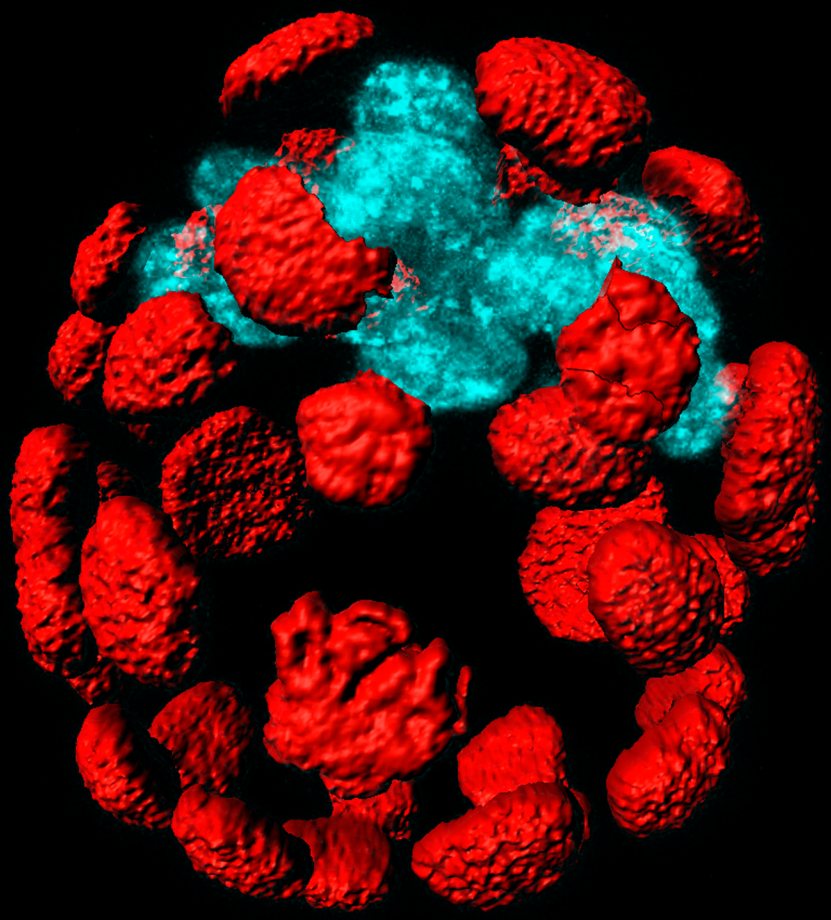
Universidad Autónoma de Madrid

Departament of Biochemistry

Understanding the first lineage choice in the embryo: Regulation of Cdx2 in the blastocyst through the Notch and Hippo pathways

Teresa Rayón Alonso

Madrid, 2014



cells
embryos
Cdx2
TEE
activity
reporter
blastocyst
Notch
CDX2
TE
blastomeres
TS
fragment
Tead4
pathway
stage
Rbpj
embryo
regulatory
enhancer
et
al
sites
blastocysts
line
TSR
RBPJ
cell
green
endogenous
mRFP
showed
wild-type
transient
preimplantation
mouse
early
transgenic
used
levels
different
trophectoderm
fragments
treated
significantly
gene
recombination
expressed
inputs
lacZ-TEE
TEAD4
morula
Rosa
driven
treatment
active
Nuclei
stages
first
H2BmRFP
transgenesis
Analysis
analysis
together
culture
DAPI
ICM
stained
compared
ES
Oct4
number
Hippo
outer
trophoblast
identified
able
CBF:H2B-Venus
transcription
upstream
TEE+
distribution
also
microinjected
locus
DNA
two
region
RO
signalling
shown
results
lines
regulation
show
cis-regulatory
pattern
binding
mRFP-TEE
inhibitor
population
red
GFP
development
detected
tested
implantation
sequence
indicate
promoter
protein
post-implantation
GFP+
TE-specific
transcriptional
activation
NICD
bars
H3K4me1
RFP

Universidad Autónoma de Madrid

Department of Biochemistry

Faculty of Medicine

Understanding the first lineage choice in the embryo: Regulation of *Cdx2* in the blastocyst through the Notch and Hippo pathways.

Teresa Rayón Alonso. BSc Biology

Doctoral thesis directed by Dr. Miguel Manzanares Fourcade

CNIC, Madrid 2014

I hereby certify that TERESA RAYON ALONSO has carried out the experimental work leading to her PhD thesis entitled "*Understanding the first lineage choice in the embryo: Regulation of Cdx2 in the blastocyst through the Notch and Hippo pathways*" under my supervision at the Centro Nacional de Investigaciones Cardiovasculares-CNIC in Madrid.

I also declare that the work presented is novel and of great importance in the field, and of sufficient quality to merit to be presented in order to obtain a PhD degree by the Universidad Autonoma de Madrid.

Madrid, 25th August 2014

Miguel Manzanares
Department of Cardiovascular Developmental & Repair
Centro Nacional de Investigaciones Cardiovasculares-CNIC
Melchor Fernandez Almagro, 3
28029 Madrid, Spain

tel: (34) 91 453 12 00, ext. 3109
fax: (34) 91 453 13 04
mmanzanares@cnic.es
http://www.cnic.es/en/desarrollo/genomica

A mi madre.

"Expect the unexpected,
and the excitement of the experiment".
Prophecies of the Analogue future. Lomography

ACKNOWLEDGEMENTS

Esta tesis doctoral es el resumen del trabajo de siete años. Muchísimas gracias por el esfuerzo y apoyo a todas y cada una de las personas que me han ayudado. Aviso a navegantes: estos agradecimientos van para largo.

En primer lugar, agradezco a Miguel Manzanares que me eligiera como becaria predoctoral después de conocerme en una entrevista en el IIB en la que yo sólo abría los ojos como platos y contaba el número de veces que decías la palabra 'bioinformática'. Has sido el mejor director de tesis que se puede tener, me has guiado en el proyecto como si el camino estuviera escrito, me has enseñado a pensar y me has apoyado en los peores momentos. Gracias por involucrarme en un montón de proyectos que me han hecho aprender lo indecible. Quiero darte especialmente las gracias por haberme dejado disfrutar de la investigación (este trabajo ha sido casi un juego). Además, te agradezco el haberme dejado ver mundo en las estancias, cursos, congresos y también permitiéndome hacer largos viajes de vacaciones a países lejanos.

El laboratorio ha cambiado mucho desde que entre allá por 2007, muchas personas han venido y se han ido y cada una de ellas forma parte de esta tesis de alguna u otra manera. Querría agradecer a Miguel Crespo su generosidad al dejarme acabar el proyecto que él empezó con su tesis. Muchas gracias por dejarme a 'cedequisdos' manga por hombro, como tú decías en tu tesis. Aprendí a microinyectar gracias a tu infinita ayuda y a pesar de mis resoplidos. Quiero agradecer a Bárbara el estar siempre ahí para solventar cualquier duda, te doy gracias por enseñarme a 'echarle piernas'. Además de una compi excepcional, muchas gracias por ser mi amiga, gracias por la ayuda infinita en Londres (sin ti puede que la estancia hubiera acabado conmigo) y gracias por las horas de charleta, bien en el parking del CNIC, en el Albertines, o por Skype. Muchas gracias a Susana por estar para, literalmente, todo. MUCHAS GRACIAS. No me olvido ni un día del ambiente que promovías en el laboratorio: abierto, alegre, profesional y trabajador. Te agradezco los mil y un momentos entrañables, las risas, todo. De las mejores cosas que me llevo de la tesis eres tú. A Melisa le doy las gracias por ser única, no imagino el día a día sin verte la cara, sin oír un 'nada que ver' o un 'recién llego'. Muchas gracias por nuestras conversaciones de ciencia y de vida, muchas gracias por las risas y por tus consejos. Sergio desde que llegó ha estado haciendo experimentos sin descanso bajo mi ala y dando frescura y alegría al labo cada día. Cuando caíste en mis manos 'de chiquito' para acabar el proyecto de fin de carrera no sabías si hacer la tesis, ahora, me temo que el gusanillo de la investigación te ha atrapado. Muchas gracias por la dedicación y el entusiasmo. Has sido el compañero de

trabajo ideal, muchas gracias por hacerme la vida más fácil y por celebrar las alegrías conmigo. El otro becario que ya lleva tiempo dando el callo es Julio. Muchas gracias por el interés, el alma y la originalidad que le pones a la vida. Conocerte hace que vivas cualquier cosa más intensamente, muchas gracias por los momentos compartidos durante la tesis y por los que vengan. Muchísimas gracias a Inma Ors por trabajar codo con codo conmigo. Llegaste al labo en un momento de la tesis donde no me daban las manos para tanto experimento e hiciste las cosas tan bien, que aprendí más yo de ti que tú de mí. Gracias por las mil horas de trabajo al día, el interés y el aguante ante tanto experimento imposible. Contigo he vivido uno de los momentos más dulces de la rampa final de la tesis, cuando por fin tanto trabajo daba sus frutos y el esfuerzo merecía la pena. A Isabel le agradezco su interés y dedicación en el trabajo, su alegría y buen humor constante y su apoyo en mi tramo final del doctorado. A Claudio le agradezco que se viniera a trabajar con los manzanares. Claudio, eres una muy buena persona y, aunque no hemos hecho experimentos juntos, tenerte alrededor da confianza. A Eva Alonso quiero agradecerle que me hiciera entender el tiempo que llevan los experimentos con ratón. Gracias a Diego por dejarme enseñarle lo poco que sabía por aquel entonces, siempre echaré de menos que no te quedaras a hacer la tesis con las manzanitas. A Eva, la cicerone que nos enseñaba bioinformática en vez de nosotros a ella, le doy las gracias por su entusiasmo y por su apoyo, aunque los veranos que has estado siempre he estado demasiado liada, tu energía positiva es contagiosa. Al resto de compis (Gonzalo, Raquel, Elena, Luis, Cris, Bea, y demás estudiantes cicerone o técnicos en prácticas varios) deciros que he aprendido mucho teniéndoos como compañeros, gracias.

Moreover, I would like to acknowledge Veronique Azuara's lab and the IRDB 1st and 2nd floor people. It was a pleasure be a member of Veronique's lab. Thank you Anne for your help with the tedious ChIP experiments, thanks Fabrice for helping me out in everything, you were my reference in London. Special thanks to Michelle, you are a great scientist and the best friend ever, I really miss you now you are so far away in SF. Special thanks to Tristán Rodriguez and the people in his lab for hosting me as if I was a visiting student in their lab. Mushísimas gracias Tristán por todo el feedback y las science discussions que tenemos cada vez que nos vemos. Also, thanks to all Stembryo members as it has been a pleasure to speak about my work in those meetings with you. I want to acknowledge too all those mates I met in meetings and with whom I've enjoyed science and life in Croatia, EMBO YIP, Consolider, Hydra and Prague.

Quiero también dar un agradecimiento muy especial los compañeros con los que coincidí en el laboratorio de Anna Bigas en Barcelona (Lluís, Vero, Vane, Julia, Jordi, Tiago, Mels, Jessica). Estuve con vosotros cuatro meses muy a gusto. Ironías de la vida, iba a hacer el doctorado con ellos estudiando Notch y al final mi tesis en Madrid

termina descubriendo un nuevo papel de la vía de Notch. Gracias a Lluís Espinosa y a Vero por darme un mini proyecto que me tuvo muy entretenida. Vero, para mí siempre serás la mejor profe del mundo, eres una crack. Gracias.

Porque siete años dan para mucho, quiero agradecer a la gente del CNIC presente y pasada toda su ayuda en cada uno de los ámbitos. A todo el departamento (Antes 3N ahora 3N-3S-2N) por su excelente ambiente de trabajo y su constante ayuda. Muchas gracias a Jose Luis por aclarar cualquier duda sobre Notch. Gracias a Bea por la ayuda en cada experimento relacionado con Notch, a Patri por la cantidad de ratones que me ha pasado que ha hecho que este trabajo tenga mucha combinación de alelos, a Juli gracias por ser una buena amiga, gracias a Gaetano, Belén, Vane, a todos los pomperos. A Abel gracias por cuidar de nuestros ratoncitos del grupo Manzanares primero y después por compartir tareas con Patri. Gracias a Nadia y Marina por dejarme probar el mundo de los peces, a Héctor gracias por confiar en mí, por ser más que un compi del curro y por llevarme a festivales. A Miguel, Cris Clavería y Susana por sus sabios consejos y por ser una fuente inagotable de ratones reporteros. A Alberto por ser una referencia. Gracias a Vero por compartir los arrechuchos finales de la tesis combinada con paper. Gracias a Aurora por tu apoyo (sobre todo en los malos momentos) y por la alegría que compartes cada día en el labo. Gracias a Jesús Mª Gomez Salinero por recordarme que el desarrollo temprano es interesante a pesar de no ser tan mainstream en el CNIC como el corazón. Gracias a Teresa, Mary, Beatriz, Ángel, Sandra y Marta por ayudar en toda la logística y hacer la vida en CNIC muy fácil.

Gracias a todos los miembros del animalario (Santi, Frank, Mercedes, Ana, Edu, Isabel, Laura, Bahía, etc.) por cuidar tan bien de los animalitos y por realizar peticiones aunque se os avisara con poco tiempo. Gracias a todos los ratones sacrificados por la causa, sois los héroes de esta tesis. Gracias a los compañeros de Transgénesis por tantas horas juntos, por las enseñanzas de Luis Miguel, el buen humor de David y la banda sonora que nos pone Chema. Estoy agradecidísima a Simon Bartlett por su gran esfuerzo en corregir (casi re-escribir) este libro. Gracias a los chicos de la cafetería (Ángel, Elena, Eva, Lola, Jose, Charo y el antiguo camarero Jose) por alimentarme y ponerme cafés durante tantos años. Gracias al equipo de microscopía (Elvira, Antonio, Antonello) por su ayuda, su paciencia y su buen humor. Gracias a Sergio Callejas por sus enseñanzas iniciáticas en las qPCRs. Gracias también al equipo de informática (Juan Carlos, Alicia y Eduardo) por arreglarme mil veces el PC.

Además, he tenido la suerte de estar acompañada en el departamento por dos de mis mejores amigos de la carrera: Dani y Guille. Gracias Dani por tantos años de compartir momentos (era el mejor de los tiempos, era el peor de los tiempos). Además de hacer la carrera juntos, me alegra mucho haber podido echarme risas en la poyata. Cuando nos mudamos a la 3S, lo primero que eché de menos fueron las conversaciones

enfrentados ordenador con ordenador. Gracias también por presentarme a Cris. Gracias Cris por estar tan cerquita (en la 2N) y tomarte unos cafés en los que da gusto poder hablar de lo primero que se nos venga a la cabeza. Con Guille el contacto volvió según se instalaron los pomperos, muchas gracias por todo tu apoyo incondicional, eres de los de para siempre.

Siguiendo con mis amigos de la uni, muchas gracias a Tefi por tu positividad, por las buenas horas que nos hemos echado siendo compis de prácticas, en flamenco, de viajes, de todo. A Oscar le agradezco mucho su amistad, ojalá que nos sigamos cuidando el uno al otro. A Albertito y Marta les doy las gracias por estar ahí, que lo mismo te ayudan para un roto que para un descosido. A Ire decirle que es un ejemplo de dedicación y a todos los demás (Bricio, Cris, Lau, Nacho, Sali) muchas gracias por compartir momentos conmigo.

Y porque la tesis no solo pasa mientras estás en el laboratorio, agradezco a todos mis amigos su compañía a lo largo de los años. Especialmente, gracias a LAS NIÑAS: gracias Irene por intentar comprender a qué me dedico. Gracias Ana por la espontaneidad que le pones a las cosas. Gracias Jara por el día a día durante el doctorado. A pesar de que me encantaría que vivieras en Madrid, Sergio, te doy las gracias por vivir en países exóticos y que nos permiten ir a visitarte. La energía para ponerle ganas a este doctorado ha venido de ser capaz de desconectar para hacer cualquier otra cosa que no tuviera que ver con la investigación. Gracias a Nuestra Peñita, a los Erasmus, gente del cole, a todos.

Teniendo en cuenta que uno es lo que es y sus circunstancias, agradezco a toda mi amplia familia (abuelos, tíos, primos y madrina Ros) lo que soy. A mis padres les agradezco su cariño y la educación que me han dado. Gracias por mandarme fuera a aprender idiomas y por permitirme dar rienda suelta a mis inquietudes. Todo, la curiosidad, la constancia, la paciencia, el trabajo duro me lo habéis enseñado vosotros. Muchas gracias mamá por tu fortaleza y tu independencia. Aunque mi padre falleció en mis tiernos 16, le agradezco que me transmitiera lo duro que hay que trabajar y lo que hay que sacrificar para conseguir tus metas. A mis hermanos Álvaro, Paloma, Carlos y a Ruth les doy las gracias por su ayuda constante, su alegría y sus ánimos. A mis tías Chelo y Mari Carmen les doy las gracias por ser las de ciencias en la familia e inculcarme el gusto por la investigación. A Clara le doy las gracias por conocerme casi como si fuéramos gemelas y por los buenos momentos.

Gracias Andrés. Muchas gracias por tu pensamiento lateral, gracias por tu humor del absurdo. Gracias por enseñarme algo cada día, gracias por hacer que quiera que se pare el tiempo y nos quedemos así, gracias por hacerme una mejor persona. Te agradezco tu apoyo y sobretodo tu infinita paciencia mientras preparaba artículos,

contestaba a revisores y escribía la tesis. Gracias por entender mi pasión por la investigación, gracias por tu interés en comprender lo que hago y gracias por compartir las alegrías y las penas de esta etapa. Gracias por todas las aventuras que aún nos esperan juntos.

SUMMARY

Developmental biology studies the process by which a single cell, the zygote, will have to divide thousands of times to generate a huge number of specialized cells that comprise the whole adult organism.

The first distinction between cell types that occurs after fertilization is the formation of the trophectoderm and the inner cell mass. These are two essential cell populations in development: the trophectoderm is the precursor of the placenta while the inner cell mass will give raise to the rest of specialized cells of the embryo and the adult organism. Before generating these two lineages, all cells are equivalent and totipotent, i.e., they are capable of forming all structures (embryonic and extraembryonic) in the embryo. The first differentiation event occurs in the early stages of development of the embryo, before implantation in the uterus at the blastocyst stage. Segregation of trophectoderm from the inner cell mass in the blastocyst is key because, firstly, embryonic cells lose their totipotency and, moreover, a new tissue is produced, the trophectoderm, which is characteristic of mammals. The trophectoderm generates the trophoblast, which after differentiating into multiple cell types forms the placenta and amniotic membranes.

Cdx2 is the key gene in the segregation between trophectoderm and the inner cell mass. We identify a regulatory element involved in the regulation of *Cdx2*. This regulatory element directs expression of a marker gene specifically to the trophectoderm. The regulatory element characterized will be crucial to understand the information that *Cdx2* receives and processes to activate and exert its functions. In addition, we find that its activity is stage-specific and does not drive expression of the marker gene in other tissues where also *Cdx2* is expressed. Finally, we find that the Notch signalling pathway is involved in trophectoderm formation together with *Tead4*, a gene recently involved in this first decision. Thus we characterize the joint and parallel regulation of Notch and *Tead4* on the identified *Cdx2* regulatory element, such that it ensures the proper development of the embryo in which compensatory mechanisms promote embryo viability.

RESUMEN

La biología del desarrollo estudia el proceso por el que una única célula, el zigoto, se divide para generar un número muy elevado de células especializadas que forman la diversidad de tejidos que encontramos en el adulto.

La primera distinción entre tipos celulares que ocurre tras la fecundación es la formación del trofectodermo y la masa celular interna. Se trata de dos tejidos esenciales en el desarrollo: el trofectodermo es el tejido precursor de la placenta mientras que la masa celular interna es el tejido que formará el resto de células especializadas del embrión y del organismo adulto. Antes de que el embrión genere los dos tejidos, todas las células que lo forman son equivalentes y totipotentes, es decir, son capaces de formar todas las estructuras (embrionarias y extraembrionarias) del embrión. Este primer evento de diferenciación ocurre antes de la implantación en el útero materno en el estadio de blastocisto. La segregación del trofectodermo y la masa celular interna en el blastocisto es clave ya que, por un lado, las células embrionarias pierden su totipotencialidad y, por otro lado, se genera un tejido, el trofectodermo, que es característico de mamíferos. Del trofectodermo se origina el trofoblasto que, tras diferenciarse a varios tipos celulares, forma la placenta y las membranas amnióticas.

Cdx2 es el gen central en la segregación del trofectodermo frente a la masa celular interna. En este trabajo identificamos un elemento regulador implicado en la función de *Cdx2*. En concreto, este elemento regulador dirige la expresión de un gen marcador de una manera restringida al trofectodermo, siendo esencial para comprender la información que recibe y procesa *Cdx2* para activarse. Además, detectamos que su actividad es estadio-específica ya que no dirige la expresión del gen marcador en otros tejidos donde *Cdx2* también se expresa. Finalmente, identificamos que la vía de señalización de Notch está implicada en la formación del trofectodermo junto con *Tead4*, un gen recientemente implicado en esta primera decisión. De esta manera caracterizamos que la regulación conjunta y en paralelo de Notch y *Tead4* se procesa sobre el elemento regulador de *Cdx2* identificado, de tal forma que se garantiza el correcto desarrollo. Estos resultados demuestran que existen mecanismos de compensación en estos estadios de desarrollo para favorecer la viabilidad del embrión.

INDEX

LIST OF FIGURES

LIST OF ACRONYMS

5TVER7	Tead4VP16 tamoxifen- inducible ES cell line
BAC	Bacterial Artificial Chromosome
bp	Base pairs
Cdx2 $^{-/-}$	*Cdx2* knockout
ChIP	Chromatin immunoprecipitation
DMSO	Dimethyl sulfoxide
dpc	days post coitum
EMFI	Mouse Embryonic Fibroblasts Inactivated
EMFI-CM	EMFI Conditioned Media
ExE	Extraembryonic Ectoderm
H2BmRFP	Histone 2B fused to the monomeric Red fluorescent protein
HUVEC	Human Umbilical Vein Endothelial Cells
ICM	Inner Cell Mass
IRES	Internal Ribosomal Entry site
kb	Kilobases
mRFP	monomeric red fluorescent protein
N1ICD	Notch1 Intracellular Domain
N1OE	Notch1 Overexpressing
***Oct4*DE**	Oct4 distal enhancer
qPCR	quantitative PCR
Rbpj $^{-/-}$	*Rbpj* knockout
Rosa ^{Notch}	*Rosa26* locus encoding N1ICD-IRES-nEGFP
Rosa ^{YFP}	*Rosa26* locus encoding YFP
Tc	Tetracycline
TE	Trophectoderm
Tead4 $^{-/-}$	*Tead4* knockout
Tead4VP16	Constitutively active *Tead4*
TEE	Trophectoderm Enhancer
TS	Trophoblast stem cells
TS_L	*lacZ*-TEE derived TS cells
TS_R	mRFP-TEE derived TS cells
ZHBTc4	*Oct4* tetracycline -repressible ES cell line

"If you don't know where you are going, any road will take you there."
Alice in Wonderland. LEWIS CARROLL.

Developmental biology studies the process by which a single cell, the zygote, divides thousands of times to generate a huge number of specialized cells that comprise the whole adult organism, in humans ~10^{12} - 10^{14} cells. The formation of a new organism during embryogenesis is a spectacular process and represents a masterpiece of temporal and spatial control of gene expression.

Placental mammalian development occurs within the mother's reproductive tract in two separate phases. Initially, the embryo develops from the one-cell to the blastocyst stage freely (Figure 1), to then attach to the uterine wall and implant in order to receive maternal nourishment and growth cues. The first preimplantation phase occurs over 4 to 5 days in the mouse and 6 days in humans. As the preimplantation embryo is independent of external requirements, it is possible to easily recapitulate preimplantation development in culture using chemically defined media without the need of added growth factors. The first specialized tissues that are defined in mammalian embryo development are the extraembryonic cell lineages, the Trophectoderm (TE) and Primitive Endoderm (PE), which will form the placenta and amniotic membranes, essential for the interchange of nutrients and other material with the maternal uterine environment. Extraembryonic structures do not contribute to the fetal tissue; they support embryo growth and provide signals for patterning.

1. Preimplantation development.

Mammalian development is a regulative model in which the cues for differentiation are formed while the embryo develops, and in which cells initially can adopt any possible fate during the first three days. These observations indicate that early mouse development is ruled by a relatively labile developmental program. How the zygote is able to retain this flexibility and at the same time produce its first different cell types is key to understanding how development proceeds.

Upon fertilization, the unicellular mammalian zygote undergoes a series of equal cell divisions (cleavage divisions) in which it increases cell number but not net size (Figure 1). It is surrounded by a glycoprotein envelop, the zona pellucida, essential in the fertilization process. The zona protects the developing zygote and constrains its space during cleavage. The cells –known as blastomeres-, are interchangeable until the eight cell stage (reviewed in Johnson and McConnell, 2004), which allows development to proceed even if blastomeres are destroyed during these early stages (Tarkowski, 1959; Suwinska et al., 2008). The first cleavage division forms the two-cell stage embryo characterized by the switch from maternal to zygotic transcription.

From the two cell stage up to the next successive two division rounds, each blastomere retains full competence to develop into any specialized cell, extraembryonic tissues included; blastomeres appear morphologically identical, sharing a spherical and symmetrical shape. As cleavage proceeds, blastomeres get smaller and the embryo as a whole adjusts its position constantly within the zona.

The first morphological differences arise at compaction: embryos increase their cellular adhesion and their surface smoothens (Figure 1). At the same time, blastomeres polarize thus losing their symmetry. In the subsequent divisions -from the 16- to 32-cell stage, some cells occupy inner positions within the embryo and remain apolar, whereas some others remain in the periphery and polarize along the apico-basal (outside-inside) axis. Tight junction maturation between outer cells from the 32-cell stage enables embryo cavitation -the generation of a large fluid-filled cavity (blastocoel) within the early embryo. Cavitation starts in essentially every apico-basally polarized outer blastomere. Cytoplasmic vesicles within the cells secrete fluid into the intercellular space, thus creating extracellular cavities. These fluid cavities locate beneath outer blastomeres and gradually merge to form bigger cavities as a result of the sealing of epithelial blastomeres. Blastocoel expansion imposes a physical constraint on the embryo, deforms blastomere shape and is resolved in the blastocyst, where the blastocoel locates at one end of the long axis (Fleming and Pickering, 1985; Motosugi et al., 2005).

At the blastocyst stage the two first lineage populations can be clearly distinguished. The trophectoderm (TE) is the epithelial external tissue surrounding the group of cells sitting on top of the cavity termed the Inner Cell Mass (ICM) from where the embryo proper will develop. The TE will give rise to extraembryonic structures, mainly the placenta (Rossant and Tam, 2009). The final event in blastocyst formation is the segregation of a monolayer of primitive endoderm on the blastocoelic surface of the ICM by 4.5 dpc, enclosing the remaining pluripotent epiblast progenitors. By this stage, the embryo is ready to implant in the uterus, and the complex events leading to gastrulation and axis formation begin.

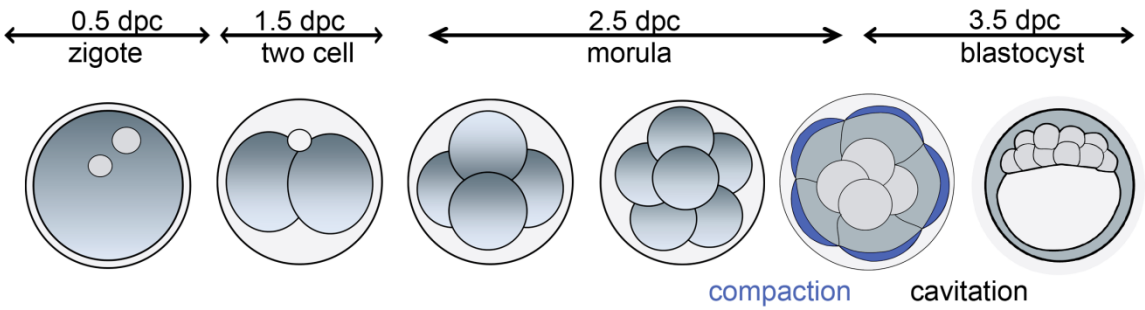


Figure 1. Stages in early preimplantation development.

Overview of mouse preimplantation development during embryonic stages 0.5 – 4.5 dpc (days post coitum). Following fertilization the early mouse embryo undergoes a series of cleavages to generate the lineages necessary for *in utero* survival. Upon compaction, apico-basal polarization of outer cells (depicted in blue) generates the first morphological difference.

The first lineage choice is thus a distinction that occurs progressively as development proceeds, starting in the morula and continuing up to the blastocyst stage. The complex mechanisms involved and the timing of this decision are still under debate.

2. Where and when are the first lineages established?

Classically, three models exist to explain TE and ICM segregation: the pre-patterning model, the inside-outside model and the polarity model (Figure 2).

In the pre-patterning model, TE and ICM are generated upon the asymmetrical distribution of molecular determinants in the oocyte in analogy to other model organisms; in this model, cells would have a fixed fate as soon as they are originated. In this context, pre-patterning arises from a fixed orientation in division planes or the sperm entry position, because so far no molecules have been found that explain lineage segregation in the regulative embryo. Nonetheless, there are some reports in which lineage tracing of cells shows a preferential allocation of lineages (Piotrowska-Nitsche and Zernicka-Goetz, 2005; Tabansky et al., 2013), supporting a bias in the lineage decision.

The inside-outside model suggests that a cell's position leads to different amounts of cell contact and different microenvironments that are interpreted to establish cell fate (Tarkowski, 1959). In this way, the inside environment is the key factor to ensuring lineage segregation, and the outside polarized epithelia confers support for inner cells and is important to ensure the inside niche. In this model, position distinguishes the

two populations, rather than the segregation of determinants, and establishes that the prospective TE protects the future pluripotent cells as soon as two populations arise in the embryo.

The polarity model suggests that the acquisition of cell polarity at the eight-cell stage is critical for lineage segregation in the next cleavage stages (Johnson and Ziomek, 1981). In the model, apolar cells generate inside cells as they divide, whereas cleavage of a polarized cell leads to the inheritance of the apical pole in the daughter cells. In this way, symmetric cleavage of polarized cells generates two daughter cells that inherit the polarizing cues thus contributing to outer cells, the future TE. In contrast, asymmetric division of a polarized cell generates two distinct daughter cells, one polar (future TE cell) and the other apolar (future ICM cell). (Summarized in Figure 2 and reviewed by Rossant, 2009 ; Wennekamp et al., 2013; Yamanaka et al., 2006).

Recent molecular and genetic studies suggest that a combination of the cell polarity and inside-outside models seems more likely to be operating, as would be inhibited by polarity in outside cells and activated by cell adhesion in inside cells (Manzanares and Rodriguez, 2013). Nevertheless, a considerable body of results is not compatible with any of the proposed models. For instance, live-imaging technology has shown cell rearrangements during preimplantation development, whereas none of the proposed models considers any change in cell allocation when outside and inside cells are formed (Watanabe et al., 2014) (Bischoff et al., 2008; McDole et al., 2011; Plusa et al., 2008). Also, careful analysis of protein positioning and quantitative analysis of mRNA expression at single-cell resolution has shown heterogeneous expression of key transcription factors (Dietrich and Hiiragi, 2007; Guo et al., 2010; Tang et al., 2009). The heterogeneous expression of transcription factors indicates that cell fate is plastic even after the first morphological differences arise. This is not compatible with models in which single events lead to a lineage choice. Additionally, the analysis of OCT4 protein kinetics indicated that differences in protein kinetics, instead of the total protein amount, might influence lineage choice (Plachta et al., 2011).

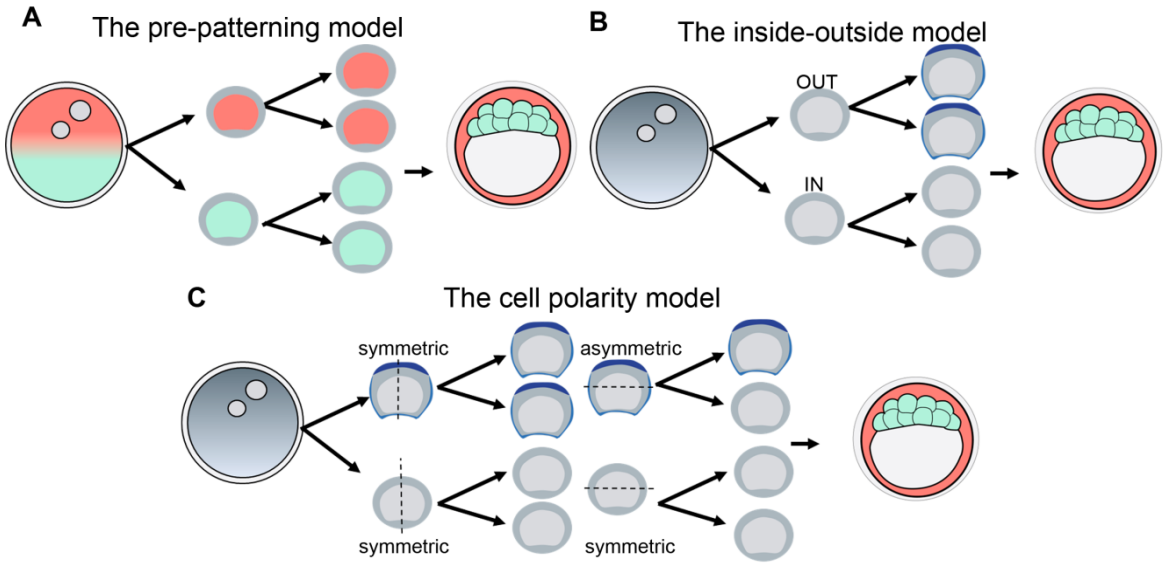


Figure 2. Models of lineage formation.

(A) **The pre-patterning model** proposes that the TE and ICM lineages are segregated on the basis of molecular determinants that are asymmetrically localized in the oocyte (exemplified in red and green gradients). During subsequent cleavages, these determinants are differentially segregated between daughter cells and determine cell fate. Single cells are depicted in dark grey and their nucleus is coloured according to their prospective fate (red, TE; green, ICM).
(B) **The inside–outside model** proposes that, at the 16-cell stage, cell position of the outside or inside in the late morula determines fate in the blastocyst. In the next division round, outer cells polarize (apico-basal domains depicted in blue). Single cells are depicted in dark grey (nucleus, light grey).
(C) **The cell polarity model** proposes that, at the eight-cell stage, cell polarity and cleavage patterns lead to the establishment of outside (blue contour) and inside (rounded grey) cells at the 16-cell stage. Symmetric cleavage of polarized cells (blue outline) produces two outside cells (vertical dashed line), whereas asymmetric cleavage produces one outside and one inside cell (horizontal dashed line in polarized cells).

