

Gene expression in the developing embryo and fetus

J. Taylor¹, H. Fairburn¹, N. Beaujean^{1,2}, R. Meehan² and L. Young³

¹Division of Gene Expression and Development, Roslin Institute, Roslin, Midlothian EH25 9PS, UK; ²Department of Biochemistry, University of Edinburgh, Hugh Robson Building, George Square, Edinburgh EH8 9XD, UK; and ³Division of Obstetrics and Gynaecology, School of Human Development, University of Nottingham, Queens Medical Centre, Nottingham NG7 2UH, UK

Determining the stage- and tissue-specific patterns of gene expression shown by the embryo and fetus will provide information about the control of normal development. Identification of alterations in these patterns associated with specific abnormal phenotypes will also be informative regarding the underlying molecular mechanisms. In addition, qualitative and quantitative changes in gene expression that deviate from the norm may provide a potential marker system for predicting future developmental defects, a system that would be particularly useful in preimplantation embryo technologies before recipient transfer. However, there are a number of important issues regarding the interpretation and relevance of many gene expression studies currently undertaken that are often not considered or are ignored. Even when rigorous methodology is applied to detect differences in gene expression, their functional significance is rarely defined. This review discusses the relevance of gene expression changes as diagnostic markers in relation to protein and epigenetic changes and indicates that gene expression studies should be rigorously designed and interpreted to yield meaningful results.

Introduction

Owing to the marked morphological changes that embryonic and early fetal cells undergo with most successive cell divisions, a highly co-ordinated and rapidly changing pattern of gene expression is required to ensure normal development. This pattern is in marked contrast to that of the adult organism in which most cells exist as differentiated lineages that maintain a baseline pattern of constitutive expression in most of their expressed genes, with a smaller component of the genome showing facultative expression changes in response to environmental stimuli. During the early phases of development, there is an entire reprogramming of the epigenetic components that regulate gene expression patterns within each cell (Bird, 2001), beginning with marked remodelling of the spermatozoa and egg nuclei into a zygotic nucleus (Renard, 1998). As the embryonic genome activates, blastomere nuclei direct differentiation of the first embryonic lineages and further epigenetic changes occur as each

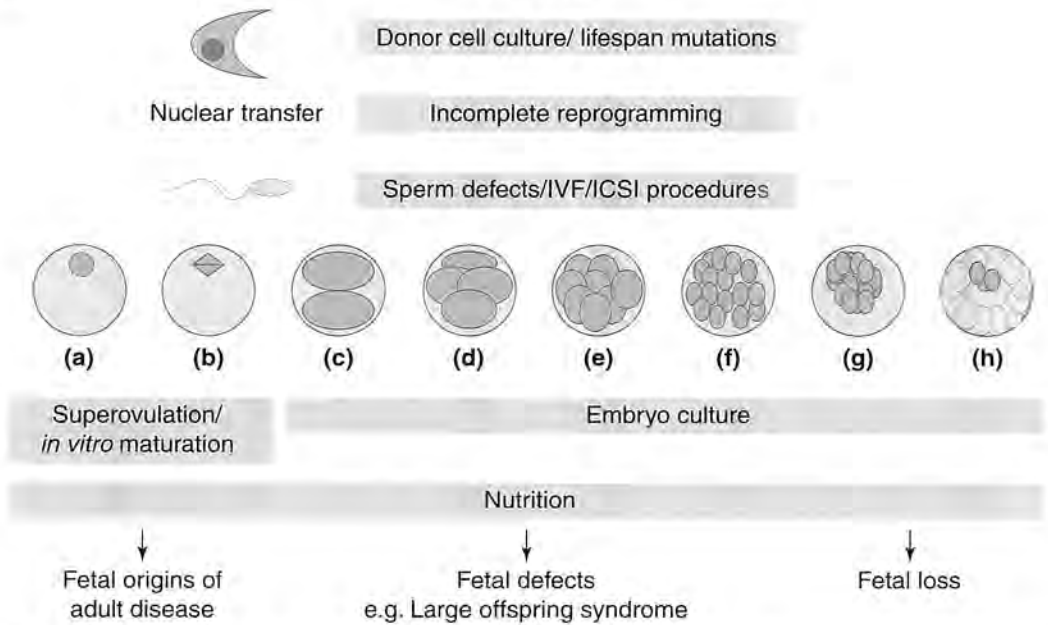


Fig. 1. Potential environmental and procedural influences on preimplantation gene expression (boxes) that may induce later developmental defects (arrows). Stages represented are (a) germinal vesicular stage oocyte, (b) mature (metaphase II) oocyte, (c) two-cell embryo, (d) four-cell embryo, (e) eight-cell embryo, (f) 16-cell embryo, (g) compact morula and (h) blastocyst with inner cell mass (green) and trophectoderm (pink) lineages.

new fetal lineage is formed. The epigenetic changes that must occur during embryogenesis are likely to be subject to nutritional and other environmental influences that programme the genome to result in the subtle changes responsible for later phenotypic effects, such as individual variation in birth weight and adult disease status (Young, 2001) (Fig. 1). More severe environmental perturbations induced by embryo technologies can induce more marked phenotypes, especially in ruminants (Young and Fairburn, 2000).

As more information on gene expression patterns during early embryogenesis becomes available, research emphasis must now be placed on unravelling the function of key genes in the developing embryo and the fetal consequences of their disruption. In addition, the nature and consequences of differences in gene expression patterns between ruminants and mice require investigation, not only to uncover developmental processes with agricultural relevance, but also to further model our understanding of human development in large animal species.

Why study gene expression?

Although functional changes in cellular processes generally occur only when the content or activation states of proteins are altered, often the first cellular response to differentiation or environmental stimuli occurs at gene transcription. Furthermore, there are a number of non-coding mRNA species that are functional only in the RNA form (Szymanski and Barciszewski, 2002). Thus, from studies of gene expression the immediate response of a cell to a change in environment can be examined. The change in environment can be due to changing stimuli during the course of normal developmental signalling, but can also be due to nutritional

influences or to non-physiological environment changes imposed by embryo technologies. Mapping the patterns of gene expression induced by each of these environmental states provides a powerful tool for assessing normal and abnormal development.

Development of the ruminant embryo and fetus is more similar to that of humans than mice in a number of respects, including delayed activation of the embryonic genome, timing and cyclic nature of follicular oogenesis, size, development and physiology of the fetus, gestation of singletons or twins and extended gestation period. Thus, study of gene expression in these species is likely to provide a more informative model for similar processes in humans.

