Spring 2020 – Systems Biology of Reproduction
Discussion Outline – Fetal Development & Birth Systems
Michael K. Skinner – Biol 475/575
CUE 418, 10:35-11:50 am, Tuesday & Thursday
April 23, 2020
Week 15

Fetal Development & Birth Systems

Primary Papers:

1. Renthal, et al. (2013) Nature Reviews Endocrinology 9:391-401
2. Glotov, et al. (2015) BMC Systems Biology 9(Suppl 2):S4
3. Pique-Regi, et al. (2019) eLife 8:e52004

Discussion

Student 12: Reference 1 above

- What are the target organs for the miRNA in pregnancy?
- What gene clusters and networks are regulated?
- What progesterone targets and impacts influence contractility?

Student 13: Reference 2 above

- What major diseases are compared with preeclampsia and why?
- What networks were identified and impact?
- What risk factors were identified?

Student 14: Reference 3 above

- What technical approach was used and types of correlations?
- What transcriptome and cellular correlations were made?
- What major insights were provided for preterm birth?

MicroRNAs—mediators of myometrial contractility during pregnancy and labour

Nora E. Renthal, Koriand'r C. Williams and Carole R. Mendelson

Abstract | The maintenance of myometrial quiescence and initiation of contractility, which lead to parturition at term and preterm, involve a shifting equilibrium between anti-inflammatory and proinflammatory signalling pathways. Progesterone (P_4), acting through the progesterone receptor (PR), has an essential and multifaceted role in the maintenance of myometrial quiescence. This effect of P_4–PR signalling is mediated, in part, by its anti-inflammatory actions and capacity to repress the expression of genes that encode proinflammatory cytokines, such as IL-1 and IL-6, and contraction-associated proteins, such as *OXTR*, *GJA1* and *PTGS2*. By contrast, increased expression of genes that ultimately lead to parturition is mediated by enhanced inflammatory and estradiol-17β (E_2) and estrogen receptor α signalling, which reduce PR function, thus further intensifying the inflammatory response. To obtain a more complete understanding of the molecular events that underlie the transition of the pregnant myometrium from a refractory to a contractile state, the roles of microRNAs, their targets, and their transcriptional and hormonal regulation have been investigated. This article reviews the actions of the miR-200 family and their P_4-regulated targets—the transcription factors ZEB1, ZEB2 and STAT5B—in the pregnant myometrium, as well as the role of miR-199a-3p and miR-214 and their mutual target PTGS2. The central role of ZEB1 as the mediator of the opposing actions of P_4 and E_2 on myometrial contractility will be highlighted.

Renthal, N. E. *et al. Nat. Rev. Endocrinol.* **9**, 391–401 (2013); published online 14 May 2013; doi:10.1038/nrendo.2013.96

Introduction

Each year, 15 million babies are born prematurely throughout the world (~11% of all live births); >1 million of these babies die because of complications of their prematurity. Preterm birth, defined as birth at <37 weeks of gestation, is the leading cause of neonatal death and the second leading cause of death in children under the age of 5 years worldwide.[1] The rate of preterm birth ranges from 5% in several European countries to 18% in parts of sub-Saharan Africa. In the USA, the incidence of preterm birth has steadily increased over the past two decades and has to date levelled off at ~12% of all live births. Notably, the incidence of preterm birth in the USA is higher among certain racial and ethnic groups compared with the general population. For example, the prematurity rate in black neonates is ~18%, whereas ~11% of white neonates are born prematurely. The reasons for this racial disparity are not understood.[2] To elucidate the underlying pathogenesis and prevent this high incidence of preterm birth, the signalling pathways that maintain quiescence of the myometrium throughout pregnancy and mediate its conversion into a synchronously contractile unit must be understood.

The maintenance of myometrial quiescence and the timing of labour involve a delicate balance between hormonal, inflammatory and physical factors that regulate integrated signalling pathways between the mother and the fetus.[3,4] MicroRNAs (miRNAs) and their regulation serve a pivotal role in the molecular events that underlie the transition of the pregnant myometrium from a refractory to a contractile state during term and preterm labour. These evolutionarily conserved regulators of gene expression play important parts in a variety of biological and pathological processes, including cell differentiation,[5–7] cancer,[8,9] immune regulation[10] and female reproduction.[11–31]

Consequently, miRNAs might serve as hormonally modulated mediators of inflammation-associated and contraction-associated gene expression in the pregnant uterus from mice to humans. Here, we review the current knowledge on the hormonal regulation of myometrial quiescence and contractility, with a focus on the role of miRNAs and their hormonal regulation during pregnancy and labour.

Myometrial quiescence during pregnancy

Quiescence of the myometrium throughout most of pregnancy is maintained by increased circulating levels of progesterone (P_4), which acts via the nuclear progesterone receptor (PR).[3] The PR maintains uterine quiescence, in part, by tethering to and antagonizing the actions of the transcription factors NF-κB and AP-1 on the expression of genes that encode proinflammatory cytokines, such as IL-1, IL-6 and IL-8, and chemokines, such as CCL2.[32–35] P_4–PR signalling also blocks NF-κB activation by increasing expression of the NF-κB inhibitor IκBα.[32]

Competing interests
The authors declare no competing interests.

Department of Pediatrics, Children's Medical Center Dallas, 1935 Medical District Drive, Dallas, TX 75235, USA (**N. E. Renthal**). Departments of Biochemistry and Obstetrics and Gynecology, the University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA (**K. C. Williams**, **C. R. Mendelson**).

Correspondence to: C. R. Mendelson carole.mendelson@utsouthwestern.edu

Key points

- Progesterone (P_4), acting via the progesterone receptor (PR), maintains uterine quiescence in part by increasing the expression of *ZEB1* and *ZEB2*, which inhibit the contraction-associated genes *OXTR* and *GJA1*
- The initiation of myometrial contractility is mediated by an increased inflammatory response, an associated increase in 17β-estradiol (E_2) and estrogen receptor (ERα) signalling and a decline in PR function
- Near term, the myometrial inflammatory response is promoted by physical and hormonal signals from mother and fetus; in preterm labour, the increased inflammatory response is commonly induced by a bacterial infection
- Expression of the miR-200 family increases in mouse and human myometrium near term and suppresses ZEB1 and ZEB2 levels, which results in the de-repression of contractile genes and increased myometrial contractility
- Increased miR-200 expression near term also inhibits STAT5B; decreased STAT5B levels de-repress 20α-hydroxysteroid dehydrogenase and increase myometrial metabolism of P_4
- Increased E_2–ERα signalling and the decline in PR function near term mediate decreased expression of ZEB1, and of miR-199a-3p and miR-214, which contributes to the induction of PTGS2

Furthermore, the PR prevents myometrial contractility by increasing expression of the zinc finger E-box-binding homeobox 1 transcription factor ZEB1, which inhibits expression of genes that encode contraction-associated proteins, such as oxytocin receptor, gap junction α-1 protein (GJA1, also known as connexin 43), and prostaglandin G,H synthase 2 (PTGS2, also known as cyclooxygenase 2), the critical enzyme in the synthesis of contractile prostaglandins, such as prostaglandin $F_2\alpha$.[29,31] Thus, PR maintains myometrial quiescence mainly by inhibiting the expression of genes associated with inflammation and contraction (Figure 1).

Initiation of labour

The inflammatory response

The initiation of labour, both at term and preterm, is associated with an upregulated inflammatory response, which is characterized by increased levels of proinflammatory cytokines in the amniotic fluid[36] and infiltration of the myometrium, cervix, and fetal membranes by neutrophils and macrophages.[37–39] In preterm labour, the intra-amniotic bacterial infection that is associated with chorioamnionitis (inflammation of the fetal membranes) can lead to increased cytokine levels in amniotic fluid and induce inflammatory cell migration.[40] By contrast, near term, increased mechanical stretch[35,41] and hormonal factors produced by the developing fetus[38,42–45] provide the inflammatory stimuli. The invading immune cells secrete cytokines and chemokines[46] that promote the activation of NF-κB and other inflammation-associated transcription factors, such as AP-1, in the myometrium[38,47] and cervix.[48–50] These activated transcription factors, in turn, inhibit PR function, which further induces the inflammatory response and expression of myometrial contractile genes,[51–54] culminating in parturition (Figure 2).

Parturition—a decline in PR function

In rodents, maternal P_4 levels in the circulation decrease precipitously near term.[55] This finding led to the hypothesis that labour at term is associated with P_4 withdrawal. By contrast, in humans and in guinea pigs, maternal circulating P_4 levels, as well as myometrial PR levels, remain elevated throughout pregnancy and into labour.[44] Nevertheless, treatment with PR antagonists, such as mifepristone (also referred to as RU-486) or onapristone, can increase cervical ripening and induce spontaneous labour or increase the sensitivity to labour induction by oxytocin or prostaglandins.[56–59] Of note, even in mice, maternal P_4 levels at term remain well above the dissociation constant (K_d) for binding to PR.[60] Thus, the progression to labour at term, which is associated with an increased inflammatory response to mechanical and

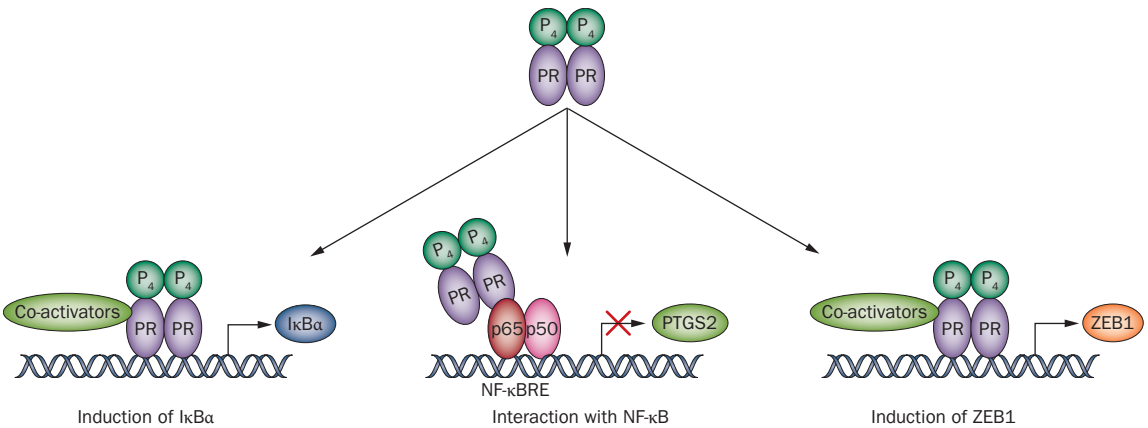

Figure 1 | P_4–PR regulation of myometrial quiescence. PR maintains myometrial quiescence, in part, by blocking activation of NF-κB and preventing its transcriptional activation of proinflammatory genes, such as *PTGS2*. The P_4–PR complex exerts this action, in part, by increasing expression of the NF-κB inhibitor IκBα, which prevents activation and nuclear translocation of the NF-κB subunits p50 and p65. PR also inhibits activation of contraction-associated genes such as *PTGS2* by interaction with NF-κB, that is, the P_4–PR complex tethers to NF-κB bound to response elements in the promoter of *PTGS2*. In addition, PR prevents myometrial contractility by increasing protein expression of ZEB1, which inhibits expression of contraction-associated genes, for example, *OXTR*, *GJA1* and *PTGS2*, by interaction with their promoter regions. Abbreviations: IκBα, NF-κB inhibitor α; NF-κB, nuclear factor κB; NF-κBRE, NF-κB response element; P_4, progesterone; PR, progesterone receptor; PTGS2, prostaglandin G,H synthase 2; ZEB1, zinc finger E-box-binding homeobox 1.

hormonal signals from the mother[61] and fetus,[38,42,62,63] might be mediated by a chain of molecular events that impair the ability of P_4 via PR to regulate target genes in the uterus that maintain myometrial quiescence.

The decline in myometrial PR function near term is caused by numerous factors, which include: a decrease in PR co-activators, for example, CREB-binding protein and the steroid receptor co-activators SRC-2 and SRC-3;[64] increased expression of inhibitory and truncated PR isoforms;[47,65] antagonism of the PR by NF-κB,[32,66,67] which is activated in the myometrium near term;[38,47] and enhanced local metabolism of P_4 to inactive products. The breakdown of P_4 into inactive metabolites is mediated by 20α-hydroxysteroid dehydrogenase (20α-HSD), an enzyme encoded by the *AKR1C1* gene in humans and the *Akr1c18* gene in mice, in the uterus[30] and by steroid 5α-reductase type 1, which is encoded by the *SRD5A1* gene, in the cervix (Figure 2).[68–70] Mice with a targeted deletion in *Srd5a1* exhibit defects in cervical ripening, which prevent delivery.[68,69] The inhibition of delivery occurs despite a precipitous decline in the levels of maternal circulating P_4 near term, which suggests that, in the mouse, local metabolism of P_4 in the cervix is also required for P_4 withdrawal and the initiation of parturition. Mice that lack 20α-HSD also manifest markedly delayed parturition,[71,72] despite a marked decline in circulating P_4 levels.[71]

Importantly, increased circulating estradiol-17β (E_2) levels[73,74] and enhanced estrogen receptor α (ERα) activity[65,75–77] near term also promote a cascade of pro-inflammatory events that contribute to the decline in PR function and initiate labour. Estrogens induce an influx of macrophages and neutrophils into the uterus and antagonize the anti-inflammatory actions of P_4 acting through PR.[65,78] Furthermore, ERα activation facilitates labour by enhancing transcription of the contraction-associated genes *OXTR*,[79] *GJA1*[80] and *PTGS2*,[31,65] which are expressed at low to undetectable levels in the uterus throughout most of pregnancy but are highly upregulated at term.[31,81,82]

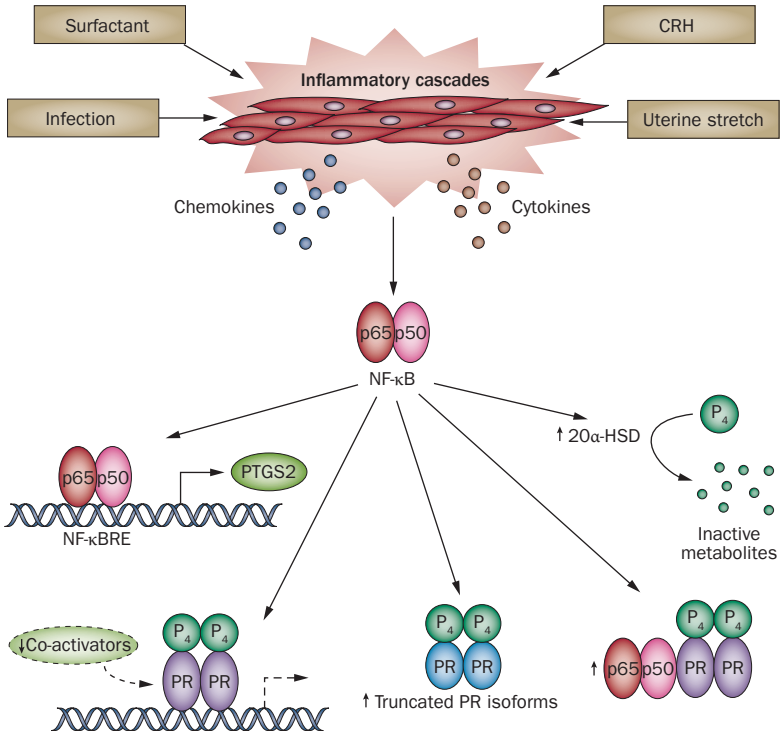

Figure 2 | The increased inflammatory response associated with parturition. Near term, a number of fetal and maternal signals, including secretion of surfactant by the fetal lung, production of hormonal factors by the placenta (such as corticotropin-releasing hormone), and mechanical stretch of the uterus, result in increased levels of proinflammatory cytokines in amniotic fluid and infiltration of the myometrium, cervix and fetal membranes by neutrophils and macrophages. The invading immune cells secrete cytokines and chemokines, promoting activation and binding of the NF-κB subunits p65 and p50 to the promoter of *PTGS2*. The increased inflammatory response can also inhibit PR function by suppressing expression of co-activators (such as CREB-binding protein and the steroid co-activators SRC-2, SRC-3), upregulating expression of truncated PR isoforms, direct interaction of NF-κB with PR, and induction of enzymes that metabolize P_4 to inactive products, such as 20α-HSD and steroid 5α-reductase type 1. Collectively, these processes result in a further induction of the inflammatory response and increased expression of myometrial contractile genes, culminating in parturition. Abbreviations: 20α-HSD, 20α-hydroxysteroid dehydrogenase; NF-κB, nuclear factor κB; NF-κBRE, NF-κB response element; P_4, progesterone; PR, progesterone receptor; PTGS2, prostaglandin G,H synthase 2.

MicroRNAs in female reproductive biology

Since their discovery almost two decades ago, miRNAs have increasingly been recognized as potent, evolutionarily conserved, post-transcriptional regulators of gene expression (Box 1). miRNAs exhibit important regulatory roles in vascular smooth muscle[83] and in female reproductive tissues, where they have been implicated in proliferation, differentiation, embryo implantation, hormone responsiveness, parturition and pathology (Table 1).[12–31,84–89] For example, conditional deletion of *Dicer1* in ovarian granulosa cells and in derivatives of the Müllerian duct (that is, oviduct, uterus and cervix) caused malformations of the uterus and sterility in mice.[14,22,24,27] Histologic examination revealed a decreased myometrial layer and fewer uterine glands. Although these findings are indicative of the importance of miRNAs for global uterine form and function, they shed little light on the roles of individual miRNAs in specific reproductive functions.

Coordinated miRNA and gene expression microarray analyses were used to identify miRNAs and targets that are differentially regulated in quiescent compared with contractile pregnant mouse myometrium.[29] Myometrial tissues from pregnant mice were removed at 15.5 days *post coitum* (dpc), when the myometrium is quiescent, and at 18.5 dpc, just before labour, which typically begins at 19 dpc. In total, 15 miRNAs were found to be significantly upregulated and downregulated ($P<0.05$), in association with a body of gestationally regulated genes.[29] Among the most markedly upregulated miRNAs was a conserved family, the miR-200 family, which was highly induced at term.[29] miRNAs that were significantly downregulated between 15.5 dpc and 18.5 dpc included the *miR-199a/214* cluster.[29,31]

The miR-200 family and its targets

The miR-200 family is comprised of five miRNAs arranged into two conserved clusters (*miR-200b/200a/429*

Box 1 | MicroRNAs—evolutionarily conserved, potent regulators of gene expression

The majority of mammalian microRNAs (miRNAs) exist throughout the genome as promoter-driven RNA polymerase II transcribed genes. Their biogenesis begins in the nucleus where they are transcribed as part of long RNA transcripts (pri-miRs). miRNAs commonly exist in polycistronic clusters, wherein multiple miRNA genes are encoded in a single primary transcript. These clustered miRNAs are often structurally similar to one another and have related biological functions. The pri-miRs are processed in the nucleus to ~70-nucleotide hairpin structures (pre-miRs) by Drosha/DGCR8, an RNase III endonuclease, and then exported to the cytoplasm where they are further processed by the RNase III Dicer to form a short miRNA duplex. Upon loading of one of the miRNA strands onto the RNA-induced silencing complex (RISC), containing Argonaute (Ago) proteins, the miRNA directs the RISC to its mRNA targets. Typically, only one of the two strands is selected as the mature, RISC-associated miRNA, whereas the other is either degraded, or expressed to a lesser extent. In some cases, both strands of the miRNA duplex become fully mature miRNAs and are incorporated into RISC complexes. In this case, the miRNAs are termed miR-XXX-3p and miR-XXX-5p. Approximately one-third of mammalian miRNAs are encoded within introns of protein-coding 'host' genes and are generated upon processing of the host RNA transcripts. The intron-encoded miRNAs are under control of the 'host' gene promoter and commonly mediate a related biological function to the host protein. This process has been reviewed in detail elsewhere.[126]

miRNA expression can be regulated at the transcriptional level[9] and at the level of their processing.[127] miRNAs typically function to repress gene expression by binding through imperfect base-pairing via their seed sequences (nucleotide 2–8 at the 5'-end) to complementary sites (seed match), which typically exist within the 3'-untranslated region of target mRNAs. Binding of the miRNA results in degradation of the mRNA target and/or inhibition of translation.[128,129] miRNAs can act as rheostats or as on–off switches of gene expression. The effect of a specific miRNA on the regulation of a target can be subtle, whereas the combined actions of several related miRNAs on the same target can have a pronounced phenotypic consequence. This phenomenon is especially true when miRNAs with similar or identical seed sequences are expressed as part of the same polycistronic transcript and/or when miRNAs encoded within the same transcript have different seed sequences that bind to the same target mRNA. It is estimated that ~1,000 miRNAs are encoded by the human genome and that these regulate approximately one-third of the expressed human genes.[5]

and *miR-141/200c*) that exist on two different chromosomes within the mouse and human genomes.[29] Each cluster is <2 kbp in length; the miRNAs in each cluster are coordinately transcribed.[90] miR-200b, miR-200c and miR-429 contain an identical seed sequence; miR-141 and miR-200a also share an identical seed sequence that differs by only a single nucleotide from that of miR-200b, miR-200c and miR-429 (Figure 3). Thus, all members of the miR-200 family probably share targets.

miR-200 targets: ZEB1 and ZEB2

From the results of a gene expression array and with the use of prediction algorithms[91] and published findings regarding miRNA–mRNA-target relationships, our group identified a pool of regulated miR-200 targets among the mRNAs downregulated at term in mice. The two most significantly downregulated miR-200 targets were the transcriptional repressors Zeb1 and Zeb2.[29] Importantly, the *Zeb1* gene was previously reported to be highly expressed in mouse myometrium, and its expression was upregulated by P_4 and PR signalling.[92,93] In myometrial tissues of pregnant mice, miR-200a,[30] miR-200b and miR-429[29] expression was significantly increased after 17.5 dpc; this timing corresponded with a reduction in Zeb1 and Zeb2 levels in the myometrium.[29] Importantly, the increased expression of miR-200b and miR-429 and decreased levels of ZEB1 and ZEB2 were also observed in the myometrium of women in labour compared with pregnant women not in labour.[29] This finding indicates that the relationship between miR-200 family members and ZEB1 and ZEB2 is conserved between mice and humans.

ZEB1 and ZEB2 were previously shown to be repressed by members of the miR-200 family in a double-negative feedback loop.[90,94–96] Furthermore, ZEB1 and ZEB2 were shown to stimulate epithelial to mesenchymal transition in cancer cells.[96–98] A clear inverse relationship exists between ZEB1 and ZEB2 and miR-200 family members in myometrium; overexpression of mimics of miR-200b and miR-429 in immortalized human myometrial cells reduced ZEB1 and ZEB2 levels, whereas transduction of mouse myometrial cells in primary culture with recombinant adenoviruses containing *Zeb1* and *Zeb2* expression vectors repressed expression of miR-200 and miR-429.[29] Moreover, using quantitative chromatin immunoprecipitation PCR (ChIP-qPCR), *in vivo* binding of Zeb1 to the *miR-200b/200a/429* promoter in mouse myometrium was observed at fairly high levels at 15.5 dpc; Zeb1 binding declined markedly at 18.5 dpc in association with the increase in miR-200b and miR-429 expression.[29]

miR-200b and miR-429 levels were increased and Zeb1 and Zeb2 mRNA and protein levels were decreased[29] in pregnant myometrium in two mouse models of premature parturition, induced either by a single subcutaneous injection of the antiprogestin and antiglucocorticosteroid mifepristone[99] or by intra-amniotic injection of the bacterial endotoxin, lipopolysaccharide. Conversely, daily injection of P_4 into pregnant mice from 15.5 dpc to 18.5 dpc, which inhibited myometrial contractile gene expression and delayed labour, specifically induced *Zeb1* expression. Surprisingly, P_4 injection had no effect on *Zeb2* expression.[29]

ZEB1 and *ZEB2* are expressed at relatively high levels in the myometrium throughout most of pregnancy, which raised the question regarding what factor(s) cause the pregnancy-associated induction of ZEB2. In studies using cultured mouse myometrial cells, *Zeb1* overexpression caused a time-dependent upregulation of Zeb2 levels.[29] This finding suggests that induction of *Zeb1* expression by P_4 and the associated inhibition of miR-200 family members, in turn, relieves suppression of *Zeb2*, allowing its subsequent induction. Notably, *Zeb1* has two putative P_4 response elements (PREs) in its promoter, whereas none are apparent in the *Zeb2* promoter. Accordingly, in co-transfection studies of human embryonic kidney (HEK293) cells with a *Zeb1*-luciferase reporter construct containing 978 bp of the *Zeb1* 5'-flanking sequence, which includes the two putative PREs, *Zeb1* promoter activity was induced by co-transfection of wild-type PR, but not by a mutant PR that contains a mutation in the DNA-binding domain.[29]

ZEB1 and ZEB2 suppress OXTR and GJA1

Two genes known to be upregulated near term in myometrium of a variety of species are *OXTR* and *GJA1*.[52,53,100–103] As *ZEB1* and *ZEB2* expression declined, *OXTR* and *GJA1* expression was temporally upregulated in mouse

and human myometrium.[29] GJA1 is a component of gap junctions in the myometrium, which mediate intercellular communication required for the synchronous contractions during labour. In mice, deletion of *Gja1* in smooth muscle leads to a substantial delay in the induction of labour.[104] Oxytocin (encoded by *OXT*) is widely accepted as a uterotonic agent; however, its role and that of the oxytocin receptor (encoded by *OXTR*) in parturition is uncertain; mice with a deletion in *Oxt* undergo normal parturition and give birth to live offspring,[105] and mice with a deletion of *Oxtr* show normal timing and duration of parturition.[106] These unexpected phenotypes might be due to a functional redundancy of the oxytocin signalling system and/or due to compensatory upregulation of other uterotonic systems, such as Ptgs2-mediated prostaglandin synthesis.

On the basis of the reciprocal temporal relationship between ZEB1, ZEB2 and contractile gene expression in pregnant mouse and human myometrium, we postulated that ZEBs suppress myometrial contractility by negatively regulating *OXTR* and *GJA1* expression. Indeed, overexpression of *ZEB1* or *ZEB2* in immortalized human myometrial cells caused a pronounced inhibition of OXTR and GJA1 mRNA levels.[29] Moreover, endogenous Zeb1 was bound at fairly high levels to E-box-containing regions of the mouse *Gja1* and *Oxtr* promoters at 15.5 dpc in pregnant mouse myometrium, whereas binding was markedly reduced at term.[29] To assess the functional roles of ZEB1 and ZEB2 on myometrial contractility, immortalized human myometrial cells transduced with *ZEB1* or *ZEB2* expression vectors or with control vectors were embedded in 3D collagen gels; the effects of oxytocin on contraction of the gels were analysed. Whereas oxytocin significantly induced contraction of collagen gel matrices embedded with cells transduced with control vectors, this action was blocked in cells transduced with *ZEB1* and *ZEB2* expression vectors. This finding indicates an inhibitory effect of these transcription factors on myometrial contractility *in vitro*.[29] Collectively, these findings support a role for the miR-200 family and ZEB1 and ZEB2 in the regulation of myometrial contractility during pregnancy and labour in mice and humans (Figure 4).

Throughout most of gestation, elevated circulating P_4 levels induce myometrial *ZEB1* expression via binding of P_4–PR to the *ZEB1* promoter. Elevated ZEB1 suppresses expression of *GJA1* and *OXTR*,[29] as well as the *miR-200* gene clusters[90,96] by binding to response elements within their promoters. The suppression of miR-200 promotes further upregulation of ZEB1 and increases expression of ZEB2. Together, ZEB1 and ZEB2 inhibit *GJA1* and *OXTR* to maintain myometrial quiescence. Near term, signals from mother and fetus, described above, together with a decline in circulating P_4[55] and/or PR function result in an increased inflammatory response within the myometrium and a further decline in PR function. Consequently ZEB1 mRNA and protein are significantly reduced. The decline in ZEB1 enables upregulation of miR-200 expression, resulting in a further suppression of ZEB1 and inhibition of ZEB2. The combined decline in ZEB1 and ZEB2 enables marked upregulation of *OXTR* and *GJA1* gene expression and the induction of myometrial contractility (Figure 4).

