

**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE PSIQUIATRIA**

VERÔNICA LUIZA VALE EUCLYDES COLOVATI

**Exposições ambientais pré-natal, mecanismos epigenéticos e
placentários como resposta e a relação com desfechos
de crescimento e neurodesenvolvimento**

**São Paulo
2020**

VERÔNICA LUIZA VALE EUCLYDES COLOVATI

**Exposições ambientais pré-natal, mecanismos epigenéticos e
placentários como resposta e a relação com desfecho
de crescimento e neurodesenvolvimento**

Tese apresentada à Faculdade de Medicina da
Universidade de São Paulo para obtenção do título
de Doutora em Ciências
Programa de Psiquiatria
Orientadora: Prof^a. Dr^a. Helena Paula Brentani

**São Paulo
2020**

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Assinatura: _____ Data ____/____/____

Catálogo na publicação
Serviço de Biblioteca e Documentação

Dados Internacionais de Catalogação na Publicação (CIP)

Preparada pela Biblioteca da
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Colovati, Verônica Luiza Vale Euclides
Exposições ambientais pré-natal, mecanismos epigenéticos e placentários como resposta e a relação com desfechos de crescimento e neurodesenvolvimento / Verônica Luiza Vale Euclides Colovati. -- São Paulo, 2020.
Tese(doutorado)--Faculdade de Medicina da Universidade de São Paulo.
Programa de Psiquiatria.
Orientadora: Helena Paula Brentani.

Descritores: 1.Epigenética 2.Efeitos tardios da exposição pré-natal 3.Desenvolvimento fetal 4.Metilação de DNA 5.Gravidez 6.Placenta 7.Células-tronco pluripotentes induzidas

USP/FM/DBD-209/20

Responsável: Erinalva da Conceição Batista, CRB-8 6755

Nome: Verônica Luiza Vale Euclides Colovati

Título: Exposições ambientais no período pré-natal, mecanismos epigenéticos e placentários como resposta e a relação com desfechos do crescimento e neurodesenvolvimento.

Tese apresentada ao Programa de Pós-Graduação em Psiquiatria da Faculdade de Medicina da Universidade de São Paulo para obtenção do título de Doutora em Ciências.

Aprovado em: ___/___/___

Banca Examinadora

Orientador: Prof. Dr. _____

Instituição: _____

Assinatura: _____

Prof. Dr _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr _____ Instituição: _____

Julgamento: _____ Assinatura: _____

DEDICATÓRIA

Aos meus amados pais, Luiz Carlos e Liete, os maiores apoiadores da minha vida. Aos meus irmãos e melhores amigos, Marcus (*in memoriam*) e Talita. Ao Fabrício, pelo amor e compreensão ao longo dessa jornada. Aos meus avós, pelo aconchego.

AGRADECIMENTOS

À orientadora, **Profª Drª Helena Brentani**, pelo brilhantismo técnico, por ser essa grande mestra e por conseguir fazer com que esse processo de crescimento acadêmico se tornasse mais leve. Minha eterna gratidão.

À **Profª. Drª. Jennifer Erwin**, por ter me recebido como aluna de doutorado sanduíche, no Lieber Institute for Brain Development (Johns Hopkins University). Agradeço ao grande aprendizado, acolhimento e estímulos frente aos novos desafios. À **toda equipe do laboratório Erwin-Paquola**, especialmente ao **Ethan**, por termos trabalhado diariamente e por tanta competência para cuidar dos nossos trofoblastos.

À toda equipe do LIM-23, vocês foram fundamentais para essa conquista. Aos companheiros do café (**Vanessa, Vinícius, Miriam, Victor, Catarina e Joana**), que faziam desses momentos antídotos quando alguma análise dava errado (frequente). Ao **Vinícius**, que além de um bioinformata excelente, se tornou um amigo. À **Vanessa**, pelos conselhos, orientação e almoços compartilhados. À **Ana Tahira**, por tantas dúvidas de bioinformática respondidas. À **Mariana Maschietto**, pela disponibilidade e orientação sempre que solicitada. À **Gisele**, por cuidar com toda competência das nossas amostras.

Aos meus amigos da pesquisa e vida gringa, **Arthur e André**. Obrigada por esses 12 meses de convivência diária, distante de nossas famílias, vocês foram meu suporte profissional e pessoal.

Aos meus tios e primos, por sempre estarem presentes. Somos uma unidade.

Aos meus sogros **Miriam** (*in memoriam*) e **Valdomiro**, pelos tantos almoços de domingo e pelo acolhimento em sua família. Aos cunhados e, especialmente, aos sobrinhos, por sempre trazerem alegria em nossos encontros.

Aos meus amigos, queridos, que acompanharam toda minha trajetória e a tornaram mais suave pelos tantos encontros: **Gerson, Jéssica, Carol, Dailey, Vanessa e Ruth**.

Ao **meu amado pai**, meu maior incentivador nessa jornada, meu grande amigo. À **minha querida mãe**, pelos sorrisos e positividade constantes. Vocês são verdadeiras joias raras. Grata a Deus, por ter me enviado para vocês.

Ao **meu irmão Marcus**, por toda leveza de viver, por me fazer sentir segura. Por ter assistido tantas apresentações de treinamento com tanta paciência e atenção. Obrigada pelo incentivo pessoal e financeiro para que eu pudesse dar início à minha carreira acadêmica. Sigo, mas faltando. À **minha irmã Tatá**, pelo humor perspicaz e por se alegrar com minhas alegrias e sofrer com minhas dores.

Ao **Fabrcício**, pelo amor construído ao longo desses anos e por compreender minha carreira e apoiar minha ausência para que eu realizasse uma meta profissional.

Aos meus **avós**, por tantos abraços calorosos, mesas fartas e apoio incondicional.

Às mulheres e filhos participantes desses estudos. Sem vocês, nada disso teria acontecido.

À **CNPQ**, pela concessão da bolsa de doutorado.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (**CAPES**) - Código de Financiamento 001

EPÍGRAFE

“Ninguém é tão sábio que nada tenha a aprender e nem tão ignorante que nada tenha a ensinar”

Blaise Pascal

SUMÁRIO

1. INTRODUÇÃO.....	21
2 OBJETIVOS.....	35
2.1 OBJETIVO GERAL.....	35
2.2 OBJETIVOS ESPECÍFICOS	35
2.2.1 Organização textual para resposta aos objetivos específicos	35
3 CAPÍTULO 3: ESTUDO ROC	37
3.1. MATERIAL E MÉTODOS ROC.....	37
3.1.1 Seleção Amostral.....	37
3.1.2 Aspectos Éticos.....	38
3.1.3 Avaliação das medidas de exposição a estresse pré-natal do estudo ROC	38
3.1.4 Dados de saúde e antropometria do recém-nascido da ROC.....	39
3.1.5 Coleta de amostra biológica (sangue do cordão umbilical - SCU) e extração do DNA da ROC.....	39
3.1.6 Análise de metilação do DNA da ROC	40
3.1.7 Análise de Bioinformática	40
3.1.6 Análise do Relógio Epigenético	42
3.1.7 Redes de metilação diferencial	42
3.1.8 Análises estatísticas da ROC	44
3.2 RESULTADOS DA ROC	46
3.2.1 Caracterização da Amostra da Coorte de Nascimento da Região Oeste (ROC)	46
3.2.2 Relógio epigenético (RE) e medidas de crescimento ao nascimento	48
3.2.4 Exposição a estresse pré-natal, RE e medidas de crescimento ao nascimento.....	50
3.2.5 Criação de grupos a partir do escore de estressores ambientais	52
3.2.6 Relógio epigenético e efeito mediador entre grupos de exposição ao estresse e crescimento ao nascimento	56
3.2.7 Análise de redes de metilação diferencial entre grupos A e B	58
3.3 DISCUSSÃO DOS DADOS DA ROC	62
4 CAPÍTULO 4: ESTUDO DE INTERVENÇÃO: ENSAIO CLÍNICO RANDOMIZADO PRIMEIROS LAÇOS.....	66
4.1 MATERIAL E MÉTODOS.....	66
4.1.1 Desenho do Ensaio Clínico Randomizado (RCT).....	66
4.1.2 Aspectos éticos	67
4.1.3 Seleção amostral: Coleta de sangue de cordão.....	67
4.1.4 Avaliação das medidas de exposição a estresse pré-natal do estudo Primeiros Laços	68

4.1.5 Coleta de amostra biológica (sangue do cordão umbilical - SCU) e extração do DNA.....	69
4.1.6 Análise de metilação.....	69
4.1.7 Análise de metilação diferencial.....	69
4.1.8 Medidas antropométricas maternas e da criança	70
4.1.9 Medidas de neurodesenvolvimento aos 12 meses de vida	70
4.1.10 A análise de rede de correlação ponderada – WGCNA	71
4.1.11 Análise de enriquecimento	71
4.1.12 Relógio epigenético para idade gestacional no nascimento	72
4.1.13 Análise Estatística.....	72
4.2 RESULTADOS PRIMEIROS LAÇOS	73
4.2.1 Características da amostra do Ensaio Clínico Randomizado “Primeiros Laços”	73
4.2.2 Identificação de diferenças de metilação entre intervenção e controles.....	78
4.2.3 Relógio Epigenético nos grupos Intervenção e Controle	84
4.2.4 Metilação do DNA como mediadora na cognição em 12 meses.....	85
4.3 DISCUSSÃO DADOS DO RCT PRIMEIROS LAÇOS	91
5 CAPÍTULO 5: PLACENTA	94
5.1 CASUÍSTICA.....	94
5.2 MÉTODOS	97
5.2.1 Cultura das células.....	97
5.2.2 Diferenciações e manutenção de células	97
5.2.3 Análises laboratoriais	98
5.2.4 Pré-processamento de dados Drop-seq.....	98
5.2.5 Análise de clusterização e expressão diferencial de genes para genes marcadores	98
5.2.6 Reconstrução de Trajetória.....	100
5.2.7 Análise de enriquecimento dos agrupamentos para genes específicos de células da placenta.....	101
5.2.8 Redes regulatórias PANDA.....	101
5.3 RESULTADOS DO CAPÍTULO 5 PLACENTA.....	102
5.3.1 Derivação de células-tronco trofoblásticas de células-tronco embrionárias humanas preparadas e células-tronco pluripotentes induzidas sem BMP4	102
5.3.2 Sequenciamento de RNA de <i>Single-cell</i> revela programação transcricional exclusiva para especificação de TS derivada de iPS	107
5.4 DISCUSSÃO PLACENTA	120
6 CONSIDERAÇÕES FINAIS	123
7 REFERÊNCIAS	124

RESUMO

Colovati VLVE. *Exposições ambientais pré-natal, mecanismos epigenéticos e placentários como resposta e a relação com desfechos de crescimento e neurodesenvolvimento* [tese]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2020.

A gestação é reconhecidamente um período de alta vulnerabilidade para o sistema nervoso central. Exposições de estresse ambiental (fumo, estresse psicossocial, tóxicos, etc.) têm sido associadas a comprometimento do crescimento à alterações e na trajetória do neurodesenvolvimento. Desse modo, intervenções na gestação têm reportado resultados promissores. Mecanismos epigenéticos, como a metilação do DNA (DNAm), estão associados tanto a exposição aos estressores gestacionais como a desfechos neonatais e vulnerabilidade para transtornos mentais. Nesse sentido, é fundamental considerar o órgão interlocutor entre mãe-feto e ambiente, a placenta. A complexidade e as especificidades celulares da placenta humana são pouco conhecidas por falta de modelos de cultura, dessa maneira, essa tese conta com 3 objetivos principais: 1) Avaliar se alterações de metilação do DNA poderiam mediar a associação entre exposição a estresse intra-útero e desfechos neonatais de crescimento, 2) Avaliar se alterações no padrão de metilação do DNA poderiam mediar a associação entre uma intervenção gestacional precoce e desfechos do neurodesenvolvimento e 3) Avaliar modelo de cultura de células-tronco de pluripotência induzidas (do inglês, *induced pluripotent stem cell*, iPS) para diferenciação em trofoblastos para compreensão das vias relacionadas ao início do desenvolvimento placentário. Microarranjos da *Illumina Methylation BeadChip 450k*, foram usados nas amostras de sangue do cordão umbilical (SCU) e analisados pelos pacotes Minfi e ChAMP. O relógio epigenético (RE) estimou a idade gestacional (IG) a partir da metilação em sítios específicos (Knight e col., 2016), e, as diferenças entre o RE e IG (Δ idade), foram obtidas para avaliar uma possível aceleração ou desaceleração frente a IG. Para responder ao objetivo 1, avaliamos dados de estresse gestacional, DNAm e as medidas antropométricas ao nascimento (inseridas no *WHO Anthro Survey Analyzer*), coletadas no estudo transversal ROC (n=83). Essas medidas foram interpretadas como *proxy* de qualidade de vida fetal. Observamos que a aceleração do RE (RE<IG) associou a menores índices de crescimento. A exposição a diferentes fatores de risco deu origem a um escore de exposição em escala crescente. Separando as mães em grupo de maior e menor exposição a estressores, mães com os maiores escores de estresse tinham crianças menores, contudo, essa associação não foi mediada pelo RE. Para responder ao objetivo 2, avaliamos dados de uma subamostra com coleta do SCU (n=29) de um Ensaio Clínico Randomizado, que propõe uma intervenção precoce na gestação, por visitas domiciliares de enfermeiras a gestantes de alta vulnerabilidade social. As medidas de

neurodesenvolvimento foram obtidas aos 12 meses de vida. Tanto sítios específicos, quanto a região diferencialmente metilada do *PF4* mediaram a associação entre a intervenção no período gestacional e o domínio cognitivo das crianças aos 12 meses. Para responder ao objetivo 3, devido às limitações de se obter células placentárias humanas no início da sua diferenciação, desenvolvemos um modelo de cultura de iPS não-dependente de BMP4 que se assemelha com células trofoblásticas, especialmente o citotrofoblasto. Concluimos que, as situações estressoras estão associadas a padrões de DNAm, e que podem ser interpretados como mediadores ou marcadores de diferentes exposições ambientais. Ademais, acreditamos que algumas lacunas podem ser preenchidas na integração das exposições ambientais, DNAm e mecanismos placentários com o novo modelo de cultura de iPS desenvolvido, especialmente associados aos comprometimentos da diferenciação de células trofoblásticas.

DESCRITORES: Epigenética; Efeitos tardios da exposição pré-natal; Desenvolvimento fetal; Metilação de DNA; Gravidez; Placenta; Células-tronco pluripotentes induzidas.

ABSTRACT

Colovati VLVE. *Prenatal environmental exposures, epigenetic and placental mechanisms as a response and the relationship with growth and neurodevelopmental outcomes* [thesis]. São Paulo: “Faculdade de Medicina, Universidade de São Paulo”; 2020.

Pregnancy is known to be a period of high vulnerability for the central nervous system development. Exposure to environmental stress (i.e, smoking, psychosocial, toxic stress) have been linked to impairment of neurodevelopment trajectory and growth. Thus, interventions in pregnancy have reported promising results. Epigenetic mechanisms, such as DNA methylation (DNAm), are associated with exposure to gestational stressors, neonatal outcomes and vulnerability to mental disorders. The hypothesis to be tested in this thesis is that changes in DNAm mediate the relationship between gestational environmental exposure and growth/neurodevelopmental outcomes. Furthermore, it is essential to consider the very beginning of the placenta differentiation in order to understand the organ that is the first to receive stress signals and assure the crosstalk between mother and fetus. The complexity and single-cell specificities of the early human placenta are not fully clarified due to the lack of reliable trophoblast culture models. Thus, this thesis has 3 main goals: 1) Verify whether changes in DNAm could mediate the association of intrauterine stress exposure and neonatal growth outcomes, 2) Assess if changes in the DNAm could mediate the association of an early gestational intervention and neurodevelopmental outcomes and, 3) Evaluate a new model of induced pluripotent stem cell culture (iPS) for differentiation into trophoblasts in order to clarify the pathways related to the onset of placental development. Genome-wide DNAm from cord blood samples was assessed using the Illumina BeadChip 450K array and analyzed by the Minfi and ChAMP packages. The epigenetic clock (EC) estimated gestational age (GA) from methylation at specific sites (Knight et al., 2016), and the differences between RE and GA (Δ age) were obtained to assess possible acceleration or deceleration compared to GA. To answer objective 1, in the cross-sectional ROC study (n = 83), we assessed data on gestational stress, DNAm and anthropometric measurements at birth (WHO Anthro Survey Analyzer software), which were interpreted as a proxy for intrauterine environment. We observed that the slowdown in the ER (RE < IG) was associated with higher growth rates. Exposure to different risk factors gave rise to an increasing exposure score. Separating mothers into groups of higher and lower exposure to stressors, mothers with the highest stress scores had smaller newborns, however, this association was not mediated by the EC. To answer objective 2, we evaluated data from a

subsample with cord blood collection (n = 29) from a Randomized Clinical Trial, which proposes an early intervention in pregnancy, through home visits by nurses to pregnant women with high social vulnerability. Neurodevelopment measures were obtained at 12 months of age. Both differentially methylated sites and the region of PF4 mediated the association between the intervention and the cognitive domain at 12 months. To answer objective 3, due to the gaps in obtaining human placental cells at the very beginning of their differentiation, we developed a non-BMP4-dependent iPS culture model that resembles trophoblastic cells, especially, the cytotrophoblast. We conclude that stressful situations are associated with DNAm variation, and that it can be interpreted as mediators or markers of different environmental exposures. In addition, we believe some gaps can be filled in the integration of environmental exposures, DNAm and placental mechanisms with the new iPS culture model developed, especially associated with the impairment of differentiation of trophoblastic cells.

DESCRIPTORS: Epigenetics; Prenatal exposure delayed effects; Fetal development; DNA methylation; Pregnancy; Placenta; Induced pluripotent stem cells.

LISTA DE FIGURAS

Figura 1 - Exposições a diferentes fatores estressores têm sido associados a desfechos na saúde em fases posteriores da vida, e a metilação do DNA, pode ser um dos mecanismos que explicam essas associações descritas.	24
Figura 2 - A placenta e resposta ao estresse. O Estresse crônico oriundos de diferentes exposições está associado com aumento de cortisol circulante que, por sua vez, altera a expressão da 11- β -hidroxiesteróide desidrogenase tipo 2 (11 β HSD2) com diminuição do controle de entrada do cortisol circulante e possível comprometimento de crescimento e neurodesenvolvimento.	31
Figura 3 - Fluxograma da apresentação desta tese com suas subdivisões em 3 capítulos que respondem os objetivos específicos 1, 2 e 3.	36
Figura 4 - Fluxograma das amostras utilizadas do estudo ROC.	38
Figura 5 - Fluxograma das análises do estudo ROC (n=83). Dados brutos (idat) foram processados e seguiram para as etapas de controle de qualidade e normalização (n=428.154). Em seguida, o relógio epigenético foi obtido pelo algoritmo de Knight e cols. (2016) e realizada a análise de metilação diferencial entre os grupos criados a partir da exposição ao estresse. .	47
Figura 6 - Gráfico de dispersão de regressão linear da variável dependente relógio epigenético e IG clínica ao nascimento, como variável independente, pelo método Capurro (p-valor=0,019).	48
Figura 7 - Gráfico de dispersão gerado de regressão linear tendo como variável independente o Δ idade (RE-IG) e de independente os índices A) IMC/Idade, B) Peso/Idade, C) Comprimento/Idade.	50
Figura 8 - Gráfico de dispersão do escore do estresse materno e A) Delta idade (p = 0,59), B) RE (p=0,930)	51
Figura 9 - <i>Boxplot</i> dos índices de crescimento (OMS, 2016), segundo os grupos criados a partir da mediana dos fatores estressores ambientais. A) Peso/Idade, B) IMC/Idade (p=0,014), C) Comprimento/Idade.	53
Figura 10 - <i>Boxplot</i> do índice de IMC/I (OMS, 2016), segundo os grupos criados a partir da mediana dos fatores estressores ambientais do estudo ROC. A) Peso/Idade, B) IMC/Idade (p=0,014), C) Comprimento/Idade.	54
Figura 11 - <i>Boxplot</i> dos índices de crescimento e circunferência abdominal de acordo com a exposição ao fumo na gestação. A) P/I, sendo peso sobre a idade, B) IMC/I, sendo IMC por idade C) C/I, sendo comprimento por idade D) Circunferência abdominal.	55
Figura 12 - <i>Boxplot</i> das medidas de RE e IG das mães fumantes e não-fumantes. A) Delta idade, B) RE em semanas, C) Idade gestacional.	56
Figura 13 - Figura esquemática da análise de mediação em que investigou-se o efeito mediador do RE e RE-IG (delta idade) sobre os índices de crescimento IMC/I, P/I, C/I. Ver Tabela 5 para coeficientes.	57

- Figura 14** - Ilustração gráfica do maior Módulo Epigenético Funcional identificado: UBE2I. As cores variam entre o estado menos metilado (amarelo) ao mais metilado (azul). 58
- Figura 15** - Conservação dos módulos do WGCNA entre os grupos A e B. A) Preservação dos módulos em ordem decrescente, de acordo com a métrica medianRank. B) Classificação dos módulos a partir da métrica Zsummary.o threshold = 10 de preservação. Nas duas subfiguras, o eixo x mostra o número de sítios pertencentes a cada módulo, e no eixo y, o score de preservação, de acordo com o medianRank e Zsummary. Em relação ao Zsummary, módulo com um score ≤ 10 são considerados pouco preservados. 60
- Figura 16** - Seleção da amostra desta tese, proveniente do Ensaio Clínico Randomizado Primeiros Laços. 68
- Figura 17**- Fluxograma do estudo com as amostras do Ensaio Clínico Randomizado, Primeiros Laços. Amostras de sangue do cordão umbilical (n = 32) foram coletadas dos participantes do Primeiros Laços (n=80). A metilação do genoma foi aplicada usando as matrizes *Infinium HumanMethylation 450k BeadChip* que foi submetida às etapas de controle de qualidade, normalização e correção para efeitos de lote. Os parâmetros do desenvolvimento neurológico foram obtidos pela avaliação de Escala de desenvolvimento Bayley III. A análise final para resposta da pergunta foi a mediação causal, tendo como variável desfecho o escore bruto do domínio cognitivo, metilação do DNA como variável mediadora e a intervenção como a variável exploratória. 77
- Figura 18** - Gráficos de Vulcão mostrando os níveis de metilação de sítios CpG nos grupos de intervenção e controle. Os eixos X e Y exibem, respectivamente, o valor delta-beta (diferenças de metilação entre os grupos) e o log dos valores p ajustados para cada sítio CpG. Pontos verdes (hipometilados) e vermelhos (hipermetilados) identificam os locais CpG com $\text{adj } P < 0,05$ e $> 5\%$ de diferença de metilação entre os grupos de intervenção e controle. 78
- Figura 19** - Análise de Preservação do Módulo. A) *Preservation Median Rank* (esquerda) e *Preservation Zsummary* (direita) dos módulos com base nos 3.090 DMPs fornecidos por WGCNA. O eixo X mostra o tamanho (número de CpGs) e o eixo Y mostra, respectivamente, a classificação mediana e as pontuações do resumo Z. Módulos com *Zsummary* ≤ 10 foram considerados menos preservados. Módulos com pontuações mais altas de *Median Rank* são aqueles com baixa preservação de acordo com esta métrica. Em relação ao *Median Rank*, o MEbrown teve exatamente a mesma pontuação do MEyellow (3) com um tamanho de módulo muito semelhante (138 CpGs e 136 CpGs), respectivamente. O gráfico de dispersão em B) e C) mostra a correlação de Pearson entre cada valor M de CpGs com seu respectivo módulo eigengene, para os grupos controle e intervenção. As amostras com a cor vermelho são as que apresentam as maiores diferenças de correlação de Pearson entre os dois grupos..... 83
- Figura 20** - A) Boxplot da Idade Gestacional (IG) entre intervenção e controle; B) *Boxplot* do relógio epigenético para a AG ($p=0,038$), C) Aceleração de Idade ($p=0,025$), D) Gráfico de dispersão do relógio epigenético e idade cronológica 84
- Figura 21** - Fluxograma das etapas que precederam a definição das variáveis de metilação do DNA para avaliação de efeito mediador entre a relação da intervenção de visitas domiciliares

realizadas por enfermeiras a gestantes adolescentes e o domínio cognitivo aos 12 meses de vida.	85
Figura 22 - Análise de sensibilidade para a mediação da DMR mapeada no gene PF4.	90
Figura 23 - Fluxograma da seleção de células únicas até análise de expressão diferencial	99
Figura 24 - Derivação de células trofoblásticas pluripotentes (TSC) a partir de células-tronco pluripotentes induzidas.	103
Figura 25 - Caracterização das linhagens celulares iPS. (A-C) Imagens imunofluorescentes da linha HiPS LIBD1C8 coradas para Sox2 (A), NANOG (B) e TRA-1-60 (C). (D-F) Imagens imunofluorescentes da linha HiPS 2014.06 coradas para Sox2 (D), NANOG (E) e TRA-1-60 (F). (G) Níveis de expressão relativa de pluripotência e marcadores específicos da camada germinativa em células espontaneamente diferenciadas da linha 2014.06 normalizados para seus respectivos níveis de expressão em HiPS indiferenciados da mesma linha celular determinada por qPCR. (H) Análise do cariótipo de 2014.06 Todas as barras de escala 50 um.	104
Figura 26 - As células TS derivadas de iPS se diferenciam em citotrofoblasto sincicial, citotrofoblastos extravilosos e estruturas semelhantes a vilosidades.	106
Figura 27 - Sequenciamento (RNAseq) temporal de <i>Single-Cell de iPS</i> até a diferenciação em TB.	108
Figura 28 - As diferenciações de trofoblasto estão enriquecidas para os tipos de células da placenta do primeiro trimestre. (A, B) Análise de enriquecimento de conjunto de genes MAGMA de diferenciações de células HiPS destacando o enriquecimento de agrupamentos de células para tipos de células identificados em placentas do primeiro trimestre. (C) Mapa térmico composto de expressão gênica normalizada dos principais marcadores únicos de tipos de células da placenta de origem fetal identificados em Vento-Tormo et al. 2018 e marcadores do trofotoderma polar (top 50) identificados em Petropulus et al. 2016	112
Figura 29 - Especificação iPS para células-tronco trofoblásticas inicia com reconexão TFAP2A sem ativação de programas de células-tronco ingênuas.	115
Figura 30 - A análise STREAM de pseudotempo das diferenciações trofoblásticas revela a suprarregulação do mesênquima estromal após o tratamento com BMP4 e a programação transcricional do trofotoderma no tratamento TS.	117
Figura 31 - ZBTB7C e retrovírus endógenos são centrais e exclusivos da rede reguladora do gene TS.	119

LISTA DE TABELAS

Tabela 1	–	Características socioeconômicas, obstétricas, clínicas e bioquímicas, de acordo com o sexo.....	46
Tabela 2	–	ROC: Associação de medidas ao nascimento com relógio epigenético (variável de interesse).....	49
Tabela 3	–	Associação de medidas ao nascimento com a somatória de estressores da ROC.....	51
Tabela 4	–	Características socioeconômicas, obstétricas, clínicas e bioquímicas, de acordo com o grupo de exposição ao estresse (n=83).....	52
Tabela 5	–	Análise de mediação entre os grupos de exposição a estresse A e B, relógio epigenético e medidas de crescimento.....	57
Tabela 6	–	Módulos Epigenéticos Funcionais (Redes de interactoma).....	59
Tabela 7	–	Tabela de enriquecimento (STRING) dos genes mais variáveis do módulo magenta da análise de co-metilação (WGCNA).....	61
Tabela 8	–	Características psíquicas, sociais e escala de neurodesenvolvimento das participantes do RCT e respectivos filhos com coleta de sangue do cordão umbilical. São Paulo, SP.....	74
Tabela 9	–	Tabela comparativa das características de mães e respectivos filhos que tiveram coleta do cordão umbilical e análise de metilação e neurodesenvolvimento aos 12 meses com a amostra inicial.....	75
Tabela 10	–	Tabela comparativa com a amostra inteira do RCT (n final =63).....	76
Tabela 11	–	Genes com maior Delta Beta (>20%) associados ao neurodesenvolvimento.....	79
Tabela 12	–	Identificação das DMPs associadas ao domínio cognitivo (R ajustado>0,4).....	80
Tabela 13	–	Identificação das DMR, média dos valores de M e associação com o domínio cognitivo.....	81
Tabela 14	–	Análise de mediação com todas as DMPs selecionadas, DMRs e RE.....	87
Tabela 15	–	Análise de mediação apenas com as variáveis de metilação de tiveram ACME (p-valor<0,05) e respectivos testes de sensibilidade (Rho).....	89

LISTA DE ABREVIATURAS E SIGLAS

ACME. Efeito de Mediação Causal

ADE. Efeito Direto

BDNF. Fator neurotrófico derivado do cérebro, do inglês *brain derived neurotrophic factors*

BMP4. Proteína Morfogenética Óssea 4

BPA. Bisfenol A

BPN. Baixo peso ao nascimento

C/I. Comprimento sobre idade

CpG. Ilhas de Citosina que precede uma Guanina

DCNTs. Doenças crônicas não-transmissíveis

DHSs. Hipersensibilidade a DNase, do inglês *DNase I hypersensitive sites*

DMP. Posições Diferencialmente Metiladas, do inglês *Differentially Methylated Positions*

DMR. Regiões Diferencialmente Metilados, do inglês *Differentially Methylated Regions*

DNMT. DNAmetiltransferases

DOHaD. Origem Desenvolvimentista da Saúde e da Doença

DP. Desvio Padrão

EGF. Fator de crescimento epidérmico, do inglês *epidermal growth factor*

FDR: Taxa de falsa descoberta

HDAC. Deacetilase de histona

HPA. Eixo hipotalâmico pituitário-adrenal

HU. Hospital universitário

IDAT. Extensão de arquivo obtido da leitura dos chips de metilação

IG. Idade Gestacional

IGR. Região intergênica

IMC. Índice de Massa Corpórea

IMC/I. Índice de Massa Corpórea sobre idade

iPS. Célula-tronco pluripotente induzida

N_Shore. Até 2000 pares de base a montante da Ilha CpG

N_Shelf. Entre 2000 a 4000 pares de base a montante da Ilha CpG

NGF. Fator de Crescimento Nervoso, do inglês, *nerve growth factor*

OMS. Organização Mundial de Saúde

Open Sea. Região fora de margens e Ilha CpG

Pb. Pares de base

PCA. Análise de Componentes Principais, do inglês *Principal Component Analysis*

PCBs. Bifenilos Policlorados

PCR: Reação em cadeia da polimerase

P/I. Peso sobre Idade

PN. Peso ao nascer

RCIU. Restrição de Crescimento Intrauterino

RE. Relógio Epigenético

RNA. Ácido ribonucleico

ROCK. Proteína quinase associada a Rho

SCU. Sangue do Cordão Umbilical

SNC. Sistema Nervoso Central

SNPs. Polimorfismos de Nucleotídeo Único

STRING. Search Tool of Retrieval of INteracting Genes

TEA. Transtorno do Espectro Autista

TDAH. Transtorno do Déficit de Atenção com Hiperatividade

TFBSs . Sítios de ligação do fator de transcrição, do inglês *intranscription factor binding site*

TGFb. Inibição do fator de crescimento beta, do inglês *Transforming Growth Factor Beta 3*

TND. Transtornos do neurodesenvolvimento

TOM. Medida de sobreposição topológica

TSC. Células-tronco pluripotentes de trofoblasto

TSS1500. 1500 pares de base do sítio de início de transcrição

TSS200. 200 pares de base do sítio de início de transcrição

UMAP. Aproximação e Projeção Uniforme de Variedades, do inglês *Uniform Manifold Approximation and Projection*

LISTA DE SÍMBOLOS

Δ . Delta

$\Delta\beta$. Diferenças dos valores de beta

Δ idade. Diferenças entre RE e IG

1. INTRODUÇÃO

Estudos epidemiológicos permitiram a observação de que experiências vivenciadas em períodos do início do desenvolvimento humano, como a vida pré-natal, explicam, em parte, o processo saúde-doença ao longo da vida (1–3). O zigoto se torna um feto com completo crescimento e desenvolvimento para a vida extrauterina em aproximadamente 40 semanas. O crescimento e desenvolvimento são exponenciais, e devido à velocidade de divisão celular, o estágio de vida fetal representa a maior taxa de crescimento humana. Nesse período intrauterino, mãe e feto, conectados pela placenta, estão vulneráveis a fatores ambientais, tais como dieta, estresse psicossocial e tóxicos, que podem causar desequilíbrios metabólicos, implicando em consequências duradouras sobre a fisiologia e estrutura do organismo (4,5). Essa vulnerabilidade pré-natal é evidenciada em modelos animais, laboratoriais e epidemiológicos que serão apresentadas ao longo desta introdução, assim como mecanismos relacionados a essas associações entre exposições maternas e desfechos para a prole. No entanto, aos mecanismos moleculares relacionados entre a exposições e desfechos ao longo da vida ainda precisam de mais esclarecimentos. Já no início do desenvolvimento embrionário o processo de diferenciação do Sistema Nervoso Central (SNC) começa a ser observado, ocorrendo um aumento da complexidade no decorrer da evolução da gestação. Nesse sentido, essa intensa plasticidade também é acompanhada por maior vulnerabilidade.

Um estudo clássico, que foi um marco para a medicina fetal, ocorreu na década de 80, a partir de dados retrospectivos de indivíduos que, durante a vida intrauterina, sofreram privação energética (em virtude da 2ª Guerra Mundial, período conhecido como Fome Holandesa). Os filhos de mulheres que não receberam a nutrição que deveriam, e apesar de não mensurado, acrescido ao potencial estresse psicossocial inerentes a uma guerra, apresentaram maior risco para o desenvolvimento doenças crônicas e psiquiátricas(1,6,7). Esses estudos deram origem ao ramo da ciência denominada Origem Desenvolvimentista da Saúde e da Doença (DOHaD), que têm evidenciado as exposições ambientais, como nutrição materna, abuso de álcool/drogas, fumo e estresse psicossocial materno como fatores de risco para doenças crônicas e transtornos do neurodesenvolvimento (TND) (8–10), sendo este último de principal interesse desta tese. O neurodesenvolvimento se caracteriza pelo desenvolvimento, crescimento e maturação do SNC e é altamente dispendioso ao metabolismo fetal. O SNC tem

considerável demanda energética e depende do ambiente proporcionado que, se inadequado, pode ser um fator limitante ao neurodesenvolvimento (11). Ainda há muito a ser compreendido concernente às vias de resposta que afetam a “programação” fenotípica em decorrência a estímulos ambientais, entretanto, alguns mecanismos têm sido propostos nesse sentido.

Para discutir possíveis mecanismos associados ao neurodesenvolvimento não explicados pela genética, é interessante percorrer as descobertas previamente descritas que ajudam a compreender onde o conhecimento científico genético se situa hoje. Há algumas décadas, o biólogo Conrad Hal Waddington ficou particularmente intrigado em desvendar como células que continham o mesmo material genético tinham destinos/diferenciação diferentes. Suas pesquisas mostraram que existia outro fator além do que hoje é conhecido como genoma humano, que estava relacionado ao fenômeno da alteração de fenótipo sem que houvesse mudanças no código genético, descrevendo hoje o que se conhece como epigenética (12). Apenas meia década depois do surgimento de uma definição vaga do que é epigenética, esse campo da biologia molecular veio amadurecendo, caracterizando-se pela dinamicidade e revisando paradigmas de herdabilidade. A definição mais atual de eventos epigenéticos refere-se a mecanismos moleculares que alteram a cromatina e interferem na expressão gênica, permitindo assim, mudanças no fenótipo, sem alteração do genótipo (13). Desse modo, as modificações de histonas, metilação do DNA, e ação dos microRNA englobam os mecanismos epigenéticos. Padrões epigenéticos diferentes foram identificados em estudos com gêmeos monozigóticos que apresentaram diferentes desfechos de saúde, reforçando assim que, além do componente genético, existe uma influência ambiental na definição do processo saúde-doença (14,15).

As modificações epigenéticas não estão associadas apenas com o processo saúde-doença, mas a partir do zigoto ocorre uma desmetilação do genoma, com exceção dos genes imprintados (fenômeno no qual certos genes são expressos apenas por um alelo, paterno ou materno, enquanto o outro é metilado) e *transposons* (podem realocar ou introduzir cópias de si mesmos em um local diferente no genoma), e uma reprogramação epigenética caracteriza e garante o desenvolvimento tecidos específicos. As modificações epigenéticas somáticas permanecem à medida que as células se dividem (e portanto, são herdadas mitoticamente). Durante o desenvolvimento um grupo especial de células migra dando origem às futuras gônadas (16). Aqui o processo de remetilação é diferente e, alterações podem ser passadas para próxima geração (17,18). Como dito, marcas epigenéticas variam de acordo com os tecidos observados, assim, alterações observadas em um tecido não necessariamente refletem as de

outros tecidos. Portanto, esses mecanismos que ocorrem no nível molecular parecem ser uma ponte que ligaria o meio ambiente a alterações fenotípicas.

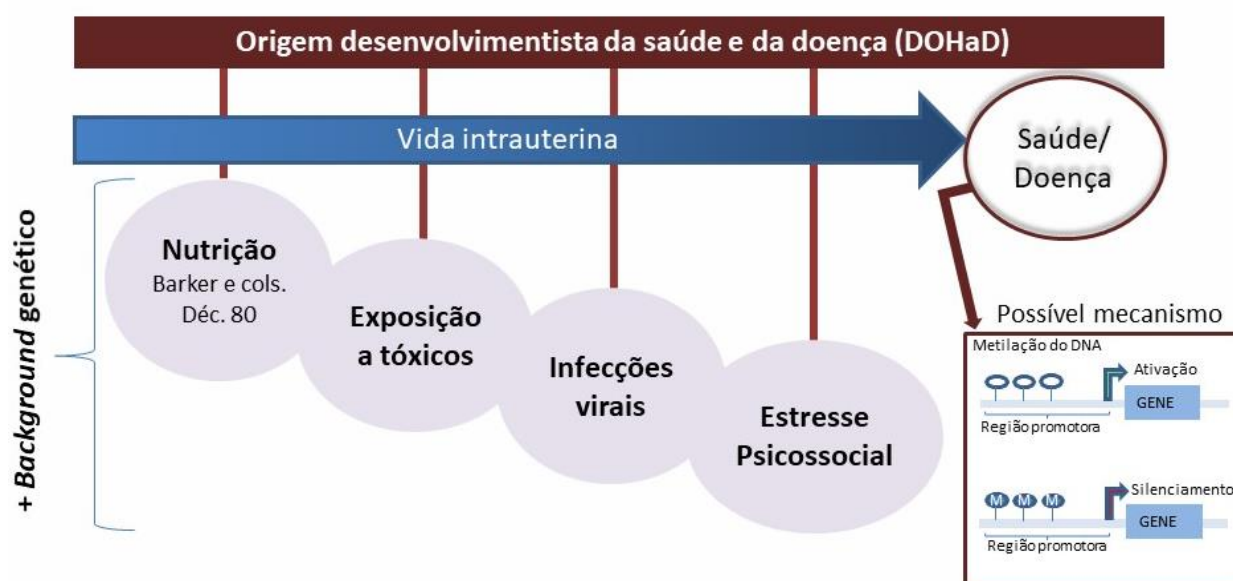
Através do avanço da ciência com metodologias de última geração e o reconhecimento da epigenética foi possível trazer-luz à explicação da evidente relação entre exposição ambiental e a alteração de um fenótipo (19,20). Hoje, o que podemos entender é que não é apenas a sequência do seu DNA que importa, mas, como o código é lido e interpretado. Em uma analogia à ortografia, é como pensar no papel do espaçamento, pontuação e formatação que permitem a leitura de palavras, sentenças e, por fim, a interpretação da informação. Palavras formam sentenças, marcas epigenéticas definem a estrutura de um cromossomo. Assim como parágrafos e capítulos altera o modo que lemos e organizamos as ideias, marcas epigenéticas definem como cada gene será lido(16). E finalmente, marcas epigenéticas poderiam ser a pontuação da sentença, definindo se um gene será lido ou não (ativo/silenciado). As marcas epigenéticas poderiam ser comparadas a pontos de exclamação como (me leia agora!), ou interrogação (quando não deve ser feita a leitura).

A conformação da cromatina pode ser dada pelos mecanismos epigenéticos tais como metilação do DNA, modificação de histonas, e de RNAs não-codificadores. Dos mecanismos epigenéticos, a doação de grupo metil ao DNA é o mais estudado por sua estabilidade (5,21,22). A catalisação da reação de adição do grupo metil ao dinucleotídeo é conduzida pelas enzimas DNAmetiltransferases (DNMT). A metilação do DNA pode alterar de forma estável a expressão gênica de uma célula, o que pode direcionar processos como diferenciação de células-tronco e *imprinting* genômico. Quanto a localização genômica, o processo de metilação pode ocorrer em diferentes regiões, mas, geralmente estão dentro do contexto de ilhas de citosina (especialmente no carbono 5 da citosina - 5mC) que precede uma guanina (ilha CpG). As ilhas são classicamente descritas como regiões do genoma de 500 a 1500 pares de base, com razão entre CG:GC igual a 0,6, ou seja, uma proporção do número observado para o número esperado de dinucleotídeos CpG acima de 0,6. Em relação a localização gênica, ilhas CpG tendem a estar localizadas mais em regiões promotoras do gene, associadas com regulação da expressão gênica, no qual 60-70% dos genes humanos estão associados a ilhas CpG (23).

Embora a metilação do DNA seja majoritariamente associada a supressão da expressão de um gene, a localização gênica onde ocorre a metilação do DNA pode tanto ser um gatilho de ativação ou de silenciamento. Geralmente, a metilação do DNA em região promotora está associada a supressão da expressão gênica, uma vez que “atrai” proteínas de ligação à metilcitosina que promovem a condensação da cromatina em conformações repressivas transcricionalmente.

Nos últimos anos, uma série de evidências têm relacionado a incorporação do grupo metil ao DNA com diversas exposições ambientais estressoras (**Figura 1**), seja como possível biomarcador, mediador ou preditor do processo saúde-doença ao longo da vida e, portanto, sendo o principal mecanismo de investigação desta tese. O universo completo de interações entre mecanismos epigenéticos, vias celulares e influências ambientais, dão origem ao conceito de Neuroepigenoma como um ponto de vista mais abrangente, que reúne dados básicos e clínicos (24).

Figura 1 - Exposições a diferentes fatores estressores têm sido associados a desfechos na saúde em fases posteriores da vida, e a metilação do DNA, pode ser um dos mecanismos que explicam essas associações descritas.



Fonte: Elaborado pela autora.

Nesse sentido, a introdução desta tese aborda primeiro, as relações entre situações estressoras vivenciadas no período de vida intrauterino e suas associações com o neurodesenvolvimento. Em seguida, será discutido evidências na literatura da relação entre exposição ao estresse ambiental e metilação do DNA. Na sequência, exploraremos o que a literatura traz quanto a metilação do DNA ser um fator mediador entre exposição ambiental durante a vida fetal e sua relação com aspectos do neurodesenvolvimento. Por fim, discutiremos a importância da placenta no contexto de programação do processo saúde-doença em fases posteriores da vida.

1.1 VARIÁVEIS AMBIENTAIS ESTRESSORAS E SUA ASSOCIAÇÃO COM CRESCIMENTO E NEURODESENVOLVIMENTO

Dada a importância da saúde intrauterina sobre a programação da saúde pós-natal, torna-se fundamental o reconhecimento dos fatores que podem representar riscos para a saúde do feto bem como identificar marcadores de qualidade de vida fetal e pós-natal. Um dos marcadores clínicos que têm sido associados à fatores de risco para transtornos do neurodesenvolvimento (TND) são as medidas de crescimento ao nascimento.(25) A seguir, serão apresentadas algumas evidências associando a exposição ao estresse pré-natal com crescimento e neurodesenvolvimento.

Quando se trata de estudos epidemiológicos, o peso ao nascimento é frequentemente adotado como um reflexo do ambiente intrauterino pois este, quando alterado, não permite um crescimento fetal adequado. Nesse sentido, uma série de estressores pré-natais foram avaliados nos últimos anos, com interesse no peso ao nascimento como desfecho. O estresse psicossocial materno, mensurado em um estudo longitudinal conduzido no estado de São Paulo, Brasil, foi associado a baixo peso ao nascimento (BPN), prematuridade e Restrição de Crescimento Intrauterino (RCIU) (26). Estima-se que 30% das mulheres grávidas relatam estresse psicossocial em suas vidas diárias, incluindo tensão no trabalho e sintomas depressivos ou de ansiedade (27). Monk e cols. avaliaram a presença de estresse em mães que reportavam, por intermédio de questionários, estresse físico e psicológico, e notaram que o grupo de mães com maior estresse psicológico tiveram filhos com mais complicações ao nascimento (28). Os autores concluíram que o suporte social foi o fator que mais diferenciou os grupos e que fetos masculinos estavam sob maiores riscos de piores desfechos. Ademais, a condição psicossocial materna é particularmente relevante quando se trata de regiões e populações de uma sociedade em vias de transição e importantes disparidades sociais, como sociedade brasileira, onde o estresse psicossocial é altamente prevalente (29–33).

O estresse pós-traumático, em revisão sistemática recente (2018), foi relacionado a BPN e menor adesão ao aleitamento materno (34) e em um estudo de meta-análise (35). Contudo, há outras revisões que concluem pouco ou nenhum efeito de estresse psíquico com BPN. É possível que interpretações individuais e instrumentos de coleta de dados, e, especialmente a heterogeneidade nas populações e circunstâncias analisadas, tenham forte relação com a dificuldade de se encontrar resultados significativos em meta-análises, ainda que os instrumentos aplicados sejam comuns. No entanto, outras exposições a estressores como hábito de fumar na gestação, tem mostrado forte relação com BPN. O peso ao nascer (PN), por sua

vez, tem sido associado a transtornos do neurodesenvolvimento (TND) (25,36). Em estudo recente no Brasil, Fink e cols., analisaram a associação de prematuridade e BPN como reflexo de estresse pré-natal e desfechos do neurodesenvolvimento aos 12 meses e constataram que ambos são fatores de risco para o crescimento e neurodesenvolvimento aos 12 meses de vida contudo, estratificando por sexo, o BPN foi mais relevante para o neurodesenvolvimento de meninos (37). Importante ressaltar que o PN segue uma curva em U em termos de risco para a saúde da criança, ou seja, tanto crianças pequenas para idade gestacional como crianças grandes para idade gestacional foram relacionadas com desfecho negativos a saúde.

Exposição a disruptores endócrinos, como o bisfenol A (BPA), encontrados em materiais plásticos, tais como mamadeiras (antes da regulamentação de 2011), os bifenilos policlorados (PCBs), os ftalatos (utilizado como aditivo para deixar o plástico mais maleável), também foram descritos como fatores de risco do neurodesenvolvimento e adiposidade da criança (38–40). Uma das formas de reatividade fisiológica ao estresse é a ativação do eixo hipotalâmico pituitário-adrenal (HPA) com consequente produção de corticosteróides. O cortisol é fundamental para o desenvolvimento e amadurecimento fetal, mas o excesso de cortisol na circulação materna pode acarretar prejuízos para estes processos assim como a programação comportamental ao longo da vida. Dada a importância do eixo HPA sobre a responsividade ao estresse, marcadores específicos desse controle central têm sido mensurados em estudos que avaliam estresse na gestação e desfechos sobre os filhos. Experimentos em animais (41) mostram que o estresse na rata prenha está relacionado ao aumento de medo, alteração do eixo HPA, e comprometimento cognitivo, especialmente funções executivas de seus filhotes. Em humanos, estudos de neuroimagem, mostram que depressão e ansiedade na gestação estão associadas com menor espessura cortical afetando importantes áreas do desenvolvimento social (42). Davis & Sandman, além da aplicação de escalas de depressão e ansiedade, mensuraram níveis de cortisol, um biomarcador de alterações do eixo HPA, e encontraram que o cortisol no início da gestação estava relacionado com o domínio cognitivo de crianças (43). Desse estudo então, é possível observar duas coisas: A) A importância do período da exposição ao estresse na linha do tempo, e B) O estresse pré-natal pode levar a mudanças fisiológicas importantes que podem ajudar a entender a possível relação entre exposição a estressores gestacionais com os TND. Contudo, acredita-se que há muito a ser explorado além do eixo HPA e estresse, uma vez que não explicam em sua totalidade os desfechos sobre o desenvolvimento da criança. Ainda, é possível achar marcas moleculares, como metilação do DNA, induzidas pelo ambiente que estão relacionadas a mudanças

fenotípicas e que nos permitam conhecer diferentes mecanismos que podem estar ligados a responsividade ao estresse.

1.2 EXPOSIÇÃO A ESTRESSORES AMBIENTAIS NA VIDA INTRAUTERINA E ASSOCIAÇÕES COM METILAÇÃO DO DNA

Aspectos cruciais do desenvolvimento, tais como proliferação, migração e diferenciação celular, são programados epigeneticamente. Durante o desenvolvimento fetal, parece ocorrer verificações constantes dos recursos disponíveis do ambiente para desenvolvimento fenotípico e se, caso observado inadequações, vias de adaptação são ativadas e muitas vezes, o potencial genético não é atingido. Se fizermos uma analogia, é como se, ao escolher um destino de um ponto de partida, exista um caminho principal programado, mas, caso esteja obstruído, outras avenidas podem ser utilizadas, que podem comprometer a chegada no ponto final de destino.

Modificações epigenéticas participam dos processos fisiológicos necessários para a formação, desenvolvimento e manutenção do SNC, e, portanto, estão associadas a desequilíbrios que podem ocorrer ao longo da vida (44). Uma série de evidências têm sido apresentadas relacionando fatores extrínsecos, neuroplasticidade e eventos epigenéticos sobre o neurodesenvolvimento. Por exemplo, utilizando-se de modelos animais com uma dieta rica em gordura foi possível associar com alterações na expressão de reguladores epigenéticos em diferentes regiões cerebrais no feto da prole, comparados ao grupo controle (45). Além da associação com ativação de vias inflamatórias, de fato, uma série de micronutrientes têm sido associados aos níveis de metilação por participarem do ciclo de um carbono tornando-os fatores que podem alterar padrões de metilação e, futuramente, representar maior risco para TND. A prematuridade foi associada a hipometilação de dois fatores neurotróficos, fator neurotrófico derivado do cérebro (*BDNF*) e fator de crescimento nervoso (*NGF*) no tecido placentário e no plasma do cordão umbilical, sugerindo mecanismos epigenéticos alterados que pode levar a alterações na programação fetal (46). Além dos fatores ambientais de exposição a partir da dieta, o estresse psicossocial também tem sido foco nos últimos anos quanto a sua associação com TND e alteração de metilação. Como referência nesse sentido, uma série de coortes podem ser citadas, como a das gestantes que vivenciam os conflitos constantes da República do Congo. Parte dessas mulheres sofrem estresse crônico e pós-traumático de guerra e foram reportados relação com peso ao nascer dos filhos e metilação diferencial em genes reguladores de mecanismos epigenéticos na mãe, tais como *DNMT1*, *DNMT3A*, *TET3* e em genes relacionados a via do cortisol no neonato (46,47).

Conforme pode ser observado nos exemplos citados anteriormente, a maior parte dos estudos avaliam os caminhos de: a) estresse pré-natal e metilação do DNA e b) metilação do DNA e desfechos do neurodesenvolvimento. Poucos relacionam a integração de ambos. Ainda, investigar a metilação do DNA pode envolver uma abordagem gene-específica, na qual se tem um gene candidato relacionado a uma via conhecida da sua hipótese, ou pode-se abordar milhares de sítios e explorar de maneira mais ampla as possíveis diferenças. Uma série de estudos, por exemplo, mostram que estresse psicossocial durante a gestação está associado a metilação do gene BDNF, e que por sua vez, está associado a um pior desfecho cognitivo. Uma outra evidência da importância da epigenética sobre o neurodesenvolvimento é sobre a via de sinalização do WNT, importante no processo da neurogênese. Por exemplo, exposição ao álcool durante a gestação pode ativar a via WNT/beta catenina e silenciar importantes genes dessa cascata, desencadeando uma hipomielinização no feto (48). A exposição de PCBs intra útero também já foi descrita por atuar na via WNT modificando o funcionamento dos canais de cálcio em neurônios fetais por intermédio do receptor RyR a partir de microRNAs, estando relacionadas a neurotoxicidade (49).

Uma meta-análise que envolveu as coortes ALSPAC e *Generation R* (2016), avaliou o estresse pré-natal nos domínios: (i) estresse da vida (por exemplo, morte na família, doença, problemas de trabalho), (ii) estresse contextual (por exemplo, dificuldades financeiras, problemas de moradia), (iii) estresse pessoal (por exemplo, psicopatologia, abuso de substâncias), e (iv) estresse interpessoal (por exemplo, dificuldades de relacionamento familiar, discussões com o parceiro). Os autores encontraram algumas regiões diferencialmente metiladas em filhos recém-nascidos de mães estressadas, mas não concluiu efeito robusto entre estresse pré-natal e metilação do DNA (50). Dois anos depois, uma revisão sistemática, em amostras de sangue do cordão umbilical encontram evidências para um único gene, o *NR3C1*, que é relacionado a produção de glicocorticóide e responsividade ao estresse (51). Existe um grande desafio na replicação desses estudos, seja pelas metodologias diferentes aplicadas, seja pela condução metodológica da análise de dados. Um estudo recente, analisou amostras da coorte do Project Viva, dos Estados Unidos, e observou um único sítio CpG hipometilado em um fator de transcrição (*ZNF575*) no cordão umbilical nos filhos de mães que usaram antidepressivos na gestação e essa associação persistiu na infância. Importaneamente, esse achado foi replicado na coorte holandesa, *Generation R* (52).

1.3 RELÓGIO EPIGENÉTICO E PLASTICIDADE FENOTÍPICA

Além de estarem associados a alterações na expressão gênica e alterações nos processos de desenvolvimento e crescimento, o nível de metilação permite estimar a idade de um indivíduo por intermédio de um “relógio epigenético”. O processo de envelhecimento humano é notável a olho nu ao longo das décadas de vida de um indivíduo, portanto, seria plausível a hipótese que marcas epigenéticas ao longo do genoma pudessem refletir esse processo, uma vez que a metilação do DNA é um mecanismo usado pelas células para controlar a expressão gênica, e esse processo também muda gradualmente à medida que envelhecemos (53). Inicialmente, pesquisadores estavam interessados na obtenção de um biomarcador de idade cronológica de fácil obtenção, uma vez que além da necessidade de entender o processo de envelhecimento ligado a inúmeras doenças, era necessário uma medida de boa acurácia para situações, por exemplo, em que não se tinha informação de idade de indivíduos, como em campos de refugiados. A criação do relógio epigenético têm evidenciado alterações no epigenoma por “envelhecimento” gerando um declínio na plasticidade fenotípica nos níveis celular e molecular (54).

Em 2013, Horvath e cols (53), publicaram seu primeiro algoritmo de estimativa da idade cronológica pelo nível de metilação de 353 sítios. Curiosamente, 24% desses sítios estão em CpGs mapeadas em elementos de resposta ao estresse (glicocorticóides). Desde então, outros algoritmos tecido-específicos foram desenvolvidos, como a idade gestacional em sangue do cordão umbilical, como de Knight e Bohlin (55,56). À partir desses dados, diferentes autores têm observado que, quando há um descompasso da estimativa do relógio biológico com a idade real do indivíduo, é possível supor uma senescência celular, que tem sido ligada a todas as causas de mortalidade não-acidentais (57).

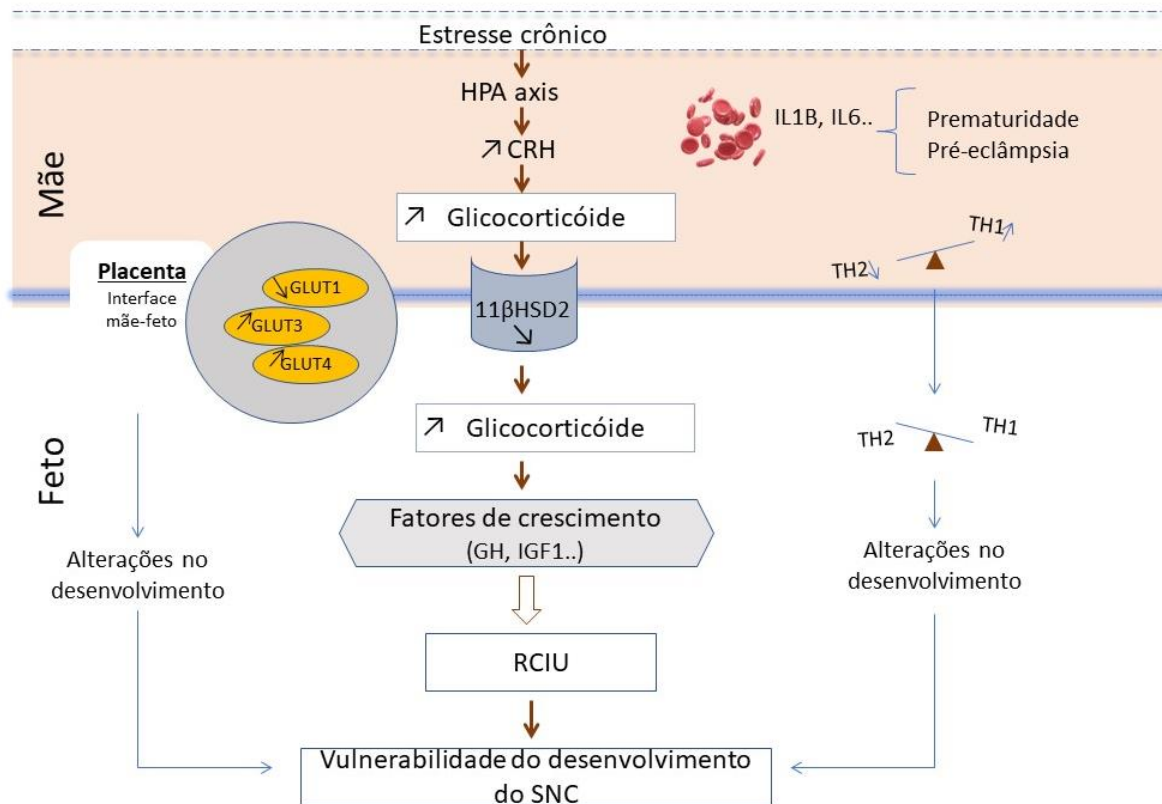
Embora a maior parte das evidências mostrem que um relógio epigenético mais acelerado, ou “envelhecido” comparado à idade cronológica de um indivíduo adulto tenham uma convergência para desfechos negativos para a saúde, quando se trata da estimativa da idade pelo perfil de metilação do recém-nascido a interpretação deve ser mais cautelosa, por existir poucos estudos investigando esse contexto. Tanto a aceleração quanto a desaceleração do relógio epigenético foram reportados com associações a diferentes exposições a estresse gestacional (58–60). Um ensaio clínico recente, (61), com população multiétnica, reconhecendo as funções sistêmicas da vitamina D e o RE como um marcador do estado cronológico celular, testou se a suplementação (4000 UI/dia) na gestação iria reduzir desacelerar o relógio epigenético (frente a idade gestacional). Aqui aceleração da idade gestacional foi positivamente

relacionada ao PN. A suplementação de vitamina D reduziu a aceleração do relógio epigenético, que apresentou um viés étnico, e foi especialmente observada em afro-americanos. Por último, ainda alguns autores trazem luz a possibilidade que embriões com perfis epigenômicos particulares têm diferentes probabilidades de sobrevivência no útero. A hipótese é de que existiria uma “seleção epigenética”. Tobi e col., reportaram em 2018, na *Cell Reports*, menor variação na metilação do DNA em fetos que sofreram adversidades nos primeiros estágios do desenvolvimento intrauterino (62).

1.4 PLACENTA: A INTERFACE QUE PERMITE A COMUNICAÇÃO ENTRE MÃE E FETO

Para entender o fenômeno de uma “programação fetal” sobre neurodesenvolvimento é importante reconhecer a placenta como um órgão neuroendócrino fundamental para garantir o desenvolvimento do feto intrauterino, que não só regula o transporte como interage com o SNC fetal e materno, alterando o sistema imune e produção de hormônios e citocinas (8,63). O papel da placenta no controle dos níveis de cortisol, por exemplo, é fundamental. Como mencionado anteriormente, o cortisol é necessário para o desenvolvimento e amadurecimento fetal, mas em excesso pode ser prejudicial. O aumento do cortisol materno inibe um importante transportador de glicose (GLUT1), prejudicando o suprimento de energia para o feto e como consequência, a restrição de crescimento intrauterino. Um dos recursos placentários de proteção fetal ocorre por intermédio da expressão do gene 11β HSD-2 que converte o cortisol em metabólitos inativos (30). Este gene parece ser hipometilado (inativado) na placenta de mães estressadas (**Figura 2**).

Figura 2 - A placenta e resposta ao estresse. O Estresse crônico oriundos de diferentes exposições está associado com aumento de cortisol circulante que, por sua vez, altera a expressão da 11- β -hidroxiesteróide desidrogenase tipo 2 (11 β HSD2) com diminuição do controle de entrada do cortisol circulante e possível restrição de crescimento intrauterino (RCIU) e neurodesenvolvimento.



Fonte: Figura adaptada de Viltart et al. (2007)(3)

No entanto, alguns descompassos durante a diferenciação e desenvolvimento placentário podem dificultar as interações entre os organismos materno e fetal com consequências duradouras para ambos. A maior parte do conhecimento sobre o início da diferenciação das células da placenta é proveniente de estudos animais, especialmente em ratos, como alternativa, face o desafio de obter placenta humana em estágios iniciais de formação. Modelos animais foram valiosos para uma série de descobertas, no entanto, apesar de algumas semelhanças, as vias de diferenciação em modelos animais não refletem inteiramente o que acontece na placenta humana e no neurodesenvolvimento (64). É certo que, especialmente a partir da sexta semana de gravidez, os estudos sobre placenta humana vêm adicionando importantes avanços (65). Por sua complexidade, a placenta possui diferentes tipos de células com funções distintas e fundamentais para garantir um pleno desenvolvimento fetal. Avanços tecnológicos, como o sequenciamento da expressão gênica de células únicas (*Single-Cell RNA*

Sequencing) em um mesmo tecido, permitiu o reconhecimento de grupos celulares com características singulares, revelando subtipos de trofoblastos e seus padrões de diferenciação. O interesse em amplificar o conhecimento e caracterizar essa diferenciação precoce se dá por acreditar-se que, grande parte das intercorrências placentárias ocorrem em fases iniciais da diferenciação, onde as células deixam de exercer funções inicialmente esperadas, comprometendo o desenvolvimento placentário e fetal (66,67). O reconhecimento desse processo poderia implicar em possíveis intervenções precoces e redução de risco para crescimento e TND (68).

Devido a desafios éticos e logísticos na obtenção de amostras de placenta humana desde o início da linhagem trofotoderma, a utilização de células-tronco pluripotentes induzida (iPS, do inglês *induced pluripotent stem cell*) de fibroblastos humanos é uma ferramenta promissora para a modelagem da placenta. Para outros órgãos e tipos de células, o iPS permite a pesquisa de células relevantes para doenças humanas anteriormente inacessíveis que são geneticamente idênticas à célula do paciente da qual derivam. Os modelos de células iPS podem ser úteis para entender a origem de doenças humanas e promover propostas terapêuticas.

A diferenciação em trofoblasto de células iPS pode acontecer por diferentes meios de culturas, que se propõem a induzir vias específicas conhecidas daquele tipo celular. Primeiro, independente do protocolo adotado, a célula pluripotente precisa inativar suas vias mantenedoras do estado de não-diferenciação. Estas vias apresentam significativa expressão dos fatores de transcrição *POU5F1*, *NANOG* e *SOX2* (69). A diferenciação em trofoblasto requer a expressão de marcadores de vias específicas, como o *GATA3*, *TFAP2A*, *ARID3A*, *HLA-G*. Majoritariamente, para indução de células parecidas com trofoblastos, os estudos utilizam um modelo com *Bone Morphogenic Protein 4* (BMP4) (70,71). No entanto, as células iPS modelam o desenvolvimento do trofotoderma e as doenças relacionadas à função do trofotoderma carecem da capacidade de gerar culturas de células em longo prazo que realmente se assemelhem às da placenta humana (referencia), reforçando a necessidade de se descrever um método de cultura eficiente e que de fato produza células mais semelhantes a trofoblastos.

1.5 DINAMICIDADE EPIGENÉTICA REFORÇA A GESTAÇÃO COMO UM PERÍODO PROPÍCIO A INTERVENÇÕES

As mudanças no padrão de metilação do DNA são geralmente interpretadas como uma resposta de uma exposição ou uma ação plástica/adaptativa do embrião (72).As evidências, especialmente sobre situações estressoras, que revelam a importância do período gestacional

para o neurodesenvolvimento aliadas ao reconhecimento da epigenética como um processo dinâmico fazem com que esse período se projete como uma janela de oportunidade para intervenções positivas, objetivando melhores desfechos ao longo da vida desse indivíduo.

O ensaio clínico randomizado (RCT, do inglês *Randomized Clinical Trial*) (73), LIFE-Moms, cuja proposta principal foi de intervenção de melhora no estilo de vida materno, resultou em menor ganho de peso gestacional e retorno ao peso pré-gestacional 12 meses pós-parto ($OR = 1,68$), mas sem diferenças antropométricas nos filhos aos 12 meses de vida. Com um olhar para a importância da maternagem o Nurse Family Partnership (Estados Unidos das Américas) (74) propõe visitas domiciliares desde o início da gestação (primigestas) por enfermeiras treinadas a dar suporte a essa futura mãe e garantir melhores desfechos para o binômio mãe e filho. Os autores observaram menor prevalência de prematuridade (8.7% vs. 12.3%) (75), adesão ao aleitamento materno (razão de prevalência ajustada = 1,20) (76) além de relatarem uma extensão positiva do impacto da intervenção apresentando menor razão de chance do segundo filho dessas mulheres nascerem com baixo peso ($OR=0,51$) (77). *Minding the Baby* (78) é outro RCT, direcionado ao suporte psicossocial de gestantes adolescentes e adultas jovens em estado de vulnerabilidade socioeconômica. Em um subgrupo analisado, os filhos de mães da intervenção apresentaram menores taxas de sobrepeso e obesidade aos 2 anos de vida ($OR=0,32$) (79). Uma intervenção na gestação baseada em *mindfulness* (atenção plena)(80) buscou analisar associação com melhora de sintomas depressivos maternos no fim da gestação bem como com desfechos obstétricos. Os autores destacaram que um nível mais alto de habilidades de atenção plena associou a nível mais baixo de sintomas depressivo. Mais importante ainda, um nível mais alto da habilidade de atenção plena de não reagir aos pensamentos perturbadores na 22ª semana de gestação foi associado a uma maior probabilidade de a criança ter peso adequado ao nascimento, em relação com BPN ($OR= 1,12$).

Existe uma plausibilidade biológica para que a metilação do DNA seja um mecanismo que explique essas associações. A maior parte dos estudos de intervenção que avaliaram o impacto sobre metilação do DNA são voltados para o aspecto nutricional materno, como intervenções dietéticas. A suplementação de ácido fólico em ensaio clínico (0.2mg ou 5 mg) (81) foi avaliada nos filhos de mulheres (média de 47 anos) e identificadas regiões diferencialmente metiladas mapeadas nos genes *HLA-DPB2*, *HLA-DPB1*, *PAX8* and *VTRNA2-1*, que estão relacionados ao desenvolvimento embrionário, proliferação celular e sistema imune. Uma revisão sistemática de ensaios clínicos randomizados (82), descreve poucas evidências em termos de suplementação de micronutrientes e metilação do DNA, enfatizando apenas que fatores estressores ambientais como IMC materno e hábito de fumar estavam

relacionados a responsividade da metilação do DNA aos estímulos nutricionais. Ademais, o uso de probióticos em uma pequena amostra de gestantes, relacionou-se com a redução de metilação do DNA, tanto da criança quanto da mãe, em genes promotores relacionados a obesidade(83).

1.6 HIPÓTESE

A hipótese desta tese é baseada em três premissas: (A) O período intrauterino representa uma alta vulnerabilidade para programação de estruturas e fisiologia do organismo, e as exposições a diferentes estressores resulta em aumento do risco de má adaptação fetal; (B) Mecanismos epigenéticos apresentam dinamismo e são influenciados pelo meio-ambiente com potencial papel mediador das relações observadas entre exposições a estresse intrauterino e medidas de crescimento e neurodesenvolvimento, e (C) É possível, partir de células-tronco pluripotentes induzidas (iPS), induzir a diferenciação em células trofoblásticas a partir de desenvolvimento de um novo modelo, auxiliando assim na compreensão de mecanismos vitais para o início do desenvolvimento placentário humano. Dessa forma, pretendemos mostrar que mecanismos epigenéticos, mais especificamente, a metilação do DNA, respondem dinamicamente aos estímulos intrauterinos e, intervenções nesse período, estão associadas a desfechos de saúde por mediação da metilação do DNA.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Avaliar a associação da metilação do DNA com exposições ambientais intrauterinas e sua relação com crescimento fetal e neurodesenvolvimento. Ainda, objetivamos criar um modelo de cultura de células que espelhe os funcionamentos das células trofoblásticas.

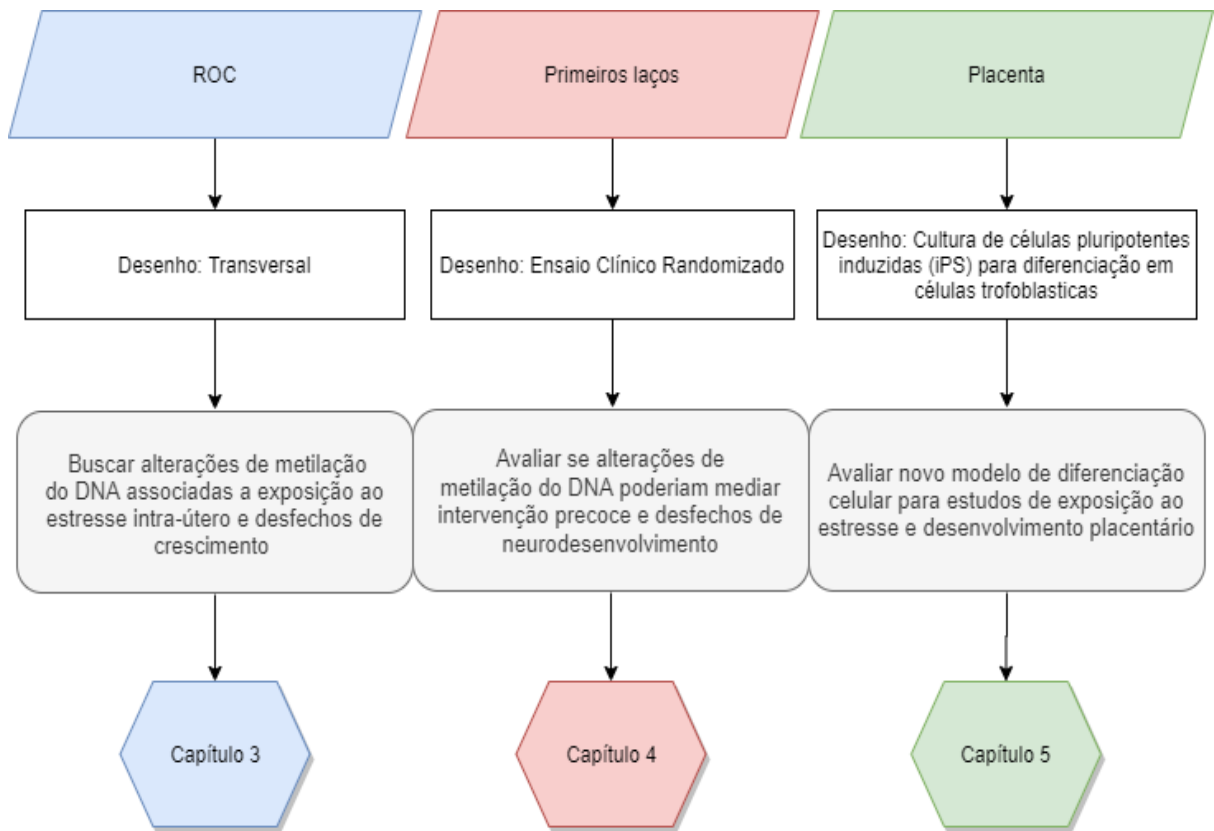
2.2 OBJETIVOS ESPECÍFICOS

- A. Buscar alterações de metilação do DNA associadas a exposição ao estresse intra-útero e desfechos de crescimento ao nascimento
- B. Avaliar se alterações de metilação poderiam mediar intervenção precoce e desfechos de neurodesenvolvimento
- C. Avaliar novo modelo de diferenciação celular para estudos de exposição ao estresse e desenvolvimento placentário

2.2.1 Organização textual para resposta aos objetivos específicos

Esta tese envolve três desenhos de estudos diferentes, dos quais, cada um deles responde a um objetivo específico. Desse modo, as seções Material e Métodos, Resultados e Discussão serão apresentadas divididas em três capítulos, organizados para responder aos objetivos supracitados (Figura 3)

Figura 3 - Fluxograma da apresentação desta tese com suas subdivisões em 3 capítulos que respondem os objetivos específicos 1, 2 e 3.



Fonte: Elaborado pela autora.

3 CAPÍTULO 3: ESTUDO ROC

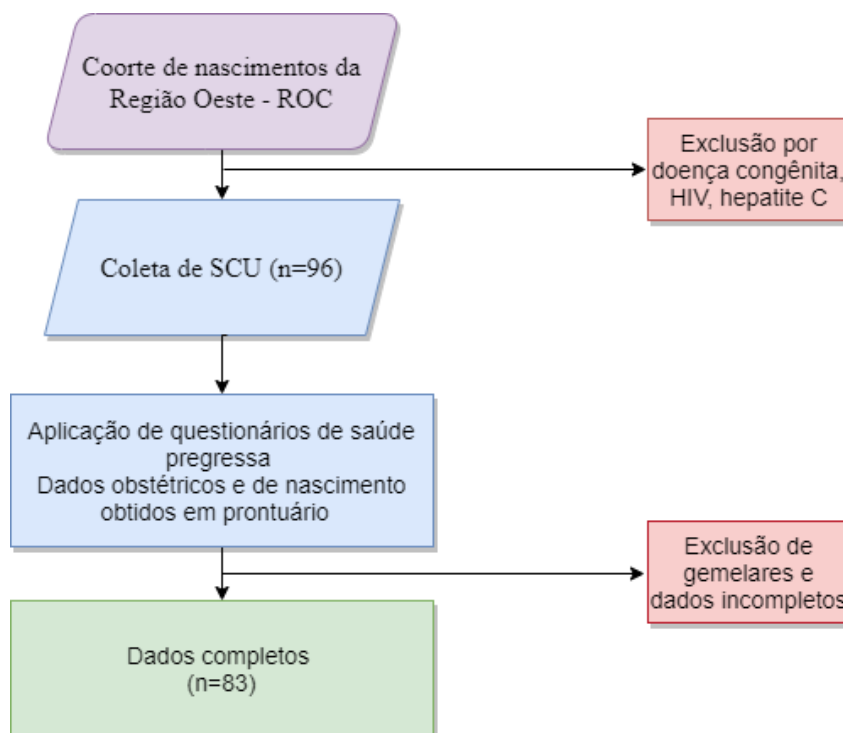
3.1. MATERIAL E MÉTODOS ROC

3.1.1 Seleção Amostral

A coorte de nascimentos da Região Oeste (ROC) conta com 900 crianças nascidas entre abril de 2012 e abril de 2014 no Hospital Universitário da Universidade de São Paulo (HU/USP). O grupo amostral desta tese consiste em uma subamostra da ROC que se diferencia por possuir coleta biológica (sangue do cordão umbilical, n =96). As gestantes foram abordadas (com o viés de seleção da necessidade da presença do pesquisador) ao dar entrada para internação no Hospital Universitário/USP para o parto, e tinham como critério de inclusão a) residir na região oeste do município de São Paulo e, b) assinar o Termo de Consentimento Livre e Esclarecido (TCLE), ou o responsável legal, se menor de idade. A coleta foi realizada no período de 6 meses da presença dos pesquisadores no hospital, sendo excluídas apenas aquelas que apresentavam doença congênita, diagnóstico de Vírus da Imunodeficiência Humana (HIV) e hepatite C. Dados obstétricos e do neonato foram obtidos a partir de prontuário médico. Foi aplicado questionário para obtenção de dados socioeconômicos, obstétricos e de situações potencialmente estressoras: estresse emocional referido, uso de drogas, hábito de fumar, estado nutricional avaliado pelo IMC pré-gestacional, ganho de peso gestacional e presença de doença psiquiátrica. Os dados avaliados neste estudo são, portanto, avaliados de maneira transversal com dados retrospectivos relacionados a gestação.

Figura 4 - Fluxograma das amostras utilizadas do estudo ROC.

Figura 4



Fonte: Elaborado pela autora.

3.1.2 Aspectos Éticos

Conforme determinações estabelecidas na resolução 466/12 do Conselho Nacional de Saúde, o presente estudo seguiu os princípios enunciados do Comitê de Ética em Pesquisa da Faculdade de Medicina da USP e Plataforma Brasil. Certificado de Apresentação para Apreciação Ética do Projeto ROC: CAAE 0104.0.198.000-10 (ANEXO 1).

3.1.3 Avaliação das medidas de exposição a estresse pré-natal do estudo ROC

As seguintes medidas de exposição a fatores de risco foram coletadas a partir do questionário aplicado por pesquisador treinado durante a internação hospitalar: dados socioeconômicos (renda e escolaridade), dados de saúde (presença de doenças), estresse emocional referido (relacionado a sentimentos na gestação e violência verbal, física ou sexual), abuso de álcool/drogas, hábito de fumar, presença de doença psiquiátrica, IMC pré-gestacional e ganho de peso gestacional. As perguntas foram respondidas durante o período de internação

hospitalar, de acordo com a disponibilidade da mãe, e no momento que elas estavam sem acompanhantes, no intuito de preservar sua privacidade.

3.1.4 Dados de saúde e antropometria do recém-nascido da ROC

Dados obstétricos e antropométricos do recém-nascido foram obtidos pelo programa de sistematização de dados clínicos dos pacientes existente no hospital em questão. Como protocolo hospitalar, logo após o nascimento, o pediatra dá o Escore de Apgar e a idade gestacional (IG) é calculada pelo método de Capurro. As medidas antropométricas coletadas ao nascimento são peso (g), comprimento (cm), circunferências cefálica, torácica e abdominal. Após obtenção desses dados, as crianças foram inseridas nas curvas de padrões de crescimento da Organização Mundial de Saúde (OMS) (84). Para inserção das crianças nas curvas de crescimento sexo-específicas, o *software* Anthro v3.2.2 foi utilizado (disponível em: <https://www.who.int/childgrowth/software/en/>). Os escores Z dos seguintes índices foram obtidos: Comprimento/ Idade (C/I), Peso/Idade (P/I), Índice de Massa Corporal/Idade (IMC/I) e Circunferência cefálica/Idade.

