Archived Unfrozen Neonatal Blood Spots Are Amenable to Quantitative Gene Expression Analysis

Peterson T. Haak \( ^{a, c, e} \) \quad \text{Julia V. Busik} \( ^{f} \) \quad \text{Eric J. Kort} \( ^{b} \) \quad \text{Maria Tikhonenko} \( ^{f} \) \quad \text{Nigel Paneth} \( ^{c, d} \) \quad \text{James H. Resau} \( ^{a, b} \)

Laboratories of \( ^{a} \) Microarray Technology and \( ^{b} \) Molecular Epidemiology, Van Andel Research Institute, Grand Rapids, Mich., Departments of \( ^{c} \) Epidemiology and \( ^{d} \) Pediatrics and Human Development, College of Human Medicine, \( ^{e} \) College of Osteopathic Medicine, and \( ^{f} \) Department of Physiology, Michigan State University, East Lansing, Mich., USA

---

**Key Words**

Guthrie \( \cdot \) Blood spot \( \cdot \) RNA \( \cdot \) Microarray \( \cdot \) qPCR \( \cdot \) Real-time PCR \( \cdot \) Neonatal screening

**Abstract**

**Background**: State laws in the USA mandate that blood be drawn from all newborn infants to screen for health-threatening conditions. These screening assays consume only a small portion of the blood samples, which are collected on filter paper ('Guthrie') cards. Many states archive unused blood spots, often in unrefrigerated storage. **Objectives**: While individual RNA transcripts have been identified from archived neonatal blood spots, no study to date has performed quantitative analysis of archived blood spot RNA. **Methods**: We demonstrate that RNA can be isolated and amplified from newborn blood spots stored unfrozen for as long as 9 years, and can be analyzed by microarray and qPCR. **Results**: Microarray assays of archived neonatal blood spots consistently detected 3,000–4,000 expressed genes with correlations of 0.90 between replicates. Blood spot mRNA is amenable to qPCR and we detected biologically relevant expression levels of housekeeping and immune-mediating genes. **Conclusions**: These experiments demonstrate the feasibility of using blood spots as a source of RNA which can be analyzed using quantitative microarray and qPCR assays. The application of these methods to the analysis of widely collected biological specimens may be a valuable resource for the study of perinatal determinants of disease development.

---

**Introduction**

Newborns are routinely screened for a variety of conditions – predominantly inborn errors of metabolism – for which early diagnosis can lead to therapy and subsequent prevention of sequelae, including permanent neurological damage [1]. Such screening programs are in place throughout the world. In the USA, mandatory screening programs exist in all 50 states; the number of conditions screened for in each state varies from 4 to 30 [2]. Within 24–48 h after birth, newborn blood is collected on filter paper cards ('Guthrie cards') and submitted for laboratory analysis using principally tandem mass spectrometry, high-performance liquid chromatography, and antibody-based assays. These assays generally con-
sume only a small portion of the available blood on the cards. The portion of the cards with remaining blood is then archived under varying conditions, and for varying durations. A recent survey found that 14 of the 51 US newborn screening programs retain newborn blood for 21 years or more [3]. These population-wide archives of biologic specimens represent a unique resource for the retrospective study of diseases of perinatal origin.

To be of use in health research, informative material in dried blood spots must persist after years of dry, room temperature storage. Proteins in blood spots are beyond detection in a matter of days [4], but nucleic acids show evidence of long-term stability. The published literature shows extensive examples of recovery and analysis of nucleic acids from dried blood spots. Many studies have involved screening for viral and bacterial DNA and RNA in dried blood spots [5–9]. The presence of infectious agents at birth may be very relevant to understanding diseases of perinatal origin, but since analysis of nucleic acids from infectious agents is relatively well documented, we focus our attention on DNA and RNA from the neonate.