Table 1 | MicroRNAs in pregnancy and labour

Tissue	Process	Subject	References
Endometrium	Embryo receptivity and implantation	miR-101a, miR-199a-3p let-7b miR-320, let-7a miR-222 miR-21 miR-96, miR-375, miR-219-5p and PR miRNA expression and endometrial receptivity miRNA and mRNA analysis with delayed implantation	Chakrabarty *et al.*[12] Fu *et al.*[13] Xia *et al.*[19,20] Qian *et al.*[23] Hu *et al.*[28] Liu *et al.*[15] Li *et al.*[16] Su *et al.*[17]
Endometrium	Endometriosis	miR-21, Dicer miR-29c miR-222, miR-17-5p miR-9, miR-34 SNPs in miRNA target sequence Differential expression of miRNAs and pathway analysis	Aghajanova & Giudice[18] Hawkins *et al.*[85] Ramon *et al.*[86] Burney *et al.*[88] Zhao *et al.*[97] Ohlsson-Teague *et al.*[89]
Uterus	Development	Dicer knockout	Hawkins *et al.*[14] Nagaraja *et al.*[22] Gonzalez & Behringer[24] Hong *et al.*[27]
Myometrium	Pregnancy versus labour	miR-200 and ZEB miR-200 and STAT5B miR-199a-3p, miR-214	Renthal *et al.*[29] Williams *et al.*[30] Williams *et al.* [31]
Cervix	Pregnancy versus labour	miR-223, miR-34b, miR-34c	Hassan *et al.*[21]
Fetal membranes	Pregnancy versus term labour	miR-338, miR-223	Montenegro *et al.*[25,26] Kim *et al.*[84]

Abbreviation: miRNA, microRNA; PR, progesterone receptor; SNP, single nucleotide polymorphism.

miR-200a and STAT5B

As mentioned above, increased metabolism of P_4 within the uterus and cervix near term may contribute to a decline in PR function that is crucial for the initiation of parturition in all mammals.[30,68,69] In fact, increased P_4 metabolism in the uterus near term has been observed in a number of species.[107–110] In myometrium of pregnant women at term, a pronounced decrease in the ratio of P_4 to 20α-dihydroprogesterone was found;[110] 20α-dihydroprogesterone is an inactive metabolite of P_4 generated by 20α-HSD, a member of the aldo-ketoreductase (AKR) superfamily.[111] 20α-HSD is encoded by *AKR1C1* in humans and by *Akr1c18* in mice. Targeted deletion of *Akr1c18* in mice caused a pronounced delay in the initiation of labour.[72] Importantly, the transcription factor signal transducer and activator of transcription 5B (STAT5B) is a P_4-responsive transcriptional repressor of the gene that encodes 20α-HSD in reproductive tissues.[72,112] Consequently, *Stat5b* deficiency in mice resulted in increased expression of ovarian 20α-HSD, decreased circulating P_4 levels and caused abortion during mid-gestation.[72] Notably, the abortion rate in *Stat5b*-deficient mice was partially corrected by combined *Akr1c18* deficiency.[72]

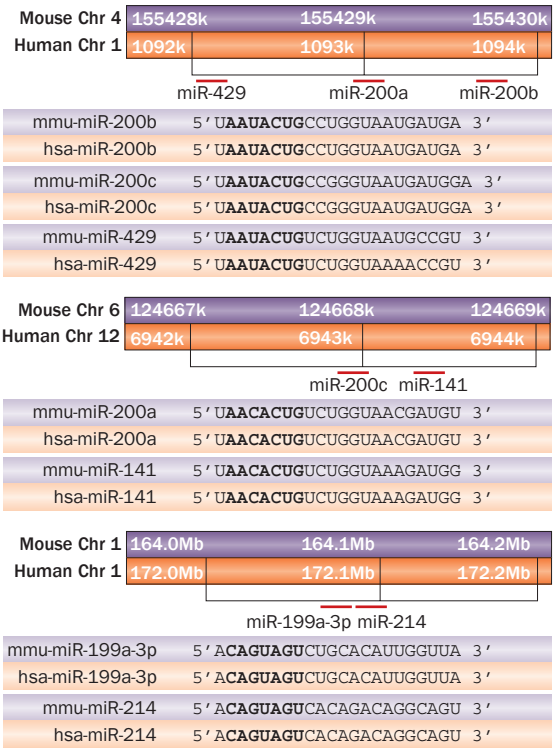

Figure 3 | Chromosomal location and seed sequences of members of the miR-200 family and the *miR-199a/214* cluster. The miR-200 family in mice and humans exists in two conserved clusters on mouse chromosome 4 and human chromosome 1 (*miR-200b/200a/429*) and on mouse chromosome 6 and human chromosome 12 (*miR-141/200c*), respectively. Family members are aligned according to their seed sequences (in bold), which are identical for miR-200b, miR-200c and miR-429 in mice and humans and differ by one nucleotide from miR-200a and miR-141, which are also identical and conserved in these species. Also shown are the locations of the *miR-199a/214* cluster in mouse and human chromosome 1 and the perfect conservation of the seed sequences of mouse and human miR-199a-3p and miR-214. Because these miRNAs are coordinately transcribed and contain distinct seed sequences, their interactions with different seed matches in the 3'-untranslated region of *PTGS2* can cooperatively exert a profound phenotypic effect. Abbreviations: Chr, chromosome; hsa, humans; miR, microRNA; mmu, mouse.

Expression of 20α-HSD in the ovarian corpus luteum remains low throughout pregnancy and increases in association with luteolysis at term;[113] hence, delayed labour in *Akr1c18*-deficient mice might have been caused by an inhibition of luteolysis and sustained circulating P_4 levels. However, in one of the *Akr1c18* knockout studies, labour was delayed in pregnant gene-targeted mice despite the fact that circulating P_4 levels declined in a similar manner as in wild-type mice.[71] This finding suggests that the decrease in ovarian P_4 production at term may not be sufficient and that actions of 20α-HSD to catalyse local metabolism of P_4 in the reproductive tract are also critical for the decline in PR function that leads to labour.

STAT5B is a target of miR-200a,[30] a member of the miR-200 family that increases dramatically at term in the myometrium of mice and humans. Remarkably, STAT5B mRNA and protein expression were reciprocally decreased in myometrial tissues of pregnant mice and humans, in association with the increase in miR-200a levels at term.[30] Moreover, these gestational changes were associated with decreased binding of endogenous Stat5b to response elements in the 5'-flanking region of the *Akr1c18* gene and an induction of 20α-HSD mRNA, protein and enzyme activity.[30] An intermediary role of miR-200a in P_4 induction of Stat5b was further suggested by the finding that P_4 treatment of ovariectomized mice inhibited miR-200a, increased Stat5b levels and inhibited 20α-HSD expression.[30] Conversely, injection of mifepristone increased miR-200a and 20α-HSD mRNA levels and inhibited *Stat5b* expression.[30] Given that the seed sequence of miR-200a is almost identical to that of miR-200b and miR-429, which are encoded within the same transcript,[29] these members of the miR-200 family might also target myometrial Stat5b and regulate 20α-HSD. However, this hypothesis remains to be confirmed. Moreover, miR-200a also targets ZEB1 and ZEB2 directly.[94] This finding suggests that increasing levels of miR-200a near term act cooperatively with miR-200b and miR-429 to inhibit ZEB expression and de-repress genes that encode proteins necessary for myometrial contractions.

These findings suggest that the elevated P_4 and PR activity in the myometrium during most of pregnancy causes induction of ZEB1 and inhibition of miR-200 expression, which is permissive for increased expression of *STAT5B*. Increased STAT5B levels maintain 20α-HSD at low concentrations and enable P_4 and PR function to remain elevated (Figure 4). The increased inflammatory response leading to term or preterm labour causes a decline in PR function and associated decrease in ZEB1 levels. This phenomenon permits upregulation of miR-200a and other miR-200 family members, an associated inhibition of *STAT5B* and induction of 20α-HSD expression and activity. The increased local metabolism of P_4 to inactive products in the myometrium contributes to the further decline in PR function and the progression to labour (Figure 4). These studies have revealed a robust positive feed-forward loop, wherein an initially modest decline in PR function and induction of miR-200 expression can escalate to an intensity that effectively reduces myometrial P_4 to levels below the K_d for binding to PR. This phenomenon culminates in a further decline of ZEB1 and ZEB2 levels, which permits the induction of contractile gene expression in response to increased NF-κB activation (Figure 4).

The *miR-199a/214* cluster and PTGS2

miRNAs that are significantly downregulated near term include miR-199a-3p, known to target the PTGS2 mRNA,[12] and miR-199a-5p, which inhibits NF-κB activation by targeting IKKβ.[114,115] Mature miR-199a-3p and miR-199a-5p are processed from the same precursor pri-miR-199a. Pri-miR-199a is synthesized as part of a 6-kb antisense transcript (Dnm3os) from the intron of the *Dnm3* gene.[116] The Dnm3os transcript, which is highly expressed in the uterus of pregnant mice,[116] also encodes miR-214, which targets PTEN[115] to activate the Akt pathway, as well as PTGS2.[31]

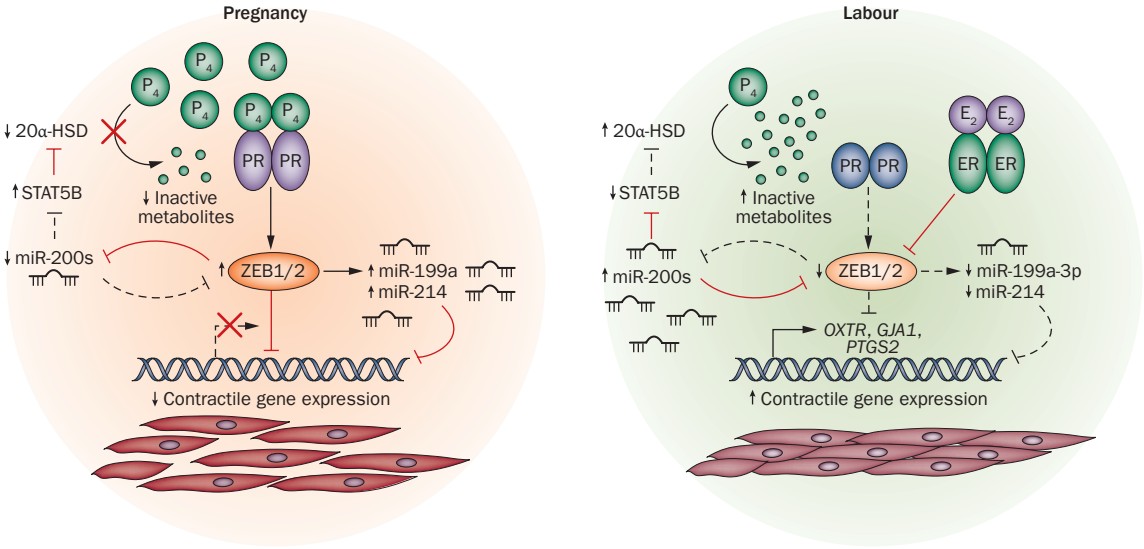

Figure 4 | Schematic representation of the pivotal role of ZEB1 and ZEB2 during pregnancy and labour. During pregnancy, increased levels of P_4 and increased PR function promote upregulation of ZEB1 levels in myometrium. ZEB1 inhibits expression of the miR-200 family and suppresses *OXTR* and *GJA1*. The decline in miR-200 levels further increases ZEB1 and ZEB2 levels, and these two proteins bind response elements upstream of the *miR-199a/214* cluster to enhance its expression, causing suppression of *PTGS2* expression and preventing synthesis of contractile prostaglandins. The decreased levels of miR-200s enable upregulation of another target, STAT5B, which inhibits expression of the gene encoding 20α-HSD, preventing the local metabolism of P_4 in myometrium. During the transition to labour, the decline in myometrial P_4 and PR function and increases in circulating E_2 and ERα activity cause downregulation of *ZEB1* expression. The decline in ZEB1 leads to induction of the miR-200 family, which further suppresses ZEB1 and ZEB2 levels, allowing upregulation of *OXTR* and *GJA1* expression. The decline in the levels of ZEB1 and ZEB2 also decreases expression of the *miR199a/214* cluster, which enables upregulation of *PTGS2* expression and increased synthesis of contractile prostaglandins. The increase in miR-200 expression inhibits STAT5B, permitting increased transcription of the gene that encodes 20α-HSD to promote increased metabolism of P_4 to inactive products in myometrium. Collectively, these molecular events contribute to the initiation of uterine contractility, leading to labour. Abbreviations: 20α-HSD, 20α-hydroxysteroid dehydrogenase; E_2, estradiol-17β; ER, estrogen receptor; GJA1, gap junction α-1 protein; miR, microRNA; OXTR, oxytocin receptor; P_4, progesterone; PR, progesterone receptor; PTGS2, prostaglandin G,H synthase 2; STAT5B, signal transducer and activator of transcription 5B; ZEB, zinc finger E-box-binding homeobox.

miR-214 was also observed to be expressed at high levels at 15.5 dpc and was downregulated in the uterus of pregnant mice at 18.5 dpc.[31] As discussed previously, NF-κB activation and expression of *PTGS2* increase in the pregnant myometrium during late gestation;[38,47,51] NF-κB signalling and PTGS-mediated prostaglandins are considered to be critical for the progression of labour. Whereas quantitative reverse transcription PCR of RNA isolated from myometrial tissues of pregnant mice at 15.5 dpc, 18.5 dpc and in active labour confirmed that miR-199a-3p and miR-214 were significantly downregulated in pregnant mouse myometrium at 18.5 dpc and during labour, a significant change in the expression of miR-199a-5p was not found.[31] For this reason, our studies have focused on the roles of miR-199a-3p and miR-214 in the regulation of PTGS2.

In pregnant mouse myometrium, Ptgs2 protein levels were significantly increased at 18.5 dpc and during labour, compared to 15.5 dpc, as miR-199a-3p and miR-214 levels reciprocally declined. By contrast, Ptgs2 mRNA levels remained low until 18.5 dpc and increased only during labour.[31] The relevance of the gestational increase in Ptgs2 protein levels is supported by the observation that prostaglandin $F_2\alpha$ levels were significantly increased in pregnant mouse uterus between 16 dpc and 18 dpc.[117] Our findings suggest that miR-199a-3p and miR-214 exert a direct effect on Ptgs2 mRNA translation rather than on mRNA stability.

This hypothesis was supported by the finding that overexpression of miR-199a-3p and miR-214 in cultured human myometrial cells decreased PTGS2 protein levels, but had no effect on PTGS2 mRNA levels.[31] Moreover, in myometrial samples from women in labour compared with those not in labour at term, in the absence of underlying infection, PTGS2 mRNA was unchanged, whereas PTGS2 protein levels were markedly increased during labour.[31] Thus, these collective findings suggest that PTGS2 expression in the pregnant myometrium is regulated at the level of mRNA translation and that the *miR-199a/214* cluster has an important role in this regulation. We view these findings to be relevant because the role of myometrial PTGS2 upregulation in the initiation of normal labour at term has been questioned.[118] This query was based on the observation that PTGS2 mRNA levels were not found to be increased in myometrial tissues from women in labour, compared with tissues from pregnant women not in labour, except in the presence of chorioamnionitis.[118] Notably, our studies revealed that myometrial levels of miR-199a-3p and miR-214 were significantly decreased in a mouse model of lipopolysaccharide-induced preterm labour, whereas Ptgs2

mRNA and protein levels were increased. Moreover, the physiological relevance of the relationship between the *miR-199a/214* cluster and PTGS2 in the regulation of myometrial contractility was further supported by the finding that miR-199a and miR-214 overexpression in cultured human myometrial cells blocked TNF-induced contractility to the same extent as the cyclooxygenase inhibitor indomethacin.[31]

As mentioned previously, P_4 and E_2 exert opposing effects on myometrial quiescence and contractility. Accordingly, E_2 treatment of ovariectomized mice suppressed, and P_4 treatment enhanced, uterine expression of miR-199a-3p and miR-214.[31] Interestingly, these opposing hormonal effects were found to be mediated by ZEB1, which is induced by P_4[29,92] and inhibited by E_2, and which activates transcription of miR199a and miR-214.[31] Thus, these findings have uncovered an intriguing pivotal role of ZEB1 as a negative regulator of the *miR-200b/200a/429* cluster and as a positive regulator of the *miR-199a/214* cluster that is under opposing control of P_4 and E_2 (Figure 4).

Future perspectives

The search for circulating biomarkers that are useful in predicting spontaneous preterm birth with the aim of clinical intervention has been ongoing for the past 40 years. In a review of the literature, of the 116 protein biomarkers analysed in 217 studies, not one provided predictive value for spontaneous preterm birth or yielded insight into the underlying pathophysiology.[119] The presence of specific miRNAs in blood and other body fluids suggests that they might serve as clinically useful biomarkers owing to a long half-life in body fluids, as well as the simplicity, accuracy and cost-effectiveness of their analysis.[120,121]

Moreover, chemically modified antisense oligonucleotides complementary to the seed sequence of miRNAs (anti-miRs or antagomiRs) have been used successfully *in vivo* to competitively inhibit miRNA function.[122–124] This finding suggests that miRNA manipulation has great potential as a future therapeutic strategy for the prevention of preterm birth. Given that members of an miRNA family share seed sequences, a single anti-miR can block the function of an entire family of miRNAs, making anti-miR therapy particularly appealing. Advances in the design of locked nucleic acid (LNA)-modified phosphorothioate oligonucleotides as highly stable miRNA antagonists that are taken up and retained by a variety of tissues for a period of up to 3 weeks[125] suggests a potential utility of these antagonists in the treatment or prevention of preterm labour and its potentially devastating consequences.

Conclusions

The studies described herein indicate that miRNAs have key collaborative roles in the hormonal control of myometrial quiescence and contractility during pregnancy and labour through regulation of contractile gene expression and PR function. They also suggest a central role for ZEB transcription factors and their unique capacity to act both as repressors and activators of gene transcription (Figure 4). Throughout most of pregnancy, when circulating P_4 levels are elevated and PR function is high, ZEB1 expression is induced to high levels. ZEB1 induction, in turn, causes suppression of the miR-200 family, which further increases ZEB1 levels and allows for de-repression of ZEB2. Together, ZEB1 and ZEB2 contribute to the maintenance of myometrial quiescence through their actions as potent transcriptional inhibitors of *OXTR* and *GJA1*. Elevated P_4–PR levels and suppression of the miR-200 family also enables upregulation of STAT5B levels, which inhibits expression of the P_4-metabolizing enzyme 20α-HSD. This inhibition permits myometrial tissue levels of P_4 to remain elevated, which increases PR function to sustain myometrial quiescence via the mechanisms just described, and suppresses NF-κB activation of inflammation-associated genes in the myometrium, such as *PTGS2*. Increased levels of ZEB1 and ZEB2 also cause upregulation of miR-199a-3p and miR-214, which directly target PTGS2 mRNA, resulting in maintained suppression of contractile prostaglandin synthesis.

Near term, the increased inflammatory response caused by uterine stretch, increased E_2 production and/or ERα activity and other signalling molecules produced by mother and fetus initiate a decline in PR function. This decrease further enhances NF-κB activation and the inflammatory response. The decreased P_4–PR function and increased E_2–ERα signalling promote a decline in *ZEB1* expression, a reciprocal increase in miR-200 expression, with further inhibition of ZEB1 and of ZEB2. This phenomenon enables the induction of *OXTR* and *GJA1* gene expression. The increased levels of miR-200s suppress STAT5B, enabling the induction of 20α-HSD, which promotes increased metabolism of P_4 to inactive products within the myometrium, contributing to the decline in PR function. The decreased levels of ZEBs also result in reduced expression of miR-199a-3p and miR-214 and thus increased PTGS2 protein expression, culminating in the synthesis of contractile prostaglandins. Together, these highly orchestrated molecular events enhance myometrial contractility and delivery of the fetus. We propose that similar molecular mechanisms mediate the induction of preterm labour, although the initiating signals (for example, bacterial infection with chorioamnionitis) are different. Taken together, this research has revealed that key miRNAs function as evolutionarily conserved, hormonally controlled modulators of inflammatory and contractile gene expression in the pregnant uterus, with a crucial role in the maintenance of pregnancy and initiation of term and preterm labour.

Review criteria

A PubMed search was performed for articles published between 1971 and 2013, using the keywords "microRNA", "miRNA", "miR", "pregnancy", "parturition", "labour", "progesterone", "progesterone receptor", "estrogen", "estrogen receptor", "inflammation", "inflammatory". Selected original research papers and review articles are discussed in this Review. All articles identified were English-language, full-text papers. We also searched the reference lists of identified articles for additional papers.

1. Blencowe, H. *et al.* National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet* **379**, 2162–2172 (2012).
2. Peltier, M. R. *et al.* Amniotic fluid and maternal race influence responsiveness of fetal membranes to bacteria. *J. Reprod. Immunol.* **96**, 68–78 (2012).
3. Mendelson, C. R. Minireview: fetal-maternal hormonal signaling in pregnancy and labor. *Mol. Endocrinol.* **23**, 947–954 (2009).
4. Shynlova, O., Lee, Y. H., Srikhajon, K. & Lye, S. J. Physiologic uterine inflammation and labor onset: integration of endocrine and mechanical signals. *Reprod. Sci.* **20**, 154–167 (2013).
5. van Rooij, E., Liu, N. & Olson, E. N. MicroRNAs flex their muscles. *Trends Genet.* **24**, 159–166 (2008).
6. Turner, M. L., Schnorfeil, F. M. & Brocker, T. MicroRNAs regulate dendritic cell differentiation and function. *J. Immunol.* **187**, 3911–3917 (2011).
7. Braun, T. & Gautel, M. Transcriptional mechanisms regulating skeletal muscle differentiation, growth and homeostasis. *Nat. Rev. Mol. Cell Biol.* **12**, 349–361 (2011).
8. Garzon, R., Calin, G. A. & Croce, C. M. MicroRNAs in cancer. *Annu. Rev. Med.* **60**, 167–179 (2009).
9. Mendell, J. T. miRiad roles for the miR-17-92 cluster in development and disease. *Cell* **133**, 217–222 (2008).
10. Boldin, M. P. & Baltimore, D. MicroRNAs, new effectors and regulators of NF-κB. *Immunol. Rev.* **246**, 205–220 (2012).
11. Hawkins, S. M., Buchold, G. M. & Matzuk, M. M. Minireview: The roles of small RNA pathways in reproductive medicine. *Mol. Endocrinol.* **25**, 1257–1279 (2011).
12. Chakrabarty, A. *et al.* MicroRNA regulation of cyclooxygenase-2 during embryo implantation. *Proc. Natl Acad. Sci. USA* **104**, 15144–15149 (2007).
13. Fu, T. Y., Lin, C. T. & Tang, P. C. Steroid hormone-regulated let-7b mediates cell proliferation and basigin expression in the mouse endometrium. *J. Reprod. Dev.* **57**, 627–635 (2011).
14. Hawkins, S. M. *et al.* Dysregulation of uterine signaling pathways in progesterone receptor-Cre knockout of dicer. *Mol. Endocrinol.* **26**, 1552–1566 (2012).
15. Liu, J. L. *et al.* Combined analysis of microRNome and 3'-UTRome reveals a species-specific regulation of progesterone receptor expression in the endometrium of rhesus monkey. *J. Biol. Chem.* **287**, 13899–13910 (2012).
16. Li, R. *et al.* MicroRNA array and microarray evaluation of endometrial receptivity in patients with high serum progesterone levels on the day of hCG administration. *Reprod. Biol. Endocrinol.* **9**, 29 (2011).
17. Su, R. W. *et al.* The integrative analysis of microRNA and mRNA expression in mouse uterus under delayed implantation and activation. *PLoS ONE* **5**, e15513 (2010).
18. Aghajanova, L. & Giudice, L. C. Molecular evidence for differences in endometrium in severe versus mild endometriosis. *Reprod. Sci.* **18**, 229–251 (2011).
19. Xia, H. F. *et al.* Temporal and spatial regulation of miR-320 in the uterus during embryo implantation in the rat. *Int. J. Mol. Sci.* **11**, 719–730 (2010).
20. Xia, H. F. *et al.* Temporal and spatial regulation of let-7a in the uterus during embryo implantation in the rat. *J. Reprod. Dev.* **56**, 73–78 (2010).
21. Hassan, S. S. *et al.* MicroRNA expression profiling of the human uterine cervix after term labor and delivery. *Am. J. Obstet. Gynecol.* **202**, 80 e1–e8 (2010).
22. Nagaraja, A. K. *et al.* Deletion of Dicer in somatic cells of the female reproductive tract causes sterility. *Mol. Endocrinol.* **22**, 2336–2352 (2008).
23. Qian, K. *et al.* Hsa-miR-222 is involved in differentiation of endometrial stromal cells in vitro. *Endocrinology* **150**, 4734–4743 (2009).
24. Gonzalez, G. & Behringer, R. R. Dicer is required for female reproductive tract development and fertility in the mouse. *Mol. Reprod. Dev.* **76**, 678–688 (2009).
25. Montenegro, D. *et al.* Differential expression of microRNAs with progression of gestation and inflammation in the human chorioamniotic membranes. *Am. J. Obstet. Gynecol.* **197**, 289 e1–e6 (2007).
26. Montenegro, D. *et al.* Expression patterns of microRNAs in the chorioamniotic membranes: a role for microRNAs in human pregnancy and parturition. *J. Pathol.* **217**, 113–121 (2009).
27. Hong, X., Luense, L. J., McGinnis, L. K., Nothnick, W. B. & Christenson, L. K. Dicer1 is essential for female fertility and normal development of the female reproductive system. *Endocrinology* **149**, 6207–6212 (2008).
28. Hu, S. *J. et al.* MicroRNA expression and regulation in mouse uterus during embryo implantation. *J. Biol. Chem.* **283**, 23473–23484 (2008).
29. Renthal, N. E. *et al.* miR-200 family and targets, ZEB1 and ZEB2, modulate uterine quiescence and contractility during pregnancy and labor. *Proc. Natl Acad. Sci. USA* **107**, 20828–20833 (2010).
30. Williams, K. C., Renthal, N. E., Condon, J. C., Gerard, R. D. & Mendelson, C. R. MicroRNA-200a serves a key role in the decline of progesterone receptor function leading to term and preterm labor. *Proc. Natl Acad. Sci. USA* **109**, 7529–7534 (2012).
31. Williams, K. C., Renthal, N. E., Gerard, R. D. & Mendelson, C. R. The microRNA (miR)-199a/214 cluster mediates opposing effects of progesterone and estrogen on uterine contractility during pregnancy and labor. *Mol. Endocrinol.* **26**, 1857–1867 (2012).
32. Hardy, D. B., Janowski, B. A., Corey, D. R. & Mendelson, C. R. Progesterone receptor (PR) plays a major anti-inflammatory role in human myometrial cells by antagonism of NF-κB activation of cyclooxygenase 2 expression. *Mol. Endocrinol.* **20**, 2724–2733 (2006).
33. Loudon, J. A., Elliott, C. L., Hills, F. & Bennett, P. R. Progesterone represses interleukin-8 and cyclo-oxygenase-2 in human lower segment fibroblast cells and amnion epithelial cells. *Biol. Reprod.* **69**, 331–337 (2003).
34. Dong, X. *et al.* p54nrb is a transcriptional corepressor of the progesterone receptor that modulates transcription of the labor-associated gene, connexin 43 (*Gja1*). *Mol. Endocrinol.* **23**, 1147–1160 (2009).
35. Shynlova, O., Tsui, P., Dorogin, A. & Lye, S. J. Monocyte chemoattractant protein-1 (CCL-2) integrates mechanical and endocrine signals that mediate term and preterm labor. *J. Immunol.* **181**, 1470–1479 (2008).
36. Cox, S. M., Casey, M. L. & MacDonald, P. C. Accumulation of interleukin-1β and interleukin-6 in amniotic fluid: a sequela of labour at term and preterm. *Hum. Reprod. Update* **3**, 517–527 (1997).
37. Thomson, A. *J. et al.* Leukocytes infiltrate the myometrium during human parturition: further evidence that labour is an inflammatory process. *Hum. Reprod.* **14**, 229–236 (1999).
38. Condon, J. C., Jeyasuria, P., Faust, J. M. & Mendelson, C. R. Surfactant protein secreted by the maturing mouse fetal lung acts as a hormone that signals the initiation of parturition. *Proc. Natl Acad. Sci. USA* **101**, 4978–4983 (2004).
39. Osman, I. *et al.* Leukocyte density and pro-inflammatory cytokine expression in human fetal membranes, decidua, cervix and myometrium before and during labour at term. *Mol. Hum. Reprod.* **9**, 41–45 (2003).
40. Rauk, P. N. & Chiao, J. P. Interleukin-1 stimulates human uterine prostaglandin production through induction of cyclooxygenase-2 expression. *Am. J. Reprod. Immunol.* **43**, 152–159 (2000).
41. Sooranna, S. R. *et al.* Mechanical stretch activates type 2 cyclooxygenase via activator protein-1 transcription factor in human myometrial cells. *Mol. Hum. Reprod.* **10**, 109–113 (2004).
42. Montalbano, A. P., Hawgood, S. & Mendelson, C. R. Mice deficient in surfactant protein A (SP-A) and SP-D or in TLR2 manifest delayed parturition and decreased expression of inflammatory and contractile genes. *Endocrinology* **154**, 483–498 (2013).
43. Shaw, G. & Renfree, M. B. Fetal control of parturition in marsupials. *Reprod. Fertil. Dev.* **13**, 653–659 (2001).
44. Challis, J. R. G., Matthews, S. G., Gibb, W. & Lye, S. J. Endocrine and paracrine regulation of birth at term and preterm. *Endocr. Rev.* **21**, 514–550 (2000).
45. Mitchell, M. D., MacDonald, P. C. & Casey, M. L. Stimulation of prostaglandin E_2 synthesis in human amnion cells maintained in monolayer culture by a substance(s) in amniotic fluid. *Prostaglandins Leukot. Med.* **15**, 399–407 (1984).
46. Romero, R. *et al.* The role of inflammation and infection in preterm birth. *Semin. Reprod. Med.* **25**, 21–39 (2007).
47. Condon, J. C., Hardy, D. B., Kovaric, K. & Mendelson, C. R. Up-regulation of the progesterone receptor (PR)-C isoform in laboring myometrium by activation of nuclear factor-κB may contribute to the onset of labor through inhibition of PR function. *Mol. Endocrinol.* **20**, 764–775 (2006).
48. Allport, V. C. *et al.* Human labour is associated with nuclear factor-κB activity which mediates cyclo-oxygenase-2 expression and is involved with the 'functional progesterone withdrawal'. *Mol. Hum. Reprod.* **7**, 581–586 (2001).
49. Lee, Y. S., Terzidou, V., Lindstrom, T., Johnson, M. & Bennett, P. R. The role of CCAAT/enhancer-binding protein β in the transcriptional regulation of COX-2 in human amnion. *Mol. Hum. Reprod.* **11**, 853–858 (2005).
50. Elliott, C. L., Allport, V. C., Loudon, J. A., Wu, G. D. & Bennett, P. R. Nuclear factor-κB is essential for up-regulation of interleukin-8 expression in human amnion and cervical epithelial cells. *Mol. Hum. Reprod.* **7**, 787–790 (2001).
51. Olson, D. M. The role of prostaglandins in the initiation of parturition. *Best Pract. Res. Clin. Obstet. Gynaecol.* **17**, 717–730 (2003).
52. Chow, L. & Lye, S. J. Expression of the gap junction protein connexin-43 is increased in the human myometrium toward term and with the onset of labor. *Am. J. Obstet. Gynecol.* **170**, 788–795 (1994).
53. Fuchs, A. R., Fuchs, F., Husslein, P. & Soloff, M. S. Oxytocin receptors in the human uterus during pregnancy and parturition. *Am. J. Obstet. Gynecol.* **150**, 734–741 (1984).
54. Soloff, M. S., Cook, D. L. Jr, Jeng, Y. J. & Anderson, G. D. In situ analysis of

