



HAL
open science

Peri-implantation stem cell models characterization : a step toward the establishment of standards

Constance Onfray

► **To cite this version:**

Constance Onfray. Peri-implantation stem cell models characterization : a step toward the establishment of standards. Cellular Biology. Nantes Université, 2023. English. NNT : 2023NANU1041 . tel-04632477

HAL Id: tel-04632477

<https://theses.hal.science/tel-04632477>

Submitted on 2 Jul 2024

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

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE DE DOCTORAT

NANTES UNIVERSITE

ECOLE DOCTORALE N° 605

Biologie-Santé

Spécialité : Biologie cellulaire, Biologie du développement

Par

Constance ONFRAY

Peri-implantation stem cell models characterization: a step toward the establishment of standards.

Thèse présentée et soutenue à Nantes, le 15 décembre 2023

Unité de recherche : CR2TI, UMR 1064

Rapporteurs avant soutenance :

Norah Fogarty Group Leader, King's college London

Edouard Lecarpentier Professeur des Université – Praticien hospitalier, Hôpital intercommunal de Créteil

Composition du Jury :

Président : Claire Rougeulle Directrice de recherche, Université Paris cité

Examineurs : Claire Rougeulle Directrice de recherches, Université Paris cité

Dir. de thèse : Laurent David Maitre de conférences des Université – Praticien hospitalier, Nantes Université

Invité(s)

Anne Camus Chargée de recherches, Nantes université

“Just keep swimming.”

—

Dory, *Finding Nemo*

Table of content

AKNOWLEDGEMENTS	6
FRENCH THESIS SUMMARY	11
Introduction	11
Résultats	14
Discussion.....	15
Perspectives	17
LIST OF FIGURES	18
TABLE OF ABBREVIATIONS	19
INTRODUCTION	21
How can we study early human development?	23
Human embryonic development: a quick overview of the beginning of life	23
Regulations of human embryo culture: International perspectives and France focus	24
Exploring Pre- and Post- implantation development: animal models	27
Developmental timing differs in mammals	27
Exploring implantation across mammalian species.....	28
Stem cell models of human embryo development	31
What is a stem cell model?	31
The different kinds of stem cells	33
Pluripotent stem cells	36
Naive pluripotent stem cells	37
Primed pluripotent stem cells.....	39
Formative pluripotency.....	41
Extended pluripotency	42
How are pluripotent stem cells generated?	43
Trophoblast stem cells	48
Trophoblast in vivo	48
Models of trophoblast.....	50
TSCs generation methods.....	53
Modeling using human stem cells models	57
Modeling using hPPSC	57

Modelling placental development using hTSCs and using them to understand principal placental illnesses	58
3D models of human peri-implantation embryo	60
Ethics and regulation linked to embryonic stem cell models	63
Guidelines ISSCR.....	63
Consent	64
Regulation (Fr)	64
RESULTS	65
Context of my work.....	65
Scientific articles submitted or published during this PhD.....	67
Induction of Human Naïve pluripotent stem cells from somatic cells	67
Deciphering hallmarks ability to stage match peri-implantation stem cell models.....	68
Naive Pluripotent and Trophoblastic Stem Cell Lines as a Model for Detecting Missing Proteins in the Context of the Chromosome-Centric Human Proteome Project.....	69
Embryonic Stem cell derivation	70
Primed Derivation	70
Naïve derivation	70
Trophoblast derivation	71
Conclusion of the derivation experiment.....	72
DISCUSSION	73
Proteomic and transcriptomic data mis alignment.....	74
Culture medium components impact on DNA methylation, X chromosome and metabolic activity.....	76
TGF b inhibition.....	76
HDAC inhibition.....	76
MEK inhibition	77
Wnt inhibition	77
Rock inhibition	78
Vitamin C.....	78
EGF.....	78
LIF	79
From the model to the embryo, are stem cell models well depicting embryonic development?	81
Reprogrammed, converted, derived: all the same results.....	84

Evaluating pluripotency state 86

PERSPECTIVES 89

 What is the link between fate and metabolism? 90

 Can embryo metabolism serve as a readout to determine its implantation potential? 91

 Does culture media have an impact on developmental and implantation potential? 93

PERSONAL OUTLOOK OF THE PHD..... 94

BIBLIOGRAPHY..... 95

ACKNOWLEDGEMENTS

I started my PhD just after the onset of the Covid-19 pandemic, a period that presented numerous challenges and uncertainties. Nonetheless, I made it to the finish line! Pursuing a PhD was one of my dreams and I would like to express my deepest gratitude to the exceptional individuals who supported and guided me throughout this incredible journey, despite the many unforeseen obstacles that arose.

First of all, I am very grateful to Dr. Norah Fogarty and Dr. Edouard Lecarpentier, the 'rapporteurs' of my thesis for their time and scientific input.

I am also grateful to Dr. Claire Rougeulle who kindly accepted to be part of my thesis jury. I would also like to thank Claire for kindly welcoming me in her lab in Paris, for me to watch RNA FISH procedure and to cultivate hTSCs.

I would like to thank Dr. Anne Camus and Dr. Peter Rugg Gunn for their support during my PhD follow-up committee. Your scientific and personal input helped me to go forward and to improve my project.

I would like to thank from the bottom of my heart my PhD supervisor, Dr. Laurent David. Your constant support, guidance, and encouragement has been invaluable throughout this entire process. From the initial stages of learning to culture stem cells to the final submission of my thesis, your presence and wisdom has been instrumental in shaping my academic growth. I highly valued the informal meetings we held, which not only served as crucial checkpoints to keep me on track academically, but also provided me with plenty of encouragement and helped me to better understand the scientific field, from the science to the individuals. I am profoundly grateful for the immeasurable contributions you made to my development. Thank you very much also for having informal meeting at your place with both our team and other European teams. It's not every day that you can play the unicorn 7 family game with European leaders.

I would like to thank all CR2Ti members but especially team 2 members, for their overall scientific, technical and moral support. Florian, Vanessa, Maïssa, we may not be in the same building but thank you for all our talks.

I would like to thank the IVF clinic: Thomas Fréour and Arnaud Reignier for welcoming me to show me the clinic, Jenna Lammers, Sophie Loubersac, especially for embryonic stem cell derivation but also for the discussions about clinical research and the scientific field.

I would like to thank Dr. Nicolas Rivron, for kindly welcoming me in Vienna to be trained to make blastoids.

I would like to thank Micropicell for letting me access their microscopes at any time (day and night, week or weekend) and for their useful advices.

I am indebted to my exceptional lab mates, whose support has been a constant source of motivation. Our informal chats, the sharing of scientific experience and other useful tips, provided a lifeline during the most challenging times. I am proud to say that we became more than just lab partners, but good friends.

Gael, merci de m'avoir formée à la culture et à la différenciation. Merci pour toutes ces discussions sur les résultats, de ta patience, de ta bienveillance et de tous tes encouragements. Je te souhaite le meilleur pour la poursuite de tes aventures, scientifiques ou non.

Alexandre, merci de tes conseils sur la culture d'embryons humains, merci pour ton aide pour les immunofluorescences, merci de ton attitude toujours positive et merci (ou pas) de tes blagues et de l'ambiance que tu apportais(es) au labo ! Merci de croire en toi mais aussi en nous, j'attends avec impatience ton deuxième doctorat !

Eva merci pour les discussions culture cell, pour les discussions scientifiques et de m'avoir donné l'occasion de découvrir ce que c'est de former quelqu'un. Merci pour Vienne, pour Mauges sur Loire, pour tous les moments passés ensemble lors de ces 2 dernières années. Merci aussi pour les discussions sur les chats, pour les potins pendant les pauses thés et pour les compromis de type barre au sol contre natation ! Et comme je

l'ai noté dans un moment de craquage, n'oublions pas ton apport inestimable dans cette thèse le jour où j'ai oublié la traduction de bottleneck.

Océane merci pour ton explication de la protéomique, pour ta gentillesse et ta sagesse. Merci (ou pas) de prendre la suite d'Alex et de nous inciter à aller manger au Trinita tous les 2 jours, merci pour les pauses thé impromptues, pour être disponible même quand tu n'es pas sur site et pour tous les soirs de garde alternée !

Merci à vous 3 d'avoir formé le groupe des totally spies, on a encore de sacrées aventures à vivre tous ensemble !

Nathan merci pour toute l'aide de compréhension de ce qu'est un « tidy » dataframe, merci pour les discussions interminables en salle de culture (tinder, la pop culture, la politique et parfois la science, tout y est passé), merci pour les pauses thés, bon courage pour la suite dans la start up nation et à bientôt à Paris !

Simon merci de m'avoir rappelé tous les 2 jours qu'il fallait que je soumette sur ENA nos données et donc de m'avoir permis de le faire, merci de tout ton apport bio-informatique que ce soit pour mon article ou pour le laboratoire, merci de toutes les discussions sur comment gérer son patrimoine et les règles des prêts en France. Bon courage, c'est toi le prochain !

Manon merci pour ton calme, ta sagesse, ta vision posée des choses et pour tous les tutos coutures et les sorties plages ! Je te souhaite beaucoup de bonheur pendant ton doctorat patiemment attendu mais amplement mérité !

Emilie, merci pour ta fraîcheur et ton apport (bienvenu il faut le dire) en tendance pour le laboratoire ! Je te souhaite plein de réussites !

Merci à tous les autres utilisateurs de la PFiPSC, présents ou passés, Emilien, Julie, merci d'être qui vous êtes !

Merci à tous pour les soirées films, les soirées pizzas ou raclette, les piques niques interrompus par les motos qui brûlent, merci pour tous ces moments passés ensemble.

Merci également du fond du cœur à toute la PF iPSC présente et passée. Isabelle Leray, Elsa Lemaitre, Caroline Chariou, Celia Courty, Sarah Tessier sans oublier, Anne Gaignerie, Coline Rogue, Carole Achard, Quentin Francheteau et Delphine Coulais. Votre gentillesse et votre aide ne sera pas oubliée, merci de gérer aussi bien la plateforme autant d'un point de vue scientifique qu'organisationnel, pour les conseils, pour les formations, pour toutes ces pauses déjeuner, de m'avoir laissée squatter votre bureau bien plus longtemps que nécessaire, de m'avoir soutenue dans mes moments de doute. On se souviendra de la soirée karaoké, de la nourriture omniprésente et des happy hours ! Vous n'êtes peut-être pas à la première place dans cette liste (non ordonnée) mais vous êtes juste avant ma famille (biologique et amicale) parce que pendant ces 3 dernières années, vous avez formé ma famille scientifique. Merci !

Un grand merci à Klaudia, Victoria et Dominic, depuis le master jusqu'au doctorat, merci pour votre soutien, pour vos messages et les moments ensemble.

Merci également à tous mes amis non scientifiques qui m'ont supportée sans forcément comprendre quand je parlais de cellules souches, de trophoblaste, de mes recherches ou du milieu de la recherche en général : Manon, Aymerik, Estelle, Julien, Alexis, Josy, Pierre, Alicia, Cécile, Arthur, Marie, Jeanne, Romain, Jules, Alexandra, Emrick, Thomas. Petits ou grands vous m'avez apporté des pauses, des rires, de la compréhension, merci.

Un immense merci à ma famille. Maman, merci de toujours me soutenir. Merci de m'avoir conseillé d'arrêter si c'est trop dur environ 1 fois tous les 3 mois pendant cette thèse, l'envie d'arrêter n'était pas là mais savoir que tu me soutiendras quoiqu'il se passe m'a aidée et rassurée. Papa, merci de m'avoir dit juste après maman que je ne pouvais pas arrêter, que c'était important parfois de mettre les efforts nécessaires dans quelque chose d'important. Merci de m'avoir soutenue. Péné, merci de me distraire quand je suis dans une spirale négative, merci d'être enjouée et de me faire bouger dans ma réflexion. Ben, merci d'être là, souvent silencieux mais toujours présent. Papy et mamie, merci de m'avoir accueillie pendant mon master, merci de me soutenir et de me demander constamment des traductions de mes travaux (c'est toujours non, désolée). Je vous aime tous très fort.

Un immense merci également à ma belle-famille qui m'a vue un jour arriver et commencer à parler de trophoblaste mais qui s'est quand même intéressée à ce que je faisais. Valérie, Dominique, Adrien, Marie, Corentin, merci pour votre soutien, de ne pas vous être offensés quand j'arrivais en retard à cause de la culture cellulaire ou que j'étais fatiguée du week-end et des expériences. Merci d'être présents, de m'aider et de m'accompagner.

Alexis, quelques mots sur une page ne suffiront pas pour exprimer à quel point ton soutien et ton amour m'aident, me rassurent et me font avancer. Merci pour tout, pour être venu me chercher le soir quand il faisait nuit noire dehors, pour m'avoir soutenue lors de ces interminables sessions de microscopie, pour avoir accepté mes retards incessants et les nuits passées à faire des allers retours au labo. Merci d'avoir écouté inlassablement mes répétitions d'oraux, merci d'avoir relu mes écrits, merci d'expliquer le trophoblaste et la recherche sur le développement péri-implantatoire à tes collègues ingé. Merci d'être toi. Je t'aime.

Et enfin, un remerciement probablement non conventionnel mais que serait cette thèse sans une mention pour mon chat. Merci Roiboos, d'être là.

FRENCH THESIS SUMMARY

Introduction

De nos jours, en France, 1 couple sur 6 est infertile. Selon l'organisation mondiale de la santé, l'infertilité correspond à l'incapacité d'avoir des enfants après 1 an de rapports sexuels réguliers. L'assistance médicale à la procréation (AMP) peut être une solution pour ces couples. Ces techniques regroupent tous les actes médicaux impliquant la manipulation des ovocytes, du sperme ou des deux en dehors du corps humain. L'utilisation de l'AMP augmente chaque année en France et dans le monde d'environ 5 à 10%. D'ici à 10 ans, on estime que ces techniques correspondront à 12 millions de naissances annuelles. Malgré ces chiffres en plein essor, les taux de succès de la fécondation in vitro (FIV ; une des techniques de l'AMP) sont relativement faibles (25-30% de naissances vivantes par cycle en moyenne). Ces taux de succès faibles sont en partie expliqués par des défauts d'implantation. En effet, 60% des embryons transférés ne s'implantent pas, faisant de l'implantation un des majeurs enjeux de la réussite de l'AMP.

Il est devenu clair, au fur et à mesure des années, qu'une meilleure compréhension du développement péri-implantatoire humain est nécessaire afin de pouvoir non seulement optimiser les conditions des cultures embryonnaires mais aussi les méthodes d'évaluation des embryons et pouvoir ainsi augmenter les taux de succès d'AMP tout en limitant les complications.

La période péri-implantatoire correspond à la période allant de 0 à 14 jours de développement, depuis la fécondation en passant par la première puis la seconde spécification des lignées embryonnaires jusqu'à avant la gastrulation. L'étude de cette période chez l'humain est régulée par de nombreuses règles en France comme dans le monde. En France, la recherche sur l'embryon est soumise à autorisation via l'agence de la biomédecine. La loi de bioéthique française a été révisée en 2021 et proscrit la création d'embryons pour la recherche, la culture d'embryons pour plus de 14 jours ou l'insertion de cellules animales dans un embryon humain.

De par le challenge logistique et légal que représente la recherche sur l'embryon humain, les chercheurs ont développé des modèles alternatifs pour étudier le développement embryonnaire.

De par les différences de timing lors du développement embryonnaire chez les différentes espèces de mammifère ainsi que les différences dans les processus d'implantation, transcrire les découvertes faites dans les modèles animaux à l'humain est complexe. La recherche sur le développement péri-implantatoire humain requiert donc des modèles spécifiques au développement humain.

Les cellules souches permettent aux chercheurs de représenter des stades et des destins cellulaires spécifiques d'une façon simplifiée et contrôlée. Elles sont cultivables indéfiniment en laboratoire. Elles peuvent être modifiées génétiquement et permettent une étude mécanistique de la différenciation des lignées embryonnaires.

Il existe de nombreux types de cellules souches pour la modélisation du développement péri implantatoire :

- Les cellules souches pluripotentes naïves représentent la lignée embryonnaire de l'épiblaste pré-implantatoire. Cette lignée donnera le futur fœtus lors du développement. Les cellules souches pluripotentes naïves ont différentes caractéristiques : elles ont une activité métabolique importante face aux autres modèles péri-implantatoires humains ; elles sont hypo-méthylées face aux autres modèles péri-implantatoires humains ; elles ont deux chromosomes X actifs ; elles peuvent se convertir en cellules souches trophoblastiques ; elles peuvent participer à des chimères inter-espèces.
- Les cellules souches pluripotentes amorcées représentent la lignée embryonnaire de l'épiblaste post-implantatoire. Cette lignée donnera le futur fœtus lors du développement. Les cellules souches pluripotentes amorcées ont différentes caractéristiques : elles ont une activité métabolique faible face aux autres modèles péri-implantatoires humains ; elles sont méthylées face aux autres modèles péri-implantatoires humains ; elles ont un seul chromosome X actif ; elles ne peuvent pas se convertir en cellules souches trophoblastiques ; elles ne peuvent pas participer à des chimères inter-espèces.

- Les cellules souches pluripotentes étendues ont été décrites comme représentant un stade de pluripotence plus précoce que les cellules pluripotentes naïves ou amorcées. Les cellules souches pluripotentes étendues ont la capacité de participer à des chimères inter-espèces et de se convertir en cellules souches trophoblastiques.
- Les cellules souches trophoblastiques représentent la lignée embryonnaire du trophoblaste post implantatoire. Cette lignée donnera le futur placenta lors du développement.

Tous ces modèles de cellules souches peuvent être obtenus selon différentes méthodes :

- Dérivation embryonnaire : cela consiste à isoler une lignée cellulaire depuis un embryon puis à la maintenir en culture pour l'utiliser lors de recherches ultérieures.
- Reprogrammation cellulaire : par l'expression forcée de facteurs de transcriptions (OCT4, SOX2, KLF4, C-Myc), des cellules somatiques peuvent être reprogrammées en cellules souches pluripotentes ou trophoblastiques.
- Conversion : des cellules pluripotentes naïves ou étendues peuvent être converties en cellules souches trophoblastiques grâce à un transfert dans le milieu de culture des cellules souches pluripotentes trophoblastiques.

Récemment, la société internationale de recherche sur les cellules souches a publié un ensemble de consignes sur la culture de cellules souches pluripotentes. Ce guide se réfère principalement à la culture de cellules souches pluripotentes amorcées et propose un ensemble de standards que la communauté devrait mettre en place dans le but d'aller vers une recherche plus crédible et fiable.

Résultats

Lors de ce travail de thèse, j'ai donc caractérisé les modèles de cellules souches péri implantatoires sur le plan transcriptomique, protéomique, épigénétique et métabolique.

Nos résultats révèlent que les cellules souches pluripotentes étendues partagent de nombreuses caractéristiques avec les cellules souches pluripotentes amorcées à l'exception de l'activité métabolique. Ce trait métabolique distinct peut expliquer leur capacité unique à se différencier directement en cellules souches trophoblastiques.

De plus, nos recherches démontrent que l'hypo-méthylation de l'ADN et une activité métabolique élevée définissent les cellules souches trophoblastiques.

Ces résultats soulignent la nécessité de considérer plusieurs caractéristiques de la pluripotence, plutôt que de s'appuyer sur un seul critère. La multiplication des caractéristiques atténue les biais de correspondance des stades développementaux.

Lors de ce travail de thèse, j'ai également pu publier un article protocolaire portant sur la reprogrammation de cellules souches pluripotentes naïves à partir de cellules somatiques.

J'ai eu la chance de participer à une collaboration nationale portant à découvrir de nouvelles protéines manquantes (des protéines jusque-là non identifiées avec les techniques de spectrométrie de masse).

Finalement, j'ai participé à la dérivation de nouvelles lignées de cellules souches embryonnaires amorcées et naïves. La dérivation de cellules souches embryonnaires naïves n'a pas abouti mais la dérivation amorcée a abouti à la dérivation d'une lignée, actuellement en cours de validation.

Ces recherches ont abouti à la publication de 3 articles scientifiques dont 2 en premier auteur.

Discussion

L'analyse des données protéomiques générées dans le cadre de mes 2 articles de thèse amène à se poser des questions au regard de la discordance qui peut être observée entre certaines analyses transcriptomiques et protéomiques. Malgré le dogme central de la biologie qui implique une concordance parfaite entre protéine et ARN, l'analyse des données expérimentales nous montre des différences notables. Ces différences peuvent être expliquées par plusieurs facteurs. Premièrement, la spectrométrie de masse reste une technologie relativement récente, avec des challenges inhérents qui rendent difficile la détection de toutes les protéines cellulaires. Certaines protéines peu abondantes ou avec des propriétés physico-chimiques particulières restent difficiles à détecter. De plus, les différences de demie vie ou des altérations post transcriptionnelles peuvent également avoir un impact. Enfin, des difficultés techniques liées à la technologie de spectrométrie de masse peuvent également impacter les conclusions. L'intégration de données protéomiques dans nos jeux de données promeut l'utilisation de ces technologies mais permet également de mieux comprendre les fonctions biologiques de nos cellules.

Le milieu de culture peut avoir un impact majeur sur différentes caractéristiques de nos modèles. De nombreux composants ont été décrits comme ayant un rôle sur la méthylation de l'ADN, l'état d'activation du chromosome X ou encore l'activité métabolique. L'analyse de la composition des milieux de culture peut permettre de prédire des propriétés de cellules cultivées dans des milieux différents. Cependant, puisque le milieu de culture n'est pas la seule cause de ces propriétés, l'analyse fonctionnelle reste la seule façon de valider les hypothèses.

Très peu d'études ont étudié l'état d'activation du chromosome X dans l'embryon humain péri-implantatoire. De même, peu d'études ont été faites sur la méthylation de l'ADN ou encore l'activité métabolique de l'embryon. Malgré l'enthousiasme généré par

les modèles de cellules souches péri-implantatoires, il est important de rappeler que leur statut épigénétique est intimement relié à la cellule d'origine dont elles proviennent. De plus, des différences notables existent entre cellules souches et embryon, tel que la présence bi-allélique de XIST sur le chromosome X. Même si nous devons nous réjouir des avancées techniques que représentent les différents modèles cellulaires à notre portée, nous devons également faire un effort de validation et de description avant de pouvoir utiliser ces modèles pour étudier le développement humain. L'effort de caractérisation réalisé dans Onfray et al doit être répété dans les nouveaux modèles 3D de blastoïdes et d'organoïdes trophoblastiques. Enfin, des études sur l'embryon humain sont également nécessaires afin d'établir une référence claire du développement humain.

Les résultats obtenus entre les différentes lignées de cellules souches trophoblastiques générées selon différentes méthodes (conversion, reprogrammation, dérivation) sont similaires. Le travail réalisé dans Onfray et al nous donne une nouvelle référence pour les modèles trophoblastiques. Avoir une comparaison en parallèle de tous les modèles de cellules souches trophoblastiques disponibles à ce jour serait d'un grand intérêt pour le domaine.

Les différences observées entre les cellules pluripotentes naïves de souris et humaines amènent des questions sur la définition de la pluripotence à travers les espèces. Il est essentiel de reconsidérer les critères utilisés pour définir la pluripotence naïve et de déterminer si les mêmes critères devraient être utilisés et appliqués dans des espèces différentes. Au global, cela souligne la nécessité d'une compréhension plus profonde des fonctions cellulaires basiques des différents types de cellules souches dans les différentes espèces pour déterminer la contribution du destin cellulaire face à la contribution de l'environnement.

Perspectives

Au cours de cette thèse, j'ai comparé les lignées cellulaires trophoblastiques, étendues, amorcées et naïves avec des analyses épigénétiques et fonctionnelles qui ont été précédemment utilisées comme caractéristiques des cellules souches pluripotentes naïves. Mes résultats clarifient le fait que les caractéristiques de la pluripotence ne sont pas prédictives les unes des autres et que la multiplication des caractéristiques atténue le biais de correspondance des stades développementaux.

Les résultats associent sans ambiguïté les cellules souches pluripotentes étendues à un destin pluripotent amorcé (Onfray et al, 2023). Néanmoins, les cellules souches pluripotentes étendues ont clairement une propension clonogénique et un taux de croissance plus élevé que les cellules souches pluripotentes amorcées. Nous avons montré que les cellules souches pluripotentes étendues ont une activité métabolique comparable aux cellules souches pluripotentes naïves, ce qui soulève un paradoxe dans l'association de caractéristiques telles que le chimérisme et la conversion trophoblastique avec le destin des cellules. Cela remet en question l'existence d'un lien entre la conversion du destin et le métabolisme.

À une échelle plus large, mes travaux de définition de l'activité métabolique des cellules souches trophoblastiques et des modèles de cellules souches pluripotentes soulèvent des questions sur l'activité métabolique de l'embryon humain. Le statut métabolique de chaque lignée au sein de l'embryon humain péri-implantatoire n'est pas encore connu. Le métabolisme de l'embryon pourrait aider à déterminer quel embryon a le meilleur potentiel d'implantation.

Enfin, tous ces travaux permettront de reconsidérer l'importance des composants fondamentaux du milieu de culture d'embryons humains et leur impact sur le développement embryonnaire et le potentiel d'implantation. Cela pourrait aider à la formulation de nouveaux milieux de culture d'embryons et donc avoir un impact sur les taux de réussite de la procréation médicalement assistée.

LIST OF FIGURES

Figure 1: Human peri-implantation development, from fertilization to early implantation

Figure 2: Top 20 investors in science and technology research and development and their human embryo laws/guidelines

Figure 3: Developmental timing differs in mammals

Figure 4: Types of implantations (From Siriwardena et al, 2022)

Figure 5: Side of implantation across mammalian species

Figure 6: Different types of pluripotent stem cells

Figure 7: Different human naïve pluripotent stem cell culture medium compositions

Figure 8: Comparison of validation of human extended pluripotent stem cell models

Figure 9: Ways to generate peri-implantation stem cells

Figure 10: Criteria for human trophoblast cells and human trophoblast stem cells

Figure 11: Comparison of validation of human trophoblast models according to Lee et al criteria

Figure 12: Comparison of human primed pluripotent stem cells conversion protocols

Figure 13: Stem cell models of human post-implantation development from Moris, 2023

Figure 14: Derivation of embryonic stem cells

Figure 15: Proteomic and transcriptomic data misalignment

Figure 16: Potential impact of culture medium components on metabolism and DNA methylation

Figure 17: Hallmarks used to evaluate naïve pluripotency in human are not predictive of each other.

TABLE OF ABBREVIATIONS

ART: assisted reproductive technology

EGA: embryonic genome activation

EPI: epiblast

EpiSC: epiblast stem cells

EPSCs: extended pluripotent stem cells

ESC: embryonic stem cells

EVT: extra villous trophoblast

FPSCs: formative pluripotent stem cells

h: human

HLA: histocompatibility leucocyte antigen

i: induced

ICM: inner cell mass

ISSCR: international society for stem cell research

IVF: in vitro fertilization

LIF: leukemia inhibitory factor

m: mouse

MEF: mouse embryonic fibroblast

OKSM: OCT4, KLF4, C-MYC, SOX2

NPSC: naive pluripotent stem cells

PPSC: primed pluripotent stem cells

PrE: primitive endoderm

PSC: pluripotent stem cells

ST: syncytiotrophoblast

TE: trophectoderm

TELC: trophectoderm like cells

TSC; trophoblast stem cells

XCI: X chromosome inactivation

INTRODUCTION

In our society, the choice to have a child is often regarded as the ultimate milestone of life. Elizabeth Stone, a popular writer even says: "Making the decision to have a child- it is momentous. It is to decide forever to have your heart go walking around outside your body". In a 2018 study examining factors contributing to relationship success, a substantial 58% of participants deemed "Having children" as "very important" (Dompnier, 2019). However, for 1 out of every 6 couples, this dream is out of reach without medical assistance due to infertility. Infertility, as defined by the World Health Organization, is a condition of the male or female reproductive system characterized by the inability to achieve pregnancy after 12 months or more of regular unprotected sexual intercourse.

Assisted reproductive technologies (ART) can be a solution for infertile couples. ART are defined as any procedure that involve the handling of eggs, sperm or both outside the human body. This includes artificial insemination, intrauterine insemination, in vitro fertilization (IVF) and ovarian stimulation (with medication). Globally, about 2-4% of pregnancies are now a result of ART, translating to approximately 8 million live births each year. Notably, the use of ART is on the rise, with an annual increase of 5-10% (Ravitsky and Kimmins, 2019). Projections indicate that ART could account for 12 million live births annually within the next decade. However, despite the increasing reliance on ART, the success rate, especially in IVF, remains modest. In 2020 in the USA, 203 164 cycles of ART were started. A cycle of ART corresponds to any process in which an ART procedure is performed, a woman has undergone ovarian stimulation or monitoring with the intent of having an ART procedure, or frozen embryos have been thawed with the intent of transferring them. Among these 203,164 cycles, there were 165,041 embryo transfers. These embryo transfers resulted in 91,453 pregnancies, 75,023 live-birth deliveries (delivery of one or more living infants), and 79,942 infants. Thus, in 2020 in the USA, the success rate of ART technics (the probability of having a live birth after a cycle of ART) was of 36% (2020 National ART summary, 2023). In France, when considering only IVF cycles, there is 25-30% live birth rate per cycle (Fiv.fr, 2019).

Multiple factors have been proposed to explain this relatively low success rate, including embryos and gametes quality, embryo culture conditions (e.g., temperature, atmosphere, culture medium composition), precocious development and most importantly, implantation failure. Indeed, 60% of transferred embryos are lost around implantation, making of implantation a major bottleneck in ART (Edwards, 2006; Polanski et al., 2014).

Altogether, it has become clear that a better understanding of human peri-implantation development is necessary as a way to both optimize embryo culture conditions, embryo evaluation methods and overall to increase ART success rate while limiting complications.

How can we study early human development?

Human embryonic development: a quick overview of the beginning of life

Development starts by the fertilization of an oocyte near the ovary by a spermatozoid. The fertilized zygote will then divide once every 24 hours roughly and start to migrate in the fallopian tube toward the uterus (Figure 1). Until the 3rd day of development, cells are identical and retain the potential to generate a complete individual (Van de Velde et al., 2008; Veiga et al., 1987). These cells are called blastomeres. At the fourth day of development, the blastomeres form a tight group of cells, with indistinct boundaries and form the morula. From this stage, differences in the position of the cells will lead to a different fate. Indeed, external cells will form the trophectoderm (TE) layer and the internal cells will form the “inner cell mass” (ICM) from where will emerge later in development the pluripotent epiblast (EPI) and the primitive endoderm (PrE). This phenomena of segregation of the trophectoderm from the ICM is called the first specification (Shahbazi, 2020) (Figure 1). Interestingly, this process is only visible at the protein level as no transcriptomic difference is yet visible at this stage (Meistermann et al., 2021).

At the same time as the first specification, a fluid-filled cavity, called the blastocoel, starts to form inside the embryo. After the cavitation process, the embryo is called a blastocyst. It is composed of a peripheric TE and of an internal ICM attached to the TE. Within this ICM, 2 distinct cell types gradually emerge: the pluripotent epiblast, destined to form the future fetus, and the primitive endoderm which will form the future yolk sac of the fetus. The emergence of the EPI and PrE from the ICM is commonly referred to as the “second specification”. Similar to the first specification, this process is dynamic and takes place in the time and space of the embryo (Figure 1). Around day 7 of development, the human embryo implants in the uterus and initiates its second week of development (Niakan et al., 2012). This period of development, often referred to as the “black box of human development” (Goedel and Lanner, 2021), poses unique challenges for study. Surprisingly, human embryo can thrive and self-organize independently in vitro, even in the absence of any cellular support. With a specialized medium known as “IVC”, human

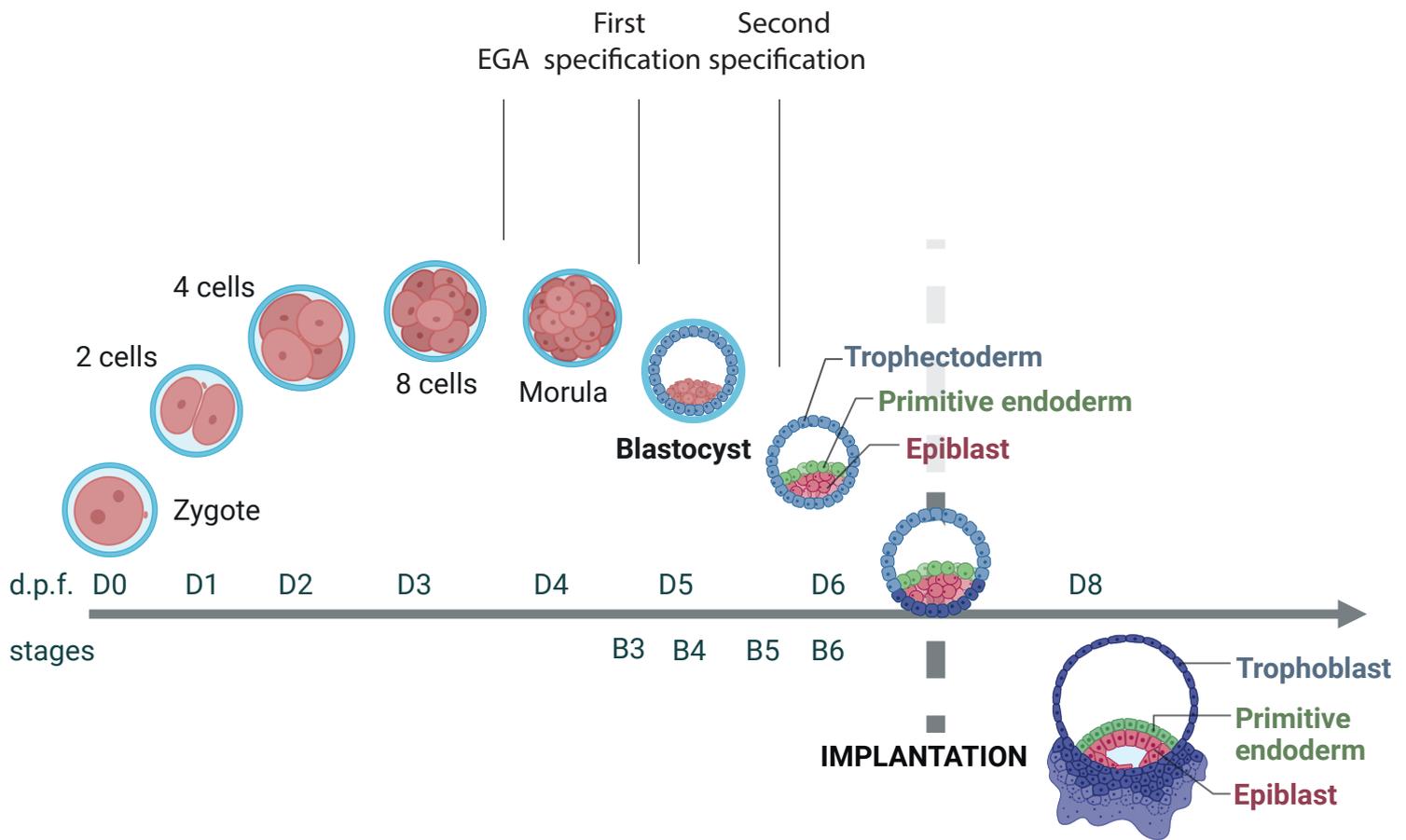


Figure 1 : Human peri-implantation development from fertilization to early implantation
 From left to right, fertilized zygote (day 0); the totipotent zygote undergoes a series of cleavages, yielding the 2-cell (day 1), 4-cell (day 2), 8-cell stages (day 3). Between day 3 and 4, the embryo initiates gene expression, using its own machinery (embryonic genome activation, EGA). On day 4, the embryo undergoes compaction, yielding the morula, at the origin of the first specification (inner cell mass - trophectoderm). By day 5, fluid filled cavities fuse and cavitation of the embryo produces the blastocyst, composed of an inner cell mass encircled by the trophectoderm layer. At day 6, the second specification takes place, segregating the epiblast from the primitive endoderm. At this stage, the embryo implants in the maternal endometrium and further differentiation of the trophoblastic lineage starts.
 Abbreviations : d.p.f. days post fertilization; EGA embryonic genome activation; D days

embryos can attach to an in vitro culture plaque without cells or matrix. They can successfully develop and form morphological structures specific to the second week of development such as the amniotic cavity or the vitellin cavity (Deglincerti et al., 2016; Shahbazi et al., 2016). To our knowledge, specification of EPI and PrE is incomplete before implantation and thus these lineages continue to mature during the 2nd week of development. Research by Deglincerti et al shows that the spatial segregation of EPI and PrE cells is not complete by day 6 but becomes evident by day 8. After 15 days of development, the embryo starts the gastrulation process, marking the transition from a simple, single layered structure to a complex multi layered embryo. Gastrulation plays a fundamental role in establishing the basic body plan to initiate the formation of various organ systems. It also marks the end of the peri-implantation period.

The development of protocols for in vitro culture of human embryos without cellular support has significantly advanced our understanding of early development. These protocols allow embryos to survive in reproducible conditions until day 13 and enable the modeling of key developmental stages. However, it's important to note that success rates remain low, with a survival rate of 12% at day 12 in our hands (Moinard et al in preparation).

Regulations of human embryo culture: International perspectives and France focus

Research on human embryos operates within a multifaceted framework defined by religious beliefs, national culture, politics, and historical context. Therefore, what is permitted or prohibited in human embryo research varies significantly from one country to another. Many countries however found a common ground in adopting the “14-day rule” which limits in vitro embryo culture to 14 days after fertilization. This limit was proposed in 2 reports on IVF: the 1979 US department of Health, Education and Welfare report and the 1984 UK Warnock report (Hurlbut, 2017).

The justification for the 14-day limit lies in the emergence of the primitive streak, which occurs around the 14th day of development. The primitive streak is considered the first

visible sign of embryonic organization and marks the last point at which twinning can occur. Notably, when this rule was established, it was not feasible to culture human embryos *in vitro* beyond the implantation stage.

Over the 20 countries with the highest gross expenditures in research and development worldwide (USA, China, Japan, Germany, South Korea, France, India, United Kingdom, Russia, Brazil, Taiwan, Italie, Canada, Spain, Turkey, Australia, Switzerland, Netherland, Sweden, Israel), 12 implemented the 14 days rules as a National law or governmental guideline (China, Japan, South Korea, France, India, United Kingdom, Taiwan, Canada, Spain, Australia, Netherland, Sweden) (Matthews and Moralí, 2020) (Figure 2).

Brazil's laws on human embryonic stem cells (hESC) research prohibit “*genetic engineering on human germ cells, human zygotes or human embryos*”. However, they do not address a development limit or other restrictions on human embryo research (Palma et al., 2015; Pranke et al., 2014).

Israel has a 1999 law banning reproductive cloning, but it does not address nor limit *in vitro* human embryo research ((Israel) Public Health Regulations (Extra-Corporeal Fertilization), KT 5035 p. 978 (1987); (Israel) Prohibition on Genetic Intervention (Human Cloning and Genetic Change in Reproductive Cells) Law, SH 1697, p. 47.; (Israel) Egg Donation Law SH 2242 p. 520 (2010)).

In the USA, although the 14 days limit has been proposed, it has not been implemented as a federal law (Hurlbut, 2017). However, federal funding for human embryo research is prohibited through the Dickey-Wicker Amendment (Hurlbut, 2017; Matthews and Rowland, 2011; Rep. Pascrell, 2019).

Switzerland has a 7 days limit ((Switzerland) Federal Act on Research Involving Embryonic Stem Cells (Stem Cell Research Act, StRA) (2003, amended 2005, 2014) 810.31 (2003)).

Germany, Russia, Italy and Turkey prohibit embryo research entirely ((Germany) Embryo Protection Act 2746, (1990).; (Germany) Stem Cell Act 2277 (2002).; German Ethics Council. Stem cell research – new challenges for the ban on cloning and treatment of artificially created germ cells? Deutscher Ethikrat, Berlin, Germany (2014).; (Turkey)

Country	Gross expenditure in research and development (2020)	Developmental limit
U.S.A	549 billions US\$	No national or nongovernmental guidelines
China	496 billions US\$	14 days post fertilization
Japan	170,9 billions US\$	14 days post fertilization
Germany	132 billions US\$	Prohibited
South Korea	91 billions US\$	14 days post fertilization
France	64,7 billions US\$	14 days post fertilization
India	49,77 billions US\$	14 days post fertilization
United Kingdom	49,3 billions US\$	14 days post fertilization
Russia	41,9 billions US\$	Prohibited
Brazil	39,9 billions US\$	No national or nongovernmental guidelines
Taiwan	39,3 billions US\$	14 days post fertilization
Italy	33,5 billions US\$	Prohibited
Canada	27,2 billions US\$	14 days post fertilization
Spain	21,9 billions US\$	14 days post fertilization
Turkey	21,7 billions US\$	Prohibited
Australia	21,2 billions US\$	14 days post fertilization
Switzerland	18,9 billions US\$	7 days post fertilization
Netherland	18,6 billions US\$	14 days post fertilization
Sweden	17,6 billions US\$	14 days post fertilization
Israel	15,4 billions US\$	No national or nongovernmental guidelines

Figure 2 : Top 20 investors in science and technology research and development and their human embryo laws/guidelines.

Regulation on Assisted Reproductive Treatment Practices and Assisted Reproductive Treatment Centers (1987).; Gurtin ZB. Et al, 2011; (Italy) Rules on medically assisted procreation 40 (2004).; (Russia) Federal Law On Biomedical Cell Products 180-FZ (2016)).

In the case of France, embryo research is permitted under specific conditions, including the use of surplus frozen embryos generated by IVF and donated to research by both parents. Importantly, it is prohibited to create embryos for research purposes, introduce cells from another species into a human embryo, or culture human embryos for more than 14 days (until the first signs of gastrulation). Research activities are rigorously overseen by the Agence de la Biomédecine, which ensures compliance with legal and ethical guidelines

The use of human embryos in research projects remains a complex challenge both from a regulatory and logistical point of view. Limited numbers, variable quality, and sample variability have prompted researchers to explore alternative models for studying embryonic development.

In the subsequent sections of this thesis introduction, we will delve into various models developed to study peri-implantation development, including mammalian embryo models, stem cell models, and advanced 3D stem cell models (e.g., "blastoids," "stembryo," etc.).

Exploring Pre- and Post- implantation development: animal models

For practical and legal reasons, a significant proportion of research on pre- and post-implantation development has initially focused on the mouse model. This led to significant progress in our understanding of lineage differentiation including crucial aspects like the first and second specification events. However, it's important to note that numerous other mammalian species offer valuable insights into peri-implantation development. Among the commonly accepted and used models are rabbit embryos, cattle embryos, and non-human primate embryos.

Developmental timing differs in mammals

While each mammalian model shares key characteristics and events with human peri-implantation development, the timing of major events varies significantly among different mammalian species.

Embryonic genome activation denotes of the initiation of gene expression, using embryonic machinery, after fertilization. In mice, this crucial event occurs at the 2-cell stage (AOKI, 2022), while in human and porcine embryos, it takes place between the 8-cell and 16-cell stages (Braude et al., 1988; Leng et al., 2019; Sirard, 2012; Tesařek et al., 1988; Vassena et al., 2011; Xue et al., 2013; Yan et al., 2013). In rabbit and bovine embryos, major embryonic genome activation occurs at the 16-cell stage (Graf et al., 2014; Sirard, 2012; Telford et al., 1990) (Figure 3).

After embryonic genome activation (EGA), embryos continue to cleave and develop. The next milestone in its development is compaction. During compaction, cells come into close contact and form a tighter structure (Firmin and Maître, 2021; Shahbazi, 2020). When it comes to compaction, timing again varies across mammalian species. In mice, compaction occurs at day 2,5 after fertilization, typically between the 8-cell and 16-cell stages (Jedrusik, 2015). In humans, compaction takes place at day 4, between the 8-cell

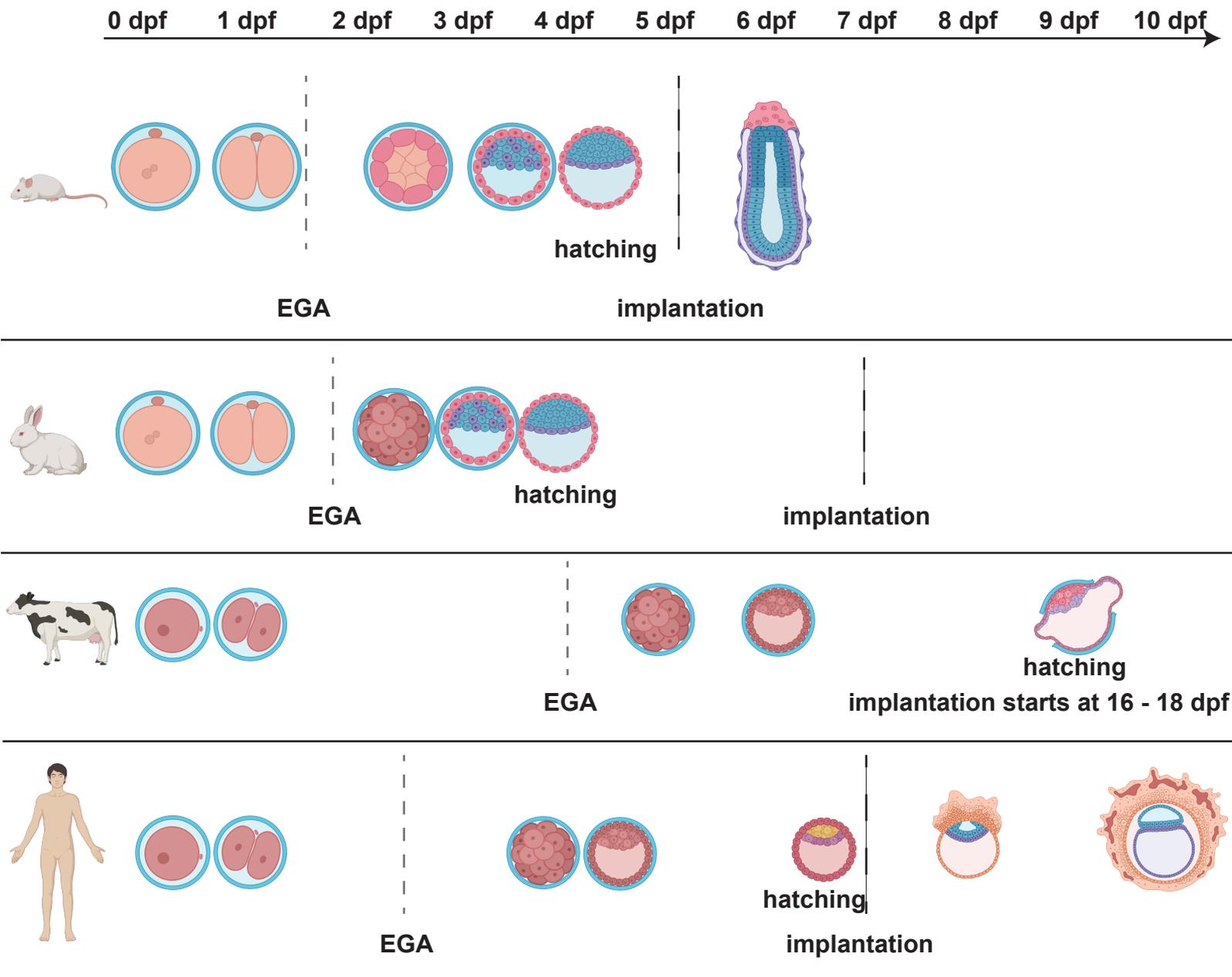


Figure 3 : Developmental timing differs in mammals.

This schematic shows development events during days after fertilization in mouse, rabbit, bovine and human embryos. Embryos are distributed along a temporal line starting with the zygote at 0,5 days post-fertilization. The timing of EGA is indicated for each animal.

Abbreviations: dpf days post fertilization, EGA embryonic genome activation.

and 16-cell stages, similar to mice (Coticchio et al., 2019; Firmin and Maître, 2021; Iwata et al., 2014; Shahbazi, 2020). In bovine embryos, compaction takes place at day 5 and at the 32 cells stage (Soom et al., 1997). Finally, in rabbit embryos, compaction occurs at day 2,5 and at the 32-64 cell stage (Johnson and Ziomek, 1981) (Figure 3).

In the compacted morula, fluid filled lumens start to form between the inner and outer cells. These lumens converge into one single lumen, leading to emergence of the blastocyst structure (Dumortier et al., 2019). Cavitation starts at the 32-cell stage in mouse, between 3 and 3,5 days post fertilization (Brinster, 1963; Wiley and Eglitis, 1981). In human, the blastocyst emerge between day 4 and 5 (13-15 hours post compaction (Meistermann et al., 2021)), around the 32-cell stage (Hertig et al., 1959; Steptoe et al., 1971). In cattle, blastulation starts at day 6, when the embryo is composed of around 100 cells (Soom et al., 1997). Finally, in rabbit embryos, blastulation happens at day 3 post fertilization (Blerkom and Manes) (Figure 3).

The blastocyst will then hatch from the zona pellucida prior to implantation. In mouse, blastocyst hatch around 4 days post conception (An et al., 2021). Human embryos hatch between day 6 and day 7 (Sathananthan et al., 2003). Bovine blastocyst hatch around 9 days post conception (Soom et al., 1997). In the rabbit, the zona pellucida disappears between day 3 and 4 and is replaced by a new layer, the neozona (Fischer et al., 1991) (Figure 3).

These differences in timing and in stages for essential timepoints in embryonic development highlight the need to study human embryonic development to uncover human specific timely events.

Exploring implantation across mammalian species

Finally, while the mechanism initiating trophoblast specification is conserved across mice, cows and humans (Gerri et al., 2022), the process of implantation itself exhibits notable variations.

Implantation encompasses several key events, including the apposition of the blastocyst to the uterine luminal epithelium, adhesion to the epithelium, penetration through the epithelium and basal lamina and invasion into the stromal vasculature (Schlafke and Enders, 1975). Apposition is mediated via the trophoblast layer. Implantation is classified in 3 main categories (Wimsatt, 1975):

- Centric implantation: The blastocyst grows large and form ample surface contact to fuse with the luminal epithelium without penetrating through it (found in rabbits, cows, pigs, sheeps, dogs and marsupials)
- Eccentric implantation: The luminal epithelium forms an invagination to surround the trophoblast (found in mice, rats and hamster).
- Interstitial implantation: Trophoblast passes through the luminal epithelium to invade the epithelium stroma and becomes imbedded into the wall of the uterus (found in human, bats, lesser apes, great apes and guinea pig) (Siriwardena and Boroviak, 2022; Wimsatt, 1975) (Figure 4).

Mouse embryos appose from the mural side of the trophoblast which is the side opposite to the epiblast. They implant through apposition, attachment and invagination of the uterine epithelium on the blastocysts (eccentric implantation) (Dey et al., 2004) (Figure 5).

Rabbit embryos lose their polar trophoblast and appose from the epiblast (Williams and Biggers, 1990). The blastocyst adheres solely to the apices of the epithelial cells (Schlafke and Enders, 1975) (Figure 5).

Bovine embryos follow a different pattern, as they do not implant before day 16 of development. They undergo gastrulation and elongation before implantation. Trophoblast cells fuse to the uterine epithelium, creating a structure known as a placentome (synepitheliochorial implantation) (Spencer and Hansen, 2015; Wooding, 1992) (Figure 5).

Non-human primate implantation varies greatly between species. Most primates appose from the polar side of the trophoblast (Enders and Lopata, 1999; Enders et al., 1983; Jones et al., 2001; Smith et al., 1987). However, some primates exhibit centric implantation such as lemur, marmoset, baboon, rhesus macaque, while other have

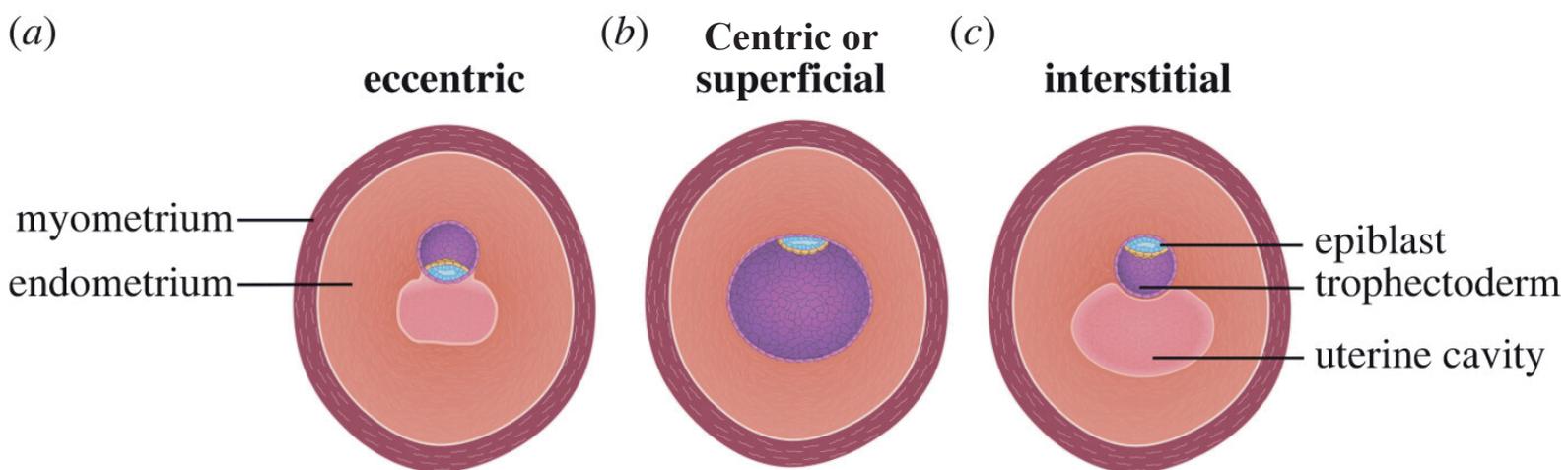


Figure 4: Types of implantation (From Siriwardena et al, 2022).

Types of implantation strategies employed by mouse, marmoset, rhesus macaque and human.

(a) Eccentric implantation: The luminal epithelium forms an invagination to surround the trophoblast. Found in mice, rats and hamster

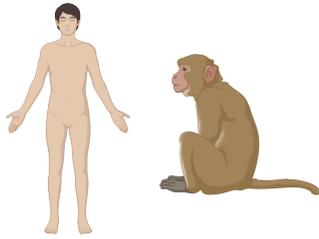
(b) Superficial implantation: The blastocyst grows large and form ample surface contact to fuse with the luminal epithelium without penetrating through it. Found in rabbits, sheep, cows, dogs and marsupials.

(c) Interstitial implantation: Trophoblast passes through the luminal epithelium to invade the epithelium stroma and becomes imbedded into the wall of the uterus. Found in human, bats, lesser apes, great apes and guinea pig.

interstitial implantation, including lesser apes, great apes (Roth, 1988). These early implantation studies in non-human primates are difficult to perform given the low fecundity of the species involved and difficulty in ascertaining early pregnancy (Lee and DeMayo, 2004; Siriwardena and Boroviak, 2022).

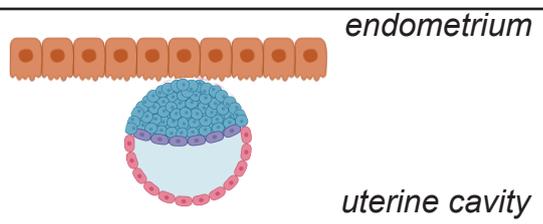
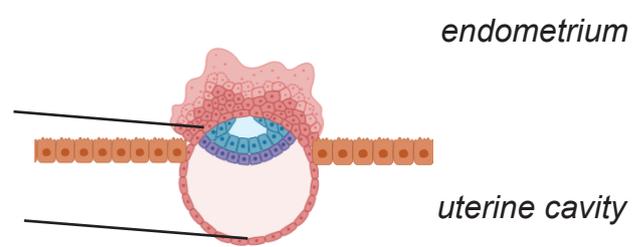
In the case of human, embryos appose from the polar side of the embryo, adjacent to the epiblast. Human embryo implants through interstitial implantation (James et al., 2012) (Figure 5).

While legal and ethical limitations have made the use of animal models crucial to understand development, predicting the extent to which findings in animal models can be translated to human development remains challenging due to differences in developmental pace and fundamental mechanisms of implantation. Considering our research problematic focus on ART and the implantation bottleneck, we require models that better represent this specific time of development in human.



Polar trophoblast

Mural trophoblast



Mural trophoblast

Polar trophoblast

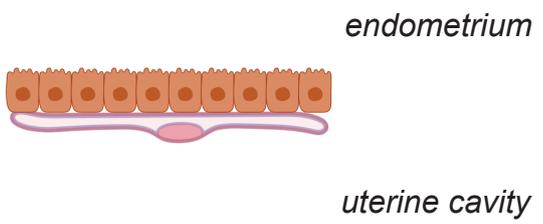
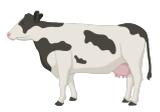
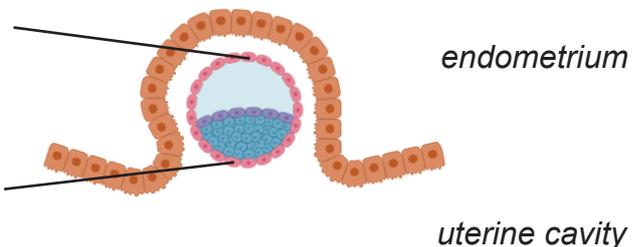


Figure 5 : Side of implantation accross mammalian species.
Schematic of implantation side across different mammalian species.
Note that marmoset do not invade.

Stem cell models of human embryo development

Cellular models enable researchers to represent a specific stage or fate of development in a simplified and controlled manner. Due to culture and in vitro conditions, there are some limitations to the use of cellular models: there is no blood system, immune system, lymphatic clearance or hormonal regulation. Therefore, it is of crucial importance to understand the complexity of the modeled system to define the culture conditions but also to determine to which point the cellular model represents our biological system.

What is a stem cell model?

The first reference to stem cell that can be found in the literature is attributed to Ernst Haeckel in 1872. He used the term 'Stammzelle', the German translation for stem cells, to describe the ancestor unicellular organism from which he thought all multicellular organisms evolved (Haeckel, 1872). In 1877, he also proposed that the fertilized egg should be called a stem cell (Haeckel, 1877). 15 years later, following studies in the *Ascaris* worm, Boveri proposed that cells along the germline lineage between the fertilized egg and committed germ cells should be called stem cells (Boveri, 1892). At the same time, Hacker identified in crustacean *Cyclops* one cell that is internalized upon gastrulation. Hacker proposed to call this cell "stem cell" after observing that this cell was capable of asymmetric division, with one daughter cell giving rise to the mesoderm and the other giving rise to the germ cells (Häcker, 1892). The term stem cells was then used by Edmund B. Wilson in his book *The cell development and inheritance* (Wilson and Wilson, 1896). In his book, he defined stem cells as the unspecialized mother cells of the germline. Thus, the term "stem cells" was first used to refer to germ cells or cells giving rise to germ cells.

The first mention to stem cells in the sense we understand it today was made by Pappenheim. He used it to describe a precursor cell capable of giving rise to both red and white blood cells (Pappenheim, 1896). It was not until the beginning of the new century

that the term was popularized as a way to refer to common precursor of the blood system (Ramalho-Santos and Willenbring, 2007).

Nowadays, according to the Oxford Language dictionary, a stem cell is defined as “an undifferentiated cell of a multicellular organism which is capable of giving rise to indefinitely more cells of the same type, and from which certain other kinds of cell arise by differentiation” (Morrison et al., 1997; Till and McCulloch, 1980; Weissman, 2000).

Pioneering works on stem cells, as defined above, took place in the 1950's. Charles Leblond characterized stem cells in situ for the first-time using autoradiography (Leblond et al., 1950). Later work from his team led to the establishment of stem cells in renewing tissues such as the skin or the intestine (Leblond, 1981; Leblond et al., 1959; Marques-Pereira and Leblond, 1965). In parallel, George Mathé and others showed that bone marrow damaged by ionizing irradiation could be regenerated from transplanted marrow cells, first in animals. In 1958, Mathé applied this discovery in human to rescue several Yugoslav physicists that were inadvertently exposed to radiations in a reactor accident. Mathé brought the physicists to Paris and infused them with donor marrow. This marked the first allogenic human bone marrow transplant and the first proof that stem cells could restore damaged environment in animal and human (Jansen, 2005; Mathe and Bernard, 1959; Mathe et al., 1959a, 1959b).

James Till and Ernest McCulloch continued this research by investigating the impact of radiation on healthy and cancer cells by grafting bone marrow on mice (Till and McCULLOCH, 1961). They observed that the more cells they injected, the higher the mice's survival rate was and that there was mass of cells growing in the spleen proportionate to the number of injected cells. They thus speculated that they might be in presence of newly formed cells. They characterized the population of cells and determined that each mass or colony, arose from a single cell (Becker et al., 1963). They then showed that every colony could generate clonally new ones (Till et al., 1964). Through their work, they provided the first demonstration of self-renewal, the first stemness property. Later, they showed that self-renewing cells from the initial colony can specialize into different types of mature blood cells, thereby characterizing the second stemness property: differentiation (Wu et al., 1967).

Collectively, this work enabled researchers to determine that stem cells play a pivotal role in tissue formation and maintenance throughout life. Later work further revealed the existence of various types of stem cells with varying differentiation potential.

