45 min incubation at room temperature (RT) in blocking buffer (20% sheep serum (#S2350-100, Biowest) in PBS). Primary antibody solutions of desired concentrations were prepared in 10% sheep serum in PBS, centrifuged for 10 min at 10000 rpm to remove potential precipitates, and added to cells/spheroids. Cells/groups of cells were incubated with the primary antibodies for 2 hours at RT, and spheroids overnight at 4 °C. After three rinses with PBS or TBS (5 min for single cells/groups, 10 min for spheroids), samples were incubated for 1 hour at RT (single cells/groups) or overnight at 4 °C (spheroids) with secondary fluorescent antibodies of the Alexa series (Molecular Probes/ Thermo Fischer Scientific), combined with the nucleic acid stain Hoechst the samples for 1 hour at RT and, then rinsed again. When coverslips were used (only in the case of single cells/groups of cells), they then were mounted using an antifade mounting media (Slowfade, #S36972, Invitrogen, Thermo Fischer Scientific). Spheroids were always stained in glass-bottom dishes; they were kept in PBS after the rinsing step without being mounted and imaged within a maximum of two days. Labeling actin cytoskeleton was carried out by adding phalloidin coupled to Alexa 647 to the primary antibody solution.

Traction force microscopy. Stock solution for soft polyacrylamide substrates of 5 kPa rigidity containing far-red fluorescent nanobeads (Bangs laboratory, #FC02F, 0.19 μm) were prepared by mixing acrylamide 40% (#A4058, Sigma) and bis-acrylamide 2% (#M1533, Sigma) in DPBS 1X (PBS without Ca & Mg, #CS1PBS01-01, eurobio SCEIENTIFIC) according to documented relative concentrations (Tse & Engler, 2010; Vignaud, Ennomani & Théry, 2014). The thin polyacrylamide-based substrate was polymerized between two different glass coverslips (#631-0162, diameter 32 millimeters, thickness No. 1, VWR) prepared as follows. The first coverslip served as a base for the gel. It was cleaned using a plasma cleaner machine, then coated with bind-silane (#GE17-1330-01, Sigma) for 3 to 5 min at RT to ensure the attachment of the gel to the coverslip. The second coverslip served to transfer the patterned extracellular matrix. It was also first plasma cleaned, then coated with 0.1 mg/ml PLL-PEG (#PLL20K-G35-PEG2K, JenKem Technology) solution for 30 min at RT to obtain a passivated surface. It was then washed with distilled water, dried, then burned for 5 minutes with UV light through a micropatterning chrome photomask (45 by 45 μm custom-designed H shapes, micropatterned onto chrome photomask by Toppan). This allowed adsorption of the collagen coating at the burned sites resulting in a micropatterned coated coverslip. Collagen type-I was added at 0.5mg/ml in 0.02 N acetic acid and left for 45 min at RT. For gel polymerization, 1 μl of 10% ammonium persulfate (#A3678, Sigma), 1 μl of TEMED (#T9281, Sigma) and 3.5 μl of the above-mentioned nanobeads were added to 165 μl of the acrylamide-bisacrylamide stock solution. 47 μl of this solution was used to put between the two coverslips for polymerization (30 min, RT). Once polymerized, the collagen-coated top coverslip was gently removed, exposing the collagen H micropatterned gel. Cells were plated on this substrate a at density of 0.5×10^5 per coverslip in a culture dish. The medium of each dish was replaced with fresh medium to wash out cells that didn't adhere to the substrate (this step avoided

ending up with cells that were not on the patterns). The dishes were then kept in the incubator overnight to allow cell division to obtain cell doublets on each H micropattern. Cells and the underneath nanobeads were imaged using an epifluorescence inverted microscope (Nikon Ti2-E2) with an Orca Flash 4.0 sCMOS camera (Hamamatsu) and a temperature control system set at 37°C controlled by the NIS element software. 40x objective (1.3NA/oil) was used as magnification. This first image served as the stressed (pulled) state of the beads. Then the cells were removed from the patterns using trypsin and another image of the same position was taken serving as the unstressed state of the beads. The displacement field analysis was done using a custom-made algorithm based on the combination of particle image velocimetry and single-particle tracking. Traction forces were calculated using Fourier transform traction cytometry with zero-order regularization (Sabass *et al.*, 2008; Milloud *et al.*, 2017). All calculations and image processing were performed using MATLAB software.

Analysis of cell doublets. To prepare isolated doublets, siRNA transfected cells were washed with 37 °C-preheated DPBS 1X twice, followed by 10 min incubation at RT with dissociation buffer. The buffer was then removed, 1 ml of complete culture medium was added to the cells, and single dissociated cells were obtained by pipetting. Control cells were labeled with Hoechst for 10 min before wash steps and dissociation. The dissociated cells were transferred to agarose-coated plastic dishes (2% agarose in PBS 1X) and put in the incubator for 10 to 15 min to allow partial reassociation in order to obtain mainly doublets. In the case of heterotypic doublets, the dissociated cells of two different conditions were gently mixed before transfer to agarose-coated plastic dishes. During reassociation time, 1 µl of a membrane dye (CellMask deep red plasma membrane stain, #C1004, Invitrogen) was added to 10 ml of complete culture medium and 1 ml of this medium was added to a glass-bottom dish that had been coated with a thin layer of 2%

agarose (50 μ l of 2% agarose added on the plasma cleaned solution glass coverslip of the dish). The reassociated cells were gently transferred to this dish and imaged by live confocal microscopy using a Dragonfly spinning disk (Andor) mounted on a Nikon inverted microscope, a 20x (0.75 NA/oil) objective. The microscope was equipped with temperature and CO₂ control. Dual-color images were obtained simultaneously using two CCD cameras (EMCCD iXon888 Life Andor). A 10 by 10 tile-scan setting was used to scan a large surface and enabling the hardware perfect focus system (PFS) allowed obtain in focus doublets within the 15 μ m z-stacks. Quantifications of the relative cortical tensions and adhesiveness were done based geometry of cell membranes at the vertices that reflect the equilibrium between forces of free cell membranes and the contact. To do so, each free membrane and the contact were defined by three points including the cell vertex. Tangents to these defined arcs were considered as their corresponding force vectors and their angles to the horizontal axis were used in two equilibrium equations to calculate the ratio between cortical tensions and the ratio between each cortical tension and the contact tension. The ratio between cortical tensions for a doublet of the same type (homotypic doublet) is expected to be on average 1 that will change in a heterotypic doublet that is used as a direct readout of relative cortical tension of the two cell types. Adhesiveness was calculated directly from the contact angle (Kashkooli *et al.*, 2021).

Confocal microscopy of fixed samples. Images of immunostained single cells/groups of cells were acquired using inverted scanning confocal microscope (Leica SP5-SMD), 63x objective (1.4NA/oil) as z-stacks (0.2 to 0.4 μ m distance between planes). Images of fixed immunostained spheroids were acquired using Nikon inverted microscope coupled to the Andor Dragonfly spinning disk, 40X water objective (1.15NA/water) with 0.6 mm working distance and as z-stacks

of 4 by 4 tile-scans (1µm distance between planes), which were stitched to generated full images of whole spheroids.

Establishment of stable lines expressing EpCAM-GFP. MCF7 cells were transfected with EpCAM-GFP plasmid that was constructed using pSBtet-pur vector (Plasmid #60507, Addgene). This vector contains tetracycline-dependent inducible promoter upstream of the inserted gene (here EpCAM-GFP) and constitutively expresses the puromycin resistance gene. After 48hrs the transfected cells were kept under puromycin (1µg/ml) (Gibco, life technologies, #A11138-03) as a selectable marker for two weeks. The medium was replaced by a puromycin-containing fresh medium to discard the non-resistant dead cells for the first two days and then at the end of the first week. After two weeks small colonies of cells were obtained that were dissociated and replated as the pool of resistant cells to induce faster cell growth. The medium still contained puromycin. After obtaining a monolayer, backups of the cells were made and kept in liquid nitrogen. These cells were used to induce EpCAM in order to study EpCAM overexpressing MCF7 cells. The medium did not contain puromycin when the cells were used in experiments. Doxycycline hyclate (from tetracycline antibiotic family, concentration 1 to 3 µM) (#D9891, Sigma) was used to induce the expression.

Statistical analysis. All statistical analyses were performed using Excel real statistics add-in. Unless otherwise stated experiments were replicated at least three times and comparisons between conditions were done using one-way ANOVA followed by Tukey-HSD post hoc test. p values met the following criteria *p < 0.05, **p < 0.01, and ***p < 0.001 and NS, not significant.

References

Baeuerle, P. A. and Gires, O. (2007) 'EpCAM (CD326) finding its role in cancer', *British Journal of Cancer*, 96(3), pp. 417–423. doi: 10.1038/sj.bjc.6603494.

Balzar M, Winter MJ, de Boer CJ, L. S. (1999) ‘The biology of the 17 – 1A antigen (Ep-CAM)’, pp. 699–712.

Barth, A. I. M., , Honesty, K. and Riedel-kruse, I. H. (2018) ‘Regulation of epithelial migration by epithelial cell adhesion molecule requires its Claudin-7 interaction domain’, pp. 1–24.

Brodland, G. W. (2002) ‘The Differential Interfacial Tension Hypothesis (DITH): a comprehensive theory for the self-rearrangement of embryonic cells and tissues.’, *Journal of biomechanical engineering*, 124(2), pp. 188–197. doi: 10.1115/1.1449491.

Canty, L. *et al.* (2017) ‘Sorting at embryonic boundaries requires high heterotypic interfacial tension.’, *Nature communications*, 8(1), p. 157. doi: 10.1038/s41467-017-00146-x.

Charras, G. and Sahai, E. (2014) ‘Physical influences of the extracellular environment on cell migration.’, *Nature reviews. Molecular cell biology*, 15(12), pp. 813–824. doi: 10.1038/nrm3897.

Charras, G. and Yap, A. S. (2018) ‘Tensile Forces and Mechanotransduction at Cell-Cell Junctions.’, *Current biology : CB*, 28(8), pp. R445–R457. doi: 10.1016/j.cub.2018.02.003.

Chaves-Pérez, A. *et al.* (2013) ‘EpCAM regulates cell cycle progression via control of cyclin D1 expression’, *Oncogene*, 32(5), pp. 641–650. doi: 10.1038/onc.2012.75.

Collins, C. and Nelson, W. J. (2015) ‘ScienceDirect Running with neighbors : coordinating cell migration and cell – cell adhesion’, *Current Opinion in Cell Biology*, 36, pp. 62–70. doi: 10.1016/j.ceb.2015.07.004.

Dai, X. *et al.* (2017) ‘J o u r n a l o f C a n c e r Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping’, 8. doi: 10.7150/jca.18457.

Deakin, N. O. and Turner, C. E. (2008) ‘Paxillin comes of age’, *Journal of Cell Science*, 121(15),

pp. 2435–2444. doi: 10.1242/jcs.018044.

Ebright, R. Y. *et al.* (2020) ‘Deregulation of ribosomal protein expression and translation promotes breast cancer metastasis.’, *Science (New York, N.Y.)*, 367(6485), pp. 1468–1473. doi: 10.1126/science.aay0939.

Fagotto, F. (2020) ‘Cell sorting at embryonic boundaries.’, *Seminars in cell & developmental biology*. England, pp. 126–129. doi: 10.1016/j.semcd.2020.07.012.

Fagotto, F. and Aslemar, A. (2020) ‘EpCAM cellular functions in adhesion and migration, and potential impact on invasion: A critical review’, *Biochimica et Biophysica Acta - Reviews on Cancer*, 1874(2), p. 188436. doi: 10.1016/j.bbcan.2020.188436.

Fagotto, F., Winklbauer, R. and Rohani, N. (2014) ‘Ephrin-Eph signaling in embryonic tissue separation.’, *Cell adhesion & migration*, 8(4), pp. 308–326. doi: 10.4161/19336918.2014.970028.

Gaber, A. *et al.* (2018) ‘EpCAM homo-oligomerization is not the basis for its role in cell-cell adhesion’, *Scientific Reports*, 8(1), pp. 1–16. doi: 10.1038/s41598-018-31482-7.

Gao, J. *et al.* (2014) ‘Epithelial-to-Mesenchymal Transition Induced by TGF- β 1 Is Mediated by AP1-Dependent EpCAM Expression in MCF-7 Cells’, (November 2013). doi: 10.1002/jcp.24802.

Gao, J. *et al.* (2015) ‘By inhibiting Ras / Raf / ERK and MMP-9 , knockdown of EpCAM inhibits breast cancer cell growth and metastasis’, 6(28).

Gaston, C. *et al.* (2021) ‘EpCAM promotes endosomal modulation of the cortical RhoA zone for epithelial organization’, *Nature Communications*, 12(1), pp. 1–22. doi: 10.1038/s41467-021-22482-9.

Guerra, E. *et al.* (2012) ‘mTrop1/Epcam Knockout Mice Develop Congenital Tufting Enteropathy

through Dysregulation of Intestinal E-cadherin/ β -catenin', *PLoS ONE*, 7(11). doi: 10.1371/journal.pone.0049302.

Insua-Rodríguez, J. and Oskarsson, T. (2016) 'The extracellular matrix in breast cancer', *Advanced Drug Delivery Reviews*, 97, pp. 41–55. doi: 10.1016/j.addr.2015.12.017.

Kashkooli, L. *et al.* (2021) 'Ectoderm to mesoderm transition by down-regulation of actomyosin contractility.', *PLoS biology*, 19(1), p. e3001060. doi: 10.1371/journal.pbio.3001060.

Keller, L., Werner, S. and Pantel, K. (2019) 'Biology and clinical relevance of EpCAM', *Cell Stress*, 3(6), pp. 165–180. doi: 10.15698/cst2019.06.188.

Koch, C. *et al.* (2020) 'Characterization of circulating breast cancer cells with tumorigenic and metastatic capacity.', *EMBO molecular medicine*, 12(9), p. e11908. doi: 10.15252/emmm.201911908.

Kramer, N. *et al.* (2013) 'In vitro cell migration and invasion assays', *Mutation Research - Reviews in Mutation Research*, 752(1), pp. 10–24. doi: 10.1016/j.mrrev.2012.08.001.

Kuhn, S. *et al.* (2007) 'A Complex of EpCAM , Claudin-7 , CD44 Variant Isoforms , and Tetraspanins Promotes Colorectal Cancer Progression', 5(June), pp. 553–568. doi: 10.1158/1541-7786.MCR-06-0384.

Leckband, D. E. and de Rooij, J. (2014) 'Cadherin adhesion and mechanotransduction.', *Annual review of cell and developmental biology*, 30, pp. 291–315. doi: 10.1146/annurev-cellbio-100913-013212.

Lei, Z. *et al.* (2012) 'EpCAM contributes to formation of functional tight junction in the intestinal epithelium by recruiting claudin proteins', *Developmental Biology*, 371(2), pp. 136–145. doi:

10.1016/j.ydbio.2012.07.005.

Lenárt, S. *et al.* (2020) ‘Trop2: Jack of All Trades, Master of None.’, *Cancers*, 12(11). doi: 10.3390/cancers12113328.

Litvinov, S. V. *et al.* (1994) ‘Ep-CAM: A human epithelial antigen is a homophilic cell-cell adhesion molecule’, *Journal of Cell Biology*, 125(2), pp. 437–446. doi: 10.1083/jcb.125.2.437.

Maetzel, D. *et al.* (2009) ‘Nuclear signalling by tumour-associated antigen EpCAM’, *Nature Cell Biology*, 11(2), pp. 162–171. doi: 10.1038/ncb1824.

Maghzal, N. *et al.* (2010) ‘The tumor-associated EpCAM regulates morphogenetic movements through intracellular signaling’, *Journal of Cell Biology*, 191(3), pp. 645–659. doi: 10.1083/jcb.201004074.

Maghzal, N. *et al.* (2013) ‘EpCAM Controls Actomyosin Contractility and Cell Adhesion by Direct Inhibition of PKC’, *Developmental Cell*, 27(3), pp. 263–277. doi: 10.1016/j.devcel.2013.10.003.

Maître, J.-L. and Heisenberg, C.-P. (2013) ‘Three functions of cadherins in cell adhesion.’, *Current biology : CB*, 23(14), pp. R626-33. doi: 10.1016/j.cub.2013.06.019.

Mayor, R. and Etienne-Manneville, S. (2016) ‘The front and rear of collective cell migration.’, *Nature reviews. Molecular cell biology*, 17(2), pp. 97–109. doi: 10.1038/nrm.2015.14.

Milloud, R. *et al.* (2017) ‘ $\alpha\beta3$ integrins negatively regulate cellular forces by phosphorylation of its distal NPXY site.’, *Biology of the cell*, 109(3), pp. 127–137. doi: 10.1111/boc.201600041.

Nakato, G. *et al.* (2020) ‘Amelioration of Congenital Tufting Enteropathy in EpCAM (TROP1)-Deficient Mice via Heterotopic Expression of TROP2 in Intestinal Epithelial Cells.’, *Cells*, 9(8).

doi: 10.3390/cells9081847.

Pan, M. *et al.* (2018) *EpCAM ectodomain EpEX is a ligand of EGFR that counteracts EGF-mediated epithelial-mesenchymal transition through modulation of phospho-ERK1/2 in head and neck cancers*, *PLoS Biology*. doi: 10.1371/journal.pbio.2006624.

Parent, S. E., Barua, D. and Winklbauer, R. (2017) ‘Mechanics of Fluid-Filled Interstitial Gaps. I. Modeling Gaps in a Compact Tissue’, *Biophysical Journal*, 113(4), pp. 913–922. doi: 10.1016/j.bpj.2017.06.062.

Parsons, J. T., Horwitz, A. R. and Schwartz, M. A. (2010) ‘Cell adhesion: Integrating cytoskeletal dynamics and cellular tension’, *Nature Reviews Molecular Cell Biology*, 11(9), pp. 633–643. doi: 10.1038/nrm2957.

Rao, C. G. *et al.* (2005) ‘Expression of epithelial cell adhesion molecule in carcinoma cells present in blood and primary and metastatic tumors’, *International Journal of Oncology*, 27(1), pp. 49–57. doi: 10.3892/ijo.27.1.49.

Rohani, N. *et al.* (2014) ‘Variable combinations of specific ephrin ligand/Eph receptor pairs control embryonic tissue separation.’, *PLoS biology*, 12(9), p. e1001955. doi: 10.1371/journal.pbio.1001955.

Sabass, B. *et al.* (2008) ‘High resolution traction force microscopy based on experimental and computational advances.’, *Biophysical journal*, 94(1), pp. 207–220. doi: 10.1529/biophysj.107.113670.

Salomon, J. *et al.* (2017) ‘Contractile forces at tricellular contacts modulate epithelial organization and monolayer integrity’, *Nature Communications*, 8. doi: 10.1038/ncomms13998.

Sankpal, N. V. *et al.* (2011) ‘Activator protein 1 (AP-1) contributes to EpCAM-dependent breast cancer invasion’, *Breast Cancer Research*, 13(6), pp. 1–13. doi: 10.1186/bcr3070.

Sankpal, N. V. *et al.* (2017) ‘A double-negative feedback loop between EpCAM and ERK contributes to the regulation of epithelial-mesenchymal transition in cancer’, *Oncogene*, 36(26), pp. 3706–3717. doi: 10.1038/onc.2016.504.

