

Psg22 expression in mouse trophoblast giant cells is associated with gene inversion and co-expression of antisense long non-coding RNAs

John M Williams^{1,2}, Melanie Ball¹, Andrew Ward² and Tom Moore¹

¹School of Biochemistry and Cell Biology, University College Cork, Western Gateway Building, Western Road, Cork, Ireland and ²Department of Biology and Biochemistry, University of Bath, Claverton Down, Bath BA2 7AY, UK

Correspondence should be addressed to T Moore; Email: t.moore@ucc.ie

Abstract

Pregnancy-specific glycoproteins (PSGs) are secreted carcinoembryonic antigen (CEA)-related cell adhesion molecules-related members of the immunoglobulin superfamily and are encoded by multigene families in species with haemochorial placentation. PSGs may be the most abundant trophoblast-derived proteins in human maternal blood in late pregnancy and there is evidence that dysregulation of PSG expression is associated with gestational pathology. PSGs are produced by syncytiotrophoblast in the human placenta and by trophoblast giant cells (TGCs) and spongiotrophoblast in rodents, and are implicated in immune regulation, angiogenesis and regulation of platelet function. PSGs are encoded by 17 genes in the mouse and ten genes in the human. While functions appear to be conserved, the typical protein domain organisation differs between species. We analysed the evolution of the mouse *Psg* genomic locus structure and report inversion of the *Psg22* gene within the locus. *Psg22* is the most abundant *Psg* transcript detected in the first half of mouse pregnancy and we identified antisense long non-coding RNA (lncRNA) transcripts adjacent to *Psg22* associated with an active local chromatin conformation. This suggests that an epigenetic regulatory mechanism may underpin high *Psg22* expression relative to the other *Psg* gene family members in TGCs.

Reproduction (2015) **149** 125–137

Introduction

Pregnancy-specific glycoproteins (PSGs) are members of a rapidly evolving multigene family (Chuong *et al.* 2010) and are the most abundant foetal proteins secreted into the maternal blood during pregnancy (Lin *et al.* 1974). Their concentration increases with gestational progression and reportedly reaches up to 200–400 µg/ml at term, far exceeding the concentrations of other placental hormones (Lin *et al.* 1974, Lee *et al.* 1979). PSGs are produced by syncytiotrophoblast (SyT) in higher primates and spongiotrophoblast (SpT) and trophoblast giant cells (TGCs) in rodents (Lei *et al.* 1992, Rebstock *et al.* 1993, Kromer *et al.* 1996). Mouse *Psg* genes are clustered on chromosome 7 in a region syntenic to human chromosome 19q13.2 (Teglund *et al.* 1994, Rettenberger *et al.* 1995). Similar to many placental hormones, PSGs are encoded by multigene families and there are ten human, 17 mouse and eight rat *PSG* genes (Rudert *et al.* 1989, www.ensembl.org). To date, PSGs have been found exclusively in species with haemochorial placentation, in which maternal blood is in direct contact with foetal trophoblast (Wu *et al.* 2008, Kammerer & Zimmermann 2010).

PSG proteins are members of the immunoglobulin (Ig) superfamily and are closely related to the predominantly membrane-anchored carcinoembryonic antigen-related cell adhesion molecules (CEACAMs; Kammerer & Zimmermann 2010). PSG and CEACAM family members possess a variable number of Ig variable (N)-like and Ig constant (C)-like domains and are heavily glycosylated (McLellan *et al.* 2005a). The Ig-like domain complements of human and rodent PSGs differ considerably. Typically, human PSGs contain one N and a variable number of C domains, whereas rodent *PSGs* typically have three or more N domains followed by a single C (subtype A) domain. All rat *PSGs*, with the exception of *PSG36* (N1-N2-N3-N4-N5-A), are of the N1-N2-N3-A domain arrangement, whereas the mouse *Psg* family has 14 members of the N1-N2-N3-A arrangement, and three (*Psg24*, *Psg30* and *Psg31*), which have an expanded structure created by duplication of N domains (McLellan *et al.* 2005b).

Notwithstanding the divergence of primate and rodent PSG protein domain complements, there is abundant evidence for conservation of function of PSGs across species. Human and mouse PSGs induce Th2 cytokine production in human and mouse monocytic and other

cells (Wessells *et al.* 2000, Snyder *et al.* 2001, Motrán *et al.* 2002, 2003, Waterhouse *et al.* 2002, Ellerman *et al.* 2003, Ha *et al.* 2005, 2008, 2010, Blois *et al.* 2012, Martínez *et al.* 2012), and human PSG1 activates TGF β 1 in a cell-free system and reduces inflammation in a mouse colitis model (Blois *et al.* 2014). Human and mouse PSGs induce tube formation by endothelial cells in matrigel, suggesting that PSGs may be pro-angiogenic *in vivo* (Lisboa *et al.* 2011, Blois *et al.* 2012). PSGs may also have an anti-thrombotic role in pregnancy because recombinant human and mouse PSGs bind platelet integrin α IIb β 3 and inhibit fibrinogen binding (Shanley *et al.* 2013). Immunomodulatory and pro-angiogenic functions of PSGs may be mediated by cell binding through a variety of receptors including the tetraspanin CD9 (Waterhouse *et al.* 2002, Ha *et al.* 2008) and heparan and chondroitin sulphates (Lisboa *et al.* 2011, Sulkowski *et al.* 2011, Blois *et al.* 2012).

Understanding the functions of PSGs and the clinical significance of their dysregulation during pregnancy is complicated by the multigenic nature of the family and evidence of extensive genomic variation at the individual level (Chang *et al.* 2013), and the lack of specific research reagents for individual family members. Nevertheless, there is suggestive evidence of alterations in PSG expression in pregnancies complicated by intrauterine growth restriction (IUGR) and preeclampsia (PE) (Pihl *et al.* 2009, Zhao *et al.* 2012, Blankley *et al.* 2013, reviewed in Moore & Dveksler (2014)). Several studies reported reduced global PSG expression in IUGR and PE (Pihl *et al.* 2009, Zhao *et al.* 2012), and a recent study, using proteomic analysis of maternal sera by mass spectrometry, has reported increased levels of PSG9 in PE (Blankley *et al.* 2013).

PSGs are expressed almost exclusively in trophoblastic tissues of the foetal component of the placenta, with no evidence of fine regulation of individual family members in different cell types or in response to physiological stimuli. However, there is some evidence for temporal modulation. *Psg* mRNA expression in the first half of mouse pregnancy is almost exclusively due to *Psg22* expression in TGCs (Wynne *et al.* 2006), and human PSGs are differentially expressed between the first trimester and the term (Shanley *et al.* 2013). PSGs may have both individual gene regulation and more global locus-wide regulation at the transcriptional level. Human *PSG* gene promoters have been analysed structurally and in reporter assays, and several transcription factors have been implicated including SP1 (Nores *et al.* 2004), Kruppel-like factors KLF4 and KLF6 (Blanchon *et al.* 2006, Racca *et al.* 2011), retinoic acid derivatives (López-Díaz *et al.* 2007, Camolotto *et al.* 2010) and ETS family members (Chamberlin *et al.* 1994, Camolotto *et al.* 2010). Many *PSG* gene promoters contain transcription factor-binding sites conserved among multiple family members, which may explain coordinated upregulation of PSGs in cultured cells undergoing replicative senescence

(Minagawa *et al.* 2005, Endoh *et al.* 2009), which may mimic terminal differentiation of trophoblast cells *in vivo*. Virtually nothing is known about transcriptional regulation of mouse *Psg* genes apart from the attenuation of *Ceacam* and *Psg* expression in Inositol-requiring enzyme 1a (*Ire1a*) and X-box-binding protein 1 (*Xbp1*) knockout mice, both of which are implicated in placental development. However, it is unclear whether these factors bind directly to *Psg* gene regulatory regions (Oikawa *et al.* 2010).

In order to achieve a deeper understanding of *PSG* gene evolution, expression and regulation, we completed the characterisation of mouse *PSG*-coding mRNA transcripts. This has identified regulated novel long non-coding RNA (lncRNA) transcripts at the mouse *Psg* locus, clarified evolutionary relationships between mouse and rat *PSG* genes and analysed the genomic context and regulation of the mouse *Psg22* gene. Our findings will facilitate rational gene targeting of the *Psg22* locus and may lead to increased understanding of regulation and function of this locus.

Materials and methods

Bioinformatics

All *PSG* sequences (genomic, coding sequences (CDS) and amino acid (aa)) were taken from public genome browsers: National Centre of Biotechnology Institute (NCBI) (<http://www.ncbi.nlm.nih.gov/>), University of California, Santa Cruz (UCSC) (<http://genome.ucsc.edu/>) and the Ensembl Genome browser (<http://www.ensembl.org/index.html>). Using these databases, a comprehensive list of all known rodent and human *PSGs* was compiled. We collated DNA sequences of all rodent *PSGs* and used sequence alignment tools to locate the mouse and rat *PSG* gene families at their respective loci. Sequence alignments were performed using online NCBI BLAST sequence alignment tool (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) and online ClustalW (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>) alignment software and, additionally, using the MEGA Molecular Evolutionary Genetics Analysis software MEGA5 (<http://www.megasoftware.net/>), we aligned individual *PSG* families for mouse, rat and human *PSG*-CDS and constructed Phylogenetic trees (neighbour-joined pairwise comparison phylogenetic trees). The evolutionary history was inferred using the Maximum Likelihood method based on the Tamura–Nei model (Tamura & Nei 1993). The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analysed (Felsenstein 1985). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches (Felsenstein 1985). All major branches yielded values of 95–100%. The scale bar represents 0.1 nucleotide substitutions per site.

