Dynamic changes in gene expression during human early embryo development: from fundamental aspects to clinical applications.

Said Assou, Imène Boumela, Delphine Haouzi, Tal Anahory, Hervé Dechaud, John de Vos, Samir Hamamah

To cite this version:


HAL Id: inserm-00512388
https://www.hal.inserm.fr/inserm-00512388
Submitted on 30 Aug 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
**Dynamic changes in gene expression during human early embryo development: from fundamental aspects to clinical applications**

Said Assou 1,2, Imène Boumela 1,2, Delphine Haouzi 1, Tal Anahory 2, Hervé Dechaud 2, John De Vos 1,2, Samir Hamamah 1,2* 

1 Biothérapie des cellules souches normales et cancéreuses INSERM : U847, Institut de recherche en biothérapie, Université Montpellier I, CHRU Montpellier, IRB - CHRU Saint-Eloi 80 Avenue Augustin Fliche 34295 MONTPELLIER Cedex 5, FR 
2 Unité d'AMP/DPI, Département de Médecine et biologie de la reproduction CHRU Montpellier, Hôpital Arnaud de Villeneuve, FR 
3 Unité de thérapie cellulaire CHRU Montpellier, FR

*Correspondence should be addressed to: Samir Hamamah <s-hamamah@chu-montpellier.fr>

Abstract

**BACKGROUND**

The first week of human embryonic development comprises a series of events that change highly specialized germ cells into undifferentiated human embryonic stem cells (hESCs) that display an extraordinarily broad developmental potential. The understanding of these events is crucial to the improvement of the success rate of in vitro fertilization (IVF). With the emergence of new technologies such as Omics, the gene expression profiling of human oocytes, embryos, and hESCs has been performed and generated a flood of data related to the molecular signature of early embryo development.

**METHODS**

In order to understand the complex genetic network that controls the first week of embryo development, we performed a systematic review and study of this issue. We performed a literature search using PubMed and EMBASE to identify all relevant studies published as original articles in English up to March 2010 (n=165). We also analyzed the transcriptome of human oocytes, embryos and hESCs.

**RESULTS**

Distinct sets of genes were revealed by comparing the expression profiles of oocytes, embryos on day-3 and hESCs, that are associated with totipotency, pluripotency and reprogramming properties, respectively. Known components of two signalling pathways (WNT and transforming growth factor-beta) were linked to oocyte maturation and early embryonic development.

**CONCLUSION**

Omics analysis provides tools for understanding the molecular mechanisms and signalling pathways controlling early embryonic development. Furthermore, we discuss the clinical relevance of using a non-invasive molecular approach to embryo selection for the single embryo transfer (SET) program.

**MESH Keywords** Embryonic Development; genetics; Embryonic Stem Cells; metabolism; Female; Gene Expression Profiling; Gene Expression Regulation, Developmental; Humans; Maternal Age; Oocytes; metabolism; Polycystic Ovary Syndrome; genetics; metabolism; Pregnancy; Pregnancy Outcome

**Author Keywords** oocyte genes; embryo genes; embryo stem cells; microarray; non-invasive approach

**Introduction**

In humans, early embryo development is a complex process that consists of sequential maturation events of the oocyte, fertilization and embryo growth (4-cell, 8-cell, morula and blastocyst). Indeed, oocytes and spermatozoa are atypical, highly specialized cell types compared to somatic cells. Yet, after fertilization, a zygote is formed, the ultimate totipotent cell that can be considered to be the ultimate undifferentiated cell type, as it gives rise to all cell types and live offspring. Totipotency persists for the very first cell doublings, from the single cell and zygote to at least the 4-cell pre-embryo. Initiation of transcription in the newly-formed embryonic genome reportedly occurs at the 4- and 8-cell stage (Braude et al., 1988). A few studies have documented key events that follow fertilization in humans such as decreases in abundance of individual mRNAs (Taylor et al., 2001), overall patterns of gene expression in individual human oocyte and preimplantation embryo stage (Dobson et al., 2004), and more generalizable transcription profiles of pooled morphologically normal human oocytes and embryos (Zhang et al., 2009). These data provide fundamental resources for understanding the genetic network controlling the early stages of human embryo development. At the morula stage, the human pre-embryo undergoes compaction, with the loss of the cellular distinction between the blastomeres. This is followed by the high expression of genes involved in tight intercellular junctions. Then, when the human embryo has reached the 35- to 65-cell stage, the trophoblast cells pump nutrients and water into the interior of the cell-sphere, forming a blastocyst, within which the inner cell mass (ICM) cells continue to proliferate. It is at the blastocyst...
stage that ICM cells are harvested for the derivation of embryonic stem (ES) cells (Reubinoff et al., 2000; Thomson et al., 1998). Pluripotent cells can be isolated, adapted and propagated indefinitely in vitro in an undifferentiated state as human embryonic stem cells (hESCs). hESCs are remarkable in their ability to generate virtually any cell type; hence they carry many hopes for cell therapy. Human mature MII oocytes, as well as hESCs, are able to achieve the feat of cell reprogramming towards pluripotency, either by somatic cell nuclear transfer or by cell fusion, respectively (Cowan et al., 2005; Hochederliniger et al., 2004; Saito et al., 2008; Sung et al., 2006).

Knowledge gathered from the field of ESCs was at the heart of the groundbreaking discovery that both mouse and human somatic cells can be reprogrammed into a pluripotent state by defined factors. The field of induced pluripotent stem cells (iPSCs) has marked a new era in stem cell research, and has also provided data pertinent to improving the understanding of pluripotency (Takahashi et al., 2007; Yu et al., 2007). Since totipotency and pluripotency are at the center of early embryonic development, comprehending their molecular mechanisms is crucial to our understanding of reproductive biology and to regenerative medicine.

DNA microarray technology is one of the most widely used and potentially revolutionary research tools derived from the human genome project (Venter et al., 2001). This technology provides a unique tool for the determination of gene expression at the level of messenger RNA (mRNA) on a genomic scale. Its capacity has opened new paths for biological investigation and generated a large number of applications (Stoughton, 2005), including the analysis of the transcriptomic profiles from early embryo development and the identification of new prognostic biomarkers for use in the in vitro fertilization (IVF) program (Hamatan et al., 2004a; Assou et al., 2008). The application of microarray technology to the analysis of human oocyte and early embryo cleavage poses specific challenges associated with the picogram levels of mRNA in a single oocyte and embryo, the plasticity of the embryonic transcriptome, the scarcity of the material and ethical considerations.

In this review we analyzed data from published reports and include our own data to define the genomic profile during early embryonic development. Once the molecular signature is established, biomarkers can be identified on a large scale, validated and tested prior to clinical applications for embryo selection to improve single embryo transfer (SET) programs. Such a research workflow should provide an understanding of the molecular and cellular mechanisms of oocyte and embryo function, as well as important insights for the development of diagnostic tests for oocyte quality and embryo competence.

Methods

A search of English-language publications from 4 computerized databases (PubMed, EMBASE, Science Direct and Ingenta Connect) was undertaken. We built an expression compendium by combining U133 Plus 2.0 (Affymetrix, Santa Clara, USA) microarray data from the US National Center for Biotechnology Information, from the Gene Expression Omnibus (GEO) through the provisional accession numbers (the GSE7234, GSE7896, GSE11450, and GSE7896 series), and from 15 samples from our own laboratory (5 pooled oocytes, 2 embryos on day-3, and 8 hESC lines). This study received institutional review board (IRB) approval, as well as French authority: Agence de la Biomédecine (ABM). The information gathered was analyzed through the GCOS 1.2 software (Affymetrix), using the default analysis settings and global scaling as the first normalization method, with a trimmed mean target intensity value (TGT) for each array arbitrarily set to 100, in agreement with the MIAME recommendations (Brazma et al., 2001). Principle component analysis (PCA) was performed using GeneSpring® software to provide a global view of how the various sample groups were related. Hierarchical clustering was carried out with CLUSTER and TREEVIEW software (Eisen et al., 1998). To uncover functional biological networks, we imported gene expression signatures into the Ingenuity Pathways Analysis (IPA) Software (Ingenuity Systems, Redwood City, CA, USA).

Results

Global analysis of the human oocytes gene expression profile

The human oocyte is equipped with an extraordinary biological competence. It can be fertilized and, at the same time, is able to reprogram the sperm chromatin, giving rise to totipotency and driving towards early embryonic development. However, the molecular mechanisms underlying oocyte competence are still largely unknown. Throughout follicle growth, the nuclear maturation of oocytes is arrested at meiotic prophase I. However, the cytoplasm progresses through a series of maturational stages, accumulating mRNAs and proteins that enable the oocyte to be fertilized and to progress through the first cleavage divisions until embryonic genes begin to be expressed. To improve the understanding of oocyte maturation and competence, microarray technology permits the assessment of changes of the global gene expression profiles in human and mouse oocytes (see Table I).

