Вавиловский журнал генетики и селекции. 2023;27(1):83-87
DOI 10.18699/VJGB-23-11

Original Russian text https://vavilovj-icg.ru/

Influence of human peripheral blood samples preprocessing on the quality of Hi-C libraries

M.M. Gridina[1] ✉, E. Vesna[1], M.E. Minzhenkova[2], N.V. Shilova[2], O.P. Ryzhkova[2], L.P. Nazarenko[3], E.O. Belyaeva[3], I.N. Lebedev[3], V.S. Fishman[1]

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Research Centre for Medical Genetics, Moscow, Russia
[3] Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia
✉ gridinam@gmail.com

Abstract. The genome-wide variant of the chromatin conformation capture technique (Hi-C) is a powerful tool for revealing patterns of genome spatial organization, as well as for understanding the effects of their disturbance on disease development. In addition, Hi-C can be used to detect chromosomal rearrangements, including balanced translocations and inversions. The use of the Hi-C method for the detection of chromosomal rearrangements is becoming more widespread. Modern high-throughput methods of genome analysis can effectively reveal point mutations and unbalanced chromosomal rearrangements. However, their sensitivity for determining translocations and inversions remains rather low. The storage of whole blood samples can affect the amount and integrity of genomic DNA, and it can distort the results of subsequent analyses if the storage was not under proper conditions. The Hi-C method is extremely demanding on the input material. The necessary condition for successfully applying Hi-C and obtaining high-quality data is the preservation of the spatial chromatin organization within the nucleus. The purpose of this study was to determine the optimal storage conditions of blood samples for subsequent Hi-C analysis. We selected 10 different conditions for blood storage and sample processing. For each condition, we prepared and sequenced Hi-C libraries. The quality of the obtained data was compared. As a result of the work, we formulated the requirements for the storage and processing of samples to obtain high-quality Hi-C data. We have established the minimum volume of blood sufficient for conducting Hi-C analysis. In addition, we have identified the most suitable methods for isolation of peripheral blood mononuclear cells and their long-term storage. The main requirement we have formulated is not to freeze whole blood.
Key words: Hi-C; human peripheral blood; blood samples storage.

For citation: Gridina M.M., Vesna E., Minzhenkova M.E., Shilova N.V., Ryzhkova O.P., Nazarenko L.P., Belyaeva E.O., Lebedev I.N., Fishman V.S. Influence of human peripheral blood samples preprocessing on the quality of Hi-C libraries. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2023;27(1):83-87. DOI 10.18699/VJGB-23-11

Влияние предварительной обработки образцов периферической крови человека на качество Hi-C библиотек

М.М. Гридина[1] ✉, Э. Весна[1], М.Е. Миньженкова[2], Н.В. Шилова[2], О.П. Рыжкова[2], Л.П. Назаренко[3], Е.О. Беляева[3], И.Н. Лебедев[3], В.С. Фишман[1]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Медико-генетический научный центр им. академика Н.П. Бочкова, Москва, Россия
[3] Томский национальный исследовательский медицинский центр Российской академии наук, Томск, Россия
✉ gridinam@gmail.com

Аннотация. Метод захвата конформации хроматина в его полногеномном варианте (Hi-C) – мощный инструмент не только для выявления закономерностей пространственной организации генома, но и для понимания влияния их нарушения на развитие заболеваний. Кроме того, метод может быть использован для детекции хромосомных перестроек, в том числе сбалансированных транслокаций и инверсий. Применение метода Hi-C для поиска хромосомных перестроек получает все более широкое распространение. Это связано с тем, что современные высокопроизводительные методы анализа генома позволяют эффективно детектировать точечные мутации и несбалансированные хромосомные перестройки. Однако чувствительность этих методов для определения сбалансированных транслокаций и инверсий остается достаточно низкой. Хранение образцов цельной крови может влиять на количество и целостность выделяемой из них геномной ДНК, а кроме того, приводить к искажению результатов последующих анализов в том случае, если хранение осуществлялось в ненадлежащих условиях. Метод Hi-C крайне требователен к исходному материалу, так как необходимым условием для его успешного применения и получения качественных данных является сохранение пространственной укладки хроматина внутри ядра. Цель нашего исследования состояла в том, чтобы определить оптимальные условия

