NIH Public Access

Author Manuscript

Obstet Gynecol. Author manuscript; available in PMC 2013 January 1.

Published in final edited form as:

Obstet Gynecol. 2012 January ; 119(1): 111–118. doi:10.1097/AOG.0b013e31823d4150.

The Amniotic Fluid Transcriptome: A Source of Novel Information About Human Development

Lisa Hui, MBBS[1,2,4], **Donna K. Slonim, PhD**[3], **Heather C. Wick, BSc**[3], **Kirby L. Johnson, PhD**[1,2], and **Diana W. Bianchi, MD**[1,2]

[1]Mother Infant Research Institute, Tufts Medical Center, Boston, MA 02111, USA

[2]Division of Genetics, Department of Pediatrics, The Floating Hospital for Children at Tufts Medical Center, Boston, MA 02111 USA

[3]Department of Computer Science, Tufts University, Medford, MA 02155, USA

[4]Discipline of Obstetrics, Gynaecology, and Neonatology, University of Sydney, Sydney, NSW 2006, Australia

Abstract

Objective—Amniotic fluid (AF) is a complex biological material that provides a unique window into the developing human. Residual AF supernatant contains cell-free fetal RNA. The objective of this study was to develop an understanding of the AF core transcriptome by analyzing the transcripts ubiquitously present in the AF supernatant of euploid midtrimester fetuses.

Methods—This was an *in silico* (computational) investigation using publicly available gene expression data previously produced by our group from 12 euploid midtrimester amniotic fluid samples. Functional analyses were performed using a web-based software analysis tool. Organ specificity was examined for each transcript using a gene expression atlas. For fetal organs not represented in the atlas, manual literature searching and the web-based software analysis tool were used to generate fetal organ-associated gene lists.

Results—There were 476 well-annotated genes present in 12 out of 12 AF samples. Functional analysis identified six physiological systems represented in the AF core transcriptome, including musculoskeletal and nervous system development and function and embryonic and organism development. Mammalian target of rapamycin signaling was identified as a key canonical pathway. Twenty-three highly organ-specific transcripts were identified; six of these are known to be highly expressed in fetal brain.

Conclusions—AF cell-free fetal RNA can provide biological information on multiple fetal organ systems. The presence of fetal-brain specific transcripts in AF suggests novel approaches to the study of developmental disorders that involve the central nervous system. The finding that mammalian target of rapamycin signaling is enriched in midtrimester fetuses may have future applications in the study of fetal growth disorders.

Corresponding author address: Lisa Hui, Mother Infant Research Institute, 800 Washington St, 9th floor Tupper Building, Boston MA 02111, Ph. +1-617-636-1468, lhui@tuftsmedicalcenter.org.

Additional information related to this article can be found in Appendices 1–12, available online at http://links.lww.com/xxx.

The preliminary results of this study were presented at the 58th annual meeting of the Society for Gynecological Investigation in Miami, 16–19th March 2011.

Financial Disclosure

The authors did not report any potential conflicts of interest.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Introduction

Amniotic fluid (AF) is a complex biological material that provides a unique window into the developing human. It is sampled for the prenatal assessment of many conditions, most commonly to obtain amniocytes for chromosome analysis. The residual AF supernatant, which is usually discarded, is a rich source of fetal cell-free DNA and RNA. Cell-free nucleic acids in AF are distinct from those in maternal blood. Firstly, AF cell-free nucleic acids are more likely to originate from the fetus itself, in contrast to circulating cell-free fetal nucleic acids, which are predominantly of trophoblast origin (1, 2). Secondly, AF cell-free fetal nucleic acids are relatively uncontaminated by maternal nucleic acids, as maternal-fetal nucleic acid trafficking is overwhelmingly unidirectional from fetus to mother (3). Finally, they are 100–200-fold more abundant in AF compared with maternal plasma, making downstream applications such as global gene expression profiling more feasible (4).

The first aim of this study was to identify transcripts that are ubiquitously present in a euploid mid-trimester AF supernatant dataset. We called this gene list the AF core transcriptome. Functional analyses were then performed on this gene list to infer the major organ systems that contribute to AF cell-free fetal RNA, and to determine the pathways of biological significance in the mid-trimester fetus. Our second aim was to identify the likely tissue sources of AF cell-free fetal RNA by using the GNF Gene Expression Atlas to map organ-specific genes. However, several important fetal organs, such as heart, kidney, intestine and skin, are not represented in the GNF Gene Expression Atlas. Therefore, our third aim was to create gene lists associated with the development of these organs, and to ascertain the presence of these genes in AF supernatant. The unifying goal of these three aims was to produce clinically-relevant information about mid-trimester gene expression in the living human fetus.

