EpCAM cellular functions in adhesion and migration, and potential impact on invasion: A critical review

François Fagotto, Azam Aslemarz

Version of Record: https://www.sciencedirect.com/science/article/pii/S0304419X20301554
Manuscript_b4e54929a0cb0e1e7050cf7efc4dddb3

EpCAM cellular functions in adhesion and migration, and potential impact on invasion: A critical review

François Fagotto[1,#] and Azam Aslemarz[1, 2]

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

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

Summary

EpCAM has long been known as a cell surface protein highly expressed in carcinomas. It has since become one of the key cancer biomarkers. Despite its high fame, its actual role in cancer development is still controversial. Beyond a flurry of correlative studies, which point either to a positive or a negative link with tumour progression, there has been surprisingly few studies on the actual cellular mechanisms of EpCAM and on their functional consequences. Clearly, EpCAM plays multiple important roles, in cell proliferation as well as in cell adhesion and migration. The two latter functions, directly relevant for metastasis, are the focus of this review. We attempt here to bring together the available experimental data to build a global coherent view of EpCAM functions. We also include in this overview EpCAM2/Trop2, the close relative of EpCAM. At the core of EpCAM (and EpCAM2/Trop2) function stands the ability to repress contractility of the actomyosin cell cortex. This activity appears to involve direct inhibition by EpCAM of members of the novel PKC family and of a specific downstream PKD-Erk cascade. We will discuss how this activity can result in a variety of adhesive and migratory phenotypes, thus potentially explaining at least part of the apparent inconsistencies between different studies. The picture remains fragmented, and we will highlight some of the conflicting evidence and the many unsolved issues, starting with the controversy around its original description as a cell-cell adhesion molecule.

1. Introduction

EpCAM is a cell surface transmembrane protein that has long been known as a major marker for carcinomas, routinely used in cancer diagnostic. Its mutation is also cause of a severe intestinal disorder, called human congenital tufting enteropathy (CTE), a disease characterized by disruption of the intestinal barrier, chronic diarrhea and inflammation. Originally identified as a cancer-derived antigen, its first detailed characterization indicated a function as a homophilic cell-cell adhesion molecule, even though it showed no sequence nor structural resemblance with any known cell adhesion molecules (CAMs), neither cadherins, nor other CAM families such as the immunoglobulin superfamily CAMs (e.g. N-CAM, ICAM, CEACAMs) or mucin-like CAMs (selectins). It was baptized Epithelial Cell Adhesion Molecule, to convey its expression restricted to epithelia [1, 2]. This adhesive function was rapidly and widely accepted by the cancer community, and for several years, only few studies had attempted to dig further into its molecular and cellular properties. Yet, three unexpected discoveries have led to question the real nature of this protein: Firstly, biochemical work led by Zoller and colleagues, identified an interaction between EpCAM and claudins (cldn), which are integral components of tight junctions [3, 4]. This interaction did not appear to be directly related to an adhesive function, but rather to recruitment to membrane domains with potential impact on cell migration [5, 6]. Secondly, EpCAM appeared to be proteolytically cleaved in cancer cell lines, producing a short cytoplasmic fragment that interacted with β-catenin and activated nuclear transcription, leading

to increased cell proliferation [7]. Finally, work in early fish and Xenopus embryos revealed that EpCAM was a major regulator of morphogenesis [8, 9]. Most surprisingly, this activity was shown to be completely independent of the assumed role as CAM, but was rather due to regulation of actomyosin contractility through intracellular signalling [9, 10]. These discoveries have boosted research on EpCAM and on its close relative EpCAM2/Trop2, and these recent years have seen a flow of information, from structural data to new potential signalling activities with impact on cell proliferation and/or cell adhesion and migration.

Nevertheless, our understanding of EpCAM and EpCAM2/Trop2 (*) function remains quite incomplete, and their actual involvement in promoting (or rather repressing) cancer malignancy is still highly controversial. This seems to be an appropriate time to put together the available data and attempt to answer two major questions:

1) Can we still consider EpCAM and EpCAM2/Trop2 as genuine CAMs, and, if not, how can one explain their well-documented effect on cell adhesion?

2) Can one come up with a coherent framework that would account for the disparate, often apparently contradictory, effects of EpCAM on cell and tissue adhesive and migratory properties?

To approach these goals, it was necessary to take a candid look at the published data, to critically re-evaluate their interpretation, also based on recent progress made in our understanding of the cellular, molecular and biophysical mechanisms controlling adhesion and migration. Another necessary step was to compile, compare and combine information on EpCAM and EpCAM2/Trop2, which, with a few exceptions, have been studied in isolation, despite their obvious high similarity and the likely high degree of redundancy.

** Note: Multiple names were given to these two proteins, including EpCAM, Tacsd1 and 2 (for tumour-associated calcium signal transducer), Trop1 and 2 (for Trophoblast cell surface antigen). In the human cancer field, EpCAM and Trop2 are the most frequently used names. However, considering that Trop2 results from duplication of the original EpCAM, and that this name is widely accepted for non-mammalian vertebrates (also in some species it is annotated as Trop or Trop-like), the names of EpCAM1 and EpCAM2 would better reflect their common origin and their high similarity. In this review, we adopt a hybrid nomenclature, using EpCAM for the former, and EpCAM2/Trop2 for the latter.*

2. General principles of cell adhesion and migration and the role of actomyosin contractility

Before discussing the role of EpCAM, it is useful to summarize some of the key principles underlying cell adhesion and cell migration.

One of the essential properties of animal cells is to be capable to adhere to neighbouring cells as well as to the extracellular matrix. Although in principle binding of any molecule exposed at the cell surface with molecules from other cells or of the matrix may produce substantial adhesive force, under physiological conditions adhesion is mediated by highly specialized CAMs. A large number of CAMs have been identified, although cadherins are by far the major contributors to cell-cell adhesion, while integrins take care of most of the adhesion with the matrix. The first general property of all CAMs is to interact with a specific partner at the surface of the adjacent cell (typically homophilic binding between cadherins) or with a component of the matrix (Fig.1). However, CAMs have additional essential properties, which are absolutely necessary to produce efficient adhesion, in particular their relationship with the actin cytoskeleton building the cell cortex. We will use cell-cell adhesion to describe these properties (Fig.1A), but the same principles are directly applicable to cell-matrix adhesion (Fig.1B).

Figure 1

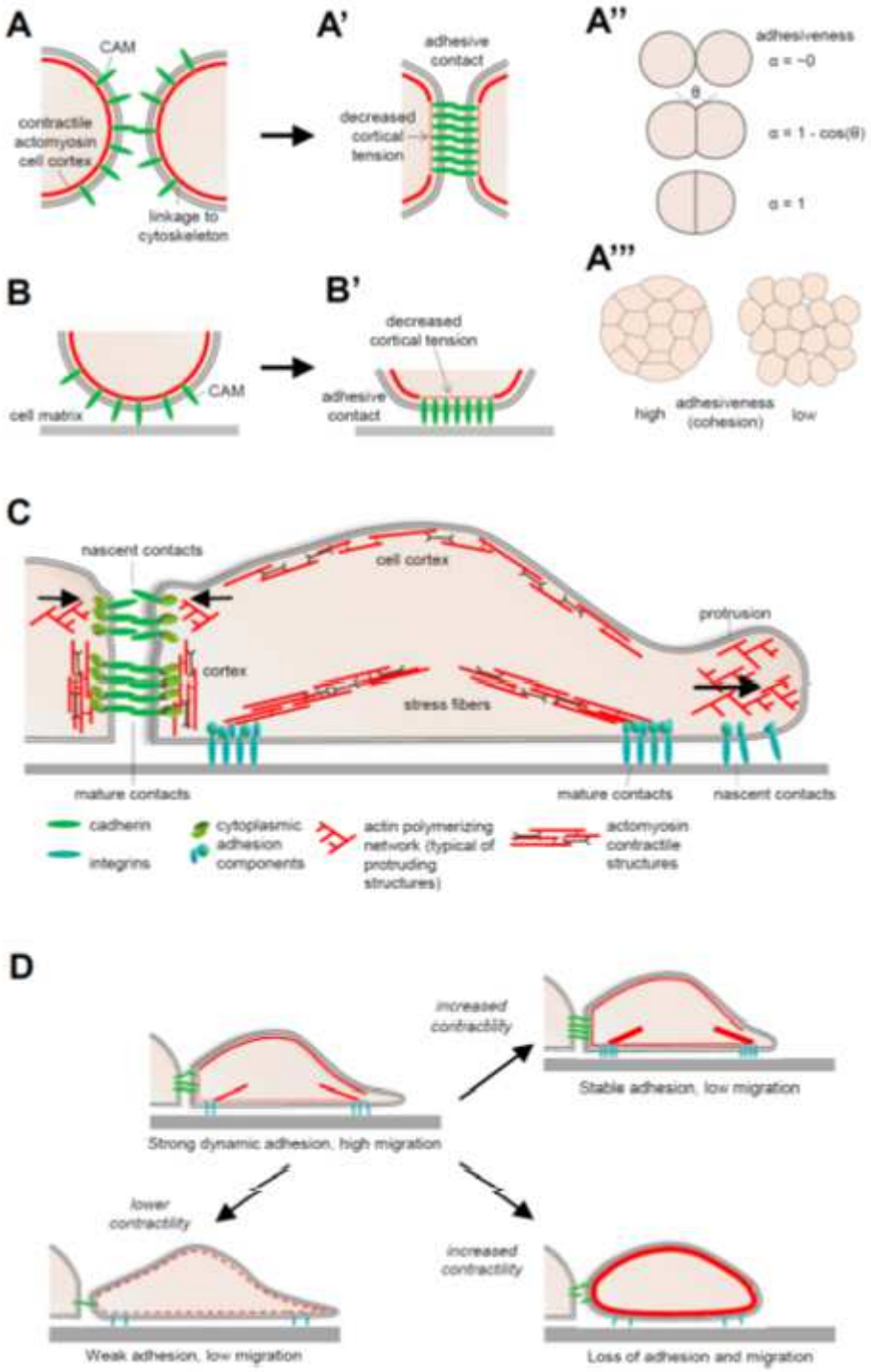

Figure 1. Principles of cell adhesion and migration and role of actomyosin contractility.

(A,B) The basic requirements for cell-cell adhesion. (A) Cell-cell adhesion. Cell adhesion molecules (CAMs) form interactions between two cell surfaces, which are called "trans-interactions". They are most often homophilic. However, these interactions are not sufficient to establish an adhesive contact. CAMs must also be anchored to the cytoskeleton in order to resist strain. Furthermore, the contractility of the actomyosin cortex is normally too strong to allow the contact to spread (A). This implies that a mechanism must exist that represses cortical tension along the contact (A'). This repression is effectively achieved by classical cadherin-based adhesive structures. The same principles are valid for integrin-based cell-matrix adhesion. (A'') Cell geometry directly reflects the underlying physical properties. When considering an ideal cell doublet in suspension, the balance between cell adhesion and cortical tension dictates the spread of the contact. The degree of "adhesiveness" is defined by the angle θ formed by the two membranes. Adhesiveness can vary from 0 (no adhesion) to 1 (maximal). (A''') The

same geometrical approach applies to groups of cells, and can be used to evaluate the actual efficiency of adhesion in cell aggregates. B) Cell matrix-adhesion. The same principles apply for spreading of a cell on a matrix substrate. (C) CAMs associate with different types of actin cytoskeleton. In regions where new contacts are being established, the stiff and contractile cortex is disassembled and replaced by a polymerizing cytoskeleton typical of expanding cellular protrusions (arrows). As adhesive contacts mature, the environing cytoskeletal is modified such that CAMs are linked to actomyosin fibres (stress fibres for integrin-based cell-matrix adhesions). (D) Changes in actomyosin contractility have diverse effects on adhesion and migration. In this example, the cell in the centre has contractility set to yield strong yet dynamic adhesion, which permits remodelling and thus effective migration. Note that migration can occur on a matrix substrate, as represented here, or relative to neighbouring cells, a process called intercellular migration. Strong repression of contractility may hamper coupling of CAMs with the cytoskeleton, compromising adhesion. If adhesion is too weak, migration may fail as well. Two possible outcomes are showed as result of increased contractility: Adhesion may be strengthened at the expense of migration speed. However, under conditions of excessive contractility, adhesion fails, cells adopting a typical round morphology.

If one considers a single cell in complete isolation, one notices that it adopts by default a spherical shape, as a result of the hydrostatic pressure of the cytoplasm, balanced by the resistance of the cell cortex (Fig.1A). The cortex is made of a dense network of heavily cross-linked actin filaments. The network is contractile, due to the activity of non-muscle myosin II, and firmly tethered to the plasma membrane. Its role is not only to resist internal cell pressure, but also to provide animal cells with a protective coat far sturdier than the bare plasma membrane. Note that despite its stiffness, the cortex is extremely dynamic, with a high turnover of its components in the sub-minute range, which allows it to be rapidly remodelled or even completely dismantled [11].

The stiffness and contractility of the cortex represents a strong barrier to adhesion (Fig.1A). Expression of cell surface "sticky molecules" alone is not sufficient to generate adhesion, because the cells would only be attached by a tiny contact. For the contact to flatten and expand, cortical contractility must be severely reduced (Fig.1A'). To take a concrete analogy, the strongest glue will do a lousy job to stick two fully inflated spherical footballs together, while this becomes an easy game once the balls are partially deflated, thus deformable. For cell-matrix adhesion (Fig.1B). The analogy would then be to glue an inflated or a deflated football on a flat surface (Fig.1B'). These considerations lead to a first central concept: Interactions between cell adhesion molecules (CAMs) (or between integrins and the matrix) are on their own unable to force the expansion of a contact against the strong resistance of the cortex. The second prerequisite to adhesion is the softening of the cortex along the nascent contact (depicted as thinner dashed cortical layer in Fig.1A'). Classical cadherins have the property to recruit multiple regulators of the actin cytoskeleton to efficiently repress actomyosin contractility. Upon initial contacts, cadherins induce a progressive "melting" of the surrounding cortex, allowing more cadherins to join and the contact to expand. Eventually, the system equilibrates when the tensions at the free cell cortex and at the contact are balanced [12-14]. The extent to which tension is reduced at the contact surface relative to the free cell surface will dictate the maximal width of the contact area at equilibrium. A direct and astonishingly simple consequence of this principle is that if one considers a contact of a cell doublet, the angle θ at the vertex (Fig.1A") yield a direct readout of the "adhesiveness" of the system [15].

Once established, the adhesive contact must be able to hold to a steady structure to effectively resist stress. The third central concept in adhesion is the absolute need for CAMs to be anchored to the cytoskeleton (Fig.1A-C) [16]. The link between adhesion molecules and the actin cytoskeleton is not direct, but involves a cytoplasmic complex of adaptor proteins. This is accomplished by α-catenin and β-catenin in cadherin-based adhesion complexes. Other adaptors, such as talin and paxillin, fulfil the

same function for integrin-based cell-matrix adhesions. At least in culture cells, the actin structures that connect to integrin-based focal adhesions form distinct cables, called "stress fibres" (Fig.1C). The actin fibres that hold cadherin adhesions are less conspicuous, because they run parallel to the cell surface, intertwined and probably continuous with the rest of the cell cortex. One should keep in mind that these adhesive structures are far from passive, but on the contrary highly dynamic and mechanosensitive, i.e. capable of responding to physical strain by complex reactions, including changes in the CAM-CAM bonds, increased CAM clustering, reinforcement of the CAM-cytoskeleton linkage, and remodelling of the cytoskeletal structures.

In summary, a CAM is defined as follows: It is a cell surface transmembrane protein that forms "trans-interactions" (1) with another protein on the surface of a neighbouring cell (called homophilic interaction if the partner is the same protein), remodels the cortical environment (2), and connects to the cytoskeleton in order to transmit forces (3).

One can certainly experimentally force non-adherent cells to "aggregate", for instance by expressing the receptor Notch and its membrane bound ligand Delta [17], yet what one gets are "grapes" of round cells that have little to do with real tissues. This comparison illustrates the clear difference between *bona fide* CAMs and other types of cell-cell interactions.

As for the cytoskeleton, actomyosin contractility has two seemingly antagonist roles in cell adhesion: By creating strong tension along the cortex, it resists cell spreading, and must be repressed to allow expansion of the contact with a neighbouring cell or with the matrix substrate (Fig.1C). Yet actomyosin structures are also required to anchor adhesive junctions and resist strain.

The two functions can be viewed, at least in coarse approximation, as acting sequentially during formation of a contact, first antagonistic to nascent adhesions, then positively reinforcing maturating contacts (Fig.1C). Nevertheless, cortical tension clearly continues to influence mature contacts, and contributes to the balance of forces exerted at adhesive sites. The schemes of Fig.1D present a few scenarios resulting from modifying myosin activity: If, for instance, contractility is too low, the contacts may fail to resist tension, and adhesion may be compromised. Increased contractility can lead to different situations: It may stimulate mechanosensing and strengthen adhesion, or, on the contrary, destabilize adhesion due to high cortical tension. Beyond these simple examples, one can easily imagine the large diversity of cell behaviour controlled by multiple parameters, from global basal contractility, to local organization of the cytoskeleton, or regulation of the cytoplasmic adaptors. These considerations are key when interpreting the effect of myosin regulators. We will discuss below this issue in the context of EpCAM function.

