Gene expression analysis in pregnant women and their infants identifies unique fetal biomarkers that circulate in maternal blood

Jill L. Maron,[1] Kirby L. Johnson,[1] Donna Slonim,[2] Chao-Qiang Lai,[3] Marco Ramoni,[4] Gil Alterovitz,[4] Zina Jarrah,[1] Zinger Yang,[2] and Diana W. Bianchi[1]

[1]Department of Pediatrics, Tufts — New England Medical Center, Boston, Massachusetts, USA. [2]Department of Computer Science, Tufts University, Medford, Massachusetts, USA. [3]Jean Mayer — United States Department of Agriculture Human Nutrition Research Center on Aging, Tufts University, Boston, Massachusetts, USA. [4]Children's Hospital Informatics Program, Harvard Medical School, Boston, Massachusetts, USA.

The discovery of fetal mRNA transcripts in the maternal circulation holds great promise for noninvasive prenatal diagnosis. To identify potential fetal biomarkers, we studied whole blood and plasma gene transcripts that were common to 9 term pregnant women and their newborns but absent or reduced in the mothers postpartum. RNA was isolated from peripheral or umbilical blood and hybridized to gene expression arrays. Gene expression, paired Student's *t* test, and pathway analyses were performed. In whole blood, 157 gene transcripts met statistical significance. These fetal biomarkers included 27 developmental genes, 5 sensory perception genes, and 22 genes involved in neonatal physiology. Transcripts were predominantly expressed or restricted to the fetus, the embryo, or the neonate. Real-time RT-PCR amplification confirmed the presence of specific gene transcripts; SNP analysis demonstrated the presence of 3 fetal transcripts in maternal antepartum blood. Comparison of whole blood and plasma samples from the same pregnant woman suggested that placental genes are more easily detected in plasma. We conclude that fetal and placental mRNA circulates in the blood of pregnant women. Transcriptional analysis of maternal whole blood identifies a unique set of biologically diverse fetal genes and has a multitude of clinical applications.

Introduction

The dynamic nature of mRNA transcripts may provide invaluable information on fetal gene expression and fetal and maternal health during pregnancy. Poon et al. (1) were the first to demonstrate that fetal expressed genes could be used for noninvasive prenatal diagnosis by identifying male-specific fetal mRNA transcripts in maternal plasma. Unlike fetal DNA, which increases proportionately in maternal plasma throughout gestation (2), mRNA transcripts vary in their expression during each trimester of pregnancy (3, 4). The previous limitations associated with the identification of fetal DNA in maternal plasma, including dependence on male gender or unique paternal polymorphisms, are theoretically eliminated with the use of mRNA transcripts.

To date, several placental and fetal mRNA transcripts have been identified in maternal plasma. Using real-time quantitative RT-PCR amplification, Ng et al. showed that human chorionic gonadotropin β subunit (*hCGB*), human placental lactogen (*PL*), and corticotropin-releasing hormone (*CRH*) transcripts could be detected in maternal plasma (5). Each gene has a pattern of expression that depends on gestational age. Fetal-derived γ-globin transcripts are elevated in the maternal circulation following elective termination of pregnancy (6). Microarray analysis has identified clinically useful placental transcripts for noninvasive fetal gene profiling (7), such as elevated *CRH* in preeclampsia (8). Each of these placental and/or fetal markers is limited in its scope and restricts our overall understanding of fetal development, fetal pathology, and complications of pregnancy.

In the present study, we utilized gene expression microarrays to perform transcriptional analyses of maternal and fetal whole blood (9). Since fetal-derived mRNA is rapidly cleared from maternal circulation following delivery (10, 11), antepartum and postpartum samples were compared with paired newborn umbilical cord blood samples to identify unique fetal biomarkers in maternal whole blood. Our hypothesis was that if specific gene transcripts were detected in pregnant women prior to delivery as well as in their infants' cord blood, yet were significantly reduced or absent in maternal blood within 24 to 36 hours of delivery, these gene transcripts could theoretically originate from the fetus (Figure 1).

In our original sample set (n = 6), a unique subset of presumed fetal transcripts emerged that differed from previously published maternal plasma profiles (7). Therefore an additional subset of pregnant women and their infants (n = 3) were enrolled. Whole blood and plasma samples were obtained for comparative microarray analyses to better understand the biology of circulating fetal nucleic acids in the maternal circulation.

Results

In each whole blood sample (n = 28), the starting amounts of total isolated RNA varied from 1 to 15 μg; 3 to 15 μg of fragmented complementary RNA (cRNA) was hybridized onto microarrays (Table 1). Mean percent present call (percentage of gene transcripts found to be present following microarray analysis) count was 26% (range, 6.2% to 38.1%). Following paired Student's t test analysis, 157 transcripts met statistical criteria ($P < 0.05$) as candidate fetal genes (Supplemental Table 1; supplemental material

Nonstandard abbreviations used: cRNA, complementary RNA; *DEFA1*, defensin α1; *MYL4*, myosin light polypeptide 4, alkali, atrial, embryonic; *TRBV19*, T cell receptor β variable 19.

Conflict of interest: D.W. Bianchi's laboratory receives sponsored research support, and D.W. Bianchi personally has stock options and receives compensation for consultation, from Living Microsystems.

Citation for this article: *J. Clin. Invest.* **117**:3007–3019 (2007). doi:10.1172/JCI29959.

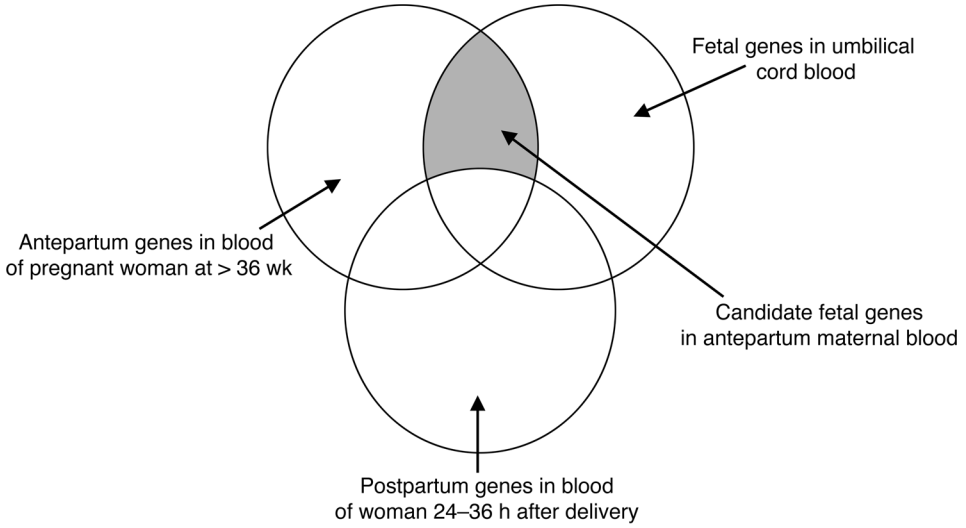

Figure 1
A Venn diagram depicting presumed fetal gene transcripts detected in antepartum maternal blood. Those gene transcripts that were detected in both the antepartum mother and fetus, but were not seen in the postpartum samples were targeted as unique fetal markers.

available online with this article; doi:10.1172/JCI29959DS1). Each transcript was evaluated for its functional role, developmental pattern of expression, and participation in key biological pathways. This resulted in a smaller subset of genes (n = 71) in which the transcript was involved in a developmental process (e.g., *ROBO4*); derived from fetal, placental, or male tissue (e.g., *PLAC1*); or associated with a physiological newborn response (e.g., *NPR1*) or had its expression limited to or highly associated with a fetus or neonate (e.g., *GDF9*) (Tables 2–6). Within this subset, there were 27 developmental genes (musculoskeletal, epidermal, and nervous system), 5 sensory perception genes (olfactory and visual), 22 genes involved in fetal physiologic function or expressed predominantly in fetal, neonatal, placental, or male tissues, and 17 immune defense genes. One gene transcript, AFFX-BioDn-3_at, a spiked control, also met statistical criteria. Presumably its presence in our list of qualifiers is explained by the expected 5% false discovery rate.

Pathway analysis identified transcripts associated with receptor proteins, antigens, and intrinsic membrane proteins. Kyoto Encyclopedia of Genes and Genomes pathway analysis suggested that transcripts involved in T cell biology were highly abundant. These included the CD3D antigen, CD19 antigen, CD22 antigen, and CD123 IL-3 receptor ($P < 0.001$). Biological Biochemical Image Database and BioCarta pathway analyses showed similar results (Table 7).

