Expression of genes encoding for proteins involved in heparan sulphate and chondroitin sulphate chain synthesis and modification in normal and malignant plasma cells

Caroline Bret^{1, 2, 3, 7}, Dirk Hose^{4,5}, Thierry Reme^{1, 2}, Anne-Catherine Sprynski^{1, 2}, Karène Mahtouk^{1, 2}, Jean-François Schved^{3,7}, Philippe Quittet⁶, Jean-François Rossi^{2, 6, 7}, Hartmut Goldschmidt⁴ and Bernard Klein^{1, 2, 7§}

¹INSERM U847, Equipe Labellisée LIGUE 2006, Montpellier, F-34295 France;

²CHU Montpellier, Institut de Recherche en Biothérapie, Montpellier, F-34295 France;

³CHU Montpellier, Laboratoire Central d'Hématologie, F-34295 France;

⁴Medizinische Klinik und Poliklinik V, Heidelberg, Germany;

⁵Nationales Centrum für Tumorerkrankungen, INF350, D-69115 Heidelberg, Germany;

⁶CHU Montpellier, Service d'Hématologie et Oncologie Médicale, F-34000 France;

⁷Université MONTPELLIER1, UFR Médecine, Montpellier, France.

[§]Corresponding Author: Pr. Bernard Klein, INSERM U847, Institut de Recherche en Biothérapie, CHU de Montpellier, 34295 Montpellier Cedex 5, France. Tel: +33(0) 4 67 33 78 88, Fax: +33(0) 4 67 33 04 59. bernard.klein@inserm.fr

Running title :

GEP of glycosaminoglycan synthesis pathway in plasma cells

Authors' information:

CB designed research, performed the experiments and wrote the paper. DH, MH, JFS, PQ, JFR and HG collected bone marrow samples and clinical data. TR, ACS and KM participated in the bioinformatics analysis of the data. DH and HG participated in the writing of the paper. BK is the senior investigator who designed research and wrote the paper.

Acknowledgements

This work was supported by grants from the Ligue Nationale Contre le Cancer (équipe labellisée LIGUE 2006), Paris, France, from INCA (n°R07001FN) and from MSCNET European strep (N°E06005FF), the Hopp-Foundation, Germany, the University of Heidelberg, Germany, the National Centre for Tumour Diseases, Heidelberg, Germany, the Tumorzentrum Heidelberg/Mannheim, Germany.

Summary

Syndecan-1 is a proteoglycan concentrating heparin-binding factors on the surface of multiple myeloma cells, playing likely a major role in multiple myeloma biology. As heparan sulphate and chondroitin sulphate are the bioactive components of syndecan-1, we analyzed the signature of genes encoding 100 proteins involved in synthesis of these chains, i.e. from precursor uptake to post-translational modifications, using Affymetrix microarrays.

Throughout the differentiation of memory B cells into plasmablasts and normal bone marrow plasma cells, we show that expression of enzymes required for heparan sulphate and chondroitin sulphate biosynthesis is increasing, in parallel with syndecan-1 expression. Sixteen genes were significantly different between normal and malignant plasma cells, nine of these genes *-EXT2, CHSY3, CSGALNACT1, HS3ST2, HS2ST1, CHST11, CSGALNACT2, HPSE, SULF2-* encoding for proteins involved in glycosaminoglycan chain synthesis or modifications. Kaplan-Meier analysis was performed in two independent series of patients: *B4GALT7, CSGALNACT1, HS2ST1* were associated with a good prognosis whereas *EXT1* was linked to a bad prognosis.

This study provides an overall picture of the major genes encoding for proteins involved in heparan sulphate and chondroitin sulphate synthesis and modifications that can be implicated in normal and malignant plasma cells.

Keywords : multiple myeloma, proteoglycans, heparan sulphate, chondroitin sulphate, glycosaminoglycan synthesis and modification pathways.

Introduction

Multiple myeloma (MM) is a B cell neoplasia characterized by accumulation of clonal plasma cells, *i.e.* multiple myeloma cells (MMC), in the bone marrow (BM). A hallmark of plasma cells is the expression of syndecan-1, a proteoglycan containing attachment sites for heparan sulphate (HS) and chondroitin sulphate (CS) chains (Rapraeger, *et al* 1985), (Kokenyesi and Bernfield 1994). In the BM of healthy individuals, plasma cells are the only cells expressing syndecan-1 (Costes, *et al* 1999). In patients with MM, all MMC express syndecan-1 with the exception of preapoptotic cells which rapidly lose this proteoglycan expression (Jourdan, *et al* 1998). In addition, among the major cell-surface heparan sulphate proteoglycans (HSPG), MMC only bear syndecan-1 (Mahtouk, *et al* 2006). Expression of chondroitin sulphate proteoglycans (CSPG) on MMC has not been fully reported.

HS chains can bind to a large diversity of proteins, including extracellular matrix proteins, growth factors, adhesion molecules, chemokines, interleukins, enzymes or plasma proteins, see (Bishop, *et al* 2007, Esko and Selleck 2002) for review. In addition, CS chains have been recognized as critical regulators of growth factor- and cytokine signalling (Maeda, *et al* 1996), (Hirose, *et al* 2001), (Fthenou, *et al* 2006).

HS and CS are linear glycosaminoglycans, glucosaminoglycans and galactosaminoglycans respectively, composed of disaccharide units. Both are covalently attached to the syndecan-1 core protein through the linkage region, a common tetrasaccharidic sequence for HS and CS (composed of one xylose unit, two galactose units and one glucuronic acid unit). As summed up in Figure 1, biosynthesis of these chains mainly takes place in the Golgi apparatus and involves three enzymatic steps, including formation of the linker tetrasaccharide, polymerization of disaccharide units and finally chemical modifications of the sugar backbone at various positions. This ultimate step generates diversity in oligosaccharide sequences, contributing to heterogeneity of binding, in particular through sulphatation pattern, see (Prydz and Dalen 2000) for review. In addition, HS chains can be further modified by exogenous

enzymes such as sulphatase 1 and 2 (SULF1 and SULF2), removing selectively 6-O sulphates (Morimoto-Tomita, *et al* 2002), and heparanase (HPSE), cleaving HS sulphate chains in small bioactive fragments (Vlodavsky, *et al* 1999). Moreover, the syndecan-1 ectodomain is rapidly shed by several metalloproteinases (MMP) as MMP-7 or MT-MMP-1 (membrane type matrix MMP-1) (Ding, *et al* 2005), (Kudo, *et al* 2007). The shedding can be accelerated by activation of thrombin receptor, a G-protein coupled receptor, activation of epidermal growth factor receptor or by others effectors as plasmin(Subramanian, *et al* 1997), (Fitzgerald, *et al* 2000).

In patients with MM, high serum levels of soluble syndecan-1 are linked with a poor prognosis, partly as an indicator of the tumour cell mass (Li, et al 2001). Knockdown of syndecan-1 gene expression in MMC injected into severe combined immune deficient mice results in decreased growth (Yang, et al 2007a). In addition, HPSE expression promotes syndecan-1 shedding into its soluble form in association with poor prognosis and tumour growth in vivo (Mahtouk, et al 2007a), (Yang, et al 2007b). Syndecan-1 can regulate activity of heparin-binding growth factors implicated in proliferation and survival of malignant plasma cells by presenting them to their specific receptors (Zimmermann and David 1999). This proteoglycan can thus enhance the binding of heparin-binding epidermal growth factor like growth factor (HB-EGF) or amphiregulin, increasing their biological activities (Mahtouk, et al 2005), (Mahtouk, et al 2006). Syndecan-1 is also critical to capture hepatocyte growth factor (HGF) and to present it to its receptor, cMet (Derksen, et al 2002). Additionally, fibroblast growth factor (FGF), which can induce the PI-3K/AKT pathway involved in myeloma tumour cells survival and proliferation, can bind syndecan-1 (Klein, et al 2003). The functional contribution of the chondroitin sulphate chains still remains to be elucidated. Nevertheless, a report shows the implication of chondroitin synthase 1 (CHSY1), a protein of CS synthesis pathway, in the interaction between MMC and BM microenvironment (Yin 2005).

In this paper, we analysed the expression of genes coding for the major enzymes involved in HS and CS chain synthesis and modifications. We also studied the expression of genes coding for proteins necessary for production and transport of the precursors: sugar nucleotides and 3'-phosphoadenosine 5'-phosphosulphate (PAPS). These building blocks are produced in the cytosol before being translocated to endoplasmic reticulum (ER) and/or Golgi lumens. These enzymatic systems and the translocators are likely to influence availability of substrates and therefore glycosaminoglycan chain synthesis (Toma, *et al* 1996), (Dick, *et al* 2008). We determined expression of genes linked with HS and CS chains biosynthesis in memory B cells, plasmablastic cells or normal bone marrow plasma cells using Affymetrix microarrays. These profiles were compared to those of purified MMC of a large panel of newly-diagnosed patients.

Patients, materials and Methods

Cell samples

The 15 XG human myeloma cell lines (HMCL) were obtained in our laboratory (Zhang, *et al* 1994), (Rebouissou, *et al* 1998), (Tarte, *et al* 1999). The other HMCL used - LP1, OPM2, RPMI 8226, SKMM and U266 – were purchased from the ATCC (Rockville, MD, USA). They were maintained in RPMI1640 (Gibco Invitrogen, France), 10% foetal bovine serum (FBS, PAA laboratory GmbH, Austria) and for the IL-6-dependant cell lines, with 2 ng/ml of IL-6 (Abcys SA, Paris, France).

Normal bone marrow plasma cells (BMPC) were obtained from healthy donors after informed consent was given. Plasma cells were purified using an autoMACS with anti-CD138 MACS microbeads (Miltenyi-Biotec, Paris, France) as indicated (Mahtouk, *et al* 2005). CD27⁺ memory B cells (MBC) and polyclonal plasmablasts (PPC, CD38⁺⁺, CD20⁻) were obtained and generated *in vitro* from purified CD19⁺ peripheral blood B cells, as previously described (Tarte, *et al* 2002).

MMC of a total of 171 patients with previously-untreated MM were included in this study after written informed consent was given (123 patients were included at the University hospital of Heidelberg (Germany) and 48 were from Montpellier (France)). These 171 patients were treated with high dose therapy (HDC) and autologous stem cell transplantation (ASCT) (Goldschmidt, *et al* 2000) and were termed in the following Heidelberg-Montpellier (HM) series. According to the Durie-Salmon classification (Durie and Salmon 1975), 15 patients were of stage IA (who progressed into therapy requiring MM), 24 of stage IIA, 2 of stage IIB, 121 of stage IIIA, and 9 of stage IIIB. According to the International Staging System (ISS), 76 patients were of stage I, 66 of stage II, 29 of stage III. 28 patients had IgA κ MM, 12 IgA λ MM, 64 IgG κ MM, 33 IgG λ MM, 1 IgD λ MM, 18 Bence-Jones κ MM, 11 Bence-Jones λ MM, and 4 non-secreting MM. We also used Affymetrix data of a cohort of

345 purified MMC from previously-untreated patients from the Arkansas Research Group (Little Rock). The patients were treated with total therapy 2 and termed in the following LR-TT2 series (Barlogie, *et al* 2006). These data are publicly available in the NIH Gene Expression Omnibus (GEO) [http://www.ncbi.nlm.nih.gov/geo/] under accession number GSE2658.

Bone marrow stromal cells (MSC) were obtained from the bone marrow of 5 healthy volunteers after informed consent was given and osteoclasts (OST) were obtained *in vitro* from peripheral blood mononuclear cells as indicated (Corre, *et al* 2007), (Moreaux, *et al* 2005). Bone marrow T cells (CD3), monocytes (CD14) and polymorphonuclear neutrophils (PMN) were purified from 5 newly-diagnosed patients after informed consent was given, as indicated (Moreaux, *et al* 2005).

Preparation of cRNA, microarray hybridization and gene expression profiling analysis

Microarray experiments were performed in the Institute of Research in Biotherapy at the Montpellier University Hospital (France) (http://irb.chumontpellier.fr/en/laboratories_microarray.html). RNA was extracted with the RNeasy Kit (Qiagen, Valencia, CA, USA), the SV-total RNA extraction kit (Promega, Mannheim, Germany) and Trizol (Invitrogen, Karlsruhe, Germany) in accordance to the manufacturer's instructions as indicated (De Vos, et al 2002). RNA was amplified and hybridized (small sample labeling protocol VII, Affymetrix, Santa Clara, CA, USA) to the human genome HG-U133A and HG-U133B or HG-U133 Plus 2.0 GeneChip oligonucleotides arrays, according to the manufacturer's instructions (Affymetrix). Fluorescence intensities were quantified and analyzed using the GCOS (GeneChip Operating Software) software (Affymetrix). Arrays were scaled to an average intensity of 100. A treshold of 1 was assigned to values <1. In the Affymetrix HG-U133 microarrays, a gene is probed by 11 pairs of perfect-match/mismatch oligonucleotides randomly spread over the chip. The signed rank MAS5 algorithm decides after scanning if the corresponding gene can be statistically declared "present" (P call) or

8

"absent" (A call), and delivers a weighted fluorescence signal. Gene expression data were analysed with our bioinformatics platform RAGE [http://rage.montp.inserm.fr/] and Amazonia [http://amazonia.montp.inserm.fr] (Assou, *et al* 2007). We also used the gene expression profile from 79 human normal tissues samples available from Hogenesch's group on a public data base (Su, *et al* 2004).

Real-time RT-PCR

RNA was extracted using the RNeasy Kit (Qiagen). We generated cDNA from 100 ng total RNA using Superscript II reverse transcriptase (Invitrogen, Cergy-Pontoise, France). For real-time reverse transcriptase–polymerase chain reaction (RT-PCR), we used Assay-on-Demand primers and probes and the TaqMan Universal Master Mix from Applied Biosystems (Courtaboeuf, France) according to the manufacturer's instructions. Gene expression of *SULF2*, *GP1*, *CHSY3* and *HS3ST2* was measured using the ABI Prism 7000 Sequence Detection System. For each sample, the cycle threshold (Ct) value for the gene of interest was determined and normalized to its respective Ct value for β 2-microglobulin (*B2M*). Then it was compared with a cell type used as a positive control (XG-11 for *SULF2*, *GP1*, *CHSY3* and one OST sample for *HS3ST2* since this gene was not expressed by HMCL) using the following formula: $100/2 \Delta \Delta Ct$, in which threshold $\Delta \Delta Ct = (Ct_{unknown} - Ct_{unknown B2M}) - (Ct_{positive control} - Ct_{B2M positive control})$. The coefficient of correlation and *P* values for the correlations between microarrays and real time RT-PCR are indicated on each graph. *P* values were determined with a Spearman test using SPSS Software. Statistical significance was obtained for P value ≤ 0.05 .

Statistical analysis

To cluster the samples according to the similarity of their gene expression patterns, we performed an unsupervised principal component analysis (PCA) with RAGE bioinformatics platform [http://rage.montp.inserm.fr/]. On the graphs related to PCA, the first principal

component (pc1) is represented by the X-axis, the second principal component (pc2) by the Y-axis. The percentage indicated on these graphs gives the information of variance for these 2 components.

SAM (significance analysis of microarrays) analysis was applied to our probe sets of interest in the different samples with 1000 permutations, a fold change of 2 and a false discovery rate of 0%.

When PCA and SAM analysis were performed with MMC data, we used data obtained from gene expression analysis of the MMC purified from samples of the 123 patients included in the hospital of Heidelberg.

To determine the link with prognosis for a given probe set, patients were ranked in 4 quartiles according to probe set expression and the event-free (EFS) or overall survivals (OAS) of patients in the lowest quartile were compared to those of patients in the highest quartile. The survival curves were plotted using the Kaplan-Meier method. The statistical significance of differences in overall survival and event-free survival between groups of patients was estimated by the log-rank test (P \leq 0.05). An event was defined as death. To avoid multiple testing correction, we retained probe sets associated with prognosis in 2 independent series of previously-untreated patients: the HM series of 171 patients and the LR-TT2 series of 345 patients.

