Title: Assessing placental maturity through histological and transcriptomic analyses in idiopathic spontaneous preterm birth.

Heather M Brockway[1*], Helen N Jones[2], William E Ackerman IV[3], Irina A Buhimschi[4], Catalin S Buhimschi[3]

Suhas G Kallapur[5], and Louis J Muglia[1]

[1] *Center for Preterm Birth, Cincinnati Children's Hospital Medical Center, Cincinnati OH 45229 USA*

[2] *Center for Fetal and Placental Research, Cincinnati Children's Hospital Medical Center, Cincinnati OH 45229 USA*

[3] *Department of Obstetrics and Gynecology, The Ohio State University, Columbus OH 43210 USA*

[4] *Center for Perinatal Research, Nationwide Children's Hospital, Columbus OH 43210 USA*

[5] *Divisions of Neonatology and Developmental Biology, David Geffen School of Medicine at the University of California, Los Angeles, CA, 90095, USA*

*Corresponding author (Heather.Brockway@cchmc.org)

Abstract:

Preterm birth (PTB) is leading contributor to infant death in the United States and globally, yet the underlying mechanistic causes are not well understood. Previous studies have suggested a role for advanced villous maturity in both spontaneous and iatrogenic preterm birth. To better understand pathological and molecular basis of idiopathic spontaneous preterm birth (isPTB), we compared placental morphology and transcriptomic analysis in carefully phenotyped cohorts of PTB due to intraamniotic infection, isPTB, and healthy term placentae. Characteristic features of precocious placental villous maturation were uniquely demonstrated in isPTB placentae. Transcriptomic analyses revealed isPTB candidate genes. These include an upregulation of three IGF binding proteins (*IGFBP1, IGFBP2,* and

IGFBP6), supporting a role for IGF signaling in isPTB. Additional Gene Ontology analyses identified alterations in biological processes such as immunological activation and programmed cell death. Our data suggest that premature placental aging may contribute to the pathogenesis of isPTB and provide a molecular basis of this subset of cases of preterm birth.

Introduction

Every year, 1 million infants die from complications resulting from their birth before 37 completed weeks of gestation. In 2017, the incidence of prematurity in the United States was 9.6%[1] and worldwide incidences reached approaching 20% in some regions[2]. While preterm birth is a multifactorial syndrome, there are two primary classifications: spontaneous preterm birth (sPTB) and iatrogenic preterm birth as a result of fetal or maternal complications. Although risk factors have been identified the underlying molecular mechanisms of truly idiopathic spontaneous preterm birth (isPTB) remain unclear[3].

The role, if any, of the placenta in isPTB is not clearly defined and remains under investigated. The placenta is a transient two-sided organ, providing a maternal/fetal interface, and its proper development and function is essential to a successful pregnancy outcome[4]. Placentation is the result of a highly complex web of molecular mechanisms originating from both the mother and the fetus, many of which are not yet fully understood even under healthy, normal conditions. Recent advances in placental transcriptomics have identified changes in gene expression and regulatory mechanisms across normal gestation[5-8]. Yet, transcriptomics of isPTB are currently limited.

During the third trimester of pregnancy, specific hallmarks of placental maturity are observed including an increased number of terminal villi, syncytial nuclear aggregates, and vasculosyncytial membranes [9-11]. Syncytial nuclear aggregates (SNA) are multi-layered aggregate of at least 10 syncytial nuclei extending out from the villous surface, that do not come in contact with other villi [9]. Terminal villi were defined as branched villi <80μm in diameter. Vasculosyncytial membranes are defined as regions in the villi where the fetal capillaries are immediately adjacent to areas of the syncytiotrophoblast free from

syncytial nuclei, thus are areas of direct diffusion [11]. These hallmarks are utilized in histological assessments to determine the maturity of the placenta, with the term placentas possessing the highest amounts of these hallmarks[9-11]. Previous histological studies of the villous trophoblast in isPTB (<37 weeks GA) without intra-amniotic infection identified at least two distinct morphological phenotypes, with the significant majority of placental samples demonstrating advanced villous maturation (AVM). The AVM samples reflected all the hallmarks of a term placenta; the remaining samples had no hallmarks of AVM[12,13]. These data suggest there are multiple subclasses of isPTB: those with or without infection and those with or without AVM. Thus, identifying and assessing placental maturity through histological and transcriptomic studies is necessary to define aberrant molecular mechanisms underlying the placental pathophysiology associated with subtypes of isPTB. The objective of this study was to identify transcriptomic signatures of placental maturity in a histologically defined cohort of idiopathic spontaneous preterm births.

Results

Study Characteristics

Maternal and fetal characteristics for the three different pregnancy outcomes included in this study are presented in Table 1. Significant differences were observed in gestational age and fetal weights between intra-amniotic infection (IAI) and isPTB samples compared to the term samples (P<0.001). Among the isPTB samples, two of neonates were small for gestational age (SGA) with a fetal weight less than the 10th percentile. Within the IAI samples, only one neonate was SGA. All term births for which there was fetal weight available (n=9) were appropriate for gestational age.

Assessment of advanced maturity in isPTB villi

Stereological assessment identified no significant differences between the isPTB and term samples in number of syncytial nuclear aggregates (SNAs) or terminal villi per high powered field (villi<80μm in

diameter) (Table 2; Figure 1A-C, F-G). One isPTB case could not be classified as there was no sample left to assess morphology. There was a trend towards no significance between the IAI and term samples in the numbers of terminal villi observed (P=0.078 Table 2) with no differences observed in the number of SNAs between IAI and term samples. Vasculosyncytial membrane counts were significantly reduced in IAI samples compared to isPTB and term samples (P=0.001 Table 2 and Figure 1). Perivillous fibrin deposits were observed in each of the sample types (Figure 1).

RNA Sequence results

All fastq files passed initial quality control assessment in FASTQC. A total of 1,246,073,145 unpaired reads were generated for the 31 samples and 761,182,139 reads (61.08%) were successfully aligned once to the human genome GRCh38 (Supplemental Tables S1 and S2). As there are multiple paralogous gene families expressed in the placenta, we opted not to include reads that mapped multiple times in our final analyses. Further quality control assessments were completed to examine sequence quality per sample (Supplemental Figure 2 A-C).

Identification of differentially expressed genes

Due to the origin of the samples and inclusion of a pre-existing dataset, we performed multiple quality control assessments, including principle component analyses (PCA) to identify potential batch effects (Supplemental Figure S1). We did not observe any significant batch effects and thus, did not remove or control for them in subsequent tests for differential gene expression. The IAI samples were entirely female in origin and had a small sample size; therefore, so we did not control for any covariate factors such as fetal sex as it could be potentially biased. We performed differential gene expression testing in three pairwise comparisons: IAI compared to term births, isPTB compared to IAI birth, and isPTB compared to term births. Genes were considered significant with an adjusted P-value of <0.1 and absolute log2 fold-change >1.5 (Figure 2A, all significant genes labeled red). We identified 160

significant genes in IAI verses term, with 117 upregulated and 43 downregulated. In the isPTB verses IAI comparison, we identified a total of 94 significant genes with 62 upregulated and 32 downregulated. Lastly, in the isPTB verses term comparison, we identified 158 significant genes with 157 upregulated and only 2 genes downregulated (Supplemental Tables S3,S4 and S5)

Categorization of maturation, gestation, and isPTB candidate genes

Differently expressed genes alone are not enough to identify transcriptomic signatures due to advanced maturation, gestational age, infection, or isPTB pathophysiology. Therefore, to identify candidate genes in each of these categories, we intersected the significant genes, both upregulated and downregulated, from each of the differential gene expression comparisons (Figure 2B). Genes categorized as infection (n=37) are represented in the intersection of IAI vs term and isPTB vs term. A previous study by Ackerman et al[14] using a subset of these data profiled genes and pathways involved in intra-amniotic infection and we did not further explore these results. We did confirm that several of the genes they identified in their study (*ACTA1, FUT9, MPO, S100A12*) were present in this category in our results. Gestational age genes (total n= 123) are represented by the intersection of IAI vs term and isPTB vs term (n=11) and the genes exclusive to IAI vs term (n=112) (Figure 2B). Maturation genes (total n= 186) are represented by the intersection of isPTB vs IAI (n=18) and the genes exclusive to isPTB v IAI (n=39) and isPTB vs term (n=129) (Figure 2B).

To further refine our analysis, we compared the expression pattern of each candidate genes in each group across all three differential expression datasets (Figure 2C). To detect a maturation signal, we examined the 186 maturation candidate genes and identified those differentially expressed in the isPTB v IAI data with an absolute log2 fold change of >1.5. We then compared their expression across the other two differential expression datasets to identify genes that a similar pattern of expression in term v IAI, but were not differentially expressed in isPTB v term. 21 genes met these criteria and are represented in a

heatmap (Figure 2; Table 3). Of these genes, 10 were upregulated and 11 were downregulated. We also were able to identify 13 of maturation genes which demonstrate differences in expression in isPTB vs IAI and term vs IAI similar to the maturation signal candidates; however, they are also differentially expressed in the isPTB vs term data (Figure 2; Table 3). Within this subset of maturation genes, 10 genes were upregulated and 4 were downregulated. Lastly, we were able to identify isPTB specific genes from the remaining maturation candidates by identifying genes in the isPTB vs term with an absolute log2 fold change of 1.5 and with a similar expression pattern in isPTB v IAI and the opposite expression or no difference in expression in term vs IAI (n=141) (Figure 3 and Table 4). Of these genes, only 2 were downregulated with the remaining 139 genes being upregulated.

