PhD degree in Systems Medicine (curriculum in Molecular Oncology)
European School of Molecular Medicine (SEMM),
University of Milan and University of Naples "Federico II"
Settore disciplinare: Bio/11

Identifying the molecular players downstream to replication stress response in early embryonic development

Negar Arghavanifard (PharmD)

The FIRC Institute of Molecular Oncology (IFOM), Milan, Italy
Matricola n. R11786

Supervisor: Prof. Vincenzo Costanzo, MD, PhD
The FIRC Institute of Molecular Oncology (IFOM), Milan, Italy

Added supervisor: Dr. Sina Atashpaz, PharmD, PhD
The FIRC Institute of Molecular Oncology (IFOM), Milan, Italy

PhD Coordinator: Prof. Giuseppe Viale

Anno accademico 2019-2020

Dedication

This thesis is dedicated to

*My **Mother and my Father's soul**, for their endless love, great support and patience.*

*My mentor, **Dr. Sina Atashpaz**, who compassionately taught me that getting PhD degree is not only about accomplishing a research project but most importantly is about self-improvement in different aspects of life…*

Acknowledgement

I would like to express my sincere gratitude to those who kindly helped me during this project.

My supervisor, **Prof. Vincenzo Costanzo** for providing me the opportunity to work in his lab.

My added supervisor, **Dr. Sina Atashpaz** for his full support, daily discussions and useful instructions.

My internal advisor, **Prof. Diego Pasini**, and my external advisor, **Prof. Andres Joaquin Lopez-Contreras** for helpful suggestions.

My internal examiner, **Prof. Saverio Minucci** (Istituto Europeo di Oncologia (IEO), Milan, Italy) and my external examiner **Dr. Travis H Stracker** (IRB-institute for research in biomedicine, Barcelona, Spain) for their precious time and kind instructions.

My research teammates, **Dr. Sara Samadi Shams** and **Andrea Gnocchi** for their supportive availability and teaching me new skills.

My best friend, **Kourosh Hayatigolkhatmi**, who rekindled my passion and creativity in research through engaging me in scientific to science fictional discussions.

My colleagues, **Dr. Anna De Antoni**, **Federica Pezzimenti**, **Dr. Lucia Falbo**, **Dr. Vincenzo Sannino**, **Anjali Mann**, **Catiana Elhkaref**, **Miguel Ramirez Otero**.

I would like to kindly thank the staff of IFOM imaging facility (**Massimiliano Garre**, **Francesca Casagrande**, **Maria Grazia Totaro**, **Emanuele Martini**, **Sara Barozzi, Sara Michela Martone** and **Serena Magni**), primary culture and preclinical model unit (**Elisa Allievi**) and cell culture facility (**Ilaria Rancati**, **Stefania Lavore**, **Eleonora Verga**, **Cinzia Cancellieri, Giuseppe Ossolengo** and **Laura Carmignani**).

I would like to kindly appreciate IFOM welcome office members (**Marina Properzi** and **Mio Sumie**) and SEMM PhD office members (**Francesca Fiore** and **Veronica Viscardi**) for their kind assistance and support.

Table of Contents

List of abbreviations

-/-	**Deficient**
•OH	**Hydroxyl radical**
^{0}C	**Celsius**
2D	**Two-dimensional**
2i	**Two inhibitors (CHIR 99021, PD 0325901)**
3D	**Three-dimensional**
53BP1	**Tumor suppressor p53-binding protein 1**
9-1-1 complex	**Rad9-Hus1-Rad1 complex**
AAD	**ATR activating domain**
ADP	**Adenosine diphosphate**
AGT	**O^{6}-alkylguanine-DNA alkyl transferase**
Amot	**Angiomotin**
APE1	**AP nuclease I**
APH	**Aphidicolin**
aPKC	**Atypical protein kinase C**
APTX	**Aprataxin**
ART	**Assisted reproductive strategies**
AT	**Ataxia Telangiectasia**
AT-rich	**Adenine and Thymine rich DNA regions**
ATM	**Ataxia telangiectasia mutated**
ATMi	**ATM inhibitor**

ATP	**Adenosine triphosphate**
ATR	**Ataxia telangiectasia and Rad3-related**
ATRi	**ATR inhibitor**
ATRIP	**ATR-interacting protein**
AVE	**Anterior VE**
Axin2	**Axis inhibition protein 2**
BER	**Base excision repair**
BMP	**Bone morphogenetic protein**
BRCA1	**Breast cancer type 1**
BSA	**Bovine serum albumin**
C	**cell**
cAMP	**Cyclic Adenosine monophosphate**
Cdc	**Cell division cycle**
CDK	**Cyclin-dependent kinase**
cDNA	**Complementary DNA**
Cdt1	**Chromatin licensing and DNA replication factor 1**
Cdx2	**Caudal type homeobox 2**
CFS	**Common fragile site**
ChIP	**Chromatin immunoprecipitation**
Chk1	**Checkpoint kinase 1**
Chk2	**Checkpoint kinase 2**
Cip1	**Cyclin-dependent kinase inhibitor 1**

Cre	**Causes recombination**
C_t	**Cycle threshold**
CtIP	**C-terminal binding protein interacting protein**
D-EPSC	**EPSCs generated by Deng and his colleagues**
d.p.c	**days post-coitum**
DAPI	**4', 6-diamidino-2-phenylindole**
Dbf4	**Dumbbell former 4 protein**
DD	**DNA damage**
DDK	**Dbf4-dependent kinase**
DDR	**DNA damage response**
DiM	**(S)- (+) -dimethindene maleate minocycline hydrochloride**
DMEM	**Dulbecco's modified Eagle's medium**
DNA	**Deoxyribonucleic acid**
DNA-PK_{cs}	**DNA-dependent protein kinase**
DNMT	**DNA Methyltransferase**
dNTP	**Deoxy nucleoside triphosphate**
DSB	**Double strand break**
DSBR	**Double-stranded brake repair**
Dux	**Double homeobox protein**
DVE	**Distal VE**
E	**Embryonic day**
e.g.	**For example**

E2F1	**E2F transcription factor 1**
E3 ligase	**Ubiquitin ligase**
EB	**Embryoid body**
ECM	**Extra-cellular matrix**
EDTA	**Ethylenediaminetetraacetic acid**
Ef1	**Eukaryotic translation elongation factor 1**
Elf5	**E74 like ETS transcription factor 5**
Eomes	**Eomesodermin (T-box brain protein 2)**
EPC	**Ectoplacental cone**
EPI	**Epiblast**
EpiSC	**Primed epiblast stem cells**
EPSC	**Expanded potential stem cell**
EPSC-B-blastoid	**Blastoids generated from EPSCs by Belmonte and his colleagues**
EPSC-Z-blastoid	**Blastoids generated from EPSCs by Zernica-Goetz and her colleagues**
EPSCM	**Extended potential stem cell medium**
Erk2	**Extracellular signal-regulated kinase 2**
ESC	**Embryonic stem cell**
esiRNA	**Endoribonuclease-prepared small interfering RNA**
ET	**ESC-TSC**
Et al.	**And others**
ET embryo	***In vitro* 3D embryo-like structures**

ETAA1	**Ewing's tumor-associated antigen 1**
etc.	**Et cetera**
ETM	**ET-embryo medium**
ExE	**Extra-embryonic ectoderm**
FA	**Fanconi anemia**
FACS	**Fluorescent activated cell sorting**
FANCM	**FA complementation group M**
FBS	**Fetal bovine serum**
Fe^{2+}	**Iron (II)**
FGF4	**Fibroblast growth factor 4**
Fgfr2	**Fibroblast growth factor receptor 2**
FNACD2	**Fanconi anemia Group D2 protein**
G1	**First gap phase**
G2	**Second gap phase**
GAPDH	**Glyceraldehyde 3-phosphate dehydrogenase**
GFP	**Green fluorescent protein**
GSK3	**Glycogen synthase kinase 3**
H2AX/ γH2AX	**H2A histone family member X/ Phosphorylated H2AX**
H_2O_2	**Hydrogen peroxide**
HDR	**Homology-directed repair**
HR	**Homologous recombination**
HRP	**Horseradish peroxidase**

hrs.	**Hours**
HU	**Hydroxy urea**
Hus	**Hydroxyurea sensitive**
i.e.	**That is**
ICL	**Inter-strand crosslink**
ICM	**Inner cell mass**
IF	**Immunofluorescent**
IgG	**Immunoglobulin G**
iPSC	**Induced pluripotent stem cell**
IR	**Ionizing radiation**
IVF	***In vitro* fertilization**
JNK	**c-Jun N-terminal kinase**
kb	**Kilobase**
KD	**Knock down**
Klf4	**Gut-enriched Krüpple-like factor 4**
KSOM	**Potassium simplex optimization medium**
L-EPSC	**EPSCs generated by Liu and his colleagues**
Lats 1/2	**Large tumor suppressor kinase 1**
LCDM	**A cocktail medium consisting of human LIF, CHIR 99021, DiM, and MiH**
Lgl	**Lethal giant larva**
LIF	**Leukemia inhibitory factor**
LOG	**Logarithm**

LoxP	**Locus of X-over P1**
M	**Mitosis phase**
m	**Mouse**
MAPK	**Mitogen-activated protein kinase**
mCherry	**A member of the mFruits family of monomeric red fluorescent proteins**
MCM 2-7	**Mini-chromosome maintenance proteins**
MDC1	**Mediator of DNA damage checkpoint protein 1**
MEF	**Mouse embryonic fibroblast**
MEK 1	**Meiotic chromosome-axis associated kinase**
MiH	**Minocycline hydrochloride**
mins	**Minutes**
MMR	**Mismatch repair**
MMS	**Methyl-methane sulfonate**
Mre11	**Double-strand break repair nuclease**
MRN complex	**Mre11-Rad50-Nbs1**
mRNA	**Messenger RNA**
NBS	**Nijmegen breakage syndrome**
Nbs1	**Nibrin**
NBSLD	**Nijmegen breakage syndrome-like disease**
NEEA	**Non-essential amino acid**
NER	**Nucleotide excision repair**

NFBD1	**Nuclear factor with BRCT domain 1**
NHEJ	**Nonhomologous end joining**
nm	**Nanometer**
Nrf2	**Neurofibromatosis 2**
O_2	**Oxygen**
Oct4	**Octamer-binding transcription factor 4**
OE	**Over expression**
Orc 1-6	**Origin recognition complexes**
P21	**Cyclin-dependent kinase inhibitor 1A**
p38	**A mitogen-activated protein kinase**
p53	**Tumor protein p53**
Par	**Polarity proteins**
PARP1	**Poly (ADP-ribose) polymerase I**
PBS	**Phosphate buffer saline**
PCA	**Principle component analysis**
PCNA	**Proliferating cell nuclear antigen**
PE	**Primitive endoderm**
PFA	**Paraformaldehyde**
PGC	**Primordial germ cell**
PGD	**Preimplantation genetic diagnosis**
PI3K	**Phosphatydilinositol-3kinase**
PI	**Placental lactogen**

Plet1	**Placenta expressed transcript 1**
PLF	**Proliferin**
PNKP	**Polynucleotide kinase 3-prime phosphatase**
pre-RC	**Pre-replicative complex**
Prl2c2	**Prolactin-2C2**
Prlpa	**Prolactin-like protein a**
qPCR	**Quantitative polymerase chain reaction**
R loops	**RNA/DNA hybrids**
R26R-H2B-mCherry	**A conditional reporter line which harbors the cDNA of fused Histone 2B-mCherry at ROSA26 locus**
Rad	**Radiation repair genes**
RBPJ	**Recombination signal binding protein for immunoglobulin Kappa J**
RF	**Replication factor**
RFB	**Replication factors barrier**
RFC	**Replication factor C**
RIPA	**Radioimmunoprecipitation assay buffer**
RNA	**Ribonucleic acid**
RNA-seq	**Bulk RNA sequencing**
RNF	**Ring finger protein**
ROS	**Reactive Oxygen species**
RPA	**Single-stranded DNA binding protein**
rpm	**Revolutions per minute**
RPMI	**Roswell park memorial institute**

RS	**Replication stress**
RSR	**Replication stress response**
RSS	**Replication start site**
RT	**Room temperature**
RT-qPCR	**Reverse transcriptase qPCR**
S	**Synthesis phase**
SC	**Stem cell**
sc	**Single cell**
Sox2	**(Sex-determining region Y)-box 2**
Src	**Proto-oncogene tyrosine-protein kinase**
SSB	**Single strand break**
SSBR	**Single-stranded break repair**
ssDNA	**Single-stranded DNA**
SSM	**Slipped strand mispairing**
STAT3	**Single transducer and activator of transcription 3**
TAZ	**Tafazzin**
Tbp	**TATA-binding protein**
TBS	**Tris-buffered saline**
TBST	**TBS buffer 1X plus 0.1% Tween 20**
Tc	**Tetracycline**
Tcfap2c	**Transcription factor AP-2 gamma**
Tdtomato	**Tandem dimeric of tomato**

TE	**Trophectoderm**
Tead4	**TEA domain transcription factor 4**
TF	**Transcription factor**
TFBM	**Transcription factor binding motif**
TGC	**Trophoblast giant cell**
TGFß1	**Transforming growth factor beta 1**
TIP60	**Histone acetyltransferase KAT5**
TLGC	**Trophoblast-like giant cells**
TLK	**Tousled-like kinase**
TLS	**Translesion synthesis**
TLSC	**Trophoblast-like stem cell**
TNKS1/2	**Tankyrases**
top	**Topoisomerase**
TopBP1	**DNA topoisomerase 2-binding protein 1**
TSC	**Trophoblast stem cell**
TSCM	**Trophoblast stem cell medium**
Tx	**Tamoxifen**
UV	**Ultraviolet**
VE	**Visceral endoderm**
WB	**Western blotting**
Wnt	**Wingless-related integration site**
WT	**Wild type**

Wwtr1	**WW domain-containing transcription regulator protein 1**
X	**Times more**
XEN	**Extra-embryonic endoderm stem cells**
XLF	**XRCC4-like factor**
XP	**Xeroderma pigmentosum**
XRCC4	**X-ray repair cross-complementing protein 4**
Yap1	**Yes-associated protein 1**
ZGA	**Zygotic genome activation**

Figures index

Abstract

Rapidly proliferating pre-implantation embryos and their derived cells, called as embryonic stem cells (ESCs) are vulnerable to various genotoxic insults. Any unrepaired damage at this critical stage can give rise to either miscarriage or later prenatal abnormalities. However, the mechanisms through which the ESCs and early-stage embryos respond to genotoxic stress is not fully studied.

Through *in vivo* chimera assay, we have recently shown that replication stress (RS)-induced ESCs contribute to the extra-embryonic compartment of the embryos in contrast to the intact cells which contribute only to the embryonic proper.

In vivo chimera assay possesses various limitations including technical and ethical challenges. Here, to overcome these limitations and to uncover the molecular mechanisms underlying our previous findings, we studied the impact of RS on *in vitro* differentiation of ESCs to the extra-embryonic lineages. Briefly, we found that RS increases the expression of key trophectoderm (TE) markers in ESCs. Moreover, under trophoblast stem cells (TSCs) culture condition, we demonstrated that RS induces the differentiation of ESCs toward TSCs and increases the number of trophoblast-like stem cells (TLSCs).

Next, to understand whether the ESC-derived TSCs (i.e., TLSCs) are functionally similar to the embryo-derived TSCs, we established an innovative combinatorial approach of ESC *in vitro* differentiation with 3D stem cell embryogenesis. Concisely, we could generate structures from the assembly of TLSCs and ESCs. Interestingly we found that these structures are characteristically similar to the structures generated from ESCs and embryo-derived TSCs. Through such approach we studied the impact of RS on the differentiation potential of ESCs. Strikingly, we found that RS-induced TLSCs are able to give rise to the higher number of embryo-like structures.

To further consolidate these findings, we studied the impact of RS on the terminal differentiation of ESCs toward trophoblast giant cells (TGCs). Surprisingly, we found that RS induces the differentiation of ESCs toward TGCs and increases the number of trophoblast-like giant cells (TLGCs).

Finally, we focused on the molecular mechanisms through which RS induces the differentiation of ESCs toward TSCs. In brief, we found the increased expression of TE genes in ESCs is mediated through ATR-Chk1 pathway and does not require the activation of ATM-Chk2 pathway. We also showed that RS increases the expression of Tead4, the master regulator of TE-specific transcriptional program. We demonstrated that the expression of Tead4 in RS-induced ESCs is not ATR-Chk1-dependent however, its binding to the transcription factor binding motifs (TFBMs) of TE genes is mediated through ATR.

Ultimately, we found that the ATR-Chk1-mediated binding of Eomes to its TFBM further enhances the expression of this TE-specific gene in RS-induced ESCs.

Overall, these findings reveal a role of ATR-Chk1-dependent transcriptional regulation of TE genes in inducing the ESCs differentiation toward TSCs in response to RS.

Introduction

Eukaryotic DNA replication and cell cycle

DNA replication is a fundamental process for cell reproduction in all organisms which ensures the faithful transmission of the duplicated genome into the daughter cells (1, 2). Complete and accurate genome segregation occurs only through synchronous coordination of DNA replication and cell cycle (3).

Cell cycle

The cell cycle is mainly composed of four phases called as G1, S, G2 and M (4). In the S phase (synthesis phase) which is followed by G1 phase (first gap phase), the cell replicates its DNA. This phase is separated by G2 phase (second gap phase) from M phase (mitosis phase) in which the segregation of the duplicated genome takes place between the progenies (5, 6). During G1 phase the cell grows and synthesizes various materials required for DNA replication and cell division (7) however, in the G2 phase the cell produces further materials needed for entering mitosis (8). The two gap phases are critical periods of time during which cell passes through restriction points. In such moments, the cell assesses the enzyme and protein sufficiency for proliferation (in G1 phase) and error-free DNA synthesis before mitosis (in G2 phase) (9).

Eukaryotic DNA replication

DNA replication is a semi-conservative phenomenon through which the cell serves its genome as a template to precisely duplicate it into two identical sets (10, 11). DNA replication begins at tens of thousands of replication start sites (RSSs) namely, replication origins which repeatedly are located through the eukaryotic linear chromosomes (12). Prior to the S phase, the pre-replicative complex (pre-RC) assembles at replication origins. These complexes act as licensing factors which ensure that DNA replicates only once in each cell cycle (13, 14). To form pre-RC, initially six conserved ATPase origin recognition complexes (Orc 1-6) mark the RSSs. Respectively, Cdc6 and Cdt1 bind to Orc 1-6 and together load the mini-chromosome maintenance (MCM 2-7) helicase onto these sites leading to DNA unwinding (15-17). Activation of pre-RC complex is the next essential step which is conducted through the kinase activity of Cdc7 in association with Dbf4 (18). The onset of S phase, Cdc6 is released and replaced by Cdc45 protein which consequently recruits DNA polymerase-α-primase and single-stranded DNA binding protein (RPA) to the replication origins (19). Subsequently, in each RSS two replication forks (RFs) form and elongate bidirectionally to synthesize new DNA and terminate where two active forks converge (20).

Cell cycle control of DNA replication

Equal segregation of chromosomes that ought to be duplicated precisely in each cell cycle is essential for maintenance of the genomic integrity (21). Any perturbation that leads to partial or over-replication of DNA could cause cell death or malignancy (22). To prevent these events, cells incorporate safety mechanisms to sequentially control the cell cycle progression (23). Of which, there are two types of serine-threonine kinases, the cyclin-dependent kinase (CDK) and the Dbf4-dependnt kinase (DDK) that tightly control the cell cycle events (24). CDKs are mainly active at two major transitions, G1/S and G2/M and regulate cellular metabolism, transcription, differentiation and DNA repair (25). CDKs have low level of kinase activity, and their activation is dependent on oscillatory cyclins. Upon activation, CDKs phosphorylate a sets of molecules required for initiation of replication (26). Briefly, CDKs play two main roles: first, they remove the pre-RCs from DNA through activation of origin firing factors so that the unlicensed origins could not initiate replication and second, highly active CDKs restricts the formation of new pre-RCs. Therefore, if required, CDKs can either leave behind the licensed origin or prevent relicensing (27). DDK is highly active during cell cycle transition to the S phase (28) and mainly phosphorylates the MCM complex (27). Higher eukaryotes have an extra safety molecule called as geminin which binds directly to Cdt1 hence, inhibits the recruitment of MCMs through Cdt1 on chromatin (29).

Cell cycle checkpoints

Cell cycle checkpoints are surveillance mechanisms that actively monitor cell cycle progression and in case of necessity, they could halt the cycle as far as its integrity is ensured (30). Mammalian checkpoints are mainly divided into five categories. First are sensor molecules including, Rad9–Hus1–Rad1 PCNA-like sliding clamp complex, Rad17–RFC clamp loading complex, and Mre11–Rad50–Nbs1 or 'MRN' nuclease complex. Second are mediators such as, BRCA1, MDC1/NFBD1, 53BP1 and Claspin. Third are apical signal transducing kinases of the phosphatidylinositol 3-kinase (PI3K) -like family namely, ATM and ATR kinases. Fourth are distal transducer kinases, Chk1 and Chk2. Lastly are effector proteins for instance, Cdc25 phosphatases, various DNA repair proteins, p53 and E2F1 transcription factors, chromatin components and regulators such as histone H2AX and tousled-like kinases (Tlks) (**Figure 1**) (31-35).

A) G1 phase checkpoint pathways

When DNA is perturbed, due to the exogenous or endogenous stimuli, G1 checkpoint network prevents cell cycle entry to the S phase. ATM/ATR and Chk1/Chk2 kinases are

critical G1 checkpoints that targets Cdc25A and p53 (34). Chk1 and Chk2 phosphorylate Cdc25A and mark it for proteasomal degradation (36). Such kinetic activity prevents the dephosphorylation process of Cdc25A on CDK2 subunit of cyclin E/CDK2 and cyclin A/CDK2 kinases. Consequently, Cdc45 cannot be loaded on chromatin which is required for DNA polymerase α recruitment hence, this process prevents replication initiation (36, 37). On the other hand, p53 is phosphorylated by both ATM/ATR and Chk1/Chk2 (34). P21 (WAF/Cip1) is the inhibitor of CDKs and transcriptionally is dependent on p53 (38). It takes several hours (hrs.) for p21 to accumulate over a threshold level to block cyclin E/CDK2. This mechanism is the alternative replacement for the acute inhibition of CDK2 through Cdc25A degradation which eventually leads to sustained or permanent cell cycle blockade (34).

B) S phase checkpoint pathways

Activation of the S phase checkpoints reversibly slows down the DNA replication progression through inhibition of new replicon initiation. In contrast to the G1 and G2, the checkpoints in the S phase do not lead to the cell cycle arrest but only delay it. Moreover, they are not dependent on p53 (34). Similar to the G1 phase, the ATM–Chk2–Cdc25A–CDK2–Cdc45 cascade acts as a key mechanism to rapidly reduce the rate of DNA replication in the S phase (39).

C) G2 phase checkpoint pathways

Entering into the G2 phase with inappropriate or incomplete replication remained from the S phase or facing with DNA damage (DD) in the G2 phase activates the related checkpoints to prevent cells from the mitotic catastrophe. The G2 checkpoints act similarly to the G1 pathways either through rapid post-translational modification of effector proteins, or sustained alteration of transcriptional programs. ATM/ATR and Chk1/Chk2-mediated subcellular localization or inhibition of Cdc25C phosphatase hampers the activity of Cyclin B/CDK1 kinase thus, prevents transition to the M phase (31). It is debated that inhibition of Cdc25 phosphatases imposes longer cell cycle arrest in G2 rather than in G1 and S phase (34).

DNA damage

DNA is intrinsically an active molecule which makes it susceptible to be constantly damaged by endogenous and exogenous agents. Endogenous damages mostly arise from chemical reaction of DNA with water (i.e., hydrolysis) or reactive Oxygen species (ROS) (i.e.,

oxidation) which are naturally found in the cells (40). However, environmental, physical and chemical agents such as ultraviolet radiation (UV), ionizing radiation (IR), alkylating agents and crosslinking agents cause exogenous DDs (41).

A) Endogenous DNA damage

1) Replication errors

Accurate DNA synthesis is relied on the exonuclease activity of DNA polymerases called as proofreading which ensures the correct insertion of complementary deoxynucleotide according to the template sequence (42). If errors escape from proofreading, the mismatch repair (MMR) pathway starts to act to increase the replication fidelity through removing mismatched dNTPs. Therefore, mutations in MMR and polymerases proofreading activity coding genes leads to mutations and related diseases (43).

2) Slipped strand mispairing (SSM)

Repeated sequences are found abundantly in eukaryotic genome. These repeated units called as microsatellite sequences are unstable hence, are susceptible to go under base insertion or deletion. These alterations, which is caused by replication slippage, can potentially shift the reading frame and lead to the certain types of genetic diseases and cancer (44).

3) Topoisomerase-associated DNA damage

Topoisomerases are proteins which control the topological state of the DNA. In other words, they can either untwist the supercoiled DNA which is generated during replication or can overwind it. To control the DNA structure, these enzymes transiently generate DNA strand breaks. Type I topoisomerase (top1) introduces single strand breaks (SSBs) however, type II enzyme forms double strand breaks (DSBs). DNA strand breaks can be produced due to the exposure to IR, alkylating agents, etc. The presence of these gaps right after the top1 binding site creates irreversible DNA cleavage and can be converted into the DD (45).

4) Oxidative DNA damage

ROS are endogenously produced in cell respiration, catabolic and anabolic metabolisms. ROS at low level are vital and act as a messenger in redox signaling reactions and also, phagocytes produce them to fight against infections (46). However, excess amount of ROS is detrimental and leads to the development of various diseases including cancer, Alzheimer's disease, Parkinson's disease, diabetes, and heart failure (41). The most reactive

ROS is •OH radical produced during Fenton's reaction of H_2O_2 with Fe^{2+} and is highly capable to damage DNA bases (47). ROS also have the potency to attack the DNA backbone and generate SSBs (48).

Spontaneous base deamination, abasic sites and DNA methylation are some of the other sources of endogenous DD (41).

B) Exogenous DNA damage

1) Ionizing radiation (IR)

DD by IR can be either direct or indirect through water radiolysis to generate ROS (49). Indeed, ROS accounts for 65% of IR-induced DD and thus, generates similar base lesions as ROS (50). Besides, IR form DSBs in case of attacking multiple DNA sites (51).

2) Ultraviolet radiation (UV)

UV can damage DNA in two ways. First, if DNA absorb UV, it can be directly excited and photochemically altered. Second, if UV cannot be absorbed, the photosensitizers transfer its energy to DNA. UV has three ranges of wavelength: UV-C (190–290 nm), UV-B (290–320 nm) and UV-A (320–400 nm). UV-C and with less efficiency, UV-B cause pyrimidine dimers. However, UV-A creates DNA breaks by the excitation of endogenous and exogenous photosensitizers (41, 52).

3) Exogenous chemical agents

- Alkylating agents are electrophilic and react with nucleophilic base ring of Nitrogen in DNA. Of note, Methyl-methane sulfonate (MMS), temozolomide and dacarbazine are among alkylating agents (41, 53).
- Hydroxy urea (HU) depletes the cellular dNTP pool through inactivation of ribonucleotide reductase. Moreover, at 37 ^{0}C it generates Nitric oxide and Hydrogen cyanide. These agents induce DNA base damage (54).
- Aphidicolin (APH) reversibly inhibits DNA polymerases α, ε and δ. It only affects the cells that are in the S phase while, the cells in other phases progress until they accumulate in the G1/S checkpoint. During the S phase, APH slows down the activity of DNA polymerase when helicase is moving forward through which it makes disconnected replicon structures. This process generates long stretches of single-stranded DNA (ssDNA) which are highly susceptible to break especially at common fragile sites (CFSs). These sites are conserved regions in mammals which are highly

unstable under replication stress (This phenomenon will be explained in the following section.). At low concentrations, APH slows down but not completely blocks the DNA polymerase activity. The slow activity of DNA polymerase leads to the generation of ssDNA with high probability at CFSs. At higher concentrations, APH stalls the DNA polymerase and arrests the cells in the S phase. In response to the formation of disconnected replicons by APH, ATR kinase quickly becomes active. Longer exposure to APH activates also ATM possibly due to the formation of DSBs in stalled RFs (53).

- Crosslinking agents inhibit the uncoiling of DNA strands and consequently RF progression through introducing covalent bounds between nucleotides, either on the same strand (i.e., intra-strand crosslink) or on the opposite strand (i.e., inter-strand crosslink (ICL)). Cisplatin and mitomycin C are intra- and inter- strand crosslinkers, respectively (53).
- Topoisomerase I and II inhibitors, such as camptothecin and etoposide, respectively halt DNA unwinding and therefore, prevent replication progression (53).

Replication stress

There is no unified explanation about replication stress (RS) or a group of markers which clearly could characterize this phenomenon. Nevertheless, any source of stress that leads to slowing or stalling the RF progression or DNA synthesis is defined as RS. In other words, RS happens due to the asynchronous activity of DNA unwinding through helicase and halted or decelerated DNA polymerase, which leads to the formation of ssDNA (53, 55).

Sources of replication stress

Cells replicate their genome in each cell cycle initiating it through large number of replication origins. This process represents a challenge for the fork passage and therefore, imposes a potential source of RS. Moreover, some regions of the genome called as “intrinsically difficult to replicate regions” are more sensitive to RS induced by endogenous or exogenous stimuli. These regions, also known as CFSs in mammals, are places where gaps and breaks accumulate under RS. CFSs are vulnerable to RS because they harbor AT-rich sequences where the fork progression is slow and frequently stalls. Also, non-canonical DNA structures and very large genes are located in these sites further enhancing their vulnerability to RS (56, 57).

A) Over- or under-firing of replication origins

Sufficient amount of replication origins which have to be sequentially fired once and only once per cell cycle is essential for the genome stability. This is modulated by the level of dNTPs. Low level of dNTP pool leads to the firing of higher number of replication origins with the slower rate to ensure the completion of replication. If the amount of dNTP reaches below the critical level, ssDNA accumulates. This happens with the higher propensity at CFSs since, they lack efficient number of origins therefore, they are unable to compensate the effect of RS (58). Over-firing of replication origins is also detrimental because, it consumes dNTPs and replication factors to a limiting level leading to destabilization of RFs. For instance, higher number of fired origins recruit massive amount of RPA to cover ssDNA however, these strands mostly remain unprotected leading to fork instability (22, 59).

B) Impediments to replication fork progression

DNA lesions, DNA-protein complexes, secondary DNA structures and highly transcribed genes can disrupt the RF progression either through altering the activity of helicase or DNA polymerase. These obstacles are often called as RF barriers (RFBs) that arise from exposure to exogenous genotoxic agents or can be caused by cellular events (60).

C) Transcription and replication collision

Transcription and replication both use DNA as template therefore, they are prone to collide. To avoid this incident, these two events are spatiotemporally separated within the nucleus. However, evidences show that still transcription interferes with DNA replication and causes genomic instability (61). During transcription RNA anneals to the template DNA strand and the non-template strand remain unpaired. This RNA/DNA hybrids (R loops) possibly induce RS (62). Moreover, coincidence of transcription and replication leads to the accumulation of positive supercoils which is another source of RS (63). Besides, in mammalian cells transcription of very large genes takes place throughout the cell cycle while, these genes are mostly located at CFSs and are replicated in the late S phase accordingly these genes are more prone to RS (59, 64).

D) Unbalanced DNA replication

Initiation of DNA replication under inappropriate metabolic situation is called as unbalanced DNA replication. For example, as discussed before, insufficient dNTP pool due to the exposure to HU leads to slower RF progression and RS (59). Furthermore, efficient chromatin assembly (e.g., histones) supply during replication is essential for the transmission of epigenetic markers to the progenies. Otherwise, any shortage in chromatin

assembly factors destabilizes the advancing RFs. For instance, inadequate newly synthesized histones slow down the rate of fork elongation and predisposes cells to RS (59, 65, 66).

DNA damage response

If left unrepaired, DD poses a threat to the cell. To combat this issue, cells have developed mechanisms - overall called as DD response (DDR) - to spot DNA lesions, signal their presence and proceed with their repair. Any defect in DNA repair machinery makes cells sensitive to DD and many of such defects are involved in the emergence of various diseases.

A) Base DNA damage repair

1) Reversal of DNA damage

A small group of DNA lesions can be directly reversed by a single repair protein in an error-free manner. There are three main direct DNA repair mechanisms. First, photolyases reverse lesions induced by UV. Second, O6-alkylguanine-DNA alkyl transferases (AGTs) and third, dioxygenases reverse O-alkylated and N-alkylated DD, respectively (67).

2) Base excision repair (BER)

Base lesions that do not highly distort the DNA helix structures are corrected through BER. This type of damage typically arises from deamination, oxidation, or methylation. Firstly, DNA glycosylase recognizes and removes the damaged base. Next, different sets of proteins are recruited to the abasic site to restore it through short- or long patch BER (68).

B) Bulky DNA damage repair

1) Nucleotide excision repair (NER)

This repair system is involved in the correction of DNA helix distorting damages. Through NER system, DNA lesions are recognized and incised on both sides. Then, the undamaged complementary strand is used as template for the synthesis of a new patch. Finally, the patch is ligated in adjacent strand (69).

2) Mismatch repair (MMR)

Errors such as incorrect base substitutions and insertion-deletion mismatches frequently occur during replication. Although DNA polymerases correct these errors, some of them

escape from such proofreading activity. MMR is a conserved pathway from bacteria to human which recognizes the escaped errors and corrects them (70).

3) Inter-strand crosslink repair

As discussed before, ICL are lesions where two bases from the opposite strands covalently bind to each other. Fanconi anemia (FA) proteins play a key role in recognizing these lesions. In response to DD two of these proteins namely, FANCM and MFH stimulate fork remodeling and generate ssDNA gaps which are followed by RPA recruitment and ATR-CHK1 signaling activation. Subsequently, activation of FA-associated molecular cascades downstream to ATR take place which finally lead to the excision of DNA lesion and its repair by translesion synthesis (TLS) polymerases (41).

