Expression and regulation of sialic acid biosynthesis and sialyltransferase genes in the mammary gland

Jovana Maksimovic

Submitted in total fulfillment of the requirements of the degree of Doctor of Philosophy

Monash Institute of Medical Research, Centre for Reproduction and Development

Faculty of Medicine, Nursing and Health Sciences

Monash University

September 2011

Notice 1

Notice 2

Abstract

Sialic acids are a family of nine-carbon, acidic monosaccharides. In mammals, they are typically located at the non-reducing terminus of oligosaccharide chains, connected either via an α2,3- or α2,6- linkage. Sialic acids can also form α2,8-linked poly-sialic acid chains on gangliosides and glycoproteins.

Milk sialylglycoconjugates can protect the gastrointestinal tract of the suckling neonate by competitively binding to invading pathogens and promoting growth of beneficial flora, and their potential role in post-natal brain development is of particular interest in human infant nutrition. Although the concentration and distribution of sialylglycoconjugates has been extensively studied in the milk of various species, the regulation of sialylation in the mammary gland, in the context of lactation, has been limited.

Sialic acid can be synthesised in all tissues to satisfy local sialylation requirements, although the final complement of sialylated structures is determined by the sialyltransferase enzymes. Sialyltransferases are a family of Golgi membrane-bound, glycosyltransferases that transfer sialic acid from CMP-sialic acid to carbohydrate acceptor groups of glycoproteins, glycolipids and free oligosaccharides, forming various sialylglycoconjugates. Their expression is primarily controlled at the level of transcription and is expedited through the use of several tissue and development-specific promoters, giving rise to numerous transcripts divergent only in their 5' untranslated regions. Thus far, 20 mammalian sialyltransferases have been cloned. Depending on the type of linkage formed and the type of sugar acceptor, they are segregated into four groups: ST6Gal, ST6GalNAc, ST3Gal and ST8Sia.

This study describes the expression of sialic acid biosynthesis and sialyltransferase genes in the lactating eutherian and marsupial mammary gland and the role of lactogenic hormones in their regulation. It also examines the developmental context in which milk

sialylglycoconjugates may be important. Analysis of mouse mammary gland lactation microarray time course data revealed that key sialic acid biosynthesis genes are up-regulated during lactation, but that only a subset of known sialyltransferase genes are expressed at all. Explant culture of mouse mammary tissue demonstrated that lactogenic hormones are involved in the regulation of the genes found to be differentially expressed during mouse lactation. These genes were subsequently characterised in the tammar wallaby and their expression assayed across the lactation cycle. A correlation was observed between the developmental state of the tammar neonate at peak mammary gland sialyltransferase gene expression and the developmental state of the mouse at birth, suggesting that an increase in sialylated milk products is particularly important during this developmental stage. Analysis of the expression of the ST6Gal I gene in the bovine mammary gland during lactation, showed a remarkably different profile of expression to that of the mouse and wallaby.

In conclusion, this study has demonstrated that transcriptional control is a key factor in the regulation of sialylation in the mammary gland during lactation and that lactogenic hormones drive the changes in gene expression. Also, given the difference in ST6Gal I gene expression between mice and cows, it is apparent that the regulation of sialylation varies between different species, possibly catering for the developmental state of the young at birth.

Acknowledgments

Many people and organisations have supported me throughout my studies. By its very nature, my project has been very independent, and although this has allowed me an incredible degree of academic freedom, it was only through the generosity of others that it was made possible. I am immensely grateful to everyone who contributed their time, effort and resources.

My supervisors Dr. Keith Savin and Prof. Kevin Nicholas have been beyond supportive throughout this endeavour. Despite geographic challenges, they always made time for discussions and meetings, which was greatly appreciated. I thank Keith and Kevin for all of their help and encouragement. I would also like to acknowledge Dr. Julie Sharp for her help with all things lab-related – I'm sure that as someone whose focus had previously been primarily bioinformatics, this project would have been over before it had started had it not been for her help. I would also like to thank A/Prof. Rosemary Horne at the Monash Institute of Medical Research for helping me navigate Monash University's processes and procedures and for dealing with my "unique" supervisory arrangements.

I am thankful to have had financial support from both the Victorian Department of Primary Industries (DPI) and Deakin University for my PhD research. I would particularly like to acknowledge the Discovery Technologies Platform in the Biosciences Research Division at DPI for adopting me as a surrogate member of the group and contributing resources to my project.

I am immeasurably grateful for all of the friendship, support and advice from colleagues at DPI Attwood and Bundoora. I have made many friends during my time there and I cannot thank them enough for making it such a fun and relaxed place to work. In particular, thanks to Carl, Charlotte, Nick, and TJ for always keeping me entertained at work!

Finally, this undertaking would not have been possible without the support of my family and friends. I thank my mum and dad for their unconditional love and for always believing that I could achieve anything I put my mind to. Also, to my sister Sandra, you have always been one of my best friends and I thank you for helping me keep everything in perspective. To my friends, thank you for your support and for reminding me that there is life outside of a PhD. To Davor, the last four years have tested our partnership, our friendship and our love. I appreciate your seemingly endless understanding and patience. You are truly one of a kind and I am incredibly lucky to have met you. I look forward to our future adventures together.

List of publications

Journal articles

Maksimovic J, Sharp JA, Nicholas KR, Cocks BG, Savin K. 2011. Conservation of the ST6Gal I gene and its expression in the mammary gland. *Glycobiology*, 21:467-481. (Chapter 2)

Conference proceedings

Maksimovic J, Savin K, Nicholas KR, Holland M (2008) Transcriptional Regulation of the β-galactoside α-2,6 Sialyltransferase I (ST6Gal I) Gene in the Bovine Mammary Gland. *Proceedings of the 5th International Symposium on Milk Genomics and Human Health*, 14-16 October 2008, Sydney, Australia. (Part of chapter 2)

Maksimovic J, Nicholas KR, Cocks BG, Savin K (2009) Developmental regulation of sialyltransferase gene expression in the murine mammary gland. *Proceedings of the 6th International Symposium on Milk Genomics and Human Health,* 28-30 September 2009, Paris, France. (Part of chapter 3)

Monash University
Monash Research Graduate School

General declaration

Declaration for a thesis based based on conjointly published or unpublished work

In accordance with Monash University Doctorate Regulation 17/ Doctor of Philosophy and Master of Philosophy (MPhil) regulations the following declarations are made:

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes 1 original paper published in a peer reviewed journal and 3 unpublished publications. The core theme of the thesis is the transcriptional control of sialylation in the lactating mammary gland. The expression and regulation of key sialic acid biosynthesis and sialyltransferase genes was investigated in the lactating mammary gland of several species and the relationship to neonate development explored. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the candidate, working within the Centre for Reproduction and Development, Monash Institute of Medical Research under the supervision of Professor Justin St. John, Professor Kevin Nicholas (External) and Dr. Keith Savin (External). The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

In the case of chapters 2, 3, 4 and 5 my contribution to the work involved the following:

For all of the studies presented in this thesis I developed the concepts with the aid of my supervisors Professor Kevin Nicholas and Doctor Keith Savin, and performed all of the laboratory experiments, subsequent analyses and data interpretation. For chapter 2, the

bovine mammary tissue used was obtained from colleagues at The University of Sydney, who had collected it for an unrelated study, which is acknowledged in the text. I performed the processing of the tissue and all subsequent experiments. The human RNA used in chapter 2 was extracted from cells harvested from human breast milk by Doctor Julie Sharp for an unrelated study. Using the human RNA, I performed all subsequent experiments that are presented in chapter 2. The microarray data used in chapter 3 had been previously published but was re-analysed and examined in a different context for the purposes of the current study. I performed all aspects of the mammary explant culture experiment presented in chapter 3 and all subsequent analyses. My supervisor, Professor Kevin Nicholas, performed the dissection of the mouse mammary glands and aided in the preparation of the explants. The tammar wallaby tissue used in chapter 5 had been collected by various individuals, over a number of years, for use in unrelated studies. I processed the tissue used in the current study and performed all of the subsequent experimental work.

Thesis chapter	**Publication title**	**Publication status**	**Nature and extent of candidate's contribution**
2	Conservation of the ST6Gal I gene and its expression in the mammary gland	Published	For this chapter I preformed all of the laboratory experiments and was responsible for the interpretation and analysis of the data and the writing of the manuscript. Dr. Julie Sharp initially collected and processed the human milk samples used in this study. The extent of my contribution was 80%
3	Differential expression of sialic acid biosynthesis and sialyltransferase genes in the mouse mammary gland during lactation	Submitted	For this chapter I analysed the publicly available microarray data and was responsible for the interpretation of the results and the writing of the manuscript. The extent of my contribution was 80%.

Thesis chapter	Publication title	Publication status	Nature and extent of candidate's contribution
4	Expression of sialic acid biosynthesis and sialyltransferase genes is regulated by lactogenic hormones in mouse mammary gland tissue culture	Submitted	For this chapter I performed the laboratory experiments and was responsible for analysis of the data, interpretation of the results and the writing of the manuscript. Prof. Nicholas was responsible for dissecting the mammary glands from the mice and aided in the preparation of the explants. The extent of my contribution was 80%.
5	Characterisation and expression of sialic acid biosynthesis and sialyltransferase genes in the lactating tammar wallaby mammary gland	Submitted	For this chapter I performed the laboratory experiments and was responsible for analysis of the data, interpretation of the results and the writing of the manuscript. The extent of my contribution was 80%. The tammar wallaby tissue used in this study had been collected previously, over a number of years, by various individuals for other studies.

I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Signed: ……

Date: ……

List of abbreviations

APR	acute phase response
BH	Benjamini-Hochberg
Blg	β-lactoglobulin
C/EBP	CCAAT-enhancer-binding proteins
ChIP-seq	chromatin immunoprecipitation sequencing
CMAS	cytidine monophospho-N-acetylneuraminic acid synthetase
CMP	cytidine monophosphate
CNRQ	calibrated normalised relative quantities
Csn1 s1	casein alpha s1
Csn1 s2a	casein alpha s2-like a
Csn1 s2b	casein alpha s2-like b
Csn2	casein beta
Csn3	casein kappa
ECR	evolutionarily conserved region
EMSA	electrophoretic mobility shift assay
F	hydrocortisone
FVB	friend leukemia virus B
Gal	β-D-galactopyranosyl
GalNAc	β-D-N-acetylgalactosaminyl
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GDP	guanosine diphosphate
GlcNAc	β-D-N-acetylglucosaminyl
GNE	UDP-GlcNAc 2-epimerase/ManNAc kinase
GR	glucocorticoid receptor
I	insulin
IL	interleukin
IPA	Ingenuity Pathways Analysis
Kdn	2-keto-3-deoxy-D-glycero-D-galactonononic acid
L	long sialylmotif

LALBA	α-lactalbumin
Ltf	lactotransferrin
ManNAc	N-Acetylmannosamine
MUC1	mucin 1
NANS	N-acetylneuraminic acid synthase
NCAM	neural cell adhesion molecule
Neu5Gc	N-glycolylneuraminic acid
Neu5NAc	N-acetylneuraminic acid
NH	no hormone
OCT	octamer transcription factor
P	prolactin
PCR	polymerase chain reaction
PMA	present/marginal/absent
PWM	positional weight matrix
qPCR	quantitative real-time PCR
RACE	rapid amplification of cDNA ends
RNA-seq	RNA sequencing
RT-PCR	reverse transcription PCR
S	short sialylmotif
ST3Gal	beta-galactoside alpha-2,3-sialyltransferase
ST6Gal	beta galactoside alpha 2,6 sialyltransferase
ST6GalNAc	(alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase
ST8Sia	alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase
STAT	signal transducer and activator of transcription
TFBS	transcription factor binding site
TSS	transcription start site
UDP	uridine diphosphate
UTR	untranslated region
VS	very short sialylmotif
WAP	whey acidic protein
YY1	yin yang 1

Table of contents

Abstract iii
Acknowledgments v
List of publications vii
General declaration viii
List of abbreviations xi
Table of contents xiii

1 Chapter one 1
1 Literature review 1
1.1 Glycan biosynthesis 1
1.1.1 Regulation of lactose synthesis 3
1.2 Sialic acid 5
1.3 Sialic acid biosynthesis 8
1.4 Sialyltransferases 10
1.4.1 ST6Gal I 11
1.5 Human milk oligosaccharides 14
1.6 Sialic acid in the nervous system 16
1.7 Sialic acid and lactation 17
1.7.1 Sialic acid, human milk and infant formula 17
1.7.2 Sialic acid in mouse and rat milk 19
1.7.3 Sialic acid in marsupial milk 21
1.8 Rationale for study 25
1.9 Aims 28
1.10 Hypothesis 29
2 Chapter two 31
2.1 Declaration for thesis chapter two 32
2.2 Conservation of the ST6Gal I gene and its expression in the mammary gland 35
2.2.1 Abstract 36
2.2.2 Introduction 37
2.2.3 Methods 40
2.2.4 Results 46
2.2.5 Discussion 62
2.2.6 References 68
3 Chapter three 77
3.1 Declaration for thesis chapter three 78
3.2 Differential expression of sialic acid biosynthesis and sialyltransferase genes in the mouse mammary gland during lactation 81
3.2.1 Abstract 82
3.2.2 Introduction 83
3.2.3 Methods 87
3.2.4 Results 91

3.2.5 Discussion 112
3.2.6 References 120
4 Chapter four 127
4.1 Declaration for thesis chapter four 128
4.2 Expression of sialic acid biosynthesis and sialyltransferase genes is regulated by lactogenic hormones in mouse mammary gland tissue culture 131
4.2.1 Abstract 132
4.2.2 Introduction 133
4.2.3 Methods 138
4.2.4 Results 142
4.2.5 Discussion 159
4.2.6 References 165
5 Chapter five 175
5.1 Declaration for thesis chapter five 176
5.2 Characterisation and expression of sialic acid biosynthesis and sialyltransferase genes in the lactating tammar wallaby mammary gland 179
5.2.1 Abstract 180
5.2.2 Introduction 181
5.2.3 Methods 186
5.2.4 Results 191
5.2.5 Discussion 206
5.2.6 References 212
6 Chapter six 223
6 Discussion and future directions 223
6.1 Transcriptional control of sialylation in the mammary gland 224
6.2 Sialic acid in neonatal development 230
6.3 Sialic acid and infant formula 234
6.4 Methodological considerations 235
6.5 Conclusions 237
7 Bibliography 239
8 Appendices 261
8.1 Published manuscript (Chapter two) 263

1 Literature review

1.1 Glycan biosynthesis

The types of glycans that are synthesised in any given cell are primarily determined by the glycosyltransferase enzymes available to assemble various monosaccharides into linear and branched glycan chains (Rini, J., Esko, J., et al. 2009). Given the vast variety of glycans present in nature, it is unsurprising that glycosyltransferases are actually a very large family of enzymes that all have the ability to transfer the monosaccharide component of a simple nucleotide sugar donor (e.g. UDP-galactose, GDP-fucose, CMP-sialic acid) to an acceptor substrate (Rini, J., Esko, J., et al. 2009).

Glycosyltransferase acceptor substrates are varied and include oligosaccharides, monosaccharides, polypeptides, lipids, small organic molecules, and DNA (Hennet, T. 2002, Rini, J., Esko, J., et al. 2009). Most glycosyltransferases operate to elongate glycan chains, acting sequentially, so that the product of one enzyme is a preferred acceptor substrate for the successive action of another, resulting in linear and/or branched glycan chains (Figure 1). Generally, glycosyltransferases are specific towards a certain nucleotide sugar donor and acceptor, and thus most glycosidic linkages are the product of a single enzyme (Hennet, T. 2002). However, there are exceptions to this rule; the β1–4 galactosyltransferase I, involved in N-acetyllactosamine formation, can operate on multiple acceptors (Furukawa, K. and Sato, T. 1999). In the mammary gland, when β1-4 galactosyltransferase I is bound to α-lactalbumin as part of the lactose synthase complex (Brodbeck, U., Denton, W.L., et al. 1967), its acceptor specificity shifts from N-acetylglucosamine to glucose, enabling the synthesis of lactose during lactation (Morrison, J.F. and Ebner, K.E. 1971, Osborne, J.C. and Steiner, R.F. 1974).

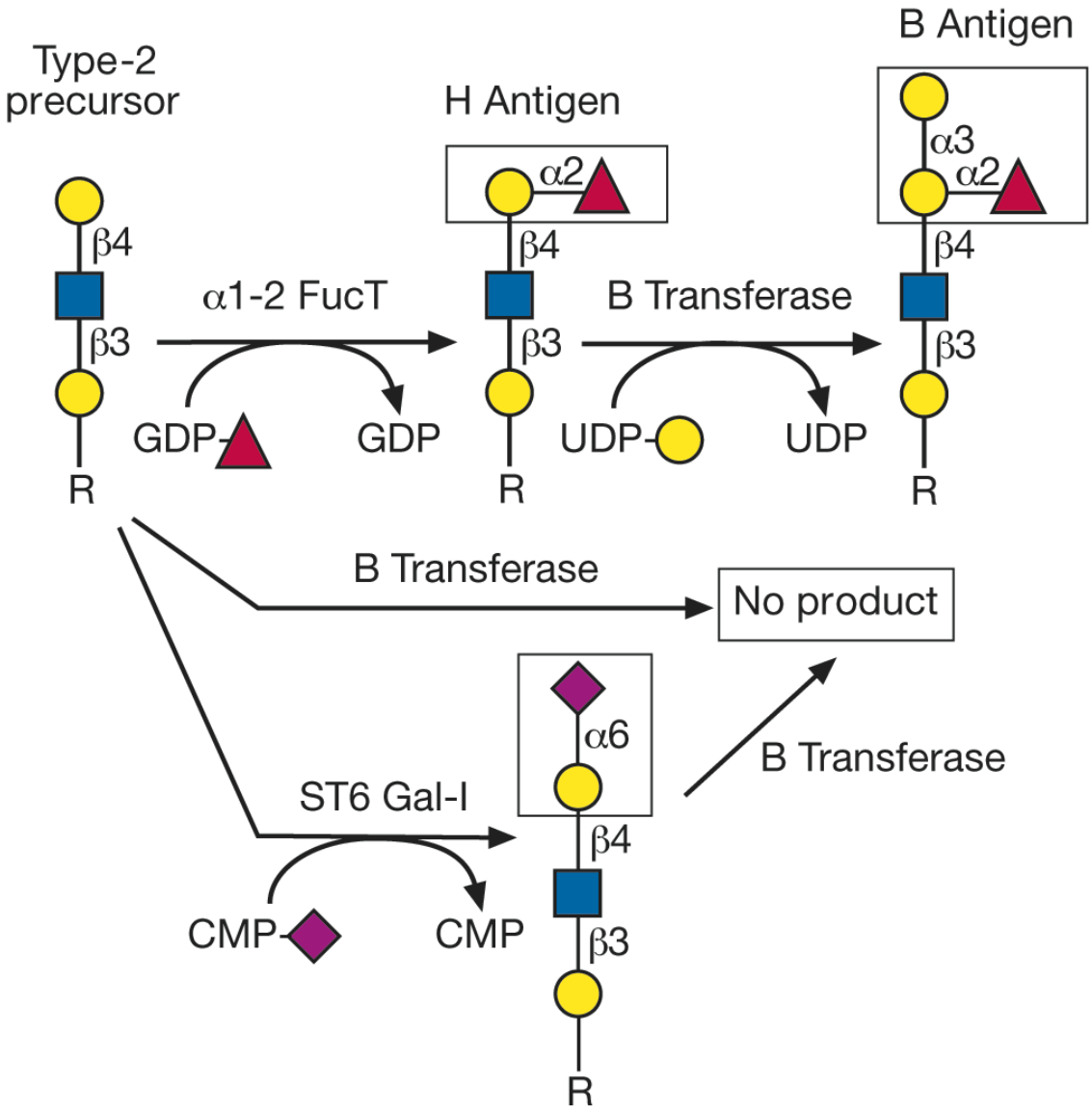


Figure 1: The strict acceptor substrate specificity of glycosyltransferases exemplified by the human B blood group α1-3 galactosyltransferase. The B transferase adds galactose to the H antigen using an α1-3 linkage (top). This enzyme can only operate on the α1-2-linked fucose modified H antigen; the B transferase cannot use the unmodified type-2 precursor as an acceptor (middle), or precursors containing sialyl residues (bottom). Yellow circle = galactose; blue sequare = N-acetylglucosamine; purple diamond = N-acetylneuraminic acid (Neu5Ac); red triangle = fucose [modified from (Rini, J., Esko, J., et al. 2009)].

1.1.1 Regulation of lactose synthesis

In the mammary gland, α-lactalbumin is an important regulator of β1-4 galactosyltransferase (Morrison, J.F. and Ebner, K.E. 1971, Osborne, J.C. and Steiner, R.F. 1974); it lowers the K_m of the enzyme for the transfer of UDP-galactose to glucose in order to synthesize lactose (Lönnerdal, B. and Lien, E.L. 2003). Lactose is the major osmotic regulator of milk volume formation and is essential for milk production (Stacey, A., Schnieke, A., et al. 1995, Stinnakre, M.G., Vilotte, J.L., et al. 1994). α-lactalbumin knockout mice were shown to be unable to feed their pups as their milk was too viscous to be removed by normal suckling (Stacey, A., Schnieke, A., et al. 1995, Stinnakre, M.G., Vilotte, J.L., et al. 1994). Alternatively, when the mouse α-lactalbumin gene was replaced with the human α-lactalbumin gene, lactation proceeded as normal. However, the human α-lactalbumin gene was found to express 15 times more mRNA than the mouse gene, which resulted in 14 times more protein, demonstrating that the concentration of α-lactalbumin is directly related to gene expression, and that the primary regulators of human α-lactalbumin gene expression are in close proximity to or within the gene itself (Stacey, A., Schnieke, A., et al. 1995).

α-lactalbumin is first expressed in the mammary gland during lactogenesis stage 1, which occurs in mid-pregnancy; milk secretion is actively suppressed by progesterone during this stage (Harigaya, T., Sakai, S., et al. 1978, Mizoguchi, Y., Kim, J.Y., et al. 1996, Neville, M.C., Morton, J., et al. 2001). During the peri partum period, the gland begins the secretion of colostrum, and subsequently milk. This second phase is lactogenesis 2, and is characterised by increased expression of the milk protein genes, including α-lactalbumin, and is accompanied by further morphological and physiological changes in the mammary gland (Hartmann, P.E. 1973, Neville, M.C., Morton, J., et al. 2001, Traurig, H.H. 1967). Progesterone withdrawal in the presence of high levels of prolactin is absolutely necessary for this phase (Humphreys, R.C., Lydon, J.P., et al. 1997, Kuhn, N.J. 2009, Ormandy, C.J., Binart, N., et al. 1997).

Pioneering work in the late 1950s and 1960s established a model system for the study of mammary differentiation *in vitro* (Elias, J.J. 1957, Elias, J.J. 1959, Elias, J.J. and Rivera, E. 1959, Nandi, S. and Bern, H.A. 1960, Rivera, E.M. and Bern, H.A. 1961, Rivera, E.M., Forsyth, I.A., et al. 1967). Chemically defined media containing insulin, glucocorticoid or aldosterone, and prolactin or human growth hormone in various combinations were examined for their effects on mouse mammary explants in 5- l0 day culture experiments. Remarkably, in the presence of insulin, adrenal steroid, and prolactin most of the epithelium appeared secretory. Similar observations were reported by Barnawell (Barnawell, E.B. 1965), using mammary explants in various species. This approach was subsequently used to demonstrate casein (Juergens, W.G., Stockdale, F.E., et al. 1965, Turkington, R.W., Juergens, W.G., et al. 1965) and α-lactalbumin synthesis (O'Keefe, E. and Cuatrecasas, P. 1974, Ono, M. and Oka, T. 1980a, Ono, M. and Oka, T. 1980b) in the presence of insulin, glucocorticoid and prolactin.

It is now generally accepted that lactation requires the presence of prolactin along with insulin and hydrocortisone (Ostrom, K.M. 1990). More recently, in pig mammary explant cultures, Dodd et al. (1994) observed maximal accumulation of α-lactalbumin in the presence of prolactin alone, which did not appear to be enhanced by insulin or corticosterone, or by a combination of insulin and corticosterone, although they conceded that it was possible that endogenous hormones from the animal could have contributed to these results. The rabbit mammary gland has been shown to respond to prolactin alone *in vivo* (Cowie, A.T., Hartmann, P.E., et al. 1969) and *in vitro* to produce casein (Devinoy, E., Houdebine, L.M., et al. 1978) and α-lactalbumin (Sankaran, L. and Topper, Y.J. 1984), although insulin, or a combination of insulin and cortisol, was found to enhance the action of prolactin particularly after prolonged culture periods (Sankaran, L. and Topper, Y.J. 1984). Prolactin is also thought to be the primary factor influencing α-lactalbumin induction in the tammar wallaby (Collet, C., Joseph, R., et al. 1991, Nicholas, K.R. and Tyndale-Biscoe, C.H. 1985). Therefore, this experimental culture model has proved useful

for determining the minimal hormonal requirements for the induction of α-lactalbumin gene expression in many species.

1.2 Sialic acid

Sialic acids are a large family of highly related, nine-carbon, acidic monosaccharides based on neuraminic acid. While “sialic acid” often refers only to N-acetylneuraminic acid (Neu5Ac) there are actually 4 “core” sialic acid structures (Neu5Ac, Neu5Gc, Kdn, and Neu) that can be modified with one or more additional substitutions (O-acetyl, O-methyl, O-sulfate, O-lactyl, or phosphate groups) at the hydroxyl groups on C-4, C-7, C-8, and C-9 or, at C-5, with the addition of a free amino group (Figure 2) (Varki, A. and Schauer, R. 2009). H typically dissociates from the carboxyl group (C-1) at physiological pH, giving sialic acids their the negative charge (Figure 2) (Varki, A. and Schauer, R. 2009). Given this vast variety of possible modifications, coupled with several distinct linkages that can be formed with other molecules, there are literally hundreds of configurations in which sialic acids exist in nature.

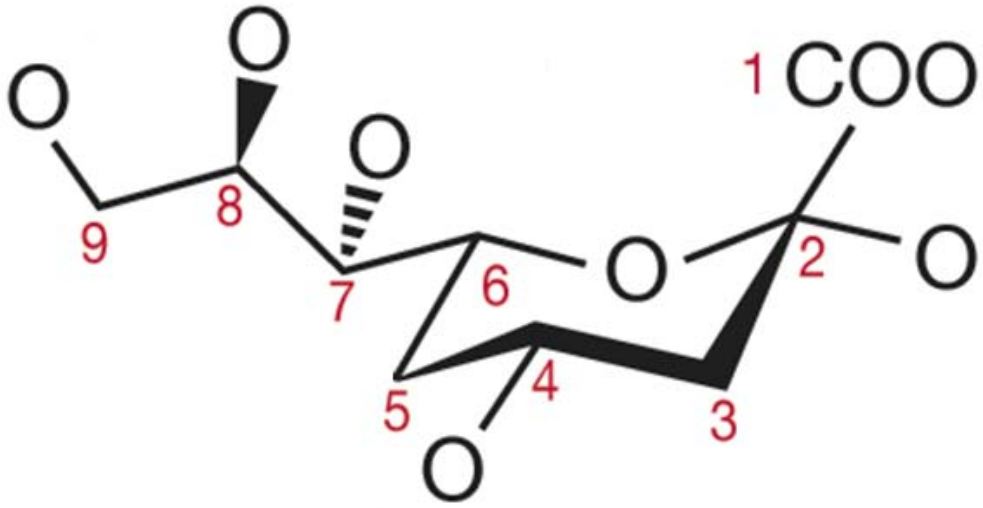


Figure 2: The nine carbon backbone common to all sialic acid molecules (α-configuration). Modifications can be made at any of the hydroxyl groups at C-4, C-7, C-8, and C-9 or, at C-5, with the addition of a free amino group. At physiological pH the carboxyl group is negatively charged [modified from Varki and Schauer (2009)].

N-acetylneuraminic acid (Neu5Ac) is the most common member of the family of molecules derived from neuraminic acid (Figure 3). In mammalian species, sialic acids are typically located at the non-reducing terminus of oligosaccharide chains, connected either via an α2,3- or α2,6- linkage to a β-D-galactopyranosyl (Gal) residue, by an α2,6-linkage to a β-D-N-acetylgalactosaminyl (GalNAc) residue or by an α2,6-linkage to a β-D-N-acetylglucosaminyl (GlcNAc) residue. Sialic acids can also form α2,8-linked poly-sialic acid chains on gangliosides and glycoproteins [reviewed in (Harduin-Lepers, A., Recchi, M.A., et al. 1995, Tsuji, S. 1996a)].

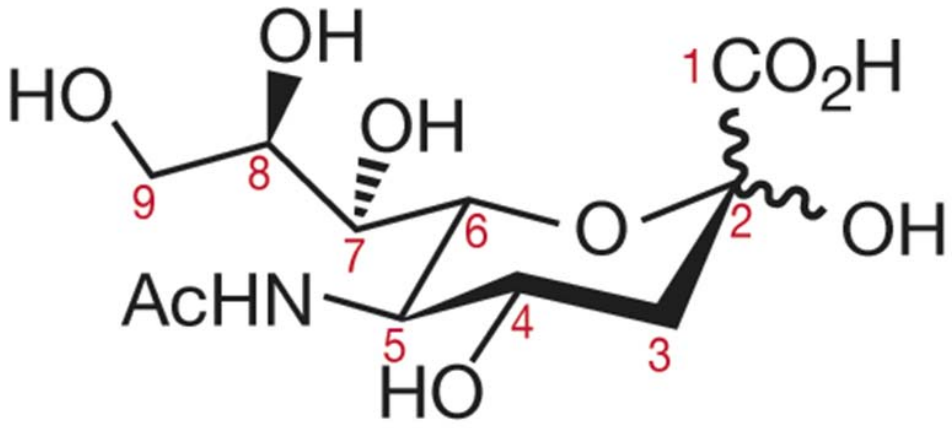


Figure 3: Structure of N-acetylneuraminic acid (Neu5Ac). All other sialic acids appear to be metabolically derived from either Neu5Ac or 2-keto-3-deoxy-D-glycero-D-galactonononic acid (Kdn, not shown). Neu5Ac is more common than Kdn in vertebrates. Glycosidically bound sialic acids in naturally occurring glycans are in the α-configuration (pictured) [modified from Varki and Schauer (2009)].

1.3 Sialic acid biosynthesis

Sialic acid is synthesised in the cytoplasm by a series of biochemical reactions beginning with UDP-GlcNAc as an initial substrate (Figure 4). GNE, the bifunctional enzyme UDP-GlcNAc 2-epimerase/ManAc kinase, first catalyses the epimerisation of UDP-GlcNAc to ManNAc and subsequently phosphorylates ManNAc to ManNAc-6-P. The epimerisation step is the first committed, and rate limiting step in sialic acid biosynthesis (Gal, B., Ruano, M.J., et al. 1997), making GNE an important regulator of sialylation. Keppler et al. (1999) demonstrated that GNE activity can be controlled at the level of transcription and that this can influence the sialylation and function of particular cell-surface molecules expressed on B-cells and myeloid cells. NANS, N-acetylneuraminic acid synthase, is responsible for converting MaNAc-6-P to NeuNAc-9-P. In the mouse, it is ubiquitously expressed in various tissues (Nakata, D., Close, B.E., et al. 2000). NeuNAc-9-P phosphatase is responsible for the dephosphorylation of NeuNAc-9-P, which ultimately gives rise to free sialic acid (Varki, A. and Schauer, R. 2009).

CMAS, cytidine monophospho-N-acetylneuraminic acid synthetase, although not directly involved in sialic acid biosynthesis, is responsible for *activating* sialic acid by converting it to CMP-sialic acid for downstream use in the Golgi apparatus. Only CMP-sialic acid, not free sialic acid, is transported into the Golgi lumen. Also, GNE is allosterically inhibited by CMP-sialic acid, but not free sialic acid (Tanner, M.E. 2005), and thus, it is the cytosolic pool of CMP-sialic acid, not free sialic acid, that modulates the activity of GNE. Consequently, CMAS is integral to regulating the rate and extent of sialylation of a cell. Furthermore, CMAS is predominantly localised in the nucleus, which implies that the rate of transporting sialic acid into the nucleus and CMP-sialic acid out of it plays a role in the regulation of cellular sialylation (Münster, A.-K., Eckhardt, M., et al. 1998).

Finally, the transfer of sialic acid from CMP-sialic acid onto glycoconjugates transiting through the eukaryotic Golgi is catalysed by a family of linkage-specific glycosyltransferases known as sialyltransferases (Varki, A. and Schauer, R. 2009).

Sialic acid can be synthesised in all tissues to satisfy local sialylation requirements, and although mechanisms do exist for the cleavage of sialic acid from sialoglycoconjugates, the extent to which different eukaryotic cell types utilise exogenous and/or internally recycled sialic acid is unknown. At the level of the organism, free sialic acid circulating in the bloodstream (derived either from cellular sources or digestive processes in the gastrointestinal tract) is quickly excreted in the urine (Varki, A. and Schauer, R. 2009).

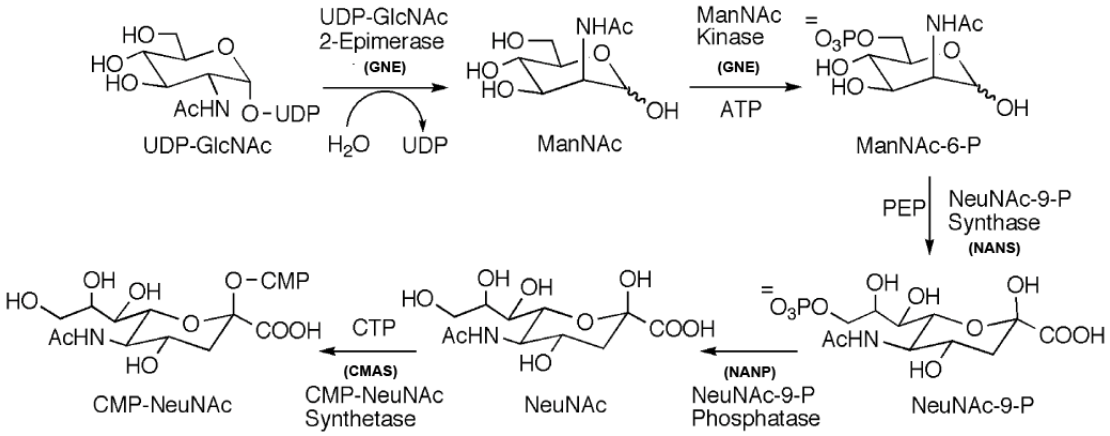


Figure 4: The mammalian sialic acid biosynthesis pathway. GNE, the bifunctional enzyme UDP-GlcNAc 2-epimerase/ManAc kinase, first catalyses the epimerisation of UDP-GlcNAc to ManNAc and subsequently phosphorylates ManNAc to ManNAc-6-P. NANS, N-acetylneuraminic acid synthase converts MaNAc-6-P to NeuNAc-9-P. NeuNAc-9-P phosphatise, NANP, then dephosphorylates NeuNAc-9-P to produce NeuNAc. Finally, NeuNAc is activated by cytidine monophospho-N-acetylneuraminic acid synthetase, CMAS, by conversion to CMP-NeuNAc [modified from Tanner (2005)].

1.4 Sialyltransferases

Sialyltransferases are a family of glycosyltransferases that transfer sialic acid from CMP-sialic acid to carbohydrate acceptor groups of glycoproteins, glycolipids and free oligosaccharides (Jeanneau, C., Chazalet, V., et al. 2004). Depending on the type of linkage formed and the type of sugar acceptor, sialyltransferases are traditionally separated into four groups: ST6Gal, ST6GalNAc, ST3Gal, ST8Sia (Tsuji, S. 1996b). There are one or more enzymes in each category and although some have broader acceptor specificities than others, the enzymes in each group only form one type of linkage between the sialic acid and the sugar acceptor (Harduin-Lepers, A., Mollicone, R., et al. 2005, Harduin-Lepers, A., Vallejo-Ruiz, V., et al. 2001).

Sialyltransferases are found in the Golgi apparatus and, like other Golgi glycosyltransferases, they have a characteristic type II architecture comprised of a short N-terminal cytoplasmic tail, a transmembrane domain followed by stem region and large C-terminal catalytic domain facing the Golgi lumen (Paulson, J.C. and Colley, K.J. 1989). The length and amino acid composition of sialyltransferase catalytic domains are reasonably conserved and the observed variations in overall protein size are usually a result of differences in the length of the stem region (Jeanneau, C., Chazalet, V., et al. 2004). Due to its highly variable amino acid sequence and apparent lack of secondary organisation, the stem region is thought to have an inherent flexibility (Donadio, S., Dubois, C., et al. 2003), although its specific function remains speculative. The presence of cysteine residues and several predicted N- and O-glycosylation sites suggests that a local conformation may exist and that it could be responsible for modulating *in vivo* acceptor specificity by exerting steric control over the catalytic domain (Legaigneur, P., Breton, C., et al. 2001).

The catalytic domains of all eukaryotic sialyltransferases contain four known conserved peptide sequences that have been named sialylmotif L (Long), S (Short), VS (Very Short) and motif III (Jeanneau, C., Chazalet, V., et al. 2004). The L-sialylmotif is primarily

involved in donor substrate binding (Datta, A.K. and Paulson, J.C. 1995), the S-sialylmotif has been demonstrated to have an effect on both donor and acceptor binding (Datta, A.K., Sinha, A., et al. 1998), whilst sialylmotif VS and motif III appear to be involved in the catalytic process (Geremia, R.A., HarduinLepers, A., et al. 1997, Jeanneau, C., Chazalet, V., et al. 2004).

Sialyltransferases are primarily regulated at the level of transcription. This transcriptional control is expedited through the use of several tissue and development-specific promoters, which give rise to numerous transcripts divergent only in their 5' untranslated regions (Svensson, E.C., Soreghan, B., et al. 1990, Takashima, S., Kono, M., et al. 2000, Wang, X., O'Hanlon, T.P., et al. 1990, Wen, D.X., Svensson, E.C., et al. 1992).

Thus far, 20 members of the murine sialyltransferase family have been cloned (Takashima, S. 2008), and a similar number of distinct human sialyltransferases have also been identified (Harduin-Lepers, A., Vallejo-Ruiz, V., et al. 2001), however, the structure of sialyltransferases has not been solved and their exact mechanism of action remains unclear. It was suggested by Takashima (2008), that due to the fundamentally important role of sialylglycoconjugates in numerous biological phenomena, and the inherent requirement for a variety of development and tissue-specific glycosylation patterns, that a precise control of this process necessitates the existence of multiple sialyltransferase genes that are co-ordinately regulated by a network of multiple promoters.

1.4.1 ST6Gal I

ST6Gal I is one of the most widely studied sialyltransferases. The sialyltransferase enzyme ST6Gal I transfers sialic acid from CMP-Sialic acid to either type 2 (Galβ1,4GlcNAc) free disaccharides or the *N*-acetyllactosamine termini of *N*- or *O*-linked oligosaccharides via an α2,6 linkage (Harduin-Lepers, A., Vallejo-Ruiz, V., et al. 2001). ST6Gal I is expressed in a variety of cell types and tissues (Aasheim, H.C., Aas-Eng, D.A., et al. 1993, Dalziel, M.,

Huang, R.Y., et al. 2001, Mercier, D., Wierinckx, A., et al. 1999, Wang, X., Vertino, A., et al. 1993), with particularly high levels of expression observed in the liver (Dalziel, M., Huang, R.Y., et al. 2001, Dalziel, M., Lemaire, S., et al. 1999). The enzyme has been successfully purified from rat, chicken and human liver and rat hepatoma cells (Bendiak, B. and Cook, G.M.W. 1982, Miyagi, T. and Tsuiki, S. 1982, Sticher, U., Gross, H.J., et al. 1991, Weinstein, J., de Souza-e-Silva, U., et al. 1982a, Weinstein, J., de Souza-e-Silva, U., et al. 1982b). A soluble form of ST6Gal I was also purified from bovine colostrum (Paulson, J.C., Beranek, W.E., et al. 1977), and was established to be the product of proteolytic cleavage of the Golgi membrane-bound enzyme.

Mercier et al. (1999) identified and extensively characterised the bovine ST6Gal I gene, which has a coding region of 1218 nucleotides. The coding sequence consists of 5 exons, the last of which contains a 3'UTR of 2.7kb. Four additional exons were identified in the 5' UTR, resulting in a total of 9 exons spanning up to 80kb of genomic DNA. The gene comprises a single open reading frame encoding a protein of 405 residues with a theoretical molecular weight of 46250Da. Structural feature predictions indicate that the enzyme is likely to have a type II membrane topology, which is typical of Golgi glycosyltransferases. The bovine ST6Gal I protein sequence exhibits a higher homology with the human enzyme than with the rat, mouse or chicken enzymes (Mercier, D., Wierinckx, A., et al. 1999).

The human ST6Gal I gene shares similarities with its bovine orthologue. It is located on human chromosome 3 and also encodes a 405 residue protein across 5 exons, which span in excess of 40kb of genomic DNA. Four additional exons have been discovered in the 5'UTR (Wang, X., Vertino, A., et al. 1993).

Transcriptional regulation of the ST6Gal I gene is coordinated by the use of several cell-specific promoters that produce transcripts with divergent 5' untranslated regions (Svensson, E.C., Soreghan, B., et al. 1990, Wang, X., O'Hanlon, T.P., et al. 1990, Wen,

D.X., Svensson, E.C., et al. 1992). The rat ST6Gal I gene, which has been extensively studied, is known to span over 80kb and has a complex network of at least four promoters that regulate its expression (Wen, D.X., Svensson, E.C., et al. 1992).

In vivo, three main families of bovine ST6Gal I transcripts have been characterised, leading to speculation of the existence of at least 3 separate promoters regulating the expression of the gene. Despite organisation of the bovine gene being similar to both human (Wang, X., Vertino, A., et al. 1993) and rat (Svensson, E.C., Soreghan, B., et al. 1990, Wang, X., O'Hanlon, T.P., et al. 1990, Wen, D.X., Svensson, E.C., et al. 1992), there are differences in the tissue distribution of its transcripts (Mercier, D., Wierinckx, A., et al. 1999).

The mouse ST6Gal I gene is known to contain at least 8 5'UTR exons in addition to exons (I)-(VI) (Dalziel, M., Huang, R.Y., et al. 2001, Dalziel, M., Lemaire, S., et al. 1999, Wuensch, S.A., Huang, R.Y., et al. 2000). Several studies have demonstrated that, despite not being translated, ST6Gal I 5'UTR exons exhibit a degree of nucleotide sequence conservation between species, with exons (I)-(VI) typically being well conserved (Aasheim, H.C., Aas-Eng, D.A., et al. 1993, Mercier, D., Wierinckx, A., et al. 1999, Wang, X., Vertino, A., et al. 1993).

The existence of a novel, lactogenic ST6Gal I promoter was proposed by the authors of a study examining the dramatic induction of ST6Gal I mRNA in the lactating mammary gland of mice. The predominant ST6Gal I mRNA species expressed in the lactating mouse mammary gland is a novel isoform containing a unique untranslated exon derived from the ST6Gal I 5'UTR, dubbed exon (L). Exon (L) is 203bp long and is located more than 48kb 5' of exon (I). Virgin mammary tissue exclusively expressed the housekeeping ST6Gal I mRNA form and probes specific to exon (L) confirmed that the exon (L)-containing transcript was present only in the lactating mammary gland and not in other tissues that were investigated (Dalziel, M., Huang, R.Y., et al. 2001).

1.5 Human milk oligosaccharides

Human milk is an extremely rich source of essential nutrients for the newborn human infant (Hamosh, M. 1999, Picciano, M.F. 1998). In addition to proteins, lipids, vitamins, minerals and antibodies, human milk provides the newborn with a suite of over 200 oligosaccharides (Ninonuevo, M.R., Park, Y., et al. 2006). After lactose and lipids, these complex sugars are the third major milk solute and are present in greater overall quantities than protein (McVeagh, P. and Brand-Miller, J. 1997). Oligosaccharides occur at concentrations of between 20 and 23g/L in human colostrum (Coppa, G.V., Pierani, P., et al. 1999) and 7 to 12g/L in mature human milk (Boehm, G. and Stahl, B. 2007). Despite this unique diversity and abundance of oligosaccharide structures (McVeagh, P. and Brand-Miller, J. 1997), their role in the human infant is poorly understood.

Human milk oligosaccharides are not digested or absorbed in the small intestine of infants (Brand-Miller, J.C., McVeagh, P., et al. 1995). Notably, many human milk oligosaccharide structures resemble the carbohydrate chains on glycoproteins or glycolipids anchored onto mucosal cell-surface membranes (Newburg, D.S. 2000). As such, numerous studies have shown that milk oligosaccharides act as decoy receptors mimicking epithelial cell-surface carbohydrates, thus protecting the gastrointestinal tract of the infant by competing with the epithelial cell ligand for pathogens (Coppa, G.V., Gabrielli, O., et al. 1990, Martin-Sosa, S., Martin, M.J., et al. 2003, Wang, B., Brand-Miller, J., et al. 2001). Thus far, more than 20 specific milk oligosaccharides are known to competitively bind to pathogens of the intestinal, respiratory and urinary tract, affecting a variety of bacteria, viruses and yeast (McVeagh, P. and Brand-Miller, J. 1997).

It has also been demonstrated that human milk oligosaccharides influence the development of the infant gastrointestinal tract, as well as the establishment of intestinal microbiota (Frank, D.N. and Pace, N.R. 2008, Savage, D.C. 1977). It has been proposed that milk oligosaccharides encourage the colonisation and proliferation of particular groups of

bacteria by acting as selective nutrients (Gibson, G.R. and Roberfroid, M.B. 1995, LoCascio, R.G., Ninonuevo, M.R., et al. 2007, Marcobal, A., Barboza, M., et al. 2010). Sela et al. (2011) showed that certain bifidobacteria can metabolise sialylated oligosaccharides, possibly contributing to their ability to establish in the infant gut. By comparing the intestinal flora of breastfed and formula-fed infants, Harmsen et al. (2000) also demonstrated the prebiotic effects of milk oligosaccharides on the gastrointestinal tract of the newborns.

The establishment of intestinal flora is essential for the proper development of the mucosal immune system. Specifically, bacterial colonisation influences morphological development of immune compartments, instigates the production of antibodies (Macpherson, A.J. and Harris, N.L. 2004) and the production of various antimicrobial proteins such as defensins (Falk, P.G., Hooper, L.V., et al. 1998). Thus, by promoting the growth of beneficial colonic flora, milk oligosaccharides also indirectly protect the breastfed infant from gastrointestinal infection (McVeagh, P. and Brand-Miller, J. 1997). Newburg (2000) noted several studies which demonstrated that breastfeeding is highly effective at protecting infants against enteric disease, particularly in the first 3-6 months of life, whilst the intestinal environment is still relatively immature.

The most abundant oligosaccharide structures are trisaccharides, usually comprised of lactose with an attached fucose or sialic acid. Fucosylated oligosaccharides are not present in the milk of most mammalian species, however, sialylated oligosaccharides are widely represented (Urashima, T., Saito, T., et al. 2001). Up to 20% of the oligosaccharide structures identified in human milk are sialylated (Ninonuevo, M.R., Park, Y., et al. 2006, Tao, N., Wu, S., et al. 2011). Sialyllactose is the best represented moiety and is found at concentrations of ~1g/L (Kunz, C. and Rudloff, S. 1993). Overall sialic acid content is highest in early lactation and decreases over time (Viverge, D. 1990). Due to their relative abundance and potentially varied roles, sialylated milk oligosaccharides are of particular interest to researchers. Sialylated milk oligosaccharides can prevent the attachment of

E.coli strains associated with neonatal meningitis and sepsis; it has been shown that these organisms are less common in the faeces of breastfed infants than formula fed infants (McVeagh, P. and Brand-Miller, J. 1997, Nakano, T., Sugawara, M., et al. 2001). As well as playing a crucial role in the protection and development of the gastrointestinal tract, it has been suggested that sialylated milk oligosaccharides facilitate post-natal brain development of the neonate (Wang, B. and Brand-Miller, J. 2003). Dietary sialic acid has been shown to enhance learning in developing piglets (Moughan, P.J., Birtles, M.J., et al. 1992, Wang, B., Yu, B., et al. 2007).

1.6 Sialic acid in the nervous system

The mammalian nervous system contains greater concentrations of sialic acid than any other organ. Within the nervous system, 65% of all sialic acid is used in the formation of gangliosides, 32% is attached to glycoproteins and only 3% is in the free form (Brunngraber, E.G., Witting, L.A., et al. 1972, Wang, B. and Brand-Miller, J. 2003). Sialic acid confers a negative charge to the oligosaccharide head group that is bound to the hydrophobic ceramide portion of gangliosides (Rosenburg, A. 1995). Gangliosides are particularly concentrated in the plasma membrane of nerve cells, especially in the vicinity of nerve endings and dendrites (McVeagh, P. and Brand-Miller, J. 1997). Human cerebral and cerebellar grey matter contain higher concentrations of gangliosides than any other tissue (Svennerholm, L., Bostrom, K., et al. 1989).

The sialic acid content of neural cell membranes is 20 times higher than in any other type of membrane, suggesting an important role for sialic acid in neural structure (Schauer, R. 1982). Furthermore, it has been proposed that ganglioside-bound sialic acid may be involved in processes such as synaptic transmission and neuronal adaptations required for storage of information, whilst sialylated glycoproteins have been implicated in learning and memory. It has also been hypothesised that sialic acid may be the receptor for

neurotransmitters in the central nervous system (Von Itzstein, M. and Thomson, R. 1997, Wang, B. and Brand-Miller, J. 2003).

According to Wang and Brand-Miller (2003), the human brain has 2-4 times greater sialic acid concentration than that of rats, mice, rabbits, sheep, cows and pigs. They go on to suggest that brain ganglioside sialic acid content has "implications in evolutionary development and intellectual capacity due to the structural and functional role of the sialic acid moieties of gangliosides and glycoprotein in the frontal cortex, and their involvement in cellular events such as cell recognition, cell-to-cell contact formation, receptor binding and modulation, immunological properties and biosignal transduction" (Wang, B. and Brand-Miller, J. 2003).

1.7 Sialic acid and lactation

1.7.1 Sialic acid, human milk and infant formula

Breastfeeding at infancy has been associated with higher later intelligence quotients in several studies (Fergusson, D.M., Beautrais, A.L., et al. 1982, Lucas, A., Morley, R., et al. 1998, Lucas, A., Morley, R., et al. 1992, Rodgers, B. 1978), and is thought to be particularly beneficial for the neural development of premature infants (Lucas, A., Morley, R., et al. 1998, Lucas, A., Morley, R., et al. 1992). Although there are many factors in milk that may contribute to cognitive development, McVeagh and Brand-Miller (1997) suggest that due to its abundance and pivotal role in the nervous system, sialic acid derived from milk oligosaccharides could be a conditional nutrient, in that breastfed infants are advantaged in having a nutritional source to supplement endogenous production.

Several animal studies have demonstrated that exogenously administered sialic acid can, to varying degrees, be incorporated into tissues (Carlson, S.E. 1985, Carlson, S.E. and House, S.G. 1986, Morgan, B.L. and Winick, M. 1981). However, the mechanism of sialic acid

cleavage and subsequent absorption across the intestinal mucosa remains unidentified (Wang, B. and Brand-Miller, J. 2003). It has also been demonstrated, in rats and guinea pigs, that, in comparison to adult levels, the capacity of the liver to synthesise sialic acid is diminished in neonates (Gal, B., Ruano, M.J., et al. 1997, Nakano, T., Sugawara, M., et al. 2001).

In humans, there is a significant burst of neural growth and development between the third trimester and 2 years post-partum (Dobbing, J. 1981). Also, studies in rats have shown that there is a major period of ganglioside accumulation during the first 3 weeks after birth (Vanier, M.T., Holm, M., et al. 1971). Interestingly, sialic acid concentration in human milk is highest in colostrum and decreases by more than 70% over the subsequent 3 months (Wang, B., Brand-Miller, J., et al. 2001). Wang and Brand-Miller et al. (2001) found that the sialic acid content of the commercial infant formulas they investigated had less than 25% of that found in mature human milk. They also noted that most of the sialic acid in the formulas (~70%) was bound to glycoproteins, whereas the majority of human milk sialic acid (69-76%) is bound to free oligosaccharides, which brings into question the bioavailability of the sialic acid present in infant formula.

Besides lactose, sialyllactose is the next most abundant carbohydrate in rat milk (Witt, W., von Nicolai, H., et al. 1979). Sialyloligosaccharides account for 3.3 g/L of human colostrum and decrease to 1.58g/L at 3 months of lactation. Sialyllactose comprises about a quarter of the colostrum sialyloligosaccharide content, with α2,6-sialyllactose at 0.59g/L and α2,3 sialyllactose at ~0.09g/L. α2,6 sialyllactose decreases to 0.24g/L by day 90 of lactation but the level of α2,3 sialyllactose level remains fairly constant throughout (Coppa, G.V., Pierani, P., et al. 1999). By contrast, bovine colostrum only contains 0.7-1.2 g/L sialyloligosaccharides, with 0.037g/L α2,6 sialyllactose and 0.33 g/L α2,3 sialyllactose (Veh, R.W., Michalski, J.-C., et al. 1981).

Sialyllactose has been shown to reduce gastrointestinal fluid accumulation induced by cholera toxin and to promote the growth of *Bifidobacterium breve* and *Bifidobacterium bifidum* in a dose dependent manner (Nakano, T., Sugawara, M., et al. 2001). Studies have also reported retention of radiolabelled sialic acid in the organs of suckling rats and mice following oral administration of radiolabelled sialyllactose (Nohle, U. and Schauer, R. 1984, Witt, W., von Nicolai, H., et al. 1979). Consequently, many authors have suggested that sialyllactose may play an important role in the protection of the gastrointestinal tract of the human infant and that it perhaps also serves as a delivery system for a large proportion of the sialic acid consumed through breastfeeding.

1.7.2 Sialic acid in mouse and rat milk

In addition to lactose, rat milk only contains a limited array of oligosaccharide structures, consisting primarily of sialyllactose and trace amounts of fucosylated lactose (Kuhn, N.J. 1972, Prieto, P.A., Mukerji, P., et al. 1995). Mouse milk oligosaccharides are also dominated by sialyllactose; the sialyllactose fraction of mouse milk is comprised of α2,3 and α2,6-sialyllactose, whilst the fucosylated fraction only includes α2,6-fucosyllactose (Prieto, P.A., Mukerji, P., et al. 1995).

The lactose content of rat milk increases threefold during lactation (Kuhn, N.J. 1972). The ultimate lactose concentration approaches the carbohydrate content of the adult rat diet, and has been suggested to serve an adaptive purpose (Kuhn, N.J. 1972). Sialyllactose is also a relatively large component of rat milk. By the fourth day of lactation, it comprises over 50% by weight of the total milk sugar, although it is only present in trace amounts by the end of lactation. In the early stages of lactation, the presence of large amounts of sialyllactose substantially raises the carbohydrate content of rat milk without undue osmotic contribution, making it a potentially efficient delivery mechanism of both energy and sialic acid (Kuhn, N.J. 1972). The lower small intestine of the suckling rat absorbs

milk immunoglobulins and other macromolecules by pinocytosis and is known to contain a very active, possibly lysosomal, sialidase. It has been suggested (Dickson, J.J. and Messer, M. 1978) that the sialyllactose found in rat milk is passively transported into the enterocytes along with other macromolecules and is subsequently digested by the lysosomal sialidase, releasing the sialic acid.

Duncan et al. (2009) found that α2,3-sialyllactose was the major oligosaccharide present in milk in early rat lactation. They found that it increased in abundance during the first week of lactation but afterwards declined steadily. α2,6-sialyllactose was found at approximately ten times lower levels than α2,3-sialyllactose and also increased in abundance throughout the first week of lactation. It then remained at a constant, albeit low level until weaning. The total amount of sialic acid in rat milk was found to be similar to the profile of α2,3-sialyllactose throughout lactation.

Duncan at al. (2009) also investigated the relationship between the sialic acid content of rat milk and the uptake, utilisation and synthesis of sialic acid in suckling pups. Sialic acid synthesis in the liver of the suckling pups was found to increase in parallel to milk sialic acid content, achieving peak activity at 5 days post-partum and maintaining it thereafter. Gene expression profiles in the neonatal colon indicated that there was a switch from sialic acid uptake and catabolism to sialic acid synthesis and utilisation that mirrored the change in sialic acid milk content from high to low abundance. These results indicate that when milk sialic acid content is high, this sialic acid is catabolised to GlcNAc in the colon, which can then be used as such or as substrate for sialic acid synthesis, and when milk sialic acid content is low the endogenous sialic acid synthesis pathway is activated in the colon (Duncan, P.I., Raymond, F., et al. 2009).

More recently, it was demonstrated by Fuhrer et al. (2010) that the sialyltransferase ST6Gal I is responsible for the synthesis of α2,6-sialyllactose in mouse milk and that ST3Gal IV is responsible for the bulk of α2,3-sialyllactose production. The limited suite of

oligosaccharide structures present in mouse milk and their relatively short gestation and lactation period make the mouse a suitable model to investigate the production of sialylated milk oligosaccharides in the mammary gland and their possible role in the suckling neonate.

1.7.3 Sialic acid in marsupial milk

Placental mammals and marsupials differ primarily in their reproductive strategies. Marsupials generally have a very short gestation and give birth to a very altricial (small and undeveloped) young. Marsupial lactation is also generally relatively long, usually much longer than for placental mammals of comparable size (Renfree, M.B. 1983, Tyndale-Biscoe, C.H. and Janssens, P.A. 1988).

The carbohydrate composition of marsupial milk is also unique in comparison to the milk of other mammals. Unlike the milk of eutherian mammals where lactose is generally the dominant carbohydrate (Jenness, R., Regehr, E.A., et al. 1964), marsupial milk contains various oligosaccharides of which free lactose is only a relatively minor component (Green, B. and Merchant, J.C. 1988). Tammar wallaby (*Macropus eugenii*) milk contains a variety of neutral oligosaccharides; a trisaccharide, β1-3,galactosyllactose (Gal(β1-3)Gal(β1-4)Glc) (Messer, M., Trifonoff, E., et al. 1980), and several homologous tetra- to hepta-saccharides (Collins, J.G., Bradbury, J.H., et al. 1981). There is also another suite of neutral oligosaccharides containing GlcNAc (Bradbury, J.H., Collins, J.G., et al. 1983, Messer, M. and Trifonoff, E. 1982). Tammar wallaby milk also contains acidic oligosaccharides that include sialic acid and a β(1–3) galactosyl sequence with two units of N-acetyllactosaminyl branches (Urashima, T., Saito, T., et al. 1994). Sialic acid is a significant component of tammar wallaby milk overall (Table 1), accounting for ~4-5% of total carbohydrate throughout most of lactation, and rising to ~9% just prior to weaning (Green, B. and Merchant, J.C. 1988).

The carbohydrate content of tammar wallaby milk changes quantitatively throughout lactation. Lactose dominates during the first week of lactation, after which it is rapidly replaced by tri- to penta-saccharides and by day 182 the milk sugars are almost exclusively higher order oligosaccharides (Green, B. and Merchant, J.C. 1988). The carbohydrates then become progressively smaller until about day 280 when all sugars are monosaccharides and overall carbohydrate concentration is markedly decreased (Green, B. and Merchant, J.C. 1988). The concentration of oligosaccharides in the milk of the tammar wallaby is also very high during most of lactation [up to 13%(w/v)] when compared with that of placental mammals (Messer, M. and Green, B. 1979, Messer, M. and Mossop, G. 1977). The evidence available to date strongly indicates that the high concentration and variety of oligosaccharides, along with the low concentration of free lactose found in tammar wallaby milk, is typical of all marsupials (Urashima, T., Saito, T., et al. 1994).

Oligosaccharides are a major milk component in species such as marsupials, monotremes, and bears (Messer, M. and Kerry, K.R. 1973, Messer, M. and Nicholas, K.R. 1991, Urashima, T., Kusaka, Y., et al. 1997, Urashima, T., Sumiyoshi, W., et al. 1999, Urashima, T., Yamashita, T., et al. 2000). In such cases it is believed that oligosaccharides must serve as energy sources, as well as delivering important monosaccharides such as sialic acid (Urashima, T., Saito, T., et al. 2001). Given that compared to other mammals, the presence of lactose is significantly diminished in the milk of marsupials, it is reasonable to make such an assumption. The tammar wallaby, like all marsupials, is extremely altricial at birth, and it has been suggested that altricial neonates have a particular requirement for milk oligosaccharides (Oftedal, O.T., Alt, G.L., et al. 1993). Furthermore, it is thought that this requirement may not solely be limited to species whose young are very small at birth compared to the size of the mother, but may specifically relate to neurological immaturity (Kunz, C., Rudloff, S., et al. 1999, Urashima, T., Saito, T., et al. 2001).

Evidence suggests that marsupials, such as the tammar wallaby, are able to digest milk oligosaccharides by a process that involves their transport into the enterocytes of the small

intestine by pinocytosis and subsequent hydrolysis to monosaccharides by lysosomal enzymes such as acid β-galactosidase, α-galactosidase, α-fucosidase, sialidase and β-N-acetylglucosaminidase (Messer, M., Crisp, E.A., et al. 1989, Urashima, T., Saito, T., et al. 2001). Given the relatively low concentration of lactose in tammar wallaby milk after the first week of lactation, it is not surprising that histochemical and biochemical evidence has demonstrated an absence of lactase activity in the brush border of the small intestine of suckling tammars and other marsupials (Crisp, E.A., Czolij, R., et al. 1987). However, a very active acid β-galactosidase was discovered in the lysosomes and the supranuclear vacuoles of the enterocytes (Walcott, P.J. and Messer, M. 1980). This enzyme was able to digest the β-(1–3) galactosides present in tammar wallaby milk. Sialidase activity has also been observed in the intestine of suckling tammars (Dickson, J.J. and Messer, M. 1978, Walcott, P.J. and Messer, M. 1980). A study by Dickson and Messer (1978) highlighted that there was a significant positive correlation between the intestinal sialidase activity of the suckling neonate of a variety of mammals and the sialic acid content of milk obtained from the same species and stage of lactation. They suggested that the primary function of intestinal sialidase of suckling mammals may be to cleave sialic acid from milk components, thereby supplying sialic acid for the synthesis of sialylglycoconjugates by the neonate. Walcott and Messer (1980) later profiled the intestinal sialidase activity of the tammar wallaby pouch young throughout the lactation cycle. They found that, following an initial increase in activity around day 70, the greatest level of intestinal sialidase activity occurred between 140 and 280 days after birth, which coincides with the period when virtually all tammar wallaby milk carbohydrates are present as higher oligosaccharides.

These results support the hypothesis that in suckling marsupials intact milk oligosaccharides are transported into the enterocytes of the small intestine, most probably by either pinocytosis or endocytosis, and are subsequently hydrolysed to their monosaccharide components by acid β-galactosidase and other lysosomal enzymes such as sialidase and N-acetylglucosaminidase.

Table 1

Monosaccharide composition of tammar wallaby milk carbohydrates, throughout the lactation cycle

Day	Galactose	Glucose	Glucosamine	Galactosamine	Sialic acid
30	71	20	4.0	< 1	5.5
130	72	19	3.8	< 1	4.6
180	80	12	3.7	< 1	4.0
220	68	20	7.4	< 1	3.6
300	51	21	14.0	5.1	9.1

Results are presented as a percentage of total carbohydrate (w/w) [adapted from Green and Merchant (1988)].

1.8 Rationale for study

This introduction has outlined what sialic acid is, how and where it is produced in the body and its importance as a component of sialylated milk oligosaccharides and sialylglycoconjugates. It has briefly reviewed the nature of oligosaccharides in human milk and their likely role in the development and protection of the infant gastrointestinal system, as well as highlighting the role of sialic acid as a component of the central nervous system and its potential as a conditional nutrient involved in infant brain development. Given the difficulty of studying these phenomena in the human, this introduction also covered the sialyloligosaccharide distribution of mouse milk and the relevant studies that have used the mouse as a model organism to investigate sialylation and its role in milk in the neonate. The tammar wallaby was introduced as an alternative model organism for the study of sialylated milk oligosaccharides, given their extreme altricial nature at birth and total dependence on milk and milk oligosaccharides as a source of nutrition and energy.

Prieto et al. (1995) demonstrated that transgenic mice expressing human α1,2-fucosyltransferase under the control of a lactogenic promoter, produced large amounts of 2'-fucosyllactose and modified glycoproteins containing the H-antigen. The authors had considered the possibility that the very low levels of fucosylated oligosaccharides otherwise observed in mouse milk may be due to factors unrelated to gene expression such as low levels of the GDP-fucose substrate. However, the large amounts of 2'-fucosyllactose produced by the transgenic mice suggest that the Golgi apparatus of the lactating mammary gland was sufficiently adaptable with regard to factors such as nucleotide availability, lactose availability, enzyme localization and enzyme secretion. It also indicates that the transcriptional regulation of sialic acid biosynthesis and sialyltransferase genes in the mammary gland is likely to be the key factor in determining the nature and extent of sialylated structures found in milk. If this is indeed the case, it has important implications in the field of infant nutrition where there is an ongoing struggle to create the best approximation of breast milk when manufacturing infant formula.

Over 200 unique oligosaccharide structures have been identified in human milk, with up to 20% being sialylated (Ninonuevo, M.R., Park, Y., et al. 2006, Tao, N., Wu, S., et al. 2011). Human milk contains significantly higher levels of sialic acid than any type of infant formula, many of which are produced using bovine milk (Wang, B. and Brand-Miller, J. 2003). Only 40 oligosaccharides have been identified in bovine milk, all of which are generally composed of shorter oligomeric chains than those found in human milk. Also, the overall concentration of oligosaccharides in bovine milk is significantly lower than in human milk (Tao, N., DePeters, E., et al. 2008). Approximately 70% of the oligosaccharides in bovine milk are sialylated, with sialic acid residues found to be both Neu5Ac and Neu5Gc, although Neu5Ac is significantly more abundant (Tao, N., DePeters, E., et al. 2008). α2,3-sialyllactose is the dominant oligosaccharide in bovine colostrum, although it gradually decreases in abundance during lactation. Overall sialylation of bovine oligosaccharides decreases from 70% in colostrum to 50% in mature milk (Tao, N., DePeters, E., et al. 2009).

Due to the relatively low amounts and dissimilar distribution of sialyloligosaccharides in bovine milk from which most infant formula is made, the infant formula industry had attempted to increase the sialyloligosaccharide content of its products by concentrating sialyllactose in bovine milk. This, however, has proven to be uneconomical (Murray Goulbourn Co-Operative Co. Limited, 2007, Personal Communication). Consequently, the ability to increase the levels of sialylated oligosaccharides produced by the bovine mammary gland is an intriguing prospect. In order to achieve this, it is essential to have a better understanding of the expression and regulation of sialic acid biosynthesis and sialyltransferase genes in the mammary gland.

It is already well documented that sialyltransferase genes are primarily regulated at the level of transcription via a complex network of multiple tissue and development-specific promoters, which give rise to numerous transcripts divergent only in their 5' untranslated regions (Svensson, E.C., Soreghan, B., et al. 1990, Takashima, S., Kono, M., et al. 2000,

Wang, X., O'Hanlon, T.P., et al. 1990, Wen, D.X., Svensson, E.C., et al. 1992). Fuhrer et al. (2010) investigated the expression of α2,3 sialyltransferase and α2,6, sialyltransferase genes in the mouse mammary gland and found that only ST6Gal I, ST3Gal I and ST3Gal IV increased in expression during lactation. Using ST3Gal I -/- and ST3Gal IV -/- knockout mice, and measuring the 3'-sialyllactose levels of their milk throughout lactation, they were able to determine that ST3Gal IV was responsible for the majority of 3'-sialyllactose production in the mouse mammary gland. The existence of a lactogenic ST6Gal I promoter was previously proposed by the authors of a study examining the dramatic induction of ST6Gal I mRNA in the lactating mammary gland of mice. The predominant ST6Gal I mRNA species expressed in the lactating mouse mammary gland is a novel isoform containing a unique untranslated exon derived from the ST6Gal I 5'UTR, dubbed exon (L). Virgin mammary tissue exclusively expressed the housekeeping ST6Gal I mRNA form and probes specific to exon (L) confirmed that the exon (L)-containing transcript was present in the lactating mammary gland and not in other tissues that were investigated (Dalziel, M., Huang, R.Y., et al. 2001). These studies were an important initial step in investigating the expression and regulation of sialyltransferase genes in the mammary gland. However, further work is required to elucidate the full suite of sialyltransferase genes expressed in the mammary gland during lactation and whether or not lactogenic hormones do indeed play a part in driving their expression. In addition, there has been no study examining the expression of sialic acid biosynthesis genes in the mammary gland. Given that most of the sialic acid used by any one tissue is produced locally (Varki, A. and Schauer, R. 2009), it follows that the expression of sialic acid biosynthesis genes in the mammary gland is an important part of ensuring the adequate supply of CMP-sialic acid for downstream sialylation. Keppler et al. (1999) demonstrated that GNE activity can be controlled at the level of transcription and that this can influence the sialylation and function of particular cell-surface molecules expressed on B-cells and myeloid cells.

Given its relatively limited suite of milk oligosaccharide structures and their relatively short gestation and lactation period, the mouse has been extensively used as a model to investigate the production of sialylated milk oligosaccharides in the mammary gland and their possible role in the suckling neonate. Recently, Fuhrer et al. (2010) successfully demonstrated that exposure to sialyllactose during infancy affected bacterial colonisation of the intestine of mice, which influenced their susceptibility to dextran sulphate sodium-induced colitis in adulthood. Although the use of mice as models for the study of sialylated milk oligosaccharides has been extensive, this species is not a particularly useful model for increasing our understanding of sialylation in other species. Their limited suite of primarily trisaccharides is far from the extensive array of high order oligosaccharide structures present in human milk. In this respect, a comparative approach using a species such as the tammar wallaby may provide additional insight as their milk includes a comprehensive suite of complex oligosaccharides. Furthermore, the extremely altricial nature of the tammar wallaby young at birth, and subsequent long lactation, provide a unique opportunity for the correlation of expression of sialic acid biosynthesis and sialyltransferase genes in the tammar wallaby mammary gland with developmental state of the suckled young.

1.9 Aims

This introduction provides an indication of the importance of sialic acid as a milk component. This thesis aims to investigate the role that transcriptional regulation of sialic acid biosynthesis and sialyltransferase genes by lactogenic hormones plays in determining the sialylated structures produced by the mammary gland during lactation. The nature of this role will be examined via a series of ancillary aims over the subsequent four chapters. Each individual aim focuses on addressing a specific aspect of the general aim. The ancillary aims are:

Chapter 2: To determine if there exists a bovine and/or human equivalent to mouse ST6Gal I exon (L), and if these species also express the ST6Gal I gene in a lactation specific manner.

Chapter 3: To determine which sialic acid biosynthesis and sialyltransferase genes are differentially expressed in the mouse mammary gland during lactation.

Chapter 4: To determine whether key lactogenic hormones are involved in regulating the expression of sialic acid biosynthesis and sialyltransferase genes that were found to be significantly differentially expressed in Chapter 3.

Chapter 5: To determine if any sialic acid biosynthesis and sialyltransferase genes found to be differentially expressed in Chapter 3 are also differentially expressed in the lactating tammar wallaby mammary gland, and to infer possible roles for their sialylated end-products based on the neonate's stage of development.

1.10 Hypothesis

The overarching hypothesis for this study is:

The complement of sialylated oligosaccharides and sialylglycoconjugates present in milk is largely determined by sialyltransferase gene expression in the mammary gland, which is controlled at the transcriptional level by the lactogenic hormones insulin, hydrocortisone and prolactin.

2 Chapter two

The mouse ST6Gal I gene is known to contain at least 8 5'UTR exons in addition to exons (I)-(VI) (Dalziel, M., Huang, R.Y., et al. 2001, Dalziel, M., Lemaire, S., et al. 1999, Wuensch, S.A., Huang, R.Y., et al. 2000). Several studies have demonstrated that, despite not being translated, ST6Gal I 5'UTR exons exhibit a degree of nucleotide sequence conservation between species, with exons (I)-(VI) typically being well conserved (Aasheim, H.C., Aas-Eng, D.A., et al. 1993, Mercier, D., Wierinckx, A., et al. 1999, Wang, X., Vertino, A., et al. 1993).

The existence of a novel, lactogenic ST6Gal I promoter was proposed by the authors of a study examining the dramatic induction of ST6Gal I mRNA in the lactating mammary gland of mice. The predominant ST6Gal I mRNA species expressed in the lactating mouse mammary gland is a novel isoform containing a unique untranslated exon derived from the ST6Gal I 5'UTR, dubbed exon (L). Exon (L) is 203bp long and is located more than 48kb 5' of exon (I). Virgin mammary tissue exclusively expressed the housekeeping ST6Gal I mRNA form and probes specific to exon (L) confirmed that the exon (L)-containing transcript was present only in the lactating mammary gland and not in other tissues that were investigated (Dalziel, M., Huang, R.Y., et al. 2001).

