Transposable Elements Continuously Remodel the Regulatory Landscape, Transcriptome, and Function of Decidual Stromal Cells

Katelyn Mika\textsuperscript{1,3} and Vincent J. Lynch\textsuperscript{2,*}

\textsuperscript{1}Department of Human Genetics, The University of Chicago, 920 East 58th Street, CLSC 319C, Chicago, Illinois 60637, USA
\textsuperscript{2}Department of Biological Sciences, University at Buffalo, SUNY, 551 Cooke Hall, Buffalo, New York 14260, USA
\textsuperscript{3}Present address: Department of Organismal Biology and Anatomy, University of Chicago, 1025 E 57th Street, Chicago, Illinois 60637, USA

*Corresponding author: E-mail: vjlynch@buffalo.edu.
Accepted: 04 November 2022

Abstract
Gene expression evolution underlies the origin, divergence, and conservation of biological characters including cell-types, tissues, and organ systems. Previously we showed that large-scale gene expression changes in decidual stromal cells (DSCs) contributed to the origins of pregnancy in eutherians and the divergence of pregnancy traits in primates and that transposable elements likely contributed to these gene expression changes. Here we show that two large waves of TEs remodeled the transcriptome and regulatory landscape of DSCs, including a major wave in primates. Genes nearby TE-derived regulatory elements are among the most progesterone responsive in the genome and play essential roles in orchestrating progesterone responsiveness and the core function of decidual cells by donating progesterone receptor binding sites to the genome. We tested the regulatory abilities of 89 TE consensus sequences and found that nearly all of them acted as repressors in mammalian cells, but treatment with a histone deacetylase inhibitor unmasked latent enhancer functions. These data indicate that TEs have played an important role in the development, evolution, and function of primate DSCs and suggest a two-step model in which latent enhancer functions of TEs are unmasked after they lose primary repressor functions.

Key words: endometrial stromal cell, decidual stromal cell, pregnancy, mammal, transposable element.

Significance
Gene expression patterns evolved very rapidly during the evolutionary origins of pregnancy in early mammals and primates. These episodes of gene expression evolution appear to be driven by the domestication of transposable elements and are linked with important processes that establish and maintain pregnancy.

Introduction
Changes in gene regulatory evolution underlie the origin, divergence, and homology of cell-types, tissues, and organs. Thus, understanding the mechanisms of gene regulatory evolution is essential for understanding how anatomical systems evolve. One hypothesis of gene regulatory evolution proposes that genes gain and lose expression domains through a multi-step accumulation of small-scale mutations, such as point mutations or indels, that either create new or destroy old transcription factor binding sites (TFBS) (Stone and Wray 2001).
At the other extreme, a gene may evolve a regulatory element in a single step through the integration and cooption of a transposable element (TE) that harbors functional TFBS (El-Deiry et al. 1992; Jordan et al. 2003; Polak and Domany 2006; Wang et al. 2007; Bourque et al. 2008; Kunarso et al. 2010; Lynch et al. 2015). While there is ample data to support both models of gene regulatory evolution (reviewed in Wray (2007); Feschotte (2008)), important questions with both models remain (Bourque et al. 2018). For example: Do TEs integrate with regulatory abilities, or as “pre-regulatory elements” that need additional mutations to acquire regulatory functions? Do the functions of TE-derived regulatory elements change, or is their evolution constrained by their functions upon integration? And does the cooption of TEs into regulatory elements occur continuously or in a single wave? Answers to these (and other) questions are essential for understanding the contribution of TEs to gene regulatory evolution.

Extant mammals span several major evolutionary transitions during the origins, divergence, and conservation of pregnancy, including the origin of new cell-types in the uterine lining (endometrium), endometrial stromal fibroblasts (ESFs), and decidual stromal cells (DSCs) (Erkenbrack et al. 2018; Wu et al. 2020; Chavan et al. 2021), which mediate many of the maternal responses to pregnancy. For example, an essential step in the establishment and maintenance of pregnancy is the differentiation (decidualization) of ESFs into DSCs in response to progesterone acting through the progesterone receptor (PGR), the second messenger cyclic AMP (cAMP) acting through protein kinase A (PKA), the transcription factor FOXO1 (Gellersen and Brosens 2003; Kajihara et al. 2013), and, in some species, to fetal signals (Gellersen et al. 2007). Decidualization evolved in the stem lineage of Eutherian mammals (Mess and Carter 2006; Kin et al. 2014, 2015) and induces large-scale gene regulatory, cellular, and physiological changes in the endometrium that are essential for successful implantation and pregnancy in many eutherians, including humans.

We have previously shown that hundreds of genes gained and lost endometrial expression coincident with the origins of pregnancy and decidualization in early mammals (Lynch et al. 2015; Marinic et al. 2021), and that a second major episode of gene expression evolution occurred in primates (Mika, Marinic, et al. 2021, Mika, Whittington, 2021). This latter wave of gene expression evolution occurred coincident with the origin of numerous primate-specific female reproductive and pregnancy traits including menstruation (Burley 1979; Strassmann 1996; Finn 1998; Erema et al. 2012), decidualization in the absence of fetal signals (spontaneous decidualization) (Gellersen and Brosens 2003; Mess and Carter 2006; Gellersen et al. 2007; Kin et al. 2016, 2015; Carter and Mess 2017), deeply invasive placentas (Pijnenborg, Vercruyssse, Brosens 2011; Pijnenborg, Vercruyssse, Carter 2011, Carter et al. 2015; Soares et al. 2018), and a derived parturition signal (Csapo 1956; Csapo and Pinto-Dantas 1965). Remarkably TEs appear to have played an important role in the origins of pregnancy through cooption into progesterone-responsive cis-regulatory elements in DSCs and may have played a similar role during the evolution of primate-specific pregnancy traits, but we excluded these evolutionarily young TEs in our previous analyses (Lynch et al. 2015; Mika, Marinic, et al. 2021). Here we expand on our previous studies on ancient mammalian TEs to all TEs and show that successive waves of TEs have been coopted into progesterone-responsive cis-regulatory elements, including a rolling wave of TE cooption in primates. Genes with regulatory elements derived from primate-specific TEs are among the most strongly differentially regulated by progesterone, have essential roles in decidualization, and likely contribute to primate-specific pregnancy traits such as spontaneous decidualization. Finally, we tested 89 consensus TE sequences, as a proxy for ancestral TE sequences, and found that nearly all have dominant repressor functions and latent enhancer functions in mammalian cells. These data suggest that primate-specific TEs played an important role in gene regulatory evolution in primate DSCs. Furthermore, our data suggest a general two-stage model of TE domestication into gene regulatory elements, whereby loss of ancestral repressor functions unmask hidden enhancer functions.

Results

Transposable Elements are Major Contributors to Regulatory Elements in DSCs

We have previously shown that Mammalian-, Therian-, and Eutherian-specific TEs played an important role in the origin of new cis-regulatory elements in DSCs during the evolution of pregnancy (Lynch et al. 2015)) and are enriched nearby genes that gained and lost expression in primate DSCs (Lynch et al. 2015; Mika, Marinic, et al. 2021). Here we expanded these studies to all classes and ages of TEs using previously generated H3K4me3 ChIP-Seq, H3K27ac ChIP-Seq, FAIRE-Seq, and DNasel-Seq data to identify promoters, enhancers, and regions of open chromatin (Lynch et al. 2015; Mika, Marinic, et al. 2021). We found that 58.7% of H3K27ac and 53.0% of H3K4me3 ChIP-Seq peaks, 42.2% of FAIRE-Seq peaks, and 67.2% of DNasel-Seq peaks overlapped annotated transposable elements (fig. 1A). Next, we annotated these TEs by their lineage specificity and found that TEs from different age classes differentially contributed to each kind of regulatory element: relatively young (i.e., Primate-specific) TEs dominated the DNasel, H3K27ac, and H3K4me3 datasets, whereas relatively ancient TEs (i.e., Eutherian-specific and older) were more common in the FAIRE dataset (fig. 1B). 427 TE families were enriched (eTE; > 1.5-fold, P ≤ 0.05, binomial test) within H3K27ac and H3K4me3 ChIP-Seq, and
Transposable Elements Continuously Remodel the Regulatory Landscape

FAIRE-, DNase-Seq peaks (fig. 1C), most of which were Eutherian- and Primate-specific (fig. 1D).

