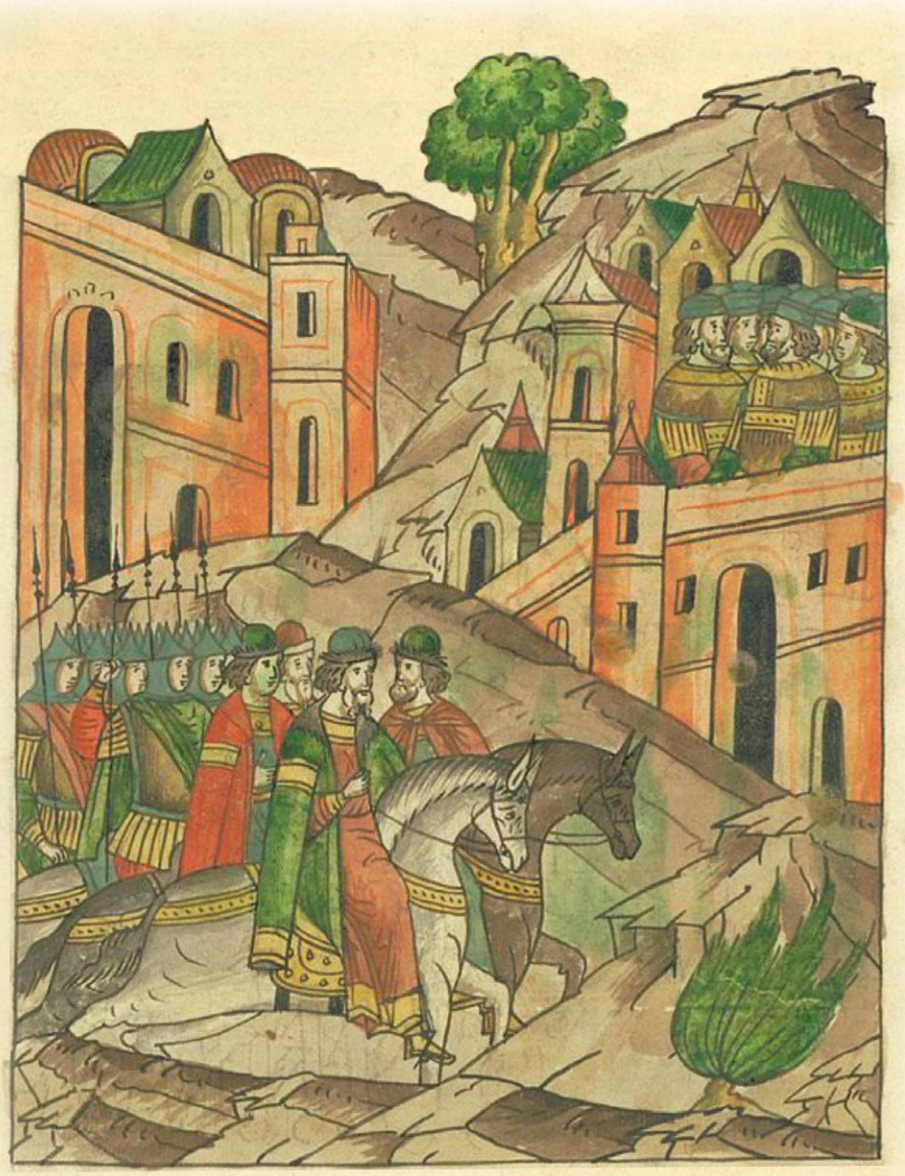
JULY–SEPTEMBER 2023 VOL. 15 № 3 (58)

ISSN 2075-8251

Acta Naturae

The Rurikids: The First Experience of Reconstructing the Genetic Portrait of the Ruling Family of Medieval Rus' Based on Paleogenomic Data



Remains of Prince Dmitry Alexandrovich from the destroyed sarcophagus of the Transfiguration Cathedral of Pereslavl

The first cervical vertebra, rear view

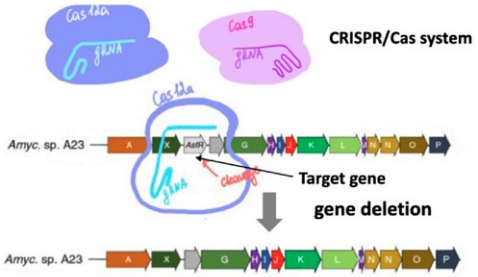
The third cervical vertebra

Fragment of the body of the thoracic vertebra

Modern Approaches to the Genome Editing of Antibiotic Biosynthetic Clusters in Actinomycetes

J. A. Buyuklyan, Yu. V. Zakalyukina, I. A. Osterman, M. V. Biryukov

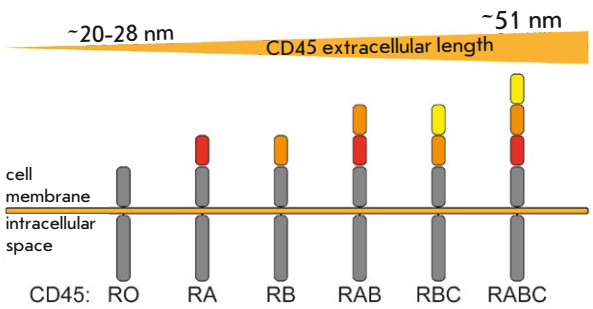
The global increase in antibiotic resistance among pathogens necessitates searching for new approaches to antibiotic design, in particular through activation of silent biosynthetic clusters in members of the phylum *Actinomycetota*. Traditional technologies are suitable for producing only a small number of secondary metabolites out of a huge variety of compounds predicted using bioinformatic methods. This paper reviews promising methods and approaches to genome editing of biosynthetic antibiotic clusters and expression of corresponding genes, which lead to synthesis of novel molecules with antibacterial activity.



Schematic representation of mutations introduced into the genome of actinomycetes using CRISPR-Casmediated genome editing

Protein Tyrosine Phosphatase CD45 As an Immunity Regulator and a Potential Effector of CAR-T therapy

D. V. Volkov, V. M. Stepanova, Y. P. Rubtsov, A. V. Stepanov, A. G. Gabibov



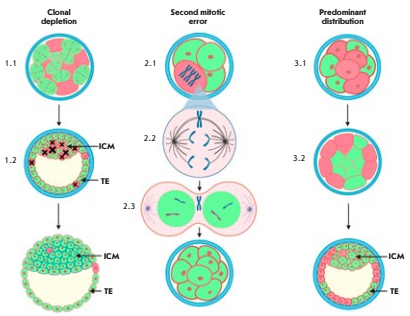
The structure and prevalence of CD45 isoforms in blood cells

The leukocyte common antigen, CD45, is a receptor tyrosine phosphatase and one of the most prevalent antigens found on the surface of blood cells. CD45 plays a crucial role in the initial stages of signal transmission from receptors of various immune cell types. Immunodeficiency, autoimmune disorders, and oncological diseases are frequently caused by gene expression disorders and imbalances in CD45 isoforms. It is of utmost importance to comprehend the structural features of CD45 and its function in regulating immune system cell activation to study oncological diseases and the impact of CD45 on lymphocytes and T cells modified by chimeric antigen receptors.

Chromosomal Aberrations As a Biological Phenomenon in Human Embryonic Development

A. D. Ivanova, M. L. Semenova

Frequent chromosomal abnormalities are a distinctive feature of early embryonic development in mammals, especially humans. Aneuploidy is considered as a contributing factor to failed embryo implantation and spontaneous abortions. In this review, the authors focus on the biological mechanisms associated with chromosomal abnormalities. In particular, the authors address the main events in oocyte meiosis that controls not only the genetic status of an unfertilized oocyte, but also further embryo viability and analyze the features of first cleavage divisions and the causes of frequent chromosomal errors in early embryonic development. In addition, the authors discuss current data on self-correction of the chromosomal status in early embryos.



Models for self-correction of the chromosomal status in mosaic embryos

ActaNaturae
JULY–SEPTEMBER 2023 VOL. 15 № 3 (58)
since april 2009, 4 times a year

Founders
Acta Naturae, Ltd,
National Research University
Higher School of Economics

Editorial Council
Editors-in-Chief: A.G. Gabibov, S.N. Kochetkov

V.V. Vlassov, P.G. Georgiev, M.P. Kirpichnikov, A.A. Makarov, A.I. Miroshnikov, V.A. Tkachuk, M.V. Ugryumov

Editorial Board
Managing Editor: V.D. Knorre

K.V. Anokhin (Moscow, Russia)
I. Bezprozvanny (Dallas, Texas, USA)
I.P. Bilenkina (Moscow, Russia)
M. Blackburn (Sheffield, England)
S.M. Deyev (Moscow, Russia)
V.M. Govorun (Moscow, Russia)
O.A. Dontsova (Moscow, Russia)
K. Drauz (Hanau-Wolfgang, Germany)
A. Friboulet (Paris, France)
M. Issagouliants (Stockholm, Sweden)
M. Lukic (Abu Dhabi, United Arab Emirates)
P. Masson (La Tronche, France)
V.O. Popov (Moscow, Russia)
I.A. Tikhonovich (Moscow, Russia)
A. Tramontano (Davis, California, USA)
V.K. Švedas (Moscow, Russia)
J.-R. Wu (Shanghai, China)
N.K. Yankovsky (Moscow, Russia)
M. Zouali (Paris, France)

Project Head: N.V. Soboleva
Editor: N.Yu. Deeva
Designer: K.K. Oparin
Art and Layout: K. Shnaider
Copy Chief: Daniel M. Medjo
Web Content Editor: O.B. Semina

Adress: 101000, Moscow, Myasnitskaya Ulitsa, 13, str. 4
Phone/Fax: +7 (495) 727 38 60
E-mail: actanaturae@gmail.com

Reprinting is by permission only.