interleukin-1-induced transcription of cox-2 and il-8 in cultured human myometrial cells. *Endocrinology* **145**, 1248–1254 (2004).
55. Virgo, B. B. & Bellward, G. D. Serum progesterone levels in the pregnant and postpartum laboratory mouse. *Endocrinology* **95**, 1486–1490 (1974).
56. Frydman, R. *et al.* Labor induction in women at term with mifepristone (RU 486): a double-blind, randomized, placebo-controlled study. *Obstet. Gynecol.* **80**, 972–975 (1992).
57. Elliott, C. L., Brennand, J. E. & Calder, A. A. The effects of mifepristone on cervical ripening and labor induction in primigravidae. *Obstet. Gynecol.* **92**, 804–809 (1998).
58. Stenlund, P. M., Ekman, G., Aedo, A. R. & Bygdeman, M. Induction of labor with mifepristone—a randomized, double-blind study versus placebo. *Acta Obstet. Gynecol. Scand.* **78**, 793–798 (1999).
59. Chwalisz, K. The use of progesterone antagonists for cervical ripening and as an adjunct to labour and delivery. *Hum. Reprod.* **9** (Suppl. 1), 131–161 (1994).
60. Pointis, G., Rao, B., Latreille, M. T., Mignot, T. M. & Cedard, L. Progesterone levels in the circulating blood of the ovarian and uterine veins during gestation in the mouse. *Biol. Reprod.* **24**, 801–805 (1981).
61. Shynlova, O., Tsui, P., Jaffer, S. & Lye, S. J. Integration of endocrine and mechanical signals in the regulation of myometrial functions during pregnancy and labour. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **144** (Suppl. 1), 2–10 (2009).
62. Toyoshima, K., Narahara, H., Furukawa, M., Frenkel, R. A. & Johnston, J. M. Platelet-activating factor. Role in fetal lung development and relationship to normal and premature labor. *Clin. Perinatol.* **22**, 263–280 (1995).
63. López Bernal, A., Newman, G. E., Phizackerley, P. J. & Turnbull, A. C. Surfactant stimulates prostaglandin E production in human amnion. *Br. J. Obstet. Gynaecol.* **95**, 1013–1017 (1988).
64. Condon, J. C., Jeyasuria, P., Faust, J. M., Wilson, J. M. & Mendelson, C. R. A decline in progesterone receptor coactivators in the pregnant uterus at term may antagonize progesterone receptor function and contribute to the initiation of parturition. *Proc. Natl Acad. Sci. USA* **100**, 9518–9523 (2003).
65. Mesiano, S. *et al.* Progesterone withdrawal and estrogen activation in human parturition are coordinated by progesterone receptor A expression in the myometrium. *J. Clin. Endocrinol. Metab.* **87**, 2924–2930 (2002).
66. Kalkhoven, E., Wissink, S., van der Saag, P. T. & van der Burg, B. Negative interaction between the RelA(p65) subunit of NF-κB and the progesterone receptor. *J. Biol. Chem.* **271**, 6217–6224 (1996).
67. Hardy, D. B., Janowski, B. A., Chen, C.-C. & Mendelson, C. R. Progesterone receptor inhibits aromatase and inflammatory response pathways in breast cancer cells via ligand-dependent and ligand-independent mechanisms. *Mol. Endocrinol.* **22**, 1812–1824 (2008).
68. Mahendroo, M. S., Cala, K. M. & Russell, D. W. 5α-reduced androgens play a key role in murine parturition. *Mol. Endocrinol.* **10**, 380–392 (1996).
69. Mahendroo, M. S., Porter, A., Russell, D. W. & Word, R. A. The parturition defect in steroid 5α-reductase type 1 knockout mice is due to impaired cervical ripening. *Mol. Endocrinol.* **13**, 981–992 (1999).
70. Andersson, S., Minjarez, D., Yost, N. P. & Word, R. A. Estrogen and progesterone metabolism in the cervix during pregnancy and parturition. *J. Clin. Endocrinol. Metab.* **93**, 2366–2374 (2008).
71. Ishida, M. *et al.* Reproductive phenotypes in mice with targeted disruption of the 20α-hydroxysteroid dehydrogenase gene. *J. Reprod. Dev.* **53**, 499–508 (2007).
72. Piekorz, R. P., Gingras, S., Hoffmeyer, A., Ihle,J. N. & Weinstein, Y. Regulation of progesterone levels during pregnancy and parturition by signal transducer and activator of transcription 5 and 20α-hydroxysteroid dehydrogenase. *Mol. Endocrinol.* **19**, 431–440 (2005).
73. Challis, J. R. Sharp increase in free circulating oestrogens immediately before parturition in sheep. *Nature* **229**, 208 (1971).
74. Buster, J. E. *et al.* Interrelationships of circulating maternal steroid concentrations in third trimester pregnancies. II. C18 and C19 steroids: estradiol, estriol, dehydroepiandrosterone, dehydroepiandrosterone sulfate, δ5-androstenediol, δ-androstenedione, testosterone, and dihydrotestosterone. *J. Clin. Endocrinol. Metab.* **48**, 139–142 (1979).
75. Wu, W. X., Myers, D. A. & Nathanielsz, P. W. Changes in estrogen receptor messenger ribonucleic acid in sheep fetal and maternal tissues during late gestation and labor. *Am. J. Obstet. Gynecol.* **172**, 844–850 (1995).
76. Mesiano, S. & Welsh, T. N. Steroid hormone control of myometrial contractility and parturition. *Semin. Cell Dev. Biol.* **18**, 321–331 (2007).
77. Welsh, T. *et al.* Estrogen receptor (ER) expression and function in the pregnant human myometrium: estradiol via ERα activates ERK1/2 signaling in term myometrium. *J. Endocrinol.* **212**, 227–238 (2012).
78. Tibbetts, T. A., Conneely, O. M. & O'Malley, B. W. Progesterone via its receptor antagonizes the pro-inflammatory activity of estrogen in the mouse uterus. *Biol. Reprod.* **60**, 1158–1165 (1999).
79. Murata, T., Narita, K., Honda, K., Matsukawa, S. & Higuchi, T. Differential regulation of estrogen receptor α and β mRNAs in the rat uterus during pregnancy and labor: possible involvement of estrogen receptors in oxytocin receptor regulation. *Endocr. J.* **50**, 579–587 (2003).
80. Piersanti, M. & Lye, S. J. Increase in messenger ribonucleic acid encoding the myometrial gap junction protein, connexin-43, requires protein synthesis and is associated with increased expression of the activator protein-1, c-fos. *Endocrinology* **136**, 3571–3578 (1995).
81. Tsuboi, K. *et al.* Uterine expression of prostaglandin H2 synthase in late pregnancy and during parturition in prostaglandin F receptor-deficient mice. *Endocrinology* **141**, 315–324 (2000).
82. Engstrøm, T. The regulation by ovarian steroids of prostaglandin synthesis and prostaglandin-induced contractility in non-pregnant rat myometrium. Modulating effects of isoproterenol. *J. Endocrinol.* **169**, 33–41 (2001).
83. Albinsson, S. *et al.* MicroRNAs are necessary for vascular smooth muscle growth, differentiation, and function. *Arterioscler. Thromb. Vasc. Biol.* **30**, 1118–1126 (2010).
84. Kim, S. Y. *et al.* miR-143 regulation of prostaglandin-endoperoxidase synthase 2 in the amnion: implications for human parturition at term. *PLoS ONE* **6**, e24131 (2011).
85. Hawkins, S. M. *et al.* Functional microRNA involved in endometriosis. *Mol. Endocrinol.* **25**, 821–832 (2011).
86. Ramon, L. A. *et al.* microRNAs expression in endometriosis and their relation to angiogenic factors. *Hum. Reprod.* **26**, 1082–1090 (2011).
87. Zhao, Z. Z. *et al.* Evaluation of polymorphisms in predicted target sites for micro RNAs differentially expressed in endometriosis. *Mol. Hum. Reprod.* **17**, 92–103 (2011).
88. Burney, R. O. *et al.* MicroRNA expression profiling of eutopic secretory endometrium in women with versus without endometriosis. *Mol. Hum. Reprod.* **15**, 625–631 (2009).
89. Ohlsson Teague, E. M. *et al.* MicroRNA-regulated pathways associated with endometriosis. *Mol. Endocrinol.* **23**, 265–275 (2009).
90. Bracken, C. P. *et al.* A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition. *Cancer Res.* **68**, 7846–7854 (2008).
91. Grimson, A. *et al.* MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol. Cell* **27**, 91–105 (2007).
92. Spoelstra, N. S. *et al.* The transcription factor ZEB1 is aberrantly expressed in aggressive uterine cancers. *Cancer Res.* **66**, 3893–3902 (2006).
93. Cochrane, D. R. *et al.* The role of miRNAs in progesterone action. *Mol. Cell. Endocrinol.* **357**, 50–59 (2012).
94. Gregory, P. A. *et al.* The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat. Cell Biol.* **10**, 593–601 (2008).
95. Postigo, A. A. & Dean, D. C. Differential expression and function of members of the zfh-1 family of zinc finger/homeodomain repressors. *Proc. Natl Acad. Sci. USA* **97**, 6391–6396 (2000).
96. Burk, U. *et al.* A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO Rep.* **9**, 582–589 (2008).
97. Vandewalle, C., Van Roy, F. & Berx, G. The role of the ZEB family of transcription factors in development and disease. *Cell. Mol. Life Sci.* **66**, 773–787 (2009).
98. Brabletz, S. *et al.* The ZEB1/miR-200 feedback loop controls Notch signalling in cancer cells. *EMBO J.* **30**, 770–782 (2011).
99. Dudley, D. J., Branch, D. W., Edwin, S. S. & Mitchell, M. D. Induction of preterm birth in mice by RU486. *Biol. Reprod.* **55**, 992–995 (1996).
100. Wu, W. X., Ma, X. H., Zhang, Q. & Nathanielsz, P. W. Characterization of topology-, gestation- and labor-related changes of a cassette of myometrial contraction-associated protein mRNA in the pregnant baboon myometrium. *J. Endocrinol.* **171**, 445–453 (2001).
101. Ou, C. W., Chen, Z. Q., Qi, S. & Lye, S. J. Increased expression of the rat myometrial oxytocin receptor messenger ribonucleic acid during labor requires both mechanical and hormonal signals. *Biol. Reprod.* **59**, 1055–1061 (1998).
102. Ou, C. W., Orsino, A. & Lye, S. J. Expression of connexin-43 and connexin-26 in the rat myometrium during pregnancy and labor is differentially regulated by mechanical and hormonal signals. *Endocrinology* **138**, 5398–5407 (1997).
103. Sparey, C., Robson, S. C., Bailey, J., Lyall, F. & Europe-Finner, G. N. The differential expression of myometrial connexin-43, cyclooxygenase-1 and -2, and Gs α proteins in the upper and lower segments of the human uterus during pregnancy and labor. *J. Clin. Endocrinol. Metab.* **84**, 1705–1710 (1999).
104. Döring, B. *et al.* Ablation of connexin43 in uterine smooth muscle cells of the mouse causes delayed parturition. *J. Cell. Sci.* **119**, 1715–1722 (2006).

105. Nishimori, K. *et al.* Oxytocin is required for nursing but is not essential for parturition or reproductive behavior. *Proc. Natl Acad. Sci. USA* **93**, 11699–11704 (1996).
106. Takayanagi, Y. *et al.* Pervasive social deficits, but normal parturition, in oxytocin receptor-deficient mice. *Proc. Natl Acad. Sci. USA* **102**, 16096–16101 (2005).
107. Puri, C. P. & Garfield, R. E. Changes in hormone levels and gap junctions in the rat uterus during pregnancy and parturition. *Biol. Reprod.* **27**, 967–975 (1982).
108. Power, S. G. & Challis, J. R. The effects of gestational age and intrafetal ACTH administration on the concentration of progesterone in the fetal membranes, endometrium, and myometrium of pregnant sheep. *Can. J. Physiol. Pharmacol.* **65**, 136–140 (1987).
109. Csapo, A. I., Eskola, J. & Tarro, S. Gestational changes in the progesterone and prostaglandin F levels of the guinea-pig. *Prostaglandins* **21**, 53–64 (1981).
110. Runnebaum, B. & Zander, J. Progesterone and 20α-dihydroprogesterone in human myometrium during pregnancy. *Acta Endocrinol. Suppl. (Copenh.)* **150**, 3–45 (1971).
111. Penning, T. M. & Drury, J. E. Human aldo-keto reductases: Function, gene regulation, and single nucleotide polymorphisms. *Arch. Biochem. Biophys.* **464**, 241–250 (2007).
112. Richer, J. K. *et al.* Convergence of progesterone with growth factor and cytokine signaling in breast cancer. Progesterone receptors regulate signal transducers and activators of transcription expression and activity. *J. Biol. Chem.* **273**, 31317–31326 (1998).
113. Stocco, C. O., Chedrese, J. & Deis, R. P. Luteal expression of cytochrome P450 side-chain cleavage, steroidogenic acute regulatory protein, 3β-hydroxysteroid dehydrogenase, and 20α-hydroxysteroid dehydrogenase genes in late pregnant rats: effect of luteinizing hormone and RU486. *Biol. Reprod.* **65**, 1114–1119 (2001).
114. Chen, J. & Nathans, J. Estrogen-related receptor β/NR3B2 controls epithelial cell fate and endolymph production by the stria vascularis. *Dev. Cell* **13**, 325–337 (2007).
115. Yin, G. *et al.* TWISTing stemness, inflammation and proliferation of epithelial ovarian cancer cells through MIR199A2/214. *Oncogene* **29**, 3545–3553 (2010).
116. Loebel, D. A., Tsoi, B., Wong, N. & Tam, P. P. A conserved noncoding intronic transcript at the mouse *Dnm3* locus. *Genomics* **85**, 782–789 (2005).
117. Olson, D. M. *et al.* Myometrial activation and preterm labour: evidence supporting a role for the prostaglandin F receptor—a review. *Placenta* **24** (Suppl. A), 47–54 (2003).
118. Havelock, J. C. *et al.* Human myometrial gene expression before and during parturition. *Biol. Reprod.* **72**, 707–719 (2005).
119. Menon, R. *et al.* Biomarkers of spontaneous preterm birth: an overview of the literature in the last four decades. *Reprod. Sci.* **18**, 1046–1070 (2011).
120. Wittmann, J. & Jäck, H. M. Serum microRNAs as powerful cancer biomarkers. *Biochim. Biophys. Acta* **1806**, 200–207 (2010).
121. Etheridge, A., Lee, I., Hood, L., Galas, D. & Wang, K. Extracellular microRNA: a new source of biomarkers. *Mutat. Res.* **717**, 85–90 (2011).
122. Frost, R. J. & Olson, E. N. Control of glucose homeostasis and insulin sensitivity by the Let-7 family of microRNAs. *Proc. Natl Acad. Sci. USA* **108**, 21075–21080 (2011).
123. van Rooij, E., Marshall, W. S. & Olson, E. N. Toward microRNA-based therapeutics for heart disease: the sense in antisense. *Circ. Res.* **103**, 919–928 (2008).
124. Montgomery, R. L. *et al.* Therapeutic inhibition of miR-208a improves cardiac function and survival during heart failure. *Circulation* **124**, 1537–1547 (2011).
125. Obad, S. *et al.* Silencing of microRNA families by seed-targeting tiny LNAs. *Nat. Genet.* **43**, 371–378 (2011).
126. Small, E. M. & Olson, E. N. Pervasive roles of microRNAs in cardiovascular biology. *Nature* **469**, 336–342 (2011).
127. Concepcion, C. P., Bonetti, C. & Ventura, A. The microRNA-17-92 family of microRNA clusters in development and disease. *Cancer J.* **18**, 262–267 (2012).
128. Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
129. Ruvkun, G. The perfect storm of tiny RNAs. *Nat. Med.* **14**, 1041–1045 (2008).

Acknowledgements
The authors' research was supported by the National Institutes of Health (NIH 5-P01-HD011149) and the March of Dimes Foundation (Prematurity Research Grant No. 21-FY11-30).

Author contributions
The authors contributed equally to all aspects of this article.

Glotov et al. BMC Systems Biology 2015, 9(Suppl 2):S4
http://www.biomedcentral.com/1752-0509/9/S2/S4

BMC Systems Biology

RESEARCH Open Access

Molecular association of pathogenetic contributors to pre-eclampsia (pre-eclampsia associome)

Andrey S Glotov[1,2], Evgeny S Tiys[3,4], Elena S Vashukova[1], Vladimir S Pakin[1], Pavel S Demenkov[3,4], Olga V Saik[3,4], Timofey V Ivanisenko[3,4], Olga N Arzhanova[1], Elena V Mozgovaya[1], Marina S Zainulina[1], Nikolay A Kolchanov[3,4], Vladislav S Baranov[1,2], Vladimir A Ivanisenko[3,4*]

From IX International Conference on the Bioinformatics of Genome Regulation and Structure\Systems Biology (BGRS\SB-2014)
Novosibirsk, Russia. 23-28 June 2014

Abstract

Background: Pre-eclampsia is the most common complication occurring during pregnancy. In the majority of cases, it is concurrent with other pathologies in a comorbid manner (frequent co-occurrences in patients), such as diabetes mellitus, gestational diabetes and obesity. Providing bronchial asthma, pulmonary tuberculosis, certain neurodegenerative diseases and cancers as examples, we have shown previously that pairs of inversely comorbid pathologies (rare co-occurrences in patients) are more closely related to each other at the molecular genetic level compared with randomly generated pairs of diseases. Data in the literature concerning the causes of pre-eclampsia are abundant. However, the key mechanisms triggering this disease that are initiated by other pathological processes are thus far unknown. The aim of this work was to analyse the characteristic features of genetic networks that describe interactions between comorbid diseases, using pre-eclampsia as a case in point.

Results: The use of ANDSystem, Pathway Studio and STRING computer tools based on text-mining and database-mining approaches allowed us to reconstruct associative networks, representing molecular genetic interactions between genes, associated concurrently with comorbid disease pairs, including pre-eclampsia, diabetes mellitus, gestational diabetes and obesity. It was found that these associative networks statistically differed in the number of genes and interactions between them from those built for randomly chosen pairs of diseases. The associative network connecting all four diseases was composed of 16 genes (*PLAT, ADIPOQ, ADRB3, LEPR, HP, TGFB1, TNFA, INS, CRP, CSRP1, IGFBP1, MBL2, ACE, ESR1, SHBG, ADA*). Such an analysis allowed us to reveal differential gene risk factors for these diseases, and to propose certain, most probable, theoretical mechanisms of pre-eclampsia development in pregnant women. The mechanisms may include the following pathways: [TGFB1 or TNFA]-[IL1B]-[pre-eclampsia]; [TNFA or INS]-[NOS3]-[pre-eclampsia]; [INS]-[HSPA4 or CLU]-[pre-eclampsia]; [ACE]-[MTHFR]-[pre-eclampsia].

Conclusions: For pre-eclampsia, diabetes mellitus, gestational diabetes and obesity, we showed that the size and connectivity of the associative molecular genetic networks, which describe interactions between comorbid diseases, statistically exceeded the size and connectivity of those built for randomly chosen pairs of diseases. Recently, we have shown a similar result for inversely comorbid diseases. This suggests that comorbid and inversely comorbid diseases have common features concerning structural organization of associative molecular genetic networks.

* Correspondence: salix@bionet.nsc.ru
[3]The Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
Full list of author information is available at the end of the article

Background

Pre-eclampsia (PE) is the leading cause of maternal and foetal morbidity and mortality. It is a pregnancy complication, predominantly occurring after 20-weeks of gestation, as well as in labour, and it is characterized by multiple organ dysfunction syndromes, including the dysfunction of the kidneys, liver, vascular and nervous systems, and the foetoplacental complex [1,2]. The general clinical symptoms of PE are oedema, proteinuria and hypertension. The clinical outcome of PE may not always be predictable. Either form of PE can be extremely insidious, rapidly progressing, and, even in the absence of one of its general symptoms, may lead to life threatening complications for the mother and foetus [3]. In 70-80% of cases, PE is secondary to an underlying disease [1]. Pre-eclampsia risk factors include cardiovascular diseases (arterial hypertension), kidney, liver and gastrointestinal tract diseases, endocrine disorders (obesity, diabetes mellitus), and autoimmune diseases (antiphospholipid syndrome) [1,3,4]. According to meta-analysis data, women with a history of PE have 1.79 times the risk of venous thromboembolism, 1.81 times the risk of stroke, 2.16 times the risk of ischemic heart disease and 3.7 times the risk of hypertensive disease, compared with women without PE [5]. Thus far, it remains unclear whether the presence of pathological processes before pregnancy predisposes one to PE, or whether defects in multiple organs and systems, induced by PE, are responsible for the development of extragenital diseases in the future. Such joint manifestations of diseases are called comorbidities [6] or syntropies [7]. Likewise, inversely comorbid [8] or dystropic [9] diseases statistically rare co-occur in patients as compared with co-occurrence that can be expected by chance. Previously, for asthma, tuberculosis, certain cancers and neurodegenerative diseases, we have shown that inversely comorbid diseases are more closely related to each other at the molecular level in comparison with randomly chosen pairs of diseases [10].

In recent years, bioinformatics methods have been widely used for modelling different pathological processes, analysing the molecular mechanisms of their development, identifying possible markers, and systematizing available data. Ample evidence regarding the influence of genetics on comorbidities has accumulated in the literature. Computer-based, text-mining methods were developed for efficient extraction of knowledge from the scientific literature. At the present time, COREMINE and MeSHOPs, which analyse the co-occurrence of biomedical terms [11,12], and STRING [13] and the MedScan system, which are based on the parsing of natural language texts [14], are widely used.

We have developed the ANDSystem, which was designed for the automated extraction of knowledge from natural language texts regarding the properties of molecular biological objects and their interactions in living systems [15]. Using this system, we have reconstructed protein-protein networks for proteins that are associated with water-salt metabolism and sodium deposition processes in healthy volunteers [16], as well as protein-protein interaction networks for *Helicobacter pylori*, which are associated with the functional divergence of *H. pylori*, isolated from patients with early gastric cancer [17]. We have also reconstructed associative networks representing molecular genetic interactions between proteins, genes, metabolites and molecular processes associated with myopia and glaucoma [18], and with cardiovascular diseases [19].

In the current study, we used the ANDSystem for the reconstruction of associative networks (the preeclampsia associome) representing molecular genetic interactions between genes associated with PE, diabetes mellitus (DM), gestational diabetes (GD) and obesity (Ob). We conducted an analysis of these networks to reveal differential and common risk factors for these diseases.

Finding pathways common to the indicated multifactorial diseases would contribute to a better understanding of the characteristic features of pre-eclampsia pathogenesis, as well as to the development of new diagnostic, preventative and therapeutic methods.

Results

Pre-eclampsia: its association, via comorbid genes, with diabetes mellitus, obesity and gestational diabetes

The main goal of the current study was to identify comorbid genes whose dysfunction or mutation represent common risk factors for diseases that are concurrent with PE. To this end, we relied on published data [3,4] regarding the four most significant and widespread pathologies concurrent with PE: DM, Ob, GD and pyelonephritis. Furthermore, using the ANDSystem and Pathway Studio software, we reconstructed associative networks (disease-protein/gene-disease) comprising interactions between the above diseases via their associated genes. Subsequently, reduction was achieved by eliminating pyelonephritis, as genes associated with nephritis were not associated with PE and the other analysed disorders. Using the ANDSystem, we identified 1,051 proteins/genes associated with PE, Ob, DM and/or GD. Using Pathway Studio, 1,138 proteins/genes were identified. The results of both programs were in good agreement regarding the number of genes in groups associated with particular diseases (Figure 1). Unfortunately, we were not able to use STRING for the reconstruction of such networks, as this program does not provide information about protein/gene-disease associations.

The number of proteins/genes common to different combinations of the examined diseases is shown in Figure 1. We assumed that comorbid diseases are more

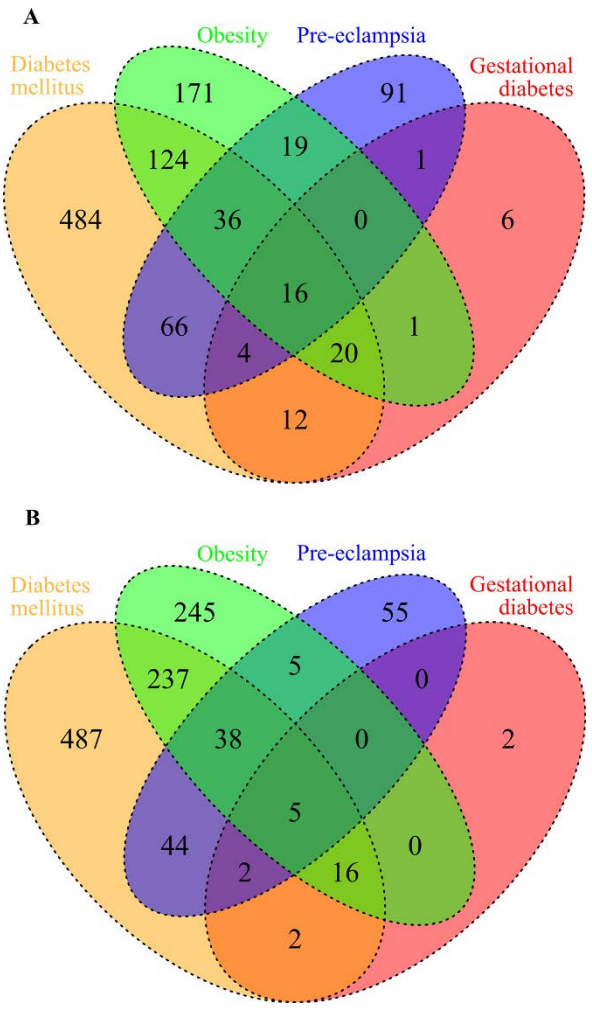

Figure 1 Venn diagram demonstrating the intersections of the lists of genes associated with the analysed diseases (PE, DM, GD, Ob) according to the ANDSystem (A) and Pathway Studio (B).

