The FEBS Journal

FEBS PRESS
science publishing by scientists

VIEWPOINT

CEMIP (HYBID, KIAA1199): structure, function and expression in health and disease

Sofia Spataro[1] (iD), Concetta Guerra[2] (iD), Andrea Cavalli[2,3] (iD), Jacopo Sgrignani[2] (iD), Jonathan Sleeman[4,5] (iD), Lina Poulain[6] (iD), Andreas Boland[6] (iD), Leonardo Scapozza[1] (iD), Solange Moll[7] and Marco Prunotto[1] (iD)

1 Institute of Pharmaceutical Sciences of Western Switzerland, University of Geneva, Switzerland
2 Institute for Research in Biomedicine, Università della Svizzera Italiana, Bellinzona, Switzerland
3 Swiss Institute of Bioinformatics, Lausanne, Switzerland
4 European Center for Angioscience (ECAS), Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany
5 Institute for Biological and Chemical Systems – Biological Information Processing (IBCS – BIP), Karlsruhe Institute for Technology (KIT), Germany
6 Department of Molecular Biology, University of Geneva, Switzerland
7 Department of Pathology, University Hospital of Geneva, Switzerland

CEMIP (cell migration-inducing protein), also known as KIAA1199 or HYBID, is a protein involved in the depolymerisation of hyaluronic acid (HA), a major glycosaminoglycan component of the extracellular matrix. CEMIP was originally described in patients affected by nonsyndromic hearing loss and has subsequently been shown to play a key role in tumour initiation and progression, as well as arthritis, atherosclerosis and idiopathic pulmonary fibrosis. Despite the vast literature associating CEMIP with these diseases, its biology remains elusive. The present review article summarises all the major scientific evidence regarding its structure, function, role and expression, and attempts to cast light on a protein that modulates EMT, fibrosis and tissue inflammation, an unmet key aspect in several inflammatory disease conditions.

Introduction

Hyaluronic acid (HA), a major glycosaminoglycan component of the extracellular matrix (ECM), is synthesised as a free linear polymer by three hyaluronan synthases (HAS) and is degraded by hyaluronidases (HYAL), cell migration-inducing protein (CEMIP) and transmembrane protein 2 (TMEM2) into fragments of different

Abbreviations
ANXA1, annexin A1; CEMIP, cell migration-inducing protein; ECM, extracellular matrix; EMT, epithelial–mesenchymal transition; GAGs, glycosaminoglycans; HA, hyaluronic acid; HAS, hyaluronan synthases; HMW, high-molecular weight; HYAL, hyaluronidases; LMW, low-molecular weight; PGs, proteoglycans; ROS, reactive oxygen species; TMEM2, transmembrane protein 2.

sizes. As a ubiquitous component of the human ECM microenvironment where it exerts a structural role, HA also regulates several biological processes including cell signalling, wound reparation, tissue regeneration and morphogenesis, inflammation, cell migration, mitosis, cancer pathogenesis and metastasis. HA-mediated biological responses are dependent upon the size and concentration of HA, its interactions with key HA-binding proteins and cell-associated receptors, and its cell context-specific signalling.

Here, we provide an overview of CEMIP, a large protein of around 153 kDa that depolymerises HA into small- and intermediate-sized fragments, and describe its structure, function as a hyaluronidase, putative role in HA-mediated biological functions, regulation by positive and negative pathways, expression in tissues, role in tumour initiation and progression, and its putative contribution to other diseases and ageing.

HA as a major component of the ECM

HA is a major glycosaminoglycan component of the interstitial ECM. The interstitial ECM is a complex molecular structure that surrounds cells in tissues that provide structural support and regulate cell behaviour in both physiological and pathological conditions [1]. The interstitial ECM is composed of two main classes of macromolecules: viscous (energy absorber) proteoglycans (PGs) and glycosaminoglycans (GAGs), such as HA, and elastic (energy storage) fibrous proteins, such as collagens, elastins, fibronectins and laminins [2]. These two classes of macromolecules confer upon the ECM, its viscoelasticity and architecture. Even though ECM is found in all types of tissues in adults, it is important to note that the composition differs depending on the tissue type. Viscous PGs fill the majority of the interstitial space within tissues in the form of a hydrated gel, for example in the brain and pancreas, whereas elastic fibrous components provide tensile strength, for example in muscle and bone [2]. HA belongs to a group of heteropolysaccharides named glycosaminoglycans (GAGs), which also includes chondroitin sulfate, dermatan sulfate, keratin sulfate, heparin sulfate and heparin [3]. HA is the most common GAG found in the tissues of humans and other vertebrates [4] and is comprised of tandem repeats of glucuronic acid and N-acetylglucosamine. It differs from other GAGs in that it is not sulfated and not synthesised by Golgi enzymes in association with proteins, but is rather produced at the inner face of the plasma membrane without any covalent bond to a protein core [3,5]. Nevertheless, HA can be covalently tethered to the serum protein inter-alpha-inhibitor by TSG6 [6]. In addition, HA can reach a very

high-molecular weight (HMW, 10^8 Da), in contrast to other GAGs that are smaller in size ($< 5 \times 10^4$ Da, usually $1.5–2 \times 10^4$ Da).

In mammals, HA is synthesised as a free linear polymer by three hyaluronan synthases named HAS1, HAS2 and HAS3. HAS are transmembrane glycosyltransferase isoenzymes whose catalytic sites are located on the inner face of the plasma membrane. Thus, growing HA chains are extruded onto the cell surface or into the ECM through the plasma membrane and HAS protein complexes [3]. HAS isoforms expression and activity are controlled differently by growth factors, cytokines and other proteins such as kinases, in a manner which appears to be cell and tissue specific. Similarly, the biochemical and synthetic properties of HAS isoforms are distinct:

- HAS1 is the least active isoenzyme and produces HMW hyaluronan from 2×10^5 to 2×10^6 Da.
- HAS2, which represents the main hyaluronan synthetic enzyme in normal adult cells, is more active than HAS1 and synthesises HA chains $> 2 \times 10^6$ Da. It regulates the developmental and reparation processes of tissue growth, and may be involved in inflammation, cancer, pulmonary fibrosis and keloid scarring [7–11].
- HAS3 is the most active isoenzyme and produces HA molecules with a MW lower than 3×10^5 Da.

HA is degraded within hours of its synthesis [12]. Extra-large (1000–10 000 kDa) native HA molecules within tissues are initially depolymerised into fragments of 10–100 kDa [13]. Most fragments are then released from the ECM, drained into lymphatic vessels and catabolised within the lymph nodes. The remaining HA fragments reach the circulation and are fully degraded within the liver. HA degradation can be mediated by HYALs—enzymes that are classified as endoglycosidases [14,15]—and through oxidative damage mediated by reactive oxygen species (ROS) [3]. With regard to the former, six HYAL gene sequences have been identified in the human genome: *HYAL-1*, *HYAL-2*, *HYAL-3*, *HYAL-4*, *PH20/SPAM1* and *HYAL-P1* [16]. However, because *HYAL-P1* is a pseudogene and both *HYAL-3* and *HYAL-4* have restricted expression patterns, these three genes are unlikely to have major roles in constitutive HA degradation *in vivo* [13]. HYAL-1, HYAL-2 and PH20/SPAM1 are the most characterised human HYAL genes [3]. HYAL-1 and HYAL-2, which are highly expressed in human somatic tissues, are acid-active with a pH optimum of 3.5 and below 4, respectively [15], and produce HA fragments of different sizes. HYAL-1 can degrade HA of any size into small hexa- and tetrasaccharides, while HYAL-2 processes HMW

HA into 20 kDa oligomers in conjunction with the cell surface HA receptor CD44 [15]. HYAL-1, which was found to regulate cell cycle progression and apoptosis, is the main HYAL expressed in cancers, and therefore, it may regulate tumour growth and angiogenesis [12]. In contrast to HYAL-1 and HYAL-2, PH20/SPAM1 shows endoglycosidase activity at both an acidic and neutral pH and behaves as a multifunctional enzyme [3].

In addition to CEMIP and the HYAL family molecules, TMEM2 was recently reported to be a cell surface transmembrane hyaluronidase in mouse organs that shares a high degree of sequence similarity with CEMIP [17]. From a structural point of view, TMEM2 has a transmembrane domain, one G8 domain, one GG domain and 3 PbH1 repeats, whereas CEMIP contains, as described below, a signal sequence, one G8 domain, two GG domains and 4 PbH1 repeats. While previous studies have indicated that the second GG domain is only present in CEMIP [17], a structural alignment of the C-terminal portion of TMEM2 (AA 1209–1355) and CEMIP (AA 1209–1353) underlines the possible presence of a second GG domain in TMEM2 (Fig. 1). Despite their structural similarities, the current (but limited) knowledge about these two proteins suggests several differences in both function and regulation. TMEM2 is a type II transmembrane protein, which traffics to the cell membrane and catabolises high-molecular-weight HA into small fragments of ~ 5 kDa [17]. Its activity is calcium dependent and requires a pH optimum between 6 and 7 [17]. In contrast to CEMIP, TMEM2-mediated HA

degradation does not rely on the participation of live cells [17]. Studies in mice also showed that TMEM2 is more highly expressed at the transcriptional level than CEMIP, both during development and in adult tissues [17].

HA degradation by the CEMIP hyaluronidase activity

CEMIP mediates HA depolymerisation [18,19] in a manner that involves clathrin-coated vesicles, internalisation into early endosomes and excretion of degraded HA molecules into the extracellular space [13]. Consistently, knock-down of clathrin heavy chain (CHC) and α-adaptin subunit of AP-2 impairs CEMIP hyaluronidase activity [13]. Additionally, deletion of the N-terminal signal peptide (amino acids 1–30) affects CEMIP trafficking and impairs its HA degradation activity [19]. Recent studies show that CEMIP-mediated HA depolymerisation requires the interaction with the Annexin-1 via its G8 domain, allowing CEMIP and Annexin-1 to colocalise at the cell membrane [20].