4) Translesion synthesis

Stalled RFs have harmful consequences for the cells. To avoid these, cells have evolved specialized polymerases to traverse the damage. These polymerases called as TLS provide the cells with extra time to repair the damage however, their activity is error-prone and is associated with the increased risk of mutagenesis (71).

C) Repair of DNA breaks

1) Single-stranded break repair (SSBR)

SSB can occur either directly (e.g., attack of ROS) or indirectly (e.g., intermediates which arise during BER). SSBR is divided into four main steps. In the first step SSBs, which are indirectly generated through AP nuclease I (APE1) activity during BER process are detected in a poly (ADP-ribose) polymerase 1 (PARP-1)-independent manner in contrast to direct breaks which require PARP1 for detection. In the second step breaks goes under end processing and are repaired by APE1, DNA polymerase β, PNKP and APTX. In the following step DNA polymerase β alone or together with DNA polymerases δ/ε fill the gaps during short- and long-patch repair, respectively. The final step is ligation which is accomplished by DNA ligase 3 in short-patch repair and by DNA ligase 1 in long-patch repair (72).

2) Double-stranded break repair (DSBR)

DSBs are repaired through two main pathways: nonhomologous end joining (NHEJ) and homologous recombination (HR). In NHEJ, 53BP1 plays a key role in recruiting the

components of repair to the break site. The Ku70/Ku80 heterodimer is the first molecule which binds to the break sites to prevent the end resection and also acts as a scaffold for other NHEJ components. Next, the DNA-PKcs, DNA ligase 4/XRCC4/XLF complex and others bind to repair the DNA. NHEJ is error-prone as this pathway does not use a homologous template to repair DNA in contrast to the HR pathway (73). In HR, the MRN complex binds to the break and recruit ATM and TIP60. ATM phosphorylates H2AX and accordingly MDC1 which act as a scaffold for ubiquitin E3 ligases RNF8 and RNF168. Both ligases ubiquitinate H2AX which provides a docking site for 53BP1 and BRCA1. HR predominantly takes place in G2/S phase where BRCA1 ubiquitinates CtIP. Concomitantly, this facilitates the recruitment of other HR components such as RPA and RAD51 to DNA. Subsequently, a set of molecules create a 3' overhangs which commit the DNA to undergo HR. Then, RAD54 removes RAD51 which leads the strand to invade the DNA template and initiates synthesis with the aid of polymerases δ, κ and ν (41).
Ultimately, if cells cannot repair the damage, they activate chronic DDR signaling to induce apoptosis or senescence (**Figure 1**) (74).

Cross talk between the cell cycle checkpoints in DDR

ATM and ATR, the upstream DDR kinases, have distinct functions and specificities in DD. ATM is activated in response to DSBs whereas, ATR plays a role in various types of DNA lesions including DSBs and SSBs. Therefore, these two kinases have cross talk at multiple levels. For example, ATM activates ATR at DSBs conversely, ATR possibly recruits ATM through phosphorylation of H2AX at stressed RFs. Moreover, ATM and ATR may phosphorylate each other, alter the function and subcellular localization of each other's regulators or share similar substrates. Additionally, these apical kinases may act equivalently and can be functional in the absence of each other. For instance, ATR could be slowly activated in the absence of ATM at DSBs. On the other hand, collapsed RFs lead to the activation of ATM when ATR signaling is impaired (75).

Replication stress response (RSR)

Formation of ssDNA at RF is a common structural determinant of RS. To solve the stress, RPA rapidly covers ssDNA which recruits ATR-interacting protein (ATRIP) and the DD checkpoint, ATR. Moreover, RPA loads Rad9–Hus1–Rad1 (9-1-1) complex through Rad17 recruitment. The 9-1-1 complex is required for the positioning and activity of TopBP1, which harbors the ATR activating domain (AAD). In parallel ETAA1, another RPA-interacting protein, activates ATR through its AAD domain. Consequently, ATR activates Chk1 and other proteins (**Figure 1**) which are necessary to stabilize the stalled fork and to

ensure rapid initiation of DNA synthesis. ATR slows down the RF progression and prevents the formation of long ssDNA by promoting the interaction of FANCD2 and MCM helicase. Next, downstream to the lesion the proteins with repriming ability initiate DNA synthesis and leave the ssDNA gap behind. Later, the gap will be filled through homology-directed repair (HDR) - mediated process or TLS polymerases which is directed by PCNA mono- or poly-ubiquitination, respectively. If left unreplicated, the ssDNA gaps could be converted to DSBs therefore, the post-replicative gap repair is crucial for genome stability (1, 76, 77).

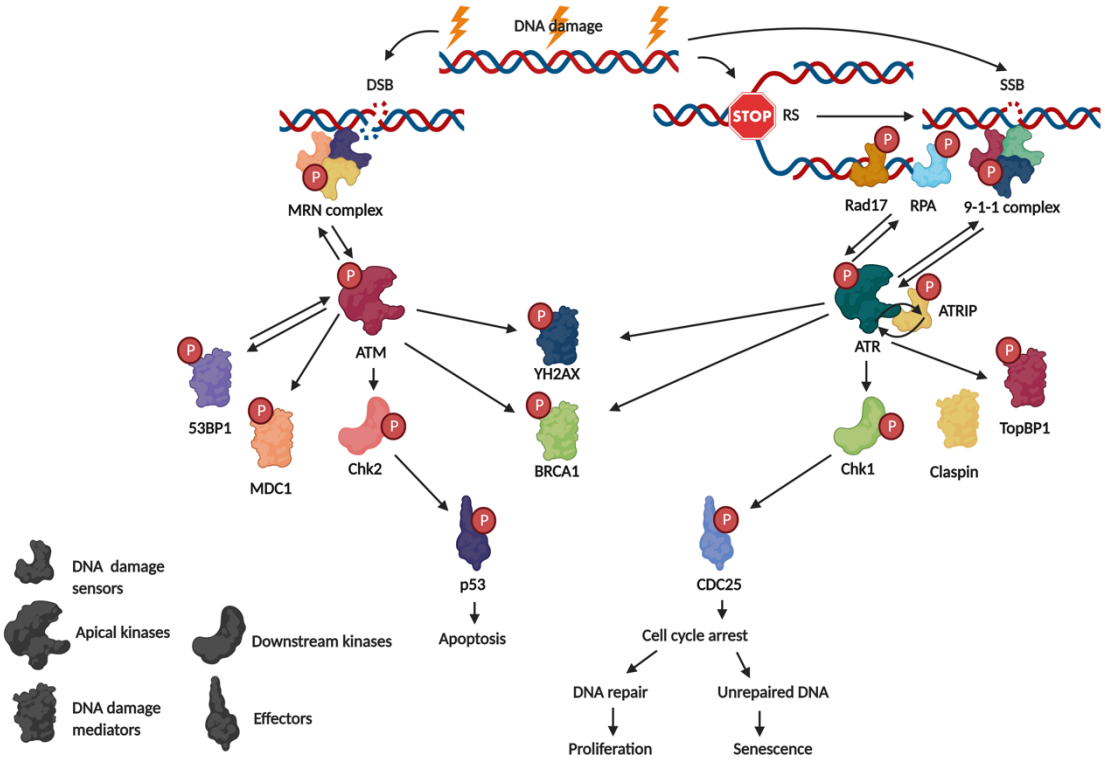


Figure 1. DNA damage response pathways

Response to DD is initiated through two main groups of sensors. The first group is 9-1–1 complex, RPA and Rad17 which detect ssDNA. The second group is MRN complex that detects DSBs. These sensors recruit the apical kinases ATR (which is bound by ATRIP) and ATM, respectively to the site of damage. These kinases, in turn, phosphorylate H2AX on Serine 139. ΓH2AX-mediated recruitment of MDC1 further amplifies DDR signaling through enhanced accumulation of the MRN complex and activation of ATM. Moreover, BRCA1 is brought at sites of damage and is phosphorylated by both ATM and ATR. This phosphorylation is required for the activity of HR components. 53BP1 is also involved to sustain the ATM-mediated DDR signaling. Consequently, ATM and ATR mainly phosphorylate downstream kinases, Chk2 and Chk1 respectively. Eventually, DDR converges on downstream effectors such as P53 and CDC25. The outcome of DDR activation is cell death by apoptosis or transient cell cycle arrest followed by DNA repair. Alternatively, in case of persistent DD the cells undergo senescence. (Based on the figures

from Sulli *et al.*, Nature reviews cancer, 2012 and Caldon, Frontiers in oncology, 2014) (78, 79).

Diseases associated with replication stress

Proper cellular response to RS is fundamental for individual's health. Otherwise, any defects in signaling pathways that respond to RS leads to various genetic disorders or cancer. Mutations in ATM and ATR checkpoints cause severe diseases called as ataxia-telangiectasia (AT) and Seckel syndrome, respectively. The AT patients have impaired voluntary movement, red lesions in skin and mucous membranes (due to the widening of the blood vessels), weakened immune system (which makes them susceptible to upper and lower respiratory infections) and lymphoid malignancies. Seckel syndrome is caused by hypomorphic mutation in *ATR* allele which leads to reduced expression of ATR protein. These patients suffer from microcephaly and mental retardation and have developmental defects. The mouse models of ATM and ATR deficiency have been developed. Despite their functional overlap in response to DD, ATR is indispensable for viability however ATM is not. Therefore, $ATM^{-/-}$ mouse model is viable while complete knock out of ATR is lethal. To circumvent this issue, Fernandez-Capetillo and colleagues developed the humanized model of ATR-Seckel syndrome mouse ($ATR^{s/s}$), which faithfully recapitulates the molecular behavior of mutated ATR in human. Beside showing severe symptoms characteristics of Seckel syndrome in mouse (e.g., dwarfism and microcephaly), interestingly, they found that mutant placentas have accumulated necrotic areas and overall loss of cellularity. Nijmegen breakage syndrome (NBS) and Nijmegen breakage syndrome-like disease (NBSLD) are other diseases caused by mutation in *Nbs1* and *Rad50* genes, respectively, which compromise the MRN complex role in ATR activation or HR. Mutations in FA proteins cause Fanconi anemia which skeletal defects, hypopigmentation, progressive bone-marrow failure, and cancer predisposition are its characteristics. Additionally, mutations in TLS polymerase lead to Xeroderma pigmentosum (XP), which the patients are vulnerable to skin cancer due to the photosensitivity. Beside genetic disorders, RS is tightly linked to cancer predisposition. Mutation of genes that code origin licensing, fork stability, TLS, DNA repair proteins and etc. are associated with cancer. Moreover, it has been reported that fragile sites are related to cancer predisposition (80-83). For the detailed list of diseases associated to RS, readers are suggested to refer to the cited reviews (14, 55).

An overview of the mouse pre-implantation embryo development

The pre-implantation stage in mammals, which begins from egg fertilization to implantation of the embryo in the uterus, is a critical period because the primary lineages segregation

takes place within this stage. Taking advantage of these knowledges in clinics, scientists could increase the success of assisted reproductive strategies (ARTs) such as *in vitro* fertilization (IVF) and preimplantation genetic diagnosis (PGD). They could also decrease the rate of early pregnancy loss and improve the derivation of stem cell lines from embryos. Our understanding about pre-implantation development mostly comes from the studies in the mouse that has been used more than a half-century as a faithful model for the human embryo (84, 85).

Early cleavage and zygotic genome activation

In the first step, the fertilized egg divides into small cells of blastomeres through a series of early cleavage (**Figure 2**). Initially, the zygote relies on the maternally deposited messenger RNA (mRNA) for protein synthesis however, after few divisions the transcription of mRNA from the zygotic genome begins. This transition is known as zygotic genome activation (ZGA) which reaches to its highest level at two-cell (2C) stage, leading to the degradation of maternal transcripts. In spite of maternal mRNA degradation, their synthesized proteins can remain at later stages which intervene with gene function analysis during pre-implantation development (84, 85).

Compaction and polarization

The blastomeres in 8C embryo go under compaction (**Figure 2**) through secretion of intercellular adhesion proteins which is an essential process for the proper lineage segregation. E-cadherin is a major component of adherens junctions, and its disruption prevents compaction. Concomitantly, the blastomeres polarize in such a way that the apical and basolateral regions become distinct and cytoplasm become reorganized. For example, Actin, polarity proteins (Par3 and Par6), and atypical PKC (aPKC) accumulate apically while basal proteins such as Par1 and lethal giant larva homolog (Lgl) localize in the basolateral domain. It seems that the cell-cell contact as well as the mutual antagonism of apical and basal proteins establish the embryo polarization (84, 85).

Asymmetric and symmetric cell divisions

To generate 32C embryo, the compacted and polarized cells of 8C embryo divide for two more rounds. The cleavage of cells can be either perpendicular (symmetric) or parallel (asymmetric) to the axis of cell polarity, which results in two daughter cells. In case of symmetric division, these cells will be located on the outside layer, while in asymmetric division one of the cells will remain on the inside of the embryo and other on the outer layer.

These divisions transform the embryo with uniform cells to the embryo with two distinct cell populations (i.e., apolar inside cells and polar outside cells) (**Figure 2**) (84, 85).

Blastocoel formation

Once embryo reaches to 32C stage, a fluid-filled cavity known as the blastocoel begins to form, which is essential for the proper development of the inner cell mass (ICM). Maintenance of blastocoel depends on the tight junctions in the outer layer of the embryo i.e., trophectoderm (TE) which form a seal to prevent water leakage. Subsequent to the formation of blastocoel at embryonic day 3.5 (E3.5), the embryo is considered as blastocyst and is ready to implant into the uterine wall within the next 24 hrs. (**Figure 2**) (84, 85).

Developmental flexibility of the early embryo

Early mammalian embryos are highly adjustable during the first three cell divisions, meaning that they can withstand any modifications including addition, removal and rearrangements of blastomeres. It is reported that the blastomeres are highly plastic and show little developmental biases. However, as soon as the emergence of the three lineages of blastocyst -namely TE, epiblast (EPI) and primitive endoderm (PE))- the developmental flexibility of the embryo diminishes (84, 85).

The first lineage decision: segregation of the TE and ICM

The first cell lineage that becomes specified in pre-implantation embryo is TE. During the 32C stage in mouse embryo, the cells become fully committed to either TE or ICM (**Figure 2**). There are two proposed models that explain how the first cell fate decision happens. First, is the inside-outside model which suggests that the fate of cells is related to their position, inside or outside layer, since they are exposed to different microenvironment and amount of cell contact in different positions. Second, is the polarity model which links the cell fate to the asymmetric cell division and generation of polar and apolar daughter cells. Irrespective of the proposed models, transcription factor (TFs) and signaling pathways certainly play a key role in cell fate specification. Cdx2 is a specific TF required for TE development and Octamer 3/4 (Oct4), Nanog, and SRY-box containing gene 2 (Sox2) are pluripotency markers which determine the ICM fate. The expression level of *Cdx2* mRNA is different in blastomeres but prior to blastocyst formation it becomes restricted to the outside cells. The same pattern is observable for pluripotency marker however, they become confined in inside cells after the blastocyst formation. Therefore, the first cell fate decision takes place through upregulation of Cdx2 and downregulation of Oct4, Nanog and Sox2 in outside cells (85-87).

The second lineage decision: segregation of the PE and EPI

This lineage segregation takes place within ICM after its specification from TE. PE is a monolayer of cells that forms at the interface of blastocoel cavity with EPI. The separation of PE and EPI occurs at E4.5 when they become committed and morphologically distinguishable (**Figure 2**). Gata4 and Gata6 are specific TFs of PE however, Nanog is EPI-specific marker. At E3.5 most cells randomly express either Gata6 or Nanog in the pattern of salt and pepper irrespective to their cell position. It seems that later these cells will be located in their right position according to the expressing TF (84-86).

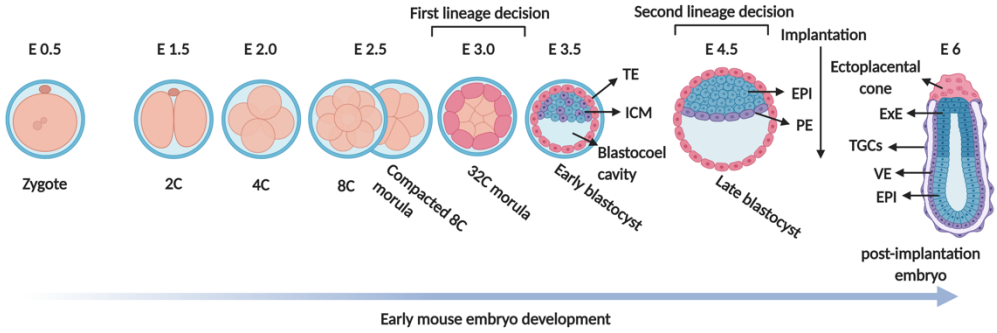


Figure 2. Early stages of the mouse embryo development

The zygote undergoes three rounds of cleavage to generate an 8C embryo. Next, the blastomeres in 8C embryo go under compaction through secretion of intercellular adhesion proteins. Then, the compacted and polarized cells of 8C embryo divide for two more rounds to generate 32C morula. During this stage, the first lineage decision takes place which leads the cells to become fully committed to either trophectoderm (TE) or inner cell mass (ICM) at E3.5. Moreover, the blastocoel cavity begins to form at 32C stage and continues to expand till late blastocyst stage. At E4.5, the second lineage segregation takes place within ICM and cells become committed to epiblast (EPI) or primitive endoderm (PE). Following the formation of the late blastocyst, the trophoblast cells start to invade into the uterus. During implantation, the mural TE cells form early trophoblast giant cells (TGCs) while the polar TE cells form diploid extra-embryonic ectoderm (ExE) cells. The ExE cells that are located at the farthest point from epiblast, further differentiate into the ectoplacental cone (EPC). PE cells further proliferate and form the visceral endoderm (VE) in post-implantation embryo. E: embryonic day, C: cell (Based on the figures from Cockburn and Rossant, JCI, 2010 and Sauvegarde *et al.*, PLOS ONE, 2016) (84, 88).

Hippo signaling pathway and TE segregation

One of the signaling pathways that plays a central role in TE specification is Hippo signaling pathway. During polarization, this pathway becomes active in the inner cells of the embryo. Briefly, in the inner cells Amot phosphorylates large tumor suppressor 1 and 2 (Lats1/2) which promotes its kinase activity. Consequently, Lats1/2 phosphorylates Yes-associated protein 1 (Yap1) which leads to its cytoplasmic localization. Conversely, this pathway is off in the outer cells. In other words, the components of polarity pathway retain Amot at the apical domain so that it cannot bind to Lats1/2 and thus, Yap1 remains unphosphorylated. The unmodified Yap1 localizes in the nucleus and acts as a co-factor for the transcription factor TEA domain family member 4 (Tead4) which is the master regulator of TE-specific transcription network. In other words, Tead4 activates the expression of Cdx2 and accordingly leads to the TE cell fate determination (89). Importantly, Tead4-null embryos cannot reach to the blastocyst stage, and their blastomeres lack the expression of canonical TE TFs such as Cdx2, Eomes and Gata3 while, they normally express ICM-specific TFs, Nanog and Oct4. In another study, Home *et al.* show that the subcellular localization of Tead4 plays a key role in the first lineage segregation in mammalian embryos. Through immunofluorescence staining of embryos at different developmental stages, they traced the subcellular localization of Tead4 protein. In 8C embryos Tead4 was abundantly found in the nucleus of blastomeres. Along with compaction, nuclear localization of Tead4 confined to the outer cells while, its localization in inner cells were mostly cytoplasmic rather than nuclear. Subsequently, in the ICM of blastocysts Tead4 were only found in the cytoplasm. The absence of Tead4 in the nucleus of ICM disrupts its chromatin occupancy leading to decreased expression of its mRNA as well as its target genes (**Figure 3**) (90).

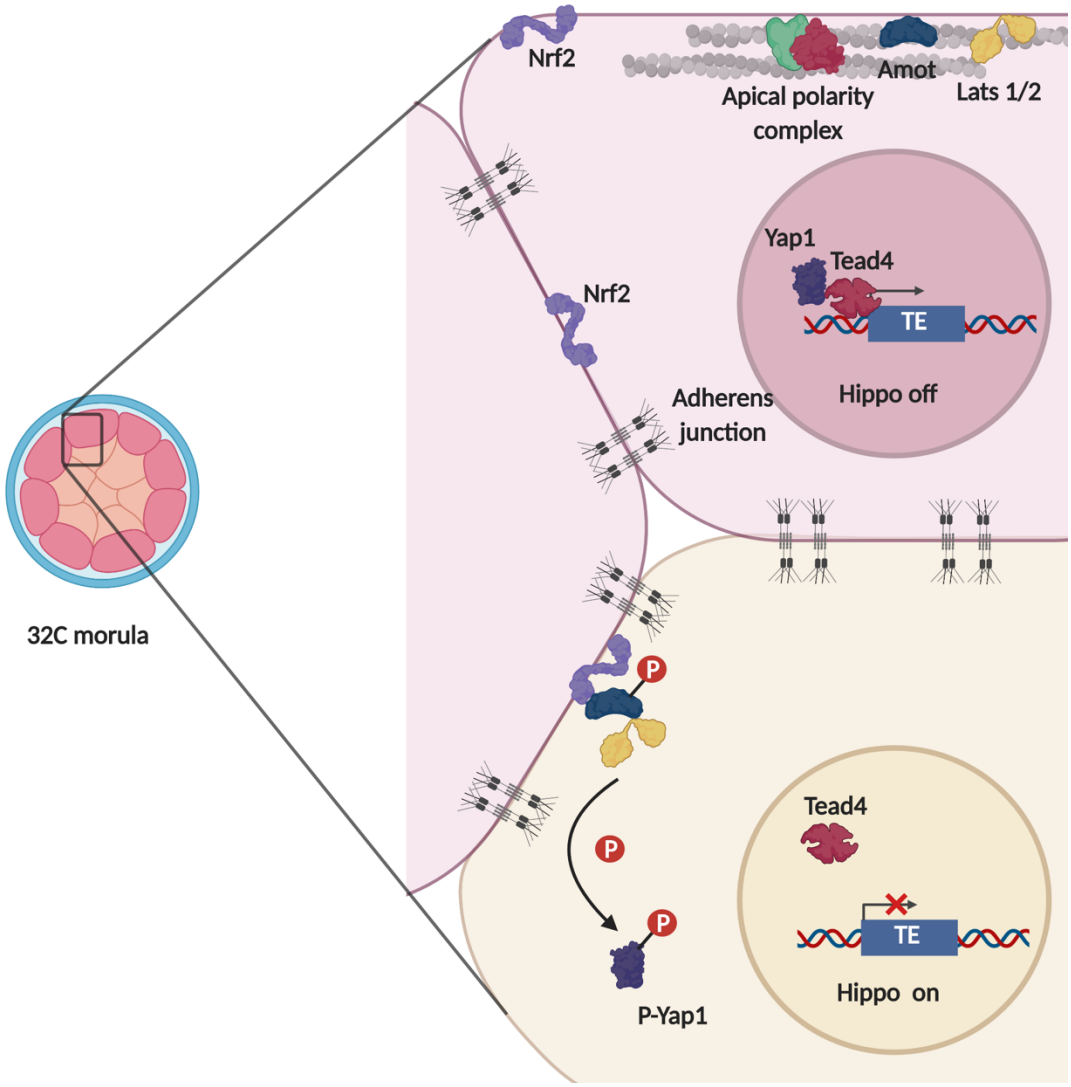


Figure 3. Hippo signaling pathway in early-stage mouse embryo

In the outer cells, the apical polarity complex retains Amot at the apical domain and thus prevents its activation. Consequently, Lats1/2 cannot be activated hence cannot phosphorylate Yap1. The unmodified Yap1 localizes in the nucleus and acts as a co-factor for Tead4. Through activated Tead4 binding to the TE transcription factor binding motifs (TFBMs), the expression of such genes increases which leads to the TE fate in the outer layer of the embryo. Conversely, in the inner cells, Amot localizes at the adherens junctions where it binds to Lats1/2 and Nrf2. This binding promotes the kinase activity of Lats1/2 which consequently it phosphorylates Yap1. The phosphorylated Yap1 is retained in the cytoplasm and goes under proteasomal degradation. Therefore, Tead4 remains inactive and cannot bind to the TE TFBMs in the inner cells. As a consequence, the inner cells acquire the ICM fate. (Based on the figures from White and Plachta, ELSEVIER, 2015 and Mo *et al.*, EMBO reports, 2014) (89, 91).

An overview of the mouse early placental development

Trophectoderm formation in blastocyst is the first emergence of future placental cells as they are able to give rise to all cell types of the placenta. The majority of placental cells descend from trophoblast cells however, the endothelial cells of fetal placental vasculature (allantois and umbilical cord) originate from extra-embryonic mesoderm which emerges at gastrulation from ICM. Following the blastocyst formation, the mural TE cells which encompass the blastocoel cavity go under endoreduplication to form early (primary) trophoblast giant cells (TGCs). These cells are responsible to penetrate the uterine and implant the embryo into the endometrium. Subsequently, the polar TE cells adjoining the ICM, proliferate in response to fibroblast growth factor four (FGF4) signal arising from the EPI and form diploid extra-embryonic ectoderm (ExE) cells. While retaining the trophoblast stem cell (TSC) potency, the ExE cells that are located at the farthest point from epiblast, further differentiate into the ectoplacental cone (EPC) (**Figure 2**). Next, the EPC marginal cells differentiate into secondary TGCs with invasive characteristics, therefore they can deeply penetrate into the endometrium to make a connection with maternal arteries. Trophoblast invasion starts from E4.5 onwards and peaks between E7.5 to E9.5.

The progressive understanding of extra-embryonic lineage development in the mouse is appreciable because, it can be used as a model to expand our knowledge in the role of extra-embryonic lineage development in pregnancy. Moreover, it is instrumental to trace the cell lineage origins of particular defects (92).

DNA damage response in early mouse embryo

The majority of embryonic loss takes place around the implantation phase. This time frame, which is mainly dedicated to the formation of extra-embryonic tissues, is called as pre-implantation period. The highest rate of embryo loss in this period indicates the vulnerability of rapidly proliferating pre-implantation embryos to genotoxic insults and the necessity of efficient stress response during this time frame. Therefore, the early embryonic development is a critical stage as any unrepaired damage can give rise to either miscarriage or later prenatal abnormalities. After fertilization, DNA repair in embryos completely relies on maternal mRNA and proteins which their insufficiency leads to early embryonic loss. Recent different studies aimed to determine the onset of zygotic DDR and point out that some of DDR pathways are defective before morula (16C) stage. This is probably for the reason that the necessity of the decreased rate of cell cycle progression to respond to DD is incompatible with highly proliferating characteristic of embryonic cells. Mutant mouse embryos show that zygotic DDR pathways are dispensable before implantation and even some after implantation as ATM or Chk2-deficient mice are viable. Conversely, embryos with ATR or

Chk1 deficiency show high level of DNA fragmentation and cannot remain after implantation. However, the mutated peri-implantation embryos have normal cell cycle progression, possibly thanks to the maternally inherited proteins. Beside cell cycle checkpoints, it is reported that HR pathway but not NHEJ pathway is essential for DSBR during pre-implantation development, since mutation in the genes related to the HR pathway such as *Nbs1* and *Rad50* is demonstrated to be lethal for the embryos (93-96). Nonetheless, the fate of embryonic cells under RS or DD during the early developmental stage is still remained to be elucidated.

Mouse embryonic stem cells as a tool to study early embryonic development

Mouse embryonic stem cells (mESCs) are derived from the EPI compartment of pre-implantation mouse blastocyst which finally gives rise to the embryo proper *in vivo*. ESCs can self-renew infinitely *in vitro*, grow clonally, recapitulate the developmental properties of EPI in culture and retain pluripotency, meaning that they can differentiate into all types of adult cells as well as the germ cells. Targeted genomic manipulation of ESCs and generation of mutant mice from these cells made a revolution in studying the gene function in mammalian biology and diseases. Therefore, ESCs can be regarded as a useful tool to study early development. ESCs were initially cultured on mitotically inactivated fibroblasts (MEFs) as a feeder layer in a media supplemented with serum. Later, it was found that the MEF layer secrets leukemia inhibitory factor (LIF) which activates STAT3 required for ESC self-renewal. Addition of two inhibitors (2i) CHIR99021 and PD0325901 (the Gsk3 and mitogen-activated protein kinases (MAPKs) inhibitors, respectively) in the media supplemented with serum and LIF facilitated the derivation and maintenance of naïve pre-implantation epiblast-like state cells from various mouse strains (97, 98).

Induced pluripotent stem cells (iPSCs)

iPSCs are the cells that are generated from mouse somatic cells (such as fibroblasts) through introduction of four TFs namely, Oct4, Sox2, Klf4 and Myc (Yamanka factors). Discovery of iPSCs was a breakthrough in science and medicine, because these cells have gene expression profile and developmental potential similar to the ESCs. iPSC technology provides a new era in the field of the stem cell-based therapy, regenerative medicine, disease modelling as well as drug discovery. iPSCs especially in three-dimensional (3D) culture system (organoids) can be used to study the pathological mechanisms of diseases, to identify effective drug candidates for the related disease and to assess their potential toxicities. Therefore, patient-specific iPSCs have the potential to develop the field of precision medicine (99).

Maintenance of the genomic stability in ESCs

mESCs are very sensitive to DD and to eliminate any damaged cells they either go under apoptosis or differentiate, but they do not become senescent. It is reported that ESCs under genotoxic stress cannot be arrested in G1 phase due to the compromised ATM-Chk2-mediated phosphorylation of Cdc25A. This property of ESCs is partly due to the Chk2 sequestration at centrosomes which makes it inaccessible to phosphorylate its substrates. Moreover, cytoplasmic localization of p53 restricts its effect on p21 accumulation. Therefore, in the absence of G1 arrest, unrepaired ESCs enter the S phase where damage could be worsened through a round of replication. Another study has revealed that ESCs harbor massive amount of ssDNA gap accumulation, reduced RF progression, frequent reversed forks and high basal level of ATR-mediated H2AX phosphorylation. The high level of γH2AX is important for the ESC self-renewal and proliferation and is not linked to the DSBR pathway however, this peculiarity is not clearly defined. Despite these features, ESCs have efficient repair mechanisms. Of note, these cells have elevated level of BER- and MMR-mediated proteins. However, in response to the high level of DD, ESCs halt the activity of NER pathway possibly to rapidly remove the damaged cells through apoptosis. Interestingly, since ESCs spend most of their time in the S phase, they predominantly use the HR-mediated pathway to repair DSBs rather than NHEJ (100, 101).

Experimental models to study the early embryonic events

A) *In vivo* mouse embryonic chimera assay

Mouse embryonic chimeras are a powerful tool to study the cell potential and lineage mutations. They can be produced through various combinations of embryonic cells i.e., aggregation of two whole 8C embryos or subsets of blastomeres from two distinct embryos or ICM with 8C embryo. In the first two types, chimerism occurs in the EPI, PE and TE. However, in the latter type due to the restricted lineage potency, chimerism takes place in EPI and PE but not TE. Beside aggregation, chimeras can be generated through introduction of foreign cells into the embryos. In this approach mostly ESCs are microinjected into the morula or blastocyst. These cells are able to contribute to the embryonic compartment however, they cannot contribute to the PE and TE. While microinjected trophoblast stem cells (TSCs) which are derived from TE of the blastocyst or from early post-implantation trophoblasts, can contribute only to the TE. Subsequently, chimera generation using genetically modified ESCs is an invaluable tool to precisely and rigorously study the biological effects of genetic alterations. Currently, both *in vitro* cell differentiation approaches and *in vivo* chimera assay are used to study the developmental potential of ESCs.

Nevertheless, the embryonic chimera provides a unique environment for injected cells to contribute into all possible lineages. In other words, through chimera assay the full range of lineage potency could be tested and in conclusion, the true extent of cell potency could be revealed. In spite of this, studying the cell through developmental potential chimera assay is associated with several limitations including practical and ethical challenges (102, 103).

B) ESCs *in vitro* differentiation to early embryonic lineages

1) Differentiation of ESCs into the primary germ layers

Beside the high capacity of ESCs for genetic modification, they provide a model of early mammalian development in culture. Furthermore, the differentiated cell types generated from ESCs could be a faithful source for cell replacement therapy. Removal of the factors from the culture which maintain ESCs at pluripotent state leads to their differentiation to three derivatives of embryonic germ layers namely, ectoderm, endoderm and mesoderm. There are three general approaches for ESC differentiation. In the first method ESCs are cultured in a monolayer on extra-cellular matrix proteins. In the second approach, ESCs differentiation takes place through direct contact with a feeder layer of stromal cells. In the last method, ESCs are directed to form 3D aggregates called as embryoid bodies (EBs). The 3D structure of EBs provide the advantage of cell-cell interactions that maybe is required at certain developmental stages (104).

2) Differentiation of ESCs into the TSCs

Although ESCs have potency to differentiate into the primary germ layers, due to the restricted developmental potential, they cannot readily differentiate into the TE lineage (104). This means that the trans-differentiation of ESCs to trophoblast lineages requires the episomal overexpression (OE) of early trophoblast TF such as Tead4, Cdx2, Eomes and Elf5 or downregulation of ESCs TFs such as Oct4. Beside the necessity of regulating specific TFs to initiate differentiation, the activation or inhibition of key signaling pathways required for differentiation is important in the extracellular environment of cultured cells (105). TSCs were successfully derived for the first time in 1998 from embryos in a media supplemented with FGF4, the signaling pathway required for the proliferation and maintenance of trophoblast cells *in vivo* (106). Afterwards, several groups attempted to differentiate ESC to TSCs using the established medium (TSCM). In the following the approach of these groups will be explained concisely.