Gene expression of the oocyte during nuclear maturation

Nuclear maturation is a fundamental developmental event that has been extensively studied for several years. However, the molecular mechanisms underlying this critical process remain incompletely understood. Many exciting biological questions surrounding this event persist. Differences in the global mRNA transcript profile have been reported between mature oocytes (metaphase II (MII)) and immature oocytes (germinal vesicle (GV) and metaphase I (MI)) in both mice and human (Assou et al., 2006; Cui et al., 2007; Wang, Q. T. et al., 2004; Yoon et al., 2006; Zhang et al., 2007). The genome-wide gene expression of human immature and mature oocytes revealed minor
modifications in the transcript profiles between GV and MI oocytes, while MII oocytes showed overexpression of more than 400 genes and a striking underexpression of more than 800 genes in comparison with less mature oocytes (Assou et al., 2006 ). The major modifications of the transcript profile appear to occur during the final stage of oocyte maturation. Components of the metaphase-promoting factor, the anaphase-promoting complex, and a number of oocyte specific genes were overexpressed in MII oocyte (Assou et al., 2006 ). Similarly, a global transcriptome comparison of GV-stage and MII-stage mouse oocytes revealed the expression of more than 12,000 genes by murine oocytes, more than 1,600 of which were upregulated in GV and 2,000 in MII oocytes (Cui et al., 2007 ). The authors reported a higher representation of genes associated with protein metabolism, the mitotic cell cycle, electron transport, fertilization, microtubule/cytoskeletal protein family, DNA replication, G protein-coupled receptors, and expression signalling in mature MII oocytes. The focus of our previously published study was to apply a microarray approach to identify new potential regulators and candidate genes involved in human oocyte maturation. In addition, (Assou et al., 2006 ) observed that genes involved in the proteasome pathway were over-represented during human oocyte maturation (Fig. 1 ), suggesting an important role in oogenesis for this gene family. This observation is of great interest in light of the recent findings on the role of the proteasome in transcription (Sztutorisz et al., 2006 ). The proteasome is known to interact with chromatin during multiple stages of transcription, through both proteolytic and non-proteolytic activities (Collins and Tansey, 2006 ). Proteolytic activities of the proteasome are important for promoting the exchange of transcription factors on chromatin, and possibly for facilitating multiple rounds of transcription initiation. Non-proteolytic activities of the proteasome are important for transcriptional elongation checkpoint and histone modification (Collins and Tansey, 2006 ). Previous studies described the BARD1 - BRCA1 heterodimer as an E3 ubiquitin ligase (Lorick et al., 1999 ; Ruffner et al., 2001 ). Gasca et al. (2007) observed that multiprotein E3 ubiquitin ligase complex containing BARD1, BRCA1 and BRCA2, termed BRCC are involved in human oocyte maturation. These proteins function predominantly in the nucleus to regulate cell cycle progression, DNA repair and gene transcription (Henderson, 2005 ). BARD1 is an important regulator of BRCA1 activity: the binding of BARD1 with BRCA1 maintains both proteins in the nucleus, consequently preventing apoptosis (Fabbro et al., 2004 ). Guillemin et al. (2009) reported that the apoptosis inhibitor protein BCL2L10 undergoes dynamic sub-cellular redistribution during human oocyte maturation. Knockdown of mouse Bcl2l10 using RNA interference (RNAi) in cultured oocytes resulted in maturation arrest at the MI stage accompanied by abnormal spindle and chromosome organizations (Yoon et al., 2009 ). These findings indicate that BCL2L10 may play a role in oocyte maturation.

Gene expression of oocytes according to female age

Oocyte quality is a main factor in human reproductive failure and has been strongly associated with advanced reproductive age. Recent studies have indicated differential gene expressions between younger and older oocytes in human (Steuerwald et al., 2007 ; Grondahl et al., 2010 ) and mouse (Hamatani et al., 2004b ; Pan et al., 2008 ). Age affects the expression of human oocyte genes in a variety of major functional categories, including cell cycle regulation, cytoskeletal structure, energy pathways, transcription initiation, and stress responses (Steuerwald et al., 2007 ). Recently, Grondahl et al. (2010) performed full-genome microarray analysis on individual oocytes of women with advanced age in comparison with a younger patient. They observed a substantial difference between younger and older oocytes at the transcriptional level of genes involved in central biological functions, from establishment and organization of the meiotic spindle checkpoint, through protein metabolism and DNA repair to the regulation of meiotic division and cell cycle. In the mouse, gene profiles involved in oxidative stress, mitochondrial function, chromatin structure, DNA methylation and genome stability are altered with old age (Hamatani et al., 2004b ). In addition, the expression profiling of young and old mouse oocytes revealed changes in the expression of several kinetochore components of the spindle assembly checkpoint (SAC); including protein kinases (e.g., Bub1, BubR1 , Aurora kinase) and Cdc20 , a critical activator of the Anaphase Promoting Complex (Pan et al., 2008 ).

Gene expression of oocytes under in vitro conditions

Intrinsic oocyte developmental competence, as assessed by development up to the blastocyst stage, has been positively associated with the site of oocyte maturation (Rizos et al., 2002 ). The site of oocyte maturation has a profound effect on oocyte quality; oocytes matured in vivo are of superior competence compared to those matured in vitro (Rizos et al., 2002 ). Microarray technologies were applied to humans and to mice to identify the differences between development in vitro versus in vivo matured oocytes (Jones et al., 2008 ; Pan et al., 2005 ; Wells and Patrizio, 2008 ). Jones et al. (2008) compared the transcriptome of oocytes matured in vivo (with relatively high developmental competence) with the transcriptome of oocytes under in vitro maturations (IVM) conditions (with low developmental competence) and identified an over-abundance of a large number of genes in human oocytes matured in vitro compared with in vivo. This study suggested, that the increase in gene expression detected in vitro could be due to either a deregulation in gene transcription or the posttranscriptional modification of genes, causing an inadequate temporal utilization of transcripts which could be translated as a developmental incompetence of any embryos resulting from these oocytes. Similarly, Zheng et al. (2005) indicated that the reduced developmental competence of non-human primate oocytes under IVM was a result of a failure of these oocytes to undergo the normal pattern of transcript silencing. Wells and Patrizio (2008) elegantly differentiated the mRNA expression in human oocytes at different maturational stages under IVM. They observed that high levels of mRNAs, proteins, substrates and nutrients are accumulated in the oocyte during maturation and are associated with oocyte developmental competence (Song et al., 2005 ; Watson. 2007 ). In addition, IVM seems
to alter the abundance of certain MII oocyte mRNAs compared with in vivo MII oocytes (Lee et al., 2008; Zheng et al., 2005). The differences in gene expression detected may assist in the design of optimized protocols for ovarian stimulation and in further refining media formulations for IVM.

**Gene expression of oocytes in PCOS patients**

Polycystic ovary syndrome (PCOS) is a good model for studying the loss of oocyte quality. Studies comparing gene expression arrays in tissues from patients with PCOS compared with normal responders have reported similar pathways for differentially expressed genes in whole ovaries (Jansen et al., 2004; Oksjoki et al., 2005) and oocytes (Wood et al., 2007). Wood et al. (2007) identified 374 genes with different mRNA transcript levels when analyzing morphologically indistinguishable oocytes from normal responders and PCOS patients (Wood et al., 2007). Many of the genes found to be differentially expressed in PCOS are involved in signal transduction, transcription, DNA and RNA processing, and the cell cycle. Of these, 68 are involved in chromosome alignment and segregation; other genes contain putative androgen receptors (Wood et al., 2007). These differences provide a partial explanation for the reduction of oocyte capacity observed in PCOS patients. Cumulus cells (CCs) play a major role in the control of oocyte metabolism, and therefore it is likely that dysfunction of these cells could play a role in PCOS.