TE-derived Regulatory Elements are Enriched in Transcription Factor Binding Sites that Regulates DSCs

To determine if TEs donated motifs for specific transcription factors, we identified over-represented TFBS within eTE-derived regions of FAIRE-seq, DNase-seq, H3K27ac ChIP-seq, and H3K4me3 ChIP-seq peaks using previously published ENCODE ChIP-seq data for 132 transcription factors as well as previously published PGR ChIP-seq data generated from human DSCs (Mazur et al. 2015). Fifty-three TFBSs were enriched within regulatory eTEs relative to genomic TFBS abundances (FDR = 0.05; fig. 2A), most notably PGR (enrichment = 10.31, FDR < 1.00 × 10^{-250}), AHR (enrichment = 2.00, FDR = 1.00 × 10^{-15}), and GATA (enrichment = 1.25, FDR = 1.30 × 10^{-20}). We also observed enrichment for the KRAB-ZFPs ZNF263 (enrichment = 1.71, FDR = 7.20 × 10^{-150}) and ZNF274 (enrichment = 1.22, FDR = 8.90 × 10^{-3}), KAP1 (also known as TRIM28; enrichment = 2.00, FDR = 7.20 × 10^{-41}), which binds KRAB-ZFPs and functions as a scaffold for the recruitment of histone modifying co-repressor complexes, and parts of the SWI/SNF chromatin remodeling complex such as BAF155 (enrichment = 2.36, FDR = 1.00 × 10^{-78}), BAF170 (enrichment = 1.93, FDR = 4.20 × 10^{-14}), INI1 (enrichment = 1.43, FDR = 9.10 × 10^{-14}), and BRG1 (enrichment = 1.38, FDR = 2.70 × 10^{-5}). These data suggest that TEs have donated binding sites for transcription factors that mediate...
decidualization such as PGR and its obligate co-factor GATA2 (Rubel et al. 2016, 2011), as well as general transcriptional repressors and chromatin modifying proteins (fig. 2B).

TE-derived Regulatory Elements Augment Ancient Progesterone Responsiveness

Our observation that specific TE families are enriched in DSC regulatory elements suggests that they may contribute to gene expression changes that occur during progesterone-induced decidualization and changes in decidualization-induced gene expression during human evolution. To test this hypothesis, we used parsimony to reconstruct the evolutionary history of gene expression in the pregnant uterus (fig. 3A) (Marinić et al. 2021), RNA-Seq data from human ESFs and DSCs to quantify gene expression changes induced by decidualization (fig. 3B), and previously published promoter capture HiC (pChHiC) data generated from human DSCs (Sakabe et al. 2020) to associate genes with putative regulatory elements (fig. 3C). Next, we used an F-test to compare the change in expression level (variance) between ESFs and DSCs between genes with and without TE-derived regulatory elements. We found that genes with eTE-derived regulatory elements were generally more strongly differentially regulated by decidualization than genes without eTE-derived regulatory elements (i.e., had greater expression variance), especially genes that were more recently recruited into endometrial expression (fig. 3D).

TE-derived Regulatory Elements Are Enriched in PGR Binding Sites

Our observations that TEs are enriched in PGR binding sites and are associated with genes that are strongly differentially expressed upon decidualization prompted us to explore the contribution of TEs to PGR binding sites in greater detail. We found that 62.8% (5344/8510) of PGR ChIP-Seq peaks in DSCs contained TEs (fig. 4A), nearly all of which are Mammalian-, Eutherian-, and Primate-specific (fig. 4B). PGR ChIP-Seq peaks, however, are almost exclusively enriched (>1.5-fold, P ≤ 0.05, binomial test) in Eutherian- and Primate-specific TEs (fig. 4C). Consistent with a functional role for TE-derived PGR binding sites in orchestrating progesterone responsiveness, genes associated with TE-derived PGR binding sites by pChHiC (fig. 4D) were more strongly differentially expressed during decidualization than genes not associated with TE-derived PGR binding sites (fig. 4E); This trend was more pronounced for recently recruited genes (fig. 4F). We also found that genes associated by pChHiC with TE-derived PGR binding sites were significantly more dysregulated by siRNA-mediated PGR knockdown in DSCs (fig. 4F) than genes not associated with TE-derived PGR binding sites (fig. 4G).

TE-derived PGR Binding Sites in Primates Regulate Genes Essential for Decidualization

To explore the functional consequences of eTE-derived PGR binding sites, we first binned eTEs into ancient mammalian or primate-specific categories based on our observation that TEs in regulatory elements are predominantly from two age classes. Next, we used the pChHiC data to associate genes with each category of eTE-derived PGR binding site and tested whether these genes were enriched in biology pathways (KEGG, Panther, Wikipathway, Reactome) using the over-representation analyses implemented in WebGestalt (Liao et al. 2019). Of the 965 genes associated with eTE-derived PGR binding sites expressed in DSCs (TPM ≥ 2), 631 were associated with primate-specific eTE-derived PGR binding sites. These genes were enriched in 113 pathways at FDR ≤ 0.10. Among the enriched pathways were
many that play important roles in decidualization and pregnancy (table 1), such as Wnt, FoxO, and prolactin signaling, and various pathways related to regulation of the cell cycle which play a critical role in the earliest stages of decidualization (table 2). In contrast, the 334 genes only associated with ancient mammalian eTE-derived PGR binding sites were not enriched in any pathway at FDR ≤ 0.10.

Consensus TEs are Repressors With Latent Enhancer Potential

It is not clear if TEs integrate with regulatory abilities, and therefore immediately function as regulatory elements, or if they integrate as “pre-regulatory elements’ that are weakly- or non-functional and require additional mutations to acquire regulatory functions. Alternatively, TEs might not initially have direct effects on the expression of nearby genes, rather they might have collateral effects on host gene transcription by being targets of TE silencing where a TE is bound by a host repressive transcription factor, thereby causing the silencing of nearby gene (Fueyo et al. 2002). To test these scenarios, we selected 89 enriched TEs, synthesized their consensus sequences (conTE), and cloned them into the pGL3-Basic[minP] luciferase reporter vector. Next, we transiently transfected human ESFs and DSCs with each conTE reporter and used a dual luciferase reporter assay to test their regulatory abilities. We found that 55 (62%) the conTE reporters functioned as repressors in ESFs while 58 (65%) functioned as repressors in DSC (FDR ≤ 0.05), in contrast only 13 (3%) and 21 (23%) had...
enhancer functions in ESFs and DSCs (fig. 5A; FDR ≤ 0.05). To test whether these effects were cell-type specific, we repeated the luciferase reporter assay in the human hepatocellular carcinoma cell line HepG2 and again observed that 59/89 (66%) were strong repressors (FDR ≤ 0.05) whereas 28 were enhancers (fig. 5A; FDR ≤ 0.05). To determine if these results were species-specific, we repeated the luciferase assay in mouse embryonic fibroblasts (MEFs) and observed that 51 (57%; FDR ≤ 0.05) of conTEs were repressors (34 were enhancers; FDR ≤ 0.05) whereas in elephant
dermal fibroblasts 21 (24%; FDR ≤ 0.05) were repressors (34 were enhancers; FDR ≤ 0.05) (fig. 5A). While some conTEs, such as LTR elements, which have strong internal promoters, had enhancer functions in all cell-types, significantly more were repressors than expected by chance in ESFs (Binomial \( P = 0.027 \)), DSCs (Binomial \( P = 1.00 \times 10^{-6} \)), HepG2 (Binomial \( P < 1.00 \times 10^{-6} \)), and MEFs (Binomial \( P = 0.027 \)).

