CSIRO PUBLISHING

www.publish.csiro.au/journals/rfd *Reproduction, Fertility and Development*, 2005, **17**, 37–45

Whole-genome approaches for large-scale gene identification and expression analysis in mammalian preimplantation embryos

James Adjaye

Max Planck Institute for Molecular Genetics, Department of Vertebrate Genomics, Ihnestrasse 73, D-14195 Berlin, Germany. Email: adjaye@molgen.mpg.de

Abstract. The elucidation, unravelling and understanding of the molecular basis of transcriptional control during preimplantion development is of utmost importance if we are to intervene and eliminate or reduce abnormalities associated with growth, disease and infertility by applying assisted reproduction. Importantly, these studies should enhance our knowledge of basic reproductive biology and its application to regenerative medicine and livestock production. A major obstacle impeding progress in these areas is the ability to successfully generate molecular portraits of preimplantation embryos from their minute amounts of RNA. The present review describes the various approaches whereby classical embryology fuses with molecular biology, high-throughput genomics and systems biology to address and solve questions related to early development in mammals.

Extra keywords: bioinformatics, cDNA libraries, embryonic cells, gene expression, microarrays, preimplantation development, RNA amplification, signalling pathways, systems biology.

Introduction

The past decade has seen a significant increase in nucleotide sequence information emanating from various large-scale sequencing efforts of various mammalian genomes. Because fertilisation, cellular growth and differentiation depend on specified genetic programmes, alterations of these signalling pathways can lead to abnormal growth and, ultimately, death. The development of various highly sensitive procedures for analysing transcriptomes at the level of single cells and also the availabilty of sequence information in the various databases (e.g. Mammalian Gene Collection (MGC), RIKEN, NCBI and Ensembl; see Table 1 for URLs), paves the route for deciphering the plethora of molecular mechanisms underlying early development.

Table 1. A list of databases and corresponding URLs

Database	URL
RIKEN	http://www.riken.go.jp
NCBI	http://www.ncbi.nlm.nih.gov/
ENSEMBL	http://www.ensembl.org
TIGR	http://www.tigr.org
KEGG	http://www.genome.jp/kegg
GenMapp	http://www.genmapp.org
Onto-Tools	http://vortex.cs.wayne.edu/Projects.html
GOAL	http://microarrays.unife.it
MGC	http://mgc.nci.nih.gov
GO consortium	http://www.geneontology.org
Genome Matrix	http://www.genome-matrix.org

The advent of reverse transcription–polymerase chain reaction (RT-PCR) has alleviated the problems associated with working with the limited amounts of embryonic RNA available for analysis. As an application, single-cell RT-PCR analysis has proved useful in examining gene expression in single cells and embryos (for reviews, see Daniels and Monk 1997; Daniels *et al*. 1998; Huntriss *et al*. 1998; Steuerwald *et al*. 1999; Wrenzycki *et al*. 2001, 2003; Lazzari *et al*. 2002). These approaches are limited by the number of genes that can be analysed in each individual cell. Because the total RNA content of a single mammalian cell is in the range of 20–40 pg and only 0.5–1.5 pg of this is mRNA (Roozemond 1976; Piko and Clegg 1982), any attempt at single-cell or embryo profiling must be capable of dealing with a total of 10^5–10^6 mRNA molecules; therefore, cDNA or mRNA amplification is unavoidable if we are to attempt to examine all, or at least the majority of, genes expressed in individual cells. For this purpose, a variety of global amplification protocols have been devised.

These amplification protocols include mRNA isolation coupled to whole-genome RT-PCR to generate large quantities of double-stranded cDNAs (uncloned libraries) from single embryos (Adjaye *et al*. 1997, 1998, 1999; Goto *et al*. 1999; Adjaye and Monk 2000; Smelt *et al*. 2000; Pitera *et al*. 2001; Ponsuksili *et al*. 2001; Tesfaye *et al*. 2003; Huntriss *et al*. 2004) and linear amplification protocols (Eberwine 1996; Brambrink *et al*. 2002; Hamatani *et al*. 2004*a*, 2004*b*; Wang *et al*. 2004; J. Adjaye *et al*. unpublished data). However, there are significant limitations and shortcomings to cDNA and mRNA amplification technologies that should be kept

 10.1071/RD04075 1031-3613/05/020037

in mind when designing studies that depend on these technologies. In both cases, the two main concerns are: (*1*) the representation of all transcripts present in the starting material in the final amplified material (i.e. loss of transcripts); and (*2*) the preservation of the relative abundances of the different transcripts. All the available technologies compromise these features to some degree.

Among the various gene expression detection methods that have also been applied to preimplantation development in human, mouse, bovine and pig are subtractive hybridisation (Rothstein *et al*. 1993; Roberts *et al*. 2000; Ponsuksili *et al*. 2002*a*, 2002*c*; Yao *et al*. 2003), expressed sequence tag (EST) generation (Rothstein *et al*. 1993; Adjaye *et al*. 1997, 1999; Sasaki *et al*. 1998; Ko *et al*. 2000; Ponsuksili *et al*. 2002*b*, 2002*c*; Uenishi *et al*. 2004), differential display (Zimmermann and Schultz 1994; Adjaye *et al*. 1998; Goto *et al*. 1999; Ponsuksili *et al*. 2002*a*, 2002*c*) serial analysis of gene expression (SAGE; Neilson *et al*. 2000; Stanton *et al*. 2002), *in situ* data mining (Rajkovic *et al*. 2001; Stanton and Green 2002; Sharov *et al*. 2003; Evsikov *et al*. 2004) and, finally, microarrays (Brambrink *et al*. 2002; Carter *et al*. 2003; Dobson *et al*. 2004; Hamatani *et al*. 2004*a*, 2004*b*; Tanaka and Ko 2004; Wang *et al*. 2004; J. Adjaye *et al*. unpublished data).

Although each technique has its own unique merit (for reviews, see Hurley *et al*. 2000; Ko 2001, 2004; Stanton *et al*. 2003), cDNA and oligonucleotide microarray technologies have gained precedence because they enable comparative whole-genome transcriptome analysis. In spite of this, there is a major drawback with the use of microarrays in comparison with the other technologies in that some low-abundant oocyte- and embryo-specific transcripts may not be represented as probes on most microarray platforms.

The present paper reviews the current methodologies for gene expression analysis during mammalian early development with special emphasis on microarray platforms.

Use of microarray platforms as a means of identifying genes and signalling pathways active during preimplantation development

A detailed description is given on how to detect differential expression between, for example, human unfertilised oocytes and embryos at the eight-cell stage of development to investigate the molecular mechanisms underlying embryonic genome activation (EGA), which is the transition from maternal to embryonic control of gene transcription (Latham and Schultz 2001).