© ACTA NATURAE, 2023

Номер подписан в печать 30 сентября 2023 г.
Тираж 15 экз. Цена свободная.
Отпечатано в типографии: НИУ ВШЭ,
г. Москва, Измайловское шоссе, 44, стр. 2

Founder and Chairman
of the Editorial Board (from 2009 to 2023)
of the journal Acta Naturae
Academician Grigoriev Anatoly Ivanovich

Indexed in PubMed, Web of Science, Scopus, and RISC

Impact Factor: 2.0 (WOS); 3.5 (Scopus)

CONTENTS

REVIEWS

RESEARCH ARTICLES

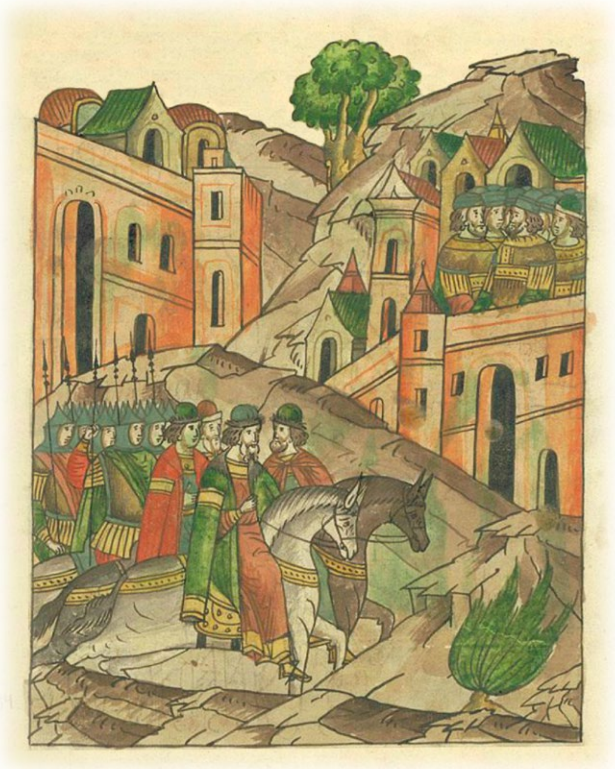


IMAGE ON THE COVER PAGE
Grand Prince Yaroslav Yaroslavich with Prince Dmitry Alexandrovich, Tovtivil Polotskiy and the Novgorodians marching under the German city of Yuriev (Tartu). ANO "Runiverse" expresses its gratitude to the "Society of Lovers of Ancient Scripts" and personally Herman L. Sterligov for providing the original illustrations from Illuminated Compiled Chronicle of Ivan IV the Terrible.
(see the article by Zhur)

Modern Approaches to the Genome Editing of Antibiotic Biosynthetic Clusters in Actinomycetes

J. A. Buyuklyan[1], Yu. V. Zakalyukina[1,2], I. A. Osterman[1,3], M. V. Biryukov[1,2*]
[1]Center for Translational Medicine, Sirius University of Science and Technology, Sochi, 354340 Russian Federation
[2]Lomonosov Moscow State University, Moscow, 119234 Russian Federation
[3]Skolkovo Institute of Science and Technology, Skolkovo, Moscow Region, 143025 Russian Federation
*E-mail: metrim@gmail.com
Received: July 13, 2023; in final form, August 19, 2023
DOI: 10.32607/actanaturae.23426
Copyright © 2023 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **Representatives of the phylum *Actinomycetota* are one of the main sources of secondary metabolites, including antibiotics of various classes. Modern studies using high-throughput sequencing techniques enable the detection of dozens of potential antibiotic biosynthetic genome clusters in many actinomycetes; however, under laboratory conditions, production of secondary metabolites amounts to less than 5% of the total coding potential of producer strains. However, many of these antibiotics have already been described. There is a continuous "rediscovery" of known antibiotics, and new molecules become almost invisible against the general background. The established approaches aimed at increasing the production of novel antibiotics include: selection of optimal cultivation conditions by modifying the composition of nutrient media; co-cultivation methods; microfluidics, and the use of various transcription factors to activate silent genes. Unfortunately, these tools are non-universal for various actinomycete strains, stochastic in nature, and therefore do not always lead to success. The use of genetic engineering technologies is much more efficient, because they allow for a directed and controlled change in the production of target metabolites. One example of such technologies is mutagenesis-based genome editing of antibiotic biosynthetic clusters. This targeted approach allows one to alter gene expression, suppressing the production of previously characterized molecules, and thereby promoting the synthesis of other unknown antibiotic variants. In addition, mutagenesis techniques can be successfully applied both to new producer strains and to the genes of known isolates to identify new compounds.**

KEYWORDS **antibiotic biosynthetic clusters, genome editing, site-directed mutagenesis, actinomycetes, antibiotics.**

ABBREVIATIONS **BGC – biosynthetic gene cluster; PCR – polymerase chain reaction; NRP – nonribosomal peptide; PKS – polyketide synthase; UDG – Uracil-DNA glycosylase; DSB – double-strand break.**

INTRODUCTION

Actinomycetota phylum members, high G–C content Gram-positive bacteria, are one of the main sources of biologically active substances [1, 2]. Modern high-throughput sequencing techniques enable the detection of dozens of biosynthetic clusters of potential antibiotics in the genomes of many actinomycetes [3]; however, the production of secondary metabolites using traditional laboratory screening techniques [4, 5], which were pioneered by Waksman in the 1940s, amounts to less than 5% of the full genetic potential of the producer strains [6, 7]. Often, these antibiotics have already been described. Some known antibiotics are frequently "rediscovered," whereas novel molecules may remain virtually invisible against the general background. The usual approaches to increase the production of novel antibiotics include the creation of optimal cultivation conditions by modifying the growth medium composition [8], co-cultivation methods [9], microfluidics methods [10], and the use of various transcription factors to activate silent genes [11, 12]. Unfortunately, these tools are non-universal for various actinomycete strains, stochastic in nature, and, therefore, they are not always success-

ful. Genetic engineering technologies are much more effective, because they provide for targeted and controllable changes in the production of target metabolites [13]. One of these technologies is mutagenesis-based genome editing of antibiotic biosynthetic clusters [14–16]. This targeted approach can alter gene expression [17] and inhibit the production of already characterized molecules, thereby facilitating the synthesis of heretofore unknown antibiotics. In addition, mutagenesis techniques can be successfully used in both new producer strains and the genes of known isolates in order to identify novel compounds.

GENOME OF ACTINOMYCETES

The genome of actinomycetes is represented by a 5- to 10-Mb circular or linear DNA molecule (*Streptomyces* spp.) with high G–C content amounting to more than 70% in some genera [18–20]. In actinomycetes—representatives of prokaryotes—implementation of genetic information, namely transcription and translation, is coupled in time and space due to the lack of internal compartmentalization of the cell [21]. The ribosome can bind to a RNA polymerase-synthesized mRNA and begin protein synthesis. In actinomycete genomes, genes encoding bioactive compounds are usually organized into biosynthetic gene clusters (BGCs) [22, 23]. BGCs are a group of two or more genes that share a common transcription start point and together encode a biosynthetic pathway for the production of a specialized metabolite. These genes contain information about the regulatory proteins that control the timing and level of expression and secretion of a particular metabolite.

There are different structural BGC classes, including non-ribosomal peptide synthetases (NRPSs), polyketide synthetases (PKSs), terpenes, and bacteriocins [24]. NRPSs and PKSs are common markers for the detection of secondary metabolites, because they synthesize structurally diverse molecules exhibiting antibiotic and immunosuppressive properties, as well as great pharmaceutical potential [25, 26]. These regions can be used to identify new antibiotic biosynthetic pathways [27] (*Fig. 1*).

According to bioinformatics data generated by the DOE Joint Genome Institute, all antibiotic producers contain dozens of potential biosynthetic clusters; i.e., they have much greater biosynthetic potential compared with that identified by routine cultivation [27].

Currently, there are various approaches to activating silent clusters [31]. They may be divided into

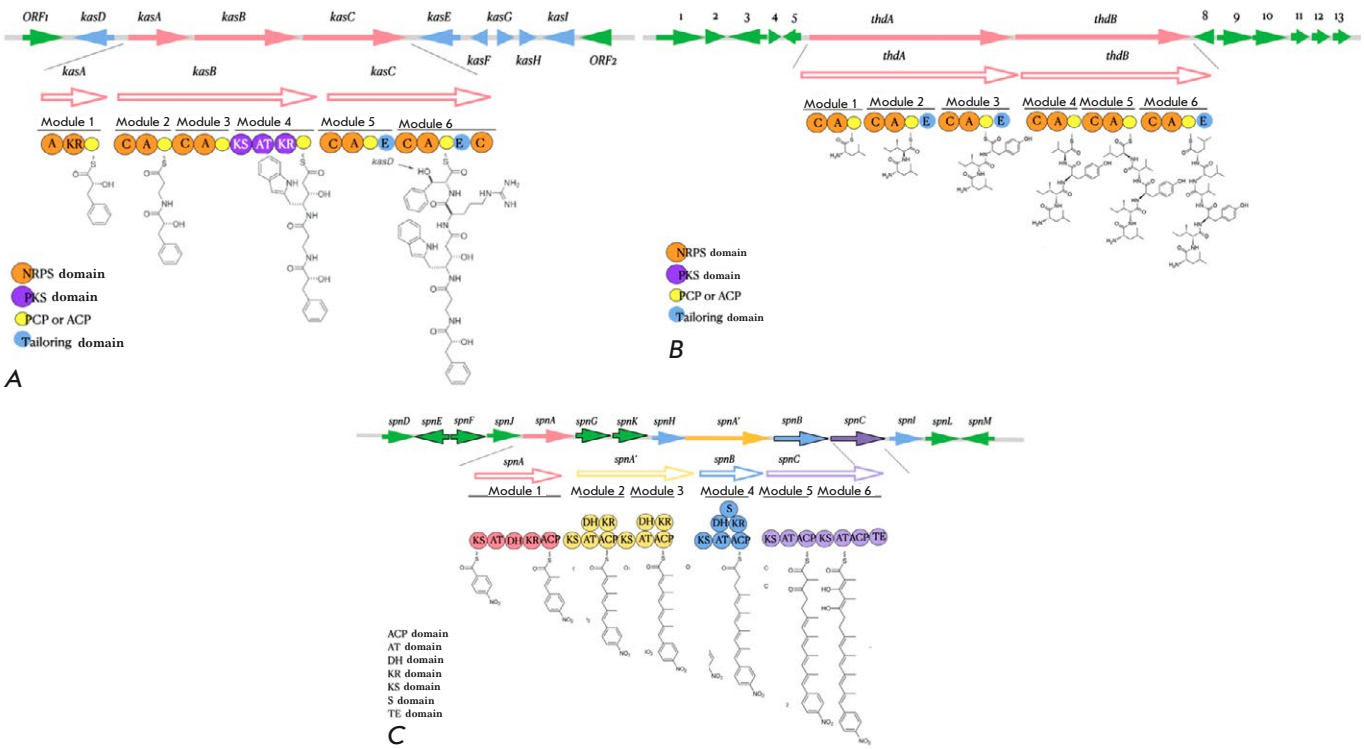


Fig. 1. Schematic representation of antibiotic biosynthesis clusters: (*A*) thermoactinoamide A from *Thermoactinomyces* sp. [28]; (*B*) kasugamycin from *Streptomyces kasugaensis* [29]; (*C*) spinosyn from *Streptomyces albus J1074* [30]