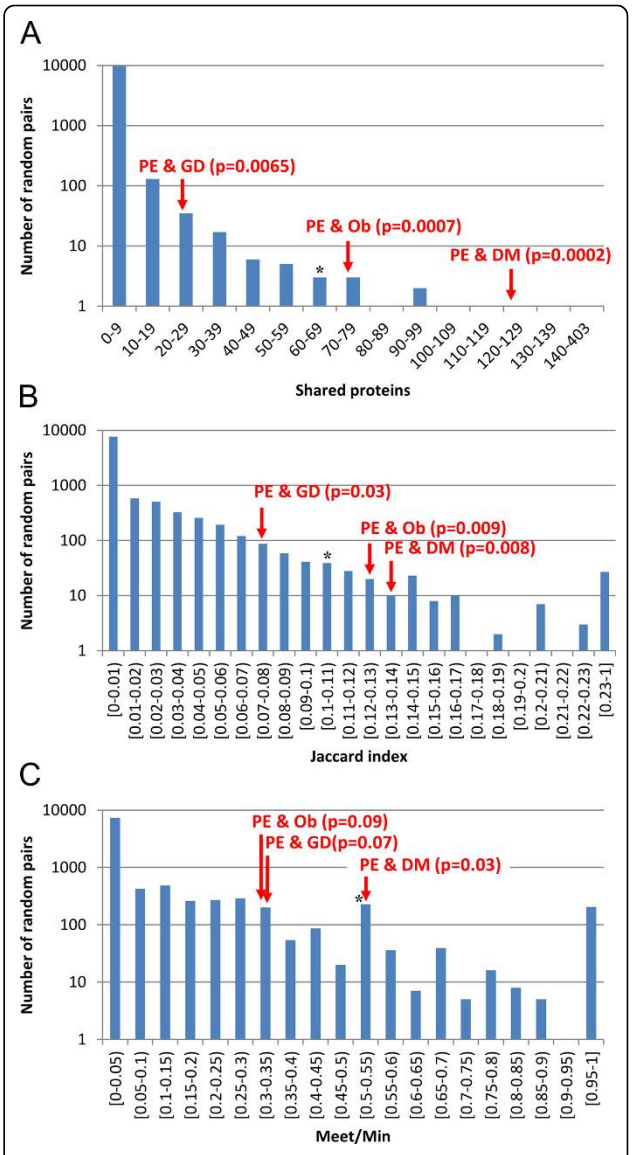

Figure 2 Comparison of analysed and random networks by intersection (A), Jaccard (B) and Meet/Min (C) indices. Bars show the distribution of the value for the features of the associative networks for randomly chosen disease pairs. Arrows indicate PE & GD, PE & DM and PE & Ob comorbid disease pairs. Asterisks indicate the position of inversely comorbid disease pairs (bronchial asthma and pulmonary tuberculosis) [10].

closely interrelated, via the common proteins/genes associated with them, as compared with randomly chosen disease pairs. To test this assumption, we calculated the distribution of three relation indices of random disease-protein/gene-disease networks built for random disease pairs: I_{AB} (number of shared proteins), J_{AB} (Jaccard index) and M_{AB} (Meet/Min). All three disease pairs (PE & DM, PE & GD, PE & Ob) were significantly connected by the I_{AB} and J_{AB} indices at $p < 0.05$ (Figure 2). Only PE & DM pair was significantly different by M_{AB} index ($p < 0.05$) from randomly generated pairs of diseases. Thus, PE and DM were found to be the most significantly associated disease pair for all three relation indices.

Next, we tested the hypothesis whether comorbid proteins/genes common to comorbid disease pairs interact more closely compared to a set of randomly chosen proteins/genes. Comparison of the associative molecular genetic networks with random ones demonstrated that the networks that describe the interactions between the comorbid proteins/genes for all three disease pairs (PE & GD, PE & DM, and PE & Ob) exhibited significantly greater connectivities than those of the random networks ($p < 0.001$).

Of particular interest was an appended analysis of the associative molecular genetic networks built for proteins associated concurrently with four comorbid diseases (PE, DM, Ob and GD). The three programs used to

build this network were the ANDSystem, Pathway Studio and STRING (Figure 3). As Figure 3A shows, the ANDSystem network comprised 32 objects: 16 proteins and 16 genes, as well as 142 interactions. The ANDSystem has an advantageous feature: an object pair can also be associated concurrently with links of several types. For this reason, the number of associated object pairs, 87, was smaller than the number of links. The ANDSystem

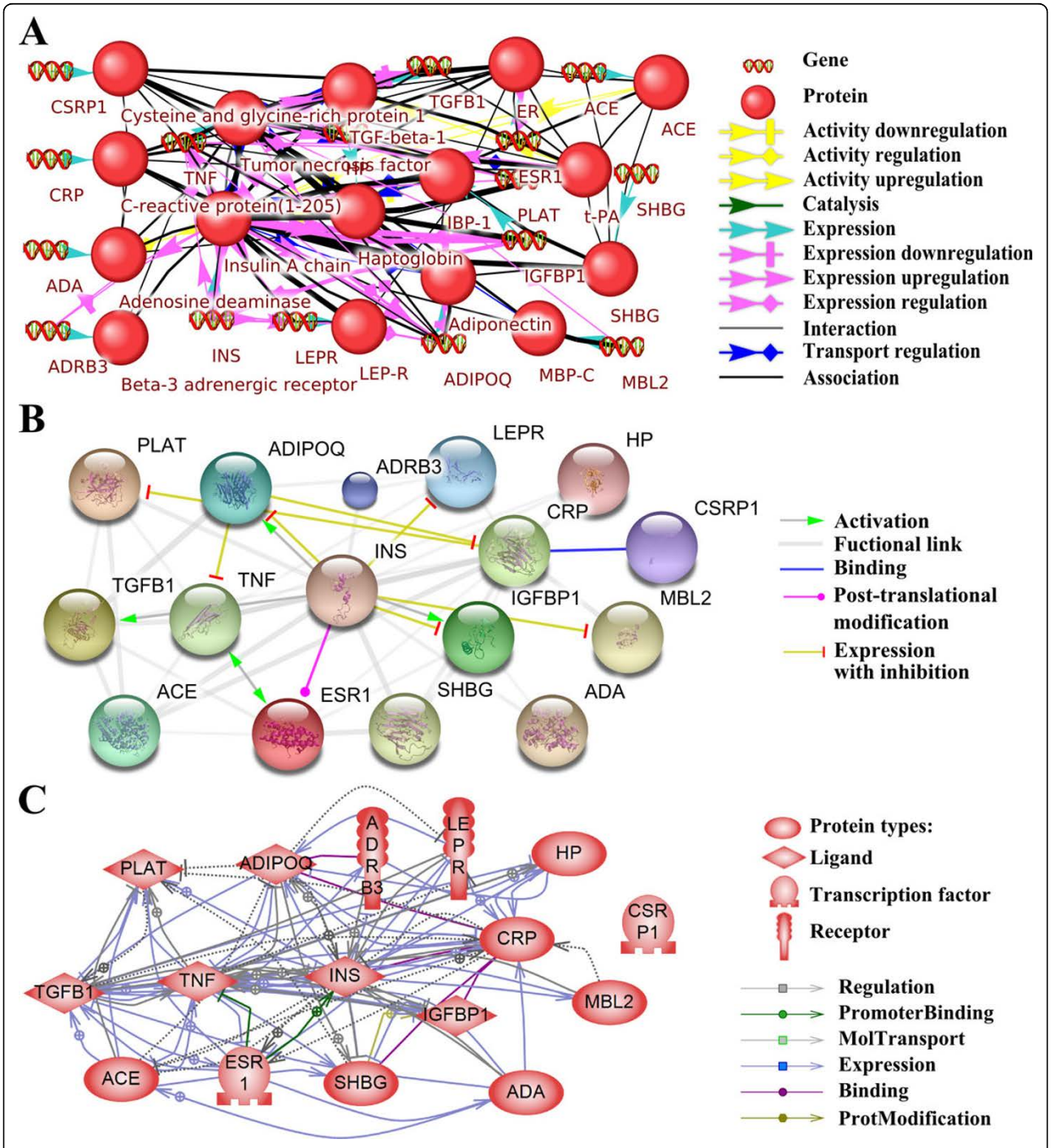

Figure 3 ANDSystem (A), STRING (B) and Pathway Studio (C) networks describing shared proteins/genes. Interactions between proteins/genes concurrently shared in PE, DM, Ob and GD are shown. In the ANDSystem network, proteins are shown as balls, and patterns with double helices designate genes. In the STRING network, proteins/genes are shown as balls (gene names are given). The colour and shape of the arrows indicate the type of association between the objects in all the networks.

represented cases of the regulation of protein activity (six links), including up-regulation (two links) and down-regulation (three links) of protein activity; gene expression regulation (37 links), including up-regulation (seven links) and down-regulation (seven links); protein-protein interactions (two links); protein transport regulation (10 links); catalysis (one link); expression (16 links) and association (70 links). To compare the ANDSystem network with those of the STRING and Pathway Studio, the ANDSystem network was transformed into a protein-protein interaction network, with links from the genes assigned to their respective proteins, while links from genes as separate vertices were deleted from the network. Such a network contained 45 interconnected protein pairs.

The STRING network (Figure 3B) contained 16 proteins/genes, and 45 gene pairs connected by 47 links, including five different types: activation (four links), expression with inhibition (seven links), binding (one link), post-translational modification (one link), and functional links (34 links). The functional links in STRING were determined on the basis of Neighbourhood in the Genome, Gene Fusions, Co-occurrence Across Genomes, Co-Expression, Experimental/Biochemical Data, Association in Curated Databases, and Co-Mentioned in PubMed Abstracts [13].

The network built by Pathway Studio (Figure 3C) contained 16 proteins/genes, and 62 pairs of genes connected by 98 links, including six different types: binding (five links), expression (55 links), molecular transport (19 links), promoter binding (two links), protein modification (one link) and regulation (16 links).

There was a significant difference between the comorbid and random networks ($p < 0.001$), not only for disease pairs, but also for the associative molecular genetic networks that describe the interactions between proteins/genes associated concurrently with all four diseases (PE, DM, GD, Ob) (Figure 3A). These results demonstrated that comorbid proteins/genes are presumably involved in shared biological processes. This can explain the increased number of interactions between proteins/genes, as compared with those for associative molecular genetic networks of randomly chosen proteins/genes. Confirmation of this hypothesis would shed light on the molecular mechanisms underlying the interactions between comorbid diseases.

Analysis of overrepresentation of Gene Ontology (GO) processes

Overrepresentation of GO biological processes was analysed for the group of proteins/genes associated with single diseases (PE, DM, GD and Ob) and pairs of diseases (PE & DM, PE & GD, PE & Ob), as well as concurrently with all four diseases. In each of these cases, more than 1,000 overrepresented processes were found (Additional file 1). Among these were a high number of quite general processes for which thousands of genes have been annotated. The connectivity rate (CR) was calculated for each process listed in Additional file 1 to check how closely the proteins/genes, which caused an overrepresentation of processes, interacted. After ranking the overrepresented biological processes according to the CRs, 313 processes had the highest CR (equal to 1) (see Additional file 1). Just as expected, generalized, nonspecific biological processes had smaller CR values in the majority of cases as compared with specialized processes involving a relatively small number of genes.

Among the overrepresented biological processes with a maximum CR were positive regulation of monooxygenase activity, regulation of fat cell differentiation, regulation of lipid metabolic process, nitric oxide and carbon monoxide metabolism, regulation of protein kinase B signalling cascade, regulation of NF-kappa B transcription factor activity, regulation of glucose metabolism and transport, regulation of cellular response to oxidative stress, regulation of cytokine production, regulation of cell cycle process and others. Thus, the use of the CR index in the GO enrichment analysis revealed the specific GO processes and lower the rank of less informative general processes.

Reconstruction of associative pathways describing potential molecular mechanisms via comorbid genes involved in overrepresented GO biological processes

The next step of the current study was to reconstruct the molecular pathways connecting PE with DM, Ob, and GD, via interactions between the specific and comorbid genes. The Pathway Discovery module of the ANDVisio software was used to trace separate pathways in the network of molecular genetic interactions associated concurrently with all four pathologies. The Pathway Discovery module was used to search for pathways in the network using patterns set by the user.

The patterns were of the following type: <PE> - <any protein/gene specific to PE> - <any comorbid protein/gene> - <any protein/gene specific to Ob or GD, or DM> - <Ob or GD, or DM>. The program chose all the pathways meeting the pattern's criteria: the starting link was PE; the second link of the chain should be one of the proteins/genes associated with PE, exceptions were proteins/genes comorbid for all four diseases (4-comorbid); the third link should be one of the 4-comorbid proteins/genes (PLAT, ADIPOQ, ADRB3, LEPR, HP, TGFB1, TNFA, INS, CRP, CSRP1, IGFBP1, MBL2, ACE, ESR1, SHBG, ADA); the fourth link should be one of the proteins specific to Ob, GD, or DM, with the exception of 4-comorbid proteins/genes. The last link should be one of the diseases (Ob, GD or DM). The total number of identified pathways was more than 50. These were combined into a single pathway network (Figure 4).

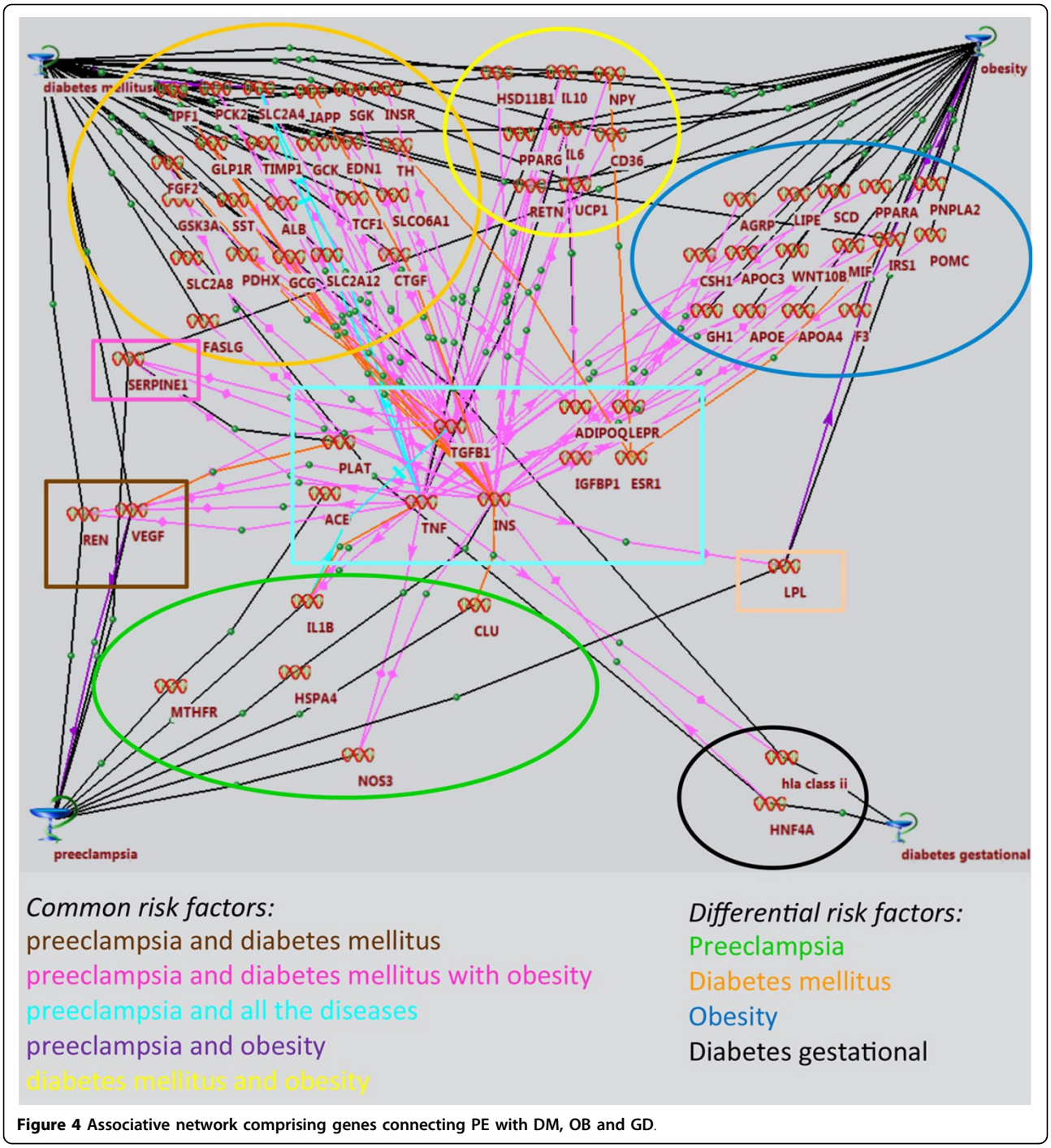

Figure 4 Associative network comprising genes connecting PE with DM, OB and GD.

Common, as well as specific, risk factors were distinguished for the following combinations of diseases: PE and DM; PE and Ob; PE and DM, Ob; PE and DM, Ob, GD; (see Figure 4). The largest number of connections was obtained for the *TNFA, TGFB1* and *INS* genes, which revealed specialized GO processes with maximum CRs, such as: «positive regulation of protein kinase B signalling», «cascade regulation of NF-kappa B transcription factor activity», «regulation of mitosis», «regulation of nuclear division», «regulation of protein secretion MAPK cascade», «positive regulation of protein transport», «regulation of protein complex assembly», «positive regulation of cell migration», «positive regulation of secretion», «positive regulation of cellular component movement», «positive regulation of organelle organization», «regulation of mitotic cell cycle»,

«regulation of immune effector process», «intracellular protein kinase cascade», «regulation of cellular component biogenesis», «regulation of cell cycle process», «regulation of organelle organization», «regulation of cell cycle» (see Additional file 1).

An associative pathway network connecting PE, via the *PLAT, ADIPOQ, LEPR, TGFB1, TNFA, INS, IGFBP1, ACE* and *ESR1* genes, with DM, OB and GD incorporated 66 genes with 167 connections (see Figure 4). Most of these connections (78) corresponded to the "association" type (shown in black). Sixty-nine of them could be referred to "expression regulation" types and 13 as "co-expression" (shown in red); eight comprised "down regulation", "degradation regulation", and "degradation downregulation" (shown in violet).

The differential network of PE risk factors included seven genes (interleukin-1-beta (IL1B), endothelial (NOS3) NO-synthase, heat shock 70 kDa protein 4 (HSPA4), apolipoprotein J (CLU) and 5,10-methylenetetrahydrofolate reductase (MTHFR).

Thus, whereas all the identified PE risk factors might be treated as potential markers of this disease, the most probable molecular mechanism underlying PE, DM, OB and GD includes the pathway starting from the *TGFB1, TNFA, INS* and *ACE* genes, through the *IL1B, NOS3, HSPA4* (HSP74), *CLU* and *MTHFR* genes, and eventually to PE.

Thus, the probable chains of molecular events on the way to combined PE, in this context, are as follows: TGFB1 or TNFA - IL1B - PE; TNFA or INS - NOS3 - PE; INS - HSPA4 or CLU - PE; ACE - MTHFR - PE.

Discussion

The associative networks analysed in this work (see Figures 1, 2, 3, 4) appeared to be significant for the understanding of the nature of PE, thereby supporting the hypothesis that PE represents a stable complex of clinical manifestations [1,3,4]. The key players in the reconstructed networks are comorbid genes which, on the one hand, contribute to the development of PE and its pathogenically related disorders, and, on the other hand, may play the role of "triggers" in the presence of other pre-eclampsia-promoting factors (genes and proteins). Comorbid genes are characteristic of many multifactorial diseases [20]. Moreover, many comorbid diseases may involve various pathophysiological mechanisms [20], and the construction of associative networks makes it possible to understand their molecular interrelations.

An analysis of reconstructed associative networks, which describe interactions between comorbid proteins/genes associated with different pair combinations of PE with DM, Ob, and GD, demonstrated that comorbid diseases differ in a statistically significant manner from random disease pairs. The differences concern both the number of common genes associated with the diseases and the interactions between such genes. The number of vertices in the comorbid networks, as well as the number of interactions between the vertices, exceeded those of random disease pairs. At the same time, the density of connections in the associative molecular genetic network describing the interactions between proteins/genes associated concurrently with all four diseases also differed significantly from those of the random networks formed by random sets of proteins/genes. Interestingly, we also observed the same regularity for inversely comorbid diseases [10]. It has been shown that the associative networks reconstructed for pairs of inversely comorbid diseases, including bronchial asthma and pulmonary tuberculosis, as well as nine pairs formed by neurodegenerative (Parkinson disease, schizophrenia, Alzheimer disease) and cancer diseases (colorectal neoplasms, prostatic neoplasms, lung neoplasms), significantly differed from the networks that describe interactions between random diseases. An example of the mutual arrangement of inversely comorbid (bronchial asthma and pulmonary tuberculosis) and comorbid diseases is shown in Figure 2.

Our current results are in many respects consistent with those of epidemiological studies worldwide. It has been amply demonstrated that the common risk factors of PE were DM, Ob and GD [1,2,21-26]. In most studies, DM is a leading risk factor, as it occurs in more than half of the women with PE [1,2,24]. Furthermore, DM is more strongly associated with a late-onset of the disease, which prevails among all the cases [24,25]. A study of twin gestations supports our reasoning. In this study, an evaluation of associated factors in PE gestations and a comparison of the incidence of pregnancy complications among twins with and without PE demonstrated that a high pregnancy body mass index (BMI) and diabetes were associated with PE [27].

We identified 16 genes encoding shared proteins in the molecular network, built using the literature- and database-mining (ANDSystem, Pathway Studio and STRING), that simultaneously connected with PE, DM, GD and Ob. Most shared genes determined in this study encode proteins controlling energy metabolism, and are associated with the immune response and inflammation.

An analysis of the associations of these genes with PE and DM, GD and Ob obtained in case-control, family-based, and meta-analyses studies, which we conducted using the HuGE Navigator, revealed that 14 of the 16 shared genes were associated with at least one of the diseases (see Table 1). Two genes (*CSRP1* and *PLAT*) had never been shown to be associated with PE and DM, GD and Ob. Four shared genes (*ACE, ADIPOQ,*

Table 1. Statistics of gene-disease associations for PE, DM, GD and Ob obtained with the HuGE Navigator

Gene name	PE	DM	GD	Ob
ACE	39	244	2	77
ADA	-	6	-	-
ADIPOQ	4	156	4	176
ADRB3	1	49	4	145
CRP	2	20	-	28
CSRP1	-	-	-	-
ESR1	7	21	-	36
HP	2	36	-	5
IGFBP1	-	7	-	5
INS	1	88	4	26
LEPR	7	35	1	154
MBL2	4	14	1	1
PLAT	2	3	-	1
SHBG	-	6	1	7
TGFB1	8	33	-	8
TNFA	24	132	5	83

The number of associations determined by case-control, family-based and meta-analysis studies are shown.

MBL2, TNFA) were found to be associated with all the diseases.

We believe that the identification of these genes in the current study is of importance because they encode proteins important for the development of diseases, as confirmed by experimental studies (Table 1).

Angiotensin-converting enzyme (ACE) plays a key role in regulating blood pressure by influencing vascular tone by activating the vasoconstrictor angiotensin II and inactivating the vasodilatory peptide bradykinin. Inter-individual differences in blood ACE levels are at least in part explained by the presence of an insertion/deletion (I/D) polymorphism in intron 16 of the ACE gene, with higher ACE levels observed in D allele carriers. The results of many studies confirmed the association of ACE polymorphism with PE [28]. Other studies have indicated that the ACE gene is a factor that contributes to the manifestation of GD [29], diabetic nephropathy and Ob [30,31].

It has been shown that polymorphisms in the adiponectin gene (*ADIPOQ*) modulate the circulating concentration of adiponectin. Abnormal adiponectin levels, as well as *ADIPOQ* polymorphisms, have been associated with PE [32]. Some variants of this gene are associated with the occurrence of GD [33], while other polymorphisms may contribute to type 2 DM risk [34] and Ob in adults [35].

Mannose-binding lectin (MBL) is involved in the maintenance of an inflammatory environment in the uterus. High MBL levels have been associated with successful pregnancies, whereas low levels are involved in PE development. Association between polymorphisms in the structural and promoter regions of the *MBL2* gene and PE have been evaluated [36]. MBL gene polymorphisms are associated with GD and with type 2 DM [37,38]; in addition, MBL deficiency may confer a risk of Ob and insulin resistance [39].

Tumour necrosis factor-alpha (TNF-α) participates in the immune response and inflammation. Many studies have showed that there is an association between the *TNFA* gene and PE among Europeans [2,40]. The -308 G−>A polymorphism of the *TNFA* promoter gene is involved in the pathophysiology of insulin resistance and GD [41]. The same polymorphism is a genetic risk factor for the development of type 2 DM [42]. Individuals who carry the -308A *TNFA* gene variant have a 23% greater risk of developing obesity compared with controls, and they showed significantly higher systolic arterial blood pressure and plasma insulin levels, supporting the hypothesis that the *TNFA* gene is involved in the pathogenesis of the metabolic syndrome [43].

The PE associome contains more links than each of the individual networks. The identified, shared genes have been classified according to GO. Such a network was needed for a GO overrepresentation analysis. The presence of processes identified by the GO analysis in the pathogenesis of PE is not surprising. The central hypothesis of our understanding of PE is that it results from ischaemia of the placenta, which in turn releases factors into the maternal circulation that are capable of inducing the clinical manifestations of the disease [2]. Multiple pathogenetic mechanisms have been implicated in this disorder, including an imbalance between angiogenic and anti-angiogenic factors, autoantibodies to the type-1 angiotensin II receptor, platelet and thrombin activation, defective deep placentation, intravascular inflammation, endothelial cell activation and/or dysfunction, and oxidative and endoplasmic reticulum stress that promote the differentiation of trophoblasts from a proliferative to an invasive phenotype, regulate cell homeostasis through their involvement in post-translational modifications and protein folding, and induce the release of proinflammatory cytokines and chemokines. Other mechanisms include hypoxia and trophoblast invasion, which down-regulate the expression of transforming growth factor β3 (TGF-β3) and hypoxia-inducible factors (HIF-1α and HIF-2α) [2,44]. These results indicated the contribution of common, non-specific, pathological processes to the development of PE, DG, GD and Ob.

In addition to the identification of common proteins/genes associated with different pathological processes, another goal of the study was to find unique markers for PE. To do so, we reconstructed potential mechanisms of molecular interactions using the ANDSystem software, a program that allows the identification of the largest number of links (see Figure 4). Although the central network

core of these pathways contained only nine common genes (*PLAT, ADIPOQ, LEPR, TGFB1, TNFA, INS, IGFBP1, ACE, ESR1*), it incorporated 68 genes with 174 connections between them, and differential factor risks of PE were identified: the *IL1B, NOS3, HSPA4, CLU* and *MTHFR* genes. The contributions of many of these genes to the pathogenesis of PE has been confirmed by numerous studies [2,45-50]. Here, we showed for the first time that these genes can be specifically involved in the pathogenesis of PE. However, it is not yet clear why these genes have a greater involvement in PE. The possible trigger mechanisms of combined PE are linked to the processes that are carried out by the products of the identified genes, namely, inflammation (*IL1B*), endothelial dysfunction (*NOS3*), heat shock and stress (*HSPA4*), stabilizing cell membranes at diverse fluid-tissue interfaces and protecting the vascular endothelium from an attack by some factors in plasma, such as active complement complexes (*CLU*), and homocysteine metabolism (*MTHFR*).

In addition, the results are of particular importance in regard to the theory of confounding assumptions as false mechanisms of genetic association when the factor is associated with a confound, but not the phenotype, and a confound, in turn, is associated with the phenotype [51,52]. The identified genes can act as such a confound.

Conclusions

The current results broaden our knowledge of the molecular mechanisms of the interactions between comorbid diseases. This reconstruction of associative molecular genetic networks that describe interactions between PE and comorbid diseases (GD, Ob, and DM) differed significantly from partner networks built for random disease pairs. Networks between PE and comorbid diseases had a larger number of genes and links between them. With this in mind, it is of interest that similar features of associative network structure have been observed for inversely comorbid diseases [10]. It can be suggested that comorbid and inversely comorbid relationships between diseases involve larger sets of closely interrelated genes larger than those for random pairs of diseases. In the future, we intend to perform a scale analysis that connects different disease pairs to detect potential comorbid/inversely comorbid diseases for all the possible disease pairs via which these diseases can interact. Reconstruction and analysis of the PE associome is useful for revealing the genetic factors involved in the pathogenesis of the disease and for identifying its differential risk factors, as well as for modelling the theoretical mechanisms of PE development in pregnant women with underlying diseases, such as DB, Ob or GD.

Methods

We used three systems that allowed the automated reconstruction of networks that describe the interactions between proteins/genes and diseases: STRING [13], Pathway Studio [14] and ANDSystem [15].

The ANDSystem was developed for the automated extraction of facts and knowledge regarding the relationships between proteins, genes, metabolites, microRNAs, cellular components, molecular processes, and their associations with diseases from published scientific texts and databases. To extract knowledge from texts in the ANDSystem, the shallow parsing method was applied. Pathway Studio is a software application developed for the navigation and analysis of biological pathways, gene regulation networks and protein interaction maps. The program uses the natural language processing approach to extract knowledge from the texts of scientific publications. STRING is a database and a web resource that contains information about protein-protein interactions (including physical and functional interactions) that is mainly based on the use of text-mining methods.