- Niwa *et al.* could generate TSC-like cells (TLSCs) through OE of Cdx2 or repression of Oct4 in their ESCs. Briefly, they introduced the tetracycline (Tc) -regulable *Oct4* transgene in ESCs which both endogenous *Oct4* alleles are disrupted to maintain ESC self-renewal. Through administration of Tc, they could successfully inhibit Oct4 expression and obtain TLSCs. Likewise, when they introduced the expression vector of the 4-hydroxy tamoxifen (Tx)-inducible *Cdx2* in their cells, they observed the emergence of TLSCs upon Tx addition (107).
- Ralston *et al.* identified Gata3 as trophoblast TF and through its OE, they could induce trophoblast fate in ESCs under TSC derivation conditions. However, they pointed out that the induction of Cdx2 generates more stable and homogenous TLSCs in comparison to Gata3 OE (108).
- Nishioka *et al.* targeted Tead4, the master regulator of TE-specific transcriptional program. Through stable expression of Tx-inducible vector consisting Tead4 DNA-binding domain in ESCs, they could generate TLSCs (109).
- Through a genome-wide screening, Ng Kit *et al.* identified Elf5, the specific TF of trophoblast, as methylated and repressed in ESCs in contrast to TSCs. Through deletion of DNA Methyltransferase (DNMT), they could successfully differentiate ESCs toward TSCs in TSCM (110).
- Erk2 is the downstream effector of Ras-MAPK signaling, which its inhibition in cultured mouse embryos reduces the expression of Cdx2 and disrupts the derivation of TE. Through OE of this pathway in ESCs, Lu *et al.* could generate TLSCs (111).
- During activation of Ras-MAPK signaling in ESCs, the upregulation of Tcfap2c TF was found in generated TLSCs. Taking advantage of this, Kuckenberg *et al.* could induce the TSC fate in ESCs through OE of Tcfap2c (112).
- Abad *et al.* obtained *in vivo* iPSCs through transitory induction of Yamanaka factors in their generated reprogrammable mice. They reported that *in vivo* iPSCs in comparison to ESCs generate higher number of colonies with TLSC morphology when they were plated in TSCM (113).

Although in all models of ESC-to-TSC differentiation cells with TSC-like characteristics emerges, the in-depth analysis shows that these cells are not phenotypically, transcriptionally, epigenetically and functionally completely identical to the embryo-derived TSCs. This means that the lineage conversion remains incomplete even in the most efficient differentiation models. This is because some groups of TSC genes fail to completely become demethylated in the emerged TLSCs. Even forced expression of these non-reprogrammed

genes cannot grant a stable TSC phenotype. Hence, the generated TLSCs retain the epigenetic memory of their ESC origin. Nevertheless, ESCs are the most plastic SC type and therefore, they seem to be the most flexible cells for the differentiation to TE lineages (105).

In all models of ESC differentiation, three criteria should be considered. First, the development of the cell type of interest should be efficient and reproducible through the established approach. Second, if required, the selection tools should be used to isolate the highly matured cells. Third, the differentiated cells must be functional similar to the bona fide cells derived from the animal both in culture and when they are transplanted into the animal models (104).

Hippo signaling pathway in ESCs and TSCs

Similar to the embryos, Hippo signaling pathway is active in ICM-derived ESCs while is inactive in TE-derived TSCs. Home *et al.* have reported that the mRNA and protein levels of Tead4 in mouse ESCs are low compared to TSCs (90). This is in consistent with the above-mentioned study which shows that OE of Tead4 induces trophoblast fate in ESCs (109). This group detected Tead4 only in the cytoplasm of ESCs while in TSCs they found it both in the nucleus and cytoplasm with higher abundance in the nucleus. Through ectopic OE of Tead4 in ESCs, they observed its localization into the nucleus, which subsequently led to Tead4 chromatin occupancy at TE-specific genes (90). On the other hand, there are several other studies showing that Tead4 is constantly nuclear in the outer and inner cells of the pre-implantation embryo, and that the Yap1 but not Tead4 cytoplasmic or nuclear localization defines the embryonic or extra-embryonic cellular fate. Concisely, nuclear Yap1 is evenly detectable at 8C blastomeres. Following embryo development, Yap1 increases in the nucleus of outside cells while it gradually decreases in the nucleus of inner cells until it becomes undetectable in the ICM of mid/late blastocysts. In the embryo inside cells, Lats1/2 mediates the phosphorylation of Yap1 which leads to its retention in the cytoplasm thus, Tead4 remains inactive and consequently inside cells adopt the ICM fate. Conversely, in the outer cells Yap1 molecules are mostly unphosphorylated, therefore they localize in the nucleus and activate Tead4 which finally lead to the TE fate (109, 114).

3) ESCs *in vitro* differentiation to the placental giant cells

Following the establishment of TSCM and successful derivation of TSCs from embryos, Tanaka *et al.* reported that upon removal of FGF4, the proliferation rate of TSCs decreases and they differentiate toward TGCs, the fundamental cells of the placenta. They also pointed out that even in the presence of FGF4 some TGCs consistently emerge at the borders of

TSCs colonies. Since these cells are trypsin-resistant, they are left behind following each passaging therefore, their number remains low (106). Several groups took advantage of this approach and reported the formation of trophoblast like giant cells (TLGCs) from TLSCs which they generated through OE of TSC-specific genes (as discussed previously).

- After generation of TLSCs through OE of Cdx2 in ESCs, Tolkunova *et al.* differentiated these cells toward TGCs upon withdrawal of FGF4 and MEF conditioned medium (115).
- Kuckenberg *et al.* reported the generation of TLGCs from ESC-derived TLSCs, which they previously generated through OE of Tcfap2c (112).
- To further characterize the differentiation potential of *in vivo* iPSCs, Abad *et al.* differentiated *in vivo* iPSCs-derived TLSCs toward TGCs. They could observe the emergence of TLGCs similar to the TSC-derived TGCs (113).
- Another study on ESC differentiation toward placenta cells is done by Hemberger *et al.* It has been reported that ESCs which are deficient in PARP1 ($Parp1^{-/-}$) form cells with TGCs morphology in teratocarcinoma-like tumors when they are injected subcutaneously into the mice. As explained previously, PARP1 plays an important role in the maintenance of genome integrity. PARP1-deficient mice are viable and fertile. However, they are very sensitive to alkylating agents and irradiation. To further investigate this observation, Hemberger and her colleagues checked the expression of TGCs markers in $Parp1^{-/-}$ tumors and confirmed the activation of TGC-related genes. Surprisingly, they reported that similar to the $Parp1^{-/-}$ ESCs, wild type (WT) ESCs form a very small population of TLGCs in tumors. Strikingly, they noticed the generation of TLGCs at the margins of both $Parp1^{-/-}$ and WT ESCs in the culture. Through in situ hybridization, they detected the expression of TGC-specific genes (*Pl1*, *Plf*, *Plfr*, *Prlpa*, and *Pl2*) in both WT and $Parp1^{-/-}$ ESCs although with significant higher level in the latter. Therefore, they suggest that the well-established ESC lines possess some cells with differentiation potential to the TE (116).

4) *In vitro* modeling of early mammalian embryogenesis

Mammalian embryogenesis events from the first cell divisions to axial patterning and later cellular differentiation take place in the uterus, which makes the embryos (especially in human) inaccessible to direct observation. Therefore, due to the complexity of studying this process, scientists have been attempting to establish *in vitro* systems to better inspect the embryonic development in a simplified and controlled manner (117).

Signaling pathways in early post-implantation development; from blastocyst implantation to gastrulation

Embryonic development requires the cooperative interactions between embryonic and extra-embryonic tissues. In the blastocyst the EPI will generate the fetus while the TE and PE give rise to the placenta and the yolk sac, respectively. Following implantation, a series of events take place that lead to the architectural reorganization of the embryo and the formation of the egg cylinder. Briefly, the PE-derived visceral endoderm (VE) secretes extra-cellular matrix (ECM) which initiates lumenogenesis in EPI and TE-derived ExE. Ultimately, the cavities generated in EPI and ExE merge to form pro amniotic cavity. Next, the crosstalk between Nodal, Wnt, and bone morphogenetic protein (BMP) signaling pathways breaks the symmetry of the embryo at the boundary of embryonic and extra-embryonic layers to induce the mesoderm and primordial germ cell (PGC) formation. In more details, the distal part of the VE (DVE) releases the inhibitors of the Nodal and Wnt signaling towards the anterior VE (AVE). While at the posterior section of the embryo the ExE expresses BMP4 to reinforce the activity of Nodal and Wnt in the EPI. Nodal induces BMP4 which in turn increases the activity of Wnt in a positive feedback loop. Collectively, these events specify the primitive streak and mesoderm formation at the posterior site of the embryo (118, 119). These observations have been largely obtained from the knockout phenotypes of genes involved in these signaling pathways (117), which helped researchers to develop *in vitro* embryogenesis approaches that will be discussed in the following sections.

I) Modeling development using 2D stem cell micropatterns

ESCs grow in colonies with different shapes and size. The density of seeded cells as well as their position within the colonies affect their differentiation to various cell lineages. To overcome these challenges, 2D micropatterns have been developed through which the shape and size of the colonies can be controlled. In this approach cells are seeded on a surface that harbors the patterns of ECM proteins and passivating materials which respectively enhances and prevents the cell adhesion. In a pioneer study human ESCs were plated on a micropatterned slides in the presence of BMP4. After 48 hrs., the TE at the borders of the colonies and the three germ layers, endoderm, mesoderm and ectoderm were radially emerged from the edges inwards. The experiments demonstrated the gradient secretion of noggin (the BMP4 inhibitor) with the highest level at the colony center and maximal accumulation of BPM4 receptors at the borders which is responsible for the gastrula-like patterning. Therefore, trough the micropatterning system the fate of the cells could be defined according to their distance from the colony edges. Later, the micropattern culture was extended to the mouse ESCs. Together with BMP4, Wnt, Nodal and FGF4 signaling

researchers could recapitulate mouse gastrulation from radially organized ESC-derived germ layers.

The ESC-only based models such as micropatterned colonies are valuable models for studying the development. As in these systems, the gene expression can be polarized in the presence of external signals which ultimately leads to different cellular fates similar to the gastrula. However, since these structures lack the extra-embryonic tissues, they do not acquire the spatiotemporal architecture of a post-implantation embryo (117-120).

II) Three-dimensional modeling of development using embryonic and extra-embryonic stem cells

a) ET-blastoids

Rivorn *et al.* developed a stem cell-based model of blastocysts. First, they cultured an average of five ESCs in microwells within which the cells can make non-adherent aggregates. After 24 hrs., they seeded TSCs on top of the ESCs aggregates and within 65 hrs. they observed the formation of trophoblast cysts. At a very low frequency some of the aggregates organized into blastocyst-like structures i.e., these structures were TSCs cyst that were engulfing ESCs mass. Through induction of cAMP and Wnt signaling pathways (CHIR99021), they could increase TSC cavitation and blastoid formation. They further enhanced Cdx2 expression in ET-blastoids through addition of TSC regulators, FGF4 and TGFβ1. When they transferred the ET-blastoids into the uterus of pseudo-pregnant mice, they observed the anastomosis of trophoblasts with the mother's vascular system. Although the implanted ET-blastoids did not fully recapitulate the genuine embryonic development, they generated various post-implantation extra-embryonic cells, especially PLF-positive TGCs. Interestingly, they found that trophospheres (TSCs cysts devoid of ESCs) have not the potential for decidualization in the uterus. This indicates that the interaction between ESC and TSC compartments in ET-blastoids is essential for TSCs implantation potential, proliferation and self-renewal as well as maintaining the ET-blastoids shape (121).

b) ET embryos

To study the early post-implantation mouse embryogenesis, in the beginning ESCs were being cultured alone to generate 3D aggregates, known as EBs. In some of the EBs, the germ layer markers were observed to be expressed regionally so that these structures were called as gastruloids. However, these structures cannot establish the embryo axes in the absence of exogenous stimuli, similar to the micropatterning system. When ESCs are embedded in ECM, they develop into a rosette-like structure resembling the EPI at the time of

implantation. Formation of these structures is followed by lumenogenesis similar to the emergence of the pro amniotic cavity in the embryo. Despite this, these structures do not recapitulate the architecture of the embryo, due to the lack of the extra-embryonic tissue. To better simulate embryogenesis in culture, Harrison *et al.* introduced TSCs into the ESCs culture. Concisely, they devised a medium that supports the co-development of ESCs and TSCs. They also embedded the cells in ECM as a substitute for the basement membrane secreted by VE which is required for the EPI polarization and lumenogenesis. Through this approach, they first observed the self-organization of ESCs into the rosette-like structures and lumenognesis in the embryonic compartment. This event was followed by cavitation in the extra-embryonic compartment. Ultimately, the embryonic and extra-embryonic cavities unified into a single cavity analogous to the pro amniotic cavity of the egg cylinder embryos. Through investigation of the morphology, size, cell numbers, and expression of lineage markers, they showed that the ET embryos are similar to the embryos at the stage of implantation to early post-implantation. They reported that about 22% of structures comprised both embryonic and extra-embryonic compartments and 90% of them had the architectural characteristics of embryos. Therefore, ET embryos recapitulate the spatial and temporal sequence of events which take place during natural embryogenesis (**Figure 4**) (118, 121, 122).

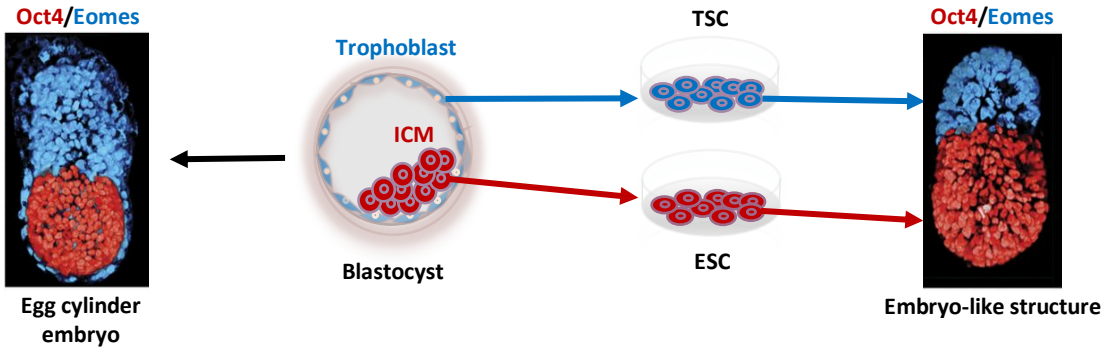


Figure 4. Assembly of embryonic and extra-embryonic stem cells to mimic embryogenesis *in vitro*

The left panel depicts natural embryo development *in vivo* from early blastocyst stage to post-implantation egg cylinder embryo. The right panel depicts the generation of embryo like structures *in vitro*. Through this approach, TSCs derived from trophoblast layer of the embryo are co-cultured with ESCs isolated from the inner cells of the embryo in a 3D matrix. The generated structure recapitulates the architectural characteristics of the egg cylinder embryo. (Adapted from Harrison *et al.*, Science, 2017) (118)

Mouse expanded potential stem cells

Recently, two groups independently reported the establishment of the mouse expanded potential stem cells (EPSCs) from embryos or ESCs.

A) L-EPSCs

Liu and his colleagues attempted to develop a culture condition that makes possible to derive and maintain cells from cleavage stage embryos (2C to 16C stages). Therefore, they hypothesized that the modulation of signaling pathways that play key roles in TE/ICM segregation might be useful to halt the blastomeres from further differentiation and to facilitate their derivation. Taking advantage of previous findings, they targeted MAPK, Src, Wnt/Hippo/tankyrases (TNKS1/2). They used Mek1 inhibitor (PD0325901), JNK Inhibitor VIII and p38 inhibitor (SB203580) to inhibit MAPK signaling. Potent pyrrolopyrimidine inhibitor (A-419259) and XAV939 were used to inhibit Src kinase and PARP family members TNKS1/2, respectively. To improve the cell culture robustness, they also included GSK3 inhibitor (CHIR99021). Together with LIF they called the cocktail of inhibitors as expanded potential stem cell medium (EPSCM). First, they could successfully derive EPSCs from cultured 8C blastomeres in EPSCM. Also, later through passaging ESCs for five times in EPSCM they could establish EPSCs. When mCherry-labeled EPSCs were injected into the morula they contributed to both ICM and TE. Moreover, when EPSCs were returned to 2i/LIF medium, these cells reverted to the ESC state. In 6.5 and 14.5 days post-coitum (d.p.c.) post-implantation chimeras were observed with the contribution of labeled EPSCs both into their embryonic proper and extra-embryonic tissues. Moreover, in some rare chimeras the contribution of EPSCs was solely in the extra-embryonic compartment, which overall indicates less efficient differentiation of EPSCs to the TE lineages. Hierarchical clustering of RNA-seq data revealed the lineage segregation of EPSCs from ESCs. Through principal component analysis (PCA) they found that EPSCs are separated from blastomeres. However, in the major development component EPSCs were in the range of 4C blastomeres. Interestingly, this group reported the establishment of TLSCs from Cdx2-GFP reporter EPSCs in a serum-free media which supports TSC maintenance and derivation, following fluorescence activated cell sorting (FACS). The TLSC-derived EPSCs proliferated well with high level of TSC gene expression and showed DNA demethylation at the *Elf5* locus (123, 124).

B) D-EPSCs

Deng and his colleagues developed a cocktail medium consisting of human LIF, CHIR 99021, (S)- (+) -dimethindene maleate (DiM), and minocycline hydrochloride (MiH)

(LCDM) which provided them with derivation of embryonic cells from blastocysts with the ability of integrating into the both embryonic and extra-embryonic compartments of chimeras (i.e., D-EPSCs). Of note, MiH acts through inhibition of PARP1 and DiM inhibits G protein coupled receptors and subsequently MAPK which is downstream to them. Next, they evaluated the functionality of injected Tdtomato^{+} D-EPSCs in embryos. To this end, they cultured chimeras with contribution of D-EPSCs into the both embryonic and extra-embryonic compartments either in 2i/LIF or TSCM and successfully derived Tdtomato^{+} D-EPSC-derived ESCs and TSCs, respectively. However, they could not directly differentiate D-EPSCs to the TSCs by culturing them in TSCM which suggests that D-EPSC-derived TSCs are originated from D-EPSC-differentiated TE cells and they cannot be generated directly from D-EPSCs *in vitro*. Moreover, they checked the developmental potential of D-EPSCs beyond the implantation stages and confirmed the contribution of them at E10.5, E12.5 and E17.5 into both embryonic and extra-embryonic lineages. Notably, these cells were observed in the layer of TGCs and were expressing proliferin (PLF), the TGC-specific marker. The transcriptome analysis of D-EPSCs showed that they are distinct from 2C-like cells and EPI (125).

Defining the developmental potential of EPSCs

Despite the evidence reported by Liu and Deng labs regarding the expanded developmental potential of EPSCs, a recent study from Rossant lab raises controversies about this feature of EPSCs. Briefly, they generated D-EPSCs and L-EPSCs with respect to the published protocols and investigated them according to the various criteria which will be explained in the following.

They performed transcriptional similarity analysis to assess which *in vivo* developmental stage resemble D-EPSCs and L-EPSCs. Using bulk RNA sequencing (RNA-seq), they studied the transcriptome shift during ESC-to-L-EPSC conversion. They observed rapid transcriptome change within the third day of conversion which indicates that ESCs readily convert to L-EPSCs. However, the transcriptome of L-EPSCs is more similar to ESCs transcriptome rather than the mouse early-stage embryo. Additionally, most of the 4-16C stage embryo genes remained silenced in L-EPSCs. To identify subpopulations with distinct transcriptional features, they generated an integrated dataset taking advantage of single-cell (sc) RNA-seq. To this end, they used scRNA-seq data of different embryonic stages from zygote to embryo at E6.75 together with naïve ESCs, primed epiblast stem cells (EpiSCs), D-EPSCs and L-EPSCs. Naïve ESCs clustered with E3.5 ICM and E4.5 EPI, while the cluster of primed EpiSCs overlapped with E5.5 and E6.75 EPI. They observed that L-EPSCs

similar to ESCs aligned with E3.5 ICM and E4.5 EPI and the cluster of D-EPSCs correlated with E5.5 EPI and EpiSCs.

Through *in vitro* differentiation of L-EPSC to the TSCs they did not observe a significant activation of TSC-specific genes. They also investigated the reprogramming efficiency of L-EPSCs and D-EPSCs in comparison to ESC to TSCs. They differentiated Cdx2-inducible L-EPSCs, D-EPSCs and ESCs toward TSCs and used flow cytometry to read out the expression of TSC-specific surface markers. They reported that in the absence or presence of Cdx2 induction there was no difference in the percentage of TLSCs generated from EPSCs or ESCs. This means EPSCs do not facilitate the reprogramming efficiency into the TSCs.

To investigate the capacity of EPSCs to contribute to the trophoblast lineages, they aggregated EPSCs and their parental ESCs with the host embryos and analyzed chimeras at E4.5, E6.25 and E12.5. At E4.5 they observed EPSCs were located in the TE in 20% of the chimeras while ESCs were localized only in the EPI. However, immunostaining for Cdx2, the TE TF, revealed that none of the EPSCs in the TE layer express this marker. At E6.25 both ESCs and EPSCs readily contributed to the EPI. In 25% of the chimeras L-EPSCs were also observed in the ExE yet, these cells were not expressing trophoblast markers, Elf5 and Tcfap2c. Interestingly, most of the mis-localized EPSCs in ExE were expressing Oct4, the EPI marker. As the placenta is composed of both trophoblast and embryo-derived cell types, they decided to further investigate the EPSCs contribution into the placenta compartments at E12.5 and found EPSCs only in the embryonic compartment of the placenta. They concluded that EPSCs do not readily give rise to the trophoblast lineages (126).

c) EPSC-B-blastoids

Taking advantage of expanded developmental potential of EPSCs, Belmonte's group established an approach through which they could generate blastoids solely from EPSCs. Briefly, they cultured EPSCs in the medium supplemented with both KSOM and ET embryo medium. They further improved the developmental condition through addition of some growth factors and finally reached to the blastoid formation efficiency of 15 and 11% using D-EPSCs and L-EPSCs, respectively. They reported that EPSC-B-blastoid formation recapitulates the cellular and molecular characteristics of early preimplantation development. Investigation of the lineage markers showed that the outer layer of EPSC-B-blastoids express TE-specific markers and the inner cells express the pluripotency factors. RNA-seq data showed that EPSC-B-blastoids transcriptionally resemble E3.5 blastocysts. Moreover, two separate clusters of ICM and TE genes as well as two clusters of PE genes were found to be expressed in EPSC-B-blastoids through scRNA-seq data analysis.

Furthermore, this group successfully reported the *in vitro* derivation of ESCs, TSCs and PE-derived extra-embryonic endoderm stem cells (XENs) from EPSC-B-blastoids. *In vivo* chimera assay also indicated the implantation and decidualization of around 7% of EPSC-B-blastoids into the uterus (127).

d) EPSC-Z-blastoids

It is reported that cells in the ET-blastoids develop into the EPI-like, TE-like lineages and low proportion of PE-like cells due to their confined development. Therefore, Goetz and her colleagues aimed to establish an approach to provide full specification of PE lineage in blastoids. They took advantage of D-EPSCs to promote the formation of embryonic and extra-embryonic cells through optimization of LCDM culture medium. Following compaction of EPSCs in microwells, they co-cultured TSCs in ET-blastoid medium under hypoxic condition (5% O_2). Through such approach, they increased the formation of cystic structures. Moreover, EPSCs could develop into the PE-like cells 45% more than ESCs in blastoids. However, they did not observe any contribution of EPSCs toward the TE cells in blastoids. Furthermore, they could not generate blastoids from EPSCs in the absence of TSCs. Through investigating the morphology, size, cell numbers and blastocyst lineage markers, they reported that EPSC-Z-blastoids resemble mid- to late-blastocyst stages. The sc transcriptome analysis revealed the presence of three blastocyst lineages in EPSC-Z-blastoids. Importantly, EPSC-Z-blastoids can decidualize into the uteri of pseudo-pregnant mice although their deciduae were smaller than those generated by genuine blastocysts (128).

Defining the blastoid-forming ability of EPSCs

As discussed previously, the developmental potential of EPSCs is controversial according to the study from Rossant lab. Besides, her group aimed to study the blastoid-forming ability of EPSCs using scRNA-seq data reported by both Belmonte and Goetz labs and aligned them with their data set. In accord with Goetz findings, they found that EPSCs can give rise to EPI and PE cells but not TE cells in EPSC-Z-blastoids. However, the cells in EPSC-B-blastoids were clustered with EPI and PE cells of E4.5 blastocysts and only 6.7% of them were grouped between TE and ExE cells. The formation of small population of TE-like cells in blastoids is interesting, as it shows the possibility of EPSCs differentiation toward TE cells under EPSC-B-blastoid forming condition. Nevertheless, these TE-like cells express very low levels of key TE markers, Cdx2, Elf5 and Gata3. The gene regulatory analysis indicated that both EPSC-B- and EPSC-Z-blastoid cells align with embryo cells. However, EPSCs could not well recapitulate the bona fide TE regulatory network in blastoids. In

conclusion, the mis-regulation of embryogenesis regulatory programs in blastoids could be a main reason that limits their further development (126).

Advantages of studying embryonic events through synthetic embryos

Knowledge about embryonic development over the time has allowed scientists to at least partially reproduce it in the culture. Under defined culture conditions, 3D synthetic embryos can be formed through stem cells self-organization. Hence, their assembly processes could be utilized as a system to unravel the embryonic developmental events. The lineage specification within the self-organization system is determined by the interactive effect of intrinsic cell characteristics and the chemical and physical property of the culture. The 3D structures generally require minimum external manipulation therefore, they are technically controllable, reproducible, experimentally accessible and cost-effective with less ethical challenges. Besides, 3D structures such as organoids and artificial embryos can present a specific phenotype or pathological and genetic abnormalities. Together with the feasibility of genome editing, 3D structures can be used in studying development, disease modelling, drug screening and regenerative medicine. Thanks to the self-organization property of 3D structures, they can host the cells from different origins, which conclusively make them as a valuable tool to study the cell fate determination. In more details, the cellular fate can be studied through two different ways in the synthetic embryos: either the donor stem cells can be introduced via self-assembly or they can be injected into the assembled structures. Self-organizing systems have also some limitations for example, they require more external instructions to reproduce organ-like structures in comparison to the *in vivo* environment. Furthermore, the size and ratio of their cellular components can be variable, and they cannot further develop after a certain level. Despite these, the advantages of 3D *in vitro* approaches outweigh their drawbacks and make them a precious model to study the development (129).

Materials and Methods

ESC culture

E14 (129P2 x OlaHsd), R1 (129X1 x 129S1), MC1 (129S6 x SvEvTac) and TCF3.2 (129S6 x C57BL/6N) mouse ESCs were seeded in feeder-free culture condition and incubated at 37 ^{0}C under 3% O_2 tension.

ESC medium composition: KnockOut DMEM (ThermoFisher, 10829–018), 15% ESC qualified Fetal Bovine Serum (FBS) (ThermoFisher, 16141–079), 2 mM L-glutamine, 1/500 home-made leukemia inhibitory factor (LIF), 0.1 mM non-essential amino acids (NEAAs), 0.1 mM β-Mercaptoethanol, 50 units/mL penicillin, 50 mg/mL streptomycin supplemented with the two inhibitors (2i); PD 0325901 (1 μM) and CHIR 99021 (3 μM).

***TCF3.2 ESCs is a kind gift from Andres Joaquin Lopez-Contreras lab, University of Copenhagen, Denmark.

TSC culture

Embryo-derived mouse TSCs, a kind gift from Janet Rossant lab (The Hospital for Sick Children, Canada), were used. TSCs were cultured in TSC medium (TSCM), which contains: 30% RPMI 1640 (Lonza, BE12-167F) supplemented with 20% ESC qualified FBS (ThermoFisher, 16141–079), 1 mM pyruvate, 2 mM L-glutamine, 100 mM β-Mercaptoethanol, 70% of conditioned medium from mitomycin-C-inactivated mouse embryonic fibroblasts (MEFs), 25 ng/mL FGF4 (R and D Systems, 235-F4-025) and 1 μg/mL heparin (Sigma-Aldrich, H3149). The medium was refreshed daily to maintain TSCs (106).

Treatments

All treatments, including APH, ATM and ATR inhibitors (ATRi and ATMi) were added into the fresh medium and the cells were kept at 37 ^{0}C under 3% O_2 tension overnight (16–17 hrs.) unless otherwise indicated. APH was used at final concentrations of 1.5 and 3 μM. ATMi (KU-55933) and ATRi (VE 822) were used at the concentrations of 10 μM and 1 μM, respectively.

RNA extraction, cDNA synthesis and qPCR

RNA was extracted using RNeasy Mini Kit (QIAGEN, 74104) and quantified by spectrophotometer (NanoDrop, ThermoFisher). DNA contaminants were removed by DNAse I digestion (QIAGEN, 79254). cDNA was prepared from 2 μg of total RNA using

iScript Advanced cDNA Synthesis Kit (Bio Rad, 1725038). qPCR assay was performed based on the standard protocol using the final working concentration of 1X SsoFast EvaGreen Supermix (Bio Rad, 1725201), 0.5 µM primer mix and 5 ng of cDNA. *Gapdh* or *Tbp* were used as internal controls to normalize the qPCR data following the $\Delta\Delta C_t$ (cycle threshold) method.

Cell immunostaining

Briefly, cells were fixed in 4% paraformaldehyde (PFA) (Sigma-Aldrich) for 20 minutes (mins) and subsequently were blocked for 1 hr in 10% FBS and 0.1% Triton-X100. Then, cells were incubated with primary antibodies at +4 ^{0}C overnight, followed by washes and incubation with secondary antibodies for 1 hr at room temperature (RT). Next, samples were mounted, and images were acquired with Leica TCS SP2 AOBS inverted confocal microscope or wide-field fluorescence microscope (Olympus AX70, Olympus). Acquired images were analyzed by ImageJ software (NIH, RRID:SCR_003070).

Flow cytometry (FACS)

ESCs were fixed and permeabilized using Cytoperm/Cytofix kit (BD Biosciences, 554714), and subsequently stained for 1 hr at RT with conjugated antibodies. Cells were washed and acquired on a FACSCalibur instrument (BD Biosciences, RRID: SCR_000401) or Attune NxT (Thermo Fisher Scientific).

Protein extraction and immunoblotting

Cells were trypsinized and washed with PBS. For the whole protein extraction, the cell pellet was lysed for 30 mins on a rotating wheel at +4 ^{0}C in RIPA buffer supplemented with protease/phosphatase inhibitor cocktail (Cell Signaling, 5872). Lysates were sonicated with a Bioruptor Sonication System (UCD200) at high power for 3 cycles of 30 seconds with one-minute breaks. Lysates were centrifuged at 13,000 revolutions per minute (rpm) for 20–30 mins and clear supernatants were transferred to new tubes. Or, the NE-PER Nuclear and Cytoplasmic Extraction Reagents (Thermo Fisher Scientific, 78833) were used for nuclear and cytoplasmic protein extraction. Protein content was quantified using Bio-Rad protein assay according to the manufacturer's instructions. For the detection of each protein, 10-35 µg of whole, nuclear or cytoplasmic protein extracts were loaded into the 4–20% Criterion TGX Precast Midi Protein Gel, 18 well (Bio Rad, 5671094). After running the protein samples on the gel, they were transferred into the Trans-Blot Turbo Midi 0.2 µm Nitrocellulose Transfer Packs (Bio Rad, 1704159) using Trans-Blot Turbo Transfer System

(Bio Rad). Next, the membrane was cut according to the size of desired proteins and were blocked in 5% milk or BSA prepared in TBS 1X buffer + 0.1% Tween 20 (TBST) for 1 hr at RT. Subsequently, membranes were incubated in the primary antibody diluted in 5% milk or BSA at +4 ^{0}C overnight. The following day, after 3 washes with TBST, membranes were incubated for 1 hr at RT with Horseradish peroxidase (HRP) secondary antibodies diluted in 5% milk or BSA. After 3 washes, membranes were incubated with Amersham ECL western blotting detection reagent (GE Life sciences, rpn2106) for 3-4 mins and were finally developed on X-ray film.

ESC *in vitro* differentiation toward TE lineages

ESCs were differentiated toward TSCs and TGCs as previously reported (Abad *et al.*, 2013 (113), Tanaka *et al.*, 1998 (106) , Ng *et al.*, 2008 (110)). Briefly, RS-induced ESCs along with intact ESCs were cultured in ESC medium and 24 hrs. post-treatment, medium was changed to the TSC medium (TSCM). The medium was refreshed daily to maintain the differentiated cells. To induce giant cell differentiation, TLSCs were split at day two on gelatin-coated plates. After 24 hrs., medium was changed to RPMI 1640 with 20% FBS, 1 mM pyruvate, 2 mM L-glutamine, 100 mM β-Mercaptoethanol in the absence of heparin and FGF4 (TSC terminal differentiation medium). The medium was changed daily for 3 days.

Preparation of MEF conditioned medium

Six million mitomycin-C-inactivated MEFs were cultured on gelatin-coated 15 cm petri dish in MEF medium which is composed of DMEM with 4.5 g/L Glucose, without L-Glutamine (Lonza, 12-614Q), 10% FBS North American, 2 mM L-glutamine, 0.1 mM non-essential amino acids, 50 units/mL penicillin, 50 mg/mL streptomycin and 1 mM pyruvate. After 24 hrs. the cells medium was changed with TSC terminal differentiation medium. In the following 72 hrs. the cells medium were collected every 24 hrs. and were replaced with the fresh medium. Finally, the mixture of the collected medium was filtered with 0.2 µm bottle top filter and were aliquoted and kept at -20 ^{0}C until further use up to six months.