two groups: the first group is based on the heterologous expression of clusters in model *Escherichia coli* or *Saccharomyces cerevisiae* strains [32, 33], and the second uses genome editing directly in the producer strains [34, 35]. Each of these approaches has its own advantages and disadvantages. In the case of heterologous expression of clusters in *E. coli* or *S. cerevisiae* strains [36], the advantages are as follows: the simplicity associated with the transformation and expression of genes in well-studied model microorganisms, which provides a means to regulate the expression level of the antibiotic synthesis genes. This control of gene expression regulation may be implemented by means of inducible or constitutive promoters. Therefore, specific metabolites would be synthesized either in the presence of inducer molecules or permanently in a heterologous strain. In addition, model organisms, in particular *Escherichia coli*, are free of endogenous secondary metabolic pathways, which allows to obviate the influence on target cluster synthesis. Despite the positive aspects of this approach, there are a number of limiting factors: first, cluster transfer is based on homologous recombination [37], whose accuracy decreases as the number of events increases. Second, there are differences in the nucleotide sequence of the triplets encoding amino acids in different organisms. This leads to an additional step associated with the generation of a codon-optimized sequence for the synthesis of the target antibiotic molecule. These manipulations are necessary to eliminate frameshifting between native strains and hosts. In addition, some techniques require their own consensus sequences, such as attP-attachment sites that mediate site-specific recombination [38], and special plasmids, which makes the procedure more complex and labor-intensive [39].

An alternative approach to activating silent clusters is genome editing directly in the producer strains. This approach introduces mutations in the original wild-type strain and controls changes directly in it [40]. These genetic manipulations enable to study the effect of a specific mutation on other metabolic pathways not involved in the biosynthesis of a particular metabolite [41]. Of course, this technique has its own disadvantages, but there are ways to avoid them, and we will discuss them below.

APPROACHES TO GENOME EDITING IN ACTINOMYCETES

Compared to "traditional" targets for genetic modification, such as *E. coli* and *S. cerevisiae* [42], actinomycetes have a complex regulatory apparatus that prevents effective, targeted transformation of their genome [43, 44]. Nonetheless, there are approaches to introducing point mutations into the genetic apparatus of producer strains. All genome editing techniques may be divided into two categories: spontaneous [45] and site-directed mutagenesis [46].

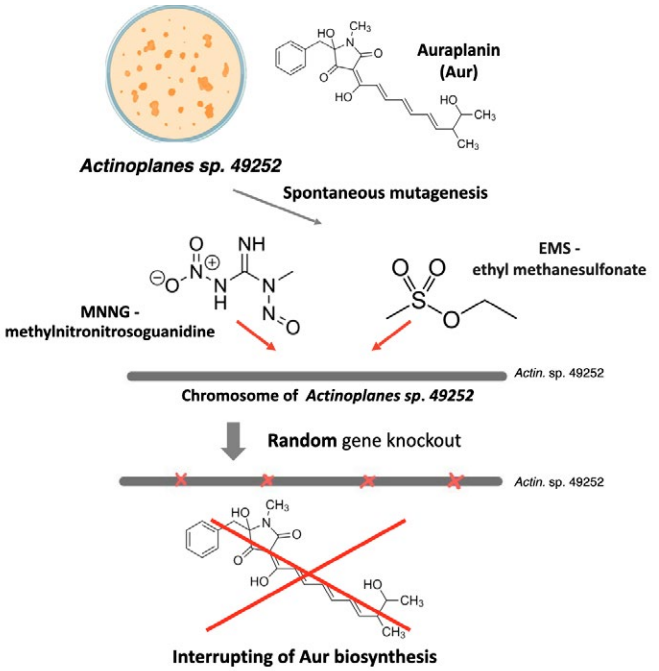


Fig. 2. Schematic representation of mutations introduced into the genome of actinomycetes using spontaneous chemical mutagenesis

Spontaneous mutagenesis

Spontaneous mutagenesis is associated with the introduction of random point mutations into DNA using a mutagen. This approach is used to solve several problems: introduction of single-nucleotide substitutions to produce new biosynthetic products [47]; an auxiliary tool for clarifying the nucleotide sequence of antibiotic biosynthetic clusters [48]. The mutagens used are methylnitronitrosoguanidine (MNNG) [49] that adds alkyl groups to the O^6 of guanine and O^4 of thymine, which leads to transition mutations between the GC and AT pairs [50], and ethyl methane sulfonate (EMS) that causes transition mutations between the GC and AT pairs [51] (*Fig. 2*). In addition to transitions of the purine and pyrimidine bases, a mutagen can also change the expression level of specific genes [52]. Because these single nucleotide substitutions are introduced randomly, alkylation/methylation of nitrogenous bases occurs in different regions of the genome. For example, modification of the promoter (regulatory region) nucleotide sequence can suppress the expression of biosynthetic gene clusters [51], and

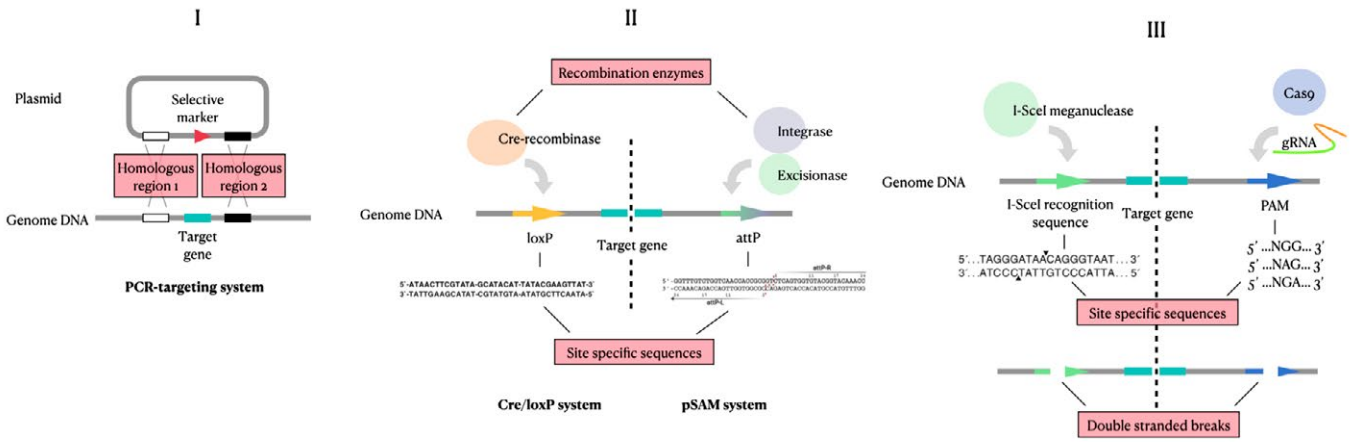


Fig. 3. Schematic representation of the molecular mechanisms underlying genetic engineering techniques for introducing mutations into the cell genome. Process I – homologous recombination; process II – site-specific recombination; process III – nuclease-induced double-strand breaks, followed by their repair

mutations in a BGC coding region can produce other genetic products and, as a consequence, new substances [48, 53].

Thus, spontaneous mutagenesis helps solve some of the problems associated with the search for new molecules, but this process is probabilistic in nature and does not guarantee reproducibility of the results; so, it cannot be used to develop a full-fledged technique for producing new antibiotics.

Site-directed mutagenesis

As mentioned above, actinomycetes implement only a small part of their biosynthetic activity and one antibiotic, such as streptothricin, can be found in every tenth isolate, while others, such as tetracycline and actinomycin D, are found at a rate of one per 100–1,000 isolates [1]. To produce novel antibiotics and their modifications using a traditional approach, such as a Waksman platform [54], it is necessary to test tens of millions of isolates. This routine approach is labor-, time-, and resources-intensive. Importantly, even known strains are a source of a huge variety of molecules with antibacterial activity [55] whose gene expression is masked by predominantly detected, known antibiotics [1].

Culp et al. proposed a concept based on the idea that disruption of the conserved biosynthetic genes of known antibiotics produced by strains may facilitate the discovery of novel metabolites whose activity has not yet been detected [56, 57]. This problem is solved using various genetic engineering tools aimed mainly at introducing deletions into the biosynthetic gene clusters of the producer strains. All these techniques may be divided into three large categories differing in their driving molecular mechanism.

Three fundamental processes are used as tools to introduce mutations: homologous recombination providing the basis for the PCR-targeting system that uses the homologous sequences required for recombination to produce deletions. The second molecular mechanism is site-specific recombination, used in the Cre-loxP recombination system and pSAM2 site-specific recombination system. The key feature is the presence of special sites: the loxP sequence for Cre recombinase and the attP sequence for the pSAM2 system. This process involves not only specific sequences, but also enzymes that perform recombination in strictly defined regions of the genome, which increases the accuracy of the process. The third process underlying site-directed mutagenesis is the introduction of double-strand breaks by nucleases, such as I-SceI meganuclease (I-SceI meganuclease-promoted recombination system) and Cas-nickase (CRISPR/Cas-based genome editing). Double-strand breaks introduced into DNA are subsequently recovered by cell repair systems (*Fig. 3*).