The presence of a lactation-specific ST6Ggal I transcript and the dramatic increase in gene expression observed during mouse lactation by Dalziel et al. (2001) suggested that α2,6-sialylated milk components may be particularly important for neonate development. Consequently, the conservation of exon (l) between several mammalian species was explored and its expression in bovine and human mammary tissue was experimentally investigated. The expression of the bovine ST6Gal I gene was also profiled in pregnancy, lactation and involution.

This chapter was published in Glycobiology 2011 Apr;21(4):467-81. Epub 2010 nov 22.

2.1 Declaration for thesis chapter two

2.1.1 Declaration by candidate

In the case of Chapter 2, the nature and extent of my contribution to the work was the following:

Nature of contribution	**Extent of contribution (%)**
For this chapter I preformed all of the laboratory experiments and was responsible for the interpretation and analysis of the data and the writing of the manuscript. Dr. Sharp initially collected and processed the human milk samples.	80%

The following co-authors contributed to the work. Co-authors who are students at Monash University must also indicate the extent of their contribution in percentage terms:

Name	**Nature of contribution**	**Extent of contribution (%)** (student co-authors only)
Dr. Julie A. Sharp	Dr. Sharp collected and processed the human milk samples and aided in experimental design, interpretation of results and writing of the manuscript.	
Prof. Kevin R. Nicholas	Prof. Nicholas aided in interpretation of results and writing of the manuscript.	
Dr. Benjamin G. Cocks	Dr. Cocks was involved in developing the initial concept for the study and aided in interpretation of results and writing of the manuscript.	
Dr. Keith W. Savin	Dr. Savin aided in interpretation of results and writing of the manuscript.	

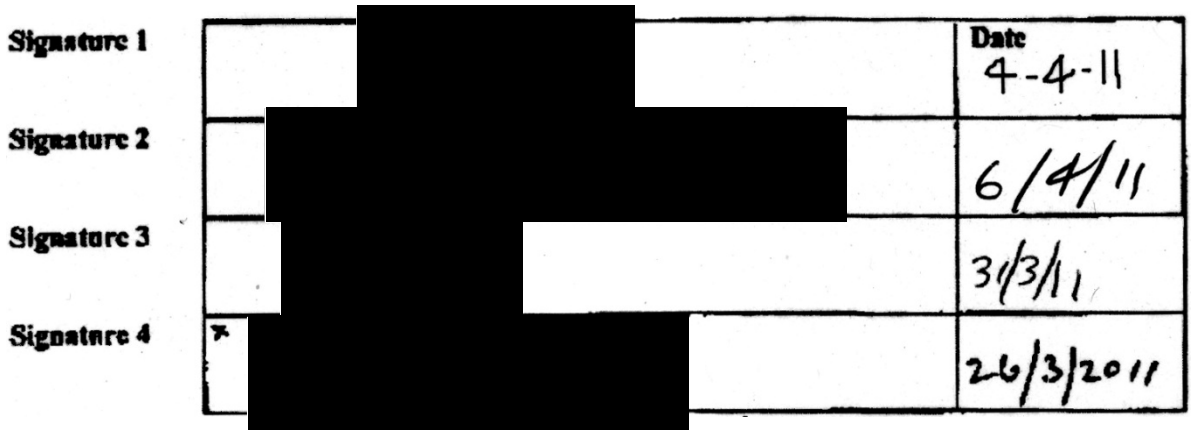
Candidate's Signature		**Date**

2.1.2 Declaration by co-authors

The undersigned hereby certify that:

(1) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.
(2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
(3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
(4) there are no other authors of the publication according to these criteria;
(5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
(6) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

Location(s)	**DPI, Biosciences Research Division, Bundoora VIC 3083**



2.2 Conservation of the ST6Gal I gene and its expression in the mammary gland

Jovana Maksimovic[1,2], Julie A. Sharp[3], Kevin R. Nicholas[3], Benjamin G. Cocks[1], Keith Savin[1]

1. Biosciences Research Division, Department of Primary Industries, Bundoora 3083, Australia
2. Centre for Reproduction and Development, Monash Institute of Medical Research, Monash University, Clayton 3168, Australia
3. Institute for Technology Research and Innovation, Deakin University, Geelong 3214, Australia

2.2.1 Abstract

Milk sialoglycoconjugates can protect the gastrointestinal tract of the suckling neonate by competitively binding to invading pathogens and promoting growth of beneficial flora, and their potential role in post-natal brain development is of particular interest in human infant nutrition. Although the concentration and distribution of sialoglycoconjugates has been extensively studied in the milk of various species, the investigation of sialyltransferase gene expression in the mammary gland, in the context of lactation, has been limited. The sialyltransferase enzyme ST6Gal I transfers sialic acid from CMP-sialic acid to type 2 (Galβ1,4GlcNAc) free disaccharides or the termini of *N*- or *O*-linked oligosaccharides using an α2,6 linkage. Expression of the ST6Gal I gene is primarily regulated at the level of transcription through the use of several cell and development-specific promoters, producing transcripts with divergent 5' untranslated regions. In the mouse mammary gland, the novel 5' UTR exon (L) appears to be associated with a drastic increase in ST6Gal I gene expression during lactation. We find that rats also possess an exon (L), suggesting conservation of this regulatory mechanism in rodents. In contrast, an exon (L)-containing transcript was not detected in lactating bovine or human mammary gland. We also observed a trend of increasing ST6Gal I gene expression in the bovine mammary gland, culminating in involution. This is in contrast to species such as mice where the greatest change in ST6Gal I gene expression occurs between pregnancy and lactation, suggesting different roles in rodents versus other mammals for α2,6 sialylated oligosaccharides present in milk.

2.2.2 Introduction

Sialoglycoconjugates are integral to various, ubiquitous biological phenomena, such as cell-cell interaction, cell migration, adhesion and metastasis (Dall'Olio, F. 2000, Hanasaki, K., Powell, L.D., et al. 1995, Varki, A. 1993, Varki, A. 2007). Interestingly, they may have been co-opted by the mammary gland to fulfil a different role. Sialoglycoconjugates found in milk can protect the gastrointestinal tract of the suckling neonate by competitively binding to invading pathogens and promoting growth of beneficial flora (Coppa, G.V., Gabrielli, O., et al. 1990, Martin-Sosa, S., Martin, M.J., et al. 2003, McVeagh, P. and Brand-Miller, J. 1997, Mouricout, M., Petit, J.M., et al. 1990, Wang, B., Brand-Miller, J., et al. 2001) and their potential role in post-natal brain development is of particular interest in human infant nutrition (McVeagh, P. and Brand-Miller, J. 1997, Wang, B. and Brand-Miller, J. 2003). Although the concentration and distribution of sialoglycoconjugates has been extensively studied in the milk of various species, the investigation of sialyltransferase gene expression in the mammary gland during lactation has been limited.

The sialyltransferase enzyme ST6Gal I transfers sialic acid from CMP-Sialic acid to either type 2 (Galβ1,4GlcNAc) free disaccharides or the *N*-acetyllactosamine termini of *N*- or *O*-linked oligosaccharides via an α2,6 linkage (Harduin-Lepers, A., Vallejo-Ruiz, V., et al. 2001). ST6Gal I is expressed in a variety of cell types and tissues (Aasheim, H.C., Aas-Eng, D.A., et al. 1993, Dalziel, M., Huang, R.Y., et al. 2001, Mercier, D., Wierinckx, A., et al. 1999, Wang, X., Vertino, A., et al. 1993), with particularly high levels of expression observed in the liver (Dalziel, M., Huang, R.Y., et al. 2001, Dalziel, M., Lemaire, S., et al. 1999). The enzyme has been successfully purified from rat, chicken and human liver and rat hepatoma cells (Bendiak, B. and Cook, G.M.W. 1982, Miyagi, T. and Tsuiki, S. 1982, Sticher, U., Gross, H.J., et al. 1991, Weinstein, J., de Souza-e-Silva, U., et al. 1982a, Weinstein, J., de Souza-e-Silva, U., et al. 1982b). A soluble form of ST6Gal I was also purified from bovine colostrum (Paulson, J.C., Beranek, W.E., et al. 1977), and was

established to be the product of proteolytic cleavage of the Golgi membrane-bound enzyme.

Mercier et al. (1999) identified and extensively characterised the bovine ST6Gal I gene, which has a coding region of 1218 nucleotides. The coding sequence consists of 5 exons, the last of which contains a 3'UTR of 2.7kb. Four additional exons were identified in the 5' UTR, resulting in a total of 9 exons spanning up to 80kb of genomic DNA. The gene comprises a single open reading frame encoding a protein of 405 residues with a theoretical molecular weight of 46250Da. Structural feature predictions indicate that the enzyme is likely to have a type II membrane topology, which is typical of Golgi glycosyltransferases. The bovine ST6Gal I protein sequence exhibits a higher homology with the human enzyme than with the rat, mouse or chicken enzymes (Mercier, D., Wierinckx, A., et al. 1999).

The human ST6Gal I gene shares similarities with its bovine orthologue. It is located on human chromosome 3 and also encodes a 405 residue protein across 5 exons, which span in excess of 40kb of genomic DNA. Four additional exons have been discovered in the 5'UTR (Wang, X., Vertino, A., et al. 1993).

Transcriptional regulation of the ST6Gal I gene is coordinated by the use of several cell-specific promoters that produce transcripts with divergent 5' untranslated regions (Svensson, E.C., Soreghan, B., et al. 1990, Wang, X., O'Hanlon, T.P., et al. 1990, Wen, D.X., Svensson, E.C., et al. 1992). The rat ST6Gal I gene, which has been extensively studied, is known to span over 80kb and has a complex network of at least four promoters that regulate its expression (Wen, D.X., Svensson, E.C., et al. 1992).

In vivo, three main families of bovine ST6Gal I transcripts have been characterised, leading to speculation of the existence of at least 3 separate promoters regulating the expression of the gene. Despite organisation of the bovine gene being similar to both human (Wang, X., Vertino, A., et al. 1993) and rat (Svensson, E.C., Soreghan, B., et al.

1990, Wang, X., O'Hanlon, T.P., et al. 1990, Wen, D.X., Svensson, E.C., et al. 1992), there are differences in the tissue distribution of its transcripts (Mercier, D., Wierinckx, A., et al. 1999).

The existence of a novel, lactogenic ST6Gal I promoter was proposed by the authors of a study examining the dramatic induction of ST6Gal I mRNA in the lactating mammary gland of mice. The predominant ST6Gal I mRNA species expressed in the lactating mouse mammary gland is a novel isoform containing a unique untranslated exon derived from the ST6Gal I 5'UTR, dubbed exon (L). Exon (L) is 203bp long and is located more than 48kb 5' of exon (I). Virgin mammary tissue exclusively expressed the housekeeping ST6Gal I mRNA form and probes specific to exon (L) confirmed that the exon (L)-containing transcript was only present in the lactating mammary gland and not in other tissues that were investigated (Dalziel, M., Huang, R.Y., et al. 2001).

The mouse ST6Gal I gene is known to contain at least 8 5'UTR exons in addition to exons (I)-(VI) (Dalziel, M., Huang, R.Y., et al. 2001, Dalziel, M., Lemaire, S., et al. 1999, Wuensch, S.A., Huang, R.Y., et al. 2000). Several studies have demonstrated that, despite not being translated, ST6Gal I 5'UTR exons exhibit a degree of nucleotide sequence conservation between species, with exons (I)-(VI) typically being well conserved (Aasheim, H.C., Aas-Eng, D.A., et al. 1993, Mercier, D., Wierinckx, A., et al. 1999, Wang, X., Vertino, A., et al. 1993).

As an initial step towards unravelling the pattern of expression of sialyltransferase genes in the lactating mammary gland, we have investigated the conservation of the 5'UTR of the ST6Gal I gene between several mammalian species. In addition, using a 5'RACE approach, we demonstrate that an equivalent to the mouse ST6Gal I exon (L) is not expressed in the lactating bovine or human mammary gland. Furthermore, expression of the bovine ST6Gal I gene is shown to be significantly increased at the onset of involution.

2.2.3 Methods

Conservation analysis

Mouse, bovine and human genomic ST6Gal I sequences, with at least an additional 100kb of 5'UTR, were obtained from Ensembl. The rat ST6Gal I genomic sequence was obtained from the Celera assembly of the *Rattus norvegicus* genome available through NCBI. Orthology of the sequences was confirmed using reciprocal best alignments and examination of chromosomal context via the Ensembl and UCSC genome browsers. The annotation of the mouse ST6Gal I gene was acquired from Ensembl in Vista format. Additional mouse 5'UTR exons were annotated by retrieving the sequences from GenBank: AF153680 (Q), AF153682 (X_2), U67989 (H), AF153684 (X_{1a} and X_{1b}), AF153683 (X_3); and aligning them to the 5'UTR of the mouse genomic ST6Gal I sequence using the Spidey mRNA-to-genomic sequence alignment program.

Conservation analysis of the mouse, bovine, human and rat ST6Gal I sequences was performed using the Vista suite of tools for comparative genomics (Dubchak, I., Brudno, M., et al. 2000, Frazer, K., Pachter, L., et al. 2004). The sequences and annotations were submitted for analysis via the mVista interface. The LAGAN algorithm (Brudno, M., Do, C.B., et al. 2003) for global multiple alignment of finished sequences was selected along with the option for translated anchoring to improve the alignment of distant homologues. All sequences were repeat-masked using species-appropriate masking. The resulting Vista plot was visualised using the Vista Browser with default parameters.

Collection of bovine mammary gland tissue and RNA extraction

The collection of the bovine mammary gland tissue used in this study was previously described in Sheehy, P.A., Della-Vedova, J.J., et al. (2004). Briefly, mammary gland tissue was collected from several Holstein-Friesian cows by surgical biopsy. The biopsies were taken from the lateral aspect of the dorsal third of the forequarter of the mammary gland. To expose the underlying mammary tissue, a vertical skin incision (60 mm) was made through the subcutaneous tissue and the lateral suspensory ligament. 5–10 g of mammary tissue was then obtained using a 40-mm elliptical incision at a depth of 10–20 mm. The tissue was then immediately placed in Medium 199 containing Earles salts (ICN, Costa Mesa CA 92626, USA), 25 mM Hepes; pH 7.4, 200 mg BSA/ml (culture grade, Sigma, St Louis MO 63178, USA), 50 units of penicillin/ml, 50 mg streptomycin/ml (Sigma), and stored on ice for transport (Sheehy, P.A., Della-Vedova, J.J., et al. 2004). Each cow was sampled during late pregnancy, ~8-23 days pre-partum; peak lactation, ~34 days post-partum; and involution, ~5 days into involution. Mammary gland tissue from four different cows, at all three time-points, was used in this analysis.

Total RNA was isolated from ~100mg of each of the 12 bovine mammary gland tissue samples using the RNeasy Lipid Tissue Mini Kit from Qiagen. The RNA samples were quantified using the Biolab NanoDrop and checked for integrity on agarose gels.

Collection of human milk samples and RNA extraction

Human milk was collected from a lactating female at 49 days following birth. Milk samples were centrifuged at 2000g, 5 min at 4^{o}C to pellet cells and milk was removed. Milk cells were resuspended in Buffer 1 of the Qiagen RNeasy Micro kit (Sydney Australia) following the manufacturer's instructions for isolation of purified total RNA.

Collection of bovine tissues and RNA extraction

A panel of tissues was collected from a lactating dairy cow that was slaughtered 34 days after calving. Total RNA was isolated from liver, abomasum, reticulum, rumen, brain stem, hind brain, lymph node and mammary gland tissue samples using the RNeasy Lipid Tissue Mini Kit from Qiagen as per the manufacturer's instructions.

5'RACE

5'RACE analysis was performed using the Roche Applied Science 5'/3' RACE Kit, 2nd Generation, as per the manufacturer's instructions. Briefly, 2µg of bovine mammary gland total RNA was annealed to the primer bST1-SP1 (5'-TAATTCCCTTTCTTCTTTTCTTTCC-3') and reverse transcribed. bST1-SP1 was designed to anneal to a region in exon (II), such that genuine reverse transcription events of ST6Gal I mRNA must span at least the exon (I)-(II) boundary. The resultant cDNA was purified using the High Pure PCR Product Purification Kit from Roche as per the instructions specified in the RACE kit. A homopolymeric A-tail was subsequently added to the 3' end of the purified cDNA. A first PCR amplification of the dA-tailed cDNA was then performed using the Promega GoTaq Green Master Mix with the ST6Gal I-specific primer bST1-SP2 (5'-GATGTCTGTTTTACTGGGTCTGG-3') and the Oligo(dT)-anchor primer provided with the RACE kit. The PCR was carried out in a MJ Research PTC-225 Pertier Thermal-Cycler at 94°C for 2 min, followed by 10 cycles at 94°C for 15 sec, 55°C for 30 sec and 72°C for 40 sec, then another 25 cycles at 94°C for 15 sec, 55°C for 30 sec and 72°C for 40 sec with an additional 20 sec of extension after each cycle, and finally 72°C for 7 min. A second PCR was then performed using the GoTaq Green Master Mix with the ST6Gal I-specific nested primer bST1-SP3 (5'-CCAAGGCTCATTCTTCTCAGG-3') and the PCR anchor primer provided with the RACE kit. The PCR parameters were 94°C for 2 min, followed by 10 cycles at 94°C for 15 sec, 60°C for 30 sec and 72°C for 40 sec, then another 25 cycles at 94°C for 15 sec, 60°C

for 30 sec and 72°C for 40 sec with an additional 20 sec of extension after each cycle, and finally 72°C for 7 min.

5'RACE was also performed on 2µg of human mammary gland total RNA as previously described. Human ST6Gal I primers reported by Wang et al. (1993) were used in the analysis; HST-P4 (5'-AACTTGATGCCTGGTCC-3') was used for reverse transcription, HST-P3 (5'-CTCTGGTTTGGCCTTGG-3') for the first PCR amplification and the nested primer HST-P5 (5'-CTGCTTCTGGCTAATC-3') for the second PCR amplification. The cycling conditions for the first PCR amplification were: 94°C for 2 min, then 10 cycles at 94°C for 15 sec, 55°C for 30 sec and 72°C for 40 sec, followed by another 25 cycles at 94°C for 15 sec, 55°C for 30 sec and 72°C for 40 sec with an additional 20 sec of extension after each cycle, and finally 72°C for 7 min. The second PCR amplification was carried out at 94°C for 2 min, 10 cycles at 94°C for 15 sec, 50°C for 30 sec and 72°C for 40 sec, followed by another 25 cycles at 94°C for 15 sec, 50°C for 30 sec and 72°C for 40 sec with an additional 20 sec of extension after each cycle, and finally 72°C for 7 min.

All products of the 5'RACE second PCR amplification were separated by electrophoresis on 2% agarose gels and the major bands subsequently excised and purified using the Qiagen QIAquick Gel Extraction Kit. The purified PCR products were then cloned into the plasmid vector pGEM-T Easy (Promega) and sequenced.

Quantitative PCR

cDNA was generated from 2µg total RNA using the Invitrogen SuperScript VILO cDNA Synthesis Kit as per the manufacturer's instructions. Quantitative PCR was performed according to instructions for using the EXPRESS SYBR GreenER qPCR SuperMix with Premixed ROX supplied by Invitrogen. All qPCR primers were designed using Primer3 (Rozen, S. and Skaletsky, H.J. 2000) and checked for primer-dimer formation, self-annealing and hairpin formation using PerlPrimer (Marshall, O. 2004) and OligoCalc (Kibbe, W.A. 2007). To detect overall ST6Gal I expression, a forward primer bST6.E2F

was designed in bovine ST6Gal I exon (II) (5'-GGTGTGCTGTGGTCTCTTCA-3') and a reverse primer in exon (III) bST6.E3R (5'-CCCACGTCTTGTTGGAATTT-3'). To distinguish the expression profiles of individual ST6Gal I transcript families, primers were designed to the unique 5'UTR exons of each bovine ST6Gal I isoform: for transcript family 1 bST6f1.E-2F (5'-GGGTCTGCTCCTGATACCAC-3') in exon (-2) and bST6f1.E1R (5'-CAAGGCTCATTCTTCTCAGGA-3') in exon (I), for transcript family 2 bST6f2.E-1F (5'-CACTACCCGGTGCTAACAAA-3') in exon (-1) and bST6f2.E0R (5'-GCAGATCTCAGCACAATTCACT-3') in exon (0), and for transcript family 3 bST6f3.E1F (5'-GCAGACTTGTCTTAGCTGATGG-3') in the extended portion of exon (I) and bST6f3.E2R (5'-CAAGGAAACCACGCTGTTCT-3') in exon (II). All primers were deliberately designed over intron-exon boundaries to facilitate detection of any genomic DNA contamination. Primers reported by (Bionaz, M. and Loor, J.J. 2007) for ribosomal protein sub-units RPS9 (RPS9.192F (5'-CCTCGACCAAGAGCTGAAG-3'), RPS9.254R (5'-CCTCCAGACCTCACGTTTGTTC-3')) and RPS15 (RPS15.405F (5'-GCAGCTTATGAGCAAGGTCGT-3'), RPS15.555R (5'-GCTCATCAGCAGATAGCGCTT-3')) and ubiquitously expressed transcript UXT (UXT.323F (5'-TGTGGCCCTTGGATATGGTT-3'), UXT.423R (5'-GGTTGTCGCTGAGCTCTGTG-3')) were used as internal controls. For all reactions, the final reaction mixture of 20µl consisted of 200 nM forward primer, 200 nM reverse primer, 2µl of 1 in 20 diluted cDNA and 10µl of the 2X EXPRESS SYBR GreenER qPCR SuperMix with Premixed ROX. PCR parameters for the Eppendorf RealPlex MasterCycler epgradients were 50^{o}C for 2 min, 95^{o}C for 10 min, 40 cycles at 95^{o}C for 15 sec then 60^{o}C for 1 min, followed by a melt-curve dissociation cycle at 95^{o}C for 15 sec then 60^{o}C for 1 min and a 20 min ramp up to 95^{o}C for 15 sec. The raw Ct values were imported into Biogazelle's qBasePlus software for calculation of Calibrated Normalized Relative Quantities (CNRQs) and their standard errors. The CRNQs were used in subsequent statistical analysis.

Statistical analysis

Significant differences in ST6Gal I gene expression in the bovine mammary gland between the stages of lactation were determined using paired t-tests, assuming unequal variances. For all analyses, a P-value of less than 0.05 was interpreted as statistically significant.

Transcription factor binding site analysis

2kb of human ST6Gal I genomic sequence upstream of, and including, exon (Y) and 2kb of bovine sequence upstream of, and including, exon (-1) were used for all TFBS analyses. All TFBS searches were restricted to lactation and APR related transcription factors. Potentially conserved transcription factor binding site positions were identified by rVista (Loots, G.G., Ovcharenko, I., et al. 2002), which makes transcription factor binding site predictions using the Match program based on TRANSFAC Professional 9.2, and determines which of the predicted sites are conserved between the species in the alignment. A core similarity value of 0.95, and matrix similarity value of 0.75 were used as cut off parameters for the TRANSFAC search. Match (Biobase) was used to search for transcription factor binding sites in each sequence individually, with default parameters for minimising false positives. The MatrixCatch (Kel, A., Kel-Margoulis, O., et al. 1999) program was used to locate potential composite elements in the human and cow sequences, also using its default parameters. This program uses a library of experimentally identified composite elements collected in the TRANSCOMPEL (Kel-Margoulis, O.V., Kel, A.E., et al. 2002) database, as well as mononucleotide position weight matrices for individual TFBS collected in TRANSFAC. Transcription start sites were predicted using TSSG (Softberry).

2.2.4 Results

Conservation analysis of the ST6Gal I gene

To investigate the level of conservation of 5'UTR exons characterised in the mouse and other mammalian species (Figure 1a) (Dalziel, M., Huang, R.Y., et al. 2001, Mercier, D., Wierinckx, A., et al. 1999, Wang, X., Vertino, A., et al. 1993, Wen, D.X., Svensson, E.C., et al. 1992, Wuensch, S.A., Huang, R.Y., et al. 2000), we compared mouse, human, bovine and rat ST6Gal I genomic sequence, including ~100kb of 5'UTR sequence, using the Vista suite of tools for comparative genomics (Frazer, K., Pachter, L., et al. 2004). Figure 1b shows the Vista conservation plot calculated from pair-wise genomic sequence alignments of the mouse ST6Gal I sequence with the other mammalian species. Translated exons (II)-(VI), are at least 75% conserved at the nucleotide sequence level for all pair-wise comparisons, with the greatest degree of conservation seen between the mouse and the rat sequence. Ubiquitously present 5'UTR exon (I) is also conserved by at least 75% for all sequence comparisons.

As expected, all mouse 5'UTR exons are highly conserved in the rat at the nucleotide sequence level, except 5'UTR exon ($X1_b$), which does not appear to be present in the rat sequence. Mouse exon (O) is particularly well conserved across all species; more than 75% identity is observed with human and rat and just under 70% with bovine. A multiple sequence alignment of mouse exon (O), rat exon (0), human exon (Z) and cow exon (0) sequences supports the Vista finding (Figure 2b).

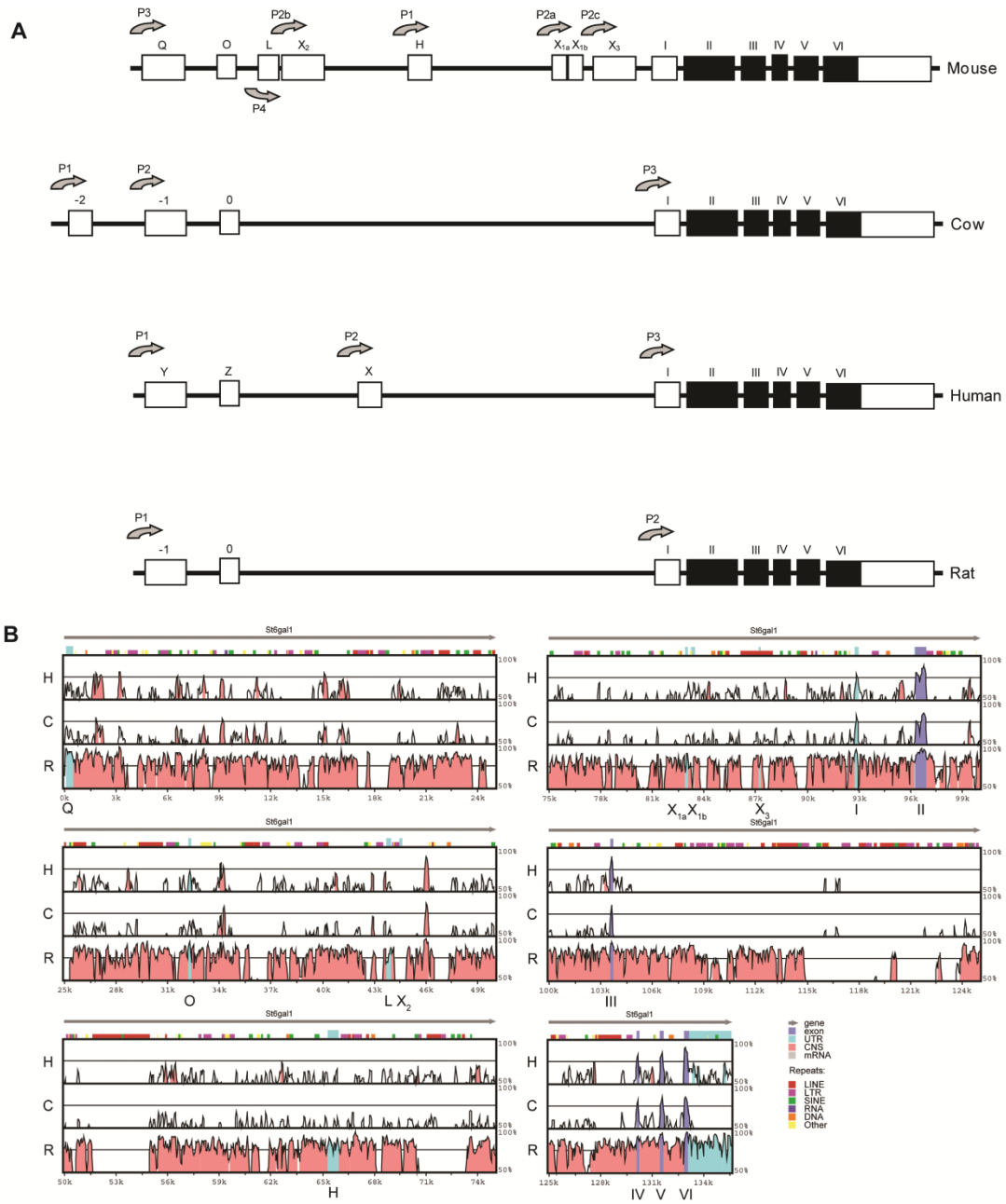


Figure 1: (a) Schematic representation of the mouse, cow, human and rat ST6Gal I genomic regions (not to scale). Translated exons are shown in black, untranslated exons are white. Putative promoter locations are denoted by arrows. (b) Vista conservation plot of the ST6Gal I genomic sequence, including ~100kb 5'UTR. The mouse ST6Gal I genomic sequence was used as the base sequence for alignment with the human (H), cow (C) and rat (R) ST6Gal I sequence. The sequences were submitted via mVista and aligned using the LAGAN algorithm for global multiple alignment of finished sequences, with the option for translated anchoring selected to improve the alignment of distant homologues. All sequences were repeat-masked using species-appropriate masking. The plot represents percent conservation between aligned sequences at a given coordinate on the base sequence, which is calculated as a windowed-average identity score for the alignment (default parameters were used for calculation). The top (100%) and bottom (50%) percentage bounds are shown to the right of every row. Regions of high conservation (>70%) are coloured as follows: exons (dark blue), UTRs (light blue) or non-coding (pink).

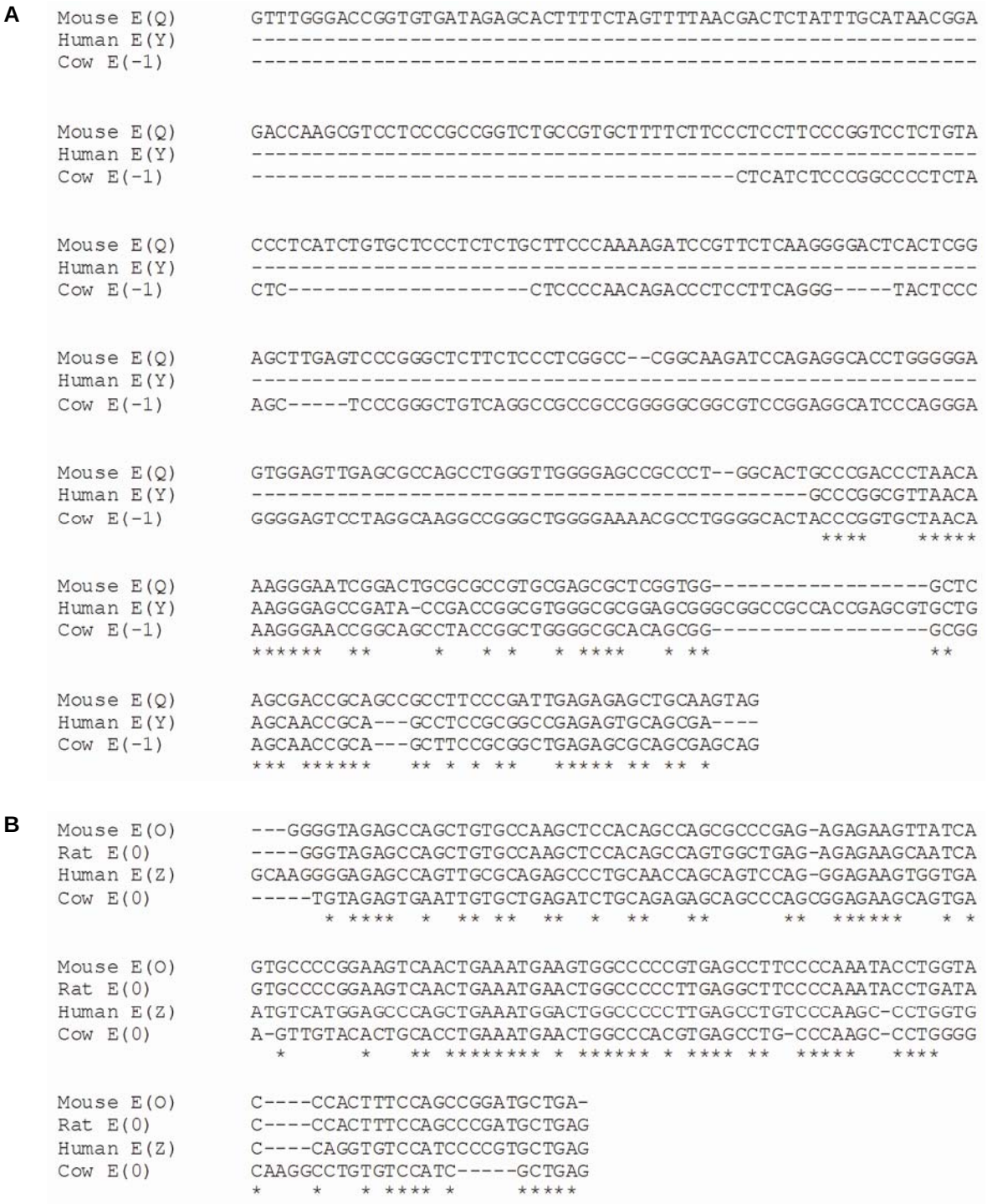


Figure 2: MUSCLE multiple sequence alignments of ST6Gal I 5'UTR exons from corresponding transcript families in mammalian species. Dashed lines denote gap in nucleotide sequence. Stars below sequence alignment indicate DNA identity. (A) Alignment of mouse exon (Q) (AF153680), human exon (Y) (NM_173216) and bovine exon (-1) (AJ006831). (B) Alignment of mouse exon (O) (AF153681), rat exon (0) (M83142), human exon (Z) (NM_173216) and cow exon (0) (AJ006832).

Our Vista analysis also shows a reasonable level of conservation in the region of mouse exon (Q), particularly between the mouse and the rat. However, a direct alignment of the mouse exon (Q) and rat exon (-1) (M83142) sequences showed only very poor sequence similarity (data not shown). The rat exon (-1) sequence was previously compared to human exon (Y) by Wang et al. (1993) who noted that the sequence similarity was not dramatic. A comparable result was reported by Mercier et al. (1999), where bovine exon (-1) was also included in the alignment. Spidey (Wheelan, S.J., Church, D.M., et al. 2001) alignments of rat exon (-1), human exon (Y) and bovine exon (-1) to the mouse ST6Gal I genomic sequence revealed that human exon (Y) and bovine exon (-1) are derived from the ST6Gal I genomic region corresponding with mouse exon (Q) but that the rat exon (-1) sequence corresponds to an area of the mouse ST6Gal I genomic sequence that is actually downstream of mouse exon (Q) and does not directly align with the mouse exon (Q) sequence (Figure 3). Mouse exon (Q) is homologous to human exon (Y) and bovine exon (-1) (Figure 2a).

Interestingly, mouse exon (H), which is exclusively expressed in the mouse liver (Dalziel, M., Lemaire, S., et al. 1999, Hu, Y.P., Dalziel, M., et al. 1997), is highly conserved in the rat genomic sequence, despite evidence that rat does not express this 5'UTR exon (Shah, S., Lance, P., et al. 1992, Vertino-Bell, A., Ren, J., et al. 1994, Wang, X., O'Hanlon, T.P., et al. 1990). There is also a very high level of sequence conservation between the mouse and the rat in the region of mouse exon (L) (Figure 4a), which is exclusively expressed in the lactating mouse mammary gland (Dalziel, M., Huang, R.Y., et al. 2001). Mouse exon (L) appears to be only moderately conserved in the human ST6Gal I genomic sequence (Figure 4c) and even less so in the bovine (Figure 4b).

Figure 3: Pair-wise alignment of mouse and rat ST6Gal I genomic sequence in the vicinity of mouse 5'UTR exon (Q), generated by Vista's LAGAN algorithm. The mouse exon (Q) sequence is highlighted in purple. Using Spidey, the rat and bovine exon (-1), and human exon (Y) sequences were mapped to the alignment. The rat exon (-1) sequence is highlighted in orange. Bovine exon (-1) maps to the mouse sequence denoted by lowercase and human exon (Y) maps to the underlined mouse sequence.

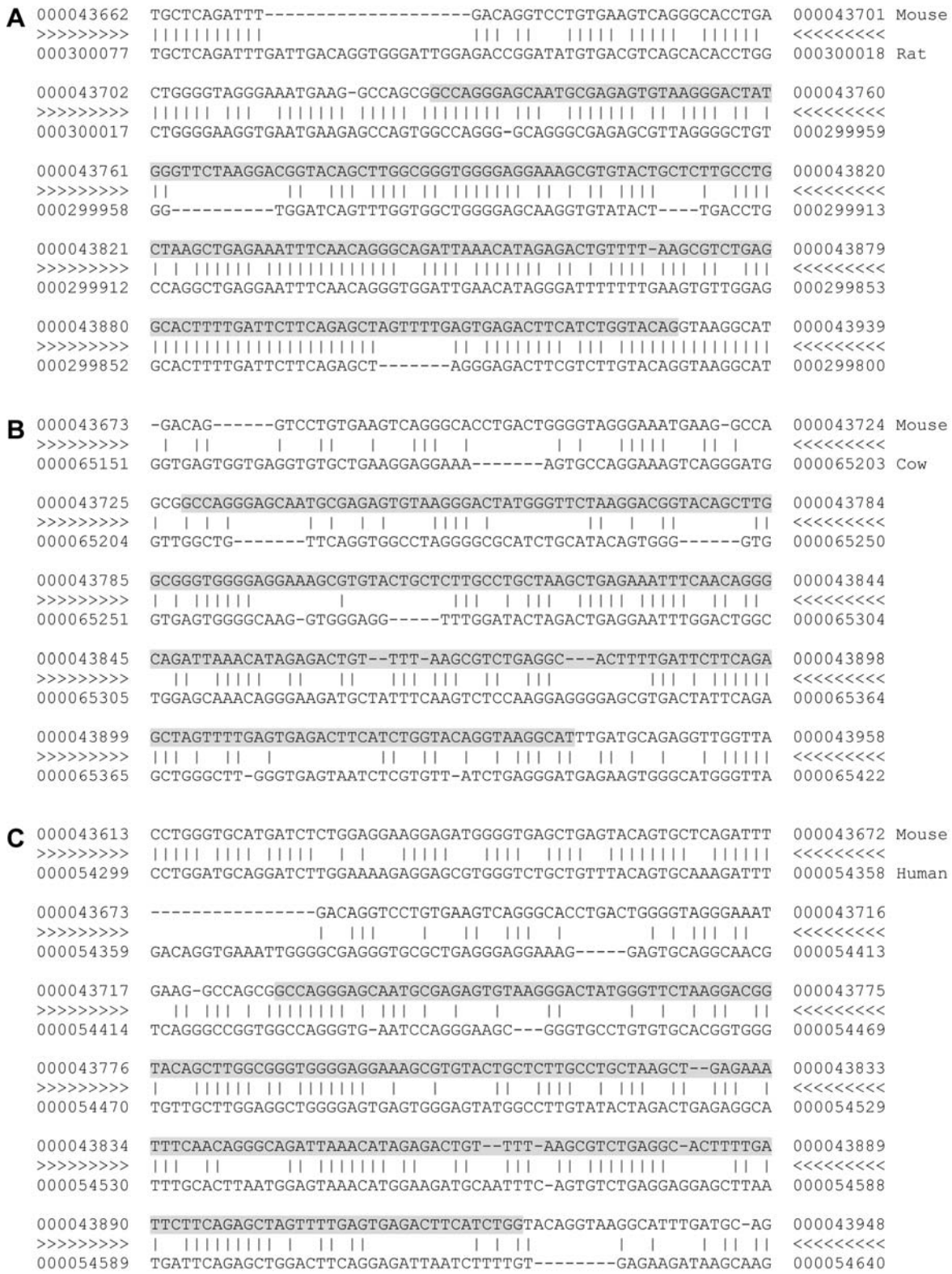


Figure 4: Pair-wise alignments of ST6Gal I genomic sequence in the vicinity of mouse 5'UTR exon (L), generated by Vista's LAGAN algorithm. The position of exon (L) in each pair-wise alignment is highlighted in grey. (A) Mouse versus rat pair-wise alignment. (B) Mouse versus cow pair-wise alignment. (C) Mouse versus human pair-wise alignment.

5'RACE of ST6Gal I transcripts in bovine and human mammary gland

Murine ST6Gal I expression is radically increased in the mammary gland during lactation (Dalziel, M., Huang, R.Y., et al. 2001). To investigate whether this phenomenon occurs in other mammals, we devised a 5' rapid amplification of cDNA ends (5'RACE) strategy to assess the nature of the ST6Gal I isoforms expressed in the mammary gland of cows and humans.

We prepared cDNA by reverse-transcription of RNA from the lactating mammary tissue of a day 34 post-partum Holstein-Friesian cow using bST1-SP1, an anti-sense primer designed to anneal to a region in ST6Gal I exon (II) (Figure 5A). The 5' ends of the synthesised cDNA were then amplified using a nested primer, bST1-SP2, which is complementary to a segment of exon (I). A further amplification was carried out with a second nested primer, bST1-SP3, which binds to a region of exon (I) further 5' than bST1-SP2. The major product, visualised on an agarose gel by SYBR Safe staining, migrated at approximately 250bp (Figure 6). RNA extracted from the mammary glands of a pregnant cow at 8 days pre-parturition and a cow 5 days into involution was subjected to the identical 5'RACE procedure. The PCR products were once again found to migrate at approximately 250bp (Figure 6). The products visible on the gel (Figure 6) were subsequently excised, purified and cloned.

Ten clones, from each stage of lactation, containing the SP3-derived PCR products were randomly selected for sequencing. All clones were confirmed to represent genuine ST6Gal I sequence as substantiated by the presence of ST6Gal I exon (I) sequence. Although all the clones contained >200bp of sequence upstream of exon (I), the sequence was not found to represent a homologous bovine form of the murine exon (L) at any of the time points investigated. All of the 5' sequences, at every stage of lactation, were found to contain bovine ST6Gal I exons (-1) and (0), representing the (-1+0) transcript family that was previously characterised by Mercier et al. (1999).

RNA was also reverse-transcribed from cells extracted from the milk of a lactating woman collected at 49 days following birth. Anti-sense primer HST-P4, complementary to a section of human ST6Gal I exon (II), was used for reverse transcription. Nested primer HST-P3, designed to anneal in exon (I), was used for the first PCR amplification and second nested primer HST-P5, which binds further 5' in exon (I), for the final amplification (Figure 5B). The major products were visualised as previously described and found to migrate at approximately 350, 320, 190 and 150bp.

All 4 of the products visible on the gel (Figure 7) were cloned. Five clones representing each HST-P5-derived product were randomly selected for subsequent sequencing. As with the bovine sequences, none of the human 5' segments were found to be homologous to murine exon (L). Furthermore, the two largest products contained human ST6Gal I exons (Y) and (Z). The smaller products were incomplete extensions of the same exon (Y+Z)-containing transcript.

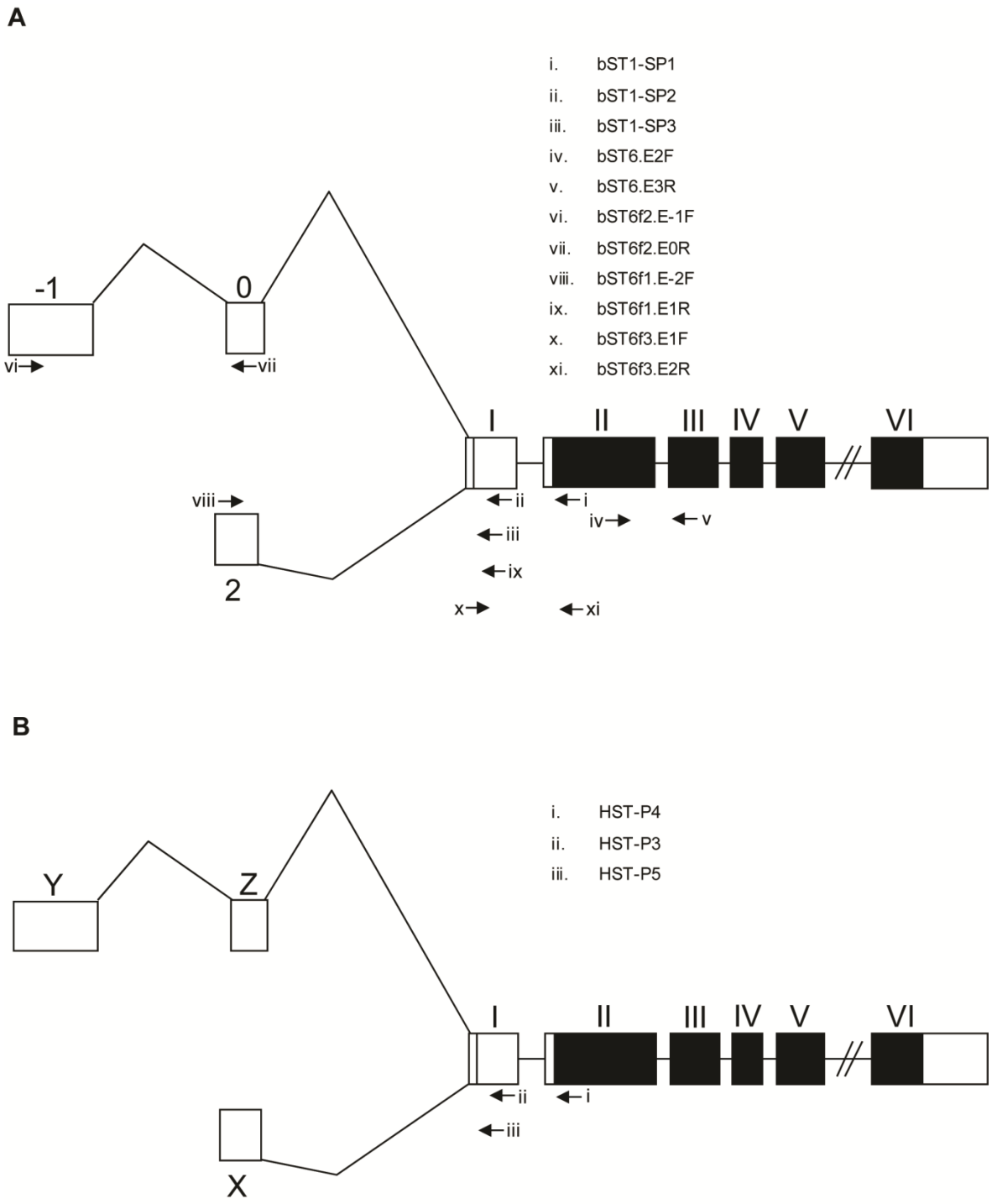


Figure 5: Schematic representations of bovine and human ST6Gal I transcript families. Translated exons are shown in black and untranslated exons in white. Primer locations are denoted by arrows. (a) Bovine ST6Gal I transcript structure. (b) Human ST6Gal I transcript structure.

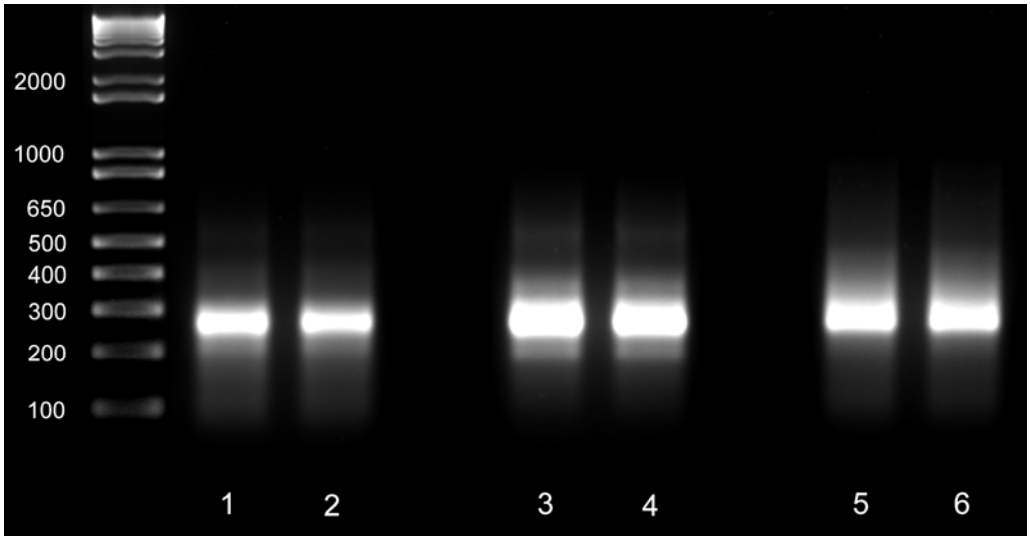


Figure 6: A 5'RACE analysis of ST6Gal I mRNA expressed in the bovine mammary gland. RNA was obtained from mammary gland tissue collected from a single animal by surgical biopsy at three stages of lactation. Lane 1, 2: 5'RACE analysis of mRNA obtained from mammary gland tissue 8 days pre-partum, generating predominantly a ~250bp product. Lane 3, 4: 5'RACE analysis of mRNA obtained from mammary gland tissue 33 days post-partum, generating predominantly a ~250bp product. Lane 5, 6: 5'RACE analysis of mRNA obtained from mammary gland tissue 5 days into involution, generating predominantly a ~250bp product. A portion of exon (I), along with exons (-1+0), representing transcript family 2, were identified by sequencing of the major product at all three stages.

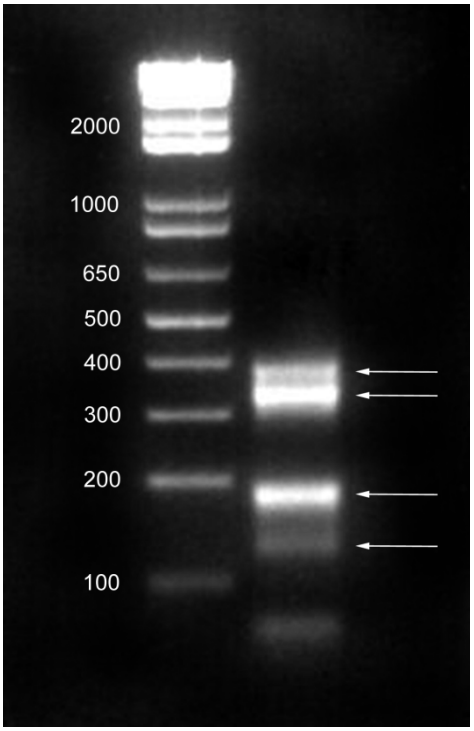


Figure 7: 5'RACE analysis of ST6Gal I mRNA expressed in the human mammary gland. RNA was obtained from cells extracted from milk collected from a woman at day 48 of lactation. The major products are indicated by arrows and migrate at approximately 350, 320, 190 and 150bp. All 4 of the products visible on the gel were sequenced. The two largest products represented human ST6Gal I exons (Y) and (Z). The smaller products were incomplete extensions of the same exon (Y+Z)-containing transcript.

ST6Gal I gene expression in the bovine mammary gland

The expression of the ST6Gal I gene in the bovine mammary gland was assessed during pregnancy, lactation and involution. Mammary gland samples, obtained by surgical biopsy from 4 Holstein-Friesian cows, at each of the 3 time points, were used for gene expression analysis. The expression level of ST6Gal I mRNA was determined for each sample using a quantitative PCR (qPCR) approach. To allow for the amplification of all ST6Gal I transcript families of the ST6Gal I gene, primers bST6.E2F and bST6.E3R were designed to anneal to regions in exon (II) and exon (III), respectively (Figure 5A). Figure 8 shows the mean normalised expression of the ST6Gal I gene in the mammary glands of 4 cows at the 3 stages of lactation sampled. Our analysis revealed a trend of increasing ST6Gal I gene expression in the bovine mammary gland between pregnancy and lactation, and lactation and involution. A statistically significant difference in ST6Gal I gene expression ($p<0.05$) was found between pregnancy and involution.

ST6Gal I gene expression in bovine tissues

The ST6Gal I gene has been shown to be more highly expressed in mouse liver than other tissues (Dalziel, M., Huang, R.Y., et al. 2001). An examination of ST6Gal I gene expression in bovine tissues showed that the kidney expressed higher levels of ST6Gal I mRNA than the brain, heart, lung or spleen (Mercier, D., Wierinckx, A., et al. 1999). We investigated the relative expression of the ST6Gal I gene in bovine liver, lactating mammary gland, brain stem, hind brain, lymph node, rumen, abomasum and reticulum by qPCR. Tissues were harvested from a lactating dairy cow that was sacrificed after 34 days of lactation. The bST6.E2F and bST6.E3R primers (Figure 5A) were used to detect the overall expression of the ST6Gal I gene. To discriminate between individual ST6Gal I transcript families, primers were designed to the unique 5'UTR exons of each bovine ST6Gal I isoform; for transcript family 1, bST6f1.E-2F in exon (-2) and bST6f1.E1R in exon (I), for transcript family 2, bST6f2.E-1F in exon (-1) and bST6f2.E0R in exon (0),

and for transcript family 3, bST6f3.E1F in the extended portion of exon (I) and bST6f3.E2R in exon (II) (Figure 5A).

ST6Gal I gene expression was significantly higher in the liver than any other tissue examined (Figure 9A). The lymph node demonstrated the next highest level of expression, relative to the liver, followed by the lactating mammary gland. Only very low levels of ST6Gal I gene expression were observed in the other tissues analysed.

Using specific primers to discriminate between the individual bovine ST6Gal I transcript families we were able to detect all 3 isoforms in all of the tissues examined. Relative to the other tissues assessed, the lymph node expressed the highest levels of transcript families 1 and 3 (Figure 9B&D). Transcript family 2 was expressed at similar levels in the brain stem, hind brain and lymph node and at lower levels in the other tissues (Figure 9C). Transcript families 2 and 3 were the best represented isoforms in the lactating mammary gland (Figure 9C&D).

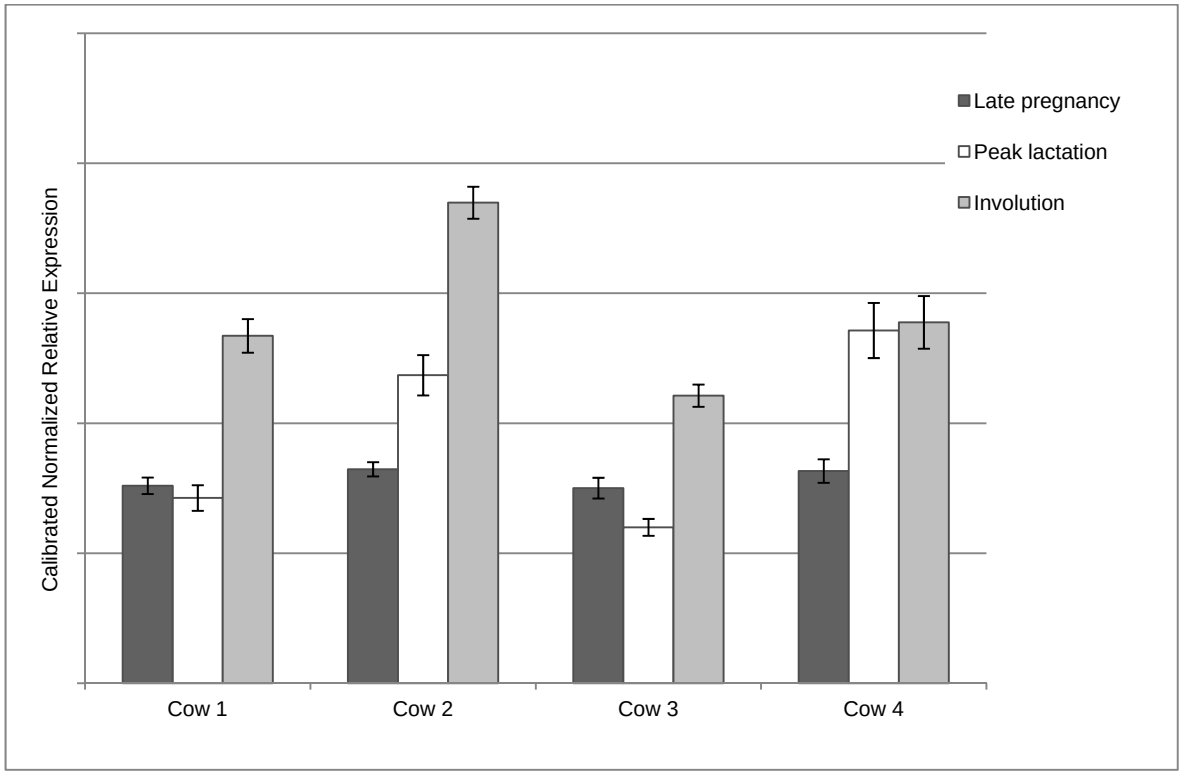


Figure 8: Relative overall expression of the ST6Gal I gene in the bovine mammary gland at three stages of lactation, as determined by qPCR. Mammary gland samples were obtained from 4 different cows by surgical biopsy at 3 stages of lactation. Each cow was sampled during late pregnancy, ~8-23 days pre-partum; peak lactation, ~34 days lactation; and involution, ~5 days into involution. Despite the increasing trend, there is no statistically significant difference in ST6Gal I gene expression in the bovine mammary gland between pregnancy and lactation, and lactation and involution; however, ST6Gal I gene expression does increase significantly between pregnancy and involution ($p < 0.05$).

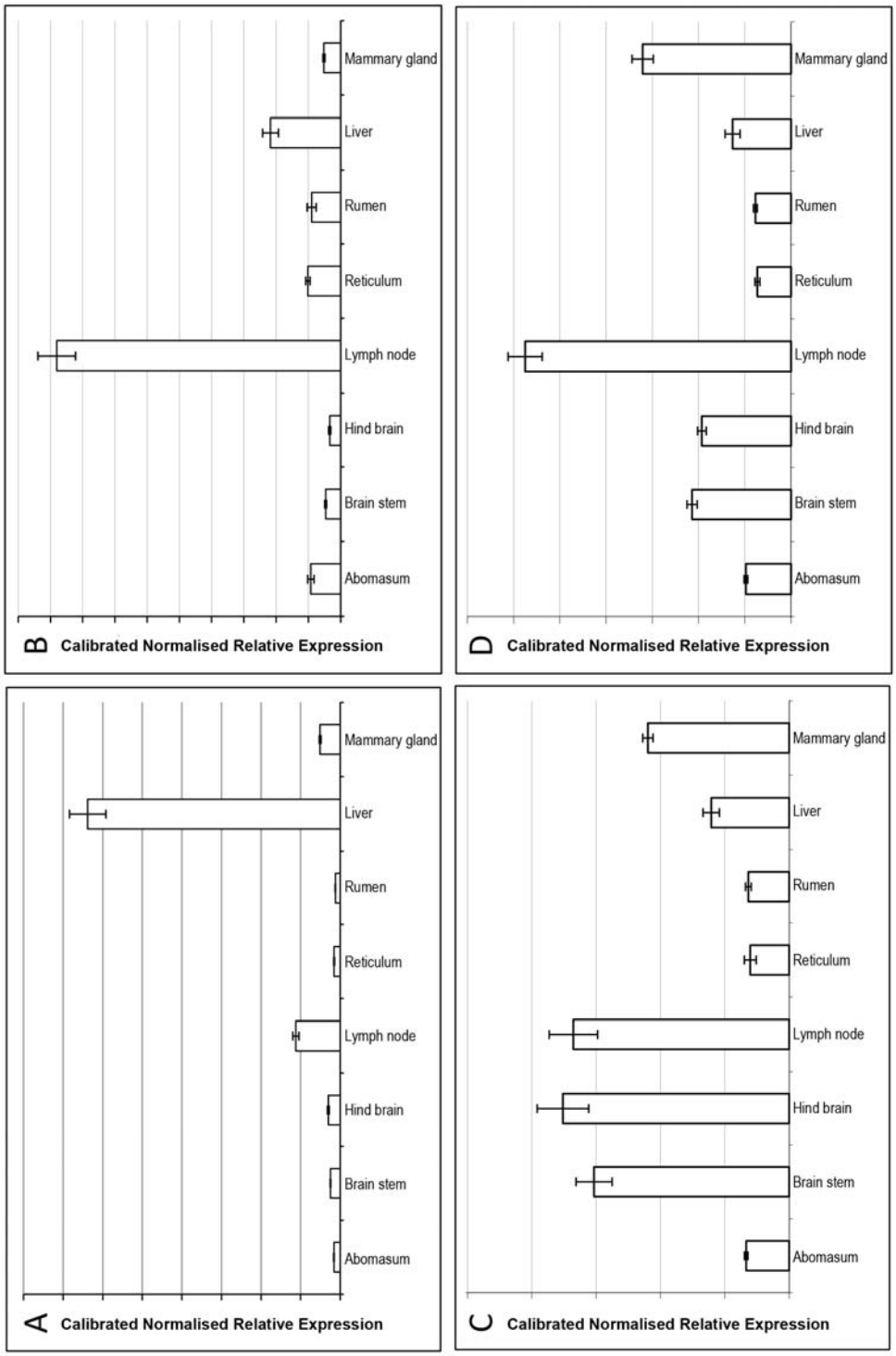


Figure 9: Relative expression of the ST6Gal I gene and its individual transcript families in different bovine tissues, as determined by qPCR. Due to different binding efficiencies of the different primers used to interrogate each transcript family, the relative expression levels of the different transcripts are not comparable. (A) Overall relative expression of the ST6Gal I gene. (B) Relative expression of ST6Gal I transcript family 1, previously documented in brain and kidney. (C) Relative expression of ST6Gal I transcript family 2, previously observed in the lung and spleen. (D) Relative expression of ST6Gal I transcript family 3, which was previously identified as the main form present in bovine tissues.

Transcription factor binding site analysis

As homologous bovine exon (-1+0)-containing and human exon (Y+Z)-containing transcripts were found to be the most well represented isoforms in the mammary gland of both species, we investigated potential similarities and differences in transcription factor binding sites (TFBS) of the putative promoter regions upstream of exon (-1) and (Y). Also, as we are primarily interested in ST6Gal I gene expression in the mammary gland, we focused our *in silico* promoter analysis on lactation-related TFBS, and TFBS involved in the acute phase response (APR), which has been implicated in early mammary gland involution (Stein, T., Morris, J., et al. 2004) and has been demonstrated to affect hepatic ST6Gal I expression (Dalziel, M., Lemaire, S., et al. 1999, Jamieson, J.C. 1992).

2kb of cow and human genomic sequence upstream of, and including, exon (-1) and (Y), was used in the TFBS analysis. A transcription start site (TSS) was predicted upstream of both exon (-1) in the cow sequence and exon (Y) in the human sequence (Figure 10). As shown in Figure 10, several lactation-related and APR TFBS were predicted to be conserved in this region. Two putative STAT5 motifs were identified, one of which is within 30bp of one of the three predicted Oct-1 sites. Two potential C/EBPβ sites were also identified, in addition to possible binding sites for Sp1 and AP1.

Both the cow and the human sequence contained a putative glucocorticoid receptor (GR) binding site, however, its position was not found to be conserved in the alignment. A GR-C/EBPβ composite element was identified in the cow sequence at -736 from the predicted TSS, but was not present in the human sequence. Conversely, the human sequence contained a potential Sp1-PU.1 composite element at -149 from the predicted TSS, which was not detected in the cow sequence.

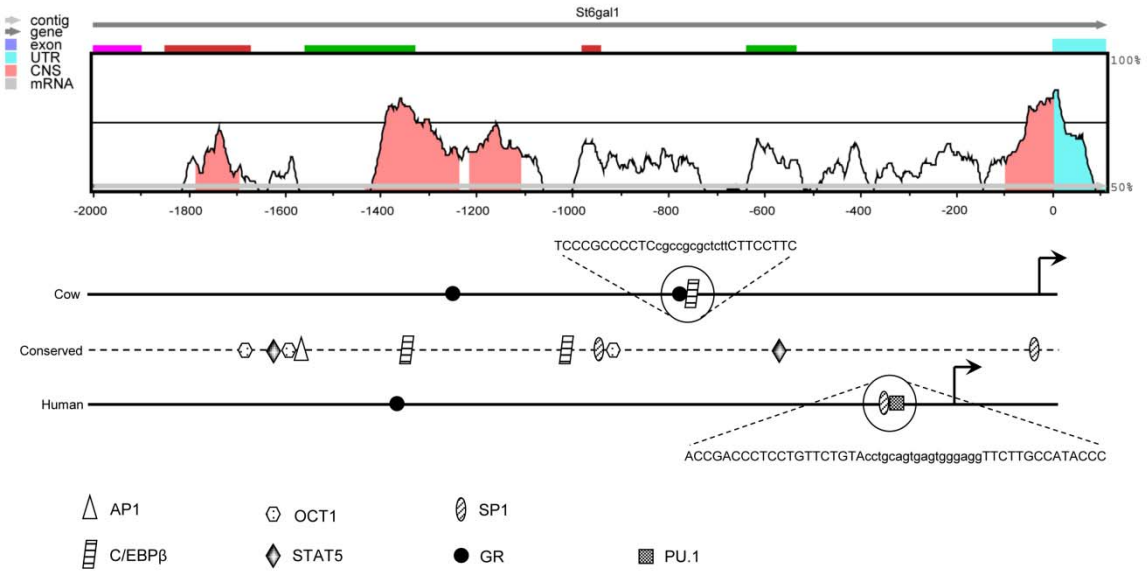


Figure 10: Cross-species comparison of the putative ST6Gal I promoter region upstream of human exon (Y) and cow exon (-1). A Vista conservation profile illustrates the level of similarity between human and cow sequence in the relevant region. 2kb human ST6Gal I genomic sequence upstream of exon (Y) was used for the alignment against 2kb cow sequence upstream of exon (-1). The top (100%) and bottom (50%) conservation percentage bounds are shown on the right of the plot. Highly conserved (>70%) non-coding sequence is shown in pink, human untranslated exon (Y) is shown in blue. Aligned underneath the conservation profile is a schematic representation of the transcription factor binding sites identified *in silico* as potentially conserved between the two species or found to be present only in the cow, or only in the human sequence. Transcription factor binding sites specifically relating to lactation or the acute phase response were targeted. Potentially conserved transcription factor binding site positions were identified by rVista. Match (Biobase) was used to search for transcription factor binding sites in each sequence individually. MatrixCatch was utilised to locate potential composite elements in the human and cow sequences. The composite elements identified are circled and their sequence is shown. The sequences corresponding to the positional weight matrix for each constituent transcription factor binding site are depicted in upper case, whilst the gap sequence is in lower case.

2.2.5 Discussion

The sequence, structure and organisation of the ST6Gal I gene has been well characterised in several mammals (Mercier, D., Wierinckx, A., et al. 1999, Svensson, E.C., Soreghan, B., et al. 1990, Wang, X., O'Hanlon, T.P., et al. 1990, Wang, X., Vertino, A., et al. 1993, Wen, D.X., Svensson, E.C., et al. 1992) and although the coding region of the gene is highly conserved between mammals, there is remarkable divergence in the pattern of expression of the different transcript families in some species (Dall'Olio, F. 2000). Despite their close evolutionary relationship, there are documented differences between rat and mouse ST6Gal I gene expression that are inconsistent with observed sequence similarity. Our analysis demonstrates that mouse exon (H) is very highly conserved in the rat genomic sequence, despite evidence that it is not expressed in rat tissues (Shah, S., Lance, P., et al. 1992, Vertino-Bell, A., Ren, J., et al. 1994, Wang, X., O'Hanlon, T.P., et al. 1990). Furthermore, an ST6Gal I transcript resembling the isoform expressed by the rat liver has been observed in human HepG2 cells (Aasheim, H.C., Aas-Eng, D.A., et al. 1993), suggesting that the liver-specific promoter upstream of exon (H) in mice may have evolved following their divergence from rats. The transcription of an additional 5' exon in mice may allow for increased translational efficiency of that particular transcript (Kozak, M. 1991a, Kozak, M. 1991b).

We also report that, although their parent transcripts exhibit similar patterns of expression, rat exon (-1) and mouse exon (Q) are not direct homologues at the genomic sequence level. The published sequence for rat exon (-1) is transcribed from a region further 3' of where mouse exon (Q), human exon (Y) and bovine exon (-1) align. This accounts for the observation by Wang et al. (1993) that although human exon (Z) and rat exon (0) were described as homologous, the sequence similarity between human exon (Y) and rat exon (-1) was not significant.

The concentration and distribution of sialoglycoconjugates in milk has been extensively studied for a variety of mammals, however, little is known about the expression of sialyltransferase genes in the mammary gland during lactation. Dalziel et al. (2001) reported the dramatic induction of ST6Gal I gene expression in the mouse mammary gland during lactation. This was achieved by the recruitment of a previously undocumented 5'UTR exon, exon (L), presumably regulated by an otherwise silent, lactogenic promoter. Our conservation analysis of the ST6Gal I 5'UTR genomic sequence between the rat and the mouse reveals that exon (L) is very highly conserved in the rat sequence. This suggests that ST6Gal I gene expression in the rat mammary gland during lactation may also be selectively elevated via the transcription of an exon (L)-containing isoform. Such a mechanism would be consistent with the observations of Ip et al. (1980), who documented an increase in the specific activity of rat ST6Gal I in the mammary gland throughout lactation, followed by a rapid return to basal values at the onset of involution. We suggest that rats and mice have evolved a mechanism to selectively up-regulate ST6Gal I in the lactating mammary gland to cater for the demand of α2,6-sialylated structures that serve as a delivery system for exogenous sialic acid. Rats and mice are born relatively immature and it has been shown that in comparison to adult levels, the liver of rat and guinea pig neonates has a diminished capacity to synthesise sialic acid (Gal, B., Ruano, M.J., et al. 1997, Nakano, T., Sugawara, M., et al. 2001). Furthermore, Dickson and Messer (1978) demonstrated a significant positive correlation between intestinal sialidase activity of suckling rats and mice, and sialic acid content of milk at the corresponding stage of lactation. They proposed that the main function of intestinal sialidase in suckling mammals is to participate in the digestion of sialylated milk components, thereby providing an exogenous supply of sialic acid for the neonate.

Based on speculation by Dalziel et al. (2001) that ST6Gal I elevation in lactating mouse mammary gland may be required to address the increased demand for the sialylation of type 2 disaccharides in milk, we investigated the presence of exon (L)-containing ST6Gal I

transcripts in human and bovine mammary gland. Interestingly, an equivalent lactation-specific transcript was not detected in either species. During pregnancy, lactation and involution we only identified the (-1+0) transcript in the bovine mammary gland, which was previously described in the lung and spleen (Mercier, D., Wierinckx, A., et al. 1999). At day 48 of lactation, the human mammary gland expressed the (Y+Z) isoform, which is thought to be the constitutively expressed transcript (Wang, X., Vertino, A., et al. 1993). Despite the documented differences in tissue distribution, the bovine (-1+0) and human (Y+Z) transcripts are homologous (Mercier, D., Wierinckx, A., et al. 1999), exhibiting a high degree of sequence conservation in the 5'UTR. Our comparative analysis of the putative promoter regions upstream of bovine exon (-1) and human exon (Y) revealed several potentially conserved TFBS that are lactation-related. However, we also identified differences that may contribute to disparate regulation of this ST6Gal I transcript between species. We propose that the difference in regulation of the ST6Gal I gene in the lactating mammary gland observed between mice, humans and cows may be related to the unique nutritional and immune requirements of the neonate of each species. Furthermore, we suggest that differences in regulation between species are likely influenced by variation in the suite of TFBS present in the proximal promoter.

Our quantitative analysis of ST6Gal I gene expression in the mammary gland throughout the bovine lactation cycle revealed that ST6Gal I gene expression significantly increases by early involution. Involution of the bovine mammary gland occurs at a slower rate than in other mammals such as rodents (Capuco, A.V. and Akers, R.M. 1999). After weaning, epithelial cells in the rat and mouse mammary gland rapidly undergo apoptosis, completely removing all alveoli within 2 weeks (Walker, N.I., Bennett, R.E., et al. 1989). This is the first step in remodelling the mammary gland to a morphology closely resembling the tissue present in the mature virgin animal. Atabai et al. (2007) noted that, particularly during rapid forced weaning, mastitis is not a significant concern in rodents. By contrast, in ruminants such as the cow, active involution may take between 20 and 30 days (Holst,

B.D., Hurley, W.L., et al. 1987). The bovine mammary gland is particularly susceptible to infection during early involution as large volumes of milk continue to accumulate in the gland for up to 4 days after weaning, creating an excellent growth medium for bacteria (O'Toole, D.K. 1995). The innate immune system plays an integral role in the defence of the mammary gland from infection at this stage (Sordillo, L.M., Shafer-Weaver, K., et al. 1997). For example, lactoferrin, a bacteriostatic iron-binding protein, is known to dramatically increase in concentration in bovine mammary secretions throughout involution (Rejman, J.J., Hurley, W.L., et al. 1989). We propose that an increase in ST6Gal I gene expression in the involuting bovine mammary gland may reflect an increased requirement for α2,6-sialylated structures to contribute to local innate immunity. Laporte et al. (2009) have demonstrated that, in a primary culture of mammary epithelial cells, bovine ST6Gal II expression increases in response to IL-6, which is known to be expressed during mastitis.