In an attempt to correlate structure and function, and based on the CEMIP model structure discussed below, it might be postulated that:

• the G8 domain is essential for HA catalytic activity and potentially interacts with multiple protein partners [20,21], some of them probably still to be discovered.
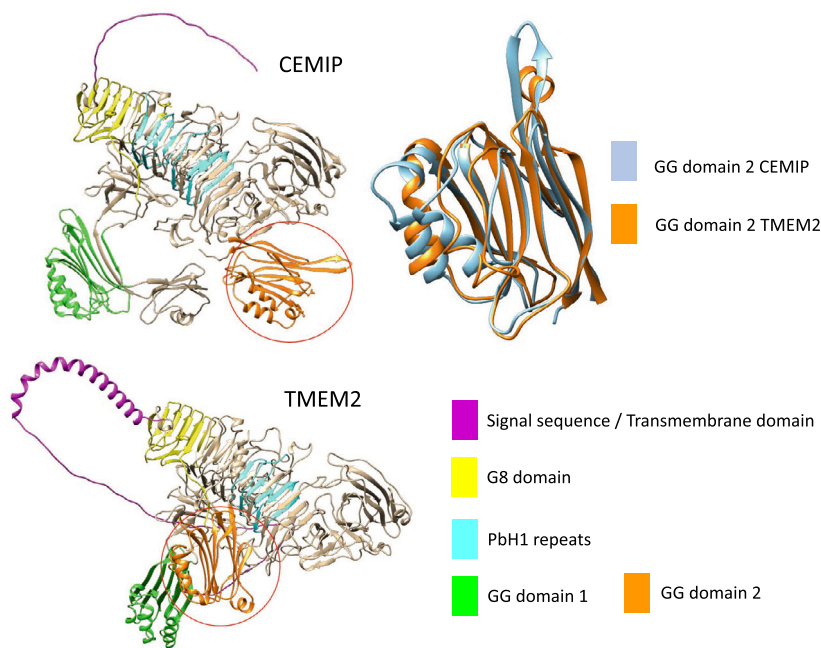
Fig. 1. Structure of CEMIP and TMEM2 with their domains (left, top and bottom). Structural alignment of AA1209-1353 (CEMIP) and AA1209-1355 (TMEM2) reveals the presence of a second GG domain in TMEM2 (right, top). From AlphaFold, modified with USCF chimera.

- the two GG domains might provide a key 'hinge' function, comparable to the XYZ domain in protein gp35 in bacteriophages. This domain is essential for low-molecular-weight-HA binding (or protein catalytic function) as suggested by mutation in position 187, and potentially also relevant in CEMIP inactivation.
- the size of CEMIP-produced HA fragments might be dependent on the integrity of what seems to be a long tunnel formed by the four PbH1 domains.

CEMIP structure: a large protein with yet unsolved atomistic structure

CEMIP is composed of 1361 amino acids. Its atomistic structure remains to date unsolved, essentially due to difficulties in expression and purification of the protein. However, as pointed out by Yamaguchi [17] and Yoshida et al. [18], sequence analyses provide a first overview of its domain architecture. CEMIP is characterised by the presence of an N-terminal signal sequence of 30 amino acids, which directs the protein to the secretion pathway. Additionally, CEMIP consists of one G8 domain, two GG domains, four PbH1 (parallel β-helix repeats) domains and several N-linked glycosylation sites identified throughout the entire sequence (Fig. 2A) [22].

The G8 domain, identified by Lang et al. in 2006 [23], is defined by the presence of eight conserved glycine residues. The domain is thought to form five consecutive β-strand pairs, although atomic coordinates of this type of domain have not yet been deposited in the protein data bank (PDB) [24]. G8 domains exist in several proteins from lower eukaryotes to animals, but are absent in plants, viruses and archaea. Notably, G8 domains have been identified in disease-related human proteins such as the polycystic kidney and hepatic disease 1 (PKHD1) protein, also known as
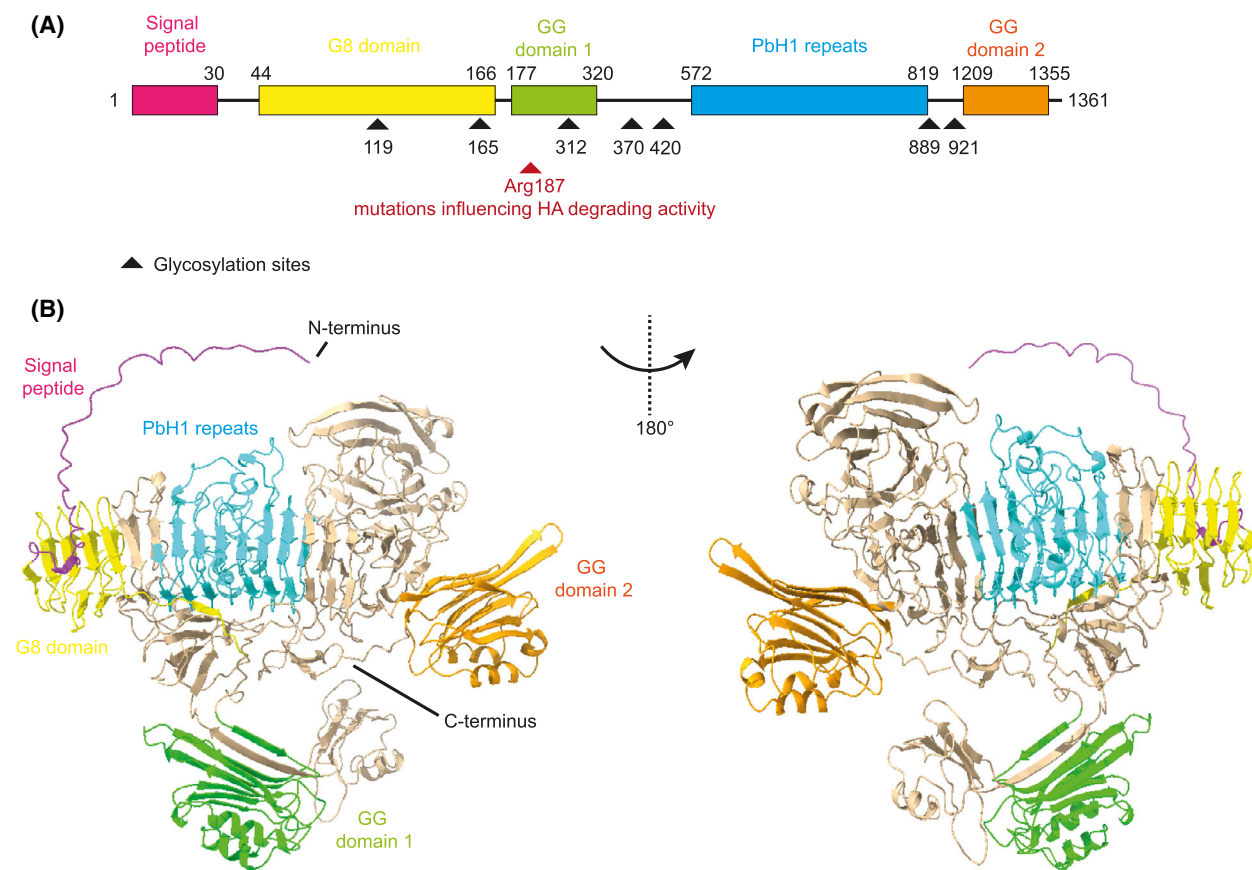
Fig. 2. CEMIP structure. (A) Schematic representation of the CEMIP secondary structure, annotated with mutation and glycosylation sites. Two examples, taken from the PDB of GG domain and PbH1 domains are shown as a cartoon for explicative purposes. (B) Atomistic model. From AlphaFold, modified with USCF chimera.

fibrocystin and polyductin, and TMEM2, also known as hyaluronoglucosaminidase [23].

The GG domain has been identified in members of the eukaryotic FAM3 superfamily (FAM3A, FAM3B, FAM3C and FAM3D), POMGnT1 (protein O-linked mannose β-1,2-N-acetylglucosaminyltransferase), TMEM2 and the phage gp35 protein [25]. The structure of POMGnT1 has been solved and deposited in the PDB [26]. The role of the GG domain is not yet fully understood. It might be involved in HA binding and/or degradation since a single mutation (R187H or R187C) in the first GG domain hinders the HA degrading activity [13,27] of CEMIP.

The PbH1 domain can be found in several polysaccharide degrading enzymes [28,29] and also in fibrocystin, a protein involved in the development of the autosomal recessive polycystic kidney disease [30]. However, the functional role of the PbH1 domain in CEMIP remains elusive.

As mentioned above, the atomic coordinates of CEMIP remain unsolved to date. Given the relevance of the three-dimensional structure of CEMIP for the design of therapeutics, we used homology modelling techniques [31] to generate a CEMIP model. To this end, we leveraged proteins already deposited in the PDB and ran a template search using the Swissmodel web-server [26]. The search indicated that a template is only available for the first GG domain (i.e. protein O-linked-mannose beta-1,2-N-acetylglucosaminyl transferase 1, PDB code 5GGP [32], percentage of identity 30%) [31]. Other regions of the protein display only a low homology with proteins of known structure (< 15–20%), which is insufficient to allow the generation of reliable models.

During the preparation of this review, Google Deepmind and the group of David Baker released two software packages, alphafold2 [33] and RoseTTAFold [34]. Both programs employ deep learning algorithms that enable prediction of protein structures at an unprecedented accuracy, even in the absence of suitable templates. We therefore analysed the recently released AI-based postulated CEMIP structure available in the alphafold2 database maintained by the EMBL (https://alphafold.ebi.ac.uk). Although deposited models are still unreliable concerning high-resolution information such as hydrogen bonds, side chain orientations or conformation of putative binding sites, these models are extremely valuable for interpreting and designing new experiments. Interestingly, the predicted CEMIP structure by alphafold2 (Fig. 2B) confirms the presence of a G8 domain (formed by 5 β-strand pairs) located immediately after the end of the signal peptide (Ala33 to Ser159) and followed by a

GG domain (Glu175 to Gln321). The model is then characterised by the presence of two small subdomains not previously identified that encompass residues Asp322 to Pro406 and Asn418 to Leu494. In the region from Ile499 to Ser916, the expected PbH1 domain is found, followed by a previously unreported domain formed by residues from Cys917 to Asp1208. The structural model further confirms the presence of the second GG domain in the C-terminal region of the protein (His1209 to Leu1361).