Normal development

During development, changes in gene expression can often occur in response to the differentiation of the surrounding cells, induced by alterations in factors they secrete or by changes in cell–cell contact. For example, the first differentiation of blastomeres into the trophectoderm lineage is initiated during compaction, when the blastomeres flatten and increase contact with neighbouring cells (Kidder, 2002). The cells on the outer region of the morula acquire polarity, expressing Na^+/K^+ ATPase (Kidder, 2002) and forming a tight junction permeability seal on their apical surface. These outer blastomeres form the trophectoderm (TE), whereas the inner cells remain non-polar and form the inner cell mass (ICM). The induced differences in gene expression that occur as the first two lineages form have been poorly established, largely due to the paucity of material and the difficulties of physically separating the lineages, especially in *in vitro* produced livestock and human embryos in which compaction is often poor and the ICM indistinct. In mice, in which ICM-derived embryonic stem (ES) cell lines and trophectoderm cell lines (Rossant, 2001) have been established, genes considered specific to the ICM include the transcription factor, *Oct 4* (Boiani *et al.*, 2002). However, at least at the protein level, *Oct 4* is expressed in both the ICM and TE cells of ruminant embryos (van Eijk *et al.*, 1999), as determined by whole mount immunostaining. The apparent species differences in a transcription factor, the presence of which is essential to retain the pluripotent phenotype of ICM-derived ES cells in mice, is both perplexing and intriguing. Furthermore, nine cell surface antigens expressed (at least at the protein level) specifically in the ICM of human blastocysts are not expressed in the equivalent mouse lineage (Henderson *et al.*, 2002). The authors of this study speculate that different genes achieve the same end-point function between the species, using different metabolic pathways and this may well apply to a range of embryonic cellular functions. More work on gene, as well as protein, expression is now required.

Instead of using immunosurgery or microdissection techniques to separate the ICM and TE lineages for gene expression studies by RT–PCR, a whole mount *in situ* hybridization approach is preferable, although technically challenging. A non-radioactive method using fluorescein labelling of oligonucleotide probes has now been optimized for use in preimplantation sheep and cattle embryos in our laboratory. Although our studies have focused on epigenetic regulator genes, the method will be generally applicable to detect the *in situ* expression of any gene (J. Taylor and L. Young, unpublished and Fig. 2). This approach will also allow determination of whether heterogeneous expression occurs between the blastomeres of an individual embryo.

Defining the processes of normal development is a prerequisite for optimizing maternal nutrition and embryo technologies at the molecular level. This information is necessary to understand agricultural problems, such as infertility in dairy cattle and early pregnancy loss. In addition, in ruminants, identifying the key genes regulating fetal growth and development,

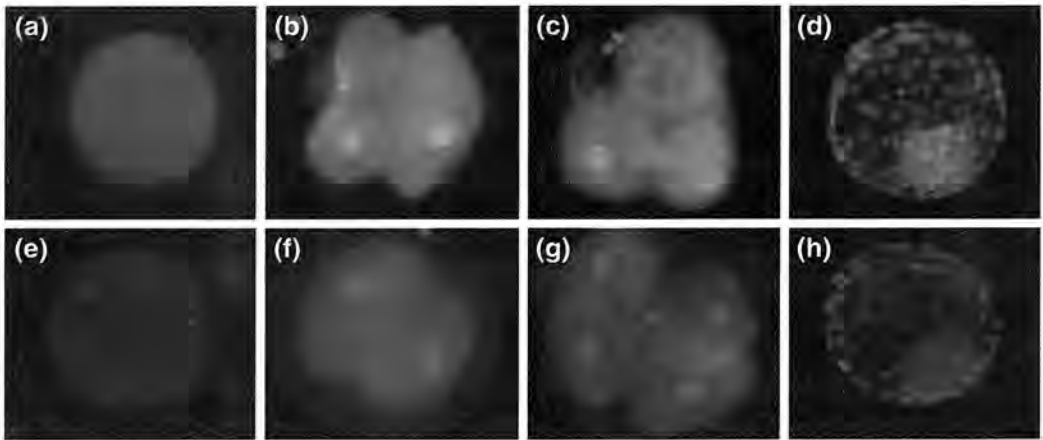


Fig. 2. Whole mount *in situ* hybridization of sheep embryos produced *in vitro* using a fluorescein labelled mixture of oligonucleotide DNA probes for the epigenetic regulator gene, DNA methyltransferase 1 (*DNMT1*). The lower panel hybridized with sense control probes shows the observed level of background and the upper panel shows *DNMT1* gene expression with antisense probes in (a and e) oocytes matured *in vitro*, (b and f) two-cell embryos, (c and g) four-cell embryos and (d and h) day 7 blastocysts.

and subsequent analysis of their allelic variations, will provide an important tool for the improved selection of desirable characteristics.

Nutritional effects and programming

Further alterations in the intrinsic developmental pattern or level of gene expression can be superimposed by the nutritional status of the dam and conceptus (Symonds *et al.*, 2001). Furthermore, the effects of different dietary regimens can vary according to factors such as maternal age (Wallace, 2000). Other environmental factors, such as maternally administered drugs or environmental toxins such as phyto-oestrogens, can also alter gene expression *in utero* (Rhind, 2002). The effects of such nutritional and environmental influences on gene expression can be immediate. Alternatively, they may be involved in 'programming' gene expression or phenotypic effects that do not occur until much later in development. The 'Barker hypothesis', indicating that fetal nutritional environments limiting fetal growth and specific organ development can predispose to adult cardiovascular disease, obesity, diabetes and perhaps even cancer, is now well-established in terms of experimental evidence (Hales and Barker, 2001). However, the underlying molecular mechanisms remain to be elucidated, and there is an ongoing debate on whether the effects are due to genetic variability in alleles of physiologically relevant genes or to direct nutritional effects initiating gene expression changes *in utero*.

Studies using sheep models of maternal nutrient manipulation have shown that a number of physiological systems in both the developing fetus and placenta have nutritionally induced alterations in gene expression (Wallace, 2000; Symonds *et al.*, 2001; Rhind, 2002). However, it is not clear whether these are initiating or downstream responses. It has been suggested that nutritional and environmental programming may occur by inducing heritable changes in the epigenetic components that regulate the structure of chromatin and hence the patterns of gene expression within cells (Fig. 3a; discussed by Young, 2001). As a result of their potent effects