Real-time RT-PCR amplification was performed on 5 of the 9 whole blood sample sets to confirm the presence of 4 gene transcripts detected on microarrays (*GAPDH*; defensin α1 [*DEFA1*], T cell receptor β variable 19 [*TRBV19*]; and myosin light polypeptide 4, alkali, atrial, embryonic [*MYL4*]). Genes were chosen randomly from an initial list of qualifiers derived from the first 6 subjects studied. There was 100% concordance between genes detected with real-time RT-PCR amplification and with microarray hybridization (Supplemental Table 2).

Initially, 10 commercially available single SNP assays were performed on cDNA synthesized from the original maternal and umbilical cord whole blood samples. Genotyping analysis of 5 women and their infants for 10 SNPs identified 1 SNP (*CD19* rs2904880) within an antepartum sample that was identical to the umbilical cord genotype, yet differed from the maternal postpartum genotype. To further substantiate our findings, an additional 10 women, representing 11 maternal-fetal pairs (one set of dizygotic twins), were enrolled in our study for the sole purpose of SNP analysis. Both genomic DNA and total RNA were obtained from each subject at the established time points. In addition to *CD19*, a SNP analysis was performed for the gene *KIR3DL2*. There were 4 informative maternal-fetal genotypes; 2 revealed evidence of fetal RNA trafficking for *KIR3DL2* (Supplemental Table 3).

In each plasma sample (n = 9), the starting amount of total isolated RNA varied from 3 to 780 ng; 2.5 to 5 μg of fragmented cDNA was hybridized onto microarrays (Table 1). The mean percent present call count for plasma cDNA was 6.8% (range, 4% to 14.2%). Following paired 2-tailed Student's t test analysis, 175 transcripts met statistical criteria (Supplemental Table 4). Again, a spiked control, AFFX-r2-Bs-dap-5_at, met statistical significance.

Overall, comparative analyses of whole blood and plasma transcripts located on the microarrays showed a range of differences (Figure 2A). Still, the interquartile range of the paired t scores contained 0 in all 3 cases, consistent with the fact that fewer than half the transcripts changed significantly in either direction between whole blood and plasma. However, in a set of 50 previously identified genes expressed in term placenta (7), 30 were shown to have significantly higher expression (adjusted $P < 0.05$) in antepartum plasma compared with corresponding antepartum whole blood using a paired 2-tailed Student's t test. A χ^2 test showed that this was significantly higher than would be expected compared with the observed rate (20%) for the full arrays ($P < 0.00001$). This contrasts strongly with the same analysis for the umbilical and postpartum comparisons, for which the differences in expression of this 50-gene set between whole blood and plasma were consistent with the distribution on the entire arrays (P = 0.45 and 0.16, respectively). The t score distribution for each of these comparisons in the 50-gene set is shown in Figure 2B.

Discussion

This study used transcriptional analyses to detect unique fetal biomarkers in maternal whole blood. Identification of these transcripts relied on primary data mining using commercially available software (NetAffx Analysis Center; Affymetrix), publicly available databases (OMIM and Entrez Gene), and publicly available expression profiles (UniGene). Pathway analysis was performed on Database for Annotation, Visualization, and Integrated Discovery (DAVID) (12). This extensive and stringent analysis allowed us to

Table 1
Starting amounts of RNA from whole blood and plasma samples

Sample	Extracted total RNA (μg)	Hybridized cRNA (μg)	Percent present call[A]
Whole blood			
1-A	10	5	6.2
1-U	15	5	23.1
1-P	10	5	34
2-A	1	15	31.1
2-U	9.3	15	22.1
2-P	3	15	34.9
3-A	1	15	31.6
3-U	5	15	29.1
3-P	1.2	15	36.9
4-A	2.4	15	26.6
4-Ua[B]	5.4	15	16.1
4-Ub	2.5	15	24
4-P	4.2	15	19.8
5-A	8.8	15	23.6
5-U	15	15	17.3
5-P	5.1	15	37.3
6-A	1.4	3	30.9
6-U	4.8	15	30
6-P	2.1	15	38.1
7-A	4	5	27.7
7-U	8.4	5	17.2
7-P	5	5	24.1
8-A	1	15	29.9
8-U	15	15	16
8-P	1	15	29
9-A	4.62	15	26
9-U	10	15	10
9-P	1.68	15	30
Plasma			
1-A	3	2.5	5
1-U	25	5	7
1-P	15	5	5
2-A	780	5	9
2-U	240	5	14.2
2-P	720	3.95	4
3-A	540	4.41	6.7
3-U	540	5	5.5
3-P	720	5	5

[A]Percentage of gene transcripts found to be present following microarray analysis. [B]Samples 4-Ua and 4-Ub are from twins. A, antepartum maternal blood; U, umbilical cord blood; P, postpartum maternal blood.

identify presumably fetal gene transcripts that not only met statistical criteria but also proved biologically plausible through expression, functional, and pathway analyses.

Our results were both surprising and encouraging. By sampling maternal blood late in pregnancy, we identified specific gene transcripts that appeared to be associated with a fetus preparing to transition from the in utero environment. Sensory perception genes, particularly of the visual system, were upregulated in the term fetus and detected in the antepartum mother (Table 2). This was unexpected. Visual pathway genes included the neural retina leucine zipper (*NRL*), which is preferentially expressed in rod photoreceptors and is essential for photoreceptor development and function, and dehydrogenase/reductase member 3 (*DHRS3*), which is involved in visual perception and retinol metabolism.

Developmental genes from the embryonic, muscular, skeletal, epidermal, and nervous systems were also identified (Tables 3 and 4). The nervous system genes constituted approximately half (48%) of the genes in this group (Table 4). Though subsequent studies profiling fetal developmental genes throughout pregnancy must be performed, it is interesting that neurodevelopmental genes predominate. It is well documented that near-term infants (32 to 36 weeks) may suffer from complications due to immaturity of the nervous system, presenting clinically as apnea of prematurity or an uncoordinated suck-swallow reflex. Thus it makes physiological sense that term fetuses would be upregulating neurodevelopmental transcripts.

Several candidate fetal-derived genes of interest were also identified, based upon their tissue expression pattern or a probable neonatal physiological response (Table 5). Of particular interest were natriuretic peptide receptor A (*NPR1*) and S100 calcium binding protein β (neural) (*S100B*). *NPR1* is likely responsible for the physiologic diuresis that occurs over the first 24 to 48 hours in all neonates, while *S100B* is involved in neurodevelopment. Overexpression of *S100B* has been associated with an increased susceptibility to perinatal hypoxia-ischemia in mice (13). This gene may be of interest for future studies, particularly in infants at risk for hypoxic-ischemic encephalopathy, a severe and often fatal disease.

Immune-mediated gene transcripts, several of which are specifically expressed in the fetus or neonate, were upregulated (Table 6). Given the intricate and delicate immune balance that must exist between a pregnant woman and her fetus (14, 15), this was not surprising. However, the immune gene transcripts identified in this study were not involved in immune tolerance but rather included genes that were involved with a host's ability to mount an inflammatory response (complement factor B [*CFB*], natural cytotoxicity triggering receptor 2 [*NCR2*], and SLAM family member 8 [*SLAMF8*]) and its ability to defend against bacteria, fungi, and viruses (bactericidal/permeability increasing protein [*BPI*], defensin α4, corticostatin [*DEFA4*], cathelicidin antimicrobial peptide [*CAMP*], cation channel, sperm associated 2 cation channel pseudogene [*CATSPER2,CATSPER2P1*], and testis-specific kinase substrate [*TSKS*]). Additionally, independent pathway analysis revealed an abundance of transcripts associated with T cells. This finding was also intriguing given that the thymus gland, the major regulator of all T cells, is much more prominent and active in the neonate compared with the mother. Though none of these genes can be considered exclusively fetal in origin, it again appears that profiling a term infant reveals upregulation of gene transcripts essential for fetal-to-newborn transition.

There are 2 statistically significant transcripts that appear to be associated with male gender (Table 5). The 2 gene transcripts (cation channel, sperm associated 2 cation channel pseudogene and testis-specific kinase substrate) are either expressed in spermatozoa or involved in spermatogenesis. Neither of these transcripts maps to the Y chromosome. It may be surprising, given the predominance of male infants in our study (70%), that more gender-specific transcripts do not appear on our list. However, there are only 47 Y-chromosome specific transcripts on the Affymetrix HGU133a microarray, and only 5 of these are expressed at higher levels in male infants than in female infants in our study, which limits our ability to detect differences in gender-specific genes in the antepartum samples. Finally, given the small number ($n = 3$) of female infants in this study, there was inadequate power to confirm any sex-specific differences in the antepartum samples.