CEMIP: putative role in HA-mediated biological functions

As mentioned above, in addition to its structural role, HA triggers different cell signalling cascades and modulates many physiological and pathological processes [1]. HA fragment size is one of the major determinants of its activity [35]. HA is rapidly depolymerised within tissues, from extra-large native molecules of 1000–10 000 kDa to HA fragments which consist of low-molecular-weight HA (LMW HA, 10–100 kDa) and much smaller HA oligosaccharides (8–30-dimer lengths). HMW-HA and LMW-HA can have antagonistic or inverse effects [36]. For example, HMW-HA can act as an anti-inflammatory molecules that promotes epithelial cell homeostasis and survival and accelerates wound healing [35], while LMW-HA has been shown to trigger inflammatory response [35], to promote activation of dendritic cells and monocyte maturation into macrophages [35] and to induce angiogenesis [37] and lymphangiogenesis [38].

In most cases, short oligosaccharides are produced by enzymatic cleavage mediated by hyaluronidases. However, it is interesting to note that LMW HA can also be directly synthesised by dysregulation of HAS enzymes, as well as by an impaired cellular metabolism that alters precursor availability [1]. Thus, it has been shown that phosphorylation of HAS2, the main hyaluronan synthetic enzyme which synthesises HMW HA, leads to the inhibition of HA secretion [1]: since HA production by HAS is an energy consuming process, a low ATP:AMP ratio, as observed in an altered cellular metabolism, activates AMPK (AMP activated protein kinase) which mediates HAS2 phosphorylation and inhibits HA secretion [1].

CEMIP has been reported to degrade HMW HA into both intermediate-sized fragments of between 35 and 50 kDa (int-HA fragments) as well as LMW-HA [5,13,39]. CEMIP-produced int-HA fragments might conceivably have specific function(s) compared with other sizes of HA. Interestingly, while small HA oligosaccharides have been intensively investigated and

have been shown to exert a number of biological effects not observed with HMW-HA [37,38], the biological activity of int-HA assembly largely remains to be investigated. Int-HA could potentially act as signalling molecules, for example in inflammation. It has been demonstrated that LMW and Int-HA fragments activate dendritic cells and macrophages via CD44 or TLR [40]. Int-HA fragments have also been reported to promote monocyte polarisation, although with discordant results, with one study reporting polarisation into M2-like macrophages [40], while the other into the M1 phenotype [41,42]. A possible explanation for these different *in vitro* results might be due to different experimental conditions. Int-HA fragments have also been demonstrated to induce wound closure, contrary to HMW-HA and LMW-HA [43], suggesting that int-HA fragments might induce a microenvironment that promotes tumorigenesis, metastasis and modulating fibrosis [44].

Regulation of CEMIP protein expression

CEMIP protein levels are regulated by cytokines, transcription factors, the Wnt/β-catenin signalling pathway, microRNAs (miRNAs), histone methylation and hypoxia (Fig. 3). However, it should be noted that the results published by different studies are often conflicting.

Several cytokines have been found to play a major role in the regulation of CEMIP. In 2017, Khoi *et al.* [45] suggested that the pro-inflammatory interleukin 1β (IL-1β) increased CEMIP transcription and migration of pancreatic ductal adenocarcinoma cells. De la Motte's group investigated the effect of interleukin 6 (IL-6) in Crohn's disease fibroblasts. The authors observed low levels of CEMIP deposition in the extracellular matrix after treatment with IL-6 [39]. However, contrasting evidence has been reported in a recent study by Sato *et al.* who demonstrated that treatment with a mixture of pro-inflammatory cytokines (TNF-α, IL-1β and IL-6) in human skin fibroblasts decreased CEMIP mRNA levels and protein expression, an effect primarily mediated by IL-1β [46].

TGF-β, a pro-fibrotic protein and inducer of cell proliferation, can regulate CEMIP expression. Deroyer *et al.* [47] identified TGF-β as a CEMIP upregulator through the Alk5/PAI-1 pathway in dedifferentiated chondrocytes and therefore as a pro-fibrotic mediator. However, another TGF-β-initiated pathway was found by Shintaro Inoue's group, in this case leading to an opposite effect in CEMIP regulation. This group investigated the TGF-β1-mediated pathway in Detroit 551 skin fibroblasts and demonstrated that TGF-β1 downstream of PI3K-Akt signalling pathway was implicated in CEMIP downregulation after TGF-β1 stimulation [48]. Consistently, we have also recently reported that TGF-β1 downregulates CEMIP expression in
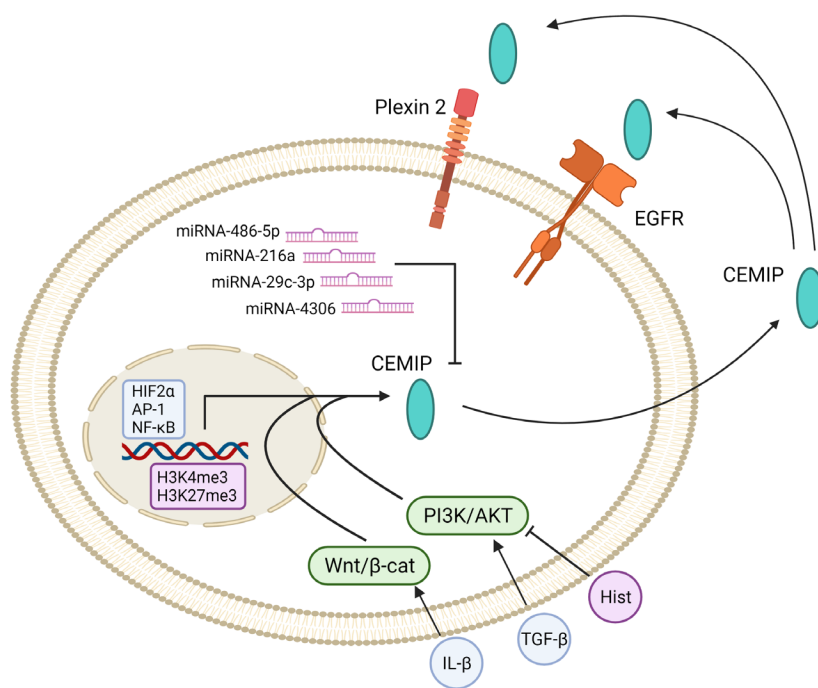
Fig. 3. Schematic representation of regulators of CEMIP protein expression.

fibroblasts [49]. The role of TGF-β in the epithelial–mesenchymal transition (EMT) will be discussed in the next section.

The coordinate roles of TMEM2 and CEMIP expressions have been analysed in human skin fibroblasts. Thus, CEMIP and TMEM2 expressions together with HA depolymerisation appeared to be tightly regulated by the action of TGF-β1 and histamine [50]: while TGF-β1 enhanced TMEM2 levels and decreased both CEMIP and HA processing, histamine was demonstrated to act in the opposite direction [50]. In the same cells, knock-down of CEMIP, but not of TMEM2, heavily impaired HA depolymerisation [50]. In addition, pro-inflammatory cytokines were shown to significantly regulate both protein levels [46]: indeed, treatment of human skin fibroblasts with IL-β1 increased TMEM2 expression, but suppressed both CEMIP and HA depolymerisation [46]. Even though CEMIP and TMEM2 are both involved in HA processing, these data clearly indicate that they act separately in different contexts and that their activity is tightly regulated at the transcriptional level through the intervention of several macromolecules.

A number of transcription factors have been shown to control CEMIP transcription. In 2014, Shostak et al. [21] identified CEMIP as an oncogene whose expression is controlled by nuclear factor-kappa B (NF-kB), which protects tumour cells from apoptosis and promotes invasiveness and the epithelial–mesenchymal transition through the EGFR pathway. Kuscu et al. [51] reported that distal NF-kB sites and the transcription factor AP-1 are required for CEMIP gene expression in fibroblast-like monkey cell lines.

Several microRNAs have been shown to negatively regulate CEMIP expression. In 2019, Jiao et al. [52] showed that miR-486-5p directly targets the 3′-UTR region of CEMIP, and exerts a significant inhibition of CEMIP mRNA and protein expression in papillary thyroid cancer cells using a luciferase reporter assay. Wang et al. [53] reported that the same miR-486-5p targets CEMIP and that its overexpression reduces the proliferation and migration of non-small-cell lung cancer cells via the EGFR pathway. Overexpression of miR-29c-3p, another miRNA that targets CEMIP, has been shown to decrease gastric cancer cells migration in vitro and in vivo via the Wnt/β-catenin and EGFR signalling pathways [54]. Similarly, overexpression of miR-216a can reduce migration and invasion of colorectal cancer cells in vitro and inhibit metastasis in vivo [55]. miR-4306 was demonstrated to regulate CEMIP expression in osteosarcoma, and LINC00958, a long noncoding RNA (lncRNA), was shown to promote tumour progression and metastasis by inhibiting

miR-4306 expression, leading to increased CEMIP expression [56].

Histone methylation can positively or negatively modulate gene transcription. Increased tri-methylation of lysine 27 on histone H3 (H3K27me3) has been associated with inactivation of CEMIP, decreased tumour cell growth and reduced migration in triple-negative breast cancer [57]. In another study, a correlation between H3K4me3 levels and CEMIP expression was demonstrated in human colon cancer cells [58]. Hypoxic conditions (see below) inhibit the histone demethylase Jarid1A, causing an increase of H3K4me3 within the CEMIP promoter, which leads to CEMIP upregulation [58].

Hypoxic conditions present in the tumour microenvironment modulate CEMIP expression. In 2015, Evensen et al. [58] observed that hypoxia-inducible factor 2α (HIF-2α) induces CEMIP expression by binding to the hypoxia response element (HRE) present in the CEMIP promoter region. Moreover, Wang et al. [59] reported that in HCC patient samples, co-expression of CEMIP and HIF-1α correlated with a poor prognosis and overall survival rate [59].