Another important principle to consider is the interplay between adhesion and migration (Fig.1D). An absolute requirement for migration is a grip to the substrate, whether the extracellular matrix or adjacent cells. Up to a level, adhesion favours migration as it allows cells to "pull" on the support. However, adhesive bonds must be dynamically undone/redone to allow the cell move further. Strong adhesion may slow down or even completely inhibit migration. Therefore, active migration requires the right level of adhesion. Again, we will come back to this important consideration when discussing the functions of EpCAM.

3. Introducing the EpCAM molecule

EpCAM is a cell surface single pass transmembrane protein typically expressed in embryonic cells and in adult epithelial tissues, where it is restricted to the basolateral domain (Fig.2A).

3.1. Structure

EpCAM is vertebrate-specific and does not show any similarity with other cell surface proteins. In fish and amphibians, EpCAM forms a single gene family on its own (duplicated gene in teleost lineage, including zebrafish and in the alloploid Xenopus laevis). In amniotes (reptiles, birds and mammals), a second gene, named Trop2 or TACSTD2, has appeared through retro-transposition [18]. EpCAM and EpCAM2/Trop2 are extremely similar, sharing 48% identity and 78% similarity in human, which is comparable to the divergence of EpCAM itself between reptiles and mammals (51% identity and 78% similarity). As discussed below, EpCAM and EpCAM2/Trop2 share all the major distinctive features and probably most cellular functions. However, that their promoters are completely unrelated as a result of retro-transposition [18]. The amino acid length of the EpCAM protein is slightly variable between species, we have here chosen human EpCAM (hEpCAM) to discuss the details of its structure. Unless mentioned otherwise, all the features described are evolutionarily conserved throughout vertebrates. The few features that are distinct between EpCAM and EpCAM2/Trop2 will be mentioned.

After removal of its signal sequence, hEpCAM is a polypeptide of 293 amino acids, with a predicted molecular weight of 35 kDa. EpCAM is heterogeneously glycosylated, migrating on SDS-PAGE with an apparent size around 40 kDa. Its extracellular domain (EpEx) is composed of three distinct domains, an N-terminal domain with a unique cysteine disulphide bonding pattern, followed by a second cysteine-rich motif with disulphide bonding pattern related to a thyroglobulin-type A1 domain, and a C-terminal domain without similarity to any other animal protein [19]. The three domains are arranged in a triangular fashion where each domain contacts the other two, resulting in a small compact extracellular domain [19].

The TM domain is rich in valine and poor in leucine, while the latter is generally the most abundant residue found in TM domains [20]. Furthermore, its sequence is evolutionary highly conserved in both EpCAM and EpCAM2/Trop2, and throughout vertebrates (unpublished), consistent with its proposed function for interactions with other membrane components (see below).

The cytoplasmic C-terminal tail is exceptionally short, from 18 to 29 amino acids in fish and birds, respectively, and 26 in hEpCAM. It can be divided in two portions, a highly conserved juxtamembrane segment rich in basic residues, and a distal segment that is more variable among vertebrate families, but includes three to four glutamic residues positioned at conserved intervals. The central part of the cytoplasmic tail of EpCAM2/Trop2 (and presumably of EpCAM), encompassing part of the juxtamembrane and part of the distal sequence, can adopt an α-helical conformation in solution [21], but the relative positions of the glutamic acids do not show any particular alignment in this structure. The mention of the existence of a C-terminal PDZ binding motif has been perpetuated in multiple reviews, yet by searching among an extensive list of validated PDZ motifs [22], we did not find any support for this claim.

Figure 2

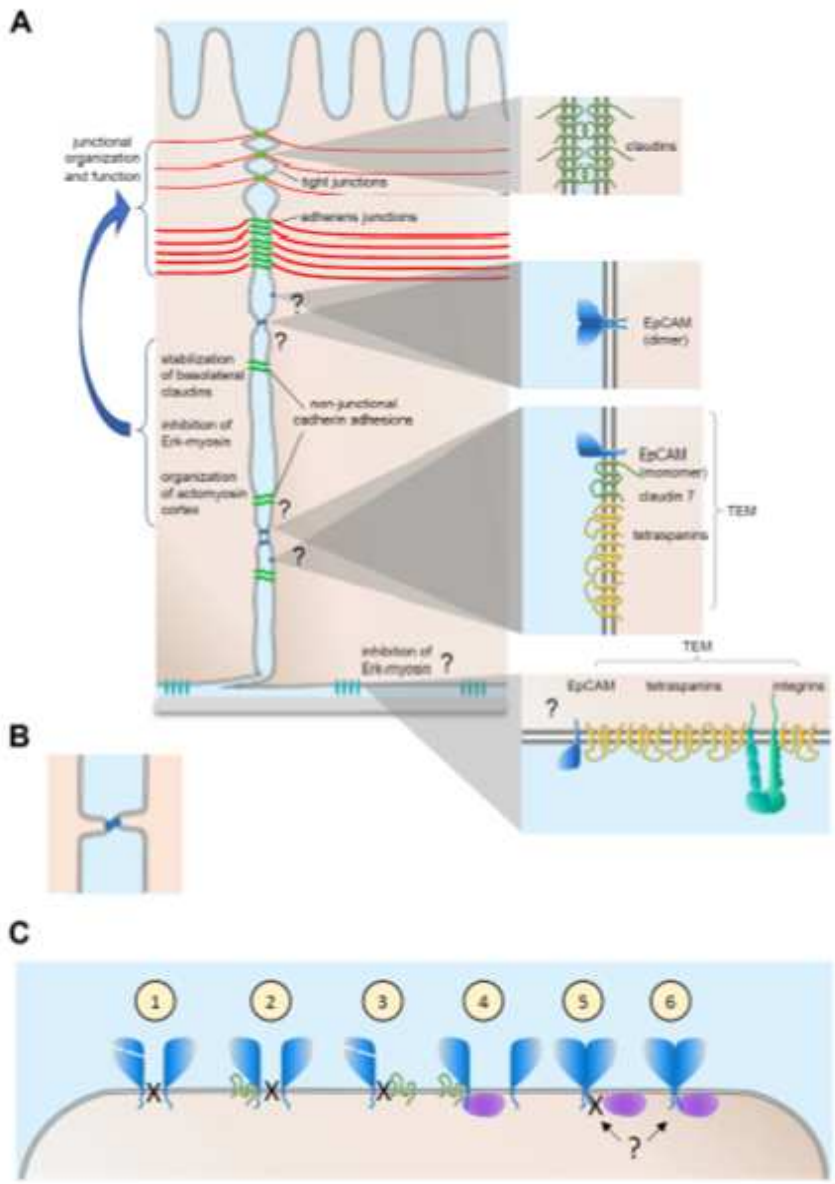

Figure 2. EpCAM as regulator of epithelial organization.

(A) General distribution of EpCAM and associated molecules in epithelial cells. The inserts represent the various proteins at scale relative to the thickness of the lipid bilayer. EpCAM localises along the lateral domain. EpCAM-cldn7 interaction appears to be responsible for the lateral distribution of a non-junctional pool of cldns. EpCAM and cldns are associated with tetraspanin membrane microdomains (TEMs). EpCAM does not appear to be associated with apical junctions (adherens junctions and tight junctions), but the detailed information about its subcellular localization is not known. Thus, the distribution of various EpCAM forms, i.e. cis-dimers and monomers associated with cldn7 and/or with TEMs, and the relative position relative to non-junctional lateral cadherin cell-cell contacts remain open questions. However, EpCAM (perhaps in association with cldns) plays an important role in regulating the actomyosin cortex through repression of the nPKC-Erk-myosin pathway. Loss of EpCAM causes a disorganization of the actomyosin network, as well as an imbalance in cldn distribution, affecting apical junctional integrity (blue arrow). EpCAM has been reported to interact with integrins, which is likely to occur within TEMs. This could account for its effect on cell-matrix adhesion and migration. However, whether EpCAM does localize to the basal side of epithelia is not known. (B) Scheme of EpCAM positive small protrusions observed at contacts between fibroblasts (L cells). (C) Relationship between dimerization, cleavage, cldn interaction, and PKC inhibition. EpCAM is at equilibrium between a monomeric form and a relatively

stable cis-homodimer. (1) The EpCAM monomer is sensitive to matriptase proteolysis. The resulting N-terminal fragment remains connected to the rest of the extracellular domain through a disulphide bridge. The cleaved form cannot form dimers. (2) EpCAM monomer can interact directly with cldn7. This interaction is incompatible with cis-homodimerization. (3) Matriptase cleavage blocks interaction with cldn7. (4) Cldn7 interaction might be required for inhibition of the nPKC-Erk pathway. (5) and (6) Whether EpCAM homodimerization is compatible with nPKC binding remains unknown.

3.2. Phosphorylation

The cytoplasmic tail of EpCAM and EpCAM2/Trop2 contain one highly conserved tyrosine (Tyr296 and 306, respectively). EpCAM2/Trop2 has also two serines, one before the tyrosine (Ser303) and one close to the tip (Ser322). Basu et al [23] did not detect any tyrosine phosphorylation, but rather serine phosphorylation, which they attributed to Ser303. A recent study on EpCAM2/Trop2 reported phosphorylation at Ser322, but not at Ser303 [24]. Both studies invoked the action of PKC [23, 24]. Consistently, phosphoproteomic data (PhosphositePlus) show little trace of tyrosine phosphorylation, nor Ser303, while Ser322 was detected in a few screens. Sequence comparison reveals limited conservation of these residues. The hypothetical significance of these modifications will be discussed below.

3.3. Glycosylation

The human sequence contains three N-linked glycosylation consensus sites, Asn74, Asn111 and Asn198 residues. The first site is only found in human EpCAM, while the third one is found in most vertebrates except bony fish and avian. The middle site is the only one conserved throughout vertebrates, suggesting an important function. However, it does not seem to be required for domain folding nor expression at the cell surface, while mutation of the Asn198 glycosylation site decreased stability of exogenous EpCAM in HEK293 cells [25]. Additional O-glycosylations have been detected in proteomic screens (PhosphositePlus), but again in non-conserved sites. EpCAM was reported to be more glycosylated in some cancers compared to wild type cells [26]. Clearly more detailed information is required about EpCAM glycosylation and its functional significance.

3.4. Dimerization

Litvinov team reported early on that EpCAM formed cis-dimers, i.e. lateral interactions between two EpCAM molecules expressed on the same membrane [27] (Fig.2). This observation was since confirmed and structural investigations of the EpEx defined the interacting interfaces [19, 28]. The resulting EpEx dimer has a compact heart-shape configuration. The two TM are predicted to be positioned close to each other [19]. Molecular dynamics have been used to propose a contribution of the TM α-helices to dimerization [19]. However, a different TM dimerization motif was proposed for EpCAM2/Trop2 [21], which is surprising, considering that most TM residues are completely conserved in all vertebrate EpCAM and EpCAM2/Trop2 proteins. Also, it is still today rather a random surmise to predict dimerization based on TM hydrophobic residues [29] (see also below). In any case, the EpEx dimer interaction appears quite strong, suggesting that EpCAM should be mostly dimeric, with only a small fraction in a monomeric form. As we will see, other factors can affect dimerization.

3.4. Proteolytic cleavage

EpCAM is a target of multiple proteolytic cleavages. As discussed below, EpCAM undergoes regulated intracellular proteolysis (RIP), involving first shedding of EpEx, followed by cleavage within the TM leading to the release of the cytoplasmic tail [7]. Multiple RIP products have been detected and analysed in detail [30-32]. RIP was also demonstrated for EpCAM2/Trop2 [33]. Note that EpCAM RIP is

thought to be mostly active in specific cancers/cancer lines that secrete high levels of ADAM metalloproteases [7, 34, 35]. One must then keep in mind that RIP is at most marginal in normal cells, while in cancer, its importance will be strongly cancer type and context-dependent [34, 36]. Independently of RIP, both EpCAM and EpCAM2/Trop2 are also cleaved by an extracellular protease called matriptase at a specific dibasic site [37, 38]. The resulting small 8kDa N-terminal fragment is not released, but remains bound to the rest of the molecule via a disulphide bond [37, 38]. Cleavage by matriptase is probably a rather general reaction, since matriptase is widely expressed and the cleavage site is perfectly conserved in EpCAM and EpCAM2/Trop2, through vertebrates. Interestingly, this cleavage is incompatible with EpCAM cis-dimerization [19] (Fig.2C). As a consequence, dimerization may protect EpCAM from proteolysis. Reciprocally, cleavage of the monomeric pool should irreversibly prevent dimerization, and potentially displace the monomer-dimer balance, favouring the former. Finally, matriptase-cleaved EpCAM (and EpCAM2/Trop2) is preferentially sent for lysosomal degradation [37, 38]. One important piece of information still missing is the actual extent of EpCAM cleavage, and the resulting turnover rate of EpCAM and EpCAM2/Trop2. In the case of RIP, cleavage is slow, while the subsequent degradation of the cleaved fragments is fast [36]. Available information about matriptase cleavage is based on long term LOF/GOF of one of the players and/or inhibition of lysosomal function. While these treatments lead to strong phenotypes, under wild type steady state conditions, only a small portion of EpCAM is found in a cleaved form [37, 38], also consistent with a low monomeric to dimeric ratio [27, 28]. What must be now established is the cause of this distribution, which could either be due to high stability of the dimeric form (therefore resistant to cleavage), or, on the contrary, to rapid degradation of the cleaved form, in which case, the apparent low abundance of cleaved/monomeric EpCAM may be misleading as to the actual rate of cleavage and subsequent degradation.

3.5. Interactions with TEMs and claudins.

Work by the Zöller group showed that EpCAM was associated with the so-called tetraspanin-enriched domains or TEMs [39] (Fig.2A,C). These are special microdomains of the plasma membrane, which can be isolated based on their resistance to extraction by mild detergents and floatation on a density gradient [40, 41]. TEMs are clearly distinct from "lipid rafts". Their assembly is not based on lipids, but on the property of tetraspanins to arrange in dynamic networks, which in turn recruit various other proteins and lipids. Tetraspanins form a large family and are expressed in multiple combinations in different cell types and tissues. TEMs are suspected to have multiple functions, though they are not yet fully understood. They clearly impact on trafficking and on protein diffusion within the membrane. Integrins are among the best-known and widespread tetraspanin interactors [40, 41]. Note that EpCAM seems to be a peripheral component of TEMs, at least based on biochemical criteria [5].

EpCAM recruitment to TEMs requires its interaction with claudin-7 (cldn7) [3, 6] (Fig.2A,C). Claudins are core transmembrane proteins of the tight junctions, which, together with occludin, insure the sealing properties of these junctions [42]. Cells express multiple claudins, which copolymerize to form heteromeric junctional stands [42]. Cldn7 was spotted because its high expression, together with EpCAM and the tetraspanin CD44, correlated with poor cancer prognosis [6]. The EpCAM-Cldn7 interaction seems to be rather stable, as it can be robustly detected by regular immunoprecipitation [3, 6, 43]. A second claudin, claudin-1, was also found as part of the complex, probably recruited by cldn7 [43]. Other tested claudins were not detected in this complex [43]. The interaction with cldn7 is essential for EpCAM association with TEMs and the dependency is reciprocal [3, 6]. Through the use of chimeric EpCAM fusion proteins, it was shown that the cytoplasmic tail and most of the extracellular domain were dispensable for binding cldn7. It was thus inferred that the interaction was mediated by the transmembrane domain. Consistently, a double point mutation in an AxxxG motif of

the hydrophobic helix, which we will call here the TM mutant, disrupted the interaction [6]. However, as already mentioned for the EpCAM dimer, whether these hydrophobic residues would be actually directly responsible for this protein-protein interaction is unclear [29]. A thorough re-investigation is needed, which should also consider a potential contribution of extracellular interactions, since a proximal segment of EpCAM was still present in the chimera constructs used to define the minimal interacting region [6]. In fact, residues immediately preceding the TM domain may be predicted to fold into an extension of the TM α-helix that would form an amphipathic helix, clearly better adapted to produce a stable cis interaction (A. Kajava and F. Fagotto unpublished observation). These detailed considerations are of some importance, because subsequent studies assumed a purely transmembrane interaction and used the EpCAM TM mutant as tool to discriminate for EpCAM-cldn7 association. In a more conservative interpretation of the available data, the TM mutant may still be used to test if a given property of EpCAM may be independent of cldn7 association, but not to extrapolate a direct role of this interaction. One should also keep in mind that some of the effects may relate on the partition of these two proteins within TEMs. The sequence requirements on the cldn7 side have not been investigated.