RECOMBINATION-BASED GENOME EDITING

PCR-Targeting System

The first-ever genome editing system, developed for *E. coli* cells, is based on homologous recombination using the λ-Red system [58]. Homologous recombination [59] is a widespread biological phenomenon that occurs in the cells of living organisms. This process is highly conserved and involves breakage and

Table 1. Selective markers for the genetic engineering of actinomycetes

No	Resistance gene	Resistance	Antibiotic	Plasmid
1.	*aac(3)IV – aminoglycoside N(3)-acetyltransferase*	Resistance to antibiotics comprising a 2-deoxy-streptamine ring	Apramycin	pCRISPomyces [63]; pStreptoBAC V [1]
2.	*aph(3)II – aminoglycoside modifying enzyme*	Resistance to aminoglycoside antibiotics	Kanamycin A and B, neomycin B and C	pCAP01 [64]; pESAC13 [65]
3.	*aadA – aminoglycoside (3") (9) adenylyltransferase*	Resistance to streptogramins and aminoglycosides	Streptomycin, spectinomycin	pIJ778 [66]
4.	*vph – phosphotransferase*	Viomycin resistance	Viomycin	pIJ780 [66]
5.	*ermE – methyltransferase - erythromycin resistance gene*	Resistance to macrolide antibiotics	Erythromycin	pBF24 [67]
6.	*hyhB – hygromycin resistance gene*	Resistance to aminoglycoside antibiotics	Hygromycin B	pBF27 N [67]

repair of double-stranded DNA (dsDNA) [60, 61]. In addition, homologous recombination is a tool for introducing point mutations into the bacterial genome [62]. This process provided the basis for developing a PCR-mediated genome editing tool that replaces the target sequence in the cell genome with an amplified fragment of the selective marker gene (*Table 1*).

For a successful homologous recombination, 2 Kb flanking sequences are required. A deletion was for the first time introduced into the geosmin biosynthetic gene cluster of *St. coelicolor A3(2)* using a PCR-mediated technique [66, 68] (*Fig. 4*).

This controlled genetic engineering enables one to generate antibiotics through combinatorial biosynthesis in the producer strain. The strain is depleted of genes of the main endogenous secondary metabolites (avermectin and filipin in *St. avermitilis*), transposon genes [69], and the IS sequences [70] that do not affect the strain growth rate but promote genome stability.

Despite successful results [66, 68], there remains limitations in PCR-mediated genome editing due to its non-universality for different actinomycete strains.

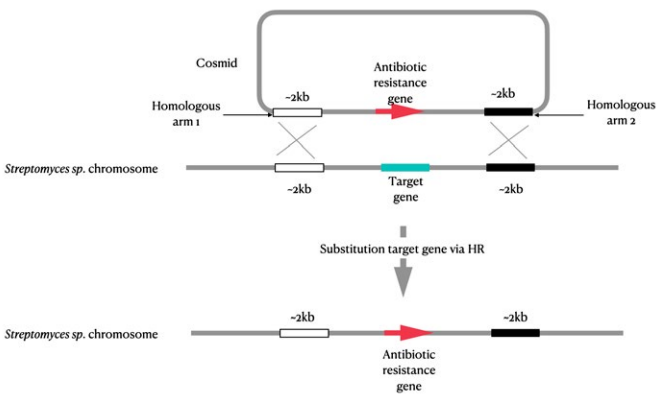


Fig. 4. Schematic representation of mutations introduced into the genome of actinomycetes using PCR-mediated genome editing

Cre-loxP recombination system

Cre-loxP recombination is used to make large deletions in the genome of bacterial cells [71, 72] using Cre recombinase [73, 74]. To introduce a mutation, two loxP (locus of crossing (x) over, P1) sequences flanking the target gene are required for site-specific Cre recombinase-mediated recombination (*Fig. 5*) [75].

The mechanism for introducing mutations involves successive recombination stages. First, two loxP sequences are introduced into the actinomycete genome in such a way as to flank the target gene. This process is mediated by two homologous recombination events [68]. Next, the Cre protein gene is expressed and the recombinase recognizes the introduced loxP sequences and performs site-specific recombination [60], leading to the deletion of the target gene. After completing the process, one of the loxP sequences is retained in the actinomycete genome.

This technique was used to produce a 1.4-Mb deletion in the geosmin biosynthetic gene cluster in the *St. avermitilis* genome [68]. This technique is more accurate than the PCR-mediated approach where recombination is controlled by the cell's internal ma-

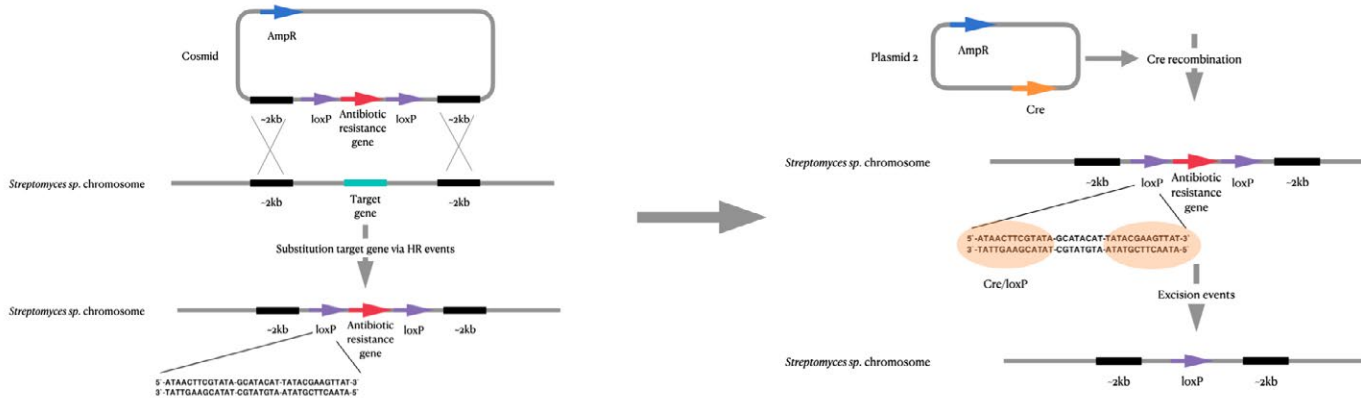


Fig. 5. Schematic representation of mutations introduced into the genome of actinomycetes using Cre/loxP-mediated genome editing

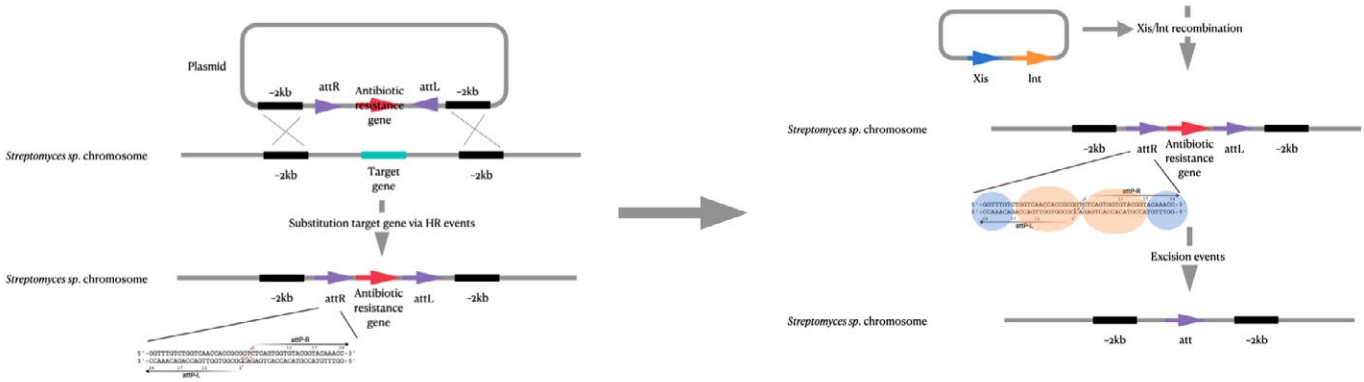


Fig. 6. Schematic representation of mutations introduced into the genome of actinomycetes using pSAM-mediated genome editing

chinery and occurs at the homologous flanking regions of the target gene [59], which is probabilistic in nature. The specificity and accuracy of the Cre/loxP approach are based on the presence of loxP flanking sequences that are specifically recognized by Cre recombinase [76]. Furthermore, the *cre* gene sequence is controlled by an inducible promoter in a separate plasmid, which provides control over the Cre recombinase expression [77]. The drawback of this system is the preservation of loxP fragments in the genome with changes in the genomic content, in addition to the target mutation–deletion.

pSAM2 site-specific recombination system

The pSAM2 system [78], like the Cre-loxP approach, is based on site-specific recombination [79]. But in this case, the specificity is associated not with recombinase activity, but with certain sequences in the genome—attachment sites attP (pSAM2 plasmid) and attB (genomic DNA of the bacterium) [80–82]. These attB sites, encoded by the non-replicative pSAM2 plasmid, are introduced into the genome of actinomycetes through homologous recombination [59]. It is noteworthy that after removal of selective pressure, the plasmid is eliminated from actinomycete cells. The Att sites introduced into the genomic sequence flank the target gene on both sides.

The introduction of a mutation using the pSAM2-based system includes the following steps: at the first step, recombination occurs at the attP/attB sites, which is accompanied by plasmid integration into the actinomycete genome. Next, the target gene is deleted by Xis excisionase. The Xis protein gene is located in a self-replicating plasmid and is eliminated from actinomycete cells when selective pressure decreases [83, 84] (*Fig. 6*).

This approach was used to delete a 90 Kb rifampicin biosynthetic cluster in *A. mediterranei DSM 40773* cells [78]. The main advantage of this technique is that mutant strains contain additional 30–40 bp inserts in the genomic sequence, which does not affect the reading frame [85].

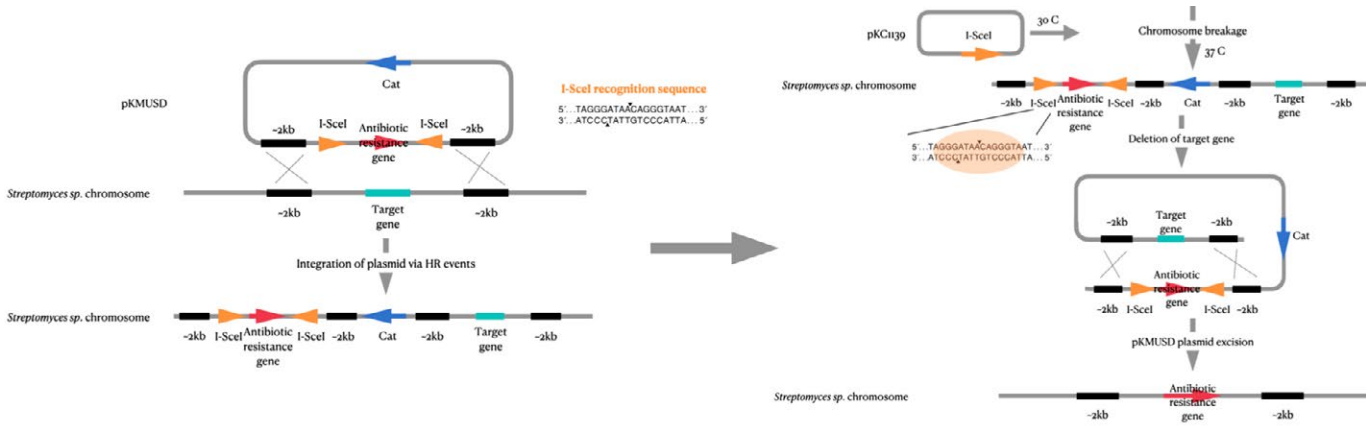


Fig. 7. Schematic representation of mutations introduced into the genome of actinomycetes using I-SceI-mediated genome editing

NUCLEASE-BASED GENOME EDITING

I-SceI meganuclease-promoted recombination system

The next genome editing technique is based on the introduction of double-strand breaks by the I-SceI meganuclease-promoted recombination system [86]. In this case, I-SceI meganuclease recognizes a unique 18 bp sequence, introduces a double-strand break, and starts the recombination process [87]. I-SceI meganuclease was first found in *S. cerevisiae* mitochondria [88].