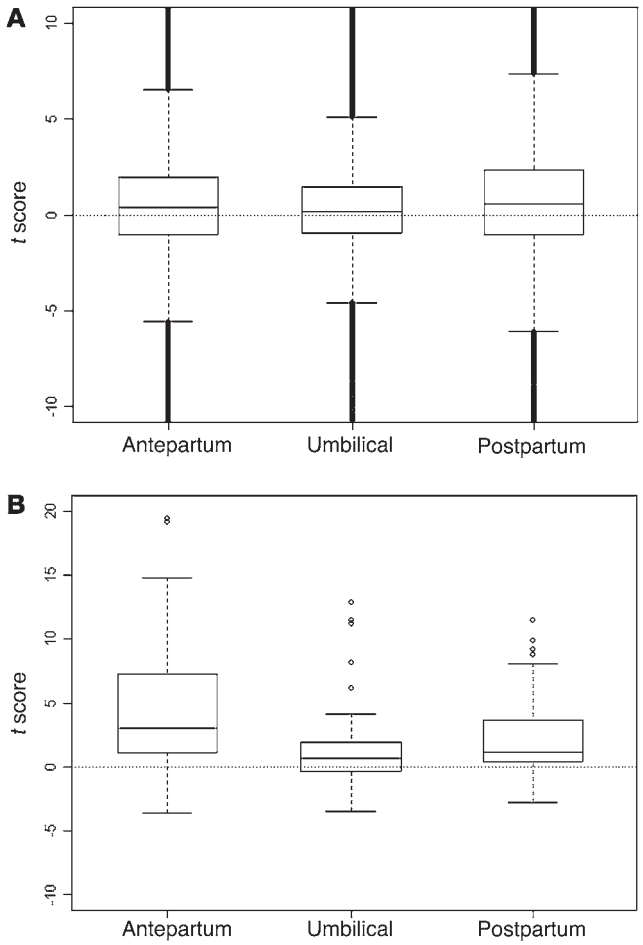

Figure 2
Boxplots of paired *t* scores comparing gene expression in plasma and whole blood samples. The boxes mark the first quartile, median, and third quartile of the *t* scores for each comparison. Vertical dotted lines represent *t* scores within 1.5 times the interquartile range. Positive *t* scores represent transcripts higher in plasma samples; negative *t* scores represent transcripts higher in whole blood. (**A**) Distribution of paired *t* scores for all 22,283 transcripts measured in antepartum, umbilical cord, and postpartum samples. What appear to be bold vertical lines are actually numerous dots. (**B**) Distribution of paired *t* scores for 50 previously identified placental transcripts. Dots indicate outliers.

The presence of fetal-specific mRNA sequences was confirmed with multiple techniques. Not only were individual transcripts identified on microarrays confirmed by real-time RT-PCR amplification (Supplemental Table 2), but there was also an independent SNP analysis that identified 3 distinct fetal transcripts for 2 different genes in maternal antepartum blood (Supplemental Table 3). Our detection of 3 unique fetal SNPs from only 5 possible informative mother-infant combinations substantiates our work and provides definitive proof that biologically diverse fetal gene transcripts are entering the maternal circulation. Additionally, the UniGene analysis performed on the 157 statistically significant genes identified 17 transcripts predominantly expressed in the fetus, 5 transcripts predominantly expressed in the embryo, 2 transcripts predominantly expressed in the neonate, and 1 gene whose expression was restricted to the embryo.

Maternal and umbilical cord whole blood samples were the source of RNA for this study. Notably, whole blood is cellular and fetal cells persist for decades within the postpartum mother (16, 17). However, the identification of fetal genes in our study was dependent upon their absence in maternal postpartum samples and is likely not representative of prior pregnancies. The use of whole blood is also known to interfere with microarray hybridization rates. Although there are commercially available kits to reduce the amount of globin genes found in whole blood, we were unable to utilize such kits because they selectively reduce β- and α-globin genes, not the predominantly fetal γ-globin gene. Thus globin reduction of this sample set would result in a preferential reduction in the maternal samples only. Our results indicate that non–globin-reduced whole blood may have mild interference with the hybridization process. Our mean percent present call was 26%, which is lower than generally established mean (40% to 45%) using the same techniques (18). Additionally, blood samples were obtained over an extended time period, resulting in the use of microarrays from several different lots. While varying lots of microarrays may increase variability of findings, it will minimize the chance of a bias or systematic error associated with a single lot (19). Finally, recent literature has emerged demonstrating the reliability and reproducibility of a single microarray, and therefore no arrays were run in multiples (20).

The overwhelming majority of the literature examining fetal nucleic acid trafficking cites the placenta as the major source of fetal transcripts and utilizes plasma for their detection (21). Although some of the gene transcripts identified in this whole blood analysis were expressed in placental tissue, including placental-specific 1 (*PLAC1*), we did not identify other previously documented placental transcripts (*PL*, *CRH*, or *hCGB*) in maternal whole blood. This was both surprising and unexpected and resulted in additional studies in a subset of women and their infants who had comparative whole blood and plasma analyses. Overall, the distributions of *t* scores represented in paired gene expression differences between whole blood and plasma did not appear to be too strongly biased. However, when specific genes were targeted, a different profile emerged. A number of placental genes previously identified in maternal plasma were present in the antepartum plasma samples but continued to elude detection in corresponding whole blood. Identification of well-established placental genes within our plasma samples validates the work of previous authors and substantiates our own plasma microarray data despite our low hybridization rate. This result suggests that placental transcripts are more readily detectable in the maternal circulation from plasma than from whole blood. The discrepancy seen between gene expression profiles in whole blood and plasma raises important questions regarding the biological mechanisms of fetal nucleic acid trafficking.

In summary, fetal as well as placental mRNAs circulate in the blood of pregnant women. This report illustrates what we believe to be a novel approach to noninvasively monitor the developing fetus. Our results have been confirmed by multiple methods. Transcriptional analysis of maternal whole blood identifies a unique set of biologically diverse fetal genes not previously recognized, which may differ from comparable plasma analysis, and suggests fetal upregulation of pathways necessary for extrauterine life. The transcripts identified can serve as a baseline to compare fetuses affected by a variety of pathologic conditions and will have a multitude of important clinical applications in prenatal diagnosis, perinatology, and neonatology.

Table 2
Candidate fetal sensory perception gene transcripts

Gene	Gene symbol	Accession no.	UNIQID	Biological function	P^A	P^B	Maternal-fetal count[C]
Dehydrogenase/ reductase member 3	DHRS3	NM_004753	202481_at	Visual perception, metabolism, retinol metabolism	< 0.0001	< 0.0001	9 of 10
Olfactory receptor, family 7, subfamily e, member 37 pseudogene[D]	OR7E37P	AW874308	217499_x_at	Perception of smell	0.0003	< 0.0001	10 of 10
Neural retina leucine zipper	NRL	NM_006177	206597_at	Regulation of transcription and rhodopsin gene activity, visual perception, response to stimulus, photoreceptor development and function, preferentially expressed in rod photoreceptors	0.005	< 0.0001	7 of 10
Tumor-associated calcium signal transducer 2	TACSTD2	J04152	202286_s_at	Signal transduction, visual perception, cell proliferation, response to stimulus, regulation of progression through cell cycle	0.02	0.0002	8 of 10
Protein phosphatase, EF-hand calcium binding domain 2	PPEF2	NM_006239	208411_x_at	Protein amino acid dephosphorylation, visual perception, response to stimulus, detection of stimulus during sensory perception	0.04	0.002	9 of 10

Transcripts associated with the visual, auditory, or olfactory systems were selected as fetal qualifiers. [A]Adjusted *P* values for antepartum vs. postpartum maternal blood samples. [B]Adjusted *P* values for umbilical cord blood vs. postpartum maternal blood samples. [C]Number of maternal-fetal pairs who met statistical criteria for gene transcript. [D]Expression pattern restricted to embryo.

Methods

Subjects. All pregnant women participating in this study presented to the Labor and Delivery Unit at Tufts – New England Medical Center, Boston, Massachusetts, USA, at 36 weeks' or more gestation for a scheduled cesarean section. Cesarean section deliveries were chosen for ease of sample attainment and to provide relatively uniform delivery conditions. Informed consent was obtained according to the protocol approved by the hospital's Institutional Review Board. Nine women and 10 infants (1 twin gestation) participated in this study. In the first 6 study participants, only whole blood was obtained; in the subsequent subjects, both whole blood and plasma were obtained. A total of 37 (9 antepartum whole blood, 3 antepartum plasma, 9 postpartum whole blood, 3 postpartum plasma, 10 newborn whole blood, 3 newborn plasma) microarrays comprise the data set. One woman had undergone in vitro fertilization; the other pregnancies were conceived naturally. All other clinical and demographic details of study subjects are given in Table 8.