All primers were designed using the online Primer-Blast Software (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>) unless otherwise stated. Primer analysis and the potential of secondary structures in primers were assessed using the online NetPrimer Software (<http://www.premierbiosoft.com/netprimer/>).

PCR primer sequences are listed in [Supplementary Table 1](#), see section on [supplementary data](#) given at the end of this article.

Materials

Chemicals were purchased from Sigma–Aldrich unless otherwise stated. Restriction enzymes were purchased from New England Biolabs (Ipswich, MA, USA). T4 DNA ligases were purchased from New England Biolabs. High Fidelity Phusion2 Hot Start Thermostable DNA Polymerase was purchased from ThermoScientific (Waltham, MA, USA). Plasmid DNA isolation, gel purification and nucleotide clean-up kits were purchased from Qiagen and used according to the manufacturer's instructions. Oligonucleotide primers for PCR were purchased from Eurofins MWG Operon (Eberberg, Germany). Bacterial medium constituents were purchased from Sigma–Aldrich. Perfectly Blunt Cloning kit and *Escherichia coli* bacterial strains used were purchased from Novagen (EMD Bioscience, Madison, MI, USA). Plastics and mammalian tissue culture materials were purchased from Starstedt (Wexford, Ireland). DNA ladders were purchased from New England Biolabs.

Mouse tissues

Mouse tissues were obtained from the Biological Services Unit, University College Cork. Mouse strains used were CD1, C57Bl6/J, 129/Sv. Embryonic (E) stage refers to the gestational age of the embryo. The morning on which the vaginal plug was found is counted as day 1 (E1) of gestation.

Cell culture

Trophoblast stem cell (TSC) lines TS-GFP, TS-EXE and TS-R26 were maintained as described previously (Tanaka 1998, Himeno *et al.* 2008). trophoblast stem (TS) were cultured in RPMI 1640 medium supplemented with 20% foetal bovine serum, 2 mM L-glutamine, 1 mM sodium pyruvate, 100 mM 2-mercaptoethanol, 50 U/ml penicillin and 50 µg/ml streptomycin. TS were seeded in T75 flasks and kept in an undifferentiated state using TS medium plus 70% mouse embryonic fibroblast conditioned medium (FCM) with 25 ng/ml FGF4 (Sigma–Aldrich) and 1 µg/ml heparin (Sigma–Aldrich). TGCs were differentiated from TS lines by culturing undifferentiated in TS medium without 70FCM+ FGF4/H for 6 days (Tanaka 1998, Hemberger *et al.* 2004, Senner & Hemberger 2010).

PCR, RT-PCR and qPCR

PCR and RT-PCR were performed using either Finnzymes Phusion Hot Start High Fidelity DNA Polymerase (ThermoScientific, F-530S) or Finnzymes Phusion 2 Hot Start High Fidelity DNA Polymerase (ThermoScientific, F-549S) on a G-Storm thermocycler (G-Storm, Somerton, Somerset, UK) in 50 µl reaction volumes. Reactions were carried out using 10 µl of 5× GC buffer, 1.2 µl dNTPs (New England Biolabs, N0447L), 1.5 µl (10 pmol) forward primer, 1.5 µl (10 pmol) reverse primer, 2 µl DMSO, 2 µl (50 mM) MgCl₂, 1 µl template and 0.5 µl (2 U/µl) Phusion DNA Polymerase, made up to 50 µl with ddH₂O. Cycling conditions were 98 °C for 3 min, and 35 cycles of 98 °C for 30 s, 60 °C

annealing for 40 s and 72 °C for 1 or 2 min, with a final extension of 72 °C for 10 min. The sizes of the PCR/RT-PCR products were confirmed by gel electrophoresis on a standard 1.5% agarose gel stained with Safeview (NBS Biologicals, Huntingdon, Cambs, UK) and visualised by exposure to u.v. light. Primer sequences of TSC differentiation marker genes: TSC marker (*Eomes*) F: 5'-TGATCATCACCAACAGGGC-3', R: 5'-ACTGTGTCTCTGA-GAAGGTG-3'; TGC marker (*Pl2*) F: 5'-TCCTTCTCTGGGCA-CTCCTGTT-3', R: 5'-CCATGAAGGCTTTTGA-AGCAAGATCA-3' and SpT marker (*Tpbpa*) F: 5'-TGAAGAGCTGAACCACTGGA-3', R: 5'-CAGGCAGTTCATATGTTGGG-3'.

qPCR was performed as described previously (Wynne *et al.* 2006), using the ABI PRISM 7900 sequence detection system (SDS) and the SYBR GREEN qPCR Kit (Applied Biosystems). RNA was extracted from cell lines and tissues using the RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. First-strand cDNA was synthesised using 1 µg total RNA in a 20 µl reaction volume using random hexamer priming and the High Capacity cDNA RT Kit (Applied Biosystems). The SYBR Green PCR Master Mix consists of Amplitaq Gold DNA polymerase, optimised PCR buffer, 25 mM MgCl₂, dNTP mix and AmpErase UNG. All qPCRs were performed in triplicate in MicroAmp Optical 384-Well Reaction Plates (Life Technologies) in a total volume of 10 µl. The following PCR protocol was used for all qPCRs: denaturation (95 °C for 10 min), amplification and quantification repeated for 40 cycles (95 °C for 15 s, 58 °C for 30 s and 60 °C for 1 min with a single fluorescence measurement), and generation of melting curve (60–95 °C with a heating rate of 1 °C/30 s and a continuous fluorescence measurement). Thereafter, PCR products were identified by generating a melting curve, which was also used to assess the occurrence of putative PCR artefacts (primer–dimers) or non-specific PCR products. Normalisation of expression levels to the housekeeping gene, hypoxanthine-guanine phosphoribosyltransferase (*Hprt*), was used to avoid discrepancies caused by variations in input RNA or in RT efficiencies. Results are described as mean *Psg* expression relative to mean *Hprt* expression. qPCR primer sequences are listed in [Supplementary Table 1](#).

Chromatin accessibility assay

Chromatin accessibility in specific genomic regions of the mouse *Psg* locus was measured using the EpiQ Chromatin Accessibility Assay Kit (Bio-Rad) according to the manufacturer's instructions. TSC lines (TS-R26 and TS-GFP) and their differentiated TGCs, MEFs and 3T3 cells were grown as described previously. *In situ* nuclease digestion was performed, cells were lysed and qPCR was performed using the Roche Lightcycler 480 system (Roche) according to the manufacturer's instructions. PCR primers used were designed according to the manufacturer's instructions and using Primer3 Software (<http://frodo.wi.mit.edu/>). PCR primer sequences are listed in [Supplementary Table 1](#). Primer efficiency was calculated using a serial dilution standard curve and efficiency was analysed using the EpiQ Chromatin Kit Data Analysis Tool software. The mouse reference (rhodopsin, *Rho*) and control (glyceraldehyde 3-phosphate dehydrogenase, *Gapdh*) gene primers used were supplied with the kit. Percentage chromatin

accessibility was quantified using the EpiQ Chromatin Kit Data Analysis Tool software supplied with the kit.

Polysome profile and sucrose gradients

TGC tissue (~20 mg) from E10 stage embryos from the CD1 mouse strain was dissected by low-power microscopy using a Nikon SMZ1500 microscope. Samples were pulverized with a mortar and pestle, precooled with liquid nitrogen, until a fine powder was obtained, adding more liquid nitrogen as necessary. Powder was lysed in 1 ml NP40 lysis buffer (20 mM Tris-HCl, pH 7.5, 250 mM NaCl, 15 mM MgCl₂, 20 mM dithiothreitol (DTT), 100 µg/ml cycloheximide, 0.5% Triton-X, 24 U/ml DNase, 20 U/ml Rnasin, 40 mM VRC and 1% NP40). Nuclei were then removed by microcentrifuging at 12 000 *g* for 10 s at 4 °C. Cytoplasmic extract was loaded onto 11 ml 10–60% sucrose gradients (10 and 60% m/v sucrose, 20 mM Tris-HCl, pH 7.5, 250 mM NaCl, 15 mM MgCl₂, 1 mM DTT and 100 µg/ml cycloheximide). Sucrose gradients were made as described previously (Stone 1973). Gradients were run for 3 h at 38 000 *g* at 4 °C in a Beckman Coulter SW41Ti Swinging bucket rotor in an ultracentrifuge with no brake applied. After centrifugation, 40×300 µl fractions were collected carefully from the top and stored at –80 °C. Total mRNA in each fraction was determined using A260/280 u.v. spectrometer. The 40 fractions were added to their neighbouring fraction to create 20 combined fractions, which were then used for RNA extraction (Prete *et al.* 2007). Fractions 20–40 were diluted with ultrapure H₂O to reduce the concentration of sucrose. Each fraction was supplemented with 30 µl of 0.5 M EDTA (pH 5.1), 30 µl of 10% SDS (to dissociate ribosomes) and 600 µl of phenol–chloroform–isoamyl alcohol mixture at acidic pH 5. Samples were vortexed and the upper aqueous phase was placed into a new tube supplemented with 60 µl 3 M NaOAc, pH 5.1, 2 µl GlycoBlue and 1 ml isopropanol. This was stored at –80 °C overnight. Fractions were thawed and microcentrifuged at 12 000 *g* for 15 min at 4 °C. The pellets were then washed with 80% EtOH, and the pellet was dried and dissolved in 50 µl H₂O. Purified RNA concentrations were determined using a u.v. spectrometer at an absorbance of A260 nm, and stored at –80 °C. RNA was used in cDNA synthesis and qPCR was used to determine which sucrose fractions contained *Psg22* transcripts. PCR primer sequences are listed in Supplementary Table 1.

