

# Factors affecting blood gene expression

## Introduction

Peripheral blood is a desirable tissue type for research in part because it is so accessible. Although it is simple to obtain, it is not simple for studying gene expression. Because blood gene expression analysis has been increasingly used in biomedical and clinical research, it is important to have a good understanding of the blood cells and factors that may affect the accuracy and consistency of gene expression measurements.

Blood accounts for about 8% of our body weight. It is made up of acellular fluid (plasma) and a mixture of multiple cell types at different stages of their life cycles. The 3 primary blood cell types are:

- Red blood cells (erythrocytes)
- White blood cells (leukocytes)
- Platelets (thrombocytes)

Each type contributes a different amount of RNA to the total RNA in blood. Common blood-processing procedures affect the component cell types, resulting in alterations of blood gene expression measurements.

## Primary blood cell types

### Red blood cells

Red blood cells (RBCs) are the most abundant cells in whole blood, constituting 44% of the blood volume. Mature RBCs lose their nucleus and organelles during maturation and thus do not contribute any RNA to the total blood RNA pool. However, immature RBCs, also known as reticulocytes, may retain some residual nucleic acids. Because reticulocytes constitute about 1% of the RBC population, residual RNA from reticulocytes contributes up to 70% of the RNA in the total blood RNA pool. The most abundant mRNA in reticulocytes is globin mRNA.

### White blood cells

White blood cells (WBCs), or leukocytes, are nucleated blood cells that carry out the immune functions of the body. Leukocytes are the most transcriptionally active cells in blood and are the focus of most blood gene expression studies. Leukocytes are made up of granulocytes, lymphocytes, and monocytes. Granulocytes can be further divided into neutrophils, basophils, and eosinophils. Lymphocytes can be divided into B cells, T cells, and natural killer (NK) cells. The number of cells in each leukocyte subtype can change significantly during disease states such as inflammation and leukemia.

### Platelets

Platelets, or thrombocytes, play a major role in blood clotting. Platelets enter the blood circulation by fragmentation of large bone marrow cells called megakaryocytes. Like red blood cells, platelets lack a nucleus and organelles. Immature platelets, called reticulated platelets, contain residual RNA from megakaryocytes. Up to 4.5% of platelet counts can be reticulated platelets. Platelet RNA therefore also contributes to the total blood RNA pool.

## Cell types in processed blood

Table 1 lists the cell types of human whole blood, the approximate number of each cell type, and its presence or absence after common processing procedures.

## Using QuantiGene Assays for blood samples

Invitrogen™ QuantiGene™ 2.0 or QuantiGene™ Plex 2.0 Assays are tools that accurately and precisely quantitate gene expression in whole blood, capturing the gene expression profile at the time of cell lysis [1]. Manipulations prior to blood cell lysis as well as data processing after performing the assay can affect the results. The most significant factors known to affect results are discussed below.

## Blood sample collection for research studies

The inter- and intra-subject variation is an important issue to consider for blood gene expression studies. Inter-subject variations are referred to as differences in demographics such as age, gender, ethnic background, health and nutritional status, metabolism, and medical history. A study of inter-subject variations in PAXgene™ reagent-stabilized whole blood using Applied Biosystems™ gene chips found that the most variable genes among 32 healthy subjects are predominantly immunoglobulin variable region or related genes [2].

Intra-subject variations come from biological influences within the body such as hormone variation and diurnal changes. Intra-subject variations in expression of ribosomal and protein synthesis genes in human blood have been reported [3]. A recent study of mouse transcriptomics suggests that about 10% of the genes in various mouse tissues have a circadian rhythm in their mRNA expression [4].

To minimize the impact of these variations on blood gene expression analysis, we recommend to include randomized samples in your studies with sufficient sampling size, and standardize collection time to ensure that pre- and post-experimental treatment collections occur at the same time within the day.

## Ex vivo gene expression and blood RNA stabilization

Accurate analysis of *in vivo* gene expression in blood cells may be complicated by changes in gene expression after phlebotomy, caused by sample collection, handling, storage, and uncontrolled coagulation [5]. Intracellular RNA might be rapidly degraded *ex vivo* by specific or nonspecific endogenous nucleases. Unintentional gene expression might be induced as a sensitive response of blood cells to environmental changes such as temperature.

The blood-stabilizing PAXgene reagent was developed to prevent RNA degradation and time-dependent *ex vivo* induction of cytokines and immediate-early response genes [5]. However, an overall gene expression pattern distinct from regular blood leukocytes has been observed in blood stabilized with PAXgene reagent [6], and it has yet to be determined whether PAXgene reagent induces expression changes in blood samples.

Note: The Invitrogen™ QuantiGene™ Sample Processing Kit for blood samples includes a procedure for blood stabilized with PAXgene reagent.