A role for milk sialoglycoconjugates in the innate immunity of the neonate has previously been suggested (Martin-Sosa, S., Martin, M.J., et al. 2003, Martin, M.J., Martin-Sosa, S., et al. 2001, Wang, B. and Brand-Miller, J. 2003, Wang, B., Brand-Miller, J., et al. 2001). The presence of α2,3-linked sialic acid on the host cell-surface is known to be essential for the successful infection of some pathogens (Karlsson, K.A. 1995) and it has been proposed that elevated levels of the α2,6-linkage may contribute to innate immunity by acting as "decoys" for invading pathogens (Gagneux, P. and Varki, A. 1999). It is conceivable, however speculative, that a similar mechanism may operate in the involuting bovine mammary gland during the period of milk stasis.

Increased expression of the ST6Gal I gene at the onset of bovine mammary gland involution may be driven by the expression of APR genes during the same period. As a result of their gene expression study of murine lactation, Stein et al. (2004) suggested a major role for the APR in the early stage of mouse mammary gland involution. The APR is

a component of the innate immune system that provides an initial mechanism of defence against disease by minimising tissue damage whilst promoting repair processes (Eckersall, P.D. 2000). Injury or infection usually stimulates acute-phase protein expression in the liver by the release of cytokines (Beutler, B. 1986, Dinarello, C.A. 1984). Jamieson et al. (1992) showed that elevation of liver and serum ST6Gal I is one component of the hepatic APR. Dalziel et al. (1999) went on to show that in mice, transcription of a novel ST6Gal I mRNA isoform containing 5'UTR exon (H) was responsible for the acute phase induction of hepatic ST6Gal I. Our analysis of the putative promoter region upstream of bovine ST6Gal I exon (-1) revealed the presence of a potential acute phase responsive composite element, GR-C/EPBβ, which may account for the increased expression of ST6Gal I observed at the onset of involution. This combination of GR and C/EPB motifs in close proximity has been shown to be necessary for the maximal induction of the APR-reactant AGP gene (Klein, E.S., DiLorenzo, D., et al. 1988, Ratajczak, T., Williams, P.M., et al. 1992, Williams, P., Ratajczak, T., et al. 1991). Furthermore, Nishio et al. (1993) demonstrated that NF-IL6, a member of the C/EPB family (Akira, S., Isshiki, H., et al. 1990), can directly interact with the glucocorticoid receptor and thus proposed that this interaction may be responsible for the synergistic activation of the rat AGP gene.

Based on our findings, we concur with previous suggestions that an exogenous supply of sialic acid may be necessary for the optimal development of mammalian neonates (Wang, B. and Brand-Miller, J. 2003), however, as Dickson and Messer (1978) proposed, we argue that the use of exogenous sialic acid may be limited to species that give birth to relatively immature young that initially have a diminished capacity for sialic acid synthesis (Gal, B., Ruano, M.J., et al. 1997, Nakano, T., Sugawara, M., et al. 2001). As such, it would be particularly interesting to investigate the expression of various sialyltransferases in the mammary gland of monotremes and marsupials, whose young are born in an extremely immature state and whose milk contains significant quantities of sialic acid (Messer, M. and Kerry, K.R. 1973). Although we suggest that mammals which give birth to relatively

well developed young may no longer have the need for direct utilisation of exogenously derived sialic acid, sialoglycoconjugates may still have a role in the innate immunity of the neonate, particularly in species such as humans where sialoglycoconjugates are a significant milk component. Additionally, we suggest that our observation of increasing ST6Gal I gene expression in the involuting mammary gland may indicate a role for α2,6-linked sialic acid during bovine mammary gland involution.

2.2.6 References

Aasheim HC, Aas-Eng DA, Deggerdal A, Blomhoff HK, Funderud S, Smeland EB. 1993. Cell-specific expression of human β-galactoside α2,6-sialyltransferase transcripts differing in the 5' untranslated region. *Eur. J. Biochem.*, 213:467-475.

Akira S, Isshiki H, Sugita T, Tanabe O, Kinoshita S, Nishio Y, Nakajima T, Hirano T, Kishimoto T. 1990. A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family. *EMBO J*, 9:1897-1906.

Atabai K, Sheppard D, Werb Z. 2007. Roles of the innate immune system in mammary gland remodeling during involution. *J. Mammary Gland Biol. Neoplasia*, 12:37-45.

Bendiak B, Cook GMW. 1982. Kinetic parameters of a β-D-galactoside α 2 leads to 6 sialyltransferase from embryonic chicken liver. *Eur. J. Biochem.*, 128:355-362.

Beutler B. 1986. Cachectin/tumor necrosis factor: an endogenous mediator of shock and inflammation. *Immunol. Res.*, 5:281-293.

Bionaz M, Loor JJ. 2007. Identification of reference genes for quantitative real-time PCR in the bovine mammary gland during the lactation cycle. *Physiol. Genomics*, 29:312-319.

Brudno M, Do CB, Cooper GM, Kim MF, Davydov E, Green ED, Sidow A, Batzoglou S. 2003. NISC Comparative sequencing program. LAGAN and Multi-LAGAN: Efficient tools for large-scale multiple alignment of genomic DNA. *Genome. Res.*, 13:721-731.

Capuco AV, Akers RM. 1999. Mammary involution in dairy animals. *J. Mammary Gland Biol. Neoplasia*, 4:137-144.

Coppa GV, Gabrielli O, Giorgi P, Catassi C, Montanari MP, et al. 1990. Preliminary study of breastfeeding and bacterial adhesion to uroepithelial cells. *The Lancet; London*, 335:569.

Dall'Olio F. 2000. The sialyl-α2,6-lactosaminyl-structure: Biosynthesis and functional role. *Glycoconj. J.*, 17:669-676.

Dalziel M, Huang RY, Dall'Olio F, Morris JR, Taylor-Papadimitriou J, Lau JTY. 2001. Mouse ST6Gal sialyltransferase gene expression during mammary gland lactation. *Glycobiology*, 11:407-412.

Dalziel M, Lemaire S, Ewing J, Kobayashi L, Lau JT. 1999. Hepatic acute phase induction of murine β-galactoside α2,6 sialyltransferase (ST6Gal I) is IL-6 dependent and mediated by elevation of Exon H-containing class of transcripts. *Glycobiology*, 9:1003-1008.

Dickson JJ, Messer M. 1978. Intestinal neuraminidase activity of suckling rats and other mammals - relationship to sialic-acid content of milk. *Biochem. J.*, 170:407-413.

Dinarello CA. 1984. Interleukin-1 and the pathogenesis of the acute-phase response. *N. Engl. J. Med.*, 311:1413-1418.

Dubchak I, Brudno M, Loots GG, Mayor C, Pachter L, Rubin EM, Frazer KA. 2000. Active conservation of noncoding sequences revealed by 3-way species comparisons. *Genome Res.*, 10:1304.

Eckersall PD. 2000. Recent advances and future prospects for the use of acute phase proteins as markers of disease in animals. *Rev. Med. Vet.*, 151:577-584.

Frazer K, Pachter L, Poliakov A, Rubin E, Dubchak I. 2004 VISTA: computational tools for comparative genomics. *Nucleic Acids Res.*, Jul 1:W273-279.

Gagneux P, Varki A. 1999. Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology*, 9:747-755.

Gal B, Ruano MJ, Puente R, GarciaPardo LA, Rueda R, Gil A, Hueso P. 1997. Developmental changes in UDP-N-acetylglucosamine 2-epimerase activity of rat and guinea-pig liver. *Comp. Biochem. Phys. B*, 118:13-15.

Hanasaki K, Powell LD, Varki A. 1995. Binding of human plasma sialoglycoproteins by the B cell-specific lectin CD22. Selective recognition of immunoglobulin M and haptoglobin. *J. Biol. Chem.*, 270:7543-7550.

Harduin-Lepers A, Vallejo-Ruiz V, Krzewinski-Recchi MA, Samyn-Petit B, Julien S, Delannoy P. 2001. The human sialyltransferase family. *Biochimie*, 83:727-737.

Holst BD, Hurley WL, Nelson DR. 1987. Involution of the bovine mammary gland: histological and ultrastructural changes. *J. Dairy Sci.*, 70:935-944.

Hu YP, Dalziel M, Lau JT. 1997. Murine hepatic β-galactoside α2,6-sialyltransferase gene expression involves usage of a novel upstream exon region. *Glycoconj. J.*, 14:407-411.

Ip C. 1980. Changes in tissue and serum sialyltransferase activities as related to proliferation and involution of the rat mammary gland. *Biochim. Biophys. Acta*, 628:249-254.

Jamieson JC. 1992. The importance of N-linked glycoproteins and dolichyl phosphate synthesis for fusion of L6 myoblasts. *Biochem. Cell Biol.*, 70:408-412.

Karlsson KA. 1995. Microbial recognition of target-cell glycoconjugates. *Curr. Opin. Struct. Biol.*, 5:622-635.

Kel-Margoulis OV, Kel AE, Reuter I, Deineko IV, Wingender E. 2002. TRANSCompel(R): a database on composite regulatory elements in eukaryotic genes. *Nucl. Acids Res.*, 30:332-334.

Kel A, Kel-Margoulis O, Babenko V, Wingender E. 1999. Recognition of NFATp/AP-1 composite elements within genes induced upon the activation of immune cells. *Journal of Molecular Biology*, 288:353-376.

Kibbe WA. 2007. OligoCalc: an online oligonucleotide properties calculator. *Nucl. Acids Res.*, 35:W43-46.

Klein ES, DiLorenzo D, Posseckert G, Beato M, Ringold GM. 1988. Sequences downstream of the glucocorticoid regulatory element mediate cycloheximide inhibition of steroid induced expression from the rat alpha 1-acid glycoprotein promoter: evidence for a labile transcription factor. *Mol Endocrinol*, 2:1343-1351.

Kozak M. 1991a. An analysis of vertebrate mRNA sequences: intimations of translational control. *J. Cell Biol.*, 115:887-903.

Kozak M. 1991b. Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem.*, 266:19867-19870.

Laporte B, Gonzalez-Hilarion S, Maftah A, Petit J-M. 2009. The second bovine β-Galactoside-α2,6-sialyltransferase (ST6Gal II): genomic organization and stimulation of its in vitro expression by IL-6 in bovine mammary epithelial cells. *Glycobiology*:1082-1093.

Loots GG, Ovcharenko I, Pachter L, Dubchak I, Rubin EM. 2002. rVista for Comparative Sequence-Based Discovery of Functional Transcription Factor Binding Sites. *Genome Research*, 12:832-839.

Marshall O. 2004. PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. *Bioinformatics* 20:2471-2472.

Martin-Sosa S, Martin MJ, Garcia-Pardo LA, Hueso P. 2003. Sialyloligosaccharides in human and bovine milk and in infant formulas: variations with the progression of lactation. *J. Dairy Sci.*, 86:52-59.

McVeagh P, Brand-Miller J. 1997. Human milk oligosaccharides: only the breast. *J. Paediatr. Child Health*, 33:281-286.

Mercier D, Wierinckx A, Oulmouden A, Gallet PF, Palcic MM, Harduin-Lepers A, Delannoy P, Petit JM, Leveziel H, Julien R. 1999. Molecular cloning, expression and exon/intron organization of the bovine β-galactoside α2,6-sialyltransferase gene. *Glycobiology*, 9:851-863.

Messer M, Kerry KR. 1973. Milk carbohydrates of echidna and platypus. *Science*, 180:201-203.

Miyagi T, Tsuiki S. 1982. Purification and characterization of β-galactoside α2,6-sialyltransferase from rat liver and hepatomas. *Eur. J. Biochem.*, 126:253-261.

Mouricout M, Petit JM, Carias JR, Julien R. 1990. Glycoprotein glycans that inhibit adhesion of Escherichia coli mediated by K99 fimbriae: treatment of experimental colibacillosis. *Infect. Immun.*, 58:98-106.

Nakano T, Sugawara M, Kawakami H. 2001. Sialic acid in human milk: composition and functions. *Acta Paediatr. Taiwan*, 42:11-17.

Nishio Y, Isshiki H, Kishimoto T, Akira S. 1993. A nuclear factor for interleukin-6 expression (NF-IL6) and the glucocorticoid receptor synergistically activate transcription of the rat α1-acid glycoprotein gene via direct protein-protein interaction. *Mol. Cell. Biol.*, 13:1854-1862.

O'Toole DK. 1995. Factors inhibiting and stimulating bacterial growth in milk: a historical perspective. Advances in Applied Microbiology Volume 40. San Diego: Academic Press. p. 45-47.

Paulson JC, Beranek WE, Hill RL. 1977. Purification of a sialyltransferase from bovine colostrum by affinity chromatography on CDP-agarose. *J. Biol. Chem.*, 252:2356-2362.

Ratajczak T, Williams PM, DiLorenzo D, Ringold GM. 1992. Multiple elements within the glucocorticoid regulatory unit of the rat α1-acid glycoprotein gene are recognition sites for C/EBP. *J. Biol. Chem.*, 267:11111-11119.

Rejman JJ, Hurley WL, Bahr JM. 1989. Enzyme-linked immunosorbent assays of bovine lactoferrin and a 39-kilodalton protein found in mammary secretions during involution. *J. Dairy Sci.*, 72:555-560.

Rozen S, Skaletsky HJ. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S editors. Bioinformatics Methods and Protocols: Methods in Molecular Biology. Totowa, NJ: Humana Press. p. 365-386.

Shah S, Lance P, Smith TJ, Berenson CS, Cohen SA, Horvath PJ, Lau JT, Baumann H. 1992. n-butyrate reduces the expression of β-galactoside α2,6- sialyltransferase in Hep G2 cells. *J. Biol. Chem.*, 267:10652-10658.

Sheehy PA, Della-Vedova JJ, Nicholas KR, Wynn PC. 2004. Hormone-dependent milk protein gene expression in bovine mammary explants from biopsies at different stages of pregnancy. *J. of Dairy Res.*, 71:135-140.

Sordillo LM, Shafer-Weaver K, DeRosa D. 1997. Immunobiology of the mammary gland. *J. Dairy Sci.*, 80:1851-1865.

Stein T, Morris J, Davies C, Weber-Hall S, Duffy M-A, Heath V, Bell A, Ferrier R, Sandilands G, Gusterson B. 2004. Involution of the mouse mammary gland is associated

with an immune cascade and an acute-phase response, involving LBP, CD14 and STAT3. *Breast Cancer Res.*, 6:R75 - R91.

Sticher U, Gross HJ, Brossmer R. 1991. Purification and characterization of α2,6-sialyltransferase from human liver. *Glycoconj. J.*, 8:45-54.

Svensson EC, Soreghan B, Paulson JC. 1990. Organization of the β-galactoside α2,6-sialyltransferase gene. Evidence for the transcriptional regulation of terminal glycosylation. *J. Biol. Chem.*, 265:20863-20868.

Varki A. 1993. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*, 3:97-130.

Varki A. 2007. Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. *Nature*, 446:1023-1029.

Vertino-Bell A, Ren J, Black JD, Lau JTY. 1994. Developmental regulation of β-galactoside α2,6-sialyltransferase in small intestine epithelium. *Dev. Biol.*, 165:126-136.

Walker NI, Bennett RE, Kerr JFR. 1989. Cell death by apoptosis during involution of the lactating breast in mice and rats. *Am. J. Anat.*, 185:19-32.

Wang B, Brand-Miller J. 2003. The role and potential of sialic acid in human nutrition. *Eur. J. Clin. Nutr.*, 57:1351-1369.

Wang B, Brand-Miller J, McVeagh P, Petocz P. 2001. Concentration and distribution of sialic acid in human milk and infant formulas. *Am. J. Clin. Nutr.*, 74:510-515.

Wang X, O'Hanlon TP, Young RF, Lau JTY. 1990. Rat β-galactoside α2,6-sialyltransferase genomic organization: alternate promoters direct the synthesis of liver and kidney transcripts. *Glycobiology*, 1:25-31.

Wang X, Vertino A, Eddy RL, Byers MG, Jani-Sait SN, Shows TB, Lau JT. 1993. Chromosome mapping and organization of the human β-galactoside α2,6-sialyltransferase gene. Differential and cell-type specific usage of upstream exon sequences in B-lymphoblastoid cells. *J. Biol. Chem.*, 268:4355-4361.

Weinstein J, de Souza-e-Silva U, Paulson JC. 1982a. Purification of a Gal β1 to 4GlcNAc α2 to 6 sialyltransferase and a Gal β1 to 3(4)GlcNAc α2 to 3 sialyltransferase to homogeneity from rat liver. *J. Biol. Chem.*, 257:13835-13844.

Weinstein J, de Souza-e-Silva U, Paulson JC. 1982b. Sialylation of glycoprotein oligosaccharides N-linked to asparagine. Enzymatic characterization of a Gal β1 to 3(4)GlcNAc α2 to 3 sialyltransferase and a Gal β1 to 4GlcNAc α2 to 6 sialyltransferase from rat liver. *J. Biol. Chem.*, 257:13845-13853.

Wen DX, Svensson EC, Paulson JC. 1992. Tissue-specific alternative splicing of the β-galactoside α2,6- sialyltransferase gene. *J. Biol. Chem.*, 267:2512-2518.

Wheelan SJ, Church DM, Ostell JM. 2001. Spidey: a tool for mRNA-to-genomic alignments. *Genome Res.*, 11:1952-1957.

Williams P, Ratajczak T, Lee SC, Ringold GM. 1991. AGP/EBP(LAP) expressed in rat hepatoma cells interacts with multiple promoter sites and is necessary for maximal glucocorticoid induction of the rat α-1 acid glycoprotein gene. *Mol. Cell. Biol.*, 11:4959-4965.

Wuensch SA, Huang RY, Ewing J, Liang X, Lau JT. 2000. Murine B cell differentiation is accompanied by programmed expression of multiple novel β-galactoside α2, 6-sialyltransferase mRNA forms. *Glycobiology*, 10:67-75.

3 Chapter Three

Given the importance of transcriptional control in the expression of sialyltransferase genes (Svensson, E.C., Soreghan, B., et al. 1990, Takashima, S., Kono, M., et al. 2000, Wang, X., O'Hanlon, T.P., et al. 1990, Wen, D.X., Svensson, E.C., et al. 1992) and the evidence of differential expression in the mammary gland during lactation (Dalziel, M., Huang, R.Y., et al. 2001), it was necessary to obtain a comprehensive picture of sialyltransferase gene expression in the lactating mammary gland, in order to gain a better understanding of the types of sialylated structures that may be particularly important as milk constituents and the role that transcriptional control plays in the process. As the mouse has been extensively utilised as a model for lactation, there are an abundance of microarray gene expression datasets available in the public domain that comprehensively cover the murine lactation cycle and can consequently be mined for data on genes involved in sialylation. Two publicly available mouse lactation microarray datasets were chosen for this study as both incorporated extensive biological replication in their experimental design, both covered a comprehensive, overlapping time course of mouse lactation and, as they were previously used as the basis for published studies, both had been extensively validated (Anderson, S.M., Rudolph, M.C., et al. 2007, Lemay, D., Neville, M., et al. 2007, Rudolph, M.C., McManaman, J.L., et al. 2003, Stein, T., Morris, J., et al. 2004, Stein, T., Salomonis, N., et al. 2007). Furthermore, the array platform used in both studies included probes for a majority of the known sialyltransferase genes. Also, as sialic acid can be synthesised in all tissues to satisfy local sialylation requirements (Varki, A. and Schauer, R. 2009), the data also facilitated the investigation of whether sialic acid biosynthesis genes increased in expression to cater for an increase in demand for sialic acid in the lactating mammary gland.

This chapter is presented in manuscript format in preparation for journal review.

3.1 Declaration for thesis chapter three

3.1.1 Declaration by candidate

In the case of Chapter 3, the nature and extent of my contribution to the work was the following:

Nature of contribution	**Extent of contribution (%)**
For this chapter I analysed the publicly available microarray data and was responsible for the interpretation of the results and the writing of the manuscript.	80%

The following co-authors contributed to the work. Co-authors who are students at Monash University must also indicate the extent of their contribution in percentage terms:

Name	**Nature of contribution**	**Extent of contribution (%)** **(student co-authors only)**
Prof. Kevin R. Nicholas	Prof. Nicholas aided in interpretation of results and writing of the manuscript.	
Dr. Julie A. Sharp	Dr. Sharp aided in interpretation of results and writing of the manuscript.	
Asoc. Prof. Christophe Lefèvre	Asoc. Prof. Lefèvre aided in the analysis of the microarray data and with interpretation of results and writing of the manuscript.	
Dr. Keith W. Savin	Dr. Savin aided in interpretation of results and writing of the manuscript.	

Candidate's Signature		**Date**

3.1.2 Declaration by co-authors

The undersigned hereby certify that:

(1) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.
(2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
(3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
(4) there are no other authors of the publication according to these criteria;
(5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
(6) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

Location(s)	**DPI, Biosciences Research Division, Bundoora VIC 3083**

Signature 1	[redacted]	**Date** 6/4/11
Signature 2	[redacted]	4-4-11
Signature 3	[redacted]	6-4-11
Signature 4	[redacted]	26/3/2011

3.2 Differential expression of sialic acid biosynthesis and sialyltransferase genes in the mouse mammary gland during lactation

Jovana Maksimovic[1,2], Kevin R. Nicholas[3], Julie A. Sharp[3], Christophe Lefèvre[3], Keith W. Savin[1]

1. Biosciences Research Division, Department of Primary Industries, Bundoora 3083, Australia
2. Centre for Reproduction and Development, Monash Institute of Medical Research, Monash University, Clayton 3168, Australia
3. Institute for Technology Research and Innovation, Deakin University, Geelong 3214, Australia

3.2.1 Abstract

Milk sialylglycoconjugates can protect the gastrointestinal tract of the suckling neonate by competitively binding to invading pathogens and promoting growth of beneficial flora, and their potential role in post-natal brain development is of particular interest in human infant nutrition. Although the concentration and distribution of sialylglycoconjugates has been extensively studied in the milk of various species, the investigation of sialyltransferase gene expression in the mammary gland, in the context of lactation, has been limited. Thus far, 20 members of the murine sialyltransferase family have been identified. Depending on the type of linkage formed and the type of sugar acceptor, they are traditionally allocated to four groups: ST6Gal, ST6GalNAc, ST3Gal and ST8Sia. Our analysis of mouse mammary gland lactation time course microarray data has revealed that key sialic acid biosynthesis genes are up-regulated during lactation, but that only a subset of known sialyltransferase genes are expressed. The UDP-GlcNAc 2-epimerase/ManNAc kinase (GNE) and N-acetylneuraminic acid synthase (NANS) genes coding for enzymes exclusively involved in sialic acid biosynthesis were found to be up-regulated in the mouse mammary gland during lactation. The ST6Gal I, ST3Gal I, ST3Gal IV and ST6GalNAc II sialyltransferase genes were also found to increase in expression during lactation, whilst ST3Gal VI and ST8Sia IV are actively down-regulated during the same period. Our findings indicate that sialylation in the mammary gland is tightly controlled at the level of transcription and appears to correlate with the phases of the mammary gland lactation cycle, suggesting that the genes governing sialylation may be regulated by lactogenic promoters.

3.2.2 Introduction

Sialic acids are a family of highly related nine-carbon, acidic monosaccharides. In mammalian species, sialic acids are typically located at the non-reducing terminus of oligosaccharide chains, connected either via an α2,3- or α2,6- linkage to a β-D-galactopyranosyl (Gal) residue, or by an α2,6-linkage to a β-D-N-acetylgalactosaminyl (GalNAc) or β-D-N-acetylglucosaminyl (GlcNAc) residue. Alternatively, sialic acid can form α2,8-linked poly-sialic acid chains on gangliosides and glycoproteins (Harduin-Lepers, A., Recchi, M.A., et al. 1995, Tsuji, S. 1996).

Sialoglycoconjugates are fundamental to various, ubiquitous biological phenomena, such as cell-cell interaction, cell migration, adhesion and metastasis (Dall'Olio, F. 2000, Hanasaki, K., Powell, L.D., et al. 1995, Varki, A. 1993, Varki, A. 2007). However, they may perform an alternative function in the lactating mammary gland. Milk sialoglycoconjugates are thought to exert a protective affect on the gastrointestinal tract of the suckling neonate by competitively binding to invading pathogens and/or promoting the growth of beneficial flora (Coppa, G.V., Gabrielli, O., et al. 1990, Martin-Sosa, S., Martin, M.J., et al. 2003, McVeagh, P. and Brand-Miller, J. 1997, Mouricout, M., Petit, J.M., et al. 1990, Wang, B., Brand-Miller, J., et al. 2001) and their possible role in post-natal brain development is of particular interest in human infant nutrition (McVeagh, P. and Brand-Miller, J. 1997, Wang, B. and Brand-Miller, J. 2003). Although the concentration and distribution of sialylated compounds has been extensively studied in the milk of various species, the investigation of sialyltransferase gene expression in the mammary gland during lactation has been limited.

Sialyltransferases are a family of Golgi membrane-bound, glycosyltransferases that transfer sialic acid from CMP-sialic acid to carbohydrate acceptor groups of glycoproteins, glycolipids and free oligosaccharides, forming various sialoglycoconjugates (Harduin-Lepers, A., Mollicone, R., et al. 2005). Their expression is primarily controlled at the level of transcription and is expedited through the use of several tissue and development-specific

promoters, giving rise to numerous transcripts divergent only in their 5' untranslated regions (Svensson, E.C., Soreghan, B., et al. 1990, Takashima, S., Kono, M., et al. 2000, Wang, X., O'Hanlon, T.P., et al. 1990, Wen, D.X., Svensson, E.C., et al. 1992).

Thus far, 20 members of the murine sialyltransferase family have been cloned [for review see Takashima (2008)]. Depending on the type of linkage formed and the type of sugar acceptor, they are traditionally allocated to four groups: ST6Gal, ST6GalNAc, ST3Gal and ST8Sia. There are one or more enzymes in each category and although some have broader acceptor specificities than others, the enzymes in each group only form one type of linkage between the sialic acid and the sugar acceptor (Harduin-Lepers, A., Mollicone, R., et al. 2005, Harduin-Lepers, A., Vallejo-Ruiz, V., et al. 2001). It was suggested by Takashima (2008), that due to the fundamentally important role of sialoglycoconjugates in numerous biological phenomena, and the inherent requirement for a variety of development and tissue-specific glycosylation patterns, that the precise control of this process necessitates the existence of multiple sialyltransferase genes that are co-ordinately regulated by a network of multiple promoters.

The expression of the mouse ST6Gal I gene is thought to be regulated by a network of at least 6 putative promoters (Dalziel, M., Huang, R.Y., et al. 2001, Dalziel, M., Lemaire, S., et al. 1999, Wuensch, S.A., Huang, R.Y., et al. 2000). One of these promoters was proposed to regulate the dramatic induction of the ST6Gal I gene in the lactating mammary gland of mice (Dalziel, M., Huang, R.Y., et al. 2001). It was reported that the predominant ST6Gal I mRNA species expressed in the lactating mouse mammary gland is a novel isoform containing the unique untranslated exon (L) derived from the ST6Gal I 5'UTR. Exon (L) is 203bp and located more than 48kb 5' of exon (I). Mammary glands from mature virgin mice exclusively produced the housekeeping form of ST6Gal I mRNA and probes specific to exon (L) confirmed that the exon (L)-containing transcript was only present in the lactating mammary gland and not in other tissues examined (Dalziel, M.,

Huang, R.Y., et al. 2001). This finding suggested the presence of a putative lactogenic promoter upstream of exon (L).

The mammary gland undergoes pronounced physiological changes throughout pregnancy and lactation [for review see (Anderson, S.M., Rudolph, M.C., et al. 2007)] and subsequently during involution when it is remodelled to its pre-pregnant state (Stein, T., Salomonis, N., et al. 2007). Using a novel bioinformatics approach to analyse a microarray time series of the mammary gland pregnancy cycle, Lemay et al. (2007) were able to derive several key principles related to the regulation of this process. Their analysis showed that almost one third of the mammary transcriptome is committed to the cellular structures required for lactation.

Given the importance of transcriptional control in mammary gland processes, as well as in the regulation of sialylation, our aim was to investigate which sialyltransferase genes are expressed in the murine mammary gland during the lactation cycle, in order to understand their role for the suckling young. In addition, we were interested in determining whether the genes encoding key enzymes involved in sialic acid biosynthesis were also up-regulated during lactation to cater for the potential increase in demand for sialic acid.

To investigate the expression of sialic acid biosynthesis and sialyltransferase genes in the mammary gland throughout the pregnancy cycle, we analysed 2 publicly available mouse lactation microarray time series (Rudolph, M.C., McManaman, J.L., et al. 2003, Stein, T., Morris, J., et al. 2004), focusing specifically on genes regulating sialic acid biosynthesis and sialylation. These two datasets were specifically chosen because each incorporated extensive biological replication in their experimental design and both covered a comprehensive, overlapping time course of mouse lactation. Furthermore, both datasets were previously used as the basis for published studies and have been extensively validated (Anderson, S.M., Rudolph, M.C., et al. 2007, Lemay, D., Neville, M., et al. 2007, Rudolph, M.C., McManaman, J.L., et al. 2003, Stein, T., Morris, J., et al. 2004, Stein, T., Salomonis, N., et al. 2007). Here we report that genes encoding enzymes responsible for sialic acid

biosynthesis and transfer are up-regulated prior to lactation and that the observed change in expression is highly correlated with lactation. In addition, we report that the transcription of a further 2 sialyltransferase genes is suppressed during lactation, suggesting a potential involvement in the development of the gland specifically during pregnancy. Finally, we identify a sialyltransferase gene that increases in expression in late pregnancy, with a further increase observed at the onset of involution codes for an enzyme potentially involved in the innate immunity of the mammary gland.

3.2.3 Methods

Microarray study design

Expression analysis focused on time course lactation data previously described in (Rudolph, M.C., McManaman, J.L., et al. 2003), which we refer to as the "primary" set. In brief, RNA was obtained from the mammary glands of FVB mice, isolated, and hybridized onto Affymetrix MG_U74Av2 arrays. The study represented a total of 10 time points across the mammary gland development cycle, with 4 biological replicates at each time point: P1, the day a vaginal plug was observed; P3, pregnancy day 3; P7, pregnancy day 7; P12, pregnancy day 12; P17, pregnancy day 17; P19, pregnancy day 19; L1, early lactation, the first day pups were observed in the cage, L2, lactation day 2; L9, lactation day 9; and I2, two days following pup removal on lactation day 9. Histologically, the proliferative stage of the mammary gland is represented by P1, P3, and P7, the secretory differentiation stage by P12, P17, and P19, early lactation by L1 and L2, full lactation by L9, and involution by I2. The raw CEL files are available from the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/), GEO Series accession number GSE8191.

As a validation tool, a second lactation dataset was utilised that was initially reported in Stein et al. (2004), and which we refer to as the "validation" set. Briefly, mammary gland RNA samples were collected from Balb/C mice, isolated, and hybridized onto Affymetrix MG_U74Av2 chips. A total of 18 time points in mammary development were represented in the study, with 3 biological replicates interrogated at each time point: V6, V10, V12, virgin 6, 10 and 12 weeks; P1, P2, P3, P8.5, P12.5, P14.5, P17.5, pregnancy days 1, 2 , 3, 8.5, 12.5, 14.5 and 17.5; L1, L3, L7, lactation days 1, 3 and 7; I1, I2, I3, I4, I20, involution days 1, 2, 3, 4 and 20. As in the Rudolph study, the Stein et al. data represents mammary gland proliferation (P1, P2, P3, P8.5), secretory differentiation (P12.5, P14.5, P17.5), early lactation (L1, L3) and full lactation (L7). However, it also includes three virgin mammary gland time points and more comprehensively describes involution at days I1, I2, I3, I4 and

I20. The raw CEL files are available from the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/), GEO Series accession number GSE12247.

Microarray quality assessment

The Bioconductor packages "affy", "gcrma" and "affyPLM" were used to produce images, histograms, box plots, degradation plots, and scatter plots to evaluate the quality of the array data.

Microarray data analysis

All microarray data analysis was performed using Bioconductor for R 2.10.0. Firstly, to ensure that the two lactation datasets were comparable, they were combined and normalised using the "gcrma" package. The combined, normalised dataset was subsequently clustered by time points using "hclust" hierarchical clustering.

Both datasets were then analysed independently using the following approach. The raw CEL files were normalised using "gcrma". The complete set of 12,488 genes was then pre-filtered using "mas5calls" – P, Present; M, Marginal; A, Absent – to remove those probes with unreliable or undetectable signals. Probes that did not return a P call for at least 4 of the 40 samples from the Rudolph dataset, and, at least 3 of the 54 samples from the Stein dataset were excluded from further analysis. Following PMA filtering, 7520 probes remained from the Rudolph dataset and 7209 probes from the Stein dataset. The reduced datasets were utilised in the subsequent analysis.

To assess how the genes of interest ranked in terms of differential expression compared to highly differentially expressed genes such as the milk proteins, the Biconductor "timecourse" package was used. This package assesses differential expression by comparing the mean time course profiles allowing for variability both within and between

the time points. The probes remaining from each dataset were ranked based on the Hotelling statistic $\tilde{t}^2$.

To find evidence of lactation-specific differential expression of the genes of interest, the 7000+ probes remaining from each dataset were analysed using the Bioconductor "limma" package. For the Rudolph data, the contrasts of interest were P12 versus L2 and L9 versus I2. To validate these results, similar contrasts from the Stein data; P12.5 versus L3 and L7 versus I1 were chosen. A target gene was deemed to be differentially expressed if, when examining comparable contrasts across the two datasets, the Benjamini and Hochberg (BH) false discovery rate adjusted p-value was less than 0.05 in both datasets and that it showed at least a 2-fold change in expression in at least one of the datasets.

As the Stein data covered a wider range of time points across mammary gland development, additional contrasts were investigated solely on the Stein data; I1 versus I4, I4 versus I20 and V6 versus P3.

Clustering

The 3703 genes from the Rudolph data that were differentially expressed between P12 and L2 with a BH adjusted p-value of less than 0.05 were used for cluster analysis by the Bioconductor "Mfuzz" fuzzy clustering package. The mean normalised expression of the biological replicates was first calculated at each time point. Since the clustering is performed in Euclidean space, the expression values of the genes were then standardised to have a mean value of zero and a standard deviation of one. The data was subsequently checked for probes containing more than 25% missing values, as the clustering algorithm does not allow for them, however, no probes were excluded based on this criteria. Five distinct "Mfuzz" clustering steps were performed using an m-value of 1.1, beginning with 8 clusters and increasing by 2, incrementally, to 16 clusters. The optimal number of clusters (8) was determined empirically based on the observed variability and redundancy

between similar clusters, supported by overlap plots, which indicate how many genes are shared by two clusters and thus describe the similarity between pairs of clusters. A transcript needed to have at least a cluster membership value of 0.75 to be considered for downstream analysis.

Co-expression analysis

The "geneRecommender" Bioconductor package was used to specifically identify genes that are co-expressed with the casein milk proteins. The package ranks the submitted genes according to how strongly they correlate with the selected query genes over the experiments for which the query genes behave in a similar fashion. For our analysis, we used all of the casein probes ("96030_at", "98814_at", "99065_at", "99130_at", "100463_at") present on the MG_U74Av2 array as the query set. The 7520 probes from the Rudolph data remaining following MAS5.0 PMA filtering were submitted to be queried by "geneRecommender", with a return cut-off of the top 300 genes found to be co-expressed with the caseins.

Functional annotation

The clusters generated using "Mfuzz" and the ranked list provided by "geneRecommender" were functionally annotated using Ingenuity Pathways Analysis (IPA) to identify the biological functions and/or diseases that were most significant to them. Molecules from the dataset that were associated with biological functions in Ingenuity's Knowledge Base were considered for the analysis. Right-tailed Fisher's exact test was used to calculate a p-value determining the probability that each biological function assigned to that dataset is due to chance alone. Functional associations were considered significant if their BH corrected p-value was less than 0.05.

3.2.4 Results

We analysed 2 publicly available microarray datasets, which will be referred to as the primary dataset (Rudolph, M.C., McManaman, J.L., et al. 2003) and the validation dataset (Stein, T., Morris, J., et al. 2004). Both datasets represent microarray analysis of the mammary gland in the mouse during pregnancy, lactation and involution. Each dataset was generated using the same Affymetrix mouse array platform, which contains 12,488 probes, representing in excess of 12,000 genes, approximately a third of the genes in the murine genome. Of the probes present on the array, 20 represent 18 unique sialyltransferase genes and a further 3 probes represent genes encoding enzymes involved in sialic acid biosynthesis (Table 1). The sialyltransferases ST6Gal II and ST8Sia VI were not represented on the arrays and are consequently not included in this study. The ST6Gal I gene is represented by 2 probe sets, both binding in different regions of the 3'UTR (Table 1). This set of 23 probes will be henceforth referred to as the target gene set. Our analysis focused on discerning any changes in the expression of these genes between pregnancy and lactation, and lactation and involution.

To confirm parity between the two different datasets, we initially combined the primary and validation data and clustered it across the pooled time course. The results of the hierarchical clustering analysis (Figure 1) show that the individual datasets do not form distinct clusters, but rather, that time points from the same stage of lactation cluster together, regardless of which dataset they originate from. Also, time points from developmental stages when the mammary gland is in a similar morphological state, e.g. early pregnancy and involution, also cluster closely. This indicates that the datasets describe similar expression patterns for the entire probe set and that similar results would be obtained from independent analysis of each dataset, thus presenting a useful means of validation.

Subsequent filtering of the individual datasets to remove any probes with an undetectable or unreliable signal identified ~8000 probes suitable for further analysis in both the primary and validation dataset. Twelve of the 23 probes from the target gene set remained following the filtering step. The target probes that were removed in the filtering step are highlighted in Table 1.

To ascertain the degree of differential expression of the target genes compared with genes known to be significantly differentially expressed during the mammary gland pregnancy, lactation and involution cycle, such as the major milk proteins, the probes in each filtered gene list were ranked based on their differential expression across the entire time course, allowing for variability both within and between time points. The major milk proteins ranked highly in both datasets (Table 2). The probes from the target gene set also ranked relatively highly; only 2 of the target probes remaining in the primary dataset and 3 in the validation dataset were ranked outside of the top 2000 (Table 2).

Table 1

Sialic acid biosynthesis and sialyltransferase genes represented on the Affymetrix mouse U74Av2 array

Affymetrix Probe ID	Symbol	Accession	Description
98593_at	Cmas	AJ006215	cytidine monophospho-N-acetylneuraminic acid synthetase
97924_at	Gne	AJ132236	glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase
104147_at	Nans	AW122052	N-acetylneuraminic acid synthase (sialic acid synthase)
99847_at	St3gal1	X73523	ST3 beta-galactoside alpha-2,3-sialyltransferase 1
100279_at	St3gal2	X76989	ST3 beta-galactoside alpha-2,3-sialyltransferase 2
97778_at	St3gal3	X84234	ST3 beta-galactoside alpha-2,3-sialyltransferase 3
95599_at	St3gal4	D28941	ST3 beta-galactoside alpha-2,3-sialyltransferase 4
98596_s_at	St3gal5	Y15003	ST3 beta-galactoside alpha-2,3-sialyltransferase 5
102208_at	St3gal6	AI153959	ST3 beta-galactoside alpha-2,3-sialyltransferase 6
94431_at	St6gal1	D16106	beta galactoside alpha 2,6 sialyltransferase 1
94432_at	St6gal1	AI117157	beta galactoside alpha 2,6 sialyltransferase 1
100382_at	St6galnac1	Y11274	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 1
99055_at	St6galnac2	X94000	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2
102899_at	St6galnac3	Y11342	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 3
160232_at	St6galnac4	AJ007310	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 4
96682_at	St6galnac4	Y15780	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 4
92403_at	St6galnac5	AB030836	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 5
100369_at	St6galnac6	AB035174	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 6
94164_at	St8sia1	X84235	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 1
100474_at	St8sia2	X99646	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 2
99504_at	St8sia3	X80502	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 3
102318_at	St8sia4	X86000	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 4
99320_at	St8sia5	X98014	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 5

All of the sialic acid biosynthesis and sialyltransferase genes represented on the Affymetrix mouse U74Av2 array used in this analysis. The rows shaded in grey indicate the probes that were removed in the initial filtering step due to undetectable or unreliable signals.

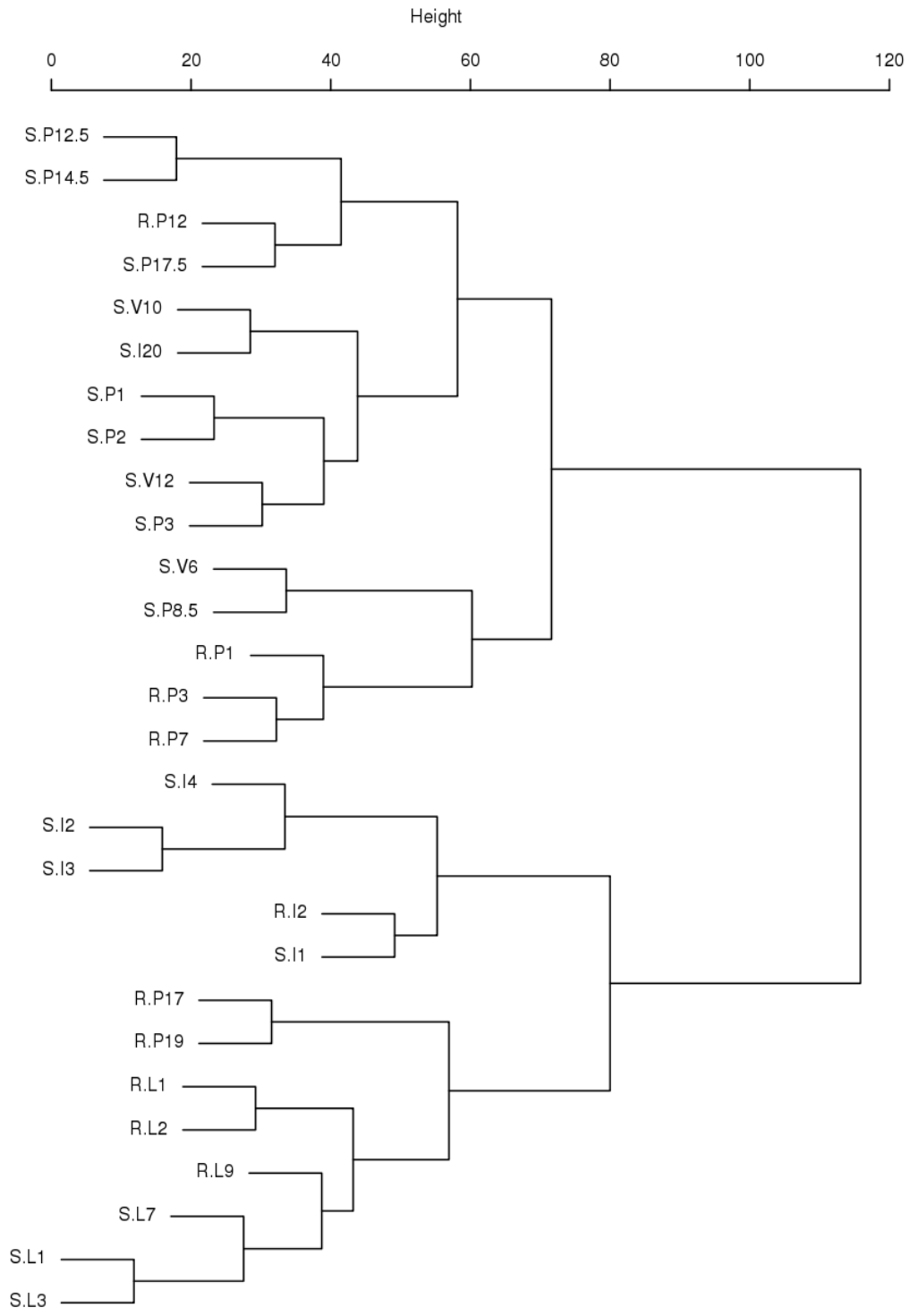


Figure 1: Dendrogram depicting the results of hierarchical clustering of combined primary and validation microarray datasets. Time points from different datasets are denoted by the first letter of the original author's surname; R (Rudolph, M.C., McManaman, J.L., et al. 2003), S (Stein, T., Morris, J., et al. 2004). In general, time points from the same stage of lactation cluster together, regardless of which dataset they originated from. Also, time points from developmental stages where the mammary gland is in a similar morphological state, e.g. early pregnancy and involution, tend to cluster closely.

Table 2a

Major milk proteins and target genes from the primary dataset ranked by overall change in expression throughout the pregnancy, lactation and involution cycle

Rank	Symbol	Description
1	Csn1s2b	casein alpha s2-like B
3	Lalba	lactalbumin, alpha
31	Ltf	lactotransferrin
147	St6gal1	beta galactoside alpha 2,6 sialyltransferase 1
355	St6galnac2	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2
421	St6gal1	beta galactoside alpha 2,6 sialyltransferase 1
504	St6galnac5	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 5
606	Wap	whey acidic protein
742	Gne	glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase
1144	Nans	N-acetylneuraminic acid synthase (sialic acid synthase)
1169	St3gal4	ST3 beta-galactoside alpha-2,3-sialyltransferase 4
1218	St8sia4	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 4
1462	St3gal1	ST3 beta-galactoside alpha-2,3-sialyltransferase 1
1864	St3gal6	ST3 beta-galactoside alpha-2,3-sialyltransferase 6
2184	Csn1s2a	casein alpha s2-like A
2533	Csn2	casein beta
2718	Csn3	casein kappa
3071	Cmas	cytidine monophospho-N-acetylneuraminic acid synthetase
3378	Csn1s1	casein alpha s1
7251	St6galnac3	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 3

Table 2b

Major milk proteins and target genes from the validation dataset ranked by overall change in expression throughout the pregnancy, lactation and involution cycle

Rank	Symbol	Description
1	Lalba	lactalbumin, alpha
2	Csn1s2b	casein alpha s2-like B
4	Wap	whey acidic protein
5	Csn1s2a	casein alpha s2-like A
8	Csn2	casein beta
19	Ltf	lactotransferrin
24	Csn1s1	casein alpha s1
89	St6gal1	beta galactoside alpha 2,6 sialyltransferase 1
136	Csn3	casein kappa
160	Nans	N-acetylneuraminic acid synthase (sialic acid synthase)
453	St6galnac2	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2
469	St6gal1	beta galactoside alpha 2,6 sialyltransferase 1
553	Gne	glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase
576	St8sia4	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 4
1106	St3gal6	ST3 beta-galactoside alpha-2,3-sialyltransferase 6
1228	St3gal1	ST3 beta-galactoside alpha-2,3-sialyltransferase 1
1239	Cmas	cytidine monophospho-N-acetylneuraminic acid synthetase
2718	St3gal4	ST3 beta-galactoside alpha-2,3-sialyltransferase 4
3696	St6galnac5	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 5
6889	St6galnac3	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 3

Differential expression analysis

Analysis focused on significant differential expression of the target gene set between pregnancy and lactation, and lactation and involution. More specifically, we were interested in gene expression differences between the beginning of secretory activation in pregnancy at around day 12 (Lemay, D., Neville, M., et al. 2007) and early lactation, as well as differences between late lactation and involution. Genes exhibiting lactation-specific expression, such as the major milk protein genes, are significantly up-regulated between pregnancy and lactation and subsequently decrease in expression between lactation and involution (Figure 2).

Six probes from the target set, representing GNE, NANS, ST3Gal I, ST3Gal IV, ST6Gal I and ST6GalNAc II (Figure 3), were found to significantly increase in expression between pregnancy and lactation (Table 3). Intriguingly, 2 additional probes from the target set, ST3Gal VI and ST8Sia IV (Figure 3), were found to be significantly down-regulated during the same period (Table 3). When comparing lactation and involution, GNE, NANS, ST3Gal I, ST3Gal IV and ST6Gal I were all significantly down-regulated by the second day of involution (Table 4).

As the validation dataset covered a more comprehensive time series during the involution phase, we further investigated the expression of the target genes during this period solely using the Stein et al. (2004) data. In a comparison of day 7 of lactation with the first day of involution (Table 5), CMAS showed an increase in expression of approximately 2-fold, whilst GNE, NANS and ST6Gal I were all significantly down-regulated. Comparison of day 1 involution with day 4 involution (Table 6), revealed a decrease in expression of CMAS and ST6GalNAc II, as well as a further down-regulation of ST6Gal I. The ST3Gal VI and ST8Sia IV genes, which were down-regulated between pregnancy and lactation, increased in expression by more than 2-fold during this period.

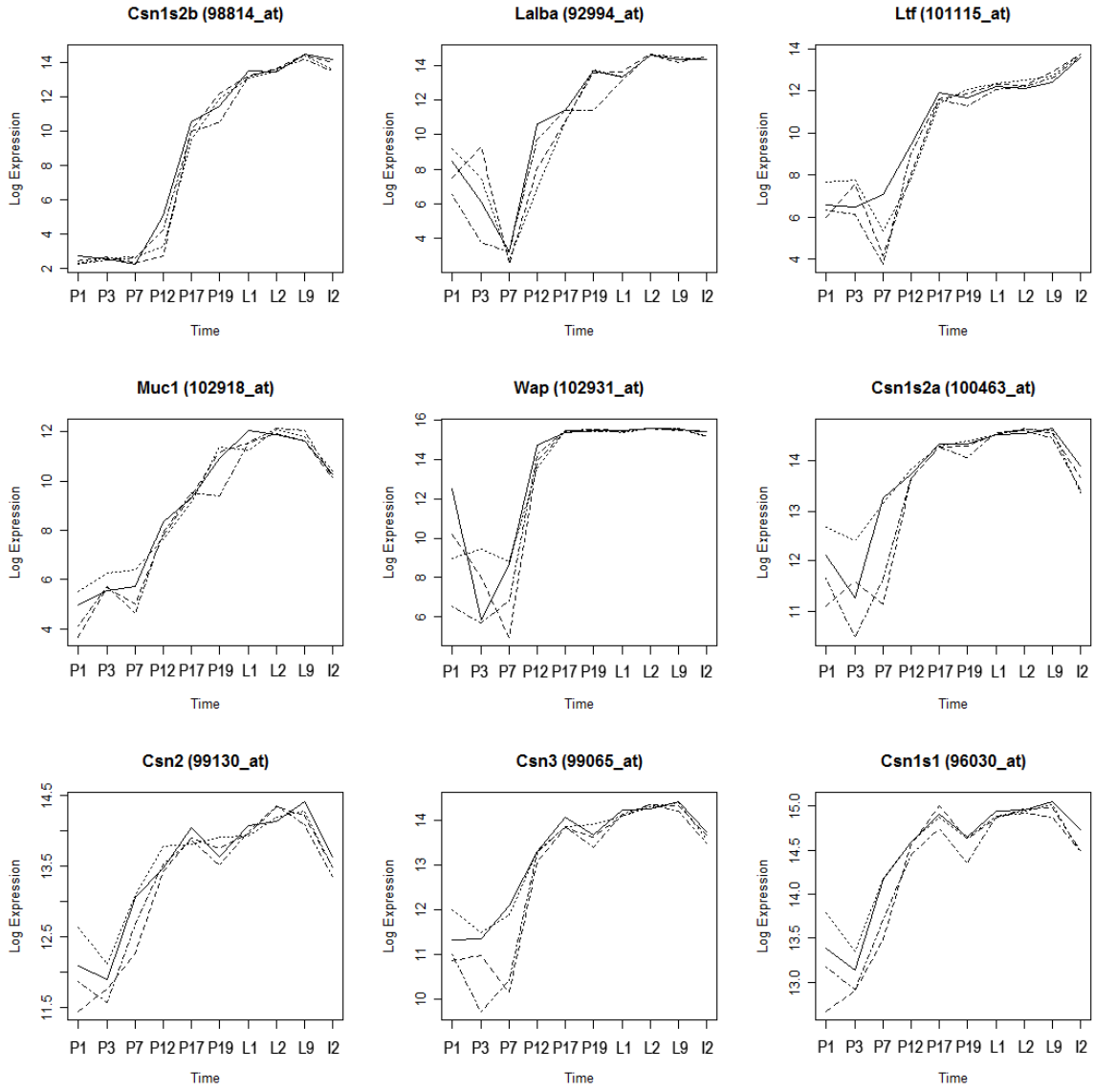


Figure 2a: Expression profiles of the major milk protein genes from the *primary* dataset. The genes exhibit the classical lactation-specific pattern of expression whereby they dramatically increase in expression in mid-pregnancy, remain very highly expressed throughout lactation and gradually decrease in expression during involution.

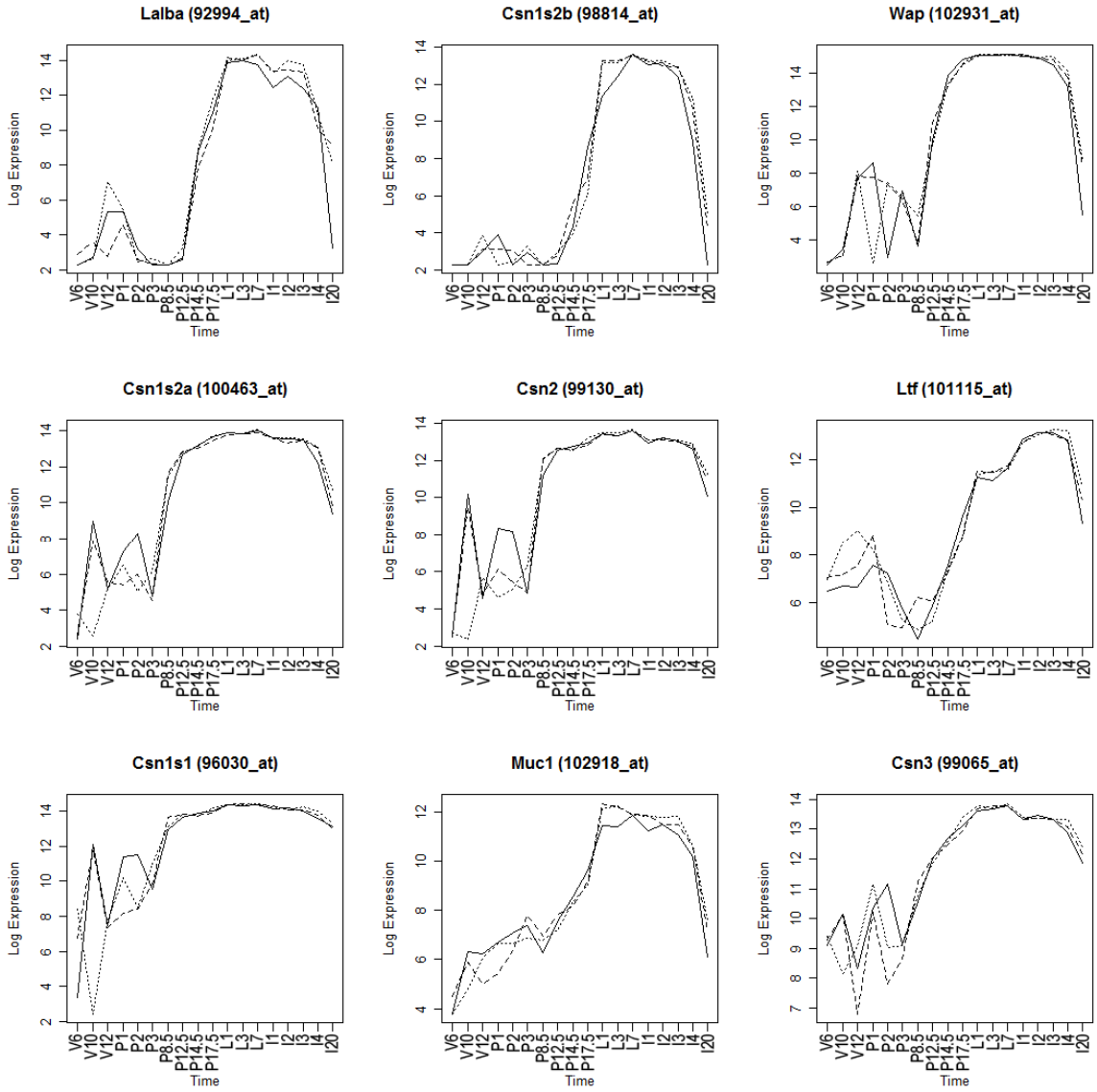


Figure 2b: Expression profiles of the major milk protein genes from the *validation* dataset. The genes exhibit the classical lactation-specific pattern of expression whereby they dramatically increase in expression in mid-pregnancy, remain very highly expressed throughout lactation and gradually decrease in expression during involution.

Table 3

Target genes significantly differentially expressed between pregnancy and lactation

Symbol	Description	Primary Dataset FC	Validation Dataset FC
Gne	glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase	1.88	5.52
Nans	N-acetylneuraminic acid synthase (sialic acid synthase)	4.43	16.02
St3gal1	ST3 beta-galactoside alpha-2,3-sialyltransferase 1	3.15	5.25
St3gal4	ST3 beta-galactoside alpha-2,3-sialyltransferase 4	2.30	1.93
St3gal6	ST3 beta-galactoside alpha-2,3-sialyltransferase 6	-1.85	-3.42
St6gal1	beta galactoside alpha 2,6 sialyltransferase 1	8.66	24.14
St6gal1	beta galactoside alpha 2,6 sialyltransferase 1	12.11	12.06
St6galnac2	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2	8.11	4.74
St8sia4	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 4	-2.07	-2.96

Probes from the target gene set found to be significantly differentially expressed between ~day 12 pregnancy and ~day 2 lactation in both datasets. All genes were significantly differentially expressed with a BH corrected p-value < 0.05 in both the primary and validation datasets. Furthermore, to be deemed biologically significant the genes had to exhibit at least a 2-fold change in expression in at least one of the datasets. Fold change is denoted by FC.

Table 4

Target genes significantly differentially expressed between lactation and involution

Symbol	Description	Primary Dataset FC	Validation Dataset FC
Gne	glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase	-2.85	-4.79
Nans	N-acetylneuraminic acid synthase (sialic acid synthase)	-2.59	-2.29
St3gal1	ST3 beta-galactoside alpha-2,3-sialyltransferase 1	-1.97	-2.11
St3gal4	ST3 beta-galactoside alpha-2,3-sialyltransferase 4	-6.37	-2.37
St6gal1	beta galactoside alpha 2,6 sialyltransferase 1	-7.62	-8.47
St6gal1	beta galactoside alpha 2,6 sialyltransferase 1	-5.02	-5.20

Probes from the target gene set found to be significantly differentially expressed between ~day 9 lactation and ~day 2 involution in both datasets. All genes were significantly differentially expressed with a BH corrected p-value < 0.05 in both the primary and validation datasets. Furthermore, to be deemed biologically significant the genes had to exhibit at least a 2-fold change in expression in at least one of the datasets. Fold change is denoted by FC.

Table 5

Target genes significantly differentially expressed between day seven of lactation and the first day of involution

Symbol	Description	FC
Cmas	cytidine monophospho-N-acetylneuraminic acid synthetase	2.06
Gne	glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase	-5.95
Nans	N-acetylneuraminic acid synthase (sialic acid synthase)	-2.87
St6gal1	beta galactoside alpha 2,6 sialyltransferase 1	-7.53
St6gal1	beta galactoside alpha 2,6 sialyltransferase 1	-7.49

Probes from the target gene set found to be significantly differentially expressed between day seven of lactation and the first day of involution in the validation dataset. All genes were significantly differentially expressed with a BH corrected p-value < 0.05 and exhibited at least a 2-fold change in expression. Fold change is denoted by FC.

Table 6

Target genes significantly differentially expressed between the first day of involution and the fourth day of involution

Symbol	Description	FC
Cmas	cytidine monophospho-N-acetylneuraminic acid synthetase	-3.08
St6gal1	beta galactoside alpha 2,6 sialyltransferase 1	-3.05
St6galnac2	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2	-2.72
St3gal6	ST3 beta-galactoside alpha-2,3-sialyltransferase 6	2.54
St8sia4	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 4	3.08

Probes from the target gene set found to be significantly differentially expressed between the first day of involution and the fourth day of involution in the validation dataset. All genes were significantly differentially expressed with a BH corrected p-value < 0.05 and exhibited at least a 2-fold change in expression. Fold change is denoted by FC.

Expression profile clustering and functional annotation

The expression patterns of all genes found to be differentially expressed between pregnancy and lactation were compared and clustered according to similarity of their expression profiles into 8 general gene expression clusters (Figure 4). The differentially expressed probes from the target gene set grouped into 4 distinct clusters (Table 7). The ST3Gal VI and ST8Sia IV genes both grouped into Cluster 1, The ST6GalNAc II gene in Cluster 2, the ST3Gal I and ST3Gal IV genes in Cluster 3 and GNE, NANS and ST6Gal I in Cluster 7.

The 4 clusters containing the genes of interest were annotated using Ingenuity Pathways Analysis (Ingenuity® Systems, www.ingenuity.com) to identify statistically significant associations between cluster of interest and molecular functions, as well as physiological system development and functions. A comparison of the associations between each cluster of interest and physiological system development and function can be seen in Figure 5. The complete list of functional annotations for each cluster is available in Additional files 1-4.

Reproductive system development and function, a category which encompasses lactation and other mammary gland processes, was most significantly associated with cluster 7, which contained the caseins as well as sialic acid biosynthesis genes GNE and NANS, and sialyltransferase ST6Gal I. This cluster is characterised by an increase in gene expression during pregnancy and lactation followed by a distinct and progressive decline between lactation and involution. Cluster 7 was significantly associated with molecular processes involved in cell morphology, cellular assembly and organisation, molecular transport, protein trafficking and lipid metabolism.

Clusters 1, 2 and 3 were all significantly associated with the physiological process of connective tissue development and function; however, the associations with various molecular and cellular functions involved in this process were unique for each cluster.

Cluster 1 had the strongest association with this category, and its constituent transcripts were largely correlated with differentiation and quantity of adipocytes, deposition of adipose tissue and fibroblast attachment, mobility and binding.

Cluster 3 contained transcripts that increased in expression at, or just prior to parturition, peaked in lactation and were dramatically down-regulated at the onset of involution (Figure 5). This cluster was highly correlated with lipid metabolism, small molecule biochemistry and vitamin and mineral metabolism. Connective tissue development and function were also significantly associated with this cluster.

The genes in Cluster 2 increased in expression in mid-pregnancy, were steadily expressed throughout lactation and up-regulated at the onset of involution. The genes in Cluster 2 also appear to have an involvement in connective tissue development and function state with a proportion of constituent transcripts involved in the morphology, arrest in growth and autophagy of fibroblasts. It is also significantly correlated with cell death and apoptotic processes, as well as the acute phase response. Transcripts involved in cell-to-cell signalling and interaction of various immune cells, such as neutrophils and macrophages, were also represented in this cluster.

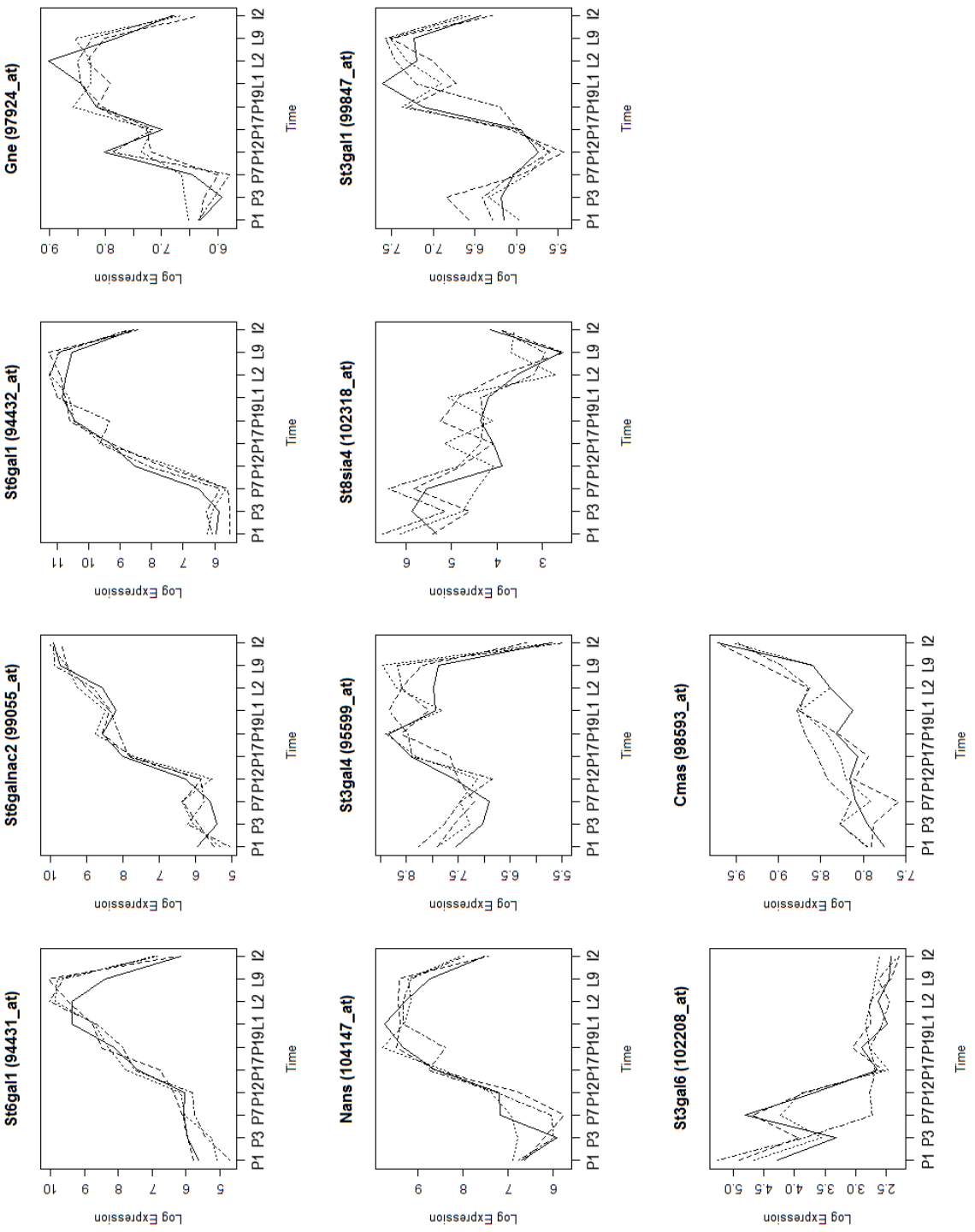


Figure 3a: Expression profiles of sialic acid biosynthesis and sialyltransferase genes found to be differentially expressed between pregnancy and lactation, and lactation and involution in the *primary* dataset. Many of the genes exhibit the classical lactation-specific pattern of expression whereby they dramatically increase in expression in mid-pregnancy, remain very highly expressed throughout lactation and gradually decrease in expression during involution. However, ST3Gal VI and ST8Sia IV display an inverse pattern of expression. Although the 2 ST6Gal I probe sets both target the 3'UTR of the ST6Gal I transcript, and thus should both detect overall ST6Gal I gene expression, the regions of the 3'UTR that the probes target appear to have different GC %, which can affect probe binding and result in slightly different expression profiles.

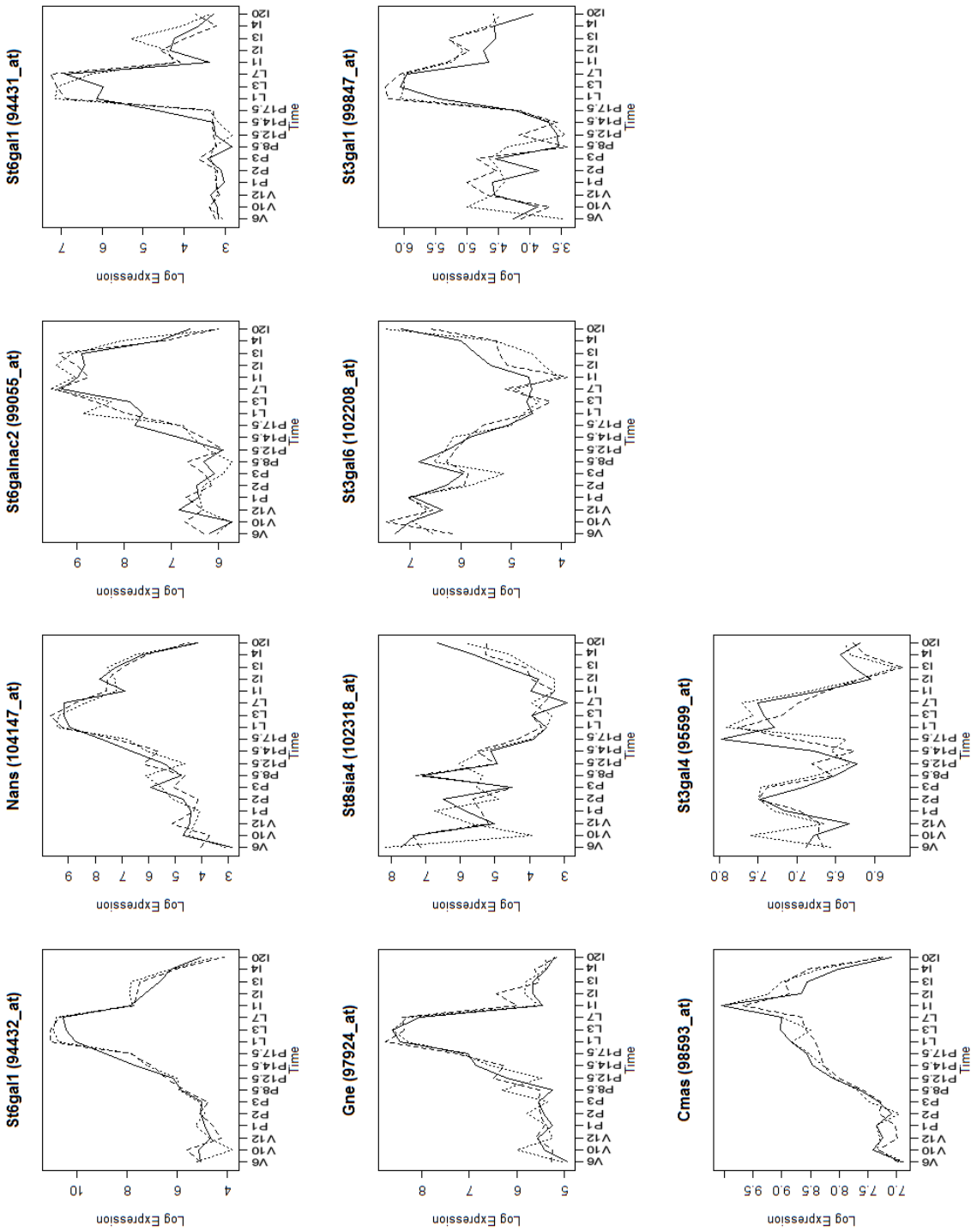


Figure 3b: Expression profiles of sialic acid biosynthesis and sialyltransferase genes found to be differentially expressed between pregnancy and lactation, and lactation and involution in the *validation* dataset. Many of the genes exhibit the classical lactation-specific pattern of expression whereby they dramatically increase in expression in mid-pregnancy, remain very highly expressed throughout lactation and gradually decrease in expression during involution. However, ST3Gal VI and ST8Sia IV display an inverse pattern of expression. Although the 2 ST6Gal I probe sets both target the 3'UTR of the ST6Gal I transcript, and thus should both detect overall ST6Gal I gene expression, the regions of the 3'UTR that the probes target appear to have different GC %, which can affect probe binding and result in slightly different expression profiles.