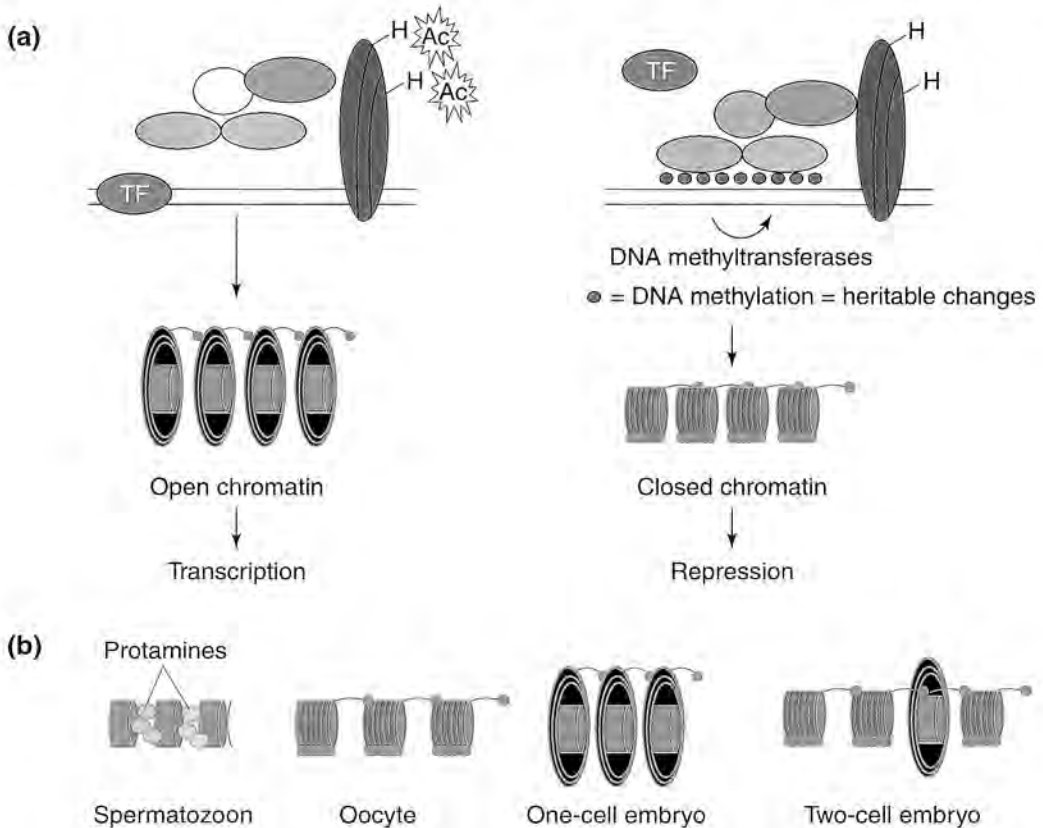


Fig. 3. The epigenetic regulation of gene activity in (a) differentiated lineages and (b) the early mouse embryo. (a) Transcriptionally active, open chromatin is formed when the complement of epigenetic regulators (represented by coloured circles and ovals) includes histone (H), acetylation (Ac) and hypomethylation, allowing transcription factor (TF) access to gene promoters (left panel). Repressed chromatin conversely excludes transcription factor access when histones are deacetylated and this repression is rendered heritable between cell cycles by the recruitment of methyl binding proteins by methylated DNA (right panel). (b) Sperm DNA is rendered transcriptionally inactive not by the formation of condensed chromatin but by tight packaging by protamines. Mature oocyte DNA is structured as condensed chromatin but this is temporarily reactivated as the spermatozoa and egg chromatin are remodelled into a zygotic nucleus in the one-cell embryo. This permissive chromatin environment probably explains the expression of tissue-specific genes at this time. By the time of zygotic genome activation in the late two-cell mouse embryo, most chromatin has been recondensed to allow only activation of those genes required for stage-specific development. The pattern of spermatozoa and egg chromatin remodelling in other species is not known.

in determining birth weight, affecting development of specific organs and predisposition to heritable epigenetic alterations, several imprinted genes have been suggested as potential candidates for initiating the fetal origins of adult disease (Young, 2001). Constancia *et al.* (2002) demonstrated that deletion of a placental-specific transcript of the imprinted fetal mitogen gene, *Igf2*, results in altered placental supply of amino acids to the fetus, resulting in fetal growth restriction in mice.

Recent evidence from rats (Kwong *et al.*, 2000) and ruminants (Sinclair *et al.*, 2000; Young, 2001) indicates that the nutritional programming of development described by the Barker hypothesis can now be extended back to the preimplantation embryo and even the

developing oocyte. Nutrients that could affect the DNA methylation epigenetic component (including methionine, homocysteine, polyamines, vitamin B12, folate and urea/ammonia (Heby, 1995; Young, 2001; Rees, 2002) may permanently alter the programmed pattern of gene expression in these early stages, when very dynamic changes of both genome-wide methylation and imprinted gene methylation occur (for a review, see Young and Fairburn, 2000; Young, 2001). This hypothesis merits urgent investigation because of the implications for both animal breeding and human health.

Embryo technologies

As most embryo technologies involve a period of development *in vitro* (or *in vivo* culture in a ligated oviduct), exposure to a non-physiological environment may also alter gene expression (Fig. 1; reviewed by Young and Fairburn, 2000; Niemann *et al.*, 2002). In addition to the effect of unusual embryo culture environments, nuclear transfer embryos are likely to be subject to a wider range of alterations due to incomplete reprogramming of the donor cell genome (Fig. 1; Wilmut *et al.*, 2002). Gene disruptions resulting from specific embryo technologies are reviewed extensively elsewhere (De Sousa *et al.*, 1998; Niemann and Wrenzycki, 2000; Daniels *et al.*, 2001; Natale *et al.*, 2001; Dieleman *et al.*, 2002; Niemann *et al.*, 2002; Rizos *et al.*, 2002; McEvoy *et al.* this issue; Wells, this issue). At least some of these changes are likely to underlie the poor developmental potential of embryos produced by *in vitro* maturation (IVM), IVF and *in vitro* embryo culture (IVC), as well as nuclear transfer (NT). Embryos subject to many gene alterations at the early embryo stage, or in a few genes that are key to early developmental processes, would be expected to perish in the pre- or peri-implantation period, accounting for the high incidence of early losses often associated with embryo technologies (Young *et al.*, 1998). However, embryos escaping with fewer gene alterations are more likely to survive to later fetal or neonatal stages.

The contention that at least some of these alterations can have profound phenotypic consequences was most marked when a high incidence of large offspring syndrome (Young *et al.*, 1998) was induced after *in vivo* derived sheep zygotes were cultured in serum-containing medium for 5 days before embryo transfer (Sinclair *et al.*, 1999). As both the gene and protein expression as well as DNA methylation (Bird, 2002) of an imprinted gene, *IGF2R*, were downregulated in the tissues of overgrown (but not of normal-sized) sheep fetuses (Young *et al.*, 2001) and as the knockout phenotype in mice is 30% oversize at birth with associated perinatal lethality and a number of specific organ defects (Lau *et al.*, 1994), it seems likely that an epigenetic change in this gene in a subset of *in vitro* cultured embryos resulted in a long-term alteration in gene expression. The particular susceptibility of imprinted genes to disruption during oogenesis and preimplantation development is discussed by Young and Fairburn (2000). However, confirmation of this hypothesis awaits the development of sufficiently sensitive and reproducible techniques to examine the methylation of single gene loci in individual embryos.

