Dynamic expression of Lrp2 pathway members reveals progressive epithelial differentiation of primitive endoderm in mouse blastocyst.

François Gerbe, Brian Cox, Janet Rossant, Claire Chazaud

Dynamic expression of Lrp2 pathway members reveals progressive epithelial differentiation of Primitive Endoderm in mouse blastocyst

François Gerbe[1], Brian Cox[2,3], Janet Rossant[2,3] and Claire Chazaud[1*].

[1] INSERM, U384, Clermont-Ferrand F-63001, France ; Université d'Auvergne, Clermont-Fd.

[2] Department of Developmental and Stem Cell Biology, The Hospital for Sick Children, Toronto, Ontario, Canada

[3] Department of Molecular and Medical Genetics, University of Toronto, 600 University Avenue,Toronto, Ontario, M5G 1X5, Canada.

*Corresponding Author :
Claire Chazaud
INSERM UMR384
Faculté de médecine
28 place H. Dunant BP38
63001 Clermont-Fd
FRANCE

Tel : +33 4 73 17 83 80
Fax : +33 4 73 27 61 32
e-mail : claire.chazaud@u-clermont1.fr

ABSTRACT

Mesenchyme to epithelium transitions are crucial to embryonic development. The early mouse embryo offers an excellent model to study epithelium formation as during the first three days of development two epithelia are formed, the Trophectoderm (TE) and the Primitive Endoderm (PrE). We have previously shown that PrE cells are determined within the blastocyst ICM long before epithelium formation. In this work, we isolated Lrp2 as a novel PrE precursor (pre-PrE) marker by using a microarray strategy that combines a transcriptome analysis of three stem cell lines and early embryos. A detailed expression analysis shows that Lrp2 expression is induced in late E3.5 embryos indicating that pre-PrE cells are progressively maturing prior to polarization into an epithelium. Furthermore, the subcellular location of Lrp2, Disabled-2 (Dab2) and Collagen-IV shows that the epithelial structure is acquired in individual cells through successive steps.

Key words : blastocyst, epithelium, cell differentiation, primitive endoderm, cell polarisation, microarray.

INTRODUCTION

During embryonic development, cells are subjected to dynamic changes in their morphology, position and interactions with their environment. In particular, the transition between two different cell states, mesenchymal and epithelial, is fundamental to organogenesis. In early mouse development, two distinct epithelia, the trophectoderm (TE) and the primitive endoderm (PrE) have differentiated by the embryonic day 4.5 (E4.5) whereas the epiblast (Epi) remains as a mass of undifferentiated cells (Yamanaka et al., 2006). TE and PrE are extraembryonic tissues that are required for nutrient exchange and have also been shown to play important roles in the induction of the antero-posterior axis (Ang and Constam, 2004; Rossant and Tam, 2004). The mechanism of epithelium formation is different in the formation of the TE and PrE (Johnson and McConnell, 2004; Yamanaka et al., 2006). After fertilization, all blastomeres of the developing embryo are equipotent until the 8-cell stage when compaction occurs. During this process, the cells polarize by maximizing their contacts through E-cadherin binding, producing the first proto-epithelium of development. During the subsequent two rounds of mitosis each cell either divides asymmetrically producing an inner and an outer cell or symmetrically producing two outer cells. Outer cells remain polarized and generate the TE epithelium with adherens and tight junctions. The etiology of PrE differentiation is not known but recent data has shown that PrE cells are already determined by the blastocyst stage (E3.5) (Chazaud et al., 2006). However at this early stage PrE precursor (pre-PrE) cells, expressing the PrE marker Gata6, have not formed an epithelium but are scattered throughout the inner cell mass, intermingled with other cells of the future Epi expressing Nanog. Lineage tracing experiments have shown that individual cells of the ICM contribute only to the Epi or PrE, suggesting that between E3.5 and E4.5 pre-PrE cells move toward the surface of the ICM to form the epithelium (Chazaud

et al., 2006). The mechanisms involved in these morphogenetic movements are currently unknown, but cell adhesion mechanisms are certainly implicated (Yamanaka et al., 2006).

Low-density lipoprotein Receptor-related Proteins (Lrp) are transmembrane receptors that share homology with mammalian LDL-Receptor (Fisher and Howie, 2006; Nykjaer and Willnow, 2002). One Lrp family member, Lrp2, also known as gp330 or Megalin (Saito et al., 1994) has been shown to be expressed during fetal development in the neuroepithelium (Assemat et al., 2005a) and the endodermal portion of the yolk sac (Assemat et al., 2005b; Drake et al., 2004; Maurer and Cooper, 2005; Yang et al., 2007). In adults, Lrp2 expression can be detected in various tissues but the protein is mainly found in absorptive epithelia, such as renal proximal tubutles, gallbladder or mammary epithelia (Fisher and Howie, 2006). Historically, Lrp family members were generally thought to only be involved in receptor-mediated endocytosis. Newer research has shown Lrp2 to have a wide variety of ligands, such as lipoproteins, proteases and protease/inhibitor complexes, plasminogen and their activators, albumin and some drugs like Gentamycin (Christensen and Birn, 2002; Hussain et al., 1999). Analysis of *Lrp2* mutant mice has highlighted an important role during embryo development (Willnow et al., 1996), implicating this receptor in the binding and regulation of morphogenic signaling pathways such as Sonic Hedgehog (Shh) and BMP4 (McCarthy et al., 2002; Spoelgen et al., 2005).

Here we show that Lrp2 is an early marker of PrE differentiation. A microarray-based screen identified this gene and other members of this signaling pathway as potential PrE specific genes. Detailed expression analysis confirms that Lrp2 is expressed in pre-PrE cells as early as E3.5. Interestingly, the timing of expression as well as the changes in sub-cellular localization of Lrp2 highlights a progressive maturation of pre-PrE cells to finally form the PrE epithelium.

MATERIALS AND METHODS

Embryo microarray

Embryos from each stage were pooled and mRNA extracted with the Micro-Fast Track isolation kit (Invitrogen, 45-0036). Two independent pools (for two independent experiments) were made for each stage, in average 106 embryos at E2.5, 61 at E3.5, 47 at E4.5, 28 at E4.8, 17 at E5.5 and 19 at E6.5 (see Fig.1A for stages description). SMART (Clontech) reverse transcription and PCR was adapted for each developmental stage to amplify cDNAs. To generate probes for array hybridization, 1 microgram of cDNA was labeled by incorporation of either Cy5 or Cy3-dCTP during random hexamer-primed primer extension in the presence of Klenow DNA polymerase (Roche) according to Livesey et al.(Livesey et al., 2000). Poly L-lysine coated slides spotted with probe sequences from the NIA mouse 15K cDNA library (Ko et al., 2000) were purchased from the University Health Network Microarray Centre (www.microarrays.ca). Labeled probes were hybridized according to Wigle et al. (Wigle et al., 2002). Slides were scanned with a Genepix Axon 4000 microarray scanner. Spot intensities were quantified and median back ground corrected with the supplied Genepix software and exported as tables. Duplicate samples were analyzed and probe spot intensities were averaged. Expression data was set as a $\log_2$ ratio of expression at E2.5. Data has been deposited at GEO under accession number (GSE8339).