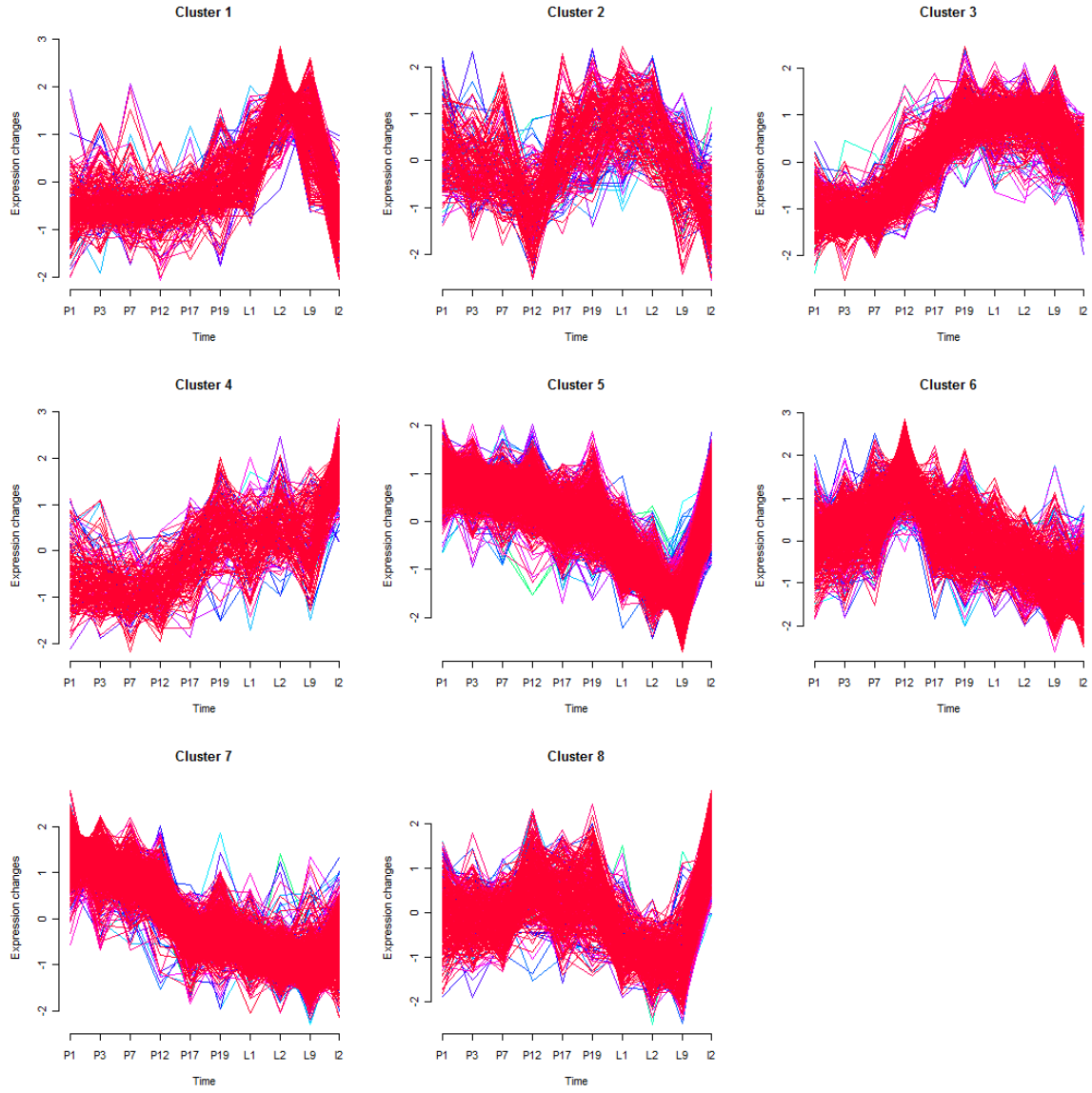


Figure 4: This figure represents the results of fuzzy clustering using 8 clusters and an FCM m-value of 1.1. The 3703 genes from the primary dataset that were differentially expressed between day 12 of pregnancy and day 2 of lactation with a BH adjusted p-value < 0.05 were clustered based on their expression profiles using the Bioconductor "Mfuzz" clustering package. Before clustering, the mean normalised expression of the biological replicates was calculated at each time point and the expression values of the genes were standardised to have a mean value of 0 and a standard deviation of 1.

Table 7

Distribution of target genes and milk proteins between expression profile clusters

Cluster	No. Transcripts	Target Genes	Milk Proteins
1	919	ST3Gal VI ST8Sia IV	-
2	200	ST6GalNAc II	Ltf Lalba
3	327	ST3Gal I ST3Gal IV	-
7	450	GNE NANS ST6Gal I	Csn1s1 Csn1s2a Csn1s2b Csn2 Csn3

Co-regulation analysis

The relationship between the genes of interest and lactation processes was further interrogated using an alternative method for inferring co-regulation of gene expression. This approach involved utilising the casein probes as a "query" to resolve potentially co-regulated candidates from the remaining probe set. The list of the top 300 candidates was functionally annotated and ranked according to their likelihood of co-regulation with the casein genes (Table 8).

The list was highly associated with reproductive system development and function, particularly lactation and related processes. It was also related to cellular assembly and organisation, particularly molecular processes such as budding of vesicles, formation of intracellular junctions, accumulation and engulfment of milk fat globules, lipid metabolism and small molecule biochemistry. The list also included transcripts involved in carbohydrate metabolism, one aspect of which is sialic acid biosynthesis.

Figure 5: Comparison of changes in association between broad physiological system development and function and the expression profile clusters containing our genes of interest. Clusters were annotated using Ingenuity Pathways Analysis (Ingenuity® Systems, www.ingenuity.com). Statistically significant associations between cluster of interest and physiological system development and functions, as well as molecular functional category, were identified. The complete list of functional annotations for each cluster is available in Additional files 1-4.

Table 8

Genes known to play a role in lactation, sialic acid biosynthesis and sialylation ranked based on likelihood of co-regulation with the casein genes

Affymetrix Probe ID	Rank	Status	Symbol	Description
99130_at	1	Within query	Csn2	casein beta
99065_at	2	Within query	Csn3	casein kappa
100463_at	3	Within query	Csn1s2a	casein alpha s2-like A
96030_at	6	Within query	Csn1s1	casein alpha s1
94432_at	15	Not within query	St6gal1	beta galactoside alpha 2,6 sialyltransferase 1
162302_f_at	22	Not within query	Folr1	folate receptor 1 (adult)
103002_at	33	Not within query	B4galt1	UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 1
102931_at	41	Not within query	Wap	whey acidic protein
104147_at	86	Not within query	Nans	N-acetylneuraminic acid synthase (sialic acid synthase)
97924_at	115	Not within query	Gne	glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase
92199_at	187	Not within query	Stat5b	signal transducer and activator of transcription 5B
92994_at	204	Not within query	Lalba	lactalbumin, alpha
95030_at	257	Not within query	Prlr	prolactin receptor
101115_at	297	Not within query	Ltf	lactotransferrin

"Within query" indicates that these were the probes used for the comparison. "Not within query" were probes found to be similar in expression profile to those used for the comparison.

3.2.5 Discussion

Our study has demonstrated that a subset of sialyltransferase genes is differentially expressed during the murine mammary gland pregnancy cycle and that the different sialyltransferases appear to have distinct roles at specific stages in mammary gland development. This suggests that the sialylation requirements of the mammary gland vary throughout pregnancy, lactation and involution and that sialylation is partially regulated at the level of transcription.

We found that, of the 18 sialyltransferase and 3 sialic acid biosynthesis genes investigated in this study, all of the sialic acid biosynthesis genes and 6 sialyltransferase genes were differentially expressed during the lactation cycle. Clustering analysis and subsequent functional annotation highlighted that the differentially expressed sialic acid biosynthesis and sialyltransferase genes are likely to be involved in different mammary gland processes depending on their pattern of expression during the relevant time course.

Two of the 3 genes encoding enzymes uniquely involved in the sialic acid biosynthesis pathway, GNE and NANS, were represented on the mouse Affymetrix array. These two genes, GNE and NANS, (Tanner, M.E. 2005) exhibited similar expression profiles. The genes significantly increased in expression in mid-pregnancy, were highly expressed during lactation and were then subsequently down-regulated at the onset of involution. GNE, the bifunctional enzyme UDP-GlcNAc 2-epimerase/ManNAc kinase, first catalyses the epimerisation of UDP-GlcNAc to ManNAc and subsequently phosphorylates ManNAc to ManNAc-6-P. The epimerisation step is the first committed, and rate limiting, step in sialic acid biosynthesis (Gal, B., Ruano, M.J., et al. 1997), making GNE an important regulator of sialylation. Keppler et al. (1999) demonstrated that GNE activity can be controlled at the level of transcription and that this can influence the sialylation and function of particular cell-surface molecules expressed on B-cells and myeloid cells. NANS, N-acetylneuraminic acid synthase, is responsible for converting MaNAc-6-P to

NeuNAc-9-P. In the mouse, it is ubiquitously expressed in various tissues (Nakata, D., Close, B.E., et al. 2000). It is likely that the up-regulation of GNE and NANS during lactation occurs to allow for the increased demand for sialic acid in downstream sialylation.

Cytidine monophospho-N-acetylneuraminic acid synthetase (CMAS) is not directly involved in sialic acid biosynthesis but is responsible for activating sialic acid for use by sialyltransferases by converting it to CMP-sialic acid (Kean, E.L. 1991). Therefore, it is integral to regulating the rate and extent of sialylation (Münster, A.-K., Eckhardt, M., et al. 1998). Furthermore, CMAS is predominantly localised in the nucleus implying a potential role of transporting sialic acid into the nucleus and CMP-sialic acid out of the nucleus in the regulation of sialylation (Münster, A.-K., Eckhardt, M., et al. 1998). Interestingly, the expression profile of CMAS did not correlate with that observed for GNE and NANS. CMAS gradually increased in expression throughout the course of pregnancy and lactation and the difference in expression levels between lactation and mid-pregnancy was not found to be statistically significant. However, a significant, greater than 2-fold, change in expression was observed between lactation and involution, followed by a down-regulation between the first and fourth day of involution. This difference in expression profile between GNE and NANS, and CMAS, implies that increasing the rate and extent of overall sialylation cannot simply be achieved by increasing transcription of all relevant genes. GNE is allosterically inhibited by CMP-sialic acid, but not free sialic acid (Tanner, M.E. 2005). Thus, we propose that the gradual increase in CMAS expression, combined with the temporary sequestration of activated CMP-sialic acid in the nucleus, serves to maximise sialic acid production in the cytosol whilst delivering a steadily increasing supply of CMP-sialic acid throughout lactation that gradually declines in involution.

ST3Gal I mediates the transfer of sialic acid to Galβ1,3GalNAc terminal disaccharides of O-linked carbohydtrates on glycoproteins and glycolipids via an α2,3 linkage (Takashima,

S. 2008). ST3Gal IV, on the other hand, prefers the Galβ1,4GlcNAc or Galβ1,3GlcNAc structure of glycoproteins and oligosaccharides as an acceptor (Takashima, S. 2008). Additionally, ST3Gal I shows a restricted pattern of expression in mice, whilst ST3Gal IV is more widely expressed (Kono, M., Ohyama, Y., et al. 1997). Our analysis showed that both ST3Gal I and ST3Gal IV were up-regulated in lactation, compared with pregnancy and involution and their expression profiles clustered with genes involved in fatty acid and cholesterol synthesis and calcium regulation.

Although ST6Gal I was up-regulated earlier in pregnancy than ST3Gal I and ST3Gal IV, it was the only other sialyltransferase gene found to be expressed in a lactation-specific manner. Its expression profile was similar to that observed for the caseins and other major milk protein genes. ST6Gal I is generally expressed in a variety of tissues (Aasheim, H.C., Aas-Eng, D.A., et al. 1993, Dalziel, M., Huang, R.Y., et al. 2001, Mercier, D., Wierinckx, A., et al. 1999, Wang, X., Vertino, A., et al. 1993), with particularly high expression observed in the liver (Dalziel, M., Huang, R.Y., et al. 2001, Dalziel, M., Lemaire, S., et al. 1999). Although a golgi transmembrane protein, a soluble form of ST6Gal I has been purified from bovine colostrum and was established to be the product of proteolytic cleavage of the Golgi enzyme (Paulson, J.C., Beranek, W.E., et al. 1977). ST6Gal I preferentially sialylates either type 2 (Galβ1,4GlcNAc) free disaccharides or the *N*-acetyllactosamine termini of *N*- or *O*-linked oligosaccharides via an α2,6 linkage (Harduin-Lepers, A., Vallejo-Ruiz, V., et al. 2001). Dalziel et al. (2001) demonstrated that the murine ST6Gal I gene increased in expression between pregnancy and lactation through the recruitment of a novel 5'UTR exon, implying the existence of an upstream lactogenic promoter, which is in agreement with the current study.

Only a modest number of free oligosaccharide structures have been identified in the milk of mice and rats. A review by Urashima et al. (2001) described 4 different oligosaccharide structures in rat milk, whilst a study by Duncan et al. (2009) detected primarily α2,3 and

α2,6-sialyllactose in the milk of rats. In the same study, the authors reported α2,3-sialyllactose to be the dominant isoform, and that it was present at 10 times the concentration observed for α2,6-sialyllactose. The authors also found that both α2,3-sialyllactose and α2,6-sialyllactose increased in abundance over the first week of lactation, after which the α2,3 form began to steadily decline, whilst the concentration of the α2,6 form remained constant until weaning. As lactose (Galβ1,4Glc) is a poor acceptor for both ST3Gal and ST6Gal sialyltransferases, except at very high concentrations (Paulson, J.C., Rearick, J.I., et al. 1977, Sherblom, A.P. and Bourassa, C.R. 1983), and given that it is a disaccharide that is unique to milk which is only detectable just prior to parturition (Anderson, S.M., Rudolph, M.C., et al. 2007), we have concluded that the up-regulation of specific ST3Gal and ST6Gal sialyltransferase genes in the lactating murine mammary gland is most likely required for the formation of α2,3 and α2,6-sialyllactose in appropriate ratios, which was recently confirmed by Fuhrer et al. (2010). This is also consistent with the results of the study by Duncan et al. (2009) who showed that overall milk sialic acid content, primarily as α2,3-sialyllactose, correlated with the expression of catabolic genes in the neonate colon, whilst sialic acid biosynthesis genes are only up-regulated in the neonate once the dietary source is diminished. Therefore, our findings are consistent with those of Duncan et al. (2009) and suggest that sialic acid biosynthesis and sialyltransferase gene regulation in the mothers' mammary gland, and the regulation of sialic acid biosynthesis and catabolism genes in the digestive tract of the neonate are complementary processes that ultimately serve to deliver milk sialic acid to the neonate so that it can be digested and used as a nutrient. Furthermore, it has been suggested that by delivering sialic acid via oligosaccharides may also serve a decoy function to prevent pathogens from infecting the immunologically fragile neonate by limiting pathogen-host interactions (Newburg, D.S., Ruiz-Palacios, G.M., et al. 2005).

ST6GalNAc II was the only member of the GalNAc α2,6-sialyltransferase (ST6GalNAc) family found to be significantly differentially expressed in the mouse mammary gland

throughout pregnancy, lactation and involution. Six members of the ST6GalNAc family have been identified in the mouse and can be classified into 2 subgroups, based on their amino acid sequence similarities, substrate specificities and gene structures (Takashima, S. 2008). ST6GalNAc I and II belong to a smaller subfamily and catalyse the addition of sialic acid to similar substrates, primarily using GalNAc-, Galβ1,3GalNAc- and Siaα2,3Galβ1,3GalNAc-structures on O-glycans of glycoproteins as acceptors (Harduin-Lepers, A., Vallejo-Ruiz, V., et al. 2001, Takashima, S. 2008). Kono et al. (2000) demonstrated that ST6GalNAc II exhibits activity towards asialo ovine submaxillary mucin, suggesting a potential role in the sialylation of mammary gland mucin.

Mucins are glycoproteins that contain a significant amount (>30%) of carbohydrate. They can be secreted, like in the intestine, nasopharyngeal and reproductive tracts or they can be transmembrane proteins (Gendler, S.J. and Spicer, A.P. 1995). Mammary gland mucins are of the latter variety. They are found in the plasma membrane on the apical surface of lactating epithelial cells and are secreted into milk as part of the milk fat globule (Patton, S., Gendler, S.J., et al. 1995). Mucin isolated from bovine milk fat globules was found to contain in excess of 50% carbohydrate, including 30% sialic acid. Due to their size, copious glycosylation and negative charge conferred by high sialic acid content, mucins create a physical barrier on the surface of epithelia that protects the cell from invading microorganisms (Gendler, S.J. and Spicer, A.P. 1995). It has also been suggested that the negative charge of cell-surface mucins may be responsible for maintaining the integrity of secretory systems by preventing adherence in ducts and alveoli so they remain open (Patton, S. 1999, Patton, S., Gendler, S.J., et al. 1995). Furthermore, ingested milk fat globule mucins have been implicated in the protection of the neonate's gastrointestinal tract by acting as competitive inhibitors against pathogens (Schroten, H., Hanisch, F.G., et al. 1992). Therefore, mucins may have multiple functions including protecting the mammary gland from infection, preventing adherence in ducts and alveoli to ensure they remain open, and acting as "decoys" for microorganisms invading the neonatal gut. Thus,

it is likely that the observed changes in ST6GalNAc II gene expression are linked to the mucin sialylation requirements of the murine mammary gland.

MUC1 is a highly evolutionarily conserved mammary epithelial transmembrane mucin (Patton, S., Gendler, S.J., et al. 1995). A study by Parry et al. (1992) demonstrated that the MUC1 gene is expressed in the mouse mammary gland throughout development but that its levels are significantly increased in mid-pregnancy and lactation. They also showed, using CID-9 mammary epithelial cell culture, that MUC1 gene expression was controlled by lactogenic hormones prolactin, insulin and hydrocortisone. Furthermore, Parry et al. (1992) observed a progressive increase in MUC1 sialylation during pregnancy and lactation which they suggested could be indicative of hormonal induction of a sialyltransferase during the course of pregnancy. The ST6GalNAc II gene is dramatically up-regulated in mid to late pregnancy and remains high throughout lactation, which correlates with the pattern of expression for the MUC1 gene observed in this study and reported by Parry et al. (1992) (Figure 2). This is consistent with a potential increase in demand for mucin sialylaton during lactation that would be required for maintaining the complex secretory structures of the mammary gland, protecting externally exposed ductal surfaces from infection and potentially providing a beneficial milk constituent to the neonate. Unlike the other sialylatransferase genes discussed, ST6GalNAc II remains highly expressed during the first four days of involution, which is correlated with the MUC1 gene expression profile and coincides with the requirement for mucin sialylation in involution when, due to the cessation of milk removal, the mammary gland is most susceptible to infections such as mastitis (O'Toole, D.K. 1995).

ST8Sia IV and ST3Gal VI were the only significantly differentially expressed sialyltransferase genes actively down-regulated during lactation and highly expressed in early pregnancy and late involution. ST8Sia IV transfers sialic acid from CMP-sialic acid to terminal sialic acid residues on oligosaccharides via an α2,8-linkage (Takashima, S.

2008) and is most widely studied for its role in synthesising polysialic acid on neural cell adhesion molecules (NCAMs) (Bruses, J.L. and Rutishauser, U. 2001). It has been demonstrated that the highly negative charge of extended polysialic acid chains causes sufficient repulsion between neighbouring membranes to counteract the binding interactions of NCAMs (Johnson, C.P., Fragneto, G., et al. 2005, Johnson, C.P., Fujimoto, I., et al. 2005). The mammary epithelium is extremely permeable throughout pregnancy due to leaky tight junctions that are sealed at lactogenesis (Linzell, J.L. and Peaker, M. 1971, Nguyen, D., Parlow, A., et al. 2001). It is conceivable that the leakiness of tight junctions during pregnancy may be related to the presence of polysialic acid in the alveolar epithelium. The gradual reduction of ST8Sia IV gene expression throughout pregnancy, with a more pronounced decrease occurring from ~day 12, may result in sufficient decrease of polysialic acid to allow closure of tight junctions at lactogenesis. Nguyen et al. (2001) demonstrated that progesterone withdrawal is the trigger for tight junction closure, which implies that progesterone may be involved in the regulation of ST8Sia IV gene expression in the murine mammary gland. ST3Gal VI forms an α2,3 linkage with the terminal type 2, Galβ1,4GlcNAc disaccharide of glycoproteins and glycolipids (Takashima, S. 2008). It is involved in the synthesis of sialyl-paragloboside, a precursor of sialyl-Lewis X determinant (Okajima, T., Fukumoto, S., et al. 1999). Sialyl-LeX determinants are ligands for at least 3 selectins (E-, P-, and L-selectins), which are cell adhesion molecules implicated in the recruitment of leukocytes to lymphoid tissues. Therefore, it is possible that ST3Gal VI is involved in the synthesis of ligands for immune cell recruitment to mammary tissue. Also, given that the ST3Gal VI gene is most highly expressed in virgin mammary gland and begins to decrease very early in pregnancy and recovers in late involution, it would appear that its associated function is most important when the mammary gland is not secreting milk.

Overall, our study supports a multi-faceted role for sialylation in the mammary gland and demonstrates that it is tightly regulated at the level of gene transcription. In summary, we

have described a coordinated system of expression of sialic acid biosynthesis and sialyltransferase genes in the murine mammary gland, which complements findings made concerning overall sialic acid content of rodent milk and the expression of sialic acid biosynthesis and catabolism genes in the digestive tract of the rodent neonate. We have also suggested that differential expression of the ST6GalNAc II gene is likely to be related to the sialylation of milk mucins, which has implications in maintenance of secretory system integrity during pregnancy and lactation and gland health in involution. Furthermore, we proposed that the expression profiles of ST8Sia IV and ST3Gal VI suggest that they may be involved in tight junction formation and leukocyte recruitment, respectively.

3.2.6 References

Aasheim HC, Aas-Eng DA, Deggerdal A, Blomhoff HK, Funderud S, Smeland EB. 1993. Cell-specific expression of human β-galactoside α2,6-sialyltransferase transcripts differing in the 5' untranslated region. *Eur. J. Biochem.*, 213:467-475.

Anderson SM, Rudolph MC, McManaman JL, Neville MC. 2007. Key stages in mammary gland development. Secretory activation in the mammary gland: it's not just about milk protein synthesis! *Breast Cancer Res*, 9:204.

Bruses JL, Rutishauser U. 2001. Roles, regulation, and mechanism of polysialic acid function during neural development. *Biochimie*, 83:635-643.

Coppa GV, Gabrielli O, Giorgi P, Catassi C, Montanari MP, et al. 1990. Preliminary study of breastfeeding and bacterial adhesion to uroepithelial cells. *The Lancet; London*, 335:569.

Dall'Olio F. 2000. The sialyl-α2,6-lactosaminyl-structure: Biosynthesis and functional role. *Glycoconj. J.*, 17:669-676.

Dalziel M, Huang RY, Dall'Olio F, Morris JR, Taylor-Papadimitriou J, Lau JTY. 2001. Mouse ST6Gal sialyltransferase gene expression during mammary gland lactation. *Glycobiology*, 11:407-412.

Dalziel M, Lemaire S, Ewing J, Kobayashi L, Lau JT. 1999. Hepatic acute phase induction of murine β-galactoside α2,6 sialyltransferase (ST6Gal I) is IL-6 dependent and mediated by elevation of Exon H-containing class of transcripts. *Glycobiology*, 9:1003-1008.

Duncan PI, Raymond F, Fuerholz A, Sprenger N. 2009. Sialic Acid Utilisation and Synthesis in the Neonatal Rat Revisited. *PLoS One*, 4:e8241.

Fuhrer A, Sprenger N, Kurakevich E, Borsig L, Chassard C, Hennet T. 2010. Milk sialyllactose influences colitis in mice through selective intestinal bacterial colonization. *The Journal of Experimental Medicine*, 207:2843-2854.

Gal B, Ruano MJ, Puente R, GarciaPardo LA, Rueda R, Gil A, Hueso P. 1997. Developmental changes in UDP-N-acetylglucosamine 2-epimerase activity of rat and guinea-pig liver. *Comp. Biochem. Phys. B*, 118:13-15.

Gendler SJ, Spicer AP. 1995. Epithelial Mucin Genes. *Annual Review of Physiology*, 57:607-634.

Hanasaki K, Powell LD, Varki A. 1995. Binding of human plasma sialoglycoproteins by the B cell-specific lectin CD22. Selective recognition of immunoglobulin M and haptoglobin. *J. Biol. Chem.*, 270:7543-7550.

Harduin-Lepers A, Mollicone R, Delannoy P, Oriol R. 2005. The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. *Glycobiology*, 15:805-817.

Harduin-Lepers A, Recchi MA, Delannoy P. 1995. 1994, the year of sialyltransferases. *Glycobiology*, 5:741-758.

Harduin-Lepers A, Vallejo-Ruiz V, Krzewinski-Recchi MA, Samyn-Petit B, Julien S, Delannoy P. 2001. The human sialyltransferase family. *Biochimie*, 83:727-737.

Johnson CP, Fragneto G, Konovalov O, Dubosclard V, Legrand JF, Leckband DE. 2005. Structural studies of the neural-cell-adhesion molecule by X-ray and neutron reflectivity. *Biochemistry*, 44:546-554.

Johnson CP, Fujimoto I, Rutishauser U, Leckband DE. 2005. Direct evidence that neural cell adhesion molecule (NCAM) polysialylation increases intermembrane repulsion and abrogates adhesion. *J Biol Chem*, 280:137-145.

Kean EL. 1991. Sialic acid activation. *Glycobiology*, 1:441-447.

Keppler OT, Hinderlich S, Langner J, Schwartz-Albiez R, Reutter W, Pawlita M. 1999. UDP-GlcNAc 2-Epimerase: A Regulator of Cell Surface Sialylation. *Science*, 284:1372-1376.

Kono M, Ohyama Y, Lee Y-C, Hamamoto T, Kojima N, Tsuji S. 1997. Mouse β-galactoside α2,3-sialyltransferases: comparison of in vitro substrate specificities and tissue specific expression. *Glycobiology*, 7:469-479.

Kono M, Tsuda T, Ogata S, Takashima S, Liu H, Hamamoto T, Itzkowitz SH, Nishimura S, Tsuji S. 2000. Redefined Substrate Specificity of ST6GalNAc II: A Second Candidate Sialyl-Tn Synthase. *Biochemical and Biophysical Research Communications*, 272:94-97.

Lemay D, Neville M, Rudolph M, Pollard K, German J. 2007. Gene regulatory networks in lactation: identification of global principles using bioinformatics. *BMC Systems Biology*, 1:56.

Linzell JL, Peaker M. 1971. The permeability of mammary ducts. *J Physiol*, 216:701-716.

Martin-Sosa S, Martin MJ, Garcia-Pardo LA, Hueso P. 2003. Sialyloligosaccharides in human and bovine milk and in infant formulas: variations with the progression of lactation. *J. Dairy Sci.*, 86:52-59.

McVeagh P, Brand-Miller J. 1997. Human milk oligosaccharides: only the breast. *J. Paediatr. Child Health*, 33:281-286.

Mercier D, Wierinckx A, Oulmouden A, Gallet PF, Palcic MM, Harduin-Lepers A, Delannoy P, Petit JM, Leveziel H, Julien R. 1999. Molecular cloning, expression and exon/intron organization of the bovine β-galactoside α2,6-sialyltransferase gene. *Glycobiology*, 9:851-863.

Mouricout M, Petit JM, Carias JR, Julien R. 1990. Glycoprotein glycans that inhibit adhesion of Escherichia coli mediated by K99 fimbriae: treatment of experimental colibacillosis. *Infect. Immun.*, 58:98-106.

Münster A-K, Eckhardt M, Potvin B, Mühlenhoff M, Stanley P, Gerardy-Schahn R. 1998. Mammalian cytidine 5′-monophosphate N-acetylneuraminic acid synthetase: A nuclear protein with evolutionarily conserved structural motifs. *Proceedings of The National Academy of Sciences of The United States of America*, 95:9140-9145.

Nakata D, Close BE, Colley KJ, Matsuda T, Kitajima K. 2000. Molecular Cloning and Expression of the Mouse N-Acetylneuraminic Acid 9-Phosphate Synthase Which Does Not Have Deaminoneuraminic Acid (KDN) 9-Phosphate Synthase Activity. *Biochemical and Biophysical Research Communications*, 273:642-648.

Newburg DS, Ruiz-Palacios GM, Morrow AL. 2005. Human milk glycans protect infants against enteric pathogens. *Annu Rev Nutr*, 25:37-58.

Nguyen D, Parlow A, Neville M. 2001. Hormonal regulation of tight junction closure in the mouse mammary epithelium during the transition from pregnancy to lactation. *J Endocrinol*, 170:347-356.

O'Toole DK. 1995. Factors inhibiting and stimulating bacterial growth in milk: a historical perspective. Advances in Applied Microbiology Volume 40. San Diego: Academic Press. p. 45-47.

Okajima T, Fukumoto S, Miyazaki H, Ishida H, Kiso M, Furukawa K, Urano T, Furukawa K. 1999. Molecular Cloning of a Novel α2,3-Sialyltransferase (ST3Gal VI) That Sialylates Type II Lactosamine Structures on Glycoproteins and Glycolipids. *Journal of Biological Chemistry*, 274:11479-11486.

Parry G, Li J, Stubbs J, Bissell M, Schmidhauser C, Spicer A, Gendler S. 1992. Studies of Muc-1 mucin expression and polarity in the mouse mammary gland demonstrate developmental regulation of Muc-1 glycosylation and establish the hormonal basis for mRNA expression. *J Cell Sci*, 101:191-199.

Patton S. 1999. Some practical implications of the milk mucins. *J Dairy Sci*, 82:1115-1117.

Patton S, Gendler SJ, Spicer AP. 1995. The epithelial mucin, MUC1, of milk, mammary gland and other tissues. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes*, 1241:407-423.

Paulson JC, Beranek WE, Hill RL. 1977. Purification of a sialyltransferase from bovine colostrum by affinity chromatography on CDP-agarose. *J. Biol. Chem.*, 252:2356-2362.

Paulson JC, Rearick JI, Hill RL. 1977. Enzymatic properties of beta-d-galactoside alpha-2-6 sialyltransferase from bovine colostrum. *Journal of Biological Chemistry*, 252:2363-2371.

Rudolph MC, McManaman JL, Hunter L, Phang T, Neville MC. 2003. Functional development of the mammary gland: use of expression profiling and trajectory clustering to reveal changes in gene expression during pregnancy, lactation, and involution. *J Mammary Gland Biol Neoplasia*, 8:287-307.

Schroten H, Hanisch FG, Plogmann R, Hacker J, Uhlenbruck G, Nobis-Bosch R, Wahn V. 1992. Inhibition of adhesion of S-fimbriated Escherichia coli to buccal epithelial cells by human milk fat globule membrane components: a novel aspect of the protective function of mucins in the nonimmunoglobulin fraction. *Infect. Immun.*, 60:2893-2899.

Sherblom AP, Bourassa CR. 1983. Specificity of submaxillary gland sialyltransferases. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 761:94-102.

Stein T, Morris J, Davies C, Weber-Hall S, Duffy M-A, Heath V, Bell A, Ferrier R, Sandilands G, Gusterson B. 2004. Involution of the mouse mammary gland is associated with an immune cascade and an acute-phase response, involving LBP, CD14 and STAT3. *Breast Cancer Res.*, 6:R75 - R91.

Stein T, Salomonis N, Gusterson B. 2007. Mammary Gland Involution as a Multi-step Process. *Journal of Mammary Gland Biology and Neoplasia*, 12:25-35.

Svensson EC, Soreghan B, Paulson JC. 1990. Organization of the β-galactoside α2,6-sialyltransferase gene. Evidence for the transcriptional regulation of terminal glycosylation. *J. Biol. Chem.*, 265:20863-20868.

Takashima S. 2008. Characterization of Mouse Sialyltransferase Genes: Their Evolution and Diversity. *Bioscience, biotechnology, and biochemistry*, 72:1155-1167.

Takashima S, Kono M, Kurosawa N, Yoshida Y, Tachida Y, Inoue M, Kanematsu T, Tsuji S. 2000. Genomic organization and transcriptional regulation of the mouse GD3 synthase gene (ST8Sia I): Comparison of genomic organization of the mouse sialyltransferase genes. *Journal of Biochemistry*, 128:1033-1043.

Tanner ME. 2005. The enzymes of sialic acid biosynthesis. *Bioorganic Chemistry*, 33:216-228.

Tsuji S. 1996. Molecular Cloning and Functional Analysis of Sialyltransferases. *J Biochem (Tokyo)*, 120:1-13.

Urashima T, Saito T, Nakamura T, Messer M. 2001. Oligosaccharides of milk and colostrum in non-human mammals. *Glycoconjugate Journal*, 18:357-371.

Varki A. 1993. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*, 3:97-130.

Varki A. 2007. Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. *Nature*, 446:1023-1029.

Wang B, Brand-Miller J. 2003. The role and potential of sialic acid in human nutrition. *Eur. J. Clin. Nutr.*, 57:1351-1369.

Wang B, Brand-Miller J, McVeagh P, Petocz P. 2001. Concentration and distribution of sialic acid in human milk and infant formulas. *Am. J. Clin. Nutr.*, 74:510-515.

Wang X, O'Hanlon TP, Young RF, Lau JTY. 1990. Rat β-galactoside α2,6-sialyltransferase genomic organization: alternate promoters direct the synthesis of liver and kidney transcripts. *Glycobiology*, 1:25-31.

Wang X, Vertino A, Eddy RL, Byers MG, Jani-Sait SN, Shows TB, Lau JT. 1993. Chromosome mapping and organization of the human β-galactoside α2,6-sialyltransferase gene. Differential and cell-type specific usage of upstream exon sequences in B-lymphoblastoid cells. *J. Biol. Chem.*, 268:4355-4361.

Wen DX, Svensson EC, Paulson JC. 1992. Tissue-specific alternative splicing of the β-galactoside α2,6- sialyltransferase gene. *J. Biol. Chem.*, 267:2512-2518.

Wuensch SA, Huang RY, Ewing J, Liang X, Lau JT. 2000. Murine B cell differentiation is accompanied by programmed expression of multiple novel β-galactoside α2, 6-sialyltransferase mRNA forms. *Glycobiology*, 10:67-75.

4 Chapter Four

Having determined that key sialic acid biosynthesis genes and certain sialyltransferase genes were expressed in an apparently lactation-specific manner, the response of these genes to lactogenic hormones in mouse mammary explant culture was investigated. Dalziel et al. (2001) had already demonstrated that a lactation-specific transcript existed for ST6Gal I containing a novel 5'UTR-derived exon, exon (L), suggesting the presence of a lactogenic promoter upstream and a potentially important role in lactation. If the sialic acid biosynthesis and sialyltransferase genes found to be differentially expressed in chapter 3 were found to be controlled by lactogenic hormones, it would further support their importance in the mouse mammary gland, particularly during lactation. It would also identify the genes that were likely to either contain lactogenic promoters or are under the influence of other lactation-related regulatory elements.

This chapter is presented in manuscript format in preparation for journal review.

4.1 Declaration for thesis chapter four

4.1.1 Declaration by candidate

In the case of Chapter 4, the nature and extent of my contribution to the work was the following:

Nature of contribution	**Extent of contribution (%)**
For this chapter I performed the laboratory experiments and was responsible for analysis of the data, interpretation of the results and the writing of the manuscript.	80%

The following co-authors contributed to the work. Co-authors who are students at Monash University must also indicate the extent of their contribution in percentage terms:

Name	**Nature of contribution**	**Extent of contribution (%)** **(student co-authors only)**
Dr. Keith W. Savin	Dr. Savin aided in interpretation of results and writing of the manuscript.	
Dr. Julie A. Sharp	Dr. Sharp aided in the explant culture, interpretation of results and writing of the manuscript.	
Prof. Kevin R. Nicholas	Prof. Nicholas initially collected the mammary glands from the mice and aided in the preparation of the explants, interpretation of the results and writing of the manuscript.	

Candidate's Signature		**Date**

4.1.2 Declaration by co-authors

The undersigned hereby certify that:

(1) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.
(2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
(3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
(4) there are no other authors of the publication according to these criteria;
(5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
(6) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

Location(s)	**DPI, Biosciences Research Division, Bundoora VIC 3083**

Signature 1	[redacted]	**Date** 4-4-11
Signature 2	[redacted]	26/3/2011
Signature 3	[redacted]	6/4/11

4.2 Expression of sialic acid biosynthesis and sialyltransferase genes is regulated by lactogenic hormones in mouse mammary gland tissue culture

Jovana Maksimovic[1,2], Keith W. Savin[2], Julie A. Sharp[3], Kevin R. Nicholas[3]

1. Biosciences Research Division, Department of Primary Industries, Melbourne 3083 Australia
2. Centre for Reproduction and Development, Monash Institute of Medical Research, Monash University, Melbourne 3168 Australia
3. Institute for Technology Research and Innovation, Deakin University, Geelong 3214 Australia

4.2.1 Abstract

Milk sialylglycoconjugates are believed to protect the gastrointestinal tract of the suckling neonate by competitively binding to invading pathogens and promoting the growth of beneficial flora, and their potential role in post partum brain development is highly relevant to human infant nutrition. Sialylated compounds are synthesised by sialyltransferases by the addition of sialic acid to the terminal position of sugar chains on glycoproteins, glycolipids and free di- or oligosaccharides. The expression of sialyltransferases is primarily controlled at the level of transcription and is expedited through the use of several tissue and development-specific promoters. Previous studies have shown that certain sialyltransferases are differentially expressed in the mouse mammary gland during lactation, suggesting lactogenic control. The current study has examined the expression of two key sialic acid biosynthesis genes and several sialyltransferase genes in mouse mammary explants cultured with the lactogenic hormones insulin, hydrocortisone and prolactin. It demonstrates that the expression of sialic acid biosynthesis genes, UDP-GlcNAc 2-epimerase/ManNAc kinase (GNE) and N-acetylneuraminic acid synthase (NANS), increases in the presence of lactogenic hormones, as do the ST3Gal I and ST6Gal I sialyltransferase genes. In contrast, the ST8Sia IV gene was found to decrease in expression in explants cultured with lactogenic hormones. Several putative binding sites for lactation-related transcription factors were also indentified in the genes that responded to the hormones. Taken collectively, the data presented supports lactogenic control of sialylation in the mouse mammary gland during lactation.

4.2.2 Introduction

Sialic acid is a nine-carbon, acidic monosaccharide and in mammalian species it is generally found at the non-reducing terminus of either oligosaccharides, glycolipids or glycoproteins [reviewed in (Harduin-Lepers, A., Recchi, M.A., et al. 1995, Tsuji, S. 1996)]. Sialylglycoconjugates are integral to a variety of biological processes such as cell-cell interaction, cell migration, adhesion and metastasis (Dall'Olio, F. 2000, Hanasaki, K., Powell, L.D., et al. 1995, Varki, A. 1993, Varki, A. 2007). However, they are required to perform a different function in the lactating mammary gland. Milk sialylglycoconjugates are believed to protect the gastrointestinal tract of the suckling neonate by competitively binding to invading pathogens and promoting the growth of beneficial flora (Coppa, G.V., Gabrielli, O., et al. 1990, Martin-Sosa, S., Martin, M.J., et al. 2003, McVeagh, P. and Brand-Miller, J. 1997, Mouricout, M., Petit, J.M., et al. 1990, Wang, B., Brand-Miller, J., et al. 2001), and their potential role in post partum brain development is highly relevant to human infant nutrition (McVeagh, P. and Brand-Miller, J. 1997, Wang, B. and Brand-Miller, J. 2003). Many studies have examined the sialylglycoconjugate content of the milk of various species, however, studies examining the gene expression and regulatory processes responsible for the final suite of structures found in milk are limited.

Sialyltransferases are Golgi membrane-bound glycosyltransferases that transfer the sialic acid moiety from the CMP-sialic acid donor to the carbohydrate acceptor groups of glycoproteins, glycolipids and free oligosaccharides (Harduin-Lepers, A., Mollicone, R., et al. 2005). The relative abundance of various sialyltransferases present in a particular cell type is largely regulated at the transcriptional level (Svensson, E.C., Soreghan, B., et al. 1990, Takashima, S., Kono, M., et al. 2000, Wang, X., O'Hanlon, T.P., et al. 1990, Wen, D.X., Svensson, E.C., et al. 1992) . This is achieved via the coordinated use of numerous tissue and development stage-specific promoters that are present in the large 5'UTR of each sialyltransferase gene, generating several transcript families that differ only in the structure of their 5' untranslated regions (Svensson, E.C., Soreghan, B., et al. 1990,

Takashima, S., Kono, M., et al. 2000, Wang, X., O'Hanlon, T.P., et al. 1990, Wen, D.X., Svensson, E.C., et al. 1992).

Twenty mouse sialyltransferase genes have been identified and cloned to date (Takashima, S. 2008). In his review, Takashima (2008) suggested that this large number of sialyltransferase genes, which are regulated by a network of multiple promoters, is essential to providing the necessary variety of development and tissue-specific glycosylation patterns. Sialyltransferases can be divided into 4 groups; ST6Gal, ST6GalNAc, ST3Gal and ST8Sia. Group membership is based on the type of linkage formed between the sialic acid and the acceptor, and the type of acceptor that is preferred (Harduin-Lepers, A., Mollicone, R., et al. 2005, Harduin-Lepers, A., Vallejo-Ruiz, V., et al. 2001).

The murine ST6Gal I gene is likely regulated by a suite of at least 6 candidate promoters (Dalziel, M., Huang, R.Y., et al. 2001, Dalziel, M., Lemaire, S., et al. 1999, Wuensch, S.A., Huang, R.Y., et al. 2000). One of these 6 putative promoters is proposed to be exclusively regulated by lactogenic hormones, which is consistent with the observed increase of ST6Gal I gene expression in the mammary gland between pregnancy and lactation (Dalziel, M., Huang, R.Y., et al. 2001). The latter report showed that the most abundant ST6Gal I transcript expressed in the lactating mouse mammary gland was a novel species containing a previously uncharacterised untranslated exon, named exon (L). Exon (L) is 203bp long and is found >48kb upstream of exon (I). The exon (L)-containing transcript was only present in the lactating mammary gland and not virgin mammary tissue, and was not found in any of the other tissues analysed. These observations led the authors to conclude that a lactogenic promoter must exist upstream of exon (L), although they did not find the usual transcription factor consensus binding motifs that are known to be associated with lactation (Dalziel, M., Huang, R.Y., et al. 2001).

The ultimate structure of the adult mammary gland, and its ability to lactate, is the culmination of an intricately controlled network of pathways and processes driven by numerous steroid and peptide hormones (Topper, Y.J. and Freeman, C.S. 1980). *In vitro* mammary explant culture has proven to be an ideal model for studying the regulation of milk protein gene expression, as the tissue essentially retains its three-dimensional structure and responds to regulatory stimuli. Furthermore, it is an isolated system that enables tight control over the endocrine environment (Menzies, K., Lee, H., et al. 2010, Nagaiah, K., Bolander, F.F., Jr., et al. 1981, Ono, M. and Oka, T. 1980). For these reasons, the mammary explant culture model has been extensively used to mimic the lactogenic process and study endocrine regulation of milk protein gene expression (Dodd, S.C., Forsyth, I.A., et al. 1994, Menzies, K., Lee, H., et al. 2010, Nagaiah, K., Bolander, F.F., Jr., et al. 1981, Nicholas, K.R. and Tyndale-Biscoe, C.H. 1985, Ono, M. and Oka, T. 1980, Sankaran, L. and Topper, Y.J. 1984, Sheehy, P.A., Della-Vedova, J.J., et al. 2004). Mammary explant culture studies in the mouse and rat have demonstrated that insulin (I), hydrocortisone (F), and prolactin (P) are absolutely necessary for the induction of the casein genes (Bolander, F.F., Jr., Nicholas, K.R., et al. 1981, Kulski, J.K., Nicholas, K.R., et al. 1983, Nagaiah, K., Bolander, F.F., Jr., et al. 1981). Studies in mice have identified that hydrocortisone and prolactin are involved in both transcription of the casein gene and stabilisation of the mRNA (Bolander, F.F., Jr., Nicholas, K.R., et al. 1981, Choi, K.M., Barash, I., et al. 2004, Kulski, J.K., Nicholas, K.R., et al. 1983, Nagaiah, K., Bolander, F.F., Jr., et al. 1981). Furthermore, mammary explant culture experiments performed by Nicholas et al. (1983) demonstrated that insulin was also essential for casein synthesis and Chomczynski et al. (1984) went on to show that insulin was essential for transcription of the β-casein gene, but not for stabilisation of its mRNA. A more recent study has shown that in the mouse mammary gland insulin plays an important role in milk protein synthesis at multiple levels (Menzies, K., Lee, H., et al. 2010).

Prolactin regulation of milk protein gene expression is expedited by its initiation of the JAK/STAT signalling cascade (Pellegrini, S. and Dusanter-Fourt, I. 1997). Numerous studies have confirmed that the hormonal and developmental control of milk protein gene expression is expedited by proximal and distal composite response elements that contain binding sites for several different transcription factors [for review see (Rosen, J.M., Wyszomierski, S.L., et al. 1999)]. These composite response elements are comprised of binding sequences for both positive and negative regulatory elements, which enables them to incorporate the effects of signal transduction pathways via protein:DNA and protein:protein interactions, resulting in distinctive temporal and spatial expression of target genes (Jiang, J. and Levine, M. 1993). STAT5, GR, C/EBPβ, YY1 and OCT-1 have all been demonstrated to be involved, in various combinations, in the transcriptional regulation of milk protein gene expression (Dong, B., Huang, C., et al. 2009, Dong, B. and Zhao, F.Q. 2007, Doppler, W., Welte, T., et al. 1995, Lechner, J., Welte, T., et al. 1997, Li, S. and Rosen, J.M. 1994, Li, S. and Rosen, J.M. 1995, Meier, V.S. and Groner, B. 1994, Raught, B., Khursheed, B., et al. 1994, Raught, B., Liao, W.S., et al. 1995, Schmitt-Ney, M., Doppler, W., et al. 1991, Seagroves, T.N., Krnacik, S., et al. 1998, Wakao, H., Gouilleux, F., et al. 1994, Zhao, F.Q., Zheng, Y., et al. 2004).

In a previous study (Maksimovic, manuscript in preparation) we demonstrated that several sialic acid biosynthesis and sialyltransferase genes are differentially expressed in the mouse mammary gland throughout pregnancy, lactation and involution. Many of the differentially expressed genes showed a distinctly lactation-specific pattern of expression that was correlated with the expression of the milk protein genes. The regulation of sialic acid biosynthesis and sialylation in the mammary gland is not understood, although the distinctly lactation-specific pattern of expression that we previously observed for several relevant genes suggests that it may involve similar endocrine and autocrine mechanisms that regulate milk protein gene expression. Here we report the expression of these sialic acid biosynthesis and sialyltransferase genes in mouse mammary explants cultured with

lactogenic hormones I, F and P. Finally, we present data on potential proximal and distal transcription factor binding sites that may be involved in the lactation-specific regulation of these genes.

4.2.3 Methods

Animals

Balb/c mice were obtained at day 12 of pregnancy from Monash Animal Services. The mice were euthanised and their mammary glands excised under sterile conditions. Mammary tissue was pooled into 2 groups (4 mice per group) for experiment 1 (E1) and 3 groups (4 mice per group) for experiment 2 (E2). The tissue was then either stored at −80 °C for analysis of target gene expression at day 12 of pregnancy or used immediately for tissue culture. All animal experimentation was performed according to the procedures of The Deakin University Animal Ethics Committee.

Tissue culture

Mammary explants were cultured as previously described by Nicholas and Tyndale-Biscoe (1985). Briefly, the explants were incubated in 5 ml Medium 199 at 37°C (5% CO_2) either without the addition of growth factors or hormones (NH), or with a combination of bovine insulin (I; 250 ng/ml, Sigma #I-5500), hydrocortisone (F; 50 ng/ml, Sigma #H-4001) and ovine prolactin (P; 250 ng/ml, National Hormone and Pituitary Program USA-oPRL-21). The media was changed every second day. The treatments added to the media included NH and IFP for E1 and NH, I, F, P, IF, FP, IFP for E2. The various treatments were performed in duplicate for each pool of explants.

Reverse Transcription PCR

Total RNA (2ug) was used to perform first strand cDNA synthesis with the SuperScript VILO cDNA Synthesis Kit, according to the instructions supplied by Invitrogen. The primers used for RT-PCR analysis of the targets genes are described in Table 1. Identical reactions, using the Promega GoTaq Green Master Mix, were set up for each target. Each final 25µl reaction mixture consisted of 400 nM forward primer, 400 nM reverse primer, 2µl of 1 in 10 diluted cDNA and 12.5µl of the 2X GoTaq Master Mix. A single reaction

tube for each target was removed from the thermal cycler during the middle of the extension phase at 28, 30, 32, 34, 36, 38 and 40 cycles. The PCR products were subsequently separated by electrophoresis on a 2% agarose gel and visually inspected to determine the range of cycles encompassing the exponential amplification phase. PCR parameters for the MJ Research PTC-225 Pertier Thermal-Cycler were: 94^{o}C for 2 min, 40 cycles at 94^{o}C for 15 sec then 60^{o}C for 1 min and 72^{o}C for 5 min. Once the optimal number of cycles for quantification was determined, the final PCR amplification was performed for each target on all samples. Each final 25µl reaction mixture consisted of 400 nM forward primer, 400 nM reverse primer, 2µl of 1 in 10 diluted cDNA and 12.5µl of the 2X GoTaq Master Mix. PCR parameters for the MJ Research PTC-225 Pertier Thermal-Cycler were: 94^{o}C for 2 min, 28 cycles at 94^{o}C for 15 sec then 60^{o}C for 1 min and 72^{o}C for 5 min. The PCR products were subsequently separated by electrophoresis on a 2% agarose gel and the image acquired using the PerkinElmer ProXPRESS 2D Proteomic Imaging System with top illumination, 100µm resolution, 480/30 excitation, 530/30 emission and an 8 second exposure. The resulting image was analysed using ImageJ (http://rsb.info.nih.gov/ij/) (Abramoff, M.D., Magelhaes, P.J., et al. 2004) and the expression of the targets normalised to the levels of GAPDH.

Statistical analysis

Significant differences between gene expression in the mammary explants cultured either with hormones (IFP) or without hormones (NH) were determined using paired t-tests. For all analyses, a P-value of less than 0.05 was interpreted as statistically significant. All results are expressed as the mean of available replicates, and the mean±S.E.M. is included in the figures.

Transcription factor binding site analysis

To characterise putative promoter regions of genes found to respond to lactogenic hormones, 2kb of mouse genomic sequence upstream of, and including, the first exon contained in the transcript of interest was used for transcription factor binding site (TFBS) analysis. The Match program by Biobase (http://www.gene-regulation.com/), which searches TRANSFAC using positional weight matrices (PWM) to locate TFBS, was used to identify putative TFBS in our sequences of interest. A core similarity of 0.95 and matrix similarity value of 0.9 was used as a cut off for the TRANSFAC search. Transcription start sites were predicted for the putative promoter region using TSSG from Softberry (http://linux1.softberry.com/).

To find putative distal STAT5 binding sites for genes found to respond to lactogenic hormones, the entire genomic sequence with an additional 20kb of 5' sequence was used for Match TFBS searching. Two different PWM, STAT and STATx, were used to identify potential STAT5 binding sites. A core and matrix similarity cut off of 0.75 and 0.95 was used for STAT, and 0.75 and 0.97, respectively, for STATx.

Table 1

Primers used for RT-PCR

Gene	Forward Primer	Reverse Primer	Product (bp)
GAPDH	CCGCCTGGAGAAACCTGCCA	TGAGAGCAATGCCAGCCCCG	170
LALBA	ACGGCAGCACAGAGTACGGA	AGAGGCATTCCTGACCACGGG	298
GNE	TGGCCTTGCAGAGGGAAGCAAA	CCCAAAGCAGTTCCAGCTGTTCGT	169
NANS	AATGCACCATGCCGCTGGAACT	TCTTCTGAAACTTAGCGCAGTCGGC	174
ST6Gal I	TGGGGCACCTACAGACAACTTCCA	TCTGGCTTCTGATACCACTGCGGA	186
ST6Gal I exon (L)	GCGTGTACTGCTCTTGCCTGCT	AGGTCTGGAGGCCAAGGCTCAT	175
ST3Gal I	TGGCTCGGCTTCCTGAGATGAAGA	TGAAAGGGCATCTCCTGGAAGGCA	113
ST3Gal IV	AAGAAGCGGGTGCGAAAAGGCT	AGGTGTAGAGCCAAGGTGATGGCTA	197
ST3Gal VI	AGGTTGCTGCCATGTCCGGT	TGCAGTTTTGAAAGGGCCGAGCT	94
ST8Sia IV	TCCGGAAGGCTGGCTCCACC	GTCCGGCGTCTGTCCAGCAC	184

All primers were designed to either span introns or at least one primer was designed across an intron/exon boundary in order to allow for identification of any genomic DNA contamination based on discrimination by product size.

4.2.4 Results

Endocrine regulation of sialic acid biosynthesis and sialyltransferase gene expression in mammary explants from mice

Using RT-PCR, RNA isolated from mouse mammary explants cultured in either no hormone (NH) or a combination of insulin, hydrocortisone and prolactin (IFP) for 4 days was analysed for expression of the α-lactalbumin (LALBA) gene and certain sialic acid biosynthesis and sialyltransferase genes (Figure 1-Figure 4). LALBA gene expression was significantly greater in explants cultured in media with IFP for 2-4 days than those cultured in NH (Figure 1), which is consistent with published studies (Palmiter, R.D. 1969). This confirmed the efficacy of the culture system and demonstrated that densitometric analysis of RT-PCR results was adequate for the detection of differences in gene expression between treatments.

The GNE and NANS genes code for UDP-N-acetyl-2-epimerase/N-acetylmannosamine kinase and N-acetylneuraminic acid synthase, respectively, which are 2 of the 3 enzymes uniquely involved in sialic acid biosynthesis [reviewed in (Tanner, M.E. 2005)]. To investigate whether lactogenic hormones play a role in regulating sialic acid biosynthesis in the lactating mouse mammary gland, we assayed the expression of the NANS and GNE genes in response to IFP. When comparing levels of GNE and NANS gene expression in mouse mammary explants cultured in IFP for 4 days and those cultured in NH, expression of both genes appeared higher in IFP treated explants than in those cultured in NH at days 2, 3 and 4 of culture (Figure 2). However, a statistically significant ($p\text{-value}<0.05$) difference in expression was only observed at day 3 and 4 for the GNE gene, and only day 4 for the NANS gene.

The ST6Gal I gene has previously been demonstrated to increase in expression in the mouse mammary gland between pregnancy and lactation, and that increase was attributed to the recruitment of a lactation-specific transcript containing the previously

uncharacterised 5'UTR exon, exon (L) (Dalziel, M., Huang, R.Y., et al. 2001). Although the authors did not find evidence of a STAT5 consensus motif in the putative proximal promoter region upstream of exon (L), the observed change in gene expression pattern suggested the likely involvement of lactogenic hormones in the regulation of its expression. Thus, we investigated both the overall expression of the ST6Gal I gene and the exon (L) containing transcript in response to the addition of IFP to the mammary explant culture media (Figure 3). Overall ST6Gal I gene expression was significantly increased in explants cultured in IFP, in comparison to NH, on days 2, 3 and 4 of culture. An extremely low level of the ST6Gal I exon (L) transcript was observed in the uncultured mammary tissue from day 12 pregnant mice (D0). However, after 4 days of culture there was a significant increase in expression of the ST6Gal I exon (L) transcript only in explants cultured in IFP.

The ST3Gal I and ST3Gal IV sialyltransferase genes have been shown to express in a lactation-specific manner (Fuhrer, A., Sprenger, N., et al. 2010), suggesting possible endocrine regulation by lactogenic hormones. A statistically significant difference in ST3Gal I gene expression between IFP and NH cultured explants was seen at days 2, 3 and 4 of culture (Figure 4). In contrast, ST3Gal IV gene expression did not vary significantly between explants cultured in IFP and NH (Figure 4).

The ST8Sia IV and ST3Gal VI sialyltransferase genes are highly expressed in early pregnancy and involution, but are actively down-regulated in mid to late pregnancy and throughout lactation (data not shown). When mouse mammary explants were cultured in IFP and NH for 4 days, expression of the ST8Sia IV gene was consistently lower in the presence of IFP than NH after 2, 3 and 4 days of culture (Figure 4). At day 4, the reduction of ST8Sia IV gene expression in IFP compared to NH is statistically significant ($p < 0.05$). Expression of the ST3Gal VI gene was lowest in the uncultured mammary tissue taken from mice in mid-pregnancy. ST3Gal VI gene expression subsequently increased in the

mouse mammary explants cultured in both NH and IFP for 4 days however, there was no statistically significant difference in its expression level between the tissue cultured in IFP versus NH on any of the days that explants were harvested (Figure 4).

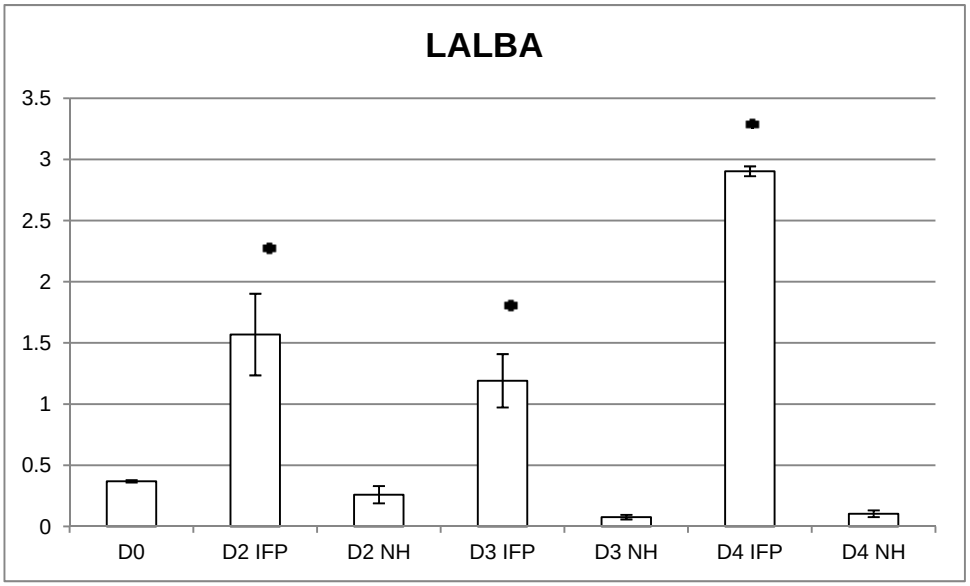


Figure 1: RT-PCR analysis of expression of the α-lactalbumin (LALBA) gene in mammary explants. Mammary explants were assayed for LALBA gene expression in uncultured mammary tissue, and after 2, 3 and 4 days of culture in the treatments indicated. The hormone treatments are abbreviated as follows: D0, uncultured tissue, I (insulin), F (hydrocortisone), P (prolactin), NH (no hormone). LALBA gene expression is expressed as relative to %GAPDH. A statistically significant difference in expression between NH and IFP on each day is denoted by a "*".

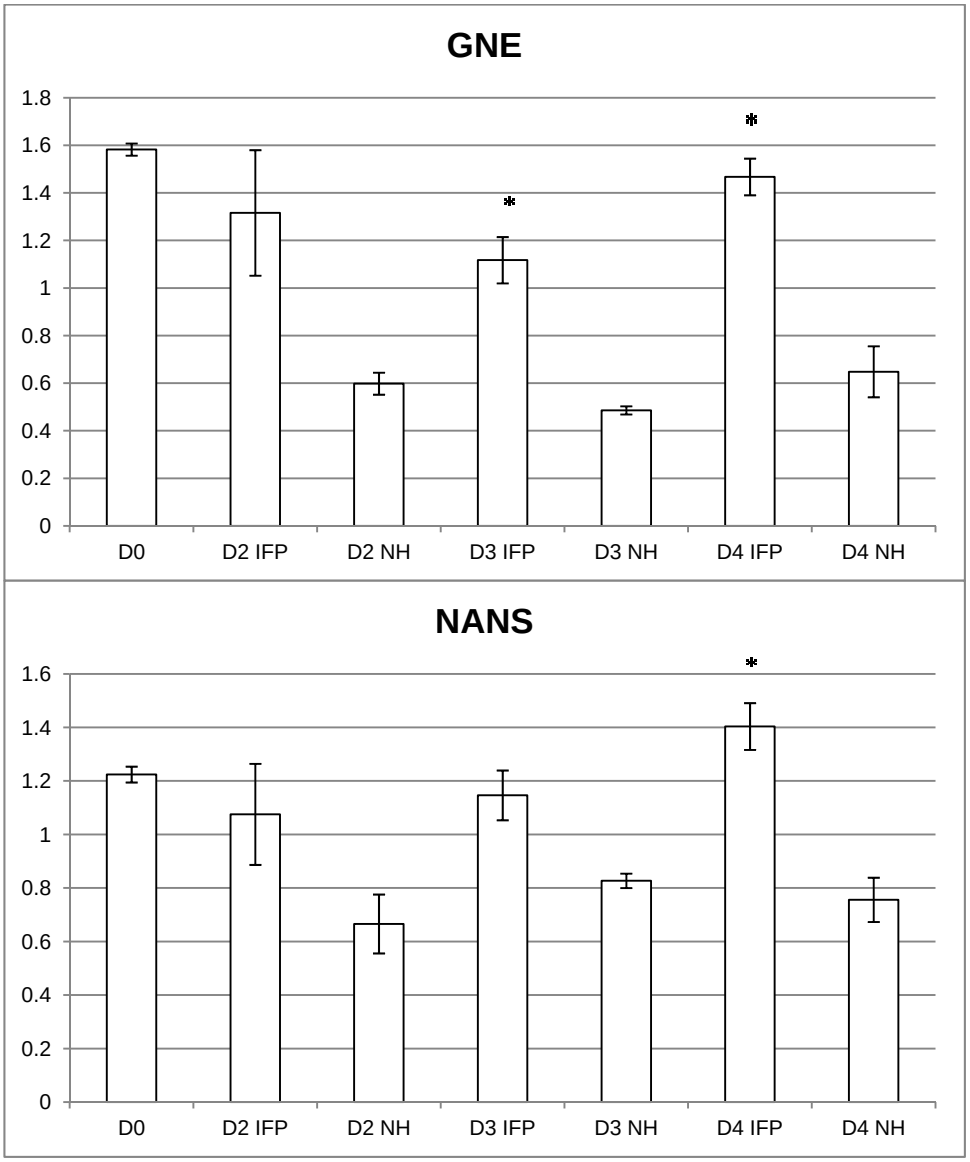


Figure 2: RT-PCR analysis of expression of sialic acid biosynthesis genes in mammary explants. Mammary explants were assayed for GNE and NANS gene expression in uncultured mammary tissue, and after 2, 3 and 4 days of culture in the treatments indicated. The hormone treatments are abbreviated as follows: D0, uncultured tissue, I (insulin), F (hydrocortisone), P (prolactin), NH (no hormone). Target gene expression is expressed as relative to %GAPDH. A statistically significant difference in expression between NH and IFP on each day is denoted by a "*".

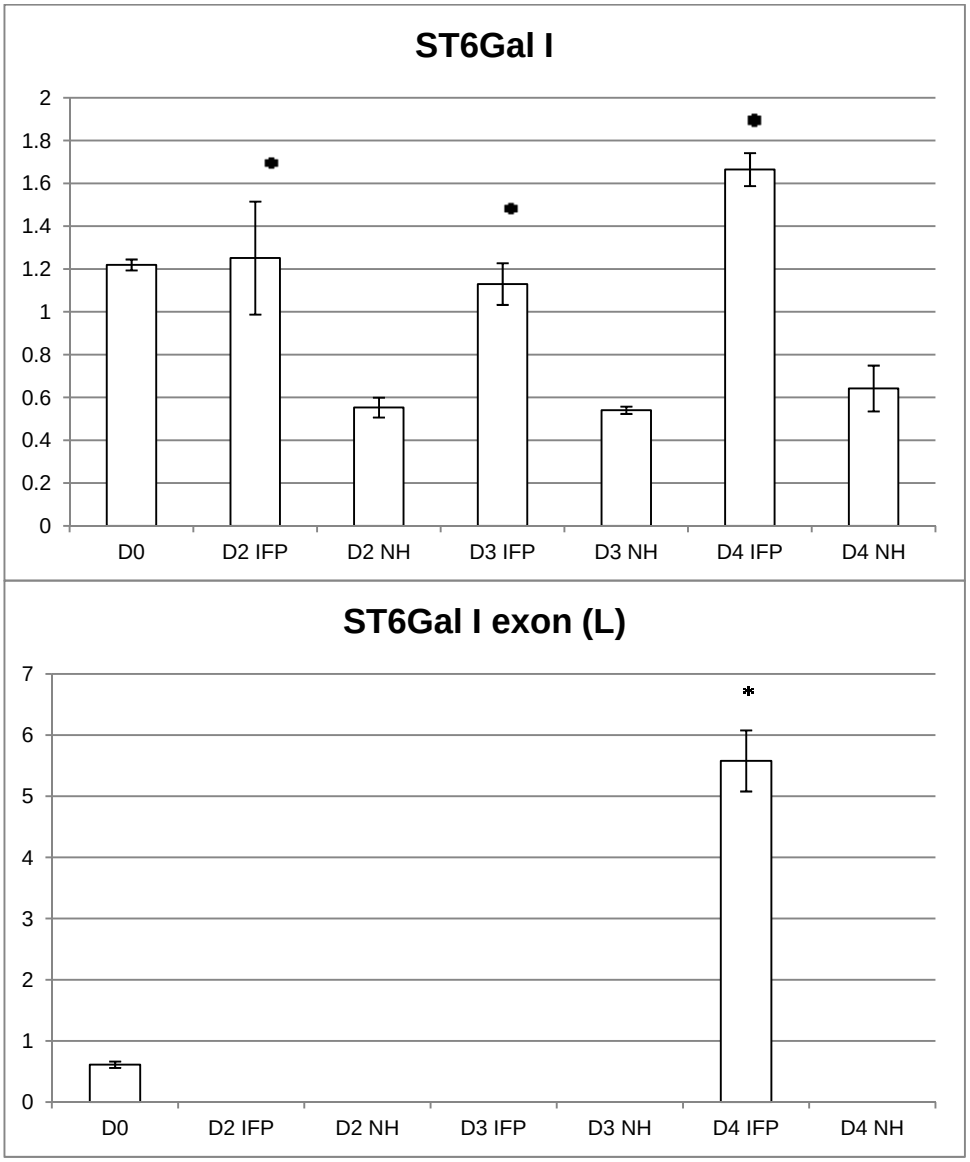


Figure 3: RT-PCR analysis of expression of the ST6Gal I gene and ST6Gal I exon (L) transcript in mammary explants. Mammary explants were assayed for ST6Gal I gene expression in uncultured mammary tissue, and after 2, 3 and 4 days of culture in the treatments indicated. The hormone treatments are abbreviated as follows: D0, uncultured tissue, I (insulin), F (hydrocortisone), P (prolactin), NH (no hormone). Target gene expression is expressed as relative to %GAPDH. A statistically significant difference in expression between NH and IFP on each day is denoted by a "*".

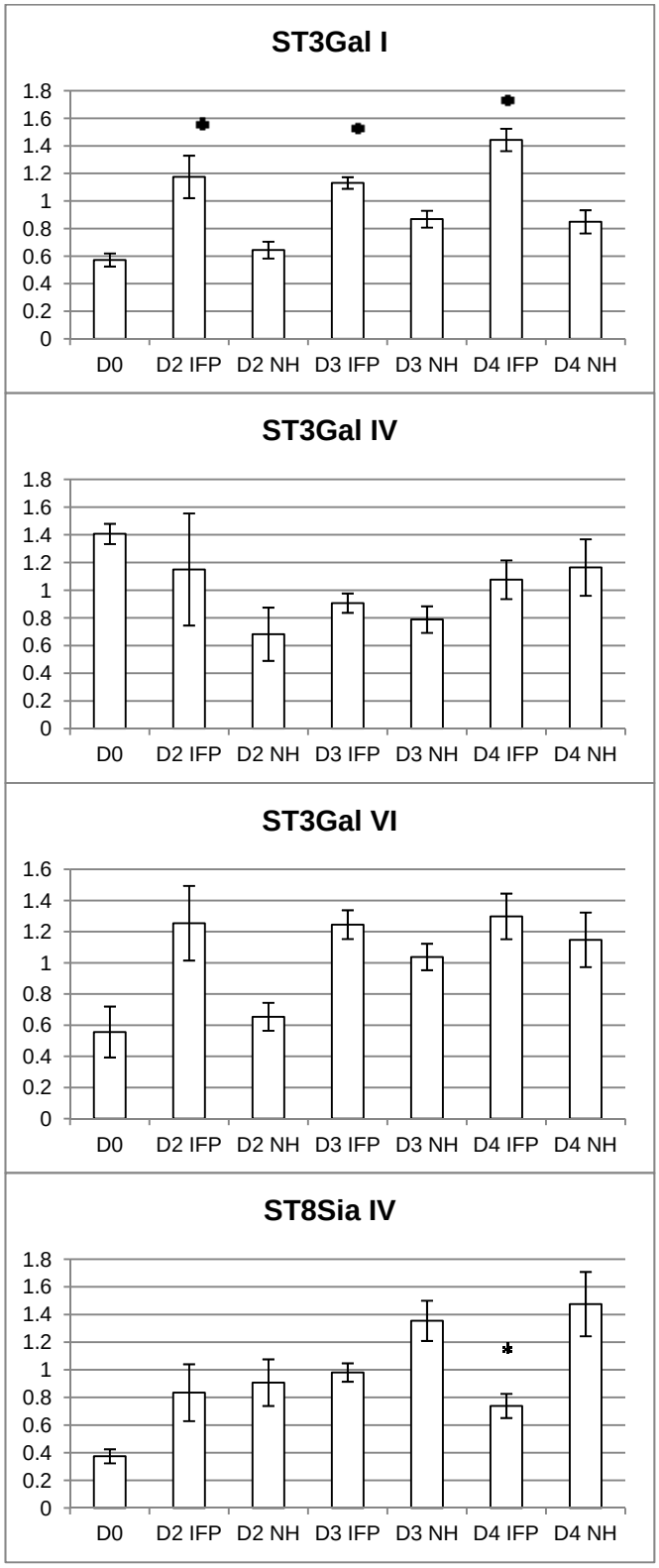


Figure 4: RT-PCR analysis of expression of sialyltransferase genes in mammary explants. Mammary explants were assayed for target gene expression in uncultured mammary tissue, and after 2, 3 and 4 days of culture in the treatments indicated. The hormone treatments are abbreviated as follows: D0, uncultured tissue, I (insulin), F (hydrocortisone), P (prolactin), NH (no hormone). Target gene expression is expressed as relative to %GAPDH. A statistically significant difference in expression between NH and IFP on each day is denoted by a "*".

Minimum hormone requirements for the induction of change in expression of sialic acid biosynthesis and sialyltransferase genes in mammary explants from mice

To determine which hormones were necessary to affect the level of gene expression observed in the previous culture experiment, the mammary explants were cultured for 4 days in every combination of I, F and P, and NH. The effectiveness of the culture system was confirmed by examining LALBA gene expression using RT-PCR. As previously reported (Palmiter, R.D. 1969), LALBA gene expression was P dependent but can also be stimulated to a degree by I. We observed a statistically significant increase ($p<0.05$) in LALBA gene expression when compared to D0 levels in explants cultured with IP and IFP (Figure 5). An increase in LALBA gene expression was also observed in IF, although this change was not statistically significant.

The GNE and NANS genes showed a similar pattern of expression in response to the various hormone treatments (Figure 6). A significant drop in expression was observed for both genes in all of the various hormone treatments when compared to D0. However, a significantly higher ($p<0.05$) level of GNE and NANS gene expression was maintained in IFP and the other P-containing combinations when compared to the negative control.

The overall expression of the ST6Gal I sialyltransferase gene was significantly elevated ($p<0.05$) in explants cultured in the presence of IFP when compared to D0 (Figure 7). ST6Gal I gene expression was maintained at D0 levels in explants cultured in FP and only showed a minor, albeit statistically significant ($p<0.05$), decrease in expression in IP and P. The exon (L) transcript of the ST6Gal I gene was significantly up-regulated ($p<0.05$) in explants cultured in IFP when compared with D0, but was not detectable in explants cultured in all other hormone combinations, except for a minor trace in IP (Figure 7).

The ST3Gal I gene increased in expression to a statistically significant level in explants cultured in FP when compared to the D0 tissue. Several other combinations also showed an

increase in ST3Gal I gene expression on D0 levels, however, these increases were not statistically significant (Figure 8).

The ST8Sia IV gene expression was increased in mouse mammary explants cultured in most hormone combinations when compared with D0 tissue. Its expression was significantly reduced from D0 levels only in mouse mammary explants cultured in IFP and IF (Figure 8).

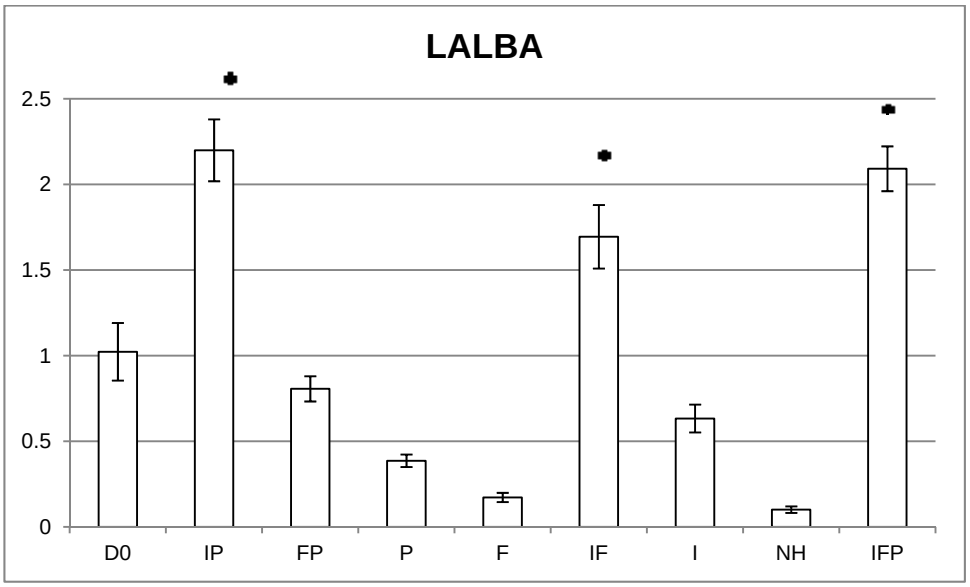


Figure 5: RT-PCR analysis of expression of the α-lactalbumin (LALBA) gene in mammary explants. Mammary explants were assayed for LALBA gene expression in uncultured mammary tissue and after 4 days of culture in the treatments indicated. The hormone treatments are abbreviated as follows: D0, uncultured tissue, I (insulin), F (hydrocortisone), P (prolactin), NH (no hormone). LALBA gene expression is expressed as relative to %GAPDH. A statistically significant increase in expression between D0 and a hormone treatment is denoted by a "*".