CEMIP expression: tissue and cell localisation

CEMIP is expressed in a wide range of tissues, with the highest levels in brain, skin, placenta, lung, testis and ovary, but its expression is notably absent in the liver, kidney and spleen [60]. CEMIP is widely expressed in the brain, in contrast to the HYAL1 and HYAL2 enzymes, which are not expressed in this organ [61]. Therefore, it is likely that CEMIP may have an important role in HA catabolism in tissues that do not express HYAL1 or HYAL2, such as the brain. Indeed, it is noteworthy that ECM molecules fill the extracellular space, which occupies up to 20% of the adult brain volume [62]. HA is important components of the brain ECM in both condensed ECM (such as perineuronal nets) and diffuse ECM (which fills perisynaptic spaces), and plays crucial roles in neuronal development, plasticity and pathophysiology [62]. Thus, HA plays a major structural role in ECM throughout all stages of brain maturation, specifically supporting HA-binding proteins and proteoglycans as a central filament of aggregates [63]. Moreover, an interesting study by Wilson et al. [64] highlighted the importance of HA in regulating the formation and function of synapses in a model of human cortical brain development. Numerous studies showed that, from the beginning of brain development, HA critically regulates neural circuit formation. At the onset

of neurulation, HA is necessary for neural crest cell migration. Later in cortical neurogenesis, HA regulates neural progenitor cell proliferation and promotes neuronal differentiation, migration and the formation of cortical layers. In addition, HA surrounds developing excitatory synapses, where it critically regulates synapse formation and the resulting balance between excitatory to inhibitory signalling [64]. Accordingly, HA removal was demonstrated to be sufficient to drive a hyperexcitable state which is characteristic of neurodevelopmental disorders, including epilepsy, intellectual disability and autism spectrum disorders. Conversely, the observation that HA decreases excitatory synapse formation could have implications for ageing and Alzheimer's disease, both of which exhibit HA accumulation and synapse loss [64].

CEMIP was first shown to be widely expressed in the brain by the pioneering work of Izumi Horokawa's team at NIH on induced cell mortality [60]. Yoshino et al. [65] subsequently showed that CEMIP is expressed in the hippocampus and cerebellum in wild-type mice, and demonstrated decreased mnemonic ability in novel object recognition in CEMIP KO mice. In particular, high levels of CEMIP transcripts have been detected in the granular cell layer of the murine hippocampus and cerebellum [65], and also during the immature period of the organ of Corti in mice, in fibrocytes of the spiral ligaments and spiral limbus, and in Deiters' cells [27]. Finally, in another study, Yoshino et al. [66] demonstrated that the dendritic spine density was significantly decreased in the dentate gyrus granule cells in CEMIP KO mice, suggesting that CEMIP-mediated HA degradation may be critical for the synaptic formation process by contributing to cognitive functions, such as learning and memory in the mouse brain. In future work, it will be interesting to investigate further the role of CEMIP in the homeostatic regulation of the HA content of the brain, and the impact this has on brain development, function and disease.

In addition to the brain, CEMIP has been frequently detected in fibroblasts. Yoshida *et al.* [13] reported CEMIP expression in dermal fibroblasts, while other groups have detected CEMIP in synovial [67] and colon fibroblasts [39]. CEMIP is also expressed in hypertrophic chondrocytes at the chondro-osseous junction [68]. We have recently reported the expression of CEMIP in mesenchymal stromal cells and have shown that CEMIP regulates the differentiation of these cells into the osteogenic and adipogenic lineages [49].

Within the cell, CEMIP has been found to localise in different cellular subcompartments, a fact that could explain its involvement in regulating different cellular

pathways and diseases. Cytoplasmic CEMIP has been identified in different cells of the human cochlea [69] and in several tumour cells such as gastric cancer [70] and hepatocellular carcinoma (HCC) [71], where its knock-down triggers ER stress-mediated apoptosis through upregulation of CHOP and ATF4 [71]. CEMIP heterogeneous localisation was observed in colorectal cancer cells, with nuclear [22], cytoplasmic [22,72] and cell membrane [72,73] localisations. Despite this heterogeneous localisation, several lines of evidence suggest a common role in colorectal cancers for CEMIP as modulator of Wnt-β catenin signalling [22,72,73] (see below, CEMIP and EMT). CEMIP has also been found in the ER [19,74] and Golgi [19] compartments, which is prototypical for a secretory protein [19] and consistent with the existence of an N-terminal signal peptide. Interestingly, this localisation pattern has been classically associated with the HA depolymerising function of CEMIP [19]. Recent evidence also links cell membrane localisation of CEMIP to its HA degradation activity [20], as CEMIP can localise to the cell membrane by binding to ANXA1 in rheumatoid arthritis fibroblast-like synoviocytes, which facilitates exogenous HA depolymerisation [20].

CEMIP, cancer and EMT

The role of CEMIP in cancer development has been recently reviewed [75], so here we briefly summarise the insights into studies on cancer cells afford regarding CEMIP function and its involvement in patho-physiological mechanisms.

CEMIP has been extensively documented in colon cancer. Pioneering studies in 2007 by Giancarlo Marra's group at the University of Zurich searched for a distinguishing gene signature of precancerous adenomatous colorectal polyps compared with the normal colonic epithelium [72]. The authors showed that modulation of the Wnt pathway was a key feature of this transformation and that, amongst all the genes identified, transcription of CEMIP was most affected. Thus, the CEMIP expression pattern, normally confined to the lower portion of normal colonic epithelial crypts, was disrupted in dysplastic glands with widespread higher expression levels at both the mRNA and the protein levels. Subsequent work demonstrated that this overexpression was also present in colorectal cancer, with a nuclear expression in the majority of UICC stage I-IV adenocarcinomas [22]. Other studies confirmed these findings, showing that increased CEMIP expression occurred in the phenotypic switch from normal to adenomatous or neoplastic lesions [76]. In addition, increased CEMIP mRNA expression was

detectable in the plasma of patients with adenomatous or neoplastic lesions [76], which was associated with poor patient prognosis [77].

CEMIP expression in other types of cancer is consistent with the observations made in colon cancer. For example, CEMIP was detected with intermediate or high expression levels in human papillomavirus cervical (pre)neoplastic lesions, whereas only marginal levels were detected in normal exocervical epithelium [21]. This pattern was also observed in pancreatic cancer (PanCa) with an overexpression in benign pancreatic intraepithelial neoplasia, as well as in PanCa in both human and genetically engineered mouse models [78]. It is noteworthy that in all these cancer studies, CEMIP expression was always confined to the epithelial cells.

CEMIP has been shown to play a central role in EMT, a developmental morphogenic program in which cells lose their epithelial characteristics and gain a mesenchymal phenotype [79]. EMT is characterised by downregulation of epithelial and upregulation of mesenchymal markers, has been implicated in tumour initiation through endowing transformed cells with stemness characteristics and has been more extensively investigated in the process in invasion and metastasis [80]. Several *in vitro* studies have reported CEMIP-induced EMT in different tumour cell lines. Thus, CEMIP silencing in non-small-cell lung cancer cell lines induced expression of E-cadherin (epithelial marker) and decreased expression of Vimentin, Snail and Twist (mesenchymal markers) [81]. Conversely, CEMIP overexpression resulted in decreased E-cadherin levels and increased expression of Vimentin, Snail and Twist [81]. In gastric cancer cell lines CEMIP knock-down resulted in decreased mRNA expression of the EMT-related markers Slug, Snail, Vimentin and Twist [82]. Another study demonstrated an increase in E-cadherin protein levels and a decrease in Slug and Vimentin levels in BCPAP cells line after CEMIP silencing [52].

Several recent studies analysed the relationship between CEMIP and EMT markers in human tumour tissues. For example, in 2018, Jiang and colleagues studied HCC samples and found a positive correlation between CEMIP expression and *N*-cadherin and Vimentin expression, and a negative one with E-cadherin expression [83]. In colon cancer samples, an inverse relationship between CEMIP localisation and E-cadherin expression [58] exists, while in cholangiocarcinoma tissue samples, a negative relationship of CEMIP with E-cadherin and a positive one with *N*-cadherin and Vimentin [59] has been found. Other evidence of CEMIP involvement in EMT includes the fact that the epithelial splicing regulatory protein 1

(ESRP1), an epithelial cell-specific regulator, has been identified as an upstream regulator of CEMIP [84].

Three main pathways have been described in the CEMIP-dependent EMT process (Fig. 2): PI3K-Akt-mediated, EGFR-mediated and Wnt/β-catenin-mediated EMT.

- *PI3K-Akt-mediated EMT*: TGF-β is a potent inducer of EMT [85], and promotes cell survival through the PI3K-Akt pathway in cancer [86]. Tang et al. [81] showed that CEMIP silencing in lung cancer cell lines decreased the levels of TGF-β, PI3K and Akt . On the contrary, CEMIP overexpression increased TGF-β, PI3K and Akt levels [81]. Similar findings have been reported in cholangiocarcinoma cell lines, where a significant downregulation of TGF-β as well as PI3K, AKT and mTOR was observed after CEMIP silencing, whereas an increase of these proteins was observed in CEMIP overexpressing cells [59].

- *EGFR-mediated EMT*: Epidermal growth factor (EGF) promotes cell growth and differentiation by binding to the epidermal growth factor receptor (EGFR) [87]. It has been reported that CEMIP activates EGFR signalling and regulates the downstream kinases in the EGF-mediated EMT pathway [53]. Tang et al. [81] observed a correlation between the expression of CEMIP and the expression of EFGR in lung cancer cell lines. In HCC cells resistant to the kinase inhibitor drug Sorafenib, high levels of CEMIP triggered EGF-induced EMT, increasing the migratory and invasive ability of these cells [88]. CEMIP upregulation in HCC parental cells induces phosphorylation of EGFR and its downstream kinases, leading to EMT, while CEMIP silencing reduces EGFR expression [88]. All in all, these data suggest that CEMIP acts as an upstream regulator of EGFR signalling.

