

**Heparan sulfate proteoglycans, tumour progression and the cancer stem cell niche.**

**\*Larisa M. Haupt and Lyn R. Griffiths**

Genomics Research Centre, Griffith University Institute for Health and Medical Research,  
Parklands Drive, Southport, Queensland 4215, Australia.

Address for correspondence:

Professor Lyn Griffiths

Genomics Research Centre

Griffith University Gold Coast Campus

Parklands Drive

Southport 4215

Queensland, Australia.

[l.griffiths@griffith.edu.au](mailto:l.griffiths@griffith.edu.au)

Ph + 61 7555 29158

Fax + 61 7555 28908

Keywords:

**Abstract:**

The cancer stem cell hypothesis states that tumours arise from cells with the ability to self-renew and differentiate into multiple cell types, and that these cells persist in tumors as a distinct population that can cause disease relapse and metastasis. The crux of this hypothesis is that these cells are the only cells capable of, by themselves, giving rise to new tumours. What proportion of a tumour consists of these stem cells, where are they localised, and how can we identify them?

### **Cancer and the cancer stem cell hypothesis:**

The majority of cancers are solid tumours, with eighty percent of these epithelial in origin and hence arising in tissues that include the breast, lung, colon, prostate and ovary [Visvader 08]. The specific cellular origin of most tumours remains largely unknown, with speculation that tumour heterogeneity may reflect not only the specific cell of origin but also the microenvironment in which the tumour arises, and it is consistent with varied responses to therapies observed across patient populations [Cancemi 08]. This tumour microenvironment or niche which includes the stroma, inflammatory cells and vasculature has a significant role in all stages of tumorigenesis. Tumours are generally characterised clinically at the gross level by histology and by their expression profile of known specific markers. This combination of detection has enabled the definition of distinct tumour subtypes, but has not yet produced distinctive markers of initiation or indeed progression. As no tumour-specific target antigens are available, monoclonal antibodies against various antigens including cytoskeletal markers, surface adhesion molecules, or growth factor receptors are used to determine tumorigenicity. Currently, there are two models for tumour heterogeneity: the cancer stem cell hypothesis and the clonal evolution model [Visvader 08, Riethdorf 08, Hwang-Versluis 08].

Cancer stem cells (CSCs) refer to a specific subset of tumour cells that have the ability to self-renew and generate the diversity of cells which make up the tumour. This model implies a hierarchical structure where these CSCs, with stem-like properties, are able to initiate and sustain tumorigenesis [Reya 01, Visvader 08, Dontu 08]. The differentiated or non-CSC cells that make up the bulk of the tumour mass have a high proliferative potential, but this is not unlimited [Dontu 08]. The first evidence for the existence of CSCs came from leukemia in which a rare subset comprising 0.1-1% of the acute myeloid tumour cell population inducing leukemia following transplantation into immunodeficient mice [Bonnet 97]. In contrast, the clonal evolution model postulates that mutant tumour cells with a growth advantage are selected and expanded, with cells in the dominant population having potential for regenerating tumour growth. The acquisition of genetic events, epigenetic and other changes, along with microenvironmental changes are plausible contributors in both models. Indeed both models are not mutually exclusive, with both likely to exist in human cancer as demonstrated by leukemia cells and models of serial transplantation (resulting in more aggressive tumours). Indeed, the characteristics that define CSCs, including their long life and high proliferative capacity may increase the susceptibility of these cells to transformation [Visvader 08, Morel 08, Dontu 08].

