Published in final edited form as:

J Pediatr. 2014 January ; 164(1): 189–191.e1. doi:10.1016/j.jpeds.2013.09.025.

Comparison of Frozen and Unfrozen Blood Spots for Gene Expression Studies

Changshuai Wei¹, Qing Lu¹, Sok Kean Khoo², Madeleine Lenski¹, Raina Fichorova³, Alan Leviton⁴, and Nigel Paneth^{1,5}

¹Department of Epidemiology and Biostatistics, College of Human Medicine, Michigan State University, East Lansing, MI

²Laboratory of Microarray Technology, Van Andel Research Institute, Grand Rapids, MI

³Laboratory of Genital Tract Biology, Department of Obstetrics, Gynecology and Reproductive Biology, Harvard Medical School and Brigham and Women's Hospital, Boston, MA

⁴Neuroepidemiology Unit, Department of Neurology, Boston Children's Hospital, Harvard University, Boston, MA

⁵Department of Pediatrics and Human Development, College of Human Medicine, Michigan State University, East Lansing, MI

Abstract

We studied gene expression in nine sets of paired newborn blood spots stored for 8–10 years in either the frozen or unfrozen state. Fewer genes were expressed in unfrozen spots, but the average correlation coefficient for overall gene expression comparing the frozen and unfrozen state was 0.771 (95% CI: [0.700–0.828]).

Newborn blood spots, archived after genetic screening in at least ten states,¹ represent a unique resource for the study of diseases with roots in the perinatal period. In Michigan, newborn spots were archived from 1987 – 2008 unfrozen, but since 2009, at –20°C.

We have shown, in archived unfrozen blood spots, that gene transcripts² and gene markers³ can be detected, and that gene expression profiles differ in children who later develop CP.⁴ In fresh blood spots, we found nearly identical gene expression in placenta and newborn blood.⁵ We here compare microarray gene expression in 8–10 year old blood spots collected from nine newborns, with one spot stored frozen and the other unfrozen.

© 2013 Mosby, Inc. All rights reserved.

Corresponding Author: Nigel Paneth, MD MPH, Fax: 517-432-7730, Phone: 517-290-5062, paneth@msu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The authors declare no conflicts of interest.

Methods

With parental permission, we obtained frozen (-70°C) blood spots stored in the ELGAN study at Sparrow Hospital, Lansing, MI in nine newborns < 28 weeks gestation at birth in 2002–2004. We also retrieved their state-archived blood spots, taken at about the same age of 1–3 days, but stored unfrozen. The ELGAN spots were kept unfrozen for days to weeks before being sent in batches to the ELGAN laboratories. For this study, the 9 frozen spots were sent in dry ice to the Laboratory of Microarray Technology at VARI, where they were stored at -70°C until RNA extraction. The state-archived blood spots were stored at ambient temperatures until RNA extraction.

Total RNA was extracted, purified and concentrated from three 3mm punches using glass-fiber filter and subjected to DNase treatment. Average mRNA yield was higher from frozen than unfrozen spots (73.4 ng vs 22.8 ng) but the range was very large (5.9 – 113.3). RNA integrity (RIN) was equivalent (unfrozen = 2.3; frozen = 2.2).

The WT-Ovation Pico RNA Amplification System (NuGEN Technologies) generated single-stranded cDNA, which was purified and labeled with Alexa Fluor 3 fluorescent dye before hybridization onto the Agilent whole human genome gene expression 8x60K microarray, which has 60,000 oligonucleotide probes (60bp probe) covering 27,958 Entrez Gene RNAs and 7,419 long intergenic non-coding RNAs. Arrays were hybridized (with the same amount of labeled cDNA from all samples) for 17 hours at 65°C and 10 rpm rotation speed, washed for 2 min each with wash buffer 1 and 2, and scanned with an Agilent G3 high-resolution scanner. Probe features were extracted from the microarray scan data using Feature Extraction software v. 10.7.3.1 (Agilent). All paired specimens were run on the same array.

Statistical analyses

After filtering and quantile normalization^{6,7}, we aggregated the expression signal at the gene level for each of the 18 specimens. We enumerated genes expressed, and calculated the correlation coefficients *within* pairs (frozen vs. unfrozen) both for the whole genome, and for three gene sets (Inflammatory [n of genes = 36]; Hypoxic [n=127]; and Thyroidal [n=140]) associated with CP.⁴ The summary correlation coefficients and their confidence intervals were estimated by using a random effects model.^{8,9} For correlation coefficients, we obtained a Z-score and variance for each pair using a Fisher transformation, summarized the Z-scores using a random effects model, and transformed these scores to the final point and interval estimates shown.