We identified gestational age candidates using the same approach, first identifying genes in the IAI vs term data with an absolute log2 fold change of > 1.5 as this represents the greatest difference in gestational ages between our samples. We then compared the expression changes in isPTB v term and isPTB v IAI (Figure 2). While we did observe differences in expression in the isPTB vs IAI comparison, this is likely due to the differences in gestational ages of these samples (isPTB 29-36 wks v IAI 25-31wks). Using this approach, we identified 32 candidate genes with 29 of the genes upregulated and 3 downregulated (Figure 2; Supplemental Table S6).

Functional classification and enrichment analyses for candidate genes

We assessed enrichment in each of the candidate gene categories. There was no enrichment for Reactome pathways or GO terms in either of the maturity categories. However, there was enrichment in the gestational age candidates for cellular components related to the extracellular region (GO:0005576). These included 19 genes including receptors and ligands for the WNT signaling pathway, cell proliferation, and inflammatory response. The candidate genes associated with isPTB physiopathology had significant pathway enrichment including the IGF signaling pathway including three IGF binding

proteins (Table 7). Furthermore, the isPTB candidate genes were enriched for GO terms associated with immunity, signaling, and regulation of blood flow (Table 5 and Supplemental Tables S7 and S8).

In addition to the enrichment analyses, we also performed functional classifications on all the candidate genes to identify additional pathways and functions (Supplemental Figures S3 and S4). The maturation candidate genes were divided into two subsets: those that showed no differences in expression (signal) between isPTB and term, and a subset that were expressed more in isPTB than term (drivers). While the candidate genes shared many biological processes, notable differences occur between the groups specifically within the metabolic processes (GO:0008152), immune system processes(GO:0002376), and locomotion (GO:0040011) (Supplemental Figure S3).

Discussion:

The numerous subclassifications of preterm birth coupled with a limited understanding of the placental role in birth timing, have made identifying the etiological underpinnings of this devastating pregnancy outcome exceedingly difficult. We previously attempted to identify molecular signatures of spontaneous PTB using publicly available microarray data[15] and identified several placental genes and pathways associated with spontaneous preterm birth. However, those analyses lacked complete covariate information, appropriate gestational age controls, and were on mixed array platforms. In this study, we overcame those limitations and combined transcriptome analyses with histological assessment of matched placental samples to assess maturity in order to elucidate the placental role in isPTB, and to identify a placental molecular signature associated with isPTB.

Morgan et al. hypothesized that placental maturity, as assessed by markers of villous maturation, and not infection may be the leading cause of both idiopathic and iatrogenic PTB[12,13]. However, only one other study has linked placental maturity to the molecular etiology of PTB. A recent study by Leavey et al. of placental maturity in pre-eclampsia (a common reason for iatrogenic preterm delivery) identified morphological and molecular similarities between PE placentas with advanced villous maturity (AVM)

and normal term placentas[16]. Furthermore, they demonstrated that AVM placentas had a shift in their molecular signature and thus appeared older than their actual gestational age at delivery. These data along with Morgan et al.[13] suggest a role for the placenta in PTB etiology, a role where AVM is potentially affecting placental output through increased terminal villi, more syncytiotrophoblast, and thus an increase in placental output in terms of secreted proteins and exosomes earlier in gestation. If placental output has a role in modulating birth timing, the AVM placentas regardless of whether they are idiopathic or iatrogenic may lead to PTB. These data also indicate that placental maturity and its associated molecular signatures may have a profound impact on how we utilize placental output to assess adverse pregnancy outcomes such as PTB and its subclassifications in the clinical setting.

In our current study, all isPTB placentas demonstrated changes in villous structure consistent with AVM along with peri-villous fibrin deposition with no significant differences observed compared to term births despite being delivered on average 5.8 weeks before term. This is consistent with previous histological observations[12,13] which defined AVM as an increase in syncytial nuclear aggregates and terminal villi. Our assessment of vasculosyncytial membranes and fibrin deposition further strengthens the interpretation of advanced placental maturity in isPTB. In contrast, the IAI samples demonstrated structural hallmarks appropriate for their gestational maturity and age, as previous analysis of those samples concluded acute inflammation was the likely cause of their preterm delivery[14]. Given that control tissues from 20-36 weeks of gestation are not available due to the ethics of mid-late-gestation sampling or termination, placentas from IAI were the most appropriate controls that we could utilize for our transcriptomic analyses. Placentas in both the isPTB and IAI groups demonstrated perivillous fibrin deposition which may impact the amount placental surface area available for nutrient and oxygen exchange, leading to reduced fetal growth and SGA neonates. However, only three neonates were observed to be SGA, two were isPTB and one IAI; therefore, we do not believe the fibrin deposition is problematic or affecting placental function.

One of the primary obstacles in utilizing placental samples for transcriptomics has always been the lack of appropriate gestational aged (GA) control placentas. In most studies, GA matched placentas are not normal and thus are not suitable controls. As Ackerman et al[14] have shown, the most parsimonious cause for the preterm birth in the IAI samples is infection, not adverse placental maturity or even placental insufficiency, making them the most appropriate controls. As such, we were able to then use them to conduct three pairwise comparisons to identify candidate genes that represented differences in maturity, gestation, and isPTB specific to their expression patterns.

The primary goal of this study was to determine if a molecular signature of maturity and isPTB could be identified from placentas demonstrating AVM. In our preliminary analyses, we classified significantly differentially expressed genes into gestational age, maturity, and isPTB candidates. Two maturity candidate genes, keratin 24 (*KRT24*) and shisa like 1 (*SHISAL1*) were significantly upregulated between the isPTB and term samples compared to IAI samples. Keratins are intermediate filament proteins expressed in epithelial cells that have a variety of roles in the cell including providing structural support to the cell and cellular mechanics[17]. While KRT24 is similar to other type I keratins, is found in epithelial cells, and is believed to be a terminal differentiation marker. Its increased expression induces senescence, autophagy, and apoptosis[18]. Given that KRT24 localizes to the villous trophoblasts[19], the increase in terminal villi observed in isPTB likely accounts for increase expression. *SHISAL1* function is currently unknown, but the shisa family of genes encode transmembrane proteins with roles in development as wingless and INT1 (WNT) and fibroblast growth factor (FGF) signaling antagonists[20] and myocyte fusion[21].

The two most downregulated candidates genes for maturity were cellular retinoic acid binding protein 1 (C*RABP1*) and lysosomal associated membranes protein family member 5 (*LAMP5*). *CRABP1* is an epigenetically modulated tumor suppressor gene with a known role in retinoic acid signaling in the placenta[22]. Its downregulation in the isPTB and term samples could indicate altered retinoic acid transport and metabolism in fetal tissues[22,23]. LAMP5 is a recently identified member of the lysosomal associated

membrane protein gene family[24]. Unlike the other broadly expressed family members, LAMP5 expression has only been observed in the brain, dendritic cells, and placental trophoblasts[19]. LAMP proteins are involved in autophagy and lysosomal formation and transport[25]. While we did not identify any specific enrichment for pathways in the maturity candidate gene list, functional analyses revealed WNT signaling and transforming growth factor beta (TGFβ) signaling as potential pathways of interest. We also examined the biological processes represented by the maturation genes although none were significantly enriched. Given that various bioprocesses these genes represent are active in placental trophoblasts, further investigations to refine their roles in placental development and maturation longitudinally are warranted.

For the gestational age candidates, growth regulating estrogen receptor binding 1 (*GREB1*) and dickkof WNT signaling pathway inhibitor (*DKK1*) are upregulated in isPTB and IAI compared to term while carboxypeptidase X, M14 family member (*CPXM2*) is most downregulated gene. GREB1 appears to be localized to the fetal endothelial cells[19] and is known to localize to maternal endometrial epithelial cells[26]. GREB1 function has not been studied in fetal villous physiology; however, in endometriosis[26] and more recently in decidualization[18]. *DKK1*, a WNT signaling antagonist, was also identified as a highly upregulated gene in our array study; however, it was not significant. While previous studies have focused on its role in the decidua, we show that DKK1 localized to the syncytiotrophoblast in isPTB with a reduction in expression in term samples as expected from the expression data. It is known that WNT signaling, is essential to placental development acting through trophoblast proliferation and inhibition of apoptosis[27,28]. It has been theorized that aberrant DKK1 expression, especially upregulation in iatrogenic PTB samples, could be associated with etiology or pathophysiology[28,29]. However, these studies as well as our previous transcriptomics study lack appropriate gestational age controls. Given our data in this study, we posit that the difference in *DKK1* expression between PTB samples is due to gestational age rather than pathology. Yet, as it does have such a strong expression pattern between gestational ages, it is worth further study to determine its precise role in normal placental maturation.