**Human oocyte-cumulus cells relationships**

The oocyte is surrounded by several cell layers, CCs, which are tightly connected to each other and to the oocyte through gap junctions such as GJA1 (Cx43) and GJA4 (Cx37), that facilitate the bi-directional traffic between the oocyte and CCs (Tanghe et al., 2002). The oocytes create and control their microenvironment by promoting differentiation of the cumulus cell phenotype through secretion of paracrine signalling factors, such as growth and differentiation factor 9 (GDF9) and bone morphogenic protein-15 (BMP15), which are members of the transforming growth factor β (TGFβ) family (Eppig et al., 1997). Absence of these two oocyte-secreted factors has been shown to cause infertility (Dong et al., 1996; Galloway et al., 2000). GDF9 show increased cumulus production of hyaluronic acid synthetase (HAS2), an enzyme responsible for the hyaluronic acid synthesis which is the major structural backbone of the cumulus extracellular matrix (Elvin et al., 1999), as well as the proteoglycan versican, another important component of the cumulus matrix (Russell et al., 2006). Other genes were induced and regulated by GDF-9 such as Peroxiredoxin 2 (PRDX2) (leyens et al., 2004) and SMAD family member 2 and 3 (SMAD2/3) (Dragovic et al., 2007). Studies on SMAD2/3 conditional knockout mice indicate that SMAD2/3 are indispensable for normal cumulus expansion (Li, Q. et al., 2008). Furthermore, CCs play an essential role, particularly for normal oocyte growth, ovulation, fertilization and embryo development (Chang et al., 2002). Active components of the cumulus matrix come from several sources: they are synthesized directly by CCs under the control of endocrine and oocyte-derived factors, secreted by mural granulosa cells (GCs), or enter the follicle from blood plasma (Eppig, 1982, 1991; Tanghe et al., 2002; Vanderhyden et al., 2003). In addition, CCs and oocytes have different metabolic needs. Oocytes themselves are unable to synthesize cholesterol and poorly metabolize glucose for energy production. Inversely, the CCs metabolize alternative substrates, such as cholesterol and glucose, which are essential for the development and function of oocyte (Eppig et al. 2005). The exact nature and diversity of oocyte-CC signalling molecules are complex and dynamic. Errors in the regulatory cumulus-oocyte complex (COC) may result in the production of oocytes unable to undergo embryo and pregnancy outcomes. The study of the CC gene expression profile offers the opportunity by a non-invasive method, to predict oocyte competence, because bidirectional communication between the CC and the oocyte are necessary for acquisition of this competence.

**Global analysis of human embryo gene expression profile**

Embryo survival requires that the oocyte be sufficiently equipped with stored maternal products and that embryonic gene expression start at the correct time. Human early embryos are highly sensitive to the environment under in vitro culture conditions. The culture conditions can alter the gene expression pattern (Lonergan et al., 2006). Several studies have explored global patterns of embryos gene expression in both mouse (Hamatani et al., 2004a; Rinaudo et al., 2004; Wang, Q. T. et al., 2004; Zeng et al., 2004) and human (Dobson et al., 2004; Li, S. S. et al., 2006; Wells et al., 2005; Zhang et al., 2009) embryos (Table 1). Gene expression profiling of early mouse embryos showed characteristic patterns of maternal RNA depletion and revealed that embryonic genome activation (EGA) happens in two phases: the first phase is zygotic genome activation (ZGA), the second mid-preimplantation gene activation (MGA), which precedes the dynamic morphological and functional changes from the morula to the blastocyst stage (Hamatani et al., 2006). Two major groups of gene activity can be defined: firstly in oocytes and first early cleavage stages and the second consisting of 4-cell to blastocyst stages, which correlates with the transition from maternal to zygotic expression. Several transcripts of molecules involved in WNT or BMP signalling pathways were identified in the embryonic transcriptome, indicating that these critical regulators of cell fate and patterning are conserved and functional at these stages of pre-implantation development. Dobson et al. (2004) reported the global patterns of gene expression in human embryos on day 2 and 3 (n=8) by using DNA microarray analysis. This study revealed that the first few days of oocyte maturation and embryo development were characterized by a significant decrease in transcript levels, suggesting that decay of RNAs associated with gamete identity is integral to embryo development (Dobson et al., 2004). Furthermore, this study revealed that developmental arrest and activation of the embryonic genome are unrelated events. Wells et al. (2005) used reverse transcription and real-time fluorescent PCR to quantify the expression patterns of nine known genes implicated in DNA damage and cell division (BRCA1, BRCA2, ATM, TP53, RB1, ...
, MAD2, BUB1, APC, and beta-Actin) at the early stages of human embryonic development. The data suggest that embryos with patterns of gene expression appropriate for their developmental stage may have superior viability to those displaying atypical gene activities (Wells et al., 2005). Li et al. (2006) used a cDNA microarray containing 9,600 transcripts to investigate 631 differently expressed genes in oocytes, as well as the 4-cell and 8-cell human embryonic stages. Numerous genes were found to be expressed at a value double that of the median value of all genes expressed; 184 genes in oocytes, 29 in 4-cell embryos, and 65 genes in 8-cell embryos. These results indicate that the expression of some zygotic genes had already occurred at the 4-cell embryonic stage (Li et al., 2006). In addition, Chen et al. (2005) observed global gene expression changes during the hatching of mouse blastocysts, an essential process for implantation. This study observed that there were 85 genes which were upregulated in blastocysts at the hatching stage. These genes included cell adhesion molecules, epigenetic regulators, stress response regulators, and immunoregulatory regulators (Chen et al., 2005). Furthermore, Adjaye et al. (2005) identified biomarker transcripts specific to inner cell mass (ICM) such as (OCT4/POUSF1, NANOG, HMGBl, and DPPA5) when comparing gene expression profiles of ICM and trophectoderm cells (TE) from human blastocysts. This data indicates that the emergence of pluripotent ICM lineages from the morula is controlled by metabolic and signalling pathways including WNT, mitogen-activated protein kinase (MAPK), transforming growth factor-β (TGF-β), NOTCH, integrin-mediated cell adhesion and apoptosis-signalling pathways (Adjaye et al., 2005).

The molecular mechanism of early embryonic development is still unclear. To identify genes that are involved in early embryonic development, we have conducted an analysis related to the expression profiles of five pooled oocytes, two embryos on day-3 and eight human embryonic stem cell (hESC) lines. In order to ascertain whether there is a specific molecular signature corresponding to each stage of early embryonic development, we performed principal component analysis (PCA) and hierarchical clustering (Figure 2). All of the oocyte, embryo and hESC samples clustered as homogeneous groups (Figure 2A) in agreement with the robustness of the Affymetrix microarrays (Irizarry et al., 2005). Hierarchical clustering on 10,000 genes delineated five major clusters of genes (Figure 2B). The hierarchical clustering dendrogram showed that hESCs are distantly related to oocytes and embryos on day-3 (Figure 2C), revealing the distinctive patterns of maternal RNA degradation and embryonic gene activation, which contribute to the dramatic morphological changes during early embryonic development. We identified genes whose expression gradually increased or decreased during early embryonic development (Figure 3A–B). Taken together, these results seem to indicate that early embryonic development, particularly totipotent cells and highly pluripotent cells (hESCs), have quite distinct genetic programs. This supports the notion of a gradual decrease of developmental potential from the pre-implantation embryo stage to stem cells.

DNA damage can have severe cellular consequences, especially in cells which are the origin of all cells of the future human being. Hence, DNA repair systems play a crucial role in early embryonic development. The pathways and strategies for DNA repair in early embryos have been reviewed elsewhere (Jaroudi et al., 2007). The importance of DNA repair gene products during development is highlighted by the phenotypes observed for human genetic disorders associated with DNA damage response (Hales, 2005). Synthesis and repair of DNA, as well as methylation of DNA, is crucial in gametogenesis, fertilization and pregnancy. Recently, DNA repair gene expression was investigated in human oocytes and blastocysts to identify the pathways involved at these stages and to detect potential differences in DNA repair mechanisms of pre- and post-embryonic genome activation (Jaroudi et al., 2009). Large numbers of repair genes were detected indicating that all DNA repair pathways are potentially functional in human oocytes and blastocysts. Expression levels of DNA repair genes at the pre- and post-EGA transcriptional level suggest differences in DNA repair mechanisms at these developmental stages (Jaroudi et al., 2009).