Our observation that TEs enriched within regulatory elements and open chromatin are also enriched in binding sites for KRAB-ZFPs, KAP1, as well as histone modifying co-repressor complexes, suggest that conTEs may function as repressors because they are recognized by the host cell anti-TE machinery and silenced (Matsui et al. 2010; Rowe et al. 2010; Feschotte and Gilbert 2012). To test this hypothesis, we repeated the luciferase reporter assay in ESFs and DSCs and treated cells with the mammalian class I and II histone deacetylase (HDAC) inhibitor trichostatin A (TSA). TSA treatment de-repressed 30/35 (86%; FDR ≤ 0.05) and 39/55 (71%; FDR ≤ 0.05) of the conTEs that were repressors in untreated ESFs and DSCs, respectively (fig. 5B). TSA treatment also unmasked the latent enhancer functions of 29 conTEs in ESFs and 26 conTEs in DSCs (e.g.: MIR3, MIRb, and MIRc, MamRep1879, and MER96b). Thus, while the majority of conTEs repress luciferase expression, their repressive abilities are likely HDAC dependent.

The preponderance of repressor functions among conTEs and their dependence on HDACs suggests that conTEs may be recognized by the KRAB-ZFPs/KAP1 TE suppression system and silenced by histone modifying co-repressor complexes (NuRD, CoREST, SWI/SNF) (Matsui et al. 2010; Rowe et al. 2010; Feschotte and Gilbert 2012).

### Table 1

Pathways in Which Genes Regulated by Primate-specific eTE-derived PGR Binding sites are Enriched

| Rank | Pathway                                      | E    | \( P \)  | FDR   | Reference         |
|------|----------------------------------------------|------|---------|-------|-------------------|
| 2    | VEGFA-VEGFR2 Signaling Pathway               | 2.60 | 2E-8    | 2.75E-5 | (Kim et al. 2013) |
| 6    | Insulin signaling pathway                    | 3.79 | 6E-7    | 2.72E-4 | (Kawamura et al. 2009) |
| 13   | TGF-beta Signaling Pathway                   | 3.52 | 6E-6    | 1.38E-3 | (Kim et al. 2005) |
| 17   | ErbB Signaling Pathway                       | 3.91 | 3E-5    | 4.28E-3 | (Klonisch et al. 2001) |
| 26   | Hippo signaling pathway                      | 3.02 | 5E-5    | 5.32E-3 | (Chen et al. 2017) |
| 27   | PDGF signaling pathway                       | 3.28 | 5E-5    | 5.41E-3 | (Schwenke et al. 2013) |
| 35   | FoxO signaling pathway                       | 3.11 | 9E-5    | 7.84E-3 | (Kajihara et al. 2013) |
| 37   | Mitophagy                                    | 4.21 | 1E-4    | 8.71E-3 | (Mestre Citrinovitz et al. 2019) |
| 38   | Oncostatin M Signaling Pathway               | 4.21 | 1E-4    | 8.71E-3 | (Fu et al. 2019) |
| 40   | TGF-beta Receptor Signaling                  | 4.56 | 1E-4    | 9.63E-3 | (Ni and Li 2017) |
| 52   | BDNF signaling pathway                       | 2.85 | 3E-4    | 1.40E-2 | (Kawamura et al. 2009) |
| 54   | EGF/EGFR Signaling Pathway                   | 2.70 | 3E-4    | 1.54E-2 | (Large et al. 2014) |
| 55   | Wnt signaling pathway                        | 2.81 | 3E-4    | 1.54E-2 | (Zhang and Yan 2016) |
| 65   | Prolactin Signaling Pathway                  | 3.60 | 4E-4    | 1.90E-2 | (Bao et al. 2007) |
| 74   | Myometrial Relaxation and Contraction Pathways | 2.63 | 6E-4    | 2.32E-2 | (Salomonis et al. 2005) |

E, enrichment ratio. \( P \), hypergeometric \( P \)-value. FDR, Benjamini–Hochberg false discovery rate. Reference, reference for role of pathway in decidualization.

### Table 2

Pathways Related to Senescence and the Cell cycle in Which Genes Regulated by Primate-specific eTE-derived PGR Binding sites are Enriched

| Rank | Pathway                                      | E    | \( P \)  | FDR   | Reference         |
|------|----------------------------------------------|------|---------|-------|-------------------|
| 11   | Cellular senescence                          | 3.25 | 6E-6    | 1.38E-3 | (Lucas et al. 2020) |
| 14   | DNA Damage Response (only ATM dependent)     | 3.73 | 1E-5    | 2.23E-3 | (Lei et al. 2012) |
| 19   | G2/M Transition                              | 2.82 | 3E-5    | 4.37E-3 | (Logan et al. 2012) |
| 20   | Mitotic G2-G2/M phases                       | 2.79 | 3E-5    | 4.70E-3 | (Logan et al. 2012) |
| 41   | Apoptosis signaling pathway                  | 3.29 | 2E-4    | 1.08E-2 | (Joswig et al. 2003) |
| 42   | Oxidative Stress                             | 5.80 | 2E-4    | 1.08E-2 | (Yu et al. 2019) |
| 59   | Mitotic G1-G1/S phases                       | 2.75 | 4E-4    | 1.79E-2 | (Logan et al. 2012) |
| 66   | Cell Cycle                                   | 2.96 | 4E-4    | 1.92E-2 | (Das 2009) |
| 67   | Senescence and Autophagy in Cancer           | 3.12 | 4E-4    | 1.92E-2 | (Deryabin et al. 2020) |
| 75   | Cell cycle                                   | 2.87 | 6E-4    | 2.32E-2 | (Das 2009) |
| 91   | Cyclin D associated events in G1             | 4.35 | 1E-3    | 3.09E-2 | (Logan et al. 2012) |
| 92   | G1 Phase                                     | 4.35 | 1E-3    | 3.09E-2 | (Logan et al. 2012) |
| 99   | p53 signaling pathway                        | 3.42 | 1E-3    | 3.45E-2 | (Deng et al. 2016) |

E, enrichment ratio. \( P \), hypergeometric \( P \)-value. FDR, Benjamini–Hochberg false discovery rate. Reference, reference for role of pathway in decidualization.
To test this hypothesis, we repeated the luciferase reporter assay in KAP1 knockout mouse embryonic fibroblasts (MEF KAP1<sup>−/−</sup>), but again observed 53 conTEs (59%; FDR ≤ 0.05) were repressors and relief of repression for only 8 (9%; FDR ≤ 0.05) conTEs (fig. 5A). To infer if repression may be mediated by KRAB-ZFPs, we took advantage of their restricted lineage specificity and tested the regulatory abilities of conTEs in chicken embryonic fibroblasts. The chicken genome only encodes 41 KRAB-ZFP, none of which are expressed in chicken embryonic fibroblasts (Addison et al. 2015). In stark contrast to the different mammalian cell-types we tested, only 18 (20%; FDR ≤ 0.05) of the conTEs functioned as repressors whereas 69 (78%; FDR ≤ 0.05) were strong enhancers in CEFs (fig. 5A).

Finally, we used multi-dimensional scaling to explore whether conTEs could be grouped into distinct clusters based on their regulatory abilities in different cell-types. We found that conTEs formed four clusters in the MDS plot, corresponding to those with weak, moderate, and strong regulatory abilities (fig. 5C).