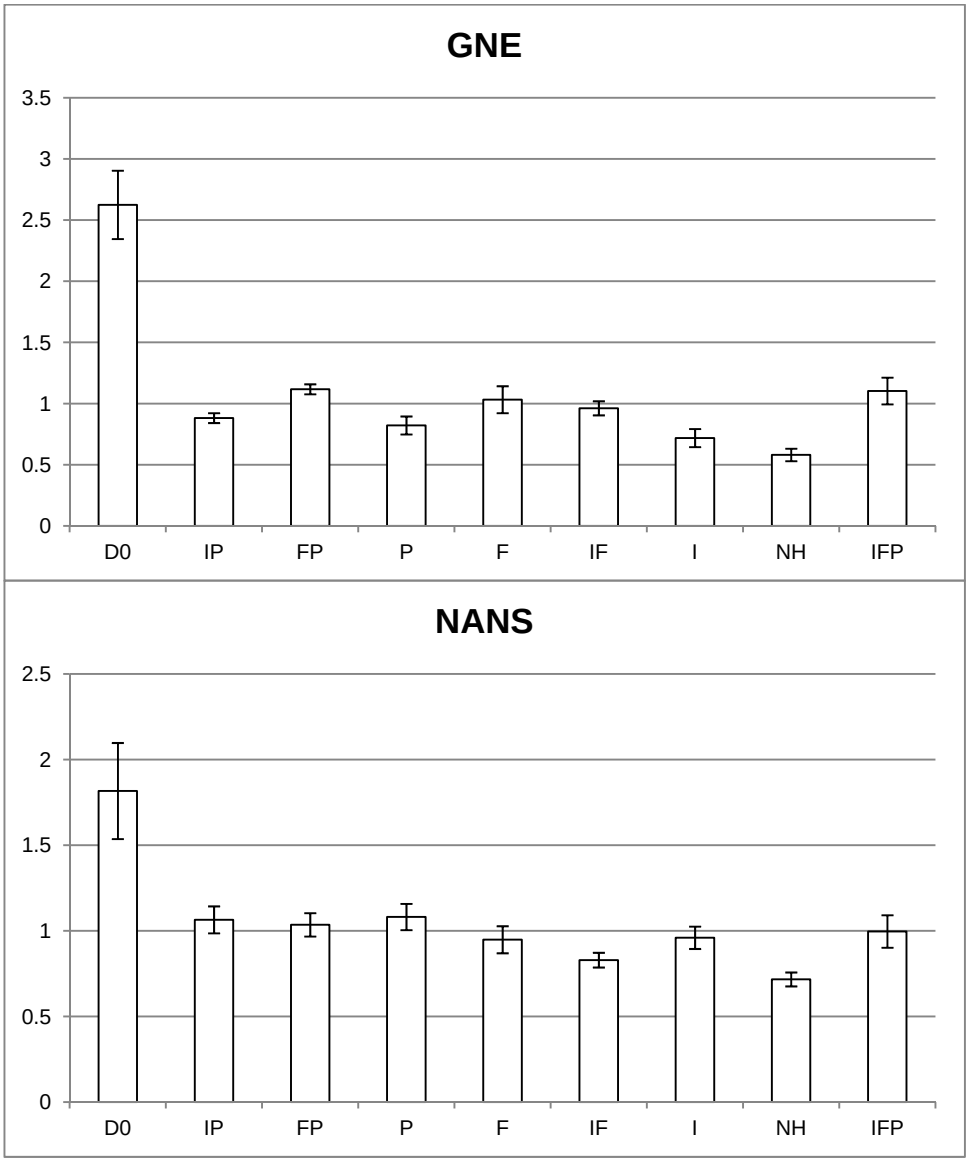


Figure 6: RT-PCR analysis of expression of sialic acid biosynthesis genes in mammary explants. Mammary explants were assayed for GNE and NANS gene expression in uncultured mammary tissue and after 4 days of culture in the treatments indicated. The hormone treatments are abbreviated as follows: D0, uncultured tissue, I (insulin), F (hydrocortisone), P (prolactin), NH (no hormone). Target gene expression is expressed as relative to %GAPDH.

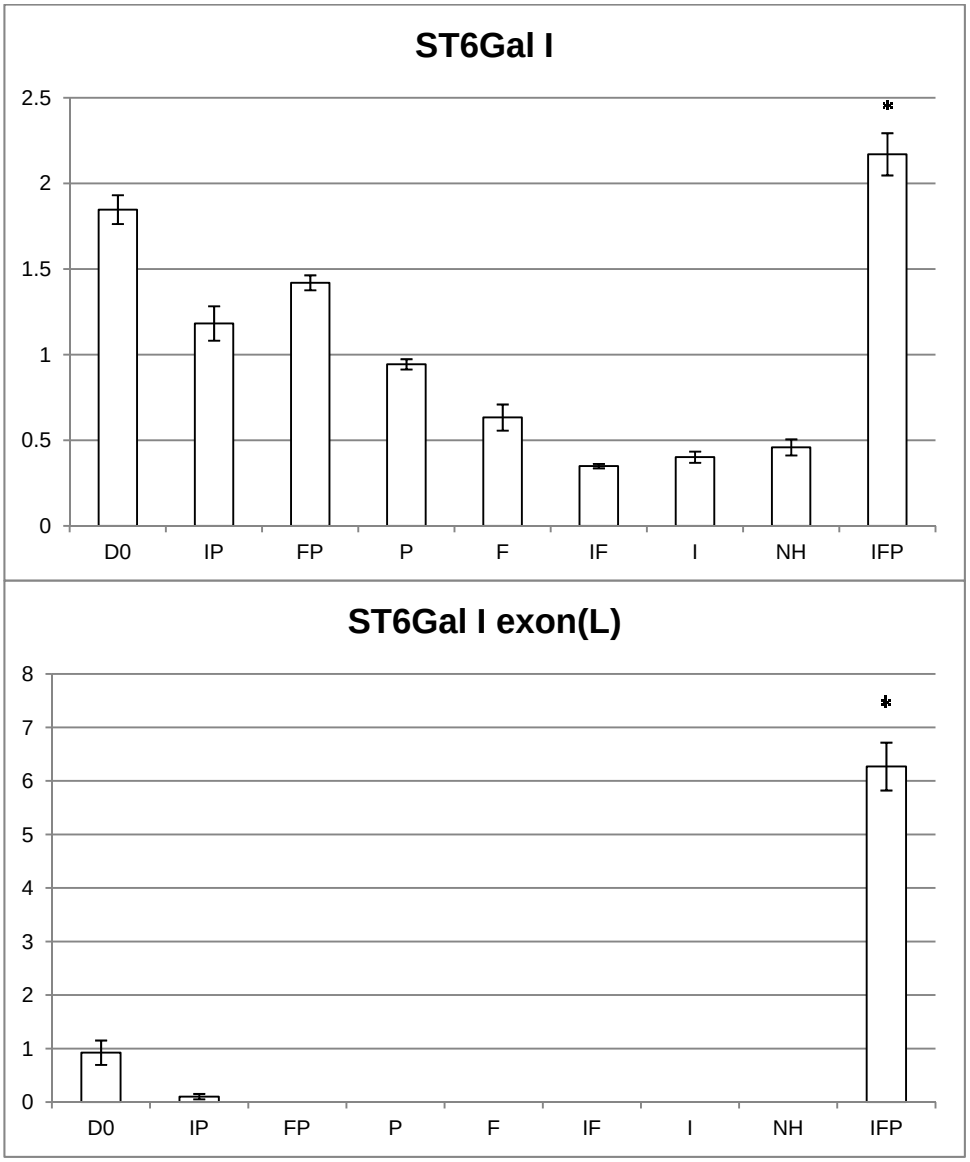


Figure 7: RT-PCR analysis of expression of the ST6Gal I gene and ST6Gal I exon (L) transcript in mammary explants. Mammary explants were assayed for ST6Gal I gene expression in uncultured mammary tissue and after 4 days of culture in the treatments indicated. The hormone treatments are abbreviated as follows: D0, uncultured tissue, I (insulin), F (hydrocortisone), P (prolactin), NH (no hormone). Target gene expression is expressed as relative to %GAPDH. A statistically significant increase in expression between D0 and a hormone treatment is denoted by a "*".

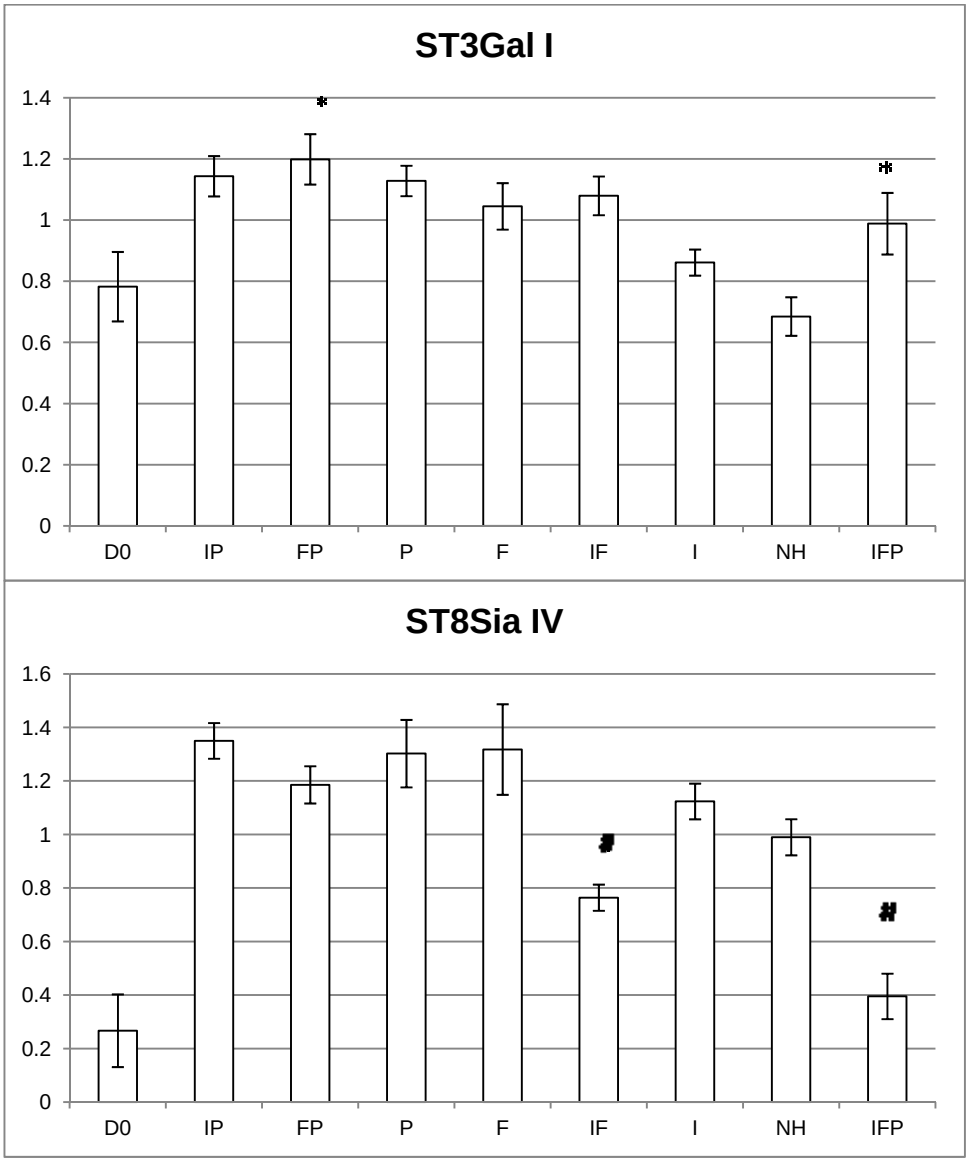


Figure 8: RT-PCR analysis of expression of sialyltransferase genes in mammary explants. Mammary explants were assayed for ST3Gal I gene expression in uncultured mammary tissue and after 4 days of culture in the treatments indicated. The hormone treatments are abbreviated as follows: D0, uncultured tissue, I (insulin), F (hydrocortisone), P (prolactin), NH (no hormone). Target gene expression is expressed as relative to %GAPDH. A statistically significant increase in ST3Gal I expression between D0 and a hormone treatment is denoted by a "*". Treatments that **did not** result in a statistically significant change in ST8Sia IV expression compared to D0 are denoted by a "#".

Characterisiation of putative promoter regions of genes responsive to lactogenic hormones

The lactating mammary gland phenotype is the result of coordinated control by numerous steroid and peptide hormones (Topper, Y.J. and Freeman, C.S. 1980). Doppler et al. (1990) demonstrated that mammary epithelial cells cultured in media with I,F and P were able to induce casein synthesis. This endocrine control has been shown to be expedited via the JAK/STAT pathway in concert with a suite of regulatory factors (Groner, B. 2002, Groner, B., Altiok, S., et al. 1994, Pellegrini, S. and Dusanter-Fourt, I. 1997). We used a bioinformatics approach to identify potential transcription factor binding sites (TFBS) for known lactogenic factors in the putative promoter regions of the genes shown to respond to lactogenic hormones (Figure 9).

Firstly we used the TSSG program by Softberry to predict candidate transcription start sites (TSS) upstream of the first transcribed exon for each gene of interest. For the ST6Gal I gene we were specifically interested in the putative promoter region upstream of exon (L), which is the first transcribed exon for the lactation-specific transcript of ST6Gal I (Dalziel, M., Huang, R.Y., et al. 2001). Due to the lack of evidence for a lactation-specific transcript of GNE, NANS, ST3Gal I and ST8Sia IV, the analysis focused on the region 5' of exon (I) for these genes. A TSS was predicted at just over 1kb from the 5' end of exon (L) in the ST6Gal I gene. ST3Gal I, on the other hand, had a candidate TSS predicted ~100bp upstream of exon (I). Similarly, a potential TSS was identified <100bp from the most 5' end of the first exon. Two TSS were predicted with similar confidence upstream of the first exon in GNE; one TFBS was predicted in close proximity to exon (I), whilst the other was predicted >500bp upstream.

Only a single candidate STAT5 binding site was identified in the ~2kb upstream of the first exon of each gene investigated (Figure 9, red). A STAT5 binding site was predicted ~700bp from the predicted TSS upstream of exon (L) in the ST6Gal I gene. The putative site was in close proximity to a candidate YY1 site, which is a known repressor of STAT5-

regulated gene transcription (Groner, B., Altiok, S., et al. 1994). A potential glucocorticoid receptor element (GR) was also found <100bp downstream of the STAT5-YY1 pair. Several putative C/EBPβ binding elements were also predicted between the STAT5 site and the candidate TSS (Figure 9). The putative STAT5 site in the GNE gene was several hundred base pairs upstream of both predicted TSS. It is situated in very close proximity to 2 potential C/EBPβ sites, and <100bp from a putative YY1 and GR site (Figure 9). Upstream of the NANS gene, the potential STAT5 site was located almost 600bp from the predicted TSS. Similarly to GNE, the NANS STAT5 site is proximal to a putative C/EBPβ and YY1 site, and within 200bp of the closest predicted GR element (Figure 9). The ST3Gal I gene had a putative STAT5 site identified within 200bp of the predicted TSS, which is only in the proximity of 3 potential Sp1 sites (Figure 9).

There were no candidate STAT5 sites predicted in the putative promoter region upstream of ST8Sia IV exon (I). However, multiple candidate Oct-1 and C/EBPβ sites were identified upstream of the predicted TSS, in addition to a potential YY1 site 688bp 5' of the TSS. Three potential GR binding sites were also identified, 2 of which directly overlapped with potential Oct-1 sites (Figure 9).

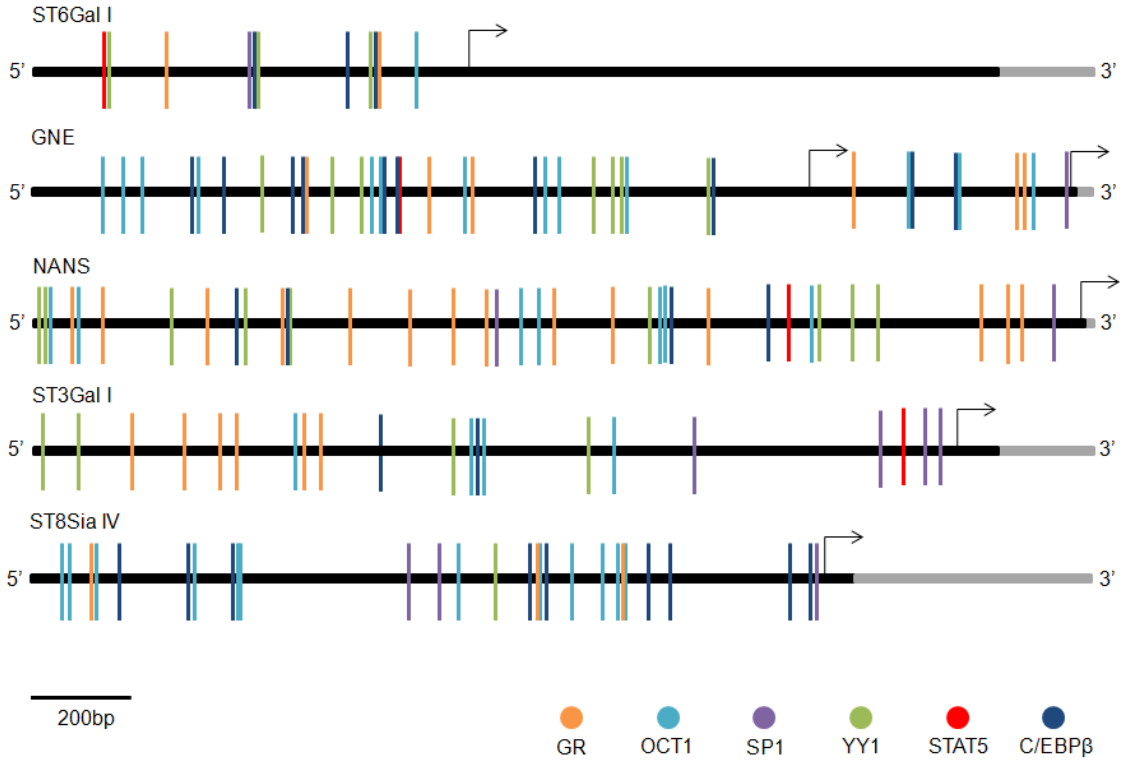


Figure 9: Predicted TSS and binding sites of STAT5 and related factors in the 2kb region upstream and including the first tanscribed exon of genes found to be responsive to lactogenic hormones. The thick black bar represents the ~2kb region upstream of the first transcribed exon, the grey bar on the right represents the position of the first transcribed exon. A scale bar is shown. For ST6Gal I, the region represented is the ~2kb region upstream of exon (L); for all other genes the region represented is upstream of exon (I). The approximate positions of putative STAT5 binding sites, and binding sites for related transcription factors are depicted by the coloured vertical lines; the legend explains the colour coding for each factor. All putative TFBS were identified using a Match search of TRANSFAC with matrix and core similarity cut off of 0.92 and 0.95.

Identification of potential distal STAT5 binding sites in genes responsive to lactogenic hormones

Nelson et al. (2006) demonstrated in their study that many functional STAT5 sites are located outside traditional promoters, consequently we used a bioinformatics approach to identify and map putative distal STAT5 sites in the entire genomic region of GNE, NANS, ST6Gal I and ST3Gal I, including an additional 20kb of 5' sequence (Figure 10). Overall, the ST6Gal I and ST3Gal I sialyltransferase genes contained more candidate STAT5 sites in the region investigated than the GNE and NANS genes. The NANS gene contained the least potential STAT5 binding sites, with one in the proximal promoter and another in the second intron (Figure 10). The GNE gene contained a total of 7 candidate STAT5 sites, 2 of which were upstream of the first exon, whilst the remaining 5 were all in introns (Figure 10). The majority of the STAT5 sites found in the ST6Gal I and ST3Gal I sialyltransferase genes were within introns, with only relatively few located upstream of the first exon. Since sialyltransferase genes are regulated by multiple promoters, the position of the "first" exon will vary between tissues and developmental states (Figure 10). In the case of lactation-specific expression of ST6Gal I, exon (L) is the "first" exon. Relative to exon (L), in addition to the candidate STAT5 site present in the putative promoter upstream of exon (L), 5 more potential STAT5 binding sites sit downstream of exon (L) in intronic regions prior to exon (I) (Figure 10).

There is evidence suggesting that tissue-specific and developmentally acting enhancers are frequently located in evolutionarily conserved regions (ECRs) (Blom van Assendelft, G., Hanscombe, O., et al. 1989, Moon, A.M. and Ley, T.J. 1990), and therefore we used Vista (Frazer, K., Pachter, L., et al. 2004) to examine the co-localisation of our putative STAT5 sites with ECRs identified in the mouse and human sequence for each gene of interest. A single putative STAT5 site downstream of exon (L) in ST6Gal I was located in an ECR (Fig. 12, denoted by "*") (Vista's criteria for determining an ECR is >70% sequence

identity over a 100bp sliding window). The presence of paired STAT5 binding sites was also investigated as they tend to be highly represented in STAT5-regulated promoters (Nelson, E.A., Walker, S.R., et al. 2006), however, only one potentially paired STAT5 in the ST3Gal I gene was identified where the 2 putative STAT5 sites were within 300bp (Figure 10, underlined).

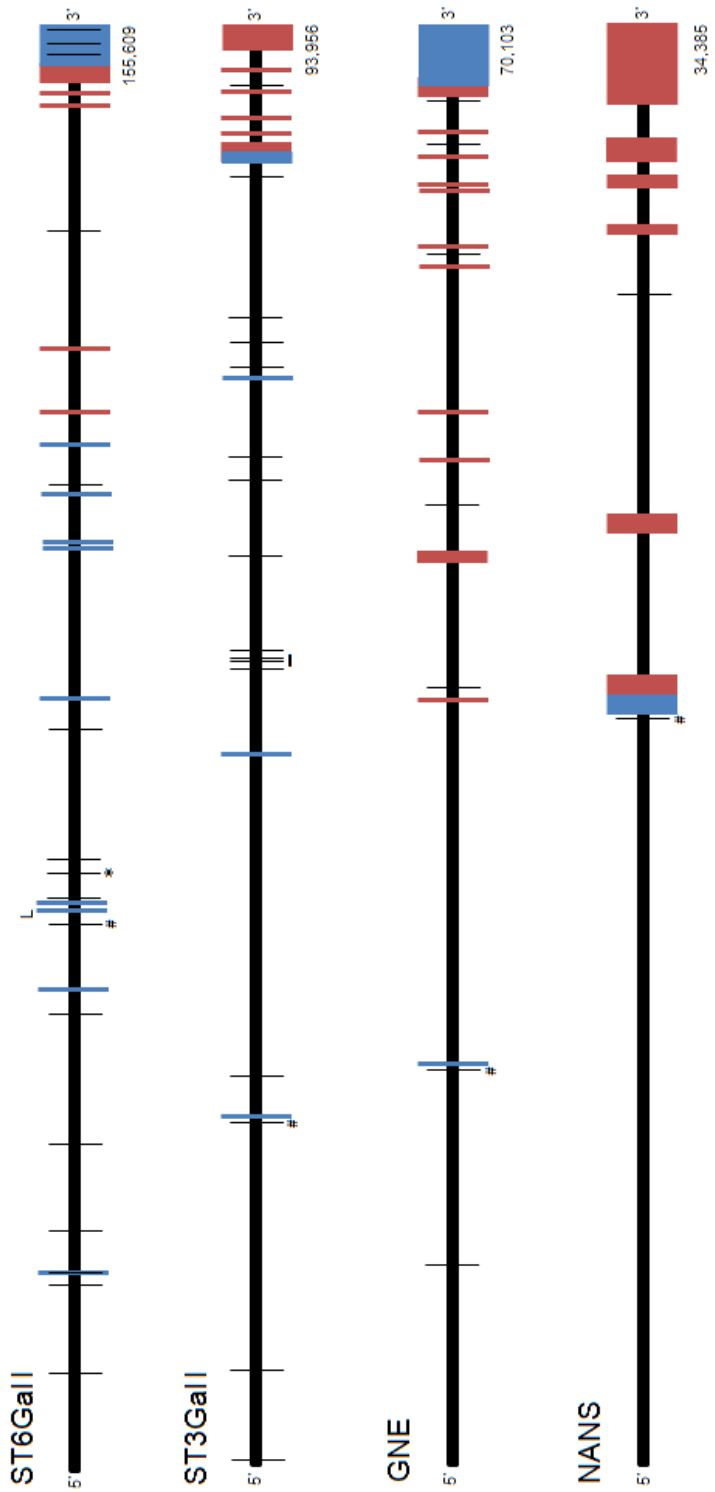


Figure 10: A schematic representation of the location of putative STAT5 binding sites in ST6Gal I, ST3Gal I, GNE and NANS loci. Each sequence includes an additional 20kb 5' of the first exon and the total length of each sequence is indicated on the right. Coding exons are depicted in red, whilst untranslated exons are shown in blue. The position of ST6Gal I exon (L) is indicated by an "L". The vertical black lines represent the putative STAT5 binding sites that were identified using the Match program. Putative STAT5 binding sites denoted by a "#", were found to be within 2kb of the predicted TSS. Sites highlighted with a "*" were found to be in an evolutionarily conserved region (ECR), whilst pairs of sites that are underlined are <300bp apart.

4.2.5 Discussion

Our study has demonstrated that the expression of genes involved in sialic acid biosynthesis and sialylation can be regulated in mammary explants cultured in media with lactogenic hormones. In these genes we also identified putative binding sites for transcription factors known to regulate the expression of milk protein genes. This data suggests that the transcriptional regulation of sialylation in the mammary gland is lactation specific and is regulated by the same signalling pathways responsible for initiating milk protein synthesis.

The current study demonstrated that the GNE, NANS, ST6Gal I and ST3Gal I genes all increased in expression in mammary explants cultured for 4 days with IFP. GNE and NANS are 2 key enzymes involved in sialic acid biosynthesis [reviewed in (Tanner, M.E. 2005)] and in a previous study we found that both the GNE and NANS genes were significantly up-regulated in mid-pregnancy, maintained a high level of expression during lactation and subsequently decreased in expression at the beginning of involution. GNE, the bifunctional enzyme UDP-GlcNAc 2-epimerase/ManNAc kinase, is responsible for the epimerisation of UDP-GlcNAc to ManNAc and the subsequent phosphorylation of ManNAc to ManNAc-6-P. The epimerisation reaction is the first committed step in sialic acid biosynthesis, and is also the rate limiting step in the pathway (Gal, B., Ruano, M.J., et al. 1997), making GNE a principal regulator of overall sialylation. GNE activity can be controlled at the transcriptional level and this has been shown to affect the sialylation and function of certain cell-surface molecules expressed on B-cells and myeloid cells (Keppler, O.T., Hinderlich, S., et al. 1999). N-acetylneuraminic acid synthase or NANS, catalyses the conversion of ManNAc-6-P to NeuNAc-9-P and in the mouse, is ubiquitously expressed in a variety of tissues (Nakata, D., Close, B.E., et al. 2000). Total sialic acid content of rat milk dramatically increases in abundance over the first week of lactation, primarily in the form of α2,3 and α2,6-sialyllactose (Duncan, P.I., Raymond, F., et al. 2009). Therefore, we propose that this increased demand for sialylation of milk products,

and thus an increased demand for sialic acid in the lactating mammary gland is catered for by the up-regulation of the GNE and NANS genes in response to lactogenic hormones, I, F and P. Unlike lactose, sialic acid is not a lactation-specific product, and therefore we expected to observe a baseline level of GNE and NANS gene expression even in explants cultured in NH. However, the increase in their expression in explants cultured in the presence of IFP is consistent with the assumption that additional sialic acid biosynthesis is required during lactation to cater for the sialylation of milk products and that it is achieved by the up-regulation of sialic acid biosynthesis genes in response to lactogenic hormones. Investigation of the minimum hormone combination required to confer an increase in GNE and NANS gene expression did not provide a conclusive result as gene expression was significantly lower in all treatments when compared with D0. P-containing treatments appeared to maintain higher levels of GNE and NANS gene expression relative to the negative control, although the results were not conclusive. The GNE and NANS gene both contained potential binding sites in their putative promoter regions for transcription factors that are known to be activated in response to lactogenic hormones. The GNE gene also contained several candidate STAT5 sites in introns that could potentially be acting as distal enhancers of transcription in response to the lactogenic hormones (Nelson, E.A., Walker, S.R., et al. 2006).

Both the ST6Gal I and ST3Gal I sialyltransferase genes were found to increase in expression in mammary explants cultured in the presence of IFP as compared to those cultured in NH. This result is consistent with increases in ST6Gal I and ST3Gal I gene expression previously observed between pregnancy and lactation (Dalziel, M., Huang, R.Y., et al. 2001, Fuhrer, A., Sprenger, N., et al. 2010), and observations made by Duncan et al. (2009) that the α2,3 and α2,6-sialyllactose content of mouse milk is dramatically elevated during the initial week of lactation post partum. According to Duncan et al. (2009) α2,3-sialyllactose is 10 times more abundant than α2,6-sialyllactose in the first week of lactation, after which it steadily declines until weaning, whilst the level of α2,6-

sialyllactose remains constant. Although lactose (Galβ1,4Glc) is not the preferred acceptor for either ST3Gal I or ST6Gal I, it has been demonstrated to be suitable at very high concentrations (Paulson, J.C., Rearick, J.I., et al. 1977, Sherblom, A.P. and Bourassa, C.R. 1983). Consequently, it is reasonable to conclude that lactation-specific up-regulation of ST3Gal I and ST6Gal I genes by lactogenic hormones is necessary to satisfy demand for the sialyllactose products of the encoded enzymes. Interestingly, although the ST3Gal I and ST3Gal IV genes are both up-regulated between pregnancy and lactation, Fuhrer et al. (2010) showed that ST3Gal IV is responsible for the majority of α2,3 sialyllactose production in early mouse lactation. However, our study did not show any evidence that the expression of the ST3Gal IV gene is affected by lactogenic hormones. Also, ST3Gal I preferentially transfers sialic acid to Galβ1,3GalNAc terminal disaccharides of O-linked carbohydrates on glycoproteins and glycolipids, via an α2,3 linkage (Takashima, S. 2008), suggesting that its primary sialylation target in the lactating mouse mammary gland may be a milk glycoprotein or glycolipid.

The ST3Gal I gene was found to contain several intronic candidate STAT5 binding sites and a single potential STAT5 binding site upstream of the exon (I) predicted TSS, although this was not found to be in close proximity to the usual cofactors. We suggest that it is probable that an as yet undiscovered lactation-specific transcript exists of ST3Gal I, given its apparent importance in murine lactation. It is likely that an accompanying lactogenic promoter region exists to regulate the expression of this gene. A study of ST3Gal I transcripts present in the lactating mouse mammary gland using 5'RACE would reveal the presence of any novel, lactation-specific transcripts and enable the localisation of a related promoter region.

Dalziel et al. (2001) have previously demonstrated the existence of a lactation-specific transcript of ST6Gal I that accounted for the increase in ST6Gal I gene expression observed between pregnancy and lactation. This lactation-specific ST6Gal I transcript

contains a unique 5'UTR exon, exon (L), which implies the presence of an upstream lactogenic promoter. This is consistent with our data which showed an overall increase in expression of the ST6Gal I gene in the presence of IFP, as well as an increase in expression of the exon (L)-containing transcript. The expression of the exon (L) transcript was only observed in the presence of IFP, which presumably accounts for the *overall* increase in ST6Gal I gene expression observed in the presence of all three hormones. Intriguingly, the increase in transcription of the exon (L)-containing ST6Gal I isoform occurred suddenly and dramatically only after 4 days of culture, although the reason for this delay is unclear. Based on our findings, we concur with Dalziel et al. (2001) that a lactogenic promoter exists in the vicinity of ST6Gal I exon (L). Our analysis of TFBS present upstream of the predicted TSS 5' of exon (L), revealed a single putative STAT5 binding site, in close proximity to YY1, which is a known repressor of milk protein gene transcription (Meier, V.S. and Groner, B. 1994, Raught, B., Khursheed, B., et al. 1994). A GR element was also found ~150bp downstream, as well as a putative C/EBPβ binding site ~300bp downstream. These have all previously been found to be essential in achieving a maximal increase in the expression of milk protein genes [for review see (Rosen, J.M., Wyszomierski, S.L., et al. 1999)]. Furthermore, we identified several potential STAT5 binding sites in ST6Gal I introns, which Nelson et al. (2006) demonstrated to be where many functional STAT5 sites are located. The first large intron downstream of exon (L), contained 4 putative STAT5 sites, one of which is found to be in a region that is evolutionarily conserved between mice, rats and humans (Fig. 12). We propose that this region may act as a distal enhancer of STAT5-mediated transcription of the ST6Gal I exon (L) isoform.

ST8Sia IV synthesises polysialic acid on neural cell adhesion molecules (NCAMs) (Bruses, J.L. and Rutishauser, U. 2001). It has been shown that the highly negative charge of large polysialic acid chains causes sufficient repulsion between neighbouring membranes to counteract the binding interactions of NCAMs (Johnson, C.P., Fragneto, G., et al. 2005, Johnson, C.P., Fujimoto, I., et al. 2005). We previously proposed that ST8Sia

IV may be playing a similar role in the mammary gland, by contributing to the leakiness of mammary epithelial tight junctions during pregnancy via the polysialylation of tight junction proteins that are present in the epithelial cell membrane. Our earlier analysis of sialyltransferase gene expression revealed that ST8Sia IV expression decreased throughout pregnancy, remained low in lactation and subsequently increased in involution (Maksimovic, manuscript in preparation). We believe that this decrease in ST8Sia IV gene expression is necessary to allow for a sufficient decrease in polysialylation facilitating the closure of the tight junctions. In the present study we have demonstrated that ST8Sia IV gene expression is decreased in mammary explant culture in the presence of lactogenic hormones I, F and P. Considering the closure of tight junctions occurs at parturition and that it is a necessary precursor to lactation (Linzell, J.L. and Peaker, M. 1971, Nguyen, D., Parlow, A., et al. 2001), it follows that ST8Sia IV may be involved in the leakiness of the mammary epithelium and the current data is consistent with its expression being regulated by the same hormones responsible for lactogenesis. We identified several potential Oct-1 and C/EBPβ sites, in addition to a putative YY1 site in the putative promoter region of the ST8Sia IV gene. YY1 is a multifunctional protein that is both an activator and repressor of transcription (Rosen, J.M., Wyszomierski, S.L., et al. 1999). It has been shown to repress casein gene expression in the absence of lactogenic hormones, but has little or no effect once hormones are present (Meier, V.S. and Groner, B. 1994). Lactogenic hormones do not affect the levels of YY1 in mammary epithelial cells (Meier, V.S. and Groner, B. 1994, Raught, B., Khursheed, B., et al. 1994). C/EBP is known to interact with YY1 (Bauknecht, T., See, R.H., et al. 1996), and it has been proposed that the protein-protein interaction between C/EBP and YY1 in the promoter region results in β-casein gene repression (Rosen, J.M., Wyszomierski, S.L., et al. 1999). The binding of YY1 and Oct-1 to the proximal promoter of IL-5 is thought to be involved in suppression of IL-5 transcription in T cells (Mordvinov, V.A., Schwenger, G.T., et al. 1999). The repressor role of YY1 on key regulatory elements is often linked to competition for binding of positive acting factors to

the same DNA, including among others the AP-1 complex (Ye, J., Cippitelli, M., et al. 1996) and Oct-1 (Mizuno, D., Takahashi, Y., et al. 2003). Any combination of these mechanisms could potentially be operating on the ST8Sia IV to reduce its expression during lactation, although further work is required to determine how lactogenic hormones are involved in this process.

The ST3Gal VI gene, which we also found to be highly expressed in early pregnancy and involution but down-regulated in mid to late pregnancy and lactation, does not appear to be directly regulated by IFP via the JAK/STAT pathway. Our current data has shown that although ST3Gal VI levels were indeed relatively low in uncultured mammary tissue obtained from mice in mid pregnancy, expression levels in explants cultured in both NH and IFP increased, suggesting that lactogenic hormones are not responsible for the down-regulation of the ST3Gal VI gene.

Taken collectively these data suggest that the up-regulation of sialic acid biosynthesis and sialyltransferase genes in the lactating mouse mammary gland is hormonally regulated, and lactation-specific, and necessary for the increase in demand for sialylation of milk products (Duncan, P.I., Raymond, F., et al. 2009).

4.2.6 References

Abramoff MD, Magelhaes PJ, Ram SJ. 2004. Image processing with ImageJ. *Biophotonics International*, 11:36-42.

Bauknecht T, See RH, Shi Y. 1996. A novel C/EBP beta-YY1 complex controls the cell-type-specific activity of the human papillomavirus type 18 upstream regulatory region. *J Virol*, 70:7695-7705.

Blom van Assendelft G, Hanscombe O, Grosveld F, Greaves DR. 1989. The beta-globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. *Cell*, 56:969-977.

Bolander FF, Jr., Nicholas KR, Van Wyk JJ, Topper YJ. 1981. Insulin is essential for accumulation of casein mRNA in mouse mammary epithelial cells. *Proc Natl Acad Sci U S A*, 78:5682-5684.

Bruses JL, Rutishauser U. 2001. Roles, regulation, and mechanism of polysialic acid function during neural development. *Biochimie*, 83:635-643.

Choi KM, Barash I, Rhoads RE. 2004. Insulin and prolactin synergistically stimulate beta-casein messenger ribonucleic acid translation by cytoplasmic polyadenylation. *Mol Endocrinol*, 18:1670-1686.

Chomczynski P, Qasba P, Topper YJ. 1984. Essential role of insulin in transcription of the rat 25,000 molecular weight casein gene. *Science*, 226:1326-1328.

Coppa GV, Gabrielli O, Giorgi P, Catassi C, Montanari MP, et al. 1990. Preliminary study of breastfeeding and bacterial adhesion to uroepithelial cells. *The Lancet; London*, 335:569.

Dall'Olio F. 2000. The sialyl-α2,6-lactosaminyl-structure: Biosynthesis and functional role. *Glycoconj. J.*, 17:669-676.

Dalziel M, Huang RY, Dall'Olio F, Morris JR, Taylor-Papadimitriou J, Lau JTY. 2001. Mouse ST6Gal sialyltransferase gene expression during mammary gland lactation. *Glycobiology*, 11:407-412.

Dalziel M, Lemaire S, Ewing J, Kobayashi L, Lau JT. 1999. Hepatic acute phase induction of murine β-galactoside α2,6 sialyltransferase (ST6Gal I) is IL-6 dependent and mediated by elevation of Exon H-containing class of transcripts. *Glycobiology*, 9:1003-1008.

Dodd SC, Forsyth IA, Buttle HL, Gurr MI, Dils RR. 1994. Hormonal induction of α-lactalbumin and β-lactoglobulin in cultured mammary explants from pregnant pigs. *Journal of Dairy Research*, 61:35-45.

Dong B, Huang C, Li D, Zhao FQ. 2009. Oct-1 functions as a transactivator in the hormonal induction of beta-casein gene expression. *Mol Cell Biochem*, 328:93-99.

Dong B, Zhao FQ. 2007. Involvement of the ubiquitous Oct-1 transcription factor in hormonal induction of beta-casein gene expression. *Biochem J*, 401:57-64.

Doppler W, Hock W, Hofer P, Groner B, Ball RK. 1990. Prolactin and Glucocorticoid Hormones Control Transcription of the β-Casein Gene by Kinetically Distinct Mechanisms. *Mol Endocrinol*, 4:912-919.

Doppler W, Welte T, Philipp S. 1995. CCAAT/enhancer-binding protein isoforms beta and delta are expressed in mammary epithelial cells and bind to multiple sites in the beta-casein gene promoter. *J Biol Chem*, 270:17962-17969.

Duncan PI, Raymond F, Fuerholz A, Sprenger N. 2009. Sialic Acid Utilisation and Synthesis in the Neonatal Rat Revisited. *PLoS One*, 4:e8241.

Frazer K, Pachter L, Poliakov A, Rubin E, Dubchak I. 2004 VISTA: computational tools for comparative genomics. *Nucleic Acids Res.*, Jul 1:W273-279.

Fuhrer A, Sprenger N, Kurakevich E, Borsig L, Chassard C, Hennet T. 2010. Milk sialyllactose influences colitis in mice through selective intestinal bacterial colonization. *The Journal of Experimental Medicine*, 207:2843-2854.

Gal B, Ruano MJ, Puente R, GarciaPardo LA, Rueda R, Gil A, Hueso P. 1997. Developmental changes in UDP-N-acetylglucosamine 2-epimerase activity of rat and guinea-pig liver. *Comp. Biochem. Phys. B*, 118:13-15.

Groner B. 2002. Transcription factor regulation in mammary epithelial cells. *Domestic Animal Endocrinology*, 23:25-32.

Groner B, Altiok S, Meier V. 1994. Hormonal regulation of transcription factor activity in mammary epithelial cells. *Molecular and Cellular Endocrinology*, 100:109-114.

Hanasaki K, Powell LD, Varki A. 1995. Binding of human plasma sialoglycoproteins by the B cell-specific lectin CD22. Selective recognition of immunoglobulin M and haptoglobin. *J. Biol. Chem.*, 270:7543-7550.

Harduin-Lepers A, Mollicone R, Delannoy P, Oriol R. 2005. The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. *Glycobiology*, 15:805-817.

Harduin-Lepers A, Recchi MA, Delannoy P. 1995. 1994, the year of sialyltransferases. *Glycobiology*, 5:741-758.

Harduin-Lepers A, Vallejo-Ruiz V, Krzewinski-Recchi MA, Samyn-Petit B, Julien S, Delannoy P. 2001. The human sialyltransferase family. *Biochimie*, 83:727-737.

Jiang J, Levine M. 1993. Binding affinities and cooperative interactions with bHLH activators delimit threshold responses to the dorsal gradient morphogen. *Cell*, 72:741-752.

Johnson CP, Fragneto G, Konovalov O, Dubosclard V, Legrand JF, Leckband DE. 2005. Structural studies of the neural-cell-adhesion molecule by X-ray and neutron reflectivity. *Biochemistry*, 44:546-554.

Johnson CP, Fujimoto I, Rutishauser U, Leckband DE. 2005. Direct evidence that neural cell adhesion molecule (NCAM) polysialylation increases intermembrane repulsion and abrogates adhesion. *J Biol Chem*, 280:137-145.

Keppler OT, Hinderlich S, Langner J, Schwartz-Albiez R, Reutter W, Pawlita M. 1999. UDP-GlcNAc 2-Epimerase: A Regulator of Cell Surface Sialylation. *Science*, 284:1372-1376.

Kulski JK, Nicholas KR, Topper YJ, Qasba P. 1983. Essentiality of insulin and prolactin for accumulation of rat casein mRNAs. *Biochem Biophys Res Commun*, 116:994-999.

Lechner J, Welte T, Tomasi JK, Bruno P, Cairns C, Gustafsson J, Doppler W. 1997. Promoter-dependent synergy between glucocorticoid receptor and Stat5 in the activation of beta-casein gene transcription. *J Biol Chem*, 272:20954-20960.

Li S, Rosen JM. 1994. Glucocorticoid regulation of rat whey acidic protein gene expression involves hormone-induced alterations of chromatin structure in the distal promoter region. *Mol Endocrinol*, 8:1328-1335.

Li S, Rosen JM. 1995. Nuclear factor I and mammary gland factor (STAT5) play a critical role in regulating rat whey acidic protein gene expression in transgenic mice. *Mol Cell Biol*, 15:2063-2070.

Linzell JL, Peaker M. 1971. The permeability of mammary ducts. *J Physiol*, 216:701-716.

Martin-Sosa S, Martin MJ, Garcia-Pardo LA, Hueso P. 2003. Sialyloligosaccharides in human and bovine milk and in infant formulas: variations with the progression of lactation. *J. Dairy Sci.*, 86:52-59.

McVeagh P, Brand-Miller J. 1997. Human milk oligosaccharides: only the breast. *J. Paediatr. Child Health*, 33:281-286.

Meier VS, Groner B. 1994. The nuclear factor YY1 participates in repression of the beta-casein gene promoter in mammary epithelial cells and is counteracted by mammary gland factor during lactogenic hormone induction. *Mol Cell Biol*, 14:128-137.

Menzies K, Lee H, Lefèvre C, Ormandy C, Macmillan K, Nicholas K. 2010. Insulin, a key regulator of hormone responsive milk protein synthesis during lactogenesis in murine mammary explants. *Functional & Integrative Genomics*, 10:87-95.

Mizuno D, Takahashi Y, Hiroi T, Imaoka S, Kamataki T, Funae Y. 2003. A novel transcriptional element which regulates expression of the CYP2D4 gene by Oct-1 and YY-1 binding. *Biochim Biophys Acta*, 1627:121-128.

Moon AM, Ley TJ. 1990. Conservation of the primary structure, organization, and function of the human and mouse beta-globin locus-activating regions. *Proc Natl Acad Sci U S A*, 87:7693-7697.

Mordvinov VA, Schwenger GT, Fournier R, De Boer ML, Peroni SE, Singh AD, Karlen S, Holland JW, Sanderson CJ. 1999. Binding of YY1 and Oct1 to a novel element that downregulates expression of IL-5 in human T cells. *J Allergy Clin Immunol*, 103:1125-1135.

Mouricout M, Petit JM, Carias JR, Julien R. 1990. Glycoprotein glycans that inhibit adhesion of Escherichia coli mediated by K99 fimbriae: treatment of experimental colibacillosis. *Infect. Immun.*, 58:98-106.

Nagaiah K, Bolander FF, Jr., Nicholas KR, Takemoto T, Topper YJ. 1981. Prolactin-induced accumulation of casein mRNA in mouse mammary explants: a selective role of glucocorticoid. *Biochem Biophys Res Commun*, 98:380-387.

Nakata D, Close BE, Colley KJ, Matsuda T, Kitajima K. 2000. Molecular Cloning and Expression of the Mouse N-Acetylneuraminic Acid 9-Phosphate Synthase Which Does Not Have Deaminoneuraminic Acid (KDN) 9-Phosphate Synthase Activity. *Biochemical and Biophysical Research Communications*, 273:642-648.

Nelson EA, Walker SR, Li W, Liu XS, Frank DA. 2006. Identification of Human STAT5-dependent Gene Regulatory Elements Based on Interspecies Homology. *Journal of Biological Chemistry*, 281:26216-26224.

Nguyen D, Parlow A, Neville M. 2001. Hormonal regulation of tight junction closure in the mouse mammary epithelium during the transition from pregnancy to lactation. *J Endocrinol*, 170:347-356.

Nicholas KR, Sankaran L, Topper YJ. 1983. A unique and essential role for insulin in the phenotypic expression of rat mammary epithelial cells unrelated to its function in cell maintenance. *Biochim Biophys Acta*, 763:309-314.

Nicholas KR, Tyndale-Biscoe CH. 1985. Prolactin-dependent accumulation of alpha-lactalbumin in mammary gland explants from the pregnant tammar wallaby (Macropus eugenii). *J Endocrinol*, 106:337-342.

Ono M, Oka T. 1980. alpha-Lactalbumin-casein induction in virgin mouse mammary explants: dose-dependent differential action of cortisol. *Science*, 207:1367-1369.

Palmiter RD. 1969. Hormonal induction and regulation of lactose synthetase in mouse mammary gland. *Biochem J*, 113:409-417.

Paulson JC, Rearick JI, Hill RL. 1977. Enzymatic properties of beta-d-galactoside alpha-2-6 sialyltransferase from bovine colostrum. *Journal of Biological Chemistry*, 252:2363-2371.

Pellegrini S, Dusanter-Fourt I. 1997. The Structure, Regulation and Function of the Janus Kinases (JAKs) and the Signal Transducers and Activators of Transcription (STATs). *European Journal of Biochemistry*, 248:615-633.

Raught B, Khursheed B, Kazansky A, Rosen J. 1994. YY1 represses beta-casein gene expression by preventing the formation of a lactation-associated complex. *Mol Cell Biol*, 14:1752-1763.

Raught B, Liao WS, Rosen JM. 1995. Developmentally and hormonally regulated CCAAT/enhancer-binding protein isoforms influence beta-casein gene expression. *Mol Endocrinol*, 9:1223-1232.

Rosen JM, Wyszomierski SL, Hadsell D. 1999. Regulation of milk protein gene expression. *Annual Review of Nutrition*, 19:407-436.

Sankaran L, Topper YJ. 1984. Prolactin-induced alpha-lactalbumin activity in mammary explants from pregnant rabbits. A role for epidermal growth factor and glucocorticoids. *Biochem J*, 217:833-837.

Schmitt-Ney M, Doppler W, Ball RK, Groner B. 1991. Beta-casein gene promoter activity is regulated by the hormone-mediated relief of transcriptional repression and a mammary-gland-specific nuclear factor. *Mol Cell Biol*, 11:3745-3755.

Seagroves TN, Krnacik S, Raught B, Gay J, Burgess-Beusse B, Darlington GJ, Rosen JM. 1998. C/EBPbeta, but not C/EBPalpha, is essential for ductal morphogenesis, lobuloalveolar proliferation, and functional differentiation in the mouse mammary gland. *Genes Dev*, 12:1917-1928.

Sheehy PA, Della-Vedova JJ, Nicholas KR, Wynn PC. 2004. Hormone-dependent milk protein gene expression in bovine mammary explants from biopsies at different stages of pregnancy. *J. of Dairy Res.*, 71:135-140.

Sherblom AP, Bourassa CR. 1983. Specificity of submaxillary gland sialyltransferases. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 761:94-102.

Svensson EC, Soreghan B, Paulson JC. 1990. Organization of the β-galactoside α2,6-sialyltransferase gene. Evidence for the transcriptional regulation of terminal glycosylation. *J. Biol. Chem.*, 265:20863-20868.

Takashima S. 2008. Characterization of Mouse Sialyltransferase Genes: Their Evolution and Diversity. *Bioscience, biotechnology, and biochemistry*, 72:1155-1167.

Takashima S, Kono M, Kurosawa N, Yoshida Y, Tachida Y, Inoue M, Kanematsu T, Tsuji S. 2000. Genomic organization and transcriptional regulation of the mouse GD3 synthase gene (ST8Sia I): Comparison of genomic organization of the mouse sialyltransferase genes. *Journal of Biochemistry*, 128:1033-1043.

Tanner ME. 2005. The enzymes of sialic acid biosynthesis. *Bioorganic Chemistry*, 33:216-228.

Topper YJ, Freeman CS. 1980. Multiple hormone interactions in the developmental biology of the mammary gland. *Physiol. Rev.*, 60:1049-1106.

Tsuji S. 1996. Molecular Cloning and Functional Analysis of Sialyltransferases. *J Biochem (Tokyo)*, 120:1-13.

Varki A. 1993. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*, 3:97-130.

Varki A. 2007. Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. *Nature*, 446:1023-1029.

Wakao H, Gouilleux F, Groner B. 1994. Mammary gland factor (MGF) is a novel member of the cytokine regulated transcription factor gene family and confers the prolactin response. *Embo J*, 13:2182-2191.

Wang B, Brand-Miller J. 2003. The role and potential of sialic acid in human nutrition. *Eur. J. Clin. Nutr.*, 57:1351-1369.

Wang B, Brand-Miller J, McVeagh P, Petocz P. 2001. Concentration and distribution of sialic acid in human milk and infant formulas. *Am. J. Clin. Nutr.*, 74:510-515.

Wang X, O'Hanlon TP, Young RF, Lau JTY. 1990. Rat β-galactoside α2,6-sialyltransferase genomic organization: alternate promoters direct the synthesis of liver and kidney transcripts. *Glycobiology*, 1:25-31.

Wen DX, Svensson EC, Paulson JC. 1992. Tissue-specific alternative splicing of the β-galactoside α2,6- sialyltransferase gene. *J. Biol. Chem.*, 267:2512-2518.

Wuensch SA, Huang RY, Ewing J, Liang X, Lau JT. 2000. Murine B cell differentiation is accompanied by programmed expression of multiple novel β-galactoside α2, 6-sialyltransferase mRNA forms. *Glycobiology*, 10:67-75.

Ye J, Cippitelli M, Dorman L, Ortaldo JR, Young HA. 1996. The nuclear factor YY1 suppresses the human gamma interferon promoter through two mechanisms: inhibition of AP1 binding and activation of a silencer element. *Mol Cell Biol*, 16:4744-4753.

Zhao FQ, Zheng Y, Dong B, Oka T. 2004. Cloning, genomic organization, expression, and effect on beta-casein promoter activity of a novel isoform of the mouse Oct-1 transcription factor. *Gene*, 326:175-187.

5 Chapter Five

Given that sialylated milk oligosaccharides have been implicated in facilitating post-natal brain development of the neonate (Wang, B. and Brand-Miller, J. 2003) and that dietary sialic acid has been shown to enhance learning in developing piglets (Moughan, P.J., Birtles, M.J., et al. 1992, Wang, B., Yu, B., et al. 2007), it was of interest to investigate the expression of sialic acid biosynthesis and sialyltransferase genes previously found to be differentially expressed in the mouse mammary gland in a species such as the tammar wallaby. The tammar is a marsupial that is born in an altricial state and is totally dependent on milk for its growth and development (Renfree, M.B. 1983, Tyndale-Biscoe, C.H. and Janssens, P.A. 1988). As it has been suggested that altricial neonates may have a particular requirement for milk oligosaccharides (Oftedal, O.T., Alt, G.L., et al. 1993), it was of particular interest to discern whether the pattern of gene expression correlated with any developmental milestones and whether any such correlation provided further clues as to the role of sialic acid in neonatal development.

This chapter is presented in manuscript format in preparation for journal review.

5.1 Declaration for thesis chapter five

5.1.1 Declaration by candidate

In the case of Chapter 5, the nature and extent of my contribution to the work was the following:

Nature of contribution	**Extent of contribution (%)**
For this chapter I performed the laboratory experiments and was responsible for analysis of the data, interpretation of the results and the writing of the manuscript. The tammar wallaby tissue used in this study had been collected previously, over a number of years, by various individuals for other studies.	80%

The following co-authors contributed to the work. Co-authors who are students at Monash University must also indicate the extent of their contribution in percentage terms:

Name	**Nature of contribution**	**Extent of contribution (%)** **(student co-authors only)**
Dr. Julie A. Sharp	Dr. Sharp aided in interpretation of results and writing of the manuscript.	
Dr. Keith W. Savin	Dr. Savin aided in interpretation of results and writing of the manuscript.	
Prof. Kevin R. Nicholas	Prof. Nicholas aided in interpretation of results and writing of the manuscript.	

Candidate's Signature		**Date**

5.1.2 Declaration by co-authors

The undersigned hereby certify that:

(7) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.
(8) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
(9) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
(10) there are no other authors of the publication according to these criteria;
(11) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
(12) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

Location(s)	**DPI, Biosciences Research Division, Bundoora VIC 3083**

Signature 1	[redacted]	**Date** 26/3/2011
Signature 2	[redacted]	4-4-11
Signature 3	[redacted]	6/4/11

5.2 Characterisation and expression of sialic acid biosynthesis and sialyltransferase genes in the lactating tammar wallaby mammary gland

Jovana Maksimovic[1,2], Keith W. Savin[2], Julie A. Sharp[3], Kevin R. Nicholas[3]

1. Biosciences Research Division, Department of Primary Industries, Melbourne 3083 Australia
2. Centre for Reproduction and Development, Monash Institute of Medical Research, Monash University, Melbourne 3168 Australia
3. Institute for Technology Research and Innovation, Deakin University, Geelong 3214 Australia

5.2.1 Abstract

Milk sialylglycoconjugates can protect the gastrointestinal tract of the suckling neonate by competitively binding to invading pathogens and promoting growth of beneficial flora, and their potential role in post-natal brain development is of particular interest in human infant nutrition. Sialylglycoconjugates are synthesised by sialyltransferase enzymes via the transfer of sialic acid from CMP-sialic acid to the terminal disaccharides of carbohydrate chains on glycoproteins, glycolipids and free oligosaccharides. Sialyltransferases are primarily regulated at the transcriptional level by the use of several tissue and development-specific promoters. Thus far, 20 sialyltransferase enzymes have been characterised in mice and previous studies have shown that certain sialyltransferases genes are differentially expressed in the mouse mammary gland during lactation. Here we report the characterisation of UDP-GlcNAc 2-epimerase/ManNAc kinase (GNE) and N-acetylneuraminic acid synthase (NANS), two key sialic acid biosynthesis genes, in the tammar wallaby (*Macropus eugenii*), along with the ST3Gal I, ST6Gal I and ST6GalNAc II sialyltransferase genes. We also demonstrate that these genes are differentially expressed in the tammar mammary gland during lactation and that their expression correlates with previously published data on the sialic acid content of tammar wallaby milk. Given that tammar milk is comprised of higher oligosaccharides for most of lactation and that their development can be equated to that of eutherians, we propose that they are an appropriate model for comparative studies concerning the effect of milk sialic acid on neonate development.

5.2.2 Introduction

Sialic acid is an acidic monosaccharide with a nine-carbon backbone. In mammals, it is typically located at the non-reducing terminus of oligosaccharides, glycoproteins or glycolipids [reviewed in (Harduin-Lepers, A., Recchi, M.A., et al. 1995, Tsuji, S. 1996)]. Sialylated structures are essential to numerous biological phenomena such as cell-cell interaction, cell migration, adhesion and metastasis (Dall'Olio, F. 2000, Hanasaki, K., Powell, L.D., et al. 1995, Varki, A. 1993, Varki, A. 2007).

The mammalian nervous system contains greater concentrations of sialic acid than any other organ. Within the nervous system, 65% of all sialic acid is used in the formation of gangliosides, 32% is attached to glycoproteins and only 3% is in the free form (Brunngraber, E.G., Witting, L.A., et al. 1972, Wang, B. and Brand-Miller, J. 2003). The sialic acid content of neural cell membranes is 20 times higher than in any other type of membrane, suggesting an important role for sialic acid in neural structure (Schauer, R. 1982). Furthermore, it has been proposed that ganglioside-bound sialic acid may be involved in processes such as synaptic transmission and neuronal adaptations required for storage of information, whilst sialylated glycoproteins have been implicated in learning and memory. It has also been hypothesised that sialic acid may be the receptor for neurotransmitters in the central nervous system (Von Itzstein, M. and Thomson, R. 1997, Wang, B. and Brand-Miller, J. 2003).

Human milk is an extremely rich source of essential nutrients for the newborn human infant (Hamosh, M. 1999, Picciano, M.F. 1998). In addition to proteins, vitamins, minerals and antibodies, human milk provides the neonate with a suite of over 200 oligosaccharides. Following lactose and lipids, these complex sugars are the third major milk solute and are present in greater overall quantities than protein (McVeagh, P. and Brand-Miller, J. 1997, Ninonuevo, M.R., Park, Y., et al. 2006). Sialylated milk oligosaccharides have been of particular interest to researchers because of their protective role in the gastrointestinal tract

of the suckling neonate which results from competitive binding of invading pathogens and their ability to promote the growth of beneficial flora (Coppa, G.V., Gabrielli, O., et al. 1990, LoCascio, R.G., Ninonuevo, M.R., et al. 2007, Martin-Sosa, S., Martin, M.J., et al. 2003, McVeagh, P. and Brand-Miller, J. 1997, Mouricout, M., Petit, J.M., et al. 1990, Sela, D.A., Li, Y., et al. 2011, Wang, B., Brand-Miller, J., et al. 2001). Sialylated milk oligosaccharides have also been implicated in facilitating post-natal brain development (Moughan, P.J., Birtles, M.J., et al. 1992, Wang, B. and Brand-Miller, J. 2003, Wang, B., Yu, B., et al. 2007).

Breastfeeding at infancy has been associated with higher later intelligence quotients in several studies (Fergusson, D.M., Beautrais, A.L., et al. 1982, Lucas, A., Morley, R., et al. 1998, Lucas, A., Morley, R., et al. 1992, Rodgers, B. 1978), and is thought to be particularly beneficial for the neural development of premature infants (Lucas, A., Morley, R., et al. 1998, Lucas, A., Morley, R., et al. 1992). Although there are many factors in milk that may contribute to cognitive development, McVeagh and Miller (1997) suggest, that due to its abundance and pivotal role in the nervous system, sialic acid derived from milk oligosaccharides could be a conditional nutrient, in that breastfed infants are advantaged in having a nutritional source to supplement endogenous production.

The lactation cycle of tammar wallaby is markedly different from that of eutherian mammals. After a relatively brief gestation, the tammar wallaby gives birth to an altrical young and begins a comparatively long lactation (Renfree, M.B. 1983, Tyndale-Biscoe, C.H. and Janssens, P.A. 1988) with the quantity and composition of the milk changing over time (Nicholas, K. 1988). The changes in milk composition concur with putative changes in the suckling behaviour of the pouch young (Nicholas, K., Simpson, K., et al. 1997). Consequently, the tammar wallaby lactation cycle was divided into 3 phases, with the second phase divided into 2 distinct parts (Nicholas, K., Simpson, K., et al. 1997, Tyndale-Biscoe, C.H. and Janssens, P.A. 1988). The first phase encompasses a 26.5 day

pregnancy (Findlay, L. 1982), whilst phase 2A, the first part of phase 2, covers the initial ~110 days after birth. Phase 2A is characterised by the permanent attachment of the pouch young to a single teat, which results in the continued lactation of this gland while the remaining 3 glands regress. The second part of phase 2, phase 2B, is marked by the surrender of the teat by the pouch young, followed by intermittent and less frequent suckling for a further ~100 days. Phase 3 typically begins at approximately day 200 when the pouch young first starts to regularly leave the pouch, and sucking becomes less frequent but more vigorous.

The observed changes in tammar wallaby milk composition are presumably necessary for the optimal growth and development of the pouch young (Dove, H. and Cork, S.J. 1989, Green, B., Merchant, J., et al. 1988, Nicholas, K. 1988, Trott, J.F., Wilson, M.J., et al. 2002). In phase 2 of lactation, tammar wallaby milk is low in protein and lipids, and very high in carbohydrates (Green, B., Merchant, J., et al. 1988, Nicholas, K. 1988), which are primarily oligosaccharides, rather than lactose, which is typically a significant component of eutherian milk (Kunz, C., Rudloff, S., et al. 2000). Like human milk, tammar wallaby milk is rich in sialylated oligosaccharides (Messer, M. and Urashima, T. 2002, Urashima, T., Saito, T., et al. 2001, Urashima, T., Saito, T., et al. 1994). Sialic acid is a significant component of tammar wallaby milk, accounting for ~4-5% of the monosaccharide composition of total carbohydrate throughout most of lactation, and rising to ~9% just prior to weaning (Green, B. and Merchant, J.C. 1988). During phase 3, the sugar molecules in tammar milk become progressively smaller, until, at about day 280, milk sugars are present only as monosaccharides. This is concomitant with an overall decrease in milk carbohydrate and an increase in protein and lipids (Green, B. and Merchant, J.C. 1988).

The emphasis on lactation in marsupial reproduction can be regarded as an alternative to the gestation-focused reproductive strategy of eutherian mammals. Therefore, including

the tammar in a comparative approach with the human and other species provides a useful approach to investigate the potential role of sialylated oligosaccharides in early development (Tyndale-Biscoe, C.H. and Janssens, P.A. 1988). During phase 2 of lactation, particularly phase 2A, the growth and development of the pouch young is solely dependent on nutrients delivered via the milk and it has been suggested that altricial neonates have a particular requirement for milk oligosaccharides (Oftedal, O.T., Alt, G.L., et al. 1993). The phenomenon of asynchronous concurrent lactation demonstrated by Nicholas (1988), suggests that local gene expression patterns determine the output of a particular mammary gland. Furthermore, a study by Trott et al. (2003) demonstrated that, irrespective of the age of the pouch young, it is the lactating tammar wallaby that regulates both milk composition and rate of milk production, thus determining pouch young growth and development. Therefore, studying sialic acid biosynthesis and sialyltransferase gene expression in the lactating tammar mammary gland will most likely provide new information on the role of sialic acid and other sialylglycoconjugates in neonatal development.

Sialic acid is synthesised by a series of dedicated enzymes from sugar precursors (Varki, A. and Schauer, R. 2009) and sialylglycoconjugates are subsequently synthesised by sialyltransferases, which are a family of Golgi membrane-bound glycosyltransferases that catalyse the transfer of sialic acid from CMP-sialic acid to carbohydrate acceptors on glycoproteins, glycolipids and free oligosaccharides (Harduin-Lepers, A., Mollicone, R., et al. 2005). Sialyltransferases are mainly regulated at the transcriptional level (Svensson, E.C., Soreghan, B., et al. 1990, Takashima, S., Kono, M., et al. 2000, Wang, X., O'Hanlon, T.P., et al. 1990, Wen, D.X., Svensson, E.C., et al. 1992) and can be divided into 4 groups, ST6Gal, ST6GalNAc, ST3Gal and ST8Sia, based on the type of linkage formed between sialic acid and the preferred acceptor (Harduin-Lepers, A., Mollicone, R., et al. 2005, Harduin-Lepers, A., Vallejo-Ruiz, V., et al. 2001). Sialic acid biosynthesis and sialyltransferase genes have not been previously characterised in the tammar wallaby. In the current study we have obtained near full-length cDNA sequences for several tammar

sialic acid biosynthesis and sialyltransferase genes and compared their predicted amino acid sequences to those of other species. We also report the expression profiles of sialic acid biosynthesis and sialyltransferase genes in the lactating wallaby mammary gland and show a correlation with previously reported data on sialic acid and sialylated structures in milk.

5.2.3 Methods

Animals

Tammar wallabies were maintained in open enclosures at the Department of Primary Industries (Attwood, VIC, Australia) and The University of Melbourne Macropod Research Facility (Wantirna, VIC, Australia), with all animals sourced originally from Kangaroo Island, South Australia. Mammary glands from pregnant - phase 1 (days -18, -21, -25), lactating - phase 2A (days 1.75, 2, 3, 5, 10, 15, 22, 37, 62, 70, 80, 100, 110), lactating - phase 2B (days 133, 150, 171, 180, 193), lactating - phase 3 (days 216, 220, 260) and involuting (day 5) tammar wallabies were excised under sterile conditions after the animals were euthanised. A single animal was sacrificed at each time point. All tissue was frozen at -80 deg C for later use. The day of gestation was scheduled by removing the pouch young early in the breeding season (February-April) to reactivate the dormant blastocysts, with births expected 26-27 days later (Findlay, L. 1982). Animal pouches were checked daily for several days around the time of the expected birth, until a pouch young was detected (lactation, day 1). The day of lactation was subsequently determined either using the known date of birth or a head length measurement of the pouch young to estimate its age (Poole, W.E. 1991). All animal experimentation was approved by and performed according to the Institutional Animal Ethics Committees.

RNA extraction

Total RNA was isolated from approximately 100mg of wallaby tissue using the RNeasy Lipid Tissue Mini Kit from Qiagen. The RNA samples were quantified using the Biolab NanoDrop and checked for integrity on agarose gels.

Reverse Transcription PCR and sequencing

A 2x coverage of the wallaby genome has been assembled and is available from Ensembl (Meug_1.0). We verified the identity of our genes of interest by designing primers using publicly available sequences from Ensembl in the vicinity of start and stop codons, allowing for amplification of the longest possible cDNA fragment (Table 1). PCR products were electrophoretically separated on 2% agarose gels and subsequently excised and direct sequenced in both the forward and reverse direction.

Total RNA (2ug) of lung, liver, heart, kidney and spleen was used to synthesise cDNA with the SuperScript III First-Strand cDNA Synthesis System and Oligo$(dT)_{15}$ (Promega), according to the instructions supplied by Invitrogen. A PCR was performed on each sample using the primers listed in Table 1. Each final 25µl reaction mixture consisted of 400 nM forward primer, 400 nM reverse primer, 2µl of 1 in 10 diluted cDNA and 12.5µl of the 2X GoTaq Master Mix (Promega). PCR parameters for the MJ Research PTC-225 Pertier Thermal-Cycler were: 95^{o}C for 2 min, 40 cycles at 95^{o}C for 30 sec, then 58^{o}C for 30 sec and 72^{o}C for 2.5 min, and finally, 72^{o}C for 5 min.

The PCR products were subsequently separated by electrophoresis on a 2% agarose gel, excised and purified and direct sequenced at the Australian Genome Research Facility (Brisbane) in forward and reverse directions. Where necessary further rounds of sequencing were performed using internal primers designed for qPCR. Contigs were constructed by combining the forward and reverse sequences using the Geneious software package (Drummond, A., Ashton, B., et al. 2010). Geneious was then used to translate the contigs representing the longest possible cDNA fragment for each gene of interest into an amino acid sequence. Amino acid sequences were aligned with their orthologues using the Genious multiple sequence alignment algorithm with default settings. Trees were built from the multiple alignments using Genious Tree Builder with the Jukes-Cantor genetic distance model, using the Neighbour-joining method.

Primer design for quantitative PCR analysis

Wherever possible, primers were designed to span introns to facilitate detection of any genomic DNA contamination. Primers for the 6 candidate reference genes were designed based on publicly available sequences; primers for target genes were designed from contigs assembled from sequenced cDNA (Table 2). Primer specificity and amplification efficiency were also validated empirically with agarose gel electrophoresis and melting curve analysis, and a standard curve analysis of a 10-fold dilution series.

Quantitative PCR

cDNA was generated from 2µg total RNA using the Invitrogen SuperScript VILO cDNA Synthesis Kit as per the manufacturer's instructions. qPCR was performed according to instructions for the use of EXPRESS SYBR GreenER qPCR SuperMix with Premixed ROX supplied by Invitrogen. Reaction mixtures of 20µl consisted of 200 nM forward primer, 200 nM reverse primer, 2µl of 1:10 diluted cDNA and 10µl of the 2X EXPRESS SYBR GreenER qPCR SuperMix with Premixed ROX. Cycling parameters for the Eppendorf RealPlex MasterCycler were 50^{o}C for 2 min, 95^{o}C for 10 min, 40 cycles at 95^{o}C for 15 sec then 60^{o}C for 1 min, followed by a melt-curve dissociation cycle at 95^{o}C for 15 sec then 60^{o}C for 1 min and a 20 min ramp up to 95^{o}C for 15 sec.

Analysis of expression stability

Expression stabilities of selected reference genes were evaluated using the publicly available geNorm VBA applet for Microsoft Excel (http://medgen.ugent.be/~jvdesomp/genorm/) (Vandesompele, J., De Preter, K., et al. 2002). geNorm determines the stability of candidate reference genes according to the similarity of their expression profile. Using a pair-wise comparison, it calculates an M-value that reflects the stability of each gene, where the gene with the largest M-value is the least stable. The raw expression data for each candidate gene was exported from the

RealPlex software in the geNorm format and an automated geNorm analysis was performed in Microsoft Excel. Based on resulting M-values and pairwise variation, V, of less than 0.15, it was determined that the 2 most stable reference genes were adequate for the normalisation of our system.

Normalisation of quantitative PCR results

The raw Ct values were imported into Biogazelle's qBasePlus software for calculation of Calibrated Normalized Relative Quantities (CNRQs) and their standard errors. The CRNQs were used in subsequent statistical analyses.

Statistical analysis

Significant differences in gene expression in the wallaby mammary gland between the phases of lactation were determined using paired t-tests, assuming unequal variances. For all analyses, a P-value of less than 0.05 was interpreted as statistically significant. All results are expressed as the mean of available replicates, and the mean±S.E.M. is included in the figures.