хранения крови для проведения последующего анализа Hi-C. Были выбраны 10 различных условий хранения образцов крови и пробоподготовки. Для каждого условия приготовлены Hi-C библиотеки и отсеквенированы, после чего оценивалось качество полученных библиотек. В результате сформулированы требования к хранению и подготовке образцов, необходимые для получения качественных Hi-C данных. Нами установлен минимальный объем образца крови, достаточный для проведения Hi-C анализа. Помимо этого, мы определили способы выделения ядерных элементов крови и их долгосрочного хранения, наиболее подходящие для последующего проведения Hi-C анализа. Основное требование, сформулированное нами, – не замораживать цельную кровь.

Ключевые слова: Hi-C; периферическая кровь человека; хранение образцов крови.

Introduction

The combination of chromatin conformation capture methods with whole genome sequencing led to the development of a simple and efficient Hi-C protocol that allows genome-wide studying of chromatin architecture (Lieberman-Aiden et al., 2009; Rao et al., 2014). In addition to the data concerning the organization and dynamics of chromatin in the nucleus, the Hi-C results showed that the relationship between three-dimensional distance in nuclear space and "nucleotide" distance in genomic coordinates can be described by a power function for all cell types. This means that chromosomal rearrangements have effects not only on the contacts frequency of regions directly located at the points of chromosome breaks, but also change the pattern of three-dimensional contacts of a wide area around the rearrangement boundary (Mozheiko, Fishman, 2019). Chromosomal rearrangements detecting methods based on the analysis of the chromatine three-dimensional organization have recently been proposed (Harewood et al., 2017; Chakraborty, Ay, 2018; Díaz et al., 2018; Fishman et al., 2018; Melo et al., 2020). These methods detect various types of rearrangements, including balanced ones, which are still difficult to detect by other methods (Hakim et al., 2012; Dong et al., 2017). In addition, information about single nucleotide variations can be obtained from Hi-C data (Mozheiko, Fishman, 2019), which is important for medical genetics.

Whole blood is a common biological starting sample for medical genetics. Proper blood samples handling is critical for genome-wide studies. Long-term storage and inadequate storage conditions lead to a decrease in the amount of isolated DNA (Nederhand et al., 2003; Malentacchi et al., 2015; Schröder, Steimer, 2018) and its degradation (Ross et al., 1990; Permenter et al., 2015). A high degree of DNA degradation is a serious problem for subsequent molecular biological analyses (Palmirotta et al., 2011; Malentacchi et al., 2015). For example, an increase in the storage time of a blood sample leads to an overestimation of the level of DNA methylation, which may be due to the different stability of methylated and unmethylated DNA (Schröder, Steimer, 2018).

The key steps of the Hi-C protocol are chromatin fragmentation and ligation. To obtain high-quality datasets, it is necessary that both of these steps take place *in nucleus*, under conditions of maximum preservation of the nucleus integrity. Thus, unlike DNA sequence analysis methods, the Hi-C method imposes additional requirements on the quality of the input material. In this regard, it seems relevant to determine the appropriate storage conditions for blood samples intended for Hi-C analysis.

Materials and methods

Peripheral human blood was collected from the antecubital vein into Vacutainer EDTA Blood Collection Tubes. Blood samples storage conditions and preprocessing steps are specified in the Table and Figure 1.

The isolation of peripheral blood mononuclear cells (PBMC) from 3 ml of whole blood was performed using one of the following methods:

- Red Blood Cell Lysis Buffer (RBCL, BioLegend) was used for lysis of erythrocytes according to the manufacturer's instructions. Then the cells were washed once with phosphate buffer saline (PBS).
- centrifugation 300 g for 10 minutes. Serum, including interphase, was transferred into PBS and centrifuged 300 g for 10 minutes.
- sedimentation method on the density gradient Histopaque-1077 Hybri-max (Sigma) according to the manufacturer's instructions.

Cryopreservation of PBMC was performed in a cell freezing medium: 10 % DMSO, 90 % KSR (Thermo Fisher Scientific). Cells were frozen at –80 °C and stored in liquid nitrogen. After thawing, the cells were washed once with PBS.