Stem cell microarray

Stem cell microarray data on Affymetrix MGU74x2A chips was obtained from the GEO database (http://www.ncbi.nlm.nih.gov/geo/) for XEN cells (GDS1763), ES and TS cells (GSE3766). Cel files were downloaded and processed in GCOS ver1.4 software (Affymetrix).

Expression data of XEN cells were taken as ratios of ES and TS cells using GCOS ver1.4 software (Affymetrix).

Data mining

TS, XEN and ES cell expression ratios were filtered to obtain XEN enriched genes based on the fold change call and expression call. A gene was considered enriched if it was called as present in both XEN cell data sets and had increased expression compared to both TS and ES cells. Data filtering was preformed in Microsoft Excel.

Comparison of array platforms

BLAST (Altschul et al., 1990) was used to reciprocally compare all of the NM, XM sequences represented on the Affymetrix MGU74v2A chip, as of March 2007, (http://www.affymetrix.com) with the ~15,000 probe sequences on the NIA mouse 15k chip, as of March 2007 (http://lgsun.grc.nia.nih.gov/). NM and XM FASTA formatted sequences were obtained from NCBI Genbank release 155. Reciprocal best matches were further filtered for only those with an e score of $1x10^{-100}$ or lower with a minimum of 98% identity.

Embryos

CD1 or ICR embryos were collected by flushing or dissecting the uterus. Noon of the vaginal plug was considered as E0.5. Embryos were collected according to Fig.1. Early E3.5 embryos were collected in the morning and correspond to expanding blastocyst with about v/v ICM/cavity. Late E3.75 embryos are expanded blastocysts collected 6 to 8 hours later

Semi quantitative RT-PCR

PolyA+ RNA was extracted from about 50 pooled embryos from E2.5, E3.5 and E4.5 stages (Fig.1A) with the Micro-FastTrack kit (Invitrogen). RNA was subjected to first strand cDNA synthesis using SuperScript III reverse transcriptase (Invitrogen). Newly synthesized cDNA was amplified using the SMART protocol (Clontech). Products of amplification were used as matrices for gene-specific PCR.: Lrp2 forward 5'-CCTTGCCAAACCCTCTGAAAAT-3'; Lrp2 reverse 5'-CACAAGGTTTGCGGTGTCTTTA-3'; Lrpap1 forward 5'-AACGCCCTCAATGAAGACAC-3'; Lrpap1 reverse 5'-TTCTGGTGGGAAATCTCCAG-3'; G3PDH forward 5'-ACCACAGTCCATGCCATCA-3'; G3PDH reverse 5'-TCCACCACCCTGTTGCTGT-3'; Dab2 forward 5'-AAAGGACATTCCCAGTGACG-3'; Dab2 reverse 5'-GAGCGAGGACAGAGGTCAAC-3'.

Fluorescent whole mount *in situ* hybridization

Whole-mount fluorescent *in situ* hybridization was performed according to Chazaud et al (2006). Probes for *Lrp2*, *Lrpap1* and *Amn* were obtained by RT-PCR as follows. PCR primers for Lrpap1: 5'-ATGGCGCCTCGAAGAGAGAGGGTCT-3' and 5'-CCAGTGGCAGGCAGGTTTATGTGAT-3' (nucleotides 7 to 1391 sequence **NM_013587**). Lrp2, primers: 5'-CACCAGTGCCTCTGTGAAGA-3' and 5'-GTCAGCATCGTACGCTTTCA-3' (nucleotides 1276 to 3069 of sequence **NM_001081088**). Amn primers:5'-GACGAGGACCTGACTGCTTT-3' and 5'-CCCGAATGGTAACAGCACTT-3', (nucleotides 537 to 1613 of sequence **NM_033603**).

Immunohistology

Whole-mount immunostaining was carried out according to Chazaud and Rossant (Chazaud and Rossant, 2006).

Antibodies

Antibodies were used as follows: goat anti-Megalin serum (1/10000, (Leheste et al., 2003); rabbit anti-Nanog (1/700, Cosmo Bio, Japan, RCAB0002P-F or 1/700, AbCam, ab21603), goat anti-Cubilin (1/200, Santa Cruz, sc-20607), mouse anti-Dab2 (1/400, BD Biosciences, 610464), rabbit anti-collagen IV (1/200, (De Arcangelis et al., 1996)). Due to a strong cross-reaction with Lrp2 antibody, zona-pellucida was removed from the embryos on Fig. 3 and 4. Nuclei were either stained with YOYO1 (Molecular Probes, 1/1000), Hoechst B-2883 (Sigma, 1/10000) or Draq5 (Alexis Biochemicals, 1/5000).

Microscope

Pictures were taken with Olympus confocal microscope (FV300, 40X oil immersion-objective with NE 1.0) or with Zeiss confocal microscope LSM510 Meta 40x oil immersion-objective with NE 1.3). LSM510, Fluoview FV1000 and Photoshop softwares were used to visualize the data. All embryo images are individual laser confocal sections.

RESULTS

Microarray analysis

Deposits of large microarray datasets afford an excellent opportunity for the data mining of candidate genes. Three stem cell lines, ES, TS and XEN, derived from the first three lineages of the blastocyst, the Epi, TE and PrE respectively, have been previously characterized by microarray analysis (Kunath et al., 2005). We reasoned that a comparison of the microarray data of all three of these stem cell lines should reveal a candidate list of genes enriched to the XEN cells and that as these were derived from the PrE, a subset may be expressed *in vivo* in the PrE. A list of 750 XEN enriched probe sets was generated by selecting probe sets with a present call in both XEN array data sets and with an increased expression relative to both TS and ES cells (See Experimental Procedures). This list included several known PrE genes such as *Gata6*, *Dab2* and *Sox17* (Chazaud et al., 2006; Kurimoto et al., 2006; Yang et al., 2002; Yasunaga et al., 2005) (Supplemental Table 1).

To further filter this large set of genes, and to assess their potential expression *in vivo* in the PrE, we filtered these candidates against a microarray series of embryo development from E2.5 to E6.0 that we generated. The stages selected were (Fig.1A) : E2.5 (equipotent blastomeres), E3.5 (induction of PrE), E4.5 (morphological differentiation of PrE), E4.75 (differentiation of extraembryonic ectoderm), E5.5 (visceral endoderm differentiation, induction of DVE/AVE) and E6.0 (mesendoderm induction). Data for each time point was collected in duplicates, normalized and averaged (see Experimental Procedures). Expression changes were set as a ratio of signal intensities relative to E2.5 (data not shown, full data set is available at GEO accession number GSE8339). We mapped between the two microarray platforms by means of reciprocal BLAST (See Experimental Procedures), which resulted in the selection of 608 probes on the embryo array (data not shown) corresponding to genes in our XEN candidate list.