Most reports on the analysis of neonatal nucleic acids from archived blood spots have focused on DNA, and several procedures for extracting DNA from filter paper have been published [10–13]. Subsequent analyses of blood spot DNA have included on-chip PCR-based microarray screening for specific gene polymorphisms [14], screening for disease markers and mutations [15–18] and post-mortem diagnosis of a disease-causing mutation [19]. Another interesting use of archived neonatal blood DNA has been to use PCR and sequencing to identify leukemic cells present at birth and other diagnostic markers in infants who subsequently developed leukemia [20–23].

As RNA is known to be more labile than DNA, it is reasonable to doubt the stability of RNA in long-term unfrozen storage. Yet peer-reviewed research shows that RNA in dried blood spots is stable after years of storage, contains transcripts indicative of specific diseases, and mirrors the gene expression in fresh blood. Blood spot RNA analysis was first published by Matsubara et al. [24] and Zhang et al. [23], who performed mutation and restriction enzyme analysis of expressed genes from spots stored for several years. Long-term RNA stability has been most strongly demonstrated by Swedish researchers who have recovered intact RNA transcripts from blood spots archived up to 21 years [25]. Several researchers show that genes expressed in archived blood samples reflect disease states. Suttorp et al. [26] detected RNA for the bcr-abl rearrangement indicative of the Philadelphia chromosome in dried blood spots. Pai et al. [27] showed that patients who lack expression of the FMR-1 gene due to fragile-X syndrome lack FMR-1 RNA in their neonatal blood spot. Both Pai et al. and Suttorp et al. verified their blood spot findings against RNA from fresh blood. We found one paper that compared RNA from a sample of fresh blood against RNA from identical blood spotted, dried, and stored for several months [28]. Maeno et al. confirmed that, after several months of room temperature storage, infection-induced cytokine mRNA in the blood spot precisely mirrored the genes expressed in the original fresh blood sample. While these publications all support the stability of blood spot RNA and its ability to reflect disease states, they only examined from 1 to 13 RNA transcripts at a time.

Researchers have yet to apply the latest tools in RNA technology to measure blood spot gene expression on a larger scale. We can find no published reports of microarray-based gene expression analysis of RNA from archived neonatal blood spots. If whole-expressed genome quantitative RNA expression can be applied to archived newborn blood spots, these specimens would constitute a remarkable resource for clinicians and researchers. More so than DNA, blood spot RNA would provide a temporal molecular snapshot of physiologic and metabolic function shortly after birth. These ‘expressed genetic profiles’ could then be retrospectively studied to identify markers and potential causal factors of childhood (and perhaps adult-onset) diseases from the perinatal period. RNA data could add further value to the already universally collected (and in some states research-available) neonatal blood spot.

In this report, we describe the results of parallel microarray-based gene expression and quantitative PCR analyses of RNA extracted from 12 archived neonatal blood spots.

**Materials and Methods**

**Selection of Neonatal Blood Spots**

In Michigan, 5 circles of 1.25 cm diameter on Guthrie cards are filled with blood obtained from a heel stick blood draw obtained between 24 and 72 h after birth. In most instances, only a portion of the first two spots are used for clinical testing. By state law in effect since 1986, the remaining blood is archived for the next 21.5 years. The Michigan Department of Community Health stores the Guthrie cards in boxes on shelves in rooms at ambient temperature. Michigan state law governing the use of these archived blood spots explicitly encourages their use for medical research, and anonymous specimens can be obtained for research.

**Archived Neonatal Blood Spots and Quantitative Gene Expression Analysis**

Neonatology 2009;95:210–216

211
Isolation, Quantification, and Amplification of RNA from Neonatal Blood Spots

We randomly selected 12 of the 85 blood spots for RNA testing; three blood spots (each representing a different neonate) from each of the years 1998, 2000, 2003, and 2004. The 73 remaining spots were set aside for future experiments. We collected eight 3-mm punches from each blood spot. Punches were soaked in 1 ml TRIzol reagent (Invitrogen) for 15 min and agitated until all blood was dissolved from the spot. Tubes were then spun in a microcentrifuge to pellet the filter paper. TRIzol reagent containing the lysed blood cells was transferred to a fresh tube with 200 l of total cDNA for each sample were set aside during the first round of amplification for quantitative PCR analysis. Amplified RNA (aRNA) was quantified using the Nanodrop ND1000 spectrophotometer.