Our laboratory has recently developed a single-embryo methylation-sensitive PCR method that is highly reproducible for detecting the degree of methylation in sheep SINE (short interspersed nucleotide elements) and satellite sequences that are repeated many times throughout the genomic DNA (H. Fairburn and L. Young, unpublished and Fig. 4a) and that we will now optimize for use in imprinted gene loci. However, as significant differences were observed among IVM, IVF zygotes cultured to the blastocyst stage in a serum-free medium and blastocysts recovered *in vivo* from superovulated dams (Fig. 4b), the choice of control embryos for any imprinting studies will have to be examined carefully. Furthermore, preliminary bisulphite sequencing of satellite I DNA from single sheep embryos has revealed considerable

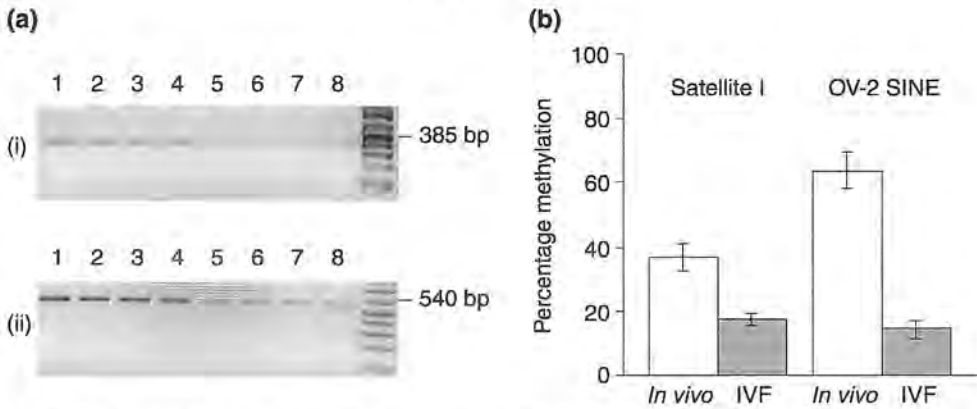


Fig. 4. (a) Methylation sensitive-PCR replicates of satellite 1 DNA in single sheep blastocysts using (i) Satellite 1 or (ii) OV-2 short interspersed nucleotide element (SINE) specific primers. Lanes 1–4 are replicate PCR products generated from single blastocyst DNA. Lanes 5–8 contain replicate PCR products generated from single blastocyst DNA digested with the methyl-sensitive enzyme *Hha* I. A single *Hha* I site is spanned by the Satellite 1 and OV-2 SINE-specific PCR primers. (b) Comparison of Satellite 1 and OV-2 SINE DNA methylation in blastocysts derived either *in vivo* or after *in vitro* oocyte maturation, fertilization and embryo culture. The mean of sequence-specific DNA methylation levels (Satellite 1 or OV-2) from embryos of a single source was calculated and is indicated as percentage values on the y axis. The standard error of the mean is indicated as bi-directional error bars on the chart. The differences in mean methylation observed for *in vivo* and IVM and IVF derived embryos are significant ($P < 0.01$) for both Satellite 1 and OV-2 SINE.

heterogeneity in the methylation status of individual cytosine residues between embryos produced in the same *in vitro* system (N. Beaujean, R. Meehan and L. Young, unpublished). Heterogeneity was also noted between the DNA copies of cells within an individual blastocyst stage embryo. Thus, as in RT-PCR analysis for the detection of gene expression in embryos, PCR-based DNA methylation analysis will require rigorous validation to generate meaningful and interpretable results.

Preimplantation gene expression changes as diagnostic markers

Many authors have suggested that identifying alterations in expression of one or more genes would provide a useful 'marker' system for predicting abnormal development in subsequent developmental stages and there is particular interest in following this approach in the preimplantation embryo before recipient transfer. However, as not all gene expression changes result in functional changes in proteins and, therefore, cellular phenotype, ultimately it is important to assess the expression and activity of the protein encoded by the gene of interest (Gerlt, 2002).

In addition, epigenetic changes induced at a particular developmental stage may not result in altered gene expression until later in development, when appropriate transcription factors are expressed (Young, 2001). Importantly, studies examining only gene or protein expression at a particular developmental stage (for example in the preimplantation embryo) will fail to detect epigenetic changes and may wrongly conclude that a perturbed tissue sample is developing normally. Such studies include genome wide analysis using for example expression arrays, mRNA differential display and proteomics.

Thus, ultimately, any reliable 'marker' system using molecular changes to assess 'normal' development will probably require assessment at the mRNA, protein and epigenetic levels

and, because of the wide range of potential phenotypic problems, is likely to require screening for a relatively large number of genes. Although PCR approaches are likely to allow the sensitivity required to determine the expression of at least one gene from a single cell biopsy, unlike in preimplantation genetic diagnosis in which all blastomeres are expected to contain the same mutation, gene expression changes initiated at the embryo stage may be subject to mosaicism, with only some cells being affected (or the degree of perturbation may vary between cells).

Functional significance

One of the major problems currently facing gene expression analysis as a 'marker' system, in both ruminants and humans, is determining the phenotypic effects of detected changes and, thus, the efficacy of the marker. Many genes found to be upregulated or downregulated or abnormally expressed or silent at a particular preimplantation stage in particular embryo production systems (for example, Niemann and Wrenzycki, 2000; Natale *et al.*, 2001, Rizos *et al.*, 2002; Dieleman *et al.*, 2002; Niemann *et al.*, 2002) are not easily correlated with the relatively high pregnancy rates often obtained in these systems or with a specific phenotype. Obviously, some observed changes may be associated with the wide range of phenotypic defects often observed in the fetus after production of bovine and sheep embryos *in vitro* (Young *et al.*, 1998), although at least some of the serum-free culture systems reported in the above studies are associated only with a very low incidence of problems. As the gene expression changes described appear to affect a large proportion of embryos, they may be more likely to induce the relatively high levels of early developmental loss or have no obvious phenotypic effect. Thus, the question arises as to whether the embryo is simply extremely plastic, being able to tolerate considerable fluctuations in gene expression, especially if they are transient.

In some cases of downregulated expression it may be that there is functional gene redundancy. If one gene is misexpressed, other genes with overlapping function can compensate. Although it may be expected that a range of expression of a particular gene could be tolerated, tolerance of the unusual absence or presence of genes apparently key to embryo development from mouse studies (such as, leukaemia inhibitory factor (LIF) and LIF receptor, Niemann and Wrenzycki, 2000; Rizos *et al.*, 2002) is more perplexing. LIF and LIF receptor have been detected in *in vitro* produced but less so or not at all in *in vivo* derived bovine embryos. The functional significance of this ectopic expression may lie in the considerable species differences now emerging in terms of gene expression. Although the function of LIF has not been tested in ruminants, it is known to be essential to maintain the undifferentiated state in mouse ICM-derived ES cells, whereas it is neither required nor expressed by human ES or EC cells (Schuringa *et al.*, 2002).