Within the isPTB candidate genes, we observed significant enrichment for genes within the IGF (insulin like growth factor) signaling pathway, specifically IGF binding proteins: *IGFBP1*, *IGFBP2*, *IGFBP6* (Tables 4 and 5). IGFBPs bind to IGF1 and IGF2 modulating their bioavailability to activate the IGF signaling pathway[30]. IGFBP1 was thought to be primarily expressed in the decidua, but we and others[31] have shown that it is present in the syncytiotrophoblast (Figure 4). IGFBP2 has been shown to be expressed in the placental villi via qPCR and western blotting[31] but its specific localization is not known. IGFBP6 is expressed in the villous trophoblasts[19]. Upregulation of each of these genes indicates potential changes in the IGF signaling that could alter the development of the placenta. The IGF signaling pathway is associated with a plethora of biological processes essential to placental growth including trophoblast migration, nutrient sensing and transport, metabolism, and proliferation through the activation of the mitogen activated protein kinase (MAPK), extracellular signal regulated (ERK), and phosphoinositide 3-kinase/mammalian target of rapamycin (PI3K/mTOR) pathways which are downstream of the IGF1R[32-34]. Previous studies have demonstrated IGF1 and 2 regulate trophoblast physiology in the developing placenta; therefore, the aberrant upregulation of these specific *IGFBP*s could alter IGF signaling and growth of the placenta via trophoblast differentiation and metabolism[33,34]. While much of the study of IGF signaling has been focused on fetal growth, we only observed two cases of SGA in our isPTB samples, with the remaining cases at or above the 20th percentile for their gestational age, suggesting the alteration in IGF signaling modulating gestational length is not affecting fetal growth. Furthermore, it is important to note that both IGFBP2 and IGFBP6 have roles independent of the IGF signaling including integrin binding modulation through nuclear factor kappa light change enhancer of activated B cells (NFκB) signaling[35], inhibition of angiogenesis[36] and could also be modulating additional biological processes outside the IGF signaling pathway.

In our previous PTB transcriptomics study[15], we identified *IGFBP1* as an upregulated but non-significant gene of interest and PI3K signaling pathway as a significant network potentially predictive of PTB. These data along with the current findings support a role for IGF signaling in isPTB independent of

aberrant fetal growth. IGFBP1 has been further linked to preterm birth as a marker of cervical ripening through analysis of vaginal fluids[37,38]. While these studies focused on the role of IGFBP1 in infection with premature rupture of membranes (PROM) or fetal growth syndromes suggesting a role for circulating decidual derived IGFBP1, the apparent upregulation of *IGFBP1* in the syncytiotrophoblast as observed in this study (Figure 4) suggests origin of circulating IGFBPs are not limited to the decidua and their role in isPTB requires further investigation.

Further Gene Ontology (GO) analyses of the isPTB candidate genes revealed enrichment for cellular components including the T-cell receptor complex, immunological synapse and various membrane associated complexes. GO enrichment for biological processes included programmed cell death, T-cell activation, and cellular defense response among others. Taken together these data suggest an enrichment for an immunological trigger of apoptosis or autophagy; however, it is unclear if this is trophoblastic in origin or a result of alterations in the villous stroma. We did not identify any enrichment for GO terms or pathways in the maturity, we did identify cellular components of interest in the gestational age candidates including several members of the WTN signaling subpathways along with cell proliferation and inflammatory response signals all of which require further analyses for their role in birth timing. We also performed functional classifications for biological processes and pathways to determine differences and similarities between the different candidate classifications. Interestingly, we observed the candidates shared numerous biological pathways, but differed in three specific areas, locomotion, metabolic, and immune processes.

In summary, this is the first study to associate placental morphological phenotypes and genome-wide transcriptome signatures in isPTB. Our work has shown not only a role for the placenta in isPTB, but that accelerated placental maturity directly affects pregnancy outcomes. The similarities of the isPTB and term placentas on both the morphological and molecular levels suggests a precocious maturation phenotype in isPTB, further demonstrating the need to precisely identify the subclassifications of spontaneous preterm birth. Identifying the molecular signatures related to the pathophysiological

differences and similarities between isPTB and term placentas will allow us to gain insight birth timing and potentially develop meaningful clinical therapeutic interventions.

Materials and Methods

Study Population

This study was approved by the Cincinnati Children's Hospital Medical Center institutional review board (#IRB 2013-2243, 2015-8030, 2016-2033). De-identified term (n=9) and idiopathic spontaneous preterm (n=8) placental villous samples and patient information were obtained from the following sources: The Global Alliance to Prevent Prematurity and Stillbirth (GAPPS) in Seattle Washington USA, the Research Centre for Women's and Infant's Health (RCWIH) at Mt Sinai Hospital Toronto Canada, and the University of Cincinnati Medical Center (UCMC). Inclusion criteria included: maternal age 18 years or older, singleton pregnancies with either normal term delivery (38-42 weeks gestation) or preterm delivery (29-36 weeks gestation) without additional complications other than idiopathic spontaneous preterm. Utilizing an RNA sequencing power calculation from[39], it was determined that 10-15 transcriptomes for this particular study. Thus, previously published RNA sequence based, placental villous transcriptomes (GEO GSE73714 term=5, isPTB=5, inter-amniotic infection=5) were also utilized[14]. Birth weight percentiles were estimated using the WHO weight percentiles calculator[40] with an US average birth weight of 3400gm[41].

Transcriptome Generation

All placental samples from GAPPS, RCWIH, and UCMC which were collected within 60 minutes of delivery and snap frozen prior to biobanking. Total RNA was prepared from placental villous samples thawed in RNAIce Later (Invitrogen) as per manufacturer instructions. Total RNA was isolated using the RNAeasy Mini Kit (Qiagen). 50-100 µg of total RNA was submitted to the University of Cincinnati Genomics, Epigenomics and Sequencing Core for RNA quality assessment and sequencing. Long RNA

total libraries were generated using a RiboZero Kit (Illumina) and sequencing was run on a Illumina High Seq 2100 system to generate single end 50bp reads at a depth of 50 million reads. Details on the collection of placental samples and generation of transcriptomes from GSE73714 can be found here[14].

RNA-sequence Analyses

To facilitate RNA sequence analyses, a local instance of the Galaxy[42] was utilized with the following tools: FASTQC (Galaxy v0.71)[43], TrimGalore! (Galaxy v0.4.3.1)[44], Bowtie2 (Galaxy v2.3.4.1)[45], and FeatureCounts (Galaxy v1.6.0.3)[46]. The quality of the raw fastq files was assessed with FASTQC and with adapters subsequently trimmed with TrimGalore. Trimmed sequences were then aligned to the University of California Santa Cruz (UCSC) human genome hg38 using Bowtie2. FeatureCounts was used to generate total read counts per gene and to generate a count matrix file to be used in differential gene expression analyses.

Differential Expression Analyses

After annotation, all non-coding transcripts were removed from the count matrix. Counts per gene were calculated and genes with less than 70 counts total across all samples were removed. The count data was then normalized using the counts per million (CPM) method to allow for various quality control analyses to ensure the data was ready for differential expression testing (Supplemental Figure S1). Differential expression tests were conducted in EdgeR (Emperical Analyses of Digital Gene Expression in R)[47] on only protein coding genes. Within EdgeR, data were normalized using TMM (trimmed means of m values)[48] to account for differences in library sizes. Comparisons for differential expression testing were as follows: IAI compared to term births, isPTB compared to term births, and isPTB compared to IAI births. Multiple corrections testing was performed using the Benjamini Hochberg method with a Q value of <0.05. Venny v2.0[49] was utilized to generate Venn diagrams and identify candidate genes for maturation, gestational age, and isPTB. Heatmaps were generated in Prism v7 (GraphPad).

348

349 *Pathway and Gene ontology Analyses*

350 Significant genes were divided into upregulated and downregulated categories then entered into the
351 Panther Pathway DB[50] for statistical overrepresentation analyses for Reactome Pathways and gene
352 ontology (GO). Fisher's Exact tests were used to determine significance and Bonferroni correction for
353 multiple comparisons. Pathways were considered significant if they had an adjusted p-value <0.05 and
354 enrichment score of >4.

355

356 *Histology and Immunohistochemistry (IHC)*

357 Serial sections were stained with Hematoxylin and Eosin (H&E) and assessed for placental maturity as
358 described below. Immunohistochemistry was performed as previously published[9]. Briefly, all slides were
359 incubated 95°C target retrieval solution for 30 minutes then washed in deionized water. Slides were then
360 incubated in 3% hydrogen peroxide for 10 minutes followed by blocking in 10% normal goat serum +1%
361 bovine albumin for 60 minutes. Primary antibodies generated in rabbit sera were diluted in phosphate
362 buffered saline (PBS): DKK1 (1:20, GeneTex GTX40056) and IGFBP1 (1: 50, GeneTex GTX31149).
363 Slides were incubated overnight at 4°C washed and incubated with biotinylated secondary antibody (anti-
364 rabbit) for 60 minutes. Antibody binding was detected using DAB and slides were counterstained with
365 hematoxylin. All slides were imaged on a Nikon Eclipse 80i microscope.

366

367 *Clinical definitions.* Gestational age was established based on last menstrual period confirmed by an
368 ultrasonographic examination prior to 20 weeks [51]. IAI was established based on analysis of the amniotic
369 fluid retrieved in sterile conditions by trans-abdominal amniocentesis. Amniotic fluid infection was
370 established by a positive Gram stain or a positive microbial culture result[52]. Preterm birth was defined as
371 delivery of the neonate <37 weeks GA[53]. Idiopathic preterm birth was established absent IAI and
372 histologic inflammation of the placenta and fetal membranes as assessed by a clinical pathologist[54].

373

374 *Morphological Assessment of Placental Maturity*

375 Syncytial nuclear aggregates (SNA) were defined as a multi-layered aggregate of at least 10 syncytial
376 nuclei extending out from the villous surface but not in contact with other villi [9]. Terminal villi were
377 defined as branched villi <80μm in diameter. Vasculosyncytial membranes are defined in[11]. Fibrin
378 deposition was quantified by a score of 0-3 with where 0 was no fibrin observed and where 3 was the
379 majority of the field containing fibrin. SNAs, terminal villi, and vasculo-syncytial membranes were
380 counted manually and fibrin deposition was scored in 20 high powered fields (hpf) by two blinded
381 reviewers.