Fig. 5—Consensus TEs are repressors with hidden enhancer potential. (A) The regulatory ability of consensus TEs was assessed using a dual luciferase reporter vector. Each box represents the mean luminescence values (log) of 4–6 replicate reporter assays. Basic[ΔminP], empty reporter vector reference. TAP2_C is a known progesterone-responsive enhancer<sup>44</sup>. The 89 TE constructs are sorted in order of decreasing age inferred by lineage specificity. Red indicates enhancer function and blue indicates repressor function (see inset color scale). ESF, Endometrial Stromal Fibroblasts; DSC, Decidual Stromal Cells; HEPG2, Human Liver Cancer; MEF Control, Mouse Embryonic Fibroblast, MEF KAP1<sup>−/−</sup>, Mouse KAP1 knockout Embryonic Fibroblast; Elephant, Elephant Dermal Fibroblast; Chicken, Chicken Embryonic Fibroblast; +TSA, Trichostatin A added to media. (B) volcano plot ESF versus ESF treated with TSA and DSC versus DSC treated with TSA (C) MDS plot of consensus TEs based on mean luminescence values (log) and grouped by K-means clustering (k = 4). K-means clustering groups TEs into four categories corresponding to TEs with strong/moderate/weak enhancer or repressor functions.
strong enhancer functions and those with repressor functions (fig. S5). Similarly, hierarchical clustering grouped conTEs with enhancer and repressor functions (supplementary fig. S1, Supplementary Material online). Consistent with our observation that the regulatory functions of conTEs were similar in different mammalian cell-types but different in CEFs, CEFs did not cluster with mammalian cell-types in an MDS plot (supplementary fig. S2, Supplementary Material online). Collectively these data indicate that while a few conTEs, mostly LTRs, have strong enhancer or promoter abilities, most conTEs function as HDAC-dependent repressors in mammalian cells. We note that these effects were independent of conTE length across all experiments indicating there is no bias for short or long TEs to act as repressors or enhancers of luciferase expression (supplementary fig. S1, Supplementary Material online).

Our observations that TEs are enriched binding sites for transcription factors important for DSC function and that consensus TEs have dominant repressor abilities and latent enhancer potential, suggest that these consensus TEs should have binding sites for repressors and activators. To explore this possibility, we used XTREME to discover over-represented motifs in conTEs. Consistent with the repressor abilities of conTEs, we found that zinc finger proteins were among the most numerous over-represented motifs, including matches binding sites for KAP1-interacting zinc finger proteins such as ZNF320, ZIC1-3, ZNF417, ZKSCAN5, and ZNF189, as well as repressors and co-repressors such as CTCF, EHF, ESRRA, ESRRB, and ESRRG (supplementary fig. S4–6, Supplementary Material online). We also found over-represented motifs that match binding sites for transcription factors important for regulating stromal cell identity and progesterone responsiveness, such as PGR, FOXO, and Hox. These data suggest that the ability of conTEs to act as repressors is dependent on pre-existing binding sites for transcription factors with repressor functions, either directly like ESRRA-G, indirectly though recruiting TE silencing repressors mediated by KRAB-ZFP containing proteins, or indirectly through regulating chromatin organization and positional information mediated by CTCF.

Discussion

Relationship to Previous Studies

We previously used a combination of comparative transcriptomics and genomics to show that ancient mammalian transposable elements, that is TEs that integrated into the genome of the last common ancestor of Mammalia, Theria, and Eutheria, were enriched within active regulatory elements and PGR binding sites in human DSCs (Lynch et al. 2015). However, in that study we focused exclusively on ancient mammalian TEs and used a computational method (GREAT), which assigns each gene a regulatory domain including a basal domain that extends 5 kb upstream and 1 kb downstream from its transcription start site and an extension up to the basal regulatory domain of the nearest upstream and downstream genes within 1 Mb, to associate TE-derived regulatory elements with genes. Here we considered all TE age classes, rather than just ancient mammalian ones, and used a direct experimental method, promoter capture HiC (pHiC), to directly associate TE-derived regulatory elements with genes expressed in human DSCs. This more expansive TE inclusion set and the direct association of genes with their long-range regulatory elements suggest transposable elements continuously remodel the regulatory landscape, transcriptome, and function of DSCs.

Two Waves of Transposable Element Co-option Remodel the Transcriptome and Regulatory Landscape of DSCs

Transposable elements are so frequently coopted into regulatory elements that it is not possible to cite all or even most studies reporting either the cooption of individual TEs or large-scale cooption of (almost) entire TE families. Previous studies, for example, have dissected in great detail the cis-regulatory element that drives extra-pituitary prolactin (PRL) in human DSCs, which is normally expressed by the pituitary and immune cell-types (Gerlo et al. 2006). These studies found that human PRL expression is initiated from an alternative promoter located 5.8 kb upstream of the canonical pituitary transcription start site which contains binding sites for transcription factors that are essential for the identity and function of DSCs including PGR, FOXO1A, ETS1, CEBPB, and FOS (Gerlo et al. 2006). Remarkably, the human decidual PRL promoter is derived from a primate-specific long terminal repeat (LTR)-like transposable element in the medium reiterated repeat (MER39) family and an upstream enhancer derived from a Eutherian-specific MER1 class of DNA transposon (MER20) (Gerlo et al. 2006). However, while there is little evidence that other MER39 elements function as promoters (Emera and Wagner 2012), MER20 elements function as enhancers in DSCs for numerous genes across the genome (Lynch et al. 2011).

These data suggest that Eutherian-specific transposable elements played a role in re-wiring the gene regulatory network during the evolution of pregnancy and decidualization. Consistent with this observation, we previously found large-scale cooption of Mammalian-, Therian-, and Eutherian-specific TEs (AncMamTEs) into progesterone-responsive cis-regulatory elements. Here we expanded on these studies and found that 427 TE families were enriched in DSC regulatory elements, nearly half of which were primate-specific. Thus, there were at least two waves of TE cooption into decidual regulatory elements—a first wave in early mammals and the second wave in primates. Remarkably,
genes associated by pHiC with TE-derived regulatory elements were significantly more responsive to progesterone than genes without TE-derived cis-regulatory elements. Furthermore, genes with primate-specific TE-derived PGR binding sites were more progesterone responsive than genes with either ancient TE-derived PGR binding sites or genes without TE-derived PGR binding sites. These data suggest that TEs may have played a role in re-wiring the progesterone-responsive gene regulatory network both during the evolution of pregnancy in early mammals and primates.

Primate-specific Transposable Elements Contribute to the Function of DSC

While eutherian mammals share a suite of traits mediated by endometrial stromal lineage cells that support prolonged pregnancies, there is also considerable variation in pregnancy traits within eutherians. Catarrhine primates, for example, have evolved spontaneous decidualization (differentiation) of ESFs into DSCs under the combined action of progesterone, cyclic adenosine monophosphate (cAMP), and other unknown maternal signals (Gellersen and Brosens 2003; Mess and Carter 2006; Gellersen et al. 2007; Kin et al. 2015, 2016; Carter and Mess 2017). Decidualization induces dramatic gene expression and functional changes (Giudice 2003; Gellersen et al. 2007; Aghajanova et al. 2011), but the molecular mechanisms that underlie the evolution of spontaneous decidualization are largely unknown. We found that coopted primate-specific TEs regulate genes in several pathways involved in decidualization and the cell cycle (Tables 1 and 2), most notably the FOXO1 signaling pathway. For example, the transcription factor FOXO1 plays a particularly important function as a key regulator of decidualization, which integrates cAMP and progesterone signaling through physical and functional interactions with the PGR and other transcription factors to direct expression of progesterone response genes (Gellersen and Brosens 2003; Takano et al. 2007; Lynch et al. 2009). Similarly, among the earliest steps in the decidualization process is cell cycle exit (Das 2009; Logan et al. 2010); progesterone initially induces cell cycle arrest at the G0/G1 checkpoint followed by arrest at the G2/M checkpoint, both of which are regulated by p53 signaling (Logan et al. 2012). These data suggest that primate-specific TEs may have played a role in the origin of spontaneous decidualization by altering the regulation of genes in the FOXO1 signaling and cell cycle regulation pathways.