Data and statistical analysis

All graphs were created using GraphPad Prism Software (GraphPad Software, Inc., La Jolla, CA, USA).

Results

Characterisation of mouse and rat *PSG* loci

Accession numbers of mouse and rat *PSG* genes, mRNA and protein sequences were collated from RefSeq (NCBI), Ensembl and UCSC genome browsers (Supplementary Table 2, see section on supplementary data given at the end of this article; mouse *Psg* gene synonyms and

GenBank accession numbers are given in Supplementary Table 3). Using these data and the BLAST program, each mouse or rat *Psg* mRNA sequence was aligned to ~2 Mb of genomic sequence from NCBI: M38:7: 17566974 to 19627308 of mouse chromosome (chr.) 7 or to ~1.3 Mb of genomic sequence from NCBI: RGSC3.4:1: 77301714 to 78604399 of rat chr. 1. From these data, we compiled updated genomic maps of mouse and rat *Psg* loci (Fig. 1A). These maps indicate that *Psg* locus structure is highly conserved between mouse and rat; the rat has eight *Psg* genes, two of which are flanked by the MHC I-like leukocyte 1 genes, *Mill1* and *Mill2*. In mouse, there are 11 genes in the *Mill1/Mill2*-flanked cluster, indicating significant gene duplication or loss between these species since a common ancestor. The *Mill1/Mill2*-flanked *Psg* genes in the mouse are all of the domain complement structure N1-N2-N3-A and, similar to the rat, are transcribed in the opposite orientation to *Psg* genes outside the *Mill1/Mill2* cluster. However, mouse *Psg22* is exceptional and is transcribed in the opposite orientation to other *Psg* genes within the *Mill1/Mill2*-flanked cluster. The inverted orientation of *Psg22* could be a C57Bl/6 strain-specific event; however, we found that a similar orientation is present in the 129/Sv strain because one of the boundaries of the cloned *Psg* locus region in PAC clone no. 647-D4 bisects the *Psg22* gene, and end-sequencing of this clone allowed us to establish the

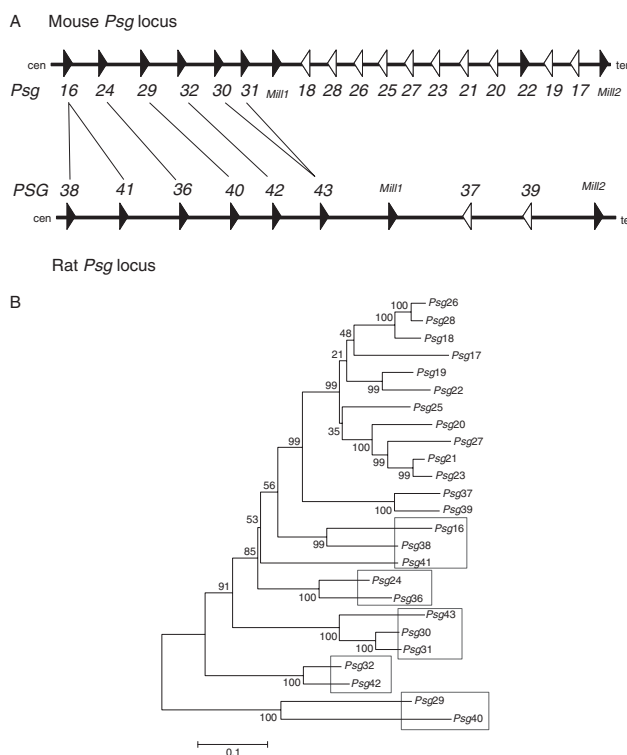


Figure 1 (A) Mouse and rat *Psg* loci genomic maps showing gene order and orthologous relationships. Gene transcriptional direction relative to centromere is indicated by arrowheads. (B) Phylogenetic tree of mouse and rat *Psg* CDS sequences. Scale bar represents 0.1 nucleotide substitutions per site.

orientation of *Psg22* relative to the rest of the locus (data not shown; PAC clone 647-D4 from RPCI-21 Female 129S6/SvEvTac Mouse PAC Library, BACPAC resources, Childrens Hospital Oakland Research Institute, USA).

There are no detectable orthologous relationships between rodent and primate *PSG* gene families. To determine whether such relationships are evident within rodent species, phylogenetic analysis was performed using full-length CDS sequences of mouse and rat *PSGs*. All mouse *Psgs* encoded by genes in the *Mill1/Mill2* cluster are located on one major branch of the neighbour-joining (N-J) tree (Fig. 1B). Similarly, in rat, *Mill1/Mill2*-flanked *Psg37* and *Psg39* branch together (Fig. 1B). A phylogenetic analysis using combined rat and mouse CDS sequences indicated strong orthology between six *PSGs* that map outside the *Mill1/Mill2*-flanked cluster (Fig. 1B). All of the major N-J tree branches yielded confidence values of 95–100%. Moreover, consistent with the sequence analysis, *PSG* gene order outside of the *Mill1/Mill2* cluster is conserved between mouse and rat (Fig. 1A).

To complete the description of the mouse *Psg* gene family, we characterised *Psg31* and *Psg32* which were previously reported as a pseudogene (LOC381852/Gm5155) and a hypothetical gene (*Psg-ps1*, *Cea6*), respectively, in NCBI databases (Beauchemin *et al.* 1999, McLellan *et al.* 2005b). Using the BLAST alignment programme, *Psg31* and *Psg32* sequences were aligned to a genomic contig as described above and their exonic structures were mapped (Fig. 2A). Expression and exonic structure of these genes were determined using a series of overlapping primer sets spanning the *Psg31* transcript, and a single primer set spanning the *Psg32* transcript (Fig. 2A and B). Using RT-PCR, we amplified these sequences from E15 placental cDNA and purified amplicons were subcloned into pSTblue-1 vector and transformed into NovaBlue Singles competent cells (Novagen, EMD Bioscience, Madison, MI, USA). Positive clones were identified by a diagnostic PCR screen of bacterial colonies and were sent in triplicate for sequencing, which confirmed expression of *Psg31* and *Psg32*. In addition, we deduced that *Psg31* has ten coding exons encoding the predicted largest mouse *PSG* protein with eight N domains and one A domain. *Psg32* structure is similar to the majority of mouse *Psgs*, with five exons encoding three protein N domains and one A domain (Supplementary Figure 1, see section on supplementary data given at the end of this article).

***Psg22* expression in TGCs is developmentally regulated and its mRNA is associated with polysomes**

Psg22 is by far the most abundant *Psg* transcript in TGCs in early pregnancy, although its expression is considerably lower than total *Psg* expression in the SpT later in pregnancy (Wynne *et al.* 2006). To analyse whether *Psg22* expression in TGCs is merely 'leaky' expression and possibly

functionally unimportant, we determined its expression and regulation using two TSC lines TS-EXE and TS-GFP, which were induced to differentiate into TGCs *in vitro*. We characterised the relative expression of different *Psg* genes using two degenerate primer sets that give unbiased amplification of all mouse *Psg* transcripts using RT-PCR, as described previously (Wynne *et al.* 2006). cDNA was synthesised from total RNA extracted from undifferentiated and differentiated TS line cultures. Purified amplicons were subcloned into pSTblue-1 vector and transformed into NovaBlue Singles competent cells (Novagen, EMD Bioscience, Madison, MI, USA). Positive clones were identified by a diagnostic PCR screen of bacterial colonies. Ten PCR clones arising from each primer set from both TS lines (in both undifferentiated and differentiated states) were sequenced (80 PCR clones in total). Undifferentiated TS lines expressed multiple *Psgs* with *Psg22* and *Psg27* predominant (Fig. 3A and B). However, following differentiation of TS lines to TGCs as confirmed by increased expression of TGC/SpT markers *Pl2* and *Tpbpa* (Supplementary Figure 2, see section on supplementary data given at the end of this article), *Psg22* is predominantly expressed (80% of total; Fig. 3C and D). We further found, using *Psg* gene-specific primers in a qPCR experiment, that TS differentiation was associated with strong upregulation of *Psg22* expression (greater than three orders of magnitude), with less dramatic changes in expression of *Psg19*, which is closely related to *Psg22*, and *Psg21* and *Psg23*, which are predominantly expressed in SpT (Fig. 4A, B, C and D). A similar pattern of high expression of *Psg22* relative to these three *Psg* genes in TGCs, and high expression of *Psg22* in TGCs relative to other placental tissues was found in a qPCR experiment using dissected placental tissues (Fig. 5A, B, C and D).

We used sucrose gradient fractionation to determine whether *Psg22* in TGCs is associated with ribosomes and therefore likely to be translated. TGC tissue was dissected from three E10 embryos and processed as described in 'Materials and methods' section. U.v. spectrometry results were plotted to produce a fractionated RNA profile. Free mRNPs (fractions 1–13), 80S RNA (fractions 16–20) and polysome-bound RNA (fractions 30–39), are evident (Fig. 6A). Fractions were combined (1+2...39+40), giving 20 in total, and fractions containing *Psg22* transcripts were determined using qPCR with *Psg22*-specific primers of cDNA made from RNA extracted from each pair of combined fractions. qPCR of the housekeeping gene, *Hprt*, was used as a control. Similar to *Hprt*, *Psg22* transcripts were enriched in fractions ~15–20, indicating significant association with ribosomes (Fig. 6B and C).