From a broader cellular perspective, the association of EpCAM with core proteins of the tight junctions may seem rather odd, since EpCAM is clearly not localized at these junctions. It turns out that, in addition of this primary site, claudins, and in particular cldn7, also localize along the non-junctional basolateral membrane [44]. It is this second pool that is thought to interact with EpCAM. Interestingly EpCAM and cldn7 depend on each other not only for incorporation into TEMs, but also for their stabilization at the basolateral membrane, as the experimental depletion of one partner leads to the downregulation of the other [43, 45]. One may speculate that their interdependent stability is related to their partition into TEMs. Whether it requires direct interaction is not clear. In one report, cldn7 levels at the membrane could be rescued by the EpCAM TM mutant [46], while in another study the mutant was much less efficient than wild type EpCAM [43]. What then about the tight junctions? The impact on these structures is an unsettled issue: EpCAM depletion was reported to cause a decrease in cldn7 primarily from the basolateral membrane, while the tight junctional pool was either unaffected [46] or even increased, correlating with increased transepithelial resistance, a readout for tighter junctional seal [43]. This led to the proposal that EpCAM regulates tight junction function by competing for cldn7. However, in EpCAM KO mice, cldn7 was undetectable [47]. The most reasonable explanation to these dissimilar results is that the distribution of cldn7 should be seen as a steady-state that depends on several parameters, including relative levels of EpCAM and cldn7, stability and turnover of the EpCAM-cldn basolateral pool (predicted to depend itself on TEM properties), as well as tight junction composition and turnover (predicted to influence cldn7 retention). Furthermore, we will see that EpCAM controls myosin activity via cell signalling, which is another indirect route through which EpCAM levels and localization can impact on cell junctions and epithelial organization (Fig.2A). Variations between cell types and experimental conditions, including degree and duration of depletions, are likely to explain the apparent discrepancies between various studies. Cleavage by matriptase, which inhibits both EpCAM cis-dimerization and EpCAM-cldn interaction [37, 38], is one extra ingredient that is bound to influence the system by strongly increasing turnover. While the influence of EpCAM on tight junction function remains unclear, we will see below strong hints for a role of cldn7 in EpCAM signalling activity. Consistent with the high similarity between EpCAM and EpCAM2/Trop2, the latter was also found to interact with cldn7 and cldn1, and to regulate their levels and distribution [43, 48].

3.6. Other interactions

There have been reports of EpCAM2/Trop2 physically interacting with the fibronectin receptor integrin α5β1, and that it is able to re-localize integrins from focal adhesions to the leading edge of prostate cancer migrating cells [49, 50]. The interaction appears specific to integrin α5β1 compared to α3β1 and αVβ3, but has not yet been further characterized. EpCAM also seems to interact with integrin β1 [51]. Note that these reports omitted to refer to a previous study from the Zoller lab, which showed that EpCAM could be found together with α3β1, but not α2β1, within TEMs [5]. Further work is needed to better characterize the nature and properties of the EpCAM-integrin association.

Another interaction often cited is that of EpCAM2/Trop2 with phosphoinositol-2-phosphate. However, this again corresponds to an unsupported claim only based on a weak sequence similarity [52].

4. Is EpCAM a cell adhesion molecule?

4.1. The original characterization

The first molecular and cellular study on EpCAM, authored by Litvinov and colleagues, showed that expression of EpCAM in mouse L fibroblasts (or L cells) induced cell aggregation in a calcium-independent manner [2]. L cells lack cadherins, and are a classical model used to characterize ectopically expressed adhesion molecules. EpCAM expression was sufficient to produce homotypic aggregates of L cells, although adhesion appeared significantly weaker than that produced by classical cadherin, as it failed to induce compaction of cell aggregates (Fig.1A'") [2]. In mixing experiments, wild type L cells were largely excluded from aggregates of EpCAM positive cells. Litvinov and colleagues went on to demonstrate that this adhesive activity required both the EpEx and the cytoplasmic tail [27, 53]. Importantly, EpEx was found to oligomerize laterally: A significant fraction of EpCAM was indeed present at the cell surface as cis-dimers and cis-tetramers [27]. Thanks to the characterization of antibodies that would recognize only monomeric, or monomeric and multimeric EpCAM, the authors could infer that EpCAM exposed at the cell membrane was largely multimeric [27]. Note that they did not find any evidence for trans-oligomerization. Nevertheless, their ultrastructural study on EpCAM localization yielded important observations [27]. Immunogold labelling was performed in both EpCAM-expressing L cells and colon cells expressing EpCAM endogenously. EpCAM distribution was discontinuous along the lateral membranes (Fig.2A). Unlike cadherins, EpCAM did not seem to concentrate in "clusters". Rather, anti-EpCAM-coupled gold particles were systematically found as isolated doublets positioned face to face across the contact. This remarkable pattern suggested that EpCAM localized to tiny contacts made by two identical structures, both containing EpCAM [27]. In epithelial cells, EpCAM was detected at closely apposed membranes (Fig.2A), while in L cells expressing ectopic EpCAM, they were observed at the tip of tiny protrusions (Fig.2B) [27]. One should highlight that the authors quite conservatively made no claim for homophilic binding. As further discussed below, an alternative model could be that these structures consist of two antiparallel heterotypic contacts (Fig.3D). Another observation was that EpCAM was detected exclusively on lateral membranes of epithelial cells, not at the basal membrane [27, 43, 54, 55]. We shall keep this in mind when discussing the potential role of EpCAM in cell-matrix adhesion and migration. Note that EpCAM, similar to cadherins, is distributed along the whole basolateral membrane in early embryonic epithelia [8-10]. These early epithelia are not yet fully differentiated and still lack a basal membrane.

Figure 3.

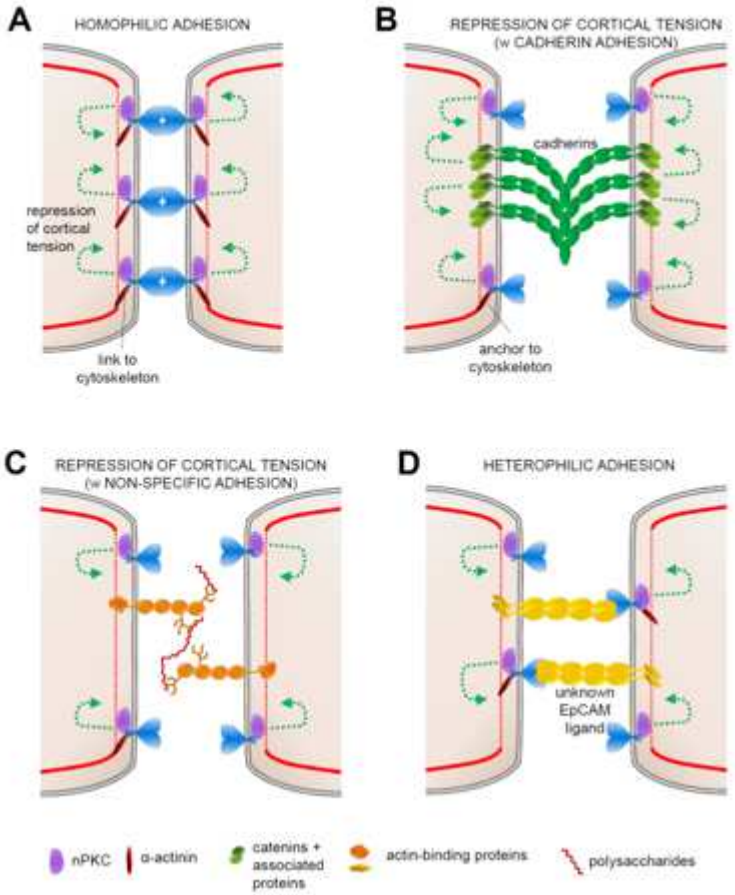

Figure 3. Hypothetical models for the pro-adhesive activity of EpCAM.

A body of evidence has involved EpCAM in cell-cell adhesion, but the mechanisms are not fully elucidated. (A) In the initial model, EpCAM was proposed to act as homophilic adhesion molecule. It would be linked to the actin cytoskeleton through direct interaction with α-actinin. For this model to be viable, a mechanism should exist to decrease cortical tension at the contact. This could be achieved by EpCAM itself through nPKC inhibition. (B) EpCAM inhibition of the nPKC-myosin pathway is an established mechanism that stimulates cadherin-based adhesion. (C) The same mechanism could be used to decrease tension sufficiently to favour cell-cell adhesion even in the absence of cadherins, via non-specific interactions between glycoproteins and extracellular components such as hyaluronan [60]. D) EpCAM could be a heterotypic cell adhesion molecule with an unknown partner. The diagrams conserve the relative sizes of the extracellular domains and distance between membrane, but the thickness of the membrane is reduced and the cortex symbolized with a single line for the sake of clarity.

Another key property of CAMs is the ability to interact with the actin cytoskeleton. Biochemical characterization of such interactions under physiological conditions is far from trivial. One classical criterion is detergent insolubility, based on the well-known insolubility of the cytoskeleton upon extraction with mild detergents. Note, however, that there are alternative causes for insolubility, including association with special lipid membrane, or, trivially, protein aggregation. At any rate, Litvinov and colleagues found that 5-20% of total EpCAM was detergent-insoluble [54]. This insoluble fraction was localized at the cell surface, and was partially decreased upon actin depolymerization by cytochalasin D treatment, supporting the proposed interaction with the cytoskeleton. Furthermore, insolubility required EpCAM cytoplasmic tail. This short cytoplasmic segment was shown to bind directly to the actin crosslinker α-actinin, providing a molecular base for EpCAM anchoring to the cytoskeleton [54]. Additional support included accumulation of α-actinin at cell contacts between EpCAM-expressing L cells [54]. Also, experiments testing the effect of partial truncations of EpEx

showed that correlation between cis-dimerization, detergent insolubility, α-actinin interaction, and the potential to promote cell adhesion [27].

The last prerequisite for adhesion is a local decrease in cortical tension. As mentioned above, the geometry of cell contacts provides a direct read-out of adhesiveness. EpCAM-expressing L cells remained quite round, with rather small cell-cell contacts, indicative of weak adhesiveness [2, 27, 54]. Yet, as we will see later, EpCAM has the capacity to downregulate actomyosin contractility [9, 10], which in principle could contribute to adhesiveness.

4.2. Conflicting evidence

Since these initial reports, the cancer community considered that EpCAM was an established homophilic CAM. However, there were from the very beginning observations that were not quite consistent with this assumption. The first and perhaps most striking case was reported by Litvinov et al [56], who found that expression of EpCAM inhibited cadherin-mediated adhesion, not only in the artificial L cell system, but also in epithelial cells. The authors interpreted these results as reflecting a competition between two parallel adhesive systems, a hypothesis that remains plausible. However, the original suggestion that the competition would involve titration of α-catenin does not hold, since EpCAM does not bind this protein [27]. There isn't currently any obvious molecule for which cadherins and EpCAM could compete for. Other studies found that EpCAM deletion/depletion had no visible effect on cell-cell adhesion for a variety of cell types (mouse embryonic stem cells and of various cancer cells) [35, 57, 58].

We obtained the first strong hint for a proadhesive activity of EpCAM that would be independent of a CAM function while investigating its role in early development of the amphibian Xenopus embryo [9, 10]. Among other observations present below, we found that cadherin levels were strongly dependent on EpCAM. A similar observation had been made in the zebrafish embryo, although this effect had still been interpreted based on a putative adhesive function of EpCAM [8]. However, we could show that EpCAM extracellular domain was fully dispensable, unambiguously pointing to the existence of an intracellular signal [9]. We succeeded at characterizing the mechanism, which involves direct inhibition of a PKC-Erk-MLCK-myosin pathway (see below) [9, 10]. Since this discovery, other reports have confirmed the importance of this inhibitory activity in multiple contexts [46, 55, 59].

What then about homophilic binding? There are currently strong doubts that this interaction exists. A recent study [28] has attempted to detect homophilic interaction using a variety of approaches, both with soluble recombinant EpEx *in vitro* and with full length EpCAM in cells. Techniques went from small angle X-ray scattering (SAXS), or chemical cross-linking and analysis of the crosslinked multimers by mass spectroscopy (MS), to bead aggregation, as well as Förster resonance energy transfer (FLIM-FRET). While these experiments confirmed the lateral cis-interaction and refined its characterization, there were no evidence for trans-interactions. One has always to be cautious when drawing a final conclusion on negative results, as there are always potential caveats. It is worth listing them here: In vitro experiments (SAXS and crosslinking) using recombinant proteins may be missing factors that would stabilize the interaction, such as orientation at the membrane or local crowding. This caveat was addressed by the authors, for instance by forcing parallel cis-dimerization and/or immobilizing EpEx on beads, but still no trans-interaction was detected. As for MS results, they yielded a relatively small number of cis-interactions (which is supposedly strong), and it is plausible that a weaker trans-interaction could be missed altogether. Finally, the FRET experiments may suffer from potential steric hindrance by the fluorescent proteins, a possibility that cannot be discarded based on the fact that FRET was detected for cis-interactions. Despite these caveats, all data pooled together raise very strong doubts about the possibility of homotypic binding [28]. Two additional consideration further supports

this negative conclusion: Firstly, even assuming the existence of EpCAM homophilic binding, this would be at best a weak interaction, and the only way then to build effective adhesion would be by concentrating EpCAM in dense clusters. Yet, the immunogold study of Balzar et al [27] argued on the contrary that EpCAM positive "contacts" were small and sparse. Another simple argument is suggested based on EpCAM dimensions: According to the structural data, EpEx sticks out of the membrane roughly 5nm [19, 21]. Accordingly, a trans-interaction would require the two membrane to be brought as close as 10nm. This is an awfully short distance. For comparison, in cadherin adhesions, the membranes are about 25-30nm apart (Fig.3B), 6-7nm in tight junctions, and the lipid bilayer itself is roughly 7nm thick. It is highly unlikely that weak EpCAM interactions could, on their own, pull the membranes that close. Electron microscopy images of EpCAM-expressing L-cells [27] showed EpCAM-positive thin digitations that do not seem compatible with a role in physical adhesion, but may rather correspond to "sensing" contacts, consistent with the EpCAM signalling functions described below.

4.3. Alternative models

In summary, currently available data tend to argue against a role as a homophilic CAM. There are other potential explanations for EpCAM activity in promoting cell adhesion (Fig.3). One possibility would be that EpCAM is a heterophilic CAM, whose partner would still be unknown (Fig.3D). This model would imply, however, an EpCAM partner widely expressed, including in fibroblasts (L cells). Another possibility would be that EpCAM has no actual CAM function, but has an indirect effect on adhesion through its ability to reduce cortical tension by inhibiting the PKC-Erk pathway (Fig.3B). This activity clearly accounts for the adhesive phenotypes observed in the Xenopus embryo [9, 10] and the intestinal epithelium [55]. Importantly, the same activity could in principle account for the aggregation of L cells, even though these cells seem devoid of CAMs: Indeed, the surface of all cells harbours a vast array of proteins and glycoproteins and interacting secreted glycans and proteoglycans. Although not often discussed, these molecules are thought to produce unspecific, yet non-negligible, cell-cell adhesion [60]. In fact, even in the absence of cadherins, most cells, show some weak degree of aggregation, including wild type L cells (A. Alsemarz and F. Fagotto, unpublished observations). We predict that non-specific adhesion would also be strongly favoured by reduced cortical tension (Fig.3C), potentially explaining the original observations by Litvinov and co-workers.

5. EpCAM signalling activities

Very early on, there were hints that EpCAM could influence cell signalling, as Ripani et al saw that anti-Trop2 antibodies stimulated transient intracellular calcium bursts [61]. While this lead has not been further pursued, evidence for other signalling activities have been reported. Although the distinction is perhaps somewhat arbitrary, we will present them in two categories, i.e. activities that seem to mostly impact on gene regulation and proliferation, and activities that act more directly on adhesion and migration. We will only briefly summarize the former, which have been reviewed recently [62], and expand on the latter, which are most relevant for the focus of this review.

5.1. Nuclear signalling

Endogenous EpCAM in carcinoma cells, and exogenous EpCAM expressed in HEK293 cells, can be cleaved by the RIP pathway, resulting first in shedding of the extracellular domain, followed by the release of the short cytoplasmic tail through the action of the γ-secretases on the transmembrane domain [7]. The released cytoplasmic peptide, called intracellular domain or ICD (Fig.4D) was found to bind to the multifunctional adaptor four-and-a-half LIM domain protein 2 (FHL2), which forms a complex with β-catenin and β-catenin transcriptional partners of the TCF/Lef1 family that binds DNA

and activates targets of the canonical Wnt-β-catenin-TCF pathway, such as c-Myc [7] and Cyclin D1 [31]. RIP of EpCAM2/Trop2 stimulates the same pathway and plays a role in stem cell self-renewal in the prostate [33]. The initial step in the cascade, i.e. EpEx shedding, can be performed either by the extracellular metalloproteinase ADAM17 (a so-called "α-secretase"), or by BACE1 ("β-secretase"). The former process is thought to be highly restricted to some cancer cell types expressing high ADAM17 levels [7]. The expression of BACE1 appears more widespread, but this enzyme has only weak activity at neutral pH. It has thus been proposed that BACE1-mediated shedding mostly occurs in the endosomal-lysosomal compartment [32]. Note that since the tumour environment tends to be acidic, extracellular EpCAM shedding by BACE1 remains possible [35]. In any case, a recent study from the Gires team quite conclusively demonstrates that the release of the intracellular fragments (it appears that there is more than one fragment) is quite slow, and their vast majority is rapidly degraded by the proteasome [36]. As a consequence, the amount of intracellular EpCAM tail available for regulation of β-catenin-dependent transcription is predicted to be rather low, irrespective of the relative efficiency of extracellular shedding. The actual impact of this low basal contribution to Wnt-β-catenin signalling remains to be determined. Huang et al hypothesize that EpCAM turnover, which is otherwise extremely slow (many hours), may be important during epithelial to mesenchymal transition (EMT). The authors propose that EpCAM rapid removal from the cell surface could release higher levels of the cytoplasmic tail before being degraded by the lysosome, contributing to increased Wnt-β-catenin signalling [36]. Regulation of shedding and intracellular tail release remains to be further investigated. In their original study, Gires et al found that incubation with the soluble EpEx stimulated cleavage [7]. Cleavage was also stimulated by cell density in culture, presumably by cell-cell contact [63]. Regulation via the EGF receptor (EGFR) was also proposed [64], but as noted by Olivier Gires, these data were misinterpreted [65] and could not be reproduced [36].