RNA whole blood isolation. Whole blood samples (7.5 cc) were obtained in 3 PaxGene (PreAnalytiX) tubes immediately prior to cesarean delivery and 24 to 36 hours postpartum and from the corresponding umbilical cord(s). Samples were stored at room temperature for 6 to 36 hours prior to RNA extraction in accordance with the PaxGene blood RNA kit (PreAnalytiX) manufacturer's protocol. During extraction, on-column DNase digestion (RNase-Free DNase Set; Qiagen) was performed to eliminate DNA contamination. After final elution of extracted total RNA, 2.0 μl from each sample were analyzed on the Bioanalyzer 2100 (Agilent) to assess the quantity and purity of each sample. Only samples with clear peaks at ribosomal 18S and 28S that had a minimum starting amount of 1 μg total RNA met our quality and quantity criteria for subsequent amplification (Table 1).

RNA plasma isolation. Whole blood was collected in 10-cc EDTA Vacutainer blood tubes (BD). Blood samples underwent an initial spin at 1,600 *g* for 10 minutes at 4°C. Harvested plasma was then transferred into microcentrifuge tubes and underwent an additional spin at 15,350 *g* for 10 minutes at 4°C to eliminate any residual cells. Plasma extraction was performed in accordance with previously published reports (22). Total RNA was eluted in 60 μl of RNase-free water. Quantity and quality of extracted total plasma RNA was assessed with the Bioanalyzer 2100 (Agilent) prior to amplification.

RNA whole blood amplification. Extracted total RNA from whole blood was amplified according to the Eberwine protocol (23) with the use of the commercially available One Step Amplification Kit (Affymetrix). Amplified cRNA was again assessed with the Bioanalyzer 2100 for purity and quantity prior to fragmentation. Globin reduction was not performed on any sample.

RNA plasma amplification. Extracted total RNA from plasma was amplified with the WT–Ovation Pico RNA Amplification System (NuGEN Technologies Inc.) in accordance with manufacturer's protocol. Amplified cDNA was assessed with the Bioanalyzer 2100 for purity and quantity. Fragmentation and biotinylation occurred with the FL–Ovation cDNA Biotin Module v2 (NuGEN Technologies Inc.) per the manufacturer's protocol.

Microarray hybridization, staining, and scanning. Approximately 15 μg of amplified and labeled cRNA from whole blood and 5 μg of amplified and labeled cDNA from plasma samples were fragmented and hybridized onto GeneChip Human Genome U133A microarrays (Affymetrix) (Table 7). Hybridization quantities of whole blood cRNA and plasma cDNA were based on each respective amplification protocol. Following hybridization, each array was washed and stained in the GeneChip Fluidics Station 400 (Affymetrix). Each array was used once. Arrays were then scanned with the GeneArray Scanner (Affymetrix), and initial analysis was performed using the GeneChip Microarray Suite 5.0 (Affymetrix).

Real-time RT-PCR. Confirmatory 1-step real-time RT-PCR was done on the extracted total RNA from 5 of the original 6 whole blood sample sets for 4 genes identified on the microarrays. The genes identified were

Table 3
Candidate fetal non-nervous system developmental gene transcripts

Gene	Gene symbol	Accession no.	UNIQID	Biological function	*P*[A]	*P*[B]	Maternal-fetal count[C]
Cysteine- and glycine-rich protein 2	CSRP2	U46006	211126_s_at	Muscle development, organ morphogenesis, cell proliferation, growth and differentiation, development	0.008	< 0.0001	7 of 10
S-phase kinase-associated protein 1A	SKP1A	NM_006930	200719_at	Ubiquitin cycle, development or maintenance of inner ear functions	0.02	< 0.0001	9 of 10
Exostoses (multiple)-like 1	EXTL1	NM_004455	206329_at	Skeletal development	0.026	< 0.0001	9 of 10
Roundabout homolog 4, magic roundabout (*Drosophila*)[D]	ROBO4	NM_019055	220758_s_at	Angiogenesis, cell differentiation, regulation of cell migration development	0.05	< 0.0001	7 of 10
Keratin 84	KRT84	Y19209	217031_at	Structural constituent of epidermis, heterodimers with type 1 keratin	0.05	0.0002	7 of 10
Sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3F	SEMA3F	U38276	209730_at	Development	0.03	0.0004	9 of 10
Dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 4	DYRK4	NM_052850	212954_at	Phosphorylation, cell proliferation, survival, and development	0.001	0.008	8 of 10
Frizzled homolog 6 (*Drosophila*)	FZD6	NM_003506	203987_at	Establishment of tissue polarity, signal transduction, signaling pathways, development	0.05	0.001	7 of 10
Musculin (activated B cell factor 1)	MSC	AF060154	209928_s_at	Regulation of transcription, development, specifically myogenesis	0.02	0.002	9 of 10
Collagen type VIII α1	COL8A1	NM_001850	214587_at	Phosphate transport, cell adhesion, somatic muscle development	0.01	0.03	8 of 10
ST6 β-galactosamide α-2,6-sialyltransferase 1	ST6GAL1	X17247	214970_s_at	Glycosylation, immune response, development, oligosaccharide metabolism	0.01	0.03	9 of 10
Integral membrane protein 2A	ITM2A	NM_004867	202747_s_at	Expressed in outer perichondrial rim of postnatal mandibular condyle	0.05	0.01	7 of 10
IL-3 receptor α (low affinity)	IL3RA	NM_002183	206148_at	Phosphorylation, development	0.03	0.02	8 of 10
Peroxisome proliferator-activated receptor δ	PPARD	BC002715	210636_at	Regulation of transcription, insulin secretion, differentiation, fatty acid transport and catabolism, apoptosis, embryo implantation, lipid metabolism, cell proliferation, nerve ensheathment, epidermis development, decidualization	0.05	0.02	8 of 10

Genes were selected based upon known developmental function. [A]Adjusted *P* values for antepartum vs. postpartum maternal blood samples. [B]Adjusted *P* values for umbilical cord blood vs. postpartum maternal blood samples. [C]Number of maternal-fetal pairs that met statistical criteria for gene transcript. [D]Embryo-specific expression pattern.

GAPDH, *DEFA1*, *TRBV19,* and *MYL4*. All experiments were performed on the PerkinElmer Applied Biosystems 7700 Sequence Detector with the TaqMan One-Step RT-PCR Master Mix Reagents Kit (Applied Biosystems Inc.) as previously described (24).

The amplification primers and probes were as follows: GAPDH forward, 5′-GAAGGTGAAGGTCGGAGTC-3′, GAPDH reverse, 5′-GAAGATGGTGATGGGATTTC-3′, GAPDH probe, 5′-(6-FAM)-CAAGCTTCCCGTTCTCAGCC-(TAMRA)-3′; TRBV19 forward, 5′-GGCCACCTTCTGGCAGAA-3′, TRBV19 reverse, 5′-AGAGCCCGTAGAACTGGACTTG-3′, TRBV19 probe, 5′-(6-FAM)-CCCCGCAACCACTTCCGCTG-(TAMRA)-3′; DEFA1 forward, 5′-CAGCCCCGGAGCAGATT-3′, DEFA1 reverse, 5′-TTTCGTCCCATGCAAGGG-3′, DEFA1 probe, 5′-(6-FAM)-CAGCGGACATCCCAGAAGTGGTTGT-(TAMRA)-3′; MYL4 forward, 5′-GGGCCTGCGTGTCTTTGA-3′, MYL4 reverse, 5′-CGTGCCGAAGCTCAGCA-3′ MYL4 probe, 5′-(6FAM)-AAGGAGAGCAATGGCACGGTCATGG-(TAMRA)-3′ (Applied Biosystems Inc.).

Calibration curves for *DEFA1*, *TRBV19*, and *MYL4* were prepared with extracted total RNA from umbilical cord blood. The stock umbilical RNA was amplified with the WT–Ovation RNA Amplification System (NuGEN Technologies Inc.) and then serially diluted 10-fold to generate a curve. The calibration curve for *GAPDH* was prepared from commercially available Total RNA (Applied Biosystems Inc.).