Recently, a self-organizing model has been put forward that incorporates all available data (Figure 3). It can also explain observations that are not considered in the previous models, such as the stochastic expression of transcription factors or the interdependence between cell polarity and the activity of signal transduction pathways involved in lineage choices (Wennekamp et al., 2013).

In this model, lineage establishment is initiated, corrected and refined by several factors, including transcription factors and sensors of morphology in cells. The model postulates that each of these factors makes a specific contribution to enforce the expression of lineage-specific genes. Importantly, their relevance varies with the context, and some factors can be dispensable or dominant according to the situation.

Another essential feature is the interconnection and interdependence of these factors to regulate each other. Thus, the self-organizing model provides a general framework that covers the physical properties of the cell, differential gene expression patterns and the localization of molecules. In conclusion, it is a working model that includes most of the parameters described so far that account for preimplantation development.

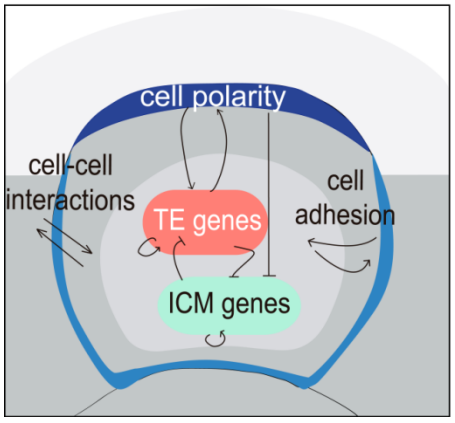


Figure 3. The self-organizing model.

In the regulative embryo, a combination of cell features drives blastomere arrangement to finally generate the TE and ICM segregation. The local microenvironment can influence the activity of gene expression (ICM gene expression is depicted in green and TE-specific gene expression in red). Also, gene expression can modify the cell's physical properties (features depicted in black). Feedback loops stabilize the state of the cell and adjust the final output.

3. Transcriptional circuitry driving cell fate in the preimplantation embryo.

Transcription factors orchestrate the destiny of cells as they are proteins that bind DNA and regulate the expression of downstream genes. Transcription factors promote a specific transcriptional program that is ultimately responsible for the identity of a cell. The specification of the first lineages occurs as a small set of transcription factors becomes restricted to distinct populations. This is a progressive process in which transcription factors appear to be operating in three separate phases: initiation, commitment and maintenance. During initiation, a homogenous population starts to show molecular heterogeneities. Differences can appear either stochastically or through subtle external cues. During initiation, as reported in other systems (Zaret and Carroll, 2011), pioneering factors that can directly bind condensed chromatin might be important for transcription and initiation of lineage determination. In the commitment phase, transcription factors respond in coordination with cellular signals and activate downstream genes, and at the same time form positive feedback loops that strengthen their expression. At these stages, lineage choice is still plastic and a cell can switch fate according to its context. In the final maintenance phase, the transcription factor regulatory network is assembled to maintain its transcriptional status and impede any unwanted lineage switch.

In the last decade, the core set of transcription factors involved in the first lineage choices in the mammalian embryo has been identified. Table A (Supplementary Table) lists the main transcription factors involved in TE and ICM segregation.

In preimplantation development, the TE and ICM appear upon compaction, the commitment phase, and it is only at the blastocyst stage that they are properly established as the two first different lineages of the embryo. The TE and ICM will only get irreversibly segregated when expression of the core transcription factors is sustained by positive feedback loops and cross-regulatory repression at the blastocyst stage. To date, the key transcription factors identified as a part of the TE regulatory network are *Cdx2, Tead4, Gata3* and *Eomes*. The key regulators of ICM establishment are *Oct4, Sox2* and *Nanog*. TE and ICM are thus the two first specialized cell types that are formed during development. Still, both tissues retain fully potency to differentiate into many tissues: the TE will differentiate into trophoblast and subsequently to most of the placental tissues of embryonic origin, and the ICM will give rise to the epiblast and primitive endoderm one day later, and form all the tissues of the fetus and the adult organism and part of the extraembryonic structures.

The homeodomain transcription factor *Cdx2* is the first marker detected (Figure 4A). It is expressed in a salt and pepper manner at the eight cell stage and progressively restricts its expression to outer cells and the TE before *Oct4*, *Nanog*, and *Sox2* expression is restricted to the ICM (Dietrich and Hiiragi, 2007). Remarkably, the *Cdx2* spatial restriction previous to that of *Oct4*, *Nanog* and *Sox2* suggests that it may be required to downregulate these three transcription factors in outer cells. Additionally, *Cdx2* mutant (*Cdx2* $^{-/-}$) embryos die at 3.5 dpc as they fail to form an expanded blastocyst that maintains its epithelial integrity. In these mutants, *Oct4* and *Nanog* are ectopically expressed and not properly restricted in outside cells at the blastocyst stage (Chawengsaksophak et al., 1997; Strumpf et al., 2005).

It has been proposed that differences in the expression levels of *Cdx2* influence cell polarity and account for TE initiation. An asymmetric distribution of *Cdx2* mRNA would result in asymmetric cell division in which daughter cells with low levels of *Cdx2* contribute to ICM, whereas cells with high levels of *Cdx2* form the TE (Jedrusik et al., 2008). However, there is evidence that neither maternal nor zygotic *Cdx2* transcripts direct the initiation of ICM/TE lineage separation (Ralston and Rossant, 2008; Wu et al., 2010) and *Cdx2* mutation does not lead to complete failure to initiate the formation of the TE epithelium of the blastocyst (Ralston and Rossant, 2008; Strumpf et al., 2005); thus, how TE fate is initiated still remains to be resolved.

The ICM marker *Oct4* is ubiquitously expressed early on. It is gradually restricted to the ICM at 4.5 dpc (Figure 4A). *Oct4* is crucial for stabilizing the ICM, such that all cells of *Oct4* deficient embryos convert to TE (Nichols et al., 1998). The pluripotency gene *Nanog* exhibits highly variable levels and gets restricted to the epiblast at 4.5 dpc (Figure 4A). *Nanog* mutants fail to generate the epiblast and their ICM only produces primitive endoderm cells (Mitsui et al., 2003). Importantly, neither factor is predictive of lineage segregation: the ubiquitous expression of *Oct4* discards the factor as a possible early marker of ICM segregation, and the variability in *Nanog* levels is independent of position within the morula. Additionally, there is no initial correlation in the mosaic expression of *Cdx2* with *Oct4* or *Nanog* (Dietrich and Hiiragi, 2007). A recent study suggests that OCT4 kinetics, rather than total transcription levels, is different before compaction among otherwise indistinguishable cells and that this may initiate TE/ICM segregation in the early mouse embryo (Plachta et al., 2011). The other main pluripotency factor *Sox2* is present together with *Oct4* in the early blastomeres (Ralston and Rossant, 2008), but unlike *Oct4*, *Sox2* remains present in TE cells (Adachi et al., 2013). *Sox2* expression has been suggested as the earliest ICM marker by transcriptional profiling (Guo et al., 2010) and *Sox2* mutants produce a failure at implantation due to defects in epiblast and extraembryonic ectoderm (Avilion et al., 2003).

After TE and ICM commitment, these expression patterns are maintained by reciprocal repression of *Oct4* by *Cdx2* in the TE (Niwa et al., 2005), and repression of *Cdx2* by *Oct4*, *Nanog*, and *Sox2* (Boyer et al., 2005; Loh et al., 2006) in the ICM, together with the auto regulation of *Oct4* and *Cdx2* (Chew et al., 2005).

In conclusion, the transcriptional circuitry regulating TE and ICM fates is crucial for the establishment of the two lineages; however, their expression does not completely correlate with the first distinction in the populations. Moreover, these regulatory networks and their cross-regulation do not explain how the outside and inner populations appear.

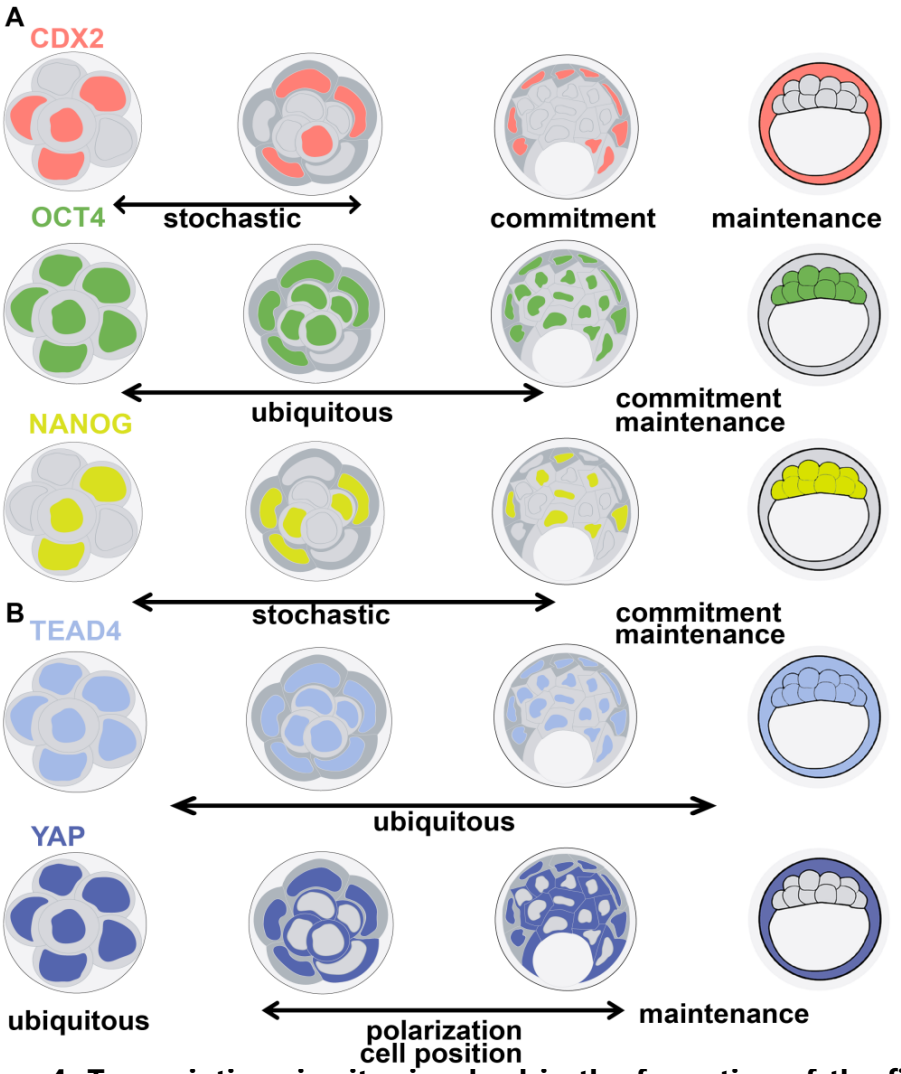


Figure 4. Transcription circuitry involved in the formation of the first lineages in the blastocyst.

(A) The initial expression of *Cdx2*, *Oct4* and *Nanog* transcription factors in the eight cell stage is not restricted to specific cell populations. At the 16-cell stage, outer (dark grey) and inner populations (light grey) arise as a consequence of compaction and polarization. **CDX2** protein is detected beginning at the eight- to 16-cell stage; its initial expression is stochastic. By the morula to early blastocyst stages, *Cdx2* expression is higher in outer, apically polarized cells. Restricted expression in TE cells is established by the blastocyst stage. **OCT4** protein is observed in all blastomeres throughout early cleavage stages. It is not until the blastocyst stage that *Oct4* is gradually downregulated in the TE and restricted to the ICM. **NANOG** is detected from the eight-cell stage. It is expressed in a mosaic fashion in cells until the blastocyst stage. It is then restricted to the ICM; next, *Nanog*-positive will segregate to form the epiblast. (B) The core components of the Hippo pathway YAP and TEAD4 translate the maturation of cellular structures into a transcriptional response. **TEAD4** is ubiquitously expressed in these stages and its TE-specificity depends on its coactivator YAP. **YAP** is ubiquitously expressed in the nuclei of the eight-cell stage embryo. Inner cells promote YAP sequestration in the cytoplasm favouring ICM formation; in contrast, polarization in outer cells allows translocation of YAP to the nucleus to promote TE specification.

4. The first lineage choice and the Hippo pathway.

The first cell fate decision in the mouse embryo takes place during the transition from morula to early blastocyst when the first morphological differences arise. The processes of compaction, polarization, and asymmetric divisions take place at the time that lineage specific transcription factors become spatially restricted and TE and ICM become segregated.

It remains unclear how compaction is initiated. Compaction is associated with the formation of adherens junctions in regions of cell-to-cell contact mediated through E-cadherin, as shown by the precocious compaction when cells are supplied with high exogenous levels (Dietrich and Hiiragi, 2007) and failure of maternal/zygotic E-cadherin mutants to compact (Stephenson et al., 2010). These mutants present an incorrect TE/ICM cell positioning, confirming that epithelial integrity is essential for the lineage choice; however, *Cdx2* expression is maintained, discarding cell adhesion mediated events as early upstream regulators of *Cdx2.* Polarization is concomitant with compaction. Polarization is the asymmetric distribution of molecules and organelles into the apical and basolateral domains. In morulae, the apical domain faces the surface whereas the basolateral domain remains inside. Polarization is key to the divergence of the first lineages, since the apical domain acts as a signal to promote differentiation of the TE. Although there is a clear relationship between polarization and cell adhesion, isolated blastomeres can form distinct apical and basal domains, suggesting that other mechanisms are responsible for polarization (Cockburn and Rossant, 2010). Additionally, in the division from the eight- to the 16- cell stage, a population of inner cell arises. This inside population is smaller than that of apolar cells, indicating that some outer cells are apolar. Thus, polarization and inside-outside positioning are two distinct processes at the eight to 16- cell stage (Anani et al., 2014).

In recent years, the molecular link between polarization and cell fate in the preimplantation embryo has started to be understood. The Hippo signalling pathway has been shown to act as a sensor of polarization in polar versus apolar cells in 16-cell embryos. Hippo pathway components redistribute in different subcellular domains in polar and apolar populations, and thus drive the activation of lineage-specific gene expression programs. The first evidence for the pathway having a role came when it was found that *Tead4* mouse mutants die at the blastocyst stage. Although TEAD4 is expressed in all cells of the blastocyst (Figure 4B), mouse mutants have defects in the specification and development of the TE, do not form a blastocoel and show decreased expression of TE markers *Cdx2* and *Gata3*. In contrast, the ICM seems unaffected in *Tead4* mutant embryos, since expression of ICM-specific *Oct4* and *Nanog* is detected in

Tead4 mutant embryos, and ES cells can be established from them (Nishioka et al., 2008; Ralston et al., 2010; Yagi et al., 2007).

The Tead family of transcription factors requires a transcriptional co-activator to stimulate downstream gene expression. The co-activator of TEAD and downstream effector of the Hippo pathway is the Yes-associated protein (YAP). This signalling pathway, which is conserved from Drosophila to mammals, is a major regulator of cell growth, proliferation, apoptosis, and is critical for cell fate decisions. Phosphorylation of YAP (and TAZ) by the Ser/Thr kinases LATS1/2 regulates subcellular localization of YAP and thereby the activation of its downstream targets (Zhao et al., 2008). In the embryo, phosphorylation of YAP results in cytoplasmic accumulation, leading to inactive TEAD4 (Figure 4B). It has been shown that phosphorylated YAP is increased in apolar cells or cells with smaller apical domains (Anani et al., 2014). Cells expressing high levels of phosphorylated YAP locate in the inside and form the ICM cells. However, in polarized cells YAP is not retained in the cytoplasm; instead it shuttles to the nucleus, binds to TEAD4 and induces *Cdx2* expression. In this way, YAP activation leads to the acquisition of a TE fate (Cockburn et al., 2013; Nishioka et al., 2009).

Importantly, the subcellular distribution of Hippo components constitutes a direct link between polarization and fate. It has been shown that the vertebrate-specific Hippo pathway component Angiomotin (Amot) is localized to the apical domain of outer cells and bound to actin where it is inactive, because it is sequestered by components of the polarity pathway. In contrast, in inner cells, Amot is found throughout the membrane, co-localized with adherens junctions from where it can mediate YAP phosphorylation together with LATS2 (Hirate et al., 2013). Similarly, LATS2 is apically localized in outside cells but evenly distributed throughout the cytoplasm of inside cells, suggesting that this protein is also sequestered, like Amot, by components of the polarity pathway. Additionally, the upstream effector of the Hippo pathway Nf2/Merlin has been shown to be required for Lats1/2-dependent YAP phosphorylation in inner cells (Figure 5). This mechanism is essential for ICM specification, since its absence leads to TE-markers being expressed in inner cells at the expense of ICM-markers irrespective of cell position (Cockburn et al., 2013).

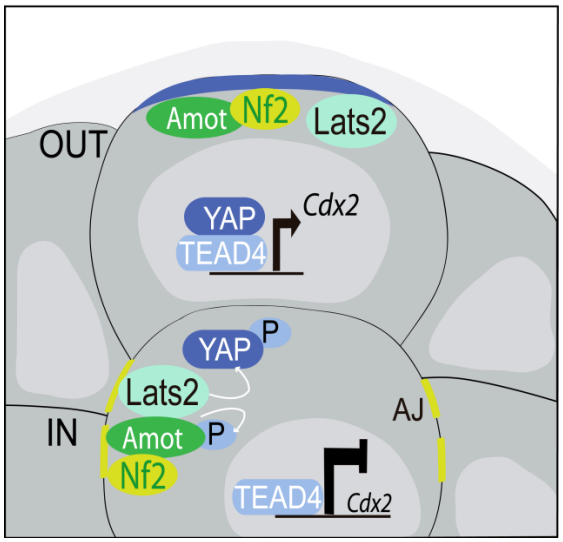


Figure 5. Differential distribution of Hippo pathway components in the first lineage choice.

Outside cells of the morula contain Amot and Lats2 sequestered in the apical domain (blue). This impedes their association and inhibits Amot phosphorylation, thus allowing YAP shuffling in the nucleus to activate, together with TEAD4, *Cdx2* and TE genes. In inside cells, Amot and Lats2 associate in the adherens junctions (AJ, green borders) where Nf2 is also present. Lats2 can phosphorylate Amot, and subsequently this complex phosphorylates YAP, sequestering it in the cytoplasm. In this way, the Hippo pathway is activated, resulting in an ICM fate.

In this way, the Hippo pathway converts morphological cues into a genetic response during the first lineage choice. Also, it has a fundamental role in switching off the expression of TE markers such as *Cdx2* in the inside population. However, the early stochastic expression of lineage specific factors operates when YAP expression is still ubiquitously nuclear (Figure 4), suggesting that there might be other mechanisms that regulate early expression of TE markers.

5. Deconstructing transcriptional circuits in the preimplantation embryo.

The zygotic genome contains all the DNA necessary to specify all the cell types in the adult. Non-coding DNA sequences contain the regulatory information relevant for the expression of a given gene. Regulation of transcription involves at least two primary cis-acting DNA sequence components: promoters and enhancers. Promoters determine where transcription begins; they function upstream and proximal to the initiation site and must contain at least the minimal sequences responsible for the assembly of the basal transcriptional machinery in any cell. In contrast, enhancers play an essential role

in driving cell-type specific gene expression; they consist of clusters of sequence-specific transcription factor binding sites that activate transcription of target genes irrespective of their distal position, ranging from several to hundreds or even thousands of kilobases (Calo and Wysocka, 2013). Whereas promoters tend to be positioned near the transcription start site, the identification of enhancers is more challenging as they can be positioned virtually anywhere in the genome and are not distinguished by any defining sequence feature.

The identification of regulatory sequences allows the use of the enhancer sequence to drive specific expression of proteins to tissues of interest (for example in Cre -loxP systems), permits their deletion to investigate in which molecular mechanisms each sequence is involved, and enables identification of the transcription factors that act directly upstream of the gene of interest. In the preimplantation embryo, the transcriptional circuitry involved in the first lineage choice has been extensively studied but the specific regulation of key transcription factors has not been comprehensively addressed. Although *Oct4*, *Nanog* and *Sox2* enhancers have been described, most of these studies have been performed in ES cells and there is scant evidence regarding their function in early embryos (Catena et al., 2004; Jiang et al., 2008; Levasseur et al., 2008; Tomioka et al., 2002).

Recently, the advent of massive parallel sequencing coupled to chromatin immunoprecipitation (ChIP-Seq) has increased the predictive capability to detect enhancer sequences. The first genome-wide approaches that characterized enhancer elements were the identification of specific histone modifications in non-coding regulatory sequences as indicators of transcriptional activity and ChIP-Seq of the enhancer-associated protein p300 (Heintzman et al., 2007; Rada-Iglesias et al., 2011; Visel et al., 2009). Today, the Encyclopedia of DNA Elements (ENCODE) project provides a comprehensive list of functional elements in the mouse and human genome, including regulatory elements that control cells and tissues in which a gene is active and elements that act at the protein and RNA levels (Stamatoyannopoulos et al., 2012), thus facilitating the identification of putative regulatory sequences. Enhancers are associated with transcription factors and thereby associate with regions of nuleosomal depletion; in this way, enhancers exhibit high sensitivity to DNA nucleases such as DNase I. Also, histones surrounding enhancer regions are enriched for H3K4me1 and H3K27ac (Creyghton et al.; Heintzman et al., 2009; Rada-Iglesias et al., 2011). Nevertheless, while highly valuable, these data provide only indirect evidence of cis-regulatory activity. In addition to this, large-scale epigenomic mapping of non-coding regulatory regions in preimplantation embryos is not feasible with current techniques. The amount of material we can retrieve compared with the quantity of chromatin that is

routinely used in this type of experiment is very limited. Thus, we still need to understand the regulatory mechanisms that underlie the variable expression of core transcription factors. It is also important to define which elements are critical for their functionality. Since *Cdx2* is a key factor for TE specification and is the first factor to be differentially expressed, we focus our study on *Cdx2* regulation. To overcome the technical limitations of working with preimplantation embryos, stem cell populations derived from the blastocyst can serve as a complementary tool for investigation of the regulatory mechanisms and networks involved in the first lineage choice during mammalian development.

6. Embryo derived stem cell populations.

An important property of the mouse blastocyst is that TE and ICM populations can be isolated and cultured *in vitro* under conditions of stemness and also can be directed to differentiate, thereby providing additional tools for studying the gene regulatory and signalling networks operating in development.

In the early 80s, mouse embryonic stem cells (ES) were derived from the epiblast and pluripotency was captured in culture (Evans and Kaufman, 1981). Additionally, trophoblast stem cells (TS cells) have been derived from the trophectoderm lineage (Tanaka et al., 1998), and extraembryonic endoderm (XEN) cells from the primitive endoderm (Kunath et al., 2005). Like other stem cells, these cells can either self-renew or differentiate into lineage-specific cell types. Importantly, all have been shown to contribute mainly to their lineage of origin in chimaeras (Beddington and Robertson, 1989; Kunath et al., 2005; Tanaka et al., 1998). ES cells can be injected into a host blastocyst and this yields chimaeric mice with ES cell contribution to all tissues of the adult organism, including functional colonization of the germline. Moreover, using embryo aggregation with tetraploid hosts, it is also possible to derive live-born mice entirely composed of ES cell derivatives. ES cells thus show an unprecedented developmental capacity after prolonged in vitro culture. In a similar fashion, when TS cells are injected into a host embryo, they colonize and contribute to the TE and the placenta. Results obtained from the analysis of mouse mutants suggest that ES and TS cells recapitulate the lineages from which they originate. ES cells express *Oct4* and cannot be derived from *Oct4* mutant embryos, suggesting they parallel the ICM (Nichols et al., 1998); analogously, no TS cells can be derived from *Cdx2* mutant embryos (Strumpf et al., 2005), implicating *Cdx2* as a main transcription factor in TS regulation and suggesting that TS cells resemble the TE.

ES cells can be derived solely from preimplantation stages since epiblast-derived cells from the post-implantation embryo, although they share some features, do not

contribute to chimaeras (Rossant, 2008; Tesar et al., 2007). Importantly, ES cells closely resemble the epiblast of the blastocyst, when the primitive endoderm has already segregated (Boroviak et al., 2014). In contrast, TS cells can be derived from blastocysts as well as from the extraembryonic ectoderm (ExE) up to 8.5 dpc of the post-implantation embryo (Tanaka et al., 1998). Additionally, the similarity of TS cells to either ExE or TE has not yet been assessed. A defined culture medium for TS cells has been developed only this year (Kubaczka et al., 2014). RNA profiling in TS cells derived at different developmental stages ($TS_{3.5}$ or $TS_{6.5}$) to compare the effect of culture media showed that TS clustered together according to their developmental stage. However, future work is needed to characterize different possible populations within TS cells.

Stem cells allow a detailed molecular study of the interactions between components of the regulatory circuitry. Manipulation in ES cells of key components can help us to better understand the first lineage choice. It has been shown that some modulation over core factors causes stem cell lines to adopt properties of other lineages despite the lineage restriction between ICM/ES and TE/TS. ES cells with downregulated *Oct4* or overexpressing *Cdx2* can grow as TS-like cells in the presence of TS growth medium. Moreover, *Cdx2*-overexpressing cells have been shown to contribute to the placental lineage in chimaeras (Niwa et al., 2000; Niwa et al., 2005). This conversion from ES to TS is seen in ES cells overexpressing *Gata3*, *Ras* or *Tead4* (Lu et al., 2008; Nishioka et al., 2009; Ralston et al., 2010). Manipulation of these factors not only changes fate. It has also been shown that under certain culture conditions, ES cells revert to a totipotent state and contribute *in vivo* to both embryonic and extraembryonic lineages (Macfarlan et al., 2012; Morgani et al., 2013). Additionally, following reprogramming of iPS cells *in vivo*, cells acquire a totipotent-like state displaying both ES and TS cell features: *in vivo* iPS cells can contribute to the TE and the placenta when injected in host embryos (Abad et al., 2013).

Stem cell lines from the first three lineages of the mouse blastocyst have not only provided an invaluable tool to understand the first lineage choice, but ES cells have also emerged as a research model in their own right. Their ability to self-renew and retain the capacity to differentiate into any adult tissue provides the perfect framework within which to understand differentiation into any cell fate, to build organs *in vitro* or to model diseases.

7. The Notch pathway and the control of cell fate.

Notch is one of the key signalling pathways during embryonic development. It functions as a general developmental tool that is used to direct cell fate and, consequently, to build an organism. Notch signalling controls cell fates through local cell interactions and is involved in processes such as stem cell pool maintenance, cell-cell interactions and epithelial-to-mesenchymal transitions (Liu et al., 2010; Perdigoto and Bardin, 2013). The implementation of a particular developmental program modulated by Notch depends on how Notch integrates its activity with other cellular factors.

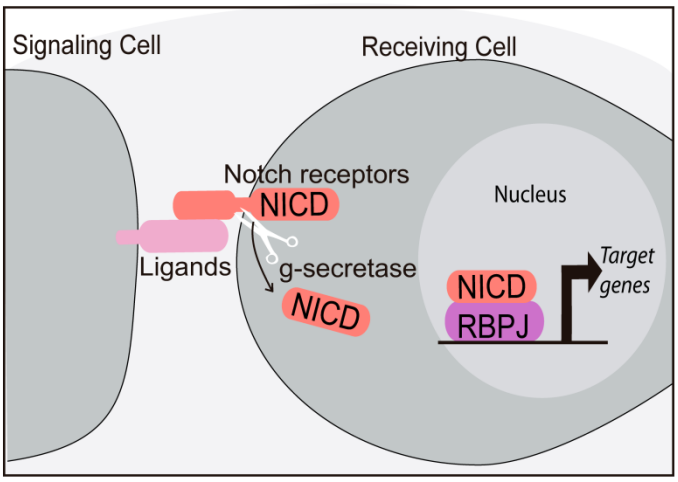


Figure 6. Core components in the Notch signalling pathway.

In the signalling cell, membrane-bound Notch ligands (Dll1, 3, 4 and Jag1, 2, in pink) interact with the Notch receptor (Notch1-4) in the adjacent cell. Ligand-receptor interaction leads to Notch receptor processing by the γ-secretase, which releases the Notch intracellular domain (NICD) in the receiving cell. NICD translocates to the nucleus and forms a transcriptional activation complex after binding to RBPJ. This ternary complex directs the transcription of a set of target genes.

In mammals, the core components of this pathway include four Notch transmembrane receptors (Notch1-4), their canonical membrane-bound ligands of the Delta-like (Dll1, Dll3, Dll4) and the Jagged family (Jag1 and Jag2), and specific co-factors such as Mastermind (Maml 1-3). Notch ligand and receptor interaction induces a cascade of proteolytic events, which involves ADAM metalloproteases and the γ-secretase complex. The γ-secretase minimally consists of four individual proteins: presenilin, nicastrin, APH-1, and PEN-2. It cleaves Notch receptor inside the receiving cell. Notch cleavage leads to the release of the intracellular domain of the Notch receptor (NICD), which then translocates to the nucleus to mediate gene activation. Notch activation results in the displacement of RBPJ repressors (such as N-CoR). The NICD binds the transcription factor RBPJ and after the recruitment of the co-activator MAML leads to transcriptional activation of target genes (Artavanis-Tsakonas and Muskavitch, 2010)

(Figure 6). Common targets of the pathway are the Hes and Hey families of transcription factors and MYC (Ntziachristos et al., 2014).

Notch-ligand interaction normally occurs between two neighbouring cells (Figure 6), where the signalling cell interacts through its ligands with the Notch receiving cell. Nonetheless, ligand-receptor interaction drives multiple outputs (Figure 7). Additionally, *Delta* and *Notch* expression in different systems indicates that individual cells can express receptor and ligand at the same time and that they can interact even if they are in the same cell, to induce cell-autonomous effects on Notch dependent signals. Furthermore, depending on the expression level, their action can be either agonistic or antagonistic (Artavanis-Tsakonas and Muskavitch, 2010).

Loss of Notch function in vertebrate development has been shown to disrupt aspects of neurogenesis, heart development, somite formation, angiogenesis, and lymphoid development (Andersson et al., 2011). Analysis of mutants has shown that a number of them die from 9.5 dpc -11.5 dpc and interestingly, the analysis of mouse strains with targeted mutations in the *Notch2*, *Notch1/4*, *Hey1/2*, *Dll4* and *Rbpj* genes has demonstrated that these genes are indispensable for proper development of the placenta, a TE derived tissue (Gasperowicz and Otto, 2008). Although mutants for Notch receptors and ligands do not show any preimplantation phenotype, there are two reports of Notch pathway components presenting preimplantation defects. *Brainiac* is a glycosyltransferase that shares some features with *Fringe* (Goode and Perrimon, 1997). It modifies Notch receptors on the extracellular domain by adding glucosides and this modulates Notch-Dll interactions. *Brainiac* null embryos die between 3.5 dpc and 4.5 dpc. Initially the null embryos appear morphologically indistinguishable from wild-type littermates and hatch properly, but they die after 48h in culture (Vollrath et al., 2001). *Notchless* (*Nle*) is another modulator of the Notch signalling pathway activity, and disruption of *Nle* results in embryonic lethality shortly after implantation due to an increase in apoptosis of the ICM (Cormier et al., 2006). Importantly, *Nle* is expressed specifically in the ICM and *Nle* mutant embryos show normal CDX2 expression. However, the specific role of *Brainiac* and *Nle* in the Notch pathway has not been described in detail yet.

Therefore, the control of cell fate during the first lineage choice relies on different regulatory mechanisms that act in combination to ensure embryo development. To better understand the TE/ICM specification process, it is crucial to link the expression of transcription factors with inputs from different signalling pathways.

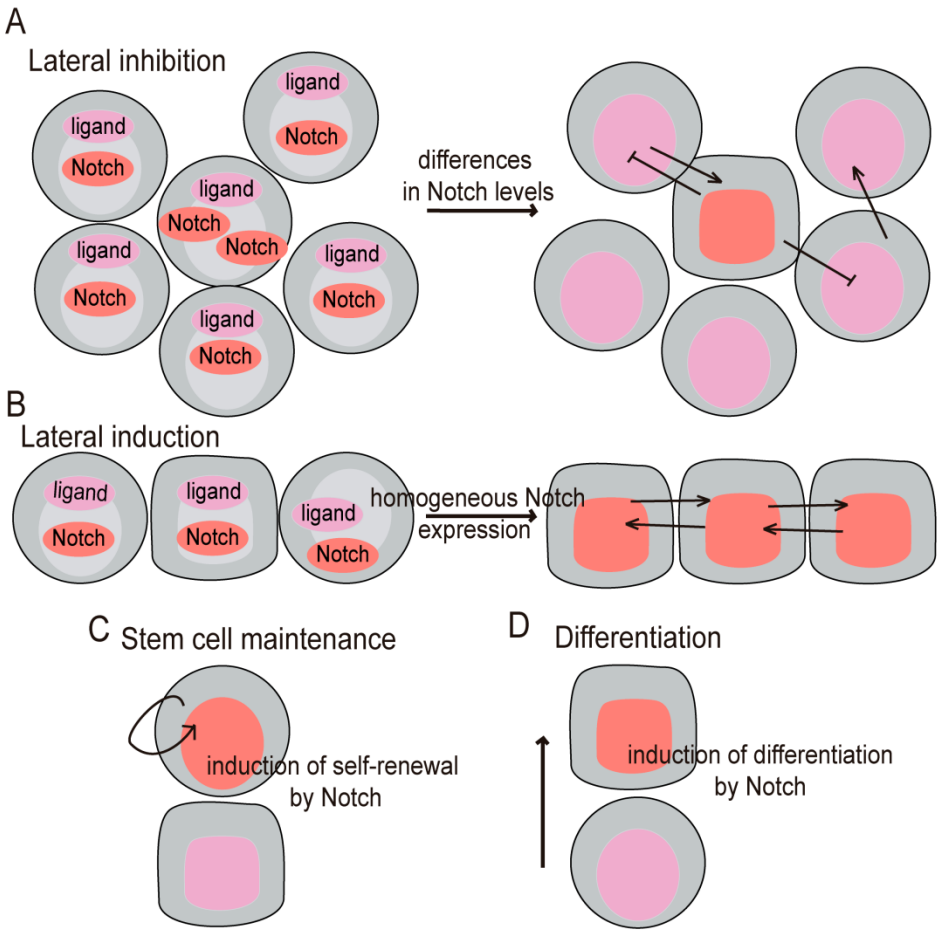


Figure 7. Notch mechanisms in development and differentiation.