The associative networks for the considered disease pairs were graphs whose vertices were diseases and human proteins/genes, while the edges were the associations between diseases and proteins.

The following indices of relation between a pair of associative networks were used: (1) the intersection index, $I_{AB} = |A \cap B|$ equal to the intersection size of protein sets A and B composed of proteins concurrently associated with diseases D_A and D_B; (2) the Jaccard index [53] was calculated as the ratio of I_{AB} to the combination of sets A and B involving at least one of the diseases D_A and D_B, $J_{AB} = \frac{I_{AB}}{|A \cup B|}$; (3) Meet/Min [54] was calculated as $M_{AB} = \frac{I_{AB}}{\min(|A|, |B|)}$, where the denominator denotes the size of the minimum of sets A and B.

The statistical significance of the relation indices for the analysed diseases in the associative networks was determined by comparing these networks with the associative ones formed by pairs of randomly chosen diseases. For such an analysis, we used the ANDSystem because this program allows the comparison of reconstructed networks with random ones generated using the ANDCell knowledge base. All the interactions between proteins, genes, metabolites, diseases and other objects described by the ANDSystem are deposited in the ANDCell knowledge base, which is a module of this system [15]. The total number of diseases described in ANDCell was 4,075; of these, 991 were not found to be associated with any human protein. Such diseases were discarded from the analysis. To build the distribution of the relation indices for random disease pairs, 10,000 random disease pairs were generated (see Additional file 2). The P-value for the analysed disease pairs was calculated as the

proportion of 10,000 random networks with the same or larger CR as in the examined pairs of diseases. The associative networks were reconstructed using the ANDSystem and Pathway Studio programs. STRING was not used for this purpose because it gave no information regarding interactions between protein/gene and diseases. The associative networks for the analysed disease pairs included only interactions of the disease-protein/gene type; the interactions between proteins/genes were discounted. As a result, to analyse the interactions between proteins/genes in the associative networks, additional protein/gene-protein/gene associative molecular genetic networks were built using the ANDSystem, Pathway Studio and STRING. The statistical significance of the connectivity of the associative molecular genetics networks built for the analysed disease pairs was also determined by comparing them with random networks. In such a case, for each analysed associative molecular genetic networks, 1,000 random networks were generated using the ANDSystem (only human proteins/genes were considered).

The statistical significance (p-value) of the difference between the connectivity of the analysed network and that of the random networks was also determined, like in the case of the associative networks, as the proportion of random networks with the same or greater number of links between the vertices compared with the number of links in the analysed network. The random molecular genetic networks were built according to the following rules. Proteins/genes considered as vertices in the random networks were taken from the ANDCell knowledge base. To ensure that the proteins/genes in the random networks were represented at a level of study close to that of the proteins/genes from the analysed networks, we considered only those random proteins/genes whose connectivity rate was the same as connectivity rate of proteins/genes from the analysed networks. The set Q_i was formed for each i-th vertex of the analysed network. Q_i was composed of all the proteins/genes from the ANDCell knowledge base having an interaction number in ANDCell equal to the protein/gene interactions in the knowledge base represented by the i-th vertex. The protein/gene for the i-th vertex of the random network was chosen by chance for the set Q_i . The links between the vertices in the random networks were set according to the interactions described in the ANDCell knowledge base.

The results of the automated extraction of information regarding the interactions between proteins/genes and diseases were tested manually. The recognition correctness of the object names in the text, as well as the presence of their interactions, was tested. The lists of shared and specific proteins were reduced by expert evaluation to retain only those participating in the pathogenesis of both diseases for shared proteins, and in the pathogenesis of either disease for specific proteins, as shown previously [10].

The BINGO tool [55] was used to evaluate the overrepresentation of the biological processes for the considered protein/gene set. The enrichment was evaluated using a hypergeometric test with the Benjamini and Hochberg FDR correction using the whole annotation as a reference set. The human Uniprot-GOA Gene Association file (release 2013_05) was used as the custom annotation file. In addition to the statistical significance of the overrepresentation, the overrepresented GO processes were characterized by the CR of the respective proteins/genes in the associative molecular genetics network built for intersection of the four studied diseases. The CR for the protein group of the examined network involved in the overrepresented GO biological process was calculated as the ratio of the number of the protein pairs connected by the network protein pairs of the given group to the number of all possible pairwise combinations of proteins of this group. As is known, the reconstruction quality of the molecular genetic networks is related frequently to the problem of the completeness of information regarding the interactions between proteins. For this reason, to build the network, we took advantage of three independent programs: ANDSystem, Pathway Studio and STRING, with their parameters set by default.

Additional material

Additional file 1: Excel spreadsheet file containing information regarding the characteristics of overrepresented biological processes.

Additional file 2: Excel spreadsheet file containing information regarding the distribution of the relation indices of the disease-protein-disease associative networks.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Expert analysis of the pathogenetic contributors (diabetes mellitus, gestational diabetes and obesity) was done by ASG, ESV, VSP, ONA, EVM, MSZ and VSB. The development of methods, programs, calculations and analyses of the structural organization of the molecular genetic networks was done by EST, PSD, OVS, TVI, NAK and VAI. All authors read and approved the final manuscript.

Acknowledgements

The work was supported in part by the Russian Science Foundation grant No. 14-24-00123 (development of methods, programs and reconstruction and analysis of the pre-eclampsia associative networks) and Saint-Petersburg State University grant No. 1.38.79.2012 (expert analysis of the pathogenetic contributors: diabetes mellitus, gestational diabetes and obesity).

Declarations

Publication of this article has been funded by the Russian Science Foundation grant No. 14-24-00123.

This article has been published as part of *BMC Systems Biology* Volume 9 Supplement 2, 2015: Selected articles from the IX International Conference on the Bioinformatics of Genome Regulation and Structure\Systems Biology (BGRS \SB-2014): Systems Biology. The full contents of the supplement are available online at http://www.biomedcentral.com/bmcsystbiol/supplements/9/S2.

Authors' details
[1]Federal State Budget scientific Institution "#34;The Research Institute of Obstetrics, Gynecology and Reproductology named after D.O. Ott", St. Petersburg, Russia. [2]Saint-Petersburg State University, St. Petersburg, Russia. [3]The Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia. [4]Novosibirsk State University, Novosibirsk, Russia.

Published: 15 April 2015

References

1. Bilano VL, Ota E, Ganchimeg T, Mori R, Souza JP: **Risk factors of pre-eclampsia/eclampsia and its adverse outcomes in low- and middle-income countries: a WHO secondary analysis.** *PLoS One* 2014, **9**:e91198.
2. Chaiworapongsa T, Chaemsaithong P, Yeo L, Romero R: **Pre-eclampsia part 1: current understanding of its pathophysiology.** *Nat Rev Nephrol* 2014.
3. Young BC, Levine RJ, Karumanchi SA: **Pathogenesis of pre-eclampsia.** *Annu Rev Pathol* 2010, **5**:173-92.
4. Duckitt K, Harrington D: **Risk factors for pre-eclampsia at antenatal booking: systematic review of controlled studies.** *BMJ* 2005, **330**:565.
5. Bellamy L, Casas J-PP, Hingorani AD, Williams DJ: **Pre-eclampsia and risk of cardiovascular disease and cancer in later life: systematic review and meta-analysis.** *BMJ* 2007, **335**:974.
6. Feinstein AR: **The pre-therapeutic classification of co-morbidity in chronic disease.** *Journal of Chronic Diseases* 1970, **23**:455-468.
7. Pfaundler M, Seht L: **Über Syntropie von Krankheitszuständen.** *Zeitschrift für Kinderheilkunde* 1921, **30**:100-120.
8. Ibáñez K, Boullosa C, Tabarés-Seisdedos R, Baudot A, Valencia A: **Molecular evidence for the inverse comorbidity between central nervous system disorders and cancers detected by transcriptomic meta-analyses.** *PLoS genetics* 2014, **10**:e1004173.
9. Freidin MB, Puzyrev VP: *Syntropic genes of allergic diseases* 2010.
10. Bragina EY, Tiys ES, Freidin MB, Koneva LA, Demenkov PS, Ivanisenko VA, Kolchanov NA, Puzyrev VP: **Insights into pathophysiology of dystropy through the analysis of gene networks: an example of bronchial asthma and tuberculosis.** *Immunogenetics* 2014, **66**:457-65.
11. Jenssen TK, Laegreid A, Komorowski J, Hovig E: **A literature network of human genes for high-throughput analysis of gene expression.** *Nat Genet* 2001, **28**:21-8.
12. Cheung WA, Ouellette BFF, Wasserman WW: **Quantitative biomedical annotation using medical subject heading over-representation profiles (MeSHOPs).** *BMC bioinformatics* 2012, **13**:249.
13. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, Mering C von, Jensen LJ: **STRING v9.1: protein-protein interaction networks, with increased coverage and integration.** *Nucleic Acids Res* 2013, **41(Database)**:D808-15.
14. Nikitin A, Egorov S, Daraselia N, Mazo I: **Pathway studio--the analysis and navigation of molecular networks.** *Bioinformatics* 2003, **19**:2155-7.
15. Demenkov PS, Ivanisenko TV, Kolchanov NA, Ivanisenko VA: **ANDVisio: a new tool for graphic visualization and analysis of literature mined associative gene networks in the ANDSystem.** *In Silico Biol* 2011, **11**:149-61.
16. Larina IM, Kolchanov NA, Dobrokhotov IV, Ivanisenko VA, Demenkov PS, Tiĭs ES, Valeeva OA, Pastushkova LK, Nikolaev EN: **[Reconstruction of associative protein networks connected with processes of sodium exchange' regulation and sodium deposition in healthy volunteers by urine proteome analysis].** *Fiziol Cheloveka* 2012, **38**:107-15.
17. Momynaliev KT, Kashin SV, Chelysheva VV, Selezneva OV, Demina IA, Serebryakova MV, Alexeev D, Ivanisenko VA, Aman E, Govorun VM: **Functional divergence of Helicobacter pylori related to early gastric cancer.** *Journal of proteome research* 2010, **9**:254-67.
18. Podkolodnaya OA, Yarkova EE, Demenkov PS: **Application of the ANDCell computer system to reconstruction and analysis of associative networks describing potential relationships between myopia and glaucoma.** *Russian Journal of Genetics: Applied Research* 2011, **1**:21-28.
19. Sommer B, Tiys ES, Kormeier B, Hippe K, Janowski SJ, Ivanisenko TV, Bragin AO, Arrigo P, Demenkov PS, Kochetov AV, Ivanisenko VA, Kolchanov NA, Hofestädt R: **Visualization and analysis of a cardio vascular disease- and MUPP1-related biological network combining text mining and data warehouse approaches.** *J Integr Bioinform* 2010, **7**:148.
20. Puzyrev VP, Freidin MB: **Genetic view on the phenomenon of combined diseases in man.** *Acta Naturae* 2009, **1**:52-7.
21. Mahaba HM, Ismail NA, El Damaty SI, Kamel HA: **Pre-eclampsia: epidemiology and outcome of 995 cases.** *J Egypt Public Health Assoc* 2001, **76**:357-68.
22. Wendland EM, Duncan BB, Belizán JM, Vigo A, Schmidt MI: **Gestational diabetes and pre-eclampsia: common antecedents?** *Arq Bras Endocrinol Metabol* 2008, **52**:975-84.
23. Schneider S, Freerksen N, Röhrig S, Hoeft B, Maul H: **Gestational diabetes and pre-eclampsia–similar risk factor profiles?** *Early Hum Dev* 2012, **88**:179-84.
24. Ornaghi S, Tyurmorezova A, Algeri P, Giardini V, Ceruti P, Vertemati E, Vergani P: **Influencing factors for late-onset pre-eclampsia.** *J Matern Fetal Neonatal Med* 2013, **26**:1299-302.
25. Lisonkova S, Joseph KS: **Incidence of pre-eclampsia: risk factors and outcomes associated with early-versus late-onset disease.** *Am J Obstet Gynecol* 2013, **209**:544-e1.
26. Dadelszen P von, Magee LA: **Pre-eclampsia: an update.** *Curr Hypertens Rep* 2014, **16**:454.
27. Lučovnik M, Tul N, Verdenik I, Novak Z, Blickstein I: **Risk factors for pre-eclampsia in twin pregnancies: a population-based matched case-control study.** *J Perinat Med* 2012, **40**:379-82.
28. Buurma AJ, Turner RJ, Driessen JH, Mooyaart AL, Schoones JW, Bruijn JA, Bloemenkamp KW, Dekkers OM, Baelde HJ: **Genetic variants in pre-eclampsia: a meta-analysis.** *Hum Reprod Update* 2013, **19**:289-303.
29. Dostálová Z, Bienertová-Vasků AJ, Vasků A, Gerychová R, Unzeitig V: **[Insertion-deletion polymorphism in the gene for angiotensin-converning enzyme (I/D ACE) in pregnant women with gestational diabetes].** *Ceska Gynekol* 2006, **71**:369-73.
30. Yu Z-YY, Chen L-SS, Zhang L-CC, Zhou T-BB: **Meta-analysis of the relationship between ACE I/D gene polymorphism and end-stage renal disease in patients with diabetic nephropathy.** *Nephrology (Carlton)* 2012, **17**:480-7.
31. Mao S, Huang S: **A meta-analysis of the association between angiotensin-converting enzyme insertion/ deletion gene polymorphism and the risk of overweight/obesity.** *J Renin Angiotensin Aldosterone Syst* 2013.
32. Machado JS, Palei AC, Amaral LM, Bueno AC, Antonini SR, Duarte G, Tanus-Santos JE, Sandrim VC, Cavalli RC: **Polymorphisms of the adiponectin gene in gestational hypertension and pre-eclampsia.** *J Hum Hypertens* 2014, **28**:128-32.
33. Low CF, Mohd Tohit ER, Chong PP, Idris F: **Adiponectin SNP45TG is associated with gestational diabetes mellitus.** *Arch Gynecol Obstet* 2011, **283**:1255-60.
34. Chu H, Wang M, Zhong D, Shi D, Ma L, Tong N, Zhang Z: **AdipoQ polymorphisms are associated with type 2 diabetes mellitus: a meta-analysis study.** *Diabetes Metab Res Rev* 2013, **29**:532-45.
35. Wu J, Liu Z, Meng K, Zhang L: **Association of adiponectin gene (ADIPOQ) rs2241766 polymorphism with obesity in adults: a meta-analysis.** *PloS One* 2014, **9**:e95270.
36. Vianna P, Silva GK Da, Santos BP Dos, Bauer ME, Dalmáz CA, Bandinelli E, Chies JA: **Association between mannose-binding lectin gene polymorphisms and pre-eclampsia in Brazilian women.** *Am J Reprod Immunol (New York, NY: 1989)* 2010, **64**:359-74.
37. Megia A, Gallart L, Fernández-Real J-MM, Vendrell J, Simón I, Gutierrez C, Richart C: **Mannose-binding lectin gene polymorphisms are associated with gestational diabetes mellitus.** *J Clin Endocrinol Metab* 2004, **89**:5081-7.
38. Muller YL, Hanson RL, Bian L, Mack J, Shi X, Pakyz R, Shuldiner AR, Knowler WC, Bogardus C, Baier LJ: **Functional variants in MBL2 are associated with type 2 diabetes and pre-diabetes traits in Pima Indians and the old order Amish.** *Diabetes* 2010, **59**:2080-5.
39. Fernández-Real JM, Straczkowski M, Vendrell J, Soriguer F, Pérez Del Pulgar S, Gallart L, López-Bermejo A, Kowalska I, Manco M, Cardona F, García-Gil MM, Mingrone G, Richart C, Ricart W, Zorzano A: **Protection from inflammatory disease in insulin resistance: the role of mannan-binding lectin.** *Diabetologia* 2006, **49**:2402-11.

40. Harmon QE, Engel SM, Wu MC, Moran TM, Luo J, Stuebe AM, Avery CL, Olshan AF: **Polymorphisms in inflammatory genes are associated with term small for gestational age and pre-eclampsia.** *Am J Reprod Immunol* 2014, **71**:472-84.
41. Chang Y, Niu XM, Qi XM, Zhang HY, Li NJ, Luo Y: **[Study on the association between gestational diabetes mellitus and tumor necrosis factor-alpha gene polymorphism].** *Zhonghua Fu Chan Ke Za Zhi* 2005, **40**:676-8.
42. Sefri H, Benrahma H, Charoute H, Lakbakbi El Yaagoubi F, Rouba H, Lyoussi B, Nourlil J, Abidi O, Barakat A: **TNF A -308G>A polymorphism in Moroccan patients with type 2 diabetes mellitus: a case-control study and meta-analysis.** *Mol Biol Rep* 2014.
43. Sookoian SC, González C, Pirola CJ: **Meta-analysis on the G-308A tumor necrosis factor alpha gene variant and phenotypes associated with the metabolic syndrome.** *Obes Res* 2005, **13**:2122-31.
44. Ehsanipoor RM, Fortson W, Fitzmaurice LE, Liao W-XX, Wing DA, Chen D-BB, Chan K: **Nitric oxide and carbon monoxide production and metabolism in pre-eclampsia.** *Reprod Sci* 2013, **20**:542-8.
45. Lachmeijer AM, Nosti-Escanilla MP, Bastiaans EB, Pals G, Sandkuijl LA, Kostense PJ, Aarnoudse JG, Crusius JB, Peña AS, Dekker GA, Arngrímsson R, Kate LP ten: **Linkage and association studies of IL1B and IL1RN gene polymorphisms in pre-eclampsia.** *Hypertens Pregnancy* 2002, **21**:23-38.
46. Serrano NC, Casas JP, Díaz LA, Páez C, Mesa CM, Cifuentes R, Monterrosa A, Bautista A, Hawe E, Hingorani AD, Vallance P, López-Jaramillo P: **Endothelial NO synthase genotype and risk of pre-eclampsia: a multicenter case-control study.** *Hypertension* 2004, **44**:702-7.
47. Chen M, Yuan Z, Shan K: **Association of apolipoprotein J gene 866C–>T polymorphism with pre-eclampsia and essential hypertension.** *Gynecol Obstet Invest* 2005, **60**:133-8.
48. Fekete A, Vér A, Bögi K, Treszl A, Rigó J: **Is pre-eclampsia associated with higher frequency of HSP70 gene polymorphisms?** *Eur J Obstet Gynecol Reprod* 2006, **126**:197-200.
49. Mütze S, Rudnik-Schöneborn S, Zerres K, Rath W: **Genes and the pre-eclampsia syndrome.** *J Perinat Med* 2008, **36**:38-58.
50. Wang XM, Wu HY, Qiu XJ: **Methylenetetrahydrofolate reductase (MTHFR) gene C677T polymorphism and risk of pre-eclampsia: an updated meta-analysis based on 51 studies.** *Arch Med Res* 2013, **44**:159-68.
51. Vanderweele TJ: **Sensitivity analysis: distributional assumptions and confounding assumptions.** *Biometrics* 2008, **64**:645-9.
52. Vanderweele TJ, Mukherjee B, Chen J: **Sensitivity analysis for interactions under unmeasured confounding.** *Stat Med* 2012, **31**:2552-64.
53. Jaccard P: **The distribution of the flora in the alpine zone.** *New Phytol* 1912, **11**:37-50.
54. Goldberg DS, Roth FP: **Assessing experimentally derived interactions in a small world.** *Proc Natl Acad Sci USA* 2003, **100**:4372-4376.
55. Maere S, Heymans K, Kuiper M: **BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks.** *Bioinformatics* 2005, **21**:3448-9.

doi:10.1186/1752-0509-9-S2-S4
Cite this article as: Glotov *et al.*: Molecular association of pathogenetic contributors to pre-eclampsia (pre-eclampsia associome). *BMC Systems Biology* 2015 **9**(Suppl 2):S4.

eLIFE elifesciences.org

RESEARCH ARTICLE

Single cell transcriptional signatures of the human placenta in term and preterm parturition

Roger Pique-Regi[1,2,3]*, Roberto Romero[1,3,4,5,6]*, Adi L Tarca[2,3,7], Edward D Sendler[1], Yi Xu[2,3], Valeria Garcia-Flores[2,3], Yaozhu Leng[2,3], Francesca Luca[1,2], Sonia S Hassan[2,8], Nardhy Gomez-Lopez[2,3,9]*

[1]Center for Molecular Medicine and Genetics, Wayne State University, Detroit, United States; [2]Department of Obstetrics and Gynecology, Wayne State University School of Medicine, Detroit, United States; [3]Perinatology Research Branch, Division of Obstetrics and Maternal-Fetal Medicine, Division of Intramural Research, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, U.S. Department of Health and Human Services, Detroit, United States; [4]Department of Obstetrics and Gynecology, University of Michigan, Ann Arbor, United States; [5]Department of Epidemiology and Biostatistics, Michigan State University, East Lansing, United States; [6]Detroit Medical Center, Detroit, United States; [7]Department of Computer Science, College of Engineering, Wayne State University, Detroit, United States; [8]Department of Physiology, Wayne State University School of Medicine, Detroit, United States; [9]Department of Immunology, Microbiology, and Biochemistry, Wayne State University School of Medicine, Detroit, United States

***For correspondence:** rpique@wayne.edu (RP-R); prbchiefstaff@med.wayne.edu (RR); ngomezlo@med.wayne.edu (NG-L)

Competing interests: The authors declare that no competing interests exist.

Funding: See page 13

Received: 19 September 2019
Accepted: 12 December 2019
Published: 12 December 2019

Reviewing editor: Stephen Parker, University of Michigan, United States

Abstract More than 135 million births occur each year; yet, the molecular underpinnings of human parturition in gestational tissues, and in particular the placenta, are still poorly understood. The placenta is a complex heterogeneous organ including cells of both maternal and fetal origin, and insults that disrupt the maternal-fetal dialogue could result in adverse pregnancy outcomes such as preterm birth. There is limited knowledge of the cell type composition and transcriptional activity of the placenta and its compartments during physiologic and pathologic parturition. To fill this knowledge gap, we used scRNA-seq to profile the placental villous tree, basal plate, and chorioamniotic membranes of women with or without labor at term and those with preterm labor. Significant differences in cell type composition and transcriptional profiles were found among placental compartments and across study groups. For the first time, two cell types were identified: 1) lymphatic endothelial decidual cells in the chorioamniotic membranes, and 2) non-proliferative interstitial cytotrophoblasts in the placental villi. Maternal macrophages from the chorioamniotic membranes displayed the largest differences in gene expression (e.g. *NFKB1*) in both processes of labor; yet, specific gene expression changes were also detected in preterm labor. Importantly, several placental scRNA-seq transcriptional signatures were modulated with advancing gestation in the maternal circulation, and specific immune cell type signatures were increased with labor at term (NK-cell and activated T-cell signatures) and with preterm labor (macrophage, monocyte, and activated T-cell signatures). Herein, we provide a catalogue of cell types and transcriptional profiles in the human placenta, shedding light on the molecular underpinnings and non-invasive prediction of the physiologic and pathologic parturition.

Introduction

Parturition is essential for the reproductive success of viviparous species (**Romero et al., 2006a**); yet, the mechanisms responsible for the onset of labor remain to be elucidated (**Norwitz et al., 1999**; **Norwitz et al., 2015**). Understanding human parturition is essential to tackle the challenge of prematurity, which affects 15 million neonates every year (**Muglia and Katz, 2010**; **Blencowe et al., 2012**; **Romero et al., 2014a**). Bulk transcriptomic studies of the cervix (**Hassan et al., 2006**; **Hassan et al., 2007**; **Hassan et al., 2009**; **Bollopragada et al., 2009**; **Dobyns et al., 2015**), myometrium (**Charpigny et al., 2003**; **Romero et al., 2014b**; **Mittal et al., 2010**; **Mittal et al., 2011**; **Chan et al., 2014**; **Stanfield et al., 2019**), and chorioamniotic membranes (**Haddad et al., 2006**; **Mittal et al., 2009**; **Nhan-Chang et al., 2010**) revealed that labor is a state of physiological inflammation; however, finding specific pathways implicated in preterm labor still remains an elusive goal. A possible explanation is that gestational tissues, and especially the placenta, are heterogeneous composites of multiple cell types, and elucidating perturbations in the maternal-fetal dialogue requires dissection of the transcriptional activity at the cell type level, which is not possible using bulk analyses. Recent microfluidic and droplet-based technological advances have enabled characterization of gene expression at single-cell resolution (scRNA-seq) (**Klein et al., 2015**; **Macosko et al., 2015**). Previous work in humans (**Tsang et al., 2017**; **Pavličev et al., 2017**; **Vento-Tormo et al., 2018**) and mice (**Nelson et al., 2016**) demonstrated that scRNA-seq can capture the multiple cell types that constitute the placenta and identify their maternal or fetal origin. Such studies showed that single-cell technology can be used to infer communication networks across the different cell types at the maternal-fetal interface (**Vento-Tormo et al., 2018**). Further, the single-cell-derived placental signatures were detected in the cell-free RNA present in maternal circulation (**Tsang et al., 2017**), suggesting that non-invasive identification of women with early-onset preeclampsia is feasible. However, these studies included a limited number of samples and did not account for the fact that different pathologies can arise from dysfunction in different placental compartments. In addition, the physiologic and pathologic processes of labor have never been studied at single-cell resolution.

Results and discussion

In this study, a total of 25 scRNA-seq libraries were prepared from three placental compartments: basal plate (BP), placental villous (PV), and chorioamniotic membranes (CAM) (**Figure 1A**). These were collected from nine women in the following study groups: term no labor (TNL), term in labor (TIL), and preterm labor (PTL). scRNA-seq libraries were prepared with the 10X Chromium system and were processed using the 10X Cell Ranger software, resulting in 79,906 cells being captured and profiled across all samples (**Supplementary file 1**). We used Seurat (**Butler et al., 2018**) to normalize expression profiles and identified 19 distinct clusters, which were assigned to cell types based on the expression of previously reported marker genes (**Tsang et al., 2017**; **Pavličev et al., 2017**; **Vento-Tormo et al., 2018**) (see Materials and methods, **Figure 1—figure supplement 1** and **Supplementary file 2–3**). The uniform manifold approximation and projection (UMAP **Becht et al., 2019**) was used to display these clusters in two dimensions (**Figure 1B**). With this approach, the local and global topological structure of the clusters is preserved, with subtypes of the major cell lineages (trophoblast, lymphoid, myeloid, stromal, and endothelial sub-clusters) being displayed proximal to each other. The trophoblast lineage reconstruction displayed in **Figure 1—figure supplement 2** shows the progression from cytotrophoblasts to either extravillous trophoblasts or syncytiotrophoblasts, which recapitulates the differentiation structure previously reported (**Tsang et al., 2017**; **Vento-Tormo et al., 2018**).

