GALNTs: master regulators of metastasis-associated epithelial-mesenchymal transition (EMT)?

Ellie-May Beaman, David R. F. Carter and Susan A. Brooks*

Department of Biological & Medical Sciences

Oxford Brookes University, Headington, Oxford, OX3 0BP, UK.

*Corresponding author: Department of Biological & Medical Sciences, Oxford Brookes University, Headington, Oxford, OX3 0BP, UK. Email sbrooks@brookes.ac.uk

Running head: GALNTs regulate epithelial-mesenchymal transition

Keywords: Epithelial-mesenchymal transition (EMT) / GalNAc O-glycosylation / metastasis / Tn antigen / UDP-N-acetylgalactosamine:polypeptide N-acetylgalactosaminyltransferases (GALNTs)

Statement of conflict of interest: The authors are Founders of, and hold shares in, MetaGuideX Ltd., a spinout company based at Oxford Brookes University with an interest in glycans as cancer biomarkers.

Abstract

In humans, the UDP-N-α-D galactosamine:polypeptide N-acetylgalactosaminyltransferases family (ppGalNAc-Ts, GalNAc-Ts or GALNTs) comprises 20 isoenzymes. They are responsible for the initial synthesis of α-GalNAc1,3-O-Ser/Thr, or Tn antigen, at initiation of mucin type O-linked glycosylation. This structure is normally extended by the further sequential action of glycosytransferases to build more complex linear or branched O-linked structures, but in cancers it is frequently left unelaborated, and its presence is often associated with poor patient prognosis. Altered levels of GALNT expression or distribution have also been extensively reported in a wide range of cancers. These changes would be predicted to result in marked alterations in GalNAc O-linked glycosylation, including altered levels of site specific O-linked glycosylation and changes in the glycan structures formed, including, potentially, exposure of truncated O-glycans such as Tn antigen. Many reports have demonstrated that altered levels of specific GALNTs have prognostic significance in cancers, or shown that they are associated with changes in cell behaviour, including proliferation, migration, invasion or growth and metastasis in animal models. We have previously reviewed how deregulation of GALNTs in several epithelial cancers is a feature of different stages metastasis. Here we consider evidence that changes in GALNT expression, and therefore consequent alterations in GalNAc O-linked glycosylation, may directly influence molecules implicated in aspects of epithelial-mesenchymal transition (EMT), a fundamental aspect of cancer metastasis, during which epithelial cancer cells lose their cell-cell junctions, apical-basal polarity and adhesive interactions with basement membrane and become mesenchymal, with a spindle-shaped morphology and increased migratory capacity.

The extensive family of UDP-N-α-D galactosamine:polypeptide N-acetylgalactosaminyltransferases

In humans, there are 20 isoenzymes in the UDP-N-α-D galactosamine:polypeptide N-acetylgalactosaminyltransferases family, also known as ppGalNAc-Ts, GalNAc-Ts or GALNTs. They are responsible for the initial synthesis of α-GalNAc1,3-O-Ser/Thr, or Tn antigen, at initiation of mucin type O-linked glycosylation. This initial structure is then usually extended by the further sequential action of glycosytransferases to build more complex linear or branched structures (see Brooks et al., 2002; Brockhausen and Stanley, 2017 for detail). The GALNTs are by far the largest known glycosyltransferase enzyme family, and stand in contrast to the situation in other types of protein glycosylation where just one or two enzymes are responsible for the addition of the glycan to the amino acid of the polypeptide (Clausen and Bennett, 1996; Bennett et al., 2012). Members of the family are designated GALNT1 to GALNT20 following the order in which they were first described. Table I gives a summary of the accession numbers and gene locations of the GALNTs, plus a note of any alternative terms by which they were reported before the CAZy nomenclature (http ://afmb.cnrs-mrs.fr/fr/CAZY/) was agreed. It is believed that all GALNTs, except for GALNT4, arose through evolution from an ancestral GALNT gene, with nine members of the GALNT family being present in *Caenorhabditis elegans*, fourteen in Drosophila species, twenty-one in fish, and nineteen in mice. They are present in all metazoans, but not in yeasts or plants (Hashimoto et al., 2009). Figure 1 illustrates the proposed phylogenetic tree for the GALNT gene family based upon subfamily clusters (reviewed by Bennett et al., 2012).

All GALNTs except GALNT20 share a type II membrane structure, illustrated in Figure 2, composed of a short N-terminal cytoplasmic tail (Smith and Lupashin, 2008), a hydrophobic membrane-spanning domain, a stem region (90 amino acids, for GALNT1, -T13, -T16; 170 amino acids for GALNT3, -T8, -T15; and 470 amino acids for GALNT5) (Bennett et al., 2012), a luminal catalytic domain (about 230 amino acids long) (Paulson and Colley, 1989) and, unique to the GALNTs, a C-terminal ricin-like lectin domain (about 120 amino acids long) which has a binding specificity for α-GalNAc monosaccharides

(Hazes, 1996; Imberty et al., 1997). GALNT10 has an additional binding pocket in the catalytic domain (Kubota et al., 2006). GALNT20 does not have the lectin domain, but does share the other structural components of the family (Bennett et al., 2012).

The reasons why so many GALNTs are required for initiation of mucin type O-glycosylation is not completely understood. However, it does not appear to represent redundancy or a duplication of function (Hassan et al., 2000). GALNTs show different, albeit partly overlapping, substrate specificities, kinetic properties and catalytic activities, have preferences for the proline in the glycosylation site being in the -1 or-3 position, exhibit differential temporal, spatial and tissue-specific expression (described below), and some GALNTs will only glycosylate proteins already glycosylated by a GalNAc (reviewed by Bennett et al., 2012; Wandall et al., 2021). This allows for fine differential regulation of O-glycosylation distribution and density. GALNTs may be categorised based on their *in vitro* substrate preferences as: “early”, which act on naked peptides/proteins, such as GALNT1 and -T2; “mid”, which have a preference for peptides neighbouring prior GalNAc glycosylated peptides, such as GALNT4; and “late”, which strictly act on areas of high GalNAc glycosylation, such as GALNT7 and -T10 (Ten Hagen et al., 1999; Hassan et al., 2000; Pratt et al., 2004; Bennett et al., 2012, and reviewed by Ten Hagen et al., 2003; de las Rivas et al., 2019).

Li et al. (2012) reported a sub-family, which they designated the Y subfamily, and which are only expressed in vertebrates. It comprises GALNT8, -T9, -T17 and -T18, and all share a common mutation in the catalytic domain, such that they have lost their transferase activity. The remaining sixteen GALNTs are termed the W subfamily. GALNT18 has been demonstrated to modulate the catalytic activity of W family GALNTs. On the basis of this, Li et al. (2012) postulate that Y subfamily GALNTs may act as molecular chaperones to W subfamily GALNTs in a similar manner to the way in which Cosmc acts as a chaperone to core 1 synthase during extension of O-linked glycans, thus facilitating a subtle level of regulation of GalNAc O-glycosylation.

There are some inconsistencies in the literature regarding the overall tissue expression patterns of the GALNTs. Toba et al. (2000) reported that GALNT1, -T2, -T4 and -T8 are ubiquitously expressed, whilst GALNT3, -T5, -T6, -T7 and -T9 are tissue-specific, with GALNT9 expression being limited to brain. Zhang et al. (2003) reported that GALNT1, -T2, -T4, -T7, -T8, and T12 are expressed ubiquitously, whilst GALNT3, -T6, -T9 and -T11 are restrictively expressed. Peng et al. (2010) reported that GALNT1, -T2, -T4, -T7 and -T8 are ubiquitously expressed, whilst GALNT3, -T5, -T6, -T9, -T10, -T11, -T12, -T13, -T14, -T15, -T16 and -T20 are more selectively expressed; they were unable to detect expression of GALNT16, -T18 and -T19. Bennett et al. (2012) reported that GALNT1, -T2 and -T7 show ubiquitous expression, while GALNT3, -T4, -T6, -T11, -T12, -T14, -T16 and -T18 show a more limited distribution, and GALNT5, -T8, -T9, -T10, -T13, -T15, -T17, -T19 and -T20 are highly restricted in expression. They also commented that the GALNT20 reported by Peng et al. (2010) was mis-identified and was actually GALNT17. However, all published reports agree that GALNT1 and -T2 are distributed ubiquitously in human tissues (Homa et al., 1993; White et al., 1995; Toba et al., 2000; Zhang et al., 2003; Peng et al., 2010; Bennett et al., 2012). The literature also indicates that some GALNTs are normally widely distributed across human tissues, whilst others show more limited and discrete patterns of expression, summarised in Table II.

Relocation of the GALNTs in cancer and exposure of cancer-associated Tn antigen

Since the GALNTs are responsible for initiation of mucin-type GalNAc O-linked glycosylation, alterations in their distribution, expression or activity would be predicted to result in alterations in this type of glycosylation including, but not exclusively, the synthesis of truncated O-linked glycans such as Tn antigen. The GALNTs are normally located throughout the Golgi apparatus (Röttger et al., 1998). However, it has been reported that in cancers GALNTs are often relocated to the endoplasmic reticulum, termed the GALA (GalNAc-T activation) pathway (Gill et al., 2011; Chia et al., 2014; 2016; 2019; Gomes et al., 2017). This results in the synthesis of proteins glycosylated with unelaborated Tn

antigen. Tn antigen is almost always further extended in normal adult tissues and therefore ordinarily present at very low levels. However, unelaborated Tn antigen is a recognised feature of around 50-90% of all cancers, where its exposure has often been reported to be associated with metastatic potential and poor patient prognosis (for example, Springer, 1984; Ju et al., 2008; Julien et al., 2012). Recently, Rømer et al. (2021) reported Tn immunolabelling of a more than 700 normal and cancer tissue samples, and found high levels in cancers arising from breast, colorectum and pancreas, but negligible or undetectable immunoreactivity in normal tissues or cancers derived from the central nervous system or mesenchymal tissue.

Exposure of Tn antigen could also be attributed to factors other than GALNT relocation to the endoplasmic reticulum. There is evidence, for example, that the pH and redox potential within the Golgi apparatus influences the location of glycosyltransferases, and therefore patterns of O-linked glycosylation (Axelsson et al., 2001; Hassinen et al., 2019). There has also been a long standing hypothesis that Tn antigen exposure in cancer could result from a failure of normal chain extension (Ju et al., 2011), arising from lack of expression of the core 1 synthase, β1,3 galactosyltransferase (C1GALT1), or loss of its molecular chaperone, Cosmc (Ju and Cummings, 2002; Ju et al., 2008). There is some evidence for this; for example, Thomas et al. (2019) showed that knockdown of Cosmc resulted in accumulation of Tn antigen and increased epithelial mesenchymal transition (EMT) a feature of metastatic cancer cells, described later, in pancreatic cancer cells. Some studies have provided evidence that epigenetic modification through hypermethylation of the Cosmc promoter results in silencing of C1GALT1 and therefore synthesis of unelaborated Tn antigen (Mi et al., 2012; Jiang et al., 2018; Xu et al., 2020a). However, core 1 is frequently observed in cancers that also feature Tn antigen (Baldus et al., 2000; Yu, 2007; Kölbl et al., 2016; Sindrewicz et al., 2016) which, if the presence of Tn antigen were due to the absence of Cosmc or core 1 synthase, would be predicted not to occur. Alternatively, decreased expression of β1,3GlcNAc-T (C3GnT), the transferase responsible for addition of a GlcNAc monosaccharide to Tn antigen to form core 3, GlcNAcβ1,3GalNAc-O-Ser/Thr, can also lead to enhanced exposure of Tn and sialyl Tn (An et al.,

2007; Vavasseur et al., 1994). Finally, both UDP-Gal and the UDP-Gal transporter are necessary for chain extension, and their absence can also result in increased levels of Tn antigen (Deutscher et al., 1986; Kingsley et al., 1986; Stanley et al., 1991). There are therefore several mechanisms by which unelaborated Tn antigen may feature in cancers, but evidence suggests that GALNT relocation to the endoplasmic reticulum, the GALA pathway, is prominent among them.

Alterations in levels of GALNTS in cancer

That exposure of Tn antigen is a feature of many cancers and is associated with poor patient prognosis is well established, and relocation of GALNTs within the secretory pathway appears to be the most important mechanism underlying its appearance. Altered levels of GALNT expression or distribution, determined by molecular approaches and by immunohistochemistry, respectively, have also been extensively reported in a broad range of cancers. These include colorectal (Kohsaki et al., 2000; Shibao et al., 2002; Guo et al., 2004; Guda et al., 2009; Abulí et al., 2011; Clarke et al., 2012; Lavrsen et al., 2018; Ubillos et al., 2018; Tang et al., 2021), gastric (Onitsuka et al., 2003; Gomes et al., 2009; Gao et al., 2013; He et al ., 2014; Xu et al., 2020b; Liu et al., 2016), renal (Kitada et al., 2013), breast (Berois et al., 2006a; Freire et al., 2006; Patani et al., 2008; Wu et al., 2010; 2021a; 2021b; Park et al., 2010; Liu et al., 2020; Huang et al., 2022), pancreatic (Li et al., 2011; Taniuchi et al., 2011; Tarhan et al., 2016; Caffrey et al., 2019), gallbladder (Miyahara et al., 2004), prostate (Landers et al., 2005), oesophageal (Ishikawa et al., 2005), lung adenocarcinoma (Gu et al., 2004; Song et al., 2020; Wang et al., 2021) and non small cell lung cancer (Dosaka-Akita et al., 2002), cervical (Peng et al., 2012; Cao et al., 2020), bladder (Ding et al., 2012), liver (Wu et al., 2011; Huang et al., 2015), oral (Lin et al., 2014; Harada et al., 2016), ovarian (Wang et al., 2014; Yang et al., 2016; Lin et al., 2017; Li et al., 2021), glioma (Sun et al., 2019), chronic lymphocytic leukaemia (Libisch et al., 2014) and neuroblastoma (Berois et al., 2006b; 2013; Ho et al., 2014). A summary of these studies is given in Table III.

Clearly, such changed expression of GALNTs would be predicted to result in both altered levels of site specific O-linked glycosylation, and, potentially, changes in the glycan structures formed, including exposure of truncated O-glycans such as Tn antigen. Several of the studies cited (indicated in Table III) attempted to correlate increased or decreased levels of specific GALNTs with altered GalNAc O-glycosylation in cancer cells. Most commonly, this was approached by using lectin pulldown from cell lysates using the lectin *Vicia villosa* agglutinin (VVA) which has a reported specificity for Tn antigen, VVA labelling of specific target proteins on western blots, and/or VVA lectin cytochemistry. However, results of such analyses should be interpreted with a degree of caution in terms of them indicating synthesis of proteins glycosylated with Tn antigen. Firstly, whilst VVA has been extensively employed to identify Tn antigen, it is also reported to recognise other glycan structures (Tollefsen and Kornfeld, 1983; Iskratsch et al., 2009). Some studies, for example, that of Libisch et al. (2014) on GALNT11 in chronic lymphocytic leukaemia, addressed this by also employing monoclonal antibodies against Tn antigen. Moreover, many of the studies cited interpreted VVA binding to either permeabilised cells or to proteins from cell lysates as evidence of increased GalNAc O-glycosylation, but this may not mean that the mature target proteins carry Tn antigen. Detarya et al. (2020), in their study of GALNT5 in cholangiocarcinoma, illustrated that whilst VVA labelling could be localised in permeabilised cells to the peri-nuclear space in cells over expressing GALNT5 – presumably representing enhanced initiation of GalNAc O-glycosylation in the Golgi apparatus or endoplasmic reticulum - no cell surface VVA labelling of mature glycoproteins was detectable. This suggests that whilst these may be O-glycosylated by the GALNT, they did not feature unelaborated Tn antigen.

As indicated in Table III, many studies have reported differences in levels of GALNTs between normal in comparison with cancer tissues. In many cases, differences were shown to have prognostic significance in clinical studies, or to be associated with altered cell behaviour *in vitro* – commonly, proliferation, migration in wound healing or transwell assays, or invasion through matrigel – or altered tumorigenicity in animal models. For example, increased expression of GALNT1 has been

reported to predict poorer survival in liver cancer patients (Huang et al., 2015). Reports suggest that high GALNT2 levels predict poor prognosis in some cancer types, such as lung adenocarcinoma (Wang et al., 2021), and glioma (Sun et al., 2019). However, they are associated with good prognosis in other cancers, such as neuroblastoma, where immunopositivity has been reported to be associated with early stage (Ho et al., 2014), reduced levels to be associated with advanced disease and shorter recurrence free survival in gastric cancer (Liu et al., 2016), and with vascular invasion and recurrence in hepatocellular carcinoma (Wu et al., 2011). Similarly, increased GALNT3 appears to be associated with good prognosis in some cancer types, and poor prognosis in others. Elevated levels of GALNT3 have been reported to predict improved survival in colorectal (Shibao et al., 2002) and gastric cancers (Onitsuka et al., 2003), and low levels to predict poor prognosis in lung adenocarcinomas and non small cell lung cancers (Gu et al., 2004; Dosaka-Akita et al., 2002). Conversely, GALNT3 positive oesophageal cancers have been reported to be more likely to have metastasised (Ishikawa et al., 2005) and GALNT3 positive renal carcinomas to be larger, feature vascular invasion and have poorer survival (Kitada et al., 2013). GALNT3 positive oral squamous carcinomas have been reported to be poorly differentiated, feature vascular and lymphatic invasion and be more likely to recur (Harada et al., 2016), and high GALNT3 levels to predict shorter progression free survival in epithelial ovarian cancers (Wang et al., 2014). Wu et al. (2021a;b) report increased GALNT4 to be associated with enhanced recurrence free survival in breast cancer. Reduced levels of GALNT5 have been reported to predict poor prognosis in gastric cancer (He et al., 2014). High levels of GALNT6 have been reported to predict improved survival in pancreatic (Li et al., 2011) and colorectal cancer (Ubillos et al., 2018), but poorer survival in breast (Liu et al., 2020), lung (Song et al., 2020) and ovarian cancer (Lin et al, 2017). Increased expression in bone marrow samples from breast cancer patients predicted subsequent recurrence in lymph node negative patients (Freire et al, 2006). Overexpression of GALNT8 has been reported to be correlated with poor patient prognosis in colorectal cancer (Tang et al., 2021) and, the opposite, downregulation, to be associated with poor outcome in breast cancer (Huang et al., 2022). Increased expression of

GALNT9 has been reported to be associated with improved survival in neuroblastoma (Berois et al., 2013). GALNT10 immunopositivity has been reported to be associated with high grade (Gao et al., 2013) and with lymph node and distant metastases and poor prognosis in gastric cancer (Xu et al., 2020b). GALNT12 levels have been reported to be lower in colorectal cancer cases with metastases (Guo et al., 2004). Increased expression of GALNT13 in patients with advanced stage neuroblastoma correlated with poor outcome (Berois et al., 2006b). High levels of GALNT14 in ovarian cancer were associated with advanced stage (Yang et al., 2016) but with low histological grade in breast cancer (Wu et al. 2010).

There is therefore strong evidence of alterations in expression of most members of the GALNT family in a wide range of cancers, which, in some instances, have been shown to have prognostic significance. However, the situation is not as simple as, for example, an increase in GALNT activity being associated with more or less aggressive cancer behaviour. We have previously reviewed how deregulation of GALNTs in several epithelial cancers is a feature of different stages of the metastatic cascade (Beaman and Brooks, 2014). Recently, we have become aware that changes in GALNT expression, and therefore consequent alterations in GalNAc O-linked glycosylation, may directly influence molecules implicated in aspects of epithelial-mesenchymal transition (EMT), a fundamental aspect of cancer metastasis.