In practice, a codon-optimized sequence of the I-SceI meganuclease gene [89, 90] and the temperature-sensitive plasmid pHZ1358 and its derivatives (pKC1139 and pJTU1278) are required to introduce deletions or substitutions into the nucleotide sequence of actinomycete strains (*Fig.* 7). In addition, insertion of 18 bp into the producer strain genome is required. This technique was used to delete the actinorhodin (*Act*) gene from *St. coelicolor A3(2)* cells [86, 91].

The process includes a series of homologous recombination events necessary to introduce an 18-nucleotide I-SceI target sequence into the actinomycete genome. These sites are encoded by the self-replicating plasmid pKMUSD. Next, the I-SceI protein gene is expressed under the control of the inducible tipA promoter [92]. Meganuclease recognizes a specific sequence in the actinomycete genome and introduces double-strand breaks that are then repaired using homologous fragments present in the cell genome [86]. It is worth noting that the I-SceI protein gene is localized in the temperature-sensitive plasmid pKC1139; therefore, after the second homologous recombination event, as the temperature rises to 36°C, the plasmid, together with the I-SceI protein gene, is eliminated [93]. This activity is controlled by the temperature-sensitive origin of pSG5 replication in the plasmid [86, 94]. This enables the introduction of deletions without additional changes in the genomic content (*Table 2*).

The major drawback of the I-SceI meganuclease-based approach is a lack of the *tipA* gene for inducing nuclease genes in some strains. In addition, this process is accompanied by double-strand DNA breaks, so errors in the repair apparatus can lead to mutations not associated with the target deletion. The positive aspects of this technique include preservation of the genomic content without additional nucleotide sequences after the completion of genetic manipulations.

CRISPR/Cas-based genome editing

Technology based on the clustered, regularly interspaced short palindromic repeats CRISPR/Cas system, in particular the CRISPR/Cas9 system, has become a promising tool for the genetic engineering of actinomycete strains [95–97].

CRISPR/Cas is a natural system for defending prokaryotic cells against foreign DNA [98–100]. This technology is widely used for genome editing in organisms from various taxonomic groups. Unlike I-SceI meganuclease-based genome editing [101], the CRISPR/Cas-based technology does not require preintegration of a unique enzyme-recognized sequence into the target genome but uses a transcribed guide RNA (sgRNA, a chimera of crRNA and tracrRNA) or crRNA alone to selectively bind Cas proteins in any genomic region [102, 103]. The Cas9/crRNA/tracrRNA complex can target any DNA sequence, known as a protospacer, provided that its 3'-end carries an appropriate trinucleotide protospacer adjacent motif (PAM) [104, 105], such as NGG (N is any nucleotide) in *Streptococcus pyogenes* [106].

Table 2. The site-directed mutagenesis techniques used in actinomycetes

No	Technique	Advantages	Disadvantages	Efficiency
1.	PCR-Targeting System	No additional tools, except PCR, are needed to introduce mutations.	Complex protocols, time-consuming procedures, universality for different actinomycete strains; deletion is accompanied by introduction of a selective marker into the genome.	Efficacy was shown only in the geosmin BGC of a model *St. coelicolor* strain.
2.	Cre-loxP Recombination System	Opportunity to delete large gene regions of about 1.4 Mb. [16]. Greater specificity due to Cre recombinase.	Time-consuming procedures, changes in genomic content, apart from the target mutation (deletion), due to preservation of loxP fragments.	A positive result was shown only for the geosmin BGC in a model *St. avermitilis* strain.
3.	pSAM2 Site-Specific Recombination System	Deletion of entire BGCs, minimal changes in the reading frame after excision of the genetic construct.	Time-consuming procedures, preservation of small sequences in the bacterial genome.	The approach is effective not only in model streptomycete strains but also in rare genera such as *Actinoplanes mediterranei*. The technique was effective in 90 Kb antibiotic BGCs (rifampicin cluster).
4.	I-SceI Meganuclease-Promoted Recombination System	Implementation of deletions without additional changes in genomic content.	Genome editing requires a codon-optimized I-SceI meganuclease gene sequence and a temperature-sensitive plasmid pKC1139.	Efficiency was shown in the actinorhodin BGC of a model *St. coelicolor* strain.
5.	Cas9-Based Genome Editing	Opportunity to introduce genomic deletions up to 30 Kb [16]; opportunity to introduce mutations into the promoter sequence.	Toxicity of the Cas9 protein due to the off-target effect; DSBs require a G-rich PAM sequence (5'-NGG-3'); Introduced DSBs cannot always be eliminated by intracellular repair systems.	Efficiency varies from 21 to 100% both in model streptomycete strains and in three members of rare genera [16]. Widely used for editing antibiotic BGCs of various lengths.
6.	Cpf1-Assisted Genome Editing	High specificity due to the need in a T-rich PAM sequence (5′-TTV-3) for introducing DSBs.	Introduced DSBs cannot always be eliminated by intracellular repair systems.	Efficiency of the system has been demonstrated in various actinomycete strains. The Gpf1 protein exhibits specificity for the T-rich PAM sequence, which reduces the off-target effect by 26%, thereby increasing the efficiency from 47 to 100% [44, 60].
7.	CRISPR-BEST (CRISPR-Base Editing SysTem)	Genome editing does not require DSBs; point mutations are introduced to create a stop codon.		The technique is applicable both to model actinomycete strains, such as *St. coelicolor*, and to members of rare genera. This technique is relatively new and has been tested in known BGCs such as actinorhodin.

The genome of streptomycetes is mainly edited with two Cas nucleases: the class 2 type II Cas9 from *Str. pyogenes* [107] and the class 2 type V Cpf1, also known as Cas12a, from *Francisella novicida* (*Fig. 8*) [103, 108, 109].

Compared with other genome editing technologies, the CRISPR/Cas system has clear advantages: high efficiency, ease of use, and rapid results [110].

Cas9-based genome editing. Based on the CIRSPR/Cas9 system, two plasmid versions have been developed for manipulating the genome of streptomycetes: pCRISPomyces-1 and pCRISPomyces-2 [63].

pCRISPomyces-1 comprises crRNA and tracrRNA gene sequences and the *cas9* gene. pCRISPomyces-2 includes the chimeric sgRNA cassette and *cas9* gene. Both plasmids use strong constitutive promoters for the expression of CRISPR/Cas elements and an optimized cas9 gene sequence for better expression in *Streptomyces* [111, 112].

Using this tool, Cobb et al. successfully achieved 20–31.4 Kb DNA deletions, including individual genes

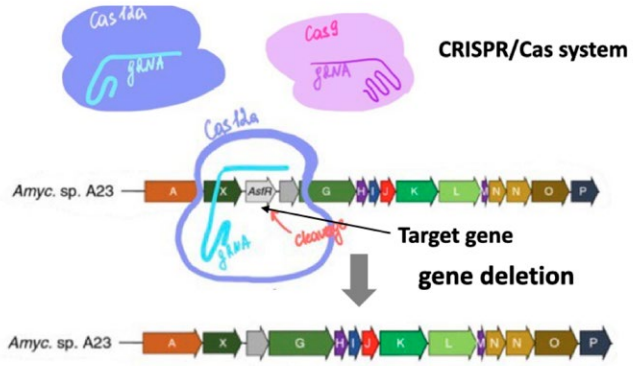


Fig. 8. Schematic representation of mutations introduced into the genome of actinomycetes using CRISPR-Cas-mediated genome editing

and clusters of antibiotic biosynthesis, with an efficiency of 21–100% in three different streptomycete species [63]. The introduction of such deletions into streptomycin and streptothricin biosynthesis clusters led to the identification of novel antibiotics in known producer strains: thiolactomycin, amicetin, phenanthroviridin, and 5-chloro-3-formylindole [113].

In addition to deletions, the CRISPR/Cas system allows for the introduction of mutations into promoters. Thus, it was possible to activate silent biosynthetic gene clusters of different classes in five *Streptomyces* strains and to identify unique metabolites, including a novel pentangular type II polyketide in *St. viridochromogenes* [114].

Despite the obvious advantages, the pCRISPomyces system has a number of disadvantages associated with the toxicity of the Cas9 protein to the bacterial cell. This is due to the cleavage of non-target DNA (off-target effect) [115, 116] and difficulty to use it in streptomycetes with a low DNA transformation efficiency [94]. Wang et al. developed a modified pWHU2653 plasmid-based CRSPR/Cas9 system with the Cas9 protein gene under the control of an inducible promoter, which provides the control over protein synthesis [117]. Also, double-strand break repair is ATP-dependent, so the *AtpD* gene encoding the β-subunit of ATP synthase was introduced into the pWHU2653 plasmid to enhance the editing efficiency (*Table 2*) [94].