This section describes a step-by-step guide as to how microarrays are fabricated using cDNA as probes, the generation of amplified RNA as labelled targets for hybridisation, image acquisition, data analysis and, finally, elucidation of signalling pathways based on the expression data using a systems biology approach (Dieterich *et al*. 2003; Ge *et al*. 2003).

Microarrays are typically composed of thousands of DNAs complementary to genes of interest (probes) and spotted onto pretreated (to enhance DNA binding) glass slides at defined positions. The thousands of cDNAs (cDNA arrays) to be spotted by an arrayer robot are derived from PCR-amplified cDNA from plasmid-transformed bacterial libraries (Schena *et al*. 1995; Lockhart *et al*. 1996; Adjaye *et al*. 2004) or chemically synthesised gene-specific oligonucleotides (oligonucleotide arrays). The bound DNAs are denatured and fixed onto the glass surface to enable hybridisation with complementary single-stranded labelled cDNA (Fig. 1).

RNA isolation, reverse transcription and whole-genome amplification

The ultimate aim in developmental biology and disease-related research is to carry out expression profiling with as few cells as possible, preferably single cells. However, a limitation of microarray-based expression profiling is the large amount of total RNA (8–20 μg) required to produce an adequate signal over noise. Therefore, the isolation of RNA from embryonic cells should incorporate an amplification procedure. The first step towards the generation of amplified RNA (aRNA) is mRNA isolation from lysed embryos or single cells using oligo-dT-linked magnetic beads first described by Adjaye *et al*. (1999) and illustrated in Fig. 2*a*.

In the second step, first-strand cDNA synthesis from RNA by reverse transcription is primed with an oligo(dT)-T7 polymerase promoter primer. Double-stranded cDNAs are then generated by PCR with primers specific for each end of the original first-strand cDNA molecules using SMART (Switching Mechanism at the 5′ end of RNA Template; Clontech, Palo Alto, CA, USA) as described in the context of preimplantation development (Brambrink *et al*. 2002; Hamatani *et al*. 2004*a*; Huntriss *et al*. 2004; Wang *et al*. 2004; J. Adjaye *et al*. unpublished data).

In a second reaction, the double-stranded cDNA is used as template for an *in vitro* transcription reaction using commercially available kits to generate million-fold amplifications of the original mRNA population of duplicate sets of unfertilised oocytes and single eight-cell preimplantation embryos. The reproducibility of the technique can be monitored by hybridising equal amounts of the duplicate sets of aRNA samples onto a chip and comparing the intensities of each gene by means of a scatter plot. A correlation coefficient (r^2) of 0.95 is normally attained using this procedure (Fig. 2*b*).

Target labelling, complex hybridisations and image acquisition

In a typical microarray experiment, two pools of RNA are compared (e.g. expression in unfertilised oocytes and eight-cell preimplantation embryos) to study embryonic genome activation (Ma *et al*. 2001; Yao *et al*. 2003). Each RNA pool is reverse transcribed using a nucleotide mixture including

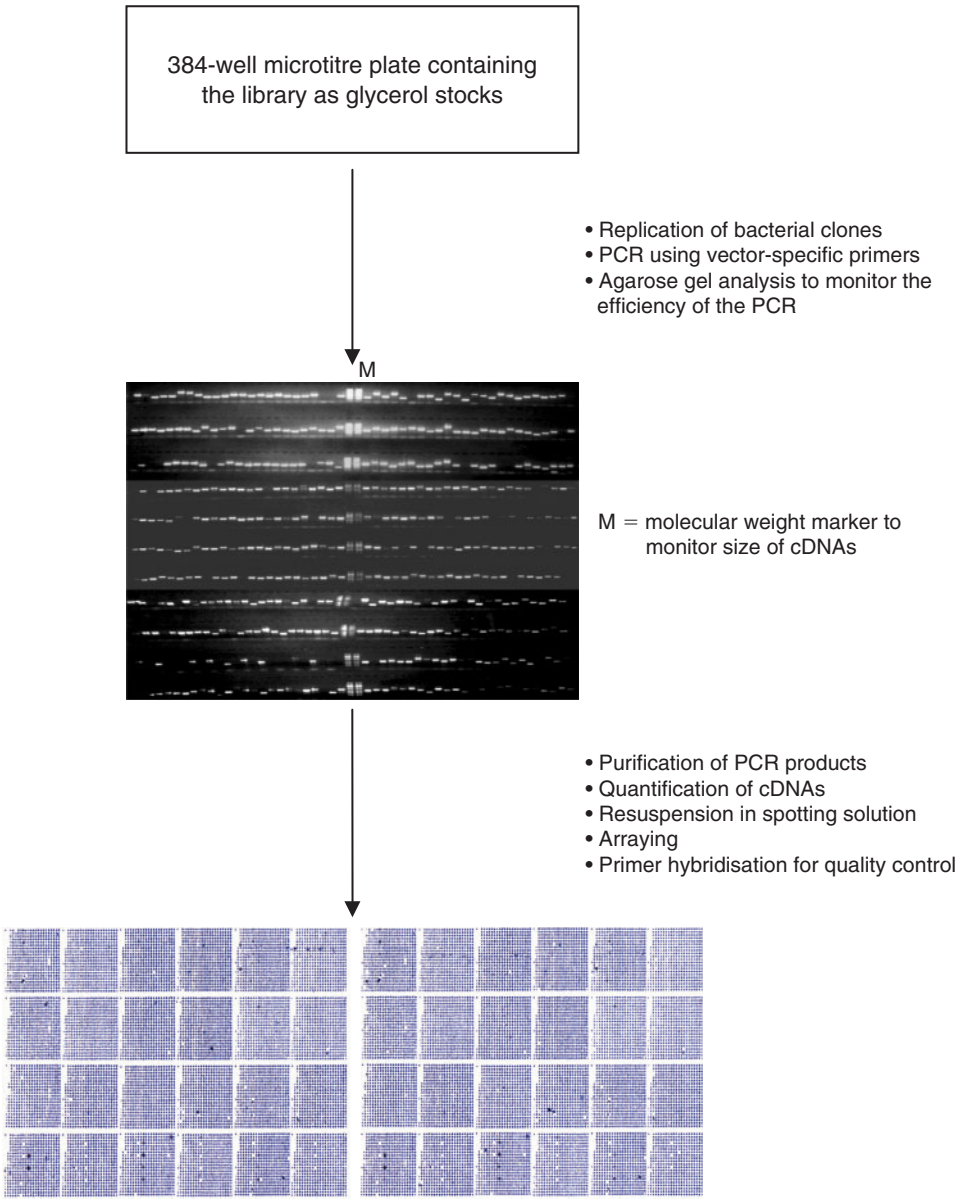

Fig. 1. Scheme illustrating the fabrication of a typical cDNA microarray chip composed of 15 500 genes (The Human ENSEMBL Chip) spotted in duplicate.