Sivagnanam, M. *et al.* (2008) ‘Identification of EpCAM as the Gene for Congenital Tufting Enteropathy’, *Gastroenterology*, 135(2), pp. 429–437. doi: 10.1053/j.gastro.2008.05.036.

Slanchev, K. *et al.* (2009) ‘The epithelial cell adhesion molecule EpCAM is required for epithelial morphogenesis and integrity during zebrafish epiboly and skin development’, *PLoS Genetics*, 5(7). doi: 10.1371/journal.pgen.1000563.

Soule, H. D. *et al.* (1973) ‘A human cell line from a pleural effusion derived from a breast carcinoma.’, *Journal of the National Cancer Institute*, 51(5), pp. 1409–1416. doi: 10.1093/jnci/51.5.1409.

Tobias Nubel, Preobraschenski, J. *et al.* (2009) ‘Claudin-7 Regulates EpCAM-Mediated Functions in Tumor Progression’, 7(March), pp. 285–300. doi: 10.1158/1541-7786.MCR-08-0200.

Tse, J. R. and Engler, A. J. (2010) ‘Preparation of hydrogel substrates with tunable mechanical properties’, *Current Protocols in Cell Biology*, (SUPPL. 47), pp. 1–16. doi: 10.1002/0471143030.cb1016s47.

Tseng, Q. *et al.* (2012) ‘Spatial organization of the extracellular matrix regulates cell-cell junction positioning’, *Proceedings of the National Academy of Sciences of the United States of America*, 109(5), pp. 1506–1511. doi: 10.1073/pnas.1106377109.

Tsubouchi, A. *et al.* (2002) 'phosphorylated paxillin in cell adhesion and migration', 159(4), pp. 673–683. doi: 10.1083/jcb.200202117.

Vannier, C. *et al.* (2013) 'Zeb1 regulates E-cadherin and Epcam (epithelial cell adhesion molecule) expression to control cell behavior in early zebrafish development', *Journal of Biological Chemistry*, 288(26), pp. 18643–18659. doi: 10.1074/jbc.M113.467787.

Vignaud, T., Ennomani, H. and Théry, M. (2014) 'Polyacrylamide Hydrogel Micropatterning', *Methods in Cell Biology*, 120, pp. 93–116. doi: 10.1016/B978-0-12-417136-7.00006-9.

Wang, J. *et al.* (2011) 'Loss of Trop2 promotes carcinogenesis and features of epithelial to mesenchymal transition in squamous cell carcinoma', *Molecular Cancer Research*, 9(12), pp. 1686–1695. doi: 10.1158/1541-7786.MCR-11-0241.

Wang, M. H. *et al.* (2018) 'Epithelial cell adhesion molecule overexpression regulates epithelial-mesenchymal transition, stemness and metastasis of nasopharyngeal carcinoma cells via the PTEN/AKT/mTOR pathway', *Cell Death and Disease*, 9(1). doi: 10.1038/s41419-017-0013-8.

Went, P. T. *et al.* (2004) 'Frequent EpCam Protein Expression in Human Carcinomas', *Human Pathology*, 35(1), pp. 122–128. doi: 10.1016/j.humpath.2003.08.026.

Winklbauer, R. (2015) 'Cell adhesion strength from cortical tension – an integration of concepts', 2, pp. 3687–3693. doi: 10.1242/jcs.174623.

Wu, C. J. *et al.* (2020) 'Matriptase Cleaves EpCAM and TROP2 in Keratinocytes, Destabilizing Both Proteins and Associated Claudins', *Cells*, 9(4). doi: 10.3390/cells9041027.

Figure legends

Figure 1. EpCAM KD stimulates cohesive collective migration of MCF7 spheroids. Spheroids of transfected MCF7 cells with control, EpCAM, Trop2, and EpCAM/Trop2 siRNA were plated on a layer of fibrillar collagen gel, let adhere for 30 minutes, then phase-contrast images were captured every 30min for 24hrs. **A-D.** Images of whole spheroids at selected time points. Scale bar = 100 μ m. **E.** Quantification of relative area increase after 24hrs. The box plots show the interquartile range (box limits), median (centerline), and min and max values without outliers (whiskers) for calculated area ratios of four independent experiments. Statistical analysis: one-way ANOVA followed by Tukey-HSD post hoc test. For all experiments presented in this study, P values are indicated as follows: * P < 0.05, ** P < 0.01, *** P < 0.001 and NS, not significant. **F.** Quantification of spheroids solidity, measured as the ratio ([Area])/([Convex area]). Quantification from the same four experiments as in E. **H, H'.** Top and orthogonal view of phalloidin labeled spheroids after 24hrs spreading. The orthogonal view allows us to visualize clear differences between the multicellular compact and tall center of the siCtrl spheroids compared to the much flatter siEpCAM and dKD spheroids. siTrop2 spheroids are the tallest. **I.** Schematic representation of typical spheroid morphologies in the top and orthogonal views. **J.** Details of typical spheroid edges. Maximal projection of 3 z planes, 1 μ m apart. Similar protrusions are observed in all conditions (arrowheads). However, siEpCAM and dKD show characteristic actin cables along the spheroid edge (arrows), while control and siTrop2 spheroids have numerous single cells and small cell groups protruding out of the main cell mass (asterisks).

Figure 2. Single-cell migration is impaired by EpCAM depletion. MCF7 cells, transfected with control, EpCAM, Trop2, and EpCAM/Trop2 siRNA, were dissociated and plated on collagen gel. 6hrs long time-lapse movies were started ~ 3hrs after plating. **A-D.** Representative cells at four

time points. Scale bar = 15 μ m. **E,F.** Quantification of velocity and distance to the origin, measured using the MetaMorph software tracking tool. Three independent experiments. Statistical analysis is one-way *ANOVA* followed by Tukey-HSD post hoc test. **G.** Quantification of cell area, obtained by z projection on stacks of confocal images of cells labeled with phalloidin (not shown). Three independent experiments.

Figure 3. EpCAM depletion prevents the detachment of single cells. A-D. Selected time frames of time-lapse movies showing edges of spheroids. Control MCF7 spheroids occasionally show single cells breaking off and migrating away (A, arrow). This almost never happens, at least within 24hrs of imaging, for siEpCAM spheroids (B), while it does occur in siTrop2 (C). **E.** Proportion of spheroids showing at least one cell or one cell group migrating away during 24hrs. Data from 84, 85, 51, and 50 spheroids (corresponding to siCtrl, siEpCAM, siTrop2, and dKD respectively) from 13 independent experiments. Statistical analysis: Kruskal-Wallis Test. **F.** Quantification of relative area increase in spheroids of EpCAM-GFP MCF7 cells. Spheroids of EpCAM-GFP stable cells were formed and migration assay was done as explained in figure 1. Conditions included non-treated, doxycycline (to induce EpCAM-GFP expression) 1 and 3 μ M concentrations treated spheroids. Data from 16, 19, and 19 spheroids of non-treated, doxycycline 1 and 3 μ M concentrations treated spheroids, respectively. Three independent experiments. Statistical analysis: one-way *ANOVA* followed by Tukey-HSD post hoc test.

Figure 4. E-cadherin and pMLC levels are elevated upon both EpCAM and Trop2 KD. A,B. Representative spinning disk confocal images of spheroids after 24hrs spreading, immunolabelled for p-MLC and E-cadherin. Levels are visualized using the “fire” pseudocolors of ImageJ. The selected z planes correspond to the widest area for each spheroid. Scale bar = 50 μ m **C,D.** Quantification of p-MLC and E-cadherin mean intensities along cell outlines, delimited using a

level threshold in ImageJ. Mean intensities were averaged for each spheroid. 8 spheroids of each condition from three independent experiments. Statistical analysis was done using one-way *ANOVA* followed by Tukey-HSD post hoc test.

Figure 5. EpCAM and Trop2 depletions differentially affect the myosin balance between cell contacts and the free cortex. Confocal images of MCF7 cells on collagen gel, co-immunolabelled for E-cadherin (red) and pMLC (green). F-actin filaments were visualized with phalloidin-Alexa647 (magenta). **A,B.** Representative images, respectively of cell pairs and small groups. Scale bars = 10 μm **C.** Quantification of peak intensities for p-MLC along free cell edges (left) and cell-cell contacts (middle), and ratio contact to the cortex (right). **D.** Quantification of E-cadherin peak intensity at cell-cell contacts (left), and corresponding ratios p-MLC to E-cadherin (middle) and phalloidin to E-cadherin (right). **E.** Quantification of F-actin, similar to pMLC (C). Statistical analysis: one-way *ANOVA* followed by Tukey-HSD post hoc test.

Figure 6. Distinct impact of EpCAM KD and Trop2 KD on focal adhesions. **A.** Confocal images of single cells on collagen gel, co-immunolabelled for phospho-Tyr118-paxillin (p-paxillin, red) and vinculin (green). F-actin filaments were visualized with phalloidin-Alexa647 (magenta). Scale bars = 5 μm . **B.** Total levels of p-paxillin (left) and vinculin (middle) accumulated in focal adhesions, per cell, normalized to siCtrl. Right: vinculin to p-paxillin ratio. **C.** p-paxillin and vinculin levels in focal adhesions, normalized to the cell surface. Five independent experiments, statistical analysis using one-way *ANOVA* followed by Tukey-HSD post hoc test.

Figure 7. Traction and cell-cell forces are increased upon EpCAM and Trop2 depletions. Traction Force Microscopy (TFM) of cell doublets on H patterns. Cells doublets were laid on H patterns coated with a thin layer of collagen on a polyacrylamide gel with stiffness of 5kPa, containing far-red fluorescence nanobeads. Cells and underneath nanobeads were imaged, then

cells were removed by trypsinization, and second images of the same positions were taken and used as a reference, from which bead displacement was measured. **A.** Diagram of force vectors obtained by TFM on H-shape micropatterns. While traction forces (black arrows) are measured from the displacement of nanobeads, the cell-cell force (red arrow) is calculated indirectly based on its counterbalancing the sum of traction forces, as shown in the equation. **B.** Representative images of micropattern-confined cell doublets for the four experimental conditions. **C.** Corresponding average maps of traction forces. **D.** Quantification of traction forces. **E.** Quantification of cell-cell forces. **F.** Cell-cell force to traction forces ratio. **G.** Quantification of junction lengths, measured from the phase-contrast images. **H.** Spheroid migration on a collagen-coated 5kPa polyacrylamide gel. Images of siCtrl and siEpCAM spheroids at four time points as in Figure 1. Scale bar = 150 μ m. Arrows point to detachments of single and small cell groups from the main mass of the control spheroid.

Figure 8. EpCAM KD increased both cortical tension and cell-cell adhesiveness in suspended cell doublets. A, A'. Diagram of a symmetric and an asymmetric cell doublet respectively, with the balance between cortical tensions at the free edges (C_{tA} and C_{tB}) and contact tension (T_{AB}). The direction of C_{tA} , C_{tB} , and T_{AB} is tangential to the membranes at the cell vertex, which allows direct calculation of the relative strengths of these tensions based on the geometry at vertices. The orange layer represents the actomyosin cortex, with its thickness symbolizing relative contractility. The curved cell-cell contact reflects unequal C_{tA} and C_{tB} tensions. **B,C.** Examples of homotypic and heterotypic doublets imaged by live confocal microscopy. The doublets were formed by mixing dissociated cells and letting them re-associate on an adhesion-free support. Membranes were labeled with CellMask Alexa Fluor 647. siCtrl cells were marked by Hoechst staining

(showed in red) prior to dissociation. Scale bar = 10 μ m. **D.** Quantification of relative cortical tensions (see text for definition). **E.** Quantification of adhesiveness (see text for definition).

Figure 9. Differential EpCAM and Trop2 localization. **A.** Confocal microscopy images of groups of wild-type MCF7 cells on collagen gel, co-immunolabelled for EpCAM (green, right), Trop2 (red, middle), and F-actin (labeled using phalloidin-Alexa647, magenta, left). EpCAM tends to be rather uniformly distributed along all membranes, while Trop2 is more punctate, and more abundant at free cell edges (arrowheads), while relatively low at cell-cell contacts (arrows). Concave arrowheads in **B** point at bright intracellular spots. **B-C.** Quantification of EpCAM or Trop2 levels, obtained from the analysis of whole z-stacks. Four membrane regions were distinguished, free cell edge, cell-cell contacts, top and bottom of the cells, from which the average intensity was measured. Values were normalized to the median intensity for cell edges. **D.** Trop2/EpCAM ratio, calculated for each compartment. Statistics: One-way ANOVA followed by Tukey-HSD post hoc test.

Figure 10. EpCAM-negative cells sort out from EpCAM-positive cells in mixed spheroids. Spheroids were formed from MCF7 cells transfected with a low amount of EpCAM siRNA, titrated to obtain a mixed population of EpCAM positive and negative cells. After 24hrs on collagen gel, spheroids were fixed, immunolabelled for EpCAM, and nuclei labeled with Hoechst. The entire spheroid was imaged by confocal microscopy. **A.** Representative example, showing a single z plane close to the bottom (widest spheroid area). EpCAM positive cells (red) tend to cluster, mostly in the center of the spheroid. Arrows: Examples of a cell expressing a high level of EpCAM. Spheroid edges (arrowheads) are largely devoid of EpCAM signals. **B.** Diagram showing the position of the highly EpCAM-positive cells as large red dots. The other nuclei are represented as small grey dots. The diagram is the superposition of a merged projection of the nuclei from the

bottom slices (20 μ m) of 3 spheroids, all centered around their centroid. The nuclei of the 3 spheroids are distinguished by their tint, light to dark red for EpCAM positive cells, grey scale for the others. Scale bar = 50 μ m. **C.** Histogram showing the distance from the centroid for total nuclei and nuclei from EpCAM-positive cells. Five spheroids from two independent experiments. Statistical analysis: Student's t-test. **D.** Example of EpCAM-positive cells at the spheroid edge. Scale bar = 50 μ m.

Supplementary figure legends

Figure S1. Efficiency of EpCAM and Trop2 depletions. Representative confocal microscopy images of MCF7 cells transfected for 96hrs with Ctrl, EpCAM, and Trop2 siRNA, immunolabelled for EpCAM and Trop2. Nuclei were stained with Hoechst (blue). The specific signal along cell membranes is undetectable in the respective siRNA condition.

Figure S2. Increased spheroid migration upon EpCAM KD is independent of cell proliferation. Spheroids were treated with 2.5 μ M mitomycin C (MMC) during the entire migration assay. **A.** Validation of mitomycin MMC efficiency by imaging EdU incorporation. At the end of the migration assay, the spheroids were incubated for 1hr with thymidine analog EdU, which efficiently incorporates into newly synthesized DNA. EdU was detected in green (see Material and Methods), while EpCAM was detected by immunofluorescence (red), and nuclei were stained with Hoechst (blue). The four panels show representative confocal microscopy images of non-treated and MMC treated spheroids of siCtrl and siEpCAM conditions. Scale bar = 15 μ m. **B.** Quantification of spheroid spreading. Three independent experiments. Statistical analysis: one-way ANOVA followed by Tukey-HSD post hoc test.

Figure S3. Increased spheroid migration upon EpCAM KD can be rescued by myosin or nPKC inhibition. Collective migration for both control spheroids and EpCAM KD was tested in the absence or presence of 5 μ M of the myosin inhibitor blebbistatin (BB) or 0.5 μ M of the nPKC inhibitor calphostin C (CC). Increased spheroid areas were normalized to untreated siCtrl. Four independent experiments. Statistical analysis: one-way *ANOVA* followed by Tukey-HSD post hoc test.

Figure S4. EpCAM and Trop2 depletions affect the geometry and polarity of cell-matrix adhesions. Single cells immunolabelled for p-paxillin and vinculin were categorized based on their general spreading morphology (round, bipolar, spread) as well as the geometry of the distribution of the focal adhesions (symmetric ~ polygonal or radial, asymmetric, or unipolar). **A-E.** Examples of cells with different geometries. **A-B.** Characteristic symmetric configurations of siEpCAM cells. **A.** Arrows show four opposed corners with large focal adhesions. **B.** Example of cell displaying a radial arrangement with multiple weak focal adhesions. **C.** In the most frequent configuration of control/wild type condition, cells are spread, but one side of the cells appears to have a deficit in focal adhesions (arrowhead), suggesting an imbalance with the other corners (arrows). **D.** Round cell, typically observed for siCtrl and siTrop2 conditions. The filled and concave arrowheads point to the two sides of the contact, respectively high and low in vinculin. **E.** Compact siTrop2 cell forming one major protrusion (arrow). **F.** Quantification. Data were compiled from two experiments.