Materials and Methods

This was an *in silico* (computational) investigation using gene expression data previously produced by our group from 12 mid-trimester AF samples. This sample size is within the range of biological replicates (10–15) recommended to achieve statistical stability and confidence in microarray experiments (5). The Tufts Medical Center Institutional Review Board approved the collection and analysis of the samples. These data are publicly available at www.ncbi.nlm.nih.gov/geo/ (GSE 33168). Eleven of the 12 samples were originally used as euploid controls in studies focusing on the pathophysiology of trisomies 18 and 21 (6, 7). In the present study, we reanalyzed the data from these 11 cases together with a previously unpublished sample to better understand gene expression in euploid fetuses. The methods of AF processing, RNA extraction, fragmentation, labeling, and hybridization have been previously described (7). In brief, RNA was extracted from 10ml of AF supernatant from 12 women undergoing fetal testing for clinical indications. Eleven of the women had testing for advanced maternal age (median maternal age 36, range 23–40 years). Three of these women had abnormal second trimester serum screening results. One fetus was a twin (chorionicity unknown). None of the fetuses had a known major sonographically-detectable anomaly at the time of amniocentesis. AF supernatant samples from 6 male and 6 female euploid fetuses (median gestational age 18 weeks, range 16 to 21 weeks) were included. The final amplified cDNA products were hybridized to Affymetrix U133 Plus 2.0 microarrays (Affymetrix, Santa Clara, CA.).

Data were normalized using the three step function from the affyPLM package in Bioconductor (version 2.8.1) (8), with ideal-mismatch background/signal adjustment, quantile normalization, and the Tukey biweight summary method (9). This summary method includes a logarithmic transformation. To obtain detection calls consistent with those

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

produced by Affymetrix' 5.0 software, we used the mas5calls function from the Bioconductor affy package.

We defined the AF core transcriptome as the list of genes corresponding to those Affymetrix probe sets that were called "present" in all twelve of the AF samples. A probe set was called "present" if it had a detection *P* value of < 0.04 (10), consistent with other published research (11).

The web-based software tool Ingenuity Pathways Analysis 9.0 (content version 3210) was used for the biological interpretation of the AF core transcriptome gene list. This pathways analysis tool uses a manually-curated repository of biological interactions and functional annotations to identify the most significantly enriched signaling pathways and biological processes represented in a given gene set. The enriched pathways for the categories "Cellular and Molecular Functions," "Physiological Systems Development and Function," and "Canonical Pathways" were reported separately. The pathways analysis software uses the right-tailed Fisher's exact test to calculate a *P* value representing the probability that a biological function not really relevant to the AF core transcriptome is reported as relevant. Pathways that contained at least one functional annotation with a *P* value < 0.01 were considered statistically significant. In addition, we applied a multiple testing correction using the Benjamini-Hochberg approach, which bounds the false discovery rate (FDR). Here, we report all biological functions with Fisher *P* < 0.01 and their corresponding FDRs.

We identified organ-specific genes in the AF core transcriptome using the Novartis Research Foundation Gene Expression Database (the GNF Gene Expression Atlas at http://biogps.gnf.org). This publicly-available atlas of the human protein-encoding transcriptome used Affymetrix Human Genome-U133A and GNF1H custom human arrays to map gene expression profiles in 78 normal human tissues (12). We chose this resource because of its coverage of normal adult and fetal tissues, high reproducibility and good correlation between transcript levels and protein abundance (13). Samples from brain, liver, and lung samples were pooled from spontaneously aborted fetuses (15–33 weeks gestation); the placentas were collected at birth. No gestational age data were available for the fetal thyroid and placental samples.

The GNF Gene Expression Atlas allowed us to assess the gene expression patterns of individual Affymetrix probe sets in the AF core transcriptome. Data were downloaded on September 2, 2011. We categorized probe sets as highly organ-specific if they mapped to a single organ with an expression value > 30 multiples of the median (MoM) and had no unrelated tissue expression greater than one third of the maximum expression level.

As several important fetal organs were not represented in the GNF Atlas, we used a combination of manual literature searching and pathways analysis software to generate lists of genes that we would expect to be expressed in mid-trimester fetal heart, kidney, bladder, skin, intestine, placenta and amnion. We then examined the AF core transcriptome for the presence of these genes as an indirect method of ascertaining potential organ contributions to AF cell-free fetal RNA. A full description of the methods used in the literature and pathways analysis searches can be found in Appendix 1, available online at http://links.lww.com/xxx.