Cpf1-assisted genome editing. Apart from the indicated disadvantages, the CRISPR/Cas9 system has a number of limitations. As mentioned earlier, the genome of actinomycetes has a high GC content [118], and recognition of the target sequence by the Cas9 protein requires a G-rich (PAM) sequence (5'-NGG-3')' e.g., 260 targets per 1,000 bp in *St. coelicolor* [119, 120]. The system, based on the Cas12a protein from *F. novicida,* gets around this limitation because double-strand breaks require a T-rich PAM sequence (5'-TTV-3) [121], which increases the specificity of the process [97]. Using Cpf1 nuclease, Yeo et al. achieved gene deletion in the 5-oxo-milbemycin A3/A4 producing strain *St. hygroscopicus* SIPI-KF, which could not be edited by Cas9 due to its high toxicity [120]. Thus, Cpf1 and alternative genome editing technologies complement current CRISPR/Cas-based tools and facilitate the discovery of novel biologically active substances in *Streptomyces* spp. and members of other actinomycete genera (*Table 2*) [93].

CRISPR-BEST (CRISPR-Base Editing SysTem). Nuclease-based genome editing techniques require the introduction of double-strand breaks in DNA, which may lead to genome instability due to failure of the repair systems [122]. David Liu developed an alternative CRISPR/Cas system-based technique that does not require DSBs. This technique generates point mutations leading to the emergence of a stop codon in the coding sequence [123]. The technique uses two types of deaminases: cytidine deaminase [124] converts cytosine to thymine, and adenine deaminase [125] leads to transitions, such as A–G and C–T. This difference was used to produce two genetic constructs: CRISPR-cBEST comprising a variant of the rat *APOBEC1* cytidine deaminase gene (r*APOBEC1*) and CRISPR-aBEST with adenine deaminase controlled by the inducible tipA (*thiostrepton-responsive activator*) promoter [124]: so, the key factor is the presence of the *tipA* gene in the target strain [125]. In addition, both plasmids contain the Cas9 nickase gene [126] and a codon-optimized sgRNA sequence [103]. The use of these plasmids leads to the expression of deaminase genes and transitions. Deamination of adenine in an A : T pair or cytosine in a C : G pair results in the formation of new pairs, I : T and U : G, in one DNA strand. During replication, uracil in the new U : G pair is recognized as thymine and inosine in the I : T pair is recognized as guanine; this discrepancy leads to the activation of cell repair systems [96, 127].

In the first case, uracil DNA glycosylase (UDG) is activated [128], triggering the excision repair mechanism [129, 130], or the original pairs are repaired by the mismatch repair system [129, 131, 132]: thus, the original pairs are repaired, and DSBs are not required in further replication processes.

It should be noted that this system has shown good results in model *St. coelicolor* strains and in *St. griseofuscus* (*Table 2*).

CONCLUSION

Genome mining and manipulations with the genome, in particular antibiotic gene clusters, represent an enormous potential in our efforts to identify new molecules that exhibit antibacterial activity. Importantly, the discovery of new BGCs in the genome of actinomycetes opens up broad opportunities for their editing; however, there are some limitations associated with these techniques and tools for changing the metabolic activity of strains.

Each of the described approaches can be used for specific genetic engineering tasks. For example, spontaneous mutagenesis is used as an additional tool to identify the BGC of a potential novel antibiotic. The introduction of random mutations into the genome of the producer strain may change the biosynthetic activity of a test metabolite, and further genomic analysis identifies the mutated gene region in the biosynthetic cluster. The key advantages of site-directed mutagenesis include its target specificity and efficiency: this approach is applied to known gene clusters in order to alter their expression and subsequently identify masked molecules in known isolates.

As stated earlier, most site-directed mutagenesis techniques, except CRISPR-BEST, involve recombination. Furthermore, recombination can use either internal systems of the cell, as in the PCR-mediated technique, or special enzymes: Cre-recombinase, Xis-excisionase, and Int-integrase. Undoubtedly, enzymes and specific enzyme-recognized sequences not only enhance precision, but also change the genomic content.

A number of site-directed mutagenesis techniques are based on the introduction of double-strand breaks, followed by DSB repair. These techniques include an I-SceI meganuclease-based system and CRISPR-Cas modifications (Cas9-based genome editing and Cpf1-assisted genome editing). All three approaches can be used to edit actinomycete BGCs. However, the genomic features of these bacteria impose a number of restrictions on the use of CRISPR-Cas9, given the off-target effect and toxicity of the Cas9 protein, and the restrictions on the use of I-SceI are due to the genome optimization associated with the generation of an 18-bp consensus meganuclease target sequence. The CRISPR-Cas system based on the Cas12 nuclease (Cpf1) recognizes a different T-rich PAM sequence, which reduces the risk of accidental double-strand breaks. In addition to the specific interaction between the nuclease and the target sequence, an important role is played by the internal cellular repair system associated with double-strand break repair.

Importantly, all these techniques require their own genetic constructs with the corresponding nucleotide sequences, which are used to transform streptomycete strains. A separate issue in all these approaches may be the low transformability of a particular strain. However, despite all the limitations of the described methodologies, they have allowed researchers to achieve good results–discovery of novel antibiotics and enhancement of the biosynthetic potential of actinomycetes. For example, in 2003, a PCR-mediated editing technique was used to perform manipulations with the geosmin cluster of a model *St. coelicolor* strain [36]. Seven years later, a Cre-recombinase-based technique was used to achieve a 1.4-Mb deletion in the geosmin cluster of a *St. avermitilis* strain [46]. When dealing with the site-specific approach, we should also mention the pSAM2 system. Despite the fact that this plasmid was generated back in 1989 [12], it was successfully applied in 2022 to introduce a mutation into the rifampicin cluster of *A. mediterranei DSM 40773* cells [87]. An I-SceI meganuclease-based approach was used to produce a mutation in the actinorhodin cluster of *St. coelicolor A3(2)* cells in 2014 [30]. The possibility of using all the described techniques for genome manipulations has so far been demonstrated only in actinomycetes; CRISPR-Cas was the most effective approach that not only demonstrated a good outcome associated with the introduction of mutations, but also identified novel molecules. For example, using a CRISPR-Cas9-based technique, the ability of previously studied streptomycin-producing strains to synthesize the novel antibiotics thiolactomycin, phenanthroviridine, and 5-chloro-3-formylindole was revealed in 2019 [21].

Further prospects for the use of genome-editing techniques are associated with the opportunity to identify novel antibiotic BGCs in the genomes of characterized strains and induce targeted mutagenesis. This requires combining predictive bioinformatics algorithms for identifying potential BGCs of secondary metabolites and reliable tools for targeted mutations of regulatory sites, introduction of inducible promoters, and deletion of repressor genes. Transfer of target antibiotic BGCs into strains more suitable for expression seems promising. This approach may be especially effective for large-scale biotechnological production, when the production of a target metabolite is increased using a specially designed, genetically engineered strain, which will increase the profitability of the production. ●

This study was supported by the Ministry of Science and Higher Education of the Russian Federation (agreement No. 075-10-2021-093; project [BTH-RND-2127]).