Cells were counted and resuspended in PBS at a concentration of 1 million cells/ml. Cell fixation, Hi-C library preparation, and data analysis were performed as described in Gridina et al. (2021) using DNase I (Thermo Fisher Scientific) or S1 nuclease (Thermo Fisher Scientific) for chromatin frag-

Storage and preprocessing conditions

#	A brief description of blood samples storage conditions and preprocessing (the time from the blood collection)
1	Less than 4 hrs; RBCLB
2	Less than 4 hrs; RBCLB; freezing KSR+DMSO
3	Less than 4 hrs; centrifugation
4	Less than 4 hrs; centrifugation; freezing KSR+DMSO
5	24 hrs RT; RBCLB
6	–20 °C 4 days; RBCLB
7	+4 °C 2 days; RBCLB
8	+4 °C 4 days; RBCLB
9	+4 °C 7 days; RBCLB
10	Less than 4 hrs; Histopaque-1077 Hybri-max; freezing KSR+DMSO

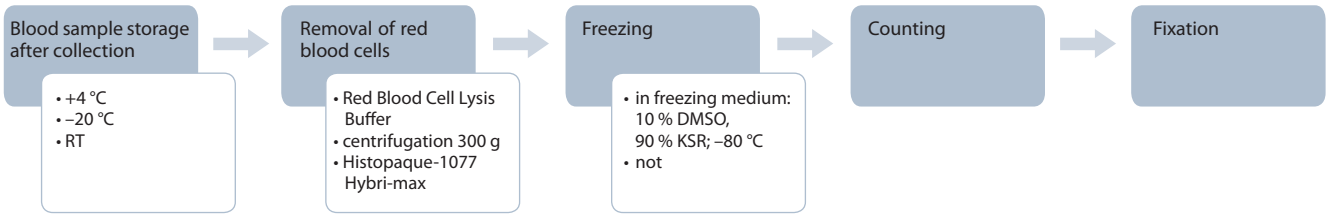

Fig. 1. Blood samples preprocessing.

mentation. HAPA Hyper prep and QIAseq® FX DNA Library Kit (Qiagen) were used for NGS libraries preparation, according to the manufacturer's instructions. The DNA concentration was measured using a Qubit 3.0 fluorimeter (Thermo Fisher Scientific). NGS libraries were sequenced on HiSeq XTen (Illumina) with 150 bp paired reads.

Results and discussion

The first Hi-C step is cells fixation with paraformaldehyde, which is necessary to preserve the native spatial organization of chromatin within the nucleus. Unfortunately, it is not always possible to deliver the sample to the laboratory for fixation on the blood collection day. We decided to systematically estimate the impact of blood storage and preprocessing conditions on the quality of the obtained Hi-C data. Ten conditions were chosen, which included: different methods of PBMC isolation from whole blood; different time and temperature of sample storage; the possibility of freezing PBMC before fixing for long-term storage (see Fig. 1 and the Table).

Although blood sampling is a minimally invasive procedure for biomedical diagnostics, it is clear that there are certain limits on the amount of blood that can be obtained from a patient. Especially if the patient is a small child, or has certain problems with the blood coagulation system. Hi-C analysis requires 1.5–2.5 million cells. Normally, 1 ml of blood contains $(4–11) \times 10^6$ cells. To test each condition, 3 ml of whole blood was taken in two replicates. The PBMC were counted (Fig. 2) after erythrocytes lysis but before cells fixation. A significantly higher number of cells were in the samples processed according to condition #3 (isolation of nuclear elements without RBCL treatment). We did not determine the proportion of living cells during counting. It is possible that dying cells were preserved in samples #3, whereas they were lysed in other cases using RBCL buffer (Brown et al., 2016) or freezing. Supporting the assumption, there were significantly less cells in samples #4 and #10 that were not treated with RBCL but were frozen than in #2.

There were no signs of hemolysis before the start of the isolation of PBMC for all samples except #6, and it was not possible to evaluate this parameter for samples #6. Hemolysis should be avoided, as it is one of the main factors negatively affecting the amount of DNA isolated from blood (Caboux et al., 2012), which may be associated with DNA degradation by nucleases released from degrading cells.

Cell conglomerates were formed in some samples during erythrocyte lysis and subsequent washings. The conglomerates were in both replicates in samples #6, #8, #9 and #10. For these samples, it was not possible to accurately count cells and aliquot them uniformly.