The Notch pathway has different roles in binary cell fate decisions. (A) **Lateral inhibition**. (Left) Within a group of cells of the same type which expresses equal amounts of ligand and receptors, a single cell varies its ligand levels. (Right) The ligand then activates the Notch signalling cascade in the neighbouring cell, thereby promoting two different fates. (B) **Lateral induction**. The Notch pathway expression of ligand and receptors in all cells favours a homogenous expression of the signalling pathway and target gene expression. (C) A cell type induces stem cell proliferation without differentiation. (D) A Notch expressing cell initiates differentiation.

OBJECTIVES

The detection of regulatory elements that drive the expression of core transcription factors can shed light on the regulatory mechanisms that confer stochastic expression and lineage commitment of these factors during the first lineage choice. A *Cdx2* enhancer previously characterized in the lab directs reporter expression to the TE, opening the possibility to dissect specific *Cdx2* regulation. With this aim in mind, we defined the following objectives:

» Identify the minimal sequence that directs TE-specific enhancer activity.
» Study the behaviour of the identified element *in vivo.*
» Find the regulatory inputs that direct enhancer activity identified upstream of *Cdx2.*
» Characterize these inputs in relation to endogenous *Cdx2* expression and TE specification.

"If you want to inspire confidence, give plenty of statistics – it does not matter that they should be accurate, or even intelligible, so long as there is enough of them."

Three Months in a Curatorship. LEWIS CARROLL.

1. Construct generation for microinjection

Cdx2 genomic regions were amplified by PCR using BAC RP245I065 as template. This BAC covers the whole intergenic region containing mouse *Cdx2* and was obtained from the BACPAC Resources Center (http://bacpac.chori.org/). The restriction enzyme strategy or primers used for PCR, together with the lengths of corresponding amplified fragments (Figure 9), were as follows:

Table 1. Primers used for fragment amplification and cloning

Fragment	Primer F	Primer R	Product size (bp)
#1	GTTTGGAGAGAAGAAAGGAG	GGGTGAAGTGAAGAAGATCAG	5428
#2	Short piece of Fragment #1 digested with ApaI		1856
#3	Long piece of Fragment #1 digested with ApaI		3572
#4	TGCTAACACAAGCTCCCTCA	AAAGCAGGGAAGAGCACTTTA	766
#5	GACTGGCTGCCTTACCAGAG	TCTTCCAAAGACGCTGGAGT	1487
#6	CACACGGATGAATTGTCTGG	AACAGGGACAGGTGAGATGG	1329
#7	GCCTAGGATGCTGACTGAGG	CCCAAGTTGGAAAGGTTTGA	809
#8	ATCTCACCTGTCCCTGTTGG	CCCTGGGTGAAGTGAAGAAG	1107

As a positive control, we used the *Pou5f1* distal enhancer element (*Oct4*DE, Figure 10) (Pernaute et al., 2010; Yeom et al., 1996). Each fragment was subcloned in pGEM-T Easy Vector and then excised and cloned into a modified pBluescript vector (Yee and Rigby, 1993) containing either a *lacZ* reporter gene or H2BmRFP reporter gene under the control of the human beta-globin minimal promoter and including an SV40 polyadenylation signal. Constructs were linearized and plasmid sequences removed before microinjection.

2. Transient transgenic analysis

For the generation of transient transgenic embryos, F1 (C57Bl/6xCBA) females were superovulated to obtain fertilized oocytes as described (Nagy, 2003). Each construct

was microinjected at 3-6 ng/µl into the pronucleus of fertilized oocytes at 0.5 dpc. Microinjected oocytes were cultured in microdrops of M16 medium (Sigma) covered with mineral oil (Sigma) at 37°C, 5% CO_2 until the blastocyst stage.

A minimum of 50 blastocysts were used to calculate the percentage of *lacZ* or H2BmRFP positive embryos per construct. When using the empty vector containing only the minimal promoter and the different reporter as a negative control, we routinely obtain low-level punctuate *lacZ* expression or weak H2BmRFP expression in approximately 10% of blastocysts. For *lacZ* staining, blastocysts were fixed in buffer containing 1% formaldehyde, 0.2% glutaraldehyde, 2 mM $MgCl_2$, 5 mM ethyleneglycoltetraacetic acid (EGTA), and 0.02% Igepal for 5 min at room temperature. After 2x 5min washes in phosphate buffered saline (PBS), blastocysts were transferred to X-Gal staining solution for 24 hours at room temperature in the dark. For H2BmRFP detection, embryos were fixed in 4% paraformaldehyde for 10 min at room temperature and either analysed for endogenous fluorescence or processed for immunostaining. To visualize nuclei, embryos were incubated in DAPI at 1µg/ml (Vector Laboratories).

3. Generation of TEE mouse lines

Three independent transgenic mouse lines where obtained for fragment #3 constructs linked to each *lacZ* or H2BmRFP. All lines reproduced the TE-restricted expression pattern in early pre-implantation stages. Genotyping was performed by PCR using SV40 and 5'-rev comp primers for the *lacZ*-TEE lines, and 5'Cdx2 and mRFP primers for the mRFP-TEE lines. Primer sequences are indicated in the table below:

Table 2. Primers used for genotyping

Primer	Sequence	Product size (bp)
SV40	TCACTGCATTCTAGTTGTGG	150
5'-rev comp	CTGATCTTCTTCACTTCACCCAG	
5'Cdx2	GCAGGTGGCTGATCTTCTTC	900
mRFP	GAGCCGTACTGGAACTGAGG	

PCR conditions were 95ºC for 5min, 35 cycles of 94 ºC for 1min, 60ºC for 1min and 72ºC for 1 min, followed by 72 ºC for 10 min. Embryos were collected from crosses of TEE-*lacZ* or TEE-mRFP males with outbred superovulated ICR females.

4. Mouse breeding

The different mouse lines used are listed in the table below:

Table 3. Mouse strains.

Alleles	Reference
Cdx2 $^{-/-}$	(Strumpf et al., 2005)
Tead4 $^{-/-}$	(Nishioka et al., 2008)
Rbpj $^{-/-}$	(Oka et al., 1995)
Rosa^{Notch}	(Murtaugh et al., 2003)
Rosa^{YFP}	(Srinivas et al., 2001)
Sox2 Cre	(Hayashi et al., 2002)

Adults were genotyped by PCR of tail-tip DNA using primers and conditions previously described for each line. For preimplantation embryos, genotyping was performed directly on individually isolated embryos, after observation in culture and X-gal or antibody staining.

Animal procedures were approved by the CNIC Animal Experimentation.

5. Embryo collection and culture

For characterization of *lacZ*-TEE and mRFP-TEE mouse lines, 0.5 DPC embryos were collected from swollen ampulas, treated with hyaluronidase (Sigma) to remove cumulus cells and cultured until the blastocyst stage at 37.5°C in 5% CO_2 in air, in M16 medium (Sigma) covered with mineral oil (Sigma). For experiments using *Tead4*, *Cdx2* and *Rbpj* mutant mouse strains, 2.5 dpc embryos were collected by flushing the oviduct through the infundibulum, and were cultured up to the blastocyst stage. For experiments with *Rosa*^{YFP} and *Rosa*^{Notch} strains, 3.5 dpc embryos were recovered by flushing uteri with M2 media (Sigma).

6. Mutagenesis

Mutated versions of fragment #6 (TEE^{RBPJ}mut, TEE^{TEAD}mut and TEE$^{RBPJ/TEAD}$mut) were generated by site-directed mutagenesis (Mutagenex Inc.). RBPJ binding sites were located according to the consensus motif and changes that abolish binding were introduced as described (Tun et al., 1994). TEAD binding sites were based on the MCAT consensus motif (5′-CATTCCA/T-3′) (Anbanandam et al., 2006). The following changes were introduced:

Table 4. Substitutions introduced in RBPJ and TEAD binding sites.

Lower case indicates the altered residues.

RBPJ binding sites	mm9 assembly position	Mut
TTCCCACCG	chr5:148124480-148124488	TTC**gg**ACCG
TGTGGGAAA	chr5:148124694-148124702	TGT**cc**GAAA
TTCCCAGGT	chr5:148124759-148124767	TTC**gg**AGGT
TTCCCACTT	chr5:148125418-148125426	TT**agg**ACTT

TEAD binding sites	mm9 assembly position	Mut
AATTCCTA	chr5:148124456-148124465	A**cggaa**TA
ATTCCAG	chr5:148125498-148125504	**cggaa**AG

7. Pharmacological inhibitor treatments

Two-cell embryos were cultured in drops of M16 medium containing the corresponding pharmacological inhibitor or control (DMSO) until the blastocyst stage. The following inhibitors and concentrations were used: 10 µM of the γ-secretase inhibitor RO4929097 (S1575, Selleckchem) (Munch et al., 2013) and 2.5 µM of the TEAD/YAP inhibitor Verteporfin (Sigma) (Liu-Chittenden et al., 2012).

8. Cell culture

Mouse embryonic fibroblasts were grown in standard media conditions and mitomycin inactivated as described (Himeno et al., 2008) to obtain mitomycin-treated primary embryonic fibroblast conditioned media (EMFI-CM) for TS cell culture.

The B1-TS cell line was established and maintained as described (Tanaka et al., 1998). TS cell lines were derived from *lacZ*-TEE or mRFP-TEE mouse strains (see below, section 13).

ZHBTc4 ES cells were cultured on gelatin-coated dishes with 1000 U/µl LIF (ESGRO-LIF; Millipore) (Niwa et al., 2005). Repression of *Oct4* (official name Pou5f1) in ZHBtc4 ES cells was induced by addition of tetracycline (Sigma) at 1µg/ml in EMFI-CM media (Tanaka et al., 1998) in the presence of $2.5x10^{-3}$µg/ml FGF4 (R&D Systems) and 1µg/ml heparin (Sigma).

5TVER7-ES cells were cultured on gelatin-coated dishes with 1000U/µl LIF as previously described (Nishioka et al., 2009). Tead4VP16ER was induced by incubation of cells for 48 h with 4-hydroxytamoxifen (Sigma) at 0.1µg/ml in EMFI-CM media in the presence of $2.5x10^{-3}$µg/ml FGF4 and 1µg/ml heparin.

9. Immunohistochemistry

Immunohistochemistry was performed as previously described (Dietrich and Hiiragi, 2007). Cells were cultured on gelatin-coated glass coverslips and fixed in PBS with 4% paraformaldehyde. Samples were permeabilised and blocked at room temperature before incubation. The following antibodies and dilutions were used: monoclonal mouse anti-CDX2 (MU392-UC, BioGenex) 1:200, rabbit polyclonal living colours DsRed (632496 Clontech) 1:500, monoclonal mouse anti-OCT4 (sc-5279, Santa Cruz) 1:200, rabbit polyclonal living colours GFP (632460 Clontech) 1:200, mouse anti-TEAD4 (ab58310 Abcam) 1:100, and rabbit anti-Cleaved NOTCH1 (Val1744) (2421, Cell Signalling Technology) 1:100. Nuclei were visualized by incubating embryos in DAPI at 1µg/ml and Rhodamin-phalloidin (Molecular Probes) 1:300, to detect F-actin in the cell membrane.

10. Quantitative-PCR

RNA was isolated from ZHBTc4 and 5TVER7 cells with the RNeasy Mini Kit (Qiagen) and then reverse transcribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). For γ-secretase inhibitor experiments, RNA from pools of 25 embryos was isolated using the Arcturus PicoPure RNA Isolation Kit (Applied Biosystems) and reverse transcribed using the Quantitect Kit (Qiagen). cDNA was used for quantitative-PCR (qPCR) with Power SYBR® Green (Applied Biosystems) in a 7900HT Fast Real-Time PCR System (Applied Biosystems). Expression of each gene was normalized to the expression of the housekeeping genes *Actin* and *Ywhaz* (Vandesompele et al., 2002). The following primers were used:

Table 5. qPCR primers

Gene	Forward Primer	Reverse Primer
Actin	CAGAAGGAGATTACTGCTCTGGCT	TACTCCTGCTTGCTGATCCACAT
Cdx2	TCAACCTCGCCACAACCTTCCC	TGGCTCAGCCTGGGATTGCT
Eomes	TTCACCTTCTCAGAGACACAGTTCAT	GAGTTAACCTGTCATTTTCTGAAGCC
Esrrb	GGACACACTGCTTTGAAGCA	ACAGATGTCTCTCATCTGGC
Fgfr2	GAGGAATACTTGGATCTACC	CTGGTGCTGTCCTGTTTGGG
Gata3	GGGTTCGGATGTAAGTCGAG	CCACAGTGGGGTAGAGGTTG
Hand1	TGCACAAGCAGGTGACCCCG	CCCTTTAATCCTCTTCTCGCCG
Nanog	CTTACAAGGGTCTGCTACTGAGATGC	TGCTTCCTGGCAAGGACCTT
Oct4	ATCAGCTTGGGCTAGAGAAGGATG	AAAGGTGTCCCTGTAGCCTCATAC
Stra13	GGTGAGCAGACTACTCCATTT	GTGCCCCACATATTCCCCAC
Ywhaz	CGTTGTAGGAGCCCGTAGGTCAT	TCTGGTTGCGAAGCATTGGG

11. Cell transfections

Transfection of TS cells was performed with Lipofectamine 2000 (Invitrogen) and 2.5μg DNA in non-adherent dishes following the manufacturer's instructions. ZHBTc4 cells were transfected in 6 well plates with Lipofectamine 2000 and 0.8μg DNA (eGFP, empty vector, *Oct4*DE or fragments I, II or III). At 6h after transfection, media was changed and tetracycline added when indicated. 5TVER7 cells were transfected similarly; cells were transfected in 12 well plates with 0.4μg DNA (eGFP, empty vector, *Oct4*DE or fragments I, II or III).

Forty-eight hours after transfection, cultures were measured by flow cytometry (LSRFortessa Flow Cytometer) or random fields per well were photographed (Zeiss) and fluorescent cells counted (ImageJ). Regulatory activity is expressed as the proportion of red (RFP^+) over total control green (GFP^+) cells transfected.

12. Histone ChIP

Histone ChIP on cross-linked chromatin from TS cells followed by quantitative PCR (qPCR) was carried out as previously described (Azuara, 2006). Briefly, 50μg of precleared cross-linked chromatin was incubated overnight with 5μg of the antibody followed by 4h incubation with Sepharose A beads. Elutes were digested with Proteinase K (0.5mg/ml) and Rnase A (0.1mg/ml) and DNA was phenol-choloroform extracted and recovered by standard precipitation with ethanol. The following antibodies were used: anti-H3K4me1 (ab8895; Abcam), anti-H3K4me3 (ab8580; Abcam), anti-H3K27me3 (07-449; Millipore), anti-H3K27ac (ab4729, Abcam), anti IgG (m7023, Sigma). Table 6 lists primers used for ChIP. ChIP data are represented as fractional differences (difference between the enrichment of the histone mark minus the IgG basal enrichment divided by the IgG value for the specific primer sets).

Table 6. Primers used for ChIP experiments

Fragment	Primer	Forward	Reverse
Actin		CCCCAACACACCTAGCAAAT	ACTGCCCCATTCAATGTCTC
Nanog		CCCAGGTTTCCCAATGTGAAG	AAAGAGTCAGACCTTGCTGCCA
Sox1		ACAAGAGGAGGCAGCGAACC	TCGCAGGTGGAAAGTTTCTCC
I	1	TCCTGGGAGATGTCGGGTGCC	CAGCTCTGGGTTCAGGCCGC
	2	CTGCCTGCCTCCTCCCTCCA	GGGCCCCCTCTGCCTACACT
II	3	GGAACGCGTCTCACCTGCCC	CCAGCAGCCCCGCGCTATTT
	4	CCCGCGCCTGCTTTGGAAGT	GCCAGCGCGTGGTGCTCTAA
	5	CACTCCGGCAGCATTGCCCA	TGGCACAGCCAGGCCACATG
promoter	promoter	CTCGACGTCTCCAGCCATTGGT	CCAGCGGCCTTACGTGATTAAC
intergenic	6	GGTGGCTTGTAGAGCTGCGGT	GGGGGCGCAACCTGGAGGTA
	7	CCCAATCTCATCAAGCTGCCTTTG	TGGAACCCTACAGGAGAACCTTTG
III	8	AGGTTCTCCACTTGCTGCGGC	GCATCCAAGCACGGAAGTGAACA
	9	GCCCCATTCACAGTCTCCAGTTACA	TGCTTCGTTCCTCACCTTCCCCA
	10	TCCCACCGAACGCAAAACAGCT	ACCGCTCCTGTGGCCCAGAA
	11	CTCGGAGGGATAAGCTCTCAAGTGT	TGCCTCTCTGGAACAACCCGGT
	12	AGGGCTGGCATCCTCGAGCA	TGTGCCACAGCTTTTGGGCT

13. TS cell derivation

TS cells were derived essentially as described. (Himeno et al., 2008; Tanaka et al., 1998). 3.5 dpc blastocysts were obtained from wt superovulated females. For the derivation TS_L cells, females were mated with *lacZ*-TEE heterozygote males. TS_R cells were derived from homozygote mRFP-TEE x wt crosses.

Blastocysts were plated in 4 well plates of $5x10^4$ inactive mouse embryonic fibroblasts (EMFIs) with EMFI-CM media supplemented with $37.5x10^{-3}$μg/ml FGF4 and 1.5μg/ml heparin. Plated blastocysts hatch, adhere to the well and form outgrowths in the dish. Depending on the individual blastocyst, outgrowths were tripsinized from days 4 to 6 and subsequently plated until TS cell colonies were clearly visible. After 20-21 days, MEFs were removed and TS-derived cells were cultured in EMFI-TS media on plastic dishes.

Twelve independent clones (TS_L) were derived from the *lacZ*-TEE line and 7 TS clones (TS_R) from the TEE-mRFP mouse line. Four TS_L and two TS_R clones were selected for further characterization.

13.1 TS_L cells genotyping

Genomic DNA was extracted by digestion in a 4well plate with 500ul of embryo lysis buffer with Proteinase K (0.1mg/ml) overnight at 55ºC. DNA was phenol-choloroform extracted and recovered by standard precipitation with ethanol. In order to detect the presence of the transgene in TS_L clones, PCR was performed with the SV40 and Cdx2-rev comp. primers (Table 2). Three (TS_L #1- #3) out of 4 clones were PCR amplified for the transgene.

13.2 TS_R and TS_L cell expression profile

TS_L and TS_R cells formed epithelial colonies as expected and were capable to self-renew and differentiate appropriately upon removal of FGF4 and heparin (Figure 8A). The TS_L and TS_R cells were further characterized by qPCR expression profiling (view section 10). TS_L and TS_R derived lines expressed *Cdx2*, *Eomes*, *Esrrb* and *Fgfr2*, factors involved in the TS pluripotency network and showed differentiation markers (*Stra13*, *Hand1*) to the same extent as TS. From the analysis, we selected TS_L #2 and TS_R #1 for further study.

13.3 TS derived cell cycle profiling

TS cells have the potency to differentiate into various cell types, the most readily detectable being the differentiation to trophoblast giant cells. Giant cells are formed by endoreduplication of DNA content. A cell cycle profiling assay was performed to

evaluate the capacity of the selected clones to differentiate in normal TS culture conditions and upon FGF4 removal.

Cell cycle was profiled by flow cytometry using a LSRFortessa Flow Cytometer (Tanaka et al., 1998). Briefly, tripsinized cells were fixed in 70% ethanol, treated with RNAse at a final concentration of 100µg/ml, and resuspended in 0.003% propidium iodide solution. DNA content profiles of TS cells (2n/G_1 phase), replicating TS cells (up to 4n/S-G_2-M) and endoreduplicating cells (4n or >4n) are shown in Figure 8E-G. For the different clones analysed (2 TS_L clones and 1 TS_R clone), ~ 20-30% of cells differentiated in normal culture conditions, whereas after 6 days of differentiation, the percentage increased up to 42%. Some giant cells might not have been included in the flow citometry due to their size thus not being able to detect all the differentiated cells.

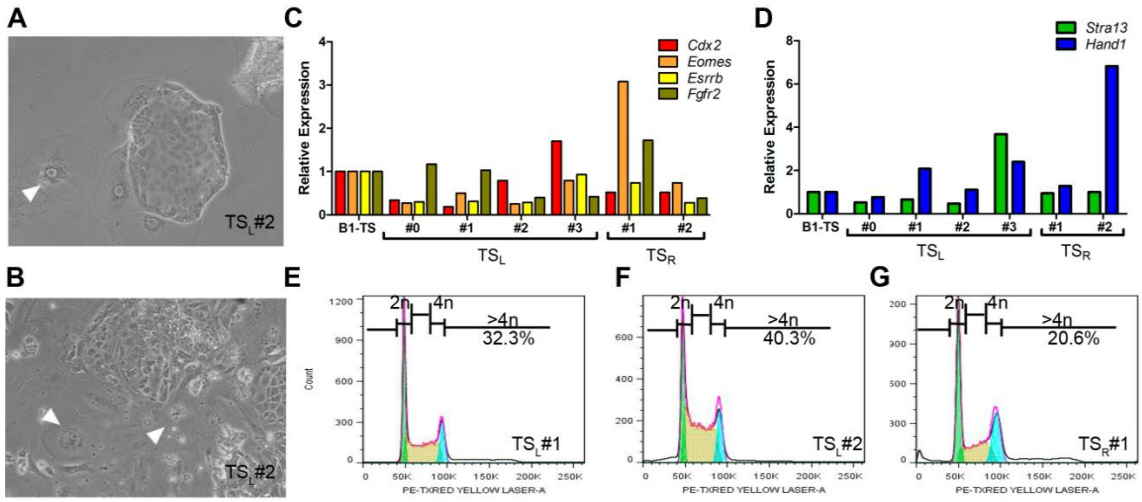


Figure 8. Characterization of TS_L and TS_R cells.

(A) Representative TS_L epithelial colony in early passages. (B) TS_L clone upon FGF4 removal subjected to differentiation. Arrowheads point to differentiated cells. Relative expression of (C) TS pluripotency and (D) differentiation markers in TS_L (#0-#3) and TS_R (#1- #2) clones. TS_L #0 was genotyped as *lacZ* negative. (E-G) DNA content in different TS_L and TS_R cells. DNA content is represented by the propidium iodide intensity.

14. TS cell infection

Lentivirus encoding the GFP reporter were generated by transient transfection of 293T cells. For transduction, 10^6 TS cells were seeded in suspension followed by adding 0.25 mL vector suspension in RPMI media (1×10^8IU/mL, MOI = 100). Eight hours post infection, vector suspension was removed and transduced cells were seeded in a 24 well dish with fresh EMFI-CM TS media and incubated at 37°C with 5% CO_2. Transduction efficiencies were evaluated 48h post-infection.

15. TS cell injection

Eight to 10 TS_R cells were microinjected into eight-cell stage embryos using standard techniques (Springer, 2011). Morulae were incubated overnight in M16 medium (Sigma) microdrops under mineral oil. Endogenous fluorescence was assessed the next day. Immunohistochemistry for CDX2 and mRFP1 was performed when indicated as in section 9.

16. Imaging and quantification

Images of transfected cells were acquired with a Zeiss Axiover 200M inverted microscope. Confocal images of cells were acquired with a Leica SpE microscope (20x or 40x objective). Confocal images of microinjected or antibody-stained embryos were acquired with a Leica SP5 confocal microscope. Images were acquired with a 63x objective and 2x zoom every 2.5µm. Images of *lacZ*-stained embryos were obtained with a Leica DMIRE2 inverted microscope. Images were prepared for figures using Adobe Photoshop CS5.

For quantification, unmodified images were analysed as previously described using IMARIS imaging software version 7.6.3 (Bitplane AG) (Dietrich and Hiiragi, 2007) (Dietrich and Hiiragi, 2007) with some modifications. Nuclei were segmented in 3D reconstructions based on DAPI staining with an 8µm isosurface. After computer segmentation, segments were inspected visually and corrected when necessary. The number of nuclei staining positive for TEE, CDX2 or GFP was evaluated visually (8-cell stage embryos) or by segmentation (IMARIS software; 2.5 dpc to 4.5 dpc). CDX2 protein level was estimated from unmodified mean fluorescence intensities within segmented nuclei. Mean DAPI fluorescence intensity was used to minimize error caused by staining and confocal imaging variability. CDX2 intensity values for each blastomere were normalized to the mean DAPI fluorescence intensity for each nucleus, and these ratios were normalized to the average mean DAPI intensity per whole embryo.

17. Statistics

Statistical analyses were performed with GraphPad Prism 5. Data are presented as means ± s.e.m or ± s.d. as indicated in the figures. Differences were considered statistically significant at $p < 0.05$. p values were calculated by t-test for comparisons of two groups, and ANOVA with Bonferroni post-test for multiple pair wise comparisons. For the data presented in Figure 24, Figure 27 and Figure 33, a Chi square test was performed.

B1-TS, 5TVER7 ES cell culture assays and RO treatments of wild type embryos were performed in collaboration with Sergio Menchero. TS_L and TS_R cell lines derivation and characterization, TS_R injections, mouse genotyping and ZHBTc4 ES cells transfections were performed in collaboration with Inma Ors and the CNIC Transgenesis Unit. Histone ChIP was carried out in Veronique Azuaras' lab at the Institute of Reproductive & Developmental Biology, Imperial College London, UK in collaboration with Anne Hellness. Vectors were produced, concentrated, and titrated by the CNIC Viral Vectors facility. TEE-mRFP line generation and *Cdx2* $^{-/-}$ breeding was performed in collaboration with Andres Nieto from the lab of Janet Rossant at the Developmental & Stem Cell Biology Department in The Hospital for Sick Children, Toronto, Canada. CBF reporter was generated and characterized in collaboration with Panagiotis Xenopoulus at Kat Hadjantonakis lab at the Sloan-Kettering Institute, New York, USA.

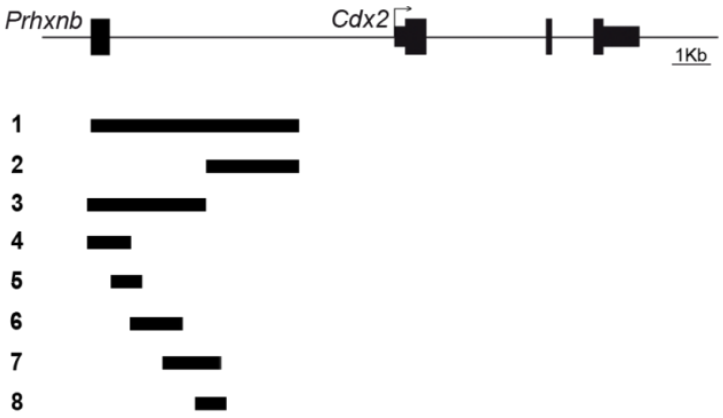
"Sometimes I've believed as many as six impossible things before breakfast."
Through the Looking-Glass. LEWIS CARROLL

1. Identification of a trophectoderm-specific *Cdx2* enhancer.

To better understand how the first lineage decision takes place in the mammalian embryo, we studied the regulation of *Cdx2*, one of the first transcription factors involved in this process. *Cdx2* is the first gene expressed in the TE and is the first transcription factor shown to be allocated to the outermost population of the compact morula (Dietrich and Hiiragi, 2007; Strumpf et al., 2005).

Screening of stage-specific functional enhancers *in vivo* by transient transgenic analysis of non-coding DNA sequences has long been used as a tool to evaluate the regulatory potential of a given sequence during development. Transient transgenesis in cultured blastocysts also allows relatively fast interrogation of enhancer capacity because enhancer potential is scored three days after DNA microinjection of the zygote, and embryos do not need to be transferred to a foster female to continue their development. Using this approach, assessment of transcriptional activity is as straightforward as in cell culture assays, but with the advantage that scoring of the cis-regulatory selected sequence is stage- and tissue-specific. Moreover, the systematic and extensive screening of cis-regulatory sequences with this method provides a convenient way to narrow down the size of a functional DNA sequence without the need to generate stable lines of all the regulatory sequences analysed.

Previous work in the lab, using transient transgenic analysis in mouse preimplantation embryos, identified a cis-regulatory element located 5' of *Cdx2* that drives reporter expression in the trophectoderm (TE) of the blastocyst (fragment #1). Unlike other elements tested in the vicinity of *Cdx2*, fragment #1 showed TE-specific activity in transient transgenic assays at the blastocyst stage, thus reproducing the endogenous *Cdx2* expression pattern. Using transient transgenesis, we were able to track reporter expression in the ICM driven by the *Oct4* distal enhancer (*Oct4*DE) (Yeom et al., 1996) (Figure 10A) and TE-specific activity of fragment #1 (Figure 9), confirming that this fragment reproduces the endogenous *Cdx2* expression pattern (Figure 10 B,C).



fragment	TE specificity	% reporter expressing
#1	+	47.00%
#2	-	2.27%
#3	+	25.42%
#4	-	6.98%
#5	-	11.11%
#6	+	21.75%
#7	-	4.00%
#8	-	9.26%

Figure 9. A cis-regulatory element upstream of *Cdx2* drives restricted expression in the trophectoderm.

Diagram of the *Cdx2* locus showing the fragments tested in transient transgenic embryos, the specific activity of the fragments in the TE, and the percentage of embryos for each construct showing reporter activity.

We then characterized the minimal DNA fragment able to reproduce the TE-restricted expression shown by fragment #1. We first divided the DNA region into two smaller regions (fragments #2 and #3) by restriction enzyme digestion with ApaI. Fragment #2 is a 1.8Kb DNA sequence located in the more proximal region to *Cdx2* gene, and fragment #3 is the remaining 3.6Kb DNA sequence located 4331 bps upstream of the *Cdx2* transcription start site. Fragment #3 directed TE-specific reporter expression, while fragment #2 was negative for cis-regulatory activity (Figure 9).

In order to find the minimal DNA region that contains Trophectoderm-specific enhancer activity (TEE), we designed smaller overlapping DNA fragments of around 800-1500bp (fragments #4 to #8) that covered fragment #3 in full (Figure 9). Fragment #6 (1329bp) preserved TE restricted expression of the reporter gene (Figure 10C). Fragments that drove TE-specific expression pattern were tested with both H2BmRFP (data not shown) and *LacZ* reporters (Figure 10C).

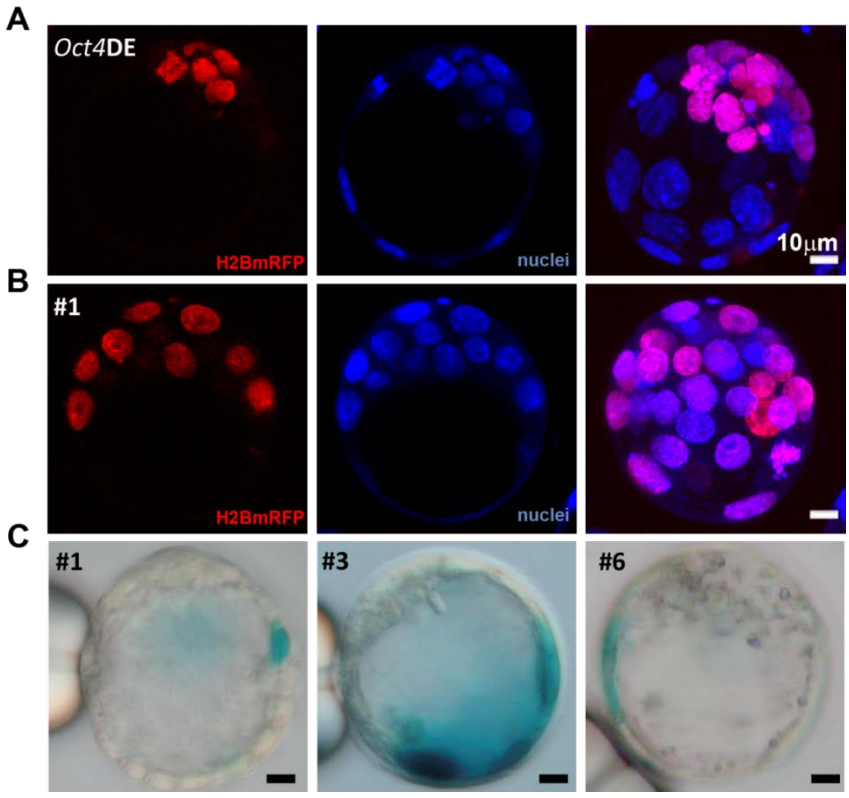


Figure 10. Fragments tested by transient transgenesis in the blastocyst.

H2BmRFP expression driven in the ICM by (A) *Oct4*DE and in the TE by (B) fragment #1 of *Cdx2*. Nuclei were stained with DAPI. Maximal projection is shown on the right. Scale bars, 10μm. (C) Representative transient transgenic embryos from fragments #1, #3 and #6 showing TE-specific *lacZ* reporter activity.

2. Characterization of the *Cdx2* TEE mouse lines.

Having identified a *Cdx2* TE-specific enhancer sequence, we generated stable mouse lines to screen for stage-specific TEE-driven activity in detail, avoiding the mosaicism inherent to transient transgenesis approaches. Mosaicism occurs because the microinjected DNA often integrates not in the zygote but after the embryo has already started to divide. Some cells will consequently have integrated the transgene whereas others will not, giving rise to genetically different cell populations. For this reason, mosaic embryos show heterogeneous reporter expression and do not reflect the possible regulatory capacity of the interrogated enhancer in every cell. In stable mouse lines, the transgene is stably integrated in the genome, and thus reporter expression will mimic endogenous regulatory capacity response in all cells.

We generated several transgenic mouse lines using fragment #3 linked either to *lacZ* or to a H2BmRFP reporter gene and studied its enhancer capacity. Four stable transgenic

lines were analysed, three *lacZ* lines (*lacZ*-TEE) and one H2BmRFP mouse line (mRFP-TEE). All the lines reproduced the TE-restricted expression pattern of the reporter gene at the blastocyst stage. We selected the H2BmRFP line *and* one of the *lacZ* lines for further analysis.