Figure 1

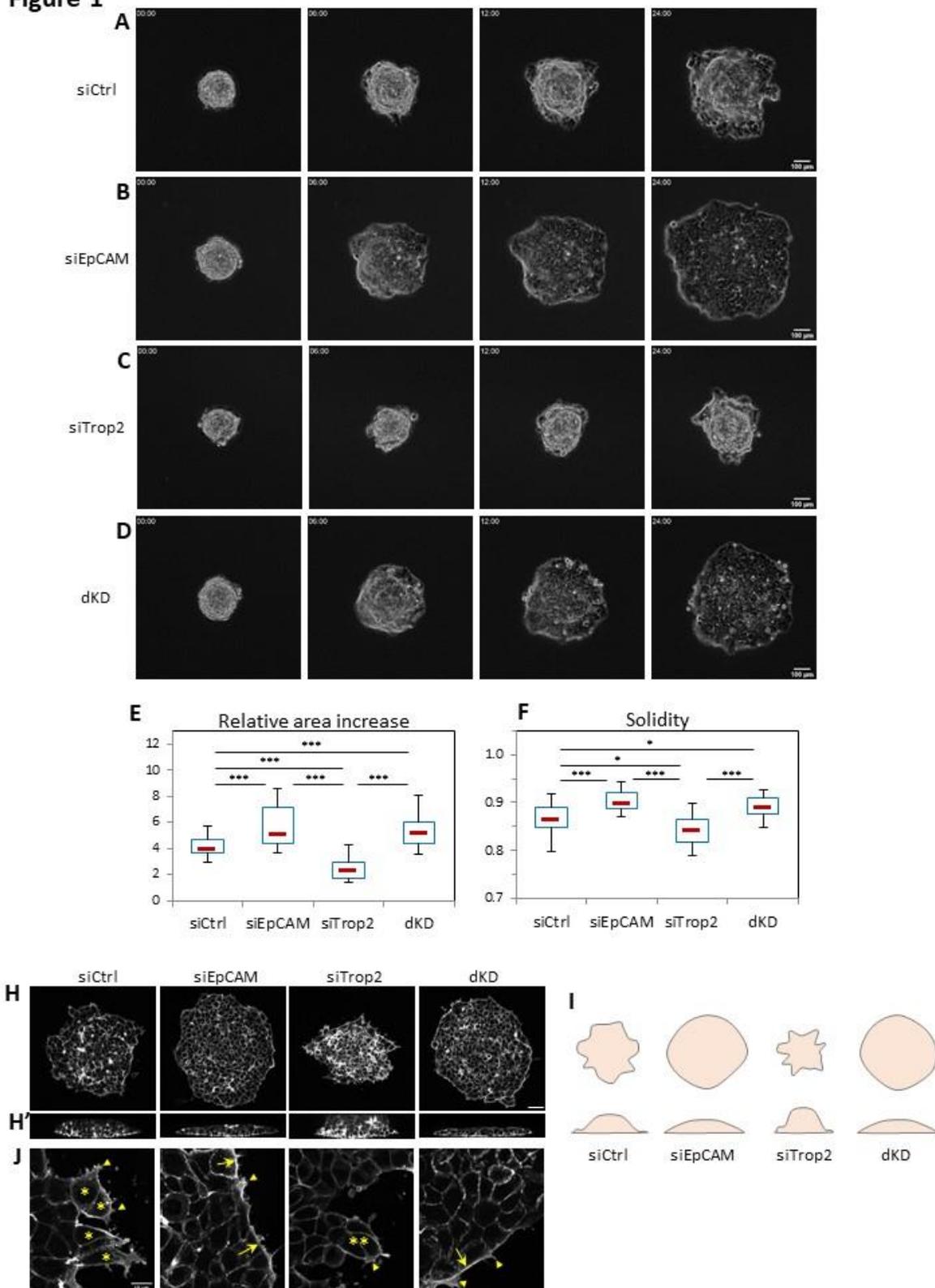


Figure 2

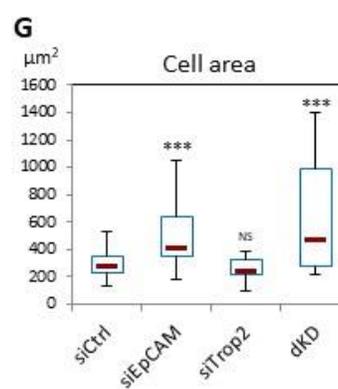
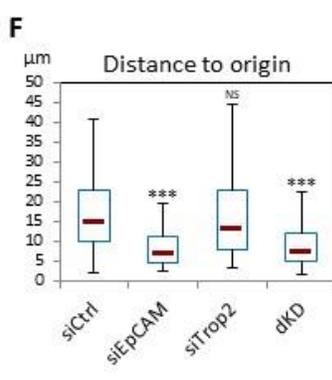
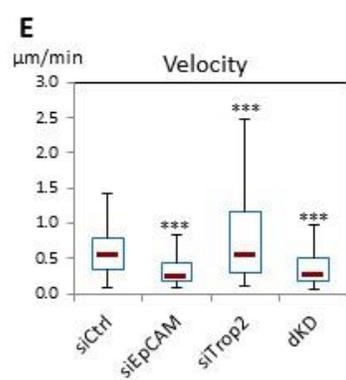
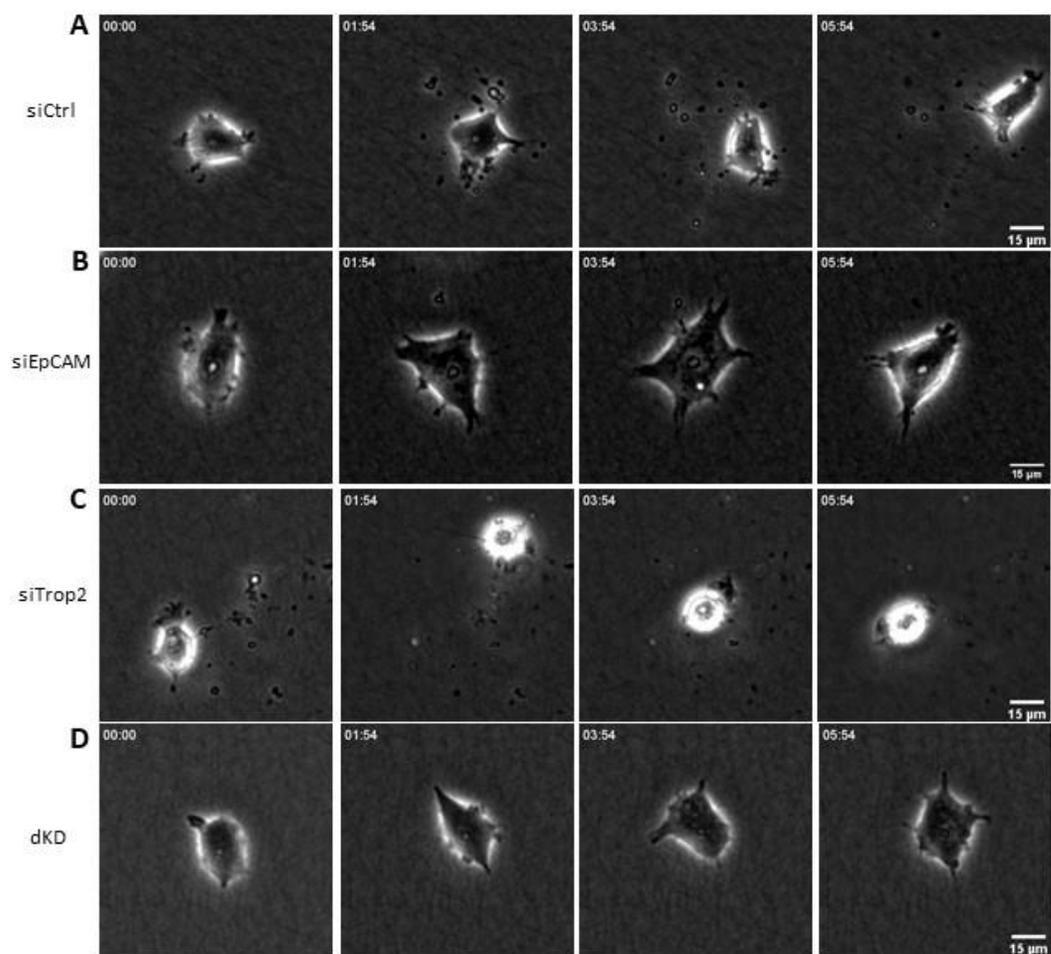


Figure 3

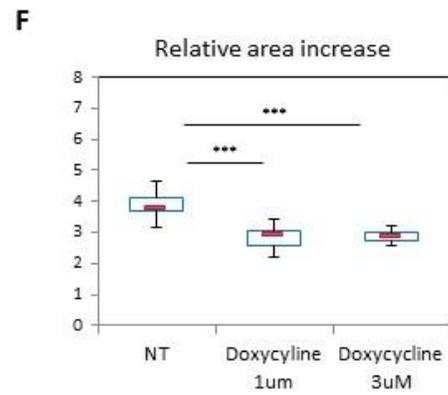
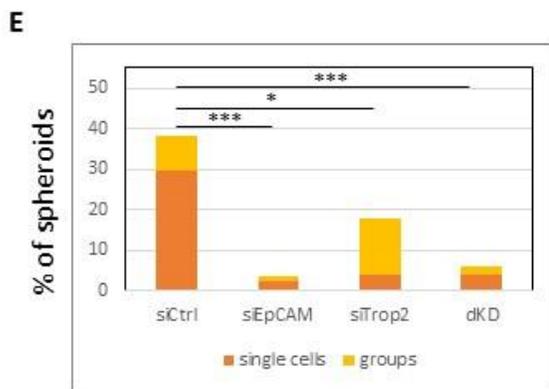
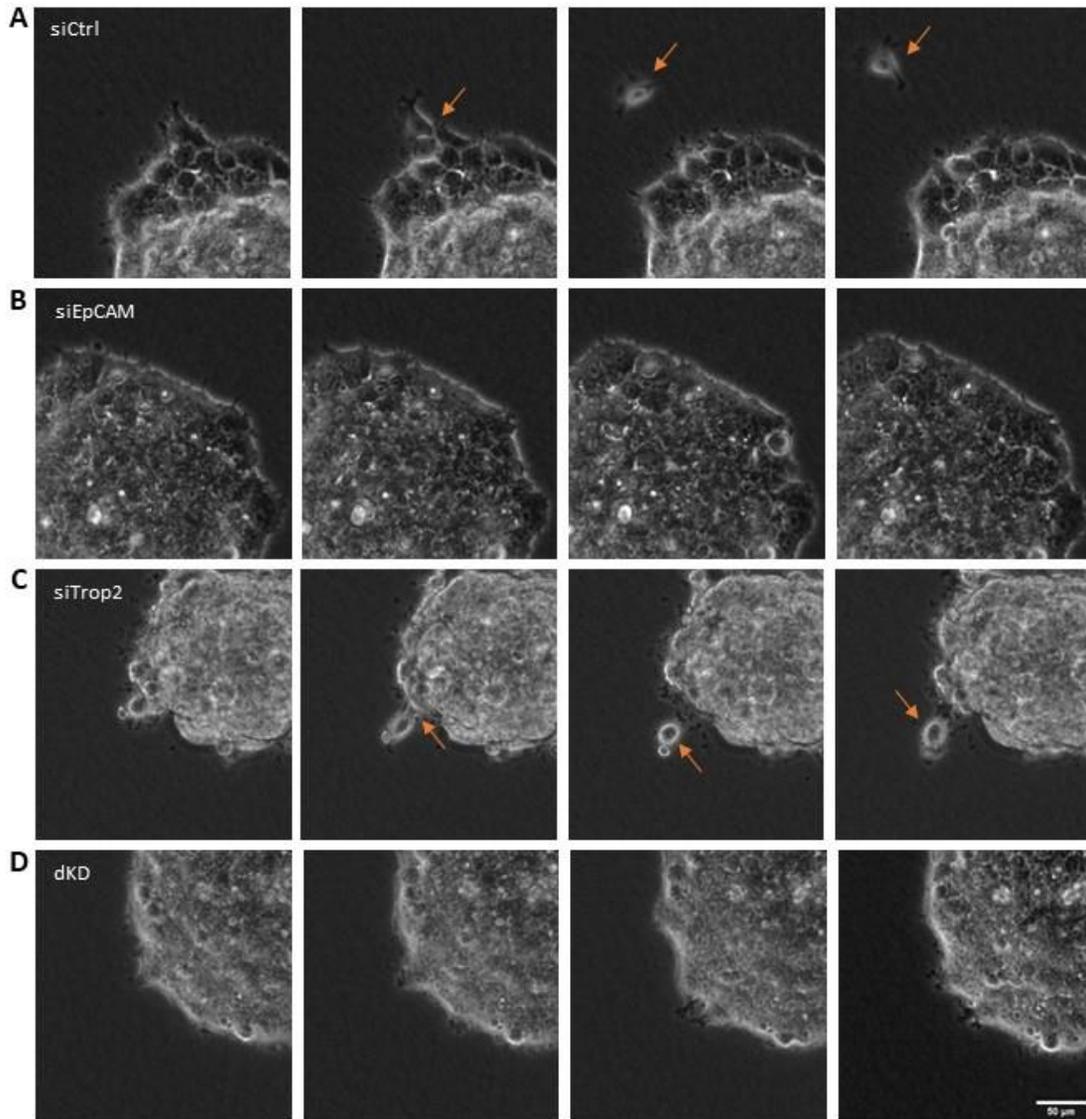


