Published in final edited form as:

J Matern Fetal Neonatal Med. 2012 December ; 25(12): 2507–2511. doi:
10.3109/14767058.2012.717126.

Performing discovery-driven neonatal research by transcriptomic analysis of routinely discarded biofluids

Jill L. Maron^{1,2}, Jessica A. Dietz¹, Christopher Parkin³, Kirby L. Johnson¹, and Diana W. Bianchi^{1,2}

¹Department of Pediatrics, The Floating Hospital for Children at Tufts Medical Center, Boston, MA 02111, USA

²Mother Infant Research Institute at Tufts Medical Center, Boston, MA 02111, USA

³Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA 02111, USA

Abstract

Objective—To perform discovery-driven research on the neonatal salivary and cord blood transcriptomes.

Methods—Two separate cohorts of infants were enrolled in this study. In one, cord blood ($n = 10$) and in the other, saliva samples ($n = 10$) were collected at term gestation. Total RNA was extracted, amplified and hybridized onto Affymetrix HG U133a gene expression microarrays. Following normalization, genes expressed in the highest quintile (80%) across all subjects in each biofluid were analyzed with Ingenuity Pathway Analysis. Over-represented pathways relating to organ specific development and physiological functions in the newborn were explored.

Results—There were 303 genes in neonatal saliva and 282 genes in umbilical cord blood that met statistical criteria. Of these, 114 were common to both biofluids. Pathway analyses revealed the important roles of redox balance, cellular proliferation, and smooth muscle relaxation. In blood, hematopoiesis and immune response pathways predominated. In saliva, pathways associated with the gastrointestinal system were highlighted.

Conclusions—Neonatal cord blood and saliva provide a wealth of transcriptomic information. These normally discarded biofluids should be considered an important source of real-time gene expression data that may elucidate key pathways in neonatal physiology and pathology.

Keywords

Blood; neonate; noninvasive monitoring; saliva; transcriptome

Correspondence: Jill L. Maron, MD, MPH, Division of Newborn Medicine, Floating Hospital for Children at Tufts Medical Center, Mother Infant Research Institute, 800 Washington Street, Box 44, Boston, MA 02111, USA. Tel: 617-636-0766. Fax: 617-636-1469. jmaron@tuftsmedicalcenter.org.

Declaration of Interest:

The authors have no conflict of interests to report.

Introduction

Advances in molecular genetics alter the ways in which physicians can monitor, diagnose, and treat patients. Large-scale genome-wide association studies (GWAS), whole genome sequencing, and microarray technology permit the identification of population based single nucleotide polymorphisms (SNPs) and splice variants that have the potential to improve diagnosis and generate novel, targeted preventative strategies that may ultimately correlate with improved clinical outcomes [1–3]. In addition, advances in transcriptomic and proteomic analyses have made correlating gene expression and protein data to real-time development and disease a reality [4–7], while the newly emerging field of metabolomics aims at identifying metabolites (i.e. amino acids and vitamins) in order to assess the metabolic status of an organism in both health and disease states [8]. However, despite the strong need to utilize this technology for the discovery of novel biomarkers in the newborn, there remains a relative paucity of genomic, transcriptomic, metabolomic and proteomic research in the neonatal population as compared to adults.

The limited sources of readily available biological specimens prevent the application of these technologies to the newborn. An ideal biofluid for discovery-driven research would be one that is routinely discarded and/or noninvasively obtained. Both neonatal saliva and umbilical cord blood meet these criteria and are a rich source of genetic material. Previously, we described the umbilical cord transcriptome and reported that, at term gestation, genes involved in redox balance were highly expressed [9]. We speculated that these genes most likely played an important role in the newborn's transitioning from the *in utero* hypoxic to the relative hyperoxic *ex utero* environment. This work demonstrated that normally discarded neonatal materials, such as cord blood, provided important real-time information about neonatal physiology [9].