3.1.5 Coleta de amostra biológica (sangue do cordão umbilical - SCU) e extração do DNA da ROC

O material biológico coletado no Centro Obstétrico do Hospital Universitário da Universidade de São Paulo (CO HU-USP) foi obtido logo após o parto, através de punção venosa do cordão umbilical, armazenado em 2 tubos *Vacutainer* de 5mL contendo EDTA. As amostras foram mantidas no laboratório do HU e posteriormente transportadas para o LIM 23 (IPq-HCFMUSP), onde realizou-se o procedimento de extração de DNA. Para obter DNA a partir do SCU utilizou-se o kit QIAamp DNA Blood Mini (Qiagen, Hilden, Germany), conforme o protocolo do fabricante. A qualidade e integridade das amostras de DNA foram avaliadas por meio da densidade óptica (DO) em espectrofotômetro (NanoDrop® 2000, Thermo Scientific, Waltham, Massachusetts, Estados Unidos) e eletroforese em gel de agarose, respectivamente. As amostras de DNA foram tratadas com bissulfito de sódio (*EZ DNA methylation kit. Zymo Research, Irvine, Califórnia, Estados Unidos*), de acordo com as recomendações do fabricante.

3.1.6 Análise de metilação do DNA da ROC

A plataforma da *HumanMethylation450* (Illumina, San Diego) foi utilizada para determinar o status de metilação do DNA genômico de forma quantitativa de 482.421 dinucleotídeos CpGs. A metodologia empregada foi idêntica para ambos os estudos descritos nesta tese e será descrita apenas uma vez. O primeiro passo para essa análise ocorre na transferência das amostras para poços de microplacas. Essas amostras são então transferidas para um *Array* onde é criada uma imagem pelo Illumina Scanner, que sucede as etapas de amplificação, fragmentação, precipitação e ressuspensão e hibridização. O HM450K *BeadChip* emprega dois importantes ensaios químicos: a) *Infinium I*, utiliza dois tipos de *beads* por sítios CpG, um para cada um dos estados metilado e não metilado e b) *Infinium II*, ambas tecnologias visam aumentar a profundidade de cobertura para análise de metilação. As especificações de cada etapa podem ser consultadas na íntegra no guia do fabricante (*Illumina's Infinium HD Assay Methylation Protocol Guide*). O sinal emitido por fluorescência de cada sonda foi obtido pelo escaneamento no iScanner SQ (Illumina). O procedimento laboratorial foi executado pelo laboratório Deoxi Biotecnologia (www.deoxi.com.br). Os sinais obtidos foram convertidos em IDAT para análises posteriores de bioinformática, explicadas a seguir.

3.1.7 Análise de Bioinformática

A análise de dados dos microarranjos de metilação foi realizada na linguagem estatística de programação 'R' utilizando os pacotes *minfi*, *ChAMP* e *limma* (v3.11) (85), disponíveis no projeto Bioconductor (<https://www.bioconductor.org/>). Os arquivos Illumina IDAT, que contém informações dos sinais verde (não-metilado) e vermelho (metilado), foram processados e, em sequência, submetidos à etapa de controle de qualidade (exclusão de p-valor <0,01) e pré-processamento, que foram realizados de acordo com as recomendações do pacote utilizado. A normalização dos dados foi obtida utilizando os métodos *Noob*, que corrige o background baseado em uma convolução exponencial normal usando as sondas *out-of-band* (86) e *Quantile*, que usa uma normalização de quantis para ajustar as intensidades de todas as sondas dos *arrays* (87). A obtenção de imagens pré e pós-normalização foi por intermédio da função "*qcReport*" que auxiliam na identificação da escolha do método de normalização. Gráficos de multidimensionalidade (MDS) também foram obtidos para obtenção da variabilidade biológica do dado. Foram excluídas as *probes* dos cromossomos X e Y, para evitar viés de identidade sexual e aquelas localizadas em Polimorfismos de Nucleotídeo Único (SNPs), que são variações genéticas que influenciam a metilação. Sondas não específicas foram excluídas pela alta

probabilidade de co-hibridização. A anotação da lâmina foi realizada a partir dos arquivos Illumina usando UCSC versão hg19 do genoma humano de referência (*IlluminaHumanMethylation450kanno.ilmn12.hg19*).

O valor de metilação foi estimado como Beta ou M (log de Beta). Sendo beta (β) = intensidade do alelo metilado / (intensidade do alelo metilado + intensidade do alelo não metilado + 100). O valor de beta varia de 0 a 1, e é amplamente utilizado para medir a porcentagem de metilação. Em condições ideais, um valor de zero indica que todas as cópias do sítio CpG na amostra estão completamente não metiladas (nenhuma molécula metilada foi medida) e um valor de um indica que todas as cópias do local foram metiladas. Um valor M próximo a 0 indica uma intensidade semelhante entre as sondas metiladas e não metiladas, assumindo que os dados de intensidade foram devidamente normalizados. Valores M positivos significam que mais moléculas são metiladas do que não metiladas, enquanto valores M negativos significam o oposto. O valor M tem sido amplamente utilizado em análises estatísticas, especificamente em dados heterocedásticos, enquanto o Beta é mais utilizado para interpretação biológica (88). O controle de vieses técnicos foi analisado com análise de decomposição singular de valores (SVD - Singular value decomposition), através da função “*champ.SVD()*” (Morris et al., 2014). Para efeitos de correção de lâmina, “*champ.runCombat()*”, do pacote *sva*. A composição de celularidade sanguínea (Células B, células T CD4, células T CD8, granulócitos, monócitos, células exterminadoras naturais e glóbulos vermelhos) foi estimada com a função “*estimateCellCounts*” para tipo celular de SCU, do pacote *minfi*. O pacote *SVA* foi utilizado para correção de heterocelularidade.

Para a análise de metilação diferencial (comparação de metilação entre grupos de exposição a estresse na gestação), foram identificados inicialmente os sítios diferencialmente metilados (DMPs, do inglês, *Differentially Methylated Positions*), obtido por modelo linear Bayesiano do pacote *limma* (89), sendo considerados aqueles com teste Benjamini-Hochberg $<0,05$. O pacote “*DMRcate*” foi utilizado para obtenção das Regiões Diferencialmente Metiladas (DMRs, do inglês, *Differentially Methylated Regions*), seguindo os critérios de mínimo 500 nucleotídeos.

3.1.6 Análise do Relógio Epigenético

O Relógio Epigenético utilizado nesta tese consiste em um algoritmo desenvolvido a partir dos dados de metilação de 148 sítios da lâmina 450K, que dão uma estimativa da IG entre 24 e 44 semanas em amostras de sangue do cordão umbilical. Esse algoritmo foi desenvolvido por Knight e cols. (2016)(55) a partir de treinamento de banco de dados, utilizando seis coortes independentes para prever a Idade Gestacional (IG) com acurácia. Em termos estatísticos, a regressão de rede elástica foi adotada para selecionar esse conjunto de 148 sítios CpG preditivos de IG de um universo de 16.838 sítios CpG. Nos conjuntos de dados de treinamento, a correlação entre o preditor de IG e o IG clinicamente estimada foi de 0,99, indicando um forte ajuste do modelo.

3.1.7 Redes de metilação diferencial

As redes de metilação diferencial foram utilizadas para complemento de possíveis explicações moleculares das diferenças entre os fenótipos e serão detalhadas nos tópicos a seguir.

3.1.7.1 Identificação de Módulos Epigenéticos Funcionais (FEM)

O algoritmo de Módulos Epigenéticos Funcionais (FEM do inglês, *Functional Epigenetic Modules*) (90) é uma extensão do EpiMod e foi adotado aqui para avaliar se existiriam redes de “interactoma” de metilação diferencial de região promotora, partindo-se do pressuposto que a metilação do sítio próximo ao TSS (*transcription start site*) seja inversamente correlacionada com a expressão diferencial. Os pesos das arestas são construídos a partir das estatísticas de associação da metilação do DNA com o fenótipo de interesse. Assim, o algoritmo EpiMod pode ser visto como um algoritmo funcional supervisionado, que usa uma rede de relações entre genes (no nosso caso, uma rede PPI, *protein protein interaction*), para identificar sub-redes onde um número significativo de genes está associado a um fenótipo de interesse (POI). Esses “pontos de acesso” são os Módulos Epigenéticos Funcionais (FEMs) (Disponível em: <http://code.google.com/p/epimods/>).

3.1.7.2 A análise de rede de correlação ponderada – WGCNA

Para as redes de co-expressão, adotamos o algoritmo de análise de rede de correlação ponderada, o (WGCNA, do inglês *Weighted Correlation Network Analysis*) que é um método amplamente utilizado para análises envolvendo expressão gênica (microarranjos ou RNA-seq) e também de metilação de DNA (91,92). Brevemente, a entrada de dados é o valor de metilação (valores M) dos sítios de dois grupos (por exemplo, caso e controle), processados em separado. Para cada um dos dois grupos, o algoritmo calcula a Correlação de Pearson entre os valores M de todos os pares de sítios. O WGCNA irá criar uma aresta entre dois sítios se eles apresentaram uma altíssima correlação (ex. 0,9), criando assim uma rede de co-metilação. para construir matrizes de correlação de pares de co-metilaçãoexpressão (uma para cada grupo). Este algoritmo assume que as redes de co-metilação/ expressão gênica apresentam topologia de uma rede livre de escala (*Scale-Free*). Para isso, então os valores de correlação são elevados a potência β (valores naturais crescentes) (β), a fim de eliminar arestas da rede que não atendam a um determinado limiar e mantendo somente as arestas originárias das mais altas correlações. A rede final é aquela resultante do valor de β que atinge maximiza um índice mínimo de independência livre de escala (o coeficiente de determinação da linha de regressão que melhor se aproxima da distribuição do grau do nó em escala log-log). Assim, os dados de metilação expressão dos grupos de caso e controle são agrupados usando medida de sobreposição topológica (TOM) e os módulos são definidos como ramos da árvore agrupada resultante. Até este ponto, os grupos de caso e de controle são processados em paralelo. Finalmente, o WGCNA usa os valores de TOM de ambos os grupos para uma análise de preservação de módulo, onde indica a preservação dos módulos nas redes de controle e de caso (com base nos sítios genes contidos em cada módulo). O resultado é uma lista de módulos (e seus respectivos sítios genes) classificados por uma pontuação de preservação (92).

3.1.8 Análises estatísticas da ROC

O banco de dados e as análises estatísticas e construção de gráficos foram realizados no programa R 3.5.0. A normalidade do dado foi avaliada pelo teste de *Shapiro-Wilk*. Para comparação de proporções das variáveis normais, os testes t de *Student* e Qui-quadrado foram aplicados. Para os não-paramétricos utilizou-se o teste de *Wilcoxon*.

Para buscar associações entre as variáveis de metilação e desfechos de crescimento ao nascimento, modelos de regressão linear foram utilizados. Como variável dependente, inserindo isoladamente as medidas de crescimento: índices da curva de crescimento da OMS de acordo com o sexo da criança em escores Z: Peso/ Idade, Comprimento/Idade, IMC/Idade e circunferências abdominais, cefálica e torácica. Variáveis obtidas a partir da análise de metilação foram analisadas como independentes.

Para avaliação da discordância entre a idade metilômica estimada (RE) e a idade gestacional clínica, criou-se uma variável a partir da subtração da idade cronológica e RE (Δ idade). Dessa forma, também foram separados em grupos aqueles que tinham um RE menor que a IG clínica e aqueles que tinham RE maior que a IG clínica.

$$\Delta \text{ idade} = \text{Relógio Epigenético} - \text{Idade Gestacional}$$

Para avaliação das medidas de exposição de estresse, consideramos a somatória de cada variável coletada por questionário de saúde da gestação com uma pontuação para a obtenção de escores de estresse. A partir dessa variável, criou-se 2 grupos (A, menor exposição e B, maior exposição) a partir da mediana para avaliar perfis que poderiam representar maior risco de comprometimento de crescimento fetal. Esses grupos foram comparados quanto às diferenças em relação aos desfechos de crescimento.

Por fim, para a análise de mediação, nossa variável exploratória foi grupos de exposição a estresse, a mediadora o RE e o descompasso do RE e os desfechos foram as medidas de crescimento. Usamos o pacote de mediação versão 3.4.4 [37] para analisar a mediação causal com bootstrap de 10.000 replicações para intervalos de confiança. Os dados do modelo de mediação para cada CpGs ou DMRs incluem estimativas e ICs de 95% dos efeitos total, direto e de mediação. O *Average Causal Effects* (ACME) afirma se a exposição ao estresse contribui negativamente para medidas de crescimento por meio de um mediador (RE ou descompasso do RE). O efeito total e o efeito direto (ADE) mostram se o estresse tem probabilidade de alterar

o crescimento intrauterino nos modelos de mediação. A função *medsens* foi aplicada para realizar a análise de sensibilidade para testar os efeitos de fatores de confusão não medidos.

3.2 RESULTADOS DA ROC

3.2.1 Caracterização da Amostra da Coorte de Nascimento da Região Oeste (ROC)

A **Tabela 1** apresenta as características maternas. A média de idade materna das participantes da coorte ROC foi de 25 anos, com IMC pré-gestacional médio de 25 kg/m², cuja classificação segundo a OMS é de sobrepeso. Oito mulheres (9,6%) relataram uso de cigarro durante a gestação, 7,2% faziam uso de medicação psiquiátrica. Quanto aos recém nascidos, 54,2% eram do sexo feminino e o peso médio e comprimento ao nascer foram de 3325g e 48,8 cm, respectivamente.

Tabela 1 - Características socioeconômicas, obstétricas, clínicas e bioquímicas, de acordo com o sexo (n=83)

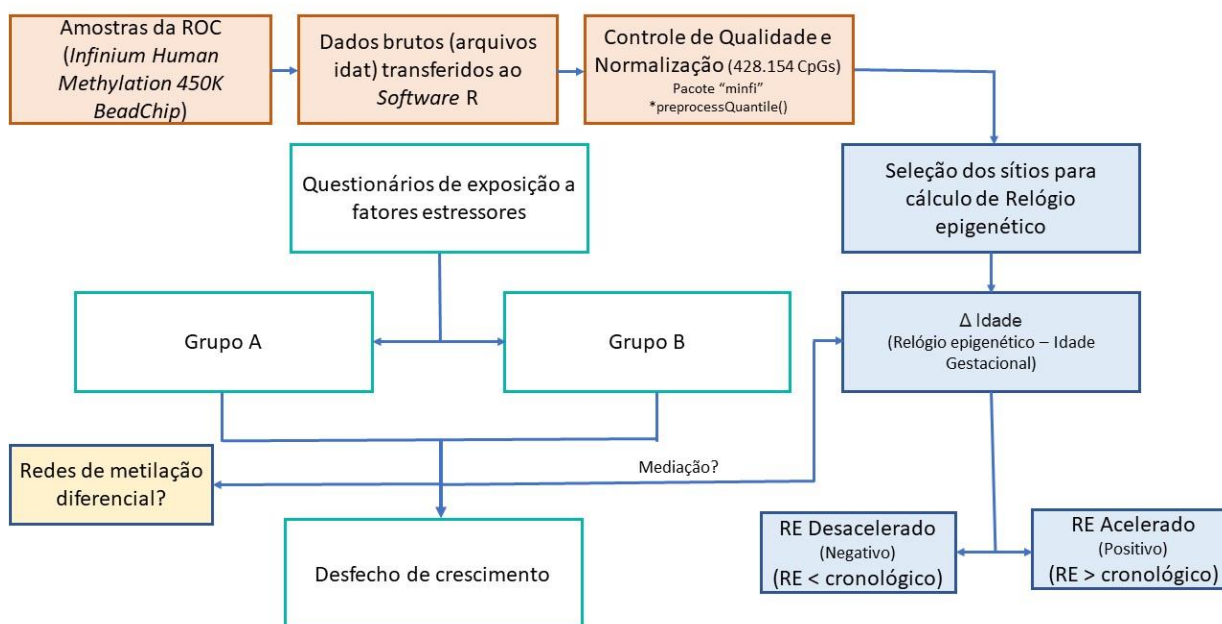
	n	%	Média (± DP)
Características Maternas			
Idade (anos)			25,6 (6,9)
IMC pré-gestacional (kg/m ²)			25,0 (5,2)
Ganho de peso na gestação (kg)			13,3 (5,3)
Sentimentos negativos na gestação (1-5)			1,8 (1,2)
Hábito de fumar	8	9,6	
Uso de medicação psiquiátrica	6	7,2	
Tipo de parto (Normal)	48	57,8	
Características do Recém-nascido			
- Sexo (Feminino)	45	54,2	
- Idade Gestacional (semanas)			39,5 (1,6)
- Peso ao Nascer (g)			3324,7 (489,3)
- Comprimento (cm)			48,8 (1,95)
- IMC/idade (escore z)			0,4 (1,1)
- Peso/idade (escore z)			0,0 (1,0)
- Comprimento/idade (escore z)			-0,5 (1,1)
- Circunferência torácica (cm)			33,3 (1,7)
- Circunferência abdominal (cm)			32,3 (2,2)
- Circunferência cefálica (cm)			34,4 (1,2)

Fonte: Elaborado pela autora.

O objetivo das análises com os dados da ROC foi responder se as exposições de estresse se diferenciavam entre as mulheres, e se estas se associariam com alterações do padrão de

metilação do DNA e medidas de crescimento. Sendo assim, os dados foram explorados em diferentes formas, tendo sempre como métrica a busca relação entre exposições ambientais na gestação e alterações epigenéticas e de crescimento. Como a metilação do DNA pode ser marcadora de exposições do ambiente fetal, apresentaremos os resultados da metilação e em sequência a primeira hipótese a ser explorada será que a idade metilômica (relógio epigenético) está relacionada a medidas de crescimento. Na figura abaixo (**Figura 5**) podemos observar o fluxograma representando as análises. Em seguida, olhamos para os dados de exposição ao estresse e fatores de crescimento. Ademais, tendo o hábito de fumar como um fator estressor amplamente investigado na literatura, também o avaliamos de maneira isolada, para confirmar se apenas esse fator poderia explicar possíveis diferenças nas medidas de crescimento. Por fim, avaliamos se haveria um papel mediador do relógio epigenético sobre as relações de exposições estressoras e medidas de crescimento.

Figura 5- Fluxograma das análises do estudo ROC (n=83). Dados brutos (idat) foram processados e seguiram para as etapas de controle de qualidade e normalização (n=428.154). Em seguida, o relógio epigenético foi obtido pelo algoritmo de Knight e cols. (2016) e realizada a análise de metilação diferencial entre os grupos criados a partir da exposição ao estresse.

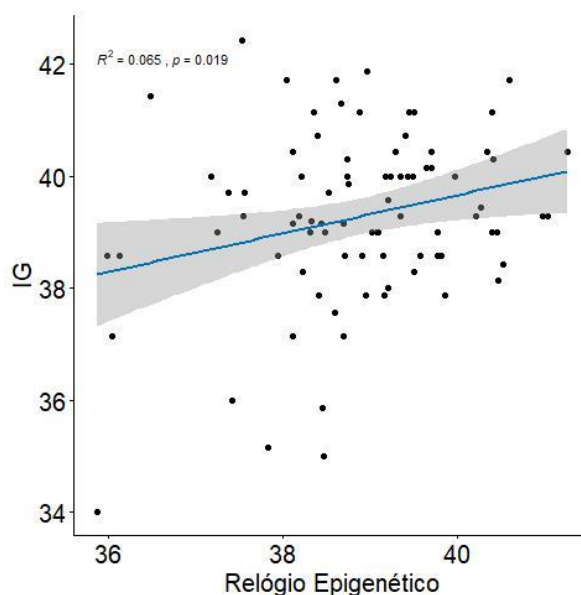


Fonte: Elaborado pela autora.

3.2.2 Relógio epigenético (RE) e medidas de crescimento ao nascimento

A discordância entre a idade cronológica e a idade estimada pela metilação pode ser indicativo de algum processo celular em descompasso, ou até representar um indicativo de senescência celular, esse último, mais especialmente descrito no adulto. O RE foi calculado a partir do algoritmo de Knight (55), cujo objetivo é estimar a idade gestacional a partir da metilação do DNA de 148 sítios específicos. Tendo em vista que a exposição ao estresse gestacional está associada a alterações do crescimento fetal, exploramos a hipótese que o descompasso entre RE e IG poderia mediar efeitos da exposição ao estresse nos desfechos do desenvolvimento fetal, através das medidas de crescimento ao nascimento. O primeiro passo antes de dar seguimento às análises principais foi avaliar se havia de fato uma associação entre o estimado pelo algoritmo do RE e a IG cronológica. Houve uma associação positiva, com um R^2 ajustado de 0,07 (**Figura 6**).

Figura 6 - Gráfico de dispersão de regressão linear da variável dependente relógio epigenético e IG clínica ao nascimento, como variável independente, pelo método Capurro (p-valor=0,019).



Fonte: Elaborado pela autora.

A média de IG estimada pelo RE foi de 38,9 (35,9 a 41,2, DP=1,2) semanas, enquanto a média de IG por prontuário médico (método Capurro) foi de 39,3 (34,0 a 42,4, DP=1,6) semanas. Para avaliar o descompasso (Δ idade) entre a IG real e estimada, o RE foi subtraído

da IG, ou seja, um valor negativo indica que $RE < IG$ e um valor positivo indica $RE > IG$. Obtivemos descompasso com média de -0,4, com valores entre -5,0 a 3,5 (DP=1,7). Dessa forma, olhamos para i) idade estimada pelo relógio epigenético (semanas) e ii) diferença aritmética entre a idade estimada pelo relógio epigenético e a idade gestacional (ambas em semanas).

Quando investigamos a variável contínua (em semanas) obtida pelo algoritmo do RE, não encontramos nenhuma associação com parâmetros de crescimento. A **Tabela 2** abaixo, mostra os resultados das regressões controladas por sexo (exceto para as curvas da OMS, por serem sexo específicas). Interessantemente, ao olharmos para o delta idade, é possível observar associações negativas com índices Peso/Idade Gestacional e Comprimento/Idade Gestacional (**Figura 7**). Isto é, quanto houve um saldo positivo na subtração das idades ($RE > IG$), os respectivos índices foram menores.

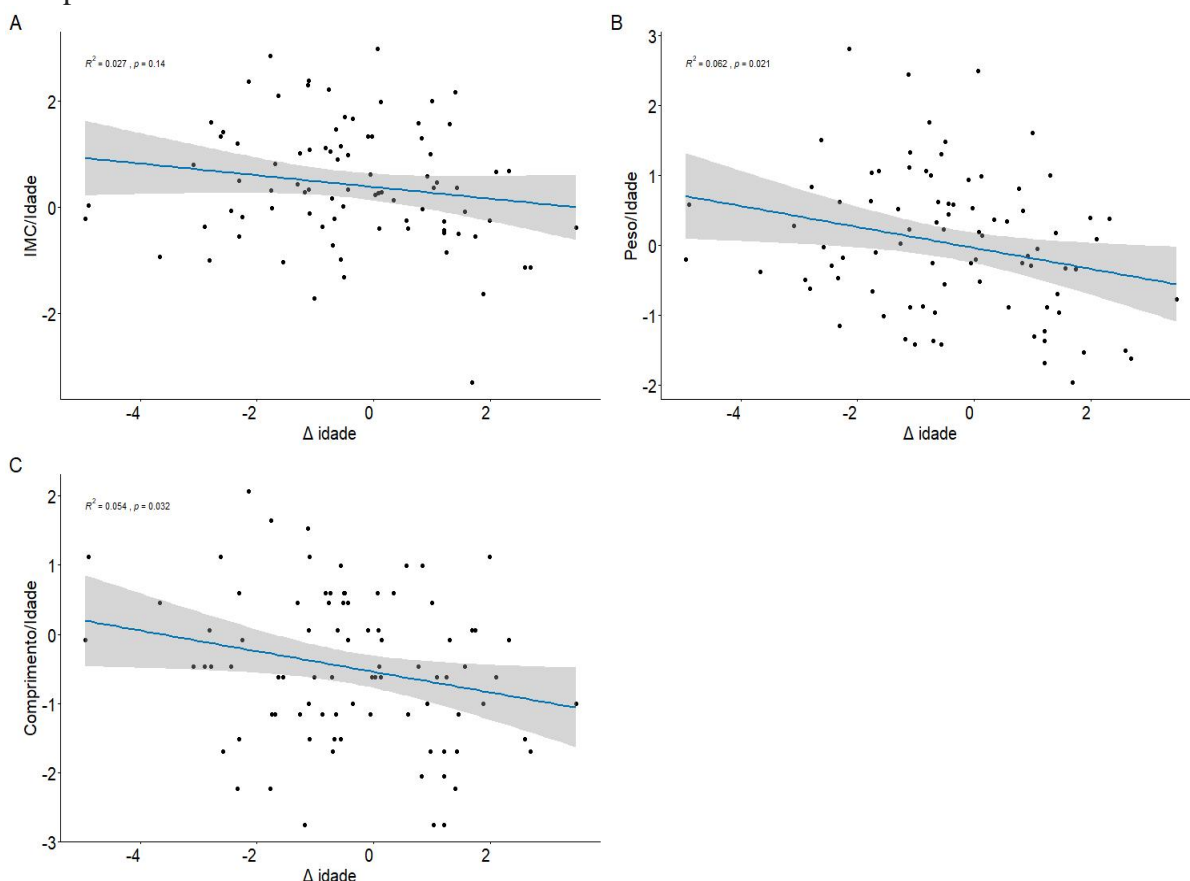
Tabela 2 - ROC: Associação de medidas ao nascimento com relógio epigenético (variável de interesse)

	Relógio Epigenético				Relógio Epigenético – Idade Gestacional			
	Coef.	R ² ajustado	P-valor	P-valor Reg	Coef.	R ² ajustado	P-valor	P-valor Reg
Peso ao Nascer*	3.28	0.00	0.209	0.451	0.00	0.00	0.209	0.451
Comprimento*	0.05	-0.02	0.467	0.766	-0.05	-0.02	0.467	0.766
Idade Gestacional*	0.34	0.04	0.020	0.059	-0.81	0.55	0.000	0.000
Circunferencia cefálica*	0.00	-0.02	0.974	0.998	0.00	-0.02	0.974	0.998
Circunferência torácica*	0.08	-0.01	0.264	0.533	-0.08	-0.01	0.264	0.533
Circunferência abdominal*	0.00	-0.02	0.977	0.999	0.00	-0.02	0.977	0.999
IMC/IG	0.13	0.00	0.246	0.246	-0.24	0.01	0.135	0.135
P/IG	0.16	0.01	0.193	0.193	-0.42	0.05	0.021	0.021
C/IG	0.08	-0.01	0.479	0.479	-0.36	0.04	0.032	0.032

*Controlados por sexo; legenda: IMC= Índice de Massa Corpórea (kg/m²), P = peso, Coef.= coeficiente, C= comprimento, IG=Idade Gestacional, I= Idade.

Fonte: Elaborado pela autora.

Figura 7 - Gráfico de dispersão gerado de regressão linear tendo como variável independente o Δ idade (RE-IG) e de dependente os índices **A)** IMC/Idade, **B)** Peso/Idade, **C)** Comprimento/Idade.



Fonte: Elaborado pela autora.

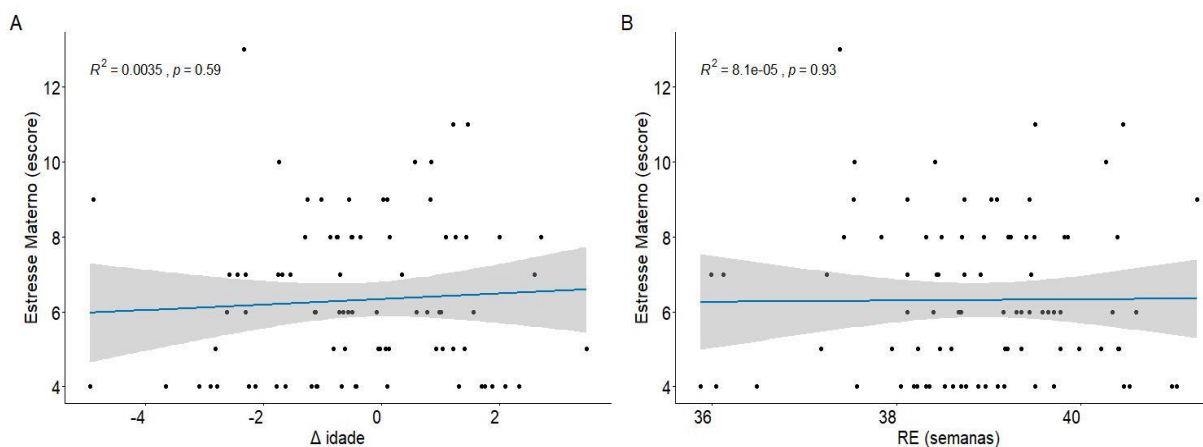
A análise mostrou associação negativa indicando que uma idade metilômica maior que a cronológica, associa-se à menor crescimento. Assumindo que um envelhecimento fetal seja decorrente da exposição ao estresse gestacional, buscamos associação entre exposição ao estresse e delta idade.

3.2.4 Exposição a estresse pré-natal, RE e medidas de crescimento ao nascimento

A maior parte do dado de exposição a estresse na gestação foi coletado de maneira dicotômica (sim e não). As variáveis de exposições ambientais coletadas nos questionários foram escolaridade, socioeconômicas (renda familiar), fumo, abuso de álcool e drogas, doença psiquiátrica, sentimentos durante a gestação (escala de 1-5) e IMC. Para a criação de escores de estresse, não incluímos as variáveis: a) renda familiar; como essas gestantes são provenientes de um hospital público da zona Oeste de São Paulo, a renda familiar apresentou pouca variabilidade e, b) IMC materno, por falta de dados de parte da amostra ($n=20$, 24,1% da

amostra total). Primeiro, olhamos para a relação entre escore de exposição a estressores e o RE-IG. Nesse sentido, nenhuma associação foi observada. Buscamos também associações entre escores de exposição ao estresse e medidas do crescimento, sem nenhum achado significativo.

Figura 8 - Gráfico de dispersão do escore do estresse materno e **A)** Delta idade ($p = 0,59$), **B)** RE ($p=0,930$)



Fonte: Elaborado pela autora.

Tabela 3. Associação de medidas ao nascimento com a somatória de estressores da ROC.

	Coefficiente	R ² ajustado	P-valor
Peso ao Nascer	-40.29	0.02	0.116
Comprimento	-0.07	-0.01	0.542
Circunferência cefálica	-0.02	-0.01	0.758
Circunferência torácica	-0.11	0.01	0.213
Circunferência abdominal	-0.11	0.00	0.308
IMC/IG	-0.09	0.01	0.137
P/IG	-0.08	0.01	0.149
C/IG	-0.03	-0.01	0.645

Fonte: Elaborado pela autora.

3.2.5 Criação de grupos a partir do escore de estressores ambientais

Com o intuito de fazer uma análise categórica, criamos dois grupos a partir da mediana do escore de exposição a estresse. Para fins práticos, chamaremos de grupo A, que concentra as mães que apresentaram menor escore de estresse e grupo B, aquelas que apresentaram maior escore de estresse.

Nessa formação dos grupos A e B pela exposição a estressores ambientais durante a vida pré-natal, observa-se que as crianças do grupo B apresentam médias menores do índice de crescimento da OMS, IMC/Idade (**Figura 9**). A **Tabela 4** abaixo apresenta a caracterização da amostra segundo os grupos de exposição a estressores ambientais na gestação.

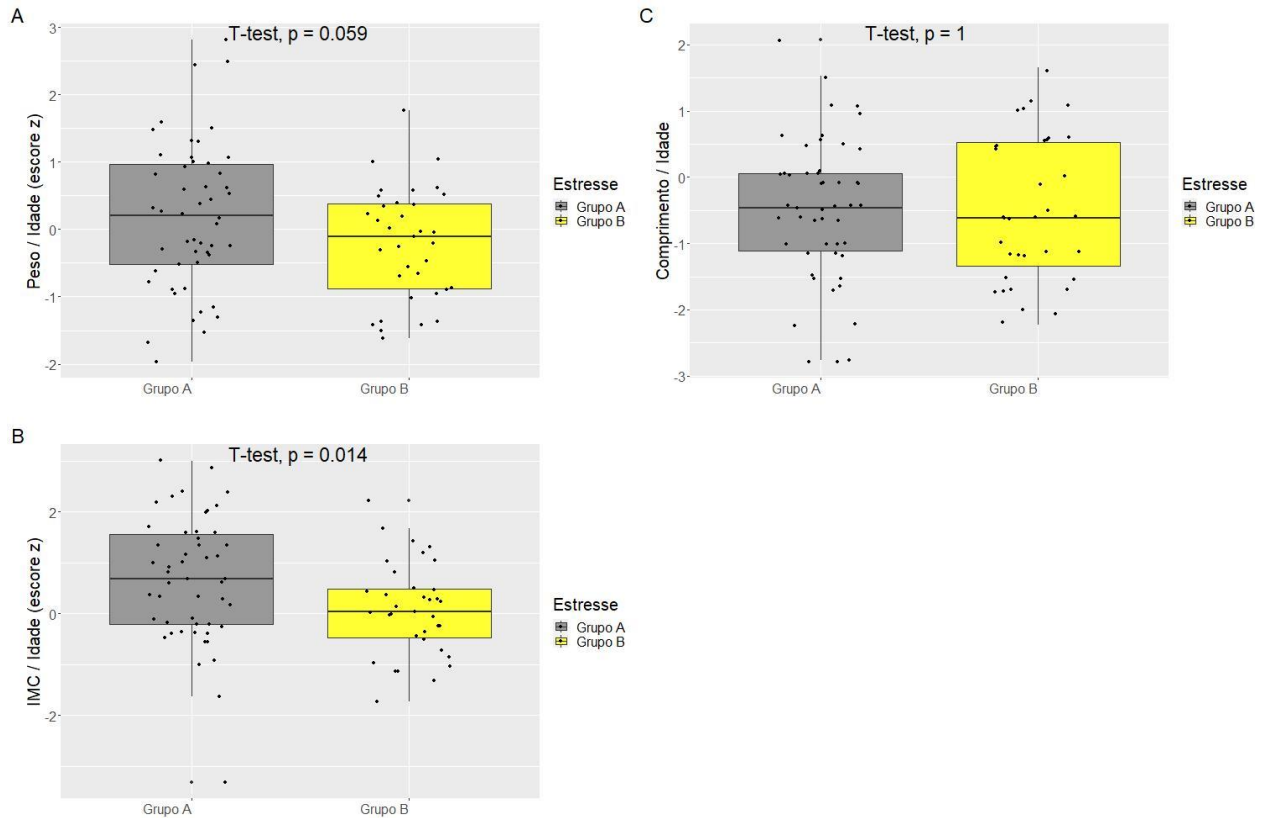
Tabela 4 - Características socioeconômicas, obstétricas, clínicas e bioquímicas, de acordo com o grupo de exposição ao estresse (n=83)

	Grupo A (n=50)			Grupo B (n=35)			p-valor
	n	%	Média (± DP)	n	%	Média(± DP)	
Características Maternas							
Idade (anos)			25.8 (7.7)			25.3 (5.6)	0.706
Renda Familiar							0.519
1-2	2			5			
3-4	23			14			
≥5	19			17			
Escolaridade (anos)							0.244
2	4			2			
3	26			17			
4	8			6			
5	8			10			
IMC pré-gestacional (kg/m ²)			25.4 (5.7)			24.3 (4.4)	0.386
Ganho de peso na gestação (kg)			12.7 (4.8)			12.5 (5.1)	0.904
Sentimentos negativos na gestação							<0.001
Hábito de fumar	3			13			<0.001
Uso de medicação psiquiátrica				7			0.004
Tipo de parto							0.999
Normal	29			20			
Características do neonato							
Sexo (meninos)	24			14			0.289
Idade gestacional (semanas)			39.3(1.6)			39.1 (1.5)	0.472
- Birth weight (g)			3397.3 (529.1)			3201.5 (395,3)	0.064
- Length (cm)			48.6 (2.1)			48.5 (2.2)	0.903
- BMI/Age (z-score)			0.7 (1.2)			0.1 (0.9)	0.019
- Cephalic circumference (cm)			34.5 (1.3)			34.4 (1.2)	0.826

*Salário mínimo no Brasil, no período = R\$ 937.00.

Fonte: Elaborado pela autora.

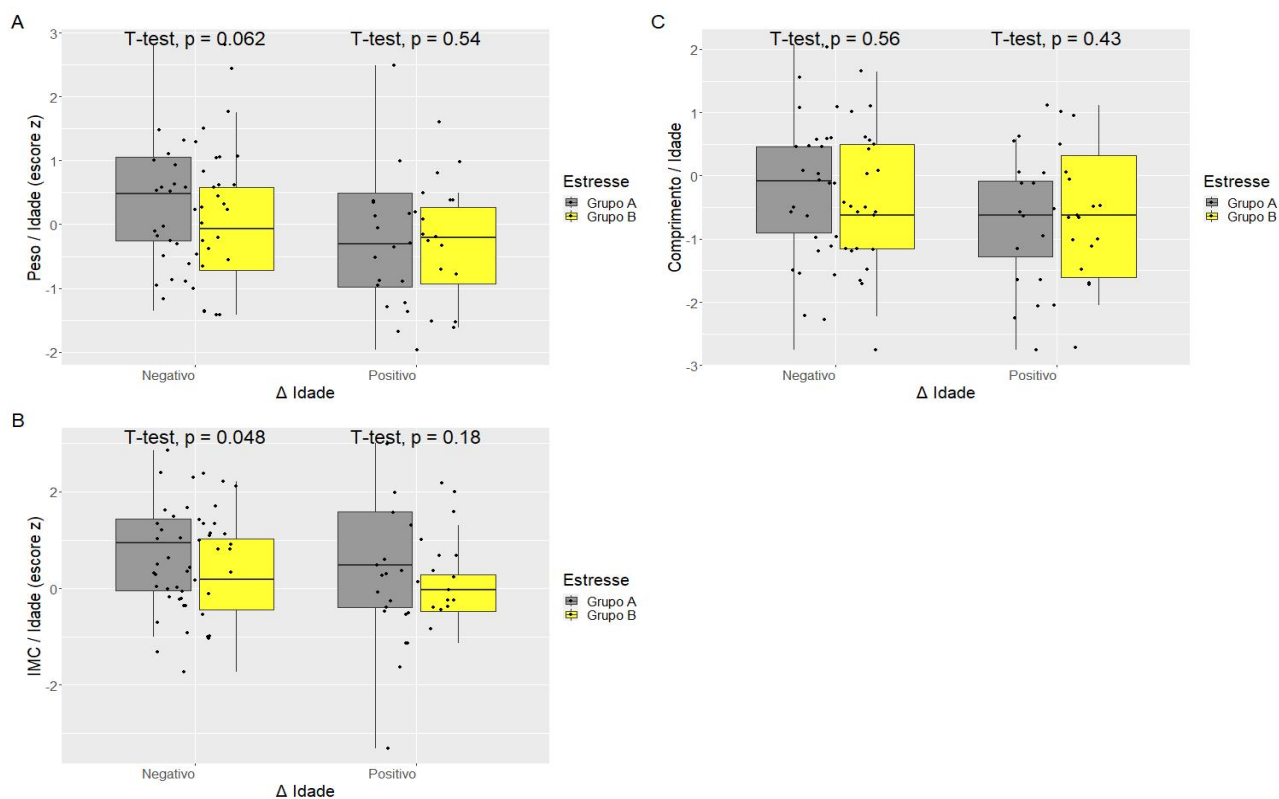
Figura 9 - Boxplot dos índices de crescimento (OMS, 2016), segundo os grupos criados a partir da mediana dos fatores estressores ambientais. **A)** Peso/Idade, **B)** IMC/Idade ($p=0,014$), **C)** Comprimento/Idade.



Fonte: Elaborado pela autora.

Buscamos saber se haveria diferenças dos grupos de estresse, desfechos neonatais e RE-IG (**Figura 10**). A única diferença estatisticamente significativa foi que para os indivíduos que tiveram $RE-IG > 0$ e maior exposição a estresse comparados aos com menor exposição apresentaram menores IMC, e Peso/Idade, mas sem diferença estatisticamente significativa.

Figura 10 - *Boxplot* do índice de IMC/I (OMS, 2016), segundo os grupos criados a partir da mediana dos fatores estressores ambientais do estudo ROC. **A)** Peso/Idade, **B)** IMC/Idade ($p=0,014$), **C)** Comprimento/Idade.

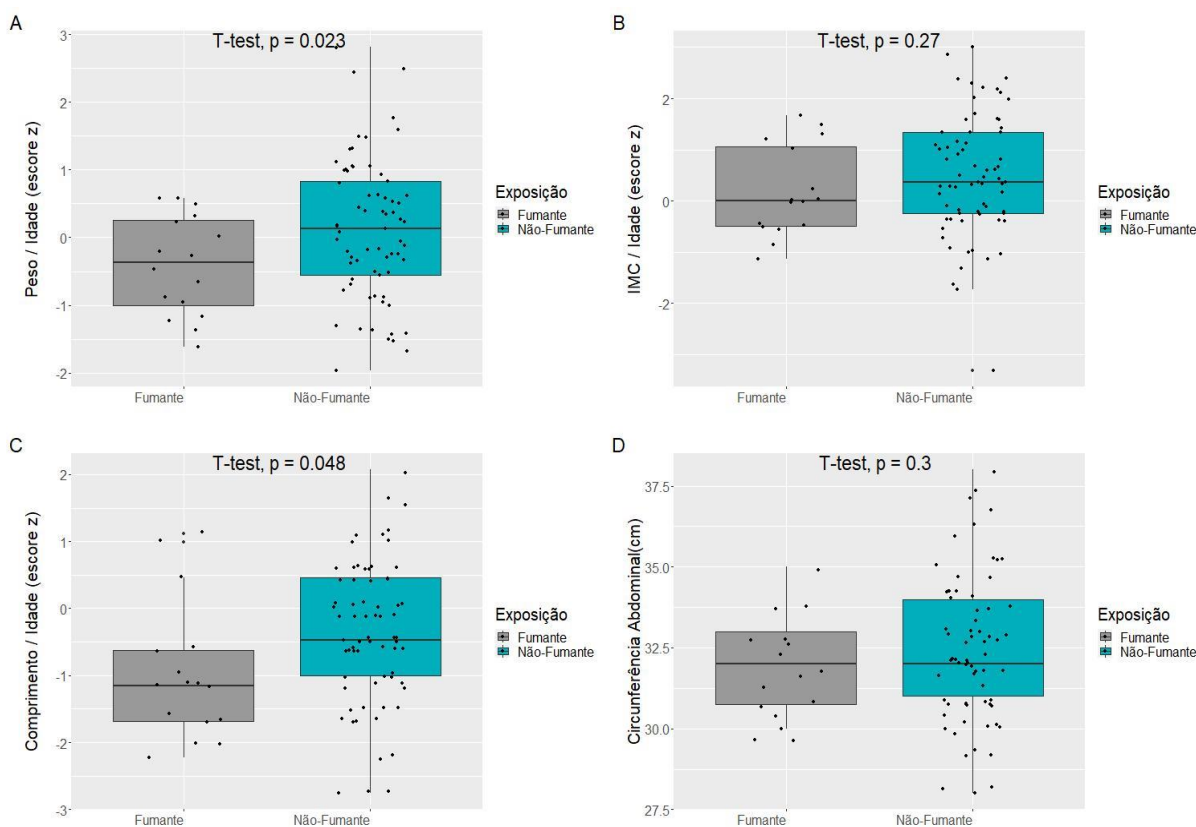


Fonte: Elaborado pela autora.

3.2.5.1 Avaliação do fumo como fator de estresse

Um dos principais fatores de exposição ao estresse na gestação com impactos negativos para a criança descritos na literatura é o fumo. Por essa razão, como utilizamos escores que somavam estressores ambientais (tóxicos e psíquicos), avaliamos se só o hábito de fumar seria responsável pelos desfechos observados nos grupos formados anteriormente, o que faria o fumo o único ou preponderante fator estressor associado com o crescimento fetal. Criamos desta forma, grupos de fumantes ou não-fumantes e as diferenças observadas foram diferentes àquelas encontradas nos grupos de escore de estresse, A e B. O peso e o comprimento, foram menores que o esperado para a idade, mas, quando olhamos para o IMC para idade, que trata da proporção de peso e comprimento, este índice não foi diferente entre os grupos (**Figura 11**). Isto é, apesar das crianças pesarem menos e serem menores, proporcionalmente, não houve alteração, como observada nos grupos A e B.

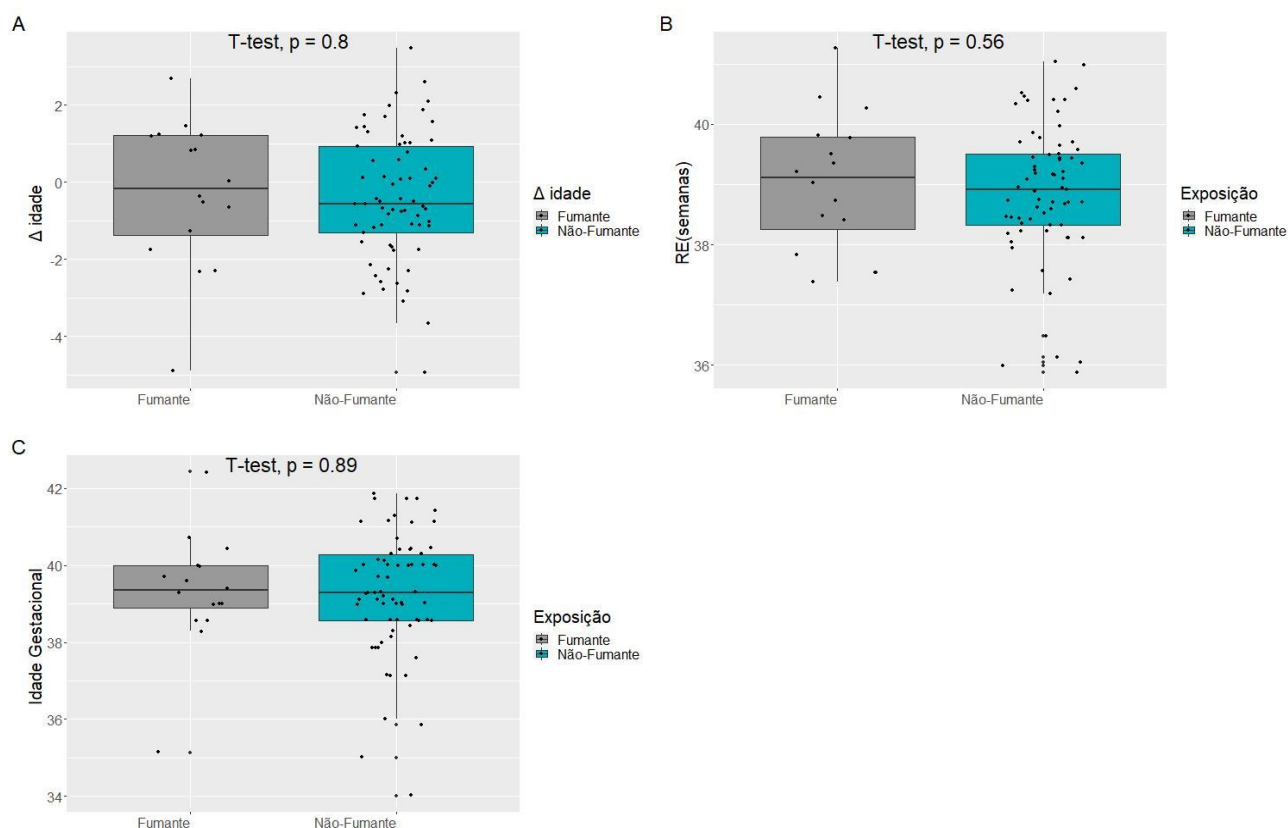
Figura 11 - *Boxplot* dos índices de crescimento e circunferência abdominal de acordo com a exposição ao fumo na gestação. **A)** P/I, sendo peso sobre a idade, **B)** IMC/I, sendo IMC por idade **C)** C/I, sendo comprimento por idade **D)** Circunferência abdominal.



Fonte: Elaborado pela autora.

Ainda no intuito investigativo do fumo como medidor de estresse importante, comparamos o relógio epigenético entre os filhos de mães que fumaram e as não-fumantes. O relógio epigenético não apresentou diferenças entre as fumantes e não-fumantes, assim como não houve diferença da IG (**Figuras 12A-C**). Portanto, essas análises foram aqui apresentadas apenas para apontar que os grupos formados não foram explicados exclusivamente pelo ato de fumar e não serão mais mencionados nesta tese, por não contemplar o objetivo de olhar para cada exposição ao estresse individualmente.

Figura 12 - *Boxplot* das medidas de RE e IG das mães fumantes e não-fumantes. **A)** Delta idade, **B)** RE em semanas, **C)** Idade gestacional.



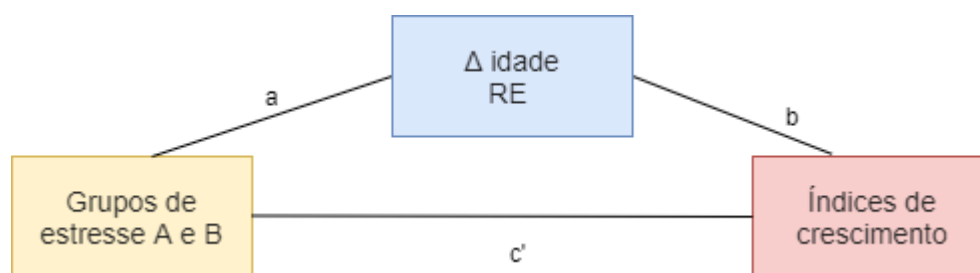
Fonte: Elaborado pela autora.

3.2.6 Relógio epigenético e efeito mediador entre grupos de exposição ao estresse e crescimento ao nascimento

Como nossa hipótese é baseada na metilação do DNA como mediadora das relações entre exposições de estresse pré natal e desfechos de crescimento, investigamos a relação mediadora do RE e do Delta Idade na relação dos grupos de exposição a estresse e os índices de crescimento (**Figura 13**). As análises foram rodadas na linguagem R, pelo pacote *mediation*. A função *mediate* nos dá os efeitos de mediação causais médios (ACME), Efeitos Diretos Médios (ADE), Efeitos Indiretos e Diretos combinados (Efeito Total). O ACME aqui é o efeito indireto de M (efeito total - efeito direto) e, portanto, esse valor nos diz se nosso efeito de mediação é significativo. O ACME, então, informa se dependendo do grupo de exposição ao estresse — maior ou menor — foi possível observar efeito mediador positivo (a probabilidade de maiores escores nos índices de crescimento) ou negativo (diminuição nos escores nos índices

de crescimento) no resultado (tamanho da criança), considerando como mediador a DNAm: RE ou Δ Idade. O efeito total e o ADE (**Tabela 5**) mostram se é provável que o estresse mude os índices de crescimento (controlando para a variável mediadora) nos modelos de mediação. Não foi encontrado papel mediador do RE ou Δ Idade sobre os índices de crescimento. Observamos efeito direto ($\beta = -0.6$, p valor =0,012) e total ($\beta = -0.6$, p valor =0,018) do RE para o índice IMC/I. Ao analisar o delta idade, encontramos um efeito direto ($\beta = -0.6$, p valor =0,012) e total ($\beta = -0.6$, p valor =0,016) sobre o índice P/I e IMC/I ($\beta = -0.4$, p valor =0,048) (Tabela 5).

Figura 13 - Figura esquemática da análise de mediação em que investigou-se o efeito mediador do RE e RE-IG (delta idade) sobre os índices de crescimento IMC/I, P/I, C/I. Ver Tabela 5 para coeficientes.



Fonte: Elaborado pela autora.

Tabela 5 - Análise de mediação entre os grupos de exposição a estresse A e B, relógio epigenético e medidas de crescimento.

	Média do Efeito Direto (ADE)			Média da mediação causal (ACME)			Efeito total		
	Coef.	95%IC	<i>p</i> -valor	Coef.	95%IC	<i>p</i> -valor	Coe f.	95%IC	<i>p</i> -valor
IMC/IDADE									
RE (s)	-0.60	-1.05 -0.13	0.012	0	-0.12 0.11	0.948	-0.60	-1.06 -0.12	0.018
Δ idade	-0.61	-1.07 -0.16	0.012	-1e-03	-0.11 0.12	0.998	-0.61	-1.09 -0.14	0.016
PESO/IDADE									
RE (s)	-4e-01	-8e-01 0	0.056	5e-06	0.31 0.12	0.95	-0.39	-0.83 0.03	0.066
Δ idade	-0.40	-0.79 -0.01	0.048	-2e-04	-0.11 0.12	1.00	-0.40	-0.81 0.01	0.064
COMPRIMENTO/IDADE									
RE (s)	0.04	-0.42 0.49	0.87	0	-0.13 0.13	0.970	0.04	-0.45 0.49	0.880
Δ idade	0.03	-0.45 0.50	0.87	0.00	-0.11 0.13	0.980	0.04	-0.48 0.54	0.880

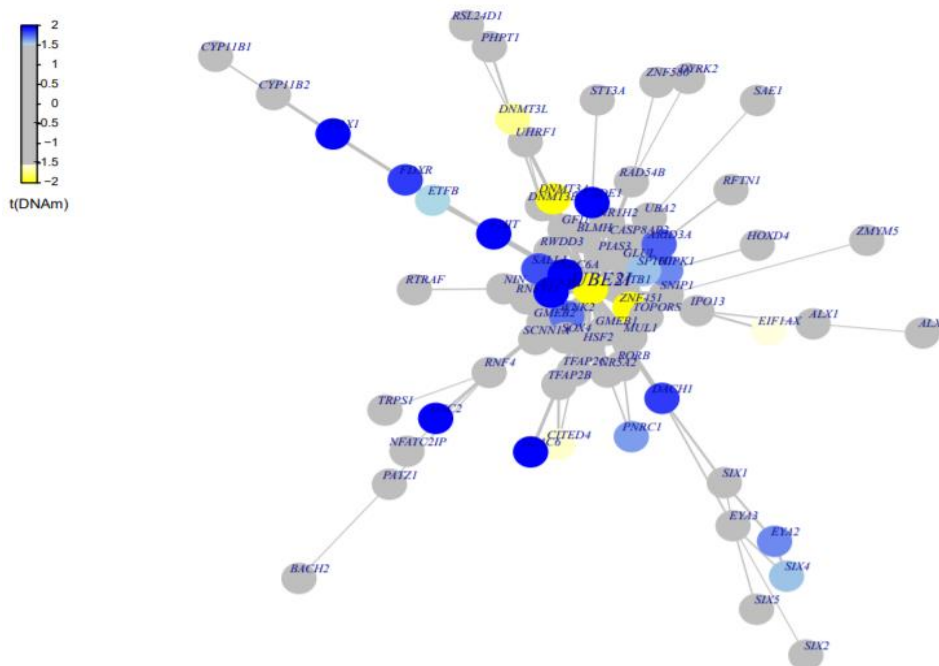
RE: Relógio Epigenético; Δ idade: RE – IG cronológica

Fonte: Elaborado pela autora.

3.2.7 Análise de redes de metilação diferencial entre grupos A e B

Como diferenças de desfecho entre os grupos de exposição a estressores na gestação foram observadas, porém não sem diferenças do RE entre elas, perguntamos então, se haveria redes de interação e de co-metilação diferencial que pudesse direcionar a uma interpretação biológica por intermédio das diferenças epigenéticas dos grupos formados exposição a estressores. Essas redes foram construídas a partir da seleção de sítios com maior variabilidade entre os grupos (p -valor ajustado $< 0,05$). A primeira rede construída pelo FEM, que basicamente cria uma rede com os sítios que estão mapeados próximos a regiões de transcrição (assumindo que a expressão gênica nessas regiões vai silenciar-se se metilado) identificou 5 módulos principais, cujas sementes foram os genes UBE2I, TLR1, TCF7L2, REN, EXOSC2 (Tabela 6). UBE2I foi a semente do maior módulo encontrado nesta análise e participa da ubiquitinação de proteínas, fundamental para inúmeros processos metabólicos envolvidos na degradação proteica no citosol (Figura 14).

Figura 14 - Ilustração gráfica do maior Módulo Epigenético Funcional identificado: UBE2I. As cores variam entre o estado menos metilado (amarelo) ao mais metilado (azul).



Fonte: Elaborado pela autora.

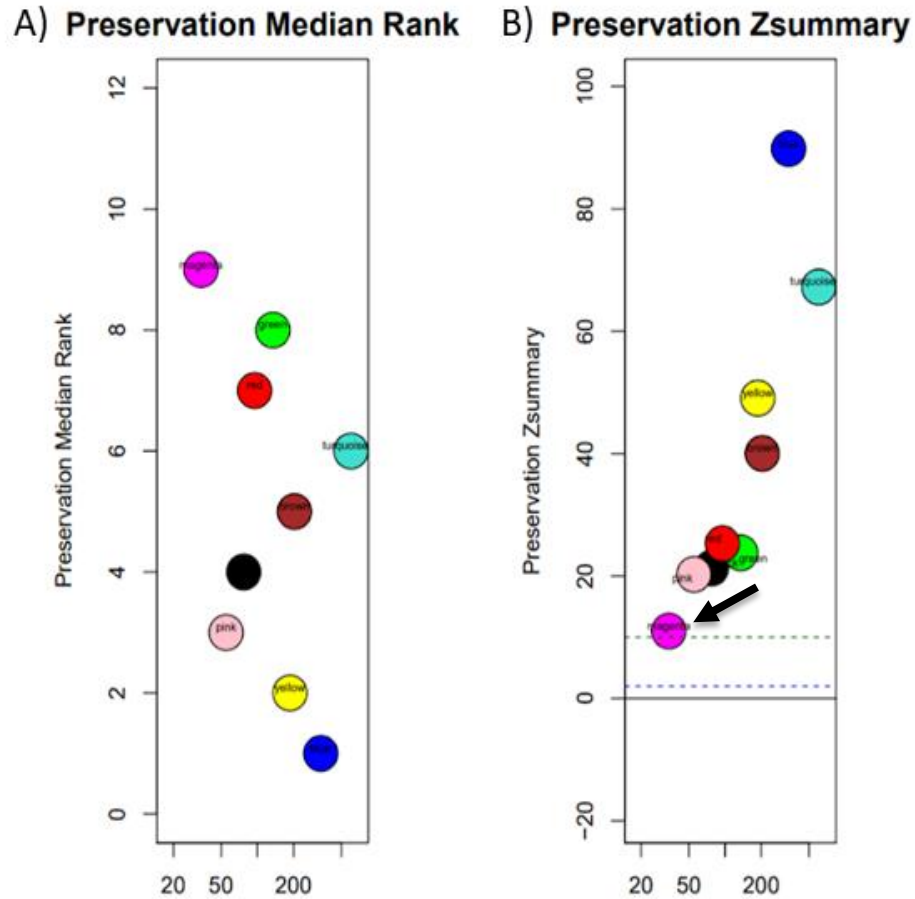
Tabela 6 - Módulos Epigenéticos Funcionais (Redes de interactoma)

Símbolo (Semente)	Tamanho	Mod	p-valor	Genes
UBE2I	74	1.864025	0.001	UBE2I, RAD54B, NR1H2, SATB1, DNMT3B, DNMT3A, ZNF451, RORB, HSF2, NIN, RNF111, SNIP1, GFI1, CSDE1, TFAP2B, RWDD3, GLUL, PIAS3, DACH1, TOPORS, HIPK1, GMEB1, SP100, RNF4, CASP8AP2, PATZ1, ALX1, ALX4, BACH2, EYA3, SIX5, SIX1, SIX4, EYA2, SIX2, ZMYM5, HOXD4
TLR1	31	1.606792	0.028	TLR2, MYD88, IRAK2, LY96, HMGB1, BGN, TIRAP, CD14, TLR6, SAA1, VCAN, TLR9, TLR5, S100A9, S100A8, L1RL1, IRAK3, PELI2, CHGB, SIGIRR, TMED1, TRAM1, RAB17, TLR7, CCN4, PLA2G1B
TCF7L2	12	2.279695	0.002	AXIN2, TNIK, KLF4, TLE2, CTNNBIP1, ADCY7, TCF7, TBXT, HIC1, TCF7L1, PRDM1
REN	21	1.844134	0.003	REN, PCSK5, PCSK1, ATP6AP2, M6PR, PCSK1N, AMH, PMCH, MCHR1, MCHR2, PCSK6, PCSK2, SCG5, CPE, PLIN3, AMHR2, ZMYND19, TRH, TRHR, FZD8, RYK
EXOSC2	18	1.653806	0.034	EXOSC2, AICDA, PARN, EXOSC4, DCP2, PDP1, C1D, MPHOSPH6, EXOSC9, EXOSC3, ZNF598, KHSRP, AKR1A1, KNSTRN, ARHGAP18, A1CF, THOP1, CDK5RAP1

Fonte: Elaborado pela autora.

Para as redes de co-expressão, empregamos o método do WGCNA, cuja proposta inicial do algoritmo foi pensada como auxílio na interpretação biológica de doenças, especialmente as complexas, ou seja, multifatoriais. Investigamos módulos que poderiam não estar conservados entre os grupos A e B (**Figura 15-B**). Não foram encontrados módulos pouco conservados, contudo, o módulo na cor magenta (indicado com uma seta) foi o menos preservado comparados aos demais (ainda assim, fora do ponto de corte para ser considerado pouco conservado =10). Como forma de investigar se haveria uma interpretação para a lista de sítios mapeados em genes presentes no módulo magenta, selecionamos os mais variáveis fizemos enriquecimento por STRING no evidenciou os processos biológicos relacionados a síntese proteica, vias imunológicas e de reconhecimento viral (**Tabela 7**).

Figura 15 - Conservação dos módulos do WGCNA entre os grupos A e B. **A)** Preservação dos módulos em ordem decrescente, de acordo com a métrica medianRank. **B)** Classificação dos módulos a partir da métrica Zsummary. o threshold = 10 de preservação. Nas duas subfiguras, o eixo x mostra o número de sítios pertencentes a cada módulo, e no eixo y, o score de preservação, de acordo com o medianRank e Zsummary. Em relação ao Zsummary, módulo com um score ≤ 10 são considerados pouco preservados.



Fonte: Elaborado pela autora.

Tabela 7 - Tabela de enriquecimento (STRING) dos genes mais variáveis do módulo magenta da análise de co-metilação (WGCNA)

ID Processo	Descrição do termo	Genes na via	FDR
GO:0006413	início translacional	7 de 142	2E-06
GO:0019081	tradução viral	3 de 13	6E-04
GO:0034645	processo biossintético de macromolécula celular	16 de 3518	2E-03
GO:0075525	terminação-reinicialização da tradução viral	2 e 5	4.2E-03
GO:0044271	processo biossintético do composto de nitrogênio celular	15 de 3528	8.4E-03
GO:0075522	início de tradução viral dependente de IRES	2 de 10	1.0E-02
GO:0006446	regulamento de iniciação translacional	3 de 71	1.3E-02
GO:0010467	expressão gênica	15 de 3733	1.3E-02
GO:0002183	iniciação translacional citoplasmática	2 de 13	1.3E-02
GO:0045947	regulação negativa da iniciação translacional	2 de 19	2.4E-02
GO:0044260	processo metabólico da macromolécula celular	19 de 6413	4.1E-02
GO:1901566	processo biossintético de composto organonitrogênio	8 de 1370	4.1E-02

Fonte: Elaborado pela autora.

3.3 DISCUSSÃO DOS DADOS DA ROC

Buscamos, com os dados da ROC, as associações entre as alterações de metilação do DNA (DNAm), exposição a estresse intrauterino e desfechos de crescimento. A obtenção dos valores de DNAm dá a alternativa de obtenção do relógio epigenético (RE), que foi a principal variável de interesse na amostra da ROC. O achado central aqui foi a associação negativa com Δ idade (RE - IG) e os índices de crescimento de peso e comprimento para idade (P/I, C/I), indicando que o RE acelerado pode estar negativamente associado ao crescimento fetal. Nossos resultados mostram ainda que, apesar dos filhos de mães mais expostas ao estresse, quando agrupadas, apresentarem menores escores dos índices IMC/I, essa associação não foi mediada pelo RE ou o descompasso do RE com IG.

Dada a natureza, ao menos parcial, da origem intrauterina de doenças crônicas não-transmissíveis, doenças psiquiátricas e TND; um número robusto de trabalhos têm adotado dados antropométricos, especialmente o peso ao nascimento (PN), como fatores de risco para esses desfechos. Por exemplo, em meta-análise recente, a cada 1 DP aumentado do PN, caiu-se o risco em 14% para desenvolvimento de transtorno de hiperatividade e déficit de atenção(93). Por uma própria questão de factibilidade, boa parte dos estudos epidemiológicos usam o PN como um “*proxy*” da qualidade do desenvolvimento intrauterino. Adotamos aqui, no entanto, os índices de crescimento propostos pela OMS (2006), por já terem peso e comprimento ajustados para sexo e idade. Dessa forma, tivemos duas formas de entender exposição a estressores: por intermédio dos parâmetros de crescimento e por questionário de exposições estressoras ambientais aplicados durante internação hospitalar por motivo do parto.

A nossa hipótese que o relógio epigenético poderia ser um mediador das relações encontradas entre estresse pré-natal e as medidas antropométricas ao nascimento não foi confirmada. Porém, ainda que não tenha um papel mediador, parece que o descompasso do relógio epigenético com a idade gestacional seja um biomarcador de algum estresse fetal, uma vez que a aceleração do RE é negativamente associada ao PN. Interpretar esse dado como reflexo de estresse intrauterino exige uma certa cautela, já que a literatura para o RE de recém-nascidos ainda é escassa. Ademais nossos resultados são uma fotografia de um ponto na linha do tempo (medidas de crescimento ao nascimento), e essas relações podem ser observadas em fases posteriores da vida. Especialmente reconhece-se que uma idade acelerada do RE em indivíduos adultos está associada a desfechos negativos, já que importantes doenças e a própria funcionalidade celular está envolvida com processos de envelhecimento.

Contudo, apesar de não ter um papel mediador, a hipótese que o relógio epigenético é um biomarcador importante foi confirmada com nossos resultados. A história do envolvimento do RE como biomarcador de saúde teve destaque em 2013, quando o pesquisador, Steve Horvath (53), publicou pela primeira vez um algoritmo que seria capaz de inferir a idade de um indivíduo por suas marcas epigenéticas. A partir daí, as evidências e o número de estudos foram crescendo, e incluiu-se uma métrica de Δ idade justamente para entender um possível descompasso naquilo que é estimado/esperado e o estado metilômico que de fato as células se encontram na linha do tempo da vida. Por exemplo, no adulto, um Δ idade de 5 representa um risco de 21% maior de mortalidade e mesmo quando corrigida para uma série de fatores (QI, educação, fatores socioeconômicos, DCNTs), esse risco ainda é 16% maior (57). Ainda, recentemente, Ross e cols, observaram que mulheres mais velhas epigeneticamente, ou seja, com aceleração do RE, tiveram período gestacional encurtado e filhos com menor peso ao nascimento (94). Contudo, discutir envelhecimento ou riscos do RE ainda no início da vida exige uma complexidade e evidências ainda desconhecidas. No entanto, as evidências no adulto e essa escassez de dados do RE em recém-nascidos nos direcionou para investigar aqui se o RE poderia nos orientar ou indicar uma alteração metilômica importante que se relacionasse com o crescimento fetal. Até porque, quando esses algoritmos são criados, parte-se do pressuposto que os sítios selecionados são menos vulneráveis ao dinamismo dos níveis de metilação. Então, é possível pensar que, quando observamos mudanças nesses sítios, elas possam indicar alterações significativas da programação celular.

Em concordância com nossos achados em relação ao descompasso do RE com medidas de crescimento, Girchenko e cols (2017) com grupo amostral de 814 mães e filhos mostrou que a aceleração do RE no nascimento foi associada a menor peso ao nascer, e entre os fatores de risco pré-natais para o RE acelerado, observaram a idade materna > 40 anos, pré-eclâmpsia e morte fetal em gravidez prévia, pré-eclâmpsia materna, tratamento com betametasona pré-natal, menor índice de Apgar de 1º minuto e sexo feminino. Aos 5 anos os filhos de mães diabéticas apresentaram aceleração do RE e maior adiposidade, indicando marcadores de risco cardiovascular. Já a desaceleração do RE, foi associada a diabetes mellitus gestacional tratada com insulina (GDM) em uma gravidez anterior e síndrome de Sjögren (95). Temos ainda os resultados da coorte PREDO, cujas crianças filhas de mães com maior estresse gestacional (sintomas depressivos) apresentaram idade metilômica desacelerada, que, por sua vez, previu prospectivamente problemas totais e internalizantes, mediando parcialmente os efeitos da depressão pré-natal materna nos problemas internalizantes, particularmente em meninos. Comparar nossos achados torna-se um desafio, tanto pelas diferenças de desenhos dos estudos,

quanto pela própria discordância de achados. No entanto, algumas observações serão destacadas, já que a aceleração metilômica pode ser tanto resultado do residual da regressão entre IG e RE, quanto pode ser a subtração aritmética de RE com IG cronológica (como empregamos no nosso estudo). Na coorte PREDO, o Δ idade foi resultado do residual da regressão. Ademais, na nossa amostra não conseguimos encontrar uma associação do RE com nenhuma medida de estresse coletada neste estudo ou com a somatória ou criação de grupos mais ou menos expostos a estressores. Alguns achados sugerem maior ou menor flexibilidade do relógio epigenético dependente do sexo, que apesar de não explorado aqui pelo tamanho amostral, deve direcionar futuros estudos na área (59,60).

Não encontramos associação entre medidas de crescimento ou RE com as exposições a estressores quando fizemos uma somatória das variáveis (fumo, abuso de álcool e drogas, doença psiquiátrica, sentimentos durante a gestação). Nosso interesse nesse sentido não era avaliar uma única exposição a estresse, já que reconhecemos que geralmente as gestantes são expostas a mais de um estressor, e que uma somatória de exposição poderia ser mais significativa que o olhar isolado do dado. Curiosamente, ao criar grupos pela mediana, vimos que mulheres mais expostas a fatores estressores tinham crianças com menores índices de IMC para idade. Isso significa que, quando olhamos para a proporção, considerando o peso e o comprimento, essas crianças expostas a maior estresse pré-natal apresentaram escores mais baixos. Reconhecemos aqui uma série de limitações quanto a criação do escore de estresse, especialmente, sendo um dado retrospectivo e a falta de instrumentos validados na busca por indicadores. Ainda assim, o fato de termos encontrados diferenças no desfecho de crescimento reforça de alguma maneira, um possível maior estresse das mães que apresentaram escores maiores (grupo B).

Dando seguimento ao objetivo de entender as diferenças de DNAm, como análise adicional, avaliamos então se as diferenças nas redes de metilação entre os grupos de exposição a estresse poderiam nos indicar possíveis interpretações biológicas para as diferenças fenotípicas entre os grupos A e B. Criamos uma rede FEM, que tem como princípio criar redes de sítios próximos a regiões transcricionais, onde se interpreta que a metilação está associada a um silenciamento gênico. O maior módulo encontrado nesta análise teve como semente o *UBE2I*, responsável pela ubiquitinação de proteínas, um processo fundamental para marcação e catalisação proteica, além de ser codificador da família de proteínas do pequeno modificador relacionado à ubiquitina (SUMO) fundamentais para regular processos celulares como transporte nuclear, transcrição, segregação cromossômica e reparo de DNA(96). O segundo módulo teve como semente o *TLR1* (do inglês, *Toll-Like Receptor 1*), que é um membro da

família dos receptores *Toll-Like* que desempenha um papel fundamental no sistema imunológico inato e adaptativo. Ademais os receptores Toll-Like parecem desempenhar importantes funções no desenvolvimento do SNC em ratos, sugerindo um papel na patofisiologia do cérebro (97). Além disso, as redes de co-metilação, mostram enriquecimento de vias biológicas de processos básicos de expressão gênica e reconhecimento viral.

Ainda, este estudo apresenta a limitação da ausência de cálculo amostral, que pode ser parcialmente justificada com a existência de outro estudo com n amostral semelhante e com resultados estatisticamente significantes (94). Ademais, o cálculo amostral para estudos de metilação enfrenta uma série de desafios, tais como: (i) a construção da lâmina. Os sítios representados correspondem a aproximadamente 2% do genoma e não foram selecionados randomicamente, por outro lado não existe um único critério de seleção, assim existem sítios muito próximos uns dos outros no genoma que se correlacionam e sítios completamente isolados; existem genes que possuem mais de dez sítios e outros que possuem apenas um, além disso as posições dentro dos genes é variável alguns no promotor outros no corpo do gene. Além disso, como os níveis de metilação são tecido-específicos, não existem padrões para todos eles. Por fim, existem sítios, por exemplo os chamados sítios VMR (*variable methylation region*), associados a SNPs que variam entre indivíduos; (ii) a dinamicidade das marcas epigenéticas e suas variações dentro de uma amostra biológica. Os sinais epigenéticos em uma amostra biológica de um indivíduo representam medidas de frequência de uma população de células (98). A literatura hoje apresenta duas propostas com dados simulados para cálculo de poder em estudos de EWAS, o que não existia quando este estudo foi proposto. Ademais, as propostas foram baseadas na lâmina da Illumina 850K, cuja estrutura foi desenvolvida posteriormente à 450K (99,100).

Concluimos com esse dado que o descompasso entre a idade metilômica e a idade cronológica pode ser um indicador de estresse pré-natal uma vez que se associou com índices de crescimento ao nascimento, contudo, não parece mediar as associações de estresse materno e desfechos de crescimento ao nascimento. Mais estudos longitudinais são necessários para a compreensão em longo prazo para as associações de descompassos entre a idade celular estimada e a cronológica. Os dados deste estudo também corroboram com os achados que mulheres mais expostas a estressores ambientais têm maior risco de comprometimento de crescimento fetal, acarretando maior risco de descompassos metabólicos na linha do tempo da vida.

4 CAPÍTULO 4: ESTUDO DE INTERVENÇÃO: ENSAIO CLÍNICO RANDOMIZADO PRIMEIROS LAÇOS

4.1 MATERIAL E MÉTODOS

4.1.1 Desenho do Ensaio Clínico Randomizado (RCT)

O Programa de Visitação para Jovens Gestantes, “Primeiros Laços” (ID clinical trials.org, 38350114.0.0000.007638350114.0.0000.0076), tem coordenação científica liderada pelo Prof. Guilherme Polanczyk (Instituto de Psiquiatria, FMUSP). O tamanho amostral foi calculado com base em diferenças esperadas de resultados de eletroencefalograma (EEG) aos seis meses, segundo dados da literatura. O Primeiros Laços é um programa de visitas domiciliares realizadas por enfermeiras treinadas às adolescentes grávidas pela primeira vez e aos seus bebês, a partir das primeiras 16 semanas de gravidez até os 24 meses de idade. Os visitantes eram enfermeiras especializadas em saúde materna ou mental e eram acompanhados semanalmente por enfermeiras seniores e psicólogos.

O RCT Primeiros Laços é baseado em três referenciais teóricos(101–103) (teoria do apego, teoria da autoeficácia, teoria do desenvolvimento bioecológico) e foi estruturado em cinco eixos (cuidados de saúde, ambiente de saúde, parentalidade e apego, rede de apoio social e familiar e projeto de vida). O programa foi desenvolvido pela equipe de coordenação científica (104) com base em programas existentes de visita domiciliar (105,106). O plano de cuidados para cada díade mãe-bebê foi cuidadosamente elaborado para fortalecer as competências maternas para um cuidado atencioso e responsivo. Um fator chave foi o estabelecimento de relacionamentos positivos entre os visitantes domiciliares e a família. Durante as visitas, as enfermeiras ajudaram os pais a desenvolver interações centradas na criança, melhorar seu vínculo com seu filho, refletir sobre sua própria história de apego e paternidade que receberam, considerar seu filho como um indivíduo com suas próprias necessidades, sentimentos e pensamentos, e melhorar suas habilidades parentais por modelagem. As visitantes incentivaram a maternidade atenta e estimularam comportamentos sensíveis, como estar atento aos sinais comunicativos da criança ou seguir a liderança da

criança. Os pais também receberam apoio para pensar de forma reflexiva. A frequência das visitas era semanal (primeiro / último mês de gravidez / puerpério), quinzenal (gestação / 2-20 meses de idade da criança) e mensal (21-24 meses de idade da criança), compondo um total de 60-62 visitas domiciliares, em um grupo inicial de 80 gestantes, com desfechos do neurodesenvolvimento de 63 crianças. O grupo controle foi composto pelas gestantes que receberam o pré-natal usual nas Unidades Básicas de Saúde próximas às respectivas residências, disponível pelo Sistema Único de Saúde.

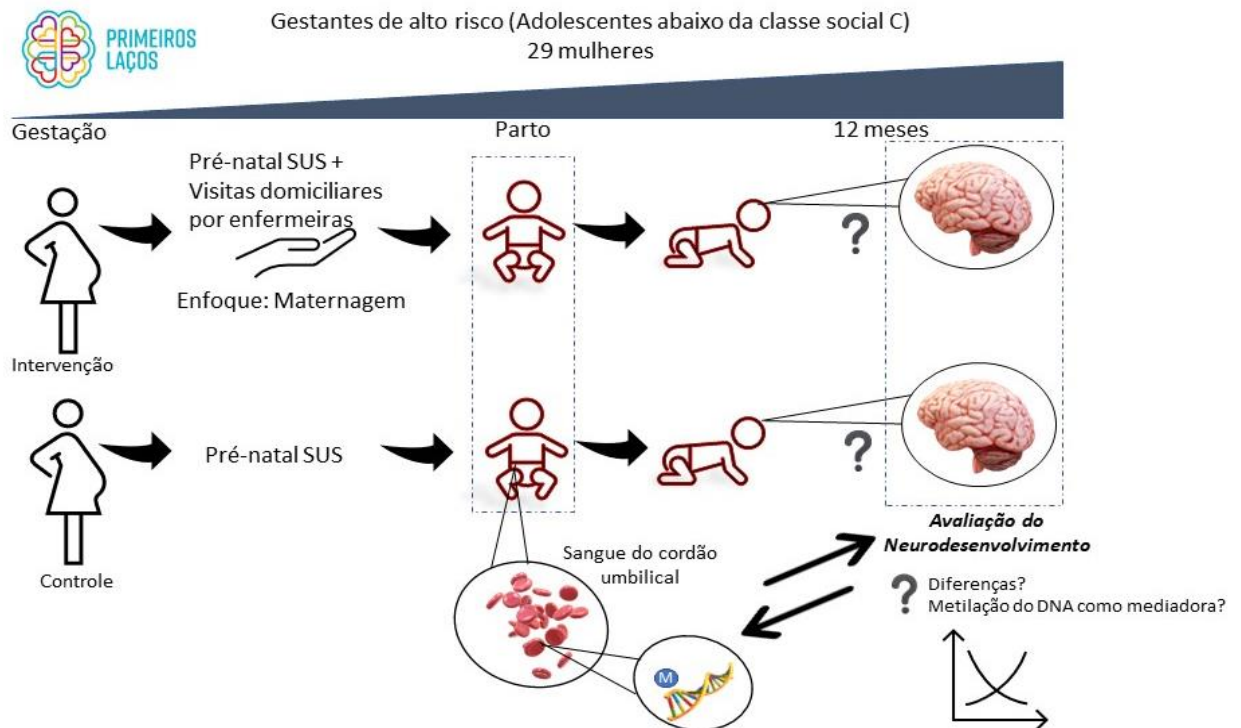
4.1.2 Aspectos éticos

Conforme determinações estabelecidas na resolução 466/12 do Conselho Nacional de Saúde, o presente estudo seguiu os princípios enunciados do Comitê de Ética em Pesquisa da Faculdade de Medicina da USP e Plataforma Brasil (CAAE): 41573015.0.0000.0065 (ANEXO). Continuidade: 41573015.0.3001.0086 (ANEXO 2).