Epigenetic regulation of human early embryonic development

Epigenetics refers to a collection of mechanisms that can cause a change in the phenotype of a cell without altering its DNA sequence (Goldberg et al., 2007). Early embryonic development is regulated epigenetically. During the development process, the epigenetic signature changes as the cell enters fertilization and/or differentiation. DNA methylation is an important epigenetic event, regulating gene expression and chromatin structure in any developmental processes including gene imprinting and embryogenesis (Kafri et al., 1992). To maintain methylated DNA, DNA methyltransferases (DNMTs) are necessary (Vassena et al., 2005). This family of enzymes is divided into two major classes: DNMT1 and DNMT3 (including DNMT3A and DNMT3B). DNMT1 enzyme is the main methyltransferase by far and its exceptional preference for hemimethylated DNA indicates its role in the maintenance of methylated status during DNA replication. Genetic studies have suggested that DNMT1 may play a role in the methylation of repeat elements in growing oocytes (Gaudet et al., 2004). Our genes expression data indicate that the DNMT1 mRNA was overexpressed in mature MII oocytes and decreased progressively in embryos on day-3 and in hESCs (Figure 3B), revealing that the oocyte-specific DNMT1 transcript is expressed in human oogenesis and persists in embryos at day-3. DNMT3B transcripts were detected in mature MII oocytes and hESCs (Assou et al., 2009), and the expression of DNMT3A was enriched in undifferentiated hESCs (Assou et al., 2007; Huntriss et al., 2004; Richards et al., 2004), indicating that the DNMT3A family methyltransferases are candidate epigenetic regulators during preimplantation development. Both DNMT3A and DNMT3B enzymes are involved in de novo methylation processes with different substrate preferences. DNMT3A acts preferentially on unmethylated DNA while DNMT3B can act on both hemimethylated and unmethylated DNA (Vassena et al., 2005). These enzymes are responsible for the establishment of DNA methylation during development (Okano et al., 1999; Watanabe et al., 2002).
As imprinting is closely related to embryogenesis, we analyzed the expression of imprinted genes in oocytes, embryos on day-3 and hESCs. We looked at the expression of forty imprinted human genes (Morison et al., 2001) in our data (the list of imprinted genes is available at http://www.otago.ac.nz/IGC). The analysis of the expression levels of these imprinted genes compared with their expression levels in mature MII oocytes, embryos on day-3, and hESCs reveals their reprogramming time points in humans. Several imprinted genes (see Table II) were expressed to high levels in the oocytes, embryos, and hESCs. The study of the methylome and the epigenome is important because minor defects can lead to serious human diseases. Indeed, several syndromes are associated with imprinting defects during preimplantation development (van der Maarel, 2008; Yamazawa et al., 2008). Beyond transcription and epigenetics, the discovery of hundreds of microRNAs (miRNAs) has defined yet another level of mRNA regulation that could have a major impact on defining the mechanisms controlling early embryonic development (Blakaj et al., 2008; Kedde et al., 2008). These short miRNAs have the ability to target and regulate the translation and stability of target mRNAs, and thus are likely to have profound effects on embryonic gene expression patterns.

Expression of known components of two pathways in human early embryo development

Known components of two signalling pathways involved in human early embryonic development, (i) the WNT signalling pathway and (ii) the TGF-β superfamily pathway, have been observed.

**WNT signalling**

The WNT gene family consists of numerous conserved glycoproteins that regulate pattern formation during embryogenesis in a wide variety of tissues (Morkel et al., 2003). Limited data is available about the role of the WNT pathway in the reproduction process. It has been shown that this pathway is operational during human and mouse pre-implantation development stages (Adjaye et al., 1999). Furthermore, it has also been reported that activation of the canonical WNT pathway is sufficient to maintain self-renewal of both human and mouse ESCs (Sato et al., 2003). We observed that the main components of the WNT pathway such as DVL2, CTNNB1, CTNN2A, CTNN1D1, LEF1, CSNK1G3, and DAAM1 were overexpressed in MII oocytes; FZD4, GSK3β, CTNN1D1, APC, SFRP1, SFRP8, CSNK2A2, and CSNK1E to be overexpressed in embryos on day-3; FZD3, as well as FZD7, SNPs, CXXC1, AXIN, and SFRP1 in hESCs. These findings imply that the WNT pathway may participate in early embryonic development. On the other hand, we observed that expression of glycogen synthase 3 kinase (GSK-3β) is downregulated in the hESCs, corroborating the reported inactivation of GSK-3β, which leads to the activation of the WNT pathway and the maintaining of the undifferentiated state of ESCs (Sato et al., 2004). In the absence of a WNT signal, the GSK-3β kinase phosphorylates β-catenin and targets it for ubiquitin-mediated destruction. Activation of the WNT pathway by a WNT ligand inhibits GSK-3β activity, resulting in the accumulation of β-catenin. Stable β-catenin can interact with DNA-binding TCF factor, at the site where the complex activates transcription of target genes. Treating ESCs with WNT or inhibiting GSK-3 activity can stimulate ESC self-renewal and support pluripotency (Sato et al., 2004). Recently, two studies (Lluis et al., 2008; Marson et al., 2008) show that WNT-β-catenin signalling stimulates nuclear reprogramming. These studies, using distinct reprogramming methods, offer insights into the mechanisms underlying the acquisition and maintenance of pluripotency.

**TGF-β superfamily**

TGF-β and its family members, including BMPs, NODAL, and ACTIVINS, have been implicated in the development of oocytes and embryos (Chang et al., 2002). Kocabas et al. (2006) confirmed the presence of TGF-β pathways in human oocytes, previously described only in mice (Mummery, 2001). TGF-β family signals play important roles in the maintenance of self-renewal and pluripotency in both human and mouse ESCs (Watabe et al., 2009). Inhibition of TGF-β/ACTIVIN/NODAL signalling by SB-431542, a chemical inhibitor of GSK-3β-catenin and targets it for ubiquitin-mediated destruction. Activation of the WNT pathway by a WNT ligand inhibits GSK-3β activity, resulting in the accumulation of β-catenin. Stable β-catenin can interact with DNA-binding TCF factor, at the site where the complex activates transcription of target genes. Treating ESCs with WNT or inhibiting GSK-3 activity can stimulate ESC self-renewal and support pluripotency (Sato et al., 2004). Recently, two studies (Lluis et al., 2008; Marson et al., 2008) show that WNT-β-catenin signalling stimulates nuclear reprogramming. These studies, using distinct reprogramming methods, offer insights into the mechanisms underlying the acquisition and maintenance of pluripotency.

**The expression program of human embryonic stem cells**

hESCs do not exist in nature but are derived from the inner cell mass (ICM) of the human blastocyst (Thomson et al., 1998) and are considered equivalent to ICM cells, having captured the phenotype of a cell type that exists only transiently in vivo. hESCs have the capacity to differentiate into the derivatives of all three germ layers: endoderm, mesoderm and ectoderm (Carpenter et al., 2003; Lu et al., 2007); thus hESCs have great potential in the field of regenerative medicine. Although artificial, hESCs represent an excellent model for...
investigating fundamental aspects of the early embryo development and the pluripotency. Numerous genomic studies have used microarray techniques to capture a detailed view of ESC gene expression (Abeyta et al., 2004; Assou et al., 2007; Beqqali et al., 2006; Bhattacharya et al., 2004; Cowan et al., 2005; Forsyth et al., 2008; Golan-Mashiach et al., 2005). These transcriptome data can generate molecular signatures of ESCs that define the individual components, and the pathways, that regulate pluripotency and maintain the undifferentiated state. We performed a meta-analysis of 38 papers on the topic (Assou et al., 2007). We determined that 1,076 genes were found to be upregulated in hESCs by at least three studies when compared to differentiated cell types. In fact, only a single gene, OCT3/4, was found to be in common to all 38 lists of hESC-specific genes. Our analysis shows that gene expression analyses are sensitive to experimental conditions, therefore requiring extreme caution during data interpretation. A cross-comparison of the gene lists generated by these efforts shows that OCT3/4, NANOG, SOX2, DNM3B, LIN28, and CD24 are commonly expressed in the majority of hESCs (Assou et al., 2007). A "core" set of transcription factors, OCT3/4, SOX2 and NANOG, is common between the hESC lines and is considered to be a general criteria for hESC characterization and pluripotency (Hoffman et al., 2005; Zaeheres et al., 2005). A list of OCT3/4, SOX2, and NANOG target genes has been determined by cross-species genomic analysis followed by chromatin immunoprecipitation and gene expression analysis (Boyer et al., 2005). Genes that are activated by OCT3/4, SOX2, and NANOG in ESCs are mostly upregulated and important for ESC self-renewal. Cross-species comparisons between human and mouse ESC transcriptomes revealed both distinct and similar features (Rao, 2004; Sato et al., 2003; Wei et al., 2005). Wei et al. (2005) compared gene expression profiles between hESC and mESC, and showed that LIFR was overexpressed in mice, but not in humans; FGF2 was overexpressed in humans, while FGF4 was abundant in mice. Epiblast derived mouse ESCs also maintain self-renewal by the dependence on FGF2, hence activated FGF signalling (Brone et al., 2007; Tesar et al., 2007). The accumulating ESC transcriptome data is now being aimed at identifying molecular markers that will be useful for quality control of ESCs for clinical applications.