fluorescently labelled nucleotides (cyanine (Cy) 3 or Cy5). Following the reverse transcription reaction and purification to eliminate unincorporated fluorescent dyes, the labelled cDNAs are mixed and hybridised onto an appropriate chip. Hybridisation typically proceeds for at least 18 h in a moist chamber placed in a hybridisation oven. After hybridisation, the slides are washed under stringent conditions to eliminate unbound labelled targets. The slides are scanned in a confocal laser scanner to generate two microarray images that correspond to the emission of each of the two fluors used during the labelling reaction. To demonstrate differential expression, the two images can be overlayed using appropriate software to generate a pseudo-colour image with yellow spots signifying equal expression levels, red spots signifying overexpression in oocyte RNA labelled with Cy5 and green spots indicating overexpression in RNA derived from the eight-cell embryo and labelled with Cy3 (Fig. 2*c*).

Quantification, normalisation and analysis of data

Data analysis usually involves a two-step procedure. In the first step, data are normalised (using a set of housekeeping genes of almost constant expression levels across cell types or using the median of the global distribution of gene expression levels in each cell type as a reference) to eliminate extrinsic influencing factors and hybridisation artefacts not attributable to the probe–target interaction. For example,

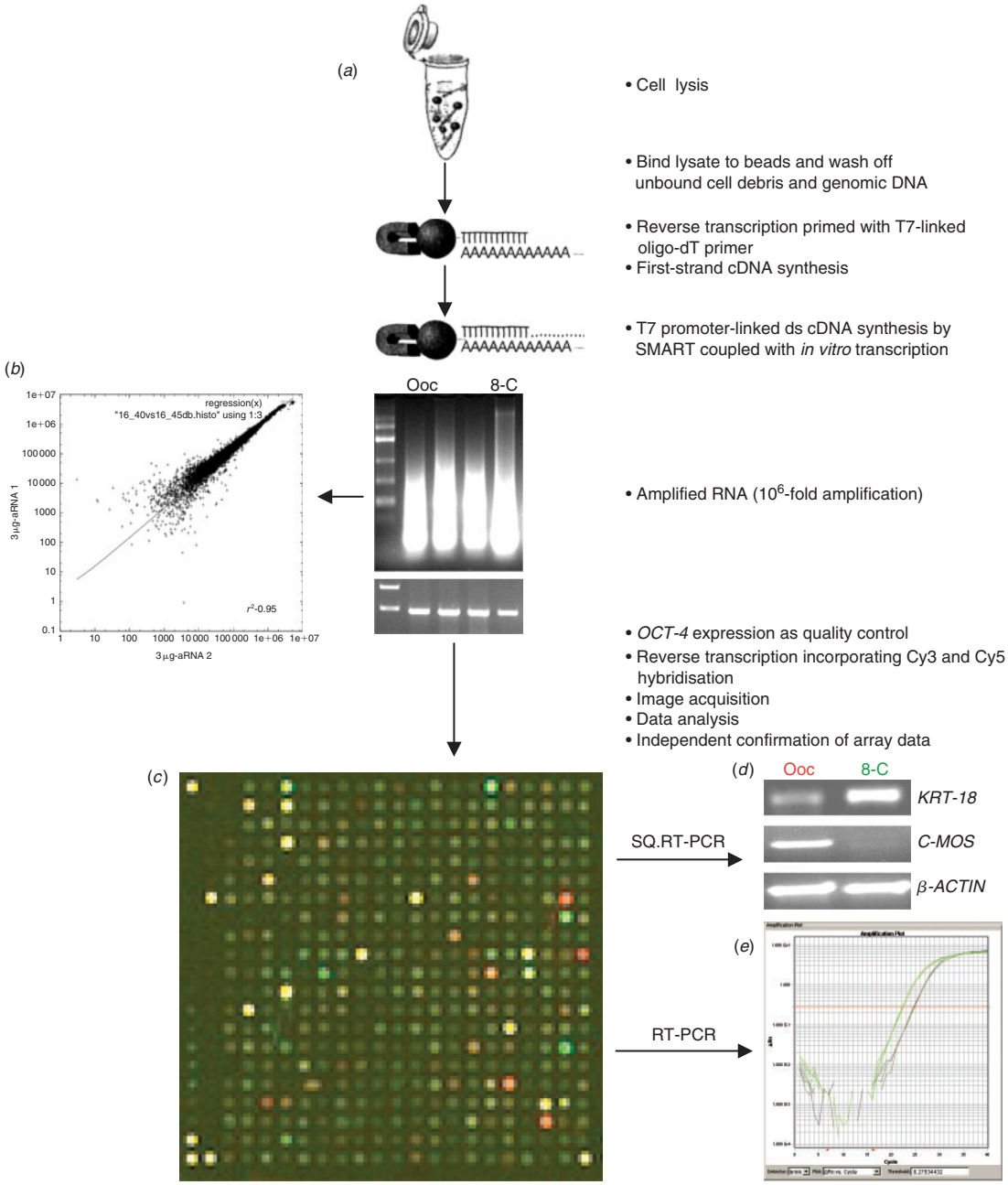

Fig. 2. An illustration of the generation of amplified RNA (aRNA), hybridisation and detection of differential gene expression. (*a*) Derivation of intact, high-quality aRNA with up to a million-fold amplification generated from duplicate sets of unfertilised oocytes and eight-cell preimplantation embryos. The quality of the RNA is confirmed by the expression of the embryonic marker gene *OCT4*. (*b*) The reproducibility of the amplification step is illustrated by the scatter-plot (using non-normalised intensities). A correlation coefficient (r^2) of 0.95 was obtained when two independently amplified RNA samples from eight-cell embryos were compared. A comparable value of 0.93 was obtained for the duplicate oocyte samples (data not shown). (*c*) A pseudo-colour image illustrating differential gene expression between unfertilised oocytes and eight-cell embryos. Red spots correspond to genes overexpressed in the oocyte, green spots to those overexpressed in the eight-cell embryo and yellow spots are of genes with comparable expression levels in the two samples. (*d*) Independent confirmation of array generated data by real-time polymerase chain reaction (PCR) and semiquantitative reverse transcription–polymerase chain reaction. The expression of *C-MOS*, a proto-oncogene encoding a serine-threonine kinase, was detected in the unfertilised oocyte only, thus suggesting a role in oocyte maturation (Adjaye *et al*. 1999). Although the cytoskeletal gene *KERATIN-18* (*KRT-18*) is more abundant in the compacted eight-cell embryo, the cytoskeletal housekeeping gene *β-actin* is expressed at similar levels. (*e*) An illustration of a typical output of the same experiment using real-time PCR. The amplification plots enable a quantitative comparison of transcript levels.

a non-linear normalisation procedure called 'LOWESS normalisation' (Cleveland *et al*. 1992) denotes a method that is descriptively known as locally weighted polynomial regression. Using this normalisation procedure, Locally Weighted Regression and Smoothing Scatter plots can be generated.