- *Wnt/β-catenin-mediated EMT*: Several studies suggest an important role of Wnt/β-catenin signalling during EMT [89,90]. In 2019, Deroyer and colleagues demonstrated that CEMIP modulates the Wnt/β-catenin signalling in human chondrocytes [47]. In gastric carcinoma cells, CEMIP knock-down reduces the expression of β-catenin, but also of c-myc and cyclinD1, two downstream players of the Wnt/β-catenin pathway [82]. These findings suggest a role for CEMIP in promoting invasion and metastasis through the Wnt/β-catenin signalling pathway [82]. Indeed, the Wnt/β-catenin signalling pathway coordinates pivotal cellular processes such as cell differentiation, proliferation, migration and many others. A study published in 2007 by Sabates-Bellver et al. [72]

examined the transcriptomes of colorectal polyp samples and found a positive correlation between CEMIP upregulation and the expression of Wnt targets. Moreover, northern blotting analysis confirmed an important decrease in CEMIP mRNA levels linked to Wnt pathway inhibition. In 2011, Birkenkamp-Demtroder et al. [22] demonstrated that lentiviral-mediated knock-down of CEMIP in colon cell lines alters the cell cycle and the Wnt-signalling pathway. Ingenuity pathways analysis (IPA) together with immunofluorescence analyses and western blotting showed a change in the expression (mostly downregulation) of 67 genes involved in Wnt/β-catenin signalling upon CEMIP knock-down. In view of these observations, it is interesting to speculate about a potential role of CEMIP and the HA fragments it produces in the activation of the Wnt/β-catenin signalling pathway, which could conceivably coherently connect CEMIP expression to cell proliferation and migration initiated by the binding of Wnt followed by β-catenin dephosphorylation.

Taken together, these observations suggest an involvement of CEMIP in EMT regulation. The publications that support this notion are summarised in Table 1.

In addition to cancer-relevant functions for CEMIP in tumour cells, several observations suggest that CEMIP might also regulate the tumour microenvironment. For example, as a major component of ECM, HA can act as a modulator of the tumour microenvironment, and thereby regulate tumour growth, angiogenesis, invasion and metastasis [91]. Given that CEMIP functions as a hyaluronidase, its involvement in cancer could conceivably be directly linked to its hyaluronidase activity in this context. Furthermore, CEMIP has been identified as a candidate gene for skin tropism of cancers [92]. Although these data are anecdotal because the study only employed a single patient-derived cell line, it is interesting to note that they suggest that CEMIP might create a permissive ECM environment that foster tumour engraftment in the skin. Moreover, CEMIP-containing exosomes have been reported to educate the brain microenvironment and thereby foster cranial metastases [93]. Possibly consistent with these findings, the growth of glioblastoma cells implanted into the mouse brain was suppressed in CEMIP KO mice compared with the situation in wild-type mice, which was associated with decreased macrophage infiltration into the brain tumours [94], again implicating CEMIP in the creation of a pro-tumour microenvironment.

CEMIP in diseases and ageing

In this section, we will summarise the evidence that CEMIP is involved in pathophysiological processes

Table 1. Summary of EMT markers involved in CEMIP silencing and upregulation.

	Epithelial phenotype	Mesenchymal phenotype	Cell line/tissue	References
CEMIP silencing	↑ E-cadherin	↓ Vimentin ↓ Snail ↓ Twist	NSCLC cell line	Tang et al. [81]
		↓ Vimentin ↓ Snail ↓ Slug ↓ Twist	GC cell line	Jia et al. [82]
	↑ E-cadherin	↓ Vimentin ↓ Slug	PTC cell lines	Jiao et al. [52]
CEMIP upregulation	↓ E-cadherin	↑ Vimentin ↑ Snail ↑ Twist	NSCLC cell line	Tang et al. [81]
	↓ E-cadherin	↑ N-cadherin ↑ Vimentin	HCC tissues	Jiang et al. [83]
	↓ E-cadherin	↑ N-cadherin ↑ Vimentin	CCA tissues	Zhai et al. [59]
EMT markers	↑ E-cadherin ↑ Cytokeratin ↑ Laminin-1 ↑ ZO-1 ↑ Occludin	↑ N-cadherin ↑ Vimentin ↑ Snail ↑ Slug ↑ Twist	NSCLC = non-small-cell lung cancer GC = gastric cancer PTC = papillary thyroid cancer HCC = hepatocellular carcinoma CCA = cholangiocarcinoma	

Light red: silencing; light green: upregulation; grey: EMT markers of the epithelial and mesenchymal phenotype.

that regulate diseases of the nervous system, inflammatory diseases, as well as the ageing/senescence process.

Diseases of the nervous system

Hearing loss

CEMIP was originally described in families that are affected by nonsyndromic hearing loss, which is associated with CEMIP mutations (R187C, R187H and H783Y) [27]. These authors showed that in mice, CEMIP was expressed specifically in Deiters' cells in the organ of Corti at post-natal day P0, before the onset of hearing, and that expression in those cells disappeared by day P7. In addition, CEMIP expression was observed in fibrocytes of the spiral ligament and the spiral limbus through to P21, when the murine cochlea matures. The *CEMIP* gene was therefore proposed to play a role in auditory development. These results were confirmed in subsequent studies [69,95,96].

Multiple sclerosis

In experimental autoimmune encephalomyelitis (EAE), an animal model for multiple sclerosis (MS) that is characterised by focal demyelinating lesions, CEMIP immunoreactivity was demonstrated to be exclusively associated with focal loci in damaged white columns of the spinal cord, and was mainly expressed by activated astrocytes that invaded damaged tissue [97]. Similar findings were observed in tissue from a patient with MS, suggesting that CEMIP expression by activated astrocytes could explain the focal HA degradation observed during MS progression and might represent a possible new therapeutic target [97].

Inflammatory diseases

Rheumatoid arthritis and osteoarthritis

Pioneering work on the role of CEMIP was conducted in 2013 by the team of Yoshida et *al*. using osteoarthritis (OA) and rheumatoid arthritis (RA) synovial tissues. They found that transcription of CEMIP was higher in OA or RA synovium than in noninflamed synovium [13]. These data suggested that CEMIP is a unique hyaladherin, with a key role in HA catabolism in the arthritic synovium. Relative to healthy controls, CEMIP expression was also found to be more strongly increased in synovial tissues with active RA compared with tissue with inactive disease, and is associated with CEMIP-dependent proliferation

of fibroblast-like synoviocytes [98]. Angiogenesis is thought to be a key event in the formation and maintenance of the pannus in RA, and CEMIP was also implicated in promoting angiogenesis *in vitro* and *in vivo* in this study [98]. Subsequently, a positive correlation between CEMIP levels and the inflammatory markers TNF-α, IL-1β and IL-6 in both serum and synovial fluids of RA patients was reported, underscoring the clinical relevance of these findings, and treatment with an anti-CEMIP antibody partially alleviated arthritis severity and reduced serum LMW-HA levels and cytokine secretion in a collagen-induced arthritis (CIA) mouse model [20], providing proof of principle of the possible therapeutic relevance of these observations.

Mechanistically, CEMIP-KO mice exhibited resistance to CIA, which could be partially rescued by intra-articular injection of vectors encoding full-length CEMIP, whereas a CEMIP mutant with an inactive G8 domain had no effect [20]. Moreover, CEMIP expression was found to be suppressed by TGFβ-1 in normal synovial fibroblasts, but was only slightly decreased in synovial fibroblasts from OA or RA patients [48]. Coupled to this, TGFβ-1 upregulated the expression HA synthases (HAS1/2), which resulted in the increased accumulation of intermediate-sized HA fragments in cultures of synovial fibroblasts from OA or RA patients compared with normal synovial fibroblasts [48]. In addition, IL-6 significantly upregulated CEMIP expression and HA-degrading activity in OA synovial fibroblasts [67]. Collectively, these data suggest a key role for CEMIP in the pathophysiology of RA.

In addition to synovial fibroblasts, CEMIP has been shown to be expressed by hypertrophic chondrocytes at the chondro-osseous junction. Thus, in CEMIP-deficient mice, an accumulation of HMW-HA with reduced angiogenesis can be observed in this context and is thought to be a key mechanism that modulates endochondral ossification during postnatal development [68]. In a mouse model of temporo-mandibular joint lesion, CEMIP expression was shown to be induced in injured mandibular condyles in association with increased HYAL2 expression [99]. In a study analysing human cartilage, CEMIP transcription was significantly higher in OA cartilage compared with control cartilage, and CEMIP was highly expressed by chondrocytes in the HA-depleted area of OA cartilage [100]. Interestingly, CEMIP immunoreactivity correlates with the Mankin score, the histopathologic severity score used to evaluate OA lesions of the cartilage. OA chondrocytes exhibit HA-degrading activity, which is abolished by knock-down of CEMIP but not by

downregulation of hyaluronidases HYAL1, HYAL2 or CD44 [100]. Unlike RA synovial fibroblasts, only TNF-α but not histamine stimulated OA chondrocytes to overexpress CEMIP. Similar results were obtained by others [47]. In addition, these authors studied CEMIP expression *in vitro* using a chondrocyte dedifferentiation model. High-throughput RNA sequencing was performed on chondrocytes after CEMIP silencing. Most of the deregulated genes were involved in cartilage turnover, mesenchymal transition and fibrosis. Moreover, CEMIP was demonstrated to be essential for chondrocytes proliferation, promoted αSMA expression, and was co-expressed *in situ* with αSMA in all OA cartilage layers. These results strongly suggest a role for CEMIP in the trans-differentiation of chondrocytes into 'chondro-myofibroblasts' which are found in OA cartilage but not in healthy cartilage.

Inflammatory bowel disease

Crohn's disease (CD) is a chronic inflammatory disease of the gastrointestinal tract. Cultured CD fibroblasts produce increased levels of CEMIP protein through an IL-6-driven autocrine mechanism compared with control colon fibroblasts [39]. This leads to CEMIP deposition in the ECM and excessive degradation of HA, generating HA fragments that contribute to gut inflammation and fibrosis. Accordingly, antibody blockade of IL-6 receptors in CD fibroblasts decreased CEMIP protein levels in the ECM, and CEMIP silencing abrogated the ability of colon fibroblasts to degrade HA.