Results

Affymetrix microarrays to investigate pathways involved in heparan sulphate and chondroitin sulphate chain synthesis and modifications

In Figure 1, proteins involved in synthesis of HS and CS chains, from precursor uptake to post-translational modifications, are listed. These proteins, including their different isoforms, are encoded by 100 genes. Out of these genes, expression of 75 might be evaluated in B and plasma cells using 123 probe sets with Affymetrix U133 Plus 2:0 microarrays (see Table SI). 25 genes could not be studied in reason of a lack of probe set (*GLUT7, UGP1*) or to defective probe sets having an absent call or a weak Affymetrix signal (below 100) in our samples and a panel of 79 human tissues (Hogenesch data base) (Su, *et al* 2004).

The first step of HS and CS synthesis corresponds to cytoplasmic synthesis of nucleotide sugars and PAPS, the universal donor of sulphate. We evaluated expression of 16 genes coding for membrane transporters (9 genes for glucose transport and 7 genes for sulphate anion transport), 19 genes coding for enzymes leading to nucleotide sugar synthesis and 2 genes coding for proteins required during sulphate activation in PAPS.

The second step leads to the translocation of these precursors within the lumen of ER and Golgi apparatus where synthesis of HS and CS takes place: 7 genes coding for nucleotide sugar transporters and 2 genes coding for PAPS transporters could be studied with U133 Affymetrix data. In these organelles, a complex set of enzyme reactions encoded by 26 out of the 75 available genes from Affymetrix chips is involved in linkage region synthesis and in HS and CS chain polymerization and modifications. We also analyzed the expression of the 3 genes coding for extracellular enzymes able to modify the pattern of sulphatation (*SULF1 and 2*) or to produce HS fragments of 10 to 20 sugar units long (*HPSE*).

Gene expression profiles associated with normal B cell and plasma cell samples

To determine whether the normal populations of memory B cells (MBC), plasmablasts (PPC) and bone marrow plasma cells (BMPC) present a specific expression profile for HS and CS synthesis and modifications pathways, an unsupervised principal component analysis (PCA) using the 123 probe sets was performed. PCA clearly delineated three distinct clusters: MBC, PPC and BMPC (Fig 2A). The first principal component accounted for 39% of variance and separated MBC from PPC and BMPC. The second principal component accounted for 24% of variance separating PPC and BMPC. We then identified the genes separating MBC from PPC + BMPC along the first PCA axis and PPC from BMPC along the second axis. We used a supervised analysis with the SAM (significance analysis of microarrays) software (1000 permutations, 2-fold change in expression with a 0% false discovery rate (FDR)) in order to identify probe sets linked to each type of cells. Results in Table I underline the weak level of expression of genes involved in each step of synthesis and modifications of HS and CS chains in MBC compared to plasma cells. PPC and BMPC cluster was associated with 33 overexpressed probe sets interrogating 25 genes and MBC cluster by 8 overexpressed probe sets interrogating 4 genes only. Out of the 25 plasma cell genes, 1 encoded for glucose transporter (SLC2A5), 2 for a sulphate anion transporter (SLC26A2, SLC26A6), 1 for an activator of sulphate in PAPS (PAPSS1), 1 for a PAPS transporter (SLC35B3), 11 for nucleotide sugar synthesis enzymes (GALK1, GALK2, PGM1, PGM2, PGM3, GFPT1, GNPNAT1, UAP1, UGDH, UGP2 and GNPDA1), 2 for a nucleotide sugar transporter (SLC35B1, SLC35A4), 2 for enzymes implicated in HS synthesis and modification (EXT1, HS2ST1) and 5 for enzymes involved in CS synthesis and modification (CHPF, CHST11, CHST12, GALNAC4S-6ST and CSGlcA-T).

Regarding genes dissociating PPC and BMPC along the 2nd PCA axis, 10 genes were increased in PPC compared to BMPC - *SLC2A3, HK2, GPI, PGM2, GFPT1, GALK2, PAPSS1, SLC35D1, SLC35B4*, and *HS2ST1* - and 12 genes were overexpressed in BMPC compared to PPC - *SULF2, PAPSS2, UAP1, GALK1, CHST12, CSGlcA-T, CSGALNACT1,*

HS3ST2, HS6ST1, SLC2A10, SLC35D2, and SLC26A11 (Table II). These data emphasize an increased expression of genes coding for proteins involved in synthesis of HS and CS chains throughout B cell differentiation, in PPC and mainly in BMPC. Thus, out of the 123 probe sets initially included, a total of 63 probe sets discriminates the 3 populations (33 + 8 probe sets discriminating B cells from plasma cells and 10 + 12 probe sets separating plasmablasts from mature plasma cells).

Gene expression profiles linked with multiple myeloma cells

Unsupervised PCA clustering of MMC together with normal MBC, PPC and BMPC delineated two groups according to first PCA component accounting for 18% of variance: a group of MMC only (principal component 1>0) and a group of normal cells and MMC (principal component 1<0) (Fig 2B). Similar clustering was obtained with the limited list of 63 probe sets reported above. Running only normal plasmablasts and mature plasma cells together with MMC yielded the same clustering. In order to identify probe sets overexpressed in MMC compared to normal plasma cells (PPC + BMPC), normal PPC or normal BMPC, a supervised analysis was realized with SAM software (1000 permutations, fold-change ≥ 2 , FDR = 0%). Expression of 16/75 genes was significantly different between BMPC and MMC. Results are shown in Table III. Twelve probe sets interrogating 11 genes were up-regulated in MMC in comparison with BMPC: HK2, GPI, GNPDA2, PAPSS1, PAPSS2, EXT2, HS2ST1, CHST11, CHSY3, HPSE and SULF2. CHSY3 gene presented the highest variance among all the 75 genes studied. CHSY3 was only expressed by MMC, unlike normal BMPC, PPC or MBC (Fig 3A) and had a present call in MMC of 105 out of 123 patients included in the Heidelberg patient cohort. Two other genes are remarkable because the encoded proteins have already documented functions in MM (SULF2) (Dai, et al 2005) or other cancers (GPI) (Watanabe, et al 1996), (Tsutsumi, et al 2003b). SULF2 gene was overexpressed 3.7 fold in MMC in comparison with BMPC (Fig 3B). GPI, whose expression was significantly enhanced in PPC, had a 2.08-fold change in expression between MMC and BMPC (Fig 3C).

Moreover, 7 probe sets interrogating 5 genes were under expressed in MMC compared to BMPC: *SLC2A3, GALK1, HS3ST2, CSGALNACT2* and *CSGALNACT1*. A loss of *HS3ST2* expression has been documented in some epithelial cancers (Tsutsumi, *et al* 2003b), and this is also the case in 39 out of the 123 MMC purified from patients of the Heidelberg group (Fig 3D).

Microarray data for these four genes of interest (*CHSY3*, *SULF2*, *GPI*, and *HS3ST2*) were validated (Fig 4) by real-time RT-PCR ($P \le 0.05$) using RNA of 7 human myeloma cell lines (HMCL) or 7 osteoclast (OST) samples (*HS3ST2* probe set is not expressed in HMCL). Supplementary information regarding genes differentially expressed between MMC and PPC or MMC and PPC+BMPC is given in Table SII.

Gene expression profiles associated with human myeloma cell lines

Unsupervised PCA clustering of primary MMC and HMCL using the 123 probe sets defined previously yielded to 2 well identified clusters, a MMC and a HMCL cluster. The first PCA component accounted for 24% and the second one for 8% (data not shown). Using SAM analysis (1000 permutations, fold change \geq 2, FDR = 0%), 47 of the 123 probe sets were differentially expressed between MMC and HMCL: 30 being overexpressed in MMC and the 17 others in HMCL (Table IV). The 30 MMC probe sets interrogated 4 genes implicated in glucose transport (*SLC2A1*, *SLC2A9*, *SLC2A10*, *SLC2A13*), 1 gene involved in sulphate anion transport (*SLC26A1*), 1 gene coding for PAPSS2, 1 gene coding for SLC35B3, 2 genes implicated in sugar metabolism (*GALK1*, *PGM3*), 1 gene coding for the nucleotide sugar transporter SLC35D2, 5 genes implicated in HS synthesis and modifications (*EXT1*, *HS2ST1*, *HS3ST2*, *HS3ST4* and *NDST1*), 1 gene coding for SULF2, and 5 genes involved in CS chain polymerization and modifications: *CSGALNACT1*, *GalNAc4S-6ST*, *CHST11*, *CHSY3* and *CHPF*. This last observation underlines an increased expression of enzymes implicated in CS chain polymerization in MMC. Among the 17 probe sets up-regulated in HMCL, 9 probe sets are linked with genes encoding for enzymes involved in sugar metabolism: *PGM2*, *GALE*,

GALK2, *GNPDA1*, *HK2*, *GPI*, *GFPT1*, *GNPNAT1*, and *GALM*. This result indicates an enhanced sugar metabolism capacity in cell lines. Other genes encoded for the SLC35B4 nucleotide sugar transporter, for B3GAT3, for HS3ST3B1, for SULF1 and for HPSE.

Expression of genes coding for extracellular enzymes involved in HS modifications

The function of HSPG, and thus of syndecan-1, can be regulated by extracellular enzymes able to modulate the structure of HS chains. Among these enzymes, 2 sulphatases (SULF1 and SULF2) are known to possess endoglucosamine-6 sulphatase activity. Another enzyme, HPSE, acts as an endoglucuronidase that can cleave HS chains, producing fragments more biologically active than native chains. Given the major role of the BM microenvironment in the pathogenesis of myeloma, see (Mitsiades, et al 2006) for review, we examined gene expression level of these 3 enzymes in bone marrow T cells (CD3), monocytes (CD14) and polymorphonuclear neutrophils (PMN) purified from 5 patients with newly diagnosed myeloma. In addition, bone marrow stromal cells (MSC) (n= 5) and OST (n=7) were generated in vitro from cells of 5 healthy donors and 7 patients with myeloma, respectively. As it can be seen in Figure 5, SULF1 which was not expressed by MMC (data not shown) was exclusively expressed by MSC. SULF2 was expressed by all types of cells studied, except for CD3 cells. Besides, an over expression of this enzyme was observed in CD14 cells (median, 1678) and PMN (median, 724) in comparison with MSC (median, 267) and OST (median, 67). Moreover, HPSE was mainly expressed by OST and CD14 cells as previously published (Mahtouk, et al 2007b). Taken together, these data suggest that the BM environment could be involved in the regulation of the structure of HS chains and thus in their function.

Probe sets linked to prognosis

Out of the 123 probe sets, 4 were associated with prognosis in the 2 independent series of previously-untreated patients, the HM series of 171 patients and the LR-TT2 series of 345

patients treated with TT2 therapy. Overall survival (OAS) and event-free survival (EFS) data are shown in Figure 6 and are summed up in Table V. *B4GALT7*, *CSGALNACT1* and *HS2ST1* expressions were correlated with a good prognosis and *EXT1* expression was associated with a bad prognosis. Expression of *B4GALT7*, *CSGALNACT1*, *HS2ST1* and *EXT1* was evaluated in the 8 MM subtypes defined by Zhan et al. (Table VI) (Zhan, *et al* 2006). The poor prognosis of *EXT1* is related to its higher expression in the poor prognosis group of patients with t(4;14) and spiked MMSET (MS). On the other hand, the frequency of the good prognosis genes is decreased in the Proliferation (PR) and MMSET (MS) bad prognosis groups.

Discussion

In the present study, we have first analysed the change in expression profiles of 75 genes that could be interrogated with Affymetrix microarrays and encoding for HS and CS chains biosynthesis and modifications pathways, throughout B cell to plasma cell differentiation. These chains are the major bioactive components of syndecan-1 that is a hallmark of plasma cell differentiation. Our work shows that, in the differentiation from MBC to PPC and BMPC, gene expression of enzymes required for HS and CS biosynthesis is increasing, in parallel with syndecan-1 expression in the cells. The expression of these 75 genes makes it possible to classify MBC, PPC and BMPC into 3 distinct and homogeneous clusters, characterized by a set of overexpressed genes. Only 4/75 genes were demonstrated to have an increased expression in MBC compared to other types of cells. These genes encode for proteins involved in precursor transport and synthesis, which are probably necessary for other metabolic processes. The 22 genes defining the PPC cluster indicate that these cells are, above all, able to enhance their production of PAPS and of nucleotide sugars, and then to promote their transport within ER and Golgi apparatus. Gene expression profile of PPC also indicates that these cells possess the majority of the different enzymes required for linkage region synthesis, polymerization and modifications of HS and CS chains, even if syndecan-1 expression in PPC is low (data not shown). Thus, PPC have the machinery to enhance their metabolic potential, with a special effect on sugar metabolism, and are then prompt to increase the synthesis of syndecan-1 HS and CS chains when differentiating into mature plasma cells (PC). The BMPC cluster is characterized by the high expression of the overall enzymes required for HS and CS synthesis and modifications in accordance with the high levels of syndecan-1 HS and CS chains in these cells. Of note, mature PC have to continuously synthesize a lot of these chains since membrane syndecan-1 has a 2-hour turnover, being rapidly cleaved by MMP into soluble syndecan-1. Among the genes that were highly overexpressed in BMPC in comparison to PPC, HS3ST2 and SULF2 encode for a sulphotransferase and for a sulphatase, respectively. These enzymes are involved in the late modifications of heparan sulphate chains. They are responsible of their structural diversity by modulating sulphatation, which governs interactions between HSPG and a broad range of proteins.

Regarding MMC, the unsupervised PCA analysis indicates that MMC are closed to BMPC, MBC and PPC being in distant clusters. Sixteen genes are however differently expressed between MMC and BMPC with a fold change ranging from 10.33 (*CHSY3*) to 2 fold (*GNPDA2*) with a 0% FDR.

Among these genes, one is involved in initial steps of precursor synthesis: *GPI*. This gene encodes for an enzyme responsible for isomerization of glucose-6-phosphate and fructose-6-phosphate. In our study, it was associated with OAS in Kaplan–Meier analysis realized from data of the LR-TT2 cohort but was not significant in HM cohort (data not shown). In LR-TT2, upper levels of expression appeared to be associated with shorter OAS, suggesting that this parameter could be linked to a bad prognosis. GPI is an ubiquitous cytosolic enzyme that plays a critical role in sugar metabolism. Interestingly, this protein can be secreted and molecular cloning and sequencing studies have ascribed multiple identities to the secreted form: autocrine motility factor, neuroleukin and maturation factor (Watanabe, *et al* 1996). GPI is thus considered as being a multifunctional protein and is involved in different malignant processes as proliferation, survival or angiogenesis (Funasaka, *et al* 2001), (Tsutsumi, *et al* 2003a). Its contribution to MM is not known but could be of particular interest.

Of note, among the genes differently expressed between MMC and BMPC, 9 genes – *EXT2*, *CHSY3*, *CSGALNACT1*, *HS2ST1*, *HS3ST1*, *CHST11*, *CSGALNACT2*, *HPSE*, *SULF2*-encode for proteins involved in HS or CS chain synthesis or modifications.

EXT2 gene encodes for exostose multiple 2, a protein involved in the first step of HS chain polymerization. Of note, *EXT2* is one of the three genetic loci, with *EXT1* and *EXT3*, mutated in hereditary multiple exostoses, an autosomal dominant disease characterized by the

formation of osteochondromas in relation with a disorganization of chondrocytes, leading to random ossification (Zak, et al 2002). The protein EXT2 combines with EXT1 to form an hetero-oligomeric complex that catalyzes the HS chain elongation (Busse, et al 2007). These enzymes do not appear to play a redundant role in HS polymerization (Zak, et al 2002). In our study, we observed that *EXT1* expression in MMC is associated with an adverse prognosis in the 2 independent patient series. This result is not in accordance with previous reports suggesting that EXT1 functions as a tumour suppressor gene in patients with hereditary multiple exostoses (Hecht, et al 1995, Piao, et al 1997). In multiple myeloma, apart from its role in HS chain polymerization, another function of the protein EXT1 should be investigated. HS chains bind to a plethora of proteins, including different growth factors involved in myeloma biology. Sulphatation pattern of glycosaminoglycan chains is a critical element responsible for diversity in oligosaccharide sequences and, as a consequence, for heterogeneity of binding. HS2ST1 encodes for a 2-O sulphotransferase (heparan sulphate 2-Osulphotransferase 1). This enzyme transfers sulphate to the C-2 position of the iduronic acid residues of HS chains. In our study, we observed an overexpression of this gene in MMC. Besides, the gene encoding for this enzyme was correlated to a good prognosis. Additionally, HS3ST2, an other enzyme implicated in HS chain modifications, is of interest. This protein exhibits heparan sulphate glucosaminyl 3-O-sulphotransferase activity and is thus able to modulate sulphatation pattern in C-3 position of the sugar backbone. Interestingly, a loss of HS3ST2 expression has been documented in some epithelial cancers (Tsutsumi, et al 2003b) and a similar result was obtained in our study when comparing BMPC and MMC, a decreased expression being obtained in this last category of cells. These observations are consistent with a major role of HS sulphatation in myeloma biology, most likely through the modulation of growth factor binding.