***LncRNA* expression and open chromatin conformation adjacent to *Psg22* is associated with differentiation of TGCs**

Psg22 is flanked by *Psg19* and *Psg20* within a cluster of 11 *Psg* genes. We hypothesised that inversion of *Psg22*

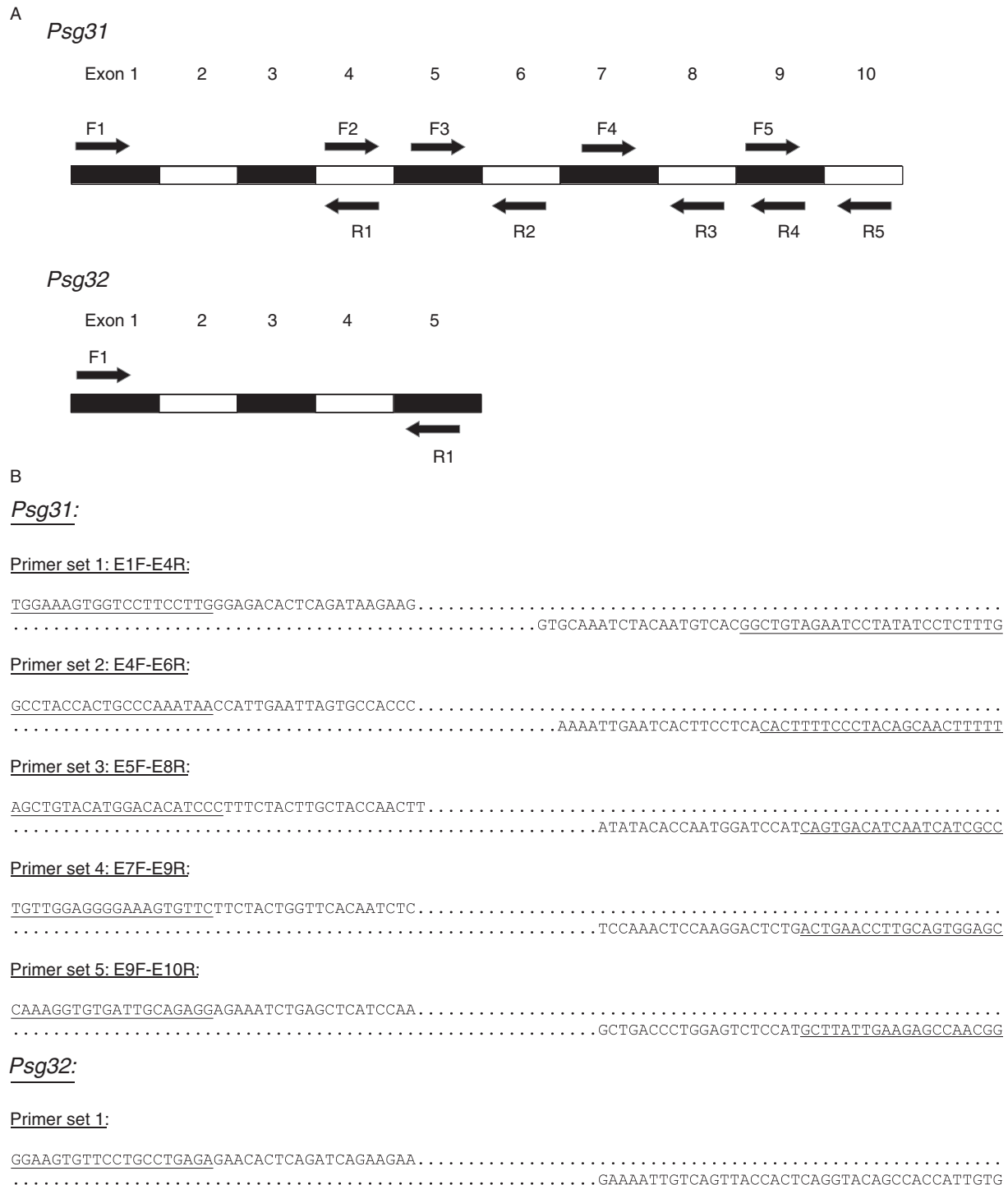


Figure 2 Structure of *Psg31* and *Psg32* mRNAs expressed in E15 CD1 mouse placenta. (A) Graphical representation of primer sites used to map *Psg31* and *Psg32* cDNAs. (B) PCR primers (underlined) and adjacent sequences derived from sequencing cloned PCR products.

on the chromosome relative to other *Psg* genes might explain its unique expression pattern in TGCs. A search of genome browsers identified a 417 bp expressed sequence tag (BY564540) ~5738 bp upstream of the *Psg22* transcriptional start site (TSS) and transcribed from the opposite strand (Fig. 7A). BLAST searching of the *Psg*

locus using the BY564540 sequence identified three regions of similarity (Fig. 7A). The highest (90%) match located ~9264 bp downstream of the *Psg22* TGA stop codon is hereafter named Blast1. Two other regions (Blast2 and Blast3) were 86 and 81% similar respectively (Fig. 7A). Blast1, downstream of *Psg22*, is in the same

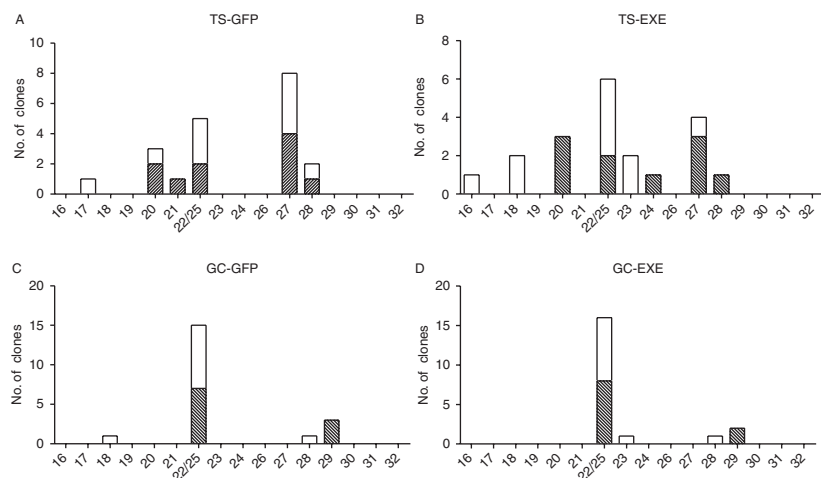


Figure 3 Relative expression of mouse *Psgs* in undifferentiated (TSC) and differentiated (TGC) trophoblast cell lines. (A) TS-GFP; (B) TS-EXE; (C) GC-GFP and (D) GC-EXE. RT-PCR performed using primer sets PSGF/R (hatched) and Psg-All2F/R (open) that give unbiased amplification of all mouse *Psgs*. Twenty cloned PCR products (ten from each primer set) were sequenced from each cell line and the results were combined. *Psg22* and *Psg25* results are combined because primer set PSGF/R does not discriminate between them; however, primer set Psg-All2F/R does, and all sequenced PCR clones were *Psg22*.

orientation as BY564540, whereas Blast2 and Blast3 are in the opposite orientation. Expression analysis of BY564540 using RT-PCR showed that it is expressed in differentiated extraembryonic lineages including EPC, TGCs and whole placenta at E13, E15 and E17, but not in undifferentiated ES cells or TSCs (Fig. 7B). Similarly, the three related regions (Blast1, Blast2 and Blast 3) are also expressed in differentiated, but not undifferentiated, TSCs (Fig. 7B). None of these four expressed sequences contains an open reading frame or microRNAs, as determined by ORF finder software at www.ncbi.nlm.nih.gov/projects/gorf/gorf.html, SMART Domain finder software at <http://smart.embl-heidelberg.de/> and scrutiny of www.mirbase.org.

We mapped the structure of the BY564540 and Blast1 transcripts to determine the extent of antisense transcription in the *Psg22* region using sets of overlapping PCR primers for RT-PCR of cDNA from E10 TGCs dissected from CD1 mice (Fig. 7C and Supplementary Table 1). Transcript lengths (assuming no splicing events) were estimated as the distance between the most 5' and 3' primers that amplified an expressed sequence and were 6148 bp (BY564540) and 6606 bp (Blast1) respectively. We attempted to use 5'-RACE PCR to precisely map the TSS of the transcripts, but were unsuccessful. However, we deduce from PCR walking experiments that BY564540 transcription initiates at the 5' UTR of *Psg22*, and Blast1 initiates <1500 bp from a U6 snRNA gene (Fig. 7D).

We used a proprietary EpiQ chromatin assay to analyse chromatin accessibility in the *Psg22* gene region to determine whether lncRNA expression is associated with enhanced chromatin accessibility. Chromatin preparations were made from TS lines TS-R26 and TS-GFP before and after differentiation to TGCs. Following nuclease digestion, PCR primers from regions of interest were used to assess the degree of chromatin accessibility using proprietary software. Primers used

spanned the *Psg22* gene TSS; the *Psg23* TSS; the BY564540 EST and Blast1 sequences, and regions 2 kb upstream and downstream of them (Fig. 8A). Percentage chromatin accessibilities (normalised to assay controls) are shown (Fig. 8B and C). Briefly, none of the regions exhibit open chromatin in undifferentiated TSC samples (Fig. 8B and C). However, following differentiation to TGCs, *Psg22* core promoter region; and BY564540 EST and Blast1 sequences and associated upstream, but not downstream, sequences all exhibit open chromatin (Fig. 8B and C).