Analysis of giant cells

Following the induction of giant cell differentiation, cells were fixed and stained with DAPI. From each sample, ten fields were randomly acquired and were analyzed with ImageJ software. Briefly, the area of cells in the acquired fields was calculated and the cells with the area of more than 600 µm^2 were considered and counted as giant cells.

ET embryo generation

ESC colonies were dissociated to single cells, and TSC or TLSC colonies were dissociated into small clumps by incubation in trypsin-EDTA. Cells were pelleted by centrifugation. ESCs were washed with PBS twice to remove 2i. Next, cells were counted and equal number of ESC and TSC or TLSC were mixed and repelleted in ET-embryo medium (ETM) (Cell guidance systems, M13-100). Then, the mixed cells were cultured on growth factor-reduced Matrigel (BD Biosciences, 356230) in μ-Slide 8 well plates (Ibidi, 80826) according to the 3D on-top approach for four days (118, 122, 130). After 48 hrs. of culturing cells, the medium was changed every 12-24 hrs. with respect to the consumption of the cells (i.e., according to the changes in the medium color).

ET embryos immunostaining

ET embryos were fixed with ice-cold 4% PFA for 15 mins at RT, then rinsed twice in PBS. Permeabilization was performed with 0.3% Triton-X-100, 0.1% Glycine in PBS for 10 mins at RT then, wells were washed once with PBS. Subsequently, primary antibody incubation was performed overnight in the blocking buffer (PBS plus 10% FBS and 1% Tween-20) at +4 ^{0}C. The next day, cells were washed twice in PBS, then incubated overnight in the secondary antibody in the blocking buffer at +4 ^{0}C. DAPI in PBS (5 mg/ml) was added before imaging. Images were acquired with Leica TCS SP2 AOBS inverted confocal microscope with 63X magnification. Acquired images were analyzed by ImageJ software (NIH, RRID:SCR_003070).

Analysis of ET embryos

To count the number of generated ET embryos, all the area of the one well in the μ-Slide was checked with the 63X magnification of the confocal microscope to find the bi-compartmentalized structures that were expressing mCherry in one compartment. Then, the whole well was acquired using 4X magnification in DAPI channel and all the structures were counted with ImageJ software. The percentage of generated ET embryos in each condition were calculated by dividing their number to the number of all structures in each well and were multiplied by 100.

EPSC generation

ESCs were plated on gelatin-coated plates in EPSC medium (EPSCM) which is composed of DMEM/F-12 medium, high glucose, no glutamine (Sigma-Aldrich, D6421), supplemented with 20% (vol/vol) KnockOut serum replacement (Invitrogen, 10828028), 1× (vol/vol) NEAAs, 1× (vol/vol) glutamine, 1× (vol/vol) penicillin–streptomycin, 0.1 mM β-

Mercaptoethanol, 1/500 home-made LIF, 1.0 μM PD0325901, 3.0 μM CHIR99021, 4.0 μM JNK inhibitor VIII (Tocris, 3222), 10.0 μM SB203580 hydrochloride (Tocris, 1402), 0.3 μM A-419259 trihydrochloride (Santa Cruz, sc-361094) and 5.0 μM XAV939 (Sigma-Aldrich, X3004) for five passages (123).

Transfection

ESCs were plated in ESC medium as described above. Lipofectamine RNAiMAX (ThermoFisher) was diluted in Opti-MEM (ThermoFisher) according to the manufacturer's protocol and incubated for 5 mins at RT. *Atr* MISSION esiRNAs (Merck, EMU213731-50UG) was diluted in Opti-MEM to the opportune concentration and then added to the Lipofectamine emulsion at 1:1 ratio. Following incubation at RT, the mixture was added to the cell's suspension in culture medium. The final concentration of esiRNAs was 60 nM. For each experiment a sample was transfected with esiRNA targeting RLUC as a negative control at the same concentration used for esiRNA. Gene expression was assessed 48 hrs. post-transfection.

Generation of ROSA26 constitutive mCherry-expressing ESCs

R26R-H2B-mCherry ESCs, a kind gift from Hiroshi Kiyonari lab (RIKEN Center for Biosystems Dynamics Research, Japan) were used. pBS513 EF1α-cre plasmid (Addgene, 11918) was purchased and maxi prepared using NucleoBond Xtra Maxi kit (MACHEREY-NAGEL, 740414.50). Next, four μg of plasmid in Opti-MEM was used with diluted Lipofectamine 2000 (ThermoFisher) in Opti-MEM according to the manufacturer's protocol for transfection of 10^5 ESCs in a 6-well plate. Following two passages, top five percent of highly expressing mCherry-positive cells were sorted, expanded and froze until further use. Non-transfected R26R-H2B-mCherry ESCs were used as a negative control for the cell sorting.

Chromatin immunoprecipitation (ChIP)

Simple ChIP Enzymatic Chromatin IP Kit (Magnetic Beads) (Cell Signalling, 9003) was used. Briefly, for each condition around two million ESCs were plated in 15 cm petri dishes. The number of required seeded plates for each condition depends on the cell viability upon treatment and should be calculated in a way to finally have around 26 million fixed cells (around 6 IPs) from each condition. qPCR assay was performed based on the standard protocol using final working concentration of 1X SsoFast EvaGreen Supermix (Bio Rad, 1725201), 0.5 μM primer mix and 1 μL of DNA. The DNA regions that were checked are:

Eomes (-) 88.3 kbs (kilo bases), *Cdx2 Intron 1*, *Fgfr2* (+) 74.2 kbs, *Gata3* 5' UTR (untranslated region). Primers that are specific to the mentioned regions were purchased according to the Home *et al.* study (90). Primers within a gene desert region (negative region) were used as a negative control. qPCR analysis was performed according to the following formulae:

C_t Adjusted input = C_t 2% Input-LOG (50;2)
ΔC_t = C_t Adjusted input-C_t Antibody of interest or IgG
Percentage of input = $100*2^{(\Delta C_t)}$

List of antibodies used in this study

Antibody	Application / Catalogue No.	Immunofluorescence	Immunoblotting	FACS	ChIP
Nanog	ab80892	1/200	_____	_____	_____
Cdx2	CM226B	1/100	_____	_____	_____
Oct4	sc-5279 ab27985	1/100 1.5/500	_____	_____	_____
Eomes	ab233345	1/400	1/1000	_____	3µL/IP
mCherry	ab167453	1/1000	_____	_____	_____
RFP	5f8-100	1/500	_____	_____	_____
E-Cadherin	13-1900	1/400	_____	_____	_____
Laminin	L9393	1/400	_____	_____	_____
Plet1	33A10	1/200	_____	_____	_____
P-Chk1	12302S	1/800	1/1000	_____	_____
Tead4	ab58310	1/1000	1/3000	_____	5µL/IP
Total Chk2	05-649	_____	1/500	_____	_____
P-Chk2	ab59408	_____	1/1000	_____	_____
Total Chk1	sc-8408	_____	1/1000	_____	_____
P-ATR	2853S	_____	1/1000	_____	_____
P-ATM	ab81292	_____	1/10000	_____	_____
GAPDH	G8795	_____	1/100000	_____	_____
Elf5-Alexa Fluor 647	ITA8395 (1mg/mL)	_____	_____	2.5µL/test	_____
Cdx2-Alexa Fluor 488	sc-393572 (0.2mg/mL)	_____	_____	2.5µL/test	_____

List of RT-qPCR primers used in this study

	Forward (5'-> 3')	Reverse (5'->3')
Elf5	ATGCTGAAGAGACCAAGACTG	TCTGACAAATTCCCACAGGTG
Cdx2	CAAGGACGTGAGCATGTATCC	GTAACCACCGTAGTCCGGGTA
Eomes	TTCACCTTCTCAGAGACACAGTTCAT	GAGTTAACCTGTCATTTTCTGAAGCC
Fgfr2	GAGGAATACTTGGATCTCACC	CTGGTGCTGTCCTGTTTGGG
Oct4	TCTTTCCACCAGGCCCCCGGCTC	TGCGGGCGGACATGGGGAGATCC
Nanog	CAAAGGATGAAGTGCAAGCG	CCAGATGCGTTCACCAGATAG
Gata3	CGGGTTCGGATGTAAGTCGA	GTAGAGGTTGCCCCGCAGT
Tead4	CGAAGGTCTGCTCATTTGG	GGTGGATGCGGTACAAATAG
Prl2c2	AGAGACAAAAGCCCCATGAG	ACTCACTAGATCGTCCAGAGG
Axin2	GGATGCTGAAGGCTCAAAG	AGGCAAATTCGTCACTCG
H19	AGCAGTGATCGGTGTCTC	TGTCATCCTCGCCTTCAG
Dlx3	AGTTCAATCTCAATGGGCTCG	CTCTTTCACCGACACTGGG
Gapdh	CTTTGTCAAGCTCATTTCCTGG	TCTTGCTCAGTGTCCTTGC
Tbp	AAGAAAGGGAGAATCATGGACC	GAGTAAGTCCTGTGCCGTAAG

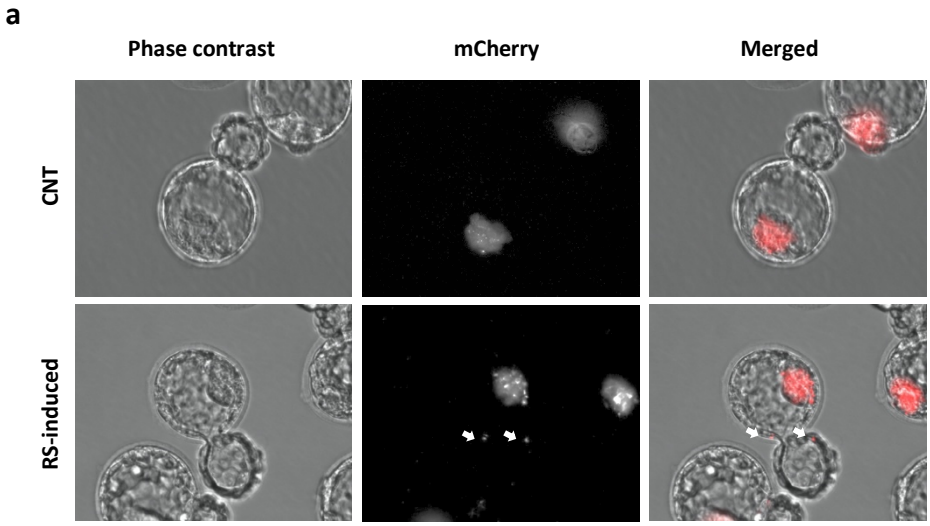
List of ChIP-qPCR primers used in this study

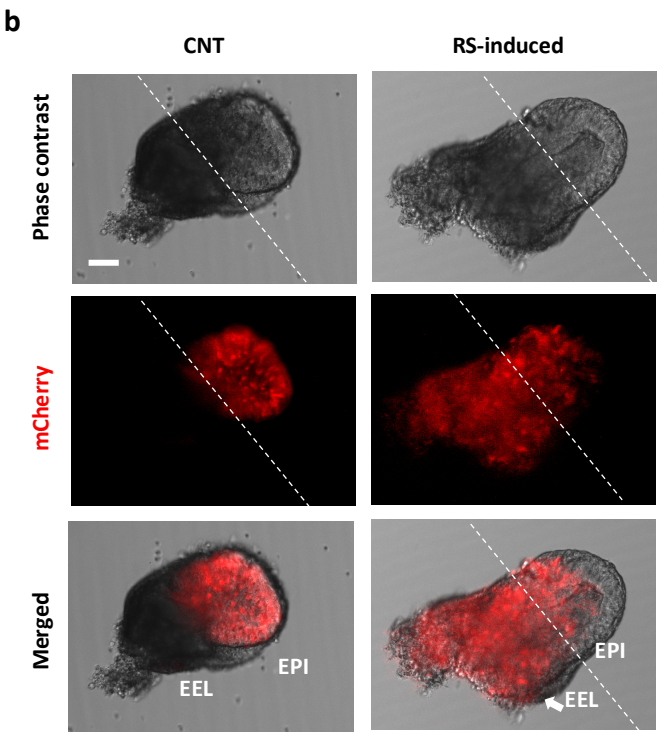
	Forward (5'-> 3')	Reverse (5'->3')
***Eomes* (-) 88.3 kb**	CACCTCTTCCAAAGCTATTGTT	CGGAGAGGAGGTTAGGGAAG
Cdx2 Intron 1	ACCCTCTCGGCCAGGAAC	CTGAATCGTGCCCTTGAGTT
***Fgfr2* (+) 74.2 kb**	TCATCTGTCGCCAGGTAAGG	GGAGGCCTGCAGATAGGTTA
***Gata3* 5' UTR**	TCTTTCTCCTCTCCCTCTCTCA	GGCGCCGTCTTGATAGTTT
Negative region	ATGCATGCCATACCTCCAGT	CCTTGTCCATGGATTGTTCACG

Results

Aim of the project

Strikingly, through *in vivo* chimera assay we have recently found that RS-induced ESCs contribute to the extra-embryonic compartment of the embryos, while unperturbed cells contribute only to the embryonic proper (**Figure 5**). These results suggest that RS induces the differentiation of ESCs toward extra-embryonic lineages to minimize the contribution of unrepaired ESCs to the embryonic compartment and later to the progeny (131). *In vivo* chimera assay possesses various limitations. Firstly, it is technically challenging, meaning that it requires extra effort and tedious procedures in comparison to the *in vitro* assays. Secondly, it is time consuming and costly. Besides, live embryos have restricted manipulation possibilities and working with them raises ethical concerns. According to these drawbacks, the *in vivo* chimera assay is overall a low throughput approach (102, 103, 129). To overcome these limitations, we aimed to further investigate our finding through *in vitro* models of ESC differentiation to the extra-embryonic lineages. During this study, we established an innovative combinatorial approach of ESC *in vitro* differentiation with 3D stem cell embryogenesis to study the impact of RS on the differentiation potential of ESCs. Next, we focused on the molecular mechanisms downstream to RSR that lead to the contribution of RS-induced ESCs to the extra-embryonic compartment.




b
CNT
RS-induced
Phase contrast
mCherry
Merged
EEL
EPI
EPI
EEL

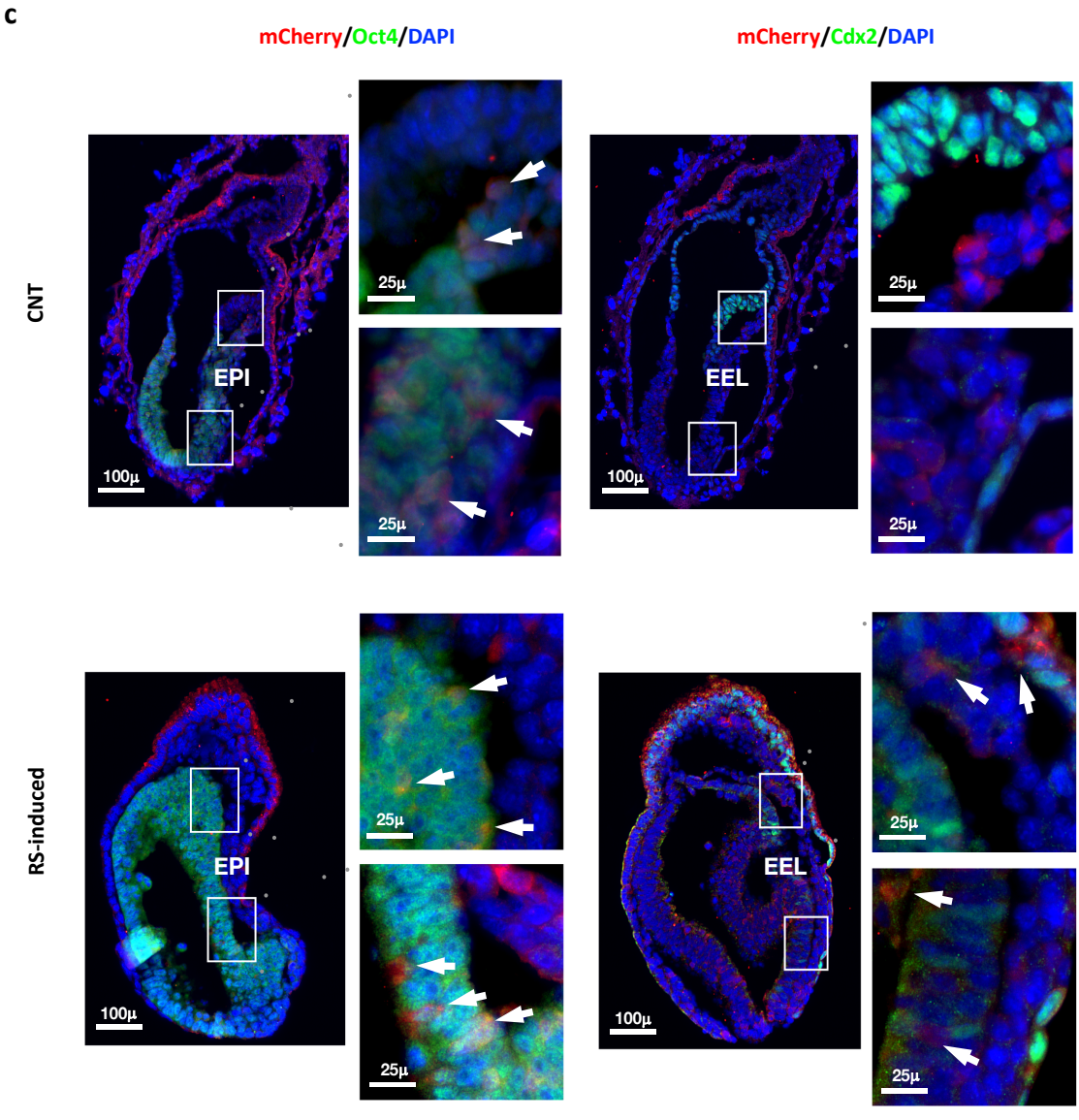

c
mCherry/Oct4/DAPI
mCherry/Cdx2/DAPI
CNT
RS-induced
EPI
EEL
EPI
EEL
100μ
25μ

Figure 5. RS-induced ESCs contribute to the extra-embryonic compartment of the embryos

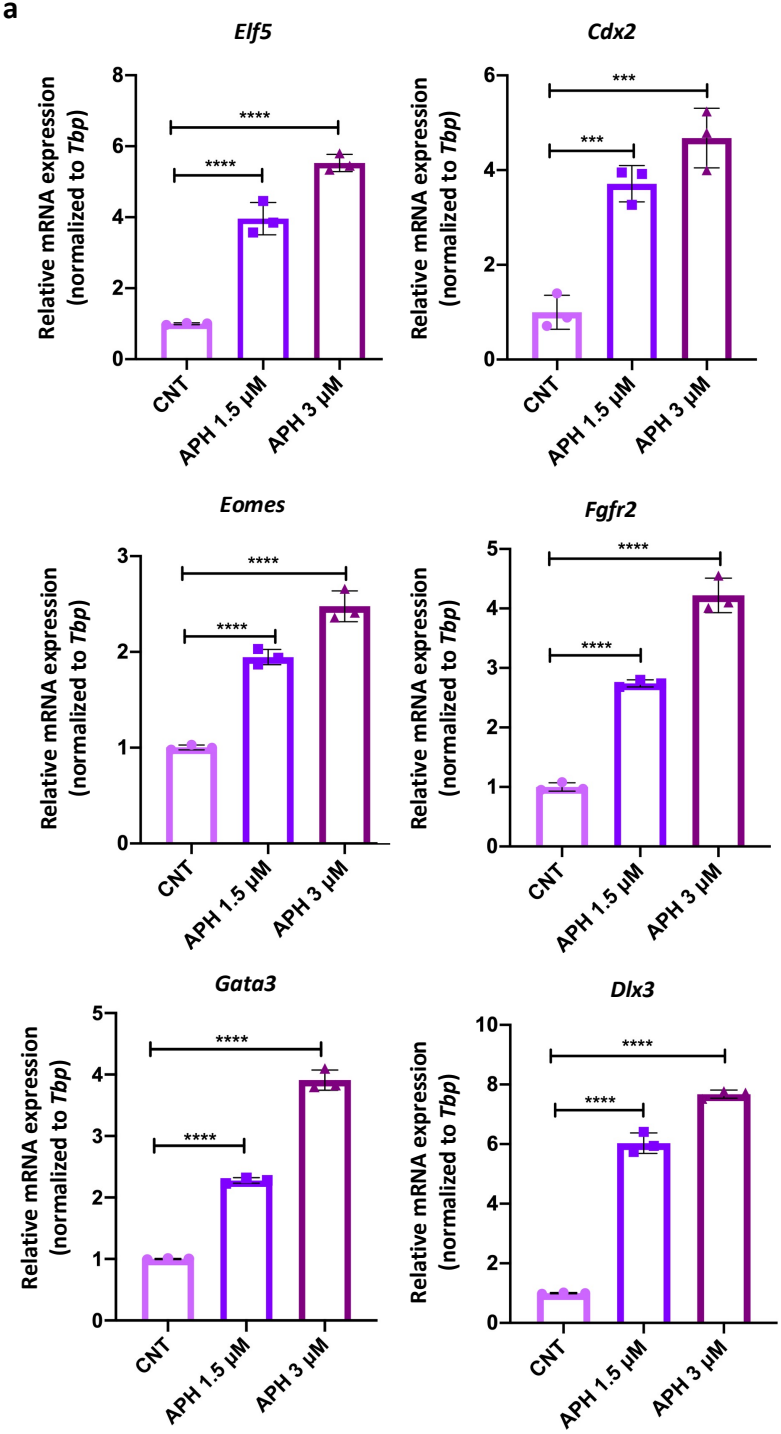
(a) Images of blastocysts displaying the contribution of mCherry-labeled ESCs to the ICM and TE layers in the presence and absence of RS (bar = 20 μm). **(b)** Images showing the contribution of injected mCherry-labeled ESCs (unperturbed or RS-induced ESCs) to the epiblast (EPI) or extra-embryonic layers (EEL) of mouse embryos at E7.5 (bar = 50 μm). **(c)** Immunostaining of mouse embryos at E7.5. Arrows in the left panel indicate the contribution of mCherry-labeled ESCs to the EPI (marked by Oct4) in intact and RS-induced conditions. Arrows in the right panel indicate the contribution of injected mCherry-labeled ESCs to the EEL (marked by Cdx2) only in RS-induced condition. (lower magnification, bar = 100 μm, higher magnification, bar = 25 μm). (Adapted from Atashpaz *et al.*, eLIFE, 2020) (131).

RS activates the expression of key TE genes in mESCs

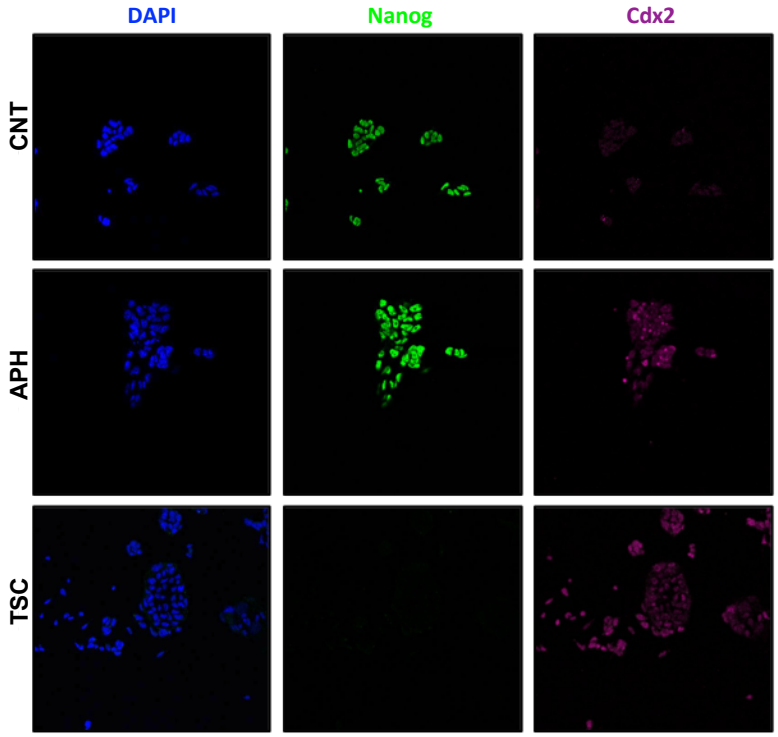
Our recent finding indicates the inductive role of RS in the extra-embryonic contribution of labeled ESCs and their co-localization with Cdx2-expressing cells (**Figure 5**). Hence, we asked what are the molecular players downstream to RSR that lead to the contribution of RS-induced ESCs to the extra-embryonic compartment. To this end, first we asked whether RS activates the expression of TE genes in ESCs *in vitro*. To find the answer, we treated mESCs with different concentrations of APH, which reversibly decreases the activity of DNA polymerases and ultimately induces RS (53). The range of APH concentration and its exposure time to the cells were carefully checked before to make sure that our treatment only induces RS to the cells without causing any DD. Interestingly, RT-qPCR analysis showed a significant increase in the expression of key TE markers upon APH treatment in a dose-dependent manner (**Figure 6a**). IF staining for Cdx2 revealed the expression of this TE marker upon APH treatment in mESCs similar to the TSCs (**Figure 6b**). Moreover, FACS analysis showed the increase in the number of Cdx2- and Elf5-positive ESCs following APH treatment. This result further confirmed the inductive role of RS in the expression of TE-specific genes in mESCs not only at transcriptional but also at translational level (**Figure 6c and d**).

These data together suggest that RS activates the expression of canonical TE markers in mESCs (**Figure 6**).

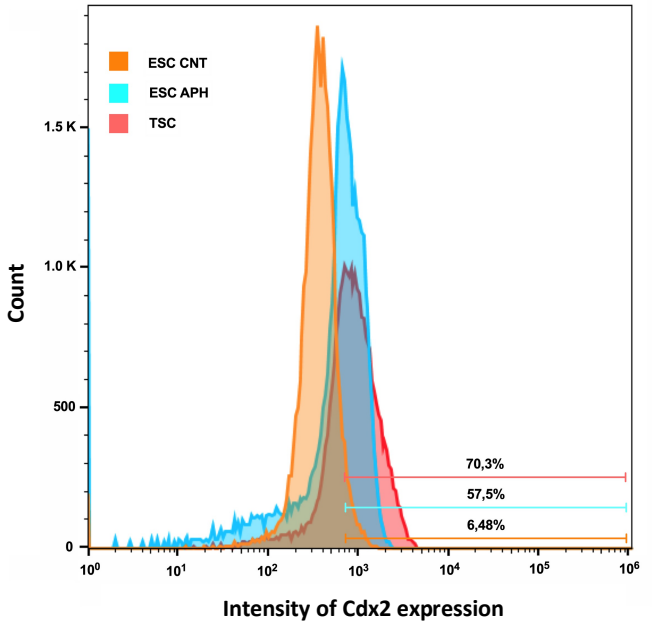
a

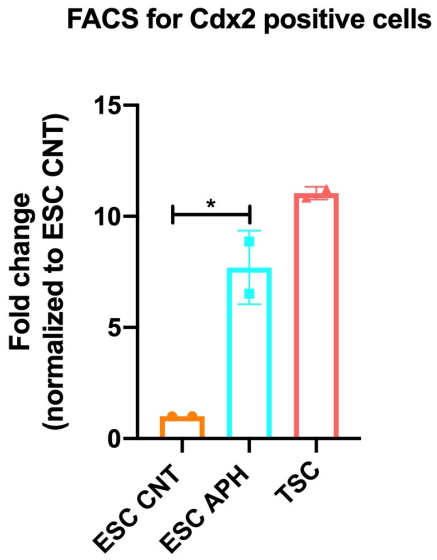

Elf5
Cdx2
Eomes
Fgfr2
Gata3
Dlx3
Relative mRNA expression (normalized to Tbp)
CNT
APH 1.5 µM
APH 3 µM
****
***

b


DAPI
Nanog
Cdx2
CNT
APH
TSC

c


ESC CNT
ESC APH
TSC
1.5 K
1.0 K
500
0
Count
70,3%
57,5%
6,48%
10^0
10^1
10^2
10^3
10^4
10^5
10^6
Intensity of Cdx2 expression


FACS for Cdx2 positive cells
Fold change
(normalized to ESC CNT)
15
10
5
0
*
ESC CNT
ESC APH
TSC

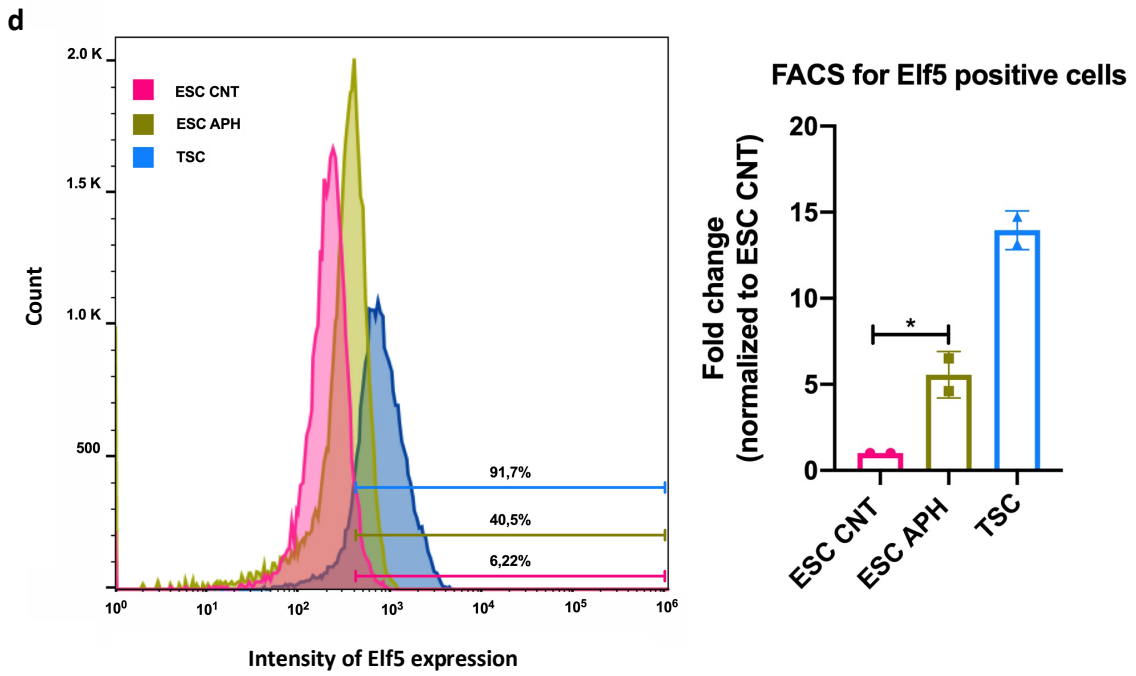


Figure 6. RS activates the expression of key TE genes in mESCs

(a) RT-qPCR analysis on mECSs treated with different concentrations of APH for key TE markers. **(b)** Immunostaining of mESCs for the TE marker (Cdx2) and pluripotency marker (Nanog) upon treatment with APH. TSCs were stained as a positive control for Cdx2. **(c)** FACS analysis on mESCs upon treatment with APH showing the percentage of Cdx2-positive cells. TSCs were used as a positive control for Cdx2 expression. Two biological replicates were analyzed per sample. **(d)** FACS analysis on mESCs upon treatment with APH showing the percentage of Elf5-positive cells. TSCs were used as a positive control for Elf5 expression. Two biological replicates were analyzed per sample. All bar plots show mean with ± SD (*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, one-way ANOVA).

RS induces the differentiation of ESCs toward TSCs and increases the number of TLSCs

Several studies have reported the successful generation of TLSCs through OE of key TSC TFs such as Cdx2, Gata3, Tead4 and Elf5 in ESCs using an established medium (TSCM). This medium is supplemented with FGF4, which is essential for the proliferation and maintenance of trophoblast cells *in vivo* (106-110). We previously showed the increased expression of TSC-specific genes under RS in mESCs (**Figure 6**) as well as the contribution of these cells to the extra-embryonic compartment of the embryos *in vivo* (**Figure 5**). With respect to these findings, we asked whether, the increased expression of TE markers in RS-induced ESCs is sufficient to induce their differentiation toward TSCs.

To this end, first we cultured ESCs in TSCM to differentiate them toward TSCs (**Figure 7a**) and characterized the generated cells. RT-qPCR analysis showed a significant decrease in the expression of ESC-specific TFs, *Oct4* and *Nanog* and a remarkable increase in the expression of TE genes including *Cdx2* and *Eomes* in the differentiated cells (**Figure 7b**). Moreover, IF staining indicated the reduced expression of Nanog and increased expression of Cdx2 in the differentiated ESCs similar to the TSCs (**Figure 7c**).

These data together confirmed the differentiation of ESCs toward TSCs, or in other words, the formation of ESC-derived TSCs (TLSCs) (**Figure 7**).

Next, to investigate the impact of RS on ESCs differentiation toward TSCs, we treated ESCs with APH and in parallel with intact cells we cultured them in TSCM (**Figure 8a**). Through various methods, we checked the expression of canonical TE markers in differentiated cells. RT-qPCR analysis revealed a significant increase in the transcript of many TE genes. Of note was *Fgfr2*, a target for FGF signaling secreted from ICM, which is required for the TE growth and specification (132). Also *Tead4*, that initiates the expression cascade of TE genes (133), as well as *Cdx2*, *Gata3* and *Elf5* were significantly increased in response to RS (**Figure 8b**). IF staining showed the emergence of Eomes-positive cells that have lost the expression of Oct4 during differentiation both in the presence and absence of APH (**Figure 8c**). Importantly, FACS analysis demonstrated the significant increase in the number of Cdx2-positive cells upon differentiation. Interestingly APH treatment remarkably increased this cell population in comparison to the intact cells (**Figure 8d**). These data are in consistent with our *in vivo* results (**Figure 5**) (131).