Results

The average present call rate for all 54,675 gene probe sets contained in the dataset was 19.8%. The total number of probe IDs that were expressed in 12/12 samples was 796 (1.46%). After excluding duplicate genes due to multiple probe sets, hypothetical genes, pseudogenes, non-protein coding genes, and genes with unknown functions, the final

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

number of well-annotated genes in the AF core transcriptome was 476 (Appendix 2, available online at http://links.lww.com/xxx). The three most common molecule types within the AF core transcriptome were enzymes, ribosomal proteins and transcription regulators.

Functional analysis of the AF core transcriptome identified six physiological systems development and function pathways, 11 canonical pathways, and 27 cellular and molecular pathways with a Fisher P value < 0.01 (Tables I, II, III). The most significantly over-represented physiological systems were skeletal and muscular system development and function, tissue development and hematological system development and function (FDR < 0.05). Individual functional annotations within these major categories include growth of muscle and developmental process of cardiac muscle. Biological processes associated with nervous system development and function were also over-represented. The most significant functional annotation in the nervous system category was neurogenesis of stem cells (P = 0.03, FDR = 0.09). The mammalian target of rapamycin (mTOR) signaling pathway, an important central regulator of cell growth and nutrient sensing, was identified as a key canonical pathway. The most significant cellular and molecular pathway was protein synthesis.

After examining the tissue expression patterns of each of the genes in the AF core transcriptome using the GNF Gene expression atlas, we identified 23 highly organ-specific genes associated with the brain, spinal cord, lung, pancreas, liver, tongue, thyroid, placenta and blood (Table IV). Seven of these genes are highly expressed by the fetal and adult central nervous system (CNS). The plakophilin 4 gene *(PKP4)* mapped to adult spinal cord, but not fetal brain. Fetal spinal cord was not represented in the GNF Atlas. We therefore confirmed *PKP4* expression in fetal spinal cord in a GEO dataset for two second trimester normal spinal cord samples (GEO accession no. GSE1481).

Fetal lung was represented in AF by the surfactant protein genes *SFTPB* and *SFTPC*, and the bronchial epithelial cell gene *stratifin*. This is consistent with both mid-trimester fetal physiology, and prior results from our laboratory (14). Gastrointestinal-specific transcripts for coagulation factor 7 and pancreatic trypsin were also present in the AF; these are known to be expressed in the mid-trimester human liver and pancreas respectively.

The small proline-rich protein genes *SPRR1B* and *SPRR2B* were found in the AF core transcriptome. We propose that these transcripts are likely to derive from the fetal skin, even though fetal skin is not represented in the GNF Atlas. These small proline-rich proteins are involved with formation of the cell envelope in keratinocytes and are highly expressed in differentiating squamous epithelium, including fetal skin. Postnatally, *SPRR* genes are known to be highly expressed in the squamous epithelium of the upper gastrointestinal tract, as reflected the GNF Atlas results (adult tongue, tonsil)(15).

We created lists of genes associated with developing heart, kidney, bladder, intestine, skin, placenta and amnion based on our Medline and pathways analysis software searches. A major caveat to these results is that most of the genes identified in the literature and pathways analysis software searches were not tissue-specific in their expression pattern, and therefore could originate from more than one fetal tissue. The analysis is summarized in Appendices 6–12, available online at available online at http://links.lww.com/xxx. These gene lists and literature references are an additional source of functional annotation for our ongoing project characterizing gene function in human development (http://dflat.cs.tufts.edu), and represent an additional resource for future studies of gene expression in human fetuses.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Discussion

In the present study we have defined the AF core transcriptome, which is comprised of 476 well-annotated genes consistently expressed in the AF supernatant of euploid mid-trimester fetuses. The multiple organ-specific transcripts and physiological systems detected in the AF core transcriptome provide strong evidence that AF cell-free fetal RNA originates from more than one cell type. Moreover, both the functional analysis and GNF Atlas mapping suggest that AF contains gene derived from the fetal nervous system. The presence of these CNS genes in the AFCT was unexpected, given the lack of direct physical contact between the fetal brain and the amniotic fluid. We speculate that these transcripts may enter the AF directly from CNS regions with relatively thin tissue barriers such as the olfactory mucosa, the tympanic membrane, or cerebral fontanelles. Alternatively, CNS genes may enter the AF indirectly through the fetal cardiovascular circulation. Another possibility is that they may be transcripts that were directly released from the CNS into the AF prior to neural tube closure during the first trimester and persisting long-term. Despite our lack of knowledge about their precise route of entry and their *in vivo* half-life, these fetal-brain specific transcripts may be a novel source of biomarkers of nervous system development that are uniquely accessible from living fetuses.