As the cDNA arrays do not have a normalized binding affinity we could not justify excluding a gene based on the magnitude of its change and therefore consider only its vector value as positive or negative relative to E2.5 before the differentiation of the PrE (Fig.1A). As previously shown (Chazaud et al., 2006) at E3.5 the pre-PrE cells are already determined and mixed with the pre-Epi cells of the ICM (Fig.1A). Later, by E4.5, this has resolved into a monolayer of PrE cells underlying the Epi, separating it from the blastocoel cavity (Fig.1A). As the embryo matures the TE and the Epi become a greater proportion of the embryo mass relative to the PrE making analysis of expression difficult at later time points. For these reasons we looked for genes enriched in XEN cells with increased expression at E3.5 (early PrE candidates, 179 probes) or E4.5 (late PrE candidates, 93 probes) relative to expression at E2.5. These candidates contained the known markers *Gata6* (early) and *Dab2* (late) (Supplemental Table 2).

Next we tested the early and late PrE candidate list for enrichment of Gene Ontology terms using the BinGO plug-in for Cytoscape and MGI GO build (April 6, 2007). This revealed many enriched terms after Benjamini & Hochberg False Discovery Rate (FDR) correction, with p values less than 0.05 (Supplemental Table 3). We noted that there was a general enrichment for genes annotated to cell component terms cytoplasm, endoplasmic reticulum, vesicles and endosomes (P-value<0.05 in all cases), but terms such as nucleus and mitochondria were not enriched. These observations were interesting as the PrE and its derivative tissues are required for nutrient exchange with the growing embryo until the placenta has developed. We were especially intrigued by the presence of *Lrp2* and its molecular chaperone *Lrpap1* (Willnow et al., 1995) (Fig. 1B). Furthermore, Dab2 , which is also present in these GO categories, is known to interact with the cytoplasmic domain of Lrp2 (Gallagher et al., 2004) (Fig. 1B). As at least three members of Lrp2 pathway were selected by our screen, we decided to carry out further studies.

We verified the microarray expression data of *Lrp2* and *Lrpap1* and *Dab2* by RT-PCR (Fig. 1C). The results confirmed our embryo microarray data that transcripts for *Lrp2* and *Lrpap1* were undetectable or barely detectable at the 8-cell morula stage and were greatly up-regulated in E3.5 and E4.5 embryos. *Dab2* did not show increased expression until E4.5, which is also consistent with our embryo microarray data. We therefore decided to examine *in situ* the expression profile of these genes.

Expression pattern of Lrp2 and associated proteins in implanted blastocysts:

Using laser scanning microscopy, we found that *Lrp2* transcripts are exclusively present in the PrE at E4.5 (n=5; Fig. 2A), and a strong expression is observed in the apical part of PrE cells (Fig. 2D, inset) by immunostaining with Lrp2 serum (n=10; Leheste et al., 2003). mRNA for *Lrpap1*, is also detected in the PrE of E4.5 blastocysts (n=4; Fig. 2B). We decided to extend our study to Cubilin and Amnionless, as these proteins are known to interact and form complexes with Lrp2 (Hammad et al., 2000; Moestrup et al., 1998; Strope et al., 2004) and all three have been shown to be co-expressed in many different tissues (Drake et al., 2004; Strope et al., 2004). Moreover, they have been shown to be expressed in PrE derivatives (Assemat et al., 2005b; Drake et al., 2004; Kalantry et al., 2001; Tomihara-Newberger et al., 1998). Confocal imaging of E4.5 embryos clearly shows Cubilin and Amnionless expressed in the PrE (n=4 and n=2 respectively; Fig. 2C, E). Furthermore, Cubilin receptor is strongly detected in the apical part of PrE cells (Fig. 2E, inset).

Lrp2 is gradually expressed in the ICM.

As PrE is already induced at E3.5, we examined whether the selected candidates are expressed in the ICM of E3.5 blastocysts. The *in situ* hybridization technique was not sensitive enough to detect *Lrp2* transcripts at that stage of development (data not shown), so we used Lrp2-specific antibodies (Leheste et al., 2003). Lrp2 is expressed in the trophoblasts

as was described before (Assemat et al., 2005b; Gueth-Hallonet et al., 1994), and we found a novel expression in some cells of the ICM, in a "salt and pepper" pattern (Fig. 3). The Lrp2 ICM-expressing cells are either facing the blastocoelic cavity, in contact or not with trophoblasts (Fig 3A), or localized deeper in the ICM (Fig. 3B). Thus, there is no preferential position for these cells within the ICM. We noticed that whereas all the trophoblasts are strongly stained, only one to two cells of E3.5 ICM are weakly expressing Lrp2. We hypothesized that this expression could increase during time, therefore we analyzed blastocysts at different time points from early morning of E3.5 to late afternoon. We noticed that older E3.75 blastocysts had more Lrp2 expressing cells (compare Fig. 3A to 3D). A thorough analysis on several embryos allowed to establish a correlation between the blastocyst stages, determined by the number of ICM cells, and the number of ICM cells expressing Lrp2 (Fig. 3E). Taken together, these data show that Lrp2 is progressively expressed in some ICM cells during blastocyst development.

Lrp2 is expressed in PrE cells.

We have previously shown by lineage tracing and gene expression experiments that ICM cells of E3.5 blastocysts are already determined to either the PrE or Epi cell lineages (Chazaud et al., 2006). As Lrp2 is strongly expressed in the PrE at E4.5, the Lrp2 expressing cells at E3.5 should correspond to pre-PrE. To determine the identity of Lrp2 expressing cells, we carried out double immunostaining for Lrp2 and Nanog, which is an Epi-specific marker (Chambers et al., 2003; Chazaud et al., 2006; Mitsui et al., 2003). As expected, Nanog and Lrp2 have exclusive expression in ICM cells at every time point analyzed (Fig. 4A). Cell counting analysis reveals that whereas the proportion of Nanog expressing cells remains nearly the same, around 42%, the amount of Lrp2 ICM-expressing cells almost doubled during blastocyst maturation (Fig. 4B). Thus, Lrp2 expressing cells belong to the pool of pre-

PrE cells, however, *Lrp2* induction is not synchronized within all PrE cells but appears to be randomly and progressively activated until the epithelium has formed.

Lrp2 subcellular localization during PrE maturation :

We noticed that Lrp2 localization within pre-PrE and PrE cells is different according to the stage and the position of the cell (Fig. 3A-C'; 5A-C). Indeed, in younger embryos (Fig. 3A) or when the pre-PrE cell is not at the blastocoelic surface (Fig. 3B-C'; 5A), Lrp2 is homogenously distributed in the cytoplasm around the nucleus. At high magnification we observed that Lrp2 localization appeared punctate and was not associated with the plasma membrane (Fig. 3C'). This was in sharp contrast to the strong polarized expression in the epithelium at E4.5 (Fig. 2D, 5B, C). We did observe cells with polarized Lrp2 expression in E3.75 embryos, but this was only in pre-PrE cells that had contact with the blastocoelic cavity (n=8; Fig. 5A, arrowhead). The Lrp2 localization in these cells corresponds to the future apical side of the PrE epithelium. We also noted that these cells had a strong increase in the level of Lrp2 expression (Fig.3C, arrowhead; 5A). Our observations imply that PrE cells can polarize individually once they reach the ICM surface (Fig. 3C; 5A) and do not form an epithelium in unison.