The different kinds of stem cells

After Till and McCulloch's discovery of hematopoietic stem cells, many other types of "adult stem cells" were subsequently identified. Adult stem cells, such as hematopoietic stem cells or intestinal stem cells are found in the post-natal body and can also be found in the umbilical cord at birth. They can be defined as "multipotent" or "unipotent".

In contrast, embryonic stem cells were not isolated from mouse embryos until 20 years later, in 1981 (Evans and Kaufman, 1981). This achievement was accomplished by Martin Evans, in collaboration with Matthew Kaufman, who later received the Nobel prize for their groundbreaking work, jointly with Mario Capecchi and Oliver Smithies in 2007. In 1998, the biologist J. Thomson derived human embryonic stem cells by plating human blastocysts and cultivating the resulting outgrowth (Thomson et al., 1998). Embryonic stem cells serve as the precursor to all cells within the adult organism. They can be defined as "pluripotent".

All stem cells share 2 basic stemness properties: self-renewal and differentiation potential.

Totipotency

Totipotency refers to the ability to give rise to an entire conceptus, encompassing both the embryo and its extraembryonic tissues. In vivo, totipotent cells correspond to the zygote up to cells of the 8-cell stage (Tarkowski, 1959; Tarkowski and Wróblewska, 1967).

Studies involving human naive pluripotent stem cells (hNPSCs) showed that these cells have the ability to convert into trophoblast stem cells (hTSC) (Castel et al., 2020; Cinkornpumin et al., 2020; Dong et al., 2020; Io et al., 2021) but also into extra embryonic

mesoderm (Pham et al., 2022); These studies suggested the presence of a totipotent like population of cells in human NPSCs. Recent research has reported the presence of 8 cell like cells in naive pluripotent stem cells culture, modeling EGA-like transcriptional program, using single cell gene expression analysis (Balaton and Pasque, 2022; Taubenschmid-Stowers et al., 2022). This stem cell population hold promises for in depth research on EGA-like transcription and enhancing our knowledge of human pluripotency and development. It is important to note that while having expression profile showing some similarities with blastomeres, the hNPSC are incapable on their own to form a blastocyst. However, hNPSCs, subjected to a specific environment, can form all cells from an embryo, including the extra-embryonic ones. As hNPSCs can form all cells of an embryo but cannot form an embryo, it raises a question about whether the definition of totipotency should be revised.

Pluripotency

Pluripotency is defined as the capacity of cells to form the entire fetus but not the extra-embryonic lineages (Gardner et al., 1985; Smith, 2009). This definition has been challenged by studies in human, demonstrating the ability of hNPSCs to give rise to trophoblast stem cells and extraembryonic mesoderm (Castel et al., 2020; Cinkornpumin et al., 2020; Dong et al., 2020; Guo et al., 2021; Io et al., 2021; Pham et al., 2022).

Pluripotent stem cells possess the unlimited capacity to divide, self-renew and differentiate into cells representing the early primary cell layers, namely *mesoderm*, *endoderm*, and *ectoderm*. In vivo, pluripotent cells have a limited presence during a specific stage of embryonic development before differentiating into more specialized tissues. Pluripotent stem cells correspond to embryonic stem cells, which are derived from human embryos.

Currently, different states of pluripotency have been described, including primed, naive, extended and formative pluripotency, which will be further explored in this thesis.

Multipotency

Multipotency is defined by the ability of stem cells to differentiate into multiple cell types associated with one specific cell lineage. Multipotent stem cells correspond to adult stem cells and are typically found in the organs or tissues in the body in which they are normally located. Some multipotent stem cells can also be found in the fetus such as the hematopoietic stem cells in the fetal liver during early development or mesenchymal stem cells in the umbilical cord.

Unipotency

Unipotency is the ability to differentiate into a single cell type. For example, unipotent stem cells include muscle stem cells or spermatogonia stem cells, which can respectively only differentiate into skeletal muscle cells and sperm cells (Visvader and Clevers, 2016).

Pluripotent stem cells

Human embryonic stem cells (hESCs) are derived from the pluripotent EPI of the embryo. The first hESC line was successfully derived in 1998 (Thomson et al., 1998), 17 years after the mouse (Evans and Kaufman, 1981; Martin, 1981). These cells possess the ability to differentiate into all three embryonic germ layers (De Los Angeles et al., 2015). Pluripotency is a highly dynamic property that evolves at different stages of peri-implantation development (Hackett and Surani, 2014). Thus, a spectrum of pluripotent stem cells models various stages of peri-implantation development (Hackett and Surani, 2014; Hackett et al., 2017; Hough et al., 2014; Nichols and Smith, 2009). It's important to note that some pluripotent states may also be artefacts of the in vitro culture environment (Gokhale et al., 2015; Ying and Smith, 2017):

- Pre implantation represented by Naive pluripotent stem cells (NPSCs) (Gafni et al., 2013; Giulitti et al., 2019; Guo et al., 2016, 2014; Kilens et al., 2018; Leitch et al., 2013; Liu et al., 2017; Pastor et al., 2016; Smith et al., 2014; Takashima et al., 2014; Theunissen et al., 2014; Ware et al., 2014; Yan et al., 2013).
- Post implantation represented by Primed pluripotent stem cells (PPSCs) (Nichols and Smith, 2009).
- Intermediate stages represented by Formative pluripotent stem cells (FPSCs) (Kinoshita et al., 2021; Rostovskaya et al., 2019) and Extended pluripotent stem cells (EPSCs) (Castel et al., 2020; Gao et al., 2019; Liu et al., 2021a; Tan et al., 2021a; Yang et al., 2017b; Zheng et al., 2021a) (Figure 6).

Pluripotent stem cells represent a major advancement in the developmental biology field and hold great promise in regenerative medicine (reviewed in Yamanaka, 2020). Thus, groundbreaking achievement of derivation of embryonic stem cells and generating pluripotent stem cells via reprogramming was awarded by the Nobel prize in respectively 2007 and 2012 (Evans and Kaufman, 1981; Takahashi and Yamanaka, 2006).

developmental stage blastoid formation potential germline induction potential interspecies chimerism potential culture medium

 hNPSC	Pre implantation epiblast	✓	✗	✓	T2iLGö ; PXGL; 5iLAF; NHSM
 hPPSC	Post implantation epiblast	✗	✗	✗	E8; mTeSR; IPS Brew ...
 hEPSC	Post implantation epiblast	✓	✗	✓	LCDM
 hFPSC	intermediate epiblast	✗	✓	✗	A _{low} XR; XAF

Figure 6 : Different types of pluripotent stem cells
 Represented here is a table with the main different types of pluripotent stem cells and their major characteristics
 Abbreviations : hNPSCs human naive pluripotent stem cells; hFPSCs human formative pluripotent stem cells;
 hPPSCs human primed pluripotent stem cells; hEPSCs human extended pluripotent stem cells

Naive pluripotent stem cells

In 1981, a significant milestone was achieved in mouse research when two separate teams reported the derivation of pluripotent stem cells from either the inner cell mass (ICM) of the intact mouse blastocyst (Evans and Kaufman, 1981; Martin, 1981). These cells exhibited the capacity to differentiate into all three germ layers, form teratomas and contribute to chimera formation with germline transmission (Bradley et al., 1984). Mouse embryonic stem cells were shown to be in an “ICM-like” state (Hanna et al., 2009a) and are referred as “naive” pluripotent stem cells. Early culture conditions of mouse ESCs culture was composed of irradiated mouse embryonic fibroblasts (MEFs), bovine serum albumin and LIF to activate STAT3 (Matsuda et al., 1999; Niwa et al., 1998; Raz et al., 1999). To improve heterogeneity and Naive transcription factors expression (Chambers et al., 2007; Hayashi et al., 2008; Toyooka et al., 2008), culture conditions then evolved to include MEK/ERK and GSK3b inhibition (known as 2i medium, (Buehr et al., 2008; Hanna et al., 2009b; Silva et al., 2008; Ying et al., 2008)). It was proposed that NPSCs are sustained primarily by preventing differentiation (Martello and Smith, 2014).

It took 33 years to derive a comparable state in human (Gafni et al., 2013; Guo et al., 2016; Pastor et al., 2016; Theunissen et al., 2014; Ware et al., 2014). Human naive pluripotent stem cells (hNPSCs) model human pre-implantation epiblast (Guo et al., 2016; Stirparo et al., 2018; Theunissen et al., 2016). They can be obtained by direct derivation from human embryo (Gafni et al., 2013; Guo et al., 2016; Ware et al., 2014), conversion from hPPSCs (Guo et al., 2017; Hanna et al., 2010; Takashima et al., 2014; Theunissen et al., 2014) or reprogramming (Giulitti et al., 2019; Kilens et al., 2018; Liu et al., 2017). Unlike mouse NPSCs (Buehr et al., 2008; Ying et al., 2008), human NPSCs do not thrive in 2i medium alone and necessitate more complex culture media. Initially, human NPSCs culture medium was composed of MAPK, GSK3 β and PKC inhibitor and LIF (Takashima et al., 2014).

More studies proposed alternative culture mediums such as 5iLAF (LIF, MEK inhibitor, BRAF inhibitor, GSK3 β inhibitor, SRC inhibitor, Activin A and β FGF), which was demonstrated as optimal for conversion or derivation of hNPSCs; or NHSM (LIF, TGF β ,

FGF2, MEK inhibitor, GSK3 β inhibitor, JNF inhibitor, p38 inhibitor, Rock inhibitor, PKC inhibitor) (Gafni et al., 2013; Theunissen et al., 2014) (Figure 7).

More recent studies have found that inhibiting the Wnt pathway (using XAV939) benefits the stabilization of naive pluripotency during induction and expansion (Bredenkamp et al., 2019). This is in direct opposition to findings in the mouse. However, human NPSCs have low expression of TCF3 (TCF7L1) and do not express ESRRB (Rostovskaya et al., 2019; Takashima et al., 2014), the key component regulated by GSK3 inhibition in mouse ESCs (Martello et al., 2012; Wray et al., 2011). This led to the development of PXGL culture medium (Lif, Wnt inhibitor, MEK inhibitor, PKC inhibitor, Rock inhibitor), which is proposed to yield more robust and stable cells than T2iLGö medium, while yielding cells with equivalent naive features (Bredenkamp et al., 2019). It was later shown that Wnt/ beta catenin pathway was a major priming agent for human NPSCs (Bayerl et al., 2021) (Figure 7).

T2iLGö, 5iLAF, NHSM and PXGL culture media all require culture of hNPSCs on a MEFs layer. A recent study also proposed to adapt NHSM culture medium to generate human NPSCs without MEFs. This adapted NHSM medium is composed of LIF, MEK inhibitor, PKC inhibitor, Wnt inhibitor, SRC inhibitor, Activin A, Rock inhibitor and P38 inhibitor (Bayerl et al., 2021) (Figure 7).

Human NPSCs are characterized by distinct features at the transcriptomic, proteomic, metabolic and epigenetic level compared to PPSCs.

Transcriptomic and proteomic features

Human NPSCs are characterized by the expression of specific pluripotency markers, present in pre-implantation epiblast:

- Classical pluripotency marker: NANOG, SOX2...
- Naïve specific markers: KLF17, DNMT3L, DPPA5...

	Tankirase inhibitor / Wnt inhibitor (XAV939)	Mek inhibitor (PD032591)	LiF	PKC inhibitor (Gö6983)	Rock inhibitor (Y27632)	GSK3 inhibitor / Wnt activator (CHIR / IM-12)	SRC inhibitor (SB590885)	Activin A	FGF2	TGFb	JNFi (SP600125)	p38i (SB203580 / BIRB0796)
PXGL (Benderkamp, 2019)	✓	✓	✓	✓	✓							
T2iLGö (Takashima, 2014)		✓	✓	✓	✓	✓						
5iLAF (Theunissen, 2014)		✓	✓		✓	✓	✓	✓	✓	✓		
NHSM (Gafni, 2013)		✓	✓	✓	✓	✓			✓	✓	✓	✓
eNHSM (Bayerl, 2019)	✓	✓	✓		✓			✓	✓			✓

Figure 7 : Different human naive pluripotent stem cells culture mediums composition
 Represented here is a table with the main culture medium of human naive pluripotent stem cells

On top of that, NPSCs express the same set of genes as human pre implantation epiblast (Castel et al., 2020; Kilens et al., 2018).

Metabolic features

Human NPSCs have been shown to have a high metabolic activity compared to primed pluripotent stem cells; not restricted to glycolysis (Gu et al., 2016; Kilens et al., 2018).

Epigenetic features

Human NPSCs have been described to be hypomethylated compared to PPSCs. On top of that, NPSCs feature 2 active X chromosomes in the female lines

Chimerism potential

Human NPSCs have been described to contribute to human-animal pre implantation chimera, although at low efficiency. This phenomenon has been attributed to a blockage in G1 phase following the chimeric insertion of the cells (Aksoy et al, 2021).

Primed pluripotent stem cells

17 years after the derivation of the first mice ESCs lines, J. Thomson was able to derive 5 human embryonic stem cell lines from embryos produced by IVF (Thomson et al., 1998). These hESCs were cultivated on a MEF layer in a medium containing serum and LIF. Optimal culture condition of human ESCs were found to differ from mouse ESCs culture conditions, as LIF alone was insufficient to maintain pluripotency in human

(Dahéron et al., 2004). Similarly, the commonly used 2i medium, which is effective for mouse ESCs, was unable to sustain human pluripotency (Hanna et al., 2010). It was shown that human ESCs require activation of Activin/Nodal and FGF signaling pathways to sustain their pluripotency (Beattie et al., 2005; James et al., 2005; Vallier et al., 2005). Despite these variations in culture requirements, human ESCs and mouse ESCs share fundamental characteristics, thus, the differences in signaling pathway necessary for their maintenance were first attributed to unknown genetic differences between species as human ESCs were also derived from the ICM.

In 2007, two separate teams reported the generation and maintenance of pluripotent stem cells from mouse post implantation blastocysts expressing core pluripotency markers Oct4, Sox2 and Nanog. These cells, referred to as mouse epiblast stem cells (mEpiSCs), were able to form teratoma in vivo and to differentiate into derivatives of the three germ layers in vitro (Brons et al., 2007; Tesar et al., 2007). mEpiSCs require FGF and Activin to maintain their pluripotent state, do not resist single cell passaging and cannot contribute to chimeras. They represent a more advanced developmental stage than mouse ESCs. This discovery proved the existence of distinct states of pluripotency and led to the distinction between naive and primed pluripotent stem cells (Nichols and Smith, 2009; Rossant, 2008). Since mEpiSCs share the similar culture requirement, morphology, molecular and epigenetic characteristics as hESCs, their discovery enabled to determine that human ESCs lines derived by Thomson are in a primed state of pluripotency (Brons et al., 2007; Davidson et al., 2015; De Los Angeles et al., 2012; Rossant, 2015; Tesar et al., 2007). It was later shown, through the derivation of mEpiSCs from mouse pre implantation blastocyst, that specific culture conditions support the maintenance of a specific stage (Guo et al., 2009; Hanna et al., 2009b; Tosolini and Jouneau, 2016).

Primed pluripotency is sustained in human by the activation of FGF and TGF β 1 signaling pathways (James et al., 2005; Weinberger et al., 2016). Unlike mouse primed pluripotent stem cells (PPSCs), human PPSCs do not require Leukemia inhibitory factor (LIF) to be stabilized (Smith, 2001). PPSCs are characterized by methylated DNA when compared to NPSCs, have one inactive X chromosome and have a low metabolic activity compared to NPSCs. It should be noted that PPSCs cannot perform oxidative phosphorylation (Guo

et al., 2014; Kilens et al., 2018; Leitch et al., 2013; Mekhoubad et al., 2012; Sahakyan et al., 2017; Smith et al., 2014; Vallot et al., 2017; Yan et al., 2013). Considering chimera contribution, hESCs have been shown to contribute to post implantation mouse embryos (early and late gastrula stage) (Mascetti and Pedersen, 2016). Transcriptomic analysis showed that hPPSCs differ from human pre-implantation epiblast (Blakeley et al., 2015; Yan et al., 2013) and represent post implantation human epiblast (Castel et al., 2020; Kilens et al., 2018).

Formative pluripotency

In 2017, it was hypothesized by A. Smith, a leader in the pluripotency field, that an intermediate state was mandatory between primed and naive PSCs to acquire competence for multi lineage induction (Smith, 2017). This state called formative phase, was derived in vivo in mouse first and then in human (Kinoshita et al., 2021; Rostovskaya et al., 2019; Smith, 2017; Wang et al., 2021; Zhu et al., 2023).

Formative stem cells exhibit notable morphological changes including cell movement and flattening, within 20 hours in adherent culture, starting from hNPSCs (Kalkan et al., 2017). Thus, there is a time interval of 24 hours or longer between the loss of naive pluripotency and lineage priming. This might be crucial for multi-lineage potential. Formative phase has been demonstrated to be longer in primate than in rodent (Nakamura et al., 2016).

In mice, one defining feature of formative epiblast cells is their ability to respond to germline induction, a trait absent in both naive and primed stage epiblast (Ohinata et al., 2009).

Contrary to primed PPSCs, formative cells are cultured without FGF supplementation. FGF supplementation in formative cells induces differentiation into primitive streak like cells expressing T (Rostovskaya et al., 2019) or in mesendoderm and neural lineage (while more slowly than mEpiSCs).

Extended pluripotency

In 2017, several teams reported the identification of an additional state of pluripotency in both mouse and human, able to contribute to pre implantation interspecies chimeras. These cells are termed extended pluripotent stem cells (EPSCs) (Yang et al., 2017a, 2017b). According to the initial publications, these cells represent a stage corresponding to cleavage stage of the mouse embryo (Yang et al., 2017a). In human, the initial publication described EPSCs as distinct from both NPSCs and PPSCs, expressing genes typically found from the oocyte to the morula stage (Yang et al., 2017b). These 2 defining characteristics along with the expression of pluripotency marker genes led to the assumption that these cells represent a more potent form of pluripotency.

In human, extended pluripotency is maintained by the inhibition of ERK, GSK3 β , histamine and muscarinic receptor and PARP1, in presence of LIF: LCDM culture medium (Yang et al., 2017b).

Since the initial publication of mouse and human EPSCs, subsequent reports showed that EPSCs can convert into trophoblast stem cells and generate blastoids, though at a lower efficiency than NPSCs (Castel et al., 2020; Fan et al., 2021; Gao et al., 2019; Sozen et al., 2021). This supports the notion that EPSCs possess a distinct pluripotency stage, different from PPSCs. However, concerns have been raised regarding the developmental stage of EPSCs. Studies in both mice and humans have indicated that EPSCs resemble later pluripotent EPI rather than earlier totipotent developmental stage (Posfai et al., 2021). In fact, study in human shown significant transcriptomic similarities between EPSCs and PPSCs (Castel et al., 2020). Mouse EPSCs were even described as lacking extensive extended potential compared to embryonic day 4,5 – embryonic day 5,5 EPI cells (Posfai et al., 2021). Moreover, the chimera experiments performed in the initial studies have raised questions, as it remains unclear whether the injected EPSCs truly divide and integrate into the host (Aksoy et al., 2021; Posfai et al., 2021). Additionally, these cells have not been comprehensively benchmarked using established naïve pluripotency hallmarks (Figure 8).

While EPSCs hold promise and open up intriguing new questions for research, their full potential for peri implantation studies awaits systematic and rigorous characterization.

	Comparison with hNPSC	Comparison with hPPSC	Comparison with embryo	Marker genes used	Epigenetic characterization	Metabolic characterisation	trophoblast potential	Chimera potential	Blastoid potential
Castel et al, 2020	✓	✓	✓	OCT4, NANOG SOX2	✗	✗	✓	✗	✗
Yang et al, 2017 (Deng team)	✗	✓	Comparison of DE genes between hPPSC and hEPSC to integrated datasets of human embryo	FOXA2, aSMAD, LHX5, OCT4, KLF4	✗	✗	✗	✓	✗
Tan et al, 2021	✗	✗	✗	✗	✗	✗	✗	✓	✗
Gao et al, 2020	✓	✓	✗	OCT4, NANOG, SOX2, REX1, SALL4	histone modifications	✗	✓	✓	✗
Sozen et al, 2021	✗	✗	✗	✗	✗	✗	✗	✗	✓
Fan et al, 2021	✗	✗	✗	✗	✗	✗	✗	✗	✓
Liu et al, 2021	✗	✓	✗	OCT4, NANOG, SOX2, KLF4	✗	✗	✓	✓	✗
Zheng et al, 2021	✗	✓	✗	OCT4, NANOG	H3K27me3 DNMT3L	✗	✓	✓	✗

Figure 8 : Comparison of validation of human extended pluripotent stem cell models. Red cross means that this criteria has not been tested in the cited article. Green tick means that this criteria has been tested in this article. Abbreviations : hNPSC human naive pluripotent stem cells; hPPSC human primed pluripotent stem cell; DE differentially expressed

How are pluripotent stem cells generated?

Derivation

Different methods are now available to generate pluripotent stem cells. To provide a chronological overview, Thomson derived the first embryonic stem cells from a human embryo in 1998, marking a significant milestone (17 years after similar work in mice by Evans and Kaufman 1981 and Martin 1981).

In 1981, two teams reported the derivation of mouse embryonic stem cells, one from either entire blastocysts (Evans and Kaufman, 1981) and the other from the inner cell mass (ICM) of mouse embryos (Martin, 1981). These cells' differentiation capacities were validated through teratoma formation and later, by generating chimeras (Bradley et al., 1984). It took 17 more years to achieve similar success in human (Thomson et al., 1998) with the establishment of 5 different stem cell lines (H1, H7, H9, H13, H14) from IVF embryos. Human ESCs are cultured on a layer of mouse embryonic fibroblast (MEF), in a culture media composed of serum and LIF. Notably, H1 and H9 cell lines are still widely used to this day all around the world and serve as the gold standard for pluripotent stem cells. These cells were validated through teratoma formation in mice.

Unlike in mice, LIF signaling is not sufficient to maintain human pluripotency and BMP4 induces differentiation of human pluripotent stem cells into the trophoblast lineage (Amita et al., 2013; Bernardo et al., 2011; Horii et al., 2016; Li et al., 2013; Xu et al., 2002). Human ESCs require the activation of Activin/Nodal and the FGF pathway to maintain pluripotency (Beattie et al., 2005; James et al., 2005; Vallier et al., 2005). PSCs derived in 1998 were later described to correspond to PPSCs, modeling post implantation epiblast.

Derivation of NPSCs from human embryos was achieved in 2014 (Gafni et al., 2013; Guo et al., 2016; Theunissen et al., 2014; Ware et al., 2014). Whether they are in a naive or primed state, derivation of ESCs involves the removal of the zona pellucida to access to

pluripotent epiblast cells within the ICM. The TE can be removed either through immunosurgery or laser dissection (Lammers et al., 2023; Solter and Knowles, 1975).

For PPSCs ESCs, ICM cells can be placed on a MEF layer intact. However, for NPSCs, it is necessary to dissociate cells before plating them on a MEF layer.

It is important to note that success rates of ESCs line derivation are relatively low, with approximately 2 in 10 embryos resulting in a successful ESC line (Ström et al., 2007, 2010) (Figure 9).

Reprogramming

Several groundbreaking discoveries paved the way for the development of somatic reprogramming into pluripotent stem cells. While it is accepted now that a cell can be “rewired” into other types of cells, it was previously thought that cells genetic material was irreversibly altered as they begin to differentiate (Briggs and King, 1952). John Gurdon frog cloning experiment in 1958 demonstrated that cells of the adult organism are genetically identical to the fertilized egg from which they originate (Gurdon et al., 1958). This experiment showed that nucleus from a differentiated cell can replace the nucleus of an egg and still give rise to an adult organism (Gurdon et al., 1958). The cloning of mammals was then achieved, first in 1984, with the cloning of a blastomere’s nucleus into a sheep oocyte and the successful live birth of 3 out of 4 lambs from the transferred blastocysts (Willadsen, 1986). In 1995, cloning of in vitro cultured cells was achieved, leading to the live birth of 2 healthy and fertile individuals: Megan and Morag sheeps (Campbell et al., 1996). The most iconic milestone in adult cell cloning occurred in 1996 when an adult mammary gland cell was cloned into a sheep oocyte, resulting in the live birth of fertile Dolly the Sheep (Wilmut et al., 1997). One unexpected development of adult cell reprogramming has been the development of companies such as ViaGen which can clone pets and animals for a considerable cost. However, the most significant aspect of mammalian adult cell reprogramming, relevant to this PhD, is the proof of concept that differentiated cells can be reprogrammed to an embryonic-like state. In this context,

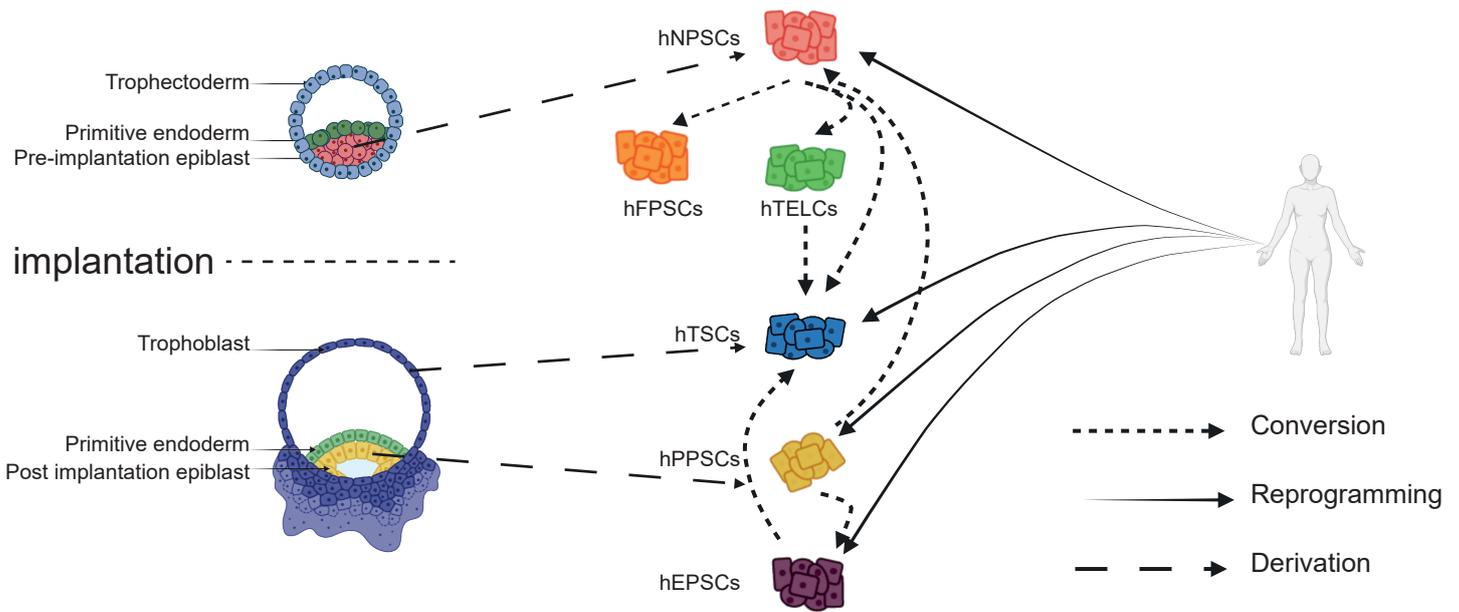


Figure 9 : Ways to generate peri-implantation stem cells

Peri-implantation stem cells can be generated through conversion from another state, reprogramming from somatic cells or derivation from embryonic lineages.

Abbreviations : hNPSCs human naive pluripotent stem cells; hFPSCs human formative pluripotent stem cells; hTELCs human trophoblast like cells; hPPSCs human primed pluripotent stem cells; hEPSCs human extended pluripotent stem cells

Yamanaka's team began investigating transcription factor to determine if their addition could reprogram somatic cells. By exploring factors important in the maintenance of ESCs identity, they identified a cocktail of 4 transcription factors: Oct4, C-Myc, Klf4 and Sox2 (OKSM factors) that are necessary and sufficient to generate pluripotent stem cells from somatic cells (Takahashi and Yamanaka, 2006). This groundbreaking study was soon replicated in human by several groups (Lowry et al., 2008; Park et al., 2007; Takahashi et al., 2007), using the same OKSM factors or alternative combinations such as OCT4, SOX2, NANOG and LIN28 (Yu et al., 2007). These achievements culminated in the Nobel prize attribution to Yamanaka and Gurdon in 2012. These new pluripotent cell lines satisfy all the original criteria proposed for human ESCs (Thomson et al., 1998), except that they are not derived from embryos. Over time, many protocols were established to generate iPSCs, such as retroviruses, lentiviruses, sendai viruses, episomes, transposons, RNA or protein carriers. iPSCs offer an ethical alternative to embryonic stem cells and enable the generation of stem cells with specific genetic background, making them crucial for regenerative medicine and the study of specific genetic diseases.

From 2017, OKSM factors were used to derive NPSCs from somatic cells (Giulitti et al., 2019; Kilens et al., 2018; Liu et al., 2017). This method of generating NPSCs also addresses the ethical issues associated with naïve ESCs. The development of this technic also paved the way to generate isogenic primed and naïve PSCs.

In 2019, OSKM factors were used to generate EPSCs from somatic cells (Castel et al., 2020; Yang et al., 2017a) (Figure 9).

Conversion

Somatic to primed pluripotency

Since the development of reprogramming protocols, research teams have been trying to reduce the viral load associated with these methods through the use of chemical

cocktails. In 2022, a team reported the successful resetting of fibroblast to primed pluripotent stem cells (Guan et al., 2022).

This small molecule cocktail was composed of the sequential use of CHIR99021, 616,452, TTNPB, Y27632, ABT869, SAG, JNKIN8, 5-azacytidine, tranylcypromine, valproic acid, DZNep, EPZ004777, UNC0379, and PD0325901. While the resetting takes 5 to 6 weeks and more comprehensive analysis and optimization are required for this protocol, it is still a significant finding in the pluripotency field.

Interestingly, the chemical cocktail inhibits DNA methylation, histone acetylation, histone methylation and other major cell signaling pathway which in tune provides an open chromatin environment, enabling the erasure of cell's history. Compared to induced pluripotent stem cells, the method minimizes the risk of tumorigenesis associated with the viral methods and opens new avenues in regenerative medicine.

Primed pluripotency to Naïve pluripotency (reset)

Several groups have reported the conversion of hNPSCs from hPPSCs through chemical resetting (Guo et al., 2017; Szczerbinska et al., 2019). There are also alternative protocols that convert PPSCs into NPSCs; however, these involve the overexpression of transgenes (Qin et al., 2016) and will not be discussed here.

Chemical resetting (Austin lab protocol, (Guo et al., 2017)) from hPPSCs to hNPSCs involves the use of ERK inhibitor, HDAC inhibitor and LIF for the initial 3 days before transitioning in ERK inhibitor, Gö and LIF for 6 days and finally transitioned to the naïve maintenance medium. This approach offers an effective means of converting hPPSCs into hNPSCs without the need for transgene overexpression.

Primed pluripotency to extended pluripotency

The initial publication on hEPSCs described a protocol for converting hPPSCs into hEPSCs using the previously mentioned LCDM medium (Yang et al., 2017a).

Trophoblast stem cells

The trophoblast plays a central role in the formation of the maternal-fetal interface and abnormalities in the trophoblast layer are likely to have dramatic consequences on placental development. These disruptions, in turn, can lead to post-natal outcomes and potentially result in chronic disease in adulthood (Burton et al., 2016). However, our understanding of the nature and prevalence of trophoblast disorder, as well as the connection of trophoblast with placental diseases such as preeclampsia, fetal growth restriction, miscarriage or choriocarcinomas remains limited. In the context of IVF, it is important to note that trophoblast disorder could affect implantation and consequently the pregnancy outcome. Therefore, gaining a better understanding of trophoblast is essential to better comprehend human development from implantation to latter stages. Since the derivation of mouse trophoblast stem cells (mTSCs) in 1998 (Tanaka et al., 1998), isolation of human TSCs has been a significant challenge in the field of developmental biology and stem cell research. Arima's lab was able in 2018 to derive hTSCs (Okoe et al., 2018). hTSCs are invaluable tools to help in the understanding of trophoblast development and trophoblast misfunctions.

Trophoblast in vivo

In vivo, the trophoblast performs many essential functions, including facilitating the physiological adaptation of the mother to immunological acceptance, providing nourishment and support to the developing embryo, delivering nutrients and oxygen to the developing fetus and clearing of waste products. To accomplish all these tasks, placental villi undergo dynamic changes throughout gestation.

The first specification in the embryo segregates the epiblast from the trophectoderm (TE) layer around day 4-5 after fertilization. The trophectoderm serve as a precursor to all trophoblast cells and thus of the placenta. The interaction between the TE, particularly the region adjacent to the epiblast known as the polar TE and the maternal uterine epithelium

enables implantation around day 6-7. Following implantation, placenta starts to develop into a complex organ composed of several distinct cell types.

First, trophoblast cells generate cytotrophoblasts, considered as *in vivo* trophoblast stem cells. Cytotrophoblast then differentiate into multinucleated syncytiotrophoblast (ST) which forms an interface layer between the mother and the fetus, for nutrient transport and gas exchange. ST cells constitute the outer layer of the trophoblast epithelium and also serve to protect the fetus from pathogens. Additionally, cytotrophoblasts differentiate into invasive extra villous trophoblast cells (EVT). These cells establish tissue connection within the developing placental-uterine interface, participate in the remodeling of spiral arteries and participate to the establishment of immune tolerance between the developing conceptus and the mother through a unique pattern of histocompatibility leukocyte antigen (HLA) expression, notably HLA-G (Knöfler et al., 2019; Turco and Moffett, 2019).

In 2016, Lee and colleagues proposed a list of criteria to define first trimester trophoblast (Lee et al., 2016). This study proposed a consensus in the field of the molecular definition of the first trimester trophoblast (Figure 10)

The criteria proposed by Lee et al include:

- The expression of KRT7, GATA3 and TFAP2C as mononuclear trophoblast;
- Hypomethylation of ELF5 promoter in human trophoblast cells;
- Expression of C19MC miRNA cluster;
- Absence of HLA-A or -B or class II molecules in any trophoblast cells
- Lack of expression of any HLA class I molecules in cytotrophoblasts
- Expression of HLA-C, -G and -E in EVT cells

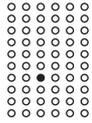
Among of these criteria, the relevance of ELF5 is less clear as its expression pattern is distinct in mouse and human (Gamage et al., 2018).

A

Expression of C19MC miRNA cluster



Hypomethylation of Elf5 promoter



Lee et al criteria for human trophoblast cells

Expression of HLA class I



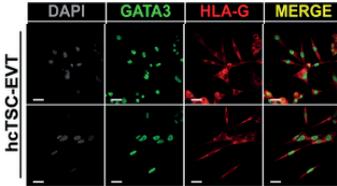
No expression of HLA class II



Expression of TFAP2C, KRT7, GATA3

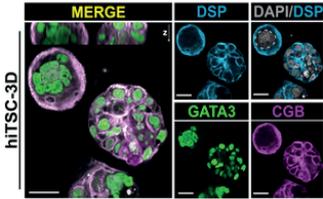
B

Differentiation into extra-villous trophoblast



Additional criteria for human trophoblast stem cells

Differentiation into syncytiotrophoblast



Long term self renewal



Figure 10 : Criteria for human trophoblast cells and human trophoblast stem cells
 (A) Criteria established by Lee et al to define human trophoblast cells
 (B) Additional criteria to define human trophoblast stem cells (immunofluorescences from Castel et al, 2020)

Models of trophoblast

Several cellular models have been employed to study human trophoblast development, each with its advantages and limitations (Figure 11).

HTR8/SV Neo cell lines

HTR8/SV Neo cell lines were established by immortalizing primary extra villous trophoblast (EVT) cells via transfection with a plasmid containing the simian virus 40 large T antigen (SV40). These cells display self-renewing abilities, featured by their capacity to form spheroid bodies (Graham et al., 1993). However, they express the pluripotency markers OCT4 and NANOG. It has been proposed that the transfection has altered the original (extravillous) trophoblast cell (Weber et al., 2013).

BeWO cell line

BeWo cell line is derived from choriocarcinoma, a malignant trophoblastic cancer. It has syncytialization and invasion abilities and can be differentiated into ST- and EVT-like cells. However, it might present abnormal features compared to endogenous trophoblast as it continuously express hCG, a marker of ST fate, while it is supposed to model CT (which do not express hCG). Moreover, any treatment of those cells leads to a decrease in ST-like morphology (decrease in fusion rate), without an accompanying decrease in hCG expression (Orendi et al., 2010; Pattillo and Gey, 1968).

Despite HTR8/SV Neo and BeWo cell lines widespread usage in the study of miscarriages and pre-eclampsia, their malignant origin is a major limitation for the study of placental physiology. On top of this, they represent an advanced developmental timing. Thus, to study early trophoblast differentiation and development, other models have been developed.

	<i>Lee et al Markers (RNA)</i>	<i>Lee et al Markers (protein)</i>	<i>Cellular controls</i>	<i>C19 miRNA cluster</i>	<i>Elf5 promoter methylation</i>	<i>ST differentiation</i>	<i>EVT differentiation</i>	<i>Long term in vitro self renewal</i>
Castel et al, 2020 hTSC	GATA3; KRT7	GATA3	hNPSC, hEPSC, hPPSC, Okae TSC	✗	✗	CGA, CGB, SDC1 (RNA), DSP, GATA3, CGB (protein), morphology	HLA-G, MMP2 ASCL2 (RNA), GATA3, HLA-G (protein) morphology	✓
Cinkorpumin et al, 2020 hTSC	✗	TFAP2C; KRT7	primary placental cells, Okae TSC	✗	✗	hCG	HL1GCB7, GCM1 (RNA), ITGA1, HLAG (protein)	✓
Dong et al, 2020 hTSC	GATA3; KRT7; TFAP2C	KRT7	hNPSC	✗	✗	CGB, SDC1 TEAD4 (RNA), SDC1, CGB (protein), morphology	MMP2 (RNA), SSC-A, HLA-G (protein), morphology	✓
Gao et al, 2020 hTSC	GATA3; TFAP2C; KRT7	✗	hEPSC	✗	✓	hCG	✗	✓
Okae et al, 2020 hTSC	✗	GATA3	primary placental cells (CT, EVT, ST)	✓	✓	ITGA6, VIM, HLA-G CDH1, SDC1, CGB (protein), morphology	ITGA6, VIM, HLA-G CDH1, SDC1, CGB (protein), morphology	✓
Xu et al, 2002 BMP4 treated hPPSC	GATA3; TFAP2C	✗	hPPSC, HELA cells	✗	✗	hCG	HLA-G	✗
Amita et al, 2013 BAP treated hPPSC	✗	KRT7	hPPSC, BMP4 treated cells	✗	✗	hCG	HLA-G	✗
Primary placental cells (CT)	GATA3; TFAP2C; KRT7	GATA3; TFAP2C; KRT7	/	✓	✓	/	/	✗
Pattillo and Gey, 1968 BeWO	✗	✗	/	✗	✗	CGB, ERVWEI, fusion	/	✓
Graham et al, 1993 HTR8/SV Neo	✗	✗	parental HTR8 cells	✗	✗	/	morphology, invasion	✓

Figure 11 : Comparison of validation of human trophoblast models according to Lee et al criteria

Red cross means that this criteria has not been tested in the cited article.

Green tick means that this criteria has been tested in this article.

Abbreviations : hTSC human trophoblast stem cells; CT cytotrophoblast; ST syncytiotrophoblast; EVT extra-villous trophoblast; BAP BMP4, A83-01, PD032591 treatment

Differentiation of human embryonic stem cells (hESC) to trophoblast-like cells by BMP4 treatment

BMP4 differentiation of hESCs to trophoblast like cells was first reported by Xu et al (Xu et al., 2002). Indeed, after 7 days of BMP4 treatment, TFAP2C and GATA3 are expressed, among other trophoblast marker genes and levels of expression of marker of pluripotency like POU5F1 are decreased (Xu et al., 2002). However, the relevance of this model has been debated as it leads to a gene expression profile that does not resemble primary trophoblast, with an up-regulation of mesoderm markers (Amita et al., 2013; Horii et al., 2016; Li et al., 2013).

New culture conditions have been proposed such as the MEF BAP treatment, which consists in MEF conditioned media, BMP4, A83-01 and PD173074, inhibitors of respectively TGF- β and FGFR signaling pathways (Amita et al., 2013). After this treatment, hESCs differentiate into a subset of invasive placental-like cells, and may provide good models for studying trophoblast functions, including invasion. This treatment suppresses the up-regulation of mesoderm markers but a difference in global gene expression still remains between the transitioned hESCs and primary trophoblasts. Most importantly, under these conditions, those cells rapidly enter terminal differentiation and cannot be propagated, limiting their utility. On top of this, the identity of those trophoblast-like cells has been put into question (Bernardo et al., 2011), as they do not fulfil established criteria for trophoblast identity (Lee et al., 2016; Seetharam et al., 2022).

Interestingly, BMP pathway is not involved in trophectoderm specification in the human blastocyst (Paepe et al., 2019) and one group reported the differentiation of hPPSCs into amnion like cells under BMP treatment (Zheng et al., 2019). Thus, it is still debated whether BMP-treated cells correspond to bona fide trophoblast like cells or to amnion like cells.

Recently, BMP4 differentiation has been used in combination with trophoblast stem cell medium to generate trophoblast stem cells from primed pluripotent stem cells.

Trophoblast stem cells

Mouse trophoblast stem cells (mTSCs) were successfully derived in 1998 (Tanaka et al., 1998). However, lack of knowledge regarding the suitable culture conditions for maintenance of human TSCs led some researcher to question their existence (Chang and Parast, 2017). In 2018, solid proof of the existence of stem-like trophoblast cells in human was reported when Okae et al derived hTSCs from blastocyst and first trimester cytotrophoblast (Okae et al., 2018). Those lines were the first trophoblast lines that met all the criteria defined by Lee et al (Lee et al., 2016) (Figure 10). These hTSCs can long-term self-renew and differentiate into the lineages of the placenta (EVT and ST) in vitro. Moreover, they were validated in vivo by injection into immuno-deficient mice, where they mimicked some key features of trophoblast invasion during implantation (Okae et al., 2018).

Morphologically, hTSCs are flat epithelial-like cells, resembling cobble stone, express key trophoblast proteins, such as GATA3, TFAP2C or KRT7. ELF5 promoter is hypomethylated and the chromosome 19 miRNA cluster C19MC is highly expressed in these cells (Lee et al., 2016; Okae et al., 2018). hTSCs can differentiate into EVT cells expressing HLA-G and ST cells expressing CGB (pregnancy hormone) (Okae et al., 2018) (Figure 10).

Following the derivation of hTSCs from blastocyst and first trimester cytotrophoblast, multiple labs including ours have developed alternative methods to generate hTSCs, including reprogramming using Yamanaka factors (OKSM) and conversion first from hNPSCs and then from hPPSCs, using a variant approach than BMP4 treatment (Castel et al., 2020; Cinkornpumin et al., 2020; Dong et al., 2020; Fogarty et al., 2021; Io et al., 2021; Liu et al., 2020) (Figure 9).

In summary, various cellular models have been employed to study human trophoblast development, each with unique characteristics and limitations. The field continues to evolve with ongoing research efforts to better understand trophoblast biology and development.

TSCs generation methods

In the past 5 years, following the first publication on human TSCs, significant progress has been made on hTSCs biology but also their generation. There are still many controversies in the field towards what is a good model of hTSCs and most importantly, what are the correct controls to determine whether the cells are hTSCs or not. This section provides an overview of the primary methods used for hTSCs derivation, conversion, and reprogramming, along with the associated challenges and controversies.

Derivation

Mouse trophoblast stem cells were derived from mouse blastocyst in 1998 (Tanaka et al., 1998). These cells were cultured in the presence of FGF4 and demonstrated the ability to differentiate into other trophoblast subtypes in vitro in the absence of FGF4. On top of this, mTSCs exclusively contribute to the trophoblast lineage in vivo in chimeras.

However, deriving human TSCs from blastocysts presented a challenge, taking 2 decades to accomplish. In 2018, H.Okae, Pr. Arima's lab successfully derived hTSCs from blastocysts and first trimester cytotrophoblasts. Culture medium for the maintenance of hTSCs is composed of EGF, vitamin C, GSK3 β inhibitor, TGF β inhibitor, HDAC inhibitor and ROCK inhibitor (ASECRiAV medium) (Okae et al., 2018).

Conversion

Following Okae's work, several labs developed conversion methods to generate hTSCs, offering a faster approach that does not involve the destruction of an embryo.

It is still not clear whether we should talk about differentiation or conversion, mostly because we do not know whether pluripotent to trophoblast conversion occurs *in vivo*.

Naïve/Extended to trophoblast

Conversion of hTSCs was successfully established from the conversion of hEPSCs in ASECRiAV medium (Castel et al., 2020; Gao et al., 2019). Shortly after, hNPSCs in 5iLAF medium were converted in hTSCs (Cinkornpumin et al., 2020; Dong et al., 2020). Finally, hNPSCs in T2iLGö were converted in hTSCs (Castel et al., 2020). Finally, more recently, other papers proposed the conversion of hTSCs from hNPSCs but through a trophectoderm (TE) like intermediate (Guo et al., 2021; Io et al., 2021).

Different validation methods and controls have been used in all these articles. It is worth mentioning that Okae's cells (either derived from blastocyst or from first trimester placenta) are used as control for hTSCs in all articles but Gao's paper converting hEPSCs into hTSCs. When considering the criteria for first trimester trophoblast established by Lee et al, with the exception of Okae's paper, not one hTSCs was checked according to all the criteria. Indeed, no article except for Okae's and Liu's checked for miRNA C19MC complex expression; no article checked for ELF5 promoter methylation except for Gao's and Okae's. While Okae, Castel and Dong checked the ability of their hTSCs to differentiate into EVT and ST on the morphology, RNA and protein level, Gao and Cinkornpumin only looked at hCG expression for ST differentiation, thus missing the information of the multinucleation. Only Castel study compared hTSCs with peri-implantation embryo datasets, thus staging hTSCs to NR2F2+ day 8-10 cytotrophoblasts. When considering the markers used, each article used a combination of markers measured at the RNA level and at the protein level.

Primed to trophoblast

It was reported in the Castel et al that direct conversion from hPPSCs to hTSCs in ASECRiAV medium was not successful. However, recent papers reported the conversion of hPPSCs into hTSCs (Jang et al., 2022; Soncin et al., 2022; Viukov et al., 2022; Wei et al., 2021; Zorzan et al., 2023). Contrary to hNPSCs and hEPSCs conversion into hTSCs, all these studies first use an initial treatment with either BMP4 or MEK inhibitor, LIF and HDAC inhibitor before converting treated hPPSCs in ASECRiAV medium (Figure 12).

When considering these results and the transient trophoblast differentiation generated by BMP4 or BAP treatment (Amita et al., 2013; Horii et al., 2016; Li et al., 2013; Xu et al., 2002), we can hypothesize that the transient trophoblast differentiation induced through BMP4 treatment is made durable through the ASECRiAV culture medium. It could be interesting to determine whether the intermediate created through BMP4 treatment actually corresponds to amnion like cells or to trophoblast cells or whether it could be a more potent intermediate able to give rise to both amnion and trophoblast. Soncin's analysis showed that hPPSCs treated with BMP4 directly convert to a TE-like fate with the induction of only one naïve marker (Soncin et al., 2022).

When considering Martello's team results (Zorzan et al., 2023), the combination of HDAC inhibitor, MEK inhibitor and LIF corresponds to the first step of resetting hPPSCs to hNPSCs before transitioning cells in a classic naive pluripotency medium (Rugg-Gunn, 2022). Another lab described the appearance of hNPSCs from 3 days into the resetting protocol (Rugg-Gunn, 2022). Thus, it is difficult to determine whether in the case of this protocol, the cells transitioned in ASECRiAV medium correspond to hNPSCs or an intermediate state between PPSCs and NPSCs.

Globally, these results indicate that the potential to engage into the trophoblast lineage is common to all hPSCs.

Starting culture medium

Protocol

hTSC comparison in culture experiment

hTSC comparison in transcriptomic experiment

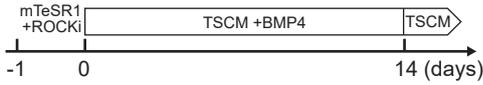
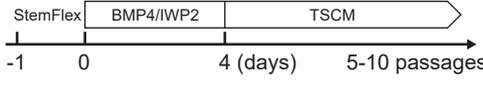
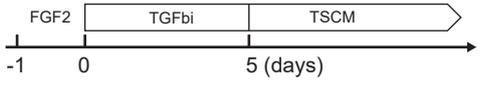
<p>Jang et al, 2022 (PNAS)</p>	<p>mTeRS1 feeder free</p>		<p>hTSC (Okae)</p>	<p>hTSC (Okae)</p>
<p>Soncini et al, 2022 (Stem cell rep)</p>	<p>StemFlex feeder free</p>		<p>Primary trophoblast line</p>	<p>Primary TSC line</p>
<p>Wei et al, 2021 (Science advance)</p>	<p>mTeRS1 feeder free</p>		<p>hTSC (Okae)</p>	<p>hTSC (Okae)</p>
<p>Viukov et al, 2022 (Stem cell rep)</p>	<p>FGF2 feeder</p>		<p>hTSC (Okae)</p>	<p>hTSC (Okae)</p>
<p>Zorzan et al, 2023 (EMBO rep)</p>	<p>E8</p>		<p>Naive converted hTSC (same study)</p>	<p>hTSC (Okae)</p>

Figure 12 : Comparison of human primed pluripotent stem cells to human trophoblast stem cells conversion protocols

Abbreviations : hTSC human trophoblast stem cells; TSCM trophoblast stem cells culture medium

Reprogramming

Alongside the generation of conversion protocols, protocols to reprogram somatic cells into hTSCs using OCT4, KLF4, SOX2, MYC factors also emerged (Castel et al., 2020; Liu et al., 2020). This enabled the discovery that during reprogramming, primed-like, naive-like and TE-like cells coexist, without exposure to any pluripotent or trophoblastic media. Direct reprogramming of fibroblasts into hTSCs holds promise for studying placental defects and generating patient-specific hTSCs for research on trophoblast dysfunction. Recently, a new protocol using *GATA2*, *GATA3*, *TFAP2C* and *KLF5* and *MYC* factors was proposed. This expands the techniques used to create cellular models of progenitor cells in the early human placenta (Fogarty et al., 2021).

In summary, significant progress has been made in recent years in the field of hTSCs, with various methods for their generation and characterization. However, challenges and controversies persist, particularly regarding the criteria for validating hTSCs identity and understanding the intermediate states during conversion and reprogramming processes.

Modeling using human stem cells models

Stem cells have emerged as crucial tools to elucidate developmental processes, conducting in depth physio-pathological investigations and undertaking studies in the field of regenerative medicine. Here, the focus will first insist on modeling with hPSCs, Yamanaka's comprehensive review in 2020 has extensively explored the pivotal role of hPSCs in the domains of regenerative medicine and physiopathology, thus it will not be developed here. Consequently, the modeling capacities of hTSCs and their potential to significantly influence the understanding of various physio-pathological mechanisms will be developed.

Modeling using hPPSC

Human pluripotent stem cells (hPSCs) offer a versatile research platform for studying various aspects of peri-implantation development, ranging from the equivalent of 8-cell stage embryos to post-implantation epiblasts. This provides researchers with invaluable tools to investigate early events in human embryonic development, epigenetic changes, X chromosome inactivation (XCI), transposable element regulation, metabolic characteristics, embryonic lineage differentiation potential, and extra-embryonic lineage differentiation potential.

Female hPSCs, provide an in vitro model to study human XCI and X chromosome erosion. Indeed, hPPSCs have one inactive X chromosome but exhibit X Chromosome erosion during long-term culture, meaning that genes from the inactive X chromosome start to be re-expressed (Geens and Chuva De Sousa Lopes, 2017; Mekhoubad et al., 2012; Wutz, 2012). hNPSCs, on the other hand, have two active X chromosomes and can offer insights into the early stages of XCI (Sahakyan et al., 2017; Vallot et al., 2017). Optimizing hPSCs models and conversion techniques can model XCI processes in vitro.

Together, the culture conditions of human naive PSCs might still require refinement to precisely correspond to the X chromosome phenotype of human epiblast cells. Nevertheless, their present status has already facilitated a deeper comprehension of the mechanisms behind X chromosome reactivation (Panda et al., 2020). With the ability to exhibit two activated X chromosomes, hNPSCs serve as cellular models for examining X-linked diseases and for the advancement of regenerative medicine.

Recent studies highlight the importance of metabolism in pluripotency regulation. Naive hPSCs exhibit distinct metabolic characteristics compared to primed hPSCs (Guo et al., 2016; Sperber et al., 2015; Takashima et al., 2014). Investigating these metabolic differences can shed light on their roles in regulating self-renewal and lineage commitment of stem cells, potentially mirroring metabolic dynamics during in vivo human development, which could in turn help to optimize culture conditions of human embryos to adapt to their metabolic activity.

Modelling placental development using hTSCs and using them to understand principal placentation illnesses

The placenta, often referred to as the most enigmatic organ in the human body, plays a pivotal role in pregnancy and fetal development. While animal studies on placental development have provided valuable insights, they cannot fully replicate human trophoblast development (see section “Exploring implantation across mammalian species”). Moreover, human primary placental cells are highly variable (depending on the placental source) and access to early samples is limited, making the study of trophoblasts development challenging. However, the recent establishment of suitable culture conditions for TSCs has opened new avenues in the field, allowing researchers to explore human trophoblast development in greater detail (Castel et al., 2020; Cinkornpumin et al., 2020; Dong et al., 2020; Liu et al., 2020; Okae et al., 2018). On top of this, recent protocol enabling the culture of 3D trophoblastic organoids enable the study of several trophoblastic cell types in parallel in a 3D environment (Karvas et al., 2022).

Both hTSCs and 3D trophoblastic organoids will be instrumental in understanding mechanisms of trophoblastic maturation and differentiation. However, they can also be used to better understand principal placentation illnesses.

Pre-eclampsia

As one of the major causes of extreme prematurity (20% of preterm birth before 32 weeks of gestation) and the cause of 76 000 maternal death worldwide each year, pre-eclampsia is one of the most frequent and severe pregnancy complications. Pre-eclampsia is characterized by a high blood pressure and proteinuria that can start from the end of the first gestational trimester in human, possibly yielding to maternal complications and pregnancy loss (nhs., 2018). It has been described that failure of invasion of maternal tissues by the EVT and incomplete syncytialisation by ST cells are causing the maternal endothelial dysfunction in this pathology (Gauster et al., 2009; Redman and Sargent, 2005). Using hTSCs and 3D trophoblastic organoids would help to understand the mechanisms leading to pre-eclampsia. One pending challenge will be to integrate trophoblastic models with endometrial and vascular cells in order to model the complexity of trophoblastic micro-environment. All this may lead to improved treatments for patients with pregnancy complications.

Infectious diseases

Infections during pregnancy can have detrimental effects on the placenta and the developing fetus. Certain viruses, categorized as TORCH infections (congenital infections of toxoplasmosis, others (Syphilis, Hepatitis B), rubella, Cytomegalovirus (CMV), and herpes simplex.), can breach the trophoblast barrier and trigger an interferon response in the placenta (reviewed in Megli and Coyne, 2022). Trophoblast stem cells help researchers investigate how these infections affect the placenta's functionality and its

impact on fetal development. For instance, understanding how viruses like Cytomegalovirus and Herpes virus interact with trophoblasts can provide insights into preventing congenital infections. Impact of Zika Virus and SARS Cov2 infection in trophoblast organoids have already been explored, enabling to conclude on the most susceptible cells to be infected by these viruses (mature for SARS-CoV-2, early human trophoblast cells for Zika Virus) (Karvas et al., 2022). This gives a platform to investigate placental barrier and could in term lead to increase knowledge and better treatment of infections during pregnancy.

Placental tumors

Placental tumors, although rare, present significant clinical challenges. These tumors, including choriocarcinomas, arise from trophoblast cells undergoing malignant transformation (reviewed in Al-Riyami et al., 2013). Studying trophoblast stem cells could help to better understand mechanisms of placental tumorigenesis.

In conclusion, trophoblast stem cells offer a versatile and powerful platform for investigating placental diseases and disorders that impact pregnancy outcomes. The knowledge gained from these studies not only advances our understanding of placental biology but also holds the potential to improve clinical approaches for managing and treating complications associated with pregnancy and fetal development.

3D models of human peri-implantation embryo

In recent years, new approaches have been developed to study 3D structure and tissue interactions: organoids. These 3D lab-grown masses of cells model some functions of organs.

The development of 3D models for studying human peri-implantation embryos represents a significant breakthrough in reproductive biology and developmental research. These models provide a unique opportunity to investigate early human development, both pre- and post-implantation, which was previously inaccessible due to ethical and technical constraints.

3D trophoblastic organoids have been developed recently (Karvas et al., 2022; Turco et al., 2018) and show comparable tissue architecture, placental hormone secretion, and capacity for long-term self-renewal as primary trophoblast organoids. They provide a methodology to model the impact of disease-associated mutations in a 3D microenvironment that reflects the cellular diversity of the first-trimester placenta.

Recently, new advances in the field enabled the generation of 3D models of human pre and post implantation embryo: blastoids (Fan et al., 2021; Kagawa et al., 2021; Liu et al., 2021c; Sozen et al., 2021; Yanagida et al., 2021). These 3D structures resemble on the transcriptomic level human embryos around implantation. They also express key markers of human peri implantation development.

Multiple protocols have documented how to generate blastoids from hNPSCs, starting from PXGL or 5iLAF culture media. However, the efficiency of blastoid formation differs significantly between protocols starting with cells cultured in PXGL or 5iLAF, with over 80% of blastoid formation in PXGL and less than 10% for 5iLAF (Kagawa et al., 2021; Yu et al., 2021). Additionally, only Kagawa et al. demonstrated advanced culture techniques using endometrial cells so far. Following Kagawa et al protocol, pre-implantation blastoids are formed via the aggregation of NPSCs for 96 hours. After 96 hours, the aggregates start to cavitate and the cells start to express markers of the different lineages in an organized manner. The external cells express trophoblast markers such as GATA3 and CDX2 while the internal cells express EPI (NANOG, OCT4) and PrE (GATA4, KLF17) markers. These blastoids can pseudo implant on endometrial cells and start expressing the pregnancy hormone hCG (Kagawa et al., 2021).

Other groups have described the generation of post implantation models and the emergence of structures resembling the 2nd week of development (Oldak et al., 2023;

Pedroza et al., 2023; Weatherbee et al., 2023). As described in a recent news and views on this subject (Moris, 2023): "Pedroza *et al.* aggregated these cells (NPSC) under specific chemical conditions to allow them to spontaneously become other cell types. Weatherbee *et al.* forced higher than normal expression of a particular set of genes so that cells were biased towards becoming the different cell types, before aggregating the cells to allow an embryo-like structure to form. Oldak *et al.* identified chemical cocktails that allowed the naive cells to assume the identity of two supporting cell types, before combining these with naive cells in defined ratios" (Figure 13).

Many studies are now converging toward the generation of integrated models enabling to model both embryo and its environment to go further in our understanding of peri-implantation development.

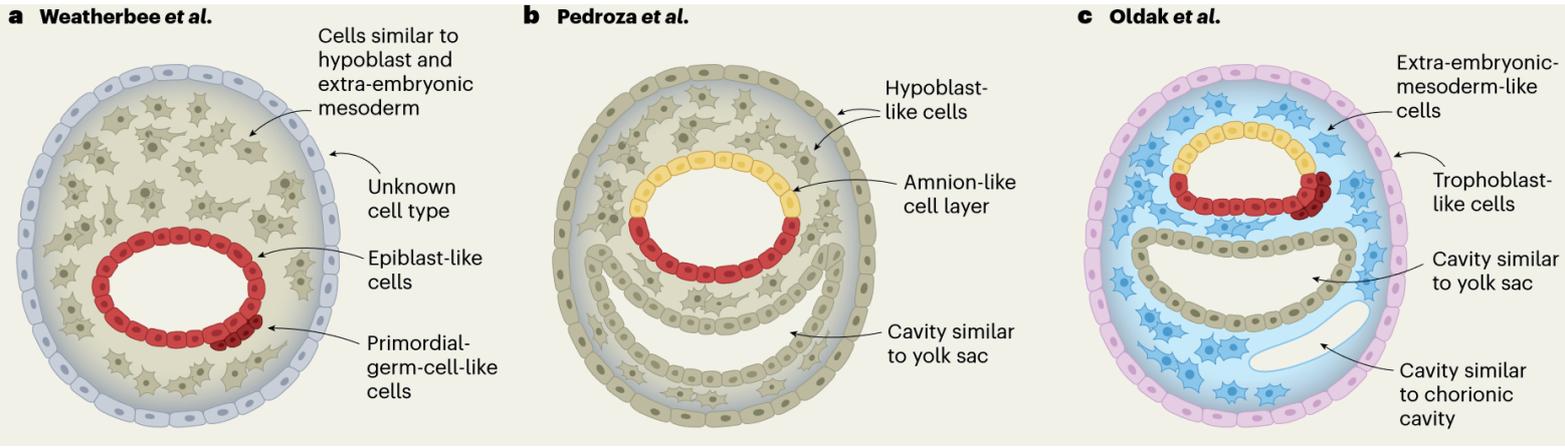


Figure 13 : Stem cell models of human post-implantation development

From Moris, 2023

Human-embryo models were created by Weatherbee et al. (a), Pedroza et al. (b) and Oldak et al. (c) from naive human stem cells, which, under the correct conditions, self-organize to form structures with features that resemble those of a real embryo at days 13 to 14 (not shown). Each model contains a layer of cells that is similar to the epiblast (which will form the embryo proper and, later, the fetus), and two of the models (a, c) contain cells that resemble the primordial germ cells (which will form sperm or eggs). The three models have various representations of supporting cell layers, such as the amnion (which will form the amniotic sac), hypoblast (which surrounds the yolk sac), extra-embryonic mesoderm (which forms the chorionic cavity that later surrounds the growing embryo) and trophoblast (which will form part of the placenta). Not all features of the models are shown, and some features vary within each model. Illustrations were adapted from individual micrographs from Fig. 3h of ref. 3, Fig. 3e of ref. 2 and Fig. 4g of ref. 1.