Analysis of the mRFP-TEE line showed that the reporter is only occasionally active in a few cells at the eight-cell stage (Figure 11A), and upon compaction starts to be present in the outer cells of the morula (Figure 11B), attaining strong activity at the 16 cell stage (Figure 11C). At the blastocyst stage, TEE-driven reporter activity is localized throughout the TE and excluded from the ICM (Figure 11D). The *Cdx2 lacZ-TEE* reporter lines revealed an identical behaviour (Figure 11E). We used the nuclear localization of the H2BmRFP reporter protein to track blastomeres up to the 3.5 dpc blastocyst. Embryos with fewer than eight cells rarely showed any sign of reporter activity; upon compaction, at the 8-cell stage, around 40% of the blastomeres were TEE+, indicative of the first onset of TEE activation, and at the 16-cell stage, 88% of blastomeres were TEE+. These results point to an early activation of the TEE through signals concomitant with compaction. The number of blastomeres expressing mRFP increased during successive divisions (Figure 12D), with TEE+ cells accounting for ~60% of total blastomeres. At the blastocyst stage, reporter activity in the H2BmRFP and *lacZ-TEE* lines was found throughout the TE, although occasionally we detected positive cells in the ICM and cells with lower levels of expression in the TE (Figure 12C). Immunohistochemical analysis of the association between reporter activity and endogenous CDX2 showed concordant expression in >75% of blastomeres (CDX2+/TEE+, CDX-/TEE-; Figure 12A-C, E), confirming that TEE activity closely matches the first onset of CDX2 expression at preimplantation stages (Dietrich and Hiiragi, 2007; Strumpf et al., 2005).

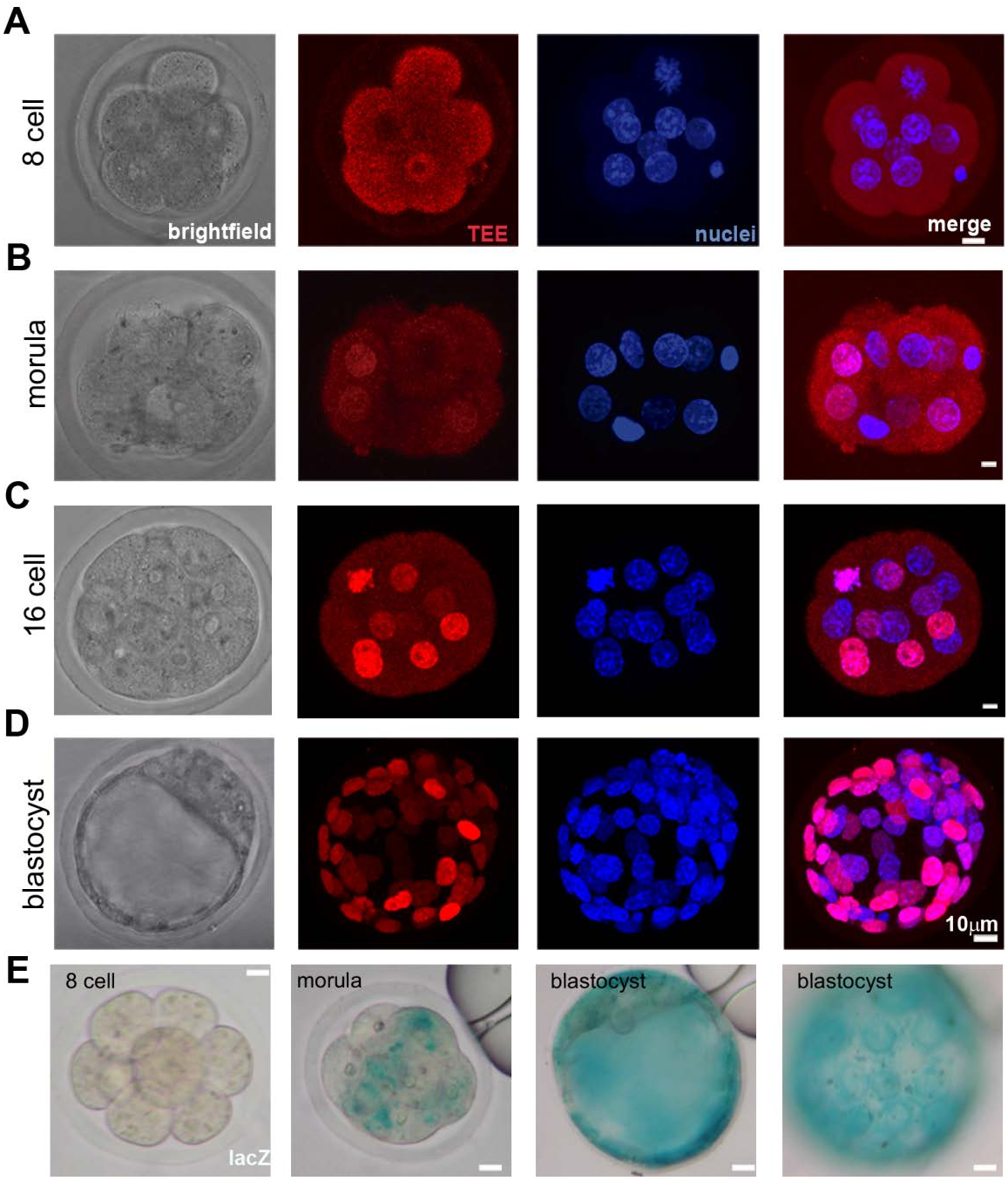


Figure 11. mRFP- and *lacZ*-TEE lines drive restricted expression in the trophectoderm.

(A-D) H2BmRFP reporter expression driven by the TEE at (A) the non-compacted 8-cell stage, (B) compacted 8-cell morula stage, (C) 16-cell stage, and (D) blastocyst stage. H2BmRFP was detected by immunohistochemistry and nuclei were stained with DAPI. (E) Temporal dynamics of a stable *Cdx2 lacZ*-TEE line at the 8-cell, morula, and blastocyst stages. Two different focal planes of the blastocyst are shown to evidence TE staining. Scale bars, 10 µm.

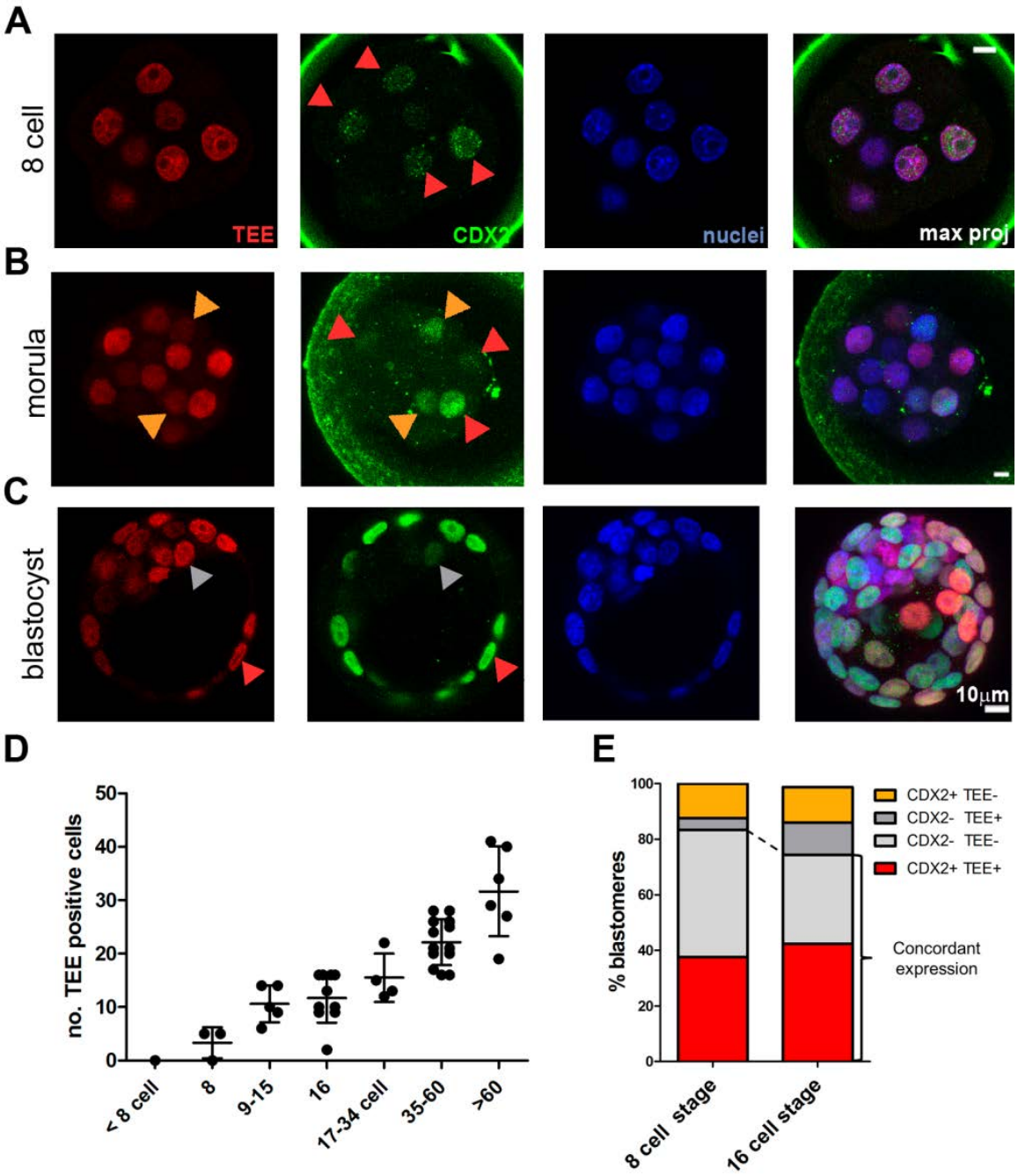


Figure 12. *Cdx2* TEE reproduces the early onset of CDX2 expression.

(A-C) Immunohistochemistry of TEE and CDX2 at the (A) eight-cell, (B) morula and the (C) blastocyst stage. Some TEE and below-normal CDX2 expression is still found in the ICM of the blastocyst (grey arrowhead). Arrowheads indicate CDX2+TEE+ (red) and CDX2+ TEE- (orange), as quantified in (E). (D) Quantification of the number of TEE-positive cells per embryo, staged by total cell number. Each dot represents an individual embryo, and bars and whiskers indicate means and standard deviations. (E) Correlation of TEE activity with endogenous CDX2 protein expression at the 8-cell and 16-cell stages.

Next, we analysed the expression pattern of the TEE mouse lines at early post-implantation stages. After implantation of the blastocyst into the uterus, the TE gives rise to all trophoblast derivatives of the conceptus: parietal trophoblast giant cells that line the implantation site, extraembryonic ectoderm (ExE), the ectoplacental cone, and later the various trophoblast cell types of the mature chorioallantoic placenta (Roper and Hemberger, 2009). *Cdx2* is expressed in trophoblast derivatives (Beck et al., 1995) (Figure 13) and is a key transcription factor for the trophoblast lineage. *Cdx2* $^{-/-}$ embryos die before implantation because the TE fails to maintain trophoblast identity and epithelial integrity, resulting in the collapse of the blastocyst (Strumpf et al., 2005).

We expected to detect TEE mediated expression in the TE derivatives where *Cdx2* is normally expressed such as the ExE of the early embryo. We dissected embryos at 6.5 dpc and 7.5 dpc from both reporter lines and surprisingly did not detect any sign of reporter activity (Figure 13). This does not coincide with the *Cdx2* endogenous expression pattern, since *Cdx2* at these stages remains expressed in the TE-derived extraembryonic ectoderm. The lack of TEE reporter activity suggests that *Cdx2* expression at this stage requires other regulatory elements.

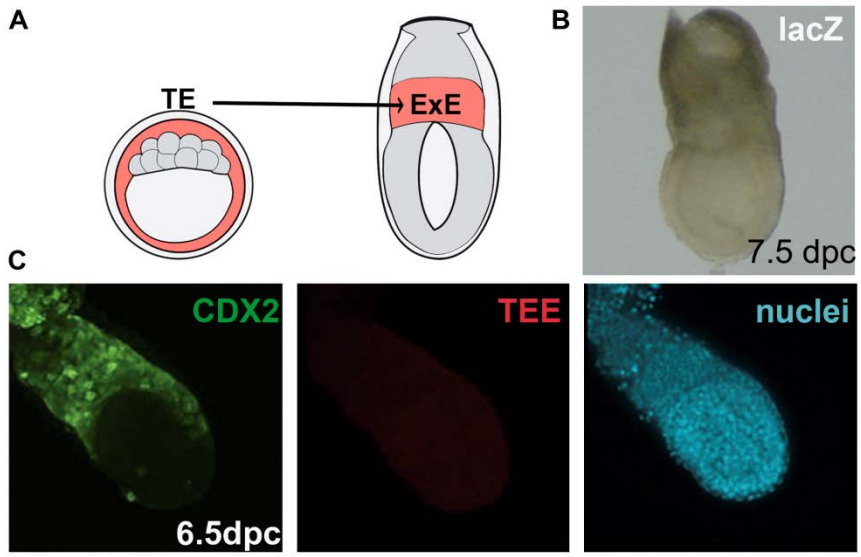


Figure 13. *Cdx2* TEE drives restricted expression in the trophectoderm but is inactive in the post-implantation embryo.

(A) *Cdx2* expression at blastocyst and post-implantation stages (red). (B) *lacZ*-TEE line reporter expression at 7.5 dpc. (C) CDX2 (green) and TEE (red) expression in the 6.5dpc post implantation embryo in the mRFP-TEE line. CDX2 and TEE are detected by immunohistochemistry with anti-CDX2 and anti-mRFP antibodies. CDX2 depicts extraembryonic ectoderm (ExE) expression, while TEE activity is absent at this stage.

3. TEE behaviour in trophoblast stem cells as a model of TEE function upon implantation.

Three different stem cell subpopulations can be derived from blastocysts (Yamanaka et al., 2006). These populations have the ability to self-renew indefinitely *in vitro* and to differentiate into the various lineages of the tissue from which they originate. Moreover, they are a valuable tool for the study of aspects of stem cell biology, such as self-renewal and differentiation. The derivation of trophoblast stem (TS) cells from the TE mirrors the formation of the trophoblast from the TE after implantation. *Cdx2* is the core component in TS cell maintenance as shown by the fact that TS cells cannot be derived from *Cdx2* mutant blastocysts, showing that CDX2 is indispensable for TS cell self-renewal (Strumpf et al., 2005), and overexpression of Cdx2 in ES forces its conversion to TS cells (Niwa et al., 2005). The equivalence of TE and TS cells is confirmed by the contribution of TS cells, when injected back into a blastocyst, to the TE and its derivatives (Tanaka et al., 1998).

We exploited the equivalence of TE and TS cells to study *Cdx2* regulation *in vitro*. To investigate the differential activity of the TEE between the TE and the ExE, we transfected wild-type TS cells with putative *Cdx2* cis-regulatory sequences, analysed chromatin status in the *Cdx2* locus and derived TS cells from the *lacZ* and mRFP mouse lines.

3.1 TS cell transfections of regulatory elements.

To understand the possible mechanisms driving TEE activity, we first examined if the TEE drives reporter expression in wild-type TS cells. For this, we generated three RFP reporter constructs from fragments in the *Cdx2* locus (Figure 14A): fragment I is a DNA sequence downstream of *Cdx2* that has been characterized as a cryptic "shadow" enhancer in adult mouse liver (Watts et al., 2011); fragment II is an enhancer in the first intron of *Cdx2*, previously identified in the lab, that possesses enhancer activity in both ICM and TE; and fragment III is the TEE. Upon transfection into TS cells, fragment III showed no significant activity above that of empty vector (mock), contrasting with its competency to drive reporter expression in blastocyst. Fragment I, despite its "shadow" enhancer capacity, did not yield detectable RFP reporter expression, suggesting that it is a tissue specific enhancer and that the regulatory landscape of mouse liver and TS cells may be completely different. Interestingly, fragment II was able to drive RFP reporter expression in TS cells. This activity is in line with transient transgenic experiments in which this fragment directed reporter expression in both ICM and TE lineages of the blastocyst. These results thus show that the TEE is unable to induce the expression of a reporter gene in TS cells (Figure 14B).

3.2 Analysis of the *Cdx2* epigenetic landscape.

Analysis of epigenetic regulation of chromatin structure has emerged as a powerful way to identify cis-regulatory elements. Certain enhancer-associated chromatin features can be effectively used to annotate them. Chromatin signature at transcriptionally active promoters is remarkably similar across all cell types, with trimethylation of histone H3 at lysine 4 (H3K4me3) being the most common modification (Heintzman et al., 2007). In contrast, active enhancers differ between cells and are marked by histone monomethylation (H3K4me1), histone acetylation (H3K27ac), binding of coactivator proteins (p300) or DNaseI hypersensitivity (DHS). Interestingly, H3K27ac distinguishes active enhancers from 'poised' H3K27ac-negative enhancers (Heintzman et al., 2007; Rada-Iglesias et al., 2011).

We profiled histone modifications in the *Cdx2* genomic region to identify the chromatin signature of the TEE in wild-type TS cells and to search for presence of other putative regulatory regions. Chromatin immunoprecipitation of histone marks followed by qPCR of selected regions in the *Cdx2* locus was used to interrogate histone modifications across the regulatory sequences analysed by transgenesis and transfection assays. We selected primers covering the TEE (amplicons 8-12), an upstream intergenic region (amplicons 6-7), the *Cdx2* promoter (promoter), the first intron (amplicons 3-5), and the region covering the 3' enhancer (amplicon 2) (Figure 14A).

For H3K4me3 ChIP, we included the housekeeping gene *Actin* as a positive control, used as an active promoter in TS cells. There are no reports showing enrichment for the H3K4me1 mark in TS-specific enhancers that we could use as positive control. As a negative control, we chose the *Nanog* promoter, which is not expressed in TS cells.

ChIP for H3K4me3 and H3K4me1 showed that, as expected, the *Cdx2* promoter region was enriched for H3K4me3 and depleted for H3K4me1. In agreement with our embryo transgenic and TS transfection assays for fragment II, we detected significant enrichment for H3K4me3 and H3K4me1 in the amplicon 4. In contrast, H3K4me1 was depleted over the TEE (amplicons 8-12), consistent with it not being active in TS cells. PCR products covering fragment I (2) were not enriched for H3K4me1 or H3K4me3 (Figure 14C). These results indicate that the TEE does not possess active histone marks. In contrast, the first intron of *Cdx2* shows enrichment for H3K4me1 and me3, in agreement with the results obtained by transient transgenesis with the same region. Although TS cells are derived from the blastocyst TE, these results suggest that *Cdx2* may be differentially regulated in the TE versus trophoblast-derived cells.

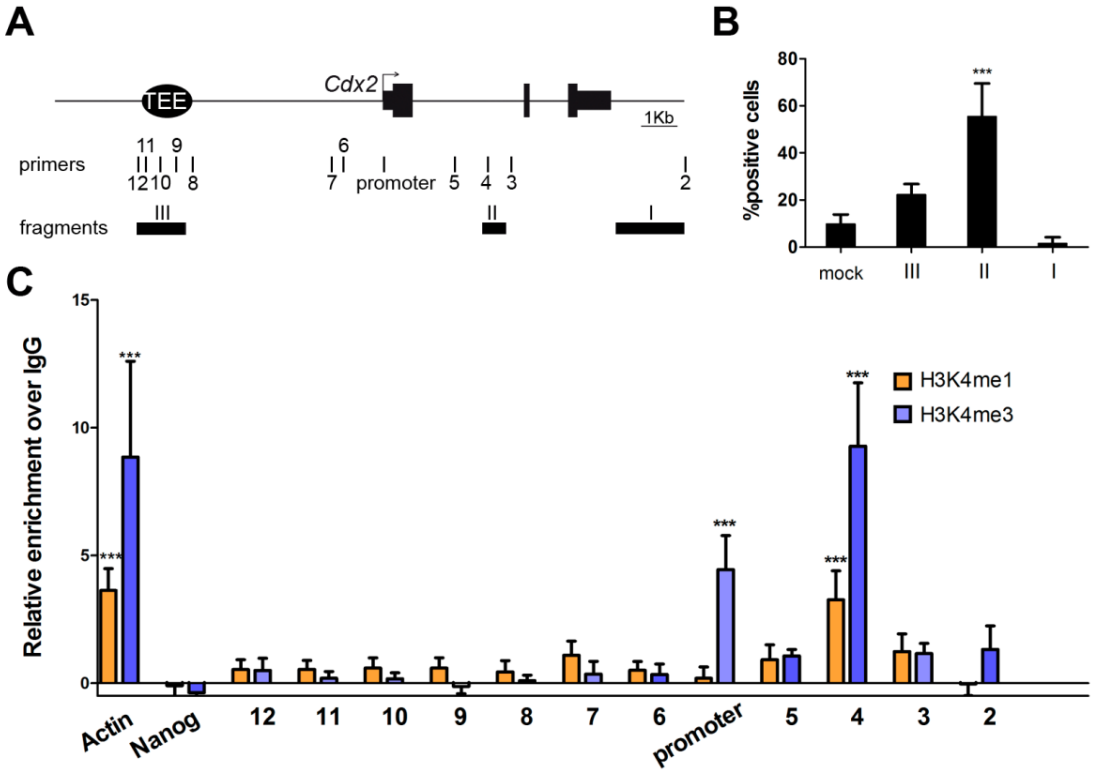


Figure 14. *Cdx2* regulatory landscape analysis in TS cells.

(A) *Cdx2* locus, showing the position of the TEE, the promoter and exons of *Cdx2*. Arabic numerals show relative positions of amplicons used for ChIP experiments (2-12); Roman numerals show the relative size and position of *Cdx2* fragments used for transfection assays (I, II and III). (B) TS transfection of different *Cdx2* regulatory fragments. Fragments were tested for regulatory capacity and compared to the expression to the empty vector (mock). (C) Relative enrichment over IgG of H3K4me1 (red) and H3K4me3 (blue) in a panel of selected *Cdx2* regions and controls in the chromatin of TS cells. Data are means ± s.e.m. n=3. ***p<0.001 by Bonferroni post-test.

3.3 TS derivation from mRFP- and *lacZ*-TEE mouse lines.

Results from experiments with TS cells are reminiscent of the TEE behaviour in the post-implantation embryo, where the mRFP- and *lacZ*-TEE mouse lines showed no sign of reporter activity in the ExE of 6.5 dpc embryos, in striking contrast to the TE-restricted expression of the TEE in blastocysts. However, TS cells are considered to be equivalent to the TE, because when TS cells are injected back into a blastocyst, they contribute to the TE and its derivatives.

We decided to derive TS cells from the mRFP- and *lacZ*- TEE mouse lines (TS_R and TS_L respectively) to analyse reporter expression driven by the TEE when the transgene is

stably integrated in the genome. In this scenario, we expected the TEE to be inactive in TS cells, as already shown in TS transfections of fragment III. These cells would be a valuable *in vitro* tool to better characterize the differences between the *in vivo* regulation of *Cdx2* expression in the trophectoderm and in the post-implantation trophoblast.

From the 19 clones we derived (twelve TS_L and seven TS_R independent clones), we verified that four TS_L and two TS_R cell clones expressed appropriate levels of *Cdx2* and of the other trophectoderm pluripotency markers, as shown by RNA expression analysis and immunohistochemistry (see section 13 of Materials and Methods). Next, we scored for reporter expression driven by the TEE. TS_L cells from the *lacZ-TEE* mouse line did not express β-galactosidase driven by the enhancer (data not shown), and neither did the TS_R cells derived from the mRFP-TEE line (Figure 15), consistent with TS transfections and chromatin status results. Hence, TS cells do not reproduce the TEE activity in the blastocyst. These results also indicate that additional inputs on the TEE may be absent or not sufficient to activate reporter expression in TS cells. They also suggest that different transcription factors may be responsible for *Cdx2* regulation in the TE versus TS cells and post-implantation stages.

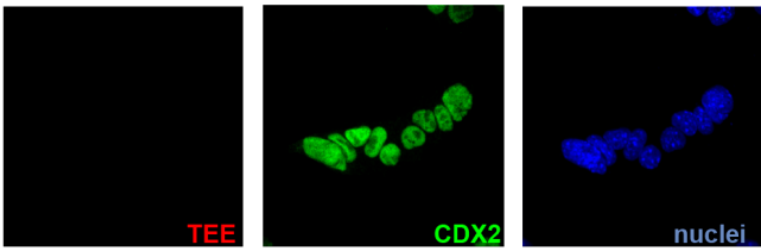


Figure 15. TS_R cells do not drive reporter activity.

TS_R cells do not show mRFP driven expression. Cells were stained for mRFP (red), CDX2 (green) and DAPI (blue).

The definitive demonstration that TS cells retain the same features as TE cells, other than their *in vitro* gene expression profile and their differentiation potential, is their ability to contribute to the TE when injected back into the preimplantation embryo. We therefore tested if TS_R cells could indeed contribute to the TE and, if so, if they re-expressed mRFP directed by the TEE when located in the embryo.

To study if TS cells where able to activate the TEE and contribute to the TE, TS_R cells were injected in the morula stage and their contribution to the TE of the blastocyst was assessed 24 hours later. After injection, injected cells mix with cells of the host, resulting

in the development of a mosaic embryo. To test if the TS_R cells were able to contribute to the TE, we stained injected embryos for mRFP and CDX2 at the blastocyst stage. TS_R cells were able to contribute to the TE, expressed CDX2, and, more importantly, were able to re-express the reporter, indicating that the enhancer is still functional (Figure 16A).

To better follow the behaviour of TS_R cells in blastocysts and to make sure that H2BmRFP re-activation was occurring, we infected TS_R cells with a lentiviral GFP construct, allowing tracking them after injection into the morula. To obtain a near-pure population of GFP-expressing cells we re-sorted TS_R cells immediately before injection into morulae (Figure 16 B, C).

Analysis of endogenous GFP and H2BmRFP reporter expression showed that TS_R cells re-expressed the nuclear red fluorescent reporter. However, this was seen in only a proportion of TS_R injected cells: some GFP+ TS_R cells were able to contribute to the blastocyst without activating H2BmRFP reporter expression (Figure 16D). These cells appeared not to be allocated to the TE, and this might be the reason why they could not turn on the enhancer. It will be important to check which TE/ICM markers these TS_R cells express in the blastocyst.

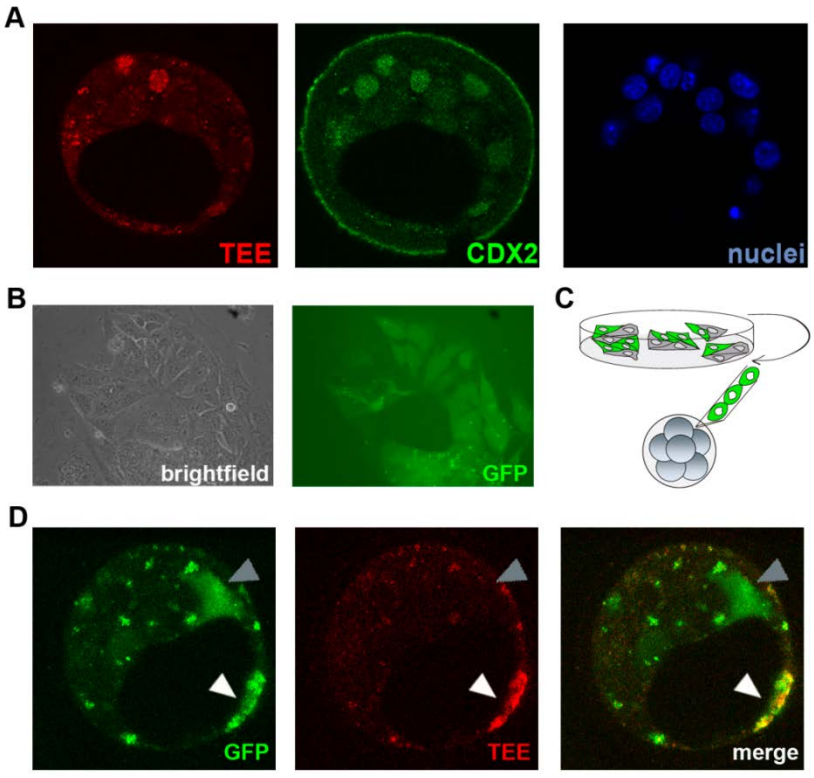


Figure 16. TS_R cells drive mRFP reporter expression in blastocysts.

(A) TS_R cells re-express the mRFP-driven reporter. Embryos were stained for mRFP (red), CDX2 (green) and DAPI (blue). (B) GFP-infected TS_R cells. Left panel shows a brightfield image of the cells. Right panel shows GFP expressing TS cells. (C) GFP+ TS_R cell injection into a morula. (D) TS_R cells re-express mRFP driven reporter. Endogenous expression of GFP (green) and H2BmRFP (red). White arrowheads point to GFP+,mRFP+ TS_R cells, grey arrowheads show GFP+,mRFP- TS_R cells.

We conclude that the TEE is inactive in cultured TS cells and that necessary inputs are lacking in the culture system, because when the cells are re-inserted into the embryo, the TEE is still capable of driving reporter expression. Intriguingly, only a proportion of TS_R cells infected with GFP were able to re-activate the transgene when injected back in morula, indicating a need for further characterization of the process. The fact that we also find a lack of reporter expression in cultured TS_R cells suggests that two different regulatory mechanisms may be operating on *Cdx2.* Hence, TS cells in culture may reflect the TEE regulatory activity shown in the post implantation embryo, pointing to a stage dependent *Cdx2* regulatory network in preimplantation versus the post-implantation embryos.

4. Searching for known upstream regulators of *Cdx2* that act through the TEE.

To better understand *Cdx2* regulation during the first lineage choice, we decided to look for transcription factors that act upstream of the TEE to direct *Cdx2* expression. Upon compaction, *Cdx2* expression is upregulated and maintained specifically in the outermost population of the morula. Stabilization of *Cdx2* in the prospective trophectoderm is triggered through a combination of inside versus outside positioning of blastomeres together with polarization and cell to cell contact cues. To date, the only transcription factor proposed to regulate *Cdx2* concomitant with compaction is TEAD4, acting downstream of Hippo (Nishioka et al., 2009). In the context of the first lineage decision, *Oct4* is involved in the maintenance of ICM versus TE fate through *Cdx2* repression around the time of implantation (Ralston et al., 2010). Both TE and ICM transcription factors show a complementary expression pattern, and it has been shown that both gene regulatory networks crosstalk. In ES cells, pluripotency expression patterns are reinforced by the reciprocal repression of *Cdx2* by *Oct4*, *Nanog* and *Sox2* in ES cells (Boyer et al., 2005; Loh et al., 2006), also, it has be shown that CDX2 binds to and transactivates its own promoter and positively regulates its expression (Chew et al., 2005). Moreover, CDX2 and OCT4 have been proposed to form a reciprocal inhibitor complex that may resemble what happens in the early embryo (Niwa et al., 2005).

4.1 Analysis of the TEE response to CDX2 and OCT4.

To test if CDX2 regulates its own expression through the TEE in the early onset of CDX2 restriction to outer cells, we examined homozygous *Cdx2* $^{-/-}$ mice. In *Cdx2* mutants blastocoel formation initiates, ICM lineage is specified, however epithelial integrity is not maintained and embryos fail to implant (Strumpf et al., 2005). In addition, mouse embryos are able to initiate TE specification and to polarize when *Cdx2* is deleted maternal-zygotically. Thus, *Cdx2* is not responsible for the first events that lead to TE formation (Blij et al., 2012).

Analysis of *Cdx2* $^{-/-}$ blastocysts revealed that TEE reporter activity was still restricted to the TE (Figure 17B). Counting of TEE+ blastomeres to rule out that the lack of *Cdx2* results in a reduced number of TEE+ cells revealed that the number of positive cells was the same as in wild-type littermates (Figure 17C), thus the TEE does not change its ability to drive reporter activity in the absence of *Cdx2* and is upstream of *Cdx2*.

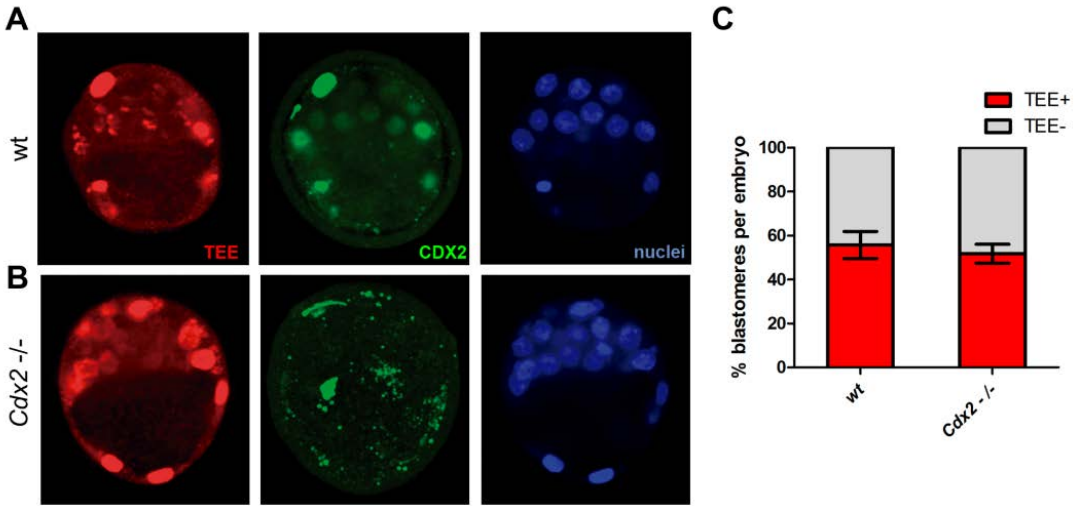


Figure 17. The TEE is not an autoregulatory element.

(A-B) Activity of the TEE, detected by immunohistochemistry with anti-mRFP (red) antibody, and immunodetection of CDX2 (green) in (A) wild-type (wt) and (B) *Cdx2* $^{-/-}$ blastocysts. Nuclei were stained with DAPI. (C) Percentage of TEE-positive cells per wt and *Cdx2* $^{-/-}$ embryo.