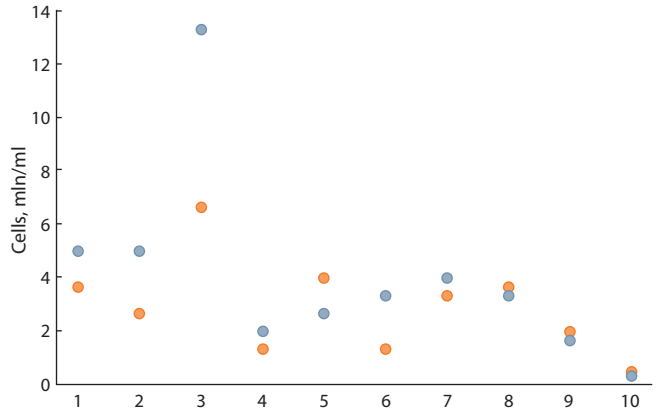

Fig. 2. The PBMC count in 1 ml of blood.

Colors indicate replicates. The horizontal axis represents the storage and preprocessing conditions described in the Table.

2.5 million fixed cells were taken to prepare Hi-C libraries. To assess the quality of Hi-C libraries (Belaghzal et al., 2017), the following controls were made: genomic DNA, DNA after chromatin fragmentation and after ligation. All controls looked accepted (Fig. 3).

We sequenced the Hi-C libraries using paired-end reads with a length of 150 bp, mapped the paired-end reads to the human hg19 genome (GRCh_37) and estimated quality metrics of Hi-C datasets. All libraries had a high proportion of aligned reads (Fig. 4, *a*).

Previously, we have shown (Gridina et al., 2021) that the most important quality metric of Hi-C datasets is the proportion of *cis* interactions (ratio *cis/all* (FF and RR orient)) (see Fig. 4, *b*). It reflects the proportion of Hi-C reads that mapped on the same chromosome among all Hi-C reads. The percentage of *cis* interactions was comparable for all libraries except samples #6 where it was 40.3 and 35.7 %. It means that these Hi-C data are not informative as most fragments ligated randomly. Blood samples #6 were frozen without a cryoprotectant and stored for 4 days at –20 °C. The observed low percentage of *cis* interactions might be due to random DNA strand breaks occurring when cells are frozen without cryoprotectants (Narayanan et al., 2001; Peng et al., 2008; Al-Salmani et al., 2011). On the other hand, this method of freezing leads to ice crystals formation inside the cell and, as a result, to the breaking of cellular and nuclear structures (Mazur, 1984). The release of DNA fragments from the nucleus and their ligation in solution can occur in any way, which leads to the formation of non-informative DNA fragments.

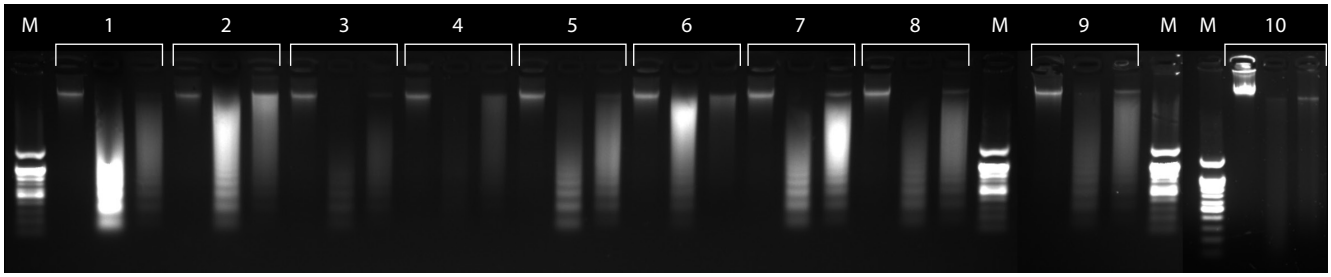

Fig. 3. Chromatin fragmentation and ligation controls in Hi-C experiments.

The numbers represent the storage and preprocessing conditions described in the Table. The order of samples: gDNA, fragmented DNA, ligated DNA. M – DNA ladder 100 bp.