**Common molecular signatures between human metaphase II oocytes and iCM**

ESCs and mature MII oocytes are able to reprogram the nuclei of fully differentiated human somatic cells and render these cells pluripotent (Cowan et al., 2005; Rideout et al., 2001). Through the comparative analysis of the transcriptome of such cells, it may be possible to achieve a better understanding of the mechanism of nuclear reprogramming, the molecular events which occur during the first week of human embryo development, and the identification of the genes involved in pluripotency initiation (Kocabas et al., 2006; Zhang et al., 2007; Assou et al., 2009). Kocabas et al. (2006) compared the transcriptome of human MII oocytes and hESCs with a reference sample consisting of a mixture of total RNA from different normal somatic cells. In this study, the authors identified 1,626 transcripts that were significantly upregulated in the hESCs relative to somatic cells. Intersection of these 1,626 transcripts enriched in hESCs with the 5,331 transcripts enriched in human MII oocytes, identified 388 common genes that are hypothesized to play an important role in oocyte reprogramming and early events associated with human early embryo development (Kocabas et al., 2006). Zhang et al. (2007) obtained a large amount of information on gene expression in human oocytes at GV stage as compared with hESCs. In total, 1,629 genes were upregulated in GV oocytes when compared with hESCs and human foreskin fibroblasts (hFIBs). This list of genes, including MATER, ZAR1, NPM2, FIGLA, some imprinted genes, and some components of four signalling pathways such as MOS-MPF, NOTCH, WNT and TGF-beta (Zhang et al., 2007), may provide good candidate genes for understanding the oocyte maturation and the early embryonic development. On the other hand, Assou et al. (2009) compared the gene expression profiles of mature MII oocytes with hESCs, and then both to somatic tissues. There was a common oocyte/hESC gene expression profile, which included genes associated with pluripotency, such as LIN28 and TGF1, and a large chromatin remodeling network (TOP2A, DNM3B, JARID2, SMARCA5, CBX1, CBX5). Interestingly, a large set of genes was also found to code for proteins involved in the ubiquitination and proteasome pathway (Assou et al., 2009). This list provides reliable candidate genes for future studies on molecular mechanisms of pluripotency and nuclear reprogramming. The zinc finger motif is a DNA binding domain dependent on a zinc ion frequently found in transcription factors. In seeking to identify genes encoding zinc finger domain proteins expressed in both mature MII oocytes and hESCs, we observed that the distribution of genes encoding these proteins on the genome is not random: they are significantly enriched in chromosomes 19 and 20 in mature MII oocytes, and in 16 and 17 in hESCs (Assou et al., 2009). As can be seen in Figure 4, seven zinc finger proteins were common to mature MII oocytes and hESCs, three were found in oocytes only (e.g. ZNF528, ZNF468, ZNF313), and three in hESCs only (e.g. ZNF268, ZNF232, ZNF423).

**Human early embryo development transcriptome data visualization through an internet interface**

Amazonia! (http://www.amazonia.transcriptome.eu) is a freely accessible web site dedicated to the visualization of large, publicly available human transcriptome data, including data on stem cells, human embryo and germinal cells, for facilitating the free exchange of information on genomics. Data is visualized as expression bar plots with a color code enabling the recognition of cell type (Assou et al., 2007; Le Carrou et al., 2010). Genes are accessed either by keywords or through lists. Most interestingly, when data is obtained using the same platform format, sample labeling, and data normalization, it is possible to combine different experiments in a single virtual experiment. Amazonia generates bar plots that combine expression in oocytes, embryos and hESCs with more than 200 germinal and somatic tissues.

**Omnis: tools for the identification of reliable biomarkers for oocyte and embryo selection**
Since the birth of Louise Brown, the first baby conceived by IVF in the UK in 1978, more than 2 million children in the world have been born as a result of assisted reproductive technology (ART). Since then, these techniques have advanced considerably. ART protocols continue to evolve with the aim of achieving higher pregnancy rates, reducing multiple pregnancies and obtaining healthier babies. However, despite the incontestable advances, the pregnancy rate is still relatively low and has not increased significantly in the last decade (Andersen et al., 2007). The selection of embryos with higher implantation potential is one of the major challenges in ART. Initially, multiple embryo transfer was used to maximize pregnancy rates. However, improved embryo quality and rising multiple pregnancy rates has resulted in the decrease in number of embryos for replacement. Therefore, selection of the “best” embryo has become crucial, particularly with elective SET being strongly recommended. There is therefore a need to develop new objective approaches for embryo selection. The classical methods to select healthy embryos under IVF and ICSI conditions are based on morphological criteria such as early embryonic cleavage, the number and size of blastomeres, fragmentation degree, and the presence of multi-nucleation at the 4- or 8-cell stages (Fenwick et al., 2002). However, most studies suggest, that embryos with proper morphological appearance alone are not sufficient to predict a successful implantation. Beyond the criteria of embryo selection, defining oocyte quality remains one of the most difficult challenges (Wang, Q. et al., 2007). Many morphologically normal embryos do not achieve implantation or spontaneously abort during pregnancy. Considering the limitation of morphologic evaluation and cytogenetic screening methods, there is now a movement towards more sophisticated, high-performance technologies and the emerging "omics" science, such as genomics, transcriptomics, proteomics, and metabolomics (Hillier, 2008). These approaches may contribute to the design of a non-invasive viability assay to assist in the embryos selection in IVF or ICSI programs. It is necessary to distinguish the invasive approaches, which are based on direct analysis of the oocyte and embryo, from non-invasive approaches (Figure 5). Can omics approaches help shape the future of ART?

**Transcriptomic approach**

Several investigations using reverse transcription followed by real-time PCR have found that specific genes display alterations in activity apparently related to oocyte or embryo quality (Dode et al., 2006; Fair et al., 2004). Despite the accuracy and sensitivity of real-time PCR, it is still limited by the restricted number of genes that can be assessed per single oocyte or embryo. Microarray technology permits the simultaneous analysis of tens of thousands of genes, and provides greater hope for the identification of new biomarkers. Transcriptomics represent a valuable approach to biomarkers development. Both technical and statistical advances are currently facilitating the integration of this approach to IVF programs. The transcriptomic approach has been used experimentally in humans on throphoblastic biopsies of blastocysts (Jones et al., 2008a). Analysis of the transcriptome of aneuploid and normal oocytes, as determined by comparative genomic hybridization (CGH) analysis of the corresponding polar bodies, has identified several differentially expressed genes with roles in chromosome segregation during meiosis, including genes affecting cell cycle checkpoints, spindle dynamics, and chromosomal movement (Fragouli et al., 2006b; Fragouli et al., 2006a; Fragouli et al., 2009; Gutierrez-Mateo et al., 2004). An indirect and attractive approach for predicting embryo and pregnancy outcomes has been recently reported by using transcriptomic data of CCs gene expression (Assou et al., 2008). The CCs are abundant and easily accessible, which makes them an ideal material to use for the potential assessment of oocyte and embryo quality. Data concerning biomarker expression in these cells should yield more precise information, improving the embryologists’ capacity to select competent embryos either for fresh replacement or for embryo freezing. By studying genes expression profile of CCs or GCs using microarrays or reverse transcription-polymerase chain reaction (RT-PCR) and quantitative RT-PCR some studies have identified candidate biomarkers for (i) oocyte quality and competence (Zhang et al., 2005; Cillo et al., 2007; Hamel et al., 2008; Adriaensen et al., 2010), (ii) early embryo development (Anderson et al., 2009; McKenzie et al., 2004; van Montfoort et al., 2008) and (iii) embryo quality and pregnancy outcome (Assou et al., 2008; Hamel et al., 2010) (Table III). Zhang et al. (2005) reported that the expression of pentraxin-3 (PTX3) in CCs was positively correlated with oocyte competence. PTX3 was initially studied for its possible role in inflammatory reaction processes (Gewurz et al., 1995). Fertilization in vivo is very low in PTX3 knockout mice (Varani et al., 2002). Feuerstein et al (2007) reported that a number of genes including prostaglandin-endoperoxide synthase-2 (PTGS2) and steroidogenic acute regulatory protein (STAR) were also negatively associated with oocyte competence. In addition, the expression of gremlin 1 (GREM1) is also indicative of embryo quality (McKenzie et al., 2004; Cillo et al., 2007; Anderson et al., 2009, Adriaensen et al., 2010). GREM1 is an antagonist of BMPs and TGF-β signalling (Hsu et al., 1998). It has been proposed that GREM1 regulates the GDF9 and BMP15 balance (Pangas et al., 2004). By using microarrays, our group reported a specific transcriptomic signature in CCs including 630 genes correlated with embryo and pregnancy outcomes. Among this gene list, we identified BCL2L11, PCK1, and NFIB (Assou et al., 2008). BCL2L11 is involved in triggering cell death in response to abnormalities, PCK1 is associated with energy production, and NFIB regulates some of the earliest processes in embryonic development. Other studies associated the gene expression profile in follicular cells (CCs or GCs) with pregnancy (Hamel et al., 2010). They found that variation of PGK1, a transferase enzyme involved in the process of glycolysis and regulation of G-protein signalling 2 (RGS2) in follicular cells are likely candidate genes that could have a utility in the prediction of pregnancy. All these studies provide supporting evidence to the concept that CCs are a promising source of reliable biomarkers for predicting oocyte quality and/or embryo and/or pregnancy outcome (Table III). These efforts will ultimately lead to improved efficiency outcomes of ART.