A different, and rather unexpected, connection was discovered between EpCAM and EGFR (Fig.4D). Indeed, in head and neck cancers, the EpEx appears to act as a ligand for EGFR and to activate Erk and Akt pathways [66]. Interestingly, EpEx stimulated proliferation but antagonized to the normal EGF-dependent activation of EMT-driving transcription factors, such as Snail and Zeb1. This work raises many interesting questions. Indeed, one may wonder whether full length EpCAM may also be capable of binding to EGFR, and whether EpEx (or EpCAM) interacts in a cis or trans orientation, in other words if this may constitute in an autocrine, juxtracrine and/or paracrine signal. Sankpal et al found an opposite role of EpCAM, which negatively regulated Erk signalling in a variety of cancer cell lines. In addition, Erk signalling repressed EpCAM expression, establishing a double-negative feedback loop [67]. In this case, the mechanism through which EpCAM controlled Erk signalling was not identified. EpCAM was also found to negatively modulate NFkB signalling [59]. In the zebrafish endoderm, EpCAM was also reported to have a positive impact on the Wnt pathway by a completely different mechanism, which involved direct interaction between EpCAM and the transmembrane protein Kremen1 [68]. Kremen1 normally forms a tertiary complex with Dickkopf and the Wnt receptor LRP6, which stimulates LRP6 internalization. EpCAM competes with formation of this complex, resulting in the stabilization of LRP6 at the cell surface and enhanced downstream signalling [68].

5.2. Signalling with impact on the cytoskeleton and cell junctions

5.2.1. PKC Inhibition

Gain- and loss-of-function (GOF and LOF) experiments in zebrafish and Xenopus embryos showed that EpCAM expression has a strong impact on cell adhesive and migratory properties [8-10]. As mentioned above, and further explained in the next section, our team demonstrated that this effect was

independent of a hypothetical adhesive function, but rather due to the ability of EpCAM to inhibit PKCs, specifically members of the class of novel PKCs (nPKCs) [10] (Fig.4).

The PKC family of serine/threonine kinases has a special mode of regulation, based on an autoinhibitory intramolecular interaction between their catalytic domain and an internal pseudosubstrate sequence. This pseudosubstrate resembles a typical substrate phosphorylation sequence, but lacks the phosphorylatable serine or threonine. At rest, PKC is in a compact configuration, such that the pseudosubstrate binds the substrate recognition site of the kinase domain, blocking access to the catalytic site. PKCs are activated by diacylglycerol (DAG) generated in the inner leaflet of the plasma membrane (calcium is also needed for activation of classical PKCs, but not nPKCs). DAG-mediated recruitment to the plasma membrane is accompanied by decompaction, and the pseudosubstrate segment is pulled away from the kinase domain, exposing the catalytic site. PKCs have a large number of cellular substrates. The signature of PKC phosphorylation sites is the presence of several basic residues, which can be positioned on either side of the serine/threonine target, or on both [69, 70]. Beyond this common feature, different PKCs appear to have distinct substrate preferences, although it is difficult to define precisely these preferences, as many validated phosphorylated sites depart considerably from the experimentally determined consensus motifs [69-71].

EpCAM turned out to function as a direct PKC inhibitor: The highly conserved juxtramembrane cytoplasmic sequence is strikingly similar to PKC substrates and pseudosubstrates [10] (Fig.4A). *In vitro* GST pulldowns and surface plasmon resonance analysis demonstrated that human and Xenopus EpCAM juxtramembrane sequences interacted directly, with high affinity (K_D ~ 50nM, i.e. similar to PKC pseudosubstrates) with two nPKCs, namely PKCδ and η, but not to classical PKCβ [10]. In vitro kinase assays confirmed that the juxtamembrane domain behaves as a pseudosubstrate, strongly inhibiting nPKC activity, but not the classical PKCβ. Consistently, experiments using a panel of inhibitors on cells and embryonic tissues, showed that EpCAM mostly antagonized nPKCs, with little to no effect on classical and atypical PKCs [9, 10].

Considering the multitude and incredible variety of PKC substrates, one may wonder about the specificity of EpCAM inhibition. Immunostaining using a generic phospho-(Ser)PKC substrate antibody showed a dramatic, cell-wide decrease in substrate phosphorylation upon EpCAM overexpression, and, conversely, a general upregulation in response to EpCAM depletion, also confirmed by Western blot [9, 10]. These results suggested that EpCAM has a global impact on PKC-dependent processes, although a thorough identification of relevant PKC targets is still missing.

Figure 4

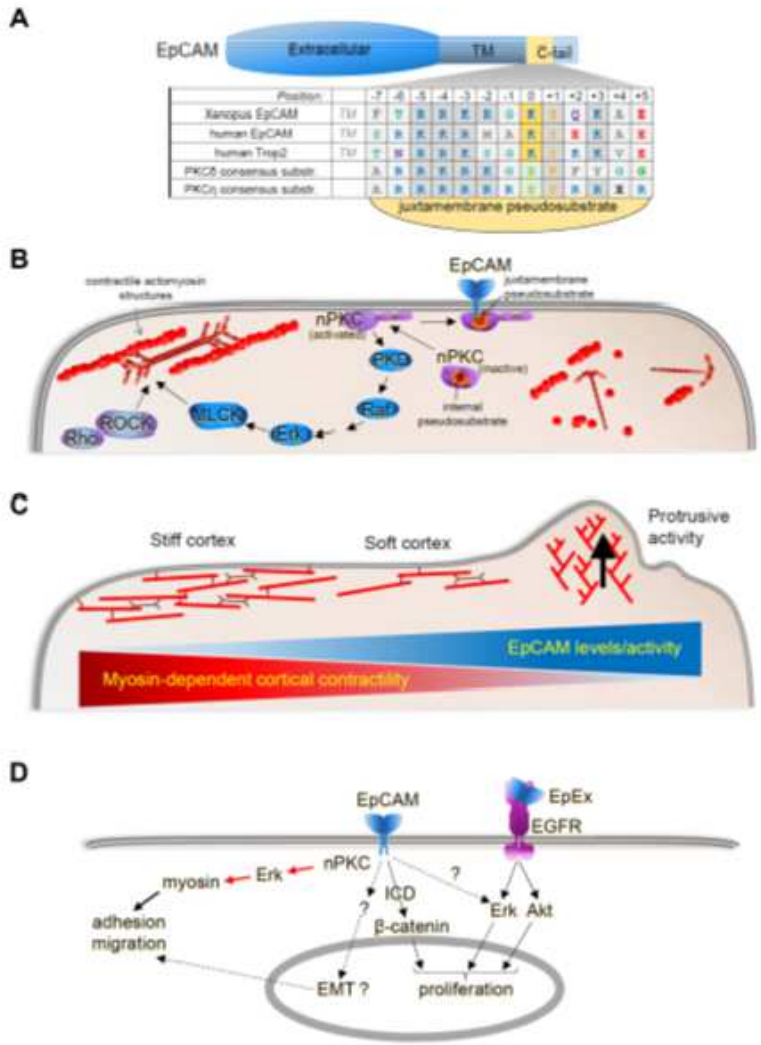

Figure 4. EpCAM represses actomyosin contractility via direct inhibition of novel PKCs.

(A) A short sequence of the EpCAM cytoplasmic tail acts as inhibitory pseudosubstrate for nPKCs. The table shows the alignment of Xenopus EpCAM, human EpCAM and Trop2 with the consensus sequences for two so-called novel PKCs, PKCδ and PKCη. The phosphorylated serine of the substrates is highlighted in green and the substitution by a lysine in EpCAM/Trop2 in yellow. Key residues of the consensus sequences are highlighted in grey. (B) Repression of the PKC-myosin pathway. PKC kinases are kept in an inactive state by autoinhibition, involving an internal pseudosubstrate sequence that masks the catalytic site. Recruitment at the membrane leads to PKC unfolding and activation. One of the pathways downstream of the novel PKCs (nPKC) involves PKD, the Raf-Erk cascade, and activation of myosin-light chain kinase (MLCK) by Erk phosphorylation. MLCK and ROCK are the two major kinases responsible for myosin II activation through phosphorylation of its regulatory light chain. EpCAM directed sequestration and inhibition of nPKCs leads to downregulation of the Erk pathway and myosin inactivation. (C) EpCAM levels control myosin-dependent contractility, softening the cell cortex and promoting protrusive activity. (D) Simplified diagram of EpCAM signalling functions. In addition of repression of the nPKC-myosin pathway, EpCAM can regulate gene expression and stimulate proliferation through various pathways. The two major routes are activation of β-catenin signalling by release of EpCAM intracellular domain (ICD), and activation of the EGF receptor (EGFR) by shedding EpCAM extracellular domain (EpEx). Other ill-defined signals are symbolized by dashed arrows and question marks, including potential stimulation of epithelial to mesenchymal transition, which in turn will regulate adhesion and migration.

Quite surprisingly, all GOF and LOF morphogenetic phenotypes observed in the Xenopus embryo could be explained by the inhibitory action of EpCAM on one particular pathway (Fig.4B). This pathway involves nPKCs, their direct target PKD/PKCμ and the Raf-Erk cascade. In turn, Erk was found to stimulate myosin II activity via myosin light chain kinase (MLCK) [10]. Consistently, EpCAM regulation of the same Erk-myosin branch has been confirmed in several other systems, including human intestinal Caco2 and SW480 cell lines [10, 55], and accounts for the function of EpCAM in maintaining intestinal epithelium homeostasis [55]. Thus, current evidence indicates that a major function of EpCAM is to moderate myosin activity (Fig.4C). We will see that this function influences cell and tissue properties in major ways.

One important outstanding question about this PKC inhibitory activity is its potential regulation. It has been reported in MDCK cells that the capacity of EpCAM to downregulate Erk and myosin was lost in the TM mutant, which was interpreted as evidence for a requirement of association with cldn7 [46]. As discussed above, the assignment to cldn interaction is equivocal, as the TM mutant may also be defective in other features, such as interaction with TEMs. Nevertheless, this result is in line with the possibility that different states of EpCAM may influence its ability to bind and inhibit PKCs, starting with its dimerization and/or its interaction with other components (Fig.2C).

5.2.2. Generalization of PKC inhibition to other membrane proteins

Before going into the physiological significance of this inhibitory activity of EpCAM, one should broach an interesting question related to the pseudosubstrate sequence. Its general characteristic, i.e. the presence of several basic amino acids, is far from unusual for a transmembrane protein: Most transmembrane helices are "clamped" and stabilized by immediately adjacent charged residues, which are predominantly basic on the cytoplasmic side, consistent with stabilizing interactions with the negatively charged phospholipids of the membrane inner leaflet [72]. Considering that the number and position of basic residues in PKC substrates is rather flexible, one may wonder how many membrane proteins other than EpCAM could be potential PKC inhibitors. Previous information on PKC substrates and on their own pseudosubstrate sequences indicated that there were additional features involved in PKC recognition, which we confirmed by systematic mutation of the EpCAM juxtamembrane region and their screen for PKC binding to PKC [10]. This juxtamembrane region appeared to be exquisitely tailored for this interaction. For instance, it contains a highly conserved Tyr297 at the +1 position, which is predicted to precisely fit into a hydrophobic pocket adjacent to the catalytic site [10]. Furthermore, the position of the EpCAM pseudosubstrate relative to the lipid bilayer is optimal to capture activated, membrane-associated PKC. Based on this analysis, we defined a consensus sequence, which was used to bioinformatically search through the human proteome for other the transmembrane proteins with a similar motif. We retrieved a list of about 40 candidate proteins. We chose a small subset for experimental validation, and found that all of them could bind nPKCs [10]. These candidates were entirely unrelated, yet, interestingly, the vast majority could be functionally grouped either as cell adhesion molecules, such as ICAM or NrCAM, or as cell-cell contact signalling molecules, such as the ephrin receptor EphA4. We thus postulated the existence of a new class of cell contact "receptors" that may downregulate PKC signalling [10]. This hypothesis awaits further experimental confirmation, which may open a new fact of signal regulation.

6. EpCAM in cell adhesion, migration and morphogenesis: the role of PKC inhibition

We have introduced the diverse roles of myosin in adhesion and migration: What is then the impact of its regulation by EpCAM on cell behaviour and on tissue properties? To start answering this question, we will first discuss the phenotypes observed in early fish and frog embryo models, and how they can

be explained at the cellular level. We will then present phenotypes reported in differentiated tissues and in mammalian cell lines, where the role of EpCAM is less understood.

6.1. EpCAM expression in early embryonic development.

While, in adults, EpCAM is typically an epithelial-specific component, its expression is much more widespread in early development. In Xenopus, EpCAM expression was investigated both at mRNA (Xenbase) and protein level [9, 10]. It is expressed in all germ layers throughout early stages, including gastrulation. It then becomes depleted from the neural field (Xenbase), while enriched in the epidermis as well as in the notochord [10]. In zebrafish, in situ hybridization showed that EpCAM is maternally inherited and ubiquitously expressed in the pre-gastrula stages [8], but at gastrulation, zygotic expression appears restricted to the superficial ectoderm layer, called the enveloping layer in fish*, expanding later to the whole epidermis [8]. In the absence of data on protein levels, one does not know whether EpCAM may still persist in other regions of the gastrula. In the mouse, information about the earliest stages (cleavage, morula, compaction) is missing, but in the blastocyst, EpCAM is expressed in embryonic stem cells, in both the epiblast and primitive endoderm lineages [57, 73], as well as in the trophectoderm (thus its alternate name for Trophoblast cell-surface antigen-1) [74]. EpCAM is then maintained in epithelial tissues, starting with ectoderm and visceral endoderm, but is prominently downregulated in the mesoderm and its derivatives, and in the neuroectoderm [73]. Interestingly, EpCAM is also expressed in germ cells throughout development and adult life [57]. EpCAM2/Trop2 expression in early mammalian embryos has not been investigated. At later stages, it is expressed in a variety of epithelia, but in distinct spatial and temporal patterns compared to EpCAM (reviewed in [18]).

**Note: Enveloping layer is the accepted term in the fish embryology field, but we prefer to refer to the superficial ectoderm layer, as it better relates to the conserved multi-layered organization of the ectoderm [75].*

6.2. EpCAM early loss-of-function phenotypes

Early embryonic EpCAM LOF phenotypes were reported in zebrafish, through the analysis of an EpCAM mutant [8], and in Xenopus, where EpCAM was depleted using antisense morpholino oligonucleotides [9, 10]. Globally, the phenotypes were strikingly similar: In both cases, loss of EpCAM led to a delay in epiboly. Epiboly is a major morphogenetic process of gastrulation, through which the ectoderm thins and spreads to eventually cover the other germ layers. It involves distinct changes in the superficial layer, which spreads mostly through cell stretching, and in the deep layer, where cells are rearranged via radial intercalation (Fig.5A). The EpCAM LOF defect was strong in Xenopus, subtler in zebrafish, but in all cases the process eventually recovered and gastrulation could be completed, allowing the observation of additional phenotypes at later stages [8-10]. In both species, post-gastrula embryos lacking EpCAM showed defects in the organization of the epidermis. In zebrafish, piles of round cells were found at the superficial layer, indicating loss of epithelial integrity [8]. The exact same phenotype was observed in Xenopus, except that it was again much more dramatic, and involved deeper layers as well: As embryos reached the late neurula stage, the epidermis started to fall apart, and eventually the entire embryo disintegrated [10] (Fig.5B). The reasons for weaker impact on zebrafish development are not known. A likely explanation is partially redundancy and/or compensation, due to the presence of a second EpCAM gene in zebrafish, called pan-epithelial glycoprotein*. In addition, the restriction of expression to the superficial ectoderm layer is consistent with EpCAM being most important for this layer, even though the deep layer is also abnormal in EpCAM fish mutants.