382

383 *Statistical Analyses*

384 Data were analyzed in Prism7.0 (GraphPad). Data were evaluated for normality and non-parametric tests
385 applied as appropriate. Non-parametric data are expressed as median and range and were analyzed by
386 Kruskal-Wallis Test ANOVA with Dunn's Multiple Comparisons. Categorical data were analyzed using
387 Fisher's Exact Test.

388

389 **Author Contributions**

390 HMB-Designed and preformed research, analyzed data, and wrote the paper

391 HNJ-Assisted in research design and advised data analyses, edited paper

392 WEA- provided RNA sequence data, advised on research design, edited paper

393 IAB - provided RNA sequence data, advised on research design, edited paper

394 CSB - provided RNA sequence data, advised on research design, edited paper

395 SGK-provided preterm birth samples, edited paper

396 LJM- Assisted in research design and advised data analyses, edited paper

397

Acknowledgements

The authors would like to express their gratitude to the patients who donated their placentas for research and to Pietro Presicce, Paranthaman Senthamarai Kannan, and Manuel Alvarez (Kallapur lab), GAPPS, and RCWIH for assisting us in obtaining the placental samples. Additionally, the authors thank the staff at the UC Genomics Core for their help generating the bulk of the transcriptomics data for this project. This work was supported by grants to LM from the March of Dimes Prematurity Research Center Ohio Collaborative Funding 07/13-6/18, The Bill and Melinda Gates Foundation: Systems Biology Approaches to Birth Timing and Prematurity OPP1113966 Funding 11/14-10/17, and the Eunice Kennedy Shriver National Institute of Child and Health & Human Development of the National Institutes of Health Award Number R01HD091527.

Data Availability

Data will be made available in GEO as per funding requirements upon manuscript acceptance for publication.

Additional information

The authors declare they have no competing interests.

References:

1 Dimes, M. o. March of Dimes Prematurity Report Card 2017. 1-4 (2017).

2 Blencowe, H. *et al.* Born too soon: the global epidemiology of 15 million preterm births. *Reprod Health* **10 Suppl 1**, S2, doi:10.1186/1742-4755-10-S1-S2 (2013).

3 Monangi, N. K., Brockway, H. M., House, M., Zhang, G. & Muglia, L. J. The genetics of preterm birth_ Progress and promise. *Seminars in Perinatology*, 1-10, doi:10.1053/j.semperi.2015.09.005 (2015).

4 Burton, G. J. & Fowden, A. L. The placenta: a multifaceted, transient organ. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **370**, 20140066-20140066, doi:10.1098/rstb.2014.0066 (2015).

5 Mikheev, A. M. *et al.* Profiling gene expression in human placentae of different gestational ages: an OPRU Network and UW SCOR Study. *Reproductive sciences (Thousand Oaks, Calif.)* **15**, 866-877, doi:10.1177/1933719108322425 (2008).

6 Roberts, C. Placental transcriptome co-expression analysis reveals conserved regulatory programs across gestation. *BMC Genomics*, 1-13, doi:10.1186/s12864-016-3384-9 (2016).

7 Sitras, V., Fenton, C., Paulssen, R., Vårtun, Å. & Acharya, G. Differences in gene expression between first and third trimester human placenta: a microarray study. *PLoS ONE* **7**, e33294 (2012).

8 Winn, V. D. *et al.* Gene Expression Profiling of the Human Maternal-Fetal Interface Reveals Dramatic Changes between Midgestation and Term. *Endocrinology* **148**, 1059-1079, doi:10.1210/en.2006-0683 (2007).

9 Jones, H. N. *et al.* Hypoplastic left heart syndrome is associated with structural and vascular placental abnormalities and leptin dysregulation. *Placenta* **36**, 1078-1086, doi:10.1016/j.placenta.2015.08.003 (2015).

10 Loukeris, K., Sela, R. & Baergen, R. N. Syncytial knots as a reflection of placental maturity: reference values for 20 to 40 weeks' gestational age. *Pediatr Dev Pathol* **13**, 305-309, doi:10.2350/09-08-0692-OA.1 (2010).

11 Sankar, K. D., Bhanu, P. S., Kiran, S., Ramakrishna, B. A. & Shanthi, V. Vasculosyncytial membrane in relation to syncytial knots complicates the placenta in preeclampsia: a histomorphometrical study. *Anatomy & Cell Biology* **45**, 86-86, doi:10.5115/acb.2012.45.2.86 (2012).

12 Morgan, T. K. Placental Insufficiency Is a Leading Cause of Preterm Labor. *NeoReviews* **15**, e518-e525, doi:10.1542/neo.15-12-e518 (2014).

13 Morgan, T. K. *et al.* Placental villous hypermaturation is associated with idiopathic preterm birth. *Journal of Maternal-Fetal and Neonatal Medicine* **26**, 647-653, doi:10.3109/14767058.2012.746297 (2013).

14 Ackerman IV, W. E. *et al.* Comprehensive RNA profiling of villous trophoblast and decidua basalis in pregnancies complicated by preterm birth following intra-amniotic infection. *Placenta* **44**, 23-33, doi:10.1016/j.placenta.2016.05.010 (2016).

15 Paquette, A. G., Brockway, H. M., Price, N. D. & Muglia, L. J. Comparative transcriptomic analysis of human placentae at term and preterm delivery. *Biology of Reproduction* **98**, 89-101, doi:10.1093/biolre/iox163 (2017).

16 Leavey, K. *et al.* Gene markers of normal villous maturation and their expression in placentas with maturational pathology. *Placenta* **58**, 52-59, doi:10.1016/j.placenta.2017.08.005 (2017).

17 Jacob, J. T., Coulombe, P. A., Kwan, R. & Omary, M. B. Types I and II Keratin Intermediate Filaments. *Cold Spring Harbor Perspectives in Biology* **10**, a018275-018217, doi:10.1101/cshperspect.a018275 (2018).

18 Camden, A. J. *et al.* Growth regulation by estrogen in breast cancer 1 (GREB1) is a novel progesterone-responsive gene required for human endometrial stromal decidualization. *MHR: Basic science of reproductive medicine* **23**, 646-653, doi:10.1093/molehr/gax045 (2017).

19 Uhlen, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419-1260419, doi:10.1126/science.1260419 (2015).

20 Pei, J. & Grishin, N. V. Unexpected diversity in Shisa-like proteins suggests the importance of their roles as transmembrane adaptors. *Cellular Signalling* **24**, 758-769, doi:10.1016/j.cellsig.2011.11.011 (2012).

21 Liu, Z., Wang, C., Liu, X. & Kuang, S. Shisa2 regulates the fusion of muscle progenitors. *Stem Cell Research* **31**, 31-41, doi:10.1016/j.scr.2018.07.004 (2018).

22 Marceau, G., Gallot, D., Lémery, D. & Sapin, V. Vol. 75 97-115 (Elsevier, 2007).

23 Das, B. C. *et al.* Retinoic acid signaling pathways in development and diseases. *Bioorganic & Medicinal Chemistry* **22**, 673-683, doi:10.1016/j.bmc.2013.11.025 (2014).

24 David, A. *et al.* BAD-LAMP defines a subset of early endocytic organelles in subpopulations of cortical projection neurons. *Journal of Cell Science* **120**, 353-365, doi:10.1242/jcs.03316 (2007).

25 Schwake, M., Schröder, B. & Saftig, P. Lysosomal Membrane Proteins and Their Central Role in Physiology. *Traffic* **14**, 739-748, doi:10.1111/tra.12056 (2013).

26 Pellegrini, C. *et al.* The expression of estrogen receptors as well as GREB1, c-MYC, and cyclin D1, estrogen-regulated genes implicated in proliferation, is increased in peritoneal endometriosis. *Fertil Steril* **98**, 1200-1208, doi:10.1016/j.fertnstert.2012.06.056 (2012).

27 Meinhardt, G. *et al.* Wingless ligand 5a is a critical regulator of placental growth and survival. *Scientific Reports*, 1-14, doi:10.1038/srep28127 (2016).

28 Zhang, Z., Li, H., Zhang, L., Jia, L. & Wang, P. Differential expression of beta-catenin and dickkopf-1 in the third trimester placentas from normal and preeclamptic pregnancies: a comparative study. *Reproductive Biology and Endocrinology* **11**, 17, doi:10.1186/1477-7827-11-17 (2013).

29 Wang, X. *et al.* Wnt/β-catenin signaling pathway in severe preeclampsia. *Journal of Molecular Histology* **49**, 317-327, doi:10.1007/s10735-018-9770-7 (2018).

30 Firth, S. M. & Baxter, R. C. Cellular Actions of the Insulin-Like Growth Factor Binding Proteins. *Endocrine Reviews* **23**, 824-854, doi:10.1210/er.2001-0033 (2002).

31 Nawathe, A. R. *et al.* Insulin-like growth factor axis in pregnancies affected by fetal growth disorders. *Clinical Epigenetics*, 1-13, doi:10.1186/s13148-016-0178-5 (2016).

32 Forbes, K. & Westwood, M. The IGF Axis and Placental Function. *Hormone Research in Paediatrics* **69**, 129-137, doi:10.1159/000112585 (2008).

33 Forbes, K., Westwood, M., Baker, P. N. & Aplin, J. D. Insulin-like growth factor I and II regulate the life cycle of trophoblast in the developing human placenta. *American Journal of Physiology-Cell Physiology* **294**, C1313-C1322, doi:10.1152/ajpcell.00035.2008 (2008).