Based upon our results in the previous study, we questioned whether a similar gene expression profile would be available from noninvasively obtained salivary samples. Rather than merely describe the neonatal term salivary transcriptome, we aimed to better define and contrast the gene expression profiles of both umbilical cord blood and saliva. Here, we prospectively performed transcriptomic analyses on saliva samples from a new cohort of term newborns, and reanalyzed the umbilical cord data, to explore the biological pathways of genes expressed at the highest quintile (> 80%) in each biofluid for all subjects. We highlight the organ specific, developmental and physiological similarities, as well as key differences, between these two biofluids to perform discovery-driven transcriptomic analysis of routinely discarded biofluids.

Methods

This study was approved by the Tufts Medical Center Institutional Review Board. Informed parental consent was obtained for each enrolled subject ($n = 20$). There is a growing consensus in the literature that a minimum of five microarray experiments are required for statistically significant results [10]. For the purpose of this study, we chose to exceed those standards and analyze 10 biological specimens collected from two groups of neonates. Saliva was collected between days of life zero to four from neonates born between 36 6/7 to

40 3/7 weeks' gestation ($n = 10$). Umbilical cord blood data was obtained at birth from infants born between 36 5/7 and 39 5/7 weeks' gestation ($n = 10$).

Salivary acquisition, RNA extraction, amplification and microarray hybridizations

Whole saliva samples were collected and processed with previously developed techniques [11]. Estimated salivary volumes varied from 10 to 50 μ l. Neonates had not eaten within 1 h of sample collection to limit the risk of breast milk or formula contamination. Samples were stored at 4°C for a minimum of 48 h before being extracted with the RNeasy Saliva Protect Mini Kit (QIAGEN, Hilden, Germany) per the manufacturer's instructions. On-column DNase digestion was performed on each sample to eliminate DNA contamination. Extracted RNA was stored at -80°C awaiting amplification.

For all salivary samples, total RNA was amplified with the use of the Ovation Pico WTA System (NuGEN® Technologies, San Carlos, CA, USA). The quantity and quality of the amplified cDNA was assessed with an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) before fragmentation and biotinylation with the Encore Biotin Module (NuGEN Technologies). For each sample, a standard concentration of 5 μ g of amplified and labeled cDNA was hybridized onto the Affymetrix GeneChip Human Genome U133a microarray, comprising approximately 14,500 genes. Following hybridization, each array was washed and stained in the GeneChip Fluidics Station 400 (Affymetrix, Santa Clara, CA, USA) and scanned with the GeneArray Scanner (Affymetrix).

Umbilical cord blood data

The umbilical cord blood microarray data used in this study has been published previously [9]. In the prior study, we only reported on the cord blood transcriptome and made no comparative analyses with other biofluids. Here, we reanalyzed these data in order to compare transcriptomic profiles.

Assessment of array quality

All salivary and cord blood microarrays were subject to quality assessment prior to being included in this new analysis. Average backgrounds, scaling factors, percent present calls and 3 $\times$ /5 $\times$ ratios were measured and analyzed using the simpleaffy package for BioConductor [12].

Data analysis: umbilical cord blood and whole saliva

Microarray data were processed using the BioConductor suite of packages for the R statistical programming language [13]. Probe summarization and background correction were carried out using the MAS5 algorithm developed at Affymetrix, along with a quantile normalization [14]. In addition, nonexperimental variation was minimized through the use of ComBat, which adjusts the data for batch effect that could be caused by environmental factors when running experiments at various times or when using different batches of reagents [15]. Prior to analysis, intensity distribution and density plots were analyzed to ensure that all arrays behaved in a similar fashion.

Probes were filtered based on expression value to remove those at the background level. They were subsequently clustered into five groups based on probe intensity across all arrays. Only genes identified in the highest quintile for both saliva and blood were considered in our analyses. Gene transcripts that were well known to be expressed in either body fluid served as positive controls for the highest quintile of expression. For the saliva samples, the genes *KRT4*, *KRT6*, *KRT13*, *SPRR1* and *SPRR3* were used as positive controls [16–18]. Each of these genes is known to be highly expressed in saliva and/or buccal mucosa. Similarly, for the umbilical cord whole blood samples, the expression levels of highly expressed hemoglobin genes *HBA1*, *HBB*, *HBE1*, *HBD* and *HBG2* served as positive controls.