**Note that for better clarity and consistency with the usual gene nomenclature in zebrafish, the two genes should be called EpCAMa and EpCAMb.*

Figure 5

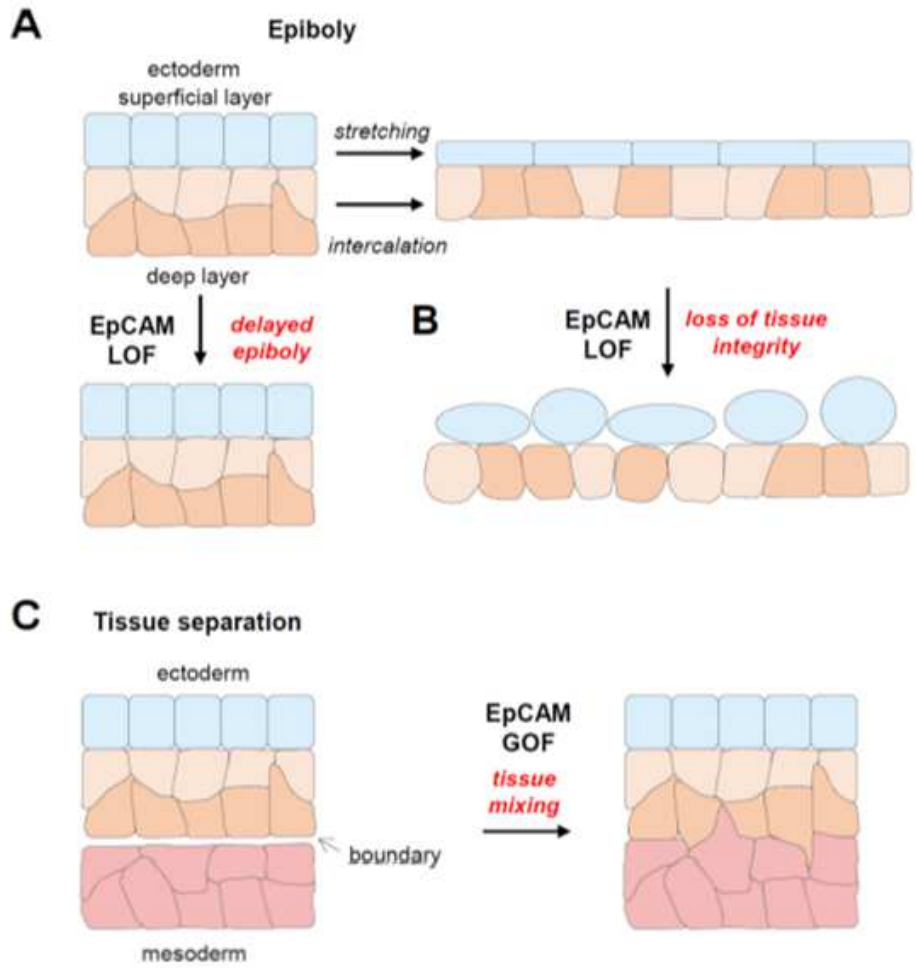

Figure 5. EpCAM and early embryonic morphogenesis

Three major phenotypes are linked to EpCAM. (A) During gastrulation, the ectoderm thins and spreads over the other tissues by the process of epiboly, which involves changes in the cell shape in the superficial layer and rearrangement of deep cells through radial intercalation. EpCAM loss-of-function (LOF) results in delayed epiboly. (B) Later during development, EpCAM LOF compromises tissue integrity. Indeed, uncontrolled cortical tension results in cell rounding and loss of cell-cell adhesion. (C) During gastrulation, the mesoderm is separated from the mesoderm by a so-called embryonic boundary. Elevated EpCAM levels (gain-of-function, GOF) induce mixing of the two tissues.

The high similarity of phenotypes was also observed at the cellular level: In both species, loss of EpCAM leads to impaired protrusive activity as well as a marked drop in cadherin levels [8-10]. The former is linked to lower motility, consistent with delayed epiboly. As for cadherin downregulation, it fitted well with the loss of cell adhesion and of tissue integrity, which led Hammerschmidt and colleagues to suggest that EpCAM acted synergistically with cadherin toward cell-cell adhesion in epithelial tissues [8, 9]. However, we will see below that experiments in Xenopus demonstrated that all these effects were due to myosin downregulation.

In the mouse, EpCAM knock out was reported to be embryonic lethal due to defects in the placenta [76], consistent with the early expression of EpCAM in the trophectoderm. However, two other studies reported that development of EpCAM null mice appeared normal, but mice died shortly after birth due to defects in the intestine closely resembling CTE [47, 77]. The cause for the inconsistency between these studies remains unresolved. The lack of early embryonic phenotype, despite strong expression in the blastocyst, may be again due to redundancy and/or compensation by EpCAM2/Trop2 (see below). EpCAM2/Trop2 knock out mice are viable and do not show any obvious developmental nor physiological phenotype, although mice are more susceptible to develop cancer [78].

Note the existence of additional LOF phenotypes, which have not been further analysed, but indicate that EpCAM is implicated in multiple morphogenetic processes. These include a decrease in proneuromast deposition in the lateral line organs of the zebrafish [79] and failure of the notochord to rearrange in a stack of coin-like cells in Xenopus (Maghzal and Fagotto, unpublished).

6.3. EpCAM gain-of-function and embryonic tissue mixing

In Xenopus, EpCAM had been originally spotted in a gain-of-function screen designed to identify molecules capable to induce mixing between ectoderm and mesoderm germ layers [9]. Indeed, during gastrulation, a boundary maintains a sharp separation between these two tissues, which is absolutely required for smooth mesoderm migration and proper gastrulation [80]. Increasing EpCAM expression led to cells of one tissue intruding in the other tissue, disrupting the boundary and interfering with gastrulation (Fig.5C). The effect was not tissue specific, since it was similarly obtained by increasing EpCAM levels either in the ectoderm or in the mesoderm. Further experiments showed that EpCAM acts as a general activator of so-called "intercellular migration", i.e. cells migrating using neighbouring cells as substrate. At the cellular level, elevated EpCAM stimulated protrusive activity and led to higher cadherin levels, perfectly mirroring the LOF phenotype [9].

6.4. EpCAM phenotypes depend on PKC/myosin regulation

It turned out that all embryonic and cellular phenotypes observed during Xenopus development could be accounted for by the nPKC/myosin inhibitory activity of EpCAM.

A model based on the putative function of EpCAM as an adhesion molecule was ruled out by the following decisive experiment [9]: An EpCAM construct lacking the whole extracellular domain was sufficient to rescue the epiboly phenotype in EpCAM-depleted embryos, and was also capable to stimulate intercellular migration, and to induce tissue mixing in GOF experiments [9]. It also induced cadherin downregulation. Conversely, an EpCAM construct lacking the intracellular tail failed to mimic wild type EpCAM. A refined analysis showed that substitution of two amino acids of the juxtamembrane region was sufficient to abrogate its activity [9, 10].

The role of nPKC inhibition was supported by multiple lines of evidence: Firstly, EpCAM LOF and GOF induced acute changes in PKC activity. PKC inhibitors mimicked EpCAM GOF and were sufficient to rescue EpCAM LOF. Furthermore, generic PKC activation using classical PMA treatment, as well as selective nPKC activation by the chemical activator Coleon U phenocopied the epiboly defect caused by EpCAM LOF, and rescued tissue separation in EpCAM GOF [9]. Beyond the early gastrula phenotypes, similar experiments using specific inhibitors and activators showed that nPKC deregulation was also responsible for loss of tissue integrity at later stages [10].

The ability of EpCAM to inhibit nPKCs also explained all effects at the cellular and molecular level, including impact on PKC signalling (e.g. PKD and Erk phosphorylation), and on cellular properties (myosin activation, actin cytoskeleton organization, cadherin levels and surface expression) [9, 10].

The same was true for the downstream target, myosin II. For instance, experimental myosin activation by co-expression of a constitutively active form of RhoA rescued the defect in ectoderm-mesoderm separation [9]. Conversely, myosin inhibition rescued EpCAM LOF. The most surprising and spectacular observation was made by placing EpCAM-depleted embryos in medium containing either a nPKC inhibitor, or the myosin inhibitor blebbistatin, a treatment that completely rescued tissue integrity and normal development [10]. This key experiment demonstrated that myosin repression was sufficient to account for EpCAM function during embryonic development.

At the cellular level, inhibition of any of the components of the nPKC-myosin cascade could rescue cadherin levels [10]. This crucial result unequivocally positioned cadherin downstream of myosin in this regulation. Note that cadherin overexpression was also sufficient to rescue tissue integrity and embryonic development [10]. This can be explained by both the direct action of cadherins in reinforcing of cell-cell adhesion and their intrinsic capacity to downregulate myosin. We proposed that loss of cell-cell adhesion upon EpCAM depletion results from a snowball effect, where myosin hyperactivation destabilizes adhesive contacts, which causes cadherin loss, which in turn leads to further contact weakening through the removal of cadherin-mediated myosin downregulation at the contact.

While similar experiments have not yet been performed in zebrafish, the regulatory pathway identified in Xenopus can readily explain the phenotypes in fish.

6.5. Impact on other systems

6.5.1 Intestinal epithelium

CTE is a rare but severe genetic disease that leads to accumulation of cell groups (tufts), associated with disruption of the intestinal epithelial barrier [81, 82]. EpCAM LOF mutations have been linked to three quarters of cases, and the connection with EpCAM is reinforced by similar defects detected in the EpCAM KO mouse [47, 77].

The EpCAM LOF phenotype was characterized at the cellular level by Delacour team [55]. The authors studied both intestinal biopsies from CTE patients and EpCAM-depleted stable line of Caco2 intestinal cells. The latter were also cultured on a micropattern 3D landscape reconstituting the geometry of the intestinal villi. They observed disruption of the epithelial integrity, severe loss of cadherin-based adherens junctions, expansion of the apical domain, and defects in the organization of the tight junctions, in particular at tricellular contacts [55]. They demonstrated that the phenotype could be attributed to exacerbated myosin activation. Note that perturbation of tricellular junctions was consistent with the fact that they are the sites of highest tension in an epithelial layer. The molecular link between EpCAM loss and myosin hyperactivation was not investigated by Salomon et al, but our team had previously shown that EpCAM depletion in Caco2 cells caused upregulation of the PCK-Erk-myosin pathway and decrease in cadherin levels [10]. One can then conclude that the role of EpCAM in the integrity of the differentiated intestinal epithelium appears to be identical to the one identified in the early Xenopus embryos.

One piece of the puzzle that has not yet been integrated is the EpCAM-cldn7 interaction. EpCAM KO and cldn7 KO in mice led to reciprocal downregulation of their partner [45, 47]. While in these studies, the CTE phenotype had been interpreted based on the EpCAM-cldn7 interaction, it is now clear that deregulation of the PKC-myosin has a major role in this phenotype. It would be important to revisit the EpCAM loss in the intestine to sort out the actual contributions of the two mechanisms. This pending issue is essential both to better understand EpCAM biology and to firmly establish the cause(s) of CTE.

Since EpCAM is expressed in all epithelia, it may seem surprising that the effect of its deletion is restricted to the intestine. An obvious explanation is that the intestine prominently lacks expression of EpCAM2/Trop2 [48], which presumably compensates in other tissues in humans with EpCAM mutations and in EpCAM KO mice. For instance, skin expresses both EpCAM genes, and their double depletion causes downregulation of cldns, while single depletions have no effect [38]. Presumably double KO could lead to tissue disruption, similar to Xenopus. To our knowledge, the only reported double KO was made in immortalized human corneal epithelial cells, and it caused defects in the epithelial barrier, although apparently the monolayer remained coherent [83]. A systematic analysis of double EpCAM1/EpCAM2 LOF is still missing.

6.5.2. Epithelial to mesenchymal transition and cancer cell lines

EpCAM is one of the typical epithelial markers that is generally lost during EMT [67, 84-86]. The same is true for EpCAM2/Trop2 [18]. This correlation, combined with the belief that EpCAM was an adhesion molecule, has naturally led to the assumption that its downregulation may be functionally linked to loss of cell adhesion and induction of migration, two landmarks of EMT. However, this simple model seems at odds with the strong correlation between EpCAM expression and malignancy. The potential relationship between EpCAM and EMT remains ambiguous, which, after all, is not that surprising considering the complexity of EpCAM biology. For instance, we have mentioned above that activation of EGFR-Erk signalling by EpEx was found to antagonize EMT in head and neck cancers [66]. In contrast, EpCAM was shown to be indispensable for TGFβ-induced EMT of MCF7 breast cancer cells [87]. Here, EpCAM expression was upregulated by TGFβ, through JNK, Jun and Fos, and EpCAM silencing prevented EMT through an uncharacterized mechanism [87]. Note that EpCAM repression could also simply be a secondary consequence of EMT, not necessarily causal to the observed changes in adhesive and migratory properties.

It is also important to note that the classical model of EMT as a central step in metastasis is being increasingly questioned. The metastatic state now appears much more dynamic. At least in some cases, it may involve a hybrid epithelial/mesenchymal phenotype [88, 89]. Furthermore, there is accumulating evidence for the existence of collective modes of invasion [90, 91]. Therefore, metastasis does not necessarily exclude maintenance of epithelial traits, including EpCAM expression. Consistently, Alix-Panabières and colleagues have isolated EpCAM-positive human circulating cancer stem cells from which they established lines that happened to be highly metastatic [92, 93]. EpCAM function should then be interrogated taking into consideration this radically new view of malignancy.

The field has not yet quite reached this stage, and most data reported on EpCAM function in cancer cell lines has been interpreted in a classical EMT perspective. Multiple reports have addressed the role of EpCAM and EpCAM2/Trop2 on adhesion and migration on various cell lines, with the aim to infer a potential invasive or on the contrary an anti-invasive function. Globally, most evidence are in favour of a proadhesive role (Table 1). In terms of migration, both promigratory and antimigratory phenotypes were reported, and no general consensus can be drawn (Table 2). EpCAM may certainly have antagonistic roles depending on the cell type and on experimental conditions. However, most of the information on this topic has remained rather superficial and difficult to interpret, except for a few well-documented studies, such as the analysis of EpCAM in MDCK cells by Barth et al [46]and the study of EpCAM2/Trop2 in embryonic fibroblasts [94]. One major issue is that most studies were solely based on two assays, i.e. migration through a filter (so-called transwell assays) and/or wound-healing after scratching a cell monolayer. Despite their popularity and apparent simplicity of their output, these assays deal with complex processes, which do not only depend on cell motility, but also on other parameters, most prominently cell-cell and cell-matrix adhesion. For instance, cell-cell adhesion is bound to antagonize cell migration in transwell assays. On the other hand, collective

migration, as monitored in the scratch assay, is ruled by completely different principles for cohesive epithelial monolayers or for low adhering mesenchymal-type cells. As for matrix-adhesion, one should remember that weak and strong adhesion may be equally detrimental to migration. In order to be correctly interpreted, these assays should be complemented with additional analysis, starting with a simple single cell migration assay, and some basic characterization at the cellular level (e.g. cadherin levels and localization, focal adhesions, cytoskeleton, epithelial organization).

Another potential caveat of studies concluding a promigratory role of EpCAM or EpCAM2/Trop2 is the systematic concomitant stimulation of proliferation, which, by affecting cell density during the course of the experiment can produce an "apparent" change in migration/invasion. An important control that should be included is to compare migration rates in the presence of a cell cycle inhibitor. Effects of cell survival and apoptosis may also obviously impact on the apparent migration rate. In the early Xenopus embryo, we could formally demonstrate that, while EpCAM LOF ultimately led to apoptosis, the latter was not the cause, but only a late consequence of loss of adhesion in this system [10]. The interpretation can be less unequivocal in long term experiments, for instance those involving establishment of stable lines, whether for ectopic expression of gene editing, since the process of selection is prone to complex modifications, which may eventually also impact on the adhesive and migratory properties. One strategy to circumvent such problems is the use of inducible cell lines, as reported by Gires and Untergasser teams [112, 113]. One study of the latter group exemplifies the importance of careful experimental design and interpretation. In this study on breast cancer cell lines, the authors showed that EpCAM expression levels had opposite effects on invasive capacities of epithelial MCF7 cells versus mesenchymal MDA-MB-231 cells. In summary, unlike the better characterized role in cell proliferation, the morphogenetic function of EpCAM and of EpCAM2/Trop2 in cancer remains poorly understood, and clearly requires a thorough re-evaluation.

6.6. Evaluating the potential diversity of EpCAM effects

We have seen that a consensus has emerged about the morphogenetic role of EpCAM in embryonic tissues and in homeostasis of the intestinal epithelium, but the situation is much less clear in cell lines and in cancer development. The following discussion aims at providing some keys to interpreting EpCAM-related phenotypes and help making predictions about its function in different situations. A first cause of potential complication is the dual action of EpCAM, through direct cytoplasmic signalling and transcriptional regulation (Fig.4D). Indeed, the latter activity, in addition to its role in promoting cell proliferation, is likely to modulate various genetic programs, which can in turn input on cell adhesion and migration. The above-mentioned cases of negative or positive impact on EMT clearly point toward such mechanisms.

Complexity is also expected at the level of cytoplasmic signalling. We have already mentioned that any component downstream of PKC inhibition can potentially modulate multiple targets, which in turn may act on the actin cytoskeleton and cell adhesion in various ways. For instance, multiple direct and indirect paths have been identified downstream of PKD1, which could contribute to its role in maintenance of the epithelial phenotype and repression of metastasis [114]. Furthermore, to add an extra layer of complexity, another member of the PKD family, PKD3, appears to have the exact opposite effect on cell migration [114]*.