Table 1

PCR and sequencing primers

Gene	Forward Primer	Reverse Primer
GNE	ATGAAGAATGGGAACAACCGAAAG	CTCTACGTCCATCACTGTACTCG
NANS	ATGCCGCTGGAGCTG	CAATAAATGCCACTTTTCTTTAAGATT
ST6Gal I	CCCAGGGACACTGTATGCTT	TTGCCCAGATGTTAGCAGTG
ST3Gal I	GGAAAAGATGGCAGCGATAA	ATGGTGCCTTTGGTCATAGC
ST6GalNAc II	GAGCCTCCTGAACACCTCAG	CCTGCTTCATGCAAATTCCT

Table 2

Primers used for qPCR

Gene	Forward Primer	Reverse Primer	Product (bp)
Blg	GCATGTGCCCACTATGTCAG	AGGGGGATAACTTCGTCTGC	154
GNE	GCCCTACAGCATCCTGTGACCA	TGCTCCCACTTCTCGCACTCCA	282
NANS	CGTAGGTCTTCCCCCAGGAGTGTT	TCATCGCCGAGATCGGGCAGAA	183
ST6Gal I	CTGGCCCTTGGTTGAGGTGCTT	GCAGCCCTTTTACATCCTCAGCCC	290
ST3Gal I	TGGTGAGTGGTCTTCCTCCCGA	CGTCGACCCATTCCTGGGGAAA	183
ST6GalNAc II	TGGACCAGACGCCTCTGCAAGT	CCGTCAGCAACATGAGAGCCCC	149
RPS15	GCGGAAGTGGAGCAGAAGAAGAAGC	AGGGAATGCTGTTTGCGCCG	164
RPL13A	TTGCATATCTGGGACGGCTGGC	ACATTCTTCTCAGCTTGCTTGCGG	151
GAPDH	GCTGCATTCACCTCAGGTCACG	GCTTTTCCATTGATCACCAGCTTCCCA	158
RPS9	TGCTCATCCGCCAGAGGCACAT	TCCCTGGCCCTTCTTGGCATT	167
RPS23	TGCCGTGGTCTTCGTACTGCT	TGGCAGAATTGGGCTGTTTGGC	187
UXT	TCGGGGGATCAGAGACTGCACA	TGAGCAGGCGGCTTTTCCGA	169

All shaded genes were tested as possible reference genes. Those eventually used as references were RPS15 and RPL13A.

5.2.4 Results

Amino acid sequence analysis

A 2x coverage of the wallaby genome has been assembled and is available from Ensembl (Meug_1.0). The current build of the wallaby genome was automatically annotated using a low-coverage gene-build method and contains a total of 17,877 genes, of which 15,290 are protein coding. Given the relatively low coverage of the wallaby genome and the presence of large gaps in many of the scaffolds, we verified the identity of our genes of interest using Sanger sequencing. Primers were designed using the sequence available from Ensembl to amplify the largest possible cDNA fragment for each gene of interest. The resulting PCR products were visualised by gel electrophoresis and subsequently excised and sequenced in both directions. Contigs were constructed by combining the forward and reverse sequences using the Geneious software package (Drummond, A., Ashton, B., et al. 2010). Geneious was then used to translate the contigs representing the longest possible cDNA fragment for each gene of interest into amino acid sequence, which was then used in multiple sequence alignments with orthologues from other species.

The GNE protein sequence is 93.5% conserved between all 6 species compared (Figure 1). This high level of conservation is consistent with GNE functionality being essential for sialylation, particularly in development, as it has been shown that inactivation of the GNE gene causes early embryonic lethality in mice (Schwarzkopf, M., Knobeloch, K.-P., et al. 2002). Of the observed amino acid sequence variations, there are several occurrences where one variant is present in the opossum and wallaby and the alternative in all other species. Given that all of the other species used in the comparison are eutherian mammals, the variation at these particular residues appears to be a differentiator between marsupial and eutherian GNE sequence. Predictably, the wallaby sequence is most closely related to the opossum sequence and the eutherian sequences are more closely related to each other than to the marsupial sequences (Figure 1). Within the eutherian group, the mouse and rat

GNE sequences are most closely related to each other and then to the Chinese hamster. The human sequence was more closely related to the Chinese hamster than to the other 2 rodents.

The NANS protein sequence is not as well conserved between the species investigated than the GNE sequence, with only 66% identity across all species (Figure 2). However, the pairwise identity between sequences is 84.9%. The majority of the amino acid sequence variation is in the second half of the protein, with the first 250 residues showing approximately 75% identity across all the species surveyed. In contrast, the remaining 109 residues are only approximately 50% identical, with overall identity dropping to about 20% for the last 33 residues. The phylogenetic analysis of the sequences shows that the wallaby and the opossum sequences are most closely related, as are the rat and the mouse sequences, the pig and the human sequences, and the African clawed frog and fugu sequences (Figure 2).

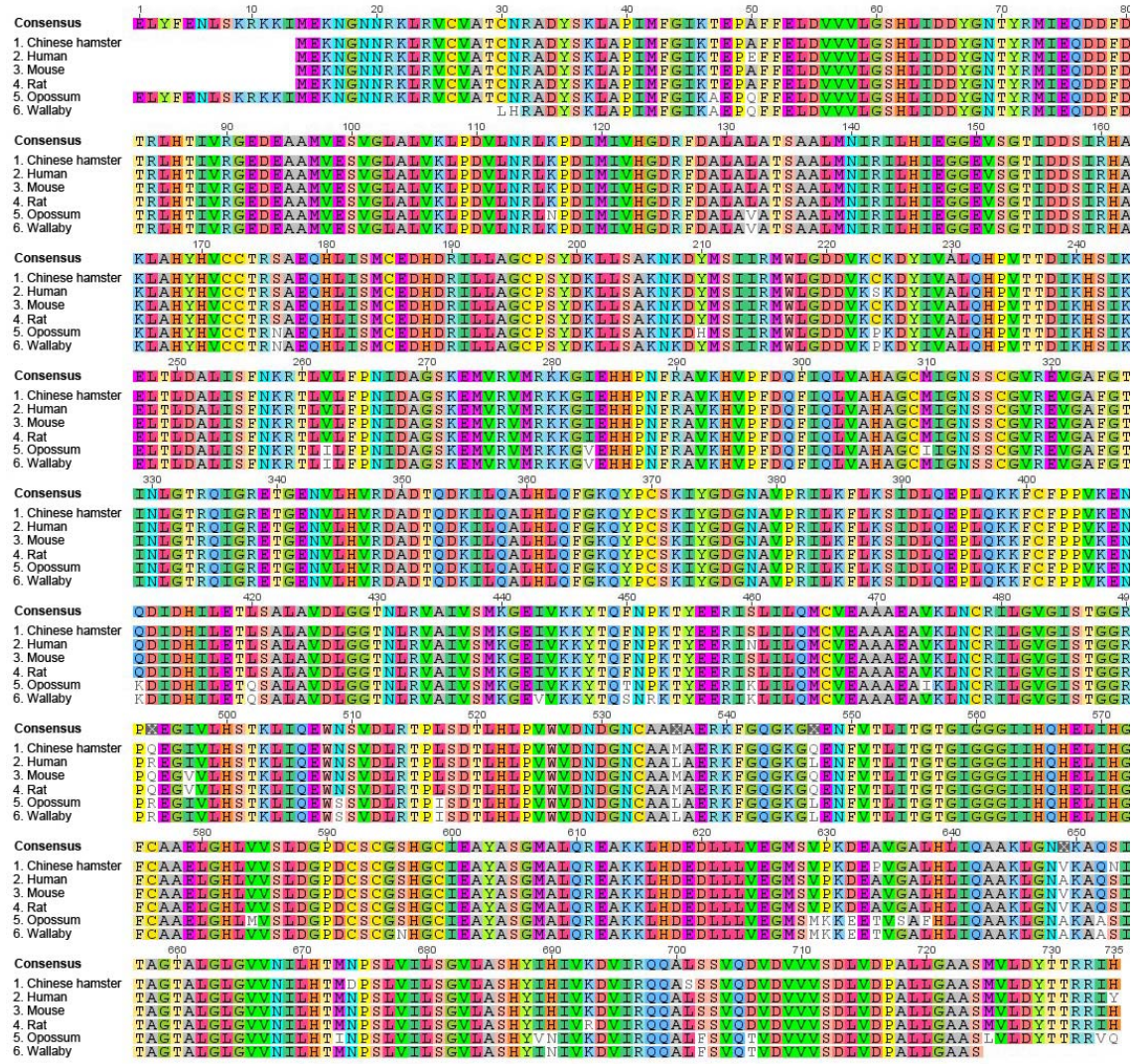
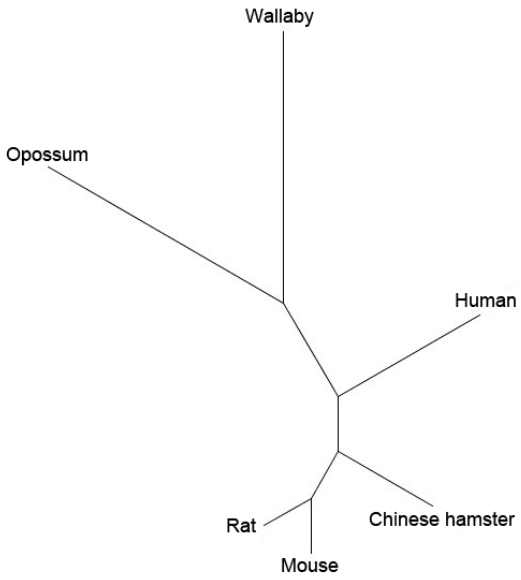
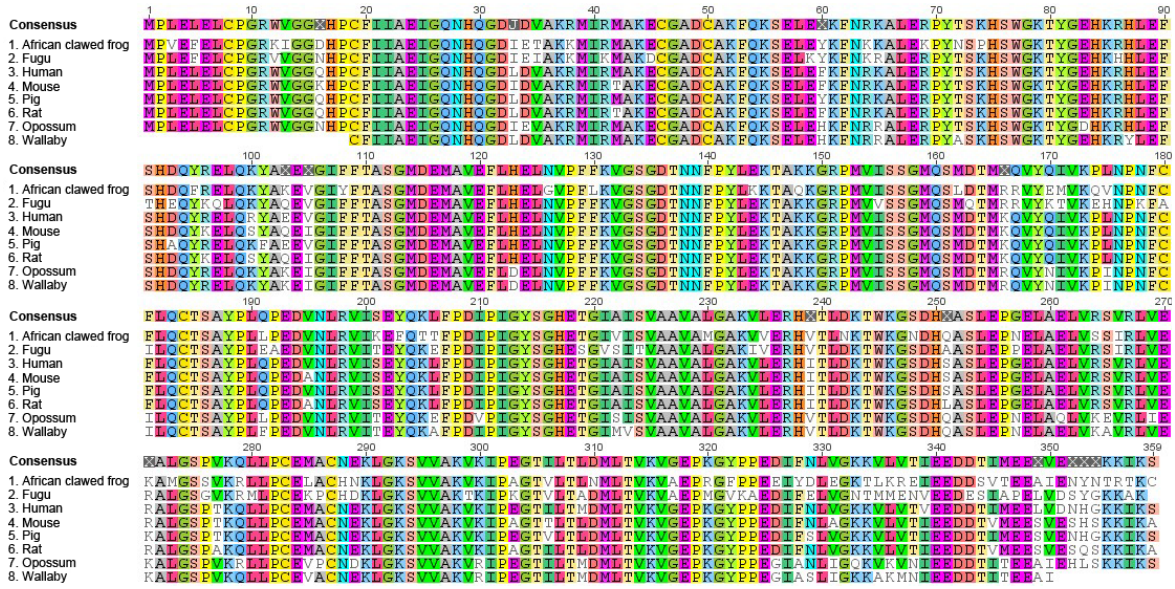


Figure 1: Predicted wallaby GNE amino acid sequence aligned with orthologous GNE sequences with Neighbour-joining tree representation of sequence phylogeny. Wallaby GNE amino acid sequence was predicted from cDNA sequence. GNE amino acid sequences of other species used in the alignment were obtained from SwissProt, except for the Opossum sequence, which was obtained from Ensembl. Amino acid sequences were aligned with their orthologues using the Genious multiple sequence alignment algorithm with default settings. Trees were built from the multiple alignments using Genious Tree Builder with the Jukes-Cantor genetic distance model, using the Neighbour-joining method. The amino acid sequence is shown in single letter code and each amino acid is assigned its own colour. White shading indicates an amino acid variant.





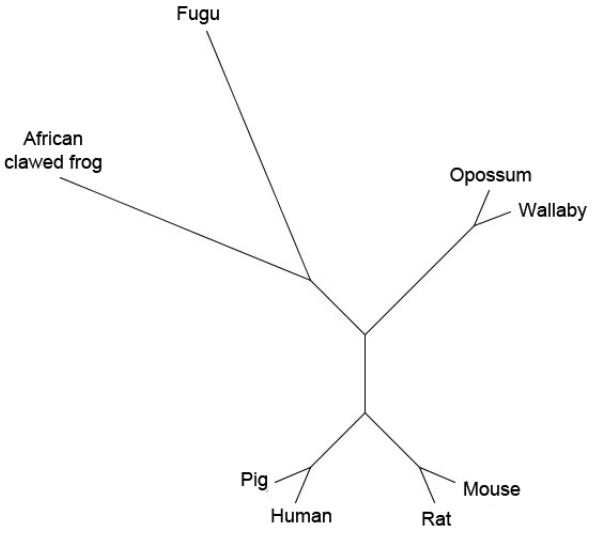


Figure 2: Predicted wallaby NANS amino acid sequence aligned with orthologous NANS sequences with Neighbour-joining tree representation of sequence phylogeny. Wallaby NANS amino acid sequence was predicted from cDNA sequence. NANS amino acid sequences of other species used in the alignment were obtained from SwissProt, except for the Opossum sequence, which was obtained from Ensembl. Amino acid sequences were aligned with their orthologues using the Genious multiple sequence alignment algorithm with default settings. Trees were built from the multiple alignments using Genious Tree Builder with the Jukes-Cantor genetic distance model, using the Neighbour-joining method. The amino acid sequence is shown in single letter code and each amino acid is assigned its own colour. White shading indicates an amino acid variant.

The overall sequence identity of the ST6Gal I sialyltransferase protein between species is a relatively low 43%; although, the sequence variability is not uniform for the entire length of the protein (Figure 3). Most of the variability is present in the first half of the ST6Gal I protein sequence where overall identity is as low as 12%. The second half of the protein sequence is more conserved, with an overall identity of approximately 48%. This observation is typical of sialyltransferases, which generally have a very low overall sequence identity (Harduin-Lepers, A., Mollicone, R., et al. 2005) across species, mainly due to differences in the length and amino acid sequence of the highly variable stem region. The length and amino acid composition of sialyltransferase catalytic domains are relatively well conserved (Jeanneau, C., Chazalet, V., et al. 2004). In the more conserved segment of the ST6Gal I protein sequence (Figure 3), we identified the three sialylmotifs, L, S and VS, previously described by Datta and Paulson (1997) and Geremia et al (1997). These motifs are the characteristic of eukaryotic sialyltransferases. Despite the high sequence variability, the phylogenetic analysis shows that, as expected, the marsupial sequences are most closely related to each other, as are the 2 rodent sequences. The human sequence is more closely related to the rodent sequences than the marsupial, and the chicken and the fugu sequences are more related to each other than those of the other species (Figure 3).

We observe a similar pattern of sequence variability in the ST3Gal I protein sequence as we do for the ST6Gal I sequence. Overall sequence identity is 41.6%, although, once again, it is the first half of the protein that contributes to the bulk of the variability, with only 17.4% identity in the region. The second half of the ST3Gal I amino acid sequence is 58.4% conserved and contains the L, S and VS sialylmotifs (Figure 4). The phylogenetic relationship between the sequences is expected, with the human and chimpanzee sequences most closely related to each other, and then to the pig. The rodent sequences also cluster together, as do the marsupials. The fugu and the chicken sequences were more related to each other than to any of the other groups (Figure 4).

Similar to ST6Gal I and ST3Gal I, sequence conservation across the ST6GalNAc II amino acid sequence averages 44%. Similar to the other sialyltransferases, sequence conservation is confined mainly to the region containing the L, S and VS sialylmotifs in the C-terminal half of the protein, where sequence identity is approximately 55% (Figure 5). For this protein our phylogenetic analysis utilised only the region containing the sialylmotifs as a large portion of the N-terminal sequence was unavailable. This analysis showed that the wallaby and opossum sequences are the most closely related, as are the mouse and rat sequences. The human and cow sequences were most closely related to each other and then to the rodent sequences. The chicken sequence was the most diverged from the other groups.

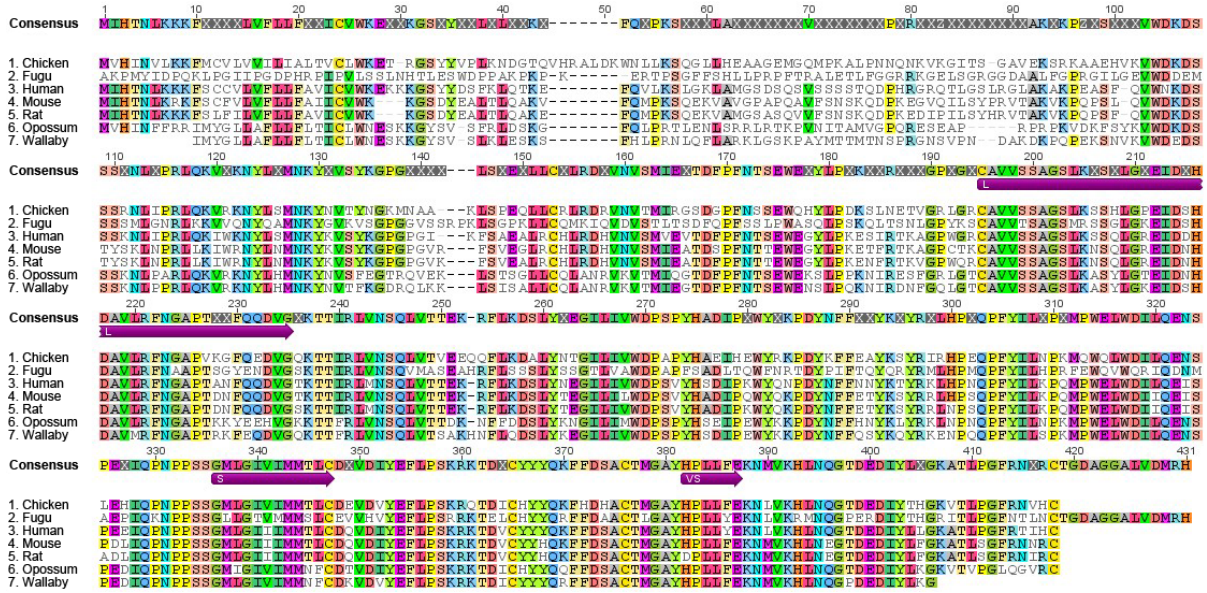
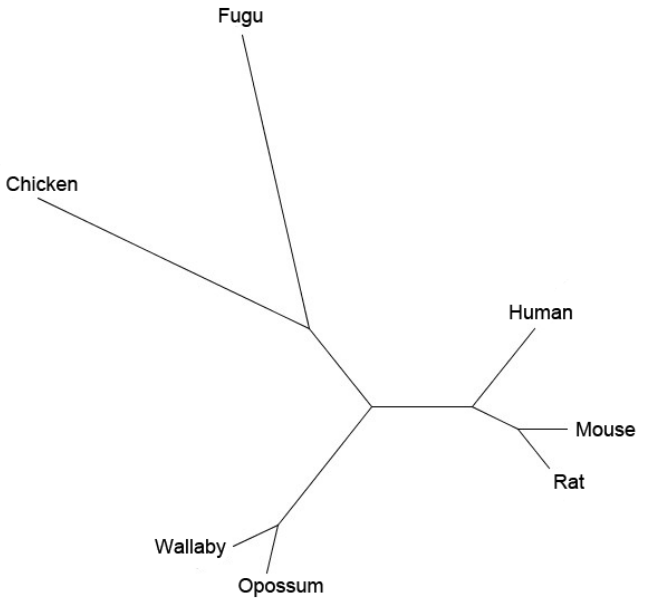


Figure 3: Predicted wallaby ST6Gal I amino acid sequence aligned with orthologous ST6Gal I sequences with Neighbour-joining tree representation of sequence phylogeny. Wallaby ST6Gal I amino acid sequence was predicted from cDNA sequence. ST6Gal I amino acid sequences of other species used in the alignment were obtained from SwissProt, except for the Opossum sequence, which was obtained from Ensembl. Amino acid sequences were aligned with their orthologues using the Genious multiple sequence alignment algorithm with default settings. Trees were built from the multiple alignments using Genious Tree Builder with the Jukes-Cantor genetic distance model, using the Neighbour-joining method. The amino acid sequence is shown in single letter code and each amino acid is assigned its own colour. White shading indicates an amino acid variant. The positions of sialylmotifs L (CxxVxxxxxLxxxxxGxxxxxxxxxxRxxxxxxx xxxxDVG), S (GxxxxxxxxxxC) and VS (HxxxxE) are highlighted by the large arrows above the sequence alignment.



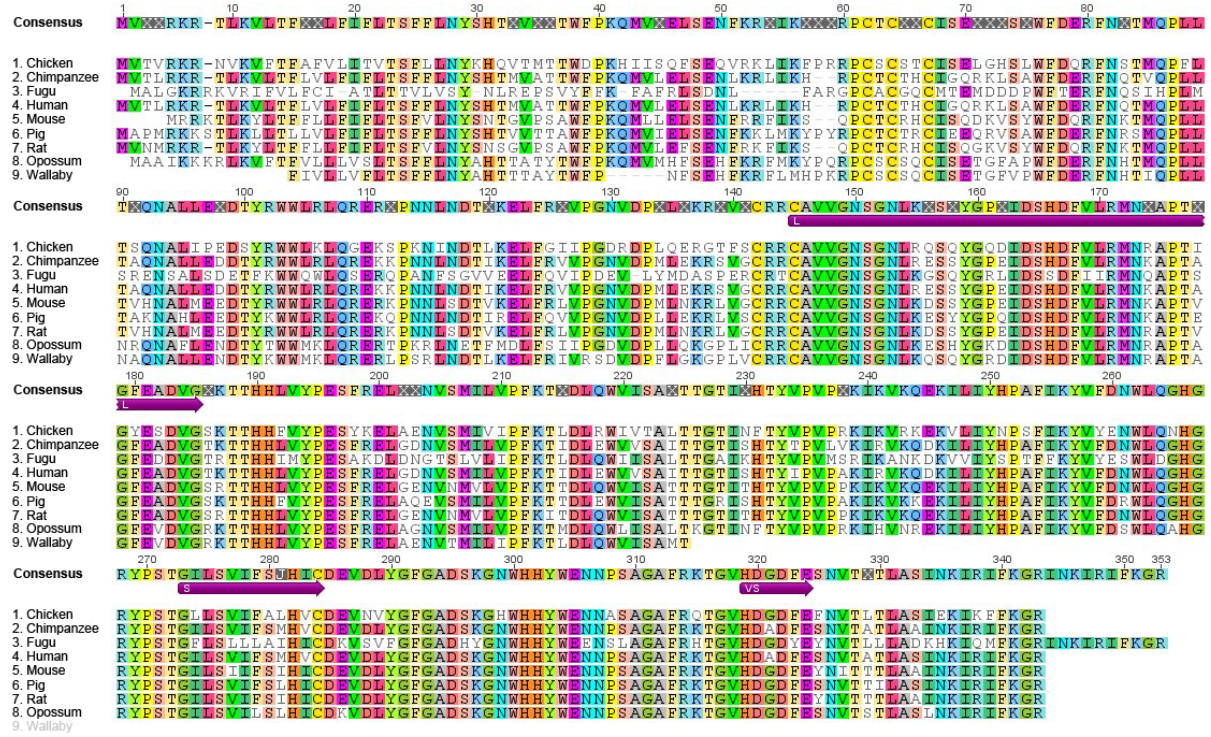
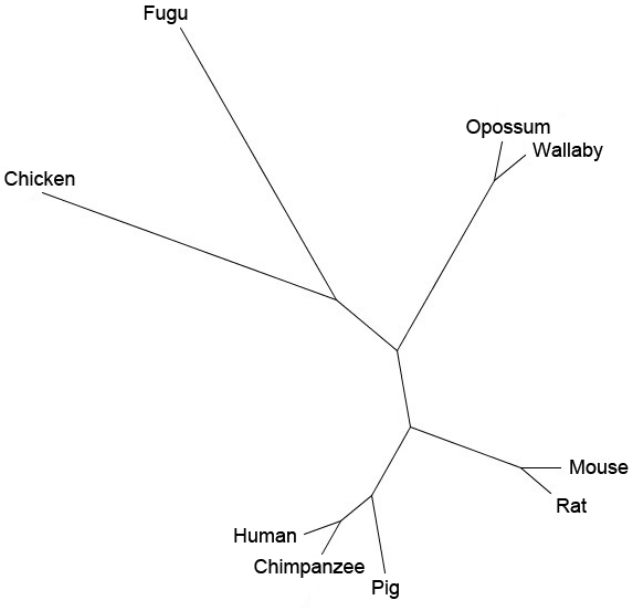
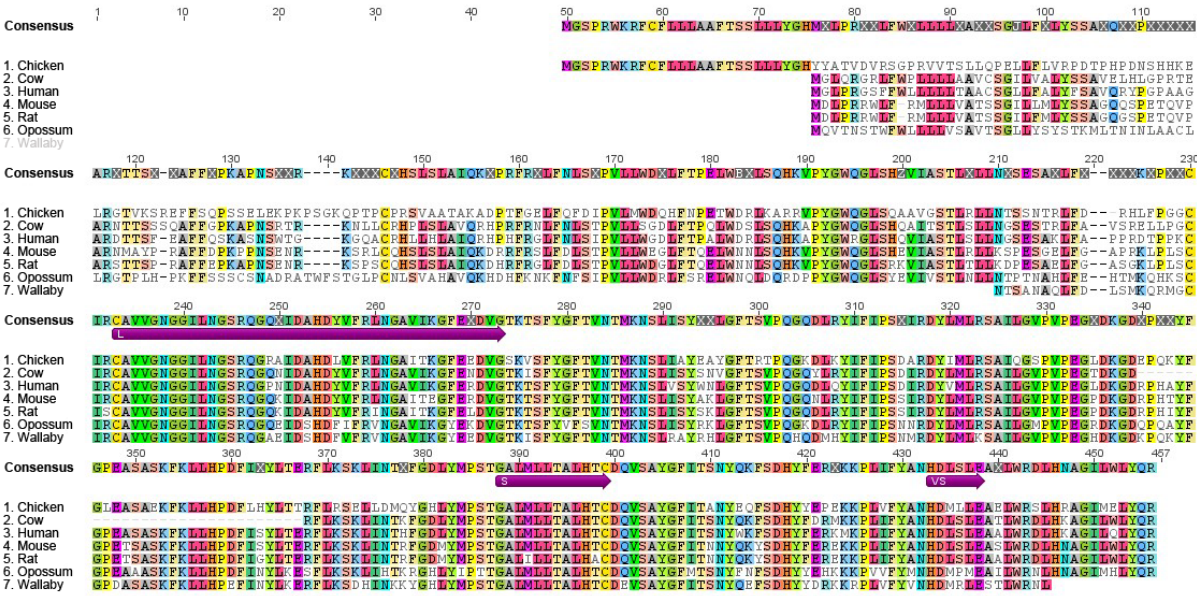


Figure 4: Predicted wallaby ST3Gal I amino acid sequence aligned with orthologous ST3Gal I sequences with Neighbour-joining tree representation of sequence phylogeny. Wallaby ST3Gal I amino acid sequence was predicted from cDNA sequence. ST3Gal I amino acid sequences of other species used in the alignment were obtained from SwissProt, except for the Opossum sequence, which was obtained from Ensembl. Amino acid sequences were aligned with their orthologues using the Genious multiple sequence alignment algorithm with default settings. Trees were built from the multiple alignments using Genious Tree Builder with the Jukes-Cantor genetic distance model, using the Neighbour-joining method. The amino acid sequence is shown in single letter code and each amino acid is assigned its own colour. White shading indicates an amino acid variant. The positions of sialylmotifs L (CxxVxxxxxLxxxxxGxxxxxxxxxxRxxxxxxxxxxDVG), S (GxxxxxxxxxxC) and VS (HxxxxE) are highlighted by the large arrows above the sequence alignment.





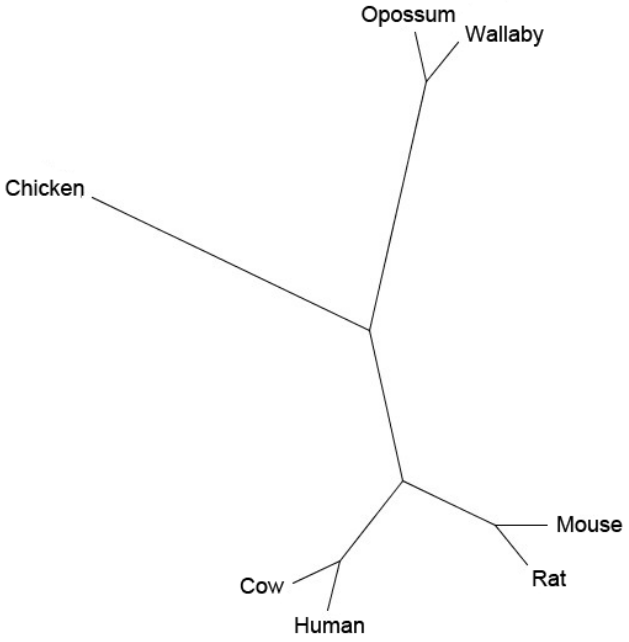


Figure 5: Predicted wallaby ST6GalNAc II amino acid sequence aligned with orthologous ST6GalNAc II sequences with Neighbour-joining tree representation of sequence phylogeny. Wallaby ST6GalNac II amino acid sequence was predicted from cDNA sequence. ST6GalNAc II amino acid sequences of other species used in the alignment were obtained from SwissProt, except for the Opossum sequence, which was obtained from Ensembl. Amino acid sequences were aligned with their orthologues using the Genious multiple sequence alignment algorithm with default settings. Trees were built from the multiple alignments using Genious Tree Builder with the Jukes-Cantor genetic distance model, using the Neighbour-joining method. The amino acid sequence is shown in single letter code and each amino acid is assigned its own colour. White shading indicates an amino acid variant. The positions of sialylmotifs L (CxxVxxxxxLxxxxxGxxxxxxxxxxRxxxxxxx xxxxDVG), S (GxxxxxxxxxxC) and VS (HxxxxE) are highlighted by the large arrows above the sequence alignment.

Gene expression analysis

The expression of sialic acid biosynthesis and sialyltransferase genes in the wallaby mammary gland was assessed throughout the lactation cycle. Mammary gland samples were obtained from several animals at each phase of wallaby lactation; phase 1 (n=3), phase 2A (n=13), phase 2B (n=5) and phase 3 (n=3). Tissue was available from only a single animal at involution (n=1). The expression level of our genes of interest was determined for each sample using quantitative PCR (qPCR). To ensure that we were capturing the overall expression level for each gene and not individual isoforms, which, for sialyltransferase genes, are heterogeneous in their 5' regions, primers were designed from the second exon or beyond and, where possible, across exon-exon boundaries.

Figure 6 shows the mean normalised expression level of the β-lactoglobulin (Blg) gene in the wallaby mammary gland during the 3 phases of lactation. Blg is an extensively studied whey protein that was used in this study as a positive control to validate the known expression pattern for the Blg gene. Blg gene expression was not detected in the wallaby mammary gland during pregnancy, which is phase 1 of the lactation cycle. The expression level of the Blg gene subsequently increased in phases 2A and 2B, and was highest in phase 3, before declining considerably in involution. This is consistent with the expression pattern for the Blg gene in wallaby mammary gland previously reported by Collet et al. (1991) and Bird et al. (1994), confirming the integrity of our qPCR analysis.

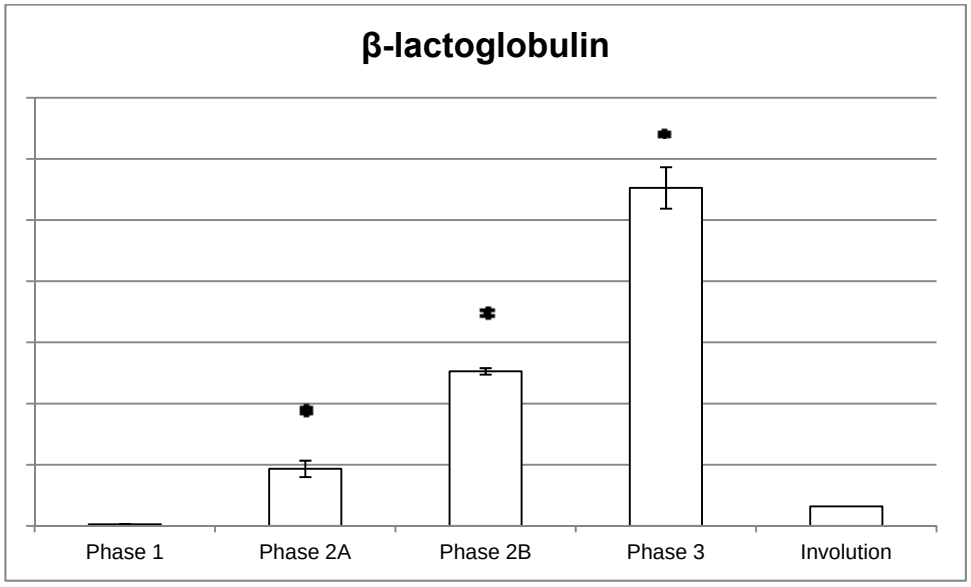


Figure 6: qPCR expression analysis of the β-lactoglobulin (Blg) gene in wallaby mammary tissue harvested throughout the lactation cycle . Phase 1 (n=3) is represented by mammary tissue from animals at 25, 21, and 18 days *before birth*, phase 2A (n=13) is represented by tissue from animals at days 1.75, 2, 3, 5, 10, 15, 22, 37, 62, 70, 80, 100 and 110 of lactation, phase 2B (n=5) is represented by tissue from animals at days 133, 150, 171, 180 and 193 of lactation, whilst phase 3 (n=3) encompassed tissue from days 216, 220 and 260 of lactation. The involution sample was taken 5 days post weaning (n=1). As each sample was obtained from a different animal, samples within each phase were treated as biological replicates. The mean calibrated normalised relative expression level of the Blg gene at each phase of lactation is shown including the standard error of the mean. A statistically significant difference in expression between each phase and the preceding phase is denoted by a "*".

GNE is a bifunctional enzyme that catalyses the first committed, and rate limiting step in sialic acid biosynthesis (Gal, B., Ruano, M.J., et al. 1997), making it a key regulator of sialylation.

Figure 7 shows that the GNE gene is expressed at a low level during phase 1, however its level of expression increases significantly ($p<0.05$) in phase 2A, and again in phase 2B. The apparent further increase in phase 3 was not statistically significant. GNE gene expression was seen to dramatically decline in involution, 5 days after weaning.

NANS is another of the key enzymes uniquely involved in sialic acid biosynthesis [reviewed in (Tanner, M.E. 2005)]. Similar to the GNE gene, the NANS gene is expressed at a low level in the wallaby mammary gland in phase 1 of lactation. The expression level of the NANS gene then significantly increases in each subsequent stage of lactation ($p<0.05$), culminating in phase 3. The NANS gene is dramatically down-regulated in the early stages of involution (Figure 7).

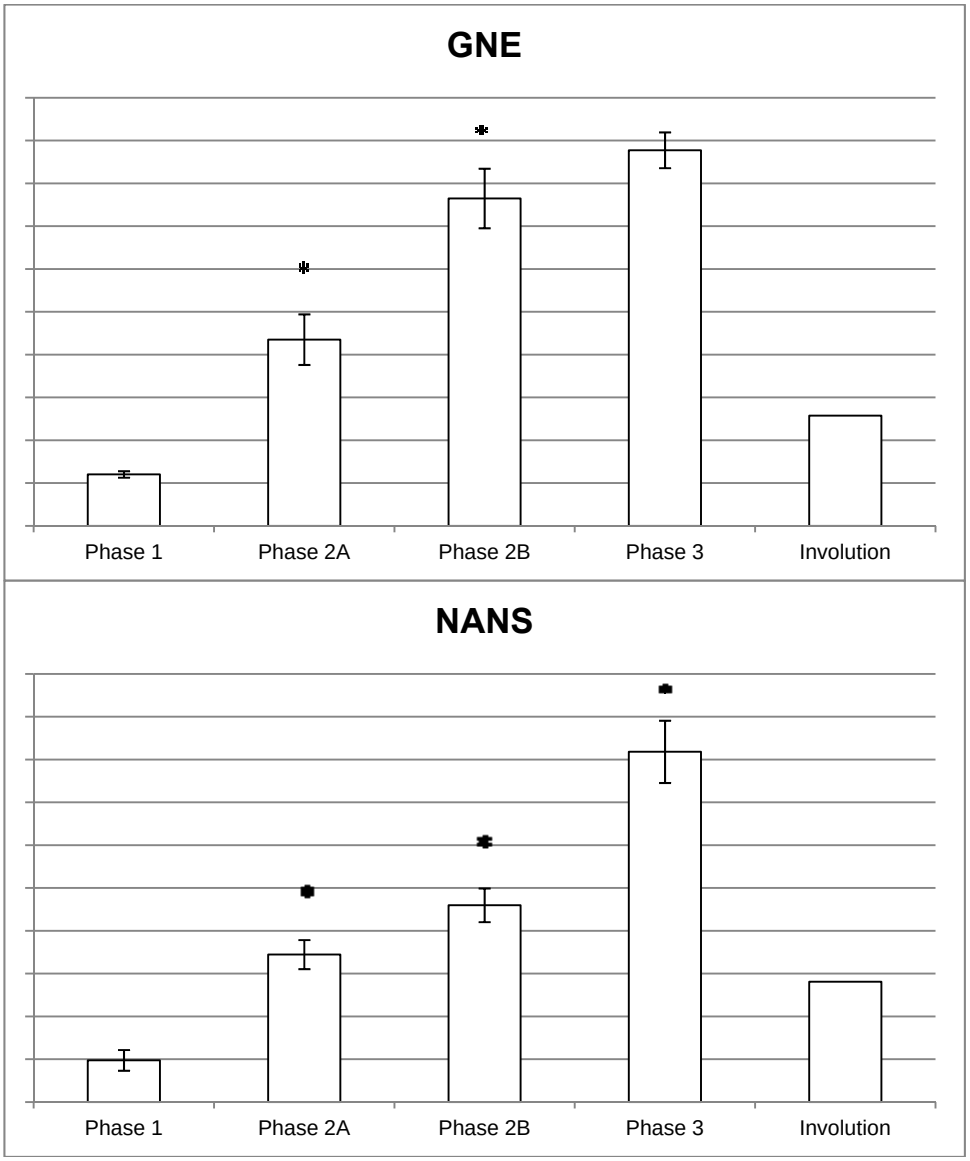


Figure 7: qPCR expression analysis of the GNE and NANS genes in wallaby mammary tissue harvested throughout the lactation cycle. Phase 1 (n=3) is represented by mammary tissue from animals at 25, 21, and 18 days *before birth*, phase 2A (n=13) is represented by tissue from animals at days 1.75, 2, 3, 5, 10, 15, 22, 37, 62, 70, 80, 100 and 110 of lactation, phase 2B (n=5) is represented by tissue from animals at days 133, 150, 171, 180 and 193 of lactation, whilst phase 3 (n=3) encompassed tissue from days 216, 220 and 260 of lactation. The involution sample was taken 5 days post weaning (n=1). As each sample was obtained from a different animal, samples within each phase were treated as biological replicates. The mean calibrated normalised relative expression level of each gene, at each phase of lactation is shown including the standard error of the mean. A statistically significant difference in expression between each phase and the preceding phase is denoted by a "*".

Figure 8 outlines the expression pattern of the ST6Gal I gene in the lactating wallaby mammary gland. It shows barely detectable levels of ST6Gal I gene expression in phase 1 of lactation, followed by a statistically significant increase in phase 2A ($p<0.05$). ST6Gal I gene expression peaks in phase 2B of the wallaby lactation cycle, followed by a down-regulation of expression in phase 3, although observed levels are still higher than those in phase 2A. As with the other genes tested, ST6Gal I gene expression declines dramatically in involution.

The α2,3-sialyllactose is the dominant oligosaccharide in mouse milk (Duncan, P.I., Raymond, F., et al. 2009). ST3Gal I catalyses the transfer of sialic acid to Galβ1,3GalNAc disaccharides via an α2,3 linkage and has been shown to be able to use lactose as an acceptor *in vitro* (Kono, M., Ohyama, Y., et al. 1997, Paulson, J.C., Rearick, J.I., et al. 1977, Sherblom, A.P. and Bourassa, C.R. 1983). In the lactating wallaby mammary gland analysed here, the ST3Gal I gene expression profile is similar to that of the GNE and NANS genes whereby its expression peaks in phase 3 of lactation (Figure 8). ST3Gal I gene expression is barely detectable in phase 1, but then gradually increases in phases 2A, 2B and 3 and its expression level decreased soon after weaning.

ST6GalNAc II primarily initiates the transfer of sialic acid to GalNAc-, Galβ1,3GalNAc-, and Siaα2,3Galβ1,3GalNAc-structures on O-glycans of glycoproteins (Harduin-Lepers, A., Vallejo-Ruiz, V., et al. 2001, Takashima, S. 2008). It has a demonstrated activity towards asialo ovine submaxillary mucin, implying a potentially vital role in the mammary gland (Kono, M., Tsuda, T., et al. 2000). In the wallaby mammary gland, the ST6GalNAc II gene is expressed at comparable levels in phase 1 and phase 3 of lactation, and involution. In comparison to phase 1, its expression significantly increases in phase 2A ($p<0.05$). Phase 2B ST6GalNAc II expression levels are also significantly higher than those observed in phase 2A ($p<0.05$) (Figure 8).

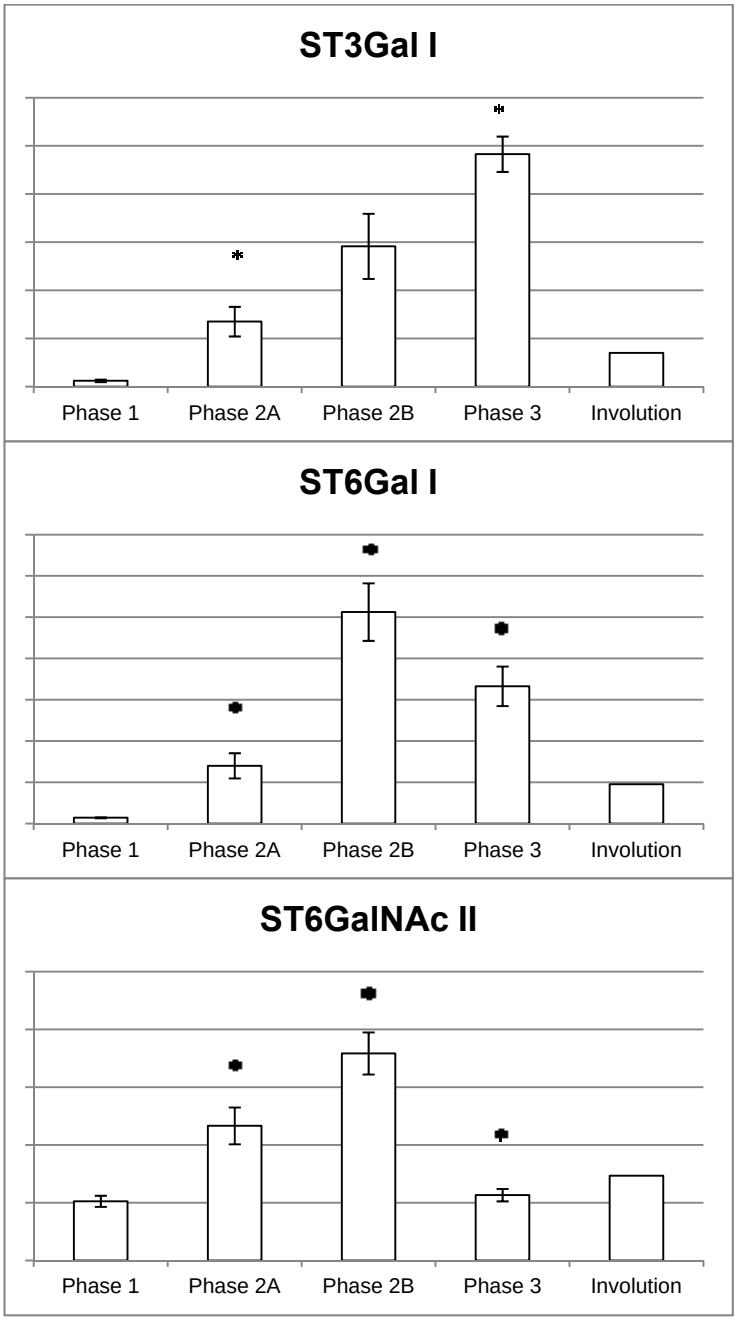


Figure 8: qPCR expression analysis of the ST6Gal I gene in wallaby mammary tissue harvested throughout the lactation cycle . Phase 1 (n=3) is represented by mammary tissue from animals at 25, 21, and 18 days *before birth*, phase 2A (n=13) is represented by tissue from animals at days 1.75, 2, 3, 5, 10, 15, 22, 37, 62, 70, 80, 100 and 110 of lactation, phase 2B (n=5) is represented by tissue from animals at days 133, 150, 171, 180 and 193 of lactation, whilst phase 3 (n=3) encompassed tissue from days 216, 220 and 260 of lactation. The involution sample was taken 5 days post weaning (n=1). As each sample was obtained from a different animal, samples within each phase were treated as biological replicates. The mean calibrated normalised relative expression level of the ST6Gal I gene at each phase of lactation is shown including the standard error of the mean. A statistically significant difference in expression between each phase and the preceding phase is denoted by a "*".

5.2.5 Discussion

We have obtained near full-length cDNA sequences for several tammar wallaby sialic acid biosynthesis and sialyltransferase genes and determined the phylogenetic relationship of the translated tammar wallaby protein sequences to those of other species. Furthermore, our study has demonstrated that these genes are differentially expressed in the mammary gland during the ~300 days of tammar wallaby lactation. All of the genes examined were either not expressed or only expressed at minimal levels during pregnancy and subsequently peaked in different phases of lactation, suggesting that the sialylglycoconjugate products of the sialyltransferase enzymes are pertinent to specific developmental stages of either the mammary gland or the suckled young.

Overall, our phylogenetic analysis of the tammar wallaby sialic acid biosynthesis and sialyltransferase enzyme amino acid sequences are in accord with the postulated evolutionary divergence of the various vertebrate taxa (Springer, M.S., Stanhope, M.J., et al. 2004). The amino acid sequence analysis of the sialyltransferase genes identified the 3 sialylmotifs (Datta, A.K. and Paulson, J.C. 1997, Geremia, R.A., HarduinLepers, A., et al. 1997), which were highly conserved across all species. However, a significant component of the proteins was not well conserved, which is most probably due to the highly variable nature of the sialyltransferase stem region (Paulson, J.C. and Colley, K.J. 1989, Takashima, S. 2008).

The GNE and NANS genes code for 2 of the 3 enzymes known to be uniquely involved in sialic acid biosynthesis (Tanner, M.E. 2005). GNE, the bifunctional enzyme UDP-GlcNAc 2-epimerase/ManAc kinase, initially catalyses the epimerisation of UDP-GlcNAc to ManNAc and then phosphorylates ManNAc to ManNAc-6-P. The epimerisation step is the first committed, and rate limiting, step in sialic acid biosynthesis (Gal, B., Ruano, M.J., et al. 1997). Subsequently, NANS, N-acetylneuraminic acid synthase is responsible for converting MaNAc-6-P to NeuNAc-9-P. Given their crucial roles in a ubiquitous

biological process, it was unsurprising that we found a high level of conservation between the GNE and NANS amino acid sequences of the tammar wallaby and those of other species. In the lactating tammar wallaby mammary gland, the GNE and NANS genes exhibited similar patterns of expression. They both showed only minimal expression in phase 1, pregnancy, but both significantly increased in expression in phases 2A and 2B, with peak expression occurring in phase 3. Given its essential role in sialic acid biosynthesis, GNE is an important regulator of sialylation. It was shown by Keppler et al. (1999) that GNE activity can be modulated at the transcriptional level and that this influenced downstream sialylation in the cell types investigated.

The pattern of GNE expression we observed suggests that sialic acid biosynthesis is up-regulated in the lactating tammar wallaby mammary gland throughout phases 2A and 2B, eventually peaking in phase 3. This is consistent with Messer and Green's (1979) published data on the monosaccharide content of tammar wallaby milk, which shows an increase in sialic acid content, as a percentage of total carbohydrate, in phase 3 of lactation. Collectively, this data suggests that the tammar young has an increasing demand for sialic acid throughout lactation. However, given that the profile of carbohydrate structures changes dramatically throughout tammar lactation (Messer, M. and Green, B. 1979), the form in which the sialic acid is delivered in the milk is likely to be quite disparate between the different lactation phases.

Lactose is the major carbohydrate during the first 5 days of tammar wallaby lactation, but is replaced by complex oligosaccharides by as early as the 6^{th} day of lactation. These oligosaccharaides remain dominant for the remainder of phase 2 of lactation (Messer, M. and Green, B. 1979, Messer, M. and Nicholas, K.R. 1991). This suggests that for the first week of lactation, sialic acid may be present either in its free form or as sialyllactose, while thereafter it is most likely a component of larger oligosaccharide structures. During the transition to phase 3, complex oligosaccharides are gradually lost (Messer, M. and Mossop,

G. 1977), leading to the secretion of solely monosaccharides, accompanied by an overall decrease in carbohydrate concentration (Messer, M. and Elliott, C. 1987, Messer, M. and Nicholas, K.R. 1991). The lack of oligosaccharide structures in phase 3 implies that sialic acid is only present in its free form at this stage, which is unusual, as phase 3 is generally thought to approximate the onset of eutherian lactation (Tyndale-Biscoe, C.H. and Janssens, P.A. 1988) and free sialic acid is not a significant component in the milk of other species (Wang, B. and Brand-Miller, J. 2003). Given that for the majority of tammar lactation, sialic acid is likely to be present as part of oligosaccharides, the tammar wallaby appears to be a unique model to investigate the role of sialyloligosaccharides in development.

As with other mammalian species, the profile of sialylated structures present in wallaby milk is undoubtedly determined by the sialyltransferases expressed in the mammary gland throughout lactation (Svensson, E.C., Soreghan, B., et al. 1990, Takashima, S., Kono, M., et al. 2000, Wang, X., O'Hanlon, T.P., et al. 1990, Wen, D.X., Svensson, E.C., et al. 1992). The ST3Gal I sialyltransferase gene showed a similar pattern of expression to the GNE and NANS genes, peaking in phase 3, in contrast to the ST6Gal I and ST6GalNAc II sialyltransferase genes which were both maximally expressed in phase 2B and then down-regulated in phase 3. The α2,3 and α2,6-linked, are the only varieties of sialylated oligosaccharides to have been described in tammar wallaby milk (Urashima, T., Saito, T., et al. 1994), although the temporal profile of their abundance has not been reported. ST3Gal I preferentially transfers sialic acid to Galβ1,3GalNAc terminal disaccharides of O-linked carbohydtrates on glycoproteins and glycolipids, via an α2,3 linkage (Takashima, S. 2008). Given that ST3Gal I gene expression peaks in phase 3 when free oligosaccharides decrease in abundance until all milk carbohydrate is present as monosaccharides (Messer, M. and Green, B. 1979), and that ST3Gal I preferentially sialylates glycoproteins and glycolipids, it is unlikely that it is responsible for the synthesis of α2,3-sialylated oligosaccharides present in tammar milk. Furthermore, the protein and

lipid content of tammar wallaby milk markedly increases during the same period (Green, B. and Merchant, J.C. 1988), adding weight to the argument that ST3Gal I may be involved in the sialylation of glycoproteins and/or glycolipids in milk.

ST6Gal I, however, has a broad acceptor specificity and can catalyse the sialylation of type 2 (Galβ1,4GlcNAc) free disaccharides, via an α2,6 linkage (Harduin-Lepers, A., Vallejo-Ruiz, V., et al. 2001). Considering that sialyltransferases are primarily regulated at the level of transcription (Svensson, E.C., Soreghan, B., et al. 1990, Takashima, S., Kono, M., et al. 2000, Wang, X., O'Hanlon, T.P., et al. 1990, Wen, D.X., Svensson, E.C., et al. 1992), we suggest that based on the gene expression data reported in the current study, the presence of α2,6-sialylated oligosaccharides in tammar wallaby milk is particularly important in phase 2B of lactation, with an additional requirement for α2,3-sialylated glycoproteins and/or glycolipids in phase 3.

Walcott and Messer (1980) profiled the intestinal sialidase activity of the tammar wallaby pouch young throughout the lactation cycle. They found that, following an initial increase in activity around day 70, the greatest level of intestinal sialidase activity occurred between 140 and 280 days after birth, which encompasses phases 2B and 3 of tammar wallaby lactation (Tyndale-Biscoe, C.H. and Janssens, P.A. 1988). This correlates with our finding that shows sialyltransferase gene expression peaking in phases 2B and 3, and also with the increase in sialic acid monosaccharide content observed in tammar wallaby milk towards the end of phase 2B and in phase 3 by Messer and Green (1979). Taken together, these data strongly imply that sialic acid is cleaved from milk oligosaccharides in the gut of the tammar neonate, thereby providing an exogenous supply of sialic acid for the synthesis of sialylglycoconjugates.

Interestingly, circulating prolactin levels also dramatically increase in the tammar wallaby towards the end of phase 2B and into phase 3 (Hinds, L.A. 1988). This is consistent with our results in mouse explants which showed that expression of GNE, NANS and ST3Gal

and ST6Gal sialyltransferase genes is regulated by lactogenic hormones (Maksimovic, manuscript in preparation), suggesting that the orthologous tammar sialic acid biosynthesis and sialyltransferase genes may also be regulated by prolactin and other lactogenic hormones.

Phase 2 of tammar wallaby lactation has no equivalent in eutherian mammals whereas phase 3 approximates eutherian lactogenesis stage 2 when copious milk secretion is initiated (Tyndale-Biscoe, C.H. and Janssens, P.A. 1988). The transition between rapid and mature brain growth phases of the tammar wallaby pouch young occurs in the period between 120 and 180 days post partum (Renfree, M.B., Holt, A.B., et al. 1982) in phase 2B of lactation. The eyes of the pouch young open at approximately day 140 and it subsequently begins to thermoregulate, stand unaided and makes its first excursion from the pouch at around 200 days after birth (Tyndale-Biscoe, C.H. and Janssens, P.A. 1988). Brain development of the mouse neonate, at approximately 10 days after birth when the opening of the eyes occurs, is at a similar stage to that of the tammar wallaby at approximately 140-160 days (Darlington, R., Dunlop, S., et al. 1999). Given the speculation regarding the role of exogenous sialic acid in the nervous system of the suckling mammalian neonate (Nakano, T., Sugawara, M., et al. 2001, Nohle, U. and Schauer, R. 1984, Wang, B. and Brand-Miller, J. 2003, Wang, B., Yu, B., et al. 2007), it is intriguing that the milk of both rats and the tammar wallaby contains high concentrations of sialylated oligosaccharides when they are both at similar stages of neural development. It is also interesting to note that the human foetus opens its eyes at approximately 6 months, whilst it is still *in utero* (Darlington, R., Dunlop, S., et al. 1999), suggesting that the functions of milk sialyloligosaccharides in relatively altricial species such as mice and wallabies, as compared to humans, may be considerably different.

The current study suggests that the sialylation of milk products in the tammar wallaby mammary gland varies throughout pregnancy, lactation and involution and similar to other

species sialylation is at least partially regulated at the level of transcription of sialyltransferases and key genes involved in sialic acid biosynthesis. The present findings correlate well with previous studies on the sialic acid content of tammar milk and the sialylated structures, as well as with reported activity levels of intestinal sialidase of the neonate (Dickson, J.J. and Messer, M. 1978, Walcott, P.J. and Messer, M. 1980). Taken collectively, a comparative approach indicates that sialylated milk products may be particularly important near the time of eye opening in the mouse and wallaby coordinated systems appear to exist in the mammary gland and neonatal gastrointestinal tract to ensure the delivery and utilisation of sialic acid from milk oligosaccharides.

Animal models with divergent adaptations to lactation have alternative reproductive strategies to commonly used laboratory and livestock models and can facilitate an improved understanding of the lactation cycle, and the relationship between milk products and the development of the young. The extremely altricial nature of the tammar wallaby young at birth, and subsequent long lactation, provide a unique opportunity for the correlation of the expression of sialic acid biosynthesis and sialyltransferase genes in the mammary gland with qualitative and quantitative assessments of milk sialyloligosaccharides and developmental state of the neonate. We propose that, given its accessibility to the researcher throughout key early developmental events, the tammar wallaby is an excellent model for the study of sialic acid biosynthesis and sialylation in the mammary gland and related processes in the neonate.

5.2.6 References

Bird PH, Hendry KA, Shaw DC, Wilde CJ, Nicholas KR. 1994. Progressive changes in milk protein gene expression and prolactin binding during lactation in the tammar wallaby (Macropus eugenii). *J Mol Endocrinol*, 13:117-125.

Brunngraber EG, Witting LA, Haberland C, Brown B. 1972. Glycoproteins in Tay-sachs disease: isolation and carbohydrate composition of glycopeptides. *Brain research*, 38:151-162.

Collet C, Joseph R, Nicholas K. 1991. A marsupial beta-lactoglobulin gene: characterization and prolactin-dependent expression. *J Mol Endocrinol*, 6:9-16.

Coppa GV, Gabrielli O, Giorgi P, Catassi C, Montanari MP, et al. 1990. Preliminary study of breastfeeding and bacterial adhesion to uroepithelial cells. *The Lancet; London*, 335:569.

Dall'Olio F. 2000. The sialyl-α2,6-lactosaminyl-structure: Biosynthesis and functional role. *Glycoconj. J.*, 17:669-676.

Darlington R, Dunlop S, Finlay B. 1999. Neural development in metatherian and eutherian mammals: Variation and constraint. *The Journal of Comparative Neurology*, 411:359-368.

Datta AK, Paulson JC. 1997. Sialylmotifs of sialyltransferases. *Indian J Biochem Biophys*, 34:157-165.

Dickson JJ, Messer M. 1978. Intestinal neuraminidase activity of suckling rats and other mammals - relationship to sialic-acid content of milk. *Biochem. J.*, 170:407-413.

Dove H, Cork SJ. 1989. Lactation in the tammar wallaby (Macropus eugenii). I. Milk consumption and the algebraic description of the lactation curve. *Journal of Zoology*, 219:385-397.

Drummond A, Ashton B, Buxton S, Cheung M, Cooper A, Heled J, Kearse M, Moir R, Stones-Havas S, Sturrock S, *et al.* 2010. Geneious v5.1. *Available from* http://www.geneious.com.

Duncan PI, Raymond F, Fuerholz A, Sprenger N. 2009. Sialic Acid Utilisation and Synthesis in the Neonatal Rat Revisited. *PLoS One*, 4:e8241.

Findlay L. 1982. The mammary glands of the tammar wallaby (Macropus eugenii) during pregnancy and lactation. *J Reprod Fertil*, 65:59-66.

Gal B, Ruano MJ, Puente R, GarciaPardo LA, Rueda R, Gil A, Hueso P. 1997. Developmental changes in UDP-N-acetylglucosamine 2-epimerase activity of rat and guinea-pig liver. *Comp. Biochem. Phys. B*, 118:13-15.

Geremia RA, HarduinLepers A, Delannoy P. 1997. Identification of two novel conserved amino acid residues in eukaryotic sialyltransferases: Implications for their mechanism of action. *Glycobiology*, 7:R5-R7.

Green B, Merchant J, Newgrain K. 1988. Milk Consumption and Energetics of Growth in Pouch Young of the Tammar Wallaby, Macropus-Eugenii. *Australian Journal of Zoology*, 36:217-227.

Green B, Merchant JC. 1988. The Composition of Marsupial Milk. In: Tyndale-Biscoe CH, Janssens PA editors. The Developing Marsupial. Berlin Heidelberg: Springer-Verlag.

Hamosh M. 1999. Breastfeeding: unravelling the mysteries of mother's milk. *Medscape General Medicine*, 1.

Hanasaki K, Powell LD, Varki A. 1995. Binding of human plasma sialoglycoproteins by the B cell-specific lectin CD22. Selective recognition of immunoglobulin M and haptoglobin. *J. Biol. Chem.*, 270:7543-7550.

Harduin-Lepers A, Mollicone R, Delannoy P, Oriol R. 2005. The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. *Glycobiology*, 15:805-817.

Harduin-Lepers A, Recchi MA, Delannoy P. 1995. 1994, the year of sialyltransferases. *Glycobiology*, 5:741-758.

Harduin-Lepers A, Vallejo-Ruiz V, Krzewinski-Recchi MA, Samyn-Petit B, Julien S, Delannoy P. 2001. The human sialyltransferase family. *Biochimie*, 83:727-737.

Hinds LA. 1988. Hormonal control of lactation. In: Tyndale-Biscoe CH, Janssens PA editors. The Developing Marsupial. Berlin, Heidelberg: Springer-Verlag.

Jeanneau C, Chazalet V, Auge C, Soumpasis DM, Harduin-Lepers A, Delannoy P, Imberty A, Breton C. 2004. Structure-function analysis of the human sialyltransferase ST3Gal I - Role of N-glycosylation and a novel conserved sialylmotif. *Journal of Biological Chemistry*, 279:13461-13468.

Keppler OT, Hinderlich S, Langner J, Schwartz-Albiez R, Reutter W, Pawlita M. 1999. UDP-GlcNAc 2-Epimerase: A Regulator of Cell Surface Sialylation. *Science*, 284:1372-1376.

Kono M, Ohyama Y, Lee Y-C, Hamamoto T, Kojima N, Tsuji S. 1997. Mouse β-galactoside α2,3-sialyltransferases: comparison of in vitro substrate specificities and tissue specific expression. *Glycobiology*, 7:469-479.

Kono M, Tsuda T, Ogata S, Takashima S, Liu H, Hamamoto T, Itzkowitz SH, Nishimura S, Tsuji S. 2000. Redefined Substrate Specificity of ST6GalNAc II: A Second Candidate Sialyl-Tn Synthase. *Biochemical and Biophysical Research Communications*, 272:94-97.

Kunz C, Rudloff S, Baier W, Klein N, Strobel S. 2000. Oligosaccharides in human milk: structural, functional, and metabolic aspects. *Annu Rev Nutr*, 20:699-722.

LoCascio RG, Ninonuevo MR, Freeman SL, Sela DA, Grimm R, Lebrilla CB, Mills DA, German JB. 2007. Glycoprofiling of bifidobacterial consumption of human milk oligosaccharides demonstrates strain specific, preferential consumption of small chain glycans secreted in early human lactation. *J Agric Food Chem*, 55:8914-8919.

Martin-Sosa S, Martin MJ, Garcia-Pardo LA, Hueso P. 2003. Sialyloligosaccharides in human and bovine milk and in infant formulas: variations with the progression of lactation. *J. Dairy Sci.*, 86:52-59.

McVeagh P, Brand-Miller J. 1997. Human milk oligosaccharides: only the breast. *J. Paediatr. Child Health*, 33:281-286.

Messer M, Elliott C. 1987. Changes in alpha-lactalbumin, total lactose, UDP-galactose hydrolase and other factors in tammar wallaby (Macropus eugenii) milk during lactation. *Aust J Biol Sci*, 40:37-46.

Messer M, Green B. 1979. Milk carbohydrates of marsupials. II. Quantitative and qualitative changes in milk carbohydrates during lactation in the tammar wallaby (Macropus eugenii). *Aust J Biol Sci*, 32:519-531.

Messer M, Mossop G. 1977. Milk Carbohydrates of Marsupials I. Partial Separation and Characterization of Neutral Milk Oligosaccharides of the Eastern Grey Kangaroo. *Australian Journal of Biological Sciences*, 30:379-388.

Messer M, Nicholas KR. 1991. Biosynthesis of marsupial milk oligosaccharides: characterization and developmental changes of two galactosyltransferases in lactating mammary glands of the tammar wallaby, Macropus eugenii. *Biochim Biophys Acta*, 1077:79-85.

Messer M, Urashima T. 2002. Evolution of Milk Oligosaccharides and Lactose. *Trends in Glycoscience and Glycotechnology*, 14:153-176.

Moughan PJ, Birtles MJ, Cranwell PD, Smith WC, Pedraza M. 1992. The piglet as a model animal for studying aspects of digestion and absorption in milk-fed human infants. *World Rev Nutr Diet*, 67:40-113.

Mouricout M, Petit JM, Carias JR, Julien R. 1990. Glycoprotein glycans that inhibit adhesion of Escherichia coli mediated by K99 fimbriae: treatment of experimental colibacillosis. *Infect. Immun.*, 58:98-106.

Nakano T, Sugawara M, Kawakami H. 2001. Sialic acid in human milk: composition and functions. *Acta Paediatr. Taiwan*, 42:11-17.

Nicholas K. 1988. Asynchronous dual lactation in a marsupial, the tammar wallaby (Macropus eugenii). *Biochem Biophys Res Commun*, 154:529–536.

Nicholas K, Simpson K, Wilson M, Trott J, Shaw D. 1997. The tammar wallaby: a model to study putative autocrine-induced changes in milk composition. *J Mammary Gland Biol Neoplasia*, 2:299-310.

Ninonuevo MR, Park Y, Yin H, Zhang J, Ward RE, Clowers BH, German JB, Freeman SL, Killeen K, Grimm R, *et al.* 2006. A strategy for annotating the human milk glycome. *J Agric Food Chem*, 54:7471-7480.

Nohle U, Schauer R. 1984. Metabolism of sialic acids from exogenously administered sialyllactose and mucin in mouse and rat. *Hoppe Seylers Z Physiol Chem*, 365:1457-1467.

Oftedal OT, Alt GL, Widdowson EM, Jakubasz MR. 1993. Nutrition and growth of suckling black bears (Ursus americanus) during their mothers' winter fast. *Br J Nutr*, 70:59-79.

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

Paulson JC, Rearick JI, Hill RL. 1977. Enzymatic properties of beta-d-galactoside alpha-2-6 sialyltransferase from bovine colostrum. *Journal of Biological Chemistry*, 252:2363-2371.

Picciano MF. 1998. Human milk: nutritional aspects of a dynamic food. *Biology of the Neonate*, 74:84-93.

Poole WE. 1991. Tables for age determination of the Kangaroo Island wallaby (tammar), Macropus eugenii, from body measurements CSIRO, Division of Wildlife and Ecology:Lyneham, A.C.T.

Renfree MB. 1983. In: Finn C editor. Oxford Reviews of Reproductive Biology. England: Oxford Univ. Press. p. 1–26.

Renfree MB, Holt AB, Green SW, Carr JP, Cheek DB. 1982. Ontogeny of the brain in a marsupial (Macropus eugenii) throughout pouch life. I. Brain growth. *Brain Behav Evol*, 20:57-71.

Schauer R. 1982. Sialic Acids, Chemistry, Metabolism, and Function Springer-Verlag:Wien, New York.

Schwarzkopf M, Knobeloch K-P, Rohde E, Hinderlich S, Wiechens N, Lucka L, Horak I, Reutter W, Horstkorte R. 2002. Sialylation is essential for early development in mice. *Proceedings of The National Academy of Sciences of The United States of America*, 99:5267-5270.

Sela DA, Li Y, Lerno L, Wu S, Marcobal AM, German JB, Chen X, Lebrilla CB, Mills DA. 2011. An infant-associated bacterial commensal utilizes breast milk sialyloligosaccharides. *Journal of Biological Chemistry*.

Sherblom AP, Bourassa CR. 1983. Specificity of submaxillary gland sialyltransferases. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 761:94-102.

Springer MS, Stanhope MJ, Madsen O, de Jong WW. 2004. Molecules consolidate the placental mammal tree. *Trends Ecol Evol*, 19:430-438.

Svensson EC, Soreghan B, Paulson JC. 1990. Organization of the β-galactoside α2,6-sialyltransferase gene. Evidence for the transcriptional regulation of terminal glycosylation. *J. Biol. Chem.*, 265:20863-20868.

Takashima S. 2008. Characterization of Mouse Sialyltransferase Genes: Their Evolution and Diversity. *Bioscience, biotechnology, and biochemistry*, 72:1155-1167.

Takashima S, Kono M, Kurosawa N, Yoshida Y, Tachida Y, Inoue M, Kanematsu T, Tsuji S. 2000. Genomic organization and transcriptional regulation of the mouse GD3 synthase gene (ST8Sia I): Comparison of genomic organization of the mouse sialyltransferase genes. *Journal of Biochemistry*, 128:1033-1043.

Tanner ME. 2005. The enzymes of sialic acid biosynthesis. *Bioorganic Chemistry*, 33:216-228.

Trott JF, Hearn CM, Renfree MB, Shaw G, Simpson KJ, Moyle RLC, Nicholas KR. 2003. Maternal regulation of milk composition, milk production, and pouch young development during lactation in the tammar wallaby (Macropus eugenii). *Biology of Reproduction*, 68:929-936.

Trott JF, Wilson MJ, Hovey RC, Shaw DC, Nicholas KR. 2002. Expression of novel lipocalin-like milk protein gene is developmentally-regulated during lactation in the tammar wallaby, Macropus eugenii. *Gene*, 283:287-297.

Tsuji S. 1996. Molecular Cloning and Functional Analysis of Sialyltransferases. *J Biochem (Tokyo)*, 120:1-13.

Tyndale-Biscoe CH, Janssens PA. 1988. Introduction. In: Tyndale-Biscoe CH, Janssens PA editors. The Developing Marsupial. Berlin, Heidelberg: Springer-Verlag.

Urashima T, Saito T, Nakamura T, Messer M. 2001. Oligosaccharides of milk and colostrum in non-human mammals. *Glycoconjugate Journal*, 18:357-371.

Urashima T, Saito T, Tsuji Y, Taneda Y, Takasawa T, Messer M. 1994. Chemical characterization of sialyl oligosaccharides isolated from tammar wallaby (Macropus eugenii) milk. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1200:64-72.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, 3:research0034.0031 - research0034.0011.

Varki A. 1993. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*, 3:97-130.

Varki A. 2007. Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. *Nature*, 446:1023-1029.

Varki A, Schauer R. 2009. Sialic Acids. In: Varki A, Cummings R, Esko J, Freeze H, Stanley P, Bertozzi C, Hart G, Etzler M editors. Essentials of Glycobiology. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

Von Itzstein M, Thomson R. 1997. Sialic acids and sialic acid recognising proteins: drug discovery targets and potential glycopharmaceuticals. *Curr Med Chem*, 4:185–210.

Walcott PJ, Messer M. 1980. Intestinal lactase (beta-galactosidase) and other glycosidase activities in suckling and adult tammar wallabies (Macropus eugenii). *Aust J Biol Sci*, 33:521-530.

Wang B, Brand-Miller J. 2003. The role and potential of sialic acid in human nutrition. *Eur. J. Clin. Nutr.*, 57:1351-1369.

Wang B, Brand-Miller J, McVeagh P, Petocz P. 2001. Concentration and distribution of sialic acid in human milk and infant formulas. *Am. J. Clin. Nutr.*, 74:510-515.

Wang B, Yu B, Karim M, Hu H, Sun Y, McGreevy P, Petocz P, Held S, Brand-Miller J. 2007. Dietary sialic acid supplementation improves learning and memory in piglets. *Am J Clin Nutr*, 85:561-569.

Wang X, O'Hanlon TP, Young RF, Lau JTY. 1990. Rat β-galactoside α2,6-sialyltransferase genomic organization: alternate promoters direct the synthesis of liver and kidney transcripts. *Glycobiology*, 1:25-31.

Wen DX, Svensson EC, Paulson JC. 1992. Tissue-specific alternative splicing of the β-galactoside α2,6- sialyltransferase gene. *J. Biol. Chem.*, 267:2512-2518.

6 Discussion and future directions

Milk sialylglycoconjugates are thought to exert a protective affect on the gastrointestinal tract of the suckling neonate by competitively binding to invading pathogens and/or promoting the growth of beneficial flora (Coppa, G.V., Gabrielli, O., et al. 1990, Martin-Sosa, S., Martin, M.J., et al. 2003, McVeagh, P. and Brand-Miller, J. 1997, Mouricout, M., Petit, J.M., et al. 1990, Wang, B., Brand-Miller, J., et al. 2001). Sialylated milk oligosaccharides can prevent the attachment of *E.coli* strains associated with neonatal meningitis and sepsis; it has been shown that these organisms are less common in the faeces of breastfed infants than formula fed infants (McVeagh, P. and Brand-Miller, J. 1997, Nakano, T., Sugawara, M., et al. 2001). It has also been shown that milk oligosaccharides encourage the colonisation and proliferation of particular groups of bacteria by acting as selective nutrients (Gibson, G.R. and Roberfroid, M.B. 1995, LoCascio, R.G., Ninonuevo, M.R., et al. 2007, Marcobal, A., Barboza, M., et al. 2010); Sela et al. (2011) demonstrated that certain bifidobacteria can metabolise sialylated oligosaccharides, possibly contributing to their ability to establish in the infant gut. Furthermore, in addition to playing a crucial role in the protection and development of the neonatal gastrointestinal tract, sialylated milk oligosaccharides have also been implicated in having a role in post-natal brain development (Wang, B. and Brand-Miller, J. 2003) as demonstrated in 2 studies which showed that dietary sialic acid may enhance learning in developing piglets (Moughan, P.J., Birtles, M.J., et al. 1992, Wang, B., Yu, B., et al. 2007). Given the important roles of sialylated milk constituents this thesis aimed to investigate the expression of sialic acid biosynthesis and sialyltransferase genes in the mammary gland and the role that lactogenic hormones play in ultimately determining the sialylated milk components produced during lactation.