Discussion

Our description of the mouse and rat *Psg* loci indicates that orthologous relationships are discernible between individual mouse and rat *Psg* genes, except within the

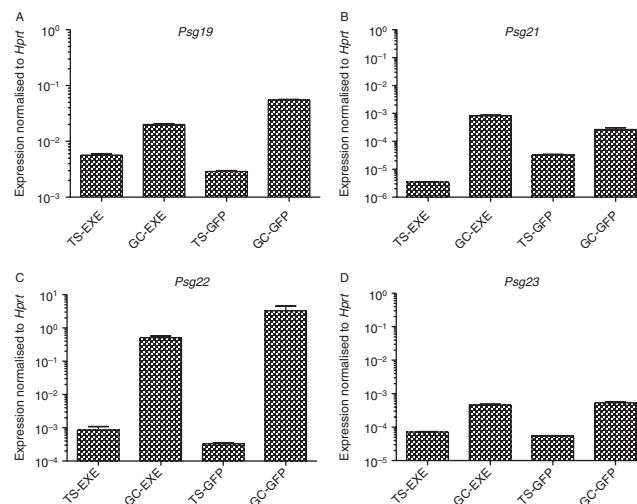


Figure 4 Expression of selected mouse *Psgs* in undifferentiated (TSC) and differentiated (TGC) cell lines measured by qRT-PCR, normalised to *Hprt* expression, and plotted on log scale. $n=3$.

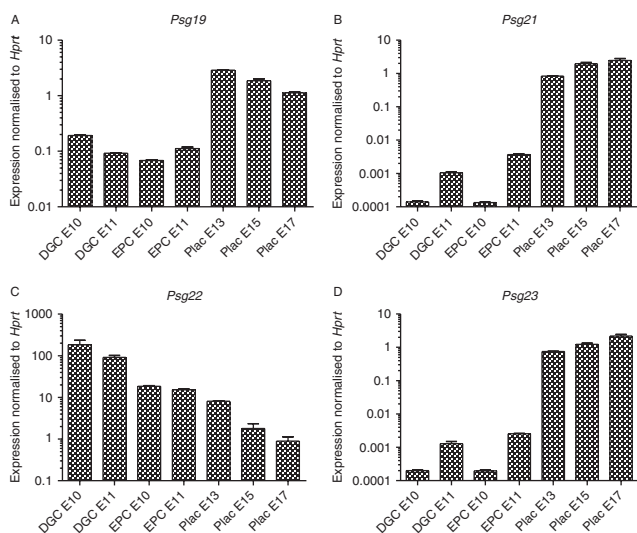


Figure 5 Expression of selected mouse *Psg*s in primary trophoblastic tissues. Dissected TGCs (DGC; E10 and E11), ectoplacental cone (EPC; E10 and E11), and whole placenta (Plac; E13, E15 and E17). *Psg* expression normalised to *Hprt* expression. $n=2$.

Mill1/Mill2-flanked cluster, where there has been considerable expansion of the mouse *Psg* gene complement. The reason for this expansion is unknown; however, it has been proposed that PSGs, similar to other placental hormones, may be subject to evolutionary selection pressures arising from maternal–foetal conflict, which could underpin gene duplication to increase gene dosage or to diversify the functionality of the locus (Haig 1993, Ball *et al.* 2004, McLellan *et al.* 2005b, Wynne *et al.* 2006, Shanley *et al.* 2013). Rapid evolution of PSGs is also evident in the lack of orthologous relationships between rodent and primate PSGs, the extreme individual variability of the human locus and the rapid evolution of human PSG-CDS (McLellan *et al.* 2005b, Kammerer & Zimmermann 2010, Chang *et al.* 2013). Interestingly, in a recent study of biomarkers of PE, PSG9, alone among human PSGs, was significantly upregulated, suggesting that individual PSGs may encode unique functions or undergo gene-specific regulation or pathological dysregulation (Blankley *et al.* 2013). PSG functional diversification could also arise from expansion of PSG protein domain number. While most rodent PSGs have a N1-N2-N3-A domain structure encoded by five exons, in this study, we found that *Psg31* has ten coding exons encoding the predicted largest mouse PSG protein comprising eight N domains and one A domain. However, it should be noted that functional studies to date have not reported differences between different mouse or human PSGs in multiple assays, suggesting that at least the currently known functions of PSGs are conserved across rodent and primate gene families (McLellan *et al.* 2005b, Shanley *et al.* 2013). Therefore, the diversification of PSG protein structure may suggest

other undetected functions for PSGs, or relaxed constraint arising from redundancy within the family. Understanding these aspects of PSG evolution will require detailed functional analysis of individual PSGs using mutagenesis in model organisms, or association of human PSG variants with pregnancy phenotypes or disorders.

We have previously demonstrated that total mouse *Psg* gene expression is markedly increased in the SpT compared with TGCs, and that the majority of *Psg* transcripts expressed in the TGCs in the first half of pregnancy are encoded by *Psg22* (Wynne *et al.* 2006). *Psg22* is reportedly proangiogenic and its expression in TGCs may promote vascular development in the placental bed or regulate the maternal vascular

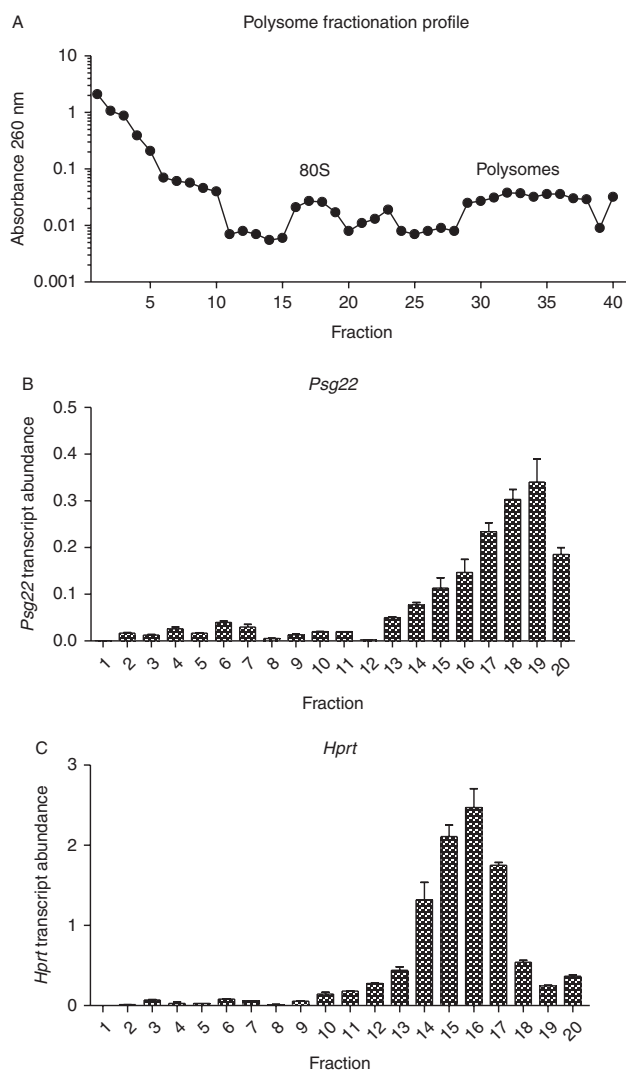


Figure 6 Association of *Psg22* mRNA with polysomes. For this experiment, 20 mg of lysed E10 dissected TGC tissue were used as a template. (A) A_{260 nm} readings of 40 sucrose gradient fractions. (B) *Psg22* transcript representation (qRT-PCR) in pooled fractions. (C) *Hprt* transcript representation in pooled fractions (positive control).