Pathway analyses and comparisons

To identify and interpret relevant biological information from each respective gene list, data were further analyzed with the use of the Ingenuity® Pathway Analysis (IPA) system (www.ingenuity.com; version 3204, release date: 27, October, 2010). IPA identifies statistically significantly over represented genes within known biological and pathological pathways [19]. Here, we utilized IPA to analyze the pathways expressed in the highest quintile of both whole blood and whole saliva, independently examine genes shared between biofluids, and highlight genes only expressed at the highest quintile in either biofluid.

Results

Array quality parameters

All quality measurements were within recommended guidelines established by Affymetrix [20]. In addition, Affymetrix control probes, such as *GAPDH*, showed similarity in terms of data quality between experiments. Altogether, this information assured a low risk of data-quality-bias in the comparative expression analyses between experiments.

The term umbilical cord and salivary transcriptomes

Subjects' demographic information is provided in Table I. After background filtering, we identified 303 genes expressed in the highest quintile in term neonatal saliva and 282 in the highest quintile of expression for term umbilical cord blood (Figure 1) out of a possible 14,500 genes on the microarray. Thus, only 2% of the genes on the array met statistical criteria and are included in this study. All positive control genes for each biofluid were present in the highest quintile of gene expression. The complete lists of all the genes and their corresponding functions expressed at the highest quintile in both umbilical cord blood and whole saliva can be found in Supplemental Tables I and II.

Genes expressed in both cord blood and saliva statistically significantly cluster into pathways associated with redox balance. These pathways include 'Oxidative Stress Response Mediated by Nrf2' (nuclear factor (erythroid-derived 2)-like 2), a well described transcription factor), 'Oxidative Stress', 'Hypoxia Inducible Factor Signaling' and 'Mitochondrial Dysfunction' ($0.001 < p < 0.003$). In addition, the G2/M Transition of the Cell Cycle was statistically significantly over represented in salivary samples ($p < 0.002$).

Gene expression similarities and differences between blood and saliva

The expression profiles of blood and saliva had 114 genes in common, resulting in 40% and 38% concordance, respectively (Figure 1). Analysis of these genes revealed clustering into the Cellular Effects of Sildenafil (Viagra)', 'Cellular Growth and Proliferation' and 'Nrf2-Mediated Oxidative Stress Response' ($10^{-5} < p < 10^{-2}$) pathways (Table II). Conversely, independent analysis of genes solely expressed in whole blood highlighted pathways associated with 'Hematopoiesis', 'Immune Cell Trafficking' and 'Antigen Presentation Pathway' ($10^{-5} < p < 10^{-2}$) (Table II). Genes unique to whole saliva clustered into pathways associated with 'Cellular Assembly and Organization' and 'Organ Development', including development of the gastrointestinal system ($10^{-6} < p < 10^{-2}$) (Table II). In total, 10 gastrointestinal pathways highlighted within the category 'Cancer', including those involved in the developing esophagus, colon, pancreas, and liver were more readily detected in neonatal whole saliva (Table III). The complete lists of pathways and their associated genes shared between and unique to cord blood and saliva can be found in Supplemental Tables IIIa, IIIb, and IIIc, respectively.

Discussion

The results of our study demonstrate that gene expression analyses of both whole blood and whole saliva provide relevant biological information about the term newborn. We chose to focus our study on those genes expressed by all infants, independent of gender, in either biofluid at the highest quintile in order to identify a discrete list of genes that represented important normal neonatal developmental and physiological processes. By also identifying genes unique to and shared between each biofluid, we have further highlighted important biological information about the neonatal salivary and blood transcriptomes that may be utilized by future investigators for prospective newborn studies. A comprehensive discussion of all the genes detected in our analyses would be beyond the scope of this manuscript. However, transcriptomic analyses of both umbilical cord blood and saliva at term reveal intricate pathways of redox balance, cellular proliferation, and smooth muscle relaxation occurring in normal newborns at birth. These data not only provide insight into relevant physiological processes, but also serve as a basis for comparative analyses between infants who fail to properly transition to the *ex utero* environment.