Epithelial mesenchymal transition in cancer metastasis

Metastasis, the dissemination of cancer cells from a primary tumour mass to form secondary lesions at distant anatomical sites, accounts for about 90% of cancer patient deaths (Sporn, 1996; Germanov et al., 2006). It is accepted that the most common route by which cancer cells disseminate is through the blood circulation; however, cancer cells may also invade new sites by direct extension into adjacent tissues or may enter the lymphatics and colonise local and regional lymph nodes. There remains some controversy as to whether lymph node metastases represent a

dead end for metastasising cancer cells, and serve only as a clinically useful proxy marker for the likelihood of further disseminated disease, or whether instead they are a staging post such that tumours in the lymph nodes can themselves seed further distant metastases (reviewed by Weinberg, 2007). Either way, metastasis results from a cascade of events which commonly includes initial tumour angiogenesis, invasion of local basement membrane and local stroma, intravasation and hematogenous /lymphatic dissemination, attachment to vascular endothelium and extravasation, and formation of a secondary tumour. Whilst a detailed description of metastasis goes beyond the scope of the current review, the key stages are illustrated in Figure 3 and it is well reviewed in the literature (for example, Brooks et al., 2010; Chaffer and Weinberg, 2011; 2015; Welch and Hurst, 2019). Throughout the metastatic cascade, tumour cells alter their phenotype and behavioural characteristics to a remarkable degree. These highly plastic changes in cancer cell morphology and motility involve reactivation of the developmental program termed epithelial-mesenchymal transition (EMT) and its reversion, mesenchymal-epithelial transition (MET) (reviewed by Micalizzi et al., 2010; Chaffer et al., 2016; Kim et al., 2017; Lu and Kang, 2019).

EMT is the trans-differentiation of epithelial cells to mesenchymal cells (Hay, 1995), and involves a number of morphological and molecular changes, summarised in Figure 4. During EMT, epithelial cells lose their characteristic apical-basal polarity through the dissolution of tight junctions. This permits the fraternisation of apical and basolateral membrane components, resulting in the end-to-end polarity associated with mesenchymal cells (Greenburg and Hay, 1982). The structural integrity of epithelial sheets dissipates through elimination of gap junctions and adherens junctions (Shtutman et al., 2006). Cell adhesion molecules, including E-cadherin and epithelial-specific integrins that mediate cell-cell and cell-basement membrane connections, respectively, are replaced with N-cadherin and $\alpha v\beta 6$ integrin, which impart a transient adhesion (Nakajima et al., 2004). In the process of becoming mesenchymal, cells gain resistance to anoikis (Frisch et al., 2013). The actin cytoskeleton of epithelial cells is re-structured and substituted with stress fibres, whilst cytokeratin intermediate filaments are exchanged for vimentin (Micalizzi et al., 2010), resulting in a

morphological change in cells from cuboidal to spindle-shaped (Boyer et al., 1989). Finally, the basement membrane is degraded through the secretion of proteases, and the motile mesenchymal cells are able to invade the extracellular matrix and surrounding stroma (Savagner, 2001).

Transcriptional regulation of tumour-associated epithelial-mesenchymal transition

The process of EMT is reviewed in depth by Lamouille et al. (2014). The process is induced in the primary tumour microenvironment by autocrine/paracrine secretion of a number of growth factors, cytokines and extracellular matrix proteins (Tse and Kalluri, 2007; Martin et al., 2010; Jing et al., 2011) produced in response to hypoxia (Jiang et al., 2011) and inflammation (López-Novoa and Nieto, 2009; Zhou et al., 2012). Binding of these EMT-inducing signals to their receptors on the cell surface triggers a myriad of intracellular signalling networks which act pleiotropically to induce EMT transcription factors.

A highly simplified diagram illustrating transcriptional regulation of some of the principle pathways involved in EMT signalling is given in Figure 5. Briefly, at the cell surface, binding of bone morphogenic proteins (BMPs) to their receptors BMPR1/2 activates intracellular SMAD proteins which initiate transcriptional induction of EMT (Frey et al., 2020). Alternatively, upon binding of the cytokine transforming growth factor-beta (TGF-β) to its receptors TGF-βR1/2, intracellular signalling can occur through either SMAD-dependant or -independent pathways (Aashaq et al., 2022). A SMAD-independent pathway is activation of the SRC homology 2 domain-containing-transforming A (SHCA), which creates a docking site for growth factor receptor-bound protein 2 (GRB2) and son of sevenless (SOS) and initiates the RAS-RAF-MEK1/2-ERK1/2 pathway (Zhang, 2009). Binding of growth factors such as epidermal growth factor (EGF) (Ackland et al., 2003), hepatocytpe growth factor (HGF) (Nagai et al., 2011) and insulin-like growth factor (IGF) (Cevenini et al., 2018) to tyrosine kinase receptors can activate the RAS-RAF-MEK1/2-ERK1/2 pathway as well as the P13K-Akt signalling cascade. Binding of Delta or Jagged to Notch receptor activates hairy and enhancer of split 1 (HES1),

inhibits phosphatase and TENsin homolog (PTEN) which is an inhibitor of the PI3K/Akt pathway (Ranganathan et al., 2011). Binding of Wnt to Frizzled (FZD) inhibits glycogen synthase kinase-3β (GSK3β), which in turn stabilises β-catenin so that it can translocate to the nucleus and initiate EMT transcription factors (Wu and Pan, 2010). Akt from the PI3K-Akt signalling pathway can also inhibit GSK3β. Binding of Sonic hedgehog (SHH) to Patched (PTCH) results in a change in structural conformation of the transmembrane protein thereby releasing Smoothened (SMO), which is then able to activate the glioma-associated (GLI) family (Riaz et al., 2019).

These EMT-inducing signals are responsible for the stimulation of EMT core regulators, which are transcription factors that regulate EMT effectors; that is, the proteins that identify cells as being either epithelial or mesenchymal, such as E-cadherin and cytokeratin (epithelial markers) and vimentin and N-cadherin (mesenchymal markers) (reviewed by Tsai and Yang, 2013). EMT core regulators fall into three groups: (1) SNAI1 (snail family transcriptional repressor 1, SNAIL) and SNAI2 (snail family transcriptional repressor 2, SLUG)) (Savagner et al., 1997; Cano et al., 2000; Bolós et al., 2003), (2) ZEB1 and ZEB2 (Comijn et al., 2001; Eger et al., 2005; Xiong et al., 2012) and (3) TWIST1 and TWIST2 (Lo et al., 2007; Vesuna et al., 2008; Casas et al., 2011; Fang et al., 2011). Once these transcription factors have been expressed and activated, each can act pleiotropically to orchestrate the EMT process (reviewed by Sánchez-Tilló et al., 2012).

Changes in GALNT expression or activity may influence cancer cell behaviour through altered O-glycosylation of many different proteins

Widespread and diverse alterations in glycosylation of myriad different molecules is reported during EMT, and have sometimes been shown to have a regulatory or functional role, as recently reviewed by Pucci et al., (2021). In the following sections, we will focus specifically on the evidence that GALNTs, through their role in regulating the initiation of O-linked GalNAc glycosylation of a wide variety of proteins, potentially influence EMT, and this may in part explain the association between

changes in GALNT levels and cancer prognosis, described previously. It is important to make the point, however, that a hugely diverse range of proteins passing through the secretory pathway could potentially be modified by the action of GALNTs, and it would be naïve to postulate that the only effect of cellular changes in GALNT levels is directly on EMT. Indeed, since many O-glycosylated proteins are extracellular matrix proteins, signalling molecules, receptors and adhesion molecules, changes in GALNTs might be predicted to influence a range of cellular mechanisms involved in cancer progression by influencing the stability or properties of these molecules (reviewed by Chia et al., 2016).

Using 'SimpleCell' technology, Steentoft et al. (2013) identified around 3000 potential GalNAc O-glycosylation sites in 600 proteins produced by cancer cells, which illustrates the potential complexity. Lavrsen et al. (2018) carried out RNAseq analysis of colon cancer cells engineered to express or not express GALNT6 and identified more than 100 genes that were associated with GALNT6 expression, and more than half were involved in cell cycle control, indicating a role in cell proliferation. Moreover, they identified a range of cell adhesion molecules that were upregulated with GALNT6 knockdown, including P-cadherin, cadherin-17, CD44 and versicon, and GALNT6 expression decreased cell adhesion in a cell dissociation assay. But in their study, they found no evidence that cells expressing GALNT6 had altered levels of EMT-associated markers (E-cadherin, N-cadherin, vimentin, fibronectin) or inducers (SNAIL, SLUG, TWIST or TGF-β), implying that GALNT6 expression alone did not induce EMT in colon cancer cells. This study therefore provides clear evidence that expression of GALNT6, which is a feature of colon cancers but not of normal colonic epithelium, resulted in loss of cell differentiation and adhesion, and increased cell proliferation, which might contribute to malignant progression and indirectly influence EMT, but did not directly contribute to EMT. Added to this, there are many studies that implicate GALNTs in contributing to cell growth (for example, Peng et al., 2012; Li et al., 2012; Ding et al., 2012; Wang et al., 2014) and immune evasion (for example, Adachi et al., 1997; Gaziel-Sovran et al 2011), which could contribute to tumorigenesis by mechanisms other than by inducing EMT.

It is also important to recognise that few of the studies cited in the following sections explored site specific O-glycosylation mediated by the GALNTs. Our broader knowledge of the precise function of site specific O-glycosylation of specific proteins still remains quite limited, largely because of how technically complex it is to investigate. However, established functions include protecting proteins from proteolytic cleavage, stabilising domain structures, and modulating binding affinity and receptor function. This rapidly expanding field was recently reviewed by Wandall et al. (2021), and they emphasised how occupancy of a single O-glycosylation site may offer highly specific and dynamic regulation of protein function. Some representative examples of where function of site specific O-glycosylation has been well studied and is established include GALNT3 glycosylation of fibroblast growth factor 23 contributing to familial tumoral calcitosis, and GALNT2 glycosylation of angiopoeitin like-3 protein contributing to high serum triglyceride and high-density lipoprotein cholesterol levels, implicated in cardiovascular disease (reviewed by Chia et al., 2016; Wandall et al., 2021).

GALNTs alter GAlNAc O-glycosylation of signal induction receptors and signal transduction pathways involved in EMT

The strongest evidence of GALNT expression regulating EMT comes from reports where GALNTs have been demonstrated to act on EMT signal induction receptors. These studies have often demonstrated altered levels of GALNTs in clinical samples of various cancers, then explained this relationship by demonstrating the effects of GALNT overexpression and/or knockdown on cell proliferation, and on cell migration and invasion. Many studies have also sought to demonstrate that GALNT expression results in either a general increase in GalNAc O-glycosylation of proteins or of increased GalNAc O-glycosylation of specific EMT-related molecules.

The EGF/HGF/IGF-PI3K/Akt signalling pathway

Huang et al. (2015) explored GALNT1 and hepatocellular carcinoma. They first demonstrated that GALNT1 levels were often elevated in hepatocellular carcinoma, and that this correlated with decreased overall patient survival. They showed that GALNT1 O-glycosylates EGFR. Whilst they detected no decrease in Tn antigen on EGFR after GALNT1 knockdown using VVA labelling, treating the cells with benzyl-α-GalNAc , which inhibits O-linked chain extension, did result in an increase in VVA labelling – signifying that GALNT1 knockdown resulted in a decrease in more extended O-glycans on EGFR. GALNT1 knockdown also increased EGFR degradation and decreased its activation, leading to decreased hepatocellular carcinoma cell migration and invasion.

Similar findings have been presented in several studies of other GALNTs, and in a range of cancers. Wu et al. (2011) found that GALNT2 was downregulated in hepatocellular carcinomas in comparison to adjacent normal tissue, and downregulation was most common in patients with vascular invasion and evidence of recurrence. Over-expression of GALNT2 in cell lines supressed their invasive and migratory potential and the growth of xenografts in mice. VVA pulldown and western blot demonstrated that EGFR was O-glycosylated by GALNT2. Lin et al. (2014) demonstrated that GALNT2 was detectable by immunohistochemistry in high-grade oral squamous carcinoma tissues, and was localised at the invasive edge of tumours. GALNT2 overexpression in oral squamous carcinoma cells enhanced their invasion and migration. Furthermore, GALNT2 knockdown decreased VVA binding to EGFR and EGF-induced phosphorylation of EGFR and Akt, and overexpression enhanced them, indicating that GALNT2 O-glycosylates EGFR. Sun et al. (2019) also explored GALNT2, this time in glioma. Again, they demonstrated that GALNT2 immunopositivity was associated with high grade, poor prognosis gliomas. GALNT2 over expression in glioma cell lines resulted in increased phosphorylation of EGFR, and an increase in its glycosylation by Tn antigen, determined by western blot and VVA lectin pull-down assays, and influenced the expression of p21, cyclin-dependent kinase 4 (CDK4), cyclinD1, MMP-2 and-9 through the EGFR/PI3K/Akt pathway, as well as enhancing cell

invasion and migration. Knockdown achieved the opposite effects. Moreover, GALNT2 silencing resulted in less aggressive tumours in a mouse model, with reduced cellular proliferation and MMP2 expression. Confusingly, Hu et al. (2018) reported contradictory results in gastric cancer cells. They reported that GALNT2 knockdown resulted in decreased Tn antigen on EGFR, determined by VVA pulldown, but also enhanced, rather than reduced, its phosphorylation, and also the phosphorylation of Akt. This was associated with enhanced migration and invasion, which was abrogated by the addition of an EGRF inhibitor or an Akt inhibitor.

Xu et al. (2021) reported that upon GALNT3 knockdown in ovarian cancer cells, levels of phosphorylated PI3K and phosphorylated Akt decreased, suggesting that GALNT3 activated the PI3K/Akt pathway. Moreover, they demonstrated that the microRNA miR-378a-3p suppressed the expression of PI3K and pAkt, which could be rescued by GALNT3. Further, they showed that miR-378a-3p binds to the 3'UTR of GALNT3 which degrades GALNT3 and suppresses its translation. Taken together, these findings suggest a role of GALNT3 in activation of the EMT-associated PI3K/Akt signalling pathway.

Lin et al. (2017) demonstrated that elevated levels of GALNT6 correlate with lymph node and distant metastasis and reduced survival in ovarian endometrioid and clear cell carcinomas, and provided evidence that this was through altered EGFR signalling, in studies reminiscent of those cited previously exploring GALNT2. Knocking down GALNT6 in epithelial ovarian cancer cells significantly reduced their mesenchymal-like properties of cellular migration and invasion, and did so by decreasing phosphorylation of EGFR, whilst VVA pull down confirmed that GALNT6 glycosylated EGFR with Tn antigen. Liu et al. (2020) also investigated GALNT6, this time in breast cancer. They showed that survival of breast cancer patients with tumours expressing high levels of GALNT6 was shorter than those with tumours expressing low levels. In cell lines, knock down of GALNT6 inhibited PI3K/Akt signalling, and this was further suppressed by the knockdown of alpha 2 macroglobulin (α2M); the opposite was observed when GALNT6 and α2M were overexpressed. They

describe α2M as a substrate of GALNT6 and propose that GALNT6 mediates O-glycosylation, and therefore stabilisation, of the α2M-activated PI3K-Akt signalling pathway in breast cancer. VVA pulldown indicated that GALNT6 knockdown reduced the amount of exposed Tn antigen present on breast cancer cells. Breast cancer cells over-expressing GALNT6 showed enhanced migratory and invasive behaviour, and those where GALNT6 was knocked down showed reduced migratory and invasive potential, indicating that GALNT6 can drive these EMT-like behaviours through the mechanisms identified.

Cao et al. (2020) demonstrated that miR-125a-5p, which has been reported to be reduced in cancer tissues in comparison to normal, directly binds to the GALNT7 gene, repressing its expression. Moreover, they provided evidence that GALNT7 promotes cervical cancer cell proliferation and invasion in a matrigel assay by activating the EGFR/PI3K/Akt pathway. This was confirmed in a mouse xenograft model where GALNT7 suppression resulted in smaller tumours. It would be reasonable to assume that the effect of GALNT7 may be through O-glycosylation and phosphorylation of EGFR, as reported above, although this was not directly assessed in the study, and, in common with the studies cited above, classical markers of EMT, such as E-cadherin and vimentin were not examined.

Supporting evidence for the role of GALNT family members in this mechanism was also provided by Li et al. (2021), who demonstrated that downregulation of GALNT14 in ovarian cancer cells significantly decreased Tn antigen glycosylation of EGFR, detected through VVA pull down. They showed that this reduced the stability of EGFR, shortened its half-life and subsequently inhibited the EMT-related EGFR/PI3K/Akt signalling pathway, although the authors did not specifically address EMT or the migratory or invasive behaviour of the cells. Instead their focus was on identifying that GALNT14 contributes to enhancing the cytotoxic action of cisplatin through glycosylating and stabilising EGFR.

Comparable evidence has been reported regarding GALNT glycosylation of type 1 insulin-like growth factor receptor (IGFR) influencing the activation of the IGF/PI3K/Akt signalling pathway. Ho et al. (2014) reported that increased GALNT2 immunoreactivity predicted early stage and favourable prognosis in neuroblastoma patients, and GALNT2 over expression resulted in smaller tumours in mice. Overexpression of GALNT2 in neuroblastoma cells resulted in increased O-glycosylation of IGFR with Tn antigen, as evidenced by VVA pull down. This prevented IGFR dimerization and subsequent downstream signalling, resulting in reduced cell growth, and reduced migration and invasion in transwell and matrigel assays.

In contrast, Liu et al. (2016) demonstrated that GALNT2 protein and mRNA were reduced in human gastric cancer samples, and that reduced GALNT2 was associated with advanced disease and shorter recurrence free survival. Knockdown of GALNT2 in gastric cancer cells increased their proliferation, and also their migration and invasion in transwell and matrigel assays. GALNT2 knockdown also promoted metastasis in a mouse model. GALNT2 knockdown in gastric cancer cells enhanced phosphorylation and reduced Tn antigen on hepatocyte growth factor receptor (HGFR, tyrosine protein kinase Met, or c-Met), as evidenced by reduced VVA binding. The authors propose that GALNT2 supresses the malignant potential of gastric cancer cells through exposure of truncated O-glycans, including Tn antigen, dysregulating HGFR activation and inhibiting HGF/PI3K/Akt signalling.

The Notch/HES1-PTEN-PI3K/Akt signalling pathway

Wang et al. (2021) reported that GALNT2 expression in lung adenocarcinoma correlated with poor patient prognosis. Moreover, GALNT2 increased the ability of lung cancer cells to form colonies, migrate in a wound healing assay and invade matrigel by activating the Notch/HES1-PTEN-PI3K/Akt signalling pathway. They observed that Notch1, -2, -3 and -4 are co-expressed with GALNT2 in lung cancer and confirmed using western blot that GALNT2 expression elevated Notch1 and -3 levels. The downstream molecules of the Notch signalling pathway, HES1 and -5, were also elevated. HES1

inhibits the tumour suppressor PTEN, which actively blocks the PI3K/Akt signalling pathway. They further demonstrated that by knocking down GALNT2, PTEN expression increased and the downstream signalling molecules decreased. The opposite was found with a cell line that was engineered to overexpress GALNT2. It is reasonable to assume that GALNT2 O-glycosylation stabilises Notch, subsequently activating HES1, thereby inhibiting PTEN and activating the PI3K/Akt signalling pathway. Indeed, the authors speculate that GALNT2 acts to glycosylate and therefore stabilise Notch, but note that they plan to investigate this in further work. Wu et al. (2021a) reported conflicting findings in relation to GALNT4 and breast cancer. They showed that analysis of patient datasets indicated that high GALNT4 was associated with enhanced, rather than reduced, recurrence free survival. Knockdown in breast cancer cell lines increased cell proliferation and the authors concluded, through comparison of differentially expressed genes in knockdown versus wild type cells, that this was through activation of Notch signalling pathway.