Even in mice, in which knockout technology is now in routine use to establish gene function, knockouts are initiated only in ES cells (derived from the ICM of the blastocyst) and so if the phenotype is lethal or affects reproduction in the first generation, it is not very easy to establish the effects on the preimplantation embryo. Both targeted additions (McCreath *et al.*, 2000) and deletions (Denning *et al.*, 2001) of genes have now been achieved in ruminant species, although the efficiencies are currently too low for routine gene function studies to be either cost-effective or justified on welfare grounds. Furthermore, as gene knockouts and knockins in ruminants currently require nuclear transfer of targeted somatic cells to introduce the genetic modification, it is currently not possible to distinguish fully the effects of the genetic modification from the defects induced by the NT process. The use of RNA interference (RNAi) for transient 'knock-down' of individual gene transcripts has been reported in

mouse oocytes and two-cell embryos (Wianny and Zernicka-Goetz, 2000). However, the need to inject the single stranded RNA into each cell needs to be circumvented before this approach can be used to study gene function at later preimplantation stages. Adaptation of a recently reported plasmid delivery system (Paul *et al.*, 2002) may facilitate these studies as well as allow more stable knock-down if integrated into the genome under control of an introduced promoter.

Obviously in at least some cases gene function will be similar in both mice and ruminants and phenotypic effects of disruption will also be similar. Both the mouse gene knockout database (http://www.informatics.jax.org/searches/allele_form.shtml) and MRC database on mouse developmental gene expression (<http://genex.hgu.mrc.ac.uk/>) can be very informative tools when choosing candidate genes for a specific phenotype or developmental process for study in ruminants. However, it is becoming increasingly obvious that the developmental expression of many genes is quite different between the species (van Eijk *et al.*, 1999; Henderson *et al.*, 2002) and this is now being mirrored by the findings of our laboratory at the epigenetic level (for example, see Wilmot *et al.*, 2002)

Do all genes expressed in the preimplantation embryo regulate development?

Chromatin remodelling

Owing to the marked remodelling of the spermatozoa and egg genomes into a functional zygotic nucleus, there are unusual epigenetic states in the preimplantation embryo that may further complicate functional analysis of gene expression patterns at this time. For many years it has been noticed that the preimplantation embryo expresses a surprising number of genes that in the adult are highly tissue specific (Zhang *et al.*, 1994; Daniels *et al.*, 1997). In ruminants, studies in our laboratory have detected unexpected expression of several transcripts, including the angiogenic gene, *VEGF*, when there is no blood supply, of placental and trophoblast genes, such as *PIGF* and *MASH2*, even in early cleavage stages before the trophoblast is formed and the normally adult liver-specific transcript form of insulin-like growth factor 2 (*IGF2*) (J. Taylor, J. Gardner and L. Young, unpublished). As these genes are generally considered to have very specific functions in specific types of cell, the role of their expression in the embryo is not clear. Furthermore, in our laboratory, a large proportion of the very diverse group of genes that have been examined in both cattle and sheep are expressed at all stages from zygote to blastocyst. This group includes genes that would be expected from studies in mice to be highly stage-specific, for example the *HSP70.1* and *EIF1A* genes associated with the onset of embryonic transcription (G. Ryan and L. Young, unpublished). A similar result for *EIF1A* has been obtained from human embryos (D. J. Bloor and S. J. Kimber, personal communication). These studies have been carried out with rigorous controls to ensure that the results reflect true expression and not any form of contamination. These surprising observations highlight the possibility that many changes in gene expression that are detected in preimplantation embryos of large mammals have no functional role or that the genes involved may not perform the same function among species. However, they can perhaps be explained in light of several studies performed so far in the mouse embryo.

Activation of the mouse embryonic genome occurs via a 'zygotic clock' mechanism, in which the control of development switches from the control of maternal transcripts stored in the developing oocyte to *de novo* synthesis of mRNA by the zygotic genome (Latham and Schultz, 2001). The major event of zygotic genome activation (ZGA; alternatively known as the maternal-zygotic transition or MZT) occurs 40 h after fertilization in mice, delaying co-ordinated transcription of the zygotic genome until the two-cell stage. It is thought that ZGA

is initially inhibited by the high concentration of negative transcriptional regulators which reduces as the zygote divides to form two smaller blastomeres, reducing the nuclear:cytoplasmic ratio. However, as transcription of some genes does occur before ZGA, it is likely that these transcriptional regulators do not regulate the entire genome.

In early embryogenesis, as in adult cells, the structure of chromatin at specific gene loci determines whether specific genes are expressed. When chromatin is tightly packed around nucleosomes, this closed conformation prevents transcription factor binding to gene promoter regions creating transcriptional repression (Fig. 3a; Bird, 2001). In mouse zygotes, the paternal pronucleus has a more open chromatin structure and is more transcriptionally active than the female pronucleus (Renard, 1998; Latham and Schultz, 2001). In fact until the late two-cell stage, mouse chromatin is maintained in an unusually permissive state (Ma *et al.*, 2001) (Fig. 3b) and this may explain why many tissue-specific genes are detected in the early embryo. Furthermore, it has been suggested that the tissue-specific genes that are activated at this time contain CpG islands, regions of DNA that are particularly prone to regulation by one epigenetic component, DNA methylation (Daniels *et al.*, 1997; Bird, 2002). At the late two-cell stage when the murine embryonic genome is activated, much of the chromatin is, paradoxically, modified to a repressive state (Fig. 3b; Renard, 1998; Latham and Schultz, 2001). The timing of this process correlates with synthesis and modification of histones that are part of the chromatin modifying machinery (Fig. 3a). However, some chromatin in the late two-cell embryo is retained in an active state. Presumably, genes actively transcribed in this region include the transcription requiring complex components that initiate ZGA (Ma *et al.*, 2001). After ZGA, the largely transcriptionally repressive state is maintained in two- to eight-cell embryos and is probably critical for the correct programming of stage-specific gene expression necessary for further development.