Overall, these data suggest that RS induces the differentiation of ESCs toward TSCs through activating the expression of canonical TE markers and increasing the number of TLSCs (**Figure 8**).

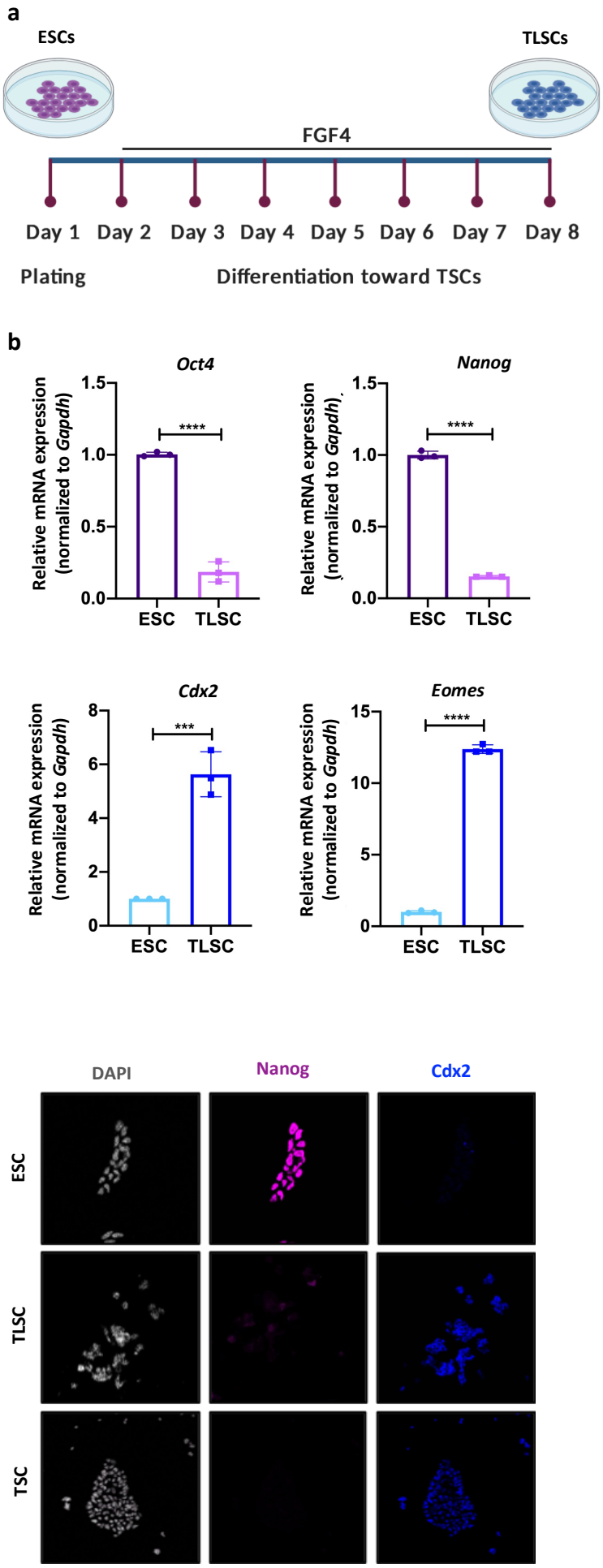
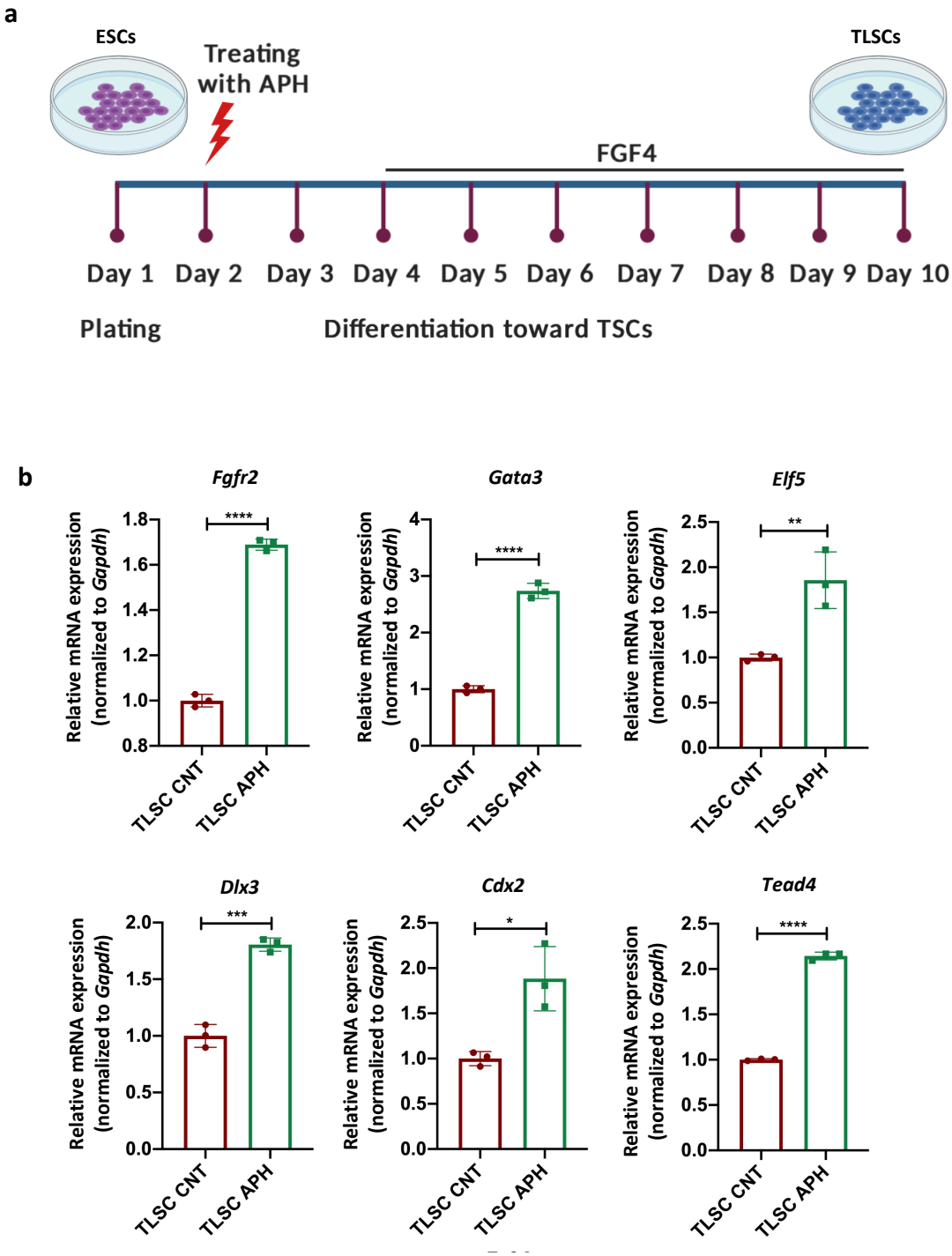
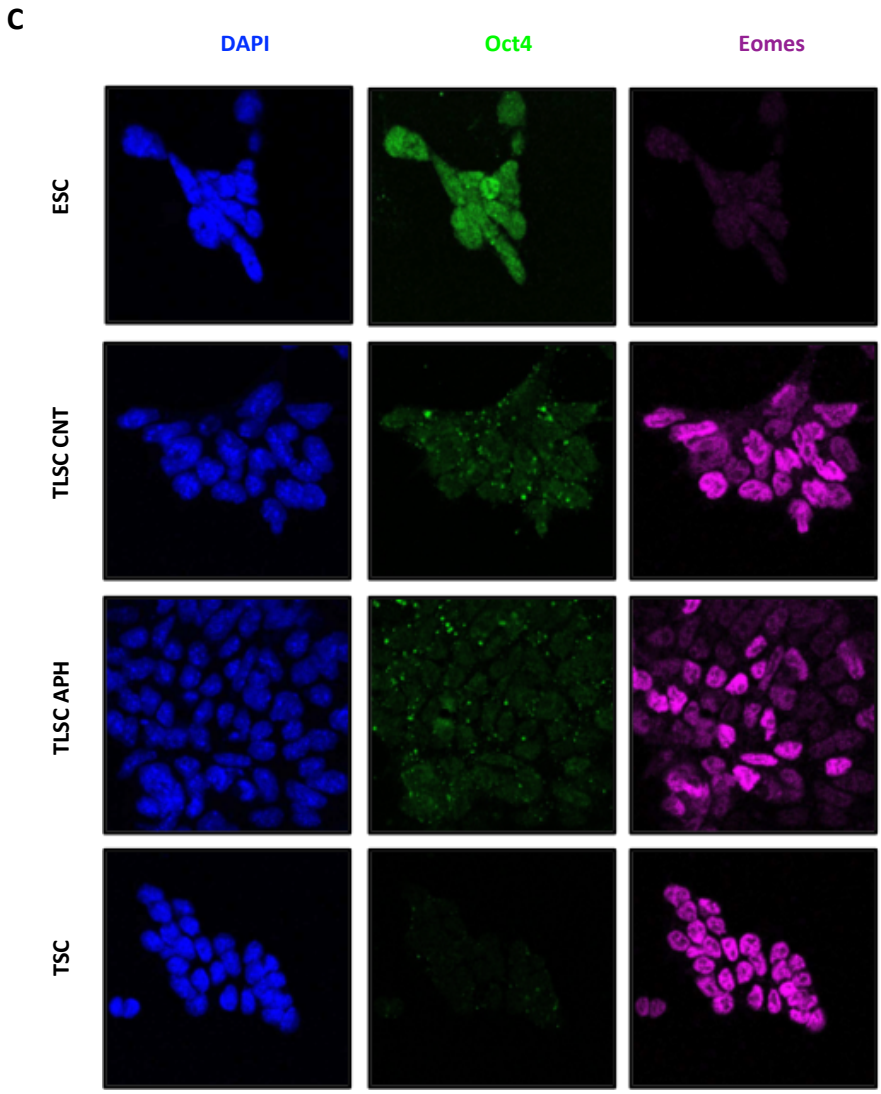

a
ESCs
TLSCs
FGF4
Day 1
Day 2
Day 3
Day 4
Day 5
Day 6
Day 7
Day 8
Plating
Differentiation toward TSCs
b
Oct4
Nanog
Cdx2
Eomes
Relative mRNA expression (normalized to Gapdh)
ESC
TLSC
****
***
c
DAPI
Nanog
Cdx2
ESC
TLSC
TSC

Figure 7. ***In vitro*** **differentiation of ESCs toward TSCs**

(a) Schematic view of ESCs differentiation toward TSCs. **(b)** RT-qPCR analysis on mECSs and TLSCs for key ESC and TSC markers. **(c)** Immunostaining of mESCs and TLSCs for the TE marker (Cdx2) and the pluripotency marker (Nanog). TSCs were stained as positive control for Cdx2 expression. All bar plots show mean with ± SD (*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, unpaired t test).




C
DAPI
Oct4
Eomes
ESC
TLSC CNT
TLSC APH
TSC

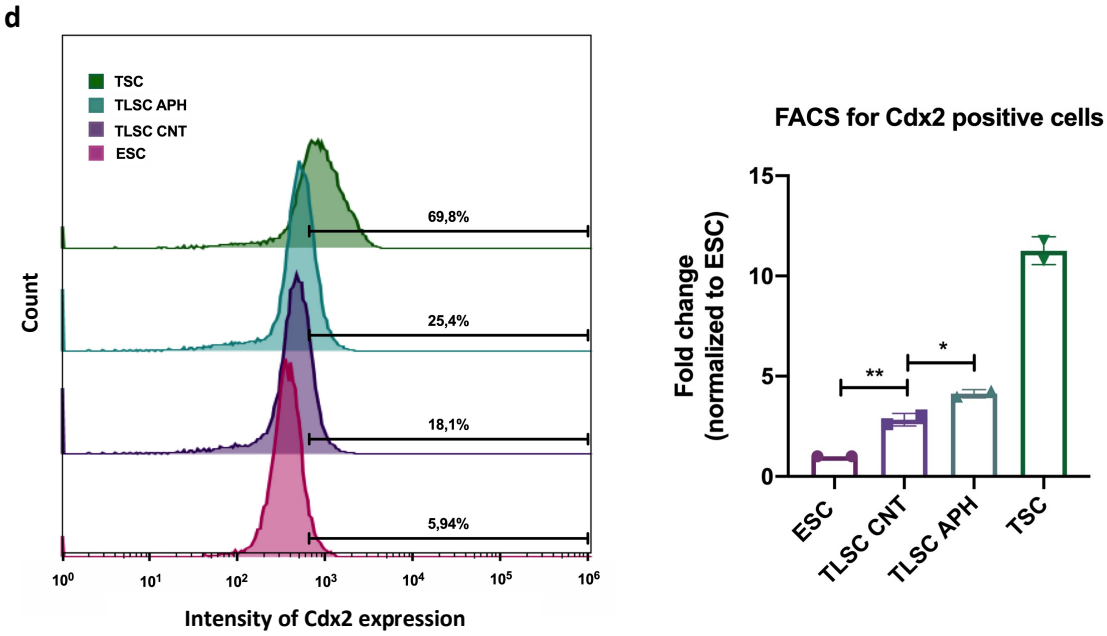

d
TSC
TLSC APH
TLSC CNT
ESC
69,8%
25,4%
18,1%
5,94%
Count
Intensity of Cdx2 expression
FACS for Cdx2 positive cells
Fold change (normalized to ESC)
**
*
ESC
TLSC CNT
TLSC APH
TSC

Figure 8. RS induces the differentiation of ESCs toward TSCs and increases the number of TLSCs

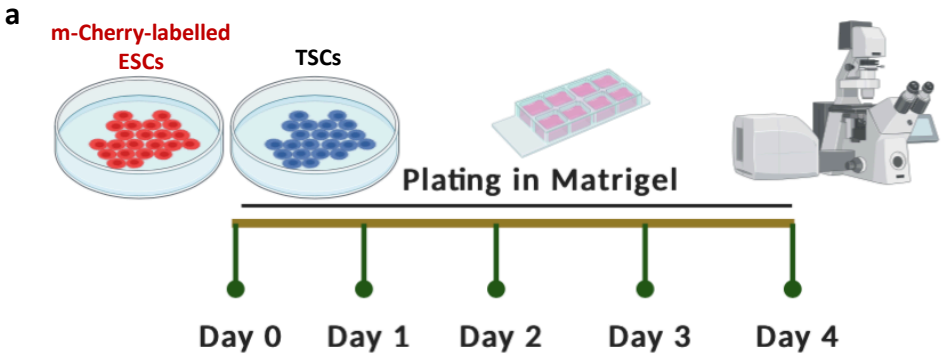
(a) Schematic view of ESCs differentiation toward TSCs in the absence or presence of APH. **(b)** RT-qPCR analysis on TLSCs for key TSC markers in the absence or presence of APH. Bar plots show mean with ± SD (*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, unpaired t test). **(c)** Immunostaining of mESCs, TLSCs (in the absence or presence of APH) and TSCs for the TE marker (Eomes) and the pluripotency marker (Oct4). TSCs were stained as a positive control for Eomes expression. **(d)** FACS analysis on TLSCs in the presence or absence of APH showing the percentage of Cdx2-positive cells. ESCs and TSCs were used as a negative and positive control for Cdx2 expression, respectively. Two biological replicates were analyzed per sample. Bar plot shows mean with ± SD (*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, one-way ANOVA).

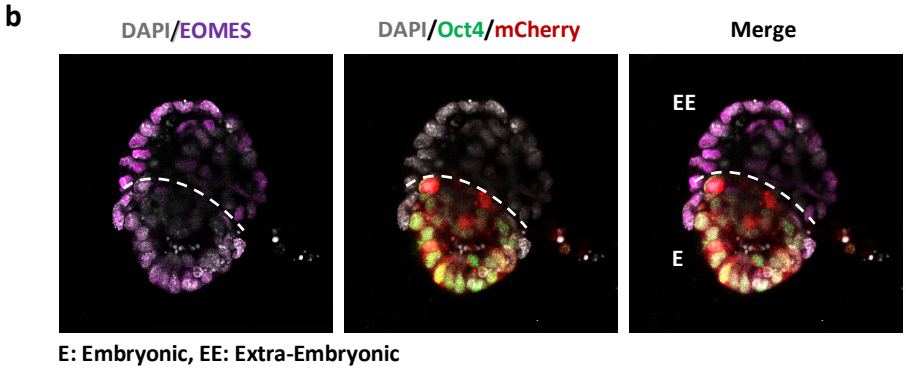
TLSCs are able to generate ET embryos similar to the embryo-derived TSCs

So far, we found that RS activates the expression of canonical TE markers in ESCs and RS induces the differentiation of ESCs toward TSCs (**Figure 6 and 8**). With respect to these findings, we wondered whether RS-induced ESCs have functionally higher differentiation potential than untreated ESCs. To find the answer, we took advantage of an *in vitro* embryogenesis approach published by Harrison *et al.* which makes possible to generate embryo-like structures (ET embryos) through co-culturing ESCs with TSCs in a dish (**Figure 4**) (118, 122). To this end, first we aimed to establish this approach through co-embedding our mCherry-labeled ESCs with TSCs in ECM in a medium (ETM) that supports the simultaneous cultivation and development of both cell lines (**Figure 9a**). It is reported that about 22% of structures are ET embryos and the rest are either ESC-ESC aggregates or TSC-TSC aggregates (118). After four days, we could successfully generate 3D structures with single central cavity and architectural structure of the reported ET embryos. Consistent with the ET embryos, our structures were expressing the embryonic marker (Oct4) and the extra-embryonic marker (Eomes) (**Figure 9b**).

Next, we asked whether TLSCs are able to generate ET embryos similar to the embryo-derived TSCs. Or in other words, whether TLSCs are functionally similar to the TSCs. Since our aim is to investigate the differentiation efficiency of ESCs to TLSCs and this process is not 100% efficient, we took advantage of ESCs that constitutively express mCherry as embryonic compartment. This makes us able to clearly investigate the differentiation potential of unlabeled ESCs to TLSCs. Initially we were using ESCs that we had infected

them with EF1α-mCherry constitutively expressing vector. However, we found mCherry silencing in some of the colonies following a number of passaging (Data not shown). This could cause misinterpretation in detecting the structures that actually are generated from the assembly of mCherry-labeled ESCs and TLSCs. Hence, we decided to establish another constitutive expression system to solve this issue. We took advantage of the R26R reporter mESCs, which harbor the cDNA coding sequence of fused H2B-mCherry preceded by the neo cassette flanked by *LoxP* sites at ROSA26 locus (**Figure 10a**). ROSA26 is an endogenous constitutively expressing locus therefore, is an ideal genomic site to integrate the sequence of interest. Through excision of the *LoxP* sites by Cre in these cells, H2B-mcherry is expressed and is detectable ubiquitously in the nucleus (134). Therefore, R26R-H2B-mCherry ESCs are a perfect system to solve the issue of mCherry silencing. To activate mCherry in these cells, we took advantage of EF1α-Cre plasmid. Following the maxi preparation of the plasmid, we transfected it into the R26R-H2B-mCherry ESCs. Next, we sorted the mCherry-positive cells for two times to acquire top five percent highly mCherry-expressing cells (**Figure 10b**). Subsequently, we performed IF staining to assess the expression of reporter protein (**Figure 10c**). Ultimately, we could successfully generate mCherry constitutively expressing ESCs which the expression of their reporter protein did not alter even after several passages (**Figure 10**).

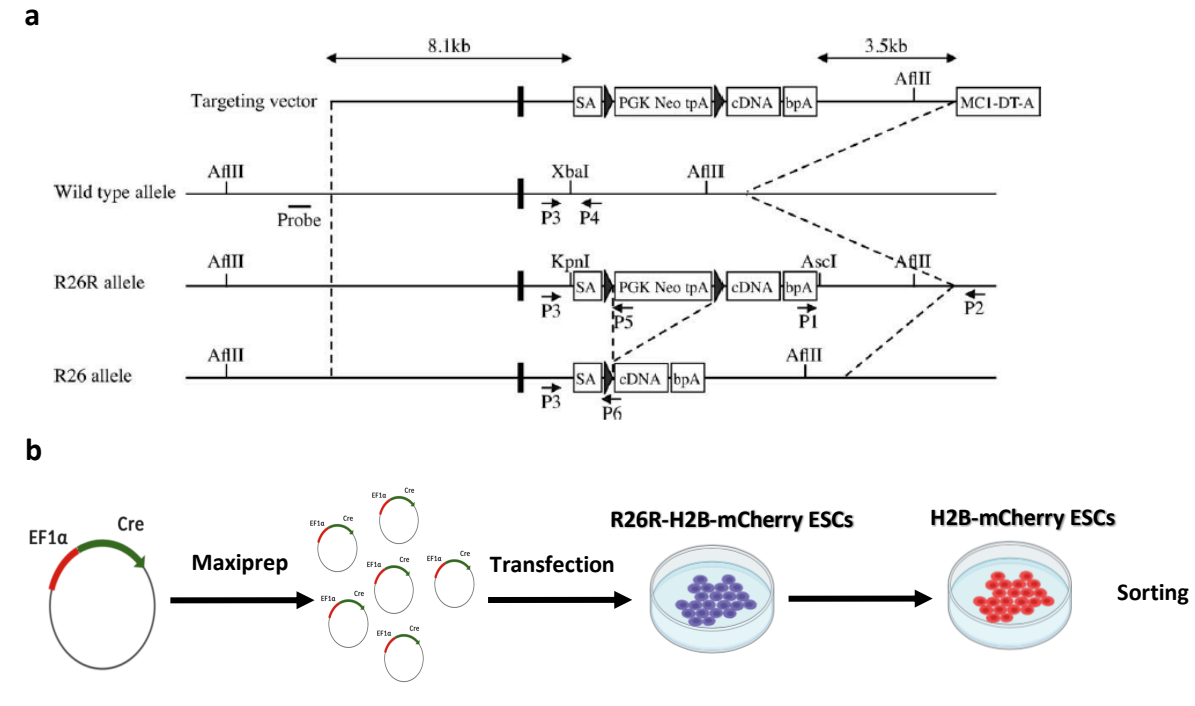




E: Embryonic, EE: Extra-Embryonic

Figure 9. Generation of ET embryos from the assembly of ESCs and TSCs

(a) Schematic view of ET embryos generation from the assembly of mCherry-labeled ESCs and TSCs. **(b)** Immunostaining of ET embryo for the embryonic marker (Oct4) and the extra-embryonic marker (Eomes).



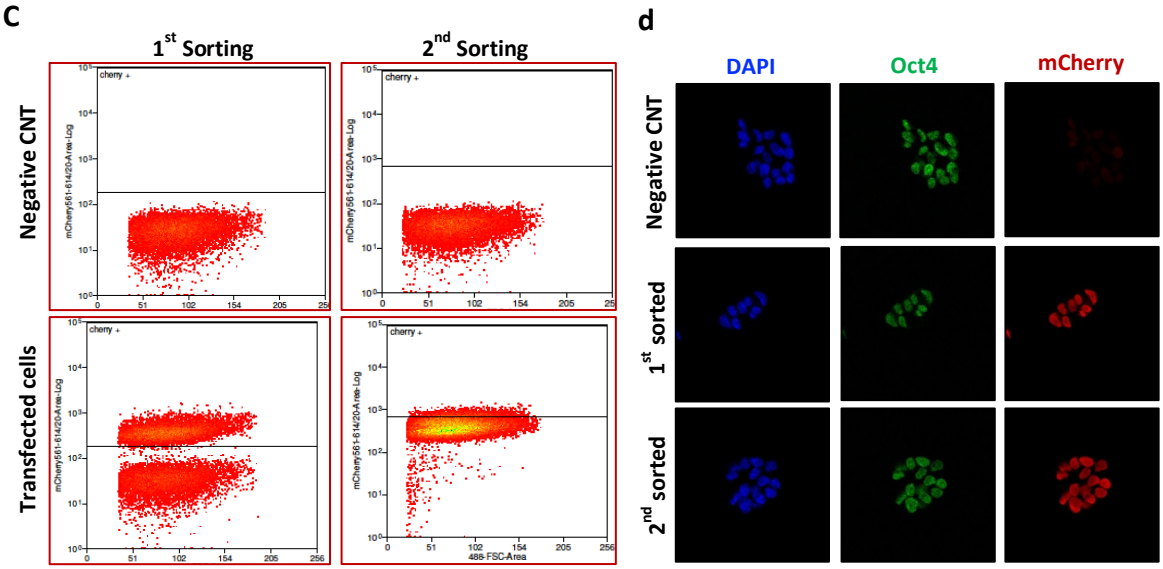


Figure 10. Generation of ROSA26 mCherry constitutively expressing ESCs

(a) Schematic view of R26R-H2B-mCherry ESCs generation (Reprinted from Abe *et al.*, Wiley, 2011) (134). **(b)** Schematic view of the generation of ROSA26 mCherry constitutively expressing ESCs. **(c)** Graphs show the sorting of transfected R26R-H2B-mCherry ESCs

with EF1α-Cre plasmid. **(d)** Immunostaining of R26R-H2B-mCherry ESCs and sorted ROSA26 mCherry constitutively expressing ESCs for ESC marker (Oct4) and mCherry.

Back to our question regarding the functionality of TLSCs, we aimed to answer whether these cells are able to generate ET embryos similar to the embryo-derived TSCs. To this end, we differentiated unlabeled ESCs toward TSCs and subsequently cultured them with labeled ESCs in a 3D matrix (**Figure 11a**). After four days, we checked the morphology of structures as well as the expression of the embryonic and extra-embryonic markers. Interestingly, we found cylindrical structures from the assembly of ESCs with TLSCs that were expressing Oct4 and Eomes in the embryonic and extra-embryonic compartments, respectively (**Figure 11b**). This finding was in line with the properties of the reported ET embryos that are generated from embryo-derived TSCs.

Next, we aimed to further characterize ESC-TLSC structures according to the two reported criteria of ET embryos. The first criterion is the formation of a single central cavity in ET embryos. Similar to what happens *in vivo*, the embryonic compartment in ET embryos goes under lumenogenesis. This is followed by cavitation in the extra-embryonic proper which finally the two cavities unify and form a single central cavity (118). Interestingly, through IF staining for the cell adhesion marker (E-cadherin), we found that ESC-TLSC structures formed a unique central cavity similar to the ET embryos (**Figure 12a**).

The Second criterion is the basement membrane displacement toward the extra-embryonic compartment and its disappearance after merging the two cavities in ET embryos. While this event does not happen in ESC-ESC aggregates. In other words, the disappearance of the basement membrane is a unique feature of ESC-TSC structures (118). IF staining for laminin, a basement membrane protein, confirmed that this protein became undetectable after four days in ESC-TLSC structures similar to the ET embryos. Nonetheless, in line with the reports, this membrane remained visible in mCherry-labeled ESC-ESC aggregates (**Figure 12b**).

These data together show that TLSCs are functionally similar to the TSCs as they are able to generate structures with the characteristics of the ET embryos (**Figure 11 and 12**).

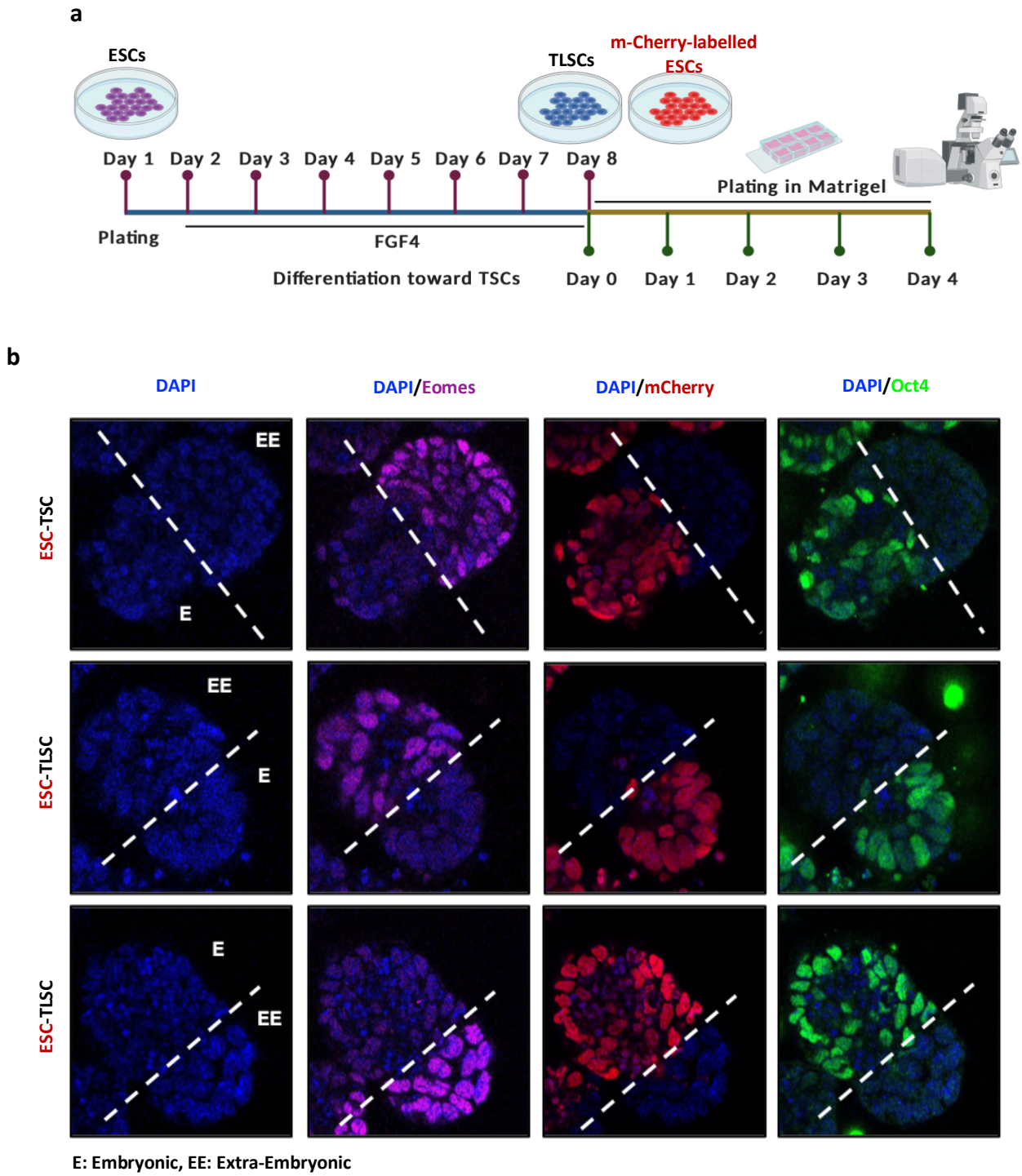
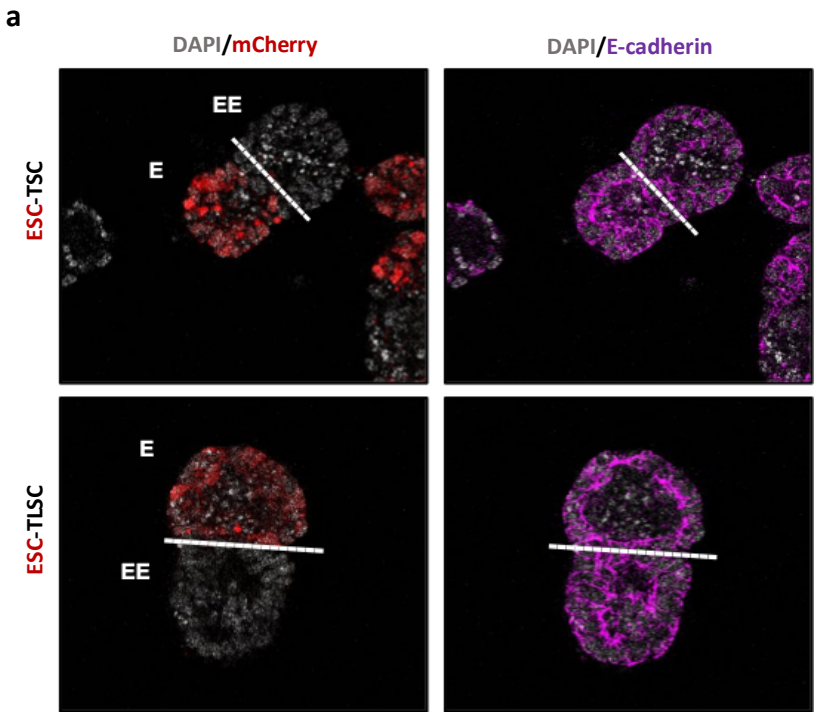


Figure 11. TLSCs generate ET embryos similar to the embryo-derived TSCs

(a) Schematic view of ET embryo generation from the assembly of TLSCs and mCherry-labeled ESCs. **(b)** Immunostaining of ET embryos generated from the assembly of mCherry-labeled ESCs with TSCs or TLSCs for the embryonic marker (Oct4) and the extra-embryonic marker (Eomes). At least two biological replicates were performed per sample.