There was substantial overlap in gene expression between both whole blood and whole saliva samples. For example, in each cohort we were able to see the effects of the acute transition at birth from a relatively hypoxic *in utero* environment to the more hyperoxic ambient air. We have previously described the redox pathways detected in umbilical cord blood associated with this transition [9]. Here, for the first time, we observed nearly identical pathways in saliva samples. Importantly, we detected these pathways in saliva up to 4 days of life, suggesting that this transition is prolonged. In addition, analysis of genes unique to either body fluid often clustered into similar pathways. For example, pathways associated with 'Infectious Disease' and 'Cell Death' were both seen in the independent analyses, suggesting that at the highest level of expression, functional analysis is significantly similar between whole blood and whole saliva. This work further extends the comparative analyses between plasma and saliva previously described in adults [21].

Extensive salivary research in adults has shown that saliva has a biomarker profile similar to plasma, and here too, we saw significant similarities in gene detection, as well as pathways, between biofluids [22,23].

Further examination of the 110 genes expressed at the highest quintile in both whole saliva and blood revealed a novel pathway, 'Effects of Sildenafil', that was of great interest. Following birth, a newborn requires a sharp decrease in pulmonary vascular resistance to allow for adequate perfusion of the pulmonary vascular bed with lung expansion [24]. Failure to make this transition after birth results in a serious and sometimes fatal condition known as persistent pulmonary hypertension, (PPHN), or hypoxic failure of the newborn. Sildenafil, a type V phosphodiesterase inhibitor, has been used for the treatment of PPHN in newborns [25–27]. Through gene expression analyses of both saliva and blood, we identified specific genes involved in this pathway that are highly expressed at birth in healthy term neonates. These genes include myosin light chains 6 and 12B (*MYL6*, *MYL12B*), beta actin (*ACTB*), gamma actin (*ACTG1*), calmodulin 2 (*CALM2*) and GNAS complex locus (*GNAS*), which work sequentially to allow for smooth muscle relaxation. We speculate that the development of PPHN in neonates may result, in part, from aberrant gene expression along this pathway. Future comparative transcriptomic studies, with the use of either whole blood or whole saliva, between healthy newborns and those who develop PPHN, will need to be conducted to determine if this hypothesis holds true.

Not surprisingly, differential gene expression between biofluids revealed that whole blood more readily detected 'hematological development', 'hematopoiesis', and 'immune cell trafficking'. Conversely, gene expression in whole saliva samples highlighted pathways involving the gastrointestinal system. Over-represented pathways in both cord blood and saliva enriched into similar profiles independent of filtering for shared and unique transcripts. These data suggest that each biofluid, although similar in their gene expression profiles, represents a subset of distinct biological processes that should be considered by the investigator when choosing a source of genetic material.

A common hurdle associated with the utilization of commercially available analytic tools, such as IPA, is that they rely on published data for annotation. Due to the relative paucity of genomic, proteomic, and transcriptomic research on pediatric patients, our data often statistically significantly cluster into more well studied adult conditions such as 'Cancer', 'Hematological Disease' and 'Respiratory Disease'. We speculate that these classifications are best explained by the active proliferation of multiple organ systems at the cellular level at term gestation, as seen by the presence of several gastrointestinal pathways within the category 'Cancer' in the neonatal salivary analysis (Table III).

A potential limitation of this study is that the data are not paired, but rather are derived from two separate groups of neonates. However, the purpose of this study was not to compare the genes readily detectable in both biofluids in the same individual. Rather, we aimed to explore each biofluid separately, hypothesizing that detection of genes involved in key physiological processes would remain consistent across newborns. Previously, we demonstrated that healthy term neonates expressed a distinct transcriptomic profile that correlated with successful neonatal transitioning from the *in utero* to *ex utero* environment.

By prospectively collecting salivary samples from a new group of neonates, we validated that assumption and demonstrated that there is a distinct gene expression profile associated with neonatal transitioning at birth that can be seen independently, repeatedly, and in different biofluids. Indeed, Romero and Tromp have argued that validation of micro-array experiments must be conducted in different sets of samples, ‘otherwise the validation will be of the microarray technique, rather than the biological problem under study’ [10].