DNA Microarray Analysis of Neonatal Blood Spot RNA

For each of our 12 samples, we performed the following procedure in triplicate. For each microarray, reverse transcription using SuperScript II (Invitrogen) was performed on 20 l of aRNA primed with random hexamers (Invitrogen) in the presence of cyanine-3 dCTP (PerkinElmer). To enable comparison within and across all replicate samples, each individual sample was compared against a common reference; equal amounts of aRNA from all 12 neonatal samples were pooled to form the common reference. This reference pool was labeled with cyanine-5 dCTP and divided for use across all arrays. Microarrays were synthesized at the Van Andel Institute and had passed internal quality assurance/quality control standards. Labeled cDNA cleanup, hybridization to 20,000 feature human cDNA arrays, and post-hybridization array washes were performed in triplicate for each sample as previously described [29, 30]. Arrays were scanned at 532 and 635 nm using standard settings in an Agilent Microarray Scanner (Agilent Technologies, Palo Alto, Calif., USA). Images were analyzed using GenePix Pro 5.0 (Axon, Union City, Calif., USA). Spots with signal lower than three times the standard deviation of the global array background in either channel were excluded from analysis. The array intensities were adjusted (normalized) using the within print-tip normalization methods as implemented in the limma package [31] for the R statistical environment [32].

Quantitative PCR Analysis of Neonatal Blood Spot RNA

The 2 l aliquot of double-stranded cDNA, representing mRNA extracted from each blood spot, was divided into 12 aliquots to permit study of four genes in triplicate from each spot. These four genes were two housekeeping/metabolic genes: mitochondrial ribosomal protein L11 (MRPL11) and hexokinase 1 (HK1); and two inflammatory genes: nuclear factor kappa B1 (NFKB1) and alpha X integrin (ITGAX). The primers used for amplification were specifically selected to amplify only RNA. In each pair of primers, either one or both were exon-spanning, and the primers were always located on different exons. The introns present in genomic DNA prevent primer binding and thus the PCR reaction cannot be initiated by contaminating DNA. Two additional steps were then used to confirm that we had amplified mRNA. First, we used commercially available genomic DNA in each experimental array and showed that amplification did not occur with DNA. Second, we applied RNAse to our experimental arrays and showed that after this treatment, amplification did not occur, indicating that we were detecting mRNA. RNAse treatment with IRB approval [MI Public Health Code 333.5431]. Individually identified spots can be obtained with parental informed consent.

After receiving IRB approvals from MSU and the Michigan Department of Community Health (MDCH), we obtained from MDCH an anonymous stratified random sample of newborn blood spots collected from 1996 to 2004. All blood spots were stored in boxes corresponding to the date of laboratory testing of the samples. Within the boxes, which are filed by birth date in the Michigan Department of Laboratories, blood spot cards are bundled in groups of 70, with most boxes containing between 8 and 20 bundles. A stratified sampling procedure was used to randomly select boxes, bundles, and cards within bundles. We selected only cards with three unused spots, or two unused spots and a third spot with only a single punch taken. If the randomly selected card was not eligible, we alternated examining cards on either side of the selected card until we found a usable card. This procedure was required less than 5% of the time. We then excised one unused 1.25-cm diameter blood spot and placed it in a plastic envelope. We aimed to obtain 10 samples per year for each of the 9 years, and using this method we were able to obtain 85 archived specimens.

Support Facility (RTSF) at MSU on an ABI Prism 7900HT Sequence Detection System. Hexokinase 1, a stably and highly expressed neonatal blood transcript, was used as a loading control in qPCR analysis. We monitored the amount of DNA produced in each PCR cycle and counted the fold changes required to obtain satisfactory binding to SYBR Green.