In the second step, several numerical characteristics are evaluated. These include signal detection values, signal reproducibility and statistical significance of differential expression. To judge whether the differential expression of a gene is statistically significant, various statistical tests can be applied. These may include Student's *t*-test, Welch test and Wilcoxon's rank sum test for each gene based on the series of signal intensities generated from the replicate hybridisations. Another means of assessing the significance of gene expression can be achieved by 'mixed models analysis' (Wolfinger *et al*. 2001). This method can accommodate a variety of experimental designs and can assess simultaneously significant differences between several types of biological samples under investigation.

For this reason, it is always important to perform at least three biological (RNA samples) and technical (independent hybridisations) replication experiments (Herwig *et al*. 2001; Quackenbush 2001; Yang *et al*. 2002; Adjaye *et al*. 2004). The relative intensity of Cy5/Cy3 is a reliable measure of the relative abundance of specific mRNAs in each sample.

To detect groups of coregulated genes in the developmental stages under investigation, cluster analyses are routinely used (Fig. 3*b*). There are various methods available: (*1*) hierarchical analysis, in which data are organised into a tree-like graph based on similarity (Eisen *et al*. 1998); (*2*) K-means, an algorithm that generates fixed-sized, flat classifications and clusters based on distance metrics for similarity (the specified K-value determines the number of clusters created; Herwig *et al*. 1999); (*3*) self-organising maps (SOMs), which are algorithms used to visualise and interpret large datasets in an unsupervised manner, thus allowing automatic data structuring (typical applications include visualisation of biological process within different cell types on a map such that their distribution indicates their relative similarities; Tamayo *et al*. 1999); and (*4*) gene shaving, which is a means of identifying distinct sets of genes with similar expression patterns (Hastie *et al*. 2000). The final classification according to biological processes will give insights into the molecular changes occurring during, for instance, the transition from totipotency (two-cell, four-cell and eight-cell stage of development) to pluripotency (inner cell mass (ICM) of the blastocyst) and embryonic genome activation.

Additional efforts should be placed on the validation of the calculated clusters by applying different numerical evaluation methods. For example, applying topological measures such as 'compactness', 'isolation' and bootstrapping. Isolation is a measure of separation between a given cluster and other clusters, whereas compactness measures the internal cohesion between the objects in the cluster. A cluster that is valid should be very isolated and compact. Bootstrap analysis is a statistical method for obtaining an estimate of error used to evaluate the reliability of a cluster. This is achieved by testing the reliability of a dataset by creating pseudoreplicate datasets by resampling. *In silico* approaches that can also be used include functional genomics data sources, such as gene ontologies, promoter information and sequence information. All these analysis modules are essential for maximal information retrieval and should be incorporated in the work flow of microarray data analysis (Fig. 3*b–d*).

Independent confirmation of expression data generated by microarray analysis

Because microarray analysis is a high-throughput procedure, false signals can be generated that may skew the final interpretation of the data. Therefore, genes that are identified as likely candidates for subsequent functional study need to be validated by alternative means. These may include one of the following: comparative semiquantitative RT-PCR, real-time PCR, Northern blots analyses, RNase protection assays or *in situ* hybridisation of the same embryos or tissues from which RNA samples were derived for the initial microarray based expression profiling. A confirmatory experiment for the differential expression of selected genes by real-time PCR or semiquantitative RT-PCR is illustrated (Fig. 2*d*,*e*).

Assimilation of miroarray data

A major limitation in the final assessment of microarray results is the lack of universal standards for the presentation of the huge datasets generated by independent investigators. In an effort to alleviate this shortfall, major efforts have been made to set standard data presentation and exchange through the establishment of a universally acceptable format, namely the minimum information about a microarray experiment (MIAME) format (Brazma *et al*. 2001).

Gene ontology and prediction of signalling pathways operative during embryogenesis

The sequencing of the genomes of human, mouse, rat and other organisms, in combination with high-throughput procedures based on microarrays and SAGE techniques, has started to yield vast amounts of expression data that are often stored in public databases. To understand gene function and dissect genetic pathways underlying preimplantation development, microarray datasets can be incorporated into available databases designed for this purpose. An example of this is the Gene Ontology (GO) consortium (http://www.geneontology.org), which has developed a systematic and standardized nomenclature for annotating genes in various organisms based on three main ontologies: molecular function, biological process and cellular component (Harris *et al*. 2004). The GO project provides the necessary information for the interpretation of the expression patterns