The TGF-β signalling pathway

Maupin et al. (2010) specifically sought to identify glycosylation-related genes associated with EMT by examining pancreatic cancer cells lines in three different model systems: (a) direct induction of EMT by treatment with TGF-β (b) classifying the morphologies of cancer cell lines as epithelial-like or mesenchymal-like and (c) an assay of cell migration where cells were seeded into a defined area and allowed to migrate for a set time period. Expression data from a list of 587 genes involved in glycosylation were analysed. The authors identified a number of 'glycogenes' that correlated with EMT-related behaviours in these models, including GALNT2 and GALNT10, which were increased in TGF-β-induced EMT, and GALNT13 which showed higher expression in migratory cells. What role(s) GALNTs play in TGF-β-induced EMT is unclear, however it would be reasonable to speculate that they would be performing a stabilising role, thereby preventing proteolysis, through O-glycosylating proteins that are involved in the TGF-β signalling cascade. Evidence from Freire-de-Lima et al. (2011),

described later, indicates that O-glycosylation of the extracellular matrix protein fibronectin may also be implicated.

It is interesting that Wu et al. (2021b) reported converse findings when they examined the relationship between GALNT4 and EMT in breast cancer cells. Having demonstrated that GALNT4 was upregulated in good prognosis, luminal type breast cancer and associated with favourable recurrence free survival, they showed that knockdown of GALNT4 was associated with mesenchymal cell morphology and over-expression with an epithelial phenotype. They also correlated GALNT4 knockdown with a decrease in classical epithelial marker E-cadherin, and increase in mesenchymal markers N-cadherin and SNAIL, and overexpression with the opposite. These findings correlated with cell behaviour in wound healing and invasion assays. Upon treatment with TGF-β, EMT was enhanced in GALNT4 knockdown cells and supressed in cells where GALNT4 was over-expressed. GALNT4 activity was associated with O-glycosylation of TGF-β type I and type II receptors, including their glycosylation with Tn antigen assessed by VVA binding, and this led to decreased TGF-β dimerization and activity. Using an HPLC-based *in vitro* O-glycosylation system followed by experiments where specific residues were mutated, they were able to identify Ser 31, corresponding to the extracellular domain of TGF-β type II receptor, as being the O-glycosylation site of GALNT4. Intriguingly, GALNT4 was shown to have a higher affinity for the potential glycosylation site on the TGF-β type II receptor site than the ubiquitously expressed GALNT2. These findings suggest a further layer of complexity in the regulation of the EMT process exerted by GALNTs.

BMP4/SMAD1 EMT activation

Huang et al. (2022) provided evidence of GALNT8 being involved in regulation of EMT activation in breast cancer through the BMP4/SMAD1 pathway. The focus of their study was to seek evidence of GALNT8 having a role in regulating estrogen receptor-α (ER-α) through BMP signalling in breast cancer. Levels of GALNT8 were assessed in publicly available breast cancer datasets, and a clear

relationship between GALNT8 downregulation, poor clinical outcome and lower ER-α expression was apparent. The relationship between downregulation of GALNT8 and reduced levels of ER-α was confirmed by immunohistochemistry of tissue samples, and in a series of breast cancer cell lines. Manipulating expression of GALNT8 resulted in changes in proliferative rate of the cells, with overexpression reducing proliferation, and knockdown enhancing it. That this was mediated through ER signalling was demonstrated through use of an ER inhibitor. However, EMT-related cell behaviour or markers of EMT were not directly assessed. Introducing a BMP/SMAD activator or inhibitor resulted in alterations in GALNT8 levels. VVA pulldown and assessment of the effects of mutating potential O-glycosylation sites of BNPR1A indicated O-GalNAc glycosylation of BNPR1A by GALNT8. Taken together, the results suggest that GALNT8 O-glycosylates BNPR1A, thus activating the BMP4/SMAD1 pathway.

Xu et al. (2020b) focussed on a different player in the BMP4/SMAD1 pathway, HOXD13. They showed that high levels of GALNT10 in gastric cancer clinical samples, as determined by qRT-PCR, were associated with lymph node and distant metastases and poor prognosis. They demonstrated that knocking down GALNT10 in gastric cancer cells reduced their proliferative activity and their invasive potential in a transwell migration assay. Bioinformatics analysis that suggested that GANT10 could combine with HOXD13 were confirmed in a dual-luciferase reporter gene experiment, which verified that HOXD13 could directly target GALNT10. These data suggest that that GALNT10 O-glycosylates and stabilises HOXD13, and that this promotes EMT-like migratory behaviour, although EMT was not specifically assessed in this study. Confusingly, contradictory evidence was provided by Xue et al. (2021) who showed that HOXD13 prevents BMP4/SMAD1 EMT activation, and therefore inhibits EMT, in prostate cancer cells. HOXD13 was demonstrated to directly bind to the *SMAD1* promoter and suppress SMAD1 transcription. HOXD13 overexpression blocked BMP4/SMAD1-induced EMT, assessed by protein levels and gene expression of E-cadherin and fibronectin after BMP4 treatment, and knockdown of HOXD13 resulted in prostate cancer cells becoming more invasive and migratory. Moreover, in analysis of patient datasets,

decreased HOXD13 expression was significantly associated with early recurrence. However, this study did not consider GALNTs or HOXD13 glycosylation directly.

GALNTs directly or indirectly modulate EMT effectors

A number of studies provide evidence that GALNT activity directly or indirectly modulates the expression of classical EMT effectors, including the epithelial marker E-cadherin and the mesenchymal marker vimentin. Song et al. (2020) showed that GALNT6 immunopositivity was associated with lymph node metastasis and poor prognosis in lung adenocarcinoma. They went on to demonstrate that knockdown of GALNT6 in lung adenocarcinoma cell lines increased the expression of E-cadherin and reduced the expression of N-cadherin and SLUG; conversely, overexpression of GALNT6 decreased E-cadherin and increased N-cadherin and SLUG. These changes were associated with changes in the cells migratory and invasive potential in a wound healing assay, and GALNT6 silencing resulted in reduced metastatic potential in a mouse xenograft model. By immunoprecipitation studies followed by VVA labelling of western blots, they showed that GALNT6 O-glycosylated and stabilised glucose-regulated protein 78 (GRP78), an endoplasmic reticulum-based molecular chaperone which promotes the folding and oligomerization of proteins. Over expression of GRP78 caused relocation of GALNT6 to the endoplasmic reticulum, potentially increasing initiation of GalNAc O-glycosylation and synthesis of Tn antigen through the GALA pathway, and also increased phosphorylation of ERK1/2. Knockdown of GALNT6 decreased expression of both GRP78 and ERK1/2, thus contributing to modulation of MEK1/2/ERK1/2 signalling.

Raghu et al. (2019) demonstrated that loss of GALNT3 was associated with a mesenchymal phenotype in breast cancer cells and in trophoblast stem cells. In trophoblast stem cells, which were the main focus of the study, GALNT3 knockdown was associated with reduced E-cadherin and increased vimentin, as well as other EMT markers that were investigated, and induction of EMT transcription factors TWIST, ZEB1, and SLUG. It also increased motility and invasiveness of cells. Re-

expression of GALNT3 restored the epithelial phenotype. GALNT3 was demonstrated by VVA labelling to O-glycosylate E-cadherin with Tn antigen, and knockdown resulted in E-cadherin being retained it in the Golgi apparatus, rather than expressed at the cell surface, thus contributing to EMT. Interestingly, Balcik-Ercin et al. (2018) reported that ZEB2 also binds to, and represses, GALNT3.

Ectonucleotide pyrophosphatase 1 (ENPP1) normally plays a role in regulating cell signalling response to insulin. Marucci et al. (2013) reported that GALNT2 knockdown in human liver cancer cells increased ENPP1 expression, decreased insulin-induced insulin receptor beta subunit phosphorylation, and also reduced Akt phosphorylation and therefore signalling. They demonstrated that GALNT2 bound to ENPP1-3'UTR mRNA and inhibited ENPP1 transcript and protein expression. The authors speculate that GALNT2 may also alter the O-glycosylation of ENPP1, but they did not directly explore this in their study. Hu et al. (2019) provided evidence that ENPP1 also functions to induce an EMT phenotype in lung cancer cells. Knocking down ENPP1 repressed E-cadherin and vimentin. Consistent with this, Detarya et al. (2020) demonstrated that knocking down GALNT5 inhibited EMT in cholangiocarcinoma cells and reduced their proliferation, invasion and migration, by inactivating Akt/ERK signalling.

GALNTs influence glycosylation of EMT-associated mucins

Overexpression of mucins (MUCs) has been demonstrated to promote tyrosine kinase signalling, activate growth and survival pathways, including the Wnt/β-catenin EMT signalling pathway, and induce loss of epithelial cell polarity (reviewed by Kufe, 2009). They are therefore of interest as attractive targets for the development of anti-cancer antibodies and drug inhibitors. MUCs are a large family of highly O-glycosylated glycoproteins (Corfield, 2015; Dhanisha et al., 2018). Several members MUC family, including MUC1, -4 and -13, are reported to be frequently over-expressed in cancers and are related to poor prognosis and metastasis (for example, Taylor-Papadimitriou et al.,

1999; Gendler, 2001; Carraway et al., 2009; Chauhan et al., 2009; Sousa et al., 2016; He et al., 2017; Kumari et al., 2018; Syrkina et al., 2019). It is well accepted that dense clustering of O-glycans on mucins contributes to their structural support and protection from degradation, and aberrant glycosylation of mucins, particularly with truncated O-glycans such as Tn antigen, have been reported to influence their stability and subcellular localisation (Altschuler et al., 2017).

Ponnusamy et al. (2010) provided clear evidence that one member of the family, MUC4, was able to upregulate N-cadherin, and thereby drive metastasis in ovarian cancer cells through EMT. Expressing MUC4 in ovarian cancer cells resulted in a dramatic change in their morphology, from adhesive and epithelial to spindle-shaped with reduced cell-cell contact and the presence of obvious filopodia, lammelopodia and microspikes. E-cadherin was decreased, whilst N-cadherin, vimentin, TWIST1, TWIST2 and SNAIL were increased. They went on to demonstrate that MUC4 expression activated Akt and ERK1/2 signalling, and that ERK1/2 phosphorylation also upregulated expression of MMP-9. Induction of MUC4 resulted in ovarian cancer cells being more metastatic in a mouse model. Whilst this study did not address MUC glycosylation, it clearly demonstrates the potential for members of the MUC family to induce EMT. Evidence from other studies have directly related GALNT and MUC expression to EMT. These studies all suggest a regulative role of GALNT expression, where upregulation of a specific GALNT acts to glycosylate and stabilise a member of the MUC family, resulting in altered expression of cell adhesion molecules and the subsequent transition towards a more mesenchymal-like phenotype and dissemination from the primary tumour mass.

A good example is the study by Wang et al. (2014) who identified increased levels of GALNT3 in epithelial ovarian cancer tissues in comparison to normal, where it was associated with shorter recurrence free survival. Knockdown in epithelial ovarian cancer cells led to reduced proliferation, invasion and migration. They carried out gene expression profiling and network and pathway analysis and found that numerous genes and pathways implicated in ovarian tumorigenesis, invasion and metastasis were downregulated upon GALNT3 suppression, while some tumour suppressor

genes were induced. GALNT3 downregulation also resulted in reduced MUC1 protein expression and they hypothesised that this was due to reduced O-glycosylation resulting in its destabilization and subsequent proteolysis of MUC1 protein due to lack of GALNT3 glycosylation activity, although this was not directly demonstrated. GALNT3 knockdown was also associated with an increase in the epithelial associated cell adhesion molecules, β-catenin and E-cadherin after GALNT3 knockdown, and the authors concluded that these molecules are normally supressed by MUC1 in cancer, supporting a role for GALNT3/MUC1 in promoting cancer cell invasion through EMT.

Park et al. (2010) made similar observations concerning GALNT6 in breast cancer. Immunohistochemistry showed that GALNT6 was strongly detectable in breast cancers but not in normal breast. Knockdown of either GALNT6 or MUC1 in breast cancer cells resulted in a change in their morphology to a more epithelial, rounded shape and enhanced proliferation. Moreover, knockdown of GALNT6 resulted in a reduction in MUC1 protein and over-expression to an increase, indicating that they are co-expressed. Knockdown of GALNT6 also resulted in an increase in epithelial markers E-cadherin and β-catentin. An *in vitro* GalNAc transferase assay and VVA lectin blotting both confirmed that GALNT6 O-glycosylates MUC1. The authors propose that upregulation of GALNT6 stabilises MUC1 through its glycosylation activity, and MUC1 then interacts cell adhesion molecules such as β-catenin and E-cadherin, resulting in the anti-adhesive effect. The same group (Tarhan et al., 2016) reported that knocking down GALNT6 in pancreatic cancer cells reduced the expression of MUC4 and EGFR2. It also resulted in an alteration in cell morphology from elongated and mesenchymal-like to epithelial-like, and switched cadherin expression from P-cadherin to E-cadherin, and reduced cell invasiveness in a transwell assay. GALNT6 was demonstrated to O-glycosylate MUC4 in an *in vitro* glycosylation assay, presumably resulting in its stabilisation, and the Thr residue within a peptide fragment that was the predominant target of GALNT6-mediated O-glycosylation was identified by MALDI-TOF mass spectrometry. It had previously been reported that MUC4 knockdown enhanced formation of an E-cadherin/β-catenin complex resulting in the

downregulation of the Wnt/β-catenin EMT signalling pathway in pancreatic cancer cells (Zhi et al., 2014).

Consistent with these observations, Wang et al. (2013) reported that knockdown of GALNT14 in epithelial ovarian cancer cells resulted in a morphological change from elongated and mesenchymal-like to rounded and epithelial-like, and the migratory capacity of the cells in a wound healing and transwell migration assay also decreased. They demonstrated that GALNT14 O-glycosylates MUC13 by assessing VVA binding to MUC13 in western blots before and after GALNT13 knockdown, and proposed that this effect was through reduced O-glycosylation, and subsequent destabilisation, of the MUC13 protein by GALNT14, in a manner analogous to that proposed regarding MUC1 and MUC4. Moreover, they demonstrated that an inhibitor of ERK1/2 supressed the expression of GALNT14 and reduced Tn antigen levels as assessed by VVA binding.

To add a further level of complexity, Caffrey et al. (2019) demonstrated that in pancreatic cancer, the MUC1 cytoplasmic tail binds to regulatory elements in the GALNT5 gene, thus downregulating its expression, and moreover that levels of MUC1 and GALNT5 were inversely correlated in pancreatic cancer tissues. This suggests that in some instances there may be a reciprocal regulatory process occurring between the GALNTs and MUCs that may also contribute to the EMT process.

GALNTs modulate the extracellular matrix and MMPs, as well as EMT markers

A number of studies have shown that GALNT activity, in addition to driving EMT through the various mechanisms described above, can also have a parallel effect on modifying the cancer-related stroma, thus potentially further facilitating cancer cell motility and invasion. GALNT6, for example, has been shown in several studies, cited earlier, to drive EMT in epithelial ovarian cancer (Lin et al., 2017), pancreatic cancer (Tarhan et al., 2016) and breast cancer (Park et al., 2010). GALNT6 has also been shown to modulate the breast cancer-associated stroma. Expression of fibronectin, a key

extracellular matrix glycoprotein, has been shown to be increased in EMT-associated processes such as wound healing (reviewed by Lenselink, 2015) and embryonic development (de Almeida et al., 2016) and is a prominent feature of breast cancer stroma (Park and Schwarzbauer, 2014; Li et al., 2019). In order for fibronectin fibrils to be assembled, cellular contractile forces are required, which stretch compact fibronectin dimers thereby exposing a cryptic heparin-binding region. Crucially, this domain binds many soluble EMT growth factors, including TGF-β1 (Griggs et al., 2017). Park et al. (2011) demonstrated that GALNT6 O-glycosylates, and thereby stabilises, fibronectin thus preventing its degradation and presumably allowing assembly into fibrils and subsequent exposure of the growth factor binding site. They also reported that overexpression of GALNT6 in a normal breast epithelial cell line resulted in a change in morphology from rounded and epithelial-like to elongated and mesenchymal-like, accompanied by a reduction in the epithelial marker E-cadherin and increase in the mesenchymal marker N-cadherin. Overexpression of GALNT6 also resulted in a disruption in the formation of an acinar morphology when the cells were grown in three dimensional culture, and invasive behaviour in Matrigel. GALNT6 expression is clinically associated with poor prognosis in breast cancer (Liu et al., 2020). These results are supported by those of Freire-de-Lima (2011), described briefly earlier, who set out to specifically examine GalNAc O-linked glycosylation during TGF-β-induced EMT, this time in prostate cancer cells. They identified that under TGF-β stimulation, oncofetal fibronectin was O-glycosylated at a specific Thr in its IIICS domain, and this was associated with a change from epithelial to fibroblastic morphology, enhanced cell motility, decreased expression of the epithelial cell marker, E-cadherin, and enhanced expression of mesenchymal markers. Moreover, knockdown of GALNT6 and GALNT3 inhibited TGF-β-induced EMT up-regulation of oncofetal fibronectin and the EMT process.

Yang et al., (2016) demonstrated that GALNT14 suppression, mediated by miR-125a, was associated with reduced invasive capacity of ovarian cancer cells, and also reduced MMP-2 and -9 activity. Consistent with this, GALNT14 was shown to be upregulated in clinical samples, and this correlated

with reduced levels of miR-125a, and also with advanced disease stage. Consistent with these observations, Huanna et al. (2015) showed that over-expression of GALNT14 in breast cancer cells increased their motility and their ability to invade extracellular matrix, and GALNT14 knockdown had the opposite effect. Increased GALNT14 was also associated decreased mRNA and protein levels of the epithelial marker E-cadherin, and increased levels of the mesenchymal markers N-cadherin and vimentin, as well as VEGF and TGF-β, and the opposite was seen in GALNT14 knockdown cells. The shift to a more mesenchymal and invasive phenotype was associated with increased MMP-2 activity. Using a luciferase assay, the authors attempted to determine whether GALNT14 upregulates MMP-2 by activating its transcription, and to identify the possible AP-1 and E1F-1 sites in the GALNT14 promoter. Results indicated that the AP-1 element is critical in achieving induction of the MMP-2 promoter, whilst E1F-1 is not. Zuo et al. (2018) went on to demonstrate, through VVA labelling of western blot, that GALNT14 glycosylates multiple proteins, including EGF-containing fibulin-like extracellular matrix protein 2 (EFEMP2) with Tn antigen and stabilises it, and this increases the invasive capacity of breast cancer cells.