A Multi-stage Model for Transposable Element Domestication

Numerous studies have shown that transposable elements have donated binding sites for transcription factors to the genome, can be bound by transcription factors, and have been coopted into cis-regulatory elements (El-Deiry et al. 1992; Jordan et al. 2003; Polak and Domany 2006; Wang et al. 2007; Kunarso et al. 2010; Lynch et al. 2011), but it generally has not been determined if TEs integrate into the genome with regulatory functions and therefore immediately function as regulatory elements or if they integrate as “pre-regulatory elements’ that are not immediately functional and require additional mutations to acquire regulatory functions. We addressed this question using luciferase assay functional tests of conTEs and found the majority of consensus TEs have enhancer abilities but these abilities are silenced in mammalian cells, perhaps by KRAB-ZFPs and/or the NURD HDAC inhibitory complex. We do see a mild enrichment within the dataset of TEs with regulatory marks in DSCs. This enrichment is likely low because 1) the elements analyzed have likely escaped repression by KRAB-ZFPs in order to be coopted into enhancers, and 2) the datasets used to identify KRAB-ZFP, KAP1, and the majority of other transcription factors, were from ENCODE and not DSCs. However, the relief of repression seen in chicken cells, which do not express KRAB-ZFPs (Addison et al. 2015), is at least coincidental evidence that KRAB-ZFPs may play a role in silencing conTEs. In embryonic stem cells, KAP1 binds the KRAB-ZFPs and coordinates the silencing of the bound TE (Matsui et al. 2010; Rowe et al. 2010; Feschotte and Gilbert 2012). However, in adult somatic tissues, the role of KAP1 is unclear and TE silencing in these tissues may be independent of KAP1 but require other housekeeping co-repressors (Matsui et al. 2010; Rowe et al. 2010; Ecco et al. 2016). In this model KRAB-ZFPs bind TEs in somatic cells, but silence TEs by recruiting repressors other than KAP1. Thus our observations suggest that silencing of the conTEs in MEFs may require KRAB-ZFPs but be KAP1 independent. These results need to be confirmed, for example, by ChIP to demonstrate conTEs are indeed bound by KRAB-ZFPs. We hypothesize that only when TEs escape this silencing regulation are they coopted by the genome to play a regulatory role (fig. 6).

Caveats and Limitations

Ideally, we would use an ancestral sequence reconstruction (ASR) of each TE to determine if TEs had ancestral regulatory abilities. However, reconstructing the earliest ancestral sequence for most TEs is not possible because most TEs do not have an outgroup to root their phylogeny and therefore we cannot identify which node is the deepest ancestor. In place of an ASR, we and others who have explored similar questions used conTEs. This introduces an obvious limitation to our inferences: the conTE may not be an accurate representation of the deepest ancestral sequence. Indeed, there is no guarantee that a consensus sequence represents a sequence that ever existed. Another possible limitation of
our approach is the use of immortalized ESFs maintained under standard tissue conditions for both functional genomic studies and luciferase assays, which may not faithfully represent in vivo functions. These limitations impact virtually every study of primate pregnancy, however, endometrial organoids and iPSC-derived ESFs are promising systems in which to study the evolution of pregnancy (Boretto et al. 2017; Turco et al. 2017; Abbas et al. 2020; Marinic et al. 2020; Rawlings et al. 2021). Another limitation of our experimental approach is that the reporter assays used an episomal vector rather than a construct inserted into the genome which may bias our results by removing context dependent effects of regulatory elements within the genome and it is possible that TSA treatment induces the activation of a general transcriptional activator leading to up-regulation of luciferase expression upon TSA treatment. Finally, we have not performed the kinds of empirical studies necessary to confirm a causal relationship between TEs and gene expression levels, however, our results provide strong circumstantial evidence that TEs have influenced the gene expression and function of decidual cells.

Conclusions

Here we demonstrate that there were two major waves of TE cooption into regulatory elements in endometrial stromal lineage cells, specifically in ancient mammals and primates, and that these TEs donated functional TFBS to the genome. Furthermore, the majority of the TEs tested have context-dependent repressor and enhancer functions, suggesting they may have integrated into the genome with regulatory abilities and that their genomic functions are cell-type dependent. Genes regulated by TE-derived regulatory elements are among the most progesterone responsive in the genome and are associated with essential functions of stromal cells and the processes of decidualization. These data suggest that TEs have played an important role in the evolution of gene regulation and function of endometrial stromal lineage cells. This may have had a particularly significant effect on primates, which have divergent pregnancy traits than other mammals.

Materials and Methods

Identification of TE Containing Regulatory Elements

We used previously published ChIP-Seq data generated from human DSCs that were downloaded from NCBI SRA and processed remotely using Galaxy (Afgan et al. 2016). ChIP-Seq reads were mapped to the human genome (GRCh37/hg19) using HISAT2 (Pertea et al. 2016; Kim et al. 2015, 2019) with default parameters and peaks called with MACS2 (Zhang et al. 2008; Feng et al. 2012) with default parameters. Samples included H3K4me3 (GSE61793), H3K27ac (GSE61793), PGR (GSE69539), the PGR A and B isoforms (GSE62475), and DNase1-Seq (GSE61793). FAIRE-Seq peaks were downloaded from the UCSC genome browser and not re-called.

To identify regulatory elements derived from transposable elements (TEs), peaks were intersected with the RepMask 3.2.7 track at the UCSC genome browser. Non-transposable element annotations were removed and corrected for fragmented annotations. To identify TEs that were significantly enriched within TE-derived peaks the TEanalysis pipeline (https://github.com/4ureliek/TEanalysis) was used with 10,000 replicates. A custom bioinformatic pipeline was used to determine the enrichment of TFBS in TEs that intersect with the DSC FAIRE-seq, DNase-seq, H3K27ac ChIP-Seq, and H3K4me3, and PGR ChIP-Seq peaks (https://github.com/4ureliek/TEanalysis) versus the genomic abundance of ChIP-Seq peaks; 10,000 bootstrap reshuflings were used to assess statistical significance. Scripts are publicly available and archived at https://github.com/4ureliek/TEanalysis. The location of ChIP-Seq peaks in hg19 ENCODE data was downloaded...
from the USCS genome browser (Txn Fac ChIP V2—Transcription Factor ChIP-Seq from ENCODE (V2)). The location of PGR ChIP-Seq peaks was obtained from GEO (GSE94036) and is available (Mazur et al. 2015). We also used previously published promoter capture HiC (pCHIC) generated from DSCs (Sakabe et al. 2020) to associate genes with regulatory elements (SDY1626) and did not reanalyze these data.

Gene Expression Data and Parsimony Reconstruction of Gene Expression Gain/Loss

We used previously generated RNA-Seq data from the pregnant or gravid uterus of amniotes (supplementary table S1, Supplementary Material online) and Kallisto version 0.42.4 (Bray et al. 2016) to quantify gene expression levels. Kallisto was run with default parameters, bias correction, 100 bootstrap replicates. Kallisto outputs are gene expression levels in transcripts per million (TPM). Our previous studies of endometrial gene expression data suggest that genes with a TPM ≤ 2 are likely from transcriptionally suppressed genes (Wagner et al. 2013). This threshold is consistent with one obtained by comparing the transcript abundance with the chromatin state of the respective gene (Hebenstreit et al. 2011) and classified genes with TPM ≥ 2 as expressed and those with TPM < 2 as not expressed. Next, we used Mesquite (v2.75) and parsimony optimization to reconstruct ancestral gene expression states and identified genes that gained or lost endometrial expression in Amniotes. Expression was classified as an unambiguous gain if a gene was not inferred as expressed at the ancestral node (state 0) but inferred as expressed in a descendent node (state 1) and vice versa for the classification of a loss from endometrial expression.