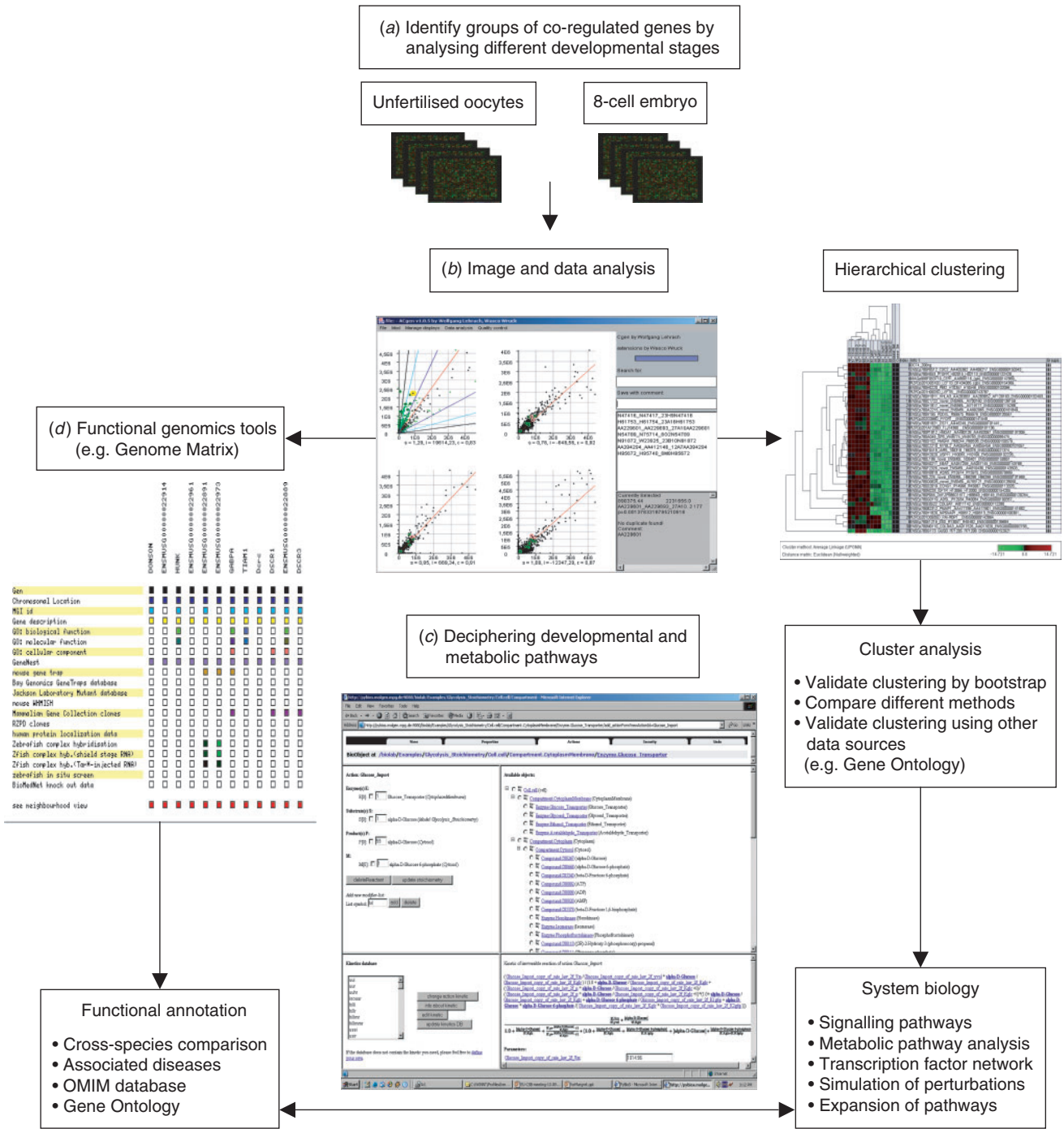

Fig. 3. A typical bioinformatics platform showing information work flow. (*a*) Hybridisation of oocyte and eight-cell-derived aRNAs on the ENSEMBL chip. The major phases of microarray data analysis are shown (*b–d*) and their connectivity in the work flow processes is indicated by arrows. Adherence to sound methodological principles ensures accurate data analysis, modelling and data storage. (*c*,*d*) Information flow via the internet maximises information retrieval from electronic databases and resources, such as the KEGG-pathway database (http://www.genome.ad.jp/kegg/), Gene Ontology database (http://www.godatabase.org) and the functional genomics database Genome Matrix (http://www.genome-matrix.org).

once each gene is associated with its related GO term(s). This, together with other integrated web-accessible data mining tools (see Table 1 for URLs) such as the Gene Ontology Automated Lexicon (GOAL) and Onto-Tools (Khatri *et al*. 2004; Volinia *et al*. 2004), can be used in association with databases such as the Kyoto Encyclopedia of Genes and Genomes (KEGG; Kanehisa *et al*. 2002) and Gene Microarray Pathway Profiler (GenMapp; Dahlquist *et al.* 2002) to map and

view gene expression data in the context of metabolic and signalling pathways operative during preimplantation development as follows: (*1*) murine embyronic genome activation (i.e. transition from maternal to embryonic control of gene expression; Evsikov *et al*. 2004); (*2*) delineation of the molecular mechanisms governing the differentiation of the morula to the ICM and trophectoderm (TE) of the human blastocyst (J. Adjaye *et al*. unpublished data); and (*3*) comparing the molecular changes resulting from *in vitro* and *in vivo* culture of embryos.

Cross-species comparative genome analysis using microarrays

Gene expression studies in mammalian species other than humans and rodents (e.g. non-human primates, bovine, sheep and porcine) will undoubtedly advance our understanding of human health and disease. However, a current lack of adequate sequence data and commercial cDNA and oligonucleotide microarrays keeps this technology beyond the reach of investigators working on these species. A potential solution to this problem is the use of cross-species hybridisations (i.e. human sequence-based arrays as tools for undertaking comparative genome expression studies using RNAs derived from other species; Evertsz *et al*. 2001; Miller *et al*. 2002; Adjaye *et al*. 2004; Ji *et al*. 2004). These studies represent critical areas of research directly related to the understanding of human diseases because non-human primates, bovine, sheep and porcine play crucial roles in organ transplantation, vaccine development, viral pathogenesis, gene therapy and a host of other human health-related technologies.

Concluding remarks

Analysis of gene expression patterns in early embryos using cDNA microarray technology should provide insights into operative gene regulatory networks and would thus be a major step forward in unravelling molecular mechanisms associated with developmental abnormalities resulting from the *in vitro* manipulation and culture of embryos (Niemann and Wrenzycki 2000). In addition, embryo-related technologies, such as the *in vitro* production of embryos for stem cell derivation and somatic nuclear transfer cloning, can only be achieved successfully once we have a clearer understanding of the molecular mechanisms underlying preimplantation development. Most importantly, our increased knowledge of transcriptional networks operative during embryogenesis should lead to advances in assisted reproduction and preimplantation genetic diagnosis in humans.

Acknowledgments

The author is greatful to Dr Pascal Kahlem and Thore Brink for reading the manuscript and providing valuable comments and suggestions; Dr Marie-Laure Yaspo and Dr Alia BenKahla for selecting the ENSEMBL clones, the RZPD for clone picking and sequencing, Dr Claus Hultschig for the arraying and Dr Ralf Herwig for advice on bioinformatics. Finally, the author thanks Professor Hans Lehrach for imparting his wisdom on functional genomics and making the undertaking of gene expression profiles during preimplantation development possible in his department. This work was funded by the German Ministry for Education and Research (BMBF) as part of the National Genome Research Network (NGFN).