6.1 Transcriptional control of sialylation in the mammary gland

This study is the first comprehensive investigation of sialic acid biosynthesis and sialyltransferase gene expression and regulation in the mammary gland during pregnancy, lactation and involution. It has described the differential expression of key sialic acid biosynthesis genes and certain sialyltransferase genes in the lactating eutherian and marsupial mammary gland and demonstrated that lactogenic hormones are likely to be important in their regulation. By comparing temporal gene expression profiles between species, it has also examined the developmental context in which milk sialylglycoconjugates may be important. Most significantly, however, this study has highlighted that transcriptional control is likely to be a major determinant of the ultimate profile of sialylated structures produced by the mammary gland irrespective of eventual species differences in milk composition.

The oligosaccharide profile of the milk of various species has been extensively studied (Asakuma, S., Akahori, M., et al. 2007, Carlson, S.E. 1985, Crisp, E.A., Cowan, P.E., et al. 1989, Messer, M., FitzGerald, P.A., et al. 1987, Messer, M. and Green, B. 1979, Messer, M. and Mossop, G. 1977, Messer, M. and Nicholas, K.R. 1991, Messer, M. and Trifonoff, E. 1982, Messer, M., Trifonoff, E., et al. 1980, Messer, M. and Urashima, T. 2002, Tao, N., DePeters, E., et al. 2009, Tao, N., Wu, S., et al. 2011, Urashima, T., Kusaka, Y., et al. 1997, Urashima, T., Sumiyoshi, W., et al. 1999, Urashima, T., Yamashita, T., et al. 2000, Wu, S., Grimm, R., et al. 2010). Despite the ultimate end product being milk, there are significant differences between placental mammals, monotremes and marsupials in the quantity and quality of their milk carbohydrates (Messer, M. and Urashima, T. 2002, Urashima, T., Saito, T., et al. 2001). Generally, for the majority of placental mammals, free lactose is the dominant milk carbohydrate (Messer, M. and Urashima, T. 2002). Conversely, the milk of monotremes and marsupials contains virtually no lactose, with the exception of the first few days of lactation (Messer, M. and Green, B. 1979, Messer, M. and Urashima, T. 2002). Monotreme milk is rich in fucose, and lactose is primarily found

as tri- and tetrasaccharides (Messer, M. 1974, Messer, M. and Kerry, K.R. 1973); whilst marsupial milk is dominated by a series of galactosyl saccharides and their derivatives that vary in magnitude from tri- to octasaccharides (Bradbury, J.H., Collins, J.G., et al. 1983, Collins, J.G., Bradbury, J.H., et al. 1981, Crisp, E.A., Cowan, P.E., et al. 1989, Messer, M., Crisp, E.A., et al. 1989, Messer, M., FitzGerald, P.A., et al. 1987, Messer, M. and Green, B. 1979, Messer, M. and Mossop, G. 1977, Messer, M. and Trifonoff, E. 1982, Messer, M., Trifonoff, E., et al. 1980, Urashima, T., Saito, T., et al. 1994). Interestingly, most monotreme milk oligosaccharides are also found in the milk of eutherian mammals but not in the milk of marsupials, suggesting that, at least with respect to milk carbohydrates, eutherian mammals may be more closely related to monotremes than to marsupials (Urashima, T., Saito, T., et al. 2001). However, despite the knowledge of the oligosaccharide composition of the milk of various species, we are still far from determining the exact role sialic acid plays in human infant nutrition and in the nutrition of other mammalian neonates. This stems from the inherent differences in the oligosaccharide composition of even closely evolutionarily related mammals, let alone divergent species, calling into question our ability to translate the findings from one species to another. Tao et al. (2011) have recently demonstrated that the oligosaccharide composition of primate milk is generally more complex and exhibits greater diversity than that of non-primate mammals with a moderate level of sialylation in the milk of all species studied, highlighting its fundamental importance to the overall function of milk oligosaccharides and the selective advantages they may afford. However, primate milk oligosaccharide composition did not necessarily cluster according to primate phylogeny, suggesting that other selective pressures were responsible for determining the quantitative and qualitative nature of primate milk oligosaccharides. The authors proposed that the milk oligosaccharides of various primates may in fact be responsible for specifically guiding the prevalent microbe regimes that are associated with particular social group sizes and states.

However, despite these vast differences in milk oligosaccharide composition between species, the common element is that the young of all mammals are suckled on milk which is produced from mammary glands. At a molecular level, the mammary gland secretes a milk containing a combination of lipids, proteins and carbohydrates with a specific nutritional and developmental purpose. Transcriptional suppression and activation has been shown to play an integral role in the transition of the mammary gland from pregnancy to lactation and involution (Anderson, S.M., Rudolph, M.C., et al. 2007, Clarkson, R., Wayland, M., et al. 2004, Clarkson, R.W.E. and Watson, C.J. 2003, Lemay, D., Neville, M., et al. 2007, Rudolph, M.C., McManaman, J.L., et al. 2003, Stein, T., Morris, J., et al. 2004, Stein, T., Salomonis, N., et al. 2007), therefore, it is reasonable to assume that the eventual oligosaccharide composition of the milk of any species is determined by the molecular processes governing mammary gland development and lactation. Prieto et al. (1995) have already demonstrated that transgenic mice expressing human α1,2-fucosyltransferase, under the control of a lactogenic promoter, can produce large amounts of 2'-fucosyllactose and modified glycoproteins containing the H-antigen. The significant amounts of 2'-fucosyllactose produced by these transgenic mice suggest that the Golgi apparatus of the lactating mammary gland is sufficiently adaptable with regard to factors such as sugar nucleotide availability, lactose availability, enzyme localization and enzyme secretion. If the mammary gland is similarly adaptable with respect to sialic acid biosynthesis and sialyltransferase substrates, then transcriptional regulation of sialic acid biosynthesis and sialyltransferase genes in the mammary gland is potentially the key factor in determining the nature and extent of milk oligosaccharide sialylation.

Given the evidence from previous studies and the results presented in this thesis demonstrating the importance of transcriptional regulation of sialylation in the lactating mammary gland, it is possible that the level of expression of sialic acid biosynthesis and sialyltransferase genes at various stages of lactation is actually an important marker for the relative importance of sialic acid at any particular stage in the development of the neonate.

By focusing on the temporal expression of key genes involved in determining the nature and extent of milk sialylation, the confounding effect of milk oligosaccharide differences between species could be effectively removed, as the production of all milk sialyloligosaccharides is governed by the same suite of genes, irrespective of species. This study has demonstrated that sialic acid biosynthesis and sialyltransferase gene expression is highest in the mouse and tammar wallaby at around the time of eye opening, suggesting that sialic acid may be particularly important during that time in development (Chapter 5).

In order to determine whether the lactation-specific expression of the ST6Gal I gene observed by Dalziel et al. (2001) occurred in other mammalian species, the current study examined the expression of the ST6Gal I gene in the bovine mammary gland during pregnancy, lactation and involution (Chapter 1). In the bovine mammary gland, ST6Gal I gene expression increased only slightly between pregnancy and lactation but significantly increased between lactation and involution and no evidence was found in either the cow or the human for the presence of the lactation-specific ST6Gal I exon (L) that was previously identified in the mouse (Dalziel, M., Huang, R.Y., et al. 2001). A striking difference was observed in the expression pattern of the ST6Gal I gene in the mouse and tammar wallaby, where expression increased between pregnancy and lactation and was rapidly down-regulated in involution, This is an intriguing and vital difference that has implications regarding the nature of the sialylated components observed in the milk of these species. It also highlights the importance of investigating expression of sialic acid biosynthesis and sialyltransferase genes across a variety of species as the observations from a single model organism may not be very widely applicable. In future studies, the same panel of bovine mammary gland biopsy samples could be utilised for a more extensive analysis of sialic acid biosynthesis and sialyltransferase gene expression in the bovine mammary gland throughout the lactation cycle. Ideally, an expanded time series would provide a more comprehensive overview of the transcriptional landscape in the bovine mammary gland and its effects on sialylation, however, the mammary biopsy collection process is invasive

and it would be difficult to obtain large numbers of samples from the same animal during the course of lactation as the risk of infection is high. A microarray approach could be used to perform a global gene expression analysis of bovine lactation that, as well as providing valuable expression data on sialic acid biosynthesis and sialyltransferase genes, would allow us to associate that information with expression data on genes in related pathways, providing us with further insight into how these genes may be regulated in the mammary gland. Alternatively, an RNA sequencing (RNA-Seq) approach would allow us to simultaneously capture the expression of any previously unidentified 5'UTR exons in sialyltransferase genes. If the sample panel can be expanded to include several biological replicates of mammary tissue from cows that have never been pregnant, this would provide baseline expression data for comparison, allowing us to identify sialyltransferase isoforms that are *only* expressed in the lactating bovine mammary gland. Furthermore, comparing bovine and mouse mammary gland lactation gene expression data with several other species representing the various mammalian infra-classes would be a powerful approach for the identification of any gene expression trends within and between classes. It could potentially highlight windows in neonatal development when the presence of sialylated milk oligosaccharides is likely to be crucial, as well as any evolutionarily conserved associations with other pathways, which may provide further insight into the complex regulatory mechanisms involved. A similar comparative approach was recently utilised to identify the folate receptor alpha as a putative regulator of milk protein synthesis (Menzies, K.K., Lefevre, C., et al. 2009).

The current study has demonstrated that lactogenic hormones are involved in the regulation of key sialic acid biosynthesis genes and certain sialyltransferase genes in the mouse mammary gland. *In silico* analysis has identified potential STAT5 transcription factor binding sites in several relevant genes. Potential binding sites for several other lactation-related transcription factors have also been identified, suggesting that the regulation of sialic acid biosynthesis and sialyltransferase genes by lactogenic hormones may be carried

out via the JAK/STAT pathway. Clearly defining the lactogenic promoter region would allow for the identification of the transcription factors involved in regulating sialyltransferase gene expression during lactation. A candidate lactogenic promoter region has already been identified in the mouse upstream of ST6Gal I exon (L) (Dalziel, M., Huang, R.Y., et al. 2001) and this study has demonstrated that the ST6Gal I exon (L) transcript is only expressed in the presence of lactogenic hormones (Chapter 4). Thus, the region upstream of murine ST6Gal I exon (L), is an ideal candidate for functional promoter studies that would provide us with more in depth knowledge of the key aspects of the transcriptional control of sialylation in the context of lactation. Furthermore, considering the size of sialyltransferase genes, particularly their expansive 5'UTR that are known to contain multiple promoters, and the fact that functional STAT5 binding sites can be found in introns or distal enhancers (Nelson, E.A., Walker, S.R., et al. 2004, Nelson, E.A., Walker, S.R., et al. 2006), a chromatin immunoprecipitation sequencing (ChIP-Seq) approach appears to be the best strategy for identifying all functional, lactation-related transcription factor binding sites in the ST6Gal I gene. Beginning with STAT5, which is a known regulator of milk protein gene expression (Chia, D.J., Varco-Merth, B., et al. 2010, Lechner, J., Welte, T., et al. 1997, Li, S. and Rosen, J.M. 1995, Magne, S., Caron, S., et al. 2003, Wakao, H., Schmitt-Ney, M., et al. 1992), potential regulatory regions of interest could be identified for further characterisation. Electrophoretic mobility shift assays (EMSA) could then be used to identify all protein complexes binding to the region of interest, by incubating the target sequence with a nuclear extract from mammary epithelial cells from the appropriate stage of lactation. The results of this proposed analysis of the mouse ST6Gal I gene would provide us with a detailed example of how a sialyltransferase gene is regulated in the lactating mammary gland. This would be an important point of reference for future studies in other species or of other sialyltransferase genes.

6.2 Sialic acid in neonatal development

The utilisation of sialyloligosaccharides by the neonate is an important area of research that requires further work. Various phenomena have been demonstrated in different species, although it is unclear if the findings in one species can be easily translated to another, given the vast differences in milk sialyloligosaccharide content and differences in reproductive and lactation strategies. For example, human milk oligosaccharides have been demonstrated to be important in the establishment of beneficial gastrointestinal flora (German, J.B., Freeman, S.L., et al. 2008, LoCascio, R.G., Ninonuevo, M.R., et al. 2007, Marcobal, A., Barboza, M., et al. 2010, Sela, D.A., Li, Y., et al. 2011, Zivkovic, A.M., German, J.B., et al.) as well as acting as binding "decoys" for potential gastrointestinal pathogens (Coppa, G.V., Gabrielli, O., et al. 1990, Martin-Sosa, S., Martin, M.J., et al. 2003, Wang, B., Brand-Miller, J., et al. 2001). Both are essential functions for the health and wellbeing of the human newborn outside of the womb. Marsupials, such as the tammar wallaby, are born in an extremely alrticial state (Renfree, M.B. 1983, Tyndale-Biscoe, C.H. and Janssens, P.A. 1988) and enter an essentially non-sterile environment. They are continuously attached to the teat for the first 100 days of lactation and do not begin to consume herbage until almost 200 days after birth (Tyndale-Biscoe, C.H. and Janssens, P.A. 1988). Lactose is only dominant in tammar milk for the first week of lactation, after which the milk sugars are almost exclusively comprised of higher order oligosaccharides. The oligosaccharide structures then become progressively smaller until approximately day 280, when all sugars are monosaccharides and overall carbohydrate concentration is markedly decreased (Green, B. and Merchant, J.C. 1988). Interestingly, the levels of microbial species reported in the tammar wallaby gut are significantly below the levels reported for human infants of a similar age (Lentle, R.G., Dey, D., et al. 2006).

It was proposed in chapter 2 of this study that the use of sialylated milk oligosaccharides by the neonate may be dependent on their relative developmental state at birth. It has previously been suggested that altricial neonates such as those of either marsupials or bears

have a particular requirement for milk oligosaccharides in general (Oftedal, O.T., Alt, G.L., et al. 1993). This may be because the osmotic pressure exerted by a certain mass of oligosaccharides in milk is lower than that exerted by an equal mass of lactose, thus allowing the delivery of greater amounts of energy without the milk becoming hyperosmotic. Furthermore, attachment of sialic acid residues at the terminus of oligosaccharide chains is a potentially efficient means of delivering an additional level of nutrition. Sialidase activity has been observed in the intestine of suckling tammars, suggesting that the neonate is able to *directly* utilise the sialic acid delivered via milk oligosaccharides (Dickson, J.J. and Messer, M. 1978, Walcott, P.J. and Messer, M. 1980). Walcott and Messer (1980) showed that, following an initial increase in activity around day 70, the greatest level of intestinal sialidase activity coincided with the period when virtually all tammar wallaby milk carbohydrates are present as higher oligosaccharides. Consequently, the authors suggested that the primary function of intestinal sialidase of suckling tammars may be to cleave sialic acid from milk components, thereby supplying sialic acid for the synthesis of sialylglycoconjugates by the neonate.

An earlier study by Dickson and Messer (1978) had highlighted that there was a significant positive correlation between the intestinal sialidase activity of the suckling neonate of rats and the sialic acid content of milk obtained at the corresponding stage of lactation. More recently, Duncan et al. (2009) examined the relationship between the sialic acid content of rat milk and the uptake, utilization and synthesis of sialic acid in suckling rat pups. They found that sialic acid concentration, as part of sialyllactose, was highest between 3 and 10 days postpartum, followed by a decrease until weaning. Their colon gene expression data strongly suggested that a change from sialic acid uptake and catabolism to sialic acid synthesis and utilization occurred during suckling, mirroring the change in milk sialic acid content. However, they found no significant change in brain sialic acid related gene expression profiles throughout the suckling period.

Despite this evidence that some species appear to be able to directly utilise sialic acid derived from milk oligosaccharides (Dickson, J.J. and Messer, M. 1978, Duncan, P.I., Raymond, F., et al. 2009, Walcott, P.J. and Messer, M. 1980) and suggestions that it may be being utilised for nervous system development (Nakano, T., Sugawara, M., et al. 2001, Wang, B. and Brand-Miller, J. 2003, Wang, B., Yu, B., et al. 2007), the actual role exogenous sialic acid has in that capacity is unclear. Given that the current study has identified the sialic acid biosynthesis and sialyltransferase genes which are differentially expressed in the mouse mammary gland during lactation, it would now be possible to create gene knockout mice that do not secrete sialylated oligosaccharides in their milk. Fuhrer et al. (2010) have already examined the consequences on the development of the gastrointestinal tract of pups suckled on ST6Gal I -/- or ST3Gal IV -/- knockout mice, however they did not examine the effects of consuming sialic acid depleted milk on development of the neonatal nervous system. By either removing the sialic acid biosynthetic capacity of the mammary gland through mammary-specific GNE gene knockout or knocking out the ST6Gal I, ST3Gal IV and ST3Gal I sialyltransferase genes, it should be possible to examine the effects on the nervous system development of pups suckled on the mammary gland of the gene knockout mothers. Mammary gland-specific knockout of genes has already demonstrated in numerous studies (Bry, C., Maass, K., et al. 2004, Cui, Y., Riedlinger, G., et al. 2004, Klover, P.J., Muller, W.J., et al. 2010) and would be appropriate in the proposed study given the importance of sialylation in other organs and physiological processes.

The development of the marsupial can be equated to that of eutherians with the period inside the pouch prior to the acquisition of homeothermy considered to be the equivalent of the *in utero* period in eutherians (Russell, E.M. 1982), providing us with a useful model to investigate both early development and lactation (Tyndale-Biscoe, C.H. and Janssens, P.A. 1988). This study has already demonstrated a correlation between previously published intestinal sialidase activity of the tammar wallaby young and expression of sialic acid

biosynthesis and sialyltransferase genes in the mammary gland of the mother (Chapter 5). The tammar wallaby could be a particularly useful model for investigating the role of milk oligosaccharide-derived sialic acid in the development of the brain and nervous system of the pouch young. The neonate is easily accessible for study during the majority of its development, providing us with the ability to accurately correlate developmental state with qualitative and quantitative measurements of milk sialic acid. However, firstly, a comprehensive analysis of the carbohydrate content of tammar wallaby milk across the lactation cycle is necessary, using current techniques such as those described by Tao et al. (2009) for profiling the bovine glycome during lactation.

As the milk glycomes of many eutherians have already been characterised using modern techniques (Barile, D., Marotta, M., et al. 2010, Ninonuevo, M.R., Park, Y., et al. 2006, Tao, N., DePeters, E., et al. 2008, Tao, N., DePeters, E., et al. 2009, Tao, N., Ochonicky, K.L., et al. 2010, Tao, N., Wu, S., et al. 2011, Zivkovic, A.M., German, J.B., et al.), it follows that to obtain a comprehensive understanding of the evolution and utilisation of sialyoligosaccharides by different mammalian species, the milk glycome of mammals representative of all classes should be characterised. The milk oligosaccharide content of monotremes has not been characterised using current technologies, despite the fact that they are a potentially rich source of information on the evolution of milk oligosaccharides and their role in development. Egg-laying monotremes are thought to be representative of ancient mammals (Griffiths, M. 1978). Their minute hatchlings are highly altricial and, like marsupial neonates, are totally dependent on milk as their only source of nutrition during suckling, a period which is also relatively long in comparison to their gestation and incubation (Griffiths, M. 1978). The echidna (*Tachyglossus aculeatus*) would be a particularly interesting subject in relation to the role of sialyloligosaccharides in development as sialyllactose is dominant amongst its suite of milk oligosaccharides (Messer, M. and Kerry, K.R. 1973).

6.3 Sialic acid and infant formula

Typically, approximately 10-30% of the 200 human milk oligosaccharide structures are sialylated (Tao, N., Wu, S., et al. 2011). Only 40 oligosaccharides have been identified in bovine milk and the overall concentration of oligosaccharides is significantly lower than in human milk (Tao, N., DePeters, E., et al. 2008). Approximately 70% of bovine milk oligosaccharides are sialylated (Tao, N., DePeters, E., et al. 2008) but given the limited diversity of structures and a lower overall oligosaccharide concentration than human milk, bovine milk is a poor approximation of human milk in terms of oligosaccharide complexity and diversity as well as *total* sialic acid content. This is noteworthy as human milk contains significantly higher levels of sialic acid than any type of infant formula, many of which are produced using bovine milk (Wang, B. and Brand-Miller, J. 2003).

Due to the relatively low amounts and dissimilar distribution of sialyloligosaccharides in bovine milk from which most infant formula is made, the infant formula industry has attempted to increase the sialyloligosaccharide content of its products by concentrating sialyllactose in bovine milk. This, however, has proven to be uneconomical (Murray Goulbourn, 2007, Personal Communication). Consequently, the ability to increase the levels of sialylated oligosaccharides in the bovine mammary gland is an intriguing prospect. In order to achieve this, a comprehensive understanding is required of the expression and regulation of sialic acid biosynthesis and sialyltransferase genes in the mammary gland, along with a detailed knowledge of the genes and regulatory mechanisms involved in glycan biosynthesis. This study has made initial progress in discerning the expression and regulation of sialic acid biosynthesis and sialyltransferase genes in model organisms and has highlighted that differences exist at the gene expression level between species that are likely to be causative of differences in eventual milk sialylation; although, it is clear that further work is required before any practical applications can be achieved.

Once the complex regulatory mechanisms behind the lactation-specific expression of sialic acid biosynthesis and sialyltransferase genes is better understood, methods for manipulating cows to produce more human-like milk in terms of sialyloligosaccharide content can be considered, in order to improve the quality of infant formula. Prieto et al. (1995) have shown that the mouse mammary gland is adaptable enough to express the human α1,2-fucosyltransferase transgene under the control of a lactogenic promoter. It is conceivable that, knowing the glycosyltransferase and sialyltransferase genes expressed in the human breast, transgenic cows could be engineered to express these genes in the mammary gland during lactation. Lysostaphin producing transgenic cows have already been engineered for resistance to intramammary *Staphylococcus aureus* infection (Donovan, D., Kerr, D., et al. 2005, Powell, A., Kerr, D., et al. 2007, Wall, R.J., Powell, A.M., et al. 2005). However, if the possibility of transgenic cows producing more "human-like" sialyloligosaccharides in their milk is to be considered, consideration also needs to be given to the oligosaccharide structures produced in the cows versus those produced in humans and how these might impinge on the utilisation of sialyloligosaccharides by the human infant.

6.4 Methodological considerations

Although the current study has made initial progress in discerning the expression and regulation of sialic acid biosynthesis and sialyltransferase genes in the mouse and tammar wallaby and has highlighted that there is a difference in the expression of the ST6Gal I gene between these species and the cow, some of the approaches utilised could be modified.

The mouse mammary explant culture experiment (Chapter 4) in the current study demonstrated that lactogenic hormones affect the expression of certain sialic acid biosynthesis and sialyltransferase genes. In the mouse mammary explant culture

experiment reported in chapter 4, mammary tissue was taken from day 10-12 pregnant mice and either immediately frozen or utilised in the culture experiment. The frozen samples were used for comparison with the cultured samples. However, during analysis, it became apparent that endogenous lactogenic hormones had already affected the expression of the genes of interest during the first 10-12 days of pregnancy. Therefore the baseline expression level of these genes, without hormone, was unknown, and the degree of stimulation by hormones may have been underestimated. An improved approach may involve first culturing all mammary tissue without hormones for several days to minimise the effect of endogenous hormones, followed by the approach taken in the current experiment. In recent experiments, mouse mammary explants have been shown to be able to survive for several days in culture without hormones and have the ability to up-regulate milk protein gene expression with the subsequent addition of lactogenic hormones (Brennan, A.J., Sharp, J.A., et al. 2008).

RT-PCR was used to determine levels of gene expression in mammary tissue following culture of explants in either a combination of lactogenic hormones or with no hormones. Densitometric analysis of bands generated by RT-PCR during the exponential amplification phase of the cDNA can be used to gauge whether a treatment has caused a gross change in gene expression (Spitali, P., Heemskerk, H., et al. 2010). However, this method only provides an indication of an expression difference and is generally not considered quantitative. To obtain an accurate, quantitative measure of the response of sialic acid biosynthesis and sialyltransferase genes to lactogenic hormones, a technique such as qPCR should be employed. A microarray-based approach would allow us to capture the response of all relevant genes, simultaneously.

This study also attempted to identify potential transcription factor binding sites that may provide a clue as to *how* sialic acid biosynthesis and sialyltransferase genes are regulated by lactogenic hormones. This was done using an *in silico* approach combined with filtering criteria based on information known about the characteristics of functional sites. However,

due to the highly degenerate nature of transcription factor binding sites, prediction programs tend to generate many false positives and although this can be mitigated by increasing stringency or applying other criteria, it is difficult to identify "real" transcription factor binding sites computationally. Although *in silico* techniques can be effectively utilised to inform and target further analyses (Nelson, E.A., Walker, S.R., et al. 2006), an experimental approach is still essential in confirming the existence of functional transcription factor binding sites.

6.5 Conclusions

In conclusion, this study has demonstrated that transcriptional control is likely to be a key factor in the regulation of sialylation in the mammary gland during lactation and that lactogenic hormones regulate the changes in sialic acid biosynthesis and sialyltransferase gene expression. Also, given the difference in ST6Gal I gene expression demonstrated between cows and other species, it is apparent that the regulation of sialylation varies between species, and is possibly related to the developmental state of the young at birth. Furthermore, this study has shown that a comparative approach could be used to determine when, in development, milk sialyloligosaccharides are most important and provides the initial data for further studies to explore the utilisation of sialylated milk oligosaccharides by different species and why such vast qualitative and quantitative differences in sialyloligosaccharide content have evolved in the milk of mammals.

7 Bibliography

2007. Murray Goulbourn Co-Operative Co. Limited. Melbourne.

Aasheim HC, Aas-Eng DA, Deggerdal A, Blomhoff HK, Funderud S, Smeland EB. 1993. Cell-specific expression of human β-galactoside α2,6-sialyltransferase transcripts differing in the 5' untranslated region. *Eur. J. Biochem.*, 213:467-475.

Anderson SM, Rudolph MC, McManaman JL, Neville MC. 2007. Key stages in mammary gland development. Secretory activation in the mammary gland: it's not just about milk protein synthesis! *Breast Cancer Res*, 9:204.

Asakuma S, Akahori M, Kimura K, Watanabe Y, Nakamura T, Tsunemi M, Arai I, Sanai Y, Urashima T. 2007. Sialyl oligosaccharides of human colostrum: changes in concentration during the first three days of lactation. *Biosci Biotechnol Biochem*, 71:1447-1451.

Barile D, Marotta M, Chu C, Mehra R, Grimm R, Lebrilla CB, German JB. 2010. Neutral and acidic oligosaccharides in Holstein-Friesian colostrum during the first 3 days of lactation measured by high performance liquid chromatography on a microfluidic chip and time-of-flight mass spectrometry. *Journal of Dairy Science*, 93:3940-3949.

Barnawell EB. 1965. A comparative study of the responses of mammary tissues from several mammalian species to hormones in vitro. *J Exp Zool*, 160:189-206.

Bendiak B, Cook GMW. 1982. Kinetic parameters of a β-D-galactoside α 2 leads to 6 sialyltransferase from embryonic chicken liver. *Eur. J. Biochem.*, 128:355-362.

Boehm G, Stahl B. 2007. Oligosaccharides from Milk. *The Journal of Nutrition*, 137:847S-849S.

Bradbury JH, Collins JG, Jenkins GA, Trifonoff E, Messer M. 1983. 13C-N.m.r. study of the structures of two branched oligosaccharides from marsupial milk. *Carbohydr Res*, 122:327-331.

Brand-Miller JC, McVeagh P, McNeil Y, Gillard B. 1995. Human milk oligosaccharides are not digested and absorbed in the small intestine of young infants. *Proceedings-Nutrition Society of Australia*, 19:44.

Brennan AJ, Sharp JA, Khalil E, Digby MR, Mailer SL, Lefevre CM, Nicholas KR. 2008. A population of mammary epithelial cells do not require hormones or growth factors to survive. *J Endocrinol*, 196:483-496.

Brodbeck U, Denton WL, Tanahashi N, Ebner KE. 1967. The Isolation and Identification of the B Protein of Lactose Synthetase as α-Lactalbumin. *Journal of Biological Chemistry*, 242:1391-1397.

Brunngraber EG, Witting LA, Haberland C, Brown B. 1972. Glycoproteins in Tay-sachs disease: isolation and carbohydrate composition of glycopeptides. *Brain research*, 38:151-162.

Bry C, Maass K, Miyoshi K, Willecke K, Ott T, Robinson GW, Hennighausen L. 2004. Loss of connexin 26 in mammary epithelium during early but not during late pregnancy results in unscheduled apoptosis and impaired development. *Dev Biol*, 267:418-429.

Carlson SE. 1985. N-acetylneuraminic acid concentrations in human-milk oligosaccharides and glycoproteins during lactation. *American Journal of Clinical Nutrition*, 41:720-726.

Carlson SE, House SG. 1986. Oral and intraperitoneal administration of N-acetylneuraminic acid - effect on rat cerebral and cerebellar N-acetylneuraminic acid. *Journal of Nutrition*, 116:881-886.

Chia DJ, Varco-Merth B, Rotwein P. 2010. Dispersed Chromosomal Stat5b-binding Elements Mediate Growth Hormone-activated Insulin-like Growth Factor-I Gene Transcription. *Journal of Biological Chemistry*, 285:17636-17647.

Clarkson R, Wayland M, Lee J, Freeman T, Watson C. 2004. Gene expression profiling of mammary gland development reveals putative roles for death receptors and immune mediators in post-lactational regression. *Breast Cancer Res*, 6:R92 - R109.

Clarkson RWE, Watson CJ. 2003. Microarray Analysis of the Involution Switch. *Journal of Mammary Gland Biology and Neoplasia*, 8:309-319.

Collet C, Joseph R, Nicholas K. 1991. A marsupial beta-lactoglobulin gene: characterization and prolactin-dependent expression. *J Mol Endocrinol*, 6:9-16.

Collins JG, Bradbury JH, Trifonoff E, Messer M. 1981. Structures of four new oligosaccharides from marsupial milk, determined mainly by 13C-n.m.r. spectroscopy. *Carbohydr Res*, 92:136-140.

Coppa GV, Gabrielli O, Giorgi P, Catassi C, Montanari MP, et al. 1990. Preliminary study of breastfeeding and bacterial adhesion to uroepithelial cells. *The Lancet; London*, 335:569.

Coppa GV, Pierani P, Zampini L, Carloni I, Carlucci A, Gabrielli O. 1999. Oligosaccharides in human milk during different phases of lactation. *Acta Paediatrica*, 88:89-94.

Cowie AT, Hartmann PE, Turvey A. 1969. The maintenance of lactation in the rabbit after hypophysectomy. *J Endocrinol*, 43:651-662.

Crisp EA, Cowan PE, Messer M. 1989. Changes in milk carbohydrates during lactation in the common brushtail possum, Trichosurus vulpecula (Marsupialia:Phalangeridae). *Reproduction, Fertility and Development*, 1:309-314.

Crisp EA, Czolij R, Messer M. 1987. Absence of beta-galactosidase (lactase) activity from intestinal brush borders of suckling macropods: implications for mechanism of lactose absorption. *Comp Biochem Physiol B*, 88:923-927.

Cui Y, Riedlinger G, Miyoshi K, Tang W, Li C, Deng C-X, Robinson GW, Hennighausen L. 2004. Inactivation of Stat5 in Mouse Mammary Epithelium during Pregnancy Reveals Distinct Functions in Cell Proliferation, Survival, and Differentiation. *Mol. Cell. Biol.*, 24:8037-8047.

Dalziel M, Huang RY, Dall'Olio F, Morris JR, Taylor-Papadimitriou J, Lau JTY. 2001. Mouse ST6Gal sialyltransferase gene expression during mammary gland lactation. *Glycobiology*, 11:407-412.

Dalziel M, Lemaire S, Ewing J, Kobayashi L, Lau JT. 1999. Hepatic acute phase induction of murine β-galactoside α2,6 sialyltransferase (ST6Gal I) is IL-6 dependent and mediated by elevation of Exon H-containing class of transcripts. *Glycobiology*, 9:1003-1008.

Datta AK, Paulson JC. 1995. The sialyltransferase "sialylmotif" participates in binding the donor substrate CMP-NeuAc. *Journal of Biological Chemistry*, 270:1497-1500.

Datta AK, Sinha A, Paulson JC. 1998. Mutation of the sialyltransferase S-sialylmotif alters the kinetics of the donor and acceptor substrates. *Journal of Biological Chemistry*, 273:9608-9614.

Devinoy E, Houdebine LM, Delouis C. 1978. Role of prolactin and glucocorticoids in the expression of casein genes in rabbit mammary gland organ culture. Quantification of casein mRNA. *Biochim Biophys Acta*, 517:360-366.

Dickson JJ, Messer M. 1978. Intestinal neuraminidase activity of suckling rats and other mammals - relationship to sialic-acid content of milk. *Biochem. J.*, 170:407-413.

Dobbing J. 1981. Nutritional growth restriction and the nervous system. The molecular basis of neuropathology. London: Edward Arnold. p. 221-223.

Dodd SC, Forsyth IA, Buttle HL, Gurr MI, Dils RR. 1994. Hormonal induction of α-lactalbumin and β-lactoglobulin in cultured mammary explants from pregnant pigs. *Journal of Dairy Research*, 61:35-45.

Donadio S, Dubois C, Fichant G, Roybon L, Guillemot J-C, Breton C, Ronin C. 2003. Recognition of cell surface acceptors by two human α-2,6-sialyltransferases produced in CHO cells. *Biochimie*, 85:311-321.

Donovan D, Kerr D, Wall R. 2005. Engineering Disease Resistant Cattle. *Transgenic Research*, 14:563-567.

Duncan PI, Raymond F, Fuerholz A, Sprenger N. 2009. Sialic Acid Utilisation and Synthesis in the Neonatal Rat Revisited. *PLoS One*, 4:e8241.

Elias JJ. 1957. Cultivation of adult mouse mammary gland in hormone-enriched synthetic medium. *Science*, 126:842-843.

Elias JJ. 1959. Effect of insulin and cortisol on organ cultures of adult mouse mammary gland. *Proc Soc Exp Biol Med*, 101:500-502.

Elias JJ, Rivera E. 1959. Comparison of the responses of normal, precancerous, and neoplastic mouse mammary tissues to hormones in vitro. *Cancer Res*, 19:505-511.

Falk PG, Hooper LV, Midtvedt T, Gordon JI. 1998. Creating and maintaining the gastrointestinal ecosystem: what we know and need to know from gnotobiology. *Microbiol Mol Biol Rev*, 62:1157-1170.

Fergusson DM, Beautrais AL, Silva PA. 1982. Breast-feeding and cognitive development in the first seven years of life. *Social Science & Medicine*, 16:1705-1708.

Frank DN, Pace NR. 2008. Gastrointestinal microbiology enters the metagenomics era. *Current opinion in gastroenterology*, 24:4-10.

Fuhrer A, Sprenger N, Kurakevich E, Borsig L, Chassard C, Hennet T. 2010. Milk sialyllactose influences colitis in mice through selective intestinal bacterial colonization. *The Journal of Experimental Medicine*, 207:2843-2854.

Furukawa K, Sato T. 1999. Beta-1,4-galactosylation of N-glycans is a complex process. *Biochim Biophys Acta*, 1473:54-66.

Gal B, Ruano MJ, Puente R, GarciaPardo LA, Rueda R, Gil A, Hueso P. 1997. Developmental changes in UDP-N-acetylglucosamine 2-epimerase activity of rat and guinea-pig liver. *Comp. Biochem. Phys. B*, 118:13-15.

Geremia RA, HarduinLepers A, Delannoy P. 1997. Identification of two novel conserved amino acid residues in eukaryotic sialyltransferases: Implications for their mechanism of action. *Glycobiology*, 7:R5-R7.

German JB, Freeman SL, Lebrilla CB, Mills DA. 2008. Human milk oligosaccharides: evolution, structures and bioselectivity as substrates for intestinal bacteria. *Nestle Nutr Workshop Ser Pediatr Program*, 62:205-218; discussion 218-222.

Gibson GR, Roberfroid MB. 1995. Dietary Modulation of the Human Colonic Microbiota: Introducing the Concept of Prebiotics. *The Journal of Nutrition*, 125:1401-1412.

Green B, Merchant JC. 1988. The Composition of Marsupial Milk. In: Tyndale-Biscoe CH, Janssens PA editors. The Developing Marsupial. Berlin Heidelberg: Springer-Verlag.

Griffiths M. 1978. The Biology of Monotremes Academic Press:New York.

Hamosh M. 1999. Breastfeeding: unravelling the mysteries of mother's milk. *Medscape General Medicine*, 1.

Harduin-Lepers A, Mollicone R, Delannoy P, Oriol R. 2005. The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. *Glycobiology*, 15:805-817.

Harduin-Lepers A, Recchi MA, Delannoy P. 1995. 1994, the year of sialyltransferases. *Glycobiology*, 5:741-758.

Harduin-Lepers A, Vallejo-Ruiz V, Krzewinski-Recchi MA, Samyn-Petit B, Julien S, Delannoy P. 2001. The human sialyltransferase family. *Biochimie*, 83:727-737.

Harigaya T, Sakai S, Kohmoto K. 1978. Induction of mammary prolactin receptors and lactose synthesis after ovariectomy in the pregnant mouse. *Endocrinol Jpn*, 25:157-161.

Harmsen HJM. 2000. Analysis of intestinal flora development in breast-fed and formula-fed infants by using molecular identification and detection methods. *Journal of Pediatric Gastroenterology and Nutrition*, 30:61-67.

Hartmann PE. 1973. Changes in the composition and yield of the mammary secretion of cows during the initiation of lactation. *J Endocrinol*, 59:231-247.

Hennet T. 2002. The galactosyltransferase family. *Cell Mol Life Sci*, 59:1081-1095.

Humphreys RC, Lydon JP, O'Malley BW, Rosen JM. 1997. Use of PRKO mice to study the role of progesterone in mammary gland development. *J Mammary Gland Biol Neoplasia*, 2:343-354.

Jeanneau C, Chazalet V, Auge C, Soumpasis DM, Harduin-Lepers A, Delannoy P, Imberty A, Breton C. 2004. Structure-function analysis of the human sialyltransferase ST3Gal I - Role of N-glycosylation and a novel conserved sialylmotif. *Journal of Biological Chemistry*, 279:13461-13468.

Jenness R, Regehr EA, Sloan RE. 1964. Comparative biochemical studies of milk. II. Dialyzable carbohydrates. *Comp Biochem Physiol*, 13:339-352.

Juergens WG, Stockdale FE, Topper YJ, Elias JJ. 1965. Hormone-dependent differentiation of mammary gland in vitro. *Proc Natl Acad Sci U S A*, 54:629-634.

Keppler OT, Hinderlich S, Langner J, Schwartz-Albiez R, Reutter W, Pawlita M. 1999. UDP-GlcNAc 2-Epimerase: A Regulator of Cell Surface Sialylation. *Science*, 284:1372-1376.

Klover PJ, Muller WJ, Robinson GW, Pfeiffer RM, Yamaji D, Hennighausen L. 2010. Loss of STAT1 from mouse mammary epithelium results in an increased Neu-induced tumor burden. *Neoplasia*, 12:899-905.

Kuhn NJ. 1972. The lactose and neuraminlactose content of rat milk and mammary tissue. *Biochemical Journal*, 130:177-180.

Kuhn NJ. 2009. Progesterone withdrawal as the lactogenic trigger in the rat. 1969. *J Mammary Gland Biol Neoplasia*, 14:327-342.

Kunz C, Rudloff S. 1993. Biological functions of oligosaccharides in human-milk. *Acta Paediatrica*, 82:903-912.

Kunz C, Rudloff S, Schad W, Braun D. 1999. Lactose-derived oligosaccharides in the milk of elephants: comparison with human milk. *British Journal of Nutrition*, 82:391-399.

Lechner J, Welte T, Tomasi JK, Bruno P, Cairns C, Gustafsson J, Doppler W. 1997. Promoter-dependent synergy between glucocorticoid receptor and Stat5 in the activation of beta-casein gene transcription. *J Biol Chem*, 272:20954-20960.

Legaigneur P, Breton C, El Battari A, Guillemot J-C, Auge C, Malissard M, Berger EG, Ronin C. 2001. Exploring the Acceptor Substrate Recognition of the Human beta-Galactoside alpha 2,6-Sialyltransferase. *J. Biol. Chem.*, 276:21608-21617.

Lemay D, Neville M, Rudolph M, Pollard K, German J. 2007. Gene regulatory networks in lactation: identification of global principles using bioinformatics. *BMC Systems Biology*, 1:56.

Lentle RG, Dey D, Hulls C, Mellor DJ, Moughan PJ, Stafford KJ, Nicholas K. 2006. A quantitative study of the morphological development and bacterial colonisation of the gut of the tammar wallaby Macropus eugenii eugenii and brushtail possum Trichosurus vulpecula during in-pouch development. *J Comp Physiol B*, 176:763-774.

Li S, Rosen JM. 1995. Nuclear factor I and mammary gland factor (STAT5) play a critical role in regulating rat whey acidic protein gene expression in transgenic mice. *Mol Cell Biol*, 15:2063-2070.

LoCascio RG, Ninonuevo MR, Freeman SL, Sela DA, Grimm R, Lebrilla CB, Mills DA, German JB. 2007. Glycoprofiling of bifidobacterial consumption of human milk oligosaccharides demonstrates strain specific, preferential consumption of small chain glycans secreted in early human lactation. *J Agric Food Chem*, 55:8914-8919.

Lönnerdal B, Lien EL. 2003. Nutritional and Physiologic Significance of α-Lactalbumin in Infants. *Nutrition Reviews*, 61:295-305.

Lucas A, Morley R, Cole TJ. 1998. Randomised trial of early diet in preterm babies and later intelligence quotient. *BMJ*, 317:1481-1487.

Lucas A, Morley R, Cole TJ, Lister G, Leeson-Payne C. 1992. Breast milk and subsequent intelligence quotient in children born preterm. *The Lancet*, 339:261-264.

Macpherson AJ, Harris NL. 2004. Interactions between commensal intestinal bacteria and the immune system. *Nat Rev Immunol*, 4:478-485.

Magne S, Caron S, Charon M, Rouyez M-C, Dusanter-Fourt I. 2003. STAT5 and Oct-1 Form a Stable Complex That Modulates Cyclin D1 Expression. *Mol. Cell. Biol.*, 23:8934-8945.

Marcobal A, Barboza M, Froehlich JW, Block DE, German JB, Lebrilla CB, Mills DA. 2010. Consumption of Human Milk Oligosaccharides by Gut-Related Microbes. *Journal of Agricultural and Food Chemistry*, 58:5334-5340.

Martin-Sosa S, Martin MJ, Garcia-Pardo LA, Hueso P. 2003. Sialyloligosaccharides in human and bovine milk and in infant formulas: variations with the progression of lactation. *J. Dairy Sci.*, 86:52-59.

McVeagh P, Brand-Miller J. 1997. Human milk oligosaccharides: only the breast. *J. Paediatr. Child Health*, 33:281-286.

Menzies KK, Lefevre C, Sharp JA, Macmillan KL, Sheehy PA, Nicholas KR. 2009. A novel approach identified the FOLR1 gene, a putative regulator of milk protein synthesis. *Mamm Genome*, 20:498-503.

Mercier D, Wierinckx A, Oulmouden A, Gallet PF, Palcic MM, Harduin-Lepers A, Delannoy P, Petit JM, Leveziel H, Julien R. 1999. Molecular cloning, expression and exon/intron organization of the bovine β-galactoside α2,6-sialyltransferase gene. *Glycobiology*, 9:851-863.

Messer M. 1974. Identification of N-acetyl-4-O-acetylneuraminyl-lactose in Echidna milk. *Biochemical Journal*, 139:415-420.

Messer M, Crisp EA, Czolij R. 1989. Kangaroos, Wallabies and Rat-Kangaroos Surry Beatty and Sons Pty Ltd:NSW, Australia.

Messer M, FitzGerald PA, Merchant JC, Green B. 1987. Changes in milk carbohydrates during lactation in the eastern quoll, Dasyurus viverrinus (Marsupialia). *Comp Biochem Physiol B*, 88:1083-1086.

Messer M, Green B. 1979. Milk carbohydrates of marsupials. II. Quantitative and qualitative changes in milk carbohydrates during lactation in the tammar wallaby (Macropus eugenii). *Aust J Biol Sci*, 32:519-531.

Messer M, Kerry KR. 1973. Milk carbohydrates of echidna and platypus. *Science*, 180:201-203.

Messer M, Mossop G. 1977. Milk Carbohydrates of Marsupials I. Partial Separation and Characterization of Neutral Milk Oligosaccharides of the Eastern Grey Kangaroo. *Australian Journal of Biological Sciences*, 30:379-388.

Messer M, Nicholas KR. 1991. Biosynthesis of marsupial milk oligosaccharides: characterization and developmental changes of two galactosyltransferases in lactating mammary glands of the tammar wallaby, Macropus eugenii. *Biochim Biophys Acta*, 1077:79-85.

Messer M, Trifonoff E. 1982. Structure of a branched tetrasaccharide from marsupial milk. *Carbohydr Res*, 102:316-320.

Messer M, Trifonoff E, Stern W, Collins JG, Bradbury JH. 1980. Structure of a marsupial-mild trisaccharide. *Carbohydr Res*, 83:327-334.

Messer M, Urashima T. 2002. Evolution of Milk Oligosaccharides and Lactose. *Trends in Glycoscience and Glycotechnology*, 14:153-176.

Miyagi T, Tsuiki S. 1982. Purification and characterization of β-galactoside α2,6-sialyltransferase from rat liver and hepatomas. *Eur. J. Biochem.*, 126:253-261.

Mizoguchi Y, Kim JY, Sasaki T, Hama T, Sasaki M, Enami J, Sakai S. 1996. Acute expression of the PRL receptor gene after ovariectomy in midpregnant mouse mammary gland. *Endocr J*, 43:537-544.

Morgan BL, Winick M. 1981. The subcellular localization of administered N-acetylneuraminic acid in the brains of well-fed and protein restricted rats. *British Journal of Nutrition*, 46:231-238.

Morrison JF, Ebner KE. 1971. Studies on Galactosyltransferase. *Journal of Biological Chemistry*, 246:3992-3998.

Moughan PJ, Birtles MJ, Cranwell PD, Smith WC, Pedraza M. 1992. The piglet as a model animal for studying aspects of digestion and absorption in milk-fed human infants. *World Rev Nutr Diet*, 67:40-113.

Mouricout M, Petit JM, Carias JR, Julien R. 1990. Glycoprotein glycans that inhibit adhesion of Escherichia coli mediated by K99 fimbriae: treatment of experimental colibacillosis. *Infect. Immun.*, 58:98-106.

Münster A-K, Eckhardt M, Potvin B, Mühlenhoff M, Stanley P, Gerardy-Schahn R. 1998. Mammalian cytidine 5′-monophosphate N-acetylneuraminic acid synthetase: A nuclear protein with evolutionarily conserved structural motifs. *Proceedings of The National Academy of Sciences of The United States of America*, 95:9140-9145.

Nakano T, Sugawara M, Kawakami H. 2001. Sialic acid in human milk: composition and functions. *Acta Paediatr. Taiwan*, 42:11-17.

Nakata D, Close BE, Colley KJ, Matsuda T, Kitajima K. 2000. Molecular Cloning and Expression of the Mouse N-Acetylneuraminic Acid 9-Phosphate Synthase Which Does Not Have Deaminoneuraminic Acid (KDN) 9-Phosphate Synthase Activity. *Biochemical and Biophysical Research Communications*, 273:642-648.

Nandi S, Bern HA. 1960. Relation between mammary-gland responses to lactogenic hormone combinations and tumor susceptibility in various strains of mice. *J Natl Cancer Inst*, 24:907-931.

Nelson EA, Walker SR, Alvarez JV, Frank DA. 2004. Isolation of Unique STAT5 Targets by Chromatin Immunoprecipitation-based Gene Identification. *Journal of Biological Chemistry*, 279:54724-54730.

Nelson EA, Walker SR, Li W, Liu XS, Frank DA. 2006. Identification of Human STAT5-dependent Gene Regulatory Elements Based on Interspecies Homology. *Journal of Biological Chemistry*, 281:26216-26224.

Neville MC, Morton J, Umemura S. 2001. Lactogenesis. The transition from pregnancy to lactation. *Pediatr Clin North Am*, 48:35-52.

Newburg DS. 2000. Oligosaccharides in human milk and bacterial colonization. *Journal of Pediatric Gastroenterology and Nutrition*, 30:S8-S17.

Nicholas KR, Tyndale-Biscoe CH. 1985. Prolactin-dependent accumulation of alpha-lactalbumin in mammary gland explants from the pregnant tammar wallaby (Macropus eugenii). *J Endocrinol*, 106:337-342.

Ninonuevo MR, Park Y, Yin H, Zhang J, Ward RE, Clowers BH, German JB, Freeman SL, Killeen K, Grimm R*, et al.* 2006. A strategy for annotating the human milk glycome. *J Agric Food Chem*, 54:7471-7480.

Nohle U, Schauer R. 1984. Metabolism of sialic acids from exogenously administered sialyllactose and mucin in mouse and rat. *Hoppe Seylers Z Physiol Chem*, 365:1457-1467.

O'Keefe E, Cuatrecasas P. 1974. Insulin receptors in murine mammary cells: comparison in pregnant and nonpregnant animals. *Biochim Biophys Acta*, 343:64-77.

Oftedal OT, Alt GL, Widdowson EM, Jakubasz MR. 1993. Nutrition and growth of suckling black bears (Ursus americanus) during their mothers' winter fast. *Br J Nutr*, 70:59-79.

Ono M, Oka T. 1980a. alpha-Lactalbumin-casein induction in virgin mouse mammary explants: dose-dependent differential action of cortisol. *Science*, 207:1367-1369.

Ono M, Oka T. 1980b. The differential actions of cortisol on the accumulation of alpha-lactalbumin and casein in midpregnant mouse mammary gland in culture. *Cell*, 19:473-480.

Ormandy CJ, Binart N, Kelly PA. 1997. Mammary gland development in prolactin receptor knockout mice. *J Mammary Gland Biol Neoplasia*, 2:355-364.

Osborne JC, Steiner RF. 1974. Interaction of the components of the lactose synthetase system. *Arch Biochem Biophys*, 165:615-627.

Ostrom KM. 1990. A review of the hormone prolactin during lactation. *Prog Food Nutr Sci*, 14:1-43.

Paulson JC, Beranek WE, Hill RL. 1977. Purification of a sialyltransferase from bovine colostrum by affinity chromatography on CDP-agarose. *J. Biol. Chem.*, 252:2356-2362.

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

Picciano MF. 1998. Human milk: nutritional aspects of a dynamic food. *Biology of the Neonate*, 74:84-93.

Powell A, Kerr D, Guthrie D, Wall R. 2007. Lactation induction as a predictor of post-parturition transgene expression in bovine milk. *Journal of Dairy Research*, 74:247-254.

Prieto PA, Mukerji P, Kelder B, Erney R, Gonzalez D, Yun JS, Smith DF, Moremen KW, Nardelli C, Pierce M, *et al.* 1995. Remodeling of Mouse Milk Glycoconjugates by Transgenic Expression of a Human Glycosyltransferase. *Journal of Biological Chemistry*, 270:29515-29519.

Renfree MB. 1983. In: Finn C editor. Oxford Reviews of Reproductive Biology. England: Oxford Univ. Press. p. 1–26.

Rini J, Esko J, Varki A. 2009. Glycosyltransferases and Glycan-processing Enzymes. In: Varki A, Cummings R, Esko J, Freeze H, Stanley P, Bertozzi C, Hart G, Etzler M editors. Essentials of Glycobiology. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

Rivera EM, Bern HA. 1961. Influence of insulin on maintenance and secretory stimulation of mouse mammary tissues by hormones in organ-culture. *Endocrinology*, 69:340-353.

Rivera EM, Forsyth IA, Folley SJ. 1967. Lactogenic activity of mammalian growth hormones in vitro. *Proc Soc Exp Biol Med*, 124:859-865.

Rodgers B. 1978. Feeding in infancy and later ability and attainment - longitudinal-study. *Developmental Medicine and Child Neurology*, 20:421-426.

Rosenburg A. 1995. Biology of Sialic Acids. New York: Plenum Press.

Rudolph MC, McManaman JL, Hunter L, Phang T, Neville MC. 2003. Functional development of the mammary gland: use of expression profiling and trajectory clustering to reveal changes in gene expression during pregnancy, lactation, and involution. *J Mammary Gland Biol Neoplasia*, 8:287-307.

Russell EM. 1982. Patterns of parental care and parental investment in marsupials. *Biol Rev Camb Philos Soc*, 57:423-486.

Sankaran L, Topper YJ. 1984. Prolactin-induced alpha-lactalbumin activity in mammary explants from pregnant rabbits. A role for epidermal growth factor and glucocorticoids. *Biochem J*, 217:833-837.

Savage DC. 1977. Microbial ecology of the gastrointestinal tract. *Annu Rev Microbiol*, 31:107-133.

Schauer R. 1982. Sialic Acids, Chemistry, Metabolism, and Function Springer-Verlag:Wien, New York.

Sela DA, Li Y, Lerno L, Wu S, Marcobal AM, German JB, Chen X, Lebrilla CB, Mills DA. 2011. An infant-associated bacterial commensal utilizes breast milk sialyloligosaccharides. *Journal of Biological Chemistry*.

Spitali P, Heemskerk H, Vossen RH, Ferlini A, den Dunnen JT, t Hoen PA, Aartsma-Rus A. 2010. Accurate quantification of dystrophin mRNA and exon skipping levels in duchenne muscular dystrophy. *Lab Invest*, 90:1396-1402.

Stacey A, Schnieke A, Kerr M, Scott A, McKee C, Cottingham I, Binas B, Wilde C, Colman A. 1995. Lactation is disrupted by alpha-lactalbumin deficiency and can be restored by human alpha-lactalbumin gene replacement in mice. *Proc Natl Acad Sci U S A*, 92:2835-2839.

Stein T, Morris J, Davies C, Weber-Hall S, Duffy M-A, Heath V, Bell A, Ferrier R, Sandilands G, Gusterson B. 2004. Involution of the mouse mammary gland is associated with an immune cascade and an acute-phase response, involving LBP, CD14 and STAT3. *Breast Cancer Res.*, 6:R75 - R91.

Stein T, Salomonis N, Gusterson B. 2007. Mammary Gland Involution as a Multi-step Process. *Journal of Mammary Gland Biology and Neoplasia*, 12:25-35.

Sticher U, Gross HJ, Brossmer R. 1991. Purification and characterization of α2,6-sialyltransferase from human liver. *Glycoconj. J.*, 8:45-54.

Stinnakre MG, Vilotte JL, Soulier S, Mercier JC. 1994. Creation and phenotypic analysis of alpha-lactalbumin-deficient mice. *Proc Natl Acad Sci U S A*, 91:6544-6548.

Svennerholm L, Bostrom K, Fredman P, Mansson JE, Rosengren B, Rynmark BM. 1989. Human brain gangliosides: developmental changes from early fetal stage to advanced age. *Biochimica et Biophysica Acta*, 1005:109-117.

Svensson EC, Soreghan B, Paulson JC. 1990. Organization of the β-galactoside α2,6-sialyltransferase gene. Evidence for the transcriptional regulation of terminal glycosylation. *J. Biol. Chem.*, 265:20863-20868.

Takashima S. 2008. Characterization of Mouse Sialyltransferase Genes: Their Evolution and Diversity. *Bioscience, biotechnology, and biochemistry*, 72:1155-1167.

Takashima S, Kono M, Kurosawa N, Yoshida Y, Tachida Y, Inoue M, Kanematsu T, Tsuji S. 2000. Genomic organization and transcriptional regulation of the mouse GD3 synthase gene (ST8Sia I): Comparison of genomic organization of the mouse sialyltransferase genes. *Journal of Biochemistry*, 128:1033-1043.

Tanner ME. 2005. The enzymes of sialic acid biosynthesis. *Bioorganic Chemistry*, 33:216-228.

Tao N, DePeters E, Freeman S, German J, Grimm R, Lebrilla C. 2008. Bovine Milk Glycome. *Journal of Dairy Science*, 91:3768-3778.

Tao N, DePeters E, German J, Grimm R, Lebrilla C. 2009. Variations in bovine milk oligosaccharides during early and middle lactation stages analyzed by high-performance liquid chromatography-chip/mass spectrometry. *Journal of Dairy Science*, 92:2991-3001.

Tao N, Ochonicky KL, German JB, Donovan SM, Lebrilla CB. 2010. Structural determination and daily variations of porcine milk oligosaccharides. *J Agric Food Chem*, 58:4653-4659.

Tao N, Wu S, Kim J, An HJ, Hinde K, Power ML, Gagneux P, German JB, Lebrilla CB. 2011. Evolutionary Glycomics: Characterization of Milk Oligosaccharides in Primates. *Journal of Proteome Research*:null-null.

Traurig HH. 1967. Cell proliferation in the mammary gland during late pregnancy and lactation. *The Anatomical Record*, 157:489-503.

Tsuji S. 1996a. Molecular Cloning and Functional Analysis of Sialyltransferases. *J Biochem (Tokyo)*, 120:1-13.

Tsuji S. 1996b. Systematic nomenclature for sialyltransferases. *Glycobiology*, 6:R5-R7.

Turkington RW, Juergens WG, Topper YJ. 1965. Hormone-dependent synthesis of casein in vitro. *Biochim Biophys Acta*, 111:573-576.

Tyndale-Biscoe CH, Janssens PA. 1988. Introduction. In: Tyndale-Biscoe CH, Janssens PA editors. The Developing Marsupial. Berlin, Heidelberg: Springer-Verlag.

Urashima T, Kusaka Y, Nakamura T, Saito T, Maeda N, Messer M. 1997. Chemical characterization of milk oligosaccharides of the brown bear, Ursus arctos yesoensis. *Biochim Biophys Acta*, 1334:247-255.

Urashima T, Saito T, Nakamura T, Messer M. 2001. Oligosaccharides of milk and colostrum in non-human mammals. *Glycoconjugate Journal*, 18:357-371.

Urashima T, Saito T, Tsuji Y, Taneda Y, Takasawa T, Messer M. 1994. Chemical characterization of sialyl oligosaccharides isolated from tammar wallaby (Macropus eugenii) milk. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1200:64-72.

Urashima T, Sumiyoshi W, Nakamura T, Arai I, Saito T, Komatsu T, Tsubota T. 1999. Chemical characterization of milk oligosaccharides of the Japanese black bear, Ursus thibetanus japonicus. *Biochim Biophys Acta*, 1472:290-306.

Urashima T, Yamashita T, Nakamura T, Arai I, Saito T, Derocher AE, Wiig O. 2000. Chemical characterization of milk oligosaccharides of the polar bear, Ursus maritimus. *Biochim Biophys Acta*, 1475:395-408.

Vanier MT, Holm M, Ohman R, Svennerholm L. 1971. Developmental profile of gangliosides in human and rat brain. *J Neurochem*, 18:581-592.

Varki A, Schauer R. 2009. Sialic Acids. In: Varki A, Cummings R, Esko J, Freeze H, Stanley P, Bertozzi C, Hart G, Etzler M editors. Essentials of Glycobiology. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

Veh RW, Michalski J-C, Corfield AP, Sander-Wewer M, Gies D, Schauer R. 1981. New chromatographic system for the rapid analysis and preparation of colostrum sialyloligosaccharides. *Journal of Chromatography A*, 212:313-322.

Viverge D. 1990. Variations in oligosaccharides and lactose in human milk during the first week of lactation. *J Pediatr Gastroenterol Nutr*, 11:361.

Von Itzstein M, Thomson R. 1997. Sialic acids and sialic acid recognising proteins: drug discovery targets and potential glycopharmaceuticals. *Curr Med Chem*, 4:185–210.

Wakao H, Schmitt-Ney M, Groner B. 1992. Mammary gland-specific nuclear factor is present in lactating rodent and bovine mammary tissue and composed of a single polypeptide of 89 kDa. *J Biol Chem*, 267:16365-16370.

Walcott PJ, Messer M. 1980. Intestinal lactase (beta-galactosidase) and other glycosidase activities in suckling and adult tammar wallabies (Macropus eugenii). *Aust J Biol Sci*, 33:521-530.

Wall RJ, Powell AM, Paape MJ, Kerr DE, Bannerman DD, Pursel VG, Wells KD, Talbot N, Hawk HW. 2005. Genetically enhanced cows resist intramammary Staphylococcus aureus infection. *Nat Biotech*, 23:445-451.

Wang B, Brand-Miller J. 2003. The role and potential of sialic acid in human nutrition. *Eur. J. Clin. Nutr.*, 57:1351-1369.

Wang B, Brand-Miller J, McVeagh P, Petocz P. 2001. Concentration and distribution of sialic acid in human milk and infant formulas. *Am. J. Clin. Nutr.*, 74:510-515.

Wang B, Yu B, Karim M, Hu H, Sun Y, McGreevy P, Petocz P, Held S, Brand-Miller J. 2007. Dietary sialic acid supplementation improves learning and memory in piglets. *Am J Clin Nutr*, 85:561-569.

Wang X, O'Hanlon TP, Young RF, Lau JTY. 1990. Rat β-galactoside α2,6-sialyltransferase genomic organization: alternate promoters direct the synthesis of liver and kidney transcripts. *Glycobiology*, 1:25-31.

Wang X, Vertino A, Eddy RL, Byers MG, Jani-Sait SN, Shows TB, Lau JT. 1993. Chromosome mapping and organization of the human β-galactoside α2,6-sialyltransferase gene. Differential and cell-type specific usage of upstream exon sequences in B-lymphoblastoid cells. *J. Biol. Chem.*, 268:4355-4361.

Weinstein J, de Souza-e-Silva U, Paulson JC. 1982a. Purification of a Gal β1 to 4GlcNAc α2 to 6 sialyltransferase and a Gal β1 to 3(4)GlcNAc α2 to 3 sialyltransferase to homogeneity from rat liver. *J. Biol. Chem.*, 257:13835-13844.

Weinstein J, de Souza-e-Silva U, Paulson JC. 1982b. Sialylation of glycoprotein oligosaccharides N-linked to asparagine. Enzymatic characterization of a Gal β1 to 3(4)GlcNAc α2 to 3 sialyltransferase and a Gal β1 to 4GlcNAc α2 to 6 sialyltransferase from rat liver. *J. Biol. Chem.*, 257:13845-13853.

Wen DX, Svensson EC, Paulson JC. 1992. Tissue-specific alternative splicing of the β-galactoside α2,6- sialyltransferase gene. *J. Biol. Chem.*, 267:2512-2518.

Witt W, von Nicolai H, Zilliken F. 1979. Uptake and distribution of orally applied N-acetyl-(14C)neuraminosyl-lactose and N-acetyl-(14C)neuraminic acid in the organs of newborn rats. *Nutrition and Metabolism*, 23:51-61.

Wu S, Grimm R, German JB, Lebrilla CB. 2010. Annotation and Structural Analysis of Sialylated Human Milk Oligosaccharides. *Journal of Proteome Research*, 10:856-868.

Wuensch SA, Huang RY, Ewing J, Liang X, Lau JT. 2000. Murine B cell differentiation is accompanied by programmed expression of multiple novel β-galactoside α2, 6-sialyltransferase mRNA forms. *Glycobiology*, 10:67-75.

Zivkovic AM, German JB, Lebrilla CB, Mills DA. Human milk glycobiome and its impact on the infant gastrointestinal microbiota. *Proceedings of the National Academy of Sciences*.

8 Appendices

Glycobiology vol. 21 no. 4 pp. 467–481, 2011
doi:10.1093/glycob/cwq185
Advance Access publication on November 22, 2010

Conservation of the *ST6Gal I* gene and its expression in the mammary gland

Jovana Maksimovic[1,2,3], Julie A Sharp[4], Kevin R Nicholas[4], Benjamin G Cocks[3], and Keith Savin[3]

[2]Centre for Reproduction and Development, Monash Institute of Medical Research, Clayton 3168, Australia; [3]Biosciences Research Division, Department of Primary Industries, Bundoora 3083, Australia; and [4]Institute for Technology Research and Innovation, Deakin University, Geelong 3214, Australia

Received on December 2, 2009; revised on November 2, 2010; accepted on November 3, 2010

Milk sialoglycoconjugates can protect the gastrointestinal tract of the suckling neonate by competitively binding to invading pathogens and promoting growth of beneficial flora, and their potential role in postnatal brain development is of particular interest in human infant nutrition. Although the concentration and the distribution of sialoglycoconjugates have been extensively studied in the milk of various species, the investigation of sialyltransferase gene expression in the mammary gland, in the context of lactation, has been limited. The sialyltransferase enzyme *ST6Gal I* transfers sialic acid from CMP-sialic acid to type 2 (Galβ1,4GlcNAc) free disaccharides or the termini of *N*- or *O*-linked oligosaccharides using an α2,6-linkage. Expression of the *ST6Gal I* gene is primarily regulated at the level of transcription through the use of several cell and development-specific promoters, producing transcripts with divergent 5′ untranslated regions (UTR). In the mouse mammary gland, the novel 5′UTR exon (L) appears to be associated with a drastic increase in *ST6Gal I* gene expression during lactation. We find that rats also possess an exon (L), suggesting conservation of this regulatory mechanism in rodents. In contrast, an exon (L)-containing transcript was not detected in the lactating bovine or human mammary gland. We also observed a trend of increasing *ST6Gal I* gene expression in the bovine mammary gland, culminating in involution. This is in contrast to species such as mice where the greatest change in *ST6Gal I* gene expression occurs between pregnancy and lactation, suggesting different roles in rodents vs. other mammals for α2,6-sialylated oligosaccharides present in milk.

Keywords: mammary gland / milk / sialic acid / sialyltransferase / ST6Gal

[1]To whom correspondence should be addressed: Tel: +61-3-8341-6221; e-mail: jovana.maksimovic@mcri.edu.au

Introduction

Sialoglycoconjugates are integral to various, ubiquitous biological phenomena, such as cell–cell interaction, cell migration, adhesion and metastasis (Varki 1993; Hanasaki et al. 1995; Dall'Olio 2000; Varki 2007). Interestingly, they may have been co-opted by the mammary gland to fulfill a different role. Sialoglycoconjugates found in milk can protect the gastrointestinal tract of the suckling neonate by competitively binding to invading pathogens and promoting growth of beneficial flora (Coppa et al. 1990; Mouricout et al. 1990; McVeagh and Miller 1997; Wang et al. 2001; Martin-Sosa et al. 2003), and their potential role in postnatal brain development is of particular interest in human infant nutrition (McVeagh and Miller 1997; Wang and Brand-Miller 2003). Although the concentration and the distribution of sialoglycoconjugates have been extensively studied in the milk of various species, the investigation of sialyltransferase gene expression in the mammary gland during lactation has been limited.

The sialyltransferase enzyme *ST6Gal I* transfers sialic acid from CMP-sialic acid to either type 2 (Galβ1,4GlcNAc) free disaccharides or the *N*-acetyllactosamine termini of *N*- or *O*-linked oligosaccharides via an α2,6-linkage (Harduin-Lepers et al. 2001). *ST6Gal I* is expressed in a variety of cell types and tissues (Aasheim et al. 1993; Wang et al. 1993; Mercier et al. 1999; Dalziel et al. 2001), with particularly high levels of expression observed in the liver (Dalziel et al. 1999; Dalziel et al. 2001). The enzyme has been successfully purified from rat, chicken and human liver and rat hepatoma cells (Bendiak and Cook 1982; Miyagi and Tsuiki 1982; Sticher et al. 1991; Weinstein et al. 1982a, 1982b). A soluble form of *ST6Gal I* was also purified from bovine colostrum (Paulson et al. 1977) and was established to be the product of proteolytic cleavage of the Golgi membrane-bound enzyme.

Mercier et al. (1999) identified and extensively characterized the bovine *ST6Gal I* gene, which has a coding region of 1218 nucleotides. The coding sequence consists of five exons, the last of which contains a 3′UTR of 2.7 kb. Four additional exons were identified in the 5′UTR, resulting in a total of nine exons spanning up to 80 kb of genomic DNA. The gene comprises a single open reading frame encoding a protein of 405 residues with a theoretical molecular weight of 46,250 Da. Structural feature predictions indicate that the enzyme is likely to have a type II membrane topology, which is typical of Golgi glycosyltransferases. The bovine *ST6Gal I* protein sequence exhibits a higher homology with the human enzyme

than with the rat, mouse or chicken enzymes (Mercier et al. 1999).

The human *ST6Gal I* gene shares similarities with its bovine ortholog. It is located on human chromosome 3 and also encodes a 405 residue protein across five exons, which span in excess of 40 kb of genomic DNA. Four additional exons have been discovered in the 5′UTR (Wang et al. 1993).

Transcriptional regulation of the *ST6Gal I* gene is coordinated by the use of several cell-specific promoters that produce transcripts with divergent 5′UTR (Svensson et al. 1990; Wang et al. 1990; Wen et al. 1992). The rat *ST6Gal I* gene, which has been extensively studied, is known to span over 80 kb and has a complex network of at least four promoters that regulate its expression (Wen et al. 1992).

In vivo, three main families of bovine *ST6Gal I* transcripts have been characterized, leading to speculation of the existence of at least three separate promoters regulating the expression of the gene. Despite organization of the bovine gene being similar to both human (Wang et al. 1993) and rat (Svensson et al. 1990; Wang et al. 1990; Wen et al. 1992), there are differences in the tissue distribution of its transcripts (Mercier et al. 1999).

The existence of a novel, lactogenic *ST6Gal I* promoter was proposed by the authors of a study examining the dramatic induction of *ST6Gal I* mRNA in the lactating mammary gland of mice. The predominant *ST6Gal I* mRNA species expressed in the lactating mouse mammary gland is a novel isoform containing a unique untranslated exon derived from the *ST6Gal I* 5′UTR, dubbed exon (L). Exon (L) is 203 bp long and is located >48 kb 5′ of exon (I). Virgin mammary tissue exclusively expressed the housekeeping *ST6Gal I* mRNA form and probes specific to exon (L) confirmed that the exon (L)-containing transcript was only present in the lactating mammary gland and not in other tissues that were investigated (Dalziel et al. 2001).

The mouse *ST6Gal I* gene is known to contain at least eight 5′UTR exons in addition to exons (I)–(VI) (Dalziel et al. 1999, 2001; Wuensch et al. 2000). Several studies have demonstrated that, despite not being translated, *ST6Gal I* 5′ UTR exons exhibit a degree of nucleotide sequence conservation between species, with exons (I)–(VI) typically being well conserved (Aasheim et al. 1993; Wang et al. 1993; Mercier et al. 1999).

As an initial step toward unraveling the pattern of the expression of sialyltransferase genes in the lactating mammary gland, we have investigated the conservation of the 5′UTR of the *ST6Gal I* gene between several mammalian species. In addition, using a 5′ rapid amplification of cDNA ends (5′RACE) approach, we demonstrate that an equivalent to the mouse *ST6Gal I* exon (L) is not expressed in the lactating bovine or the human mammary gland. Furthermore, expression of the bovine *ST6Gal I* gene is shown to be significantly increased at the onset of involution.

Results

Conservation analysis of the ST6Gal I *gene*

To investigate the level of conservation of 5′UTR exons characterized in the mouse and other mammalian species (Figure 1A; Wen et al. 1992; Wang et al. 1993; Mercier et al. 1999; Wuensch et al. 2000; Dalziel et al. 2001), we compared mouse, human, bovine and rat *ST6Gal I* genomic sequence, including ∼100 kb of 5′UTR sequence, using the Vista suite of tools for comparative genomics (Frazer et al. 2004). Figure 1B shows the Vista conservation plot calculated from pairwise genomic sequence alignments of the mouse *ST6Gal I* sequence with the other mammalian species. Translated exons (II)–(VI) are at least 75% conserved at the nucleotide sequence level for all pairwise comparisons, with the greatest degree of conservation seen between the mouse and the rat sequences. Ubiquitously present 5′UTR exon (I) is also conserved by at least 75% for all sequence comparisons.

As expected, all mouse 5′UTR exons are highly conserved in the rat at the nucleotide sequence level, except 5′UTR exon ($X1_b$), which does not appear to be present in the rat sequence. Mouse exon (O) is particularly well conserved across all species; more than 75% identity is observed with human and rat and just under 70% with bovine. A multiple sequence alignment of mouse exon (O), rat exon (0), human exon (Z) and cow exon (0) sequences supports the Vista finding (Figure 2B).