To test if OCT4-mediated repression of *Cdx2* in the ICM occurs via the TEE, we took advantage of a genetically modified ES cell line (ZHBTc4) in which *Oct4* is repressed by tetracycline. Upon tetracycline (Tc) addition, ES cells switch off the gene regulatory pluripotency program, convert to TS cells, and upregulate *Cdx2*. It has been suggested that this transdifferentiation switch could mimic the first lineage choice and the early

distinction between TE and ICM (Niwa et al., 2000). We took advantage of this system to address if derepression of TE genes through inactivation of *Oct4* could lead to TEE-driven mRFP activity.

To set up the correct conditions, we tested different Tc doses (data not shown) and also tested *Oct4* repression with Tc in ES or TS culture media. By immunohistochemistry, we were able to detect a downregulation of OCT4 upon drug administration (Figure 18A, C) and an upregulation of CDX2 48h post induction in EMFI-TS cultured cells (Figure 18B, D). Expression profiling of *Oct4*, *Nanog* and *Cdx2* by qPCR 48h after Tc administration detected a strong reduction of *Oct4* and *Nanog*, concomitant with upregulation of *Cdx2*. The upregulation of *Cdx2* was even stronger when Tc was added and cells were cultured in EMFI-TS medium, probably because this medium favours TS cell self-renewal and maintenance (Figure 18E).

Next, we transfected ZHBTc4 ES cells with either empty vector, *Oct4*DE as positive control, or fragments from the *Cdx2* regulatory locus previously used in the TS transfection assay (Figure 14, fragments I, II and III). Transfection of the selected regulatory sequences into ES cells after 48h in culture showed that *Oct4*DE directed robust expression of RFP, and that fragment II was moderately active in ES medium, in line with the results obtained in the lab by transient transgenesis of embryos in which this fragment is active in the ICM. When *Oct4* was repressed by Tc and EMFI-TS medium added, activities of *Oct4*DE and fragment II were strongly reduced; in contrast, no RFP activity could be detected in neither the fragment III (TEE) nor fragment I (Figure 18F).

We therefore conclude that the TEE does not respond to the changes of gene expression patterns observed when cells switch from ES to a TS-like state upon *Oct4* repression, and hence the TEE is not the regulatory element responsible for the mutual maintenance and cross-repression through *Cdx2* and *Oct4*.

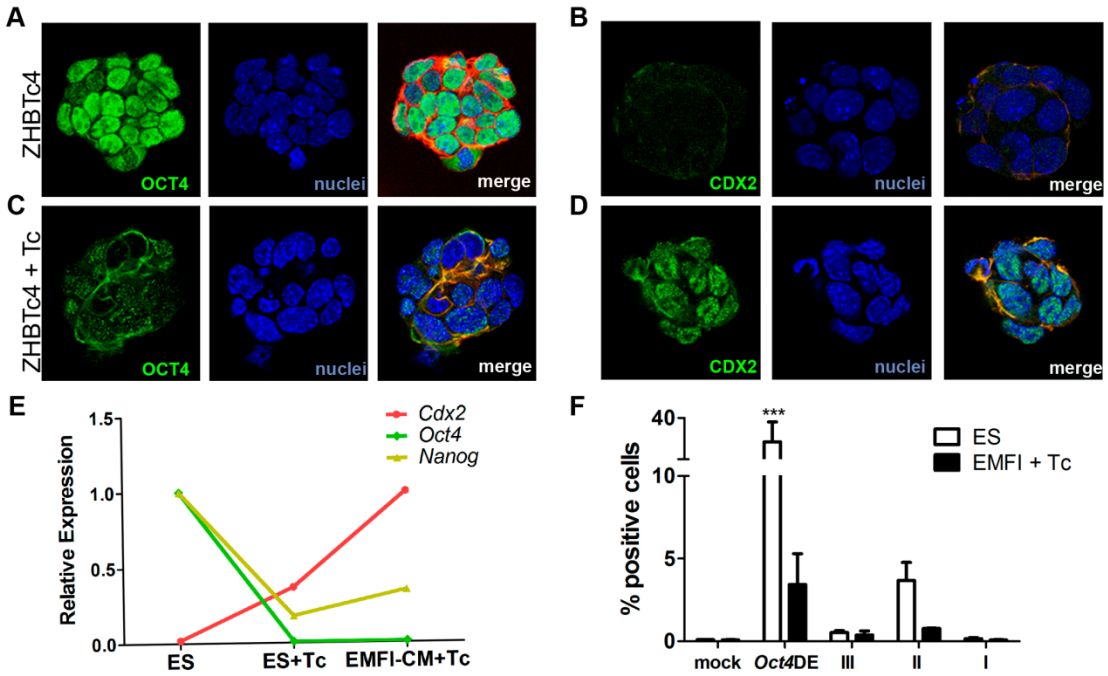


Figure 18. The TEE does not respond to OCT4.

OCT4 and CDX2 expression in ZHBTc4 cell line in ES medium or upon tetracycline induction in EMFI-TS medium. (A-C) OCT4 (green) staining (B-D) CDX2 staining (green) in ZHBtc4 cells; (Tc, 48h). Nuclei were stained with DAPI and F-actin with Rhodamine-phalloidin (red) to detect cell membrane. (E) Relative expression of *Cdx2*, *Oct4* and *Nanog* in the ZHBTc4 cell line in different culture conditions with and without Tc. (F) Percentage of ZHBTc4 cells that show reporter expression in transient transfections upon Tc addition. Data are means ± s.e.m. n=3. ***p<0.001, by Student's t-test versus mock.

4.2 Analysis of the mRFP-TEE line response to the Hippo pathway.

Having ruled out an action of *Cdx2* and *Oct4* on the enhancer, we next analysed the regulation of the TEE by the Hippo pathway. The Hippo pathway is responsible for *Cdx2* regulation upon polarization in preimplantation stages; however, whether *Tead4* regulates *Cdx2* directly is still not known as the *Cdx2* trophoblast enhancer has not been identified (Nishioka et al., 2009).

To check to what extent the TEE is a YAP/TEAD response element, we bred the mRFP reporter line to the *Tead4* $^{-/-}$ background (Nishioka et al., 2008). Once the Hippo pathway is turned off, YAP shuttles to the nucleus where it binds to TEAD4. TEAD4 is the transcriptional effector of the pathway and its absence impedes the expression of target genes such as *Cdx2*. Analysis of the TEE in *Tead4* mutants revealed that it was fully active in homozygous *Tead4* $^{-/-}$ 3.5 dpc embryos, and even conserved preferential activity in outer blastomeres (Figure 19B).

Consistent with previous reports (Kaneko and DePamphilis, 2013; Nishioka et al., 2008), we also found that some *Tead4* $^{-/-}$ embryos retained CDX2 expression (Figure 19C). Moreover, *Cdx2* stochastic expression has been identified in eight-cell embryos where TEAD4 and YAP expression is ubiquitous (Dietrich and Hiiragi, 2007; Hirate et al., 2012), suggesting that additional inputs are necessary for the early expression of *Cdx2*. Furthermore, we did not detect significant differences in total cell number or RFP+ blastomeres per embryo between wild-type, heterozygous and homozygous embryos (Figure 19D), precluding an effect of the loss of *Tead4* on the total number of cells in which the TEE is active.

To investigate this further, we took advantage of the 5TVER7 ES cell line, which stably expresses a tamoxifen-inducible active form of the TEAD4 protein (Tead4VP16ER). This active form of the protein contains the Tead4 DNA binding domain fused to the transcriptional activator of the herpes virus VP16 and the estrogen receptor domain (ER). Upon tamoxifen induction, the fusion protein translocates to the nucleus and activates its target genes. When introduced in ES cells, they activate *Cdx2* expression and acquire a TS cell phenotype (Nishioka et al., 2009). We transfected 5TVER7 cells with Fragments II, III (TEE) or with the *Oct4DE* as a control, and treated them with tamoxifen in EMFI-TS media. After 48 hours in these conditions, qPCR analysis showed robust activation of *Cdx2* and downregulation of *Oct4*, accompanied by a marked decrease in *Oct4*DE activity (Figure 19E and F). Despite the forced expression of the active TEAD4 form, we did not observe a strong effect on fragment III (TEE) activity compared with culture in ES cell culture medium without tamoxifen, although treatment did increase activity slightly but significantly compared with control empty vector (Figure 19F). Interestingly, fragment II showed a relative sustained activity in both conditions suggesting a Hippo-independent regulatory mechanism. These experiments provide additional evidence that TEAD4 alone is not sufficient to fully activate the TEE.

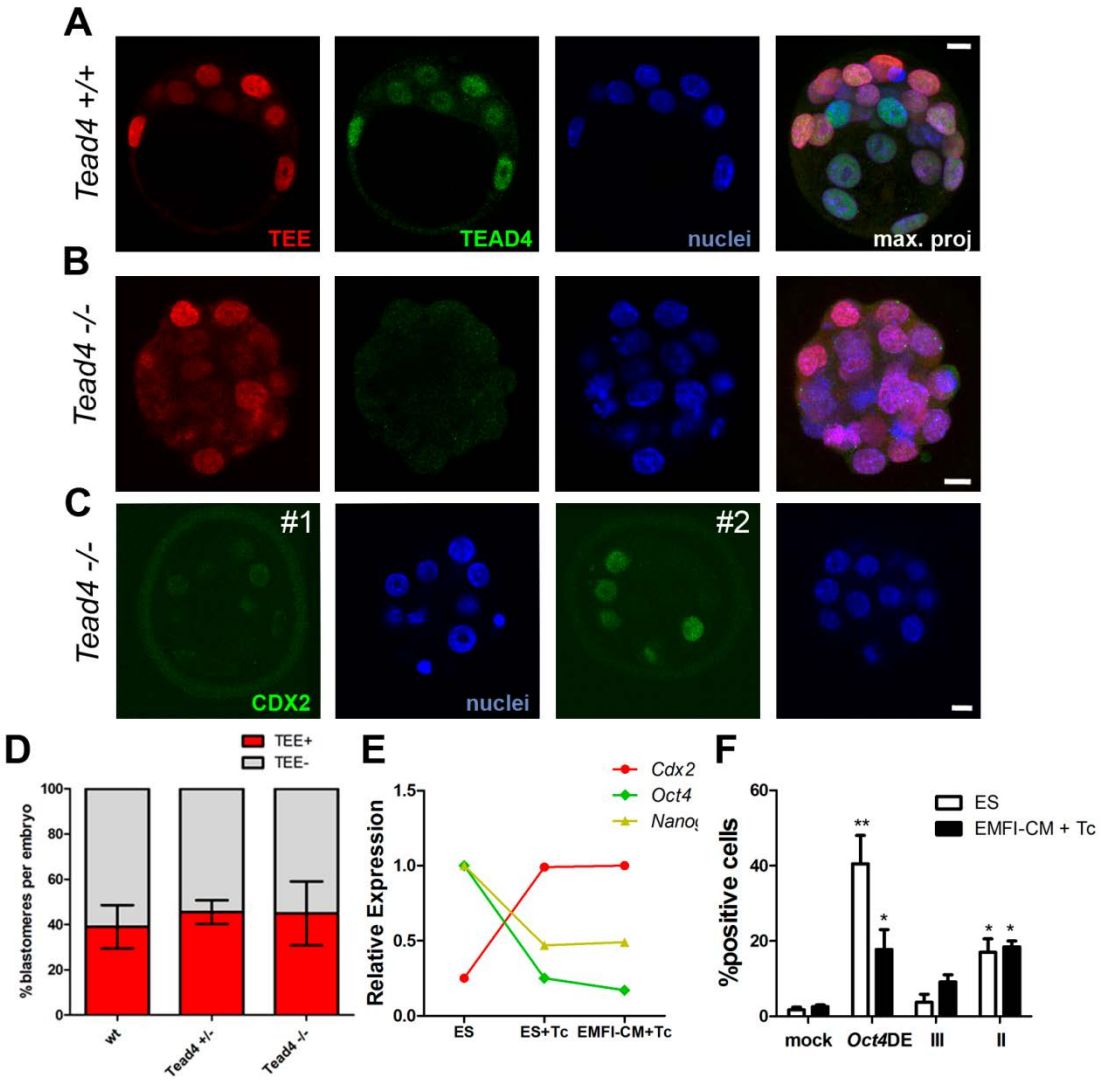
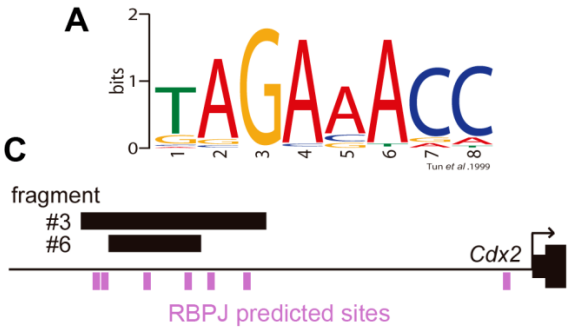


Figure 19. The Hippo pathway is not sufficient to regulate the TEE.

(A-B) TEE activity and TEAD4 immunodetection in (A) wild-type (wt) and (B) *Tead4* $^{-/-}$ embryos. (C) CDX2 expression in two *Tead4* $^{-/-}$ embryos from the same litter. Nuclei were stained with DAPI. Maximal projections of merged images are shown in the right panels. Scale bars, 10 µm. (D) Percentage of TEE-positive cells per embryo in wt, *Tead4* $^{+/-}$ and *Tead4* $^{-/-}$ blastocysts. (E) Relative expression of *Cdx2*, *Oct4* and *Nanog* in the 5TVER7 and (F) percentage of 5TVER7 cells that activate the TEE in response to tamoxifen (Tx). Data are means ± s.e.m. n=3. **p<0.01, *p<0.05 by Student's t-test compared to mock.

5. Searching for unknown upstream regulators of *Cdx2* that act through the TEE.

The above results suggested that there must be other pathways responsible for the early expression of *Cdx2*. To identify other transcriptional inputs acting on the TEE, we looked for transcription factors that may relate the TEE to its early activation. This analysis was carried out *in silico*, using the FIMO (Find Individual Motif Occurences) software program. We searched the whole upstream region of *Cdx2* up to the next gene, a distance of 9.6kb (Figure 9), for putative binding sites specifically enriched on the TEE. In the search we included RBPJ, the transcriptional effector of the Notch pathway (Kopan and Ilagan, 2009). Interestingly, over the 9.6kb span, six out of seven RBPJ predicted binding sites were included in the TEE sequence (Figure 20). The remaining predicted site was at the *Cdx2* promoter.



location	Strand	Start	End	p-value	Matched Sequence
TEE	+	3792	3810	2.47e-05	GTTCGGTGGGAACAATTAA
TEE	–	3575	3593	5.12e-05	CCAGTGTGGGAAACCGGCC
TEE	+	4628	4646	0.000119	ATAGGGTGGGAGTTTTGAG
TEE	–	1881	1899	0.000144	GCCCGGTGGGAGCTCTGGA
OUT	+	7806	7824	0.000182	AAGAGGTGGGAATTGAGGG
TEE	+	2843	2861	0.000191	GAGCAGTGAGAACAAAGTG
TEE	–	1994	2012	0.0002	ATGAAGTGGGAGATGAAAG

Figure 20. Identification of putative RBPJ binding sites.

(A) RBPJ motif logo generated to identify the putative binding sites according to the position-specific scoring matrix from the literature (Tun et al., 1994) (B) High scoring motif occurrences in the 9.6kb region for RBPJ sites identified by FIMO. TEE location, strand, p-value, position and sequence are indicated. (C) Location in the *Cdx2* locus of the predicted RBPJ binding sites (pink). Identified fragments that contain TEE activity as in Figure 9 are shown. The majority of RBPJ sites are included in the TEE.

5.1 Characterization of the Notch pathway in early preimplantation.

Compaction brings about a distinction between two cell populations: an external population of polarized cells and an apolar inner cell population. Position and polarization are key features of early development and the role of Notch-mediated cell-to-cell signalling and cell fate decisions during development is well established (Koch et al., 2013), although no role in early preimplantation has been assigned. There is scant evidence available for the expression of Notch pathway components at this stage

(Cormier et al., 2004; Wang et al., 2004), and no studies have addressed their differential localization. To investigate the role of Notch pathway in preimplantation development, we studied its spatial localization in the early mouse embryo.

To examine the activation status of the Notch pathway during preimplantation stages we used a mouse line containing a CBF:H2B-Venus reporter, which contains multiple RBPJ binding sites driving the expression of a nuclear-targeted Venus fluorescent protein and that faithfully reports the activation state of the pathway (Nowotschin et al., 2013). The reporter was active in a few cells at the 8-cell stage, showed stochastic expression in morula stage and its activity became progressively restricted to outer cells by the 3.5-dpc blastocyst stage (Figure 21A-C). In addition, we examined reporter activity together with endogenous CDX2 expression in early blastocysts of this line, finding that CBF:H2B-Venus reporter colocalized with CDX2 in the nascent TE (Figure 21D). To confirm reporter specificity, we treated embryos from CBF:H2B-Venus reporter line with the gamma-secretase inhibitor RO4929097 (RO). RO is a well characterized and widely used tool that interferes with Notch signalling by blocking the processing of the receptor (Munch et al., 2013). Treatment of embryos from this line with the inhibitor resulted in a strong downregulation of the reporter (Figure 21E).

Finally, we studied the endogenous distribution of the NOTCH1 (N1ICD) intracellular domain. N1ICD is the active processed form of Notch receptor. Upon interaction with the ligand, NOTCH receptor is cleaved and N1ICD translocates into the nucleus where it binds RBPJ and activates its target genes (Figure 6). Immunofluorescence detection of N1ICD in embryos showed that it is localized to the nucleus only in the outer cells of the blastocyst (Figure 21F). Together, these data show that the Notch signalling pathway is active specifically in the developing TE.

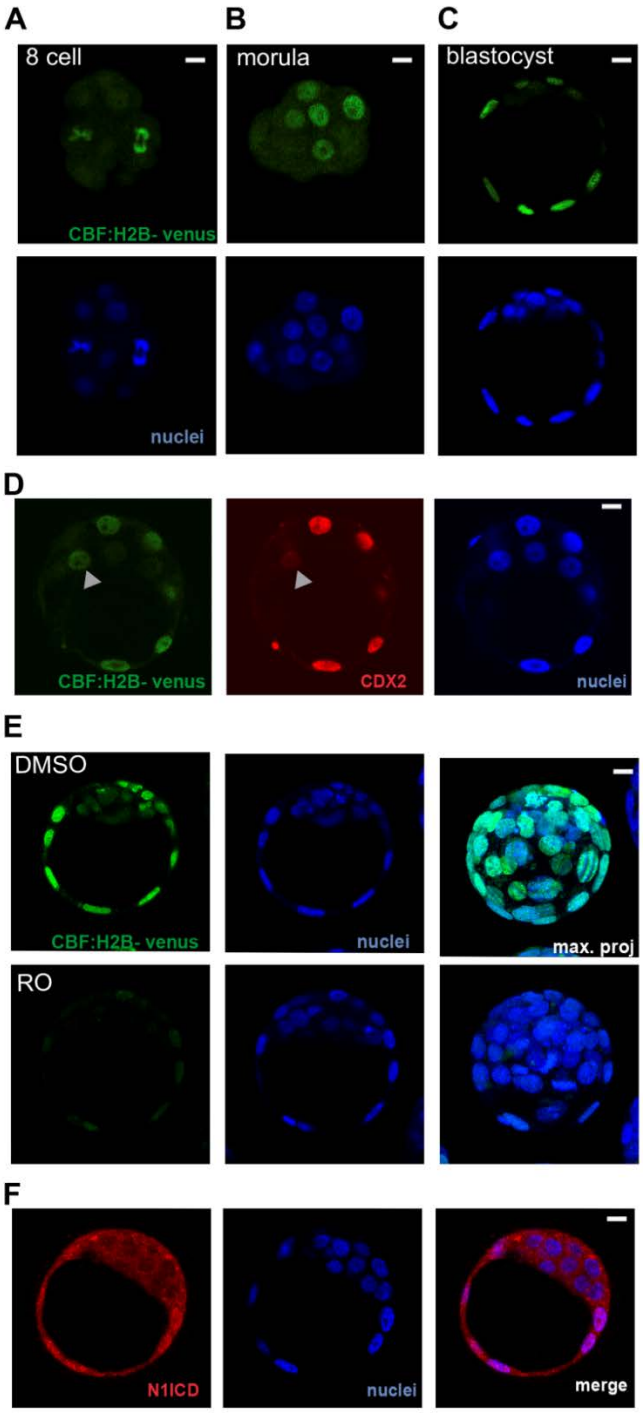


Figure 21. The Notch signalling pathway is active in the trophectoderm.

(A-C) CBF:H2B-Venus reporter expression at (A) the non-compacted 8-cell stage, (B) morula stage, and (C) blastocyst stage. (D) Detection of CBF:H2B-Venus reporter (green) and CDX2 (red) at the blastocyst stage. Grey arrowhead points to a low level CDX2-expressing cell. (E) Treatment of embryos from the CBF:H2B-Venus line with the g-secretase inhibitor RO4929097 (RO) downregulates reporter activity as compared to controls incubated in DMSO. (F) Immunodetection of NOTCH1 intracellular domain (N1ICD, red). Nuclei were stained with DAPI. Scale bars, 10 μm

5.2 Notch activity in TS cells.

The TE-specific expression of active Notch shown above suggested that Notch components might be differently distributed in TS cells versus ES cells. Surprisingly, CBF:H2B-Venus ES cells still express the reporter in a subpopulation of cells (Nowotschin et al., 2013), although in this mouse strain no reporter expression is detected in the ICM, from where the ES cells are derived (Figure 21). Moreover, all Notch pathway components are expressed in ES and in TS cells (Cormier et al., 2004), and Jagged1 receptor has been immunodetected in ES cells (Lowell et al., 2006). To try to resolve the discrepancy between the expression pattern we observe in the ICM of the blastocyst and findings in cultured cells, we studied N1ICD expression in blastocyst-derived ES and TS cells.

As a positive control, we therefore used human umbilical vein endothelial cells (HUVEC), a widely used model for the study of endothelial-cell function and pathology that strongly express Notch pathway components (Figure 22A) (Shawber and Kitajewski, 2004; Uyttendaele et al., 2001). In TS cells, we could detect N1ICD in a limited number of nuclei (Figure 22B). To validate N1ICD expression in other TS cell lines, we stained for N1ICD in TS_R cells, detecting the same expression pattern (Figure 22C). Strikingly, staining of the active form of *Notch1* was heterogeneous in both TS cell lines, while in the blastocyst, an even distribution of N1ICD was found throughout the TE. In contrast, and contrary to what has been shown with the CBF:H2B-Venus reporter, no specific signal was detected in ES cells (Figure 22D). This result suggests that ES cells might express other Notch receptors that are not present in the ICM, or that the expression levels of the fluorescent protein are stronger and thus more easily detectable than endogenous N1ICD.

Together, this data indicate that analysis of N1ICD in blastocyst-derived stem cells does not mimic completely its expression pattern in blastocysts. Thus, to better understand the Notch pathway in preimplantation stages, ES and TS cells do not seem a consistent *in vitro* tool.

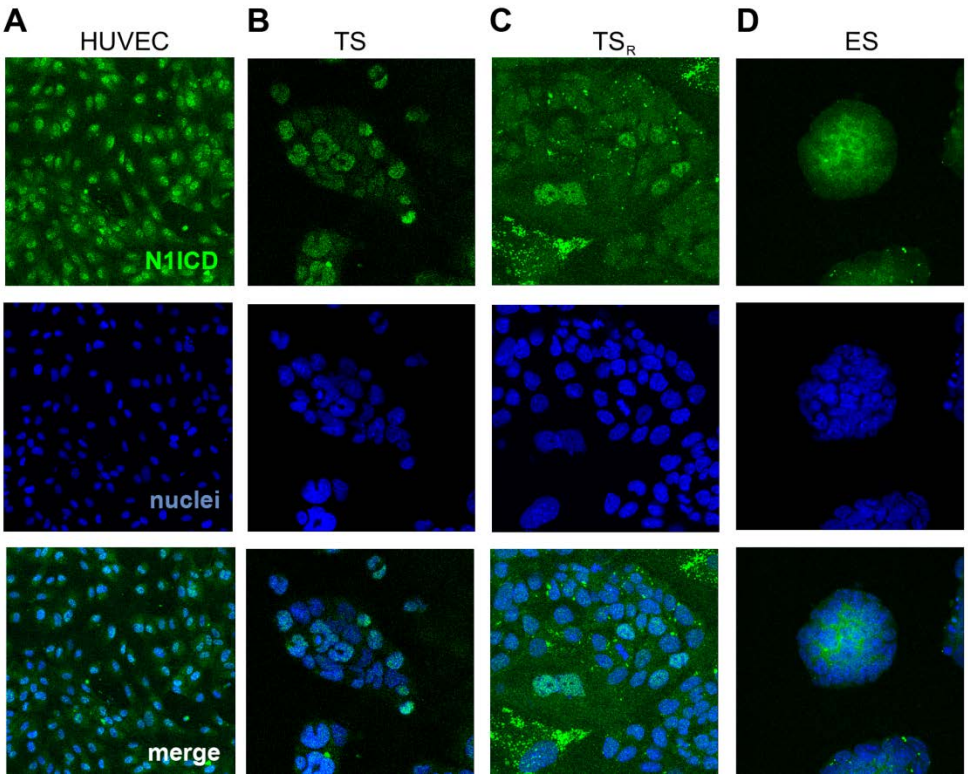


Figure 22. TS cells contain N1ICD heterogeneous levels.

(A-D) Immunodetection of N1ICD (green) in (A) HUVEC, (B) TS cells, (C) TS_R cells and (D) ES cells. N1ICD is present in all (A) HUVEC nuclei, but only a few nuclei (B) TS and (C) TS_R cells. (D) ES cells show no sign of expression. Nuclei were stained with DAPI and the merge image is shown.

5.3 Functionality of the identified TEE RBPJ and TEAD binding sites.

To characterize the possible direct link between RBPJ (the Notch pathway) and *Cdx2*, we analysed the minimal 1.3 kb TEE sequence (fragment #6, Figure 9) for bona fide RBPJ sites, detecting four high-confidence sites for RBPJ. Given that we detected a slight but significant increase in TEE activity when Tead4VP16 was overexpressed (refers to section 4.2), we also searched for putative TEAD binding sites (Anbanandam et al., 2006) and identified two.

To investigate the function of the identified sites in the minimal TEE fragment (fragment #6), we generated transient transgenic embryos expressing the wild-type minimal TEE or versions mutated in the four RBPJ sites alone (TEE^{RBPJ}mut), the two TEAD sites alone (TEE^{TEAD}mut) or in the RBPJ and the TEAD sites ($TEE^{RBPJ/TEAD}$mut) (Figure 23A). The mutated constructs contained substitutions of the minimal nucleotides that have been shown to abolish RBPJ binding (Tun et al., 1994). We then tested the capacity of each

construct to drive reporter expression by transient transgenesis. Transgenic embryos carrying the TEE^{RBPJ}mut or the TEE^{TEAD}mut version retained TE lineage-restricted activity as was the case for the TEE, while TEE activity was abolished when the RBPJ and TEAD sites were together mutated (Figure 23B).

The lack of response of the individual RBPJ and TEAD sites suggests that the enhancer responds in a redundant fashion in the blastocyst and that the activity is only reduced when a combination of mutated binding sites is used.

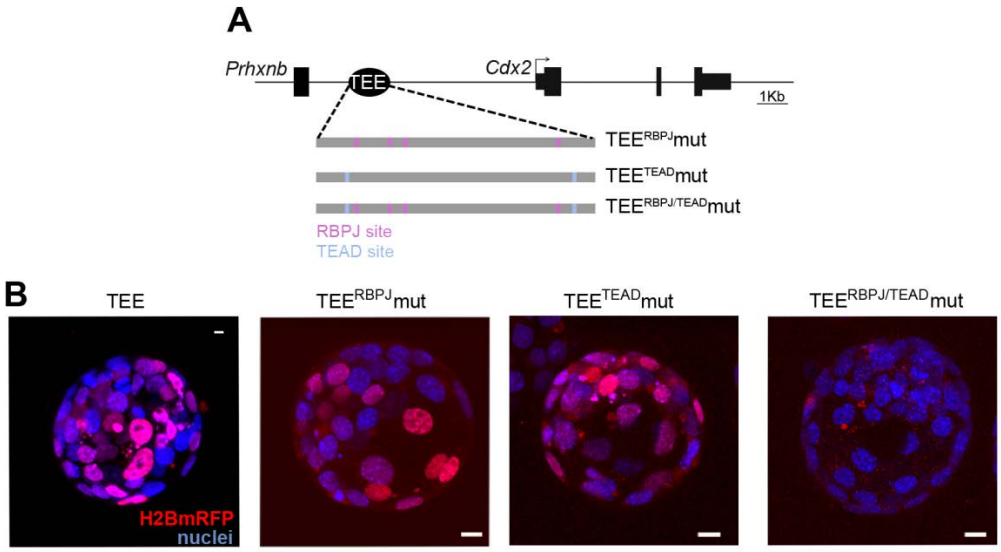


Figure 23. RBPJ together with TEAD binding sites are necessary for the activity of the Cdx2 TEE.

(A) The different TEE fragment #6 versions with the identified putative binding sites for RBPJ (pink) and for TEAD (blue). (B) Activity of wild type TEE, TEE^{RBPJ}mut, TEE^{TEAD}mut and TEE$^{RBPJ/TEAD}$mut constructs in transient transgenic embryos. In TEE$^{RBPJ/TEAD}$mut embryos, no TE-specific activity was observed. Nuclei were stained with DAPI. Scale bars, 10 μm.

As shown above, the TEE is able to respond with similar efficiency when the mutated sites for either RBPJ or TEAD are present. In this scenario, the input from the remaining pathway is sufficient to induce TEE driven reporter expression. To study the effect on the TEE of each pathway when the other input is missing, we microinjected mutated TEE constructs and treated the embryos with pharmacological blockers of the Notch and Hippo signalling pathways. The day after microinjection, surviving 2-cell stage embryos were divided in two pools and treated either with the solvent used to prepare the drug (vehicle) or with the pharmacological inhibitor and cultured for three days (Figure 24A). Embryos that reached the blastocyst stage where scored for RFP

expression and classified as TE specific, unspecific (TE+ICM), or weak in cases were embryos showed less than four RFP+ cells or generally low RFP intensity.

Notch pathway disruption in embryos microinjected with TEE^{TEAD}mut by treatment with the gamma-secretase inhibitor RO4929097 (RO) strongly reduced the number of transgenic embryos showing TE specific expression compared with microinjected controls incubated with DMSO (Figure 24B, D). This result confirms that attenuation of Notch signalling diminishes activity of a TEE that lacks functional TEAD sites.

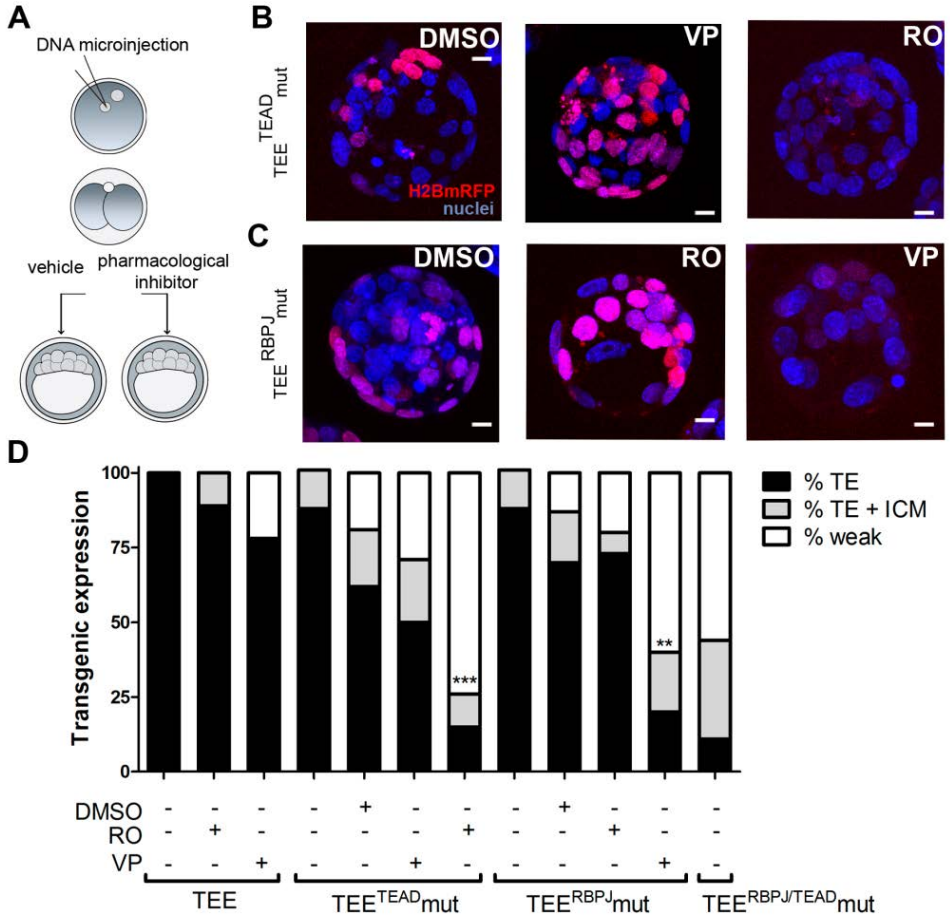


Figure 24. Notch and Hippo are responsible for full enhancer activity through the identified binding sites.

(A) Experimental setup for inhibitor or vehicle treatment of embryos from the same batch of microinjected embryos. (B) TEE^{TEAD}mut activity in transient transgenic embryos treated with DMSO, the TEAD/YAP inhibitor VP or the γ-secretase inhibitor RO. (C) TEE^{RBPJ}mut activity in transient transgenic embryos treated with DMSO, RO or VP. (D) Percentage of transgenic expression (TE, TE + ICM, or weak) in embryos microinjected with the TEE, TEE^{TEAD}mut or TEE^{RBPJ}mut versions and treated with the indicated inhibitor. **p<0.01, ***p<0.001 by Chi squared test.

As for the Hippo pathway, we used verteporfin (VP), a small molecule that inhibits TEAD–YAP association (Liu-Chittenden et al., 2012). To confirm that VP disrupts Hippo signalling as expected, we treated embryos from the mRFP-TEE reporter line and examined reporter activity together with endogenous CDX2 expression (Figure 25). A high percentage of VP treated embryos phenocopied *Tead4* $^{-/-}$ embryos (Figure 19), with strong reduction of CDX2 but retention of TEE activity (Figure 25).