Results

On average, 15,864 genes were expressed in the frozen samples, 11,897 in the unfrozen samples, representing, respectively, 75% and 56% of all 22,266 genes in the array (Table I). Of genes expressed in the frozen samples, 69% (95% CI, [57%, 81%]) were detected in the unfrozen sample. A larger fraction of genes expressed in the unfrozen samples were expressed in the frozen (93%, 95% CI, [90%, 95%]). Variability in the number of genes expressed was small for frozen spots (range: 14,589 to 16,744), but genes expressed in the

unfrozen spots ranged more widely (7,017 to 17,911). Two infants (#6 and #8) had especially few genes expressed in their unfrozen samples (7,017 and 7,454) even though the N of genes expressed in their frozen samples (15,984 and 15,829) was nearly identical to the mean of all frozen samples (15,864).

The average pairwise correlation of all genes expressed in both the frozen and unfrozen samples was 0.771; 95% CI [0.700, 0.828] (Table II). In the two infants with few genes expressed, correlations between the frozen and unfrozen samples were lower (0.674 and 0.635). The mean difference in expression level in the paired samples, expressed as log₂ fold changes, (unfrozen minus frozen) was not significantly different from zero (−0.485; 95% CI [−1.175, 0.205]) (Table III; available at www.jpeds.com).

Just 4,316 genes were expressed in all 18 specimens (available from authors upon request), indicating considerably heterogeneity in genes expressed in infants. The average pairwise correlation coefficient for these genes was 0.777; 95% CI [0.712, 0.828].

Findings in the three gene sets of interest paralleled those for the whole genome. Of the 301 genes in these sets, 263 (87%) were expressed in the frozen sample and 189 (62%) in the unfrozen sample. Of genes expressed in frozen samples, 70% (95% CI, [69%, 83%]) were detectable in unfrozen samples. The within-pair correlation coefficients for expression of the gene sets of interest were nearly identical to those found for the whole genome (0.768, 95% CI: [0.686, 0.830]) (Table II). As in the whole genome analysis, the expression levels for the three gene sets within pairs (unfrozen minus frozen) did not differ (−0.56, 95% CI: [−1.222, 0.097]) (Table III). The two infants whose unfrozen spots showed fewer genes overall, also had fewer genes expressed in the gene sets of interest and lower correlations between the two specimen types (Table II)

Discussion

After 8–10 years of storage, 30% of genes detectable in blood spots at −70°C were not detectable in unfrozen spots prepared at the same time, but genes that were expressed were very highly correlated in frozen and unfrozen spots, indicating that unfrozen spots generally provided a reasonable approximation of the frozen state, both for all genes expressed and for select gene sets of interest. In specimens with greater loss of gene signal, the degree of approximation was reduced. We cannot know from these data the relative value of storage at −20°C compared with −70°C, nor whether these findings extend beyond very premature infants, although morbidity or gestation age-dependent regulation of mRNA decay or stability factors such as microRNA are poorly understood and should be studied in the preterm and term newborn.

In studies comparing cases and controls, archived spots should be comparable in age and storage conditions, minimizing case-control differences in loss of gene expression. In that scenario, power would be reduced, but false positive results unlikely.

Although blood spots frozen −70°C provide a more complete picture of perinatal gene expression, archived unfrozen newborn blood spots should still prove useful for interrogating perinatal gene expression in childhood disorders.

Acknowledgments

Funded by National Institutes of Health (R01 NS055101 to N.P.). The frozen blood spot specimens were collected under the auspices of the ELGAN study (U01NS040069 to A.L.). The unfrozen blood spot specimens were provided by Michigan Department of Community Health.

References

1. Olney RS, Moore CA, Ojodu JA, Lindegren ML, Hannon WH. Storage and use of residual dried blood spots from state newborn screening programs. *J Pediatr*. 2006; 148:618–622. [PubMed: 16737872]
2. Haak PT, Busik JV, Kort EJ, Tikhonenko M, Paneth N, Resau JH. Archived Unfrozen Neonatal Blood Spots Are Amenable to Quantitative Gene Expression Analysis. *Neonatology*. 2009; 95:210–216. [PubMed: 18799893]
3. Resau JH, Ho NT, Dykema K, Faber MS, Busik JV, Nickolov RZ, et al. Evaluation of Sex-Specific Gene Expression in Archived Dried Blood Spots (DBS). *International Journal of Molecular Sciences*. 2012; 13:9599–9608. [PubMed: 22949818]
4. Ho NT, Furge K, Fu W, Busik J, Khoo SK, Lu Q, et al. Gene expression in archived newborn blood spots distinguishes infants who will later develop cerebral palsy from matched controls. *Pediatr Res*. 2013; 73:450–456. [PubMed: 23269123]
5. Slaughter J, Wei C, Korzeniewski SJ, Lu Q, Beck JS, Khoo SK, et al. High correlations in gene expression between paired umbilical cord blood and neonatal blood of healthy newborns on Guthrie cards. *J Matern Fetal Neonatal Med*. 2013 May 14. [Epub ahead of print].
6. Patterson TA, Lobenhofer EK, Fulmer-Smentek SB, Collins PJ, Chu TM, Bao WJ, et al. Performance comparison of one-color and two-color platforms within the MicroArray Quality Control (MAQC) project. *Nat Biotechnol*. 2006; 24:1140–1150. [PubMed: 16964228]
7. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*. 2003; 19:185–193. [PubMed: 12538238]
8. Dersimonian R, Laird N. Metaanalysis in Clinical-Trials. *Controlled Clin Trials*. 1986; 7:177–188. [PubMed: 3802833]
9. Brockwell SE, Gordon IR. A comparison of statistical methods for meta-analysis. *Stat Med*. 2001; 20:825–840. [PubMed: 11252006]