4.1.3 Seleção amostral: Coleta de sangue de cordão

Com o intuito de máxima captação, as gestantes receberam um cartão que foi anexado à caderneta de pré-natal. No cartão do projeto havia o telefone para contato das pesquisadoras para que fossem informadas pela própria gestante ou acompanhante quando as gestantes entrassem em trabalho de parto. Ademais, dentro do hospital também havia um treinamento e um cartão informativo para comunicação com as pesquisadoras caso a gestante não tivesse conseguido ou tivesse se esquecido de comunicar às pesquisadoras responsáveis pela coleta da placenta e cordão umbilical. Assim que recebiam este aviso, as pesquisadoras deste projeto iam ao centro obstétrico e realizavam a coleta do material biológico. A **Figura 16** esquematiza a seleção amostral e objetivos do estudo.

Figura 16 - Seleção da amostra desta tese, proveniente do Ensaio Clínico Randomizado Primeiros Laços.



Fonte: Elaborado pela autora.

4.1.4 Avaliação das medidas de exposição a estresse pré-natal do estudo Primeiros Laços

Diferentes instrumentos validados foram utilizados para obtenção dos dados de estresse pré-natal. As gestantes foram avaliadas em dois momentos: 1º) Na inserção da gestante no estudo (*baseline*), 2º) Na 30ª semana gestacional, com o emprego dos mesmos instrumentos em cada abordagem (exceto dados socioeconômicos). Para avaliação de dados socioeconômicos utilizou-se o Critério de Classificação Econômica Brasil – ABEP (107), a Escala Brasileira de Segurança Alimentar (EBIA) (108), renda familiar e educação. Para avaliação psíquica foram utilizadas o Inventário de Depressão e o de Ansiedade de Beck (BDI, BAI)(109,110) e Questionários sobre Traumas na Infância (CTQ)(111). Foram coletados dados de doenças prévias e/ou atuais e abuso de álcool/drogas, hábito de fumar e, no momento da entrevista, peso e altura materna. Os questionários foram aplicados por psicólogas contratadas e previamente treinadas.

4.1.5 Coleta de amostra biológica (sangue do cordão umbilical - SCU) e extração do DNA

Os procedimentos foram idênticos à metodologia da ROC apresentada em materiais e métodos do **capítulo 3**.

4.1.6 Análise de metilação

A análise de metilação foi idêntica à metodologia da ROC apresentada no capítulo 3 desta tese. Brevemente, o pacote Minfi versão 1.26.2 (112) foi usado para controle de qualidade das amostras, bem como para aplicar a normalização de quantis (113) e para calcular valores de M (log do valores Beta). Removemos as sondas que continham SNPs no local CpG, aquelas localizadas nos cromossomos sexuais e aquelas conhecidas por co-hibridizarem com mais de uma sequência genômica, resultando em 429.174 sondas. O pacote ChAMP versão 2.10.2 (114) foi usado para identificar e ajustar os valores M para efeitos de lote e para sexo. A composição celular da amostra foi avaliada usando o pacote FlowSorted.CordBlood.450k implementado no minfi, mas nenhuma diferença significativa foi encontrada ao comparar os grupos de intervenção e controle.

4.1.7 Análise de metilação diferencial

Para a análise de metilação diferencial (comparação de metilação entre grupos de controle e intervenção), foram identificados inicialmente os sítios diferencialmente metilados (DMPs, do inglês, *Differentially Methylated Positions*), obtido por modelo linear Bayesiano do pacote limma (Ritchie et al., 2015) a partir dos valores de M. Considerados como DMPs os sítios com teste Benjamini-Hochberg $<0,05$. O pacote “*DMRcate*” foi utilizado para obtenção das DMPs e de Regiões Diferencialmente Metiladas (DMRs, do inglês, *Differentially Methylated Regions*), conforme implementado no ChAMP. Depois de identificar os DMPs, os sítios foram posteriormente anotados para obter informações sobre os *enhancers*, sítios de ligação do fator de transcrição (TFBSs, do inglês *intranscription factor binding site*) e sítios de hipersensibilidade a DNase I (DHSs, do inglês DNase I hypersensitive sites). Para mapear para *enhancers*, usamos as anotações completas de EnhancerAtlas para “*Fetal Brain*” e “*Fetal Placenta*”. Para TFBSs e DHSs, usamos *wgEncodeRegDnaseClusteredV3* e

wgEncodeRegTfbsClusteredV3 do *Genome Browser's Data Integrator* (<http://genome.ucsc.edu/cgi-bin/hgIntegrator>).

Foram consideradas sequências de DMRs de 300 bp contendo pelo menos sete sítios CpG e apresentando diferenças com $FDR \leq 0,05$. Para identificar o trofoblasto putativo, foram utilizados vários recursos de bioinformática disponíveis publicamente. O navegador do genoma UCSC foi usado para encontrar trofoblastos conservados em HMR e a proximidade de regiões ENCODE trofoblastos putativas aos sítios CpG ensaios via sequenciamento de chip. A presença de trofoblastos putativos flanqueando um determinado local CpG foi ainda avaliada usando PhysBinder [33], com o nível de limite definido para precisão média, e MotifMap

4.1.8 Medidas antropométricas maternas e da criança

Para as medidas antropométricas maternas, o peso pré-gestacional foi referido pela mãe. Para a aferição de estatura, o estadiômetro portátil SECA 213 foi utilizado.

As medidas antropométricas ao nascimento, peso, comprimento, circunferência cefálica, torácica e abdominal seguiram o protocolo adotado no Hospital Universitário da USP e foram obtidas por prontuário médico. Aos 12 meses o peso foi obtido por balança pediátrica Seca (precisão 10g). O comprimento da criança foi mensurado no infantômetro (precisão de 0,1 cm). Para o diagnóstico do estado nutricional, utilizou-se os critérios da Organização Mundial da Saúde (OMS) de 2006, de acordo com a classificação dos índices antropométricos de escore Z de peso para idade (P/I), de estatura para idade (C/I) e de índice de massa corporal para idade (IMC/I) (Onis, De, 2006; WHO, 1995; World Health Organisation (WHO), 1995).

4.1.9 Medidas de neurodesenvolvimento aos 12 meses de vida

Finalmente, para o desfecho principal desta tese, as Escalas Bayley III de Desenvolvimento Infantil (BSID-III) (115,116) foram conduzidas por psicólogos treinados que avaliaram as características do neurodesenvolvimento dos bebês aos 12 meses de idade. Os cinco domínios do neurodesenvolvimento foram considerados separadamente: Cognitivo, Linguagem Receptiva, Linguagem Expressiva, Habilidade Motora Fina e Habilidade Motora Grossa.

4.1.10 A análise de rede de correlação ponderada – WGCNA

Para as redes de co-expressão, adotamos o algoritmo de análise de rede de correlação ponderada o (WGCNA, do inglês Weighted Correlation Network Analysis) que é um método amplamente utilizado para análises envolvendo expressão gênica (microarranjos ou RNA-seq) e também de metilação de DNA(91,92). Brevemente, a entrada de dados é o valor de metilação (valores M) dos sítios de dois grupos (por exemplo, caso e controle), processados em separado. Para cada um dos dois grupos, o algoritmo calcula a Correlação de Pearson entre os valores M de todos os pares de sítios. O WGCNA irá criar uma aresta entre dois sítios se eles apresentaram uma altíssima correlação (ex. 0.9), criando assim uma rede de co-metilação. para construir matrizes de correlação de pares de co-metilação/expressão (uma para cada grupo). O Este algoritmo assume que as redes de co-metilação/expressão gênica apresentam topologia de uma rede livre de escala (Scale-Free). Para isso, então os valores de correlação são elevados a potência β (valores naturais crescentes) (β), a fim de eliminar arestas da rede que não atendam a um determinado limiar e mantendo somente as arestas originárias das mais altas correlações. A rede final é aquela resultante do valor de β que atinge maximiza um índice mínimo de independência livre de escala (o coeficiente de determinação da linha de regressão que melhor se aproxima da distribuição do grau do nó em escala log-log). Assim, os dados de metilação expressão dos grupos de caso e controle são agrupados usando medida de sobreposição topológica (TOM) e os módulos são definidos como ramos da árvore agrupada resultante. Até este ponto, os grupos de caso e controle são processados em paralelo. Finalmente, o WGCNA usa os valores de TOM de ambos os grupos para uma análise de preservação de módulo, onde indica a preservação dos módulos nas redes de controle e de caso (com base nos sítios/genes contidos em cada módulo). O resultado é uma lista de módulos (e seus respectivos sítios/genes) classificados por uma pontuação de preservação (92).

4.1.11 Análise de enriquecimento

Selecionando os DMPs mapeadas em genes, usamos o pacote *clusterprofiler* (117) para avaliar o enriquecimento para a Ontologia Genética (Processo Biológico, Componente Celular e Função Molecular) e vias de sinalização (KEGG). Também aplicamos a *Over Representation Analysis (ORA)* em genes únicos usando o genoma humano como *background*. Um valor de $FDR \leq 0,05$ foi usado como limite de significância.

4.1.12 Relógio epigenético para idade gestacional no nascimento

Para calcular a idade do DNAm (55), os níveis de metilação de 148 locais CpGs diferencialmente metilados foram submetidos a uma regressão de rede elástica baseada em seis conjuntos de dados de treinamento, descritos em métodos do capítulo 3 desta tese.

4.1.13 Análise Estatística

As análises estatísticas foram realizadas em R v. 3.5.0. Os testes qui-quadrado e t de *Student* foram aplicados para comparar os fenótipos entre os grupos intervenção e controle. Para realizar a análise de mediação, definimos os seguintes critérios no local, região ou nível do módulo CpG: (1) Apenas DMPs associados à cognição foram testados como mediadores. DMPs relacionados à cognição ($R^2 \text{ adj} > 0,4$) foram definidos por um modelo de regressão linear usando valores M de DMPs como variável dependente e o escore do domínio cognitivo e sexo como variáveis independentes. (2) A partir de dados de co-expressão, todos os locais CpG que eram hubs do módulo de co-metilação não preservado. (3) Cada DMR foi representado como as médias dos valores M de todos os locais CpG dentro da região específica. (4) Relógio epigenético como valores brutos.

Usamos o pacote de *mediation* versão 3.4.4 (118) para analisar a mediação causal com bootstrap de 10.000 replicações para intervalos de confiança (IC). Os dados do modelo de mediação para cada CpGs ou DMRs incluem estimativas e ICs de 95% dos efeitos total, direto e de mediação. O Average Causal Effects (ACME) afirma se a intervenção afeta positivamente (a probabilidade de aumentar o escore cognitivo) ou negativamente (diminui o escore cognitivo) o resultado (escore cognitivo) por meio de um mediador (DNAm, de acordo com os sítios selecionados). O efeito total e o efeito direto médio mostram se a intervenção tem probabilidade de alterar a cognição nos modelos de mediação. A função *medsens* foi aplicada para realizar a análise de sensibilidade para testar os efeitos de fatores de confusão não medidos.

4.2 RESULTADOS PRIMEIROS LAÇOS

Os resultados apresentados abaixo resultaram em produção de um artigo científico em fase de submissão (ANEXO 3)

4.2.1 Características da amostra do Ensaio Clínico Randomizado “Primeiros Laços”

Coletamos sangue do cordão umbilical (SCU) de 32 mães / bebês para o estudo. Das 32 crianças com coleta do SCU, três não realizaram o teste de neurodesenvolvimento e foram excluídas das análises, resultando em 13 no grupo controle e 16 no grupo intervenção, com total de 29 crianças. Uma explicação breve, para exemplificar dificuldades inerentes a todo projeto de pesquisa: tínhamos como a nossa proposta inicial também avaliar a expressão gênica da placenta dessas mulheres e fazer uma integração desses dados, para tanto, também realizamos coleta de tecido placentário de face fetal e materna. Contudo, a análise de controle de qualidade do tecido estava abaixo do ideal para obtenção da expressão gênica e, portanto, não houve análise posterior desse dado. Quanto à caracterização dos grupos controle e intervenção, não houve diferenças quanto às medidas socioeconômicas, tabagismo, IMC, escalas psiquiátricas entre os grupos de mães como esperado pela seleção randômica, e não houve diferenças quanto às medidas antropométricas nem diferenças obstétricas ou antropométricas ao nascimento. Aos 12 meses, o grupo intervenção apresentou escores mais altos nos domínios motor e cognitivo brutos (**Tabela 8**).

Como esse grupo amostral é proveniente do RCT, caracterizando-se como subamostra ($n = 80 \times 29$), comparamos fenótipos maternos e infantis entre todos os participantes do RCT e aqueles com dados do sangue do cordão umbilical, que compreendem a amostra desta tese. Fizemos isso justamente para verificar se, apesar do n amostral reduzido, a subamostra representa as características do estudo de origem. Encontramos um escore maior do Inventário de Ansiedade de Beck ($p = 0,04$) e um número maior de meninos ($p = 0,01$) (**Tabela 9**), no entanto, não houve diferenças nas proporções entre intervenção e controle. Assim, embora os domínios cognitivos e de motor grosso tenham sido diferentes entre os grupos intervenção e controle nesta subamostra, apenas analisamos o escore cognitivo como resultado, porque este foi concordantemente diferente em toda a coorte do RCT. Ressaltamos aqui que olhamos para o dado de intervenção aos 12 meses de vida e que objetivo desta tese não foi avaliar a efetividade da intervenção, e sim, o papel mediador da DNAm na relação entre intervenção e desfecho. Dados robustos dos resultados da efetividade da intervenção Primeiros Laços serão

publicados posteriormente. A **Tabela 10**, apresenta as 63 avaliações de neurodesenvolvimento aos 12 meses de idade e compara com as crianças analisadas aqui (n=29) (**Tabela 10**). O desenho do estudo pode ser visualizado na **Figura 17**.

Tabela 8 – Características psíquicas, sociais e escala de neurodesenvolvimento das participantes do RCT e respectivos filhos com coleta de sangue do cordão umbilical. São Paulo, SP.

		Controle (n=13)		Intervenção (n=16)		
Recrutamento						
Mães	n	%	média (DP)	n	%	média (DP) P valor
IMC pré-gestacional (kg/m ²)			21.2(3.9)			22.5(6.7) 0.527
ABEP (score)			18.1 (6.4)			17.6 (5.6) 0.825
Inventário de Depressão de Beck - BDI			14.0 (6.9)			13.3 (6.3) 0.804
Inventário de Depressão de ansiedade – BAI			10.4 (5.9)			16.1 (9.2) 0.179
Renda familiar (Salário mínimo)			4.7(1.6)			4.4 (1.4) 0.626
Ao nascimento						
Idade gestacional (semanas)			39.3 (1.2)			39.4 (1.4) 0.989
Peso ao nascer (g)			3120.3 (365.4)			3168.2 (409.2) 0.728
IMC/Idade (z-scores)			-0.231 (0.9)			0.109 (0.9) 0.320
Neurodesenvolvimento (12 meses)						
Sexo (meninas)	6	42.9		8	50.0	0.964
Bayley III						
- Cognitivo			41.5 (3.0)			44.2 (4.0) 0.025
- Motor Fino			28.2 (2.2)			30.1 (2.8) 0.051
- Motor Grosso			41.2 (4.0)			44.8 (4.1) 0.026
- Linguagem expressiva			13.1 (3.6)			15.3 (3.4) 0.111
-Linguagem Receptiva			12.5 (2.3)			13.5 (4.5) 0.429

Fonte: Elaborado pela autora.

Tabela 9 – Tabela comparativa das características de mães e respectivos filhos que tiveram coleta do cordão umbilical e análise de metilação e neurodesenvolvimento aos 12 meses com a amostra inicial.

	RCT amostra inicial (n=80)			Amostra analisada (n=29)			
Momento de inclusão no estudo (dados maternos)							
	n	%	Média (DP)	n	%	Média (DP)	P valor
IMC pré-gestacional (kg/m ²)			24.7 (3.6)			25.3 (4.5)	0.718
ABEP (pontuação)			18.2 (5.8)			18.3 (4.4)	0.937
Inventário de Depressão de Beck - BDI			11.3 (7.0)			13.6 (3.3)	0.192
Inventário de Depressão de ansiedade – BAI			8.9 (6.4)			12.4 (8.0)	0.043
Escala de Neurodesenvolvimento (12 meses)							
Sexo (meninas)			19 (51.3)			14 (43.8)	0.012
Bayley III							
- Cognitivo			42.7 (3.9)			43.2 (4.6)	0.627
- Motor Fino			28.9 (2.5)			30.0 2.6)	0.734
- Motor Grosso			42.3 (5.6)			43.2 (4.1)	0.481
- Linguagem expressiva			12.7 (2.3)			13.8 (3.6)	0.661
-Linguagem Receptiva			13.8 (3.3)			14.3 (3.2)	0.583

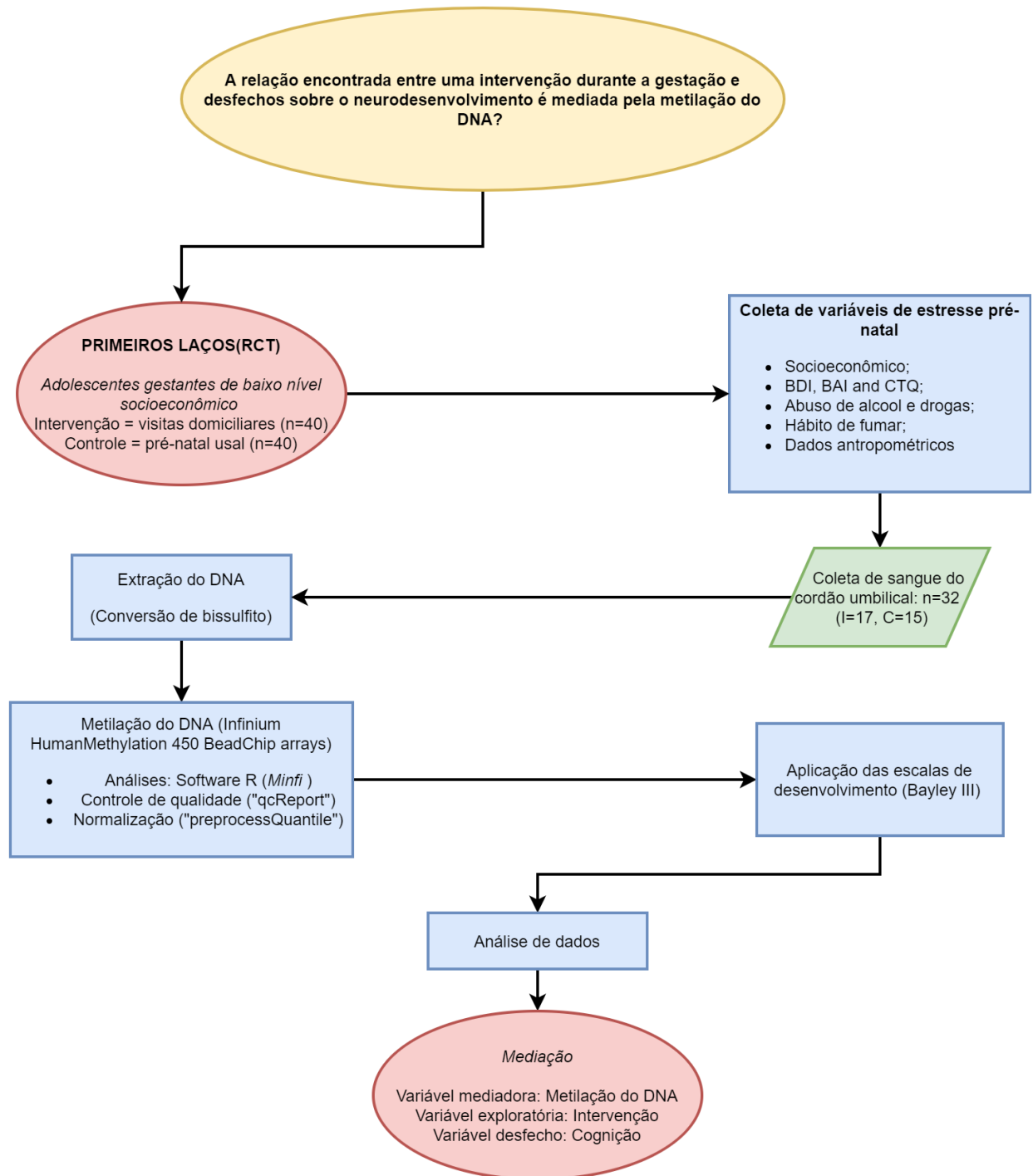
Fonte: Elaborado pela autora.

Tabela 10 – Tabela comparativa com a amostra inteira do RCT (n inicial=80), e (n com desfechos do neurodesenvolvimento=63).

		Controle		Intervenção			
Entrada no estudo (dados maternos)							
	n	%	Média (DP)	n	%	Média (DP)	P valor
IMC pré-gestacional (kg/m ²)			22.2 (3.3)			23.0 (4.7)	0.527
ABEP (score)			18.5 (5.8)			17.8 (4.4)	0.825
Inventário de Depressão de Beck - BDI			11.7 (7,0)			12.8 (8.0)	0.534
Inventário de Depressão de ansiedade – BAI			9.4 (6.4)			11.4 (8,3)	0.244
Características do neonato (ao nascer)							
Idade gestacional (semanas)			39.3 (1.2)			39.3 (1.4)	0.989
BMI/age (z-scores)			-0.24 (0.9)			-0.00 (0.9)	0.457
Escala de Neurodesenvolvimento (12 meses)							
Bayley III							
- Cognitivo			41.6 (2.3)			44.4 (3.6)	0.015
- Motor Fino			28. 8 (2.6)			30.0 (2.5)	0.075
- Motor Grosso			41.2 (5.5)			44.8 (4.1)	0.089
- Linguagem expressiva			13.1 (3.7)			15.3 (3.0)	0.111
Linguagem Receptiva			13.0 (2.3)			15.1 (3.6)	0.011

Fonte: Elaborado pela autora.

Figura 17- Fluxograma do estudo com as amostras do Ensaio Clínico Randomizado, Primeiros Laços. Amostras de sangue do cordão umbilical (n = 32) foram coletadas dos participantes do Primeiros Laços (n=80). A metilação do genoma foi aplicada usando as matrizes *Infinium HumanMethylation 450k BeadChip* que foi submetida às etapas de controle de qualidade, normalização e correção para efeitos de lote. Os parâmetros do desenvolvimento neurológico foram obtidos pela avaliação de Escala de desenvolvimento Bayley III. A análise final para resposta da pergunta foi a mediação causal, tendo como variável desfecho o escore bruto do domínio cognitivo, metilação do DNA como variável mediadora e a intervenção como a variável exploratória.

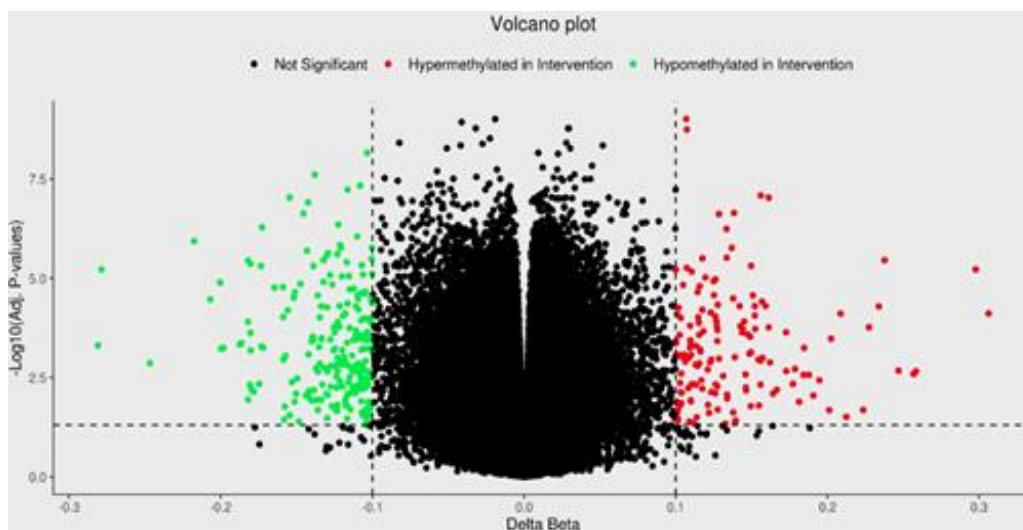


Fonte: Elaborado pela autora.

4.2.2 Identificação de diferenças de metilação entre intervenção e controles

Ao comparar os níveis de DNAm do SCU entre os grupos intervenção e controle, encontramos 3.090 DMPs, sendo 1.855 (60%) sítios CpG hipometilados no grupo de intervenção (**Figura 18**). Os 3.090 DMPs foram localizados em 1.267 genes, sendo 217 mapeados dentro de 1.500 pares de base do local de início da transcrição. Dezenove DMPs apresentaram delta-beta > 20%, seis mapeados para genes previamente associados ao neurodesenvolvimento como IL18RAP, OR2L13 e CACNA1A (**Tabela 11**). Todos os processos percorridos na análise diferencial e seleção das variáveis para análise de mediação final estão representados na **Figura 19** e serão detalhados a seguir.

Figura 18 - Gráficos de Vulcão mostrando os níveis de metilação de sítios CpG nos grupos de intervenção e controle. Os eixos X e Y exibem, respectivamente, o valor delta-beta (diferenças de metilação entre os grupos) e o log dos valores p ajustados para cada sítio CpG. Pontos verdes (hipometilados) e vermelhos (hipermetilados) identificam os locais CpG com adj P <0,05 e > 5% de diferença de metilação entre os grupos de intervenção e controle.



Fonte: Elaborado pela autora.

Tabela 11 - Genes com maior Delta Beta (>20%) previamente associados ao neurodesenvolvimento

CpG	p-valor ajustado	Delta Beta	CHR	Gene	Relação gene	Relação genoma	DHS	Enhancer	Gene previamente descrito com associações com o neurodesenvolvimento
cg26191747	1.43E-07	-0.28	10	<i>MSMB</i>	TSS200	<i>opensea</i>	NA	NA	
cg04548204	5.65E-07	-0.22	12	<i>KLRG1</i>	3'UTR	<i>opensea</i>	NA	NA	
cg24891846	9.15E-07	0.24	19	<i>CACNA1A</i>	Body	<i>island</i>	NA	Sim	Genome Biol. 2012 Jun 15;13(6): R43
cg04985582	5.53E-06	0.23	21	<i>ABCC13</i>	TSS200	<i>opensea</i>	NA	Sim	
cg00413089	6.07E-06	-0.2	7		<i>IGR</i>	<i>island</i>	Sim	NA	
cg21177396	6.32E-06	0.3	9		<i>IGR</i>	<i>opensea</i>	NA	Sim	
cg12744031	8.97E-06	-0.21	7		<i>IGR</i>	<i>island</i>	Sim	NA	
cg11076954	1.20E-05	0.21	17	<i>SLC43A2</i>	Body	<i>shore</i>	NA	Sim	Transl Psychiatry. 2019;9(1):152
cg09329516	2.00E-05	0.23	17	<i>SLC43A2</i>	Body	<i>shore</i>	NA	Sim	Transl Psychiatry. 2019;9(1):152
cg04131969	7.10E-05	0.31	2	<i>MYADM</i>	Body	<i>shore</i>	NA	NA	PNAS. 2002 Nov 26; 99(24): 15729–15734*
cg03938978	0.000195	-0.28	2	<i>IL18RAP</i>	Body	<i>opensea</i>	NA	NA	J Psychiatr Res. 2016 Mar; 74:10-6.
cg15070894	0.000288	0.2	6	<i>HCG4P6</i>	TSS1500	<i>island</i>	NA	NA	BMC Med Genomics 12, 105 (2019).
cg04028570	0.000423	-0.25	1	<i>OR2L13</i>	TSS200	<i>island</i>	Sim	NA	Front Neurol. 2015; 6: 107.
cg26536949	0.000985	0.26	17		<i>IGR</i>	<i>island</i>	NA	Sim	
cg12036633	0.001135	0.25	15		<i>IGR</i>	<i>opensea</i>	Sim	Sim	
cg05528899	0.001413	0.26	17		<i>IGR</i>	<i>island</i>	NA	Sim	
cg01941243	0.008665	0.2	1		<i>IGR</i>	<i>shelf</i>	NA	NA	
cg19384241	0.011159	0.22	2		<i>IGR</i>	<i>opensea</i>	NA	Sim	
cg24851651	0.019317	0.21	11	<i>CCS</i>	Body	<i>shelf</i>	NA	NA	

* O achado dos autores foi com o gene *MYADM*. Nossa análise apontou diferenças no pseudogene do *MYADM*

Fonte: Elaborado pela autora.

Ainda, como o desfecho principal neste estudo foi o domínio cognitivo, submetemos a lista das 3090 DMPs quanto a sua associação potencial e poder de explicação (R^2 ajustado) com o escore cognitivo dessas crianças aos 12 meses de idade. A regressão linear múltipla foi controlada por sexo, tendo o escore do domínio cognitivo como desfecho de interesse e o valor de cada sítio como variável independente. Dezoito sítios CpG apresentados R^2 ajustado $> 0,4$ (**Tabela 12**), sendo 11 CpG mapeados para os genes: C10orf118, C11orf58, C6orf174, GRIA1, LHFPL4, LOC100240726, MYOM2, RB1, STXBP6, TBL1XR1 e TOR1AIP2. Para alinhar a interpretação dos dados das CpGs não anotados para um gene, usando a anotação do ENCODE (Enciclopédia de Elementos do DNA, acesso em: <https://www.encodeproject.org/>), encontramos uma CpG (cg19610750) localizado em um sítio de hipersensibilidade à DNase que foi anotado para múltiplos fatores de transcrição, incluindo CTCF.

Tabela 12 - Identificação das DMPs associadas ao domínio cognitivo (R ajustado $> 0,4$).

DMP	Gene	adj.P.Val	delta Beta	CHR	L. Gene	I. genoma	Enhancer	P.valor Reg	R ² aj.
cg01600222		0.002	0.07	6	IGR	shore	NA	1.50E-05	0.49
cg03648020		0.001	-0.07	1	IGR	shelf	NA	4.70E-06	0.54
cg03930153	TBL1XR1	0.003	-0.07	3	5'UTR	opensea	TRUE	6.60E-05	0.43
cg05405872	C11orf58	>0.001	-0.06	11	Body	shore	NA	7.30E-05	0.42
cg06003184	C6orf174	0.003	-0.07	6	3'UTR	opensea	NA	4.30E-05	0.45
cg07761912	TOR1AIP2	0.003	-0.06	1	5'UTR	island	NA	3.00E-05	0.46
cg08174789	MYOM2	0.000	0.07	8	Body	island	NA	4.70E-05	0.44
cg09656541	C10orf118	0.024	0.05	10	5'UTR	shore	NA	9.60E-05	0.42
cg11734401	STXBP6	0.001	0.07	14	5'UTR	shore	NA	6.20E-06	0.52
cg11964314		>0.001	-0.05	12	IGR	opensea	TRUE	3.20E-05	0.45
cg13389575	RB1	0.002	0.09	13	Body	shore	NA	1.10E-04	0.40
cg16652817		0.001	0.06	7	IGR	island	NA	3.50E-05	0.45
cg19610750		0.033	0.06	2	IGR	shelf	NA	1.90E-07	0.63
cg21164545	LOC100240726	0.001	0.06	20	TSS1500	opensea	NA	1.10E-04	0.40
cg22798121	GRIA1	>0.001	-0.05	5	Body	opensea	TRUE	2.00E-05	0.48
cg23223755	LHFPL4	>0.001	0.05	3	TSS1500	shore	NA	5.60E-05	0.43
cg25487135		>0.001	0.07	12	IGR	opensea	TRUE	9.90E-05	0.41
cg26975848		0.001	-0.05	2	IGR	opensea	TRUE	7.10E-05	0.42

Nosso próximo passo então foi explorar as 21 DMRs identificadas, na qual, 15 (71%) estavam hipometiladas no grupo intervenção. Quanto ao mapeamento das regiões anotadas em genes, 14 DMRs foram localizadas (*HLA-DPB1*, *ABAT*, *ANKRD30B*, *PLEKHH3*, *ABCC13*, *OR2L13*, *LOC650226*, *LOC441666*, *PAX8*, *BOLL*, *RUFY1*, *LOC149837*, *HCG4P6*, *PF4* e *TACSTD2*), principalmente próximas ao local de iniciação / transcrição do respectivo gene (**Tabela 13**). As DMRs foram testadas quanto à sua associação potencial com o escore cognitivo dessas crianças aos 12 meses de idade por regressão linear múltipla controlada por sexo usando o escore do domínio cognitivo como resultado, e *PF4* foi a único DMR com um valor p significativo (0,006) com a β (SE) = -5,48 (1,8) e R^2 adj = 0,30 associado ao desfecho cognitivo em 12 meses.

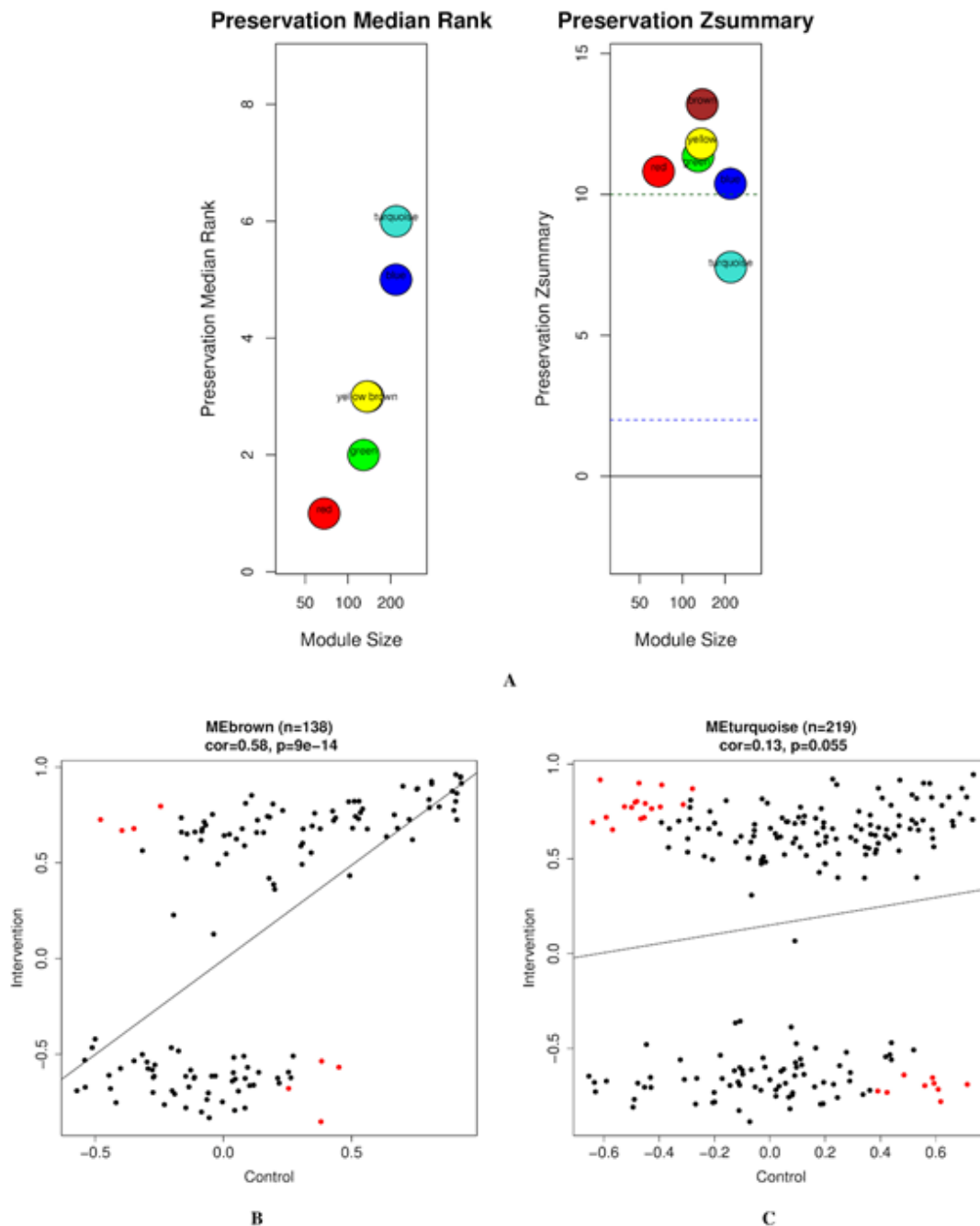
Tabela 13 - Identificação das DMR, média dos valores de M e associação com o domínio cognitivo

DMR	Coordenadas	CpGs	Valores M (média)	Total	Gene	β (SE)	p- valor	R ² aj.
DMR_1	chr6:33047944-33048919	19	1.7 (0.4)	975	HLA-DPB1	-1.89 (2.9)	0.522	0.08
DMR_2	chr16:8806359-8807043	13	-0.9 (0.3)	684	ABAT	1.6 (3.9)	0.684	0.07
DMR_3	chr7:158749941-158751591	9	0.9 (0.5)	1650		-3.34 (2.5)	0.198	0.13
DMR_4	chr18:14747661-14748439	11	-1.5 (0.3)	778	ANKRD30B	3.35 (3.5)	0.342	0.10
DMR_5	chr17:40822424-40824361	9	-0.5 (0.3)	1937	PLEKHH3	-6.03 (3,1)	0.066	0.19
DMR_6	chr1:248099757-248101009	13	-0.4 (0.9)	1252	OR2L13	-0.67(1,12)	0.559	0.08
DMR_7	chr21:15645649-15646635	7	1.7 (0.4)	986	ABCC13	-2.95 (2,6)	0.265	0.11
DMR_8	chr6:76203225-76203675	7	-0.0 (0.5)	450		3.54 (2,0)	0.096	0.17
DMR_9	chr19:10736006-10736448	8	2.4 (0.3)	442		-1.00 (3,2)	0.759	0.07
DMR_10	chr7:56515510-56516309	10	-0.6 (0.3)	799	LOC650226	-4.15 (2.4)	0.098	0.16
DMR_11	chr10:42862876-42863594	9	-1.7 (0.5)	718	LOC441666	-2.1(1.8)	0.247	0.12
DMR_12	chr2:113992762-113993313	8	1.5 (0.6)	551	PAX8	0.03 (1.5)	0.983	0.07
DMR_13	chr2:198650752-198651576	8	-2.4 (0.5)	824	BOLL	-3.65(1.9)	0.066	0.19
DMR_14	chr11:67383377-67384040	8	-0.0 (0.3)	663		-3.68 (2.3)	0.120	0.15
DMR_15	chr2:105852798-105853796	8	1.8 (0.4)	998		-1.79 (2.1)	0.395	0.09
DMR_16	chr5:178986131-178986906	9	-2.0 (0.7)	775	RUFY1	<u>1.13 (1,1)</u>	0.309	0.10
DMR_17	chr20:5485144-5485511	8	-0.3 (0.4)	367	LOC149837	1.60 (1.9)	0.414	0.09
DMR_18	chr1:247802703-247803166	7	-0.4 (0.5)	463		-2.66 (1.7)	0.122	0.15
DMR_19	chr6:29894050-29894737	7	-1.7 (0.5)	687	HCG4P6	3.71 (1,6)	0.493	0.08
DMR_20	chr4:74847646-74847829	7	-0.3 (0.4)	183	PF4	-5.48 (1,8)	0.006	0.30
DMR_21	chr1:59043070-59043370	7	-0.6 (0.6)	300	TACSTD2	0.19 (1.3)	0.887	0.07

Fonte: Elaborado pelos autores.

Ademais, buscamos redes de co-metilação dos 3.090 DMPs, que apontaram para seis módulos: MEblue (218 CpGs), MEbrown (138 CpGs), MEgreen (129 CpGs), MERed (68 CpGs), METurquoise (219 CpGs) e MEyellow (136 CpGs). METurquesa ($Z_{summary} = 7,42$, $ME_{cor} = 0,13$, $p\text{-value} = 0,055$) apresentou baixo escore de conservação (**Figura 19**). O módulo METurquoise foi composto por 126 sítios CpG mapeados para 96 genes (de 219). A análise de enriquecimento sugeriu o envolvimento da organização do filamento de actina, organização do citoesqueleto de actina, organização da fibra supramolecular e processos biológicos da morfogênese do tubo epitelial (Gene Ontology). Este resultado não foi mantido após o ajuste de correção pelo FDR. Quatro sites CpGs foram classificados como nós de hub na rede METurquoise: um localizado no corpo do CNTNAP1 e três sites CpG localizados no TSS200 do FILIP1.

Figura 19 - Análise de Preservação do Módulo. **A)** *Preservation Median Rank* (esquerda) e *Preservation Zsummary* (direita) dos módulos com base nos 3.090 DMPs fornecidos por WGCNA. O eixo X mostra o tamanho (número de CpGs) e o eixo Y mostra, respectivamente, a classificação mediana e as pontuações do resumo Z. Módulos com *Zsummary* ≤ 10 foram considerados menos preservados. Módulos com pontuações mais altas de *Median Rank* são aqueles com baixa preservação de acordo com esta métrica. Em relação ao *Median Rank*, o MEBrown teve exatamente a mesma pontuação do MEyellow (3) com um tamanho de módulo muito semelhante (138 CpGs e 136 CpGs), respectivamente. O gráfico de dispersão em **B)** e **C)** mostra a correlação de Pearson entre cada valor M de CpGs com seu respectivo módulo eigengene, para os grupos controle e intervenção. As amostras com a cor turquesa são as que apresentam as maiores diferenças de correlação de Pearson entre os dois grupos.

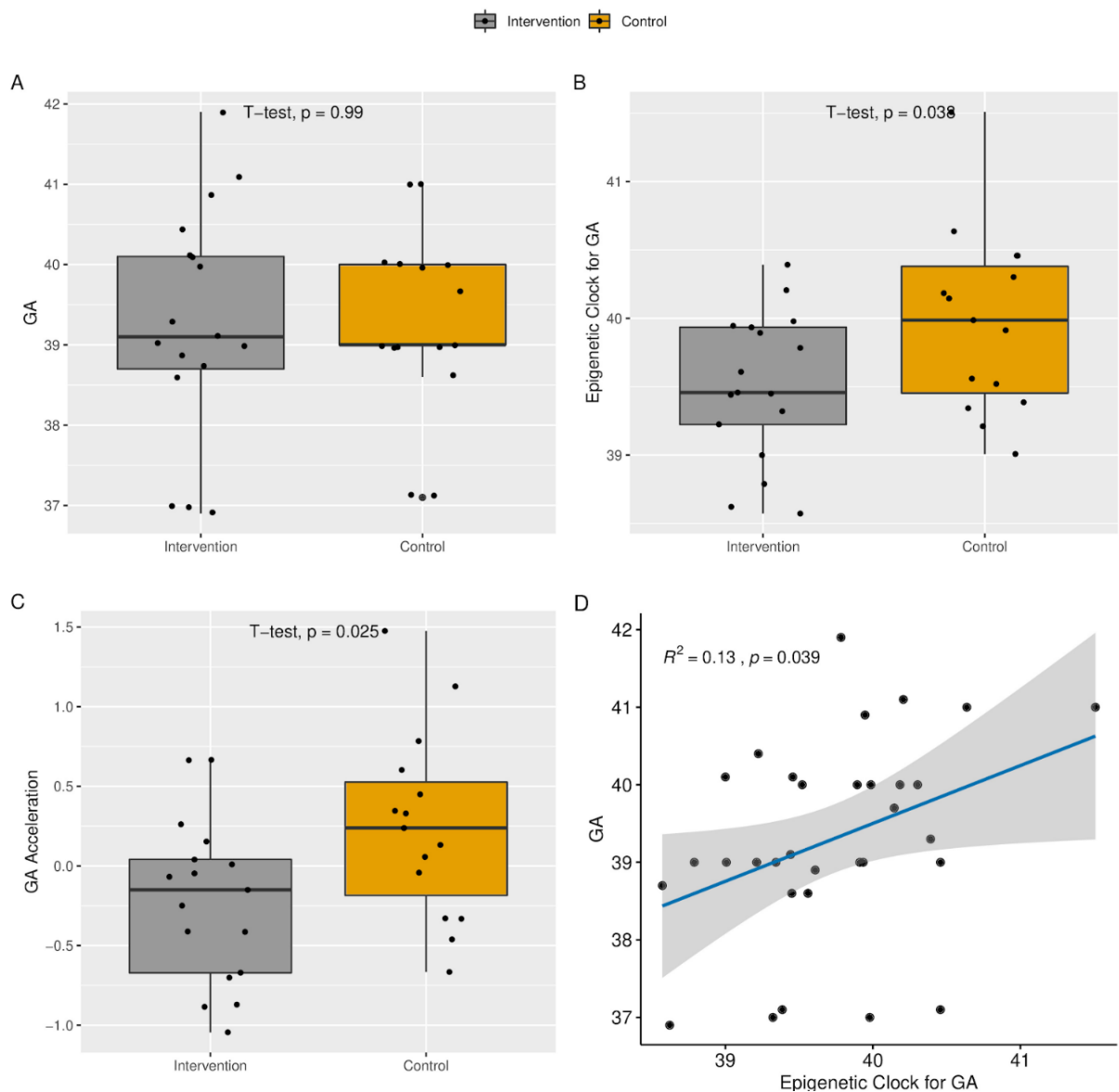


Fonte: Elaborado pelos autores.

4.2.3 Relógio Epigenético nos grupos Intervenção e Controle

Os grupos intervenção e controle não apresentaram diferenças na idade gestacional cronológica ao nascimento (Tabela 9, 39,4 X 39,3, intervenção e controle, respectivamente), porém o grupo intervenção apresentou menor média de idade gestacional epigenética em relação ao grupo controle (39,5 x 40,1 semanas, respectivamente) (**Figura 20**). A regressão logística controlada por sexo e com a idade gestacional como variável desfecho indicou coeficiente de -0,35 para o grupo intervenção ($p = 0,016$).

Figura 20 - A) Boxplot da Idade Gestacional (IG) entre intervenção e controle; B) *Boxplot* do relógio epigenético para a AG ($p=0,038$), C) Aceleração de Idade ($p=0,025$), D) Gráfico de dispersão do relógio epigenético e idade cronológica

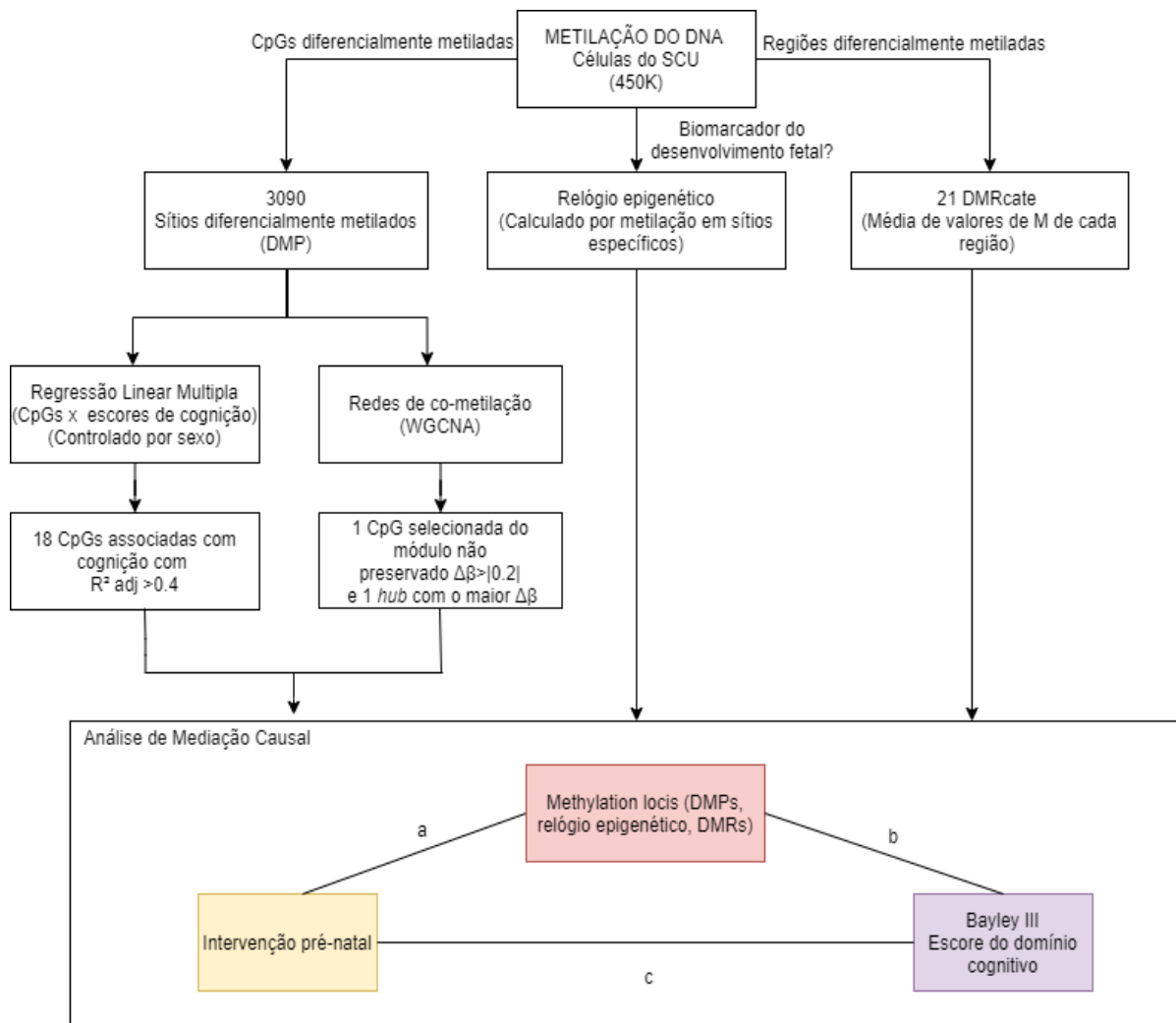


Fonte: Elaborado pelos autores.

4.2.4 Metilação do DNA como mediadora na cognição em 12 meses

Análises de mediação causal foram aplicadas para avaliar se os níveis de DNAm podem ser um mediador entre a intervenção e a cognição dos bebês aos 12 meses de idade. A **Figura 21** mostra os critérios de seleção de variáveis para testar a mediação que selecionou (1) 18 sítios CpG associados ao domínio cognitivo ($\text{adj. } R^2 > 0,4$), (2) 4 sítios CpG que eram hubs no módulo de co-metilação menos preservado (3) 21 DMRs, considerando os valores M médios de todos os CpGs na região, (4) relógio epigenético como valores brutos.

Figura 21 - Fluxograma das etapas que precederam a definição das variáveis de metilação do DNA para avaliação de efeito mediador entre a relação da intervenção de visitas domiciliares realizadas por enfermeiras a gestantes adolescentes e o domínio cognitivo aos 12 meses de vida.



Fonte: Elaborado pela autora.

Encontramos 11 DMPs apresentando ACME significativa e mediação de proporção, com seis localizados em C11orf58, TOR1AIP2, MYOM2, C10orf118, STXBP6 e LOC100240726. Além disso, três DMRs localizados em BOLL, LOC650226 e PF4 apresentaram ACME significativo, mas apenas PF4 apresentou mediação de proporção significativa. Um CpG do FILIP1 tinha um ACME significativo, mas não uma proporção significativa de mediação. Por fim, a idade gestacional dada pelo relógio epigenético não apresentou efeito de mediação.

Tabela 14 - Análise de mediação com todas as DMPs selecionadas, DMRs e RE. (continua)

Gene	Relação à ilha	Relação ao gene	Efeito Direto (ADE)			Efeito de mediação causal (ACME)			Efeito Total			
			Coef.	95%IC	<i>p</i> -valor	Coef.	95%IC	<i>p</i> -valor	Estimate			
DMP												
cg01600222		shore	IGR	-0.34	-3.28	2.76	0.810	3.5	1.41	6.05	<2e-16	3.16
cg03648020		shelf	IGR	-0.64	-3.39	1.88	0.620	3.81	1.74	6.6	<2e-16	3.17
cg03930153	TBL1XR1	opensea	5'UTR	-0.02	-2.97	2.87	0.980	3.07	3.07	6.08	<2e-16	3.05
cg05405872	C11orf58	shore	Body	-2.00	-5.42	1.48	0.260	5.15	2.55	8.23	<2e-16	3.15
cg06003184	C6orf174	opensea	3'UTR	0.09	-2.92	3.07	0.990	3.03	1.1	5.45	0.002	3.11
cg07761912	TOR1AIP2	island	5'UTR	-0.19	-2.88	2.20	0.890	3.33	1.33	5.89	<2e-16	3.13
cg08174789	MYOM2	island	Body	-1.30	-3.80	1.23	0.310	4.37	2.16	7.01	<2e-16	3.06
cg09656541	C10orf118	shore	5'UTR	0.70	-1.53	3.24	0.580	2.29	0.63	4.34	0.010	2.99
cg11734401	STXBP6	shore	5'UTR	-0.72	-3.07	1.63	0.560	3.79	1.78	6.11	<2e-16	3.07
cg11964314		opensea	IGR	-2.36	-5.67	0.79	0.150	5.51	3.1	8.29	<2e-16	3.15
cg13389575	RB1	shore	Body	0.38	-2.44	3.36	0.820	2.76	1.04	4.89	<2e-16	3.15
cg16652817		island	IGR	-0.40	-3.20	2.47	0.760	3.54	1.42	6.17	<2e-16	3.13
cg19610750		shelf	IGR	-0.02	-2.26	2.27	0.970	3.12	0.81	5.75	0.010	3.10
cg21164545	LOC100240726	opensea	TSS1500	-0.96	-4.06	2.44	0.570	4.13	1.78	6.95	<2e-16	3.16
cg22798121	GRIA1	opensea	Body	-2.01	-5.36	1.43	0.240	5.02	2.2	8.37	<2e-16	3.00
cg23223755	LHFPL4	shore	TSS1500	-1.13	-4.21	1.86	0.460	4.25	2.04	6.93	<2e-16	3.12
cg25487135		opensea	IGR	-3.41	-7.78	0.86	0.110	6.45	3.49	9.98	<2e-16	3.04
cg26975848		opensea	IGR	-0.03	-2.81	2.74	0.100	3.14	1.26	5.57	<2e-16	3.11

Tabela 14 - Análise de mediação com todas as DMPs selecionadas, DMRs e RE. (conclusão)

Gene	Relação à ilha	Relação ao gene	Efeito Direto (ADE)			Efeito de mediação causal (ACME)			Efeito Total Estimate			
			Coef.	95%IC	p-valor	Coef.	95%IC	p-valor				
WGCNA												
cg08860498	CNTNAP1	island	Body	1.00	-3.60	5.33	0.688	2.19	-1.60	6.10	0.294	3.15
cg12569246	FILIP1	opensea	TSS200	5.85	1.51	10.13	0.004	-2.74	-6.00	-0.11	0.038	3.11
cg26466508	FILIP1	opensea	TSS200	4.41	-1.29	9.76	0.11	-1.35	-5.67	3.12	0.508	3.05
cg18729664	FILIP1	opensea	TSS200	5.45	1.12	9.69	0.02	-2.45	-5.47	-0.04	0.050	3.00
DMR												
chr2:198650752-198651576	BOLL			0.61	-2.88	4.21	0.748	2.56	0.25	5.61	0.042	3.16
chr7:56515510-56516309	LOC650226			4.69	1.09	7.77	0.008	-1.55	-3.78	-0.11	0.036	3.15
chr4:74847646-74847829	PF4			1.41	-1.04	4.05	0.284	1.75	0.15	4.01	0.022	3.16
RE												
				2.81	-0.63	6.34	0.106	0.29	-1.87	2.85	0.776	3.108

Coef.= coeficiente, IC= Intervalo de Confiança

Fonte: Elaborado pela autora.

Tabela 15. Análise de mediação apenas com as variáveis de metilação de tiveram ACME (p-valor<0,05) e respectivos testes de sensibilidade (Rho).

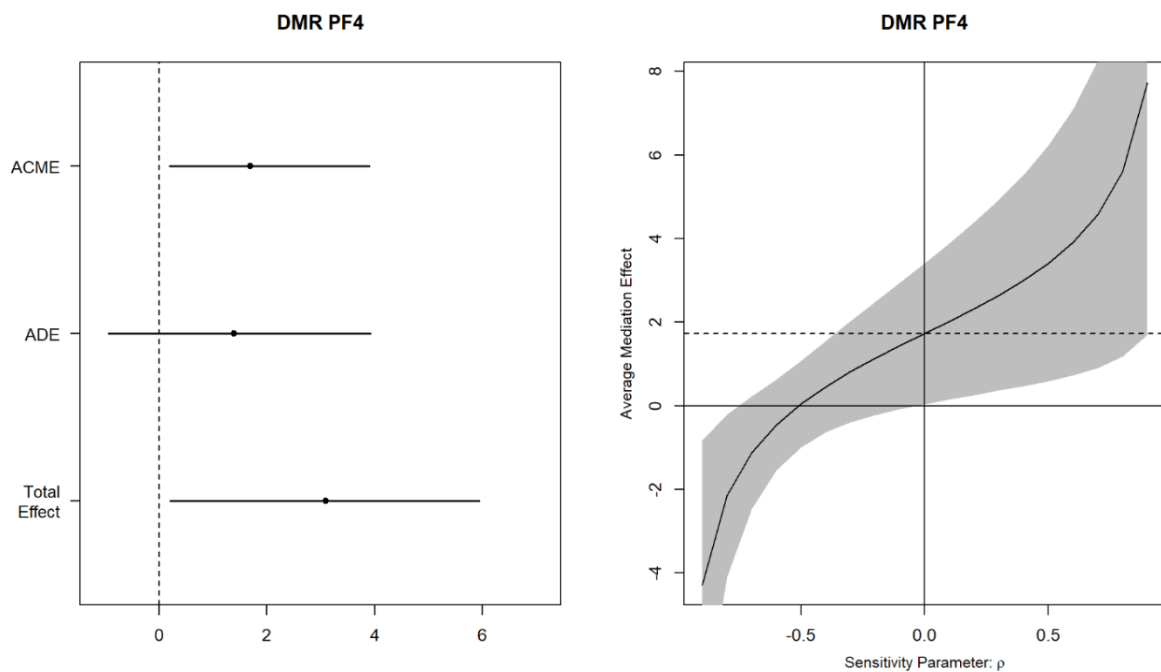
Gene	Metilação ~randomização ¹		Metilação ~ cognição ²		Efeito Direto (ADE)			Efeito de mediação causal (ACME)			Efeito Total			Rho				
	Coef.	p-valor	Coef.	p-valor	Coef.	95%IC	p-valor	Coef.	95%IC	p-valor	Coef.	95%IC	p-valor		ACME=0			
DMP																		
cg01600222	-0.57	0.002	-0.09	2E-05	-0.34	-3.28	2.76	0.810	3.5	1.41	6.05	<2E-16	3.16	0.16	6.3	0.04	-0.6	
cg03648020	0.44	0.002	0.07	5E-06	-0.64	-3.39	1.88	0.620	3.81	1.74	6.60	<2E-16	3.17	0.25	6.13	0.04	0.7	
cg05405872	C11orf58	0.40	0.000	0.05	7E-05	-2.00	-5.42	1.48	0.260	5.15	2.55	8.23	<2E-16	3.15	0.42	5.92	0.02	0.6
cg07761912	TOR1AIP2	0.59	0.003	0.09	3E-05	-0.19	-2.88	2.20	0.890	3.33	1.33	5.89	<2E-16	3.13	0.23	5.84	0.04	0.6
cg08174789	MYOM2	-0.74	>0.001	-0.09	5E-05	-1.30	-3.80	1.23	0.310	4.37	2.16	7.01	<2E-16	3.06	0.17	5.92	0.04	-0.6
cg09656541	C10orf118	-0.35	0.020	-0.06	1E-04	0.70	-1.53	3.24	0.580	2.29	0.63	4.34	0.010	2.99	0.26	6.02	0.03	-0.6
cg11734401	STXBP6	-0.53	0.002	-0.08	6E-06	-0.72	-3.07	1.63	0.560	3.79	1.78	6.11	<2E-16	3.07	0.19	6.03	0.04	-0.7
cg11964314		0.36	1E-05	0.04	3E-05	-2.36	-5.67	0.79	0.150	5.51	3.1	8.29	<2E-16	3.15	0.32	5.84	0.03	0.7
cg21164545	LOC100240726	-0.44	>0.001	-0.05	1E-04	-0.96	-4.06	2.44	0.570	4.13	1.78	6.95	<2E-16	3.16	0.12	6.25	0.04	-0.6
cg25487135		-0.65	5E-07	-0.07	1E-04	-3.41	-7.78	0.86	0.110	6.45	3.49	9.98	<2E-16	3.04	0.06	5.83	0.04	-0.6
cg26975848		0.30	0.003	0.04	7E-05	-0.03	-2.81	2.74	0.100	3.14	1.26	5.57	<2E-16	3.11	0.07	6.12	0.04	0.6
DMR																		
chr4:74847646-74847829	PF4	-0.31	0.046	-5.48	0.006	1.41	-1.04	4.05	0.284	1.75	0.15	4.01	0.022	3.16	0.32	6.35	0.03	-0.5

Coef.= coeficiente, IC= Intervalo de Confiança, ¹ Regressar linear entre o mediador e intervenção, ² Regressão linear entre o mediador e desfecho.

Fonte: Elaborado pela autora.

Para todas as análises com mediação significativa (ACME p-valor $>0,05$) nós realizamos as análises de sensibilidade. Essa etapa é recomendada por estarmos analisando dados que estão susceptíveis a outras variáveis não abordadas aqui, que são chamadas de confundidoras. A análise de sensibilidade é justamente para indicada justamente para avaliar o quanto outras variáveis não mensuradas poderiam interferir na mediação fazendo-a não significativa diante do universo de tantas outras. Para a análise de sensibilidade uma das formas aritméticas é o cálculo do produto dos coeficientes de determinação de ambos os modelos de regressão. Para a DMR PF4, por exemplo, o efeito indireto será zero quando os fatores de confusão da relação mediador-resposta juntos explicam 25% ou mais da variação residual (Figura 22). Para todas as variáveis mediadoras, aplicamos essa métrica e todas reforçaram que de fato, os sítos e DMRs com ACME significativo, a metilação do DNA foi mediadora ainda considerando as variáveis desconhecidas não contempladas aqui.

Figura 22 - Análise de sensibilidade para a mediação da DMR mapeada no gene PF4.



Fonte: Elaborado pela autora.

4.3 DISCUSSÃO DADOS DO RCT PRIMEIROS LAÇOS

Neste estudo, demonstramos que as mudanças epigenéticas atuam como mediadoras dos efeitos de uma intervenção de visita domiciliar conduzida por enfermeiras durante a gravidez sobre os escores cognitivos de bebês aos 12 meses de idade em uma amostra de alto risco participante de um ensaio clínico randomizado.

Relatamos diferenças de DNAm entre os grupos de intervenção e controle em locais CpG, DMRs e níveis de rede de co-metilação. Comparando as análises de metilação, encontramos alguns CpGs e genes sobrepostos, embora nenhum atuasse como mediador. *ABCC1* tinha um DMR e um DMP com $\Delta\beta > 20\%$ ($\text{adjP} = 5,5\text{E-}6$). *OR2L13* também teve um DMR e um DMP com $\Delta\beta > 20\%$ ($\text{adjP} \leq 0,001$). *OR2L13* foi encontrado por múltiplas análises independentes de estudos EWAS [38]. *OR2L13* pertence à via do sistema sensorial olfatório, que está relacionada ao prejuízo social para crianças com desenvolvimento típico [39]. Em nosso estudo, o local CpG (cg20507276) localizado em *OR2L13* foi menos metilado no grupo de intervenção (média $\Delta\beta = -0,17$), mas foi encontrado hipermetilado (média $\Delta\beta = 0,18$) em indivíduos com transtorno do espectro do autismo [40]. O cg08860498 mapeado em *CNTNAPI* é um hub do módulo de rede menos preservado, MEtuquoise, e mostrou cinco DMPs hipometilados significativos (cg07871971, cg16308533, cg12559685, cg27616541, cg08860498). O *CNTNAPI* está envolvido na formação de axônios mielinizados e na manutenção da estabilidade das conexões adjacentes [41,42]. Dois DMPs, cg23223755 e cg21164545, localizados em LHPL4 e LOC100240726, respectivamente, estão associados à cognição e pertencem ao módulo turquesa (MEtuquoise). A transmissão inibitória rápida no cérebro é mediada principalmente por receptores ionotrópicos GABAA (GABAARs). LH4 forma um complexo tripartido estável com GABAARs e neuroligin-2 no cérebro. Além disso, o LH4 é necessário para a localização sináptica de GABAARs e transmissão sináptica inibitória no hipocampo [43]. DMPs localizados em *ANKRD30B* (cg02105152, cg23703062, cg21293934, cg16584393, cg02992546) também pertencem ao módulo MEtuquoise e foram considerados um DMR. Encontramos uma DMR no *PAX8*, um gene que teve papel descrito no desenvolvimento do mesencéfalo ao se ligar a sequências regulatórias EN2 [44] e foi associado ao estado neurocomportamental no nascimento [45].

Analisamos o relógio epigenético, uma medida indireta da senescência celular, considerando que o envelhecimento pode começar antes mesmo do nascimento [46]. O grupo intervenção apresentou menor idade obtida pelo relógio epigenético. Bebês nascidos de mães que fumaram durante a gestação apresentaram risco três vezes maior de envelhecimento do que

bebês de mães não fumantes [47], mas isso não poderia explicar o resultado, uma vez que não foi encontrada diferença em relação ao tabagismo entre os grupos. Além disso, mais estudos longitudinais são necessários para entender como interpretar esses resultados e se o relógio epigenético poderia prever resultados de saúde mais tarde na vida.

Todos os DMPs testados apresentaram efeito indireto no escore cognitivo aos 12 meses medidos pela escala de Bayley. Considerando a localização dos sítios CpG com ACME significativo, mas com mediação de proporção não significativa, *TBL1XR1* e *GRIA1* foram previamente associados ao neurodesenvolvimento. *TBL1XR1* (cg03930153) recruta β -catenina para genes alvo de Wnt que ativam a transcrição. Mutações *TBL1XR1* foram associadas a condições neuropsiquiátricas [48,49], *GRIA1* (cg22798121) tem um papel na transmissão sináptica glutamatérgica, estando relacionada à plasticidade sináptica e distúrbios do neurodesenvolvimento [50,51].

Três das 11 DMPs com um efeito de mediação foram mapeados para genes envolvidos na manutenção e funcionalidade dos neurônios: *C10orf118* (cg05405872), *STXBP6* (cg11734401) e *TOR1AIP2* (cg07761912). *C10orf118* e *STXBP6* estão envolvidos com o tráfego de vesículas e sinalização [52], uma vez que mantêm a organização funcional da célula e, portanto, desempenham um papel em processos neuronais, como a localização de proteínas dendríticas [53]. *TOR1AIP1* codifica cofatores para a ATPase Torsina 1A, que tem sido implicada na manutenção da fisiologia celular normal. Estudos recentes também demonstraram que as Torsinas são importantes no desenvolvimento de neurônios, sugerindo um papel durante a diferenciação neuronal [54]. Os três DMPS estão localizados no 5'UTR aumentando a probabilidade de sua metilação estar associada com a regulação negativa da expressão gênica. Assim, sugerimos que a melhora do escore cognitivo pela intervenção é mediada por uma diminuição na expressão de *TOR1AIP2* e um aumento na expressão de *C10orf118* e *STXBP6*, com a primeira mediando 70% do efeito total. Além disso, cg19610750, que medeia o efeito de intervenção na cognição, está localizado em uma região intergênica onde existem locais de ligação para os fatores de transcrição CTCF e EGR1, ambos envolvidos com o desenvolvimento fetal [55,56]

Em relação aos DMRs, o PF4 DMR hipometilado mediou 55% do efeito total na associação entre intervenção e escores cognitivos. PF4 codifica um membro da família de quimiocinas CXC que inicia uma cascata de transdução de sinal de funções agudas e retardadas em monócitos humanos, incluindo fagocitose, explosão respiratória, sobrevivência e secreção de citocinas [57]. O PF4 foi o centro de uma análise de rede de coexpressão comparando indivíduos expostos ao estresse psicossocial inicial a controles [58]. O módulo foi enriquecido

para atividade de quimiocinas e para processos biológicos funcionais de vesículas secretoras. Além disso, a superexpressão de PF4 contribui para um efeito anti-envelhecimento em camundongos idosos que tomam peptídeos de colágeno [59]. Descrevemos o PF4 como hipometilado no grupo intervenção, o que pode estar associado à superexpressão gênica e sugere que pode estar contribuindo para a menor aceleração da idade observada e melhores escores cognitivos nesse grupo.

Em comparação com toda a amostra, a subamostra com dados do cordão umbilical foi mais frequentemente meninos ($p = 0,01$) e mães apresentaram maiores escores no Inventário de Ansiedade de Beck ($p = 0,04$). Mas não houve diferenças entre os grupos na subamostra. Mesmo assim, tivemos um poder estatístico reduzido devido ao menor tamanho da amostra. Também é importante enfatizar que um critério de corte permissivo de 0,05 para correção múltipla foi usado para encontrar sítios diferencialmente metilados. No entanto, ele poderia gerar mais falsos positivos e nós atenuamos selecionando apenas os locais CpG para a análise de mediação associada ao resultado que explicava mais de 40% do modelo de regressão. Também reconhecemos que as mudanças epigenéticas no sangue do cordão umbilical não refletem necessariamente o cérebro fetal. No entanto, o tecido substituto pode ser informativo mesmo se o tecido investigado não estiver diretamente envolvido no fenótipo de interesse, pois eles podem refletir a exposição às influências ambientais (foi demonstrado, por exemplo, para os efeitos da subnutrição pré-natal [9], produtos químicos, adversidade psicossocial [60] ou combate [61]). Assim, embora a variação no tecido periférico possa não refletir diretamente a variação epigenética no cérebro, ela ainda pode representar biomarcadores utilizáveis, pois reflete uma exposição que leva ao fenótipo em um tecido diferente [62].

Mesmo com algumas limitações, os achados apresentados têm implicações importantes, pois documentamos os efeitos epigenéticos de uma intervenção psicossocial durante a gravidez, o que é relevante para entender os mecanismos biológicos relacionados à intervenção (Insel & Gogtay, 2014), *JAMA Psychiatry* 2014; 71 (7): 745-6.

Em conclusão, este estudo piloto fornece suporte para diferenças de DNAm em genes envolvidos com o neurodesenvolvimento, comparando bebês cujas mães receberam visitas de enfermeiras treinadas durante a gravidez com aqueles que não receberam a intervenção. Especificamente, descobrimos que a associação entre a intervenção e o escore cognitivo aos 12 meses pode ser mediada pelo DNAm.

5 CAPÍTULO 5: PLACENTA

5.1 CASUÍSTICA

A placenta é um órgão materno-fetal que tem o início do seu desenvolvimento no começo da implantação do blastocisto ~~em~~ e seu tempo de vida limitado ao fim da gestação. Durante esse período de tempo, a placenta irá se desenvolver e se diferenciar com o intuito de preservar o crescimento e desenvolvimento fetal. Nesse sentido, desde o início da formação da placenta, das diferenciações em células trofoblásticas, existe um fino equilíbrio que, se interrompido, pode impactar sobre a vida em formação. Uma série de evidências sugerem que o início da placentação pode fornecer informações valiosas para a compreensão de diferentes desfechos na saúde do indivíduo.

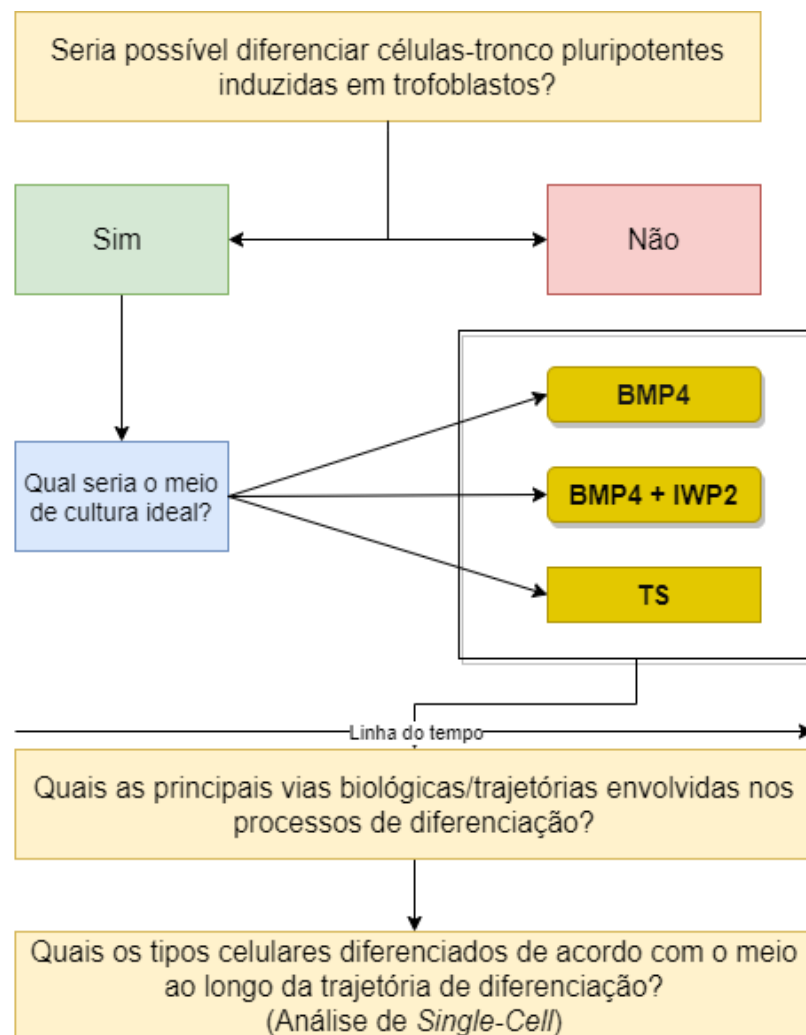
Em virtude disso, uma série de estudos têm sido direcionados para reconhecer as vias de diferenciação e condições estressoras que poderiam desequilibrar esse processo. O primeiro evento de diferenciação durante o desenvolvimento embrionário humano se dá pela especificação do trofotoderma e da massa celular interna. O trofotoderma medeia a implantação do blastocisto no útero e é o precursor de todas as células trofoblásticas da placenta (119). Quando o blastocisto alcança o epitélio do endométrio, células trofoblásticas irão se originar da camada mais externa do blastocisto, a trofotoderma: citotrofoblasto, trofoblastos extravilosos e sinciotrofoblasto. Quando o citotrofoblasto se proliferam, novas células são formadas, perdendo suas membranas e formando células multinucleadas, os sinciotrofoblastos. Os citotrofoblastos liberam enzimas proteolíticas, enquanto os sinciotrofoblastos vão se moldando, permitindo que o blastocisto invada o endométrio. Uma lacuna é formada entre os dois tipos celulares, que logo será preenchida com vasos maternos e glândulas secretoras, criando dessa forma uma rede de suporte para garantir crescimento e desenvolvimento. Na segunda semana, pequenas projeções do citotrofoblasto começam a invadir em direção aos sinciotrofoblasto formando as primeiras vilosidades coriônicas. No início da terceira semana, o mesoderma extraembrionário cresce dentro dessas vilosidades formando um tecido interconectado. No final da terceira semana se forma a vilosidade coriônica terciária, os citotrofoblastos migram ao topo formando uma “membrana” externa e, até a quarta

semana, a circulação fetal começa a ser estabelecida. O sinciciotrofoblasto de vilosidades flutuantes representa a unidade de transporte da placenta humana, as vilosidades de ancoragem da placa basal da placenta formam outro tipo de trofoblasto diferenciado, o denominado trofoblasto extraviloso invasivo (EVT). Já com 2 semanas após a fertilização, dois tipos de EVTs podem ser detectados dentro do compartimento uterino materno, CTBs intersticiais (iCTBs), colonizando o estroma decidual e CTBs endovasculares (eCTBs). Ademais, os iCTBs acumulam na parede do vaso muscular promovendo sua elastólise e degradação, onde macrófagos deciduais e células uterinas *natural killer* também contribuem para este processo. Assim que a circulação útero-placentária é estabelecida, entre a 10^a e a 12^a semana de gravidez, as vilosidades placentárias são banhadas em sangue materno e, portanto, são denominadas vilosidades flutuantes (120). As artérias umbilicais criam braços nas vilosidades coriônicas formando uma rede de capilares que permitem que as trocas gasosas comecem a ocorrer, com oxigênio e nutrientes provenientes da circulação materna, onde esse orquestrado equilíbrio relacionado a angiogênese se torna cada vez mais imperativo para garantir a nutrição fetal que, por sua vez, a cada fase se torna mais exigente. Cada passo de diferenciação e crescimento das células placentárias importam (121). Algumas complicações têm sido descritas relacionadas ao início do desenvolvimento da placenta, tais como defeito no processo angiogênico, impedindo um fluxo adequado de nutrientes e oxigênio, falhas nas interações entre células trofoblástica e vasos intrauterinos, desequilíbrio de células do sistema imunológico/inflamatório(120). E é sobre o reconhecimento de vias biológicas fundamentais ao início do desenvolvimento placentário que existem lacunas a serem preenchidas.

Atualmente, há uma compreensão limitada das trajetórias iniciais de desenvolvimento das diferenciações trofoblásticas humanas (Rossant, 2018). Existe uma dificuldade na utilização de modelos animais, uma vez que já foi demonstrado que a expressão gênica relacionada à placenta e a reprodução são os que mais se divergem em modelos humanos e animais. Os dados mais robustos que se tem na literatura do início do desenvolvimento placentário (antes da 6^a semana, especialmente) são provenientes de modelos animais, contudo, como mencionado, não refletem exatamente o que acontece no desenvolvimento placentário humano. É nesse sentido que propusemos desenvolver um modelo, a partir de células pluripotentes induzidas (iPS, do inglês, *induced Pluripotent Stem Cell*) para que estudos relacionados ao desenvolvimento placentário elucidem melhor o início da vida e trajetórias de diferenciação. As perguntas relacionadas a esse trabalho podem ser observadas na **Figura 22**, abaixo). Esse projeto foi desenvolvido no *Lieber Institute of Brain Development*, da *Johns Hopkins University*, em oportunidade de bolsa contemplada pelo programa PDSE da

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) por 12 meses. Os resultados deste projeto resultaram em submissão de artigo científico da Cell Stem Cell (CELL-STEM-CELL-D-20-00696, ANEXO 4), cuja aluna de doutorado em questão divide primeira autoria com dois colegas de equipe.