Haak/Busik/Kort/Tikhonenko/Paneth/Resau
Results

Isolation, Quantification, and Amplification of RNA from Neonatal Blood Spots

Each column in figure 1 depicts the average yield of RNA from three different samples within a given year. Although one sees a slight decrease in total RNA and first round aRNA when comparing older versus newer blood spots, differences in yield are negligible by the second round of amplification. The 12 samples averaged 160 ± 13 μg aRNA. First round amplification products should meet the requirements for most commercial microarray platforms (9, 11, 21 and 23 μg, from oldest to newest spots), while the second round of RNA amplification provides enough aRNA to perform multiple microarrays.

DNA Microarray Analysis of Neonatal Blood Spot RNA

As shown in figure 2, the average number of expressed genes detected between all years and all neonatal blood spots was 3,480 of 20,000 total. This ranged from an average of 3,017 genes for blood spots from 1998 and 4,067 genes for blood spots from 2004. The average correlation within triplicate assays for all neonatal blood samples was 0.904 ± 0.042.

Quantitative PCR Analysis of Neonatal Blood Spot RNA

We amplified four genes in triplicate from each of our 12 blood spots, for a total of 36 assays per gene and 144 total qPCR reactions. Because of limited starting material, some samples did not have enough material for triplicates. Standard deviations were not calculated for these samples.

Figure 3 (main window) shows saturation curves for individual assays. HK1, a metabolic gene expressed in higher number in neonatal erythrocytes [37], gave Ct values before 35 cycles for 26/36 assays (72%) and was present in all (12/12) samples. MRPL11 gave Ct values before 35 cycles for 24/36 assays (67%, 11/12 samples). ITGAX, also known as leukocyte adhesion receptor, reached Ct before 35 cycles in 13/36 assays (36%, 9/12 samples). As ITGAX can mediate cell-cell interactions during inflam-
Neonatology 2009;95:210–216

Information [38], we would not expect it to be expressed in all subjects. The same is true of qPCR assays on the pro-inflammatory cytokine NFκB1; only five of our twelve blood spot samples (6/36 aliquots) reached Ct for this cytokine between 20 and 35 cycles.

The dissociation curves (fig. 3; inset) for the amplification products of each assay show dissociation at only one temperature, indicating that – where there was amplification – our primers amplified a single transcript with a uniform melting point. This shows a lack of non-specific amplification and further validates the specificity of our pPCR measurements.

In table 1, values with standard deviations indicate there was adequate starting material for triplicate assays. Relative expression levels, after normalization to HK1, show biologically plausible results. As is expected, NFκB and ITGAX were not expressed in all samples and relative expression levels were lower for these inflammatory genes than for MRPL11. One of our randomly selected blood spots (1998_B) showed relatively higher levels of NFκB1 when compared to the other four NFκB1 expressing blood spots.

Concordance between Microarray and Quantitative PCR for Blood Spot RNA

Our qPCR experiment was not designed to validate the microarray results, but rather was performed in parallel as a proof-of-principle study. Because of this we used different sets of primers for microarray and qPCR analysis, and were unable to target qPCR to measure the genes reported as most highly misregulated by microarray. However, even with this caveat, we observed a very high (79 ± 22%) concordance between genes measured as present by microarray and genes measured as present by
qPCR. Concordance of RNA transcript presence was highest for the housekeeping genes HXK and MRPL11 (100 and 92% agreement, respectively) and lower for ITGAX and NFkB (75 and 50% agreement, respectively). To enable accurate microarray validation in the future we will use the exact same sets of primers, based on microarray sequences, for qPCR assays.