Notably, our study is the first one to focus on several genes involved in CS chain synthesis and modification in myeloma. *CSGALNACT1*, also known as chondroitin sulphate N-acetylgalactosaminyltransferase 1, is involved in the initiation of CS chain synthesis (Sato, *et*

al 2003). Our results show that overexpression of this gene is associated with a good prognosis in HM and LR-TT2 groups. CHSY3 gene encodes for a glycosyltransferase implicated in CS synthesis (Yada, et al 2003). Our data show that CHSY3 is not expressed by normal BMPC but by MMC with a high level of expression. Moreover, CHSY3 correlated with a good prognosis in Kaplan-Meier analysis realized from EFS data of the HM cohort and from OAS data of the LR-TT2 cohort (data not shown). Recently, it was shown that coexpression of CHSY3 with CHPF (CHPF gene is expressed by MMC, data not shown) confers a CS polymerization activity (Izumikawa, et al 2007). Another enzyme involved in CS sulphatation, CHST11, was overexpressed in MMC too. We can speculate that these enzymes could extend and modify CS chains in MMC conferring on them more controlled growth rate and better prognosis. They could be involved in MM biology and clinical response through various proteoglycans. This remains to be elucidated. To examine expression of proteoglycans with CS chains in our samples, a supervised analysis was realized with SAM software as described previously (see Table SIII, Table SIV and Table SV). We first compared MBC samples to the group of PPC+BMPC. The MBC population was characterized by an overexpression of SPOCK2 (a gene encoding for testican 2) whereas the genes APLP2 and SRGN (a gene encoding for serglycin) were overexpressed in the cluster PPC+BMPC. Of note, the secretion of serglycin by HMCL and MMC has already been described (Theocharis, et al 2006). When we compared BMPC and MMC, we found that VCAN and COL9A2 genes were overexpressed in normal plasma cells whereas none of the genes encoding for CSPG was overexpressed in MMC.

Two MMC genes encode for enzymes involved in HS chain modifications: *SULF2* and *HPSE*. *SULF2* gene is already overexpressed in BMPC compared to PPC and its expression increases in MMC. Recent data have shown that SULF2, and the first member of the corresponding endosulphatase family SULF1, act as strong inhibitors of myeloma tumour growth *in vivo* (Dai, *et al* 2005). In our study, *SULF1* was not expressed by normal plasma cells or MMC but by MSC, which further underlines the role of the myeloma

20

microenvironment. Of note, *SULF1* is expressed by some HMCL emphasizing that these HMCL can express environment genes, making it possible to escape from environment dependence. This is the case for IL-6 (Jernberg-Wiklund, *et al* 1992), HB-EGF (De Vos, *et al* 2001), (Wang, *et al* 2002) or for heparanase (Mahtouk, *et al* 2007a). This last enzyme cleaves HS chain into bioactive 10-20 saccharide fragments. Transduction of *HPSE* gene in HMCL confers a growth advantage *in vivo* (Yang, *et al* 2005). One mechanism is that heparanase can control syndecan-1 gene expression and syndecan-1 shedding into a soluble form that also confers MMC growth advantage (Yang, *et al* 2007b).

Among the genes whose expression is associated to prognosis in our study, B4GALT7 encodes for a glycosyltransferase (beta 1,4-galactosyltransferase I) involved in HS chain polymerization in the *trans*-Golgi apparatus. In addition, this enzyme can be located at the cell membrane and functions as a receptor for extracellular glycoside ligands by binding terminal N-acetylglucosamine residues (Rodeheffer and Shur 2002). Mutations of this gene resulting in absent or defective enzyme activity have been associated with Ehlers-Danlos syndrome (progeroid variant) (Furukawa and Okajima 2002). Additionally, several studies have suggested a contribution of B4GALT7 to cancer development and metastasis. Expression of this enzyme was increased in astrocytomas in comparison with normal brain tissue (Xu, et al 2001). An overexpression of this enzyme has been observed in hepatoma in comparison with normal liver tissues, the gene B4GALT7 being under the control of the viral regulatory protein Hbx. Besides, in this model, an up-regulation of B4GALT7 was associated with hepatoma cell growth in vivo and in vitro (Wei, et al 2008). In other studies, the surface activity of the enzyme was a determining factor for invasive behaviour of murine melanoma cell lines whereas activity of the Golgi enzyme remained unchanged in cells with different metastatic and invasive capacities (Johnson and Shur 1999, Passaniti and Hart 1990). The same observation was made for metastatic lung cancer cells, an up-regulation of B4GALT7 being under the control of the transcription factor E1AF (Zhu, et al 2005). In the current study, the result related to B4GALT7 challenges these data, this gene being linked to good prognosis. B4GALT7 is involved in the initial step of linkage region synthesis in Golgi apparatus. As this step is shared by HS and CS chains biosynthesis, one hypothesis is that *B4GALT7* overexpression promotes production of proteoglycans other than syndecan-1 and is able to control myeloma growth. The cellular location of B4GALT7 has not been described in myeloma cells and could be studied to understand the role of this enzyme in myeloma biology.

In conclusion, this study shows an increased expression of enzymes required for HS and CS biosynthesis in parallel with syndecan-1 expression through B cell differentiation from MBC to BMPC. Of interest, some of these genes correlate with prognosis, which suggests their important function in MM biology. Additionally, this study underlines enzymes implicated in CS synthesis and modification. Their implication in MM could be linked to CS proteoglycans expressed by plasma cells.

List of abbreviations

Acetyl-CoA: acetyl-Coenzyme A; ADP: adenosine di-phosphate; ASCT: autologous stem cell transplantation; ATP: adenosine tri-phosphate; B2M: β2-microglobulin; BM: bone marrow; BMPC: bone marrow plasma cell; CD14 : monocyte; CD3 : bone marrow T cell; CHSY1: chondroitin synthase 1; CS: chondroitin sulphate; CSPG: chondroitin sulphate proteoglycan: Ct: cycle threshold; EFS: event-free survival; ER: endoplasmic reticulum; FDR: false discovery rate; FGF: fibroblast growth factor; Fru: fructose; Gal: galactose; GalNAc: Nacetyl galactosamine; GEP: gene expression profile; GlcN: glucosamine; GlcNAc: Nacetylglucosamine; GCOS: GeneChip Operating Software (Affymetrix); Glc: glucose; HB-EGF: heparin-binding epidermal growth factor like growth factor; HDC: high dose therapy; HGF: hepatocyte growth factor; HM: Heidelberg-Montpellier series of patients; HS: heparan sulphate; HSPG: heparan-sulphate proteoglycan; LR-TT2: Little Rock series of patients treated with total therapy 2; MBC: memory B cell; MM: multiple myeloma; MMC: multiple myeloma cell; MMP: metalloproteinase; MS: MMSET (multiple myeloma SET domain); MSC: bone marrow stromal cell; NAD: nicotinamide adenine dinucleotide; OAS: overall survival; OST: osteoclast; PAPS: 3'-phosphoadenosine 5'-phosphosulphate; PC: plasma cell; pc1: first principal component; pc2: second principal component; PCA: principal component analysis; PI3K: phosphatidylinositol 3-kinase; PMN: polymorphonuclear neutrophil; PPC: plasmablast; PPi: inorganic phosphate; PR: proliferation; SAM: significance analysis of microarrays; UDP: uridine di-phosphate; UTP: uridine tri-phosphate; Xyl: xylose.

References

- Assou, S., Le Carrour, T., Tondeur, S., Strom, S., Gabelle, A., Marty, S., Nadal, L., Pantesco, V., Reme, T., Hugnot, J.P., Gasca, S., Hovatta, O., Hamamah, S., Klein, B. & De Vos, J. (2007) A meta-analysis of human embryonic stem cells transcriptome integrated into a web-based expression atlas. *Stem Cells*, 25, 961-973.
- Barlogie, B., Tricot, G., Rasmussen, E., Anaissie, E., van Rhee, F., Zangari, M., Fassas, A., Hollmig, K., Pineda-Roman, M., Shaughnessy, J., Epstein, J. & Crowley, J. (2006) Total therapy 2 without thalidomide in comparison with total therapy 1: role of intensified induction and posttransplantation consolidation therapies. *Blood*, **107**, 2633-2638.
- Bishop, J.R., Schuksz, M. & Esko, J.D. (2007) Heparan sulphate proteoglycans fine-tune mammalian physiology. *Nature*, **446**, 1030-1037.
- Busse, M., Feta, A., Presto, J., Wilen, M., Gronning, M., Kjellen, L. & Kusche-Gullberg, M. (2007) Contribution of EXT1, EXT2, and EXTL3 to heparan sulfate chain elongation. *J Biol Chem*, **282**, 32802-32810.
- Corre, J., Mahtouk, K., Attal, M., Gadelorge, M., Huynh, A., Fleury-Cappellesso, S., Danho, C., Laharrague, P., Klein, B., Reme, T. & Bourin, P. (2007) Bone marrow mesenchymal stem cells are abnormal in multiple myeloma. *Leukemia*, 21, 1079-1088.
- Costes, V., Magen, V., Legouffe, E., Durand, L., Baldet, P., Rossi, J.F., Klein, B. & Brochier, J. (1999) The Mi15 monoclonal antibody (anti-syndecan-1) is a reliable marker for quantifying plasma cells in paraffin-embedded bone marrow biopsy specimens. *Human Pathology*, **30**, 1405-1411.
- Dai, Y., Yang, Y., MacLeod, V., Yue, X., Rapraeger, A.C., Shriver, Z., Venkataraman, G., Sasisekharan, R. & Sanderson, R.D. (2005) HSulf-1 and HSulf-2 are potent inhibitors of myeloma tumor growth in vivo. *J Biol Chem*, **280**, 40066-40073.
- De Vos, J., Couderc, G., Tarte, K., Jourdan, M., Requirand, G., Delteil, M.C., Rossi, J.F., Mechti, N. & Klein, B. (2001) Identifying intercellular signaling genes expressed in malignant plasma cells by using complementary DNA arrays. *Blood*, **98**, 771-780.
- De Vos, J., Thykjaer, T., Tarte, K., Ensslen, M., Raynaud, P., Requirand, G., Pellet, F., Pantesco, V., Reme, T., Jourdan, M., Rossi, J.F., Orntoft, T. & Klein, B. (2002) Comparison of gene expression profiling between malignant and normal plasma cells with oligonucleotide arrays. *Oncogene*, **21**, 6848-6857.
- Derksen, P.W., Keehnen, R.M., Evers, L.M., van Oers, M.H., Spaargaren, M. & Pals, S.T. (2002) Cell surface proteoglycan syndecan-1 mediates hepatocyte growth factor binding and promotes Met signaling in multiple myeloma. *Blood*, **99**, 1405-1410.
- Dick, G., Grondahl, F. & Prydz, K. (2008) Overexpression of the 3'-phosphoadenosine 5'phosphosulfate (PAPS) transporter 1 increases sulfation of chondroitin sulfate in the apical pathway of MDCK II cells. *Glycobiology*, **18**, 53-65.
- Ding, K., Lopez-Burks, M., Sanchez-Duran, J.A., Korc, M. & Lander, A.D. (2005) Growth factor-induced shedding of syndecan-1 confers glypican-1 dependence on mitogenic responses of cancer cells. *J Cell Biol*, **171**, 729-738.
- Durie, B.G. & Salmon, S.E. (1975) A clinical staging system for multiple myeloma. Correlation of measured myeloma cell mass with presenting clinical features, response to treatment, and survival. *Cancer*, **36**, 842-854.
- Esko, J.D. & Selleck, S.B. (2002) Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annual Review of Biochemistry*, **71**, 435-471.
- Fitzgerald, M.L., Wang, Z., Park, P.W., Murphy, G. & Bernfield, M. (2000) Shedding of syndecan-1 and -4 ectodomains is regulated by multiple signaling pathways and mediated by a TIMP-3-sensitive metalloproteinase. *J Cell Biol*, **148**, 811-824.
- Fthenou, E., Zafiropoulos, A., Tsatsakis, A., Stathopoulos, A., Karamanos, N.K. & Tzanakakis, G.N. (2006) Chondroitin sulfate A chains enhance platelet derived growth

factor-mediated signalling in fibrosarcoma cells. *International Journal of Biochemistry and Cell Biology*, **38**, 2141-2150.

- Funasaka, T., Haga, A., Raz, A. & Nagase, H. (2001) Tumor autocrine motility factor is an angiogenic factor that stimulates endothelial cell motility. *Biochem Biophys Res Commun*, 285, 118-128.
- Furukawa, K. & Okajima, T. (2002) Galactosyltransferase I is a gene responsible for progeroid variant of Ehlers-Danlos syndrome: molecular cloning and identification of mutations. *Biochim Biophys Acta*, **1573**, 377-381.
- Goldschmidt, H., Egerer, G. & Ho, A.D. (2000) Autologous and allogeneic stem cell transplantation in multiple myeloma. *Bone Marrow Transplant*, **25 Suppl 2**, S25-26.
- Hecht, J.T., Hogue, D., Strong, L.C., Hansen, M.F., Blanton, S.H. & Wagner, M. (1995) Hereditary multiple exostosis and chondrosarcoma: linkage to chromosome II and loss of heterozygosity for EXT-linked markers on chromosomes II and 8. *American Journal of Human Genetics*, 56, 1125-1131.
- Hirose, J., Kawashima, H., Yoshie, O., Tashiro, K. & Miyasaka, M. (2001) Versican interacts with chemokines and modulates cellular responses. *J Biol Chem*, **276**, 5228-5234.
- Izumikawa, T., Uyama, T., Okuura, Y., Sugahara, K. & Kitagawa, H. (2007) Involvement of chondroitin sulfate synthase-3 (chondroitin synthase-2) in chondroitin polymerization through its interaction with chondroitin synthase-1 or chondroitin-polymerizing factor. *Biochem J*, **403**, 545-552.
- Jernberg-Wiklund, H., Pettersson, M., Carlsson, M. & Nilsson, K. (1992) Increase in interleukin 6 (IL-6) and IL-6 receptor expression in a human multiple myeloma cell line, U-266, during long-term in vitro culture and the development of a possible autocrine IL-6 loop. *Leukemia*, 6, 310-318.
- Johnson, F.M. & Shur, B.D. (1999) The level of cell surface beta1,4-galactosyltransferase I influences the invasive potential of murine melanoma cells. *J Cell Sci*, **112** (**Pt 16**), 2785-2795.
- Jourdan, M., Ferlin, M., Legouffe, E., Horvathova, M., Liautard, J., Rossi, J.F., Wijdenes, J., Brochier, J. & Klein, B. (1998) The myeloma cell antigen syndecan-1 is lost by apoptotic myeloma cells. *Br.J Haematol.*, **100**, 637-646.
- Klein, B., Tarte, K., Jourdan, M., Mathouk, K., Moreaux, J., Jourdan, E., Legouffe, E., De Vos, J. & Rossi, J.F. (2003) Survival and proliferation factors of normal and malignant plasma cells. *Int J Hematol*, **78**, 106-113.
- Kokenyesi, R. & Bernfield, M. (1994) Core protein structure and sequence determine the site and presence of heparan sulfate and chondroitin sulfate on syndecan-1. *J Biol Chem*, **269**, 12304-12309.
- Kudo, T., Takino, T., Miyamori, H., Thompson, E.W. & Sato, H. (2007) Substrate choice of membrane-type 1 matrix metalloproteinase is dictated by tissue inhibitor of metalloproteinase-2 levels. *Cancer Sci*, **98**, 563-568.
- Li, X., Zhang, X., Lu, Z. & Klein, B. (2001) [Relative study of soluble syndecan-1 and prognosis of patients with multiple myeloma]. *Zhonghua Xue Ye Xue Za Zhi*, **22**, 303-305.
- Maeda, N., Nishiwaki, T., Shintani, T., Hamanaka, H. & Noda, M. (1996) 6B4 proteoglycan/phosphacan, an extracellular variant of receptor-like protein-tyrosine phosphatase zeta/RPTPbeta, binds pleiotrophin/heparin-binding growth-associated molecule (HB-GAM). *J Biol Chem*, **271**, 21446-21452.
- Mahtouk, K., Cremer, F.W., Reme, T., Jourdan, M., Baudard, M., Moreaux, J., Requirand, G., Fiol, G., De Vos, J., Moos, M., Quittet, P., Goldschmidt, H., Rossi, J.-F., Hose, D. & Klein, B. (2006) Heparan sulphate proteoglycans are essential for the myeloma cell growth activity of EGF-family ligands in multiple myeloma. *Oncogene*, 25, 7180-7191.
- Mahtouk, K., Hose, D., Raynaud, P., Hundemer, M., Jourdan, M., Jourdan, E., Pantesco, V., Baudard, M., De Vos, J., Larroque, M., Moehler, T., Rossi, J.F., Reme, T.,

Goldschmidt, H. & Klein, B. (2007a) Heparanase influences expression and shedding of syndecan-1, and its expression by the bone marrow environment is a bad prognostic factor in multiple myeloma. *Blood*.