**Note that the Erk-MLCK-myosin pathway is not mentioned in this reference.*

Figure 6

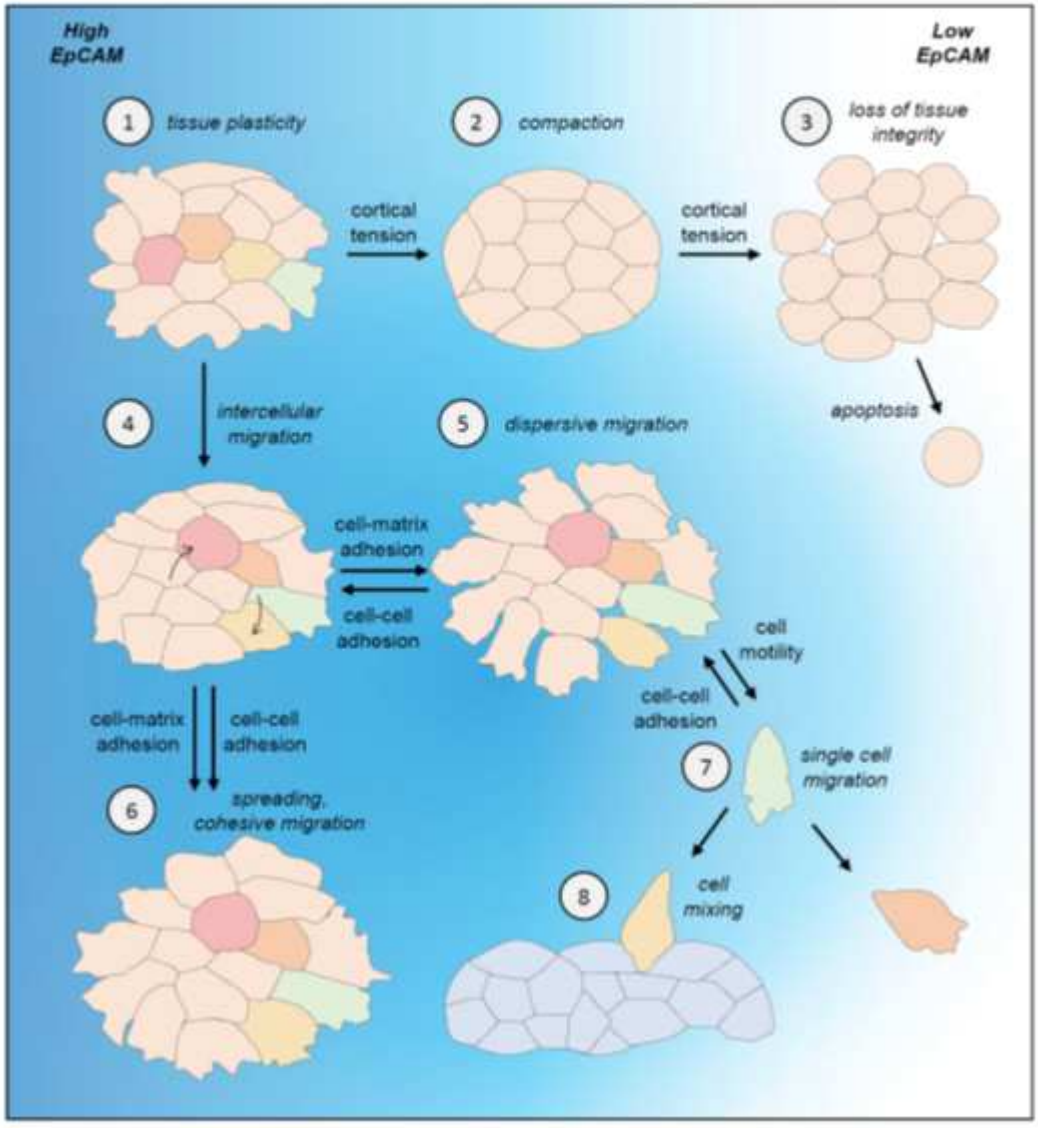

Figure 6.

General model for the role of EpCAM in regulating morphogenetic behaviours

The picture represents examples of behaviours resulting from the modulation of the four parameters dictating tissue rheology, i.e. cortical tension, motility, cell-cell adhesion, and cell-matrix adhesion. All four parameters can be influenced by EpCAM, as well as by multiple other regulators. EpCAM levels are symbolized by graded intensity of blue in the background. (1-8): Illustration of the variety of behaviours that can be produced by changes in EpCAM levels combined with other factors influencing adhesion and cell motility (see main text).

Keeping in mind these multiple potential alternative pathways, we will limit our discussion to the nPKC-myosin branch, firstly because it does appear as a major conserved role in tissue morphogenesis, but also because it nicely illustrates the range of effects that may be produced by this single activity. One traditionally opposes migration and adhesion as two antagonistic mechanisms, especially in the field of cancer invasion and EMT. Thus, the capacity of EpCAM to stimulate simultaneously protrusive activity/cell motility and cell-cell adhesion may appear contradictory. This property can be explained, however, based on the principles of tissue rheology presented at the beginning of this review. We have

seen indeed that it is determined by the balance of four major parameters, i.e. cortical tension, cell-cell adhesion, cell-matrix adhesion and cell motility. We have also seen that cortical tension is antagonistic to both adhesion and motility. EpCAM seems to primarily target myosin involved in cortical contractility, which is in principle sufficient to account for its proadhesive and promigratory activity. However, the actual impact of EpCAM will be influenced by other factors in different cell types and under different conditions, leading to a large spectrum of morphogenetic behaviours. We have attempted to illustrate a few simplified scenarios that may result from the action of EpCAM (levels represented by the blue intensity of the background), combined with additional changes in adhesion and motility (Fig.6). (1-3) Changes in cortical tension in response to EpCAM levels can modulate tissue rheology, ranging from soft dynamic (1) to compact and rigid (2), up to loss of adhesion in the most extreme case (3). One should here emphasize the importance of the context: One may predict that high EpCAM will be required to maintain dynamics of tissues with strong basal tension, and will be crucial to preserve the integrity of such tissues when exposed to high stress, which is typically the case of the epidermis and the intestine. Tissue with lower basal tension will be unlikely to undergo such dramatic disruption in the absence of EpCAM. Here, EpCAM will be more likely to fine tune tissue plasticity, favouring in particular intercellular migration (4). Hence, for instance, the tissue mixing phenotype and the epiboly defect observed in the early embryos. The balance between cell-cell and cell-matrix adhesion will also crucially influence tissue behaviour: If matrix adhesion dominates, the balance may be tilted toward a looser tissue configuration and a more dispersive mode of migration (5). Alternatively, there may be tissues where both cell-cell and matrix adhesion are high, resulting in a mode of coherent collective migration, which is the case of the Xenopus mesoderm (6). We have seen that EpCAM also plays a role in single cell migration (7). This action may be related to its association with TEMs and/or with integrins, but one currently lacks sufficient information to draw conclusion about the actual function in this context. Context-dependent tensile and adhesive properties are likely to explain the lack of overt effects of EpCAM LOF on adhesion and migration in some cell types [35]. Finally, EpCAM can stimulate intermingling of embryonic tissues [9]. One may then extrapolate that it could also contribute to heterotypic interactions between cell types during cell invasion (8), although this has not yet been experimentally addressed.

7. Open questions and concluding remarks

We have already highlighted several aspects of EpCAM biology that need to be clarified. We will here highlight the major general questions specifically related to its role in adhesion and migration.

7.1. Regulation of EpCAM morphogenetic activity

One obvious question concerning nPKC inhibition is whether it is a constitutive activity, or whether it may be controlled in time and space. There is no direct evidence so far for direct regulatory mechanisms, but a few pieces of the puzzle can be gathered from available data. While most cell surface receptors are activated by lateral clustering, there isn't a priori a clear rationale for such a requirement, at least if one assumes that binding of the juxtamembrane domain to the catalytic domain of a PKC is sufficient for its inhibition. On the contrary, indirect data suggest that this activity requires either interaction with cldn7, and/or with TEMs [46], and since such interaction appears to be mutually exclusive with EpCAM dimerization (Fig. 2), dimerization could be in fact inhibitory. A counterargument comes from the fact that PKC binding was demonstrated *in vitro* using a fusion between PKC cytoplasmic tail and GST, which is well-known to dimerize [10]. This issue needs to be resolved by more direct experiments. Another related question is to determine whether PKC inhibition is widespread along the whole plasma membrane, or concentrated in specific subcellular regions. Obvious possible locations would be cell-cell contacts and focal adhesions, which are precisely the sites where downregulation of cortical tension would have the most impact on promoting adhesion.

7.2. Relationship between various activities

So far, most studies have focused on a single function, either related to cell proliferation and transcriptional regulation (β-catenin, Erk), or to direct cellular processes, such as nPKC inhibition and myosin regulation, or interactions with cldns and the potential effect on tight junctions. A good example is the case of integrity of the intestinal epithelium, which had been first explained based on the EpCAM-cldn7 interaction [47], and subsequently by myosin regulation [55]. Conceivably, the defects in epithelial organization caused by myosin deregulation could also indirectly account from cldn mislocalization. Alternatively, EpCAM may contribute to epithelial function through two parallel mechanisms.

Similarly, we need to understand the apparently antagonistic relationship between Erk downregulation by the nPKC inhibitory pathway, and Erk upregulation via EGFR. Both pathways could well be genuinely antagonist. On the other hand, in a different model, one may pose that Erk inhibition may be locally restricted to the cell surface, while the EGFR-Erk pathway may target nuclear activity. Such specialization can be easily conceived assuming spatially and functionally distinct signalling complexes.

Addressing this type of question will require separation-of-function mutants which would lack specifically one of the interactions. The binding specificity to nPKC has been dissected in quite some detail [10], but nothing is known about binding to FHL2 and β-catenin, nor to EGFR. We already mentioned the requirement for better mutants to discriminate interaction with cldns.

7.3. EpCAM and EpCAM2/Trop2

EpCAM and its twin EpCAM2/Trop2 have very similar sequences, and most of the characteristics for which both of them were tested turned out to be conserved, spanning from biochemical features (e.g. cleavage by matriptase, interaction with cldns and with integrins), to stimulation proliferation, and strong link to poor cancer prognosis. Yet only few studies have directly compared their properties and their function in adhesion and migration [38, 83]. Validation of the PKC-Erk-myosin pathway in EpCAM2/Trop2 is prominently missing, although quite likely based on the juxtamembrane sequence.

EpCAM and EpCAM2 are likely to be redundant for most of their functions, as suggested by the fact that the EpCAM LOF phenotype in mice is restricted to the intestine, which happens to be one of the rare epithelia that only expresses EpCAM. This redundancy was directly demonstrated in keratinocytes, where the double depletion was required to observe cldn downregulation [38]. On the other hand, the simple fact of the specific exclusion of EpCAM2 from the intestine should be viewed as indicative of qualitative functional differences. This hypothesis is fully consistent with the strict conservation of detailed sequence specificities that unambiguously differentiate mammalian EpCAM from EpCAM2/Trop2. The only clear difference reported so far is the phosphorylation by PKC of a serine located at the C-terminal tip of EpCAM2/Trop2, which is absent in EpCAM [24]. Experiments using amino acid substitutions to mimic constitutively phosphorylated or dephosphorylated states argued in favour of a role of phosphorylation in stimulating cell migration [24]. However, the corresponding residue in EpCAM is an asparagine, which constitutes a conservative substitution in terms of hydrophilicity, suggesting as possible alternative explanation that this residue is important independently of a potential phosphorylation. Mutation of this residue in EpCAM should discriminate between these possibilities. Lastly, whether EpCAM and EpCAM2/Trop2 may be able to form heterodimers, or at least to be incorporated in the same TEMs, has never been addressed. We suspect that this is likely to occur. If so, the next question would be to look for potential functional specificities.

7.3. Single versus collective migration

The spectacular phenotypes occurring at the tissue level in embryonic systems and in the intestinal model should be a strong incentive for cancer biologists to resolutely move toward the analysis of EpCAM function in multicellular systems, and in particular in collective migration, as it is likely that important clues to understand the role of EpCAM in cancer progression will be revealed.

Conclusion

Our view of EpCAM has considerably evolved over the past decade, revealing multiple mechanisms through which it can influence cellular functions. Among them, regulation of morphogenetic behaviour through PKC inhibition and myosin downregulation appears to be a major direct pathway. Scattered circumstantial evidence suggest that EpCAM may have other ways to control adhesion and migration, which await further characterization. Moreover, the EpCAM-dependent nuclear activities are likely to indirectly impact on these parameters, in particular by modulating the genetic program driving EMT. While its original role as CAM is not being supported by recent data, it remains an intriguing possibility that should still be seriously considered.

Irrespective of whether it may indeed act as a bona fide homophilic or perhaps heterophilic CAM, we have seen that its proadhesive property can be explained purely by myosin regulation through intracellular signalling. Furthermore, we have discussed how this activity can, in principle, lead to a variety of behaviours in a cell type and context dependent manner. One should stress that the capacity of EpCAM to coordinate increased cell adhesion and cell motility is not a common effect of myosin regulators, and represents a rather unique and remarkable feature of this molecule, particularly fitted to stimulate intercellular migration, and thus tissue plasticity.

Acknowledgements

We would like to thanks David Rozema for careful reading and helpful comments on the manuscript. FF research is supported by ANR grant ANR-14-ACHN-0004–ICM, the Labex EpiGenMed, and ARC Fundation grant 206766.

Competing interests

The authors declare that no competing interests exist.

References

[1] S.V. Litvinov, H.A. Bakker, M.M. Gourevitch, M.P. Velders, S.O. Warnaar, Evidence for a role of the epithelial glycoprotein 40 (Ep-CAM) in epithelial cell-cell adhesion, Cell Adhes Commun, 2 (1994) 417-428.
[2] S.V. Litvinov, M.P. Velders, H.A. Bakker, G.J. Fleuren, S.O. Warnaar, Ep-CAM: a human epithelial antigen is a homophilic cell–cell adhesion molecule., J. Cell Biol., 125 (1994) 437–446.
[3] S. Kuhn, M. Koch, T. Nubel, M. Ladwein, D. Antolovic, P. Klingbeil, D. Hildebrand, G. Moldenhauer, L. Langbein, W.W. Franke, J. Weitz, M. Zoller, A complex of EpCAM, claudin-7, CD44 variant isoforms, and tetraspanins promotes colorectal cancer progression, Mol Cancer Res, 5 (2007) 553-567.
[4] M. Ladwein, U.F. Pape, D.S. Schmidt, M. Schnolzer, S. Fiedler, L. Langbein, W.W. Franke, G. Moldenhauer, M. Zoller, The cell-cell adhesion molecule EpCAM interacts directly with the tight junction protein claudin-7, Exp Cell Res, 309 (2005) 345-357.
[5] C. Claas, J. Wahl, D.J. Orlicky, H. Karaduman, M. Schnolzer, T. Kempf, M. Zoller, The tetraspanin D6.1A and its molecular partners on rat carcinoma cells, Biochem J, 389 (2005) 99-110.
[6] T. Nubel, J. Preobraschenski, H. Tuncay, T. Weiss, S. Kuhn, M. Ladwein, L. Langbein, M. Zoller, Claudin-7 regulates EpCAM-mediated functions in tumor progression, Mol Cancer Res, 7 (2009) 285-299.
[7] D. Maetzel, S. Denzel, B. Mack, M. Canis, P. Went, M. Benk, C. Kieu, P. Papior, P.A. Baeuerle, M. Munz, O. Gires, Nuclear signalling by tumour-associated antigen EpCAM, Nat Cell Biol, 11 (2009) 162-171.
[8] K. Slanchev, T.J. Carney, M.P. Stemmler, B. Koschorz, A. Amsterdam, H. Schwarz, M. Hammerschmidt, The Epithelial Cell Adhesion Molecule EpCAM Is Required for Epithelial Morphogenesis and Integrity during Zebrafish Epiboly and Skin Development, PLoS Genet, 5 (2009) e1000563.
[9] N. Maghzal, E. Vogt, W. Reintsch, J.S. Fraser, F. Fagotto, The tumor associated EpCAM regulates morphogenetic movements through intracellular signaling., J Cell Biol, 119 (2010) 645-659.
[10] N. Maghzal, H.A. Kayali, N. Rohani, A.V. Kajava, F. Fagotto, EpCAM Controls Actomyosin Contractility and Cell Adhesion by Direct Inhibition of PKC, Dev Cell, 27 (2013) 263-277.
[11] M. Kelkar, P. Bohec, G. Charras, Mechanics of the cellular actin cortex: From signalling to shape change, Curr Opin Cell Biol, 66 (2020) 69-78.
[12] J.D. Amack, M.L. Manning, Knowing the boundaries: extending the differential adhesion hypothesis in embryonic cell sorting, Science, 338 (2012) 212-215.