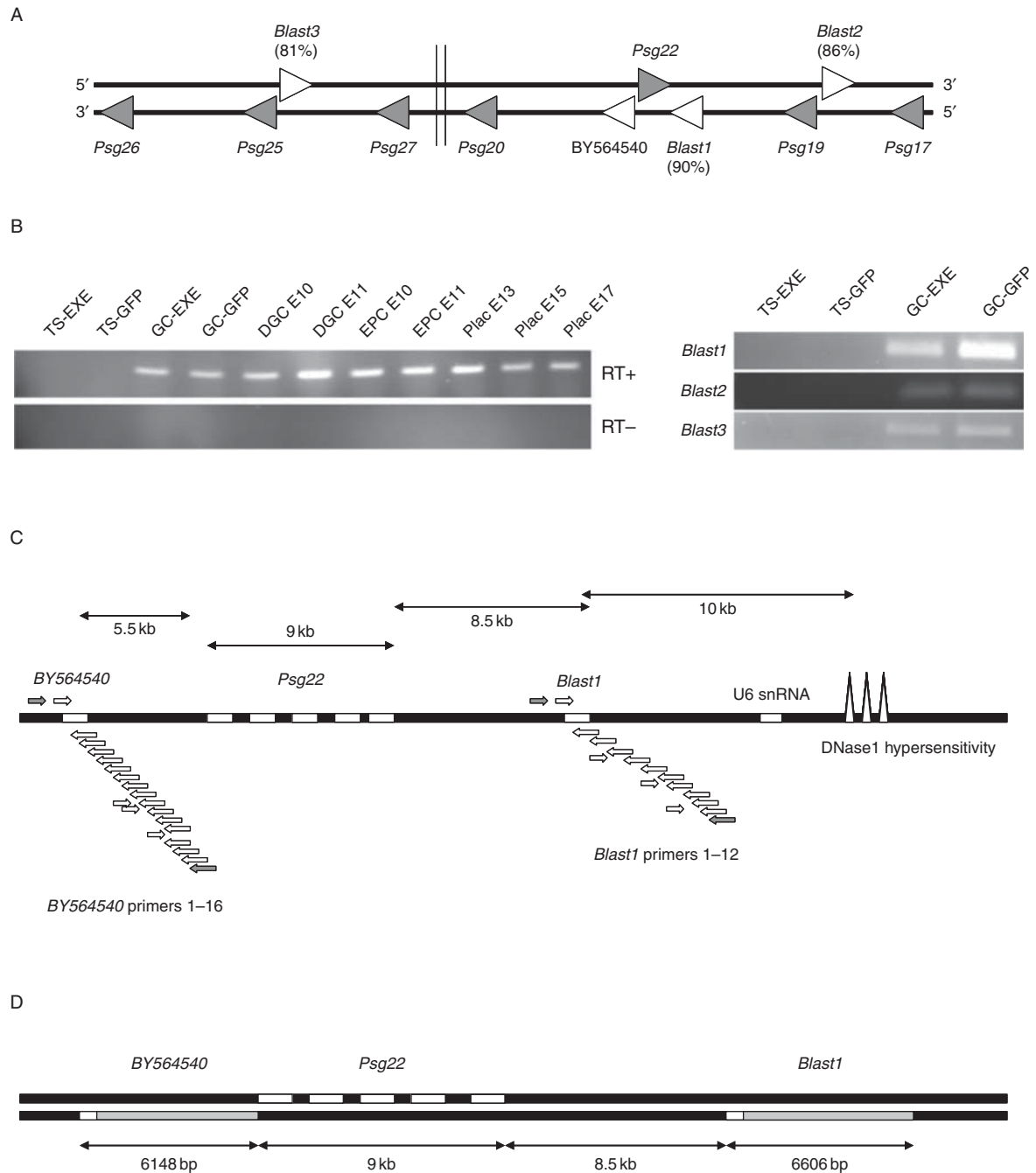


Figure 7 Mapping BY564540 EST and related sequences. (A) Three regions (Blast1–3) contain sequences homologous to BY564540 EST region. Directionality indicated by arrowhead orientation. Degree of similarity to BY564540 given in parentheses (figure not to scale). (B) BY564540 transcript expression analysis by RT-PCR in TSC and TGCs (panel on left). Expression analysis by RT-PCR of Blast1–3 regions in TSC and TGCs (panel on right). (C) Schematic of *Psg22* locus, indicating RT-PCR primers used to map transcription in BY564540 and Blast1 regions. Filled primers did not detect transcription; open primers did. Primer sequences are listed in Supplementary Table 1. (D) Schematic of *Psg22* locus indicating extent of BY564540 (6148 bp) and Blast1 transcription (6606 bp). In (C) and (D), white bars are expressed sequences: *Psg22* exons 1–5; BY564540 EST region and homologous Blast1 region; U6snRNA. In (D), grey bars are expressed sequences mapped upstream of BY564540 and Blast1.

endothelium (Blois *et al.* 2012). However, little is known about *Psg22* regulation or whether its expression in TGCs is functionally significant. In this study, we found that *Psg22* is strongly induced upon differentiation of TSC to TGCs *in vitro*, compared with its closely related

paralogue *Psg19*, or to the closely related *Psg21* and *Psg23* gene pair, which are relatively highly expressed in SpT compared with TGCs. We found that the *Psg22* locus has an inverted orientation relative to other *Psg* genes, within the *Mill1/Mill2* *Psg* cluster, which we

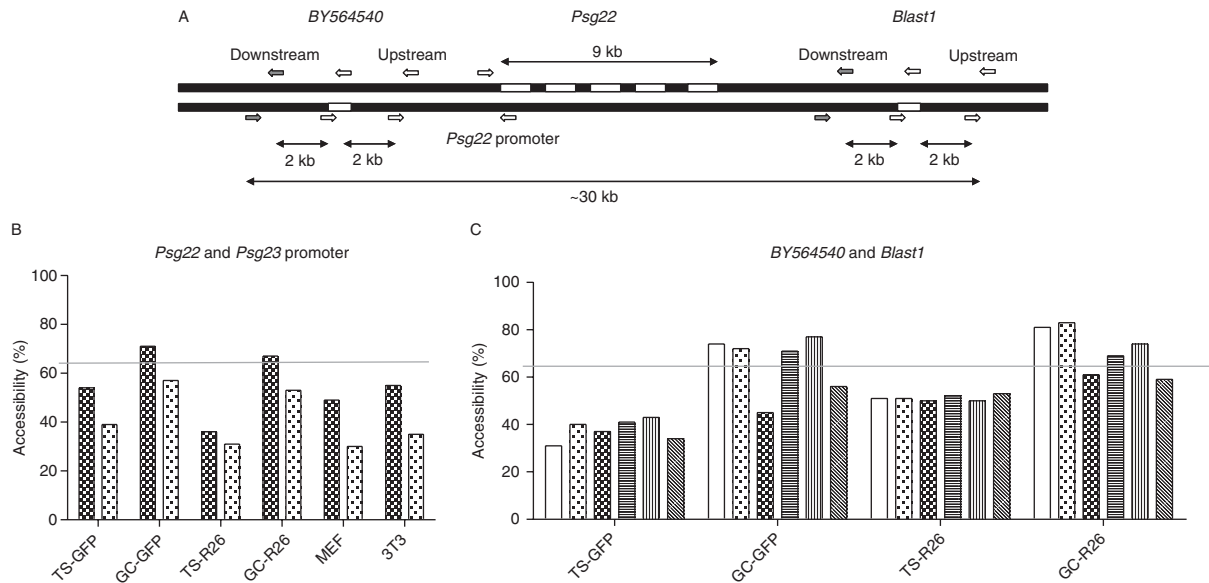


Figure 8 Quantification of chromatin accessibility associated with BY564540 and Blast1 transcripts in TSC and TGCs. (A) Locations of primers used in chromatin accessibility assay. Open chromatin indicated by open primer pairs and closed chromatin by filled primer pairs. (B) Quantification of chromatin accessibility in *Psg22* and *Psg23* promoter regions. Dark bar, *Psg22*; light bar *Psg23*. MEF, primary mouse embryonic fibroblast cells; 3T3, mouse embryonic fibroblast cell line. (C) Quantification of chromatin accessibility in BY564540 region (2 kb upstream, open bar; EST region, speckled bar; 2 kb downstream, filled bar); and Blast1 region (2 kb upstream, horizontal hatch; homologous region to BY564540, vertical line; 2 kb downstream, diagonal hatch). Bio-Rad EpiQ Chromatin Accessibility Assay was used with specific primer pairs producing < 300 bp amplicons, as per (A). In (B) and (C), horizontal line at 65% indicates threshold score for open chromatin. $n=3$.

speculate may explain its unique regulation relative to other mouse *Psgs*. This configuration is present in both the C57Bl/6 and 129/Sv strains, suggesting that it is not a recent laboratory mouse strain mutation. Allied to this, we found evidence of extensive non-coding RNA transcription (lncRNA) associated with conserved sequences flanking the *Psg22* gene, elements that may have undergone duplication arising from the inversion event. Moreover, using an *in vitro* TSC differentiation model, we showed that, similar to *Psg22*, these lncRNAs are developmentally regulated.

To further understand the significance of lncRNA expression at the *Psg22* locus, we analysed chromatin structure in the *Psg22* core promoter and in regions upstream and downstream of the BY564540 and Blast1 lncRNA transcriptional units. Consistent with specific developmental regulation of *Psg22* and associated lncRNA expression, we found that chromatin accessibility increased following differentiation of TSCs to TGCs. Open chromatin is generally associated with permissive or active gene promoters and other regulatory elements such as enhancers, and with lncRNA expression (Cheung & Lau 2005, Amaral & Mattick 2008, Mattick *et al.* 2009, Ma *et al.* 2012, Rinn & Chang 2012). Significantly, 2 kb regions upstream, but not downstream, of lncRNAs exhibited increased sensitivity to nuclease following TSC differentiation, suggesting that lncRNA expression is associated with specific regulatory regions, rather than being a reflection of a global *Psg*

locus-wide increase in chromatin accessibility. Moreover, a region spanning the TSS of the *Psg23* gene, which is highly expressed in SpT but not in TGCs, did not exhibit increased accessibility in these experiments indicating that the chromatin response to TSC differentiation to TGCs may be specific to *Psg22* and associated lncRNAs and not a general feature of the locus (Fig. 8B).

Our attempts to define the lncRNA TSS using 5'-RACE PCR were unsuccessful. Therefore, we used series of overlapping PCR primers in genome walking experiments to identify the approximate locations of lncRNA TSS. The location of the BY564540 TSS in the *Psg22* 5' UTR indicates that this transcript partly overlaps the *Psg22* transcript and suggests that the open chromatin conformation associated with the expression of the BY564540 may facilitate access of transcriptional regulators to the *Psg22* promoter. The fact that the BY564540 transcript is apparently non-coding, lacks a polyA tail, has concordant expression with, and is transcribed antisense to, *Psg22*, suggests that it may function as an enhancer RNA (eRNA; De Santa *et al.* 2010, Kim *et al.* 2010, Rinn & Chang 2012). Blast analysis of BY564540 EST identified similar regions within the *Psg* locus. The Blast1 region downstream of *Psg22* and two other regions located upstream of *Psg19* (Blast2) and upstream of *Psg25* (Blast3) are all expressed in TGCs but not TSCs, and may contribute to enhancer-mediated *Psg* expression. The exact mechanism of these putative eRNAs is yet to be elucidated but may involve

an RNA:RNA hybrid of lncRNA with a nascent coding transcript (Bonasio *et al.* 2010, Hung & Chang 2010).

Our evidence of specific regulation of *Psg22* during TGC differentiation suggests that it is involved in regulating placentation or maternal physiology. Although lacking a specific anti-*Psg22* antibody, we nevertheless showed that *Psg22* in TGCs is associated with polysomes suggesting that it is translated. These observations all point to a role for *Psg22* expression in regulating placentation or other aspects of pregnancy, rather than merely being leaky or coincidental expression from this complex locus.