Relocation of GALNT1 to the endoplasmic reticulum in liver cancer cells through the GALA pathway, described earlier, was reported by Nguyen et al. (2017) to result in increased GalNAc O-glycosylation of MMP-14, which was determined using a number of methods. These included VVA pulldown and western blot, and metabolic labelling with an azide-modified analogue of GalNAc, GalNAz. Also, to determine the presence of more extended O-glycans, VVA pulldown after neuraminidase treatment of MMP-14 from cells where Cosmc had been knocked down was performed, in addition to pulldown and western blot analysis using lectins from peanut and *Datura stramonium*. The GALNT1-glycosylated form of MMP-14 facilitated increased tumour growth and metastatic ability of the cancer cells in a mouse model, associated with enhanced extracellular matrix degradation and increased capacity to invade into nearby organs, in comparison to cancer cells expressing the Golgi apparatus located form. The authors confirmed previous reports that MMP14 was O-glycosylated within the hinge domain because mutants deficient in this domain showed reduced glycosylation.

Taken together, these observations point to a complex interplay between regulation of glycosylation, through GALNT expression, controlling EMT and also mediating the integrity of the stroma surrounding the tumour, further contributing to invasion and metastasis.

Conclusions

Glycosylation of proteins may influence their binding properties, activity, expression level, and stability. The interplay of the multiple members of the GALNT family contribute to subtle and fine control of GalNAc O-glycosylation of proteins. There are numerous reports in the literature of the normal fine control of GALNTs being dysregulated in a wide range of cancers, and this would be predicted to dramatically alter GalNAc O-glycosylation of a diverse range of proteins, potentially influencing cancer cell biology significantly. Changed levels of GALNT expression or activity have been clearly and repeatedly associated with altered cancer prognosis in humans and tumour behaviour in animal models. Moreover, the clear functional effect of over-expression or knockdown of specific GALNTs in cancer cells leading to profound changes in their behaviour, including migration and invasion, clearly points to their glycosylating, and thereby altering the function of, molecules implicated in epithelial-mesenchymal transition, the process by which stationary, adhesive epithelial cancer cells undergo dramatic transformation to become mesenchymal, motile and invasive. In many instances, studies have provided strong evidence of global changes in GalNAc O-glycosylation as a result of GALNT expression, and often that specific molecular players in the complex process of EMT are glycosylated by the GALNT under scrutiny. However, what is often missing is analysis of the glycan structures produced and, critically, of site-specific glycosylation. What is clear is that GALNTs are master regulators of GalNAc O-linked glycosylation and that their dysregulation in cancer has a profound effect on cancer cell behaviour through, amongst other mechanisms, altering glycosylation and therefore properties of the myriad players in the complex mechanisms of EMT. Further exploration of the specifics of their action and effect may open up new avenues for a better

understanding of cancer biology and therefore, potentially, development if more effective interventions.

List of Abbreviations

BMP: Bone morphogenic protein

BMPR: Bone morphogenic protein receptor

C1GALT1: β1,3 galactosyltransferase or core 1 synthase

C3GnT: β1,3GlcNAc-T, core 3 synthase

EFEMP2: epidermal growth factor-containing fibulin-like extracellular matrix protein 2

EGF: Epidermal growth factor

EGFR: Epidermal growth factor receptor

EMT: Epithelial mesenchymal transition

ENPP1: Ectonucleotide pyrophosphatase 1

ERK: Extracellular signal-regulation kinase

FZD: Frizzled

Gal: galactose

Glc: glucose

GALA: GalNAc-T activation

GalNAc: N-acetylgalactosamine

GALNT: UDP-N-α-D galactosamine:polypeptide N-acetylgalactosaminyltransferase (also known as ppGalNAc-T, GalNAc-T)

GlcNAc: N-acetylglucosamine

GLI: glioma-associated

GRB2: growth factor receptor-bound protein 2

Hes 1: Hairy and enhancer of split 1

HGF: hepatocyte growth factor

HOXD13: Homeobox D13

IGF: insulin-like growth factor

α2M: Alpha-2 macroglobin

MALDI-TOF

MET: Mesenchymal epithelial transition

MUCs: Mucins

PI3K: Phosphoinositide 3-kinase

PTCH: patched

PTEN: Phosphatase and TENsin homolog

q-PCR: quantitative polymerase chain reaction

RT-PCR: reverse transcription polymerase chain reaction

Ser: Serine

SHCA: SRC homology 2 domain-containing-transforming A

SHH: sonic hedgehog

SMAD: Suppressor of mothers against decapentaplegic

SMO: smoothened

SNP: Single nucleotide polymorphism

SOS: son of sevenless

TGF-β: Transforming growth factor-beta

Thr: Threonine

Tn: Thomsen nouvelle

UDP: Uridine diphosphate

VVA: *Vicia villosa* agglutinin

Wnt: Wingless-related integration site

ZEB1: Zinc finger E-box-binding homeobox 1

ZEB2: Zinc finger E-box-binding homeobox 2

References

Aashaq S, Batool A, Mir SA, Beigh MA, Andrabi KI, Shah ZA. 2022. TGF-β signaling: A recap of SMAD-independent and SMAD-dependent pathways. *J Cell Physiol.* 237(1):59-85.

Abulí A, Fernández-Rozadilla C, Alonso-Espinaco V, Muñoz J, Gonzalo V, Bessa X, Gonzálex D, Clofent J, Cubiella J, Morillas JD, Rigau J, Latorre M, Fernández-Bañares F, Peña E, Riestra S, Payá A, Jover R, Xicola RM, Llor X, Carvajal-Carmona L, Villanueva CM, Moreno V, Piqué JM, Carracedo A, Castells A, Andreu M, Ruiz-Ponte C, Castellví-Bel S. 2011. Case-study for colorectal cancer genetic susceptibility in EPICOLON: previously identified variants and mucins. *BMC Cancer*. 11:339.

Adachi T, Hinoda Y, Nishimori I, Adachi M, Imai K. 1997. Increased sensitivity of gastric cancer cells to natural killer and lymphokine-activated killer cells by antisense suppression of *N*-acetylgalactosaminyltransferase. *J Immunol*. 159: 2645-2651.

Altschuler Y, Kinlough CL, Poland PA, Bruns JB, Apodaca G, Weiscz OA, Hughey RP. 2017. Clathrin-mediated endocytosis of MUC1 is modulated by its glycosylation state. *Molecular Biology of the Cell* 11(3): 819-831.

An G, Wei B, Xia B, McDaniel JM, Ju T, Cummings RD, Braun J, Xia L. 2007. Increased susceptibility to colitis and colorectal tumors in mice lacking core 3-derived O-glycans. *J Exp Med.* 204:1417-1429.

Axelsson MAB, Karlsson NG, Steel DM, Ouwendijk J, Nilsson T, Hansson GC. 2001. Neutralization of pH in the Golgi apparatus causes redistribution of glycosyltransferases and changes in the O-glycosylation of mucins. *Glycobiology* 11(8): 633-644.

Balcik-Ercin P, Cetin M, Yalim-Camci I, Odabas G, Tokay N, Sayan AE, Yagci T. 2018. Genome-wide analysis of endogenously expressed ZEB2 binding sites reveals inverse correlations between ZEB2 and GalNAc-transferase GALNT3 in human tumors. *Cell Oncol.* 41:379-393.

Baldus SE, Zirbes TK, Hanisch FG, Kunze D, Shafizadeh ST, Nolden S, Mönig SP, Schneider PM, Karsten U, Thiele J, Hölscher AH, Dienes HP. 2000. Thomsen-Friedenreich antigen presents as a prognostic factor in colorectal carcinoma: A clinicopathologic study of 264 patients. *Cancer*. 88:1536-1543.

Beaman E-M, Brooks SA. 2014. The extended ppGalNAc-T family and their functional involvement in the metastatic cascade. *Histol Histopathol.* 29:293-304.

Bennett EP, Hassan H, Clausen H. 1996. cDNA cloning and expression of a novel human UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase, GalNAc-T3. *J Biol Chem.* 271:17006-17012.

Bennett EP, Hassan H, Mandel U, Mirgorodskaya E, Roepstorff P, Burchell J, Taylor-Papadimitriou J, Hollingsworth MA, Merkx G, van Kessel AG, Eiberg H, Steffensen R, Clausen H. 1998. Cloning of a human UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase that complements other GalNAc-transferases in complete O-glycosylation of the MUC1 tandem repeat. *J Biol Chem.* 273:30472-30481.

Bennett EP, Hassan H, Mandel U, Hollingsworth MA, Akisawa N, Ikematsu Y, Merkx G, van Kessel AG, Olofsson S, Clausen H. 1999a. Cloning and characterization of a close homologue of human UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase-T3, designated GalNAc-T6. Evidence for genetic but not functional redundancy. *J Biol Chem.* 274:25362-25370.

Bennett EP, Hassan H, Hollingsworth MA, Clausen H. 1999b. A novel human UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase, GalNAc-T7, with specificity for partial GalNAc-glycosylated acceptor substrates. *FEBS Lett.* 460:226-30.

Bennett EP, Mandell U, Clausen H, Gerken TA, Fritz TA, Tabak LA. 2012. Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. *Glycobiology*. 22:736-756.

Berois N, Mazal D, Ubillos L, Trajtenberg F, Nicolas A, Sastre-Garau X, Magdelenat H, Osinaga E. 2006a. UDP-N-acetyl-d-galactosamine: polypeptide N-acetylgalactosaminyltranferase-6 as a new immunohistochemical breast cancer marker. *J Histochem Cytochem.* 54:317-328.

Berois N, Blanc E, Ripoche H, Mergui X, Trajtenberg F, Cantais S, Barrois M, Dessen P, Kågedal B, Bénard J, Osinaga E, Raguénez G. 2006b. ppGalNAc-T13: a new molecular marker of marrow involvement in neuroblastoma. *Clin Chem*. 52:1701-1712.

Berois N, Gattolliat C, Barrios E, Capandeguy L, Douc-Rasy S, Valteau-Couanet D, Bénard J, Osinaga E. 2013. GALNT9 gene expression is a prognostic marker in neuroblastoma patients. *Clin Chem.* 59:225-233.

Bolós V, Peinado H, Pérez-Moreno MA, Fraga MF, Esteller M, Cano A. 2003. The transcription factor Slug represses E-cadherin and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. *J Cell Sci.* 116:499-511.

Boyer B, Tucker G, Valles A, Franke W, Thiery J. 1989. Rearrangements of desmosomal and cytoskeletal proteins during the transition from epithelial to fibroblastoid organization in cultured rat bladder carcinoma cells. *J Cell Biol.* 109:1495-1509.

Brockhausen I, Stanley P. 2017. Chapter 10. O-GalNAc glycans. In Essentials of Glycobiology [Internet]. 3rd edition. https://www.ncbi.nlm.nih.gov/books/NBK453030/

Brooks SA, Dwek MV, Schumacher U. 2002. Chapter 5 O-linked (mucin-type) glycoproteins. In Functional and molecular glycobiology. Oxford, UK: Bios Scientific Publishers Ltd. Pp 89-115.

Brooks SA, Lomax-Browne HJ, Carter TM, Kinch CE, Hall DMS. 2010. Molecular interactions in cancer cell metastasis. *Acta Histochem*. 112:3-25.

Caffrey T, Sagar S, Thomas D, Lewallen ME, Hollingsworth MA, Radhakrishnan P. 2019. The glycoprotein mucin-1 negatively regulates GalNAc transferase 5 expression in pancreatic cancer. *FEBS Lett.* 593:2751-2761.

Cano A, Pérez-Moreno MA, Rodrigo I, Locascio A, Blanco MJ, del Barrio MG, Portillo F, Nieto MA. 2000. The transcription factor Snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat Cell Biol.* 2:76-83.

Cao Q, Wang N, Ren L, Tian J, Yang S, Cheng H. 2020. miR-125a-5p post-transcriptionally suppresses GALNT7 to inhibit proliferation and invasion in cervical cancer cells via the EGFR/PI3K/AKT pathway. *Cancer Cell Int.* 20:117.

Carraway KL, Theodoropoulos G, Kozloski GA, Carothers Carraway CA. 2009. Muc4/MUC4 functions and regulation in cancer. *Future Oncol.* 5:1631-1640.

Casas E, Kim J, Bendesky A, Ohno-Machado L, Wolfe CJ, Yang J. 2011. Snail2 is an essential mediator of Twist1-induced epithelial-mesenchymal transition and metastasis. *Cancer Res*. 71:245-254.

Cevenini A, Orrù S, Mancini A, Alfieri A, Buono P, Imperlini E. 2018. Molecular Signatures of the Insulin-like Growth Factor 1-mediated Epithelial-Mesenchymal Transition in Breast, Lung and Gastric Cancers. *Int J Mol Sci,* 19(8):2411.

Chaffer CL, Weinberg RA. 2011. A perspective on cancer cell metastasis. *Science.* 331:1559-1564.

Chaffer CL, Weinberg RA. 2015. How does multistep tumorigenesis really proceed? *Cancer Discov.* 5:22-24.

Chaffer CL, San Juan BP, Lim E, Weinberg RA. 2016. EMT, cell plasticity and metastasis. *Cancer Metastasis Rev.* 35:645-654.

Chauhan SC, Vannatta K, Ebeling MC, Vinayek N, Watanabe A, Pandey KK, Bell MC, Koch MD, Aburatani H, Lio Y, Jaggi M. 2009. Expression and functions of transmembrane mucin MUC13 in ovarian cancer. *Cancer Res*. 69:765-774.

Cheng L, Tachibana K, Zhang Y, Guo JM, Kahori Tachibana K, Kameyama A, Wang H, Hiruma T, Iwasaki H, Togayama A, Kudo T, Narimatsu H. 2002. Characterisation of a novel human UDP-GalNAc transferase, pp-GalNAc-T10. *FEBS Lett.* 531:115-121.

Cheng L, Tachibana K, Iwasaki H, Kameyama A, Zhang Y, Kubota T, Hinima T, Tachibana K, Kudo T, Guo JM, Narimatsu H. 2004. Characterisation of a novel human UDP-GalNAc transferase, PP-GalNAc-T15. *FEBS Lett.* 566:17-24.

Chia J, Tham KM, Gill DJ, Bard-Chapeau EA, Bard FA. 2014. ERK8 is a negative regulator of O-GalNAc glycosylation and cell migration. *Elife*. 3:e01828

Chia J, Goh G, Bard F. 2016. Short O-GalNAc glycans: regulation and role in tumor development and clinical perspectives. *Biochim Biophys Acta.* 1860:1623-1639.

Chia J, Tay F, Bard F. 2019. The GalNAc-T activation (GALA) pathway: drivers and markers. *PLoS One.* 14: e0214118.

Clarke E, Green RC, Green JS, Mahoney K, Parfrey PS, Younghusband HB, Woods MO. 2012. Inherited deleterious variant in GALNT12 are associated with CRC susceptibility. *Hum Mutat.* 33:1056-1058.

Clausen H, Bennett EP. 1996. A family of UDP-GalNAc: polypeptide N-acetylgalactosaminyl-transferases control the initiation of mucin-type O-linked glycosylation. *Glycobiology*. 10:635-646.

Comijn J, Berx G, Vermassen P, Verschueren K, van Grunsven L, Bruyneel E, Mereel M, Huylebroeck D, van Roy F. 2001. The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Mol Cell.* 7:1267-1278.

Corfield AP. 2015. Mucins: a biologically relevant glycan barrier in mucosal protection. *Biochim Biophys Acta.* 1850:236-52.

de Almeida PG, Pinheiro GG, Nunes AM, Gonçalves AB, Thorsteinsdóttir S. 2016. Fibronectin assembly during early embryo development: A versatile communication system between cells and tissues. *Dev Dyn.* 245(4):520-35.

de Las Rivas M, Lira-Navarrete E, Gerken TA, Hurtado-Guerrero R. 2019. Polypeptide GalNAc-Ts: from redundancy to specificity. *Curr Opin Struct Biol*. 56:87-96.

Detarya M, Sawanyawisuth K, Aphivatanasiri C, Chuangchaiya S, Saranaruk P, Sukprasert L, Silsirivanit A, Araki N, Wongkham S, Wongkham C. 2020. The O-GalNAcylating enzyme GALNT5 mediates carcinogenesis and progression of cholangiocarcinoma via activation of AKT/ERK signalling. *Glycobiology.* 30:312-324.

Deutscher SL, Hirschberg CB. 1986. Mechanism of galactosylation in the Golgi-apparatus of a Chinese-hamster ovary cell mutant deficient in translocation of UDP-galactose across Golgi vesicle membranes. *J Biol Chem.* 261:96-100.

Dhanisha SS, Guruvayoorappan C, Drishya S, Abeesh P. 2018. Mucins: structural diversity, biosynthesis, its role in pathogenesis and as possible therapeutic targets. *Crit Rev Oncol Hematol.* 122:98-122.

Ding M, Wang H, Wang J, Zhan H, Zuo Y, Yang D, Liu J, Wang W, Ke C, Yan R. 2012. ppGalNAc T1 as a potential novel marker for human bladder cancer. *Asian Pac J Cancer Prev.* 13:5653-5657.

Dosaka-Akita H, Kinoshita I, Yamazaki K, Izumi H, Itoh T, Katoh H, Nishimura M, Matsuo K, Yamada Y, Kohno K. 2002. N-acetylgalactosaminyl transferase-3 is a potential new marker for non-small cell lung cancers. *Br .J Cancer.* 87:751-755.

Eger A, Aigner K, Sonderegger S, Dampier B, Oehler S, Schreiber M, Berx G, Cano A, Beug H, Foisner R. 2005. DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. *Onocogene*. 24:2375-2385.

Fang X, Cai Y, Liu J, Wang Z, Wu Q, Zhang Z, Yang CJ, Yuan L, Ouyang G. 2011. Twist2 contributes to breast cancer progression by promoting an epithelial-mesenchymal transition and cancer stem-like cell self-renewal. *Oncogene.* 30:4707-4720.

Freire T, Berois N, Sóñora C, Varangot M, Barrios E, Osinaga E. 2006. UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 6 (ppGalNAc-T6) mRNA as a potential

new marker for detection of bone marrow-disseminated breast cancer cells. *Int J Cancer.* 119:1383-1388.

Freire-de-Lima L, Gelfenbeyn K, Ding Y, Mandel U, Clausen H, Handa K, Hakomori SI. 2011. Involvement of O-glycosylation defining oncofetal fibronectin in epithelial-mesenchymal transition process. *Proc Natl Acad Sci USA*. 108:17690–17695.

Frey P, Devisme A, Schrempp M, Andrieux G, Boerries M, Hecht A. 2020. Canonical BMP Signaling Executes Epithelial-Mesenchymal Transition Downstream of SNAIL1. *Cancers (Basel).* Apr 21;12(4):1019.

Frisch SM, Schaller M, Cieply B. 2013. Mechanisms that link the oncogenic epithelial-mesenchymal transition to suppression of anoikis. *J Cell Sci.* 126:21-29.

Gao Y, Liu Z, Feng J, Sun Q, Zhang B, Zheng W, Ma W. 2013. Expression pattern of polypeptide N-acetylgalactosaminyltransferase-10 in gastric carcinoma. *Oncol Lett.* 5:113-116.

Gaziel-Sovran A, Segura MF, Di Micco R, Collins M K, Hanniford D, Vega-Saenz de Miera E, Rakus JF, Dankert JF, Shang S, Kerbel RS, Bhardwaj N, Shao Y, Darvishian F, Zavadil J, Erlebacher A, Mahal LK, Osman I, Hernando E. 2011. miR-30b/30d regulation of GalNAc transferases enhances invasion and immunosuppression during metastasis. *Cancer Cell*. 20:104-118.