E: Embryonic, EE: Extra-Embryonic

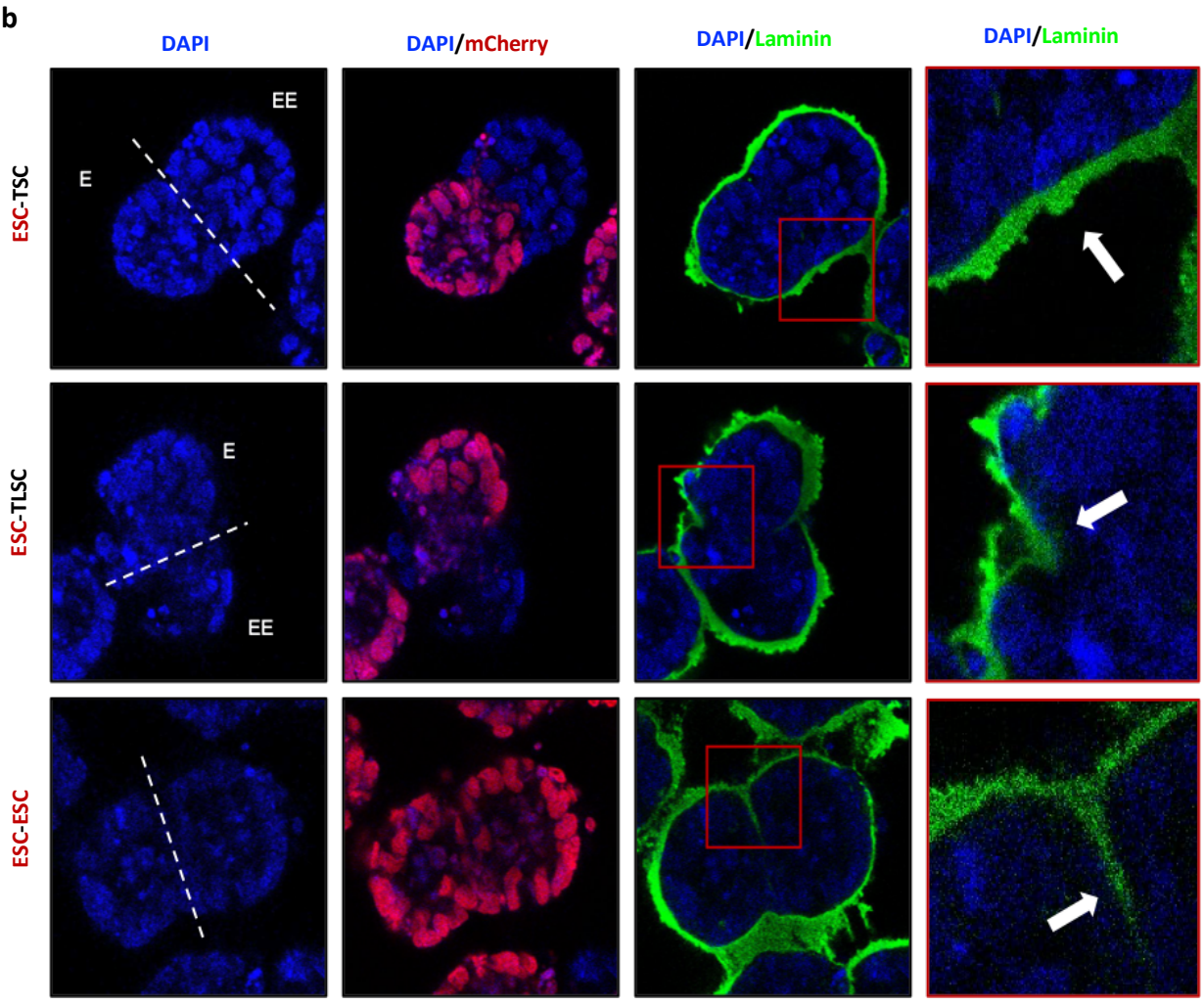


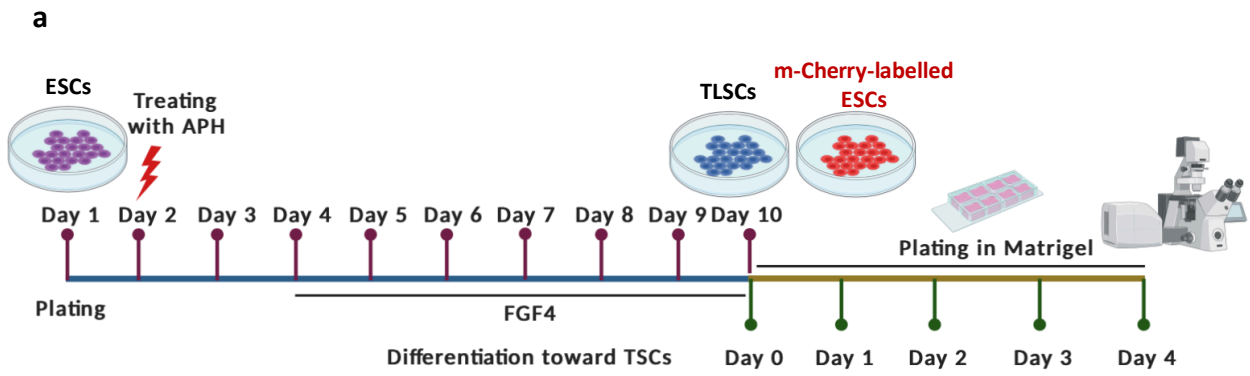
Figure 12. Characterization of ESC-TLSC structures

(a) Immunostaining of ESC-TLSC structures and ET embryos for the cell adhesion marker (E-cadherin). At least two biological replicates were performed per sample. **(b)** Immunostaining of ESC-TLSC structures, ET embryos and ESC-ESC aggregates for the basement membrane (Laminin). At least two biological replicates were performed per sample.

RS-induced TLSCs are able to give rise to the higher number of ET embryos

Following the generation of ET embryos from the assembly of ESCs with TLSCs **(Figure 11 and 12)**, we wondered about the differentiation potential of RS-induced ESCs and asked whether RS-induced TLSCs are functionally able to give rise to a higher number of ET embryos. To find the answer, we treated ESCs with APH and in parallel with untreated cells we differentiated them toward TSCs. Subsequent to the generation of TLSCs, we co-cultured them with mCherry-labeled ESCs to generate ET embryos (**Figure 13a**). On the fourth day, we fixed the structures and immunostained for the embryonic and extra-embryonic markers. RS-induced TLSCs were functional similar to the intact TLSCs as they could generate ET embryos (**Figure 13b**). Strikingly, through counting the number of structures we found significantly higher percentage of ET embryos in RS-induced condition (**Figure 13c**). In line with our previous results, indicating that RS increases the number of TLSCs (**Figure 8**), RS-induced TLSCs generated higher number of ET embryos.

In conclusion, these data indicate that RS-induced ESCs have functionally higher differentiation potential than untreated ESCs, since RS-induced ESC-derived TSCs could functionally give rise to a higher number of ET embryos than untreated ESC-derived TSCs (**Figure 13**).


a
ESCs
Treating with APH
TLSCs
m-Cherry-labelled ESCs
Day 1
Day 2
Day 3
Day 4
Day 5
Day 6
Day 7
Day 8
Day 9
Day 10
Plating in Matrigel
Plating
FGF4
Differentiation toward TSCs
Day 0
Day 1
Day 2
Day 3
Day 4

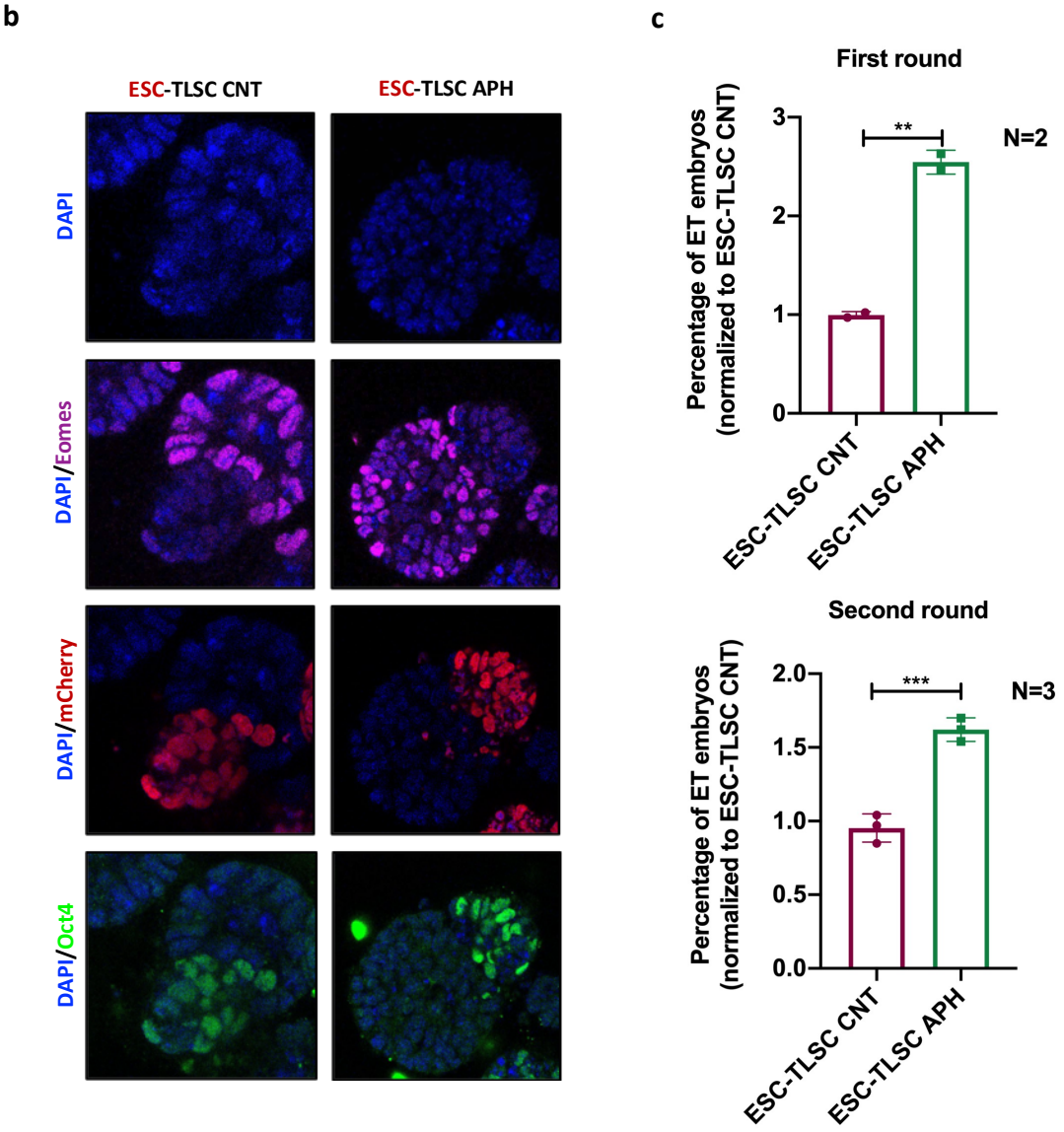

b
ESC-TLSC CNT
ESC-TLSC APH
DAPI
DAPI/Eomes
DAPI/mCherry
DAPI/Oct4
c
First round
Percentage of ET embryos (normalized to ESC-TLSC CNT)
**
N=2
ESC-TLSC CNT
ESC-TLSC APH
Second round
Percentage of ET embryos (normalized to ESC-TLSC CNT)
***
N=3
ESC-TLSC CNT
ESC-TLSC APH

Figure 13. RS-induced TLSCs are able to give rise to a higher number of ET embryos **(a)** Schematic view of ET embryo generation from the assembly of ESCs and TLSCs in the presence or absence of APH. **(b)** Immunostaining of ET embryos generated from the assembly of ESCs and TLSCs in the presence or absence of APH for the embryonic marker (Oct4) and the extra-embryonic marker (Eomes). At least two biological replicates were performed per sample. **(c)** Graphs showing the percentage of ET embryo generation from the assembly of ESCs and TLSCs in the absence or presence of APH in two separate biological replicates. N shows the number of technical replicates in each round. All bar plots show mean with ± SD (*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, unpaired t test).

RS induces the terminal differentiation of ESCs toward TGCs and increases the number of TLGCs

To further consolidate our finding regarding the higher differentiation potential of RS-induced ESCs than untreated ESCs **(Figure 13)**, we aimed to investigate the differentiation potential of ESCs to terminal TE lineages, namely TGCs under RS. To this end we took advantage of the published approach used by several groups to differentiate TLSCs to TLGCs (112, 113, 115, 116). In this approach following the removal of FGF4 and MEF conditioned medium, TSCs as well as TLSCs differentiate to TGCs (106). Through this approach, we first differentiated RS-induced ESCs together with intact cells to TLSCs and following withdrawal of FGF4, we differentiated these cells toward TGCs (**Figure 14a**). Then, we checked the expression of TGC-specific markers as well as the number of generated TLGCs. Phase contrast imaging confirmed the formation of TLGCs both in the presence and absence of APH (**Figure 14b, upper panel**). Moreover, IF staining showed the expression of placenta expressed transcript 1 protein (Plet1) in the generated TLGCs (**Figure 14b, lower panel**). Interestingly, through the cell counting we found that the number of TLGCs in RS-induced condition was significantly higher than these cells in untreated condition (**Figure 14c**) (131). In other words, RS remarkably increased the number of ESC-derived TGCs. Also, RT-qPCR analysis revealed the significant increase in the expression of *Prl2c2* (a TGC-specific gene) in RS-induced TLGCs (**Figure 14d**).

These data together suggest that RS induces the terminal differentiation of ESCs toward TGCs and increases the number of TLGCs. These evidence are in line with the inductive role of RS in the differentiation of ESCs toward TSCs, the progenitor of TGCs (**Figure 14**).

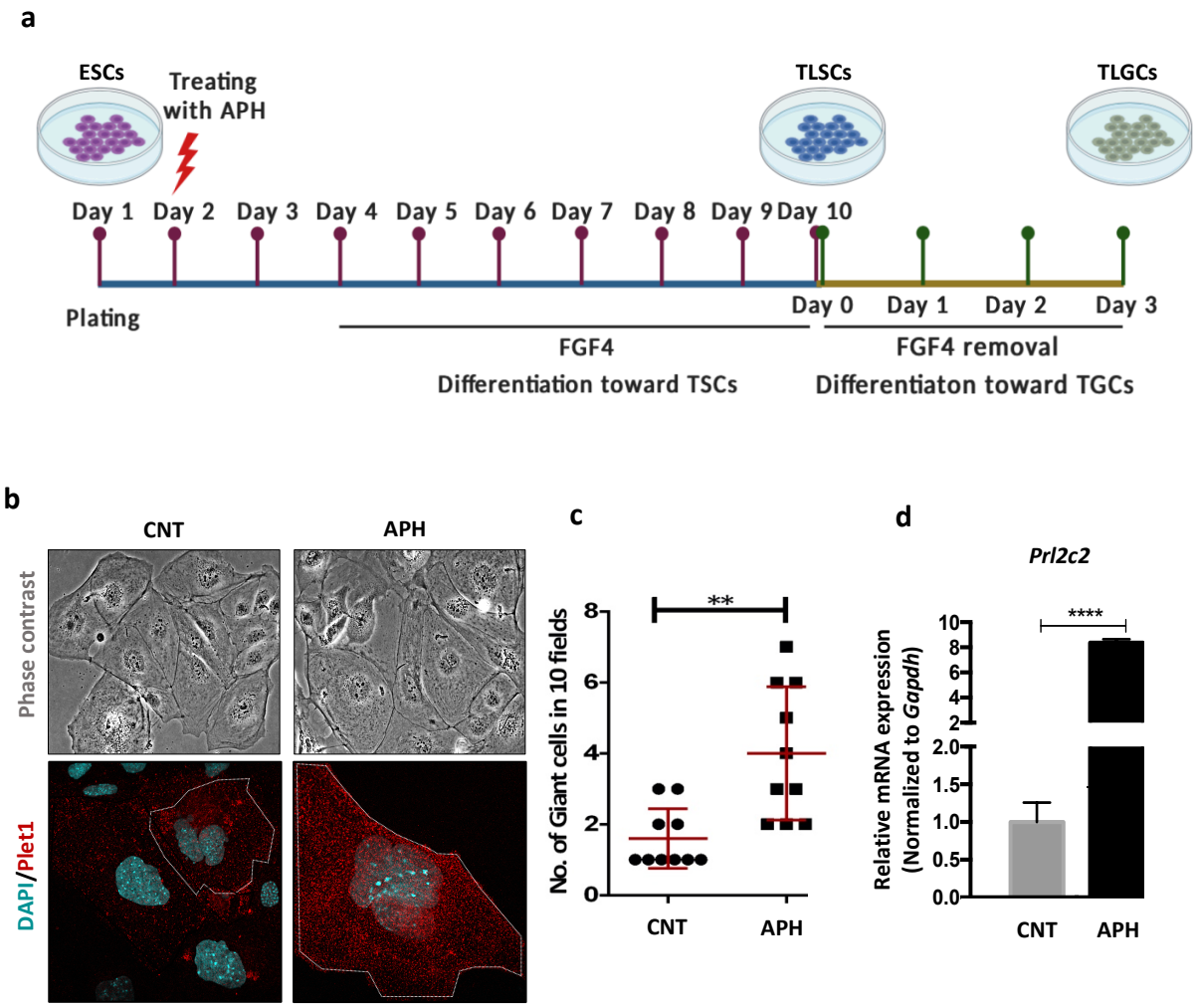


Figure 14. RS induces the terminal differentiation of ESCs toward TGCs and increases the number of TLGCs

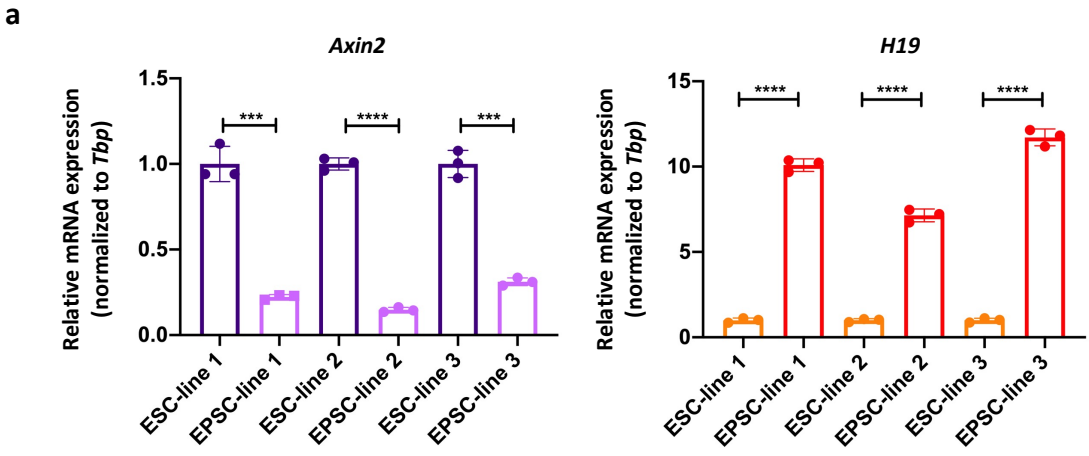
(a) Schematic view of ESCs differentiation toward TLGCs. **(b)** Phase contrast and immunofluorescence images of TLGCs for the TGC marker (Plet1) in the presence or absence of APH (Adapted from Atashpaz *et al.*, eLIFE, 2020) (131). **(c)** Graph showing the number of generated TLGCs from ESCs in the absence or presence of APH. **(d)** RT-qPCR analysis on TLGCs for the key TGC marker (*Prl2c2*) in the absence or presence of APH (Adapted from Atashpaz *et al.*, eLIFE, 2020) (131). All Bar plots show mean with ± SD (*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, unpaired t test).

RS does not activate the expression of key TE genes in EPSCs

As we found that RS activates the expression of TE-specific genes in ESCs and induces their differentiation toward TE lineages (**Figure 6 and 8**), we wondered whether activation of such TE genes could be mediated through RS in cleavage stage embryonic cells, namely EPSCs. These cells have been recently derived from 8C blastomeres and also ESCs through

culturing them in a medium which its components modulate signaling pathways that play key roles in TE/ICM segregation (EPSCM). When injected into the embryos, these cells are reported to contribute to both into the embryonic and extra-embryonic tissues. Moreover, it has been shown that EPSCs and ESCs are hierarchically segregated cell clusters in RNA-seq data analysis. Besides, the major development component revealed that EPSCs are in the range of 4C blastomeres (123, 124). To this end, we asked whether the effect of RS on the earlier-stages embryonic cells, in other words EPSCs, would be similar to the ESCs. To find the answer, we took advantage of the published protocol by Liu *et al.* which makes possible to generate L-EPSCs through passaging ESCs for five times in EPSCM. Significant reduction of *Axin2* mRNA level and remarkable increase of *H19* transcript level are the two important criteria of EPSC emergence from ESCs in EPSCM (123, 124). Therefore, we aimed to generate EPSCs from three different ESCs with distinct genetic backgrounds. RT-qPCR analysis showed considerable decrease in the expression of *Axin2* as well as the significant increase in the level of *H19* mRNA (**Figure 15a**). This result demonstrated the successful generation of L-EPSCs from all ESC lines. Next, we checked whether induction of RS could increase the expression of TE genes in EPSCs. To this end, we treated EPSCs along with ESCs with APH. Interestingly, RT-qPCR analysis demonstrated the significant activation of TE-specific genes such as *Cdx2*, *Elf5* and *Eomes* in ESCs while, this was not found in EPSCs (**Figure 15b**).

These data suggest that RS does not activate the expression of TE markers in EPSCs and in other words, the RSR-mediated activation of TE-specific genes is ESC-stage specific (**Figure 15**).



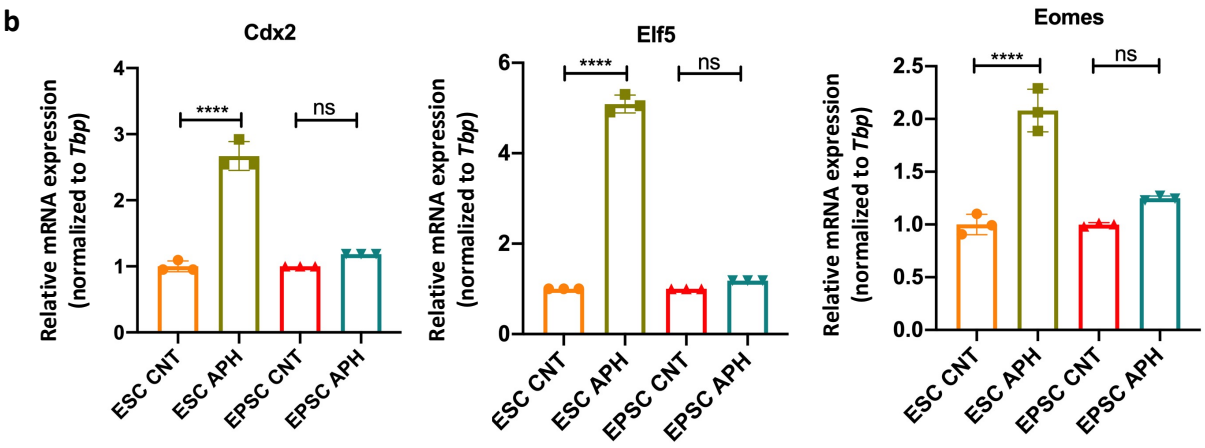


Figure 15. RS does not activate the expression of key TE genes in EPSCs

(a) RT-qPCR analysis on ESCs and EPSCs for *Axin2* and *H19*. Bar plots show mean with ± SD (*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, unpaired t test). **(b)** RT-qPCR analysis on ESCs and EPSCs for the key TE markers (*Cdx2*, *Eomes* and *Elf5*) in the presence or absence of APH. Bar plots show mean with ± SD (*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, one-way ANOVA).

ATR-Chk1-mediated RSR activates the expression of TE genes in ESCs

Next, we aimed to gain insight to the molecular mechanisms through which RSR increases the expression of TE transcriptional network in ESCs. To this end, we treated ESCs with APH and at the same time we inhibited DDR pathways with specific ATMi and ATRi. Interestingly, RT-qPCR analysis showed that upon induction of RS the expression of canonical TE genes such as *Eomes* and *Elf5* significantly increased and only inhibition of ATR but not ATM could decrease the expression of such genes (**Figure 16a**). We further consolidated this finding through knocking down (KD) of *Atr* in both untreated and APH-treated ESCs as RT-qPCR analysis confirmed the ATR-dependent expression of *Eomes* (**Figure 16b**). Moreover, western blot (WB) analysis confirmed the RT-qPCR results showing that the increased expression of Eomes in RS-induced ESCs was ATR- but not ATM-dependent (**Figure 16c and d**). Interestingly increased expression of Eomes was mediated through phosphorylation of Chk1, which is downstream to the ATR pathway. While Chk2 phosphorylation was not found in APH-treated cells.

These results together demonstrate that activation of TE-specific genes in RS-induced ESCs is ATR-Chk1 but not ATM-Chk2-dependent. Concisely, ATR-Chk1-mediated RSR activates the expression of TE genes in ESCs (**Figure 16**).

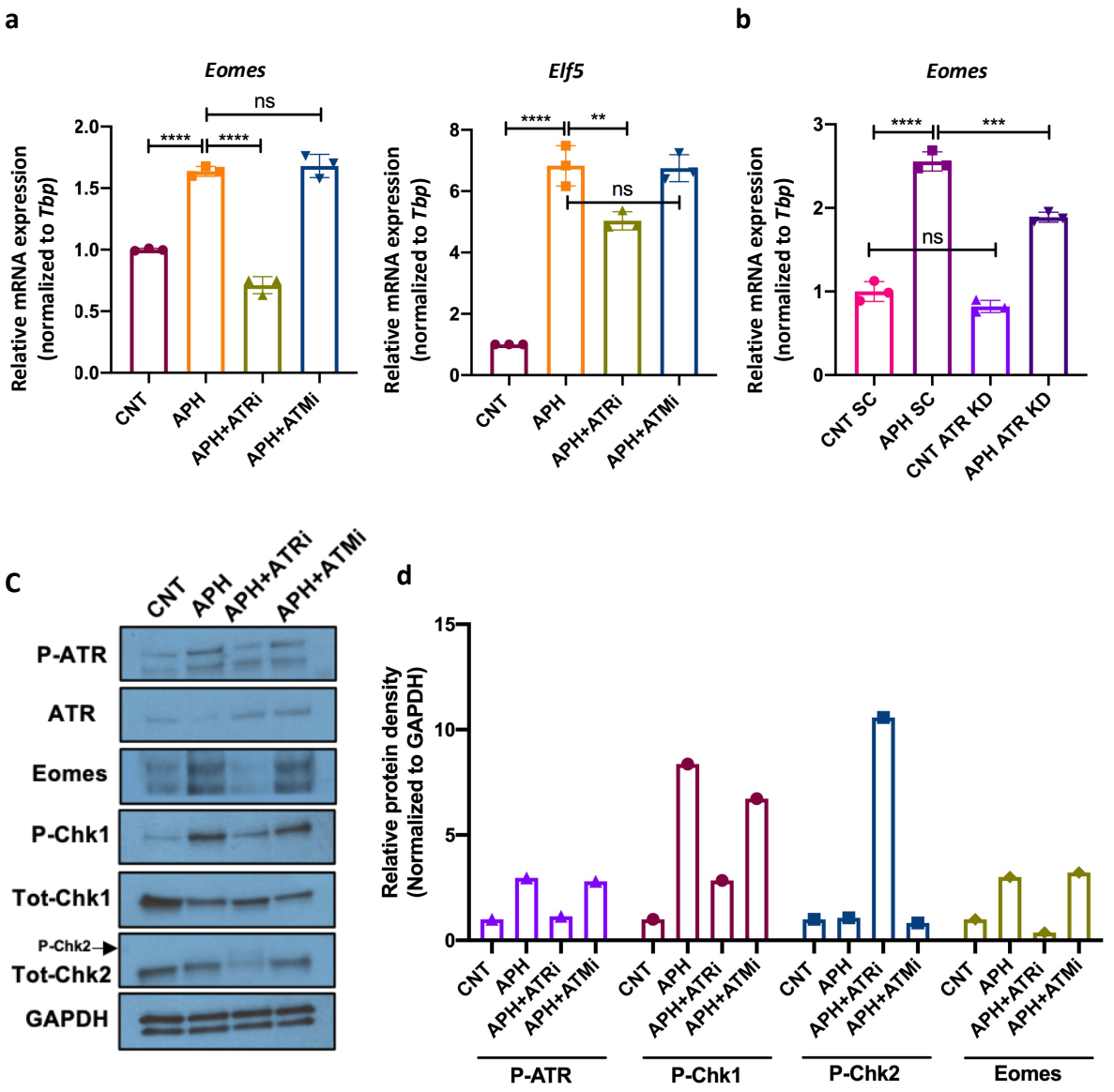


Figure 16. ATR-Chk1-mediated RSR activates the expression of TE genes in ESCs

(a) RT-qPCR analysis on ESCs for key the TE markers (*Eomes* and *Elf5*) treated with APH and ATR/ATM inhibitors. **(b)** RT-qPCR analysis on ESCs for the key TE marker (*Eomes*) treated with APH and/or esiRNA against *Atr*. **(c)** Immunoblot for the phosphorylation status of the key DDR kinases (ATR, Chk1 and Chk2) and expression of the key TE marker (Eomes) upon treatment with APH and ATM/ATR inhibitor in ESCs. **(d)** Graph shows the protein density of phosphorylated ATR, phosphorylated Chk1 and phosphorylated Chk2 and the key TE marker (Eomes) upon treatment with APH and ATM/ATR inhibitor in ESCs. All bar plots show mean with ± SD (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$, one-way ANOVA).

RS activates the expression of TE genes through ATR-Chk1-mediated binding of Tead4 to the transcription factor-binding motifs of TE-specific genes

Next, we aimed to identify the molecular mechanism downstream to ATR-Chk1 pathway that increases the expression of TE transcriptional network in ESCs. To this end, we focused on Tead4, the master regulator of TE-specific transcription network which plays a central role in TE/ICM segregation in the embryos. Home *et al.* reported that the subcellular localization of Tead4 plays a key role in this first lineage segregation. Briefly, in 8C embryos they found Tead4 abundantly expressed in the nucleus of blastomeres. However, after compaction they observed Tead4 in the nucleus of the outer cells, while in the ICM they mostly found it in the cytoplasm. This group detected Tead4 only in the cytoplasm of ESCs, while in TSCs they observed it both in the nucleus and cytoplasm with higher abundance in the nucleus. Therefore, the absence of Tead4 in the nucleus of ESCs prevents its chromatin binding, leading to the decreased expression of TE genes in these cells. Besides, this group showed that the mRNA and protein levels of Tead4 in mouse ESCs is low compared to TSCs (90). With respect to the Home *et al.* study, we wondered whether RS alters the subcellular localization of Tead4 in ESCs. To find the answer, we immunostained intact and RS-induced ESCs along with TSCs using a Tead4 antibody. In contrast to the Home *et al.* report, we found Tead4 not only in the cytoplasm but also in the nucleus of ESCs. However, in line with the Home *et al.* study, we found Tead4 in both the nucleus and cytoplasm of TSCs albeit with higher abundance in the nucleus. Interestingly, RS increased the level of Tead4 in both compartments of ESCs. Even so, the expression of Tead4 in RS-induced ESCs had foci pattern which was different from its diffused pattern in TSCs (**Figure 17a**). According to the immunostaining results, we could not conclude that the increased level of nuclear Tead4 in RS-induced ESCs is related to its subcellular localization. In other words, the Tead4 cytoplasm to nucleus repositioning was not detectable. This is firstly because the level of Tead4 increased both in the nucleus and cytoplasm of RS-induced ESCs. Secondly, the Tead4 is already observable in the nucleus of untreated ESCs. To further elucidate the mechanisms underlying the Tead4 increase, we checked the level of Tead4 transcript and protein in RS-induced ESCs and untreated cells through RT-qPCR and immunoblotting. Consistent with immunostaining results, interestingly we found that the level of Tead4 transcript and protein significantly increased upon APH treatment in ESCs (**Figure 17b-d**).

Taken together, these results show that in response to RS the Tead4 expression increases in ESCs (**Figure 17**).