The lack of specific markers used to define most normal stem cells and their progeny highlights the difficulty in isolating cells of tumour initiation, which is likely to include CSCs and their progeny [Visvader 08]. The identification of these cells within the tumour combined with extensive *in vivo* analysis to accurately determine the frequency of CSCs within solid tumours would provide targets for directed therapeutics. It follows that eliminating these differentiated progenies while sparing the cancer stem cells will ultimately result in relapse [Dontu 08]. To date, several cell surface markers have proved useful for the isolation of tumours enriched for CSCs, including CD133 (PROM1), CD44, CD24 and epithelial cell adhesion molecule (EpCAM, also known as epithelial specific antigen (ESA) and TACSTD1). Although useful in some contexts, more definitive markers are required as many antigens, including CD44, CD24 and CD133 heterogeneous between and within tumours and widely expressed in many different types of CSCs as well as normal stem cells [Honeth 08, Bloushtain-Qimron 08, Visvader 08]. It is also important to keep the use of the current suite of cell surface markers in context. Although markers of tumourigenic cells in established cell lines, they may differ from CSCs in primary tumours. Thus, observations made in cancer cell lines must be extended to primary tumours and in contextual models in order to validate their significance [Ivascu 06, Visvader 08]. As an example, recent work demonstrates the previously described CD44<sup>+</sup>CD24<sup>-</sup> stem cell phenotype is associated with basal-type breast cancers in human patients; in particular BRCA1 inherited cancers, but this has been shown not to correlate with clinical outcome. The use of tumour cells cultured in spheroids and other three-dimensional systems, allows the multicellular *in vivo* format to be mimicked *in vitro*, and have already demonstrated morphological, functional and gene expression differences with their traditional two-dimensional cell cultures [Wright 08, Ivascu 06, Sutherland 86, Angello 82, Streuli 91, Ghosh 05, Kenny 07]. These findings caution that the success of our efforts in translating cancer stem cell research into clinical practice depends on the thorough characterisation of CSCs and their surrounding microenvironment [Dontu 08].

For epithelial malignancies, the epithelial-mesenchymal transition (EMT) is considered to be a crucial event in the metastatic process, which involves disruption of epithelial cell homeostasis and the acquisition of a migratory mesenchymal phenotype. In many epithelial tumours, an EMT appears to be controlled by canonical pathways such as the Wnt and TGF $\beta$  pathways, both of which can be aberrantly activated during neoplasia [Birchmeier 94, Thiery 03, Blick 08, Visvader 08]. It is conceivable that CSCs may have also play a role in the formation of a particular niche for metastasis, for example primary tumour cells can generate a pre-metastatic niche by recruiting haematopoietic progenitor cells to tumour-specific niches. Evidence for this comes from work demonstrating that the site of transplantation influences tumour initiating frequency in xenograft and *in vivo* models [Forsberg 05, Spradling 01,

Visvader 08]. Critical to improving cancer models currently in use is the improved understanding of the cellular niche including the elucidation of markers specific for tumour type.

### **Heparan Sulfate Proteoglycans (HSPG)s:**

Proteoglycan (PGs) are a large family of proteins characteristically composed of a core protein to which one or more glycosaminoglycans (GAGs) attach. They include the heparan sulfate (HS) proteoglycans (HSPGs) and chondroitin sulfate (CS) proteoglycans (CSPGs). CS side chains are predominantly found on matrix localised PGs, whereas membrane PGs such as glypicans and syndecans contain mostly HS side chains. Precursor HS chains are synthesised in the golgi as non-sulfated copolymers attached to a HSPG core protein. Following chain initiation, several modification steps are required to complete the specific HS sulfation pattern. Polymerisation is the first step of HS biosynthesis, the beginning of chain assembly, and the commitment of the PG toward a HSPG and not a CSPG. The alternating addition of glucuronic acid (GlcA) and N-acetylglucosamine (GlcNAc) residues is catalysed by members of the *Ext* family (Ext1 and Ext2). Following completion of chain assembly, further sequential modifications occur within the golgi. The NDST (N-deacetylase/N-sulfotransferase) family of enzymes removes acetyl groups from the GlcNAc residues, generating free amino groups sulfated via N-sulfotransferase activity. All subsequent modifications of the HS chain depend on the presence of these residues, making NDSTs responsible for the overall design of the polysaccharide [Sugahara 02, Grobe 02].