The thermal cycle profile for all transcripts was as follows: the reaction was initiated at 48°C for 30 minutes for the uracil N-glycosylase to

Table 4
Candidate fetal nervous system developmental gene transcripts

Gene	Gene symbol	Accession no.	UNIQID	Biological function	P^A	P^B	Maternal-fetal count[C]
Neuralized-like (*Drosophila*)	NEURL	NM_004210	204888_s_at	Nervous system development, determination of cell fate in neurogenic region of embryo	< 0.0001	0.008	9 of 10
Par-3 partitioning defective 3 homolog (*C. elegans*)	PARD3	NM_019619	221526_x_at	Cell division, maintenance of cell polarity, PKC activation, axonogenesis	0.04	< 0.0001	8 of 10
MAP/microtubule affinity-regulating kinase 4	MARK4	AB049127	221560_at	Cytoskeleton organization and biogenesis, phosphorylation, nervous system development, apoptosis, cell proliferation, Wnt receptor signaling	0.0005	0.0001	9 of 10
Cholinergic receptor, muscarinic 5[D]	CHRM5	NM_012125	221347_at	Signal transduction, nervous system development, adenylate cyclase–inhibiting pathway, cell proliferation	0.007	0.0001	8 of 10
Neuregulin 1	NRG1	L12260	208232_x_at	Nervous system/embryonic development, transcription, cell differentiation	0.03	0.0002	7 of 10
S100 calcium binding protein β (neural)[D]	S100B	BC001766	209686_at	Energy reserve metabolism, calcium ion homeostasis, apoptosis, axonogenesis, cell proliferation, complement activation, cytokine and protein biosynthesis, neuronal synaptic plasticity, astrocyte activation	0.008	0.0006	8 of 10
Myelin associated glycoprotein	MAG	X59350	38521_at	Myelination, cell adhesion, antimicrobial response	0.0007	0.04	8 of 10
Acetylcholinesterase (Yt blood group)	ACHE	M55040	210332_at	Nervous system/muscle development, synaptogenesis, replication, cell adhesion/ proliferation, wound response, neurotransmitter catabolism, protein metabolism, regulation of protein secretion	0.04	0.001	8 of 10
Embryonic lethal, abnormal vision, Drosophila–like 3 (Hu antigen C)	ELAVL3	NM_001420	206338_at	Nervous system development, cell differentiation, development	0.02	0.002	7 of 10
Disabled homolog 1 (*Drosophila*)	DAB1	NM_021080	220611_at	Nervous system development, cell differentiation, development	0.03	0.003	7 of 10
Amyloid β (A4) precursor protein-binding, family A, member 2 (X11-like)	APBA2	NM_005503	209871_s_at	Nervous system development, protein transport	0.009	0.007	8 of 10
Neuropilin 1	NRP1	NM_001024628	212298_at	Angiogenesis, nervous system development, axon guidance, signal transduction, organ morphogenesis, cell adhesion, proliferation, and differentiation	0.03	0.02	7 of 10
Transcription factor activating enhancer binding protein 2β)	TFAP2B	AL031224	215686_x_at	Transcription regulation, nervous system development, cell proliferation during embryonic development	0.04	0.03	5 of 10

Genes were selected based upon known nervous system developmental function. [A]Adjusted *P* values for antepartum vs. postpartum maternal blood samples. [B]Adjusted *P* values for umbilical cord blood vs. postpartum maternal blood samples. [C]Number of maternal-fetal pairs who met statistical criteria for gene transcript. [D]Fetal-specific expression pattern.

act, followed by reverse transcription at 60°C for 30 minutes. Following a 5-minute denaturing cycle at 95°C, 40 cycles of PCR were performed with 20 seconds of denaturing at 94°C, then 1 minute at 60°C for annealing and extension.

SNP analysis. SNP analysis was performed both retrospectively on stored RNA samples and prospectively from genomic DNA and total RNA in whole blood collected from women and infants specifically for SNP identification in both cases. cDNA was synthesized from total RNA using the WT–Ovation RNA Amplification System (NuGEN Technologies Inc.). SNPs were chosen if their sequence was contained within a genomic DNA coding region and was found within the coding region identified on the HGU133a Affymetrix gene expression microarray. Additionally, each SNP

Table 5
Candidate fetal gene transcripts chosen for their likely involvement in fetal physiologic function

Gene	Gene symbol	Accession no.	UNIQID	Biological function	P^A	P^B	Maternal-fetal count[C]
Cyclin B2	CCNB2	NM_004701	202705_at	Cell cycle regulation, translational initiation, protein biosynthesis, spermatogenesis	< 0.0001	0.002	8 of 10
Transcription factor 4[D]	TCF4	NM_003199	212382_at	DNA-dependent transcription regulation, negative regulation of transcription from RNA polymerase II promoter	0.0001	0.0001	9 of 10
IQ motif containing C[D]	IQCC	NM_018134	206650_at	Unknown	0.04	0.0002	8 of 10
Placenta-specific 1[E]	PLAC1	NM_021796	219702_at	Placenta development	0.01	0.0003	9 of 10
mRNA sequence, IMAGE clone 446411[F]		H49077	215477_at	Unknown	0.05	0.0003	8 of 10
Natriuretic peptide receptor A/guanlyate cyclase A	NPR1	NM_000906	204648_at	Fluid secretion, blood pressure regulation, vascular permeability, angiogenesis, diuresis, natriuresis, vasodilation	0.010	0.0005	8 of 10
Testis-specific kinase substrate	TSKS	NM_021733	220545_s_at	Testicular physiology, protein binding, spermatogenesis	0.02	0.0006	7 of 10
Histone 1, H4e[F]	HIST1H4E	NM_003545	206951_at	Chromatin architecture, phosphoinositide -mediated signaling	0.0008	0.04	8 of 10
Glutamate dehydrogenase 2[G]	GLUD2	BC005111	210447_at	Electron transport, amino acid metabolism	0.03	0.0009	9 of 10
Hydroxysteroid (17-β) dehydrogenase 6	HSD17B6	U89281	37512_at	Androgen biosynthesis, catabolism, and metabolism	0.001	0.02	7 of 10
Solute carrier family 8 (sodium-calcium exchanger), member 2[F]	SLC8A2	NM_015063	215267_s_at	Sodium and calcium ion transport, cell communication	0.03	0.001	8 of 10
Transmembrane 6 superfamily member 2[D]	TM6SF2	AK024515	216736_at	Unknown	0.01	0.004	7 of 10
Elastase 2, neutrophil[F]	ELA2	NM_001972	206871_at	Inflammatory and UV response, chemokine biosynthesis, chemotaxis, regulation of IL-8, SMC proliferation	0.004	0.03	10 of 10
Acidic (leucine-rich) nuclear phosphoprotein 32 family, member C[F]	ANP32C	NM_012403	208538_at	Tumor suppressor activity	0.03	0.004	7 of 10
Solute carrier organic anion transporter family, member 1B3[F]	SLCO1B3	NM_019844	206354_at	Ion and organic anion transport	0.02	0.0043	7 of 10
Glycerophosphodiester phosphodiesterase domain containing 2[D]	GDPD2	NM_017711	220291_at	Glycerol metabolism	0.008	0.007	8 of 10
Pyruvate kinase liver and rbc[F]	PKLR	M15465	210451_at	Glycolysis	0.008	0.03	8 of 10
Somatostatin receptor 2	SSTR2	NM_001050	214597_at	Cell-cell signaling, nutrient digestion, cell proliferation	0.009	0.01	7 of 10
Signal-regulatory protein γ[F]	SIRPG	NM_018556	220485_s_at	Intracellular signaling, cell-cell signaling and proliferation	0.02	0.01	7 of 10
Growth differentiation factor 9[F]	GDF9	NM_005260	221314_at	TGF-β receptor signaling, female gamete generation	0.02	0.04	6 of 10
Receptor-associated protein of the synapse[F]	RAPSN	BC004196	211570_s_at	Synaptic transmission	0.03	0.04	6 of 10
Cation channel, sperm associated 2 cation channel pseudogene[F]	CATSPER2, CATSPER2P1	NM_054020, NR_002318	217588_at	Ion transport[H]	0.04	0.03	4 of 10

[A]Adjusted *P* values for antepartum vs. postpartum maternal blood samples. [B]Adjusted *P* values for umbilical cord blood vs. postpartum maternal blood samples. [C]Number of maternal-fetal pairs who met statistical criteria for gene transcript. [D]Embryo-specific expression pattern. [E]Placental/fetal-specific expression pattern. [F]Fetal-specific expression pattern. [G]Neonate-specific expression pattern. [H]Specific to spermatozoa, particularly the flagellum.