Such extensive studies of transcriptional regulation have not yet been undertaken in ruminant embryos, although indications of marked and unexpected differences from mice are now emerging. The transcriptionally active male pronucleus loses its DNA methylation shortly after fertilization in mice (Oswald *et al.*, 2000), whereas the inactive female pronucleus retains both DNA methylation and other epigenetic components associated with inactive chromatin (Arney *et al.*, 2002). In marked contrast, the male pronucleus of sheep shows virtually no demethylation in the first cell cycle (Wilmut *et al.*, 2002). However, whether the male sheep pronucleus is transcriptionally inactive remains to be established. One suggested role of demethylating the paternal pronucleus in mice is to open the chromatin sufficiently to allow maternal factors access to the DNA to modify paternal imprints, a process obviously facilitated by the unique separation of the parental genomes for this brief period in the first cell cycle (Oswald *et al.*, 2000). Such species differences in epigenetic processes, despite morphological similarity in preimplantation development, may underlie the increased susceptibility of ruminant embryos to embryo culture-induced defects, such as early loss and large offspring syndrome.

As ZGA is delayed until the 8–16-cell stage in sheep and cattle embryos, the permissive expression of a range of genes not specifically required to regulate preimplantation developmental stage transition may occur for longer in this species. This could explain our unexpected detection of transcripts, such as *HSP70.1* and *EIF1A*, that from mouse studies would be expected to be detected only at the 8–16-cell stage. It may also explain why at least some nutrition or embryo culture-induced alterations in gene expression can occur in early development without apparent effect. Perhaps only genes with very strong enhancers or promoters, or those regulated by specific transcription factors and epigenetic states, have functional significance in early ruminant development. Interpretation of many gene expression results at this stage can be confounded by the use of relatively high cycle numbers of PCR, or often

by the use of cDNA from pooled embryos, such that low, but functionally insignificant, gene expression is often detected.

Stage-specific gene expression

One major recent study has finally started to shed some light on stage-specific gene expression in the preimplantation embryo, at least in the mouse. In a large-scale sequencing project, Ko *et al.* (2000) produced cDNA libraries from each mouse developmental stage between oocyte and blastocyst, using *in vivo* derived embryos to avoid any effects of culture *in vitro*. After examining the sequence identity of transcripts detected with a database of known mouse sequences, 9718 genes were identified that were expressed between fertilization and the blastocyst stages. Thus, about one third of the genes in the mouse genome are expressed during the first few days of life, perhaps a surprisingly high proportion considering that only two lineages are present by this time. Furthermore, half of the genes expressed were novel, that is, had not been identified previously in other tissues or developmental stages. Thus, a considerable proportion of murine genes are apparently dedicated to regulating early embryonic development. However, in light of the tissue-specific gene expression and other issues discussed above, it remains to be established whether all of the early genes transcribed are functional in blastomere development. In the study by Ko *et al.* (2000), expression of 75% of the genes identified was low during preimplantation development and this is consistent with the induction of a repressive chromatin environment. Within this environment, only genes with strong promoter or enhancer regions are able to transcribe at high levels (Ma *et al.*, 2001; Latham and Schultz, 2001) and genes with low expression may not be translated or have no functional significance.

Of the genes identified, 3.2% showed complex expression patterns, increasing and decreasing at various stages. Only 0.1% of the expressed genes was detected at all preimplantation stages and presumably includes housekeeping genes. In contrast, 17% of genes showed expression at only a single stage, and 2.5, 3.57 and 1.66% were specific to the fertilized egg, eight-cell and blastocyst stages, respectively. These stage-specific genes may well be the key genes required to advance embryos between developmental stages and cell divisions. Of the expressed genes, 3.4% were detected only in the early stages and were presumably mRNAs inherited from the maternal ooplasm. Finally, 2.1% of the genes that were not initially detected at fertilization continued to be expressed after they had been activated.

Such large-scale sequencing of cDNA libraries requires huge resources in terms of embryos, staff, cost, time and bioinformatics that it is unlikely to be undertaken in ruminant species. However, the continuing optimization of array technology for use in early bovine developmental stages (see Sirard *et al.*, this issue) promises to provide the large-scale analysis of stage-specific expression that is not feasible at present. Unfortunately, it is not only the lack of biological material or sequence information in ruminants that hinders array analysis, but also the nature of the available biological material. In cows, embryos produced *in vitro* are most easily obtained but even when embryos produced *in vivo* are used for studies in both cows and sheep, superovulation is almost always applied. The cyclic nature of follicle development and dominance means that ruminant embryos are harvested from a range of follicle sizes and health status states. Thus, the composition of maternally stored factors in oocytes from some smaller or atretic follicles is probably different from that of large, healthy follicles (Dieleman *et al.*, 2002). This heterogeneity is likely to result in direct inter-embryo variation in early gene expression patterns. If some of the heterogeneity is in epigenetic components needed to remodel chromatin and control transcription around ZGA, there may also

be indirect effects. This problem highlights the need to work towards using well-replicated single embryos instead of heterogeneous pools to yield interpretable data.

Using the fetus to identify key genes

In terms of developmental consequences of embryo technologies or environmental influences such as maternal nutrition, a complimentary approach to identifying the causes of observed phenotypes is to work back from the fetal stage where an obvious effect is observed. This can be particularly relevant in, for example, problems such as large offspring syndrome, in which only a proportion of the embryos transferred to recipients will result in detected phenotypes and so embryo expression studies (especially when pools are used) do not have the benefit of direct correlation with individual phenotype. Furthermore, the phenotypes observed can be very variable, with a range of organ defects among individuals that may or may not be associated with large size (Young *et al.*, 1998). As problems in a wide range of genes are likely to be responsible for the variations in phenotype, selecting a group of individuals with a consistent phenotype such as large size, abnormal kidneys, immunological problems and polyhydramnios and comparing affected relevant tissues with stage-matched controls provides a much more robust means of identifying the causes underlying those specific defects. However, for this approach it is essential to have sufficient control and abnormal individuals with the specific phenotype of interest to detect significant differences at the molecular level (see Young *et al.*, 2001). With each phenotype, gene disruptions can either be identified using a candidate gene approach if candidates from, for example, mouse knockouts or human disorders are available, or by broad screening methods such as microarrays, subtractive hybridization and differential display. One possible drawback to this approach is that identified gene disruptions may be well downstream of the initiating genes that were perturbed in the embryo and dissecting the initial pathways to provide markers in the embryo may be difficult. However, if, as suggested by our work on the fetal oversize phenotype, the initial event is an epigenetic disruption in a gene (for example, a DNA methylation change in imprinted genes), then the stable inheritance of such changes will facilitate confirmation that a change detected in the fetus was indeed initiated in the embryo. Although this approach is currently limited by the ability reproducibly to detect epigenetic changes at single loci in single embryos, the technologies to allow this are continually being improved.