Table 1

Number and percent of genes expressed in frozen and unfrozen specimens

Pair number	genes expressed in unfrozen sample (N)	% of all 21,266 genes in array	genes expressed in frozen sample (N)	% of all 21,266 genes in array	% of all genes expressed in frozen that were also expressed in unfrozen	% of all genes expressed in unfrozen that were also expressed in frozen
1	13,009	61.2%	15,786	74.2%	76.3%	92.6%
2	10,277	48.3%	15,588	73.3%	62.0%	94.1%
3	12,929	60.8%	14,589	68.6%	81.9%	92.4%
4	11,640	54.7%	17,372	81.7%	64.9%	96.9%
5	11,550	54.3%	15,285	71.9%	68.8%	91.0%
6	7,454	35.1%	15,984	75.2%	43.9%	94.1%
7	15,287	71.9%	15,606	73.4%	88.8%	90.7%
8	7,017	33.0%	15,829	74.4%	42.7%	96.4%
9	17,911	84.2%	16,744	78.7%	94.0%	87.9%
Column mean	11,897	55.9%	15,865	74.6%	69.3%	92.9%

Table 2

Correlation Coefficients of gene expression for each frozen-unfrozen pair

Pairs	Whole genome			Three gene sets of interest		
	N of genes expressed in both spots	R*	95% CI	N of genes expressed in both spots	R*	95% CI
1	12,049	0.782	0.775–0.789	197	0.837	0.789–0.874
2	9,667	0.716	0.706–0.726	175	0.775	0.708–0.828
3	11,948	0.896	0.892–0.900	194	0.902	0.872–0.926
4	11,274	0.736	0.727–0.744	183	0.683	0.597–0.754
5	10,513	0.738	0.729–0.747	180	0.705	0.624–0.772
6	7,017	0.674	0.661–0.686	109	0.617	0.484–0.721
7	13,858	0.882	0.878–0.886	213	0.859	0.819–0.890
8	6,763	0.635	0.620–0.649	127	0.557	0.424–0.667
9	15,738	0.750	0.743–0.757	264	0.769	0.714–0.814
Mean**	10,981	0.771	0.700–0.828	182	0.768	0.686–0.830

* Correlation coefficients of gene expression among commonly expressed genes in frozen and unfrozen pairs

** Average values of point and interval estimate for correlations are obtained by using a random-effect meta-analysis model and a Fisher transformation.

Table III

Additional statistics for each pair

whole genome				
pair	Pws_diff	Pws_sd	Pws_L95	Pws_U95
1	0.0267976	1.044432	-2.02025	2.073847
2	-0.8745799	1.22428	-3.27412	1.524965
3	-0.3120615	0.749434	-1.78093	1.156802
4	-1.8395134	1.275882	-4.3402	0.661169
5	-1.2632413	1.203479	-3.62202	1.095534
6	-1.1340645	1.391089	-3.86055	1.59242
7	-0.1594933	0.752218	-1.63381	1.314828
8	-0.3136327	1.602922	-3.4553	2.828036
9	0.2588777	1.170426	-2.03512	2.552871
summary	-0.4853291	0.351965	-1.17517	0.204511
3 gene sets				
1	0.1727529	0.816749	-1.42805	1.773551
2	-0.7380517	1.000885	-2.69975	1.223648
3	-0.2387602	0.67582	-1.56334	1.085823
4	-2.3963275	1.421549	-5.18251	0.389857
5	-1.6052553	1.31615	-4.18486	0.974352
6	-1.5994686	1.554078	-4.6454	1.446468
7	-0.3875994	0.783097	-1.92244	1.147243
8	-0.9967203	1.703444	-4.33541	2.341969
9	-0.3148442	1.205413	-2.67741	2.047722
summary	-0.5621628	0.336403	-1.2215	0.097176

#Pws_diff represents the averaged pair-wise difference for gene expression without normalization

#Pws_sd represents the standard deviation of pair-wise difference for gene expression without normalization

#Pws_L95 and Pws_U95 represents the lower bound and upper bound of 95% CI for pair-wise difference for gene expression without

#percentages or averaged differences for 9 pairs are summarized by meta-analysis using random effect model