In conclusion, neonatal cord blood and saliva provide a wealth of transcriptomic information that may be analyzed and explored in newborn medicine. Both biofluids are obtained noninvasively and routinely discarded, although saliva has the advantage of being able to be obtained serially over time without inflicting additional trauma. This benefit may prove necessary in certain clinical studies. Regardless of the biofluid, real-time gene expression data, as it relates to development and disease can further our understanding of newborn medicine, identify novel biomarkers, and potentially generate targeted therapies for diseases that affect this population.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank the families who graciously participated in this work, as well as the nurses and staff of the Departments of Obstetrics and Gynecology and Pediatrics at Tufts Medical Center and the Floating Hospital for Children.

Funding for this research was provided by NICHD K08HD059819-03 and The Charles H. Hood Foundation awarded to JLM, and NICHD R01 HD042503-07 awarded to DWB.

References

1. Bonnefond A, Durand E, Sand O, De Graeve F, Gallina S, Busiah K, Lobbens S, et al. Molecular diagnosis of neonatal diabetes mellitus using next-generation sequencing of the whole exome. *PLoS ONE*. 2010; 5:e13630. [PubMed: 21049026]
2. Resch B, Müller WD. Interleukin-6-174 CC polymorphism is associated with clinical chorioamnionitis and cerebral palsy. *Ann Neurol*. 2010; 66:663–670.
3. Mailaparambil B, Krueger M, Heizmann U, Schlegel K, Heinze J, Heinzmann A. Genetic and epidemiological risk factors in the development of bronchopulmonary dysplasia. *Dis Markers*. 2010; 29:1–9. [PubMed: 20826912]
4. De Paepe ME, Greco D, Mao Q. Angiogenesis-related gene expression profiling in ventilated preterm human lungs. *Exp Lung Res*. 2010; 36:399–410. [PubMed: 20718599]
5. Maron JL, Johnson KL, Rocke DM, Cohen MG, Liley AJ, Bianchi DW. Neonatal salivary analysis reveals global developmental gene expression changes in the premature infant. *Clin Chem*. 2010; 56:409–416. [PubMed: 19959617]
6. Bujold E, Romero R, Kusanovic JP, Erez O, Gotsch F, Chaiworapongsa T, Gomez R, et al. Proteomic profiling of amniotic fluid in preterm labor using two-dimensional liquid separation and mass spectrometry. *J Matern Fetal Neonatal Med*. 2008; 21:697–713. [PubMed: 19012186]
7. Oggé G, Romero R, Lee DC, Gotsch F, Than NG, Lee J, Chaiworapongsa T, et al. Chronic chorioamnionitis displays distinct alterations of the amniotic fluid proteome. *J Pathol*. 2011; 223:553–565. [PubMed: 21294128]
8. Fanos V, Antonucci R, Barberini L, Noto A, Atzori L. Clinical application of metabolomics in neonatology. *J Matern Fetal Neonatal Med*. 2012; 25(Suppl 1):104–109. [PubMed: 22339399]