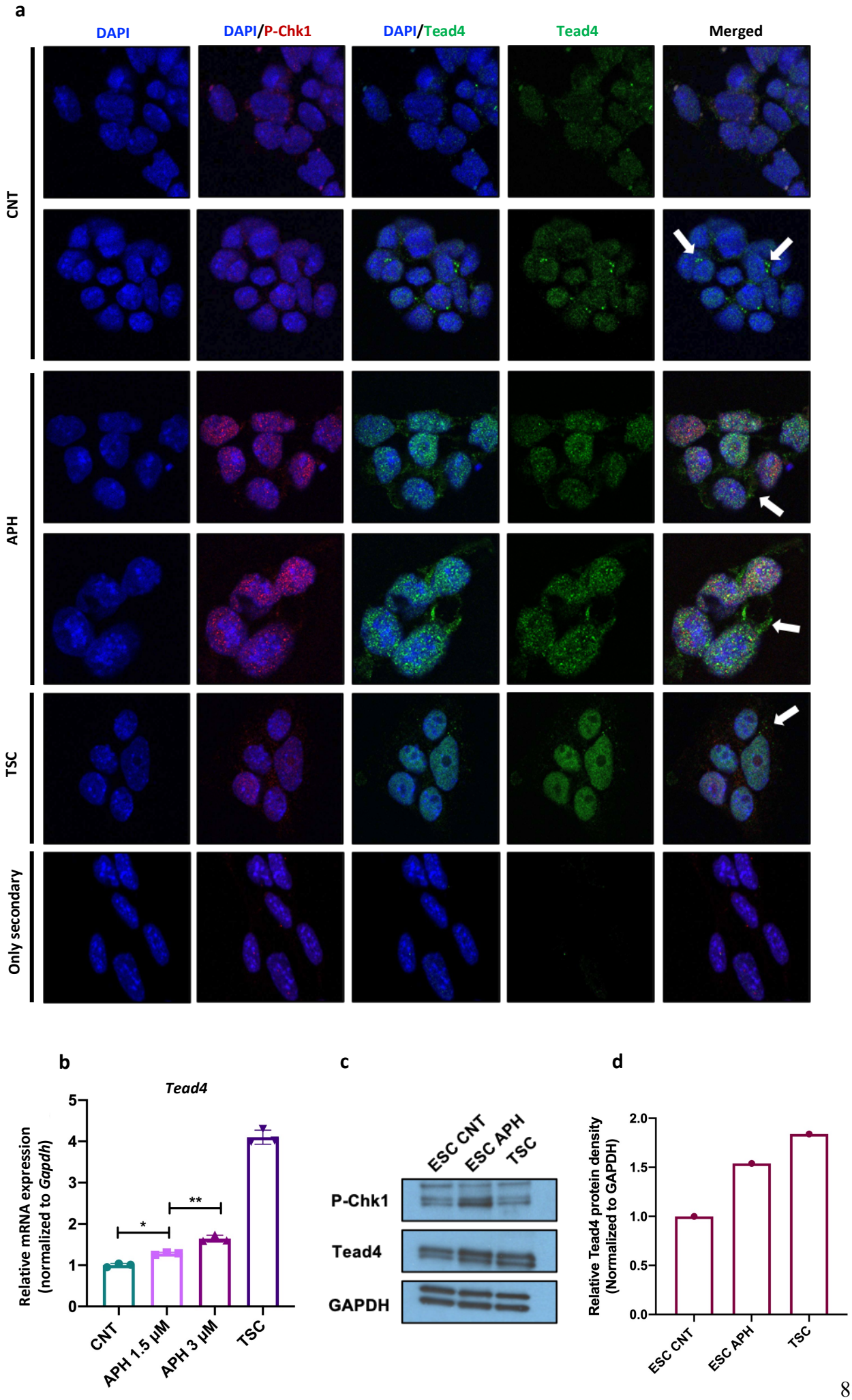

a
DAPI
DAPI/P-Chk1
DAPI/Tead4
Tead4
Merged
CNT
APH
TSC
Only secondary
b
Tead4
Relative mRNA expression (normalized to Gapdh)
CNT
APH 1.5 μM
APH 3 μM
TSC
c
ESC CNT
ESC APH
TSC
P-Chk1
Tead4
GAPDH
d
Relative Tead4 protein density (Normalized to GAPDH)
ESC CNT
ESC APH
TSC

Figure 17. RS increases the expression of Tead4 in ESCs

(a) Immunostaining of ESCs for Tead4 and P-Chk1 in the presence and absence of APH. TSCs were used as a positive control for Tead4 expression. The white arrows show the Tead4 presence in the cytoplasm. **(b)** RT-qPCR analysis on ESCs for *Tead4* upon treatment with different concentrations of APH. **(c)** Immunoblot for the phosphorylation status of Chk1 and Tead4 expression upon treatment with APH in ESCs. The protein extract of TSCs were used as a positive control for Tead4 expression. **(d)** Graph shows the protein density of Tead4 in untreated ESCs, APH-treated ESCs and TSCs.

Next, we asked whether the increased expression of Tead4 in RS-induced ESCs is mediated through ATR-Chk1 pathway. Hence, we treated ESCs with APH and inhibited DDR pathways with ATRi and ATMi. According to the RT-qPCR and WB analysis, we found that the increased expression of Tead4 is neither ATR-Chk1 nor is ATM-Chk2 dependent (**Figure 18**).

These data together suggest that although similar to the TE-specific genes the Tead4 expression increases in response to RS, however in contrast to the TE-specific genes the increased expression of Tead4 is not mediated through ATR-Chk1 pathway (**Figure 18**).

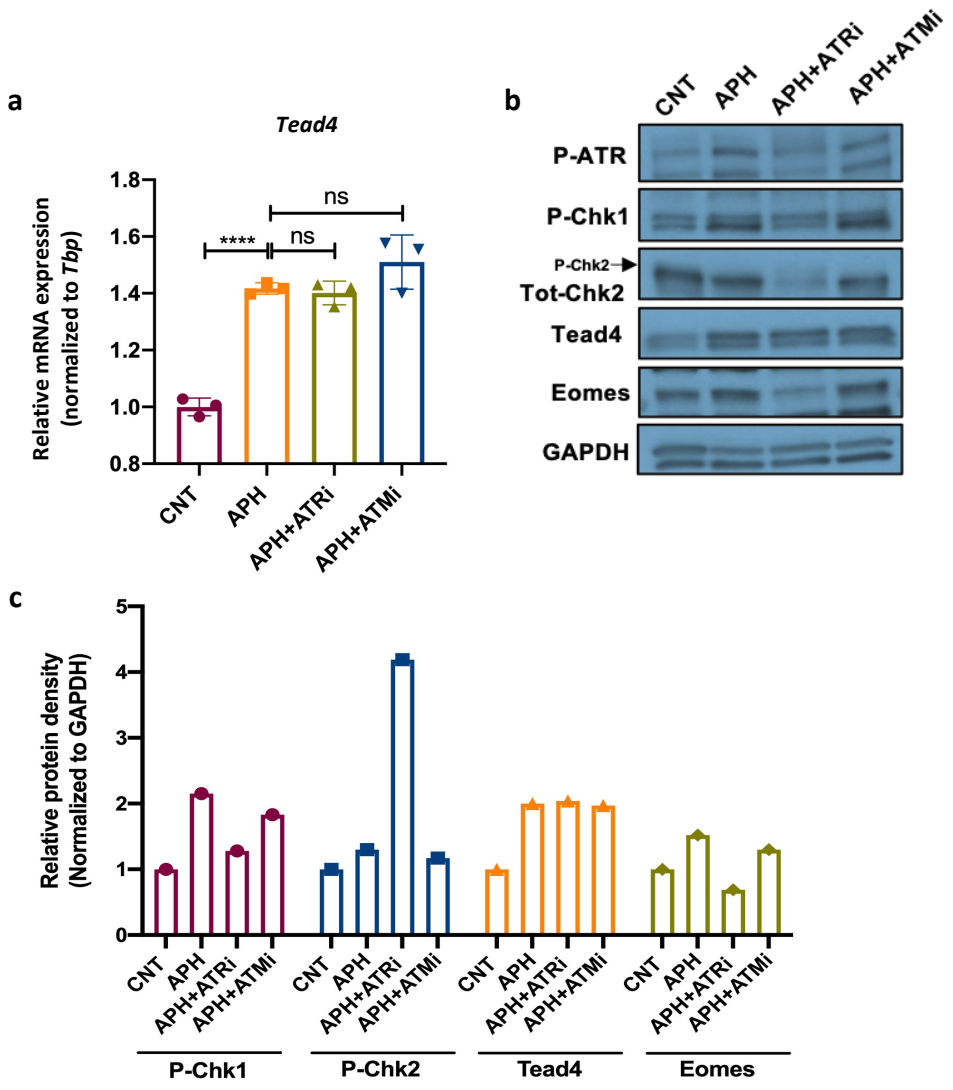


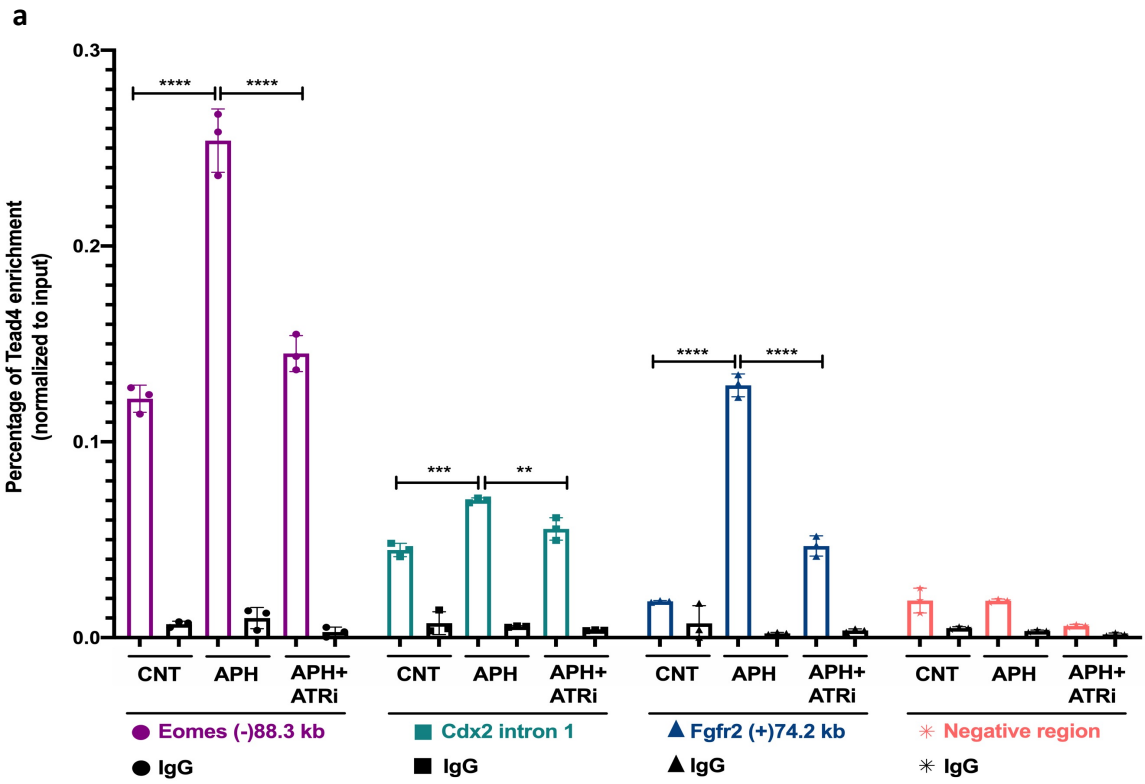
Figure 18. RS increases the expression of Tead4 independently of ATR-Chk1 pathway in ESCs

(a) RT-qPCR analysis on ESCs for *Tead4* upon treatment with APH and ATR/ATM inhibitor. **(b)** Immunoblot for the phosphorylation status of key DDR kinases (ATR, Chk1 and Chk2) and expression of the key TE marker (Eomes) and Tead4 upon treatment with APH and ATM/ATR inhibitor. **(c)** Graph shows the protein density of phosphorylated Chk1 and Chk2 and the key TE marker (Eomes) and Tead4 upon treatment with APH and ATM/ATR inhibitor in ESCs.

Subsequently, with respect to our findings and according to the Nishioka *et al.* study which shows that OE of Tead4 induces trophoblast fate in ESCs (109), we wondered whether ATR-Chk1-mediated expression of TE genes in ESCs is Tead4-dependent in spite of the fact that the expression of Tead4 per se is not. Tead4 is reported to bind to the transcription factor-

binding motifs (TFBMs) of TE-specific genes in TSCs (90). Therefore, we wondered whether RS increases the expression of TE-specific genes through altering the enrichment of Tead4 to the TFBMs of such genes in ESCs. Through chromatin immunoprecipitation (ChIP) analysis, strikingly we found that the Tead4 binding to the TFBMs of canonical TE-specific genes such as *Eomes*, *Fgfr2*, *Cdx2* and *Gata3* significantly increased and, interestingly, this binding was ATR-dependent (**Figure 19**).

In conclusion, these data suggest that RS induces TE cell fate determination in ESCs through a two-step mechanism. Firstly, RS increases the level of Tead4 mRNA and protein independently of ATR-Chk1 activation. Secondly, the ATR-CHK1-mediated RSR provokes the occupancy of Tead4 at TFBMs of TE-specific genes (**Figure 17-19 and 21**).



b

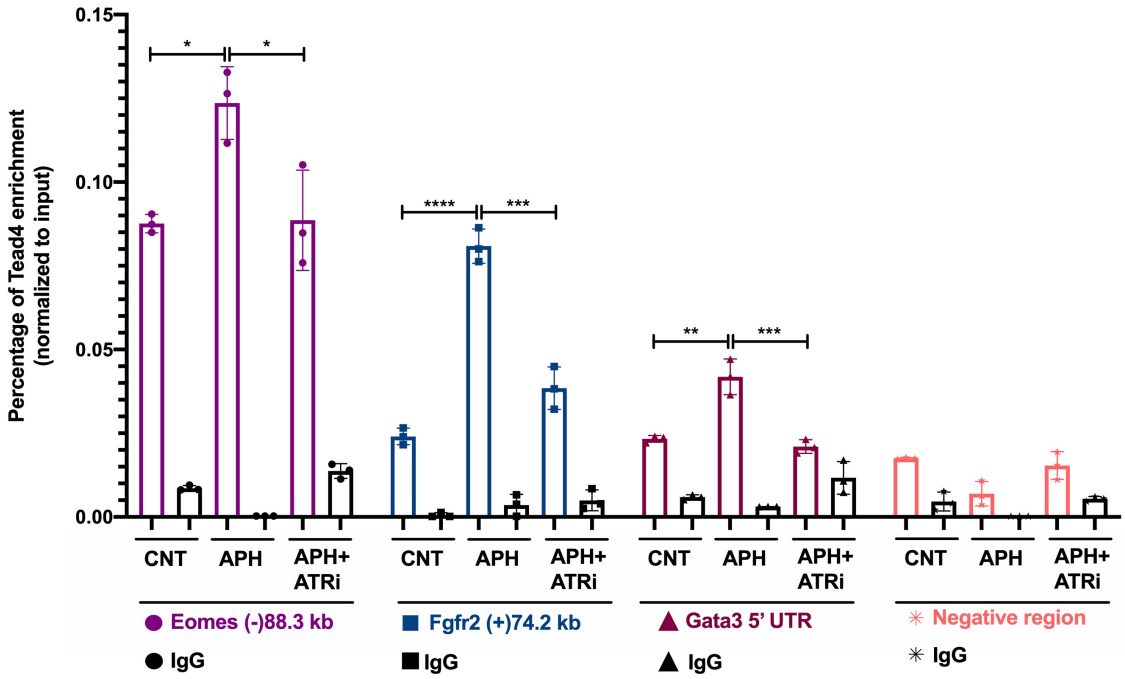


Figure 19. RS activates the expression of TE genes through ATR-Chk1-mediated binding of Tead4 to the transcription factor-binding motifs of TE-specific genes

(a) and **(b)** Graphs show the Tead4 enrichment at the TFBMs of key TE genes (*Eomes*, *Cdx2* and *Fgfr2*) and (*Eomes*, *Fgfr2 and Gata3*) upon APH treatment and ATR inhibition. The graphs show the data of two separate biological replicates. All bar plots show mean with ± SD (*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, one-way ANOVA).

ATR-Chk1-mediated binding of Eomes to its transcription factor binding motif enhances its expression in RS-induced ESCs

Kidder and Palmer have reported that Eomes transcriptionally regulates its gene as well as some TE-specific genes through binding to their promoters (135). As we found that the expression of Eomes in RS-induced ESCs is ATR-dependent, we wondered whether binding of Eomes to its TFBM is mediated through ATR. To find the answer, we performed ChIP analysis and checked the occupancy of Eomes at its TFBM. Surprisingly, we found that Eomes binding significantly increased to its TFBM in response to RS, and this binding was ATR-dependent (**Figure 20**).

This result shows that Eomes binding to its own TFBM is enriched through ATR-Chk1-mediated RSR. This is similar to Tead4, which its binding to the TFBM of TE genes is ATR-Chk1-dependent. Consequently, the increased binding of Eomes to its TFBM further enhances its expression in RS-induced ESCs (**Figure 20 and 21**).

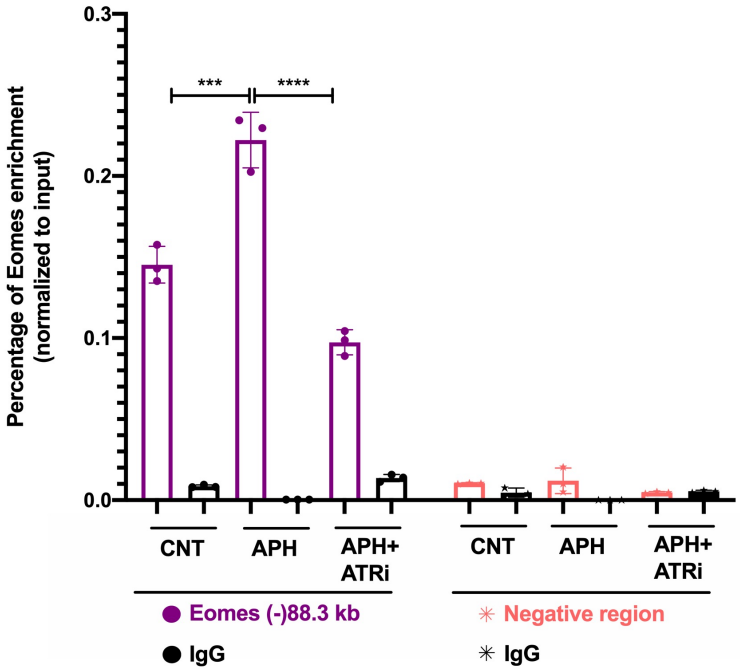


Figure 20. ATR-Chk1-mediated binding of Eomes to its transcription factor binding motif enhances its expression in RS-induced ESCs

Graph shows the Eomes enrichment at the TFBMs of its gene upon APH treatment and ATR inhibition. Bar plot shows mean with ± SD (*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, one-way ANOVA).

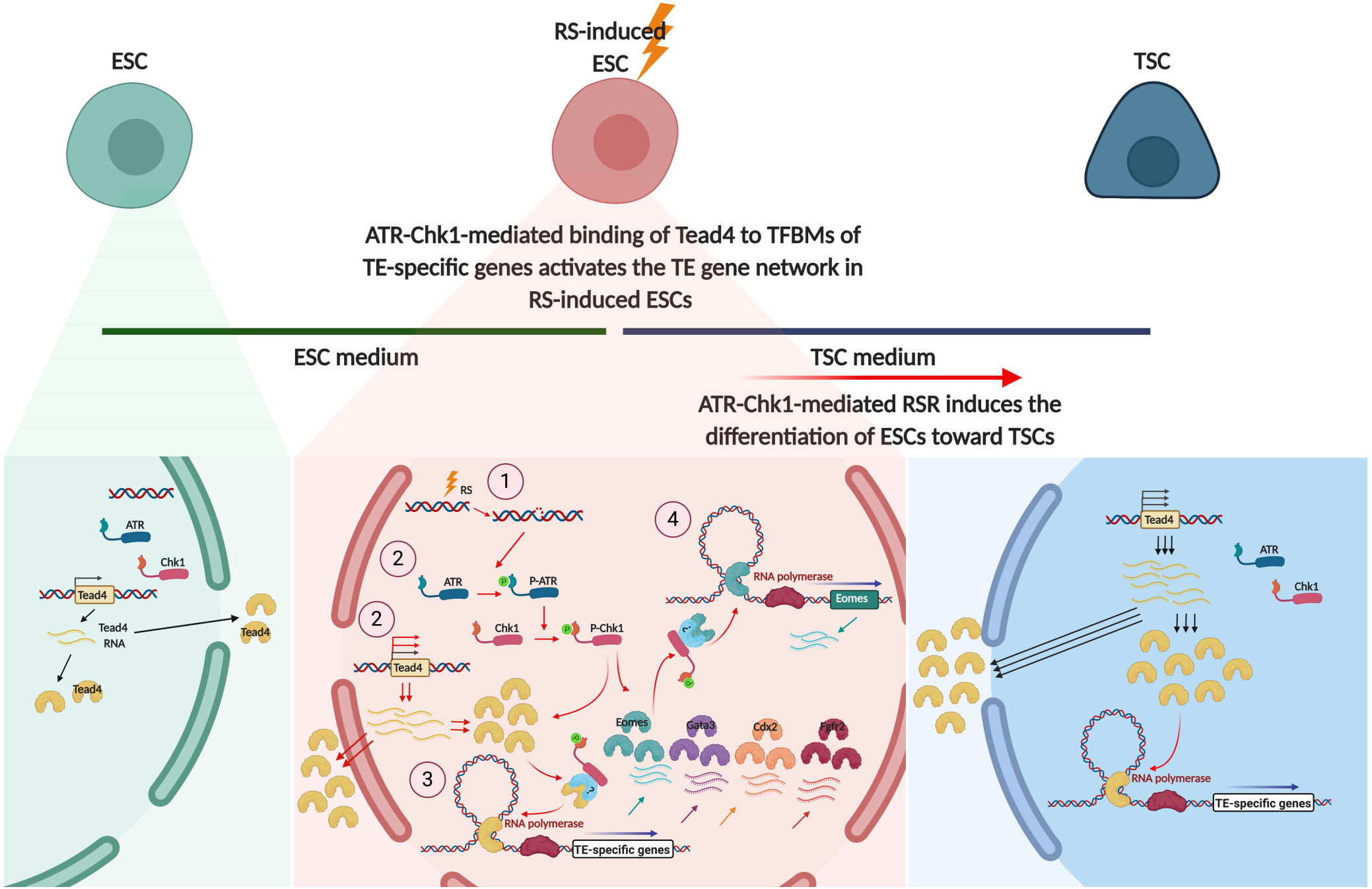

ESC
RS-induced ESC
TSC
ATR-Chk1-mediated binding of Tead4 to TFBMs of TE-specific genes activates the TE gene network in RS-induced ESCs
ESC medium
TSC medium
ATR-Chk1-mediated RSR induces the differentiation of ESCs toward TSCs
ATR
Chk1
Tead4
Tead4 RNA
Tead4
Tead4
RS
1
2
ATR
P-ATR
2
Chk1
P-Chk1
Tead4
4
RNA polymerase
Eomes
Eomes
Gata3
Cdx2
Fgfr2
3
RNA polymerase
TE-specific genes
Tead4
ATR
Chk1
RNA polymerase
TE-specific genes

Figure 21. Schematic model of ATR-Chk1-mediated RSR in ESCs

ESCs (green) express Tead4 at low level, while TSCs (blue) highly express this TF. Subsequently, Tead4 molecules bind to the TFBMs of TE genes and activate their expression in TSCs. Under RS conditions (red) (1), ESCs activate ATR and Chk1 through their phosphorylation (2). In parallel, RS-induced cells increase the expression of Tead4 (2). Subsequently, activated ATR-Chk1 pathway stimulates the binding of highly expressed Tead4 molecules to the TFBMs of TE genes. Hence, this pathway induces the expression of TE genes in ESCs (3). Besides, the ATR-Chk1-mediated binding of Eomes to its TFBM further enhances its expression in RS-induced ESCs (4). Under TSC culture condition, the ATR-Chk1-mediated RSR induces the differentiation of ESCs toward TSCs.

***All the schematic diagrams are created with the ©BioRender unless otherwise they are indicated to be adopted or reprinted from a reference.

***All the graphs are created with the Prism software.

***All the FACS histograms are created with the FlowJo software.

Discussion

Pre-implantation embryos and their derived cells, called as ESCs, are vulnerable to various genotoxic stresses due to the rapid proliferation. It has been reported that in response to DD, ESCs undergo apoptosis or differentiation, but not senescence, to be eliminated from the pluripotent pool (94, 101, 131, 136). Although induced differentiation of adult stem cells by DD has been reported by several groups (137, 138), the DD-induced differentiation of ESCs have remained elusive. We have recently found, through *in vivo* chimera assay, that in response to RS ESCs contribute to the extra-embryonic compartment of the embryos and co-localize with Cdx2-expressing cells (**Figure 5**). This phenomenon probably restricts the incorporation of defective cells in the embryonic proper and later to the progeny through diverting them toward the extra-embryonic tissues (131).

Since working with live embryos has several limitations, ranging from ethical issues to restricted manipulation possibilities, during this study we established *in vitro* models of ESC differentiation to the extra-embryonic lineages to further investigate the fate of ESCs in response to RS. Briefly, in line with our *in vivo* findings we revealed that first, RS activates the expression of key TE genes in mESCs and second, under TSC culture conditions, RS induces the differentiation of ESCs toward TSCs which ultimately gives rise to the higher number of TLSCs (**Figure 6 and 8**). These results are in consistent with several studies which have reported the successful generation of TLSCs through OE of key TSC TFs in ESCs (106-110).

Beside observing a higher number of generated TLSCs and accordingly significant activation of TE-specific markers in response to RS, we checked whether TLSCs are functionally similar to the embryo-derived TSCs. Harrison *et al.* have reported an *in vitro* 3D stem cell embryogenesis approach, which makes possible to recapitulate embryogenesis in a dish (**Figure 4**). Concisely, through co-culturing ESCs with TSCs in a Matrigel-coated plate they could generate embryo-like structures (ET embryos). Through such structures, they could observe the spatial and temporal sequences of events happening during embryogenesis (118). To check the functionality of our TLSCs, we took advantage of this method and established an innovative combinatorial approach of ESC *in vitro* differentiation with 3D stem cell embryogenesis. In other words, we co-embedded our mCherry-labeled ESCs with generated TLSCs and interestingly noticed similar sequential events of embryogenesis (i.e., formation of a single central cavity and laminin disappearance) with TLSCs comparable to the TSCs (**Figure 11 and 12**). Through this approach, we strikingly

found that RS-induced TLSCs give rise to a higher number of ET embryos (**Figure 13**). This finding is likely linked to the inductive impact of RS on the differentiation of ESCs toward TSCs, which leads to the generation of higher number of TLSCs. In other words, the higher is the number of TLSCs, the higher would be the possibility of their interaction with ESCs to generate the ET embryos. Conclusively, we found higher number of these structures with RS-induced TLSCs. These findings are in line with our *in vivo* findings (**Figure 5**) and further strengthen our hypothesis regarding the safeguard mechanism during embryogenesis through which the embryo limits the incorporation of defective cells in the embryonic compartment by diverting them toward the extra-embryonic tissues (131). The majority of embryo loss takes place around the implantation period. This is due to the vulnerability of rapidly proliferating pre-implantation embryos to the genotoxic insults (95, 96). Hence, the established *in vitro* combinatorial approach could serve as a model to study the impact of genotoxic stress during early embryogenesis. Moreover, the fate of various ESCs with deficiency in DDR pathways could be studied through this system. For instance, ATR-deficient Seckel ESCs, which have dramatically lower level of ATR kinase than the WT ESCs (81) and CHK1 haploinsufficient ESCs, which express considerably lower level of CHK1 than WT ESCs (139) could be investigated within this approach to unravel their fate during early embryonic stages.

Besides, the combinatorial 3D embryogenesis approach could be used as a platform to model diseases related to the early embryonic development. Furthermore, the fate of patient-derived iPSCs could be investigated through this approach. In other words, iPSCs with different genetic backgrounds or genetically modified ESCs could be either first differentiated *in vitro* and then co-cultured with WT ESCs in a 3D matrix. Or they could be co-embedded with WT TLSCs or TSCs to generate ET embryos. Careful investigation of the morphology, size, expressing markers as well as the number of generated structures could hopefully unravel the mechanisms of diseases related to such iPSCs or genetically modified ESCs. Similar to the 3D organoids, the ET embryos could be controlled with minimum external manipulation. Moreover, in comparison to the *in vivo* studies, they are reproducible, accessible, cost-effective with fewer ethical challenges (129). Therefore, the ET embryos with a specific phenotype or pathological and genetic abnormalities could be used to identify effective drug candidates and to assess their potential toxicities. Ultimately, such studies hopefully could decrease the rate of early pregnancy loss as well as the prenatal abnormalities.

It is critical to note that, for investigating the differentiation potential of ESCs to TLSCs through ET embryos, we took advantage of mCherry-expressing ESCs as embryonic compartment. These cells carry the coding sequence of fused H2B-mCherry reporter at an endogenous constitutively expressed locus called as ROSA26 locus. Since this sequence is flanked by *LoxP* sites (134), through their excision by Cre, we could ubiquitously express H2B-mCherry and establish an expressing system with the least probability of reporter silencing (**Figure 10**). This system circumvents the problem of mCherry silencing that initially we had with ESCs that were infected with EF1α-mCherry constitutively expressing vector. The EF1α-mCherry silencing is due to the random integration of the virus in different parts of the genome. Therefore, the risk of epigenetic silencing of the reporter gene after repeated cell passages is considerably high in this system.

We also gained further insight into the RSR-mediated differentiation of ESCs toward terminal TE lineages i.e., TGCs. Interestingly, we found that RS not only induces the initial differentiation of ESCs toward TSCs but also provokes their terminal differentiation toward TGCs. Hence, RS increases the number of generated ESC-derived TGCs (**Figure 14**). According to the several reports, through withdrawal of FGF4 and MEF conditioned medium from TSC culture, these cells differentiate to TGCs (106). These cells are a fundamental component of the placenta as they deeply penetrate into the endometrium and make connection with maternal arteries (92). Removal of FGF4 and MEF conditioned medium from TLSC culture also caused the formation of TLGCs, and strikingly, RS further enhanced the formation of giant cells (**Figure 14**). This finding confirms that TLSCs behave similar to the embryo-derived TSCs. Since under terminal differentiation culture conditions, TLSCs activate TGC-specific markers comparable to the terminally differentiated TSCs and acquire the morphology of actual TGCs.

As the generation and maintenance of earlier-stages embryonic cells i.e., EPSCs have been recently reported (123-125), we wondered whether the RSR in such cells would be similar to ESCs. Briefly, EPSCs have been recently obtained through targeting the signaling pathways that play key roles in TE/ICM segregation in the ESC culture. EPSCs (expanded potential stem cells) take their name from their capability to contribute to both the embryonic proper and the extra-embryonic tissues when they are injected into the embryos (123-125). RNA-seq data analysis reveals that EPSCs cluster are hierarchically segregated from ESC cluster and they are in the range of 4C blastomeres according to the major development component analysis (123, 124). With respect to the facilitated contribution of EPSCs to the extra-embryonic tissue, we were expecting a remarkable activation of TE-specific genes in

response to RS in such cells rather than ESCs. However, interestingly in contrast to the ESCs, we did not find any increase in the expression of TE markers in RS-induced EPSCs (**Figure 15**). This suggests that the RSR-mediated activation of TE-specific genes is ESC-stage specific. On the other hand, a study from Rossant lab raises a controversy about the expanded potential property of EPSCs. Through bulk and scRNA-seq, they showed that the transcriptome of EPSCs is distinct from the blastomeres and is closer to the EPI and EPiSCs. Moreover, they did not observe the activation of most of the 4-16C stage embryo genes in EPSCs. In addition, they did not report a significant activation of TSC-specific genes following *in vitro* differentiation of EPSC to the TSCs. They further studied the developmental potential of EPSCs through their injection into the embryos and investigated their contribution to the extra-embryonic tissues at different embryonic stages. At E4.5 and E6.25 although they found EPSCs in the extra-embryonic compartment, these cells were not expressing TE-specific markers such as Cdx2, Elf5 and Tcfap2c. However, most of them were expressing Oct4, the ESC-specific marker. Therefore, they suggest that, when EPSCs are injected into the embryos they mostly mis-localize in the extra-embryonic tissues and do not readily contribute to such tissues because of their expanded developmental potential (126). Although we could successfully generate embryonic cells with similar transcriptional and morphological characteristics of EPSCs reported by Yang *et al.* (123, 124), we could not find any alteration in the expression of TE-specific genes in response to RS in such cells (**Figure 15**). This finding raises two possibilities: first, according to the major development component analysis of Yang *et al.* these cells are in the range of 4C blastomeres but do not respond to RS similar to the ESCs (123, 124); second, as claimed by Posfai *et al.* these cells are distinct from blastomeres and are more at later developmental stages than ESCs (i.e., EPI and EPiSCs) (126). Consequently, they have a completely different global gene expression pattern compared to the ESCs (140). Therefore, they cannot activate TE markers similar to the ESCs under RS. Ultimately, regardless of the actual developmental stage of EPSCs, they are distinct from ESCs. Moreover, due to the fact that the RSR-mediated activation of TE-specific genes does not happen in EPSCs in contrast to the ESCs, we suggest that this response is ESC-stage specific.

We also focused on the DDR mechanisms through which RS could activate the expression of TE gene network in ESCs. Interestingly, we found that the activation of TE markers in RS-induced ESCs is ATR-Chk1 but not ATM-Chk2-dependent (**Figure 16**). Next, we aimed to go into more depth in studying the mechanisms triggered by ATR-Chk1 pathway that lead to the activation of TE genes. Therefore, we focused on one of the signaling pathways that plays a central role in TE/ICM segregation in the embryos, the Hippo signaling pathway.