The cell type composition differed both among placental compartments (**Figure 1C**) and due to the presence of physiologic and pathologic processes of labor (i.e. term in labor and preterm labor) (**Figure 1D**). While extravillous trophoblasts (EVT) were present in all three compartments, cytotrophoblasts (CTB) were especially pervasive in the placental villi, which is explained by the fact that CTBs are abundant in the parenchyma of the placentas. CTBs were also present in the basal plate since this placental compartment is adjacent to the placental villi (**Figure 1A**). The phenotypic similarities between trophoblasts in proximity to the decidua parietalis (layer attached to the chorioamniotic membranes) and those found in the basal plate have been previously documented

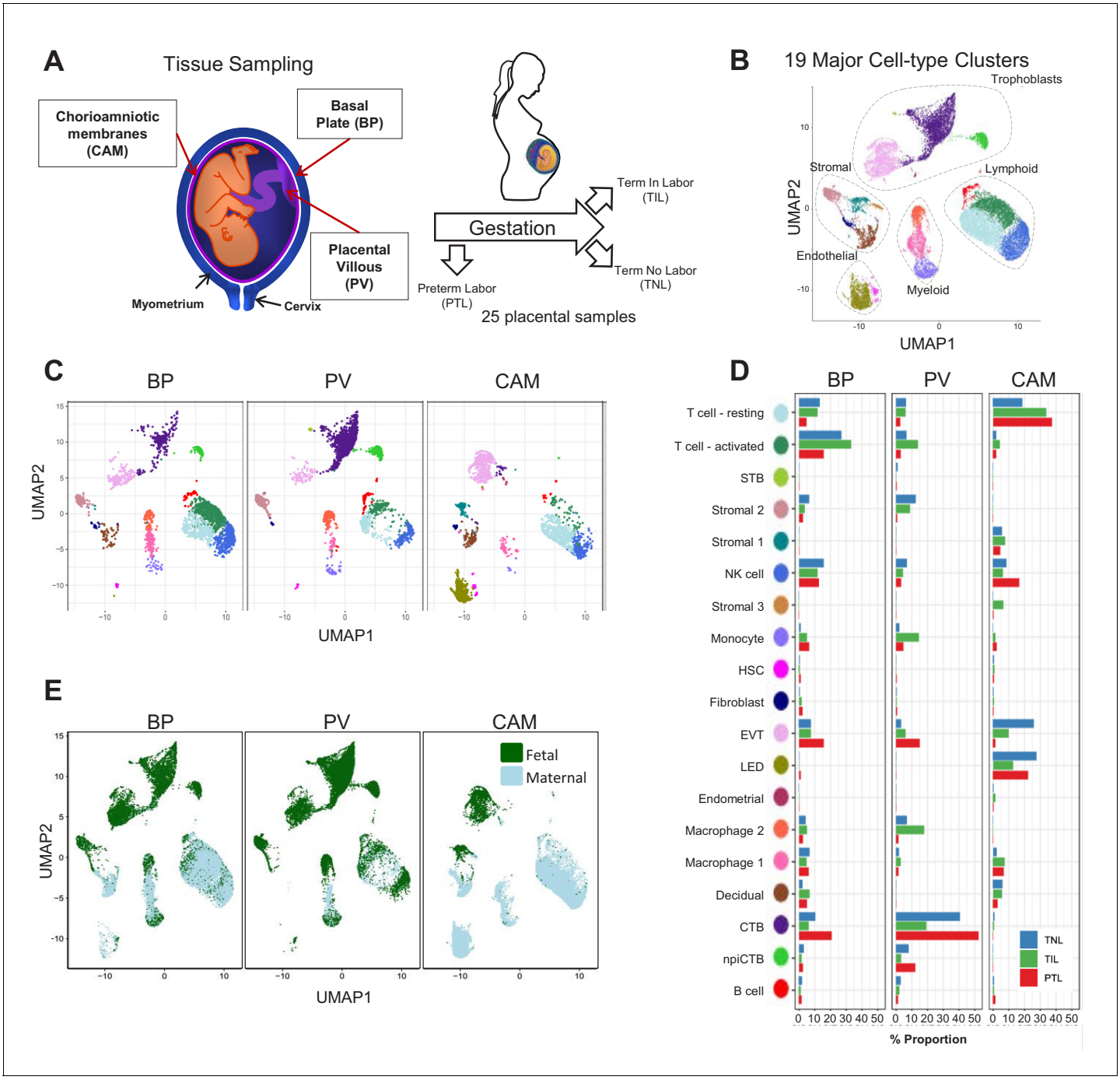

Figure 1. Transcriptional map of the placenta in human parturition. (**A**) Study design illustrating the placental compartments and study groups. (**B**) Uniform Manifold Approximation Plot (UMAP), where dots represent single cells and are colored by cell type. (**C**) Distribution of single-cell clusters by placental compartments. (**D**) Average proportions of cell types by placental compartments and study groups. (**E**) Distribution of single cells by maternal or fetal origin. STB, Syncytiotrophoblast; EVT, Extravillous trophoblast; CTB, cytotrophoblast; HSC, hematopoietic stem cell; npiCTB, non proliferative interstitial cytotrophoblast; LED, lymphoid endothelial decidual cell.

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. Heatmap of the top gene expression markers defining each cell-type.

Figure supplement 2. UMAP plot highlighting the trophoblast cell-types and their inferred differentiation path using slingshot R package.

Figure supplement 3. Single marker gene expression UMAP plot for genes differentially expressed between CTB and npiCTB.

Figure supplement 4. Analysis of the fetal/maternal origin of the cell-types based on data from three pregnancies with a male fetus.

Figure supplement 5. Alluvial diagram showing the correspondence between our final curated cluster labels and automated cell-labeling methods.

Figure 1 continued on next page

Figure 1 continued

Figure supplement 6. Heatmap showing the correspondence between our final curated cluster labels and automated cell-labeling methods.
Figure supplement 7. Uniform Manifold Approximation Plot (UMAP), where dots representing single cells and color represents Seurat predicted cell type labels.
Figure supplement 8. Doublet analysis by DoubletFinder.

(***Genbačev et al., 2015***; ***Garrido-Gomez et al., 2017***). Importantly, non-proliferative interstitial cytotrophoblasts (npiCTB) were identified for the first time in the placental villi as forming a distinct cluster. This new cluster was also observed in the basal plate, but not in the chorioamniotic membranes, suggesting that this type of trophoblast has specific functions in the placental tree. Lineage reconstruction by slingshot (***Street et al., 2018***) revealed that npiCTBs are likely derived from conventional CTBs (***Figure 1—figure supplement 2***). The non-proliferative nature of npiCTBs was evidenced by the reduced expression of genes involved in cell proliferation such as *XIST*, *DDX3X*, and *EIF1AX* (***Figure 1—figure supplement 3***). npiCTBs displayed an increased expression of *PAGE4* (***Figure 1—figure supplement 3***), a gene expressed by CTBs isolated from pregnancy terminations (***Genbacev et al., 2011***), suggesting that this type of trophoblast cell is present earlier in gestation. As expected, trophoblast cell types were of fetal origin, and decidual cells present in the basal plate (including the decidua basalis) and chorioamniotic membranes (including the decidua parietalis) were of maternal origin (***Figure 1E*** and ***Figure 1—figure supplement 4***).

In terms of immune cell types, the chorioamniotic membranes largely contained lymphoid and myeloid cells of maternal origin, including T cells (mostly in a resting state), NK cells, and macrophages (***Figure 1C and E*** and ***Figure 1—figure supplement 4***). In contrast, the basal plate included immune cells of both maternal and fetal origin, such as T cells (mostly in an active state), NK cells, and macrophages. The placental villi contained more fetal than maternal immune cells, namely monocytes, macrophages, T cells, and NK cells. Two macrophage subsets were found in placenta compartments: macrophage 1 of maternal origin that was predominant in the chorioamniotic membranes, and macrophage 2 of fetal origin that was mainly present in the basal plate and placental villi. Together with previous single cell studies of early pregnancy (***Vento-Tormo et al., 2018***), these results highlight the complexity and dynamics of the immune cellular composition of the placental tissues, including the maternal-fetal interface (i.e. decidua), from early gestation to term or preterm delivery.

Importantly, a new lymphatic endothelial decidual (LED) cell type of maternal origin was identified in the chorioamniotic membranes, forming a distinct transcriptional cluster that was separate from other endothelial cell-types (***Figure 1C and E***). LED cells were rarely observed in the basal plate and were completely absent in the placental villous. Similar to other endothelial cell types, LED cells highly expressed *CD34*, *CDH5*, *EDNRB*, *PDPN*, and *TIE1* (***Figure 2—figure supplement 1***). The signature genes of this novel cell type were enriched for pathways involving cell-cell and cell-surface interactions at the vascular wall, extracellular matrix organization (***Figure 2—figure supplement 2***), tight junction, and focal adhesion (***Figure 2—figure supplement 3***), indicating that LEDs possess the machinery required to mediate the influx of immune cells into the chorioamniotic membranes. Immunostaining confirmed the co-expression of LYVE1 (lymphatic marker) and CD31 (endothelial molecule marker) in the vessels of the decidua parietalis of the chorioamniotic membranes, but not in the basal plate or placenta (***Figure 2A***). The co-localization of LYVE1 and CD31 proteins (i.e. LED cells) in the chorioamniotic membranes is shown in ***Figure 2B*** and ***Figure 2—video 1***. LED cells also expressed the common endothelial cell marker *CD34* (***Figure 2C***, green arrow). *LYVE1* was also expressed by the fetal macrophages present in the placental villi and basal plate (***Figure 2C***, red arrow), yet the protein was only visualized by immunostaining in immune cells located in the villous tree (***Figure 2A***, red arrows). This finding conclusively shows the presence of lymphatic vessels in the decidua parietalis of the chorioamniotic membranes, providing a major route for maternal lymphocytes (e.g. T cells) infiltrating the maternal-fetal interface (***Arenas-Hernandez et al., 2019***).

For cell types that were present in more than one placental compartment, major differences in gene expression were identified across locations, indicative of further specialization of cells depending on the unique physiological functions of each microenvironment (***Figure 3—figure supplement 1*** and ***Supplementary file 4***). Differences in the transcriptional profiles were particularly large for

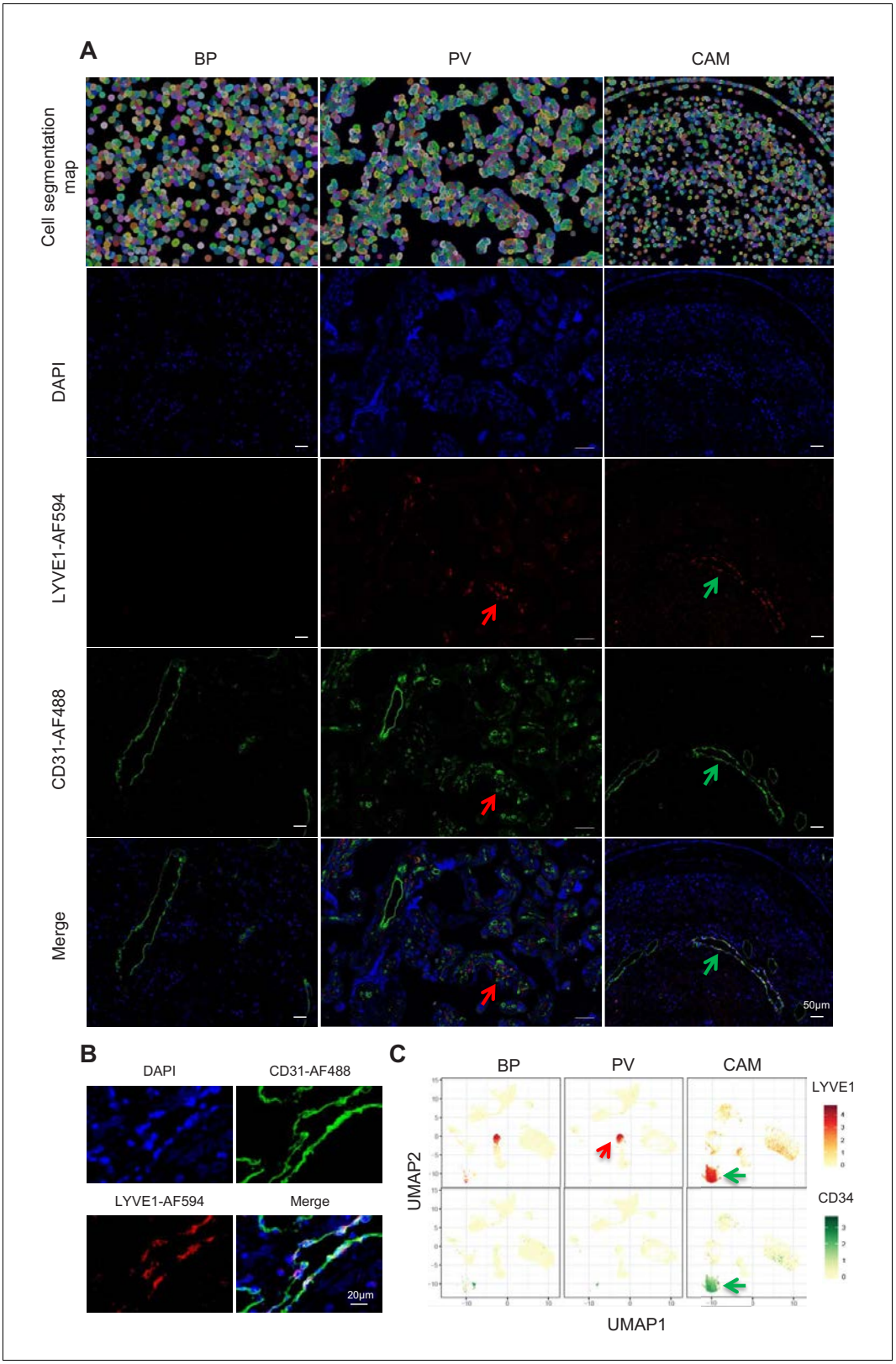

Figure 2. Identification of LED cells in the chorioamniotic membranes. (**A**) Cell segmentation map (built using the DAPI nuclear staining) and immunofluorescence detection of LYVE-1 (red) and CD31 (green) in the basal plate (BP), placental villi (PV), and chorioamniotic membranes (CAM). Red arrows point to fetal macrophages expressing LYVE1 but not CD31, and green arrows indicate lymphatic endothelial decidual cells (LED cells)

Figure 2 continued on next page

Figure 2 continued

expressing both LYVE1 and CD31. (**B**) Co-expression of LYVE1 and CD31 (i.e. LED cells) in the chorioamniotic membranes. (**C**) Single-cell expression UMAP of LYVE-1 (red) and CD34 (green) in the placental compartments.

The online version of this article includes the following video and figure supplement(s) for figure 2:

Figure supplement 1. Single marker gene expression UMAP plot for genes that are more highly expressed in lymphatic endothelial decidual (LED) cells.

Figure supplement 2. Clusterprofiler dot plot showing the ReactomeDB Pathways enriched for genes that define each cell-type.

Figure supplement 3. Clusterprofiler dot plot showing the Kegg Pathways enriched for genes that define each cell-type.

Figure 2—video 1. Video with the 3D reconstruction of the lymphatic endothelium in the decidua present in the CAM compartment.
https://elifesciences.org/articles/52004#fig2video1

maternal macrophages as well as EVTs, NK cells, and T cells in the chorioamniotic membranes compared to the other compartments. Genes differentially expressed in the chorioamniotic membranes were enriched for interleukin and Toll-like receptor signaling as well as for the NF-κB and TNF pathways (***Figure 3—figure supplements 2–4***). These results are consistent with previous reports showing a role for these mediators in the inflammatory process of labor (***Romero et al., 1989a***; ***Romero et al., 1990b***; ***Romero et al., 1992a***; ***Romero et al., 1990a***; ***Santhanam et al., 1991***; ***Romero et al., 1993***; ***Romero et al., 1991***; ***Hsu et al., 1998***; ***Keelan et al., 1999***; ***Young et al., 2002***; ***Osman et al., 2003***; ***Kim et al., 2004***; ***Abrahams et al., 2004***; ***Kumazaki et al., 2004***; ***Koga et al., 2009***; ***Belt et al., 1999***; ***Yan et al., 2002***; ***Lindström and Bennett, 2005***; ***Vora et al., 2010***; ***Romero, 1989b***; ***Romero et al., 1992b***; ***Lonergan et al., 2003***). Conversely, the placental villous and basal plate were more similar to each other, with most differentially expressed genes (DEG) between these compartments being noted in fibroblasts (335 DEG, $q < 0.1$ and fold change >2) (***Figure 3—figure supplements 1*** and ***5–10***). DEGs in the placental villous fibroblasts showed enrichment in smooth muscle contraction, the apelin and oxytocin signaling pathways (***Figure 3—figure supplement 9***), while DEGs in CAM fibroblasts were enriched in elastic fiber formation and extracellular matrix pathways (***Figure 3—figure supplement 2***). The latter finding indicates that the same cell type (e.g. fibroblasts) may have distinct functions in different microenvironments of the placenta.

Next, we assessed changes due to term and preterm labor in each cell type (***Supplementary file 5***). The largest number of DEGs between the term labor and term no labor groups were observed in the maternal macrophages (macrophage 1), followed by the EVT (144 and 37, respectively, $q < 0.1$; ***Figure 3A***). The largest number of DEGs between the preterm labor and term labor groups were observed in EVT and CTB (37 and 33, respectively, $q < 0.1$; ***Figure 3A***). ***Figure 3B*** displays the gene expression changes between TIL and TNL or PTL and TNL that are shared between the two labor groups, representing the common pathway of parturition (defined as the anatomical, physiological, biochemical, endocrinological, immunological, and clinical events that occur in the mother and/or fetus in both term and preterm labor ***Romero et al., 2006b***). Non-shared differences in gene expression with labor at term and in preterm labor were mostly observed in trophoblast cell types such as CTB and EVT as well as in stromal cells (***Figure 3C***). Some of these changes may be explained by the unavoidable confounding effect of gestational age since placentas from women without labor in preterm gestation cannot be obtained in the absence of pregnancy complications. Specifically, the expression of *NFKB1* by maternal macrophages was higher in women with term labor compared to non-labor controls, and this increase was further accentuated in preterm labor (***Figure 3D***). Consistent with the induction of the NFκB pathway, the labor-associated DEGs in macrophages involved biological processes such as activation of immune response and regulation of pro-inflammatory cytokine production (***Figure 3—figure supplement 11A***). These results are in line with previous studies showing that decidual macrophages undergo an M1-like macrophage polarization (i.e. pro-inflammatory phenotype) during term and preterm labor (***Xu et al., 2016***).

When comparing the effect sizes between the PTL/TNL and TIL/TNL juxtapositions on the same gene and cell type, positive correlations were observed for most of the placental cell types (***Figure 3E***). Genes displaying differential effects in term and preterm labor are mostly found in trophoblast cell types (see off-diagonal points in the scatter plot), which may be explained by the phenomenon of gene expression decoherence (***Lea et al., 2019***). This lack of proper correlation between biomarkers to their expected normal relationships is commonly found in pathological

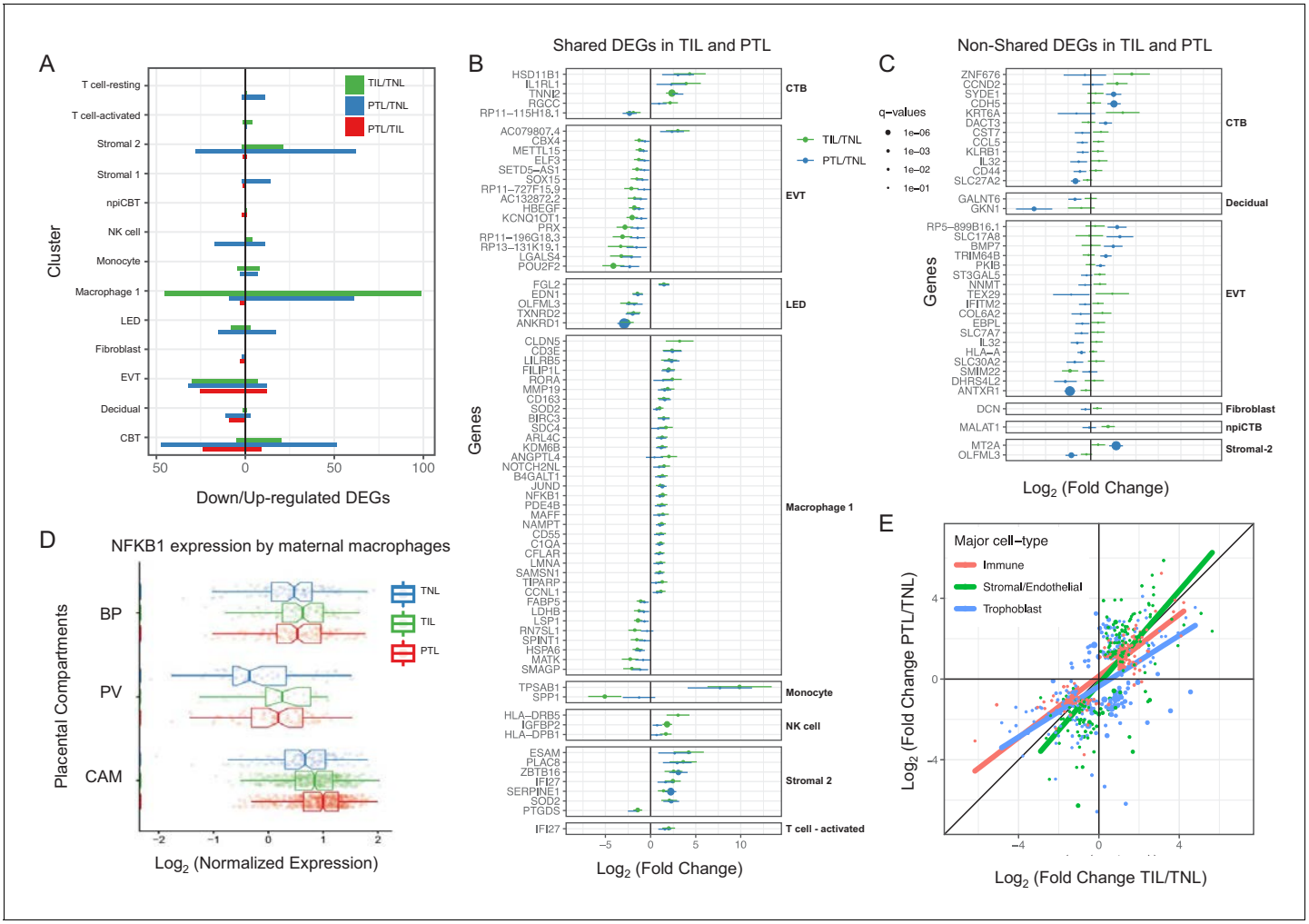

Figure 3. Cell type specific expression changes in term and preterm labor. (**A**) Number of differentially expressed genes (DEGs) among study groups (TNL, term no labor; TIL, term in labor; PTL, preterm labor) by direction of change. Shared (**B**) and non-shared (**C**) expression changes in term labor and preterm labor relative to the term no labor group ($q < 0.01$). The length of each whisker represents the 95% confidence interval. (**D**) The expression of *NFKB1* by maternal macrophages in the placental compartments (BP, basal plate; PV, placental villous; CAM, chorioamniotic membranes) and study groups. The notch represents the 95% confidence interval of the median. (**E**) Differences and similarities in expression changes with preterm labor and term labor by three major cell types (immune, stromal/endothelial, and trophoblast cells).

The online version of this article includes the following figure supplement(s) for figure 3:

Figure supplement 1. Stacked bar plot summarizing differentially expressed genes across compartments for a cell types that are present on all three of them.

Figure supplement 2. Clusterprofiler dot plot showing the ReactomeDB Pathways enriched for genes that are significantly more highly expressed in the CAM compartment relative to the other compartments for each cell-type.

Figure supplement 3. Clusterprofiler dot plot showing the Kegg Pathways enriched for genes that are significantly more highly expressed in the CAM compartment relative to the other compartments for each cell-type.

Figure supplement 4. Clusterprofiler dot plot showing gene ontology (GO) terms enriched for genes that are significantly more highly expressed in the CAM compartment relative to the other compartments for each cell-type.

Figure supplement 5. Clusterprofiler dot plot showing the ReactomeDB Pathways enriched for genes that are significantly more highly expressed in the BP compartment relative to the other compartments for each cell-type.

Figure supplement 6. Clusterprofiler dot plot showing the Kegg Pathways enriched for genes that are significantly more highly expressed in the BP compartment relative to the other compartments for each cell-type.

Figure supplement 7. Clusterprofiler dot plot showing gene ontology (GO) terms enriched for genes that are significantly more highly expressed in the BP compartment relative to the other compartments for each cell-type.

Figure supplement 8. Clusterprofiler dot plot showing the ReactomeDB Pathways enriched for genes that are significantly more highly expressed in the PV compartment relative to the other compartments for each cell-type.

Figure 3 continued on next page

Figure 3 continued

Figure supplement 9. Clusterprofiler dot plot showing the Kegg Pathways enriched for genes that are significantly more highly expressed in the PV compartment relative to the other compartments for each cell-type.

Figure supplement 10. Clusterprofiler dot plot showing gene ontology (GO) terms enriched for genes that are significantly more highly expressed in the PV compartment relative to the other compartments for each cell-type.

Figure supplement 11. Clusterprofiler dot plot showing ReactomeDB pathways enriched using gene set enrichment analysis (GSEA) for genes differentially expressed in term labor compared to term no labor condition.

conditions. Lastly, in EVT the DEGs with labor were enriched for genes implicated in cellular response to stress, including the WNT and NOTCH pathways, as well as cell cycle checkpoints (***Figure 3—figure supplement 11B***), further supporting the hypothesis that the cellular senescence pathway (i.e. cell cycle arrest) is implicated in the physiologic (***Behnia et al., 2015***; ***Polettini et al., 2015***) and pathologic (***Hirota et al., 2010***; ***Gomez-Lopez et al., 2017***) processes of labor.

To demonstrate the translational value of single-cell RNA signatures derived from the placenta, we conducted an in silico analysis in public datasets (***Tarca et al., 2019***; ***Paquette et al., 2018***) to test whether the single-cell signatures could be non-invasively monitored in the maternal circulation throughout gestation (***Figure 4A***). Previous studies have correlated bulk mRNA expression in the maternal circulation with gestational age at blood draw (***Tarca et al., 2019***; ***Al-Garawi et al., 2016***), risk for preterm birth (***Paquette et al., 2018***; ***Heng et al., 2014***; ***Sirota et al., 2018***; ***Knijnenburg et al., 2019***), or both (***Heng et al., 2016***; ***Ngo et al., 2018***). First, using whole blood bulk RNAseq data, we quantified the expression of single-cell signatures in the maternal circulation. We found that the expression of the single-cell signatures of macrophages, monocytes, NK cells, T cells, npiCTB, and fibroblasts is modulated with advancing gestational age (***Figure 4B–C***, ***Figure 4—figure supplement 1A***). These results validate the T-cell and monocyte signature changes with gestational age that were previously reported (***Tsang et al., 2017***; ***Tarca et al., 2019***); yet, here we show that novel placental single-cell signatures (e.g., npiCTB and fibroblast) can also be non-invasively monitored in maternal circulation (***Figure 4—figure supplement 1A***). In addition, for the first time, we report that the expression of the single-cell signatures of NK-cells and activated T-cells were upregulated in women with spontaneous labor at term compared to gestational-age matched controls without labor (***Figure 4D***). Furthermore, we found that the average expression of the single-cell signatures of macrophages, monocytes, activated T cells, and fibroblasts were increased in the circulation of women with preterm labor and delivery compared to gestational age-matched controls (24–34 weeks of gestation) (***Figure 4E*** and ***Figure 4—figure supplement 1B***). These findings are in line with previous reports indicating a role for these immune cell types in the pathophysiology of preterm labor (***Arenas-Hernandez et al., 2019***; ***Hamilton et al., 2012***; ***Shynlova et al., 2013***; ***Gomez-Lopez et al., 2016***).

Conclusion

In summary, this study provides evidence of differences in cell type composition and transcriptional profiles among the basal plate, placental villi, and chorioamniotic membranes, as well as between the pathologic and physiologic processes of labor at single-cell resolution. Using scRNAseq technology, two novel cell types were identified in the chorioamniotic membranes and placental villi. In addition, we showed that maternal macrophages and extravillous trophoblasts are the cell types with the most transcriptional changes during the process of labor. Importantly, many of the genes differentially expressed in these cell-types replicate for both conditions of labor. This result shows that we have enough statistical power to detect the changes in gene expression with a large effect size that are general or a common molecular pathway in parturition; yet, additional studies are needed to characterize the different etiologies of the preterm labor syndrome. Lastly, we report that maternal and fetal transcriptional signatures derived from placental scRNA-seq are modulated with advancing gestation and are markedly perturbed with term and preterm labor in the maternal circulation. These results highlight the potential of single-cell signatures as biomarkers to non-invasively monitor the cellular dynamics during pregnancy and to predict obstetrical disease. The current study represents the most comprehensive single-cell analysis of the human placental transcriptome in physiologic and pathologic parturition.