Figura 22 - Principais pontos de investigação relacionados à utilização de células-tronco pluripotentes induzidas para diferenciação em células trofoblásticas



Fonte: Elaborado pela autora.

5.2 MÉTODOS

5.2.1 Cultura das células

As análises de bancada foram realizadas pelo técnico-superior do laboratório Erwin, Ethan Tietze, cuja submissão de artigo científico consiste em primeira autoria compartilhada. Placas de 6 poços com cultura PSC 739 foram enriquecidas com fator de crescimento reduzido Cultrex (1 mg / 12 ml DMEM / F12) a 37 °C por pelo menos uma hora. Células Tronco e células pluripotentes induzidas de fibroblastos humanos foram mantidos em placas revestidas em meio Stemflex. As células foram confluídas em pequenos grupos usando solução de Versene e uma proporção de divisão de 1: 10-1: 12. Para criopreservação, as células foram suspensas em meio PSC e temporariamente armazenadas em um freezer a -80 °C antes de serem transferidas para nitrogênio líquido para armazenamento de longo prazo.

5.2.2 Diferenciações e manutenção de células

As placas de 6 poços foram revestidas com fator de crescimento reduzido Cultrex (1 mg / 12 ml DMEM / F12) a 37 °C por pelo menos uma hora. As iPSC foram colhidas em pequenos grupos usando solução Versene sem centrifugação, semeadas em uma proporção de passagem de 1:12 e cultivadas em 2 ml de meio Stemflex. Após 24 horas, as células foram lavadas com DPBS e 2 ml de meio apropriado foram adicionados a cada um se diferenciando bem.

Para diferenciação TS, o meio TS [DMEM / F12 com Glutamax suplementado com 2-mercaptoetanol 0,1 mM, FBS 0,2%, BSA 0,3%, suplemento ITS-X a 1%, ácido L-ascórbico 1,5 µg / ml, EGF 50 ng / ml 2 µM de CHIR99021, 0,5 µM de A83-01, 1 µM de SB431542, 0,8 mM de VPA e 5 µM de Y27632] foi adicionado a cada poço.

Para diferenciações baseadas em BMP4, 10ng / ml de BMP4 e / ou 2µM IWP2 foram adicionados ao meio de diferenciação basal (DMEM / F12 com Glutamax suplementado com ITS e ácido L-ascórbico 2-fosfato de magnésio). A cada 24 horas, 2 ml de meio fresco foram adicionados a cada poço. As células foram coletadas nos dias 2, 4 e 6 para sequenciamento de RNA.

Para a condição TS, após 6 dias de diferenciação, as células foram passadas usando TrypLE expresso em uma proporção de divisão de 1: 3 e plaqueadas em novas placas revestidas com Cultrex ou col IV (5 µg / ml). Após a primeira passagem, as células receberam suplementos a cada 48 horas e divididas na proporção de 1: 3-1: 6 a cada 3-4 dias. Células TS maduras entre

as passagens 10-20 foram usadas para todas as experiências, a menos que indicado de outra forma. As células foram cultivadas em incubadoras umidificadas com 5% de CO₂ e 20% de O₂. Para culturas hipóxicas, os níveis de O₂ foram reduzidos para 5%. Para a criopreservação, as células TS foram suspensas em Cell Banker 1 e armazenadas temporariamente em um freezer a -80 °C antes de serem transferidas para nitrogênio líquido para armazenamento de longo prazo.

5.2.3 Análises laboratoriais

As colorações por imunofluorescência, Sequenciamento de *Single-cell* por Dropseq, controle de qualidade e sequenciamento estão descritos no artigo submetido (ANEXO 3).

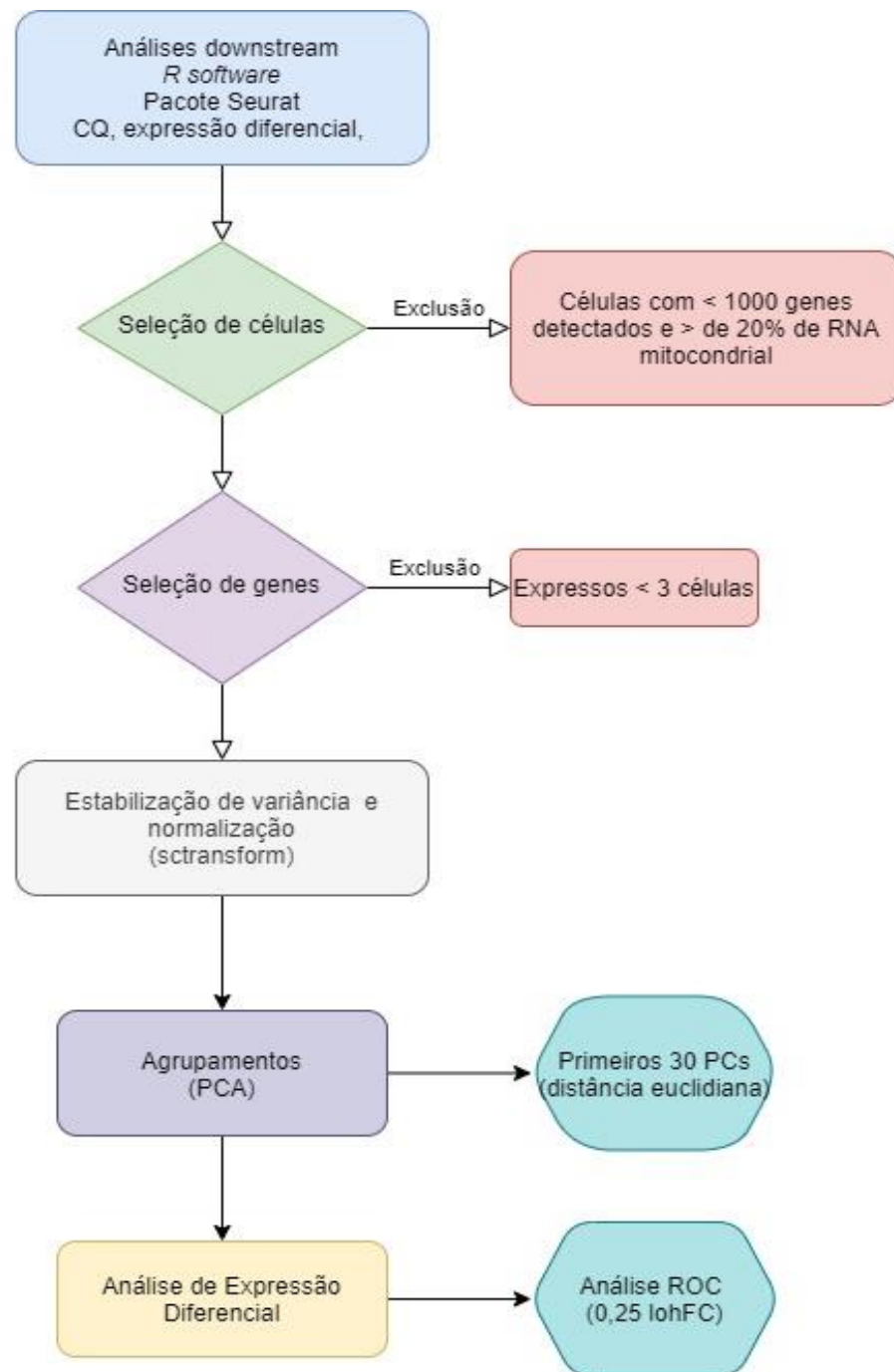
5.2.4 Pré-processamento de dados Drop-seq

Os dados de sequenciamento brutos foram pré-processados seguindo o *pipeline* “Drop-seq Alignment Cookbook” v2.0.0 encontrado em <https://github.com/broadinstitute/Drop-seq/releases/> e descrito em(122). Resumidamente, as leituras emparelhadas foram filtradas para remover pares de leitura com qualquer base com qualidade inferior a 10 em códigos de barras celulares e moleculares. Adaptadores SMART na extremidade 5' e caudas poliA na extremidade 3' com 6 ou mais pb foram removidos do segundo par e, em seguida, alinhados ao genoma humano de referência (GRCh38) usando HISAT2 v2.1.0 (Kim et al., 2015) com as configurações padrão. As leituras mapeadas exclusivamente foram mantidas e agrupadas por código de barras da célula. Códigos de barras UMI com distância de hamming = 1 foram mesclados dentro de cada célula, e o número de transcritos de cada gene foi definido com base no número total de cada sequência UMI distinta, para aquele gene, em uma determinada célula.

5.2.5 Análise de clusterização e expressão diferencial de genes para genes marcadores

Todas as análises *downstream* foram desenvolvidas juntamente com bioinformata André Barbosa (compartilha primeira autoria). Para o controle de qualidade, agrupamento e expressão diferencial foram realizadas usando o pacote Seurat (123) v3.0 R e estão apresentadas no fluxograma abaixo (**Figura 23**)

Figura 23 - Fluxograma da seleção de células únicas até análise de expressão diferencial



Fonte: Elaborado pela autora.

Os dados foram inicialmente filtrados para remover células de baixa qualidade e genes detectados de muito baixo nível. Descartamos células com menos de 1000 genes detectados e com mais de 20% do conteúdo de genes mitocondriais. Genes expressos em menos de 3 células também foram removidos. A estabilização e normalização da variância foram realizadas usando

a abordagem *sctransform*(124), ajustando o tamanho da biblioteca e a porcentagem de genes mitocondriais em cada célula. Para agrupar as células, as distâncias das células foram definidas com 30 primeiros PCs usando uma abordagem baseada em gráfico implementada em Seurat, que constrói brevemente um gráfico do vizinho mais próximo K849 (KNN) com base na distância euclidiana no espaço PCA, com posterior refino de pesos de borda usando similaridade de *Jaccard*. Os *clusters* (agrupamentos) foram definidos usando o algoritmo padrão Seurat — um gráfico vizinho mais próximo compartilhado (SNN) baseado em k-vizinhos mais próximos com otimização de modularidade usando o algoritmo Louvain original. Marcadores gênicos de *clusters* de células foram definidos usando análise ROC para apenas genes com pelo menos 0,25 logFC de diferença média entre os *clusters*. As visualizações gráficas de *clusters* e expressão gênica foram feitas usando Seurat e ggplot2 em R.

5.2.6 Reconstrução de Trajetória

As medições de uma única célula em larga escala abriram oportunidades sem precedentes para estudar processos dinâmicos, como diferenciação ou resposta a estímulos. Os métodos de inferência de trajetória são ferramentas importantes para descrever esses processos com base em medições instantâneas. A reconstrução de trajetória foi feita por pacote STREAM (Pacote na linguagem Python) (125). O STREAM é uma ferramenta de inferência de trajetória capaz de reconstruir trajetórias e inferir uma espécie de pseudotempo, a partir de diferentes tipos de dados e quando vários pontos de ramificação estão presentes. Ademais, não requer conhecimento prévio, como pontos no tempo, célula inicial ou o número de eventos de ramificação para reconstruir as trajetórias. O dado é obtido a partir dos genes mais variáveis. Ainda, genes que são considerados marcadores de cada braço do tempo, são detectados pelo algoritmo nomeados de Folha (variação de expressão, baseado em *score-z* e, confirmado por teste de *Kruskal-Wallis*). A análise de enriquecimento *Gene Ontology* foi feita nesses genes nomeados de Folhas (*Leaf genes*), com os 100 genes mais significantes obtidos por teste hipergeométrico (pacote do R, GOstats v2.48.0). A projeção em imagem em círculo com as vias biológicas foi realizada em pacote do R, GOplot v1.0.2. Além disso, o algoritmo reconhece genes chamados de transição, e para essa definição esse gene tem que estar presente em 20% de um grupo e 80% do outro. A média de expressão gênica é calculada para cada grupo formado, com *log2 fold change* maior que 0,25 entre os grupos.

5.2.7 Análise de enriquecimento dos agrupamentos para genes específicos de células da placenta

Para caracterizar os perfis de expressão dos 19 *clusters* detectados neste trabalho, analisamos a sobreposição com genes específicos de células identificadas por estudos anteriores de *Single-cell*. Usamos dois conjuntos de genes candidatos. Um conjunto de 38 listas de genes específicos de células (30 genes cada) foi extraído da tabela suplementar 2 do paper do Vento-Tormo et al. 2018(65). Essas listas de genes abrangem tecidos maternos (decídua e sangue) e fetais (placenta). Outro conjunto de 14 listas de genes específicos de células foi extraído da tabela suplementar 1 de (Liu et al. 2018), que também abrange tecidos análogos.

5.2.8 Redes regulatórias PANDA

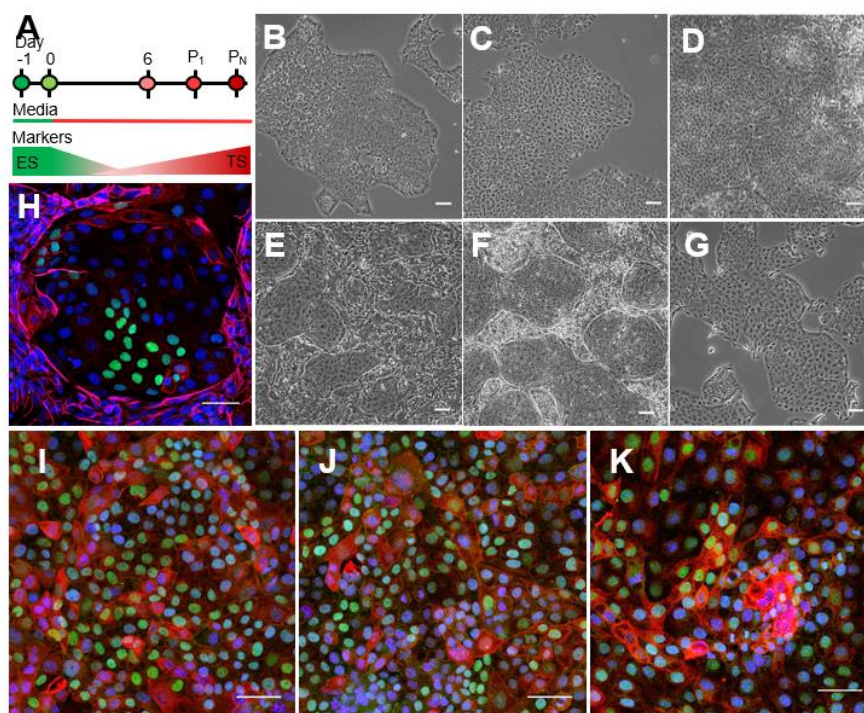
A análise de redes regulatórias fez parte da tese de doutorado do aluno de Bioinformática André Barbosa, orientado pela Prof^a Helena Brentani e será brevemente descrita aqui para compreensão da interpretação biológica, proposta desse capítulo da tese. Redes reguladoras de genes foram construídas usando o pacote pandaR v1.14.0 R que integra vários tipos de dados para inferir interações diretas (bordas) entre FT (Fator de Transcrição) e genes alvo. O algoritmo do PANDA inicia com uma rede regulatória anterior que pode ser construída mapeando locais de ligação de FT ao genoma e refina esta rede inicial integrando dados de expressão de genes alvo e dados de FT — e rede de PPI do STRING v11.0 (baixado de <https://string-db.org/>) (do inglês, *Protein-Protein Interaction networks*). A ideia principal por trás do algoritmo é que genes-alvo de um determinado FT são provavelmente co-expressos e os FTs que interagem uns com os outros são mais propensos a regular um grupo semelhante de genes-alvo. Iterativamente, o algoritmo refina a estrutura inicial da rede e infere uma rede regulatória de consenso final. Uma rede regulatória foi construída para cada *cluster* usando como rede inicial um mapa de ligação do motif FT (126), disponível em: <https://sites.google.com/a/channing.harvard.edu/kimberlyglass/tools/resources>. A visualização gráfica, a *clusterização* de genes e a análise de enriquecimento de GO de *clusters* de genes dentro de redes foram realizadas usando Cytoscape v3.7.2 e ClueGO v2.5.4 (127,128).

5.3 RESULTADOS DO CAPÍTULO 5 PLACENTA

5.3.1 Derivação de células-tronco trofoblásticas de células-tronco embrionárias humanas preparadas e células-tronco pluripotentes induzidas sem BMP4

Uma série de limitações têm sido discutidas em termos dos modelos de iPS para diferenciação em células trofoblásticas. O meio de cultura com BMP4 tem sido o mais estudado na diferenciação e, por isso, o usamos aqui para comparar com nossa proposta de ativação de outras vias de diferenciação. Sugerimos isso especialmente porque BMP4 não é expresso no TB pré-implantação, e não temos evidências se na ausência de BMP4 seria possível desenvolver TB. Uma das vias alternativas a ativação de BMP4 seria pela ativação da WNT (do inglês, wingless/Integrated) e do fator de crescimento epidérmico (EGF) e inibição do fator de crescimento beta (TGF β), deacetilase de histona (HDAC) e Proteína quinase associada a Rho (ROCK). Os resultados deste estudo mostram que a ativação de WNT e EGF e inibição de TGF β , HDAC e ROCK tiveram uma resposta positiva tanto para o tempo de diferenciação quanto para marcadores de trofoblastos. Vinte e quatro horas após a passagem de confluência da linha de células-tronco embrionárias humanas H1 (129) em condições de células-tronco livres de alimentação, as colônias ES demonstraram empacotamento apertado em cúpula com bordas definidas características de células iniciadas (**Figura 24A-B**). Quarenta e oito horas após a passagem, as células foram trocadas diretamente para meio TSC. Por imagem de campo claro durante os 6 dias subsequentes, observamos que as células proliferaram rapidamente e algumas células adotaram uma aparência de paralelepípedo (**Figura 24 A-D**). Após a passagem, emergiram duas populações morfolologicamente distintas.

Figura 24 - Derivação de células trofoblásticas pluripotentes (TSC) a partir de células-tronco pluripotentes induzidas.



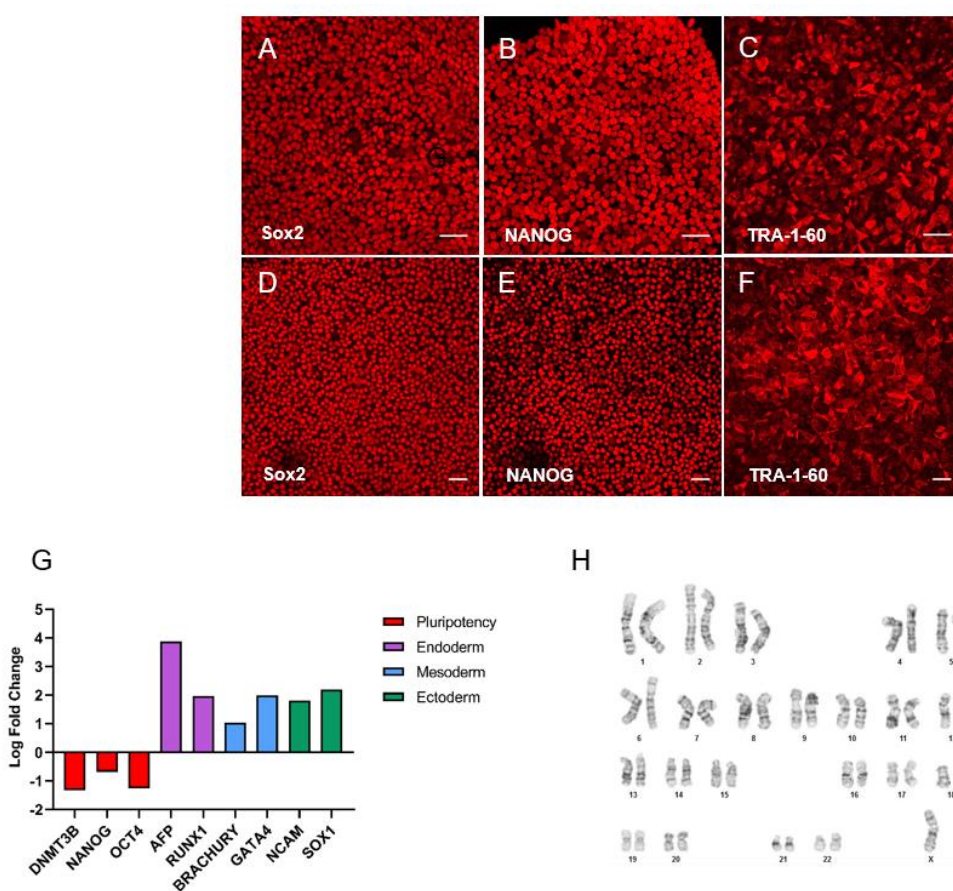
Fonte: Elaborado pelos autores.

As células epiteliais continuaram a proliferar e se tornaram morfológicamente semelhantes às células trofoblásticas derivadas de embriões humanos primários (Figura 24G). A imunofluorescência revelou que as subpopulações das células da colônia expressaram TP63 e/ou KRT7 e expressaram minimamente VIM, enquanto as células fibroblásticas circundantes expressaram fortemente VIM (Figura 24H). Ao longo de várias passagens, as células epiteliais preferencialmente se expandem e atingem ficarem mais semelhantes ao trofoblasto entre as passagens 5-7. Estas células trofoblásticas pluripotentes expressam amplamente TP63 e KRT7 (Figura 24-K), sem expressar 167 dos genes de pluripotência SOX2 e NANOG. Em resumo, descobrimos que a ativação do fator de crescimento Wnt e do EGF e a inibição do TGF β , HDAC) e da proteína quinase associada a Rho (ROCK) diretamente nas linhas iPS iniciadas resulta em especificação eficiente de células-tronco trofoblásticas.

O estado de pluripotência varia na propensão que uma célula tem em se diferenciar em outros tipos celulares(130). Dessa forma, investigamos se as TSC sem BMP4 poderia especificar em uma série de células pluripotentes humanas. Para reconhecimento de uma célula com TB, os marcadores VIM, KRT7 e TP63 foram utilizados em linhas iPS reprogramadas a partir de fibroblastos dérmicos por sendai, denominado 2014.06 (**Figura 25**), e duas linhas

celulares iPS derivadas de fibroblasto post-mortem reprogramadas por epissona, denominadas LIBD7c6 e LIBD1c8 (131). Para confirmar a capacidade de proliferar como células semelhantes a TS, as linhas de TS derivadas de iPS foram submetidas por até 32 passagens e, ainda mantiveram a expressão de KRT7 e TP63.

Figura 25 - Caracterização das linhagens celulares iPS. (A-C) Imagens imunofluorescentes da linha HiPS LIBD1C8 coradas para Sox2 (A), NANOG (B) e TRA-1-60 (C). (D-F) Imagens imunofluorescentes da linha HiPS 2014.06 coradas para Sox2 (D), NANOG (E) e TRA-1-60 (F). (G) Níveis de expressão relativa de pluripotência e marcadores específicos da camada germinativa em células espontaneamente diferenciadas da linha 2014.06 normalizados para seus respectivos níveis de expressão em HiPS indiferenciados da mesma linha celular determinada por qPCR. (H) Análise do cariótipo de 2014.06. Todas as barras de escala 50 μ m.



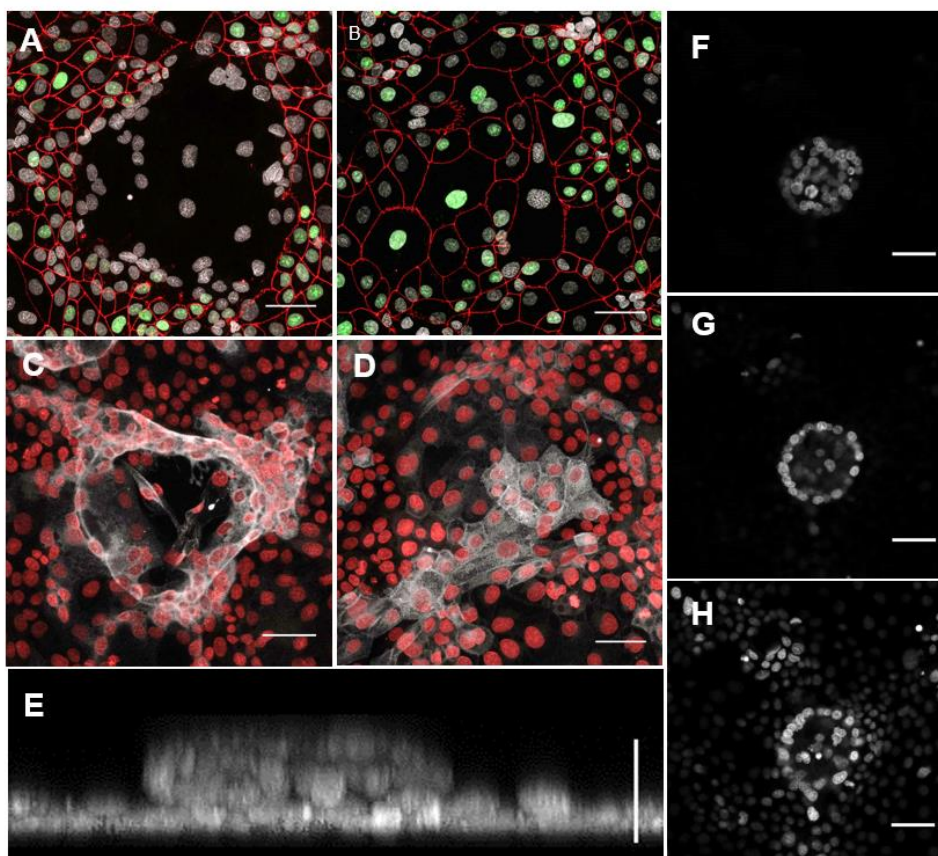
Fonte: Elaborado pelos autores.

Para investigar o potencial de maturação do TSC derivado de iPS, avaliamos a capacidade das células TS derivadas de iPS de se diferenciarem em tipos de células trofodermas maduras. As células-tronco trofoblásticas proliferativas são células-tronco bipotenciais com capacidade de se diferenciarem em citotrofoblastos sinciciais e

citotrofblastos extravilosos (132). O citotrofblasto sincicial é uma camada epitelial multinucleada terminalmente diferenciada que se infiltra no endométrio materno. As células multinucleadas se formam a partir de múltiplas fusões de células citotrofblasticas. Quando as células TS cresceram até alta densidade, observamos a formação espontânea de células multinucleadas por microscopia de campo claro.

As células sinciciais multinucleadas são núcleos negativos para Ki-67 rodeados pelas células positivas da proteína ZO-1 de junção apertada (**Figura 26A e B**). Observou-se grandes áreas de núcleos fundidos de formação sincicial (**Figura 26A**). Uma das características dos sinciciotrofblastos é que produzem e secretam altos níveis de gonadotrofina coriônica humana (*hCGB3*), um hormônio essencial para a diferenciação, imunossupressão e invasão placentária dos citotrofblastos (133). Observamos a expressão de *hCGB3* em regiões de culturas diferenciadas de citotrofblastos, indicando diferenciação em sinciciotrofblastos (**Figura 26C**). Os trofblastos extravilosos expressam o antígeno leucocitário humano G (HLA-G), a principal molécula envolvida na evasão do sistema imunológico materno após a invasão placentária do endométrio. Também foram detectadas regiões de células *HLA-G* positivas, indicando que as células TS derivadas de iPS são capazes de diferenciação em citotrofblasto extraviloso (**Figura 26D**). A expressão de *hCGB3* e HLA-G foi específica para células TS derivadas de iPS e não detectada em células iPS.

Figura 26 - As células TS derivadas de iPS se diferenciam em citotrofoblasto sincicial, citotrofoblastos extravilosos e estruturas semelhantes a vilosidades.



Fonte: Elaborado pelos autores.

Como é sabido que o início da diferenciação ocorre em baixas concentrações de oxigênio, os meios foram investigados quanto à resposta frente a condições de hipóxia. Wakeland e cols. (134) mostraram *in vivo* que a diferenciação de sinciciotrofoblasto depende de baixas concentrações de oxigênio, dessa forma fizemos uma transição de uma atmosfera de 20% a 5% de oxigênio. Corroborando com os achados anteriores, a baixa concentração de oxigênio fez com que as células tivessem inibição por contato mínima e estruturas semelhantes a vilosidades 3D que se projetam da superfície da monocamada da placa de cultura (**Figura 26 E-H**). Quando visualizadas por microscopia confocal, essas estruturas semelhantes a vilosidades têm aproximadamente 50 microns de altura. Em resumo, o TSC derivado de iPS diferenciou-se em citotrofoblastos vilosos positivos para *HLA-G* e citotrofoblastos sinciciais positivos para *hCGB3*.

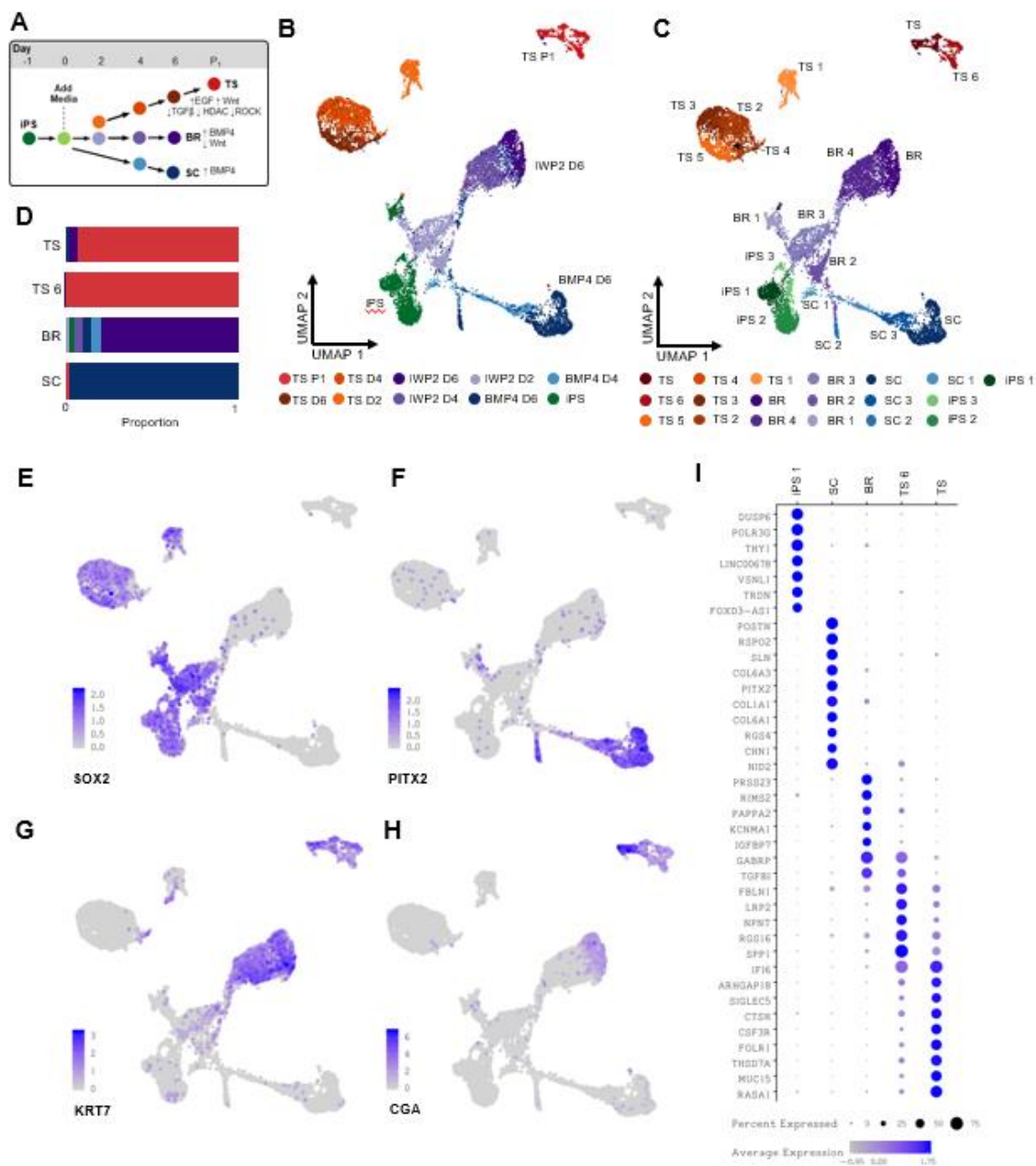
5.3.2 Sequenciamento de RNA de *Single-cell* revela programação transcricional exclusiva para especificação de TS derivada de iPS

Para definir os eventos moleculares envolvidos na especificação de células iPS para trofotoderme, realizamos análise de RNAseq de *Single-cell* durante uma passagem de tempo (dias 0, 2, 4, 6, 8) de células iPS humanas especificadas para células trofoblásticas placentárias por três condições distintas:

- 1) adição de BMP4 (135),
- 2) adição de BMP4 e inibição de Wnt com o inibidor de molécula pequena IWP2(136),
- 3) ativação da condição de células-tronco pluripotentes de trofoblasto (TSC) de EGF e Wnt, inibição de TGFbeta, HDAC e sinalização ROCK para gerar células-tronco trofoblásticas (**Figura 27A**)

O sequenciamento foi gerado para células iPS antes da diferenciação (dia 0), para BMP4 nos dias 4 e 6, para BMP4 + IWP2 nos dias 2, 4 e 6, e nos dias 2, 4, 6 e 8 para a condição de TS, resultando em 10 transcriptomas de uma única célula (**Figura 27**). Na condição TS, o dia 8 é um ponto de tempo dois dias após a primeira passagem anotada como p1. Nas condições BMP4 e BMP4 + IWP2, nenhuma célula proliferativa sobreviveu à passagem, portanto, o ponto de tempo p1 não foi sequenciado. Usando o método de sequenciamento de célula única baseado em gotículas altamente paralelas Drop-Seq (122), 9.821 células de alta qualidade foram obtidas após a remoção de células com menos de 1000 genes detectados e mais de 20% do mapeamento mitocondrial. A proporção média de leituras de mapeamento para genes mitocondriais em todas as células foi de 3,9%, indicando boa viabilidade. O número total de genes detectados variou entre 21804 para BMP4 + IWP2 dia 4 a 27134 para TS dia 4 ponto de tempo com entre 563 a 1572 células de alta qualidade por ponto de tempo.

Figura 27 - Sequenciamento (RNaseq) temporal de *Single-Cell de iPS* até a diferenciação em TB.



Fonte: Elaborado pelos autores.

No intuito de explorar os dados de uma maneira imparcial, realizamos a redução de dimensionalidade para os genes mais variáveis em todas as células com incorporação de Análise de Componentes Principais (PCA) e Aproximação e Projeção Uniforme de Variedades (UMAP). De maneira geral, as células se separaram por estágio de diferenciação e condição, com dimensão 2 de UMAP majoritariamente capturando a condição de diferenciação (**Figura 27B**). Para identificar grupos de células transcricionalmente semelhantes, realizamos uma

análise de agrupamento baseada em gráficos de k vizinhos mais próximos usando a distância euclidiana dos 30 primeiros PCs em Seurat (123). Dessa maneira, identificamos 19 *agrupamentos*, e os *agrupamentos* continham principalmente um único ponto no tempo e condição de mídia (**Figura 27B-D**). Nomeamos então os *agrupamentos* de acordo com os processos de diferenciação através da identificação de marcadores em cada um deles, e para facilitar a compreensão, identificaremos aqui o dia da cultura para uma interpretação de linha do tempo cronológica (além do tempo de diferenciação estimado por algoritmo - pseudotempo, que será discutido mais adiante). Os agrupamentos iPS 1, 2 e 3 continham principalmente células iPS com pequenas contribuições desde os primeiros dias de todas as três condições de diferenciação.

As células da condição BMP4+IWP-2 foram encontradas predominantemente em um grupo de cinco grupos heterogêneos denominados BR (responsivos a BMP). BR1 é composto por partes aproximadamente iguais de células iPS e células tratadas com BMP4/IWP-2 dia 2. As células BMP4/IWP2 constituíram a maioria das células nos outros agrupamentos BR ao lado de uma população menor de células BMP4. BR2 e BR3 foram predominantemente derivados de células BMP4+IWP2 do dia 2, enquanto o dia 4 contribuiu para BR4 e o dia 6 contribuiu para BR. As células da condição BMP4 formaram quase exclusivamente quatro agrupamentos denominados SC (célula estromal). SC1 e SC3 eram predominantemente compostos por células do dia 4 da diferenciação, enquanto SC2 e SC eram predominantemente células do dia 6. Além disso, o *agrupamento* SC continha uma pequena contribuição de células originárias do dia de diferenciação final da condição TS denominada TSP1 (TS, passagem 1). As células da diferenciação TS povoaram quase exclusivamente 7 agrupamentos, denominados TS. TS1 era composto exclusivamente por células do dia 2 e TS 2-5 era composto por uma mistura de células dos dias 2-6. O dia TS final (TSP1) separado em dois grupos denominados TS e TS6, o que confirma as observações de campo claro de 2 populações de células. Curiosamente, 14 células de BMP4 + IWP2 d6 e 4 células de BMP4 d6 contribuem para o *agrupamento* TS final (**Figura 27D**).

Para entender como os perfis transcricionais mudam durante a diferenciação, perguntamos quais células expressam genes canônicos pluripotentes, específicos de trofotoderma e mesoderme. O fator de transcrição de células-tronco (SOX2) foi o mais expresso no estágio iPS e está ausente dos agrupamentos BR, TS e SC mais diferenciados (**Figura 27E**). A expressão de SOX2 é mantida em um nível mais alto nos dias iniciais das condições de TS em comparação com o tratamento com BMP+WNT e BMP e ausente nas células TS mais diferenciadas. O fator de transcrição do homeodomínio pareado 2 (PITX2) é

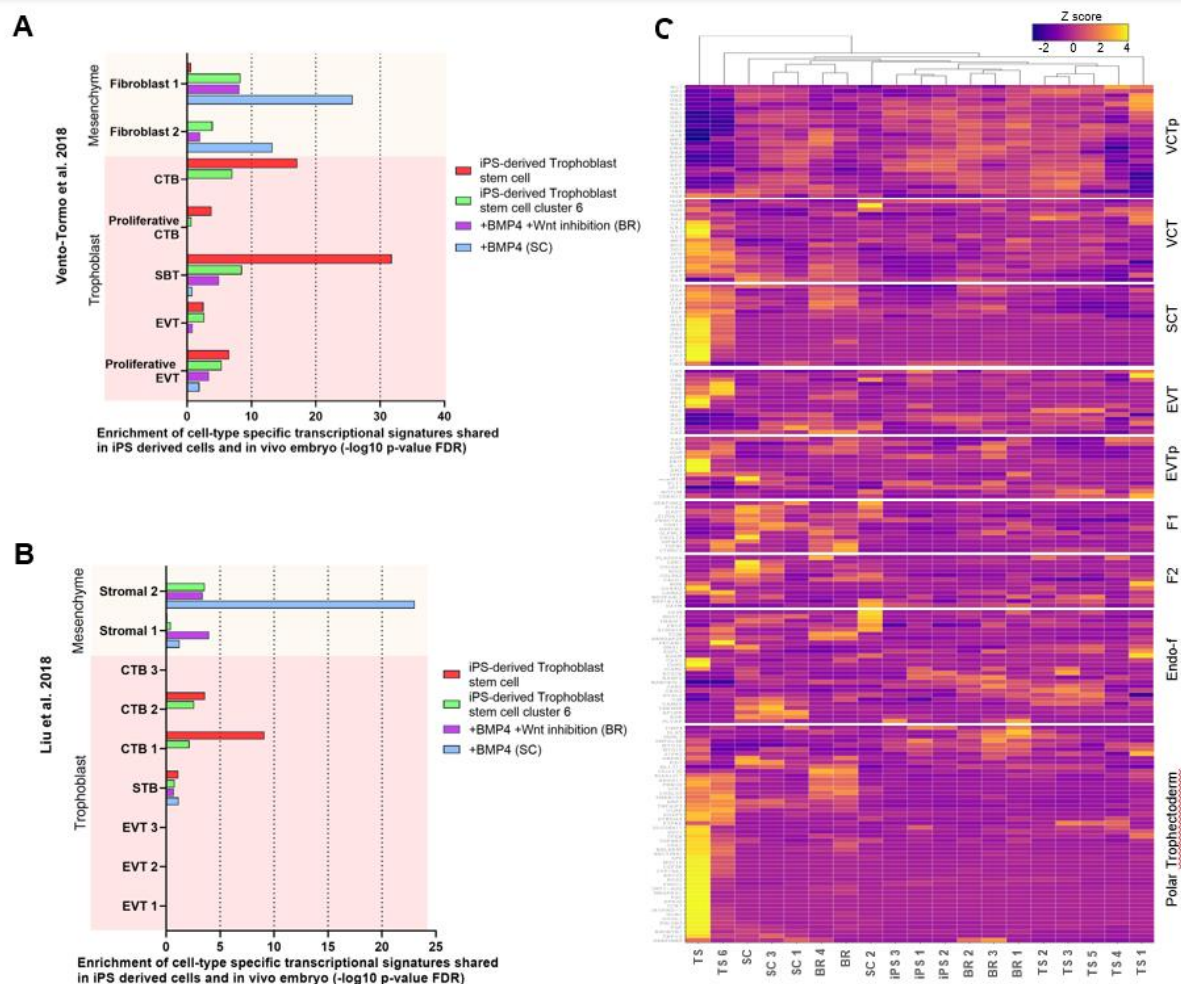
altamente expresso nos estágios posteriores da diferenciação de BMP4, sugerindo que as condições de BMP4 geram células do estroma placentário (Figura 3F). Em RNAseq de célula única da placenta humana, PITX2 foi identificado como o gene marcador superior para um tipo de célula comunicante fetal estromal F1 por Vento-Tormo (65,137). A queratina 7 (KRT7) e a subunidade alfa da gonadotrofina coriônica humana (CGA), são expressas em citotrofoblastos derivados de células-tronco e na placenta in vivo (138). Indicando a especificação da programação de transcrição do trofotoderma, encontramos que KRT7 e CGA são expressos nos estágios posteriores da condição BMP + WNTi e condição TS (**Figura 27G-H**). No geral, ambas as células tratadas com TS e BMP4 + WNTi demonstraram uma expressão aumentada para a maioria dos marcadores de TB, incluindo GATA3 e TFAP2C, enquanto as condições de BMP4 sozinhas demonstraram uma expressão mais associada ao mesoderma, com expressão significativa de GATA4, TBXT e PDGFRA. Em camundongos, o homeobox 2 do tipo Caudal (CDX2) está envolvido na segregação das linhagens ICM e trofotoderma no estágio de blastocisto, reprimindo OCT4 / POU5F1 e NANOG no TE. A hiper-expressão de CDX2 em células-tronco embrionárias de camundongo causa diferenciação para células-tronco do trofoblasto (139). Em humanos, a expressão de CDX2 é iniciada após a formação de blastocisto e tem padrões de expressão variáveis no trofotoderma (140). Semelhante às células TS derivadas de embriões humanos primários, CDX2 não é expresso no estado TS diferenciado, e há expressão esparsa de CDX2 em d4 de todas as diferenciações.

O próximo passo foi investigar quais genes são preferencialmente expressos em cada *agrupamento* pelo teste não paramétrico de soma de postos de Wilcoxon (p -valor aj. $<0,05$; $\log_{FC} > 0,25$) (**Figura 27I**). Definimos marcadores de população de células comparando a expressão gênica de cada população de células em comparação com todas as outras células. DUSP6 e THY1 são genes primários do estado das células-tronco e estão entre os principais genes marcadores identificados nas células iPS. Os genes marcadores de agrupamento SC incluem genes de sinalização WNT, RSPO2 um agonista de WNT (141) e PITX2 que interage com a sinalização WNT e regula a expressão de colágeno durante o desenvolvimento (142). Os marcadores do *agrupamento* BR incluem PRSS23, uma serina protease envolvida na transição epitelial para mesenquimal dependente de caracol (143) e reguladores da sinalização do fator de crescimento semelhante à insulina IGFBP7 (144) e PAPPA2 (145). Os marcadores dos *agrupamentos* TS incluem genes envolvidos na sinalização de EGFR NPNT (146), IFI6 (147), e MUC15 (148). Além disso, ARHGAP18 é uma proteína ativadora de Rho GTPase que funciona como um efetor a jusante da sinalização de YAP (149) recentemente implicada na especificação HiPS para células-tronco trofoblásticas bipotentes (150).

Em resumo, TSC, BMP4 e BMP4 com inibição de WNT resultaram em tipos celulares diversos. A ativação de WNT e EGF e a inibição de TGF β , HDAC e ROCK especificam iPSC para citotrofoblastos, BMP4 sozinha induz destinos mesenquimais e a inibição de WNT no contexto de BMP4 gera uma pequena porcentagem de células TS. Fornecemos um navegador da web dos perfis de expressão de *Single-Cell* (<http://erwinpaquolalab.libd.org/publications/singlecellTS>) e os genes expressos diferencialmente para todas as comparações de agrupamento em pares.

Dando seguimento a interpretação das diferenciações, seria importante que comparássemos as células encontradas com as descritas na placenta humana. Existe um debate sobre a autenticidade dos trofoblastos diferenciados de iPSC ou célula trofoblástica pluripotente. Por essa razão, os transcriptomas de cada um dos agrupamentos encontrados aqui foram comparados com bancos de dados robustos, publicados por Vento-Tormo e cols (2018) e Liu e cols. (2018). A análise de enriquecimento usando Análise de expressão específica de célula (CSEA) mostrou que genes específicos dos diferentes tipos de célula foram preferencialmente expressos em cada tipo de célula / agrupamento dos 19 agrupamentos pela estatística de probabilidade de índice de especificidade (pSI) $p < 0,05$. Em seguida, testamos se os genes específicos do tipo de célula previamente identificados por estudos de *Single-Cell* da placenta humana estão hiper-representados nos genes específicos do tipo de célula de nossa especificação iPSC por teste hipergeométrico e aplicamos a correção de Bonferroni para comparações múltiplas, considerando todos os listas de genes [$\alpha = 0,05 / (19 \times (38 + 14)) = 5,1 \times 10^{-5}$]. O estudo de Vento-Tormo identificou um conjunto de 38 tipos de células de tecidos maternos (decídua e sangue) e fetais (placenta) do primeiro trimestre e relatou 30 genes específicos de tipos celulares (65). Descobrimos que o perfil transcricional de todos os agrupamentos iPSC e agrupamentos de pontos de tempo dias 2-4 não teve enriquecimento significativo para qualquer um dos agrupamentos de tipo de célula da placenta. Em contraste, o agrupamento TS derivado de iPSC mais maduro foi altamente enriquecido para o citotrofoblasto sincicial (SCT) e agrupamento de citotrofoblasto viloso (VCT) ($p = \log_{10}^{-32}$ e \log_{10}^{-17}) (Figura 4A).

Figura 28 - As diferenciações de trofoblasto estão enriquecidas para os tipos de células da placenta do primeiro trimestre. (A, B) Análise de enriquecimento de conjunto de genes MAGMA de diferenciações de células HiPS destacando o enriquecimento de agrupamentos de células para tipos de células identificados em placentas do primeiro trimestre. (C) Mapa térmico composto de expressão gênica normalizada dos principais marcadores únicos de tipos de células da placenta de origem fetal identificados em Vento-Tormo et al. 2018 e marcadores do trofotoderma polar (top 50) identificados em Petropolis et al. 2016



Fonte: Elaborado pelos autores.

O perfil de expressão do *cluster* TS derivado de iPS foi altamente específico para SCT e VCT, sem enriquecimento significativo para qualquer outro tipo de célula na interface materno-fetal. O *agrupamento* SC D6 BMP4 foi mais enriquecido nos dois *agrupamentos* de fibroblastos (F1 e F2) ($p = \log_{10}-26$ e $p = \log_{10}-13$). Vento-Tormo descreveu esses fibroblastos placentários como células estromais mesenquimais de origem fetal que derivam do endoderma primitivo que expressa GATA4, GATA6, PDGFR α e SOX17. As condições BMP4 + IWP2 (BR) não foram enriquecidas para qualquer tipo de célula placentária específico e, em vez disso, tiveram um enriquecimento mais fraco para vários agrupamentos, incluindo citotrofoblasto sincicial (SCT) e tipo de célula endodérmica F1 derivada da decídua materna e dS1. Em suma,

as células TS derivadas de iPS são altamente enriquecidas para a expressão de genes específicos para citotrofoblasto sincicial e citotrofoblasto viloso de placenta humana ex vivo. Para o enriquecimento do banco de dados de Liu e cols (8 semanas de gestação) (**Figura 28B**). Mais uma vez, encontramos um forte enriquecimento para o TS derivado de iPS (citotrofoblasto, log10-9). As condições de BMP4 sozinha e BMP4 + IWP2 foram enriquecidas para duas células mesodermas diferentes, identificadas por Liu et al (log10-23 e log10-4). Para visualizar os padrões de expressão específicos do tipo de célula da placenta entre os tipos de células derivadas de iPS, traçamos um mapa de calor (*Heatmap*) dos principais marcadores únicos dos tipos de células fetais identificados em Vento-Tormo e marcadores de trofotoderma polar identificados em (Petropoulos et al., 2016) (**Figura 28C**). Consistente com a Análise de Enriquecimento de Conjunto de Genes, observamos alta expressão para a maioria dos genes marcadores únicos previamente identificados em citotrofoblastos de vilosidades do trofotoderma polar e sinciciotrofoblasto. Também comparamos os perfis de expressão de todos os *agrupamentos* celulares derivados de iPS com a linhagem previamente identificada e os genes marcadores de conjuntos de dados pré-implantação (Stirparo et al 2018). Novamente, encontramos a maior expressão dos genes do trofoblasto KRT18, TEAD3, GATA3 e GATA2 em TS derivado de iPS, com expressão moderada em células BR. Os genes do endoderma primitivo SOX17, GATA4, GATA6 foram os mais expressos em células SC. Quando comparados com as células trofoblásticas primárias (151), confirmamos que as células TS expressaram o nível mais alto de citotrofoblasto e genes específicos de EVT. Em resumo, as células iPS, especificadas por meio de condições TS, têm um perfil transcricional altamente homólogo aos citotrofoblastos retirados diretamente da placenta humana, enquanto a diferenciação BMP4 gera células mesenquimais da placenta.

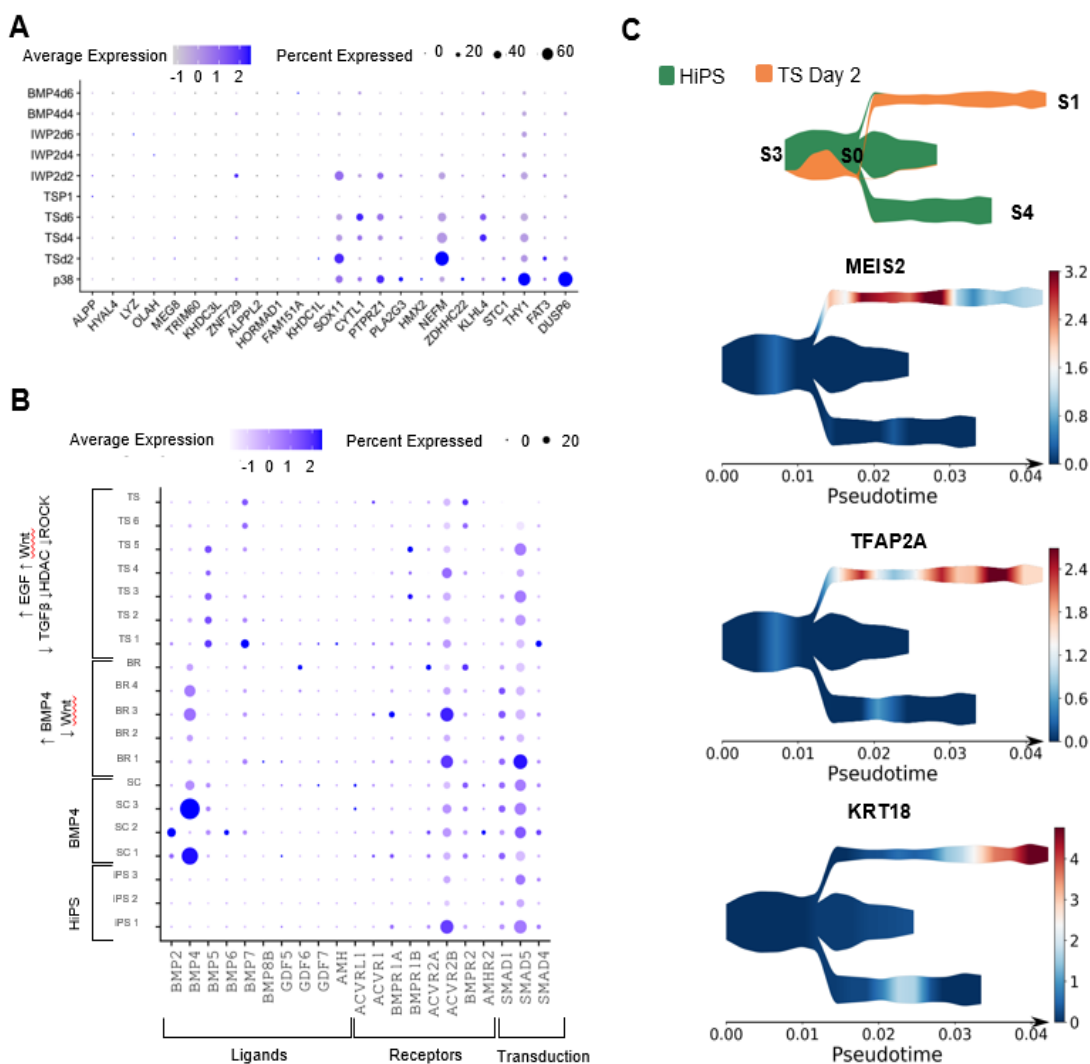
Um estudo prévio mostrou que as células-tronco iniciadas são restritas em sua potência e incapazes de se diferenciar em TSC, enquanto as células-tronco ingênuas se diferenciam prontamente em TSC porque estão prontas para especificar o trofotoderma (61). Portanto, investigamos se as células tratadas com iPS ou TSCM iniciadas adotam uma programação de células-tronco ingênuas durante a diferenciação. Para 12 genes marcadores identificados anteriormente (152), os transcritos ingênuos estavam quase ausentes. Cinco genes *HORMAD1*, *ALPPL2*, *KHD3L*, *TRIM60* e *HYALA* não tiveram expressão observada em nenhuma célula (Figura 5A, Tabela S4). Menos de 1% das células para qualquer condição tinham expressão detectável para seis outros genes marcadores *ALPP*, *OLAH*, *LYZ*, *MEG8*, *KDC1L* e *FAM151A*. *ZNF729* foi o único gene expresso em 0,51% das células TSp1 a 14,69% do IWP2d2. Em contraste, os genes marcadores de iniciação foram expressos de forma consistente em

agrupamentos iPS e muitas vezes ao longo do estado de diferenciação. Por exemplo, *DUSP6* foi detectado em 74,87% das células iPS. Em suma, a transição de iPS iniciada para citotrofoblastos não passou pelo estado de ingenuidade.

Como a via BMP é a mais explorada para diferenciação de TB, buscamos toda a lista de sinalização de BMP para ver se estavam presentes. Nos agrupamentos TS, houve pouca expressão de BMP4, enquanto na cultura com BMP4 e BMP4+IWP2 observou-se maior expressão. BMP7 foi mais expresso em células tratadas com TSCM em comparação com iPS (BMP5: 0- 0,67% de iPS vs 24,72% de células TS1 e BMP7: 1,01-4,71% de iPS vs 29,06% de células TS1). As proteínas efetoras da via BMP, SMAD1, 4 e 5 são expressas nos primeiros dias de todas as condições e estão ausentes do *agrupamento* TS maduro. Concluímos que a sinalização BMP5 e BMP7 padrões de especificação TS derivada de iPS ao invés de BMP4.

Para a avaliação dos genes mais significantes para especificação de TS nos estágios iniciais, apenas células do dia 2 foram avaliadas na inferência de trajetória de *Single-Cell* e estimativa de pseudotempo (STREAM). Inicialmente, o STREAM identificou um braço com iPS e TS dia 2. Também identificou braços contendo células iPS (S0-S2, S0-S4), refletindo o contínuo de pluripotência, bem como um ramo de dia 2 (S0-S1) refletindo a saída da pluripotência (**Figura 29C**).

Figura 29 - Especificação iPS para células-tronco trofoblásticas inicia com reconexão TFAP2A sem ativação de programas de células-tronco ingênuas.



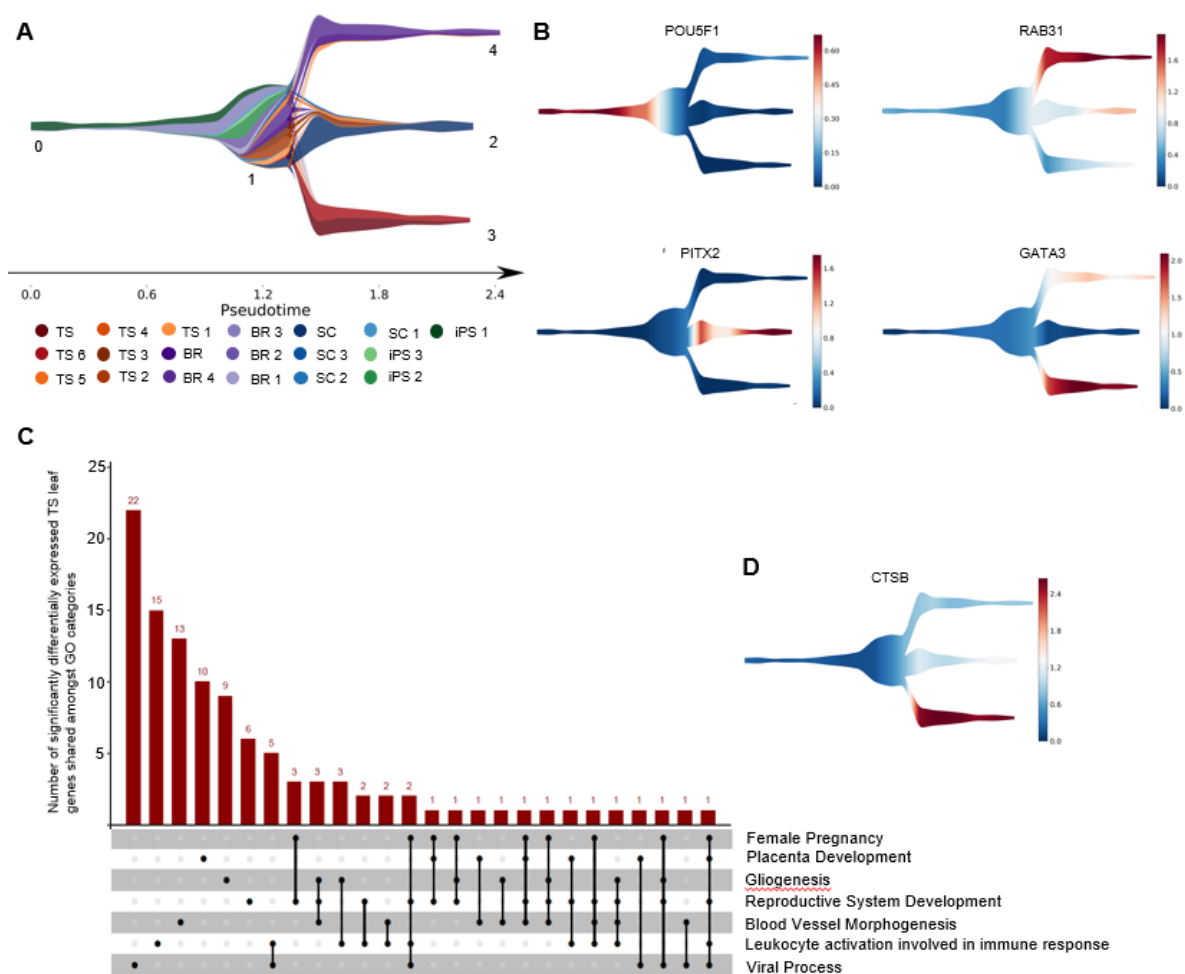
Fonte: Elaborado pelos autores.

No braço S3-S0, o nó S3 era predominantemente composto por células iPS, enquanto as células TS do dia 2 são mais abundantes em relação a S0 (**Figura 29C**). Em seguida, identificamos os genes que contribuem para a especificação, identificando aqueles altamente expressos dentro de um braço que estão correlacionados com a transição ao longo dos braços da especificação (ver métodos). Os principais genes de transição correlacionados positivamente ao longo do ramo S3-S0 incluem *LRRN3* que amplifica a sinalização de MAPK através de *EGFR*(153), *MEIS2* no qual codifica para uma proteína homeobox TALE que, por sua vez, promove a progressão do tumor através da ação combinada de duas isoformas, aumentando a translocação de YAP para o núcleo, ativando a via WNT / b-catenina (154), *BMP5*, *BMPRII* e *TFAP2A*, um fator de transcrição chave para a supressão da pluripotência e expressão de

genes associados ao trofoblasto (155) (**Figura 29C**). Esses genes foram expressos não só ao longo do braço S3-S0, mas também ao longo do braço S0-S1. O braço S0-S1 captura a transição do estado pluripotente para o estado TS com maior expressão de citoqueratinas (*KRT18*, *KRT19*, *KRT8*) e proteínas de ligação de cálcio (*S100A11*, *S100A10*) conhecidas por serem expressas no trofotoderma (156). Também encontrado como um gene de transição superior, o *TMSB4X* está implicado na diminuição da pluripotência de células trofoblásticas progenitoras de placentas humanas de primeiro trimestre, aumentando a atividade de NOTCH1 (Huang et al., 2016; Lee et al., 2018; Lv et al., 2013). Em resumo, a programação transcricional do trofotoderma ocorre no início da especificação e utiliza fatores regulatórios associados ao trofotoderma previamente identificados, incluindo *TMSB4X*, *YAP*, *BMP5* e *TFAP2A*.

Para entender as mudanças transcricionais associadas à trajetória completa do iPS para a especificação do trofotoderma, repetimos a análise STREAM com o conjunto de dados completo do iPS para o dia 8. O STREAM identificou 4 pontos de ramificação principais entre todas as células (**Figura 30**).

Figura 30 - A análise STREAM de pseudotempo das diferenciações trofoblásticas revela a suprarregulação do mesênquima estromal após o tratamento com BMP4 e a programação transcricional do trofotoderma no tratamento TS.



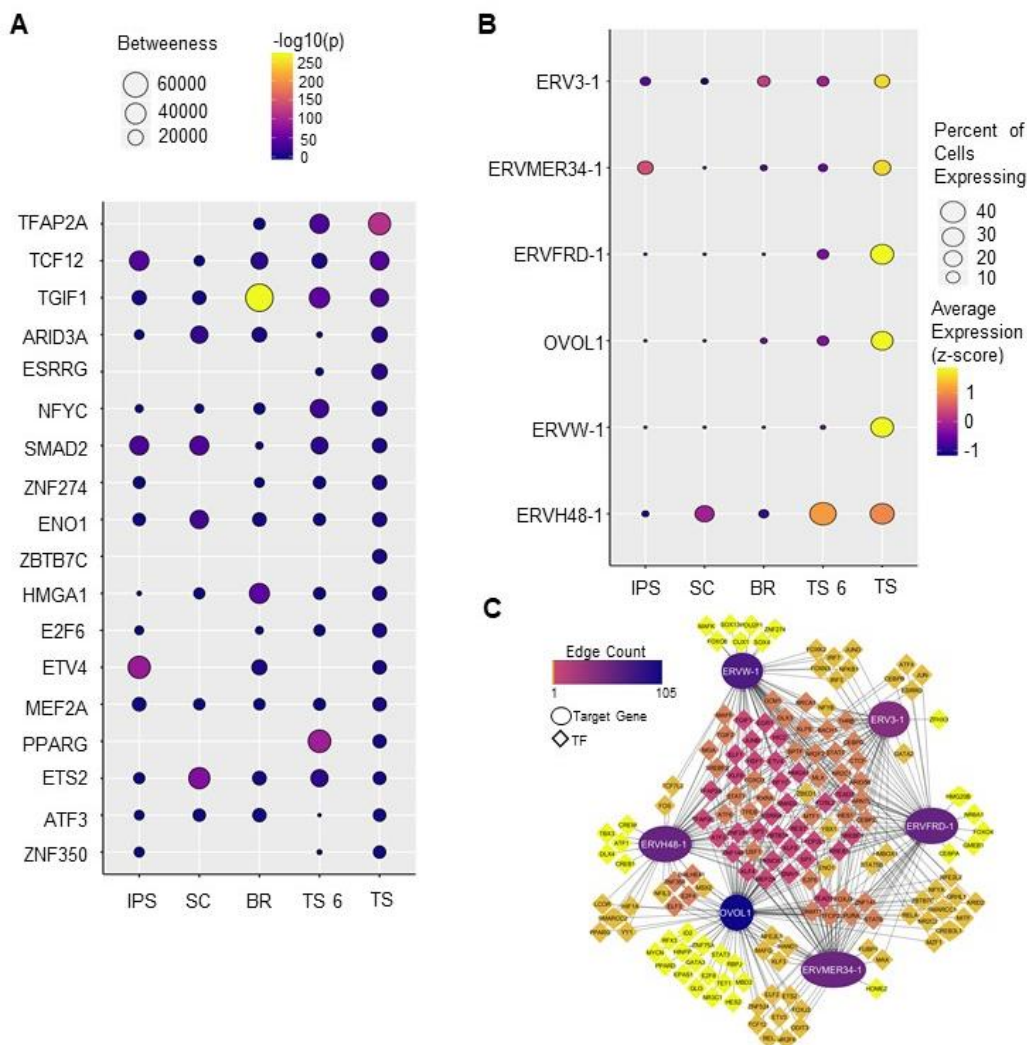
Fonte: Elaborado pelos autores.

Os primeiros dias de diferenciação, para todas as condições, foram misturados no pseudotempo, demonstrando um estado de transcrição semelhante nas fases iniciais de comprometimento de IPS quando comparados no contexto das células dos dias 6 e 8. Semelhante à projeção UMAP (**Figura 27B-C.**), o STREAM inferiu a mistura de dois estágios iPS com células dos tratamentos do dia 2. Ademais, o STREAM inferiu três pontos principais de ramificação que se separaram com base nas células mais diferenciadas de cada condição. Para revelar os estados transcricionais que contribuem para a ramificação, realizamos análise de expressão diferencial para identificar genes altamente expressos dentro de um ramo que são diferencialmente expressos em comparação com todos os outros ramos, denominado análise de detecção de gene de folha em STREAM (ver métodos) (**Figura 30B**). Para o ponto zero de ramificação iPS, genes específicos de estado pluripotente, incluindo *DNMT3A* e *POU5F1* /

OCT4, foram contrastantes na análise de expressão diferencial. Para o ramo de células expostas a BMP4 (SC), os marcadores do mesoderma, incluindo *PITX2*, foram identificados como diferencialmente expressos. Para o ramo BMP + IWP2, genes associados à placenta foram identificados incluindo *RAB31*, uma GTPase altamente expressa na placenta que regula o tráfego da membrana intracelular de *EGFR* ligado ao receptor (157). Para o ramo TS, genes significativamente diferencialmente expressos foram associados com comprometimento trofotoderma e implantação e pré-eclâmpsia, incluindo *GATA3*. Realizamos uma análise de enriquecimento de *Gene Ontology* para processos biológicos usando os 300 principais genes Folha expressos de forma significativa diferencial para o ramo TS. As categorias GO significativamente super-representadas incluíram desenvolvimento placentário, gravidez, processos virais (**Figura 30C-D**).

Encontramos que um conjunto de genes envolvidos em processos virais são regulados positivamente nos últimos pontos de tempo de diferenciação de TS (**Figura 30C**). Isso ecoa um crescente corpo de evidências demonstrando que os retrovírus endógenos desempenham um papel fundamental na diversificação evolutiva da placenta de mamíferos (158). Portanto, nos interessou investigar qual seria a expressão de genes humanos derivados de ERV (do inglês, *Endogenous Retroviruses*) durante a especificação do trofotoderma. Encontramos especificamente, hiper-expressão de cinco genes derivados de ERV no tipo de célula TS mais diferenciado (**Figura 31A**).

Figura 31 - ZBTB7C e retrovírus endógenos são centrais e exclusivos da rede reguladora do gene TS.



Fonte: Elaborado pelos autores.

Em sequência, nos interrogamos onde estariam os genes ERV na rede regulatória de TS. O uso dessa metodologia foi parte da tese do aluno André Rocha Barbosa (Teses USP, Bioinformática, 2020), em que ele propôs o uso de uma rede que liga um fator de transcrição e genes-alvo usando o algoritmo de passagem de atributos entre redes para assimilação de dados (PANDA). A rede PANDA integra informações de dados de *motif* de sequência de FT, expressão gênica e interação proteína-proteína (PPI) em uma abordagem de passagem de mensagem (159). Resumidamente, o algoritmo da PANDA faz a suposição de que os genes direcionados pelo mesmo FT têm maior probabilidade de serem co-expressos do que os genes que são regulados por um FT diferente. Observamos que as interações ERV (fatores de transcrição + genes derivados de ERV) são proeminentes apenas no *agrupamento* TS mais diferenciado. Descobrimos que as interações regulatórias para cinco genes derivados de ERV

(*ERVH48-1*, *ERV3-1*, *ERVMER34-1*, *ERVW-1* e *ERVFRD-1*) são altamente exclusivas para a rede do *agrupamento* TS, enquanto apenas 0 ou 1 gene derivado de ERV foi encontrado nos outros agrupamentos de células (**Figura 31B**). Os *agrupamentos* SC e SC 3 (células BMP4) exibiram interações para o gene *ERVH48-1*, e os *agrupamentos* BR 1, BR 3 e BR 4 (células BMP4 + IWP2) apresentaram o gene *ERV3-1*. Curiosamente, os genes englobados nas interações regulatórias de ERV de células TS derivadas de iPS foram enriquecidos em vias biológicas relacionadas ao metabolismo hormonal, diferenciação celular e ao sistema imunológico (valor de p ajustado $<0,01$) (**Figura 31C**). Além disso, observamos que as interações regulatórias do ERV também são enriquecidas para categorias como desenvolvimento da placenta, diferenciação de células trofodérmicas e formação de sinciotrofoblasto. Para entender a importância das interações ERV em sua sub-rede, perguntamos a seguir qual é a força das arestas ERV no contexto de rede do *agrupamento* TS. Descobrimos que os pesos das interações regulatórias de ERV foram altamente classificados entre todas as conexões de rede (*Wilcoxon Rank* soma $p = 1,7e-31$, valor p permutado = 0,001). Basicamente, descobrimos que os genes derivados do ERV participam de uma sub-rede regulatória, especificamente nas células TS.

5.4 DISCUSSÃO PLACENTA

Existe um debate quanto à capacidade de se formar células de linhagem de trofodermis a partir de células-tronco humanas. Aqui, demonstramos que as células-tronco humanas preparadas retêm uma ampla potência para gerar linhagens extra embrionárias que são transcricionalmente muito semelhantes à placenta. Em termos de vias de diferenciação, o principal achado deste estudo foi que células diferenciadas em trofoblastos oriundas de iPS não dependem do BMP4 para diferenciação, e, em vez disso, inibem as vias de *EGF* e *TGF- β* , e ativam a via WNT. O Valproato de Sódio é um ácido graxo de cadeia curta que atua como um modificador epigenético, usado para promover a reprogramação de fibroblastos humanos para células iPS, e sua administração inibiu as histonas desacetilases (HDAC). As células TS derivadas de iPS são altamente semelhantes em termos de transcrição aos citotrofoblastos encontrados na placenta humana. As células TS proliferam rapidamente, exibem núcleos de tamanhos muito variáveis, tornam-se rapidamente multinucleadas, conforme observado no trofoblasto pós-implantação inicial, e expressam fortemente proteínas extravilosas e sincitiotrofoblásticas associadas (HLA-G, HCG- β) em subpopulações. Após oito dias de

especificação, as células TS derivadas de iPS exibem uma rede de transcrição exclusivamente ativa de retrovírus endógenos humanos semelhantes ao trofoblasto *in vivo*. Demonstramos que, em condições de hipóxia, as células TS derivadas de iPS criam estruturas semelhantes a vilosidades 3D, condições que provavelmente serão úteis para futuros modelos organóides 3D.

Nossos resultados são consistentes com o modelo de que o estado inicial está em um contínuo de potência, ao invés de um estado fixo e restrito (160). As diferenças metodológicas e o uso frequente de BMP4 provavelmente fundamentam as diferenças relatadas anteriormente na capacidade de células-tronco pluripotentes iniciadas e ingênuas de gerar TSCs. Confirmamos que o modelo BMP4 pode gerar células com características de trofoblasto, mas nas condições propostas (TS) aqui, as células especificadas para BMP4 são menos semelhantes aos citotrofoblastos *in vivo*. Após a derivação recente de células primárias em trofoblásticas humanas (132,150), algumas tentativas foram feitas para derivar TSCs de células tronco/iPS de uma série de estados pluripotentes (150). Mischler et al. (2019) (150) usou células tronco embrionárias iniciadas com um protocolo híbrido que combinou um tratamento de três dias de BMP4 e SB431542 em cultura de longo prazo em meio TM4 (CYM554, um agonista S1PR3, CHIR99021, A83-01 e FGF10) para gerar uma população semelhante a CDX2 + TSC (132,150). Quando transferidas para a mídia TSCM, essas células perderam a expressão de CDX2 e adotaram características de TSCs. Omitindo TM4 com transferência direta de células tratadas com BMP4 para TSCM resultou em um enriquecimento gradual de células epiteliais comparável com sua população semelhante a CDX2-TSC. Curiosamente, esse processo reflete a progressão da diferenciação que observamos quando tratamos diretamente o iPS preparado com o TSCM. Nossos dados sugerem que BMP4 não está envolvido nesta transição. As células semelhantes a TSC também foram geradas a partir de células iPS iniciadas usando um sistema de cultura de micro-malha com meio de células-tronco pluripotentes iniciadas seguido por cultura em TSCM(161), embora a comparação dessas células com células da placenta *in vivo* seja menos extensa. Em contrapartida, Dong et al. 2020 relatou que apenas células tronco ingênuas e não iniciadas poderiam gerar células semelhantes a TSC por exposição direta a TSCM (162). Os estados de pluripotência ingênuo ou iniciado apresentam propriedades ligeiramente diferentes. O estado ingênuo representa o estado celular da massa celular interna do blastocisto pré-implantação, enquanto o estado inicializado é representativo das células epiblasticas pós-implantação. As células-tronco com potencial expandido humano (EPSCs) também podem gerar células semelhantes a TSC quando diferenciadas em TSCM (163). Interessantemente, o grupo de Gao não conseguiu de gerar células semelhantes a trofoblastos pluripotentes usando células-tronco embrionárias ingênuas ou iniciadas em condições

semelhantes. Aqui, demonstramos que o contínuo de pluripotência presente no estado inicializado permite prontamente a especificação para trofotoderma.

A diferenciação robusta de células iPS iniciadas para trofotoderma permite a modelagem de doenças da placenta a partir de células específicas do indivíduo, sem a necessidade de obtenção do tecido primário no primeiro trimestre. As respostas adaptativas da placenta diante situações adversas podem alterar o desenvolvimento fetal e programar o feto para doenças posteriores, incluindo diabetes e transtornos do neurodesenvolvimento(67). Seja como uma barreira protetora, ou um componente do sistema imunológico inato, os citotrofoblastos e seus progenitores desempenham um papel importante desde o início da gestação em respostas a estressores pré-natais. Sugerimos que, o modelo aqui desenvolvido a partir de iPS com diferenciação para células trofoblásticas poderá ser uma valiosa ferramenta para contribuir na compreensão de mecanismos envolvidos nas respostas um ambiente adverso.

6 CONSIDERAÇÕES FINAIS

A proposta desta tese tinha como objetivo geral avaliar a metilação do DNA e sua associação com exposições estressoras gestacionais e desfechos de crescimento neonatais e do neurodesenvolvimento, assim como desenvolver um modelo de diferenciação de placenta para estudar estas relações. Observamos que o Relógio Epigenético acelerado ao nascimento associou-se negativamente com os índices de crescimento ao nascimento, porém não comprovamos que a associação entre exposição ao estresse gestacional e desfechos de crescimento é mediada por alterações da metilação. Mostramos que o efeito de intervenções psicossociais em desfechos do neurodesenvolvimento pode ser mediado pela metilação diferencial do DNA. Especificamente, confirmamos a existência de sítios e regiões diferencialmente metiladas entre os grupos controle e intervenção e especialmente, confirmamos que a associação entre a intervenção realizada neste estudo e desfecho cognitivo foi mediada pela metilação diferencial do DNA. Ademais, buscamos criar um modelo de cultura de placenta reconhecendo que as respostas a situações adversas envolvem fatores epigenéticos e o desenvolvimento das células trofoblásticas. Confirmamos que foi possível criar um modelo de cultura (TS) semelhante ao trofoblasto.

7 REFERÊNCIAS

1. Barker DJ, Martyn CN. The maternal and fetal origins of cardiovascular disease. *J Epidemiol Community Health* [Internet]. 1992 Feb [cited 2017 Jul 13];46(1):8–11. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/1573367>
2. Hales CN, Barker DJ, Clark PM, Cox LJ, Fall C, Osmond C, et al. Fetal and infant growth and impaired glucose tolerance at age 64. *BMJ* [Internet]. 1991 Oct 26 [cited 2017 Jul 13];303(6809):1019–22. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/1954451>
3. Viltart O, Vanbesien-Mailliot CCA. Impact of Prenatal Stress on Neuroendocrine Programming. *Rev Artic TheScientificWorldJOURNAL* [Internet]. 2007 [cited 2020 Sep 23];7:1493–537. Available from: www.thescientificworld.com
4. Goldstein JA, Norris SA, Aronoff DM. DOHaD at the intersection of maternal immune activation and maternal metabolic stress: a scoping review. *J Dev Orig Health Dis* [Internet]. 2017 Jun 15 [cited 2017 Aug 16];8(03):273–83. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28196555>
5. Bale TL, Baram TZ, Brown AS, Goldstein JM, Insel TR, McCarthy MM, et al. Early life programming and neurodevelopmental disorders. *Biol Psychiatry* [Internet]. 2010 Aug 15 [cited 2018 Apr 23];68(4):314–9. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0006322310005275>
6. Barker D, Osmond C. Infant Mortality, Childhood Nutrition, and Ischaemic Heart Disease in England and Wales. *Lancet*. 1986;327(8489):1077–81.
7. Kim DR, Bale TL, Epperson CN. Prenatal Programming of Mental Illness: Current Understanding of Relationship and Mechanisms. *Curr Psychiatry Rep* [Internet]. 2015 Feb 24 [cited 2018 Apr 23];17(2):5. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25617041>
8. Gabory A, Roseboom TJ, Moore T, Moore LG, Junien C. Placental contribution to the origins of sexual dimorphism in health and diseases: sex chromosomes and epigenetics. *Biol Sex Differ* [Internet]. 2013 Mar 21 [cited 2017 Jul 13];4(1):5. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23514128>
9. van Dongen J, Nivard MG, Willemsen G, Hottenga J-J, Helmer Q, Dolan C V., et al.