References

Adjaye, J., and Monk, M. (2000). Transcription of homeobox-containing genes in cDNA libraries derived from human oocytes and preimplantation embryos. *Mol. Hum. Reprod.* **6**, 707–711. doi:10.1093/MOLEHR/6.8.707

Adjaye, J., Daniels, R., Bolton, V., and Monk, M. (1997). cDNA libraries from single human preimplantation embryos. *Genomics* **46**, 337–344. doi:10.1006/GENO.1997.5117

Adjaye, J., Daniels, R., and Monk, M. (1998). The construction of cDNA libraries from single human preimplantation embryos and their use in the study of gene expression during development. *J. Assist. Reprod. Genet.* **15**, 344–348. doi:10.1023/A:1022565115741

Adjaye, J., Bolton, V., and Monk, M. (1999). Developmental expression of specific genes detected in high quality cDNA libraries from single human preimplantation embryos. *Gene* **237**, 373–383. doi:10.1016/S0378-1119(99)00329-7

Adjaye, J., Herwig, R., Herrmann, D., Wruck, W., BenKahla, A., Brink, T. C., Nowak, M., Carnwath, J. W., Hultschig, C., Niemann, H., and Lehrach, H. (2004). Cross-species hybridisations of human and bovine orthologous genes on high density cDNA microarrays. *BMC Genomics* **5**, 83. doi:10.1186/1471-2164-5-83

Brambrink, T., Wabnitz, P., Halter, R., Klocke, R., Carnwath, J., Kues, W., Wrenzycki, C., Paul, D., and Niemann, H. (2002). Application of cDNA arrays to monitor mRNA profiles in single preimplantation mouse embryos. *Biotechniques* **33**, 376–385.

Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C. A., Causton, H. C., Gaasterland, T., Glennison, P., Holstege, F. C., Kim, I. F., Markowitz, V., Matese, J. C., Parkinson, H., Robinson, A., Sarkans, U., Schulze-Kremer, S., Stewart, J., Taylor, R., Vilo, J., and Vingron, M. (2001). Minimum information about a microarray experiment (MIAME): toward standards for microarray data. *Nat. Genet.* **29**, 365–371. doi:10.1038/NG1201-365

Carter, M. G., Hamatani, T., Sharov, A. A., Carmack, C. E., Qian, Y., Aiba, K., Ko, N. T., Dudekula, D. B., Brzoska, P. M., Hwang, S. S., and Ko, M. S. (2003). *In situ*-synthesized novel microarray optimized for mouse stem cell and early developmental expression profiling. *Genome Res.* **13**, 1011–1021. doi:10.1101/GR.878903

Cleveland, W. S., Grosse, E., and Shyu, W. M. (1992). Local regression models. In 'Statistical Models'. (Eds J. M. Chambers and T. J. Hastie.) pp. 309–376. (Wadsworth/Brooks Cole: Pacific Grove, CA, USA.)

Dahlquist, K. D., Salomonis, N., Vranizan, K., Lawlor, S. C., and Conklin, B. R. (2002). GenMAPP, a new tool for viewing and analysing microarray data on biological pathways. *Nat. Genet.* **31**, 19–20. doi:10.1038/NG0502-19

Daniels, R., and Monk, M. (1997). Gene expression in human preimplantation embryos. In 'Molecular Genetics of Early Human Development'. (Eds T. Strachan, S. Lindsay and D. I. Wilson.) pp. 155–170. (BIOS Scientific Publishers: Oxford, UK.)

Daniels, R., Adjaye, J., Bolton, V., and Monk, M. (1998). Detection of a novel splice variant of the HPRT gene in human oocytes and

preimplantation embryos. Implications for an RT-PCR based preimplantation diagnosis of Lesch Nyhan syndrome. *Mol. Hum. Reprod.* **4**, 785–789. doi:10.1093/MOLEHR/4.8.785

Dieterich, C., Herwig, R., and Vingron, M. (2003). Exploring potential targets of signaling pathways by predicting conserved transcription factor binding sites. *Bioinformatics* **19**, 50–56.

Dobson, A. T., Raja, R., Abeyta, M. J., Taylor, T., Shen, S., Haqq, C., and Pera, R. (2004). The unique transcriptome through day 3 of human preimplantation development. *Hum. Mol. Genet.* **13**, 1461–1470. doi:10.1093/HMG/DDH157

Eberwine, J. (1996). Amplification of mRNA populations using aRNA generated from immobilized oligo(dT)-T7 primed cDNA. *Biotechniques* **20**, 584–589.

Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl Acad. Sci. USA* **95**, 14 863–14 868. doi:10.1073/PNAS.95.25.14863

Evertsz, E. M., Au-Young, J., Ruvolo, M. V., Lim, A. C., and Reynolds, M. A. (2001). Hybridization cross-reactivity within homologous gene families on glass microarrays. *Biotechniques* **31**, 1182–1192.

Evsikov, A. V., De Vries, W. N., Peaston, A. E., Radford, E. E., Fancher, K. S., Chen, F. H., Blake, J. A., Bult, C. J., Latham, K. E., Solter, D., and Knowles, B. B. (2004). Systems biology of the 2-cell mouse embryo. *Cytogenet. Genom. Res.* **105**, 2–4.

Ge, H., Walhout, A. J. M., and Vidal, M. (2003). Integrating 'omic' information: a bridge between genomics and systems biology. *Trends Genet.* **19**, 551–560. doi:10.1016/J.TIG.2003.08.009

Goto, T., Adjaye, J., Rodeck, C., and Monk, M. (1999). Identification of human primordial germ cell specific transcribed genes by differential display. *Mol. Hum. Reprod.* **5**, 851–860. doi:10.1093/MOLEHR/5.9.851

Hamatani, T., Carter, M. G., Sharov, A. A., and Ko, M. S. H. (2004*a*). Dynamics of global gene expression changes during mouse preimplantation development. *Dev. Cell* **6**, 117–131. doi:10.1016/S1534-5807(03)00373-3

Hamatani, T., Daikoku, T., Wang, H., Matsumoto, H., Carter, M. G., Ko, M. S., and Dey, S. K. (2004*b*). Global gene expression analysis identifies molecular pathways distinguishing blastocyst dormancy and activation. *Proc. Natl Acad. Sci. USA* **101**, 10 326–10 331. doi:10.1073/PNAS.0402597101

Harris, M. A., Clark, J., Ireland, A., Lomax, J., Ashburner, M., *et al*. (2004). Gene Ontology Consortium. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* **32**, D258–D261. doi:10.1093/NAR/GKH066

Hastie, T., Tibshirani, R., Eisen, M. B., Alizadeh, A., Levy, R., Staudt, L., Chan, W. C., Botstein, D., and Brown, P. (2000). 'Gene shaving' as a method for identifying distinct sets of genes with similar expression patterns. *Genome Biol.* **1**, 3.1–3.21.