We have previously demonstrated that the tetraspanin CD9, a receptor for several mouse Psgs including *Psg22*, is strongly expressed in the decidua and maternal vasculature in early mouse gestation (Wynne *et al.* 2006). A polyclonal antiserum raised against *Psg23* that is likely to cross-react with multiple Psgs (although this has not been extensively tested), stained maternal vascular endothelium in histological sections of early implantation sites, consistent with binding of *Psg22* to CD9 expressed on maternal tissues (Wynne *et al.* 2006). We have previously reported, using embryo transfer experiments of WT embryos into *Cd9* null mice, that CD9 was not essential for successful pregnancy (Wynne *et al.* 2006). However, a recent report has indicated that maternal CD9 is critical for uterine repair following parturition, and therefore we speculate that putative binding of *Psg22* to CD9 may regulate endometrial function or repair. This report, combined with our current findings, suggest *Psg22* as a strong candidate for further functional studies including gene targeting to produce a null mutant.

Supplementary data

This is linked to the online version of the paper at <http://dx.doi.org/10.1530/REP-14-0390>.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This work was supported by a Science Foundation Ireland (SFI) Principal Investigator Award and a Higher Education Authority (HEA-PRTL13) grant to T Moore. J M Williams was a recipient of a Marie Curie EU FP6 Early Stage Research Training Fellowship in Regenerative Medicine at the University of Bath.

Acknowledgements

The authors thank Dr Myriam Hemberger and Dr Tilo Kunath for providing TS cell lines, and Dr Kim Moorwood, Dr Mike Cowley and Dr Dmitry Andreev for technical advice.

References

- Amaral PP & Mattick JS 2008 Noncoding RNA in development. *Mammalian Genome* **19** 454–492. (doi:10.1007/s00335-008-9136-7)
- Ball M, McLellan AS, Collins B, Coadwell J, Stewart F & Moore T 2004 An abundant placental transcript containing an IAP-LTR is allelic to mouse pregnancy-specific glycoprotein 23 (*Psg23*): cloning and genetic analysis. *Gene* **325** 103–113. (doi:10.1016/j.gene.2003.10.001)
- Beauchemin N, Draber P, Dveksler G, Gold P, Gray-Owen S, Grunert F, Hammarström S, Holmes KV, Karlsson A, Kuroki M *et al.* 1999 Redefined nomenclature for members of the carcinoembryonic antigen family. *Experimental Cell Research* **252** 243–249. (doi:10.1006/excr.1999.4610)
- Blanchon L, Nores R, Gallot D, Marceau G, Borel V, Yang VW, Bocco JL, Lémyer D, Panzetta-Dutari G & Sapin V 2006 Activation of the human pregnancy-specific glycoprotein PSG-5 promoter by KLF4 and Sp1. *Biochemical and Biophysical Research Communications* **343** 745–753. (doi:10.1016/j.bbrc.2006.03.032)
- Blankley RT, Fisher C, Westwood M, North R, Baker PN, Walker MJ, Williamson A, Whetton AD, Lin W, McCowan L *et al.* 2013 A label-free selected reaction monitoring workflow identifies a subset of pregnancy specific glycoproteins as potential predictive markers of early-onset pre-eclampsia. *Molecular & Cellular Proteomics* **12** 3148–3159. (doi:10.1074/mcp.M112.026872)
- Blois SM, Tirado-González I, Wu J, Barrientos G, Johnson B, Warren J, Freitag N, Klapp BF, Irmak S, Ergun S *et al.* 2012 Early expression of pregnancy-specific glycoprotein 22 (PSG22) by trophoblast cells modulates angiogenesis in mice. *Biology of Reproduction* **86** 191. (doi:10.1095/biolreprod.111.098251)
- Blois SM, Sulkowski G, Tirado-González I, Warren J, Freitag N, Klapp BF, Rifkin D, Fuss I, Strober W & Dveksler GS 2014 Pregnancy-specific glycoprotein 1 (PSG1) activates TGF- β and prevents dextran sodium sulfate (DSS)-induced colitis in mice. *Mucosal Immunology* **7** 348–358. (doi:10.1038/mi.2013.53)
- Bonasio R, Tu S & Reinberg D 2010 Molecular signals of epigenetic states. *Science* **330** 612–616. (doi:10.1126/science.1191078)
- Camolotto S, Racca A, Rena V, Nores R, Patrino LC, Genti-Raimondi S & Panzetta-Dutari GM 2010 Expression and transcriptional regulation of individual pregnancy-specific glycoprotein genes in differentiating trophoblast cells. *Placenta* **31** 312–319. (doi:10.1016/j.placenta.2010.01.004)
- Chamberlin ME, Lei KJ & Chou JY 1994 Subtle differences in human pregnancy-specific glycoprotein gene promoters allow for differential expression. *Journal of Biological Chemistry* **269** 17152–17159.
- Chang CL, Semyonov J, Cheng PJ, Huang SY, Park JI, Tsai HJ, Lin CY, Gruzner F, Soong YK, Caj JJ *et al.* 2013 Widespread divergence of the CEACAM/PSG genes in vertebrates and humans suggests sensitivity to selection. *PLoS ONE* **8** e61701. (doi:10.1371/journal.pone.0061701)
- Cheung P & Lau P 2005 Epigenetic regulation by histone methylation and histone variants. *Molecular Endocrinology* **19** 563–573. (doi:10.1210/me.2004-0496)
- Chuong EB, Tong W & Hoekstra HE 2010 Maternal–fetal conflict: rapidly evolving proteins in the rodent placenta. *Molecular Biology and Evolution* **27** 1221–1225. (doi:10.1093/molbev/msq034)
- De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK, Muller H, Ragoussis J, Wei CL & Natoli G 2010 A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biology* **8** 5. (doi:10.1371/journal.pbio.1000384)
- Ellerman DA, Ha CT, Primakoff P, Myles DG & Dveksler GS 2003 Direct binding of the ligand PSG17 to CD9 requires a CD9 site essential for sperm–egg fusion. *Molecular Biology of the Cell* **14** 5098. (doi:10.1091/mbc.E03-04-0244)
- Endoh M, Kobayashi Y, Yamakami Y, Yonekura R, Fujii M & Ayusawa D 2009 Coordinate expression of the human pregnancy-specific glycoprotein gene family during induced and replicative senescence. *Biogerontology* **10** 213–221. (doi:10.1007/s10522-008-9173-3)
- Felsenstein J 1985 Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39** 783. (doi:10.2307/2408678)
- Gardner RL 1983 Origin and differentiation of extraembryonic tissues in the mouse. *International Review of Experimental Pathology* **24** 63–133.
- Ha CT, Waterhouse R, Wessells J, Wu JA & Dveksler GS 2005 Binding of pregnancy-specific glycoprotein 17 to CD9 on macrophages induces secretion of IL-10, IL-6, PGE₂, and TGF- β 1. *Journal of Leukocyte Biology* **77** 948–957. (doi:10.1189/jlb.0804453)