As substrate recognition by most enzymes during HS chain modification is generally dependant on structural modifications introduced in previous reactions [Lindahl 98], there appears to be coordinated temporal interaction. The final HS chain contains clusters of N- and O-sulfated sugar residues separated by non-sulfated regions of varying length, with ligand binding properties of the HS chain determined by the density of sulfated domains [Grobe 02, Lindahl 98, Baldwin 08]. HSPGs are ubiquitous to the cell surface and in the extracellular matrix (ECM), including basement membranes, where they mediate cell-cell and cell-ECM interactions [Lamoureux 07]. The exact mechanism of these interactions is still not fully elucidated, but is thought to include regulation of the activity and stability of the ligands [Hacker 05, Habuchi 04]. As HS sequences are not directly encoded by genes, but are created by an elaborate post-translational biosynthesis [Sugahara 02, Gallagher 92], the tissue specificity of these interactions are due, at least in part, to the fine structure of HS characterised by the sulfation pattern [Baldwin 08, Habuchi 04].

The sulfation of HS has been associated with roles in both differentiation and [Habuchi 04, Ling 06] transformation of cells [Jayson 98, Safaiyan 98]. Examples include mutated or inactivated Ext1 resulting in inappropriate bone growth [Hecht 97]; sulfateless, the *Drosophila* NDST analog and its effect on multiple signalling pathways mediated by wingless, hedgehog, and the FGF/FGFR, breathless and branchless [The 99, Lin 99]; and developmental studies in embryonic brain demonstrating a HS-mediated growth factor activity switch from FGF2 to FGF1, accompanied by alterations in 6-O-sulfation patterns, total chain length, and the number of sulfated domains of the HS species [Lindahl 98, Brickman 98, Feyzi 98]. These data demonstrate the specificity of the HS produced by the cell may be mediated by distinct stimuli and hence can be up or down regulated dependent on the cell type and its niche [Kjellen 83].

### **HSPGs, the niche and cancer:**

The extracellular matrix (ECM) is a reservoir of cellular and non-cellular material including binding proteins and growth factors that mediate cellular processes [Bi 06, DeClerck 04]. The stromal cell components of the ECM are integral participants in these signaling mechanisms. Stromal cells are a diverse population consisting of fibroblasts, smooth muscle cells, endothelial cells and others. These cells provide a scaffold to the developing stem and progenitor cells, but also produce transmembrane ligands, extracellular matrix components and soluble proteins [Schofield 99, Cattaruzza 08, Habuchi 04, Ling 06]. Within the ECM, levels of gene expression vary, but the presence of specific combinations of growth factors and their receptors, may prime certain cells for induction when and if they are exposed to specific combinations in this microenvironment or niche. Due in part to their high negative charge, HS chains of PGs interact with a large number of ligands including growth factors and morphogens (FGF, Wnts, VEGFs), their receptors (FGFRs), as well as enzymes and enzyme inhibitors, lipases, apolipoproteins, ECM structural molecules (collagen, fibronectin, laminin) and plasma proteins [Uchimura 06, Kreuger 01, Ashikari-Hada 04, Feyzi 97].

It has been well established that the ECM and its constituents modulate cellular phenotypes, cell-cell interactions and signalling cascades. Recent evidence suggests that PGs may sequester growth factors through their core proteins to simultaneously modulate growth factor receptor activity and signalling efficacy. Dependant on type, expression pattern, and accessibility to ligands, PGs can either promote or inhibit cell proliferation and hence tumourigenesis. This tumour growth-promoting or tumour suppressing-activity may be dependent on any aspect or indeed the combination of the specific HSPG (the core protein, its ligand/growth factor specificity or the tumour cell characteristics). PGs localised in the ECM may be upregulated in the tumour cells themselves and participate in the cells response to

growth factor stimuli. Indeed GAG dysregulation in cancer has been correlated with prognosis in several malignancies [Cattaruzza 08, Frankel 08]. It is unclear whether these specific interactions occur at the level of the ECM or at the plasma membrane [Cattaruzza 08].

Members of the glypican and syndecan families of HSPGs have been implicated in the promotion of local tumour cell growth in some cancer tissues, but to inhibit tissue invasion and metastasis in others [Sorensen 08, Cattaruzza 08]. The mechanism/s by which this duality occurs is still unclear. Previous reports have demonstrated expression of glypican-1 is induced in human pancreatic tumour cells and surrounding fibroblasts, and in breast cancer cells, with the mitogenic response of the pancreatic tumour cells to FGF2 and EGF abrogated by specific downregulation of this HSPG [Ding 05, Cattaruzza 08]. Syndecan-4 overexpression has been shown to enhance FGF2 signalling, and is thought to regulate erb-B2 and erb-B3 levels in colon carcinoma. In terms HSPG interactions, syndecan-1 shedding is involved in the initial FGF2 response for glypican-dependent cells and is thought to modulate the long-term FGF2 response of glypican-independent cells [Cattaruzza 08].