Ethics and regulation linked to embryonic stem cell models

Guidelines ISSCR

When researching on pubmed using the search terms “human pluripotent stem cells”, the result by year section indicates that the number of articles related to this topic peaked in 2021 with a total of 4,110 articles. This indicates that human pluripotent stem cells continue to be a current and active research field.

To establish standardized practices within the Stem cell community, the international society for stem cells research (ISSCR) has issued recommendations, aimed at defining minimum characterization and reporting criteria. These guidelines are intended for scientists, students and technicians in basic research laboratories working with human stem cells. The goal is to enhance rigor and reproducibility of stem cell research. The importance of universally accepted quality standards is emphasized in order to elevate the credibility and reliability of stem cell-related studies.

As outlined in the ISSCR Standards document, the key recommendations include:

- Generation of master cell bank before any experiments or distribution, at the lowest passage possible.
- Clear authentication and identification of cell lines
- Transgene free and micro-organism free: Cell lines and master and working banks should be transgene free (when applicable) and not contaminated with any micro-organisms. This should be checked regularly, as well as the acquisition of culture-acquired genetic changes.
- Xenograft assay are not required to check for pluripotency.
- Evaluation of developmental state, undifferentiated status and differentiation potential should be performed.
- Consideration of variability of cell lines should be considered in the experimental design (multiple independent clones from isogenic lines but also different genetic

backgrounds). Researchers should clearly elucidate their experimental designs in their publications

These criteria have gained widespread acceptance within the stem cell scientific community and are primarily applied in the context of primed pluripotency. However, we believe that these criteria should be adapted and published for all peri-implantation stem cell models to ensure comprehensive and standardized practices.

Consent

Consent forms that were collected for the derivation of embryonic cell lines back in 1998, with H9 and H1 being among the most commonly used cell lines, are no longer suitable for contemporary research needs. New techniques, such as those involving blastoids, necessitate the development of updated consent forms to align with current research requirements and ethical standards.

Regulation (Fr)

Following the revision of bioethics laws in 2021 (loi du 2 août 2021 relative à la bioéthique, 2021), human embryonic stem cells (hESCs) in France are now subject to a declarative regime, as opposed to the previous authorization regime. The Agence de la Biomédecine holds the authority to oppose the use of hESCs in research if the research lacks a medical purpose, is deemed scientifically irrelevant, or is found to be ethically problematic. These regulatory changes reflect the evolving landscape of stem cell research and its ethical considerations within France.

RESULTS

Context of my work

When I first started this project in January 2020, during my 6 months master internship, we only had access to 2D cellular models of human peri-implantation development. In the lab, we had previously published a paper on isogenic reprogramming fibroblasts into hNPSCs and hPPSCs (Kilens et al., 2018). We were on the verge of publishing our study on reprogramming fibroblast and converting hNPSC into hTSCs (Castel et al., 2020). hEPSCs model were generated by reprogramming at the same time than induced hTSCs, but remained uncharacterized (Gao et al., 2019; Liu et al., 2021c). Although hallmarks of naive and primed pluripotency were described, there were no established guidelines considering the banking and use of those cells yet. Additionally, human blastoids were not yet developed.

Considering the array of available models, it was imperative to characterize them before proceeding with more mechanistic studies. I initiated this project during my Master 2 internship but was quickly interrupted by Covid-19 pandemic. I stopped all experiments until the beginning of my PhD mid-September when I could resume my work only to be stopped once more by the 2nd French lockdown, from November to December 2020. I had to stop all experiments again as there were limitations on the number of people authorized to conduct experiment in each lab. During this time away from the lab, I wrote and submitted a protocol paper on hNPSCs reprogramming from somatic cells (Onfray et al., 2022).

I was able to resume in January 2021; however, plastic and reagent shortage impaired the culture and experiment conditions. This setback delayed my work for several months. Experiments were prioritized in order to manage plastic distribution to every lab from our institute. I then initiated the metabolic characterization of our stem cell models. In 2020, Agilent updated their Seahorse 24 analyzer, rendering the Nantes' analyzer obsolete.

Because maintenance was impaired during Covid, I had to try to adapt my experiment on a Seahorse XFp analyzer. Although informative, the number of wells available with this analyzer (8, including 1 control well) was insufficient for my experiment. This, once again, delayed my work for several months before finding a new technic to study metabolism, using a fluorescence reader available in Nantes.

Nevertheless, in the span of those 3 years, numerous manuscripts were published describing alternative methods to generate hTSCs, including chemical, transient naive conversion, transcription factors based, conversion from hPPSCs. However, none of these manuscripts looked at additional functional readouts. Moreover, hEPSCs were used to generate “developmental models” such as blastoids, yet they have not been comprehensively characterized. Consequently, our analysis remains highly relevant for the field and even more so than 3 years ago, given the significant volume of manuscripts published in the field. We think it is an essential step toward the broader use of both 2D and 3D models of peri-implantation development.

As part of our analysis, we decided to investigate protein content in our stem cell models through mass spectrometry analysis. This not only enabled us to have a deeper level of understanding of our models on a more functional level but also allowed us to participate to the human proteome project. This project aims to identify missing proteins, meaning that they are predicted to exist but have not yet been detected through mass spectrometry (Girard et al., 2023).

Scientific articles submitted or published during this PhD
Induction of Human Naïve pluripotent stem cells from somatic cells



Induction of Human Naïve Pluripotent Stem Cells from Somatic Cells

Constance Onfray, Jia Ping Tan, Stéphanie Kilens, Xiaodong Liu, Jose Polo, and Laurent David

Abstract

Generating patient-specific stem cells representing the onset of development has become possible since the discovery of somatic cell reprogramming into induced pluripotent stem cells. However, human pluripotent stem cells are generally cultured in a primed pluripotent state: they are poised for differentiation and represent a stage of development corresponding to post-implantation epiblast. Here, we describe a protocol to reprogram human fibroblasts into naive pluripotent stem cells by overexpressing the transcription factors *OCT4*, *SOX2*, *KLF4*, and *c-MYC* using Sendai viruses. The resulting cells represent an earlier stage of development that corresponds to pre-implantation epiblast. We also discuss validation methods for human naive pluripotent stem cells.

Key words Naive pluripotent stem cells, Primed pluripotent stem cells, Somatic cell reprogramming, Yamanaka factors

1 Introduction

Pluripotent stem cells (PSC), whether derived from human blastocysts [1] or reprogrammed from somatic cells [2, 3], have the ability to self-renew and the potency to give rise to any cell types of the three germ layers. In mammals, pluripotency is considered to exist in a continuum of states between inner cell mass (ICM) progenitors and gastrulation. Within this continuum, two attractor states are particularly studied, which are termed naive and primed pluripotent states. The naive state of pluripotency corresponds to the pre-implantation epiblast and is characterized by specific transcriptomic, proteomic, metabolic, and epigenetic features. For instance, naive PSCs feature two active X-chromosomes in female cells, DNA hypomethylation, the expression of specific genes, such

Constance Onfray and Jia Ping Tan contributed equally to this work.

Peter Rugg-Gunn (ed.), *Human Naïve Pluripotent Stem Cells*, Methods in Molecular Biology, vol. 2416, https://doi.org/10.1007/978-1-0716-1908-7_4,

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2022

as *KLF17*, *DNMT3L*, and *DPPA5*, and a high metabolic activity that is not restricted to glycolysis. The primed pluripotent state on the other hand corresponds to the post-implantation epiblast, which is a later developmental state. Female primed PSC lines have an inactive X-chromosome, and the cells cannot perform oxidative phosphorylation [4–11]. Both cell states can be used to study human early development and also to model diseases and to screen drugs. Their different characteristics make them complementary in their applications. Indeed, it has been proposed that specific pluripotent states might be needed for the success of chimeras [12–14].

Naive PSCs (NPSCs) can be generated via derivation from blastocysts [7, 8]. However, the use of cells of embryonic origin comes with ethical issues, tight regulation, limited accessibility, and restricted genomic diversity. An alternative method is to convert primed-state embryonic or induced pluripotent stem cells (iPSC) to NPSCs. Generating patient-specific naive iPSC by conversion would need to first reprogram somatic cells into primed iPSC, and then convert those primed iPSC into naive iPSC, which is a process that would take several months. One possibility to overcome those limitations is to directly reprogram somatic cells into NPSCs. Indeed, this approach enables the generation of induced NPSCs (iNPSCs) from a specific genetic background within a relatively short time frame without traversing primed pluripotency.

Several protocols have been proposed to reprogram somatic cells into iNPSCs [5, 15, 16] (*also see* Chapter 5). The reprogramming strategy also gives the opportunity to generate human isogenic PSCs featuring different states of pluripotency in one experiment. This could help to understand the biological properties of the different pluripotent states in humans and also to study early human development. Furthermore, it circumvents the limitations of derived and converted NPSCs.

Here, we present (1) a detailed protocol to reprogram somatic cells into NPSCs via transduction of OSKM (OCT4, SOX2, KLF4, and *c-MYC*) and culture in T2iLGö medium [5, 15], and (2) methods that can be used to validate the naive pluripotency of reprogrammed iNPSCs.

Reprogramming of fibroblast cells into NPSCs starts by a mesenchymal-to-epithelial transition, with expression of pluripotency markers, such as NANOG [17]. The full set of naive pluripotency-associated transcription factors are subsequently expressed, establishing the pluripotency gene regulatory network [18]. Among the last changes to occur includes the reactivation of the second X-chromosome of female cells. The use of T2iLGö medium instead of the classical KnockOut Serum Replacement/FGF2 culture conditions enable the reprogramming of cell to the naive pluripotent state.

We now have a good understanding of human naive pluripotency as those cells have been thoroughly investigated. Moreover, readouts used to validate naive pluripotency have been tested in human preimplantation embryos, which are the gold standard. A combination of methods can be used to validate human NPSCs: transcriptomic profiling, marker expression, metabolic activity, lncRNA profiling, epigenetic profiling, X-chromosome biallelic expression in female cells, and the ability of cells to differentiate into human trophoblast stem cells [19–22].

In conclusion, overexpression of OSKM factors allows the reprogramming of cells not only to the primed state of pluripotency in human, but also to the naive state of pluripotency. Recent studies demonstrated that OSKM overexpression in human cells also opens a path toward trophoblast stem cell reprogramming [18]. It is now possible to reprogram patient cells into all cell types, including preimplantation pluripotent and trophoblast stem cells.

2 Materials

2.1 Reagents for Sendai Virus Transduction

1. CytoTune-iPS 2.0 Sendai reprogramming kit [polycistronic vector *KLF4-OCT4-SOX2* (KOS), *C-MYC* (M), *KLF4* (K)].

2.2 General Cell Culture Reagents

1. Dulbecco's phosphate-buffered saline (DPBS) without calcium and magnesium.
2. Gelatin solution (0.1%). Place a vial of 2% gelatin solution (Merck) in the water bath at 37 °C. Once melted, add 25 mL of the dissolved 2% gelatin solution to 475 mL of DPBS without calcium and magnesium. Mix the solution and store for up to 6 months at 4 °C.
3. Matrigel hESC-qualified matrix (Corning).
4. TrypLE Express.
5. Mycoplasma Detection Kit.

2.3 Reagents for the Culture Media

1. N2B27 base medium: 50% DMEM/F-12, 50% Neurobasal medium, 1× N-2 Supplement, 1× B-27 Supplement (serum-free), 1× MEM non-essential amino acids, 2 mM GlutaMAX, 0.1 mM 2-mercaptoethanol, 50 µg/mL bovine serum albumin, 0.5% penicillin–streptomycin (10,000 U/mL).
2. T2iLGö medium: N2B27 base medium with 20 ng/mL LIF, 1 µM PD0325901, 1 µM CHIR99021, 5 µM Gö6983, 10 µM Y27632 (*see Note 1*).

3. Fibroblast medium: High-glucose DMEM, 2 mM GlutaMAX-I, 10% fetal bovine serum (FBS), 1 mM sodium pyruvate, 1 × MEM non-essential amino acids, 0.5% penicillin-streptomycin (10,000 U/mL).
4. TeSR-E7 medium for reprogramming.

2.4 Inactivated Mouse Embryonic Fibroblasts (iMEFs)

1. MEFs can be derived from embryonic day 13.5 mouse embryos or purchased commercially and then mitotically inactivated by irradiation or mitomycin treatment.

2.5 Human Fibroblasts

Human fibroblasts can be obtained by culturing skin biopsies. Alternatively, they can be purchased (e.g., we have used fibroblast obtained from Lonza) or ordered from ATCC (e.g., we use BJ1 fibroblasts as positive controls). Fibroblasts should be used at a passage number as early as possible, to limit the appearance of genetic abnormalities that can occur over passaging.

2.6 Equipment for iPSC Colony Subcloning

1. Dissection microscope.
2. Accutase.

3 Methods

3.1 Reprogramming (Fig. 1)

3.1.1 Thawing Human Fibroblasts (Day-5)

1. Coat tissue culture dishes with gelatin solution. Incubate dishes at room temperature for at least 20 min.
2. Transfer 5 mL of fibroblast medium to a tube.
3. Thaw a cryovial containing human fibroblasts in the water bath until just thawed. Transfer the cells into the tube containing fibroblast medium (prepared in **step 2**). Pellet the cells by centrifugation at $170 \times g$ for 5 min.
4. Remove the supernatant and resuspend the cells in 10 mL of fibroblast medium. Count the cells using a hemocytometer or an automated cell counter.
5. Aspirate gelatin solution from dishes prepared in **step 1**. Seed the fibroblasts at a minimal density of 1×10^4 cells per cm^2 of tissue culture plate surface area. This will allow expansion of the fibroblasts before reprogramming.
6. Evenly disperse the cells within the dish and incubate the cells at 37 °C, 5% CO₂.
7. On the following day (day-4), test the fibroblasts for the presence of mycoplasma using a Mycoplasma Detection Kit.

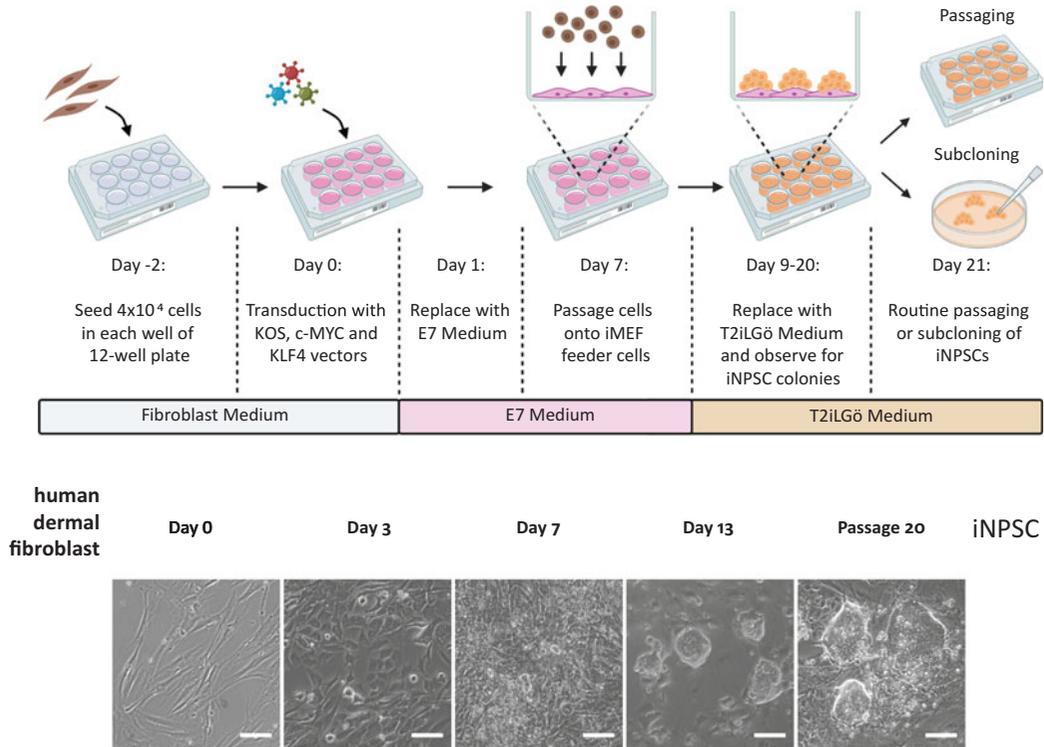


Fig. 1 Schematic depicting reprogramming of human fibroblasts into iNPC. Bright field images depicting the cell morphological changes during each stage of the process are presented underneath each step of the protocol. Between day 9 and 20, cells can be either cultured in 12-well plates or in 35 mm dishes. 35 mm dishes are easier to pick colonies from. Scale bar = 100 μm

3.1.2 Preparation of Human Fibroblasts for Transduction (Day-2) (See **Notes 2 and 3**)

1. Coat a 12-well plate with Matrigel solution (*see Note 4*). Incubate the plate at room temperature for at least 20 min.
2. Aspirate the medium from the human fibroblasts (prepared in Subheading 3.1) and wash the cells with 1 mL of DPBS to remove traces of culture medium.
3. Add 350 μL of TrypLE Express to the cells. Incubate the dish for 5 min at 37 $^{\circ}\text{C}$.
4. Add 700 μL of fibroblast medium to inactivate the TrypLE and pipette up and down for 3–5 times to dissociate the fibroblasts.
5. Collect the cell suspension into a tube. Pellet the cells by centrifugation at $170 \times g$ for 5 min.
6. Remove the supernatant and resuspend the human fibroblasts in 1 mL of fibroblast medium. Count the cells using a hemocytometer or an automated cell counter.
7. Aspirate the Matrigel solution from the 12-well plate. Seed 4×10^4 human fibroblasts per well in 1 mL of fibroblast medium (*see Note 2*). Distribute the cells evenly and incubate the plate at 37 $^{\circ}\text{C}$, 5% CO_2 .

8. On the following day (Day-1), change the medium for fresh fibroblast medium.

3.1.3 Sendai Virus Transduction (Starting on Day 0)

1. Harvest the human fibroblasts from one well of the 12-well plate following **steps 2–6** in Subheading **3.1.2**. Count the cells using a hemocytometer or an automated cell counter to estimate the cell number for transduction. We typically obtain around 100,000 fibroblasts per well.
2. Transduce the cells in the wells of interest using the CytoTune-iPS 2.0 Sendai Reprogramming Kit with three vectors (polycistronic vector *KLF4-OCT4-SOX2* (KOS), *C-MYC* (M), *KLF4* (K)) at 5 (KOS), 5 (M), 3 (K) multiplicity of infection (MOI) (*see Note 5*). Follow the manufacturer's guidelines for the cell transduction procedure and adhere to local safety requirements for working with the viruses.
3. Evenly disperse the viral particles and incubate at 37 °C, 5% CO₂.
4. On the following day (Day 1) remove the fibroblast medium from the cells and replace with TeSR-E7 medium (*see Note 6*). Over the next 5 days (Days 2–6), monitor the cells under a microscope and replace the medium every other day with TeSR-E7 medium (*see Note 6*).

3.1.4 iMEF Preparation (Day 6) (See **Note 7**)

1. Coat 35 mm dishes with 1 mL of gelatin solution. Incubate the dishes at room temperature for at least 20 min.
2. Transfer 5 mL of fibroblast medium into a tube.
3. Thaw a cryovial of iMEF in a water bath until just thawed.
4. Add 1 mL of fibroblast medium to the iMEF and transfer the cells into the tube containing MEF medium (prepared in **step 2**).
5. Pellet the cells by centrifugation at $170 \times g$ for 5 min. Remove the supernatant and resuspend the cells in 10 mL fibroblast medium. Count the cells using a hemocytometer or an automated cell counter.
6. Aspirate the gelatin solution from the 35 mm dishes (prepared in **step 1**). Seed 2.5×10^5 iMEF per 35 mm dish in 1.5 mL of fibroblast medium. Evenly distribute the cells in the dish and incubate the dish for 24 h at 37 °C, 5% CO₂.

3.1.5 Passaging the Transduced Cells (Day 7)

1. Remove the TeSR-E7 medium from the transduced cells in the 12-well plate. Wash the cells with 1 mL of DPBS per well to remove traces of culture medium.
2. Add 350 µL of TrypLE per well. Incubate the plate for 5 min at 37 °C.

3. Add 700 μL of TeSR-E7 medium to inactivate the TrypLE and pipette up and down for 3–5 times to dissociate the fibroblasts.
4. Collect the cell suspension into a tube. You can pool cells from the same line of fibroblasts. Pellet the cells by centrifugation at $170 \times g$ for 5 min.
5. Remove the supernatant and resuspend the cells in 1 mL of TeSR-E7 medium. Count the cells using a hemocytometer or an automated cell counter.
6. Seed one-third of the dissociated cells per 35 mm dishes pre-coated with iMEFs ($\sim 1 \times 10^5$ transduced cells or 12×10^3 transduced cells/ cm^2) that was prepared in Subheading 3.1.4. Evenly distribute the cells in the dishes and incubate the cells at 37°C , 5% CO_2 .
7. On the following day (Day 8), replace the TeSR-E7 medium from the cells in the 35 mm dishes with fresh TeSR-E7 medium.
8. On Day 9, aspirate the TeSR-E7 medium from the cells in the 35 mm dishes and replace with 1.5 mL of T2iLGö medium freshly supplemented with 10 μM Y27632 (*see Note 8*).

3.1.6 Routine Culture (Days 10–20)

1. Replace the medium everyday with T2iLGö medium freshly supplemented with 10 μM Y27632 and monitor the dishes for the appearance of iNPSC colonies (*see Note 9*).
2. Passage the cells as follows (*see Notes 8 and 10*).
3. One day before passaging the cells, prepare iMEF-coated dishes as described in Subheading 3.1.4 (*see Note 7*).
4. On the day of the passaging, remove the T2iLGö medium from the 35 mm dishes. Wash the cells with 1 mL of DPBS per dish to remove traces of the culture medium.
5. Add 350 μL of TrypLE per 35 mm dish. Incubate the cells for 5 min at 37°C .
6. Add 700 μL of T2iLGö medium to each dish. Collect the cell suspension into a tube.
7. Pellet the cells by centrifugation at $170 \times g$ for 5 min. Remove the supernatant and resuspend the cells in 1 mL of T2iLGö medium freshly supplemented with 10 μM Y27632. Count the cells (*see Note 11*).
8. Seed 2×10^5 dissociated cells in 1.5 mL of T2iLGö medium freshly supplemented with 10 μM Y27632 onto new 35 mm dishes pre-coated with iMEF.
9. Single-cell subcloning can also be performed at this step to obtain transgene-free cell lines more quickly (*see Note 12*). Identify a dome-shaped iNPSC colony using a dissection microscope. Isolate the colony using a pipette tip (**step 1**,

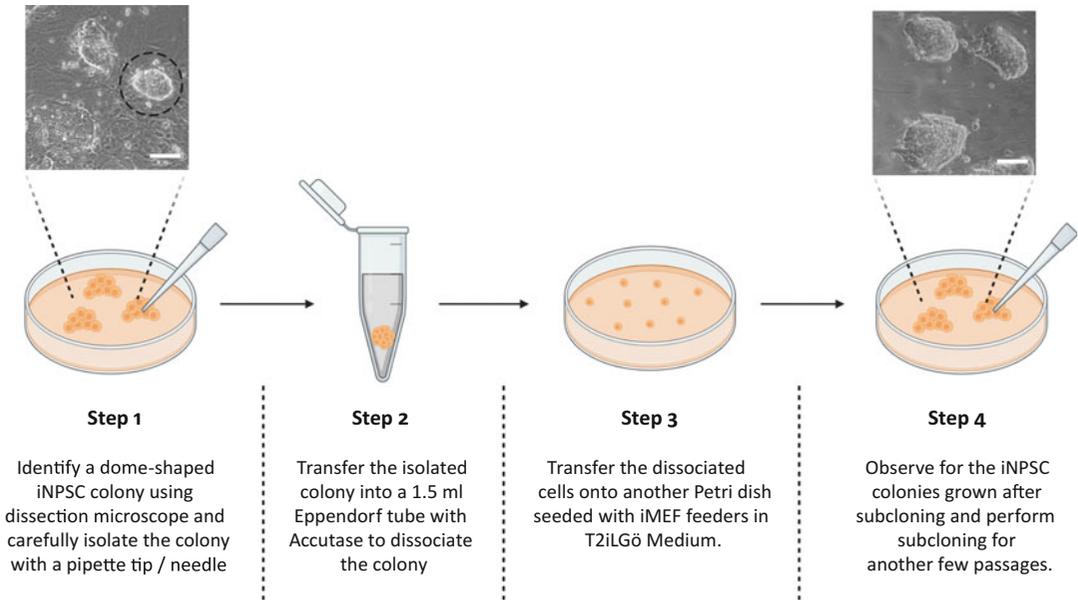


Fig. 2 Schematic depicting the subcloning process step by step with bright field pictures depicting the initial and final cultures of the cells. One NPSC colony out of three is circled on the first phase image. This colony is isolated using a pipette tip. On the second phase image, three new colonies have arisen from the circled dissociated colony from the first phase image. Scale bar = 100 μm

Fig. 2). Transfer the colony into a 1.5 mL Eppendorf tube and dissociate it with 0.5 mL of accutase (**step 2**, Fig. 2). Transfer the dissociated cells onto a new 35 mm dish pre-coated with iMEF (**step 3**, Fig. 2). Monitor the colony growth and repeat these steps for at least the five subsequent passages before testing for the presence of Sendai virus (**step 4**, Fig. 2).

10. Every 4–5 days afterward, repeat the passaging step described in Subheading 3.1.6 or subclone a colony using a dissection microscope (Fig. 2).

**3.2 After Passage 15:
Cell Validation**

After 15 regular passages or five single-cell subcloning passages (Fig. 2), confirmation of the absence of Sendai virus and validation of the naive state of the cells should be performed (*see* **Notes 12** and **13**). Cells at an earlier passage usually still contain the Sendai virus.

**3.2.1 Sendai Virus
Detection**

Perform Reverse Transcription-quantitative PCR (RT-qPCR) using primers that can detect the Sendai virus genome and transgene (*see* **Note 14**). It is advised to use transduced cells at Day 7 as a positive control. Cells are generally negative after 15 passages.

**3.2.2 Naive State
Validation**

We recommend to validate NPSCs through a combination of techniques. We often start by RNA-sequencing or RT-qPCR for specific markers for naive pluripotency. iNPSCs express the core

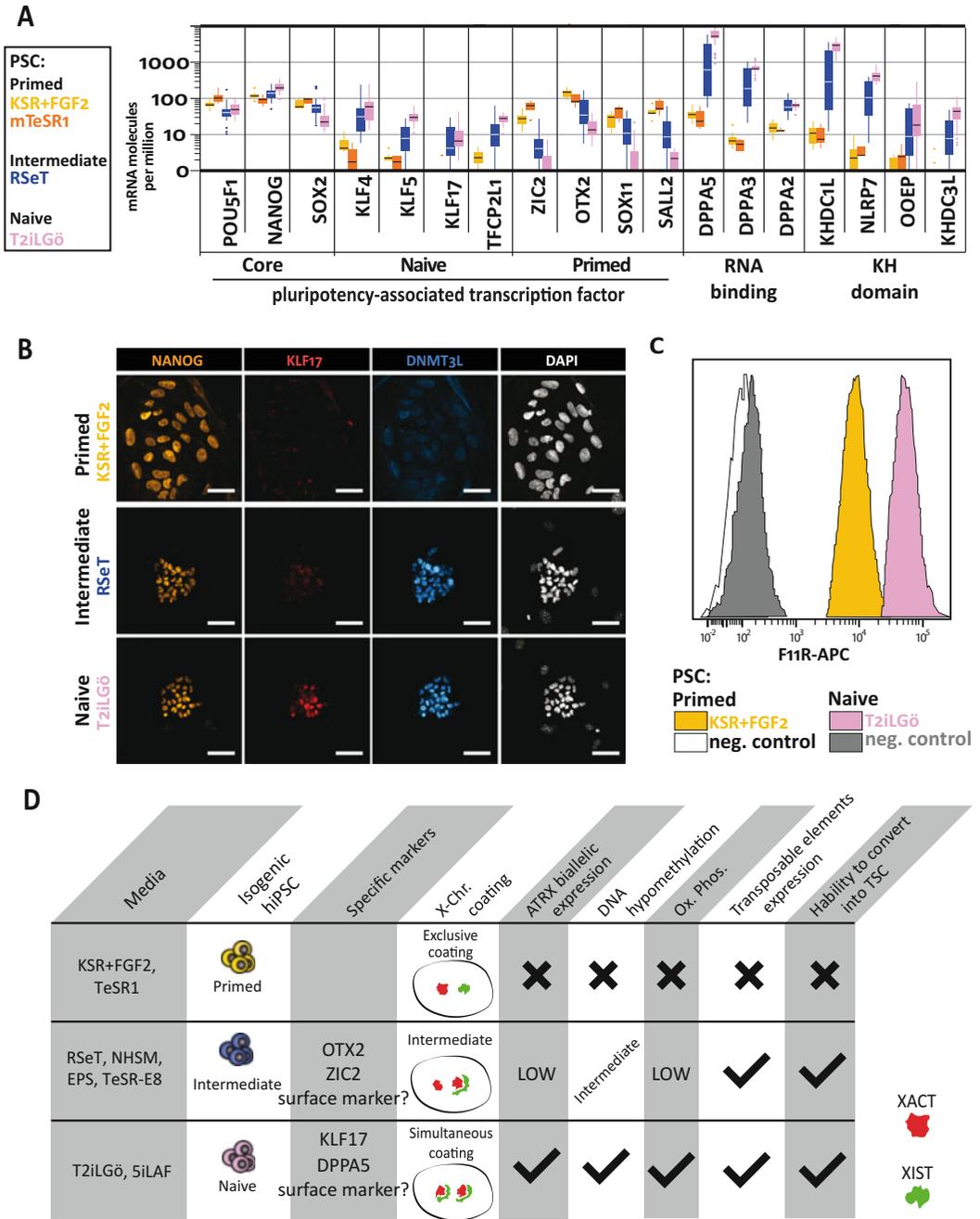


Fig. 3 Validation of human pluripotency states. (a) Expression levels of indicated genes in primed, intermediate, and naive PSCs. Of note, the expression levels are expressed as the number of RNA molecules per million of RNA molecules. The scale is log 10. (b) Sample immunofluorescence showing staining for NANOG, KLF17, and DNMT3L in primed, intermediate, and naive PSC. Note that intermediate cells express DNMT3L but not KLF17. Scale bar = 50 μ m. (c) Flow cytometry histograms showing F11R-APC signal in hiPSC. Negative controls are isogenic antibodies. (d) Schematic summary of the readouts that can distinguish human pluripotency states. (Data from Kilens et al. [5] and Liu et al. [15] were reused with permission from Nature Publishing Group)

transcription factors, OCT4 (POU5F1), NANOG, and SOX2, and generally express higher *NANOG* levels and lower *SOX2* levels compared to primed PSCs. iNPSCs also express specific transcription factors, such as *KLF17*. Conversely, primed PSCs express *OTX2*, *ZIC2*, and *SALL2* (Fig. 3a). NPSCs also display extremely high levels of *DPPA5* and *KHDC1L*: each of those genes could represent up to 0.5% of total transcripts per cell (Fig. 3a). As a follow up, we routinely perform *KLF17* immunofluorescence [5, 8, 15] or flow cytometry analysis of F11R that is expressed specifically in NPSCs [15] (Fig. 3b and c). We include additional readouts when thoroughly validating NPSCs (Fig. 3d). Metabolically, NPSCs have high oxidative phosphorylation and glycolysis activity [5, 8]. Intermediate PSCs, such as human pluripotent cells with some features of naive pluripotent cells, have specific lipid metabolic activity [23]. Epigenetically, NPSCs are globally DNA demethylated [7] (*see also* Chapter 11) and the two X-chromosomes of female cells are coated with both XIST and XACT long non-coding RNA, and express both alleles of X-linked genes [4–6] (*see also* Chapter 15). Finally, we can assess the ability of human PSCs to convert into trophoblast stem cells (*see also* Chapter 7). Indeed, NPSCs can readily convert into trophoblast, whereas intermediate cells can convert at a lower rate and primed cells cannot convert [20–22].

4 Notes

1. T2iLGö medium (without Y-27632) should be kept at 4 °C and used within 3 days. Y-27632 should be added daily.
2. Seed the number of wells that you plan to reprogram, plus one or several extra well(s) for counting.
3. To avoid genomic abnormalities or cell senescence acquired during long-term fibroblast culturing, fibroblasts used for reprogramming are recommended to be passage 3–5 or less.
4. Refer to the manufacturer's instructions for the Matrigel solution dilution as the concentration is lot-dependent.
5. An alternative multiplicity of infection can be used: Liu and colleagues used a 5, 5, 6 multiplicity of infection (MOI) instead of 5, 5, 3 (ref. 15). This alternative MOI can be used if reprogramming efficiency is too low with the 5, 5, 3 MOI. The manufacturer of the CytoTune™-iPS 2.0 Sendai Reprogramming kit also states that an MOI of 10, 10, 3 or 10, 10, 6 can be used to increase the efficiency of reprogramming.
6. As an alternative to TeSR-E7 medium, the reprogramming process up to day 7 can be performed in fibroblast medium,

with the culture medium being switched to T2iLGö medium at day 8 for the generation of iPSCs [15].

7. iMEF can be prepared either by irradiation or by treatment with mitomycin. Irradiation can be done up to 3 days in advance for culture use. Paynter and colleagues describe a protocol to generate iMEF [24]. Our protocol to prepare mitomycin-inactivation MEF is detailed in the manuscript by Castel and colleagues [20].
8. At day 9, one alternative is to generate isogenic primed and naïve iPSC lines by dividing the cells into different batches, each grown in their respective medium. T2iLGö medium can also be switched for another naïve pluripotency medium, such as RSeT, NHSM, 5iLAF, or PXGL. Depending on the media used from day 9, the time necessary before transferring the newly formed colonies onto a new plate will vary (between 15 and 21 days).
9. During reprogramming and culture, differentiation events in the cultures can be observed. Monitor the cells daily with a phase contrast microscope and if differentiation is spotted, aspirate the differentiated cell patches during the medium change.
10. Do not let the colonies overgrow. A maximum size is around 40 cells per colony. Passage the cells regularly in a range of density that leave them not too dense or sparse. With our cell lines, re-plating cells at 2×10^4 cells per cm^2 every 4–5 days is ideal.
11. We routinely obtain between 5×10^4 and 6×10^4 cells per cm^2 every 4–5 days.
12. The manufacturer of the CytoTune™-iPS 2.0 Sendai Reprogramming Kit suggests single colony subcloning during the first few passages to reduce the length of time that is necessary to obtain transgene-free clones.
13. T2iLGö medium can lead to karyotype abnormalities, such as tetraploidy. Karyotyping should be performed before banking the cells (*see* Chapter 17). 5iLAF-generated cells have many similarities to T2iLGö cells; however, it has been reported that reprogramming using 5iLAF medium also led to karyotypic anomalies [7, 15, 25].
14. Sendai virus primer sequences are available in the CytoTune™-iPS 2.0 Sendai Reprogramming Kit User Guide (available for RT-qPCR, real-time RT-PCR, and Taqman probes).

References

1. Thomson JA, Itskovitz-Eldor J, Shapiro SS et al (1998) Embryonic stem cell lines derived from human blastocysts. *Science* 282: 1145–1147. <https://doi.org/10.1126/science.282.5391.1145>
2. Takahashi K, Tanabe K, Ohnuki M et al (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131: 861–872. <https://doi.org/10.1016/j.cell.2007.11.019>
3. Yu J, Vodyanik MA, Smuga-Otto K et al (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318:1917–1920. <https://doi.org/10.1126/science.1151526>
4. Vallot C, Patrat C, Collier AJ et al (2017) XACT noncoding RNA competes with XIST in the control of X chromosome activity during human early development. *Cell Stem Cell* 20: 102–111. <https://doi.org/10.1016/j.stem.2016.10.014>
5. Kilens S, Meistermann D, Moreno D et al (2018) Parallel derivation of isogenic human primed and naive induced pluripotent stem cells. *Nat Commun* 9:360. <https://doi.org/10.1038/s41467-017-02107-w>
6. Sahakyan A, Kim R, Chronis C et al (2016) Human naive pluripotent stem cells model X chromosome dampening and X inactivation. *Cell Stem Cell* 20:87–101. <https://doi.org/10.1016/j.stem.2016.10.006>
7. Pastor WA, Chen D, Liu W et al (2016) Naive human pluripotent cells feature a methylation landscape devoid of blastocyst or germline memory. *Cell Stem Cell* 18:323–329. <https://doi.org/10.1016/j.stem.2016.01.019>
8. Guo G, von Meyenn F, Santos F et al (2016) Naive pluripotent stem cells derived directly from isolated cells of the human inner cell mass. *Stem Cell Reports* 6:437–446. <https://doi.org/10.1016/j.stemcr.2016.02.005>
9. Guo H, Zhu P, Yan L et al (2014) The DNA methylation landscape of human early embryos. *Nature* 511:606–610. <https://doi.org/10.1038/nature13544>
10. Smith ZD, Chan MM, Humm KC et al (2014) DNA methylation dynamics of the human pre-implantation embryo. *Nature* 511:611–615. <https://doi.org/10.1038/nature13581>
11. Gu W, Gaeta X, Sahakyan A et al (2016) Glycolytic metabolism plays a functional role in regulating human pluripotent stem cell state. *Cell Stem Cell* 19:476–490. <https://doi.org/10.1016/j.stem.2016.08.008>
12. Theunissen TW, Friedli M, He Y et al (2016) Molecular criteria for defining the naive human pluripotent state. *Cell Stem Cell* 19:502–515. <https://doi.org/10.1016/j.stem.2016.06.011>
13. Gafni O, Weinberger L, Mansour A et al (2013) Derivation of novel human ground state naive pluripotent stem cells. *Nature* 504: 282–286. <https://doi.org/10.1038/nature12745>
14. Wu J, Platero-Luengo A, Sakurai M et al (2017) Interspecies chimerism with mammalian pluripotent stem cells. *Cell* 168:473–486. <https://doi.org/10.1016/j.cell.2016.12.036>
15. Liu X, Nefzger CM, Rossello FJ et al (2017) Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. *Nat Methods* 14: 1055–1062. <https://doi.org/10.1038/nmeth.4436>
16. Giulitti S, Pellegrini M, Zorzan I et al (2019) Direct generation of human naive induced pluripotent stem cells from somatic cells in microfluidics. *Nat Cell Biol* 21:275–286. <https://doi.org/10.1038/s41556-018-0254-5>
17. David L, Polo JM (2014) Phases of reprogramming. *Stem Cell Res* 12:754–761. <https://doi.org/10.1016/j.scr.2014.03.007>
18. Liu X, Ouyang JF, Rossello FJ et al (2020) Reprogramming roadmap reveals route to human induced trophoblast stem cells. *Nature* 586:101–107. <https://doi.org/10.1038/s41586-020-2734-6>
19. Boroviak T, Nichols J (2017) Primate embryogenesis predicts the hallmarks of human naive pluripotency. *Development* 144:175–186. <https://doi.org/10.1242/dev.145177>
20. Castel G, Meistermann D, Bretin B et al (2020) Induction of human trophoblast stem cells from somatic cells and pluripotent stem cells. *Cell Rep* 33(8):108419. <https://doi.org/10.1016/j.celrep.2020.108419>
21. Cinkornpumin JK, Kwon SY, Guo Y et al (2020) Naive human embryonic stem cells can give rise to cells with a trophoblast-like transcriptome and methylome. *Stem Cell Reports* 15(1):198–213. <https://doi.org/10.1016/j.stemcr.2020.06.003>
22. Dong C, Beltcheva M, Gontarz P et al (2020) Derivation of trophoblast stem cells from naive human pluripotent stem cells. *elife* 9:e52504. <https://doi.org/10.7554/eLife.52504>
23. Cornacchia D, Zhang C, Zimmer B et al (2019) Lipid deprivation induces a stable, naive-to-primed intermediate state of

- pluripotency in human PSCs. *Cell Stem Cell* 25:120–136. <https://doi.org/10.1016/j.stem.2019.05.001>
24. Paynter JM, Chen J, Liu X et al (2019) Propagation and maintenance of mouse embryonic stem cells. *Methods Mol Biol* 1940:33–45. https://doi.org/10.1007/978-1-4939-9086-3_3
25. Theunissen TW, Powell BE, Wang H et al (2014) Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell* 15: 471–487. <https://doi.org/10.1016/j.stem.2014.07.002>

Deciphering hallmarks ability to stage match peri-implantation stem cell models

1 **Deciphering hallmark combination distinct for peri-implantation stem cell models**

2
3 Constance Onfray¹, Simon Chevolleau¹, Eva Moinard¹, Océane Girard¹, Kasturi Mahadik²,
4 Ryan Allsop³, Grigorios Georgolopoulos³, Régis Lavigne^{7,8}, Ophélie Renoult⁵, Irene Aksoy⁶,
5 Jean-François Ouimette², Thomas Fréour^{1,9,10}, Claire Pecqueur⁵, Charles Pineau^{7,8}, Vincent
6 Pasque³, Claire Rougeulle², Laurent David^{1,4}.

7
8 1. Nantes Université, CHU Nantes, Inserm, CR2TI, F-44000, Nantes, France

9 2. Université Paris Cité, CNRS, Epigenetics and Cell Fate, F-75013 Paris, France

10 3. KU Leuven - University of Leuven, Department of Development and Regeneration, Leuven
11 Institute for Single Cell Omics and Leuven Stem Cell Institute, Herestraat 49, B-3000 Leuven,
12 Belgium

13 4. Nantes Université, CHU Nantes, Inserm, CNRS, BioCore, F-44000 Nantes, France

14 5. Nantes Université, CNRS, Inserm, CRCI2NA, F-44000 Nantes, France

15 6. Univ Lyon, Université Lyon 1, Inserm, Stem Cell and Brain Research Institute U1208, F-69500
16 Bron, France

17 7. Univ Rennes, Inserm, EHESP, Irset (Institut de Recherche en Santé, Environnement et
18 Travail) - UMR_S 1085, F-35000 Rennes, France

19 8. Univ Rennes, CNRS, Inserm, Biosit UAR 3480 US_S 018, Protim Core Facility, F-35000
20 Rennes, France

21 9. Department of Obstetrics, Gynecology and Reproductive Medicine, Dexeus University
22 Hospital, 08028 Barcelona, Spain

23 10. CHU Nantes, Service de Biologie de la Reproduction, F-44000 Nantes, France

24
25 Correspondence: laurent.david@univ-nantes.fr

26 Lead Contact: laurent.david@univ-nantes.fr

27 28 **Summary**

29
30 Development of new 2D and 3D models of human development such as trophoblast stem cells,
31 gastruloids or blastoids widened possibilities to study early timepoints of development and
32 brightened up ever so slightly the black box of human development. While opening new
33 horizons, the cell sources of those models need proper benchmarking to clarify which hallmark
34 is associated with which lineage and developmental stage. Here, we propose a thorough
35 characterization of pluripotent and trophoblastic stem cell models by transcriptomic, proteomic,
36 epigenetic and metabolic approaches. Extended pluripotent stem cells are similar to primed
37 pluripotent stem cells for most criteria, except metabolic activity, which might explain their
38 ability to convert directly into trophoblast stem cells. We show that trophoblast stem cells are
39 hypo-methylated and that they have a high metabolic activity. Our results clarify the fact that
40 hallmarks of pluripotency are not predictive of each other and have to be used in combination.
41 Multiplying hallmarks alleviate stage matching bias.

42 43 **Keywords**

44 Epiblast; trophectoderm; trophoblast; pluripotency; cell fate; peri-implantation development;
45 hallmarks

46 **Introduction**

47

48 The discovery and popularization of organoids, complex stem cell-based and integrated models
49 call for a strong effort to characterize and standardize those models. Community efforts led the
50 revision of ISSCR standards for basic stem cell research. While most of the community's effort
51 have been focused on primed pluripotency, the recent development of blastoids and other stem
52 cell models of peri-implantation demonstrates the importance of applying the same efforts to
53 all stem cell models, including peri-implantation models (Rugg-Gunn et al., 2023).

54

55 A flurry of cellular systems has been developed to model the different lineages of the human
56 embryo. In 1998, pluripotent stem cells (PSCs) were derived from human blastocysts (Thomson
57 et al., 1998) and were later shown to correspond to the post-implantation epiblast (10 to 14 dpf),
58 hence to a primed state of pluripotency. Culture conditions to capture cells corresponding to
59 pre-implantation epiblast (6 to 9 dpf) were later developed, allowing naive pluripotent stem
60 cells to be derived from human blastocysts (Chen et al., 2015; Guo et al., 2016; Hanna et al.,
61 2010; Takashima et al., 2014; Theunissen et al., 2014). Later work then generated naive
62 pluripotent stem cells (NPSCs) by resetting primed pluripotent stem cells (PPSCs) or through
63 reprogramming using OSKM (Gafni et al., 2013; Giulitti et al., 2019; Kilens et al., 2018; Liu
64 et al., 2017; Takahashi et al., 2007; Yu et al., 2007).

65

66 Naive pluripotent stem cells have the closest transcriptomic profile to human preimplantation
67 epiblast. They express specific genes such as KLF17, DNMT3L and DPPA5 (Kilens et al 2018).
68 They are hypo-methylated and have a higher metabolic activity compared to primed pluripotent
69 stem cells. Female cell lines also have both X chromosomes active (Guo et al., 2014; Kilens et
70 al., 2018; Leitch et al., 2013; Linneberg-Agerholm et al., 2019; Pastor et al., 2016; Sahakyan et
71 al., 2017; Smith et al., 2014; Takashima et al., 2014; Theunissen et al., 2016; Vallot et al., 2017;
72 Yan et al., 2013). All these characterization efforts led to the establishment of consensus
73 hallmarks for the assessment of naive pluripotency (De Los Angeles et al., 2015). Recently,
74 new markers have been added to the previously established hallmarks of naive pluripotency:
75 the ability to make human-animal chimeras, to differentiate into the trophoblast lineage and to
76 form blastoids (Castel et al., 2020; Cinkornpumin et al., 2020; Gafni et al., 2013; Kagawa et
77 al., 2022; Liu et al., 2020, 2021b; Sozen et al., 2021; Theunissen et al., 2014, 2016; Wu et al.,
78 2017; Yanagida et al., 2021; Yu et al., 2021).

79

80 An additional state of pluripotency has been captured *ex vivo* and reported to contribute to
81 chimeras, although at low efficiency, and to convert into extra-embryonic cells: extended PSCs
82 or EPSCs (Castel et al., 2020; Gao et al., 2019; Sozen et al., 2021; Tan et al., 2021; Yang et al.,
83 2017). It has been proposed that these pluripotent stem cells with extended or expanded
84 potential resemble 2C-like cells and represent a new, more potent state of stemness (Yang et
85 al., 2017, Gao et al., 2019; Liu et al., 2021a). However, several groups have pointed out the
86 limit of the chimera experiments performed in EPSCs studies (Posfai et al., 2021, Aksoy et al.,
87 2021). Since EPSCs are not actually colonizing tissues, one could propose that these cells have
88 the ability to survive in an ectopic environment, but do not actually chimerise with the host
89 embryos. Additionally, these cells have not been properly benchmarked using naive
90 pluripotency hallmarks.

91

92 Trophoblast stem cells (TSCs) have recently been derived from human blastocyst, but also from
93 first trimester placenta (Okoe et al., 2018). Our team and others also observed that naive PSCs
94 can engage into the trophoblast fate and convert into TSCs (Castel et al., 2020; Dong et al.,
95 2020; Liu et al., 2020). Alternative culture medium to the original ASECRiAV medium (Okoe

96 et al., 2018) that enables to maintain TSCs in culture was also developed: ACE medium (Io et
97 al., 2021). While TSCs have been proposed to correspond to cytotrophoblast emerging around
98 7 to 9 d.p.f. (Castel et al., 2020), the validation mostly relied on transcriptomic analysis. Further
99 work is needed to precisely characterize trophoblast stem cells and their correspondence with
100 the human embryo but also to determine the correspondence between cells cultured in
101 ASECRiAV and ACE medium. In addition, pre-implantation trophoderm engagement of
102 naive PSCs has been observed upon inhibition of NODAL and ERK pathways (Guo et al., 2021;
103 Io et al, 2021), but again, the cells were mostly characterized at the transcriptomic level.

104

105 The wide range of stem cell models gives the unique opportunity to clarify which hallmark is
106 associated with which lineage and developmental stage. Here, we have characterized human
107 EPSCs and TSCs in parallel with NPSCs and PPSCs through analyses of transcriptomic,
108 proteomic, epigenetic and metabolic features. Parallel comparison of peri-implantation stem
109 cell models over different characteristics will significantly contribute to the establishment of
110 standards for these models, as previously done for primed pluripotency.

111

112 **Results**

113

114 **Transcriptomic comparison of pluripotent and trophoblast stem cell models**

115 In order to compare transcriptomic features of pre- and post-implantation models, we co-
116 analyzed RNAseq datasets from naive pluripotent stem cells (NPSCs), primed pluripotent stem
117 cells (PPSCs), extended pluripotent stem cells (EPSCs), trophoblast stem cells (TSCs) and
118 trophoderm like cells (TELCs) (**Figure 1A, Figure sup 1A**). To avoid biases associated with
119 sequencing platforms, we obtained RNA directly from the groups that generated the models:
120 NPSCs (Guo et al., 2016), ACE-TSCs (Io et al., 2021), placenta- and blastocyst-derived
121 ASECRiAV-TSCs (Okoe et al., 2018) and compared them to 6 ASECRiAV-TSCs lines
122 generated in house: 3 induced TSCs from somatic cells, 3 converted TSCs from NPSCs. We
123 also included NPSC undergoing TSC conversion (Castel et al., 2022) and trophoderm-like
124 cells (TELCs) generated through differentiation of NPSC treated with A8301 and PD0325901
125 (AP) up to 6 days, as published in Guo et al, 2021. Finally, we included 2 EPSC lines generated
126 in-house (Castel et al., 2020), 4 PPSCs lines (3 in-house and H9 hESCs) and 7 NPSCs lines (6
127 in-house (Kilens et al., 2018) and HNES1 line (Guo et al., 2016)(**Table sup3**).

128

129 We performed a Pearson correlation analysis on all samples which revealed 4 main groups of
130 samples: 1/ all TSCs, from Okoe et al, Castel et al, later mentioned as ASECRiAV-TSCs,
131 together with TSCs from Io et al, later referred to as ACE-TSCs; 2/ some TSCs together with
132 TELCs; 3/ all NPSCs together with some intermediates of TELCs differentiation; 4/ EPSCs
133 together with PPSCs (**Figure 1A**). We also performed a correlation analysis after grouping
134 samples per culture condition, which confirmed correlation between TELCs with both TSCs
135 and NPSCs (**Figure sup 1A**).

136

137 **TELCs recapitulate human TE specification molecular aspects**

138 TELCs have been proposed to recapitulate human TE maturation (Guo et al., 2021). To further
139 study this aspect, we analyzed three markers of TE fate progression that we recently described
140 in human embryos: GATA3, the earliest marker of TE specification (Gerri et al., 2020), CDX2,
141 that comes up at B3 blastocyst stage (Kagawa et al., 2022; Niakan and Eggan, 2013), followed
142 by NR2F2 upon TE maturation at the polar side (Meistermann et al., 2021). We measured the
143 expression levels of these key markers of trophoblastic fate progression using bulk RNA
144 sequencing. We observed GATA3 expression 48h after induction of differentiation. We
145 observed CDX2 expression transiently peaks after 72h of differentiation and NR2F2 expression

146 coming up after 72h of differentiation (**Figure 1B**). The correlation of TELCs with NPSCs is
147 puzzling, therefore we analyzed heterogeneity of the population by immunofluorescence
148 analysis for GATA3 (TE marker) and NANOG (EPI marker). Within 120h of differentiation,
149 we observed 2 patterns of TELCs: GATA3 positive cells, and a small subset of cells progressing
150 in TE fate with expression of CDX2 or NR2F2 (**Figure 1 C, D, Figure sup 1B**). We conclude
151 that cells acquire GATA3 first, and based on transcriptomic data, CDX2 second and NR2F2
152 last, as in blastoids (Kagawa et al, 2021). Additionally, immunofluorescence quantification
153 showed that a significant proportion of cells (30 to 40%) did not engage toward the TE fate,
154 which is consistent with recent reports (Osnato et al., 2021; Zijlmans et al., 2022). Given the
155 high heterogeneity of TELCs in our hands, we excluded this model for further population-based
156 analysis.

157

158 **DNA methylation levels are lower in TSCs and NPSCs compared to PSCs and EPSCs**

159 As a proxy to characterize the chromatin features of stem cell models for human peri-
160 implantation development, we analyzed expression levels of the DNMT gene family. This
161 revealed that DNMT3L is mostly expressed in NPSCs (Zijlmans et al., 2022) and that
162 DNMT3B, responsible for DNA methylation, is more expressed in PPSCs and EPSCs.
163 Interestingly, TSCs have the lowest expression of DNMT3A and have low expression of
164 DNMT3B, like NPSCs. However, TSCs express low levels of DNMT3L, like PPSCs. On
165 another hand, EPSCs have a similar profile than PPSCs (**Figure 2A**). To validate our
166 transcriptomic data, we used mass spectrometry data as an orthogonal validation dataset. To do
167 so, we acquired proteomic profile of EPSCs and PPSCs to complete our previously described
168 dataset of NPSCs and TSCs (Girard et al, 2023). Mass spectrometry data independent analysis
169 (DIA) detected respectively 8740, 8659 and 9141 proteins in NPSCs, PPSCs and EPSCs. An
170 in-depth analysis was performed on the peptide score, reflecting the sum of areas under the
171 curve for all the fragments corresponding to one peptide-precursor. This analysis reflects
172 protein expression. Protein expression of DNMT1, DNMT3A, DNMT3B and DNMT3L by
173 mass spectrometry confirmed the transcriptomic analysis, associating the EPSCs with PPSCs
174 and confirming the low levels of DNMT3A, DNMT3B and DNMT3L in TSCs (**Figure 2B**).

175

176 We then quantified global DNA methylation levels, as it has proven to be a robust way to assess
177 naive vs primed pluripotency (De Los Angeles et al., 2015; Kilens et al., 2018; Pastor et al.,
178 2016; Takashima et al., 2014; Theunissen et al., 2016). We quantified 5-methylcytosine (5mC)
179 by mass spectrometry and showed that cells are organized in two samples groups. On one hand,
180 we identified PPSCs and EPSCs linking EPSCs with primed pluripotency (**Figure 2C**). On the
181 other hand, NPSCs and TSCs were hypomethylated when compared to PPSCs and EPSCs
182 (**Figure 2C**). This complements the observation by Okae et al and collaborators that TSCs are
183 globally hypomethylated relative to cytotrophoblasts (Cinkornpumin et al., 2020; Okae et al.,
184 2018). Altogether, our investigation precisely specifies the relative methylation levels between
185 all peri-implantation stem cell models and associate EPSCs with the primed state of
186 pluripotency.

187

188 **TSCs and EPSCs have one inactivated X chromosome, contrasting with NPSCs.**

189 The X chromosome activity status has been shown as one of the most stringent criteria that
190 distinguishes NPSCs and PPSCs, with female NPSCs carrying 2 active X chromosomes (Xa),
191 whereas female PPSCs are characterized by the presence of one active and one inactive X (Xi)
192 (Kilens et al., 2018; Vallot et al., 2017). This can be monitored by RNA-FISH of X-linked
193 genes such as HUWE1 or XACT that demarks X transcriptional activity (Sahakyan et al., 2017;
194 Vallot et al., 2017). RNA-FISH for the long non-coding RNA XIST, the trigger of XCI, can
195 also be informative, although in NPSCs, presence of XIST is unlinked from silencing (Vallot

196 et al., 2017). In addition, progressive loss of XIST expression and other XCI hallmarks occurs
197 spontaneously upon culturing PPSCs, a process called XCI erosion (Vallot et al., 2015,
198 Sahakyan et al., 2016); erosion is for example accompanied by the re-expression of certain
199 genes on the X including XACT. Erosion of XCI has been associated with the presence of
200 GSK3 inhibitor in the culture medium (Cloutier et al., 2022). As we are using culture medium
201 with GSK3 inhibitor, we included the analysis of XACT lncRNA to take this potential bias into
202 account in our analysis.

203 Analysis and quantification of XIST, XACT and HUWE1 patterns in NPSCs, PPSCs, EPSCs
204 and TSCs allowed us to determine the activity status of each chromosome (table sup4). The
205 XaXa status is determined by bi-allelic expression of HUWE1 and XACT associated with the
206 presence of XIST from at least one chromosome. The XaXi status is defined by monoallelic
207 expression of HUWE1 together with XIST accumulation from the other X chromosome.
208 Finally, the XaXe status is inferred from co-presence of XACT and HUWE1 on both
209 chromosomes and lack of XIST expression (table sup4) (**Figure 2D**). We observed that NPSCs
210 display 2 active Xs in more than 50% of the cells (**Figure 2E, Figure sup 1D**). On the other
211 hand, TSCs, like PPSCs, show inactivation of one X chromosome in more than 80% of the
212 cells, corroborating recent findings obtained in trophoblast organoids (Karvas et al 2022). Of
213 note, we observed that TSCs do not express XACT, in line with the hypothesis of Vallot et al
214 that XACT is lost upon TE specification in pre-implantation embryos (Vallot et al 2017).
215 Finally, EPSCs are either XaXi or XaXe (**Figure 2D and E**). Of note, one female EPSC line
216 was massively XaXe whereas the other one was XaXi (**Figure sup 1D**). In all cases, EPSCs are
217 in a post-XCI state, confirming their association with the primed state of pluripotency.
218 Altogether, our analysis revealed distinct X chromosome activity status in peri-implantation
219 stem cell models that resemble their *in vivo* counterpart.

220
221

222 **Metabolism-related genes distinguish EPSCs from PPSCs.**

223 DNA methylation and X-chromosome inactivation associate EPSCs with primed pluripotency.
224 However, we observed that EPSCs are more clonogenic and proliferate faster than PPSCs, as
225 previously reported (Yang et al., 2017). Additionally, we and others have shown that EPSCs
226 are able to directly convert into TSCs when switched to ASECRiAV medium (Castel et al.,
227 2020; Liu et al., 2017), whereas PPSCs require a priming treatment (Mischler et al., 2021; Wei
228 et al., 2021; Jang et al., 2022; Soncin et al., 2022 ; Viukov et al., 2022, Zorzan et al., 2022).

229

230 Pluripotency marker analysis of 3 datasets from 3 independent groups generating EPSCs was
231 performed. We looked at the expression of naive markers DPPA3, KLF4, KLF5, KLF17,
232 TFPC2L1, ZFP42 and ZFP57 as well as the expression of core pluripotency markers: FGF4,
233 GDF3, NANOG, POU5F1, SALL4, SOX2, TDGF1 and UTF1. In our study, EPSCs and PPSCs
234 express those genes similarly (**Figure sup 2A**). Of note, reanalysis of EPSCs transcriptomic
235 data from the original study by Yang et al. (Yang et al., 2017) revealed lower expression of
236 NANOG, POU5F1, UTF1 and ZFP42 and KLF5, compared to PPSCs (from the same study,
237 **Figure sup 2B**). We also reanalyzed Aksoy et al (Aksoy et al., 2021) transcriptomic data.
238 Looking at the pluripotency markers, we found that Aksoy's EPSCs express less NANOG,
239 DPPA3, GDF3, TDGF1, ZFP42, KLF5, SALL4 and POU5F1 when compared to PPSCs (from
240 the same study). On the other hand, NPSCs from the same study express more DPPA3, KLF5,
241 KLF17, TFPC2L1 and ZFP57 than PPSCs and EPSCs, as expected (**Figure sup 2C**).

242

243 To further decipher the link between EPSCs and PPSCs, we performed differential gene
244 expression analysis and plotted the results on a MA plot (**Figure 3A**). Among the 180
245 differentially expressed genes between our EPSCs and PPSCs, 110 are significantly

246 downregulated genes and less than 20 are expressed above 20 mRNA molecules per million of
247 mRNA molecules (see Mat & Met for cut-off). Among the most differentially expressed genes,
248 PIR and PKIB are potentially involved in nucleic acid homeostasis and MYH14 represents an
249 alternative myosin. Further studies are necessary to uncover the link between these pathways
250 and the survival of EPSCs. On the other hand, 70 genes are significantly upregulated in EPSCs
251 and 8 of them belong to the MT1/2 family (out of 14 members detected in our analysis) (**Figure**
252 **3B**). Additionally, we performed a pair-wise comparison of EPSCs with NPSCs and PPSCs
253 with NPSCs, which confirmed that the most differentially expressed genes are the same in both
254 comparisons (**Figure sup 3A, B**). We compared the differentially expressed genes between
255 EPSCs and PPSCs in our dataset and in Yang et al. and Aksoy et al. datasets. We found no
256 differential expression of these genes in these two other datasets, showing discrepancies
257 between cell types under similar culture conditions between laboratories (**Figure sup 3C**).