**Proteomic approach**
Proteomics describes changes in the expression of proteins translated from a single genome. The human proteome, estimated at over a million proteins, is highly diverse and dynamic. The analysis of the protein composition of the embryos poses a particular challenge because of the low concentration and wide dynamic range of the population of molecules which need to be analyzed; unlike mRNA, proteins cannot be amplified for analysis. Recent advances in mass spectrometry have led to the development of methods sensitive enough to allow the examination of the secretome of single oocytes or embryos. The secretome has been defined as the subset of secreted or consumed proteins that can be found in the environment where the oocyte or embryo grows. It is now possible to obtain comprehensive protein profiles from limited amounts of complex biological fluids including IVF spent culture mediums. The secretome of an individual human embryo has been previously analyzed by surface-enhanced laser desorption and ionization time-of-flight (SELDI-TOF) mass spectrometry (MS) technology, and distinctive secretome signatures were observed at each embryonic developmental stage from fertilization to the blastocyst stage (Katz-Jaffe et al., 2006 ; Katz- Jaffe et al., 2009 ). However, no clear biomarkers of embryo viability and implantation and maturation have yet been reported. Dominguez et al. (2010) compared blastocyst conditioned media with control medium using protein microarrays, thus revealing increased expression of interleukin 10 (IL-10) and soluble tumor necrosis factor (TNFα) receptor 1 (TNFR1) and decreased expression of stem cell factor (SCF) and chemokine ligand 13 (CXCL13) (Dominguez et al., 2010). Other studies have reported a link between the level of cytokines in human follicular fluid (hFF) and the implantation potential of the embryo derived from the oocyte of this follicle. Using a bead-based multiplex sandwich immunoassay (Luminex), Ledee et al. (2008) demonstrated that only granulocyte colony-stimulating factors (G-CSF) measured in individual hFF samples together with 26 other cytokines was particularly elevated in the hFF corresponding to the embryos with highest implantation potential (Ledee et al., 2008). In addition, significantly higher levels of interleukin 10 (IL10) and interferon (IFN-γ) were found in hFF from oocytes that generated early cleavage embryos. However, Luminex technology is complex and expensive. Further studies will help to confirm the advantageous applications of proteomics testing in the field of IVF.

Metabolomic approach

Metabolomics aims at the comprehensive and quantitative analysis of metabolites in biological samples. This approach is presently being applied to human embryos (Botros et al., 2008 ; Seli et al., 2008, 2010) and oocytes (Singh et al., 2007). Because of the chemical diversity of cellular metabolites, no single analytical platform can measure the metabolome in a cell. This is due to the absence of methods to amplify metabolites, the labile nature of many metabolites, and their chemical heterogeneity and complexity. Thus, measuring the metabolome is a considerable analytical challenge. Rapid technological advances in MS and nuclear magnetic resonance (NMR) have enabled the simultaneous detection of a wide range of small molecules. Recently, the development of a screening technology using Raman and near infrared (NIR) spectroscopy were used to detect biomarkers in culture medium with algorithms generated for positive, as compared to negative, IVF outcomes (Scott et al., 2008 ; Seli et al., 2007). NIR analysis is not typically used for target metabolite identification, but is used for overall spectral profile comparisons. Recent studies using a SET program on days-2 and -3 showed higher mean viability scores for embryos that resulted in a pregnancy with fetal heart activity, compared with those without (Seli et al., 2008; Vergouw et al., 2008). They showed that metabolomic profiling of culture media from embryos was independent of morphology. Large-scale, prospective studies are necessary to relate the metabolic profile of medium to the developmental potential of human oocytes and embryos. In addition, the ideal metabolomic profile should be confirmed by several independent platforms.

Conclusion

This review highlights the recent advances in elucidating the molecular complexity in early embryonic development. The knowledge of the global pattern of gene expression is important for understanding critical regulatory pathways involved in the crucial steps of early embryonic development. Certain gene expression profiles help to unravel the mystery of these developmental stages. Further understanding of the biological role of these genes may expand our knowledge of the oocyte maturation, fertilization, chromatin remodeling, totipotency, pluripotency, and early morphogenesis steps. The practical implications of compiling gene expression information on human oocytes and embryos would be enormous since it could potentially help us to understand and solve problems related to infertility.

In recent years, the emergence of new technologies ‘Omics’ such as microarrays have already started to be used in the IVF/ICSI program. The use of microarray technology in the analysis of early embryonic development poses specific challenges associated with the picogram levels of mRNA in a single cell/embryo. Another challenge for applying microarrays to early embryonic development is the high degree of expression plasticity seen in early stage embryos.

The ideal DNA microarrays would potentially give rise to diagnostics for assessing human embryo quality in IVF program. The application of transcriptomic, proteomic and metabolomic approaches have greatly broadened our understanding of early human development. The challenge now is to correlate gene/protein/metabolite function and regulation to specific events in early embryonic development. These tools may ultimately lead to non-invasive tests for oocyte or embryo quality revealing previously hidden information concerning both oocyte and embryonic developmental competence. Once fully validated, these new approaches are expected to improve...
References:

Acknowledgements:

Funding

We thank the direction of the University-Hospital of Montpellier, the Association Francaise contre les Myopathies (AFM), Vitrolife, Genervier, and Ferring Pharmaceutical Companies for their support.

Page 1020


Okano M, Bell DW, Haber DA, Lina E. DNA methyltransferase Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999; 99: 247 - 57


Richards M, Tan SP, Tan JH, Chan WK, Bongso A. The transcriptome profile of human embryonic stem cells as defined by SAGE. Stem Cells. 2004; 22: 51 - 64


Figure 1
Upregulation of proteasome genes during oocyte maturation. Histograms show signal values of six proteasome genes (PSMA2, PSMA5, PSMD2, PSMD9, PSMD11 and PSMG1) in each stage of oocyte maturation. Gene expression is measured by pan-genomic HG-U133 Plus 2.0 Affymetrix oligonucleotides microarrays, and the signal intensity for each gene is shown on the Y-axis as arbitrary units determined by the GCOS 1.2 software (Affymetrix). GV; germinal vesicle; MI, metaphase I; MII, metaphase II.

Figure 2
Principal component analysis (PCA) and hierarchical clustering of all samples from different developmental stages. (A) PCA distributes samples into a three-dimensional space based on the variances in gene expression; samples that have similar trends in their gene expression profiles will cluster together in the PCA plot. This analysis, using GeneSpring® software, resulted in a clear segregation of the fifteen replicated samples into three clusters, corresponding to human mature oocytes, human embryos, and hESCs. Each colored point represents a sample, characterized by the gene expression of all probe sets (54,675) on the Affymetrix HG U133 Plus 2 array. The first, second, and third principal components are displayed on the X-, Y- and Z-axes, respectively. (B) The expression signatures of samples were visualized by hierarchical clustering on the 10,000 probe sets with the highest variation coefficient. The colors indicate the relative expression levels of each gene, with red indicating an expression above median, green indicating expression under median and black representing median expression. Cluster (a) was a group of genes differentially overexpressed in oocytes, including DAZL, ZP1, ZP2, ZP3, ZP4, AURKA and HOXA7. Cluster (b) featured genes that were detected in both oocyte and embryo samples, such as NALP4, DPPA5, and ACTL8. Cluster (c) grouped genes that were detected in both embryo and hESC samples, such as NANOG, SALL4, and ANXA2. Cluster (d) included genes differentially overexpressed in hESCs, such as POU5F1, CD24, SOX2, FZD7 and ZIC3. Cluster (e) assembled genes overexpressed in day-3 embryos, such as CCNA1, H3F3B and FGF9. (C) The dendogram shows that all replicates are clustered by their appropriate stage. All the replicate samples of the hESC group self-cluster into one branch. Both oocyte and embryo samples self-cluster into another branch (dotted box) that further divides into two major sub-branches and into which all the replicates from oocytes and embryos self-cluster.
Figure 3
Expression of selected genes in human oocytes, and embryos on day-3 and hESCs. Figure show signal values of 18 genes that are gradually increased (A) or decreased (B) during early embryonic development. (C) Characterization of TGF-β signalling pathway during early embryonic development. Genes shown in red are upregulated in human oocytes or in embryos on day-3 or in hESCs. Examples of genes overexpressed (red) in each stage are indicated in boxes.