REFERENCES

1. Aminov R. // Bioc. Pharm. 2017. V. 133. P. 4–19.
2. Barka E.A., Vatsa P., Sanchez L., Gaveau-Vaillant N., Jacquard C., Klenk H.P., Clément C., van Wezel G.P. // Microbiol. Mol. Biol. Rew. 2016. V. 80. № 1. P. 1–43.
3. Wright G.D. // ACS Infec. Dis. 2015. V. 1. № 2. P. 80–84.
4. Cox G., Sieron A., King A.M., De Pascale G., Pawlowski A.C., Koteva K., Wright G.D. // Cell. Chem. Biol. 2017. V. 24. № 1. P. 98–109.
5. Blin K., Shaw S., Steinke K., Villebro R., Ziemert N., Lee S.Y., Medema M.H., Weber T. // Nucl. Acid. Res. 2019. V. 47. № W1. P. W81–W87.
6. Baltz R.H., Miao V., Wrigley S.K. // Nat. Prod. Rep. 2005. V. 22. № 6. P. 717–741.
7. Ochi K. // Microbiol. Res. 2017. V. 1. P. 189–203.
8. Tomm H.A., Ucciferri L., Ross A.C. // J. Ind. Microbiol. Biotech. 2019. V. 46. № 9–10. P. 1381–1400.
9. Matilla M.A. // Microbiol. Biotech. 2022. V. 15. № 2. P. 392–394.
10. Maglangit F., Fang Q., Kyeremeh K., Sternberg J.M., Ebel R., Deng H.A. // Mol. 2020. V. 25. № 2. P. 256.
11. Onaka H. // J. Antib. 2017. V. 70. № 8. P. 865–870.
12. Wang B., Guo F., Dong S.H., Zhao H. // Nat. Chem. Bio. 2019. V. 15. № 2. P. 111–114.
13. Zou X., Wang L., Li Z., Luo J., Wang Y., Deng Z., Du S., Chen S. // Med. Res. Rev. 2018. V. 38. № 1. P. 229–260.
14. Mitousis L., Thoma Y., Musiol-Kroll E.M. // Antib. 2020. V. 9. № 8. P. 494.
15. Rutledge P.J., Challis G.L. // Nat. Rev. Microbiol. 2015. V. 13. № 8. P. 509–523.
16. Ochi K. // J. Antib. 2017. V. 70. № 1. P. 25–40.
17. Mao D., Okada B.K., Wu Y., Xu F., Seyedsayamdost M.R. // Cur. Opin. Microbiol. 2018. V. 45. P. 156.
18. Cimmino T., Metidji S., Labas N., Le Page S., Musso D., Raoult D., Rolain J.M. // New. Microbiol. Infec. 2016. V. 12. P. 1–5.
19. Cornell C.R., Marasini D., Fakhr M.K. // Fron. Microbiol. 2018. V. 9. P. 2282.
20. Ventura M., Canchaya C., Tauch A., Chandra G., Fitzgerald G.F., Chater K.F., van Sinderen D. // Microbiol. Mol. Biol. Rev. 2007. V. 71. № 3. P. 495–548.
21. Blaha G.M., Wade J.T. // An. Rev. Gen. 2022. V. 56. P. 187–205.
22. Martin J.F. // J. Indst. Microbiol. 1992. V. 9. P. 73–90.
23. Walsh C.T., Fischbach M.A. // J. Amer. Chem. Soc. 2010. V. 132. № 8. P. 2469–2493.
24. Donadio S., Monciardini P., Sosio M. // Nat. Prod. Rep. 2007. V. 24. № 5. P. 1073–1109.
25. Schwecke T., Aparicio J.F., Molnar I., König A., Khaw L.E., Haydock S.F., Oliynyk M., Caffrey P., Cortés J., Lester J.B. // Proc. Natl. Acad. Sci. USA. 1995. V. 92. № 17. P. 7839–7843.
26. Tillett D., Dittmann E., Erhard M., von Döhren H., Börner T., Neilan B.A. // Chem. Biol. 2000. V. 7. № 10. P. 753–764.
27. Ziemert N., Jensen P.R. // Meth. Enzym. 2012. V. 517. P. 161–182.
28. Della Sala G., Mangoni A., Costantino V., Teta R. // Fron. Chem. 2020. V. 8. P. 397.
29. Zhu C., Kang Q., Bai L., Cheng L., Deng Z. // App. Micr. Biot. 2016. V. 100. P. 1811–1821.
30. An Z., Tao H., Wang Y., Xia B., Zou Y., Fu S., Fang F., Sun X., Huang R, Xia Y., et al. // Synt. Syst. Biot. 2021. V. 6. № 4. P. 292–301.
31. Seyedsayamdost M.R. // Proc. Natl. Acad. Sci. USA. 2014. V. 111. № 20. P. 7266–7271.
32. Li T., Du Y., Cui Q., Zhang J., Zhu W., Hong K., Li W. // Mar. Drug. 2013. V. 11. № 2. P. 466–488.
33. Myronovskyi M., Luzhetskyy A. // Nat. Prod. Rep. 2019. V. 36. № 9. P. 1281–1294.
34. Tocchetti A., Donadio S., Sosio M. // FEMS Microbiol. Lett. 2018. V. 365. № 9. P. fny064.
35. Zhang J.J., Yamanaka K., Tang X., Moore B.S. // Meth. Enzym. 2019. V. 621. P. 87–110.
36. Pyeon H.R., Nah H.J., Kang S.H., Choi S.S., Kim E.S. // Microbiol. Cell Fact. 2017. V. 16. № 1. P. 1–9.
37. Li X., Heyer W.D. // Cell. Res. 2008. V. 18. № 1. P. 99–113.
38. Astumian J.H., Waldman A.S., Scocca J.J. // J. Bact. 1989. V. 171. № 3. P. 1747–1750.
39. Hume B.C., D'Angelo C., Smith E.G., Stevens J.R., Burt J., Wiedenmann J. // Sci. Rep. 2015. V. 5. № 1. P. 1–8.
40. Zhu H., Sandiford S.K., van Wezel G.P. // J. Indst. Microbiol. Biotech. 2014. V. 41. № 2. P. 371–386.
41. Sprusansky O., Zhou L., Jordan S., White J., Westpheling J. // J. Bact. 2003. V. 185. № 20. P. 6147–6157.
42. Orr-Weaver T.L., Szostak J.W., Rothstein R.J. // Proc. Natl. Acad. Sci. USA. 1981. V. 78. № 10. P. 6354–6358.
43. Bernheim A., Sorek R. // Natr. Rev. Micr. 2020. V. 18. № 2. P. 113–119.
44. Sànchez J., Barbès C., Hernandez A., de los Reyes Gavilàn C.R.G., Hardisson C. // Cand. J. Microbiol. 1985. V. 31. № 10. P. 942–946.
45. Smith K.C. // Mut. Res. Rev. Gen. Toxic. 1992. V. 277. № 2. P. 139–162.
46. Riley L.A., Guss A.M. // Biotech. Biof. 2021. V. 14. P. 1–17.
47. Volff J.N., Altenbuchner J. // Mol. Microbiol. 1998. V. 27. № 2. P. 239–246.
48. Stonesifer J., Baltz R.H. // Proc. Natl. Acad. Sci. USA. 1985. V. 82. № 4. P. 1180–1183.
49. Gichner T., Velemínský J. // Mut. Res. Rev. Gen. Toxic. 1982. V. 99. № 2. P. 129–242.
50. Thiab R.S., Jasim H.M. // Iraqi J. Biotech. 2009. V. 8. № 1. P. 496–504.
51. Sega G.A.A // Mut. Res. Rev. Gen. Toxic. 1984. V. 134. № 2–3. P. 113–142.
52. Drake J.W., Baltz R.H. // Ann. Rew. Biotech. 1976. V. 45. № 1. P. 11–37.
53. Ikenaga M., Ichikawa-Ryo H., Kondo S. // J. Mol. Biol. 1975. V. 92. № 2. P. 341–356.
54. Lewis K. // Nature. 2012. V. 485. № 7399. P. 439–440.
55. De Simeis D., Serra S. // Antibiot. 2021. V. 10. № 5. P. 483.
56. Culp E.J., Yim G., Waglechner N., Wang W., Pawlowski A.C., Wright G.D. // Nat. Biotech. 2019. V. 37. № 10. P. 1149–1154.
57. Letzel A.C., Pidot S.J., Hertweck C.A. // Natr. Prod. Rep. 2013. V. 30. № 3. P. 392–428.
58. Datsenko K.A., Wanner B.L. // Proc. Natl. Acad. Sci. USA. 2000. V. 97. № 12. P. 6640–6645.
59. Petit M.A. // Trend. Micr. 2009. V. 17. № 6. P. 226–232.
60. Smith G.R. // Microbiol. Mol. Biol. Rev. 2012. V. 76. № 2. P. 217–228.
61. Szostak J.W., Orr-Weaver T.L., Rothstein R.J., Stahl F.W. // Cell. 1983. V. 33. № 1. P. 25–35.
62. Thomason L., Court D.L., Bubunenko M., Costantino N., Wilson H., Datta S., Oppenheim A. // Curr. Protoc. Mol.

Biol. 2007. V. 78. № 1.16. P. 1–1.16.
63. Cobb R.E., Wang Y., Zhao H. // ACS Synt. Biol. 2015. V. 4. № 6. P. 723–728.
64. Yamanaka K., Reynolds K.A., Kersten R.D., Ryan K.S., Gonzalez D.J., Nizet V., Dorrestein P.C., Moore B.S. // Proc. Natl. Acad. Sci. USA. 2014. V. 111. № 5. P. 1957–1962.
65. Custodio A.B., Alcantara E.P. // Asia-Pacific J. Mol. Biol. Biotech. 2019. V. 27. № 2. P. 56–63.
66. Gust B., Challis G.L., Fowler K., Kieser T., Chater K.F. // Proc. Natl. Acad. Sci. USA. 2003. V. 100. № 4. P. 1541–1546.
67. Fayed B., Ashford D.A., Hashem A.M., Amin M.A., El Gazayerly O.N., Gregory M.A., Smith M.C. // App. Envir. Microbiol. 2015. V. 81. № 24. P. 8402–8413.
68. Komatsu M., Uchiyama T., Ōmura S., Cane D.E., Ikeda H. // Proc. Natl. Acad. Sci. USA. 2010. V. 107. № 6. P. 2646–2651.
69. Babakhani S., Oloomi M. // J. Bas. Microbiol. 2018. V. 58. № 11. P. 905–917.
70. Ooka T., Ogura Y., Asadulghani M., Ohnishi M., Nakayama K., Terajima J., Watanabe H., Hayashi T. // Gen. Res. 2009. V. 19. № 10. P. 1809–1816.
71. Campo N., Daveran-Mingot M.L., Leenhouts K., Ritzenthaler P., Le Bourgeois P. // App. Envir. Microbiol. 2002. V. 68. № 5. P. 2359–2367.
72. Yoon Y.G., Cho J.H., Kim S.C. // Gen. Anals. Biom. Engin. 1998. V. 14. № 3. P. 89–95.
73. Hallet B., Sherratt D.J. // FEMS Microbiol. Rew. 1997. V. 21. № 2. P. 157–178.
74. Jayaram M., Grainge I. // Dynm. Bact. Gen. 2005. V. 2. № 2. P. 33–81.
75. Kühn R., Torres R. // Transg. Techn. Princ. Prot. 2002. V. 180. P. 175–204.
76. Tian X., Zhou B. // J. Biol. Chem. 2021. V. 296. P. 1–15.
77. Marx C.J., Lidstrom M.E. // Biotech. 2002. V. 33. № 5. P. 1062–1067.
78. Santos L.D., Caraty-Philippe L., Darbon E., Pernodet J.L. // Microorg. 2022. V. 10. № 4. P. 828.
79. Parks A.R., Peters J.E., Wells R.D. // Mol. Life. Sci. 2018. V. 1. P. 119–127.
80. Boccard F., Smokvina T., Pernodet J.L., Friedmann A., Guerineau M. // EMBO J. 1989. V. 8. № 3. P. 973–980.
81. Raynal A., Friedmann A., Tuphile K., Guerineau M., Pernodet J.L. // Microbiol. 2002. V. 148. № 1. P. 61–67.
82. Smokvina T., Mazodier P., Boccard F., Thompson C.J., Guérineau M. // Gene. 1990. V. 94. № 1. P. 53–59.
83. Combes P., Till R., Bee S., Smith M.C. // J. Bact. 2002. V. 184. № 20. P. 5746–5752.
84. Raynal A., Tuphile K., Gerbaud C., Luther T., Guérineau M., Pernodet J.L. // Mol. Microbiol. 1998. V. 28. № 2. P. 333–342.
85. August P.R., Tang L., Yoon Y.J., Ning S., Müller R., Yu T.W., Taylor M., Hoffmann D., Kim C.G., Zhang X., et al. // Chem. Biol. 1998. V. 5. № 2. P. 69–79.
86. Fernández-Martínez L.T., Bibb M.J. // Sci. Rep. 2014. V. 4. № 1. P. 1–6.
87. Ouedraogo J.P., Arentshorst M., Nikolaev I., Barends S., Ram A.F. // App. Micr. Biotech. 2015. V. 99. P. 10083–10095.
88. Monteilhet C., Perrin A., Thierry A., Colleaux L., Dujon B. // Nucl. Acid. Res. 1990. V. 18. № 6. P. 1407–1413.
89. Niu Y., Tenney K., Li H., Gimble F.S. // J. Mol. Biol. 2008. V. 382. № 1. P. 188–202.
90. Siegl T., Petzke L., Welle E., Luzhetskyy A. // App. Microbiol. Biotech. 2010. V. 87. P. 1525–1532.
91. Lu Z., Xie P., Qin Z.//Acta Biochim. Biophys. Sin. 2010. V. 42. № 10. P. 717–721.
92. Murakami T., Holt T.G., Thompson C.J. // J. Bact. 1989. V. 171. № 3. P. 1459–1466.
93. Zhao Y., Li G., Chen Y., Lu Y. // Biomol. 2020. V. 10. № 5. P. 734.
94. Muth G., Nußbaumer B., Wohlleben W., Pühler A. // Mol. Genl. Gen. MGG. 1989. V. 219. P. 341–348.
95. Li D., Zhou H., Zeng X. // Cell Biol. Toxic. 2019. V. 35. P. 403–406.
96. Li L., Wei K., Zheng G., Liu X., Chen S., Jiang W., Lu Y. // App. Envir. Microbiol. 2018. V. 84. № 18. P. 18.
97. Tao W., Yang A., Deng Z., Sun Y. // Front. Microbiol. 2018. V. 9. P. 1660.
98. Barrangou R., Marraffini L. // Cell. 2014. V. 54. P. 234–244.
99. Barrangou R., Brouns S.J., Charpentier E., Haft D.H., Horvath P., Moineau S., Mojica F.J., Terns R.M., Terns M.P., White M.F., et al. // Nat. Rev. Microbiol. 2015. V. 13. № 11. P. 722–736.
100. Yosef I., Goren M.G., Qimron U. // Nucl. Acid. Res. 2012. V. 40. № 12. P. 5569–5576.
101. Ouedraogo J.P., Arentshorst M., Nikolaev I., Barends S., Ram A.F. // App. Microbiol. Biotech. 2015. V. 99. P. 10083–10095.
102. Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E.A. // Sci. 2012. V. 337. № 6096. P. 816–821.
103. Knott G.J., Doudna J.A. // Sci. 2018. V. 361. № 6405. P. 866–869.
104. Gleditzsch D., Pausch P., Müller-Esparza H., Özcan A., Guo X., Bange G., Randau L. // RNA Biol. 2019. V. 16. № 4. P. 504–517.
105. Karvelis T., Gasiunas G., Young J., Bigelyte G., Silanskas A., Cigan M., Siksnys V. // Gen. Biol. 2015. V. 16. P. 1–13.
106. Yñigez-Gutierrez A.E., Bachmann B.O. // J. Med. Chem. 2019. V. 62. № 18. P. 8412–8428.
107. Wei Y., Terns R.M., Terns M.P. // Gen. Devel. 2015. V. 29. № 4. P. 356–361.
108. Fonfara I., Richter H., Bratovič M., Le Rhun A., Charpentier E. // Nature. 2016. V. 532. № 7600. P. 517–521.
109. Wiedenheft B., Sternberg S.H., Doudna J.A. // Nature. 2012. V. 482. № 7385. P. 331–338.
110. Kormanec J., Rezuchova B., Homerova D., Csolleiova D., Sevcikova B., Novakova R., Feckova L. // App. Microbiol. Biotech. 2019. V. 103. P. 5463–5482.
111. Leskiw B.K., Bibb M.J., Chater K.F. // Mol. Microbiol. 1991. V. 5. № 12. P. 2861–2867.
112. Shao Z., Rao G., Li C., Abil Z., Luo Y., Zhao H. // ACS Synt. Biol. 2013. V. 2. № 11. P. 662–669.
113. Culp E.J., Yim G., Waglechner N., Wang W., Pawlowski A.C., Wright G.D. // Nature Biotech. 2019. V. 37. № 10. P. 1149–1154.
114. Zhang M.M., Wong F.T., Wang Y., Luo S., Lim Y.H., Heng E., Ang E.L., Zhao H. // Nature Chem. Biol. 2017. V. 13. № 6. P. 607–609.
115. Fu Y., Foden J.A., Khayter C., Maeder M.L., Reyon D., Joung J.K., Sander J.D. // Nature Biotech. 2013. V. 31. № 9. P. 822–826.
116. Hsu P.D., Scott D.A., Weinstein J.A., Ran F.A., Konermann S., Agarwala V., Li Y., Fine E.J., Wu X., Shalem O., et al. // Nature Biotech. 2013. V. 31. № 9. P. 827–832.
117. Wang K., Zhao Q.W., Liu Y.F., Sun C.F., Chen X.A., Burchmore R., Burgess K., Li Y.Q., Mao X.M. // Front. Bioengin. Biotech. 2019. V. 7. P. 304.