259 To further validate our EPSCs lines at the protein level, we performed a mass spectrometry
260 analysis on NPSCs, PPSCs and EPSCs. 91 of the 180 differentially expressed genes were
261 detected in the proteomics results. The expression pattern of the proteins more abundant in the
262 EPSCs mirrors the transcriptomic expression of the associated genes (**Figure 3C**). However,
263 there was more variability for the genes predominantly expressed in the PPSCs. Interestingly,
264 the specific expression of the MT1A, MT1E, MT1F, MT1G, MT1JP, MT1L, MT1M, MT2A
265 genes in the EPSCs (**Figure 3B**) prompted us to perform further characterization of metabolic-
266 linked genes in our stem cell models.

268 **Mitochondrial genes expression profile and metabolic activity distinguish EPSCs from** 269 **PPSCs**

270 We previously showed that electron transport chain-coding genes can distinguish naive from
271 primed PSCs (Kilens et al., 2018). We analyzed the expression levels of the proteins composing
272 the electron transport chain of the mitochondria. Of the 94 genes on our list, 81 were detected
273 by mass spectrometry in all our samples. At the transcriptome level, NPSCs stand out with
274 specific components of complex 1 (NADH dehydrogenase) while TSCs express overall low
275 levels of electron transport chains components. However, in the proteomic analysis, EPSCs
276 seem to be much closely related to NPSCs. Moreover, Complex 5 (ATP synthase) proteins also
277 seem to be prevalent in TSCs (**Figure 4A**). Although intriguing, this data is in line with several
278 reports showing discordance between mRNA and protein levels in several mammalian systems
279 (Edfors et al., 2016; Schwanhäusser et al., 2011) (Berg et al., 2023).

281 Subsequently, to link phenotypic measurements to molecular signatures, we measured oxygen
282 consumption rate (OCR) and extracellular acidification rate (ECAR) using mitoXpress and
283 pHXtra kits (**Figure 4B, Figure sup 4 A-D**). The presentation of data as OCR vs ECAR showed
284 that TSCs have an oxygenation rate and an extracellular acidification rate that are two times
285 higher than NPSCs and up to 4 times higher than PPSCs. Among PSCs, PPSCs have the lowest
286 metabolic activity, and EPSCs are similar to NPSCs for this criterion, correlating with the
287 expression of the MT1/2 expression (**Figure 4B, Figure sup 4 A-D**).

289 **Discussion**

291 Here, we report the comparison of TSCs, EPSCs, PPSCs and NPSCs cell lines with epigenetic
292 and functional readouts that have been previously used as hallmarks of NPSCs. Our results
293 clarify the fact that hallmarks of pluripotency are not predictive of each other. One current issue
294 in the human peri-implantation development field is the lack of biological reference. For
295 example, the metabolic status of each lineage within the human peri-implantation embryo is not

296 known yet. Considering stem cell characterization, another issue is that most hallmarks are
297 relative. A broad array of samples needs to be assessed to have a clear view and draw
298 conclusion. The sole comparison of 2 models is not sufficient to draw strong conclusions about
299 fate or stage. All in all, multiplying hallmarks alleviates stage matching biases.

300
301 Additional readouts unambiguously associated the EPSCs with a primed pluripotent fate.
302 Nevertheless, EPSCs clearly have a higher clonogenic propensity and growth rate than PPSCs.
303 We showed that EPSCs have a metabolic activity comparable to NPSCs, which raises a paradox
304 in the association of hallmarks such as chimerism and trophoblastic conversion with the fate of
305 the cells. Indeed, the ability of EPSCs to contribute to monkey and mouse chimeras (Liu et al.,
306 2021a; Tan et al., 2021; Yang et al., 2017; Zheng et al., 2021), to convert to TSCs (Castel et al.,
307 2020; Liu et al., 2021a) and to self-assemble in blastocyst-like 3D structures (Fan et al., 2021;
308 Sozen et al., 2021), questions the link between these assays with human PSCs and the
309 conclusions we draw about fate. This discrepancy could be linked to the “black-and-white” way
310 we report results, *e.g.*, determining that a cell can or cannot contribute to chimerism or convert
311 into TSCs. As outlined in the ISSCR standards for stem cell research, a more transparent and
312 accurate way would be to acknowledge that indeed EPSCs can survive in animal embryos but
313 do not invade as much as other cell state, or that EPSCs convert into TSCs after direct media
314 transfer but at a lower rate than NPSCs. The specific behavior of EPSCs seems to uncouple
315 survival ability to fate and developmental fate matching (Aksoy et al., 2021, Castel et al., 2020).
316 Nevertheless, the enhanced clonogenicity and growth rate of EPSCs offer new opportunities to
317 understand the link between these features and pluripotency, which could open new
318 perspectives for large-scale experiments with pluripotent cells, but also help us to explore the
319 chimerism mechanisms and thus improve them.

320
321 The characterization of TSCs and TELCs clarifies the use of each model and provides a new
322 reference for trophoblast models regarding their hallmarks. TSCs have an inactivated X and
323 would correspond to cells that have low global DNA methylation levels, comparable to
324 cytotrophoblast before 10 days post-fertilization and epiblast before 8 days post-fertilization
325 (Santos et al., 2010; Zhou et al., 2019). NPSCs to TSCs differentiation would also be an
326 interesting model to study the methylation waves in the embryo. Finally, defining TSCs
327 metabolic activity will help to design media that would better support post-implantation
328 development, which needs to be improved for both human embryos and blastoids.

329
330 Altogether, systematic comparison of stem cell models is a powerful way to learn new features
331 of peri-implantation development together with the hallmarks specific of each fate and stage.
332 The set of hallmarks we used enabled a clearer characterization of TSCs and EPSCs along
333 NPSCs and PPSCs. The variation between states offers the opportunity to decipher links
334 between functions that are difficult to uncouple, such as cell cycle and fate potential. Detailed
335 hallmarks also instruct on the relevance of each model to human development. A better
336 understanding of human peri-implantation development using both 2D and 3D models will
337 further deepen our understanding of early pregnancy and help to design and optimize *in vitro*
338 fertilization techniques.

339 340 **Acknowledgements**

341 We thank our colleagues N. Rivron, A. Camus and P. Rugg-Gun for their insights. We thank
342 Y. Takashima for RNA samples of ACE-TSCs. This work was supported by Theramex and
343 ANR-22-CE91-0010. We thank the GenoBIRD, MicroPICell and iPSC core facilities, all
344 supported by IBiSA and Biogenouest, for the use of their resources and technical support.
345 MicroPICell is a member of the national infrastructure France-Bioimaging (ANR-10-INBS-

346 04). Bird is a member of Institut Français de Bioinformatique (IFB, ANR-11-INBS-0013). We
347 thank P. Aubert for the access to Synergy H1 plate reader. We thank M. Schwalfenberg
348 (Agilent) for her support for the analysis of the MitoXpress and pHXtra kits. We thank Bart
349 Ghesquière and the Metabolomics Expertise Center (VIB-KU Leuven). Research in the Pasque
350 laboratory was supported by the Research Foundation– Flanders (FWO; Odysseus Return Grant
351 G0F7716N to V.P.; FWO grants G0C9320N and G0B4420N to V.P.), FWO EOS grant G0I7822N
352 to V.P. and the KU Leuven Research Fund (C1 grant C14/21/119 to V.P.), and FWO PhD
353 fellowships to R.N.A. (11L0722N). KM was supported by a fellowship from the French
354 Medical Research Foundation (FRM SPF201809006928) and the LabEx ‘Who Am I?’ (ANR-
355 11-LABX-0071). Research in the Rougeulle laboratory was supported by the Agence Nationale
356 pour la Recherche (ANR-14-CE10-0017 and ANR-19-CE12-0018-01) and by Ligue Nationale
357 contre le Cancer.

358

359 **Author contributions**

360 CO and LD designed the study and wrote the manuscript with input from all authors. CO and
361 EM performed IF analysis and cell culture experiments. SC and CO performed bioinformatics
362 analysis. KM, JFO and CR performed the FISH experiment and helped with the interpretation.
363 IA provided the H9 EPSC line. VP, GG and RA performed the methylation experiments. CO,
364 OR and CP performed the metabolism experiment. CO, OG, RL and CP performed the mass
365 spectrometry experiment. All authors approved the final version of the manuscript.

366

367 **Declaration of interests**

368 Authors declare no conflict of interest.

369

370 Further information and requests for resources and reagents should be directed to and will be
371 fulfilled by the corresponding author L. David (laurent.david@univ-nantes.fr).

372 **Figure legends**

373

374 **Figure 1. Transcriptomic analysis of human PSC, EPSC, NPSC, TSC, TELC lines.**

375 (A) Heatmap of Pearson correlation coefficients of PSCs, EPSCs, NPSCs, TSCs lines along with
376 TELC differentiation from NPSCs and TSCs lines from Io et al study Samples are clustered from
377 the Euclidian distance of correlation, by a hierarchical clustering using Ward's method.

378 (B) Gene expression levels of indicated lineage markers are shown for NPSCs, the different
379 days of differentiation of NPSCs into TELC and TSCs lines. The NPSCs and TSCs lines are
380 included as control. Expression levels are given as number of transcripts per million of mRNA
381 molecules. In each boxplot, the top and bottom of the box represent the third and first
382 quartile, respectively; the band represents the median (second quartile); and error bars show
383 the interquartile range (IQR) (lower bound: $Q1 - 1.5 \times IQR$; upper bound: $Q3 + 1.5 \times IQR$).
384 Pvalues from differential gene expression analysis were re-used for boxplots. Asterisks
385 indicate statistical significance of the difference compared to NPSCs: *pvalue <0.05.

386 (C) Immunofluorescence images of day 4 and day 5 of TELC differentiation from NPSCs stained
387 for trophoblast-associated transcription factors GATA3 and NR2F2 and pluripotency-
388 associated transcription factor NANOG. Nuclei were stained with DAPI. Scale bar = 100 μ m.

389 (D) Immunofluorescence images of day 4 and day 5 of TELC differentiation from NPSC stained
390 for trophoblast-associated transcription factors GATA3 and CDX2 and pluripotency-associated
391 transcription factor NANOG. Nuclei were stained with DAPI. Scale bar = 100 μ m.

392

393

394 **Figure 2. DNA methylation and X chromosome coating of PSC, EPSC, NPSC and TSC**

395 (A) Gene expression levels of indicated genes are shown for NPSCs, PPSCs, EPSCs, TELCs and
396 TSCs lines. The NPSCs and PPSCs lines are included as control. Expression levels are given as
397 number of transcripts per million of mRNA molecules. In each boxplot, the top and bottom of
398 the box represent the third and first quartile, respectively; the band represents the median
399 (second quartile); and error bars show the interquartile range (IQR) (lower bound: $Q1 - 1.5 \times$
400 IQR ; upper bound: $Q3 + 1.5 \times IQR$). Pvalues from differential gene expression analysis were re-
401 used for boxplots. Asterisks indicate statistical significance of the difference compared to
402 PPSCs: *pvalue <0.05.

403 (B) Heatmap of relative expression of DNMT1, DNMT3A, DNMT3B and DNMT3L proteins.

404 (C) 5mC content is expressed as the percentage of 5mC in the total pool of cytosine for the
405 indicated cell lines. Significance levels were determined using a Kruskal Wallis test, followed
406 by a Dunn comparison. Asterisks indicate statistical significance of the difference: * pvalue
407 <0.05.

408 (D) mRNA FISH analysis for XIST, XACT and HUWE1. Scale bar = 10 μ m.

409 (E) Quantification of XaXa, XaXi and XaXe patterns. More than 100 cells were investigated for
410 their nuclear expression for each cell line represented.

411

412

413 **Figure 3. Refining gene signature distinguishing EPSCs and PSCs**

414 (A) MA plots of EPSCs and PPSCs represent the log₂ fold Change of gene by their mean
415 obtained from differential gene expression analysis between two cell type annotations. Genes
416 are colored if their adjusted p-value is under 0.05 and if the fold change is greater than 2 or
417 lower than -2.

418 (B) Gene expression levels of indicated genes are shown for NPSCs, PPSCs and EPSCs lines. The
419 NPSCs and PPSCs lines are included as control. Expression levels are given as number of
420 transcripts per million of mRNA molecules. In each boxplot, the top and bottom of the box
421 represent the third and first quartile, respectively; the band represents the median (second
422 quartile); and error bars show the interquartile range (IQR) (lower bound: $Q1 - 1.5 \times IQR$;
423 upper bound: $Q3 + 1.5 \times IQR$). P values from differential gene expression analysis were re-used
424 for boxplots. Asterisks indicate significance compared to PPSCs: * p value <0.05.
425 (C) Heatmap of relative expression of gene (left) (analyzed by DGE-seq) and associated
426 proteins (right) (analyzed by mass spectrometry) differentially expressed between EPSCs and
427 PPSCs samples.

428

429 **Figure 4. Metabolic activity of PSC, EPSC, NPSC and TSC**

430 (A) Heatmap of relative expression of gene (left) (analyzed by DGE-seq) and associated
431 proteins (right) (analyzed by mass spectrometry) of electron transport chains in EPSCs and
432 PPSCs samples. Genes were classified by mitochondrion complex and hierarchically clustered.
433 (B) Oxygen consumption rate and extracellular acidification rate of NPSCs, PPSCs, TSCs and
434 EPSCs were measured using MitoXpress and pHXtra kits. This figure presents 4 technical
435 replicates.

436

437 **Figure sup 1. Associated with Figure 1 and 2**

438 (A) Heatmap of Pearson correlation coefficients of PSCs, EPSCs, NPSCs, TSCs lines along with
439 TELC differentiation from NPSCs and TSCs lines from Io et al study. Samples are clustered from
440 the Euclidean distance of correlations, by a hierarchical clustering using Ward's method and
441 split according to their cell type (NPSC, PPSC, EPSC, TELC, ASECRiAV-TSC, ACE-TSC).
442 (B) Quantification of GATA3, CDX2 and NR2F2 levels from day 4, 5 and 6 of NPSCs
443 differentiation into TELCs.
444 (C) 5mC content is expressed as the percentage of 5mC in the total pool of cytosine for the
445 indicated cell lines in the indicated replicate. Significance levels were determined using a
446 Kruskal Wallis test, followed by a Dunn comparison. Asterisks indicate statistical significance
447 of the difference: * pvalue <0.05.
448 (D) Quantification of XaXa, XaXi and XaXe patterns in each replicate. More than 100 cells were
449 investigated for their nuclear expression for each cell line represented.

450

451 **Figure sup 2. Associated with Figure 3**

452 (A) Gene expression levels of indicated lineage markers are shown for NPSCs, PPSCs and EPSCs
453 lines from our DGE-seq dataset. The NPSCs and PPSCs lines are included as control. Expression
454 levels are given as number of transcripts per million of mRNA molecules. In each boxplot, the
455 top and bottom of the box represent the third and first quartile, respectively; the band
456 represents the median (second quartile); and error bars show the interquartile range (IQR)
457 (lower bound: $Q1 - 1.5 \times IQR$; upper bound: $Q3 + 1.5 \times IQR$). Pvalues from differential gene
458 expression analysis were re-used for boxplots. Asterisks indicate statistical significance of the
459 difference compared to PPSCs: *pvalue <0.05.
460 (B) Gene expression levels of indicated lineage markers are shown for PPSCs and EPSCs lines
461 from Yang et al study. The PPSCs lines are included as control. Expression levels are given as
462 number of reads per kilobase Million. In each boxplot, the top and bottom of the box
463 represent the third and first quartile, respectively; the band represents the median (second
464 quartile); and error bars show the interquartile range (IQR) (lower bound: $Q1 - 1.5 \times IQR$;

465 upper bound: $Q3 + 1.5 \times IQR$). Pvalues from differential gene expression analysis were re-used
466 for boxplots. Asterisks indicate statistical significance of the difference compared to PPSCs:
467 *pvalue <0.05.

468 (C) Gene expression levels of indicated lineage markers are shown for NPSCs, PPSCs and EPSCs
469 lines from Aksoy et al study. The PPSCs lines are included as control. Expression levels are
470 given as number of transcripts per million of mRNA molecules. In each boxplot, the top and
471 bottom of the box represent the third and first quartile, respectively; the band represents the
472 median (second quartile); and error bars show the interquartile range (IQR) (lower bound: $Q1$
473 $- 1.5 \times IQR$; upper bound: $Q3 + 1.5 \times IQR$). Pvalues from differential gene expression analysis
474 were re-used for boxplots. Asterisks indicate statistical significance of the difference
475 compared to PPSCs: *pvalue <0.05.

476

477 **Figure sup 3. Associated with Figure 3**

478 (A) MA plots of NPSCs and EPSCs represent the log₂ fold Change of gene by their mean
479 obtained from differential gene expression analysis between two cell type annotations. Genes
480 are colored if their adjusted p-value is under 0.05 and if the fold change is greater than 2 or
481 lower than -2.

482 (B) MA plots of NPSCs and PPSCs represent the log₂ fold Change of gene by their mean
483 obtained from differential gene expression analysis between two cell type annotations. Genes
484 are colored if their adjusted p-value is under 0.05 and if the fold change is greater than 2 or
485 lower than -2.

486 (C) Heatmap of relative expression of gene differentially expressed between EPSCs and PPSCs
487 samples in (from left to right) our dataset, Yang et al dataset and Aksoy dataset.

488

489 **Figure sup 4. Associated with Figure 4**

490 (A-D) Oxygen consumption rate and extracellular acidification rate of NPSCs, PPSCs, TSCs and
491 EPSCs were measured using MitoXpress and pHXtra kits. Each graphs represents one iteration
492 of the experiment. Each dot represents one replicate in this iteration.

493

494

495 **STAR METHODS**

496

497 **RESOURCE AVAILABILITY**

498

499 **Lead Contact**

500 Further information and requests should be directed to the Lead Contact, Laurent DAVID

501 (laurent.david@univ-nantes.fr).

502

503

504 **Materials Availability**

505 This study did not generate new unique reagents.

506

507 **Data and code availability**

508 All original data have been deposited on European Nucleotide Archive under accession number

509 PRJEB63637.

510 The source code can be retrieved by following the links below.

511 scRNAseq alignment pipeline:

512 https://gitlab.univ-nantes.fr/E114424Z/SingleCell_Align

513 Data preprocessing script and analysis scripts are available at the following link:

514 <https://gitlab.univ-nantes.fr/E198672Y/onfray-et-al-2022>.

515 All other parts of the code are available upon request.

516

517

518 **Experimental model and subject details**

519

520 **Cell lines**

521 All cell lines used in this study are described in **Table sup3**.

522 In brief, induced naïve, primed or extended PSC lines were reprogrammed from fibroblasts:

523 L71 from a 51-year-old healthy man; L80 from a 57-year-old healthy woman; MIPS220 from

524 a healthy female in her 30's (Castel et al, Kilens et al). human embryonic stem cells H9 (WA09)

525 were imported and used with authorization RE17-007R from the French oversight committee,

526 Agence de la Biomédecine. H9-EPSC were generated by I. Aksoy (ref). Naïve H9 ere generated

527 in C. Rougeulle lab (Vallot et al). TSCs lines we used were generated by Okae et al., or Castel

528 et al.

529

530 For TELC differentiation experiments we used (lesquelles) NPSCs lines from Kilens et al.

531

532

533 **METHODS DETAILS**

534

535 **Tissue culture - maintenance**

536 All cell lines were cultured at 37 °C, either under hypoxic (5% O₂, 5% CO₂) or normoxic

537 conditions (20% O₂, 5% CO₂) as indicated. Culture medium was daily replaced. 10 μM

538 Y27632 (Axon Medchem) was added to the culture medium upon single-cell seeding of all

539 human stem cells. PXX indicates passage number. All cell lines were tested negative for

540 mycoplasma using the MycoAlert kit (LONZA, LT07-318).

541

542 Mouse embryonic fibroblasts (MEFs) were prepared from E13.5 pups that were decapitated,

543 eviscerated, dissociated with 0.25% trypsin, 0.1% EDTA and plated in MEF medium [DMEM

544 high glucose (Thermo Scientific), Glutamax 1:100 (GIBCO), 0.5% of penicillin–streptomycin

545 (Life Technologies)] on 0.1% gelatin-coated plates. MEFs were mitotically inactivated using
546 0.01mg/ml mitomycin C (Sigma-Aldrich) to be used as feeder cells. MEF isolation was
547 performed in compliance with the French law and under supervision of the UTE animal core
548 facility, Nantes Université.

549
550 TSCs were cultured on MEF feeder cells in ASECRiAV medium (Okoe et al., 2018)
551 [DMEM/F12 (GIBCO) supplemented with 0.1mM 2-mercaptoethanol (GIBCO), 0.2% FBS,
552 0.5% penicillin-streptomycin, 0.3% Bovine Serum Albumin (BSA, Sigma-Aldrich), 1%
553 Insulin-Transferrin-Selenium-Ethanolamine supplement (ITS-X, GIBCO), 1.5 mg/ml L-
554 ascorbic acid (Sigma-Aldrich), 50 ng/ml hEGF (Miltenyi Biotec), 2 μ M CHIR99021 (Axon
555 Medchem), 0.5 μ M A83-01 (Tocris), 1 μ M SB431542 (Tocris), 0.8 mM valproic acid (Sigma-
556 Aldrich) and 5 μ M Y27632]. TSCs could be passaged with TrypLE (5-10 min, 37°C, Life
557 Technologies) every 4 to 5 days at a cell density between 1.04×10^4 and 2.08×10^4 cells per
558 cm^2 . TSCs were routinely cultured at 37°C in hypoxic conditions.

559
560 NPSCs were cultured on MEF feeder cells in t2iLGöY medium (Takashima et al., 2014)
561 [DMEM/F12 supplemented with 1% N2 (GIBCO), 1% B27 (GIBCO), 1% non-essential amino
562 acids, 1% GlutaMAX (GIBCO), 0.1 mM 2-mercaptoethanol, 50 μ g/ml BSA, 0.5% penicillin-
563 streptomycin, 1 μ M CHIR99021, 1 μ M PD0325901 (Axon Medchem), 20 ng/ml mLIF
564 (Miltenyi Biotec), 5 μ M Gö6983 (Axon Medchem) and 10 μ M Y27632] or PXGL(Bredenkamp
565 et al., 2019) medium [47.5% Neurobasal medium (GIBCO) and 47.5% DMEM/F12 (GIBCO)
566 supplemented with 1mM N2 (GIBCO), 2mM B27 (GIBCO), 1mM GlutaMAX (GIBCO), 1mM
567 non-essential amino acids (GIBCO), 0.33% BSA, 1mM sodium pyruvate (GIBCO), 0.1% 2-
568 mercaptoethanol (GIBCO), 0.5% penicillin-streptomycin, 1 μ M PD0325901 (Axon Medchem),
569 2 μ M XAV939 (Axon medchem), 2 μ M Gö6983 (Axon medchem), 10ng/mL hLif (Peprotech),
570 10 μ M Y27632 (Axon Medchem)]. NPSCs were passaged every 4 to 5 days at a cell density of
571 2.08×10^4 cells per cm^2 using TrypLE (5 min, 37°C, Life Technologies). NPSCs were
572 routinely cultured at 37 °C in hypoxic conditions.

573
574 EPSCs were cultured on MEF feeder cells in LCDM medium(Yang et al., 2017) [48%
575 DMEM/F12 and 48% Neurobasal (GIBCO) supplemented with 0.5% N2 supplement, 1% B27
576 supplement minus vitamin A (GIBCO), 1% non-essential amino acids, 0.1 mM 2-
577 mercaptoethanol, 0.5% penicillin-streptomycin, 5% knockout serum replacement (KSR,
578 GIBCO), 10 ng/ml human LIF (Miltenyi Biotec), 1 μ M CHIR99021, 2 μ M (S)-(+)-
579 Dimethindene maleate (Tocris) and 2 μ M Minocycline hydrochloride (Tocris), 1 μ M IWR-
580 endo-1 (Miltenyi Biotec) and 2 μ M Y-27632]. EPSCs were passaged every 4 to 5 days at a 1:20
581 to 1:40 split ratio using TrypLE (5 min, 37°C, Life Technologies). EPSCs were routinely
582 cultured at 37°C in normoxic conditions.

583
584 Primed PSCs were cultured on matrigel 0.1% in mTeSR1 medium (StemCell technologies).
585 Colonies were manually divided every 5 to 6 days for passage, following 3min in gentle cell
586 dissociation reagent (Miltenyi Biotec) and seeded as small clumps in a new MW6 well. Primed
587 PSCs were routinely cultured at 37°C in normoxic conditions.

588

589 **Tissue culture - TELC induction of NPSCs**

590 Differentiation of NPSCs into trophectoderm-like cells was performed according to Guo et al.
591 protocol (Guo et al., 2021): NPSCs are passaged with TrypLE (5 min, 37°C, Life Technologies)
592 and plated in 24-well plates on Geltrex (0,5mL per cm^2 , Gibco) at 1:1 ratio. Before plating, a
593 MEFs exclusion is performed on a 0,1% gelatine (Sigma) coated 6-well plate for 30 min in
594 PXGL + Y27632. At day -1, the cells are cultured in PXGL + Y27632. From day 0, the medium

595 is changed into N2B27 supplemented with 1 μ M PD0325901 (Axon Medchem) and 1 μ M A83-
596 01 (Tocris). Cells are cultured for up to 5 days at 37°C in hypoxic conditions.

597

598 **DNA methylation**

599 We tested DNA methylation in three batches composed of, for batch 1: 3 technical replicates
600 of NPSCs (M2A8), 2 technical replicates of EPSCs (E80), 3 biological replicates with each 2
601 technical replicates for TSC (CT30, AV03, AV23), 3 technical replicates of PPSCs (MIPS220).
602 For batch 2: 2 technical replicates of NPSCs (M2A8), 1 technical replicate of EPSCs (E80), 3
603 biological replicates with each 2 technical replicates for TSC (CT30, AV03, AV23), 1 technical
604 replicate of PPSCs (MIPS220). For batch 3: 2 technical replicates of NPSCs from (Zijlmans et
605 al., 2022), 3 biological replicates of EPSCs (E80, E71, H9), 1 technical replicate for TSC
606 (AV23), 2 biological replicates of PPSCs (H9, MIPS220). For mass spectrometry analysis of
607 DNA methylation, DNA was extracted using the genomic DNA columns (Qiagen). 1 μ g of
608 genomic DNA was analyzed using liquid chromatography triple-quadrupole mass spectrometry
609 (KU Leuven Metabolomics Core). The concentration (μ M) of Cytosine (unmodified), 5mC and
610 5-hydroxymethyl-cytosine (5hmC) were obtained using standard curves of known C, 5mC and
611 5hmC amounts. The percentage of 5mC or 5hmC in DNA was obtained by calculating the ratio
612 of 5mC or 5hmC to the total pool of C.

613

614 **RNA-FISH**

615 RNA-FISH was performed as previously described (Kilens et al, 2018). Briefly, cells were fixed
616 between 24h and 50 h post seeding in 3% paraformaldehyde for 10 min at room temperature.
617 Cells were permeabilized in CSK buffer supplemented with 1 mM EGTA, 0.5% Triton and
618 VRC (200 μ M) for 5 min on ice. After 3 washes in 70% EtOH, cells were dehydrated in 90%
619 and 100% EtOH and incubated overnight with probes at 37 °C. After three 50%
620 formaldehyde/2 \times SSC washes and three 2 \times SSC washes at 42 °C for 4 min, coverslips were
621 mounted in Vectashield plus DAPI. SpectrumGreen or SpectrumRed-labeled probes (Vysis)
622 were generated by nick translation for human XIST (10 kb Exon 5-6, gift from Dr. Edith Heard,
623 EMBL, Germany), XACT (RP11-35D3, BACPAC) and HUWE1 (RP11-42M11, BACPAC
624 Resource). Images were acquired on an inverted Nikon A1 confocal microscope, according to
625 the Shannon–Nyquist sampling rate. mRNA expression of XIST, XACT and HUWE1 are
626 manually counted in more than 100 cells per cell line: 100 cells are randomly chosen for each
627 sample. The number of dots per cell for each channel is quantified. Dots are considered when
628 present on more than 2 z-planes. Dots on different channels are considered on the same
629 chromosome only when less than 2 μ m from each other. Cells with a pattern that appears to
630 show more than two distinct dots in at least one channel are considered as “chromosomal
631 abnormalities”. Cells with a pattern without HUWE1 are pooled in “Others”. Cells with no
632 visible dots are pooled in “Nothing”. At least two biological replicates are done for each cell
633 line (table sup4).

634

635 **Oxygen consumption rate and extracellular acidification rate**

636 To assess metabolic activity of the human peri-implantation cellular models, we measured the
637 oxygen consumption rate and the extracellular acidification rate.

638 Oxygen consumption rate was measured by measuring fluorescence signal coupled with oxygen
639 concentration, using MitoXpress kit (Agilent). Extracellular acidification rate was measured by
640 measuring fluorescence signal coupled with pH variation, using pHXtra kit (Agilent). Both
641 analyses were performed conjointly. Slopes of variation of fluorescence over 30min were
642 extracted for the analysis. Synergy H1 plate reader, with dual read TR-F was used for the
643 measures.

644 The experiment was conducted 4 times, on 4 technical replicates, and included 2 biological
645 replicates for EPSCs (E80, E71), 3 for TSCs (CT30, AV03, AV23), with 1 biological replicate
646 of NPSCs (M2A8) and 3 biological replicates of PPSCs (LON80, LON71, MIPS220) as
647 controls

648 Before seeding, black sides clear bottom plates were pretreated for 20min with 2 µg/mL of Cell-
649 Tak Cell and tissue adhesive (Corning).

650 NPSCs, PPSCs, EPSCs and TSCs were dissociated using TrypLE for 5min at 37°C. NPSCs,
651 EPSCs and TSCs were incubated on gelatin (0.1%) for 30 min at 37 °C to remove feeder cells.
652 Cells were resuspended in DMEM (Sigma Aldrich) supplemented with 10 mM glucose, 2 mM
653 glutamine, 2 mM pyruvate and pH was adjusted to 7.4 using NaOH.

654 A cell seeding density titration experiment was performed before the first experiment to
655 determine the optimal quantity of cells for the experiments. Respectively, NPSCs, PPSCs,
656 EPSCs and TSCs were seeded at 0.5625×10^6 , 1.125×10^6 , 0.5156×10^6 and 0.5156×10^6
657 cells per cm².

658 To control that measurements were performed appropriately, we included controls to check the
659 maximal respiratory capacity (FCCP, 0.75 µM), inhibition of oxidative phosphorylation
660 (Antimycin-A, 2 µM and Rotenone, 1 µM), and inhibition of glycolysis (2DG, 50mM). Those
661 controls validated that our measurements were within detection limits and metabolic capacity
662 of the cells (data not shown).

663

664 **3'SRP**

665 Total RNA molecules were extracted from cells with RNeasy-Mini Kits (Qiagen). Protocol of
666 3' SRP RNA sequencing was performed as previously described in (Charpentier et al., 2021).
667 Libraries were then sequenced on a NovaSeq 6000 (Illumina). Data were aligned along the
668 human genome reference (hg19) and a count matrix was generated by counting sample specific
669 UMI associated with genes for each sample. Samples with less than 200 000 UMI and less than
670 5000 genes expressed were excluded of the analysis. Then, a batch correction between samples
671 of different experiments was applied. A Principal Component Analysis (PCA) was performed
672 in order to visualized samples repartition by reducing the number of dimensions. Correlation
673 between samples were assessed with Pearson's linear correlation heatmaps. Higher correlations
674 are marked in yellow and lower correlations are in red. Differentially expressed genes between
675 conditions were calculated using R package Deseq2 (Bioconductor) by first applying a variance
676 stabilizing transformation (vst). Genes with adjusted p-value inferior to 0.05 and with a fold
677 change superior to 2 or inferior to -2 were considered as differentially expressed genes. Gene
678 expressions were visualized with heatmaps that were generated by center genes expression.
679 Finally, pathways analysis was performed: R package "Fgsea" and databases such as Kegg,
680 Reactome and Gene Ontology were used to identify significantly enriched or depleted groups
681 of genes in each condition.

682

683 Of note, 3'SRP, the RNAseq method we are using, allows us to correlate the expression level
684 with the likeness of protein to be expressed (Girard et al., 2023). Indeed, up to 70.45% genes
685 are identified by MS/MS analysis when their expression level is above 20 mRNA molecules
686 per million of mRNA molecules.

687

688 **Immunostaining.**

689 For immunofluorescence (IF) analysis, cells were fixed at room temperature using 4%
690 paraformaldehyde for 15 min. Samples were then permeabilized for 60 min at room temperature
691 with IF buffer [phosphate-buffered saline (PBS), 0.2% Triton, 10% FBS], which also served as
692 a blocking solution. Samples were incubated with primary antibodies overnight at 4 °C. The
693 following antibodies were used: anti-GATA3 (1:300, R&D® AF2605), anti-NR2F2 (1:300,

694 Abcam® ab211776), anti-CDX2 (1:300, Abcam® ab157524). Incubation with secondary
695 antibodies was performed for 2 h at room temperature along with 4',6-diamidino-2-
696 phenylindole (DAPI) nuclei staining. Confocal immunofluorescence images were acquired
697 with A1-SIM Nikon® confocal microscope. Optical sections of 0.5-1 µm-thick were collected.
698 Images were processed using Volocity® visualization software and Fiji software (<http://fiji.sc>).
699

700 **MASS SPECTROMETRY**

701 **Cell culture for mass spectrometry**

702 For all samples (except EPSCs) we performed 2 types of sample preparation: Lysis after TrypLE
703 dissociation or Lysis without TrypLE dissociation.

704
705 NPSCs: 1 day prior lysis (4-5 days after seeding), NPSCs, were dissociated using TrypLE for
706 5min at 37°C. NPSCs were incubated on gelatin (0.1%) for 1h at 37 °C to remove feeder cells.
707 NPSCs were plated overnight respectively on a 0.1% geltrex coated plate (1/1 ratio). The next
708 day, cells were rinsed with PBS-/- before being lysed using the iST kit (PreOmics GmbH,
709 Planegg, Germany) following the manufacturer's instructions.

710 The lysis day, another well of NPSCs was dissociated using TrypLE for 5min at 37°C. NPSCs
711 were incubated on gelatin (0.1%) for 1h at 37 °C to remove feeder cells. Then, NPSCs cells
712 were rinsed with PBS -/- before being lysed using the iST kit (PreOmics GmbH, Planegg,
713 Germany) following the manufacturer's instructions.

714
715 TSCs: 1 day prior lysis (4-5 days after seeding) TSCs were dissociated using TrypLE for 5min
716 at 37°C. TSCs were incubated on gelatin (0.1%) for 1h at 37 °C to remove feeder cells. TSC
717 were plated overnight respectively on a 3 µg/mL vitronectin and 1 µg/mL laminin coated plate
718 (1/1 ratio). The next day, cells were rinsed with PBS-/- before being lysed using the iST kit
719 (PreOmics GmbH, Planegg, Germany) following the manufacturer's instructions.

720 The lysis day, another well of TSCs was dissociated using TrypLE for 5min at 37°C. TSCs
721 were incubated on gelatin (0.1%) for 1h at 37 °C to remove feeder cells. Then, TSCs cells were
722 rinsed with PBS -/- before being lysed using the iST kit (PreOmics GmbH, Planegg, Germany)
723 following the manufacturer's instructions.

724
725 PPSCs: The lysis day, PPSCs were dissociated using TrypLE for 5min at 37°C. PPSCs were
726 rinsed with PBS-/- before being lysed using the iST kit (PreOmics GmbH, Planegg, Germany)
727 following the manufacturer's instructions.

728 The lysis day, another well of PPSCs was rinsed with PBS -/- before being lysed using the iST
729 kit (PreOmics GmbH, Planegg, Germany) following the manufacturer's instructions.

730
731 EPSCs: The lysis day, EPSCs were dissociated using TrypLE for 5min at 37°C. EPSCs were
732 incubated on gelatin (0.1%) for 1h at 37 °C to remove feeder cells. Then, EPSCs were rinsed
733 with PBS -/- before being lysed using the iST kit (PreOmics GmbH, Planegg, Germany)
734 following the manufacturer's instructions.

735 **Protein extraction and digestion**

736 Samples were thawed and lysed (denatured, reduced and alkylated) for 10min at 95°C and then
737 Trypsin/LysC digested for 3h at 37°C. Purification of peptides was then carried out at room
738 temperature on a spin cartridge, and peptides were finally eluted with the iST Fractionation
739 Add-on (PreOmics GmbH, Planegg, Germany) in three fractions in 10µL of a LC-loaded buffer.

740 Simultaneously, a protein assay has been realized to quantify proteins present in the samples.
741 Once purified, the three fractions of each cell type (hNPSCs, hPSCs, hEPSCs and hTSCs
742 samples) were prepared for mass spectrometry injection at approximately $3\mu\text{g}$ of protein in
743 $10\mu\text{L}$.

744 **Nanoliquid Chromatography Coupled with Tandem Mass Spectrometry (NanoLC-** 745 **MS/MS)**

746 The sample from Girard et al, 2023 (Girard et al., 2023) were analyzed in Data-Dependent
747 Analysis (DDA) and Parallel Accumulation Serial Fragmentation (PASEF) mode to generate
748 the spectral library. Each sample of enzymatically digested plasma proteins (about 200 to
749 300ng) were separated on a $75\mu\text{m} \times 250\text{mm}$ IonOpticks Aurora 3 C18 column (Ion Opticks Pty
750 Ltd., Bundoora, Australia). A gradient of reverse phase buffer (Buffer A: 0.1% formic acid, 2%
751 acetonitrile, 97.9% H_2O ; Buffer B: 0.1% formic acid, 99.9% acetonitrile) was run on a
752 nanoElute UHPLC System (Bruker Daltonik GmbH, Bremen, Germany) at a flow rate of
753 $250\text{nL}/\text{min}$ at 50°C controlled by HyStar software (v6.0.30.0, Bruker Daltonik). The liquid
754 chromatography (LC) run lasted for 80min. A starting concentration of 2% buffer B increasing
755 to 13% over the first 42 minutes was first performed and buffer B concentrations were increased
756 up to 20% at 65min; 30% at 70min; 85% at 75min and finally 85% for 5min to wash the column.
757 The temperature of the ion transfer capillary was set at 180°C . Ions were accumulated for
758 100ms, and mobility separation was achieved by ramping the entrance potential from -160V
759 to -20V within 114ms. MS and MS/MS mass spectra were acquired with average resolutions
760 50.000 FWHM full width at half maximum (with a m/z range of 100 to 1700), respectively. To
761 enable the PASEF method, precursor m/z and mobility information was first derived from full
762 scan TIMS-MS experiments (with a mass range of m/z 100–1700). The quadrupole isolation
763 width was set to 2 and 3 Th and, for fragmentation, the collision energies varied between 31
764 and 52 eV depending on the precursor mass and charge. TIMS, MS operation and PASEF were
765 controlled and synchronized using the control instrument software OtofControl 6.2.5 (Bruker
766 Daltonik). LC-MS/MS data were acquired using the PASEF method as described previously
767 (Banliat et al, 2019) with a total cycle time of 1.31s, including 1 TIMS MS scan and 10 PASEF
768 MS/MS scans. The 10 PASEF scans (100ms each) contained, on average, 12 MS/MS scans per
769 PASEF scan. Ion mobility-resolved mass spectra, nested ion mobility vs. m/z distributions, as
770 well as summed fragment ion intensities were extracted from the raw data file with
771 DataAnalysis 6.0 (Bruker Daltonik) (Banliat et al., 2020).

772 The three fractions per samples were then analysed individually in diaPASEF mode. Each
773 tryptic peptide sample, of approximately 400-500ng each, was analyzed under the same
774 conditions as described above. These included the same analytical conditions (identical
775 instrumentation, type of separation column and gradient length) and analysis on the same
776 instrument (timsTOF Pro; Bruker Daltonik GmbH, Bremen, Germany)). For the development
777 of the diaPASEF method, we used a method with an adapted instrument firmware to perform
778 data-independent isolation of data from several 25 m/z wide precursor windows, also called
779 segments, in a single TIMS separation (107.5ms). We used a method with two boxes per
780 segment in each 107.5ms diaPASEF scan, i.e. a total of thirty-two segments and sixty-four
781 boxes, of which sixteen of these scans perfectly cover the diagonal area of doubly and triply
782 charged peptides in the m/z and ion mobility output range. MS and MS/MS data were collected
783 over the m/z range 100 - 1700 and over the mobility range from $1/\text{K}0 = 0.6\text{Vs.cm}^{-2}$ to $1/\text{K}0 =$
784 1.6Vs.cm^{-2} . During each data collection, each TIMS cycle was 1.25 seconds long and
785 comprised 1 MS and 22 cycles of diaPASEF MS/MS segments, comprising 2, 3 or 4 boxes, to
786 cover a total of 64 boxes defined in the acquisition method. The collision energy was increased
787 linearly with mobility from 68eV at $1/\text{K}0=1.6\text{Vs.cm}^{-2}$ to 25 eV at $1/\text{K}0=0.6\text{Vs.cm}^{-2}$.

788 **MS Data processing**

789 Ion mobility resolved mass spectra, nested ion mobility versus m/z distributions, and fragment
790 ion intensity sums were extracted from the raw data file with DataAnalysis 6.0 (Bruker
791 Daltonik). The signal-to-noise ratio was increased by summing the individual TIMS scans.
792 Mobility peak positions and half-peak widths were determined on the basis of extracted ion
793 mobilograms (EIM, $\pm 0.05\text{Da}$) using the peak detection algorithm implemented in the
794 DataAnalysis software. Feature detection was also performed using DataAnalysis 6.0 software;
795 stored at the raw data level.

796 **Data analysis – Hybrid library generation**

797 For the project library, the DDA raw files were analyzed in Spectronaut software version 16
798 (Biognosys, Schlieren, Switzerland), using the Pulsar search engine integrated into the
799 Spectronaut software, and a search schema with default settings to generate respective spectral
800 library. The calibration search was dynamic and MS1, MS2 correction factor was 1. Data were
801 searched against the UniProt KB Human database (20,594 sequences, downloaded on February,
802 2023), with trypsin/P as the protease with up to one missed cleavage. To account for post-
803 translational modifications and chemical labelling settings, carbamidomethylation of cysteine
804 residues was defined as a fixed modification, and methionine oxidation and acetylation of
805 Lysines and acetylation of protein N-termini were defined as variable modifications. An FDR
806 less than 1% was ensured on precursor, peptide and protein level.
807 Additionally, the DIA files from the individual's samples, based on raw files, were searched in
808 the same way as described above, to generate a combination of DDA and DIA in a so-called
809 "hybrid libraries".

810 **Library search of DIA data**

811 The raw files from individual samples and acquired in DIA were then used again for the DIA
812 analysis. The files were analyzed with Spectronaut using the previously generated hybrid
813 libraries and default settings, and allowed quantification of the precursors, peptides and
814 proteins. The results were filtered by a 1% FDR on precursor, peptide and protein level using a
815 target-decoy approach, which corresponds to a Q value ≤ 0.01 (Bruderer et al., 2017)
816 Quantification data were then normalized by Spectronaut software to take into account the
817 overall acquisition heterogeneity between samples. Given the number of samples analyzed (less
818 than 500 individuals), the type of data normalization carried out for the whole dataset was a
819 local regression normalization described by Callister et al. 2006.

820
821 The LC-MS data, libraries, results tables and Spectronaut projects of the different analysis have
822 been deposited to the ProteomeXchange Consortium (Deutsch et al., 2017) via the PRIDE
823 (Vizcaíno et al., 2013) partner repository with the dataset identifier PXD043712. The
824 Spectronaut projects can be viewed using the free Spectronaut viewer
825 (www.biognosys.com/technology/spectronaut-viewer).

826 **QUANTIFICATION AND STATISTICAL ANALYSIS**

827 **DGE-Seq data preprocessing**

829 Read pairs used for analysis matched the following criteria: all 16 bases of the first read had
830 quality scores of at least 10 and the first 6 bases correspond exactly to a designed well-specific
831 barcode. The second reads were aligned to RefSeq human mRNA sequences (hg19) using bwa
832 version 0.7.17. Reads mapping to several transcripts of different genes or containing more

833 than 3 mismatches with the reference sequences were filtered out from the analysis. DGE
834 profiles were generated by counting for each sample the number of unique UMIs associated
835 with each RefSeq genes. DGE-sequenced samples were acquired from 8 sequencing runs.
836 Samples were retained if the number of UMIs was superior to 50000 and the number of
837 expressed genes above 6000, a total of 386 samples passed those cutoffs. Also, genes have
838 been filtered by keeping only a set of over-dispersed genes determined. To pick these, the co-
839 efficient of variation of each gene from the normalized adjusted expression was fitted by the
840 mean expression of each gene, using a LOESS method. Genes with a positive residual for the
841 regression were marked as over-dispersed. This leads to a total of 23885 genes.

842

843 **Transcriptomic analyses**

844 The 8 runs were merged using ComBat(gmail.com> et al., 2022) (Leek et al. 2022) from the R
845 library “sva”. Technical replicates between batches were used as references for batch-effect
846 correction, but only samples from 4 runs were kept the others are out of the scope of this
847 article. Each gene expression of the corrected values was subtracted by the minimum of the
848 gene expression before the batch correction. This step does not change the relative expression
849 of genes; however, it permits an easier interpretation of the expression values as minimums
850 cannot be less than zero. Finally, each set of technical replicates were
851 merged.

852

853 Heatmaps were computed using complexheatmap R package (2.6.2)(Gu, 2022; Gu et al.,
854 2016); samples were clustered from the Euclidean distance of expression by a hierarchical
855 clustering using Ward’s method. Differential gene expression analysis was performed with
856 deseq2 R package (1.34.0)(Love et al., 2014) in combination with log fold change shrinkage
857 function from apeglm R package (1.16.0)(Zhu et al., 2019). Deseq2 was used with raw counts
858 expression matrix and the corresponding design was cell type plus the run information. P-
859 values were adjusted based on an alpha threshold of 0.1.

860 MAplots was constructed with “ggmaplot” function from ggpubr R package (0.4.0), name of
861 genes was printed if the genes have an adjusted pvalue under 0.05 and a log fold change
862 superior to 2 or inferior to 2, except for the comparison between naïve pluripotent stem cells
863 versus primed pluripotent stem cells, where the number of upregulated genes was low
864 enough to be printed. Boxplots were computed with ggplot R package (3.3.3) and pvalues
865 were re-used from differential gene expression analysis.

866

867

868 **Supplemental spreadsheets legends**

869

870 **Spreadsheet S1. Antibodies**

871 Details of primary and secondary antibodies used for this study.

872

873 **Spreadsheet S2. Differentially Expressed genes**

874 Details of the differentially expressed genes between all conditions described in this study.

875

876 **Spreadsheet S3. Cell lines**

877 Details of cell lines used in this study

878

879 **Spreadsheet S4. X chromosome**

880 Details of the XaXa, XaXe, XaXi countings

881

882 **Spreadsheet S5. Proteomic data**

883 Details of the detected proteins by Mass spectrometry

884

885 **References**

- 886 Aksoy, I., Rognard, C., Moulin, A., Marcy, G., Masfaraud, E., Wianny, F., Cortay, V.,
887 Bellemin-Ménard, A., Doerflinger, N., Dirheimer, M., et al. (2021). Apoptosis, G1 Phase
888 Stall, and Premature Differentiation Account for Low Chimeric Competence of Human and
889 Rhesus Monkey Naive Pluripotent Stem Cells. *Stem Cell Reports* 16, 56–74.
890 <https://doi.org/10.1016/j.stemcr.2020.12.004>.
- 891 Banliat, C., Tsikis, G., Labas, V., Teixeira-Gomes, A.-P., Com, E., Lavigne, R., Pineau, C.,
892 Guyonnet, B., Mermillod, P., and Saint-Dizier, M. (2020). Identification of 56 Proteins
893 Involved in Embryo–Maternal Interactions in the Bovine Oviduct. *Int J Mol Sci* 21, 466.
894 <https://doi.org/10.3390/ijms21020466>.
- 895 Berg, P.R. van den, Bérenger-Currias, N.M.L.P., Budnik, B., Slavov, N., and Semrau, S.
896 (2023). Integration of a multi-omics stem cell differentiation dataset using a dynamical model.
897 *PLOS Genetics* 19, e1010744. <https://doi.org/10.1371/journal.pgen.1010744>.
- 898 Bredenkamp, N., Yang, J., Clarke, J., Stirparo, G.G., Meyenn, F. von, Dietmann, S., Baker,
899 D., Drummond, R., Ren, Y., Li, D., et al. (2019). Wnt Inhibition Facilitates RNA-Mediated
900 Reprogramming of Human Somatic Cells to Naive Pluripotency. *Stem Cell Reports* 13, 1083–
901 1098. <https://doi.org/10.1016/j.stemcr.2019.10.009>.
- 902 Bruderer, R., Bernhardt, O.M., Gandhi, T., Xuan, Y., Sondermann, J., Schmidt, M., Gomez-
903 Varela, D., and Reiter, L. (2017). Optimization of Experimental Parameters in Data-
904 Independent Mass Spectrometry Significantly Increases Depth and Reproducibility of Results.
905 *Mol Cell Proteomics* 16, 2296–2309. <https://doi.org/10.1074/mcp.RA117.000314>.
- 906 Castel, G., Meistermann, D., Bretin, B., Firmin, J., Blin, J., Loubersac, S., Bruneau, A.,
907 Chevolleau, S., Kilens, S., Chariau, C., et al. (2020). Induction of Human Trophoblast Stem
908 Cells from Somatic Cells and Pluripotent Stem Cells. *Cell Reports* 33, 108419.
909 <https://doi.org/10.1016/j.celrep.2020.108419>.
- 910 Charpentier, E., Cornec, M., Dumont, S., Meistermann, D., Bordron, P., David, L., Redon, R.,
911 Bonnaud, S., and Bihouée, A. (2021). 3' RNA sequencing for robust and low-cost gene
912 expression profiling (Protocol Exchange).
- 913 Chen, H., Aksoy, I., Gonnot, F., Osteil, P., Aubry, M., Hamela, C., Rognard, C., Hochard, A.,
914 Voisin, S., Fontaine, E., et al. (2015). Reinforcement of STAT3 activity reprogrammes human
915 embryonic stem cells to naive-like pluripotency. *Nat Commun* 6, 7095.
916 <https://doi.org/10.1038/ncomms8095>.
- 917 Cinkornpumin, J.K., Kwon, S.Y., Guo, Y., Hossain, I., Sirois, J., Russett, C.S., Tseng, H.-W.,
918 Okae, H., Arima, T., Duchaine, T.F., et al. (2020). Naive Human Embryonic Stem Cells Can
919 Give Rise to Cells with a Trophoblast-like Transcriptome and Methylome. *Stem Cell Reports*
920 15, 198–213. <https://doi.org/10.1016/j.stemcr.2020.06.003>.
- 921 De Los Angeles, A., Ferrari, F., Xi, R., Fujiwara, Y., Benvenisty, N., Deng, H., Hochedlinger,
922 K., Jaenisch, R., Lee, S., Leitch, H.G., et al. (2015). Hallmarks of pluripotency. *Nature* 525,
923 469–478. <https://doi.org/10.1038/nature15515>.
- 924 Deutsch, E.W., Csordas, A., Sun, Z., Jarnuczak, A., Perez-Riverol, Y., Ternent, T., Campbell,
925 D.S., Bernal-Llinares, M., Okuda, S., Kawano, S., et al. (2017). The ProteomeXchange
926 consortium in 2017: supporting the cultural change in proteomics public data deposition.
927 *Nucleic Acids Res* 45, D1100–D1106. <https://doi.org/10.1093/nar/gkw936>.
- 928 Dong, C., Beltcheva, M., Gontarz, P., Zhang, B., Popli, P., Fischer, L.A., Khan, S.A., Park,
929 K., Yoon, E.-J., Xing, X., et al. (2020). Derivation of trophoblast stem cells from naïve human
930 pluripotent stem cells. *ELife* 9, e52504. <https://doi.org/10.7554/eLife.52504>.
- 931 Edfors, F., Danielsson, F., Hallström, B.M., Käll, L., Lundberg, E., Pontén, F., Forsström, B.,
932 and Uhlén, M. (2016). Gene-specific correlation of RNA and protein levels in human cells
933 and tissues. *Mol Syst Biol* 12, 883. <https://doi.org/10.15252/msb.20167144>.

934 Fan, Y., Min, Z., Alsolami, S., Ma, Z., Zhang, E., Chen, W., Zhong, K., Pei, W., Kang, X.,
935 Zhang, P., et al. (2021). Generation of human blastocyst-like structures from pluripotent stem
936 cells. *Cell Discov* 7, 1–14. <https://doi.org/10.1038/s41421-021-00316-8>.

937 Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma,
938 Y., Viukov, S., Maza, I., Zviran, A., et al. (2013). Derivation of novel human ground state
939 naive pluripotent stem cells. *Nature* 504, 282–286. <https://doi.org/10.1038/nature12745>.

940 Gao, X., Nowak-Imialek, M., Chen, X., Chen, D., Herrmann, D., Ruan, D., Chen, A.C.H.,
941 Eckersley-Maslin, M.A., Ahmad, S., Lee, Y.L., et al. (2019). Establishment of porcine and
942 human expanded potential stem cells. *Nat Cell Biol* 21, 687–699.
943 <https://doi.org/10.1038/s41556-019-0333-2>.

944 Gerri, C., McCarthy, A., Alanis-Lobato, G., Demtschenko, A., Bruneau, A., Loubersac, S.,
945 Fogarty, N.M.E., Hampshire, D., Elder, K., Snell, P., et al. (2020). Initiation of a conserved
946 trophoctoderm program in human, cow and mouse embryos. *Nature* 587, 443–447.
947 <https://doi.org/10.1038/s41586-020-2759-x>.

948 Girard, O., Lavigne, R., Chevolleau, S., Onfray, C., Com, E., Schmit, P.-O., Chapelle, M.,
949 Fréour, T., Lane, L., David, L., et al. (2023). Naive Pluripotent and Trophoblastic Stem Cell
950 Lines as a Model for Detecting Missing Proteins in the Context of the Chromosome-Centric
951 Human Proteome Project. *J. Proteome Res.* 22, 1148–1158.
952 <https://doi.org/10.1021/acs.jproteome.2c00496>.

953 Giulitti, S., Pellegrini, M., Zorzan, I., Martini, P., Gagliano, O., Mutarelli, M., Ziller, M.J.,
954 Cacchiarelli, D., Romualdi, C., Elvassore, N., et al. (2019). Direct generation of human naive
955 induced pluripotent stem cells from somatic cells in microfluidics. *Nat Cell Biol* 21, 275–286.
956 <https://doi.org/10.1038/s41556-018-0254-5>.

957 gmail.com>, J.T.L. <jtleek at, bu.edu>, W.E.J. <wej at, jhsph.edu>, H.S.P. <hiparker at,
958 jhmi.edu>, E.J.F. <ejfertig at, jhsph.edu>, A.E.J. <ajaffe at, gmail.com>, Y.Z. <zhangyuqing
959 pkusms at, princeton.edu>, J.D.S. <jstorey at, and gmail.com>, L.C.T. <lcolladotor at (2022).
960 sva: Surrogate Variable Analysis (Bioconductor version: Release (3.16)).

961 Gu, Z. (2022). Complex heatmap visualization. *IMeta* 1, e43. <https://doi.org/10.1002/imt2.43>.

962 Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations
963 in multidimensional genomic data. *Bioinformatics* 32, 2847–2849.
964 <https://doi.org/10.1093/bioinformatics/btw313>.

965 Guo, G., von Meyenn, F., Santos, F., Chen, Y., Reik, W., Bertone, P., Smith, A., and Nichols,
966 J. (2016). Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human
967 Inner Cell Mass. *Stem Cell Reports* 6, 437–446. <https://doi.org/10.1016/j.stemcr.2016.02.005>.

968 Guo, G., Stirparo, G.G., Strawbridge, S.E., Spindlow, D., Yang, J., Clarke, J., Dattani, A.,
969 Yanagida, A., Li, M.A., Myers, S., et al. (2021). Human naive epiblast cells possess
970 unrestricted lineage potential. *Cell Stem Cell* 28, 1040-1056.e6.
971 <https://doi.org/10.1016/j.stem.2021.02.025>.

972 Guo, H., Zhu, P., Yan, L., Li, R., Hu, B., Lian, Y., Yan, J., Ren, X., Lin, S., Li, J., et al.
973 (2014). The DNA methylation landscape of human early embryos. *Nature* 511, 606–610.
974 <https://doi.org/10.1038/nature13544>.

975 Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat,
976 J., Carey, B.W., and Jaenisch, R. (2010). Human embryonic stem cells with biological and
977 epigenetic characteristics similar to those of mouse ESCs. *Proceedings of the National*
978 *Academy of Sciences* 107, 9222–9227. <https://doi.org/10.1073/pnas.1004584107>.

979 Io, S., Kabata, M., Iemura, Y., Semi, K., Morone, N., Minagawa, A., Wang, B., Okamoto, I.,
980 Nakamura, T., Kojima, Y., et al. (2021). Capturing human trophoblast development with
981 naive pluripotent stem cells in vitro. *Cell Stem Cell* 28, 1023-1039.e13.
982 <https://doi.org/10.1016/j.stem.2021.03.013>.

983 Kagawa, H., Javali, A., Khoei, H.H., Sommer, T.M., Sestini, G., Novatchkova, M., Scholte op
984 Reimer, Y., Castel, G., Bruneau, A., Maenhoudt, N., et al. (2022). Human blastoids model
985 blastocyst development and implantation. *Nature* *601*, 600–605.
986 <https://doi.org/10.1038/s41586-021-04267-8>.

987 Karvas, R.M. Stem-cell-derived trophoblast organoids model human placental development
988 and susceptibility to emerging pathogens. *25*. .

989 Kilens, S., Meistermann, D., Moreno, D., Chariou, C., Gaignerie, A., Reignier, A., Lelièvre,
990 Y., Casanova, M., Vallot, C., Nedellec, S., et al. (2018). Parallel derivation of isogenic human
991 primed and naive induced pluripotent stem cells. *Nat Commun* *9*, 1–13.
992 <https://doi.org/10.1038/s41467-017-02107-w>.

993 Leitch, H.G., McEwen, K.R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W.,
994 Nashun, B., Knezovich, J.G., Smith, A., et al. (2013). Naive pluripotency is associated with
995 global DNA hypomethylation. *Nat Struct Mol Biol* *20*, 311–316.
996 <https://doi.org/10.1038/nsmb.2510>.

997 Linneberg-Agerholm, M., Wong, Y.F., Romero Herrera, J.A., Monteiro, R.S., Anderson,
998 K.G.V., and Brickman, J.M. (2019). Naïve human pluripotent stem cells respond to Wnt,
999 Nodal and LIF signalling to produce expandable naïve extra-embryonic endoderm.
1000 *Development* *146*, dev180620. <https://doi.org/10.1242/dev.180620>.

1001 Liu, B., Chen, S., Xu, Y., Lyu, Y., Wang, J., Du, Y., Sun, Y., Liu, H., Zhou, H., Lai, W., et al.
1002 (2021a). Chemically defined and xeno-free culture condition for human extended pluripotent
1003 stem cells. *Nat Commun* *12*, 3017. <https://doi.org/10.1038/s41467-021-23320-8>.

1004 Liu, X., Nefzger, C.M., Rossello, F.J., Chen, J., Knaupp, A.S., Firas, J., Ford, E., Pflueger, J.,
1005 Paynter, J.M., Chy, H.S., et al. (2017). Comprehensive characterization of distinct states of
1006 human naive pluripotency generated by reprogramming. *Nat Methods* *14*, 1055–1062.
1007 <https://doi.org/10.1038/nmeth.4436>.

1008 Liu, X., Ouyang, J.F., Rossello, F.J., Tan, J.P., Davidson, K.C., Valdes, D.S., Schröder, J.,
1009 Sun, Y.B.Y., Chen, J., Knaupp, A.S., et al. (2020). Reprogramming roadmap reveals route to
1010 human induced trophoblast stem cells. *Nature* *586*, 101–107. <https://doi.org/10.1038/s41586-020-2734-6>.

1011 Liu, X., Tan, J.P., Schröder, J., Aberkane, A., Ouyang, J.F., Mohenska, M., Lim, S.M., Sun,
1012 Y.B.Y., Chen, J., Sun, G., et al. (2021b). Modelling human blastocysts by reprogramming
1013 fibroblasts into iBlastoids. *Nature* *591*, 627–632. <https://doi.org/10.1038/s41586-021-03372-y>.

1014 y.

1015 Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and
1016 dispersion for RNA-seq data with DESeq2. *Genome Biology* *15*, 550.
1017 <https://doi.org/10.1186/s13059-014-0550-8>.

1018 Meistermann, D., Bruneau, A., Loubersac, S., Reignier, A., Firmin, J., François-Campion, V.,
1019 Kilens, S., Lelièvre, Y., Lammers, J., Feyeux, M., et al. (2021). Integrated pseudotime
1020 analysis of human pre-implantation embryo single-cell transcriptomes reveals the dynamics of
1021 lineage specification. *Cell Stem Cell* *28*, 1625-1640.e6.
1022 <https://doi.org/10.1016/j.stem.2021.04.027>.

1023 Niakan, K.K., and Eggan, K. (2013). Analysis of human embryos from zygote to blastocyst
1024 reveals distinct gene expression patterns relative to the mouse. *Developmental Biology* *375*,
1025 54–64. <https://doi.org/10.1016/j.ydbio.2012.12.008>.