Evidence that mTOR signaling is active in mid-trimester human fetuses is another significant finding in live human pregnancies. The mTOR pathway is a central regulator of cell growth, integrating extracellular signals from amino acids, growth factors such as IGF-1, energy, and stress (16). This pathway is increasingly recognized as an important placental mechanism for influencing fetal growth in response to nutrient availability (17, 18). Prior work from our laboratory has documented the presence of the mTOR gene in mid-trimester AF using NanoArray PCR (19). The present study confirms that the mTOR signaling pathway as a whole is significantly overrepresented in the AF core transcriptome. This illustrates the ability of AF cell-free fetal RNA to contribute translational human data to results from animal and *in vitro* work. The detection of fetal mTOR signaling in AF may have clinical significance in future studies of abnormal fetal growth.

The presence of lung gene transcripts for the surfactant proteins B and C was an expected finding given established knowledge about fetal lung development (20), the contribution of lung fluid to AF volume, and previous research on AF cell-free fetal RNA (14). The detection of organ-specific genes from the liver and pancreas, and the less tissue-specific mucin genes associated with the intestine are similarly biologically plausible, given their developmental expression and direct communication between the fetal gastrointestinal tract and the amniotic cavity (21–24).

Surprisingly, we did not find significant enrichment of renal system development and function in the pathways analysis. There were no kidney-specific genes identified using the GNF Atlas. This is despite the substantial contribution of fetal urine to AF, and the presence of cell-free fetal nucleic acids in maternal urine (25). While there were some kidney and bladder-associated genes in the AF core transcriptome identified from the literature and pathways analysis software searches, these were all non-organ specific genes that are widely expressed in all fetal tissues, such as the fibroblast growth factor receptor 1.

The absence of gene transcripts specific to kidney highlights one of the major limitations of our study. The reliance on the presence of “organ-specific genes” to identify putative organ sources is subject to an inherent bias in the GNF Atlas due to the range of tissues sampled and the variable number of organ-specific genes in each organ. The variability in tissue specificity has been quantified by Dezso *et al*, who determined that fetal brain has a relatively high number of organ-specific genes, while the fetal kidney has relatively few

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

(26). This variation may explain both the high numbers of fetal brain transcripts and the absence of fetal kidney-specific genes found in the AF core transcriptome gene list. Similarly, annotation biases may be present in the pathways analysis software database that may under-represent renal development. Presently, there are no definitive methods to determine the origins of cell-free gene transcripts within AF. Despite these limitations, we considered the tissue-expression patterns of the universally-expressed genes and the functional analyses as valuable approaches to understanding the biology of AF cell-free fetal RNA.

Another limitation of our study is the lack of clinical follow up on the women that contributed the AF supernatant samples. The most common indication for amniocentesis was advanced maternal age. While all of the fetuses were euploid, we were not able to collect long-term obstetric outcome data as our samples were anonymized. Our results should therefore be interpreted with caution, as they may not be completely representative of uncomplicated pregnancies.

Clinically, there are numerous potential applications for AF cell-free fetal RNA in fetal medicine. Functional genomic analysis is a powerful tool for understanding both normal physiology and disease. Prior studies of human development have mainly relied on postnatal tissue specimens (27, 28) or extraembryonic prenatal samples (29). AF supernatant has the unique property of being a pure fetal sample that can provide gene expression information on live fetuses without posing unacceptable risks to the pregnancy. If reproducible gene expression profiles can be demonstrated for specific phenotypes, then these could be used to learn more about the pathophysiology of fetal diseases and to identify biomarkers. This approach has already been successfully used to study fetuses with trisomies 18 and 21 and has generated candidate therapies for translational research (6, 7). Furthermore, the presence of fetal brain-specific transcripts suggests novel approaches to the study of developmental disorders that involve the CNS.

In conclusion, this study provides a detailed examination of AF cell-free fetal RNA from euploid mid-trimester pregnancies and suggests future potential applications for fetal gene expression studies. Our broad approach, using multiple publicly-available resources, enabled us to discover significant biological processes represented in the AF core transciptome, determine putative organ sources of AF cell-free fetal RNA, and identify genes of specific biological interest that represent potential fetal biomarkers. This work develops the concept of AF as a “summary fluid” derived from multiple organs and lays the foundation for future studies of gene expression in abnormal fetal development.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Keiko Koide MD, a former member of our laboratory, who produced the original microarray datasets; and Janet Cowan PhD, who provided the amniotic fluid supernatant samples and subject data.