Gendler SJ. 2001. MUC1, the renaissance molecule. *J Mammary Gland Biol Neoplasia.* 6:339-53.

Germanov E, Berman J, Guernsey D. 2006. Current and future approaches for the therapeutic targeting of metastasis. *Int J Mol Med.* 18:2025-2036.

Gill DJ, Clausen H, Bard F. 2011. Location, location, location: new insights into O-GalNAc protein glycosylation. *Trends Cell Biol.* 21:149-158.

Gomes J, Mereiter S, Magalhães A, Reis CA. 2017. Early GalNAc O-glycosylation: pushing the tumor boundaries. *Cancer Cell*. 32:544-545.

Gomes J, Marcos NT, Berois N, Osinaga E, Magalhães A, Pinto-de-Sousa J, Almeida R, Gärtner F, Reis CA. 2009. Expression of UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase-6 in gastric mucosa, intestinal metaplasia, and gastric carcinoma, *J Histochem Cytochem.* 57: 79-86.

Greenburg G, Hay ED. 1982. Epithelia suspended in collagen gels can lose polarity and express characteristics of migrating mesenchymal cells. *J Cell Biol.* 95:333-339.

Griggs LA, Hassan NT, Malik RS, Griffin BP, Martinez BA, Elmore LW, Lemmon CA. 2017. Fibronectin fibrils regulate TGF-β1-induced epithelial-mesenchymal transition. *Matrix Biol.* 60-61:157-175.

Gu C, Oyama T, Osaki T, Li J, Takenoyama M, Izumi H, Sugio K, Kohno K, Yasumoto K. 2004. Low expression of polypeptide GalNAc N-acetylgalactosaminyl transferase-3 in lung adenocarcinoma: impact on poor prognosis and early recurrence. *Br J Cancer*. 90:436-442.

Guda K, Moinova H, He J, Jamison O, Ravi L, Natale L, Lutterbaugh J, Lawrence E, Lewis S, Willson JKV, Lowe JB, Wiesner GL, Parmigiani G, Barnholtz-Sloan J, Dawson DW, Velculescu VE, Kinzler KW, Papadopoulos N, Vogelstein B, Willis J, Gerken TA, Markowitz SD. 2009. Inactivating germ-line and somatic mutations in polypeptide N-acetylgalactosaminyltransferase 12 in human colon cancers. *Proc Natl Acad Sci USA.* 106:12921-12925.

Guo J, Chen H, Wang G, Zhang Y, Narimatsu H. 2004. Expression of UDP-GalNAc:polypeptide N-acetylgalactosaminyltrnasferase-12 in gastric and colonic cancer cell lines and in human colorectal cancer. *Oncology.* 67:271-276.

Guo JM, Zhang Y, Cheng L, Iwasaki H, Wang H, Kubota T, Tachibana K, Narimatsu H. 2002. Molecular cloning and characterisation of a novel member of the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase family, pp-GalNAc-T12. *FEBS Lett.* 524:211-218.

Hagen FK, Ten Hagen KG, Beres TM, Balys MM, VanWuyckhuyse C, Tabak LA. 1997. cDNA cloning and expression of a novel UDP-*N*-acetyl-D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase. *J Biol Chem.* 272(21):13843-13848.

Harada Y, Izumi H, Noguchi H, Kuma A. Kawatsu Y, Kimura T, Kitada S, Uramoto H, Wang K-Y, Sasaguri Y, Hijioka H, Miyawaki A, Oya R, Nakayama T, Kohno K, Yamada S. 2016 . Strong expression of polypeptide Nacetylgalactosaminyltransferase 3 independently predicts shortened disease-free survival in patients with early stage oral squamous cell carcinoma. *Tumor Biol.* 37:1357–1368.

Hashimoto K, Tokimatsu T, Kawano S. Yoshizawa AC, Okuda S, Goto S, Kanehisa M. 2009. Comprehensive analysis of glycosyltransferases in eukaryotic genomes for structural and functional characterization of glycans. *Carbohydr Res*. 344:881–887.

Hassan H, Reis CA, Bennett EP, Mirgorodskaya E, Roepstorff P, Hollingsworth MA, Burchell J, Taylor-Papadimitriou J, Clausen H. 2000. The lectin domain of UDP-N-acetyl-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase-T4 directs its glycopeptide specificities. *J Biol Chem.* 275(49):38197-38205.

Hassinen A, Khoder-Agha F, Khosrowabadi E, Mennerich D, Harrus D, Noel M, Dimova EY, Glumoff T, Harduin-Lepers A, Kietzmann T, Kellokumpu S. 2019. A Golgi-associated redox switch regulates catalytic activation and cooperative functioning of ST6Gal-I with B4GalT-I. *Redox Biol*. 24:101182.

Hay ED. 1995. An overview of epithelio-mesenchymal transformation. *Acta Anat*. 154:8-20.

Hazes B. 1996. The (QxW)3 domain: a flexible lectin scaffold. *Protein Sci*. 5:1490-1501.

He H, Shen Z, Zhang H, Wang X, Tang Z, Xu J, Sun Y. 2014. Clinical significance of polypeptide N-acetylgalactosaminyl transferase-5 (GalNAc-T5) expression in patients with gastric cancer. *Br J Cancer*. 110(8):2021-2029.

He L, Qu L, Wei L, Chen Y, Suo J. 2017. Reduction of miR-132-3p contributes to gastric cancer proliferation by targeting MUC13. *Mol Med Rep.* 15:3055-3061.

Ho WL, Chou CH, Jeng YM, Lu MY, Yang YL, Jou ST, Lin DT, Chang HH, Lin KH, Hsu WM, Huang MC. 2014. GALNT2 suppresses malignant phenotypes through IGF-1 receptor and predicts favourable prognosis in neuroblastoma. *Oncotarget.* 5:12247-59.

Homa FL, Hollander T, Lehman DJ, Thomsen DR, Elhammer AP. 1993. Isolation and expression of a cDNA clone encoding a bovine UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase. *J Biol Chem.* 268:12609-12616.

Hu M, Guo W, Liao Y, Xu D, Sun B, Song H, Wang T, Kuang Y, Jing B, Li K, Ling J, Yao F, Deng J. 2019. Dysregulated ENPP1 increases the malignancy of human lung cancer by inducing epithelial-mesenchymal transition phenotypes and stem cell features. *Am J Cancer Res*. 9:134-144.

Hu WT, Yeh CC, Liu SY, Huang MC, Lai IR. 2018. The O-glycosylating enzyme GALNT2 suppresses the malignancy of gastric adenocarcinoma by reducing EGFR activities. *Am J Cancer Res.* 8:1739-1751.

Huang MJ, Hu RH, Chou CH, Hsu CL, Liu YW, Huang J, Hung JS, Lai IR, Juan HF, Yu SL, Wu YM, Huang MC. 2015. Knockdown of GALNT1 suppresses malignant phenotype of hepatocellular carcinoma by suppressing EGFR signalling. *Oncotarget*. 6:5650-65.

Huang T, Wu Q, Huang H, Zhang C, Wang L, Wang L, Liu Y, Li W, Zhang J, Liu Y. 2022. Expression of GALNT8 and O-glycosylation of BMP receptor 1A suppress breast cancer cell proliferation by upregulating ERα levels. *Biochim Biophys Acta Gen Subj.* 1866(1):130046.

Huanna T, Tao Z, Xiangfei W, Longfei A, Yuanyuan X, Jianhua W, Cuifang Z, Manjing J, Wenjing C, Shaochuan Q, Feifei X, Naikang L, Jinchao Z, Chen W. 2015. GALNT14 mediates tumour invasion and migration in breast cancer cell MCF-7. *Mol Carcinogenesis.* 54:1159-71.

Imberty A, Piller V, Piller F, Breton C. 1997. Fold recognition and molecular modelling of a lectin-like domain in UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases. *Protein Eng.* 10:1353-1356.

Ishikawa M, Kitayama J, Kohno K, Nagawa H. 2005. The expression pattern of UDP-N-acetyl-α-D-galactosamine-polypeptide N-acetyl-galactosaminyl transferase-3 in squamous cell carcinoma of the esophagus. *Pathobiology*. 72:139-145.

Iskratsch T, Braun A, Paschinger K, Wilson IB. 2009. Specificity analysis of lectins and antibodies using remodeled glycoproteins *Anal Biochem* 386(2):133–146.

Jiang J, Tang YL, Liang XH. 2011. EMT: a new vision of hypoxia promoting cancer progression. *Cancer Biol Ther*. 11:714-23.

Jiang Y, Liu Z, Xu F, Dong X, Yurong C, Hu Y, Gao T, Liu J, Yang L, Jia X, Qian H, Wen T, An G. 2018. Aberrant O-glycosylation contributes to tumorigenesis in human colorectal cancer. *J Cell Mol Med.* 22:4875–4885.

Jing Y, Han Z, Zhang S, Liu Y, Wei L. 2011. Epithelial-Mesenchymal Transition in tumor microenvironment. *Cell Biosci.* 1:29.

Ju T, Cummings RD. 2002. A unique molecular chaperone *Cosmc* required for activity of the mammalian core 1β3-galactosyltransferase. *Proc Natl Acad Sci USA.* 99:16613-16618.

Ju T, Aryal RP, Stowell CJ, Cummings RD. 2008. Regulation of protein O-glycosylation by the endoplasmic reticulum-localised molecular chaperone Cosmc. *J Cell Biol.* 182:531-542.

Ju T, Otto VI, Cummings RD. 2011. The Tn antigen-structural simplicity and biological complexity. *Angew Chem Int Ed.* 50:1770-1791.

Julien S, Videira PA, Delannoy P. 2012. Sialyl-Tn in cancer: (how) did we miss the target? *Biomolecules*. 2:435-466.

Kim DH, Xing T, Yang Z, Dudek R, Lu Q, Chen YH. 2017. Epithelial mesenchymal transition in embryonic development, tissue repair and cancer: a comprehensive overview. *J Clin Med.* 7:1.

Kingsley DM, Kozarsky KF, Hobbie L, Krieger M. 1986. Reversible defects in O-linked glycosylation and LDL receptor expression in a UDP-Gal/UDP-GalNAc 4 epimerase deficient mutant. *Cell.* 44:749-759.

Kitada S, Yamada S, Kuma A, Ouchi S, Tasaki T, Nebeshima A, Noguchi H, Wang K-Y, Shimajiri, Nakano R, Izumi H, Kohno K, Matsumoto T, Sasaguri Y. 2013. Polypeptide N-acetylgalactosaminyl transferase 3 independently predicts high-grade tumours and poor prognosis in patients with renal cell carcinomas. *Br J Cancer.* 109: 472–481.

Kohsaki T, Nishimori I, Nakayama H, Miyazaki E, Enzan H, Nomoto M, Hollingsworth MA, Onishi S. 2000. Expression of UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase isozymes T1 and T2 in human colorectal cancer. *J Gastroenterol.* 35:840-848.

Kölbl AC, Jeschke U, Friese K, Andergassen U. 2016. The role of TF- and Tn-antigens in breast cancer metastasis. *Histol Histopathol.* 31:613-621.

Kubota T, Shiba T, Sugioka S, Furukawa S, Sawaki H, Kato R, Wakatsuki S, Narimatsu H. 2006. Structural basis of carbohydrate transfer activity by human UDP-GalNAc:polypeptide alpha-N-acetylgalactosaminyltransferase (pp-GalNAc-T10). *J Mol Biol.* 359:708-727.

Kufe DW. (2009) Mucins in cancer: function, prognosis and therapy. *Nat Rev Cancer.* 9 874-885.

Kumari S, Khan S, Gupta SC, Kashyap VK, Yallapu MM, Chauhan SC, Jaggi M. 2018. MUC13 contributes to rewiring of glucose metabolism in pancreatic cancer *Oncogenesis*. 7:19.

Lamouille S, Xu J, Derynck R. 2014. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol.* 15:178-96.

Landers KA, Burger MJ, Tebay MA, Purdie DM, Scells B, Samaratunga H, Lavin MF, Gardiner RA. 2005. Use of multiple biomarkers for a molecular diagnosis of prostate cancer. *Int J Cancer.* 114:950-956.

Lavrsen K, Dabelsteen S, Vakhrushev SY, Levann AMR, Haue AD, Dylander A, Mandel U, Hansen L, Frödin M, Bennett EP, Wandall HH. 2018. De novo expression of human polypeptide N-acetylgalactosaminyltransferase 6 (GalNAc-T6) in colon adenocarcinoma inhibits the differentiation of colonic epithelium. *J Biol Chem.* 293(4):1298-1314.

Lenselink EA. 2015. Role of fibronectin in normal wound healing. *Int Wound J.* 12(3):313-6.

Li B, Shen W, Peng H, Li Y, Chen F, Zheng L, Xu J, Jia L. 2019. Fibronectin 1 promotes melanoma proliferation and metastasis by inhibiting apoptosis and regulating EMT. *Onco Targets Ther.* 12:3207-3221.

Li HW, Liu MB, Jiang X, Song T, Feng SX, Wu JY, Deng PF, Wang XY. 2021. GALNT14 regulates ferroptosis and apoptosis of ovarian cancer through the EGFR/mTOR pathway. *Future Oncol*. 18 (2):149-161.

Li X, Wang J, Li W, Xu Y, Shao D, Xie Y, Xie W, Kubota T, Narimatsu H, Zhang Y. 2012. Characterisation of ppGalNAc-T18, a member of the vertebrate-specific Y subfamily of UDP-N-acetyl-α-D-galactosamine:polypeptide N-acetylgalactosaminyltransferases. *Glycobiology.* 22:602-615.

Li Z, Yamada S, Inenaga S, Imamura T, Wu Y, Wang K-Y, Shimajiri S, Nakano R, Izumi H, Kohno K, Sasaguri Y. 2011. Polypeptide *N*-acetylgalactosaminyltransferase 6 expression in pancreatic cancer is an independent prognostic factor indicating better overall survival. *Br J Cancer.* 104:1882-18.

Libisch MG, Casás M, Chiribao M, Moreno P, Cayota A, Osinaga E, Oppezzo P, Robello C. 2014. GALNT11 as a new molecular marker in chronic lymphocytic leukemia. *Gene.* 533(1):270-279.

Lin MC, Huang MJ, Liu CH, Yang TL, Huang MC. 2014. GALNT2 enhances migration and invasion of oral squamous cell carcinoma by regulating EGFR glycosylation and activity. *Oral Oncol.* 50:478-84.

Lin TC, Chen ST, Huang MC, Huang J, Hsu CL, Juan HF, Lin HH, Chen CH. 2017. GALNT6 expression enhances aggressive phenotypes of ovarian cancer cells by regulating EGFR activity. *Oncotarget.* 8:42588-42601.

Liu C, Li Z, Xu L, Shi Y, Zhang X, Shi S, Hou K, Fan Y, Li C, Wang X, Zhou L, Liu Y, Qu X, Che X. 2020. GALNT6 promotes breast cancer metastasis by increasing mucin-type O-glycosylation of α2M. *Aging* 12:11794-11811.

Liu SY, Shun CT, Hung KY, Juan HF, Hsu CL, Huang MC, Lai IR. 2016. Mucin glycosylating enzyme GALNT2 suppresses malignancy in gastric adenocarcinoma by reducing MET phosphorylation. *Oncotarget.* 7:11251-11262.

Lo HW, Hsu SC, Xia W, Cao X, Shin JY, Wei Y, Abbruzzese JL, Hortobagyi GN, Hung MC. 2007. Epidermal growth factor receptor cooperates with signal transducer and activator of transcription 3 to induce epithelial-mesenchymal transition in cancer cells via up-regulation of TWIST gene expression. *Cancer Res.* 67:9066-9076.

López-Novoa JM, Nieto MA. 2009. Inflammation and EMT: an alliance towards organ fibrosis and cancer progression. *EMBO Mol Med.* 1:303-14.

Lu W, Kang Y. 2019. Epithelial-mesenchymal plasticity in cancer progression and metastasis. *Dev Cell*. 49:361-374.

Martin FT, Dwyer RM, Kelly J, Khan S, Murphy JM, Curran C, Miller N, Hennessy E, Dockery P, Barry FP, O'Brien T, Kerin MJ. 2010. Potential role of mesenchymal stem cells (MSCs) in the breast tumour microenvironment: stimulation of epithelial to mesenchymal transition (EMT). *Breast Cancer Res Treat*. 124:317-26.

Marucci A, Cozzolino F, Dimatteo C, Monti M, Pucci P, Trischitta V, Di Paola R. 2013. Role of GALNT2 in the modulation of ENPP1 expression, and insulin signalling and action: GALNT2: a novel modulator of insulin signalling. *Biochim Biophys Acta.* 1833:1388-1395.

Maupin KA, Sinha A, Eugster E, Miller J, Ross J, Paulino V, Keshamouni VG, Tran N, Berens M, Webb C, Haab BB. 2010. Glycogene expression alterations associated with pancreatic cancer epithelial-mesenchymal transition in complementary model systems. *PLoS One*. 5:e13002.

Mi R, Song L, Wang Y, Ding X, Zeng J, lehoux S, Aryal RP, Wang J, Crew VK, van Die I, Chapman AB, Cummings RD, Ju T. 2012 Epigenetic silencing of the chaperone Cosmc in human leukocytes expressing Tn antigen. *J Biol Chem*. 287(49):41523-41533.

Micalizzi DS, Farabaugh SM, Ford HL. 2010. Epithelial-mesenchymal transition in cancer: parallels between normal development and tumor progression. *J Mammary Gland Biol Neoplasia.* 15:117–134.

Miyahara N, Shoda J, Kawamoto T, Furukawa M, Ueda T, Todoroki T, Tanaka N, Matsuo K, Yamada Y, Kohno K, Irimura T. 2004. Expression of UDP-N¬-acetyl-α-D-galactosaminyltranferase isozyme 3 in the subserosal layer correlates with postsurgical survival of pathological tumor stage 2 carcinoma of the gallbladder. *Clin Cancer Res.* 10:2090-2099.

Nakajima S, Doi R, Toyoda E, Tsuji S, Wada M, Koizumi M, Tulachan SS, Ito D, Kami K, Mori T, Kawaguchi Y, Fujimoto K, Hosotani R, Imamura M. 2004. N-cadherin expression and epithelial-mesenchymal transition in pancreatic carcinoma. *Clin Cancer Res.* 10:4125-4133.

Nakamura N, Toba S, Hirai M, Morishita S, Mikami T, Konishi M, Itoh N, Kurasaka A. 2005. Cloning and expression of a brain-specific putative UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase gene. *Biol Pharm Bull.* 28:429-433.

Nguyen AT, Chia J, Ros M, Hui KM, Saltel F, Bard F. 2017. Organelle specific O-glycosylation drives MMP14 activation, tumor growth, and metastasis. *Cancer Cell*. 32:639-653.

Onitsuka K, Shibao K, Nakayama Y, Minagawa N, Hirata K, Izumi H, Matsuo K, Nagata N, Kitazato K, Kohno K, Itoh H. 2003. Prognostic significance of UDP-N-acetyl-α-D-galactosamine:polypeptide N-

acetylgalactosaminyltransferase-3 (GalNAc-T3) expression in patients with gastric carcinoma. *Cancer Sci.* 94:32-36.

Park J, Schwarzbauer JE. 2013. Mammary epithelial cell interactions with fibronectin stimulate epithelial-mesenchymal transition. *Oncogene.* 33:1649-1657.