1026 Okae, H., Toh, H., Sato, T., Hiura, H., Takahashi, S., Shirane, K., Kabayama, Y., Suyama,
1027 M., Sasaki, H., and Arima, T. (2018). Derivation of Human Trophoblast Stem Cells. *Cell*
1028 *Stem Cell* *22*, 50-63.e6. <https://doi.org/10.1016/j.stem.2017.11.004>.

1029 Osnato, A., Brown, S., Krueger, C., Andrews, S., Collier, A.J., Nakanoh, S., Quiroga
1030 Londoño, M., Wesley, B.T., Muraro, D., Brumm, A.S., et al. (2021). TGFβ signalling is
1031

1032 required to maintain pluripotency of human naïve pluripotent stem cells. *ELife* *10*, e67259.
1033 <https://doi.org/10.7554/eLife.67259>.

1034 Pastor, W.A., Chen, D., Liu, W., Kim, R., Sahakyan, A., Lukianchikov, A., Plath, K.,
1035 Jacobsen, S.E., and Clark, A.T. (2016). Naive Human Pluripotent Cells Feature a Methylation
1036 Landscape Devoid of Blastocyst or Germline Memory. *Cell Stem Cell* *18*, 323–329.
1037 <https://doi.org/10.1016/j.stem.2016.01.019>.

1038 Rugg-Gunn, P.J., Moris, N., and Tam, P.P.L. (2023). Technical challenges of studying early
1039 human development. *Development* *150*, dev201797. <https://doi.org/10.1242/dev.201797>.

1040 Sahakyan, A., Kim, R., Chronis, C., Sabri, S., Bonora, G., Theunissen, T.W., Kuoy, E.,
1041 Langerman, J., Clark, A.T., Jaenisch, R., et al. (2017). Human Naive Pluripotent Stem Cells
1042 Model X Chromosome Dampening and X Inactivation. *Cell Stem Cell* *20*, 87–101.
1043 <https://doi.org/10.1016/j.stem.2016.10.006>.

1044 Santos, F., Hyslop, L., Stojkovic, P., Leary, C., Murdoch, A., Reik, W., Stojkovic, M.,
1045 Herbert, M., and Dean, W. (2010). Evaluation of epigenetic marks in human embryos derived
1046 from IVF and ICSI. *Human Reproduction* *25*, 2387–2395.
1047 <https://doi.org/10.1093/humrep/deq151>.

1048 Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and
1049 Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature* *473*,
1050 337–342. <https://doi.org/10.1038/nature10098>.

1051 Smith, Z.D., Chan, M.M., Humm, K.C., Karnik, R., Mekhoubad, S., Regev, A., Eggan, K.,
1052 and Meissner, A. (2014). DNA methylation dynamics of the human preimplantation embryo.
1053 *Nature* *511*, 611–615. <https://doi.org/10.1038/nature13581>.

1054 Sozen, B., Jorgensen, V., Weatherbee, B.A.T., Chen, S., Zhu, M., and Zernicka-Goetz, M.
1055 (2021). Reconstructing aspects of human embryogenesis with pluripotent stem cells. *Nat*
1056 *Commun* *12*, 5550. <https://doi.org/10.1038/s41467-021-25853-4>.

1057 Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka,
1058 S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined
1059 Factors. *Cell* *131*, 861–872. <https://doi.org/10.1016/j.cell.2007.11.019>.

1060 Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F.,
1061 Clarke, J., Mansfield, W., et al. (2014). Resetting Transcription Factor Control Circuitry
1062 toward Ground-State Pluripotency in Human. *Cell* *158*, 1254–1269.
1063 <https://doi.org/10.1016/j.cell.2014.08.029>.

1064 Tan, T., Wu, J., Si, C., Dai, S., Zhang, Y., Sun, N., Zhang, E., Shao, H., Si, W., Yang, P., et
1065 al. (2021). Chimeric contribution of human extended pluripotent stem cells to monkey
1066 embryos *ex vivo*. *Cell* *184*, 2020–2032.e14. <https://doi.org/10.1016/j.cell.2021.03.020>.

1067 Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan,
1068 Z.P., Maetzel, D., Ganz, K., Shi, L., et al. (2014). Systematic Identification of Culture
1069 Conditions for Induction and Maintenance of Naive Human Pluripotency. *Cell Stem Cell* *15*,
1070 471–487. <https://doi.org/10.1016/j.stem.2014.07.002>.

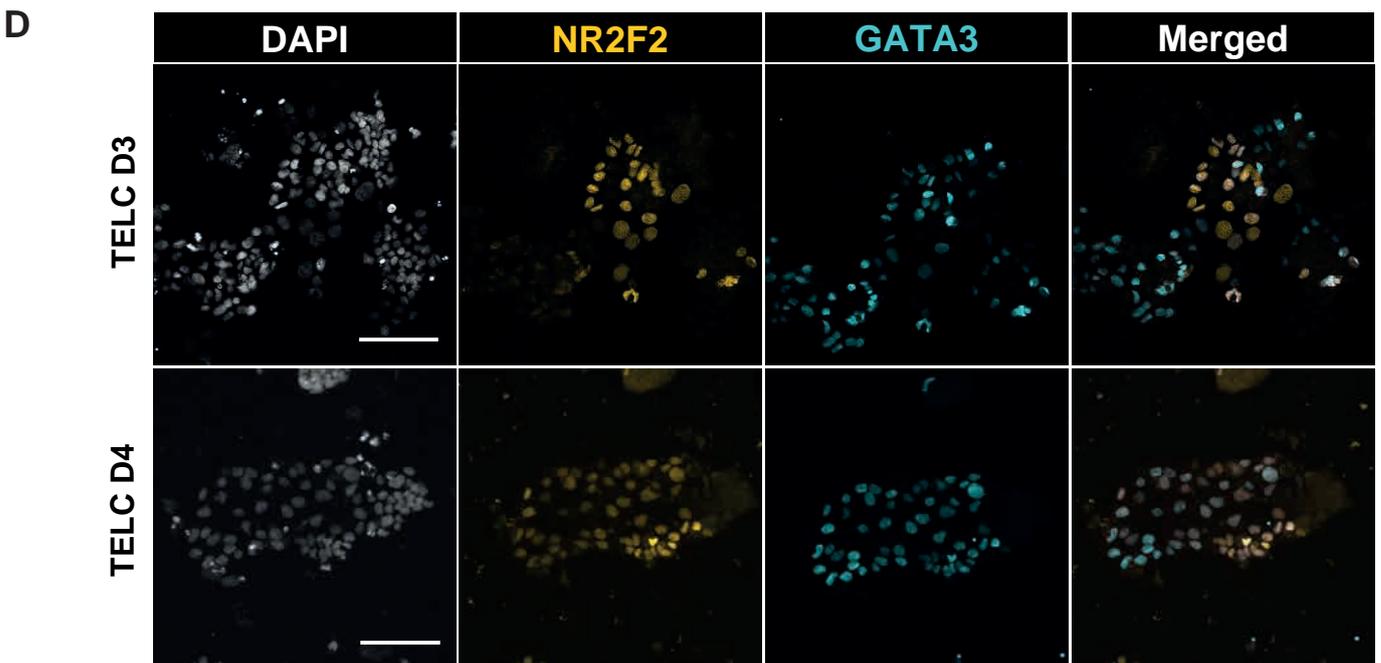
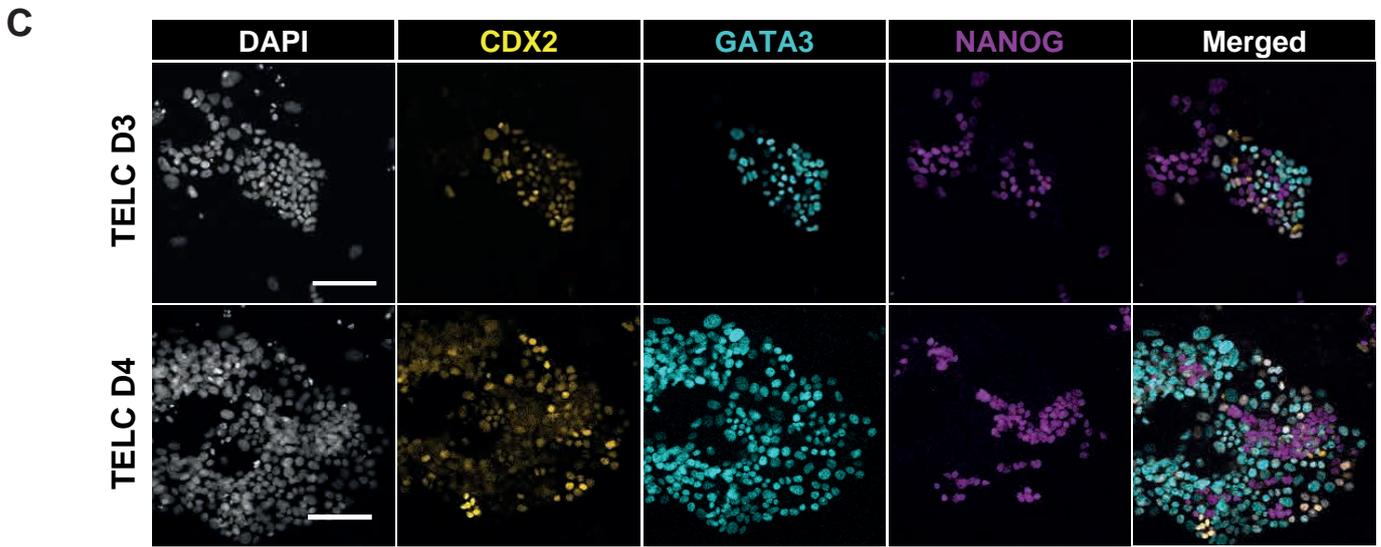
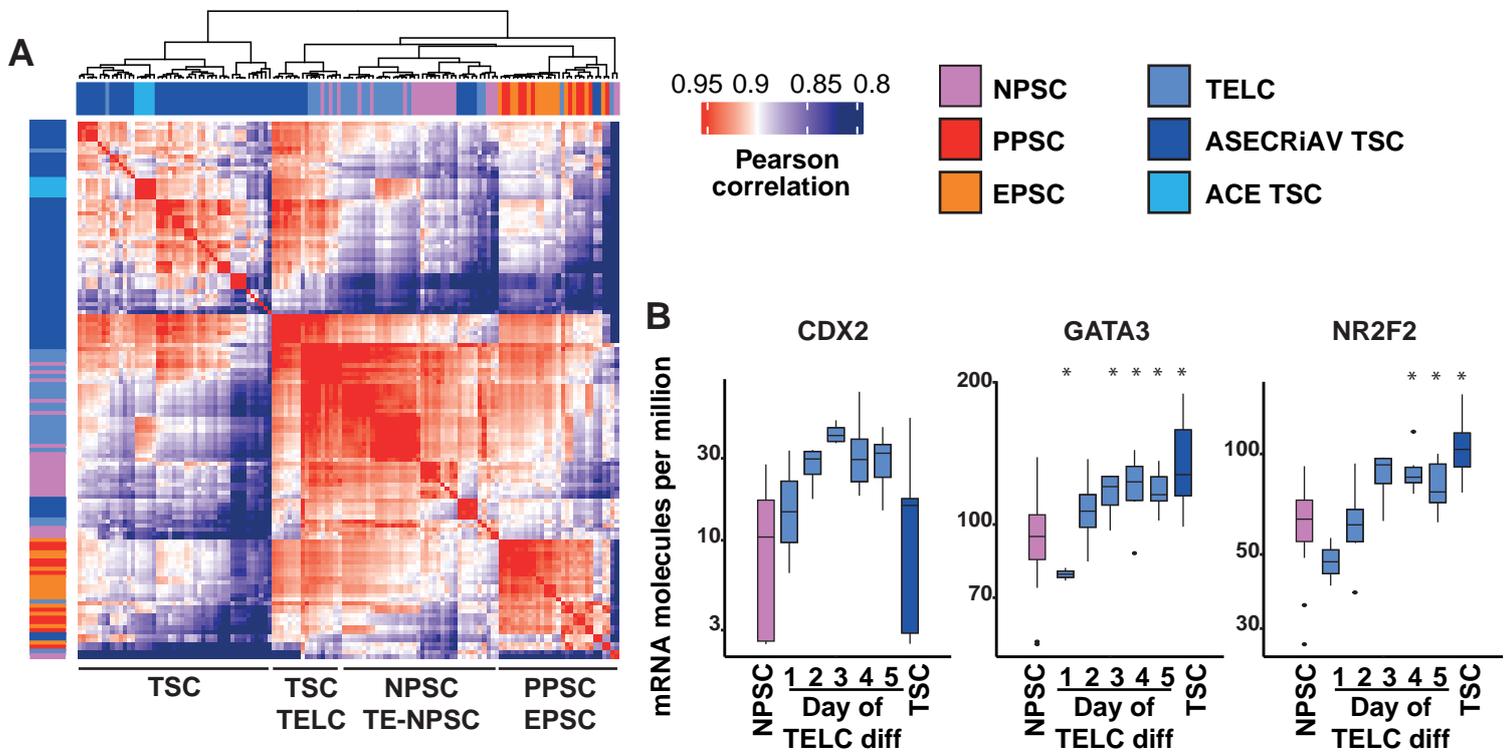
1071 Theunissen, T.W., Friedli, M., He, Y., Planet, E., O’Neil, R.C., Markoulaki, S., Pontis, J.,
1072 Wang, H., Iouranova, A., Imbeault, M., et al. (2016). Molecular Criteria for Defining the
1073 Naive Human Pluripotent State. *Cell Stem Cell* *19*, 502–515.
1074 <https://doi.org/10.1016/j.stem.2016.06.011>.

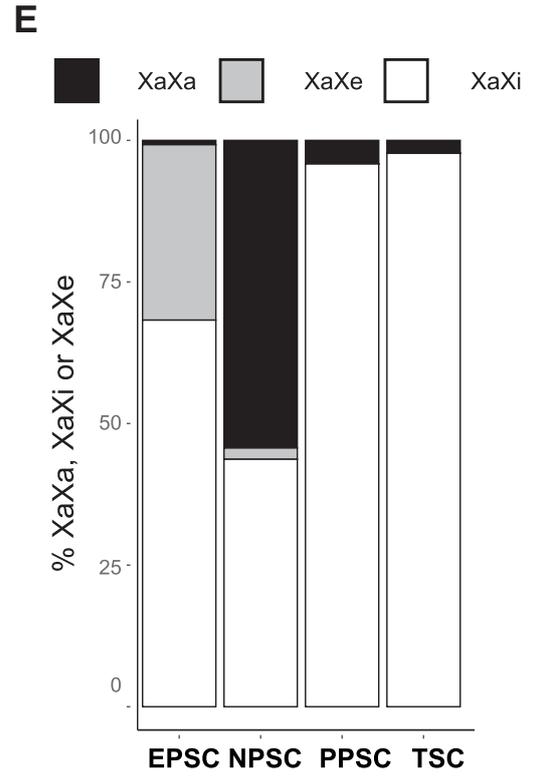
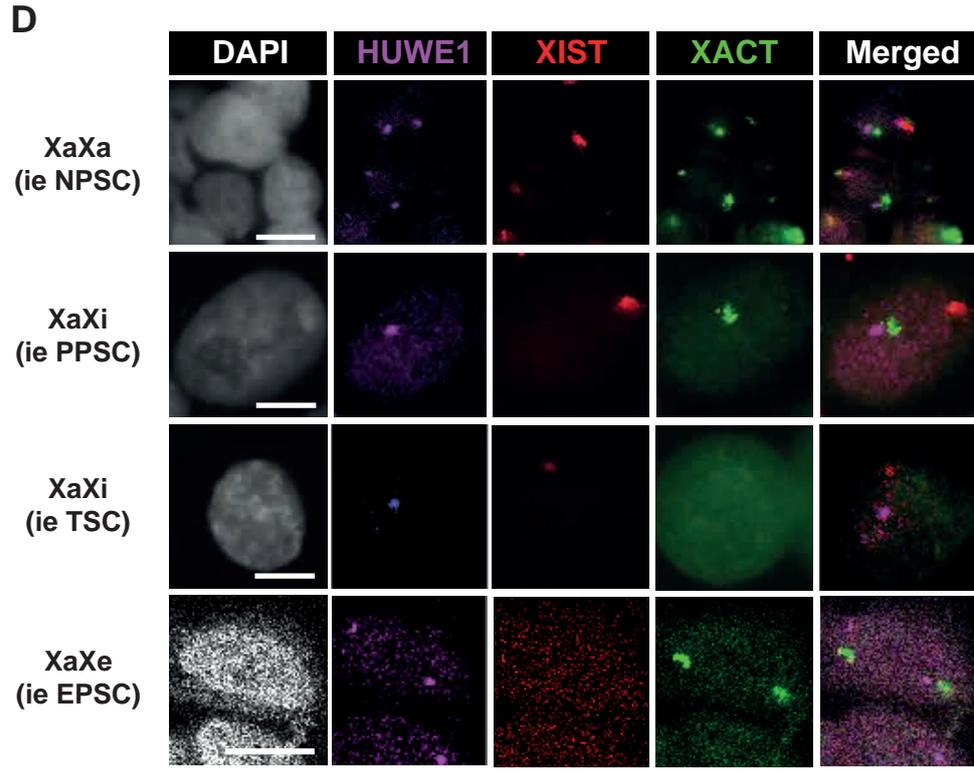
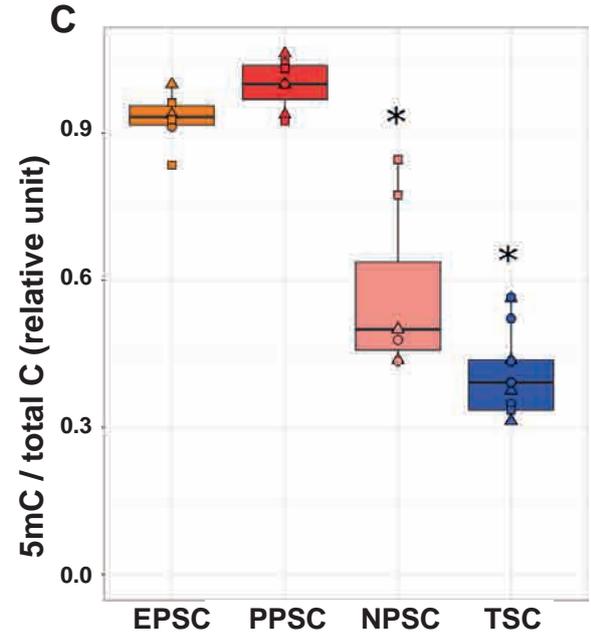
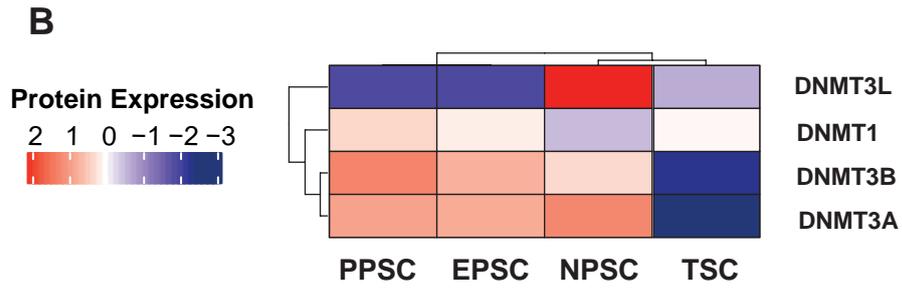
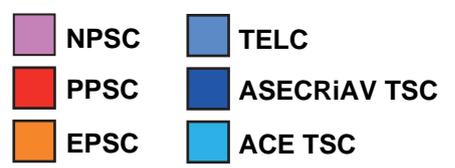
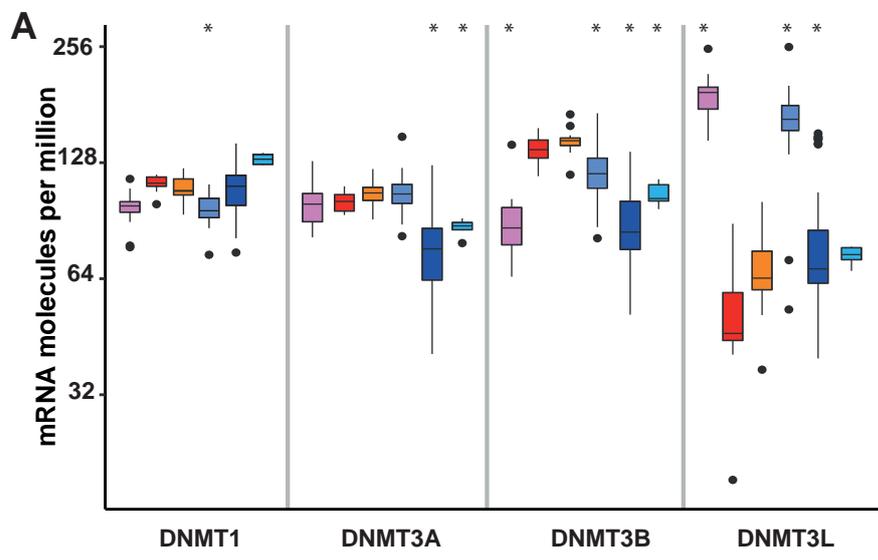
1075 Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall,
1076 V.S., and Jones, J.M. (1998). Embryonic Stem Cell Lines Derived from Human Blastocysts.
1077 *Science* *282*, 1145–1147. <https://doi.org/10.1126/science.282.5391.1145>.

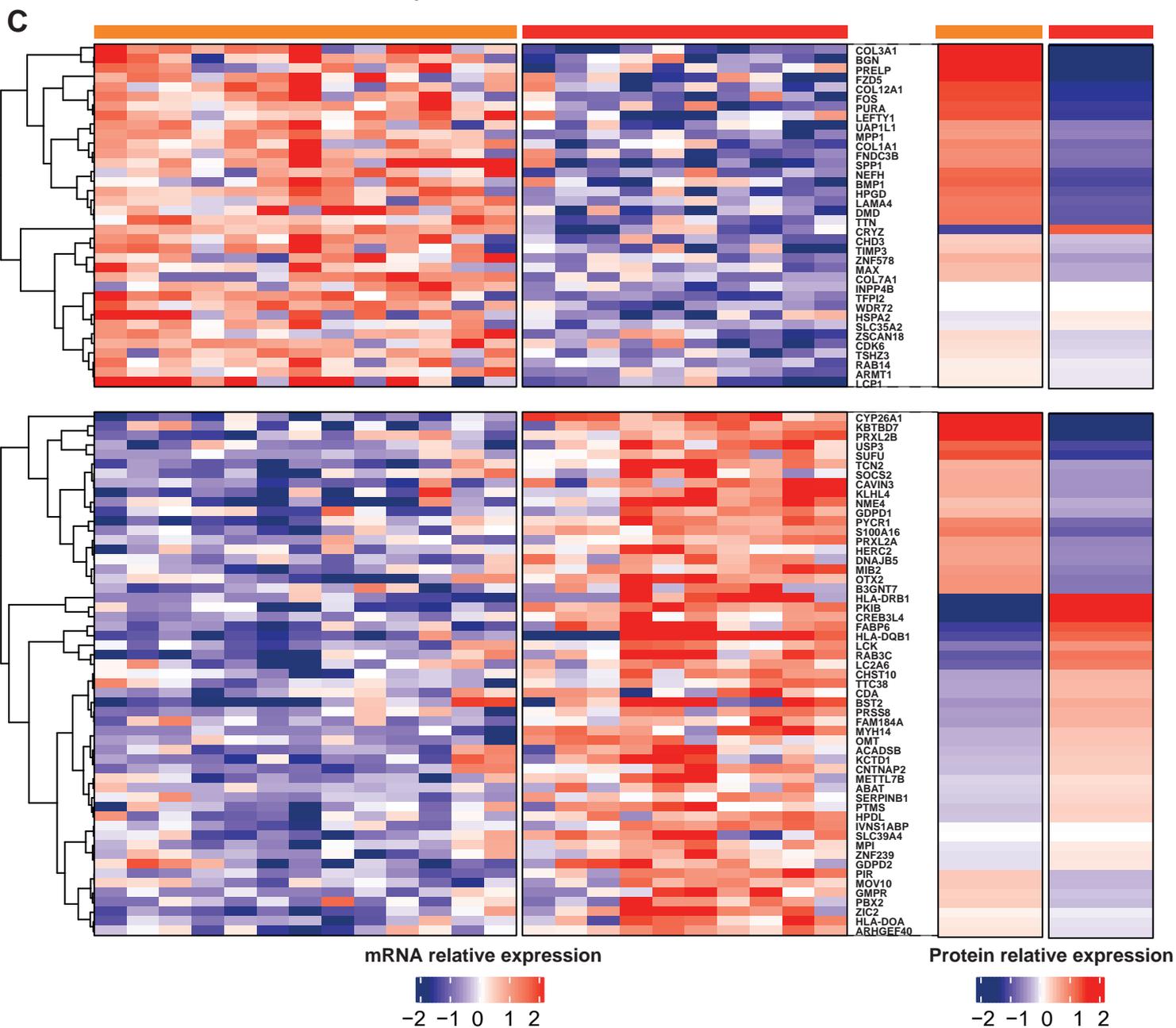
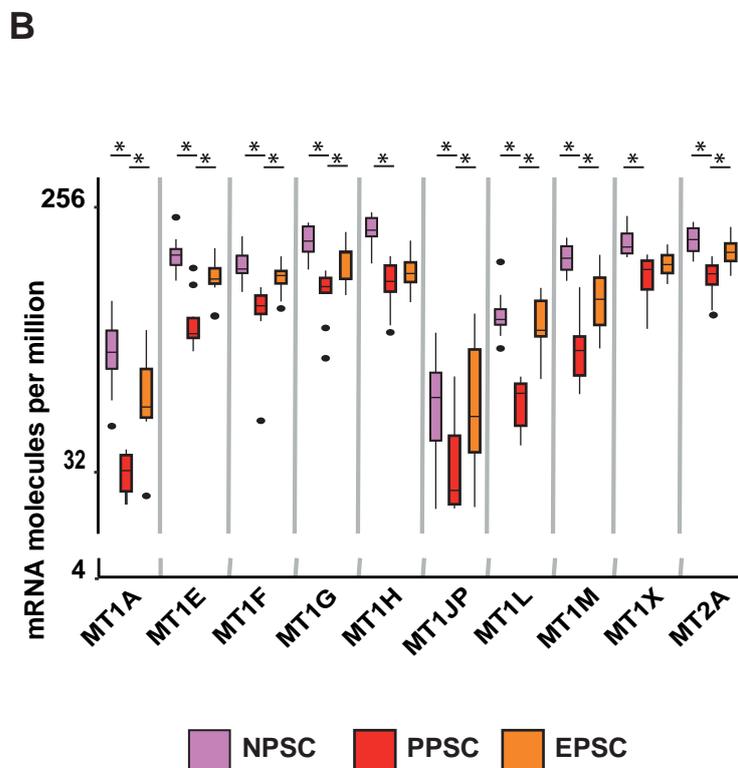
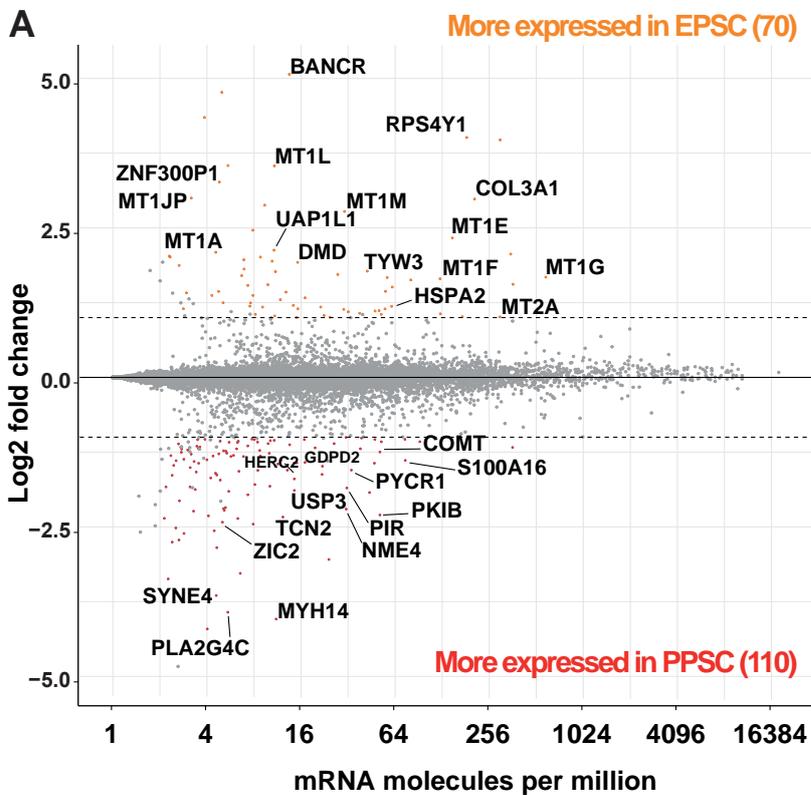
1078 Vallot, C., Patrat, C., Collier, A.J., Huret, C., Casanova, M., Liyakat Ali, T.M., Tosolini, M.,
1079 Frydman, N., Heard, E., Rugg-Gunn, P.J., et al. (2017). XACT Noncoding RNA Competes
1080 with XIST in the Control of X Chromosome Activity during Human Early Development. *Cell*
1081 *Stem Cell* *20*, 102–111. <https://doi.org/10.1016/j.stem.2016.10.014>.

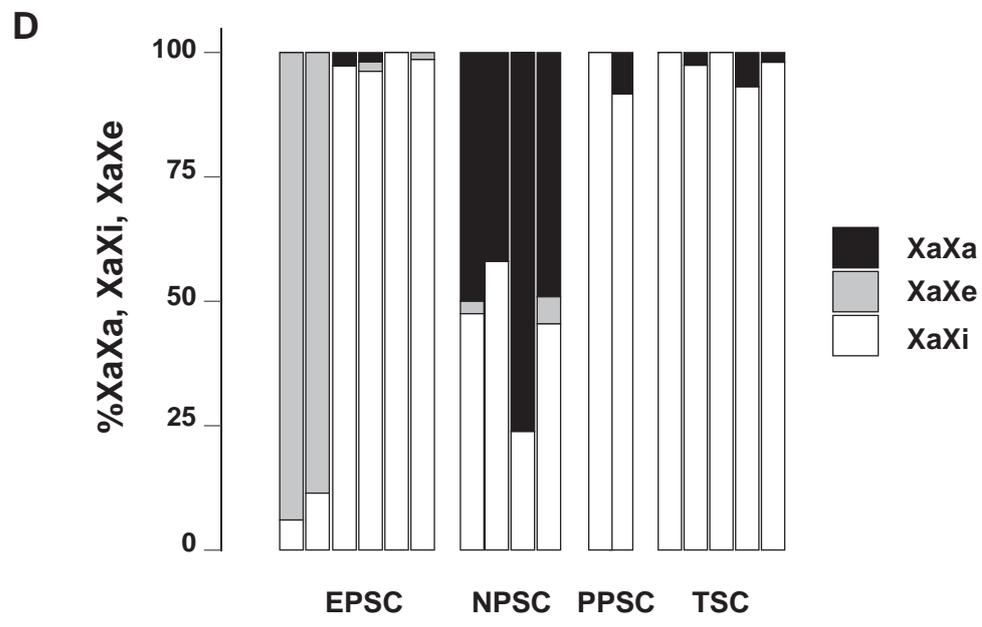
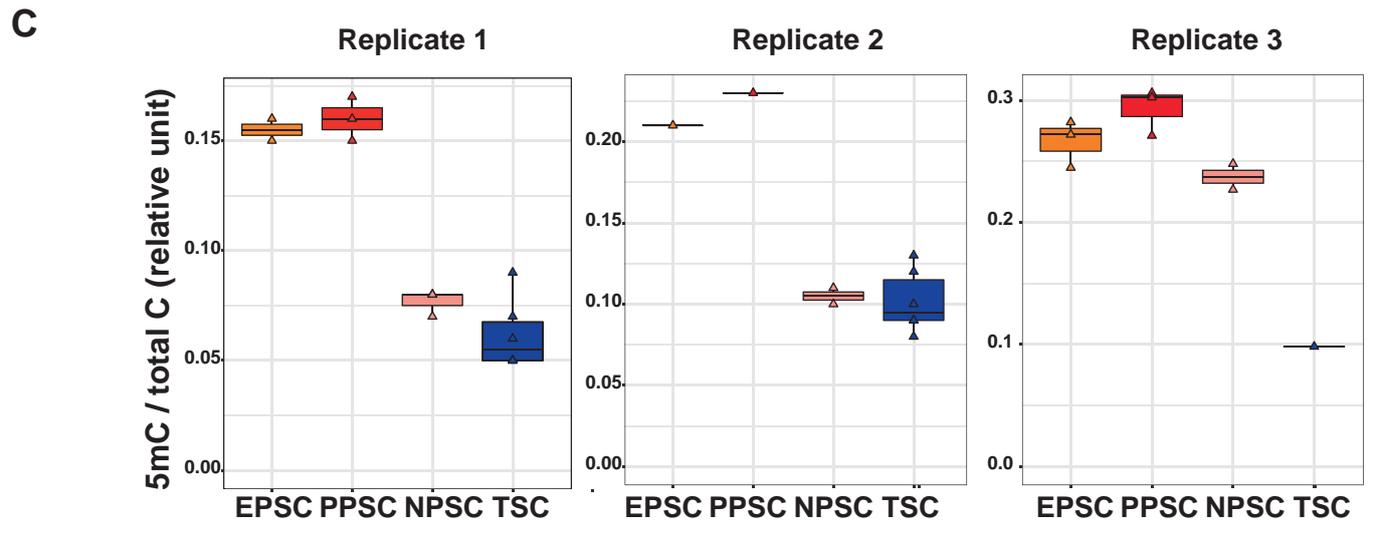
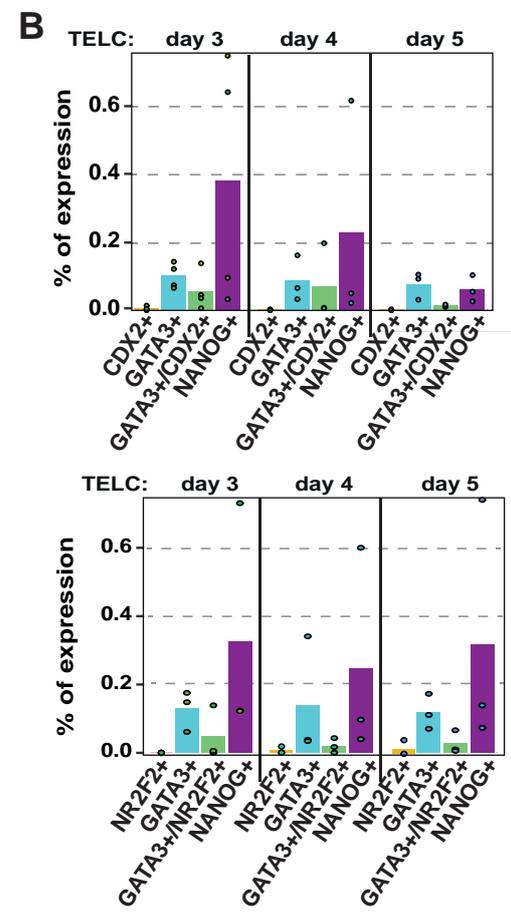
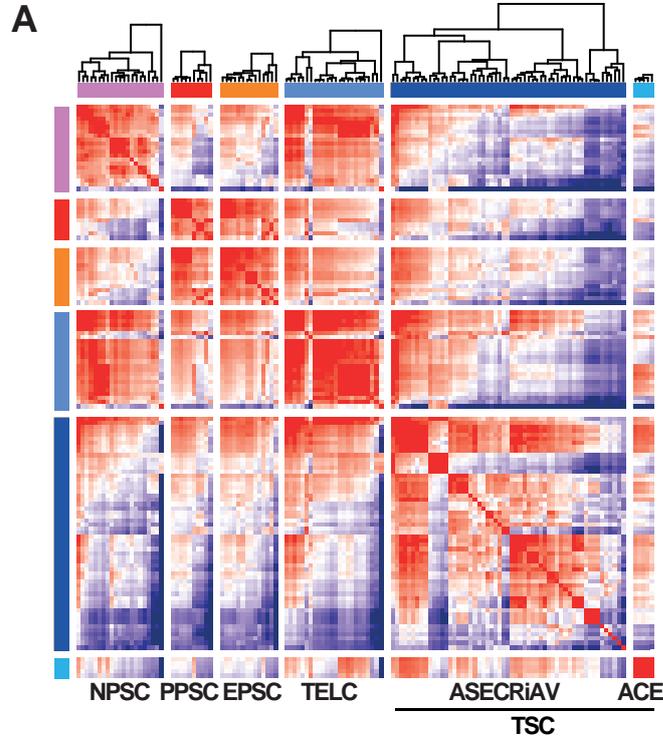
1082 Vizcaíno, J.A., Côté, R.G., Csordas, A., Dianes, J.A., Fabregat, A., Foster, J.M., Griss, J.,
1083 Alpi, E., Birim, M., Contell, J., et al. (2013). The Proteomics Identifications (PRIDE)
1084 database and associated tools: status in 2013. *Nucleic Acids Res* *41*, D1063–D1069.
1085 <https://doi.org/10.1093/nar/gks1262>.
1086 Wu, J., Platero-Luengo, A., Sakurai, M., Sugawara, A., Gil, M.A., Yamauchi, T., Suzuki, K.,
1087 Bogliotti, Y.S., Cuello, C., Morales Valencia, M., et al. (2017). Interspecies Chimerism with
1088 Mammalian Pluripotent Stem Cells. *Cell* *168*, 473–486.e15.
1089 <https://doi.org/10.1016/j.cell.2016.12.036>.
1090 Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J., et
1091 al. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic
1092 stem cells. *Nat Struct Mol Biol* *20*, 1131–1139. <https://doi.org/10.1038/nsmb.2660>.
1093 Yanagida, A., Spindlow, D., Nichols, J., Dattani, A., Smith, A., and Guo, G. (2021). Naive
1094 stem cell blastocyst model captures human embryo lineage segregation. *Cell Stem Cell* *28*,
1095 1016–1022.e4. <https://doi.org/10.1016/j.stem.2021.04.031>.
1096 Yang, Y., Liu, B., Xu, J., Wang, J., Wu, J., Shi, C., Xu, Y., Dong, J., Wang, C., Lai, W., et al.
1097 (2017). Derivation of Pluripotent Stem Cells with In Vivo Embryonic and Extraembryonic
1098 Potency. *Cell* *169*, 243–257.e25. <https://doi.org/10.1016/j.cell.2017.02.005>.
1099 Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie,
1100 J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced Pluripotent Stem Cell Lines
1101 Derived from Human Somatic Cells. *Science* *318*, 1917–1920.
1102 <https://doi.org/10.1126/science.1151526>.
1103 Yu, L., Wei, Y., Duan, J., Schmitz, D.A., Sakurai, M., Wang, L., Wang, K., Zhao, S., Hon,
1104 G.C., and Wu, J. (2021). Blastocyst-like structures generated from human pluripotent stem
1105 cells. *Nature* *591*, 620–626. <https://doi.org/10.1038/s41586-021-03356-y>.
1106 Zheng, R., Geng, T., Wu, D.-Y., Zhang, T., He, H.-N., Du, H.-N., Zhang, D., Miao, Y.-L.,
1107 and Jiang, W. (2021). Derivation of feeder-free human extended pluripotent stem cells. *Stem*
1108 *Cell Reports* *16*, 2410–2414. <https://doi.org/10.1016/j.stemcr.2021.07.019>.
1109 Zhou, F., Wang, R., Yuan, P., Ren, Y., Mao, Y., Li, R., Lian, Y., Li, J., Wen, L., Yan, L., et
1110 al. (2019). Reconstituting the transcriptome and DNA methylome landscapes of human
1111 implantation. *Nature* *572*, 660–664. <https://doi.org/10.1038/s41586-019-1500-0>.
1112 Zhu, A., Ibrahim, J.G., and Love, M.I. (2019). Heavy-tailed prior distributions for sequence
1113 count data: removing the noise and preserving large differences. *Bioinformatics* *35*, 2084–
1114 2092. <https://doi.org/10.1093/bioinformatics/bty895>.
1115 Zijlmans, D.W., Talon, I., Verhelst, S., Bendall, A., Van Nerum, K., Javali, A., Malcolm,
1116 A.A., van Knippenberg, S.S.F.A., Biggins, L., To, S.K., et al. (2022). Integrated multi-omics
1117 reveal polycomb repressive complex 2 restricts human trophoblast induction. *Nat Cell Biol*
1118 *24*, 858–871. <https://doi.org/10.1038/s41556-022-00932-w>.

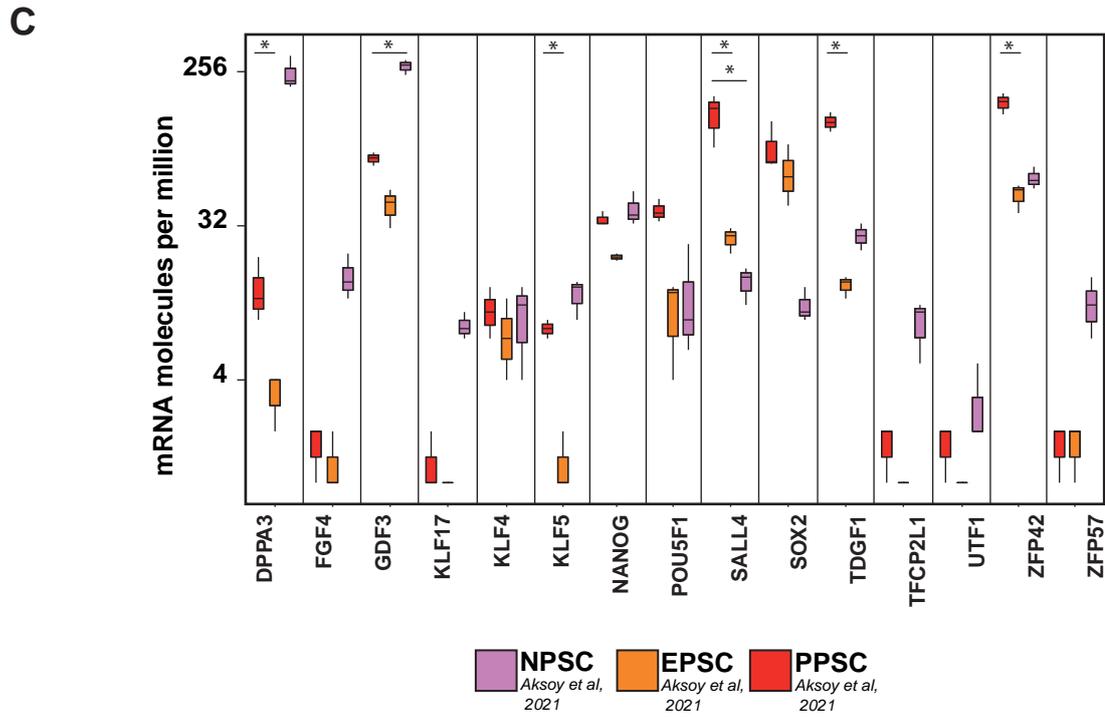
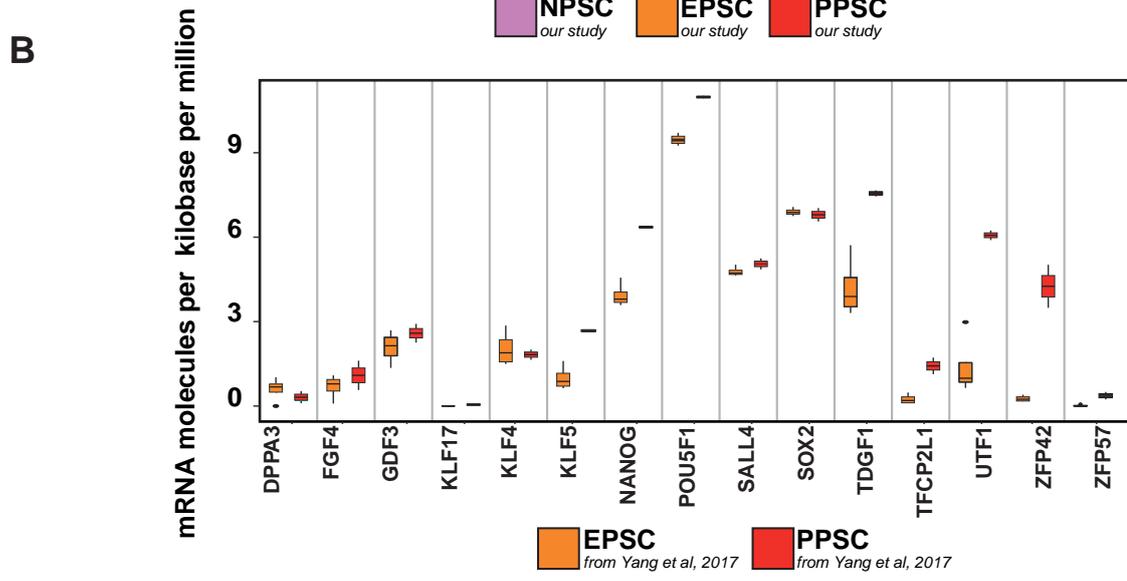
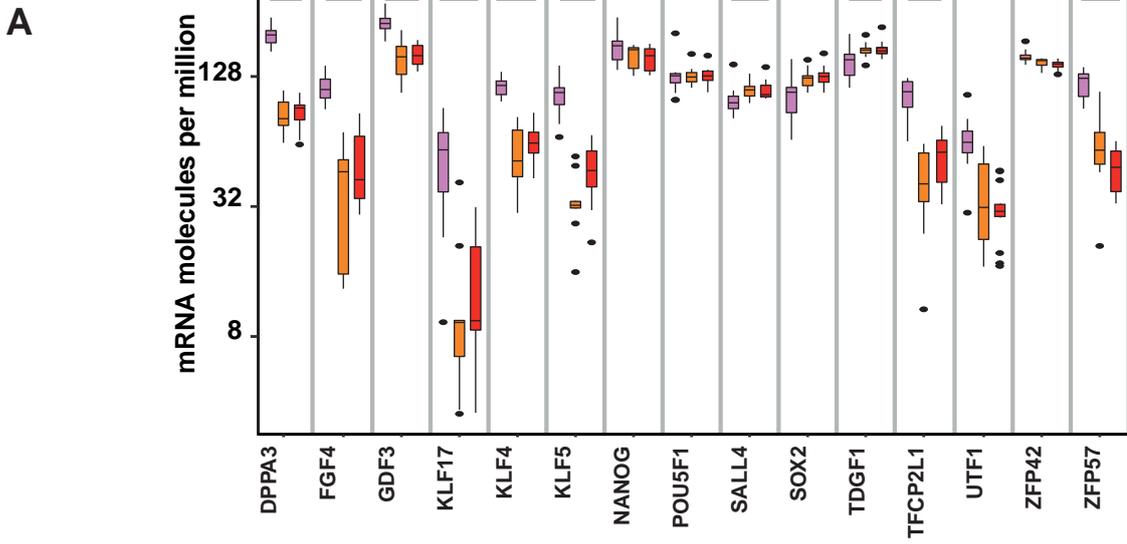
1119

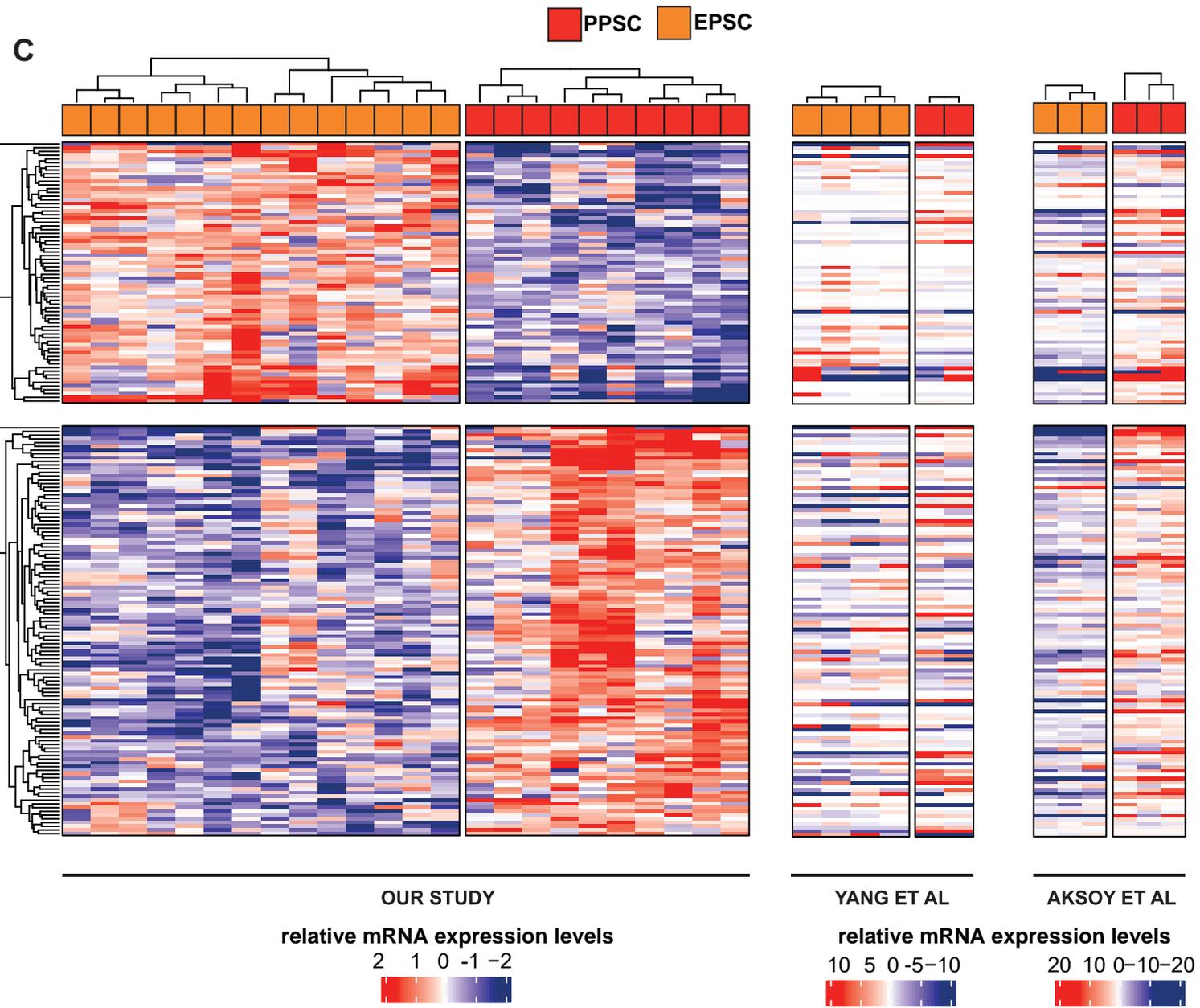
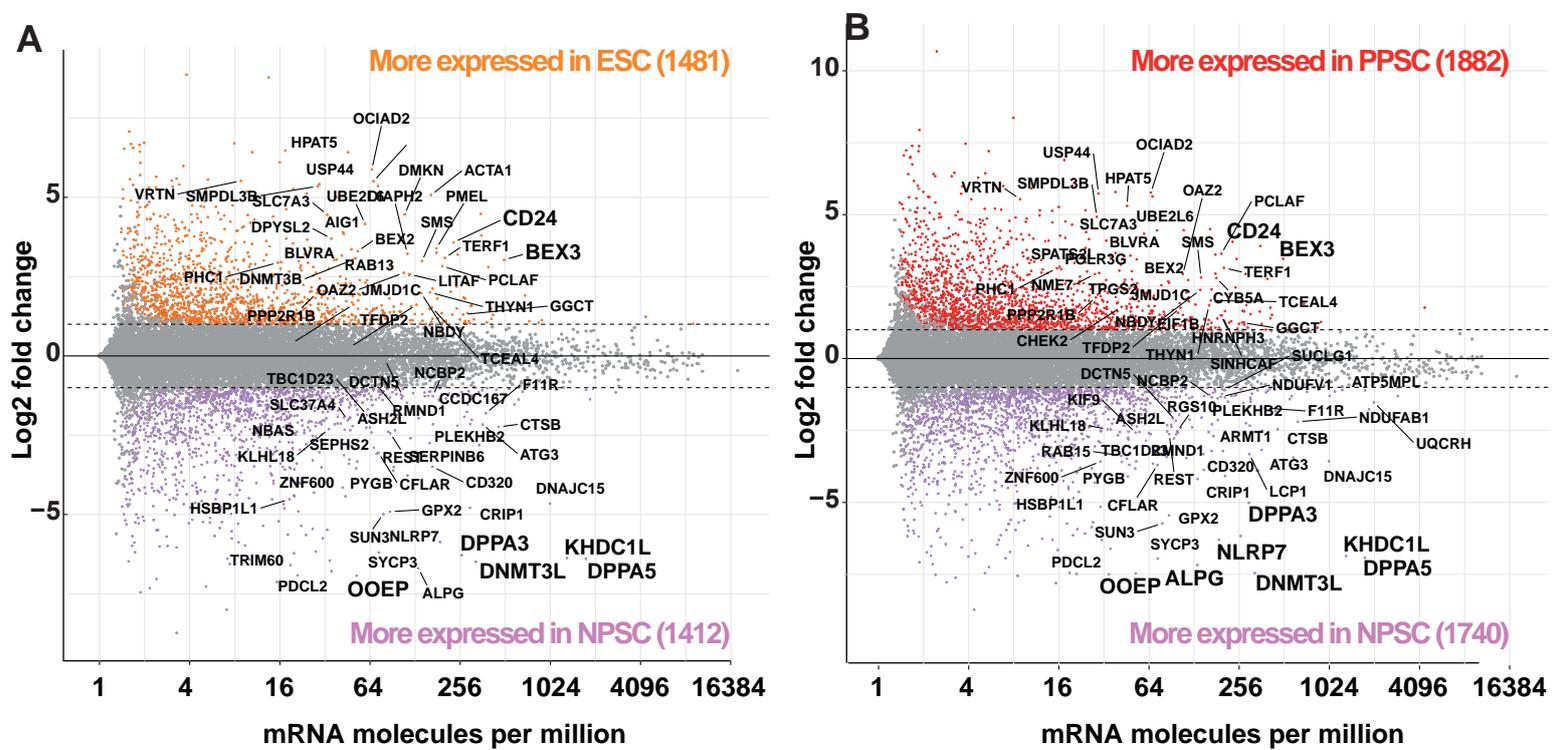




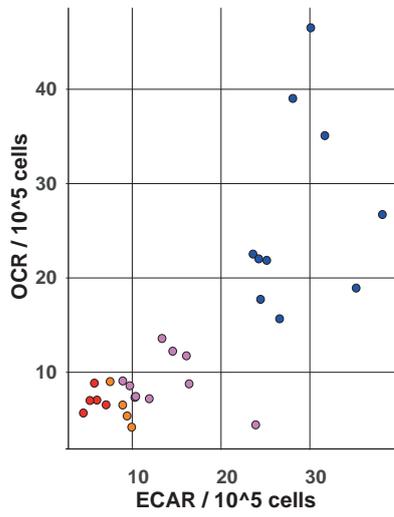




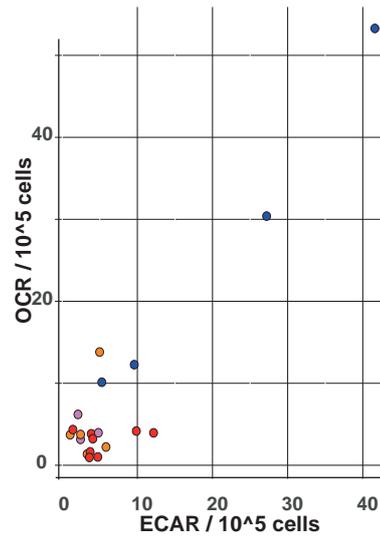




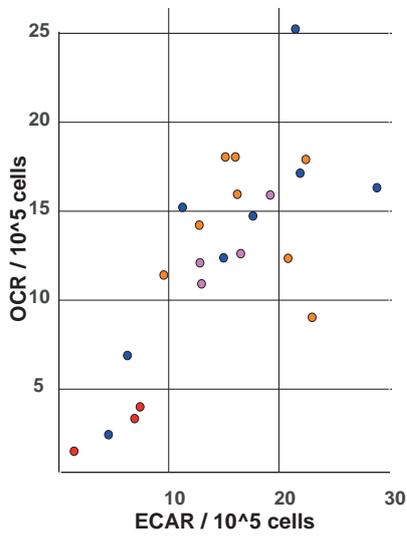
Onfray et al, sup Figure 3

A

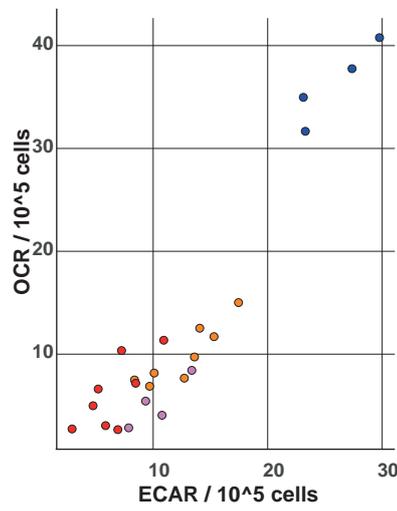
Replicate 1

B

Replicate 2

C

Replicate 3

D

Replicate 4

● TSC ● NPSC ● PPSC ● EPSC

Naive Pluripotent and Trophoblastic Stem Cell Lines as a Model for Detecting Missing Proteins in the Context of the Chromosome-Centric Human Proteome Project

Naive Pluripotent and Trophoblastic Stem Cell Lines as a Model for Detecting Missing Proteins in the Context of the Chromosome-Centric Human Proteome Project

Océane Girard ¹, Régis Lavigne ^{3,4}, Simon Chevolleau ¹, Constance Onfray ¹, Emmanuelle Com ^{3,4}, Pierre-Olivier Schmit ⁵, Manuel Chapelle ⁵, Thomas Fréour ^{1,6,7}, Lydie Lane ⁸, Laurent David ^{1,2}, and Charles Pineau ^{3,4,*}

¹ Nantes Université, CHU Nantes, Inserm, CR2TI, UMR 1064, F-44000 Nantes, France

² Nantes Université, CHU Nantes, Inserm, CNRS, BioCore, F-44000 Nantes, France

³ Univ Rennes, Inserm, EHESP, Irset (Institut de Recherche en Santé, Environnement et Travail) - UMR_S 1085, F-35000 Rennes, France

⁴ Univ Rennes, CNRS, Inserm, Biosit UAR 3480 US_S 018, Protim Core Facility, F-35000 Rennes, France

⁵ Bruker Daltonique SA, 34 rue de l'Industrie, F-67166 Wissembourg cedex, France

⁶ CHU Nantes, Service de Biologie de la Reproduction, F-44000 Nantes, France

⁷ Department of Obstetrics, Gynecology and Reproductive Medicine, Dexeus University Hospital, Barcelona, Spain

⁸ CALIPHO Group, SIB Swiss Institute of Bioinformatics and University of Geneva, Geneva, Switzerland

Corresponding author: charles.pineau@inserm.fr

Keywords: human proteome project, missing proteins, human naive pluripotent stem cells, human trophoblastic stem cells

Project Name: Proteomics of human naive pluripotent and trophoblastic stem cells

Project accession: PXD035768

Project DOI: 10.6019/PXD035768

ABSTRACT

The Chromosome-centric Human Proteome Project (C-HPP) aims at identifying the proteins as gene products encoded by the human genome, characterizing their isoforms and functions. The existence of products has now been confirmed for 93.2% of the genes at the protein level. The remaining mostly correspond to proteins of low abundance or difficult to access. Over the past years, we have significantly contributed to the identification of missing proteins in the human spermatozoa. We pursue our search in the reproductive sphere with a focus on early human embryonic development. Pluripotent cells, developing into the fetus, and trophoblast cells, giving rise to the placenta, emerge during the first weeks. This emergence is a focus of scientists working in the field of reproduction, placentation and regenerative medicine. Most knowledge has been harnessed by transcriptomic analysis. Interestingly, some genes are uniquely expressed in those cells, giving the opportunity to uncover new proteins that might play a crucial role in setting up the molecular events underlying early embryonic development. Here, we analyzed naive pluripotent and trophoblastic stem cells and discovered 4 new missing proteins, thus contributing to the C-HPP. The mass spectrometry proteomics data was deposited on ProteomeXchange under the dataset identifier PXD035768.

INTRODUCTION

The Human Proteome Project (HPP) is the flagship initiative of the global Human Proteome Organization (HUPO). Its main objective is to catalog proteins as gene products encoded by the human genome and credibly identify these essentially, but not entirely, by mass spectrometry¹. In 2020, the HPP celebrated its 10th anniversary with a major achievement, *i.e.*, a high stringency blueprint of the Human proteome detailing the detection of over 90% of all predicted human proteins². It has since increased this effort to 93.2% in 2022. In this context, the number of experimentally validated proteins (PE1) is annually updated by neXtProt, the reference protein knowledge-base for the HPP (www.nextprot.org) and has now reached 18,407 whereas the actual count of missing proteins (MPs) scored as PE2 or 3 or 4, now stands at 1,343.