Supported by the *Eunice Kennedy Shriver* National Institute for Child Health & Human Development [R01 HD 42053-08 to Dr. Bianchi and R01 HD 058880 to D.K.S]; the University of Sydney Medical School [Albert S McKern Research Scholarship to Dr. Hui]; and the Royal Australian and New Zealand College of Obstetricians and Gynaecologists Research Foundation [Fotheringham Fellowship to Dr. Hui.]. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the funding institutions.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

References

1. Orozco AF, Bischoff FZ, Horne C, Popek E, Simpson JL, Lewis DE. Hypoxia-induced membrane-bound apoptotic DNA particles: potential mechanism of fetal DNA in maternal plasma. Ann N Y Acad Sci. 2006; 1075:57–62. [PubMed: 17108192]
2. Maron JL, Bianchi DW. Prenatal diagnosis using cell-free nucleic acids in maternal body fluids: a decade of progress. Am J Med Genet C Semin Med Genet. 2007; 145C:5–17. [PubMed: 17299735]
3. Sekizawa A, Yokokawa K, Sugito Y, Iwasaki M, Yukimoto Y, Ichizuka K, et al. Evaluation of bidirectional transfer of plasma DNA through placenta. Hum Genet. 2003; 113:307–10. [PubMed: 12884005]
4. Bianchi DW, LeShane ES, Cowan JM. Large amounts of cell-free fetal DNA are present in amniotic fluid. Clin Chem. 2001; 47:1867–9. [PubMed: 11568107]
5. Pavlidis P, Li Q, Noble WS. The effect of replication on gene expression microarray experiments. Bioinformatics. 2003; 19:1620–7. [PubMed: 12967957]
6. Koide K, Slonim DK, Johnson KL, Tantravahi U, Cowan JM, Bianchi DW. Transcriptomic analysis of cell-free fetal RNA suggests a specific molecular phenotype in trisomy 18. Hum Genet. 2011; 129:295–305. [PubMed: 21152935]
7. Slonim DK, Koide K, Johnson KL, Tantravahi U, Cowan JM, Jarrah Z, et al. Functional genomic analysis of amniotic fluid cell-free mRNA suggests that oxidative stress is significant in Down syndrome fetuses. Proc Natl Acad Sci U S A. 2009; 106:9425–9. [PubMed: 19474297]
8. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 2004; 5:R80. [PubMed: 15461798]
9. Bolstad, BM. Low level analysis of high-density oligonucleotide array data: background, normalization and summarization. University of California; Berkeley: 2004.
10. Affymetrix. Affymetrix technical note: new statistical algorithms for monitoring gene expression on GeneChip probe Arrays. Santa Clara, CA: 2001.
11. Li Y, Zhou X, St John MA, Wong DT. RNA profiling of cell-free saliva using microarray technology. J Dent Res. 2004; 83:199–203. [PubMed: 14981119]
12. Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, et al. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A. 2004; 101:6062–7. [PubMed: 15075390]
13. Kislinger T, Cox B, Kannan A, Chung C, Hu P, Ignatchenko A, et al. Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. Cell. 2006; 125:173–86. [PubMed: 16615898]
14. Larrabee PB, Johnson KL, Lai C, Ordovas J, Cowan JM, Tantravahi U, et al. Global gene expression analysis of the living human fetus using cell-free messenger RNA in amniotic fluid. JAMA. 2005; 293:836–42. [PubMed: 15713773]
15. Lee CH, Marekov LN, Kim S, Brahim JS, Park MH, Steinert PM. Small proline-rich protein 1 is the major component of the cell envelope of normal human oral keratinocytes. FEBS Lett. 2000; 477:268–72. [PubMed: 10908733]
16. Wullschleger S, Loewith R, Hall MN. TOR signaling in growth and metabolism. Cell. 2006; 124:471–84. [PubMed: 16469695]
17. Wen H, Abbasi S, Kellems R, Xia Y. mTOR: A placental growth signaling sensor. Placenta. 2005; 26:S63–S9. [PubMed: 15837070]
18. Roos S, Powell TL, Jansson T. Placental mTOR links maternal nutrient availability to fetal growth. Biochem Soc Trans. 2009; 37:295–8. [PubMed: 19143650]
19. Massingham LJ, Johnson KL, Bianchi DW, Pei S, Peter I, Cowan JM, et al. Proof of Concept Study to Assess Fetal Gene Expression in Amniotic Fluid by NanoArray PCR. J Mol Diagn. 2011; 13:565–70. [PubMed: 21827969]
20. Liley HG, White RT, Warr RG, Benson BJ, Hawgood S, Ballard PL. Regulation of messenger RNAs for the hydrophobic surfactant proteins in human lung. J Clin Invest. 1989; 83:1191–7. [PubMed: 2539393]