Unfortunately, use of this approach to identify problems in a population of clones unselected for specific phenotypes, and without an appropriate number of stage-matched controls, is unlikely to be a sufficiently sensitive experimental design to detect appropriate differences. In identifying downregulated *IGF2R* as the likely cause of fetal oversize in large offspring syndrome, we would not have detected significant differences in the expression of this gene had we compared our controls with all of the (large and normal) fetuses derived from the perturbing embryo culture treatments, in which only 30–50% of the fetuses were oversized (Sinclair *et al.*, 1999; Young *et al.*, 2001). Examination of *IGF2R* in normal-sized fetuses derived from the same cultures (and in some case the same microdrops; Carolan *et al.*, 1998) that produced oversized fetuses demonstrated no significant difference between the AI-derived control fetuses (Young *et al.*, 2001). This finding provided further evidence that the *IGF2R* effect specifically contributed to the increase in fetal growth. Other phenotypes observed in this experiment, such as placental differences, abnormal kidney development and polyhydramnios, may well be due to disruptions in different genes and remain to be investigated. Furthermore, as some of the large fetuses we examined demonstrated *IGF2R*

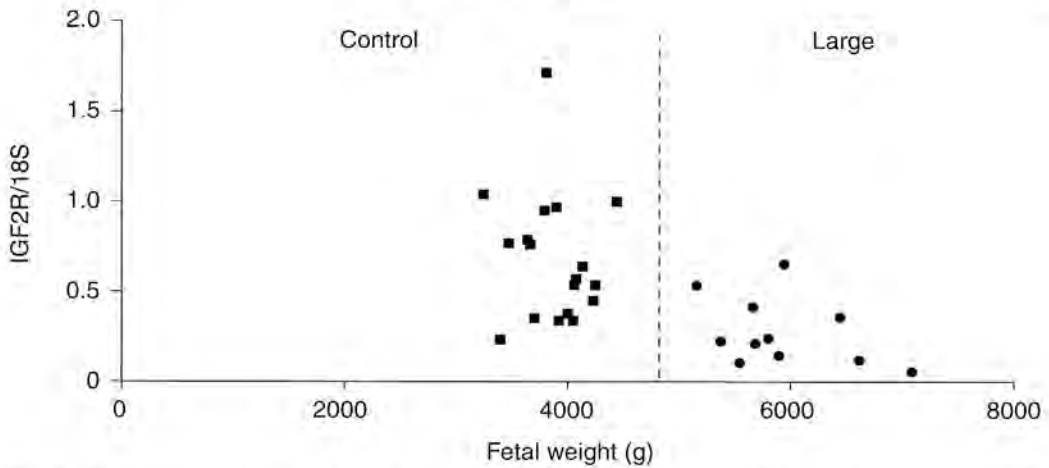


Fig. 5. Relationship between sheep liver expression of insulin-like growth factor 2 receptor (*IGF2R*) in the fetus (relative to *18S*) and fetal weight at day 125 of gestation after 5-day culture of zygotes to the blastocyst stage in serum-containing media.

expression within the range of the control group (Fig. 5), other genes may contribute to the oversized phenotype. We are currently investigating other fetal growth related candidates, both imprinted and non-imprinted, especially genes directly or indirectly involved in the complex IGF2 axis.

Even if an epigenetic change induced in the embryo is responsible for a fetal phenotype, a related change in gene expression may not be detected in all developmental stages. Although heritable throughout subsequent cell cycles, epigenetic changes in DNA methylation will only result in altered gene expression at stages when appropriate transcription factors for that gene are expressed. Thus, failure to detect a difference in expression of a candidate gene when examining only a single timepoint may not be conclusive of lack of change at another stage. This further highlights the need to combine gene expression analysis with protein, epigenetic and other functional measures.

Conclusion

Profiling the expression patterns of preimplantation stages has much to teach us about development. However, rigorous controls are required to ensure the validity of reported information since data generated in embryos nears the sensitivity limits of current PCR technology. Although RT-PCR provides a very powerful tool for analysing gene expression, it can be plagued by procedural pitfalls and inadequate experimental design and, thus, rigorous validations should be both applied and reported in publications.

There are some paradoxes between the expected need for tightly regulating stage-specific gene expression in preimplantation development and the observed non-stage-specific expression of many genes in ruminants. In addition, several genes normally associated with very specific later-stage tissues and processes are commonly detected in the early embryo. These unexpected observations are probably the result of the unique epigenetic remodelling that the preimplantation embryo undergoes to form a zygotic genome from sperm and egg genomes with very diverse epigenetic states. Thus, the relevance of specific induced gene expression changes, in terms of developmental consequences, requires additional information on gene function. Combining gene expression analysis with studies of protein expression, epigenetic status and possibly metabolic or other measures of function of the gene will be necessary

before a truly reliable system of developmental diagnostic markers can be developed for specific changes initiated in the preimplantation embryo. Further studies examining development of specific organ and physiological processes in the ruminant conceptus will expedite understanding of the increasing variety of changes likely to be induced in the embryo.

Work in our laboratory was supported by DEFRA, EC and BBSRC.