Figure 4

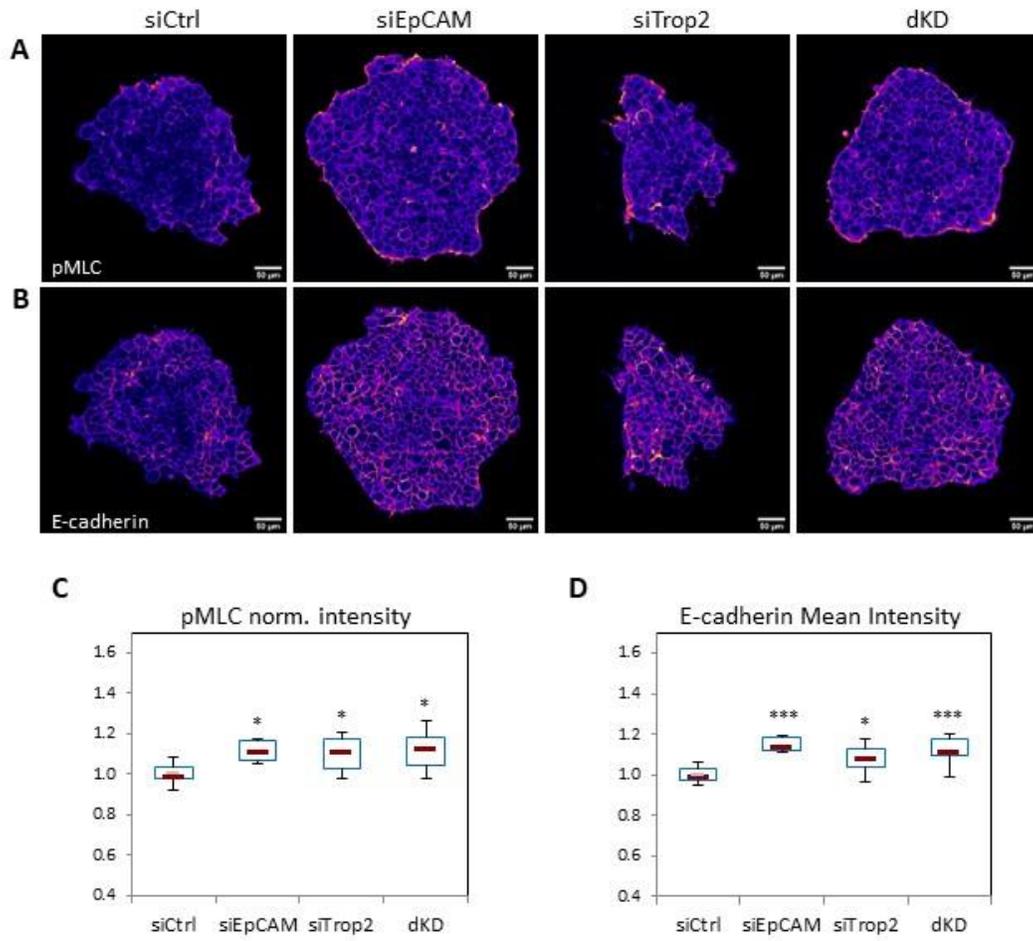


Figure 5

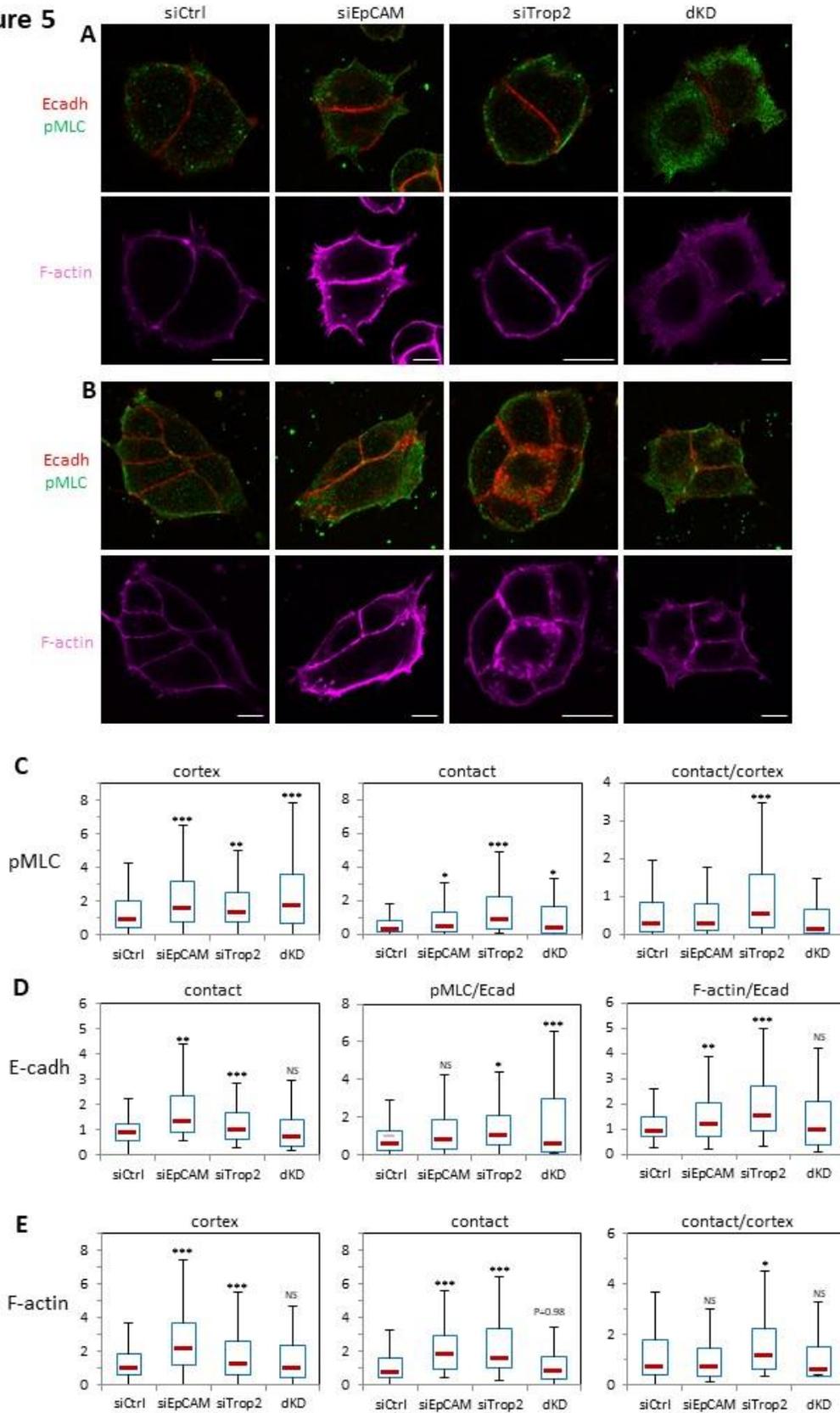


Figure 6

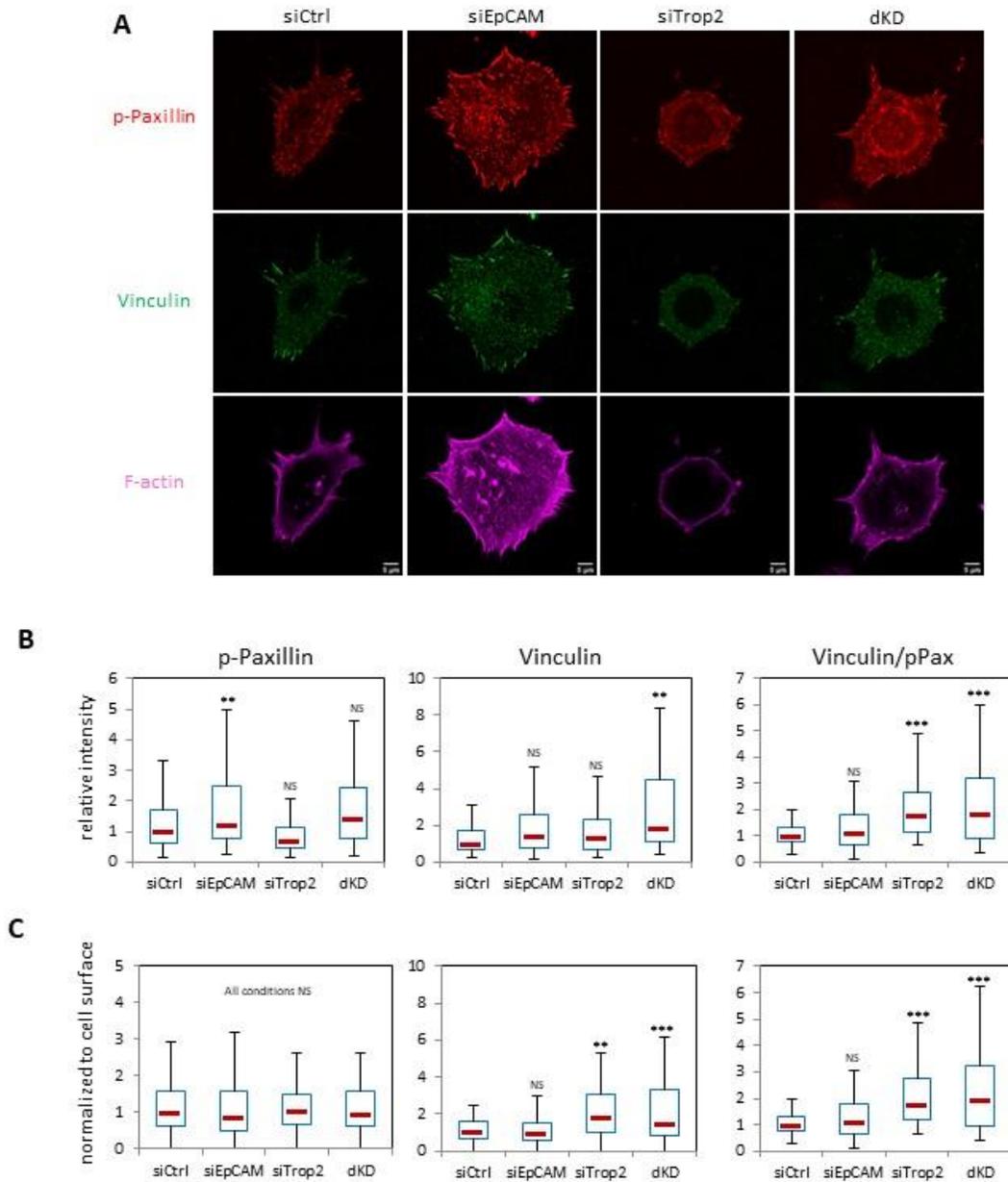


Figure 7

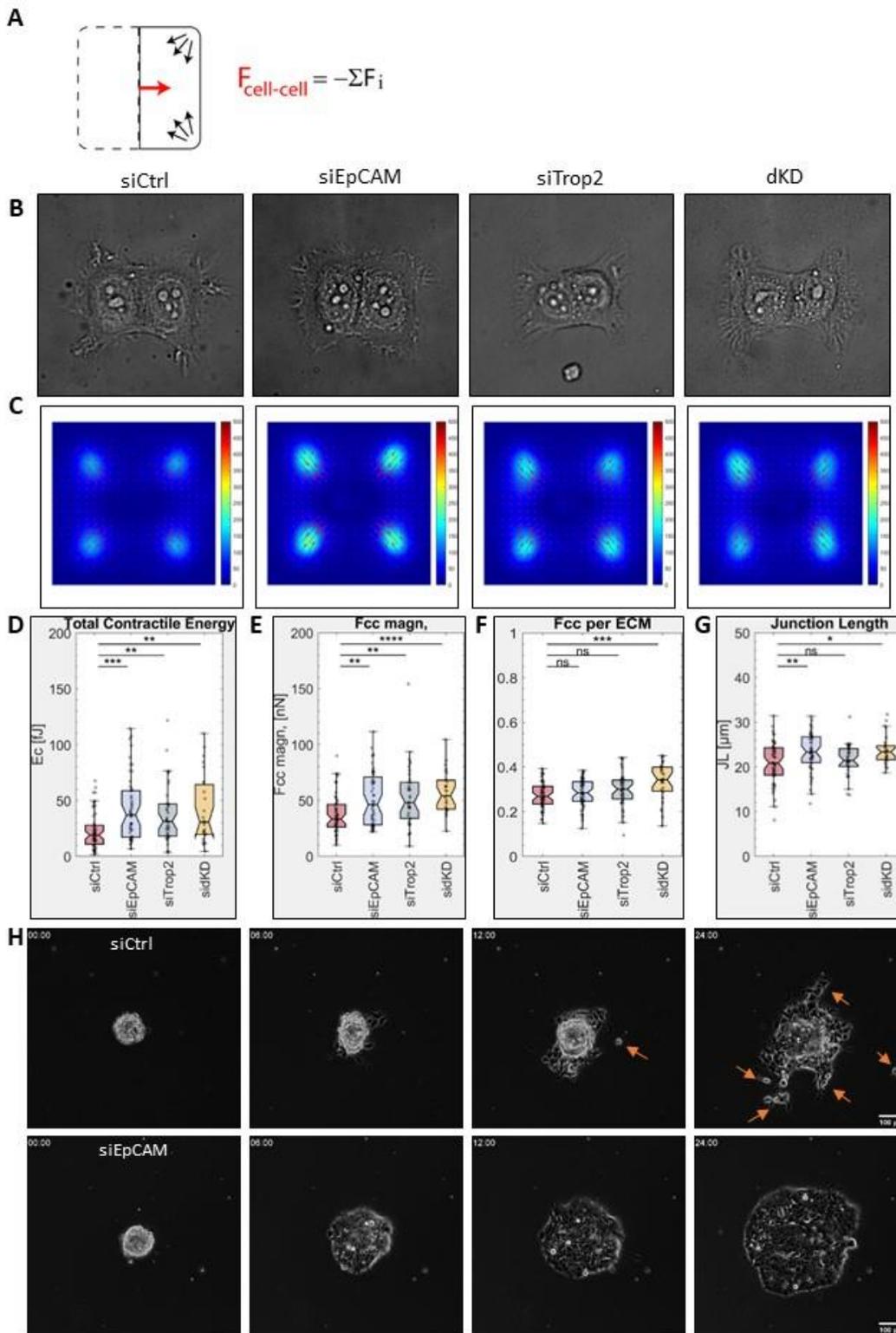


Figure 8

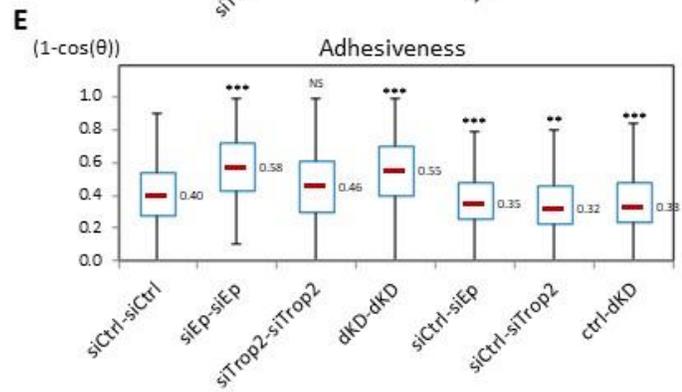
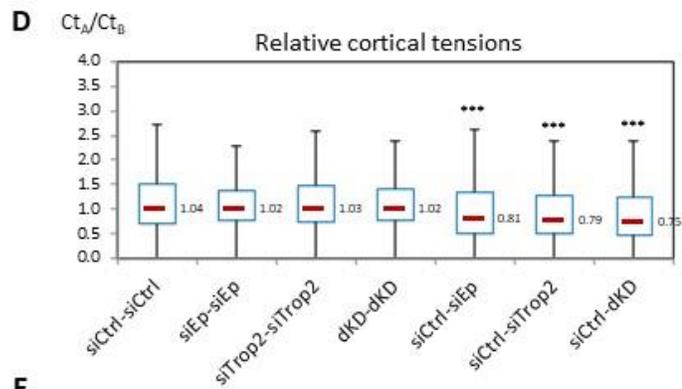
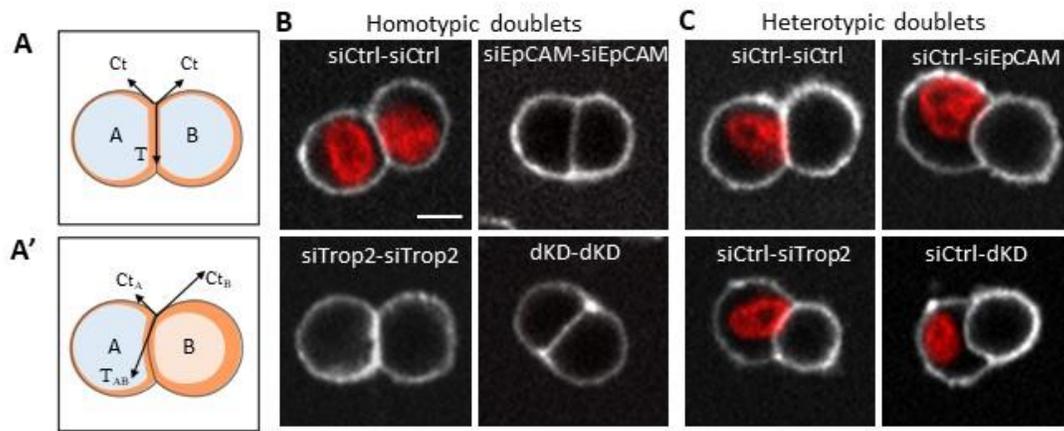


Figure 9

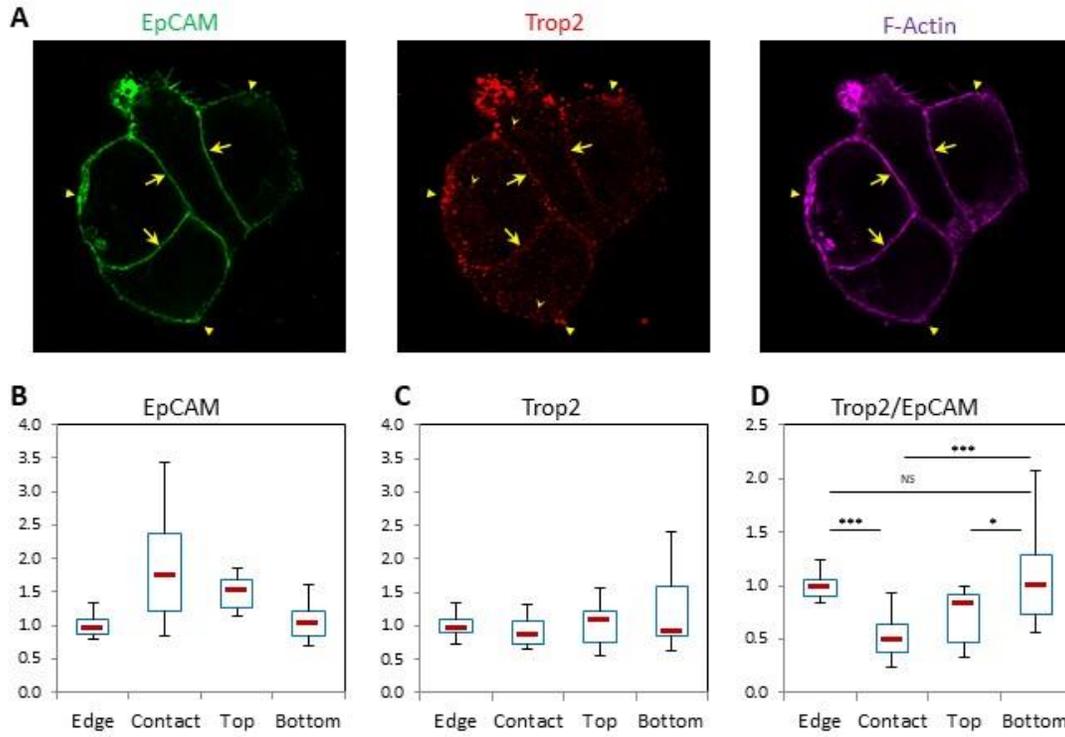


Figure 10

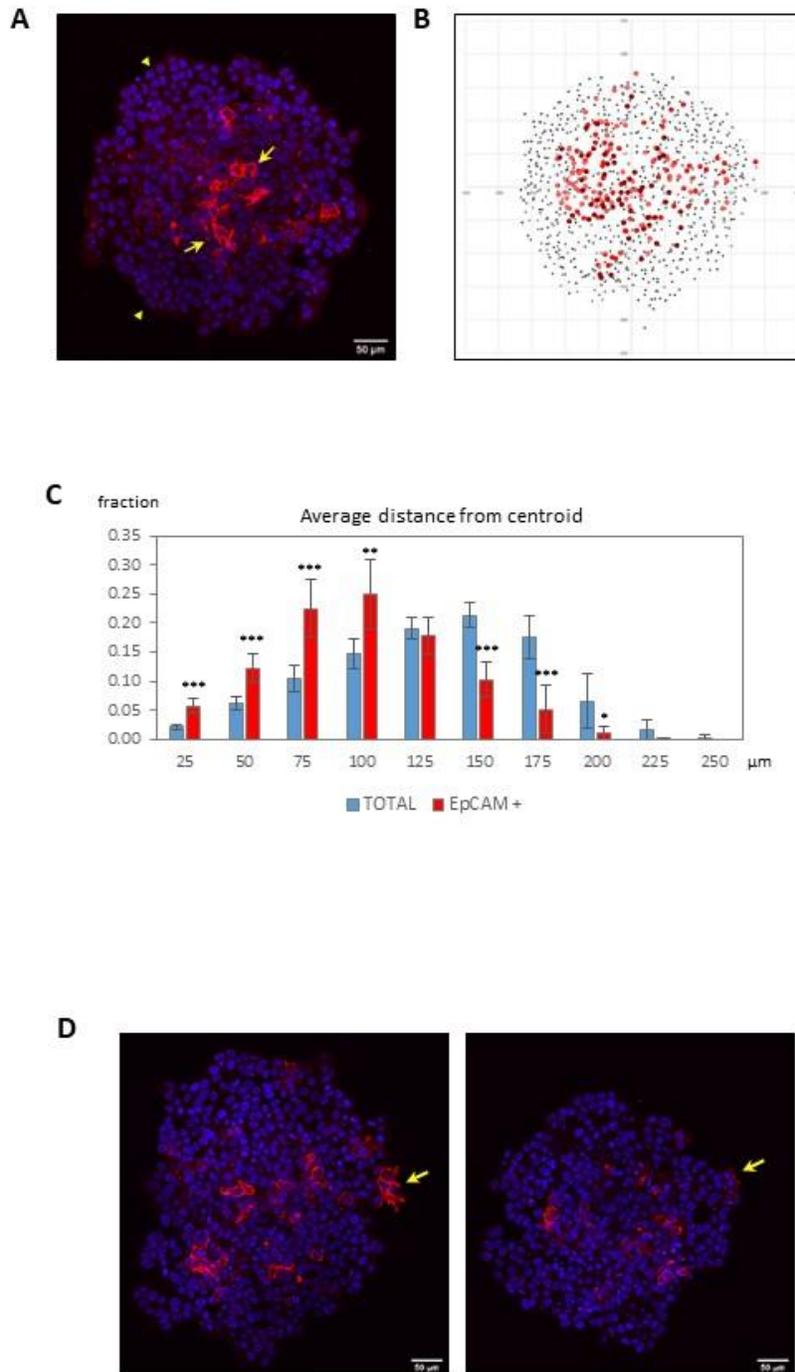


Figure S1

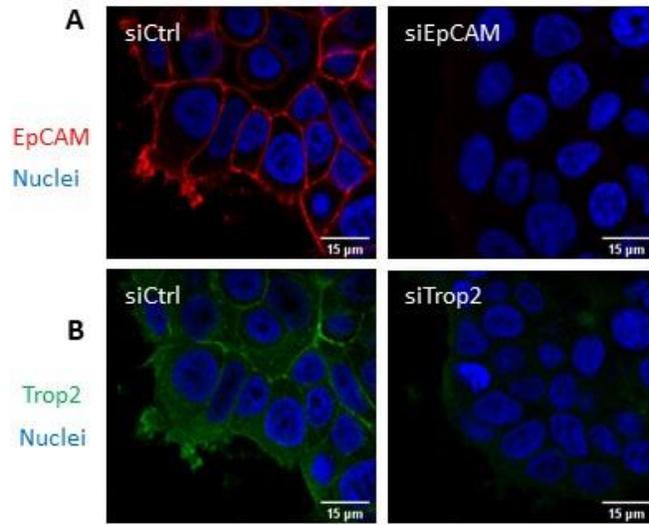


Figure S2

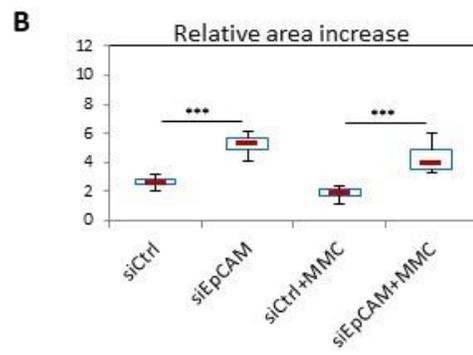
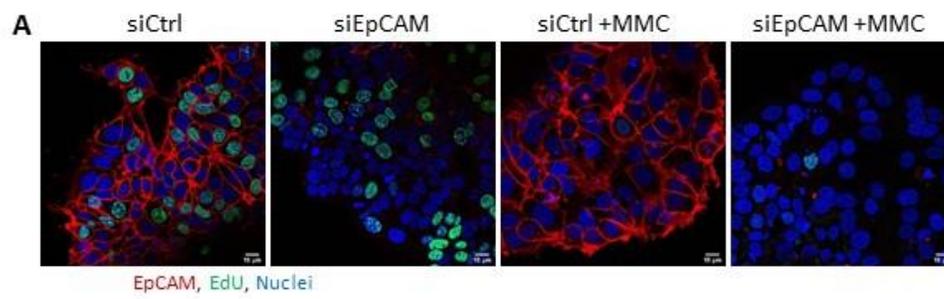


Figure S3

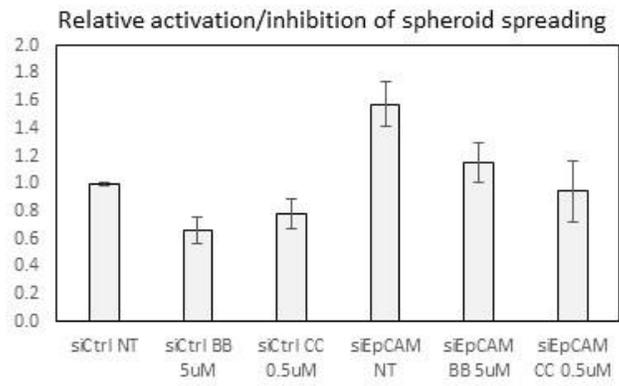
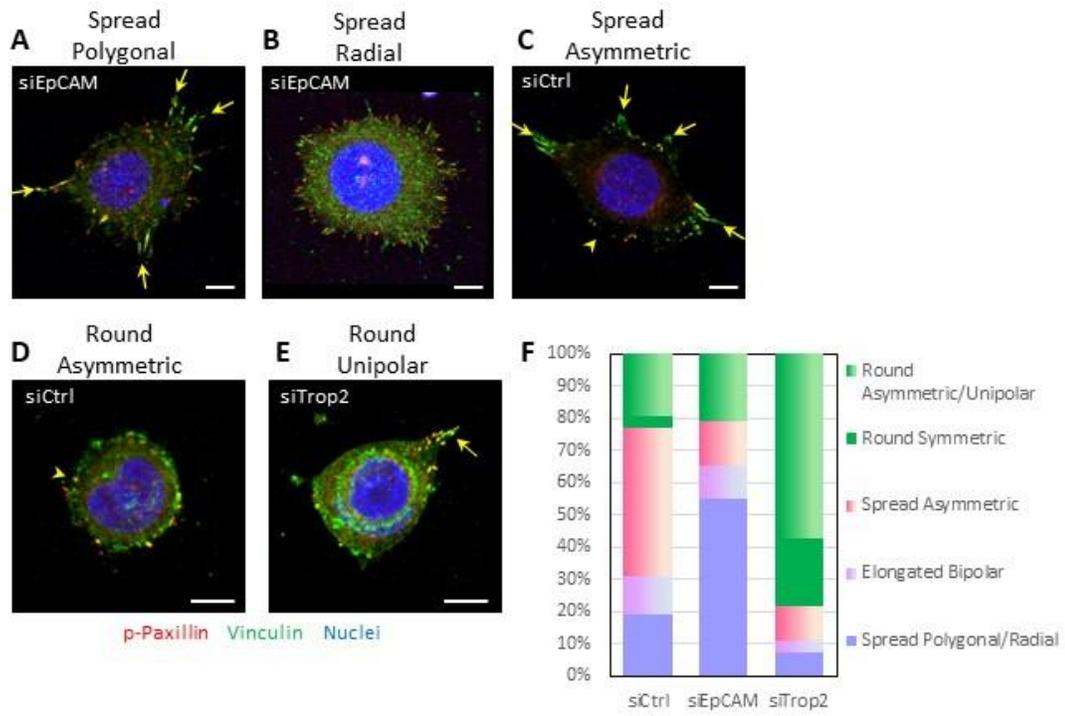


Figure S4



Chapter III

Limited significance of the in situ proximity ligation assay

Azam Alsemarz^{1,2}, Paul Lasko² and François Fagotto^{1,2}

¹CRBM, University of Montpellier and CNRS, Montpellier 34293, France

and ²Dept. of Biology, McGill University, Montreal, QC, Canada H3A1B1.