34 Hiden, U., Glitzner, E., Hartmann, M. & Desoye, G. Insulin and the IGF system in the human placenta of normal and diabetic pregnancies. *Journal of Anatomy* **215**, 60-68, doi:10.1111/j.1469-7580.2008.01035.x (2009).

35 Holmes, K. M. *et al.* Insulin-like growth factor-binding protein 2-driven glioma progression is prevented by blocking a clinically significant integrin, integrin-linked kinase, and NF-κB network. *Proceedings of the National Academy of Sciences* **109**, 3475-3480, doi:10.1073/pnas.1120375109 (2012).

36 Bach, L. A. Recent insights into the actions of IGFBP-6. *Journal of Cell Communication and Signaling* **9**, 189-200, doi:10.1007/s12079-015-0288-4 (2015).

37 Kallioniemi, H., Rahkonen, L., Heikinheimo, O. & Paavonen, J. Early pregnancy vaginal fluid phosphorylated insulin-like growth factor binding protein-1 predicts preterm delivery. *Prenatal Diagnosis* **362**, 1-6, doi:10.1002/pd.4072 (2013).

38 Mešić Đogić, L., Mićić, D., Omeragić, F., Kovač, R. & Fazlagić, S. IGFBP-1 marker of cervical ripening and predictor of preterm birth. *Medicinski glasnik : official publication of the Medical Association of Zenica-Doboj Canton, Bosnia and Herzegovina* **13**, 118-124, doi:10.17392/856-16 (2016).

39 Hart, S. N., Therneau, T. M., Zhang, Y., Poland, G. A. & Kocher, J.-P. Calculating Sample Size Estimates for RNA Sequencing Data. *Journal of Computational Biology* **20**, 970-978, doi:10.1089/cmb.2012.0283 (2013).

40 Hadlock, F. P., Harrist, R. B. & Martinez-Poyer, J. In utero analysis of fetal growth: a sonographic weight standard. *Radiology* **181**, 129-133, doi:10.1148/radiology.181.1.1887021 (1991).

41 Gardosi, J., Francis, A., Turner, S. & Williams, M. Customized growth charts: rationale, validation and clinical benefits. *Am J Obstet Gynecol* **218**, S609-s618, doi:10.1016/j.ajog.2017.12.011 (2018).

42 Boekel, J. *et al.* Multi-omic data analysis using Galaxy. *Nature Publishing Group* **33**, 137-139, doi:10.1038/nbt.3134 (2015).

43 Andrews, S. *FASTQC; a quality control tool for highthroughput sequence data,* http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (2010).

44 Kruger, F. *TrimGalore! A tool to automate quality and adapter trimming for highthroughput sequencing*, http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/ (2012).

45 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357, doi:10.1038/nmeth.1923 (2012).

46 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930, doi:10.1093/bioinformatics/btt656 (2014).

47 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2009).

48 Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* **11**, doi:10.1186/gb-2010-11-3-r25 (2010).

49 Oliveros, J. *Venny. An interactive tool for comparing lists with Venn's diagrams.*, http://bioinfogp.cnb.csic.es/tools/venny/index.html (2007).

50 Mi, H. & Thomas, P. Vol. 563 123-140 (Humana Press, 2009).

51 Committe on Obstetric Practice American Institute of Ultrasound in Medicine society for Maternal-Fetal, M. in *Obstetrics and gynecology* Vol. 129 967-968 (2017).

52 Buhimschi, C. S. *et al.* Proteomic profiling of the amniotic fluid to detect inflammation, infection, and neonatal sepsis. *PLoS Medicine* **4**, 84-94, doi:papers3://publication/doi/10.1371/journal.pmed.0040018 (2007).

53 Committee on Practice Bulletins—Obstetrics, T. A. C. o. O. & Gynecologists. in *Obstetrics and gynecology* Vol. 120 964-973 (2012).

54 Buhimschi, I. A. *et al.* Multidimensional Proteomics Analysis of Amniotic Fluid to Provide Insight into the Mechanisms of Idiopathic Preterm Birth. *PLOS ONE* **3**, e2049-2011, doi:papers3://publication/doi/10.1371/journal.pone.0002049 (2008).

558 **Figure Legends**

559 **Figure 1: Normal and advanced villous maturation.** A-C) Representative micrographs of normal
560 mature placental villi at term delivery (39.2-42.2 weeks), D-E) Representative micrographs of placental
561 villi from intraamniotic infected placentas delivered (25.4-31.3 weeks). Perivillous fibrin deposition (blue
562 arrows) was observed along with an increase in syncytial nuclear aggregates (green arrows). F-H)
563 Representative micrographs of advanced villous maturation in idiopathic spontaneous preterm deliveries
564 (36.6-36.1 weeks). Terminal villi were numerous along with an increase in the number of syncytial
565 aggregates and vasculosyncytial membranes. Perivillous fibrin deposition was also observed. All images
566 were captured at 20X magnification and scale bar = 100μm.

567

568 **Figure 2: A comparative approach to identifying molecular signatures of maturation and gestation.**
569 Differentially expressed genes were identified using pairwise comparisons EdgeR. Red dots represent
570 significant genes that have an absolute log2 fold-change of 1.5 and Benjamini Hochberg adjusted P-value
571 of <0.1. Blue lines represent log2 fold-change of 1.5 and dotted line represents a raw P-value <0.05. B) A
572 comparative analysis to identify genes categorized as gestational age, infection, and maturation. The Venn
573 diagram represents the intersection of significant genes from Panel A pairwise comparisons IAI vs term
574 n=160 total genes, isPTB vs IAI n=94 total genes, and isPTB vs term n=158 total genes. Genes categorized
575 as infection (n=37) are represented in the intersection of IAI vs term and isPTB vs term. Gestational age
576 genes (total n= 123) are represented by the intersection of IAI vs term and isPTB vs term (n=11) and the
577 IAI vs term genes exclusive to IAI vs term (blue shaded area). Villous Maturation genes (total n= 186)
578 are represented by the intersection of isPTB vs IAI (n=18) and the genes exclusive to isPTB v IAI shaded
579 in yellow (n=39) and isPTB vs term shaded in green (n=158). C) Identification of candidate genes for
580 maturation and gestational age. We compared genes with significant differential expression across all
581 three pairwise comparisons to identify candidate genes with differential expression in isPTB and term
582 compared to IAI but with no change or the opposite expression pattern in isPTB compared to term

583 (maturation signal) or with a minimal expression difference between isPTB compared to term (maturation
584 drivers). Gestational age candidates were identified by comparing isPTB and IAI samples to term.
585 Differences observed in isPTB compared to IAI are likely due to the overall differences in gestational age
586 (Table 1).
587
588 **Figure 3: Identification of isPTB specific genes.** We compared genes with significant differential
589 expression across all three pairwise comparisons to identify candidate genes with differential expression
590 in isPTB compared to term and IAI and with the either the opposite expression pattern or no change in
591 expression in term v IAI. 104 genes met this criteria and are shown in heatmaps.
592
593 **Figure 4: Placental tissue in isPTB demonstrates increased DKK1 and IGFBP1 expression.** DKK1
594 and IGFBP demonstrate localization the syncytiotrophoblast in the control term births with increased
595 expression in isPTB. Images are taken at 40x magnification and scale bar = 50μm
596
597
598
599
600
601
602
603
604
605
606
607

608

Table 1: Clinical characteristics of the placental villous samples included in this study

Characteristics	Intra-Amniotic Inection Births (IAI Births)	Idiopathic Spontaneous Preterm Births (isPTB)	Term Births	P-values
Number of samples	5	12	14	
Maternal Age	35(21-41)	28(18-39)	31 (19-37)	NS[1]
Gravidity	3(2-5)	2(1-5)	2(1-3)	0.058[1]
Parity	1(0-4)	1(0-2)	2(0-5)	NS[1]
Gestational Age	26(25-31)**	33(29-36)*	39(38-42)	<0.0001[1]
Fetal sex (% female)	5(100%)	6(50%)	7(50%)	NS[2]
Fetal weight (grams)	830(680-1830)**	2062(1450-2722)*	3505 (3360-4690)	<0.0001[1]
Birth weight percentile	44(1-60)	33(3-60)*	72(40-99)	0.0008[1]
SGA %	20%	16.6%	0	
Delivery type				
Cesarean (%)	2(40%)	5(41.7%)	7(50%)	NS[2]
Vaginal (%)	3(60%)	7(58.3%)	7(50%)	
Infection Status (% Positive)	5(100%)*	0(0%)	1(7%)	<0.005[2]

[1]Kruskal-Wallis Test ANOVA with Dunn's Multiple Comparisons [2]Fisher's Exact Test
*P<0.005 **P<0.0001 vs Term Births Abbreviations: SGA Small for gestational age, NS Not significant

Table 2: Morphological assessment of placental villous samples

	IAI Births(n=5)	isPTB (n=11)	Term Births (n=14)	P-value
SNA/pf (range)	0.7(0.4-1.15)	1.03(0.4-2.45)	1.29(0.25-2.2)	0.093
Total termial Villi <80um (range)	58(46-96)	103(57-126)	102(44-157)	0.078
Vasculo-syncytial membranes/hpf(range)	0.59(0.1-1.35)*	2.02(0.45-5.1)	3.86(1.45-10)	0.0010

Kruskal-Wallis Test ANOVA with Dunn's Multiple Comparisons *P<0.005
Abbreviations: isPTB idiopathic spontaneous preterm birth