There are many reasons, often combined, why missing proteins still lack evidence, either because they are expressed at low-copy numbers, because of their biology (*e.g.*, time- or stress-dependent, restricted to specific cell types or during pathophysiological situations) or because of detection limits by mass spectrometry due to some peculiar physicochemical properties (*e.g.*, hydrophobicity, basicity). In such regard, Lane and collaborators suggested that the proteins that have been systematically missed might be restricted to a few unusual organs or cell types, particularly the testis³. Unsurprisingly, the very high number of testis-specific genes supported the hypothesis that the testis was a promising organ in which to search for missing proteins. In a series of works, we succeeded in identifying in spermatozoa over 250 missing proteins whose expression is restricted and/or specific to the post-meiotic germ cell lineage⁴⁻⁷. Our mining of publicly available transcriptomic expression datasets indicates that both the male and female reproductive sphere organs are sources for detecting a significant number of missing proteins⁷. However, it can be anticipated that most of these will be particularly difficult to access due to their physicochemical properties (membrane proteins, defensins, etc.) or to narrow windows of expression.

In the present work, we pursue our search for missing proteins in the reproductive sphere with a focus on the early human embryonic development.

During pre-implantation development, the early human embryo successively undergoes two main morphological events, *i.e.*, compaction and cavitation. Alongside these morphological events and embryo growth, two cell fates specification are required at the morula and blastocyst stages. The first cell fate decision in the morula segregates the outside trophectoderm (TE) cells from inner cell mass (ICM). Subsequently, the second cell fate decision at the early blastocyst stage processes ICM cells to form the pluripotent epiblast (EPI) and primitive endoderm (PrE), precursors of the embryo and yolk sac, respectively⁸ (**Figure 1**). Proteomic analysis of human embryos has been scarce up to now, not only owing to the technical challenge, but also to ethical and legal issues, resulting in a limited access to human embryos for research.

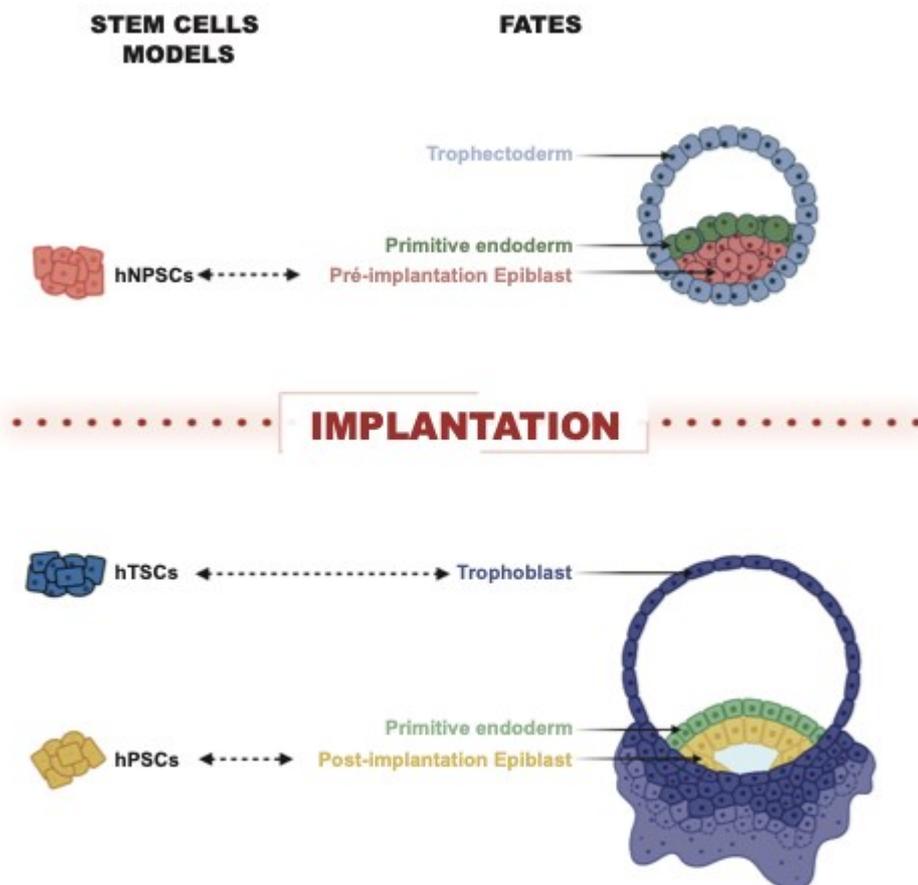


Figure 1: Schematic representation showing the correlation between embryo cell fate and stem cells. On the right, the three cell fates in the embryo before and after implantation with the trophectoderm (blue), the epiblast (pink and yellow) and the primitive endoderm (green). On the left, the stem cell models corresponding to the cell

fate with the hNPSCs (pink) as equivalent to the pre-implantation epiblast, the hTSCs (blue) as equivalent to the trophoctoderm and the hPSCs (yellow) as equivalent to the post-implantation epiblast (*Figure created with BioRender.com*).

We thus hypothesized that missing proteins might be accessible in human embryos. To overcome the above-mentioned challenges, stem cell models are commonly used as a proxy for human embryos. In order to study human peri-implantation, 3 main stem cell types are used. The first is naive pluripotent stem cells (hNPSCs) with the characteristics of the preimplantation epiblast⁹. The second is trophoblastic stem cells (hTSCs) with the characteristics of post-implantation cytotrophoblasts¹⁰. Those stem cells models are often compared to the primed pluripotent stem cells (hPSCs), not studied here, which represent post-implantation epiblast, a later stage of development than the hNPSCs⁹ (**Figure 1**).

Here we describe a strategy combining the analysis of RNASeq datasets and MS-based experiments to identify and validate missing proteins in hNPSCs and hTSCs. “We validated 4 missing proteins with at least 2 non nested unique peptides, according to the HPP guidelines¹¹, whereas single peptides were detected for 6 additional proteins. The use of the PaSER 2022 database search platform (Bruker Daltonik GmbH, Bremen, Germany) was also shown to be relevant and crucial for validating several unique peptides out of the MS proteomics data.

MATERIALS AND METHODS

Cell culture

All cell lines were cultured on a feeder cell layer, under hypoxic conditions (5%O₂, 5%CO₂). Culture medium was replaced daily. 10µM Y27632 (Axon Medchem, Groningen, The Netherlands) was added to the culture medium upon thawing and passaging of the cells.

hTSCs were cultured in hTSCs medium [DMEM/F12 (Gibco, ThermoFisher, Les Ulis, France) supplemented with 0.1mM 2-mercaptoethanol (Gibco), 0.2% FBS, 0.5% penicillin-streptomycin, 0.3% Bovine Serum Albumin (BSA, Sigma-Aldrich, St Louis, United States), 1% Insulin-Transferrin-Selenium-Ethanolamine supplement (ITS-X, Gibco), 1.5 mg/ml L-ascorbic acid (Sigma-Aldrich), 50 ng/ml hEGF (Miltenyi Biotec, Bergisch Gladbach, Germany), 2 µM CHIR99021 (Axon Medchem), 0.5 µM A83-01 (Tocris, Bristol, United Kingdom), 1 µM SB431542 (Tocris), 0.8 mM valproic acid (Sigma-Aldrich) and 5 µM Y27632]. hiTSCs (human induced naive pluripotent stem cells) were passaged with TrypLE (5 min, 37°C, Life Technologies Corporation, Carlsbad, United States) every 4 to 5 days at a cell density between 1.04*10⁴ and 2.08 *10⁴ cells per cm².

hNPSCs were cultured in PXGLY medium [47.5% Neurobasal medium (GIBCO) and 47.5% DMEM/F12 (Gibco) supplemented with 1mM N2 (GIBCO), 2mM B27 (Gibco), 1mM glutamax (Gibco), 1mM non-essential amino acids (GIBCO), 0.45% BSA, 1mM sodium pyruvate (Gibco), 0.1% 2-mercaptoethanol (Gibco), 0.5% penicillin-streptomycin, 1µM PD0325901 (Axon Medchem), 2µM XAV939 (Axon medchem), 2µM Gö6983 (Axon medchem), 10ng/mL hLif (PeproTech, Neuilly-sur-Seine, France), 10µM Y27632 (Axon Medchem)]. hNPSCs were passaged every 4 to 5 days with TrypLE (5 min, 37°C, Life Technologies) at a cell density of 2.08 *10⁴ cells per cm².

Prior to the experiment, cells were submitted to two dissociation and lysis protocols. In the first one, cells were dissociated using TrypLE (5min, 37°C), feeders were removed by incubation on 0.1% gelatin-coated plates layer for 1 hour and cells were lysed with the iST Kit (PreOmics GmbH, Planegg, Germany) following the manufacturer's instructions. In the second one, cells were plated overnight after feeder removal on respectively 0.1% geltrex coated plate for hNPSCs or on 3µg/mL

vitronectin and 1 μ g/mL laminin coated plate for hTSCs. 24 hour after seeding, cells were processed with the iST Kit.

Protein Extraction, Digestion, and Liquid Chromatography—Tandem Mass Spectrometry (LC-MS/MS) Analyses

Briefly, samples were thawed and lysed (denatured, reduced and alkylated) for 10 min at 95°C then Trypsin/LysC digested for 3hours at 37°C. Purification of peptides was then carried out at room temperature on spin cartridge and peptides were finally eluted in 10 μ L of LC-load buffer. Simultaneously, a protein assay has been realized to precisely know the quantity of proteins present in the samples. Once purified, hTSCs and hNPSCs samples were peptide assayed to prepare for mass spectrometry injection at an amount of 10 μ g in 10 μ L.

Approximately 200ng each of tryptic peptides samples were separated onto a 75 μ m x 250mm IonOpticks Aurora 2 column (Ion Opticks Pty Ltd, Australia) packed with a 120 Å pore, 1.6 μ m particle size C18 beads. A reversed-phase gradient of basic buffers (Buffer A: 0.1% formic acid, 98% H₂O MilliQ, 2% acetonitrile ; Buffer B: 0.1% formic acid, 100% acetonitrile) was run on a NanoElute HPLC System (Bruker Daltonik) at a flow rate of 400 nl/min at 50°C. The liquid chromatography (LC) run lasted for 120 min with a starting concentration of 2% of buffer B increasing to 15% over the initial 60 min, a further increase in concentration to 25% over 30 min, then to 37% in ten minutes, and finally to 95% in ten minutes again. This elution gradient was followed by a 95% wash during ten minutes and re-equilibration.

Temperature of the separation column was regulated thanks to the Sonation column oven. The constant column temperature makes measurement significantly more accurate and higher column temperatures could also be applied allowing us to apply the flow of 400 nl/min while maintaining the same pressure. The NanoElute HPLC system was coupled online to a Tims TOF Pro mass spectrometer (Bruker Daltonik) with a CaptiveSpray ion source (Bruker Daltonik). The CaptiveSpray nanoflow ESI source was directly attached to a vacuum inlet capillary via a short capillary extension heated using the instrument's drying gas. High voltage for the electrospray (ESI) process was applied to the vacuum capillary inlet, whereas the sprayer was kept at ground. Temperature of the ion transfer capillary was set at 180°C. The spray type was automatically mechanically aligned on the axis with the capillary inlet without the need for any adjustment. Ions were accumulated for 114 ms, and mobility separation

was achieved by ramping the entrance potential from -160 V to -20 V within 114 ms. The acquisition of the MS mass spectra with the TIMS TOF Pro is done with an average resolution of 60 000 FWHM (mass range 100-1700 m/z). To enable the PASEF method, precursor m/z and mobility information was first derived from full scan TIMS-MS experiments (with a mass range of m/z 100-1700). Resulting quadrupole mass, collision energy and switching times were automatically transferred to the instrument controller as a function of the total cycle time. The quadrupole isolation width was set to 2 and 3 Th and, for fragmentation, the collision energies varied between 31 and 52 eV depending on precursor mass and charge. TIMS, MS operation and PASEF were controlled and synchronized using the control instrument software OtofControl 6.2 (Bruker Daltonik). LC-MS/MS data were acquired using the PASEF method with a total cycle time of 1.28 s, including 1 TIMS MS scan and 10 PASEF MS/MA scans. The 10 PASEF scans (110ms each) contain on average 12 MS/MS scans per PASEF scan. In addition, the most abundant precursors which could have been sequenced in previous scan cycles are dynamically excluded from re-sequencing. The acquisition of the MS/MS mass spectra with the TIMS TOF Pro is also done with an average resolution of 50 000 FWHM (mass range 100-1700 m/z).

Protein identification

Ion mobility resolved mass spectra, nested ion mobility vs m/Z distributions, as well as summed fragment ion intensities were extracted from the raw data file with DataAnalysis 6.0 (Bruker Daltonik). Signal-to-noise (S/N) ratio were increased by summations of individual TIMS scans. Mobility peak positions and peak half-widths were determined based on extracted ion mobilograms (± 0.05 Da) using the peak detection algorithm implemented in the DataAnalysis software. Features detection were also performed using DataAnalysis 6.0 software and exported in .mgf format.

Peptides identification were performed with the Mascot search engine (version 2.6.2, Matrix Sciences), applying the previously described search parameters and using its automatic decoy database search to calculate the false discovery rate (FDR)^{4,7}. The search database was the complete human proteome homo sapiens UP000005640 from UniProtKB 2022_02 restricted to one protein sequence per gene. Briefly, 1 miscleavage for trypsin was allowed and mass tolerance of peptides and fragments was established at 15 ppm and 0,05 Da. Moreover, mass modifications of peptides

are taken into account. For fixed modifications, carbamidomethylation of cysteines and for variable modifications, oxidations and acetylation of lysines and N-term proteins were considered. After interrogations on Mascot, data processing was performed using the Proline software (version 2.1.0). All the results of the queries performed on Mascot were imported into Proline with a subset threshold of 1. After importation, the results were validated with a peptide pretty rank of 1, an FDR for PSM of 1% on adjusted e-value and an FDR for protein set of 1% with a standard scoring.

Using the Mascot search engine, the total number of expected false positives for hNPSCs classically dissociated from cultures dishes with trypsin, was 0.06% for PSM, 0.09% for peptide set and 0.51% for protein set. For hNPSCs directly dissociated from culture plates with the lysis buffer, the total number of expected false positives was 0.06% for PSM, 0.09% for peptide set and 0.52% for protein set. The merge of the two hNPSCs conditions provided an expected false positive rate of 0.12% for PSM and peptide set and 0.77% for protein set. Similarly, the total number of expected false positives for hTSCs classically dissociated from cultures dishes with trypsin, was 0.06% for PSM, 0.09% for peptide set and 0.53% for protein set. For hTSCs directly dissociated from culture plates with the lysis buffer, the total number of expected false positives was 0.06% for PSM, 0.09% for peptide set, 0.52% for protein set. The merge of the two hTSCs conditions provided an expected false positive rate of 0.13% for PSM and peptide set and 0.85% for protein set.

On top of the standard Mascot search, all datasets were searched using the Graphics Processing Units (GPU)-based PaSER 2022 V3 solution (Bruker). PaSER was configured to use the ProLucid search engine¹² with the complete human proteome homo sapiens UP000005640 from UniProtKB 2022_02 restricted to one protein sequence per gene and with modification definitions as the ones used with Mascot. With PaSER, the acetylation of lysines and N-term proteins was parameterized at the peptide level and not at the protein level like Mascot. The mass tolerance of peptides for PaSER was fixed at 30 ppm. Both protein and peptide FDR thresholds were set to 1%. TIMScore was enabled to allow the use of the peptide Collisional Cross Section (CCS) during the scoring process.

When using the PaSER search engine, the total number of expected false positives for hNPSCs classically dissociated from cultures dishes with trypsin, was 0.22% for PSM, 0.25% for peptide set, 0.99% for protein set, whereas for hNPSCs directly

dissociated from culture plates with the lysis buffer, the total number of expected false positives was 0.30% for PSM, 0.36% for peptide set, 0.98% for protein set. For hTSCs classically dissociated from cultures dishes with trypsin, the total number of expected false positives was 0.31% for PSM, 0.36% for peptide set, 0.99% for protein set, whereas for hTSCs directly dissociated from culture plates with the lysis buffer, the total number of expected false positives was 0.25% for PSM, 0.33% for peptide set, 0.98% for protein set.

An essential component to TIMScore is defining the deviation between experimental and predicted CCS values. Machine learning was used in order to accurately predict CCS values from a peptide's primary amino acid sequence. A training dataset of hundreds of thousands of tryptic and phosphorylated peptides was used, where the dataset included peptides of doubly, triply and quadruply charge states. A transformer model of peptide CCS was developed from this training set. The model was tested for accuracy against an independent dataset it had previously not seen. For doubly, triply and quadruply charged peptides the accuracy in predicting a peptide CCS from the primary amino acid sequence was 95% for tryptic peptides. Upon setting up the parameters file, *in silico* peptide candidates are sent to the CCS prediction model to generate a predicted CCS value. The PaSER search algorithm (in our case, ProLucid) is run as normal and the search algorithm compares the predicted and measured CCS values and calculates a correlation score, namely TIMScore for the top 5 peptide candidates for each spectra.

Peptide and protein identifications summaries were generated for each sample with both the Mascot and the PaSER search engines. Datasets were further analyzed in accordance with version 3.0 the HPP data interpretation guidelines¹¹. Finally peptide-to-protein mappings were checked using the neXtProt uniqueness checker¹³.

3'SRP datasets

3'seq-RNA Profiling (3'SRP)¹⁴ datasets were obtained from Kilens et al (2018) and Castel et al (2020). Concerning the plots made from 3'SRP datasets, the values lower than 0.1 mRNA molecules per million of mRNA molecules have been arbitrarily represented as 0.1 mRNA molecules per million of mRNA molecules for convenience.

Data Availability

The mass spectrometry proteomics data, including raw files and identification files, form a complete submission with the ProteomeXchange Consortium¹⁵. Data were submitted via the PRIDE partner repository under the dataset identifiers PXD035768 and 10.6019/PXD035768".

RESULTS & DISCUSSION

The aim of the present work was to pursue our ongoing project of possibly detecting numerous missing proteins in human tissues and cells of the reproductive sphere. Here, we assessed whether hNPSCs and hTSCs were relevant biological material for searching for missing proteins by analyzing expression in these cells of 1343 mRNAs corresponding to referenced missing proteins in the latest neXtProt release (2022.02.25).

hNPSCs and hTSCs are recently discovered cell types^{16,17}. They are also representative of a unique developmental time, with specific features: hNPSCs are modeling the pre-implantation epiblast, that need to remain pluripotent and proliferate to form the source of the fetus, while hTSCs represent cytotrophoblast, the “placental stem cells”. We surmised that those unique and transient developmental stages might use specific gene sets, potentially restricted to this developmental stage. As a consequence, we started by analyzing 3'SRP datasets^{10,18} corresponding to the two stem cell lines. 3'SRP uses unique molecular identifiers (UMI) to correct errors in quantification of mRNA, hence reflecting the likeliness of a gene to be expressed at a level that corresponds to protein expression, and not background. This was for example the case for DPPA5 which is often detected by qPCR in both hNPSCs and hPSCs, but that is only detected by western blot when the expression levels are around 5000 mRNA molecules per million of mRNA molecules (in hNPSCs) and not detected when expression levels are around 10 molecules per million of mRNA molecules¹⁸.

811 genes corresponding to 1343 missing proteins were unambiguously annotated in transcriptomic datasets. Out of these 811 genes, only 14 genes expressed in hNPSCs (**Figure 2A**) and 9 genes expressed in hTSCs (**Figure 2B**) had abundance over the threshold for credible expression, which is 20 mRNA molecules per million of RNAs. We organized proteins based on their relative expression in hNPSCs or hTSCs. We also displayed expression levels in hPSCs to highlight that the identified missing proteins are globally specific of hNPSCs and might therefore correspond to the preimplantation epiblast (**Figure 2C**).

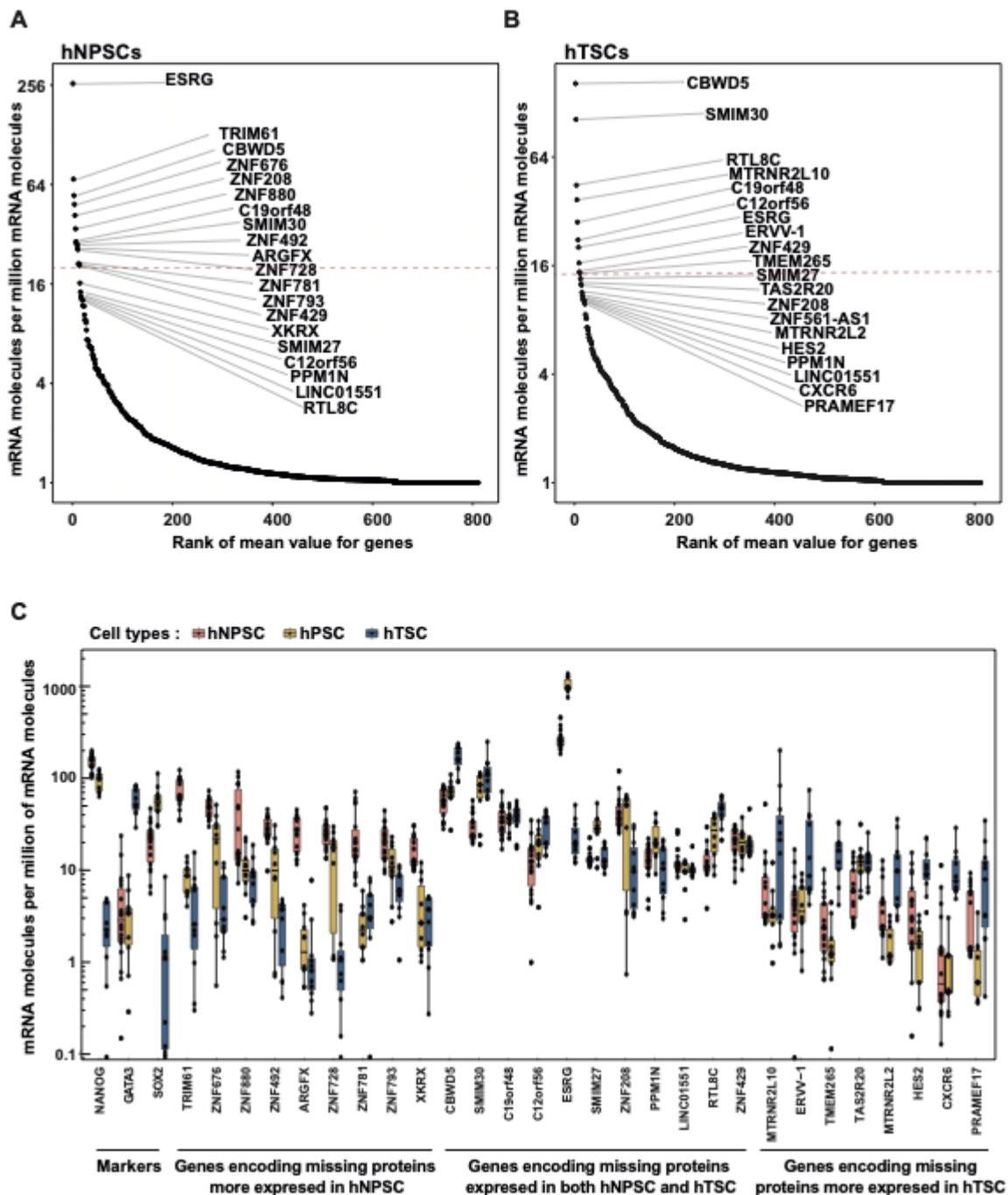


Figure 2: Transcriptomic expression levels of missing proteins in hNPSCs and hTSCs. (A, B). Distribution of average transcriptomic expression levels ($\log_2(\text{mRNA molecules per million of mRNA molecules}+1)$) for hNPSCs (**A**) and hTSCs (**B**). We highlighted the top 20 genes; for hNPSCs: ESRG, TRIM61, CBWD5, ZNF676, ZNF208, ZNF880, C19orf48, SMIM30, ZNF492, ARGFX, ZNF728, ZNF781, ZNF793, ZNF429, XKRX, SMIM27, C12orf56, PPM1N, LINC01551 and RTL8C; for hTSCs: CBWD5, SMIM30, RTL8C, MTRNR2L10, C19orf48, c12orf56, ESRG, ERVV-1,

ZNF429, TMEM265, SMIM27, TAS2R20, ZNF208, ZNF561-AS1, MTRNR2L2, HES2, PPM1N, LINC01551, CXCR6 and PRAMEF17.

(C) Gene expression levels of the 20 most expressed MPs in hNPSCs (pink), hPSCs (yellow) and hTSCs (blue). NANOG and SOX2 are markers of pluripotency, with higher expression in hNPSCs and hPSCs, respectively. GATA3 is a marker of hTSCs. TRIM61, ZNF676, ZNF880, ZNF492, ARGFX, ZNF728, ZNF781, ZNF793, XKRX are in the 20 most expressed MPs in hNPSCs when RTL8C, MTRNR2L10, ERVV-1, ZNF429, TMEM265, TAS2R20, MTRNR2L2, HES2, CXCR6, PRAMEF17 are in the 20 most expressed MPs in hTSCs. RTL8C, CBWD5, SMIM30, C19orf48, C12orf56, ESRG, SMIM27, ZNF208, PPM1N, LINC01551 are present in the 2 top 20 MPs most expressed (hTSCs and hNPSCs). For each box, the median, the first and third quartile are displayed.

Interestingly, the 9 genes with highest expression levels in hNPSCs but not in hTSCs are also 3 to 30 times more expressed in hNPSCs than hPSCs, therefore of great potential interest to understand the specificity of preimplantation vs postimplantation EPI (**Figure 2C, Supplementary Table 1**).

Total protein digests from cells dissociated with two different protocols were analyzed by MS/MS using Mascot and Proline search engines. This analysis allowed the identification of 44247 peptides for hNPSCs and 44618 peptides for hTSCs, further mapping to 5150 and 5253 proteins respectively. We analyzed the gene expression levels of genes corresponding to proteins identified by MS/MS. This showed that in both hNPSCs and hTSCs, genes that were expressed over 20 mRNA molecules per million of mRNA molecules had 70% MS identification for hNPSCs and 66% MS identification for hTSCs (**Supplementary Figure 1**). This further confirmed the likelihood of identifying missing proteins in hNPSCs and hTSCs (**Figure 2A, B**).

In a second time the same datasets were processed using PaSER with the objective to catch new peptides that would not have been seen by Mascot. When using the PaSER search engine, 58446 peptides from 5906 proteins were identified in hNPSCs classically dissociated from cultures dishes with trypsin, and 51273 peptides from 5297 proteins were identified in hNPSCs directly dissociated from culture plates with the lysis buffer. As many as 61339 peptides from 6073 proteins were identified in

hTSCs dissociated from culture dishes with trypsin, and 51408 peptides from 5513 proteins in hTSCs directly dissociated from culture plates with the lysis buffer. Detailed information on the proteins identified from our MS datasets is reported in **Supplementary Table 2**.

The list of all identified proteins in our study was searched for missing proteins against neXtProt data (release 2022-02-25). The overall workflow for the detection and validation of missing proteins was used as previously described⁶. Applying this workflow, we produced a list of 16 and 18 “candidate missing proteins” entries in hNPSCs and hTSCs respectively.

In hNPSCs, UQCRHL, CTAGE15, MAP1LC3B2, ZNF98, ZNF732, ZNF728, ZNF208, ZNF804B, ZNF117, ZNF676, ZNF492, RGPD1, CPSF4L, TRIM61, NANOGP8 and WASH2P were considered. In hTSCs, the missing proteins examined correspond to 18 genes: PPIAL4E, PPIAL4C, PPIAL4D, PPIAL4F, TRBV18, RGPD1, CTAGE6, CTAGE15, MAP1LC3B2, DDTL, CGB1, CGB2, CGB7, ZNF732, FBXO47, WASH2P, OR1M1 and OR5M8.

These two subsets were further analyzed in line with version 3.0 of the HPP mass spectrometry data interpretation guidelines¹¹. Peptides smaller than 9 amino acids in length were removed and peptide-to-protein mappings were checked using the neXtProt uniqueness checker¹³, considering alternative mappings by taking into account the 9.7 million single amino acid variants currently available in neXtProt. As a result, 4 missing proteins could be validated in hNPSCs with at least 2 unique, non-nested peptides of at least 9 amino acids, but none in hTSCs. In addition, single “one-hit wonder” peptides uniquely mapping to 6 other missing proteins could be proposed to the community. Full details (description, number of unique peptides, chromosome location, etc.) of this analysis are reported in **Table 1**.

Table 1: List of Missing Proteins (PE2-PE4) identified in hNPSCs.

Table 1. List of Missing proteins (PE2-PE4) identified in the two datasets

Localization	UniProt ID	Gene ID	Chromosome location	Peptide sequence	Peptide identified by	Spectrum quality's	PE level	Protein name	Functional note	Number of validated peptides	SAAV
	P00KX0	ZNF728	19p12	DFNGSSHLTHK	PaSER+Mascot iProline	Medium	PE2	Zinc Finger Protein 728	Involved in regulation of transcription by RNA polymerase II	5	No
				AFSWVSVLKKHK	MascotProline	Low					No
				IGCTNVDECK	PaSER	High					No
				AFWSSRLSEHK	PaSER	Medium					No
				YANIFKCSMSK	PaSER	Low					No
INPSCs	Q5EBN2	TRIM61	4q32.3	FISNPQLGSLTEIAK	PaSER+Mascot iProline	High	PE2	PuIative tripartite motif-containing protein 61	Ubiquitin protein ligase activity	2	No
				LEEYNAPWK	PaSER	Medium					No
	Q5NTQ3	ZNF676	19p12	GFSSVSTLNTHK	PaSER	High	PE2	Zinc Finger Protein 676	Involved in regulation of transcription by RNA polymerase II	2	No
				AFWSSILTEHKIHTGEEK	PaSER	Low					No
	A4D1E1	ZNF804B	7q21.13	DFSVLAK3NHISMTSK	PaSER	Low	PE2	Zinc Finger Protein 804B	Enables metal ion binding	2	No
				CQEQSSNVEISSNSCK	PaSER	Low					No

All peptides but three could only be identified thanks to a search via the PaSER engine (Table 1). The capacity of the PaSER search pipeline to detect more proteins can be attributed to the use of the peptides's CCS value to add an extra scoring dimension (TIMScore). The true benefit of TIMScore can be realized during the peptide-validation and FDR estimation steps of the proteomics pipeline. In a non-CCS enabled algorithm, such as Mascot, only two dimensions can be utilized to estimate the FDR rate, and

so a discriminate line is fit to a 1% error to distinguish forward and reverse peptide candidates. With TIMScore, and the extra CCS dimension, the peptide-candidates can be vectorized in 3-dimensions, allowing a discriminate contoured plane to be applied to achieve the same 1% error. Applying a discriminate plane provides increased accuracy and precision, helping to validate formerly poorly scoring PSMs in the standard two dimensions. Thus, the key effect of TIMScore is derived from the additional dimension of CCS that it provides in assigning true positives from decoy peptide sequences. TIMScore works in a bidirectional fashion, boosting the confidence of borderline peptides under strict FDR thresholds while simultaneously lowering the probability score of a peptide candidate such that it falls below the level of detection. Additionally, the probability score differentiates ambiguous PSMs where the traditional search score cannot distinguish between the first and second (or more) best candidates. Spectra corresponding to peptides exclusively identified using PaSER in this study are provided as **Supplemental material 1**.

Fluorescence immunocytochemical studies on hNPSCs in suspension were undertaken using HPA antibodies to provide orthogonal evidence of the expression of TRIM61 and ZNF728 selected based on our data mining process. However, results were not convincing enough using both antibodies and the approach requires further optimization before drawing any conclusion.

We then carried out a rapid search of the literature and knowledge bases to highlight the potential relevance of the identified missing proteins to the reproduction field. Putative tripartite motif-containing protein 61 (TRIM61) is a RING finger domain protein that is predicted to bind four zinc cations. Many proteins containing a RING finger play a key role in the ubiquitination pathway and TRIM61 is predicted to possess a ubiquitin protein ligase activity¹⁹. According to the Human Protein Atlas (<https://www.proteinatlas.org>), the gene is expressed in a large array of tissues and the protein may be localized in the cytoplasm, nucleoli fibrillar center and endoplasmic reticulum. In the mouse, Trim61, also called Rnf35, is transcribed temporally in the preimplantation mouse embryo, predominantly at the two-cell embryonic stage. However, the gene is permanently silenced before the blastocyst stage of development. It is thus supposedly implicated in zygotic gene expression²⁰. Huang and collaborators have later demonstrated that Rnf35 was actively transcribed from

the newly formed embryonic genome between the late 1-cell and 2-cell stages of early development²¹. Finally, the Trim61 mRNA was shown to bind to Cpeb, a sequence-specific RNA-binding protein that regulates polyadenylation-induced translation, that controls oocyte growth and follicle development in the mouse²². The pool of information gathered here on TRIM61 is an example of what should be provided for all PE1 proteins and centralized on a specific repository in the frame of the newly launched HPP “Grand Project”.

The three other missing proteins identified in the present study are classical C2H2 type zinc-finger proteins, *i.e.*, ZNF728, ZNF676 and ZNF804B. They contain one (ZNF804B) or several (ZNF728, ZNF676) zinc-finger domains that consist of short 30 amino acid motifs making tandem contacts with a target molecule. ZNF motifs are stabilized by one or more zinc ions. In numerous C2H2 type zinc-finger proteins, the motif mediates direct interaction with DNA. ZNF728 and ZNF676 are hominoid specific proteins that were predicted to be DNA-binding transcription factors modulating the transcription of specific gene sets transcribed by the RNA polymerase II¹⁹. Additionally, they both contain a Krüppel associated box (KRAB) domain, typically found in transcription repressors. ZNF676, and possibly ZNF728, have been recently shown to repress the transcriptional activity of a subset of ERV-embedded regulatory sequences active during gametogenesis and early development of the egg²³.

Using our 3'SRP datasets analysis^{10,18} corresponding to hNPSCs cells, we observed a concordance between our proteomics and transcriptomics analyses (**Figure 3**).

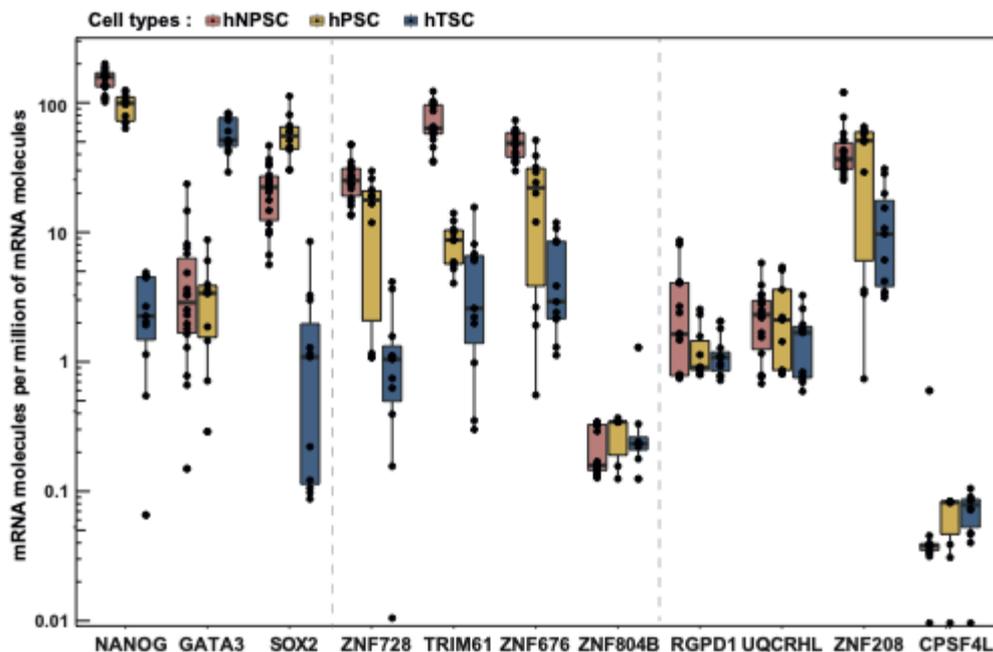


Figure 3. Gene expression levels of missing proteins identified for hNPSCs, hPSCs and hTSCs. Expression levels of the 8 MPs specifically studied in hNPSCs (pink), hPSCs (yellow) and hTSCs (blue). NANOG, GATA3 and SOX2 are included as controls, as in Figure 2. For each box, the median, the first and third quartile are displayed.

Indeed, three MPs identified (TRIM61, ZNF676, ZNF728) at protein level are present in the top 11 MPs most expressed at transcriptomic level. We also observed a difference in expression between hNPSCs and hTSCs/hPSCs with greater transcriptomic expression in hNPSCs than in hTSCs and hPSCs for these three MPs. That is consistent with our proteomic analysis by mass spectrometry. Of note, for our last identified MPs, *i.e.*, ZNF804B, the transcriptomic expression is low in hNPSCs, hTSCs and hPSCs. Additional unique peptides for this protein need to be identified for reinforcing its identification (**Table 1**).

Six other proteins (PE2) (RGPD1, UQCRHL, ZNF208, CPSF4L, TRBV18 and OR5M8) were detected with only one distinct uniquely mapping peptide of length ≥ 9 amino acids (**Table 2**). For these 6 MPs, no additional peptide considered unique but smaller than 9 amino acids could be identified in our datasets. As a consequence, we must be particularly careful with these peptide identifications. Thus, the unambiguous

validation of the corresponding MPs will rely on the identification of other unique peptides.

Table 2: List of “one-hit wonder” proteins to be further validated.

Table 2. List of “one-hit wonder” proteins to be further validated

Localization	UniProt ID	Gene ID	Chromosome location	Peptide sequence	Peptide identified by	Spectrum quality ^a	PE level	Protein name	Functional note	Number of validated peptides	SAAV
	POCJD0	RGPD1	2p11.2	MNVVMGFNTDR	PaSER	Low	PE2	Ran-binding protein 2-like 9	Contributes to GTPase activator activity	1	No
	A0A069LP95	UQCRHL	1p36.21	ERLELYDEHVSSR	PaSER	Medium	PE2	Cytochrome b-c1 complex subunit 6-like, mitochondrial	Involved in mitochondrial electron transport, ubiquinol to cytochrome c	1	No
HNPSCs	Q43345	ZNF208	19p12	WSSTLSYHK	PaSER	Medium	PE2	Zinc Finger Protein 208	Involved in regulation of transcription by RNA polymerase II	1	No
	A6NMK7	CPSF4L	17q25.1	MVVCKRMLR	PaSER	Medium	PE2	Positive cleavage and polyadenylation factor subunit 4-like protein	Enables metal ion and RNA binding	1	No
	A0A067X0M5	TRBV18	7q34	FMVYLOKNIIDSGMPK	PaSER	Low	PE2	T cell receptor beta variable 18	Involved in adaptive response	1	No
hTSCs	Q8NGP6	OR6M8	11q12.1	ELSMKIVFS	PaSER	Low	PE2	Olfactory receptor 5M8	Enables G protein-coupled receptor activity and enables olfactory receptor activity	1	No

For most of these proteins, limited information is available in the literature. The RANBP2-like and GRIP domain-containing protein 1 (RGPD1) is expressed in a large

array of tissues, group enriched in the cervix, placenta and testis, and mainly localized in nuclear membranes. The protein is predicted to contribute to GTPase activator activity¹⁹. Of note is that this protein is identical to RGP2 (NX_P0DJD1) except in the first 16 amino acids. The peptide identified in the present work (aa 1-10) is the only one that allows the two proteins to be distinguished. Further validation of this protein in the frame of the HPP will be impossible with the current guidelines. This is clearly a case that justifies a derogation in the current HPP mass spectrometry data identification guidelines or the addition of a dedicated paragraph in its future version.

UQCRHL, the Cytochrome b-c1 complex subunit 6-like, mitochondrial protein is expressed in a large set of tissues and cell-type enriched in cardiomyocytes. The protein is predicted to be localized in the mitochondria inner membrane where it would be a component of the respiratory chain complex III and catalyze the oxidation of ubiquinol by oxidized cytochrome c1. Yet another unique peptide for UQCRHL has been reported in Peptide Atlas. Thanks to our contribution, this protein should be validated in the coming future.

ZNF208 is another Krüppel C2H2-type zinc-finger protein with low tissue specificity and that can be considered as a transcription factor. It is among the highest expressed transcripts in hNPSCs and hTSCs. Of note is that another unique peptide for ZNF208 has been reported in Peptide Atlas. Again, thanks to our contribution, this protein should be soon validated.

CPSF4L, the Putative cleavage and polyadenylation specificity factor subunit 4-like is predicted to be a RNA-binding protein involved in pre-mRNA cleavage required for polyadenylation¹⁹. CPSF4L mRNA expression level is surprisingly low, questioning whether this protein is present in the hNPSCs or whether its expression is difficult to detect by 3'SRP.

TRBV18 is the variable region 18 of the T cell receptor beta chain located in the plasma membrane. This region is responsible for recognizing fragments of antigen as peptides bound to major histocompatibility complex (MHC) molecules.

Finally, OR5M8 is the olfactory receptor 5M8. It is predicted to be a G protein-coupled receptor, involved in the detection of chemical stimuli and sensory perception of smell. The unique peptide identified in our study corresponds to a cytoplasmic domain of the protein. Considering no olfactory receptor has been so far unambiguously identified by mass spectrometry in the frame of the HPP, additional effort is planned in our

laboratory to further analyze hNPSCs and hTSCs protein extracts. As a matter of fact, other unique peptides corresponding to extracellular domains of OR5M8 could potentially be obtained through a controlled sample digestion by some other enzymes. Interestingly there is another long unique peptide reported by MassIVE: ESVEQGKMVAVFYTTVIPMLNLIYSLRNKNVKEALIK (mzspec:PXD022531;j7912_PDIA6.mzXML:scan:9162:ESVEQGKMVAVFYTTVIPMLNLIYSLRNKNVKEALIK/3). Of note is that the mRNA corresponding to TRBV18 and OR5M8 were not or barely detected in the list of 24,849 transcripts generated during our transcriptomic analysis; which eventually makes the identification of these unique peptides questionable and calls for a cross-validation by the trans-Proteomics Pipeline. As far as OR5M8 is concerned, as the identification of the unique peptide was good, the only credible explanation is that its mRNA was at the limit of detection of the 3'seq-RNA profiling approach used.

The present study shows that we have reached lower technical limits for identifying missing proteins in mass spectrometry but also in sample preparation. In the present work, the use of the TIMScore and peptide Collisional Cross Section (CCS), only made possible on a Trapped Ion Mobility Spectrometry instrument (*i.e.*, Tims TOF Pro; Bruker Daltonik), was shown to be a valuable additional feature that strengthens peptide identification by mass spectrometry. Thus, the PaSER search engine will be systematically used in our future studies.

As regards sample preparation, interestingly, MPs evidenced in this work are generally recognized as discrete gene products, *e.g.*, transcription factors. We are thus continuing our efforts to identify missing proteins in hNPSCs and hTSCs extracts thanks to additional enrichment strategies offering better resolution to favor the selective extraction of membranous, membrane-bound and nuclear proteins. The protein digestion tool recently developed by neXtProt (www.nextprot.org/tools/protein-digestion) can also be used wisely to determine enzymes alternative to trypsin and select the experimental conditions that would yield additional unique peptides to confidently identify missing proteins.

Among the top 20 missing proteins expected to be found in hNPSCs and hTSCs cellular extracts based on mRNA expression, several were not identified in the present study. Interestingly, according to the literature and to HUGO Gene Nomenclature

Committee curators, the LINC01551 and ZNF561-AS1 genes are probably not protein coding. Additionally, the expression of several proteins cannot be validated using the current HPP guidelines. A few examples are CBWD5 that has only 1 amino acid difference with CNWD3; the mature form of SMIM30 is 35 amino acids long and MTRNR2L10 and MTRNR2L2 are also short proteins <30 amino acids.

We should then focus on proteins such as ESRG that are highly expressed and should be found even with conventional trypsin digestion. Sample preparation will be key here to access the missing proteins that could not be seen in the first round. Additionally, we plan to use targeted mass assays using a PRM acquisition approach to target unique peptides matching the current HPP data interpretation guidelines.

CONCLUSION

We demonstrate here that early development stages harbor unique missing proteins. Indeed, this is the first time that these 4 proteins have been found in particularly understudied samples. They might be involved in events restricted to the naive stem cells lineage, and have a crucial role in setting up the molecular events that underlie early embryonic development. Identifying the transcription factors involved in the establishment and maintenance of human naive pluripotency is an important focus for the field. Knowing that those transcription factors exist at the protein level will support further biological investigations, and the development of gene-editing and invalidation approaches to create knock-out cell lines (e.g., CRISPR-Cas9). Moreover, identification of those transcription factors will also trigger their further study in human embryos.

In this paper several peptides unique to missing proteins were identified thanks to the PaSER search engine. This new validation strategy, supported by the PaSER search engine, makes use of the correlation between theoretical (predicted) and measured CCS value for each peptide as an extra scoring dimension. This correlation is combined to the more classical fragmentation-based correlation pattern to rescue out of the 1% FDR plane some of the peptides showing a low-quality MS/MS fragmentation pattern but with a good CCS correlation. The feasibility of an accurate CCS value prediction for peptides has been demonstrated recently and this feature is now exploited as an extra filtering value for peptide candidates along the identification

process, and is also used in the FDR calculation process. All presented peptides have passed a 1% peptide FDR threshold.

The HPP mass spectrometry data interpretation guidelines version 3.0 do not take into consideration CCS values. To date, CCS values can only be generated when mass spectrometry analyses are performed on a mass spectrometer that is equipped for ion mobility. Yet major constructors have launched ion mobility instruments. It appears to us that ion mobility is the future of mass spectrometry in proteomics. As a consequence, the use of CCS values to interpret spectra and validate peptide should be taken into consideration and discussed when preparing the next version of the guidelines.

We focused our study on the prominent stem cell types modeling peri-implantation, but given the technical progresses of mass spectrometry for proteomics in recent years, the analysis of primed PSCs could also lead to the identification of another set of missing proteins.

AUTHORS CONTRIBUTIONS

CP and LD co-coordinated the study. CP conceived and designed the HPP mass spectrometry experiments and analyses. SC processed transcriptomics data. CO produced stem cell lines. OG performed the proteomics sample fractionation and preparation. RL performed MS/MS analyses. RL, EC processed and analyzed the MS/MS datasets. MC and POS performed PaSER analyses. OG and EC performed the bioinformatics analyses. OG, CP, LD and LL performed data/ literature mining of the identified proteins and selected candidates. OG, SC and LD prepared the figures, tables, and supporting Information. OG, CP and LD drafted the manuscript.

Corresponding Author *

Charles Pineau: charles.pineau@inserm.fr, Tel: +33 (0)2 2323 5072

ORCID number: 0000-0002-7461-5433

ACKNOWLEDGMENTS

O. Girard holds a BIRTH GRANT 2021 fellowship. This work was also supported by structural grants from Biogenouest, Infrastructures en Biologie Santé et Agronomie (IBiSA), and the Conseil Régional de Bretagne awarded to C.P. We are grateful to Cecilia Lindskog (Human Protein Atlas, Uppsala, Sweden) for the gift of TRIM61 and ZNF728 antibodies.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ABBREVIATIONS

C-HPP, chromosome-centric human proteome project; HPP, Human Proteome Project; HUPO, Human Proteome Organization; PE1, existence based on evidence at protein level; MPs, missing proteins; PE2, existence based on evidence at transcript level; TE, trophectoderm; ICM, inner cell mass; EPI, pluripotent epiblast; PrE, primitive endoderm; hPSCs, human primed stem cells; hNPSCs, human naive pluripotent stem cells; hiTSCs, human induced naive pluripotent stem cells; hTSCs, human trophoblastic stem cells; FDR, false discovery rate; PSM, peptide spectrum match; CCR, collisional cross section; BSA, Bovine Serum Albumin; 3'SRP, 3'seq-RNA Profiling; UMI, unique molecular identifiers; RPL, Ribosomal Protein Large; RPS, ribosomal protein small.

REFERENCES

- (1) Legrain, P.; Aebersold, R.; Archakov, A.; Bairoch, A.; Bala, K.; Beretta, L.; Bergeron, J.; Borchers, C. H.; Corthals, G. L.; Costello, C. E.; Deutsch, E. W.; Domon, B.; Hancock, W.; He, F.; Hochstrasser, D.; Marko-Varga, G.; Salekdeh, G. H.; Sechi, S.; Snyder, M.; Srivastava, S.; Uhlén, M.; Wu, C. H.; Yamamoto, T.; Paik, Y.-K.; Omenn, G. S. The Human Proteome Project: Current State and Future Direction. *Mol Cell Proteomics* **2011**, *10* (7), M111.009993. <https://doi.org/10.1074/mcp.M111.009993>.
- (2) Adhikari, S.; Nice, E. C.; Deutsch, E. W.; Lane, L.; Omenn, G. S.; Pennington, S. R.; Paik, Y.-K.; Overall, C. M.; Corrales, F. J.; Cristea, I. M.; Van Eyk, J. E.; Uhlén, M.; Lindskog, C.; Chan, D. W.; Bairoch, A.; Waddington, J. C.; Justice, J. L.; LaBaer, J.; Rodriguez, H.; He, F.; Kostrzewa, M.; Ping, P.; Gundry, R. L.; Stewart, P.; Srivastava, S.; Srivastava, S.; Nogueira, F. C. S.; Domont, G. B.; Vandenbrouck, Y.; Lam, M. P. Y.; Wennersten, S.; Vizcaino, J. A.; Wilkins, M.; Schwenk, J. M.; Lundberg, E.; Bandeira, N.; Marko-Varga, G.; Weintraub, S. T.; Pineau, C.; Kusebauch, U.;

- Moritz, R. L.; Ahn, S. B.; Palmblad, M.; Snyder, M. P.; Aebersold, R.; Baker, M. S. A High-Stringency Blueprint of the Human Proteome. *Nat Commun* **2020**, *11* (1), 5301. <https://doi.org/10.1038/s41467-020-19045-9>.
- (3) Lane, L.; Bairoch, A.; Beavis, R. C.; Deutsch, E. W.; Gaudet, P.; Lundberg, E.; Omenn, G. S. Metrics for the Human Proteome Project 2013–2014 and Strategies for Finding Missing Proteins. *J. Proteome Res.* **2014**, *13* (1), 15–20. <https://doi.org/10.1021/pr401144x>.
 - (4) Jumeau, F.; Com, E.; Lane, L.; Duek, P.; Lagarrigue, M.; Lavigne, R.; Guillot, L.; Rondel, K.; Gateau, A.; Melaine, N.; Guével, B.; Sergeant, N.; Mitchell, V.; Pineau, C. Human Spermatozoa as a Model for Detecting Missing Proteins in the Context of the Chromosome-Centric Human Proteome Project. *J. Proteome Res.* **2015**, *14* (9), 3606–3620. <https://doi.org/10.1021/acs.jproteome.5b00170>.
 - (5) Carapito, C.; Duek, P.; Macron, C.; Seffals, M.; Rondel, K.; Delalande, F.; Lindskog, C.; Fréour, T.; Vandenbrouck, Y.; Lane, L.; Pineau, C. Validating Missing Proteins in Human Sperm Cells by Targeted Mass-Spectrometry- and Antibody-Based Methods. *J. Proteome Res.* **2017**, *16* (12), 4340–4351. <https://doi.org/10.1021/acs.jproteome.7b00374>.
 - (6) Vandenbrouck, Y.; Lane, L.; Carapito, C.; Duek, P.; Rondel, K.; Bruley, C.; Macron, C.; Gonzalez de Peredo, A.; Couté, Y.; Chaoui, K.; Com, E.; Gateau, A.; Hesse, A.-M.; Marcellin, M.; Méar, L.; Mouton-Barbosa, E.; Robin, T.; Burlet-Schiltz, O.; Cianferani, S.; Ferro, M.; Fréour, T.; Lindskog, C.; Garin, J.; Pineau, C. Looking for Missing Proteins in the Proteome of Human Spermatozoa: An Update. *J. Proteome Res.* **2016**, *15* (11), 3998–4019. <https://doi.org/10.1021/acs.jproteome.6b00400>.
 - (7) Melaine, N.; Com, E.; Bellaud, P.; Guillot, L.; Lagarrigue, M.; Morrice, N. A.; Guével, B.; Lavigne, R.; Velez de la Calle, J.-F.; Dojahn, J.; Pineau, C. Deciphering the Dark Proteome: Use of the Testis and Characterization of Two Dark Proteins. *J. Proteome Res.* **2018**, *17* (12), 4197–4210. <https://doi.org/10.1021/acs.jproteome.8b00387>.
 - (8) Gerri, C.; McCarthy, A.; Alanis-Lobato, G.; Demtschenko, A.; Bruneau, A.; Loubersac, S.; Fogarty, N. M. E.; Hampshire, D.; Elder, K.; Snell, P.; Christie, L.; David, L.; Van de Velde, H.; Fouladi-Nashta, A. A.; Niakan, K. K. Initiation of a Conserved Trophectoderm Program in Human, Cow and Mouse Embryos. *Nature* **2020**, *587* (7834), 443–447. <https://doi.org/10.1038/s41586-020-2759-x>.
 - (9) Pera, M. F.; Rossant, J. The Exploration of Pluripotency Space: Charting Cell State Transitions in Peri-Implantation Development. *Cell Stem Cell* **2021**, *28* (11), 1896–1906. <https://doi.org/10.1016/j.stem.2021.10.001>.
 - (10) Castel, G.; Meistermann, D.; Bretin, B.; Firmin, J.; Blin, J.; Loubersac, S.; Bruneau, A.; Chevolleau, S.; Kilens, S.; Chariou, C.; Gaignerie, A.; Francheteau, Q.; Kagawa, H.; Charpentier, E.; Flippe, L.; François-Campion, V.; Haider, S.; Dietrich, B.; Knöfler, M.; Arima, T.; Bourdon, J.; Rivron, N.; Masson, D.; Fournier, T.; Okae, H.; Fréour, T.; David, L. Induction of Human Trophoblast Stem Cells from Somatic Cells and Pluripotent Stem Cells. *Cell Rep* **2020**, *33* (8), 108419. <https://doi.org/10.1016/j.celrep.2020.108419>.
 - (11) Deutsch, E. W.; Lane, L.; Overall, C. M.; Bandeira, N.; Baker, M. S.; Pineau, C.; Moritz, R. L.; Corrales, F.; Orchard, S.; Van Eyk, J. E.; Paik, Y.-K.; Weintraub, S. T.; Vandenbrouck, Y.; Omenn, G. S. Human Proteome Project Mass Spectrometry Data Interpretation Guidelines 3.0. *J. Proteome Res.* **2019**, *18* (12), 4108–4116. <https://doi.org/10.1021/acs.jproteome.9b00542>.
 - (12) Xu, Y.; Zhao, J.; Ren, Y.; Wang, X.; Lyu, Y.; Xie, B.; Sun, Y.; Yuan, X.; Liu, H.; Yang, W.; Fu, Y.; Yu, Y.; Liu, Y.; Mu, R.; Li, C.; Xu, J.; Deng, H. Derivation of Totipotent-like Stem Cells with Blastocyst-like Structure Forming Potential. *Cell Res* **2022**. <https://doi.org/10.1038/s41422-022-00668-0>.
 - (13) Schaeffer, M.; Gateau, A.; Teixeira, D.; Michel, P.-A.; Zahn-Zabal, M.; Lane, L. The NeXtProt Peptide Uniqueness Checker: A Tool for the Proteomics Community. *Bioinformatics* **2017**, *33* (21), 3471–3472. <https://doi.org/10.1093/bioinformatics/btx318>.
 - (14) Charpentier, E.; Cornec, M.; Dumont, S.; Meistermann, D.; Bordron, P.; David, L.; Redon, R.; Bonnaud, S.; Bihouée, A. 3' RNA Sequencing for Robust and Low-Cost Gene Expression Profiling; preprint; Protocol Exchange, 2021. <https://doi.org/10.21203/rs.3.pex-1336/v1>.
 - (15) Vizcaino, J. A.; Deutsch, E. W.; Wang, R.; Csordas, A.; Reisinger, F.; Rios, D.; Dianes, J. A.; Sun, Z.; Farrah, T.; Bandeira, N.; Binz, P.-A.; Xenarios, I.; Eisenacher, M.; Mayer, G.; Gatto, L.; Campos, A.; Chalkley, R. J.; Kraus, H.-J.; Albar, J. P.; Martinez-Bartolomé, S.; Apweiler, R.; Omenn, G. S.; Martens, L.; Jones, A. R.; Hermjakob, H. ProteomeXchange Provides Globally Coordinated Proteomics Data Submission and Dissemination. *Nat Biotechnol* **2014**, *32* (3), 223–226. <https://doi.org/10.1038/nbt.2839>.
 - (16) Takashima, Y.; Guo, G.; Loos, R.; Nichols, J.; Ficuz, G.; Krueger, F.; Oxley, D.; Santos, F.; Clarke, J.; Mansfield, W.; Reik, W.; Bertone, P.; Smith, A. Resetting Transcription Factor Control Circuitry

- toward Ground-State Pluripotency in Human. *Cell* **2014**, *158* (6), 1254–1269. <https://doi.org/10.1016/j.cell.2014.08.029>.
- (17) Okae, H.; Toh, H.; Sato, T.; Hiura, H.; Takahashi, S.; Shirane, K.; Kabayama, Y.; Suyama, M.; Sasaki, H.; Arima, T. Derivation of Human Trophoblast Stem Cells. *Cell Stem Cell* **2018**, *22* (1), 50–63.e6. <https://doi.org/10.1016/j.stem.2017.11.004>.
- (18) Kilens, S.; Meistermann, D.; Moreno, D.; Chariou, C.; Gaignerie, A.; Reignier, A.; Lelièvre, Y.; Casanova, M.; Vallot, C.; Nedellec, S.; Flippe, L.; Firmin, J.; Song, J.; Charpentier, E.; Lammers, J.; Donnart, A.; Marec, N.; Deb, W.; Bihouée, A.; Le Caignec, C.; Pecqueur, C.; Redon, R.; Barrière, P.; Bourdon, J.; Pasque, V.; Soumillon, M.; Mikkelsen, T. S.; Rougeulle, C.; Fréour, T.; David, L.; Milieu Intérieur Consortium. Parallel Derivation of Isogenic Human Primed and Naive Induced Pluripotent Stem Cells. *Nat Commun* **2018**, *9* (1), 360. <https://doi.org/10.1038/s41467-017-02107-w>.
- (19) Gaudet, P.; Livstone, M. S.; Lewis, S. E.; Thomas, P. D. Phylogenetic-Based Propagation of Functional Annotations within the Gene Ontology Consortium. *Brief Bioinform* **2011**, *12* (5), 449–462. <https://doi.org/10.1093/bib/bbr042>.
- (20) Chen, H.-H.; Liu, T. Y.-C.; Li, H.; Choo, K.-B. Use of a Common Promoter by Two Juxtaposed and Intronless Mouse Early Embryonic Genes, Rnf33 and Rnf35: Implications in Zygotic Gene Expression. *Genomics* **2002**, *80* (2), 140–143. <https://doi.org/10.1006/geno.2002.6808>.
- (21) Huang, C.-J.; Wu, S.-C.; Choo, K.-B. Transcriptional Modulation of the Pre-Implantation Embryo-Specific Rnf35 Gene by the Y-Box Protein NF-Y/CBF. *Biochem J* **2005**, *387* (Pt 2), 367–375. <https://doi.org/10.1042/BJ20041364>.
- (22) Racki, W. J.; Richter, J. D. CPEB Controls Oocyte Growth and Follicle Development in the Mouse. *Development* **2006**, *133* (22), 4527–4537. <https://doi.org/10.1242/dev.02651>.
- (23) Iouranova, A.; Grun, D.; Rossy, T.; Duc, J.; Coudray, A.; Imbeault, M.; de Tribolet-Hardy, J.; Turelli, P.; Persat, A.; Trono, D. KRAB Zinc Finger Protein ZNF676 Controls the Transcriptional Influence of LTR12-Related Endogenous Retrovirus Sequences. *Mobile DNA* **2022**, *13* (1), 4. <https://doi.org/10.1186/s13100-021-00260-0>.



(Figure created with BioRender.com)

FOR TOC ONLY

Embryonic Stem cell derivation

As my project was developing, I was aware of the blastoid models being available soon and I was interested in setting them up for a follow-up project. To do so, a critical aspect is to have a good, clearly consented, hNPSCs line to generate blastoids. I therefore contributed to the derivation of new human embryonic stem cell (hESCs) lines. The goal was first to directly derive naive hESCs. We also wanted to derive primed hESCs, in order to also have a primed alternative with appropriate consents.

Primed Derivation

For the primed PSCs derivation, I used 4 embryos in 2 experiments (2 embryos per experiments) and the Lerou protocol (Lerou, 2011) (Figure 14)

Briefly, expanded blastocysts were laser dissected to separate EPI from TE. The EPI was then plated on mouse embryonic fibroblasts (MEFs) for culture before being passed when colonies formed. Due to colony collapse after 1 passage for 1 embryo, it was decided to change the culture media from KSR-FGF2 to iPSC Brew medium 7 days after the start of the experiment for the other.

Out of 4 embryos, 2 attached and 1 yielded 1 cell line after transition into iPSC Brew culture medium.

Since the derived hPPSCs have a normal karyotype at P5 (SNP analysis), we have decided to attempt to convert them to hNPSCs and hTSCs.