Gene Expression Data From Human ESFs and DSCs

We used previously published RNA-Seq gene expression data generated from human primary human ESFs treated for 48 hr with control non-targeting and PGR-targeting siRNA prior to decidualization stimulus for 72 hr (GSE94036) to compare treatments (ESF/DSC and DSC + control siRNA/DSC + PGR siRNA) and gene sets (+TE−TE). Data were downloaded from NCBI SRA and processed remotely using the Galaxy platform (https://usegalaxy.org/; Version 20.01). RNA-Seq datasets were transferred from SRA to Galaxy using the Download and Extract Reads in FASTA/Q format from the NCBI SRA tool (version 2.10.4 + galaxy1). We used HISAT2 (version 2.1.0 + galaxy5) to align reads to the Human hg38 reference genome using single- or paired-end options depending on the dataset and unstranded reads, and report alignments tailored for transcript assemblers including StringTie. Transcripts were assembled and quantified using StringTie (v1.3.6) (Pertea et al. 2016, 2015), with reference file to guide assembly and the “reference transcripts only” option, and output count files for differential expression with DESeq2/edgeR/limma-voom. Differentially expressed genes were identified using DESeq2 (Love et al. 2014) (version 2.11.40.6 + galaxy1). The reference file for StringTie guided assembly was wgEncodeGencodeBasicV33. Comparisons between Log2 fold change in gene expression between ESFs and DSCs were estimated by identifying genes with and without TE-derived regulatory elements and comparing the variance in fold change expression between these two sets with an F-test in R.

Over-Representation Analyses

We used WebGestalt v. 2019 (Liao et al. 2019) to identify enriched ontology terms using over-representation analysis (ORA). We used ORA to identify enriched terms for three pathway databases (KEGG, Reactome, and Wikipathway), three disease databases (Disgenet, OMIM, and GLAD4U), and a custom database of genes implicated in preterm birth by GWAS. The preterm birth gene set was assembled from the NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog), including genes implicated in GWAS with either the ontology terms “Preterm Birth” (EFO_0003917) or “Spontaneous Preterm Birth” (EFO_0006917), as well as two recent preterm birth GWAS (Warrington et al. 2019; Sakabe et al. 2020) using a genome-wide significant P-value of 9 × 10−6. The custom gmt file used to test for enrichment of preterm birth associated genes is included as a supplementary data file, Supplementary Material online to figure 2 (Source data 1).

Transposable Element Reporter Vectors

To generate luciferase reporter vectors for functional testing, we selected 79 of the 427 TE families enriched within DSC regulatory regions as marked H3K4me3 ChIP-seq, H3K27ac ChIP-seq, DNasel-seq, and FAIRE-Seq datasets. These 79 TEs also met the following criteria: 1) At least 1 element from every lineage; and 2) represented all four classes of TEs (LTR, LINE, SINE, and DNA). Consensus sequences for these elements were taken from the database Dfam (Wheeler et al. 2013). 10 additional TEs also found in the enriched 427 were then chosen that were unique to Old World Monkeys or younger and their consensus sequences were also obtained from Dfam (Wheeler et al. 2013). The total set of elements were biased towards, but not limited to, the DNA class of transposable elements. These 89 consensus sequences were then synthesized by Genscript and cloned into the pGL3-Basic vector (Promega) with an added minimal promoter (pGL3Basic[minP]) and are available in the supplementary materials, Supplementary Material online.

Cell Lines

Human hTERT-immortalized ESFs were purchased from ATCC (CRL-4003), and their identity has been authenticated by the Lynch lab to
be mycoplasma free. KAP1 knockout MEFs and Flox/Flox control MEFs were gifts from D. Trono (Ecole Polytechnique Fédérale de Lausanne), their identity has been authenticated by the Trono lab and was determined by the Lynch lab to be mycoplasma free. Chicken embryonic fibroblast cells were purchased from ATCC (CRL-12203), their identity has been authenticated by ATCC, and was determined by the Lynch lab to be mycoplasma free. African elephant fibroblasts were a gift from the San Diego Frozen Zoo, their identity has been authenticated by the San Diego Frozen Zoo and were determined by the Lynch lab to be mycoplasma free.

Cell Culture and Luciferase Assays
ESF, KAP1 knockout MEFs, and Flox/Flox control MEFs were maintained in phenol red free DMEM (Gibco) supplemented with 10% charcoal stripped fetal bovine serum (CSFBS; Gibco), 1x ITS (Gibco), 1% sodium pyruvate (Gibco), and 1% L-glutamine (Gibco). Chicken embryonic fibroblast cells (ATCC CRL-12203) and HEPG2 cells were maintained in phenol red containing DMEM + Glutamax (Gibco) supplemented with 10% FBS (Gibco) and Normocin (InviviGen). Elephant fibroblasts were maintained in 1:1 MEM (Corning cellgro) supplemented with 10% FBS, 1% penstrep (Gibco), 1x sodium pyruvate (Gibco), and 1x L-glutamine (Gibco) to FGM-2 (Lonzia), made per manufacturer’s instructions. Confluent cells in 96-well plates in 80µl of Opti-MEM (Gibco) were transfected with 100 ng of the TE containing luciferase plasmid and 10 ng of the pRl-null renilla vector (Promega) with 0.1µl PLUS reagent (Invitrogen) and 0.25µl of Lipofectamine LTX (Invitrogen) in 20µl Opti-MEM. The cells were incubated in the transfection mixture for 6 hrs and the media was replaced with the maintenance media overnight. Decidualization of ESFs was then induced by incubating the cells in the decidualization media: DMEM with phenol red (Gibco), 2% CSFBS (Gibco), 1% sodium pyruvate (Gibco), 0.5 mM 8-Br-cAMP (Sigma), and 1 µM of the progesterone analog medroxyprogesterone acetate (Sigma) for 48 hr. ESFs (decidualization controls) were incubated in the decidualization control media (phenol red free DMEM (Gibco), 2% CSFBS (Gibco), and 1% sodium pyruvate (Gibco) instead for 48 hrs. For trichostatin A (TSA; Tocris Bioscience) trials, 1µM TSA was added to all the media from plating through decidualization. After decidualization for ESFs and DSCs or after 48 hrs from transfection for other cell types, Dual Luciferase Reporter Assays (Promega) were started by incubating the cells for 15 min in 20 µl of 1 x passive lysis buffer. Luciferase and renilla activity were then measured using the Glomax multi+ detection system (Promega). Luciferase activity values were standardized by the renilla activity values and background activity values as determined by measuring luminescence from the pGL3-Basic[minP] plasmid with no insert. Each luciferase experiment was replicated in 4–6 independent experiments. To identify significant expression shifts, we performed Wilcoxon tests on the data, and adjusted by the Benjamini and Hochberg (1995) method for multiple testing; significance was determined by having an adjusted P-value ≤ 0.05.

Data Exploration and MDS
We used classical MDS to explore the structure of luciferase assay data. MDS is a multivariate data analysis method that can be used to visualize the similarity/dissimilarity between samples by plotting data points onto two-dimensional plots. MDS returns an optimal solution that represents the data in a two-dimensional space, with the number of dimensions (k) specified a priori. Classical MDS preserves the original distance metric, between data points, as well as possible. MDS was performed using the vegan R package (Oksanen et al. 2019) with four reduced dimensions. Luciferase assay data were grouped using K-means clustering with K=2–6, K=4 optimized the number of distinct clusters and cluster memberships.

Motif Discovery in conTEs
We used XSTREME (Motif Discovery and Enrichment Analysis) version 5.5.0, implemented in the MEME suite of tools (https://meme-suite.org/meme/) to identify sequence motifs that were over-represented in conTE sequences using three runs: 1) Default settings; 2) Default settings relaxed to allow for up to 15 motifs; and 3) To specifically search for binding sites similar to PGR, we note that the PGR binding site is identical to the NR3C2 binding site. Discovered motifs were matched to known motifs in the JASPAR 2022 database. Exact search parameters are given in the output file supplementary figures S4–6, Supplementary Material online.

Supplementary material
Supplementary data are available at Genome Biology and Evolution online (http://www.gbe.oxfordjournals.org/).