Table 3: Candidate genes associated with advanced villous maturity in isPTB

Gene symbol	Description	Log2 Fold Change[1]	Adjusted P-value*
	Signal Genes		
KRT24	keratin 24	10.38	0.0891
KRT6A	keratin 6A	7.76	0.0931
ZBED2	zinc finger BED-type containing 2	5.63	0.0850
AVPR1A	arginine vasopressin receptor 1A	4.91	0.0977
MPZL2	myelin protein zero like 2	3.69	0.0677
TEKT2	tektin 2	3.44	0.0783
CDH2	cadherin 2	3.15	0.0505
OR2T10	olfactory receptor family 2 subfamily T member 10	2.63	0.0846
HTRA4	HtrA serine peptidase 4	2.24	0.0432
S100A2	S100 calcium binding protein A2	2.12	0.0704
SLC5A9	solute carrier family 5 member 9	-1.54	0.0934
HPSE2	heparanase 2 (inactive)	-1.57	0.0353
PGLYRP1	peptidoglycan recognition protein 1	-1.62	0.0575
ZNF560	zinc finger protein 560	-1.70	0.0633
C2orf40	chromosome 2 open reading frame 40	-1.73	0.0850
HTR1B	5-hydroxytryptamine receptor 1B	-1.75	0.0018
GPR12	G protein-coupled receptor 12	-1.87	0.0926
AMELX	amelogenin, X-linked	-1.91	0.0989
SCG2	secretogranin II	-1.99	0.0353
DNAAF3	dynein axonemal assembly factor 3	-2.04	0.0894
CRABP1	cellular retinoic acid binding protein 1	-2.12	0.0999
	Driver genes		
SHISAL1	shisa like 1	5.20	0.0856
FGFBP1	fibroblast growth factor binding protein 1	5.05	0.0821
ICOS	inducible T cell costimulator	4.61	0.0835
CYP7B1	cytochrome P450 family 7 subfamily B member 1	4.59	0.0298
HOXD10	homeobox D10	4.58	0.0516
SLC16A9	solute carrier family 16 member 9	3.54	0.0977
TNFRSF18	TNF receptor superfamily member 18	3.32	0.0935
MGST1	microsomal glutathione S-transferase 1	2.80	0.0374
FOSL1	FOS like 1, AP-1 transcription factor subunit	2.77	0.0902
SPOCK1	SPARC/osteonectin, cwcv and kazal like domains proteoglycan 1	2.38	0.0483
KHDC1	KH domain containing 1	-1.51	0.0298
CYP4F12	cytochrome P450 family 4 subfamily F member 12	-1.79	0.0560
ANKRD30A	ankyrin repeat domain 30A	-2.16	0.0435
LAMP5	lysosomal associated membrane protein family member 5	-4.49	0.0516

[1] Values from the isPTB verses IAI comparison *Benjamini Hochberg correction for multiple comparisons

Table 4: Candidate genes specific to isPTB

Gene symbol	Description	Log2 Fold Change[1]	Adjusted P-value[*]
PAEP	progestagen associated endometrial protein	7.28	0.00001
IGFBP1	insulin like growth factor binding protein 1	6.17	0.00006
RORB	RAR related orphan receptor B	6.15	0.00001
NDP	NDP, norrin cystine knot growth factor	5.80	0.00001
CHRDL1	chordin like 1	5.75	0.00013
PRL	prolactin	5.68	0.00003
SCGB1D2	secretoglobin family 1D member 2	5.66	0.03211
JCHAIN	joining chain of multimeric IgA and IgM	5.18	0.00067
EPYC	epiphycan	4.88	0.00143
GSTA1	glutathione S-transferase alpha 1	4.65	0.02587
LBP	lipopolysaccharide binding protein	4.63	0.00787
HEPH	hephaestin	4.62	0.00012
UGT2B7	UDP glucuronosyltransferase family 2 member B7	4.49	0.04936
FOXL2NB	FOXL2 neighbor	4.16	0.00582
PIGR	polymeric immunoglobulin receptor	4.12	0.02485
FGB	fibrinogen beta chain	4.06	0.04506
RXFP1	relaxin/insulin like family peptide receptor 1	4.02	0.00010
FAM19A4	family with sequence similarity 19 member A4, C-C motif chemokine like	3.97	0.00582
DIRAS2	DIRAS family GTPase 2	3.93	0.01842
RBP4	retinol binding protein 4	3.92	0.00017
GNLY	granulysin	3.91	0.00004
SCARA5	scavenger receptor class A member 5	3.90	0.00006
MEDAG	mesenteric estrogen dependent adipogenesis	3.87	0.00004
CXCL9	C-X-C motif chemokine ligand 9	3.78	0.04135
PLA2G2D	phospholipase A2 group IID	3.72	0.07457
PDZK1IP1	PDZK1 interacting protein 1	3.62	0.00087
LRRN4CL	LRRN4 C-terminal like	3.60	0.00045
ITGAD	integrin subunit alpha D	3.46	0.00036
CNR1	cannabinoid receptor 1	3.43	0.00086
IGFBP6	insulin like growth factor binding protein 6	3.40	0.00522
WT1	Wilms tumor 1	3.37	0.00143
CSF3	colony stimulating factor 3	3.35	0.01058
PRUNE2	prune homolog 2	3.23	0.00080
CDHR1	cadherin related family member 1	3.19	0.03039
RELN	reelin	3.12	0.01473
FAM19A5	family with sequence similarity 19 member A5, C-C motif chemokine like	3.02	0.07215
DRAXIN	dorsal inhibitory axon guidance protein	2.96	0.07852

CD2	CD2 molecule	2.94	0.01984
BRINP2	BMP/retinoic acid inducible neural specific 2	2.93	0.00820
KLRC1	killer cell lectin like receptor C1	2.93	0.00582
CLIC6	chloride intracellular channel 6	2.87	0.09012
NTRK1	neurotrophic receptor tyrosine kinase 1	2.87	0.01134
KCND2	potassium voltage-gated channel subfamily D member 2	2.84	0.03211
SLPI	secretory leukocyte peptidase inhibitor	2.82	0.00396
RASD1	ras related dexamethasone induced 1	2.71	0.00448
GZMB	granzyme B	2.61	0.00287
MMP1	matrix metallopeptidase 1	2.61	0.05125
MAG	myelin associated glycoprotein	2.60	0.09009
NCAM1	neural cell adhesion molecule 1	2.58	0.00889
WDR72	WD repeat domain 72	2.57	0.04582
MZB1	marginal zone B and B1 cell specific protein	2.56	0.05713
POU2AF1	POU class 2 associating factor 1	2.56	0.04936
EOMES	eomesodermin	2.55	0.01473
SLC18A2	solute carrier family 18 member A2	2.54	0.03211
BHMT2	betaine--homocysteine S-methyltransferase 2	2.54	0.00787
CP	ceruloplasmin	2.54	0.00582
ADCY1	adenylate cyclase 1	2.54	0.00186
GPR174	G protein-coupled receptor 174	2.50	0.01830
GZMA	granzyme A	2.49	0.05337
IRF4	interferon regulatory factor 4	2.48	0.01734
EPDR1	ependymin related 1	2.48	0.00068
CD8A	CD8a molecule	2.46	0.01699
GAS1	growth arrest specific 1	2.44	0.00004
COL4A4	collagen type IV alpha 4 chain	2.42	0.02472
TRPC4	transient receptor potential cation channel subfamily C member 4	2.38	0.01393
GZMH	granzyme H	2.37	0.03028
CD38	CD38 molecule	2.36	0.00251
DUSP2	dual specificity phosphatase 2	2.35	0.01545
SH2D1A	SH2 domain containing 1A	2.33	0.09336
AADAC	arylacetamide deacetylase	2.32	0.04506
RAMP1	receptor activity modifying protein 1	2.32	0.00010
LEFTY2	left-right determination factor 2	2.31	0.05736
GLB1L2	galactosidase beta 1 like 2	2.28	0.00881
LSAMP	limbic system-associated membrane protein	2.28	0.04837
GBP5	guanylate binding protein 5	2.27	0.01393
TNFRSF8	TNF receptor superfamily member 8	2.24	0.00522
PCBP3	poly(rC) binding protein 3	2.24	0.01545
OMD	osteomodulin	2.23	0.00186
DEFB1	defensin beta 1	2.22	0.04506
FASLG	Fas ligand	2.21	0.07457