### Discussion

In this study we used two different laboratory methods – DNA microarray and quantitative PCR – to demonstrate the feasibility of quantitative analysis of RNA from archived blood spots. Both methods demonstrated that the RNA obtained from archived blood spot samples is usable for quantitative gene expression analysis. Our results show that consistent, quantitative RNA measurements can be obtained from neonatal blood spot samples archived at room temperature for up to 9 years.

Further investigation is warranted to determine if improvements in RNA isolation and amplification will increase the quantity and quality of data from blood spots. Targeted removal of globin RNA could unmask more genes in the samples [39]. We realize the threat that DNA contamination can play when using small amounts of RNA and have performed experiments showing the efficacy of treatment with RNase-free DNAse (data not shown). We plan to apply this treatment to all future experiments. Future experiments will also utilize commercially available Agilent arrays; these arrays require less input RNA, have inherent QC and QA considerations, and provide more standardized measurements that will facilitate across-laboratory comparisons. Agilent microarrays will also enable us to design qPCR primers that correspond to the oligonucleotide probes used on the array, enabling qPCR validation of our microarray results. As we only consumed a portion of the available blood spots (8 of 14 3 mm punches), we will preserve at least part of the spot for future assays, including viral and bacterial screening by PCR.

Newborn metabolic screening continues to be a vital and successful clinical assay of neonates. By teaming with state health departments and utilizing existing blood spot archives, we hope to improve our understanding of diseases that are not immediately apparent at birth, but have roots in the perinatal period. In conclusion, these findings correlate with previously published research showing that, in addition to being a source of informative DNA, RNA from archived neonatal blood spots is amenable to quantitative gene expression measurement. Measuring the relative abundance of thousands of expressed genes from universally collected neonatal blood spots may open new avenues of research into perinatal markers, determinants of disease development, and add value to an underutilized biologic archive.

### Acknowledgements

The authors would like to thank Todd Lydic and Shirali Gaurang Patel for RNA work, and Kyle Furge and Karl Dykema of the VARI Bioinformatics core for their analysis contributions and guidance. We also appreciate the efforts of Madeleine Lenski at Michigan State University and Harry Hawkins at the Michigan Department of Community Health for neonatal sample acquisition.

Partial support for this work was provided by the Office of Research, College of Human Medicine, Michigan State University, as well as generous support from the Jay and Betty Van Andel Foundation (J.H.R., E.J.K., and P.T.H.). P.T.H. and E.J.K. are also supported through a T32 training grant in perinatal epidemiology from the NIH.

---

**Table 1. Neonatal blood spot qPCR expression levels relative to hexokinase 1**

<table>
<thead>
<tr>
<th>Year</th>
<th>Sample No.</th>
<th>2−ΔΔCt compared to hexokinase1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MRPL11</td>
</tr>
<tr>
<td>1998</td>
<td>1998_A</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>1998_B</td>
<td>1.94±1.63</td>
</tr>
<tr>
<td></td>
<td>1998_C</td>
<td>0.66±0.01</td>
</tr>
<tr>
<td>2000</td>
<td>2000_A</td>
<td>1.36±1.91</td>
</tr>
<tr>
<td></td>
<td>2000_B</td>
<td>0.69±0.54</td>
</tr>
<tr>
<td></td>
<td>2000_C</td>
<td>0.81±0.39</td>
</tr>
<tr>
<td>2003</td>
<td>2003_A</td>
<td>0.29±0.40</td>
</tr>
<tr>
<td></td>
<td>2003_B</td>
<td>1.61</td>
</tr>
<tr>
<td></td>
<td>2003_C</td>
<td>0.56±0.04</td>
</tr>
<tr>
<td>2004</td>
<td>2004_A</td>
<td>1.92±0.70</td>
</tr>
<tr>
<td></td>
<td>2004_B</td>
<td>0.22±0.11</td>
</tr>
<tr>
<td></td>
<td>2004_C</td>
<td>0.59±0.67</td>
</tr>
</tbody>
</table>

---

Archived Neonatal Blood Spots and Quantitative Gene Expression Analysis