Table 6
Immunologic gene transcripts

Gene	Gene symbol	Accession no.	UNIQID	Biological function	P^A	P^B	Maternal-fetal count[C]
Cathepsin G	CTSG	NM_001911	205653_at	Proteolysis, immune response, killing of pathogens	0.004	< 0.0001	9 of 10
SLAM family member 8[D]	SLAMF8	NM_020125	219386_s_at	Immune response, lymphocyte activation	0.005	< 0.0001	9 of 10
Cathelicidin antimicrobial peptide	CAMP	U19970	210244_at	Defense response to pest, pathogen, or parasite	0.014	< 0.0001	8 of 10
CREB/ATF bZIP transcription factor	CREBZF	NM_001039618	213584_s_at	Regulation of gene expression, viral response	< 0.0001	0.04	8 of 10
Interleukin 23, α subunit p19	IL23A	NM_016584	217328_at	Inflammatory and immune response, response to virus, tissue development, innate immune response	0.008	0.0002	8 of 10
Killer cell Ig-like receptor	KIR3DL2	AF263617	211688_x_at	Immune response, NK cell activation and negative regulation, cellular defense response, signal transduction	0.009	0.0002	9 of 10
CD19 molecule	CD19	NM_001770	206398_s_at	Cellular defense response, signal transduction	0.007	0.0006	10 of 10
CD3d molecule δ (CD3-TCR complex)	CD3D	NM_000732	213539_at	Protein complex assembly, signal transduction, T cell activation and selection	0.0009	0.05	9 of 10
Killer cell Ig-like receptor	KIR2DL4	AF276292	211242_x_at	Cellular defense response, signal transduction	0.001	0.0028	8 of 10
Interferon-α2	IFNA2	NM_000605	211338_at	Induction of apoptosis, defense response to virus, inflammatory response, signal transduction, cell-cell signaling	0.02	0.001	7 of 10
Complement factor B	CFB	AF349679	211920_at	Proteolysis, complement activation, immune response	0.002	0.004	7 of 10
Defensin α4, corticostatin[E]	DEFA4	NM_001925	207269_at	Xenobiotic metabolism, defense response to pest, pathogen or parasite	0.04	0.002	9 of 10
Bactericidal/permeability-increasing protein	BPI	NM_001725	205557_at	Immune response, defense response to pest, pathogen or parasite	0.009	0.01	8 of 10
Tubulin, β2C	TUBB2C	NM_006088	213726_x_at	Microtubule-based movement, NK cell–mediated cytotoxicity, protein polymerization	0.03	0.01	9 of 10
Chemokine (C motif) ligand 2[E]	XCL2	U23772	206366_x_at	Chemotaxis, immune response, signal transduction, cell-cell signaling, circulation, calcium ion homeostasis, antimicrobial response	0.01	0.04	8 of 10
Complement factor H-related 2	CFHR2	NM_005666	206910_x_at	Immune response	0.02	0.02	9 of 10
Natural cytotoxicity triggering receptor 2[E]	NCR2	NM_004828	221074_at	Defense response, signal transduction	0.03	0.02	7 of 10

Transcripts represent both immune tolerance and immune defense. [A]Adjusted *P* values for antepartum vs. postpartum maternal blood samples. [B]Adjusted *P* values for umbilical cord blood vs. postpartum maternal blood samples. [C]Number of maternal-fetal pairs who met statistical criteria for gene transcript. [D]Neonate-specific expression pattern. [E]Fetal-specific expression pattern.

had to demonstrate a high polymorphism rate (>10%) across multiple populations and be included in a gene transcript that was identified in our antepartum subjects no less than 90% of the time. Identification of SNPs was performed with the ABI 7900HT genotyping system based on TaqMan SNP genotyping technology (Applied Biosystems Inc.). Genomic DNA, when available, and cDNA samples were genotyped for each SNP assay using Applied Biosystems TaqMan SNP genotyping system (25).

Selection of candidate fetal genes. Microarray data analysis was performed in R 2.3.0 using the Affy and Multtest packages in Bioconductor 1.8 (http://www.bioconductor.org/) (26). All of the arrays were normalized as a set using the quantile normalization method. Two sets of paired Student's *t* tests were performed on the (base 2) log-transformed data from the whole blood samples, 1 comparing the antepartum and postpartum women, and 1 comparing the umbilical cord blood and the postpartum mothers. In the

Table 7
Pathway analysis

Gene	Accession no.	Pathway	Pathway source	*P*
T cell receptor β locus	U66059	T cell anergy	BBID	< 0.0001
CD3D antigen, δ polypeptide (TIT3 complex)	NM_000732	Activation of CSK by cAMP-dependent protein kinase, IL-12– and STAT4-dependent signaling in TH1 development, CTL-mediated immune response, IL-17 signaling, HIV-induced T cell apoptosis, T cell receptor signaling, LCK and FYN tyrosine kinases in initiation of TCR activation, T cell receptor and CD3 complex, cytotoxic T cell surface molecules, T helper cell surface molecules, role TOB in T cell activation	BioCarta	< 0.0001
Complement factor B	AF349679	Complement	BioCarta	< 0.0001
Neuregulin 1	L12260	Neuregulin receptor degradation protein 1	BioCarta	< 0.0001
Arginine vasopressin (neurophysin II, antidiuretic hormone, diabetes insipidus, neurohypophyseal)	NM_000490	Control of skeletal myogenesis by HDAC and calcium/calmodulin-dependent kinase	BioCarta	< 0.0001
Aspartate β-hydroxylase	NM_020164	Hypoxia-inducible factor in the cardiovascular system	BioCarta	< 0.0001
Interleukin 3 receptor α (low affinity)	NM_002183	Regulation of BAD phosphorylation, IL-3 signaling	BioCarta	< 0.0001
Killer cell immunoglobulin-like receptor, 2 domains, long cytoplasmic tail 4	AF276292	Antigen processing and presentation, NK cell–mediated cytotoxicity	KEGG	< 0.0001
Cathepsin G	NM_001911	Neuroactive ligand-receptor interaction	KEGG	< 0.0001
Sema domain, immunoglobulin domain (IG), short basic domain, secreted (semaphorin) 3F	U38276	Axon guidance	KEGG	< 0.0001
CD22 antigen	X59350	Cell adhesion molecule, hematopoietic cell lineage, B cell receptor signaling	KEGG	< 0.0001
Cholinergic receptor, muscarinic 5	NM_012125	Calcium signaling, neuroactive ligand-receptor interaction, regulation of actin cytoskeleton	KEGG	< 0.0001
CD19 antigen	NM_001770	Hematopoietic cell lineage, B cell receptor signaling	KEGG	< 0.0001
Myo-inositol oxygenase	AL096767	Inositol phosphate metabolism	KEGG	< 0.0001
Exostoses (multiple) like 1	NM_004455	Heparin sulfate biosynthesis	KEGG	< 0.0001
CD3D antigen, δ polypeptide (TIT3 complex)	NM_000732	Hematopoietic cell lineage, T cell receptor signaling	KEGG	< 0.0001
Solute carrier family 8 (sodium-calcium exchanger), member 2	NM_015063	Calcium signaling	KEGG	< 0.0001
Mannosyl (β-1,4-)-glycoprotein β-1,4-N-acetylglucosaminyltransferase	AL022312	N-glycan biosynthesis	KEGG	< 0.0001
Complement factor B	AF349679	Complement and coagulation cascades	KEGG	< 0.0001
Acetylcholinesterase (Yt blood group)	M55040	Glycerophospholipid metabolism	KEGG	< 0.0001
γ-Aminobutyric acid (GABA) A receptor, β3	NM_000814	Neuroactive ligand-receptor interaction	KEGG	< 0.0001
Prostaglandin D2 synthase 21 kDa (brain)	BC005939	Arachidonic acid metabolism	KEGG	< 0.0001
Killer cell immunoglobulin-like receptor, 3 domains, long cytoplasmic tail 2	AF263617	Antigen processing and presentation, NK cell–mediated cytotoxicity	KEGG	< 0.0001
Arginine vasopressin (neurophysin II, antidiuretic hormone, diabetes insipidus, neurohypophyseal)	NM_000490	Neuroactive ligand-receptor interaction	KEGG	< 0.0001
Desmoglein 2	NM_001943	Cell communication	KEGG	< 0.0001
Arylsulfatase A	NM_000487	Glycophospholipid metabolism	KEGG	< 0.0001
Galactose-3-O-sulfotransferase 4	NM_024637	Cysteine metabolism	KEGG	< 0.0001
Interferon, α2	M54886	Cytokine-cytokine receptor interaction, antigen processing and presentation, TLR signaling, JAK-STAT signaling, NK cell–mediated cytotoxicity	KEGG	< 0.0001
Interleukin 3 receptor α (low affinity)	NM_002183	Cytokine-cytokine receptor interaction, apoptosis, JAK-STAT signaling, hematopoietic cell lineage	KEGG	< 0.0001
ST6 β-galactosamide α-2,6-sialyl transferase 1	X17247	N-glycan biosynthesis	KEGG	< 0.0001
γ-Glutamyltransferase-like activity 1	NM_004121	Taurine, hypotaurine, selenoamino acid, cyanoamino, glutathione, and arachidonic acid metabolism	KEGG	< 0.0001
Calcium channel, voltage-dependent α1I subunit	AL022312	MAPK signaling, calcium signaling	KEGG	< 0.0001
Natriuretic peptide receptor A/Guanlyate cyclase A (atrionatriuretic peptide receptor A)	NM_000906	Purine metabolism, gap junction, long-term depression	KEGG	< 0.0001
Frizzled homolog 6 (*Drosophila*)	NM_003506	WNT signaling	KEGG	< 0.0001
Somatostatin receptor 2	BC000256	Neuroactive ligand-receptor interaction	KEGG	< 0.0001
Defensin α4, corticostatin	NM_001925	Neuroactive ligand-receptor interaction	KEGG	< 0.0001