As a further test of cell polarity we analyzed the expression of the epithelium marker Collagen-IV, which is a major component of basal lamina. Collagen-IV has a strong expression in the basal lamina of the PrE and mural trophectoderm at E4.5 (Supplemental Fig.1). In E3.75 embryos, we observed Collagen-IV expression associated with Lrp2 in cells either deep within the ICM or at the blastocoelic surface (n=6; Supplemental Fig.1). This is evidence that pre-PrE cells are preparing for epithelialization by producing basement membrane components.

Dab2 has been recently reported to position Lrp2 at the cell surface and also was shown to be necessary for Lrp2 and Cubilin trafficking (Maurer and Cooper, 2005; Yang et

al., 2007). *Dab2* mutant embryos have a defective visceral endoderm (Morris et al., 2002) or a stronger phenotype as no PrE epithelium differentiates (Yang et al., 2002). Interestingly, these mutants are able to differentiate pre-PrE cells, as they expressed the PrE marker *D.Biflorus* agglutinin (Yang et al., 2002), but these cells are unable to move to the surface of the ICM to form an epithelium. This phenotype could be recapitulated *in vitro*, as embryoid bodies with ES cells knocked down for *Dab2* express *Gata4* (Rula et al., 2007). As observed later in the VE, Lrp2 and Dab2 are co-expressed on the apical surface of the PrE epithelium at E4.5 (Fig. 5B,C). By E3.5, we were not able to detect any Dab2 expression within the ICM, despite a strong amplification with the tyramid system (data not shown). The lack of detection is probably due to the very low amount of the protein as the RNA was also undetectable at E3.5 by RT-PCR (Fig.1C) and had no significant change in our microarray analysis (Supp. table 2). Nevertheless, at E3.75, in some PrE that cells have reached the surface, Dab2 was induced and co-localized on the apical side of the cell with Lrp2 (n=2; Fig. 5A, arrowhead).

DISCUSSION

We have recently shown that the ICM of E3.5 blastocysts is not homogenous but composed of cells either fated to be Epi or PrE (Chazaud et al., 2006; Rossant et al., 2003). This finding was also supported by a recent microarray study analyzing the expression of single cells in the ICM (Kurimoto et al., 2006). However, only *Nanog* and *Gata6* genes have been reported to be expressed within the embryo in pre-Epi and pre-PrE cells respectively (Chazaud et al., 2006). Our aim in this current study was to use microarray data to find genes involved in PrE differentiation, thus we focused on embryonic stages E2.5 through E4.75 that comprise PrE induction and epithelialization. However, embryos contain derivatives of all

three cell lineages so we required a means of sorting our lineage specific gene expression patterns. Three previously derived stem cell lines, ES, TS and Xen, correspond to the three cell lineages of the mouse blastocyst, Epi, TE and PrE respectively as they can be derived from the blastocyst, express many genes in common and have similar behavior in stem cell/embryo chimeras (Beddington and Robertson, 1989; Kunath et al., 2005; Tanaka et al., 1998; Yamanaka et al., 2006). To identify genes specifically expressed in the PrE, we compared our microarray data set to previously published microarray datasets of the three stem cell lines ES, TS and Xen cells. We first selected a list of candidate PrE genes as those that had enriched expression in Xen cells, relative to TS and ES. These candidates were then further filtered for their expression *in vivo* by comparison to our early embryo microarray dataset. This selected a set of 271 PrE candidate genes. This list contained several known PrE markers, such as *Gata4*, *Gata6*, *Pdgfra*, *Dab2*, and *Itga5*. More broadly, an analysis of Gene Ontology (GO) terms revealed enrichment for genes of endocytic pathways. This is reflective of the biological function of the visceral and parietal endoderm (PrE derivative tissues), as a nutrient exchange interface and support for the rapidly proliferating Epi.

Among candidate PrE genes detected by the microarray, the Lrp2 pathway was especially evident as its chaperone Lrpap1 was present, as was Dab2. Dab2 is an adaptor protein that binds to Lrp2 (Gallagher et al., 2004) and is known to be involved in PrE differentiation (Maurer and Cooper, 2005; Morris et al., 2002; Yang et al., 2002). As Lrp2 mutant embryos exhibit a late postimplantation phenotype (Willnow et al., 1996), this gene does not play an obligate role in PrE formation. Thus Dab2 must be interacting with other proteins to mediate its essential role in the formation of the PrE epithelium. We show that other components of the Lrp2 pathway, such as Cubilin, Amnionless (Assemat et al., 2005b; Kalantry et al., 2001), are all specifically expressed in the mouse PrE at E4.5 along with Lrp2, *Lrpap1* and Dab2 (Yang et al., 2002). The expression of Dab2, Cubilin and Amnionless were

all apparent after the epithelium has formed, suggesting that this endocytic machinery is functional at that stage. It is interesting to note that Cubilin and Lrp2 have both been reported to be expressed in the TE by E3.5 and even earlier in the morula (Gueth-Hallonet et al., 1994; Assemat et al., 2005; and data not shown) but expression is completely lost in the trophoblasts by E4.5, when both are strongly expressed in the PrE. This is suggestive of a relay between first the TE and then the PrE in the absorption and provision of nutrients to the developing epiblast.

We previously observed *Gata6* and *Nanog* to be expressed uniquely in earlier stages in pre-PrE and pre-Epi cells of the E3.5 blastocyst. Lrp2 was also expressed in a few cells of the E3.5 ICM and double staining with Nanog clearly identifies these cells as pre-PrE cells as they are not co-localized to Nanog expressing cells. Thus Lrp2 is a new pre-PrE cells marker that confirms the "salt and pepper" model of PrE and Epi induction (Chazaud et al., 2006; Rossant et al., 2003). Unfortunately, we are unable to co-localize Lrp2 and Gata6 due to antibody incompatibility. Interestingly, the proportion of Lrp2 expressing cells increases while the embryo matures. Lrp2 expression was randomly activated in cells that could be located either deep in the ICM or closer to the blastocoelic surface. This activation could be driven by Gata6, which is expressed in pre-PrE cells, as Gata and Sox response elements are found in Lrp2 promoter and intronic regions (data not shown). It is intriguing that Lrp2 is not expressed in all PrE cells, but its activation could be following Gata6 activity. We had previously noticed that the level of expression among the Gata6 labeled cells is not homogenous (Chazaud et al., 2006; C.C., unpublished observation). Nevertheless, we cannot exclude that other independent factors such as Sox7 or Sox17, might be required to activate *Lrp2* expression.