The ageing/senescence process

Skin

In the skin, the metabolism of HA is highly regulated. Ageing leads to chronic low-grade inflammation, which is characterised by elevated levels of pro-inflammatory cytokines [46]. CEMIP is expressed in dermal fibroblasts [13] and in photoexposed skin [101,102], and is directly correlated with skin roughness in the papillary dermis [101]. Interestingly, mast cells were significantly enriched in the photoaged skin and were frequently associated with CEMIP-positive fibroblasts [102]. Furthermore, treatment of skin fibroblasts with a pro-inflammatory cytokine mixture (TNF-α, IL-1β and IL-6) or with IL1β alone suppressed HA depolymerisation through downregulation of CEMIP expression, which was associated with permanently increased HA levels in the culture medium due to the upregulation of HAS2 [46]. These results

might suggest that cytokine-stimulated fibroblasts in unexposed skin increase the amounts of dermal HMW-HA to protect surrounding tissues against short-term inflammation. Notably, IL-6 and IL-8 individually had no effect on HA metabolism in human skin fibroblasts, and their effect was only observed as part of a cocktail [46]. By contrast, and as mentioned above, IL-6 has been reported to promote HA depolymerisation in synovial fibroblasts by inducing CEMIP expression [67]. This suggests that CEMIP might be differently regulated according to the cell- or tissue type.

Aortic valve disease

Aortic valve disease (AVD) is one of the leading causes of cardiovascular mortality. Ageing is a significant clinical risk factor, and calcific AVD is the most common type of AVD occurring in 2% of the aged population [103]. During latent AVD, abnormal expression of HA and of its synthesising/degrading enzymes has been observed [103]. Physiologically, the middle spongiosa layer of aortic valve consists primarily of proteoglycans and glycosaminoglycans, and provides lubrication and dampening functions when the valve leaflet flexes opens and closes. In normal human aortic valve tissue, CEMIP expression is present at all postnatal stages and increases with age, being mostly expressed inside the cells at the young stage and in ECM at advanced age. Increased CEMIP expression was demonstrated in calcified aortic valves, especially around calcified nodules in the fibrosa layer [102]. Finally, *in vitro* studies reported that CEMIP expression, which was detected in porcine aortic valve interstitial cells, was modulated by stiffness, as less stiff substrates were observed to downregulate CEMIP expression [104].

Conclusion

This review summarises all the major scientific evidence reported in the literature so far regarding the structure, role and mechanism of CEMIP. Collectively, CEMIP emerges as an important protein loosening the ECM and thus creating a permissive environment for EMT and cellular migration. CEMIP expression and function is associated with several types of cancer cell, while in diseases other than cancer, its cellular expression is largely limited to mesenchymal cells.

Despite all the scientific observations reported here, CEMIP biology remains, in our opinion, enigmatic. Our current knowledge of CEMIP biology is dominated by the vast body of literature that has

investigated cancer-relevant roles for CEMIP expression or absence, for example in the context of tumour proliferation and invasion. CEMIP might exert part of its role via a structural function, however, as we have detailed above. CEMIP also exerts a unique enzymatic activity, producing, alone or in coordination with other proteins, a range of int-HA and LMW-HA fragments. This role might be partially or fully distinct from the structural one. Additional loss and gain of function studies are required to investigate the functional relationship between the HA fragments produced by CEMIP, and the physiological and pathophysiological processes that are regulated by CEMIP. Moreover, the role of CEMIP, and the HA fragments it produces might exert quite different effects in different cellular settings [100] and disease contexts [20,78,92,94].

In our opinion, the literature provides compelling evidence that CEMIP plays a role in modulating fibrosis (for example through the modulation of CEMIP activity by TGF-β1 and histamine), an observation that can guide future mechanistic studies in simple cellular systems [105]. Once additional structural and functional knowledge have been generated, this can be leveraged to identify compounds capable of modulating the CEMIP protein function *in vivo*, Indeed, we have recently reported that highly sulfated HA is a potent inhibitor of the CEMIP hyaluronidase activity [49]. These and further next-generation compounds are expected to have a strong potential for therapeutic application, for example through modulating tissue inflammation, an unmet clinical need for a range of inflammatory disease conditions.

Acknowledgement

Conflict of interest

The authors declare no conflict of interest.

Author contributions

SS, CG, SM and MP conceived and designed the manuscript. SS and LP produced figures and tables. SS, CG, MP and JS wrote the manuscript. AC, AB, LS, JSl and SM critically revised the text. All authors contributed to the final version of this manuscript.

Data availability statement

Data derived from public domain resources.

References

1 Vigetti D, Viola M, Karousou E, De Luca G, Passi A. Metabolic control of hyaluronan synthases. *Matrix Biol*. 2014;**35**:8–13.

2 Padhi A, Nain AS. ECM in differentiation: a review of matrix structure, composition and mechanical properties. *Ann Biomed Eng*. 2020;**48**:1071–89.

3 Fallacara A, Baldini E, Manfredini S, Vertuani S. Hyaluronic acid in the third millennium. *Polymers (Basel)*. 2018;**10**:701.

4 Bayer IS. Hyaluronic acid and controlled release: a review. *Molecules*. 2020;**25**:2649.

5 Yoshida H, Nagaoka A, Nakamura S, Sugiyama Y, Okada Y, Inoue S. Murine homologue of the human KIAA1199 is implicated in hyaluronan binding and depolymerization. *FEBS Open Bio*. 2013;**3**:352–6.

6 Milner CM, Tongsoongnoen W, Rugg MS, Day AJ. The molecular basis of inter-alpha-inhibitor heavy chain transfer on to hyaluronan. *Biochem Soc Trans*. 2007;**35**:672–6.

7 Supp DM, Hahn JM, McFarland KL, Glaser K. Inhibition of hyaluronan synthase 2 reduces the abnormal migration rate of keloid keratinocytes. *J Burn Care Res*. 2014;**35**:84–92.

8 Volpi N, Schiller J, Stern R, Soltes L. Role, metabolism, chemical modifications and applications of hyaluronan. *Curr Med Chem*. 2009;**16**:1718–45.

9 Zhang Z, Tao D, Zhang P, Liu X, Zhang Y, Cheng J, et al. Hyaluronan synthase 2 expressed by cancer-associated fibroblasts promotes oral cancer invasion. *J Exp Clin Cancer Res*. 2016;**35**:181.

10 Li Y, Liang J, Yang T, Monterrosa Mena J, Huan C, Xie T, et al. Hyaluronan synthase 2 regulates fibroblast senescence in pulmonary fibrosis. *Matrix Biol*. 2016;**55**:35–48.

11 Zhang H, Tsang JY, Ni YB, Chan SK, Chan KF, Cheung SY, et al. Hyaluronan synthase 2 is an adverse prognostic marker in androgen receptor-negative breast cancer. *J Clin Pathol*. 2016;**69**:1055–62.

12 Stern R. Hyaluronidases in cancer biology. *Semin Cancer Biol*. 2008;**18**:275–80.

13 Yoshida H, Nagaoka A, Kusaka-Kikushima A, Tobiishi M, Kawabata K, Sayo T, et al. KIAA1199, a deafness gene of unknown function, is a new hyaluronan binding protein involved in hyaluronan depolymerization. *Proc Natl Acad Sci USA*. 2013;**110**:5612–7.

14 Kaul A, Short WD, Wang X, Keswani SG. Hyaluronidases in human diseases. *Int J Mol Sci*. 2021;**22**:3204.

15 Zadnikova P, Sinova R, Pavlik V, Simek M, Safrankova B, Hermannova M, et al. The degradation of hyaluronan in the skin. *Biomolecules*. 2022;**12**:251.

16 Csoka AB, Frost GI, Stern R. The six hyaluronidase-like genes in the human and mouse genomes. *Matrix Biol*. 2001;**20**:499–508.

17 Yamamoto H, Tobisawa Y, Inubushi T, Irie F, Ohyama C, Yamaguchi Y. A mammalian homolog of the zebrafish transmembrane protein 2 (TMEM2) is the long-sought-after cell-surface hyaluronidase. *J Biol Chem*. 2017;**292**:7304–13.

18 Yoshida H, Okada Y. Role of HYBID (hyaluronan binding protein involved in hyaluronan depolymerization), alias KIAA1199/CEMIP, in hyaluronan degradation in Normal and photoaged skin. *Int J Mol Sci*. 2019;**20**:5804.

19 Yoshida H, Nagaoka A, Nakamura S, Tobiishi M, Sugiyama Y, Inoue S. N-terminal signal sequence is required for cellular trafficking and hyaluronan-depolymerization of KIAA1199. *FEBS Lett*. 2014;**588**:111–6.

20 Zhang W, Yin G, Zhao H, Ling H, Xie Z, Xiao C, et al. Secreted KIAA1199 promotes the progression of rheumatoid arthritis by mediating hyaluronic acid degradation in an ANXA1-dependent manner. *Cell Death Dis*. 2021;**12**:102.

21 Shostak K, Zhang X, Hubert P, Goktuna SI, Jiang Z, Klevernic I, et al. NF-kappaB-induced KIAA1199 promotes survival through EGFR signalling. *Nat Commun*. 2014;**5**:5232.

22 Birkenkamp-Demtroder K, Maghnouj A, Mansilla F, Thorsen K, Andersen CL, Oster B, et al. Repression of KIAA1199 attenuates Wnt-signalling and decreases the proliferation of colon cancer cells. *Br J Cancer*. 2011;**105**:552–61.

23 He QY, Liu XH, Li Q, Studholme DJ, Li XW, Liang SP. G8: a novel domain associated with polycystic kidney disease and non-syndromic hearing loss. *Bioinformatics*. 2006;**22**:2189–91.

24 Bernstein FC, Koetzle TF, Williams GJB, Meyer JEF, Brice MD, Rodgers JR, et al. The protein data bank: a computer-based archival file for macromolecular structures. *J Mol Biol*. 1977;**112**:535–42.