Verteporfin treatment on TEE^{RBPJ}mut microinjected embryos significantly reduced TE specificity (Figure 24C, D), showing that interfering with the transcriptional activity of YAP and TEAD4 when the RBPJ sites are missing decreases enhancer activity. As expected, treatment of TEE^{RBPJ}mut microinjected embryos with RO, or TEE^{TEAD}mut microinjected embryos with VP had no significant effect on the TE specificity of reporter activity (Figure 24).

Taken together, these results show that transcriptional inputs from both pathways are responsible for full enhancer activity through the specific binding sites identified, acting in a redundant fashion *in vivo*.

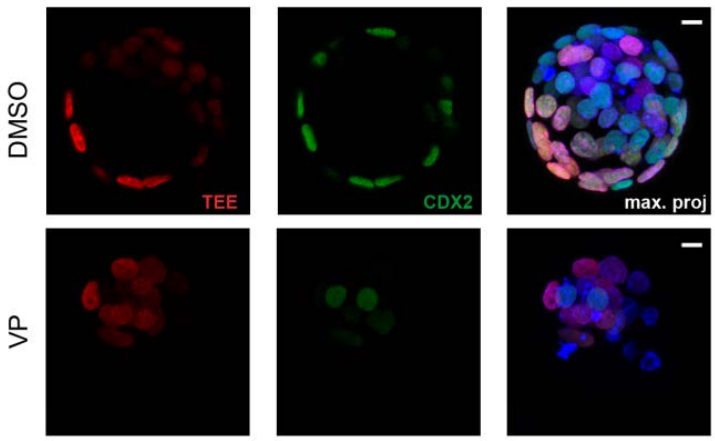


Figure 25. Verteporfin treatment phenocopies *Tead4* $^{-/-}$.

Effect of the TEAD–YAP inhibitor verteporfin (VP) on treated mRFP-TEE embryos from the 2-cell stage. Immunodetection of TEE activity (red) and CDX2 (green). Nuclei were stained with DAPI. Maximal projections of merged images are shown in the right panels.

Scale bars, 10 μm.

6. Does the Notch signalling pathway together with *Tead4* regulate *Cdx2* trough the TEE?

To further explore the regulation of the TEE by Notch, we tested its activity in the absence of Notch signalling by breeding the mRFP-TEE line in the *Rbpj* $^{-/-}$ background. Absence of the transcription factor *Rbpj* impedes the recruitment of the coactivator complex to activate Notch target genes, thus making the cell unresponsive to Notch signalling (Oka et al., 1995).

Consistent with the activity of TEE ^{RBPJ}mut in the TE (Figure 23), *Rbpj* $^{+/-}$ and *Rbpj* $^{-/-}$ embryos showed normal TEE-driven expression, as was the case for *Tead4* $^{+/-}$ and *Tead4* $^{-/-}$ embryos (Figure 26A-C, Figure 19). These embryos also had normal numbers of cells and of TEE+ blastomeres, although *Rbpj* $^{-/-}$ embryos showed a notable but statistically non-significant tendency to have fewer TEE+ cells than wild-type embryos (Figure 27A). Altogether, these results demonstrate that inactivation of a single pathway is not sufficient to abolish the ability of the TEE to drive RFP expression and suggest that the TEE compensates the lack of one of the inputs to ensure its activity.

In previous results, we observed a combinatorial effect of Hippo and Notch, with TEE activity lost only when both RBPJ and TEAD sites were mutated. Also, TE specificity was significantly reduced when single TEE mutated versions were pharmacological disrupted for the converse pathway (see 5.3). To further test for interaction between Notch and Hippo on the TEE, we generated embryos of the mRFP-TEE line containing different combinations of *Tead4* and *Rbpj* mutant alleles. Double heterozygote embryos (*Rbpj* $^{+/-}$;*Tead4* $^{+/-}$) contained significantly fewer than normal TEE+ cells (Figure 26D, Figure 27A) and this effect was more marked in *Rbpj* $^{-/-}$;*Tead4* $^{+/-}$ embryos, where only a few mRFP cells were detected in the TE (Figure 26E, Figure 27A). This was not due to disruption of overall cell number or the inside-outside distribution of blastomeres in these embryos, as these parameters showed no significant difference compared with wild type, *Tead4* $^{+/-}$, *Rbpj* $^{+/-}$ or *Rbpj* $^{-/-}$ blastocysts (Figure 27B).The reduction in RFP+ blastomeres in double heterozygotes suggests that activation of the TEE requires minimum threshold amounts of TEAD4 and RBPJ, since the lack of one copy of each is enough to reduce the number of TEE+ cells. Unexpectedly, we did not recover any double homozygote knock-out embryos at 3.5 dpc and the frequency distribution did not follow Mendelian ratios, suggesting that the combined lack of these two factors causes the death of embryos at an earlier stage (Figure 27C).

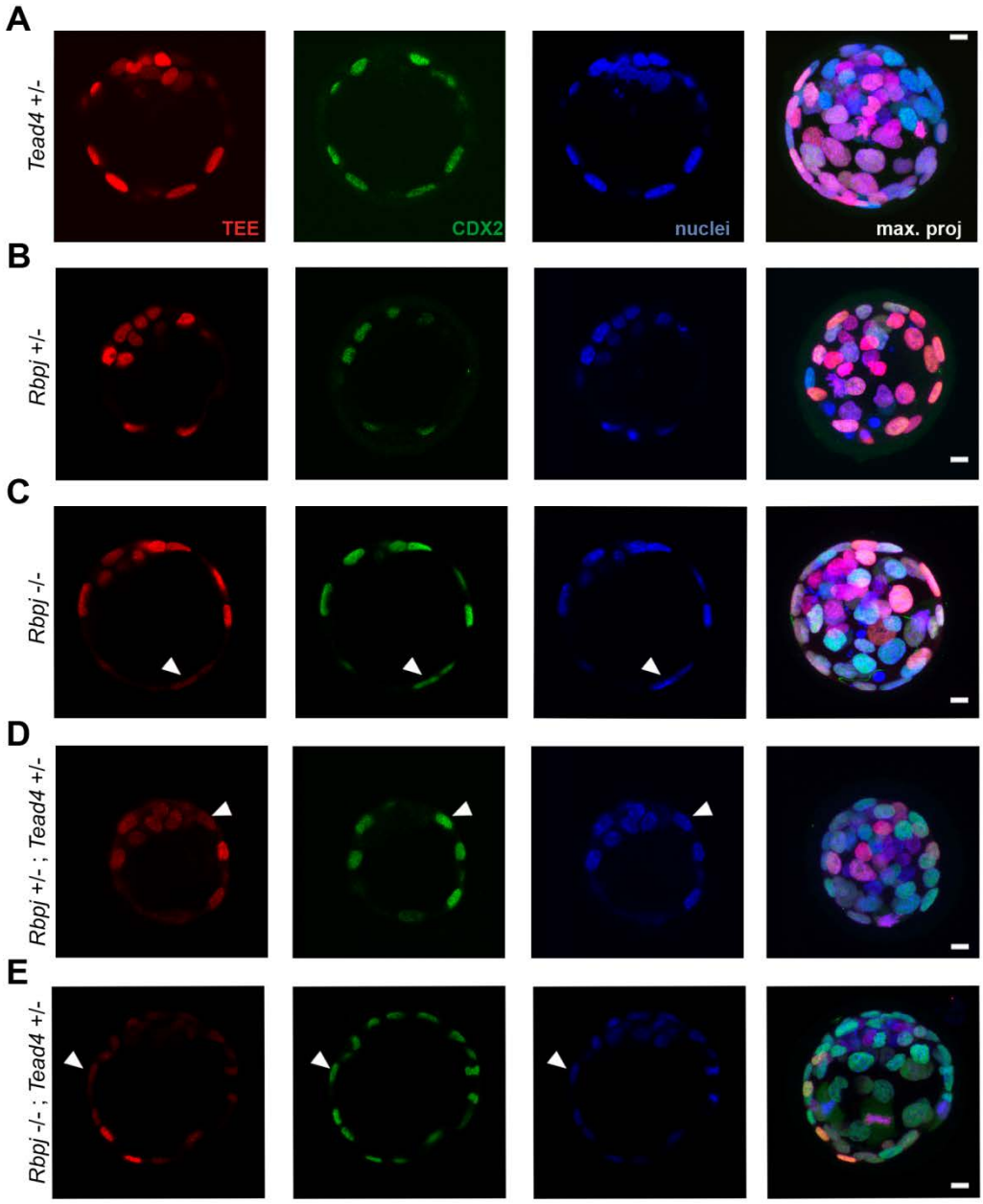


Figure 26. TEE activity and *Cdx2* expression require transcriptional inputs from Notch and Hippo.

(A-E) TEE activity and CDX2 immunodetection in (A) *Tead4* $^{+/-}$, (B) *Rbpj* $^{+/-}$, (C) *Rbpj* $^{-/-}$, (D) *Rbpj* $^{+/-}$;*Tead4* $^{+/-}$, and (E) *Rbpj* $^{-/-}$;*Tead4* $^{+/-}$ mutant embryos. Nuclei were stained with DAPI. Arrowheads in (C-E) indicate TEE-, CDX2+ outer blastomeres. Maximal projections of merged images are shown in the right panels. Scale bars, 10 µm.

Analysis of the TEE enhancer regulation by the Notch pathway and *Tead4* showed a reduction in TEE+ cells and no double homozygous knock-out could be detected, suggesting an unexpected early role of Notch in combination with *Tead4* in

preimplantation development. However, quantification showed no difference in the total number of CDX2+ cells per embryo in the different genotypes (Figure 27B). Of note, *Rbpj* $^{-/-}$, *Rbpj* $^{+/-}$;*Tead4* $^{+/-}$ and *Rbpj* $^{-/-}$;*Tead4* $^{+/-}$ mutants contained blastomeres in which we could detect outer blastomeres expressing CDX2 but not the TEE (arrowheads Figure 26C-E), indicating that *Cdx2* regulation does not solely rely on the TEE regulatory sequence.

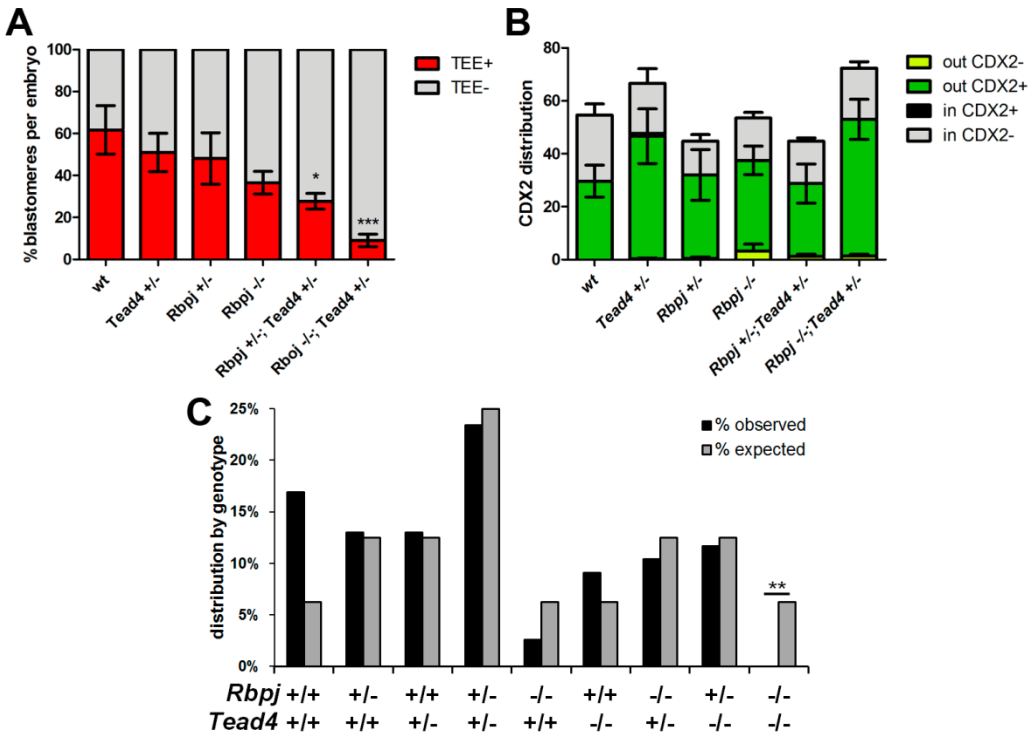


Figure 27. Rbpj and Tead4 happloinsuficient embryos show reduced TEE+ cells and double knock-outs die before 3.5 dpc.

(A) Percentage of TEE-positive cells per embryo in wild type blastocysts and in *Tead4* $^{+/-}$, *Rbpj* $^{+/-}$, *Rbpj* $^{-/-}$,*Rbpj* $^{+/-}$;*Tead4* $^{+/-}$ and *Rbpj* $^{-/-}$;*Tead4* $^{+/-}$ allelic combinations. *p<0.05, ***p<0.001 by Bonferroni post-test. (B) Average number of inside and outside cells positive or negative for CDX2 in wild-type, *Tead4* $^{+/-}$, *Rbpj* $^{+/-}$, *Rbpj* $^{-/-}$, *Rbpj* $^{+/-}$;*Tead4* $^{+/-}$ and *Rbpj* $^{-/-}$;*Tead4* $^{+/-}$ allelic combinations quantified in Figure 28.(C) Distribution (%) of embryos for the different allelic combinations. The distribution does not follow Mendelian ratios; **p<0.005 by chi-square test.

Taking into account that the TEE is a faithful *Cdx2* regulatory element and that the TEE responds to both Notch signalling and *Tead4*, we thought that cell number differences in the mRFP-TEE line when we compared the different genotypes could reflect subtle endogenous changes in *Cdx2* levels. We therefore used our segmented data on single

embryos to quantify CDX2 mean intensity per blastomere. Quantification of CDX2 protein per blastomere revealed significantly below-normal expression in *Rbpj* $^{-/-}$,*Rbpj* $^{+/-}$;*Tead4* $^{+/-}$ and *Rbpj* $^{-/-}$;*Tead4* $^{+/-}$ embryos, but not in *Rbpj* or *Tead4* single heterozygotes (Figure 28A). Notch signalling and RBPJ are thus necessary for proper expression of endogenous *Cdx2* in the embryo, in cooperation with TEAD4.

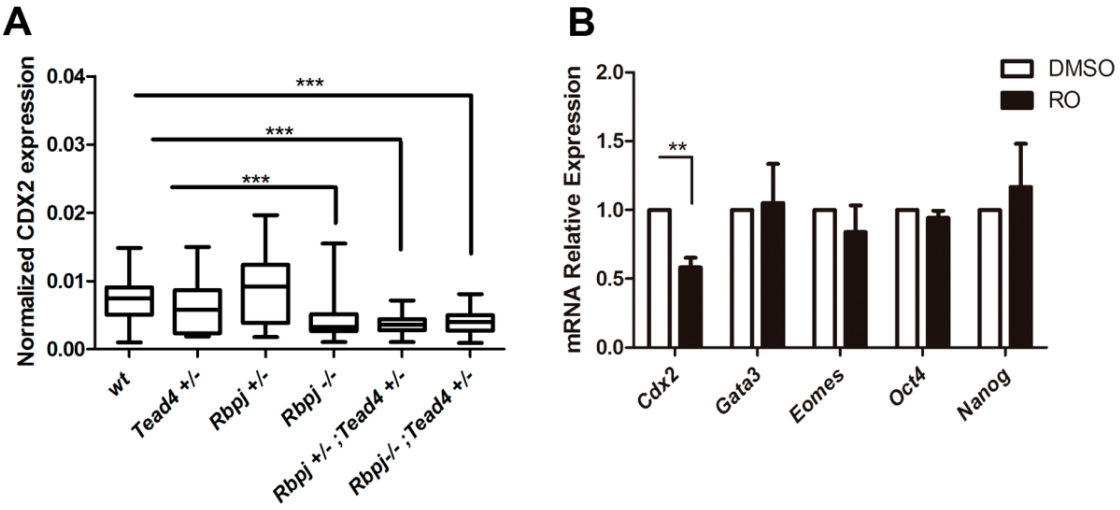


Figure 28. *Cdx2* expression requires transcriptional inputs from Notch and Hippo.

(A) Quantified CDX2 expression in outer cells of wt blastocysts and in *Tead4* $^{+/-}$, *Rbpj* $^{+/-}$, *Rbpj* $^{-/-}$, *Rbpj* $^{+/-}$*;Tead4* $^{+/-}$ *and Rbpj* $^{-/-}$*;Tead4* $^{+/-}$ allelic combinations. Boxes span the 25th to the 75th percentile, internal horizontal lines indicate median values, and whiskers show minima and maxima. ***p<0.001 by Bonferroni post-test. (B) Relative expression of *Cdx2*, *Gata3*, *Eomes*, *Oct4* and *Nanog* in pools of 25 embryos treated from 2-cell until blastocyst stage with DMSO or RO. Data are means ± s.e.m. **p<0.01 by Student's t-test.

We confirmed the effect of Notch on *Cdx2* expression by treating pools of embryos from the 2-cell to blastocyst stages with the γ-secretase inhibitor RO, which significantly decreased mRNA expression of *Cdx2.* We also examined changes in expression of other TE-expressed genes, finding that neither *Gata3* nor *Eomes* changed significantly (Figure 28A). This might reflect the earlier requirement of *Cdx2* and the fact that *Gata3* and *Eomes* only function later in TE development (Ralston et al., 2010; Strumpf et al., 2005). We did not detect any change in *Oct4* or *Nanog*. This result confirms that Notch inhibition specifically downregulates *Cdx2* expression and that no alteration in the pluripotency transcription factors can be detected.

To verify the response of the TEE to the Notch pathway together with *Tead4*, we pharmacologically repressed TEE activity by treating embryos with RO together with VP

(for RO and VP information, see results page 61). Consistent with the genetic analysis, combined treatment of embryos from the mRFP-TEE transgenic line with both inhibitors strongly reduced TEE activity and endogenous CDX2 expression (Figure 29) confirming the genetic analysis results.

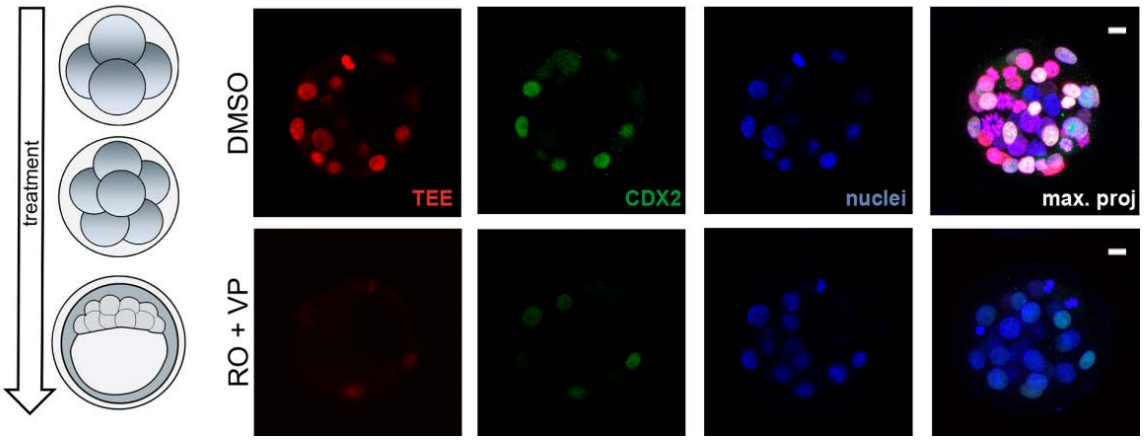


Figure 29. TEE activity is reduced by VP + RO treatment.

TEE activity (red) and CDX2 immunodetection (green) in embryos from the mRFP line treated with DMSO or RO+VP from 4-cell stage until blastocyst stage (left panel). Maximal projections of merged images are shown in the right panels. Nuclei were stained with DAPI. Scale bars, 10 µm.

Our results hence show that the Notch and Hippo pathways act together in the preimplantation embryo, affecting TEE activity, CDX2 expression and embryo viability.

7. Analysis of a Notch gain of function model.

To further characterize the role of Notch in trophectoderm specification and to complement the analysis of Notch pathway loss-of-function, we used a gain-of-function model to examine the effects of forcing Notch expression in atypical spatial and temporal patterns at the stage of the first lineage choice. For this, we overexpressed the active form of NOTCH1 in the preimplantation embryo by crossing a R26-stop-N1ICD-ires-nEGFP line (*Rosa*^{Notch}) (Murtaugh et al., 2003) with a line carrying a maternal Sox2-Cre allele (Hayashi et al., 2002) (Figure 30A). In the *Rosa*^{Notch} line, in the absence of Cre recombinase, transcription of the modified locus is blocked by the STOP sequence; in the presence of Cre, the Neo/STOP cassette will be deleted, yielding heritable, constitutive coexpression of Notch1 intracellular domain (N1ICD) and nuclear enhanced GFP (nEGFP). Targeting of the permissive *Rosa26* locus using the *Rosa26* endogenous promoter and a STOP fragment flanked by loxP sites generates a ubiquitous yet moderate level of expression upon loxP recombination (Srinivas et al.,

2001). The *Sox2*-Cre allele has been shown to be an efficient epiblast driver at 6.5 dpc and maternal Sox2-Cre recombination can induce full germline recombination (Hayashi et al., 2002).

7.1 *Cdx2* expression analysis of *Rosa*^{Notch} overexpressing blastomeres.

To overexpress active Notch starting in the zygote, we crossed *Sox2*-Cre females with *Rosa*^{Notch} males so that the recombination would start upon fertilization. Most blastocysts obtained from this cross were viable (80%, n=58), as shown by the presence of a blastocoele and proper CDX2 expression in outer cells (Figure 31B). This indicates that embryos are able to develop until the blastocyst stage when Notch pathway is constitutively active.

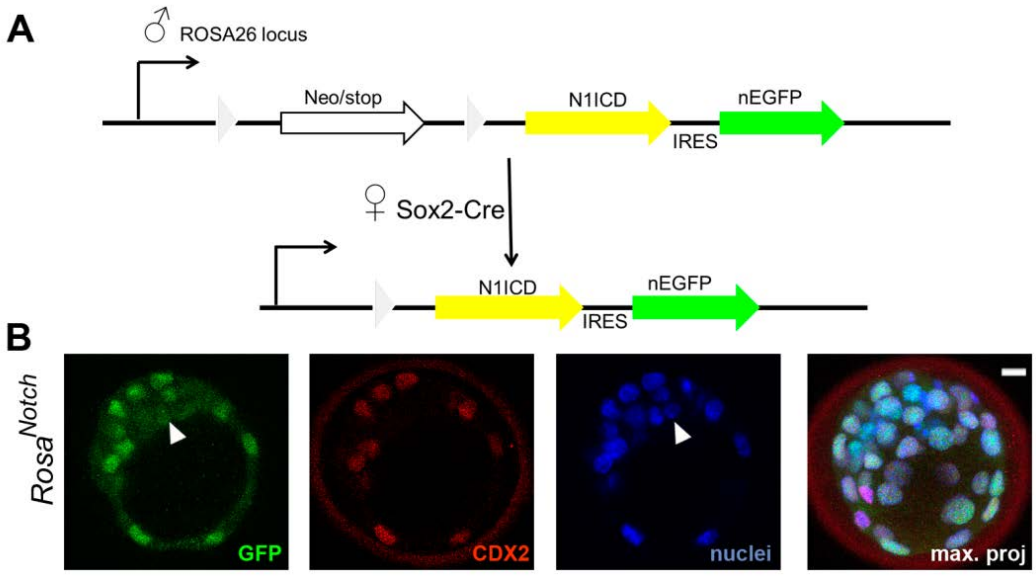


Figure 30. *Rosa* ^{Notch} gain of function model.

(A) Breeding strategy for the ♂R26-stop-N1ICD-ires-EGFP X ♀Sox2-Cre cross (B) Expression of the GFP reporter from the *Rosa* ^{Notch} line (green) and CDX2 (red) when recombined by maternal Sox2-Cre detected by immunochemistry. White arrowheads mark non-recombined cells. Nuclei were stained with DAPI. Scale bars, 10 μm.

As the *Sox2*-Cre is provided maternally, CRE is already expressed in the oocyte and recombination occurs right after the *Rosa* ^{Notch} allele is delivered. Although this strategy predicts uniform recombination in all cells of the blastocyst, we unexpectedly found mosaic expression of the reporter at 3.5-4.0 dpc (23 % GFP- blastomeres; arrowheads in Figure 31B, Figure 31B,C).The degree of mosaicism in the *Rosa* ^{Notch} embryos ranged from zero recombined cells to fully recombined blastocysts (Figure 31B). To check for mosaic recombination, we genotyped individual embryos for the presence or absence

of *Neomycin* (*Neo*). The Neo/STOP fragment flanked by loxP sites is recombined upon CRE expression and consequently a PCR product will only be detected when recombination has not occurred. We confirmed the presence of the non-recombined allele in blastocysts obtained from this cross by PCR (Figure 31A), indicating mosaicism. Moreover, to rule out the possibility that the mosaicism was due to undetectable GFP expression due to lower efficiency of GFP IRES-mediated translation compared with N1ICD translation, we used a R26-stop-YFP reporter (*Rosa* ^{YFP}) (Srinivas et al., 2001) that does not contain an IRES (internal ribosome entry site). In this case, we were also able to detect mosaic recombination when crossed to the maternal *Sox2*-Cre allele (Figure 31C).

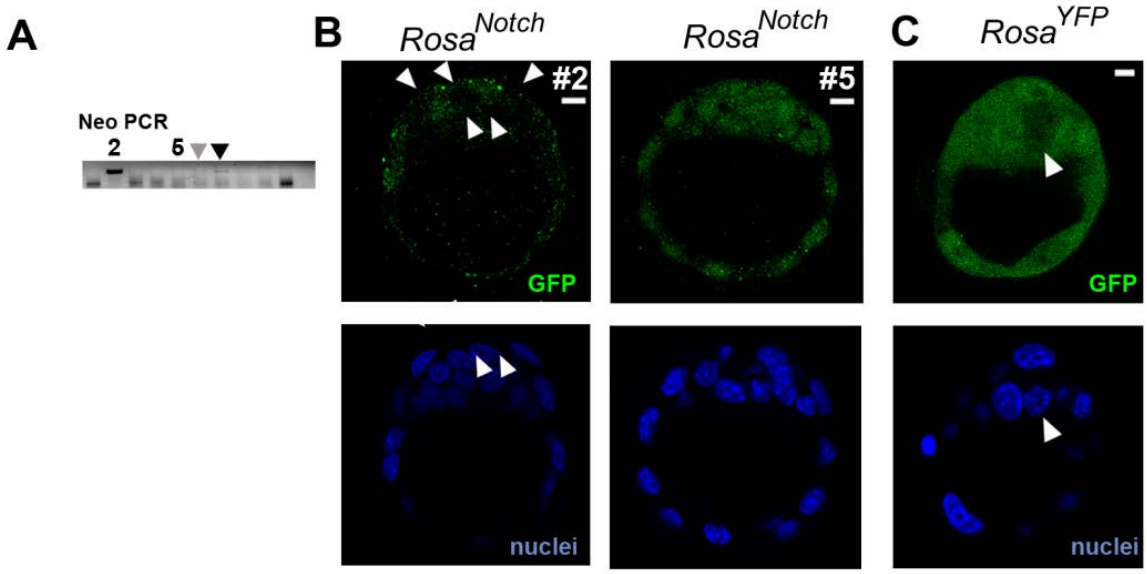


Figure 31. Mosaicism of maternal Sox2-Cre activity in the blastocyst.

(A) Detection of mosaic recombination in blastocysts by PCR of Neo. Arrowheads indicate the varying degree of detection of the non-recombined allele in different embryos, from low (grey) to high (black); embryo 2, and 5 are shown in (B). (B) Example of *Rosa* ^{Notch} blastocyst with a very high proportion of non-recombined cells (#2) and of a blastocyst with recombination occurring in all cells (#5). (C) Mosaicism of GFP reporter expression (green) from the *Rosa* ^{YFP} line when recombined by maternal Sox2-Cre. White arrowheads mark non-recombined cells. Nuclei in were stained with DAPI. Scale bars, 10 μm.

The fact that embryos contained N1ICD-overexpressing, GFP-labelled blastomeres confronted to blastomeres expressing endogenous levels of the pathway prompted us to analyse whether GFP+ and GFP- populations differ in their levels of *Cdx2* expression. This allowed us to directly interrogate if Notch overexpression increases *Cdx2* endogenous levels. We first plotted the mean CDX2 intensity for non-overexpressing

(N1OE-) and N1ICD-overexpressing (N1OE+) blastomeres. We included blastomeres in the ICM, which would be expected to show low levels of CDX2, and TE blastomeres in outer positions that would be expected to have higher levels of CDX2, independently of NOTCH1 overexpression. Quantification of CDX2 staining at the blastocyst stage showed that blastomeres overexpressing N1ICD have higher levels of CDX2 protein irrespective of their position (Figure 32A). We then analysed only outer cells (those that would normally express *Cdx2*) and the same pattern was found (Figure 32B), showing that in a population where CDX2 levels are already high, Notch overexpression induces even higher levels of expression. These results complement the earlier observation of reduced CDX2 in *Rbpj* mutant embryos (Figure 28), and confirm that the Notch pathway directly regulates *Cdx2* expression.

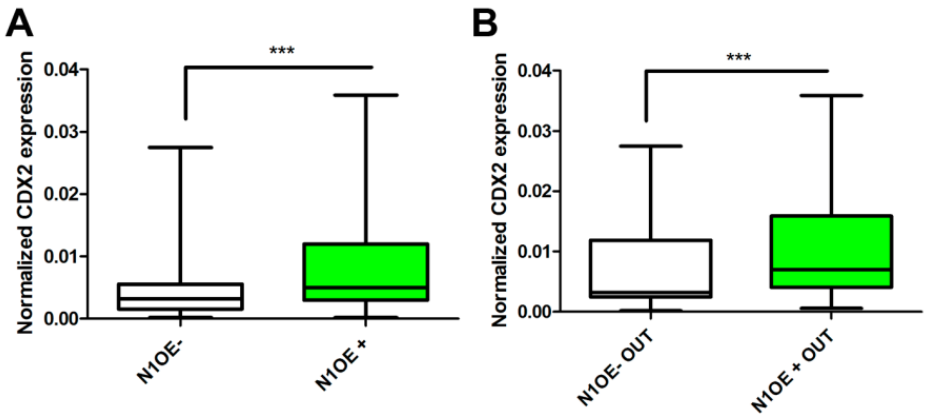


Figure 32. Notch overexpression regulates *Cdx2* expression.

(A) Quantified CDX2 expression in all N1ICD-overexpressing (N1OE+; green) and non-overexpressing (N1OE-) blastomeres in embryos generated from the ♂R26-stop-N1ICD-ires-EGFP X ♀Sox2-Cre cross. (B) Quantified CDX2 expression in outer N1OE+ and N1OE- blastomeres. Boxes span the 25th to the 75th percentile, internal horizontal lines indicate median values, and whiskers show minima and maxima. ***p<0.001 by Student's t-test.

7.2 *Rosa* ^{Notch} overexpressing blastomeres allocate preferentially in the TE of the blastocyst.

In this set of experiments, we also noticed that the TE of *Rosa* ^{Notch} embryos contained more blastomeres than wild-type embryos and it appeared that GFP+ blastomeres were more often localized to the TE. We took advantage of the mosaicism to analyse the spatial distribution of blastomeres overexpressing N1ICD in more detail. We compared wild-type and *Rosa* ^{Notch} embryos that did not show any difference in total cell number (Figure 33A) and quantified wild-type, N1OE- and N1OE+ blastomeres allocated to the outside or inside as a percentage of all blastomeres. Whereas GFP-

(N1OE-) blastomeres from mosaic embryos showed an inside-outside distribution similar to that of blastomeres from wild-type blastocysts, GFP+ (N1OE+) blastomeres were significantly more abundant in outside positions (Figure 33B). These results suggest that N1ICD-overexpressing blastomeres not only express higher levels of CDX2 but also have a higher probability of being allocated to the TE.

We next wanted to understand the distribution of N1OE+ versus N1OE- blastomeres per embryo. In wild-type embryos, we found that around 65% of the blastomeres are located outside. In contrast, in *Rosa* ^{Notch} embryos, the percentage was significantly higher (75%), confirming that *Rosa* ^{Notch} embryos have more outer, TE cells. Comparison of the proportions of outer and inner GFP+ cells in the *Rosa* ^{Notch} embryos showed that, N1OE+ cells were overrepresented in the TE population compared with inner cells (81.3% vs 64%; Figure 33C). These results indicate that blastomeres in which the Notch pathway is active localize preferentially to the outer TE population, and that cells can be directed to this population by forced overexpression of the active form of NOTCH1.

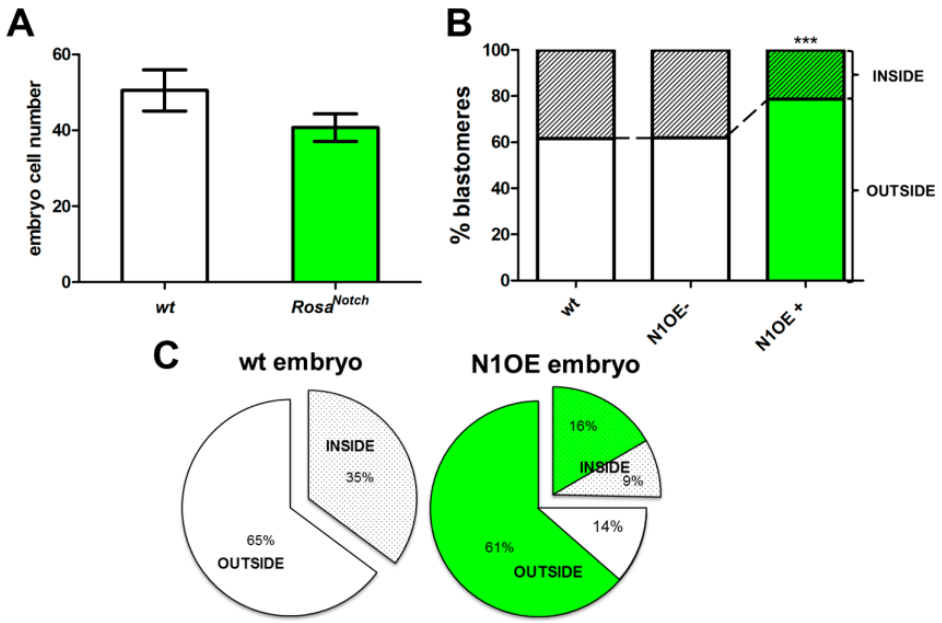


Figure 33. Notch instructs cells to adopt an outer position in the blastocyst.