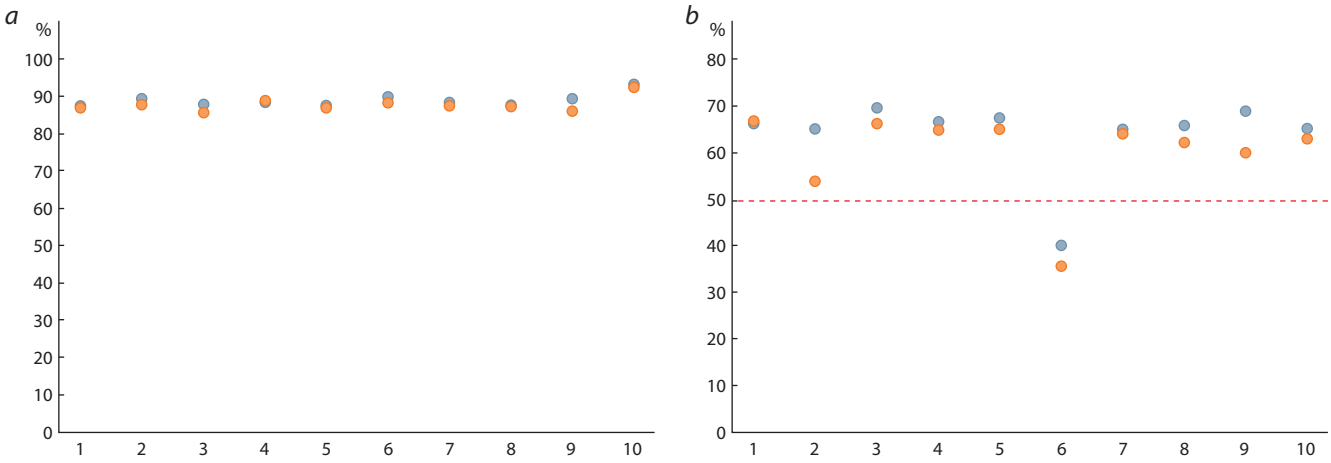

Fig. 4. Quality metrics of Hi-C datasets: *a*, aligned reads; *b*, *cis* interactions.

Colors indicate replicates. The horizontal axes represent the storage and preprocessing conditions described in the Table.

Conclusions

We systematically evaluated various blood samples storage and preprocessing conditions in this work.

As a result, we formulated the following recommendations for the storage and preprocessing of blood samples for Hi-C analysis:

- If it is not possible to deliver the sample on the blood collection day, the samples can be stored at +4 °C for a minimum of 7 days.
- It is better to lyse red blood cells with RBCL buffer before cryopreservation.
- 1–2 ml of whole blood is sufficient (in a person without signs of leukopenia), but if the sample is going to be stored for more than 48 hours, the volume should be increased up to 4–6 ml.
- Never freeze whole blood.

References

Al-Salmani K., Abbas H.H., Schulpen S., Karbaschi M., Abdalla I., Bowman K.J., So K.K., Evans M.D., Jones G.D., Godschalk R.W., Cooke M.S. Simplified method for the collection, storage, and comet assay analysis of DNA damage in whole blood. *Free Radic. Biol. Med.* 2011;51(3):719-725. DOI 10.1016/j.freeradbiomed.2011.05.020.

Belaghzal H., Dekker J., Gibcus J.H. Hi-C 2.0: an optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods*. 2017;123:56-65. DOI 10.1016/j.ymeth.2017.04.004.

Brown W.E., Hu J.C., Athanasiou K.A. Ammonium-chloride-potassium lysing buffer treatment of fully differentiated cells increases cell purity and resulting neotissue functional properties. *Tissue Eng. Part C Methods*. 2016;22(9):895-903. DOI 10.1089/ten.tec.2016.0184.