- Mahtouk, K., Hose, D., Raynaud, P., Hundemer, M., Jourdan, M., Jourdan, E., Pantesco, V., Baudard, M., De Vos, J., Larroque, M., Moehler, T., Rossi, J.F., Reme, T., Goldschmidt, H. & Klein, B. (2007b) Heparanase influences expression and shedding of syndecan-1, and its expression by the bone marrow environment is a bad prognostic factor in multiple myeloma. *Blood*, **109**, 4914-4923.
- Mahtouk, K., Hose, D., Reme, T., De Vos, J., Jourdan, M., Moreaux, J., Fiol, G., Raab, M., Jourdan, E., Grau, V., Moos, M., Goldschmidt, H., Baudard, M., Rossi, J.F., Cremer, F.W. & Klein, B. (2005) Expression of EGF-family receptors and amphiregulin in multiple myeloma. Amphiregulin is a growth factor for myeloma cells. *Oncogene*, 24, 3512-3524.
- Mitsiades, C.S., Mitsiades, N.S., Munshi, N.C., Richardson, P.G. & Anderson, K.C. (2006) The role of the bone microenvironment in the pathophysiology and therapeutic management of multiple myeloma: interplay of growth factors, their receptors and stromal interactions. *European Journal of Cancer*, **42**, 1564-1573.
- Moreaux, J., Cremer, F.W., Reme, T., Raab, M., Mahtouk, K., Kaukel, P., Pantesco, V., De Vos, J., Jourdan, E., Jauch, A., Legouffe, E., Moos, M., Fiol, G., Goldschmidt, H., Rossi, J.F., Hose, D. & Klein, B. (2005) The level of TACI gene expression in myeloma cells is associated with a signature of microenvironment dependence versus a plasmablastic signature. *Blood*, **106**, 1021-1030.
- Morimoto-Tomita, M., Uchimura, K., Werb, Z., Hemmerich, S. & Rosen, S.D. (2002) Cloning and characterization of two extracellular heparin-degrading endosulfatases in mice and humans. *J Biol Chem*, **277**, 49175-49185.
- Passaniti, A. & Hart, G.W. (1990) Metastasis-associated murine melanoma cell surface galactosyltransferase: characterization of enzyme activity and identification of the major surface substrates. *Cancer Res*, **50**, 7261-7271.
- Piao, Z., Kim, H., Jeon, B.K., Lee, W.J. & Park, C. (1997) Relationship between loss of heterozygosity of tumor suppressor genes and histologic differentiation in hepatocellular carcinoma. *Cancer*, **80**, 865-872.
- Prydz, K. & Dalen, K.T. (2000) Synthesis and sorting of proteoglycans. *J Cell Sci*, **113 Pt 2**, 193-205.
- Rapraeger, A., Jalkanen, M., Endo, E., Koda, J. & Bernfield, M. (1985) The cell surface proteoglycan from mouse mammary epithelial cells bears chondroitin sulfate and heparan sulfate glycosaminoglycans. *J Biol Chem*, **260**, 11046-11052.
- Rebouissou, C., Wijdenes, J., Autissier, P., Tarte, K., Costes, V., Liautard, J., Rossi, J.F., Brochier, J. & Klein, B. (1998) A gp130 Interleukin-6 Transducer-Dependent SCID Model of Human Multiple Myeloma. *Blood.*, **91**, 4727-4737.
- Rodeheffer, C. & Shur, B.D. (2002) Targeted mutations in beta1,4-galactosyltransferase I reveal its multiple cellular functions. *Biochim Biophys Acta*, **1573**, 258-270.
- Sato, T., Gotoh, M., Kiyohara, K., Akashima, T., Iwasaki, H., Kameyama, A., Mochizuki, H., Yada, T., Inaba, N., Togayachi, A., Kudo, T., Asada, M., Watanabe, H., Imamura, T., Kimata, K. & Narimatsu, H. (2003) Differential roles of two Nacetylgalactosaminyltransferases, CSGalNAcT-1, and a novel enzyme, CSGALNACT2. Initiation and elongation in synthesis of chondroitin sulfate. *J Biol Chem*, **278**, 3063-3071.
- Su, A.I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K.A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G., Cooke, M.P., Walker, J.R. & Hogenesch, J.B. (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci U S A*, **101**, 6062-6067.

- Subramanian, S.V., Fitzgerald, M.L. & Bernfield, M. (1997) Regulated shedding of syndecan-1 and -4 ectodomains by thrombin and growth factor receptor activation. *J Biol Chem*, **272**, 14713-14720.
- Tarte, K., De Vos, J., Thykjaer, T., Zhan, F., Fiol, G., Costes, V., Reme, T., Legouffe, E., Rossi, J.F., Shaughnessy, J., Jr., Orntoft, T.F. & Klein, B. (2002) Generation of polyclonal plasmablasts from peripheral blood B cells: a normal counterpart of malignant plasmablasts. *Blood*, **100**, 1113-1122.
- Tarte, K., Zhang, X.G., Legouffe, E., Hertog, C., Mehtali, M., Rossi, J.F. & Klein, B. (1999) Induced Expression of B7-1 on Myeloma Cells Following Retroviral Gene Transfer Results in Tumor-Specific Recognition by Cytotoxic T Cells. *J Immunol.*, 163, 514-524.
- Theocharis, A.D., Seidel, C., Borset, M., Dobra, K., Baykov, V., Labropoulou, V., Kanakis, I., Dalas, E., Karamanos, N.K., Sundan, A. & Hjerpe, A. (2006) Serglycin constitutively secreted by myeloma plasma cells is a potent inhibitor of bone mineralization in vitro. *J Biol Chem*, **281**, 35116-35128.
- Toma, L., Pinhal, M.A., Dietrich, C.P., Nader, H.B. & Hirschberg, C.B. (1996) Transport of UDP-galactose into the Golgi lumen regulates the biosynthesis of proteoglycans. J Biol Chem, 271, 3897-3901.
- Tsutsumi, S., Hogan, V., Nabi, I.R. & Raz, A. (2003a) Overexpression of the autocrine motility factor/phosphoglucose isomerase induces transformation and survival of NIH-3T3 fibroblasts. *Cancer Res*, **63**, 242-249.
- Tsutsumi, S., Yanagawa, T., Shimura, T., Fukumori, T., Hogan, V., Kuwano, H. & Raz, A. (2003b) Regulation of cell proliferation by autocrine motility factor/phosphoglucose isomerase signaling. *J Biol Chem*, **278**, 32165-32172.
- Vlodavsky, I., Friedmann, Y., Elkin, M., Aingorn, H., Atzmon, R., Ishai-Michaeli, R., Bitan, M., Pappo, O., Peretz, T., Michal, I., Spector, L. & Pecker, I. (1999) Mammalian heparanase: gene cloning, expression and function in tumor progression and metastasis. *Nat Med*, 5, 793-802.
- Wang, Y.D., De Vos, J., Jourdan, M., Couderc, G., Lu, Z.Y., Rossi, J.F. & Klein, B. (2002) Cooperation between heparin-binding EGF-like growth factor and interleukin-6 in promoting the growth of human myeloma cells. *Oncogene*, **21**, 2584-2592.
- Watanabe, H., Takehana, K., Date, M., Shinozaki, T. & Raz, A. (1996) Tumor cell autocrine motility factor is the neuroleukin/phosphohexose isomerase polypeptide. *Cancer Res*, 56, 2960-2963.
- Wei, Y., Liu, D., Zhou, F., Ge, Y., Xu, J., Yun, X., Gu, J. & Jiang, J. (2008) Identification of beta-1,4-galactosyltransferase I as a target gene of HBx-induced cell cycle progression of hepatoma cell. J Hepatol, 49, 1029-1037.
- Xu, S., Zhu, X., Zhang, S., Yin, S., Zhou, L., Chen, C. & Gu, J. (2001) Over-expression of beta-1,4-galactosyltransferase I, II, and V in human astrocytoma. *Journal of Cancer Research and Clinical Oncology*, **127**, 502-506.
- Yada, T., Sato, T., Kaseyama, H., Gotoh, M., Iwasaki, H., Kikuchi, N., Kwon, Y.D., Togayachi, A., Kudo, T., Watanabe, H., Narimatsu, H. & Kimata, K. (2003) Chondroitin sulfate synthase-3. Molecular cloning and characterization. *J Biol Chem*, 278, 39711-39725.
- Yang, Y., Macleod, V., Bendre, M., Huang, Y., Theus, A.M., Miao, H.Q., Kussie, P., Yaccoby, S., Epstein, J., Suva, L.J., Kelly, T. & Sanderson, R.D. (2005) Heparanase promotes the spontaneous metastasis of myeloma cells to bone. *Blood*, **105**, 1303-1309.
- Yang, Y., MacLeod, V., Dai, Y., Khotskaya-Sample, Y., Shriver, Z., Venkataraman, G., Sasisekharan, R., Naggi, A., Torri, G., Casu, B., Vlodavsky, I., Suva, L.J., Epstein, J., Yaccoby, S., Shaughnessy, J.D., Jr., Barlogie, B. & Sanderson, R.D. (2007a) The syndecan-1 heparan sulfate proteoglycan is a viable target for myeloma therapy. *Blood*, **110**, 2041-2048.

- Yang, Y., Macleod, V., Miao, H.Q., Theus, A., Zhan, F., Shaughnessy, J.D., Jr., Sawyer, J., Li, J.P., Zcharia, E., Vlodavsky, I. & Sanderson, R.D. (2007b) Heparanase enhances syndecan-1 shedding: a novel mechanism for stimulation of tumor growth and metastasis. *J Biol Chem*, 282, 13326-13333.
- Yin, L. (2005) Chondroitin synthase 1 is a key molecule in myeloma cell-osteoclast interactions. *J Biol Chem*, **280**, 15666-15672.
- Zak, B.M., Crawford, B.E. & Esko, J.D. (2002) Hereditary multiple exostoses and heparan sulfate polymerization. *Biochim Biophys Acta*, **1573**, 346-355.
- Zhan, F., Huang, Y., Colla, S., Stewart, J.P., Hanamura, I., Gupta, S., Epstein, J., Yaccoby, S., Sawyer, J., Burington, B., Anaissie, E., Hollmig, K., Pineda-Roman, M., Tricot, G., van Rhee, F., Walker, R., Zangari, M., Crowley, J., Barlogie, B. & Shaughnessy, J.D., Jr. (2006) The molecular classification of multiple myeloma. *Blood*, **108**, 2020-2028.
- Zhang, X.G., Gaillard, J.P., Robillard, N., Lu, Z.Y., Gu, Z.J., Jourdan, M., Boiron, J.M., Bataille, R. & Klein, B. (1994) Reproducible obtaining of human myeloma cell lines as a model for tumor stem cell study in human multiple myeloma. *Blood.*, 83, 3654-3663.
- Zhu, X., Jiang, J., Shen, H., Wang, H., Zong, H., Li, Z., Yang, Y., Niu, Z., Liu, W., Chen, X., Hu, Y. & Gu, J. (2005) Elevated beta1,4-galactosyltransferase I in highly metastatic human lung cancer cells. Identification of E1AF as important transcription activator. J Biol Chem, 280, 12503-12516.
- Zimmermann, P. & David, G. (1999) The syndecans, tuners of transmembrane signaling. *Faseb J*, **13 Suppl**, S91-S100.

Figure titles and legends

Fig 1. Pathways for HS and CS chains biosynthesis: from precursor up-take to

extracellular modifications

The first step corresponds to monosaccharide and sulphate uptake through specialized transport systems (GLUT, SLC26 and SLC13 families). These precursors are activated in UDP-sugars and PAPS in the cytosol via a set of enzymes (GALM, GALK, GALT, GALE, PGM1, PGM2, UGP2, UGDH, UXS1, GFPT, GNPDA, GNPNAT1, PGM3 and UAP1 for nucleotide sugars synthesis and PAPSS1 and 2 for sulphate activation). UDP-sugars and PAPS are actively transported into the Golgi lumen and/or RE (SLC35 family). Thereafter, glycosaminoglycan synthesis is initiated. A linkage region is produced: a xylose, two galactose and one glucuronic acid residues are added on an attachment site of the B4GALT7, B3GALT6 proteoglycan core (XYLT, and B3GAT3). Then, the glycosaminoglycan lineage depends on the addition of the fifth nucleotide sugar. Addition of a glucosaminyl residue by EXTL2 generates glycosaminoglycans (HS chains), whereas a galactosaminyl residue added by CSGALNACT1 creates galactosaminoglycans (CS chains). For these two types of chains, enzymes catalyse elongation and polymerization (EXT1/EXT2 for HS chains, CSGlcA-T, CSGALNACT2, CHSY1, CHSY3 and CHPF for CS chains). As the chains grow, they are modified at various positions: deacetylation/N-sulphatation of glucosamine by NDST, epimerisation of glucuronic acid to iduronic acid by GLCE and Osulphatation in various positions by HS2ST, HS6ST or HS3ST in HS chains and sulphatation by CHST11, 12, 13, CHST3, GalNac4S-6ST and UST in various sites in CS chains. Finally, proteoglycans bearing HS and CS chains can be transported to plasma membrane. Extracellular enzymes, as SULF1 and 2 which exhibit sulphatase activity on HS chains removing sulphate from the 6-position of glucosamine and HPSE able to cleave HS in little fragments of 10 to 20 saccharidic units, are responsible of the latest modifications.

When gene expression is associated to prognosis in our data, proteins are colour-coded in green for good genes and in red for bad genes.



Figure 1

Fig 2. A global view of variation of gene expression in MBC, PPC, BMPC and MMC using PCA

Unsupervised PCA using the list of 75 genes involved in HS and CS biosynthesis pathway was performed to obtain a global view of the variation in gene expression among MBC, PPC, BPMC and MMC.

(A). MBC, PPC and BMPC were separated in 3 distinct clusters by a PCA. The first principal component is represented by the X-axis (pc1, variance: 39%), the second principal component by the Y-axis (pc2, variance: 24%). According to pc1, PPC and BMPC clustered together on the left side and MBC clustered separately on the right side. The second component makes it possible to separate BMPC from PPC.

(B) PCA clustering of MMC together with normal MBC, PPC and BMPC delineated two groups according to first PCA component (pc1, variance: 18%): a group of MMC only with a pc1 value ≥ 0 and a group of normal cells and MMC with a pc1 value < 0.



Figure 2A



Figure 2B

Fig 3. Expression data of *CHSY3*, *SULF2*, *GPI* and *HS3ST2* in MBC, PPC, BMPC, MMC and HMCL

Histograms show the probe set signal corresponding to *CHSY3* (3A), *SULF2* (3B), *GPI* (3C) and *HS3ST2* (3D) gene expression on the Y-axis as arbitrary units determined by the Affymetrix software from 6 MBC, 7 PPC, 7 BMPC, MMC of 123 patients of the Heidelberg series and 20 HMCL. Filled histograms indicate a "present" Affymetrix call and white ones an "absent" call in the sample. When several probe sets were available for a same gene, we used the probe set with the highest variance.