**Table 1. Human whole blood cell types.**

Cell types		Approximate number in 1 $\mu$ L of whole blood	Blood processing				
			Anticoagulated whole blood	PAXgene reagent	RBC-lysed blood	Buffy coat (leukocyte preparation)	Peripheral blood mononuclear cell (PBMC) preparation
Red blood cells	Erythrocytes						
	• Mature • Reticulocytes	5 x 10 <sup>6</sup> 5 x 10 <sup>4</sup>	√	√	—	√ (Reduced number)	—
Platelets		3 x 10 <sup>5</sup>	√	√	√	√ (Reduced number)	—
Leukocytes	Granulocytes						
	• Neutrophils • Basophils • Eosinophils	5,000 35 150	√	√	√	√	—
	Lymphocytes						
	• B cells • T cells • NK cells	300 1,600 300	√	√	√	√	√
Monocytes		250	√	√	√	√	√

## Blood processing and sample preparation

Traditionally, blood gene expression analysis is performed on leukocytes that are fractionated from whole blood to remove interference from red blood cells. Buffy coat preparations contain most of the leukocytes in blood with small amounts of contaminating red blood cells and platelets, whereas PBMCs contain only mononucleated cells in blood (lymphocytes and monocytes), with very few other cell types. Therefore, the gene expression pattern from fractionated blood will be different from whole blood. This is not only because of differences in blood cell composition, but also because fractionation can reportedly alter the expression of select genes within blood cells [7].

RNA processing steps such as reverse transcription, labeling, and amplification can introduce further variability. Reverse transcription is especially problematic, as the type and sequence of primers and the choice of reverse transcriptase enzyme can influence the final gene expression data [8]. This has been an unresolved issue for technologies such as microarray analysis and RT-PCR, which rely on reverse transcription.

QuantiGene 2.0 and QuantiGene Plex 2.0 Assays are the first products on the market that are capable of measuring gene expression directly from whole blood. It involves only blood cell lysis, no blood or RNA processing. The presence of excess red blood cell proteins and RNAs such as globin mRNA do not interfere with the RNA measurement. The QuantiGene Assays directly measure the concentration of target RNA instead of its derivatives such as cDNA or rRNA.

When comparing results from QuantiGene 2.0 or QuantiGene Plex 2.0 Assays to those from other technologies, bear in mind the following potential problems with other methods and technologies:

- Some might require blood purification (erythrocyte lysis, for example) steps prior to RNA extraction
- Some might be subject to interference from globin mRNA
- Most rely on reverse transcriptase to convert mRNA to cDNA

## Data normalization

Normalizing data between samples is recommended because it corrects for small variations in the total number of cells. Typically, sample data are normalized to the expression of an invariant housekeeping gene. However, blood is unique in that it is one of the most variable tissue types in the body. The relative proportions of the different types of blood cells often vary significantly from time to time and from subject to subject, even though the total number of blood cells may not change substantially.

### Example

Suppose two otherwise similar individuals have similar total WBC counts but a 5-fold difference in monocyte counts. (The normal range of monocytes in WBCs is 1–5%.) If a monocyte-specific target gene is assayed and compared, such as a monocyte-specific chemokine induced by an immune stimulation, the induction levels may appear to differ by 5-fold if normalized to GAPDH, but will actually be the same if normalized to a monocyte-specific marker such as CD14. Therefore, it may be necessary to normalize data to common housekeeping genes, blood cell type-specific markers, or both.

## References

1. Zheng Z, Luo Y, McMaster GK (2006) Sensitive and quantitative measurement of gene expression directly from a small amount of whole blood. *Clin Chem* 52:1294–302.
2. Fan H, Hegde PS (2005) The transcriptome in blood: challenges and solutions for robust expression profiling. *Curr Mol Med* 5:3–10.
3. Whitney AR, Diehn M, Popper SJ, et al. (2003) Individuality and variation in gene expression patterns in human blood. *Proc Natl Acad Sci USA* 100:1896–901.
4. Rudic RD, McNamara P, Reilly D, et al. (2005) Bioinformatic analysis of circadian gene oscillation in mouse aorta. *Circulation* 112:2716–24.
5. Rainen L, Oelmueller U, Jurgensen S, et al. (2002) Stabilization of mRNA expression in whole blood samples. *Clin Chem* 48:1883–90.
6. Feezor RJ, Baker HV, Mindrinos M, et al. (2004) Whole blood and leukocyte RNA isolation for gene expression analyses. *Physiol Genomics* 19:247–54.
7. Tamul KR, Schmitz JL, Kane K, et al. (1995) Comparison of the effects of Ficoll-Hypaque separation and whole blood lysis on results of immunophenotypic analysis of blood and bone marrow samples from patients with hematologic malignancies. *Clin Diagn Lab Immunol* 2:337–42.
8. Stahlberg A, Hakansson J, Xian X, et al. (2004) Properties of the reverse transcription reaction in mRNA quantification. *Clin Chem* 50:509–15.

Find out more at [thermofisher.com/quantigene](http://thermofisher.com/quantigene)

**ThermoFisher**  
SCIENTIFIC