[13] M. Michael, A.S. Yap, The regulation and functional impact of actin assembly at cadherin cell-cell adhesions, Seminars in cell & developmental biology, 24 (2013) 298-307.
[14] R. Winklbauer, Cell adhesion strength from cortical tension - an integration of concepts, J Cell Sci, 128 (2015) 3687-3693.
[15] S.E. Parent, D. Barua, R. Winklbauer, Mechanics of Fluid-Filled Interstitial Gaps. I. Modeling Gaps in a Compact Tissue, Biophys J, 113 (2017) 913-922.
[16] R.M. Mege, N. Ishiyama, Integration of Cadherin Adhesion and Cytoskeleton at Adherens Junctions, Cold Spring Harb Perspect Biol, 9 (2017).
[17] R.G. Fehon, P.J. Kooh, I. Rebay, C.L. Regan, T. Xu, M.A. Muskavitch, S. Artavanis-Tsakonas, Molecular interactions between the protein products of the neurogenic loci Notch and Delta, two EGF-homologous genes in Drosophila, Cell, 61 (1990) 523-534.
[18] A.R. McDougall, M. Tolcos, S.B. Hooper, T.J. Cole, M.J. Wallace, Trop2: from development to disease, Dev Dyn, 244 (2015) 99-109.
[19] M. Pavsic, G. Guncar, K. Djinovic-Carugo, B. Lenarcic, Crystal structure and its bearing towards an understanding of key biological functions of EpCAM, Nature communications, 5 (2014) 4764.
[20] M.N. Mbaye, Q. Hou, S. Basu, F. Teheux, F. Pucci, M. Rooman, A comprehensive computational study of amino acid interactions in membrane proteins, Sci Rep, 9 (2019) 12043.
[21] M. Pavsic, G. Ilc, T. Vidmar, J. Plavec, B. Lenarcic, The cytosolic tail of the tumor marker protein Trop2--a structural switch triggered by phosphorylation, Sci Rep, 5 (2015) 10324.
[22] R. Tonikian, Y. Zhang, S.L. Sazinsky, B. Currell, J.H. Yeh, B. Reva, H.A. Held, B.A. Appleton, M. Evangelista, Y. Wu, X. Xin, A.C. Chan, S. Seshagiri, L.A. Lasky, C. Sander, C. Boone, G.D. Bader, S.S. Sidhu, A specificity map for the PDZ domain family, PLoS Biol, 6 (2008) e239.
[23] A. Basu, D.M. Goldenberg, R. Stein, The epithelial/carcinoma antigen EGP-1, recognized by monoclonal antibody RS7-3G11, is phosphorylated on serine 303, Int J Cancer, 62 (1995) 472-479.
[24] Y. Mori, K. Akita, K. Ojima, S. Iwamoto, T. Yamashita, E. Morii, H. Nakada, Trophoblast cell surface antigen 2 (Trop-2) phosphorylation by protein kinase C alpha/delta (PKCalpha/delta) enhances cell motility, J Biol Chem, 294 (2019) 11513-11524.
[25] M. Munz, K. Fellinger, T. Hofmann, B. Schmitt, O. Gires, Glycosylation is crucial for stability of tumour and cancer stem cell antigen EpCAM, Front Biosci, 13 (2008) 5195-5201.
[26] C. Pauli, M. Munz, C. Kieu, B. Mack, P. Breinl, B. Wollenberg, S. Lang, R. Zeidler, O. Gires, Tumor-specific glycosylation of the carcinoma-associated epithelial cell adhesion molecule EpCAM in head and neck carcinomas, Cancer Lett, 193 (2003) 25-32.
[27] M. Balzar, I.H. Briaire-de Bruijn, H.A.M. Rees-Bakker, F.A. Prins, W. Helfrich, L. de Leij, G. Riethmuller, S. Alberti, S.O. Warnaar, G.J. Fleuren, S.V. Litvinov, Epidermal Growth Factor-Like Repeats Mediate Lateral and Reciprocal Interactions of Ep-CAM Molecules in Homophilic Adhesions, Mol. Cell. Biol., 21 (2001) 2570-2580.
[28] A. Gaber, S.J. Kim, R.M. Kaake, M. Bencina, N. Krogan, A. Sali, M. Pavsic, B. Lenarcic, EpCAM homo-oligomerization is not the basis for its role in cell-cell adhesion, Sci Rep, 8 (2018) 13269.
[29] E. Li, W.C. Wimley, K. Hristova, Transmembrane helix dimerization: beyond the search for sequence motifs, Biochim Biophys Acta, 1818 (2012) 183-193.
[30] U. Schnell, J. Kuipers, B.N. Giepmans, EpCAM proteolysis: new fragments with distinct functions?, Biosci Rep, 33 (2013) e00030.
[31] A. Chaves-Perez, B. Mack, D. Maetzel, H. Kremling, C. Eggert, U. Harreus, O. Gires, EpCAM regulates cell cycle progression via control of cyclin D1 expression, Oncogene, 32 (2013) 641-650.
[32] M. Hachmeister, K.D. Bobowski, S. Hogl, B. Dislich, A. Fukumori, C. Eggert, B. Mack, H. Kremling, S. Sarrach, F. Coscia, W. Zimmermann, H. Steiner, S.F. Lichtenthaler, O. Gires, Regulated intramembrane proteolysis and degradation of murine epithelial cell adhesion molecule mEpCAM, PLoS One, 8 (2013) e71836.
[33] T. Stoyanova, A.S. Goldstein, H. Cai, J.M. Drake, J. Huang, O.N. Witte, Regulated proteolysis of Trop2 drives epithelial hyperplasia and stem cell self-renewal via beta-catenin signaling, Genes Dev, 26 (2012) 2271-2285.
[34] C. Driemel, H. Kremling, S. Schumacher, D. Will, J. Wolters, N. Lindenlauf, B. Mack, S.A. Baldus, V. Hoya, J.M. Pietsch, P. Panagiotidou, K. Raba, C. Vay, D. Vallbohmer, U. Harreus, W.T. Knoefel, N.H. Stoecklein, O. Gires, Context-dependent adaption of EpCAM expression in early systemic esophageal cancer, Oncogene, 33 (2014) 4904-4915.
[35] T. Tsaktanis, H. Kremling, M. Pavsic, R. von Stackelberg, B. Mack, A. Fukumori, H. Steiner, F. Vielmuth, V. Spindler, Z. Huang, J. Jakubowski, N.H. Stoecklein, E. Luxenburger, K. Lauber, B. Lenarcic, O. Gires, Cleavage and cell adhesion properties of human epithelial cell adhesion molecule (HEPCAM), J Biol Chem, 290 (2015) 24574-24591.
[36] Y. Huang, A. Chanou, G. Kranz, M. Pan, V. Kohlbauer, A. Ettinger, O. Gires, Membrane-associated epithelial cell adhesion molecule is slowly cleaved by gamma-secretase prior to efficient proteasomal degradation of its intracellular domain, J Biol Chem, 294 (2019) 3051-3064.
[37] C.J. Wu, X. Feng, M. Lu, S. Morimura, M.C. Udey, Matriptase-mediated cleavage of EpCAM destabilizes claudins and dysregulates intestinal epithelial homeostasis, J Clin Invest, 127 (2017) 623-634.
[38] C.J. Wu, M. Lu, X. Feng, G. Nakato, M.C. Udey, Matriptase Cleaves EpCAM and TROP2 in Keratinocytes, Destabilizing Both Proteins and Associated Claudins, Cells, 9 (2020).
[39] D.S. Schmidt, P. Klingbeil, M. Schnolzer, M. Zoller, CD44 variant isoforms associate with tetraspanins and EpCAM., Exp. Cell Res., 297 (2004) 329-347.
[40] M.E. Hemler, Tetraspanin functions and associated microdomains, Nat Rev Mol Cell Biol, 6 (2005) 801-811.

[41] S. Charrin, S. Jouannet, C. Boucheix, E. Rubinstein, Tetraspanins at a glance, J Cell Sci, 127 (2014) 3641-3648.
[42] J. Piontek, S.M. Krug, J. Protze, G. Krause, M. Fromm, Molecular architecture and assembly of the tight junction backbone, Biochim Biophys Acta Biomembr, 1862 (2020) 183279.
[43] C.J. Wu, P. Mannan, M. Lu, M.C. Udey, Epithelial cell adhesion molecule (EpCAM) regulates claudin dynamics and tight junctions, J Biol Chem, 288 (2013) 12253-12268.
[44] C.M. Van Itallie, J.M. Anderson, Claudin interactions in and out of the tight junction, Tissue Barriers, 1 (2013) e25247.
[45] H. Tanaka, M. Takechi, H. Kiyonari, G. Shioi, A. Tamura, S. Tsukita, Intestinal deletion of Claudin-7 enhances paracellular organic solute flux and initiates colonic inflammation in mice, Gut, 64 (2015) 1529-1538.
[46] A.I.M. Barth, H. Kim, I.H. Riedel-Kruse, Regulation of epithelial migration by epithelial cell adhesion molecule requires its Claudin-7 interaction domain, PLoS One, 13 (2018) e0204957.
[47] Z. Lei, T. Maeda, A. Tamura, T. Nakamura, Y. Yamazaki, H. Shiratori, K. Yashiro, S. Tsukita, H. Hamada, EpCAM contributes to formation of functional tight junction in the intestinal epithelium by recruiting claudin proteins, Dev Biol, 371 (2012) 136-145.
[48] M. Nakatsukasa, S. Kawasaki, K. Yamasaki, H. Fukuoka, A. Matsuda, M. Tsujikawa, H. Tanioka, M. Nagata-Takaoka, J. Hamuro, S. Kinoshita, Tumor-associated calcium signal transducer 2 is required for the proper subcellular localization of claudin 1 and 7: implications in the pathogenesis of gelatinous drop-like corneal dystrophy, Am J Pathol, 177 (2010) 1344-1355.
[49] M. Trerotola, D.L. Jernigan, Q. Liu, J. Siddiqui, A. Fatatis, L.R. Languino, Trop-2 promotes prostate cancer metastasis by modulating beta(1) integrin functions, Cancer Res, 73 (2013) 3155-3167.
[50] M. Trerotola, K.K. Ganguly, L. Fazli, C. Fedele, H. Lu, A. Dutta, Q. Liu, T. De Angelis, L.W. Riddell, N.A. Riobo, M.E. Gleave, A. Zoubeidi, R.G. Pestell, D.C. Altieri, L.R. Languino, Trop-2 is up-regulated in invasive prostate cancer and displaces FAK from focal contacts, Oncotarget, 6 (2015) 14318-14328.
[51] J. Yang, T. Isaji, G. Zhang, F. Qi, C. Duan, T. Fukuda, J. Gu, EpCAM associates with integrin and regulates cell adhesion in cancer cells, Biochem Biophys Res Commun, 522 (2020) 903-909.
[52] T. El Sewedy, M. Fornaro, S. Alberti, Cloning of the murine TROP2 gene: conservation of a PIP2-binding sequence in the cytoplasmic domain of TROP-2, Int J Cancer, 75 (1998) 324-330.
[53] M. Balzar, H.A. Bakker, I.H. Briaire-de-Bruijn, G.J. Fleuren, S.O. Warnaar, S.V. Litvinov, Cytoplasmic tail regulates the intercellular adhesion function of the epithelial cell adhesion molecule, Mol Cell Biol, 18 (1998) 4833-4843.
[54] M. Balzar, F.A. Prins, H.A. Bakker, G.J. Fleuren, S.O. Warnaar, S.V. Litvinov, The structural analysis of adhesions mediated by Ep-CAM, Exp Cell Res, 246 (1999) 108-121.
[55] J. Salomon, C. Gaston, J. Magescas, B. Duvauchelle, D. Canioni, L. Sengmanivong, A. Mayeux, G. Michaux, F. Campeotto, J. Lemale, J. Viala, F. Poirier, N. Minc, J. Schmitz, N. Brousse, B. Ladoux, O. Goulet, D. Delacour, Contractile forces at tricellular contacts modulate epithelial organization and monolayer integrity, Nature communications, 8 (2017) 13998.
[56] S.V. Litvinov, M. Balzar, M.J. Winter, H.A. Bakker, I.H. Briaire-de Bruijn, F. Prins, G.J. Fleuren, S.O. Warnaar, Epithelial cell adhesion molecule (Ep-CAM) modulates cell-cell interactions mediated by classic cadherins, J Cell Biol, 139 (1997) 1337-1348.
[57] R. Anderson, K. Schaible, J. Heasman, C. Wylie, Expression of the homophilic adhesion molecule, Ep-CAM, in the mammalian germ line, J Reprod Fertil, 116 (1999) 379-384.
[58] Y.T. Hsu, P. Osmulski, Y. Wang, Y.W. Huang, L. Liu, J. Ruan, V.X. Jin, N.B. Kirma, M.E. Gaczynska, T.H. Huang, EpCAM-Regulated Transcription Exerts Influences on Nanomechanical Properties of Endometrial Cancer Cells That Promote Epithelial-to-Mesenchymal Transition, Cancer Res, 76 (2016) 6171-6182.
[59] N.V. Sankpal, T.P. Fleming, W.E. Gillanders, EpCAM modulates NF-kappaB signaling and interleukin-8 expression in breast cancer, Mol Cancer Res, 11 (2013) 418-426.
[60] R. Winklbauer, Dynamic cell-cell adhesion mediated by pericellular matrix interaction - a hypothesis, J Cell Sci, 132 (2019).
[61] E. Ripani, A. Sacchetti, D. Corda, S. Alberti, Human Trop-2 is a tumor-associated calcium signal transducer., Int. J. Cancer, 76 (1998) 671-676.
[62] O. Gires, M. Pan, H. Schinke, M. Canis, P.A. Baeuerle, Expression and function of epithelial cell adhesion molecule EpCAM: where are we after 40 years?, Cancer metastasis reviews, (2020).
[63] S. Denzel, D. Maetzel, B. Mack, C. Eggert, G. Barr, O. Gires, Initial activation of EpCAM cleavage via cell-to-cell contact, BMC Cancer, 9 (2009) 402.
[64] Y.T. Hsu, P. Osmulski, Y. Wang, Y.W. Huang, L. Liu, J. Ruan, V.X. Jin, N.B. Kirma, M.E. Gaczynska, T.H. Huang, EGFR-Dependent Regulated Intramembrane Proteolysis of EpCAM-Response, Cancer Res, 77 (2017) 1777.
[65] O. Gires, EGFR-Dependent Regulated Intramembrane Proteolysis of EpCAM-Letter, Cancer Res, 77 (2017) 1775-1776.
[66] M. Pan, H. Schinke, E. Luxenburger, G. Kranz, J. Shakhtour, D. Libl, Y. Huang, A. Gaber, M. Pavsic, B. Lenarcic, J. Kitz, M. Jakob, S. Schwenk-Zieger, M. Canis, J. Hess, K. Unger, P. Baumeister, O. Gires, 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 Biol, 16 (2018) e2006624.
[67] N.V. Sankpal, T.P. Fleming, P.K. Sharma, H.J. Wiedner, W.E. Gillanders, A double-negative feedback loop between EpCAM and ERK contributes to the regulation of epithelial-mesenchymal transition in cancer, Oncogene, 36 (2017) 3706-3717.