Caboux E., Lallemand C., Ferro G., Hémon B., Mendy M., Biessy C., Sims M., Wareham N., Britten A., Boland A., Hutchinson A., Siddiq A., Vineis P., Riboli E., Romieu I., Rinaldi S., Gunter M.J., Peeters P.H.M., van der Schouw Y.T., Travis R., Bueno-de-Mesquita H.B., Canzian F., Sánchez M.-J., Skeie G., Olsen K.S., Lund E., Bilbao R., Sala N., Barricarte A., Palli D., Navarro C., Panico S., Redondo M.L., Polidoro S., Dossus L., Boutron-Ruault M.C., Clavel-Chapelon F., Trichopoulou A., Trichopoulos D., Lagiou P., Boeing H., Fisher E., Tumino R., Agnoli C., Hainaut P. Sources of pre-analytical variations in yield of DNA extracted from blood samples: analysis of 50,000 DNA samples in EPIC. *PLoS One*. 2012;7(7):e39821. DOI 10.1371/journal.pone.0039821.

Chakraborty A., Ay F. Identification of copy number variations and translocations in cancer cells from Hi-C data. *Bioinformatics*. 2018;34(2):338-345. DOI 10.1093/bioinformatics/btx664.

Díaz N., Kruse K., Erdmann T., Staiger A.M., Ott G., Lenz G., Vaquerizas J.M. Chromatin conformation analysis of primary patient tissue using a low input Hi-C method. *Nat. Commun*. 2018;9(1):4938. DOI 10.1038/s41467-018-06961-0.

Dong Z., Wang H., Chen H., Jiang H., Yuan J., Yang Z., Wang W.-J., Xu F., Guo X., Cao Y., Zhu Z., Geng C., Cheung W.C., Kwok Y.K., Yang H., Leung T.Y., Morton C.C., Cheung S.W., Choy K.W. Identification of balanced chromosomal rearrangements previously unknown among participants in the 1000 Genomes Project: implications for interpretation of structural variation in genomes and the

future of clinical cytogenetics. *Genet. Med*. 2018;20(7):697-707. DOI 10.1038/gim.2017.170. Epub 2017. Nov. 2.

Fishman V.S., Salnikov P.A., Battulin N.R. Interpreting chromosomal rearrangements in the context of 3-dimentional genome organization: a practical guide for medical genetics. *Biochemistry* (*Mosc.*). 2018;83(4):393-401. DOI 10.1134/S0006297918040107.

Gridina M., Mozheiko E., Valeev E., Nazarenko L.P., Lopatkina M.E., Markova Z.G., Yablonskaya M.I., Voinova V.Y., Shilova N.V., Lebedev I.N., Fishman V. A cookbook for DNase Hi-C. *Epigenetics Chromatin*. 2021;14(1):15. DOI 10.1186/s13072-021-00389-5.

Hakim O., Resch W., Yamane A., Klein I., Kieffer-Kwon K.-R., Jankovic M., Oliveira T., Bothmer A., Voss T.C., Ansarah-Sobrinho C., Mathe E., Liang G., Cobell J., Nakahashi H., Robbiani D.F., Nussenzweig A., Hager G.L., Nussenzweig M.C., Casellas R. DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. *Nature*. 2012;484(7392):69-74. DOI 10.1038/nature10909.

Harewood L., Kishore K., Eldridge M.D., Wingett S., Pearson D., Schoenfelder S., Collins V.P., Fraser P. Hi-C as a tool for precise detection and characterisation of chromosomal rearrangements and copy number variation in human tumours. *Genome Biol*. 2017; 18(1):125. DOI 10.1186/s13059-017-1253-8.

Lieberman-Aiden E., van Berkum N.L., Williams L., Imakaev M., Ragoczy T., Telling A., Amit I., Lajoie B.R., Sabo P.J., Dorschner M.O., Sandstrom R., Bernstein B., Bender M.A., Groudine M., Gnirke A., Stamatoyannopoulos J., Mirny L.A., Lander E.S., Dekker J. Comprehensive mapping of long range interactions reveals folding principles of the human genome. *Science*. 2009;326(5950):289-293. DOI 10.1126/science.1181369.

Malentacchi F., Ciniselli C.M., Pazzagli M., Verderio P., Barraud L., Hartmann C.C., Pizzamiglio S., Weisbuch S., Wyrich R., Gelmini S. Influence of pre-analytical procedures on genomic DNA integrity in blood samples: the SPIDIA experience. *Clin. Chim. Acta*. 2015;440: 205-210. DOI 10.1016/j.cca.2014.12.004.

Mazur P. Freezing of living cells: mechanisms and implications. *Am. J. Physiol*. 1984;247(3):C125-C142. DOI 10.1152/ajpcell.1984.247.3.C125.