Figure 3C



Figure 3D

Fig 4. Validation of Affymetrix microarray data of CHSY3, GPI, SULF2 and HS3ST2 genes

Gene expression of *CHSY3* (4A), *SULF2* (4B) and *GPI* (4C) were assayed with real-time RT-PCR in 7 HMCL samples and normalized with *B2M*. XG-11 was used as a positive control. Gene expression of *HS3ST2* (4D) was studied in 7 OST and normalized with *B2M*. One OST sample was used as a positive control. The *P* values of the correlation between microarrays and real time RT-PCR data were determined with a Spearman test and are indicated in the figure.



real-time RT-PCR expression

Figure 4A



real-time RT-PCR expression

Figure 4B



real-time RT-PCR expression

Figure 4C



Figure 4D

Fig 5. Expression of extracellular enzymes implicated in modifications of HS chains in BM environment

The histograms show the Affymetrix signals of 3 probe sets interrogating *SULF1* (5A), *SULF2* (5B) and *HPSE* (5C) expression in 5 cell populations of the BM environment: MSC, OST, CD14, PMN and CD3 cells. Filled histograms indicate a "present" Affymetrix call and white ones an "absent" call in the sample. When several probe sets were available for a same gene, we used the probe set with the highest variance.







Figure 5B



Figure 5C

Fig 6. Overall survival and event-free survival related to 4 genes in 2 independent patient series

Data are Kaplan-Meier survival curves and/or event-free survival curves of patients with the highest or lowest expression (quartile) of *B4GALT7* (6A), *CSGALNACT1* (6B), *HS2ST1* (6C) and *EXT1* (6D). The HM series comprises 171 patients and the LR-TT2 series 345 patients. *P* values are indicated in each panel.





Figure 6A



Figure 6B





Figure 6C



Figure 6D

Table I: SAM-defined overexpressed genes in MBC and in PPC plus BMPC

| | | Gene name | Function | Probe set | Fold-change | Median in MBC | Median in PPC | Median in BMPC |
|-----------------|---|-----------------------|-------------------------------|-------------|-------------|---------------|---------------|----------------|
| | | SLC2A3 | Glucose transport | 202499_s_at | 7.24 | 2738 | 474 | 146 |
| | | SLC2A3 /// SLC2A14 | Glucose transport | 222088_s_at | 5.46 | 567 | 154 | 50 |
| | SAM-defined genes overexpressed in MBC | SLC2A3 | Glucose transport | 202497_x_at | 5.29 | 863 | 357 | 28 |
| | | SLC2A3 | Glucose transport | 236571_at | 4.40 | 304 | 63 | 82 |
| | | SLC35D2 | Nucleotide sugar transport | 213082_s_at | 3.80 | 132 | 6 | 50 |
| | | SLC2A3 /// SLC2A14 | Glucose transport | 216236_s_at | 2.86 | 536.5 | 182 | 175 |
| | | UST | CS chain modification | 205139_s_at | 2.78 | 74.5 | 20 | 33 |
| | | SLC2A1 | Glucose transport | 201250_s_at | 2.42 | 482 | 212 | 191 |
| | | PGM3 | Sugar metabolism | 210041_s_at | 28.25 | 8.5 | 231 | 191 |
| | | CHST12 | CS chain modification | 218927_s_at | 17.17 | 33.5 | 405 | 1025 |
| | | CHST11 | CS chain modification | 226372_at | 15.25 | 12.5 | 350 | 207 |
| | | SLC35B3 | PAPS transport | 222691_at | 11.16 | 32.5 | 737 | 823 |
| | | GALK2 | Sugar metabolism | 205219_s_at | 6.48 | 8.5 | 162 | 53 |
| | | HS2ST1 | HS chain modification | 203284_s_at | 6.26 | 60 | 487 | 374 |
| | | UGDH | Sugar metabolism | 203343_at | 4.96 | 71.5 | 592 | 413 |
| | | UAP1 | Sugar metabolism | 209340_at | 4.94 | 270.5 | 797 | 1855 |
| | | GNPDA1 | Sugar metabolism | 202382_s_at | 4.69 | 21.5 | 133 | 182 |
| | | PGM3 | Sugar metabolism | 221788_at | 4.49 | 34.5 | 106 | 206 |
| | | GALNAC4S-6ST | CS chain modification | 203066_at | 4.31 | 292.5 | 1329 | 1499 |
| Comparison MBC | | GNPNAT1 | Sugar metabolism | 225853_at | 4.27 | 78.5 | 480 | 261 |
| versus PPC plus | SAM-defined | SLC2A5 | Glucose transport | 204430_s_at | 4.11 | 68 | 373 | 202 |
| bin c | | SLC26A2 | Sulphate anion transport | 224959_at | 4.08 | 65 | 266 | 357 |
| | | CSGlcA-T | CS chain polymerization | 55093_at | 3.98 | 77.5 | 214 | 539 |
| | | SLC35B1 | Nucleotide sugar transport | 202433_at | 3.84 | 195.5 | 944 | 773 |
| | genes overexpressed in | GFPT1 | Sugar metabolism | 202722_s_at | 3.82 | 199 | 811 | 634 |
| | PPC plus BMPC | SLC26A6 | Sulphate anion transport | 221572_s_at | 3.69 | 25 | 80 | 114 |
| | | SLC35A4 | Nucleotide sugar transport | 224626_at | 3.35 | 77 | 259 | 265 |
| | | SLC26A2 | Sulphate anion transport | 205097_at | 3.23 | 43.5 | 97 | 185 |
| | | HS2ST1 | HS chain modification | 203283_s_at | 3.16 | 54.5 | 236 | 110 |
| | | GALK1 | Sugar metabolism | 229458_s_at | 3.06 | 26.5 | 35 | 158 |
| | | CHPF | CS chain modification | 202175_at | 3.02 | 69.5 | 136 | 265 |
| | | CSGlcA-T | CS chain polymerization | 221799_at | 2.91 | 79.5 | 214 | 273 |
| | | UGP2 | Sugar metabolism | 205480_s_at | 2.86 | 444.5 | 1506 | 1426 |
| | | PGM2 | Sugar metabolism | 225366_at | 2.71 | 48.5 | 154 | 100 |
| | | GFPT1 | Sugar metabolism | 202721_s_at | 2.69 | 94.5 | 381 | 118 |
| | | PAPSS1 | PAPS synthesis | 209043_at | 2.56 | 316 | 1321 | 470 |
| | | HS2ST1 | HS chain modification | 203285_s_at | 2.49 | 72 | 206 | 153 |
| | | PGM1 | Sugar metabolism | 201968_s_at | 2.46 | 260 | 687 | 652 |
| | | EXT1 | HS chain polymerization | 201995_at | 2.20 | 82 | 152 | 217 |
| | | PGM2 | Sugar metabolism | 223738_s_at | 2.17 | 84 | 199 | 127 |
| | | CHST12 | CS chain modification | 222786_at | 2.00 | 113 | 203 | 248 |

A supervised analysis was performed with the SAM software (1000 permutations) in order to compare the gene expression profile of the MBC population and of the group PPC plus BMPC (these two populations were grouped according to PCA result shown in Figure 2A). This analysis identified probe sets related to each type of cells. Only probe sets with a 2-fold change in expression at least and a 0% false discovery rate were retained. Genes were ranked according to the fold-change.

Table II: SAM-defined overexpressed genes in PPC and BMPC

| | | Gene name | Function | Probe set | Fold- change | Median in MBC | Median in PPC | Median in BMPC |
|------------|------------------------|-----------------------|-------------------------------|-------------|-----------------|------------------|------------------|-------------------|
| | | SLC2A3 | Glucose transport | 202497_x_at | 11.79 | 863 | 357 | 28 |
| | | SLC35B4 | Nucleotide sugar transport | 238418_at | 10.67 | 3 | 30 | 3 |
| | | HK2 | Sugar metabolism | 202934_at | 8.73 | 335.5 | 716 | 66 |
| | | PGM2 | Sugar metabolism | 225367_at | 3.73 | 350.5 | 1293 | 311 |
| | | SLC2A3 | Glucose transport | 202498_s_at | 3.47 | 147.5 | 106 | 34 |
| | | GPI | Sugar metabolism | 208308_s_at | 3.31 | 393 | 1131 | 314 |
| | SAM-defined | GFPT1 | Sugar metabolism | 202721_s_at | 3.19 | 94.5 | 381 | 118 |
| | genes overexpressed | SLC2A3 /// SLC2A14 | Glucose transport | 222088_s_at | 3.00 | 567 | 154 | 50 |
| | IN PPC | PAPSS1 | PAPS synthesis | 209043_at | 2.68 | 316 | 1321 | 470 |
| | | SLC35B4 | Nucleotide sugar transport | 225881_at | 2.65 | 137 | 376 | 142 |
| | | SLC2A3 | Glucose transport | 202499_s_at | 2.58 | 2738 | 474 | 146 |
| | | SLC35D1 | Nucleotide sugar transport | 209711_at | 2.50 | 184 | 277 | 132 |
| comparison | | GALK2 | Sugar metabolism | 205219_s_at | 2.47 | 8.5 | 162 | 53 |
| PPC versus | | HS2ST1 | HS chain modification | 203283_s_at | 2.02 | 54.5 | 236 | 110 |
| BMPC | | SULF2 | HS chain modification | 224724_at | 37.00 | 81 | 6 | 383 |
| | | SLC2A10 | Glucose transport | 221024_s_at | 22.15 | 3 | 3 | 89 |
| | | PAPSS2 | PAPS synthesis | 203060_s_at | 7.03 | 22.5 | 18 | 94 |
| | | SULF2 | HS chain modification | 233555_s_at | 6.34 | 22.5 | 18 | 94 |
| | | SLC26A11 | Sulphate anion transport | 226679_at | 4.32 | 211.5 | 61 | 252 |
| | SAM-defined | HS3ST2 | HS chain modification | 219697_at | 4.19 | 100 | 54 | 188 |
| | genes | CSGALNACT1 | CS chain polymerization | 219049_at | 3.50 | 727 | 418 | 1469 |
| | in BMPC | GALK1 | Sugar metabolism | 229458_s_at | 3.43 | 26.5 | 35 | 158 |
| | | SLC35D2 | Nucleotide sugar transport | 213083_at | 3.21 | 321.5 | 71 | 227 |
| | | HS6ST1 | HS chain modification | 225263_at | 2.74 | 146 | 101 | 275 |
| | | CHST12 | CS chain modification | 218927_s_at | 2.34 | 33.5 | 405 | 1025 |
| | | CSGlcA-T | CS chain polymerization | 55093_at | 2.29 | 77.5 | 214 | 539 |
| | | UAP1 | Sugar metabolism | 209340_at | 2.22 | 270.5 | 797 | 1855 |

A supervised analysis was performed with the SAM software in order to compare the gene

expression profile of PPC and BMPC as indicated in Table I legend.

Table III: SAM-defined overexpressed genes in BMPC and in MMC

| | | Gene name | Function | Probe set | Fold-change | Median in PCC | Median in BMPC | Median in MMC |
|--------------------|--------------------------|-----------------------|-------------------------|-------------|-------------|------------------------|------------------------|------------------------|
| | | CHSY3 | CS chain polymerization | 242100_at | 10.33 | 5 | 26 | 218 |
| | | SULF2 | HS chain modification | 224724_at | 3.67 | 6 | 383 | 1042 |
| | | SULF2 | HS chain modification | 233555_s_at | 3.17 | 5 | 69 | 207 |
| | | HK2 | Sugar metabolism | 202934_at | 2.84 | 716 | 66 | 174 |
| | | CHST11 | CS chain modification | 226368_at | 2.62 | 189 | 121 | 273 |
| | SAM-defined genes | PAPSS2 | PAPS synthesis | 203058_s_at | 2.32 | 3 | 51 | 87 |
| | overexpressed in | HS2ST1 | HS chain modification | 230465_at | 2.32 | 53 (range : 8-64) | 39 (range : 31-58) | 68 (range : 1-596) |
| | MMC | PAPSS1 | PAPS synthesis | 209043_at | 2.16 | 1321 | 470 | 1006 |
| | | GPI | Sugar metabolism | 208308_s_at | 2.08 | 1131 | 314 | 633 |
| Comparison | | HPSE | HS chain modification | 219403_s_at | 2.07 | 57 (range: 11 – 85) | 40 (range: 12 – 81) | 57 (range: 1 – 895) |
| MMC versus BMPC | | EXT2 | HS chain polymerization | 202012_s_at | 2.06 | 188 | 193 | 374 |
| | | GNPDA2 | Sugar metabolism | 227022_at | 2.01 | 103 | 72 | 125 |
| | | GALK1 | Sugar metabolism | 229458_s_at | 4.23 | 35 | 158 | 26 |
| | | SLC2A3 | Glucose transport | 202499_s_at | 3.54 | 474 | 146 | 36 |
| | | SLC2A3 /// SLC2A14 | Glucose transport | 216236_s_at | 2.93 | 182 | 175 | 50 |
| | SAM-defined genes | HS3ST2 | HS chain modification | 219697_at | 2.43 | 54 | 188 | 74 |
| | overexpressed in BMPC | CSGALNACT2 | CS chain polymerization | 218871_x_at | 2.39 | 753 | 963 | 374 |
| | | CSGALNACT2 | CS chain polymerization | 222235_s_at | 2.38 | 410 | 558 | 228 |
| | | CSGALNACT1 | CS chain polymerization | 219049_at | 2.25 | 418 | 1469 | 530 |

A supervised analysis was performed with the SAM software in order to compare the gene expression profile of the BMPC and MMC populations and to underline probe sets related to each type of cells as indicated in Table I.

Table IV: SAM-defined overexpressed genes in MMC and HMCL

| | | Gene name | Function | Probe set | Fold- change | Median in MMC | Median in HMCL |
|------------|---------------|-------------------------|--------------------------------------|----------------------------|-----------------|---------------------------------|-----------------------|
| | | SULF1 | HS chain modification | 212353_at | 15.10 | 8 | 18 (range: 1 - 1303) |
| | | HS3ST3B1 | HS chain modification | 227361 at | 8 46 | 4 | 156 |
| | | SULF1 | HS chain modification | 212354 at | 6.18 | 14 | 15 (range: 1 – 693) |
| | | SLC35B4 | Nucleotide sugar | 238418_at | 5.88 | 7 | 92 |
| | | нк2 | Sugar metabolism | 202934 at | 4 68 | 174 | 827 5 |
| | | HPSE | HS chain modification | 222881 at | 3.89 | 13 | 25.5 (range: 1 – 445) |
| | | GNPDA1 | Sugar metabolism | 202382 s at | 3.22 | 169 | 475 |
| | SAM-defined | GALK2 | Sugar metabolism | 205219 s at | 2.81 | 84 | 211 |
| | genes | GALE | Sugar metabolism | 202528 at | 2.78 | 54 | 129.5 |
| | overexpressed | GALM | Sugar metabolism | | 2.70 | 65 | 206.5 |
| | in HMCL | HPSE | HS chain modification | 219403_s_at | 2.55 | 57 | 127.5 |
| | | GNPNAT1 | Sugar metabolism | 225853_at | 2.48 | 242 | 577 |
| | | GFPT1 | Sugar metabolism | 202721_s_at | 2.42 | 167 | 399.5 |
| | | SLC35B4 | Nucleotide sugar transport | 225881_at | 2.42 | 219 | 581.5 |
| | | B3GAT3 | Synthesis of linkage region | 203452_at | 2.18 | 88 | 172.5 |
| | | GPI | Sugar metabolism | 208308_s_at | 2.12 | 633 | 1465 |
| | | PGM2 | Sugar metabolism | 225367_at | 2.07 | 482 | 1140 |
| | | SLC35D2 | Nucleotide sugar | 213083 at | 27.43 | 176.5 | 3 |
| | | SLC35D2 | transport Nucleotide sugar | 213082_s_at | 21.87 | 53.5 (range: 2 - 693) | 3 |
| | | CSCALNACT1 | transport CS abain nolymonization | 210040 at | 11.72 | 521 | 20 |
| | | GALKI | Sugar metabolism | 219049_at | 6.44 | 26 5 (range: 1 - 181) | 30 |
| comparison | | CHSV3 | CS chain polymerization | 229438_s_at 242100 at | 6.37 | 20.5 (Talige: 1 - 181) 214 5 | 25 |
| MMC versus | | SULF2 | HS chain modification | 233555 s at | 6.07 | 214.5 | 30 |
| HMCL | | SUC2A10 | Glucose transport | 2233555_s_at | 5.95 | 45 5 (range: 1 - 686) | 3 |
| | | SULF2 | HS chain modification | 224724 at | 5.88 | 1034 | 146 |
| | | HS3ST2 | HS chain modification | 219697 at | 4.90 | 74.5 | 17 |
| | | SLC2A9 | Glucose transport | 219991 at | 4.61 | 101 | 23 |
| | | EXT1 | HS chain polymerization | | 4.54 | 69.5 | 15 |
| | | CHST11 | CS chain modification | 226368_at | 3.93 | 270 | 86 |
| | | HS3ST4 | HS chain modification | 228206_at | 3.87 | 79.5 | 19 |
| | CAM J.C. | EXT1 | HS chain polymerization | 239227_at | 3.60 | 63.5 | 16 |
| | SAM-defined | NDST1 | HS chain modification | 202607_at | 3.52 | 49.5 (range: 6 - 252) | 18 |
| | overexpressed | PAPSS2 | PAPS synthesis | 203060_s_at | 3.36 | 105.5 | 14 |
| | in MMC | EXT1 | HS chain polymerization | 237310_at | 3.35 | 46.5 (range: 7 - 222) | 15 |
| | | HS2ST1 | HS chain modification | 230465_at | 3.25 | 64.5 | 25 |
| | | PAPSS2 | PAPS synthesis | 203058_s_at | 3.21 | 89 | 12 |
| | | EXT1 | HS chain polymerization | 232174_at | 3.13 | 110 | 35 |
| | | PAPSS2 | PAPS synthesis | 203059_s_at | 2.92 | 63.5 | 19 |
| | | SLC35D2 | Nucleotide sugar transport | 231437_at | 2.55 | 41.5 (range: 4 - 121) | 15 |
| | | SLC26A1 | Sulphate anion transport | 205058_at | 2.36 | 35 (range: 4 - 104) | 15 |
| | | CHST11 | CS chain modification | 219634_at | 2.33 | 163 | 84 |
| | | SLC2A1 | Glucose transport | 201250_s_at | 2.33 | 189 | 83 |
| | | SLC2A13 /// C12ORF40 | Glucose transport | 227176_at | 2.29 | 102.5 | 54 |
| | | PGM3 | Sugar metabolism | 221788_at | 2.18 | 182.5 | 60 |
| | | CHPF | CS chain polymerization | 202175_at | 2.10 | 181.5 | 102 |
| | | SLC35B3 | PAPS transport | 231003_at | 2.05 | /0 | 32 |
| | | GALNAC4S-6ST | CS chain modification | 244874_at | 2.01 | 68 | 30 |