Strikingly, Lrp2 subcellular localization changed according to the maturity of the embryo and the position of the pre-PrE cell within the ICM. Lrp2 positive cells of young, expanding blastocysts (E3.5) are not polarized as indicated by the localization of Lrp2 even when lying at the blastocoelic surface. In older embryos (E3.75), the cells that have joined the surface of the blastocoelic cavity adopt an epithelial character by expressing Collagen-IV and remodeling the localization of Lrp2 with a stronger accumulation towards the blastocoelic cavity. The localization of Lrp2 to the apical membrane perfectly matches the beginning of Dab2 expression. Interestingly, we did note that individual pre-PrE cells can polarize at E3.75 if they are located at the surface of the ICM, despite being adjacent to pre-Epi cells. A similar observation was made during PrE differentiation in embryoid bodies (EBs) (Rula et al., 2007). However there are differences between *in vitro* and *in vivo* models. EBs are made of aggregates of ES cells that all have an Epi character at the beginning. In EBs inside cells seem to stop differentiating PrE cells and remain Epi cells once PrE epithelium is fully formed (Rula et al., 2007). In blastocysts, ICM cells are already committed to Epi or PrE and thus pre-Epi cells do not produce PrE cells. Once pre-PrE cells are determined and start differentiation, cell sorting towards the surface and epithelium formation involve similar mechanisms, requiring cell adhesion molecules such as Laminin C1, Integrin β1 or Maspin (Fassler and Meyer, 1995; Gao et al., 2004; Smyth et al., 1999; Stephens et al., 1995) or possibly Collagen-IV as we found it expressed in a subset of ICM cells. Also, the role of TE is not known in this process. Potentially, cues from the polar trophectoderm must repel pre-PrE cells, otherwise there would be PrE all around the Epi like in EBs. However, this mechanism is complicated by the presence of Trophectodermal Processes (TP) (Fleming et al., 1984) that separate the ICM from the blastocoelic cavity. Currently it is not known what the role of TP in the polarization of the PrE is or if polarization of factors such as Dab2 precedes or proceeds retraction of the TP.

Dab2 is known to be involved in two phases of PrE development. It is first required for positioning pre-PrE cells at the surface of the ICM (Rula et al., 2007; Yang et al., 2002), although the exact mechanisms remain unknown, mutation of a Dab2 interaction partner Integrin β1 also leads to PrE formation failure (Fassler and Meyer, 1995; Stephens et al., 1995). Second, once the epithelium has formed, Dab2 is involved in endocytosis, together with Lrp2 (Maurer and Cooper, 2005; Morris et al., 2002). In this study, we could not detect Dab2 protein or RNA before epithelium formation, in agreement with our microarray analysis. Therefore if Dab2 is required at the "cell positioning" phase it must be in very low amounts. Like many PrE genes, the level of Dab2 is greatly increased at E4.5 and it was clearly detected on the apical surface of the epithelium. In our experiments it was always co-localized with Lrp2 when pre-PrE cells start to polarize, in accordance with data showing the recruitment of Lrp2 by Dab2 in embryoid bodies (Yang et al., 2007). It is striking to note that in Dab2 knocked-down ES cells Lrp2 has no apical membrane localization (Yang et al., 2007), similar to our observations of the inner ICM pre-PrE cells. This suggests that Dab2 may recruit Lrp2 to the apical membrane once the pre-PrE cell has reached the ICM surface. However, Lrp2 polarization in epithelial cells is not strictly dependent on Dab2, as Lrp2 in the TE is polarized despite an absence of Dab2 expression. Thus other factors can recruit Lrp2 to the membrane and the transient expression of these proteins in the TE could reflect different mechanisms.

Clearly the events driving PrE differentiation and epithelialization are not yet fully understood, however the addition of these novel markers will be beneficial to the continued study of PrE differentiation and polarization. Importantly, we confirm with novel markers that PrE starts differentiating in cells of the ICM in an apparent random fashion. The data presented here show that PrE cells progressively acquire epithelial properties (Fig. 6) during blastocyst maturation. Once determined as PrE, the cell will gradually start to accumulate

structural proteins of the epithelium such as Collagen-IV or Laminins without undergoing epithelialization. Collagen and Laminins are required for basement membrane formation once the epithelium has formed but their earlier expression may aide in cell sorting of the PrE to the surface of the blastocoelic cavity. Concurrently, the pre-PrE cells inside the ICM prepare for their future functional activity by starting to express and accumulate proteins from the endocytic pathway.

ACKNOWLEDGMENTS

Authors want to thank Dr. Willnow and Dr. Simon-Assmann for the generous gift of anti-Lrp2 serum and anti Collagen-IV antibodies respectively. We thank Y. Yamanaka for a critical reading of the manuscript. FG was supported by the "Fondation pour La Recherche Médicale" (FRM). BC was supported by a CIHR Doctoral Research Fellowship. This work was supported by Avenir (INSERM) and EpiPro (INCA) grants.

REFERENCES

Altschul, S. F., et al., 1990. Basic local alignment search tool. J Mol Biol. 215, 403-10.

Ang, S. L., Constam, D. B., 2004. A gene network establishing polarity in the early mouse embryo. Semin Cell Dev Biol. 15**,** 555-61.

Assemat, E., et al., 2005a. Overlapping expression patterns of the multiligand endocytic receptors cubilin and megalin in the CNS, sensory organs and developing epithelia of the rodent embryo. Gene Expr Patterns. 6**,** 69-78.

Assemat, E., et al., 2005b. Expression and role of cubilin in the internalization of nutrients during the peri-implantation development of the rodent embryo. Biol Reprod. 72**,** 1079-86.

Beddington, R. S., Robertson, E. J., 1989. An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. Development. 105**,** 733-7.

Chambers, I., et al., 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell. 113**,** 643-55.

Chazaud, C., Rossant, J., 2006. Disruption of early proximodistal patterning and AVE formation in Apc mutants. Development. 133**,** 3379-87.

Chazaud, C., et al., 2006. Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Dev Cell. 10**,** 615-24.

Christensen, E. I., Birn, H., 2002. Megalin and cubilin: multifunctional endocytic receptors. Nat Rev Mol Cell Biol. 3**,** 256-66.

De Arcangelis, A., et al., 1996. Inhibition of laminin alpha 1-chain expression leads to alteration of basement membrane assembly and cell differentiation. J Cell Biol. 133**,** 417-30.

Drake, C. J., et al., 2004. Differential distribution of cubilin and megalin expression in the mouse embryo. Anat Rec A Discov Mol Cell Evol Biol. 277**,** 163-70.

Fassler, R., Meyer, M., 1995. Consequences of lack of beta 1 integrin gene expression in mice. Genes Dev. 9**,** 1896-908.

Fisher, C. E., Howie, S. E., 2006. The role of megalin (LRP-2/Gp330) during development. Dev Biol. 296**,** 279-97.

Fleming, T. P., et al., 1984. Trophectodermal processes regulate the expression of totipotency within the inner cell mass of the mouse expanding blastocyst. J Embryol Exp Morphol. 84**,** 63-90.

Gallagher, H., et al., 2004. The adaptor disabled-2 binds to the third psi xNPxY sequence on the cytoplasmic tail of megalin. Biochimie. 86**,** 179-82.