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

21. Hassan HJ, Leonardi A, Chelucci C, Mattia G, Macioce G, Guerriero R, et al. Blood coagulation factors in human embryonic-fetal development: preferential expression of the FVII/tissue factor pathway. Blood. 1990; 76:1158–64. [PubMed: 1698100]
22. Sarkar SA, Kobberup S, Wong R, Lopez AD, Quayum N, Still T, et al. Global gene expression profiling and histochemical analysis of the developing human fetal pancreas. Diabetologia. 2008; 51:285–97. [PubMed: 18094957]
23. Buisine MP, Devisme L, Savidge TC, Gespach C, Gosselin B, Porchet N, et al. Mucin gene expression in human embryonic and fetal intestine. Gut. 1998; 43:519–24. [PubMed: 9824580]
24. Buisine MP, Devisme L, Maunoury V, Deschodt E, Gosselin B, Copin MC, et al. Developmental mucin gene expression in the gastroduodenal tract and accessory digestive glands. I. Stomach. A relationship to gastric carcinoma. J Histochem Cytochem. 2000; 48:1657–66. [PubMed: 11101634]
25. Shekhtman EM, Anne K, Melkonyan HS, Robbins DJ, Warsof SL, Umansky SR. Optimization of transrenal DNA analysis: detection of fetal DNA in maternal urine. Clin Chem. 2009; 55:723–9. [PubMed: 19181739]
26. Dezso Z, Nikolsky Y, Sviridov E, Shi W, Serebriyskaya T, Dosymbekov D, et al. A comprehensive functional analysis of tissue specificity of human gene expression. BMC Biol. 2008; 6:49. [PubMed: 19014478]
27. Kho AT, Bhattacharya S, Tantisira KG, Carey VJ, Gaedigk R, Leeder JS, et al. Transcriptomic analysis of human lung development. Am J Respir Crit Care Med. 2010; 181:54–63. [PubMed: 19815808]
28. Madsen-Bouterse SA, Romero R, Tarca AL, Kusanovic JP, Espinoza J, Kim CJ, et al. The transcriptome of the fetal inflammatory response syndrome. Am J Reprod Immunol. 2010; 63:73–92. [PubMed: 20059468]
29. Farina A, Morano D, Arcelli D, De Sanctis P, Sekizawa A, Purwosunu Y, et al. Gene expression in chorionic villous samples at 11 weeks of gestation in women who develop preeclampsia later in pregnancy: implications for screening. Prenat Diagn. 2009; 29:1038–44. [PubMed: 19731222]

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1

Physiological Systems Development and Function in the Amniotic Fluid Core Transcriptome

Category	No. of Genes	P^*	False Discovery Rate†
Skeletal and muscular system development and function	17	3.85×10^{-4}- 2.44×10^{-2}	2.85×10^{-2} - 1.59×10^{-1}
Tissue development	20	3.85×10^{-4}- 2.44×10^{-2}	2.85×10^{-2} - 1.59×10^{-1}
Hematological system development and function	6	7.37×10^{-4} - 2.44×10^{-2}	3.66×10^{-2} - 1.59×10^{-1}
Nervous system development and function	20	3.45×10^{-3} - 2.44×10^{-2}	9.05×10^{-2} - 1.59×10^{-1}
Embryonic development	28	9.78×10^{-3} - 2.44×10^{-2}	1.59×10^{-1} - 1.59×10^{-1}
Organism development	50	9.78×10^{-3} - 2.44×10^{-2}	1.59×10^{-1} - 1.59×10^{-1}

[*] Right-tailed Fisher *P* and

[†] false discovery rate ranges for individual functional annotations within each category.

See Appendix 3 (http://links.lww.com/xxx) for gene lists and individual functional annotations with Fisher $P < 0.01$.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Canonical Pathways in the Amniotic Fluid Core Transcriptome

Canonical Pathways	No. of Genes	P^*	False Discovery Rate
Mammalian target of rapamycin signaling	13	4.86×10^{-5}	1.21×10^{-2}
Regulation of actin-based motility by rho	9	2.05×10^{-4}	2.54×10^{-2}
Cell cycle: G2/M DNA damage checkpoint regulation	6	6.44×10^{-4}	4.15×10^{-2}
PI3K/AKT signaling	10	6.7×10^{-4}	4.15×10^{-2}
Semaphorin signaling in neurons	6	1.6×10^{-3}	7.96×10^{-2}
Integrin signaling	12	3.48×10^{-3}	1.4×10^{-1}
ERK5 signaling	6	3.95×10^{-3}	1.77×10^{-1}
Clathrin-mediated endocytosis signaling	10	5.95×10^{-3}	1.77×10^{-1}
Glycolysis and gluconeogenesis	7	6.44×10^{-3}	1.77×10^{-1}
Molecular mechanisms of cancer	16	9.09×10^{-3}	2.07×10^{-1}
p70S6K signaling	8	9.19×10^{-3}	2.07×10^{-1}