Correspondence to François Fagotto, email: francois.fagotto@crbm.cnrs.fr

Abstract

In situ proximity ligation assay (isPLA) is an increasingly popular technique that aims at detecting the close proximity of two molecules in fixed samples using two primary antibodies. The maximal distance between the antibodies required for producing a signal is 40 nm, which is lower than optical resolution and approaches the macromolecular scale. Therefore, isPLA may provide refined positional information, and is commonly used as supporting evidence for direct or indirect protein-protein interaction. However, we show here that this method is inherently prone to false interpretations, yielding positive and seemingly ‘specific’ signals even for totally unrelated antigens. We discuss the difficulty to produce adequate specificity controls. We conclude that isPLA data should be considered with extreme caution.

Introduction

A major challenge of cell biology is the ability to detect protein-protein interactions in situ. The current methods of choice are FRET and related techniques. However, this type of approach involves the expression of tagged fusion proteins, has limited sensitivity, and often requires extensive optimization. The in situ proximity ligation assay (isPLA) has thus appeared as a very attractive, easy and ready to use alternative (Fredriksson *et al.*, 2002; Gullberg *et al.*, 2004; Söderberg *et al.*, 2006, 2008; Bellucci *et al.*, 2014; Blokzijl *et al.*, 2014; Bagchi, Fredriksson and Wallén-Mackenzie, 2015; Greenwood *et al.*, 2015; Lutz *et al.*, 2017; Jalili *et al.*, 2018). Its principle is based on the immunodetection of two antigens with a pair of primary antibodies raised in different species (Fig.1A). The two primary antibodies are then recognized by two species-specific secondary antibodies, called PLA probes, each linked to a unique short DNA strand. When the two PLA probes are in close proximity, the DNA strands can be used to recruit two additional

connector oligonucleotides, which are ligated to form a DNA circle. This allows the synthesis of a single-stranded rolling circle PCR product, composed of hundreds of concatenated complements of the DNA circle, which is then visualized using a fluorescently labeled complementary oligonucleotide probe. The maximal distance allowing this reaction is 40 nm, which is not quite small enough to demonstrate a physical interaction between the two antigens, but sufficient to support a very close ‘proximity’. It is certainly below the limit of optical resolution, thus potentially much more informative than classical fluorescence colocalization experiments. The potential applications of this method are huge since it can be used in principle with any pair of antibodies, allowing co-detection of any endogenous antigen, including posttranslational modifications, such as specific phosphorylated sites. This flexibility explains its increasing popularity, in fields as diverse as cell biology, pharmacology, immunology, virology, proteomics, biomarkers for cancer, pathogen diagnostic, or even astrobiology (Weibrecht *et al.*, 2010; Blokzijl *et al.*, 2014; Bunse *et al.*, 2015; Greenwood *et al.*, 2015; Lipovsky *et al.*, 2015; Gomes, Sierra and Devi, 2016; Mereiter *et al.*, 2016; Debaize *et al.*, 2017; Lutz *et al.*, 2017; Jalili *et al.*, 2018; Johnson *et al.*, 2018). Unfortunately, when attempting to apply this approach, we were surprised to obtain robust positive results for pairs of antigens that, although partially yielding overlapping immunofluorescence signals, could not possibly establish any direct or indirect interaction. A closer consideration of the principle of isPLA suggested the possibility that indeed a positive signal may be generated for any pair of antigens, provided that a subset of the primary antibodies would happen to bind sufficiently close to each other. Yet, the goal of a proximity assay should be to reveal the presence of two antigens within structures in the tens of nanometer range (a macromolecular complex, a vesicle, or a membrane subdomain), and to discriminate between these

relevant cases and a mere random proximity, for instance, the close but fortuitous encounter of a plasma membrane protein and an unrelated soluble cytosolic protein (Fig.1B).

We, therefore, evaluated the capacity of isPLA to differentiate between an actual interaction *versus* random proximity (Fig.2B,B'). We unambiguously conclude that isPLA yields positive signals in a variety of conditions that are irrelevant both in terms of protein-protein interactions or refined subcellular localization. In fact, attempts to set adequate specificity controls as performed in our study, and which would be necessary to support the existence of a biologically relevant proximity, were, to our knowledge, not implemented in other studies. As currently performed, isPLA is most likely to yield a large number of false positives and to lead to unsupported conclusions.

Results

In the series of tests presented here, we set to evaluate the occurrence of PLA positive signals for different situations, comparing pairs of antibodies against known interacting proteins and pairs recognizing unrelated antigens. The PLA results were analyzed both in terms of distribution of the fluorescent spots and spot density. An additional classical double immunofluorescence labeling was performed on the same samples (Fig.2A), which allowed to directly verify the distribution of the two primary antibodies and compare relative intensities between conditions. The standard negative control for isPLA is the omission of one of the primary antibodies. In our hands, such control always yielded little or no PLA reaction, but as discussed below, this cannot be considered as an acceptable control.

In this survey, we tested PLA for well-established residents of three cellular structures, the plasma membrane (E-cadherin), cytoskeleton (α -tubulin), and nucleus (transcription factor Sox9). We used ectopically expressed soluble GFP as a randomly distributed protein. As a general marker

for the plasma membrane, we expressed mGFP, i.e. GFP fused to a sequence that becomes palmitoylated and thus efficiently anchored to the inner leaflet of the plasma membrane. GFP and mGFP were detected with a specific anti-GFP antibody. Note that using ectopic GFP enabled us to monitor isPLA for a wide range of expression levels and immunofluorescence intensities, which could be compared to signals obtained after labeling endogenous proteins.

We also included in our tests non-specific IgGs. These are normally used as a classical negative control in traditional immunofluorescence, but they are also well known to broadly label the cytoplasm and the nucleus, to an extent that depends on the concentration used. We thus used them here as an example of widespread non-specific antibody labeling. Upon adequate titration, they yielded an immunofluorescence signal in the range obtained with the other antibodies (Figs.3, S2, and S3). We stress here this important distinction between controls for specific immunolabelling or specific proximity. For the latter purpose, any type of negative control, which removes the binding of one of the primary antibodies (e.g. omitting the antibody, knocking out the antigen, competing with the corresponding peptide) will give trivially a low or blank signal, but will not give information about the specificity of the positive reaction. Instead, one needs to be able to evaluate side by side isPLA produced by a candidate pair of interactors, as depicted in Fig.1B, or produced non-specifically by random adjacent antibodies present in the region of interest at similar levels, as depicted in Fig.1B'. To our knowledge, this type of control has never been implemented for isPLA.

We compared isPLA between E-cadherin and its direct cytoplasmic interactor β -catenin, E-cadherin and mGFP, or E-cadherin and soluble GFP (Fig.2 and suppl. Fig.S1). The anti-E-cadherin antibody recognized an epitope on the extracellular domain. We found that the plasma membrane was decorated with PLA signal in all three cases (Fig.2, arrows). The density of PLA spots

appeared to be roughly proportional to the intensity of the β -catenin/GFP immunofluorescence signals, thus to the density of the primary antibodies. This relationship was quantified through line scans along the membrane (Fig.S1H-I), and the data were plotted as PLA density as a function of green fluorescence intensity signal, representing the relative levels of anti- β -catenin or anti-GFP antibodies (Fig.S1J): While the frequency of PLA spots tended to be on average slightly higher for the E-cadherin- β -catenin pair, and lowest for the E-cadherin-GFP pair, the three distributions largely overlapped (enlargement in Fig.S1J). We concluded that for similar levels of primary antibodies, PLA could not effectively discriminate between the specific E-cadherin- β -catenin interaction and the other conditions.

We then tested isPLA for tubulin and soluble GFP or mGFP, or non-specific IgGs (Fig.3 and Fig.S2). Robust PLA was obtained in all three cases. The spots decorated microtubules in the first and last conditions (Fig.3A,C), consistent with the widespread distribution of the anti-GFP and non-specific IgG signals. For the tubulin-mGFP pair, they concentrated along the plasma membrane, where the two antibodies mostly overlapped (Fig.3B). Again, the density of PLA spots was related to the global antibody levels (e.g. compare cells 1 and 2, in Figure 3A, which express different GFP levels), although in detail the PLA position did not necessarily correlate with sites of highest immunofluorescence signal. Standard negative controls, i.e. cells not expressing GFP (Fig.3A,B and Fig.S2A,B), or omission of the anti-tubulin antibody (Fig.S2D) gave little to no PLA signal.

Finally, we compared isPLA for antibodies raised against Sox9 and Sam68, a nuclear factor recently shown to physically interact with Sox9 (Girardot *et al.*, 2018) (Fig.4A), or for anti-Sox9 and a non-specific antibody (Fig.4B). An additional interest of the Sox9-Sam68 interaction was its proposed enrichment in the peripheral region of the nucleus, based on isPLA (Girardot *et al.*,

2018). We reproduced this distribution (Fig.4A,A', histogram in panel C), which was in stark contrast with the immunofluorescence staining indicating a relatively homogenous distribution of the primary antibodies (Fig.4A", quantification in Fig.S3F). Strikingly, however, a very similar pattern was observed when the anti-Sam68 antibody was replaced with non-specific IgGs (Fig.4B,B',D), indicating that it did not represent a specific sub-nuclear site of interaction (discussed below). Note that PLA spots were also abundant in the cytoplasm signal, which was surprising considering the rather low cytoplasmic immunofluorescence signals for Sox9 and Sam68 (see discussion). Negative controls omitting anti-Sox9 or anti-Sam68 antibodies were blank for PLA (suppl. Fig.3D,E).

Discussion

The series of tests presented here show that the proximity ligation assay can produce positive reactions under a variety of situations, including conditions that bear little to no significance in terms of actual specific proximity between macromolecules (Fig.5). This is a critical issue, which, retrospectively, appears inherent with the principle of the assay: Because the annealing of the probes requires the right distance and positioning of two antibodies, as well as the contribution of multiple reagents, the PLA reaction is intrinsically stochastic, with a probability increasing with, among other factors, the local density of the antibodies. It has been assumed that such high density reflects a specific concentration of two antigens, resulting from their physical interaction or their localization to the same subcellular structure. These are certainly favorable conditions for PLA, but the reaction may also be generated with any pair of antigens, even if they clearly do not interact, as shown here for GFP and cadherin or tubulin, the only apparent condition being a partial overlap of their distribution (Fig.1B' and Fig.5). The resulting PLA pattern will primarily be determined by this overlap, rather than by the actual distribution of the antigen, and

the density of spots generated in these overlapping regions will depend directly on the density of bound antibodies, not necessarily on a specific local accumulation of the two antigens. This easily explains the striking, seemingly “specific”, PLA decoration of membranes or microtubules obtained in our tests that used broadly distributed antigens. The illusion of a “specific proximity” may appear particularly convincing in cases where the two antibodies would appear by immunofluorescence to mark preferentially the same discrete structure (e.g. centrosome, cilium, or nucleolus, see symbolized yellow circle in Fig.1B,B’). These considerations lead us to conclude that in its current form, isPLA does not bring more information than classical immunofluorescence co-localization.

To be able to extract more useful information from isPLA, it would be imperative to set controls and criteria that would convincingly define a meaningful “proximity”. As mentioned above, classical controls only verify the specificity of the antibody, not of the proximity reaction. The standard negative control of the isPLA method, i.e. simple omission of one of the two primary antibodies (Bellucci *et al.*, 2014; Bagchi, Fredriksson and Wallén-Mackenzie, 2015; Debaize *et al.*, 2017), is insufficient, as it only controls for the potential non-specific binding of the secondary antibodies (and here PLA probes), which is generally very low anyway. We have confirmed here that indeed little to no PLA reaction is observed under this condition. Another previously suggested control is the use of a sample missing one of the antigens, either naturally or through knock-out/knock-down (Bellucci *et al.*, 2014; Bagchi, Fredriksson and Wallén-Mackenzie, 2015), but again, this control can only validate the specificity of the antibodies, not of the isPLA.

One clearly needs much more stringent tests. One possibility would be to compare the density of PLA spots to the levels of primary antibodies localizing at the structure of interest, taking as a reference point a “negative control” condition (suppl. Fig.S1). Unfortunately, we readily noticed

serious obstacles to implementing such control. A major problem is that the assay relies on antibody binding, which, unlike for instance FRET, does not necessarily faithfully reflect the position and levels of the molecules under investigation. Indeed, antibody binding varies widely, depending on affinity, sensitivity to fixation, and antigen availability/masking. In practice, we do not believe there is any objective criterion that would allow us to set appropriate antibody dilutions in order to quantitatively compare “positive” and “negative control” signals.

Our tests reveal additional levels of complexity and highlight the difficulty in drawing conclusions from PLA experiments. We were surprised that PLA for E-cadherin- β -catenin, which forms a 1:1 complex, was only marginally more effective than PLA for the “random” E-cadherin-mGFP pair (Fig.2 and Fig.S1). It is quite easy to conceive that despite the known direct interaction of E-cadherin with β -catenin, the two epitopes targeted by the antibodies may not be in the optimal configuration for efficient PLA. Note indeed that in addition to the absolute distance between the epitopes, their relative orientation (and therefore the position of the antibodies) may also influence the probability of a positive PLA reaction. Partial antigen masking within the dense adhesive structures could also decrease the probability of simultaneous binding of the two antibodies to the same cadherin-catenin complex. In any case, this example clearly shows that isPLA using a well-characterized pair of interacting partners may not be as efficient as originally expected.

The influence of complex parameters such as accessibility and orientation is supported by the line scans presented in supplemental Fig.S1, which show that PLA spots do not necessarily coincide with local peaks of antibody concentration. This is a general observation, which we made for all conditions tested (not shown). Furthermore, the high frequency of cytoplasmic PLA spots in the case of Sox9-Sam68 (Fig.4), despite the strong accumulation of both antibodies in the nucleus also argues in favor of this hypothesis.

Our analysis of nuclear PLA raises an additional issue. PLA enrichment at the periphery of this organelle, initially reported for the Sox9-Sam68 pair, was also observed for the non-specific anti-Sox9/non-specific IgG pair (Fig.4 and suppl. Fig.S3), which suggests the existence of an intrinsic bias independent of specific protein-protein interactions. We think that it may be linked to the density of the nuclear content, which, although not sufficient to significantly restrict the diffusion of primary and secondary antibodies (Fig.S3G), could be a more serious obstacle for the PLA reaction, which requires that the simultaneous convergence of multiple components (oligonucleotides, ligase, polymerase, fluorescent probes) on the same spot. While the nucleus is arguably the densest and largest structure of the cell, similar considerations about limiting diffusion and accessibility may apply in subtler ways to other parts of the cell.

In summary, we conclude that in its present form, isPLA cannot be trusted as a source of information about localization/interaction at the subcellular level.

Material and methods

Cells and transfections.

Human breast cancer MCF7 cells (ATCC #HTB-22) were grown in Dulbecco's Modified Eagle's medium (Life technologies, #31966047) with 10% FBS (Life technologies, #10500064). Cells were transfected with pCS2-eGFP or pCS2-mGFP (Maghzal *et al.*, 2010) using jetPRIME transfection reagent (Ozyme, #POL114-07), according to manufacturer instructions. For imaging, MCF7 cells were seeded on 12 mm No. 1.5 coverslips coated with 50ug/ml collagen type I (Corning, #354236).

Antibodies:

Antigen	Species	Conc. ($\mu\text{g/ml}$)	Supplier	Cat. number
GFP	Mouse	1.25	Life technologies	A11120
E-cadherin	Rabbit	0.33	Cell Signaling	3195S
β -catenin	Mouse	0.83	Invitrogen	13-8400
α -tubulin	Mouse	2.5	Sigma	T6199
non-specific IgG	Rabbit	20	Santa Cruz Biothechnology	SC-2027
SOX9	Mouse	1.0	Sigma	1.0
Sam68	Rabbit	4	Santa Cruz Biothechnology	SC-333

Proximity Ligation Assay and Immunofluorescence:

Cells were fixed for 10 min with 3.7% paraformaldehyde (EMS #15714) in PHEM buffer (60 mM Pipes, 25 mM Hepes, 8 mM EGTA 4 mM MgCl_2), followed by 10 min permeabilization with 1% Triton X100 (Applichem. Panreac., #A4975) in phosphate buffer saline, and blocking with 20% sheep serum in phosphate buffer saline for 1h. Primary antibodies in blocking buffer serum for 2h were added.

Proximity ligation assay was performed using the Duolink kit (Sigma-Aldrich DUO92102), according to manufacturer's protocol, using Orange red reagent (excitation 554 nm, emission 579 nm). After completion of the protocol, two additional secondary antibodies, coupled to green (Alexa488, ThermoFisher) and far red (Alexa 647, ThermoFisher) dyes, were added at $5\mu\text{g/ml}$, with the goal to directly detect each of the two primary antibodies. Samples were finally counterstained with Hoechst 33258 ($5\mu\text{g/ml}$, Invitrogen, 33342). Images of the four color signals were collected using a Leica SP5 laser scanning confocal microscope, with a 63X NA objective.