Melo U.S., Schöpflin R., Acuna-Hidalgo R., Mensah M.A., Fischer-Zirnsak B., Holtgrewe M., Klever M.-K., Türkmen S., Heinrich V., Pluym I.D., Matoso E., Bernardo de Sousa S., Louro P., Hülsemann W., Cohen M., Dufke A., Latos-Bieleńska A., Vingron M., Kalscheuer V., Quintero-Rivera F., Spielmann M., Mundlos S. Hi-C identifies complex genomic rearrangements and TAD-shuffling in developmental diseases. *Am. J. Hum. Genet*. 2020;106(6):872-884. DOI 10.1016/j.ajhg.2020.04.016.

Mozheiko E.A., Fishman V.S. Detection of point mutations and chromosomal translocations based on massive parallel sequencing of enriched 3C libraries. *Russ. J. Genet*. 2019;55(10):1273-1281. DOI 10.1134/S1022795419100089.

Narayanan S., O'Donovan M.R., Duthie S.J. Lysis of whole blood *in vitro* causes DNA strand breaks in human lymphocytes. *Mutagenesis*. 2001;16(6):455-459. DOI 10.1093/mutage/16.6.455.

Nederhand R.J., Droog S., Kluft C., Simoons M.L., De Maat M.P.M., Investigators of the EUROPA trial. Logistics and quality control for DNA sampling in large multicenter studies. *J. Thromb. Haemost*. 2003;1(5):987-991. DOI 10.1046/j.1538-7836.2003.00216.x.

Palmirotta R., Ludovici G., De Marchis M.L., Savonarola A., Leone B., Spila A., De Angelis F., Della Morte D., Ferroni P., Guadagni F. Preanalytical procedures for DNA studies: the experience of the interinstitutional multidisciplinary BioBank (BioBIM). *Biopreserv. Biobank*. 2011;9(1):35-45. DOI 10.1089/bio.2010.0027.

Peng L., Wang S., Yin S., Li C., Li Z., Wang S., Liu Q. Autophosphorylation of H2AX in a cell-specific frozen dependent way. *Cryobiology*. 2008;57(2):175-177. DOI 10.1016/j.cryobiol.2008.06.005.

Permenter J., Ishwar A., Rounsavall A., Smith M., Faske J., Sailey C.J., Alfaro M.P. Quantitative analysis of genomic DNA degradation in whole blood under various storage conditions for molecular diagnostic testing. *Mol. Cell. Probes*. 2015;29(6):449-453. DOI 10.1016/j.mcp.2015.07.002.

Rao S.S.P., Huntle M.H., Durand N.C., Stamenova E.K., Bochkov I.D., Robinson J.T., Sanborn A.L., Machol I., Omer A.D., Lander E.S., Aiden E.L. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. 2014;159(7):1665-1680. DOI 10.1016/j.cell.2014.11.021.

Ross K.S., Haites N.E., Kelly K.F. Repeated freezing and thawing of peripheral blood and DNA in suspension: effects on DNA yield and integrity. *J. Med. Genet*. 1990;27(9):569-570. DOI 10.1136/jmg.27.9.569.

Schröder C., Steimer W. gDNA extraction yield and methylation status of blood samples are affected by long-term storage conditions. *PLoS One*. 2018;13(2):e0192414. DOI 10.1371/journal.pone.0192414.

ORCID ID

M.M. Gridina orcid.org/0000-0002-7972-5949
E. Vesna orcid.org/0000-0003-3480-3963
M.E. Minzhenkova orcid.org/0000-0001-5458-0408
N.V. Shilova orcid.org/0000-0002-0641-1084
I.N. Lebedev orcid.org/0000-0002-0482-8046
V.S. Fishman orcid.org/0000-0002-5573-3100

Acknowledgements. The study was supported by the Russian Science Foundation, project No. 22-24-00190. Samples from the biocollection "Biobank of the population of Northern Eurasia" of the Research Institute of Medical Genetics, Tomsk NRMC were used in the study.
We thank Alexandra Yan and Artem Nurislamov.

Conflict of interest. The authors declare no conflict of interest.

Received November 8, 2022. Revised December 29, 2022. Accepted December 30, 2022.