Gao, F., et al., 2004. Maspin plays an essential role in early embryonic development. Development. 131**,** 1479-89.

Gueth-Hallonet, C., et al., 1994. Gp330 is specifically expressed in outer cells during epithelial differentiation in the preimplantation mouse embryo. Development. 120**,** 3289-99.

Hammad, S. M., et al., 2000. Megalin acts in concert with cubilin to mediate endocytosis of high density lipoproteins. J Biol Chem. 275**,** 12003-8.

Hussain, M. M., et al., 1999. The mammalian low-density lipoprotein receptor family. Annu Rev Nutr. 19**,** 141-72.

Johnson, M. H., McConnell, J. M., 2004. Lineage allocation and cell polarity during mouse embryogenesis. Semin Cell Dev Biol. 15**,** 583-97.

Kalantry, S., et al., 2001. The amnionless gene, essential for mouse gastrulation, encodes a visceral-endoderm-specific protein with an extracellular cysteine-rich domain. Nat Genet. 27**,** 412-6.

Ko, M. S., et al., 2000. Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development. Development. 127**,** 1737-49.

Kunath, T., et al., 2005. Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. Development. 132**,** 1649-61.

Kurimoto, K., et al., 2006. An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. Nucleic Acids Res. 34**,** e42.

Leheste, J. R., et al., 2003. Hypocalcemia and osteopathy in mice with kidney-specific megalin gene defect. Faseb J. 17**,** 247-9.

Livesey, F. J., et al., 2000. Microarray analysis of the transcriptional network controlled by the photoreceptor homeobox gene Crx. Curr Biol. 10**,** 301-10.

Maurer, M. E., Cooper, J. A., 2005. Endocytosis of megalin by visceral endoderm cells requires the Dab2 adaptor protein. J Cell Sci. 118**,** 5345-55.

McCarthy, R. A., et al., 2002. Megalin functions as an endocytic sonic hedgehog receptor. J Biol Chem. 277**,** 25660-7.

Mitsui, K., et al., 2003. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell. 113**,** 631-42.

Moestrup, S. K., et al., 1998. The intrinsic factor-vitamin B12 receptor and target of teratogenic antibodies is a megalin-binding peripheral membrane protein with homology to developmental proteins. J Biol Chem. 273**,** 5235-42.

Morris, S. M., et al., 2002. Dual roles for the Dab2 adaptor protein in embryonic development and kidney transport. Embo J. 21**,** 1555-64.

Nykjaer, A., Willnow, T. E., 2002. The low-density lipoprotein receptor gene family: a cellular Swiss army knife? Trends Cell Biol. 12**,** 273-80.

Rossant, J., et al., 2003. Lineage allocation and asymmetries in the early mouse embryo. Philos Trans R Soc Lond B Biol Sci. 358**,** 1341-8; discussion 1349.

Rossant, J., Tam, P. P., 2004. Emerging asymmetry and embryonic patterning in early mouse development. Dev Cell. 7**,** 155-64.

Rula, M. E., et al., 2007. Cell autonomous sorting and surface positioning in the formation of primitive endoderm in embryoid bodies. Genesis. 45**,** 327-338.

Saito, A., et al., 1994. Complete cloning and sequencing of rat gp330/"megalin," a distinctive member of the low density lipoprotein receptor gene family. Proc Natl Acad Sci U S A. 91**,** 9725-9.

Smyth, N., et al., 1999. Absence of basement membranes after targeting the LAMC1 gene results in embryonic lethality due to failure of endoderm differentiation. J Cell Biol. 144**,** 151-60.

Spoelgen, R., et al., 2005. LRP2/megalin is required for patterning of the ventral telencephalon. Development. 132**,** 405-14.

Stephens, L. E., et al., 1995. Deletion of beta 1 integrins in mice results in inner cell mass failure and peri-implantation lethality. Genes Dev. 9**,** 1883-95.

Strope, S., et al., 2004. Mouse amnionless, which is required for primitive streak assembly, mediates cell-surface localization and endocytic function of cubilin on visceral endoderm and kidney proximal tubules. Development. 131**,** 4787-95.

Tanaka, S., et al., 1998. Promotion of trophoblast stem cell proliferation by FGF4. Science. 282**,** 2072-5.

Tomihara-Newberger, C., et al., 1998. The amn gene product is required in extraembryonic tissues for the generation of middle primitive streak derivatives. Dev Biol. 204**,** 34-54.

Wigle, D. A., et al., 2002. Molecular profiling of non-small cell lung cancer and correlation with disease-free survival. Cancer Res. 62**,** 3005-8.

Willnow, T. E., et al., 1995. Functional expression of low density lipoprotein receptor-related protein is controlled by receptor-associated protein in vivo. Proc Natl Acad Sci U S A. 92**,** 4537-41.

Willnow, T. E., et al., 1996. Defective forebrain development in mice lacking gp330/megalin. Proc Natl Acad Sci U S A. 93**,** 8460-4.

Yamanaka, Y., et al., 2006. Cell and molecular regulation of the mouse blastocyst. Dev Dyn. 235**,** 2301-14.

Yang, D. H., et al., 2007. Disabled-2 is an epithelial surface positioning gene. J Biol Chem. 282**,** 13114-22.

Yang, D. H., et al., 2002. Disabled-2 is essential for endodermal cell positioning and structure formation during mouse embryogenesis. Dev Biol. 251**,** 27-44.

Yasunaga, M., et al., 2005. Induction and monitoring of definitive and visceral endoderm differentiation of mouse ES cells. Nat Biotechnol. 23**,** 1542-50.

FIGURE LEGENDS

Figure 1. (A) Representation of the stages studied in the microarray. AVE, Anterior Visceral Endoderm; DVE, Distal Visceral Endoderm (yellow); EC, Ectoplacental Cone (pink); Epi, Epiblast (blue); ExE, Extraembryonic Ectoderm (red); MS, mesendoderm (light gray); PA, ProAmniotic Cavity; PA, Parietal Endoderm (green dots); PrE, Primitive endoderm (green); TE, Trophectoderm (orange); VE, Visceral Endoderm (green). (B). Table displaying candidate genes enriched for GO terms. (C). Expression analysis of *Lrp2*, *Lrpap1* and *Dab2* by RT-PCR at indicated stages. *G3pdh* is used as reference for sample normalization.

Figure 2. Expression pattern of members of the LRP2 pathway in E4.5 embryos by fluorescent whole mount *in situ* hybridization (A-C) and fluorescent whole mount immunostaining (D, E). (A-C) *Lrp2*, *Lrpap1* and *Amn* mRNA detection. (D, E) LRP2 and Cubilin (CUB) protein localization (red). Insets are magnifications of corresponding framed cells. Nuclei (green) were stained with YOYO1.