A supervised analysis was performed with the SAM software in order to compare the gene expression profile of the MMC and the HMCL and to determine probe sets linked to each type of cell as indicated in Table I.

| Gene Probe set Good/bad gene | | B4GALT7 53076_at good | CSGALNACT1 219049_at good | HS2ST1 203285_s_at good | <i>EXT1</i> 239227_at <i>bad</i> |
|---------------------------------------|---------|-----------------------------|---------------------------------|-------------------------------|--|
| Heidelberg-Montpellier series (n=171 | EFS (P) | 0.00070 | 0.02080 | 0.03650 | NS |
| rieuenberg-wiontpenier series (n=171) | OAS (P) | NS | 0.03400 | NS | 0.00940 |
| Little Rock series | EFS (P) | 0.00888 | 0.00230 | 0.03050 | NS |
| (n=345) | OAS (P) | 0.00280 | 0.00160 | NS | 0.03230 |

prognosis in HM and LR-TT2 series of patients.

Data are the *P* values of the 4 probe sets - *B4GALT7*, *CSGALNACT1*, *HS2ST1* and *EXT1* - having a significant EFS and/or OAS in both HM and LR-TT2 series. When different probe sets interrogating the same gene were associated to prognosis, the probe set with the highest variance was selected. "NS" (not significant") indicates that Kaplan-Meier analysis is not significant.

| | | Patients subgroup | | | | | | | | |
|-------------|---|-------------------|--|------------------------------------|---|-------|---|------------------------------------|---|-------|
| | | All patients | PR | LB | MS | HY | CD1 | CD2 | MF | MY |
| Probe set | | 100% | 8.4% | 9.0% | 12.1% | 18.8% | 6.3% | 11.9% | 6.0% | 27.5% |
| 53076_at | frequency of patients with B4GALT7 ^{high} | 25% | 3.4% (P = 10 ⁻⁵) | 9.7% | 7.1% (<i>P</i> = 5.10 ⁻⁴) | 32.3% | 45.5% (<i>P</i> = 2.10 ⁻³) | 36.6% | 15.0% | 31.6% |
| 219049_at | frequency of patients with CSGALNACT1 ^{high} | 25% | 3.4% (<i>P</i> = 10 ⁻⁵) | 45.2% (P = 2.10 ⁻³) | 19.0% | 12.3% | 45.5% (P = 2.10 ⁻³) | 65.9% (P = 6.10°) | 30.0% | 12.6% |
| 203085_s_at | frequency of patients with HS2ST1 ^{high} | 25% | 3.4% (P = 10 ⁻⁵) | 32.3% | 7.1% (<i>P</i> = 5.10 ⁻⁴) | 27.7% | 31.8% | 78.0% (P = 6.10 ⁻⁴) | 0.0% (P =9.10 ⁻⁸) | 15.8% |
| 239227_at | frequency of patients with EXT1 ^{high} | 25% | 27.6% | 22.6% | 45.2% (P = 2.10 ⁻³) | 15.4% | 27.3% | 22.0% | 15.0% | 23.2% |

molecular classification of multiple myeloma

For each probe set, patients were classified in 4 quartiles according to probe set signal and we examined the frequency of patients belonging to the highest quartile in the 8 molecular subgroups defined by Zhan et al(Zhan, *et al* 2006): PR: proliferation, LB: low bone disease, MS: MMSET, HY: hyperdiploid, CD1 and CD2: CCND1/CCND3, MF: MAF/MAFB and MY: myeloid group. Data are shown in bold and italic when the percentage of patients with *B4GALT7^{high}*, *CSGALNACT1* ^{high}, *HS2ST1* ^{high} or *EXT1* ^{high} MMC in a given group was significantly different from that in all patients with a Chisquare test ($P \le 0.05$).

Supporting information

Table SI. Genes linked to HS and CS synthesis and modification

| | Gene name | Gene description | Probe set | References |
|-------------------------|-----------|---|-------------------------------------|--------------------------|
| | SLC2A1 | solute carrier family 2 (facilitated glucose transporter); member 1 (GLUT1) | 201250_s_at | (Wood and Trayhurn 2003) |
| | SLC2A2 | solute carrier family 2 (facilitated glucose transporter); member 2 (GLUT2) | absent or weak Affymetrix signal | " |
| | | | 202499_s_at | " |
| | | | 202497_x_at | " |
| | SI C242 | solute carrier family 2 (facilitated glucose transporter); | 216236_s_at | " |
| | SLC2A5 | member 3 (GLUT3) | 222088_s_at | " |
| | | | 236571_at | " |
| | | | 202498_s_at | " |
| | SLC2A4 | solute carrier family 2 (facilitated glucose transporter); member 4 (GLUT4) | absent or weak Affymetrix signal | " |
| | SLC2A5 | solute carrier family 2 (facilitated glucose/fructose | 204430_s_at | " |
| Glucose transporters | | transporter); member 5 (GLUT5) | 204429_s_at | " |
| | SLC2A6 | solute carrier family 2 (facilitated glucose/fructose transporter); member 6 (GLUT6) | 220091_at | " |
| | SLC2A7 | solute carrier family 2 (facilitated glucose/fructose transporter); member 7 (GLUT7) | no probe set available | " |
| | SLC2A8 | solute carrier family 2 (facilitated glucose/fructose transporter); member 8 (GLUT8) | 218895_at | " |
| | SLC2A9 | solute carrier family 2 (facilitated glucose transporter); member 9 (GLUT9) | 219991_at | " |
| | SLC2A10 | solute carrier family 2 (facilitated glucose transporter); member 10 (GLUT10) | 221024_s_at | " |
| | SLC2A11 | solute carrier family 2 (facilitated glucose transporter); member 11 (GLUT11) | 232167_at | " |
| | SLC2A12 | solute carrier family 2 (facilitated glucose transporter); member 12 (GLUT12) | absent or weak Affymetrix signal | " |
| | SLC2A13 | solute carrier family 2 (facilitated glucose transporter); member 13 (H(+)-myo-inositol symporter, HMIT) | 227176_at | " |

| | Gene name | Gene description | Probe set | References |
|---------------------------------|-----------|---|-------------------------------------|--|
| | HK1 | hexokinase 1 | 200697_at | (Wilson 2003) |
| | HK2 | hexokinase 2 | 202934_at | " |
| | НКЗ | hexokinase 3 | absent or weak Affymetrix signal | " |
| | PGM1 | phosphoglucomutase 1 | 201968_s_at | (Drago, et al 1991) |
| | | | 225367_at | " |
| | PGM2 | phosphoglucomutase 2 | 223738_s_at | " |
| | | | 225366_at | " |
| | PCM2 | phosphoglucomutese 2 | 221788_at | " |
| | rGMS | pnospnoglucomulase 5 | 210041_s_at | " |
| | PGM5 | phosphoglucomutase 5 | absent or weak Affymetrix signal | (Edwards, et al 1995) |
| | UGP1 | UDP-glucose pyrophosphorylase 1 | no probe set available | (Bulter and Elling 1999) |
| | UGP2 | UDP-glucose pyrophosphorylase 2 | 205480_s_at | " |
| | UGDH | UDP-glucose dehydrogenase | 203343_at | " |
| | 11981 | LIDD aluqueonata dagarhovylasa 1 | 225583_at | " |
| Sugar | UASI | | 219675_s_at | " |
| metabolism, nucleotide sugar | GALE | UDP-galactose-4-epimerase | 202528_at | (Novelli and Reichardt 2000) |
| synthesis | GALK1 | galactokinase 1 | 229458_s_at | |
| | GALK2 | galactokinase 2 | 205219_s_at | " |
| | CALT | calastasa 1 nkonhata uridulultransfarasa | 203179_at | " |
| | UAL1 | galaciose-1-phosphale unityryntansierase | 232708_at | " |
| | CAIM | colocion mutarotaca (aldaca 1 animaraca) | 235256_s_at | (Timson and Reece 2003) |
| | GALM | galaciose mutarotase (algose 1-epimerase) | 234974_at | " |
| | UAP1 | UDP-N-acetylglucosamine pyrophosphorylase 1 | 209340_at | (Szumilo, et al 1996) |
| | GPI | glucose phosphate isomerase | 208308_s_at | (Achari, et al 1981) |
| | GFPT1 | glutamine-fructose-6-phosphate transaminase 1 | 202722_s_at | (Ghosh, et al 1960), (Whitmore, et al 1995) |
| | | | 202721_s_at | " |
| | GFPT2 | glutamine-fructose-6-phosphate transaminase 2 | absent or weak Affymetrix signal | (Ghosh, et al 1960), (Oki, et al 1999) |
| | GNPDA1 | glucosamine-6-phosphate deaminase 1 | 202382_s_at | (Shevchenko, et al 1998) |
| | GNPDA2 | glucosamine_6-nhosnhate deaminase 2 | 227022_at | " |
| | GITIDIL | giteosanine-o-piospilate deaninase 2 | 230149_at | " |
| | GNPNAT1 | glucosamine-phosphate N-acetyltransferase 1 | 225853_at | (Davidson, et al 1957) |

| | Gene name | Gene description | Probe set | References |
|----------------------|-----------|--|-------------|----------------------------|
| | SLC35A2 | solute carrier family 35 (UDP-galactose transporter); member | 209326_at | (Ishida and Kawakita 2004) |
| | ~~~~~ | A2 | 207439_s_at | " |
| | \$1,03543 | solute carrier family 35 (UDP-N-acetylglucosamine (UDP- | 226894_at | " |
| | SECISAS | GlcNAc) transporter); member A3 | 206770_s_at | " |
| | \$1,C35D1 | solute carrier family 35 (UDP-glucuronic acid/UDP-N- | 209711_at | " |
| | 5205501 | acetylgalactosamine dual transporter); member D1 | 209712_at | " |
| Nucleotide sugars | SLC35A4 | solute carrier family 35; member A4 (UDP-galactose transporter) | 224626_at | " |
| transporters | | | 225881_at | " |
| | SLC35B4 | solute carrier family 35; member B4 (UDP-Xylose/N- acetylglucosamine transporter) | 225882_at | " |
| | | | 238418_at | " |
| | SLC35B1 | solute carrier family 35; member B1 (UDP-galactose transporter related) | 202433_at | " |
| | | | 213083_at | " |
| | SLC35D2 | solute carrier family 35; member D2 (UDP-N- acetylglucosamine transporter) | 231437_at | " |
| | | | 213082_s_at | " |

| | Gene name | Gene description | Probe set | References |
|----------------|-----------|--|-------------------------------------|-------------------------------|
| | SLC13A1 | solute carrier family 13 (sodium/sulphate symporter); member 1 (NaSi-1) | absent or weak Affymetrix signal | (Markovich and Murer 2004) |
| | SIC13AA | solute carrier family 13 (sodium/sulphate symporter); | 224758_at | " |
| | SLC15A4 | member 4 (SUT-1) | 238273_at | " |
| | SLC26A1 | solute carrier family 26 (sulphate transporter); member 1 (SAT-1) | 205058_at | (Markovich 2001) |
| | SLC26A2 | | 224959_at | " |
| Sulphoto onion | | solute carrier family 26 (sulphate transporter); member 2 (DTDST) | 205097_at | " |
| | | | 224963_at | " |
| | SLC26A3 | solute carrier family 26; member 3 (DRA/CLD) | absent or weak Affymetrix signal | |
| transporters | SLC26A4 | solute carrier family 26; member 4 (PENDRIN) | absent or weak Affymetrix signal | |
| | SLC26A5 | solute carrier family 26; member 5 (PRESTIN) | 238020_at | " |
| | SLC26A6 | solute carrier family 26; member 6 (PAT-1) | 221572_s_at | " |
| | SLC26A7 | solute carrier family 26; member 7 | 238214_at | " |
| | SLC26A8 | solute carrier family 26; member 8 | absent or weak Affymetrix signal | |
| | SLC26A9 | solute carrier family 26; member 9 | absent or weak Affymetrix signal | |
| | SLC26A10 | solute carrier family 26; member 10 | absent or weak Affymetrix signal | |
| | SLC26A11 | solute carrier family 26; member 11 | 226679_at | " |

| | Gene name | Gene description | Probe set | References |
|----------|-----------|---|-------------|------------------|
| | PAPSS1 | 3'-phosphoadenosine 5'-phosphosulphate synthase 1 | 209043_at | (Xu, et al 2000) |
| Sulphate | | | 203058_s_at | " |
| PAPS | PAPSS2 | 3'-phosphoadenosine 5'-phosphosulphate synthase 2 | 203059_s_at | " |
| | | | 203060_s_at | " |

| | Gene name | Gene description | Probe set | References |
|----------------------|-------------|---|-----------|------------------------|
| | SLC35B2 | solute carrier family 35; member B2 (PAPST1) | 224716_at | (Kamiyama, et al 2003) |
| PAPS transporters | SI C 25 P 2 | solute corrier family 25: member P2 (DADST2) | 222691_at | (Kamiyama, et al 2006) |
| | SLC55B5 | solute carrier failing 55, member B5 (FAFS12) | 231003_at | " |

| | Gene name | Gene description | Probe set | References |
|----------------|------------------|---|----------------------------------|---|
| Sunthesis of | XYLT1 | xylosyltransferase I | absent or weak Affymetrix signal | (Kamiyama, et al 2006) |
| | XYLT2 | xylosyltransferase II | 231550_at | (Gotting, <i>et al</i> 2000), (Cuellar, <i>et al</i> 2007) |
| | B4GALT7 | xylosylprotein beta 1;4-galactosyltransferase; polypeptide 7 | 53076_at | (Furukawa and Okajima 2002) |
| linkage region | | (galactosyltransferase 1) | 222191_s_at | " |
| | B3GALT6 | UDP-Gal:betaGal beta 1;3-galactosyltransferase polypeptide 6 (putative galactosyltransferase II) | 225733_at | (Bai, et al 2001) |
| | B3GAT3 beta-1;3- | hata 1.2 aluguranyiltransfarasa 2 (alugurangayiltransfarasa I) | 35179_at | (Pedersen, et al 2000) |
| | | beta-1,5-gluculonymansterase 5 (gluculonosymansterase 1) | 203452_at | " |