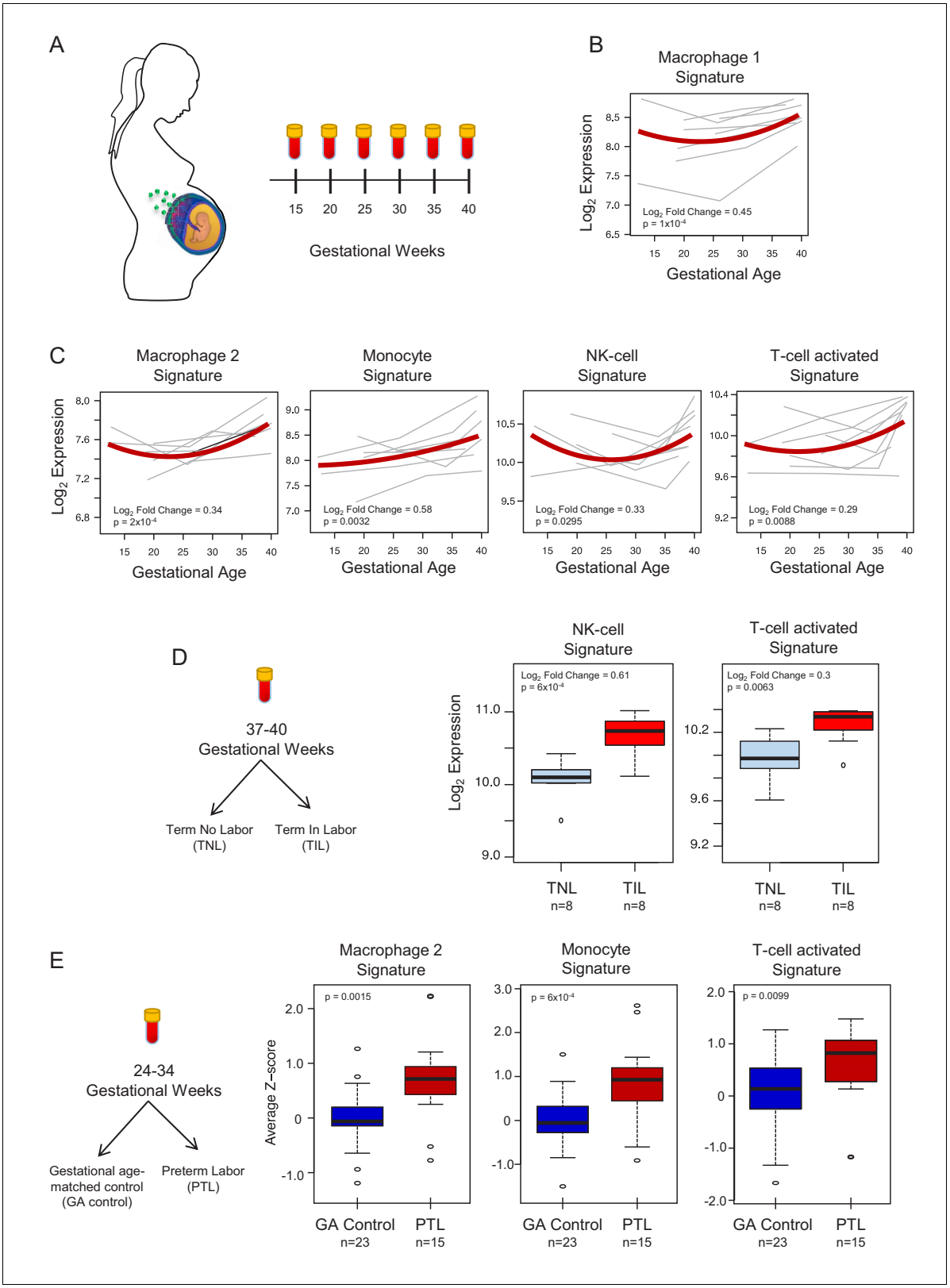

Figure 4. In silico analysis to quantify scRNA-seq signatures in the maternal circulation. (**A**) Diagram of the longitudinal study used to generate bulk RNAseq data (GSE114037) (***Tarca et al., 2019***) to evaluate changes in scRNA-seq signatures with advancing gestation. Whole blood samples were collected throughout gestation from women who delivered at term. (**B and C**) Variation of scRNA-seq signature expression in the maternal circulation with advancing gestation. (**D**) Diagram of the cross-sectional study used to generate bulk RNAseq data (GSE114037) to evaluate changes in scRNA-seq

Figure 4 continued on next page

Figure 4 continued

signatures with labor at term (***Tarca et al., 2019***). Differences in the expression of scRNA-seq signatures between women with spontaneous labor at term (TIL) and term no labor controls (TNL). (**E**) Diagram of the cross-sectional study used to generate bulk RNAseq data (GSE96083) to evaluate changes in scRNA-seq signatures in preterm labor (***Paquette et al., 2018***). Differences in the expression of scRNA-seq signatures between women with spontaneous preterm labor (PTL) and gestational-age matched controls (GA control).
The online version of this article includes the following figure supplement(s) for figure 4:

Figure supplement 1. Quantification of scRNA-seq signatures in maternal circulation (continued from main ***Figure 4***).

Materials and methods

Sample collection and processing, single-cell preparation, library preparation, and sequencing

Human subjects

Immediately after delivery, placental samples [the villi, basal plate (including the decidua basalis) and chorioamniotic membranes (including the decidua parietalis)] were collected from women with or without labor at term or preterm labor at the Detroit Medical Center, Wayne State University School of Medicine (Detroit, MI). Labor was defined by the presence of regular uterine contractions at a frequency of at least two contractions every 10 min with cervical changes resulting in delivery. Women with preterm labor delivered between 33–35 weeks of gestation whereas those with term labor delivered between 38–40 weeks of gestation (***Supplementary file 6***). The collection and use of human materials for research purposes were approved by the Institutional Review Boards of the Wayne State University School of Medicine. All participating women provided written informed consent prior to sample collection.

Single-cell preparation

Cells from placental villi, basal plate, and chorioamniotic membranes were isolated by enzymatic digestion, using previously described protocols with modifications (***Tsang et al., 2017***; ***Xu et al., 2015***). Briefly, placental tissues were homogenized using a gentleMACS Dissociator (Miltenyi Biotec, San Diego, CA) either in an enzyme cocktail from the Umbilical Cord Dissociation Kit (Miltenyi Biotec) or in collagenase A (Sigma Aldrich, St. Louis, MO). After digestion, homogenized tissues were washed with ice-cold 1X phosphate-buffered saline (PBS) and filtered through a cell strainer (Fisher Scientific, Durham, NC). Cell suspensions were then collected and centrifuged at 300 x g for 5 min. at 4˚C. Red blood cells were lysed using a lysing buffer (Life Technologies, Grand Island, NY). Next, cells were washed with ice-cold 1X PBS and resuspended in 1X PBS for cell counting, which was performed using an automatic cell counter (Cellometer Auto 2000; Nexcelom Bioscience, Lawrence, MA). Lastly, dead cells were removed from the cell suspensions using the Dead Cell Removal Kit (Miltenyi Biotec) and cells were counted again using an automatic cell counter.

Single-cell preparation using the 10x genomics platform

Viable cells were used for single-cell RNAseq library construction using the Chromium Controller and Chromium Single Cell 3' version two kit (10x Genomics, Pleasanton, CA), following the manufacturer's instructions. Briefly, viable cell suspensions were loaded into the Chromium Controller to generate gel beads in emulsion (GEM) with each GEM containing a single cell as well as barcoded oligonucleotides. Next, the GEMs were placed in the Veriti 96-well Thermal Cycler (Thermo Fisher Scentific, Wilmington, DE) and reverse transcription was performed in each GEM (GEM-RT). After the reaction, the complementary cDNA was cleaned using Silane DynaBeads (Thermo Fisher Scentific) and the SPRIselect Reagent kit (Beckman Coulter, Indianapolis, IN). Next, the cDNAs were amplified using the Veriti 96-well Thermal Cycler and cleaned using the SPRIselect Reagent kit. Indexed sequencing libraries were then constructed using the Chromium Single Cell 3' version two kit, following the manufacturer's instructions.

Library preparation

cDNA was fragmented, end-repaired, and A-tailed using the Chromium Single Cell 3' version two kit, following the manufacturer's instructions. Next, adaptor ligation was performed using the

Chromium Single Cell 3' version two kit followed by post-ligation cleanup using the SPRIselect Reagent kit to obtain the final library constructs, which were then amplified using PCR. After performing a post-sample index double-sided size selection using the SPRIselect Reagent kit, the quality and quantity of the DNA were analyzed using the Agilent Bioanalyzer High Sensitivity chip (Agilent Technologies, Wilmington, DE). The Kapa DNA Quantification Kit for Illumina platforms (Kapa Biosystems, Wilmington, MA) was used to quantify the DNA libraries, following the manufacturer's instructions.

Sequencing

Sequencing of the single-cell libraries was performed by NovoGene (Sacramento, CA) using the Illumina Platform (HiSeq X Ten System).

Immunofluorescence

Samples of the chorioamniotic membranes, placenta villi, and decidua basal plate were embedded in Tissue-Tek Optimum Cutting Temperature (OCT) compound (Miles, Elkhart, IN) and snap-frozen in liquid nitrogen. Ten-µm-thick sections of each OCT-embedded tissue were cut using the Leica CM1950 (Leica Biosystems, Buffalo Grove, IL). Frozen slides were thawed to room temperature, fixed with 4% paraformaldehyde (Electron Microscopy Sciences, Hatfield, PA), and washed with 1X PBS. Non-specific background signals were blocked using Image-iT FX Signal Enhancer (Life Technologies) followed by blocking with antibody diluent/blocker (Perkin Elmer, Waltham, MA) for 30 min. at room temperature. Slides were then incubated with the rabbit anti-LYVE-1 antibody (Novus Biologicals, Centennial, CO) and the Flex mouse anti-human CD31 antibody (clone JC70A, Dako North America, Carpinteria, CA) for 90 min. at room temperature. Following washing with 1X PBS and blocking with 10% goat serum (SeraCare, Milford, MA), the slides were incubated with secondary goat anti-rabbit IgG–Alexa Fluor 594 (Life Technologies) and goat anti-mouse IgG–Alexa Fluor 488 (Life Technologies) for 30 min. at room temperature. Finally, the slides were washed and coverslips were mounted using ProLong Gold Antifade Mountant with DAPI (Life Technologies). Immunofluorescence was visualized using a confocal fluorescence microscope (Zeiss LSM 780; Carl Zeiss Microscopy GmbH, Jena, Germany) at the Microscopy, Imaging, and Cytometry Resources Core at the Wayne State University School of Medicine. Tile scans were performed from the chorioamniotic membranes, placental villi, and basal plate and the complete imaging fields were divided into six-by-six quadrants.

scRNA-seq data analyses

Raw fastq files obtained from Novogene were processed using Cell Ranger version 2.1.1 from 10X Genomics. First, sequence reads for each library (sample) were aligned to the hg19 reference genome using the STAR (***Dobin et al., 2013***) aligner, and expression of gene transcripts documented in the ENSEMBL database (Build 82) were determined for each cell. Gene expression was determined by the number of unique molecular identifiers (UMI) observed per gene (QC metrics are shown in ***Supplementary file 7***). Second, data were aggregated and down-sampled to take into account differences in sequencing depth across libraries using Cell Ranger Aggregate to obtain gene by cell expression data. Third, Seurat (***Butler et al., 2018***) was used to further clean and normalize the data. Then, only data from cells with a minimum of 200 detected genes, and from genes expressed in at least 10 cells were retained. Cells expressing mitochondrial genes at a level of >10% of total gene counts were also excluded, resulting in 77,906 cells and 25,803 genes (summary in ***Supplementary file 1***). Gene read counts were normalized with the Seurat 'NormalizeData' function (normalization.method = LogNormalize, scale.factor = 10,000). Genes showing significant variation across cells were selected based on 'LogVMR' dispersion function and 'FindVariableGenes'. Ribosomal and mitochondrial genes were next removed, yielding 3147 highly variable genes which were subsequently analyzed using Seurat 'RunPCA' function to obtain the first 20 principal components. Clustering was done using Seurat 'FindClusters' function based on the 20 PCAs (resolution of 0.7). Visualization of the cells was performed using Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP) algorithm as implemented by the Seurat 'runUMAP' function and using the first 20 principal components.

Assigning cell type labels to single-cell clusters (Appendix 1)

Multiple methods were utilized to label the cell clusters identified by Seurat. First, marker genes showing distinct expression in individual cell clusters compared to all others were identified using the Seurat FindAllMarkers function with default parameters (***Supplementary file 3***). Marker genes with significant specificity to each cluster (***Figure 1—figure supplement 1*** and ***Supplementary file 3***) were compared to those reported elsewhere (***Tsang et al., 2017***; ***Pavličev et al., 2017***). We also used previous known markers used by our group and https://www.proteinatlas.org/ to manually curate the labels. Further, the xCell (http://xcell.ucsf.edu/#) (***Aran et al., 2017***) tool was utilized to compare the pseudo-bulk expression signatures of the initial clusters to those of known cell types.

Additionally, we compared our manually curated cluster cell type labels to those derived from two automated cell labeling methods: SingleR (***Aran et al., 2019***) and Seurat (***Stuart et al., 2019***), using a human cell atlas reference and the placenta single-cell data in early pregnancy (***Vento-Tormo et al., 2018***) (see Appendix 1 for more details, ***Figure 1—figure supplements 5–7***). Finally, we used the R package DoubletFinder (***McGinnis et al., 2019***) (https://github.com/chris-mcginnis-ucsf/DoubletFinder) to identify potential doublets. None of our clusters were impacted by doublets (***Figure 1—figure supplement 8***).

Identification of cell-type maternal/fetal origin

We used two complementary approaches to determine the maternal/fetal origin of each cell-type. First, we used the samples derived from pregnancies where the neonate was male (3/9 cases, 8/25 samples) and we derived a fetal index based on the sum of all the reads mapping to genes on the Y chromosome relative to the total number of reads mapping to genes on the autosomes (***Figure 1—figure supplement 4***). The second method was based on genotype information derived from the scRNA-seq reads that overlap to known genetic variants from the 1000 Genomes reference panel using the freemuxlet approach implemented in popscle (***Figure 1E***). The freemuxlet approach extends the demuxlet (***Kang et al., 2018***) method, which can be useful for cases in which separate genotype information for each individual is not available. The software available at https://github.com/statgen/popscle/ was used with the '–nsample 2' option to map each cell barcode to one of the two possible genomes: fetal or maternal. The trophoblast cells are of fetal origin; therefore, we used this information to determine the fetal genome.

Trophoblast trajectory analysis

We used the slingshot R package (***Street et al., 2018***) to reconstruct the trophoblast cell lineages from our single-cell gene expression data. This method works by building a minimum spanning tree across clusters of cells and has been reviewed as one of the most accurate tools for this task (***Saelens et al., 2019***). This analysis focused on the trophoblast cell-types (STB, CTB, EVT, and npiCTB), in which we used as input the unmerged cluster labels (i.e., four sub-clusters for CTB, and two for EVT) and the matrix of cell embedding in UMAP (see ***Figure 1—figure supplement 2***).

Differential gene expression

To identify genes differentially expressed among locations (independent of study group), we created a pseudo-bulk aggregate of all the cells of the same cell-type. Only cell types with a minimum of 100 cell in each location were considered in this analysis. Differences in cell type specific expression were estimated using negative binomial models implemented in DESeq2 (***Love et al., 2014***), including a fixed effect for each individual and location. The distribution of p-values for DEGs between pairs of compartments was assessed using a qq-plot to ensure the statistical models were well calibrated (***Supplementary file 3***). To detect DEGs across study groups we aggregated read counts across locations for each cell-type cluster, excluding cell-types with less than 100 cells in each study group (15 clusters). Differences in cell-type specific expression among study groups were estimated using negative binomial models implemented in Deseq2. Differential gene expression was inferred based on FDR adjusted p-value (q-value <0.1) and fold change >2.0.

Gene ontology and pathway enrichment analyses

The clusterProfiler (***Yu et al., 2012***) package in R was utilized for the identification and visualization of enriched pathways among differentially expressed genes identified as described above. The

functions 'enrichGO', 'enrichKEGG', and 'enrichPathway' were used to identify over-represented pathways based on the Gene Ontology (GO), KEGG, and Reactome databases, respectively. Similar enrichment analyses were also conducted using Gene Set Enrichment Analysis (GSEA) (***Subramanian et al., 2005***) which does not require selection of differentially expressed genes as a first step. Significance in all enrichment analyses were based on $q < 0.05$.

In silico quantification of single-cell signatures in maternal whole blood mRNA

Analysis of transcriptional signatures with advancing gestation and with labor at term

Whole-blood samples collected longitudinally (12 to 40 weeks of gestation) from women with a normal pregnancy who delivered at term with (TIL) (n = 8) or without (TNL) (n = 8) spontaneous labor, were profiled using DriverMap and RNA-Seq, as previously described (***Tarca et al., 2019***) and data were available as GSE114037 dataset in the Gene Expression Omnibus. The $\log_2$ normalized read counts were averaged over the top genes (up to 20, ranked by decreasing fold change) distinguishing each cluster from all others as described above (single-cell signature). Whole blood single-cell signature expression in patients with three longitudinal samples was modeled using linear mixed-effects models with quadratic splines in order to assess the significance of changes with gestational age. Differences in single-cell signature expression associated with labor at term (TIL vs. TNL) were assessed using two-tailed equal variance t-tests. In both analyses, adjustment for multiple signature testing was performed using the false discovery rate method, with $q < 0.1$ being considered significant.

Analysis of transcriptional signatures in preterm labor

Whole blood RNAseq gene expression profiles from samples collected at 24–34 weeks of gestation were previously described (***Paquette et al., 2018***) and data were available as GSE96083 dataset in the Gene Expression Omnibus. The study included samples from 15 women with preterm labor who delivered preterm, and 23 gestational age matched controls. Log_2 transformed pseudo read count data were next transformed into Z-scores based on mean and standard deviation estimated in the control group. Single cell signatures were quantified as the average of Z-scores of member genes and compared between groups using a two-tailed Wilcoxon test. Adjustment for multiple signature testing was performed using the false discovery rate method, with $q < 0.1$ being considered a significant result.

Data and materials availability

The scRNA-seq data reported in this study has been submitted to NIH dbGAP repository (accession number phs001886.v1.p1). All other data used in this study are already available through Gene Expression Omnibus (accession identifiers GSE114037 and GSE96083) and through ArrayExpress (E-MTAB-6701). All software and R packages used herein are detailed in the Materials and methods. Scripts detailing the analyses are also available at https://github.com/piquelab/sclabor. To enable further exploration of the results we have also provided a Shiny App in Rstudio available at: http://placenta.grid.wayne.edu/.

Additional information

Funding

Funder	Grant reference number	Author
Eunice Kennedy Shriver National Institute of Child Health and Human Development	HHSN275201300006C	Roberto Romero
Wayne State University	Perinatal Research Initiative	Nardhy Gomez-Lopez Adi L Tarca

The funders had no role in study design, data collection and interpretation.

Author contributions

Roger Pique-Regi, Resources, Data curation, Software, Formal analysis, Supervision, Investigation, Visualization, Methodology, Project administration; Roberto Romero, Conceptualization, Resources, Supervision, Funding acquisition, Project administration; Adi L Tarca, Resources, Software, Formal analysis, Investigation, Visualization, Methodology; Edward D Sendler, Formal analysis, Investigation, Visualization; Yi Xu, Investigation, Methodology, Project administration; Valeria Garcia-Flores, Methodology, Experiments; Yaozhu Leng, Validation, Investigation, Visualization, Methodology; Francesca Luca, Resources, Methodology; Sonia S Hassan, Resources, Funding acquisition, Project administration; Nardhy Gomez-Lopez, Conceptualization, Resources, Data curation, Supervision, Funding acquisition, Validation, Investigation, Visualization, Methodology, Project administration

Author ORCIDs

Roger Pique-Regi https://orcid.org/0000-0002-1262-2275
Roberto Romero http://orcid.org/0000-0002-4448-5121
Adi L Tarca https://orcid.org/0000-0003-1712-7588
Francesca Luca http://orcid.org/0000-0001-8252-9052
Nardhy Gomez-Lopez https://orcid.org/0000-0002-3406-5262

Ethics

Human subjects: The collection and use of human materials for research purposes were approved by the Institutional Review Boards of the Wayne State University School of Medicine 040302M1F. All participating women provided written informed consent prior to sample collection. Data sharing certification (dbGaP phs001886.v1.p1) has been provided (see data availibility section).

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.52004.sa1
Author response https://doi.org/10.7554/eLife.52004.sa2

Additional files

Supplementary files

- Supplementary file 1. Summary of the scRNA-seq libraries prepared. Each row summarizes each 10X Genomics scRNA-seq library prepared and processed in this study: sample ID, number of cells detected after filtering, location of the tissue (BP = basal plate, PV = Placental Villi, CAM = chorioamniotic membranes), pregnancy condition (TNL = term no labor, TIL = term in labor, PTL = preterm labor), gender of the neonate, and total number of UMIs detected.
- Supplementary file 2. Summary of cell count by cell-type, location and condition. Each row summarizes the total number of cells of each cell-type as determined by Seurat and split by pregnancy condition (TNL = term no labor, TIL = term in labor, PTL = preterm labor), or location of the tissue (BP = basal plate, PV = Placental Villi, CAM = chorioamniotic membranes).
- Supplementary file 3. Marker Genes identified for each cell-type. The columns represent: 1) Cluster or cell-type name, 2) Ensembl gene identifier, 3) Gene symbol, 4) pct.1: percentage of cells in this cluster where the feature is detected, 5) pct.2: percentage of cells in other clusters where the feature is detected, 6) log fold-change of the average expression between this cluster and the rest, 7) Nominal p-value, 8) Adjusted p-value (Bonferroni).
- Supplementary file 4. Genes differentially expressed across compartments for each common cell-type. The columns represent: 1) Cluster or cell-type name, 2) Comparison groups or contrast (i.e., BP vs PV, BP vs CAM, and CAM vs PV), 3) Ensembl gene identifier, 4) Gene symbol, 5) baseMean gene baseline expression as calculated by DESeq2, 6) log2 Fold Change of the first group in column two versus the second group, 7) Standard error estimated for the log2 Fold Change, 8) Nominal p-value, 9) q-value or adjusted p-value to control for FDR. Only rows with $q < 0.2$ are reported.

- Supplementary file 5. Genes differentially expressed across conditions for each cell-type. The columns represent: 1) Cluster or cell-type name, 2) Comparison groups or contrast (i.e., TNL vs TIL, TIL vs PTL), 3) Ensembl gene identifier, 4) Gene symbol, 5) baseMean gene baseline expression as calculated by DESeq2, 6) log2 Fold Change of the first group in column two versus the second group, 7) Standard error estimated for the log2 Fold Change, 8) Nominal p-value, 9) q-value or adjusted p-value to control for FDR. Only rows with q < 0. two are reported.
- Supplementary file 6. Summary of the sample demographics included in this study. Data are given as medians with interquartile ranges (IQR) or as percentages (n/N). [a]One sample missing data.
- Supplementary file 7. Summary of the QC metrics for the scRNA-seq libraries prepared. Each row represents a library, and each column a QC metric reported by the 10X Cellranger software.
- Transparent reporting form

Data availability

Protected Human subjects data deposited in dbGaP phs001886.v1.p1 Data from other sources detailed in manuscript.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Pique-Regi R, Romero R, Tarca AL, Sendler ED, Xu Y, Garcia-Flores V, Leng Y, Luca F, HassanSS, Gomez-Lopez N	2019	Single Cell Transcriptional Signatures of the Human Placenta in Term and Preterm Parturition	https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001886.v1.p1	dbGaP, phs001886.v1.p1

The following previously published datasets were used:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Tarca AL, Romero R, Gomez-Lopez N, Hassan SS, Chenchik A	2018	Targeted sequencing based maternal whole blood expression changes with gestational age and labor in normal pregnancy	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE114037	NCBI Gene Expression Omnibus, GSE114037
Paquette AG, Shynlova O, Kibschull M, Price ND, Lye SJ	2017	Genome Scale Analysis of miRNA and mRNA regulation during preterm labor	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE96083	NCBI Gene Expression Omnibus, GSE96083
Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, Park JE, Stephenson E, Polański K, Goncalves A, Gardner L, Holmqvist S, Henriksson J, Zou A, Sharkey AM, Millar B, Innes B, Wood L, Wilbrey-Clark A, Payne RP, Ivarsson MA, Lisgo S, Filby A, Rowitch DH, Bulmer JN, Wright GJ, Stubbington MJT, Haniffa M, Moffett A, Teichmann SA	2018	Reconstructing the human first trimester fetal-maternal interface using single cell transcriptomics - 10x data	https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-6701/	ArrayExpress, E-MTAB-6701

References

Abrahams VM, Bole-Aldo P, Kim YM, Straszewski-Chavez SL, Chaiworapongsa T, Romero R, Mor G. 2004. Divergent trophoblast responses to bacterial products mediated by TLRs. *The Journal of Immunology* **173**: 4286–4296. DOI: https://doi.org/10.4049/jimmunol.173.7.4286, PMID: 15383557

Al-Garawi A, Carey VJ, Chhabra D, Mirzakhani H, Morrow J, Lasky-Su J, Qiu W, Laranjo N, Litonjua AA, Weiss ST. 2016. The role of vitamin D in the transcriptional program of human pregnancy. *PLOS ONE* **11**:e0163832. DOI: https://doi.org/10.1371/journal.pone.0163832, PMID: 27711190

Aran D, Hu Z, Butte AJ. 2017. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biology* **18**:220. DOI: https://doi.org/10.1186/s13059-017-1349-1, PMID: 29141660

Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, Chak S, Naikawadi RP, Wolters PJ, Abate AR, Butte AJ, Bhattacharya M. 2019. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nature Immunology* **20**:163–172. DOI: https://doi.org/10.1038/s41590-018-0276-y, PMID: 30643263

Arenas-Hernandez M, Romero R, Xu Y, Panaitescu B, Garcia-Flores V, Miller D, Ahn H, Done B, Hassan SS, Hsu CD, Tarca AL, Sanchez-Torres C, Gomez-Lopez N. 2019. Effector and activated T cells induce preterm labor and birth that is prevented by treatment with progesterone. *The Journal of Immunology* **202**:2585–2608. DOI: https://doi.org/10.4049/jimmunol.1801350, PMID: 30918041

Becht E, McInnes L, Healy J, Dutertre C-A, Kwok IWH, Ng LG, Ginhoux F, Newell EW. 2019. Dimensionality reduction for visualizing single-cell data using UMAP. *Nature Biotechnology* **37**:38–44. DOI: https://doi.org/10.1038/nbt.4314

Behnia F, Taylor BD, Woodson M, Kacerovsky M, Hawkins H, Fortunato SJ, Saade GR, Menon R. 2015. Chorioamniotic membrane senescence: a signal for parturition? *American Journal of Obstetrics and Gynecology* **213**:359.e1–35359. DOI: https://doi.org/10.1016/j.ajog.2015.05.041, PMID: 26025293

Belt AR, Baldassare JJ, Molnár M, Romero R, Hertelendy F. 1999. The nuclear transcription factor NF-kappaB mediates interleukin-1beta-induced expression of cyclooxygenase-2 in human myometrial cells. *American Journal of Obstetrics and Gynecology* **181**:359–366. DOI: https://doi.org/10.1016/S0002-9378(99)70562-4, PMID: 10454683

Blencowe H, Cousens S, Oestergaard MZ, Chou D, Moller A-B, Narwal R, Adler A, Vera Garcia C, Rohde S, Say L, Lawn JE. 2012. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *The Lancet* **379**:2162–2172. DOI: https://doi.org/10.1016/S0140-6736(12)60820-4

Bollopragada S, Youssef R, Jordan F, Greer I, Norman J, Nelson S. 2009. Term labor is associated with a core inflammatory response in human fetal membranes, Myometrium, and cervix. *American Journal of Obstetrics and Gynecology* **200**:104.e1–10104. DOI: https://doi.org/10.1016/j.ajog.2008.08.032

Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. 2018. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology* **36**:411–420. DOI: https://doi.org/10.1038/nbt.4096

Chan YW, van den Berg HA, Moore JD, Quenby S, Blanks AM. 2014. Assessment of myometrial transcriptome changes associated with spontaneous human labour by high-throughput RNA-seq. *Experimental Physiology* **99**: 510–524. DOI: https://doi.org/10.1113/expphysiol.2013.072868, PMID: 24273302

Charpigny G, Leroy MJ, Breuiller-Fouché M, Tanfin Z, Mhaouty-Kodja S, Robin P, Leiber D, Cohen-Tannoudji J, Cabrol D, Barberis C, Germain G. 2003. A functional genomic study to identify differential gene expression in the preterm and term human myometrium. *Biology of Reproduction* **68**:2289–2296. DOI: https://doi.org/10.1095/biolreprod.102.013763, PMID: 12606369

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**:15–21. DOI: https://doi.org/10.1093/bioinformatics/bts635, PMID: 23104886

Dobyns AE, Goyal R, Carpenter LG, Freeman TC, Longo LD, Yellon SM. 2015. Macrophage gene expression associated with remodeling of the prepartum rat cervix: microarray and pathway analyses. *PLOS ONE* **10**: e0119782. DOI: https://doi.org/10.1371/journal.pone.0119782, PMID: 25811906

Garrido-Gomez T, Ona K, Kapidzic M, Gormley M, Simón C, Genbacev O, Fisher SJ. 2017. Severe pre-eclampsia is associated with alterations in cytotrophoblasts of the smooth chorion. *Development* **144**:767–777. DOI: https://doi.org/10.1242/dev.146100, PMID: 28232601

Genbacev O, Donne M, Kapidzic M, Gormley M, Lamb J, Gilmore J, Larocque N, Goldfien G, Zdravkovic T, McMaster MT, Fisher SJ. 2011. Establishment of human trophoblast progenitor cell lines from the chorion. *Stem Cells* **29**:1427–1436. DOI: https://doi.org/10.1002/stem.686, PMID: 21755573

Genbačev O, Vićovac L, Larocque N. 2015. The role of chorionic cytotrophoblasts in the smooth chorion fusion with parietal decidua. *Placenta* **36**:716–722. DOI: https://doi.org/10.1016/j.placenta.2015.05.002, PMID: 26003500