ABLIM2	actin binding LIM protein family member 2	2.20	0.01021
ARC	activity regulated cytoskeleton associated protein	2.20	0.06836
SPTBN4	spectrin beta, non-erythrocytic 4	2.17	0.02472
SMOX	spermine oxidase	2.16	0.00068
DNASE1L3	deoxyribonuclease 1 like 3	2.16	0.05553
MAOB	monoamine oxidase B	2.15	0.01065
ITK	IL2 inducible T cell kinase	2.14	0.01858
PLIN1	perilipin 1	2.12	0.01545
CSDC2	cold shock domain containing C2	2.12	0.02278
CD248	CD248 molecule	2.12	0.01163
TMEM59L	transmembrane protein 59 like	2.12	0.03746
SLFN12L	schlafen family member 12 like	2.11	0.02239
C3	complement C3	2.11	0.00582
F2R	coagulation factor II thrombin receptor	2.10	0.00010
CD3G	CD3g molecule	2.08	0.05553
UCHL1	ubiquitin C-terminal hydrolase L1	2.07	0.02188
LCK	LCK proto-oncogene, Src family tyrosine kinase	2.05	0.00704
ABI3BP	ABI family member 3 binding protein	2.05	0.00523
OXTR	oxytocin receptor	2.04	0.00134
C11orf96	chromosome 11 open reading frame 96	2.02	0.00522
TNFRSF9	TNF receptor superfamily member 9	2.01	0.08424
PZP	PZP, alpha-2-macroglobulin like	2.01	0.07355
SAMD3	sterile alpha motif domain containing 3	2.01	0.02241
HSPB6	heat shock protein family B (small) member 6	1.99	0.00437
IKZF3	IKAROS family zinc finger 3	1.99	0.00228
CTSW	cathepsin W	1.99	0.04227
ISLR	immunoglobulin superfamily containing leucine rich repeat	1.98	0.00086
KCNB1	potassium voltage-gated channel subfamily B member 1	1.97	0.01593
SAXO2	stabilizer of axonemal microtubules 2	1.96	0.03543
EGR3	early growth response 3	1.93	0.01977
RGS1	regulator of G protein signaling 1	1.93	0.01858
THBS2	thrombospondin 2	1.92	0.08708
ALDH1A2	aldehyde dehydrogenase 1 family member A2	1.92	0.00221
SLC47A1	solute carrier family 47 member 1	1.92	0.02583
SH2D2A	SH2 domain containing 2A	1.87	0.04419
PTGIS	prostaglandin I2 synthase	1.84	0.07949
MAATS1	MYCBP associated and testis expressed 1	1.84	0.00413
SCML4	Scm polycomb group protein like 4	1.83	0.04554
HSD11B1	hydroxysteroid 11-beta dehydrogenase 1	1.81	0.02188
KCNH1	potassium voltage-gated channel subfamily H member 1	1.80	0.05341
CHST2	carbohydrate sulfotransferase 2	1.75	0.00522
IGFBP2	insulin like growth factor binding protein 2	1.74	0.02723
KCNQ3	potassium voltage-gated channel subfamily Q member 3	1.74	0.01473

RUNX3	runt related transcription factor 3	1.73	0.03062
TNC	tenascin C	1.71	0.05109
PRDM1	PR/SET domain 1	1.71	0.00582
ZAP70	zeta chain of T cell receptor associated protein kinase 70	1.70	0.05856
ADRA2C	adrenoceptor alpha 2C	1.70	0.04135
CD7	CD7 molecule	1.68	0.04030
SSC5D	scavenger receptor cysteine rich family member with 5 domains	1.67	0.05461
FNDC4	fibronectin type III domain containing 4	1.67	0.00303
EHF	ETS homologous factor	1.66	0.01964
SPHK1	sphingosine kinase 1	1.64	0.02634
ESPNL	espin like	1.63	0.09012
CPXM1	carboxypeptidase X, M14 family member 1	1.59	0.02587
LIF	LIF, interleukin 6 family cytokine	1.57	0.09012
ATOH8	atonal bHLH transcription factor 8	1.53	0.02543
FHOD3	formin homology 2 domain containing 3	1.53	0.06551
ERICH5	glutamate rich 5	-1.53	0.08983
GABRB1	gamma-aminobutyric acid type A receptor beta1 subunit	-1.60	0.00921

[1] Values from the isPTB verses term comparison *Benjamini Hochberg correction for multiple comparisons

Table 5: Enrichment analyses for isPTB candidate genes

Reactome pathways	Fold Enrichment	Raw P-value	Adjusted P-value[1]	Genes
Regulation of Insulin-like Growth Factor (IGF) transport and uptake by Insulin-like Growth Factor Binding Proteins (IGFBPs) (R-HSA-381426)	35.28	7.41E-07	1.47E-03	*MMP1, IGFBP6, IGFBP1, IGFBP2, GZMH*
GO cellular component complete	**Fold Enrichment**	**Raw P-value**	**Adjusted P-value[1]**	**Genes**
T cell receptor complex (GO:0042101)	23.4	4.14E-04	4.33E-02	*CD8A, ZAP70,CD3G*
immunological synapse (GO:0001772)	17.43	1.20E-04	1.70E-02	*LCK, GZMB, GZMA, ZAP70*
plasma membrane raft (GO:0044853)	8.8	8.46E-05	1.29E-02	*TRPC4,PTGIS, CD8A,FASLG, F2R, KCND2*
membrane microdomain (GO:0098857)	5.83	1.57E-06	5.21E-04	*LCK, TRPC4, PTGIS, CD8A, CNR1, ARC, FASLG, ZAP70, MAG, F2R, ADCY1, KCND2*
membrane raft (GO:0045121)	5.83	1.57E-06	4.46E-04	*LCK, TRPC4, PTGIS, CD8A, CNR1, ARC, FASLG, ZAP70, MAG, F2R, ADCY1, KCND2*
membrane region (GO:0098589)	5.63	2.24E-06	5.56E-04	*LCK, TRPC4, PTGIS, CD8A, CNR1, ARC, FASLG, ZAP70, MAG, F2R, ADCY1, KCND2*
extracellular matrix (GO:0031012)	5.08	6.14E-06	1.11E-03	*LEFTY2, CD248, OMD, NCAM1, THBS2, NDP, MMP1, TNC, SLPI, SSC5D, ABI3BP, COL4A4*

GO molecular function complete	Fold Enrichment	Raw P-value	Adjusted P-value[1]	Genes
signaling receptor binding (GO:0005102)	2.46	8.44E-06	3.93E-02	*LEFTY2,CSF3,DEFB1,LCK,LBP,CLIC6,FGB,NDP,GABRB1,CD8A,CD2,CXCL9,LIF,PRL,IGFBP6, C3,IGFBP1,ADRA2C, NTRK1,FASLG,IGFBP2, ITK,MAG,RELN,CD3G, UCHL1,F2R,JCHAIN*

GO biological process complete	Fold Enrichment	Raw P-value	Adjusted P-value[1]	Genes
programmed cell death involved in cell development (GO:0010623)	49.39	6.21E-05	1.86E-02	*DNASE1L3, NTRK1, FASLG*

regulation of norepinephrine secretion (GO:0014061)	26.15	3.09E-04	4.83E-02	*OTXR, KCNB1, ADRA2C*
positive regulation of vasoconstriction (GO:0045907)	22.45	5.30E-06	2.85E-03	*CD38, FGB, OXTR, ADRA2C, F2R*
alpha-beta T cell differentiation (GO:0046632)	15.12	3.06E-05	1.09E-02	*ITK, ZAP70, EOMES, IRF4, PLA2G2D*
antibacterial humoral response (GO:0019731)	14.82	2.13E-04	3.83E-02	*DEFB1, FGB, SLPI, JCHAIN*
cellular defense response (GO:0006968)	12.77	6.48E-05	1.91E-02	*GNLY, LBP, CXCL9, SH2D1A, ITK*
alpha-beta T cell activation (GO:0046631)	12.56	7.00E-05	1.92E-02	*ITK, ZAP70, EOMES, IRF4, PLA2G2D*
positive regulation of blood circulation (GO:1903524)	10.58	1.49E-04	3.29E-02	*CD38, FGB, OTXR, ARDA2C, F2R*
regulation of vasoconstriction (GO:0019229)	10.44	1.59E-04	3.36E-02	*CD38, FGB, OXTR, ADRA2C, F2R*
response to retinoic acid (GO:0032526)	9.69	1.19E-05	5.30E-03	*RBP4, RORB, BRINP2, CD38, TNC, IGFBP2, ALKDH1A2*