Herwig, R., Poustka, A. J., Müller, C., Bull, C., Lehrach, H., and O'Brien, J. (1999). Large-scale clustering of cDNA fingerprinting data. *Genome Res.* **9**, 1093–1105. doi:10.1101/GR.9.11.1093

Herwig, R., Aanstad, P., Clark, M., and Lehrach, H. (2001). Statistical evaluation of differential expression on cDNA nylon arrays with replicated experiments. *Nucleic Acids Res.* **29**, 117E. doi:10.1093/NAR/29.23.E117

Huntriss, J., Daniels, R., Bolton, V. N., and Monk, M. (1998). Imprinted expression of SNRPN in human preimplantation embryos. *Am. J. Hum. Genet.* **63**, 1009–1014. doi:10.1086/302039

Huntriss, J., Hinkins, M., Oliver, B., Harris, S. E., Beazley, J. C., Rutherford, A. J., Gosden, R. G., Lanzendorf, S. E., and Picton, H. M. (2004). Expression of mRNAs for DNA methyltransferases and methyl-CpG-binding proteins in the human female germ line, preimplantation embryos, and embryonic stem cells. *Mol. Reprod. Dev.* **67**, 323–336. doi:10.1002/MRD.20030

Hurley, J., Huntriss, J., and Adjaye, J. (2000). Molecular approaches to the study of gene expression during human preimplantation development. *Hum. Fertil. (Camb.)* **3**, 48–51.

Ji, W., Zhou, W., Gregg, K., Yu, N., and Davis, S. (2004). A method for cross-species gene expression analysis with high-density oligonucleotide arrays. *Nucleic Acids Res.* **32**, e93. doi:10.1093/NAR/GNH084

Kanehisa, M., Goto, S., Kawashima, S., and Nakaya, A. (2002). The KEGG databases at GenomeNet. *Nucleic Acids Res.* **30**, 42–46. doi:10.1093/NAR/30.1.42

Khatri, P., Bhavsar, P., Bawa, G., and Draghici, S. (2004). Onto-Tools: an ensemble of web-accessible, ontology-based tools for the functional design and interpretation of high-throughput gene expression experiments. *Nucleic Acids Res.* **32**, D449–D456. doi:10.1093/NAR/GKH086

Ko, M. S. H. (2001). Embryogenomics: developmental biology meets genomics. *Trends Biotechnol.* **19**, 511–518. doi:10.1016/S0167-7799(01)01806-6

Ko, M. S. H. (2004). Embryogenomics of pre-implantation mammalian development: current status. *Reprod. Fertil. Dev.* **16**, 79–85. doi:10.1071/RD03080

Ko, M. S. H., Kitchen, J. R., Wang, X., Threat, T. A., Hasegawa, A., Sun, T., Grahovac, M. J., Kargul, G. J., Lim, M. K., Cui, Y., Sano, Y., Tanaka, T., Liang, Y., Mason, S., Paonessa, P. D., Sauls, A. D., DePalma, G. E., Sharara, R., Rowe, L. B., Eppig, J., Morrell, C., and Doi, H. (2000). Large scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development. *Development* **127**, 1737–1749.

Latham, K. E., and Schultz, R. M. (2001). Embryonic genome activation. *Front. Biosci.* **6**, D748–D759.

Lazzari, L., Wrenzycki, C., Herrmann, D., Duchi, R., Kruip, T., Niemann, H., and Galli, C. (2002). Cellular and molecular deviations in bovine *in vitro* produced embryos are related to the Large Offspring syndrome. *Biol. Reprod.* **67**, 767–775.

Lockhart, D. J., Dong, H., Bryne, M. C., Folletie, M. T., Gallo, M. V., Chee, M. S., Mittmann, M., Want, C., Kobayashi, M., Horton, H., and Brown, E. L. (1996). DNA expression monitoring by hybridisation to high-density oligonucleotide arrays. *Nat. Biotechnol.* **14**, 1675–1680. doi:10.1038/NBT1296-1675

Ma, J., Svoboba, P., Schultz, R. M., and Stein, P. (2001). Regulation of zygotic gene activation in the preimplantation mouse embryo: global activation and repression of gene expression. *Biol. Reprod.* **64**, 1713–1721.

Miller, N. A., Gong, Q., Bryan, R., Ruvolo, M., Turner, L. A., and LaBrie, S. T. (2002). Cross-hybridization of closely related genes on high-density macroarrays. *Biotechniques* **32**, 620–625.

Neilson, L., Andalibi, A., Kang, D., Coutifaris, C., Strauss, F. J., III, Stanton, J. A., and Green, D. P. (2000). Molecular phenotype of the human oocyte by PCR-SAGE. *Genomics* **63**, 13–24. doi:10.1006/GENO.1999.6059

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. doi:10.1016/S0093-691X(99)00237-X

Piko, L., and Clegg, K. B. (1982). Quantitative changes in total RNA, total poly(A), and ribosomes in early mouse embryos. *Dev. Biol.* **89**, 362–378.

Pitera, J., Milla, P., Scambler, P., and Adjaye, J. (2001). Cloning of HOXD1 from unfertilised human oocytes and expression analyses during murine oogenesis and embryogenesis. *Mech. Dev.* **109**, 377–381. doi:10.1016/S0925-4773(01)00530-5

Ponsuksili, S., Wimmers, K., Adjaye, J., and Schellander, K. (2001). Expression of homeobox-containing genes in cDNA libraries derived from bovine unfertilised oocytes and preimplantation embryos. *Mol. Reprod. Dev.* **60**, 297–301. doi:10.1002/MRD.1091

Ponsuksili, S., Tesfaye, D., El-Halawany, N., Schellander, K., and Wimmers, K. (2002*a*). Stage-specific expressed sequence tags obtained during preimplantation bovine development by differential display RT-PCR and suppression subtractive hybridization. *Prenat. Diagn.* **12**, 1135–1142.

Ponsuksili, S., Wimmers, K., Adjaye, J., and Schellander, K. (2002*b*). A source for expression profiling in single preimplantation bovine embryos. *Theriogenology* **57**, 1611–1624. doi:10.1016/S0093-691X(02)00661-1

Ponsuksili, S., Tesfaye, D., El-Halawany, N., Schellander, K., and Wimmers, K. (2002*c*). Stage-specific expression sequence tags obtained during preimplantation bovine development by differential display RT-PCR and suppression subtractive hybridisation. *Prenat. Diagn.* **22**, 1135–1142. doi:10.1002/PD.501

Quackenbush, J. (2001). Computational analysis of microarray data. *Nat. Rev. Genet.* **2**, 418–427. doi:10.1038/35076576

Rajkovic, A., Yan, M. S. C., Klysik, M., and Matuk, M. (2001). Discovery of germ cell-specific transcripts by expressed sequence tag database analysis. *Fertil. Steril.* **76**, 550–554. doi:10.1016/S0015-0282(01)01966-5

Roberts, C., Barnes, F. L., Hue, I., and Sirard, M. A. (2000). Subtractive hybridisation used to identify mRNA associated with the maturation of bovine oocytes. *Mol. Reprod. Dev.* **57**, 167–175. doi:10.1002/1098-2795(200010)57:2<167::AID-MRD8>3.0.CO;2-P

Roozemond, R. C. (1976). Ultramicrochemical determination of nucleic acids in individual cells using the Zeiss UMSP-I microspectrophotometer. Application to isolated rat hepatocytes of different ploidy classes. *Histochem. J.* **8**, 625–638.