Our Vista analysis also shows a reasonable level of conservation in the region of mouse exon (Q), particularly between the mouse and the rat. However, a direct alignment of the mouse exon (Q) and the rat exon (−1) (M83142) sequences showed only very poor sequence similarity (data not shown). The rat exon (−1) sequence was previously compared with the human exon (Y) by Wang et al. (1993) who noted that the sequence similarity was not dramatic. A comparable result was reported by Mercier et al. (1999), where bovine exon (−1) was also included in the alignment. The Spidey (Wheelan et al. 2001) alignments of rat exon (−1), human exon (Y) and bovine exon (−1) to the mouse *ST6Gal I* genomic sequence revealed that human exon (Y) and bovine exon (−1) are derived from the *ST6Gal I* genomic region corresponding to mouse exon (Q) but that the rat exon (−1) sequence corresponds to an area of the mouse *ST6Gal I* genomic sequence that is actually the downstream of mouse exon (Q) and does not directly align with the mouse exon (Q) sequence (Figure 3). Mouse exon (Q) is homologous to human exon (Y) and bovine exon (−1) (Figure 2A).

Interestingly, mouse exon (H), which is exclusively expressed in the mouse liver (Hu et al. 1997; Dalziel et al. 1999), is highly conserved in the rat genomic sequence, despite evidence that rat does not express this 5′UTR exon (Wang et al. 1990; Shah et al. 1992; Vertino-Bell et al. 1994). There is also a very high level of sequence conservation between the mouse and the rat in the region of mouse exon (L) (Figure 4A), which is exclusively expressed in the lactating mouse mammary gland (Dalziel et al. 2001). Mouse exon (L) appears to be only moderately conserved in the human *ST6Gal I* genomic sequence (Figure 4C) and even less so in the bovine (Figure 4B).

5′RACE of ST6Gal I *transcripts in bovine and human mammary gland*

Murine *ST6Gal I* expression is radically increased in the mammary gland during lactation (Dalziel et al. 2001). To

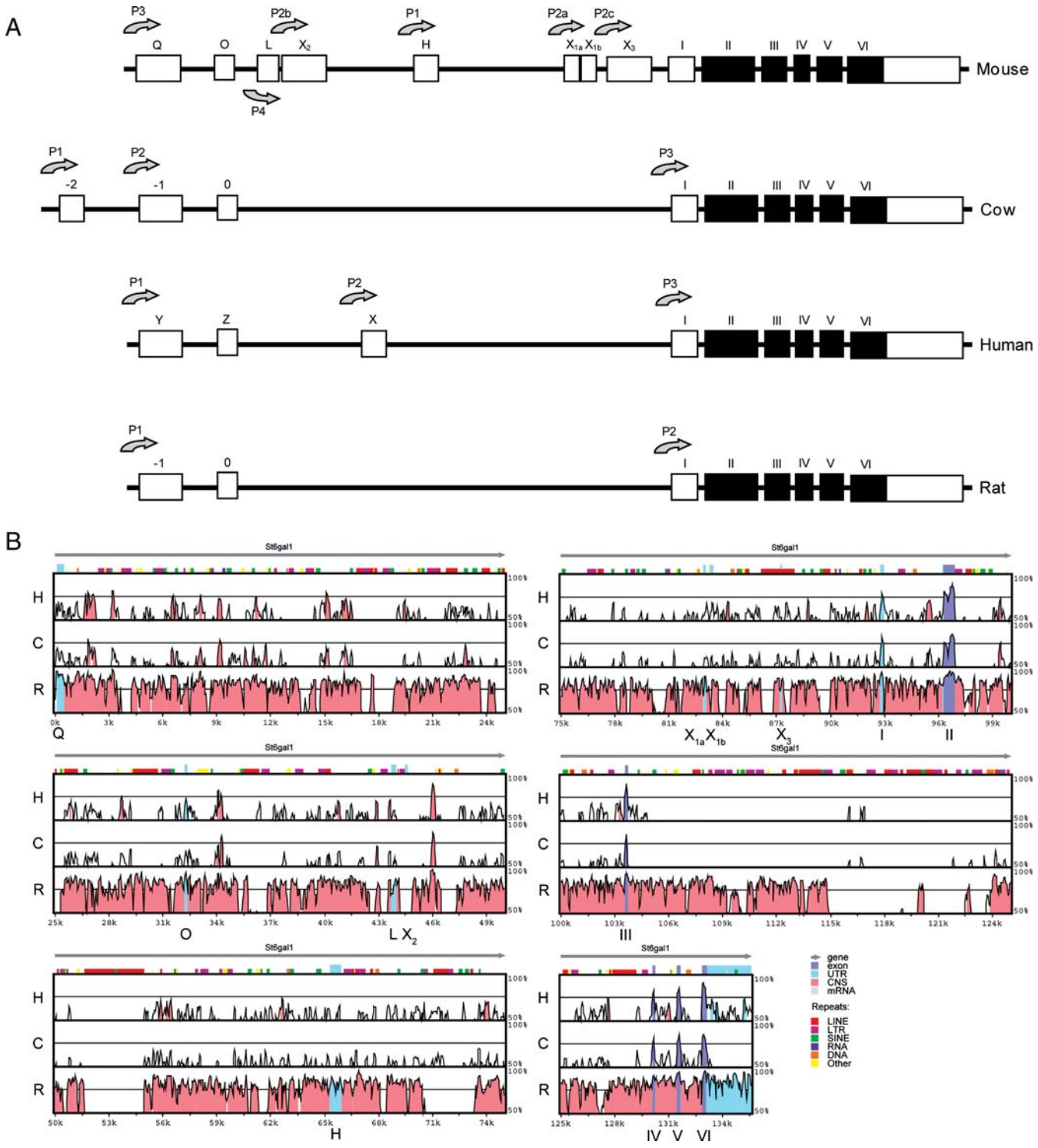


Fig. 1. (**A**) Schematic representation of the mouse, cow, human and rat *ST6Gal I* genomic regions (not to scale). Translated exons are shown in black and untranslated exons are in white. Putative promoter locations are denoted by arrows. (**B**) Vista conservation plot of the *ST6Gal I* genomic sequence, including ~100 kb 5′UTR. The mouse *ST6Gal I* genomic sequence was used as the base sequence for alignment with the human (H), cow (C) and rat (R) *ST6Gal I* sequence. The sequences were submitted via mVista and aligned using the LAGAN algorithm for the global multiple alignment of finished sequences, with the option for translated anchoring selected to improve the alignment of distant homologs. All sequences were repeat-masked using species-appropriate masking. The plot represents percent conservation between the aligned sequences at a given coordinate on the base sequence, which is calculated as a windowed-average identity score for the alignment (default parameters were used for calculation). The top (100%) and the bottom (50%) percentage bounds are shown to the right of every row. Regions of high conservation (>70%) are colored as follows: exons (dark blue), UTRs (light blue) or noncoding (pink).

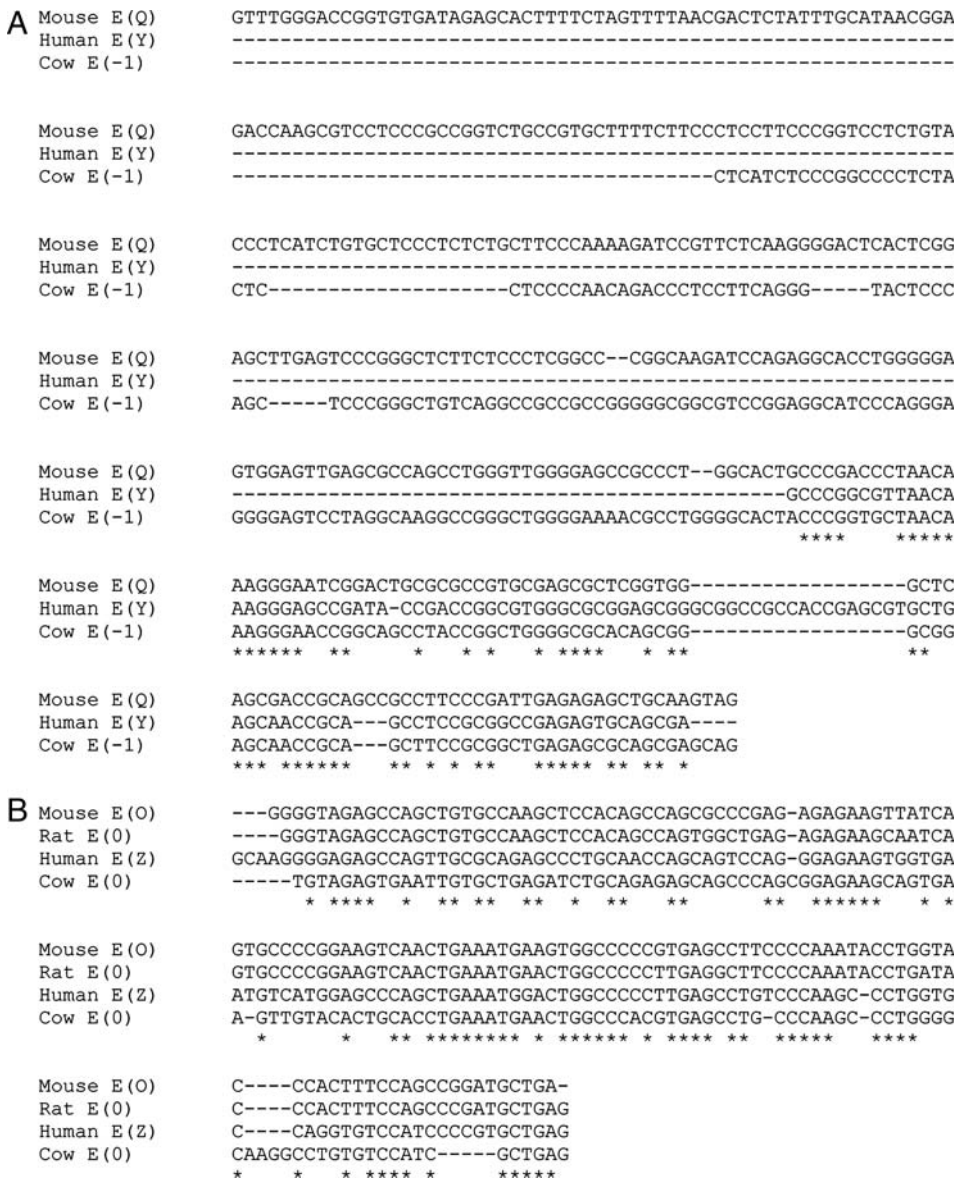


Fig. 2. MUSCLE multiple sequence alignments of *ST6Gal I* 5′UTR exons from corresponding transcript families in mammalian species. Dashed lines denote gap in the nucleotide sequence. Stars below sequence alignment indicate DNA identity. (**A**) Alignment of mouse exon (Q) (AF153680), human exon (Y) (NM_173216) and bovine exon (−1) (AJ006831). (**B**) Alignment of mouse exon (O) (AF153681), rat exon (0) (M83142), human exon (Z) (NM_173216) and cow exon (0) (AJ006832).

investigate whether this phenomenon occurs in other mammals, we devised a 5′RACE strategy to assess the nature of the *ST6Gal I* isoforms expressed in the mammary gland of cows and humans.

We prepared cDNA by reverse transcription of RNA from the lactating mammary tissue of a day 34 postpartum Holstein–Friesian cow using bST1-SP1, an antisense primer designed to anneal to a region in *ST6Gal I* exon (II) (Figure 5A). The 5′ ends of the synthesized cDNA were then amplified using a nested primer, bST1-SP2, which is complementary to a segment of exon (I). A further amplification was carried out with a second nested primer, bST1-SP3, which binds to a region of exon (I) further 5′ than bST1-SP2. The major product, visualized on an agarose gel by SYBR Safe staining, migrated at ∼250 bp (Figure 6). RNA extracted from the mammary glands of a pregnant cow at 8 days preparturition and a cow 5 days into involution was subjected to the identical 5′RACE procedure. The polymerase chain reaction (PCR) products were once again found to migrate at ∼250 bp (Figure 6). The products visible on the gel (Figure 6) were subsequently excised, purified and cloned.

Ten clones, from each stage of lactation, containing the SP3-derived PCR products were randomly selected for sequencing. All clones were confirmed to represent genuine *ST6Gal I* sequence as substantiated by the presence of *ST6Gal I* exon (I) sequence. Although all the clones contained >200 bp of sequence upstream of exon (I), the sequence was not found to represent a homologous bovine

Fig. 3. Pairwise alignment of mouse and rat *ST6Gal I* genomic sequence in the vicinity of mouse 5′UTR exon (Q), generated by Vista's LAGAN algorithm. The mouse exon (Q) sequence is highlighted in purple. Using Spidey, the rat and bovine exon (−1) and human exon (Y) sequences were mapped to the alignment. The rat exon (−1) sequence is highlighted in orange. Bovine exon (−1) maps to the mouse sequence denoted by lowercase and human exon (Y) maps to the underlined mouse sequence.

form of murine exon (L) at any of the time points investigated. All the 5′ sequences, at every stage of lactation, were found to contain bovine *ST6Gal I* exons (−1) and (0), representing the (−1 + 0) transcript family that was previously characterized by Mercier et al. (1999).

RNA was also reverse-transcribed from cells extracted from the milk of a lactating woman collected at 49 days following birth. Antisense primer HST-P4, complementary to a section of human *ST6Gal I* exon (II), was used for reverse transcription. Nested primer HST-P3, designed to anneal in exon (I), was used for the first PCR amplification and second nested primer HST-P5, which binds further 5′ in exon (I), for the final amplification (Figure 5B). The major products were visualized as previously described and found to migrate at ∼350, 320, 190 and 150 bp.

All four of the products visible on the gel (Figure 7) were cloned. Five clones representing each HST-P5-derived product were randomly selected for subsequent sequencing. As with

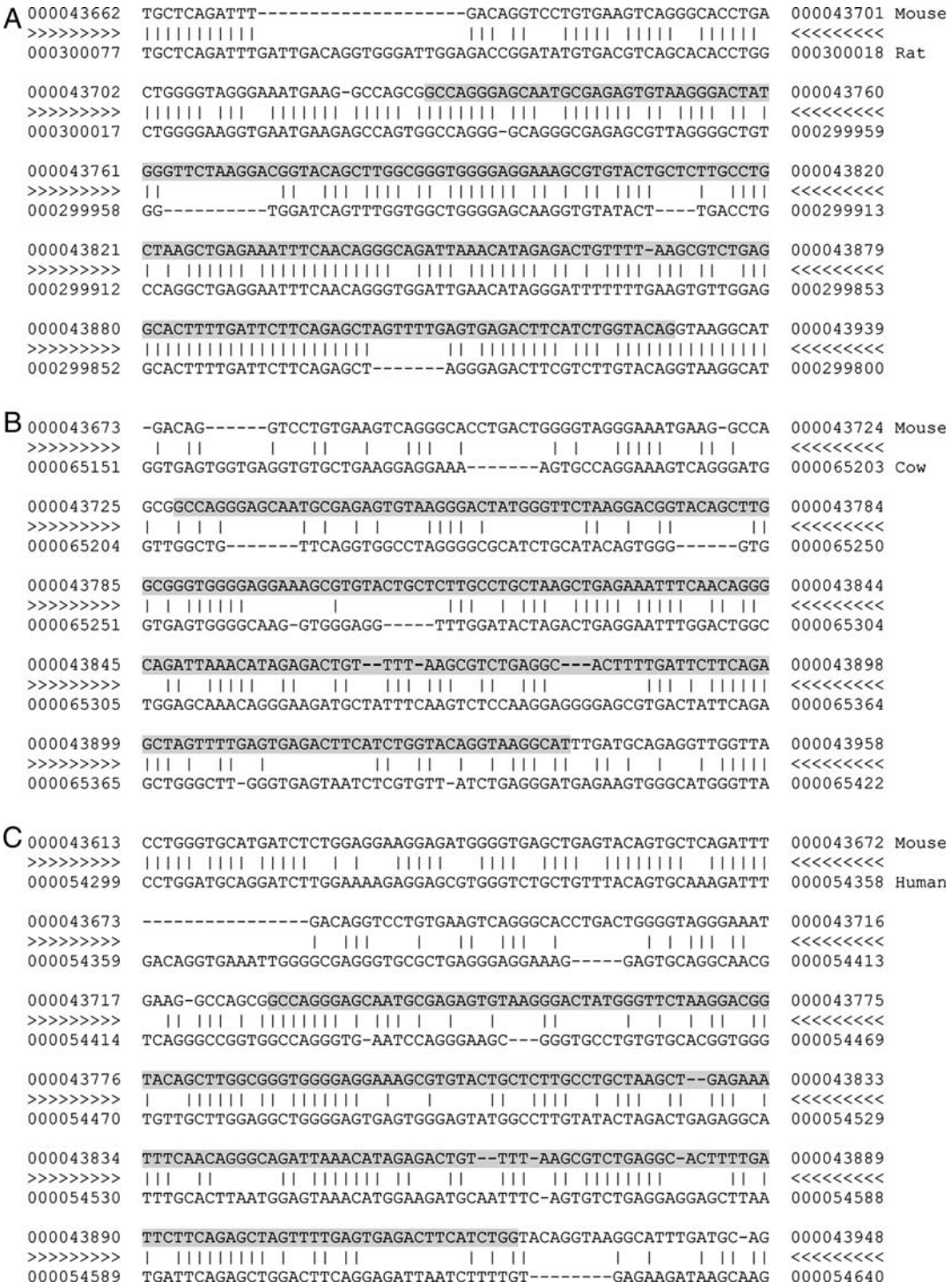


Fig. 4. Pairwise alignments of *ST6Gal I* genomic sequence in the vicinity of mouse 5′UTR exon (L), generated by Vista's LAGAN algorithm. The position of exon (L) in each pairwise alignment is highlighted in gray. (**A**) Mouse vs. rat pairwise alignment. (**B**) Mouse vs. cow pairwise alignment. (**C**) Mouse vs. human pairwise alignment.

the bovine sequences, none of the human 5′ segments were found to be homologous to murine exon (L). Furthermore, the two largest products contained human *ST6Gal I* exons (Y) and (Z). The smaller products were incomplete extensions of the same exon (Y + Z)-containing transcript.

ST6Gal I *gene expression in the bovine mammary gland*

The expression of the *ST6Gal I* gene in the bovine mammary gland was assessed during pregnancy, lactation and involution. Mammary gland samples, obtained by surgical biopsy from the four Holstein–Friesian cows, at each of the three

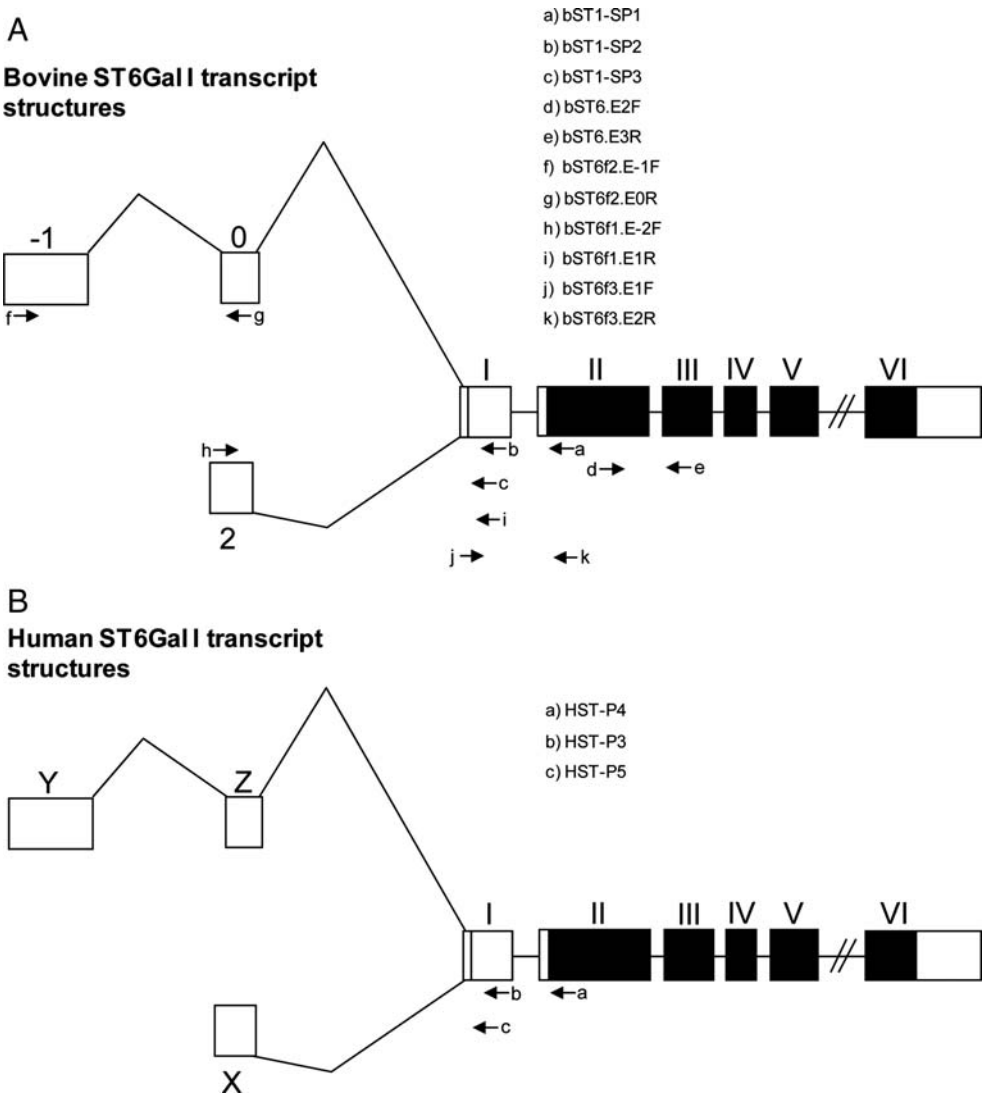


Fig. 5. Schematic representations of bovine and human *ST6Gal I* transcript families. Translated exons are shown in black and untranslated exons in white. Primer locations are denoted by arrows. (**A**) Bovine *ST6Gal I* transcript structure. (**B**) Human *ST6Gal I* transcript structure.

time points, were used for gene expression analysis. The expression level of *ST6Gal I* mRNA was determined for each sample using a quantitative PCR (qPCR) approach. To allow for the amplification of all *ST6Gal I* transcript families of the *ST6Gal I* gene, primers bST6.E2F and bST6.E3R were designed to anneal to regions in exons (II) and (III), respectively (Figure 5A). Figure 8 shows the mean normalized expression of the *ST6Gal I* gene in the mammary glands of four cows at the three stages of lactation sampled. Our analysis revealed a trend of increasing *ST6Gal I* gene expression in the bovine mammary gland between pregnancy and lactation, and lactation and involution. A statistically significant difference in *ST6Gal I* gene expression ($P < 0.05$) was found between pregnancy and involution.

ST6Gal I *gene expression in bovine tissues*

The *ST6Gal I* gene has been shown to be more highly expressed in the mouse liver than other tissues (Dalziel et al. 2001). An examination of *ST6Gal I* gene expression in bovine tissues showed that the kidney expressed higher levels of *ST6Gal I* mRNA than the brain, heart, lung or spleen (Mercier et al. 1999). We investigated the relative expression of the *ST6Gal I* gene in the bovine liver, lactating mammary gland, brain stem, hind brain, lymph node, rumen, abomasum and reticulum by qPCR. Tissues were harvested from a lactating dairy cow that was sacrificed at 34 days postpartum. The bST6.E2F and bST6.E3R primers (Figure 5A) were used to detect the overall expression of the *ST6Gal I* gene. To discriminate between the individual *ST6Gal I* transcript families, primers were designed to the unique 5′UTR exons of each bovine *ST6Gal I* isoform; for transcript family 1, bST6f1.E-2F in exon (−2) and bST6f1.E1R in exon (I), for transcript family 2, bST6f2.E-1F in exon (−1) and bST6f2.E0R in exon (0), and for transcript family 3, bST6f3.E1F in the extended portion of exon (I) and bST6f3.E2R in exon (II) (Figure 5A).

ST6Gal I gene expression was significantly higher in the liver than any other tissue examined (Figure 9A). The lymph node demonstrated the next highest level of expression,

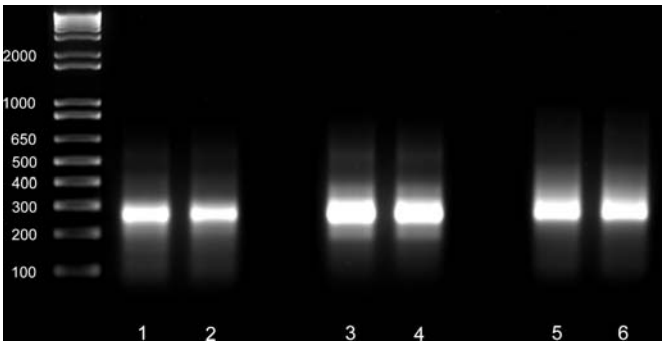


Fig. 6. A 5′RACE analysis of *ST6Gal I* mRNA expressed in the bovine mammary gland. RNA was obtained from mammary gland tissue collected from a single animal by surgical biopsy at three stages of lactation. Lanes 1 and 2: 5′RACE analysis of mRNA obtained from mammary gland tissue 8 days prepartum, generating predominantly a ∼250 bp product. Lanes 3 and 4: 5′RACE analysis of mRNA obtained from mammary gland tissue 33 days postpartum, generating predominantly a ∼250 bp product. Lanes 5 and 6: 5′ RACE analysis of mRNA obtained from mammary gland tissue 5 days into involution, generating predominantly a ∼250 bp product. A portion of exon (I), along with exons (−1 + 0), representing transcript family 2, was identified by sequencing of the major product at all three stages.

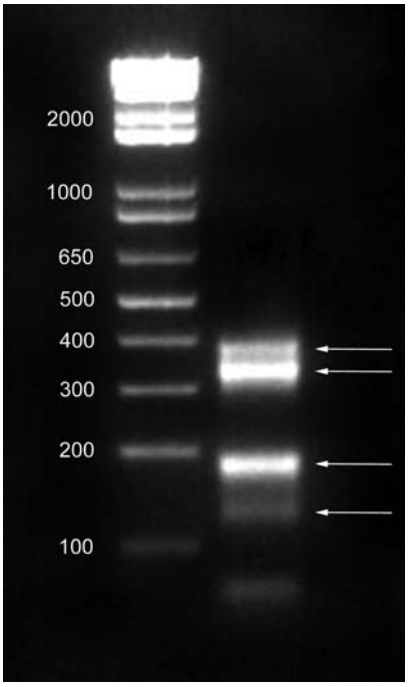


Fig. 7. 5′RACE analysis of *ST6Gal I* mRNA expressed in the human mammary gland. RNA was obtained from cells extracted from milk collected from a woman at day 48 of lactation. The major products are indicated by arrows and migrate at ∼350, 320, 190 and 150 bp. All four of the products visible on the gel were sequenced. The two largest products represented human *ST6Gal I* exons (Y) and (Z). The smaller products were incomplete extensions of the same exon (Y + Z)-containing transcript.

relative to the liver, followed by the lactating mammary gland. Only very low levels of *ST6Gal I* gene expression were observed in the other tissues analyzed.

Using specific primers to discriminate between the individual bovine *ST6Gal I* transcript families, we were able to detect all three isoforms in all the tissues examined. Relative to the other tissues assessed, the lymph node expressed the highest levels of transcript families 1 and 3 (Figure 9B and D). Transcript family 2 was expressed at similar levels in the brain stem, hind brain and lymph node and at lower levels in the other tissues (Figure 9C). Transcript families 2 and 3 were the best represented isoforms in the lactating mammary gland (Figure 9C and D).

Transcription factor-binding site analysis

As homologous bovine exon (−1 + 0)-containing and human exon (Y + Z)-containing transcripts were found to be the most well represented isoforms in the mammary gland of both species, we investigated potential similarities and differences in transcription factor-binding sites (TFBS) of the putative promoter regions upstream of exons (−1) and (Y). Also, as we are primarily interested in *ST6Gal I* gene expression in the mammary gland, we focused our in silico promoter analysis on lactation-related TFBS, and TFBS involved in the acute-phase response (APR), which has been implicated in early mammary gland involution (Stein et al. 2004) and has been demonstrated to affect hepatic *ST6Gal I* expression (Jamieson 1992; Dalziel et al. 1999).

In the TFBS analysis, 2 kb of cow and human genomic sequence upstream of, and including, exons (−1) and (Y) was used. A transcription start site (TSS) was predicted the upstream of both exon (−1) in the cow sequence and exon (Y) in the human sequence (Figure 10). As shown in Figure 10, several lactation-related and APR TFBS were predicted to be conserved in this region. Two putative STAT5 motifs were identified, one of which is within 30 bp of one of the three predicted Oct-1 sites. Two potential C/EBPβ sites were also identified, in addition to possible binding sites for Sp1 and AP1.

Both the cow and the human sequences contained a putative glucocorticoid receptor (GR) binding site; however, its position was not found to be conserved in the alignment. A GR- C/EBPβ composite element was identified in the cow sequence at −736 from the predicted TSS, but was not present in the human sequence. Conversely, the human sequence contained a potential Sp1-PU.1 composite element at −149 from the predicted TSS, which was not detected in the cow sequence.

Discussion

The sequence, structure and organization of the *ST6Gal I* gene have been well characterized in several mammals (Svensson et al. 1990; Wang et al. 1990; Wen et al. 1992; Wang et al. 1993; Mercier et al. 1999), and although the coding region of the gene is highly conserved between mammals, there is remarkable divergence in the pattern of expression of the different transcript families in some species (Dall'Olio 2000). Despite their close evolutionary relationship, there are documented differences between rat and mouse *ST6Gal I* gene expression that are inconsistent with observed sequence similarity. Our analysis demonstrates that mouse

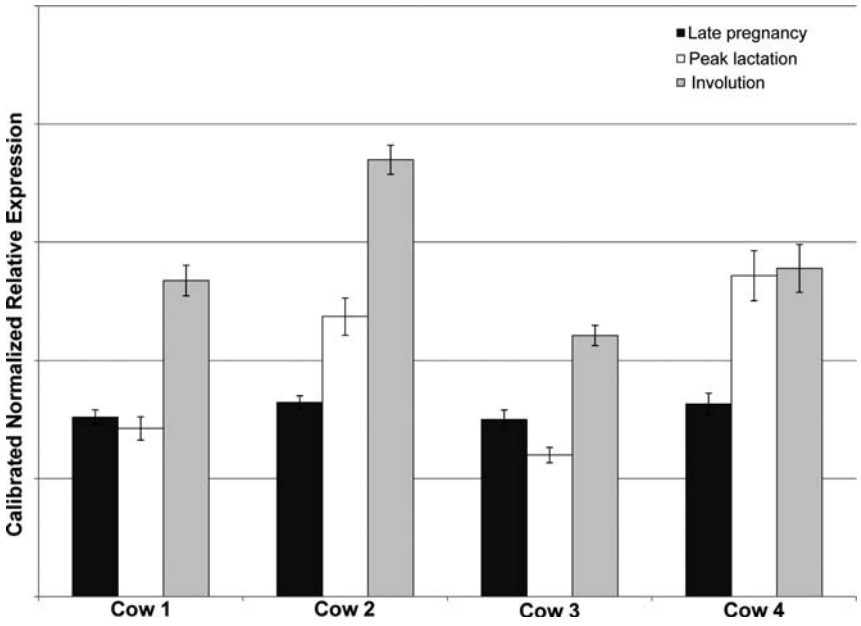


Fig. 8. Expression of *ST6Gal I* in the bovine mammary gland at three stages of lactation. Mammary gland samples were obtained from four different cows by surgical biopsy at three stages of lactation. Each cow was sampled during late pregnancy, ~8–23 days prepartum; peak lactation, ~34 days postpartum; and involution, ~5 days postweaning. Despite the increasing trend, there is no statistically significant difference in *ST6Gal I* gene expression in the bovine mammary gland between pregnancy and lactation, and lactation and involution; however, *ST6Gal I* gene expression does increase significantly between pregnancy and involution ($P < 0.05$).

exon (H) is very highly conserved in the rat genomic sequence, despite evidence that it is not expressed in rat tissues (Wang et al. 1990; Shah et al. 1992; Vertino-Bell et al. 1994). Furthermore, an *ST6Gal I* transcript resembling the isoform expressed by the rat liver has been observed in human HepG2 cells (Aasheim et al. 1993), suggesting that the liver-specific promoter upstream of exon (H) in mice may have evolved following their divergence from rats. The transcription of an additional 5′ exon in mice may allow for increased translational efficiency of that particular transcript (Kozak 1991a, 1991b).

We also report that, although their parent transcripts exhibit similar patterns of expression, rat exon (−1) and mouse exon (Q) are not direct homologs at the genomic sequence level. The published sequence for rat exon (−1) is transcribed from a region further 3′ of where mouse exon (Q), human exon (Y) and bovine exon (−1) align. This accounts for the observation by Wang et al. (1993) that although human exon (Z) and rat exon (0) were described as homologous, the sequence similarity between human exon (Y) and rat exon (−1) was not significant.

The concentration and the distribution of sialoglycoconjugates in milk have been extensively studied for a variety of mammals; however, little is known about the expression of sialyltransferase genes in the mammary gland during lactation. Dalziel et al. (2001) reported the dramatic induction of *ST6Gal I* gene expression in the mouse mammary gland during lactation. This was achieved by the recruitment of a previously undocumented 5′UTR exon, exon (L), presumably regulated by an otherwise silent, lactogenic promoter. Our conservation analysis of the *ST6Gal I* 5′UTR genomic sequence between the rat and the mouse reveals that exon (L) is very highly conserved in the rat sequence. This suggests that *ST6Gal I* gene expression in the rat mammary gland during lactation may also be selectively elevated via the transcription of an exon (L)-containing isoform. Such a mechanism would be consistent with the observations of Ip (1980), who documented an increase in the specific activity of rat *ST6Gal I* in the mammary gland throughout lactation, followed by a rapid return to basal values at the onset of involution. We suggest that rats and mice have evolved a mechanism to selectively up-regulate *ST6Gal I* in the lactating mammary gland to cater for the demand of α2,6-sialylated structures that serve as a delivery system for exogenous sialic acid. Rats and mice are born relatively immature and it has been shown that in comparison to adult levels, the liver of rat and guinea pig neonates has a diminished capacity to synthesize sialic acid (Gal et al. 1997; Nakano et al. 2001). Furthermore, Dickson and Messer (1978) demonstrated a significant positive correlation between intestinal sialidase activity of suckling rats and mice, and sialic acid content of milk at the corresponding stage of lactation. They proposed that the main function of intestinal sialidase in suckling mammals is to participate in the digestion of sialylated milk components, thereby providing an exogenous supply of sialic acid for the neonate.

On the basis of speculation by Dalziel et al. (2001) that *ST6Gal I* elevation in the lactating mouse mammary gland may be required to address the increased demand for the sialylation of type 2 disaccharides in milk, we investigated the presence of exon (L)-containing *ST6Gal I* transcripts in human and bovine mammary gland. Interestingly, an equivalent lactation-specific transcript was not detected in either species. During pregnancy, lactation and involution, we only

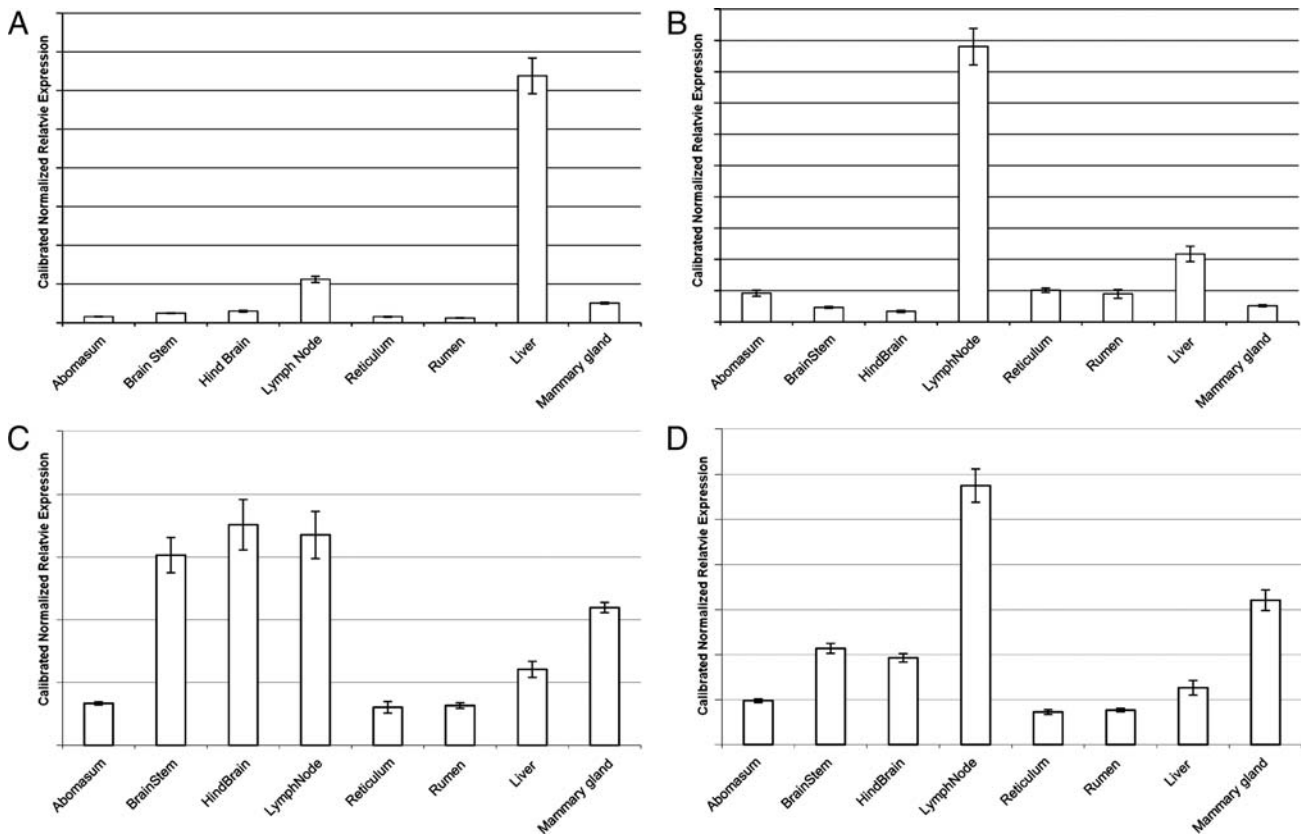


Fig. 9. Relative expression of the *ST6Gal I* gene and its individual transcript families in different bovine tissues, as determined by qPCR. As with Northern analysis, using different probes for each transcript family, relative levels between the different transcripts are not comparable by this approach. (**A**) Overall relative expression of the *ST6Gal I* gene. (**B**) Relative expression of *ST6Gal I* transcript family 1, previously documented in the brain and kidney. (**C**) Relative expression of *ST6Gal I* transcript family 2, previously observed in the lung and spleen. (**D**) Relative expression of *ST6Gal I* transcript family 3, which was previously identified as the main form present in the bovine tissues.

identified the (−1 + 0) transcript in the bovine mammary gland, which was previously described in the lung and spleen (Mercier et al. 1999). At day 48 of lactation, the human mammary gland expressed the (Y + Z) isoform, which is thought to be the constitutively expressed transcript (Wang et al. 1993). Despite the documented differences in tissue distribution, the bovine (−1 + 0) and human (Y + Z) transcripts are homologous (Mercier et al. 1999), exhibiting a high degree of sequence conservation in the 5′UTR. Our comparative analysis of the putative promoter regions upstream of bovine exon (−1) and human exon (Y) revealed several potentially conserved TFBS that are lactation-related. However, we also identified differences that may contribute to disparate regulation of this *ST6Gal I* transcript between species. We propose that the difference in regulation of the *ST6Gal I* gene in the lactating mammary gland observed between mice, humans and cows may be related to the unique nutritional and immune requirements of the neonate of each species. Furthermore, we suggest that differences in regulation between species are likely influenced by variation in the suite of TFBS present in the proximal promoter.

Our quantitative analysis of *ST6Gal I* gene expression in the mammary gland throughout the bovine lactation cycle revealed that *ST6Gal I* gene expression significantly increases by early involution. Involution of the bovine mammary gland occurs at a slower rate than in other mammals such as rodents (Capuco and Akers 1999). After weaning, epithelial cells in the rat and mouse mammary gland rapidly undergo apoptosis, completely removing all alveoli within 2 weeks (Walker et al. 1989). This is the first step in remodeling the mammary gland to a morphology closely resembling the tissue present in the mature virgin animal. Atabai et al. (2007) noted that, particularly during rapid forced weaning, mastitis is not a significant concern in rodents. In contrast, in ruminants such as the cow, active involution may take between 20 and 30 days (Holst et al. 1987). The bovine mammary gland is particularly susceptible to infection during early involution as large volumes of milk continue to accumulate in the gland for up to 4 days after weaning, creating an excellent growth medium for bacteria (O'Toole 1995). The innate immune system plays an integral role in the defense of the mammary gland from infection at this stage (Sordillo et al. 1997). For example, lactoferrin, a bacteriostatic iron-binding protein, is known to dramatically increase in concentration in bovine mammary secretions throughout involution (Rejman et al. 1989). We propose that an increase in *ST6Gal I* gene expression in the involuting bovine mammary gland may reflect an increased requirement for α2,6-sialylated structures to contribute to

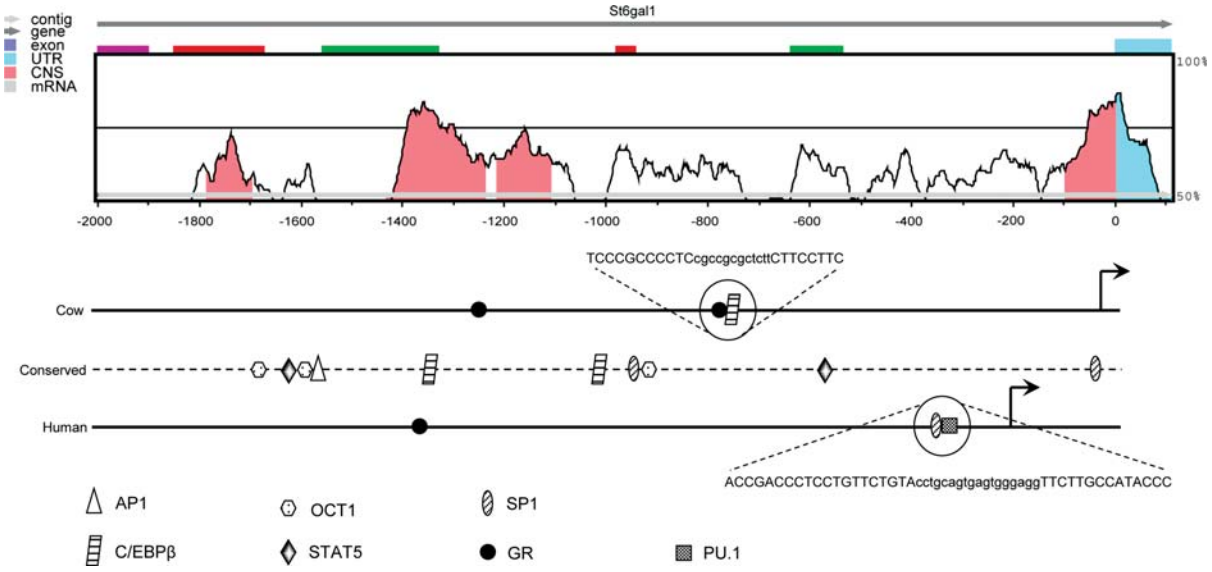


Fig. 10. Cross-species comparison of the putative *ST6Gal I* promoter region upstream of human exon (Y) and cow exon (−1). AVista conservation profile illustrates the level of similarity between human and cow sequences in the relevant region. A 2 kb human *ST6Gal I* genomic sequence upstream of exon (Y) was used for the alignment against a 2 kb cow sequence upstream of exon (−1). The top (100%) and the bottom (50%) conservation percentage bounds are shown on the right of the plot. Highly conserved (>70%) noncoding sequence is shown in pink, human untranslated exon (Y) is shown in blue. Aligned underneath the conservation profile is a schematic representation of the TFBSs identified in silico as potentially conserved between the two species or found to be present only in the cow, or only in the human sequence. TFBSs specifically relating to lactation or the APR were targeted. Potentially conserved TFBS positions were identified by rVista. Match (Biobase) was used to search for TFBSs in each sequence individually. MatrixCatch was utilized to locate potential composite elements in the human and the cow sequences. The composite elements identified are circled and their sequence is shown. The sequences corresponding to the positional weight matrix for each constituent TFBS are depicted in upper case, whereas the gap sequence is in lower case.

local innate immunity. Laporte et al. (2009) have demonstrated that, in a primary culture of mammary epithelial cells, bovine *ST6Gal II* expression increases in response to IL-6, which is known to be expressed during mastitis.

A role for milk sialoglycoconjugates in the innate immunity of the neonate has previously been suggested (Martin et al. 2001; Wang et al. 2001; Martin-Sosa et al. 2003; Wang and Brand-Miller 2003). The presence of α2,3-linked sialic acid on the host cell surface is known to be essential for the successful infection of some pathogens (Karlsson 1995), and it has been proposed that elevated levels of the α2,6-linkage may contribute to innate immunity by acting as "decoys" for invading pathogens (Gagneux and Varki 1999). It is conceivable, however speculative, that a similar mechanism may operate in the involuting bovine mammary gland during the period of milk stasis.

Increased expression of the *ST6Gal I* gene at the onset of bovine mammary gland involution may be driven by the expression of APR genes during the same period. As a result of their gene expression study of murine lactation, Stein et al. (2004) suggested a major role for the APR in the early stage of mouse mammary gland involution. The APR is a component of the innate immune system that provides an initial mechanism of defense against disease by minimizing tissue damage while promoting repair processes (Eckersall 2000). Injury or infection usually stimulates acute-phase protein expression in the liver by the release of cytokines (Dinarello 1984; Beutler 1986). Jamieson (1992) showed that elevation of the liver and serum *ST6Gal I* is one component of the hepatic APR. Dalziel et al. (1999) went on to show that in mice, transcription of a novel *ST6Gal I* mRNA isoform containing 5′UTR exon (H) was responsible for the acute-phase induction of hepatic *ST6Gal I*. Our analysis of the putative promoter region upstream of bovine *ST6Gal I* exon (−1) revealed the presence of a potential APR composite element, GR-C/EPBβ, which may account for the increased expression of *ST6Gal I* observed at the onset of involution. This combination of GR and C/EPB motifs in close proximity has been shown to be necessary for the maximal induction of the APR-reactant AGP gene (Klein et al. 1988; Williams et al. 1991; Ratajczak et al. 1992). Furthermore, Nishio et al. (1993) demonstrated that NF-IL6, a member of the C/EPB family (Akira et al. 1990), can directly interact with the GR and thus proposed that this interaction may be responsible for the synergistic activation of the rat AGP gene.

On the basis of our findings, we concur with previous suggestions that an exogenous supply of sialic acid may be necessary for the optimal development of mammalian neonates (Wang and Brand-Miller 2003), however, as Dickson and Messer (1978) proposed, we argue that the use of exogenous sialic acid may be limited to species that give birth to relatively immature young that initially have a diminished capacity for sialic acid synthesis (Gal et al. 1997; Nakano et al. 2001). As such, it would be particularly interesting to investigate the expression of various sialyltransferases in the mammary gland of monotremes and marsupials, whose young are born in an extremely immature state and whose milk contains significant quantities of sialic acid (Messer and Kerry 1973). Although we suggest that mammals which give birth to relatively well developed young may no longer have the

need for direct utilization of exogenously derived sialic acid, sialoglycoconjugates may still have a role in the innate immunity of the neonate, particularly in species such as humans where sialoglycoconjugates are a significant milk component. Additionally, we suggest that our observation of increasing *ST6Gal I* gene expression in the involuting mammary gland may indicate a role for α2,6-linked sialic acid during bovine mammary gland involution.

Materials and methods

Conservation analysis

Mouse, bovine and human genomic *ST6Gal I* sequences, with at least an additional 100 kb of 5′UTR, were obtained from Ensembl. The rat *ST6Gal I* genomic sequence was obtained from the Celera assembly of the *Rattus norvegicus* genome available through NCBI. Orthology of the sequences was confirmed using reciprocal best alignments and the examination of chromosomal context via the Ensembl and UCSC genome browsers. The annotation of the mouse *ST6Gal I* gene was acquired from Ensembl in the Vista format. Additional mouse 5′UTR exons were annotated by retrieving the sequences from GenBank: AF153680 (Q), AF153682 (X_2), U67989 (H), AF153684 (X_{1a} and X_{1b}) and AF153683 (X_3); and aligning them to the 5′UTR of the mouse genomic *ST6Gal I* sequence using the Spidey mRNA-to-genomic sequence alignment program.

Conservation analysis of the mouse, bovine, human and rat *ST6Gal I* sequences was performed using the Vista suite of tools for comparative genomics (Dubchak et al. 2000; Frazer et al. 2004). The sequences and annotations were submitted for analysis via the mVista interface. The LAGAN algorithm (Brudno et al. 2003) for the global multiple alignment of finished sequences was selected along with the option for translated anchoring to improve the alignment of distant homologs. All sequences were repeat-masked using species-appropriate masking. The resulting Vista plot was visualized using the Vista Browser with default parameters.

Collection of bovine mammary gland tissue and RNA extraction

The collection of the bovine mammary gland tissue used in this study was previously described in Sheehy et al. (2004). Briefly, the mammary gland tissue was collected from several Holstein–Friesian cows by surgical biopsy. Each cow was sampled during late pregnancy, ~8–23 days prepartum; peak lactation, ~34 days postpartum; and involution, ~5 days post-weaning. Mammary gland tissue from four different cows, at all three time-points, was used in this analysis.

Total RNA was isolated from ~100 mg of each of the 12 bovine mammary gland tissue samples using the RNeasy Lipid Tissue Mini Kit from Qiagen. The RNA samples were quantified using the Biolab NanoDrop and checked for integrity on agarose gels.

Collection of human milk samples and RNA extraction

Human milk was collected from a lactating female at 49 days following birth. Milk samples were centrifuged at 2000 g, 5 min at 4°C to pellet cells and milk was removed. Milk cells were resuspended in buffer 1 of the Qiagen RNeasy Micro Kit (Sydney, Australia) following the manufacturer's instructions for isolation of purified total RNA.

Collection of bovine tissues and RNA extraction

A panel of tissues was collected from a lactating dairy cow that was slaughtered 34 days following birth. Total RNA was isolated from liver, abomasum, reticulum, rumen, brain stem, hind brain, lymph node and mammary gland tissue samples using the RNeasy Lipid Tissue Mini Kit from Qiagen as per the manufacturer's instructions.

5′-Rapid amplification of cDNA ends

5′RACE analysis was performed using the Roche Applied Science 5′/3′ RACE Kit (Sydney, Australia), second generation, as per the manufacturer's instructions. Briefly, 2 μg of bovine mammary gland total RNA was annealed to the primer bST1-SP1 (5′-TAATT CCCTTTCTTCTTTTCTTTCC-3′) and reverse transcribed. bST1-SP1 was designed to anneal to a region in exon (II), such that genuine reverse transcription events of *ST6Gal I* mRNA must span at least the exon (I)–(II) boundary. The resultant cDNA was purified using the High Pure PCR Product Purification Kit from Roche as per the instructions specified in the RACE Kit. A homopolymeric A-tail was subsequently added to the 3′ end of the purified cDNA. A first PCR amplification of the dA-tailed cDNA was then performed using the Promega (Sydney, Australia) GoTaq Green Master Mix with the *ST6Gal I*-specific primer bST1-SP2 (5′-GATGTCTGTTTTACTGGGTCTGG-3′) and the Oligo(dT)-anchor primer provided with the RACE Kit. The PCR was carried out in an MJ Research PTC-225 Pertier Thermal-Cycler at 94°C for 2 min, followed by 10 cycles at 94°C for 15 s, 55°C for 30 s and 72°C for 40 s, then another 25 cycles at 94°C for 15 s, 55°C for 30 s and 72°C for 40 s with an additional 20 s of extension after each cycle, and finally 72°C for 7 min. A second PCR was then performed using the GoTaq Green Master Mix with the *ST6Gal I*-specific nested primer bST1-SP3 (5′-CCAAGGCTCATTC TTCTCAG G-3′) and the PCR anchor primer provided with the RACE Kit. The PCR parameters were 94°C for 2 min, followed by 10 cycles at 94°C for 15 s, 60°C for 30 s and 72°C for 40 s, then another 25 cycles at 94°C for 15 s, 60°C for 30 s and 72°C for 40 s with an additional 20 s of extension after each cycle, and finally 72°C for 7 min.

5′RACE was also performed on 2 μg of human mammary gland total RNA as described previously. Human *ST6Gal I* primers reported by Wang et al. (1993) were used in the analysis; HST-P4 (5′-AACTTGATGCCTGGTCC-3′) was used for reverse transcription, HST-P3 (5′-CTCTGGTTTGGCCT TGG-3′) for the first PCR amplification and the nested primer HST-P5 (5′-CTGCTTCTGGCTAATC-3′) for the second PCR amplification. The cycling conditions for the first PCR amplification were: 94°C for 2 min, then 10 cycles at 94°C for 15 s, 55°C for 30 s and 72°C for 40 s, followed by another 25 cycles at 94°C for 15 s, 55°C for 30 s and 72°C for 40 s with an additional 20 s of extension after each cycle, and finally 72°C for 7 min. The second PCR amplification was carried out at 94°C for 2 min, 10 cycles at 94°C for

15 s, 50°C for 30 s and 72°C for 40 s, followed by another 25 cycles at 94°C for 15 s, 50°C for 30 s and 72°C for 40 s with an additional 20 s of extension after each cycle and finally 72°C for 7 min.

All products of the 5′RACE second PCR amplification were separated by electrophoresis on 2% agarose gels and the major bands subsequently excised and purified using the Qiagen QIAquick Gel Extraction Kit. The purified PCR products were then cloned into the plasmid vector pGEM-T Easy (Promega) and sequenced.

Quantitative PCR

cDNA was generated from 2 µg of total RNA using the Invitrogen SuperScript VILO cDNA Synthesis Kit (Melbourne, Australia) as per the manufacturer's instructions. qPCR was performed according to the instructions for using the EXPRESS SYBR GreenER qPCR SuperMix with Premixed ROX supplied by Invitrogen. All qPCR primers were designed using Primer3 (Rozen and Skaletsky 2000) and checked for primer–dimer formation, self-annealing and hairpin formation using PerlPrimer (Marshall 2004) and OligoCalc (Kibbe 2007). To detect overall *ST6Gal I* expression, a forward primer bST6.E2F was designed in bovine *ST6Gal I* exon (II) (5′-GGTGTGCTGTGGTCTCTTCA-3′) and a reverse primer in exon (III) bST6.E3R (5′-CCCACGTCTTGTTGGAATTT-3′). To distinguish the expression profiles of individual *ST6Gal I* transcript families, primers were designed to the unique 5′UTR exons of each bovine *ST6Gal I* isoform: for transcript family 1 bST6f1.E-2F (5′-GGGTCTGCTCCTGATACCAC-3′) in exon (−2) and bST6f1.E1R (5′-CAAGGCTCATTCTTCTCA GGA-3′) in exon (I), for transcript family 2 bST6f2.E-1F (5′-CACTAC CCGGTGCTAACAAA-3′) in exon (−1) and bST6f2.E0R (5′-GCAGATCTCAGCACAATTCACT-3′) in exon (0), and for transcript family 3 bST6f3.E1F (5′-GCAGA CTTGTCTTA GCTGATGG-3′) in the extended portion of exon (I) and bST6f3.E2R (5′-CAAGGAAACCACGCTGT TCT-3′) in exon (II). All primers were deliberately designed over intron–exon boundaries to facilitate detection of any genomic DNA contamination. Primers reported by Bionaz and Loor (2007) for ribosomal protein subunits RPS9 [RPS9.192F (5′-CCTCGACC AAGAGCTGAAG-3′), RPS9. 254R (5′-CCTCCAGACCT CACGTTTGTTC-3′)] and RPS15 [RPS15.405F (5′-GCA GCTTATGAGCAAGGTCGT-3′), RPS15.555R (5′-GCTCATC AGCAGATAGCGCTT-3′)] and ubiquitously expressed transcript UXT (UXT.323F (5′-TGT GGCCCTTGGATATG GTT-3′), UXT.423R (5′-GGTTGTCGC TGAGCTCTGTG-3′)] were used as internal controls. For all reactions, the final reaction mixture of 20 µL consisted of 200 nM forward primer, 200 nM reverse primer, 2 µL of 1 in 20 diluted cDNA and 10 µL of the 2× EXPRESS SYBR GreenER qPCR SuperMix with Premixed ROX. PCR parameters for the Eppendorf RealPlex MasterCycler epgradients were 50°C for 2 min, 95°C for 10 min, 40 cycles at 95°C for 15 s then 60°C for 1 min, followed by a melt-curve dissociation cycle at 95°C for 15 s then 60°C for 1 min and a 20 min ramp up to 95°C for 15 s. The raw Ct values were imported into Biogazelle's qBasePlus software for calculation of calibrated normalized relative quantities (CNRQs) and their standard errors. The CNRQs were used in subsequent statistical analysis.

Statistical analysis

Significant differences in *ST6Gal I* gene expression in the bovine mammary gland between the stages of lactation were determined using the paired *t*-tests, assuming unequal variances. For all analyses, a *P*-value of <0.05 was interpreted as statistically significant.

TFBS analysis

For all TFBS analyses, 2 kb of human *ST6Gal I* genomic sequence upstream of, and including, exon (Y) and 2 kb of bovine sequence upstream of, and including, exon (−1) were used. All TFBS searches were restricted to lactation and APR-related transcription factors. Potentially conserved TFBS positions were identified by rVista (Loots et al. 2002), which makes TFBS predictions using the Match program based on TRANSFAC Professional 9.2, and determines which of the predicted sites are conserved between the species in the alignment. A core similarity value of 0.95 and matrix similarity value of 0.75 were used as cut off parameters for the TRANSFAC search. Match (Biobase) was used to search for TFBSs in each sequence individually, with default parameters for minimizing false positives. The MatrixCatch (Kel et al. 1999) program was used to locate potential composite elements in the human and cow sequences, also using its default parameters. This program uses a library of experimentally identified composite elements collected in the TRANSCOMPEL (Kel-Margoulis et al. 2002) database, as well as mononucleotide position weight matrices for individual TFBS collected in TRANSFAC. Transcription start sites were predicted using TSSG (Softberry).

Funding

This work was supported by the Victorian Department of Primary Industries.

Acknowledgments

We are grateful for the technical support provided by staff from the Discovery Technologies platform of the Biosciences Research Division at the Victorian Department of Primary Industries.

Conflict of interest

None declared.

Abbreviations

APR, acute-phase response; CNRQ, calibrated normalized relative quantities; GR, glucocorticoid receptor; PCR, polymerase chain reaction; qPCR, quantitative polymerase chain reaction; RACE, rapid amplification of cDNA ends; TFBS, transcription factor binding site; TSS, transcription start site; UTR, untranslated regions.

References

Aasheim HC, Aas-Eng DA, Deggerdal A, Blomhoff HK, Funderud S, Smeland EB. 1993. Cell-specific expression of human β-galactoside

α2,6-sialyltransferase transcripts differing in the 5′ untranslated region. *Eur J Biochem*. 213:467–475.
Akira S, Isshiki H, Sugita T, Tanabe O, Kinoshita S, Nishio Y, Nakajima T, Hirano T, Kishimoto T. 1990. A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family. *EMBO J*. 9:1897–1906.
Atabai K, Sheppard D, Werb Z. 2007. Roles of the innate immune system in mammary gland remodeling during involution. *J Mammary Gland Biol Neoplasia.* 12:37–45.
Bendiak B, Cook GMW. 1982. Kinetic parameters of a β-D-galactoside α2→6 sialyltransferase from embryonic chicken liver. *Eur J Biochem*. 128:355–362.
Beutler B. 1986. Cachectin/tumor necrosis factor: An endogenous mediator of shock and inflammation. *Immunol Res*. 5:281–293.
Bionaz M, Loor JJ. 2007. Identification of reference genes for quantitative real-time PCR in the bovine mammary gland during the lactation cycle. *Physiol Genomics.* 29:312–319.
Brudno M, Do CB, Cooper GM, Kim MF, Davydov E, Green ED, Sidow A, Batzoglou S. 2003. NISC comparative sequencing program. LAGAN and multi-LAGAN: Efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res*. 13:721–731.
Capuco AV, Akers RM. 1999. Mammary involution in dairy animals. *J Mammary Gland Biol Neoplasia.* 4:137–144.
Coppa GV, Gabrielli O, Giorgi P, Catassi C, Montanari MP, et al. 1990. Preliminary study of breastfeeding and bacterial adhesion to uroepithelial cells. *Lancet.* 335:569.
Dall'Olio F. 2000. The sialyl-α2,6-lactosaminyl-structure: Biosynthesis and functional role. *Glycoconj J*. 17:669–676.
Dalziel M, Huang RY, Dall'Olio F, Morris JR, Taylor-Papadimitriou J, Lau JTY. 2001. Mouse ST6Gal sialyltransferase gene expression during mammary gland lactation. *Glycobiology.* 11:407–412.
Dalziel M, Lemaire S, Ewing J, Kobayashi L, Lau JT. 1999. Hepatic acute phase induction of murine β-galactoside α2,6 sialyltransferase (ST6Gal I) is IL-6 dependent and mediated by elevation of Exon H-containing class of transcripts. *Glycobiology.* 9:1003–1008.
Dickson JJ, Messer M. 1978. Intestinal neuraminidase activity of suckling rats and other mammals - relationship to sialic-acid content of milk. *Biochem J*. 170:407–413.
Dinarello CA. 1984. Interleukin-1 and the pathogenesis of the acute-phase response. *N Engl J Med*. 311:1413–1418.
Dubchak I, Brudno M, Loots GG, Mayor C, Pachter L, Rubin EM, Frazer KA. 2000. Active conservation of noncoding sequences revealed by 3-way species comparisons. *Genome Res*. 10:1304.
Eckersall PD. 2000. Recent advances and future prospects for the use of acute phase proteins as markers of disease in animals. *Rev Med Vet*. 151:577–584.
Frazer K, Pachter L, Poliakov A, Rubin E, Dubchak I. 2004. VISTA: Computational tools for comparative genomics. *Nucleic Acids Res*. 32: W273–W279.
Gagneux P, Varki A. 1999. Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology.* 9:747–755.
Gal B, Ruano MJ, Puente R, GarciaPardo LA, Rueda R, Gil A, Hueso P. 1997. Developmental changes in UDP-N-acetylglucosamine 2-epimerase activity of rat and guinea-pig liver. *Comp Biochem Phys B*. 118:13–15.
Hanasaki K, Powell LD, Varki A. 1995. Binding of human plasma sialoglycoproteins by the B cell-specific lectin CD22. Selective recognition of immunoglobulin M and haptoglobin. *J Biol Chem*. 270:7543–7550.
Harduin-Lepers A, Vallejo-Ruiz V, Krzewinski-Recchi MA, Samyn-Petit B, Julien S, Delannoy P. 2001. The human sialyltransferase family. *Biochimie*. 83:727–737.
Holst BD, Hurley WL, Nelson DR. 1987. Involution of the bovine mammary gland: Histological and ultrastructural changes. *J Dairy Sci*. 70:935–944.
Hu YP, Dalziel M, Lau JT. 1997. Murine hepatic β-galactoside α2,6-sialyltransferase gene expression involves usage of a novel upstream exon region. *Glycoconj J*. 14:407–411.
Ip C. 1980. Changes in tissue and serum sialyltransferase activities as related to proliferation and involution of the rat mammary gland. *Biochim Biophys Acta*. 628:249–254.
Jamieson JC. 1992. The importance of N-linked glycoproteins and dolichyl phosphate synthesis for fusion of L6 myoblasts. *Biochem Cell Biol*. 70:408–412.
Karlsson KA. 1995. Microbial recognition of target-cell glycoconjugates. *Curr Opin Struct Biol*. 5:622–635.
Kel A, Kel-Margoulis O, Babenko V, Wingender E. 1999. Recognition of NFATp/AP-1 composite elements within genes induced upon the activation of immune cells. *J Mol Biol*. 288:353–376.
Kel-Margoulis OV, Kel AE, Reuter I, Deineko IV, Wingender E. 2002. TRANSCompel(R): A database on composite regulatory elements in eukaryotic genes. *Nucleic Acids Res*. 30:332–334.
Kibbe WA. 2007. OligoCalc: An online oligonucleotide properties calculator. *Nucleic Acids Res*. 35:W43–W46.
Klein ES, DiLorenzo D, Posseckert G, Beato M, Ringold GM. 1988. Sequences downstream of the glucocorticoid regulatory element mediate cycloheximide inhibition of steroid induced expression from the rat alpha 1-acid glycoprotein promoter: Evidence for a labile transcription factor. *Mol Endocrinol*. 2:1343–1351.
Kozak M. 1991a. An analysis of vertebrate mRNA sequences: Intimations of translational control. *J Cell Biol*. 115:887–903.
Kozak M. 1991b. Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J Biol Chem*. 266:19867–19870.
Laporte B, Gonzalez-Hilarion S, Maftah A, Petit J-M. 2009. The second bovine β-galactoside-α2,6-sialyltransferase (ST6Gal II): Genomic organization and stimulation of its in vitro expression by IL-6 in bovine mammary epithelial cells. *Glycobiology.* 19:1082–1093.
Loots GG, Ovcharenko I, Pachter L, Dubchak I, Rubin EM. 2002. rVista for comparative sequence-based discovery of functional transcription factor binding sites. *Genome Res.* 12:832–839.
Marshall O. 2004. PerlPrimer: Cross-platform, graphical primer design for standard, bisulphite and real-time PCR. *Bioinformatics*. 20:2471–2472.
Martin MJ, Martin-Sosa S, Garcia-Pardo LA, Hueso P. 2001. Distribution of Bovine Milk Sialoglycoconjugates During Lactation. *J Dairy Sci*. 84:995–1000.
Martin-Sosa S, Martin MJ, Garcia-Pardo LA, Hueso P. 2003. Sialyloligosaccharides in human and bovine milk and in infant formulas: Variations with the progression of lactation. *J Dairy Sci*. 86:52–59.
McVeagh P, Miller JB. 1997. Human milk oligosaccharides: Only the breast. *J Paediatr Child Health*. 33:281–286.
Mercier D, Wierinckx A, Oulmouden A, Gallet PF, Palcic MM, Harduin-Lepers A, Delannoy P, Petit JM, Leveziel H, Julien R. 1999. Molecular cloning, expression and exon/intron organization of the bovine β-galactoside α2,6-sialyltransferase gene. *Glycobiology.* 9:851–863.
Messer M, Kerry KR. 1973. Milk carbohydrates of echidna and platypus. *Science*. 180:201–203.
Miyagi T, Tsuiki S. 1982. Purification and characterization of β-galactoside α2,6-sialyltransferase from rat liver and hepatomas. *Eur J Biochem*. 126:253–261.
Mouricout M, Petit JM, Carias JR, Julien R. 1990. Glycoprotein glycans that inhibit adhesion of Escherichia coli mediated by K99 fimbriae: Treatment of experimental colibacillosis. *Infect Immun*. 58:98–106.
Nakano T, Sugawara M, Kawakami H. 2001. Sialic acid in human milk: Composition and functions. *Acta Paediatr Taiwan*. 42:11–17.
Nishio Y, Isshiki H, Kishimoto T, Akira S. 1993. A nuclear factor for interleukin-6 expression (NF-IL6) and the glucocorticoid receptor synergistically activate transcription of the rat α1-acid glycoprotein gene via direct protein-protein interaction. *Mol Cell Biol*. 13:1854–1862.
O'Toole DK. 1995. Factors inhibiting and stimulating bacterial growth in milk: A historical perspective. *Advances in Applied Microbiology*, Vol. 40. San Diego: Academic Press. p. 45–47.
Paulson JC, Beranek WE, Hill RL. 1977. Purification of a sialyltransferase from bovine colostrum by affinity chromatography on CDP-agarose. *J Biol Chem*. 252:2356–2362.
Ratajczak T, Williams PM, DiLorenzo D, Ringold GM. 1992. Multiple elements within the glucocorticoid regulatory unit of the rat α1-acid glycoprotein gene are recognition sites for C/EBP. *J Biol Chem*. 267:11111–11119.
Rejman JJ, Hurley WL, Bahr JM. 1989. Enzyme-linked immunosorbent assays of bovine lactoferrin and a 39-kilodalton protein found in mammary secretions during involution. *J Dairy Sci*. 72:555–560.
Rozen S, Skaletsky HJ. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, editors. *Bioinformatics Methods and Protocols: Methods in Molecular Biology.* Totowa (NJ): Humana Press. p. 365–386.
Shah S, Lance P, Smith TJ, Berenson CS, Cohen SA, Horvath PJ, Lau JT, Baumann H. 1992. *n*-Butyrate reduces the expression of β-galactoside α2,6- sialyltransferase in Hep G2 cells. *J Biol Chem*. 267:10652–10658.
Sheehy PA, Della-Vedova JJ, Nicholas KR, Wynn PC. 2004. Hormone-dependent milk protein gene expression in bovine mammary explants from biopsies at different stages of pregnancy. *J Dairy Res*. 71:135–140.

Sordillo LM, Shafer-Weaver K, DeRosa D. 1997. Immunobiology of the mammary gland. *J Dairy Sci*. 80:1851–1865.

Stein T, Morris J, Davies C, Weber-Hall S, Duffy M-A, Heath V, Bell A, Ferrier R, Sandilands G, Gusterson B. 2004. Involution of the mouse mammary gland is associated with an immune cascade and an acute-phase response, involving LBP, CD14 and STAT3. *Breast Cancer Res*. 6: R75–R91.

Sticher U, Gross HJ, Brossmer R. 1991. Purification and characterization of α2,6-sialyltransferase from human liver. *Glycoconj J*. 8:45–54.

Svensson EC, Soreghan B, Paulson JC. 1990. Organization of the β-galactoside α2,6-sialyltransferase gene. Evidence for the transcriptional regulation of terminal glycosylation. *J Biol Chem*. 265:20863–20868.

Varki A. 1993. Biological roles of oligosaccharides: All of the theories are correct. *Glycobiology*. 3:97–130.

Varki A. 2007. Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. *Nature*. 446:1023–1029.

Vertino-Bell A, Ren J, Black JD, Lau JTY. 1994. Developmental regulation of β-galactoside α2,6-sialyltransferase in small intestine epithelium. *Dev Biol*. 165:126–136.

Walker NI, Bennett RE, Kerr JFR. 1989. Cell death by apoptosis during involution of the lactating breast in mice and rats. *Am J Anat*. 185:19–32.

Wang B, Brand-Miller J. 2003. The role and potential of sialic acid in human nutrition. *Eur J Clin Nutr*. 57:1351–1369.

Wang B, Brand-Miller J, McVeagh P, Petocz P. 2001. Concentration and distribution of sialic acid in human milk and infant formulas. *Am J Clin Nutr*. 74:510–515.

Wang X, O'Hanlon TP, Young RF, Lau JTY. 1990. Rat β-galactoside α2,6-sialyltransferase genomic organization: Alternate promoters direct the synthesis of liver and kidney transcripts. *Glycobiology*. 1:25–31.

Wang X, Vertino A, Eddy RL, Byers MG, Jani-Sait SN, Shows TB, Lau JT. 1993. Chromosome mapping and organization of the human β-galactoside α2,6-sialyltransferase gene. Differential and cell-type specific usage of upstream exon sequences in B-lymphoblastoid cells. *J Biol Chem*. 268:4355–4361.

Weinstein J, de Souza-e-Silva U, Paulson JC. 1982a. Purification of a Gal β1 to 4GlcNAc α2 to 6 sialyltransferase and a Gal β1 to 3(4)GlcNAc α2 to 3 sialyltransferase to homogeneity from rat liver. *J Biol Chem*. 257:13835–13844.

Weinstein J, de Souza-e-Silva U, Paulson JC. 1982b. Sialylation of glycoprotein oligosaccharides N-linked to asparagine. Enzymatic characterization of a Gal β1 to 3(4)GlcNAc α2 to 3 sialyltransferase and a Gal β1 to 4GlcNAc α2 to 6 sialyltransferase from rat liver. *J Biol Chem*. 257:13845–13853.

Wen DX, Svensson EC, Paulson JC. 1992. Tissue-specific alternative splicing of the β-galactoside α2,6-sialyltransferase gene. *J Biol Chem*. 267: 2512–2518.

Wheelan SJ, Church DM, Ostell JM. 2001. Spidey: A tool for mRNA-to-genomic alignments. *Genome Res*. 11:1952–1957.

Williams P, Ratajczak T, Lee SC, Ringold GM. 1991. AGP/EBP(LAP) expressed in rat hepatoma cells interacts with multiple promoter sites and is necessary for maximal glucocorticoid induction of the rat α-1 acid glycoprotein gene. *Mol Cell Biol*. 11:4959–4965.

Wuensch SA, Huang RY, Ewing J, Liang X, Lau JT. 2000. Murine B cell differentiation is accompanied by programmed expression of multiple novel β-galactoside α2, 6-sialyltransferase mRNA forms. *Glycobiology*. 10:67–75.