9. Maron JL, Johnson KL, Parkin C, Iyer L, Davis JM, Bianchi DW. Cord blood genomic analysis highlights the role of redox balance. *Free Radic Biol Med*. 2010; 49:992–996. [PubMed: 20566327]
10. Romero R, Tromp G. High-dimensional biology in obstetrics and gynecology: functional genomics in microarray studies. *Am J Obstet Gynecol*. 2006; 195:360–363. [PubMed: 16890547]
11. Dietz JA, Johnson KL, Wick HC, Bianchi DW, Maron JL. Optimal techniques for mRNA extraction from neonatal salivary supernatant. *Neonatology*. 2012; 101:55–60. [PubMed: 21791940]
12. Wilson CL, Miller CJ. Simpleaffy: a BioConductor package for Affymetrix Quality Control and data analysis. *Bioinformatics*. 2005; 21:3683–3685. [PubMed: 16076888]
13. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol*. 2004; 5:R80. [PubMed: 15461798]
14. Rafael, A.; Irizarry, RA.; Gautier, L.; Cope, LM. An R package of oligonucleotide array statistical analysis.. In: Parmigiani, G.; Garrett, ES.; Irizarry, RA.; Zegar, SL., editors. *The Analysis of Gene Expression Data: Methods and Software*. Springer; New York, USA: 2003. p. 102-118.
15. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007; 8:118–127. [PubMed: 16632515]
16. Bloor BK, Su L, Shirlaw PJ, Morgan PR. Gene expression of differentiation-specific keratins (4/13 and 1/10) in normal human buccal mucosa. *Lab Invest*. 1998; 78:787–795. [PubMed: 9690557]
17. Manconi B, Cabras T, Pisano E, Nemolato S, Inzitari R, Iavarone F, Fanali C, et al. Characterization of two isoforms of human SPRR3 from saliva of preterm human newborn and autaptic fetal oral mucosa, parotid and submandibular gland samples. *Biochem Biophys Res Commun*. 2010; 398:477–481. [PubMed: 20599699]
18. Gibbs S, Fijneman R, Wiegant J, van Kessel AG, van De Putte P, Backendorf C. Molecular characterization and evolution of the SPRR family of keratinocyte differentiation markers encoding small proline-rich proteins. *Genomics*. 1993; 16:630–637. [PubMed: 8325635]
19. Jiménez-Marín A, Collado-Romero M, Ramirez-Boo M, Arce C, Garrido JJ. Biological pathway analysis by ArrayUnlock and Ingenuity Pathway Analysis. *BMC Proc*. 2009; 3(Suppl 4):S6. [PubMed: 19615119]
20. Bolstad, BM.; Collin, F.; Brettschneider, J.; Simpson, K.; Cope, L.; Irizarry, RA.; Speed, TP. Quality assessment of Affymetrix genechip data.. In: Gentleman, R.; Carey, VJ.; Huber, W.; Irizarry, RA.; Dudoit, S., editors. *Springer; New York, USA: Bioinformatics and computational biology solutions using R and Bioconductor..* 2005. p. 33-49.
21. Yan W, Apweiler R, Balgley BM, Boontheung P, Bundy JL, Cargile BJ, Cole S, et al. Systematic comparison of the human saliva and plasma proteomes. *Proteomics Clin Appl*. 2009; 3:116–134. [PubMed: 19898684]
22. Wong DT. Salivary Diagnostics: Amazing as it might seem, doctors can detect and monitor diseases using molecules found in a sample of spit. *Am Sci*. 2008; 96:37–43. [PubMed: 19750202]
23. Zhang L, Xiao H, Wong DT. Salivary biomarkers for clinical applications. *Mol Diagn Ther*. 2009; 13:245–259. [PubMed: 19712004]
24. Gao Y, Raj JU. Regulation of the pulmonary circulation in the fetus and newborn. *Physiol Rev*. 2010; 90:1291–1335. [PubMed: 20959617]
25. Steinhorn RH, Kinsella JP, Pierce C, Butrous G, Dilleen M, Oakes M, Wessel DL. Intravenous sildenafil in the treatment of neonates with persistent pulmonary hypertension. *J Pediatr*. 2009; 155:841–847. e1. [PubMed: 19836028]
26. Spillers J. PPHN: is sildenafil the new nitric? A review of the literature. *Adv Neonatal Care*. 2010; 10:69–74. [PubMed: 20386371]
27. Baquero H, Soliz A, Neira F, Venegas ME, Sola A. Oral sildenafil in infants with persistent pulmonary hypertension of the newborn: a pilot randomized blinded study. *Pediatrics*. 2006; 117:1077–1083. [PubMed: 16585301]

Figure 1.

Venn diagram depicting gene expression in whole blood and whole saliva. In our analysis, we considered the highest quintile of genes expressed in cord blood (dashed lines) and saliva (dots), as well as those genes that were shared ($n = 114$) and unique to cord blood ($n = 168$) and saliva ($n = 189$).

Table I

Subject demographic information.