| Gene name | | Gene description | Probe set | References |
|----------------|-----------------------------|-----------------------------------|----------------------------------|------------------------|
| | EXTL2 | exostoses (multiple)-like 2 | 209537_at | (Kitagawa, et al 1999) |
| | EXTL1 | exostoses (multiple)-like 1 | absent or weak Affymetrix signal | (Kim, et al 2001) |
| | EXTL3 | EXTL3 exostoses (multiple)-like 3 | | " |
| | EXT1 exostoses (multiple) 1 | 201995_at | (Zak, et al 2002) | |
| | | exostoses (multiple) 1 | 215206_at | " |
| HS chain | | | 232174_at | " |
| polymerization | | | 234634_at | " |
| | | | 237310_at | " |
| | | | 239227_at | " |
| | | | 242126_at | " |
| | EVTY | avastasas (multipla) 2 | 202012_s_at | " |
| | EAT2 | exosioses (inutiple) 2 | 202013_s_at | " |

| | Gene name | Gene description | Probe set | References |
|----------------|------------|---|-------------|---|
| | CSGALNACT1 | GGALNACT1 chondroitin beta1;4 N-acetylgalactosaminyltransferase 1 | | (Sato, et al 2003), (Prabhakar and Sasisekharan 2006) |
| | CSCI-A T | chondroitin sulphate glucuronyltransferase | 55093_at | (Gotoh, et al 2002) |
| | 65011-1 | enondrottin surpliate glueuronyaransierase | 221799_at | " |
| CS chain | CSGALNACT2 | chondroitin sulphote GalNAcT 2 | 218871_x_at | (Sato, et al 2003) |
| porymerization | | cholulotuli sulphate GalivAc1-2 | 222235_s_at | " |
| | CHSY1 | carbohydrate (chondroitin) synthase 1 | 203044_at | (Kitagawa, et al 2001) |
| | CHPF | chondroitin polymerizing factor | 202175_at | (Kitagawa, et al 2003) |
| | CHSY3 | chondroitin sulphate synthase 3 (chondroitin synthase-2) | 242100_at | (Yada, et al 2003), (Izumikawa, et al 2007) |

| | Gene name | Gene description | Probe set | References |
|---------------|-----------|--|-------------------------------------|---|
| | NDST1 | N-deacetylase/N-sulphotransferase (heparan glucosaminyl) 1 | 202607_at | (Aikawa, et al 2001) |
| | NDST2 | N-deacetylase/N-sulphotransferase (heparan glucosaminyl) 2 | 203916_at | " |
| | NDST3 | N-deacetylase/N-sulphotransferase (heparan glucosaminyl) 3 | absent or weak Affymetrix signal | |
| | NDST4 | N-deacetylase/N-sulphotransferase (heparan glucosaminyl) 4 | absent or weak Affymetrix signal | " |
| | GLCE | glucuronyl C5-epimerase | 213552_at | (Crawford, et al 2001) |
| | | | 203285_s_at | (Xu, et al 2007) |
| | US2ST1 | hanaran sulahata 2.0 sulahatranafarasa l | 203283_s_at | " |
| | 1152511 | neparan surpliate 2-0-surplion ansierase 1 | 203284_s_at | " |
| | | | 230465_at | " |
| | HS6ST1 | heparan sulphate 6-O-sulphotransferase 1 | 225263_at | (Habuchi, et al 1998), (Esko and Selleck 2002) |
| | HS6ST2 | heparan sulphate 6-O-sulphotransferase 2 | absent or weak Affymetrix signal | " |
| HS chain | HS6ST3 | heparan sulphate 6-O-sulphotransferase 3 | 232276_at | " |
| modifications | HS3ST1 | heparan sulphate (glucosamine) 3-O-sulphotransferase 1 | absent or weak Affymetrix signal | (Esko and Selleck 2002) |
| | HS3ST2 | heparan sulphate (glucosamine) 3-O-sulphotransferase 2 | 219697_at | " |
| | HS3ST3A1 | heparan sulphate (glucosamine) 3-O-sulphotransferase 3A1 | absent or weak Affymetrix signal | " |
| | HS3ST3B1 | heparan sulphate (glucosamine) 3-O-sulphotransferase 3B1 | 227361_at | " |
| | HS3ST4 | heparan sulphate (glucosamine) 3-O-sulphotransferase 4 | 228206_at | " |
| | HS3ST5 | heparan sulphate (glucosamine) 3-O-sulphotransferase 4 | absent or weak Affymetrix signal | (Chen, et al 2003) |
| | SULF1 | sulphatase l | 212353_at | (Morimoto-Tomita, et al 2002) |
| | | | 212354_at | |
| | SULF2 | subhatase 2 | 224724_at | " |
| | SULL'2 | sulplatase 2 | 233555_s_at | " |
| | UDSE | Цапологоо | 219403_s_at | (Vlodavsky, et al 1999) |
| | пгзе | reparanase | 222881_at | " |

| | Gene name | Gene description | Probe set | References | |
|---------------|-----------|---|----------------------------------|-----------------------|--|
| | | | 226368_at | (Hiraoka, et al 2000) | |
| | CHST11 | carbohydrate (chondroitin 4) sulphotransferase 11 | 219634_at | " | |
| | | | 226372_at | " | |
| | CHST12 | carbohydrate (chondroitin 4) sulphotransferase 12 | 218927_s_at | " | |
| CS chain | | | 222786_at | " | |
| modifications | CHST13 | carbohydrate (chondroitin 4) sulphotransferase 13 | absent or weak Affymetrix signal | (Kang, et al 2002) | |
| | CHST3 | carbohydrate (chondroitin 6) sulphotransferase 3 | absent or weak Affymetrix signal | (Fukuta, et al 1998) | |
| | GalNAc4S- | B cell PAG associated protein | 203066_at | (Ohtake, et al 2001) | |
| | 6ST | b cen KAO associated protein | 244874_at | " | |
| | UST | uronyl-2-sulphotransferase (2OST) | 205139_s_at | (Ohtake, et al 2005) | |

These tables list the 100 genes linked to HS and CS synthesis and modification. The expression of 75 of these 100 genes might be evaluated in B and plasma cells using 123 probe sets of Affymetrix U133 Plus 2:0 microarrays. 25 genes could not be studied owing to a lack

of probe set (2 genes) or to a defective probe set characterized by low (< 100) or absent signal

in our samples and on a large panel of human tissues (Su, et al 2004).

Table SII: SAM-defined overexpressed genes in PPC and in PPC plus BMPC versus MMC

| | | Gene name | Function | Probe set | Fold-change | Median in PCC | Median in BMPC | Median in MMC |
|--------------------------|---|--------------------|-------------------------------|-------------|-------------|---------------|----------------|-------------------------|
| | | SULF2 | HS chain modification | 224724_at | 135.83 | 6 | 383 | 1042 |
| | | PAPSS2 | PAPS synthesis | 203058_s_at | 35.86 | 3 | 51 | 87 |
| | | SLC2A10 | Glucose transport | 221024_s_at | 25.37 | 3 | 89 | 43 (range : 1-686) |
| | | CHSY3 | CS chain polymerization | 242100_at | 20.19 | 5 | 26 | 218 |
| | | SULF2 | HS chain modification | 233555_s_at | 20.09 | 5 | 69 | 207 |
| | | PAPSS2 | PAPS synthesis | 203060_s_at | 12.48 | 18 | 94 | 105 |
| | SAM-defined genes | SLC35D2 | Nucleotide sugar transport | 213082_s_at | 5.01 | 6 | 50 | 54 (range : 2-693) |
| | MMC | SLC26A11 | Sulphate anion transport | 226679_at | 3.82 | 61 | 252 | 199 (range : 40-887) |
| | | PAPSS2 | PAPS synthesis | 203059_s_at | 2.83 | 24 | 72 | 63 (range : 2-179) |
| | | UAP1 | Sugar metabolism | 209340_at | 2.76 | 797 | 1855 | 2184 |
| Comparison BBC versus | | HS3ST4 | HS chain modification | 228206_at | 2.65 | 28 | 69 | 79 |
| MMC | | EXTL2 | HS chain polymerization | 209537_at | 2.45 | 144 | 181 | 298 |
| | | EXT2 | HS chain polymerization | 202012_s_at | 2.32 | 188 | 193 | 374 |
| | SAM-defined genes overexpressed in PPC | SLC2A3 | Glucose transport | 202497_x_at | 11.31 | 357 | 28 | 14 |
| | | SLC2A3 | Glucose transport | 202499_s_at | 9.13 | 474 | 146 | 36 |
| | | SLC2A3 | Glucose transport | 202498_s_at | 6.43 | 106 | 34 | 15 |
| | | SLC2A3 /// SLC2A14 | Glucose transport | 222088_s_at | 5.96 | 154 | 50 | 28 |
| | | SLC2A6 | Glucose transport | 220091_at | 3.65 | 106 | 27 | 19 |
| | | SLC2A3 /// SLC2A14 | Glucose transport | 216236_s_at | 3.09 | 182 | 175 | 50 |
| | | HK2 | Sugar metabolism | 202934_at | 3.07 | 716 | 66 | 174 |
| | | PGM2 | Sugar metabolism | 225367_at | 2.41 | 1293 | 311 | 482 |
| | | CHSY3 | CS chain polymerization | 242100_at | 13.67 | 5 | 26 | 218 |
| | | SULF2 | HS chain modification | 224724_at | 7.15 | 6 | 383 | 1042 |
| | | SULF2 | HS chain modification | 233555_s_at | 5.47 | 5 | 69 | 207 |
| | | PAPSS2 | PAPS synthesis | 203058_s_at | 4.36 | 3 | 51 | 87 |
| | SAM-defined genes | PAPSS2 | PAPS synthesis | 203060_s_at | 3.11 | 18 | 94 | 105 |
| | overexpressed in MMC | SLC35D2 | Nucleotide sugar transport | 213082_s_at | 2.85 | 6 | 50 | 54 (range : 2-693) |
| Comparison PPC plus | | HS2ST1 | HS chain modification | 230465_at | 2.34 | 53 | 39 | 68 (range : 1-596) |
| BMPC versus MMC | | SLC2A10 | Glucose transport | 221024_s_at | 2.19 | 3 | 89 | 43 (range : 1-686) |
| | | EXT2 | HS chain polymerization | 202012_s_at | 2.18 | 188 | 193 | 374 |
| | | EXTL2 | HS chain polymerization | 209537_at | 2.08 | 144 | 181 | 298 |
| | | SLC2A3 | Glucose transport | 202497_x_at | 6.14 | 357 | 28 | 14 |
| | SAM-defined genes | SLC2A3 | Glucose transport | 202498_s_at | 4.14 | 106 | 34 | 15 |
| | plus BMPC | SLC2A3 /// SLC2A14 | Glucose transport | 216236_s_at | 3.01 | 182 | 175 | 50 |
| | | GALK1 | Sugar metabolism | 229458_s_at | 2.73 | 35 | 158 | 26 |

A supervised analysis was performed with the SAM software (1000 permutations) in order to compare the gene expression profile of the PPC population and of the group PPC plus BMPC

versus MMC. This analysis identified probe sets related to each type of cells. Only probe sets with a 2-fold change in expression at least and a 0% false discovery rate were retained. Genes were ranked according to the fold-change.

Table SIII. Genes encoding for chondroitin sulphate proteoglycans and theircorresponding probe sets

| Gene name | Gene description | Probe set |
|-----------|-------------------------------------|---|
| | | 205679_x_at |
| AGC1 | aggrecan 1 | 207692_s_at |
| | | 217161_x_at |
| | | 212283_at |
| AGRN | agrin | 212285_s_at |
| | | 217419_x_at |
| | | 208248_x_at |
| | | 208702_x_at |
| 4.01.02 | omulaid measurean like meatain 2 | 208703_s_at |
| AF LF 2 | anyou precusor like protein 2 | 208704_x_at |
| | | 211404_s_at |
| | | 228520_s_at |
| | | 200602_at |
| APP | amyloid beta (A4) precursor protein | 211277_x_at |
| | | 214953_s_at |
| | | 219107_at |
| | | 221623_at |
| BCAN | brevican | 223632_s_at |
| | | 223633_s_at |
| | | 91920_at |
| | | 201261_x_at |
| BGN | biglycan | 201262_s_at |
| | | 213905_x_at |
| | | 204489_s_at |
| | | 204490_s_at |
| | | 209835_x_at |
| | | 210916_s_at |
| | | 212014_x_at |
| CD44 | CD44 antigen | 212063_at |
| | | 216056_at |
| | | 216062_at |
| | | 217523_at |
| | | 229221_at |
| | | 234411_x_at |
| COL12A1 | collagen XII | Weak Affymetrix signal (always below 100 in our samples and in Hogenesh data base) |
| COL14A1 | collagen XIV | Weak Affymetrix signal (always below 100 in our samples and in Hogenesh data base) |
| COL15A1 | collagen XV | 203477_at |
| COL9A1 | collagen IX alpha l | Weak Affymetrix signal (always below 100 in our samples and in Hogenesh data base) |
| COL9A2 | collagen IX alpha 2 | 213622_at |
| COL9A3 | collagen IX alpha 3 | 204724_s_at |
| VCAN | versican | 204619_s_at |
| | | 204620_s_at |
| | | 211571_s_at |
| | | 215646_s_at |

| | | 221731_x_at | | |
|-------------------------|--|---|--|--|
| CSDC2 | | 205143_at | | |
| CSPG5 | neurocan | 205143_at | | |
| CSDC4 | chondroitin sulphate proteoglycan 4 (melanoma | 204736_s_at | | |
| CSPG4 | associated) | 214297_at | | |
| CEPC5 | | 205344_at | | |
| CSPGS | neurogiycan C | 39966_at | | |
| | | 209257_s_at | | |
| CSPG6 | bamacan | 209258_s_at | | |
| | | 209259_s_at | | |
| | | 201893_x_at | | |
| DCN | d e contra | 209335_at | | |
| DCN | decorin | 211813_x_at | | |
| | | 211896_s_at | | |
| ESM1 | endothelial cell-specific molecule 1 | Weak Affymetrix signal (always below 100 in our samples and in Hogenesh data base) | | |
| | | 201654_s_at | | |
| HSPG2 | periecan | 201655_s_at | | |
| LOC84664 /// CSPG4LY | chondroitin sulphate proteoglycan 4 like | Weak Affymetrix signal (always below 100 in our samples and in Hogenesh data base) | | |
| LUM | lumican | 201744_s_at | | |
| SPCN | | 201858_s_at | | |
| SKGN | proteogrycan i secretory granule | 201859_at | | |
| PTPRZ1 | protein tyrosine phosphatase; receptor-type Z polypeptide 1 | 204469_at | | |
| SPOCK | testican | 202363_at | | |
| CDOCK2 | tertions 2 | 202523_s_at | | |
| SPOCK2 | testican 2 | 202524_s_at | | |
| SPOCK3 | testican 3 | 235342_at | | |
| | | 204731_at | | |
| TGFBR3 | betaglycan | 226625_at | | |
| | | 242145_at | | |
| | | 203887_s_at | | |
| THBD | thrombomodulin | 203888_at | | |
| | | 237252_at | | |
| TNR | tenascin R | 206990_at | | |

This table lists the genes encoding for chondroitin sulphate proteoglycans (CSPG) (Oohira, *et al* 2000), (Wegrowski and Maquart 2006), (Probstmeier, *et al* 2000) examined in this study and their corresponding probe sets in Affymetrix U133 Plus 2:0 microarrays. As previously indicated, we did not consider data related to probe sets characterized by low (< 100) or absent signal in our samples and on a large panel of human tissues (Hogenesh data base) (Su, *et al* 2004).

| | | Gene name | Probe set | Fold-change | Median in MBC | Median in PPC | Median in BMPC |
|--|---|-----------|-------------|-------------|---------------|---------------|----------------|
| | SAM-defined genes | SPOCK2 | 202524_at | 9.73 | 972 | 59 | 129 |
| | overexpressed in | SPOCK2 | 202523_s_at | 3.52 | 177 | 59 | 43 |
| Comparison MBC versus PPC plus BMPC | мвс | CSPG6 | 209258_s_at | 3.36 | 143 | 24 | 40 |
| | SAM-defined genes overexpressed in PPC | APLP2 | 208702_x_at | 2.68 | 177 | 328 | 612 |
| | | APLP2 | 211404_s_at | 2.20 | 168 | 237 | 534 |
| | plus Bivir C | SRGN | 201858_s_at | 2.11 | 452 | 875 | 1210 |

Table SIV. SAM-defined overexpressed CSPG genes in MBC and in PPC plus BMPC

A supervised analysis was performed with the SAM software using the list of genes encoding for CSPG indicated in Table SIII. The aim was to compare the gene expression profile of the MBC samples and the PPC plus BMPC samples and to determine probe sets linked to each type of cell. Only probe sets with a 2-fold change in expression at least and a 0% false discovery rate with 1000 permutations were retained. Genes were ranked according to the fold-change.