Acknowledgments

We thank David Rozema and Dominique Hemlinger for critical reading of the manuscript. F.F. research is supported by the ANR grant ANR-14-ACHN-0004-ICM and by the Labex EpiGenMed.

References

- Bagchi, S., Fredriksson, R. and Wallén-Mackenzie, Å. (2015) ‘In Situ Proximity Ligation Assay (PLA).’, *Methods in molecular biology (Clifton, N.J.)*, 1318, pp. 149–159. doi: 10.1007/978-1-4939-2742-5_15.
- Bellucci, A. *et al.* (2014) ‘The “in situ” proximity ligation assay to probe protein-protein interactions in intact tissues.’, *Methods in molecular biology (Clifton, N.J.)*, 1174, pp. 397–405. doi: 10.1007/978-1-4939-0944-5_27.
- Blokzijl, A. *et al.* (2014) ‘Protein biomarker validation via proximity ligation assays.’, *Biochimica et biophysica acta*, 1844(5), pp. 933–939. doi: 10.1016/j.bbapap.2013.07.016.
- Bunse, L. *et al.* (2015) ‘Proximity ligation assay evaluates IDH1R132H presentation in gliomas.’, *The Journal of clinical investigation*, 125(2), pp. 593–606. doi: 10.1172/JCI77780.
- Debaize, L. *et al.* (2017) ‘Optimization of proximity ligation assay (PLA) for detection of protein interactions and fusion proteins in non-adherent cells: application to pre-B lymphocytes.’, *Molecular cytogenetics*, 10, p. 27. doi: 10.1186/s13039-017-0328-2.
- Fredriksson, S. *et al.* (2002) ‘Protein detection using proximity-dependent DNA ligation assays.’, *Nature biotechnology*, 20(5), pp. 473–477. doi: 10.1038/nbt0502-473.

Girardot, M. *et al.* (2018) ‘SOX9 has distinct regulatory roles in alternative splicing and transcription.’, *Nucleic acids research*, 46(17), pp. 9106–9118. doi: 10.1093/nar/gky553.

Gomes, I., Sierra, S. and Devi, L. A. (2016) ‘Detection of Receptor Heteromerization Using In Situ Proximity Ligation Assay.’, *Current protocols in pharmacology*, 75, pp. 2.16.1-2.16.31. doi: 10.1002/cpph.15.

Greenwood, C. *et al.* (2015) ‘Proximity assays for sensitive quantification of proteins.’, *Biomolecular detection and quantification*, 4, pp. 10–16. doi: 10.1016/j.bdq.2015.04.002.

Gullberg, M. *et al.* (2004) ‘Cytokine detection by antibody-based proximity ligation.’, *Proceedings of the National Academy of Sciences of the United States of America*, 101(22), pp. 8420–8424. doi: 10.1073/pnas.0400552101.

Jalili, R. *et al.* (2018) ‘Streamlined circular proximity ligation assay provides high stringency and compatibility with low-affinity antibodies.’, *Proceedings of the National Academy of Sciences of the United States of America*, 115(5), pp. E925–E933. doi: 10.1073/pnas.1718283115.

Johnson, S. S. *et al.* (2018) ‘Fingerprinting Non-Terran Biosignatures.’, *Astrobiology*, 18(7), pp. 915–922. doi: 10.1089/ast.2017.1712.

Lipovsky, A. *et al.* (2015) ‘Application of the proximity-dependent assay and fluorescence imaging approaches to study viral entry pathways.’, *Methods in molecular biology (Clifton, N.J.)*, 1270, pp. 437–451. doi: 10.1007/978-1-4939-2309-0_30.

Lutz, M. I. *et al.* (2017) ‘Novel approach for accurate tissue-based protein colocalization and proximity microscopy.’, *Scientific reports*, 7(1), p. 2668. doi: 10.1038/s41598-017-02735-8.

Maghzal, N. *et al.* (2010) ‘The tumor-associated EpCAM regulates morphogenetic movements

through intracellular signaling’, *Journal of Cell Biology*, 191(3), pp. 645–659. doi: 10.1083/jcb.201004074.

Mereiter, S. *et al.* (2016) ‘Glycomic Approaches for the Discovery of Targets in Gastrointestinal Cancer.’, *Frontiers in oncology*, 6, p. 55. doi: 10.3389/fonc.2016.00055.

Söderberg, O. *et al.* (2006) ‘Direct observation of individual endogenous protein complexes in situ by proximity ligation.’, *Nature methods*, 3(12), pp. 995–1000. doi: 10.1038/nmeth947.

Söderberg, O. *et al.* (2008) ‘Characterizing proteins and their interactions in cells and tissues using the in situ proximity ligation assay.’, *Methods (San Diego, Calif.)*, 45(3), pp. 227–232. doi: 10.1016/j.ymeth.2008.06.014.

Weibrecht, I. *et al.* (2010) ‘Proximity ligation assays: a recent addition to the proteomics toolbox.’, *Expert review of proteomics*, 7(3), pp. 401–409. doi: 10.1586/epr.10.10.

Figure legends

Figure 1. A. Principle of the in-situ proximity ligation assay coupled to double immunofluorescence labeling. After incubation with two primary antibodies, secondary antibodies with PLA probes are added and the PLA reaction is performed (details of the reaction omitted). The reaction can only occur if the two antigens are closer than 40nm. In a subsequent step, regular fluorescently labeled secondary antibodies are added in order to determine the distribution and levels of each primary antibody. In the experiments presented here, we used the orange, fluorescent PLA reagent (emission at 579 nm), and green (Alexa488-conjugated), and far-red (Alexa647-conjugated) secondary antibodies. Nuclei were counterstained with Hoechst (not shown). **B, B'. Specific versus fortuitous proximity.** Assuming that two antibodies yield a colocalization pattern (yellow spot in the cell drawing), we ask whether isPLA can further discriminate between specific proximity due to the association of the two antigens within a protein complex (or a subcellular structure, in the order of few tens of nanometers) or fortuitous proximity due to the high antigen/antibody global concentrations within the region observed by immunofluorescence. In B', the green antibody marks a protein associated with the subcellular structure, the red antibody recognizes a randomly distributed antigen. Antigens, secondary antibodies, and probes are omitted for clarity's sake.

Figure 2. isPLA signal at the plasma membrane. A. isPLA for endogenous E-cadherin and β -catenin. **B, C.** isPLA using anti-E-cadherin and anti-GFP antibodies on cells ectopically expressing membrane-targeted GFP (mGFP)(B) or soluble GFP (GFP)(C). In all cases, a positive isPLA signal was observed along the plasma membrane (arrows). **D-F.** High magnification from A-C. Arrows point to spots along the plasma membrane. Arrowheads in F point to examples of PLA

corresponding to cytoplasmic, E-cadherin-positive structures. See supplemental Figure S1 for additional examples, negative controls, and quantifications. All images were captured by laser scanning confocal microscopy using the exact same settings. Scale bars, A-C, 5 μ m; D-F, 2 μ m.

Figure 3. isPLA signal at microtubules.

A. isPLA for endogenous α -tubulin and soluble GFP. Detail of two cells expressing moderate (cell 1) and weak levels (cell 2) of soluble GFP (pseudocolors included for better comparison, see Fig.S2 for full fields). Numerous spots are observed along microtubules. Spots are denser in the cell expressing higher levels of GFP. **B.** isPLA for endogenous α -tubulin and membrane-targeted GFP (mGFP). Detail of two cells expressing mGFP. Numerous spots are found at the cell periphery, where most of the tubulin and mGFP signals overlap. **C.** isPLA using anti- α -tubulin and a non-specific rabbit serum. The dilution of the serum was adjusted as to yield a non-specific signal of an intensity comparable to the intensity for the anti-GFP antibody. Microtubules are decorated by PLA spots. Scale bars, 5 μ m.

Figure 4. isPLA in the nucleus.

A,B. isPLA for Sox9, chosen as an example of nuclear protein, and either Sam68, a candidate nuclear interacting partner (A), or non-specific serum (NI)(B). For both conditions, PLA spots were observed both in the nucleus (arrows) and in the cytoplasm (arrowheads). Nuclear spots are mostly found at the periphery of the nuclei (see Fig.S3 for more details). **C.** Negative control, anti-Sox9 antibody omitted. Scale bars, 5 μ m.

Figure 5. Summary diagram. Panels A-F show the general cellular distribution of the PLA signal (orange dark-circled dots), and the primary antibodies (green and red). Panels A'-E' represent the detailed situation generating PLA in each case. Secondary antibodies and probes are omitted for clarity's sake. isPLA yields positive signals not only for biologically meaningful pairs of antigens,

here E-cadherin and β -catenin (A), but potentially for any pair of antigens, provided that a small fraction of the two primary antibodies bind sufficiently close to allow for the ligation reaction (B-F). The pattern thus obtained may then give the illusion of a “specific interaction”, for example along the plasma membrane (B,C), at the intersection between microtubules and the plasma membrane (D), or decorating microtubules (E,F), even when one of the targeted antigens is widely distributed, such as in the case of soluble GFP (C,E,F). The density of the PLA spots depends on the levels of primary antibodies (E and F), which are dictated by antigen abundance, antibody concentrations, and additional parameters (see discussion).

Supplementary figure legends

Figure S1. isPLA signal at the plasma membrane.

A-E. Comparison of isPLA for endogenous E-cadherin with β -catenin and with different levels of ectopic mGFP/GFP. A,B and D are as in Figure 1.

F,G. Negative controls (omission of anti-E-cadherin, or anti- β -catenin primary antibodies).

H,I. Examples of line scans along plasma membranes. Plots represent red (anti-E-cadherin) and green (anti- β -catenin or anti-GFP) fluorescence intensities. Arrows indicate the position of PLA spots. **J.** PLA spot densities along plasma membranes plotted as a function of green fluorescence intensity.

Figure S2. isPLA signal along microtubules. A-D. Merge of the four channels. A'-D'. Merge of tubulin, PLA, and Hoechst signals. A''-D''. PLA signal alone. A'''-D'''. Green (GFP or non-specific) signals displayed as pseudocolors. **A,B.** isPLA for α -tubulin and GFP (A,B). **C.** isPLA for α -tubulin and non-immune serum (C). **D.** Negative control (omission of anti-GFP antibody). B and C are full fields of images of Fig.2.

Figure S3. isPLA signal in the nucleus.

A,B. Two examples of siPLA Sox9-Sam68. A corresponds to the full field of A in Figure 3. **C.** siPLA Sox9-non-specific serum (NS). Full field of B in Figure 3. **D.** Negative control (omission of anti-Sam68). **E,F.** Histogram of distribution of PLA spots in nuclei, measured in μm from the edge of the nucleus. Median values are 1.1 μm for Sox9-Sam58 and 0.8 μm for Sox9-NI. The average nucleus diameter was $\sim 8 \mu\text{m}$. **G.** Superposition of PLA spot distribution (as in E) with average fluorescence intensity profiles for Hoechst, anti-Sox9, and anti-Sam68.