- Ha CT, Waterhouse R, Warren J, Zimmermann W & Dveksler GS 2008 N-glycosylation is required for binding of murine pregnancy-specific glycoproteins 17 and 19 to the receptor CD9. *American Journal of Reproductive Immunology* **59** 251–258. (doi:10.1111/j.1600-0897.2007.00573.x)
- Ha CT, Wu J, Irmak S, Lisboa FA, Dizon AM, Warren J, Ergun S & Dveksler G 2010 Human pregnancy specific β -1-glycoprotein 1 (PSG1) has a potential role in placental vascular morphogenesis. *Biology of Reproduction* **83** 27–35. (doi:10.1095/biolreprod.109.082412)
- Haig D 1993 Genetic conflicts in human pregnancy. *Quarterly Review of Biology* **4** 495–532. (doi:10.1086/418300)
- Hemberger M, Hughes M & Cross JC 2004 Trophoblast stem cells differentiate *in vitro* into invasive trophoblast giant cells. *Developmental Biology* **271** 362–371. (doi:10.1016/j.ydbio.2004.03.040)
- Himeno E, Tanaka S & Kunath T 2008 Isolation and manipulation of mouse trophoblast stem cells. *Current Protocols in Stem Cell Biology* **1** (Unit 1) E4. (doi:10.1002/9780470151808.sc01e04s7)
- Hung T & Chang HY 2010 Long noncoding RNA in genome regulation: prospects and mechanisms. *RNA Biology* **7** 582–585. (doi:10.4161/rna.7.5.13216)
- Kammerer R & Zimmermann W 2010 Coevolution of activating and inhibitory receptors within mammalian carcinoembryonic antigen families. *BMC Biology* **8** 12. (doi:10.1186/1741-7007-8-12)
- Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S *et al.* 2010 Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465** 182–187. (doi:10.1038/nature09033)
- Kromer B, Finkenzeller D, Wessells J, Dveksler G, Thompson J & Zimmermann W 1996 Coordinate expression of splice variants of the murine pregnancy-specific glycoprotein (PSG) gene family during placental development. *European Journal of Biochemistry/FEBS* **242** 280–287. (doi:10.1111/j.1432-1033.1996.0280r.x)
- Lee JN, Grudzinskas JG & Chard T 1979 Circulating levels of pregnancy proteins in early and late pregnancy in relation to placental tissue concentration. *British Journal of Obstetrics and Gynaecology* **86** 888–890. (doi:10.1111/j.1471-0528.1979.tb10717.x)
- Lei KJ, Sartwell AD, Pan CJ & Chou JY 1992 Cloning and expression of genes encoding human pregnancy-specific glycoproteins. *Journal of Biological Chemistry* **267** 16371–16378.
- Lin T, Halbert S & Spellacy W 1974 Measurement of pregnancy-associated plasma proteins during human gestation. *Journal of Clinical Investigation* **54** 576–582. (doi:10.1172/JCI107794)
- Lisboa FA, Warren J, Sulkowski G, Aparicio M, David G, Zudaire E & Dveksler GS 2011 Pregnancy-specific glycoprotein 1 induces endothelial tubulogenesis through interaction with cell surface proteoglycans. *Journal of Biological Chemistry* **286** 7577–7586. (doi:10.1074/jbc.M110.161810)
- López-Díaz F, Nores R, Panzetta-Dutari G, Slavin D, Prieto C, Koritschoner NP & Bocco JL 2007 RXR α regulates the pregnancy-specific glycoprotein 5 gene transcription through a functional retinoic acid responsive element. *Placenta* **28** 898–906. (doi:10.1016/j.placenta.2007.03.002)
- Ma H, Hao Y, Dong X, Gong Q, Chen J, Zhang J & Tian W 2012 Molecular mechanisms and function prediction of long noncoding RNA. *Scientific World Journal* **1** 1–11. (doi:10.1100/2012/541786)
- Martínez FF, Knubel CP, Sánchez MC, Cervi L & Motrán CC 2012 Pregnancy-specific glycoprotein 1a activates dendritic cells to provide signals for Th17-, Th2-, and Treg-cell polarization. *European Journal of Immunology* **42** 1573–1584. (doi:10.1002/eji.201142140)
- Mattick JS, Amaral PP, Dinger ME, Mercer TR & Mehler MF 2009 RNA regulation of epigenetic processes. *BioEssays* **31** 51–59. (doi:10.1002/bies.080099)
- McLellan AS, Zimmerman W & Moore T 2005a Conservation of pregnancy-specific glycoprotein (PSG) N domains following independent expansions of the gene families in rodents and primates. *BMC Evolutionary Biology* **5** 39. (doi:10.1186/1471-2148-5-39)
- McLellan AS, Fischer B, Dveksler G, Hori T, Wynne F, Ball M, Okumura K, Moore T & Zimmerman W 2005b Structure and evolution of the mouse pregnancy-specific glycoprotein (Psg) gene locus. *BMC Genomics* **6** 4. (doi:10.1186/1471-2164-6-4)
- Minagawa S, Nakabayashi K, Fujii M, Scherer SW & Ayusawa D 2005 Early BrdU-responsive genes constitute a novel class of senescence-associated genes in human cells. *Experimental Cell Research* **304** 552–558. (doi:10.1016/j.yexcr.2004.10.036)
- Moore T & Dveksler GS 2014 Pregnancy-specific glycoproteins: complex gene families regulating maternal–fetal interactions. *International Journal of Developmental Biology* **58** 273–280. (doi:10.1387/ijdb.130329gd)
- Motrán CC, Díaz FL, Gruppi A, Slavin D, Chatton B & Bocco JL 2002 Human pregnancy-specific glycoprotein 1a (PSG1a) induces alternative activation in human and mouse monocytes and suppresses the accessory cell-dependent T cell proliferation. *Journal of Leukocyte Biology* **72** 512–521.
- Motrán CC, Díaz FL, Montes CL, Bocco JL & Gruppi A 2003 *In vivo* expression of recombinant pregnancy-specific glycoprotein 1a induces alternative activation of monocytes and enhances Th2-type immune response. *European Journal of Immunology* **33** 3007–3016. (doi:10.1002/eji.200323993)
- Nores R, Blanchon L, López-Díaz F, Bocco JL, Patrito LC, Sapin V & Panzetta-Dutari GM 2004 Transcriptional control of the human pregnancy-specific glycoprotein 5 gene is dependent on two GT-boxes recognized by the ubiquitous specificity protein 1 (Sp1) transcription factor. *Placenta* **1** 9–19. (doi:10.1016/S0143-4004(03)00213-3)
- Oikawa D, Akai R & Iwawaki T 2010 Positive contribution of the IRE1 α –XBP1 pathway to placental expression of CEA family genes. *FEBS Letters* **584** 1066–1070. (doi:10.1016/j.febslet.2010.02.012)
- Pihl K, Larsen T, Laursen I, Krebs L & Christiansen M 2009 First trimester maternal serum pregnancy-specific β -1-glycoprotein (SP1) as a marker of adverse pregnancy outcome. *Prenatal Diagnosis* **29** 1256–1261. (doi:10.1002/pd.2408)
- Prete MD, Vernal R & Dolznig H 2007 Isolation of polysome-bound mRNA from solid tissues amenable for RT-PCR and profiling experiments. *RNA* **13** 414–421. (doi:10.1261/rna.79407)
- Racca AC, Camolotto SA, Ridano ME, Bocco JL, Genti-Raimondi S & Panzetta-Dutari GM 2011 Krüppel-like factor 6 expression changes during trophoblast syncytialization and transactivates β hCG and PSG placental genes. *PLoS ONE* **6** 7. (doi:10.1371/journal.pone.0022438)
- Rebstock S, Lucas K, Weiss M, Thompson J & Zimmermann W 1993 Spatiotemporal expression of pregnancy-specific glycoprotein gene rncGM1 in rat placenta. *Developmental Dynamics* **198** 171–181. (doi:10.1002/aja.1001980303)
- Rettenberger G, Klett C, Zechner U, Kunz J, Vogel W & Hameister H 1995 Visualization of the conservation of synteny between humans and pigs by heterologous chromosomal painting. *Genomics* **26** 372–378. (doi:10.1016/0888-7543(95)80222-8)
- Rinn JL & Chang HY 2012 Genome regulation by long noncoding RNAs. *Annual Review of Biochemistry* **81** 145–166. (doi:10.1146/annurev-biochem-051410-092902)
- Rudert F, Zimmermann W & Thompson JA 1989 Intra- and interspecies analyses of the carcinoembryonic antigen (CEA) gene family reveal independent evolution in primates and rodents. *Journal of Molecular Evolution* **29** 126–134. (doi:10.1007/BF02100111)
- Senner CE & Hemberger M 2010 Regulation of early trophoblast differentiation – lessons from the mouse. *Placenta* **31** 944–950. (doi:10.1016/j.placenta.2010.07.013)
- Shanley DK, Kiely PA, Golla K, Allen S, Martin K, O’Riordan RT, Ball M, Aplin JD, Singer BB, Caplice N *et al.* 2013 Pregnancy-specific glycoproteins bind integrin α IIb β 3 and inhibit the platelet-fibrinogen interaction. *PLoS ONE* **8** e57491. (doi:10.1371/journal.pone.0057491)
- Snyder SK, Wessner DH, Wessells J, Waterhouse R, Wahl LM, Zimmermann W & Dveksler G 2001 Pregnancy-specific glycoproteins function as immunomodulators by inducing secretion of IL-10, IL-6 and TGF- β 1 by human monocytes. *American Journal of Reproductive Immunology* **45** 205–216. (doi:10.1111/j.8755-8920.2001.450403.x)
- Stone AB 1973 A simplified method for preparing sucrose gradients. *Journal of Biochemistry* **137** 117–118.
- Sulkowski GN, Warren J, Ha CT & Dveksler GS 2011 Characterization of receptors for murine pregnancy specific glycoproteins 17 and 23. *Placenta* **32** 603–610. (doi:10.1016/j.placenta.2011.05.008)
- Tamura K & Nei M 1993 Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* **10** 512–526.

- Tanaka S** 1998 Promotion of trophoblast stem cell proliferation by FGF4. *Science* **282** 2072–2075. (doi:10.1126/science.282.5396.2072)
- Teglund S, Olsen A, Khan WN, Frangsmyr L & Hammarström S** 1994 The pregnancy-specific glycoprotein (PSG) gene cluster on human chromosome 19: fine structure of the 11 PSG genes and identification of 6 new genes forming a third subgroup within the carcinoembryonic antigen (CEA) family. *Genomics* **23** 669–684. (doi:10.1006/geno.1994.1556)
- Waterhouse R, Ha CT & Dveksler G** 2002 Murine CD9 is the receptor for pregnancy-specific glycoprotein 17. *Journal of Experimental Medicine* **195** 277–282. (doi:10.1084/jem.20011741)
- Wessells J, Wessner DH, Parsells R, White K, Finkenzeller D, Zimmermann W & Dveksler G** 2000 Pregnancy specific glycoprotein 18 induces IL-10 expression in murine macrophages. *European Journal of Immunology* **30** 1830–1840. (doi:10.1002/1521-4141(200007)30:7<1830::AID-IMMU1830>3.0.CO;2-M)
- Wu J, Johnson BL, Chen Y, Ha CT & Dveksler GS** 2008 Murine pregnancy-specific glycoprotein 23 induces the proangiogenic factors transforming-growth factor β 1 and vascular endothelial growth factor a in cell types involved in vascular remodeling in pregnancy. *Biology of Reproduction* **79** 1054–1061. (doi:10.1095/biolreprod.108.070268)
- Wynne F, Ball M, McLellan AS, Dockery P, Zimmerman W & Moore T** 2006 Mouse pregnancy-specific glycoproteins: tissue-specific expression and evidence of association with maternal vasculature. *Reproduction* **131** 721–732. (doi:10.1530/rep.1.00869)
- Zhao L, Triche EW, Walsh KM, Bracken MB, Saitlas AF, Hoh J & Dewan AT** 2012 Genome-wide association study identifies a maternal copy-number deletion in PSG11 enriched among preeclampsia patients. *BMC Pregnancy and Childbirth* **12** 61. (doi:10.1186/1471-2393-12-61)

Received 1 August 2014

First decision 5 September 2014

Revised manuscript received 17 October 2014

Accepted 30 October 2014