- Genetic and environmental influences interact with age and sex in shaping the human methylome. *Nat Commun* [Internet]. 2016 Apr 7 [cited 2017 Jan 23];7:11115. Available from: <http://www.nature.com/doi/10.1038/ncomms11115>
10. O'Donnell KJ, Meaney MJ. Fetal Origins of Mental Health: The Developmental Origins of Health and Disease Hypothesis. *Am J Psychiatry*. 2016;appi.ajp.2016.1.
 11. Skoyles J. Human metabolic adaptations and prolonged expensive neurodevelopment: A review. *Nat Preced* [Internet]. 2008 May 5 [cited 2020 Sep 23];1–1. Available from: <https://www.nature.com/articles/npre.2008.1856.1>
 12. Felsenfeld G. A brief history of epigenetics. *Cold Spring Harb Perspect Biol* [Internet]. 2014 Jan [cited 2020 Sep 25];6(1). Available from: </pmc/articles/PMC3941222/?report=abstract>
 13. Cavalli G, Heard E. Advances in epigenetics link genetics to the environment and disease [Internet]. Vol. 571, *Nature*. Nature Publishing Group; 2019 [cited 2020 Sep 24]. p. 489–99. Available from: <https://doi.org/10.1038/s41586-019-1411-0>
 14. Sandin S, Lichtenstein P, Kuja-Halkola R, Larsson H, Hultman CM, Reichenberg A. The familial risk of autism. *JAMA - J Am Med Assoc* [Internet]. 2014 May 7 [cited 2020 Sep 24];311(17):1770–7. Available from: <https://jamanetwork.com/>
 15. Casey KF, Levesque ML, Szyf M, Ismaylova E, Verner MP, Suderman M, et al. Birth weight discordance, DNA methylation, and cortical morphology of adolescent monozygotic twins. *Hum Brain Mapp* [Internet]. 2017 Apr 1 [cited 2020 Sep 24];38(4):2037–50. Available from: <https://pubmed.ncbi.nlm.nih.gov/28032437/>
 16. Messerschmidt DM, Knowles BB, Solter D. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos [Internet]. Vol. 28, *Genes and Development*. Cold Spring Harbor Laboratory Press; 2014 [cited 2020 Sep 24]. p. 812–28. Available from: </pmc/articles/PMC4003274/?report=abstract>
 17. McRae AF, Powell JE, Henders AK, Bowdler L, Hemani G, Shah S, et al. Contribution of genetic variation to transgenerational inheritance of DNA methylation. *Genome Biol* [Internet]. 2014 May 29 [cited 2020 Sep 24];15(5):R73. Available from: <http://genomebiology.biomedcentral.com/articles/10.1186/gb-2014-15-5-r73>
 18. Manikkam M, Guerrero-Bosagna C, Tracey R, Haque MM, Skinner MK. Transgenerational actions of environmental compounds on reproductive disease and identification of epigenetic biomarkers of ancestral exposures. *PLoS One* [Internet]. 2012 [cited 2017 Jul 13];7(2):e31901. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22389676>

19. Bouwland-Both MI, van Mil NH, Tolhoek CP, Stolk L, Eilers PHC, Verbiest MMPJ, et al. Prenatal parental tobacco smoking, gene specific DNA methylation, and newborns size: the Generation R study. *Clin Epigenetics* [Internet]. 2015 Dec 11 [cited 2016 Jul 23];7(1):83. Available from: <http://www.clinicalepigeneticsjournal.com/content/7/1/83>
20. al-Haddad BJS, Oler E, Armistead B, Elsayed NA, Weinberger DR, Bernier R, et al. The fetal origins of mental illness [Internet]. Vol. 221, *American Journal of Obstetrics and Gynecology*. Mosby Inc.; 2019 [cited 2020 Sep 23]. p. 549–62. Available from: <https://pubmed.ncbi.nlm.nih.gov/31207234/>
21. Choi S-W, Friso S. Epigenetics: A New Bridge between Nutrition and Health. *Adv Nutr An Int Rev J*. 2010;1(1):8–16.
22. Dall’Aglio L, Muka T, Cecil CAM, Bramer WM, Verbiest MMPJ, Nano J, et al. The role of epigenetic modifications in neurodevelopmental disorders: A systematic review [Internet]. Vol. 94, *Neuroscience and Biobehavioral Reviews*. Elsevier Ltd; 2018 [cited 2020 Jul 9]. p. 17–30. Available from: <https://pubmed.ncbi.nlm.nih.gov/30067938/>
23. CpG Island - an overview | ScienceDirect Topics [Internet]. [cited 2020 Sep 24]. Available from: <https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/cpg-island>
24. Cariaga-Martínez A, Gutiérrez K, Alelú-Paz R. The Vast Complexity of the Epigenetic Landscape during Neurodevelopment: An Open Frame to Understanding Brain Function. *Int J Mol Sci* [Internet]. 2018 May 1 [cited 2018 Jul 10];19(5):1333. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29723958>
25. Inoue H, Ochiai M, Sakai Y, Yasuoka K, Tanaka K, Ichiyama M, et al. Neurodevelopmental outcomes in infants with birth weight ≤ 500 g at 3 years of age. *Pediatrics* [Internet]. 2018 Dec 1 [cited 2020 Sep 24];142(6). Available from: <https://pubmed.ncbi.nlm.nih.gov/30446630/>
26. Rondó PHC, Ferreira RF, Nogueira F, Ribeiro MCN, Lobert H, Artes R. Maternal psychological stress and distress as predictors of low birth weight, prematurity and intrauterine growth retardation. *Eur J Clin Nutr* [Internet]. 2003 Feb 1 [cited 2020 Sep 24];57(2):266–72. Available from: www.nature.com/ejcn
27. Loomans EM, Van Dijk AE, Vrijkotte TGM, Van Eijsden M, Stronks K, Gemke RBBJ, et al. Psychosocial stress during pregnancy is related to adverse birth outcomes: Results from a large multi-ethnic community-based birth cohort. *Eur J Public Health* [Internet]. 2013 Jun 1 [cited 2020 Sep 24];23(3):485–91. Available from: <http://www.ncbi.nlm.nih.gov>

28. Walsh K, McCormack CA, Webster R, Pinto A, Lee S, Feng T, et al. Maternal prenatal stress phenotypes associate with fetal neurodevelopment and birth outcomes. *Proc Natl Acad Sci U S A* [Internet]. 2019 Nov 26 [cited 2020 Sep 24];116(48):23996–4005. Available from: <https://pubmed.ncbi.nlm.nih.gov/31611411/>
29. Kim D, Saada A. The social determinants of infant mortality and birth outcomes in Western developed nations: a cross-country systematic review. *Int J Environ Res Public Health* [Internet]. 2013 Jun 5 [cited 2017 Jul 13];10(6):2296–335. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23739649>
30. Weinstock M. The potential influence of maternal stress hormones on development and mental health of the offspring. *Brain Behav Immun*. 2005;19(4):296–308.
31. Monk C, Spicer J, Champagne FA. Linking prenatal maternal adversity to developmental outcomes in infants: the role of epigenetic pathways. *Dev Psychopathol* [Internet]. 2012 Nov [cited 2016 Jul 31];24(4):1361–76. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23062303>
32. Shephard E, Fatori D, Mauro LR, de Medeiros Filho M V., Hoexter MQ, Chiesa AM, et al. Effects of Maternal Psychopathology and Education Level on Neurocognitive Development in Infants of Adolescent Mothers Living in Poverty in Brazil. *Biol Psychiatry Cogn Neurosci Neuroimaging* [Internet]. 2019 Oct 1 [cited 2020 Jul 9];4(10):925–34. Available from: <https://pubmed.ncbi.nlm.nih.gov/31345780/>
33. Blair C, Raver CC. Poverty, Stress, and Brain Development: New Directions for Prevention and Intervention [Internet]. Vol. 16, *Academic Pediatrics*. Elsevier Inc.; 2016 [cited 2020 Sep 24]. p. S30–6. Available from: </pmc/articles/PMC5765853/?report=abstract>
34. Cook N, Ayers S, Horsch A. Maternal posttraumatic stress disorder during the perinatal period and child outcomes: A systematic review [Internet]. Vol. 225, *Journal of Affective Disorders*. Elsevier B.V.; 2018 [cited 2020 Sep 24]. p. 18–31. Available from: <https://pubmed.ncbi.nlm.nih.gov/28777972/>
35. Lima SAM, El Dib RP, Rodrigues MRK, Ferraz GAR, Molina AC, Neto CAP, et al. Is the risk of low birth weight or preterm labor greater when maternal stress is experienced during pregnancy? A systematic review and meta-analysis of cohort studies. *PLoS One* [Internet]. 2018 Jul 1 [cited 2020 Sep 24];13(7). Available from: </pmc/articles/PMC6061976/?report=abstract>
36. Lynch JL, Gibbs BG. Birth Weight and Early Cognitive Skills: Can Parenting Offset the Link? *Matern Child Health J* [Internet]. 2017 Jan 1 [cited 2020 Sep 22];21(1):156–

67. Available from: <https://pubmed.ncbi.nlm.nih.gov/27469110/>
37. Fink G, Andrews KG, Brentani H, Grisi S, Ferrer APS, Brentani A. Overall and Sex-Specific Associations between Fetal Adversity and Child Development at Age 1 Year: Evidence from Brazil. *Am J Epidemiol* [Internet]. 2018 Nov 1 [cited 2020 Sep 24];187(11):2324–31. Available from: <https://pubmed.ncbi.nlm.nih.gov/29982368/>
38. Lyall K, Schmidt RJ, Hertz-Picciotto I. Maternal lifestyle and environmental risk factors for autism spectrum disorders. *Int J Epidemiol* [Internet]. 2014 Apr [cited 2018 May 16];43(2):443–64. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24518932>
39. Keil KP, Lein PJ. DNA methylation: a mechanism linking environmental chemical exposures to risk of autism spectrum disorders? *Environ Epigenetics* [Internet]. 2016 Jan 30 [cited 2018 May 16];2(1):dvv012. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27158529>
40. Braun JM. Early-life exposure to EDCs: Role in childhood obesity and neurodevelopment [Internet]. Vol. 13, *Nature Reviews Endocrinology*. Nature Publishing Group; 2017 [cited 2020 Sep 24]. p. 161–73. Available from: <https://pubmed.ncbi.nlm.nih.gov/27857130/>
41. Weinstock M. The long-term behavioural consequences of prenatal stress [Internet]. Vol. 32, *Neuroscience and Biobehavioral Reviews*. *Neurosci Biobehav Rev*; 2008 [cited 2020 Sep 24]. p. 1073–86. Available from: <https://pubmed.ncbi.nlm.nih.gov/18423592/>
42. Sandman CA, Buss C, Head K, Davis EP. Fetal exposure to maternal depressive symptoms is associated with cortical thickness in late childhood. *Biol Psychiatry* [Internet]. 2015 [cited 2020 Sep 24];77(4):324–34. Available from: <https://pubmed.ncbi.nlm.nih.gov/25129235/>
43. Davis EP, Sandman CA. The timing of prenatal exposure to maternal cortisol and psychosocial stress is associated with human infant cognitive development. *Child Dev* [Internet]. 2010 Jan [cited 2020 Sep 24];81(1):131–48. Available from: <https://pubmed.ncbi.nlm.nih.gov/20331658/>
44. Ortuño-Sahagún D, Schliebs R, Pallàs M. Editorial: Epigenetic Mechanisms Regulating Neural Plasticity. *Front Cell Neurosci* [Internet]. 2019 Apr 24 [cited 2020 Sep 24];13:118. Available from: <https://www.frontiersin.org/article/10.3389/fncel.2019.00118/full>
45. Glendining KA, Fisher LC, Jasoni CL. Maternal high fat diet alters offspring

- epigenetic regulators, amygdala glutamatergic profile and anxiety. *Psychoneuroendocrinology* [Internet]. 2018 Oct 1 [cited 2020 Sep 24];96:132–41. Available from: <https://pubmed.ncbi.nlm.nih.gov/29940426/>
46. Dhobale M V., Pisal HR, Mehendale SS, Joshi SR. Differential expression of human placental neurotrophic factors in preterm and term deliveries. *Int J Dev Neurosci* [Internet]. 2013 Dec [cited 2020 Sep 24];31(8):719–23. Available from: <https://pubmed.ncbi.nlm.nih.gov/24076518/>
 47. Clukay CJ, Hughes DA, Rodney NC, Kertes DA, Mulligan CJ. DNA methylation of methylation complex genes in relation to stress and genome-wide methylation in mother–newborn dyads. *Am J Phys Anthropol* [Internet]. 2018 Jan 1 [cited 2020 Sep 24];165(1):173–82. Available from: <https://pubmed.ncbi.nlm.nih.gov/29028111/>
 48. Vangipuram SD, Lyman WD. Ethanol Affects Differentiation-Related Pathways and Suppresses Wnt Signaling Protein Expression in Human Neural Stem Cells. *Alcohol Clin Exp Res* [Internet]. 2012 May [cited 2020 Sep 24];36(5):788–97. Available from: <https://pubmed.ncbi.nlm.nih.gov/22150777/>
 49. Lesiak A, Zhu M, Chen H, Appleyard SM, Impey S, Lein PJ, et al. The environmental neurotoxicant PCB 95 promotes synaptogenesis via ryanodine receptor-dependent miR132 upregulation. *J Neurosci* [Internet]. 2014 Jan 15 [cited 2020 Sep 24];34(3):717–25. Available from: <https://www.jneurosci.org/content/34/3/717>
 50. Rijlaarsdam J, Pappa I, Walton E, Bakermans-Kranenburg MJ, Mileva-Seitz VR, Rippe RCA, et al. An epigenome-wide association meta-analysis of prenatal maternal stress in neonates: A model approach for replication. *Epigenetics* [Internet]. 2016 Feb 1 [cited 2020 Sep 24];11(2):140–9. Available from: <http://dx.doi.org/10.1080/15592294.2016.1145329>
 51. Sosnowski DW, Booth C, York TP, Amstadter AB, Kliwer W. Maternal prenatal stress and infant DNA methylation: A systematic review [Internet]. Vol. 60, *Developmental Psychobiology*. John Wiley and Sons Inc.; 2018 [cited 2020 Sep 24]. p. 127–39. Available from: <https://pubmed.ncbi.nlm.nih.gov/29344930/>
 52. Cardenas A, Faleschini S, Cortes Hidalgo A, Rifas-Shiman SL, Baccarelli AA, Demeo DL, et al. Prenatal maternal antidepressants, anxiety, and depression and offspring DNA methylation: Epigenome-wide associations at birth and persistence into early childhood. *Clin Epigenetics* [Internet]. 2019 Mar 29 [cited 2020 Sep 24];11(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/30925934/>
 53. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*

- [Internet]. 2013 [cited 2018 Sep 17];14(10):R115. Available from:
<http://www.ncbi.nlm.nih.gov/pubmed/24138928>
54. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing [Internet]. Vol. 19, *Nature Reviews Genetics*. Nature Publishing Group; 2018 [cited 2020 Sep 24]. p. 371–84. Available from:
<https://pubmed.ncbi.nlm.nih.gov/29643443/>
 55. Knight AK, Craig JM, Theda C, Bækvad-Hansen M, Bybjerg-Grauholm J, Hansen CS, et al. An epigenetic clock for gestational age at birth based on blood methylation data. *Genome Biol*. 2016 Oct 7;17(1).
 56. Bohlin J, Håberg SE, Magnus P, Reese SE, Gjessing HK, Magnus MC, et al. Prediction of gestational age based on genome-wide differentially methylated regions. *Genome Biol* [Internet]. 2016 Oct 7 [cited 2020 Sep 24];17(1). Available from:
<https://pubmed.ncbi.nlm.nih.gov/27717397/>
 57. Chen BH, Marioni RE, Colicino E, Peters MJ, Ward-Caviness CK, Tsai PC, et al. DNA methylation-based measures of biological age: Meta-analysis predicting time to death. *Aging (Albany NY)* [Internet]. 2016 [cited 2020 Sep 23];8(9):1844–65. Available from: [/pmc/articles/PMC5076441/?report=abstract](https://pubmed.ncbi.nlm.nih.gov/27717397/)
 58. Girchenko P, Lahti J, Czamara D, Knight AK, Jones MJ, Suarez A, et al. Associations between maternal risk factors of adverse pregnancy and birth outcomes and the offspring epigenetic clock of gestational age at birth. *Clin Epigenetics* [Internet]. 2017 May 8 [cited 2020 Sep 22];9(1). Available from:
<https://pubmed.ncbi.nlm.nih.gov/28503212/>
 59. Suarez A, Lahti J, Czamara D, Lahti-Pulkkinen M, Knight AK, Girchenko P, et al. The Epigenetic Clock at Birth: Associations With Maternal Antenatal Depression and Child Psychiatric Problems. *J Am Acad Child Adolesc Psychiatry* [Internet]. 2018 May 1 [cited 2020 Sep 23];57(5):321-328.e2. Available from:
<https://pubmed.ncbi.nlm.nih.gov/29706161/>
 60. Tekola-Ayele F, Workalemahu T, Gorfu G, Shrestha D, Tycko B, Wapner R, et al. Sex differences in the associations of placental epigenetic aging with fetal growth. *Aging (Albany NY)* [Internet]. 2019 Aug 15 [cited 2020 Sep 23];11(15):5412–32. Available from: <https://pubmed.ncbi.nlm.nih.gov/31395791/>
 61. Chen L, Wagner CL, Dong Y, Wang X, Shary JR, Huang Y, et al. Effects of Maternal Vitamin D3 Supplementation on Offspring Epigenetic Clock of Gestational Age at Birth: A Post-hoc Analysis of a Randomized Controlled Trial. *Epigenetics* [Internet].

- 2020 Aug 2 [cited 2020 Sep 24];15(8):830–40. Available from:
<https://pubmed.ncbi.nlm.nih.gov/32089064/>
62. Tobi EW, van den Heuvel J, Zwaan BJ, Lumey LH, Heijmans BT, Uller T. Selective Survival of Embryos Can Explain DNA Methylation Signatures of Adverse Prenatal Environments. *Cell Rep* [Internet]. 2018 Dec 4 [cited 2020 Sep 24];25(10):2660-2667.e4. Available from: <https://pubmed.ncbi.nlm.nih.gov/30517855/>
 63. Mikkola K, Robinson AM, Zucchi FC, Robbins JC, Babenko O, Kovalchuk O, et al. Neurodevelopmental Outcome at 5 Years of Age of a National Cohort of Extremely Low Birth Weight Infants Who Were Born in 1996-1997. *Pediatrics* [Internet]. 2005 Dec 1 [cited 2017 Jan 6];116(6):1391–400. Available from:
<http://pediatrics.aappublications.org/cgi/doi/10.1542/peds.2005-0171>
 64. Gamage TKJB, Chamley LW, James JL. Stem cell insights into human trophoblast lineage differentiation. *Hum Reprod Update* [Internet]. 2016 [cited 2020 Sep 23];23(1):77–103. Available from: <https://pubmed.ncbi.nlm.nih.gov/27591247/>
 65. Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, et al. Single-cell reconstruction of the early maternal–fetal interface in humans. *Nature* [Internet]. 2018 Nov 15 [cited 2020 Sep 23];563(7731):347–53. Available from:
<https://pubmed.ncbi.nlm.nih.gov/30429548/>
 66. Gabory A, Roseboom TJ, Moore T, Moore LG, Junien C. Placental contribution to the origins of sexual dimorphism in health and diseases: sex chromosomes and epigenetics. *Biol Sex Differ* [Internet]. 2013 Mar 21 [cited 2017 Mar 13];4(1):5. Available from:
<http://www.ncbi.nlm.nih.gov/pubmed/23514128>
 67. Myatt L. Placental adaptive responses and fetal programming. In: *Journal of Physiology* [Internet]. *J Physiol*; 2006 [cited 2020 Sep 23]. p. 25–30. Available from:
<https://pubmed.ncbi.nlm.nih.gov/16469781/>
 68. Bale TL. Sex differences in prenatal epigenetic programming of stress pathways. *Stress* [Internet]. 2011 Jul [cited 2016 Jul 31];14(4):348–56. Available from:
<http://www.ncbi.nlm.nih.gov/pubmed/21663536>
 69. Strumpf D, Mao C-A, Yamanaka Y, Ralston A, Chawengsaksophak K, Beck F, et al. Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development* [Internet]. 2005 May 1 [cited 2019 Apr 26];132(9):2093–102. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15788452>
 70. Pick M, Azzola L, Mossman A, Stanley EG, Elefanty AG. Differentiation of Human Embryonic Stem Cells in Serum-Free Medium Reveals Distinct Roles for Bone

- Morphogenetic Protein 4, Vascular Endothelial Growth Factor, Stem Cell Factor, and Fibroblast Growth Factor 2 in Hematopoiesis. *Stem Cells* [Internet]. 2007 Sep [cited 2020 Sep 23];25(9):2206–14. Available from: <https://pubmed.ncbi.nlm.nih.gov/17556598/>
71. Chen B, Dodge ME, Tang W, Lu J, Ma Z, Fan CW, et al. Small molecule-mediated disruption of Wnt-dependent signaling in tissue regeneration and cancer. *Nat Chem Biol* [Internet]. 2009 Feb 6 [cited 2020 Sep 23];5(2):100–7. Available from: <https://pubmed.ncbi.nlm.nih.gov/19125156/>
 72. Tobi EW, Slieker RC, Luijk R, Dekkers KF, Stein AD, Xu KM, et al. DNA methylation as a mediator of the association between prenatal adversity and risk factors for metabolic disease in adulthood. *Sci Adv* [Internet]. 2018 Jan 1 [cited 2020 Jul 9];4(1):eaao4364. Available from: <http://advances.sciencemag.org/>
 73. Phelan S, Clifton RG, Haire-Joshu D, Redman LM, Van Horn L, Evans M, et al. One-year postpartum anthropometric outcomes in mothers and children in the LIFE-Moms lifestyle intervention clinical trials. *Int J Obes* [Internet]. 2020 Jan 1 [cited 2020 Sep 24];44(1):57–68. Available from: <https://www.nature.com/articles/s41366-019-0410-4>
 74. Dawley K, Loch J, Bindrich I. The nurse-family partnership [Internet]. Vol. 107, *American Journal of Nursing*. Am J Nurs; 2007 [cited 2020 Sep 24]. p. 60–7. Available from: <https://pubmed.ncbi.nlm.nih.gov/18075344/>
 75. Thorland W, Currie DW. Status of Birth Outcomes in Clients of the Nurse-Family Partnership. *Matern Child Health J* [Internet]. 2017 May 1 [cited 2020 Sep 24];21(5):995–1001. Available from: <https://pubmed.ncbi.nlm.nih.gov/28105544/>
 76. Thorland W, Currie D, Wiegand ER, Walsh J, Mader N. Status of Breastfeeding and Child Immunization Outcomes in Clients of the Nurse–Family Partnership. *Matern Child Health J* [Internet]. 2017 Mar 1 [cited 2020 Sep 24];21(3):439–45. Available from: <https://pubmed.ncbi.nlm.nih.gov/28092061/>
 77. Holland ML, Groth SW, Smith JA, Meng Y, Kitzman H. Low birthweight in second children after nurse home visiting. *J Perinatol* [Internet]. 2018 Dec 1 [cited 2020 Sep 24];38(12):1610–9. Available from: <https://www.nature.com/articles/s41372-018-0222-8>
 78. Slade A, Holland ML, Ordway MR, Carlson EA, Jeon S, Close N, et al. Minding the Baby®: Enhancing parental reflective functioning and infant attachment in an attachment-based, interdisciplinary home visiting program. *Dev Psychopathol* [Internet]. 2019 [cited 2020 Jul 9];32(1):123–37. Available from:

- <https://pubmed.ncbi.nlm.nih.gov/30636649/>
79. Ordway MR, Sadler LS, Holland ML, Slade A, Close N, Mayes LC. A home visiting parenting program and child obesity: A randomized trial. *Pediatrics* [Internet]. 2018 Feb 1 [cited 2020 Sep 24];141(2). Available from: <https://pubmed.ncbi.nlm.nih.gov/29339565/>
 80. Nyklíček I, Truijens SEM, Spek V, Pop VJM. Mindfulness skills during pregnancy: Prospective associations with mother's mood and neonatal birth weight. *J Psychosom Res* [Internet]. 2018 Apr 1 [cited 2020 Sep 24];107:14–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/29502758/>
 81. Richmond RC, Sharp GC, Herbert G, Atkinson C, Taylor C, Bhattacharya S, et al. The long-term impact of folic acid in pregnancy on offspring DNA methylation: Follow-up of the Aberdeen folic acid supplementation trial (FAST). *Int J Epidemiol* [Internet]. 2018 Jun 1 [cited 2020 Sep 24];47(3):928–37. Available from: <https://academic.oup.com/ije/article/47/3/928/4931210>
 82. Andraos S, de Seymour JV, O'Sullivan JM, Kussmann M. The Impact of Nutritional Interventions in Pregnant Women on DNA Methylation Patterns of the Offspring: A Systematic Review. *Mol Nutr Food Res* [Internet]. 2018 Dec 1 [cited 2020 Sep 24];62(24):1800034. Available from: <http://doi.wiley.com/10.1002/mnfr.201800034>
 83. Vähämäki S, Laiho A, Lund R, Isolauri E, Salminen S, Laitinen K. The impact of probiotic supplementation during pregnancy on DNA methylation of obesity-related genes in mothers and their children. *Eur J Nutr* [Internet]. 2019 Feb 1 [cited 2020 Sep 24];58(1):367–77. Available from: <https://link.springer.com/article/10.1007/s00394-017-1601-1>
 84. WHO Child Growth Standards based on length/height, weight and age. *Acta Paediatr (Oslo, Norw 1992) Suppl* [Internet]. 2006 Apr [cited 2020 Jul 9];450:76–85. Available from: <https://pubmed.ncbi.nlm.nih.gov/16817681/>
 85. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* [Internet]. 2014 May 15 [cited 2016 Jul 23];30(10):1363–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24478339>
 86. Triche TJ, Weisenberger DJ, Van Den Berg D, Laird PW, Siegmund KD. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Res* [Internet]. 2013 Apr [cited 2016 Jul 23];41(7):e90. Available from:

- <http://www.ncbi.nlm.nih.gov/pubmed/23476028>
87. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile Within Array Normalization for Illumina Infinium HumanMethylation450 BeadChips. *Genome Biol* [Internet]. 2012 Jun 15 [cited 2017 Jan 23];13(6):R44. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22703947>
 88. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* [Internet]. 2010 Nov 30 [cited 2020 Sep 24];11(1):587. Available from: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-11-587>
 89. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* [Internet]. 2015 Apr 20 [cited 2016 Jul 23];43(7):e47. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25605792>
 90. Jiao Y, Teschendorff AE. The FEM R package: Identification of Functional Epigenetic Modules [Internet]. 2016 [cited 2020 Sep 20]. Available from: <http://cran.r-project.org/>.
 91. Langfelder P, Horvath S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* [Internet]. 2008 Dec 29 [cited 2020 Jul 9];9. Available from: <https://pubmed.ncbi.nlm.nih.gov/19114008/>
 92. Sant'Anna Feltrin A, Tahira AC, Simões SN, Brentani H, Martins DC. Assessment of complementarity of WGCNA and NERI results for identification of modules associated to schizophrenia spectrum disorders. *PLoS One* [Internet]. 2019 Jan 1 [cited 2020 Sep 20];14(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/30645614/>
 93. Momany AM, Kamradt JM, Nikolas MA. A Meta-Analysis of the Association Between Birth Weight and Attention Deficit Hyperactivity Disorder [Internet]. Vol. 46, *Journal of Abnormal Child Psychology*. Springer New York LLC; 2018 [cited 2020 Sep 24]. p. 1409–26. Available from: <https://pubmed.ncbi.nlm.nih.gov/29159441/>
 94. Ross KM, Carroll JE, Horvath S, Hobel CJ, Coussons-Read ME, Dunkel Schetter C. Epigenetic age and pregnancy outcomes: GrimAge acceleration is associated with shorter gestational length and lower birthweight. *Clin Epigenetics* [Internet]. 2020 Aug 6 [cited 2020 Sep 23];12(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/32762768/>
 95. Girchenko P, Lahti J, Czamara D, Knight AK, Jones MJ, Suarez A, et al. Associations

- between maternal risk factors of adverse pregnancy and birth outcomes and the offspring epigenetic clock of gestational age at birth. *Clin Epigenetics* [Internet]. 2017 Dec 8 [cited 2020 Jul 9];9(1):49. Available from: <http://clinicalepigeneticsjournal.biomedcentral.com/articles/10.1186/s13148-017-0349-z>
96. Gareau JR, Lima CD. The SUMO pathway: Emerging mechanisms that shape specificity, conjugation and recognition [Internet]. Vol. 11, *Nature Reviews Molecular Cell Biology*. NIH Public Access; 2010 [cited 2020 Sep 22]. p. 861–71. Available from: </pmc/articles/PMC3079294/?report=abstract>
 97. Kaul D, Habel P, Derkow K, Krüger C, Franzoni E, Wulczyn FG, et al. Expression of toll-like receptors in the developing brain. *PLoS One* [Internet]. 2012 May 30 [cited 2020 Sep 22];7(5):37767. Available from: </pmc/articles/PMC3364272/?report=abstract>
 98. Tsai PC, Bell JT. Power and sample size estimation for epigenome-wide association scans to detect differential DNA methylation. *Int J Epidemiol* [Internet]. 2015 Aug 1 [cited 2020 Sep 25];44(4):1429–41. Available from: </pmc/articles/PMC4588864/?report=abstract>
 99. Graw S, Henn R, Thompson JA, Koestler DC. pwrEWAS: a user-friendly tool for comprehensive power estimation for epigenome wide association studies (EWAS). *BMC Bioinformatics* [Internet]. 2019 Dec 29 [cited 2020 Sep 25];20(1):218. Available from: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-019-2804-7>
 100. Mansell G, Gorrie-Stone TJ, Bao Y, Kumari M, Schalkwyk LS, Mill J, et al. Guidance for DNA methylation studies: Statistical insights from the Illumina EPIC array. *BMC Genomics* [Internet]. 2019 May 14 [cited 2020 Sep 25];20(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/31088362/>
 101. Bowlby J. The making and breaking of affectional bonds. I. Aetiology and psychopathology in the light of attachment theory. An expanded version of the Fiftieth Maudsley Lecture, delivered before the Royal College of Psychiatrists, 19 November 1976. *Br J Psychiatry* [Internet]. 1977 [cited 2020 Aug 11];130(3):201–10. Available from: </core/journals/the-british-journal-of-psychiatry/article/making-and-breaking-of-affectional-bonds/12BE02CC4F59067D79FC64534E36FC5E>
 102. Bandura A. Self-efficacy: Toward a unifying theory of behavioral change. *Psychol Rev* [Internet]. 1977 Mar [cited 2020 Aug 11];84(2):191–215. Available from: </record/1977-25733-001>

103. Bronfenbrenner U, Morris PA. The Bioecological Model of Human Development. In: Handbook of Child Psychology [Internet]. John Wiley & Sons, Inc.; 2007 [cited 2020 Aug 11]. Available from:
<https://onlinelibrary.wiley.com/doi/full/10.1002/9780470147658.chpsy0114>
104. Fracolli LA, Reticena K de O, de Abreu FCP, Chiesa AM. The implementation of a home visits program focused on parenting: An experience report. Rev da Esc Enferm [Internet]. 2018 [cited 2020 Jul 9];52. Available from:
<https://pubmed.ncbi.nlm.nih.gov/30156655/>
105. Olds DL. The nurse-family partnership: An evidence-based preventive intervention. Infant Ment Health J [Internet]. 2006 Jan 1 [cited 2020 Aug 11];27(1):5–25. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1002/imhj.20077>
106. Slade A, Sadler L, De Dios-Kenn C, Webb D, Currier-Ezepchick J, Mayes L. Minding the baby: A reflective parenting program [Internet]. Vol. 60, Psychoanalytic Study of the Child. Yale University Press; 2005 [cited 2020 Aug 11]. p. 74–100. Available from: <https://pubmed.ncbi.nlm.nih.gov/16649676/>
107. Critério Brasil - ABEP [Internet]. [cited 2020 Jul 9]. Available from:
<http://www.abep.org/criterio-brasil>
108. Escala Brasileira de Insegurança Alimentar – EBIA: análise psicométrica de uma dimensão da Segurança Alimentar e Nutricional - Acervo Social [Internet]. [cited 2020 Sep 24]. Available from: <https://fpabramo.org.br/acervosocial/estante/escala-brasileira-de-inseguranca-alimentar-ebia-analise-psicometrica-de-uma-dimensao-da-seguranca-alimentar-e-nutricional/>
109. Gomes-Oliveira MH, Gorenstein C, Neto FL, Andrade LH, Wang YP. Validação da versão Brasileira em Português do Inventário de Depressão de Beck-II numa amostra da comunidade. Rev Bras Psiquiatr [Internet]. 2012 [cited 2020 Jul 9];34(4):389–94. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-44462012000400005&lng=en&nrm=iso&tlng=en
110. Henrique Gomes-Oliveira M, Gorenstein C, Lotufo Neto F, Helena Andrade L, Pang Wang Y, Gomes-Oliveira M. Revista Brasileira de Psiquiatria Validation of the Brazilian Portuguese version of the Beck Depression Inventory-II in a community sample Validação da versão brasileira em português do Inventário de Depressão de Beck-II numa amostra da comunidade. Off J Brazilian Psychiatr Assoc. 2012;34(4):389–94.
111. Grassi-Oliveira R, Stein LM, Pezzi JC. Tradução e validação de conteúdo da versão em

- português do Childhood Trauma Questionnaire. *Rev Saude Publica* [Internet]. 2006 Apr [cited 2020 Sep 24];40(2):249–55. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-89102006000200010&lng=en&nrm=iso&tlng=pt
112. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays [Internet]. [cited 2020 Jul 9]. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4016708/>
 113. Fortin J-P, Triche TJ, Hansen KD, Hancock J. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. [cited 2020 Jul 9]; Available from: <https://academic.oup.com/bioinformatics/article-abstract/33/4/558/2666344>
 114. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, et al. ChAMP: Updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics*. 2017;33(24):3982–4.
 115. Bayley N. BAYLEY-III (BAYLEY SCALES OF INFANT AND TODDLER DEVELOPMENT). 3rd ed. San Antonio, TX: PsychCorp; 2006.
 116. Madaschi V, Mecca TP. Bayley-III Scales of Infant and Toddler Development: Transcultural Adaptation and Psychometric Properties 1 Article. [cited 2020 Jul 9];26(64):189–97. Available from: www.scielo.br/paideia
 117. Yu G, Wang LG, Han Y, He QY. ClusterProfiler: An R package for comparing biological themes among gene clusters. *Omi A J Integr Biol* [Internet]. 2012 May 1 [cited 2020 Jul 13];16(5):284–7. Available from: </pmc/articles/PMC3339379/?report=abstract>
 118. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. Mediation: R package for causal mediation analysis. *J Stat Softw* [Internet]. 2014 Aug 1 [cited 2020 Jul 9];59(5):1–38. Available from: <https://waseda.pure.elsevier.com/en/publications/mediation-r-package-for-causal-mediation-analysis>
 119. Placenta Development - an overview | ScienceDirect Topics [Internet]. [cited 2020 Sep 23]. Available from: <https://www.sciencedirect.com/topics/medicine-and-dentistry/placenta-development>
 120. Gabor N, Dunk C, James J, Knöfler M, Pollheimer J, Vondra S, et al. Regulation of Placental Extravillous Trophoblasts by the Maternal Uterine Environment. *Front Immunol* | www.frontiersin.org [Internet]. 2018 [cited 2020 Sep 23];9:2597. Available from: www.frontiersin.org

121. Knöfler M, Haider S, Saleh L, Pollheimer J, Gamage TKJB, James J. Human placenta and trophoblast development: key molecular mechanisms and model systems [Internet]. Vol. 76, Cellular and Molecular Life Sciences. Birkhauser Verlag AG; 2019 [cited 2020 Sep 23]. p. 3479–96. Available from: <https://pubmed.ncbi.nlm.nih.gov/31049600/>
122. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets.
123. Stuart T, Butler A, Hoffman P, Stoeckius M, Smibert P, Satija R, et al. Comprehensive Integration of Single-Cell Data Resource Comprehensive Integration of Single-Cell Data. *Cell*. 2019;177:1888-1902.e21.
124. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* [Internet]. 2019 Dec 23 [cited 2020 Sep 23];20(1):296. Available from: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-019-1874-1>
125. Chen H, Albergante L, Hsu JY, Lareau CA, Lo Bosco G, Guan J, et al. Single-cell trajectories reconstruction, exploration and mapping of omics data with STREAM. *Nat Commun* [Internet]. 2019 Dec 1 [cited 2020 Sep 23];10(1):1–14. Available from: <https://doi.org/10.1038/s41467-019-09670-4>
126. Sonawane AR, Platig J, Fagny M, Chen CY, Paulson JN, Lopes-Ramos CM, et al. Understanding Tissue-Specific Gene Regulation. *Cell Rep* [Internet]. 2017 Oct 24 [cited 2020 Sep 23];21(4):1077–88. Available from: </pmc/articles/PMC5828531/?report=abstract>
127. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res* [Internet]. 2003 Nov [cited 2020 Sep 23];13(11):2498–504. Available from: </pmc/articles/PMC403769/?report=abstract>
128. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* [Internet]. 2009 [cited 2020 Sep 23];25(8):1091–3. Available from: </pmc/articles/PMC2666812/?report=abstract>
129. Thomson JA. Embryonic stem cell lines derived from human blastocysts. *Science* (80-) [Internet]. 1998 Nov 6 [cited 2020 Sep 23];282(5391):1145–7. Available from: <https://pubmed.ncbi.nlm.nih.gov/9804556/>
130. Osafune K, Caron L, Borowiak M, Martinez RJ, Fitz-Gerald CS, Sato Y, et al. Marked

- differences in differentiation propensity among human embryonic stem cell lines. *Nat Biotechnol* [Internet]. 2008 Mar [cited 2020 Sep 23];26(3):313–5. Available from: <https://pubmed.ncbi.nlm.nih.gov/18278034/>
131. Sawada T, Benjamin KJM, Brandtjen AC, Tietze E, Allen SJ, Paquola ACM, et al. Generation of four postmortem dura-derived iPS cell lines from four control individuals with genotypic and brain-region-specific transcriptomic data available through the BrainSEQ consortium. *Stem Cell Res* [Internet]. 2020 Jul 1 [cited 2020 Sep 23];46:101806. Available from: <https://jhu.pure.elsevier.com/en/publications/generation-of-four-postmortem-dura-derived-ips-cell-lines-from-fo>
 132. Okae H, Toh H, Sato T, Hiura H, Takahashi S, Shirane K, et al. Derivation of Human Trophoblast Stem Cells. *Cell Stem Cell* [Internet]. 2018 Jan 4 [cited 2020 Sep 24];22(1):50–63.e6. Available from: <https://pubmed.ncbi.nlm.nih.gov/29249463/>
 133. Cole LA. Biological functions of hCG and hCG-related molecules [Internet]. Vol. 8, *Reproductive Biology and Endocrinology. Reprod Biol Endocrinol*; 2010 [cited 2020 Sep 23]. Available from: <https://pubmed.ncbi.nlm.nih.gov/20735820/>
 134. Wakeland AK, Soncin F, Moretto-Zita M, Chang CW, Horii M, Pizzo D, et al. Hypoxia Directs Human Extravillous Trophoblast Differentiation in a Hypoxia-Inducible Factor–Dependent Manner. *Am J Pathol* [Internet]. 2017 Apr 1 [cited 2020 Sep 23];187(4):767–80. Available from: [https://pubmed.ncbi.nlm.nih.gov/27397702/?report=abstract](https://pubmed.ncbi.nlm.nih.gov/27397702/)
 135. Xu RH, Chen X, Li DS, Li R, Addicks GC, Glennon C, et al. BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nat Biotechnol* [Internet]. 2002 Dec 1 [cited 2020 Sep 23];20(12):1261–4. Available from: <https://www.nature.com/articles/nbt761>
 136. Horii M, Bui T, Touma O, Cho HY, Parast MM. An Improved Two-Step Protocol for Trophoblast Differentiation of Human Pluripotent Stem Cells. *Curr Protoc Stem Cell Biol* [Internet]. 2019 [cited 2020 Sep 23];50(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/31479595/>
 137. Liu Y, Fan X, Wang R, Lu X, Dang YL, Wang H, et al. Single-cell RNA-seq reveals the diversity of trophoblast subtypes and patterns of differentiation in the human placenta. *Cell Res* [Internet]. 2018 Aug 1 [cited 2020 Sep 23];28(8):819–32. Available from: <https://doi.org/10.1038/s41422-018-0066-y>
 138. Yabe S, Alexenko AP, Amita M, Yang Y, Schust DJ, Sadovsky Y, et al. Comparison of syncytiotrophoblast generated from human embryonic stem cells and from term

- placentas. *Proc Natl Acad Sci U S A* [Internet]. 2016 May 10 [cited 2020 Sep 23];113(19):E2598–607. Available from: <https://pubmed.ncbi.nlm.nih.gov/27051068/>
139. Erwin JA, del Rosario B, Payer B, Lee JT. An ex vivo model for imprinting: Mutually exclusive binding of Cdx2 and Oct4 as a switch for imprinted and random X-inactivation. *Genetics* [Internet]. 2012 Nov 1 [cited 2020 Sep 23];192(3):857–68. Available from: [/pmc/articles/PMC3522163/?report=abstract](https://pubmed.ncbi.nlm.nih.gov/22163313/)
 140. Niakan KK, Eggan K. Analysis of human embryos from zygote to blastocyst reveals distinct gene expression patterns relative to the mouse. *Dev Biol* [Internet]. 2013 Mar 1 [cited 2020 Sep 23];375(1):54–64. Available from: <https://pubmed.ncbi.nlm.nih.gov/23261930/>
 141. Kazanskaya O, Glinka A, del Barco Barrantes I, Stannek P, Niehrs C, Wu W. R-Spondin2 is a secreted activator of Wnt/ β -catenin signaling and is required for *Xenopus* myogenesis. *Dev Cell* [Internet]. 2004 Oct [cited 2020 Sep 23];7(4):525–34. Available from: <https://pubmed.ncbi.nlm.nih.gov/15469841/>
 142. Hendee KE, Sorokina EA, Muheisen SS, Reis LM, Tyler RC, Markovic V, et al. PITX2 deficiency and associated human disease: Insights from the zebrafish model. *Hum Mol Genet* [Internet]. 2018 May 15 [cited 2020 Sep 23];27(10):1675–95. Available from: <https://pubmed.ncbi.nlm.nih.gov/29506241/>
 143. Parikshak NN, Luo R, Zhang A, Won H, Lowe JK, Chandran V, et al. Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell* [Internet]. 2013 Nov 21 [cited 2016 Jul 23];155(5):1008–21. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24267887>
 144. Oh Y, Nagalla SR, Yamanaka Y, Kim HS, Wilson E, Rosenfeld RG. Synthesis and characterization of insulin-like growth factor-binding protein (IGFBP)-7: Recombinant human mac25 protein specifically binds IGF-I and -II. *J Biol Chem*. 1996;271(48):30322–5.
 145. Overgaard MT, Bold HB, Laursen LS, Sottrup-Jensen L, Conover CA, Oxvig C. Pregnancy-associated Plasma Protein-A2 (PAPP-A2), a Novel Insulin-like Growth Factor-binding Protein-5 Proteinase. *J Biol Chem* [Internet]. 2001 Jun 15 [cited 2020 Sep 23];276(24):21849–53. Available from: <https://pubmed.ncbi.nlm.nih.gov/11264294/>
 146. Arai C, Yoshizaki K, Miyazaki K, Saito K, Yamada A, Han X, et al. Nephronectin plays critical roles in Sox2 expression and proliferation in dental epithelial stem cells via EGF-like repeat domains. *Sci Rep* [Internet]. 2017 Mar 27 [cited 2020 Sep 23];7.

- Available from: <https://pubmed.ncbi.nlm.nih.gov/28345658/>
147. Meyer K, Kwon YC, Liu S, Hagedorn CH, Ray RB, Ray R. Interferon- α inducible protein 6 impairs EGFR activation by CD81 and inhibits hepatitis C virus infection. *Sci Rep* [Internet]. 2015 Mar 11 [cited 2020 Sep 23];5. Available from: </pmc/articles/PMC4355636/?report=abstract>
 148. Wang RY, Chen L, Chen HY, Hu L, Li L, Sun HY, et al. MUC15 inhibits dimerization of EGFR and PI3K-AKT signaling and is associated with aggressive hepatocellular carcinomas in patients. *Gastroenterology* [Internet]. 2013 [cited 2020 Sep 23];145(6). Available from: <https://pubmed.ncbi.nlm.nih.gov/23933603/>
 149. Porazinski S, Wang H, Asaoka Y, Behrndt M, Miyamoto T, Morita H, et al. YAP is essential for tissue tension to ensure vertebrate 3D body shape. *Nature* [Internet]. 2015 May 14 [cited 2020 Sep 23];521(7551):217–21. Available from: <https://pubmed.ncbi.nlm.nih.gov/25778702/>
 150. Mischler A, Karakis V, Mahinthakumar J, Carberry C, San Miguel A, Rager J, et al. Two distinct trophectoderm lineage stem cells from human pluripotent stem cells. *bioRxiv* [Internet]. 2019 Sep 8 [cited 2020 Sep 23];762542. Available from: <https://doi.org/10.1101/762542>
 151. Suryawanshi H, Morozov P, Straus A, Sahasrabudhe N, Max KEA, Garzia A, et al. A single-cell survey of the human first-trimester placenta and decidua. *Sci Adv* [Internet]. 2018 Oct 31 [cited 2020 Sep 23];4(10):4788. Available from: </pmc/articles/PMC6209386/?report=abstract>
 152. Messmer T, von Meyenn F, Savino A, Santos F, Mohammed H, Lun ATL, et al. Transcriptional Heterogeneity in Naive and Primed Human Pluripotent Stem Cells at Single-Cell Resolution. *Cell Rep* [Internet]. 2019 Jan 22 [cited 2020 Sep 24];26(4):815-824.e4. Available from: <https://pubmed.ncbi.nlm.nih.gov/30673604/>
 153. Fukamachi K, Matsuoka Y, Ohno H, Hamaguchi T, Tsuda H. Neuronal leucine-rich repeat protein-3 amplifies MAPK activation by epidermal growth factor through a carboxyl-terminal region containing endocytosis motifs. *J Biol Chem* [Internet]. 2002 Nov 15 [cited 2020 Sep 23];277(46):43549–52. Available from: <https://pubmed.ncbi.nlm.nih.gov/12297494/>
 154. Guan L, Li T, Ai N, Wang W, He B, Bai Y, et al. MEIS2C and MEIS2D promote tumor progression via Wnt/ β -catenin and hippo/YAP signaling in hepatocellular carcinoma. *J Exp Clin Cancer Res* [Internet]. 2019 Oct 17 [cited 2020 Sep 23];38(1):417. Available from:

- <https://jccr.biomedcentral.com/articles/10.1186/s13046-019-1417-3>
155. Krendl C, Shaposhnikov D, Rishko V, Ori C, Ziegenhain C, Sass S, et al. GATA2/3-TFAP2A/C transcription factor network couples human pluripotent stem cell differentiation to trophoblast with repression of pluripotency. *Proc Natl Acad Sci U S A* [Internet]. 2017 Nov 7 [cited 2020 Sep 23];114(45):E9579–88. Available from: <https://pubmed.ncbi.nlm.nih.gov/29078328/>
 156. Abd El-Aleem SA, Dekker L V. Assessment of the cellular localisation of the annexin A2/S100A10 complex in human placenta. *J Mol Histol* [Internet]. 2018 Oct 1 [cited 2020 Sep 23];49(5):531–43. Available from: <https://pubmed.ncbi.nlm.nih.gov/30143909/>
 157. Chua CEL, Tang BL. Engagement of the small gtpase rab31 protein and its effector, early endosome antigen 1, is important for trafficking of the ligand-bound epidermal growth factor receptor from the early to the late endosome. *J Biol Chem* [Internet]. 2014 [cited 2020 Sep 23];289(18):12375–89. Available from: <https://pubmed.ncbi.nlm.nih.gov/24644286/>
 158. Johnson WE. Origins and evolutionary consequences of ancient endogenous retroviruses [Internet]. Vol. 17, *Nature Reviews Microbiology*. Nature Publishing Group; 2019 [cited 2020 Sep 24]. p. 355–70. Available from: <https://pubmed.ncbi.nlm.nih.gov/30962577/>
 159. Glass K, Huttenhower C, Quackenbush J, Yuan G-C. Passing Messages between Biological Networks to Refine Predicted Interactions. Semsey S, editor. *PLoS One* [Internet]. 2013 May 31 [cited 2020 Sep 23];8(5):e64832. Available from: <https://dx.plos.org/10.1371/journal.pone.0064832>
 160. Weinberger L, Ayyash M, Novershtern N, Hanna JH. Dynamic stem cell states: Naive to primed pluripotency in rodents and humans [Internet]. Vol. 17, *Nature Reviews Molecular Cell Biology*. Nature Publishing Group; 2016 [cited 2020 Sep 23]. p. 155–69. Available from: <https://pubmed.ncbi.nlm.nih.gov/26860365/>
 161. Li Z, Kurosawa O, Iwata H. Establishment of human trophoblast stem cells from human induced pluripotent stem cell-derived cystic cells under micromesh culture. *Stem Cell Res Ther* [Internet]. 2019 Aug 7 [cited 2020 Sep 23];10(1):245. Available from: <https://stemcellres.biomedcentral.com/articles/10.1186/s13287-019-1339-1>
 162. Dong C, Beltcheva M, Gontarz P, Zhang B, Popli P, Fischer LA, et al. Derivation of trophoblast stem cells from naïve human pluripotent stem cells. *Elife* [Internet]. 2020 Feb 1 [cited 2020 Sep 23];9. Available from:

<https://pubmed.ncbi.nlm.nih.gov/32048992/>

163. Gao X, Nowak-Imialek M, Chen X, Chen D, Herrmann D, Ruan D, et al. Establishment of porcine and human expanded potential stem cells. *Nat Cell Biol* [Internet]. 2019 Jun 1 [cited 2020 Sep 23];21(6):687–99. Available from: <https://pubmed.ncbi.nlm.nih.gov/31160711/>



Verônica Luiza Vale Euclides Colovati

Bolsista de Doutorado do CNPq


Endereço para acessar este CV: <http://lattes.cnpq.br/2380817302197904>

ID Lattes: **2380817302197904**

Última atualização do currículo em 17/09/2019

Possui graduação em Nutrição pelo Centro Universitário São Camilo (2009). Mestre em Ciências pela Universidade de São Paulo (2014). Tem experiência na área de Nutrição, atuando principalmente nos seguintes temas: neurodesenvolvimento, nutrição, programação fetal, epigenética e composição corporal. Realizou o doutorado sanduíche por 1 ano no Lieber Institute for Brain Development, Johns Hopkins University (2019). Atualmente é doutoranda no Instituto de Psiquiatria da Faculdade de Medicina da Universidade São Paulo. **(Texto informado pelo autor)**

Identificação

Nome	Verônica Luiza Vale Euclides Colovati
Nome em citações bibliográficas	EUCLYDES, V.L.V.;EUCLYDES, V L V;EUCLYDES, V;EUCLYDES, VERÔNICA V;EUCLYDES, VERÔNICA V.;EUCLYDES, VERÔNICA L.V.
Lattes iD	 http://lattes.cnpq.br/2380817302197904

Endereço

Endereço Profissional	Universidade de São Paulo, Faculdade de Medicina da Universidade de São Paulo. Rua Doutor Ovídio Pires de Campos,785 Cerqueira César 05403010 - São Paulo, SP - Brasil Telefone: (11) 26616978
------------------------------	--

Formação acadêmica/titulação

2017	Doutorado em andamento em Doutorado em Ciências da Saúde. Universidade de São Paulo, USP, Brasil. com período sanduíche em Johns Hopkins University (Orientador: Jennifer Erwin). Título: Exposição ao estresse gestacional, alterações transcricionais na placenta e diferença entre os sexos; genes associados ao imprintoma são importantes? Orientador:  Helena Paula Brentani. Bolsista do(a): Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq, Brasil. Palavras-chave: Epigenética; Metilação do DNA; programação fetal; Desenvolvimento intrauterino. Grande área: Ciências Biológicas Grande Área: Ciências Biológicas / Área: Genética / Subárea: Epigenética.
2012 - 2014	Mestrado em Nutrição Humana Aplicada (Conceito CAPES 3). Universidade de São Paulo, USP, Brasil. Título: Relação das concentrações de adiponectina, leptina e zinco- α -2- glicoproteína no cordão umbilical com a composição corporal de lactentes nos quatro primeiros meses de vida,Ano de Obtenção: 2014. Orientador:  Patrícia Helen de Carvalho Rondó. Bolsista do(a): Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq, Brasil. Palavras-chave: Programação metabólica; Composição corporal; Nutrição; citocinas. Grande área: Ciências da Saúde
2010 - 2011	Especialização em Nutrição Clínica em Pediatria. (Carga Horária: 400h). Centro de Apoio e Pesquisa em Pediatria, CAEPP, Brasil. Título: Programação fetal: efeitos do retardo de crescimento intra-uterino no

2005 - 2009

desenvolvimento de doenças crônicas não transmissíveis na vida adulta.
Orientador: Patrícia Zamberlan.
Graduação em Nutrição.
Centro Universitário São Camilo, CUSC, Brasil.
Título: Alterações nos níveis de leptina e suas consequências na fisiopatogenia da Bulimia Nervosa.
Orientador: Roberta Medeiros.

Formação Complementar

2017 - 2017

Longitudinal analyses in life course epidemiology. (Carga horária: 30h).
Faculdade de Medicina da Universidade de São Paulo, FMUSP, Brasil.

Atuação Profissional

Universidade de São Paulo, USP, Brasil.

Vínculo institucional

2017 - Atual

Vínculo: Bolsista, Enquadramento Funcional: Doutoranda, Regime: Dedicção exclusiva.

Vínculo institucional

2012 - 2014

Vínculo: Bolsista, Enquadramento Funcional: Mestranda, Regime: Dedicção exclusiva.

Instituto de Psiquiatria da Faculdade de Medicina da Universidade São Paulo, IPQFMUSP, Brasil.

Vínculo institucional

2015 - Atual

Vínculo: Colaborador, Enquadramento Funcional: Livre

Atividades

07/2015 - Atual

Pesquisa e desenvolvimento , Laboratório de Investigação Médica LIM 23, .
Linhas de pesquisa
Estudo para verificação da dosagem de cortisol salivar e capilar como medida de avaliação do estresse crônico em mães e crianças
O efeito do Programa de Visitação para Jovens Gestantes sobre o desenvolvimento infantil: um estudo piloto

Faculdade de Saúde Pública da Universidade de São Paulo, FSPUSP, Brasil.

Vínculo institucional

2011 - 2011

Outras informações

Vínculo: Monitoria, Enquadramento Funcional: Monitora, Carga horária: 8
Monitoria voluntária na disciplina de Fisiopatologia da Nutrição da Faculdade de Saúde Pública da Universidade São Paulo, sob supervisão da Profa. Patrícia Helen de Carvalho Rondó.

Atividades

08/2011 - 12/2011

Outras atividades técnico-científicas , Faculdade de Saúde Pública da Universidade de São Paulo, Faculdade de Saúde Pública da Universidade de São Paulo.
Atividade realizada
Organização e monitoramento de disciplina.

Linhas de pesquisa

1. Estudo para verificação da dosagem de cortisol salivar e capilar como medida de avaliação do estresse crônico em mães e crianças
2. O efeito do Programa de Visitação para Jovens Gestantes sobre o desenvolvimento infantil: um estudo piloto

Projetos de pesquisa

2015 - Atual

Primeiros Laços

Descrição: Ensaio clínico randomizado para testar um programa de visitaçao domiciliar por enfermeiras intitulado Primeiros Laços para mães adolescentes que vivem em condições adversas na cidade de São Paulo. Oitenta jovens gestantes de baixa renda foram randomizadas para dois grupos: intervenção (N=40) e controle (cuidados usuais em saúde n=40). As participantes foram avaliadas por entrevistadoras cegas às 8-16 semanas de gestação (linha de base), 30 semanas de gestação, 3, 6 e 12 e 24 meses após o parto. O Programa Primeiros Laços foi implementado por enfermeiras treinadas e teve como objetivos atuar em cinco domínios: (1) Saúde e assistência social, (2) Saúde ambiental, (3)

Curso de vida; (4) Habilidades parentais; (5) Família e rede social de apoio. O estudo aliou o método experimental com avaliações sistemáticas longitudinais que compreenderam os domínios psicológicos, sociais, ambientais, psicopatológicos, desenvolvimentais, genéticos, epigenéticos e neurobiológicos da mãe e da criança ao longo do período pré e pós-natal. A pesquisa teve financiamento Grand Challenges Canada, Fundação Maria Cecília Souto Vidigal, Bill & Melinda Gates Foundation, Companhia Brasileira de Metalurgia e Mineração, Conselho Nacional de Desenvolvimento Científico e Tecnológico e Fundação de Amparo à Pesquisa do Estado de São Paulo..

Situação: Em andamento; Natureza: Pesquisa.

Integrantes: Verônica Luiza Vale Euclides Colovati - Integrante / Aloísio Felipe-Silva - Integrante / Guilherme Vanoni Polanczyk - Integrante / Euripedes Constantino Miguel - Coordenador / Helena Brentani - Integrante / Alicia Matijasevich - Integrante / Angélica Braz Simões - Integrante / Anna Maria Chiesa - Integrante / Bacy Bylik - Integrante / Alexandre A. Ferraro - Integrante / Carlos Tadashi Yoshizaki - Integrante.

Financiador(es): Grand Challenges Canada - Auxílio financeiro / Fundação Maria Cecília Souto Vidigal - Auxílio financeiro.

Áreas de atuação

1. Grande área: Ciências da Saúde / Área: Nutrição.
2. Grande área: Ciências Biológicas / Área: Genética / Subárea: Epigenética.

Idiomas

Inglês Compreende Bem, Fala Bem, Lê Bem, Escreve Bem.

Prêmios e títulos

2017 2017 Hugh Gurling Memorial Award, XXV World Congress of Psychiatric Genetics.




Produções

Produção bibliográfica

Artigos completos publicados em periódicos

Ordenar por

Ordem Cronológica

1. ABREU, LAÍSA R. S. ; SHIRLEY, MEGHAN K. ; CASTRO, NATÁLIA P. ; **EUCLYDES, VERÔNICA V.** ; BERGAMASCHI, DENISE P. ; LUZIA, LIANIA A. ; CRUZ, ANA M. ; RONDÓ, PATRÍCIA H. C. . Gestational diabetes mellitus, pre-pregnancy body mass index, and gestational weight gain as risk factors for increased fat mass in Brazilian newborns. *PLoS One* **JCR**, v. 14, p. e0221971, 2019.
2.  **EUCLYDES, V.L.V.**; CASTRO, NATÁLIA P. ; REHDER, L. ; BRITO, C. ; RIBEIRO, L. ; SIMOES, F. A. ; REQUENA, G. ; LUZIA, LIANIA A. ; RONDÓ, P H C . Cord blood concentrations of leptin, zinc- α 2-glycoprotein and adiponectin, and adiposity gain in the first three months of life. *NUTRITION* **JCR**, v. 54, p. 89-93, 2018.
3.  CASTRO, NATÁLIA P. ; **EUCLYDES, V** ; SIMÕES, FERNANDA FA. ; REHDER, L. ; BRITO, C. ; LUZIA, LIANIA A. ; DEVAKUMAR, D. ; RONDÓ, P H C . The Relationship between Maternal Plasma Leptin and Adiponectin Concentrations and Newborn Adiposity. *Nutrients (Basel)* **JCR**, v. 9, p. 182, 2017.
4.  MASCHIETTO, M. ; BASTOS, L. ; TAHIRA, A. C. ; BASTOS, E. P. ; **EUCLYDES, V** ; BRENTANI, A. ; FINK, G. ; BAUMONT, A. ; FELIPE-SILVA, A. ; FRANCISCO, R. P. V. ; GOUVEIA, G. ; GRISI, S. J. E. ; ESCOBAR, A. M. U. ; MOREIRA-FILHO, C. A. ; POLANCZYK, G. V. ; MIGUEL, E. C. ; BRENTANI, H. . Sex differences in DNA methylation of the cord blood are related to sex-bias psychiatric diseases. *Scientific Reports* **JCR**, v. 17, p. 44547, 2017.
5. SIMÕES, FERNANDA FA. ; CASTRO, NATÁLIA P. ; **EUCLYDES, VERÔNICA V.** ; LUZIA, LIANIA A. ; PAIVA, ADRIANA A. ; RONDÓ, PATRÍCIA HC. . Maternal adiposity and maternal and cord blood concentrations of vitamin D [25(OH)D3]. *Clinical Nutrition Experimental*, v. 9, p. 3-12, 2016.
6. SACCO, M R ; DE CASTRO, N P ; **EUCLYDES, V L V** ; SOUZA, J M ; RONDÓ, P H C . Response to Dr Andersen et al.. *European Journal of Clinical Nutrition* **JCR**, v. 68, p. 282-283, 2013.
7.  SACCO, M R ; DE CASTRO, N P ; **EUCLYDES, V L V** ; SOUZA, J M ; RONDÓ, P H C . Birth weight, rapid weight gain in infancy and markers of overweight and obesity in childhood. *European Journal of Clinical Nutrition* **JCR**, v. 67, p. 1147-1153, 2013.

Capítulos de livros publicados

1. DE CASTRO, N P ; **EUCLYDES, V.L.V** ; SIMOES, F. A. ; RONDÓ, P H C . Nutrição na Gestação e Lactação. Nutrição, Fundamentos e Aspectos Atuais. 3ed.São Paulo: Atheneu, 2013, v. , p. 1-477.

Apresentações de Trabalho

1. **EUCLYDES, VERÔNICA L.V.**. A importância da nutrição no TEA. 2019. (Apresentação de Trabalho/Conferência ou palestra).
2. WILEY, K. ; POLANCZYK, G. V. ; MIGUEL, E. C. ; **EUCLYDES, V L V** ; BRENTANI, H. . Associations between hair cortisol during pregnancy and methylation of hypothalamic-pituitary-adrenal axis regulating genes in cord blood of Brazilian infants. 2018. (Apresentação de Trabalho/Conferência ou palestra).
3. **EUCLYDES, VERÔNICA V** ; TAHIRA, A. C. ; BASTOS, L. ; MASCHIETTO, M. ; BRENTANI, H. . A circunferência abdominal de meninas está associada à metilação do gene mediador de estresse na gestação NR3C1. 2017. (Apresentação de Trabalho/Congresso).
4. **EUCLYDES, VERÔNICA V** ; CASTRO, NATÁLIA P. ; LUZIA, LIANIA A. ; RONDÓ, PATRÍCIA HC. . O aleitamento materno exclusivo até o 4º mês de vida está associado a um menor percentual de massa gorda de meninas, mas não de meninos. 2017. (Apresentação de Trabalho/Congresso).
5. **EUCLYDES, V.L.V** ; BARBOSA, A. R. ; BASTOS, L. ; POLANCZYK, G. V. ; MIGUEL, E. C. ; TAHIRA, A. C. ; MASCHIETTO, M. ; BRENTANI, H. . Home visits by nurses to pregnant adolescents of high social vulnerability in Brazil and the impact on the epigenetic of the newborn. 2017. (Apresentação de Trabalho/Simpósio).
6. **EUCLYDES, V.L.V** ; BARBOSA, A. R. ; BASTOS, L. ; BRENTANI, A. ; FINK, G. ; TAHIRA, A. C. ; MIGUEL, E. C. ; POLANCZYK, G. V. ; MASCHIETTO, M. ; BRENTANI, H. . Prenatal stress and epigenetic profile in neonates. 2017. (Apresentação de Trabalho/Congresso).
7. **EUCLYDES, V.L.V** ; DE CASTRO, N P ; REHDER, L. ; BRITO, C. . 13º Congresso Nacional da Sociedade Brasileira de Alimentação e Nutrição. 2015. (Apresentação de Trabalho/Congresso).

Demais tipos de produção técnica

1. **EUCLYDES, V.L.V**. Nutrigenômica na Obesidade. 2014. (Curso de curta duração ministrado/Outra).

Eventos

Participação em eventos, congressos, exposições e feiras

1. 7º Fórum Nacional de Alimentação Escolar. 2011. (Simpósio).
2. Atendimento Nutricional: Personal Diet e Gestão em Consultórios. 2010. (Outra).
3. 10º Congresso Nacional da Sociedade Brasileira de Alimentação e Nutrição. ANÁLISE DO PERFIL DE FUNCIONÁRIOS DE UM GRANDE BANCO BRASILEIRO PARTICIPANTES DE UM PROGRAMA CORPORATIVO DE ATIVIDADE FÍSICA. 2009. (Congresso).
4. CONNUT?2008 ? VIII Congresso Nacional de Nutrição e Tecnologia. 2008. (Congresso).
5. Curso de nutrição em terapia nutricional em cuidados paliativos em pediatria. 2008. (Outra).
6. IV Simpósio da LUNSC. 2008. (Simpósio).
7. SIMPONUTRI 2008. 2008. (Simpósio).
8. V Simpósio LUNSC: Alimentação orgânica: Origem segura como garantia de qualidade do produto final e sua implicação na harmonia ambiental e saúde. 2008. (Simpósio).
9. XV Simpósio de nutrição em cardiologia/ XXIX Congresso da sociedade de cardiologia do estado de São Paulo. 2008. (Congresso).
10. II Congresso Mineiro de Alimentação e Nutrição. 2007. (Congresso).
11. I Jornada acadêmica do curso de Nutrição da UNIVIÇOSA. 2006. (Simpósio).

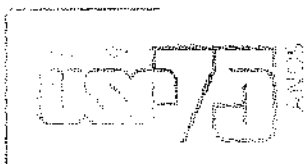
Organização de eventos, congressos, exposições e feiras

1. **EUCLYDES, V.L.V**. IV Simpósio da Liga Universitária de Nutrição São Camilo- Alimentação escolar e aplicação da educação alimentar como fator indispensável para promoção da saúde. 2008. (Outro).
2. ★ **EUCLYDES, V.L.V**. V Simpósio LUNSC. 2008. (Outro).

Outras informações relevantes

Participou como coordenadora operacional do projeto Consciência Alimentar, uma parceria entre Estado e Município de São Paulo. Atualmente, participa de projetos de pesquisas relacionados a epigenética e é bolsista de Desenvolvimento Tecnológico e Industrial B, na Faculdade de Medicina da USP, LIM 23.

[Imprimir currículo](#)



São Paulo, 16 de abril de 2012.

Il^{mo}(a). S^{ra}(a).

Profa. Dra. Alexandra Valeria Maria Brentani

Superintendência do Hospital Universitário
UNIVERSIDADE DE SÃO PAULO

REFERENTE: **Projeto de Pesquisa** “Avaliação do impacto do Programa de Saúde da Família no desenvolvimento da criança” – **Registro CEP-HU/USP:** 1191/12 – **CAAE:** 01604312.1.0000.0065

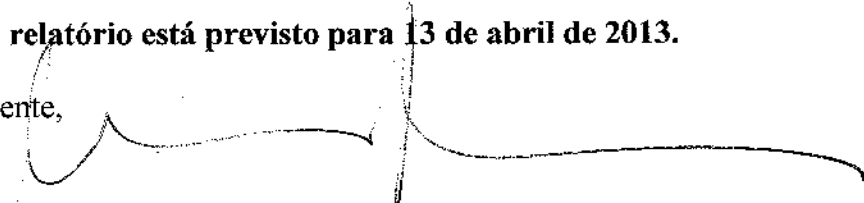
Prezado(a) Senhor(a)

O Comitê de Ética em Pesquisa do Hospital Universitário da Universidade de São Paulo, em reunião ordinária realizada no dia 13 de abril de 2012, analisou o Projeto de Pesquisa acima citado, considerando-o como **APROVADO**, bem como o seu **Termo de Consentimento Livre e Esclarecido**.

Lembramos que cabe ao pesquisador elaborar e apresentar a este Comitê, relatórios anuais (parciais ou final, em função da duração da pesquisa), de acordo com a Resolução nº 196/96 do Conselho Nacional de Saúde, inciso IX.2, letra “e”.

O primeiro relatório está previsto para 13 de abril de 2013.

Atenciosamente,



Dr. Maurício Seckler
Coordenador do Comitê de Ética em Pesquisa
Hospital Universitário da USP

USP - FACULDADE DE
MEDICINA DA UNIVERSIDADE
DE SÃO PAULO - FMUSP



PARECER CONSUBSTANCIADO DO CEP

DADOS DA EMENDA

Título da Pesquisa: O efeito do Programa de Visitação para Jovens Gestantes sobre o desenvolvimento infantil: um estudo piloto

Pesquisador: Eurípedes Constantino Miguel Filho

Área Temática: Genética Humana:

(Trata-se de pesquisa envolvendo Genética Humana que não necessita de análise ética por parte da CONEP;);

Versão: 7

CAAE: 41573015.0.0000.0065

Instituição Proponente: Faculdade de Medicina da Universidade de São Paulo

Patrocinador Principal: FUNDAÇÃO MARIA CECILIA SOUTO VIDIGAL
Bill & Melinda Gates Foundation
CNPQ

DADOS DO PARECER

Número do Parecer: 2.286.606

Apresentação do Projeto:

Projeto multidisciplinar, envolvendo vários profissionais, inicialmente aprovado pelo CEP em abril de 2015 e com apoio financeiro de agências nacionais e internacionais. A presente solicitação refere-se ao pedido de:

- a) ampliar o tamanho amostral (para um total de 200 primigestas jovens), o que aumentará o poder estatístico do projeto;
- b) ampliar o tempo de seguimento - incluindo as avaliações nos meses 18 e 24 de vida além dos previamente programados (3,6 e 12 meses) e
- c) incluir coleta de sangue e saliva aos 24 meses de idade (além das coletas relatadas em períodos anteriores e descritas na versão anterior do projeto).

Não houve alterações na metodologia do ensaio clínico, apenas aumento da amostra e do período de avaliação.

Objetivo da Pesquisa:

Não se aplica - Sem alterações

Endereço: DOUTOR ARNALDO 251 21º andar sala 36

Bairro: PACAEMBU

CEP: 01.246-903

UF: SP

Município: SAO PAULO

Telefone: (11)3893-4401

E-mail: cep.fm@usp.br

Continuação do Parecer: 2.286.606

Avaliação dos Riscos e Benefícios:

Sem alterações (já aprovado pelo CEP).

Comentários e Considerações sobre a Pesquisa:

Não se aplica

Considerações sobre os Termos de apresentação obrigatória:

O TCLE foi atualizado e modificado para incorporar as novas solicitações

Conclusões ou Pendências e Lista de Inadequações:

Projeto multidisciplinar, multiprofissional, de complexa execução, original e de grande relevância para a área. Foi pela primeira vez aprovado pelo CEP em 2015, e portanto encontra-se em andamento, mas no momento solicitou-se aumento do tamanho amostral e prolongamento dos períodos de avaliação.

Considerações Finais a critério do CEP:

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_971922_E5.pdf	28/08/2017 18:53:20		Aceito
Outros	TA_Versao_6_Ago_2017.pdf	28/08/2017 18:52:12	Eurípedes Constantino Miguel Filho	Aceito
Outros	Carta_Alteracoes_Emenda_Ago_2017.pdf	28/08/2017 18:51:20	Eurípedes Constantino Miguel Filho	Aceito
Outros	Carta_anuencia_SMS_2017.pdf	28/08/2017 18:51:00	Eurípedes Constantino Miguel Filho	Aceito
Projeto Detalhado / Brochura Investigador	Projeto_versao_6_Ago_2017.pdf	28/08/2017 18:49:58	Eurípedes Constantino Miguel Filho	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_versao_6_Ago_2017.pdf	28/08/2017 18:49:40	Eurípedes Constantino Miguel Filho	Aceito
Folha de Rosto	Programa Visitação Gestantes_Folha Rosto.pdf	06/02/2015 11:49:06		Aceito

Situação do Parecer:

Endereço: DOUTOR ARNALDO 251 21º andar sala 36

Bairro: PACAEMBU

CEP: 01.246-903

UF: SP

Município: SAO PAULO

Telefone: (11)3893-4401

E-mail: cep.fm@usp.br

USP - FACULDADE DE
MEDICINA DA UNIVERSIDADE
DE SÃO PAULO - FMUSP



Continuação do Parecer: 2.286.606

Aprovado

Necessita Apreciação da CONEP:

Não

SAO PAULO, 20 de Setembro de 2017

Assinado por:
Maria Aparecida Azevedo Koike Folgueira
(Coordenador)

Endereço: DOUTOR ARNALDO 251 21º andar sala 36

Bairro: PACAEMBU

CEP: 01.246-903

UF: SP


Município: SAO PAULO

Telefone: (11)3893-4401

E-mail: cep.fm@usp.br


DETALHAR PROJETO DE PESQUISA**— DADOS DA VERSÃO DO PROJETO DE PESQUISA**

Título da Pesquisa: O efeito do Programa de Visitação para Jovens Gestantes sobre o desenvolvimento infantil: um estudo piloto
Pesquisador Responsável: Eurípedes Constantino Miguel Filho
Área Temática: Genética Humana:
 (Trata-se de pesquisa envolvendo Genética Humana que não necessita de análise ética por parte da CONEP;)
Versão: 3
CAAE: 41573015.0.0000.0065
Submetido em: 07/12/2015
Instituição Proponente: Faculdade de Medicina da Universidade de São Paulo
Situação da Versão do Projeto: Aprovado
Localização atual da Versão do Projeto: Pesquisador Responsável
Patrocinador Principal: FUNDAÇÃO MARIA CECILIA SOUTO VIDIGAL
 Bill & Melinda Gates Foundation




Comprovante de Recepção:  PB_COMPROVANTE_RECEPCAO_638583**— LISTA DE PESQUISADORES DO PROJETO**

CPF/Documento ^	Nome ^	Atribuição	E-mail ^	Currículo	Tipo de Análise ^	Ação
905.025.900-68	Guilherme Vanoni Polaczyk	Contato Científico, Equipe do Projeto	gvp@usp.br	Lattes CV	PROPONENTE	
055.457.648-10	Eurípedes Constantino Miguel Filho	Contato Público, Pesquisador principal	ecmiguel@usp.br	Lattes CV	PROPONENTE	
772.072.976-53	Rossana Pulcineli Vieira Francisco	Equipe do Projeto	rpvf@uol.com.br	Lattes CV	PROPONENTE	
045.085.548-14	Helena Paula Brentani	Equipe do Projeto	helena.brentani@gmail.com	Lattes CV	PROPONENTE	
089.432.056-43	Verônica Luiza Vale Euclydes Colovati	Equipe do Projeto	veronicaeuclydes@hotmail.com	Lattes CV	PROPONENTE	
134.413.438-67	Alexandre Archanjo Ferraro	Equipe do Projeto	ferraro@usp.br	Lattes CV	PROPONENTE	
050.973.238-07	Anna Maria Church	Equipe do Projeto	amchiesa@usp.br	Lattes CV	PROPONENTE	
103.010.758-08	Lislaine Aparecida Fracoli	Equipe do Projeto	lislaine@usp.br	Lattes CV	PROPONENTE	
162.992.948-41	Bacy Fleitlich-Bilyk	Equipe do Projeto	bacy@uol.com.br	Lattes CV	PROPONENTE	
170.106.828-10	Mariza Marie Fujita	Equipe do Projeto	nageishi@uol.com.br	Lattes CV	PROPONENTE	

 Ocorrência 1 a 10 de 14 registro(s)
 
— LISTA DE COMITÊS DE ÉTICA DO PROJETO

Comitê de Ética ^	Tipo de Vínculo ^	Ação
65 - Faculdade de Medicina da Universidade de São Paulo - FMUSP	COORDENADOR	
76 - Hospital Universitário da Universidade de São Paulo - HU/USP	Co-participante	
86 - Secretaria Municipal da Saúde de São Paulo - SMS/SP	Co-participante	

— LISTA DE INSTITUIÇÕES DO PROJETO

CNPJ da Instituição ^	Razão Social ^	Tipo de Instituição ^	Comitê de Ética ^	Ação
63025530008512	Hospital Universitário da Universidade de São Paulo	Co-participante	76 - Hospital Universitário da Universidade de São Paulo - HU/USP	
4639500000139	Secretaria Municipal da Saúde de São Paulo - SMS/SP	Co-participante	86 - Secretaria Municipal da Saúde de São Paulo - SMS/SP	
63.025.530/0018-52	Faculdade de Medicina da Universidade de São Paulo	PROPONENTE	65 - Faculdade de Medicina da Universidade de São Paulo - FMUSP	

— LISTA DE PROJETOS RELACIONADOS

Tipo ^	CAAE ^	Versão ^	Pesquisador Responsável ^	Comitê de Ética ^	Instituição ^	Origem ^	Última Avaliação ^	Situação ^	Ação
P	41573015.0.0000.0065	3	Eurípedes Constantino Miguel Filho	65 - Faculdade de Medicina da Universidade de São Paulo - FMUSP	Faculdade de Medicina da Universidade de São Paulo	PO	E1	Aprovado	
Pc	41573015.0.3001.0086	2	Eurípedes Constantino Miguel Filho	86 - Secretaria Municipal da Saúde de São Paulo - SMS/SP	Faculdade de Medicina da Universidade de São Paulo	E1	Ec1	Em Avaliação Ética	
Pc		1	Eurípedes Constantino Miguel Filho	76 - Hospital Universitário da Universidade de São Paulo - HU/USP	Faculdade de Medicina da Universidade de São Paulo	E1	Pouco tempo	Em Recepção e Validação Documental	

ANEXO 3

Clinical Epigenetics

Title: DNA methylation mediates the effects of a home-visiting intervention during pregnancy over infant's cognitive scores at 12 months old: results of a randomized controlled trial.