References

- Arney KL, Bao S, Bannister AJ, Kouzarides T and Surani MA (2002) Histone methylation defines epigenetic asymmetry in the mouse zygote *International Journal of Developmental Biology* **46** 317–320
- Bird A (2001) Methylation talk between histones and DNA *Science* **294** 2113–2115
- Bird A (2002) DNA methylation patterns and epigenetic memory *Genes and Development* **16** 6–21
- Boiani M, Eckardt S, Scholer HR and McLaughlin KJ (2002) Oct4 distribution and level in mouse clones: consequences for pluripotency *Genes and Development* **16** 1209–1219
- Carolan C, McEvoy TG, Staines ME, Sinclair KD, Robinson JJ and Broadbent PJ (1998) Variation in the expression of fetal oversize in day 125 ovine fetuses following transfer of *in vitro* cultured embryos *Theriogenology* **49** 229
- Constancia M, Hemberger M, Hughes J *et al.* (2002) Placental-specific IGF-II is a major modulator of placental and fetal growth *Nature* **417** 945–948
- Daniels R, Lowell S, Bolton V and Monk M (1997) Transcription of tissue-specific genes in human preimplantation embryos *Human Reproduction* **12** 2251–2256
- Daniels R, Hall VJ, French AJ, Korfiatis NA and Trounson AO (2001) Comparison of gene transcription in cloned bovine embryos produced by different nuclear transfer techniques *Molecular Reproduction and Development* **60** 281–288
- Denning C, Burl S, Ainslie A *et al.* (2001) Deletion of the alpha(1,3)galactosyl transferase (GGTA1) gene and the prion protein (PrP) gene in sheep *Nature Biotechnology* **19** 559–562
- De Sousa PA, Caveney A, Westhusin ME and Watson AJ (1998) Temporal patterns of embryonic gene expression and their dependence on oogenetic factors *Theriogenology* **49** 115–128
- Dieleman SJ, Hendriksen PJ, Viuff D *et al.* (2002) Effects of *in vivo* prematuration and *in vivo* final maturation on developmental capacity and quality of pre-implantation embryos *Theriogenology* **57** 5–20
- Gerlt JA (2002) 'Fishing' for the functional proteome *Nature Biotechnology* **20** 786–787
- Hales CN and Barker DJ (2001) The thrifty phenotype hypothesis *British Medical Bulletin* **60** 5–20
- Heby O (1995) DNA methylation and polyamines in embryonic development and cancer *International Journal of Developmental Biology* **39** 737–757
- Henderson JK, Draper JS, Baillie HA, Fishel S, Thomson JA, Moore H and Andrews PW (2002) Preimplantation human embryos and embryonic stem cells show comparable expression of stage-specific embryonic antigens *Stem Cells* **20** 329–337
- Kidder GM (2002) Trophectoderm development and function: the roles of Na⁺/K⁺-ATPase subunit isoforms *Canadian Journal of Physiology and Pharmacology* **80** 110–115
- Ko MS, Kitchen JR, Wang X *et al.* (2000) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development *Development* **127** 1737–1749
- Kwong WY, Wild AE, Roberts P, Willis AC and Fleming TP (2000) Maternal undernutrition during the preimplantation period of rat development causes blastocyst abnormalities and programming of postnatal hypertension *Development* **127** 4195–4202
- Latham KE and Schultz RM (2001) Embryonic genome activation *Frontiers in Biosciences* **6** D748–D759
- Lau MM, Stewart CE, Liu Z, Bhatt H, Rotwein P and Stewart CL (1994) Loss of the imprinted IGF2/cation-independent mannose 6-phosphate receptor results in fetal overgrowth and perinatal lethality *Genes and Development* **8** 2953–2963
- McCreath KJ, Howcroft J, Campbell KH, Colman A, Schnieke AE and Kind AJ (2000) Production of gene-targeted sheep by nuclear transfer from cultured somatic cells *Nature* **405** 1066–1069
- Ma J, Svoboda P, Schultz RM and Stein P (2001) Regulation of zygotic gene activation in the preimplantation mouse embryo: global activation and repression of gene expression *Biology of Reproduction* **64** 1713–1721
- Natale DR, De Sousa PA, Westhusin ME and Watson AJ (2001) Sensitivity of bovine blastocyst gene expression patterns to culture environments assessed by differential display RT-PCR *Reproduction* **122** 687–693
- Niemann H and Wrenzycki C (2000) Alterations of expression of developmentally important genes in preimplantation bovine embryos by *in vitro* culture conditions: implications for subsequent development *Theriogenology* **53** 21–34
- Niemann H, Wrenzycki C, Lucas-Hahn A, Brambrink T, Kues WA and Carnwath JW (2002) Gene expression patterns in bovine *in vitro*-produced and nuclear transfer-derived embryos and their implications for early development *Cloning Stem Cells* **4** 29–38

- Oswald J, Engemann S, Lane N, Mayer W, Olek A, Fundele R, Dean W, Reik W and Walter J (2000) Active demethylation of the paternal genome in the mouse zygote *Current Biology* **10** 475–478
- Paul CP, Good PD, Winer I and Engelke DR (2002) Effective expression of small interfering RNA in human cells *Nature Biotechnology* **20** 505–508
- Rees WD (2002) Manipulating the sulfur amino acid content of the early diet and its implications for long-term health *Proceedings of the Nutrition Society* **61** 71–77
- Renard JP (1998) Chromatin remodelling and nuclear reprogramming at the onset of embryonic development in mammals *Reproduction, Fertility and Development* **10** 573–580
- Rhind SM (2002) Endocrine disrupting compounds and farm animals: their properties, actions and routes of exposure *Domestic Animal Endocrinology* **23** 179–187
- Rizov D, Lonergan P, Boland MP, Arroyo-Garcia R, Pintado B, de la Fuente J and Gutierrez-Adan A (2002) Analysis of differential messenger RNA expression between bovine blastocysts produced in different culture systems: implications for blastocyst quality *Biology of Reproduction* **66** 589–595
- Rossant J (2001) Stem cells from the mammalian blastocyst *Stem Cells* **19** 477–482
- Schuringa JJ, van der Schaaf S, Vellenga E, Eggen BJ and Kruijer W (2002) LIF-induced STAT3 signaling in murine versus human embryonal carcinoma (EC) cells *Experimental Cell Research* **274** 119–129
- Sinclair KD, McEvoy TG, Maxfield EK, Maltin CA, Young LE, Wilmut I, Broadbent PJ and Robinson JJ (1999) Aberrant fetal growth and development after *in vitro* culture of sheep zygotes *Journal of Reproduction and Fertility* **116** 177–186
- Sinclair KD, Young LE, Wilmut I and McEvoy TG (2000) *In utero* overgrowth in ruminants following embryo culture: lessons from mice and a warning to men *Human Reproduction* **15** Supplement 5 68–86
- Symonds ME, Budge H, Stephenson T and McMillen IC (2001) Fetal endocrinology and development—manipulation and adaptation to long-term nutritional and environmental challenges *Reproduction* **121** 853–862
- Szymanski M and Barciszewski J (2002) Beyond the proteome: non-coding regulatory RNAs *Genome Biology* **3** Reviews 0005
- van Eijk MJ, van Rooijen MA *et al.* (1999) Molecular cloning, genetic mapping, and developmental expression of bovine POU5F1 *Biology of Reproduction* **60** 1093–1103
- Wallace JM (2000) Nutrient partitioning during pregnancy: adverse gestational outcome in overnourished adolescent dams *Proceedings of the Nutrition Society* **59** 107–117
- Wianny F and Zernicka-Goetz M (2000) Specific interference with gene function by double-stranded RNA in early mouse development *Nature Cell Biology* **2** 70–75
- Wilmut I, Beaujean N, De Sousa PA, Dinnyes A, King TJ, Paterson LA, Wells DN and Young LE (2002) Somatic cell nuclear transfer *Nature* **419** 583–586
- Young LE (2001) Imprinting of genes and the Barker hypothesis *Twin Research* **4** 307–317
- Young LE and Fairburn HR (2000) Improving the safety of embryo technologies: possible role of genomic imprinting *Theriogenology* **53** 627–648
- Young LE, Sinclair KD and Wilmut I (1998) Large offspring syndrome in cattle and sheep *Reviews of Reproduction* **3** 155–163
- Young LE, Fernandes K, McEvoy TG *et al.* (2001) Epigenetic change in IGF2R is associated with fetal overgrowth after sheep embryo culture *Nature Genetics* **27** 153–154
- Zhang X, Kidder GM, Zhang C, Khamsi F and Armstrong DT (1994) Expression of plasminogen activator genes and enzymatic activities in rat preimplantation embryos *Journal of Reproduction and Fertility* **101** 235–240