Table SV. SAM-defined overexpressed CSPG genes in BMPC and MMC

| _ | | Gene name | Probe set | Fold-change | Median in BMPC | Median in MMC |
|----------------------------------|--|-----------|-------------|-------------|----------------|---------------|
| | | VCAN | 204619_s_at | 4.63 | 141 | 25 |
| | SAM-defined genes | VCAN | 221731_x_at | 3.17 | 281 | 60 |
| Comparison BMPC versus MMC | overexpressed in BMPC | VCAN | 204620_s_at | 2.47 | 271 | 74 |
| | | COL9A2 | 213622_at | 2.26 | 239 | 82 |
| | | APP | 211277_x_at | 2.05 | 118 | 66 |
| | SAM-defined genes overexpressed in MMC | | | None | | |

A supervised analysis was performed with the SAM software, as indicated in Table SIV.

References linked to additional data

- Achari, A., Marshall, S.E., Muirhead, H., Palmieri, R.H. & Noltmann, E.A. (1981) Glucose-6phosphate isomerase. *Philosophical Transactions of the Royal Society of London*. *Series B: Biological Sciences*, **293**, 145-157.
- Aikawa, J., Grobe, K., Tsujimoto, M. & Esko, J.D. (2001) Multiple isozymes of heparan sulphate/heparin GlcNAc N-deacetylase/GlcN N-sulfotransferase. Structure and activity of the fourth member, NDST4. *The Journal of Biological Chemistry*, 276, 5876-5882.
- Bai, X., Zhou, D., Brown, J.R., Crawford, B.E., Hennet, T. & Esko, J.D. (2001) Biosynthesis of the linkage region of glycosaminoglycans: cloning and activity of galactosyltransferase II, the sixth member of the beta 1,3-galactosyltransferase family (beta 3GalT6). *The Journal of Biological Chemistry*, **276**, 48189-48195.
- Bulter, T. & Elling, L. (1999) Enzymatic synthesis of nucleotide sugars. *Glycoconjugate Journal*, **16**, 147-159.
- Chen, J., Duncan, M.B., Carrick, K., Pope, R.M. & Liu, J. (2003) Biosynthesis of 3-Osulphated heparan sulphate: unique substrate specificity of heparan sulphate 3-Osulfotransferase isoform 5. *Glycobiology*, **13**, 785-794.
- Crawford, B.E., Olson, S.K., Esko, J.D. & Pinhal, M.A. (2001) Cloning, Golgi localization, and enzyme activity of the full-length heparin/heparan sulphate-glucuronic acid C5-epimerase. *The Journal of Biological Chemistry*, **276**, 21538-21543.
- Cuellar, K., Chuong, H., Hubbell, S.M. & Hinsdale, M.E. (2007) Biosynthesis of chondroitin and heparan sulphate in chinese hamster ovary cells depends on xylosyltransferase II. *The Journal of Biological Chemistry*, **282**, 5195-5200.
- Davidson, E.A., Blumenthal, H.J. & Roseman, S. (1957) Glucosamine metabolism. II. Studies on glucosamine 6-phosphate N-acetylase. *The Journal of Biological Chemistry*, **226**, 125-133.
- Drago, G.A., Hopkinson, D.A., Westwood, S.A. & Whitehouse, D.B. (1991) Antigenic analysis of the major human phosphoglucomutase isozymes: PGM1, PGM2, PGM3 and PGM4. *Annals of Human Genetics*, **55**, 263-271.
- Edwards, Y.H., Putt, W., Fox, M. & Ives, J.H. (1995) A novel human phosphoglucomutase (PGM5) maps to the centromeric region of chromosome 9. *Genomics*, **30**, 350-353.
- Esko, J.D. & Selleck, S.B. (2002) Order out of chaos: assembly of ligand binding sites in heparan sulphate. *Annual Review of Biochemistry*, **71**, 435-471.
- Fukuta, M., Kobayashi, Y., Uchimura, K., Kimata, K. & Habuchi, O. (1998) Molecular cloning and expression of human chondroitin 6-sulfotransferase. *Biochimica et Biophysica Acta*, **1399**, 57-61.
- Furukawa, K. & Okajima, T. (2002) Galactosyltransferase I is a gene responsible for progeroid variant of Ehlers-Danlos syndrome: molecular cloning and identification of mutations. *Biochimica et Biophysica Acta*, **1573**, 377-381.
- Ghosh, S., Blumenthal, H.J., Davidson, E. & Roseman, S. (1960) Glucosamine metabolism.
 V. Enzymatic synthesis of glucosamine 6-phosphate. *The Journal of Biological Chemistry*, 235, 1265-1273.
- Gotoh, M., Yada, T., Sato, T., Akashima, T., Iwasaki, H., Mochizuki, H., Inaba, N., Togayachi, A., Kudo, T., Watanabe, H., Kimata, K. & Narimatsu, H. (2002) Molecular cloning and characterization of a novel chondroitin sulphate glucuronyltransferase that transfers glucuronic acid to N-acetylgalactosamine. *The Journal of Biological Chemistry*, 277, 38179-38188.
- Gotting, C., Kuhn, J., Zahn, R., Brinkmann, T. & Kleesiek, K. (2000) Molecular cloning and expression of human UDP-d-Xylose:proteoglycan core protein beta-dxylosyltransferase and its first isoform XT-II. *Journal of Molecular Biology*, **304**, 517-528.

- Habuchi, H., Kobayashi, M. & Kimata, K. (1998) Molecular characterization and expression of heparan-sulphate 6-sulfotransferase. Complete cDNA cloning in human and partial cloning in Chinese hamster ovary cells. *The Journal of Biological Chemistry*, **273**, 9208-9213.
- Hiraoka, N., Nakagawa, H., Ong, E., Akama, T.O., Fukuda, M.N. & Fukuda, M. (2000) Molecular cloning and expression of two distinct human chondroitin 4-Osulfotransferases that belong to the HNK-1 sulfotransferase gene family. *The Journal* of Biological Chemistry, 275, 20188-20196.
- Ishida, N. & Kawakita, M. (2004) Molecular physiology and pathology of the nucleotide sugar transporter family (SLC35). *Pflugers Archiv. European Journal of Physiology*, 447, 768-775.
- Izumikawa, T., Uyama, T., Okuura, Y., Sugahara, K. & Kitagawa, H. (2007) Involvement of chondroitin sulphate synthase-3 (chondroitin synthase-2) in chondroitin polymerization through its interaction with chondroitin synthase-1 or chondroitinpolymerizing factor. *The Biochemical Journal*, **403**, 545-552.
- Kamiyama, S., Sasaki, N., Goda, E., Ui-Tei, K., Saigo, K., Narimatsu, H., Jigami, Y., Kannagi, R., Irimura, T. & Nishihara, S. (2006) Molecular cloning and characterization of a novel 3'-phosphoadenosine 5'-phosphosulphate transporter, PAPST2. *The Journal of Biological Chemistry*, **281**, 10945-10953.
- Kamiyama, S., Suda, T., Ueda, R., Suzuki, M., Okubo, R., Kikuchi, N., Chiba, Y., Goto, S., Toyoda, H., Saigo, K., Watanabe, M., Narimatsu, H., Jigami, Y. & Nishihara, S. (2003) Molecular cloning and identification of 3'-phosphoadenosine 5'phosphosulphate transporter. *The Journal of Biological Chemistry*, 278, 25958-25963.
- Kang, H.G., Evers, M.R., Xia, G., Baenziger, J.U. & Schachner, M. (2002) Molecular cloning and characterization of chondroitin-4-O-sulfotransferase-3. A novel member of the HNK-1 family of sulfotransferases. *The Journal of Biological Chemistry*, 277, 34766-34772.
- Kim, B.T., Kitagawa, H., Tamura, J., Saito, T., Kusche-Gullberg, M., Lindahl, U. & Sugahara, K. (2001) Human tumor suppressor EXT gene family members EXTL1 and EXTL3 encode alpha 1,4- N-acetylglucosaminyltransferases that likely are involved in heparan sulphate/ heparin biosynthesis. *Proceeding of the National Academy of Sciences of the United States of America*, **98**, 7176-7181.
- Kitagawa, H., Izumikawa, T., Uyama, T. & Sugahara, K. (2003) Molecular cloning of a chondroitin polymerizing factor that cooperates with chondroitin synthase for chondroitin polymerization. *The Journal of Biological Chemistry*, **278**, 23666-23671.
- Kitagawa, H., Shimakawa, H. & Sugahara, K. (1999) The tumor suppressor EXT-like gene EXTL2 encodes an alpha1, 4-N-acetylhexosaminyltransferase that transfers Nacetylgalactosamine and N-acetylglucosamine to the common glycosaminoglycanprotein linkage region. The key enzyme for the chain initiation of heparan sulphate. *The Journal of Biological Chemistry*, **274**, 13933-13937.

Kitagawa, H., Uyama, T. & Sugahara, K. (2001) Molecular cloning and expression of a human chondroitin synthase. *The Journal of Biological Chemistry*, **276**, 38721-38726.

- Markovich, D. (2001) Physiological roles and regulation of mammalian sulphate transporters. *Physiological Reviews*, **81**, 1499-1533.
- Markovich, D. & Murer, H. (2004) The SLC13 gene family of sodium sulphate/carboxylate cotransporters. *Pflugers Archiv. European Journal of Physiology*, **447**, 594-602.

Morimoto-Tomita, M., Uchimura, K., Werb, Z., Hemmerich, S. & Rosen, S.D. (2002) Cloning and characterization of two extracellular heparin-degrading endosulfatases in mice and humans. *The Journal of Biological Chemistry*, **277**, 49175-49185.

Novelli, G. & Reichardt, J.K. (2000) Molecular basis of disorders of human galactose metabolism: past, present, and future. *Molecular Genetics and Metabolism*, **71**, 62-65.

Ohtake, S., Ito, Y., Fukuta, M. & Habuchi, O. (2001) Human N-acetylgalactosamine 4sulphate 6-O-sulfotransferase cDNA is related to human B cell recombination activating gene-associated gene. *The Journal of Biological Chemistry*, **276**, 43894-43900.

- Ohtake, S., Kimata, K. & Habuchi, O. (2005) Recognition of sulfation pattern of chondroitin sulphate by uronosyl 2-O-sulfotransferase. *The Journal of Biological Chemistry*, **280**, 39115-39123.
- Oki, T., Yamazaki, K., Kuromitsu, J., Okada, M. & Tanaka, I. (1999) cDNA cloning and mapping of a novel subtype of glutamine:fructose-6-phosphate amidotransferase (GFAT2) in human and mouse. *Genomics*, **57**, 227-234.
- Oohira, A., Matsui, F., Tokita, Y., Yamauchi, S. & Aono, S. (2000) Molecular interactions of neural chondroitin sulphate proteoglycans in the brain development. *Archives of Biochemistry and Biophysics*, **374**, 24-34.
- Pedersen, L.C., Tsuchida, K., Kitagawa, H., Sugahara, K., Darden, T.A. & Negishi, M. (2000) Heparan/chondroitin sulphate biosynthesis. Structure and mechanism of human glucuronyltransferase I. *The Journal of Biological Chemistry*, **275**, 34580-34585.
- Prabhakar, V. & Sasisekharan, R. (2006) The biosynthesis and catabolism of galactosaminoglycans. *Advances in Pharmacology*, **53**, 69-115.
- Probstmeier, R., Braunewell, K. & Pesheva, P. (2000) Involvement of chondroitin sulphates on brain-derived tenascin-R in carbohydrate-dependent interactions with fibronectin and tenascin-C. *Brain Research*, **863**, 42-51.
- Sato, T., Gotoh, M., Kiyohara, K., Akashima, T., Iwasaki, H., Kameyama, A., Mochizuki, H., Yada, T., Inaba, N., Togayachi, A., Kudo, T., Asada, M., Watanabe, H., Imamura, T., Kimata, K. & Narimatsu, H. (2003) Differential roles of two N-acetylgalactosaminyltransferases, CSGalNAcT-1, and a novel enzyme, CSGALNACT2. Initiation and elongation in synthesis of chondroitin sulphate. *The Journal of Biological Chemistry*, 278, 3063-3071.
- Shevchenko, V., Hogben, M., Ekong, R., Parrington, J. & Lai, F.A. (1998) The human glucosamine-6-phosphate deaminase gene: cDNA cloning and expression, genomic organization and chromosomal localization. *Gene*, **216**, 31-38.
- Su, A.I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K.A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G., Cooke, M.P., Walker, J.R. & Hogenesch, J.B. (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. *Proceeding of the National Academy of Sciences of the United States of America*, **101**, 6062-6067.
- Szumilo, T., Zeng, Y., Pastuszak, I., Drake, R., Szumilo, H. & Elbein, A.D. (1996) Purification to homogeneity and properties of UDP-GlcNAc (GalNAc) pyrophosphorylase. *The Journal of Biological Chemistry*, **271**, 13147-13154.
- Timson, D.J. & Reece, R.J. (2003) Identification and characterisation of human aldose 1epimerase. *FEBS Letter*, **543**, 21-24.
- Vlodavsky, I., Friedmann, Y., Elkin, M., Aingorn, H., Atzmon, R., Ishai-Michaeli, R., Bitan, M., Pappo, O., Peretz, T., Michal, I., Spector, L. & Pecker, I. (1999) Mammalian heparanase: gene cloning, expression and function in tumor progression and metastasis. *Nature Medicine*, 5, 793-802.
- Wegrowski, Y. & Maquart, F.X. (2006) Chondroitin sulphate proteoglycans in tumor progression. *Advances in Pharmacology*, **53**, 297-321.
- Whitmore, T.E., Mudri, S.L. & McKnight, G.L. (1995) Physical mapping of the human glutamine:fructose-6-phosphate amidotransferase gene (GFPT) to chromosome 2p13. *Genomics*, 26, 422-423.
- Wilson, J.E. (2003) Isozymes of mammalian hexokinase: structure, subcellular localization and metabolic function. *Journal of Experimental Biology*, **206**, 2049-2057.
- Wood, I.S. & Trayhurn, P. (2003) Glucose transporters (GLUT and SGLT): expanded families of sugar transport proteins. *British Journal of Nutrition*, **89**, 3-9.
- Xu, D., Song, D., Pedersen, L.C. & Liu, J. (2007) Mutational study of heparan sulphate 2-Osulfotransferase and chondroitin sulphate 2-O-sulfotransferase. *The Journal of Biological Chemistry*, 282, 8356-8367.

- Xu, Z.H., Otterness, D.M., Freimuth, R.R., Carlini, E.J., Wood, T.C., Mitchell, S., Moon, E., Kim, U.J., Xu, J.P., Siciliano, M.J. & Weinshilboum, R.M. (2000) Human 3'phosphoadenosine 5'-phosphosulphate synthetase 1 (PAPSS1) and PAPSS2: gene cloning, characterization and chromosomal localization. *Biochemical and Biophysical Research Communications*, 268, 437-444.
- Yada, T., Sato, T., Kaseyama, H., Gotoh, M., Iwasaki, H., Kikuchi, N., Kwon, Y.D., Togayachi, A., Kudo, T., Watanabe, H., Narimatsu, H. & Kimata, K. (2003) Chondroitin sulphate synthase-3. Molecular cloning and characterization. *The Journal* of Biological Chemistry, 278, 39711-39725.
- Zak, B.M., Crawford, B.E. & Esko, J.D. (2002) Hereditary multiple exostoses and heparan sulphate polymerization. *Biochimica et Biophysica Acta*, **1573**, 346-355.