Verônica Luiza Vale Euclides¹, e-mail: veronicaeuclides@usp.br, wrote the paper, performed data collection, bioinformatic and mediation analysis

Vinicius Daguano Gastaldi¹, e-mail: vinidg@gmail.com=performed bioinformatic analysis

Arthur Sant'Anna Feltrin⁷, e-mail: s.feltrin@ufabc.edu.br=performed bioinformatic analysis

Gisele Gouveia¹, e-mail: gisele.rgouveia@gmail.com= all laboratories procedures

Hugo Cogo = supervised mediation analysis

Aloísio Felipe-Silva⁴, e-mail: cord blood collection

Rossana Pulcinelli cord blood collection

Eurípedes Constantino Miguel¹, e-mail: ecmiguel@usp.br performed clinical intervention and data collection

Guilherme Vanoni Polanczyk¹, e-mail: polanczyk.guilherme@gmail.com performed clinical intervention and data collection

Anna Chiesa <amchiesa@usp.br>, performed clinical intervention and data collection

Lislaine Fracolli <lislaine@usp.br>performed clinical intervention and data collection

Alicia Matijasevitch <amatiija@yahoo.com>performed clinical intervention and data collection

Alexandre Ferraro <ferraro@usp.br>performed clinical intervention and data collection

Adriana Argeu <adrianaargeu@gmail.com>performed clinical intervention and data collection

Mariana Maschietto⁶, e-mail: marianamasc@gmail.com contribute to methylation analyses and wrote the paper

Helena Paula Brentani^{1*}, e-mail: helena.brentani@gmail.com – conceive the work, coordinated all the study, revised bioinformatics and mediation analysis, wrote the paper Corresponding author

All authors revised the manuscript

¹Institute of Psychiatry, University of São Paulo, Medical School, SP, Brazil;

²Department of Anthropology, Yale University, Connecticut, USA;

³School of Nursing, University of São Paulo, Medical School, SP, Brazil;

³Department of Pediatrics, University of São Paulo, Medical School, SP, Brazil;

⁴Department of Pathology, University Hospital of Sao Paulo, SP, Brazil;

⁵Department of Obstetrics and Gynecology, University of São Paulo Medical School, SP, Brazil

⁶Research Center, Boldrini Children's Hospital, Campinas, Brazil;

⁷Center for Mathematics, Computation and Cognition, Federal University of ABC (UFABC),
Santo André, SP, Brazil

Main text

Tables

Table 1. Characteristics of cohort

Table 2. Mediation analysis

Figures

Figure 1. Intervention group presented a smaller epigenetic gestational age mean compared to the control group (39.5 x 40.1 weeks, respectively).

Figure 2. the variable selection criteria to test mediation which selected

Supplemental Material

Tables

Supplemental Table 1. Characteristics of the samples from this study compared to the original RCT “First Ties” (n= 80x29)

Supplemental Table 2. Characteristics of completed neurodevelopmental measures of RCT “First Ties” (n=64)

Supplemental Table 3. DMPs identification

Supplemental Table 3A. DMPs with delta-beta >20%

Supplemental Table 4. CpGs associated with the cognitive outcome at 12 months adj. ($R^2 >0.4$)

Supplemental Table 5. DMRs identification and respective mean M-values.

Supplemental Table 6. DMRs associated with the cognitive domain at 12 months

Supplemental Table 7. Co-methylation analyses modules identification

Supplemental Table 8. All variables tested in the mediation analysis

Figures

Supplemental Figure 1. study design

Supplemental Figure 2 3,090 DMPs with 1,855 (60%) hypomethylated CpG sites in the intervention group

Supplemental figure 2. Co- methylation modules and its conservation score

Abstract -

Introduction and Aims: Prenatal interventions in high risk mother groups could impact positively the brain development. The crosstalk between stress exposure and maternal and fetus individuals may be mediated by epigenetic mechanisms, including DNA methylation involved with the programming mechanism of the offspring. We hypothesize that providing maternal support by nurses' home visiting to a young high social vulnerability pregnant group could improve neurodevelopmental outcomes through mediation of DNA methylation.

Methods and Results: Eighty low income pregnant adolescents participated in a pilot randomized clinical intervention study (ClinicalTrials.gov, Identifier: NCT02807818). In a subsample of 16 neonates from intervention (I) and 13 matched controls (C), cord blood DNAm was interrogated using the Illumina Infinium Human Methylation 450 BeadChip arrays. We found 3,090 differentially methylated CpG sites (DMPs), 21 differentially methylated regions (DMRs) and one co-methylated module not conserved between groups. Gestational age estimated by methylation showed lower means in the intervention group. At 12 months of age, neurodevelopment was assessed by the Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III) reporting a higher cognitive domain score in the intervention group. Mediation analysis was tested with CpGs at a site, region, co-methylated module levels and the epigenetic clock. DMPs with higher R^2_{adjusted} (>0.4) associated with cognitive domain were tested regarding the mediation hypotheses pointing to 13 DMPs and the DMR located at *PF4* showing causal mediation effects between intervention and cognitive domain at 12 months of age.

Conclusion: The results showed differences of DNA methylation at site, region, co-methylation module and epigenetic clock between groups. Specific DMPs and PF4 DMR, associated with synaptic maintenance seem to mediate the association between the intervention and the cognitive domain at 12 months of age. Future research should replicate our findings in a larger sample and evaluate if these epigenetic changes are persistent.

Keywords: DNA methylation - mediation analysis- prenatal care - maternal-fetal relations

Introduction

Developmental Origins of Health and Disease theory explores the role of prenatal and perinatal experiences in health-disease throughout life [1]. Environmental exposures during prenatal life such as maternal diet, alcohol/drug abuse, smoking, preeclampsia, gestational diabetes and psychosocial stress could result in sustained effects in the molecular mechanisms programming offspring health, and contribute to neurodevelopmental disorders risk [2]. Epigenetic mechanisms have been related to environmental exposures, specially DNA methylation (DNAm), which has been suggested as a biomarker, and a possible mediator in health outcomes [3–5]. The epigenomic state can be established through behavioral programming as reported in animal models where poor levels of maternal care altered the offspring epigenome at a glucocorticoid receptor (GR) gene promoter in the hippocampus [6]. It is also reported that the environmental input can be mediated by DNAm on important genes associated with the neural circuitry components modulating the brain plasticity [7].

Observational studies in humans documented that specific patterns of DNA methylation (DNAm) are associated with prenatal stressful environments and subsequent behaviors. For example, war trauma and chronic stress during gestation was associated with specific sites of the gene BDNF in cord blood, placental tissue, and maternal venous blood [8]. The prenatal famine exposure predicted changes in metabolic markers in adulthood which were mediated by DNAm in genes related to the metabolism [9]. DNAm levels of specific single CpG sites and/or entire regions is frequently associated with prenatal stress and brain development, indicating the relevance of further understanding this mechanism [10,11]. In this direction, a recent development is the estimation of gestational age based on DNAm levels of specific CpGs. Estimated gestational age has been extensively associated to prenatal stress and, despite little explored in longitudinal studies, might be related to infant health outcomes [12].

Based on the concept of increased plasticity during intrauterine life, psychosocial interventions during pregnancy have been tested and demonstrated to ensure better conditions for the child's development [13]. Indeed, meta-analysis have demonstrated that

psychosocial interventions in early pregnancy promote positive developmental outcomes [14]. Nevertheless, there are very few randomized controlled trials testing psychosocial interventions during pregnancy aiming to mitigate the effect of stressors over the fetus that investigate epigenetic mechanisms underlying its positive effects [15]. The Nurse Family Partnership (NFP) lead to variation in DNAm in the blood at 27 years of age, independent of sex, ancestry, cellular heterogeneity and a polygenic risk index for major psychiatric disorders, but dependent of some lifestyle factors, such as smoking [16].

Few studies have integrated clinical trial studies, epigenetic and phenotypic data in order to explicitly test the role of DNAm as a potential biological mediator of environmental effects on health outcomes [3,17] Most of the research has been in high-income countries, but the effects may be different and often even greater in low-middle income countries, because of greater stress due to high levels of interpersonal violence, natural disasters, poverty, and potential interactions with infections [18], food insufficiency, dietary deficiencies, crowded housing, and extremes of temperature [19]. Aiming to address this important gap in the literature, we conducted a pilot study to investigate if DNAm changes at CpG site, region and network modules could mediate the effects of the home-visiting program “Primeiros Laços” to pregnant teenagers living in extreme social vulnerability on infants’s neurodevelopmental outcomes at 12 months of age.

Methodology

Design

A randomized controlled trial was conducted to test the effects of “Primeiros Laços” program in improving mother-infant relationship (NCT02807818) [20].

Participants

The study included first-time pregnant adolescents living in an urban area characterized by high rates of poverty and urban violence in Sao Paulo, Brazil. Eighty women were randomized to the intervention (n=40) and to the control (received the usual prenatal care, n=40) groups. The inclusion criteria were (1) women between 14 to 19 years that were (2) pregnant for the first time, (3) between 8 and 16 weeks’ gestation, (4)

living in western regions of São Paulo, and (5) low socioeconomic status [classes C–D/E according to the Criteria for Economic Classification assessed by the Brazilian Association of Research Companies questionnaire[21]. Exclusion criteria were (1) mental, visual, or hearing disabilities. Randomization was stratified according to type of basic health unit and years of grandmother's schooling.

Intervention

Primeiros Laços is a home-visiting program delivered by trained nurses to first-time pregnant adolescents and their infants, starting from the first 16 weeks of pregnancy until the child is aged 24 months. The visitors were nurses specialized in maternal or mental health and were supervised weekly by senior nurses and psychologists. Primeiros Laços is based on three theoretical frameworks (*Attachment theory*, Bowlby et al 1979; *Self-Efficacy theory*, Bandura, 1977; *Bioecological Development theory*, Bronfenbrenner, 1998) and was structured in five axes (*health care, health environment, parenting and attachment, social and family network, life project*). The program was developed by our team (Chiesa et al., 2008) based on existing home-visiting programs (Olds, 2006; Slade et al., 2005). The care plan for each mother-infant dyad was carefully designed to strengthen maternal competences for warm and responsive care. A key factor was establishing positive relationships between home visitors and the family. During the visits, nurses helped parents develop child-centered interactions, improve their bond with their infant, reflect on their own attachment history and parenting they received, consider their child as an individual with his/her own needs, feelings, and thoughts, and improve their parenting skills by modeling. The visitor encouraged attuned parenting and stimulated sensitive behaviors, such as being attentive to the child's communicative signals or following the child's lead. Parents were also given support to think in a reflective way. Frequency of visits was weekly (first/last month of pregnancy/puerperium), biweekly (gestation/2-20 months of child's age) and monthly (21-24 months of child's age). Mothers were expected to receive 60-62 visits in total.

Care-as-usual

Control participants received care-as-usual from the Sistema Único de Saúde (SUS), Brazil's public health system [22], according to national guidelines of the

Ministry of Health that are in line with the World Health Organization guidelines. Prenatal and postnatal care is delivered by primary healthcare units free of charge and focuses on preventive interventions, early detection of gestational risk, and referral to specialized health services in case of high-risk pregnancies. Participants from the intervention group also had access to these public health services.

Behavior and neurodevelopmental outcomes

To assess socio-economic conditions, the following instruments were used: Brazilian Scale of Food Security (EBIA) [23], monthly family income and maternal education. Also, we collected data for alcohol/drug abuse and smoking habits. Anthropometric data were used for nutritional status evaluation. To evaluate depression and anxiety, The Beck Depression (BDI) and Anxiety Inventory (BAI) [24,25], were collected at baseline and at 30th gestation week. Infant weight and length were measured at birth and at 12 months by trained professionals. The values (z-score) of the infant's growth parameters were based on the growth standards of the World Health Organization (WHO)[26].

Finally, the Bayley III Scales of Infant Development (BSID-III)[27,28] were administered by trained psychologists at 12 months old to evaluate neurodevelopmental characteristics of the infants. The five neurodevelopmental domains were considered separately: Cognitive, Receptive Language, Expressive Language, Fine Motor Skills and Gross Motor Skills.

Sample processing

From 80 pregnant women enrolled in the study, 32 had the baby umbilical cord blood samples collected in EDTA vacutainer tube (B.D. Scientific) immediately after delivery. DNA was extracted with the *QIAamp DNA Mini Kit* (Qiagen). NanoDrop (Thermo Fisher Scientific) and Qubit (Thermo Fisher Scientific) were used to assess DNA purity and quantity, respectively. Bisulfite converted DNA (EZ DNA Methylation kit, Zymo Research Corp) was hybridized in the Human Methylation 450 BeadChip arrays (HM450K, Illumina), following the Illumina Infinium HD methylation protocol. Both experiments were processed by Deoxi Biotecnologia (www.deoxi.com) according to the manufacturer's instructions. Raw data were extracted by the iScan SQ scanner

(Illumina) using GenomeStudio software (v.2011.1), with the methylation module v.1.9.0 (Illumina), into IDAT files, which were used for further analyses.

Minfi package version 1.26.2 [29] was used for quality control of the samples as well as to apply quantile normalization [30] and to calculate M-values. We removed probes that contained SNPs at the CpG site, those located at the sex chromosomes, and those known to co-hybridize to more than one genomic sequence resulting in 429,174 probes. ChAMP package version 2.10.2 [31] was used to identify and adjust the M-values for batch effects and for sex. Cell composition for the sample was assessed using the FlowSorted.CordBlood.450k package implemented in minfi but no significant differences were found when comparing the intervention and control groups.

Differential Methylation Analysis

Differential methylation sites and differentially methylated regions (DMRs) were calculated from M-values. Differentially methylated CpG sites (adjusted p-value ≤ 0.05 , Benjamini-Hochberg method) and DMRs were obtained using the DMP and DMRcate functions [32] respectively, as implemented in ChAMP. After identifying the DMPs, they were further annotated to obtain information regarding enhancers, transcription factor binding sites (TFBSs) and DNase I Hypersensitivity sites (DHSs). For enhancers we used the complete annotations from EnhancerAtlas for Fetal Brain and Fetal Placenta. For TFBSs and DHSs we used the wgEncodeRegDnaseClusteredV3 and wgEncodeRegTfbsClusteredV3 tracks from Genome Browser's Data Integrator (<http://genome.ucsc.edu/cgi-bin/hgIntegrator>).

We considered DMRs sequences of 300 bp containing at least seven CpG sites and showing differences of $FDR \leq 0.05$. To identify putative TFB, several publicly available bioinformatics resources were used. UCSC genome browser was used to find HMR conserved TFBs and the proximity of putative ENCODE TFB regions to the CpG sites assayed via chip sequencing. The presence of putative TFBs flanking a given CpG site was further assessed using PhysBinder [33], with the threshold level set to average precision, and MotifMap

Weighted Correlation Network Analysis

The Weighted Correlation Network Analysis (WGCNA) package [34] was used to compare co-methylation networks from intervention and control groups using DMPs. From each group, Pearson correlation between samples M-values were used to construct two correlation matrices, with the estimated network power value ($\beta = 15$) was used as a scale-free independence index, resulting in an R^2 of 0.90 and 0.96 for interventional and control groups, respectively. For both networks, we used a $\text{deepsplit} = 3$, $\text{mergecuthigh} = 0.25$ and minimum module size = 30.

Z-summary was used to evaluate preservation between networks (100 permutations). Modules with a low or moderate level of preservation are those with Z-summary ≤ 10 , which indicates large topological variation [35]. Considering that the Z-summary indicates a topological variation between each group, modules with a low Z-summary value indicate the CpGs with the largest differences in the co-methylation pattern. CpG sites with a kME correlation ≥ 0.9 were considered hubs of that module.

Enrichment analysis

Selecting the DMPs mapped into genes, we used the clusterprofiler package [<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3339379/>] to evaluate enrichment for Gene Ontology (Biological Process, Cellular Component and Molecular Function) and signaling pathways (KEGG). We also applied the Over Representation Analysis (ORA) on unique genes using the human genome as background. A FDR value ≤ 0.05 was used as the significance threshold.

Epigenetic Clock for Gestational Age at Birth

To calculate the DNAm age, methylation levels of 148 CpGs sites were submitted to an elastic net regression based in six training datasets, described elsewhere [36].

Statistical Analysis

Statistical analyses were performed in R v. 3.5.0. Chi-square and t-tests were applied to compare the phenotypes between intervention and control groups. To perform the mediation analysis, we defined the following criteria at CpG site, region, or module level: (1) Only DMPs associated with cognition were tested as mediators. DMPs related

to cognition (R^2 adj >0.4) were defined by a linear regression model using M-values of DMPs as dependent variable and the cognitive domain score and sex as independent variables. (2) From co-expression data, all CpG sites that were hubs from the not preserved co-methylation module. (3) Each DMR was represented as the M-values means of all CpG sites within the specific region. (4) Epigenetic clock as raw values.

We used the mediation package version 3.4.4 [37] to analyze causal mediation with bootstrap of 10,000 replications for confidence intervals. Data from mediation model for each CpGs or DMRs include estimates and 95% CIs of total, direct, and mediation effects. The Average Causal Effects (ACME) states whether intervention positively (the probability of increasing cognition score) or negatively (decreases the cognitive score) affects the outcome (cognitive score) through a mediator (DNAm, according to the selected ones). Total effect and Average Direct Effect show if the intervention is likely to change cognition in the mediation models. The function *medsens* was applied to perform sensibility analysis to test effects of unmeasured confounders.

RESULTS

Characteristics of the sample

We collected cord blood from 32 infants. Three infants had incomplete neurodevelopmental measures or dropped out from the study, resulting in 13 infants in the control group and 16 infants in the intervention group. There were no differences concerning socioeconomic, smoking, pre-pregnancy BMI, anxiety, and depressive symptoms between the mothers' groups as well as no differences regarding anthropometric measurements nor obstetric differences at birth between infants' groups. In comparison to the whole sample, the sub-sample with cord blood information were more frequently boys ($p=0.01$) and mothers presented higher scores on Beck Anxiety Inventory ($p=0.04$). But there were no differences between groups in the subsample.

. At 12 months, the intervention group presented higher scores for gross motor and cognitive domains (**Table 1**).

Thus, although the cognitive and the gross motor skills were different between intervention and control groups in this subsample, we only analyzed the cognitive score

as an outcome because it was concordantly different in comparison to the whole cohort from the RCT. RCT “First Ties” had 63 neurodevelopmental evaluations at 12 months of age (**Supplementary Table 2**). The study design can be visualized in **Supplemental Figure 1**.

Identification of methylation differences between intervention and controls

By comparing DNAm levels of cord blood, intervention and control groups presented 3,090 DMPs, being 1,855 (60%) hypomethylated CpG sites in the intervention group (**Supplemental Figure 2**). The 3,090 DMPs were located at 1,267 genes, being 217 mapped within 1,500 bp of the transcriptional start site (**Supplemental table 3**). Nineteen DMPs had presented delta-beta >20%, six mapped to genes previously associated with neurodevelopment as *IL18RAP*, *OR2L13* and *CACNA1A* (**Supplemental table 3A**).

The 3,090 DMPs were tested for their potential association with the cognitive score of these infants at 12 months of age by multiple linear regression controlled by sex using the cognitive domain score as outcome. Eighteen CpG sites presented adj. $R^2 > 0.4$ (**Supplemental table 4**), being 11 CpG mapped to the genes: *C10orf118*, *C11orf58*, *C6orf174*, *GRIA1*, *LHFPL4*, *LOC100240726*, *MYOM2*, *RB1*, *STXBP6*, *TBL1XR1* and *TOR1AIP2* (**Supplemental table 4**). From the CpGs not annotated to a gene, using ENCODE annotation, we found one CpG site (cg19610750) located on a DNase hypersensitivity site that was annotated for multiple transcription factors, including *CTCF*.

We identified 21 DMRs, 15 (71%) hypomethylated in the intervention group. Fifteen DMRs were located at genes (*HLA-DPB1*, *ABAT*, *ANKRD30B*, *PLEKHH3*, *ABCC13*, *LOC650226*, *LOC441666*, *PAX8*, *BOLL*, *RUFY1*, *LOC149837*, *HCG4P6*, *PF4* and *TACSTD2*), mostly located next to the transcriptional start site and/or promoter region of the respective gene (Suppl table 5). The DMRs were tested for their potential association with the cognitive score of these infants at 12 months of age by multiple linear regression controlled by sex using the cognitive domain score as outcome and *PF4* was the unique DMR with a significant p-value (0.006) with a β (SE) = -5.48(1.8) and R^2 adj=0.30 associated with the cognitive outcome at 12 months (Supplemental table 6).

Co-methylation analyses of the 3,090 DMPs pointed to six modules: MEblue (218 CpGs), MEbrown (138 CpGs), MEgreen (129 CpGs), MERed (68 CpGs), METurquoise (219 CpGs) and MEyellow (136 CpGs). Annotation of CpG sites as well as the respective modules are available at **Supp. Material dmp_new-WITH-wgcnamodules.txt**. METurquoise ($Z_{\text{summary}} = 7.42$, $ME_{\text{cor}} = 0.13$, $p\text{-value} = 0.055$) presented low conservation score (**Supplemental figure 2 3**). METurquoise had 126 CpG sites (out of 219) mapped to 96 genes. Enrichment analysis suggested the involvement of actin filament organization, actin cytoskeleton organization, supramolecular fiber organization and epithelial tube morphogenesis biological processes (Gene Ontology). This result was not maintained after the multiple correction adjustment (FDR). Four CpGs sites were classified as hub nodes in the METurquoise network: one located at the body of *CNTNAP1* and three CpG sites located on the TSS200 of *FILIP1*.

Gestational age estimated by methylation (GAm) point to a younger age of newborns receiving the intervention

Intervention and control groups did not present differences in the chronological gestational age at birth (**Table 1**, 39.4 X 39.3, intervention and control, respectively), nevertheless the intervention group presented a smaller epigenetic gestational age mean compared to the control group (39.5 x 40.1 weeks, respectively) (**Figure 1**). Logistic regression controlled by sex and with gestational age as outcome variable indicated a coefficient of -0.35 towards intervention group ($p = 0.016$).

DNAm show mediating effect on cognition at 12 months

Causal mediation analyses were applied to assess whether DNAm levels might be a mediator between intervention and infants score cognition at 12 months of age. (**Supplemental Table 7**). Figure 2 shows the variable selection criteria to test mediation which selected (1) 18 CpG sites associated with cognitive domain ($\text{adj. } R^2 > 0.4$), (2) 4 CpG sites that were hubs in the less preserved co-methylation module (3) 21 DMRs, considering the averaged M-values of all CpGs within the region, (4) epigenetic clock as raw values.

We found 11 DMPs presenting significant ACME and proportion mediation, with six located at *C11orf58*, *TOR1AIP2*, *MYOM2*, *C10orf118*, *STXBP6* and *LOC100240726*.

Moreover, three DMRs located at *BOLL*, *LOC650226* and *PF4* presented significant ACME but only *PF4* presented significant proportion mediation. One CpG from *FILIP1* had a significant ACME but not a significant proportion of mediation. Finally, the gestational age given by the epigenetic clock did not show mediation effect.

DISCUSSION

In this study we demonstrated that epigenetic changes act as mediators of the effects of a nurse-led home-visitation intervention during pregnancy over infants' cognitive scores at 12 months of age in a high-risk sample participant of a randomized controlled trial.

We report DNAm differences between intervention and control groups at CpG sites, DMRs and co-methylation network levels. Comparing the methylation analyses, we found some overlapping CpGs and genes, although none acting as a mediator. *ABCC1* had a DMRs and a DMP with delta-beta >20% (adjP=5.5E-6). *OR2L13* also had a DMR and one DMP with delta-beta >20% (adjP ≤0.001). *OR2L13* was found by multiple independent analysis from EWAS studies [38]. *OR2L13* belongs to the olfactory sensory system pathway, which is related to social impairment for typically developing children [39]. In our study, the CpG site (cg20507276) located in *OR2L13* was less methylated in the intervention group (mean $\Delta\beta$ = -0.17) but was found hypermethylated (mean $\Delta\beta$ = 0.18) in individuals with autism spectrum disorder [40]. The cg08860498 mapped in *CNTNAP1* is a hub from the less preserved network module, MEtuquoise, and it showed five significant hypomethylated DMPs (cg07871971, cg16308533, cg12559685, cg27616541, cg08860498). *CNTNAP1* is involved in the formation of myelinated axons and maintaining the stability of adjacent connections [41,42]. Two DMPs, cg23223755 and cg21164545, located at *LHPL4* and *LOC100240726*, respectively, are associated with cognition and belong to the MEtuquoise. Fast inhibitory transmission in the brain is mediated mostly by ionotropic GABA_A receptors (GABAARs). LH4 forms a stable tripartite complex with GABAARs and neuroligin-2 in the brain. Moreover, LH4 is required for the synaptic localization of GABAARs and inhibitory synaptic transmission in the hippocampus [43]. DMPs located at *ANKRD30B* (cg02105152, cg23703062,

cg21293934, cg16584393, cg02992546) also belong to the MEtuquoise module and were considered a DMR. *PAX8*, that also had a DMR, present a role in mid-hindbrain development by binding to *EN2* regulatory sequences[44] and was associated with neurobehavioral status at birth [45].

We analyzed the epigenetic clock, an indirect measure of cellular senescence, considering that aging could begin even before birth [46]. The intervention group presented a lower age obtained by epigenetic clock. Infants born from mothers that smoked during gestation presented three times higher risk of aging compared to infants from nonsmokers mothers[47] but this could not explain the result once no difference was found regarding smoking between the groups. Moreover, more longitudinal studies are needed to understand how to interpret these results and if the epigenetic clock could predict health outcomes later in life.

All tested DMPs presented an indirect effect on the cognitive score at 12 months measured by the Bayley scale. Considering the location of CpG sites with significant ACME, but non-significant proportion mediation, *TBLIXR1* and *GRIA1* were previously associated with neurodevelopment. *TBLIXR1* (cg03930153) recruits β -catenin to Wnt target genes that activate transcription. *TBLIXR1* mutations were associated to neuropsychiatric conditions [48,49], *GRIA1* (cg22798121) has a role on glutamatergic synaptic transmission being related to synaptic plasticity and neurodevelopmental disorders[50,51].

Three out of 11 DMPs with a mediation effect were mapped to genes involved neuron maintenance and functionality: *C10orf118* (cg05405872), *STXBP6* (cg11734401) and *TOR1AIP2* (cg07761912). *C10orf118* and *STXBP6* are involved with vesicle trafficking and signaling [52] as they maintain the functional organization of the cell and thus, play a role in neuronal processes, such as localization of dendritic proteins [53]. *TOR1AIP1* encodes cofactors for the ATPase TorsinA which has been implicated in maintaining normal cellular physiology. Recent studies have also demonstrated that Torsins are important in developing neurons, suggesting a role during neuronal differentiation [54]. The three DMPS are located at the 5'UTR increasing the probability that their methylation is associated with downregulation of gene expression. Thus, we

suggest that the cognitive score improvement by the intervention is mediated by a decrease in *TORIAIP2* expression and an increase in the expression of *C10orf118* and *STXBP6*, with the former mediating 70% of the total effect. Also, cg19610750, that mediates the intervention effect on cognition, is located at an intergenic region where there are binding sites for the transcriptional factors *CTCF* and *EGRI*, both involved with fetal development [55,56]

Regarding the DMRs, the hypomethylated *PF4* DMR mediated 55% of the total effect in the association between intervention and cognitive scores. *PF4* encodes a member of the CXC chemokine family that initiates a signal transduction cascade of acute and delayed functions in human monocytes, including phagocytosis, respiratory burst, survival, and the secretion of cytokines [57]. *PF4* was a hub in a coexpression network analysis comparing early psychosocial stress exposed individuals to controls [58]. The module was enriched for chemokine activity and for secretory vesicles functional biological processes. Moreover, *PF4* overexpression contributes to an anti-aging effect in aged mice taking collagen peptides [59]. We describe *PF4* as hypomethylated in the intervention group, which could be associated with gene overexpression and suggest that it could be contributing to the lower age acceleration observed and the better cognitive scores in this group.

In comparison to the whole sample, the sub-sample with cord blood information were more frequently boys ($p=0.01$) and mothers presented higher scores on Beck Anxiety Inventory ($p=0.04$). But there were no differences between groups in the subsample. Even though we had a reduced statistical power due to the smaller sample size. It is also important to stress that a permissive cutoff criterion of 0.05 for multiple correction was used to find differentially methylated sites. However, it could generate more false positives and we attenuated by selecting only CpG sites for the mediation analysis associated with the outcome that explained more than 40% of the regression model. We also recognize that epigenetic changes in the umbilical cord blood do not necessarily direct reflect the fetal brain. Nevertheless, surrogate tissue can be informative even if the investigated tissue is not directly involved in the phenotype of interest, as they might reflect the exposure to environmental influences (it has been shown, e.g., for the effects of prenatal sub nutrition [9], chemicals, psychosocial adversity [60] or combat

[61]. Thus, although variation in peripheral tissue may not directly reflect epigenetic variation in the brain, they may still represent usable biomarkers as they reflect an exposure that leads to the phenotype in a different tissue [62].

Even with some limitations the presented findings have important implications as we documented the epigenetic effects of a psychosocial intervention during pregnancy, which is relevant to understand the biological mechanisms related to the intervention (Insel & Gogtay, 2014), *JAMA Psychiatry* 2014; 71(7): 745-6.

In conclusion, this pilot study provides support for DNAm differences in genes involved with neurodevelopment by comparing infants whose mothers received visits by trained nurses during pregnancy to those who did not receive the intervention. Specifically, we found that the association between the intervention and the cognitive score at 12 months can be mediated by DNAm.

References

1. Wadhwa PD, Buss C, Entringer S, Swanson JM. Developmental origins of health and disease: Brief history of the approach and current focus on epigenetic mechanisms [Internet]. *Semin. Reprod. Med.* NIH Public Access; 2009 [cited 2020 Jul 9]. p. 358–68. Available from: [/pmc/articles/PMC2862635/?report=abstract](https://pubmed.ncbi.nlm.nih.gov/19271200/)
2. DeSocio JE. Epigenetics, maternal prenatal psychosocial stress, and infant mental health. *Arch Psychiatr Nurs* [Internet]. W.B. Saunders; 2018 [cited 2020 Jul 9];32:901–6. Available from: <https://pubmed.ncbi.nlm.nih.gov/30454636/>
3. Felix JF, Cecil CAM. Population DNA methylation studies in the Developmental Origins of Health and Disease (DOHaD) framework. *J Dev Orig Health Dis* [Internet]. Cambridge University Press; 2019 [cited 2020 Jul 9]. p. 306–13. Available from: <https://pubmed.ncbi.nlm.nih.gov/30101736/>
4. Exploring a causal role of DNA methylation in the relationship between maternal vitamin B12 during pregnancy and child’s IQ at age 8, cognitive performance and educational attainment: a two-step Mendelian randomization study [Internet]. [cited 2020 Jul 9]. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5703349/>
5. Lee J, Kalia V, Perera F, Herbstman J, Li T, Nie J, et al. Prenatal airborne polycyclic aromatic hydrocarbon exposure, LINE1 methylation and child development in a Chinese cohort. *Environ Int* [Internet]. Elsevier Ltd; 2017 [cited 2020 Jul 9];99:315–20. Available from: <https://pubmed.ncbi.nlm.nih.gov/28027800/>
6. Weaver ICG. Epigenetic programming by maternal behavior and pharmacological intervention. *Nature versus nurture: let’s call the whole thing off.* *Epigenetics* [Internet]. 2007 [cited 2017 Jul 13];2:22–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17965624>
7. Epigenetic and Neural Circuitry Landscape of Psychotherapeutic Interventions - PubMed [Internet]. [cited 2020 Jul 9]. Available from: <https://pubmed.ncbi.nlm.nih.gov/29226124/>

8. Kertes DA, Bhatt SS, Kamin HS, Hughes DA, Rodney NC, Mulligan CJ. BDNF methylation in mothers and newborns is associated with maternal exposure to war trauma. *Clin Epigenetics* [Internet]. Springer Verlag; 2017 [cited 2020 Jul 9];9:68. Available from:
<https://clinicalepigeneticsjournal.biomedcentral.com/articles/10.1186/s13148-017-0367-x>
9. Tobi EW, Slieker RC, Luijk R, Dekkers KF, Stein AD, Xu KM, et al. DNA methylation as a mediator of the association between prenatal adversity and risk factors for metabolic disease in adulthood. *Sci Adv* [Internet]. American Association for the Advancement of Science; 2018 [cited 2020 Jul 9];4:eaao4364. Available from:
<http://advances.sciencemag.org/>
10. Rijlaarsdam J, van IJzendoorn MH, Verhulst FC, Jaddoe VW V., Felix JF, Tiemeier H, et al. Prenatal stress exposure, oxytocin receptor gene (*OXTR*) methylation, and child autistic traits: The moderating role of *OXTR* rs53576 genotype. *Autism Res* [Internet]. John Wiley and Sons Inc.; 2017 [cited 2020 Jul 9];10:430–8. Available from:
<http://doi.wiley.com/10.1002/aur.1681>
11. Viuff AC, Sharp GC, Rai D, Henriksen TB, Pedersen LH, Kyng KJ, et al. Maternal depression during pregnancy and cord blood DNA methylation: findings from the Avon Longitudinal Study of Parents and Children. *Transl Psychiatry* [Internet]. Nature Publishing Group; 2018 [cited 2020 Jul 9];8:244. Available from:
<http://www.bristol.ac.uk/alspac/researchers/>
12. Girchenko P, Lahti J, Czamara D, Knight AK, Jones MJ, Suarez A, et al. Associations between maternal risk factors of adverse pregnancy and birth outcomes and the offspring epigenetic clock of gestational age at birth. *Clin Epigenetics* [Internet]. Springer Verlag; 2017 [cited 2020 Jul 9];9:49. Available from:
<http://clinicalepigeneticsjournal.biomedcentral.com/articles/10.1186/s13148-017-0349-z>
13. Britto PR, Lye SJ, Proulx K, Yousafzai AK, Matthews SG, Vaivada T, et al. Nurturing care: promoting early childhood development [Internet]. *Lancet*. Lancet Publishing Group; 2017 [cited 2020 Aug 11]. p. 91–102. Available from:

<https://pubmed.ncbi.nlm.nih.gov/27717615/>

14. Aboud FE, Yousafzai AK. Global Health and Development in Early Childhood. *Annu Rev Psychol* [Internet]. Annual Reviews Inc.; 2015 [cited 2020 Aug 11];66:433–57. Available from: <http://www.annualreviews.org/doi/10.1146/annurev-psych-010814-015128>

15. Fisher PA, Beauchamp KG, Roos LE, Noll LK, Flannery J, Delker BC. The Neurobiology of Intervention and Prevention in Early Adversity. *Annu Rev Clin Psychol* [Internet]. Annual Reviews Inc.; 2016 [cited 2020 Aug 11];12:331–57. Available from: <http://www.annualreviews.org/doi/10.1146/annurev-clinpsy-032814-112855>

16. O'Donnell KJ, Chen L, MacIsaac JL, McEwen LM, Nguyen T, Beckmann K, et al. DNA methylome variation in a perinatal nurse-visitation program that reduces child maltreatment: A 27-year follow-up. *Transl Psychiatry* [Internet]. Nature Publishing Group; 2018 [cited 2020 Jul 9];8:15. Available from: <https://github.com/MeaneyLab/PRSoS>

17. Barker ED, Walton E, Cecil CAM. Annual Research Review: DNA methylation as a mediator in the association between risk exposure and child and adolescent psychopathology [Internet]. *J. Child Psychol. Psychiatry Allied Discip.* Blackwell Publishing Ltd; 2018 [cited 2020 Jul 9]. p. 303–22. Available from: <https://pubmed.ncbi.nlm.nih.gov/28736860/>

18. Goldstein JA, Norris SA, Aronoff DM. DOHaD at the intersection of maternal immune activation and maternal metabolic stress: A scoping review. *J. Dev. Orig. Health Dis.* Cambridge University Press; 2017. p. 273–83.

19. Herba CM, Glover V, Ramchandani PG, Rondon MB. Maternal depression and mental health in early childhood: an examination of underlying mechanisms in low-income and middle-income countries [Internet]. *The Lancet Psychiatry.* Elsevier Ltd; 2016 [cited 2020 Jul 9]. p. 983–92. Available from: <https://pubmed.ncbi.nlm.nih.gov/27650772/>

20. Fracolli LA, Reticena K de O, de Abreu FCP, Chiesa AM. The implementation of a home visits program focused on parenting: An experience report. *Rev da Esc Enferm* [Internet]. Escola de Enfermagem de Universidade de Sao Paulo; 2018 [cited 2020 Jul 9];52. Available from: <https://pubmed.ncbi.nlm.nih.gov/30156655/>
21. Critério Brasil - ABEP [Internet]. [cited 2020 Jul 9]. Available from: <http://www.abep.org/criterio-brasil>
22. Castro MC, Massuda A, Almeida G, Menezes-Filho NA, Andrade MV, de Souza Noronha KVM, et al. Brazil's unified health system: the first 30 years and prospects for the future [Internet]. *Lancet*. Lancet Publishing Group; 2019 [cited 2020 Aug 11]. p. 345–56. Available from: <http://www.thelancet.com/article/S0140673619312437/fulltext>
23. Pé Rez-Escamilla R, Segall-Corrê AM, Maranha LK, De Fá M, Sampaio A, Marín-Leó L, et al. Community and International Nutrition An Adapted Version of the U.S. Department of Agriculture Food Insecurity Module Is a Valid Tool for Assessing Household Food Insecurity in Campinas, Brazil 1 [Internet]. *J. Nutr.* 2004. Available from: <https://academic.oup.com/jn/article-abstract/134/8/1923/4688851>
24. Quintão S, Delgado AR, Prieto G. Validity Study of the Beck Anxiety Inventory (Portuguese version) by the Rasch Rating Scale Model Estudo de Validade da Versão Portuguesa do Beck Anxiety Inventory Mediante o Modelo Rasch Rating Scale.
25. Gomes-Oliveira MH, Gorenstein C, Neto FL, Andrade LH, Wang YP. Validation of the Brazilian Portuguese Version of the Beck Depression Inventory-II in a community sample. *Rev Bras Psiquiatr* [Internet]. Associação Brasileira de Psiquiatria; 2012 [cited 2020 Jun 6];34:389–94. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1516444612000487>
26. de Onis M, Garza C, Onyango AW, Rolland-Cachera MF. WHO growth standards for infants and young children. *Arch. Pediatr.* Elsevier Masson SAS; 2009. p. 47–53.
27. Madaschi V, Mecca TP. Bayley-III Scales of Infant and Toddler Development: Transcultural Adaptation and Psychometric Properties 1 Article. [cited 2020 Jul 9];26:189–97. Available from: www.scielo.br/paideia

28. Bayley N. BAYLEY-III (BAYLEY SCALES OF INFANT AND TODDLER DEVELOPMENT). 3rd ed. San Antonio, TX: PsychCorp; 2006.
29. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* [Internet]. Oxford University Press; 2014 [cited 2016 Jul 23];30:1363–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24478339>
30. Fortin J-P, Triche TJ, Hansen KD, Hancock J. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. [cited 2020 Jul 9]; Available from: <https://academic.oup.com/bioinformatics/article-abstract/33/4/558/2666344>
31. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, et al. ChAMP: Updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics*. 2017;33:3982–4.
32. Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K, V Lord R, et al. De novo identification of differentially methylated regions in the human genome. *Epigenetics and Chromatin* [Internet]. BioMed Central Ltd.; 2015 [cited 2020 Jul 9];8. Available from: <https://pubmed.ncbi.nlm.nih.gov/25972926/>
33. Broos S, Soete A, Hooghe B, Moran R, Van Roy F, De Bleser P. PhysBinder: improving the prediction of transcription factor binding sites by flexible inclusion of biophysical properties. [cited 2020 Jul 9]; Available from: <http://bioit.dnbr.ugent.be/>
34. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* [Internet]. 2008 [cited 2018 Apr 23];9:559. Available from: <http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-9-559>
35. Langfelder P, Luo R, Oldham MC, Horvath S. Is my network module preserved and reproducible? *PLoS Comput Biol* [Internet]. Public Library of Science; 2011 [cited 2020 Jul 9];7:1001057. Available from: www.ploscompbiol.org
36. Knight AK, Craig JM, Theda C, Bækvad-Hansen M, Bybjerg-Grauholm J, Hansen

CS, et al. An epigenetic clock for gestational age at birth based on blood methylation data. *Genome Biol.* BioMed Central Ltd.; 2016;17.

37. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. Mediation: R package for causal mediation analysis. *J Stat Softw* [Internet]. American Statistical Association; 2014 [cited 2020 Jul 9];59:1–38. Available from: <https://waseda.pure.elsevier.com/en/publications/mediation-r-package-for-causal-mediation-analysis>

38. Dall’Aglia L, Muka T, Cecil CAM, Brammer WM, Verbiest MMPJ, Nano J, et al. The role of epigenetic modifications in neurodevelopmental disorders: A systematic review [Internet]. *Neurosci. Biobehav. Rev.* Elsevier Ltd; 2018 [cited 2020 Jul 9]. p. 17–30. Available from: <https://pubmed.ncbi.nlm.nih.gov/30067938/>

39. Hilton CL, Harper JD, Kueker RH, Lang AR, Abbacchi AM, Todorov A, et al. Sensory responsiveness as a predictor of social severity in children with high functioning autism spectrum disorders. *J Autism Dev Disord* [Internet]. *J Autism Dev Disord*; 2010 [cited 2020 Jul 9];40:937–45. Available from: <https://pubmed.ncbi.nlm.nih.gov/20108030/>

40. Wong CCY, Meaburn EL, Ronald A, Price TS, Jeffries AR, Schalkwyk LC, et al. Methyloomic analysis of monozygotic twins discordant for autism spectrum disorder and related behavioural traits. *Mol Psychiatry* [Internet]. Nature Publishing Group; 2014 [cited 2020 Jul 9];19:495–503. Available from: www.nature.com/mp

41. Low KJ, Stals K, Caswell R, Wakeling M, Clayton-Smith J, Donaldson A, et al. Phenotype of CNTNAP1: A study of patients demonstrating a specific severe congenital hypomyelinating neuropathy with survival beyond infancy. *Eur J Hum Genet* [Internet]. Nature Publishing Group; 2018 [cited 2020 Jul 9];26:796–807. Available from: <https://doi.org/10.1038/s41431-018-0110-x>

42. Poliak S, Peles E. The local differentiation of myelinated axons at nodes of ranvier [Internet]. *Nat. Rev. Neurosci.* European Association for Cardio-Thoracic Surgery; 2003 [cited 2020 Jul 9]. p. 968–80. Available from: <https://pubmed.ncbi.nlm.nih.gov/14682359/>

43. Yamasaki T, Hoyos-Ramirez E, Martenson JS, Morimoto-Tomita M, Tomita S. GARLH Family Proteins Stabilize GABAA Receptors at Synapses. *Neuron* [Internet]. Cell Press; 2017 [cited 2020 Jul 9];93:1138-1152.e6. Available from: [/pmc/articles/PMC5347473/?report=abstract](https://pubmed.ncbi.nlm.nih.gov/27111111/)
44. Song DL, Chalepakis G, Gruss P, Joyner AL. Two Pax-binding sites are required for early embryonic brain expression of an Engrailed-2 transgene. *Development*. 1996;122.
45. Green BB, Kappil M, Lambertini L, Armstrong DA, Guerin DJ, Sharp AJ, et al. Expression of imprinted genes in placenta is associated with infant neurobehavioral development. *Epigenetics* [Internet]. Taylor and Francis Inc.; 2015 [cited 2020 Jul 9];10:834–41. Available from: <http://www.tandfonline.com/doi/full/10.1080/15592294.2015.1073880>
46. Lewis RM, Cleal JK, Hanson MA. Review: Placenta, evolution and lifelong health. *Placenta* [Internet]. Placenta; 2012 [cited 2020 Jul 9]. Available from: <https://pubmed.ncbi.nlm.nih.gov/22205051/>
47. R J, W C, F L, H L. Infant's DNA Methylation Age at Birth and Epigenetic Aging Accelerators. *Biomed Res Int* [Internet]. Biomed Res Int; 2016 [cited 2020 Jul 9];2016. Available from: <https://pubmed.ncbi.nlm.nih.gov/28058257/>
48. Nishi A, Numata S, Tajima A, Zhu X, Ito K, Saito A, et al. De novo non-synonymous TBL1XR1 mutation alters Wnt signaling activity. *Sci Rep* [Internet]. Nature Publishing Group; 2017 [cited 2020 Jul 9];7:1–8. Available from: www.nature.com/scientificreports/
49. Chow J, Jensen M, Amini H, Hormozdiari F, Penn O, Shifman S, et al. Dissecting the genetic basis of comorbid epilepsy phenotypes in neurodevelopmental disorders. *Genome Med* [Internet]. BioMed Central Ltd.; 2019 [cited 2020 Jul 9];11:65. Available from: <https://genomemedicine.biomedcentral.com/articles/10.1186/s13073-019-0678-y>
50. Salpietro V, Dixon CL, Guo H, Bello OD, Vandrovцова J, Efthymiou S, et al. AMPA receptor GluA2 subunit defects are a cause of neurodevelopmental disorders. *Nat Commun* [Internet]. Nature Publishing Group; 2019 [cited 2020 Jul 9];10:1–16.

Available from: <https://doi.org/10.1038/s41467-019-10910-w>

51. Ang G, McKillop LE, Purple R, Blanco-Duque C, Peirson SN, Foster RG, et al. Absent sleep EEG spindle activity in GluA1 (Gria1) knockout mice: relevance to neuropsychiatric disorders. *Transl Psychiatry* [Internet]. Nature Publishing Group; 2018 [cited 2020 Jul 9];8:154. Available from: www.jax.org/strain/019011
52. Scales SJ, Hesser BA, Masuda ES, Scheller RH. Amisyn, a novel syntaxin-binding protein that may regulate SNARE complex assembly. *J Biol Chem* [Internet]. *J Biol Chem*; 2002 [cited 2020 Jul 9];277:28271–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/12145319/>
53. Al-Bassam S, Xu M, Wandless TJ, Arnold DB. Differential trafficking of transport vesicles contributes to the localization of dendritic proteins. *Cell Rep* [Internet]. *Cell Rep*; 2012 [cited 2020 Jul 9];2:89–100. Available from: <https://pubmed.ncbi.nlm.nih.gov/22840400/>
54. Rampello AJ, Prophet SM, Schlieker C. The role of torsin AAA+ proteins in preserving nuclear envelope integrity and safeguarding against disease [Internet]. *Biomolecules*. MDPI AG; 2020 [cited 2020 Jul 9]. Available from: <https://pubmed.ncbi.nlm.nih.gov/32204310/>
55. Strong E, Butcher DT, Singhanian R, Mervis CB, Morris CA, De Carvalho D, et al. Symmetrical Dose-Dependent DNA-Methylation Profiles in Children with Deletion or Duplication of 7q11.23. *Am J Hum Genet* [Internet]. Cell Press; 2015 [cited 2020 Jul 9];97:216–27. Available from: <https://pubmed.ncbi.nlm.nih.gov/26166478/>
56. Duclot F, Kabbaj M. The role of early growth response 1 (EGR1) in brain plasticity and neuropsychiatric disorders [Internet]. *Front. Behav. Neurosci.* Frontiers Media S.A.; 2017 [cited 2020 Jul 9]. Available from: <https://pubmed.ncbi.nlm.nih.gov/28321184/>
57. Kasper B, Brandt E, Brandau S, Petersen F. Platelet Factor 4 (CXC Chemokine Ligand 4) Differentially Regulates Respiratory Burst, Survival, and Cytokine Expression of Human Monocytes by Using Distinct Signaling Pathways. *J Immunol* [Internet]. The American Association of Immunologists; 2007 [cited 2020 Jul

9];179:2584–91. Available from:

<http://www.jimmunol.org/content/179/4/2584><http://www.jimmunol.org/content/179/4/2584.full#ref-list-1>

58. Dieckmann L, Cole S, Kumsta R. Stress genomics revisited: gene co-expression analysis identifies molecular signatures associated with childhood adversity. *Transl Psychiatry* [Internet]. Springer Nature; 2020 [cited 2020 Jul 9];10:1–11. Available from: <https://doi.org/10.1038/s41398-020-0730-0>

59. Song H, Zhang L, Luo Y, Zhang S, Li B. Effects of collagen peptides intake on skin ageing and platelet release in chronologically aged mice revealed by cytokine array analysis. *J Cell Mol Med* [Internet]. Blackwell Publishing Inc.; 2018 [cited 2020 Jul 9];22:277–88. Available from: <https://pubmed.ncbi.nlm.nih.gov/28922541/>

60. Lyall K, Schmidt RJ, Hertz-Picciotto I. Maternal lifestyle and environmental risk factors for autism spectrum disorders. *Int J Epidemiol* [Internet]. 2014 [cited 2018 May 16];43:443–64. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24518932>

61. Rutten BPF, Vermetten E, Vinkers CH, Ursini G, Daskalakis NP, Pishva E, et al. Longitudinal analyses of the DNA methylome in deployed military servicemen identify susceptibility loci for post-traumatic stress disorder. *Mol Psychiatry* [Internet]. Nature Publishing Group; 2018 [cited 2020 Jul 9];23:1145–56. Available from: <http://finzi.psych.upenn.edu/library/minfi/>

62. Kumsta R. The role of epigenetics for understanding mental health difficulties and its implications for psychotherapy research. *Psychol Psychother Theory, Res Pract* [Internet]. Wiley-Blackwell; 2019 [cited 2020 Jul 9];92:190–207. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1111/papt.12227>

TABLES AND FIGURES

Main text (Table 2 is too big to fit here – attached as second file)

Table 1 – Characteristics of mothers and their infants in control and intervention groups.

	Control (n=13)			Intervention (n=16)			
	Baseline						
Mothers	n	%	mean (SD)	n	%	mean (SD)	P valor
Pre-pregnancy BMI (kg/m ²)			21.2(3.9)			22.5(6.7)	0.527
ABEP (score)			18.1 (6.4)			17.6 (5.6)	0.825
Beck Depression Inventory			14.0 (6.9)			13.3 (6.3)	0.804
Beck Anxiety Inventory			10.4 (5.9)			16.1 (9.2)	0.179
Family Income			4.7(1.6)			4.4 (1.4)	0.626
Infant characteristics (at birth)							
Gestational age (weeks)			39.3 (1.2)			39.4 (1.4)	0.989
Birth weight (g)			3,120.3 (365.4)			3,168.2(409.2)	0.728
BMI/age (z-scores)			-0.231 (0.9)			0.109 (0.9)	0.320
Neurodevelopmental outcomes (12 months)							
Sex (female)	6	42.9		8	50.0		0.964
Bayley III domains							
- Cognitive			41.5 (3.0)			44.2 (4.0)	0.025
- Fine Motor skills			28.2 (2.2)			30.1 (2.8)	0.051

- Gross Motor skills	41.2 (4.0)	44.8 (4.1)	0.026
- Expressive language	13.1 (3.6)	15.3 (3.4)	0.111
-Receptive language	12.5 (2.3)	13.5 (4.5)	0.429

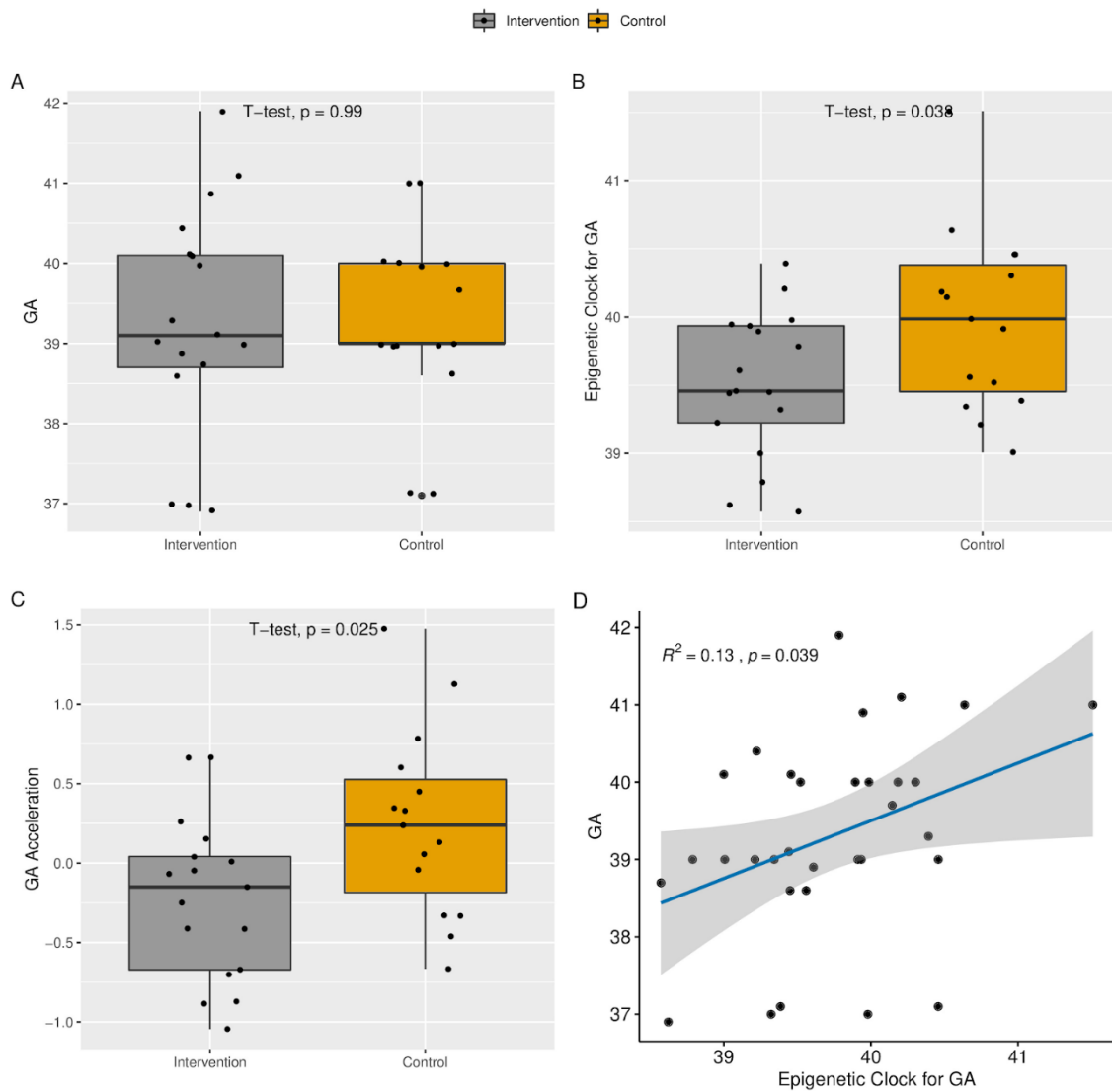


Figure 1. a) Boxplot of the Gestational Age (GA) at delivery, b) Boxplot of the epigenetic clock for GA, c) Age Acceleration.

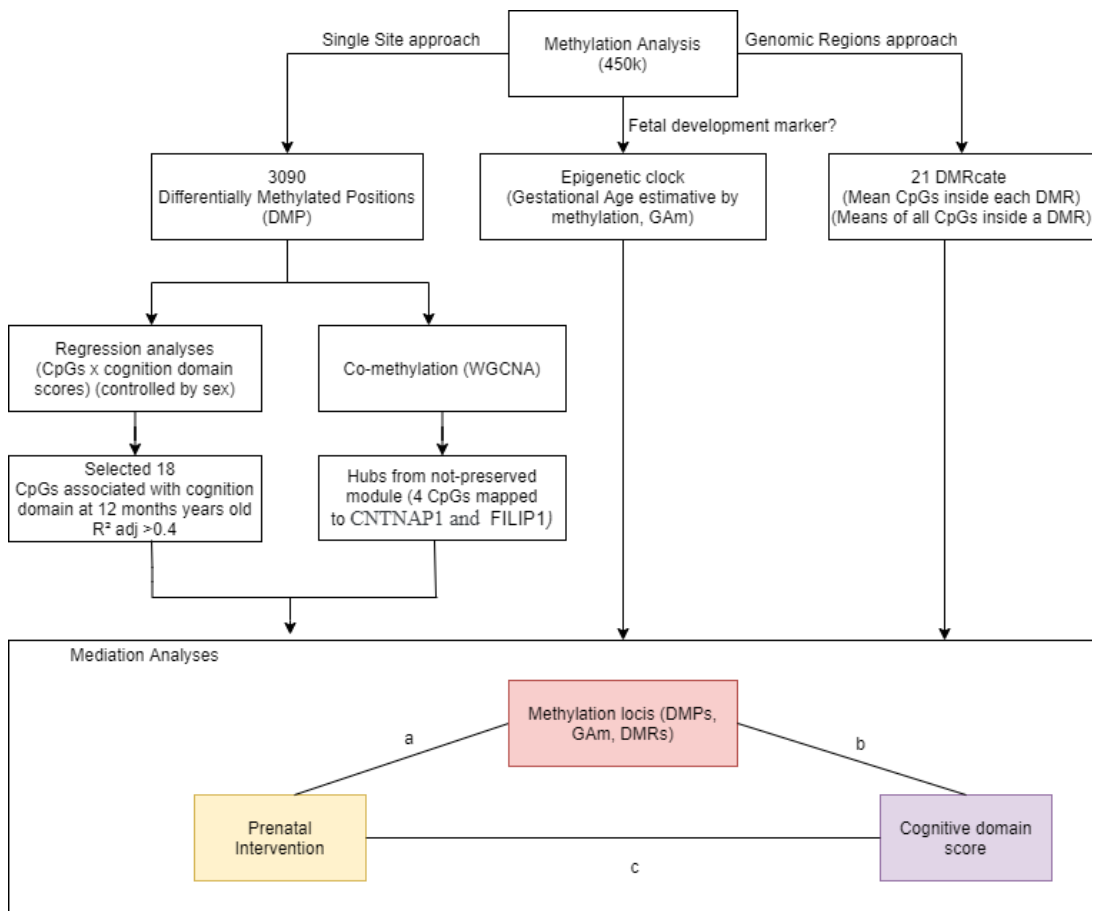


Figure 2. DNAm variables approaches to select variables to be tested in causal mediation analysis.

Supplemental Table 1 – Summary of characteristics comparing the mother/infant who had cord blood collection (n=29) to the sample that this paper comes from (n=80)

	RCT initial sample (n=80)			Cord Blood collection (n=29)			
	Baseline						
	n	%	mean (SD)	n	%	mean (SD)	P value
Pre-pregnancy BMI (kg/m ²)			24.7 (3.6)			25.3 (4.5)	0.718
ABEP (score)			18.2 (5.8)			18.3 (4.4)	0.937
Beck Depression Inventory			11.3 (7.0)			13.6 (3.3)	0.192
Beck Anxiety Inventory			8.9 (6.4)			12.4 (8.0)	0.043
	Neurodevelopmental outcomes (12 months)						
Sex (female)			19 (51.3)			14 (43.8)	0.012
Bayley III domains							
Cognitive			42.7 (3.9)			43.2 (4.6)	0.627
Fine Motor skills			28.9 (2.5)			30.0 (2.6)	0.734
Gross Motor skills			42.3 (5.6)			43.2 (4.1)	0.481
Expressive language			12.7 (2.3)			13.8 (3.6)	0.661
Receptive language			13.8 (3.3)			14.3 (3.2)	0.583

Supplemental Table 2 – Characteristics of mothers and their infants in control and intervention groups from initial sample (n=80), and neurodevelopment outcomes (n=63).

	Control			Intervention			
	Baseline						
Mother	n	%	mean (SD)	n	%	mean (SD)	P valor
Pre-pregnancy BMI (kg/m ²)			22.2 (3.3)			23.0 (4.7)	0.527
ABEP (score)			18.5 (5.8)			17.8 (4.4)	0.825
Beck Depression Inventory			11.7 (7,0)			12.8 (8.0)	0.534
Beck Anxiety Inventory			9.4 (6.4)			11.4 (8,3)	0.244
Infant characteristics (at birth)							
Gestational age (weeks)			39.3 (1.2)			39.3 (1.4)	0.989
BMI/age (z-scores)			-0.24 (0.9)			-0.00 (0.9)	0.457
Neurodevelopmental outcomes (12 months)							
Bayley III domains							
Cognitive			41.6 (2.3)			44.4 (3.6)	0.015
Fine Motor skills			28. 8 (2.6)			30.0 (2.5)	0.075
Gross Motor skills			41.2 (5.5)			44.8 (4.1)	0.089
Expressive language			13.1 (3.7)			15.3 (3.0)	0.111
Receptive language			13.0 (2.3)			15.1 (3.6)	0.011

Supplemental Table 3

Supplemental Table 3A. CpGs with Delta Beta >20%

CpG	adj.P.Val	$\Delta\beta$ Value	CHR	gene	Feature	cgi	WGCNA Module	Previously related to neurodevelopment
cg26191747	1.43E-07	-0.28	10	MSMB	TSS200	opensea	grey	
cg04548204	5.65E-07	-0.22	12	KLRG1	3'UTR	opensea	grey	
cg24891846	9.15E-07	0.24	19	CACNA1A	Body	island	grey	Genome Biol. 2012 Jun 15;13(6):R43
cg04985582	5.53E-06	0.23	21	ABCC13	TSS200	opensea	grey	
cg00413089	6.07E-06	-0.20	7		IGR	island	grey	
cg21177396	6.32E-06	0.30	9		IGR	opensea	grey	
cg12744031	8.97E-06	-0.21	7		IGR	island	grey	
cg11076954	1.20E-05	0.21	17	SLC43A2	Body	shore	yellow	Transl Psychiatry. 2019;9(1):152
cg09329516	2.00E-05	0.23	17	SLC43A2	Body	shore	yellow	Transl Psychiatry. 2019;9(1):152
cg04131969	7.10E-05	0.31	2	MYADML	Body	shore	grey	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC137784/
cg08938978	0.000195	-0.28	2	IL18RAP	Body	opensea	grey	J Psychiatr Res. 2016 Mar;74:10-6.
cg15070894	0.000288	0.20	6	HCG4P6	TSS1500	island	yellow	BMC Med Genomics 12, 105 (2019).
cg04028570	0.000423	-0.25	1	ORZL13	TSS200	island	yellow (HUB)	Front Neurol. 2015; 6: 107.
cg26536949	0.000985	0.26	17		IGR	island	grey	
cg12086633	0.001135	0.25	15		IGR	opensea	grey	
cg05528899	0.001413	0.26	17		IGR	island	grey	
cg01941243	0.008665	0.20	1		IGR	shelf	grey	
cg19384241	0.011159	0.22	2		IGR	opensea	turquoise	
cg24851651	0.019317	0.21	11	CCS	Body	shelf	grey	

Supplemental Table 4. DMPs associated with cognitive domain (adj. R^2 >0.4)

DMP	Gene	adj.P.Val	deltaBeta	CHR	feature	cgi	Enhancer	P.value Reg	adj.R ²
cg01600222		0.002	0.07	6	IGR	shore	NA	1.50E-05	0.49
cg03648020		0.001	-0.07	1	IGR	shelf	NA	4.70E-06	0.54
cg03930153	TBL1XR1	0.003	-0.07	3	5'UTR	opensea	TRUE	6.60E-05	0.43
cg05405872	C11orf58	>0.001	-0.06	11	Body	shore	NA	7.30E-05	0.42
cg06003184	C6orf174	0.003	-0.07	6	3'UTR	opensea	NA	4.30E-05	0.45
cg07761912	TOR1AIP2	0.003	-0.06	1	5'UTR	island	NA	3.00E-05	0.46
cg08174789	MYOM2	0.000	0.07	8	Body	island	NA	4.70E-05	0.44
cg09656541	C10orf118	0.024	0.05	10	5'UTR	shore	NA	9.60E-05	0.42
cg11734401	STXBP6	0.001	0.07	14	5'UTR	shore	NA	6.20E-06	0.52
cg11964314		>0.001	-0.05	12	IGR	opensea	TRUE	3.20E-05	0.45
cg13389575	RB1	0.002	0.09	13	Body	shore	NA	1.10E-04	0.40
cg16652817		0.001	0.06	7	IGR	island	NA	3.50E-05	0.45
cg19610750		0.033	0.06	2	IGR	shelf	NA	1.90E-07	0.63
cg21164545	LOC100240726	0.001	0.06	20	TSS1500	opensea	NA	1.10E-04	0.40
cg22798121	GRIA1	>0.001	-0.05	5	Body	opensea	TRUE	2.00E-05	0.48
cg23223755	LHFPL4	>0.001	0.05	3	TSS1500	shore	NA	5.60E-05	0.43
cg25487135		>0.001	0.07	12	IGR	opensea	TRUE	9.90E-05	0.41
cg26975848		0.001	-0.05	2	IGR	opensea	TRUE	7.10E-05	0.42

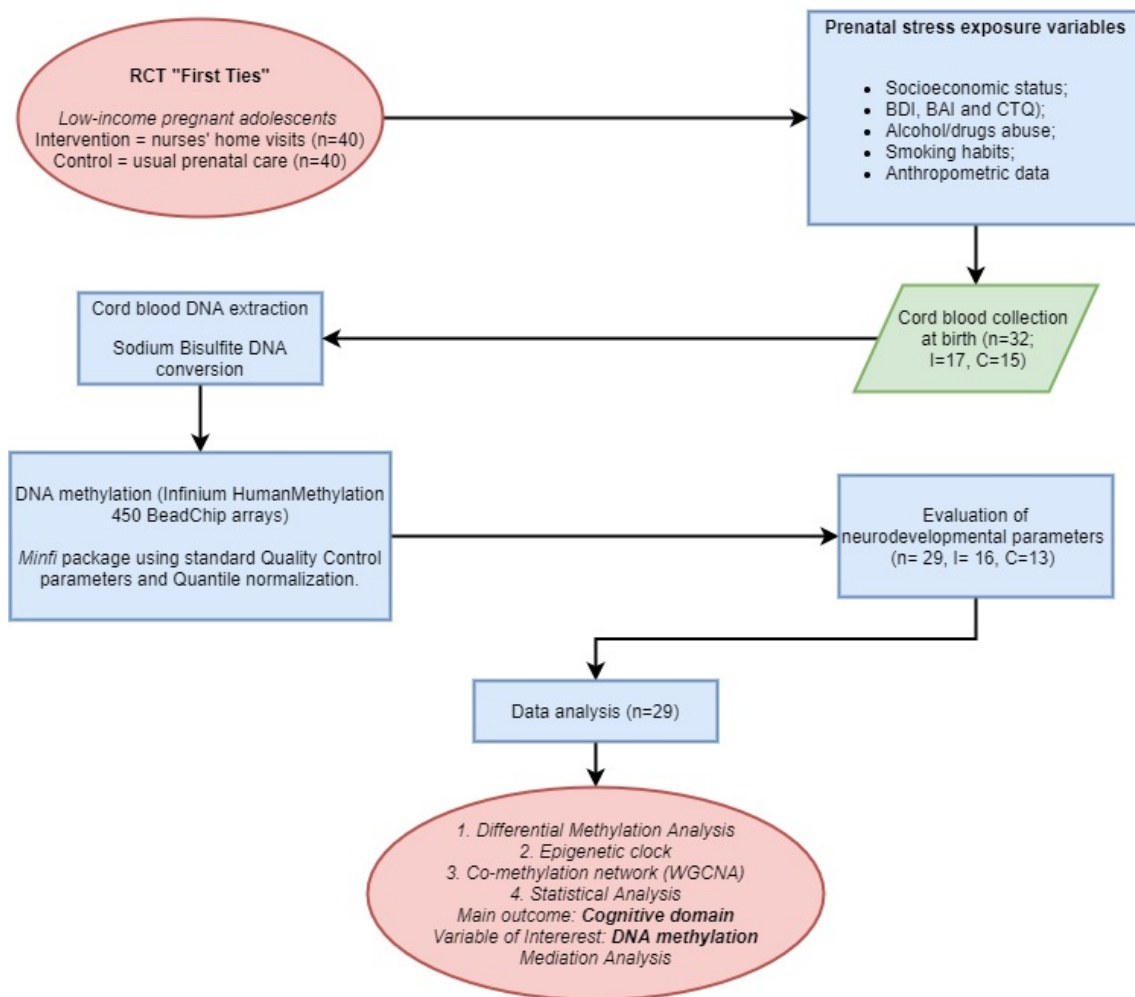
Supplemental Table 5. DMRs identification and respective association with cognitive domain.

Supplemental Table 5. Characterization of DMRs, Mean M-values and association with cognitive domain								
ID	Coordinates	CpGs	Mean M Values	Size	Gene	(SE)	P value	R ² adj.
DMR_1	chr6:33047944-33048919	19	1.7 (0.4)	975	HLA-DPB1	-1.89 (2.9)	0.522	0.08
DMR_2	chr16:8806359-8807043	13	-0.9 (0.3)	684	ABAT	1.6 (3.9)	0.684	0.07
DMR_3	chr7:158749941-158751591	9	0.9 (0.5)	1650		-3.34 (2.5)	0.198	0.13
DMR_4	chr18:14747661-14748439	11	-1.5 (0.3)	778	ANKRD30B	3.35 (3.5)	0.342	0.10
DMR_5	chr17:40822424-40824361	9	-0.5 (0.3)	1937	PLEKHH3	-6.03 (3.1)	0.066	0.19
DMR_6	chr1:248099757-248101009	13	-0.4 (0.9)	1252	OR2L13	-0.67(1,12)	0.559	0.08
DMR_7	chr21:15645649-15646635	7	1.7 (0.4)	986	ABCC13	-2.95 (2,6)	0.265	0.11
DMR_8	chr6:76203225-76203675	7	-0.0 (0.5)	450		3.54 (2,0)	0.096	0.17
DMR_9	chr19:10736006-10736448	8	2.4 (0.3)	442		-1.00 (3.2)	0.759	0.07
DMR_10	chr7:56515510-56516309	10	-0.6 (0.3)	799	LOC650226	-4.15 (2.4)	0.098	0.16
DMR_11	chr10:42862876-42863594	9	-1.7 (0.5)	718	LOC441666	-2.1(1.8)	0.247	0.12
DMR_12	chr2:113992762-113993313	8	1.5 (0.6)	551	PAX8	0.03 (1.5)	0.983	0.07
DMR_13	chr2:198650752-198651576	8	-2.4 (0.5)	824	BOLL	-3.65(1.9)	0.066	0.19
DMR_14	chr11:67383377-67384040	8	-0.0 (0.3)	663		-3.68 (2.3)	0.120	0.15
DMR_15	chr2:105852798-105853796	8	1.8 (0.4)	998		-1.79 (2.1)	0.395	0.09
DMR_16	chr5:178986131-178986906	9	-2.0 (0.7)	775	RUFY1	1.13 (1.1)	0.309	0.10
DMR_17	chr20:5485144-5485511	8	-0.3 (0.4)	367	LOC149837	1.60 (1.9)	0.414	0.09
DMR_18	chr1:247802703-247803166	7	-0.4 (0.5)	463		-2.66 (1.7)	0.122	0.15
DMR_19	chr6:29894050-29894737	7	-1.7 (0.5)	687	HCG4P6	3.71 (1,6)	0.493	0.08
DMR_20	chr4:74847646-74847829	7	-0.3 (0.4)	183	PF 4	-5.48 (1,8)	0.006	0.30
DMR_21	chr1:59043070-59043370	7	-0.6 (0.6)	300	TACSTD2	0.19 (1.3)	0.887	0.07

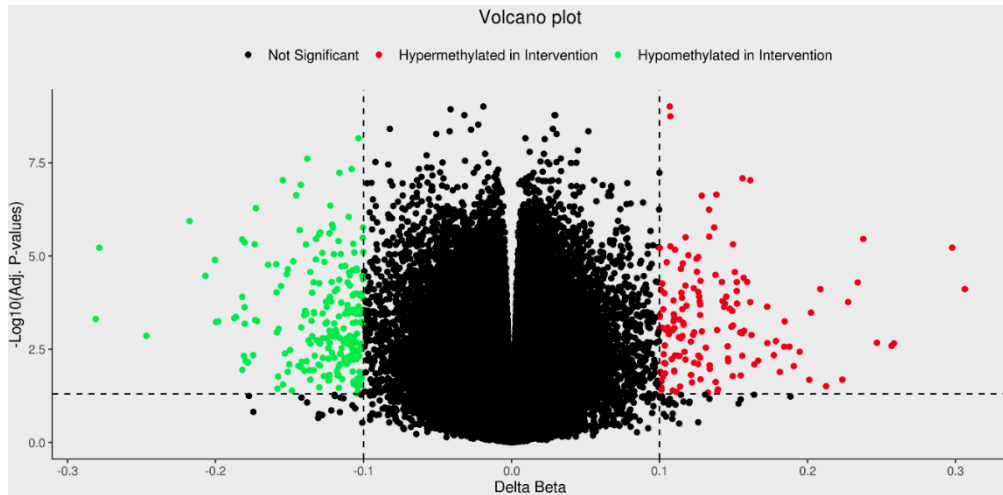
Supplemental Material for WGCNA analysis (co-methylation modules)

Supplemental Table 6. All variables tested in the mediation analysis

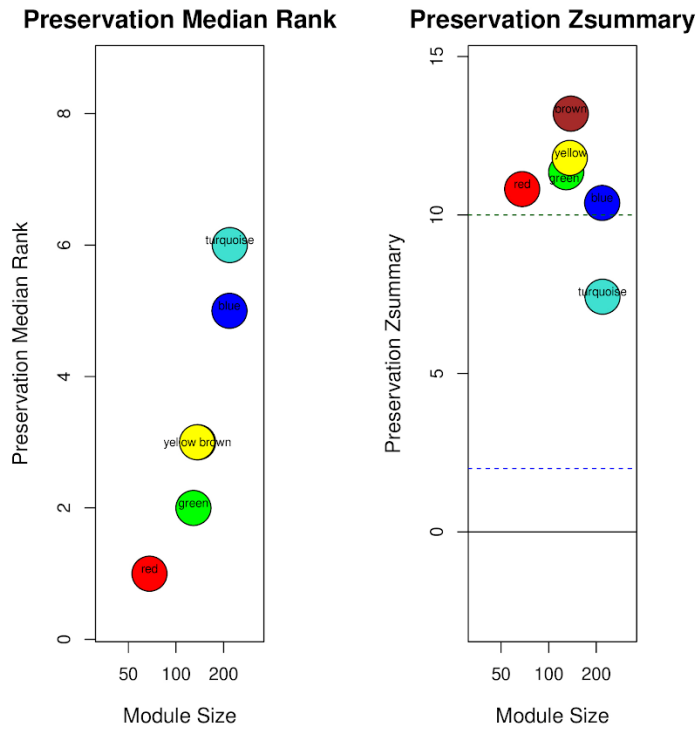
	Gene	Relation island	Feature	Average Direct Effect (ADE)			Average Causal Mediation (ACME)			Total Effect	Prop. mediated		
				Estimate	95%CI	p-value	Estimate	95%CI	p-value	Estimate	p-value		
DMP													
cg01600222		shore	IGR	-0.34	-3.28	2.76	0.810	3.5	1.41	6.05	<2e-16	3.16	0.040
cg03648020		shelf	IGR	-0.64	-3.39	1.88	0.620	3.81	1.74	6.6	<2e-16	3.17	0.036
cg03930153	TBL1XR1	opensea	5'UTR	-0.02	-2.97	2.87	0.980	3.07	3.07	6.08	<2e-16	3.05	0.050
cg05405872	C11orf58	shore	Body	-2.00	-5.42	1.48	0.260	5.15	2.55	8.23	<2e-16	3.15	0.022
cg06003184	C6orf174	opensea	3'UTR	0.09	-2.92	3.07	0.990	3.03	1.1	5.45	0.002	3.11	0.050
cg07761912	TOR1AIP2	island	5'UTR	-0.19	-2.88	2.20	0.890	3.33	1.33	5.89	<2e-16	3.13	0.040
cg08174789	MYOM2	island	Body	-1.30	-3.80	1.23	0.310	4.37	2.16	7.01	<2e-16	3.06	0.040
cg09656541	C10orf118	shore	5'UTR	0.70	-1.53	3.24	0.580	2.29	0.63	4.34	0.010	2.99	0.030
cg11734401	STXBP6	shore	5'UTR	-0.72	-3.07	1.63	0.560	3.79	1.78	6.11	<2e-16	3.07	0.040
cg11964314		opensea	IGR	-2.36	-5.67	0.79	0.150	5.51	3.1	8.29	<2e-16	3.15	0.030
cg13389575	RB1	shore	Body	0.38	-2.44	3.36	0.820	2.76	1.04	4.89	<2e-16	3.15	0.050
cg16652817		island	IGR	-0.40	-3.20	2.47	0.760	3.54	1.42	6.17	<2e-16	3.13	0.050
cg19610750		shelf	IGR	-0.02	-2.26	2.27	0.970	3.12	0.81	5.75	0.010	3.10	0.050
cg21164545	LOC100240726	opensea	TSS1500	-0.96	-4.06	2.44	0.570	4.13	1.78	6.95	<2e-16	3.16	0.040
cg22798121	GRIA1	opensea	Body	-2.01	-5.36	1.43	0.240	5.02	2.2	8.37	<2e-16	3.00	0.060
cg23223755	LHFPL4	shore	TSS1500	-1.13	-4.21	1.86	0.460	4.25	2.04	6.93	<2e-16	3.12	0.050
cg25487135		opensea	IGR	-3.41	-7.78	0.86	0.110	6.45	3.49	9.98	<2e-16	3.04	0.040
cg26975848		opensea	IGR	-0.03	-2.81	2.74	0.100	3.14	1.26	5.57	<2e-16	3.11	0.040
Co-methylation module hubs													
cg08860498	CNTNAP1	island	Body	1.00	-3.60	5.33	0.688	2.19	-1.60	6.10	0.294	3.15	0.298
cg12569246	FILIP1	opensea	TSS200	5.85	1.51	10.13	0.004	-2.74	-6.00	-0.11	0.038	3.11	0.064
cg26466508	FILIP1	opensea	TSS200	4.41	-1.29	9.76	0.11	-1.35	-5.67	3.12	0.508	3.05	0.504
cg18729664	FILIP1	opensea	TSS200	5.45	1.12	9.69	0.02	-2.45	-5.47	-0.04	0.050	3.00	0.110
DMR													
chr2:198650752-198651576	BOLL			0.61	-2.88	4.21	0.748	2.56	0.25	5.61	0.042	3.16	0.078
chr7:56515510-56516309	LOC650226			4.69	1.09	7.77	0.008	-1.55	-3.78	-0.11	0.036	3.15	0.066
chr4:74847646-74847829	PF4			1.41	-1.04	4.05	0.284	1.75	0.15	4.01	0.022	3.16	0.046
GAm													
				2.81	-0.63	6.34	0.106	0.29	-1.87	2.85	0.776	3.108	0.790



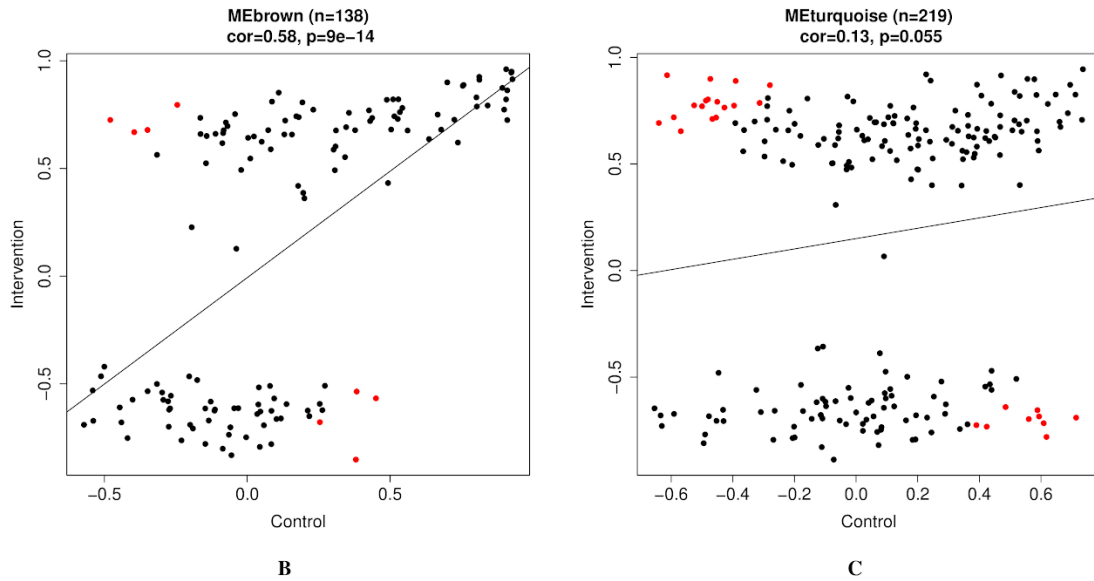
Supplemental Figure 1. Flowchart of the study. Cord blood samples (n=32) were collected from the RCT First Ties (n=80). Neurodevelopmental parameters were obtained by following the Bayley III evaluation. Genome-wide methylation was evaluated in Infinium HumanMethylation 450k BeadChip arrays following quality control, normalization and correction for batch effects steps. Downstream analysis composed by a) Differential Methylation Analysis, b) Epigenetic Clock and c) DNAm variables (DMPs, DMRs, epigenetic clock and co-methylation network) were tested if they presented mediational effect of the relationship between Intervention and neurodevelopmental outcomes.