This pathway is composed of three major components Lats1/2, Yap1, Yap1-related protein Wwtr1 (TAZ) and Tead4. According to the several studies, in the outer cells of the embryo the interaction of transcription coactivators (Yap1/TAZ) with Tead4 activates this master regulator of TE network. While phosphorylation of Yap1 in the inner cells of the embryo prevents the activation of Tead4 and subsequently leads to the ICM fate (**Figure 3**) (87, 90, 109, 141). However, there are some debates about how Tead4 activation determines the TE fate. In details, according to Nishioka *et al.*, Hirate *et al.*, Sasaki and several others, Tead4 is constantly nuclear both in the outer and inner cells of the embryo. However, Lats1/2-mediated phosphorylation of Yap1 regulates Yap1 nuclear localization and consequently Tead4 activation. Briefly, in the inner cells of the embryo Lats1/2 phosphorylates Yap1 promoting its retention in the cytoplasm. Therefore, Yap1 cannot enter the nucleus and thus cannot bind to the Tead4 to form the Yap1-Tead4 active complex. Consequently, Tead4 cannot bind to the TFBMs of TE genes and they remain silent in the inner cells. On the contrary, in the outer cells, the unphosphorylated Yap1 accumulates in the nucleus, forms a complex with Tead4 and initiates the expression of TE gene network (87, 109, 114). In contrast to this model, Home *et al.* report another observation which suggests that Tead4 subcellular localization but not Yap1 regulates the TE/ICM segregation. In their study, Home *et al.* found Tead4 only in the cytoplasm of ESCs. While in TSCs they detected it both in the nucleus and cytoplasm with higher abundance in the nucleus. Therefore, they suggest that the nuclear accumulation of Tead4 in the outside cells determines the TE fate (90). To check the expression of Tead4 as well as the possible alteration of its subcellular localization in response to RS, we performed immunostaining and interestingly found that the expression of Tead4 increases in RS-induced ESCs (**Figure 17a**). It is therefore possible that the high level of Tead4 induced by RS initiates the expression of TE gene network in ESCs similar to the TSCs. This is in consistent with the Nishioka *et al.* study in which they show that the OE of Tead4 induces trophoblast fate in the ESCs (109). In line with the Home *et al.* observation we found Tead4 both in the cytoplasm and nucleus of TSCs. However, in contrast to this study as well as Nishioka *et al.* and Hirate *et al.* studies we could detect Tead4 in both the nucleus and cytoplasm of intact ESCs. Strikingly, we found that RS increases the level of this protein both in the cytoplasm and nucleus of ESCs (**Figure 17a**). As Tead4 is already present in the nucleus of untreated ESCs and since, RS increases both the nuclear and cytoplasmic level of this protein, we could not conclude that RS alters the subcellular localization of Tead4 from cytoplasm to the nucleus in ESCs. Nevertheless, our immunostaining result shows that the Tead4 expression in RS-induced ESCs has foci pattern whereas, in TSCs Tead4 has diffused pattern throughout the nucleus (**Figure 17a**). The different patterns of Tead4 expression in RS-induced ESCs and TSCs requires further

investigation. Since increased level of nuclear Tead4 in response to RS is not likely to be related to its subcellular localization, we wondered whether it is transcriptionally regulated. Upon APH treatment, we showed that the mRNA and accordingly the protein level of Tead4 remarkably increases in a dose-dependent manner (**Figure 17b-d**) similar to the increased expression of canonical TE TFs such as Eomes, Elf5 and others in ESCs (**Figure 6a and 16**). Next, we checked whether the expression of Tead4 is ATR-Chk1-dependent similar to the expression of other TE markers. Although the expression of Tead4 increases in APH-treated ESCs, neither inhibition of ATR nor ATM could revert its level to the basal level (**Figure 18**). This finding suggests that in response to RS regardless of active or inhibited DDR signaling pathways (ATM-Chk2 and ATR-Chk1), ESCs increase the expression of Tead4 to keep themselves ready for the activation of TE gene network and subsequently differentiate toward TSCs. As Tead4 binds to the TFBMs of TE genes (87, 90, 142, 143) and we found that the expression of TE genes in RS-induced ESCs is ATR-Chk1-dependent, we wondered whether the binding of Tead4 to the TFBMs of TE genes is mediated through the activation of ATR-Chk1 pathway. Strikingly, through ChIP-qPCR analysis we showed that the occupancy of Tead4 at the TFBMs of TE genes in RS-induced ESCs significantly increases. Interestingly, this binding is dependent on the activation of ATR pathway (**Figure 19**). Therefore, under RS conditions, Tead4 is highly expressed in ESCs to readily bind to the TE genes through ATR-Chk1 pathway to induce their differentiation toward TSCs. We think that similar to the *in vitro* condition, the ATR-Chk1-Tead4-mediated differentiation of inner cells to the TE lineages could take place in RS-induced embryos to minimize the contribution of defective cells to the embryonic proper and later to the progeny. It seems that Tead4 level in the nucleus of ESCs has to reach to a certain level to be able to bind to the TE genes and activates their expression as, in contrast to TSCs and RS-induced ESCs, the low level of Tead4 in untreated ESCs could not trigger the expression of TE genes. How does RS increase the expression of Tead4, and how the activation of the ATR-Chk1 pathway increases the binding of Tead4 to the TFBMs of TE genes in ESCs remains to be elucidated. Studies have identified molecules which acts as co-activator of Tead4 in Hippo signaling pathway. Yap1 is one of these mediators, upon dephosphorylation it is translocated to the nucleus and activates Tead4. TAZ, the paralog of Yap1, also is one of the co-factors which works similar to and in parallel to Yap1 to activate Tead4 (144). Further investigation is required to identify the mediators of the Hippo signaling pathway that could play a role downstream to the ATR-Chk1 pathway in increasing the binding of Tead4 to the TE genes in ESCs. It is also suggested to study whether ATR activation plays a role in the activation of other signaling pathways that regulate the cell fate specification during embryonic development. There are pathways that work in parallel to the Hippo signaling pathway and

activate the TE gene network in the embryos. For instance, the Notch signaling pathway is reported to play a role in the activation of Cdx2 through the binding of its TF, RBPJ to the enhancer of this key TE gene (142).

Kidder and Palmer study showed that Eomes transcriptionally regulates its gene through binding to its TFBM (135). With respect to this, we interestingly found that Eomes significantly enriches to its TFBM through ATR-Chk1-mediated RSR (**Figure 20**). This finding suggests that beside ATR-Chk1-mediated enrichment of Tead4 to the TFBMs of Eomes and canonical TE genes, the enhanced occupancy of Eomes to its promoter further activates its expression in RS-induced ESCs. Hence, this mechanism further prepares unrepaired ESCs to differentiate toward TSCs to probably more efficiently remove such cells from the embryonic pool.

We have recently reported for the first time that activation of ATR-Chk1 pathway stimulates the transition of ESCs to 2C-like cells (131). These cells, which have the functional and transcriptional characteristics of totipotent 2C-stage embryos, emerge transiently in the ESC culture (145-147). This transition is reported to be required in ESC culture to enhance their genome integrity (148). Moreover, several others and we have shown that 2C-like cells have expanded developmental potential and can contribute to both the embryonic and extra-embryonic tissues (131, 145-147). In depth, we reported that the activation of ATR-Chk1 pathway increases the expression of *Dux* gene through its post-transcriptional regulation (131). This gene is the master regulator of ZGA in placental mammals (149, 150). We think that in response to RS, ESCs take advantage of ATR-Chk1-mediated increase of 2C-like cells as a safeguard mechanism to not only enhance the genome integrity of the developing embryo but also to restrict the contribution of unrepaired cells to the embryonic proper through diverting them toward the extra-embryonic tissues (131). The activation of ATR-Chk1 kinases increases the expression of 2C-like genes and TE genes in ESCs through two separate parallel pathways. Since, neither KD of *Tead4* affected the increased expression of Dux nor KD of *Dux* altered the level of highly expressed TE genes in APH-treated ESCs (data not shown). It is likely that embryos take advantage of robust parallel pathways to restrict the contribution of defective ESCs to the embryonic proper. However, the physiological significance of our findings needs to be further explored.

In summary, our study introduced an innovative *in vitro* combinatorial 3D embryogenesis approach through which the impact of genotoxic stress as well as the fate of various ESCs with deficiency in DDR pathways during early embryogenesis could be studied. Moreover,

we unraveled a molecular mechanism which ESCs exploit to respond efficiently to the RS during early embryonic development (**Figure 21**)

References

1. Mazouzi A, Velimezi G, Loizou JI. DNA replication stress: causes, resolution and disease. Exp Cell Res. 2014;329(1):85-93.
2. Bell SP, Dutta A. DNA replication in eukaryotic cells. Annu Rev Biochem. 2002;71:333-74.
3. Murray H, Koh A. Multiple regulatory systems coordinate DNA replication with cell growth in Bacillus subtilis. PLoS Genet. 2014;10(10):e1004731.
4. Hartwell LH, Weinert TA. Checkpoints: controls that ensure the order of cell cycle events. Science. 1989;246(4930):629-34.
5. Gonzalez MA, Tachibana KE, Laskey RA, Coleman N. Control of DNA replication and its potential clinical exploitation. Nat Rev Cancer. 2005;5(2):135-41.
6. Rhind N, Russell P. Signaling pathways that regulate cell division. Cold Spring Harb Perspect Biol. 2012;4(10).
7. Pardee AB. G1 events and regulation of cell proliferation. Science. 1989;246(4930):603-8.
8. Kousholt AN, Menzel T, Sorensen CS. Pathways for genome integrity in G2 phase of the cell cycle. Biomolecules. 2012;2(4):579-607.
9. Foijer F, Te Riele H. Restriction beyond the restriction point: mitogen requirement for G2 passage. Cell Div. 2006;1:8.
10. Norris V. Does the Semiconservative Nature of DNA Replication Facilitate Coherent Phenotypic Diversity? J Bacteriol. 2019;201(12).
11. Primo LMF, Teixeira LK. DNA replication stress: oncogenes in the spotlight. Genet Mol Biol. 2019;43(1 suppl 1):e20190138.
12. Ritzi M, Knippers R. Initiation of genome replication: assembly and disassembly of replication-competent chromatin. Gene. 2000;245(1):13-20.
13. Chong JP, Blow JJ. DNA replication licensing factor. Prog Cell Cycle Res. 1996;2:83-90.
14. Munoz S, Mendez J. DNA replication stress: from molecular mechanisms to human disease. Chromosoma. 2017;126(1):1-15.
15. Nishitani H, Morino M, Murakami Y, Maeda T, Shiomi Y. Chromatin fractionation analysis of licensing factors in mammalian cells. Methods Mol Biol. 2014;1170:517-27.
16. Sclafani RA, Holzen TM. Cell cycle regulation of DNA replication. Annu Rev Genet. 2007;41:237-80.
17. Yuan Z, Riera A, Bai L, Sun J, Nandi S, Spanos C, et al. Structural basis of Mcm2-7 replicative helicase loading by ORC-Cdc6 and Cdt1. Nat Struct Mol Biol. 2017;24(3):316-24.
18. Masai H, Arai K. Regulation of DNA replication during the cell cycle: roles of Cdc7 kinase and coupling of replication, recombination, and repair in response to replication fork arrest. IUBMB Life. 2000;49(5):353-64.
19. Stoeber K, Tlsty TD, Happerfield L, Thomas GA, Romanov S, Bobrow L, et al. DNA replication licensing and human cell proliferation. J Cell Sci. 2001;114(Pt 11):2027-41.
20. Gambus A. Termination of Eukaryotic Replication Forks. Adv Exp Med Biol. 2017;1042:163-87.
21. Ovejero S, Bueno A, Sacristan MP. Working on Genomic Stability: From the S-Phase to Mitosis. Genes (Basel). 2020;11(2).
22. Blow JJ, Gillespie PJ. Replication licensing and cancer--a fatal entanglement? Nat Rev Cancer. 2008;8(10):799-806.
23. Lemmens B, Hegarat N, Akopyan K, Sala-Gaston J, Bartek J, Hochegger H, et al. DNA Replication Determines Timing of Mitosis by Restricting CDK1 and PLK1 Activation. Mol Cell. 2018;71(1):117-28 e3.

24. Zegerman P. Evolutionary conservation of the CDK targets in eukaryotic DNA replication initiation. Chromosoma. 2015;124(3):309-21.
25. Morgan DO. Cyclin-dependent kinases: engines, clocks, and microprocessors. Annu Rev Cell Dev Biol. 1997;13:261-91.
26. Kelly TJ, Brown GW. Regulation of chromosome replication. Annu Rev Biochem. 2000;69:829-80.
27. Takeda DY, Dutta A. DNA replication and progression through S phase. Oncogene. 2005;24(17):2827-43.
28. Cheng L, Collyer T, Hardy CF. Cell cycle regulation of DNA replication initiator factor Dbf4p. Mol Cell Biol. 1999;19(6):4270-8.
29. Cook JG, Chasse DA, Nevins JR. The regulated association of Cdt1 with minichromosome maintenance proteins and Cdc6 in mammalian cells. J Biol Chem. 2004;279(10):9625-33.
30. Kastan MB, Bartek J. Cell-cycle checkpoints and cancer. Nature. 2004;432(7015):316-23.
31. Abraham RT. Cell cycle checkpoint signaling through the ATM and ATR kinases. Genes Dev. 2001;15(17):2177-96.
32. Bartek J, Lukas J. Chk1 and Chk2 kinases in checkpoint control and cancer. Cancer Cell. 2003;3(5):421-9.
33. Donzelli M, Draetta GF. Regulating mammalian checkpoints through Cdc25 inactivation. EMBO Rep. 2003;4(7):671-7.
34. Lukas J, Lukas C, Bartek J. Mammalian cell cycle checkpoints: signalling pathways and their organization in space and time. DNA Repair (Amst). 2004;3(8-9):997-1007.
35. Petrini JH, Stracker TH. The cellular response to DNA double-strand breaks: defining the sensors and mediators. Trends Cell Biol. 2003;13(9):458-62.
36. Mailand N, Falck J, Lukas C, Syljuasen RG, Welcker M, Bartek J, et al. Rapid destruction of human Cdc25A in response to DNA damage. Science. 2000;288(5470):1425-9.
37. Falck J, Petrini JH, Williams BR, Lukas J, Bartek J. The DNA damage-dependent intra-S phase checkpoint is regulated by parallel pathways. Nat Genet. 2002;30(3):290-4.
38. Sherr CJ, Roberts JM. CDK inhibitors: positive and negative regulators of G1-phase progression. Genes Dev. 1999;13(12):1501-12.
39. Bartek J, Lukas J. Mammalian G1- and S-phase checkpoints in response to DNA damage. Curr Opin Cell Biol. 2001;13(6):738-47.
40. Reuter S, Gupta SC, Chaturvedi MM, Aggarwal BB. Oxidative stress, inflammation, and cancer: how are they linked? Free Radic Biol Med. 2010;49(11):1603-16.
41. Chatterjee N, Walker GC. Mechanisms of DNA damage, repair, and mutagenesis. Environ Mol Mutagen. 2017;58(5):235-63.
42. Kunkel TA. Evolving views of DNA replication (in)fidelity. Cold Spring Harb Symp Quant Biol. 2009;74:91-101.
43. Manhart CM, Alani E. DNA replication and mismatch repair safeguard against metabolic imbalances. Proc Natl Acad Sci U S A. 2017;114(22):5561-3.
44. Viguera E, Canceill D, Ehrlich SD. Replication slippage involves DNA polymerase pausing and dissociation. EMBO J. 2001;20(10):2587-95.
45. Pourquier P, Pommier Y. Topoisomerase I-mediated DNA damage. Adv Cancer Res. 2001;80:189-216.
46. Yang Y, Bazhin AV, Werner J, Karakhanova S. Reactive oxygen species in the immune system. Int Rev Immunol. 2013;32(3):249-70.
47. Imlay JA, Chin SM, Linn S. Toxic DNA damage by hydrogen peroxide through the Fenton reaction in vivo and in vitro. Science. 1988;240(4852):640-2.
48. Henner WD, Grunberg SM, Haseltine WA. Enzyme action at 3' termini of ionizing radiation-induced DNA strand breaks. J Biol Chem. 1983;258(24):15198-205.

49. Desouky O, Ding N, Zhou G. Targeted and non-targeted effects of ionizing radiation. Journal of Radiation Research and Applied Sciences. 2015;8(2):247-54.
50. Vignard J, Mirey G, Salles B. Ionizing-radiation induced DNA double-strand breaks: a direct and indirect lighting up. Radiother Oncol. 2013;108(3):362-9.
51. Iliakis G. The role of DNA double strand breaks in ionizing radiation-induced killing of eukaryotic cells. Bioessays. 1991;13(12):641-8.
52. Rastogi RP, Richa, Kumar A, Tyagi MB, Sinha RP. Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair. J Nucleic Acids. 2010;2010:592980.
53. Vesela E, Chroma K, Turi Z, Mistrik M. Common Chemical Inductors of Replication Stress: Focus on Cell-Based Studies. Biomolecules. 2017;7(1).
54. Nazaretyan SA, Savic N, Sadek M, Hackert BJ, Courcelle J, Courcelle CT. Replication Rapidly Recovers and Continues in the Presence of Hydroxyurea in Escherichia coli. J Bacteriol. 2018;200(6).
55. Zeman MK, Cimprich KA. Causes and consequences of replication stress. Nat Cell Biol. 2014;16(1):2-9.
56. Debatisse M, Le Tallec B, Letessier A, Dutrillaux B, Brison O. Common fragile sites: mechanisms of instability revisited. Trends Genet. 2012;28(1):22-32.
57. Ozeri-Galai E, Lebofsky R, Rahat A, Bester AC, Bensimon A, Kerem B. Failure of origin activation in response to fork stalling leads to chromosomal instability at fragile sites. Mol Cell. 2011;43(1):122-31.
58. Poli J, Tsaponina O, Crabbe L, Keszthelyi A, Pantesco V, Chabes A, et al. dNTP pools determine fork progression and origin usage under replication stress. EMBO J. 2012;31(4):883-94.
59. Magdalou I, Lopez BS, Pasero P, Lambert SA. The causes of replication stress and their consequences on genome stability and cell fate. Semin Cell Dev Biol. 2014;30:154-64.
60. Lambert S, Carr AM. Impediments to replication fork movement: stabilisation, reactivation and genome instability. Chromosoma. 2013;122(1-2):33-45.
61. Wei X, Samarabandu J, Devdhar RS, Siegel AJ, Acharya R, Berezney R. Segregation of transcription and replication sites into higher order domains. Science. 1998;281(5382):1502-6.
62. Aguilera A, Garcia-Muse T. R loops: from transcription byproducts to threats to genome stability. Mol Cell. 2012;46(2):115-24.
63. Tuduri S, Crabbe L, Conti C, Tourriere H, Holtgreve-Grez H, Jauch A, et al. Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. Nat Cell Biol. 2009;11(11):1315-24.
64. Helmrich A, Ballarino M, Tora L. Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. Mol Cell. 2011;44(6):966-77.
65. Clemente-Ruiz M, Gonzalez-Prieto R, Prado F. Histone H3K56 acetylation, CAF1, and Rtt106 coordinate nucleosome assembly and stability of advancing replication forks. PLoS Genet. 2011;7(11):e1002376.
66. Mejlvang J, Feng Y, Alabert C, Neelsen KJ, Jasencakova Z, Zhao X, et al. New histone supply regulates replication fork speed and PCNA unloading. J Cell Biol. 2014;204(1):29-43.
67. Yi C, He C. DNA repair by reversal of DNA damage. Cold Spring Harb Perspect Biol. 2013;5(1):a012575.
68. Krokan HE, Bjoras M. Base excision repair. Cold Spring Harb Perspect Biol. 2013;5(4):a012583.
69. Spivak G. Nucleotide excision repair in humans. DNA Repair (Amst). 2015;36:13-8.
70. Hsieh P, Yamane K. DNA mismatch repair: molecular mechanism, cancer, and ageing. Mech Ageing Dev. 2008;129(7-8):391-407.

71. Goodman MF, Woodgate R. Translesion DNA polymerases. Cold Spring Harb Perspect Biol. 2013;5(10):a010363.
72. Caldecott KW. Single-strand break repair and genetic disease. Nat Rev Genet. 2008;9(8):619-31.
73. Rodgers K, McVey M. Error-Prone Repair of DNA Double-Strand Breaks. J Cell Physiol. 2016;231(1):15-24.
74. Jackson SP, Bartek J. The DNA-damage response in human biology and disease. Nature. 2009;461(7267):1071-8.
75. Marechal A, Zou L. DNA damage sensing by the ATM and ATR kinases. Cold Spring Harb Perspect Biol. 2013;5(9).
76. Bass TE, Luzwick JW, Kavanaugh G, Carroll C, Dungrawala H, Glick GG, et al. ETAA1 acts at stalled replication forks to maintain genome integrity. Nat Cell Biol. 2016;18(11):1185-95.
77. Berti M, Vindigni A. Replication stress: getting back on track. Nat Struct Mol Biol. 2016;23(2):103-9.
78. Caldon CE. Estrogen signaling and the DNA damage response in hormone dependent breast cancers. Front Oncol. 2014;4:106.
79. Sulli G, Di Micco R, d'Adda di Fagagna F. Crosstalk between chromatin state and DNA damage response in cellular senescence and cancer. Nat Rev Cancer. 2012;12(10):709-20.
80. Lavin MF, Shiloh Y. The genetic defect in ataxia-telangiectasia. Annu Rev Immunol. 1997;15:177-202.
81. Murga M, Bunting S, Montana MF, Soria R, Mulero F, Canamero M, et al. A mouse model of ATR-Seckel shows embryonic replicative stress and accelerated aging. Nat Genet. 2009;41(8):891-8.
82. O'Driscoll M. Mouse models for ATR deficiency. DNA Repair (Amst). 2009;8(11):1333-7.
83. Xu Y, Ashley T, Brainerd EE, Bronson RT, Meyn MS, Baltimore D. Targeted disruption of ATM leads to growth retardation, chromosomal fragmentation during meiosis, immune defects, and thymic lymphoma. Genes Dev. 1996;10(19):2411-22.
84. Cockburn K, Rossant J. Making the blastocyst: lessons from the mouse. J Clin Invest. 2010;120(4):995-1003.
85. Rossant J. Making the Mouse Blastocyst: Past, Present, and Future. Curr Top Dev Biol. 2016;117:275-88.
86. Chazaud C, Yamanaka Y. Lineage specification in the mouse preimplantation embryo. Development. 2016;143(7):1063-74.
87. Sasaki H. Mechanisms of trophectoderm fate specification in preimplantation mouse development. Dev Growth Differ. 2010;52(3):263-73.
88. Sauvegarde C, Paul D, Bridoux L, Jouneau A, Degrelle S, Hue I, et al. Dynamic Pattern of HOXB9 Protein Localization during Oocyte Maturation and Early Embryonic Development in Mammals. PLoS One. 2016;11(10):e0165898.
89. Mo JS, Park HW, Guan KL. The Hippo signaling pathway in stem cell biology and cancer. EMBO Rep. 2014;15(6):642-56.
90. Home P, Saha B, Ray S, Dutta D, Gunewardena S, Yoo B, et al. Altered subcellular localization of transcription factor TEAD4 regulates first mammalian cell lineage commitment. Proc Natl Acad Sci U S A. 2012;109(19):7362-7.
91. White MD, Plachta N. How adhesion forms the early mammalian embryo. Curr Top Dev Biol. 2015;112:1-17.
92. Woods L, Perez-Garcia V, Hemberger M. Regulation of Placental Development and Its Impact on Fetal Growth-New Insights From Mouse Models. Front Endocrinol (Lausanne). 2018;9:570.
93. Artus J, Cohen-Tannoudji M. Cell cycle regulation during early mouse embryogenesis. Mol Cell Endocrinol. 2008;282(1-2):78-86.

94. Houlard M, Artus J, Cohen-Tannoudji M. Cell cycle checkpoints and DNA damage response in early mouse embryo. 2009. p. 223-44.
95. Jaroudi S, SenGupta S. DNA repair in mammalian embryos. Mutat Res. 2007;635(1):53-77.
96. Yang Y, Bolnick A, Shamir A, Abdulhasan M, Li Q, Parker GC, et al. Blastocyst-Derived Stem Cell Populations under Stress: Impact of Nutrition and Metabolism on Stem Cell Potency Loss and Miscarriage. Stem Cell Rev Rep. 2017;13(4):454-64.
97. Blair K, Wray J, Smith A. The liberation of embryonic stem cells. PLoS Genet. 2011;7(4):e1002019.
98. Niwa H. Mouse ES cell culture system as a model of development. Dev Growth Differ. 2010;52(3):275-83.
99. Shi Y, Inoue H, Wu JC, Yamanaka S. Induced pluripotent stem cell technology: a decade of progress. Nat Rev Drug Discov. 2017;16(2):115-30.
100. Ahuja AK, Jodkowska K, Teloni F, Bizard AH, Zellweger R, Herrador R, et al. A short G1 phase imposes constitutive replication stress and fork remodelling in mouse embryonic stem cells. Nat Commun. 2016;7:10660.
101. Giachino C, Orlando L, Turinetto V. Maintenance of genomic stability in mouse embryonic stem cells: relevance in aging and disease. Int J Mol Sci. 2013;14(2):2617-36.
102. Mascetti VL, Pedersen RA. Contributions of Mammalian Chimeras to Pluripotent Stem Cell Research. Cell Stem Cell. 2016;19(2):163-75.
103. Tam PP, Rossant J. Mouse embryonic chimeras: tools for studying mammalian development. Development. 2003;130(25):6155-63.
104. Keller G. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. Genes Dev. 2005;19(10):1129-55.
105. Cambuli F, Murray A, Dean W, Dudzinska D, Krueger F, Andrews S, et al. Epigenetic memory of the first cell fate decision prevents complete ES cell reprogramming into trophoblast. Nat Commun. 2014;5:5538.
106. Tanaka S, Kunath T, Hadjantonakis AK, Nagy A, Rossant J. Promotion of trophoblast stem cell proliferation by FGF4. Science. 1998;282(5396):2072-5.
107. Niwa H, Toyooka Y, Shimosato D, Strumpf D, Takahashi K, Yagi R, et al. Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. Cell. 2005;123(5):917-29.
108. Ralston A, Cox BJ, Nishioka N, Sasaki H, Chea E, Rugg-Gunn P, et al. Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. Development. 2010;137(3):395-403.
109. Nishioka N, Inoue K, Adachi K, Kiyonari H, Ota M, Ralston A, et al. The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. Dev Cell. 2009;16(3):398-410.
110. Ng RK, Dean W, Dawson C, Lucifero D, Madeja Z, Reik W, et al. Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. Nat Cell Biol. 2008;10(11):1280-90.
111. Lu CW, Yabuuchi A, Chen L, Viswanathan S, Kim K, Daley GQ. Ras-MAPK signaling promotes trophectoderm formation from embryonic stem cells and mouse embryos. Nat Genet. 2008;40(7):921-6.
112. Kuckenberg P, Buhl S, Woynecki T, van Furden B, Tolkunova E, Seiffe F, et al. The transcription factor TCFAP2C/AP-2gamma cooperates with CDX2 to maintain trophectoderm formation. Mol Cell Biol. 2010;30(13):3310-20.
113. Abad M, Mosteiro L, Pantoja C, Canamero M, Rayon T, Ors I, et al. Reprogramming in vivo produces teratomas and iPS cells with totipotency features. Nature. 2013;502(7471):340-5.
114. Hirate Y, Cockburn K, Rossant J, Sasaki H. Tead4 is constitutively nuclear, while nuclear vs. cytoplasmic Yap distribution is regulated in preimplantation mouse embryos. Proc Natl Acad Sci U S A. 2012;109(50):E3389-90; author reply E91-2.

115. Tolkunova E, Cavaleri F, Eckardt S, Reinbold R, Christenson LK, Scholer HR, et al. The caudal-related protein cdx2 promotes trophoblast differentiation of mouse embryonic stem cells. Stem Cells. 2006;24(1):139-44.
116. Hemberger M, Nozaki T, Winterhager E, Yamamoto H, Nakagama H, Kamada N, et al. Parp1-deficiency induces differentiation of ES cells into trophoblast derivatives. Dev Biol. 2003;257(2):371-81.
117. Siggia ED, Warmflash A. Modeling Mammalian Gastrulation With Embryonic Stem Cells. Curr Top Dev Biol. 2018;129:1-23.
118. Harrison SE, Sozen B, Christodoulou N, Kyprianou C, Zernicka-Goetz M. Assembly of embryonic and extraembryonic stem cells to mimic embryogenesis in vitro. Science. 2017;356(6334).
119. Shahbazi MN, Siggia ED, Zernicka-Goetz M. Self-organization of stem cells into embryos: A window on early mammalian development. Science. 2019;364(6444):948-51.
120. Metzger JJ, Simunovic M, Brivanlou AH. Synthetic embryology: controlling geometry to model early mammalian development. Curr Opin Genet Dev. 2018;52:86-91.
121. Rivron NC, Frias-Aldeguer J, Vrij EJ, Boisset JC, Korving J, Vivie J, et al. Blastocyst-like structures generated solely from stem cells. Nature. 2018;557(7703):106-11.
122. Harrison SE, Sozen B, Zernicka-Goetz M. In vitro generation of mouse polarized embryo-like structures from embryonic and trophoblast stem cells. Nat Protoc. 2018;13(7):1586-602.
123. Yang J, Ryan DJ, Lan G, Zou X, Liu P. In vitro establishment of expanded-potential stem cells from mouse pre-implantation embryos or embryonic stem cells. Nat Protoc. 2019;14(2):350-78.
124. Yang J, Ryan DJ, Wang W, Tsang JC, Lan G, Masaki H, et al. Establishment of mouse expanded potential stem cells. Nature. 2017;550(7676):393-7.
125. Yang Y, Liu B, Xu J, Wang J, Wu J, Shi C, et al. Derivation of Pluripotent Stem Cells with In Vivo Embryonic and Extraembryonic Potency. Cell. 2017;169(2):243-57 e25.
126. Posfai E, Schell JP, Janiszewski A, Rovic I, Murray A, Bradshaw B, et al. Defining totipotency using criteria of increasing stringency. bioRxiv. 2020:2020.03.02.972893.
127. Li R, Zhong C, Yu Y, Liu H, Sakurai M, Yu L, et al. Generation of Blastocyst-like Structures from Mouse Embryonic and Adult Cell Cultures. Cell. 2019;179(3):687-702 e18.
128. Sozen B, Cox AL, De Jonghe J, Bao M, Hollfelder F, Glover DM, et al. Self-Organization of Mouse Stem Cells into an Extended Potential Blastoid. Dev Cell. 2019;51(6):698-712 e8.
129. Xia Y, Izpisua Belmonte JC. Design Approaches for Generating Organ Constructs. Cell Stem Cell. 2019;24(6):877-94.
130. Lee GY, Kenny PA, Lee EH, Bissell MJ. Three-dimensional culture models of normal and malignant breast epithelial cells. Nat Methods. 2007;4(4):359-65.
131. Atashpaz S, Samadi Shams S, Gonzalez JM, Sebestyen E, Arghavanifard N, Gnocchi A, et al. ATR expands embryonic stem cell fate potential in response to replication stress. Elife. 2020;9.
132. Haffner-Krausz R, Gorivodsky M, Chen Y, Lonai P. Expression of Fgfr2 in the early mouse embryo indicates its involvement in preimplantation development. Mech Dev. 1999;85(1-2):167-72.
133. Latos PA, Hemberger M. From the stem of the placental tree: trophoblast stem cells and their progeny. Development. 2016;143(20):3650-60.
134. Abe T, Kiyonari H, Shioi G, Inoue K, Nakao K, Aizawa S, et al. Establishment of conditional reporter mouse lines at ROSA26 locus for live cell imaging. Genesis. 2011;49(7):579-90.
135. Kidder BL, Palmer S. Examination of transcriptional networks reveals an important role for TCFAP2C, SMARCA4, and EOMES in trophoblast stem cell maintenance. Genome Res. 2010;20(4):458-72.

136. Tichy ED. Mechanisms maintaining genomic integrity in embryonic stem cells and induced pluripotent stem cells. Exp Biol Med (Maywood). 2011;236(9):987-96.
137. Santos MA, Faryabi RB, Ergen AV, Day AM, Malhowski A, Canela A, et al. DNA-damage-induced differentiation of leukaemic cells as an anti-cancer barrier. Nature. 2014;514(7520):107-11.
138. Schneider L, Pellegatta S, Favaro R, Pisati F, Roncaglia P, Testa G, et al. DNA damage in mammalian neural stem cells leads to astrocytic differentiation mediated by BMP2 signaling through JAK-STAT. Stem Cell Reports. 2013;1(2):123-38.
139. Liu Q, Guntuku S, Cui XS, Matsuoka S, Cortez D, Tamai K, et al. Chk1 is an essential kinase that is regulated by Atr and required for the G(2)/M DNA damage checkpoint. Genes Dev. 2000;14(12):1448-59.
140. Joo JY, Choi HW, Kim MJ, Zaehres H, Tapia N, Stehling M, et al. Establishment of a primed pluripotent epiblast stem cell in FGF4-based conditions. Sci Rep. 2014;4:7477.
141. Lin KC, Park HW, Guan KL. Regulation of the Hippo Pathway Transcription Factor TEAD. Trends Biochem Sci. 2017;42(11):862-72.
142. Rayon T, Menchero S, Nieto A, Xenopoulos P, Crespo M, Cockburn K, et al. Notch and hippo converge on Cdx2 to specify the trophectoderm lineage in the mouse blastocyst. Dev Cell. 2014;30(4):410-22.
143. Tomikawa J, Takada S, Okamura K, Terao M, Ogata-Kawata H, Akutsu H, et al. Exploring trophoblast-specific Tead4 enhancers through chromatin conformation capture assays followed by functional screening. Nucleic Acids Res. 2020;48(1):278-89.
144. Varelas X. The Hippo pathway effectors TAZ and YAP in development, homeostasis and disease. Development. 2014;141(8):1614-26.
145. Choi YJ, Lin CP, Risso D, Chen S, Kim TA, Tan MH, et al. Deficiency of microRNA miR-34a expands cell fate potential in pluripotent stem cells. Science. 2017;355(6325).
146. Ishiuchi T, Enriquez-Gasca R, Mizutani E, Boskovic A, Ziegler-Birling C, Rodriguez-Terrones D, et al. Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. Nat Struct Mol Biol. 2015;22(9):662-71.
147. Macfarlan TS, Gifford WD, Driscoll S, Lettieri K, Rowe HM, Bonanomi D, et al. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature. 2012;487(7405):57-63.
148. Zalzman M, Falco G, Sharova LV, Nishiyama A, Thomas M, Lee SL, et al. Zscan4 regulates telomere elongation and genomic stability in ES cells. Nature. 2010;464(7290):858-63.
149. De Iaco A, Planet E, Coluccio A, Verp S, Duc J, Trono D. DUX-family transcription factors regulate zygotic genome activation in placental mammals. Nat Genet. 2017;49(6):941-5.
150. Hendrickson PG, Dorais JA, Grow EJ, Whiddon JL, Lim JW, Wike CL, et al. Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. Nat Genet. 2017;49(6):925-34.