Figure 4
Chromosomal distribution of genes encoding zinc finger domain proteins were expressed significantly differently in human mature oocytes (light violet) as compared to hESCs (light pink) as well as to those commonly expressed in mature human oocytes and hESCs (light yellow). The selected genes were retrieved in lists previously published in (Assou et al., 2009).
Figure 5
Different direct or indirect approaches suggested for oocyte or embryo selection (transcriptomic, proteomic and metabolomic approaches).
### Table I
Microarray studies of oocytes and embryos

<table>
<thead>
<tr>
<th>Techniques</th>
<th>Samples</th>
<th>Number of identified genes</th>
<th>Targets</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>HG-U133 Plus 2.0 array (Affymetrix)</td>
<td>Oocytes</td>
<td>1,361 transcripts expressed in oocytes</td>
<td>Study oocyte transcriptomes</td>
<td>Bermudez et al., 2004</td>
</tr>
<tr>
<td>HG-U133 Plus 2.0 array (Affymetrix)</td>
<td>Oocytes</td>
<td>1,514 overexpressed in oocytes compared with cumulus cells</td>
<td>Understanding of the mechanisms regulating oocyte maturation</td>
<td>Assou et al., 2006</td>
</tr>
<tr>
<td>HG-U133 Plus 2.0 array (Affymetrix)</td>
<td>Oocytes</td>
<td>5,331 transcripts enriched in metaphase II oocytes relative to somatic cells</td>
<td>Comprehension of genes expressed in in vivo matured oocytes</td>
<td>Kocabas et al., 2006</td>
</tr>
<tr>
<td>HG-U133 Plus 2.0 array (Affymetrix)</td>
<td>Oocytes</td>
<td>10,183 genes were expressed in germinal vesicle</td>
<td>Study of global gene expression in human oocytes at the later stages of folliculogenesis (germinal vesicle stage)</td>
<td>Zhang et al., 2007</td>
</tr>
<tr>
<td>HG-U133 Plus 2.0 array (Affymetrix)</td>
<td>Oocytes</td>
<td>Of the 8123 transcripts expressed in the oocytes, 374 genes showed significant differences in mRNA abundance in PCOS oocytes</td>
<td>Understanding of PCOS</td>
<td>Wood et al., 2007</td>
</tr>
<tr>
<td>HG-U133 Plus 2.0 array (Affymetrix)</td>
<td>Oocytes</td>
<td>---</td>
<td>Identify new potential regulators and marker genes which are involved in oocyte maturation</td>
<td>Gasca et al., 2007</td>
</tr>
<tr>
<td>HG-U133 Plus 2.0 array (Affymetrix)</td>
<td>Oocytes</td>
<td>283 genes found in the case report sample</td>
<td>Identify molecular abnormalities in metaphase II (MII) oocytes</td>
<td>Gasca et al., 2008</td>
</tr>
<tr>
<td>Whole Genome Bioarrays printed with 54,840 discovery probes representing 18,055 human genes and an additional 29,378 human expressed sequence tags (EST)</td>
<td>Oocytes</td>
<td>2,000 genes were identified as expressed at more than 2-fold higher levels in oocytes matured in vitro than those matured in vivo or in vitro</td>
<td>Analysis of gene expression profile of oocytes following maturation in vivo</td>
<td>Jones et al., 2008b</td>
</tr>
<tr>
<td>Applied Biosystems Human Genome Survey Microarray (32,878 60-mer oligonucleotide)</td>
<td>Oocytes</td>
<td>Germinal vesicle, in vivo-MII, and IVM-MII oocytes expressed 12,219; 9,735; and 8,510 genes, respectively</td>
<td>Characterized the patterns of gene expression in germinal vesicle stage and meiosis II oocytes matured in vitro or in vivo</td>
<td>Wells and Patrizio, 2008</td>
</tr>
<tr>
<td>HG-U133 Plus 2.0 array (Affymetrix)</td>
<td>Oocytes</td>
<td>342 genes showed a significantly different expression level between the two age groups (women aged, 36 years (younger) and mature oocytes aged 37–39 years (older))</td>
<td>Investigate the effect of age on gene expression profile in mature oocytes</td>
<td>Grondahl et al., 2010</td>
</tr>
<tr>
<td>Two cDNA microarrays, each containing about 20,000 Oocytes targets (representing in total ~29,778 independent genes and embryos according to Unigene Build 155).</td>
<td>Oocytes</td>
<td>1,896 significant changes in expression following fertilization through day 3 of development</td>
<td>Global analysis of the pre-implantation embryo transcriptome</td>
<td>Dobson et al., 2004</td>
</tr>
<tr>
<td>cDNA microarrays containing 9,600 cDNA spots</td>
<td>Oocytes</td>
<td>184; 29; and 65 genes were overexpressed in oocytes, 4- and 8-cell embryos, respectively</td>
<td>Identify differential expression profiles of genes in single oocytes, 4- and 8-cell preimplantation embryos</td>
<td>Li et al., 2004</td>
</tr>
<tr>
<td>Assay Type</td>
<td>Sample Type</td>
<td>Affiliation</td>
<td>Description</td>
<td>Year</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Genome Survey Microarrays V2.0 (Applied Biosystems)</td>
<td>Oocytes and embryos</td>
<td>Jaroudi et al., 2009</td>
<td>Oocytes 107 DNA repair genes were detected in oocytes and embryos.</td>
<td>2009</td>
</tr>
<tr>
<td>HG-U133 Plus 2.0 array (Affymetrix)</td>
<td>Oocytes and embryos</td>
<td>Zhang et al., 2009</td>
<td>Oocytes 5,477 transcripts differentially expressed into transition from Study of global gene expression in human pre-implantation embryos.</td>
<td>2009</td>
</tr>
<tr>
<td>Mouse oocytes and embryos</td>
<td>Oocytes</td>
<td>Hamatani et al., 2004b</td>
<td>530 genes showed statistically significant expression changes.</td>
<td>2004b</td>
</tr>
<tr>
<td>NIA 60-mer oligo microarray (Agilent)</td>
<td>Oocytes</td>
<td>Hamatani et al., 2004b</td>
<td>Identification of gene expression profiles of oocytes obtained from aged mice with oocytes from young mice.</td>
<td>2004b</td>
</tr>
<tr>
<td>Genome Array MOE430 A and B chips (Affymetrix)</td>
<td>Oocytes</td>
<td>Pan et al., 2005</td>
<td>Study of global gene expression profiles of early folliculogenesis in primordial, primary, and secondary follicles.</td>
<td>2005</td>
</tr>
<tr>
<td>7.4 K GenePlore TwinChip</td>
<td>Oocytes</td>
<td>Yoon et al., 2006</td>
<td>Oocytes 214 genes that are upregulated between primordial and primary follicles.</td>
<td>2006</td>
</tr>
<tr>
<td>Affymetrix 430 v2.0 GeneChips</td>
<td>Oocytes</td>
<td>Su et al., 2007</td>
<td>3,002 genes changed significantly during GV-to-MII transition.</td>
<td>2007</td>
</tr>
<tr>
<td>Applied Biosystems Mouse Genome survey Array (32,996 Oocytes)</td>
<td>Oocytes</td>
<td>Cui et al., 2007</td>
<td>1,682 genes with more highly expression in GV-stage oocytes than in MII-stage oocytes.</td>
<td>2007</td>
</tr>
<tr>
<td>NIA 15k cDNA microarray</td>
<td>Embryos</td>
<td>Tanaka et al., 2004</td>
<td>428 genes upregulated and 748 downregulated in blastocyst.</td>
<td>2004</td>
</tr>
<tr>
<td>NIA 60-mer oligo microarray (Agilent)</td>
<td>Embryos</td>
<td>Hamatani et al., 2004a</td>
<td>4,000 genes changed in expression over 5-fold.</td>
<td>2004a</td>
</tr>
<tr>
<td>Affymetrix 25-mer DNA</td>
<td>Embryos</td>
<td>Rinaudo, P. F. et al., 2006</td>
<td>248 genes were differentially expressed between the two-cell embryos and two-cell block embryos.</td>
<td>2006</td>
</tr>
<tr>
<td>GenePlore TwinChip Mouse 7.4 K elements (containing 7,410 different genes)</td>
<td>Embryos</td>
<td>Jeong et al., 2006</td>
<td>248 genes were differentially expressed between the two-cell embryos and two-cell block embryos.</td>
<td>2006</td>
</tr>
<tr>
<td>Affymetrix 430 2.0 GeneChips</td>
<td>Embryos</td>
<td>Giritharan et al., 2007</td>
<td>1,912 was statistically different in the IVF cohort when compared with the in vivo control embryos.</td>
<td>2007</td>
</tr>
<tr>
<td>Affymetrix 25-mer DNA</td>
<td>Oocytes</td>
<td>Wang et al., 2007</td>
<td>Analysis of gene activity during pre-implantation development.</td>
<td>2007</td>
</tr>
</tbody>
</table>
and embryos

MG_U74Av2, MOE430A, and MOE430B GeneChips Oocytes 9,414 genes differentially expressed in oocytes

(affymetrix)

NIA 60-mer oligo microarray (Agilent) Oocytes --- a n d embryos

Analysis of global patterns of gene expression during Zeng et pre-implantation development al., 2004)

Identification of a large number of genes and multiple ( signaling pathways involved at each developmental stage of Hamatani et al., pre-implantation embryos 2006)

---

Table II
Imprinted genes abundantly expressed in human oocytes, embryos on day-3 and hESCs