(A) Average cell number in wild-type (wt) and *Rosa*^{Notch} blastocysts. (B) Inside/outside distribution of all N1OE+ and N1OE- blastomeres compared with the distribution of wild type (wt) blastomeres. ***p<0.001 by Chi squared test. (C) Distribution per embryo of inner (grey) and outer cells (white) in wild-type (wt) and N1ICD-overexpressing embryos. For N1OE embryos, the contribution of GFP+ cells (green) to each population is also shown.

In this work, we have been able to identify for the first time a *Cdx2* TE-specific enhancer and by generating transgenic mouse lines, we have been able to follow the dynamics of activation of the cis-regulatory sequence during the early stages of mammalian development and in TE-derived TS cells. Moreover, we have attempted to define how this element is regulated in order to better understand *Cdx2* regulation. By analysing known and unknown inputs upstream of *Cdx2*, we have found that Notch is a previously unidentified upstream regulator of *Cdx2*. Analysis of Notch loss- and gain-of-function models has revealed a novel role for the Notch signalling pathway during early preimplantation development.

"If it was so, it might be; and if it were so, it would be; but as it isn't, it ain't. That's logic."

Through the Looking-Glass. LEWIS CARROLL

1. Two-stage regulation model for *Cdx2.*

The mammalian zygote possesses the intrinsic capacity to initiate the process leading to differentiation and formation of an adult organism without any environmental cue. This has been proven in experiments where a single blastomere was isolated from the embryo. In these surgical experiments, the remaining blastomeres in the embryo compensated the loss and continued to develop normally. Moreover, the isolated blastomere was able to give rise to a wide variety of cell types. The ability of embryonic cells to change their fates to compensate for the missing parts during embryogenesis has been termed regulative development. The commitment of a cell depends upon the conditions in which the cell finds itself and there is no pre-patterning. Crosstalk within cells in the embryo is thus crucial for proper embryo progression.

Despite knowledge gained in the past few years, we still lack a precise understanding of the mechanisms that describe how pattern and lineages emerge in the early mouse embryo. We have a good understanding of the core set of transcription factors responsible for the establishment and maintenance of the first lineages (Cockburn and Rossant, 2010; Nishioka et al., 2009; Rossant and Tam, 2009), but little is known about how the expression of these factors themselves is initiated and then limited to specific subpopulations of cells in the embryo. Lineage restriction of genes such as *Cdx2*, a key factor in the TE lineage, is thought to be triggered by a combination of stochastic early expression together with morphological cues provided by the process of compaction and polarization of outside cells.

Stochastic expression in preimplantation development has been described as the capacity of single cells to express a random combination of key genes involved in the first lineage choice (*Cdx2*, *Oct4*, *Sox2* and *Nanog*). Immunohistochemistry on fixed embryos throughout development showed heterogeneous expression of the pluripotency regulatory networks, revealing no direct correlation among them (Dietrich and Hiiragi, 2007). Although this approach provides an interesting insight into the mechanisms that underlie the expression pattern of the factors, it still relies on the availability and sensitivity of the antibodies used. Stochastic expression has also been confirmed by single blastomere transcriptome profiling at early preimplantation stages. RNA profiling revealed that the core factors that subsequently become cell-type restricted in the blastocyst are co-expressed at high

levels in most blastomeres of the 16-cell embryo, and it is not until the 32-cell stage that cells are assigned to a TE or ICM lineage (Guo et al., 2010). The heterogeneous expression of the lineage transcription factors confers plasticity to the regulative embryo. The fate choice is taken only when the factors are stably maintained in cells.

An issue that remains unresolved is the *in vivo* dynamics of each transcription factor, in other words whether the apparent heterogeneity of the transcription factors at early stages arises from their random upregulation and downregulation within a blastomere. Although the study of OCT4 kinetics has revealed an important role in lineage patterning (Plachta et al., 2011), there is still a scarcity of tools with which to track transcription factor dynamics within embryos.

The identification of non-coding regulatory sequences for the main transcription factors would permit a better understanding of the dynamic expression profile of this set of genes and may serve as an approach to decipher the heterogeneity of each transcription factor within a blastomere. Also, the characterization of the minimal enhancer sequence responsible for full regulatory activity allows interrogation of the sequence and identification of putative binding sites for factors that may link the heterogeneity we see in the expression pattern with morphological and topological cues in the embryo. There have been some reports identifying enhancers for the core pluripotency factors *Oct4 Sox2* and *Nanog* (Catena et al., 2004; Jiang et al., 2008; Levasseur et al., 2008; Tomioka et al., 2002; Yeom et al., 1996; Yoshimizu et al., 1999). However, most of the work has been performed on ES cells and their capacity to direct expression in preimplantation development remains to be characterized. Two different enhancers have been identified for *Oct4* (Yeom et al., 1996; Yoshimizu et al., 1999). Strikingly, only one of them is able to direct reporter expression to the ICM of the blastocyst (the distal enhancer, *Oct4*DE). In contrast, the other enhancer is able to direct reporter expression in the epiblast and ES cells but not in the blastocyst. This suggests that the regulatory mechanisms that operate *in vitro* may not be equivalent to the mechanisms that operate *in vivo*.

To better understand how the combinations of inputs affect the expression of the core set of transcription factors, we dissected the regulatory landscape of *Cdx2*. There have been a number of reports addressing *Cdx2* regulation, but none of them were able to identify a TE-specific regulatory sequence (Benahmed et al., 2008; Gaunt et al., 2005; Wang and Shashikant, 2007). Here, we characterize the TEE, an enhancer that drives specific expression in the TE and that is expressed early during preimplantation development. Not only is the TEE found to be TE specific, but TEE mouse reporter lines also closely match the early onset of endogenous *Cdx2*

expression, highlighting its capacity to integrate resident regulatory signals. We cannot rule out an unspecific behaviour of the mouse reporter line due to positional integration effects of the transgene; however, the number of TEE mouse lines generated showing similar behaviour indicate that TEE reporter lines can be used as a read out of early events during the first lineage choice.

Surprisingly, when we analysed TEE dynamics in TE-derived tissues, we found that it is not able to drive expression in the ExE of the post-implantation embryo, where *Cdx2* is robustly expressed. The lack of TEE activity in the ExE led us to carefully analyse *Cdx2* regulation in TS cells. In this context, the TEE cannot drive reporter expression when it is transfected into wild-type TS cells, and neither is it able to do so in TS_L and TS_R cell cultures, in which no reporter expression was detected.

Histone mark profiling of the *Cdx2* locus in wild-type TS cells reveals that the endogenous TEE is depleted of active marks and, in contrast, a region located in the first intron of *Cdx2* shows positive regulatory capacity (Fragment II), suggesting that this sequence might be responsible for *Cdx2* regulation in TS cells. This element is able to drive reporter expression in TS and ES cells, confirming its enhancer potential. Additionally, this fragment was previously reported in the lab to drive reporter activity in the blastocyst in transient transgenic assays in both ICM and TE. Nonetheless, its role in extraembryonic regulation of *Cdx2* has yet to be characterized as it has only been shown to direct expression to mesoderm and neural tube in the post-implantation embryo (Wang and Shashikant, 2007). Preliminary data in the lab has shown that the fragment is able to direct extraembryonic mesoderm expression (data not shown), thus confirming its ability to drive expression to ICM-derived tissues.

Altogether, these results indicate that the TEE holds TE-restricted capacity to drive reporter expression, whereas Fragment II activity is ubiquitous in the embryo. The fact that Fragment II retains activity in the whole blastocyst needs to be further investigated. It could be explained if inputs responsible for its regulatory ability appear equally in both tissues, leading to a general activation of the cis regulatory sequence, or if the DNA region is competent to respond to distinct inputs in different contexts.

Here, we extensively tested the regulatory capacity of discrete cis-regulatory regions. In this way have we been able to identify regulatory elements that underlie *Cdx2* regulation, but the regulatory mechanisms behind the regulatory landscape-switch between TE and ExE remain unknown. The combinatory activity of several regulatory sequences may explain *Cdx2* expression in the post-implantation

embryo. Generation of constructs that link TEE and Fragment II sequences will help us to study if we reproduce *Cdx2* extraembryonic expression pattern completely or, in contrast, there are other cis-regulatory elements that remain to be further identified (Figure 34).

During cleavage stages there is a transition from a stochastic expression of lineage factors to progressively restricted expression in their respective territories. In the first lineage choice, CDX2 is highly expressed in the outer population and defines the TE; in contrast, OCT*4*, SOX*2* and NANOG will persist in the inside population and specify the ICM. This process involves at least two clear mechanisms to resolve the first lineage decision: the activation and reinforcement of the early expression of key transcription factors, followed by maintenance mechanisms that will ensure the lineage choice.

The ability of the TEE to drive reporter expression concomitantly with the activation of endogenous *Cdx2* expression indicates that the TEE sequence is responsible for receiving signals that activate *Cdx2* at early stages. Not only does the TEE activity disappear once the TE starts to differentiate, but the TEE is also inactive in TS cells, in which *Cdx2* is responsible for the self-renewal of the stem cell population (Strumpf et al., 2005). This suggests that the TEE is not involved in the maintenance of *Cdx2* expression in TS-derived tissues or in the self-renewing capacity of TS cells.

It is tempting to propose a two-stage regulatory model for *Cdx2*. In this model, the earlier establishment of *Cdx2* expression would be mediated by the TEE, by reading stochastic and topological cues in the first lineage choice; in contrast, *Cdx2* expression at later stages would be maintained by other regulatory sequences active both in TS cells and in the post-implantation embryo, such as fragment II (Figure 34).

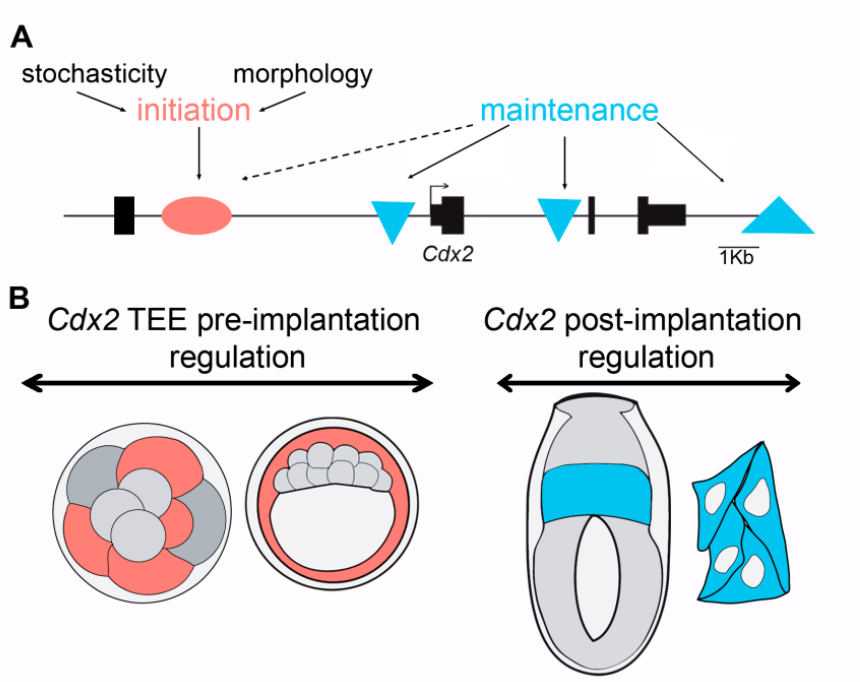


Figure 34. Two-stage regulation model of *Cdx2*.

(A) Diagram of the *Cdx2* locus showing cis-regulatory elements responsible for the initiation of *Cdx2* expression (TEE, red circle) and later *Cdx2* maintenance (blue triangles). Stochastic and morphology signals are integrated through the TEE whereas the regulatory element involved in *Cdx2* maintenance remains to be unidentified. The dashed line denotes the possible action of the TEE as a cis-regulatory element in combination with other sequences in later stages. (B) Proposed model for *Cdx2* regulation in extraembryonic lineages. In early preimplantation development, the TEE regulates initial *Cdx2* heterogeneous expression and restriction to the TE (in red). Upon implantation, *Cdx2* regulation shifts to different cis-regulatory elements that maintain CDX2 expression in the ExE or TS cells.

2. Notch and Hippo converge on *Cdx2* regulation through the TEE.

The Hippo signalling pathway has emerged as a fundamental player in the first lineage decision. This pathway translates topological cues related to compaction (cell adhesion and polarization) into a transcriptional response, its activation state differing between the prospective TE and the ICM according to the polarization of outer cells (Anani et al., 2014; Cockburn et al., 2013; Hirate et al., 2013). In the compact eight-cell stage embryo, YAP is in the nucleus and is able to interact with its transcriptional partner TEAD4, which is expressed in every cell. During the successive divisions, YAP is phosphorylated through LATS1/2 kinases and retained in the cytoplasm of inner cells, preventing its transcriptional activity. The

transcriptional mediator of Hippo signalling, TEAD4, has been characterized as an upstream regulator of *Cdx2* at these early stages.

Although it has been shown that the Hippo pathway regulates *Cdx2* expression, we were able to detect occasionally *Cdx2* expressed in *Tead4* mutants, similar to what others have reported (Kaneko and DePamphilis, 2013; Nishioka et al., 2009). Also, the finding that initiation of *Cdx2* expression is independent of compaction (Stephenson et al., 2010; Wu et al., 2010) and cell adhesion (Lorthongpanich et al., 2013) suggests that further inputs are involved in *Cdx2* regulation. We find that the TEE enhancer is still active and restricted to outer cells in *Tead4* knockout embryos, confirming that other unknown regulatory mechanisms must regulate *Cdx2* trough the TEE.

In an effort to dissect how the Hippo pathway controls the TEE, we also tested if forced activation of the Hippo pathway in outer cells silenced the TEE. Forced expression of the Hippo pathway induces YAP phosphorylation and thus prevents it from shuffling to the nucleus to activate target genes. For these experiments, embryos of the mRFP-TEE line were injected with *Lats2* RNA, which turns the Hippo pathway on and results in decreased expression of CDX2 in outer cells (Nishioka et al., 2009). This did not influence the expression of the reporter in the TE, confirming that the TEE does not respond exclusively to the activation state of the Hippo pathway in the blastocyst (results not shown).

In the search for novel upstream regulators of *Cdx2*, dissection of the TEE sequence identified Notch as a candidate signalling cascade for *Cdx2* regulation. Interestingly, as was the case for the Hippo pathway (Kaneko et al., 2007; Nishioka et al., 2008), we found that CDX2 levels are decreased in *Rbpj* mutants but the lack of Notch is not sufficient to turn TEE activity off. The requirements for Hippo and Notch pathways in the early embryo are clearly distinct, since *Tead4* null embryos arrest development whereas *Rbpj* knockouts live up to 9.5dpc. Nonetheless, the relevance of the Notch pathway in cooperation with *Tead4* is demonstrated by the finding that *Rbpj* and *Tead4* double heterozygote mice have low CDX2 expression and compromised TEE capacity. Importantly, no double homozygous mutants for *Tead4* and *Rbpj* were detected. Compound mutants for *Rbpj* and *Tead4* have below-normal CDX2 expression levels per blastomere but do not show differences in cell number, suggesting that additional regulatory inputs act on *Cdx2* in the blastocyst. A possible mechanism is auto-regulation of *Cdx2* (Barros et al., 2011; Cockburn and Rossant, 2010; Niwa et al., 2005), which would maintain and stabilize *Cdx2* expression levels.

Analysis of the response of the TEE to Hippo and Notch showed that RBPJ and TEAD compensate for each other, since TEE activity is reduced only when the binding sites for one pathway have been mutated and the converse pathway is inhibited, or when binding sites for both are mutated. Thus, Notch and Hippo operate such that the TEE is capable of driving reporter expression if at least one of its upstream regulators is on. With this finding, not only have we identified Notch as a novel upstream regulator of *Cdx2*, but we also show the convergence of Notch and Hippo on a specific regulatory sequence. The structure of the *Cdx2* TE-enhancer described here can be used to screen for cis-regulatory elements with a similar organization of RBPJ and TEAD sites and thereby identify other components of the TE-specification gene regulatory network.

Previous reports of Notch and Hippo interaction have shown that one pathway is downstream of the other in *Drosophila* (Chen et al., 2011; Graves et al., 2012) and in liver development and cancer (Tschaharganeh et al., 2013; Yimlamai et al., 2014). It will be important to study if there is any other crosstalk of the pathways during pre-implantation development; in this regard, it is interesting that in *Drosophila* wing epithelial cells, inactivation of Hippo signalling results in apical accumulation of the Notch receptor (Genevet et al., 2009). Moreover, it should be noted that phenotypes of mutations of Notch pathway components are sensitive to gene dosage (Murtomaki et al., 2013; Nus et al., 2011) (Boyer-Di Ponio et al., 2007).

The results of this study show that the Hippo and Notch pathways act through the TEE and converge to activate the expression of *Cdx2* during the first lineage specification in the blastocyst. In the developing TE, the Hippo pathway is off and YAP translocates to the nucleus and binds to TEAD4 to activate target genes. At the same time, the Notch signalling pathway is active in the TE, and N1ICD is therefore in the nucleus bound to RBPJ and activating *Cdx2* expression (Figure 35).

Consistent with what we found in the blastocyst, N1ICD is heterogeneously expressed in TS cells, whereas we do not detect N1ICD expression in ES cells. In striking contrast, ES cells derived from the H2B:CBF-Venus reporter line still express the reporter (Nowotschin et al., 2013). We did not assess the presence of other Notch receptors so we cannot rule out the possibility that these (Notch2-4) are active in ES cells. Additionally, persistent expression of the reporter in ES cells may reflect the fact that ES cells are not equivalent to ICM cells, and in fact represent epiblast cells from 4.5 dpc blastocysts (Boroviak et al., 2014).

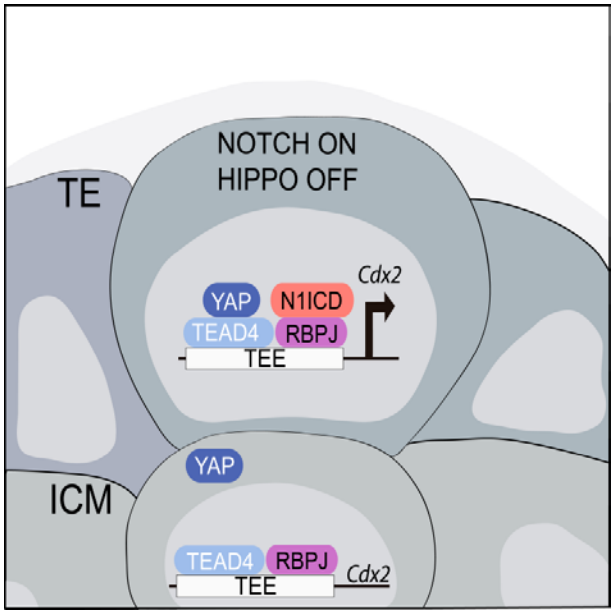


Figure 35. Notch and Hippo converge on *Cdx2* to specify trophectoderm lineage.

Representation of the signalling pathways that act through the TEE to regulate *Cdx2*. In TE cells, the Hippo pathway is OFF and thus unable to retain YAP in the cytoplasm, and the Notch pathway is ON. YAP coactivator and NOTCH1 intracellular domain (N1ICD) respectively bind to TEAD4 and RBPJ to activate *Cdx2* transcription.

Overexpression of TE factors (*Cdx2*, *Gata3*, *Ras, Tead4*) in ES cells transdifferentiates them to TS cells (Lu et al., 2008; Nishioka et al., 2009; Niwa et al., 2005; Ralston et al., 2010). However, inhibition of Hippo signalling in ES cells, while resulting in a loss of pluripotency, does not appear to convert ES cells into a TE fate, but rather promotes differentiation to endodermal (Lian et al., 2010; Tamm et al., 2011). As ES cells do not reproduce the blastocyst environment, it is conceivable that YAP responds in a different manner. Recently, cues from the extracellular matrix, cell adhesion sites, cell shape and the actomyosin cytoskeleton have been found to converge on the regulation of YAP (Halder et al., 2012). Notch over-activation in ES cells does not induce any phenotypic change; it is only upon withdrawal of self-renewal stimuli that Notch overexpression induces a rapid and exclusive differentiation into the neural lineage (Lowell et al., 2006). This does not contradict our findings, but rather indicates that Notch operates differently in ES cells versus the ICM, as is also the case for the Hippo pathway.

3. Notch signalling in the first lineage choice.

During the first fate choice, there is a crucial shift from a random expression of transcription factors to a coordinated response to morphological cues provided by the physical properties of the embryo. At the time of compaction, interconnected features such as polarization, cell adhesion and cell to cell interaction are integrated into the modulation of the transcriptional program of single blastomeres (Cockburn and Rossant, 2010; Hirate et al., 2013). What is really intriguing is that at the transcriptional level, not all blastomeres exposed to near-identical cues (for instance, blastomeres in outside positions) respond equally, while blastomeres exposed to different morphological cues (for example outside versus inner blastomeres) sometimes do, as shown by the stochastic expression of the core transcription factors (Dietrich and Hiiragi, 2007; Ralston and Rossant, 2008). This observation is also supported by the fact that outer cells as early as the 16-cell stage are a heterogeneous population, with some fated to contribute exclusively to the TE and others able to form both the TE and ICM (Anani et al., 2014; Watanabe et al., 2014). In this context, the adoption of either TE or ICM programs is a regulative decision, and the final output of key transcription factors may rely on subtle differences within blastomeres. It is thus vital to characterize adjusting mechanisms by which key transcription factors respond to single inputs but modulate their expression coordinately. In this complex framework, understanding of compensation, cooperation and redundancy among signalling pathways will help us to better dissect the first lineage choice.

Cdx2 exhibits two distinct patterning phases that eventually lead to its TE-restricted expression. In the first phase, molecular differences between blastomeres are generated (initiation of *Cdx2* expression and establishment of variability among blastomeres), and in the second phase, molecular signature responds to cell allocation within the embryo (Dietrich and Hiiragi, 2007). It has become clear that the second patterning phase of *Cdx2* relies on the Hippo pathway, which connects positioning and transcription factors (Cockburn and Rossant, 2010; Hirate et al., 2013). We have identified the Notch pathway as a signalling mechanism that modulates TE specification, and we show that Notch not only subtly alters *Cdx2* levels, but also influences the fate of blastomeres by driving them to outer positions in the embryo when overexpressed. It could be argued that Notch simply elevates *Cdx2* levels and that this is the trigger for cell relocation (Jedrusik et al., 2008). However, the fact that blastomeres mutant for *Cdx2* can adopt an outer position argues against this possibility (Ralston and Rossant, 2008). Hence, we propose that the Notch pathway could be involved in the initiation of *Cdx2* expression. In this

way, Notch and Hippo would cooperate in *Cdx2* regulation in an overlapping temporal framework in which Notch would introduce the first random transcriptional expression among cells and Hippo would be responsible for fixing the levels according to the cells' molecular signature (positional and polarization state) (Figure 36). The fact that Notch pathway mouse mutants (*Rbpj*, *Notch1-4* receptors and ligands) do not show a severe preimplantation phenotype may indicate compensatory roles among this family of receptors. Nevertheless, two reports have described deleterious defects during early development in the Notch pathway components *Notchless* (Cormier et al., 2006) and *Brainiac* (Vollrath et al., 2001), supporting a role for this signalling pathway.

Notch signalling can operate through two possible mechanisms: lateral inhibition, which generates heterogeneity from an otherwise homogenous group of cells (Artavanis-Tsakonas et al., 1999), and lateral induction, which achieves local homogeneity of cells expressing receptors and ligands and enables cells to create a sharp gene expression boundary where no such discontinuity existed before (Lewis, 1998) (Figure 7).

Lateral inhibition operates such that the signalling cell and the adjacent receiving cell induce opposite transcriptional effects. Among apparently equivalent neighbours expressing both ligand and receptor, a small increase in ligand in one cell relative to its neighbour could favour its adoption of the signalling role. The signalling cell induces Notch activation on the adjacent cell. As a result, Notch reduces ligand expression in the receiving cell, preventing the activation of Notch in the signalling cell and thus generating an opposite response. Feedback then amplifies and consolidates the differences between receptor and ligand. In this way, the adjacent cells will manifest complementary expression patterns that lead to the generation of distinct populations in an initially identical field. In the early embryo, the stochastic expression of *Cdx2*, *Sox2* and *Nanog* precedes the first lineage choice and there is still not a good understanding of how this noisy system starts. Lateral inhibition could drive the first wave of stochastic *Cdx2* expression in a polarization- and adhesion-independent manner. In this context, random expression of Notch would elevate *Cdx2* expression accordingly and downregulate the expression of Notch ligands within the cell. Cell-to-cell interactions through lateral inhibition instruct the neighbouring cell to have the opposite effect and thus generate the first heterogeneities in an equivalent population (Figure 36). In this scenario, Notch ligand and receptor distributions should be complementary, and it will also be interesting to analyse the number of CDX2+ adjacent cells and how CDX2 correlates with early N1ICD expression.

If lateral induction functions in the blastocyst, it would lead to cells of the TE acting as a coherent group, in other words a developmental field. Sustained activity of the pathway in outer cells ensures that they maintain a TE fate, allowing for correct specification of blastomeres that reposition during cell division and normalizing the early stochastic expression pattern –underpinning Hippo pathway function. Validation of lateral induction requires testing of whether Notch itself inhibits or induces expression of the Notch-activating ligand. So far, we have identified Dll1 expression in cells of the trophectoderm at the blastocyst stage (results not shown), but other components need to be further investigated. The observed co-expression of the activated Nocth1 receptor and the ligand Dll1 in cells of the trophectoderm suggests that Notch acts in the blastocyst through a mechanism of lateral induction rather than through lateral inhibition, although no mechanism can yet be ruled out (Figure 36).

As previously mentioned, Notch receptor activity relies on cell-to-cell signalling interaction through its ligands. In our study, we have limited our analysis to Notch function in the blastocyst. It will be interesting to assess the requirement for Notch pathway in the morula, during the first wave of expression of core transcription factors, to test for a Notch-mediated TE-specification model. In this model, lateral inhibition is responsible for the early stochastic expression of transcription factors, and later the stabilization of the lineages occurs as a result of the convergence of Hippo and Notch signalling through lateral induction (Figure 36). If this is the case, we predict that Notch mutants will not show *Cdx2* stochastic expression early on but that the reinforcement signals through Hippo would still raise *Cdx2* levels so that the regulative embryo develops. This view fits with the observation that *Cdx2* levels were reduced but still enough to specify the TE in *Rbpj* mutants. The model could also explain why the double *Tead4*; *Rbpj* mutant embryos die: they would lack both the initiation of random *Cdx2* expression and the reinforcement mechanisms needed to stabilize *Cdx2* expression. A clearer understanding of whether Notch has an overall effect on the TE transcriptional program could be gained from genome-wide RNA analysis of different Notch mutants and gain-of-function strains. Also, inducible Cre/loxP strategies that allow us to generate mosaic expression of Notch receptors and ligands will help to unravel the role of Notch in TE specification.

If we consider the embryo as a self-organizing system (Wennekamp et al., 2013), lineage establishment is initiated, corrected and refined by a combination of intrinsic and extrinsic cues. In this scenario, the robustness of TE establishment is conferred through the initiation and fine tuning of combinatorial inputs that drive the expression of TE-lineage markers such as *Cdx2.* The overlapping activities of

Notch and Hippo pathways are likely integrated via regulatory elements such as the TEE, and could underlie the regulatory capacities of the mammalian embryo and the initial stochastic lineage-specific transcriptional programs. Future work will provide information about how modulation of Notch together with other signalling pathways leads to the differentiation of the totipotent cell, and how symmetry is first broken during mammalian development.

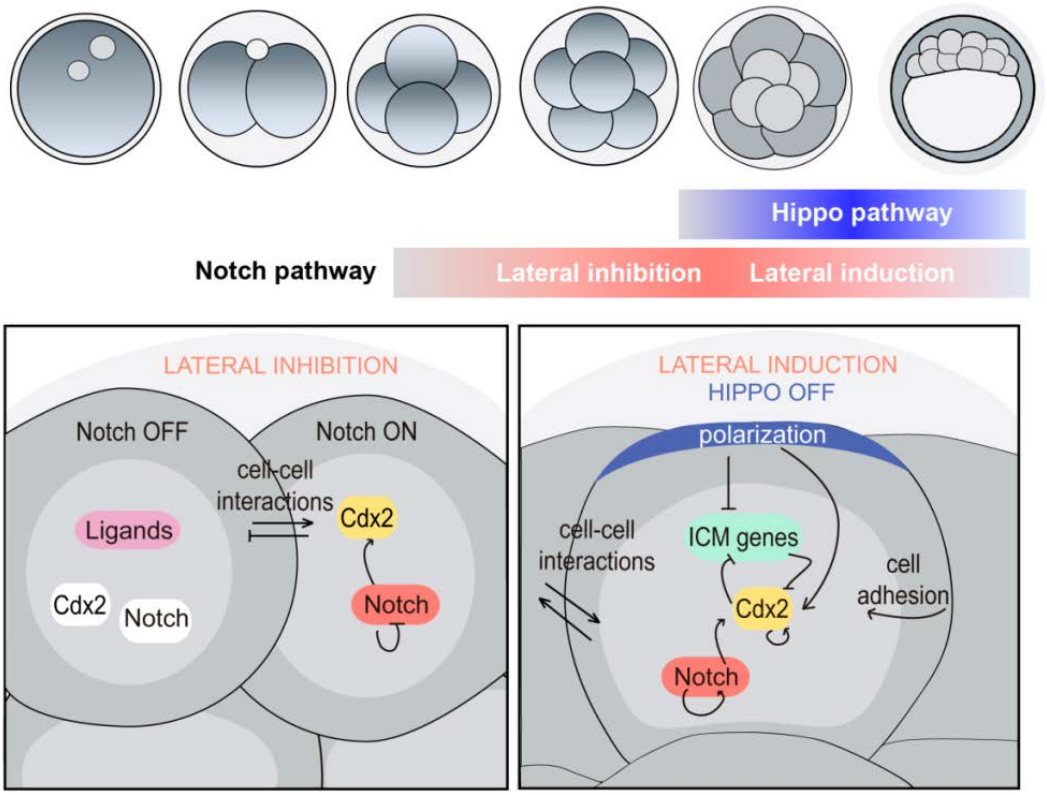


Figure 36. Temporal control of TE specification through the Notch and Hippo pathways.

Proposed temporal patterning phases for TE specification. In the early phase, the Notch pathway generates molecular differences between blastomeres and upregulates *Cdx2* through lateral inhibition. In the second phase, the Hippo pathway translates topological cues related to compaction into a transcriptional response, while Notch reinforces *Cdx2* expression in outer cells via lateral induction.

"And at last we've got to the end of this ideal racecourse!
Now that you accept A and B and C and D, of course you accept Z."
What the Tortoise Said to Achilles. LEWIS CARROLL.

CONCLUSIONS

- We have identified a 1.3 kb trophectoderm-specific enhancer (TEE) located 3 kb upstream of the mouse *Cdx2* gene.
- Expression of reporter genes driven by the TEE reproduce the early activation of *Cdx2* expression at the morula stage and trophectoderm restricted expression in the blastocyst.
- The TEE does not drive expression in trophectoderm-derived tissues of the post-implantation embryo or in trophoblast stem cells, suggesting a two-stage regulatory mechanism for *Cdx2* expression in the early embryo.
- The TEE does not respond only to *Tead4*, indicating that other regulatory inputs are responsible for *Cdx2* early expression.
- The Notch pathway directly regulates *Cdx2* expression in combination with *Tead4* through the TEE.
- Activity of the Notch pathway is restricted to the trophectoderm of the blastocyst.
- The Notch pathway, together with *Tead4*, is essential for preimplantation development.
- Notch overexpression influences the position of cells in the blastocyst and drives them to preferentially allocate to the trophectoderm.

CONCLUSIONES

- Hemos identificado un *enhancer* de 1,3 kilobases específico de trofectodermo (el TEE), localizado 3 kilobases por delante del gen *Cdx2*.
- Las expresión de genes reporteros dirigida por el TEE reproduce la activación temprana de la expresión de *Cdx2* en el estadio de morula y su expresión restringida en el trofectodermo del blastocisto.
- El TEE no dirige la expresión en tejidos derivados del trofectodermo en estadios de post-implantación ni en células troncales del trofoblasto. Esto sugiere un mecanismo de regulación en dos etapas para la expresión de *Cdx2* en el embrión temprano.
- El TEE no responde únicamente a *Tead4*, lo que indica que otras vías regulan la expresión temprana de *Cdx2*.
- La vía de Notch regula directamente la expresión de *Cdx2* junto con *Tead4* través del TEE.
- La actividad de la vía de Notch está restringida al trofectodermo del blastocisto.
- La vía de Notch, junto con *Tead4*, es esencial para el desarrollo en preimplantación.
- La sobreexpresión de Notch influye en la posición de las células en el blastocisto y las orienta preferencialmente al trofectodermo.

BIBLIOGRAPHY

Abad, M., Mosteiro, L., Pantoja, C., Canamero, M., Rayon, T., Ors, I., Grana, O., Megias, D., Dominguez, O., Martinez, D.*, et al.* (2013). Reprogramming in vivo produces teratomas and iPS cells with totipotency features. Nature *502*, 340-345.

Adachi, K., Nikaido, I., Ohta, H., Ohtsuka, S., Ura, H., Kadota, M., Wakayama, T., Ueda, H.R., and Niwa, H. (2013). Context-dependent wiring of Sox2 regulatory networks for self-renewal of embryonic and trophoblast stem cells. Mol Cell *52*, 380-392.

Anani, S., Bhat, S., Honma-Yamanaka, N., Krawchuk, D., and Yamanaka, Y. (2014). Initiation of Hippo signalling is linked to polarity rather than to cell position in the pre-implantation mouse embryo. Development.