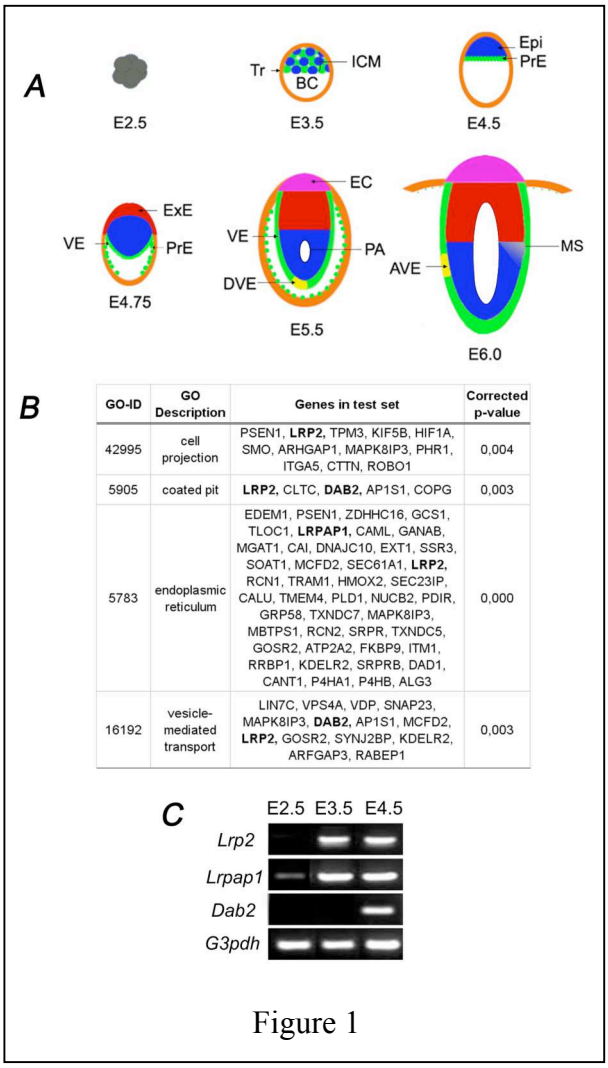
Figure 3. (A- D) Fluorescent whole mount immunostaining of LRP2 (red) with blastocysts collected at different time points. Nuclei of the ICM are stained with YOYO1 (green). Yellow asterisks indicate LRP2 labelled cells. The arrowheads is pointing towards a polarized cell. (C') Higher magnification of a LRP2 labelled cell (C, inset). (E) Number of LRP2 expressing cells in the ICM compared to the total number of ICM cells of embryos at different stages. N, nucleus.

Figure 4. (A, B, C) Double fluorescent whole mount immunostaining for LRP2 and Nanog in three embryos. The number of ICM nuclei is 24 in A, 26 in B and 27 in C. Nuclei are stained

with Hoechst and cell boundaries are visualized with an immunostaining for E-cadherin. In (B), the embryo is observed along the abembryonic – embryonic axis. (D). Table and diagrams showing the numbers and percentage of LRP2 (LRP2 +) and Nanog (Nanog +) expressing cells in early (total 15-25 ICM cells) and late (26-32) blastocysts.

Figure 5. (A, B, C). Double fluorescent whole mount immunostaining for LRP2 (red) and DAB2 (green) on E3.75 (A) and E4.5 (B,C) blastocysts. In (A) compare the expression between unpolarised (arrow) versus polarised (arrowhead) cells. (C) Higher magnification of an epithelial cell in (B). Nuclei were stained with Draq5 (blue).

Figure 6. Model for PrE epithelialization. At E3.5, the ICM is randomly composed of pre-Epi and pre-PrE cells. TPs are protecting the ICM from the blastocoelic cavity. No cells are polarised. Few PrE cells express LRP2. At E3.75, the ICM is still randomly composed of pre-Epi and pre-PrE cells. There are no data about the presence of TPs. More pre-PrE cells inside the ICM co-express LRP2 and Collagen-IV. Pre-PrE cells at the blastocoelic surface individualy polarise with an apical accumulation of LRP2 and Dab2. At E4.5 PrE epithelium is formed with an obvious basement membrane labelled by Collagen-IV. Dab2 and LRP2 are localised apically in all the cells.



GO-ID	GO Description	Genes in test set	Corrected p-value
42995	cell projection	PSEN1, **LRP2,** TPM3, KIF5B, HIF1A, SMO, ARHGAP1, MAPK8IP3, PHR1, ITGA5, CTTN, ROBO1	0,004
5905	coated pit	**LRP2,** CLTC, **DAB2,** AP1S1, COPG	0,003
5783	endoplasmic reticulum	EDEM1, PSEN1, ZDHHC16, GCS1, TLOC1, **LRPAP1,** CAML, GANAB, MGAT1, CAI, DNAJC10, EXT1, SSR3, SOAT1, MCFD2, SEC61A1, **LRP2,** RCN1, TRAM1, HMOX2, SEC23IP, CALU, TMEM4, PLD1, NUCB2, PDIR, GRP58, TXNDC7, MAPK8IP3, MBTPS1, RCN2, SRPR, TXNDC5, GOSR2, ATP2A2, FKBP9, ITM1, RRBP1, KDELR2, SRPRB, DAD1, CANT1, P4HA1, P4HB, ALG3	0,000
16192	vesicle-mediated transport	LIN7C, VPS4A, VDP, SNAP23, MAPK8IP3, **DAB2,** AP1S1, MCFD2, **LRP2,** GOSR2, SYNJ2BP, KDELR2, ARFGAP3, RABEP1	0,003