Rothstein, J. L., Johnson, D., Jessee, J., Skowronski, J., DeLoia, J. A., Solter, D., and Knowles, B. B. (1993). Construction of primary and subtracted cDNA libraries from early embryos. *Methods Enzymol.* **225**, 587–610. doi:10.1016/0076-6879(93)25038-4

Sasaki, N., Nagaoka, S., Itoh, M., Izawa, M., Konno, H., Carninci, P., Yoshiki, A., Kusakabe, M., Moriuchi, T., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y. (1998). Characterization of gene expression in mouse blastocyst using single-pass sequencing of 3995 clones. *Genomics* **49**, 167–179. doi:10.1006/GENO.1998.5209

Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467–470.

Sharov, A. A., Piao, Y., Matoba, R., Dudekula, D. B., Qian, Y., *et al*. (2003). Transcriptome analysis of mouse stem cells and early embryos. *PLoS Biol.* **1**, E74. doi:10.1371/JOURNAL.PBIO.0000074

Smelt, V. A., Upton, A., Adjaye, J., Payton, M., Boukouvala, S., Johnson, N., Mardon, H. J., and Sim, E. (2000). Expression of arylamine *N*-acetyltransferases in pre-term placentals and in human pre-implantation embryos. *Hum. Mol. Genet.* **9**, 1101–1107. doi:10.1093/HMG/9.7.1101

Stanton, J. L., and Green, D. P. (2002). A set of 1542 mouse blastocyst and pre-blastocyst genes with well-matched human homologues. *Mol. Hum. Reprod.* **8**, 149–166. doi:10.1093/MOLEHR/8.2.149

Stanton, J. L., Bascand, M., Fisher, L., Quinn, M., Macgregor, A., and Green, D. P. (2002). Gene expression profiling of human GV oocytes: an analysis of a profile obtained by Serial Analysis of Gene Expression (SAGE). *J. Reprod. Immunol.* **53**, 193–201. doi:10.1016/S0165-0378(01)00093-6

Stanton, J. L., Macgregor, A. B., and Green, D. P. (2003). Gene expression in the mouse preimplantation embryo. *Reproduction* **125**, 457–468. doi:10.1530/REP.0.1250457

Steuerwald, N., Cohen, J., Herrera, R. J., and Brenner, C. A. (1999). Analysis of gene expression in single oocytes and embryos by real-time rapid cycle fluorescence monitored RT-PCR. *Mol. Hum. Reprod.* **5**, 1034–1039. doi:10.1093/MOLEHR/5.11.1034

Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E. S., and Golub, T. R. (1999). Interpreting patterns of gene expression with self-organizing maps: methods and applications to hematopoietic differentiation. *Proc. Natl Acad. Sci. USA* **96**, 2907–2912. doi:10.1073/PNAS.96.6.2907

Tanaka, T. S., and Ko, M. S. H. (2004). A global view of gene expression in the preimplantation mouse embryo: morula versus blastocyst. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **115**, S85–S91. doi:10.1016/J.EJOGRB.2004.01.026

Tesfaye, D., Ponsuksili, S., Wimmers, K., Gilles, M., and Schellander, K. (2003). Identification and quantification of differentially expressed transcripts in *in vitro*-produced bovine preimplantation stage embryos. *Mol. Reprod. Dev.* **66**, 105–114. doi:10.1002/MRD.10338

Uenishi, H., Eguchi, T., Suzuki, K., Sawazaki, T., Toki, D., Shinkai, H., Okumura, N., Hamasima, N., and Awata, T. (2004). PEDE (Pig EST Data Explorer): construction of a database for ESTs derived from porcine full-length cDNA libraries. *Nucleic Acids Res.* **32**, D484–D488. doi:10.1093/NAR/GKH037

Volinia, S., Evangelisti, R., Francioso, F., Arcelli, D., Carella, M., and Gasparini, P. (2004). GOAL: automated Gene Ontology analysis of expression profiles. *Nucleic Acids Res.* **32**, D492–D499.

Wang, Q. T., Piotrowska, K., Ciemerych, M. A., Milenkovic, L., Scott, M. P., Davis, R. W., and Zernicka-Goetz, M. (2004). A genome-wide study of gene activity reveals developmental signaling pathways in the preimplantation mouse embryo. *Dev. Cell* **6**, 133–144. doi:10.1016/S1534-5807(03)00404-0

Wolfinger, R. D., Gibson, G., Wolfinger, E. D., Bennett, L., Hamadeh, H., Bushel, P., Afshari, C., and Paules, R. S. (2001). Assessing gene significance from cDNA microarray expression data via mixed models. *J. Comput. Biol* **8**, 625–637. doi:10.1089/106652701753307520

Wrenzycki, C., Herrmann, D., Keskintepe, L., Martins, A., Sirisathien, S., Brackett, B., and Niemann, H. (2001). Effects of basic culture medium and protein supplementation on mRNA expression in preimplantation bovine embryos. *Hum. Reprod.* **16**, 893–901. doi:10.1093/HUMREP/16.5.893

Wrenzycki, C., Herrmann, D., and Niemann, H. (2003). Timing of blastocyst expansion affects spatial mRNA expression patterns of genes in bovine blastocysts produced *in vitro*. *Biol. Reprod.* **68**, 2073–2080.

Yang, Y. H., Dudoit, S., Luu, P., Lin, D. M., Peng, V., Ngai, J., and Speed, T. P. (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* **30**, 15E. doi:10.1093/NAR/30.4.E15

Yao, Y.-Q., Xu, J.-S., Lee, W. M., Yeung, W. S. B., and Lee, K.-F. (2003). Identification of mRNAs that are up-regulated after fertilization in the murine zygote by suppression subtractive hybridization. *Biochem. Biophys. Res. Commun.* **304**, 60–66. doi:10.1016/S0006-291X(03)00537-0

Zimmermann, I. W., and Schultz, R. M. (1994). Analysis of gene expression in the preimplantation embryo: use of mRNA differential display. *Proc. Natl Acad. Sci. USA* **91**, 5456–5460.