Supplemental Figure 2. Gráficos de Vulcão mostrando os níveis de metilação de sítios CpG em grupos de intervenção e controle. Os eixos X e Y exibem, respectivamente, o valor delta-beta (diferenças de metilação entre os grupos) e o log dos valores p ajustados para cada sítio CpG. Pontos verdes e vermelhos identificam os locais CpG com $\text{adjP} < 0,05$ e $> 5\%$ de diferença de metilação entre os grupos de intervenção e controle.



A



B

C

Supplemental Figure 3. Module Preservation Analysis. **a)** Preservation Median Rank (left) and Preservation Zsummary (right) of the modules based on the 3,090 DMPs as given by WGCNA. X axis shows the size (number of CpGs) and the y axis shows respectively, the Median Rank and Zsummary scores. Modules with a Zsummary ≤ 10 were considered to be less preserved. Modules with higher Median Rank scores are the ones with low preservation according to this metric. In relation to the Median Rank, the MEbrown had the exact same score as MEyellow (3) with a very similar module size (138 CpGs and 136 CpGs) respectively. The scatterplot on **b)** and **c)** shows the Pearson correlation between each CpGs M-value with its respective module eigengene, for control and intervention groups. Samples with the red color are the ones with the largest Pearson correlation differences between the two groups.

Declarations:

Ethics approval and consent to participate

Consent for publication

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

Competing interests

The authors declare that they have no competing interests.

Funding

This work is supported by Grand Challenges Canada (GCC), Fundação Maria Cecília Souto Vidigal, Bill & Mellinda Gates Foundation, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 Liga Solidária.

Authors' contributions

Acknowledgements

We thank all study participants and the research team members for their valuable contributions.

ANEXO 4

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

Single-cell analysis of human trophoblast stem cell specification reveals activation of fetal cytotrophoblast expression programs including coronavirus associated host factors and human endogenous retroviruses

Authors:

Ethan Tietze^{*a}, Andre Rocha Barbosa^{*a,b,c}, Veronica Euclides^{*a,b,h}, Hyeon Jin Cho^a, Yong Kyu Lee^a, Arthur Feltrin^{a,d}, Joyce van de Leemput^{a,\$}, Pasquale Di Carlo^{a,i}, Tomoyo Sawada^{a,b}, Kynon J. Benjamin^{a,b,f}, Helena Brentani^{c,h}, Joel E. Kleinman^{a,f}, Thomas M. Hyde^{a,b,f}, Daniel R. Weinberger^{a,b,e,f,g}, Gianluca Ursini^{a,e}, Ronald McKay^{a,e}, Apua C.M. Paquola^{a,b}, Joo Heon Shin^{a,b}, and Jennifer A. Erwin^{#a,b,e}

^a Lieber Institute for Brain Development, Baltimore, Maryland, USA

^b Department of Neurology, Johns Hopkins School of Medicine, Baltimore, Maryland, USA

^c Inter-institutional Graduate Program on Bioinformatics, University of São Paulo, São Paulo, SP, Brazil

^d Center for Mathematics, Computation and Cognition, Federal University of ABC, Santo André, SP, Brazil

^e Department of Neuroscience, Johns Hopkins School of Medicine, Baltimore, Maryland, USA

^f Department of Psychiatry & Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA

^g Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA

^h Department of Psychiatry, University of Sao Paulo, Medical School, São Paulo, Brazil

ⁱ Department of Basic Medical Science, Neuroscience, and Sense Organs – University of Bari Aldo Moro, Bari, Italy

*Equal contributor

Lead contact and corresponding author: Jennifer.Erwin@libd.org

\$ Present address: Center for Precision Disease Modeling and Division of Endocrinology, Diabetes and Nutrition, Department of Medicine, University of Maryland School of Medicine, 670 West Baltimore Street, Baltimore, MD, 21201, USA

35 **Summary**

36

37 The human placenta is increasingly a focus of research related to early child development and
38 the impact of maternal hyperimmune states. The ability to model human trophoblast disease
39 states from human pluripotent stem cells, the nature of human pluripotent stem cell potency and
40 the mechanisms regulating human trophoblast specification remains poorly understood.
41 Recent work suggests that only the naive state can give rise to trophoblast and that primed
42 iPSC generate mixed amniotic and mesoderm lineages. Here we identify conditions that
43 efficiently drive the specification of primed iPSC to trophoblast, named Trophoblast Stem
44 Cell (TSC). iPSC-derived-TSC share transcriptional, morphological and functional characteristics
45 with human *in vivo* cytotrophoblasts including activation of human endogenous retroviruses,
46 expression of COVID-19 associated host factors and generation of multinucleated
47 syncytiotrophoblasts with a large fusion index. At high densities in 5% O₂, iPSC-derived-TSC form
48 villi-like structures and express extravillous and syncytiotrophoblast proteins HCG- β and HLA-G.
49 Using temporal single cell RNAseq, we define the molecular changes associated with
50 specification under three separate conditions: 1) BMP4, 2) BMP4 and inhibition of WNT, 3)
51 activation of EGF and WNT, inhibition of TGF β , HDAC and ROCK signaling (named TSC).
52 With 9,821 high-quality single cell transcriptomes, we find that BMP4 gives rise to mesenchymal
53 cells while TS conditions lacking exogenous BMP4 generate a stable proliferating cell type that
54 is highly similar to six week placenta cytotrophoblasts. TFAP2A programs the specification of
55 primed iPSC cells to TSC without transitioning through a naive state. TSC specification
56 independent of exogenous BMP4 will allow for robust and reproducible studies of the
57 cytotrophoblast component of human placenta.

58

59 **Introduction**

60

61 The first cellular specification occurs early in human development when blastomeres divide
62 asymmetrically to generate the outer trophoblast and the inner cell mass of the pre-
63 implantation embryo. Upon implantation, primitive syncytial trophoblast invades and remodels
64 the maternal decidual environment before specifying into the cytotrophoblast and
65 syncytiotrophoblast layers found in placenta villi. Currently, there is a limited understanding of
66 the early developmental trajectories of human trophoblast differentiations (Rossant, 2018).
67 Human models are essential to understanding how the human placenta is built. Placental
68 structures and interaction of placental cells with the maternal environment vary tremendously
69 across mammals. Comparative genome analysis demonstrates that genes involved in placenta
70 and reproduction are amongst the most divergent between human and mouse (Mouse Genome
71 Sequencing Consortium et al., 2002). The evolutionary tinkering between species is driven by
72 endogenous retroviruses (ERVs) ((Johnson, 2019)). Now comprising 8% of the human genome,
73 human endogenous retroviruses are retroviruses acquired by ancestral humans which then
74 integrate into the genome and are passed onto the next generation (Lander et al., 2001). ERV-
75 derived proteins have been repeatedly co-opted for placental functions to promote cell-cell
76 fusion and formation of syncytial trophoblast by endogenous retrovirus group W member 1
77 (*ERVWE1* or *syncytin-1*) (Blond et al., 1999, 2000; Cornelis et al., 2013) and *ERVFRD-1* (Blaise
78 et al., 2003). ERVs harbor regulatory sequences that function as enhancers to control placental

79 gene expression, such as a primate-specific ERV which functions as a placenta-specific
80 enhancer for corticotropin-releasing hormone (CRH), a hormone linked to the control of birth
81 timing in humans (Blond et al., 2000; Dunn-Fletcher et al., 2018). Therefore, human cellular
82 systems are crucial to understand the many human and primate-specific properties of human
83 placenta.

84
85 Human induced pluripotent stem cells are powerful models to understand causal genetic
86 mechanisms of human disease. For organs such as the brain, rapid progress has led to iPSC
87 models with increasing complexity and utility (Mansour et al., 2020). For the human placenta,
88 which is increasingly implicated as a risk factor for developmental disorders associated with a
89 maternal hyperimmune state, the ability to faithfully model human trophoderm disease states
90 from human pluripotent stem cells is a matter of debate. Several model systems have been
91 developed to differentiate primed and naive iPSC to trophoderm. Soon after the derivation of
92 the first human stem cells, Bone Morphogenetic Protein 4 (BMP4) was shown to differentiate
93 human stem cells to syncytial cytotrophoblasts (Xu et al., 2002). Since then, various BMP4-
94 based methods have been described including the BAP protocol (BMP4, A83-01 an alk-5
95 inhibitor and PD173074 and FGF and VEGF inhibitor)(Roberts et al., 2018) and a combination
96 of BMP4 and IWP2 (Horii et al., 2019), an inhibitor that blocks the biosynthesis of mature WNT
97 proteins (Chen et al., 2009). Primed iPSC differentiations have almost exclusively included the
98 use of BMP4, despite the fact that single cell sequencing of the peri-implantation human embryo
99 found the BMP signaling pathway and SMAD protein signal transduction to be enriched in
100 primitive endoderm, not trophoderm (Blakeley et al., 2015; Petropoulos et al., 2016; Stirparo
101 et al., 2018). BMP4 induces differentiation of mesendoderm of the embryo proper (Goldman et
102 al., 2009; Schneider et al., 2003; Yang et al., 2008; Zhang et al., 2008) and also extraembryonic
103 endoderm (Vallier et al., 2009) in addition to upregulating some genes expressed in
104 trophoblasts. TB differentiations using these model systems generate heterogeneous cells with
105 mesodermal contamination and fail to maintain a trophoblast stem cell population (Bernardo et
106 al., 2011; Drukker et al., 2012; Erb et al., 2011; Yabe et al., 2016). Models inconsistently
107 manipulate morphogens including BMP and WNT to generate trophoblast-like cells that are
108 characterized with a limited set of genes. Recent single cell analyses of the early placenta have
109 characterized the expression of these genes in multiple placental cell types (Liu et al., 2018;
110 Suryawanshi et al., 2018; Vento-Tormo et al., 2018). As many of these genes are either
111 expressed in many different cell types or not detected at all, their findings call into question the
112 confidence with which specific iPSC derived placental cell types can be identified. Therefore, it is
113 a matter of debate which pluripotent states can give rise to trophoderm and how similar these
114 trophoderm cells are to *in vivo* placenta.

115
116 Understanding if and how induced pluripotent stem cells can generate human trophoderm is
117 central to understanding human pluripotent stem cell potency and the mechanisms regulating
118 human trophoderm specification. In the present study, we first ask if BMP4 is required for
119 primed pluripotent stem cell specification to trophoderm. In the absence of BMP4, we find
120 that activation of wingless/Integrated (Wnt) and epidermal growth factor (EGF) while inhibiting
121 Transforming Growth Factor Beta (TGF β) histone deacetylase (HDAC) and Rho-associated
122 protein kinase (ROCK) efficiently specify multiple primed pluripotent stem cell lines to

123 trophoctoderm. These factors were previously discovered to derive and grow trophoblast stem
124 cells from blastocysts and the early placenta (Okae et al., 2018). We demonstrate that iPS-
125 derived TS cells are capable of self-renewal for at least 30 passages, differentiate into syncytial
126 cytotrophoblasts and villus cytotrophoblasts and generate villi-like structures in low oxygen. We
127 used temporal single-cell sequencing analysis to elucidate the transcriptional landscape of this
128 specification compared to previously implemented protocols involving BMP4. We find that BMP4
129 conditions alone largely generate extraembryonic mesoderm cells. BMP4 combined with
130 inhibition of WNT shifts differentiation away from a mesodermal fate while upregulating
131 trophoblast associated lineage markers but cells are transcriptionally distinct from *ex vivo*
132 cytotrophoblasts. Trophoblast cells generated by TS conditions are transcriptionally highly
133 similar to *ex vivo* cytotrophoblasts and uniquely upregulate endogenous retroviral genes.
134 Specification of primed iPS cells to TS occurs without entering a naive state and *TFAP2A* and
135 *ESRRG* are central genes regulating the specification. Defining the molecular mechanisms
136 specifying the trophoctoderm is key to understanding placental associated diseases and
137 improving iPS-derived models.

138

139 **Results:**

140

141 **Derivation of trophoblast stem cells (TS) from primed human embryonic stem cells and** 142 **induced pluripotent stem cells without BMP4**

143

144 In search of alternative methods to specify iPS cells to trophoctoderm, we reasoned that primed
145 human pluripotent stem cells might be capable of specification to trophoblast stem cells directly
146 by activating wingless/Integrated (WNT) and epidermal growth factor (EGF) and inhibiting
147 Transforming Growth Factor Beta (TGF β) histone deacetylase (HDAC) and Rho-associated
148 protein kinase (ROCK). While previously described primed iPS differentiations to trophoctoderm
149 have added BMP4, BMP4 is not expressed in the early pre-implantation trophoblast and it is not
150 known if iPS can differentiate to trophoctoderm in the absence of BMP4 ((Blakeley et al., 2015;
151 Petropoulos et al., 2016; Stirparo et al., 2018). Activation of WNT and EGF and inhibition of
152 TGF β , HDAC and ROCK enable the long-term culture of human trophoblast stem cells derived
153 from primary embryos (Okae et al., 2018). 24 hours after clump passage of the human
154 embryonic stem cell line H1 (Thomson et al., 1998) in primed feeder-free stem cell conditions,
155 ES colonies demonstrated domed tight packing with defined edges characteristic of primed cells
156 (Figure 1A and B). 48 hours post passage, cells were switched directly to TSC media. By
157 brightfield imaging over the subsequent 6 days, we observed that cells proliferated rapidly and
158 adopted a flatter appearance with some cells adopting a cobblestone appearance (Figure 1A-
159 D). After passaging, two morphologically distinct populations emerged. Circular colonies with an
160 epithelial-like appearance could be seen surrounded by phase bright fibroblastic cells (Figure
161 1E,F). The inner epithelial-like cells continued to proliferate and appeared morphologically
162 similar to TS cells derived from primary human embryos (Figure 1G). Immunofluorescence
163 revealed that subpopulations of the colony cells expressed TP63 and/or KRT7 and minimally
164 expressed VIM while surrounding fibroblastic cells strongly expressed VIM (Figure 1H). Over
165 the course of several passages, the epithelial-like cells preferentially expand and reach purity
166 between passage 5-7. These TSC widely express both TP63 and KRT7 (Figure 1I-K,

167 Supplemental Figure 1) without expressing the pluripotency genes SOX2 and NANOG
168 (Supplemental Figure 1). In summary, we find that activating wingless/Integrated (Wnt) and
169 epidermal growth factor (EGF) and inhibiting Transforming Growth Factor Beta (TGF β) histone
170 deacetylase (HDAC) and Rho-associated protein kinase (ROCK) directly in primed iPS lines
171 results in efficient specification of trophoblast stem cells.

172
173 Pluripotent stem cell lines vary in their propensity to differentiate to different cell types (Osafune
174 et al., 2008). Therefore, we asked if TSC conditions without exogenous BMP4 could specify a
175 variety of primed human stem cell lines. We confirmed TS specification by VIM, KRT7 and TP63
176 immunofluorescence on iPS lines reprogrammed from dermal fibroblasts by sendai, named
177 2014.06 (Supplemental Figure 2), and two postmortem dura fibroblast-derived iPS cell lines
178 reprogrammed by episome, named LIBD7c6 and LIBD1c8 (Sawada et al., 2020)(Figure 1 H-K).
179 To confirm the ability to proliferate as TS-like cells, iPS-derived TS lines were passaged for up
180 to 32 passages and maintained KRT7 and TP63 expression (Figure 1K).

181 182 **Hypoxia induces differentiation of iPS-derived TS cells into villi-like structures composed** 183 **of differentiated syncytial cytotrophoblasts and villus cytotrophoblasts**

184
185 To investigate the maturation potential of iPS-derived-TSC, we evaluated the ability of iPS-
186 derived TS cells to differentiate to mature trophoblast cell types. Proliferative trophoblast
187 stem cells are bipotential stem cells with the capacity to differentiate to syncytial
188 cytotrophoblasts and extravillous cytotrophoblasts (Okada et al 2018). Syncytial cytotrophoblast
189 is a terminally differentiated multinucleated epithelial layer that infiltrates the maternal
190 endometrium. The multinucleated cell forms from multiple cytotrophoblast cell fusions. When TS
191 cells were grown to high density, we observed the spontaneous formation of multinucleated
192 cells by brightfield microscopy. We confirmed that multinucleated syncytial cells are Ki-67
193 negative nuclei surrounded by the tight-junction protein ZO-1 positive cells (Figure 2A and B).
194 We observed large areas of syncytial formation fused nuclei (Figure 2A and S3).
195 Syncytiotrophoblasts produce and secrete high levels of human chorionic gonadotropin
196 (hCG β), a hormone essential for cytotrophoblast differentiation, immuno-suppression, and
197 placental invasion (Cole, 2010). We observe expression of hCG β in regions of the
198 differentiated cytotrophoblast cultures, indicating differentiation to syncytiotrophoblasts (Figure
199 2C). Extravillous trophoblasts express human leukocyte antigen G (HLA-G), the main molecule
200 involved in maternal immune system evasion upon placental invasion of the endometrium. We
201 also readily detected regions of HLA-G positive cells, indicating iPS-derived TS cells are
202 capable of differentiation to extravillous cytotrophoblast (Figure 2D). Expression of hCG β and
203 HLA-G was specific to iPS-derived TS cells and not detected in iPS cells (Supplemental Figure
204 1I).

205
206 To further induce differentiation of the iPS-derived-TSC, we tested if cells will respond to altering
207 oxygen concentrations. Hypoxia is a natural cue that stimulates syncytial cytotrophoblast
208 differentiation *in vivo* (Wakeland et al., 2017), therefore we transitioned a high density TS cell
209 culture from atmospheric oxygen of 20% to 5%. In 5% oxygen, these cells demonstrate minimal
210 contact inhibition and form 3D villi-like structures that protrude from the monolayer surface of

211 the culture dish (Figure 2 E-H). When imaged by confocal microscopy, these villi-like structures
212 are approximately 50 microns in height. In summary, iPS-derived-TSC readily differentiate into
213 both HLA-G positive villus cytotrophoblasts and hCGB3 positive syncytial cytotrophoblasts.

214

215 **Single cell RNA sequencing reveals unique transcriptional programs for iPS-derived TS** 216 **specification**

217

218 To define the molecular events involved in the specification of iPS cells to trophoctoderm, we
219 performed temporal single cell RNAseq analysis of human iPS cells specified to placental
220 trophoblast cells by three separate conditions: 1) addition of BMP4 (Xu et al., 2002), 2) addition
221 of BMP4 and inhibition of Wnt with the small molecule inhibitor IWP2 (Horii et al., 2019), 3)
222 Trophoblast Stem Cell condition (TSC) activation of EGF and Wnt, inhibition of TGFbeta, HDAC
223 and ROCK signaling to generate trophoblast stem cells (Figure 3A). Sequencing was generated
224 for iPS cells before differentiation (day 0), for BMP4 on days 4 and 6, for BMP4+IWP2 on days
225 2, 4 and 6, and on days 2, 4, 6 and 8 for TS condition, resulting in 10 single-cell transcriptomes
226 (Figure 3A). In TS condition, the day 8 is a timepoint two days after the first passaging
227 annotated as p1. In the BMP4 and BMP4 + IWP2 conditions, no proliferative cells survived
228 passaging, therefore the p1 time point was not sequenced. Using the highly parallel droplet
229 based single cell sequencing method Drop-Seq (Macosko et al., 2015), 9,821 high-quality cells
230 were obtained after removing cells with less than 1000 genes detected and more than 20% of
231 mitochondrial mapping. The average ratio of mapping reads for mitochondrial genes in all cells
232 were 3.9%, indicating good viability (Supplemental Figure 4A). The total number of genes
233 detected ranged between 21804 for BMP4 + IWP2 day 4 to 27134 for TS day 4 (Supplemental
234 Figure 4B) time point with between 563 to 1572 high quality cells per timepoint (Supplemental
235 Figure 4C).

236

237 In order to explore the data in an unbiased manner, we performed dimensionality reduction for
238 the most variable genes across all cells with Principal Component Analysis (PCA) and Uniform
239 Manifold Approximation and Projection (UMAP) embedding. Generally, single cells separated by
240 differentiation stage and condition, with UMAP dimension 2 generally capturing the
241 differentiation condition (Figure 3B). To identify transcriptionally similar groups of cells, we
242 performed a graph-based clustering analysis of k nearest neighbors using the euclidean
243 distance of the 30 first PCs in Seurat (Stuart et al., 2019). We identified 19 clusters, and clusters
244 contained mostly a single time-point and media condition (Figure 3B-D and S4D and Table S1).
245 IPS clusters 1, 2 and 3 mostly contained IPS cells with slight contributions from the first days of
246 all three differentiation conditions. Cells from the BMP4+ IWP-2 condition were found
247 predominantly in a group of five heterogeneous clusters termed BR (BMP responsive). BR1 was
248 composed of roughly equal parts iPS cells and BMP4 + IWP-2 day 2 treated cells. BMP4+IWP2
249 cells made up the majority of the cells in the other BR clusters alongside a smaller population of
250 BMP4 cells. BR2 and BR3 were predominantly derived from BMP4+IWP2 day 2 cells, while day
251 4 contributed to BR4, and day 6 contributed to BR. Cells from the BMP4 condition almost
252 exclusively formed four clusters termed SC (Stromal cell). SC1 and SC3 were predominantly
253 composed of cells from day 4 of the differentiation while SC2 and SC were predominantly from
254 day 6 cells. In addition, cluster SC contained a small contribution of cells originating from the

255 final differentiation day of the TS condition termed TSP1 (TS passage 1). Cells from the TS
256 differentiation almost exclusively populated 7 clusters termed TS. TS1 was exclusively
257 composed of Day 2 cells and TS 2-5 were composed of a mixture of cells from days 2-6. The
258 final TS day (TSP1) separated into two clusters named TS and TS6, which confirms the
259 brightfield observations of 2 cell populations. Interestingly, 14 cells from BMP4+IWP2 d6 and 4
260 cells from the BMP4 d6 contribute to the final TS cluster (Figure 3D, Table S1).

261
262 To understand how transcriptional profiles change during the differentiation, we asked which
263 cells express canonical pluripotent, trophoblast and mesoderm specific genes. We find that
264 the stem cell transcription factor *SOX2* is highest expressed in the iPS stage and is absent from
265 the most differentiated BR, TS and SC clusters (Figure 3E). *SOX2* expression is maintained at a
266 higher level in the initial days of the TS conditions compared to BMP WNTi and BMP treatment
267 and absent from the most differentiated TS cells. The Paired-like homeodomain transcription
268 factor 2 *PITX2* is highly expressed in the later stages of BMP4 differentiation, suggesting that
269 BMP4 conditions generate placental stromal cells (Figure 3F). In single cell RNAseq of the
270 human placenta, *PITX2* was identified as the top marker gene for a stromal fetal communicating
271 cell type F1 by Vento-Tormo (Liu et al., 2018; Vento-Tormo et al., 2018). Keratin 7, *KRT7*, and
272 human chorionic gonadotropin subunit alpha, *CGA*, are expressed in both stem cell derived
273 cytotrophoblasts and *in vivo* placenta (Yabe et al., 2016). Indicating specification to
274 trophoblast transcriptional programs, we find both *KRT7* and *CGA* are expressed in the later
275 stages of the BMP+ WNTi condition and TS condition (Figure 3G-H). Overall, both TS and
276 BMP4+ WNTi treated cells demonstrated an increased expression for most TB markers
277 including *GATA3* and *TFAP2C* (Figure S5A) while the BMP4 alone conditions demonstrated a
278 more mesoderm associated expression of *GATA4*, *TBXT* and *PDGFRA* (Figure S5B). In mouse,
279 the Caudal-type homeobox 2 (*CDX2*) is involved in the segregation of the ICM and
280 trophoblast lineages at the blastocyst stage by repressing *OCT4/POU5F1* and *NANOG* in
281 the TE. Overexpression of *CDX2* in mouse embryonic stem cells causes differentiation to
282 trophoblast stem cells (Erwin et al., 2012; Niwa et al., 2005). In humans, *CDX2* expression is
283 initiated after blastocyst formation and has variable expression patterns in trophoblast
284 (Niakan and Eggan, 2013). Similar to TS cells derived from primary human embryos, *CDX2* is
285 not expressed in the differentiated TS state, and there is sparse expression of *CDX2* at d4 of all
286 differentiations (Figure S5A).

287
288 Next we asked which genes are preferentially expressed in each cluster by non-parametric
289 Wilcoxon rank sum test (adj.p-value < 0.05; logFC > 0.25)(Figure 3I, Table S2). We defined cell-
290 population markers by comparing gene expression of each cell population compared to all other
291 cells. *DUSP6* and *THY1* are primed stem cell state genes and are among the top marker genes
292 identified in the iPS cells. The SC cluster marker genes includes WNT signaling genes *RSPO2*
293 a WNT agonist (Kazanskaya et al., 2004) and *PITX2* which interacts with both WNT signaling
294 and regulates collagen expression during development (Hendee et al., 2018). Markers of the BR
295 cluster include *PRSS23* a serine protease involved in Snail-dependent Epithelial to
296 Mesenchymal Transition (Chen et al., 2013) and regulators of Insulin-like growth factor signaling
297 *IGFBP7* (Oh et al., 1996) and *PAPPA2* (Overgaard et al., 2001). Markers of the TS clusters
298 include genes involved in EGFR signaling *NPNT* (Arai et al., 2017), *IFI6* (Meyer et al., 2015),

299 and *MUC15* (Wang et al., 2013). Additionally, *ARHGAP18* is a Rho GTPase activating protein
300 that functions as a downstream effector of YAP signaling (Porazinski et al., 2015) recently
301 implicated in the specification HiPS to bipotent trophoblast stem cells (Mischler et al., 2019).

302
303 In sum, TSC, BMP4 and BMP4 with inhibition of WNT pattern different cell types. Activation of
304 WNT and EGF and inhibition of TGFb, HDAC and ROCK specify primed iPSC to
305 cytotrophoblasts, BMP4 alone induces mesenchymal fates and inhibition of WNT in the context
306 of BMP4 generates a small percentage of TS cells. For use by the community, we provide a
307 web browser of the single cell expression profiles
308 (<http://erwinpaquolalab.libd.org/publications/singlecellTS>) and the differentially expressed genes
309 for all pairwise cluster comparisons (Table S3).

310

311 **iPS-derived trophoblast stem cells resemble first trimester placental cytotrophoblasts** 312 **while BMP4 differentiation generates placental mesenchymal cells**

313

314 We next asked how similar are the iPS-derived placental cells to *ex vivo* cells found in human
315 placenta. There is considerable debate about the authenticity of iPS derived trophoblast cells
316 and primary trophoblast stem cells, therefore we set out to directly compare the single cell
317 transcriptional profiles to *ex vivo* datasets. To characterize the expression profiles of the 19
318 clusters, we analyzed the overlap with cell type-specific genes identified by two published
319 single-cell studies of early human placenta (Liu et al., 2018; Vento-Tormo et al., 2018). We
320 performed gene set enrichment analyses using Cell-Specific Expression Analysis (CSEA) (Xu et
321 al., 2014). We identified cell type specific genes that were preferentially expressed in each cell
322 type/cluster of the 19 clusters from iPS-derived cells by the specificity index probability (pSI)
323 statistic at thresholds of $p < 0.05$. Next, we tested whether cell-type specific genes previously
324 identified by single cell studies of human placenta are over-represented in the cell-type specific
325 genes from our iPS specification by hypergeometric test and applied the Bonferroni correction
326 for multiple comparisons, considering all the tested gene lists [$\alpha = 0.05 / (19 \times (38+14)) = 5.1 \times 10^{-5}$].
327 Vento-Tormo identified a set of 38 cell types from maternal (decidua and blood) and fetal
328 tissues (placenta) of first trimester placenta and reported 30 cell-type specific genes per cell
329 type (Vento-Tormo et al., 2018). We found that the transcriptional profile from all iPS clusters
330 and clusters from d2-d4 timepoints did not have significant enrichment for any of the placental
331 cell type clusters (Figure S6A). In contrast, the most mature iPS-derived TS cluster was highly
332 enriched for the syncytial cytotrophoblast (SCT) and villous cytotrophoblast (VCT) cluster
333 ($p = \log_{10}^{-32}$ and \log_{10}^{-17}) (Figure 4A). The expression profile of the iPS-derived TS cluster was
334 highly specific for SCT and VCT with no significant enrichment for any other cell-type in the
335 maternal fetal interface. The D6 BMP4 cluster SC was most enriched in the two fibroblast
336 clusters (F1 and F2) ($p = \log_{10}^{-26}$ and $p = \log_{10}^{-13}$). Vento-Tormo described these placental
337 fibroblasts as mesenchymal stromal cells of fetal origin that derive from the primitive endoderm
338 expressing GATA4, GATA6, PDGFRa, and SOX17 (Figure S6 and S7). The BMP4 + IWP2
339 conditions (BR) were not strongly enriched for any specific placental cell type and instead had
340 weaker enrichment for several clusters including syncytial cytotrophoblast (SCT), and maternal
341 decidua derived F1 endodermal cell type and dS1. In sum, iPS-derived TS cells are highly

342 enriched for expression of genes specific to syncytial cytotrophoblast and villous cytotrophoblast
343 of *ex vivo* human placenta.

344
345 Next we compared the expression profiles to early placental dataset from the 8 week and 24
346 week placenta (Liu et al., 2018) (Figure 4B). Again, we found strong enrichment for the iPS-
347 derived TS cells in the eight week fusion competent CTB1 (cytotrophoblast, \log_{10}^{-9}). BMP4 alone
348 and BMP4 + IWP2 conditions were enriched for two different mesoderm cells identified by Liu et
349 al (\log_{10}^{-23} and \log_{10}^{-4}). To visualize placental cell-type specific expression patterns among the
350 iPS-derived cell types, we plotted a heatmap of the top unique markers of the fetal cell types
351 identified in Vento-Tormo and polar trophoderm markers identified in (Petropoulos et al.,
352 2016) (Figure 4C). Consistent with the Gene Set Enrichment Analysis, we observed high
353 expression for a majority of unique marker genes previously identified in polar trophoderm
354 villus cytotrophoblasts (VCT) and syncytial cytotrophoblasts (SCT). We also compared the
355 expression profiles of all iPS-derived cellular clusters to previously identified lineage and marker
356 genes from pre-implantation datasets (Stirparo et al 2018, Figure S7A). Again, we found the
357 highest expression of trophoblast genes KRT18, TEAD3, GATA3 and GATA2 in iPS derived TS,
358 with moderate expression in BR cells. Primitive endoderm genes SOX17, GATA4, GATA6 were
359 highest expressed in SC cells. When compared to primary trophoblast cells (Suryawanshi et al
360 2019, Figure S7B), we confirmed that TS cells expressed the highest level of cytotrophoblast
361 and EVT specific genes. In summary, iPS cells specified via TS conditions have a
362 transcriptional profile highly homologous to cytotrophoblasts taken directly from human placenta
363 while BMP4 differentiation generates placental mesenchymal cells.

364 365 **Primed iPS specify to trophoblast by activation of TFAP2A without activation of naive** 366 **stem cell programs**

367
368 It has been reported that primed stem cells are restricted in their potency and unable to
369 differentiate to TSC while naive stem cells readily differentiate to TSC because they are poised
370 to specify to trophoderm (Dong et al., 2020). Therefore, we asked if the primed iPS or TSCM
371 treated cells adopt a naive stem cell program during differentiation. For 12 previously identified
372 naive marker genes (Messmer et al., 2019), naive transcripts were nearly absent. Five genes
373 HORMAD1, ALPPL2, KHD3L, TRIM60, and HYAL4 had no observed expression in any cell
374 (Figure 5A, Table S4). Less than 1% of cells for any condition had detectable expression for six
375 other marker genes ALPP, OLAH, LYZ, MEG8, KDC1L, and FAM151A. ZNF729 was the only
376 gene expressed in 0.51% of TSp1 cells to 14.69% of IWP2d2. In contrast, primed marker genes
377 were consistently expressed in iPS clusters and often throughout the differentiation state. For
378 example, DUSP6 was detected in 74.87% of the iPS cells, indicating that the iPS initiated in the
379 primed state. In sum, primed iPS transition to cytotrophoblasts without transitioning through a
380 naive state.

381
382 We asked how BMP signaling contributes to the differentiation by evaluating BMP family
383 signaling in all samples. Low percentages of cells expressed BMP4 in TS conditions, while
384 BMP4 was significantly upregulated in several cell clusters including Day 4 BMP4 and BMP4 +
385 IWP2 treated cells compared to iPS and TS cell clusters (Figure 5B, Table S4). BMP5 and

386 BMP7 were expressed in a larger portion of TSCM treated cells compared to iPS (BMP5: 0-
387 0.67% iPS vs 24.72% TS1 cells and BMP7: 1.01-4.71% iPS vs 29.06% TS1 cells). The BMP
388 effector proteins SMAD1,4, and 5 is expressed in the first days of all conditions and is absent
389 from the mature TS cluster. We conclude that BMP5 and BMP7 signaling patterns iPS-derived
390 TS specification instead of BMP4.

391
392 To identify the most significant gene expression changes associated with the initial stages of TS
393 specification, we studied a subset of the sample using only cells from the iPS and TS day 2
394 samples and applied a single-cell trajectory inference and pseudotime estimation (STREAM) to
395 the single cell transcriptome data (Chen et al., 2019). Briefly, single cells were ordered along
396 probabilistic trajectories and a numeric value referred to as pseudotime was assigned to each
397 cell to indicate how far it progresses along the dynamic differentiation. STREAM identifies an
398 initial branch composed of both iPS and TS day2 cells (S3-S0). This level of analysis also
399 identified branches containing iPS cells (S0-S2, S0-S4), reflecting the pluripotency continuum,
400 as well as a TS day 2 (S0-S1) branch reflecting the exit from pluripotency (Figure 5C and Table
401 S5). On branch S3-S0, node S3 was predominantly composed of iPS cells while TS day 2 cells
402 are more abundant towards S0 (Figure 5C, S8A,B). We next identified the transcriptional states
403 that contribute to the specification by identifying highly expressed genes within a branch that are
404 correlated with the transition along the specification branches (see methods). Top positively
405 correlated transition genes along branch S3-S0 include *LRRN3* which amplifies MAPK signaling
406 through EGFR (Fukamachi et al., 2002), *MEIS2* which codes for a TALE homeobox protein that
407 promotes tumor progression through the combined action of two isoforms increasing YAP
408 nuclear translocation and activating the WNT/b-catenin pathway (Guan et al., 2019), *BMP5*,
409 *BMPR1B* and *TFAP2A* a transcription factor key to the suppression of pluripotency and
410 expression of trophoblast associated genes (Krendl et al., 2017) (Figure 5C, S8C). These genes
411 were expressed not only along branch S3-S0 but also along branch S0-S1 (Figure S8C). The
412 branch S0-S1 captures the transition out of the pluripotent state and into the TS state with
413 upregulation of cytokeratins (*KRT18*, *KRT19*, *KRT8*) and calcium binding proteins (*S100A11*,
414 *S100A10*) known to be expressed in trophoblast (Abd El-Aleem and Dekker, 2018; Liu et al.,
415 2006) (Figure S8D). Also found as a top transition gene, *TMSB4X* is implicated in stemness of
416 progenitor trophoblast cells of first trimester human placentas by increasing NOTCH1 activity
417 (Huang et al., 2016; Lee et al., 2018; Lv et al., 2013). In summary, trophoblast transcriptional
418 programming occurs early in the specification and utilizes previously identified trophoblast
419 associated regulatory factors including *TMSB4X*, *YAP*, *BMP5*, and *TFAP2A*.

420
421 **STREAM pseudotime analysis of trophoblast differentiations reveals upregulation of**
422 **stromal mesenchyme upon BMP4 treatment and trophoblast transcriptional**
423 **programming upon TS treatment.**

424
425 To understand the transcriptional changes associated with the full trajectory of iPS to
426 trophoblast specification, we repeated the STREAM analysis with the complete iPS to d8
427 dataset. STREAM identifies 4 main branching points among all single cells (Figure 6A). The
428 early differentiation days for all conditions were mixed in pseudotime demonstrating a similar
429 transcriptional state at the initial phases of iPS commitment when compared in the context of d6

430 and d8 cells. Similar to the UMAP projection (Figure 3B,C), STREAM inferred mixing of two iPS
431 stages with cells from the d2 treatments. STREAM inferred three main branching points which
432 separated based on the most differentiated cells from each condition. To reveal the
433 transcriptional states that contribute to the branching, we carried out differential expression
434 analysis to identify highly expressed genes within a branch that are differentially expressed
435 compared to all other branches, termed leaf gene detection analysis in STREAM (see methods)
436 (Figure 6B, Table S6). For the zero IPS branch point, pluripotent-specific genes including
437 *DNMT3A* and *POU5F1/OCT4* were highly differentially expressed. For the branch of cells
438 exposed to BMP4 (SC), mesoderm markers including *PITX2* were identified as differentially
439 expressed. For the BMP+IWP2 branch, placental associated genes were identified including
440 *RAB31*, a small GTPases highly expressed in placenta that regulates intracellular membrane
441 trafficking of ligand bound *EGFR* (Chua and Tang, 2014). For the TS branch, significantly
442 differentially expressed genes were associated with trophoctoderm commitment and
443 implantation and preeclampsia including *GATA3*. We performed a Gene Ontology enrichment
444 analysis for biological processes using the top 300 significantly differentially expressed leaf
445 genes for the TS branch. Significantly overrepresented GO categories included placental
446 development, female pregnancy, viral processes (Figure 6 C,D).

447 448 **Human endogenous retrovirus derived genes participate in a regulatory subnetwork** 449 **specifically within iPS-derived trophoblast stem cells**

450
451 We found the set of genes involved in viral-like processes are upregulated in the later TS
452 differentiation timepoints (Figure 6C). This echoes an increasing body of evidence
453 demonstrating that endogenous retroviruses play a fundamental role in the evolutionary
454 diversification of the mammalian placenta (Johnson, 2019). Therefore, we asked what is the
455 expression of human ERV-derived genes during trophoctoderm specification? We found specific
456 upregulation of five ERV-derived genes in the most differentiated TS cell type (Figure 7A).

457
458 To understand the place of ERV-derived genes in the regulatory TS regulatory network, we
459 analyzed the participation of ERV-derived genes in a gene regulatory network. We built a
460 transcription factor and target gene network model using the Passing Attributes between
461 Networks for Data Assimilation (PANDA) algorithm to all clusters among the conditions. PANDA
462 integrates information from TF-sequence-motif data, gene expression and protein-protein
463 interaction (PPI) in a message-passing approach (Glass et al., 2013). Briefly, PANDA's
464 algorithm makes the assumption that genes targeted by the same TF are more likely to be co-
465 expressed than genes that are regulated by a different TF. Using a TF-sequence motif as initial
466 network model, the algorithm refines and updates the model by computing the congruence
467 between two measures: 1) Concordance in co-expression of genes presenting motif for a same
468 TF, and 2) Concordance in the regulatory profile of a set of TFs that interact with each other via
469 Protein-protein interaction data. Therefore, the likelihood of interaction between a TF and a
470 target gene is given by its edge weight, which states for the coordination of targets' expression
471 by a TF complex. PANDA outputs Z-score normalized edge weights, therefore positive and
472 negative weights can be seen as likely and unlikely interactions respectively. One network was

473 created for each previously identified cell cluster based on the expression profile of the top 100
474 cells in each cluster (based on PC1) with genes expressed in at least 10% cells.
475 We found that ERV interactions (Transcription Factors + ERV-derived genes) are prominent in
476 the most differentiated TS cluster and in no other cell clusters. We found regulatory interactions
477 for five ERV-derived genes (*ERVH48-1*, *ERV3-1*, *ERVMER34-1*, *ERVW-1* and *ERVFRD-1*) are
478 highly unique to the TS cluster network, while only 0 or 1 ERV-derived gene was found in the
479 other cell clusters (Figure 7B, S9A). Clusters SC and SC 3 (BMP4 cells) exhibited interactions
480 for gene *ERVH48-1*, and clusters BR 1, BR 3 and BR 4 (BMP4+IWP2 cells) presented the gene
481 *ERV3-1*. Interestingly, genes encompassed in ERV regulatory interactions of iPS derived TS
482 cells were enriched in biological pathways related to hormone metabolism, cell differentiation
483 and the immune system (adj. p value < 0.01) (Figure 7C, Table S7). In addition, we observed
484 ERV regulatory interactions are also enriched for categories such as placenta development,
485 trophoctodermal cell differentiation and syncytium formation (Table S7). To understand the
486 significance of the ERV-subnetwork interactions, we next asked what is the strength of ERV
487 edges in the network context of the TS cluster? We found that weights of ERV regulatory
488 interactions were highly ranked among all network connections (Wilcoxon Rank Sum
489 pval=1.6778e⁻³¹, permuted p value=0.001). In sum, we find that ERV-derived genes participate
490 in a regulatory subnetwork specifically within TS cells.

491

492 **Coronavirus associated host factors are expressed in iPS-derived TS cells and could** 493 **potentially model cytotrophoblast involvement in infection**

494

495 We next asked if these iPS-derived cells could be relevant to model the cytotrophoblast
496 involvement in infection. Infections of pregnant women can severely impact both the developing
497 fetus and pregnant women. Trophoblast cells at the maternal-fetal interphase play a crucial role
498 in the impact of these infections. Syncytiotrophoblasts are the barrier that protects the fetus from
499 some infections (Parry et al., 1997). The permissiveness of trophoblast cells to a specific
500 infection largely determines if an infection reaches the fetus. Pathogens such as *Toxoplasma*
501 *gondii*, Rubella, herpesvirus, cytomegalovirus, and Zika virus can cross the placenta, infect the
502 fetus and cause congenital disease ([Stegmann and Carey, 2002](#); [Coyne and Lazear, 2016](#)). It is
503 still too soon to know the complete impact of the severe acute respiratory syndrome coronavirus
504 2 (SARS-CoV-2) and the associated coronavirus disease 2019 (COVID-19) on pregnant women
505 and fetus. But it has been established that the placenta can be infected by SARS-CoV-2
506 (Vivanti et al., 2020). Emerging data suggest that placental SARS-CoV-2 infections are rare
507 (Vivanti et al., 2020), perhaps because the CoV-2 host entry substrate angiotensin-converting
508 enzyme 2 (ACE2) and the serine protease TMPRSS2 are minimally expressed in the placenta
509 (Pique-Regi et al., 2020). Although ACE2 and TMPRSS2 co-expression are rare in the maternal
510 fetal interface, alternative entry receptors such as BSG (Aguiar et al., 2020). When considering
511 expression analysis of confirmed and potential SARS-CoV-2 host factors, Singh et al reported
512 that the cytotrophoblast is the cell type most susceptible to coronavirus infection in the first
513 trimester placenta while confirming that *ACE2* and *TMPRSS2* are lowly expressed (Singh et al.,
514 2020).

515

516 To evaluate the potential tropism of SARS-CoV-2 in iPS-derived TS cells, we profiled SARS-
517 CoV-2 and coronavirus associated factors. Overall, iPS-derived TS cells expressed coronavirus
518 associated factors in similar patterns compared to *ex vivo* cytotrophoblasts (Singh et al., 2020),
519 (Pique-Regi et al., 2020). *ACE2* and *TMPRSS2* are minimally expressed (Figure 8A, Table S8)
520 and putative alternative entry receptors *BSG*, *HSPA5* and *DPP4* are more expressed. Other
521 potential entry proteases *CTSB*, *CTSL* and *FURIN* are all expressed at a higher level (Figure
522 8B, Table S8). The placenta also protects the fetus from infection by the active transfer of
523 maternal antibodies to the fetus. Syncytiotrophoblasts transfer protective humoral maternal IgG
524 to the fetus via the neonatal Fc receptor encoded by the *FCGRT* gene (Leach et al., 1996).
525 Consistent with the findings that iPS-derived TS cells are transcriptionally similar to
526 cytotrophoblasts and syncytiotrophoblasts, iPS-derived TS cells also express *FCGRT*. In sum,
527 iPS-derived TS cells are highly transcriptionally similar to *ex vivo* cytotrophoblasts and
528 syncytiotrophoblasts with respect to placental host factors involved in infection. Therefore, iPS-
529 derived TS cells may be useful to model the involvement of human cytotrophoblasts in infection.

530

531 Discussion

532

533 Here we demonstrate that primed iPS cells are specified to trophoblasts without BMP4 and
534 instead with EGF, TGF- β inhibitors, a WNT activator and Valproic acid induced HDAC inhibition.
535 iPS-derived TS cells are highly transcriptionally similar to cytotrophoblasts found in human
536 placenta. TS cells proliferate rapidly, display very variably sized nuclei, readily become
537 multinucleated as observed in early post-implantation trophoblast and strongly express
538 extravillous and syncytiotrophoblast associated proteins (HLA-G, HCG- β) in sub populations.
539 After eight days of specification, iPS-derived TS cells display a uniquely active transcriptional
540 network of human endogenous retroviruses similar to *in vivo* trophoblast. We demonstrate that
541 in hypoxic conditions, iPS-derived TS cells create 3D villi-like structures, conditions which will
542 likely be useful for future 3D organoid models.

543

544 The ability of primed human stem cells to differentiate into true trophoblast lineages is under
545 debate. Here we demonstrate that primed human stem cells retain a broad potency to generate
546 extra embryonic lineages which are transcriptionally highly similar to placenta. Our results are
547 consistent with the model that the primed state is on a continuum of potency, rather than a fixed
548 and restricted state (Weinberger et al., 2016). Methodological differences and the frequent use
549 of BMP4 likely underlie previously reported differences in the ability of primed and naive
550 pluripotent stem cells to generate TSCs. We confirm that the BMP4 model can generate cells
551 with key features of trophoblast, but in the defined conditions used here BMP4 specified cells
552 are less similar to *in vivo* cytotrophoblasts. Following the recent derivation of human TSCs from
553 primary cells (Mischler et al., 2019; Okae et al., 2018), several attempts have been made to
554 derive TSCs from ES/iPS from a range of pluripotent states. Mischler et al. 2019 used primed
555 ES cells with a hybrid protocol that combined a three day treatment of BMP4 and SB431542
556 with long term culture in TM4 media (CYM554 a S1PR3 agonist, CHIR99021, A83-01, and
557 FGF10) to generate a CDX2+ TSC-like population (Mischler et al., 2019; Okae et al., 2018).
558 When transferred to TSCM media, these cells lost CDX2 expression and adopted
559 characteristics of TSCs. Omitting TM4 media and directly transferring BMP4 treated cells to

560 TSCM resulted in a gradually enriching population of epithelial cells comparable with their
561 CDX2- TSC-like population. Interestingly, this process mirrors the differentiation progression we
562 observed when we directly treated primed iPS with TSCM. Our data suggest that BMP4 is not
563 involved in this transition. TSC-like cells have also been generated from primed iPS cells using
564 a micromesh culture system with primed pluripotent stem cell media followed by culture in
565 TSCM (Li et al., 2019) although the comparison of these cells to *in vivo* placental cells is less
566 extensive. In contrast, Dong et al. 2020 reported that only naive and not primed ES could
567 generate TSC-like cells by direct exposure to TSCM (Dong et al., 2020). Human expanded
568 potential stem cells (EPSCs) can also generate TSC-like cells when differentiated in TSCM
569 (Gao et al., 2019). Interestingly, this group was unable to generate TSC-like cells using naive or
570 primed ES cells under similar conditions. Here, we demonstrate that the continuum of
571 pluripotency present in the primed state readily allows for specification to trophoblast.

572
573 The robust differentiation of primed iPS cells to trophoblast allows for modeling of placental
574 diseases from patient-specific cells without first trimester primary tissue. Placental adaptive
575 responses to insults can alter fetal development and program the fetus for later disease
576 including diabetes and neurodevelopmental disorders (Myatt, 2006). As both the protective
577 barrier and a component of the innate immune system, cytotrophoblasts and their progeny play
578 important roles in the outcome of insults (Guleria and Pollard, 2000). Recent evidence that
579 Covid19 leads to a hyperimmune state has implications for placental biology and fetal
580 development, potentially analogous to the effects of other hyperimmune states during
581 pregnancy (Al-Haddad et al., 2019). Genetic variation among individuals influences how a
582 pathogen affects the fetus. iPS models of this genetic variation could be harnessed for
583 therapeutic intervention. We anticipate that iPS-derived TS cells will be a powerful model to
584 study the placental contribution to disease and the genetic regulation of human placental
585 evolution.

586 587 **Acknowledgements**

588 Supported by funding from the Lieber Institute for Brain Development and a NARSAD Young
589 Investigator Grant from the Brain & Behavior Research Foundation to J.A.E. The authors are
590 grateful for the financial support of the Coordenação de Aperfeiçoamento de Pessoal de Nível
591 Superior (CAPES) - Financing code 001 (Andre Barbosa, Veronica Euclides and Arthur Feltrin).
592 We thank Bill Ulrich for the single cell RNAseq browser.

593 594 **Declaration of Interests**

595 The authors declare no competing interests.

596

597

598 **Figures and Figure Legends**

599

600 **Figure 1: Derivation Of Trophoblast Stem Cell Like (TS) Cells From Human Pluripotent**
601 **Stem Cells.**

602 **A.** Schematic of protocol used to derive TS cells from human pluripotent stem cells (PSCs) by
603 activating wingless/Integrated (WNT) and epidermal growth factor (EGF) and inhibiting
604 Transforming Growth Factor Beta (TGF β) histone deacetylase (HDAC) and Rho-associated
605 protein kinase (ROCK).

606 **B-G.** Representative phase contrast images of differentiating TS cells (ES line H1). (B) Images
607 of differentiating cells on days 2 (B), 4 (C), and 6 (D) after addition of differentiation (TS) media.
608 **E,F.** Images of TS cell morphology 24 (E) and 72 (F) hours following passage (P1) on day 6. (G)
609 Mature TS lines generated from ES line H1 after 10 passages.

610 **H.** Immunofluorescence for VIM (magenta), KRT7 (red), TP63 (green), and Hoechst 33342
611 (blue) in TS cells after 1 passage (P₀) (iPS line LIBD7c6).

612 **I-K.** Confocal images of KRT7 (red), TP63 (green), and DAPI (blue) for TS cells generated from
613 passage 11 H1 (I) and 2014.06 (J) and passage 32 LIBD1C8 (K).

614 All phase contrast images scale bars indicate 100 μ m. All immunofluorescence images scale
615 bars indicate 50 μ m.

616

617 **Figure 2: iPS-derived TS Cells Differentiate Into Syncytial Cytotrophoblast, Extra Villous**
618 **Cytotrophoblasts and Villi-like Structures.**

619 **A,B.** Immunofluorescent images of mature TS line LIBD1C8 stained for ZO-1 (red), KI-67
620 (green), and DAPI (grey).

621 **C,D.** Immunofluorescent images of mature TS line LIBD1C8 stained for DAPI (red) and hCG (C)
622 or HLA-G (D) (grey).

623 **E-H.** Villi like structures spontaneously form when cultured in 5% O₂. (E) Three dimensional
624 reconstruction of a "villi" from mature TS line 2014.06 using immunofluorescence images of
625 nuclei visualized with Hoechst 33342. Virtual sections at the base (H), midpoint (G), and top (F)
626 of the structure.

627 All scale bars indicate 50 μ m.

628

629 **Figure 3: Temporal Single Cell RNA Sequencing Of iPS To Trophoblast Differentiation**

630 **A.** Experimental schematic of trophoblast differentiation experiment. HiPS cells were plated and
631 allowed to adhere to culture dishes for 24 hours before addition of differentiation media. Cells
632 were collected for single-cell RNA sequencing at indicated time points.

633 **B,C.** Dimensionality reduction for the most variable genes across all cells with Principal
634 Component Analysis (PCA) and Uniform Manifold Approximation and Projection (UMAP)
635 embedding of 9,821 single cell transcriptomes from three distinct media treatments with
636 groupings based on (B) sample identity or (C) gene expression clusters calculated by k nearest
637 neighbors using the Euclidean distance of the 30 first PCs which identifies 19 clusters. Cells
638 from the BMP4 condition almost exclusively formed four clusters termed SC (Stromal cell
639 indicated in blue). Cells from the BMP4+ IWP-2 condition were found predominantly in a group
640 of five heterogeneous clusters termed BR (BMP responsive indicated in purple). iPS indicated in
641 green and TS clusters are indicated in orange.

642 **D.** Plot of the proportion of cells from each differentiation condition which are present in the final
643 k-means clusters TS, TS6, BR and SC. Color represents the differentiation condition scale in 3B
644 BMP4+IWP2: purple, BMP4: blue, iPS: green, TS: orange. Cell proportions for all clusters can
645 be found in supplemental table S1.

646 **E-H.** Single cell visualization of gene expression of selected markers of HiPS cells (E), placental
647 stromal cells (F), and trophoblasts (G,H) on UMAP projections. Grey dots indicate cells with no
648 expression and blue scale indicates normalized expression.

649 **I.** Dot plot displaying differentially upregulated marker genes for the most mature cell clusters
650 from each treatment SC (BMP4), BR (BMP4+IWP2), and TS 6, TS (TS media) compared to the
651 average expression of all other clusters. Size of dot indicates the percent of cells with
652 expression and intensity of the blue indicates the average expression level.

653

654 **Figure 4: Trophoblast Differentiations Are Enriched For First Trimester Placental Cell** 655 **Types**

656 **A,B.** Gene set enrichment analysis of most mature iPS-derived cell clusters (SC (Stromal Cell,
657 BMP4), BR (BMP responsive, BMP4+IWP2), and TS 6, TS (Trophoblast stem cell media) to cell
658 types identified in previous single cell sequencing of first trimester placenta. Cell types

659 **C.** Composite heatmap of normalized gene expression of the top unique markers of placental
660 cell types of fetal origin identified in Vento-Tormo et al. 2018 and polar trophoblast markers
661 (top 50) identified in Petropoulos et al. 2016.

662

663 **Figure 5: Primed iPS Specification To Trophoblast Stem Cells Initiates With TFAP2A** 664 **Rewiring Without Activation Of Naïve Stem Cell Programs**

665 **A.** Dotplot displaying BMP signaling in each cellular cluster organized by the treatment identity
666 of the majority cells. See table S3 for source data.

667 **B.** Dotplot displaying markers of naïve and primed human pluripotent stem cells for BMP4,
668 BMP4 + IWP2, and TS cell differentiations by sample. See table S3 for source data.

669 **C.** STREAM pseudotime trajectory of HiPS differentiation for HiPS and TS D2 cells visualizing
670 cells proportion along the smoothed pseudotime axis (orange represents TS, green represents
671 HiPS). Nodes are labeled with numbers 0-4. Branches are defined as the cells between 2 nodes
672 (Branch 1,4). Bottom: Top ranked (by STAT) genes with expression significantly correlated with
673 the pseudotime transition of the branches. MEIS2, TFAP2A, and less differentiated TS d2 to
674 more differentiated TSd2 (S0-S1) KRT18 (STAT 0.692). Ranks 2-4 for these branches is found
675 in figure S4 C-D. Additional graphs showing the branch assignment with cells and a cell level
676 representation of the pseudotime with the clusters Figure S9 and table S4.

677

678 **Figure 6: STREAM Pseudotime Analysis Of Trophoblast Differentiations Reveals** 679 **Upregulation Of Stromal Mesenchyme Upon BMP4 Treatment and Trophoblast** 680 **Transcriptional Programming Upon TS Treatment.**

681 **A.** STREAM pseudotime trajectory of HiPS differentiation including all samples highlighting the
682 distribution of each previously described cell cluster along pseudotime.

683 **B.** Average gene expression of example genes which are significantly differentially expressed
684 for each branch compared to other branches. S0 initial differentiation branch: 1284 significant
685 genes, POU5F1: p value = 1.032932e-297 by Kruskal-Wallis H-test. S4 branch dominated by BR

686 (BMP4+IWP2 cells): 538 significant genes, RAB3 p value = 0 by Kruskal–Wallis H-test. S2
687 branch. S2 branch dominated by SC (BMP4+ cells): 1395 significant genes, PITX2 p value = 0
688 by Kruskal–Wallis H-test and top ranked. S3 branch dominated by TS cells. 3855 significant
689 genes, GATA3 p value = 0 by Kruskal–Wallis H-test and top ranked gene. The complete list of
690 significant genes see Supplemental Table S5.

691 **C.** Upset plot of shared genes within significant Gene Ontology enrichment terms generated
692 from the top 300 leaf genes for branch 1-3 (specification to TS). Gliogenesis p=7.92e-06,
693 Leukocyte activation involved in immune response p=1.72e-05, Reproductive system
694 development p=4.05e-05, Female pregnancy p=7.95e-05, Viral Process p=.0004, Blood Vessel
695 Morphogenesis p=.003, Placenta Development p=.008

696 **D.** Gene expression of CTSB along the pseudotime trajectory. CTSB is a significant leaf gene
697 for branch 1-3 that is annotated as associated with all GO processes listed in (6C).
698

699 **Figure 7: Endogenous Retrovirus-derived Genes Participate In A Gene Regulatory** 700 **Subnetwork Uniquely In TS Cells.**

701
702 **A.** Dotplot of HERV expression in most mature cell clusters from each treatment and the HiPS
703 cluster 1. Size of ball indicates the percent of cells expressing the gene and color represents the
704 z-normalized expression value.

705 **B.** Cytoscape representation of the top transcription factor partners of ERVs detected in TS and
706 SC. 164 TFs and 6 Targets. See Supplemental table S7.

707 **C.** Visual Representation of GO terms from most significant partners of ERVs shown in panel B
708 ($p < .0001$). Specifics of programs and settings to obtain these results. Cytoscape 3.7.2, cluego
709 v2.5.4, cluepedia version v1.5.4, go term fusion, levels 3-8, kappa .35, graph of highest
710 significance within each group, input of 164 TFs from panel B supplemental table with analysis
711 output and details of GO analysis includes all go terms off the
712

713 **Figure 8: Coronavirus Associated Host Factors Are Expressed In iPS-derived TS Cells**

714
715 Dotplot of Coronavirus associated host factor expression in the iPS-derived TS cell clusters TS
716 and TS6. Size of ball indicates the percent of cells expressing the gene and color represents the
717 average expression value.

718 **A.** Coronavirus associated host factors expressed in 0-10% of cells.

719 **B.** Coronavirus associated host factors expressed in >10% of cells.
720

721 **Supplemental Tables**

722 Table S1: Number of cells from each condition/day that are assigned to the cluster

723 Table S2: Differentially expressed genes in each cluster where a single cluster is compared to
724 all other clusters

725 Table S3: Differentially expressed genes analyzed by pairwise comparisons of one cluster
726 compared to another cluster

727 Table S4: Expression analysis of naive, primed and BMP signaling during differentiation

728 Table S5: Stream pseudotime analysis of transition and differentially expressed leaf genes for
729 iPS to TSd2

730 Table S6: Stream pseudotime analysis of transition and differentially expressed leaf genes for
731 all cells
732 Table S7: ERV gene regulatory network and associated GO enrichment.
733 Table S8: Expression analysis of coronavirus associated factors.
734
735

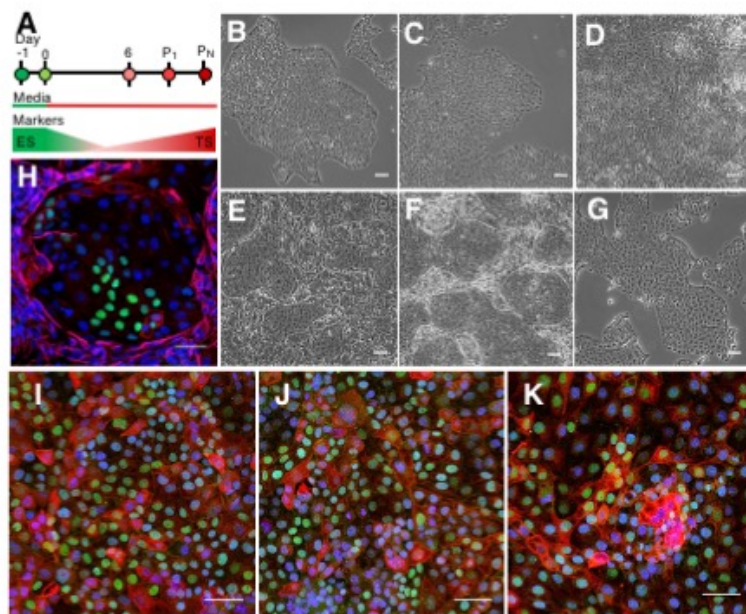


Fig. 1: Derivation of trophoblast stem cell like (TS) cells from human pluripotent stem cells.

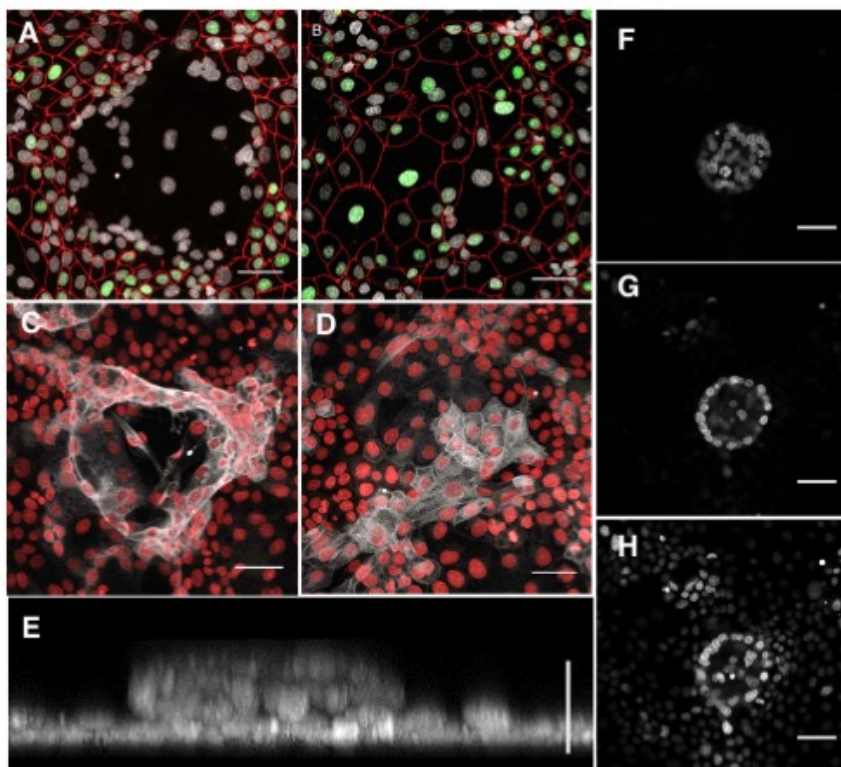


Figure 2: iPS-derived TS cells differentiate into syncytial cytotrophoblast, extra villous cytotrophoblasts and villi-like structures.

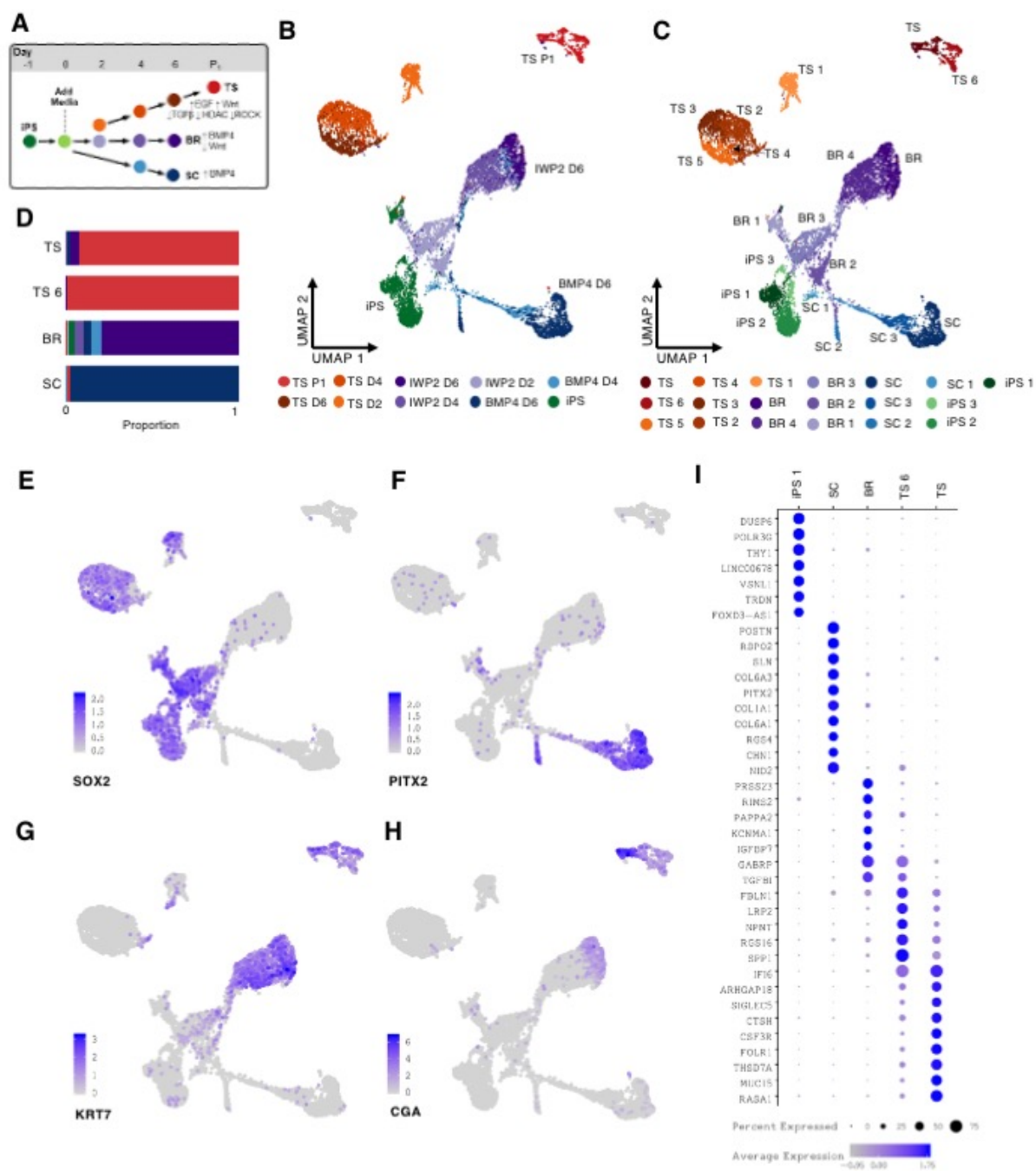


Figure 3: Temporal single Cell RNA Sequencing of iPS to trophoblast differentiation

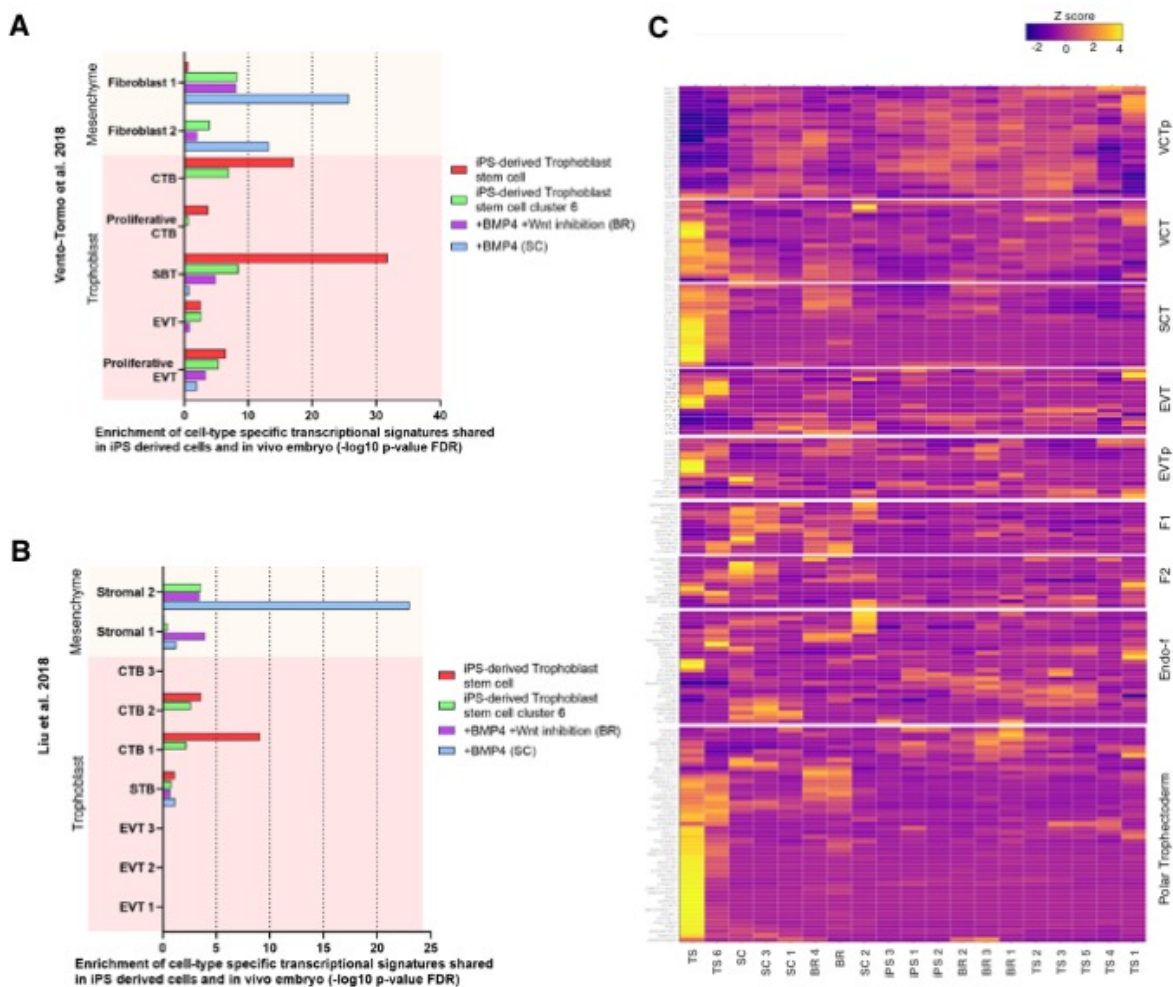


Figure 4: Trophoblast differentiations are enriched for first trimester placental cell types.

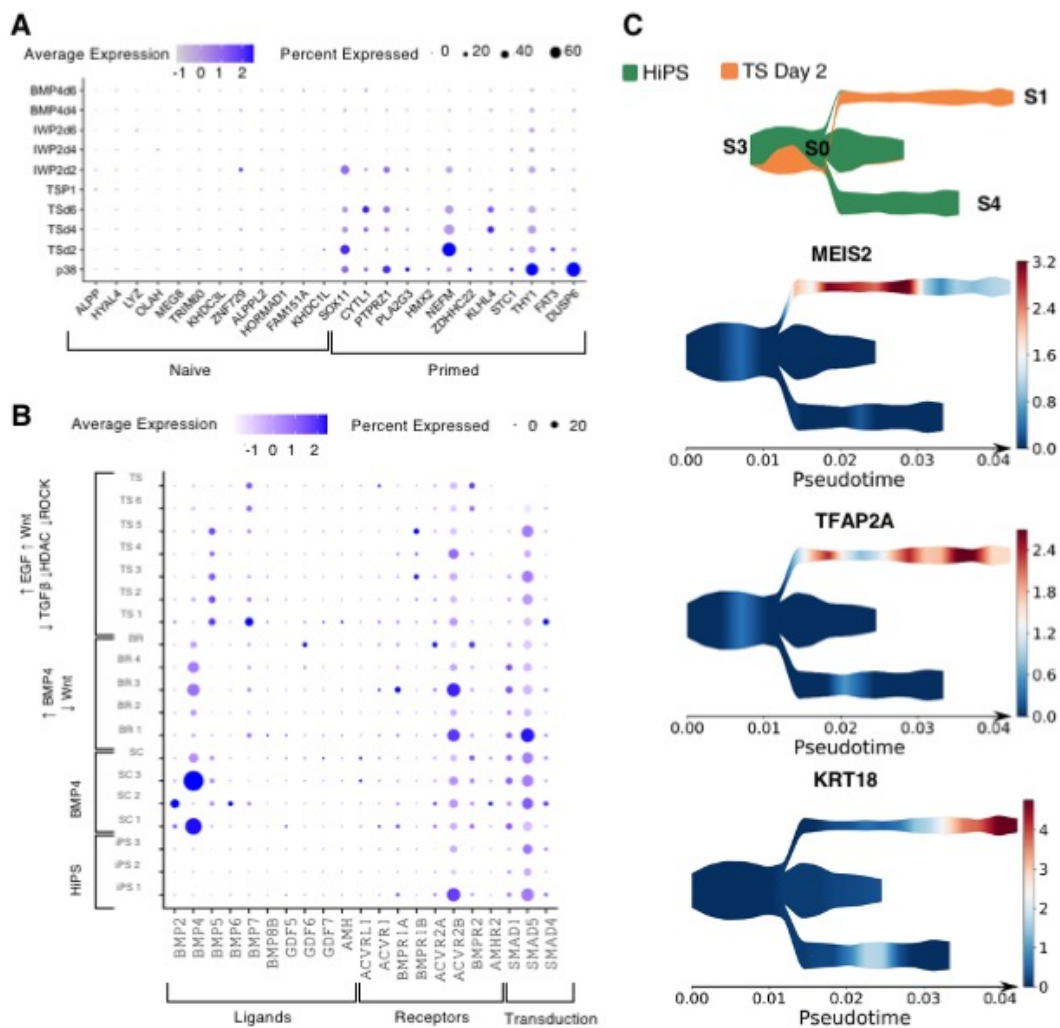


Figure 5: Primed iPS specification to trophoblast stem cells initiates with TFAP2A rewiring without activation of naïve stem cell programs

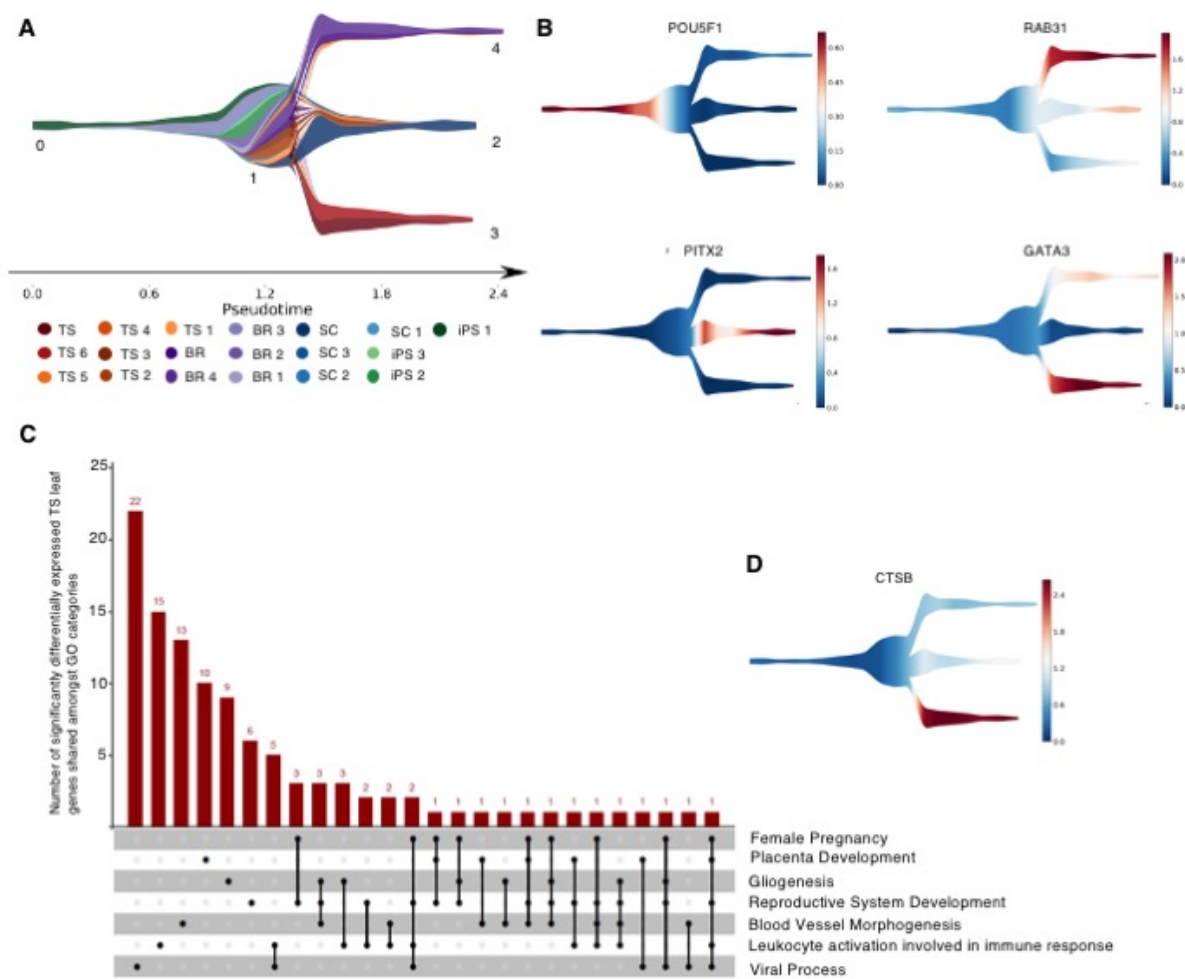


Figure 6: STREAM pseudotime analysis of trophoblast differentiations reveals upregulation of stromal mesenchyme upon BMP4 treatment and trophectoderm transcriptional programming up TS treatment.