25 Guo J, Cheng H, Zhao S, Yu L. GG: a domain involved in phage LTF apparatus and implicated in human MEB and non-syndromic hearing loss diseases. *FEBS Lett*. 2006;**580**:581–4.

26 Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018;**46**:W296–303.

27 Abe S, Usami S, Nakamura Y. Mutations in the gene encoding KIAA1199 protein, an inner-ear protein expressed in Deiters' cells and the fibrocytes, as the cause of nonsyndromic hearing loss. *J Hum Genet*. 2003;**48**:564–70.

28 Garron ML, Cygler M. Uronic polysaccharide degrading enzymes. *Curr Opin Struct Biol*. 2014;**28**:87–95.

29 Yamaguchi Y, Yamamoto H, Tobisawa Y, Irie F. TMEM2: a missing link in hyaluronan catabolism identified? *Matrix Biol*. 2019;**78-79**:139–46.

30 Ward CJ, Hogan MC, Rossetti S, Walker D, Sneddon T, Wang X, et al. The gene mutated in autosomal recessive polycystic kidney disease encodes a large, receptor-like protein. *Nat Genet*. 2002;**30**:259–69.

31 Muhammed MT, Aki-Yalcin E. Homology modeling in drug discovery: overview, current applications, and future perspectives. *Chem Biol Drug Des*. 2019;**93**:12–20.

32 Kuwabara N, Manya H, Yamada T, Tateno H, Kanagawa M, Kobayashi K, et al. Carbohydrate-binding domain of the POMGnT1 stem region modulates O-mannosylation sites of α-dystroglycan. *Proc Natl Acad Sci USA*. 2016;**113**:9280–5.

33 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;**596**:583–9.

34 Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*. 2021;**373**:871–6.

35 Garantziotis S, Savani RC. Hyaluronan biology: a complex balancing act of structure, function, location and context. *Matrix Biol*. 2019;**78-79**:1–10.

36 Jing W, DeAngelis PL. Synchronized chemoenzymatic synthesis of monodisperse hyaluronan polymers. *J Biol Chem*. 2004;**279**:42345–9.

37 Schmaus A, Klusmeier S, Rothley M, Dimmler A, Sipos B, Faller G, et al. Accumulation of small hyaluronan oligosaccharides in tumour interstitial fluid correlates with lymphatic invasion and lymph node metastasis. *Br J Cancer*. 2014;**111**:559–67.

38 Bauer J, Rothley M, Schmaus A, Quagliata L, Ehret M, Biskup M, et al. TGFbeta counteracts LYVE-1-mediated induction of lymphangiogenesis by small hyaluronan oligosaccharides. *J Mol Med (Berl)*. 2018;**96**:199–209.

39 Soroosh A, Albeiroti S, West GA, Willard B, Fiocchi C, de la Motte CA. Crohn's disease fibroblasts overproduce the novel protein KIAA1199 to create proinflammatory hyaluronan fragments. *Cell Mol Gastroenterol Hepatol*. 2016;**2**:358–368.e4.

40 Zhang B, Du Y, He Y, Liu Y, Zhang G, Yang C, et al. INT-HA induces M2-like macrophage differentiation of human monocytes via TLR4-miR-935 pathway. *Cancer Immunol Immunother*. 2019;**68**:189–200.

41 Kuang DM, Wu Y, Chen N, Cheng J, Zhuang SM, Zheng L. Tumor-derived hyaluronan induces

formation of immunosuppressive macrophages through transient early activation of monocytes. *Blood*. 2007;**110**:587–95.

42 Sokolowska M, Chen LY, Eberlein M, Martinez-Anton A, Liu Y, Alsaaty S, et al. Low molecular weight hyaluronan activates cytosolic phospholipase A2alpha and eicosanoid production in monocytes and macrophages. *J Biol Chem*. 2014;**289**:4470–88.

43 Ghazi K, Deng-Pichon U, Warnet JM, Rat P. Hyaluronan fragments improve wound healing on in vitro cutaneous model through P2X7 purinoreceptor basal activation: role of molecular weight. *PLoS One*. 2012;**7**:e48351.

44 Winkler J, Abisoye-Ogunniyan A, Metcalf KJ, Werb Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nat Commun*. 2020;**11**:5120.

45 Kohi S, Sato N, Koga A, Matayoshi N, Hirata K. KIAA1199 is induced by inflammation and enhances malignant phenotype in pancreatic cancer. *Oncotarget*. 2017;**8**:17156–63.

46 Sato S, Mizutani Y, Yoshino Y, Masuda M, Miyazaki M, Hara H, et al. Pro-inflammatory cytokines suppress HYBID (hyaluronan (HA) -binding protein involved in HA depolymerization/KIAA1199/CEMIP) -mediated HA metabolism in human skin fibroblasts. *Biochem Biophys Res Commun*. 2021;**539**:77–82.

47 Deroyer C, Charlier E, Neuville S, Malaise O, Gillet P, Kurth W, et al. CEMIP (KIAA1199) induces a fibrosis-like process in osteoarthritic chondrocytes. *Cell Death Dis*. 2019;**10**:103.

48 Nagaoka A, Yoshida H, Nakamura S, Morikawa T, Kawabata K, Kobayashi M, et al. Regulation of hyaluronan (HA) metabolism mediated by HYBID (hyaluronan-binding protein involved in HA depolymerization, KIAA1199) and HA synthases in growth factor-stimulated fibroblasts. *J Biol Chem*. 2015;**290**:30910–23.

49 Schmaus A, Rothley M, Schreiber C, Moller S, Rosswag S, Franz S, et al. Sulfated hyaluronic acid inhibits the hyaluronidase CEMIP and regulates the HA metabolism, proliferation and differentiation of fibroblasts. *Matrix Biol*. 2022;**109**:173–91.

50 Yoshino Y, Goto M, Hara H, Inoue S. The role and regulation of TMEM2 (transmembrane protein 2) in HYBID (hyaluronan (HA)-binding protein involved in HA depolymerization/ KIAA1199/CEMIP)-mediated HA depolymerization in human skin fibroblasts. *Biochem Biophys Res Commun*. 2018;**505**:74–80.

51 Kuscu C, Evensen N, Kim D, Hu YJ, Zucker S, Cao J. Transcriptional and epigenetic regulation of KIAA1199 gene expression in human breast cancer. *PLoS One*. 2012;**7**:e44661.

52 Jiao X, Ye J, Wang X, Yin X, Zhang G, Cheng X. KIAA1199, a target of MicoRNA-486-5p, promotes papillary thyroid cancer invasion by influencing epithelial-mesenchymal transition (EMT). *Med Sci Monit*. 2019;**25**:6788–96.

53 Wang A, Zhu J, Li J, Du W, Zhang Y, Cai T, et al. Downregulation of KIAA1199 by miR-486-5p suppresses tumorigenesis in lung cancer. *Cancer Med*. 2020;**9**:5570–86.

54 Wang L, Yu T, Li W, Li M, Zuo Q, Zou Q, et al. The miR-29c-KIAA1199 axis regulates gastric cancer migration by binding with WBP11 and PTP4A3. *Oncogene*. 2019;**38**:3134–50.

55 Zhang D, Zhao L, Shen Q, Lv Q, Jin M, Ma H, et al. Down-regulation of KIAA1199/CEMIP by miR-216a suppresses tumor invasion and metastasis in colorectal cancer. *Int J Cancer*. 2017;**140**:2298–309.

56 Zhou Y, Mu T. LncRNA LINC00958 promotes tumor progression through miR-4306/CEMIP axis in osteosarcoma. *Eur Rev Med Pharmacol Sci*. 2021;**25**:3182–99.

57 Hsieh IY, He J, Wang L, Lin B, Liang Z, Lu B, et al. H3K27me3 loss plays a vital role in CEMIP mediated carcinogenesis and progression of breast cancer with poor prognosis. *Biomed Pharmacother*. 2020;**123**:109728.

58 Evensen NA, Li Y, Kuscu C, Liu J, Cathcart J, Banach A, et al. Hypoxia promotes colon cancer dissemination through up-regulation of cell migration-inducing protein (CEMIP). *Oncotarget*. 2015;**6**:20723–39.

59 Zhai X, Wang W, Ma Y, Zeng Y, Dou D, Fan H, et al. Serum KIAA1199 is an advanced-stage prognostic biomarker and metastatic oncogene in cholangiocarcinoma. *Aging (Albany NY)*. 2020;**12**:23761–77.

60 Michishita E, Garces G, Barrett JC, Horikawa I. Upregulation of the KIAA1199 gene is associated with cellular mortality. *Cancer Lett*. 2006;**239**:71–7.

61 Csoka AB, Scherer SW, Stern R. Expression analysis of six paralogous human hyaluronidase genes clustered on chromosomes 3p21 and 7q31. *Genomics*. 1999;**60**:356–61.

62 Miyata S, Kitagawa H. Formation and remodeling of the brain extracellular matrix in neural plasticity: roles of chondroitin sulfate and hyaluronan. *Biochim Biophys Acta Gen Subj*. 2017;**1861**:2420–34.

63 Rauch U. Extracellular matrix components associated with remodeling processes in brain. *Cell Mol Life Sci*. 2004;**61**:2031–45.

64 Wilson E, Knudson W, Newell-Litwa K. Hyaluronan regulates synapse formation and function in developing neural networks. *Sci Rep*. 2020;**10**:16459.

65 Yoshino Y, Ishisaka M, Tsuruma K, Shimazawa M, Yoshida H, Inoue S, et al. Distribution and function of hyaluronan binding protein involved in hyaluronan depolymerization (HYBID, KIAA1199) in the mouse central nervous system. *Neuroscience*. 2017;**347**:1–10.

66 Yoshino Y, Shimazawa M, Nakamura S, Inoue S, Yoshida H, Shimoda M, et al. Targeted deletion of HYBID (hyaluronan binding protein involved in hyaluronan depolymerization/ KIAA1199/CEMIP) decreases dendritic spine density in the dentate gyrus through hyaluronan accumulation. *Biochem Biophys Res Commun.* 2018;**503**:1934–40.