Naïve derivation

For naive derivation, I used 3 embryos in 2 experiment (1 experiment with 1 embryo and 1 experiment with 2 embryos) following the Strawbridge et al, 2022 protocol (Strawbridge et al., 2022) (Figure 14).

Briefly, expanded blastocysts were laser dissected to separate EPI from TE. The EPI was then dissociated before being plated on MEFs for culture before being passed when colonies formed.

Out of 3 embryos, all yielded attached single cells that grew and formed colonies. I could not pass the colonies and obtain cell lines.

There was cell survival in the early timepoints of the derivation. However, upon passage, all cells were lost. It is worth mentioning that when looking at the literature, success rates in derivation of hPSCs directly from embryos are of about 1/10 (Ström et al., 2010).

Drawing from our experiences with primed cell derivation, it became evident that the choice of culture medium significantly impacted cell survival. Specifically, in the case of primed derivation, switching from KSR-FGF2 to iPSC Brew medium led to a notable increase in cell proliferation and ultimately enabled the derivation of a cell line.

Trophoblast derivation

For trophoblast derivation, I used the remnants of TE from both naive and primed derivations, working with four embryos in a single experiment and adapting the Okae protocol (Okae et al., 2018) (Figure 14).

Briefly, expanded blastocysts were laser dissected to separate EPI from TE. The TE was then dissociated before being plated for culture before being passed when colonies formed for 2 TE. For The 2 other TE, The TE was then plated on MEFs for culture before being passed when colonies formed.

Among the four TE samples, two briefly attached, but the cells subsequently detached within the following eight days, preventing successful passaging.

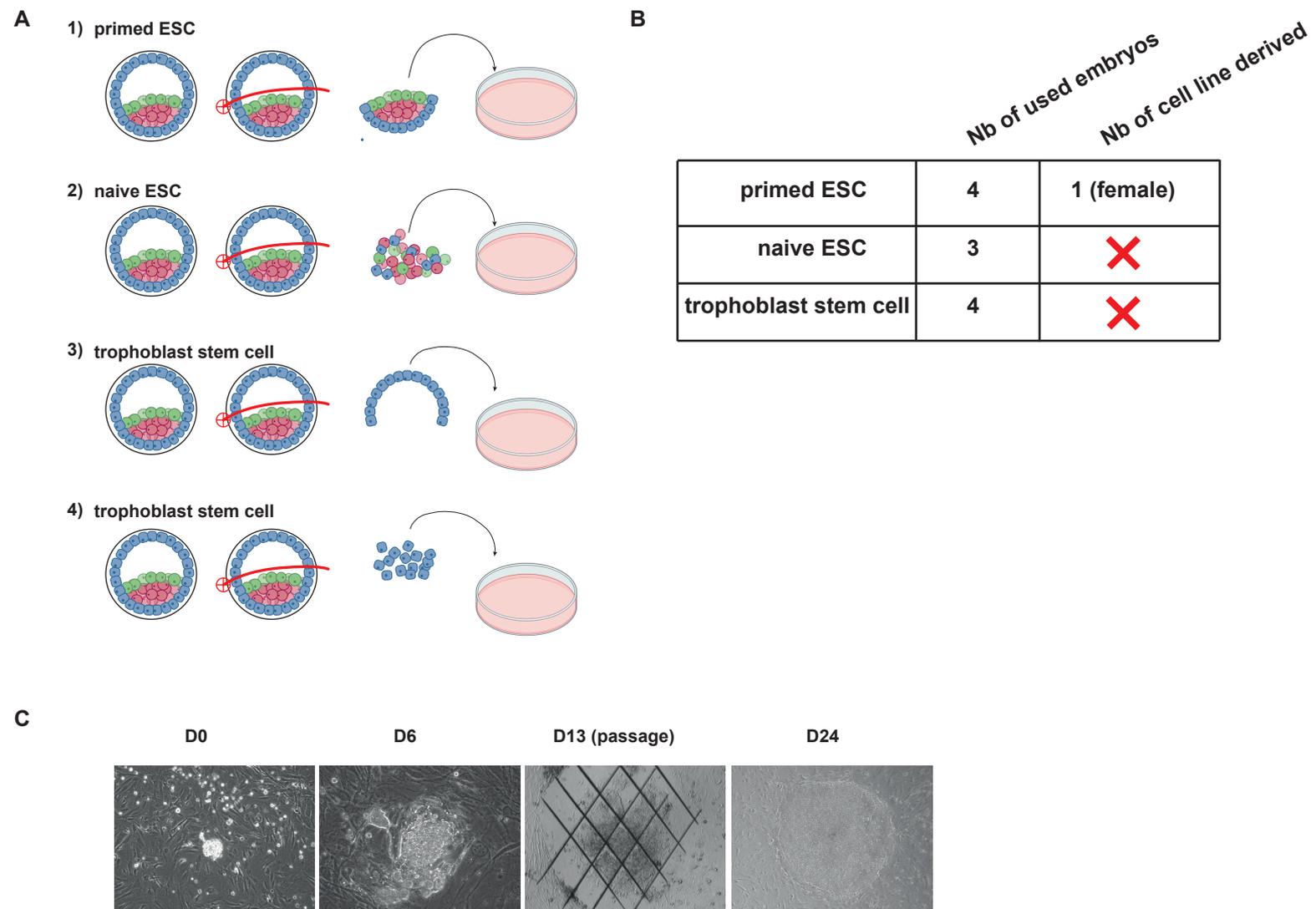


Figure 14 : Derivation of embryonic stem cells

(A) Different protocols used for embryonic stem cell derivation

1- Protocol for human primed pluripotent stem cell derivation (applied on 4 embryos). After recovery of B5 blastocysts, inner cell mass and trophectoderm layers are segregated using laser dissection. The inner cell mass is then transferred in a culture plate and subsequently cultured.

2- Protocol for human naive pluripotent stem cell derivation (applied on 3 embryos). After recovery of B5 blastocysts, inner cell mass and trophectoderm layers are segregated using laser dissection. The inner cell mass is then dissociated to obtain single cells. Cells are then transferred in a culture plate and subsequently cultured.

3- Protocol for human trophoblast stem cell derivation (applied on 2 embryos). After recovery of B5 blastocysts, inner cell mass and trophectoderm layers are segregated using laser dissection. The trophectoderm is then transferred in a culture plate and subsequently cultured.

4- Alternative protocol for human trophoblast stem cell derivation (applied on 2 embryos). After recovery of B5 blastocysts, inner cell mass and trophectoderm layers are segregated using laser dissection. The trophectoderm is then dissociated to obtain single cells. Cells are then transferred in a culture plate and subsequently cultured.

(B) Table of number of embryos used and number of cell lines successfully derived depending on the type of embryonic stem cells

(C) Representative images (bright field) of primed pluripotent stem cell derivation.

Abbreviations : D days; ESC embryonic stem cells; Nb number

Conclusion of the derivation experiment

Given the challenges encountered in the naive derivation process, it is prudent to repeat the derivation, placing a heightened emphasis on the choice of culture medium. Up until this point, attempts have been made to directly convert cells into PXGL culture medium. One potential solution could involve initially deriving cells in T2iLGö culture medium or the combination of T2iLGö and PXGL culture medium, as employed by the Nichols team (Strawbridge et al., 2022).

Furthermore, improvements can be made by preparing our own N2 and B27 supplements. Several research groups have reported that these two components can have a substantial impact on the culture and derivation of naive pluripotent stem cells.

DISCUSSION

My PhD research was centered on characterizing stem cell lines that model pre- and post-implantation development of human embryos.

In my upcoming discussion, I will go back to the specific points raised from my experiments.

Proteomic and transcriptomic data mis alignment

The Central Dogma of molecular biology as stated by Francis Crick in 1957 indicates that once “information has passed into protein it cannot get out again. Information here means the precise determination of sequence, either of bases in the nucleic acid or of amino acid residues in the protein” (Crick, 1958). Watson later encapsulated this idea by stating that genetic information flows only in one direction, from DNA to RNA to protein or directly from RNA to protein (Watson et al., 1965). Considering both versions of the dogma, it is generally assumed that proteins and mRNA levels corresponds. However, this statement is not always true when comparing bulk RNA sequencing and bulk mass spectrometry data (Figure 15). The discrepancy observed between transcriptomic and proteomic landscapes in the cells can be explained by several factors.

First, mass spectrometry remains a relatively recent technology with numerous technical challenges that hinder the comprehensive detection of all proteins. As of today, there are still 1381 “missing” proteins- predicted to exist as yet not detected through mass spectrometry (Girard et al., 2023). The low abundance of some proteins, as well as biology-specific factors like time or stress dependent expression, restriction or cell types or pathophysiological conditions can further impede detection. Additionally, peculiar biochemical properties (hydrophobicity...) can also affect detection capacities.

Some studies have also shown that correlation between mRNA and protein expression can be due to different half-lives and post transcription machinery. Physical properties and the structure of the mRNA have also been shown to impact translational efficiency. Finally, the ribosome density and the cell cycle can also affect mRNA-protein ratio (Haider and Pal, 2013).

Finally, a notable technical limitation of mass spectrometry is that, although it can detect tens of thousands of peptide like features in a single human cells, only a small fraction of them can be assigned to amino acid sequences (Slavov, 2021). This not only affects our conclusion of proteomics data alone but also affects our conclusion when comparing proteomic data with transcriptomic data.

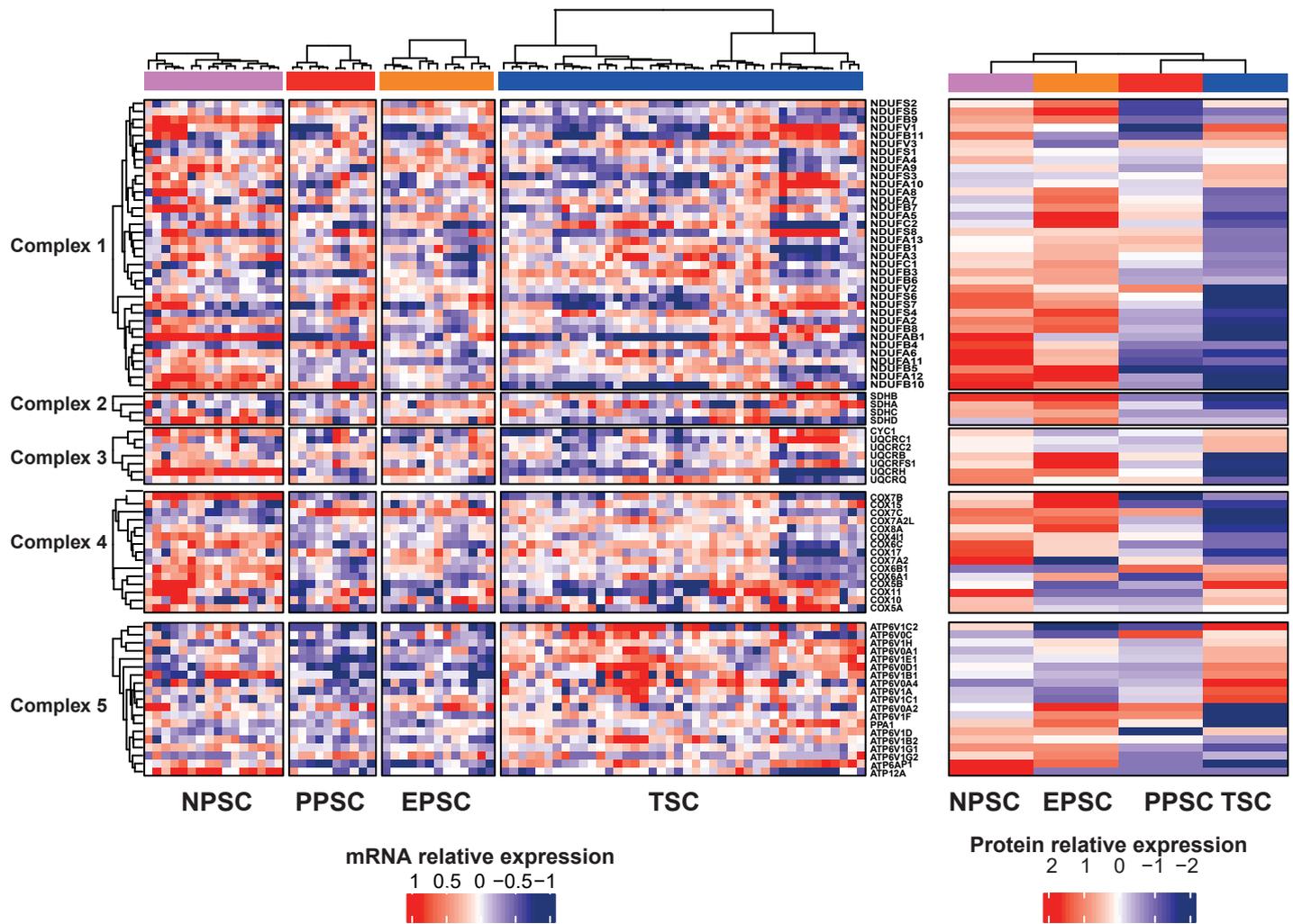


Figure 15 : Proteomic and transcriptomic data misalignment

From Onfray et al, 2023 (Figure 4A)

Heatmap of relative expression of gene (left) (analyzed by DGE-seq) and associated proteins (right) (analyzed by mass spectrometry) of electron transport chains in EPSCs and PPSCs samples. Genes were classified by mitochondrion complex and hierarchically clustered.

Currently, transcriptomic data are more prevalent than proteomic data, however, unlike transcriptome, the proteome offers insights into the dynamic of protein products within our datasets, thereby bringing us closer to understand biological function.

By integrating more proteomic data in our studies, we not only promote the utilization and development of such tools but also provide a powerful platform for multi-OMICS characterization which will serve as a resource to uncover more about the biological significance of our stem cell models.

Culture medium components impact on DNA methylation, X chromosome and metabolic activity

Components of culture media impact different signaling pathways, consequently impacting affect different properties of the cells. Considering these potential effects of the culture medium is crucial in our analysis (Figure 16).

TGF b inhibition

A83-01, a potent inhibitor of activin receptor-like kinase (ALK) including ALK5 (type I transforming growth factor- β receptor), ALK4 (type IB activin receptor), and ALK7 (type I NODAL receptor). TGF β signaling has been shown to induce DNA methylation in epithelial to mesenchymal transition in cancer cells (Cardenas et al., 2014). Direct effect of A83-01 on DNA methylation has not been tested. Based on TGF β studies, we can hypothesize that A83-01 is correlated with DNA hypomethylation.

A83-01 is part of ASECRiAV and ACE culture medium. This correlates DNA methylation results showing that hTSCs cultured in ASECRiAV are hypomethylated compared to hPPSCs.

HDAC inhibition

VPA was shown to cause DNA hypomethylation through demethylation. It was proposed that VPA increase demethylase accessibility to DNA (Detich et al., 2003). VPA has also been shown to reduce glycolysis in neuroblastoma. In terms of mechanism, VPA hinders the expression of E2F1, a transcription factor, leading to the suppression of subsequent glycolytic genes, including glucose-6-phosphate isomerase (GPI) and phosphoglycerate kinase 1 (PGK1) (Fang et al., 2019).

Culture medium component	present in culture medium					effect	
	LCDM	T2iLGö	PXGL	ASECRiAVKSR/FGF2		metabolism	DNA methylation
LIF	✓	✓	✓			Drives glycolysis	
CHIR99021	✓	✓		✓		Drives oxidative phosphorylation	Drives DNA hypomethylation
IWR1	✓						Drives DNA methylation
DIM	✓						
Y27632	✓	✓	✓	✓		Drives oxidative phosphorylation	
XAV939			✓				
PD0325901		✓	✓				Drives DNA hypomethylation
Gö6983		✓	✓				
EGF				✓		Drives glycolysis	Drives DNA methylation
Vitamin C				✓		Reduce glycolysis	Drives DNA hypomethylation
SB431542				✓			
A83-01				✓			Drives DNA hypomethylation ?
Activin A					✓		
VPA				✓		Reduce glycolysis	Drives DNA hypomethylation
FGF					✓		Drives DNA hypomethylation
MIH	✓						

Figure 16 : Potential impact of culture medium components on metabolism and DNA methylation

VPA is part of ASECRiAV culture medium. While hTSCs are hypomethylated compared to cells without VPA in their culture medium, they have a higher metabolic activity, including higher glycolytic activity than cells without VPA.

MEK inhibition

MEK inhibition causes dose dependent impairment of maintenance of methylation. Culture in MEK inhibitor and LIF triggered extensive hypomethylation in mouse ESCs, more pronounced than in the classical 2i media (Spindel et al., 2021). It was proposed in the same article that both MEK and GSK3 β inhibition inhibit de novo methylation and that this effect is seen from 0,4 μ M and above of MEK inhibitor concentration. It is worth noting that reduced MEK inhibitor concentration yields more proliferating and more stable hNPSCs in 5iLAF culture medium (Di Stefano et al., 2018).

MEK inhibitor PD0325901 is part of T2iLGö, PXGL and 5iLAF culture medium. hNPSCs hypomethylation correlates with PD0325901 presence in their culture medium.

Wnt inhibition

CHIR99021 is an inhibitor of GSK3 β . GSK3 β is a serine/threonine kinase that is a key inhibitor of the WNT pathway; therefore, CHIR99021 functions as a WNT activator.

CHIR99021 has been shown to drive DNA hypomethylation (in particular when combined with MEK inhibitor such as PD0325901). GSK3 inhibitors also diminishes XIST coating in PPSCs, thus activating X erosion (Cloutier et al., 2022). CHIR99021 has also been shown to drive oxidative phosphorylation (Ma et al., 2019).

While T2iLGö, ASECRiAV, ACE and LCDM culture media contains CHIR99021, PXGL culture medium does not. We did not see differences between T2iLGö and PXGL hNPSCs levels of expression of XIST. However, one hEPSCs line was completely eroded. This cell

line was converted from hPPSCs (H9). One hypothesis could be that the hPPSCs used were eroded or that the conversion treatment triggered the X erosion.

Rock inhibition

Y27632 is a selective ROCK inhibitor which has been shown to increase oxidative phosphorylation in corneal endothelial cells (Ho et al., 2022). ROCK inhibition leads to enhanced mitochondrial respiration and overexpression of electron transport chain components through upregulation of AMP-activated protein kinase pathway.

Y27632 is a component of LCDM, PXGL, T2iLGö and ASECRiAV culture media. This correlates with the fact that cells cultured in LCDM, PXGL, T2iLGö and ASECRiAV have an overall higher metabolic activity than hPPSC (cultured in culture media without ROCK inhibitor).

Vitamin C

L-ascorbic acid, a cofactor of TET enzymes, promotes DNA demethylation through 5hmC formation (Brabson et al., 2021).

L-ascorbic acid has also been shown to reduce glycolysis in cancer (Vuyyuri et al., 2013).

L-ascorbic acid is present in ASECRiAV and mTeSR culture media. This could affect TSCs hypomethylation. However, it is worth noting that TET gene levels are the same between ASECRiAV and ACE TSCs.

EGF

EGF has been shown to increase DNMT activity and DNA methylation in cancer cells (Samudio-Ruiz and Hudson, 2012). EGF has also been shown to promote glycolysis in

triple negative breast cancer cells, through PKM2 phosphorylation, which increases pyruvate formation (Lim et al., 2016).

EGF is part of ASECriAV and ACE culture medium. EGF presence in the culture medium does not correlate with increased DNA methylation in hTSCs as they are hypomethylated compared to hPPSCs.

LIF

LIF/STAT3 signaling has been shown to induce hypomethylation via metabolic reconfiguration through decreased production of alpha ketoglutarate from glutamine, which increases DNMT3A/B expression and DNA methylation in mouse ESCs (Betto et al., 2021). LIF has also been found to drive glycolysis in cancer (Yue et al., 2022).

LIF is a component of T2iLGö, PXGL and LCDM culture media. It is worth mentioning that cells cultured in T2iLGö, PXGL and LCDM in our study show the same level of DNMT3A expression, while showcasing distinct DNA methylation levels.

Considering all this, it is quite surprising to see that hEPSCs share the same methylation levels as hPPSCs (Onfray et al., 2023). Indeed, LCDM medium is composed of CHIR99021 and LIF, both associated with DNA hypomethylation.

We did not find any significant differences in hNPSCs culture in PXGL or T2iLGö culture medium. However, it was reported in the mouse embryo that different culture medium affects DNA methylation (Uysal et al., 2022).

TSCs have a very high metabolic activity with both a high oxidative phosphorylation and a high glycolytic ability. However, when looking at their culture media components, L-ascorbic acid and VPA presence could reduce glycolysis in TSCs. It would be interesting to determine whether glycolysis in TSCs happens through the same pathways as in cancer cells. One other explanation is that the increased pH we measure in our experiment is not a result of lactate production through glycolysis but through normal oxygenation of cells. It is worth mentioning that all results have been normalized over

100 000 cells and that controls have been performed (2DG, CCCP, Oligomycin, Rotenone, Antimycin). It is also worth mentioning that during our assessment of the metabolic activity of our stem cell models (Onfray et al., 2023), we maintained the cells in a minimum media, with the same amount of glucose, pyruvate and glutamine, without small molecules and inhibitors for 4 hours before the experiment. Considering that the metabolite flux in mammalian cells is of 0.01mM/s, those few hours should be sufficient for any media-induced metabolism to be evened out, thus leaving us to only measure cells intrinsic metabolic potential. However, some studies have tried to cultivate naive or pluripotent cells for a few days in basal medium without any changes in gene expression (Rostovskaya et al., 2019). This leads to question what is a good control for metabolic activity assessment and how can we distinguish intrinsic metabolic activity and culture induced metabolic activity.

No functional tests have been made yet comparing ACE hTSCs to ASECRiAV hTSCs but, taking into account culture medium composition, one could hypothesize that they should not differ in metabolic activity from ASECRiAV hTSCs but they could have more DNA methylation than ASECRiAV hTSCs given the absence of several small molecules linked with DNA hypomethylation.

Overall, analysis of culture medium composition can help us to form hypothesis over the characteristics of stem cells cultured in different conditions. However, culture media is not the only player. Intrinsic properties of cells and cell-cell dialog also impact deeply their properties, thus rendering the conclusions difficult. This highlights the importance of functional analysis as made in Onfray et al. Functional test are the only way to validate or invalidate our hypothesis made on the base of the state of cells and their culture media. This also highlights that the basal conditions our cells are cultivated in can affect the results of experiments and should be reported to avoid confusion.

From the model to the embryo, are stem cell models well depicting embryonic development?

In Onfray et al, we characterized in parallel hNPSCs, hPPSCs, hEPSCs and hTSCs on the transcriptomic, epigenetic and metabolic level. The set of hallmarks we used enabled a clearer characterization of hTSCs and hEPSCs along with hNPSCs and hPPSCs. Detailed hallmarks also instruct on the relevance of each model to human development.

Only few studies have so far addressed the timing of X Chromosome Inactivation (XCI) in human embryos. It was described that in human embryos, pre inactive X chromosomes accumulate both XIST and XACT long non coding RNAs on both X chromosomes. Around day 6, an expression imbalance arises between the 2 X chromosomes, which can be interpreted as the initiation of XCI (Vallot et al., 2017). However, it was shown by transcriptomic analysis that upregulation and inactivation of the X chromosome is not fully completed at day 12 in human female embryos (Zhou et al., 2019). It is not known yet when X chromosome inactivation is complete in human and analysis at the single cell level to detect spatial arrangements of transcripts with technics such as RNA FISH is still missing in human post implantation embryos.

The lineage-specific DNA methylation dynamics around implantation remain largely unknown in human embryos. In general, it was demonstrated that EPI, PE and TE experienced strong genome re-methylation during implantation, however asynchronous (Zhou et al., 2019).

The metabolic activity of the human embryo has been measured (Brinster, 1973) and overall, the embryo goes from a relatively inactive metabolic tissue at ovulation to a rapidly metabolizing tissue at implantation. However, the metabolic activity of each lineage at the peri implantation period is not known yet and post implantation metabolic activity has not been explored either.

Overall, due to technical and ethical limitations, few hallmarks used to characterize peri-implantation stem cell models have been studied in the peri implantation human embryo.

Stem-cell-based embryo models generate much excitement as they offer a window into an early phase of human development that has remained largely inaccessible to scientific investigation.

However, it is important to keep in mind that the epigenetic status of stem cell models is inextricably linked to that of the stem cell from which they originate. For example, it has been demonstrated *in vitro* that models composed from cells with established XCI will likely maintain XCI during *in vitro* differentiation (Patel et al., 2017). Additionally, some important differences exist between stem cell models and the human embryo. Indeed, biallelic coating of XIST in the entire cell population is lacking in hNPSCs, contrary to what is seen in the human embryo. On top of that, unlike what happens in human development, XCI is biased after hNPSCs differentiation. Besides, while both hNPSCs and preimplantation epiblast cells share a genome wide hypomethylation status, hNPSCs were shown to lose methylation at imprinted differentially methylated regions compared to human pre implantation epiblast and do not regain methylation at these loci upon re-priming (Pastor et al., 2016; Theunissen et al., 2016).

Although we should celebrate the technical advances represented by the plethora of models available to study human peri implantation development, such advances must also give us pause for thought. An independent and precise effort of model validation and description, using several readouts commonly used to measure fate and state is necessary before being able to use all these models to study human development. Indeed, as showed in Onfray et al, hallmarks of pluripotency are not predictive of each other and the sole use of 1 hallmark is not sufficient to determine precisely fate and state of the cells. We provided an important resource in Onfray et al by comparing 2D models of peri implantation development on the transcriptomic, proteomic, epigenetic and metabolic level. This must be applied to all models, even 3D models, such as blastoids (Fan et al., 2021; Kagawa et al., 2021; Liu et al., 2021c; Sozen et al., 2021; Yanagida et al., 2021), trophoblastic organoids (Karvas et al., 2022; Turco et al., 2018) or integrated models of human development (e.g., post implantation models integrated with endometrial cells (Cai et al., 2023)).

Finally, extensive embryo studies are necessary to establish a true and clear reference of what happens during human peri implantation development. As summarized by George E.P.Box, “All models are wrong but some are useful” and a better comprehension of the models available and of the initial object of study will help to better understand which model is better suited to answer specific questions.

Reprogrammed, converted, derived: all the same results

Analysis of converted, reprogrammed and derived hTSCs conducted in parallel reveals no significant differences between the different methods of hTSCs generation. All 3 methods of generation yielded cells with the potential for both ST and EVT differentiation along with the expression of the same markers of trophoblastic fate: GATA3, KRT7, VRGLL1 (Castel et al., 2020). Moreover, all 3 methods of generation produced cells with similar level of DNA methylation, one inactive X chromosome (in the case of female lines) and a high metabolic activity (Onfray et al., 2023). It is important to note, however, that while we compared the 3 different generation methods, all hTSCs were cultivated in the same condition: ASECRiAV medium, on MEFs feeder layer. Converted and reprogrammed hTSCs were generated in house, while derived hTSCs were generated in Arima's lab (Okae et al., 2018).

Other culture medium for hTSCs have been described such as the ACE culture medium (Io et al., 2021). Notable, this culture medium lacks Vitamin C and MEK inhibitor when compared to ASECRiAV culture medium (Okae et al., 2018). Cells cultured in ACE culture medium correlate on the transcriptomic level with hTSCs cultured in ASECRiAV medium (Onfray et al., 2023). Cells in ACE and ASECRiAV medium express DNMT1, DNMT3A, DNMT3B and DNMT3L at similar levels, suggesting that hTSCs cultured in ACE and ASECRiAV medium may exhibit comparable methylation levels although this has not been tested yet.

It would be of interest to test whether conversion of hNPSCs into hTSCs is easier and quicker in ACE medium than in ASECRiAV.

Recently, hTSCs were also generated through conversion of hPPSCs. These cells are cultured in ASECRiAV culture medium and have been compared to derived hTSCs in most study (Figure 12). However, their characterization primarily relies on transcriptomic data for now.

Work conducted in Onfray et al already provides a new reference for trophoblast models regarding their hallmarks. It would be of great interest for the field to have a side-by-side

comparison of all hTSCs converted from hPPSCs to other types of hTSCs. Such a comparison would clarify whether hTSCs converted from hPPSCs exhibit the same features as other hTSCs.

Evaluating pluripotency state

When comparing hallmarks characterizing naive pluripotency in mouse to those characterizing naive pluripotency in human, several distinctions become apparent between species.

First, in mice, the gold standard to assess naive pluripotency is the ability to extensively contribute to interspecies chimera. In contrast, in human, NPSCs can contribute to interspecies chimeras but the success rate varies between studies and is overall very limited. For example, the contribution of human NPSCs to porcine embryos corresponds to less than one human cell in 10000. In mouse embryos, it never exceeds 5%. It has been described that human NPSCs (as well as non-human primate NPSCs) do not stay mitotically active during embryo colonization and thus differentiate prematurely (Aksoy et al., 2021). Notably, both mouse and human EPSCs can contribute to interspecies chimeras, albeit at different levels even though they are transcriptionally associated to primed pluripotency (Onfray et al., 2023; Posfai et al., 2021). This raises important concerns considering chimera formation as a hallmark of naive pluripotency and highlight the need to distinguish between the survival ability and the developmental fate and state.

Second, mouse NPSCs lack extra-embryonic potential while human NPSCs can convert into trophoblast (Guo et al., 2021; Io et al., 2021). Furthermore, human NPSCs, unlike their mouse counterparts, can also convert into hTSCs (Castel et al., 2020; Cinkornpumin et al., 2020; Dong et al., 2020). Of note, both mouse and human EPSCs exhibit extended developmental potency toward extra embryonic lineages (Castel et al., 2020; Gao et al., 2019). While mouse EPSCs extra embryonic potential has been validated through invasion of trophoblast layer in chimera, it has not been validated through conversion into TSCs though, unlike in human.

These differences raise questions about what defines naive pluripotency across species.

Finally, one significant recent advance in human peri implantation field has been the development of blastoids (Fan et al., 2021; Kagawa et al., 2021; Liu et al., 2021c; Sozen et al., 2021; Yanagida et al., 2021; Yu et al., 2021). These 3D aggregates of NPSCs model

pre implantation B4-B6 blastocysts and are composed of the 3 lineages of human pre implantation embryo: epiblast, trophoblast and primitive endoderm. Due to this multi lineage formation ability, it was proposed lately that the ability to make blastoids is a hallmark of human naïve pluripotency (Zhou et al., 2023). It is important to note that mouse NPSCs cannot generate blastoids on their own (Rivron, 2018).

Given these discrepancies, it's essential to reconsider the relevance of the hallmarks used to evaluate naive pluripotency and whether the same criteria should apply across different species.

An important consideration to answer this question is that analysis of pluripotent stem cell models in parallel over transcriptomic, proteomic, epigenetic and metabolic hallmarks revealed that hallmarks are not predictive of each other and that the use of only one hallmark is not sufficient to predict the state of the cells (Onfray et al., 2023) (Figure 17). As such, it is interesting to note that X chromosome activity seems to be one of the most stringent criteria to assess naïve pluripotency in human. However, this criterion can only be applied in female lines.

To address whether we should be using the same hallmarks across different species, it is crucial to examine the differences in culture systems used in NPSCs of different species. Mouse NPSCs are maintained in 2i culture medium, while human NPSCs are maintained in T2iLGö, PXGL, 5iLAF or eNHSM culture medium. In the case of rabbit pluripotency, VALGoX medium is used. In cynomolgus, 4CL media allowed generation of implantation-competent blastoids (Li et al., 2023). All these different culture mediums use different inhibitor and small molecules to maintain cells in a state that is considered to be comparable. However, as discussed earlier, culture medium can significantly impact different characteristics such as DNA methylation, metabolic activity or X chromosome inactivation, which can, in turn, have an impact in differentiation potential. Therefore, there are two effects that are linked: state of the cells and impact of environment on the cellular processes, that might not be linked to the fate state. The problem is that maintaining naive stem cells from different species under the same culture conditions is technically impossible in the state of the art, making it difficult to differentiate between fate and cellular fitness. To address this, extensive comparisons would be needed between each NPSCs

Trophoblast conversion potential
(without priming)

Blastoid formation potential

Interspecies chimerism potential

DNA hypomethylation

X chromosome activation

Metabolic activity

Developmental stage modelled

 hNPSC	✓	✓	✓	✓	✓	++	Pre-implantation epiblast
 hPPSC	✗	✗	✗	✗	✗	—	Post-implantation epiblast
 hEPSC	✓	✓	✓	✗	✗	+	Post-implantation epiblast
 hTSC				✓	✗	+++	Post-implantation trophoblast

Figure 17 : Hallmarks used to evaluate naive pluripotency in human are not predictive of each others.

Red cross means that the cells do not have this hallmark

Green tick means that the cells do have this hallmarks

Abbreviations : hNPSC human naive pluripotent stem cells; hPPSC human primed pluripotent stem cell;

hEPSC human extended pluripotent stem cells; hTSC human trophoblast stem cells

type in each culture medium available. Moreover, resetting protocols have been limiting: it might be easier to transit naive cells in PXGL to the other media used in other species than directly reset primed cells in those media.

Altogether, this underscores the need for a deeper understanding of basic cellular functions in all these culture media to decipher the contribution of fate versus the environmental contribution (including metabolism, cell cycle regulation...).

PERSPECTIVES

During this PhD, I compared hTSCs, hEPSCs, hPPSCs and hNPSCs cell lines with epigenetic and functional readouts that have been previously used as hallmarks of hNPSCs. My results clarify the fact that hallmarks of pluripotency are not predictive of each other and that multiplying hallmarks alleviates stage matching biases.

Readouts unambiguously associated the hEPSCs with a primed pluripotent fate (Onfray et al., 2023). Nevertheless, hEPSCs clearly have a higher clonogenic propensity and growth rate than hPPSCs. We showed that hEPSCs have a metabolic activity comparable to hNPSCs, which raises a paradox in the association of hallmarks such as chimerism and trophoblastic conversion with the fate of the cells. This questions if there is a link between fate conversion and metabolism.

On a broader scale, my work in defining hTSCs and pluripotent stem cell models' metabolic activity leads to questions considering human embryo metabolic activity. The metabolic status of each lineage within the human peri-implantation embryo is not known yet. A better knowledge of human embryo metabolism could help to determine which embryo has the best implantation potential.

Finally, all this work will help to reconsider the importance of basic components of human embryo culture medium and their impact on embryonic development and implantation potential. This could help for the formulation of new embryo culture medium, which could also in turn impact ART success rates.

In this section, I will explain potential experimental plan for these perspectives.

What is the link between fate and metabolism?

It was proposed that hEPSCs represent a distinct state of pluripotency to hPPSCs due to their ability (even though limited) to form inter species chimeras (Liu et al., 2021b; Tan et al., 2021b; Yang et al., 2017c; Zheng et al., 2021b) and to convert to hTSCs without pre-treatments (Castel et al., 2020; Dong et al., 2020). We showed in Onfray et al that hPPSCs and hEPSCs are similar on the epigenetic, transcriptomic and proteomic level, among pluripotent stem cells. However, hEPSCs have a higher metabolic activity than hPPSCs.

It would be interesting to see if perturbation of metabolic activity could impact conversion potential ability. To do so, we could set up experiments where we disrupt glycolysis or oxidative phosphorylation in hEPSCs and then assess their hTSCs conversion potential. The enhanced clonogenicity and growth rate of hEPSCs offer new opportunities to understand the link between metabolic activity and pluripotency, and the link between metabolic activity and fate transition.

Then, by increasing hPPSCs metabolic activity, we could monitor whether they can convert into hTSCs without pretreatment and more efficiently. It would also be interesting to monitor metabolic activity of BMP4 and BAP treated hPPSCs as these treatments induce hTSCs conversion from hPPSCs.

Questions concerning metabolism and fate conversion potential in 2D models also bounce back toward questions over embryo metabolism. Indeed, if metabolism can affect fate conversion potential and chimera potential, could metabolism also affect embryonic survival? Is a good metabolic activity sufficient for the embryo to thrive?

While 2D models can help to unravel the role of metabolic activity in fate conversion, more studies in pre-implantation embryos are necessary before being able to answer whether metabolic activity could have an impact on embryonic survival.

Can embryo metabolism serve as a readout to determine its implantation potential?

Mouse embryo's metabolism has been shown to impact their implantation potential. Mouse blastocysts have a high level of glycolysis, even though they possess functional mitochondria and exhibits the highest oxygen consumption rates of the preimplantation stage (Houghton et al., 1996). In human blastocysts, more than 50% of the glucose consumed is not oxidized but is converted to lactate (Gardner and Leese, 1987; Gott et al., 1990). It was showed that at implantation, embryos are in a close-to-anoxic environment (1,5 to 5,3% of oxygen in uterus lumen at implantation time and absence of vasculature at implantation site) (Fischer and Bavister, 1993), thus promoting glycolysis and lactate formation at the moment of implantation. As development proceeds, the implanting blastocyst has a glycolysis-dependent metabolism, in which 90% of the consumed glucose forms lactate (Clough and Whittingham, 1983). It was hypothesized by Gardner in 2015 that this specific metabolism was helping the blastocyst at the time of implantation. It was proposed that high level of lactate, an end-product of aerobic glycolysis, produced by the blastocyst would acidify the microenvironment at the time of implantation (up to 100 mM lactic acid produced by human embryos over 24 hours around the time of implantation, while 1-3 mM lactate in blood and in resting tissues) (Gardner, 2015; Hunt et al., 2007). It was hypothesized that this low pH, high lactate environment allowed endometrial breakdown, angiogenesis and immunoregulation, thus facilitating implantation (Gardner, 2015). Of note, it was shown in mouse that acidification of uterine tissues is essential for implantation. Indeed, prevention of uterine acidification using baliformycin A1 (inhibits V-ATPase, which regulates the acidity of the environment) results in a dose dependent disruption of implantation (Xiao et al., 2017). pH around the human blastocyst has not been investigated yet. Given the calculated levels of lactate produced by the blastocyst during implantation, further reductions in the pH in the immediate vicinity of the blastocyst are predicted.

Further studies are needed in human to determine if and how metabolic activity can affect implantation

To investigate if metabolic activity could be a marker to determine implantation potential, we could measure metabolites such as lactate levels in IVF embryo culture medium before transfer and verify if high lactate levels in the culture medium corresponds to better chances of implantation.

To go further, we could study how trophoblast cells affect their microenvironment by using trophoblastic organoids or post implantation blastoids on endometrial cells and by measuring microenvironment perturbation through mass spectrometry for example, to determine the impact of metabolic activity in post implantation development. These studies will then have to be validated by experiments in post implantation embryos.

Overall, investigating how metabolic activity can impact implantation potential or development leads to questions concerning embryo culture medium basal component, as some components can have an impact on metabolic activity.

Does culture media have an impact on developmental and implantation potential?

IVF embryos are cultured for 5 to 6 days in IVF culture medium before transfer. Optimizing embryo culture condition could impact their developmental and implantation potential.

First, it would be interesting to address IVF culture media composition. Nowadays in clinic, IVF culture medium are commercially bought. Several companies exist but they do not disclose the culture medium composition. This has already been explored by mass spectrometry analysis but many components remain unknown (Tarahomi et al., 2019). We could explore this further through mass spectrometry analysis, using a new technology enabling a better characterization of protein (Data Independent Analysis) to better determine protein composition of culture media. This would serve as a reference for further studies aimed at improving IVF culture media composition.

However, while interesting, exact culture media composition is not necessary to undertake studies aimed to improve embryo development and implantation rates.

To better understand what drives proper development, we could set up studies to modulate media, through supplementation and activation or inhibition of specific signaling pathways and assess embryo development. To perform this in a high throughput manner, we could use in a first approach the blastoid model (Fan et al., 2021; Kagawa et al., 2021; Liu et al., 2021c; Sozen et al., 2021; Yanagida et al., 2021; Yu et al., 2021). New models of integrated blastoids with endometrial cells (Cai et al., 2023) could also help to model implantation in a high throughput manner. By investigating which signaling pathway increases blastoid formation and implantation rates, we could then apply interesting signaling pathway in human embryos and validate the findings.

PERSONNAL OUTLOOK OF THE PHD

During the 3 years of my PhD, I developed experimental skills, including stem cell culture, molecular biology, metabolism experiment. Additionally, I developed crucial scientific skill such as data interpretation, formulation and execution of an experimental plan, scientific writing and presentation of my work. My stem cell skills offered me the opportunities to collaborate with C. Rougeulle lab (X Chr activity in NPSC and embryos), Charles Pineau lab (proteomics analysis of stem cell lines) and the hiPSC core facility (for ESC lines derivation). I was able to use technics going from stem cell culture to Seahorse experiment, and RNA FISH. I found great satisfaction in sharing my knowledge and mentoring new students. I had the opportunity to present my work in SY-stem conference in Vienna, in FSSCR in Strasbourg, and in “Spotlight on Stem cells” in Nantes, where I received the best oral presentation award. I was also able to join popularization of science events such as “My PhD in 180 seconds” (jury award), or events in high schools in front of students (“Declics”, “comptoir des sciences”).

I am profoundly grateful for all the opportunities this PhD gave me.

BIBLIOGRAPHY

Aksoy, I., Rognard, C., Moulin, A., Marcy, G., Masfaraud, E., Wianny, F., Cortay, V., Bellemin-Ménard, A., Doerflinger, N., Dirheimer, M., et al. (2021). Apoptosis, G1 Phase Stall, and Premature Differentiation Account for Low Chimeric Competence of Human and Rhesus Monkey Naive Pluripotent Stem Cells. *Stem Cell Rep.* *16*, 56–74. <https://doi.org/10.1016/j.stemcr.2020.12.004>.

Al-Riyami, N., Al-Hadabi, R., Al-Dughaiishi, T., and Al-Riyami, M. (2013). Placental Tumour. *Sultan Qaboos Univ. Med. J.* *13*, E459–E462. .

Amita, M., Adachi, K., Alexenko, A.P., Sinha, S., Schust, D.J., Schulz, L.C., Roberts, R.M., and Ezashi, T. (2013). Complete and unidirectional conversion of human embryonic stem cells to trophoblast by BMP4. *Proc. Natl. Acad. Sci. U. S. A.* *110*, E1212–E1221. <https://doi.org/10.1073/pnas.1303094110>.

An, L., Liu, Y., Li, M., Liu, Z., Wang, Z., Dai, Y., Presicce, G.A., and Du, F. (2021). Site specificity of blastocyst hatching significantly influences pregnancy outcomes in mice. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* *35*, e21812. <https://doi.org/10.1096/fj.202100653R>.

AOKI, F. (2022). Zygotic gene activation in mice: profile and regulation. *J. Reprod. Dev.* *68*, 79–84. <https://doi.org/10.1262/jrd.2021-129>.

Balaton, B.P., and Pasque, V. (2022). Human 8-cell-like cells discovered. *Cell Stem Cell* *29*, 347–348. <https://doi.org/10.1016/j.stem.2022.01.015>.

Bayerl, J., Ayyash, M., Shani, T., Manor, Y.S., Gafni, O., Massarwa, R., Kalma, Y., Aguilera-Castrejon, A., Zerbib, M., Amir, H., et al. (2021). Principles of signaling pathway modulation for enhancing human naive pluripotency induction. *Cell Stem Cell* *28*, 1549-1565.e12. <https://doi.org/10.1016/j.stem.2021.04.001>.

Beattie, G.M., Lopez, A.D., Bucay, N., Hinton, A., Firpo, M.T., King, C.C., and Hayek, A. (2005). Activin A maintains pluripotency of human embryonic stem cells in the absence of feeder layers. *Stem Cells Dayt. Ohio* *23*, 489–495. <https://doi.org/10.1634/stemcells.2004-0279>.

Becker, A.J., McCULLOCH, E.A., and Till, J.E. (1963). Cytological Demonstration of the Clonal Nature of Spleen Colonies Derived from Transplanted Mouse Marrow Cells. *Nature* *197*, 452–454. <https://doi.org/10.1038/197452a0>.

Bernardo, A.S., Faial, T., Gardner, L., Niakan, K.K., Ortmann, D., Senner, C.E., Callery, E.M., Trotter, M.W., Hemberger, M., Smith, J.C., et al. (2011). BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages. *Cell Stem Cell* *9*, 144–155. <https://doi.org/10.1016/j.stem.2011.06.015>.

Betto, R.M., Diamante, L., Perrera, V., Audano, M., Rapelli, S., Lauria, A., Incarnato, D., Arboit, M., Pedretti, S., Rigoni, G., et al. (2021). Metabolic control of DNA methylation in naive pluripotent cells. *Nat. Genet.* *53*, 215–229. <https://doi.org/10.1038/s41588-020-00770-2>.

Blakeley, P., Fogarty, N.M.E., Snell, P., Christie, L., Robson, P., and Niakan, K.K. (2015). Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. 34. .

Blerkom, J.V., and Manes, C. Development of Preimplantation Rabbit Embryos in Vivo and in Vitro.

Boveri, T. (1892). Befruchtung ... (J.F. Bergmann).

Brabson, J.P., Leesang, T., Mohammad, S., and Cimmino, L. (2021). Epigenetic Regulation of Genomic Stability by Vitamin C. *Front. Genet.* 12, 675780. <https://doi.org/10.3389/fgene.2021.675780>.

Bradley, A., Evans, M., Kaufman, M.H., and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* 309, 255–256. <https://doi.org/10.1038/309255a0>.

Braude, P., Bolton, V., and Moore, S. (1988). Human gene expression first occurs between the four- and eight-cell stages of preimplantation development. *Nature* 332, 459–461. <https://doi.org/10.1038/332459a0>.

Bredenkamp, N., Yang, J., Clarke, J., Stirparo, G.G., von Meyenn, F., Dietmann, S., Baker, D., Drummond, R., Ren, Y., Li, D., et al. (2019). Wnt Inhibition Facilitates RNA-Mediated Reprogramming of Human Somatic Cells to Naive Pluripotency. *Stem Cell Rep.* 13, 1083–1098. <https://doi.org/10.1016/j.stemcr.2019.10.009>.

Briggs, R., and King, T.J. (1952). Transplantation of living nuclei from blastula cells into enucleated frogs' eggs *. *Proc. Natl. Acad. Sci.* 38, 455–463. <https://doi.org/10.1073/pnas.38.5.455>.

Brinster, R.L. (1963). A Method for in vitro cultivation of mouse ova from two-cell to blastocyst. *Exp. Cell Res.* 32, 205–208. [https://doi.org/10.1016/0014-4827\(63\)90093-4](https://doi.org/10.1016/0014-4827(63)90093-4).

Brons, I.G.M., Smithers, L.E., Trotter, M.W.B., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., et al. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 448, 191–195. <https://doi.org/10.1038/nature05950>.

Buehr, M., Meek, S., Blair, K., Yang, J., Ure, J., Silva, J., McLay, R., Hall, J., Ying, Q.-L., and Smith, A. (2008). Capture of Authentic Embryonic Stem Cells from Rat Blastocysts. 1287–1298. .

Burton, G.J., Fowden, A.L., and Thornburg, K.L. (2016). Placental Origins of Chronic Disease. *Physiol. Rev.* 96, 1509–1565. <https://doi.org/10.1152/physrev.00029.2015>.

Cai, Y., Li, N., and Li, H. (2023). Combining Endometrial Assembloids and Blastoids to Delineate the Molecular Roadmap of Implantation. *Stem Cell Rev. Rep.* 19, 1268–1282. <https://doi.org/10.1007/s12015-023-10527-z>.

Campbell, K.H., McWhir, J., Ritchie, W.A., and Wilmut, I. (1996). Sheep cloned by nuclear transfer from a cultured cell line. *Nature* 380, 64–66. <https://doi.org/10.1038/380064a0>.

Cardenas, H., Vieth, E., Lee, J., Segar, M., Liu, Y., Nephew, K.P., and Matei, D. (2014). TGF- β induces global changes in DNA methylation during the epithelial-to-mesenchymal transition in ovarian cancer cells. *Epigenetics* 9, 1461–1472. <https://doi.org/10.4161/15592294.2014.971608>.

- Castel, G., Meistermann, D., Bretin, B., Firmin, J., Blin, J., Loubersac, S., Bruneau, A., Chevolleau, S., Kilens, S., Chariou, C., et al. (2020). Induction of Human Trophoblast Stem Cells from Somatic Cells and Pluripotent Stem Cells. *Cell Rep.* *33*, 108419. <https://doi.org/10.1016/j.celrep.2020.108419>.
- Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. *Nature* *450*, 1230–1234. <https://doi.org/10.1038/nature06403>.
- Chang, C.-W., and Parast, M.M. (2017). Human trophoblast stem cells: Real or not real? *Placenta* *60 Suppl 1*, S57–S60. <https://doi.org/10.1016/j.placenta.2017.01.003>.
- Cinkornpumin, J.K., Kwon, S.Y., Guo, Y., Hossain, I., Sirois, J., Russett, C.S., Tseng, H.-W., Okae, H., Arima, T., Duchaine, T.F., et al. (2020). Naive Human Embryonic Stem Cells Can Give Rise to Cells with a Trophoblast-like Transcriptome and Methylome. *Stem Cell Rep.* *15*, 198–213. <https://doi.org/10.1016/j.stemcr.2020.06.003>.
- Clough, J.R., and Whittingham, D.G. (1983). Metabolism of [¹⁴C]glucose by postimplantation mouse embryos in vitro. *Development* *74*, 133–142. <https://doi.org/10.1242/dev.74.1.133>.
- Cloutier, M., Kumar, S., Buttigieg, E., Keller, L., Lee, B., Williams, A., Mojica-Perez, S., Erliandri, I., Rocha, A.M.D., Cadigan, K., et al. (2022). Preventing erosion of X-chromosome inactivation in human embryonic stem cells. *Nat. Commun.* *13*, 2516. <https://doi.org/10.1038/s41467-022-30259-x>.
- Coticchio, G., Lagalla, C., Sturmey, R., Pennetta, F., and Borini, A. (2019). The enigmatic morula: mechanisms of development, cell fate determination, self-correction and implications for ART. *Hum. Reprod. Update* *25*, 422–438. <https://doi.org/10.1093/humupd/dmz008>.
- Crick, F.H. (1958). On protein synthesis. *Symp. Soc. Exp. Biol.* *12*, 138–163. .
- Dahéron, L., Opitz, S.L., Zaehres, H., Lensch, M.W., Andrews, P.W., Itskovitz-Eldor, J., and Daley, G.Q. (2004). LIF/STAT3 signaling fails to maintain self-renewal of human embryonic stem cells. *Stem Cells Dayt. Ohio* *22*, 770–778. <https://doi.org/10.1634/stemcells.22-5-770>.
- Davidson, K.C., Mason, E.A., and Pera, M.F. (2015). The pluripotent state in mouse and human. *Dev. Camb. Engl.* *142*, 3090–3099. <https://doi.org/10.1242/dev.116061>.
- De Los Angeles, A., Loh, Y.-H., Tesar, P.J., and Daley, G.Q. (2012). Accessing naïve human pluripotency. *Curr. Opin. Genet. Dev.* *22*, 272–282. <https://doi.org/10.1016/j.gde.2012.03.001>.
- De Los Angeles, A., Ferrari, F., Xi, R., Fujiwara, Y., Benvenisty, N., Deng, H., Hochedlinger, K., Jaenisch, R., Lee, S., Leitch, H.G., et al. (2015). Hallmarks of pluripotency. *Nature* *525*, 469–478. <https://doi.org/10.1038/nature15515>.
- Deglinerti, A., Croft, G.F., Pietila, L.N., Zernicka-Goetz, M., Siggia, E.D., and Brivanlou, A.H. (2016). Self-organization of the in vitro attached human embryo. *Nature* *533*, 251–254. <https://doi.org/10.1038/nature17948>.
- Detich, N., Bovenzi, V., and Szyf, M. (2003). Valproate induces replication-independent active DNA demethylation. *J. Biol. Chem.* *278*, 27586–27592. <https://doi.org/10.1074/jbc.M303740200>.

- Dey, S.K., Lim, H., Das, S.K., Reese, J., Paria, B.C., Daikoku, T., and Wang, H. (2004). Molecular cues to implantation. *Endocr. Rev.* *25*, 341–373. <https://doi.org/10.1210/er.2003-0020>.
- Di Stefano, B., Ueda, M., Sabri, S., Brumbaugh, J., Huebner, A.J., Sahakyan, A., Clement, K., Clowers, K.J., Erickson, A.R., Shioda, K., et al. (2018). Reduced MEK inhibition preserves genomic stability in naïve human embryonic stem cells. *Nat. Methods* *15*, 732–740. <https://doi.org/10.1038/s41592-018-0104-1>.
- Dompnier, N. (2019). Avoir des enfants : entre devoir social et épanouissement conjugal. In *La France des valeurs*, (Presses universitaires de Grenoble), pp. 169–173.
- Dong, C., Beltcheva, M., Gontarz, P., Zhang, B., Popli, P., Fischer, L.A., Khan, S.A., Park, K., Yoon, E.-J., Xing, X., et al. (2020). Derivation of trophoblast stem cells from naïve human pluripotent stem cells. *eLife* *9*, e52504. <https://doi.org/10.7554/eLife.52504>.
- Dumortier, J.G., Le Verge-Serandour, M., Tortorelli, A.F., Mielke, A., de Plater, L., Turlier, H., and Maître, J.-L. (2019). Hydraulic fracturing and active coarsening position the lumen of the mouse blastocyst. *Science* *365*, 465–468. <https://doi.org/10.1126/science.aaw7709>.
- Edwards, R.G. (2006). Human implantation: the last barrier in assisted reproduction technologies? *Reprod. Biomed. Online* *13*, 887–904. [https://doi.org/10.1016/S1472-6483\(10\)61039-5](https://doi.org/10.1016/S1472-6483(10)61039-5).
- Enders, A.C., and Lopata, A. (1999). Implantation in the marmoset monkey: Expansion of the early implantation site. *Anat. Rec.* *256*, 279–299. [https://doi.org/10.1002/\(SICI\)1097-0185\(19991101\)256:3<279::AID-AR7>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1097-0185(19991101)256:3<279::AID-AR7>3.0.CO;2-O).
- Enders, A.C., Hendrickx, A.G., and Schlafke, S. (1983). Implantation in the rhesus monkey: Initial penetration of endometrium. *Am. J. Anat.* *167*, 275–298. <https://doi.org/10.1002/aja.1001670302>.
- Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* *292*, 154–156. <https://doi.org/10.1038/292154a0>.
- Fan, Y., Min, Z., Alsolami, S., Ma, Z., Zhang, E., Chen, W., Zhong, K., Pei, W., Kang, X., Zhang, P., et al. (2021). Generation of human blastocyst-like structures from pluripotent stem cells. *Cell Discov.* *7*, 81. <https://doi.org/10.1038/s41421-021-00316-8>.
- Fang, E., Wang, J., Hong, M., Zheng, L., and Tong, Q. (2019). Valproic acid suppresses Warburg effect and tumor progression in neuroblastoma. *Biochem. Biophys. Res. Commun.* *508*, 9–16. <https://doi.org/10.1016/j.bbrc.2018.11.103>.
- Firmin, J., and Maître, J.-L. (2021). Morphogenesis of the human preimplantation embryo: bringing mechanics to the clinics. *Semin. Cell Dev. Biol.* *120*, 22–31. <https://doi.org/10.1016/j.semcdb.2021.07.005>.
- Fischer, B., and Bavister, B.D. (1993). Oxygen tension in the oviduct and uterus of rhesus monkeys, hamsters and rabbits. *J. Reprod. Fertil.* *99*, 673–679. <https://doi.org/10.1530/jrf.0.0990673>.
- Fischer, B., Mootz, U., Denker, H.-W., Lambertz, M., and Beier, H.M. (1991). The dynamic structure of rabbit blastocyst coverings. *Anat. Embryol. (Berl.)* *183*, 17–27. <https://doi.org/10.1007/BF00185831>.

Fiv.fr (2019). Le taux de réussite des FIV • Fiv.fr.

Fogarty, N.M.E., Abdelbaki, A., McCarthy, A., Devito, L., Chen, A.E., Munusamy, P., Blakeley, P., Elder, K., Snell, P., Christie, L., et al. (2021). Direct reprogramming of human embryonic to trophoblast stem cells. *2021.08.18.456785*. <https://doi.org/10.1101/2021.08.18.456785>.

Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A., et al. (2013). Derivation of novel human ground state naive pluripotent stem cells. *Nature* *504*, 282–286. <https://doi.org/10.1038/nature12745>.

Gamage, T.K.J.B., Schierding, W., Hurley, D., Tsai, P., Ludgate, J.L., Bhoothpur, C., Chamley, L.W., Weeks, R.J., Macaulay, E.C., and James, J.L. (2018). The role of DNA methylation in human trophoblast differentiation. *Epigenetics* *13*, 1154–1173. <https://doi.org/10.1080/15592294.2018.1549462>.

Gao, X., Nowak-Imialek, M., Chen, X., Chen, D., Herrmann, D., Ruan, D., Chen, A.C.H., Eckersley-Maslin, M.A., Ahmad, S., Lee, Y.L., et al. (2019). Establishment of porcine and human expanded potential stem cells. *Nat. Cell Biol.* *21*, 687–699. <https://doi.org/10.1038/s41556-019-0333-2>.

Gardner, D.K. (2015). Lactate production by the mammalian blastocyst: Manipulating the microenvironment for uterine implantation and invasion? *Bioessays* *37*, 364–371. <https://doi.org/10.1002/bies.201400155>.

Gardner, D.K., and Leese, H.J. (1987). Assessment of embryo viability prior to transfer by the noninvasive measurement of glucose uptake. *J. Exp. Zool.* *242*, 103–105. <https://doi.org/10.1002/jez.1402420115>.

Gardner, R.L., Lyon, M.F., Evans, E.P., and Burtenshaw, M.D. (1985). Clonal analysis of X-chromosome inactivation and the origin of the germ line in the mouse embryo. *J. Embryol. Exp. Morphol.* *88*, 349–363.

Geens, M., and Chuva De Sousa Lopes, S.M. (2017). X chromosome inactivation in human pluripotent stem cells as a model for human development: back to the drawing board? *Hum. Reprod. Update* *23*, 520–532. <https://doi.org/10.1093/humupd/dmx015>.

Gerri, C., McCarthy, A., Scott, G.M., Regin, M., Brumm, S., Simon, C.S., Lee, J., Montesinos, C., Hassitt, C., Hockenhuil, S., et al. (2022). A conserved role of Hippo signaling in initiation of the first lineage specification event across mammals (*Developmental Biology*).

Girard, O., Lavigne, R., Chevolleau, S., Onfray, C., Com, E., Schmit, P.-O., Chapelle, M., Fréour, T., Lane, L., David, L., et al. (2023). Naive Pluripotent and Trophoblastic Stem Cell Lines as a Model for Detecting Missing Proteins in the Context of the Chromosome-Centric Human Proteome Project. *J. Proteome Res.* *22*, 1148–1158. <https://doi.org/10.1021/acs.jproteome.2c00496>.

Giulitti, S., Pellegrini, M., Zorzan, I., Martini, P., Gagliano, O., Mutarelli, M., Ziller, M.J., Cacchiarelli, D., Romualdi, C., Elvassore, N., et al. (2019). Direct generation of human naive induced pluripotent stem cells from somatic cells in microfluidics. *Nat. Cell Biol.* *21*, 275–286. <https://doi.org/10.1038/s41556-018-0254-5>.

Goedel, A., and Lanner, F. (2021). A peek into the black box of human embryology. *Nature* *600*, 223–224. <https://doi.org/10.1038/d41586-021-03381-x>.

Gokhale, P.J., Au-Young, J.K., Dadi, S., Keys, D.N., Harrison, N.J., Jones, M., Soneji, S., Enver, T., Sherlock, J.K., and Andrews, P.W. (2015). Culture adaptation alters transcriptional hierarchies among single human embryonic stem cells reflecting altered patterns of differentiation. *PLoS One* 10, e0123467. <https://doi.org/10.1371/journal.pone.0123467>.

Gott, A.L., Hardy, K., Winston, R.M.L., and Leese, H.J. (1990). Non-invasive measurement of pyruvate and glucose uptake and lactate production by single human preimplantation embryos. *Hum. Reprod.* 5, 104–108. <https://doi.org/10.1093/oxfordjournals.humrep.a137028>.

Graf, A., Krebs, S., Heininen-Brown, M., Zakhartchenko, V., Blum, H., and Wolf, E. (2014). Genome activation in bovine embryos: Review of the literature and new insights from RNA sequencing experiments. *Anim. Reprod. Sci.* 149, 46–58. <https://doi.org/10.1016/j.anireprosci.2014.05.016>.

Graham, C.H., Hawley, T.S., Hawley, R.G., MacDougall, J.R., Kerbel, R.S., Khoo, N., and Lala, P.K. (1993). Establishment and characterization of first trimester human trophoblast cells with extended lifespan. *Exp. Cell Res.* 206, 204–211. <https://doi.org/10.1006/excr.1993.1139>.

Gu, W., Gaeta, X., Sahakyan, A., Chan, A.B., Hong, C.S., Kim, R., Braas, D., Plath, K., Lowry, W.E., and Christofk, H.R. (2016). Glycolytic Metabolism Plays a Functional Role in Regulating Human Pluripotent Stem Cell State. *Cell Stem Cell* 19, 476–490. <https://doi.org/10.1016/j.stem.2016.08.008>.

Guan, J., Wang, G., Wang, J., Zhang, Z., Fu, Y., Cheng, L., Meng, G., Lyu, Y., Zhu, J., Li, Y., et al. (2022). Chemical reprogramming of human somatic cells to pluripotent stem cells. *Nature* 605, 325–331. <https://doi.org/10.1038/s41586-022-04593-5>.

Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., and Smith, A. (2009). Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Dev. Camb. Engl.* 136, 1063–1069. <https://doi.org/10.1242/dev.030957>.

Guo, G., von Meyenn, F., Santos, F., Chen, Y., Reik, W., Bertone, P., Smith, A., and Nichols, J. (2016). Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass. *Stem Cell Rep.* 6, 437–446. <https://doi.org/10.1016/j.stemcr.2016.02.005>.

Guo, G., von Meyenn, F., Rostovskaya, M., Clarke, J., Dietmann, S., Baker, D., Sahakyan, A., Myers, S., Bertone, P., Reik, W., et al. (2017). Epigenetic resetting of human pluripotency. *Development* 144, 2748–2763. <https://doi.org/10.1242/dev.146811>.

Guo, G., Stirparo, G.G., Strawbridge, S.E., Spindlow, D., Yang, J., Clarke, J., Dattani, A., Yanagida, A., Li, M.A., Myers, S., et al. (2021). Human naive epiblast cells possess unrestricted lineage potential. *Cell Stem Cell* 28, 1040–1056.e6. <https://doi.org/10.1016/j.stem.2021.02.025>.

Guo, N., Li, Y., Ai, J., Gu, L., Chen, W., and Liu, Q. (2014). Two different concentrations of oxygen for culturing precompaction stage embryos on human embryo development competence: a prospective randomized sibling-oocyte study. *Int. J. Clin. Exp. Pathol.* 7, 6191–6198. .

Gurdon, J.B., Elsdale, T.R., and Fischberg, M. (1958). Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. *Nature* 182, 64–65. <https://doi.org/10.1038/182064a0>.

Häcker, V. (1892). Die Kerntheilungsvorgänge bei der Mesoderm und Entodermbildung von Cyclops. *Arch. Für Mikrosk. Anat.* 39, 556–581. <https://doi.org/10.1007/BF02961538>.

Hackett, J.A., and Surani, M.A. (2014). Regulatory principles of pluripotency: from the ground state up. *Cell Stem Cell* 15, 416–430. <https://doi.org/10.1016/j.stem.2014.09.015>.

Hackett, J.A., Kobayashi, T., Dietmann, S., and Surani, M.A. (2017). Activation of Lineage Regulators and Transposable Elements across a Pluripotent Spectrum. *Stem Cell Rep.* 8, 1645–1658. <https://doi.org/10.1016/j.stemcr.2017.05.014>.

Haeckel, E.H.P.A. (1872). *Natürliche Schöpfungsgeschichte* (Verlag von Georg Keimer).

Haeckel, E.H.P.A. (1877). *Anthropogenie* (W. Engelmann).

Haider, S., and Pal, R. (2013). Integrated Analysis of Transcriptomic and Proteomic Data. *Curr. Genomics* 14, 91–110. <https://doi.org/10.2174/1389202911314020003>.

Hanna, J., Markoulaki, S., Mitalipova, M., Cheng, A.W., Cassady, J.P., Staerk, J., Carey, B.W., Lengner, C.J., Foreman, R., Love, J., et al. (2009a). Metastable Pluripotent States in NOD-Mouse-Derived ESCs. *Cell Stem Cell* 4, 513–524. <https://doi.org/10.1016/j.stem.2009.04.015>.

Hanna, J., Saha, K., Pando, B., van Zon, J., Lengner, C.J., Creighton, M.P., van Oudenaarden, A., and Jaenisch, R. (2009b). Direct cell reprogramming is a stochastic process amenable to acceleration. 595–601. .

Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat, J., Carey, B.W., and Jaenisch, R. (2010). Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. 9222–9227. .

Hayashi, K., de Sousa Lopes, S.M.C., Tang, F., Lao, K., and Surani, M.A. (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell* 3, 391–401. <https://doi.org/10.1016/j.stem.2008.07.027>.

Hertig, A.T., Rock, J., Adams, E.C., and Menkin, M.C. (1959). THIRTY-FOUR FERTILIZED HUMAN OVA, GOOD, BAD AND INDIFFERENT, RECOVERED FROM 210 WOMEN OF KNOWN FERTILITY : A Study of Biologic Wastage in Early Human Pregnancy. *Pediatrics* 23, 202–211. <https://doi.org/10.1542/peds.23.1.202>.

Ho, W.-T., Chang, J.-S., Chen, T.-C., Wang, J.-K., Chang, S.-W., Yang, M.-H., Jou, T.-S., and Wang, I.-J. (2022). Inhibition of Rho-associated protein kinase activity enhances oxidative phosphorylation to support corneal endothelial cell migration. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* 36, e22397. <https://doi.org/10.1096/fj.202101442RR>.

Horii, M., Li, Y., Wakeland, A.K., Pizzo, D.P., Nelson, K.K., Sabatini, K., Laurent, L.C., Liu, Y., and Parast, M.M. (2016). Human pluripotent stem cells as a model of trophoblast differentiation in both normal development and disease. *Proc. Natl. Acad. Sci. U. S. A.* 113, E3882–3891. <https://doi.org/10.1073/pnas.1604747113>.

Hough, S.R., Thornton, M., Mason, E., Mar, J.C., Wells, C.A., and Pera, M.F. (2014). Single-Cell Gene Expression Profiles Define Self-Renewing, Pluripotent, and Lineage Primed States of Human Pluripotent Stem Cells. *Stem Cell Rep.* 2, 881–895. <https://doi.org/10.1016/j.stemcr.2014.04.014>.

Houghton, F.D., Thompson, J.G., Kennedy, C.J., and Leese, H.J. (1996). Oxygen consumption and energy metabolism of the early mouse embryo. *Mol. Reprod. Dev.* 44, 476–485. [https://doi.org/10.1002/\(SICI\)1098-2795\(199608\)44:4<476::AID-MRD7>3.0.CO;2-I](https://doi.org/10.1002/(SICI)1098-2795(199608)44:4<476::AID-MRD7>3.0.CO;2-I).

Hunt, T.K., Aslam, R.S., Beckert, S., Wagner, S., Ghani, Q.P., Hussain, M.Z., Roy, S., and Sen, C.K. (2007). Aerobically Derived Lactate Stimulates Revascularization and Tissue Repair via Redox Mechanisms. *Antioxid. Redox Signal.* 9, 1115–1124. <https://doi.org/10.1089/ars.2007.1674>.

Hurlbut, J.B. (2017). *Experiments in Democracy: Human Embryo Research and the Politics of Bioethics* (Columbia University Press).

Io, S., Kabata, M., Iemura, Y., Semi, K., Morone, N., Minagawa, A., Wang, B., Okamoto, I., Nakamura, T., Kojima, Y., et al. (2021). Capturing human trophoblast development with naive pluripotent stem cells in vitro. *Cell Stem Cell* 28, 1023–1039.e13. <https://doi.org/10.1016/j.stem.2021.03.013>.

Iwata, K., Yumoto, K., Sugishima, M., Mizoguchi, C., Kai, Y., Iba, Y., and Mio, Y. (2014). Analysis of compaction initiation in human embryos by using time-lapse cinematography. *J. Assist. Reprod. Genet.* 31, 421–426. <https://doi.org/10.1007/s10815-014-0195-2>.

James, D., Levine, A.J., Besser, D., and Hemmati-Brivanlou, A. (2005). TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Dev. Camb. Engl.* 132, 1273–1282. <https://doi.org/10.1242/dev.01706>.

James, J.L., Carter, A.M., and Chamley, L.W. (2012). Human placentation from nidation to 5 weeks of gestation. Part I: What do we know about formative placental development following implantation? *Placenta* 33, 327–334. <https://doi.org/10.1016/j.placenta.2012.01.020>.

Jang, Y.J., Kim, M., Lee, B.-K., and Kim, J. (2022). Induction of human trophoblast stem-like cells from primed pluripotent stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2115709119. <https://doi.org/10.1073/pnas.2115709119>.

Jansen, J. (2005). The First Successful Allogeneic Bone-Marrow Transplant: Georges Mathé. *Transfus. Med. Rev.* 19, 246–248. <https://doi.org/10.1016/j.tmr.2005.02.006>.

Jedrusik, A. (2015). Making the first decision: lessons from the mouse. *Reprod. Med. Biol.* 14, 135–150. <https://doi.org/10.1007/s12522-015-0206-8>.

Johnson, M.H., and Ziomek, C.A. (1981). The foundation of two distinct cell lineages within the mouse morula. *Cell* 24, 71–80. [https://doi.org/10.1016/0092-8674\(81\)90502-X](https://doi.org/10.1016/0092-8674(81)90502-X).

Jones, C.J.P., Elena Ortíz, M., Croxatto, H.B., Manzur, A., Slevin, G., and Aplin, J.D. (2001). Muc1 and Glycan Expression in the Oviduct and Endometrium of a New World Monkey, *Cebus apella*. *Biol. Reprod.* 64, 1535–1544. <https://doi.org/10.1095/biolreprod64.5.1535>.

- Kagawa, H., Javali, A., Khoei, H.H., Sommer, T.M., Sestini, G., Novatchkova, M., Scholte op Reimer, Y., Castel, G., Bruneau, A., Maenhoudt, N., et al. (2021). Human blastoids model blastocyst development and implantation. *Nature* 1–6. <https://doi.org/10.1038/s41586-021-04267-8>.
- Kalkan, T., Olova, N., Roode, M., Mulas, C., Lee, H.J., Nett, I., Marks, H., Walker, R., Stunnenberg, H.G., Lilley, K.S., et al. (2017). Tracking the embryonic stem cell transition from ground state pluripotency. *Dev. Camb. Engl.* *144*, 1221–1234. <https://doi.org/10.1242/dev.142711>.
- Karvas, R.M., Khan, S.A., Verma, S., Yin, Y., Kulkarni, D., Dong, C., Park, K.-M., Chew, B., Sane, E., Fischer, L.A., et al. (2022). Stem-cell-derived trophoblast organoids model human placental development and susceptibility to emerging pathogens. *Cell Stem Cell* *29*, 810-825.e8. <https://doi.org/10.1016/j.stem.2022.04.004>.
- Kilens, S., Meistermann, D., Moreno, D., Chariou, C., Gaignerie, A., Reignier, A., Lelièvre, Y., Casanova, M., Vallot, C., Nedellec, S., et al. (2018). Parallel derivation of isogenic human primed and naive induced pluripotent stem cells. *Nat. Commun.* *9*, 360. <https://doi.org/10.1038/s41467-017-02107-w>.
- Kinoshita, M., Barber, M., Mansfield, W., Cui, Y., Spindlow, D., Stirparo, G.G., Dietmann, S., Nichols, J., and Smith, A. (2021). Capture of Mouse and Human Stem Cells with Features of Formative Pluripotency. *Cell Stem Cell* *28*, 453-471.e8. <https://doi.org/10.1016/j.stem.2020.11.005>.
- Knöfler, M., Haider, S., Saleh, L., Pollheimer, J., Gamage, T.K.J.B., and James, J. (2019). Human placenta and trophoblast development: key molecular mechanisms and model systems. *Cell. Mol. Life Sci. CMLS* *76*, 3479–3496. <https://doi.org/10.1007/s00018-019-03104-6>.
- L, B.R. (1973). Nutrition and metabolism of the ovum, zygote and blastocyst. *Handb. Physiol.* *11*, 165–185. .
- Lammers, J., Loubersac, S., David, L., Freour, T., and Reignier, A. (2023). Developmental-stage specific single-cell human embryo dissociation. *STAR Protoc.* *4*, 102363. <https://doi.org/10.1016/j.xpro.2023.102363>.
- Leblond, C.P. (1981). The life history of cells in renewing systems. *Am. J. Anat.* *160*, 113–158. <https://doi.org/10.1002/aja.1001600202>.
- Leblond, C.P., Wilkinson, G.W., Bélanger, L.F., and Robichon, J. (1950). Radio-autographic visualization of bone formation in the rat. *Am. J. Anat.* *86*, 289–341. <https://doi.org/10.1002/aja.1000860205>.
- Leblond, C.P., Messier, B., and Kopriwa, B. (1959). Thymidine-H3 as a tool for the investigation of the renewal of cell populations. *Lab. Investig. J. Tech. Methods Pathol.* *8*, 296–306; discussion 306-308. .
- Lee, K.Y., and DeMayo, F.J. (2004). Animal models of implantation. *Reproduction* *128*, 679–695. <https://doi.org/10.1530/rep.1.00340>.
- Lee, C.Q.E., Gardner, L., Turco, M., Zhao, N., Murray, M.J., Coleman, N., Rossant, J., Hemberger, M., and Moffett, A. (2016). What Is Trophoblast? A Combination of Criteria Define Human First-Trimester Trophoblast. *Stem Cell Rep.* *6*, 257–272. <https://doi.org/10.1016/j.stemcr.2016.01.006>.

- Leitch, H.G., McEwen, K.R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J.G., Smith, A., et al. (2013). Naive pluripotency is associated with global DNA hypomethylation. *Nat. Struct. Mol. Biol.* *20*, 311–316. <https://doi.org/10.1038/nsmb.2510>.
- Leng, L., Sun, J., Huang, J., Gong, F., Yang, L., Zhang, S., Yuan, X., Fang, F., Xu, X., Luo, Y., et al. (2019). Single-Cell Transcriptome Analysis of Uniparental Embryos Reveals Parent-of-Origin Effects on Human Preimplantation Development. *Cell Stem Cell* *25*, 697-712.e6. <https://doi.org/10.1016/j.stem.2019.09.004>.
- Lerou, P. (2011). Embryonic stem cell derivation from human embryos. *Methods Mol. Biol. Clifton NJ* *767*, 31–35. https://doi.org/10.1007/978-1-61779-201-4_3.
- Li, J., Zhu, Q., Cao, J., Liu, Y., Lu, Y., Sun, Y., Li, Q., Huang, Y., Shang, S., Bian, X., et al. (2023). Cynomolgus monkey embryo model captures gastrulation and early pregnancy. *Cell Stem Cell* *30*, 362-377.e7. <https://doi.org/10.1016/j.stem.2023.03.009>.
- Li, Y., Moretto-Zita, M., Soncin, F., Wakeland, A., Wolfe, L., Leon-Garcia, S., Pandian, R., Pizzo, D., Cui, L., Nazor, K., et al. (2013). BMP4-directed trophoblast differentiation of human embryonic stem cells is mediated through a $\Delta Np63+$ cytotrophoblast stem cell state. *Dev. Camb. Engl.* *140*, 3965–3976. <https://doi.org/10.1242/dev.092155>.
- Lim, S.-O., Li, C.-W., Xia, W., Lee, H.-H., Chang, S.-S., Shen, J., Hsu, J.L., Raftery, D., Djukovic, D., Gu, H., et al. (2016). EGFR Signaling Enhances Aerobic Glycolysis in Triple-Negative Breast Cancer Cells to Promote Tumor Growth and Immune Escape. *Cancer Res.* *76*, 1284–1296. <https://doi.org/10.1158/0008-5472.CAN-15-2478>.
- Liu, B., Chen, S., Xu, Y., Lyu, Y., Wang, J., Du, Y., Sun, Y., Liu, H., Zhou, H., Lai, W., et al. (2021a). Chemically defined and xeno-free culture condition for human extended pluripotent stem cells. *Nat. Commun.* *12*, 3017. <https://doi.org/10.1038/s41467-021-23320-8>.
- Liu, B., Chen, S., Xu, Y., Lyu, Y., Wang, J., Du, Y., Sun, Y., Liu, H., Zhou, H., Lai, W., et al. (2021b). Chemically defined and xeno-free culture condition for human extended pluripotent stem cells. *Nat. Commun.* *12*, 3017. <https://doi.org/10.1038/s41467-021-23320-8>.
- Liu, X., Nefzger, C.M., Rossello, F.J., Chen, J., Knaupp, A.S., Firas, J., Ford, E., Pflueger, J., Paynter, J.M., Chy, H.S., et al. (2017). Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. *Nat. Methods* *14*, 1055–1062. <https://doi.org/10.1038/nmeth.4436>.
- Liu, X., Ouyang, J.F., Rossello, F.J., Tan, J.P., Davidson, K.C., Valdes, D.S., Schröder, J., Sun, Y.B.Y., Chen, J., Knaupp, A.S., et al. (2020). Reprogramming roadmap reveals route to human induced trophoblast stem cells. *Nature* *586*, 101–107. <https://doi.org/10.1038/s41586-020-2734-6>.
- Liu, X., Tan, J.P., Schröder, J., Aberkane, A., Ouyang, J.F., Mohenska, M., Lim, S.M., Sun, Y.B.Y., Chen, J., Sun, G., et al. (2021c). Modelling human blastocysts by reprogramming fibroblasts into iBlastoids. *Nature* *591*, 627–632. <https://doi.org/10.1038/s41586-021-03372-y>.
- Lowry, W.E., Richter, L., Yachechko, R., Pyle, A.D., Tchieu, J., Sridharan, R., Clark, A.T., and Plath, K. (2008). Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 2883–2888. <https://doi.org/10.1073/pnas.0711983105>.

M, G., G, M., K, O., and B, H. (2009). Factors involved in regulating trophoblast fusion: potential role in the development of preeclampsia. *Placenta 30 Suppl A*. <https://doi.org/10.1016/j.placenta.2008.10.011>.

Ma, Y., Ma, M., Sun, J., Li, W., Li, Y., Guo, X., and Zhang, H. (2019). CHIR-99021 regulates mitochondrial remodelling via β -catenin signalling and miRNA expression during endodermal differentiation. *J. Cell Sci. 132*, jcs229948. <https://doi.org/10.1242/jcs.229948>.

Marques-Pereira, J.P., and Leblond, C.P. (1965). MITOSIS AND DIFFERENTIATION IN THE STRATIFIED SQUAMOUS EPITHELIUM OF THE RAT ESOPHAGUS. *Am. J. Anat. 117*, 73–87. <https://doi.org/10.1002/aja.1001170106>.

Martello, G., and Smith, A. (2014). The nature of embryonic stem cells. *Annu. Rev. Cell Dev. Biol. 30*, 647–675. <https://doi.org/10.1146/annurev-cellbio-100913-013116>.

Martello, G., Sugimoto, T., Diamanti, E., Joshi, A., Hannah, R., Ohtsuka, S., Göttgens, B., Niwa, H., and Smith, A. (2012). Esrrb Is a Pivotal Target of the Gsk3/Tcf3 Axis Regulating Embryonic Stem Cell Self-Renewal. 491–504. .

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. 78*, 7634–7638. <https://doi.org/10.1073/pnas.78.12.7634>.

Mascetti, V.L., and Pedersen, R.A. (2016). Human-Mouse Chimerism Validates Human Stem Cell Pluripotency. *Cell Stem Cell 18*, 67–72. <https://doi.org/10.1016/j.stem.2015.11.017>.

Mathe, G., and Bernard, J. (1959). [Attempts at treatment of grafted leukemia 1210 by x-irradiation followed by transfusion of normal hematopoietic cells (isologous or homologous, myeloid or lymphoid, adult or embryonic)]. *Rev. Fr. Etud. Clin. Biol. 4*, 442–446. .

Mathe, G., Thomas, E.D., and Ferrebee, J.W. (1959a). The restoration of marrow function after lethal irradiation in man: a review. *Transplant. Bull. 6*, 407–409. <https://doi.org/10.1097/00006534-195910000-00034>.

Mathe, G., Bernard, J., Schwarzenberg, L., Larrieu, M.J., Lalanne, C.M., Dutreix, A., Denoix, P.F., Surmont, J., Schwarzmann, V., and Ceoara, B. (1959b). [Trial treatment of patients afflicted with acute leukemia in remission with total irradiation followed by homologous bone marrow transfusion]. *Rev. Fr. Etud. Clin. Biol. 4*, 675–704. .

Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T., and Yokota, T. (1999). STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. *EMBO J. 18*, 4261–4269. <https://doi.org/10.1093/emboj/18.15.4261>.

Matthews, K.R., and Moralí, D. (2020). National human embryo and embryoid research policies: a survey of 22 top research-intensive countries. *Regen. Med. 15*, 1905–1917. <https://doi.org/10.2217/rme-2019-0138>.

Matthews, K.R., and Rowland, M.L. (2011). Stem cell policy in the Obama age: UK and US perspectives. *Regen. Med. 6*, 125–132. <https://doi.org/10.2217/rme.10.92>.

Megli, C.J., and Coyne, C.B. (2022). Infections at the maternal–fetal interface: an overview of pathogenesis and defence. *Nat. Rev. Microbiol.* *20*, 67–82. <https://doi.org/10.1038/s41579-021-00610-y>.

Meistermann, D., Bruneau, A., Loubersac, S., Reignier, A., Firmin, J., François-Campion, V., Kilens, S., Lelièvre, Y., Lammers, J., Feyeux, M., et al. (2021). Integrated pseudotime analysis of human pre-implantation embryo single-cell transcriptomes reveals the dynamics of lineage specification. *Cell Stem Cell* *28*, 1625-1640.e6. <https://doi.org/10.1016/j.stem.2021.04.027>.

Mekhoubad, S., Bock, C., de Boer, A.S., Kiskinis, E., Meissner, A., and Eggan, K. (2012). Erosion of dosage compensation impacts human iPSC disease modeling. *Cell Stem Cell* *10*, 595–609. <https://doi.org/10.1016/j.stem.2012.02.014>.

Moris, N. (2023). Stem cells used to model a two-week-old human embryo. *Nature* *622*, 469–470. <https://doi.org/10.1038/d41586-023-03150-y>.

Morrison, S.J., Shah, N.M., and Anderson, D.J. (1997). Regulatory Mechanisms in Stem Cell Biology. *Cell* *88*, 287–298. [https://doi.org/10.1016/S0092-8674\(00\)81867-X](https://doi.org/10.1016/S0092-8674(00)81867-X).

Nakamura, T., Okamoto, I., Sasaki, K., Yabuta, Y., Iwatani, C., Tsuchiya, H., Seita, Y., Nakamura, S., Yamamoto, T., and Saitou, M. (2016). A developmental coordinate of pluripotency among mice, monkeys and humans. *Nature* *537*, 57–62. <https://doi.org/10.1038/nature19096>.

Niakan, K.K., Han, J., Pedersen, R.A., Simon, C., and Pera, R.A.R. (2012). Human pre-implantation embryo development. *Dev. Camb. Engl.* *139*, 829–841. <https://doi.org/10.1242/dev.060426>.

Nichols, J., and Smith, A. (2009). Naive and Primed Pluripotent States. *Cell Stem Cell* *4*, 487–492. <https://doi.org/10.1016/j.stem.2009.05.015>.

Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev.* *12*, 2048–2060. <https://doi.org/10.1101/gad.12.13.2048>.

Ohinata, Y., Ohta, H., Shigeta, M., Yamanaka, K., Wakayama, T., and Saitou, M. (2009). A Signaling Principle for the Specification of the Germ Cell Lineage in Mice. *571–584*.

Okae, H., Toh, H., Sato, T., Hiura, H., Takahashi, S., Shirane, K., Kabayama, Y., Suyama, M., Sasaki, H., and Arima, T. (2018). Derivation of Human Trophoblast Stem Cells. *Cell Stem Cell* *22*, 50-63.e6. <https://doi.org/10.1016/j.stem.2017.11.004>.

Oldak, B., Wildschutz, E., Bondarenko, V., Comar, M.-Y., Zhao, C., Aguilera-Castrejon, A., Tarazi, S., Viukov, S., Pham, T.X.A., Ashouokhi, S., et al. (2023). Complete human day 14 post-implantation embryo models from naive ES cells. *Nature* *622*, 562–573. <https://doi.org/10.1038/s41586-023-06604-5>.

Onfray, C., Tan, J.P., Kilens, S., Liu, X., Polo, J., and David, L. (2022). Induction of Human Naïve Pluripotent Stem Cells from Somatic Cells. *Methods Mol. Biol. Clifton NJ* *2416*, 39–51. https://doi.org/10.1007/978-1-0716-1908-7_4.

Onfray, C., Chevolleau, S., Moinard, E., Girard, O., Mahadik, K., Allsop, R.N., Georgolopoulos, G., Lavigne, R., Renoult, O., Aksoy, I., et al. (2023). Deciphering Hallmark Combination Distinct for Peri-Implantation Stem Cell Models. <https://doi.org/10.2139/ssrn.4521994>.

- Orendi, K., Gauster, M., Moser, G., Meiri, H., and Huppertz, B. (2010). The choriocarcinoma cell line BeWo: syncytial fusion and expression of syncytium-specific proteins. *Reprod. Camb. Engl.* *140*, 759–766. <https://doi.org/10.1530/REP-10-0221>.
- Paepe, C.D., Aberkane, A., Dewandre, D., Essahib, W., Sermon, K., Geens, M., Verheyen, G., Tournaye, H., and Velde, H.V. de (2019). BMP4 plays a role in apoptosis during human preimplantation development. *Mol. Reprod. Dev.* *86*, 53–62. <https://doi.org/10.1002/mrd.23081>.
- Palma, V., Pitossi, F.J., Rehen, S.K., Touriño, C., and Velasco, I. (2015). Stem cell research in Latin America: update, challenges and opportunities in a priority research area. *Regen. Med.* *10*, 785–798. <https://doi.org/10.2217/rme.15.44>.
- Panda, A., Zylicz, J.J., and Pasque, V. (2020). New Insights into X-Chromosome Reactivation during Reprogramming to Pluripotency. *Cells* *9*, 2706. <https://doi.org/10.3390/cells9122706>.
- Pappenheim, A. (1896). Ueber Entwicklung und Ausbildung der Erythroblasten. *Arch. Für Pathol. Anat. Physiol. Für Klin. Med.* *145*, 587–643. <https://doi.org/10.1007/BF01969901>.
- Park, I.-H., Zhao, R., West, J.A., Yabuuchi, A., Huo, H., Ince, T.A., Lerou, P.H., Lensch, M.W., and Daley, G.Q. (2007). Reprogramming of human somatic cells to pluripotency with defined factors. *141–146*. .
- Pastor, W.A., Chen, D., Liu, W., Kim, R., Sahakyan, A., Lukianchikov, A., Plath, K., Jacobsen, S.E., and Clark, A.T. (2016). Naive Human Pluripotent Cells Feature a Methylation Landscape Devoid of Blastocyst or Germline Memory. *Cell Stem Cell* *18*, 323–329. <https://doi.org/10.1016/j.stem.2016.01.019>.
- Patel, S., Bonora, G., Sahakyan, A., Kim, R., Chronis, C., Langerman, J., Fitz-Gibbon, S., Rubbi, L., Skelton, R.J.P., Ardehali, R., et al. (2017). Human Embryonic Stem Cells Do Not Change Their X Inactivation Status during Differentiation. *Cell Rep.* *18*, 54–67. <https://doi.org/10.1016/j.celrep.2016.11.054>.
- Pattillo, R.A., and Gey, G.O. (1968). The establishment of a cell line of human hormone-synthesizing trophoblastic cells in vitro. *Cancer Res.* *28*, 1231–1236. .
- Pedroza, M., Gassaloglu, S.I., Dias, N., Zhong, L., Hou, T.-C.J., Kretzmer, H., Smith, Z.D., and Sozen, B. (2023). Self-patterning of human stem cells into post-implantation lineages. *Nature* *622*, 574–583. <https://doi.org/10.1038/s41586-023-06354-4>.
- Pham, T.X.A., Panda, A., Kagawa, H., To, S.K., Ertekin, C., Georgolopoulos, G., van Knippenberg, S.S.F.A., Allsop, R.N., Bruneau, A., Chui, J.S.-H., et al. (2022). Modeling human extraembryonic mesoderm cells using naive pluripotent stem cells. *Cell Stem Cell* *29*, 1346-1365.e10. <https://doi.org/10.1016/j.stem.2022.08.001>.
- Polanski, L.T., Baumgarten, M.N., Quenby, S., Brosens, J., Campbell, B.K., and Raine-Fenning, N.J. (2014). What exactly do we mean by ‘recurrent implantation failure’? A systematic review and opinion. *Reprod. Biomed. Online* *28*, 409–423. <https://doi.org/10.1016/j.rbmo.2013.12.006>.
- Posfai, E., Schell, J.P., Janiszewski, A., Rovic, I., Murray, A., Bradshaw, B., Yamakawa, T., Pardon, T., El Bakkali, M., Talon, I., et al. (2021). Evaluating totipotency using criteria of increasing stringency. *Nat. Cell Biol.* *23*, 49–60. <https://doi.org/10.1038/s41556-020-00609-2>.

- Pranke, P., Chagastelles, P., and Sperling, L.E. (2014). The Current State of Research with Human Pluripotent Stem Cells in Brazil. *Stem Cells Dev.* 23, 20–23. <https://doi.org/10.1089/scd.2014.0320>.
- Qin, H., Hejna, M., Liu, Y., Percharde, M., Wossidlo, M., Blouin, L., Durruthy-Durruthy, J., Wong, P., Qi, Z., Yu, J., et al. (2016). YAP Induces Human Naive Pluripotency. *Cell Rep.* 14, 2301–2312. <https://doi.org/10.1016/j.celrep.2016.02.036>.
- Ramalho-Santos, M., and Willenbring, H. (2007). On the Origin of the Term “Stem Cell.” *Cell Stem Cell* 1, 35–38. <https://doi.org/10.1016/j.stem.2007.05.013>.
- Ravitsky, V., and Kimmins, S. (2019). The forgotten men: rising rates of male infertility urgently require new approaches for its prevention, diagnosis and treatment. *Biol. Reprod.* 101, 872–874. <https://doi.org/10.1093/biolre/iox161>.
- Raz, R., Lee, C.-K., Cannizzaro, L.A., d’Eustachio, P., and Levy, D.E. (1999). Essential role of STAT3 for embryonic stem cell pluripotency. *Proc. Natl. Acad. Sci. U. S. A.* 96, 2846–2851. .
- Redman, C.W., and Sargent, I.L. (2005). Latest advances in understanding preeclampsia. *Science* 308, 1592–1594. <https://doi.org/10.1126/science.1111726>.
- Rep. Pascrell, B. (2019). Text - H.R.1865 - 116th Congress (2019-2020): Further Consolidated Appropriations Act, 2020.
- Rivron, N. (2018). Formation of blastoids from mouse embryonic and trophoblast stem cells.
- Rossant, J. (2008). Stem cells and early lineage development. *Cell* 132, 527–531. <https://doi.org/10.1016/j.cell.2008.01.039>.
- Rossant, J. (2015). Mouse and human blastocyst-derived stem cells: vive les differences. *Dev. Camb. Engl.* 142, 9–12. <https://doi.org/10.1242/dev.115451>.
- Rostovskaya, M., Stirparo, G.G., and Smith, A. (2019). Capacitation of human naïve pluripotent stem cells for multi-lineage differentiation. *Dev. Camb. Engl.* 146, dev172916. <https://doi.org/10.1242/dev.172916>.
- Roth, V.L. (1988). Vertebrate Fetal Membranes. Comparative Ontogeny and Morphology; Evolution; Phylogenetic Significance; Basic Functions; Research Opportunities. Harland W. Mossman. *Q. Rev. Biol.* 63, 89–89. <https://doi.org/10.1086/415763>.
- Rugg-Gunn, P.J. (2022). Induction of Human Naïve Pluripotency Using Chemical Resetting. *Methods Mol. Biol. Clifton NJ* 2416, 29–37. https://doi.org/10.1007/978-1-0716-1908-7_3.
- Sahakyan, A., Kim, R., Chronis, C., Sabri, S., Bonora, G., Theunissen, T.W., Kuoy, E., Langerman, J., Clark, A.T., Jaenisch, R., et al. (2017). Human Naive Pluripotent Stem Cells Model X Chromosome Dampening and X Inactivation. *Cell Stem Cell* 20, 87–101. <https://doi.org/10.1016/j.stem.2016.10.006>.
- Samudio-Ruiz, S.L., and Hudson, L.G. (2012). Increased DNA methyltransferase activity and DNA methylation following epidermal growth factor stimulation in ovarian cancer cells. *Epigenetics* 7, 216–224. <https://doi.org/10.4161/epi.7.3.19273>.

- Sathananthan, D.H., Gunasheela, S., and Menezes, J. (2003). Mechanics of human blastocyst hatching in vitro. *Reprod. Biomed. Online* 7, 228–234. [https://doi.org/10.1016/S1472-6483\(10\)61757-9](https://doi.org/10.1016/S1472-6483(10)61757-9).
- Schlafke, S., and Enders, A.C. (1975). Cellular Basis of Interaction Between Trophoblast and Uterus at Implantation. *Biol. Reprod.* 12, 41–65. <https://doi.org/10.1095/biolreprod12.1.41>.
- Seetharam, A.S., Vu, H.T.H., Choi, S., Khan, T., Sheridan, M.A., Ezashi, T., Roberts, R.M., and Tuteja, G. (2022). The product of BMP-directed differentiation protocols for human primed pluripotent stem cells is placental trophoblast and not amnion. *Stem Cell Rep.* 17, 1289–1302. <https://doi.org/10.1016/j.stemcr.2022.04.014>.
- Shahbazi, M.N. (2020). Mechanisms of human embryo development: from cell fate to tissue shape and back. *Dev. Camb. Engl.* 147, dev190629. <https://doi.org/10.1242/dev.190629>.
- Shahbazi, M.N., Jedrusik, A., Vuoristo, S., Recher, G., Hupalowska, A., Bolton, V., Fogarty, N.N.M., Campbell, A., Devito, L., Ilic, D., et al. (2016). Self-organisation of the human embryo in the absence of maternal tissues. *Nat. Cell Biol.* 18, 700–708. <https://doi.org/10.1038/ncb3347>.
- Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W., and Smith, A. (2008). Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol.* 6, e253. <https://doi.org/10.1371/journal.pbio.0060253>.
- Sirard, M.-A. (2012). Factors Affecting Oocyte and Embryo Transcriptomes. *Reprod. Domest. Anim.* 47, 148–155. <https://doi.org/10.1111/j.1439-0531.2012.02069.x>.
- Siriwardena, D., and Boroviak, T.E. (2022). Evolutionary divergence of embryo implantation in primates. *Philos. Trans. R. Soc. B Biol. Sci.* 377, 20210256. <https://doi.org/10.1098/rstb.2021.0256>.
- Slavov, N. (2021). Driving Single Cell Proteomics Forward with Innovation. *J. Proteome Res.* 20, 4915–4918. <https://doi.org/10.1021/acs.jproteome.1c00639>.
- Smith, A. (2009). Design principles of pluripotency. *EMBO Mol. Med.* 1, 251–254. <https://doi.org/10.1002/emmm.200900035>.
- Smith, A. (2017). Formative pluripotency: the executive phase in a developmental continuum. *Dev. Camb. Engl.* 144, 365–373. <https://doi.org/10.1242/dev.142679>.
- Smith, A.G. (2001). Embryo-derived stem cells: of mice and men. *Annu. Rev. Cell Dev. Biol.* 17, 435–462. <https://doi.org/10.1146/annurev.cellbio.17.1.435>.
- Smith, C.A., Moore, H.D.M., and Hearn, J.P. (1987). The ultrastructure of early implantation in the marmoset monkey (*Callithrix jacchus*). *Anat. Embryol. (Berl.)* 175, 399–410. <https://doi.org/10.1007/BF00309853>.
- Smith, Z.D., Chan, M.M., Humm, K.C., Karnik, R., Mekhoubad, S., Regev, A., Eggan, K., and Meissner, A. (2014). DNA methylation dynamics of the human preimplantation embryo. *Nature* 511, 611–615. <https://doi.org/10.1038/nature13581>.

- Solter, D., and Knowles, B.B. (1975). Immunosurgery of mouse blastocyst. *Proc. Natl. Acad. Sci. U. S. A.* 72, 5099–5102. <https://doi.org/10.1073/pnas.72.12.5099>.
- Soncin, F., Morey, R., Bui, T., Requena, D.F., Cheung, V.C., Kallol, S., Kittle, R., Jackson, M.G., Farah, O., Chousal, J., et al. (2022). Derivation of functional trophoblast stem cells from primed human pluripotent stem cells. *Stem Cell Rep.* 17, 1303–1317. <https://doi.org/10.1016/j.stemcr.2022.04.013>.
- Soom, A.V., Boerjan, M.L., Bols, P.E.J., Vanroose, G., Lein, A., Coryn, M., and Kruif, A. de (1997). Timing of Compaction and Inner Cell Allocation in Bovine Embryos Produced in Vivo after Superovulation. *Biol. Reprod.* 57, 1041–1049. <https://doi.org/10.1095/biolreprod57.5.1041>.
- Sozen, B., Jorgensen, V., Weatherbee, B.A.T., Chen, S., Zhu, M., and Zernicka-Goetz, M. (2021). Reconstructing aspects of human embryogenesis with pluripotent stem cells. *Nat. Commun.* 12, 5550. <https://doi.org/10.1038/s41467-021-25853-4>.
- Spencer, T.E., and Hansen, T.R. (2015). Implantation and Establishment of Pregnancy in Ruminants. In *Regulation of Implantation and Establishment of Pregnancy in Mammals: Tribute to 45 Year Anniversary of Roger V. Short's "Maternal Recognition of Pregnancy,"* R.D. Geisert, and F.W. Bazer, eds. (Cham: Springer International Publishing), pp. 105–135.
- Sperber, H., Mathieu, J., Wang, Y., Ferreccio, A., Hesson, J., Xu, Z., Fischer, K.A., Devi, A., Detraux, D., Gu, H., et al. (2015). The metabolome regulates the epigenetic landscape during naive-to-primed human embryonic stem cell transition. *Nat. Cell Biol.* 17, 1523–1535. <https://doi.org/10.1038/ncb3264>.
- Spindel, J., Krueger, C., Krueger, F., Papachristou, E.K., Kishore, K., D'Santos, C.S., and Reik, W. (2021). The distinct effects of MEK and GSK3 inhibition upon the methylome and transcriptome of mouse embryonic stem cells. 2021.11.18.469000. <https://doi.org/10.1101/2021.11.18.469000>.
- Stephens, P.C., Edwards, R.G., and Purdy, J.M. (1971). Human Blastocysts grown in Culture. *Nature* 229, 132–133. <https://doi.org/10.1038/229132a0>.
- Stirparo, G.G., Boroviak, T., Guo, G., Nichols, J., Smith, A., and Bertone, P. (2018). Correction: Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast (doi: 10.1242/dev.158501). *Development* 145, dev169672. <https://doi.org/10.1242/dev.169672>.
- Strawbridge, S.E., Clarke, J., Guo, G., and Nichols, J. (2022). Deriving Human Naïve Embryonic Stem Cell Lines from Donated Supernumerary Embryos Using Physical Distancing and Signal Inhibition. *Methods Mol. Biol. Clifton NJ* 2416, 1–12. https://doi.org/10.1007/978-1-0716-1908-7_1.
- Ström, S., Inzunza, J., Grinnemo, K.-H., Holmberg, K., Matilainen, E., Strömberg, A.-M., Blennow, E., and Hovatta, O. (2007). Mechanical isolation of the inner cell mass is effective in derivation of new human embryonic stem cell lines. *Hum. Reprod. Oxf. Engl.* 22, 3051–3058. <https://doi.org/10.1093/humrep/dem335>.
- Ström, S., Holm, F., Bergström, R., Strömberg, A.-M., and Hovatta, O. (2010). Derivation of 30 human embryonic stem cell lines—improving the quality. *In Vitro Cell. Dev. Biol. Anim.* 46, 337–344. <https://doi.org/10.1007/s11626-010-9308-0>.

Szczerbinska, I., Gonzales, K.A.U., Cukuroglu, E., Ramli, M.N.B., Lee, B.P.G., Tan, C.P., Wong, C.K., Rancati, G.I., Liang, H., Göke, J., et al. (2019). A Chemically Defined Feeder-free System for the Establishment and Maintenance of the Human Naive Pluripotent State. *Stem Cell Rep.* *13*, 612–626. <https://doi.org/10.1016/j.stemcr.2019.08.005>.

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *663–676*.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* *131*, 861–872. <https://doi.org/10.1016/j.cell.2007.11.019>.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., et al. (2014). Resetting Transcription Factor Control Circuitry toward Ground-State Pluripotency in Human. *Cell* *158*, 1254–1269. <https://doi.org/10.1016/j.cell.2014.08.029>.

Tan, T., Wu, J., Si, C., Dai, S., Zhang, Y., Sun, N., Zhang, E., Shao, H., Si, W., Yang, P., et al. (2021a). Chimeric contribution of human extended pluripotent stem cells to monkey embryos ex vivo. *Cell* *184*, 2020–2032.e14. <https://doi.org/10.1016/j.cell.2021.03.020>.

Tan, T., Wu, J., Si, C., Dai, S., Zhang, Y., Sun, N., Zhang, E., Shao, H., Si, W., Yang, P., et al. (2021b). Chimeric contribution of human extended pluripotent stem cells to monkey embryos ex vivo. *Cell* *184*, 2020–2032.e14. <https://doi.org/10.1016/j.cell.2021.03.020>.

Tanaka, S., Kunath, T., Hadjantonakis, A.-K., Nagy, A., and Rossant, J. (1998). Promotion of Trophoblast Stem Cell Proliferation by FGF4. *Science* *282*, 2072–2075. <https://doi.org/10.1126/science.282.5396.2072>.

Tarahomi, M., Vaz, F.M., Straalen, J.P. van, Schrauwen, F.A.P., Wely, M. van, Hamer, G., Repping, S., and Mastenbroek, S. (2019). The composition of human preimplantation embryo culture media and their stability during storage and culture. *Hum. Reprod.* *34*, 1450–1461. <https://doi.org/10.1093/humrep/dez102>.

Tarkowski, A.K. (1959). Experiments on the development of isolated blastomers of mouse eggs. *Nature* *184*, 1286–1287. <https://doi.org/10.1038/1841286a0>.

Tarkowski, A.K., and Wróblewska, J. (1967). Development of blastomeres of mouse eggs isolated at the 4- and 8-cell stage. *J. Embryol. Exp. Morphol.* *18*, 155–180.

Taubenschmid-Stowers, J., Rostovskaya, M., Santos, F., Ljung, S., Argelaguet, R., Krueger, F., Nichols, J., and Reik, W. (2022). 8C-like cells capture the human zygotic genome activation program in vitro. *Cell Stem Cell* *29*, 449–459.e6. <https://doi.org/10.1016/j.stem.2022.01.014>.

Telford, N.A., Watson, A.J., and Schultz, G.A. (1990). Transition from maternal to embryonic control in early mammalian development: A comparison of several species. *Mol. Reprod. Dev.* *26*, 90–100. <https://doi.org/10.1002/mrd.1080260113>.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D.G. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448, 196–199. <https://doi.org/10.1038/nature05972>.

Tesařek, J., Kopečný, V., Plachot, M., and Mandelbaum, J. (1988). Early morphological signs of embryonic genome expression in human preimplantation development as revealed by quantitative electron microscopy. *Dev. Biol.* 128, 15–20. [https://doi.org/10.1016/0012-1606\(88\)90261-8](https://doi.org/10.1016/0012-1606(88)90261-8).

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L., et al. (2014). Systematic Identification of Culture Conditions for Induction and Maintenance of Naive Human Pluripotency. *Cell Stem Cell* 15, 471–487. <https://doi.org/10.1016/j.stem.2014.07.002>.

Theunissen, T.W., Friedli, M., He, Y., Planet, E., O’Neil, R.C., Markoulaki, S., Pontis, J., Wang, H., Iouranova, A., Imbeault, M., et al. (2016). Molecular Criteria for Defining the Naive Human Pluripotent State. *Cell Stem Cell* 19, 502–515. <https://doi.org/10.1016/j.stem.2016.06.011>.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science* 282, 1145–1147. <https://doi.org/10.1126/science.282.5391.1145>.

Till, J.E., and McCULLOCH, E.A. (1961). A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiat. Res.* 14, 213–222. .

Till, J.E., and McCulloch, E.A. (1980). Hemopoietic stem cell differentiation. *Biochim. Biophys. Acta BBA - Rev. Cancer* 605, 431–459. [https://doi.org/10.1016/0304-419X\(80\)90009-8](https://doi.org/10.1016/0304-419X(80)90009-8).

Till, J.E., McCulloch, E.A., and Siminovitch, L. (1964). A STOCHASTIC MODEL OF STEM CELL PROLIFERATION, BASED ON THE GROWTH OF SPLEEN COLONY-FORMING CELLS. *Proc. Natl. Acad. Sci. U. S. A.* 51, 29–36. <https://doi.org/10.1073/pnas.51.1.29>.

Tosolini, M., and Jouneau, A. (2016). From Naive to Primed Pluripotency: In Vitro Conversion of Mouse Embryonic Stem Cells in Epiblast Stem Cells. *Methods Mol. Biol. Clifton NJ* 1341, 209–216. https://doi.org/10.1007/7651_2015_208.

Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K., and Niwa, H. (2008). Identification and characterization of subpopulations in undifferentiated ES cell culture. *Dev. Camb. Engl.* 135, 909–918. <https://doi.org/10.1242/dev.017400>.

Turco, M.Y., and Moffett, A. (2019). Development of the human placenta. *Dev. Camb. Engl.* 146, dev163428. <https://doi.org/10.1242/dev.163428>.

Turco, M.Y., Gardner, L., Kay, R.G., Hamilton, R.S., Prater, M., Hollinshead, M.S., McWhinnie, A., Esposito, L., Fernando, R., Skelton, H., et al. (2018). Trophoblast organoids as a model for maternal-fetal interactions during human placentation. *Nature* 564, 263–267. <https://doi.org/10.1038/s41586-018-0753-3>.

- Uysal, F., Kahveci, S., Sukur, G., and Cinar, O. (2022). Embryo culture media differentially alter DNA methylating enzymes and global DNA methylation in embryos and oocytes. *J. Mol. Histol.* *53*, 63–74. <https://doi.org/10.1007/s10735-021-10038-6>.
- Vallier, L., Alexander, M., and Pedersen, R.A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J. Cell Sci.* *118*, 4495–4509. <https://doi.org/10.1242/jcs.02553>.
- Vallot, C., Patrat, C., Collier, A.J., Huret, C., Casanova, M., Liyakat Ali, T.M., Tosolini, M., Frydman, N., Heard, E., Rugg-Gunn, P.J., et al. (2017). XACT Noncoding RNA Competes with XIST in the Control of X Chromosome Activity during Human Early Development. *Cell Stem Cell* *20*, 102–111. <https://doi.org/10.1016/j.stem.2016.10.014>.
- Van de Velde, H., Cauffman, G., Tournaye, H., Devroey, P., and Liebaers, I. (2008). The four blastomeres of a 4-cell stage human embryo are able to develop individually into blastocysts with inner cell mass and trophectoderm. *Hum. Reprod.* *23*, 1742–1747. <https://doi.org/10.1093/humrep/den190>.
- Vassena, R., Boué, S., González-Roca, E., Aran, B., Auer, H., Veiga, A., and Izpisua Belmonte, J.C. (2011). Waves of early transcriptional activation and pluripotency program initiation during human preimplantation development. *Dev. Camb. Engl.* *138*, 3699–3709. <https://doi.org/10.1242/dev.064741>.
- Veiga, A., Calderon, G., Barri, P.N., and Coroleu, B. (1987). Pregnancy after the replacement of a frozen-thawed embryo with >50% intact blastomeres. *Hum. Reprod.* *2*, 321–323. <https://doi.org/10.1093/oxfordjournals.humrep.a136542>.
- Visvader, J.E., and Clevers, H. (2016). Tissue-specific designs of stem cell hierarchies. *Nat. Cell Biol.* *18*, 349–355. <https://doi.org/10.1038/ncb3332>.
- Viukov, S., Shani, T., Bayerl, J., Aguilera-Castrejon, A., Oldak, B., Sheban, D., Tarazi, S., Stelzer, Y., Hanna, J.H., and Novershtern, N. (2022). Human primed and naïve PSCs are both able to differentiate into trophoblast stem cells. *Stem Cell Rep.* <https://doi.org/10.1016/j.stemcr.2022.09.008>.
- Vuyyuri, S.B., Rinkinen, J., Worden, E., Shim, H., Lee, S., and Davis, K.R. (2013). Ascorbic Acid and a Cytostatic Inhibitor of Glycolysis Synergistically Induce Apoptosis in Non-Small Cell Lung Cancer Cells. *PLoS ONE* *8*, e67081. <https://doi.org/10.1371/journal.pone.0067081>.
- Wang, X., Xiang, Y., Yu, Y., Wang, R., Zhang, Y., Xu, Q., Sun, H., Zhao, Z.-A., Jiang, X., Wang, X., et al. (2021). Formative pluripotent stem cells show features of epiblast cells poised for gastrulation. *Cell Res.* *31*, 526–541. <https://doi.org/10.1038/s41422-021-00477-x>.
- Ware, C.B., Nelson, A.M., Mecham, B., Hesson, J., Zhou, W., Jonlin, E.C., Jimenez-Caliani, A.J., Deng, X., Cavanaugh, C., Cook, S., et al. (2014). Derivation of naïve human embryonic stem cells. *Proc. Natl. Acad. Sci.* *111*, 4484–4489. <https://doi.org/10.1073/pnas.1319738111>.
- Watson, J., Baker, T., Bell, S., Gann, A., Levine, M., and Losick, R. (2013). *Molecular Biology of the Gene* (Boston Munich: Pearson).

Weatherbee, B.A.T., Gantner, C.W., Iwamoto-Stohl, L.K., Daza, R.M., Hamazaki, N., Shendure, J., and Zernicka-Goetz, M. (2023). Pluripotent stem cell-derived model of the post-implantation human embryo. *Nature* 622, 584–593. <https://doi.org/10.1038/s41586-023-06368-y>.

Weber, M., Knoefler, I., Schleussner, E., Markert, U.R., and Fitzgerald, J.S. (2013). HTR8/SVneo cells display trophoblast progenitor cell-like characteristics indicative of self-renewal, repopulation activity, and expression of “stemness-” associated transcription factors. *BioMed Res. Int.* 2013, 243649. <https://doi.org/10.1155/2013/243649>.

Wei, Y., Wang, T., Ma, L., Zhang, Y., Zhao, Y., Lye, K., Xiao, L., Chen, C., Wang, Z., Ma, Y., et al. (2021). Efficient derivation of human trophoblast stem cells from primed pluripotent stem cells. *Sci. Adv.* 7, eabf4416. <https://doi.org/10.1126/sciadv.abf4416>.

Weinberger, L., Ayyash, M., Novershtern, N., and Hanna, J.H. (2016). Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nat. Rev. Mol. Cell Biol.* 17, 155–169. <https://doi.org/10.1038/nrm.2015.28>.

Weissman, I.L. (2000). Stem Cells: Units of Development, Units of Regeneration, and Units in Evolution. *Cell* 100, 157–168. [https://doi.org/10.1016/S0092-8674\(00\)81692-X](https://doi.org/10.1016/S0092-8674(00)81692-X).

Wiley, L.M., and Eglitis, M.A. (1981). Cell surface and cytoskeletal elements: Cavitation in the mouse preimplantation embryo. *Dev. Biol.* 86, 493–501. [https://doi.org/10.1016/0012-1606\(81\)90207-4](https://doi.org/10.1016/0012-1606(81)90207-4).

Willadsen, S.M. (1986). Nuclear transplantation in sheep embryos. *Nature* 320, 63–65. <https://doi.org/10.1038/320063a0>.

Williams, B.S., and Biggers, J.D. (1990). Polar trophoblast (Raubert’s layer) of the rabbit blastocyst. *Anat. Rec.* 227, 211–222. <https://doi.org/10.1002/ar.1092270210>.

Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J., and Campbell, K.H. (1997). Viable offspring derived from fetal and adult mammalian cells. *Nature* 385, 810–813. <https://doi.org/10.1038/385810a0>.

Wilson, E.B., and Wilson, E.B. (1896). *The cell in development and inheritance* (New York: The Macmillan company).

Wimsatt, W.A. (1975). Some Comparative Aspects of Implantation. *Biol. Reprod.* 12, 1–40. <https://doi.org/10.1095/biolreprod12.1.1>.

Wooding, F.B.P. (1992). The synepitheliochorial placenta of ruminants: Binucleate cell fusions and hormone production. *Placenta* 13, 101–113. [https://doi.org/10.1016/0143-4004\(92\)90025-O](https://doi.org/10.1016/0143-4004(92)90025-O).

Wray, J., Kalkan, T., Gomez-Lopez, S., Eckardt, D., Cook, A., Kemler, R., and Smith, A. (2011). Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat. Cell Biol.* 13, 838–845. <https://doi.org/10.1038/ncb2267>.

Wu, A.M., Till, J.E., Siminovitch, L., and McCulloch, E.A. (1967). A cytological study of the capacity for differentiation of normal hemopoietic colony-forming cells. *J. Cell. Physiol.* 69, 177–184. <https://doi.org/10.1002/jcp.1040690208>.

- Wutz, A. (2012). Epigenetic alterations in human pluripotent stem cells: a tale of two cultures. *Cell Stem Cell* 11, 9–15. <https://doi.org/10.1016/j.stem.2012.06.012>.
- Xiao, S., Li, R., El Zowalaty, A.E., Diao, H., Zhao, F., Choi, Y., and Ye, X. (2017). Acidification of uterine epithelium during embryo implantation in mice. *Biol. Reprod.* 96, 232–243. <https://doi.org/10.1095/biolreprod.116.144451>.
- Xu, R.-H., Chen, X., Li, D.S., Li, R., Addicks, G.C., Glennon, C., Zwaka, T.P., and Thomson, J.A. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nat. Biotechnol.* 20, 1261–1264. <https://doi.org/10.1038/nbt761>.
- Xue, Z., Huang, K., Cai, C., Cai, L., Jiang, C., Feng, Y., Liu, Z., Zeng, Q., Cheng, L., Sun, Y.E., et al. (2013). Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* 500, 593–597. <https://doi.org/10.1038/nature12364>.
- Yamanaka, S. (2020). Pluripotent Stem Cell-Based Cell Therapy—Promise and Challenges. *Cell Stem Cell* 27, 523–531. <https://doi.org/10.1016/j.stem.2020.09.014>.
- Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J., et al. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* 20, 1131–1139. <https://doi.org/10.1038/nsmb.2660>.
- Yanagida, A., Spindlow, D., Nichols, J., Dattani, A., Smith, A., and Guo, G. (2021). Naive stem cell blastocyst model captures human embryo lineage segregation. *Cell Stem Cell* 28, 1016–1022.e4. <https://doi.org/10.1016/j.stem.2021.04.031>.
- Yang, J., Ryan, D.J., Wang, W., Tsang, J.C.-H., Lan, G., Masaki, H., Gao, X., Antunes, L., Yu, Y., Zhu, Z., et al. (2017a). Establishment of mouse expanded potential stem cells. *Nature* 550, 393–397. <https://doi.org/10.1038/nature24052>.
- Yang, Y., Liu, B., Xu, J., Wang, J., Wu, J., Shi, C., Xu, Y., Dong, J., Wang, C., Lai, W., et al. (2017b). Derivation of Pluripotent Stem Cells with In Vivo Embryonic and Extraembryonic Potency. *Cell* 169, 243–257.e25. <https://doi.org/10.1016/j.cell.2017.02.005>.
- Yang, Y., Liu, B., Xu, J., Wang, J., Wu, J., Shi, C., Xu, Y., Dong, J., Wang, C., Lai, W., et al. (2017c). Derivation of Pluripotent Stem Cells with In Vivo Embryonic and Extraembryonic Potency. *Cell* 169, 243–257.e25. <https://doi.org/10.1016/j.cell.2017.02.005>.
- Ying, Q.-L., and Smith, A. (2017). The Art of Capturing Pluripotency: Creating the Right Culture. *Stem Cell Rep.* 8, 1457–1464. <https://doi.org/10.1016/j.stemcr.2017.05.020>.
- Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature* 453, 519–523. <https://doi.org/10.1038/nature06968>.
- Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318, 1917–1920. <https://doi.org/10.1126/science.1151526>.

Yu, L., Wei, Y., Duan, J., Schmitz, D.A., Sakurai, M., Wang, L., Wang, K., Zhao, S., Hon, G.C., and Wu, J. (2021). Blastocyst-like structures generated from human pluripotent stem cells. *Nature* 591, 620–626. <https://doi.org/10.1038/s41586-021-03356-y>.

Yue, X., Wang, J., Chang, C., Liu, J., Yang, X., Zhou, F., Qiu, X., Bhatt, V., Guo, J.Y., Su, X., et al. (2022). Leukemia inhibitory factor drives glucose metabolic reprogramming to promote breast tumorigenesis. *Cell Death Dis.* 13, 1–12. <https://doi.org/10.1038/s41419-022-04820-x>.

Zheng, R., Geng, T., Wu, D.-Y., Zhang, T., He, H.-N., Du, H.-N., Zhang, D., Miao, Y.-L., and Jiang, W. (2021a). Derivation of feeder-free human extended pluripotent stem cells. *Stem Cell Rep.* 16, 1686–1696. <https://doi.org/10.1016/j.stemcr.2021.06.001>.

Zheng, R., Geng, T., Wu, D.-Y., Zhang, T., He, H.-N., Du, H.-N., Zhang, D., Miao, Y.-L., and Jiang, W. (2021b). Derivation of feeder-free human extended pluripotent stem cells. *Stem Cell Rep.* 16, 2410–2414. <https://doi.org/10.1016/j.stemcr.2021.07.019>.

Zheng, Y., Xue, X., Shao, Y., Wang, S., Esfahani, S.N., Li, Z., Muncie, J.M., Lakins, J.N., Weaver, V.M., Gumucio, D.L., et al. (2019). Controlled modelling of human epiblast and amnion development using stem cells. *Nature* 573, 421–425. <https://doi.org/10.1038/s41586-019-1535-2>.

Zhou, F., Wang, R., Yuan, P., Ren, Y., Mao, Y., Li, R., Lian, Y., Li, J., Wen, L., Yan, L., et al. (2019). Reconstituting the transcriptome and DNA methylome landscapes of human implantation. *Nature* 572, 660–664. <https://doi.org/10.1038/s41586-019-1500-0>.

Zhou, J., Hu, J., Wang, Y., and Gao, S. (2023). Induction and application of human naive pluripotency. *Cell Rep.* 42, 112379. <https://doi.org/10.1016/j.celrep.2023.112379>.

Zhu, P., Zhang, B., Sun, R., Wang, J., Liu, Z., Liu, X., Yan, M., Cui, Y., Sha, J., and Yuan, Y. (2023). Derivation of new pluripotent stem cells from human extended pluripotent stem cells with formative features and trophectoderm potential. *Cell Prolif.* e13480. <https://doi.org/10.1111/cpr.13480>.

Zorzan, I., Betto, R.M., Rossignoli, G., Arboit, M., Drusin, A., Corridori, C., Martini, P., and Martello, G. (2023). Chemical conversion of human conventional PSCs to TSCs following transient naive gene activation. *EMBO Rep.* 24, e55235. <https://doi.org/10.15252/embr.202255235>.

(2018). Pre-eclampsia.

(2021). LOI n° 2021-1017 du 2 août 2021 relative à la bioéthique (1).

(2023). 2020 National ART Summary | CDC.

HEW support of research involving human in vitro fertilization and embryo transfer.

Human Fertilisation: Warnock Report (Hansard, 31 October 1984).

Lei n° 11.105.

Titre : Caractérisation des modèles de cellules souches péri-implantatoire : une étape vers l'établissement de standards.

Mots clés : Cellules souches, Péri-implantation, Développement

Résumé : L'avènement de nouveaux modèles 2D et 3D pour le développement humain, notamment les cellules souches trophoblastiques, les gastruloïdes et les blastoïdes, a considérablement élargi les possibilités d'investigation des premiers événements du développement, éclairant progressivement le domaine énigmatique du développement humain. Si ces innovations ouvrent de nouvelles perspectives, il est devenu indispensable d'établir des référentiels bien définis pour les cellules sources de ces modèles. Cette thèse vise à proposer une caractérisation complète des modèles de cellules souches pluripotentes et trophoblastiques en employant une combinaison d'approches transcriptomiques, protéomiques, épigénétiques et métaboliques.

Nos résultats révèlent que les cellules souches pluripotentes étendues partagent de nombreuses caractéristiques avec les cellules souches pluripotentes amorcées, à l'exception de l'activité métabolique. Ce trait métabolique distinct peut expliquer leur capacité unique à se différencier directement en cellules souches trophoblastiques.

De plus, nos recherches démontrent que l'hypo-méthylation de l'ADN et une activité métabolique élevée définissent les cellules souches trophoblastiques. Ces résultats soulignent la nécessité de considérer plusieurs caractéristiques de la pluripotence, plutôt que de s'appuyer sur un seul critère. La multiplication des caractéristiques atténue les biais de correspondance des stades développementaux.

Title: Peri-implantation stem cell models characterization: a step toward the establishment of standards.

Keywords: Stem Cells, Peri-implantation, Development

Abstract: The advent of novel 2D and 3D models for human development, including trophoblast stem cells, gastruloids, and blastoids, has significantly expanded opportunities for investigating early developmental events, gradually illuminating the enigmatic realm of human development. While these innovations have ushered in new prospects, it has become essential to establish well-defined benchmarks for the cell sources of these models. This PhD aims to propose a comprehensive characterization of pluripotent and trophoblastic stem cell models by employing a combination of transcriptomic, proteomic, epigenetic, and metabolic approaches. Our findings reveal that extended pluripotent stem cells share many characteristics with primed pluripotent stem cells, with the exception of metabolic activity.

This distinct metabolic trait may account for their unique ability to directly differentiate into trophoblast stem cells.

Furthermore, our research demonstrates that DNA hypo-methylation and high metabolic activity define trophoblast stem cells. These results underscore the necessity of considering multiple hallmarks of pluripotency, rather than relying on single criteria. Multiplying hallmarks alleviate stage matching bias