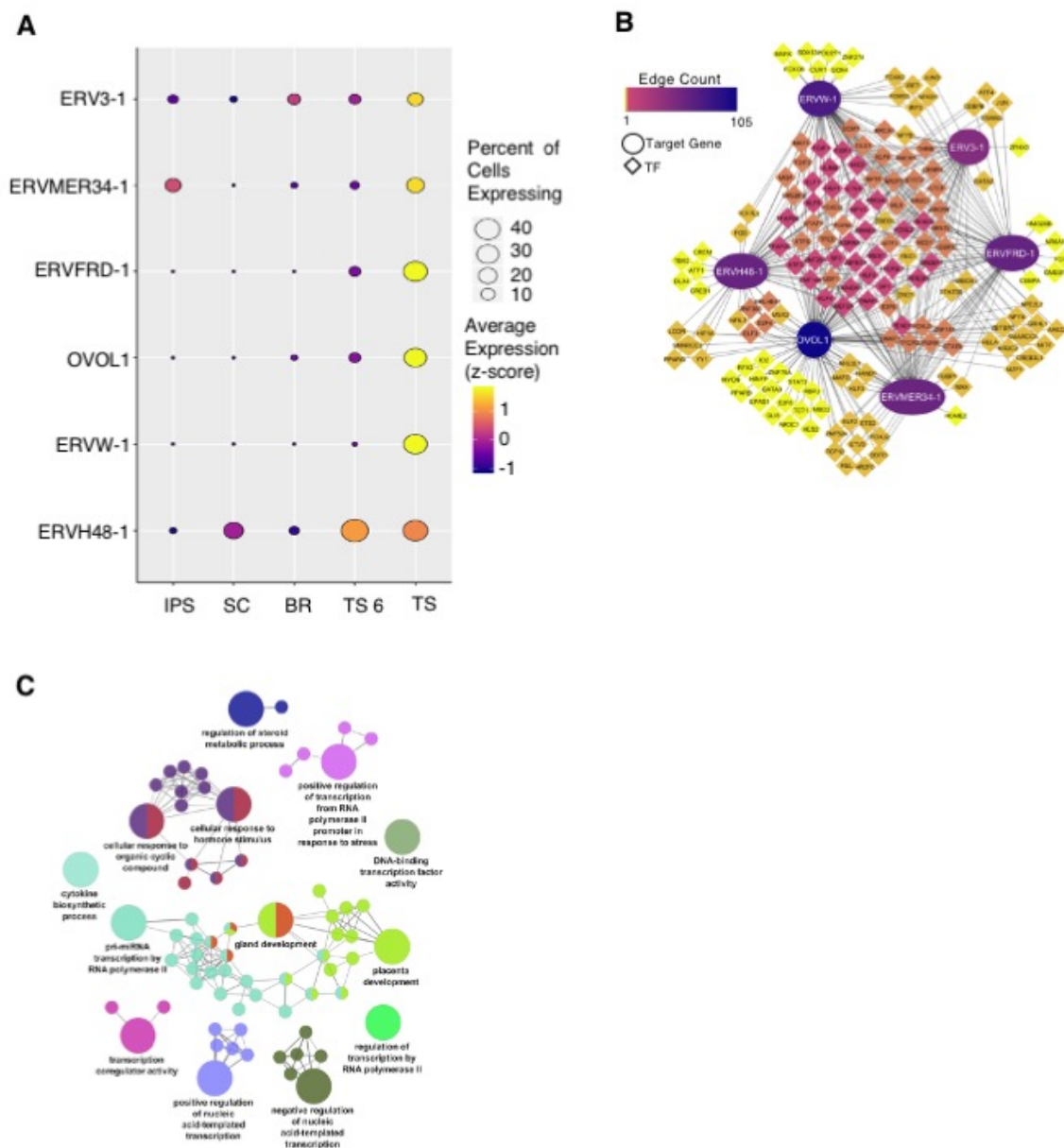


Figure 7: Endogenous retrovirus-derived genes participate in a gene regulatory subnetwork uniquely in TS cells.

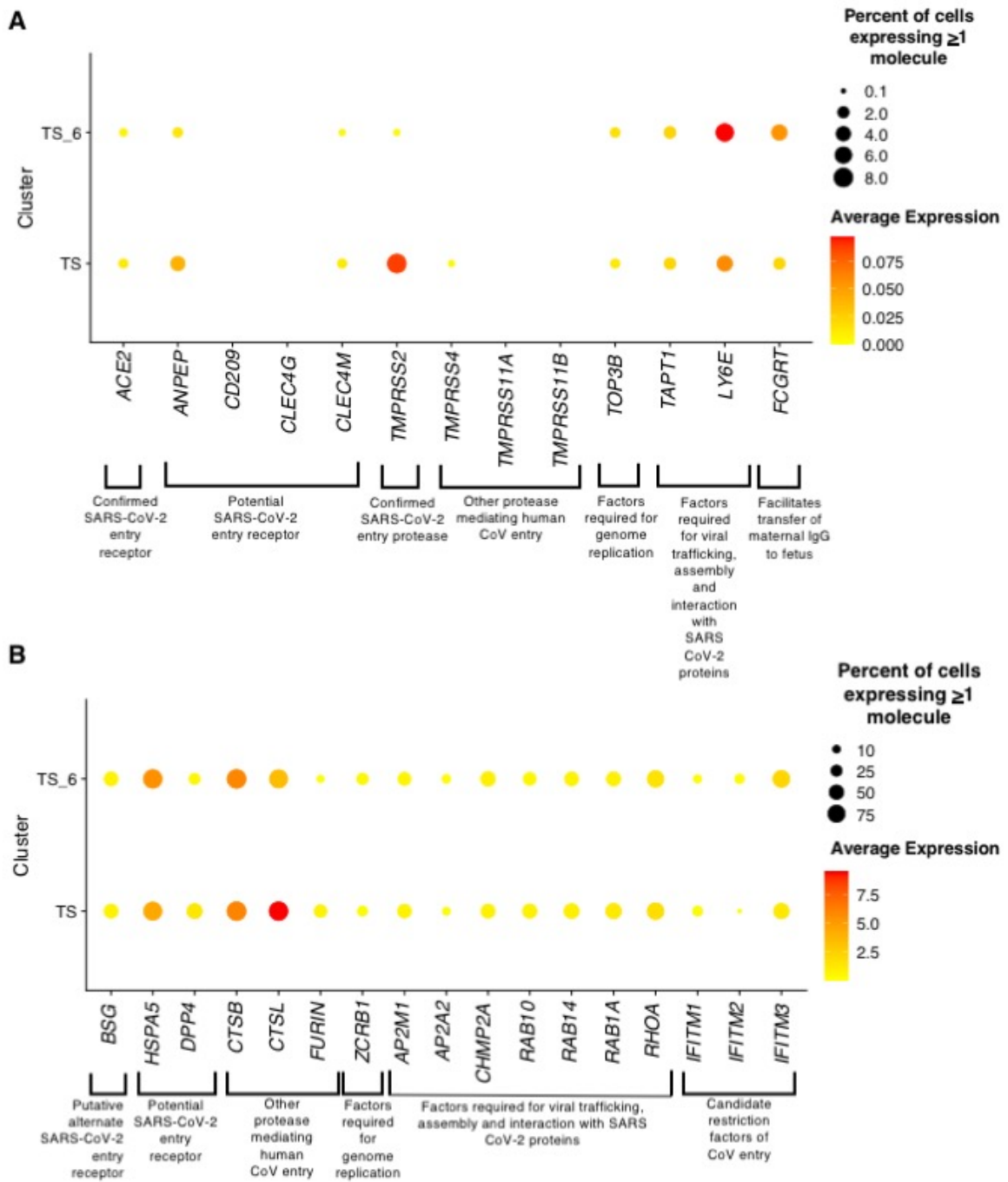


Figure 8: Coronavirus associated host factors are expressed in iPS-derived TS cells

Supplemental Figures

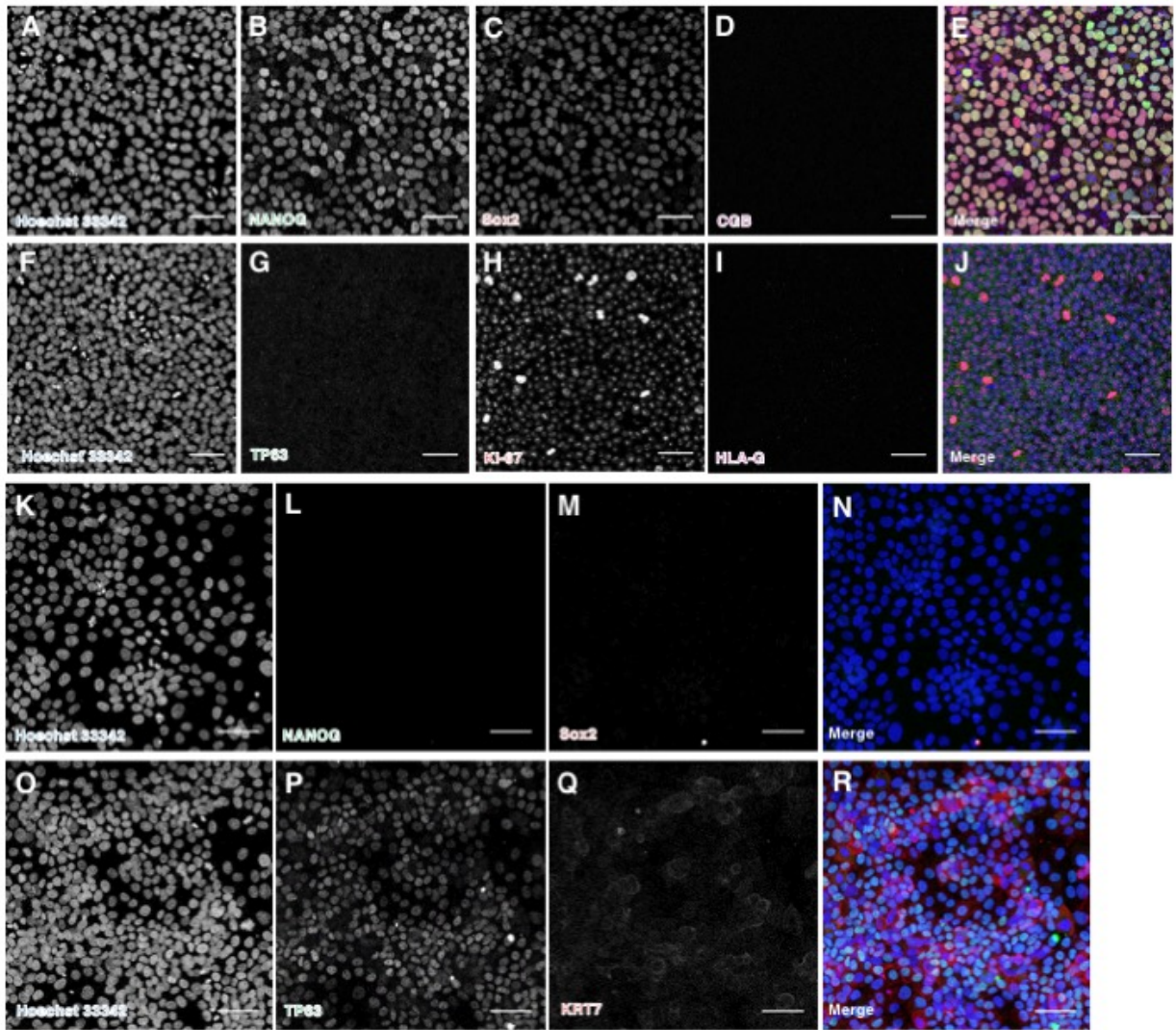


Fig. S1: Differentiated TS cells lose expression of pluripotency markers while gaining expression of trophoblast specific markers.

(A-E) Immunofluorescent images of HiPS line LIBD7C6 stained for Hoechst 33342 (A), NANOG (B), and Sox2 (C), CGB (D), and in composite (E).

(F-J) Immunofluorescent images of HiPS line LIBD7C6 stained for Hoechst 33342 (F), TP63(G), and KI-67 (H), HLA-G (I), and in composite (J).

(K-N) Immunofluorescent images of TS line 2014.06 stained for Hoechst 33342 (K), NANOG(L), and Sox2 (M), and in composite (N).

(O-R) Immunofluorescent images of TS line 2014.06 stained for Hoechst 33342 (O), TP63 (P), KRT7 (Q), and in composite (R).

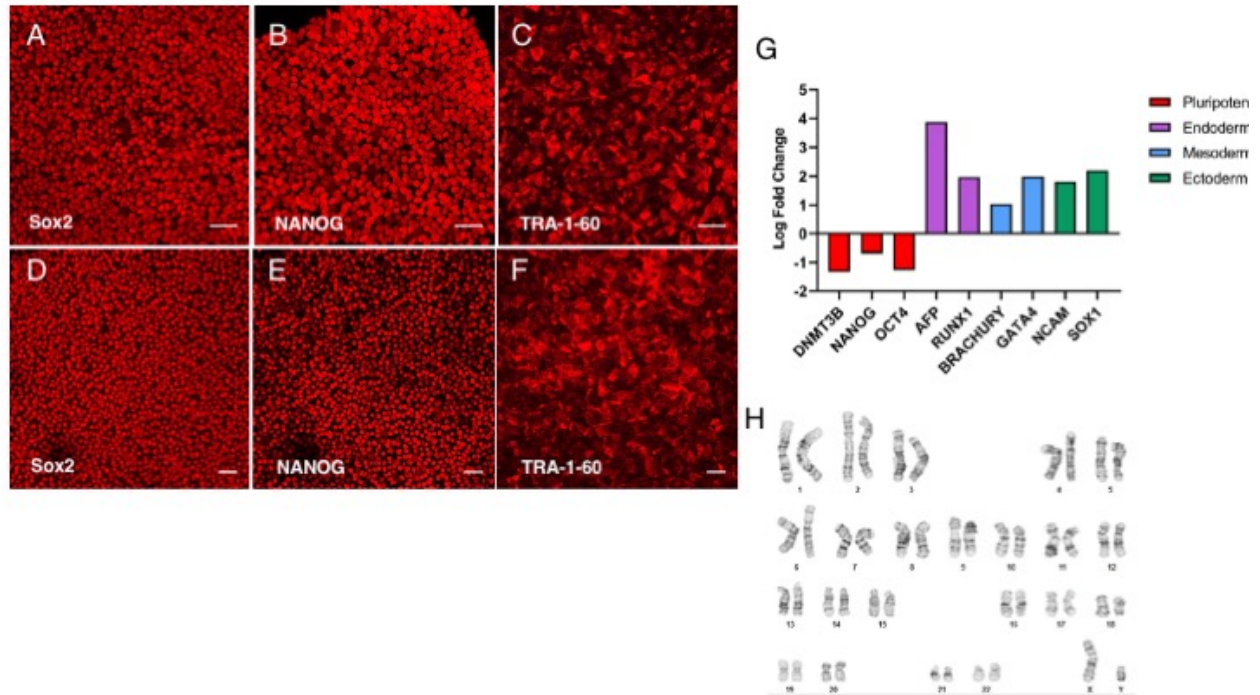


Fig. S2: iPS cell line characterization.

(A-C) Immunofluorescent images of HiPS line LIBD1C8 stained for Sox2 (A), NANOG (B), and TRA-1-60 (C).

(D-F) Immunofluorescent images of HiPS line 2014.06 stained for Sox2 (D), NANOG (E), and TRA-1-60 (F).

(G) Relative expression levels of pluripotency and germ layer specific markers in spontaneously differentiated cells of line 2014.06 normalized to their respective expression levels in undifferentiated HiPS of the same cell line determined by qPCR.

(H) Karyotype analysis of 2014.06

All scale bars 50 μ m.

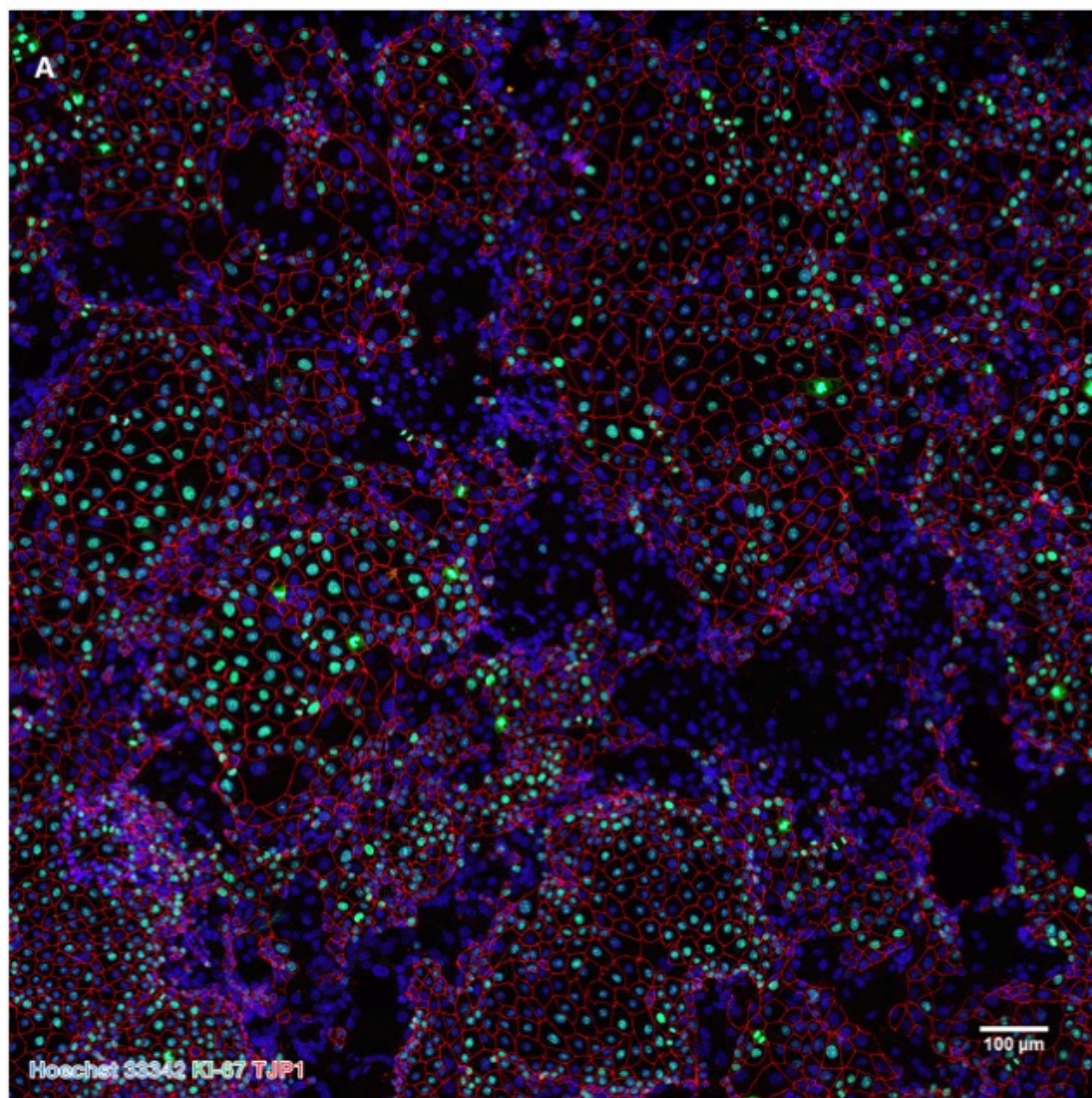


Fig. S3: Tile Scan Confluent TS LIBD1C8 P20 demonstrating extensive syncytium

(A) Tile scan of confluent TS line LIBD1C8 passage 20 grown on Collagen IV and stained for Hoechst 33342 (blue), KI67 (green), and TJP1 (red).

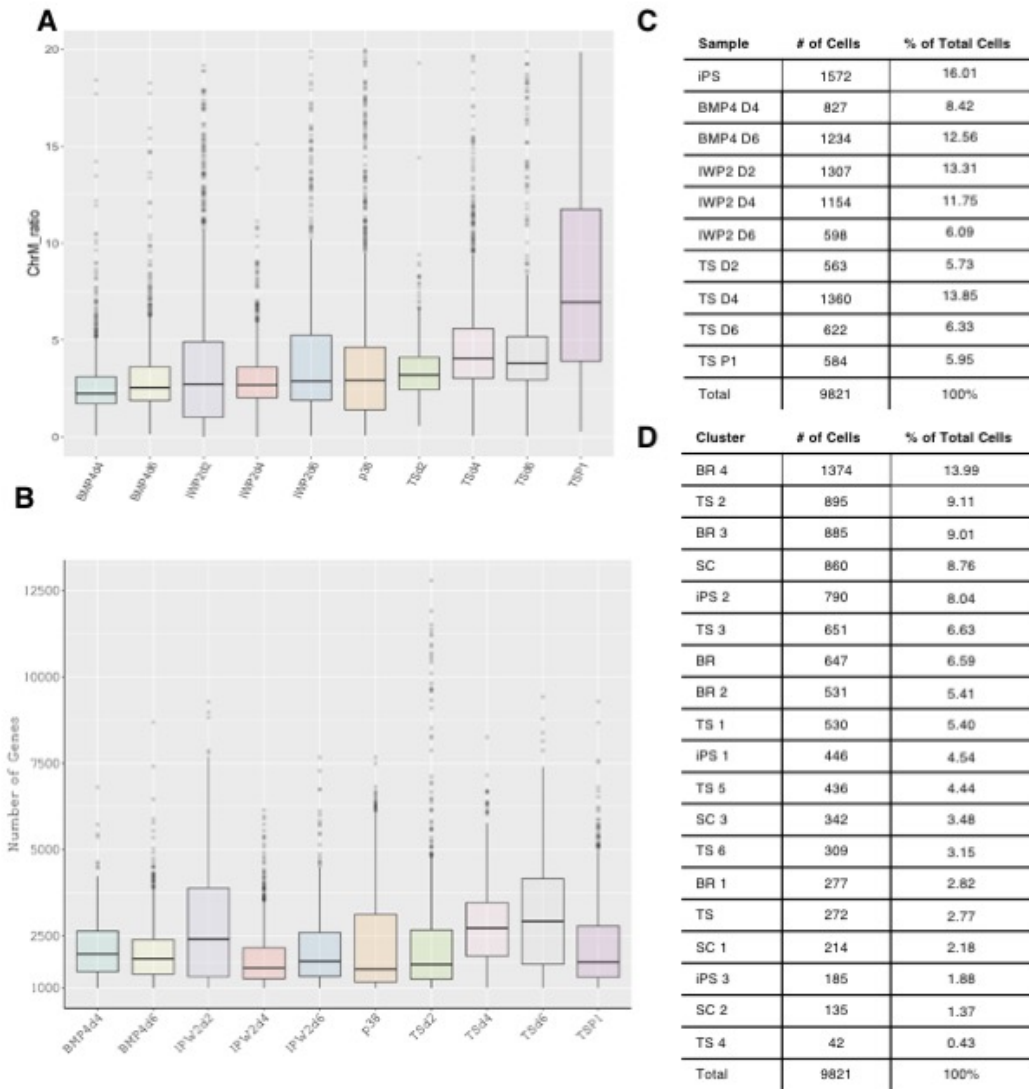


Fig. S4: Single Cell RNAseq Quality Control.

(A) Boxplot of Mitochondrial Mapping rate per droplet in the final data set.

(B) Boxplot of number of detected genes per droplet in the final data set.

(C-D) Summary tables of cell number and percent composition by experimental sample (C) and cluster identity (D).

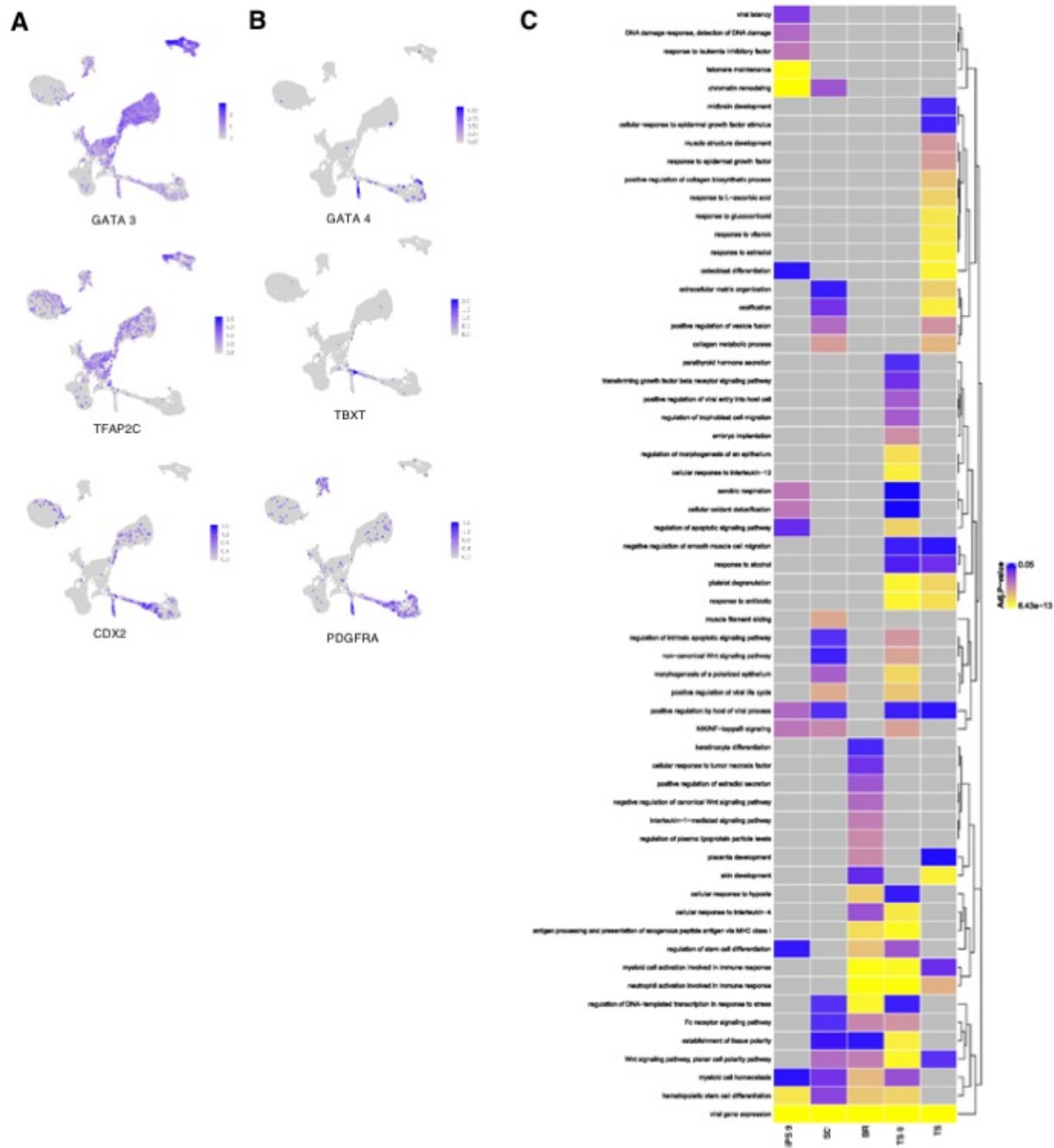


Fig. S5: BMP4 model comparison.

(A-B) Highlighted expression of trophoblast (A) and mesoderm (B) associated transcription factors at single cell resolution.

(C) Heatmap of selected GO terms (p-value < 0.05) generated from the 300 most highly expressed genes for each of the indicated mature cell types.

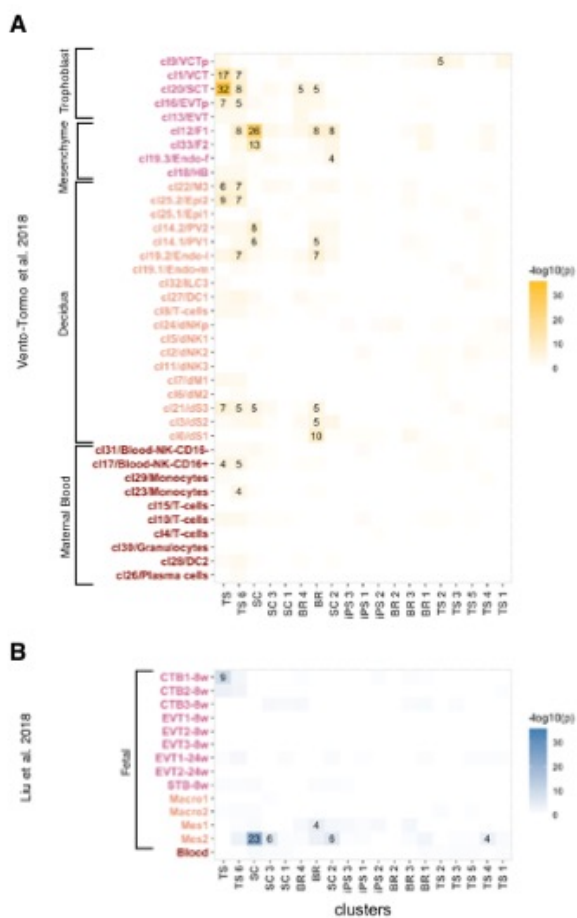
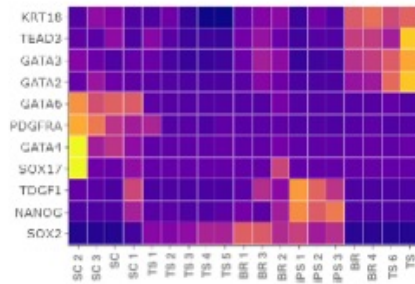


Fig. S6: Trophoblast differentiations are enriched for first trimester placental cell types.

(A,B) Gene set enrichment analysis of HiPS cell differentiations highlighting enrichment of cell clusters to cell types identified in first trimester placentas. Color and number in each square indicates the $-\log_{10}(p)$ Bonferroni Corrected hypergeometric test for enrichment of cell-specific genes in IPS-derived clusters and previous placental publications.

A



B

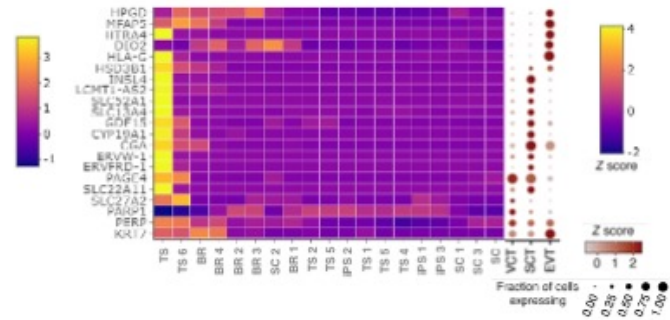


Fig. S7: Lineage specification and cellular cluster identity based on marker gene profiles.

- (A) Heatmap of detected high confidence markers for epiblast, primitive endoderm, and trophoblast from Stirparo et al. 2018.
- (B) Heatmap of trophoblast specific markers from Suryawanshi et al. 2018 including modified dotplot from their publication for primary Villous Cytotrophoblasts (VCT), Extravillous Trophoblasts (EVT), and Syncytial Trophoblasts (SCT).

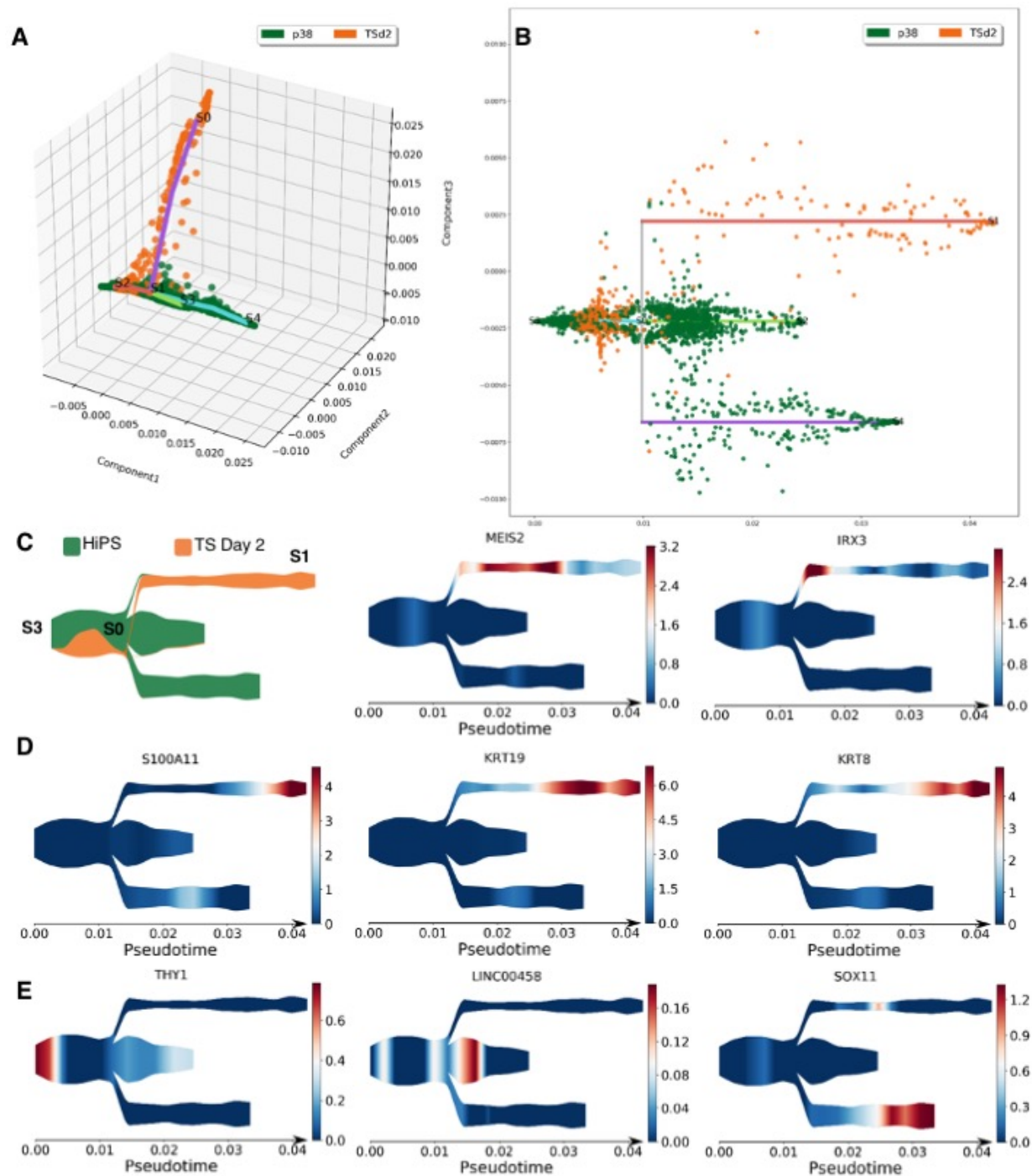


Fig. S8: Developmental trajectory of HiPS and after two days of differentiation in TS media.

(A-B) Visualization of branch assignment using the first 3 SE components (A) and resulting subway map representation of the reconstructed pseudotime trajectory using HiPS and TS D2 cells samples (B).

(C-D) Averaged expression of significant transition genes along the pseudotime trajectory ordered left to right by rank (2-4) for branch S3-S0 (C) and S0-S1 (D).

(E) Averaged expression of selected primed pluripotency markers expressed along pseudotime.

A



Figure S9: PANDA network analysis reveals unique regulation of endogenous Retroviruses in TS cells.

(A) ERV48-1 connected network identified by PANDA in SC cells.

736 **Methods**

737

738 *PSC culture*

739 6-well plates were coated with reduced growth factor Cultrex (1mg/12 ml DMEM/F12) at 37°C
740 for at least one hour. Human ES and iPS were maintained on coated plates in Stemflex media.
741 Media was changed every 48hrs in accordance with manufacturer recommendations. Cells
742 were passaged in small clusters using Versene solution and a split ratio of 1:10-1:12. For
743 cryopreservation, cells were suspended in PSC media and temporarily stored in a deep freezer
744 at -80°C before being transferred to liquid nitrogen for long term storage.

745

746 *Differentiations and Cell Maintenance*

747 6-well plates were coated with reduced growth factor Cultrex (1mg/12 ml DMEM/F12) at 37°C
748 for at least one hour. PSC were harvested in small clusters using Versene solution without
749 centrifugation, seeded at a passage ratio of 1:12, and cultured in 2 mL of Stemflex medium.
750 After 24 hours, cells were washed with DPBS and 2 mL of appropriate media was added to
751 each differentiating well. For TS differentiation, TS medium [DMEM/F12 with Glutamax
752 supplemented with 0.1 mM 2-mercaptoethanol, 0.2% FBS, 0.3% BSA, 1% ITS-X supplement,
753 1.5 µg/ml L-ascorbic acid, 50 ng/ml EGF, 2 µM CHIR99021, 0.5 µM A83-01, 1 µM SB431542,
754 0.8 mM VPA and 5 µM Y27632] was added to each well. For BMP4 based differentiations 10
755 ng/ml BMP4 and/or 2µM IWP2 was added to basal differentiation medium (DMEM/F12 with
756 Glutamax supplemented with ITS and L-ascorbic acid 2-phosphate magnesium) was added to
757 each well. Every 24 hours 2 mL of fresh medium was added to each well. Cells were collected
758 on day 2, 4, and 6 for RNA sequencing. For the TS condition, after 6 days of differentiation cells
759 were passaged using TrypLE express at a split ratio of 1:3 and plated on new Cultrex or col IV
760 (5 ug/mL) coated plates. After the first passage, cells were fed every 48 hours and split at a ratio
761 of 1:3-1:6 every 3-4 days. Mature TS cells between passages 10-20 were used for all
762 experiments unless otherwise noted. Cells were cultured in humidified incubators 5% CO₂ and
763 20% O₂. For hypoxic cultures, O₂ levels were reduced to 5%. For cryopreservation, TS cells
764 were suspended in Cell Banker 1 and temporarily stored in a deep freezer at -80°C before
765 being transferred to liquid nitrogen for long term storage.

766

767 *Immunofluorescence Staining*

768 Cells were washed in Tris Buffered Saline (TBS), fixed with 4% Paraformaldehyde for ten
769 minutes, and washed three times in TBS. Cells were then incubated in blocking solution (3%
770 horse serum, 0.1% Triton-X in TBS) for one hour at room temperature before being incubated
771 with primary antibodies in blocking solution overnight at 4°C. Cells were then rinsed twice and
772 incubated with secondary antibodies in blocking solution for two hours at room temperature.
773 After two washes in TBS cells were counterstained with Hoechst 33342 and rinsed twice in
774 TBS. Cells were imaged using a Zeiss LSM770 or LSM780 confocal microscope. The
775 brightness and/or contrast was globally adjusted for some images.

776

777 *Single Cell sequencing by Dropseq*

778 *Dissociation of Cells for Dropseq*

779 Human iPS cells were grown to 80-90% confluency. After washing cells with DPBS (Thermo
780 Fisher Scientific 14190250) cells were incubated with Versene (Thermo Fisher Scientific

781 15040066) solution for approximately 10 minutes at 37 °C and gently washed up and down with
782 a P1000 pipette to generate a single cell suspension. Cells were then diluted in Stemflex media
783 and centrifuged at 200g for 4 min before being resuspended in 1 mL of 1X PBS, pH 7.4 (Gibco,
784 # 10010023) with 0.01% BSA (Invitrogen, # AM2616) at the concentration of 100 cells/μL.

785

786 *Droplet generation*

787 Barcoded beads (Chemgenes, # Macosko-2011-10) were washed with 30 mL of ethanol (Sigma
788 Aldrich, # E7023-500ML), then removed all ethanol after centrifugation at 1,000g for 1 minute.
789 Washed beads were suspended in 30 mL of TE-TW (10 mM Tris pH8.0, 1 mM EDTA, 0.5%
790 SDS) and spun down at 1,000g for 1 minute. TE-TW was discarded and re-suspended in 20 mL
791 of TE-TW. 100 μm strainer (Corning, Cat# 352360) was used for size selection of beads.
792 Beads were suspended as the concentration of 120 beads/μL in 1 mL of the lysis buffer. To
793 generate droplets, cells, beads, and oil (BioRad, #186-4006) were loaded into the 125-μm
794 microfluidic device after aquapel flush treatment (FlowJEM). The cells and barcoded beads
795 suspension were flown at 2 mL/h each and at 13 mL/h for oil.

796

797 *Droplet breakage, reverse transcription, and exonuclease I treatment*

798 Droplets were collected into the 50 mL tubes and 30 mL of 6X SSC was added, then broken
799 with 1 mL of perfluorooctanol (Sigma, # 370533). Collected beads were washed with 1 mL of 6X
800 SSC twice and with 300 μL of RT buffer. On removing the RT buffer, reverse transcription
801 mixture was added into the beads. To synthesize cDNA, beads in RT mix were incubated with
802 rotation at room temperature for 30 minutes and 42 °C for 90 minutes. The reverse transcription
803 mixture was prepared as in the reference, Macosko, E.Z. et. al., Cell 161, 1202-1214 (2015).
804 Beads were washed once with 1 mL of TE-SDS, twice with 1 mL of TE-TW, and with 1 mL of 10
805 mM Tris pH 8.0. Exonuclease I treatment was performed to remove excess bead primers at 37
806 °C for 45 minutes with rotation followed by washing beads once with 1 mL of TE-SDS, twice
807 with 1 mL of TE-TW, and with 1 mL of molecular grade water.

808

809 *PCR amplification, clean-up with SPRI beads, and library preparation*

810 Unique barcode STAMPs (Single-cell Transcriptomes Attached to MicroParticles) were
811 resuspended in 1 mL H₂O, and loaded on a hemocytometer (Incyto, # DHC-N01). Aliquots of
812 3,000 beads in 24.6 μL of H₂O were amplified with 0.8 μM SMART PCR primer (IDT, 5'-AAG
813 CAG TGG TAT CAA CGC AGA GT-3') and 2X KAPA HiFi Hotstart Ready Mix (Fisher Scientific,
814 # KK2602) using the following PCR condition: (step 1) 95 °C for 3 min, (step 2) four cycles of:
815 98 °C for 20 sec, 65 °C for 45 sec, 72 °C for 3 min, (step 3) 10 cycles of: 98 °C for 20 sec, 67 °C
816 for 20 s, 72 °C for 3 min, and (step 4) finally, 72 °C for 5 min.

817 Purified cDNA with 0.6X SPRI beads were analyzed with High Sensitive DNA kit (Agilent, #
818 5067-4626), then tagmented and amplified using Nextera XT Library Preparation Kit (Illumina, #
819 FC-131-1096) by manufacturer's manual for sequencing.

820

821 *Sequencing*

822 All libraries were checked its quality with Agilent High Sensitive DNA chips before sequencing.
823 Each library was diluted at 3 nM for an Illumina HiSeq 3000. We used HiSeq SBS 50 cycle kits
824 for sequence (paired-end): 21-bp for read 1 with Custom Read1 primer (IDT, 5'-

825 GCCTGTCCGCGGAAGCAGTGGTATCAACGCAGAGTAC-3'), 21-bp for read 2, and 8-bp for
826 index 1.

827

828 *Preprocessing of Drop-seq data*

829 Raw sequencing data was preprocessed using the pipeline “Drop-seq Alignment Cookbook”
830 v2.0.0 found at <https://github.com/broadinstitute/Drop-seq/releases/> and described in (Macosko
831 et al., 2015). Briefly, paired-end reads were filtered in order to remove read pairs with any base
832 with quality of less than 10 in both cellular and molecular barcodes. SMART adapters at 5' end
833 and polyA tails at 3' end with 6 or more bp were removed from the second pair, and then
834 aligned to the reference human (GRCh38) genome using HISAT2 v2.1.0 (Kim et al., 2015) with
835 the default settings. Uniquely mapped reads were retained and grouped by cell barcode. UMI
836 barcodes with a hamming distance = 1 were merged within each cell, and the number of
837 transcripts of each gene was defined based on the total number of each distinct UMI sequence,
838 for that gene, in a given cell.

839

840 *Clustering and Differential Gene Expression analysis for Marker Genes*

841 All downstream analysis for quality control, clustering, and differential expression were
842 performed using Seurat (Stuart et al., 2019) v3.0 R package. Data was initially filtered to remove
843 low quality cells and very low detected genes. We discarded cells with less than 1000 genes
844 detected and with more than 20% of mitochondrial gene content. Genes expressed in less than
845 3 cells were also removed. Variance stabilizing and normalization was performed using
846 sctransform (Hafemeister and Satija, 2019) approach adjusting for library size and the
847 percentage of mitochondrial genes in each cell. To cluster the cells, cell distances were defined
848 with 30 first PCs using a graph-based approach implemented in Seurat, that briefly build a K-
849 nearest neighbor (KNN) graph based on euclidean distance in PCA space, with subsequent
850 refining of edge weights using Jacard similarity. Clusters were defined using Seurat default
851 algorithm – a shared nearest neighbor (SNN) graph based on k-nearest neighbors with
852 modularity optimization using original Louvain algorithm. Gene markers of cell clusters were
853 defined using ROC analysis for only genes with at least 0.25 logFC average difference between
854 the clusters. Graphic visualizations of clusters and gene expression were done using Seurat
855 and ggplot2 in R.

856

857 *Single-cell trajectory reconstruction*

858 Reconstruction of cell differentiation trajectories for cell clusters was performed with STREAM
859 (Single cell Trajectories Reconstruction, Exploration and Mapping) (Chen et al., 2019) Python
860 package. The expression data of high quality cells (9,821 cells) was library size normalized,
861 scaled, natural-log transformed and used as input for STREAM. Data dimension reduction was
862 done with Spectral Embedding (SE) based on most variable genes found with LOESS fitting.
863 Branches representing trajectory structure were learned based on 3 SE dimensions.
864 Marker genes of each branch were defined by Leaf gene detection analysis implemented in the
865 STREAM package. Initially, the average expression of each gene is calculated for all branches.
866 Based on the average expressions, Z-scores are computed and genes are picked as candidates
867 when greater than 1.5. Then a Kruskal–Wallis H-test is calculated for each candidate to test the
868 difference of gene expression among branches. Finally, the Conover test for multiple

869 comparison is computed for the differentially expressed candidates (Kruskal-Wallis p-value <
870 0.01) between all branches, and genes with Conover p-value < 0.01 are picked as a Leaf gene
871 for that branch. GO enrichment analysis of branch markers were performed for top 100 leaf
872 genes using hypergeometric test with GOstats v2.48.0 R package. Circos plot with GO terms
873 was done using GOplot v1.0.2 R package.

874 Transition genes were found by performing transition gene analysis implemented in STREAM.
875 The algorithm starts ordering cells based on their inferred pseudotime and splits them in two
876 groups. One group contains 20% of cells closer to the pseudotime start point, and the other is
877 composed of all remaining cells (80%). Gene average expressions are calculated for each
878 group and genes with a log2 fold change greater than 0.25 between groups are picked as
879 candidate genes. Then, Spearman's rank correlation is computed between inferred pseudotime
880 and gene expression of the candidates. Genes with correlation coefficient above 0.4 are
881 reported as transition genes.

882 Overrepresentation analysis was performed by using the algorithm MSET (Eisinger et al., 2013)
883 implemented in R, using 10000 permutations. The list of transcription factors related to
884 trophoblast differentiation used is described in (Krendl et al., 2017).

885

886 *Clusters enrichment analysis for placental cell-specific genes.*

887 To characterize the expression profiles of the 19 clusters detected in this work, we analyzed the
888 overlap with cell-specific genes identified by previous single-cell studies. We used two sets of
889 candidate genes. A set of 38 cell-specific gene lists (30 gene each) was drawn from the
890 supplementary table 2 of (Vento-Tormo et al. 2018). These gene lists span both maternal
891 (decidua and blood) and fetal tissues (placenta). Another set of 14 cell-specific gene lists was
892 drawn from the supplementary table 1 of (Liu et al. 2018), which also span analogous tissues.

893

894 To profile the expression patterns of the 19 clusters detected in this work, we applied the Cell-
895 Specific Expression Analysis (CSEA) as implemented in the pSI R package (Xu et al., 2014).
896 This approach compares each profile with all the others and identifies genes expressed in one
897 group but not in the others, calculates a score (specificity index, SI) for each gene and attributes
898 a statistical level of significance (pSI). Here, we used the normalized expression matrix to
899 compute cluster-level median gene expression for each gene, which defined 19 cluster profiles.
900 Then, cluster profiles were used to compute SI and pSI for each gene, within each profile.
901 Cluster-specific gene lists were obtained applying the statistical threshold, pSI < 0.05.

902

903 Finally, we tested whether cell-specific genes from previous studies are over-represented in our
904 clusters. We used the hypergeometric test and applied the Bonferroni correction for multiple
905 comparisons, considering all the tested gene lists [$\alpha = 0.05 / (19 \times (38+14)) = 5.1 \times 10^{-5}$].

906

907 *PANDA regulatory networks*

908 Gene regulatory networks were constructed using the pandaR v1.14.0 R package which
909 integrates multiple types of data to infer direct interactions (edges) between TFs and target
910 genes. PANDA initiates with a prior regulatory network that can be built by mapping TF binding
911 sites to the genome, and refines this initial network integrating gene expression data of target
912 genes and TF PPI data. The main idea behind the algorithm is that target genes from a given

913 TF are likely co-expressed, and TFs that interact with each other are more likely to regulate a
914 similar group of target genes. These two assumptions are used to infer the edge weight of each
915 TF-target interaction, and basically reflects the congruence between the regulatory profile of a
916 TF with target gene co-expression. Iteratively the algorithm refine the initial network structure
917 and infer a final consensus regulatory network.

918 A regulatory network was built for each cluster using as initial network a TF-motif binding map
919 described in (Sonawane et al., 2017) and downloaded in
920 (<https://sites.google.com/a/channing.harvard.edu/kimberlyglass/tools/resources>). Gene
921 expression data was library size normalized, scaled, and natural-log transformed. Genes
922 expressed in less than 10% of the cells of each cluster were filtered out. PPI network of TFs
923 was built using STRING database v11.0 (downloaded from <https://string-db.org/>) and score
924 interactions were divided by 1000 to initiate the PANDA.
925 Edges with negative weights estimated by PANDA were discarded in order to explore only
926 interactions with greater evidence. The relative importance of individual genes in the network
927 topology was addressed with the betweenness centrality measure calculated using the lgraph
928 v1.2.4 R package. For each cluster, the significance of betweenness values were calculated
929 based on the betweenness probability distribution of 100 random networks constructed with the
930 same number of edges of that cluster. Graphic visualization, gene clusterization, and GO
931 enrichment analysis of gene clusters within networks were performed using Cytoscape v3.7.2
932 (Shannon et al., 2003) and ClueGO v2.5.4.ENRICHMENT/RANKING ANALYSIS (Bindea et al.,
933 2009).

934

935 *Software and Analysis:*

936 Analysis pipelines are publicly deposited on Github:
937 <https://github.com/apuapaquola?tab=repositories>

938

939 *Datasets:*

940 Single cell RNA sequencing data is deposited at:

941

942

943 **Key Reagents and Resources**

944

Reagent or Resource Type	Name	Supplier	Product Number	Additional information
Antibody	anti-hCGb (mouse monoclonal)	Abcam	ab9582	1:100
Antibody	anti-HLAG (mouse monoclonal)	Abcam	ab52455	1:100
Antibody	anti-KI67 (rabbit polyclonal)	Abcam	ab15580	1:250
Antibody	anti-KRT7 (mouse monoclonal)	ThermoFisher Scientific	MA5-11986	1:1000
Antibody	anti-TJP1 (mouse monoclonal)	ThermoFisher Scientific	33-9100	1:100
Antibody	anti-TP63 (goat polyclonal)	Novus Biologicals	AF1916	1:100
Antibody	anti-VIM (rabbit monoclonal)	Abcam	ab92547	1:500
Antibody	Donkey anti-mouse IgG AF-488	Jackson Immunoresearch Laboratories	715-545-151	1:500
Antibody	Donkey anti-mouse IgG Cy3	Jackson Immunoresearch Laboratories	715-165-151	1:500
Antibody	Donkey anti-rabbit IgG AF-488	Jackson Immunoresearch Laboratories	711-545-152	1:500
Antibody	Donkey anti-rabbit IgG Cy3	Jackson Immunoresearch Laboratories	711-005-152	1:500
Antibody	Donkey anti-rabbit IgG AF-647	Jackson Immunoresearch Laboratories	711-605-152	1:500
Antibody	Donkey anti-goat IgG Cy3	Jackson Immunoresearch Laboratories	705-165-147	1:500
Chemical, Peptide, or Recombinant Protein	2-mercaptoethanol	ThermoFisher Scientific	21985023	
Chemical, Peptide, or Recombinant Protein	Bmp4	Peprtech	120-05-5ug	

Chemical, Peptide, or Recombinant Protein	Bovine Serum Albumin	Cell Signaling Technology	9998S	
Chemical, Peptide, or Recombinant Protein	Cell Banker 1 Cryopreservation Media	Amsbio	11888	
Chemical, Peptide, or Recombinant Protein	Collagen IV	Corning	354233	
Chemical, Peptide, or Recombinant Protein	Cultrex Reduced Growth Factor Basement Membrane Extract, PathClear	R&D Systems	3433-005-01	
Chemical, Peptide, or Recombinant Protein	DMEM/F12, Glutamax supplement	ThermoFisher Scientific	10565018	
Chemical, Peptide, or Recombinant Protein	DPBS, no calcium, no magnesium	ThermoFisher Scientific	14190250	
Chemical, Peptide, or Recombinant Protein	Epidermal Growth Factor	Peprtech	AF-100-15	
Chemical, Peptide, or Recombinant Protein	Fetal Bovine Serum	ThermoFisher Scientific	16140071	
Chemical, Peptide, or Recombinant Protein	Gelatin	Sigma-Aldrich	G1393-100mL	
Chemical, Peptide, or Recombinant Protein	Hoechst 33342	ThermoFisher Scientific	H3570	
Chemical, Peptide, or Recombinant Protein	Insulin-transferrin-selenium	Corning	25-800-CR	
Chemical, Peptide, or Recombinant Protein	ITS-X	ThermoFisher Scientific	51500056	
Chemical, Peptide, or Recombinant Protein	IWP2	BioGems	6866167-1mg	
Chemical, Peptide, or Recombinant Protein	L-Ascorbic Acid	BioGems	2520691	

Chemical, Peptide, or Recombinant Protein	L-ascorbic acid 2-phosphate magnesium	Sigma-Aldrich	A8960-5G	
Chemical, Peptide, or Recombinant Protein	SB431542	BioGems	3014193	
Chemical, Peptide, or Recombinant Protein	StemFlex Media	ThermoFisher Scientific	A3349401	
Chemical, Peptide, or Recombinant Protein	TrypLE express	ThermoFisher Scientific	12604013	
Chemical, Peptide, or Recombinant Protein	Valproic Acid Sodium salt	BioGems	1066656	
Chemical, Peptide, or Recombinant Protein	Y-27632	BioGems	1293823	
Cell Line	2014.06 Human Induced Pluripotent Stem Cells	Lieber Institute	N/A	
Cell Line	CF-1 Mouse Embryonic Fibroblasts	ThermoFisher Scientific	A34181	
Cell Line	LIBD1C8 Human Induced Pluripotent Stem Cells	Sawada et al., 2020	N/A	Generated in lab
Cell Line	LIBD6C2 Human Induced Pluripotent Stem Cells	Sawada et al., 2020	N/A	Generated in lab
Cell Line	LIBD7C6 Human Induced Pluripotent Stem Cells	Sawada et al., 2020	N/A	Generated in lab
Cell Line	WA-01 (H1) Human Embryonic Stem Cells	Wicell	WB0197	

945

946

947

948

949 Bibliography

- 950 Abd El-Aleem, S.A., and Dekker, L.V. (2018). Assessment of the cellular localization of the
951 annexin A2/S100A10 complex in human placenta. *J. Mol. Histol.* *49*, 531–543.
- 952 Aguiar, J.A., Tremblay, B.J.-M., Mansfield, M.J., Woody, O., Lobb, B., Banerjee, A.,
953 Chandiramohan, A., Tiessen, N., Cao, Q., Dvorkin-Gheva, A., et al. (2020). Gene expression
954 and in situ protein profiling of candidate SARS-CoV-2 receptors in human airway epithelial cells
955 and lung tissue. *Eur. Respir. J.*
- 956 Al-Haddad, B.J.S., Oler, E., Armistead, B., Elsayed, N.A., Weinberger, D.R., Bernier, R., Burd,
957 I., Kapur, R., Jacobsson, B., Wang, C., et al. (2019). The fetal origins of mental illness. *Am. J.*
958 *Obstet. Gynecol.* *221*, 549–562.
- 959 Arai, C., Yoshizaki, K., Miyazaki, K., Saito, K., Yamada, A., Han, X., Funada, K., Fukumoto, E.,
960 Haruyama, N., Iwamoto, T., et al. (2017). Nephronectin plays critical roles in Sox2 expression
961 and proliferation in dental epithelial stem cells via EGF-like repeat domains. *Sci. Rep.* *7*, 45181.
- 962 Bernardo, A.S., Faial, T., Gardner, L., Niakan, K.K., Ortmann, D., Senner, C.E., Callery, E.M.,
963 Trotter, M.W., Hemberger, M., Smith, J.C., et al. (2011). BRACHYURY and CDX2 mediate
964 BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and
965 extraembryonic lineages. *Cell Stem Cell* *9*, 144–155.
- 966 Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.-H.,
967 Pagès, F., Trajanoski, Z., and Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher
968 functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* *25*, 1091–
969 1093.
- 970 Blaise, S., de Parseval, N., Bénit, L., and Heidmann, T. (2003). Genomewide screening for
971 fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on
972 primate evolution. *Proc Natl Acad Sci USA* *100*, 13013–13018.
- 973 Blakeley, P., Fogarty, N.M.E., del Valle, I., Wamaitha, S.E., Hu, T.X., Elder, K., Snell, P.,
974 Christie, L., Robson, P., and Niakan, K.K. (2015). Defining the three cell lineages of the human
975 blastocyst by single-cell RNA-seq. *Development* *142*, 3151–3165.
- 976 Blond, J.L., Besème, F., Duret, L., Bouton, O., Bedin, F., Perron, H., Mandrand, B., and Mallet,
977 F. (1999). Molecular characterization and placental expression of HERV-W, a new human
978 endogenous retrovirus family. *J. Virol.* *73*, 1175–1185.
- 979 Blond, J.L., Lavillette, D., Cheynet, V., Bouton, O., Oriol, G., Chapel-Fernandes, S., Mandrand,
980 B., Mallet, F., and Cosset, F.L. (2000). An envelope glycoprotein of the human endogenous
981 retrovirus HERV-W is expressed in the human placenta and fuses cells expressing the type D
982 mammalian retrovirus receptor. *J. Virol.* *74*, 3321–3329.
- 983 Chen, B., Dodge, M.E., Tang, W., Lu, J., Ma, Z., Fan, C.-W., Wei, S., Hao, W., Kilgore, J.,
984 Williams, N.S., et al. (2009). Small molecule-mediated disruption of Wnt-dependent signaling in
985 tissue regeneration and cancer. *Nat. Chem. Biol.* *5*, 100–107.
- 986 Chen, H., Albergante, L., Hsu, J.Y., Lareau, C.A., Lo Bosco, G., Guan, J., Zhou, S., Gorban,

- 987 A.N., Bauer, D.E., Aryee, M.J., et al. (2019). Single-cell trajectories reconstruction, exploration
988 and mapping of omics data with STREAM. *Nat. Commun.* *10*, 1903.
- 989 Chen, I.-H., Wang, H.-H., Hsieh, Y.-S., Huang, W.-C., Yeh, H.-I., and Chuang, Y.-J. (2013).
990 PRSS23 is essential for the Snail-dependent endothelial-to-mesenchymal transition during
991 valvulogenesis in zebrafish. *Cardiovasc. Res.* *97*, 443–453.
- 992 Chua, C.E.L., and Tang, B.L. (2014). Engagement of the small GTPase Rab31 protein and its
993 effector, early endosome antigen 1, is important for trafficking of the ligand-bound epidermal
994 growth factor receptor from the early to the late endosome. *J. Biol. Chem.* *289*, 12375–12389.
- 995 Cole, L.A. (2010). Biological functions of hCG and hCG-related molecules. *Reprod. Biol.*
996 *Endocrinol.* *8*, 102.
- 997 Cornelis, G., Heidmann, O., Degrelle, S.A., Vernochet, C., Lavielle, C., Letzelter, C., Bernard-
998 Stoecklin, S., Hassanin, A., Mulot, B., Guillomot, M., et al. (2013). Captured retroviral envelope
999 syncytin gene associated with the unique placental structure of higher ruminants. *Proc Natl*
1000 *Acad Sci USA* *110*, E828-37.
- 1001 Dong, C., Beltcheva, M., Gontarz, P., Zhang, B., Popli, P., Fischer, L.A., Khan, S.A., Park, K.-
1002 M., Yoon, E.-J., Xing, X., et al. (2020). Derivation of trophoblast stem cells from naïve human
1003 pluripotent stem cells. *Elife* *9*.
- 1004 Drukker, M., Tang, C., Ardehali, R., Rinkevich, Y., Seita, J., Lee, A.S., Mosley, A.R., Weissman,
1005 I.L., and Soen, Y. (2012). Isolation of primitive endoderm, mesoderm, vascular endothelial and
1006 trophoblast progenitors from human pluripotent stem cells. *Nat. Biotechnol.* *30*, 531–542.
- 1007 Dunn-Fletcher, C.E., Muglia, L.M., Pavlicev, M., Wolf, G., Sun, M.-A., Hu, Y.-C., Huffman, E.,
1008 Tumukuntala, S., Thiele, K., Mukherjee, A., et al. (2018). Anthropoid primate-specific retroviral
1009 element THE1B controls expression of CRH in placenta and alters gestation length. *PLoS Biol.*
1010 *16*, e2006337.
- 1011 Eisinger, B.E., Saul, M.C., Driessen, T.M., and Gammie, S.C. (2013). Development of a
1012 versatile enrichment analysis tool reveals associations between the maternal brain and mental
1013 health disorders, including autism. *BMC Neurosci.* *14*, 147.
- 1014 Erb, T.M., Schneider, C., Mucko, S.E., Sanfilippo, J.S., Lowry, N.C., Desai, M.N., Mangoubi,
1015 R.S., Leuba, S.H., and Sammak, P.J. (2011). Paracrine and epigenetic control of trophoblast
1016 differentiation from human embryonic stem cells: the role of bone morphogenic protein 4 and
1017 histone deacetylases. *Stem Cells Dev.* *20*, 1601–1614.
- 1018 Erwin, J.A., del Rosario, B., Payer, B., and Lee, J.T. (2012). An ex vivo model for imprinting:
1019 mutually exclusive binding of Cdx2 and Oct4 as a switch for imprinted and random X-
1020 inactivation. *Genetics* *192*, 857–868.
- 1021 Fukamachi, K., Matsuoka, Y., Ohno, H., Hamaguchi, T., and Tsuda, H. (2002). Neuronal
1022 leucine-rich repeat protein-3 amplifies MAPK activation by epidermal growth factor through a
1023 carboxyl-terminal region containing endocytosis motifs. *J. Biol. Chem.* *277*, 43549–43552.
- 1024 Gao, X., Nowak-Imialek, M., Chen, X., Chen, D., Herrmann, D., Ruan, D., Chen, A.C.H.,
1025 Eckersley-Maslin, M.A., Ahmad, S., Lee, Y.L., et al. (2019). Establishment of porcine and
1026 human expanded potential stem cells. *Nat. Cell Biol.* *21*, 687–699.

- 1027 Glass, K., Huttenhower, C., Quackenbush, J., and Yuan, G.-C. (2013). Passing messages
1028 between biological networks to refine predicted interactions. *PLoS ONE* 8, e64832.
- 1029 Goldman, D.C., Bailey, A.S., Pfaffle, D.L., Al Masri, A., Christian, J.L., and Fleming, W.H.
1030 (2009). BMP4 regulates the hematopoietic stem cell niche. *Blood* 114, 4393–4401.
- 1031 Guan, L., Li, T., Ai, N., Wang, W., He, B., Bai, Y., Yu, Z., Li, M., Dong, S., Zhu, Q., et al. (2019).
1032 MEIS2C and MEIS2D promote tumor progression via Wnt/ β -catenin and hippo/YAP signaling in
1033 hepatocellular carcinoma. *J. Exp. Clin. Cancer Res.* 38, 417.
- 1034 Guleria, I., and Pollard, J.W. (2000). The trophoblast is a component of the innate immune
1035 system during pregnancy. *Nat. Med.* 6, 589–593.
- 1036 Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell
1037 RNA-seq data using regularized negative binomial regression. *Genome Biol.* 20, 296.
- 1038 Hendee, K.E., Sorokina, E.A., Muheisen, S.S., Reis, L.M., Tyler, R.C., Markovic, V., Cuturilo,
1039 G., Link, B.A., and Semina, E.V. (2018). PITX2 deficiency and associated human disease:
1040 insights from the zebrafish model. *Hum. Mol. Genet.* 27, 1675–1695.
- 1041 Horii, M., Bui, T., Touma, O., Cho, H.Y., and Parast, M.M. (2019). An Improved Two-Step
1042 Protocol for Trophoblast Differentiation of Human Pluripotent Stem Cells. *Curr. Protoc. Stem
1043 Cell Biol.* 50, e96.
- 1044 Huang, D., Wang, S., Wang, A., Chen, X., and Zhang, H. (2016). Thymosin beta 4 silencing
1045 suppresses proliferation and invasion of non-small cell lung cancer cells by repressing Notch1
1046 activation. *Acta Biochim Biophys Sin (Shanghai)* 48, 788–794.
- 1047 Johnson, W.E. (2019). Origins and evolutionary consequences of ancient endogenous
1048 retroviruses. *Nat. Rev. Microbiol.* 17, 355–370.
- 1049 Kazanskaya, O., Glinka, A., del Barco Barrantes, I., Stannek, P., Niehrs, C., and Wu, W. (2004).
1050 R-Spondin2 is a secreted activator of Wnt/beta-catenin signaling and is required for *Xenopus*
1051 myogenesis. *Dev. Cell* 7, 525–534.
- 1052 Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low
1053 memory requirements. *Nat. Methods* 12, 357–360.
- 1054 Krendl, C., Shaposhnikov, D., Rishko, V., Ori, C., Ziegenhain, C., Sass, S., Simon, L., Müller,
1055 N.S., Straub, T., Brooks, K.E., et al. (2017). GATA2/3-TFAP2A/C transcription factor network
1056 couples human pluripotent stem cell differentiation to trophectoderm with repression of
1057 pluripotency. *Proc Natl Acad Sci USA* 114, E9579–E9588.
- 1058 Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar,
1059 K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome.
1060 *Nature* 409, 860–921.
- 1061 Leach, J.L., Sedmak, D.D., Osborne, J.M., Rahill, B., Lairmore, M.D., and Anderson, C.L.
1062 (1996). Isolation from human placenta of the IgG transporter, FcRn, and localization to the
1063 syncytiotrophoblast: implications for maternal-fetal antibody transport. *J. Immunol.* 157, 3317–
1064 3322.
- 1065 Lee, C.Q.E., Turco, M.Y., Gardner, L., Simons, B.D., Hemberger, M., and Moffett, A. (2018).

- 1066 Integrin $\alpha 2$ marks a niche of trophoblast progenitor cells in first trimester human placenta.
1067 *Development* 145.
- 1068 Liu, A.-X., Jin, F., Zhang, W.-W., Zhou, T.-H., Zhou, C.-Y., Yao, W.-M., Qian, Y.-L., and Huang,
1069 H.-F. (2006). Proteomic analysis on the alteration of protein expression in the placental villous
1070 tissue of early pregnancy loss. *Biol. Reprod.* 75, 414–420.
- 1071 Liu, Y., Fan, X., Wang, R., Lu, X., Dang, Y.-L., Wang, H., Lin, H.-Y., Zhu, C., Ge, H., Cross,
1072 J.C., et al. (2018). Single-cell RNA-seq reveals the diversity of trophoblast subtypes and
1073 patterns of differentiation in the human placenta. *Cell Res.* 28, 819–832.
- 1074 Li, Z., Kurosawa, O., and Iwata, H. (2019). Establishment of human trophoblast stem cells from
1075 human induced pluripotent stem cell-derived cystic cells under micromesh culture. *Stem Cell*
1076 *Res. Ther.* 10, 245.
- 1077 Lv, S., Cheng, G., Zhou, Y., and Xu, G. (2013). Thymosin beta4 induces angiogenesis through
1078 Notch signaling in endothelial cells. *Mol. Cell. Biochem.* 381, 283–290.
- 1079 Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas,
1080 A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression
1081 Profiling of Individual Cells Using Nanoliter Droplets. *Cell* 161, 1202–1214.
- 1082 Mansour, A.A., Schafer, S.T., and Gage, F.H. (2020). Cellular complexity in brain organoids:
1083 Current progress and unsolved issues. *Semin. Cell Dev. Biol.*
- 1084 Messmer, T., von Meyenn, F., Savino, A., Santos, F., Mohammed, H., Lun, A.T.L., Marioni, J.C.,
1085 and Reik, W. (2019). Transcriptional Heterogeneity in Naive and Primed Human Pluripotent
1086 Stem Cells at Single-Cell Resolution. *Cell Rep.* 26, 815-824.e4.
- 1087 Meyer, K., Kwon, Y.-C., Liu, S., Hagedorn, C.H., Ray, R.B., and Ray, R. (2015). Interferon- α
1088 inducible protein 6 impairs EGFR activation by CD81 and inhibits hepatitis C virus infection. *Sci.*
1089 *Rep.* 5, 9012.
- 1090 Mischler, A., Karakis, V., Mahinthakumar, J., Carberry, C., San Miguel, A., Rager, J., Fry, R.,
1091 and Rao, B.M. (2019). Two distinct trophectoderm lineage stem cells from human pluripotent
1092 stem cells. *BioRxiv*.
- 1093 Mouse Genome Sequencing Consortium, Waterston, R.H., Lindblad-Toh, K., Birney, E.,
1094 Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., et al.
1095 (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* 420, 520–
1096 562.
- 1097 Myatt, L. (2006). Placental adaptive responses and fetal programming. *J Physiol (Lond)* 572,
1098 25–30.
- 1099 Niakan, K.K., and Eggan, K. (2013). Analysis of human embryos from zygote to blastocyst
1100 reveals distinct gene expression patterns relative to the mouse. *Dev. Biol.* 375, 54–64.
- 1101 Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J.
1102 (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell*
1103 123, 917–929.
- 1104 Oh, Y., Nagalla, S.R., Yamanaka, Y., Kim, H.S., Wilson, E., and Rosenfeld, R.G. (1996).

- 1105 Synthesis and characterization of insulin-like growth factor-binding protein (IGFBP)-7.
1106 Recombinant human mac25 protein specifically binds IGF-I and -II. *J. Biol. Chem.* *271*, 30322–
1107 30325.
- 1108 Okae, H., Toh, H., Sato, T., Hiura, H., Takahashi, S., Shirane, K., Kabayama, Y., Suyama, M.,
1109 Sasaki, H., and Arima, T. (2018). Derivation of human trophoblast stem cells. *Cell Stem Cell* *22*,
1110 50-63.e6.
- 1111 Osafune, K., Caron, L., Borowiak, M., Martinez, R.J., Fitz-Gerald, C.S., Sato, Y., Cowan, C.A.,
1112 Chien, K.R., and Melton, D.A. (2008). Marked differences in differentiation propensity among
1113 human embryonic stem cell lines. *Nat. Biotechnol.* *26*, 313–315.
- 1114 Overgaard, M.T., Boldt, H.B., Laursen, L.S., Sottrup-Jensen, L., Conover, C.A., and Oxvig, C.
1115 (2001). Pregnancy-associated plasma protein-A2 (PAPP-A2), a novel insulin-like growth factor-
1116 binding protein-5 proteinase. *J. Biol. Chem.* *276*, 21849–21853.
- 1117 Parry, S., Holder, J., and Strauss, J.F. (1997). Mechanisms of trophoblast-virus interaction. *J.*
1118 *Reprod. Immunol.* *37*, 25–34.
- 1119 Petropoulos, S., Edsgård, D., Reinius, B., Deng, Q., Panula, S.P., Codeluppi, S., Plaza Reyes,
1120 A., Linnarsson, S., Sandberg, R., and Lanner, F. (2016). Single-Cell RNA-Seq Reveals Lineage
1121 and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell* *165*, 1012–1026.
- 1122 Pique-Regi, R., Romero, R., Tarca, A.L., Luca, F., Xu, Y., Alazizi, A., Leng, Y., Hsu, C.-D., and
1123 Gomez-Lopez, N. (2020). Does the human placenta express the canonical cell entry mediators
1124 for SARS-CoV-2? *Elife* *9*.
- 1125 Porazinski, S., Wang, H., Asaoka, Y., Behrndt, M., Miyamoto, T., Morita, H., Hata, S., Sasaki,
1126 T., Krens, S.F.G., Osada, Y., et al. (2015). YAP is essential for tissue tension to ensure
1127 vertebrate 3D body shape. *Nature* *521*, 217–221.
- 1128 Roberts, R.M., Ezashi, T., Sheridan, M.A., and Yang, Y. (2018). Specification of trophoblast
1129 from embryonic stem cells exposed to BMP4. *Biol. Reprod.* *99*, 212–224.
- 1130 Rossant, J. (2018). Genetic control of early cell lineages in the mammalian embryo. *Annu. Rev.*
1131 *Genet.* *52*, 185–201.
- 1132 Sawada, T., Benjamin, K.J.M., Brandtjen, A.C., Tietze, E., Allen, S.J., Paquola, A.C.M.,
1133 Kleinman, J.E., Hyde, T.M., and Erwin, J.A. (2020). Generation of four postmortem dura-derived
1134 iPS cell lines from four control individuals with genotypic and brain-region-specific transcriptomic
1135 data available through the BrainSEQ consortium. *Stem Cell Res.* *46*, 101806.
- 1136 Schneider, M.D., Gaussin, V., and Lyons, K.M. (2003). Tempting fate: BMP signals for cardiac
1137 morphogenesis. *Cytokine Growth Factor Rev.* *14*, 1–4.
- 1138 Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N.,
1139 Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated
1140 models of biomolecular interaction networks. *Genome Res.* *13*, 2498–2504.
- 1141 Singh, M., Bansal, V., and Feschotte, C. (2020). A single-cell RNA expression map of human
1142 coronavirus entry factors. *BioRxiv*.
- 1143 Sonawane, A.R., Platig, J., Fagny, M., Chen, C.-Y., Paulson, J.N., Lopes-Ramos, C.M., DeMeo,

- 1144 D.L., Quackenbush, J., Glass, K., and Kuijjer, M.L. (2017). Understanding Tissue-Specific Gene
1145 Regulation. *Cell Rep.* *21*, 1077–1088.
- 1146 Stirparo, G.G., Boroviak, T., Guo, G., Nichols, J., Smith, A., and Bertone, P. (2018). Integrated
1147 analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-
1148 implantation epiblast. *Development* *145*.
- 1149 Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y.,
1150 Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell
1151 Data. *Cell* *177*, 1888-1902.e21.
- 1152 Suryawanshi, H., Morozov, P., Straus, A., Sahasrabudhe, N., Max, K.E.A., Garzia, A., Kustagi,
1153 M., Tuschl, T., and Williams, Z. (2018). A single-cell survey of the human first-trimester placenta
1154 and decidua. *Sci. Adv.* *4*, eaau4788.
- 1155 Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S.,
1156 and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science*
1157 *282*, 1145–1147.
- 1158 Vallier, L., Touboul, T., Chng, Z., Brimpari, M., Hannan, N., Millan, E., Smithers, L.E., Trotter,
1159 M., Rugg-Gunn, P., Weber, A., et al. (2009). Early cell fate decisions of human embryonic stem
1160 cells and mouse epiblast stem cells are controlled by the same signalling pathways. *PLoS ONE*
1161 *4*, e6082.
- 1162 Vento-Tormo, R., Efremova, M., Botting, R.A., Turco, M.Y., Vento-Tormo, M., Meyer, K.B., Park,
1163 J.-E., Stephenson, E., Polański, K., Goncalves, A., et al. (2018). Single-cell reconstruction of the
1164 early maternal-fetal interface in humans. *Nature* *563*, 347–353.
- 1165 Vivanti, A.J., Vauloup-Fellous, C., Prevot, S., Zupan, V., Suffee, C., Do Cao, J., Benachi, A.,
1166 and De Luca, D. (2020). Transplacental transmission of SARS-CoV-2 infection. *Nat. Commun.*
1167 *11*, 3572.
- 1168 Wakeland, A.K., Soncin, F., Moretto-Zita, M., Chang, C.-W., Horii, M., Pizzo, D., Nelson, K.K.,
1169 Laurent, L.C., and Parast, M.M. (2017). Hypoxia Directs Human Extravillous Trophoblast
1170 Differentiation in a Hypoxia-Inducible Factor-Dependent Manner. *Am. J. Pathol.* *187*, 767–780.
- 1171 Wang, R.-Y., Chen, L., Chen, H.-Y., Hu, L., Li, L., Sun, H.-Y., Jiang, F., Zhao, J., Liu, G.-M.-Y.,
1172 Tang, J., et al. (2013). MUC15 inhibits dimerization of EGFR and PI3K-AKT signaling and is
1173 associated with aggressive hepatocellular carcinomas in patients. *Gastroenterology* *145*, 1436-
1174 48.e1.
- 1175 Weinberger, L., Ayyash, M., Novershtern, N., and Hanna, J.H. (2016). Dynamic stem cell states:
1176 naive to primed pluripotency in rodents and humans. *Nat. Rev. Mol. Cell Biol.* *17*, 155–169.
- 1177 Xu, R.-H., Chen, X., Li, D.S., Li, R., Addicks, G.C., Glennon, C., Zwaka, T.P., and Thomson,
1178 J.A. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nat.*
1179 *Biotechnol.* *20*, 1261–1264.
- 1180 Xu, X., Wells, A.B., O'Brien, D.R., Nehorai, A., and Dougherty, J.D. (2014). Cell type-specific
1181 expression analysis to identify putative cellular mechanisms for neurogenetic disorders. *J.*
1182 *Neurosci.* *34*, 1420–1431.

- 1183 Yabe, S., Alexenko, A.P., Amita, M., Yang, Y., Schust, D.J., Sadovsky, Y., Ezashi, T., and
1184 Roberts, R.M. (2016). Comparison of syncytiotrophoblast generated from human embryonic
1185 stem cells and from term placentas. *Proc Natl Acad Sci USA* *113*, E2598-607.
- 1186 Yang, L., Soonpaa, M.H., Adler, E.D., Roepke, T.K., Kattman, S.J., Kennedy, M., Henckaerts,
1187 E., Bonham, K., Abbott, G.W., Linden, R.M., et al. (2008). Human cardiovascular progenitor
1188 cells develop from a KDR+ embryonic-stem-cell-derived population. *Nature* *453*, 524–528.
- 1189 Zhang, P., Li, J., Tan, Z., Wang, C., Liu, T., Chen, L., Yong, J., Jiang, W., Sun, X., Du, L., et al.
1190 (2008). Short-term BMP-4 treatment initiates mesoderm induction in human embryonic stem
1191 cells. *Blood* *111*, 1933–1941.