67 Shiozawa J, de Vega S, Cilek MZ, Yoshinaga C, Nakamura T, Kasamatsu S, et al. Implication of HYBID (hyaluronan-binding protein involved in hyaluronan depolymerization) in hyaluronan degradation by synovial fibroblasts in patients with knee osteoarthritis. *Am J Pathol.* 2020;**190**:1046–58.

68 Shimoda M, Yoshida H, Mizuno S, Hirozane T, Horiuchi K, Yoshino Y, et al. Hyaluronan-binding protein involved in hyaluronan depolymerization controls endochondral ossification through hyaluronan metabolism. *Am J Pathol.* 2017;**187**:1162–76.

69 Usami S, Takumi Y, Suzuki N, Oguchi T, Oshima A, Suzuki H, et al. The localization of proteins encoded by CRYM, KIAA1199, UBA52, COL9A3, and COL9A1, genes highly expressed in the cochlea. *Neuroscience.* 2008;**154**:22–8.

70 Matsuzaki S, Tanaka F, Mimori K, Tahara K, Inoue H, Mori M. Clinicopathologic significance of KIAA1199 overexpression in human gastric cancer. *Ann Surg Oncol.* 2009;**16**:2042–51.

71 Liu J, Han P, Gong J, Wang Y, Chen B, Liao J, et al. Knockdown of KIAA1199 attenuates growth and metastasis of hepatocellular carcinoma. *Cell Death Discov.* 2018;**4**:102.

72 Sabates-Bellver J, Van der Flier LG, de Palo M, Cattaneo E, Maake C, Rehrauer H, et al. Transcriptome profile of human colorectal adenomas. *Mol Cancer Res.* 2007;**5**:1263–75.

73 Tiwari A, Schneider M, Fiorino A, Haider R, Okoniewski MJ, Roschitzki B, et al. Early insights into the function of KIAA1199, a markedly overexpressed protein in human colorectal tumors. *PLoS One.* 2013;**8**:e69473.

74 Evensen NA, Kuscu C, Nguyen HL, Zarrabi K, Dufour A, Kadam P, et al. Unraveling the role of KIAA1199, a novel endoplasmic reticulum protein, in cancer cell migration. *J Natl Cancer Inst.* 2013;**105**:1402–16.

75 Liu J, Yan W, Han P, Tian D. The emerging role of KIAA1199 in cancer development and therapy. *Biomed Pharmacother.* 2021;**138**:111507.

76 LaPointe LC, Pedersen SK, Dunne R, Brown GS, Pimlott L, Gaur S, et al. Discovery and validation of molecular biomarkers for colorectal adenomas and cancer with application to blood testing. *PLoS One.* 2012;**7**:e29059.

77 Fink SP, Myeroff LL, Kariv R, Platzer P, Xin B, Mikkola D, et al. Induction of KIAA1199/CEMIP is associated with colon cancer phenotype and poor patient survival. *Oncotarget.* 2015;**6**:30500–15.

78 Suh HN, Jun S, Oh AY, Srivastava M, Lee S, Taniguchi CM, et al. Identification of KIAA1199 as a biomarker for pancreatic intraepithelial neoplasia. *Sci Rep.* 2016;**6**:38273.

79 Thiery JP, Sleeman JP. Complex networks orchestrate epithelial-mesenchymal transitions. *Nat Rev Mol Cell Biol.* 2006;**7**:131–42.

80 Santamaria PG, Moreno-Bueno G, Portillo F, Cano A. EMT: present and future in clinical oncology. *Mol Oncol.* 2017;**11**:718–38.

81 Tang Z, Ding Y, Shen Q, Zhang C, Li J, Nazar M, et al. KIAA1199 promotes invasion and migration in non-small-cell lung cancer (NSCLC) via PI3K-Akt mediated EMT. *J Mol Med (Berl).* 2019;**97**:127–40.

82 Jia S, Qu T, Wang X, Feng M, Yang Y, Feng X, et al. KIAA1199 promotes migration and invasion by Wnt/beta-catenin pathway and MMPs mediated EMT progression and serves as a poor prognosis marker in gastric cancer. *PLoS One.* 2017;**12**:e0175058.

83 Jiang Z, Zhai X, Shi B, Luo D, Jin B. KIAA1199 overexpression is associated with abnormal expression of EMT markers and is a novel independent prognostic biomarker for hepatocellular carcinoma. *Onco Targets Ther.* 2018;**11**:8341–8.

84 Kikuchi M, Yamashita K, Waraya M, Minatani N, Ushiku H, Kojo K, et al. Epigenetic regulation of ZEB1-RAB25/ESRP1 axis plays a critical role in phenylbutyrate treatment-resistant breast cancer. *Oncotarget.* 2016;**7**:1741–53.

85 Xu J, Lamouille S, Derynck R. TGF-beta-induced epithelial to mesenchymal transition. *Cell Res.* 2009;**19**:156–72.

86 Zhang L, Zhou F, ten Dijke P. Signaling interplay between transforming growth factor-beta receptor and PI3K/AKT pathways in cancer. *Trends Biochem Sci.* 2013;**38**:612–20.

87 Sabbah DA, Hajjo R, Sweidan K. Review on epidermal growth factor receptor (EGFR) structure, signaling pathways, interactions, and recent updates of EGFR inhibitors. *Curr Top Med Chem.* 2020;**20**:815–34.

88 Xu Y, Xu H, Li M, Wu H, Guo Y, Chen J, et al. KIAA1199 promotes sorafenib tolerance and the metastasis of hepatocellular carcinoma by activating the EGF/EGFR-dependent epithelial-mesenchymal transition program. *Cancer Lett.* 2019;**454**:78–89.

89 Basu S, Cheriyamundath S, Ben-Ze'ev A. Cell-cell adhesion: linking Wnt/beta-catenin signaling with partial EMT and stemness traits in tumorigenesis. *F1000Res.* 2018;**7**:F1000 Faculty Rev-1488.

90 Teeuwssen M, Fodde R. Wnt signaling in ovarian cancer stemness, EMT, and therapy resistance. *J Clin Med.* 2019;**8**:1658.

91 Chanmee T, Ontong P, Itano N. Hyaluronan: a modulator of the tumor microenvironment. *Cancer Lett*. 2016;**375**:20–30.

92 Koh Y, Kim D, Jung WJ, Ahn KS, Yoon SS. Revealing genomic profile that underlies tropism of myeloma cells using whole exome sequencing. *Int J Genomics*. 2015;**2015**:675379.

93 Rodrigues G, Hoshino A, Kenific CM, Matei IR, Steiner L, Freitas D, et al. Tumour exosomal CEMIP protein promotes cancer cell colonization in brain metastasis. *Nat Cell Biol*. 2019;**21**:1403–12.

94 Tsuji S, Nakamura S, Yamada T, de Vega S, Okada Y, Inoue S, et al. HYBID derived from tumor cells and tumor-associated macrophages contribute to the glioblastoma growth. *Brain Res*. 2021;**1764**:147490.

95 Usami S, Wagatsuma M, Fukuoka H, Suzuki H, Tsukada K, Nishio S, et al. The responsible genes in Japanese deafness patients and clinical application using invader assay. *Acta Otolaryngol*. 2008;**128**:446–54.

96 Hosoya M, Fujioka M, Okano H, Ogawa K. Distinct expression pattern of a deafness gene, KIAA1199, in a primate cochlea. *Biomed Res Int*. 2016;**2016**:1781894.

97 Marella M, Jadin L, Keller GA, Sugarman BJ, Frost GI, Shepard HM. KIAA1199 expression and hyaluronan degradation colocalize in multiple sclerosis lesions. *Glycobiology*. 2018;**28**:958–67.

98 Yang X, Qiu P, Chen B, Lin Y, Zhou Z, Ge R, et al. KIAA1199 as a potential diagnostic biomarker of rheumatoid arthritis related to angiogenesis. *Arthritis Res Ther*. 2015;**17**:140.

99 Shinohara T, Izawa T, Mino-Oka A, Mori H, Iwasa A, Inubushi T, et al. Hyaluronan metabolism in overloaded temporomandibular joint. *J Oral Rehabil*. 2016;**43**:921–8.

100 Shimizu H, Shimoda M, Mochizuki S, Miyamae Y, Abe H, Chijiiwa M, et al. Hyaluronan-binding protein involved in hyaluronan depolymerization is up-regulated and involved in hyaluronan degradation in human osteoarthritic cartilage. *Am J Pathol*. 2018;**188**:2109–19.

101 Yoshida H, Nagaoka A, Komiya A, Aoki M, Nakamura S, Morikawa T, et al. Reduction of hyaluronan and increased expression of HYBID (alias CEMIP and KIAA1199) correlate with clinical symptoms in photoaged skin. *Br J Dermatol*. 2018;**179**:136–44.

102 Yoshida H, Aoki M, Komiya A, Endo Y, Kawabata K, Nakamura T, et al. HYBID (alias KIAA1199/ CEMIP) and hyaluronan synthase coordinately regulate hyaluronan metabolism in histamine-stimulated skin fibroblasts. *J Biol Chem*. 2020;**295**:2483–94.

103 Krishnamurthy VK, Stout AJ, Sapp MC, Matuska B, Lauer ME, Grande-Allen KJ. Dysregulation of hyaluronan homeostasis during aortic valve disease. *Matrix Biol*. 2017;**62**:40–57.

104 Puperi DS, O'Connell RW, Punske ZE, Wu Y, West JL, Grande-Allen KJ. Hyaluronan hydrogels for a biomimetic spongiosa layer of tissue engineered heart valve scaffolds. *Biomacromolecules*. 2016;**17**: 1766–75.

105 Galarza TE, Taquez Delgado MA, Mohamad NA, Martin GA, Cricco GP. Histamine H4 receptor agonists induce epithelial-mesenchymal transition events and enhance mammosphere formation via Src and TGF-beta signaling in breast cancer cells. *Biochem Pharmacol*. 2020;**180**:114177.