Gene	Accession no.	Pathway	Pathway source	*P*
Chemokine (C motif) ligand 1	U23772	Cytokine-cytokine receptor interaction	KEGG	< 0.0001
Natural cytotoxicity triggering 2 receptor	NM_004828	NK cell–mediated cytotoxicity	KEGG	< 0.0001
Receptor-associated protein of the synapse, 43 kDa	BC004196	Regulation of apoptosis, postsynaptic regulation	BioCarta	0.0007
Cytochrome p450, family 2, subfamily E, polypeptide 1	AF182276	Mechanism of acetaminophen activity and cytotoxicity, nuclear receptors in lipid metabolism and toxicity	BioCarta	0.0007
Spectrin α, non-erythrocytic 1 (α-fodrin)	AK026484	Negative effector of FAS and TNF, synaptic proteins at the synaptic junction, induction of apoptosis, FAS signaling (CD95), TNF receptor 1 signaling, uCalpain and friends in cell spread	BioCarta	0.0007
Spectrin β, non-erythrocytic 1	NM_178313	Synaptic proteins at the synaptic junction	BioCarta	0.0007
Cytochrome p450, family 2, subfamily E, polypeptide 1	AF182276	Arachidonic and linoleic acid metabolism, metabolism of xenobiotics by cytochrome P450	KEGG	0.0007
Spectrin, α, non-erythocytic 1 (α-fodrin)	AK026484	Tight junction	KEGG	0.0007
Dehydrogenase/reductase (SDR family) member 3	NM_004753	γ-Hexachlorocyclohexane, bisphenol A, 1- and 2-methylnaphthalene, ethylbenzene, limonene and pinene degradation, benzoate degradation via CoA ligation	KEGG	0.0007
Peroxisome proliferative activated receptor δ	BC002715	Nuclear receptors in lipid metabolism and toxicity, basic mechanism of action PPARA, WNT signaling	BioCarta	0.002
Disable homolog 1 (*Drosophila*)	NM_021080	Reelin signaling	BioCarta	0.002
Par-3 partitioning defective 3 homolog (*C. elgans*)	NM_019619	Neuroactive ligand-receptor interaction, adherens and tight junctions	KEGG	0.002
Peroxisome proliferative activated receptor δ	BC002715	WNT signaling	KEGG	0.002
Neuropilin 1	NM_001024628	Axon guidance	KEGG	0.002
CD19 antigen	NM_001770	Cell development	BBID	0.002
Musculin (activated B cell factor 1)	AF060154	Cell activation	BBID	0.004
Upstream transcription factor 2, c-fos interacting	NM_003367	Early response to FAC mast cells	BBID	0.004
Activating transcription factor 4 (TAX-responsive enhancer element B67)	AL022312	Oxidative stress–induced gene expression via NRF2	BioCarta	0.02
Fanconi anemia, complementation group L	NM_018062	Tryptophan metabolism	KEGG	0.02
Acetyl–coenzyme A carboxylase β	NM_001093	Fatty acid biosynthesis, pyruvate and propanoate metabolism, insulin signaling, adipocytokine signaling	KEGG	0.02
Myotubularin related protein 2	NM_016156	Fructose, mannose, thiamine, riboflavin, vitamin B6, nicotinate, and nicotinamide metabolism	KEGG	0.02
Mevalonate (diphospho) decarboxylase	NM_002461	Biosynthesis of steroids	KEGG	0.02
Glutamate dehydrogenase 1	BC005111	Arginine catabolism	BioCarta	0.05
Eukaryotic translation initiation factor 2B, subunit 5ε, 82 kDa	U23028	Regulation of EIF2, AKT/MTOR, VEGF	BioCarta	0.05
Glutamate dehydrogenase 1	BC005111	Urea cycle and metabolism of amino sugars, D-glutamate, arginine, proline, and nitrogen	KEGG	0.05
Serine hydroxymethyltransferase 1 (soluble)	Y14488	Glycine, serine, threonine, cyanoamino, acid, and methane metabolism, lysine degradation, one carbon pool by folate	KEGG	0.05
S-phase kinase-associated protein 1A (p19A)	NM_006930	Ubiquitination, cell cycle	BBID	0.10
S-phase kinase-associated protein 1A (p19A)	NM_006930	ER-associated degradation, cyclin E destruction, regulation of p27 phosphorylation during cell cycle progression, E2F1 destruction	BioCarta	0.10
S-phase kinase-associated protein 1A (p19A)	NM_006930	Cell cycle, ubiquitin-mediated proteolysis, WNT signaling, TGF-β signaling	KEGG	0.10
Protoporphyrinogen oxidase	NM_000309	Porphyrin and chlorophyll metabolism	KEGG	0.14
Pyruvate kinase, liver, and rbc	M15465	Glycolysis	BioCarta	0.15
Pyruvate kinase, liver, and rbc	M15465	Glycolysis/gluconeogenesis, purine and pyruvate metabolism, carbon fixation, insulin signaling, type II diabetes mellitus, maturity onset diabetes	KEGG	0.15
Adenosine monophosphate deaminase (isoform E)	NM_000480	Purine metabolism	KEGG	0.15
Keratin, hair, basic 4	Y19209	Cell communication	KEGG	0.23
Dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 4	AF263541	Inositol phosphate, nicotinate, and nicotinamide metabolism, benzoate degradation via CoA ligation, phosphatidylinositol signaling	KEGG	0.30
Nuclear receptor co-repressor 2	S83390	CARM1 and regulation of the estrogen receptor, MAPK inactivation of SMRT co-repressor	BioCarta	0.32
Nuclear receptor co-repressor 2	S83390	NOTCH signaling	KEGG	0.32
Cyclin B2	NM_004701	Cell cycle, breast tumor growth	BioCarta	0.39
Cyclin B2	NM_004701	Cell cycle	KEGG	0.39
Adenylate kinase 1	NM_000476	Purine metabolism, cholera infection	KEGG	0.53

BBID, Biological Biochemical Image Database; KEGG, Kyoto Encyclopedia of Genes and Genomes.

Table 8
Patient demographic information

Subject	Maternal age (yr)	Ethnicity	Reproductive history	Preexisting medical conditions	Gestational age (wk)	Pregnancy complications	Reason for Cesarean section	Gender
1	22	African American	G2P1	None	39	None	Repeat	Female
2	32	Caucasian	G3P2	None	39	None	Repeat	Female
3	31	Asian	G2P1	Graves' disease, thyroidectomy	39	None	Repeat	Male
4	39	Caucasian	G1P0	Carpel tunnel syndrome, gastroesophageal reflux	36.9	Discordant growth of twin B	Multiple gestation	Male, Male
5	32	African American	G3P2	None	39	History of prior preterm labor, treated with Progesterone IM	Repeat	Male
6	32	Caucasian	G1P0	Heart murmur, mild mitral regurgitation	37.5	History of uterine fibroids, uterine myomectomy	Labor with prior uterine surgery	Male
7	24	African American	G4P2	Tobacco use	39.7	Higher than average MSAFP results (at risk for Down syndrome)	Repeat	Male
8	25	Asian	G2P1	None	39.3	None	Repeat	Female
9	37	Caucasian	G6P5	Obesity	39.3	Diet-controlled gestational diabetes	Repeat	Male

MSAFP, maternal serum α-fetoprotein test; G, gravidity; P, parity.

case of the twin pregnancy, each twin was compared to their mother as if there were 2 independent maternal-fetal pairs. The *P* values of these tests were adjusted for multiple testing using the Benjamini-Hochberg false discovery rate approach (27). Candidate fetal biomarkers were selected if they differed significantly (adjusted *P* values of less than 0.05) in both comparisons (antepartum vs. postpartum and umbilical vs. postpartum) and if the expression levels were lower in the postpartum samples in both cases (corresponding to having paired *t* scores less than 0 in each case). There were 157 transcripts that met these criteria.