Sample type								
Saliva					Umbilical cord blood			
Gestational age (weeks)	Gender	Day of life	Apgars	Ethnicity	Gestational age (weeks)	Gender	Apgars	Ethnicity
36 (6/7)	Female	1	9, 9	Caucasian	36 (5/7)	Male	9, 9	Caucasian
37	Female	2	*	Asian	36 (5/7)	Male	8, 9	Caucasian
37	Female	1	8, 9	Caucasian	37 (3/7)	Male	9, 9	Caucasian
38 (6/7)	Male	3	8, 9	Caucasian	39	Female	9, 9	Caucasian
39 (3/7)	Male	1	4, 9	Caucasian	39	Male	8, 9	Asian
39 (4/7)	Female	3	9, 9	Caucasian	39 (1/7)	Female	8, 9	African American
39 (4/7)	Female	4	9, 9	Asian	39 (1/7)	Male	8, 9	Hispanic
39 (6/7)	Female	2	9, 9	Caucasian	39 (2/7)	Male	8, 9	Caucasian
40 (2/7)	Male	4	9, 9	African American	39 (2/7)	Female	8, 9	Asian
40 (3/7)	Female	3	7, 9	Caucasian	39 (5/7)	Male	9, 9	African American

* Home birth; Apgar scores unavailable.

Table II

Functional pathways in term saliva and umbilical cord blood, associated p values, and number of transcripts in each pathway.

	Shared pathways	Unique blood pathways	Unique saliva pathways
Physiological system development and function	Connective tissue development and function ($0.006 < p < 0.03$; $n = 6$)	Hematological system development and function ($10^{-5} < p < 0.03$; $n = 15$)	Hair and skin development and function ($10^{-10} < p < 0.01$; $n = 24$)
	Embryonic development ($0.006 < p < 0.03$; $n = 3$)	Hematopoiesis ($10^{-5} < p < 0.02$; $n = 27$)	Hematological system development and function ($10^{-6} < p < 0.01$; $n = 34$)
	Endocrine system development and function ($0.006 < p < 0.01$; $n = 1$)	Tissue morphology ($10^{-5} < p < 0.02$; $n = 15$)	Immune cell trafficking ($10^{-6} < p < 0.009$; $n = 23$)
	Hematological system development and function ($0.006 < p < 0.04$; $n = 8$)	Nervous system development and function ($10^{-4} < p < 0.02$; $n = 26$)	Organ development ($10^{-5} < p < 0.004$; $n = 9$)
	Organismal development ($0.006 < p < 0.01$; $n = 2$)	Immune cell trafficking ($10^{-3} < p < 0.02$; $n = 55$)	Cardiovascular system development and function ($10^{-4} < p < 0.01$; $n = 7$)
Molecular and cellular functions	Protein synthesis ($10^{-25} < p < 0.02$; $n = 33$)	Protein synthesis ($10^{-7} < p < 0.03$; $n = 27$)	Cellular development ($10^{-11} < p < 0.01$; $n = 56$)
	RNA post-transcriptional modification ($10^{-10} < p < 0.006$; $n = 9$)	Small molecule biochemistry ($10^{-6} < p < 0.02$; $n = 41$)	Cell death ($10^{-10} < p < 0.01$; $n = 75$)
	Cellular assembly and organization ($10^{-8} < p < 0.04$; $n = 15$)	Cell death ($10^{-5} < p < 0.02$; $n = 49$)	Cellular growth and proliferation ($10^{-7} < p < 0.01$; $n = 75$)
	Cellular growth and proliferation ($10^{-5} < p < 0.04$; $n = 33$)	Post-translational modification ($10^{-5} < p < 0.02$; $n = 23$)	Cellular movement ($10^{-6} < p < 0.01$; $n = 49$)
	Cellular development ($10^{-5} < p < 0.04$; $n = 29$)	Free radical scavenging ($10^{-4} < p < 0.02$; $n = 13$)	Cell-to-cell signaling and interaction ($10^{-6} < p < 0.01$; $n = 41$)
Canonical pathways	Cellular effects of Sildenafil (Viagra) ($p < 0.0002$)	Antigen presentation pathway ($p < 0.0006$)	Integrin signaling ($p < 10^{-7}$)
	Clathrin-mediated endocytosis signaling ($p < 0.0005$)	Mitochondrial dysfunction ($p < 0.002$)	Regulation of actin-based motility by Rho ($p < 10^{-4}$)
	RhoA signaling ($p < 0.0006$)	EIF2 signaling ($p < 0.002$)	Phospholipase C signaling ($p < 10^{-4}$)
	Nrf2-mediated oxidative stress response ($p < 0.0009$)	Virus entry via endocytic pathways ($p < 0.002$)	Germ cell-sertoli cell junction signaling ($p < 10^{-4}$)
	Regulation of actin-based motility by Rho ($p < 0.0002$)	Oxidative phosphorylation ($p < 0.004$)	14-3-3-mediated signaling ($p < 10^{-3}$)
Diseases and disorders	Cancer ($10^{-10} < p < 0.04$; $n = 49$)	Organismal injury and abnormalities ($10^{-5} < p < 0.03$; $n = 15$)	Cancer (including gastrointestinal development/disease) ($10^{-13} < p < 0.01$; $n = 89$)
	Genetic disorder ($10^{-10} < p < 0.03$; $n = 38$)	Infectious disease ($10^{-5} < p < 0.01$; $n = 27$)	Infectious disease ($10^{-7} < p < 0.01$; $n = 34$)
	Respiratory disease ($10^{-10} < p < 0.03$; $n = 14$)	Respiratory disease ($10^{-5} < p < 0.03$; $n = 15$)	Respiratory disease ($10^{-7} < p < 0.01$; $n = 28$)
	Inflammatory response ($10^{-7} < p < 0.04$; $n = 5$)	Hematological disease ($10^{-5} < p < 0.03$; $n = 26$)	Cardiovascular disease ($10^{-7} < p < 0.01$; $n = 25$)
	Connective tissue disorders ($10^{-6} < p < 0.01$; $n = 27$)	Skeletal and muscular disorders ($10^{-5} < p < 0.03$; $n = 55$)	Developmental disorder ($10^{-7} < p < 0.01$; $n = 23$)