[68] H. Lu, J. Ma, Y. Yang, W. Shi, L. Luo, EpCAM is an endoderm-specific Wnt derepressor that licenses hepatic development, Dev Cell, 24 (2013) 543-553.
[69] A.C. Newton, Protein kinase C: structural and spatial regulation by phosphorylation, cofactors, and macromolecular interactions, Chem Rev, 101 (2001) 2353-2364.
[70] K. Nishikawa, A. Toker, F.J. Johannes, Z. Songyang, L.C. Cantley, Determination of the specific substrate sequence motifs of protein kinase C isozymes, J Biol Chem, 272 (1997) 952-960.
[71] S.F. Steinberg, Distinctive activation mechanisms and functions for protein kinase Cdelta, Biochem J, 384 (2004) 449-459.
[72] M. Bogdanov, W. Dowhan, H. Vitrac, Lipids and topological rules governing membrane protein assembly, Biochim Biophys Acta, 1843 (2014) 1475-1488.
[73] S. Sarrach, Y. Huang, S. Niedermeyer, M. Hachmeister, L. Fischer, S. Gille, M. Pan, B. Mack, G. Kranz, D. Libl, J. Merl-Pham, S.M. Hauck, E. Paoluzzi Tomada, M. Kieslinger, I. Jeremias, A. Scialdone, O. Gires, Spatiotemporal patterning of EpCAM is important for murine embryonic endo- and mesodermal differentiation, Sci Rep, 8 (2018) 1801.
[74] M. Lipinski, D.R. Parks, R.V. Rouse, L.A. Herzenberg, Human trophoblast cell-surface antigens defined by monoclonal antibodies, Proc Natl Acad Sci U S A, 78 (1981) 5147-5150.
[75] F. Fagotto, Tissue segregation in the early vertebrate embryo, Seminars in cell & developmental biology, (2020).
[76] K. Nagao, J. Zhu, M.B. Heneghan, J.C. Hanson, M.I. Morasso, L. Tessarollo, S. Mackem, M.C. Udey, Abnormal placental development and early embryonic lethality in EpCAM-null mice, PLoS One, 4 (2009) e8543.
[77] E. Guerra, R. Lattanzio, R. La Sorda, F. Dini, G.M. Tiboni, M. Piantelli, S. Alberti, mTrop1/Epcam knockout mice develop congenital tufting enteropathy through dysregulation of intestinal E-cadherin/beta-catenin, PLoS One, 7 (2012) e49302.
[78] J. Wang, K. Zhang, D. Grabowska, A. Li, Y. Dong, R. Day, P. Humphrey, J. Lewis, R.D. Kladney, J.M. Arbeit, J.D. Weber, C.H. Chung, L.S. Michel, Loss of Trop2 promotes carcinogenesis and features of epithelial to mesenchymal transition in squamous cell carcinoma, Mol Cancer Res, 9 (2011) 1686-1695.
[79] E.J. Villablanca, A. Renucci, D. Sapede, V. Lec, F. Soubiran, P.C. Sandoval, C. Dambly-Chaudiere, A. Ghysen, M.L. Allende, Control of cell migration in the zebrafish lateral line: implication of the gene "tumour-associated calcium signal transducer," tacstd, Dev Dyn, 235 (2006) 1578-1588.
[80] F. Fagotto, The cellular basis of tissue separation, Development, 141 (2014) 3303-3318.
[81] O. Goulet, J. Salomon, F. Ruemmele, N.P. de Serres, N. Brousse, Intestinal epithelial dysplasia (tufting enteropathy), Orphanet J Rare Dis, 2 (2007) 20.
[82] M. Sivagnanam, J.L. Mueller, H. Lee, Z. Chen, S.F. Nelson, D. Turner, S.H. Zlotkin, P.B. Pencharz, B.Y. Ngan, O. Libiger, N.J. Schork, J.E. Lavine, S. Taylor, R.O. Newbury, R.D. Kolodner, H.M. Hoffman, Identification of EpCAM as the gene for congenital tufting enteropathy, Gastroenterology, 135 (2008) 429-437.
[83] P. Xu, C. Kai, S. Kawasaki, Y. Kobayashi, K. Yamamoto, M. Tsujikawa, R. Hayashi, K. Nishida, A New in Vitro Model of GDLD by Knocking Out TACSTD2 and Its Paralogous Gene EpCAM in Human Corneal Epithelial Cells, Transl Vis Sci Technol, 7 (2018) 30.
[84] B.T. van der Gun, L.J. Melchers, M.H. Ruiters, L.F. de Leij, P.M. McLaughlin, M.G. Rots, EpCAM in carcinogenesis: the good, the bad or the ugly, Carcinogenesis, 31 (2010) 1913-1921.
[85] T.M. Gorges, I. Tinhofer, M. Drosch, L. Rose, T.M. Zollner, T. Krahn, O. von Ahsen, Circulating tumour cells escape from EpCAM-based detection due to epithelial-to-mesenchymal transition, BMC Cancer, 12 (2012) 178.
[86] K.A. Hyun, G.B. Koo, H. Han, J. Sohn, W. Choi, S.I. Kim, H.I. Jung, Y.S. Kim, Epithelial-to-mesenchymal transition leads to loss of EpCAM and different physical properties in circulating tumor cells from metastatic breast cancer, Oncotarget, 7 (2016) 24677-24687.
[87] J. Gao, Q. Yan, J. Wang, S. Liu, X. Yang, Epithelial-to-mesenchymal transition induced by TGF-beta1 is mediated by AP1-dependent EpCAM expression in MCF-7 cells, J Cell Physiol, 230 (2015) 775-782.
[88] M.K. Jolly, S.A. Mani, H. Levine, Hybrid epithelial/mesenchymal phenotype(s): The 'fittest' for metastasis?, Biochim Biophys Acta Rev Cancer, 1870 (2018) 151-157.
[89] I. Pastushenko, C. Blanpain, EMT Transition States during Tumor Progression and Metastasis, Trends Cell Biol, 29 (2019) 212-226.
[90] K. Campbell, J. Casanova, A common framework for EMT and collective cell migration, Development, 143 (2016) 4291-4300.
[91] P. Friedl, J. Locker, E. Sahai, J.E. Segall, Classifying collective cancer cell invasion, Nature cell biology, 14 (2012) 777-783.
[92] L. Cayrefourcq, T. Mazard, S. Joosse, J. Solassol, J. Ramos, E. Assenat, U. Schumacher, V. Costes, T. Maudelonde, K. Pantel, C. Alix-Panabieres, Establishment and characterization of a cell line from human circulating colon cancer cells, Cancer Res, 75 (2015) 892-901.
[93] A. Soler, L. Cayrefourcq, T. Mazard, A. Babayan, P.J. Lamy, S. Assou, E. Assenat, K. Pantel, C. Alix-Panabieres, Autologous cell lines from circulating colon cancer cells captured from sequential liquid biopsies as model to study therapy-driven tumor changes, Sci Rep, 8 (2018) 15931.
[94] A.R. McDougall, S.B. Hooper, V.A. Zahra, T.J. Cole, C.Y. Lo, T. Doran, M.J. Wallace, Trop2 regulates motility and lamellipodia formation in cultured fetal lung fibroblasts, Am J Physiol Lung Cell Mol Physiol, 305 (2013) L508-521.
[95] W.A. Osta, Y. Chen, K. Mikhitarian, M. Mitas, M. Salem, Y.A. Hannun, D.J. Cole, W.E. Gillanders, EpCAM Is Overexpressed in Breast Cancer and Is a Potential Target for Breast Cancer Gene Therapy, Cancer Res, 64 (2004) 5818-5824.

[96] M. Trebak, G.E. Begg, J.M. Chong, E.V. Kanazireva, D. Herlyn, D.W. Speicher, Oligomeric state of the colon carcinoma-associated glycoprotein GA733-2 (Ep-CAM/EGP40) and its role in GA733-mediated homotypic cell–cell adhesion., J. Biol. Chem., 276 (2001) 2299–2309.
[97] F. Thuma, M. Zoller, EpCAM-associated claudin-7 supports lymphatic spread and drug resistance in rat pancreatic cancer, Int J Cancer, 133 (2013) 855-866.
[98] M.R. Gaiser, T. Lammermann, X. Feng, B.Z. Igyarto, D.H. Kaplan, L. Tessarollo, R.N. Germain, M.C. Udey, Cancer-associated epithelial cell adhesion molecule (EpCAM; CD326) enables epidermal Langerhans cell motility and migration in vivo, Proc Natl Acad Sci U S A, 109 (2012) E889-897.
[99] A. Martowicz, J. Rainer, J. Lelong, G. Spizzo, G. Gastl, G. Untergasser, EpCAM overexpression prolongs proliferative capacity of primary human breast epithelial cells and supports hyperplastic growth, Mol Cancer, 12 (2013) 56.
[100] N.V. Sankpal, M.W. Willman, T.P. Fleming, J.D. Mayfield, W.E. Gillanders, Transcriptional repression of epithelial cell adhesion molecule contributes to p53 control of breast cancer invasion, Cancer Res, 69 (2009) 753-757.
[101] M.W. Chen, K.T. Hua, H.J. Kao, C.C. Chi, L.H. Wei, G. Johansson, S.G. Shiah, P.S. Chen, Y.M. Jeng, T.Y. Cheng, T.C. Lai, J.S. Chang, Y.H. Jan, M.H. Chien, C.J. Yang, M.S. Huang, M. Hsiao, M.L. Kuo, H3K9 histone methyltransferase G9a promotes lung cancer invasion and metastasis by silencing the cell adhesion molecule Ep-CAM, Cancer Res, 70 (2010) 7830-7840.
[102] Q. Fan, J.C. Cheng, X. Qiu, H.M. Chang, P.C. Leung, EpCAM is up-regulated by EGF via ERK1/2 signaling and suppresses human epithelial ovarian cancer cell migration, Biochem Biophys Res Commun, 457 (2015) 256-261.
[103] J. Zhang, H. Ma, L. Yang, H. Yang, Z. He, Silencing of the TROP2 gene suppresses proliferation and invasion of hepatocellular carcinoma HepG2 cells, J Int Med Res, 47 (2019) 1319-1329.
[104] T. Liu, Y. Liu, X. Bao, J. Tian, Y. Liu, X. Yang, Overexpression of TROP2 predicts poor prognosis of patients with cervical cancer and promotes the proliferation and invasion of cervical cancer cells by regulating ERK signaling pathway, PLoS One, 8 (2013) e75864.
[105] J. Wang, R. Day, Y. Dong, S.J. Weintraub, L. Michel, Identification of Trop-2 as an oncogene and an attractive therapeutic target in colon cancers, Mol Cancer Ther, 7 (2008) 280-285.
[106] P. Zhao, Z. Zhang, TNF-alpha promotes colon cancer cell migration and invasion by upregulating TROP-2, Oncol Lett, 15 (2018) 3820-3827.
[107] R. Cubas, S. Zhang, M. Li, C. Chen, Q. Yao, Trop2 expression contributes to tumor pathogenesis by activating the ERK MAPK pathway, Mol Cancer, 9 (2010) 253.
[108] X. Li, S. Teng, Y. Zhang, W. Zhang, X. Zhang, K. Xu, H. Yao, J. Yao, H. Wang, X. Liang, Z. Hu, TROP2 promotes proliferation, migration and metastasis of gallbladder cancer cells by regulating PI3K/AKT pathway and inducing EMT, Oncotarget, 8 (2017) 47052-47063.
[109] Q.Z. Gu, A. Nijiati, X. Gao, K.L. Tao, C.D. Li, X.P. Fan, Z. Tian, TROP2 promotes cell proliferation and migration in osteosarcoma through PI3K/AKT signaling, Mol Med Rep, 18 (2018) 1782-1788.
[110] X. Sun, G. Xing, C. Zhang, K. Lu, Y. Wang, X. He, Knockdown of Trop2 inhibits proliferation and migration and induces apoptosis of endometrial cancer cells via AKT/beta-catenin pathway, Cell Biochem Funct, 38 (2020) 141-148.
[111] H. Guan, Z. Guo, W. Liang, H. Li, G. Wei, L. Xu, H. Xiao, Y. Li, Trop2 enhances invasion of thyroid cancer by inducing MMP2 through ERK and JNK pathways, BMC Cancer, 17 (2017) 486.
[112] M. Munz, C. Kieu, B. Mack, B. Schmitt, R. Zeidler, O. Gires, The carcinoma-associated antigen EpCAM upregulates c-myc and induces cell proliferation, Oncogene, 23 (2004) 5748-5758.
[113] A. Martowicz, G. Spizzo, G. Gastl, G. Untergasser, Phenotype-dependent effects of EpCAM expression on growth and invasion of human breast cancer cell lines, BMC Cancer, 12 (2012) 501.
[114] P. Storz, Protein kinase D1: gatekeeper of the epithelial phenotype and key regulator of cancer metastasis?, Br J Cancer, 118 (2018) 459-461.

Table 1. Experimental evidence for an impact of EpCAM on cell-cell adhesion:

Abbreviations: Cell/tissue type: 1°, primary; Ad, Adenocarcinoma; Br, breast; BrC, breast cancer; Ca, carcinoma; CrC, colorectal cancer; EndC, Endometrial cancer; Ep, epithelial; OvC, ovarian cancer; ScC, Squamous Cell Carcinoma
Assays: Aggreg, aggregation assay; EP, Embryonic phenotype; insol, insolubility; Mph, cell morphology; TI, Tissue integrity
Proposed adhesion function: +, proadhesive; -, anti-adhesive; o, no effect.
Other: Adh, adhesion; Cad, cadherin; ctn, catenin; dst, downstream; rec, recombinant

Cell/tissue type	Assay	EpCAM GOF/LOF	Phenotype	Proposed adh. fct	Ref
L cells (fibroblasts) L153S (BrC)	Aggreg	Expression	Stimulation	+	[2]
E-Cad-expressing L cells HCA (Br ep)	Aggreg	Expression	Inhibition	-	[56]
Xenopus embryonic cells (multiple tissues)	EP, TI, Mph Cad levels	Depletion	Inhibition	+	[9, 10]
Zebrafish ectoderm	EP, TI, Mph Cad levels	Depletion	Inhibition	+	[8]
Human intestinal biopsies	TI, Mph Cad levels	Mutation	Inhibition	+	[55]
Caco2 (intestinal)	TI, Mph Cad levels	Depletion	Inhibition	+	[55]
Mouse ES cells on fibroblasts	Heterotypic adh ES on fibroblasts	Expression in fibroblasts	No effect	o	[57]
MCF7, MCF10a, MDA-231, 361, 453 (BrC)	Cad/ctn detergent insol α-ctn expression	Depletion	Increased Increased	+	[95]
RC-6 (Br), MCF7, T47D (BrC), LS-180 (CrC), COV-362 (OvC)	Aggreg	Antibody incubation	Inhibition	+	[1]
Caco-2 and Colo-205 (CrC)	Aggreg	Competition by rec EpCAM	Inhibition	+	[96]
ASML pancreatic Ad	Adh on monolayer	Depletion	Inhibition	+	[97]
Kyse-30 (ScC)	C-C adhesion C-Matrix	Depletion	No effect No effect	o	[35]
RL95-2 (EndC)	C-Matrix, AFM	KO	No effect, but increased elasticity	o	[58]
Skin Langerhans Cells	In vivo (to keratinocytes)	Depletion	Decreased	+	[98]

Table 2. Experimental evidence for an impact of EpCAM on cell migration:

Abbreviations: Cell/tissue type: 1°, primary; Ad, Adenocarcinoma; Br, breast; BrC, breast cancer; Ca, carcinoma; CrC, colorectal cancer; CvC, cervical cancer; EndC, Endometrial cancer; Ep, epithelial; HepC, hepatocellular carcinoma; OvC, ovarian cancer; PrC, prostate cancer; ScC, Squamous Cell Carcinoma
Assays: CellM, cellular markers; EP, Embryonic phenotype; ICM, intercellular migration; Mph, cell morphology; SCM, single cell migration; TrW, transwell; WH, wound healing
Proposed migratory function: +, promigratory; -, antimigratory; o, no effect.
Other: Adh, adhesion; cldn7, claudin-7; dst, downstream; FA, focal adhesions

Cell/tissue type	Assay	EpCAM GOF/LOF	Phenotype	Proposed migr. fct	Proposed mechanism	Ref
Xenopus embryo tissues	EP, ICM, Mph	Overexpression Depletion	Increased Decreased	+	Inhibition of anti-migratory PKC-myosin	[9]
Zebrafish lateral line	EP	Depletion	No effect	o		[79]
1° human Br ep	TrW	Overexpression	Decreased	-		[99]
MDCK (kidney ep)	WH	Overexpression Depletion	Decreased Increased	-	promigratory Erk-myosin activity	[46]
HEK293 (kidney ep)	WH	Co-expression EpCAM + cldn7	Increased	+		[6]
MCF7 (BrC ep)	WH	Overexpression Depletion	Increased Decreased	+	EMT-mediated Required dst of TGFβ	[87]
MCF7, MCF10a, MDA-231, 361, 453 (BrC)	TrW	Depletion	Decreased	+		[95]
MCF10A (+EGF) (Br ep)	TrW	Depletion	Increased	-	Erk inhibition	[67]
MDA-MD-231 (BrC)	TrW	Depletion	Decreased	+	Stimulation of NFkB and IL8	[59]
MCF10CA1a (BrC) HCT116 (CrC)	TrW	Depletion	Decreased	+		[100]
ASML (Pancreatic Ad)	TrW + WH	Depletion	Increased	-	Antagonized by cldn7 association	[97]
Kyse-30 (ScC)	WH	Depletion	Increased	-	Through EMT	[34]
Fadu, Kyse30 (ScC)	WH	Addition of EpEX	Decreased	-	Inhibition of EGF-dep EMT	[66]
CW-2 (CrC) A431 (Epidermoid Ca)	TrW	Depletion	Decreased	+	Loss of matrix adh Erk/Akt/FAK inhibition	[51]
CL1-5 (Lung Ad)	TrW	Depletion	Decreased	+		[101]
SKOV3 (OvC)	TrW	Overexpression	Decreased	-		[102]
OVCAR4 (OvC)	TrW	Depletion	Increased	+		[102]
EpCAM2/Trop2				+		
fetal lung fibroblasts	WH SCM Mph, CellM	Depletion Overexpression	Decreased Increased LOF defects: Mph, FA	+	Erk activation	[94]
Immortalized keratinocytes	WH	Mouse KO	Increased	-	Erk and Src activation, no EMT	[78]
PC3-2, DU145 (PrC)	TrW +SCM	Depletion Overexpression	Decreased Increased	+	Interaction w integrins, FAK relocalization	[49, 50]
HepG2 (HepC)	TrW + WH	Depletion	Decreased	+		[103]
Siha, CaSki , Hela, C33A (CvC)	TrW + WH	Depletion Overexpression	Increased	+	Erk activation, EMT	[104]
BE (CrC)	TrW	Overexpression	Increased	+		[105]
HCT-116 (CrC)	TrW + WH	Depletion	Decreased	+	Trop2 upregulation by TNFa via Erk	[106]
Mouse pancreatic Ca	WH	Overexpression	Increased	+	Erk activation	[107]
GBC-SD, SGC-996 (Gallbladder cancer)	TrW	Overexpression Depletion	Increased Decreased	+	PI3K/AKT	[108]
MG63, MNNG/HOS (osteosarcoma)	TrW + WH	Depletion Overexpression	Decreased Increased	+	PI3K/AKT	[109]
KLE, Ishikaw (EndC)	TrW	Depletion	Decreased	+	EMT and AKT/β-ctn	[110]
K1, FTC-133, 8505C (thyroid cancer)	TrW + WH	Overexpression Depletion	Increased Depletion	+	ERK1/2 and JNK	[111]
Skin Langerhans Cells	In vivo	Depletion	Decreased	+	Decreased adhesion to keratinocytes	[98]