Anbanandam, A., Albarado, D.C., Nguyen, C.T., Halder, G., Gao, X., and Veeraraghavan, S. (2006). Insights into transcription enhancer factor 1 (TEF-1) activity from the solution structure of the TEA domain. Proc Natl Acad Sci U S A *103*, 17225-17230.

Andersson, E.R., Sandberg, R., and Lendahl, U. (2011). Notch signalling: simplicity in design, versatility in function. Development *138*, 3593-3612.

Artavanis-Tsakonas, S., and Muskavitch, M.A. (2010). Notch: the past, the present, and the future. Curr Top Dev Biol *92*, 1-29.

Artavanis-Tsakonas, S., Rand, M.D., and Lake, R.J. (1999). Notch signalling: cell fate control and signal integration in development. Science *284*, 770-776.

Auman, H.J., Nottoli, T., Lakiza, O., Winger, Q., Donaldson, S., and Williams, T. (2002). Transcription factor AP-2gamma is essential in the extra-embryonic lineages for early postimplantation development. Development *129*, 2733-2747.

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev *17*, 126-140.

Azuara, V. (2006). Profiling of DNA replication timing in unsynchronized cell populations. Nat Protoc *1*, 2171-2177.

Barros, R., da Costa, L.T., Pinto-de-Sousa, J., Duluc, I., Freund, J.N., David, L., and Almeida, R. (2011). CDX2 autoregulation in human intestinal metaplasia of the stomach: impact on the stability of the phenotype. Gut *60*, 290-298.

Beck, F., Erler, T., Russell, A., and James, R. (1995). Expression of Cdx-2 in the mouse embryo and placenta: possible role in patterning of the extra-embryonic membranes. Dev Dyn *204*, 219-227.

Beddington, R.S., and Robertson, E.J. (1989). An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. Development *105*, 733-737.

Benahmed, F., Gross, I., Gaunt, S.J., Beck, F., Jehan, F., Domon-Dell, C., Martin, E., Kedinger, M., Freund, J.N., and Duluc, I. (2008). Multiple regulatory regions control the complex expression pattern of the mouse Cdx2 homeobox gene. Gastroenterology *135*, 1238-1247, 1247 e1231-1233.

Bischoff, M., Parfitt, D.E., and Zernicka-Goetz, M. (2008). Formation of the embryonic-abembryonic axis of the mouse blastocyst: relationships between orientation of early cleavage divisions and pattern of symmetric/asymmetric divisions. Development *135*, 953-962.

Blij, S., Frum, T., Akyol, A., Fearon, E., and Ralston, A. (2012). Maternal Cdx2 is dispensable for mouse development. Development *139*, 3969-3972.

Boroviak, T., Loos, R., Bertone, P., Smith, A., and Nichols, J. (2014). The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. Nat Cell Biol *16*, 516-528.

Boyer-Di Ponio, J., Wright-Crosnier, C., Groyer-Picard, M.T., Driancourt, C., Beau, I., Hadchouel, M., and Meunier-Rotival, M. (2007). Biological function of mutant forms of JAGGED1 proteins in Alagille syndrome: inhibitory effect on Notch signalling. Hum Mol Genet *16*, 2683-2692.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G.*, et al.* (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947-956.

Calo, E., and Wysocka, J. (2013). Modification of enhancer chromatin: what, how, and why? Mol Cell *49*, 825-837.

Catena, R., Tiveron, C., Ronchi, A., Porta, S., Ferri, A., Tatangelo, L., Cavallaro, M., Favaro, R., Ottolenghi, S., Reinbold, R.*, et al.* (2004). Conserved POU binding DNA sites in the Sox2 upstream enhancer regulate gene expression in embryonic and neural stem cells. J Biol Chem *279*, 41846-41857.

Cockburn, K., Biechele, S., Garner, J., and Rossant, J. (2013). The Hippo pathway member Nf2 is required for inner cell mass specification. Curr Biol *23*, 1195-1201.

Cockburn, K., and Rossant, J. (2010). Making the blastocyst: lessons from the mouse. J Clin Invest *120*, 995-1003.

Cormier, S., Le Bras, S., Souilhol, C., Vandormael-Pournin, S., Durand, B., Babinet, C., Baldacci, P., and Cohen-Tannoudji, M. (2006). The murine ortholog of notchless, a direct regulator of the notch pathway in Drosophila melanogaster, is essential for survival of inner cell mass cells. Mol Cell Biol *26*, 3541-3549.

Cormier, S., Vandormael-Pournin, S., Babinet, C., and Cohen-Tannoudji, M. (2004). Developmental expression of the Notch signalling pathway genes during mouse preimplantation development. Gene Expr Patterns *4*, 713-717.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A.*, et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell *113*, 643-655.

Chawengsaksophak, K., James, R., Hammond, V.E., Kontgen, F., and Beck, F. (1997). Homeosis and intestinal tumours in Cdx2 mutant mice. Nature *386*, 84-87.

Chen, H.J., Wang, C.M., Wang, T.W., Liaw, G.J., Hsu, T.H., Lin, T.H., and Yu, J.Y. (2011). The Hippo pathway controls polar cell fate through Notch signalling during Drosophila oogenesis. Dev Biol *357*, 370-379.

Chew, J.L., Loh, Y.H., Zhang, W., Chen, X., Tam, W.L., Yeap, L.S., Li, P., Ang, Y.S., Lim, B., Robson, P.*, et al.* (2005). Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. Mol Cell Biol *25*, 6031-6046.

Dietrich, J.E., and Hiiragi, T. (2007). Stochastic patterning in the mouse pre-implantation embryo. Development *134*, 4219-4231.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154-156.

Fleming, T.P., and Pickering, S.J. (1985). Maturation and polarization of the endocytotic system in outside blastomeres during mouse preimplantation development. J Embryol Exp Morphol *89*, 175-208.

Frum, T., Halbisen, M.A., Wang, C., Amiri, H., Robson, P., and Ralston, A. (2013). Oct4 cell-autonomously promotes primitive endoderm development in the mouse blastocyst. Dev Cell *25*, 610-622.

Gasperowicz, M., and Otto, F. (2008). The notch signalling pathway in the development of the mouse placenta. Placenta *29*, 651-659.

Gaunt, S.J., Drage, D., and Trubshaw, R.C. (2005). cdx4/lacZ and cdx2/lacZ protein gradients formed by decay during gastrulation in the mouse. Int J Dev Biol *49*, 901-908.

Genevet, A., Polesello, C., Blight, K., Robertson, F., Collinson, L.M., Pichaud, F., and Tapon, N. (2009). The Hippo pathway regulates apical-domain size independently of its growth-control function. J Cell Sci *122*, 2360-2370.

Goode, S., and Perrimon, N. (1997). Brainiac and fringe are similar pioneer proteins that impart specificity to notch signalling during Drosophila development. Cold Spring Harb Symp Quant Biol *62*, 177-184.

Graves, H.K., Woodfield, S.E., Yang, C.C., Halder, G., and Bergmann, A. (2012). Notch signalling activates Yorkie non-cell autonomously in Drosophila. PLoS One *7*, e37615.

Guo, G., Huss, M., Tong, G.Q., Wang, C., Li Sun, L., Clarke, N.D., and Robson, P. (2010). Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. Dev Cell *18*, 675-685.

Halder, G., Dupont, S., and Piccolo, S. (2012). Transduction of mechanical and cytoskeletal cues by YAP and TAZ. Nat Rev Mol Cell Biol *13*, 591-600.

Hayashi, S., Lewis, P., Pevny, L., and McMahon, A.P. (2002). Efficient gene modulation in mouse epiblast using a Sox2Cre transgenic mouse strain. Mech Dev *119 Suppl 1*, S97-S101.

Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W.*, et al.* (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature *459*, 108-112.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A.*, et al.* (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet *39*, 311-318.

Himeno, E., Tanaka, S., and Kunath, T. (2008). Isolation and manipulation of mouse trophoblast stem cells. Curr Protoc Stem Cell Biol *Chapter 1*, Unit 1E 4.

Hirate, Y., Cockburn, K., Rossant, J., and Sasaki, H. (2012). Tead4 is constitutively nuclear, while nuclear vs. cytoplasmic Yap distribution is regulated in preimplantation mouse embryos. Proc Natl Acad Sci U S A *109*, E3389-3390; author reply E3391-3382.

Hirate, Y., Hirahara, S., Inoue, K., Suzuki, A., Alarcon, V.B., Akimoto, K., Hirai, T., Hara, T., Adachi, M., Chida, K.*, et al.* (2013). Polarity-dependent distribution of angiomotin localizes Hippo signalling in preimplantation embryos. Curr Biol *23*, 1181-1194.

Home, P., Ray, S., Dutta, D., Bronshteyn, I., Larson, M., and Paul, S. (2009). GATA3 is selectively expressed in the trophectoderm of peri-implantation embryo and directly regulates Cdx2 gene expression. J Biol Chem *284*, 28729-28737.

Jedrusik, A., Parfitt, D.E., Guo, G., Skamagki, M., Grabarek, J.B., Johnson, M.H., Robson, P., and Zernicka-Goetz, M. (2008). Role of Cdx2 and cell polarity in cell allocation and specification of trophectoderm and inner cell mass in the mouse embryo. Genes Dev *22*, 2692-2706.

Jiang, J., Chan, Y.S., Loh, Y.H., Cai, J., Tong, G.Q., Lim, C.A., Robson, P., Zhong, S., and Ng, H.H. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. Nat Cell Biol *10*, 353-360.

Johnson, M.H., and McConnell, J.M. (2004). Lineage allocation and cell polarity during mouse embryogenesis. Semin Cell Dev Biol *15*, 583-597.

Johnson, M.H., and Ziomek, C.A. (1981). The foundation of two distinct cell lineages within the mouse morula. Cell *24*, 71-80.

Kaneko, K.J., and DePamphilis, M.L. (2013). TEAD4 establishes the energy homeostasis essential for blastocoel formation. Development *140*, 3680-3690.

Kaneko, K.J., Kohn, M.J., Liu, C., and DePamphilis, M.L. (2007). Transcription factor TEAD2 is involved in neural tube closure. Genesis *45*, 577-587.

Koch, U., Lehal, R., and Radtke, F. (2013). Stem cells living with a Notch. Development *140*, 689-704.

Kopan, R., and Ilagan, M.X. (2009). The canonical Notch signalling pathway: unfolding the activation mechanism. Cell *137*, 216-233.

Kubaczka, C., Senner, C., Arauzo-Bravo, M.J., Sharma, N., Kuckenberg, P., Becker, A., Zimmer, A., Brustle, O., Peitz, M., Hemberger, M*., et al.* (2014). Derivation and Maintenance of Murine Trophoblast Stem Cells under Defined Conditions. Stem Cell Reports *2*, 232-242.

Kunath, T., Arnaud, D., Uy, G.D., Okamoto, I., Chureau, C., Yamanaka, Y., Heard, E., Gardner, R.L., Avner, P., and Rossant, J. (2005). Imprinted X-inactivation in extra-

embryonic endoderm cell lines from mouse blastocysts. Development *132*, 1649-1661.

Le Bin, G.C., Munoz-Descalzo, S., Kurowski, A., Leitch, H., Lou, X., Mansfield, W., Etienne-Dumeau, C., Grabole, N., Mulas, C., Niwa, H.*, et al.* (2014). Oct4 is required for lineage priming in the developing inner cell mass of the mouse blastocyst. Development *141*, 1001-1010.

Levasseur, D.N., Wang, J., Dorschner, M.O., Stamatoyannopoulos, J.A., and Orkin, S.H. (2008). Oct4 dependence of chromatin structure within the extended Nanog locus in ES cells. Genes Dev *22*, 575-580.

Lewis, J. (1998). Notch signalling and the control of cell fate choices in vertebrates. Semin Cell Dev Biol *9*, 583-589.

Lian, I., Kim, J., Okazawa, H., Zhao, J., Zhao, B., Yu, J., Chinnaiyan, A., Israel, M.A., Goldstein, L.S., Abujarour, R.*, et al.* (2010). The role of YAP transcription coactivator in regulating stem cell self-renewal and differentiation. Genes Dev *24*, 1106-1118.

Lin, S.C., Wani, M.A., Whitsett, J.A., and Wells, J.M. (2010). Klf5 regulates lineage formation in the pre-implantation mouse embryo. Development *137*, 3953-3963.

Liu-Chittenden, Y., Huang, B., Shim, J.S., Chen, Q., Lee, S.J., Anders, R.A., Liu, J.O., and Pan, D. (2012). Genetic and pharmacological disruption of the TEAD-YAP complex suppresses the oncogenic activity of YAP. Genes Dev *26*, 1300-1305.

Liu, J., Sato, C., Cerletti, M., and Wagers, A. (2010). Notch signalling in the regulation of stem cell self-renewal and differentiation. Curr Top Dev Biol *92*, 367-409.

Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J.*, et al.* (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat Genet *38*, 431-440.

Lorthongpanich, C., Messerschmidt, D.M., Chan, S.W., Hong, W., Knowles, B.B., and Solter, D. (2013). Temporal reduction of LATS kinases in the early preimplantation embryo prevents ICM lineage differentiation. Genes Dev *27*, 1441-1446.

Lowell, S., Benchoua, A., Heavey, B., and Smith, A.G. (2006). Notch promotes neural lineage entry by pluripotent embryonic stem cells. PLoS Biol *4*, e121.

Lu, C.W., Yabuuchi, A., Chen, L., Viswanathan, S., Kim, K., and Daley, G.Q. (2008). Ras-MAPK signalling promotes trophectoderm formation from embryonic stem cells and mouse embryos. Nat Genet *40*, 921-926.

Luo, J., Sladek, R., Bader, J.A., Matthyssen, A., Rossant, J., and Giguere, V. (1997). Placental abnormalities in mouse embryos lacking the orphan nuclear receptor ERR-beta. Nature *388*, 778-782.

Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature *487*, 57-63.

Manzanares, M., and Rodriguez, T.A. (2013). Development: Hippo signalling turns the embryo inside out. Curr Biol *23*, R559-561.

Martello, G., Sugimoto, T., Diamanti, E., Joshi, A., Hannah, R., Ohtsuka, S., Gottgens, B., Niwa, H., and Smith, A. (2012). Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. Cell Stem Cell *11*, 491-504.

McDole, K., Xiong, Y., Iglesias, P.A., and Zheng, Y. (2011). Lineage mapping the pre-implantation mouse embryo by two-photon microscopy, new insights into the segregation of cell fates. Dev Biol *355*, 239-249.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631-642.

Morgani, S.M., Canham, M.A., Nichols, J., Sharov, A.A., Migueles, R.P., Ko, M.S., and Brickman, J.M. (2013). Totipotent embryonic stem cells arise in ground-state culture conditions. Cell Rep *3*, 1945-1957.

Morin-Kensicki, E.M., Boone, B.N., Howell, M., Stonebraker, J.R., Teed, J., Alb, J.G., Magnuson, T.R., O'Neal, W., and Milgram, S.L. (2006). Defects in yolk sac vasculogenesis, chorioallantoic fusion, and embryonic axis elongation in mice with targeted disruption of Yap65. Mol Cell Biol *26*, 77-87.

Motosugi, N., Bauer, T., Polanski, Z., Solter, D., and Hiiragi, T. (2005). Polarity of the mouse embryo is established at blastocyst and is not prepatterned. Genes Dev *19*, 1081-1092.

Munch, J., Gonzalez-Rajal, A., and de la Pompa, J.L. (2013). Notch regulates blastema proliferation and prevents differentiation during adult zebrafish fin regeneration. Development *140*, 1402-1411.

Murtaugh, L.C., Stanger, B.Z., Kwan, K.M., and Melton, D.A. (2003). Notch signalling controls multiple steps of pancreatic differentiation. Proc Natl Acad Sci U S A *100*, 14920-14925.

Murtomaki, A., Uh, M.K., Choi, Y.K., Kitajewski, C., Borisenko, V., Kitajewski, J., and Shawber, C.J. (2013). Notch1 functions as a negative regulator of lymphatic endothelial cell differentiation in the venous endothelium. Development *140*, 2365-2376.

Nagy, A., M. Gertsensten, K. Vintersten, and R. Behringer. (2003). Manipulating the mouse embryo: a laboratory manual. (New York., Cold Spring Harbor Laboratory Press).

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95*, 379-391.

Nishioka, N., Inoue, K., Adachi, K., Kiyonari, H., Ota, M., Ralston, A., Yabuta, N., Hirahara, S., Stephenson, R.O., Ogonuki, N.*, et al.* (2009). The Hippo signalling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. Dev Cell *16*, 398-410.

Nishioka, N., Yamamoto, S., Kiyonari, H., Sato, H., Sawada, A., Ota, M., Nakao, K., and Sasaki, H. (2008). Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. Mech Dev *125*, 270-283.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet *24*, 372-376.

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. Cell *123*, 917-929.

Nowotschin, S., Xenopoulos, P., Schrode, N., and Hadjantonakis, A.K. (2013). A bright single-cell resolution live imaging reporter of Notch signalling in the mouse. BMC Dev Biol *13*, 15.

Ntziachristos, P., Lim, J.S., Sage, J., and Aifantis, I. (2014). From fly wings to targeted cancer therapies: a centennial for notch signalling. Cancer Cell *25*, 318-334.

Nus, M., MacGrogan, D., Martinez-Poveda, B., Benito, Y., Casanova, J.C., Fernandez-Aviles, F., Bermejo, J., and de la Pompa, J.L. (2011). Diet-induced aortic valve disease in mice haploinsufficient for the Notch pathway effector RBPJK/CSL. Arterioscler Thromb Vasc Biol *31*, 1580-1588.

Oka, C., Nakano, T., Wakeham, A., de la Pompa, J.L., Mori, C., Sakai, T., Okazaki, S., Kawaichi, M., Shiota, K., Mak, T.W.*, et al.* (1995). Disruption of the mouse RBP-J kappa gene results in early embryonic death. Development *121*, 3291-3301.

Pandolfi, P.P., Roth, M.E., Karis, A., Leonard, M.W., Dzierzak, E., Grosveld, F.G., Engel, J.D., and Lindenbaum, M.H. (1995). Targeted disruption of the GATA3 gene causes severe abnormalities in the nervous system and in fetal liver haematopoiesis. Nat Genet *11*, 40-44.

Perdigoto, C.N., and Bardin, A.J. (2013). Sending the right signal: Notch and stem cells. Biochim Biophys Acta *1830*, 2307-2322.

Pernaute, B., Canon, S., Crespo, M., Fernandez-Tresguerres, B., Rayon, T., and Manzanares, M. (2010). Comparison of extraembryonic expression of Eomes and Cdx2 in pregastrulation chick and mouse embryo unveils regulatory changes along evolution. Dev Dyn *239*, 620-629.

Piotrowska-Nitsche, K., and Zernicka-Goetz, M. (2005). Spatial arrangement of individual 4-cell stage blastomeres and the order in which they are generated correlate with blastocyst pattern in the mouse embryo. Mech Dev *122*, 487-500.

Plachta, N., Bollenbach, T., Pease, S., Fraser, S.E., and Pantazis, P. (2011). Oct4 kinetics predict cell lineage patterning in the early mammalian embryo. Nat Cell Biol *13*, 117-123.

Plusa, B., Piliszek, A., Frankenberg, S., Artus, J., and Hadjantonakis, A.K. (2008). Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. Development *135*, 3081-3091.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279-283.

Ralston, A., Cox, B.J., Nishioka, N., Sasaki, H., Chea, E., Rugg-Gunn, P., Guo, G., Robson, P., Draper, J.S., and Rossant, J. (2010). Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. Development *137*, 395-403.

Ralston, A., and Rossant, J. (2008). Cdx2 acts downstream of cell polarization to cell-autonomously promote trophectoderm fate in the early mouse embryo. Dev Biol *313*, 614-629.

Roper, S., and Hemberger, M. (2009). Defining pathways that enforce cell lineage specification in early development and stem cells. Cell Cycle *8*, 1515-1525.

Rossant, J. (2008). Stem cells and early lineage development. Cell *132*, 527-531.

Rossant, J. (2009). Reprogramming to pluripotency: from frogs to stem cells. Cell *138*, 1047-1050.

Rossant, J., and Tam, P.P. (2009). Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. Development *136*, 701-713.

Russ, A.P., Wattler, S., Colledge, W.H., Aparicio, S.A., Carlton, M.B., Pearce, J.J., Barton, S.C., Surani, M.A., Ryan, K., Nehls, M.C.*, et al.* (2000). Eomesodermin is required for mouse trophoblast development and mesoderm formation. Nature *404*, 95-99.

Shawber, C.J., and Kitajewski, J. (2004). Notch function in the vasculature: insights from zebrafish, mouse and man. Bioessays *26*, 225-234.

Springer, ed. (2011). Advanced Protocols for Animal Transgenesis, an ISTT Manual.

Srinivas, S., Watanabe, T., Lin, C.S., William, C.M., Tanabe, Y., Jessell, T.M., and Costantini, F. (2001). Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. BMC Dev Biol *1*, 4.

Stamatoyannopoulos, J.A., Snyder, M., Hardison, R., Ren, B., Gingeras, T., Gilbert, D.M., Groudine, M., Bender, M., Kaul, R., Canfield, T.*, et al.* (2012). An encyclopedia of mouse DNA elements (Mouse ENCODE). Genome Biol *13*, 418.

Stephenson, R.O., Yamanaka, Y., and Rossant, J. (2010). Disorganized epithelial polarity and excess trophectoderm cell fate in preimplantation embryos lacking E-cadherin. Development *137*, 3383-3391.

Strumpf, D., Mao, C.A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development *132*, 2093-2102.

Suwinska, A., Czolowska, R., Ozdzenski, W., and Tarkowski, A.K. (2008). Blastomeres of the mouse embryo lose totipotency after the fifth cleavage division: expression of Cdx2 and Oct4 and developmental potential of inner and outer blastomeres of 16- and 32-cell embryos. Dev Biol *322*, 133-144.

Tabansky, I., Lenarcic, A., Draft, R.W., Loulier, K., Keskin, D.B., Rosains, J., Rivera-Feliciano, J., Lichtman, J.W., Livet, J., Stern, J.N.*, et al.* (2013). Developmental bias in cleavage-stage mouse blastomeres. Curr Biol *23*, 21-31.

Tamm, C., Bower, N., and Anneren, C. (2011). Regulation of mouse embryonic stem cell self-renewal by a Yes-YAP-TEAD2 signalling pathway downstream of LIF. J Cell Sci *124*, 1136-1144.

Tanaka, S., Kunath, T., Hadjantonakis, A.K., Nagy, A., and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. Science *282*, 2072-2075.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A.*, et al.* (2009). mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods *6*, 377-382.

Tarkowski, A.K. (1959). Experiments on the development of isolated blastomers of mouse eggs. Nature *184*, 1286-1287.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448*, 196-199.

Tiemann, U., Marthaler, A.G., Adachi, K., Wu, G., Fischedick, G.U., Arauzo-Bravo, M.J., Scholer, H.R., and Tapia, N. (2014). Counteracting Activities of OCT4 and KLF4 during Reprogramming to Pluripotency. Stem Cell Reports *2*, 351-365.

Tomioka, M., Nishimoto, M., Miyagi, S., Katayanagi, T., Fukui, N., Niwa, H., Muramatsu, M., and Okuda, A. (2002). Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. Nucleic Acids Res *30*, 3202-3213.

Tschaharganeh, D.F., Chen, X., Latzko, P., Malz, M., Gaida, M.M., Felix, K., Ladu, S., Singer, S., Pinna, F., Gretz, N.*, et al.* (2013). Yes-associated protein up-regulates Jagged-1 and activates the Notch pathway in human hepatocellular carcinoma. Gastroenterology *144*, 1530-1542 e1512.

Tun, T., Hamaguchi, Y., Matsunami, N., Furukawa, T., Honjo, T., and Kawaichi, M. (1994). Recognition sequence of a highly conserved DNA binding protein RBP-J kappa. Nucleic Acids Res *22*, 965-971.

Uyttendaele, H., Ho, J., Rossant, J., and Kitajewski, J. (2001). Vascular patterning defects associated with expression of activated Notch4 in embryonic endothelium. Proc Natl Acad Sci U S A *98*, 5643-5648.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol *3*, RESEARCH0034.

Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F.*, et al.* (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature *457*, 854-858.

Vollrath, B., Fitzgerald, K.J., and Leder, P. (2001). A murine homologue of the Drosophila brainiac gene shows homology to glycosyltransferases and is required for preimplantation development of the mouse. Mol Cell Biol *21*, 5688-5697.

Wang, Q.T., Piotrowska, K., Ciemerych, M.A., Milenkovic, L., Scott, M.P., Davis, R.W., and Zernicka-Goetz, M. (2004). A genome-wide study of gene activity reveals developmental signalling pathways in the preimplantation mouse embryo. Dev Cell *6*, 133-144.

Wang, W.C., and Shashikant, C.S. (2007). Evidence for positive and negative regulation of the mouse Cdx2 gene. J Exp Zool B Mol Dev Evol *308*, 308-321.

Watanabe, T., Biggins, J.S., Tannan, N.B., and Srinivas, S. (2014). Limited predictive value of blastomere angle of division in trophectoderm and inner cell mass specification. Development *141*, 2279-2288.

Watts, J.A., Zhang, C., Klein-Szanto, A.J., Kormish, J.D., Fu, J., Zhang, M.Q., and Zaret, K.S. (2011). Study of FoxA pioneer factor at silent genes reveals Rfx-repressed enhancer at Cdx2 and a potential indicator of esophageal adenocarcinoma development. PLoS Genet *7*, e1002277.

Wennekamp, S., Mesecke, S., Nedelec, F., and Hiiragi, T. (2013). A self-organization framework for symmetry breaking in the mammalian embryo. Nat Rev Mol Cell Biol *14*, 452-459.

Wu, G., Gentile, L., Fuchikami, T., Sutter, J., Psathaki, K., Esteves, T.C., Arauzo-Bravo, M.J., Ortmeier, C., Verberk, G., Abe, K.*, et al.* (2010). Initiation of trophectoderm lineage specification in mouse embryos is independent of Cdx2. Development *137*, 4159-4169.

Yagi, R., Kohn, M.J., Karavanova, I., Kaneko, K.J., Vullhorst, D., DePamphilis, M.L., and Buonanno, A. (2007). Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. Development *134*, 3827-3836.

Yamanaka, Y., Ralston, A., Stephenson, R.O., and Rossant, J. (2006). Cell and molecular regulation of the mouse blastocyst. Dev Dyn *235*, 2301-2314.

Yee, S.P., and Rigby, P.W. (1993). The regulation of myogenin gene expression during the embryonic development of the mouse. Genes Dev *7*, 1277-1289.

Yeom, Y.I., Fuhrmann, G., Ovitt, C.E., Brehm, A., Ohbo, K., Gross, M., Hubner, K., and Scholer, H.R. (1996). Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. Development *122*, 881-894.

Yimlamai, D., Christodoulou, C., Galli, G.G., Yanger, K., Pepe-Mooney, B., Gurung, B., Shrestha, K., Cahan, P., Stanger, B.Z., and Camargo, F.D. (2014). Hippo pathway activity influences liver cell fate. Cell *157*, 1324-1338.

Yoshimizu, T., Sugiyama, N., De Felice, M., Yeom, Y.I., Ohbo, K., Masuko, K., Obinata, M., Abe, K., Scholer, H.R., and Matsui, Y. (1999). Germline-specific expression of the Oct-4/green fluorescent protein (GFP) transgene in mice. Dev Growth Differ *41*, 675-684.

Zaret, K.S., and Carroll, J.S. (2011). Pioneer transcription factors: establishing competence for gene expression. Genes Dev *25*, 2227-2241.

Zhao, B., Ye, X., Yu, J., Li, L., Li, W., Li, S., Lin, J.D., Wang, C.Y., Chinnaiyan, A.M., Lai, Z.C.*, et al.* (2008). TEAD mediates YAP-dependent gene induction and growth control. Genes Dev *22*, 1962-1971.

SUPPLEMENTARY TABLES

Table A. Key transcription factors implicated in ICM and TE differentiation.

Name	Initial expression and localization	Knock-out phenotype	Gain of function phenotype	Reference
TE specific transcription factor				
Cdx2	mRNA found at eight-cell stage. Nuclear protein detected at eight-cell stage. Mosaic early expression, TE-restricted at 32-cell stage	Die at blastocyst stage (zygotic and maternal/zygotic mutants). Epithelial integrity loss. Cavitation fail. TE expresses ICM markers.	Overexpression in ES cells turns them to TS cells.	(Blij et al., 2012; Chawengsaksophak et al., 1997; Dietrich and Hiiragi, 2007; Jedrusik et al., 2008; Niwa et al., 2005; Ralston and Rossant, 2008; Strumpf et al., 2005; Wu et al., 2010)
Tead4	Muclear protein detected at 4 cell stage. Ubiquitous (TE+ICM).	Embryos do not form a blastocyst. Loss of *Cdx2* and *Gata3* expression.	Overexpression of *Tead4* active form in ES cells turns them to TS cells.	(Nishioka et al., 2009; Nishioka et al., 2008; Yagi et al., 2007)
Yap	Detected at 4 cell stage Nuclear in outer cells, cytoplasmic in inside cells. Cytoplasmic in ICM, nuclear in TE.	Die at 8.5dpc. Defects in Yolk Sac, chorioallantoic fusion, and embryonic axis elongation.	Overexpression in inner cells of blastocyst induces *Cdx2* expression.	(Morin-Kensicki et al., 2006; Nishioka et al., 2009)
Gata3	Nuclear protein detected at 16 cell stage. Outer cells, prospective TE cells and TE.	Die at 11.5dpc from massive internal bleeding and placental defects.	Overexpression in ES cells turns them to TS cells.	(Home et al., 2009; Pandolfi et al., 1995; Ralston et al., 2010)
Eomes	Nuclear protein detected at blastocyst stage. TE specific.	Embryos fail to implant. TE differentiation blocked. TE still expresses *Cdx2.*	Overexpression in ES cells turns them to TS cells but they spontaneously differentiate.	(Niwa et al., 2005; Ralston and Rossant, 2008; Russ et al., 2000; Strumpf et al., 2005)
Klf5	Nuclear protein	Embryos do not	Expression in ES	(Lin et al., 2010)

	detectable at two cell stage. Ubiquitous early. Strong in TE, weak in ICM.	form blastocysts. Loss of *Cdx2* but TE and ICM specified	cells permits LIF independent self-renewal.	
Tcfap2c	Suggested maternal expression. Late morula and blastocyst (not analysed earlier). Ubiquitous early and late expression (TE+ICM).	Die at 7.5 dpc with extraembrionary defects. *Cdx2* and *Eomes* reduced expression.	N/A	(Auman et al., 2002)
ICM genes				
Oct4	Maternal expression. Nuclear zygotic protein detected at eight cell stage. Ubiquitous initially; ICM restricted at 4.5dpc.	Morphological normal blastocyst. ICM expresses TE markers at 4.5dpc. Role in Epi/PE segregation	OCT4 overexpression in tissues facilitates reprogramming by interfering with the somatic transcriptional network.	(Frum et al., 2013; Le Bin et al., 2014; Nichols et al., 1998; Tiemann et al., 2014)
Sox2	Nuclear protein detected at eight cell stage. Mosaic expression. In blastocysts Cytoplasmic in TE and nuclear in ICM.	Homozygotes for targeted null mutations implant but fail to develop an egg cylinder or epiblast, and die shortly thereafter.	If overexpressed in ES or TS safeguards the stem cell state (together with some other specific sets of genes).	(Adachi et al., 2013; Avilion et al., 2003)
Nanog	Nuclear protein detected at eight cell stage. Mosaic early expression. ICM/epiblast restricted.	Required for EPI formation and germ line reprogramming.	Expression in ES cells permits LIF independent self-renewal.	(Chambers et al., 2003; Dietrich and Hiiragi, 2007; Mitsui et al., 2003)

Esrrb	ICM expression (Not assessed earlier). In post-implantation development located in the Extraembryonic ectoderm.	Die around 9.5dpc. Failure of the chorion to develop and subsequent placental defect.	Expression in ES cells permits LIF independent self-renewal.	(Luo et al., 1997; Martello et al., 2012)

Table B. Distribution (%) of embryos for the different allelic combinations.

Rbpj *Tead4*	+/+ +/+	+/+ +/-	+/- +/+	+/- +/-	+/+ -/-	-/- +/+	+/- -/-	-/- +/-	-/- -/-	total
no.embryos (%)	13 (16,88%)	10 (12,98%)	10 (12,98%)	18 (23,37%)	2 (2,59%)	7 (9,09%)	8 (10,39%)	9 (11,69%)	0	77
expected %	6,25%	12,50%	12,50%	25,00%	6,25%	6,25%	12,50%	12,50%	6,25%	100,00%

PUBLICATIONS

This work has led to the publication of the following article:

Rayon T., Menchero S., Nieto A., Xenopoulos P., Crespo M., Cockburn K., Canon S., Sasaki H., Hadjantonakis A.K., de la Pompa J.L., Rossant J., Manzanares, M (2014). *Notch and Hippo converge on Cdx2 to specify the trophectoderm lineage in the mouse blastocyst*. **Dev Cell**. 2014 Aug 25;30 (4)410-22.

The collaboration in other research projects during the development of the thesis has resulted in the following publications:

Pernaute B, Cañon S, Crespo M, Fernandez-Tresguerres B, **Rayon T**, Manzanares M (2010). Comparison of extraembryonic expression of Eomes and Cdx2 in pre-gastrulation chick and mouse embryo unveil regulatory changes along evolution. **Dev Dyn** 239, 620-629.

Fernandez-Tresguerres B, Cañon S, **Rayon T**, Pernaute B, Crespo M, Torroja C, Manzanares M (2010). *Evolution of the mammalian embryonic pluripotency gene regulatory network.* **PNAS** 107, 19955-19960.

Peralta M, Steed E, Harlepp S, González-Rosa JM, Monduc F, Ariza-Cosano A, Cortés A, **Rayón T,** Gómez-Skarmeta JL, Zapata A, Vermot J, Mercader N (2013). *Heartbeat-driven pericardiac fluid forces contribute to epicardium morphogenesis*. **Curr Biol**. 2013 Sep 23;23(18):1726-35.

Abad M, Mosteiro L, Pantoja C, Cañamero M, **Rayon T**, Ors I, Graña O, Megías D, Domínguez O, Martínez D, Manzanares M, Ortega S, Serrano M (2013). *Reprogramming in vivo produces teratomas and iPS cells with totipotency features*. **Nature**. 2013 Oct 17;502(7471):340-5