Figure 1

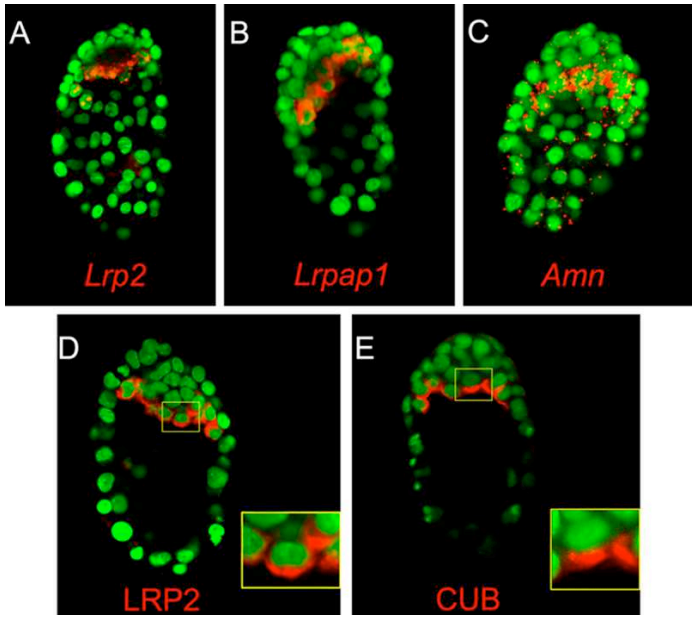


Figure 2

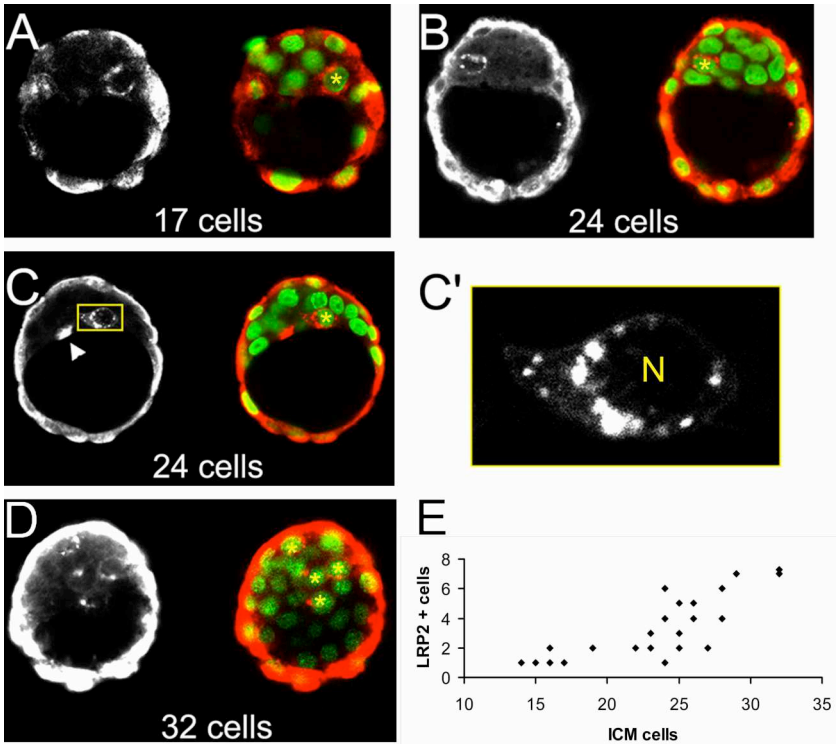
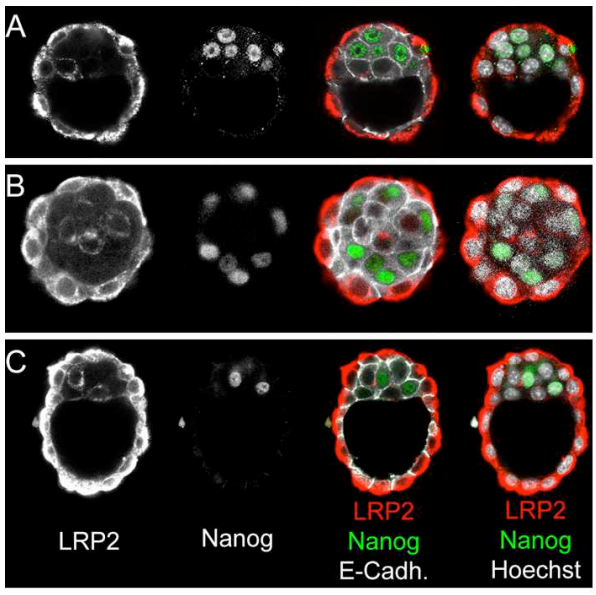


Figure 3



D

Early

Embryo	ICM cell	LRP2 +	Nanog +	Undetermined
1	16	2	6	8
2	15	1	5	9
3	14	1	6	7
4	24	4	11	9
5	25	5	12	8
6	25	3	11	11
7	25	3	12	10
Total	144	19	63	62
	%	13	44	43

Late

Embryo	ICM cell	LRP2 +	Nanog +	Undetermined
8	26	4	12	10
9	26	5	14	7
10	31	7	10	14
11	32	7	13	12
12	27	4	10	13
13	29	7	9	13
14	28	6	13	9
Total	199	40	81	78
	%	20	41	39

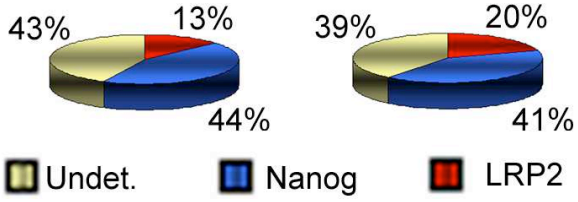


Figure 4

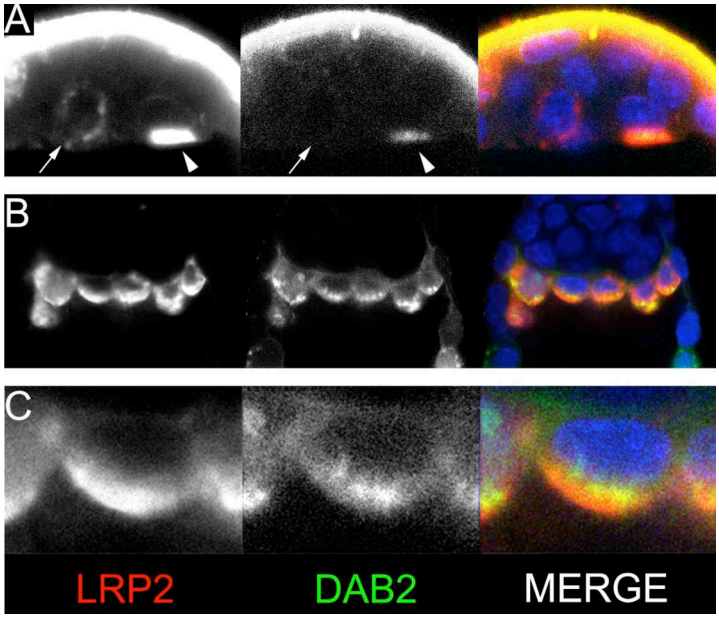


Figure 5

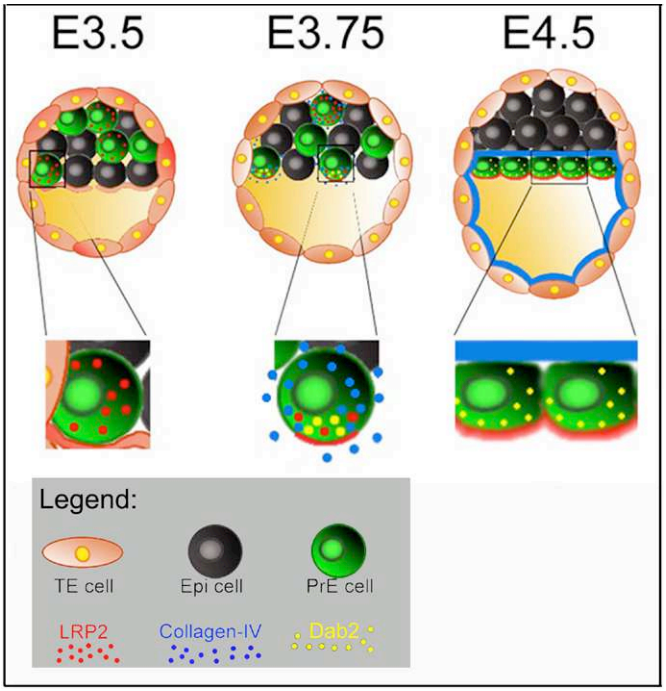


Figure 6