Gomez-Lopez N, Romero R, Arenas-Hernandez M, Ahn H, Panaitescu B, Vadillo-Ortega F, Sanchez-Torres C, Salisbury KS, Hassan SS. 2016. In vivo T-cell activation by a monoclonal αcd3ε antibody induces preterm labor and birth. *American Journal of Reproductive Immunology* **76**:386–390. DOI: https://doi.org/10.1111/aji.12562, PMID: 27658719

Gomez-Lopez N, Romero R, Plazyo O, Schwenkel G, Garcia-Flores V, Unkel R, Xu Y, Leng Y, Hassan SS, Panaitescu B, Cha J, Dey SK. 2017. Preterm labor in the absence of acute histologic chorioamnionitis is

characterized by cellular senescence of the chorioamniotic membranes. *American Journal of Obstetrics and Gynecology* **217**:592.e1–59592. DOI: https://doi.org/10.1016/j.ajog.2017.08.008

Haddad R, Tromp G, Kuivaniemi H, Chaiworapongsa T, Kim YM, Mazor M, Romero R. 2006. Human spontaneous labor without histologic chorioamnionitis is characterized by an acute inflammation gene expression signature. *American Journal of Obstetrics and Gynecology* **195**:394–405. DOI: https://doi.org/10.1016/j.ajog.2005.08.057, PMID: 16890549

Hamilton S, Oomomian Y, Stephen G, Shynlova O, Tower CL, Garrod A, Lye SJ, Jones RL. 2012. Macrophages infiltrate the human and rat decidua during term and preterm labor: evidence that decidual inflammation precedes labor. *Biology of Reproduction* **86**:39. DOI: https://doi.org/10.1095/biolreprod.111.095505, PMID: 22011391

Hassan SS, Romero R, Haddad R, Hendler I, Khalek N, Tromp G, Diamond MP, Sorokin Y, Malone J. 2006. The transcriptome of the uterine cervix before and after spontaneous term parturition. *American Journal of Obstetrics and Gynecology* **195**:778–786. DOI: https://doi.org/10.1016/j.ajog.2006.06.021, PMID: 16949412

Hassan SS, Romero R, Tarca AL, Draghici S, Pineles B, Bugrim A, Khalek N, Camacho N, Mittal P, Yoon BH, Espinoza J, Kim CJ, Sorokin Y, Malone J. 2007. Signature pathways identified from gene expression profiles in the human uterine cervix before and after spontaneous term parturition. *American Journal of Obstetrics and Gynecology* **197**:250.e1–25250. DOI: https://doi.org/10.1016/j.ajog.2007.07.008

Hassan SS, Romero R, Tarca AL, Nhan-Chang CL, Vaisbuch E, Erez O, Mittal P, Kusanovic JP, Mazaki-Tovi S, Yeo L, Draghici S, Kim JS, Uldbjerg N, Kim CJ. 2009. The transcriptome of cervical ripening in human pregnancy before the onset of labor at term: identification of novel molecular functions involved in this process. *The Journal of Maternal-Fetal & Neonatal Medicine* **22**:1183–1193. DOI: https://doi.org/10.3109/14767050903353216, PMID: 19883264

Heng YJ, Pennell CE, Chua HN, Perkins JE, Lye SJ. 2014. Whole blood gene expression profile associated with spontaneous preterm birth in women with threatened preterm labor. *PLOS ONE* **9**:e96901. DOI: https://doi.org/10.1371/journal.pone.0096901, PMID: 24828675

Heng YJ, Pennell CE, McDonald SW, Vinturache AE, Xu J, Lee MW, Briollais L, Lyon AW, Slater DM, Bocking AD, de Koning L, Olson DM, Dolan SM, Tough SC, Lye SJ. 2016. Maternal whole blood gene expression at 18 and 28 weeks of gestation associated with spontaneous preterm birth in asymptomatic women. *PLOS ONE* **11**:e0155191. DOI: https://doi.org/10.1371/journal.pone.0155191, PMID: 27333071

Hirota Y, Daikoku T, Tranguch S, Xie H, Bradshaw HB, Dey SK. 2010. Uterine-specific p53 deficiency confers premature uterine senescence and promotes preterm birth in mice. *Journal of Clinical Investigation* **120**:803–815. DOI: https://doi.org/10.1172/JCI40051, PMID: 20124728

Hsu CD, Meaddough E, Aversa K, Copel JA. 1998. The role of amniotic fluid L-selectin, GRO-alpha, and interleukin-8 in the pathogenesis of intraamniotic infection. *American Journal of Obstetrics and Gynecology* **178**:428–432. DOI: https://doi.org/10.1016/S0002-9378(98)70414-4, PMID: 9539502

Kang HM, Subramaniam M, Targ S, Nguyen M, Maliskova L, McCarthy E, Wan E, Wong S, Byrnes L, Lanata CM, Gate RE, Mostafavi S, Marson A, Zaitlen N, Criswell LA, Ye CJ. 2018. Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nature Biotechnology* **36**:89–94. DOI: https://doi.org/10.1038/nbt.4042, PMID: 29227470

Keelan JA, Marvin KW, Sato TA, Coleman M, McCowan LM, Mitchell MD. 1999. Cytokine abundance in placental tissues: evidence of inflammatory activation in gestational membranes with term and preterm parturition. *American Journal of Obstetrics and Gynecology* **181**:1530–1536. DOI: https://doi.org/10.1016/S0002-9378(99)70400-X, PMID: 10601939

Kim YM, Romero R, Chaiworapongsa T, Kim GJ, Kim MR, Kuivaniemi H, Tromp G, Espinoza J, Bujold E, Abrahams VM, Mor G. 2004. Toll-like receptor-2 and -4 in the chorioamniotic membranes in spontaneous labor at term and in preterm parturition that are associated with chorioamnionitis. *American Journal of Obstetrics and Gynecology* **191**:1346–1355. DOI: https://doi.org/10.1016/j.ajog.2004.07.009, PMID: 15507964

Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, Peshkin L, Weitz DA, Kirschner MW. 2015. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**:1187–1201. DOI: https://doi.org/10.1016/j.cell.2015.04.044, PMID: 26000487

Knijnenburg TA, Vockley JG, Chambwe N, Gibbs DL, Humphries C, Huddleston KC, Klein E, Kothiyal P, Tasseff R, Dhankani V, Bodian DL, Wong WSW, Glusman G, Mauldin DE, Miller M, Slagel J, Elasady S, Roach JC, Kramer R, Leinonen K, et al. 2019. Genomic and molecular characterization of preterm birth. *PNAS* **116**:5819–5827. DOI: https://doi.org/10.1073/pnas.1716314116, PMID: 30833390

Koga K, Cardenas I, Aldo P, Abrahams VM, Peng B, Fill S, Romero R, Mor G. 2009. Activation of TLR3 in the trophoblast is associated with preterm delivery. *American Journal of Reproductive Immunology* **61**:196–212. DOI: https://doi.org/10.1111/j.1600-0897.2008.00682.x, PMID: 19239422

Kumazaki K, Nakayama M, Yanagihara I, Suehara N, Wada Y. 2004. Immunohistochemical distribution of Toll-like receptor 4 in term and preterm human placentas from normal and complicated pregnancy including chorioamnionitis. *Human Pathology* **35**:47–54. DOI: https://doi.org/10.1016/j.humpath.2003.08.027, PMID: 14745724

Lea A, Subramaniam M, Ko A, Lehtimäki T, Raitoharju E, Kähönen M, Seppälä I, Mononen N, Raitakari OT, Ala-Korpela M, Pajukanta P, Zaitlen N, Ayroles JF. 2019. Genetic and environmental perturbations lead to regulatory decoherence. *eLife* **8**:e40538. DOI: https://doi.org/10.7554/eLife.40538, PMID: 30834892

Lindström TM, Bennett PR. 2005. The role of nuclear factor kappa B in human labour. *Reproduction* **130**:569–581. DOI: https://doi.org/10.1530/rep.1.00197, PMID: 16264088

Lonergan M, Aponso D, Marvin KW, Helliwell RJ, Sato TA, Mitchell MD, Chaiworapongsa T, Romero R, Keelan JA. 2003. Tumor necrosis factor-related apoptosis-inducing ligand (TRAIL), TRAIL receptors, and the soluble receptor osteoprotegerin in human gestational membranes and amniotic fluid during pregnancy and labor at term and preterm. *The Journal of Clinical Endocrinology & Metabolism* **88**:3835–3844. DOI: https://doi.org/10.1210/jc.2002-021905, PMID: 12915677

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**:550. DOI: https://doi.org/10.1186/s13059-014-0550-8, PMID: 25516281

Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK, Regev A, McCarroll SA. 2015. Highly parallel Genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**:1202–1214. DOI: https://doi.org/10.1016/j.cell.2015.05.002, PMID: 26000488

McGinnis CS, Murrow LM, Gartner ZJ. 2019. DoubletFinder: doublet detection in Single-Cell RNA sequencing data using artificial nearest neighbors. *Cell Systems* **8**:329–337. DOI: https://doi.org/10.1016/j.cels.2019.03.003, PMID: 30954475

Mittal P, Romero R, Mazaki-Tovi S, Tromp G, Tarca AL, Kim YM, Chaiworapongsa T, Kusanovic JP, Erez O, Than NG, Hassan SS. 2009. Fetal membranes as an interface between inflammation and metabolism: increased aquaporin 9 expression in the presence of spontaneous labor at term and chorioamnionitis. *The Journal of Maternal-Fetal & Neonatal Medicine* **22**:1167–1175. DOI: https://doi.org/10.3109/14767050903019692, PMID: 19916714

Mittal P, Romero R, Tarca AL, Gonzalez J, Draghici S, Xu Y, Dong Z, Nhan-Chang CL, Chaiworapongsa T, Lye S, Kusanovic JP, Lipovich L, Mazaki-Tovi S, Hassan SS, Mesiano S, Kim CJ. 2010. Characterization of the myometrial transcriptome and biological pathways of spontaneous human labor at term. *Journal of Perinatal Medicine* **38**:617–643. DOI: https://doi.org/10.1515/jpm.2010.097, PMID: 20629487

Mittal P, Romero R, Tarca AL, Draghici S, Nhan-Chang C-L, Chaiworapongsa T, Hotra J, Gomez R, Kusanovic JP, Lee D-C, Kim CJ, Hassan SS. 2011. A molecular signature of an arrest of descent in human parturition. *American Journal of Obstetrics and Gynecology* **204**:177.e15. DOI: https://doi.org/10.1016/j.ajog.2010.09.025

Muglia LJ, Katz M. 2010. The enigma of spontaneous preterm birth. *New England Journal of Medicine* **362**:529–535. DOI: https://doi.org/10.1056/NEJMra0904308, PMID: 20147718

Nelson AC, Mould AW, Bikoff EK, Robertson EJ. 2016. Single-cell RNA-seq reveals cell type-specific transcriptional signatures at the maternal–foetal interface during pregnancy. *Nature Communications* **7**:e11414. DOI: https://doi.org/10.1038/ncomms11414

Ngo TTM, Moufarrej MN, Rasmussen MH, Camunas-Soler J, Pan W, Okamoto J, Neff NF, Liu K, Wong RJ, Downes K, Tibshirani R, Shaw GM, Skotte L, Stevenson DK, Biggio JR, Elovitz MA, Melbye M, Quake SR. 2018. Noninvasive blood tests for fetal development predict gestational age and preterm delivery. *Science* **360**:1133–1136. DOI: https://doi.org/10.1126/science.aar3819, PMID: 29880692

Nhan-Chang C-L, Romero R, Tarca AL, Mittal P, Kusanovic JP, Erez O, Mazaki-Tovi S, Chaiworapongsa T, Hotra J, Than NG, Kim J-S, Hassan SS, Kim CJ. 2010. Characterization of the transcriptome of chorioamniotic membranes at the site of rupture in spontaneous labor at term. *American Journal of Obstetrics and Gynecology* **202**:462.e1–46462. DOI: https://doi.org/10.1016/j.ajog.2010.02.045

Norwitz ER, Robinson JN, Challis JR. 1999. The control of labor. *New England Journal of Medicine* **341**:660–666. DOI: https://doi.org/10.1056/NEJM199908263410906, PMID: 10460818

Norwitz ER, Bonney EA, Snegovskikh VV, Williams MA, Phillippe M, Park JS, Abrahams VM. 2015. Molecular regulation of parturition: the role of the decidual clock. *Cold Spring Harbor Perspectives in Medicine* **5**:a023143. DOI: https://doi.org/10.1101/cshperspect.a023143, PMID: 25918180

Osman I, Young A, Ledingham MA, Thomson AJ, Jordan F, Greer IA, Norman JE. 2003. Leukocyte density and pro-inflammatory cytokine expression in human fetal membranes, decidua, cervix and myometrium before and during labour at term. *Molecular Human Reproduction* **9**:41–45. DOI: https://doi.org/10.1093/molehr/gag001, PMID: 12529419

Paquette AG, Shynlova O, Kibschull M, Price ND, Lye SJ. 2018. Comparative analysis of gene expression in maternal peripheral blood and monocytes during spontaneous preterm labor. *American Journal of Obstetrics and Gynecology* **218**:345.e1–34345. DOI: https://doi.org/10.1016/j.ajog.2017.12.234

Pavličev M, Wagner GP, Chavan AR, Owens K, Maziarz J, Dunn-Fletcher C, Kallapur SG, Muglia L, Jones H. 2017. Single-cell transcriptomics of the human placenta: inferring the cell communication network of the maternal-fetal interface. *Genome Research* **27**:349–361. DOI: https://doi.org/10.1101/gr.207597.116, PMID: 28174237

Polettini J, Behnia F, Taylor BD, Saade GR, Taylor RN, Menon R. 2015. Telomere fragment induced amnion cell senescence: a contributor to parturition? *PLOS ONE* **10**:e0137188. DOI: https://doi.org/10.1371/journal.pone.0137188, PMID: 26397719

Romero R, Brody DT, Oyarzun E, Mazor M, Wu YK, Hobbins JC, Durum SK. 1989a. Infection and labor. III. Interleukin-1: a signal for the onset of parturition. *American Journal of Obstetrics and Gynecology* **160**:1117–1123. DOI: https://doi.org/10.1016/0002-9378(89)90172-5, PMID: 2786341

Romero R. 1989b. Infection and labor. IV. Cachectin-tumor necrosis factor in the amniotic fluid of women with intraamniotic infection and preterm labor. *Am J Obstet Gynecol* **161**:336–341.

Romero R, Avila C, Santhanam U, Sehgal PB. 1990a. Amniotic fluid interleukin 6 in preterm labor. Association with infection. *Journal of Clinical Investigation* **85**:1392–1400. DOI: https://doi.org/10.1172/JCI114583

Romero R, Parvizi ST, Oyarzun E, Mazor M, Wu YK, Avila C, Athanassiadis AP, Mitchell MD. 1990b. Amniotic fluid interleukin-1 in spontaneous labor at term. *The Journal of Reproductive Medicine* **35**:235–238. PMID: 2325034

Romero R, Ceska M, Avila C, Mazor M, Behnke E, Lindley I. 1991. Neutrophil attractant/activating peptide-1 / interleukin-8 in term and preterm parturition. *American Journal of Obstetrics and Gynecology* **165**:813–820. DOI: https://doi.org/10.1016/0002-9378(91)90422-N

Romero R, Mazor M, Brandt F, Sepulveda W, Avila C, Cotton DB, Dinarello CA. 1992a. Interleukin-1 alpha and interleukin-1 beta in preterm and term human parturition. *American Journal of Reproductive Immunology* **27**: 117–123. DOI: https://doi.org/10.1111/j.1600-0897.1992.tb00737.x, PMID: 1418402

Romero R, Mazor M, Sepulveda W, Avila C, Copeland D, Williams J. 1992b. Tumor necrosis factor in preterm and term labor. *American Journal of Obstetrics and Gynecology* **166**:1576–1587. DOI: https://doi.org/10.1016/0002-9378(92)91636-O

Romero R, Yoon BH, Kenney JS, Gomez R, Allison AC, Sehgal PB. 1993. Amniotic fluid interleukin-6 determinations are of diagnostic and prognostic value in preterm labor. *American Journal of Reproductive Immunology* **30**:167–183. DOI: https://doi.org/10.1111/j.1600-0897.1993.tb00618.x, PMID: 8311926

Romero R, Tarca AL, Tromp G. 2006a. Insights into the physiology of childbirth using transcriptomics. *PLOS Medicine* **3**:e276. DOI: https://doi.org/10.1371/journal.pmed.0030276, PMID: 16752954

Romero R, Espinoza J, Kusanovic JP, Gotsch F, Hassan S, Erez O, Chaiworapongsa T, Mazor M. 2006b. The preterm parturition syndrome. *BJOG: An International Journal of Obstetrics & Gynaecology* **113**:17–42. DOI: https://doi.org/10.1111/j.1471-0528.2006.01120.x

Romero R, Dey SK, Fisher SJ. 2014a. Preterm labor: one syndrome, many causes. *Science* **345**:760–765. DOI: https://doi.org/10.1126/science.1251816, PMID: 25124429

Romero R, Tarca AL, Chaemsaithong P, Miranda J, Chaiworapongsa T, Jia H, Hassan SS, Kalita CA, Cai J, Yeo L, Lipovich L. 2014b. Transcriptome interrogation of human myometrium identifies differentially expressed sense-antisense pairs of protein-coding and long non-coding RNA genes in spontaneous labor at term. *The Journal of Maternal-Fetal & Neonatal Medicine* **27**:1397–1408. DOI: https://doi.org/10.3109/14767058.2013.860963, PMID: 24168098

Saelens W, Cannoodt R, Todorov H, Saeys Y. 2019. A comparison of single-cell trajectory inference methods. *Nature Biotechnology* **37**:547–554. DOI: https://doi.org/10.1038/s41587-019-0071-9, PMID: 30936559

Santhanam U, Avila C, Romero R, Viguet H, Ida N, Sakurai S, Sehgal PB. 1991. Cytokines in normal and abnormal parturition: elevated amniotic fluid interleukin-6 levels in women with premature rupture of membranes associated with intrauterine infection. *Cytokine* **3**:155–163. DOI: https://doi.org/10.1016/1043-4666(91)90037-E, PMID: 1888885

Shynlova O, Nedd-Roderique T, Li Y, Dorogin A, Nguyen T, Lye SJ. 2013. Infiltration of myeloid cells into decidua is a critical early event in the labour cascade and post-partum uterine remodelling. *Journal of Cellular and Molecular Medicine* **17**:311–324. DOI: https://doi.org/10.1111/jcmm.12012, PMID: 23379349

Sirota M, Thomas CG, Liu R, Zuhl M, Banerjee P, Wong RJ, Quaintance CC, Leite R, Chubiz J, Anderson R, Chappell J, Kim M, Grobman W, Zhang G, Rokas A, England SK, Parry S, Shaw GM, Simpson JL, Thomson E, et al. 2018. Enabling precision medicine in neonatology, an integrated repository for preterm birth research. *Scientific Data* **5**:e180219. DOI: https://doi.org/10.1038/sdata.2018.219, PMID: 30398470

Stanfield Z, Lai PF, Lei K, Johnson MR, Blanks AM, Romero R, Chance MR, Mesiano S, Koyutürk M. 2019. Myometrial transcriptional signatures of human parturition. *Frontiers in Genetics* **10**:185. DOI: https://doi.org/10.3389/fgene.2019.00185, PMID: 30988671

Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, Purdom E, Dudoit S. 2018. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**:477. DOI: https://doi.org/10.1186/s12864-018-4772-0, PMID: 29914354

Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Hao Y, Stoeckius M, Smibert P, Satija R. 2019. Comprehensive integration of Single-Cell data. *Cell* **177**:1888–1902. DOI: https://doi.org/10.1016/j.cell.2019.05.031

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *PNAS* **102**:15545–15550. DOI: https://doi.org/10.1073/pnas.0506580102, PMID: 16199517

Tarca AL, Romero R, Xu Z, Gomez-Lopez N, Erez O, Hsu CD, Hassan SS, Carey VJ. 2019. Targeted expression profiling by RNA-Seq improves detection of cellular dynamics during pregnancy and identifies a role for T cells in term parturition. *Scientific Reports* **9**:848. DOI: https://doi.org/10.1038/s41598-018-36649-w, PMID: 30696862

Tsang JCH, Vong JSL, Ji L, Poon LCY, Jiang P, Lui KO, Ni YB, To KF, Cheng YKY, Chiu RWK, Lo YMD. 2017. Integrative single-cell and cell-free plasma RNA transcriptomics elucidates placental cellular dynamics. *PNAS* **114**:E7786–E7795. DOI: https://doi.org/10.1073/pnas.1710470114, PMID: 28830992

Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, Park JE, Stephenson E, Polański K, Goncalves A, Gardner L, Holmqvist S, Henriksson J, Zou A, Sharkey AM, Millar B, Innes B, Wood L, Wilbrey-Clark A, Payne RP, et al. 2018. Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* **563**:347–353. DOI: https://doi.org/10.1038/s41586-018-0698-6, PMID: 30429548

Vora S, Abbas A, Kim CJ, Summerfield TL, Kusanovic JP, Iams JD, Romero R, Kniss DA, Ackerman WE. 2010. Nuclear factor-kappa B localization and function within intrauterine tissues from term and preterm labor and

cultured fetal membranes. *Reproductive Biology and Endocrinology* **8**:8. DOI: https://doi.org/10.1186/1477-7827-8-8, PMID: 20100341

Xu Y, Plazyo O, Romero R, Hassan SS, Gomez-Lopez N. 2015. Isolation of leukocytes from the human Maternal-fetal interface. *Journal of Visualized Experiments*:e52863. DOI: https://doi.org/10.3791/52863, PMID: 26067211

Xu Y, Romero R, Miller D, Kadam L, Mial TN, Plazyo O, Garcia-Flores V, Hassan SS, Xu Z, Tarca AL, Drewlo S, Gomez-Lopez N. 2016. An M1-like macrophage polarization in decidual tissue during spontaneous preterm labor that is attenuated by rosiglitazone treatment. *The Journal of Immunology* **196**:2476–2491. DOI: https://doi.org/10.4049/jimmunol.1502055

Yan X, Sun M, Gibb W. 2002. Localization of nuclear factor-kappa B (NF kappa B) and inhibitory factor-kappa B (I kappa B) in human fetal membranes and decidua at term and preterm delivery. *Placenta* **23**:288–293. DOI: https://doi.org/10.1053/plac.2002.0789, PMID: 11969339

Young A, Thomson AJ, Ledingham M, Jordan F, Greer IA, Norman JE. 2002. Immunolocalization of proinflammatory cytokines in Myometrium, Cervix, and fetal membranes during human parturition at Term1. *Biology of Reproduction* **66**:445–449. DOI: https://doi.org/10.1095/biolreprod66.2.445

Yu G, Wang L-G, Han Y, He Q-Y. 2012. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology* **16**:284–287. DOI: https://doi.org/10.1089/omi.2011.0118

Appendix 1

Cell type labeling procedures

Multiple methods and resources were utilized to label the clusters identified by Seurat. First, we used the function FindAllMarkers to identify the genes with significant changes in expression between each cluster and the rest of the cells using min.pct=0.33 and requiring FDR adjusted q < 10% and a $\log$(FC)> 0.5 to determine significance. Clusters with no significant differences at this threshold were merged resulting in a total of 19 clusters. For each cluster, we generated a pseudo-bulk gene expression profile and xCell (http://xcell.ucsf.edu/) (***Aran et al., 2017***) was used to compare the gene expression signatures of our clusters with those of known cell types to the default n = 64 xCell reference which includes immune cells, progenitor, epithelial, and extracellular matrix cells. Eight of the original clusters clearly identified with known cell types in the xCell reference panel that includes T-cell, B-cell, Macrophage, HSC, Fibroblast and Monocyte.

The next method we used is by comparing the marker genes identified by Seurat FindAllMarkers that passed the threshold to previously published scRNAseq marker genes (***Tsang et al., 2017***; ***Pavličev et al., 2017***; ***Vento-Tormo et al., 2018***) and common known markers used by our group and others https://www.proteinatlas.org/search/placenta (***Figure 1—figure supplement 1***). This resolved many of our placental (non-immune) cell clusters in the following cell types: cytotrophoblast, extravillous trophoblast, syncytiotrophoblast, decidual, endothelial, and stromal cells. To further resolve genes differentially expressed between clusters in close proximity to each other (e.g., T-cell subtypes), we ran Seurat FindMarkers function to contrast gene expression between each cluster pair, and determined as differentially expressed genes those showing a minimum logFC> 0.25 and q < 0.1. Using this analysis, we were able to label two subgroups of T cells as activated or resting. Clusters that were distinct but could not be clearly separated into well-known cell sub-types or cellular states were assigned a number (e.g., Stromal-1, Stromal-2). Some of the differences between these clusters are likely due to the maternal/fetal origin of each cell type as shown in ***Figure 1B*** (i.e., Macrophage 1 is likely maternal and Macrophage 2 is likely fetal) as shown by genotype analysis freemuxlet (see Materials and methods). Additionally, we used DoubletFinder https://github.com/chris-mcginnis-ucsf/DoubletFinder (***McGinnis et al., 2019***) to identify doublet cells and to ensure that none of our clusters were confounded by doublets (***Figure 1—figure supplement 8***).

Finally, we also compared our manually curated cell type identification to that derived from automated cell labeling methods SingleR (***Aran et al., 2019***) and Seurat (***Stuart et al., 2019***), (see ***Figure 1—figure supplement 5*** and ***Figure 1—figure supplement 6***). Automated annotation provides a convenient way of transferring biological knowledge across datasets, thereby reducing the burden of interpreting clusters, but it is important to manually curate the cell labels using well established biological knowledge. If the reference database is not specific for the same tissue or similar conditions, this could lead to incorrect assignments. For SingleR, we used the vignette detailed in https://bioconductor.org/packages/devel/bioc/vignettes/SingleR/inst/doc/SingleR.html using the human primary cell atlas (HPCA) reference provided by SingleR and the human placenta first trimester (HPFT) single cell data made available by another group (***Vento-Tormo et al., 2018***) downloaded from https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-6701/. For Seurat, we used only the latter reference and the standard workflow detailed in https://satijalab.org/seurat/v3.1/integration.html, and we removed any cell label with a max score > 0.001 (with almost identical results if the threshold was 0.01 or 0.0001). Similarly, we only used the pruned labels provided by SingleR.

Lymphoid cell types

Four of our clusters correspond to lymphocyte cell-types: B-cell, NK-cell, T-cell activated, and T-cell resting. The cluster labeled as B-cell has an xCell score of 0.88 and express very highly *CD79A*. The automated labeling methods also clearly identify this cluster as B-cell when using the HPCA reference, while it was identified as Plasma cell when using the HPFT reference

panel, as no cell type is labeled as B-cell in HPFT and Plasma cell would make sense as a close match. The cluster labeled NK-cell express very highly *GNLY* and *NKG7* genes, and is also very well matched to NK-cell in the HPCA reference or one of the many NK cell types in HPFT (***Vento-Tormo et al., 2018***), which was a major focus of that study that also enriched for more rare NK cell types as they have a very important role in first trimester pregnancy, but here we only see evidence for one NK-cell cluster in (***Figure 1—figure supplement 7***). Our two clusters labeled as T-cells also closely matched the T-cell types for both reference panels and had xCell scores > 0.5, but only one T cell type is provided by those reference panels. Here, our two clusters differed in some of the genes being expressed that showed that one of the clusters was more active as indicated by signaling factors such as pro-inflammatory cytokine *TNF* and AP-1 factors such as *FOSL* and *JUNB*.

Myeloid cell types

Three of our clusters closely matched myeloid cell types: Macrophage 1, Macrophage 2, and Monocyte. Each of these clusters closely matched to their respective cell types (xCell score > 0.8) and also when using SingleR and Seurat automated label transfer from both reference panels. Macrophage 2, which seemed to be of fetal origin, matched the Hofbauer cell type from the HPFT reference (***Vento-Tormo et al., 2018***), which are fetal resident macrophages found in the human placenta.

Trophoblasts and other cell types

The major trophoblast cell types (CTB, EVT, and STB) expressed the markers that were defined in ***Tsang et al. (2017)***. The newly identified npiCTB also expressed the canonical CTB markers, but had a significantly higher expression of PAGE4 and decreased expression of *DDX3X*, *EIF1AX*, and *XIST* that indicate a non-proliferative state. Using automated cell labeling methods, CTB matched with VCT as defined in HPFT (***Vento-Tormo et al., 2018***), except for a small proportion that matched the SCT profile in HPFT (***Vento-Tormo et al., 2018***) (***Figure 1—figure supplement 7***). This finding may be due to differences in the expression profile of the trophoblast cells types between early and late pregnancy. The SCT in the reference panel (first trimester placental scRNA-seq data) may also include the profile of the transient stage between CTB and STB. This is supported by the trajectory analysis shown in Figure (***Figure 1—figure supplement 2***). Our EVT and STB clusters matched the labels from the automated method using the HPFT reference panel. Other small clusters showing stromal cells matched related cell types described in HPFT (***Vento-Tormo et al., 2018***).