Figure 1

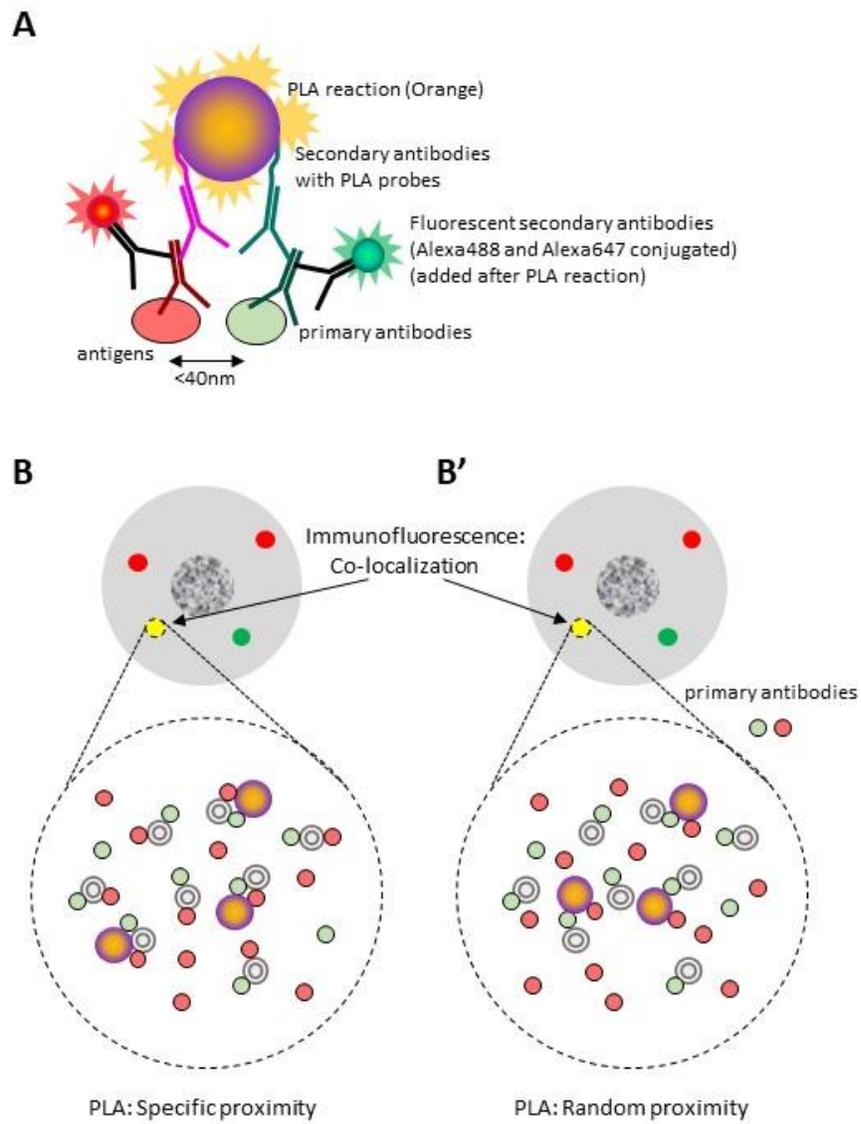


Figure 2

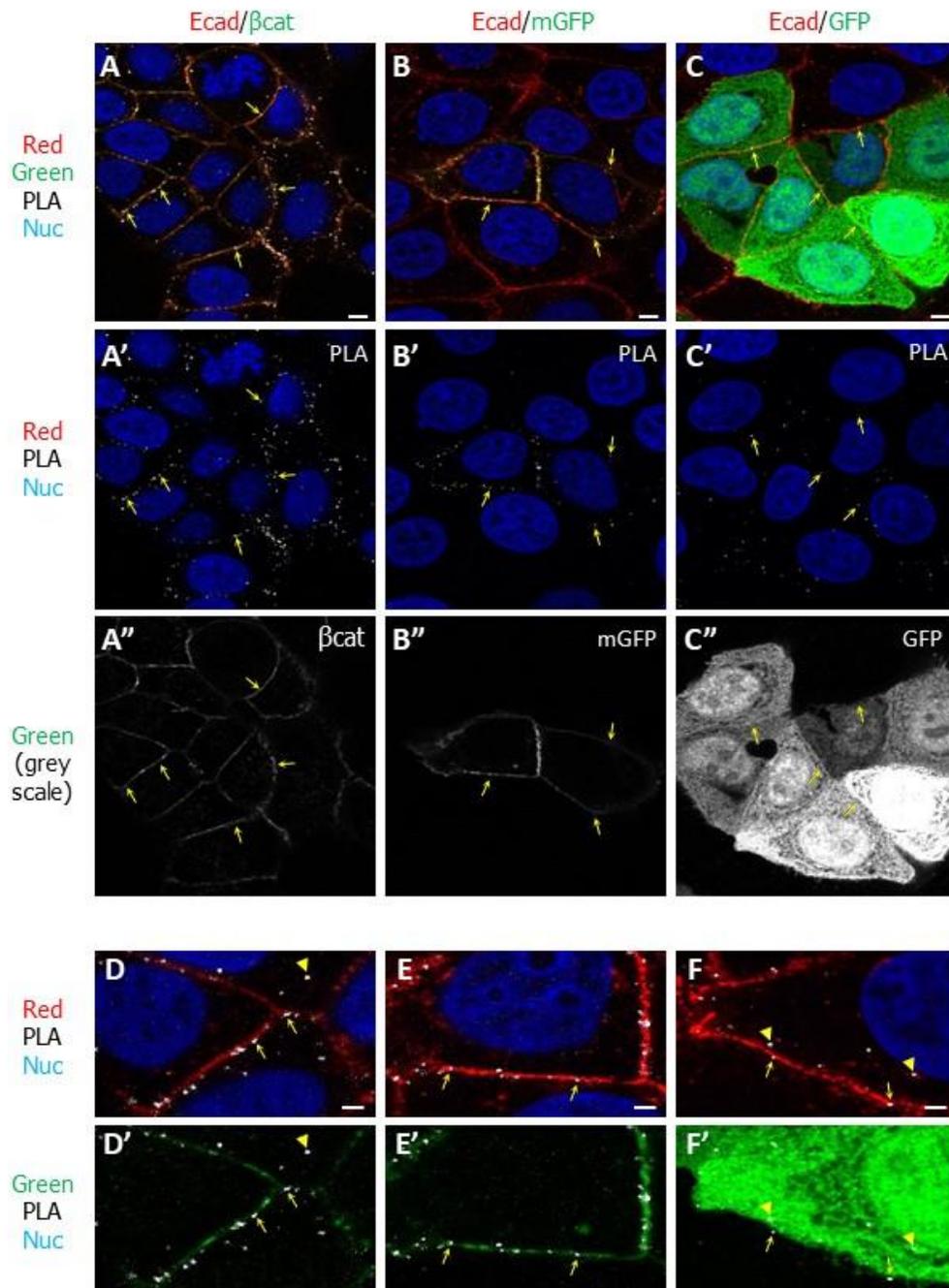


Figure 3

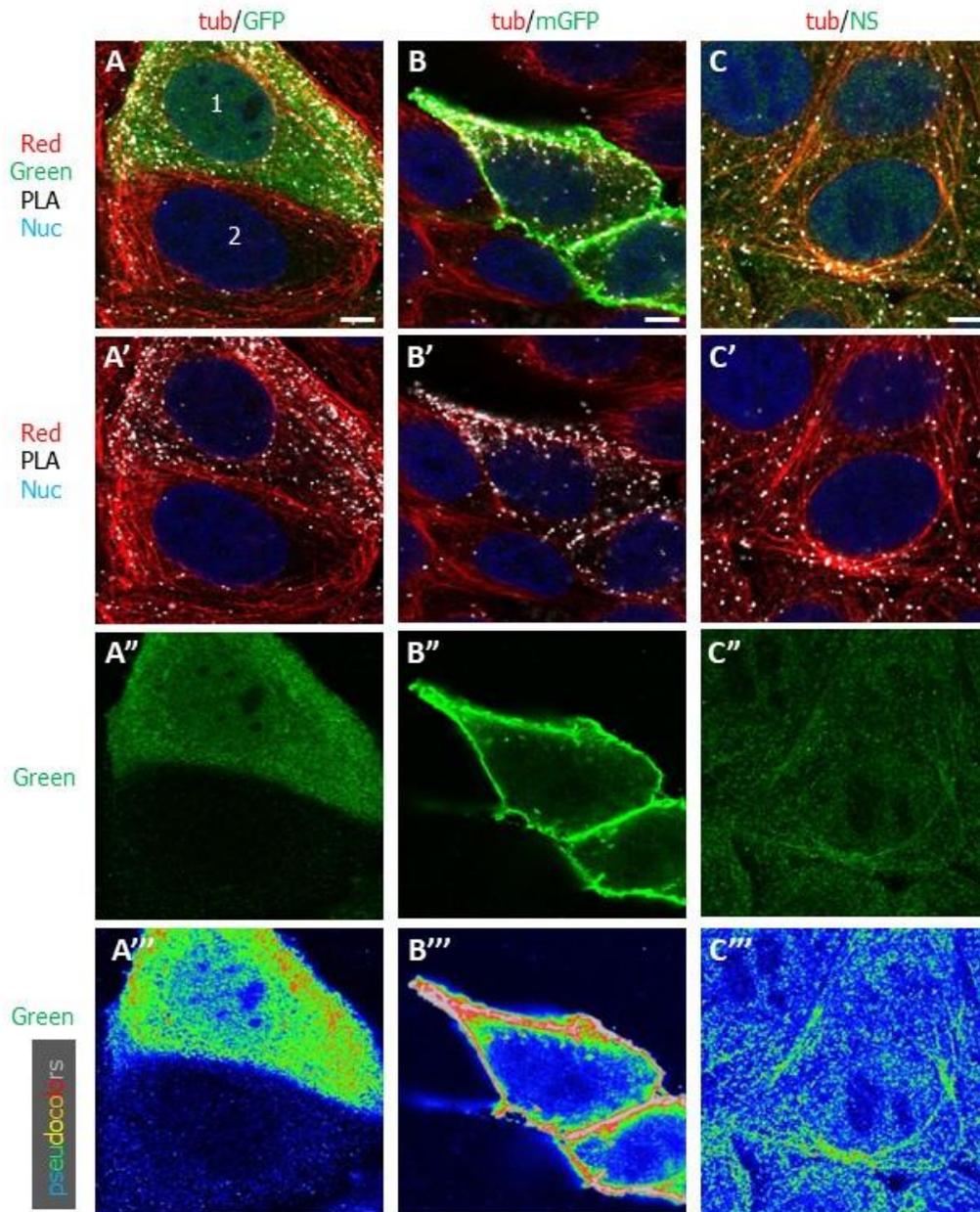


Figure 4

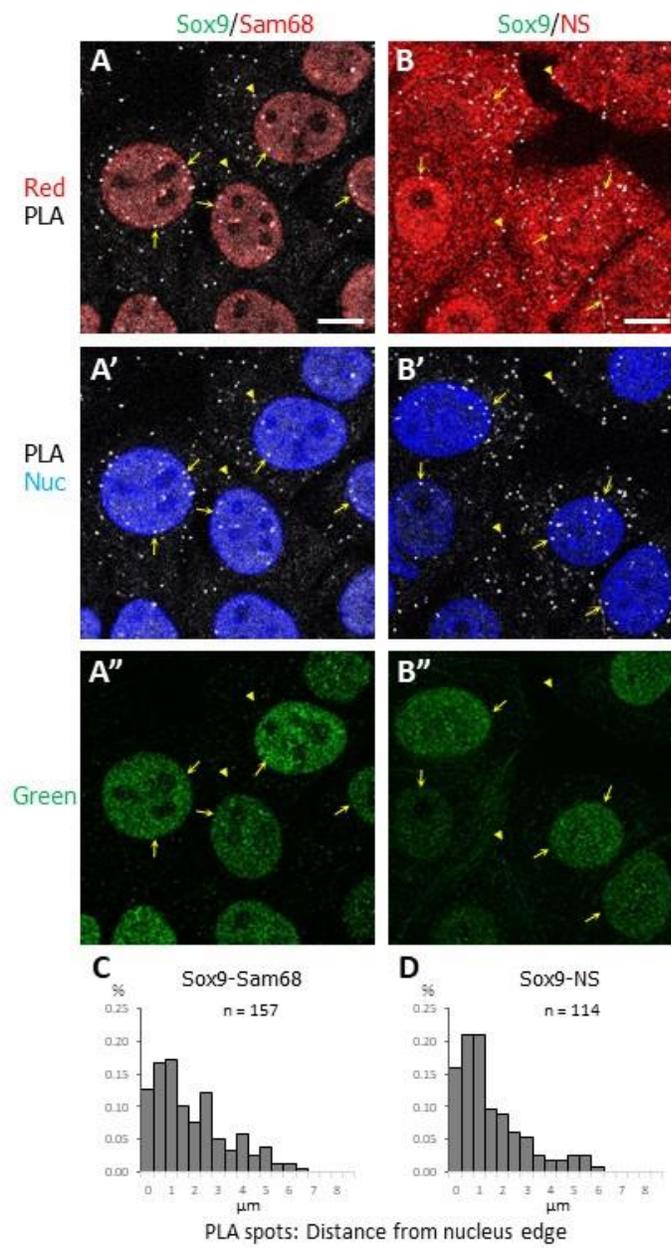
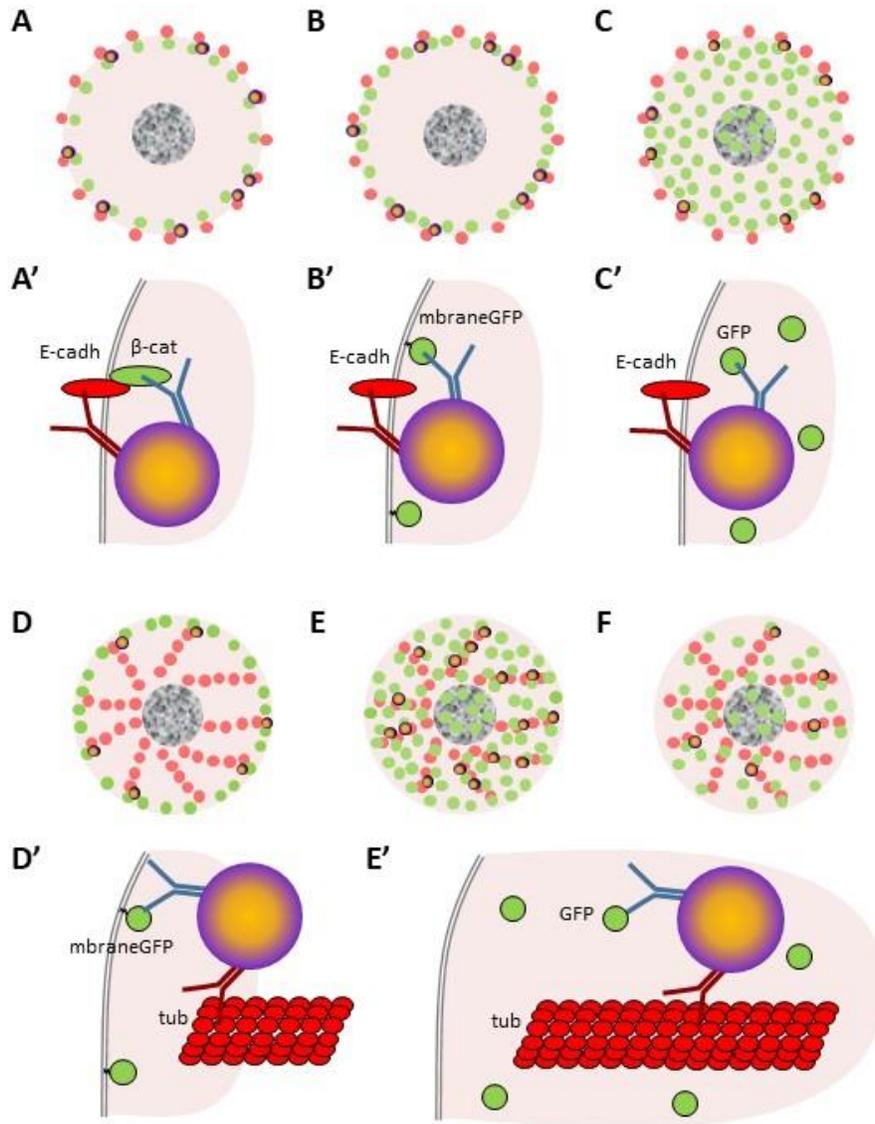
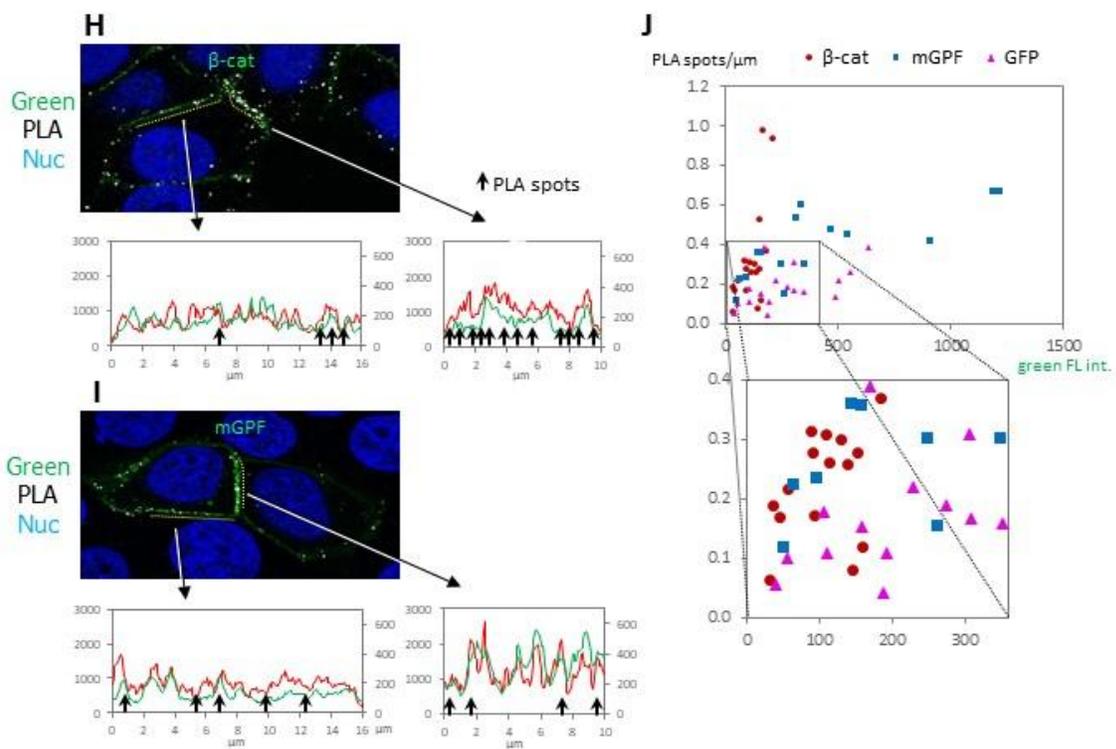
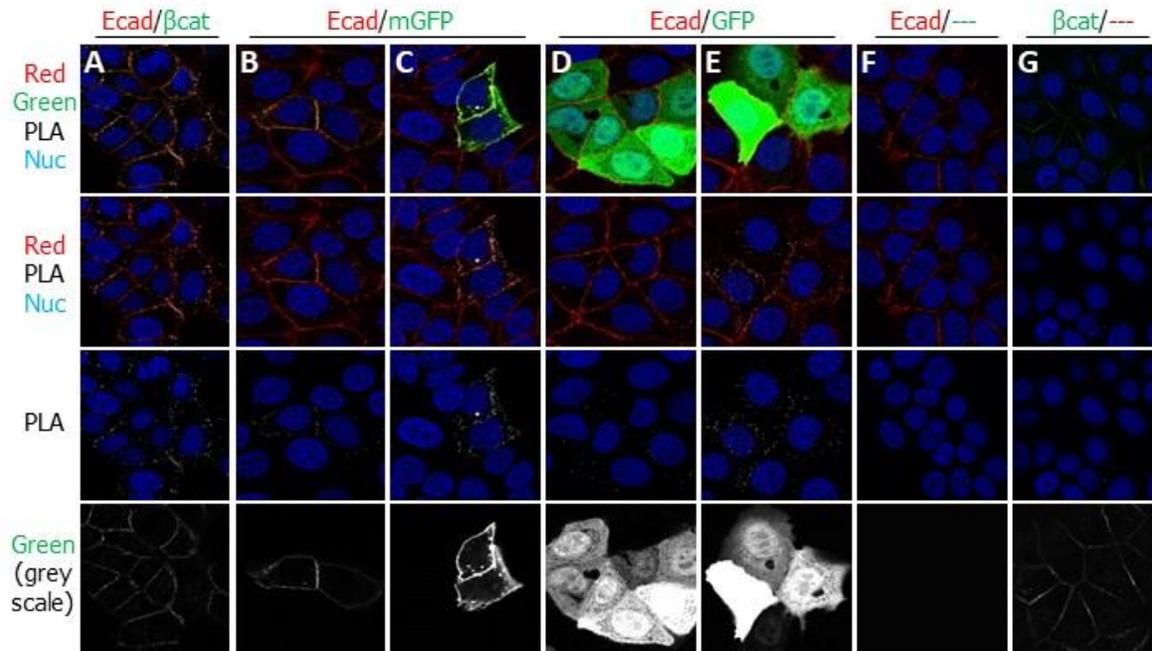


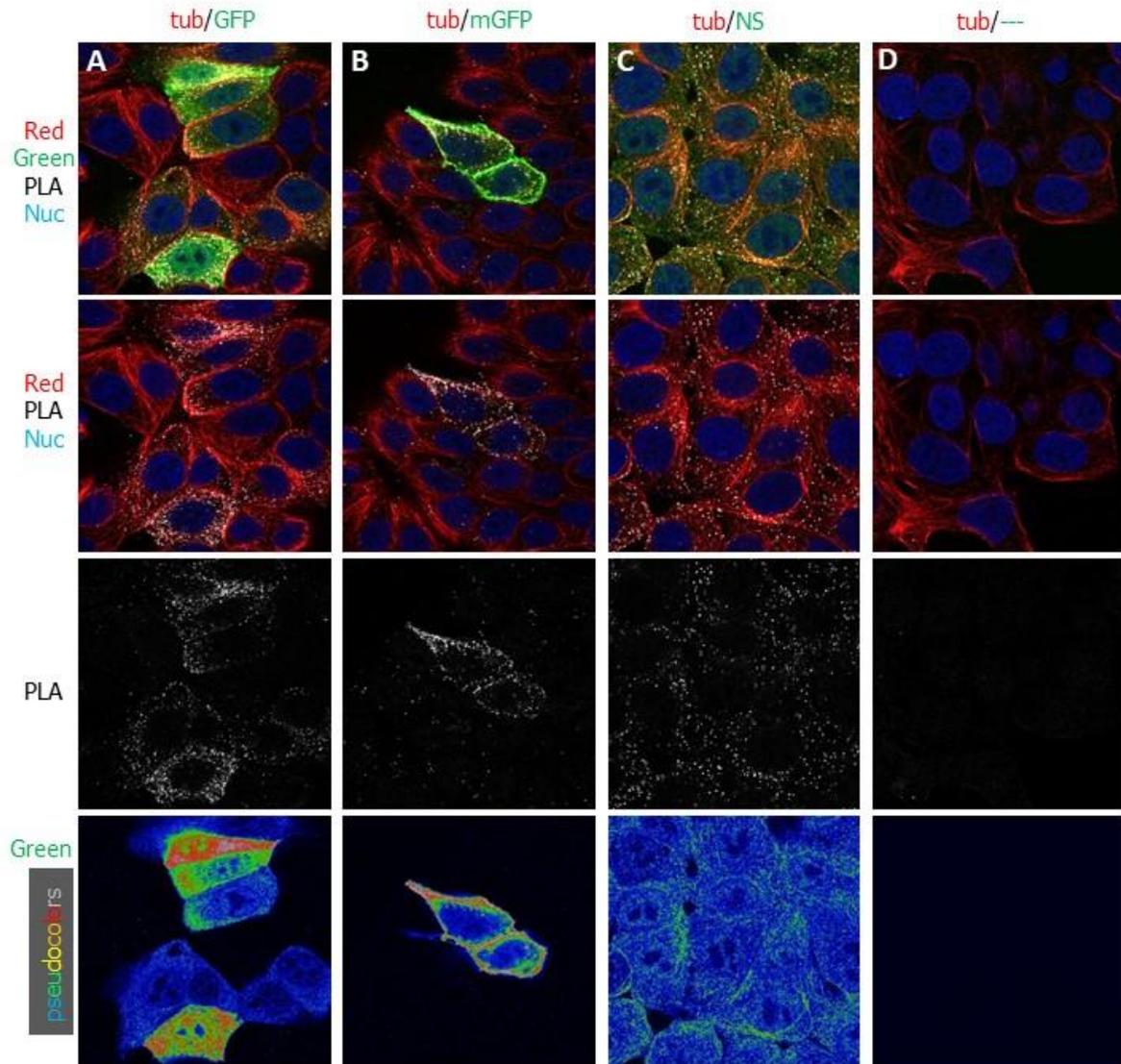
Figure 5



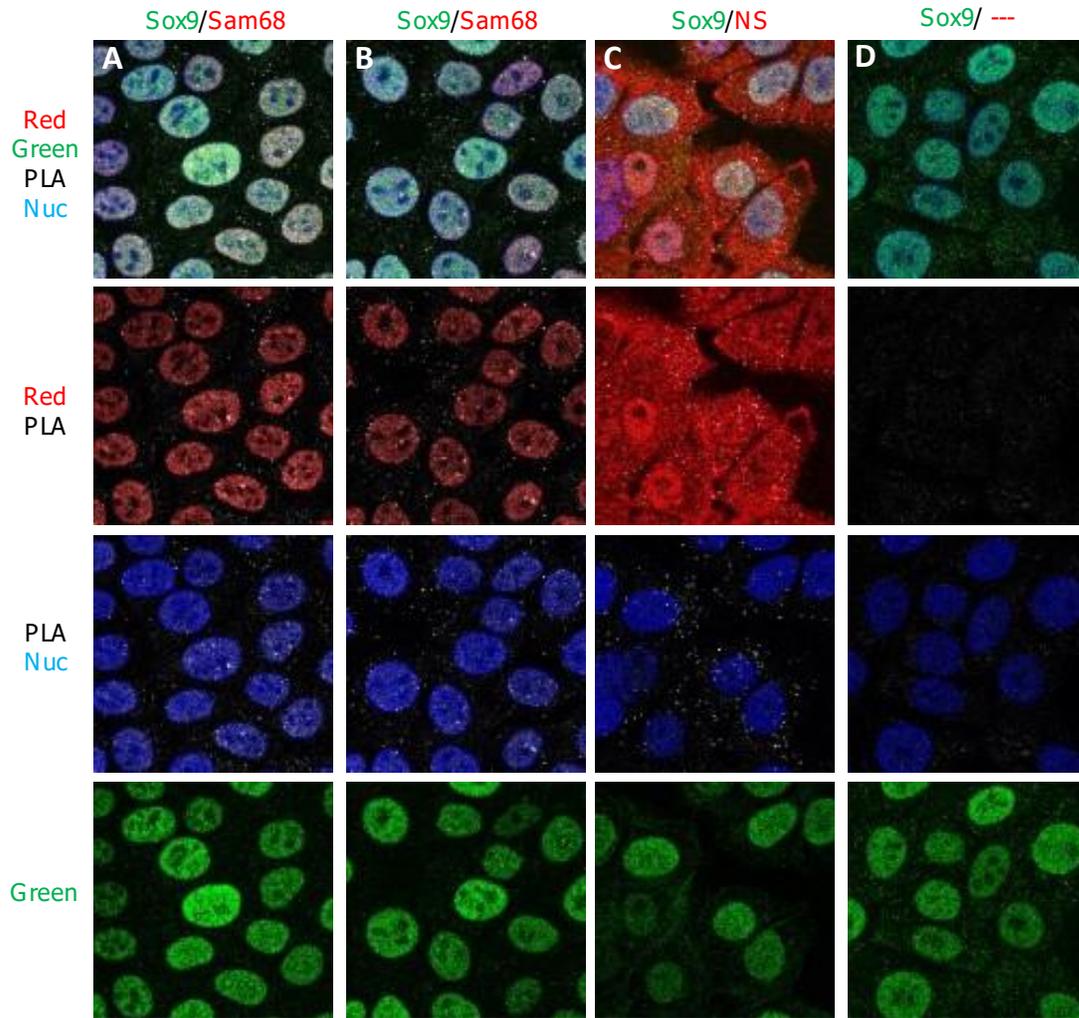
Supplemental Figure S1



Supplemental Figure S2



Supplemental Figure S3

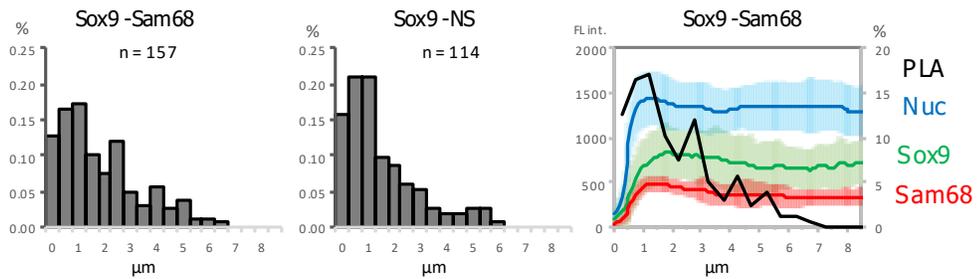


E

F

G

PLA spots: Distance from nucleus edge



CHAPTER IV

Final discussion

Discussion

The work presented in this thesis aimed to understand the function of EpCAM as a regulator of cell adhesion and migration in the context of human carcinoma MCF7 cells. In this final chapter, I summarize the main findings, I further discuss how these data may contribute to our understanding of EpCAM role in cancer progression, and I point out some remaining open questions that could be addressed in the future.