[1]Fishers Exact test with Bonferroni correction for multiple comparisons

Figure 1

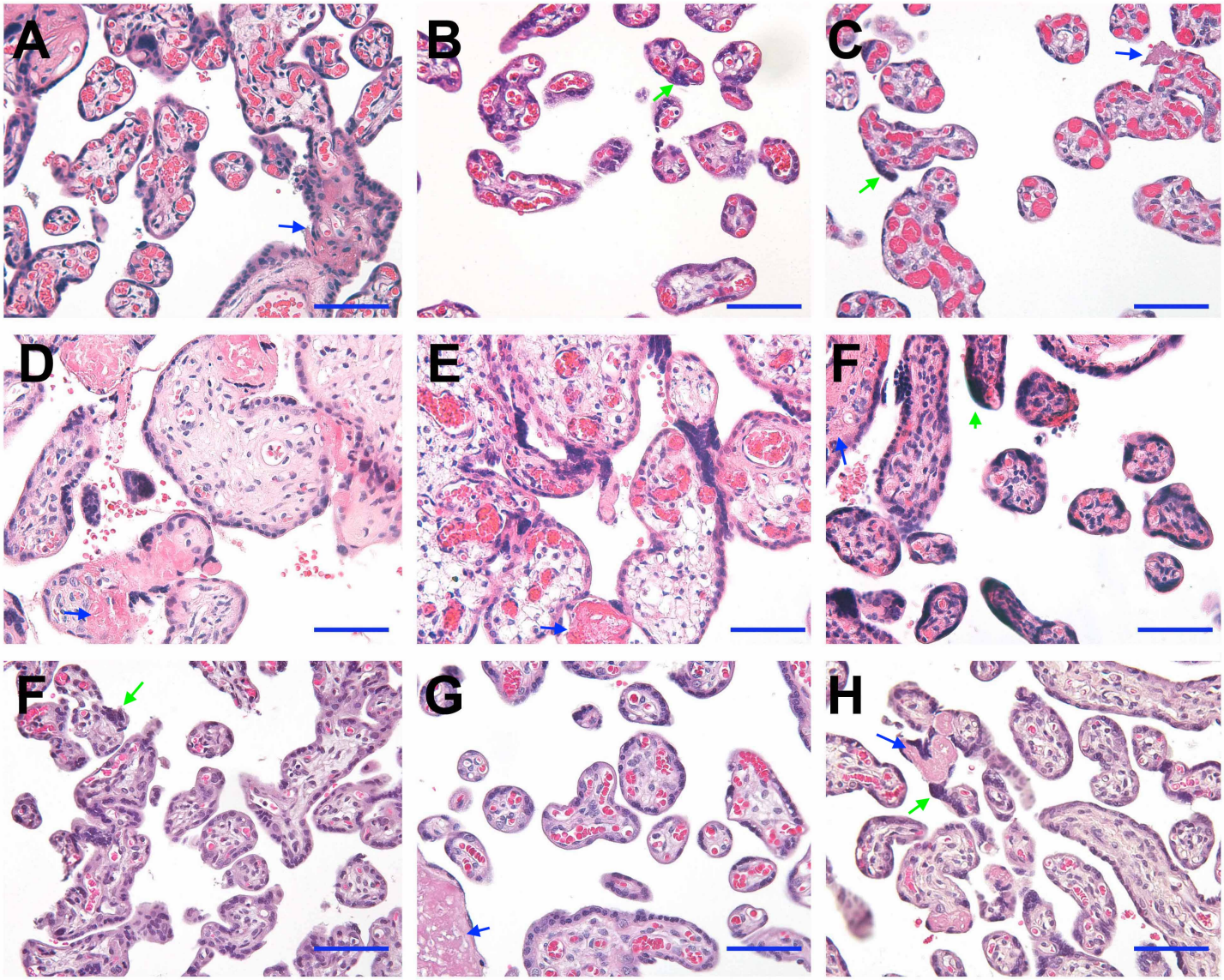

Figure 2

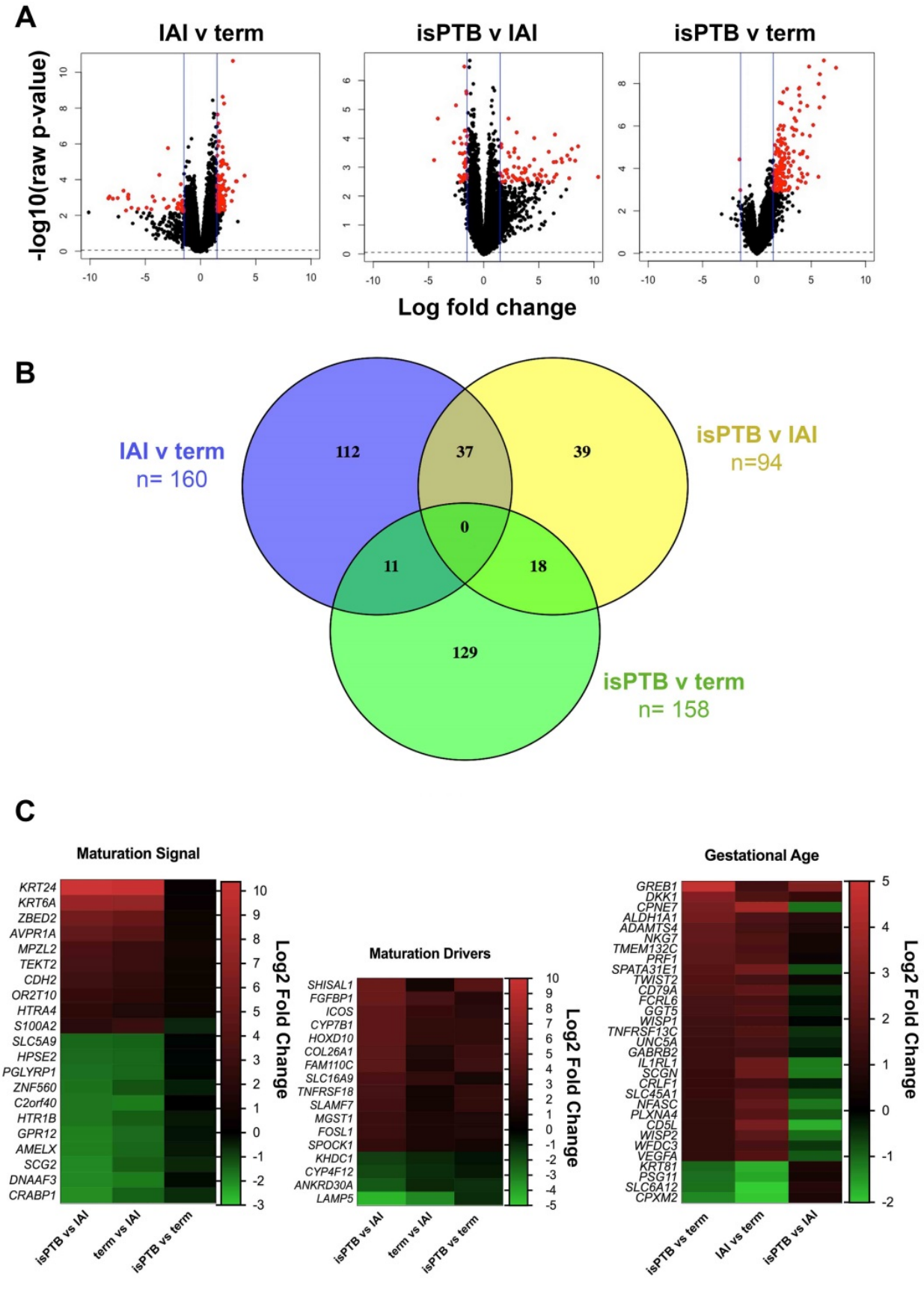

Figure 3

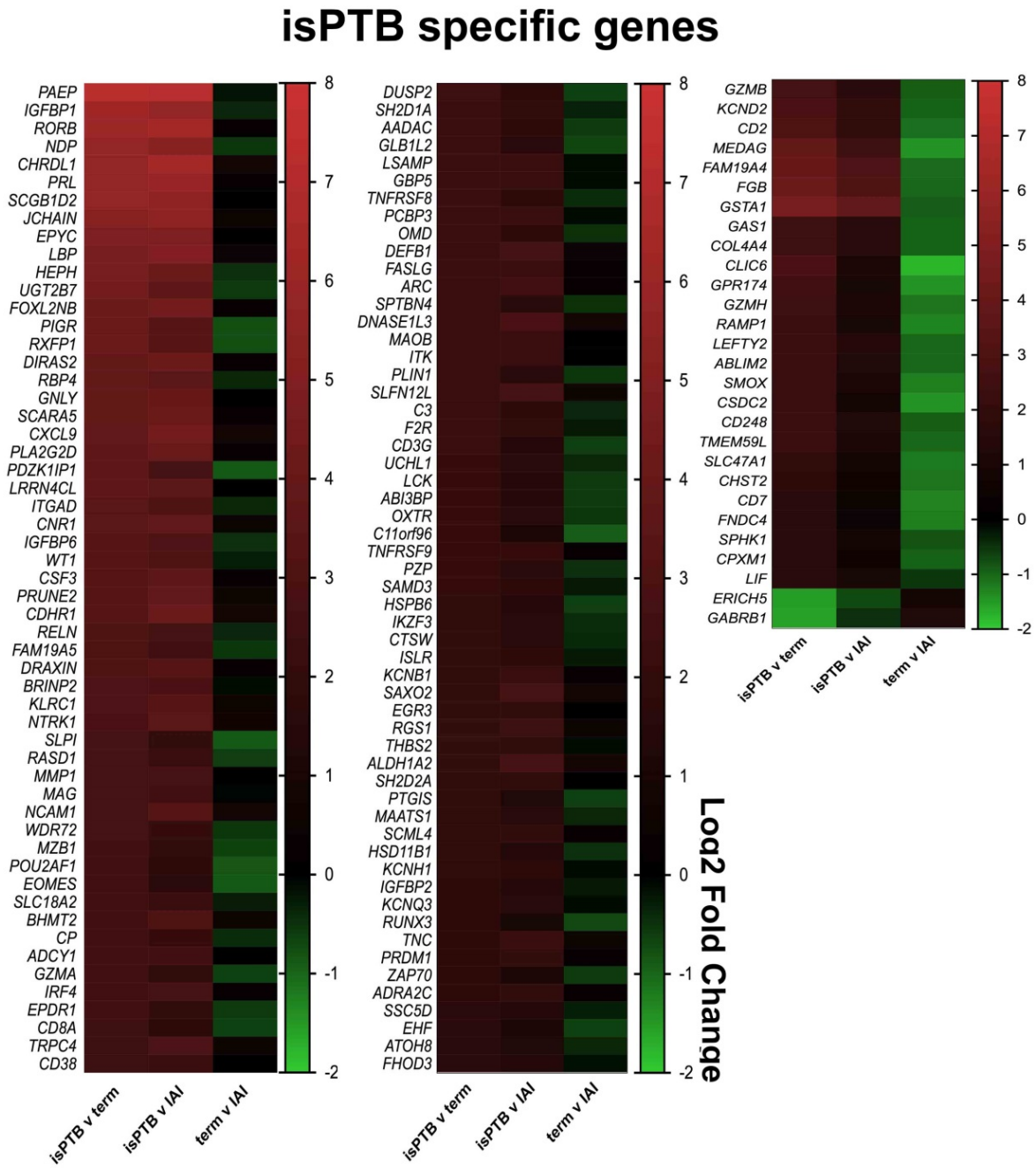

Figure 4