<table>
<thead>
<tr>
<th>Oocytes</th>
<th>Embryos (Day-3)</th>
<th>hESCs</th>
</tr>
</thead>
<tbody>
<tr>
<td>paternally</td>
<td>Maternally</td>
<td>paternally</td>
</tr>
<tr>
<td>SNRPN, TP73L, GRB10, H19, KCNQ1, CDKN1C, INPP5F, KCNQ1OT1, TP73L, GRB10, CPA4, KCNK9, PHLD2A, PEG10, SGCE, NDN, SNRPN, H19, GRB10, PPPI9A, PLAG1, OSBPL5, UBE3, PRIM2A, MEG3</td>
<td>ZIM2, PEG3, PON1</td>
<td>OSBPL5, UBE3A, ZNF597</td>
</tr>
<tr>
<td>paternally</td>
<td>Maternally</td>
<td>paternally</td>
</tr>
<tr>
<td>SNRPN, TP73L, GRB10, H19, KCNQ1, CDKN1C, INPP5F, KCNQ1OT1, TP73L, GRB10, CPA4, KCNK9, PHLD2A, PEG10, SGCE, NDN, SNRPN, H19, GRB10, PPPI9A, PLAG1, OSBPL5, UBE3, PRIM2A, MEG3</td>
<td>ZIM2, PEG3, PON1</td>
<td>OSBPL5, UBE3A, ZNF597</td>
</tr>
</tbody>
</table>

Table III
Potential follicular cells biomarkers correlated to oocyte quality, embryo competence or pregnancy outcome. CCs: cumulus cells, GCs: granulosa cells

<table>
<thead>
<tr>
<th>Biomarkers</th>
<th>Name</th>
<th>Function</th>
<th>Samples</th>
<th>Approaches</th>
<th>Outcome</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTK3</td>
<td>Pentraxin-3</td>
<td>involved in inflammatory reaction</td>
<td>Pooled CCs</td>
<td>Microarrays (4 chips)</td>
<td>Positively correlated with oocyte competence</td>
<td>Zhang, X. et al., 2005</td>
</tr>
<tr>
<td>PCK1, BCL2L11, NFIB</td>
<td>Phosphoenolpyruvate carboxykinase 1, BCL2-like 11 (apoptosis facilitator), Nuclear factor I/B</td>
<td>PCK1: associated with energy production BCL2L11: involved in triggering cell death in response to abnormalities NFIB: regulates some of the earliest processes in embryonic development</td>
<td>CCs from Microarrays</td>
<td>Predict embryo and pregnancy outcomes</td>
<td>Assou et al., 2008</td>
<td></td>
</tr>
<tr>
<td>CYP19A1, SERPINE2</td>
<td>Cytochrome P450, family 19, subfamily A, polypeptide 1, Serpin peptidase inhibitor clade E member 2, Cell division cycle 42, Ferredoxin 1</td>
<td>CYP19A1: metabolizes androgen into estradiol-17β SERPINE2: involved in apoptosis and (pooled) chromatin condensation CDC42: involved in apoptosis, transcriptional activation, cell proliferation and cell polarity FDX1 and HSD3B1: responsible for progesterone synthesis</td>
<td>CCs from Mural GCs and CCs (4 chips)</td>
<td>Predict pregnancy outcomes</td>
<td>Hamel et al., 2008</td>
<td></td>
</tr>
<tr>
<td>CCND2, CXCR4, GPX3, Cyclin D2, Chemokine (C-X-C motif) receptor 4, Glutathione peroxidase 3, 3, CCND2: plays an role in proliferation of CCs and CCs (cell cycle regulator) HSPB1, DVL3, DHCR7, Heat shock protein 1, Dishevelled, dsh homolog 3 7- dehydrocholesterol reductase, Catenin, delta 1, Tripartite motif- containing 28</td>
<td>CXX4 and GPX3: have a role in relieve hypoxic stress HSPB1: act as co-repressor of estrogen signaling DVL3: plays a role in angiogenesis</td>
<td>CCs from Mural GCs and CCs (4 chips)</td>
<td>Predict pregnancy outcomes</td>
<td>van Montfoort et al., 2008</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Page 19/20
<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAS2, GREM1</td>
<td>Hyaluronan synthase 2, Gremlin 1</td>
</tr>
<tr>
<td>STAR, AREG, Cx43, PTGS2, SCD1 and SCD5</td>
<td>Steroidogenic acute regulatory protein, Amphiregulin, Connexin 43, Prostaglandin-endoperoxide synthase 2, Stearoyl-CoA desaturase 1 and 5</td>
</tr>
<tr>
<td>BDNF, GREM1</td>
<td>Brain-derived neurotrophic factor, Gremlin 1</td>
</tr>
<tr>
<td>PGK1, RGS2 and RGS3, CDC42</td>
<td>Phosphoglycerate kinase 1, Regulator of G-protein signaling 2 and 3, Cell division cycle 42</td>
</tr>
<tr>
<td>VCAN, RPS6KA2, ALCAM, GREM1</td>
<td>Versican, Ribosomal protein S6 kinase polypeptide 2, Activated leukocyte cell adhesion molecule, Gremlin 1</td>
</tr>
</tbody>
</table>

**DHCR7**: involved in progesterone and estrogen synthesis  
**CTNND1**: has a role in adhesion  
**TRIM28**: has a role in DNA repair  
**HAS2**: involved in extracellular matrix (ECM) formation  
**GREM1**: antagonist of BMP and TGF-β

**CCs from individual eggs** from RT-PCR  
**STAR**: regulated the cholesterol transport into the inner mitochondrial  
**AREG**: act as mediators of LH  
**Cx43**: permit the transfer of metabolites for growth and development and maintenance of meiotic arrest of the oocyte  
**PTGS2**: involved in inflammation and mitogenesis  
**SCD**: involved in biosynthesis of monounsaturated fatty acids from saturated acids  

**HAS2, GREM1**  
**STAR, AREG, Cx43, Steroidogenic acute regulatory protein, Amphiregulin, Connexin 43, PTGS2, SCD1 and SCD5**  
**HAS2**: involved in extracellular matrix CCs from RT-PCR individual eggs  
**GREM1**: antagonist of BMP and TGF-β

**Positively associated (Cillo et al., 2007)**

**STAR**: regulated the cholesterol transport CCs from RT-PCR individual eggs  
**AREG**: act as mediators of LH  
**Cx43**: permit the transfer of metabolites for growth and development and maintenance of meiotic arrest of the oocyte  
**PTGS2**: involved in inflammation and mitogenesis  
**SCD**: involved in biosynthesis of monounsaturated fatty acids from saturated acids  

**Negatively associated (Feuerstein et al., 2007)**

**HAS2, PTGS2, GREM1**  
**BDNF, GREM1**  
**PGK1, RGS2 and RGS3, CDC42**  
**VCAN, RPS6KA2, ALCAM, GREM1**

**CCs from Quantitative RT-PCR**  
**HAS2, PTGS2, GREM1**  
**BDNF, GREM1**  
**PGK1, RGS2 and RGS3, CDC42**  
**VCAN, RPS6KA2, ALCAM, GREM1**

**Positively associated (McKenzie et al., 2004)**

**Negatively associated (Feuerstein et al., 2007)**

**BDNF, GREM1**  
**PGK1, RGS2 and RGS3, CDC42**  
**VCAN, RPS6KA2, ALCAM, GREM1**

**CCs from Quantitative RT-PCR**  
**BDNF**: neurotrophic factor playing a role in regulation of stress response  
**PGK1**: involved in glycolysis  
**RGS**: hydrolyzed GTP to GDP  
**VCAN**: plays a central role in tissue morphogenesis and maintenance  
**RPSK6KA2**: involved in the EGF signaling cascade  
**ALCAM**: involved in immune response  

**Mural GCs Quantitative RT-PCR**  
**BDNF**: neurotrophic factor playing a role in regulation of stress response  
**PGK1**: involved in glycolysis  
**RGS**: hydrolyzed GTP to GDP  
**VCAN**: plays a central role in tissue morphogenesis and maintenance  
**RPSK6KA2**: involved in the EGF signaling cascade  
**ALCAM**: involved in immune response  

**CCs from Quantitative RT-PCR**  
**BDNF**: neurotrophic factor playing a role in regulation of stress response  
**PGK1**: involved in glycolysis  
**RGS**: hydrolyzed GTP to GDP  
**VCAN**: plays a central role in tissue morphogenesis and maintenance  
**RPSK6KA2**: involved in the EGF signaling cascade  
**ALCAM**: involved in immune response  

**Associated with pregnancy (Anderson et al., 2009)**

**associated with Oocyte competence and embryo development (McKenzie et al., 2004)**

**Quantitative RT-PCR**  
**Correlated with oocyte maturity, low fragmentation and embryo development (Adriaenssen et al., 2010)**