Park JH, Nishidate T, Kijima K, Ohashi T, Takegawa K, Fujikane T, Hirata K, Nakamura Y, Katagiri T. 2010. Critical roles of mucin 1 glycosylation by transactivated polypeptide N-acetylgalactosaminyltransferase 6 in mammary carcinogenesis. *Cancer Res.* 70:2759-69.

Park JH, Katagiri T, Chung S, Kijima K, Nakamura Y. 2011. Polypeptide N-acetylgalactosaminyltransferase 6 disrupts mammary acinar morphogenesis through O-glycosylation of fibronectin. *Neoplasia.* 13:320-6.

Patani N, Jiang W, Mokbel K. 2008. Prognostic utility of glycosyltransferase expression in breast cancer. *Cancer Genomics Proteomics.* 5:333-340.

Paulson JC, Colley KJ. 1989. Glycosyltransferases - structure, localisation, and control of cell type-specific glycosylation. *J Biol Chem.* 264:17615-17618.

Peng C, Togayachi A, Kwon YD, Xie C, Wu G, Zou X, Sato T, Ito H, Tachibana K, Kubota T, Noce T, Narimatsu H, Zhang Y. 2010. Identification of a novel human UDP-GalNAc transferase with unique catalytic activity and expression profile. *Biochem Biophys Res Comm.* 402:680-686.

Peng RQ, Wan HY, Li HF, Liu M, Li X, Tang H. 2012. MicroRNA-214 suppresses growth and invasiveness of cervical cancer cells by targeting UDP-N-acetyl-α-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 7. *J Biol Chem.* 287(17):14301-14309.

Ponnusamy MP, Lakshmanan I, Jain M, Das S, Chakraborty S, Dey P, Batra SK. 2010. MUC4 mucin-induced epithelial to mesenchymal transition: a novel mechanism for metastasis of human ovarian cancer cells. *Oncogene.* 29:5741-5754.

Pratt MR, Hang HC, Ten Hagen KG, Rarick J, Gerken TA, Tabak LA, Bertozzi CR. 2004. Deconvoluting the functions of polypeptide N-alpha-acetylgalactosaminyltransferase family members by glycopeptide substrate profiling. *Chem Biol.* 11(7):1009-1016.

Pucci M, Malagolini N, Dall'Olio F. 2021. Glycobiology of the epithelial to mesenchymal transition. *Biomedicines*. 9:770.

Raghu D, Mobley RJ, Shendy NAM, Perry CH, Abell AN. 2019. GALNT3 maintains the epithelial state in trophoblast stem cells. *Cell Rep.* 26:3684-3697.

Ranganathan P, Weaver KL, Capobianco AJ. 2011. Notch signalling in solid tumours: a little bit of everything but not all the time. *Nat Rev Cancer.* 11(5):338-51.

Raman J, Guan Y, Perrine CL, Gerken TA, Tabak LA. 2012. UDP-N-acetyl-a-D-galactosamine:polypeptide N-acetylgalactosaminyltransferases: completion of the family tree. *Glycobiology.* 22:768-777.

Riaz SK, Ke Y, Wang F, Kayani MA, Malik MFA. 2019. Influence of SHH/GLI1 axis on EMT mediated migration and invasion of breast cancer cells. *Sci Rep.* 9(1):6620.

Rømer TB, Aasted MKM, Dabelsteen S, Groen A,Schnabel J, Tan E, Pedersen JW, Haue AD, Wandall HH. 2021. Mapping of truncated O-glycans in cancers of epithelial and non-epithelial origin*. Br J Cancer.* 125:1239–1250.

Röttger S, White J, Wandall HH, Olivo JC, Stark A, Bennett EP, Whitehouse C, Berger EG, Clausen H, Nilsson T. 1998. Localisation of three human polypeptide GalNAc-transferases in HeLa cell suggests initiation of O-linked glycosylation throughout the Golgi apparatus. *J Cell Sci.* 111:45-60.

Sánchez-Tilló E, Liu Y, de Barrios O, Siles L, Fanlo L, Cuatrecasas M, Darling DS, Dean DC, Castells A, Postigo A. 2012. EMT-activating transcription factors in cancer: beyond EMT and tumor invasiveness. *Cell Mol Life Sci.* 69:3429-3456.

Savagner P, Yamada KM, Thiery JP. 1997. The zinc-finger protein Slug causes desmosome dissociation, an initial and necessary step for growth factor-induced epithelial-mesenchymal transition. *J Cell Biol.* 137:1403-1419.

Savagner, P. 2001. Leaving the neighbourhood: molecular mechanisms involved during epithelial-mesenchymal transition. *Bioessays.* 23:912-923.

Schwientek T, Bennett EP, Flores C, Thacker J, Hollmann M, Reis CA, Behrens J, Mandel U, Keck B, Schäfer MA, Haselmann K, Zubarew B, Roepstorff P, Burchell JM, Taylor-Papadimitriou J, Hollingsworth MA, Clausen H. 2002. Functional conservation of subfamilies of putative UDP-N-acetylgalactosamine:polypeptide N-acetylgalactosaminyltransferases in Drosophila, Caenorhabditis elegans, and mammals. One subfamily composed of I(2)35Aa is essential in Drosophila. *J Biol Chem.* 277:22623-22638.

Shibao K, Izumi H, Nakayama Y, Ohta R, Nagata N, Nomoto M, Matsuo K, Yamada Y, Kitazato K, Itoh H, Kohno K. 2002. Expression of UDP-N-acetyl-α-D-galactosamine-polypeptide GalNAc N-acetylgalactosaminyl transferase-3 in relation to differentiation and prognosis in patients with colorectal carcinoma. *Cancer*. 94:1939-1946.

Shtutman M, Levina E, Ohouo P, Baig M, Roninson IB. 2006. Cell adhesion molecule L1 disrupts E-cadherin-containing adherens junctions and increases scattering and motility of MCF7 breast carcinoma cells. *Cancer Res.* 66:11370-11380.

Sindrewicz P, Lian LY, Yu LG. 2016. Interaction of the oncofetal Thomsen-Friedenreich antigen with galectins in cancer progression and metastasis. *Front Oncol.* 6:79.

Smith RD, Lupashin VV. 2008. Role of the conserved oligomeric Golgi (COG) complex in protein glycosylation. *Carbohydr Res*. 343:2024-2031.

Song J, Liu W, Wang J, Hao J, Wang Y, You X, Du X, Zhou Y, Ben J, Zhang X, Ye M, Wang Q. 2020. GALNT6 promotes invasion and metastasis of human lung adenocarcinoma cells through O-glycosylating chaperone protein GRP78. *Cell Death Dis*. 11:352.

Sousa AM, Grandgenett PM, David L, Almeida R, Hollingsworth MA, Santos-Silva F. 2016. Reflections on MUC1 glycoprotein: the hidden potential of isoforms in carcinogenesis. *APMIS*. 124:913-924.

Sporn MB. 1996. The war on cancer. *Lancet.* 347:1377-1381.

Springer GF. 1984. T and Tn, general carcinoma autoantigens. *Science*. 224:1198-1206.

Stanley P, Sundaram S, Sallustio S. 1991. A subclass of cell surface carbohydrates revealed by CHO mutant with two glycosylation mutations. *Glycobiology*. 1:307-314.

Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT-BG, Lavrsen K, Dabelsteen S, Pedersen NB, Marcos-Silva L, Gupta R, Bennett EP, Mandel U, Brunak S, WandallH, Levery SB,Clausen H. 2013. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J.* 32: 1478–1488.

Sun Z, Xue H, Wei Y, Wang C, Yu R, Wang C, Wang S, Xu J, Qian M, Meng Q, Li G. 2019. Mucin O-glycosylating enzyme GALNT2 facilitates the malignant character of glioma by activating the EGFR/PI3K/Akt/mTOR axis. *Clin Sci.* 133:1167-1184.

Syrkina MS, Vassetzky Y.S, Rubtsov MA. 2019. MUC1 story: great expectations, disappointments and the renaissance. *Curr Med Chem.* 26:554-563.

Tang X, Ruan H, Dong L, Li S, Wu Z, Guan M. 2021. lncRNA GAU1 induces GALNT8 overexpression and potentiates colorectal cancer progression. *Gastroenterol Res Pract.* 5960821.

Taniuchi K, Cerny RL, Tanouchi A, Kohno K, Kotani N, Honke K, Saibara T, Hollingsworth MA. 2011. Overexpression of GalNAc-transferase GalNAc-T3 promotes pancreatic cancer cell growth. *Oncogene.* 30(49):4843-4854.

Tarhan YE, Kato T, Jang M, Haga Y, Ueda K, Nakamura Y, Park JH. 2016. Morphological changes, cadherin switching, and growth suppression in pancreatic cancer by GALNT6 knockdown. *Neoplasia*. 18:265-272.

Taylor-Papadimitriou J, Burchell J, Miles DW, Dalziel M. 1999. MUC1 and cancer. *Biochim Biophys Acta.* 1455:301-13.

Ten Hagen KG, Hagen FK, Balys MM, Beres TM, Van Wuyckhuyse B, Tabak LA. 1998. Cloning and expression of a novel, tissue specifically expressed member of the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase family. *J Biol Chem.* 273(42):27749-27754.

Ten Hagen KG, Tetaert D, Hagen FK, Richet C, Beres TM, Gagnon J, Balys MM, VanWuyckhuyse B, Bedi GS, Degand P, Tabak LA. 1999. Characterization of a UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase that displays glycopeptide N-acetylgalactosaminyltransferase activity. *J Biol Chem.* 274(39):27867-27874.

Ten Hagen KG, Fritz TA, Tabak LA. 2003. All in the family: the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases. *Glycobiology.* 13(1):1R-16R.

Thomas D, Sagar S, Caffrey T, Grandgenett PM, Radhakrishnan P. 2019. Truncated O-glycans promote epithelial-to-mesenchymal transition and stemness properties of pancreatic cancer cells. *J Cell Mol Med.* 23:6885–6896.

Toba S, Tenno M, Konishi M, Mikami T, Itoh N, Kurosaka A. 2000. Brain-specific expression of a novel human UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase (GalNAc-T9). *Biochem Biophys Acta.* 1493:264-268.

Tollefsen SE, Kornfeld R 1983. Isolation and characterization of lectins from *Vicia villosa*. Two distinct carbohydrate binding activities are present in seed extracts. *J Biol Chem* 258 (8):5165–5171.

Tse JC, Kalluri R. 2007. Mechanisms of metastasis: epithelial-to-mesenchymal transition and contribution of tumor microenvironment. *J Cell Biochem*. 101:816-29.

Tsai JH, Yang J. 2013. Epithelial-mesenchymal plasticity in carcinoma metastasis. *Genes Dev.* 27:2192-2206.

Ubillos L, Berriel E, Mazal D, Victoria S, Barrios E, Osinaga E, Berois N. 2018. Polypeptide-GalNAc-T6 expression predicts better overall survival in patients with colon cancer. *Oncol Lett.* 16(1):225-234.

Vavasseur F, Dole K, Yang J, Matta KL, Myerscough N, Corfield A, Paraskeva C, Brockhausen I. 1994. O-glycan biosynthesis in human colorectal adenoma cells during progression to cancer. *Eur J Biochem.* 222:415-424.

Vesuna F, van Diest P, Chen JH, Raman V. 2008. Twist is a transcriptional repressor of E-cadherin gene expression in breast cancer. *Biochem Biophys Res Commun*. 367:235-241.

Wandall HH, Nielsen MAI, King-Smith S, de Haan N, Bagdonaite I. 2021. Global functions of O-glycosylation: promises and challenges in O-glycobiology. *The FEBS Journal*. 288: 7183–7212.

Wang H, Tachibana K, Zhang Y, Iwasaki H, Kameyama A, Cheng L, Guo JM, Hiruma T, Togayachi A, Kudo T, Kikuchi N, Narimatsu T. 2003. Cloning and characterisation of a novel UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase, pp-GalNAc-T14. *Biochem Biophys Res Commun.* 300:738-744.

Wang R, Yu C, Zhao D, Wu M, Yang Z. 2013. The mucin-type glycosylating enzyme polypeptide N-acetylgalactosaminyltransferase 14 promotes the migration of ovarian cancer by modifying mucin 13. *Oncol Rep.* 30:667-676.

Wang W, Sun R, Zeng L, Chen Y, Zhang N, Cao S, Deng S, Meng X, Yang S. 2021. GALNT2 promotes cell proliferation, migration, and invasion by activating the Notch/Hes1-PTEN-PI3K/Akt signalling pathway in lung adenocarcinoma. *Life Sci.* 276:119439.

Wang ZQ, Bachvarova M, Morin C, Plante M, Gregoire J, Renaud MC, Sebastianelli A, Bachvarov D. 2014. Role of the polypeptide N-acetylgalactosaminyltransferase 3 in ovarian cancer progression: possible implications in abnormal mucin O-glycosylation. *Oncotarget*. 5:544-560.

Weinberg RA. 2007. Moving out: invasion and metastasis. Chapter 14 in The biology of cancer. New York: Garland Science. pp587-654.

Welch DR, Hurst DR. 2019. Defining the hallmarks of metastasis. *Cancer Res*. 79:3011-3027.

White KE, Lorenz B, Evans WE, Meitinger T, Strom TM, Econs MJ. 2000. Molecular cloning of a novel human UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase, GalNAc-T8, and analysis as a candidate autosomal dominant hypophosphatemic rickets (ADHR) gene. *Gene.* 246:347-356.

White T, Bennett EP, Takio K, Sørensen T, Bonding N, Clausen H. 1995. Purification and cDNA cloning of a human UDP-N-acetyl-alpha-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase. *J Biol Chem.* 270:24156-24165.

Wu C, Guo X, Wang W, Wang Y, Shan Y, Zhang B, Song W, Ma S, Ge J, Deng H, Zhu M. 2010. N-acetylgalactosaminyltransferase-14 as a potential biomarker for breast cancer by immunohistochemistry. *BMC Cancer*. 10:123.

Wu D, Pan W. 2010. GSK3: a multifaceted kinase in Wnt signaling. *Trends Biochem Sci.* 35(3):161-8.

Wu Q, Xie X, Zhang K, Niang B, Liu Y, Zhang C, Huang T, Huang H, Li W, Zhang J, Liu Y. 2021a. Reduced expression of ppGalNAc-T4 promotes proliferation of human breast cancer cells. *Cell Biol Int*. 45(2):320-333.

Wu Q, Zhang C, Zhang K, Chen Q, Wu S, Huang H, Huang T, Zhang N, Wang X, Li W, Liu Y, Zhang J. 2021b. ppGalNAc-T4-catalyzed O-glycosylation of TGF-β type II receptor regulates breast cancer cells metastasis potential. *J Biol Chem.* 296:100119.

Wu YM, Liu CH, Hu RH, Huang M-J, Lee JJ, Chen CH, Huang J, Lai HS, Lee P-H, Hsu W-M, Huang HC, Huang MC. 2011. Mucin glycosylating enzyme GALNT2 regulates the malignant character of hepatocellular carcinoma by modifying the EGF receptor. *Cancer Res.* 71:7270–7279.

Xiong H, Hong J, Du W, Lin YW, Ren LL, Wang YC, Su WY, Wang JL, Cui Y, Wang ZH, Fang JY. 2012. Roles of STAT3 and ZEB1 proteins in E-cadherin down-regulation and human colorectal cancer epithelial-mesenchymal transition. *J Biol Chem.* 287:5819-5832.

Xu F, Wang D, Cui J, Li J, Jiang H. 2020a. Demethylation of the Cosmc promoter alleviates the progression of breast cancer through downregulation of the Tn and sialyl-Tn antigens. *Cancer Management and Research.* 12: 1017–1027.

Xu F, Shangguan X, Pan J, Yue Z, Shen K, Ji Y, Zhang W, Zhu Y, Sha J, Wang Y, Fan L, Dong B, Wang Q, Xu G, Wu YL, Li N, Xu R, Zhang JB, Ming H, Zhang Y. 2020b. GALNT10 promotes the proliferation and metastatic ability of gastric cancer and reduces 5-fluorouracil sensitivity by activating HOXD13. *Eur Rev Med Pharmacol Sci.* 24:11610-11619.

Xu Z, Jin H, Duan X, Liu H, Zhao X, Fan S, Wang Y, Yao T. 2021. LncRNA PSMA3-AS1 promotes cell proliferation, migration, and invasion in ovarian cancer by activating the PI3K/Akt pathway via the miR-378a-3p/GALNT3 axis. *Environ Toxicol.* 36 (12):2562-2577.

Xue W. 2021. HOXD13 suppresses prostate cancer metastasis and BMP4-induced epithelial-mesenchymal transition by inhibiting SMAD1. *Int J Cancer.* 148:3060-3070.

Yang J, Li G, Zhang K. 2016. MiR-125a regulates ovarian cancer proliferation and invasion by repressing GALNT14 expression. *Biomed Pharmacother.* 80:381-387.

Yu LG. 2007. The oncofetal Thomsen-Friedenreich carbohydrate antigen in cancer progression. *Glycoconj J.* 24:411-420.

Zhang Y, Iwasaki H, Wang H, Kudo T, Kalka TB, Hennet T, Kubota T, Cheng L, Inaba N, Gotoh M, Togayachi A, Guo J, Hisatomi H, Nakajima K, Nishihara S, Nakamura M, Marth JD, Narimatsu H. 2003. Cloning and characterisation of a new human UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase, designated pp-GalNAc-T13, that is specifically expressed in neurons and synthesises GalNAc alpha-serine/threonine antigen. *J Biol Chem.* 278:573-584.

Zhang YE. 2009. Non-Smad pathways in TGF-beta signaling. *Cell Res.* 19(1):128-39.

Zhou C, Liu J, Tang Y, Liang X. 2012. Inflammation linking EMT and cancer stem cells. *Oral Oncol.* 48:1068-75.

Zhi X, Tao J, Xie K, Zhu Y, Li Z, Tang J, Wang W, Xu H, Zhang J, Xu Z. 2014. Muc4-induced nuclear translocation of β-catenin: a novel mechanism for growth, metastasis and angiogenesis in pancreatic cancer. *Cancer Lett*. 346(1):104-113.

Zuo T, Shan J, Liu Y, Xie R, Yu X, Wu C. 2018. EFEMP2 mediates GALNT14-dependent breast cancer cell invasion. *Transl Oncol.* 11:346-352.

Legends to Figures (figures appended at end)

Figure 1. Phylogenetic tree depicting the evolutionary divergence of the GALNT family from the hypothesised ancestral GALNT gene. GALNTs with * are termed sub-family Y and have been proposed to act as chaperones for the other GALNT family members. Adapted from Bennett et al. (2012).

Figure 2. Schematic representation of the shared type II membrane structure of the GALNTs. The luminal portion consists of a catalytic domain, a ricin-like lectin domain and a variable length stem region. They are held in place by a hydrophobic transmembrane domain followed by a short N-terminal tail. Adapted from Beaman and Brooks (2014).