*Right-tailed Fisher *P* value

See Appendix 4 (http://links.lww.com/xxx) for canonical pathways gene lists.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table 3

Molecular and Cellular Functions in the Amniotic Fluid Core Transcriptome

Category	No. of Genes	P^*	False Discovery Rates†
Protein synthesis	80	2.44×10^{-51} - 1.91×10^{-2}	9.21×10^{-48} - 1.59×10^{-1}
RNA posttranscriptional modification	41	2.69×10^{-8} - 2.44×10^{-2}	2.03×10^{-5} - 1.59×10^{-1}
Gene expression	82	5.04×10^{-6} - 2.44×10^{-2}	1.46×10^{-3} - 1.59×10^{-1}
RNA trafficking	4	5.04×10^{-6} - 5.04×10^{-6}	1.46×10^{-3} - 1.46×10^{-3}
Cell cycle	62	9.24×10^{-6} - 2.44×10^{-2}	2.49×10^{-3} - 1.59×10^{-1}
Cell death	104	1.28×10^{-5} - 2.44×10^{-2}	3.02×10^{-3} - 1.59×10^{-1}
DNA replication, recombination, and repair	38	5.67×10^{-5} - 2.44×10^{-2}	9.11×10^{-3} - 1.59×10^{-1}
Posttranslational modification	18	1.24×10^{-4} - 2.44×10^{-2}	1.47×10^{-2} - 1.59×10^{-1}
Protein degradation	5	1.24×10^{-4} - 1.24×10^{-4}	1.47×10^{-2} - 1.47×10^{-2}
Cellular assembly and organization	60	2.73×10^{-4} - 2.44×10^{-2}	2.34×10^{-2} - 1.59×10^{-1}
Protein trafficking	24	2.86×10^{-4} - 3.45×10^{-3}	2.4×10^{-2} - 9.05×10^{-2}
Cellular growth and proliferation	117	4.35×10^{-4} - 2.44×10^{-2}	3.1×10^{-2} - 1.59×10^{-1}
Cellular movement	63	5.27×10^{-4}- 2.44×10^{-2}	3.41×10^{-2} - 1.59×10^{-1}
Molecular transport	39	5.49×10^{-4} - 2.44×10^{-2}	3.41×10^{-2} - 1.59×10^{-1}
Cell morphology	44	5.94×10^{-4} - 2.44×10^{-2}	3.41×10^{-2} - 1.59×10^{-1}
Cellular development	58	5.94×10^{-4} - 2.44×10^{-2}	3.41×10^{-2} - 1.59×10^{-1}
Energy production	3	5.94×10^{-4} - 2.44×10^{-2}	3.41×10^{-2} - 1.59×10^{-1}
Nucleic acid metabolism	13	5.94×10^{-4} - 2.44×10^{-2}	3.41×10^{-2} - 1.59×10^{-1}
Small molecule biochemistry	35	5.94×10^{-4} - 2.44×10^{-2}	3.41×10^{-2} - 1.59×10^{-1}
Antigen presentation	4	7.37×10^{-4} - 2.44×10^{-2}	3.66×10^{-2} - 1.59×10^{-1}
Cell-to-cell signaling and interaction	15	7.37×10^{-4} - 2.44×10^{-2}	3.66×10^{-2} - 1.59×10^{-1}
Cellular function and maintenance	35	8.34×10^{-4} - 2.44×10^{-2}	4.04×10^{-2} -1.59×10^{-1}
Carbohydrate metabolism	16	1.01×10^{-3} - 2.44×10^{-2}	4.4×10^{-2} -1.59×10^{-1}
Lipid metabolism	13	1.01×10^{-3} - 2.44×10^{-2}	4.4×10^{-2} -1.59×10^{-1}
Cell signaling	7	2.69×10^{-3} - 2.44×10^{-2}	7.82×10^{-2} -1.59×10^{-1}
Protein folding	6	4.3×10^{-3} - 9.34×10^{-3}	1.04×10^{-1} -1.59×10^{-1}
Cellular compromise	10	6.38×10^{-3} - 2.44×10^{-2}	1.31×10^{-1} -1.59×10^{-1}

*Right-tailed Fisher *P* and false discovery rate ranges for individual functional annotations within each category.