Acknowledgements
The authors thank A. Kapusta for assistance with running the TE enrichment program, D. Trono (Ecole Polytechnique Fédérale de Lausanne) for providing the KAP KO and control MEFs, C. Brown (University of Pennsylvania) for providing the HEPG2 cells, and N. Sakabe (University at Chicago) for providing the pChIC interaction data. This study was supported by a grant...
from the March of Dimes (March of Dimes Prematurity Research Center to principal investigator V.J.L.), a Burroughs Welcome Fund Preterm Birth Initiative grant (1013760, to principal investigator V.J.L.), and a Burroughs Welcome Fund Next Gen Pregnancy Research Grant (1022513, to principal investigator V.J.L.). The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

**Author Contributions**
K.M. and V.J.L. collaboratively conceived, designed, interpreted data, and wrote this paper, as well as analyzed the expression shifts of genes associated with TE-derived regulators or TE-derived PGR binding sites in response to progesterone signaling and did the binding site enrichment analyses. K.M. analyzed the RNA-Seq data, performed the gene expression calls, completed the parsimony reconstruction, and prepped, conducted, and analyzed the luciferase assays and associated cell culture. V.J.L. identified the TE containing regulatory elements.

**Data Availability**
All data necessary to reproduce the results reported in this study are publicly available or available in the supplementary materials, including ChIP-Seq data for H3K4me3 (GSE61793), H3K27ac (GSE61793), PGR (GSE69539), the PGR A and B isoforms (GSE62475), and DNase1-Seq and FAIRE-Seq (GSE61793), peaks were downloaded from the UCSC genome browser and not re-called, promoter capture HiC data (SDY1626), and RNA-Seq gene expression data (GSE94036). All other data are provided as supplementary materials.

**Literature Cited**
Abbas Y, et al. 2020. Generation of a three-dimensional collagen scaffold-based model of the human endometrium. Interface Focus 10:20190079.
Addison JB, et al. 2015. KAP1 Promotes proliferation and metastatic progression of breast cancer cells. Cancer Res 75:344–355.
Afgan E, Baker D, et al. 2016. The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Res 44:W3–W10.
Aghajanova L, et al. 2011. Unique transcriptome, pathways, and networks in the human endometrial fibroblast response to progesterone in endometriosis. Biol Reprod 84:801–815.
Bao L, et al. 2007. Decidual prolactin silences the expression of genes detrimental to pregnancy. Endocrinology 148:2326–2334.
Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol 57:289–300.
Boretto M, et al. 2017. Development of organoids from mouse and human endometrium showing endometrial epithelium physiology and long-term expandability. Development 144:1775–1786.
Bourque G, et al. 2008. Evolution of the mammalian transcription factor binding repertoire via transposable elements. Genome Res 18:1752–1762.
Bourque G, et al. 2018. Ten things you should know about transposable elements. Genome Biol 19:199.
Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol 34:525–527.
Burley N. 1979. The evolution of concealed ovulation. Am Nat 114:835–858.
Carter AM, Mess AM. 2017. Placentation in the colugos cyancephalus volans and galeopterus variegatus (dermoptera) and the transition from labyrinthine to villous placentation in primates. Placenta 55:47–53.
Carter EB, Bishop KC, Goetzinger KR, Tuuli MG, Cahill AG. 2015. The impact of chorionocity on maternal pregnancy outcomes. Am J Obstet Gynecol 213:390.e1–390.e7.
Chavan AR, et al. 2021. Evolution of embryo implantation was enabled by the origin of decidual stromal cells in eutherian mammals. Mol Biol Evol 38:1060–1074.
Chen H, et al. 2017. YAP Mediates human decidualization of the uterine endometrial stromal cells. Placenta 53:30–35.
Csapo A. 1956. Progesterone “block.” Am J Anat 98:273–291.
Csapo A, Pinto-Dantas CA. 1965. The effect of progesterone on the human uterus. Proc Natl Acad Sci U S A 54:1069–1076.
Das SK. 2009. Cell cycle regulatory control for uterine stromal cell decidualization in implantation. Reprod Conv Engl 137:889–899.
Deng W, et al. 2016. P53 coordinates decidual sestrin 2/AMPK/mTORC1 signaling to govern parturition timing. J Clin Invest 126:2941–2954.
Denyapin P, Griukova A, Nikolsky N, Borodkina A. 2020. The link between endometrial stromal cell senescence and decidualization in female fertility: the art of balance. Cell Mol Life Sci CMLS 77:1357–1370.
Ecco G, et al. 2016. Transposable elements and their KRAB-ZFP controllers regulate gene expression in adult tissues. Dev Cell 36:611–623.
El-Deiry WS, Kern SE, Pieterpol JA, Kinzler KW, Vogelstein B. 1992. Definition of a consensus binding site for p53. Nat Genet 1:45–49.
Emera D, Romero R, Wagner G. 2012. The evolution of menstruation: a new model for genetic assimilation. BioEssays 34:26–35.
Emera D, Wagner GP. 2012. Transformation of a transposon into a derived prolactin promoter with function during human pregnancy. Proc Natl Acad Sci 109:11246–11251.
Erkenbrack EM, et al. 2018. The mammalian decidual cell evolved from a cellular stress response. PLoS Biol 16:e2005594.
Feng J, Liu T, Qin B, Zhang Y, Liu XS. 2012. Identifying ChIP-seq enrichment using MACS. Nat Protoc 7:1728–1740.
Feschotte C, Gilbert C. 2012. Endogenous viruses: insights into viral evolution and impact on host biology. Nat Rev Genet 13:283–296.
Feschotte C. 2008. Transposable elements and the evolution of regulatory networks. Nat Rev Genet 9:397–405.
Finn CA. 1998. Menstruation: a nonadaptive consequence of uterine morphology and long-term expandability. Development 144:1775–1786.
Gerlo S, Davis JR, Mager DL, Kooijman R. 2006. Prolactin in man: a tale of two promoters. Bioessays 28:1051–1055.
Giudice LC. 2003. Elucidating endometrial function in the post-genomic era. Hum Reprod Update 9:223–235.
Hebenstreit D, et al. 2011. RNA Sequencing reveals two major classes of gene expression levels in metazoan cells. Mol Syst Biol 7:497.
Imbeault M, Hellebois P-Y, Trono D. 2017. KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. Nature 543:550–554.
Jordan IK, Rogozin IB, Glazko GV, Koonin EV. 2003. Origin of a substantial fraction of human regulatory sequences from transposable elements. Trends Genet 19:68–72.
Joswig A, Gabriel H-D, Klubsch M, Winterhager E. 2003. Apoptosis in uterine epithelium and decidua in response to implantation: evidence for different pathways. Reprod Biol Endocrinol RBE 1:44.
Kajihara T, Brosens JJ, Ishihara O. 2013. The role of FOXO1 in the decidual transformation of the endometrium and early pregnancy. Med Mol Morphol 46:61–68.
Kawamura K, et al. 2009. Brain-derived neurotrophic factor promotes implantation and subsequent placental development by stimulating trophoblast cell growth and survival. Endocrinology 150:3774–3782.
Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12:357–360.
Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 37:907–915.
Kim M, et al. 2013. VEGF-A regulated by progesterone governs uterine angiogenesis and vascular remodelling during pregnancy. EMBO Mol Med 5:1415–1430.
Kim MR, et al. 2005. Progesterone-dependent release of transforming growth factor-beta1 from epithelial cells enhances the endometrial decidualization by turning on the smad signalling in stromal cells. Mol Hum Reprod 11:801–808.
Kin K, et al. 2016. The transcriptomic evolution of mammalian pregnancy: gene expression innovations in endometrial stromal fibroblasts. Genome Biol Evol 8:2459–2473.
Kin K, Mizia J, Wagner GP. 2014. Immunohistological study of the endometrial stromal fibroblasts in the opossum, monodelphis domestica: evidence for homology with eutherian stromal Fibroblasts. Biol Reprod 90:1–12.
Kin K, Nnamani MC, Lynch VJ, Mihailetes E, Wagner GP. 2015. Cell-type phylogenetics and the origin of endometrial stromal cells. Cell Rep 10:1398–1409.
Klonsich T, et al. 2001. Epidermal growth factor-like ligands and erbB genes in the peri-implantation rabbit uterus and Blastocyst1. Biol Reprod 64:1835–1844.
Kunarro G, et al. 2010. Transposable elements have rewired the core regulatory network of human embryonic stem cells. Nat Genet 42:631–634.
Large MJ, et al. 2014. The epidermal growth factor receptor critically regulates endometrial function during early pregnancy. PLoS Genet 10:e1004451.
Lei W, et al. 2012. Progesterone and DNA damage encourage uterine cell proliferation and decidualization through up-regulating ribonucleotide reductase 2 expression during early pregnancy in mice. J Biol Chem 287:15174–15192.
Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. 2019. Webgestalt 2019: a web-based gene set analysis toolkit with revamped UIs and APIs. Nucleic Acids Res 47:W199–W205.
Logan PC, Ponnampalam AP, Rahnama F, Lobie PE, Mitchell MD. 2010. The effect of DNA methylation inhibitor 5-aza-2′-deoxycytidine on human endometrial stromal cells. Hum Reprod Oxf Engl 25:2859–2869.
Logan PC, Steiner M, Ponnampalam AP, Mitchell MD. 2012. Cell cycle regulation of human endometrial stromal cells during decidualization. Reprod Sci Thousand Oaks Calif 19:883–894.
Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550.
Lucas ES, et al. 2020. Recurrent pregnancy loss is associated with a prosenesical decidual response during the peri-implantation window. Commun Biol 3:37.
Lynch VJ, Brayer K, Gellersen B, Wagner GP. 2009. FoxoA1 cooperate to regulate decidual prolactin expression: towards inferring the core transcriptional regulators of decidual genes. PLoS One 4:e6845.
Lynch VJ, et al. 2015. Ancient transposable elements transformed the uterine regulatory landscape and transcriptome during the evolution of mammalian pregnancy. Cell Rep 10:551–561.
Lynch VJ, Leclerc RD, May G, Wagner GP. 2011. Transposon-mediated reprogramming of uterine regulatory networks contributed to the evolution of pregnancy in mammals. Nat Genet 43:1154–1159.
Mariani C, Mika K, Chigurupati S, Lynch VJ. 2021. Evolutionary transcriptomics implicates HAND2 in the origins of implantation and regulation of gestation length. eLife 10:e61257.
Mariani MC, Rana S, Lynch VJ. 2020. Derivation of endometrial gland organoids from term placenta. Placenta 101:75–79.
Matsui T, et al. 2010. Proviiral silencing in embryonic stem cells requires the histone methyltransferase ESET. Nature 464:927–931.
Mazur EC, et al. 2015. Progesterone receptor transcriptome and cis-acting regulatory sequences in decidualized human endometrial stromal cells. Endocrinology 156:2239–2253.
Messer A, Carter AM. 2006. Evolutionary transformations of fetal membrane characters in eutheria with special reference to afrotheria. J Exp Zool B Mol Dev Evol 308B:140–163.
Mestre Citrinovits AC, Strowitzki T, Germeyer A. 2019. Decreased autophagy impairs decidualization of human endometrial stromal cells: a role for ATG proteins in endometrial physiology. Int J Mol Sci 20:E3066.
Mika K, et al. 2021. Evolutionary transcriptomics implicates new genes and pathways in human pregnancy and adverse pregnancy outcomes. eLife 10:e69584.
Mika K, Whittington CM, McAllan BM, Lynch VJ. 2021. Gene expression phylogeny and ancestral transcriptome reconstruction resolves major transitions in the origins of pregnancy. eLife 11:e74297.
Ni N, Li Q. 2017. TGFβ superfamily signaling and uterine decidualization. Reprod Biol Endocrinol RBE 15:84.
Oksanen J, et al. 2019. Vegan: community ecology package. Available from: https://cran.r-project.org/web/packages/vegan/index.html
Pertea M, et al. 2015. Stringtie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol 33:290–295.
Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. 2016. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc 11:1650–1667.
Pijnenborg R, Vercruysse L, Brossens I. 2011. Deep placentalization. Best Pract Res Clin Obstet Gynaecol Placental Bed & Maternal–Fetal Disorders 25:273–285.
Pijnenborg R, Vercruysse L, Carter AM. 2011. Deep trophoblast invasion and spiral artery remodelling in the placental bed of the chimpanzee. Placenta 32:400–408.
Polak P, Domany E. 2006. Alu elements contain many binding sites for transcription factors and may play a role in regulation of developmental processes. BMC Genomics 7:133.
Rawlings TM, Makwana K, Tryfonsos M, Lucas ES. 2021. Organoids to model the endometrium: implantation and beyond. Reprod Fertil Dev 2:R85–R101.
Rowe HM, et al. 2010. KAP1 Controls endogenous retroviruses in embryonic stem cells. Nature 463:237–240.
Rubel CA, et al. 2011. Gata2 is a master regulator of endometrial function and progesterone signaling. Biol Reprod 85:179.
Rubel CA, et al. 2016. A Gata2-dependent transcription network regulates uterine progesterone responsiveness and endometrial function. Cell Rep 17:1414–1425.

