Effect of Long-term Storage in TRIzol on Microarray-Based Gene Expression Profiling

Wencai Ma¹, Michael Wang¹, Zhi-Qiang Wang¹, Luhong Sun¹, David Graber¹,², Jairo Matthews¹, Richard Champlin¹,²,³, Qing Yi¹,³, Robert Z. Orlowski¹,³, Larry W. Kwak¹,³, Donna M. Weber¹, Sheeba K. Thomas¹, Jatin Shah¹, Steven Kornblau¹,², and Richard E. Davis¹

Abstract

Background: Although TRIzol is widely used for preservation and isolation of RNA, there is suspicion that prolonged sample storage in TRIzol may affect array-based gene expression profiling (GEP) through premature termination during reverse transcription.

Methods: GEP on Illumina arrays compared paired aliquots (cryopreserved or stored in TRIzol) of primary samples of multiple myeloma (MM) and acute myeloid leukemia (AML). Data were analyzed at the “probe level” (a single consensus value) or “bead level” (multiple measurements provided by individual beads).

Results: TRIzol storage does not affect standard probe-level comparisons between sample groups: different preservation methods did not generate differentially expressed probes (DEP) within MM or AML sample groups, or substantially affect the many DEPs distinguishing between these groups. Differences were found by gene set enrichment analysis, but these were dismissible because of instability with permutation of sample labels, unbalanced restriction to TRIzol aliquots, inconsistency between MM and AML groups, and lack of biological plausibility. Bead-level comparisons found many DEPs within sample pairs, but most (73%) were <2-fold changed. There was no consistent evidence that TRIzol causes premature reverse transcription termination. Instead, a subset of DEPs were systematically due to increased signals in TRIzol-preserved samples from probes near the 5′ end of transcripts, suggesting better mRNA preservation with TRIzol.

Conclusions: TRIzol preserves RNA quality well, without a deleterious effect on GEP. Samples stored frozen with and without TRIzol may be compared by GEP with only minor concern for systematic artifacts.

Impact: The standard practice of prolonged sample storage in TRIzol is suitable for GEP. Cancer Epidemiol Biomarkers Prev; 19(10); 2445–52. ©2010 AACR.

Introduction

The TRIzol reagent from Invitrogen is a monophasic solution of phenol and guanidine isothiocyanate, based on the single-step RNA isolation method developed by Chomczynski and Sacchi (1). TRIzol is widely used for the isolation of RNA, and investigators often use it for preservation as well, placing fresh samples into TRIzol for freezing and storage at −80°C, and then thawing the samples later for completion of the RNA isolation procedure. However, there is undocumented suspicion that prolonged storage of tissues before RNA isolation, even at −80°C, can cause chemical modification (depuration) of RNA. In theory, this could result in early termination during reverse transcription of mRNA molecules, potentially affecting gene expression profiling (GEP) more strongly for transcripts with probes located farther from the 3′ end. We decided to investigate effects of long-term TRIzol storage by comparing paired aliquots of primary tumor samples collected for research purposes, and frozen in TRIzol or viably frozen with DMSO.

Materials and Methods

Samples

Bone marrow aspirate samples of acute myeloid leukemia (AML) were collected and processed by the Leukemia Sample Bank at M. D. Anderson Cancer Center between 1997 and 2006 under an Institutional Review Board approved protocol and stored in liquid nitrogen freezers. Informed consent was obtained in accordance with the Declaration of Helsinki. Primary tumor samples of multiple myeloma (MM) have been similarly collected.

Authors’ Affiliations: ¹The Myeloma Tissue and Leukemia Satellite Sample Banks, Departments of Lymphoma and Myeloma and ²Stem Cell Transplantation, and ³Center for Cancer Immunology Research, University of Texas M. D. Anderson Cancer Center, Houston, Texas

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Corresponding Author: R. Eric Davis, Department of Lymphoma and Myeloma, M. D. Anderson Cancer Center, SCHR B1.2015, Box 903, 7455 Fannin Street, Houston, TX 77030. Phone: 713-745-8797; Fax: 713-792-0332. E-mail: redavis1@mdanderson.org

doi: 10.1158/1055-9965.EPI-10-0565

©2010 American Association for Cancer Research.
for processing and storage under an Institutional Review Board approved protocol by the Myeloma Tissue Bank since its inception in 2007. From these banks we selected samples of AML (Ficoll-purified, CD3-depleted, and CD19-depleted) and MM (CD138+) for which paired aliquots had been stored frozen in TRIZol (“Tri”) or cryopreservation medium (“Cryo”; RPMI + 20% FCS + 10% DMSO) since the time of initial isolation (range, 1-9 y).

RNA isolation, cRNA generation, and microarray data generation

Cryo aliquots were quickly thawed, washed, and placed in TRIZol, and then total RNA was isolated from all aliquots according to the TRIZol manufacturer’s instructions (Invitrogen). RNA quality was assessed with a 2100 Bioanalyzer (Agilent). RNA from 12 sample pairs (6 AML, 6 MM), all with an RNA integrity number >7, was further purified with Qiagen RNeasy columns. From 300 ng of total RNA, the Illumina TotalPrep RNA Amplification kit (Ambion) was used to generate amplified biotinylated cRNA after reverse transcription by the Eberwine procedure. cRNA (750 ng) was hybridized overnight to Illumina HT-12 BeadArrays, which were then washed and stained with streptavidin-Cy3 (Amersham-Pharmacia Biotech) according to the Illumina protocol. Arrays were scanned on a BeadArray Reader (Illumina) at the Biomarker Core Laboratory Services of the University of Texas Health Science Center in Houston. The six MM Tri/Cryo paired samples were hybridized to one HT-12 array, and the six AML paired samples to another array.

Probe-level GEP data processing

Specific transcripts within the biotinylated cRNA are measured by fluorescent imaging after direct hybridization to HT-12 bead arrays, which contain 12 arrays per slide, each with an average of 15 beads for each of 48,803 probes measuring 37,846 annotated genes and additional transcripts. Raw measurements of the intensity of each bead were captured directly and processed as “bead-level” for the MM samples, as described below. For both MM and AML samples, measurements were processed as “probe-level” data by GenomeStudio software (Illumina). GenomeStudio checks that a probe has ≥3 beads present on the array (if not, the probe is considered to be missing), does a local background subtraction for each bead, and then condenses bead-level data to a single probe-level value per probe by removing outliers >3 median absolute deviations from the median, recalculating the mean of the remaining values (2). Raw probe-level values were extracted from GenomeStudio, without the use of its correction or normalization options, but with the use of its option for imputing missing values. Values for each were then corrected with the use of the model-based background correction method, which uses values for negative control probes to estimate and remove the nonspecific signal component for each transcript probe, with nonparametric estimation (3). Corrected values from the two arrays were combined into an integrated array, containing 12 Tri/Cryo pairs, and quantile normalized (4). Nonsignificantly detected probes are defined by GenomeStudio as having intensity no greater than that of a negative control probe (P < 0.05, based on the normal distribution, and the mean and SD of negative control probes (5). Intensities of nonsignificantly detected probes in a particular sample aliquot were reset to the P = 0.05 value of the negative control distribution (i.e., the detection threshold) so that fold-change differences between paired significantly and nonsignificantly detected probes were kept to a conservative minimum. We further excluded control probes and the ∼20% of probes whose sequences are not perfect or good matches to actual transcripts (6) and log2-transformed the probe-level data.

Bead-level GEP data processing

To process bead-level data from the MM sample array and take advantage of the statistical power inherent in many measurements of each transcript (i.e., one for each bead), we used a procedure that makes use of open-source and original software, modified from the approach of Dunning et al. (7):

1. Extract bead level data.
2. Do local background correction.
3. Discard outlier beads >3 median absolute deviations from the median of transcript probes.
4. Do model-based background correction of bead values.
5. Discard transcript probes with <3 remaining beads.
6. Do quantile normalization of bead values for the entire integrated array.
7. Exclude control probes (for biotin, housekeeping genes, Cy3 hybridization, labeling, low-stringency hybridization, and negative controls).
8. Exclude the ∼20% of probes that are not perfect or good matches to actual transcripts (6).
9. Do log2 transformation of bead-level data.

Comparative analysis of TRIZol effect

Exploratory analysis to find differentially expressed probes (DEP) was done separately for probe-level and bead-level data. Probe-level data from groups of MM and AML samples were compared by t-tests with false discovery rate correction for multiple comparisons (8), and included paired t-tests when appropriate. For bead-level data, comparisons were made between Tri/Cryo pairs for each of the six MM samples by t-tests with false discovery rate control, excluding probe comparisons in which the bead summary value (defined as the mean) for both arrays of a pair was beneath the detection threshold (calculated similarly to the GenomeStudio method).

Analysis of TRIZol effect relative to probe location

For all probes, we determined the distance between the probe target sequence and the 5′ end of its target transcript,
the “5′ end distance,” which is simply the start site nucleotide number of the probe (from “PROBE_START” in GenomeStudio). We also calculated the distance between the probe target sequence and the 3′ end, the “3′ end distance,” by subtracting the start site nucleotide number from the full length of the mRNA (from human.rna.gpff, downloaded from the National Center for Biotechnology Information). These were then used to examine a positional effect of TRIzol by three methods of analysis:

1. Bead-level DEPs were divided into Up-DEPs and Down-DEPs according to whether the bead summary value was higher or lower, respectively, in Tri versus Cryo. The Wilcoxon rank sum test was then used to determine whether the distribution of 3′ end distances differed between the Up-DEPs and Down-DEPs.

2. Differences in probe intensity between Tri and Cryo sample pairs were examined as a function of the probe distance from the 5′ or 3′ transcript end. For a given sample pair, absolute differences between probe-level intensity values for all probes with at least one significantly detected value were summed in a running cumulative score with increasing distance from the 5′ or 3′ end.

3. Pairs of probes recognizing the same transcript were examined on whether the ratio of their probe-level intensities (5′:3′) for Tri arrays differed from those for the corresponding Cryo arrays. For each gene, it was required that all four measurements (two probes each for both Tri and Cryo) were significantly detected. For genes in which there were >2 probes meeting these criteria, a single comparison was made between the most 5′ probe and the most 3′ probe.

**Results**

**TRIzol helps to preserve RNA quality**

The RNA integrity number, a measure of RNA quality generated by the Agilent 2100 Bioanalyzer, was ≥9 for all 12 AML Tri, 6 of 12 AML Cryo, all 6 MM Tri, and 3 of 6 MM Cryo samples (Supplementary Table S1). Six pairs of AML samples with high RNA integrity number values and all six MM pairs were processed further to generate cRNA.

**TRIzol does not negatively affect cRNA generation**

Bioanalyzer-generated histograms assessing cRNA quality were similar for all sample aliquots, suggesting no early termination of reverse transcription due to TRIzol (data not shown). cRNA quantity was generally slightly higher from Tri samples, but Cryo samples also yielded well more than the 750 ng required for array hybridization (data not shown).

**Probe-level comparisons show little effect of TRIzol storage**

Table 1 shows the results of t-tests and paired t-tests comparing probe-level data from groups of samples, arranged by disease (MM or AML) and preservation method (Tri or Cryo). At a false discovery rate of 0.1, no DEPs were found by within-disease comparisons, whether by group or paired t-tests, whereas many DEPs were found by between-disease comparisons, with substantial overlap (Supplementary Table S2). We also compared within-disease gene expression profiles between Tri and Cryo aliquots by gene set enrichment analysis, an increasingly used method to identify sets of genes that implicate biological processes or other explanations of differences, even when fold changes of individual genes are too low to be significant by t-tests (9). Perhaps due to this lesser requirement for fold change, a large number of gene sets were enriched among genes more highly expressed in Tri aliquots of the six MM samples (data not shown), even with a false discovery rate of 0.25 commonly used for gene set enrichment analysis. However, the validity of these gene set enrichment analysis findings was questioned by several observations: the enriched gene sets did not support a plausible or consistent biological interpretation; they were substantially affected by permutation of sample labels; no gene sets were enriched among genes less highly expressed in Tri aliquots of MM samples; and of the few gene sets enriched in the six AML samples, only two were also enriched in MM samples.

<table>
<thead>
<tr>
<th>t-test</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Down-DEPs</th>
<th>Up-DEPs</th>
<th>Total probes</th>
<th>Total DEPs, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paired MM/Tri</td>
<td>MM/Cryo</td>
<td>0</td>
<td>0</td>
<td>13,976</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Paired AML/Tri</td>
<td>AML/Cryo</td>
<td>0</td>
<td>0</td>
<td>15,972</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Group MM/Tri</td>
<td>MM/Cryo</td>
<td>0</td>
<td>0</td>
<td>13,976</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Group AML/Tri</td>
<td>AML/Cryo</td>
<td>0</td>
<td>0</td>
<td>15,972</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Group MM/Tri</td>
<td>AML/Tri</td>
<td>3,613</td>
<td>2,803</td>
<td>15,990</td>
<td>40.13</td>
<td></td>
</tr>
<tr>
<td>Group MM/Tri</td>
<td>AML/Cryo</td>
<td>3,341</td>
<td>2,697</td>
<td>16,016</td>
<td>37.70</td>
<td></td>
</tr>
<tr>
<td>Group MM/Cryo</td>
<td>AML/Cryo</td>
<td>3,468</td>
<td>2,995</td>
<td>15,971</td>
<td>40.47</td>
<td></td>
</tr>
<tr>
<td>Group MM/Cryo</td>
<td>AML/Tri</td>
<td>3,704</td>
<td>3,093</td>
<td>16,033</td>
<td>42.39</td>
<td></td>
</tr>
</tbody>
</table>
We also considered the effect of length of storage in TRIzol, which for AML samples varied from 2 to 9 years (all MM samples had been stored for <2 y). The correlation coefficient for pairs of AML samples, considering all eligible probes (those with at least one significantly detected value), ranged from 0.957 to 0.988, with no consistent effect of storage time. There was a slight decrease in the slope of the Cryo/Tri regression line over time, from 0.982 to 0.956.

**Bead-level comparison of MM samples shows an effect of TRIzol on GEP results**

To test further whether TRIzol storage has an effect on GEP results, we then used the greater statistical power of comparisons based on processed bead-level data, available only for the six MM samples. As shown in Table 2, we found a surprisingly large number of DEPs for all eligible probes when Tri/Cryo MM sample pairs were compared by t-tests. Comparisons of GEP data usually consider not just the number of DEPs but their fold change. As shown in Fig. 1, the log2 magnitude of fold change in DEPs between MM pairs was generally small, with the majority being <1 (i.e., either a 2-fold increase or decrease). Figure 1 also shows that with some exceptions, the DEPs of highest fold change were mostly in the Up-DEP group (i.e., those higher in the Tri aliquot).

To assess whether these differences in bead-level data were “real” as opposed to being an artifact of excessive statistical power, we subdivided bead-level data for eligible probes, with at least 10 beads on each array, randomly assigning the beads into two subarrays. Comparisons showed many DEPs for between-sample subarrays but not for within-sample subarrays, confirming the difference in data between sample pairs (Table 3).

Next, to determine whether the particular DEPs observed were systematic (i.e., occurring in multiple sample comparisons more frequently than expected from chance), we used hypergeometric distribution tests to examine the overlap of DEPs found in pairs of sample comparisons.

The number of eligible probes used for the DEP determinations differed between samples, and evaluation of the likelihood significance of overlap was done with the use of the most conservative denominator (i.e., the intersection of eligible probes from each sample). The overlap of DEPs was highly significant for almost all pairwise comparisons of samples 1, 2, 4, and 6 (Supplementary Table S3). Although further analysis showed that the number of overlapping DEPs between these four samples declined considerably as the number of samples compared was increased (data not shown), and DEPs were generally not overlapping in combinations involving samples 3 or 5 (Supplementary Table S3), these results suggest that there was some systematic basis for the DEPs found by bead-level comparisons.

To understand the systematic basis of DEP overlap in samples 1, 2, 4, and 6, we examined DEP intensities with respect to probe location. The distance from the transcript 3′ end potentially provides information about the efficiency of reverse transcription and, therefore, whether the “TRIzol hypothesis” of premature termination of reverse transcription is correct. Table 2 shows that for all sample pairs except #5, the Up-DEPs (higher in TRIzol) and Down-DEPs differed significantly in their 3′ end distance. In pairs 1, 2, 4, and 6, the Up-DEPs were located farther from the 3′ end of their target transcript. This is the opposite of what is expected if the TRIzol hypothesis of premature reverse transcription termination is correct. An alternative hypothesis is that the effect of probe location may be due to the positional tendency of mRNA degradation. This is not straightforward because mRNA degradation can be a specific process by which mRNA stability is regulated in living cells or a potentially artificial process occurring in cells damaged during preservation and handling. Furthermore, in living eukaryotes, general pathways of mRNA degradation can proceed from either end, although the most frequently cited pathway involves initial shortening of the 3′ poly(A) tail followed by decapping of the 5′ methylguanosine cap, exposing the mRNA.

### Table 2. Number of DEPs found in bead-level comparisons between paired MM samples, by t-tests with FDR <0.1

<table>
<thead>
<tr>
<th>Pair</th>
<th>Up-DEPs (%)</th>
<th>Down-DEPs (%)</th>
<th>Farther from 3′ end</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1,212 (8.8)</td>
<td>1,130 (8.2)</td>
<td>Up-DEPs</td>
<td>&lt;2.2e-16</td>
</tr>
<tr>
<td>2</td>
<td>745 (5.6)</td>
<td>1,126 (8.4)</td>
<td>Up-DEPs</td>
<td>1.9e-13</td>
</tr>
<tr>
<td>3</td>
<td>1,174 (8.6)</td>
<td>2,594 (19.0)</td>
<td>Down-DEPs</td>
<td>&lt;2.2e-16</td>
</tr>
<tr>
<td>4</td>
<td>821 (7.0)</td>
<td>712 (6.0)</td>
<td>Up-DEPs</td>
<td>2.5e-5</td>
</tr>
<tr>
<td>5</td>
<td>164 (1.2)</td>
<td>109 (0.8)</td>
<td>Down-DEPs</td>
<td>0.06</td>
</tr>
<tr>
<td>6</td>
<td>894 (6.6)</td>
<td>767 (5.6)</td>
<td>Up-DEPs</td>
<td>&lt;2.2e-16</td>
</tr>
</tbody>
</table>

*Number of DEPs in the indicated direction: “Up” means higher mean value in Tri than in Cryo, “Down” means lower value in Tri. Value in parentheses is the percentage of DEPs among the total number of eligible filtered probes.

†Indicates which DEP set (Up or Down) is significantly farther from the transcript 3′ end.

‡P-value of the Wilcoxon rank test in comparing Up-DEPs vs Down-DEPs in distance from the transcript 3′ end.
body to 5′-to-3′ exonuclease digestion (10-13). In theory, initial shortening of the 3′ poly(A) tail would apply to all transcripts regardless of length, and the effect on the GEP methods used here would be a global reduction on cRNA yield. However, 5′-to-3′ exonuclease digestion would have a biased effect, reducing the intensity for probes near the 5′ transcript end. Perhaps as a result, the GEP literature generally states that mRNA degradation starts at the 5′ end of transcripts (14, 15), and the 3′:5′ intensity ratios for multiprobed “housekeeping” genes on Affymetrix arrays have been used for assessing RNA quality (16-19). These observations specify an alternate hypothesis that there is less mRNA degradation in TRizol-preserved samples than in cryopreserved samples, with better preservation of the 5′ portions of transcripts. This is consistent with the Bioanalyzer RNA integrity number values, which were higher in the Tri samples. By this hypothesis, the preservation method with less degradation is expected to have artifactual Up-DEPs for probes closer to the 5′ end of transcripts, which on average are farther from the transcript 3′ end, and this is what was observed for Tri in samples 1, 2, 4, and 6.

Table 3. Number of DEPs found after subdividing bead-level data

<table>
<thead>
<tr>
<th></th>
<th>1A vs 1B</th>
<th>2A vs 2B</th>
<th>1A vs 2A</th>
<th>1B vs 2A</th>
<th>1A vs 2B</th>
<th>1B vs 2B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1</td>
<td>0</td>
<td>0</td>
<td>482</td>
<td>519</td>
<td>619</td>
<td>584</td>
</tr>
<tr>
<td>Pair 2</td>
<td>1</td>
<td>0</td>
<td>460</td>
<td>506</td>
<td>422</td>
<td>474</td>
</tr>
<tr>
<td>Pair 3</td>
<td>0</td>
<td>0</td>
<td>1,182</td>
<td>1,089</td>
<td>1,093</td>
<td>1,137</td>
</tr>
<tr>
<td>Pair 4</td>
<td>0</td>
<td>0</td>
<td>291</td>
<td>268</td>
<td>261</td>
<td>222</td>
</tr>
<tr>
<td>Pair 5</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>15</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Pair 6</td>
<td>0</td>
<td>0</td>
<td>370</td>
<td>384</td>
<td>362</td>
<td>405</td>
</tr>
</tbody>
</table>

NOTE: For each MM sample pair and each eligible probe with at least 10 bead values in each pair, Tri and Cryo GEP data (designated as “1” or “2,” respectively) were each randomly assigned to either “A” or “B” subarrays. Subarrays were then compared by t-tests with FDR <0.1.
normalization would not eliminate this effect, although it may reduce its magnitude.

To test this alternate hypothesis further, we compared probe-level normalized intensities for all eligible probes on whether the Tri or Cryo value was higher, as a function of the distance from the 5′ end of the corresponding transcripts. A running cumulative score of the intensity differences shows that Tri values were higher near the 5′ end for MM samples 1, 2, 4, and 6 (Fig. 2), supporting the alternate hypothesis for these sample pairs. Sample 5 showed little position effect, whereas sample 3 showed a mixed effect of position. Position effects were inconsistent when using the 3′ end distance and among AML samples (data not shown).

In a third analysis of the relationship between probe position and preservation method, we examined probe-level data for pairs of DEPs that measured the same transcript. For each sample probe pair and preservation method (Tri or Cryo), we divided the intensity of the probe farther from the 3′ end (5′) by that of the probe nearer to the 3′ end (3′). The rationale behind this analysis is that the same object (i.e., the expressed transcript) is being measured by the two probes; if the 5′:3′ ratio is different between the two preservation methods, it provides evidence for differences in the lengths of cRNA generated from the two samples. We limited this analysis to the 325 pairs of probes that targeted the same transcript and exceeded the detection threshold for all 48 measurements (12 samples, 2 arrays each, 2 probes per transcript). For these 325 pairs, we computed the 5′:3′ intensity ratio for each of the 12 samples and preservation methods, and then compared the median of the ratios of Tri arrays with that of Cryo arrays. A total of 209 pairs had higher median ratios for Tri than Cryo, which was significantly more than the 116 pairs with higher Cryo ratios ($P = 1.4E-07$ by the binomial distribution). We then compared the 5′:3′ Tri and Cryo median ratios for each of the 325 probes, creating a “ratio of ratios” (RR). For the 209 probes with a higher Tri 5′:3′ ratio (RR₁), we divided the Tri array ratio by that of the Cryo array, and for the 116 probes with a higher Cryo 5′:3′ ratio (RR₂), we divided the Cryo ratio by that of the Tri array. Figure 3 shows that RR₁ is consistently higher than RR₂. Therefore, not only is the RR more often higher for Tri arrays than for
Cryo arrays (209 versus 116), the magnitude of its increases is greater than those of its decreases. Again, this is the opposite of what is predicted by the TRIzol hypothesis of premature reverse transcription termination and supports the alternative hypothesis that there is less mRNA degradation in TRIzol-preserved samples than in cryopreserved samples.

**Discussion**

Our results support three conclusions:

1. **TRIzol preservation should not produce artifactual differences in the usual way by which GEP is done.** Gene expression data are usually at the probe level (i.e., a single value per probe in each sample) and are commonly analyzed by comparisons between groups of samples (e.g., with the use of t-tests). We found that comparisons between groups of paired samples differing in preservation method, even with the use of paired t-tests, failed to generate DEPs at a false discovery rate of 0.1. We found significant enrichment of gene sets by gene set enrichment analysis, which is more sensitive to changes of low fold change; however, certain unusual features of these gene sets (sensitivity to permutation, lack of biological plausibility, etc.) should serve to prevent their being interpreted as true differences.

2. **TRIzol preservation may produce differences by bead-level comparisons with samples not preserved in TRIzol, but those differences are mostly of low fold magnitude.** The number of DEPs in paired Tri versus Cryo bead-level comparisons was surprisingly high, although generally of low fold magnitude. Splitting of the data confirmed that this was not the result of high-power bead-level comparison (i.e., there were actual differences in the bead-level data). Because we did not do technical replicates with the same sample and preservation method, we cannot exclude that the origin of these differences is methodologic (i.e., introduced at or downstream of the point of RNA isolation in the GEP process). However, that seems unlikely given the evidence, at least for some samples, that differences were related to probe position. Therefore, our data suggest that bead-level comparisons may be made, with caution (use of fold-change threshold, etc.) and consideration of potential position-dependent effects, between samples irrespective of whether or not they have been preserved in TRIzol.

3. **To the extent that bead-level differences between samples attributable to preservation method (with TRIzol versus without TRIzol) are consistent and of higher fold magnitude, they may be the result of better RNA preservation with TRIzol.** When we did observe a consistent TRIzol effect in bead-level comparisons, it supported the hypothesis that TRIzol has a beneficial effect on RNA quality, with a particular effect on enhancing the intensity of signals from probes near the 5′ end of transcripts. At the least, this is evidence against the hypothesis of depurination and premature reverse transcriptase termination. We compared TRIzol-preserved specimens with their cryopreserved counterparts, the latter of which unavoidably had to undergo thawing and washing before being placed in TRIzol for RNA isolation. This provided an opportunity for RNA degradation not experienced by the TRIzol-preserved specimens, and therefore, differences may have resulted from this process rather than during prolonged low-temperature storage. A better comparison would have been between TRIzol aliquots and aliquots snap-frozen and stored without TRIzol, and then thawed in TRIzol. However, such paired samples were not available to us, and we wanted to test the effects of long-term storage in the “real-world” context of samples from tissue banks whose quality has been shown in studies of other types. Our results suggest that rather than being a potential source of artifacts in GEP analysis, sample preservation in TRIzol may be an optimal method and should be continued by those who have been using it for many years.

**Disclosure of Potential Conflicts of Interest**

No potential conflicts of interest were disclosed.

**Grant Support**

NIH grants 2P01 CA55164-05A1 and P50 CA100632-01 for the sample/tissue bank cores.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 05/29/2010; revised 07/09/2010; accepted 07/11/2010; published OnlineFirst 08/30/2010.
Cancer Epidemiology, Biomarkers & Prevention

Effect of Long-term Storage in TRIzol on Microarray-Based Gene Expression Profiling

Wencai Ma, Michael Wang, Zhi-Qiang Wang, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-10-0565

Supplementary Material
Access the most recent supplemental material at:
http://cebp.aacrjournals.org/content/suppl/2010/08/30/1055-9965.EPI-10-0565.DC1

Cited articles
This article cites 18 articles, 5 of which you can access for free at:
http://cebp.aacrjournals.org/content/19/10/2445.full.html#ref-list-1

Citing articles
This article has been cited by 4 HighWire-hosted articles. Access the articles at:
/content/19/10/2445.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.