Table III

Gastrointestinal pathways highlighted in saliva samples.

Pathway	<i>p</i> value	Genes
Digestive organ tumor	1.0×10^{-6}	<i>ANXA1, ANXA2, CAMK2N1, CD59, CDKN1A, CTTN, DNAJA1, DSTN, EZR, FKBP1A, HNRNPK, IDI1, IL1RN, KRT9, LCN2, MT2A, PTMA, SFN, TACSTD2, TM4SF1, TUBA1A, TUBB2C, YWHAE</i>
Pancreatic cancer	3.9×10^{-4}	<i>CD59, CDKN1A, FKBP1, KRT19, LCN2, S100A2, S100P, TUBA1A, TUBB2C, YWHAE</i>
Biliary tract tumor	7.8×10^{-4}	<i>ANXA1, ANXA2, FKBP1A, TM4SF1, TUBA1A, TUBB2C</i>
Esophageal carcinoma/cancer	4.0×10^{-3}	<i>CDKN1A, CTTN, EZR, TACSTD2, TUBA1A, TUBB2C</i>
Gastric cancer	7.5×10^{-3}	<i>CDKN1A, CTNNA1, DSTN, FKBP1A, IL1RN, MAL, TUBA1A, TUBB2C</i>
Stomach tumor	9.8×10^{-3}	<i>DSTN, FKBP1A, IL1RN, TUBA1A, TUBB2C</i>
Colon cancer	2.9×10^{-3}	<i>ANXA11, AQP3, ATP2A2, B4GALT1, CAMK2N1, CDKN1A, DNAJA1, FKBP1A, IDI1, PTMA, TUBA1A, TUBB2C, YWHAE</i>
Pancreatic adenocarcinoma/carcinoma	2.7×10^{-3}	<i>CD59, CDKN1A, FKBP1A, KRT19, LCN2</i>
Colorectal tumor	7.0×10^{-3}	<i>CAMK2N1, DNAJA1, IDI1, PTMA, TUBA1A, TUBB2C, YWHAE</i>
Gastrointestinal tumor	7.3×10^{-3}	<i>DSTN, EZR, FKBP1A, IL1RN, TUBA1A, TUBB2C</i>