Publicly available databases (OMIM, http://www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM; PubMed, http://www.ncbi.nlm.nih.gov/sites/entrez; and NetAffx, http://www.affymetrix.com/analysis/index.affx) describing the functional role and expression pattern of all statistically significant gene transcripts (Supplemental Table 2) were then manually reviewed to determine whether each gene potentially played a role in fetal development and well being.

Further information regarding the developmental profiles of the 157 transcripts was obtained from the publicly available database UniGene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=unigene). Any transcript that had either been exclusively identified in the embryo or highly associated with embryos, fetuses, or neonates was identified as a candidate fetal gene transcript (qualifier). From the combined review of the annotation sources above and expression patterns in UniGene, a list of 71 candidate fetal markers was selected.

To determine if the gene transcripts that met statistical significance were consistent across maternal-fetal pairs, normalized log-transformed data were further analyzed (Supplemental Table 5). For each transcript, antepartum, umbilical, and postpartum values were examined across each maternal-fetal pair. Transcripts were considered positive if both the antepartum and umbilical values were greater than the postpartum sample. In rare instances, if either the antepartum or umbilical value was higher than the postpartum sample, but the other value was equal to the postpartum sample, this too was considered positive. The twin gestation was considered as 2 independent maternal-fetal pairs. Expression count values can be found in Tables 2–6 and Supplemental Table 1.

Pathway analysis. The 157 transcripts were also analyzed for potential involvement in biological pathways. Using DAVID 2007 software (http://david.abcc.ncifcrf.gov/), the probability that each transcript would be found within a pathway compared with chance alone was calculated. These calculations were also done using pathways from Kyoto Encyclopedia of Genes and Genomes, BioCarta, and Biological Biochemical Image Database. The results are shown in Table 7.

Analysis of placental genes in the plasma samples. Plasma microarray data were analyzed similarly to whole blood data as described above. A subset of 50 transcripts expressed by full-term placentas (7) was then selectively analyzed to determine if different expression profiles existed between whole blood and plasma obtained at the same time point. Comparisons between the whole blood and plasma array data were performed using paired Student's *t* tests, adjusted as described above. χ^2 tests were used to ensure that the fraction of genes that were higher in plasma in the placental gene set was not simply an array normalization artifact.

Acknowledgments

The support and assistance of Yue Shao, Jim Foster, Victor Wong, Mamta Mohan, Xian Adiconis, and Helen Stroh are gratefully acknowledged. This work was funded by NIH grants R01 HD42053 to D.W. Bianchi and 5T15 LM007092 and R01 HG003354 to M. Ramoni.

Received for publication August 4, 2006, and accepted in revised form July 2, 2007.

Address correspondence to: Jill L. Maron, Tufts – New England Medical Center, 750 Washington Street, Boston, Massachusetts 02111, USA. Phone: (617) 636-1468; Fax: (617) 636-1469; E-mail: jmaron@tufts-nemc.org.

1. Poon, L.L.M., Leung, T.N., Lau, R.K., and Lo, Y.M.D. 2000. Presence of fetal RNA in maternal plasma. *Clin. Chem.* **46**:1832–1834.
2. Lo, Y.M.D., et al. 1998. Quantitative analysis of fetal DNA in maternal plasma and serum: implications for noninvasive prenatal diagnosis. *Am. J. Hum. Genet.* **62**:768–775.
3. Costa, J.M., et al. 2003. Fetal expressed gene analysis in maternal blood: A new tool for noninvasive study of the fetus. *Clin. Chem.* **49**:981–983.
4. Qinyu, G., Quanjun, L., Yunfei, B., Wen, T., and Lu, Z. 2005. A semi-quantitative microarray method to detect fetal RNAs in maternal plasma. *Prenat. Diagn.* **25**:912–918.
5. Ng, E.K.O., et al. 2003. mRNA of placental origin is readily detectable in maternal plasma. *Proc. Natl. Acad. Sci. U. S. A.* **100**:4748–4753.
6. Wataganara, T., et al. 2004. Circulating cell-free fetal nucleic acid analysis may be a novel marker of fetomaternal hemorrhage after elective first-trimester termination of pregnancy. *Ann. N. Y. Acad. Sci.* **1022**:129–134.
7. Tsui, N.B.Y., et al. 2004. Systematic micro-array based identification of placental mRNA in maternal plasma: towards non-invasive prenatal gene expression profiling. *J. Med. Genet.* **41**:461–467.
8. Ng, E.K.O., et al. 2003. The concentration of circulating corticotrophin-releasing hormone mRNA in maternal plasma is increased in preeclampsia. *Clin. Chem.* **49**:727–731.
9. Rainen, L., et al. 2002. Stabilization of mRNA expression in whole blood samples. *Clin. Chem.* **48**:1883–1890.
10. Chiu, R.W.K., et al. 2006. Time profile of appearance and disappearance of circulating placenta-derived mRNA in maternal plasma. *Clin. Chem.* **52**:313–316.
11. Tsui, N.B., Ng, E.K., and Lo, Y.M. 2002. Stability of endogenous and added RNA in blood specimens, serum, and plasma. *Clin. Chem.* **48**:1647–1653.
12. Dennis, G., Jr., et al. 2003. DAVID: Database for annotation, visualization, and integrated discovery. *Genome Biol.* **4**:P3.
13. Wainwright, M.S., et al. 2004. Increased susceptibility of S100B transgenic mice to perinatal hypoxia-ischemia. *Ann. Neurol.* **56**:61–67.
14. Kelemen, K., Paldi, A., Tinneberg, H., Torok, A., and Szekeres-Bartho, J. 1998. Early recognition of pregnancy by the maternal immune system. *Am. J. Reprod. Immunol.* **39**:351–355.
15. Wood, G.W. 1994. Is restricted antigen presentation the explanation for fetal allograft survival? *Immunol. Today.* **15**:15–18.
16. Bianchi, D.W., Flint, A.F., Pizzimenti, M.F., Knoll, J.H., and Latt, S.A. 1990. Isolation of fetal DNA from nucleated erythrocytes in maternal blood. *Proc. Natl. Acad. Sci. U. S. A.* **87**:3279–3283.
17. Bianchi, D.W., Zickwolf, G.K., Weil, G.J., Sylvester, S., and DeMaria, M.A. 1996. Male fetal progenitor cells persist in maternal blood for as long as 27 years postpartum. *Proc. Natl. Acad. Sci. U. S. A.* **93**:705–708.
18. Affymetrix Inc. 2005. Globin reduction protocol: A method for processing whole blood RNA samples for improved array results. http://www.affymetrix.com.
19. Bukowski, R., Hankins, G.D.V., Saade, G.R., Anderson, G.D., and Thornton, S. 2006. Labor-associated gene expression in the human uterine fundus, lower segment, and cervix. *PLoS. Med.* **3**:918–930.
20. Shi, L., et al. 2006. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat. Biotechnol.* **24**:1151–1161.
21. Maron, J.L., and Bianchi, D.W. 2007. Prenatal diagnosis using cell-free nucleic acids in maternal body fluids: a decade of progress. *Am. J. Med. Genet.* **145**:5–17.
22. Ng, E.K.O., et al. 2002. Presence of filterable and nonfilterable mRNA in the plasma of cancer patients and healthy individuals. *Clin. Chem.* **48**:1212–1217.
23. Van Gelder, R.N., et al. 1990. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc. Natl. Acad. Sci. U. S. A.* **87**:1663–1667.
24. Livak, K.J. 1999. Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *Genet. Anal.* **14**:143–149.
25. Gentleman, R.C., et al. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**:R80.
26. Draghici, S., Khatri, P., Martins, R.P., Ostermeier G.C., and Krawetz, S.A. 2003. Global functional profiling of gene expression. *Genomics.* **81**:98–104.
27. Benjamini, Y., et al. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. (Ser. A).* **B75**:290–300.