Chapter II- Summary of findings

The first and main finding/contribution of this thesis was the discovery that depleting EpCAM levels in MCF7 cancer cells impacted their individual and collective migratory behavior in opposite ways, preventing single-cell migration, but stimulating migration of cells as groups. In addition, the induced collective migration occurred in a very coherent mode that prevented detachment of individual and/or small groups of cells from the main mass of cells, which, even though relatively infrequent, occurred in control EpCAM expressing cells. We showed that EpCAM depleted cells were more contractile and at the same time highly adhesive (cell-cell contacts). The former was consistent with its role as a negative regulator of myosin activity, while the high adhesiveness related to increased cadherin recruitment at cell-cell contacts, a typical mechanism of reinforcement in response to tension (Charras and Yap, 2018). These features account for the changes in collective migration resulting from EpCAM depletion. As for the inhibition of single-cell migration, we did not find obvious signs in the contacts with the extracellular matrix which could account for this effect. However, we could explain it by the fact that these cells seemed to fail to establish the front-rear polarity typically required for migration.

Another main contribution of this work was to include and study the only other member of EpCAM family, Trop2, in parallel with EpCAM. This type of direct comparison is a serious gap in the field. While these two molecules have been assumed to be largely redundant, we found that Trop2 loss-of-function gave the exact opposite cellular phenotypes: Trop2 depletion stimulated individual cell migration but prevented migration of cells collectively. One important result of further characterization of these phenotypes was the demonstration that Trop2-depletion also resulted in increased contractility, thus confirming the close similarity with EpCAM in its molecular regulation of myosin activity. This similarity in molecular activity but opposite cellular phenotypes provided us with a very interesting situation, which helped us to identify the parameters that controlled cell behavior downstream of myosin up/downregulation. We could indeed determine differences in the balance of myosin activity and tension between the free cortex, cell contacts, and matrix adhesions. In addition, we found that the reinforcement of cadherin contacts was weaker in the case of Trop2 depletion. This partial differences in tensile and adhesive reactions coincide with partial differential localization of EpCAM and Trop2. We propose that while both molecules share a general function in moderating global actomyosin cortical contractility, they also tend to preferentially control myosin pools involved in distinct myosin-dependent reactions, with EpCAM targeting more cell adhesion reinforcement, Trop2 more tension exerted on cell-matrix adhesion. The combination of these various partial differences appears to account for the opposite EpCAM and Trop2 loss-of-function phenotypes. They also explain the opposite single cell and collective behaviors, the former being influenced by the balance between cortical tension and matrix adhesion, while in spheroids, the properties of cell-cell contacts become a dominant factor.

Together, the data presented in this thesis provide important information about the impact of EpCAM and Trop2 expression in cancer cells and offer predictions for conditions that may favor either their pro-invasive or on the contrary anti-invasive, potential.

Chapter II- Future perspectives

Understanding the mechanisms underlying the dramatic differences in EpCAM loss-of-function tissue phenotypes between various experimental systems, *Xenopus* embryonic tissues, intestinal epithelium, and MCF7 spheroids, remains a central question that should be pursued. As discussed above, simple parameters such as the intensity of myosin upregulation could in principle account for diametrically opposed phenotypic outputs, consistent with the extreme contractile reaction of embryonic tissues in response to EpCAM depletion. Yet, this simple model is not quite satisfactory, in particular, because Trop2 depletion in MCF7 cells reduces collective migration, but double EpCAM/Trop2 depletion rescues the efficient spheroid spreading obtained with single EpCAM depletion. Therefore, there is clearly more to be found behind these effects.

Among other alternative or complementary hypotheses, one aspect that seems consistent with these distinct phenotypes is the differential localization of the EpCAM and Trop2 in human cells, as reported and discussed in chapter II. One obvious next step would be to determine the cause for this difference. The question is fully open: It could be due, for instance, preferential interaction with other membrane components, either proteins or membrane microdomains, but it could also result from differential stability at the membrane: I mentioned in chapter II that Trop2 is abundant in intracellular compartments (endosomes and lysosomes, data not shown), which may reflect a lower stability, perhaps specifically at cell-cell contacts. There is a phosphorylation site (Ser 303) in the cytoplasmic tail of Trop2, absent in EpCAM (Fagotto & Aslemarz, 2020). Whether

it might be involved in localization/stability could be readily tested by expressing phosphomimetic and phosphodeficient mutants. Note that we could not find evidence in MCF7 cells for the recently proposed mechanism of EpCAM action via EpCAM-RhoA fast co-recycling (Gaston *et al.*, 2021).

One other related parameter that may influence subcellular localization could be expression levels: EpCAM is abnormally highly expressed in breast cancer cells, which might then saturate its “normal” localization (e.g. in submembrane domains) and occupy unusual locations, which may change its activity, or simply the balance of myosin contractility. Previous work in *Xenopus* embryos showed that EpCAM stimulated intracellular migration only at moderate levels, but inhibited it at higher levels (Maghzal *et al.*, 2010). The latter situation could well be precisely the rheological “regime” reached by EpCAM in wild-type MCF7 cells. This would be also consistent with the fact that overexpressing EpCAM in MCF7 cells reduced even more their spheroid collective migration. Thus, one interesting perspective would be to investigate the collective migratory properties of MCF7 cells with more moderate EpCAM levels. Fine-tuning EpCAM levels could be easily done by coupling siRNA with inducible (siRNA resistant) EpCAM expression, for which I have already established the tools. It would also be interesting to investigate cells that naturally express lower EpCAM, in particular breast epithelial cell lines, such as MCF10A. Note that I tested a potential cross-regulation of EpCAM and Trop2 expression levels. I observed a modest but reproducible increase in EpCAM in Trop2 depleted cells and no effect of EpCAM depletion on Trop2 levels (data not shown). While these preliminary data do not support an important role of such cross-regulation in MCF7 cells, they should be further validated and could be relevant under different conditions and/or for other cell types.

While differences in EpCAM expression levels are without contest an important parameter to take into account, it remains they could hardly explain the opposite effect of Trop2 depletion,

and the rescued phenotype in the double depletion observed in our experiments. Note that we have estimated relative protein expression levels at the membrane of MCF7 cells and concluded that Trop2 is expressed 5 to at most 10% of EpCAM levels (data not shown). These considerations argue that there must be qualitative differences between human EpCAM and Trop2, but also that the two molecules must crosstalk when co-expressed. One sign for such cross-interaction lays in the following striking correlation: The “anti-cohesive” EpCAM LOF phenotype occurs in *Xenopus* and in human intestine, i.e. the two systems where only EpCAM is expressed. In MCF7 cells, where both EpCAM and Trop2 are expressed, EpCAM LOF gives an opposite “cohesive” phenotype, while Trop2 LOF appears to go in the same direction as EpCAM LOF in the other systems. This may suggest that EpCAM and Trop2 in amniotes have diverged to share out specialized functions. Specialization can only be very partial, however, as both have maintained largely similar properties, and can clearly compensate, as discussed in Chapter II. The distinct, although overlapping, subcellular localization uncovered in this study is one such property. Although it may be sufficient to explain the different phenotypes, there is certainly more to be discovered, starting with the molecular mechanisms that control these localizations and could well also influence other aspects of EpCAM and Trop2 biology.

MCF7 cells, with the clear cut opposite effects of EpCAM and Trop2 depletion, both for single-cell and collective migration, should be a very good model to go deeper in the side-by-side characterization of EpCAM and Trop2, with the aim to identify the molecular mechanisms responsible for their overlapping as well as their specific properties. Besides the above-mentioned potential role of phosphorylation, another aspect to explore is the role of the extracellular domain. While this domain is dispensable for global myosin repression (Maghzal *et al.*, 2010), it is required to rescue tissue integrity (Maghzal *et al.*, 2013). The strong doubts casted on the direct role of

EpCAM in adhesion (Gaber *et al.*, 2018) do not eliminate a possible weak homophilic interaction, which would be consistent with the original data by Litvinov team (Litvinov *et al.*, 1994). Such interactions could for instance contribute to stabilizing EpCAM at cell-cell contacts, and perhaps even influence its signaling activity. Alternatively, the EpCAM extracellular domain may be important for interaction with other cell surface components. In any case, differences in the extracellular domain may also account for the different properties of EpCAM and Trop2.

An important related open question is the characterization of the downstream pathway(s). Rescue of the siEpCAM spheroid spreading phenotype by calphostin demonstrates that nPKCs are involved, similar to what is shown in *Xenopus* embryonic tissues. In the latter system, inhibition of the nPKC-PKD-Erk pathway was sufficient to account for the function of EpCAM (Maghzal *et al.*, 2010, 2013). The fact that the response of cellular adhesive structures to the high contractility was different in MCF7 cells raises the possibility that other pathways may also be involved. For instance, PKCs are known to have a multitude of targets, some of them directly impinging on the actomyosin cytoskeleton (Rosse *et al.*, 2010, Maghzal *et al.* 2013). Conceivably, distinct pathways/targets may be preferentially affected in response to EpCAM and Trop2 depletion, with different impacts on adhesiveness and migration. Determining to what extent the canonical nPKC-PKD-Erk-myosin pathway contributes to the various phenotypes observed in MCF7 cells will be an obvious first step in addressing this question, as well as identifying other targets and their potential differential regulation by EpCAM or Trop2.

In our current work, we aimed to design experimental approaches close to *in vivo* conditions relevant for cancer. There is clearly room for further exploration of various parameters. Substrate rigidity, for instance, is key for cell behavior (Keely and Nain, 2015). This was confirmed in our experiments where the low cohesiveness of MCF7 spheroids became more

evident on polyacrylamide gels compared to pure collagen gels, and I have preliminary data suggesting that it varies with the rigidity of polyacrylamide gels. It will be interesting to further characterize EpCAM-dependent cohesiveness and migration on a range of substrate rigidities, as well as different collagen densities.

More generally, broadening the present studying to different cancer cell lines would be an important contribution to the understanding of EpCAM (and Trop2) role in cancer invasion. During this work, I have set and optimized a range of tests for a wide scale from single cells to large multicellular spheroids that could be immediately applied to any other cell line/type.

Although EpCAM is restricted to epithelia and studying cells as groups seems to be more relevant, our study suggests that dissecting distinct cellular parameters at the level of single cells, cell doublets, and small groups of cells gives important information to interpret the more complex collective behavior. Moreover, the properties of individual cells are, on their own, very relevant in the context of cancer metastasis, in order to predict the potential for cells to detach from the primary tumor and initiate invasion. My results show that EpCAM (and Trop2) levels have a completely different impact on the morphogenetic properties of cell populations and on the migration of individual cells. Expanding this analysis to other cell lines may help settle the controversial role of EpCAM as pro- or anti-invasive.

Our observation of sorting of EpCAM positive from EpCAM negative cells is very exciting considering the potential impact on understanding the role of EpCAM in real tumors. While we concluded that the sorting was a result of differential cortical tension, adhesiveness, and also differential migratory behavior of EpCAM positive and negative cells, the underlying mechanisms driving this behavior could more be complex and deserves further investigation. Also, it would be important to study such mixed populations beyond 24hrs, in particular, to determine their long-

term configuration. One may predict, for instance, that while the bulk of EpCAM-positive cells may remain trapped in the core of the spheroid, the few EpCAM-positive cells that may be “carried” away by the collective migration of the bulk of EpCAM-negative cells may ultimately escape and become “invasive”.

As a final note, one should highlight that while cancer cell lines are generally accepted to be used *in vitro* as cancer models, they do not necessarily accurately reflect the cell properties during the process of primary metastasis. A current focus in the field of cancer is to study the actual circulating tumor cells (CTCs). A recent key step has been the establishment of CTC-derived stable lines (Soler *et al.*, 2018). Since these studies have revealed a strong correlation between CTCs malignancy and epithelial markers, including EpCAM, these lines could be used as a highly relevant cell model to study the role of EpCAM in invasion, as well as in the ability to form secondary tumors. Answering this question would be of high impact both for validating the relevance of EpCAM expression in the context of diagnostic, and for a better understanding of the role of this molecule in cancer progression. All the experimental settings optimized in this thesis could be immediately transferred to such CTC models.

Chapter III-Summary of findings

We had originally planned to use the “proximity ligation assay” (PLA) for our study on EpCAM, in particular, to characterize potential subcellular locations for EpCAM-PKC interaction. However, we realized by running specificity controls that the technique is intrinsically prone to produce false positive results. Specifically, a PLA reaction will be produced whenever two antibodies bind close enough from each other, whether this event reflects a specific, physiologically relevant proximity of the corresponding antigens, or it is simply fortuitous, thus

without biological significance. Surprisingly, among the hundreds of publications that presented PLA data, none ever presented the appropriate controls. Our work, which involved a systematic analysis of this issue, and yielded indisputable conclusions, constitutes an important methodological contribution. It calls for looking at published conclusions based on PLA with high skepticism. It should also be a strong incentive for the development of better techniques for in situ investigations of molecular interactions.

Chapter III-Future perspectives

As explained in the related chapter the negative controls that are usually used in this assay e.g., excluding one of the primary antibodies or one of the targets, only serve as negative control for the specificity of the secondary antibodies. In this regard, the PLA reaction is highly specific indeed, and one can conclude that it does reflect the proximity of two primary antibodies. However, the real question to solve is whether this proximity is meaningful. Answering this question would require finding a different types of control conditions. One possibility would be to compare the PLA signal from the assumed “specific” reaction, i.e. of the two antibodies targeting the pair of antigens under investigation, with the PLA signal from a control “random” reaction. What is meant by “random” is that the pair of antibodies used in the control reaction would also show an overlapping distribution when observed by regular immunofluorescence, ideally to a similar extent as the “test” pair but would not be expected to specifically bind in close proximity. Thus, a proper PLA experiment should be systematically coupled to indirect immunofluorescence, as shown in Chapter III, in order to determine the distribution of all primary antibodies, and should include antibody titrations, in order to achieve comparable conditions. This type of approach should improve the confidence that a PLA signal is indeed meaningful as reflecting specific proximity, and thus also potentially specific molecular interactions.

References

- Charras, G. and Yap, A. S. (2018) 'Tensile Forces and Mechanotransduction at Cell-Cell Junctions.', *Current biology : CB*, 28(8), pp. R445–R457. doi: 10.1016/j.cub.2018.02.003.
- Fagotto, F. and Aslemarz, A. (2020) 'EpCAM cellular functions in adhesion and migration, and potential impact on invasion: A critical review', *Biochimica et Biophysica Acta - Reviews on Cancer*, 1874(2), p. 188436. doi: 10.1016/j.bbcan.2020.188436.
- Gaber, A. *et al.* (2018) 'EpCAM homo-oligomerization is not the basis for its role in cell-cell adhesion', *Scientific Reports*, 8(1), pp. 1–16. doi: 10.1038/s41598-018-31482-7.
- Gaston, C. *et al.* (2021) 'EpCAM promotes endosomal modulation of the cortical RhoA zone for epithelial organization', *Nature Communications*, 12(1), pp. 1–22. doi: 10.1038/s41467-021-22482-9.
- Keely, P. and Nain, A. (2015) 'Capturing relevant extracellular matrices for investigating cell migration.', *F1000Research*, 4. doi: 10.12688/f1000research.6623.1.
- Litvinov, S. V. *et al.* (1994) 'Ep-CAM: A human epithelial antigen is a homophilic cell-cell adhesion molecule', *Journal of Cell Biology*, 125(2), pp. 437–446. doi: 10.1083/jcb.125.2.437.
- Maghzal, N. *et al.* (2010) 'The tumor-associated EpCAM regulates morphogenetic movements through intracellular signaling', *Journal of Cell Biology*, 191(3), pp. 645–659. doi: 10.1083/jcb.201004074.
- Maghzal, N. *et al.* (2013) 'EpCAM Controls Actomyosin Contractility and Cell Adhesion by Direct Inhibition of PKC', *Developmental Cell*, 27(3), pp. 263–277. doi: 10.1016/j.devcel.2013.10.003.

Rosse, C. *et al.* (2010) ‘PKC and the control of localized signal dynamics’, *Nature Reviews Molecular Cell Biology*, 11(2), pp. 103–112. doi: 10.1038/nrm2847.

Soler, A. *et al.* (2018) ‘Autologous cell lines from circulating colon cancer cells captured from sequential liquid biopsies as model to study therapy-driven tumor changes.’, *Scientific reports*, 8(1), p. 15931. doi: 10.1038/s41598-018-34365-z.