†false discovery ranges for individual functional annotations within each category.

See Appendix 5 (http://links.lww.com/xxx) for functional annotations with false discovery ranges < 0.05.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table 4

Organ-Specific Genes in the Amniotic Fluid Core Transcriptome

Probe Set ID	Gene Symbol	Gene Name	Tissue with maximum expression value in GNF Atlas	Entrez Gene
207300_s_at	*F7*	coagulation factor VII (serum prothrombin conversion accelerator)	adult liver	Vitamin K- dependent blood coagulation factor
207608_x_at	*CYP1A2*	cytochrome P450, family 1, subfamily A, polypeptide 2	adult liver	Member of the cytochrome P450 superfamily of enzymes
214354_x_at	*SFTPB*	surfactant protein B	adult lung	Surfactant protein essential for postnatal lung function
215454_x_at	*SFTPC*	surfactant protein C	adult lung	Surfactant protein essential for postnatal lung function
213940_s_at	*FNBP1*	formin binding protein 1	B lymphoblasts	Membrane tubulation with reorganization of the actin cytoskeleton during endocytosis
221923_s_at	*NPM1*	nucleophosmin (nucleolar phosphoprotein B23, numatrin)	B lymphoblasts	Involved in diverse cellular processes such as ribosome biogenesis, cell proliferation, and regulation of tumor suppressors p53/TP53 and ARF.
33322_i_at, 33323_r_at	*SFN*	stratifin	bronchial epithelial cells	Adapter protein. When bound to KRT17, regulates protein synthesis and epithelial cell growth
201853_s_at	*CDC25B*	cell division cycle 25 homolog B (S. pombe)	$CD4^{+}$ T cells	Tyrosine protein phosphatase which functions as a dosage-dependent inducer of mitotic progression.
217878_s_at	*CDC27*	cell division cycle 27 homolog (S. cerevisiae)	$CD71^{+}$ early erythroid	Component of the anaphase promoting complex/cyclosome
212774_at	*ZNF238*	zinc finger protein 238	cerebellum peduncles	Regulation of survival of postmitotic cortical neurons
207030_s_at	*CSRP2*	cysteine and glycine-rich protein 2	fetal brain	Development of the embryonic vascular system,
209570_s_at	*D4S234E*	DNA segment on chromosome 4 (unique) 234 expressed sequence	fetal brain	Neuron-specific protein family member 1, dopamine receptor signaling pathway
215017_s_at	*FNBP1L*	formin binding protein 1-like	fetal brain	Membrane tubulation with reorganization of the actin cytoskeleton during endocytosis
218330_s_at	*NAV2*	neuron navigator 2	fetal brain	all-trans retinoic acid-responsive gene with role in neuronal development and neurite outgrowth
214395_x_at	*EEF1D*	eukaryotic translation elongation factor 1 delta (guanine nucleotide exchange protein)	fetal thyroid	Encodes a subunit of the elongation factor-1 complex, enzymatic delivery of aminoacyl tRNAs to the ribosome
216470_x_at	*PRSS1/2/3*	protease, serine, 1 (trypsin 1)/ protease, serine, 2 (trypsin 2)/ protease, serine, 3	pancreas	Trypsin family of serine proteases
201278_at	*DAB2*	disabled homolog 2, mitogen-responsive phosphoprotein (Drosophila)	placenta	Component of the CSF-1 signal transduction pathway

Probe Set ID	Gene Symbol	Gene Name	Tissue with maximum expression value in GNF Atlas	Entrez Gene
201681_s_at	*DLG5*	discs, large homolog 5 (Drosophila)	placenta	Maintenance of the structure of epithelial cells, transmission of extracellular signals to the membrane and cytoskeleton
218309_at	*CAMK2N1*	calcium/calmodulin-dependent protein kinase II inhibitor 1	prefrontal cortex	Potent and specific inhibitor of CaM-kinase II
201928_at	*PKP4*	plakophilin 4	spinal cord	Regulation of junctional plaque organization and cadherin function
204971_at	*CSTA*	cystatin A (stefin A)	tongue	Precursor protein of cornified cell envelope in keratinocytes, role in epidermal development and maintenance
205064_at	*SPRR1B*	small proline-rich protein 1B	tongue	Cross-linked envelope protein of keratinocytes
208539_x_at	*SPRR2B*	small proline-rich protein 2B	tongue	Cross-linked envelope protein of keratinocytes

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript