Evaluating Alteration of Blood Transcriptome Affected by Storage Conditions Using RNA Sequencing

Eun Jung Koh  
Hanyang University

So Yeon Yu  
Hanyang University

Seung Jun Kim  
GeneTrigger

Eun-II Lee  
Korea University

Seung Yong Hwang (✉ syhwang@hanyang.ac.kr)  
Department of Bio-nanotechnology, Hanyang University, Ansan, Republic of Korea 2Department of Molecular and Life Science, Hanyang University, Ansan, Republic of Korea  
https://orcid.org/0000-0002-9046-8496

Research article

Keywords: RNA stability, Differentially Expressed Genes, RNA sequencing, Whole blood, Quantitative Polymerase Chain Reaction

DOI: https://doi.org/10.21203/rs.3.rs-67616/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

Background

Whole blood is one of the most widely utilized human samples in biological research and is useful for analysing the mechanisms of diverse bio-molecular phenomena. However, owing to its fluidic properties, whole blood is relatively unstable in the frozen state compared to other biopsy samples. Because RNA is structurally unstable, sample damage can severely affect RNA quality, thereby reducing its usability. This study aimed to assess the quality of RNA prepared from blood stored at different temperatures and times prior to freezing, as well as the effect of freezer storage time.

Results

The quality of the RNA derived from different blood samples was assessed by determining the RNA integrity number and RNA sequencing to identify genes (|fold-change (FC)| > 1.5, p-value < 0.05, false discovery rate (FDR) < 0.05) that were differentially expressed between the differently prepared RNA samples. We found that improper sample handling critically influenced both RNA quality and gene expression patterns. In particular, storing blood at room temperature over 12 h before freezing led to RNA degradation. Differential gene expression analysis revealed that expression of the CXCR1 gene was substantially reduced when using impaired RNA.

Conclusions

This study emphasizes the importance of proper sample management for obtaining reliable downstream application outcomes and suggests the CXCR1 gene as a candidate screening marker for RNA damage caused by improper sample handling.

Background

Population-based cohort studies are carried out worldwide to understand the bio-molecular association between various stimulants and their outcomes (e.g., causality between environmental exposure and the effect on human health) [1]. Cohort studies often use human-derived samples which are stored for decades for future use. Deep-freezing is one of the most common methods for storing samples in a stable manner so that their physical integrity can be maintained.

Whole blood is one of the most widely utilized biopsies from which nucleic acids can be obtained for omics studies. Although blood can be accessed in a minimally invasive manner compared to other samples, unprocessed blood is highly vulnerable to the effects of the freezing process because of its fluidic properties. Thus, several strategies for blood acquisition are used to prevent damage to the blood after its collection and freezing. These include using RNA-protective collection tubes and separating blood into peripheral blood mononuclear cells or serum after collection and before freezing. However, these additional processes often have temporal, spatial, financial, or personnel limitations. In cohort
studies, whole blood has been typically collected in EDTA-coated tubes, which protect the blood from clotting without substantially altering its composition \textit{in vitro}. However, since RNA studies are receiving increasing attention, it has become apparent that there are critical limitations to freezing blood collected in EDTA-coated tubes. It is not until about 20 years ago that RNA stabilizing tools have been introduced \cite{2}.

RNA is a nucleic acid that contains the genetic information required to direct the synthesis of its encoded protein. Thus, RNA is an indispensable tool in molecular biology research. However, RNA is unstable and vulnerable to physical damage, whereas DNA is relatively stable \cite{3}. Low-quality RNA crucially affects downstream experiments such as microarrays or next-generation sequencing (NGS) \cite{4}. The quality of a sample is affected from the time of sample collection through processing, as well as by specific properties intrinsic to the sample itself. Therefore, it is important to manage samples properly at each step to maintain the integrity of the original sample \cite{5}.

This study was conducted to emphasize the importance of proper blood sample handling so that blood-derived RNA can be used for its intended downstream applications and provide data that can be reliably interpreted. To achieve this, the effects of several factors related to the blood freezing process, including storage time and temperature prior to freezing and post-freezing storage time, on blood-derived RNA quality and subsequent impact on gene expression, as assessed by RNA sequencing, were determined.

\section*{Results}

\subsection*{RNA quality assessment}

There were no large differences in RNA yield between the EDTA group and PAXgene group, whereas the RNA yield in the EPR group was up to 2-fold lower than other two groups. In contrast to RNA yield, assessment of RNA quality revealed interesting results. The impact of the type of collection tube on RNA quality was obvious. As the PAXgene tube contains RNA-stabilizing reagents, RNAs from the PAXgene group had RNA integrity number (RIN) of around 7, which is an appropriate quality for downstream applications. In contrast, RNAs in the EDTA group had RIN of around 2 or were not available (N/A), making them unusable for any further experiments. Once frozen, RNA in blood collected in the EDTA tube was damaged regardless of the freezing duration. Unexpectedly, however, the RNA quality in the EPR group derived from the EDTA group showed an RIN of $\geq 6$, which is acceptable for further studies. Although blood is collected in an EDTA tube and frozen, its RNA can be saved from damage to some extent using an additional correction step. The yield and RIN for the RNA in each group are summarized in Table 1.
Table 1
RNA yield and RIN summary for the three groups. All values are shown as the mean ± S.D (standard deviation).

|                  | EDTA group           | PAXgene group         | EPR group           |
|------------------|----------------------|-----------------------|---------------------|
|                  | Yield (ng)           | RIN                   | Yield (ng)          | RIN                   | Yield (ng)          | RIN                   |
| Fresh blood      | 333.3 ± 63.0         | 6.1 ± 0.5             | 302.7 ± 73.0        | 7.57 ± 0.2            |                      |                      |
| Frozen for a month |                     |                       |                     |                       |                      |                      |
| Immediate freezing | 450 ± 88.6         | 2.6 ± N/A             | 450.7 ± 44.7        | 6.9 ± 0.2             |                      |                      |
| 4°C for 12 h     | 394 ± 41             | N/A                   | 322.7 ± 35.9        | 7.1 ± 0.1             |                      |                      |
| 4°C for 48 h     | 283.3 ± 41.2         | N/A                   | 356.7 ± 26.6        | 7.1 ± 0.2             |                      |                      |
| RT for 12 h      | 323.3 ± 38.2         | N/A                   | 334 ± 36.7          | 7.3 ± 0.0             |                      |                      |
| RT for 48 h      | 419.3 ± 211.2        | 2.6 ± N/A             | 356.7 ± 54          | 6.8 ± 0.1             |                      |                      |
| Frozen for 6 months |                 |                       |                     |                       |                      |                      |
| Immediate freezing | 463.5 ± 6.4         | N/A                   | 329.3 ± 44.2        | 7.3 ± 0.2             |                      |                      |
| 4°C for 12 h     | 333 ± 25.5           | N/A                   | 456 ± 96.2          | 6.8 ± 0.1             |                      |                      |
| 4°C for 48 h     | 370.5 ± 142.1        | N/A                   | 308 ± 81.7          | 7.1 ± 0.2             |                      |                      |
| RT for 12 h      | 535.5 ± 44.5         | 1.6 ± N/A             | 389.3 ± 47.7        | 6.9 ± 0.4             |                      |                      |
| RT for 48 h      | 544.5 ± 61.5         | N/A                   | 394.7 ± 206         | 6.7 ± 0.2             |                      |                      |
| Frozen for 12 months |             |                       |                     |                       |                      |                      |
| Immediate freezing | 444 ± 79.5         | 7.6 ± 0.2             | 220.5 ± 57.3        | 6.3 ± 0.1             |                      |                      |
| 4°C for 12 h     | 354 ± 109.5          | 7.6 ± 0.2             | 276 ± 17            | 6.3 ± 0.1             |                      |                      |
| 4°C for 48 h     | 478 ± 61.8           | 7.6 ± 0.1             | 144 ± 12.7          | 6.1 ± 0.7             |                      |                      |
| RT for 12 h      | 434 ± 58.5           | 7.3 ± 0.2             | 366 ± 17            | 6.2 ± 0.4             |                      |                      |
| RT for 48 h      | 434 ± 97             | 6.9 ± 0.1             | 226.5 ± 74.2        | 6.2 ± 0.1             |                      |                      |
The significance of RNA quality changes was assessed by ANOVA with Bonferroni correction. The EDTA group was excluded from this analysis. In the PAXgene group, two variables were individually restricted to control variables in the two analysis sets: The significance of RNA quality changes affected by long-term freezing was compared by controlling incubation before freezing and vice versa. The former analysis revealed that RNA isolated from blood frozen for more than a month significantly differed in quality compared to Fresh RNA (p < 0.05, Fig. 1a) and the latter analysis revealed that incubation at a certain temperature or duration before freezing affects RNA quality, particularly at RT for 48 h (p < 0.05, Fig. 1b).

In the EPR group, the significance of RNA quality changes based on the incubation time before freezing was assessed. There was no significance between samples (p = 0.986), indicating that the degree of RNA recovery using PAXgene reagent was not affected by the incubation conditions.

**Differentially expressed gene (DEG) analysis**

RNA sequencing was performed to evaluate the consequences of RNA quality changes on further experiment and interpretation. RNAs from the PAXgene and EPR group was sequenced, whereas those in the EDTA group were not.

The read counts were quantile normalized before DEG analysis. Principal component analysis (PCA) showed that the PAXgene and EPR group were distinctly separated from one another. Although the PAXgene group was tightly clustered, the EPR group was distanced from the PAXgene group and distributed more sporadically (Fig. 2).

DEG analysis using several sets from the PAXgene and EPR group was then conducted. For the PAXgene group, Fresh RNA was used as a control sample and the FC was calculated. For the EPR group, there were 3 analysis sets: comparison within the EPR group; between the 12-month frozen PAXgene and EPR group; and between the Fresh RNA and EPR group.

Genes for which the |FC| relative to control was higher than 1.5 were classified as DEGs (p < 0.05, FDR by Benjamini-Hochberg < 0.05). Several DEGs were identified in the PAXgene group, but no significant relationships between the DEGs were observed. In contrast, the three analysis sets in the EPR group showed dynamic gene expression patterns. In the first and second analysis sets, 68 and 80 DEGs were identified in RT48H. The third set showed much more dramatic gene expression patterns overall, with 88−3604 of DEGs identified (Table 2, additional file 1). Although the RNA quality of the EDTA group was rescued by using PAXgene reagent, the actual gene expression patterns were not comparable to those in the PAXgene group.
Table 2
Number of DEGs in the PAXgene and EPR groups for each analysis sets

|                          | IMM | C12H | C48H | RT12H | RT48H |
|--------------------------|-----|------|------|-------|-------|
| **PAXgene group - Frozen for a month** | 7   | 0    | 1    | 8     | 0     |
| Up                       | 0   | 0    | 2    | 5     | 0     |
| Total DEG                | 7   | 0    | 3    | 13    | 0     |
| **PAXgene group - Frozen for 6 months** | 0   | 0    | 0    | 0     | 0     |
| Up                       | 0   | 0    | 0    | 0     | 0     |
| Down                     | 0   | 0    | 0    | 0     | 0     |
| Total DEG                | 0   | 0    | 0    | 0     | 0     |
| **PAXgene group - Frozen for 12 months** | 2   | 0    | 0    | 0     | 0     |
| Up                       | 3   | 2    | 0    | 0     | 0     |
| Down                     | 5   | 2    | 0    | 0     | 0     |
| Total DEG                | 5   | 2    | 0    | 0     | 0     |
| **EPR group – SET1 (Comparison within EPR group)** | 1   | 0    | 3    | 43    |
| Up                       | 0   | 0    | 3    | 25    |
| Down                     | 1   | 0    | 6    | 68    |
| **EPR group – SET2 (12-month frozen PAXgene group vs. EPR group)** | 1   | 0    | 2    | 23    |
| Up                       | 3   | 0    | 4    | 3     |
| Down                     | 4   | 0    | 4    | 5     |
| Total DEG                | 4   | 0    | 4    | 5     |
| **EPR group – SET3 (Fresh RNA vs. EPR group)** | 88  | 108  | 162  | 139   | 1790  |
| Up                       | 237 | 251  | 399  | 264   | 1814  |
| Down                     | 325 | 359  | 561  | 403   | 3604  |

Based on these DEGs, the variables were hierarchically clustered. The EPR group showed the largest separation from the other groups. RT48H of EPR group was distinguished from the other groups (Fig. 3).
Physical (e.g., time or temperature) or chemical (e.g., PAXgene reagent) effects on the samples significantly impacted the gene expression profile, which cannot be detected by RIN evaluation.

DEGs induced in the EPR group were compared to assess whether there is consistent tendency in gene expression. In the third analysis set, 79 genes were differentially expressed in common (Fig. 4a). The RT48Hs across the three analysis sets showed seven genes in common (Fig. 4b). Common DEGs from the third analysis set and RT48H sets were narrowed down to a gene referred to as CXCR1 (Fig. 4c). Interestingly, the FC for CXCR1 in the EPR group was highly dynamic. Although FCs in the PAXgene group were moderate ranging from −1.5 to 1.17, the EPR group FCs ranged from −195.08 to 2.16. Some of these values did not reach statistical significance but appeared to be meaningful. Particularly, the RT48H samples showed dramatic changes of up to nearly 200-fold (Fig. 5, additional file 2).

**CXCR1 validation**

The dynamic expression changes in CXCR1 were verified by qRT-PCR. The primer sets for CXCR1 and GAPDH, the housekeeping gene, are shown in Table 3.

| Sequence | Tm (°C) | Amplicon size (bp) |
|----------|---------|--------------------|
| **Housekeeping gene - GAPDH** | | |
| Forward | 5'-ATGAGAAGTATGACAAACAGCC-3' | 58.8 | 113 |
| Reverse | 5'-GAGTCCTTCCACGATACCA-3' | 59.1 |
| **Gene of interest – CXCR1** | | |
| Forward | 5'-TGTGGAAGGTGGAGAACAC-3' | 60.2 | 133 |
| Reverse | 5'-CAATGATGTGAGAGCAGTTACT-3' | 59.2 |

CXCR1 and GAPDH were amplified from the Fresh RNA, 12-month frozen PAXgene group, and EPR group. GAPDH was amplified to a similar extent in each group with an average Ct value of around 20. In contrast, the Ct of CXCR1 in the EPR group was around 31, which was higher than those in the Fresh RNA and PAXgene groups with average Ct values of around 24. Differential expression of CXCR1 relative to GAPDH in the Fresh RNA was normalized to 1 using the $2^{-\Delta\Delta Ct}$ method and those in the PAXgene group and EPR group were compared. Expression of CXCR1 in the PAXgene group was approximately 1 fold compared to Fresh RNA, whereas CXCR1 expression in the EPR group was 0.01–0.03-fold relative to Fresh RNA, indicating much lower expression (Fig. 6). These data confirm the RNA sequencing results and show that the expression pattern of CXCR1 is dynamically affected by the RNA status.

**Discussion**
Vulnerability of blood and RNA quality

Because of the accelerated development of technologies and significant decrease in costs, omics studies to understand bio-molecular alterations caused by specific stimulations or conditions have become widely performed [6]. It is important to guarantee the quality of the starting materials when determining the mechanisms or causality of bio-molecular phenomena.

Blood is one of the most valuable samples for various analyses but its fluidic properties can lower its integrity under freezing conditions. Other factors that arise from the moment of sample acquisition to the freezing step can also affect the integrity of blood samples. The inherent structural instability and its high vulnerability to many of these factors have particularly important effects on RNA integrity.

For various reasons, such as water crystallization, whole blood cells often burst upon freezing [7]. Upon thawing, different intracellular enzymes including RNases are released from the cell and directly cause RNA degradation. PAXgene reagent purposely induces cell bursting but immediately inactivates RNases; these are the main mechanisms for minimizing RNA degradation in frozen blood. Given that RNA isolated from blood frozen for 12 months was slightly degraded, it appears that PAXgene reagent cannot protect RNA indefinitely. However, the extent of damage was not sufficient to affect the utility of the RNA in downstream applications. In contrast, collection of blood into EDTA tubes, which does not contain RNA stabilizing reagent, showed high levels of RNA degradation. Unexpectedly, adding PAXgene reagent later protected the RNA to some extent [8, 9]. In addition to the type of collection tubes, the temperature and duration before freezing and duration of long-term freezing affected RNA damage.

Typically, the concentration of RNA is determined by measuring the absorbance at 260 nm with a spectrophotometer. However, the signal absorbance at 260 nm is independent of whether the nucleic acid is intact. Thus, although it appeared that there was no large difference in RNA yield between the EDTA group and PAXgene group, the results do not reveal the usefulness of the RNA in the EDTA group. The peak images from the bioanalyzer support this observation. By measuring bioanalyzer, RNAs from EDTA group have accumulated form of peak, indicating RNA degradation. In contrast, those in the PAXgene and EPR group had two narrow sharp peaks in the size of 18S and 28S rRNA, respectively (Additional file 3).

Although RIN is certainly an intuitive and reliable indicator for evaluating RNA quality, additional verification steps are necessary to determine whether RNA can be utilized reliably in downstream applications. In this study, RNA quality and its impact on gene expression patterns were compared beyond the RIN. Simply considering the RIN values, RNAs in the EPR group appear to be useful for omics analyses. However, when sequencing results are also considered, the RNAs in the EPR group do not appear to be completely comparable with those in the PAXgene group. Even within the EPR group, which showed similar RNA quality levels, the gene expression patterns differed greatly. These results imply that the stability of RNAs is affected by factors that cannot be explained only by RIN. Presumably, these factors arise as a result of damage that cannot be controlled by PAXgene reagent or may have occurred because of unexpected chemical effects of mixing reagent on unstable blood.
DEG analysis revealed that the temperature and duration of incubation before freezing or duration of long-term freezing impact RNA degradation, consequently affecting gene expression patterns. Particularly, storing blood at room temperature for over 12 h significantly increases the risk of damage to the sample and RNA integrity. As confirmed in the PCA plot and heatmap (Figs. 3 and 4), RNAs isolated from the blood kept at RT for 12 and 48 h showed significant variations even within the EPR group. Although not distinguishable in the PAXgene group, both the storage temperature and duration are critical factors affecting sample integrity and its ultimate applicability. Therefore, samples should be frozen as soon as possible, preferably within 12 h of storage at 4 °C. To obtain consistent reliable data that can be interpreted, conditions should be maintained as similar as possible between experiments because of the possibility of fluctuations in gene expression levels caused by diverse variables [10].

In DEG analysis, the expression levels of the CXCR1 gene showed dynamic changes in accordance with the condition of the RNA. The CXCR1 gene encodes a member of the G-protein coupled receptor family and is involved in neutrophil activation. Abnormal regulation of expression levels of this gene has been reported to cause diseases such as cancers. Because the blood was derived from a single healthy person in this study, it is unlikely that these expression patterns are related to diseases but rather can be regarded as the result of damage to the RNA because of inappropriate sample handling.

qRT-PCR analysis verified that the CXCR1 gene was expressed at lower levels in the EPR group. This demonstrates that RNA instability critically influences CXCR1 expression. Additionally, the EPR group contained several genes that were also differentially expressed, although they were not found to be common DEGs. Presumably, mixing PAXgene reagent with frozen blood collected in EDTA tubes does not completely reverse molecular damage to the RNA. CXCR1 as well as several other genes are likely subject to damage or degradation by physical or chemical stimuli. The dynamic expression of the CXCR1 gene supports that improper handling of original samples can critically influence RNA quality and its downstream applicability.

On the basis of these results, if there is suspicion on the potential applicability of an RNA sample, and RIN evaluation is not sufficient, a decision can be made regarding the usability of the sample by screening for the expression pattern of specific genes such as CXCR1. Thus, CXCR1 is a candidate indicator for determining the applicability of an RNA sample. Therefore, caution should be exercised when analysing or defining this gene as a biomarker under certain diseases or conditions [11] to clarify whether differential expression of this gene arises from diseases or RNA instability.

One limitation of this study is that only one subject was evaluated. Therefore, further studies of larger sample sizes are needed to verify our results. If possible, several genes for screening sample quality or applicability should be identified and used to assess sample quality, providing a foundation for ensuring that accurate results are obtained in research studies in the clinical or industrial fields.

**Conclusions**
Improper sample handling affects sample quality and can result in inappropriate use of compromised samples, leading to incorrect data interpretation. In particular, in multi-centre cohort studies, it is crucial to establish standardized and strict guidelines for managing samples. This study provides an approach for establishing a proper method for preparing RNA from frozen blood samples to obtain reliable experimental or analytical data.

Methods

Recruitment of a single volunteer

Whole blood was obtained from a single healthy donor. The use of a single donor allowed for assessment of different sample preparation conditions that are not confounded by factors derived from inter-individual variations. Informed consent was obtained from the donor and the study was approved from the Institutional Review Board from Korea University (IRB No. KUIRB-2018-0037-01).

Variables

Several variables have been proposed to impact RNA quality, including type of blood collection tube, temperature and duration of incubation before freezing, and duration of long-term freezing. Two types of collection tubes, a trace element EDTA tube (BD Biosciences, Franklin lakes, NJ, USA, EDTA tube) and a PAXgene Blood RNA tube (PreAnalytiX, Hombrechtikon, Switzerland, PAXgene tube), were used. The inside wall of the EDTA tube is coated with an anti-coagulant but there are no other reagents for RNA stabilization. In contrast, the PAXgene tube contains a proprietary fluidic RNA stabilizing reagent, of which 6.9 mL is intended to be mixed with 2.5 mL of blood. Once mixed with blood, this stabilizing reagent protects RNA in the blood sample from degradation.

The temperature and duration of incubation before freezing were set to mimic real-life samples collection conditions. Blood samples contained in the two types of tube were incubated at 4°C or room temperature (RT) for 12 and 48 h before freezing. Some samples were frozen immediately without incubation after sufficiently mixing the blood with the additives in the tubes.

After incubation, the samples were frozen at -80°C for 1, 6, and 12 months. This study is somewhat limited with respect to the effect of time because storage times longer than 12 months were not analysed. The workflow of this study is shown in Fig. 7a.

Sample collection and storage

Blood in the tube was inverted to ensure proper mixing with the additives contained in each tube. An appropriate volume of blood was dispensed into the vials as follows: 300 µL of blood for the EDTA tube; and 1,128 µL for the PAXgene tube, corresponding to 300 µL of blood and 828 µL of reagent when considering that 2.5 mL of blood is mixed with 6.9 mL of fluidic reagent in the PAXgene tube. Two or
three dispensed vials were assigned to each condition as technical replicates for repetitive experiments. The vials were incubated as described above and frozen for 1–12 months.

**RNA isolation and quality assessment**

RNA was periodically isolated from the blood collected in the EDTA tube and PAXgene tube (EDTA group and PAXgene group, respectively). A QIAamp RNA Blood Mini Kit (Qiagen, Hilden, Germany) and PAXgene Blood RNA Kit (PreAnalytix) were used to isolate RNA from EDTA group and PAXgene group, respectively. RNA isolated from fresh blood collected in a PAXgene tube was used as one the control sample (Fresh RNA). RNA was isolated according to the respective manufacturer’s instructions, which briefly involved cell lysis, homogenization, washing, and elution steps.

In particular, blood collected in an EDTA tube and frozen for 12 months was mixed with PAXgene reagent upon thawing and incubated overnight at RT, after which RNA was isolated using the PAXgene Blood RNA Kit (EPR group) [8]. RNA concentration was measured with a Nanodrop 2000 (Thermo Fisher Scientific, Waltham, MA, USA), and RIN were determined with a 2100 Bioanalyzer and an RNA 6000 Pico Kit (Agilent, Santa Clara, CA, USA). RIN is a numeric indicator that ranges from 1 to 10 and estimates the degree of RNA degradation. A higher RIN value indicates that the RNA is more intact [12]. RIN values for the RNAs were compared and the significance of their differences was statistically assessed by an analysis of variance (ANOVA) and Bonferroni correction (p < 0.05, 95% significance level).

**RNA sequencing and analysis**

In general, it is suggested that RNAs with RINs over 7 are preferred for downstream applications such as NGS. Against this general guideline, some publications have stated that RNAs with an RIN around 5 can also be used without causing any significant bias over high-quality RNA [13]. With reference to these thresholds of RIN 5–7, RNAs were selected for RNA sequencing using the Ion AmpliSeq platform (Thermo Fisher Scientific). Two or three libraries were constructed for each sample according to the manufacturer’s instructions. Ten nanograms of each RNA was reverse-transcribed to cDNA using a SuperScript™ VILO™ cDNA Synthesis Kit (Invitrogen, Carlsbad, CA, USA). From this cDNA, approximately 20,000 known genes were amplified using random primers for targeted RNA sequencing. Adapters were ligated to each end of the amplified cDNA. Libraries were further amplified and purified and then sequenced using an ION Torrent S5 sequencer. Raw data were converted into numeric counts using Transcriptome Analysis Console (TAC) software (Thermo Fisher Scientific). Replicated counts for each sample were averaged.

DEG analyses for PAXgene group and EPR group were conducted as follows: Fresh RNA vs. PAXgene group; comparison within the EPR group (SET1); 12-month frozen PAXgene group vs. EPR group (SET2); and Fresh RNA vs. EPR group (SET3). Fresh RNA is considered as an ideal sample used for comparison with RNAs that have been subjected to physical stimulation (e.g. temperature and duration prior to freezing and long-term freezing) to understand how these factors affect the RNA status and its applicability. Periodically isolated RNAs from PAXgene group were compared with Fresh RNA individually. In contrast, the EPR group contained three analysis sets. In addition to physical stimulation, the EPR
group was subjected to an additional chemical stimulus of mixing of PAXgene reagent, making it necessary to perform numerous analyses. First, comparison within the EPR group was performed to evaluate whether mixing reagent causes notable variations even in the EPR group (SET1). Second, the 12-month frozen PAXgene group and EPR group were compared (SET2). Because these samples were treated under the same conditions except for the type of tube, this analysis revealed the effect of mixing PAXgene reagent with blood considered to be damaged. Third, we compared the Fresh RNA and EPR group (SET3). As described above, Fresh RNA is considered as the ideal RNA sample because it is undamaged and minimally affected. Therefore, this sample can be used to assess the impact of additional process of mixing PAXgene reagent on the recovery of damaged samples and their applicability. The DEG analysis scheme is shown in Fig. 7b.

FCs relative to the control were calculated for each analysis set. Genes for which the absolute FC value was more than 1.5 were selected as DEGs (p < 0.05, FDR < 0.05).

**Quantitative reverse transcription polymerase chain reaction (qRT-PCR)**

The sequence of the a DEG and GAPDH, a housekeeping gene were downloaded in FASTA format from the National Centre for Biotechnology Information (NCBI). The sequences were aligned using the ClustalW Multiple alignment function in BioEdit software. Primer sets were designed by incorporating several factors including melting temperature (T_m), GC content, and amplicon size. The design of the primers was finalized by assessing the possibility of dimer formation or non-specific binding using the IDT OligoAnalyzer tool from Integrated DNA Technologies (Coralville, IA, USA) and primer blast from NCBI [14].

For qRT-PCR, RNA was reverse-transcribed into cDNA and the genes of interest and housekeeping gene were amplified using specific primers using a CFX96 (Bio-Rad, Hercules, CA, USA) and SYBR green dye. Relative gene expression levels were quantified using the ΔΔC_t method [15].

**List Of Abbreviations**

FC
Fold-change; FDR:False discovery rate; CXCR1:C-X-C chemokine receptor type 1;
EDTA:Ethylenediaminetetraacetic acid; NGS:Next-generation sequencing; RIN:RNA integrity number;
DEG:Differentially expressed gene; PCA:Principal component analysis; GAPDH:Glyceraldehyde 3-phosphate dehydrogenase; RT:Room temperature; ANOVA:Analysis of variance; cDNA:Complementary DNA; qRT-PCR:quantitative reverse-transcription polymerase chain reaction; NCBI:National Centre for Biotechnology Information; T_m:melting temperature.

**Declarations**

_Ethics approval and consent to participate_
This study was approved by Institutional Review Board of Korea University (IRB No. KUIRB-2018-0037-01).

Informed consent was obtained from the donor who participated in this study.

**Consent for publication**

The participant and all authors agreed for publication

**Availability of data and materials**

All data generated or analysed during this study are included in this published article and its supplementary information files.

**Conflicts of interest**

All the authors declare that they have no competing interests.

**Funding**

This study was supported by Korea Environment Industry & Technology Institute (KEITI) through "the Environmental Health Action Program", funded by Korea Ministry of Environment (MOE) (2017001360005).

**Authors' contributions**

All authors are involved in the study design. Koh EJ carried out the experiments, data analysis, and drafted the manuscript. Yu SY carried out experiments and provided consultation on data analysis. Kim SJ provided consultation on interpretation of experiment results and data analysis. Lee EI provided access to crucial research components, including recruitment of a sample donor and IRB application for this study and provided consultation on revisions of the manuscript. Hwang SY supervised throughout the study and provided consultation on revisions of the manuscript. All authors read and approved the final manuscript.

**References**

1. Jeong KS, Kim S, Kim WJ, Kim H-C, Bae J, Hong Y-C, et al. Cohort profile: Beyond birth cohort study—The Korean CChildren's ENvironmental health Study (Ko-CHENS). 2019;172:358–66.

2. Augello FA, Rainen L, Walenciak M, Oelmüller U, Wyrich R, Bastian H. Method and device for collecting and stabilizing a biological sample. Google Patents; 2003.

3. Shabihkhani M, Lucey GM, Wei B, Mareninov S, Lou JJ, Vinters HV, et al. The procurement, storage, and quality assurance of frozen blood and tissue biospecimens in pathology, biorepository, and biobank settings. 2014;47(4–5):258–66.
4. Opitz L, Salinas-Riester G, Grade M, Jung K, Jo P, Emons G, et al. Impact of RNA degradation on gene expression profiling. 2010;3(1):36.

5. Elliott P, Peakman TCJjoe. The UK Biobank sample handling and storage protocol for the collection. processing archiving of human blood urine. 2008;37(2):234–44.

6. Koh EJ, Hwang SY. Multi-omics approaches for understanding environmental exposure and human health. Molecular Cellular Toxicology. 2019;15(1):1–7.

7. Yang J, Sui X, Wen C, Pan C, Zhu Y, Zhang J, et al. A hemocompatible cryoprotectant inspired by freezing-tolerant plants. Colloids Surf B. 2019;176:106–14.

8. Beekman JM, Reischl J, Henderson D, Bauer D, Ternes R, Peña C, et al. Recovery of microarray-quality RNA from frozen EDTA blood samples. 2009;59(1):44–9.

9. Kim J-H, Jin H-O, Park J-A, Chang YH, Hong YJ, Lee JKJS. Comparison of three different kits for extraction of high-quality RNA from frozen blood. 2014;3(1):76.

10. Donohue DE, Gautam A, Miller S-A, Srinivasan S, Abu-Amara D, Campbell R, et al. Gene expression profiling of whole blood: A comparative assessment of RNA-stabilizing collection methods. 2019;14(10).

11. Kågedal B, Lindqvist M, Fanebäck M, Lenner L, Peterson CJCC, Medicine L. Failure of the PAXgene™ Blood RNA System to maintain mRNA stability in whole blood. 2005;43(11):1190–2.

12. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. 2006;7(1):1–14.

13. Shen Y, Li R, Tian F, Chen Z, Lu N, Bai Y, et al. Impact of RNA integrity and blood sample storage conditions on the gene expression analysis. OncoTargets therapy. 2018;11:3573.

14. young Kwon N, Ahn JJ, Kim J-H, young Kim S, Lee JH, Kwon J-H, et al. Rapid Subtyping and Pathotyping of Avian Influenza Virus using Chip-based RT-PCR. 2019;13(4):333–40.

15. Livak KJ, Schmittgen TDJm. Analysis of relative gene expression data using real-time quantitative PCR and the \(2^{-\Delta\Delta CT}\) method. 2001;25(4):402–8.

Figures
Figure 1

ANOVA and Bonferroni correction. Significance of changes in RNA qualities affected by (a) long-term freezing and (b) temperature and duration of incubation before freezing.
Figure 2

PCA plot. Quantile normalized counts for the PAXgene and EPR group were distinctly separated into two clusters
Figure 3

Heatmap of DEGs. Each component was clustered using the Euclidean method. Fresh RNA and 12-month frozen PAXgene group showed similar expression patterns. The pattern in the EPR group was slightly distinguished from the Fresh RNA and 12-month frozen PAXgene groups. The RT12H and RT48H in the EPR group was clearly distinguishable from the other groups. Capital P in the bracket at each label indicates the PAXgene group and E indicates the EPR group.
Figure 4

Venn diagrams of DEGs in the EPR group. (a) Seventy-nine genes showed common differential expression in the third analysis set. (b) Seven genes in RT48Hs over three analysis sets. (c) CXCR1 gene was found in an overlap analysis of (a) and (b)
Figure 5

Expression of CXCR1 in each group. FCs of CXCR1 which are less than -1.5 are coloured in blue and larger than -1.5 in grey. Characters in brackets of each comparison indicates each DEG analysis set: 1M for the 1 month-frozen PAXgene group, 6M for the 6-month frozen PAXgene group, 12M for the 12-month frozen PAXgene group
Figure 6

qPCR results. (a) Average Ct values for CXCR1 in each group. (b) Relative quantitation of CXCR1 in the PAXgene and EPR group compared to Fresh RNA.
Figure 7

Workflow of the study. (a) Whole blood was collected into two types of tubes and distributed into several vials in proper amounts. These vials were subjected to different incubation protocols followed by freezing for various durations of long-term storage. (b) DEG analysis of the PAXgene and EPR group. DEG analysis was conducted for each freezing-duration (1, 6, and 12 months) in the PAXgene group. The EPR group had three analysis sets for multifarious comprehension: comparison (1) within the EPR group (SET1); (2) between 12-month frozen in the PAXgene group and EPR group (SET2); and (3) between fresh RNA isolated from fresh blood and EPR group (SET3). The control for each DEG analysis is coloured in blue, and comparative ones are noted in grey. (Fresh RNA; RNA isolated from fresh blood collected in PAXgene tube, IMM; immediately frozen without incubation after sufficiently inverting blood with additives in two types of tube, C12H; 4 °C for 12 hours, C48H; 4 °C for 48 h, RT12H; room temperature for 12 h, and RT48H; room temperature for 48 h)

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile3.pdf
- Additionalfile2.xlsx
- Additionalfile1.pdf