Figure 3. The sequence of events in metastasis. (1) Angiogenesis, the development of tumoural blood supply. (2) Dissemination of cancer cells from the primary tumour mass by invasion of local basement membrane and migration through stroma. (3) Intravasation of cancer cell into a blood (or lymphatic) vessel followed by (4) hematogenous (or lymphatic) dissemination, where, in the case of hematogenous dissemination, the cancer cells can form rafts with platelets and white blood cells. (5) Adhesion to the vascular endothelium. (6) Extravasation out of the blood (or lymphatic) vessel. (7) Invasion through stroma. (8) Formation a secondary tumour, where (1) angiogenesis is induced again. Cancer cell plasticity through EMT/MET is a feature of multiple stages during metastasis. In the primary tumour, cancer cells will be epithelial-like; to be able to detach from the primary tumour mass, the cell transdifferentiates to mesenchymal-like. The mesenchymal-like cancer cell can invade through basement membrane, surrounding stroma and then intravasate into a blood (or lymphatic) vessel. During haematogenous (or lymphatic) dissemination, the cancer cell transdifferentiates into epithelial-like and can adhere to the endothelium and initiate extravasation. The epithelial-like

cancer cell transdifferentiates into mesenchymal-like allowing it to invade the local stroma and migrate to a secondary site. At the secondary site, the mesenchymal-like cancer cell transdifferentiates into epithelial-like and can give rise to a secondary tumour.

Figure 4. During epithelial mesenchymal transition (EMT), epithelial cells down-regulate cell-cell adhesion structures. Loss of adherins junctions and tight junctions alter the cell polarity from apico-basal to anterior-posterior, changing the cell morphology from cuboidal to spindle-shaped, and allowing apical and basolateral membrane components to merge. The actin cytoskeleton is transformed into stress fibres which accumulate at cell protrusions. Epithelial intermediate filaments, the cytokeratins, are replaced by vimentin. The basement membrane is dissolved allowing motile mesenchymal cells to invade the surrounding stroma. Cells become resistant to anoikis. Epithelial-specific markers including E-cadherin are replaced by mesenchymal N-cadherin and $\alpha v\beta 6$ integrin. The process is reversible, with motile mesenchymal cells reverting to stationary, adhesive epithelial cells, termed mesenchymal epithelial transition (MET).

Figure 5. Transcriptional regulation of tumour associated epithelial mesenchymal transition (EMT): EMT is induced by autocrine/paracrine secretion of growth factors, cytokines and extracellular matrix proteins, which are secreted in response to hypoxia and inflammation. Binding of these EMT inducing signals to their specific receptors on the cell surface triggers a myriad of intracellular signalling networks which can act pleiotropically to induce EMT transcription factors. These regulate EMT effectors, which determine the epithelial, mesenchymal or metastable phenotype of the cell. Adapted from Lamouille et al. (2014). Abbreviations: BMP - bone morphogenic protein; BMPRs – bone morphogenic protein receptors; TGF-b - transforming growth factor beta; TGF-bRs transforming growth factor beta receptors; SHCA – SRC homology 2 domain-containing-transforming A GRB2 – growth factor receptor-bound protein 2; SOS – son of sevenless; EGF –

epidermal growth factor; HGF – hepatocytpe growth factor; IGF – insulin-like growth factor; HES1 – hairy and enhancer of split 1; PTEN – phosphatase and TENsin homolog; FZD – Frizzled; GSK3b - glycogen synthase kinase-3β ; Sonic hedgehog;SHH – PTCH – Patched; SMO – Smoothened; GLI – glioma-associated; ZEB – zinc-finger E-box-binding; MMPs – matrix metalloproteinases

Table I. Members of the human GALNT family, which add the first GalNAc monosaccharide in mucin type O-linked glycosylation, with alternative nomenclature and gene location

GALNT CAZy nomenclature (http ://afmb.cnrs-mrs.fr/fr/CAZY/)	**Has also been known as**	**Accession number**	**Chromosome locus**	**References**
GALNT1		X85018	18q12.1	White et al., 1995
GALNT2		X85019	1q41-q42	White et al., 1995
GALNT3		X92689	2q24-q31	Bennett et al., 1996
GALNT4	POC1B	Y08564	12q21.33	Bennett et al., 1998
GALNT5		AJ245539	2q24.1	Ten Hagen et al., 1998
GALNT6		Y08565	12q13	Bennett et al., 1999a;
GALNT7		AJ002744	4q34.1	Bennett et al., 1999b
GALNT8		AJ271385	12p13.3	White et al., 2000
GALNT9		AB040672	12q24.33	Toba et al., 2000
GALNT10		AJ505950	5q33.2	Cheng et al., 2002
GALNT11		Y12434	7q36.1	Schwientek et al., 2002
GALNT12		AJ32365	9q22.33	Guo et al., 2002
GALNT13		AJ505991	2q24.1	Zhang et al., 2003
GALNT14	FLJ12691	Y09324	2q23.1	Wang et al., 2003
GALNT15	GALNTL2, GALNT13, GALNT7, PIH5	NM_054110	3q25.1	Cheng et al., 2004

GALNT16	GLANTL1, KIAA1130	AJ505951	14q24.1	Peng et al., 2010
GALNT17	GALNTL6, GALNT20	AJ626725	4q34.1	Raman et al., 2012
GALNT18	GALNTL4, GALNT15, MGC71806	AJ626724	11p15.3	Raman et al., 2012
GALNT19	GALNTL3, GALNT16, GALNT20, WBSCR17	AJ626726	7q11.23	Nakamura et al., 2005
GALNT20	GALNTL5, GALNT15, WBSCR17	NM_145292	7p36.1	Raman et al., 2012

Table II. Reported tissue distribution of GALNT family members

GALNT	Summary description of tissue distribution	Detailed description of tissue distribution	References
GALNT1	Ubiquitous		Homa et al., 1993; White et al., 1995; Toba et al., 2000; Zhang et al., 2003; Peng et al., 2010; Bennett et al., 2012
GALNT2	Ubiquitous		Homa et al., 1993; White et al., 1995; Toba et al., 2000; Zhang et al., 2003; Peng et al., 2010; Bennett et al., 2012
GALNT3	Widely	Highly expressed in pancreas and testis; weakly in kidney, prostate, ovary, intestine and colon; low levels in placenta, lung and spleen; also found in germ cells	Bennett et al., 1996; Bennett et al., 1999a Bennett et al., 2012
GALNT4	Widely	Highly expressed in sublingual salivary gland, stomach, colon, small intestine, lung, cervix, and uterus; lower levels in kidney, liver, heart, brain, spleen and ovary	Hagen et al., 1997; Bennett et al., 1998; Ten Hagen et al., 1998
GALNT5	Widely	Sublingual salivary gland; weak expression in stomach, small intestine and colon	Ten Hagen et al., 1998
GALNT6	Widely	Placenta and trachea; weakly in brain and pancreas	Bennett et al., 1999a
GALNT7	Widely	Pancreas, uterus, retina, kidney, small intestine, omentum, stomach and central nervous system	Bennett et al., 1999b Bennett et al., 2012
GALNT8	Limited	Testis and colon	White et al., 2000; Raman et al., 2012
GALNT9	Limited	Brain and spinal cord	Toba et al., 2000
GALNT10	Widely	Highly expressed in small intestine; less so in stomach, pancreas, ovary, thyroid gland and spleen	Cheng et al., 2002

GALNT11	Widely	Highly expressed in kidney; weakly expressed in most other organs	Schwientek et al., 2002
GALNT12	Widely	Digestive organs; highly expressed in small intestine, stomach, pancreas, and colon. Expressed to a limited extent in the testis, thyroid gland and spleen	Guo et al., 2002
GALNT13	Limited	Neurons, adult cerebellum, cerebral cortex, and whole brain	Zhang et al., 2003
GALNT14	Wide		Bennett et al. 2012
GALNT15		Highly expressed in small intestine, placenta, spleen, cerebral cortex, and ovary, less so in uterus, mammary gland, stomach, cerebellum, and whole brain	Cheng et al., 2004
GALNT16	Wide	Highly expressed in heart, brain and testis, and in germ cells	Identified as T20 by Peng et al., 2010 but actually T16 according to Bennett et al., 2012; Raman et al., 2012
GALNT17	Limited	Brain and testis	Peng et al., 2010; Raman et al., 2012
GALNT18	Limited	Highly expressed in testis, thyroid and adult brain	Raman et al., 2012
GALNT19	Limited	Cerebellum, cerebral cortex and heart	Nakamura et al., 2005; Raman et al., 2012
GALNT20	Limited	Testis	Raman et al, 2012

Table III. Summary of major studies linking altered GALNT levels in cancers with prognosis, cancer cell behaviour, GalNAc O-glycosylation and EMT

GALNT	Cancer type	Changes reported in Cancer	Associated altered GalNAc O-glycosylation	Reference
GALNT1	Colorectal	Increased prevalence in cancer compared to normal determined by immunohistochemistry	Not assessed	Kohsaki et al., 2000
	Bladder	Increased expression in cancer compared to normal determined by DNA microarray; Knockdown inhibited cell proliferation and tumour growth in mice	Not assessed	Ding et al., 2012
	Liver	Increased expression associated with poorer survival. Knockdown in hepatocellular carcinoma cells reduces cell migration and invasion due to reduced O-glycosylation of EGFR	Knockdown decreases O-glycosylation of EGFR, determined by VVA labelling after benzyl-α-GalNAc treatment	Huang et al., 2015
GALNT2	Colorectal	Increased prevalence in cancer compared to normal determined by immunohistochemistry.	Not assessed	Kohsaki et al., 2000
	Hepatocellular	Downregulated in cancer compared to normal determined by immunohistochemistry and RT-PCR, associated with vascular invasion and recurrence. Over-expression in cancer cell lines supresses invasion and migration and supresses tumour growth in mice	GALNT2 O-glycosylates EGFR with Tn antigen, determined by VVA pulldown and labelling on western blot	Wu et al., 2011
	Oral squamous	Detectable at invasive edge of high grade clinical tumours, assessed by immunohistochemistry. Knockdown in oral squamous	GALNT2 O-glycosylates EGFR with Tn antigen, as determined by	Lin et al., 2014

		cell carcinoma cells reduced cell migration and invasion by reduced O-glycosylation of EGFR	VVA binding to western blots	
	Neuroblastoma	Immunopositivity associated with early stage and favourable prognosis, and over-expression resulted in smaller tumours in mice. Overexpression reduced cell migration and invasion by increased O-glycosylation of IGF-1R preventing its dimerization	Overexpression increased Tn antigen on IGF-1R, determined by VVA pull down	Ho et al., 2014
	Gastric	Reduced levels, determined by immunohistochemistry, associated with advanced disease and shorter recurrence free survival. Knockdown in gastric adenocarcinoma cells increased cell migration and invasion by reducing O-glycosylation and enhancing phosphorylation of HGFR, and promoted metastasis in a mouse model	Knockdown reduced VVA binding to HGFR in western blot	Liu et al., 2016
	Glioma	Immunopositivity associated with high grade and poor prognosis. Knockdown in glioma cells reduced cell migration and invasion, and less aggressive tumours in a mouse model, by reduced O-glycosylation of EGFR	GALNT2 O-glycosylates EGFR with Tn antigen, as determined by VVA binding on western blot	Sun et al., 2019
	Lung adenocarcinoma	High levels correlate with poor prognosis determined by immunohistochemistry and in patient datasets. Overexpression activates the Notch/Hes1-PTEN-PI3K/Akt signalling pathway	Not assessed	Wang et al., 2021
GALNT3	Colorectal	Strong presence in patient samples as determined by	Not assessed	Shibao et al., 2002

	immunohistochemistry significantly associated with enhanced 5 year survival		
Lung non small cell	Low levels, assessed by immunohistochemistry, associated with loss of differentiation, high rates of cell division and poor survival	Not assessed	Dosaka-Akita et al., 2002
Gastric	Weak presence in patient samples determined by immunohistochemistry is an independent prognostic marker of improved 5 year survival	Not assessed	Onitsuka et al., 2003
Gallbladder	Increased expression in cancer compared to normal as determined by qPCR. Presence of diffuse type localisation in subserosal layer determined by immunohistochemistry associated with poor prognosis	Not assessed	Miyahara et al., 2004
Lung adenocarcinoma	Low level determined by immunohistochemistry is a significant independent marker of poor prognosis	Not assessed	Gu et al., 2004
Prostate	Increased presence in malignant compared to benign hyperplasia determined by qPCR and immunohistochemistry.	Not assessed	Landers et al., 2005
Oesophageal squamous	Increased presence in cancer compared to dysplasia and normal determined by immunohistochemistry. Associated with lymph node metastasis.	Not assessed	Ishikawa et al., 2005
Pancreatic	High presence in pancreatic cancer patient samples determined by immunohistochemistry. Over expressed in pancreatic cancer cell lines compared to normal	Not assessed	Taniuchi et al., 2011

		determined by qPCR and WB analysis.		
	Renal	Immunopositivity associated with large tumour size, vascular invasion and poor survival; correlation with β-catenin immunopositivity but not with E-cadherin positivity.	Not assessed	Kitada et al., 2013
	Ovarian epithelial	High levels in cancer compared to normal tissues, determined by immunohistochemistry, correlated with shorter progression fee survival. Knockdown in epithelial ovarian cancer cells resulted in proteolysis of MUC1 through reduced O-glycosylation, and reduced proliferation, invasion and migration.	Not assessed	Wang et al., 2014
	Oral squamous	Immunopositivity associated with poor differentiation, lymphatic and vascular invasion and recurrence.	Not assessed	Harada et al., 2016
GALNT4	Breast	Associated with enhanced recurrence free survival in patient datasets. Knockdown in breast cancer cell lines increased cell proliferation through activation of Notch signalling pathway.	Glycosylation by Tn antigen, assessed by VVA lectin blotting, decreased in multiple proteins after knockdown	Wu et al., 2021a
	Breast	Upregulated in luminal type cancers, and associated with improved recurrence free survival in patient datasets. Knockdown in breast cancer cell lines decreased O-glycosylation of TGFβ type I and type II receptors which led to increased dimerization and activity.	Tn antigen presence on TGFβ type I and type II receptors reduced after knockdown, as assessed by VVA pulldown. Ser31 on the extracellular domain of TGFβ type II receptor identified as being the O-	Wu et al., 2021b

			glycosylation site of GALNT4 by HPLC-based *in vitro* O-glycosylation	
GALNT5	Gastric	Reduced levels, determined by immunohistochemistry, is an independent predictor of poor prognosis.	Not assessed	He et al., 2014
	Pancreatic	Reduced levels, determined by immunohistochemistry, in cancer compared to normal, and inversely correlated with MUC1 immunopositivity. Downregulated by MUC1 in cancer cells through binding to regulatory elements.	Not assessed	Caffrey et al., 2019
GALNT6	Breast	Immunopositivity rarely seen in normal but frequently in cancers, and significantly associated with stage 1 disease.	Not assessed	Berois et al., 2006a
	Breast	Increased expression in bone marrow samples from cancer patients and cell lines as determined using RT-PCR; significantly associated with subsequent recurrence in lymph node negative cases.	Not assessed	Freire et al., 2006
	Breast	Significantly higher presence in cancer samples compared to normal, determined by immunohistochemistry.	Not assessed	Patani et al., 2008
	Gastric	Detected in normal, metaplastic and cancer samples using immunohistochemistry. In cancer, associated with vascular invasion.	Not assessed	Gomes et al., 2009
	Breast	Immunohistochemistry reveals strong presence in cancer but not normal tissues. Knockdown in breast cancer cells reduces MUC1 stability	VVA pull down and western blot confirm GalNAc O-glycosylation of MUC1	Park et al., 2010

	due to decreased O-glycosylation.		
Pancreatic	Independent prognostic indicator for pancreatic cancer, determined by immunohistochemistry; associated with good or moderate degree of differentiation, absence of vascular invasion and early stage; immunopositivity predicts improved survival.	Not assessed	Li et al., 2011
Pancreatic	Highly expressed in clinical samples, assessed by analysis of publicly available datasets. Knockdown in pancreatic cancer cells causes decrease in P-cadherin and increase in E-cadherin and β-catenin and changes in morphology and cytoskeleton, also reduced invasiveness in transwell assay. Knockdown reduces expression of MUC4 and EGFR presumably through reduced O-glycosylation.	*In vitro* O-glycosylation assay confirms GALNT6 glycosylates MUC4 peptide, assessed by MALDI-TOF mass spectrometry, and identifies the 10^{th} Thr within a peptide fragment as being the predominant O-glycosylation target of GALNT6	Tarhan et al., 2016
Ovarian epithelial	Elevated levels of GALNT6 correlate with lymph node and distant metastasis and reduced survival, determined by immunohistochemistry. Knockdown in epithelial ovarian cancer cells reduces cell migration and invasion by decreased O-glycosylation of EGFR.	VVA pull down confirmed that GALNT6 glycosylated EGFR with Tn antigen	Lin et al., 2017
Colorectal	Presence, determined by immunohistochemistry, in cancer samples predicts significantly better survival; undetectable in normal mucosa.	Not assessed	Ubillos et al., 2018
Colon	Strong presence in patient colon adenocarcinoma	Not assessed	Lavrsen et al., 2018

		samples compared to normal tissue, determined by immunohistochemistry and transcriptomic analysis		
	Breast	Poor breast cancer patient survival associated with increased expression, determined by analysis of publicly available datasets and immunohistochemistry. Knockdown in breast cancer cell lines inhibits PI3K/Akt signalling, presumably through reduced O-glycosylation of signalling pathway activator α2M, and reduces invasion in a transwell assay.	VVA pull down and western blot confirm GalNAc O-glycosylation of breast cancer cells by GALNT6	Liu et al., 2020
	Lung adenocarcionoma	Immunopositivity associated with metastasis and poor prognosis in clinical samples. Over expression enhanced migration and invasion in a wound healing and matrigel assay and silencing inhibited metastasis in a mouse xenograft model. Knockdown increased expression of E-cadherin and reduced expression of N-cadherin and Slug. Overexpression of GRP78 caused relocation to ER thereby increasing O-glycosylation of ERK1/2.	VVA-lectin Western blot analysis showed enhanced O-glycosylation of GRP78	Song et al., 2020
GALNT7	Cervical	Upregulated in cervical cancer patient samples as determined by qPCR. Repression by miR-214 reduces cell proliferation, migration, and invasiveness.	Not assessed	Peng et al., 2012
	Cervical	Overexpression, assessed by RT-qPCR, identified in clinical samples and cell lines; promotes cell proliferation and invasion by activating the EGFR/PI3K/Akt pathway in cervical cancer cells through O-glycosylation of EGFR.	Not assessed	Cao et al., 2020

GALNT8	Colorectal	Up-regulated in cancer tissues in comparison with normal, determined by RNASeq and immunohistochemistry, and co-expressed with GAU1. Expression associated with poor prognosis. Over-expression in cell lines enhanced colony formation and proliferation.	Not assessed	Tang et al., 2021
	Breast	mRNA and immunohistochemistry reveal downregulation is associated with poor clinical outcome and lower ERα expression. Overexpression in cell lines reduces proliferation. Downregulation is associated with reduction in Erα through inhibition of the BMP/Smad/RUNX2 axis in breast cancer cell lines.	VVA and peanutlectin labelling confirm increased GALNT8 associated with O-GalNAcylation of total proteins and of BMPR1A	Huang et al., 2022
GALNT9	Neuroblastoma	Expression determined by RT-PCR significantly associated with high overall survival	Not assessed	Berois et al., 2013
GALNT10	Gastric	Greater immunopositivity in cancer than adjacent normal tissue; immunopositivity associated with higher grade.	Not assessed	Gao et al., 2013
	Gastric	High levels in clinical samples, as determined by qRT-PCR, associated with lymph node and distant metastases and poor prognosis. Knockdown in cancer cells reduces proliferation and migration in a transwell assay, and also HOXD13 levels.	Not assessed	Xu et al., 2020b
GALNT11	Chronic lymphocytic leukaemia	Overexpressed in B-cells of chronic lymphocytic leukaemia patients as determined by qPCR.	Flow cytometry using VVA and monoclonal antibodies against Tn and T antigen demonstrated a low density of Tn antigen	Libisch et al., 2014