Sakabe NJ, et al. 2020. Transcriptome and regulatory maps of decidua-derived stromal cells inform gene discovery in preterm birth. Sci Adv 6:eabc8696.

Salomonis N, et al. 2005. Identifying genetic networks underlying myometrial transition to labor. Genome Biol 6:R12.

Schwenke M, et al. 2013. Control of human endometrial stromal cell motility by PDGF-BB, HB-EGF and trophoblast-secreted factors. PloS One 8:e54336.

Soares MJ, Varberg KM, Iqbal K. 2018. Hemochorial placentation: development, function, and adaptations.1. Biol Reprod 99:196–211.

Stone JR, Wray GA. 2001. Rapid evolution of cis-regulatory sequences via local point mutations. Mol Biol Evol 18:1764–1770.

Strassmann BI. 1996. The evolution of endometrial cycles and menstruation. Q Rev Biol 71:181–220.

Takano M, et al. 2007. Transcriptional cross talk between the forkhead transcription factor forkhead box O1A and the progesterone receptor coordinates cell cycle regulation and differentiation in human endometrial stromal cells. Mol Endocrinol Baltim Md 21:2334–2349.

Turco MY, et al. 2017. Long-term, hormone-responsive organoid cultures of human endometrium in a chemically defined medium. Nat Cell Biol 19:568–577.

Wagner GP, Kin K, Lynch VJ. 2013. A model based criterion for gene expression calls using RNA-seq data. Theory Biosci 132:159–164.

Wang T, et al. 2007. Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53. Proc Natl Acad Sci 104:18613–18618.

Warrington NM, et al. 2019. Maternal and fetal genetic effects on birth weight and their relevance to cardio-metabolic risk factors. Nat Genet 51:804–814.

Wheeler TJ, et al. 2013. Dfam: a database of repetitive DNA based on profile hidden Markov models. Nucleic Acids Res 41:D70–D82.

Wray GA. 2007. The evolutionary significance of cis-regulatory mutations. Nat Rev Genet 8:206–216.

Wu L, Stadtmauer DJ, Maziarz J, Wagner G. 2020. Decidual cell differentiation is evolutionarily derived from fibroblast activation. bioRxiv preprint. Available from: https://www.biorxiv.org/content/10.1101/2020.12.18.423527v2

Yu H-F, et al. 2019. HB-EGF Ameliorates oxidative stress-mediated uterine decidualization damage. Oxid Med Cell Longev 2019:6170936.

Zhang Q, Yan J. 2016. Update of Wnt signaling in implantation and decidualization. Reprod Med Biol 15:95–105.

Zhang Y, et al. 2008. Model-based analysis of ChIP-Seq (MACS). Genome Biol 9:R137.

Associate editor: Dr. Ellen Pritham