However, the expression profile of the core proteins (syndecans, glypicans etc) along with the side-chain modification enzymes (NDSTs, Exts etc) has yet to be fully determined in most human cancers. The growth factor signalling associated with tumorigenesis has been postulated to be controlled, at least on one level, by an autoregulatory loop that may involve HSPGs expression levels on the cell surface and their role in the endocytosis of receptor/ligand complexes [Cattaruzza 08]. The aberrant distribution of PGs on tumour lesions could be correlated with the possibility that the stromal compartment of the lesion may also be enriched in these macromolecules. Patterns of HSPG expression are believed to mirror those of ligands that require HSPGs to elicit cellular responses. As a result, the presentation and/or release of growth factors may aid in establishing the co-ordination of growth factors and the required HSPG to facilitate and direct tumour cell motility [Cattaruzza 08]. Increased knowledge of the tumour niche and its role in the initiation, proliferation and progression of tumours is one of the major objectives in cancer research [Pucci-Minafra 08].

### **HSPGs and Metastasis:**

Cancers are diagnosed and classified histologically as invasive or non-invasive. Invasive carcinomas extend into the surrounding ECM, whilst in non-invasive carcinoma tumour cells are confined to the primary tumour. Currently, most solid primary tumours are resected, with these patients at risk from metastatic relapse as a result of residual tumour cells [Riethdorf 08]. During the invasive phase of malignant tumours, the metastatic cells break into the basal

lamina and enter the underlying connective tissue, which concurrently undergoes extensive modifications [Cancemi 08]. The processes of intravastation from the primary lesion, avoidance of immune surveillance and extravastation at the secondary site defines metastasis as an active process. Malignant tumour cells degrade ECM constituents essentially involving proteolytic enzymes produced either by the tumour cells themselves or by the surrounding stromal cells. The identification of the cellular origin of genetic and microenvironmental changes involved in the initiation, progression and the malignant conversion; the understanding of processes occurring during the establishment of secondary tumour sites; along with strategies to limit the cancer to the primary site are of medical and social interest. As ubiquitous members of the ECM microenvironment and hence cancer cell niche, HSPGs and their regulators are candidates for a central role in these processes.

At the beginning of the invasive phase, several cascade processes start to be activated, both at the cellular and extracellular levels. The host stroma undergoes modifications of the ECM composition and recruitment of host cells [Cancemi 08]. The process of matrix degradation Disassembly of the ECM for reorganisation during the processes of tissue remodeling and cell movement, including cancer progression, metastasis and angiogenesis is operated by the complex machinery of extracellular proteases, including members of the matrix metalloproteinase family and heparanase. Heparanase is the predominant mammalian enzyme degrading HS and cleaves HS side-chains into large molecular weight fragments. With heparanase gene expression shown to tightly correlate with enzyme activity [Grobe 02, Gotte 06], recent data has emerged linking heparanase overexpression with prostate cancers, gastric cancers and lymph node metastasis, including a tight correlation with grade and levels of heparanase expression [Lerner 08, Xie 08, Hoffmann 08]. *In vivo*, heparanase has been associated with remodelling of the bone microenvironment in both the developmental and metastatic context [Spiegel 08, Zhou 08]. In breast cancers, heparanase has been shown to be induced by tamoxifen [Cohen 07] and implicated in both primary tumours and metastasis [Theodoro 07]. HSPGs are well-suited to play a pivotal role in these processes, as syndecans are actively shed from the cell surface by heparanase, potentially providing the required HSPG mediated ligand/receptor interactions for downstream signalling [Cattaruzza 08]. A better understanding of these biological events of not only cancer initiation, but progression and metastasis is fundamental to biomarker discovery and their application [Cancemi 08].