GALNT12	Colorectal	Down regulated in colorectal cancer cell lines and tissues compared to normal as determined by RT-PCR, lower in cases with metastases.	Not assessed	Guo et al., 2004
	Colon	Inactivating somatic and germline mutations found in cancer samples.	Not assessed	Guda et al., 2009
	Colorectal	Two deleterious variants identified that are highly likely to predispose to hereditary colorectal cancer.	Not assessed	Clarke et al., 2012
GALNT13	Neuroblastoma	Detected by QRT-PCR in bone marrow samples from stage 4 patients but not controls and highly correlated with poor outcome	Not assessed	Berois et al., 2006b
GALNT14	Breast	Presence increased in cancer compared to normal, determined by immunohistochemistry, and is associated with low histological grade	Not assessed	Wu et al., 2010
	Ovarian	Upregulated in cancers, as determined by western blot and immunohistochemistry, and correlates with stage. Suppression in ovarian cancer cells causes a reduction in invasion and in MMP2 and -9 activity.	Not assessed	Yang et al., 2016
	Ovarian	Overexpressed in cisplatin resistant cancers, determined by analysis of publicly available datasets and western blot. Downregulation significantly decreases O-glycosylation of EGFR leading to inhibition of the EGFR/PI3K/Akt signalling pathway	EGFR glycosylated with Tn antigen, determined by VVA pulldown	Li et al., 2021
GALNT15	Colorectal	SNPs analysis of cases versus controls identified GALNT15 as	Not assessed	Abulí et al., 2011

	predicting increased cancer risk
GALNT16 to GALNT20	Not reported

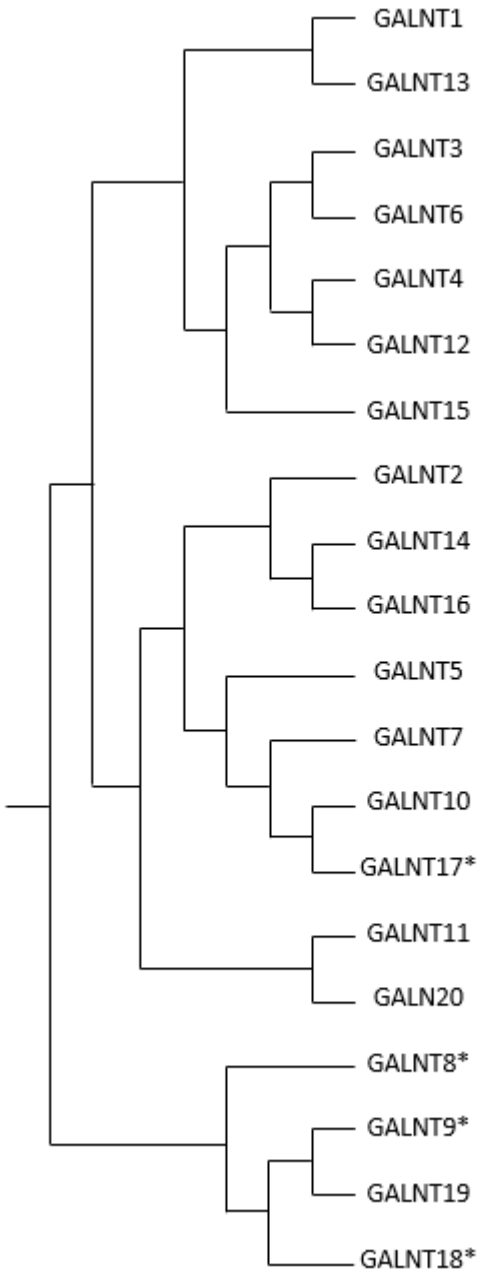


Figure 1. Phylogenetic tree depicting the evolutionary divergence of the GALNT family from the hypothesised ancestral GALNT gene. GALNTs with * are termed sub-family Y and have been proposed to act as chaperones for the other GALNT family members. Adapted from Bennett et al. (2012).

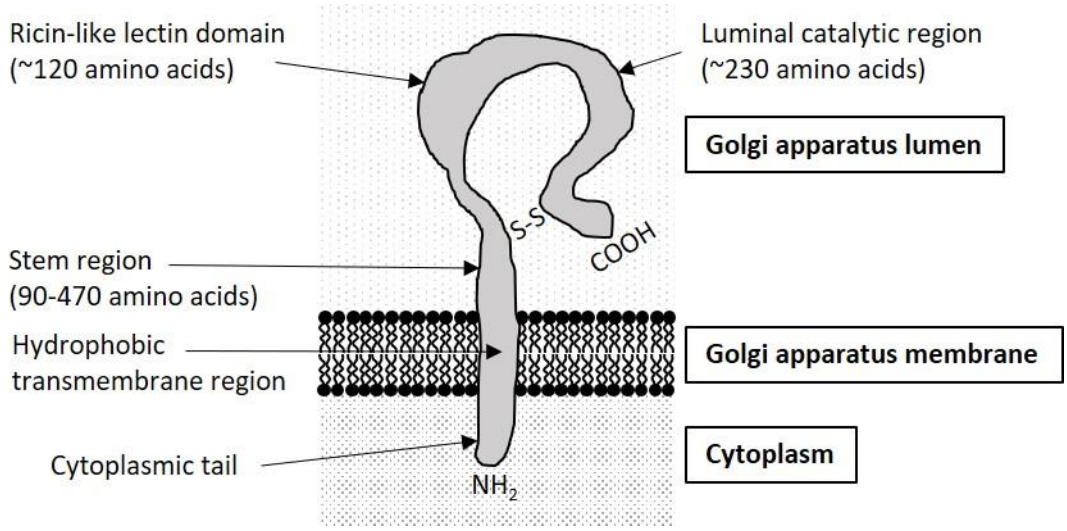


Figure 2. Schematic representation of the shared type II membrane structure of the GALNTs. The luminal portion consists of a catalytic domain, a ricin-like lectin domain and a variable length stem region. They are held in place by a hydrophobic transmembrane domain followed by a short N-terminal tail. Adapted from Beaman and Brooks (2014).

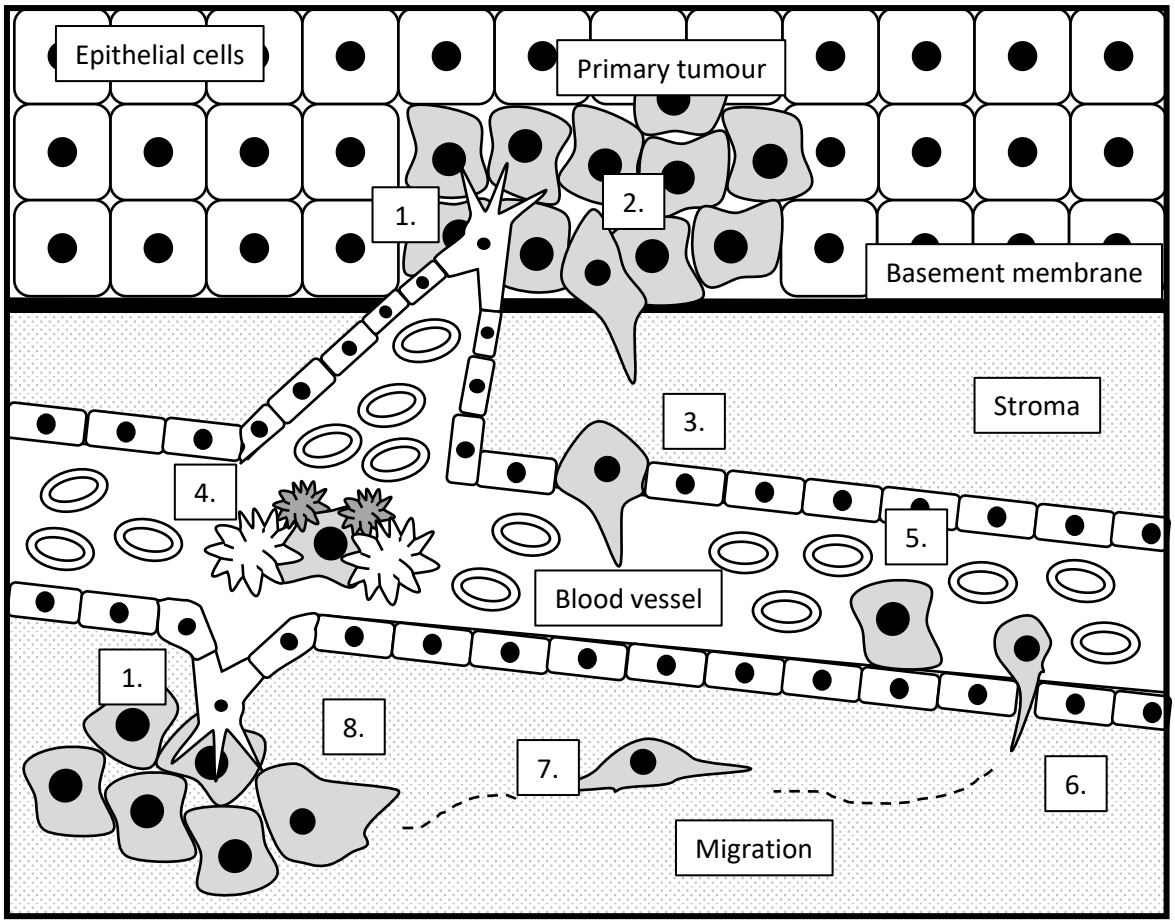
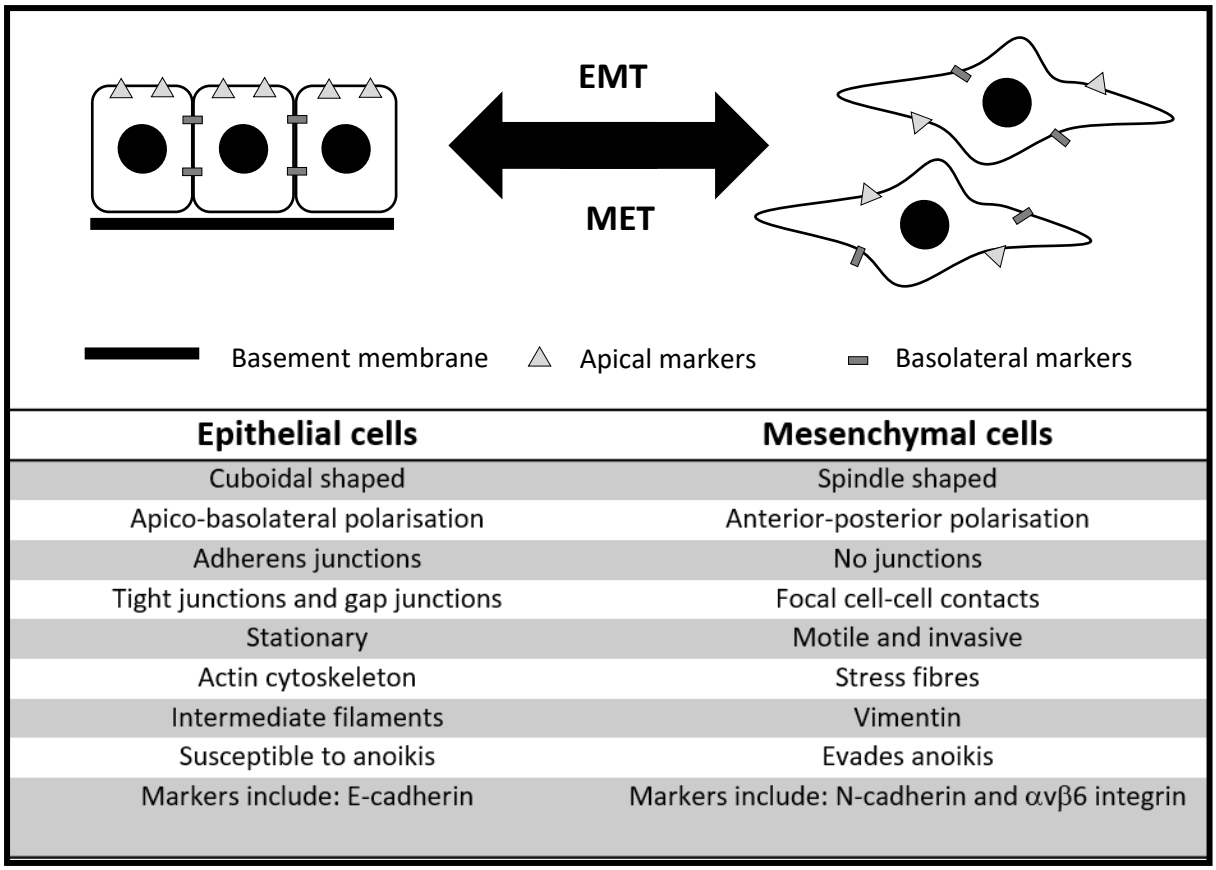


Figure 3. The sequence of events in metastasis. (1) Angiogenesis, the development of tumoural blood supply. (2) Dissemination of cancer cells from the primary tumour mass by invasion of local basement membrane and migration through stroma. (3) Intravasation of cancer cell into a blood (or lymphatic) vessel followed by (4) hematogenous (or lymphatic) dissemination, where, in the case of hematogenous dissemination, the cancer cells can form rafts with platelets and white blood cells. (5) Adhesion to the vascular endothelium. (6) Extravasation out of the blood (or lymphatic) vessel. (7) Invasion through stroma. (8) Formation a secondary tumour, where (1) angiogenesis is induced again. Cancer cell plasticity through EMT/MET is a feature of multiple stages during metastasis. In the primary tumour, cancer cells will be epithelial-like; to be able to detach from the primary tumour mass, the cell transdifferentiates to mesenchymal-like. The mesenchymal-like cancer cell can invade through basement membrane, surrounding stroma and then intravasate into a blood (or lymphatic) vessel. During haematogenous (or lymphatic) dissemination, the cancer cell transdifferentiates into epithelial-like and can adhere to the endothelium and initiate extravasation. The epithelial-like cancer cell transdifferentiates into mesenchymal-like allowing it to invade the local stroma and migrate to a secondary site. At the secondary site, the mesenchymal-like cancer cell transdifferentiates into epithelial-like and can give rise to a secondary tumour.



Epithelial cells	Mesenchymal cells
Cuboidal shaped	Spindle shaped
Apico-basolateral polarisation	Anterior-posterior polarisation
Adherens junctions	No junctions
Tight junctions and gap junctions	Focal cell-cell contacts
Stationary	Motile and invasive
Actin cytoskeleton	Stress fibres
Intermediate filaments	Vimentin
Susceptible to anoikis	Evades anoikis
Markers include: E-cadherin	Markers include: N-cadherin and αvβ6 integrin

Figure 4. During epithelial mesenchymal transition (EMT), epithelial cells down-regulate cell-cell adhesion structures. Loss of adherins junctions and tight junctions alter the cell polarity from apico-basal to anterior-posterior, changing the cell morphology from cuboidal to spindle-shaped, and allowing apical and basolateral membrane components to merge. The actin cytoskeleton is transformed into stress fibres which accumulate at cell protrusions. Epithelial intermediate filaments, the cytokeratins, are replaced by vimentin. The basement membrane is dissolved allowing motile mesenchymal cells to invade the surrounding stroma. Cells become resistant to anoikis. Epithelial-specific markers including E-cadherin are replaced by mesenchymal N-cadherin and αvβ6 integrin. The process is reversible, with motile mesenchymal cells reverting to stationary, adhesive epithelial cells, termed mesenchymal epithelial transition (MET).

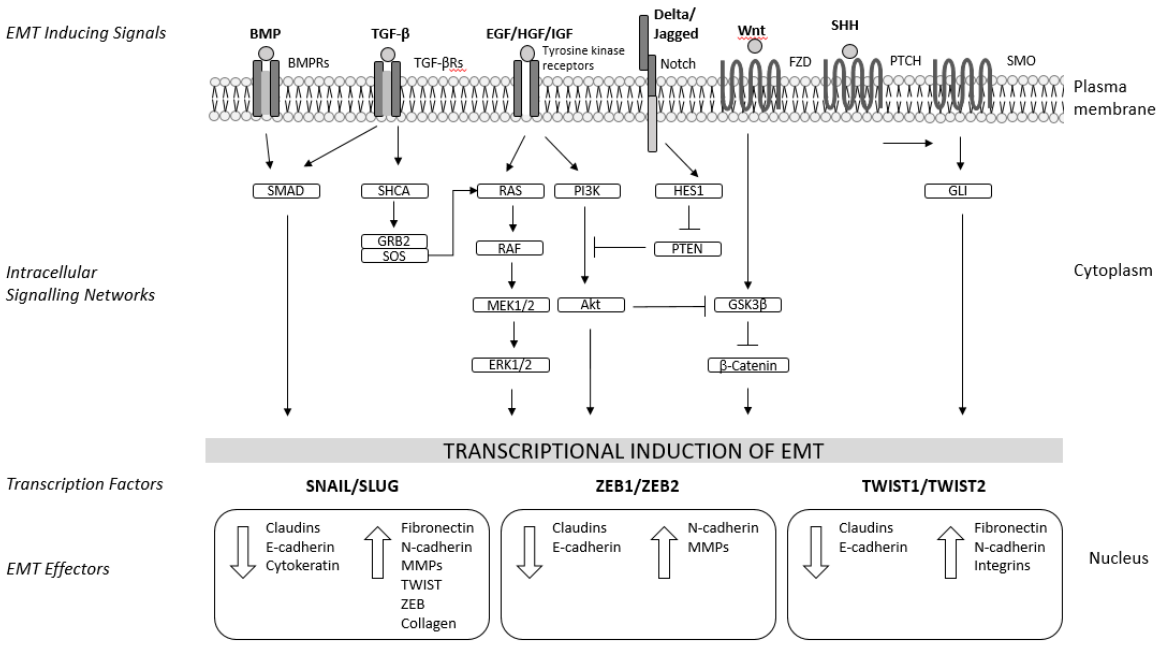


Figure 5. Transcriptional regulation of tumour associated epithelial mesenchymal transition (EMT): EMT is induced by autocrine/paracrine secretion of growth factors, cytokines and extracellular matrix proteins, which are secreted in response to hypoxia and inflammation. Binding of these EMT inducing signals to their specific receptors on the cell surface triggers a myriad of intracellular signalling networks which can act pleiotropically to induce EMT transcription factors. These regulate EMT effectors, which determine the epithelial, mesenchymal or metastable phenotype of the cell. Adapted from Lamouille et al. (2014). Abbreviations: BMP - bone morphogenic protein; BMPRs – bone morphogenic protein receptors; TGF-b - transforming growth factor beta; TGF-bRs transforming growth factor beta receptors; SHCA – SRC homology 2 domain-containing-transforming A GRB2 – growth factor receptor-bound protein 2; SOS – son of sevenless; EGF-epidermal growth factor; HGF – hepatocytpe growth factor; IGF – insulin-like growth factor; HES1 – hairy and enhancer of split 1; PTEN – phosphatase and TENsin homolog; FZD – Frizzled; GSK3b - glycogen synthase kinase-3β ; Sonic hedgehog;SHH – PTCH – Patched; SMO – Smoothened; GLI – glioma-associated; ZEB – zinc-finger E-box-binding; MMPs – matrix metalloproteinases.