118. Cimmino T., Metidji S., Labas N., Le Page S., Musso D., Raoult D., Rolain J.M. // New Microb. New Infect. 2016. V. 12. P. 1–5.
119. Karvelis T., Gasiunas G., Young J., Bigelyte G., Silanskas A., Cigan M., Siksnys V. // Gen. Biol. 2015. V. 16. P. 1–13.
120. Yeo W.L., Heng E., Tan L.L., Lim Y.W., Lim Y.H., Hoon S., Zhao H., Zhang M.M., Wong F.T. // Biotech. Bioengin. 2019. V. 116. № 9. P. 2330–2338.
121. Chen P., Zhou J., Wan Y., Liu H., Li Y., Liu Z., Wang H., Lei J., Zhao K., Zhang Y., Wang Y., Zhang X., Yin L. // Gen. Biol. 2020. V. 21. № 1. P. 1–13.
122. Hoff G., Bertrand C., Piotrowski E., Thibessard A., Leblond P. // Sci. Rep. 2018. V. 8. № 1. P. 1–11.
123. Tong Y., Whitford C.M., Robertsen H.L., Blin K., Jørgensen T.S., Klitgaard A.K. // Proc. Natl. Acad. Sci. USA. 2019. V. 116. P. 20366–20375.
124. Holmes D.J., Caso J.L., Thompson C.J. // EMBO J. 1993. V. 12. № 8. P. 3183–3191.
125. Murakami T., Holt T.G., Thompson C.J. // J. Bact. 1989. V. 171. № 3. P. 1459–1466.
126. Gaded V., Anand R. // RSC Advan. 2018. V. 8. № 42. P. 23567–23577.
127. Sklenak S., Yao L., Cukier R.I., Yan H. // J. Amer. Chem. Soc. 2004. V. 126. № 45. P. 14879–14889.
128. Venkatesh J., Kumar P., Krishna P.S.M., Manjunath R., Varshney U. // J. Biol. Chem. 2003. V. 278. № 27. P. 24350–24358.
129. Morita R., Nakane S., Shimada A., Inoue M., Iino H., Wakamatsu T., Fukui K., Nakagawa N., Kuramitsu S. // J. Nucl. Acid. 2010. V. 2010. P. 1–32.
130. Wozniak K.J., Simmons L.A. // Nat. Rev. Microbiol. 2022. V. 20. № 8. P. 465–477.
131. Schaaper R.M. // J. Biol. Chem. 1993. V. 268. № 32. P. 23762–23765.
132. Schormann N., Ricciardi R., Chattopadhyay D. // Prot. Sci. 2014. V. 23. № 12. P. 1667–1685.

Protein Tyrosine Phosphatase CD45 As an Immunity Regulator and a Potential Effector of CAR-T therapy

D. V. Volkov, V. M. Stepanova, Y. P. Rubtsov, A. V. Stepanov, A. G. Gabibov*
M.M. Shemyakin and Yu.A. Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Moscow, 117997 Russian Federation
*E-mail: gabibov@gmail.com
Received: August 10, 2023; in final form, September 12, 2023
DOI: 10.32607/actanaturae.25438
Copyright © 2023 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **The leukocyte common antigen CD45 is a receptor tyrosine phosphatase and one of the most prevalent antigens found on the surface of blood cells. CD45 plays a crucial role in the initial stages of signal transmission from receptors of various immune cell types. Immunodeficiency, autoimmune disorders, and oncological diseases are frequently caused by gene expression disorders and imbalances in CD45 isoforms. Despite extensive research into the structure and functions of CD45, the molecular mechanisms behind its role in transmitting signals from T-cell receptors and chimeric antigen receptors remain not fully understood. It is of utmost importance to comprehend the structural features of CD45 and its function in regulating immune system cell activation to study oncological diseases and the impact of CD45 on lymphocytes and T cells modified by chimeric antigen receptors.**
KEYWORDS **CD45, T lymphocytes, B lymphocytes, T cell receptor, cancer, chimeric antigen receptor.**

INTRODUCTION

Human protein tyrosine phosphatase (PTP) CD45 is encoded by the *PTPRC* gene containing 35 exons that have been characterized, and three of those (exons 4–6) [1, 2] contain homologous enhancers and silencers of alternative pre-mRNA splicing. Despite the theoretically expected significant versatility of possible variants, only six CD45 isoforms have been identified in humans: RO (exons 3, 7, and 8), RA (exons 3, 4, 7, and 8), RB (exons 3, 5, 7, and 8), RAB (exons 3, 4, 5, 7, and 8), RBC (exons 3, 5–8), and RABC (exons 3–8) (*Fig. 1A*). CD45 isoforms are found on all cells of hematopoietic lineage (except for anucleate erythrocytes and platelets); the level of CD45 correlates with the degree of cell differentiation [3, 4] (*Fig. 1B*).

The extracellular part of CD45 consists of five structural regions. The N-terminal region is long and heavily glycosylated. This very region is responsible for what receptor isoform it is. The remaining regions of the extracellular CD45 domain common to all the isoforms are the three fibronectin type III domains and the region carrying five conserved cysteine residues. Importantly, CD45 isoform regulates the sensitivity of T cells to activation upon antigen recognition. It is assumed that due to its bulky shape and structural rigidity, CD45 is expelled from the central region as the immune synapsis (IS) is formed and as the membranes of antigen-presenting cells (APCs) and T cells approach each other [5]. The number and type of CD45 isoforms in T cells vary depending on their differentiation degree: large CD45 isoforms are predominantly found in naïve and dormant cells. In turn, activated T cells synthesize CD45 isoforms in which either most or all the domains encoded by variable exons are absent [2]. The CD45 glycoprotein contains a single transmembrane domain (*Fig. 1C*) and three intracellular ones: the wedge domain D1, and D2. Frederick and colleagues demonstrated that only the proximal domain D1 exhibits phosphatase activity (previously, it was found that D2 is required for phosphatase functioning) [6, 7].

CD45 FUNCTIONS IN IMMUNE CELLS

The role and functions of CD45 in T cells

The involvement of CD45 in the activation of immune cells was first demonstrated for the T-cell receptor (TCR) signaling pathway. An analysis of T cells lacking CD45 expression showed that this phosphatase is essential at the initial stage of signal transduction from TCR [8] (*Fig. 2*). In a non-activated T cell, CD45 dephosphorylates protein tyrosine kinase (PTK) Lck

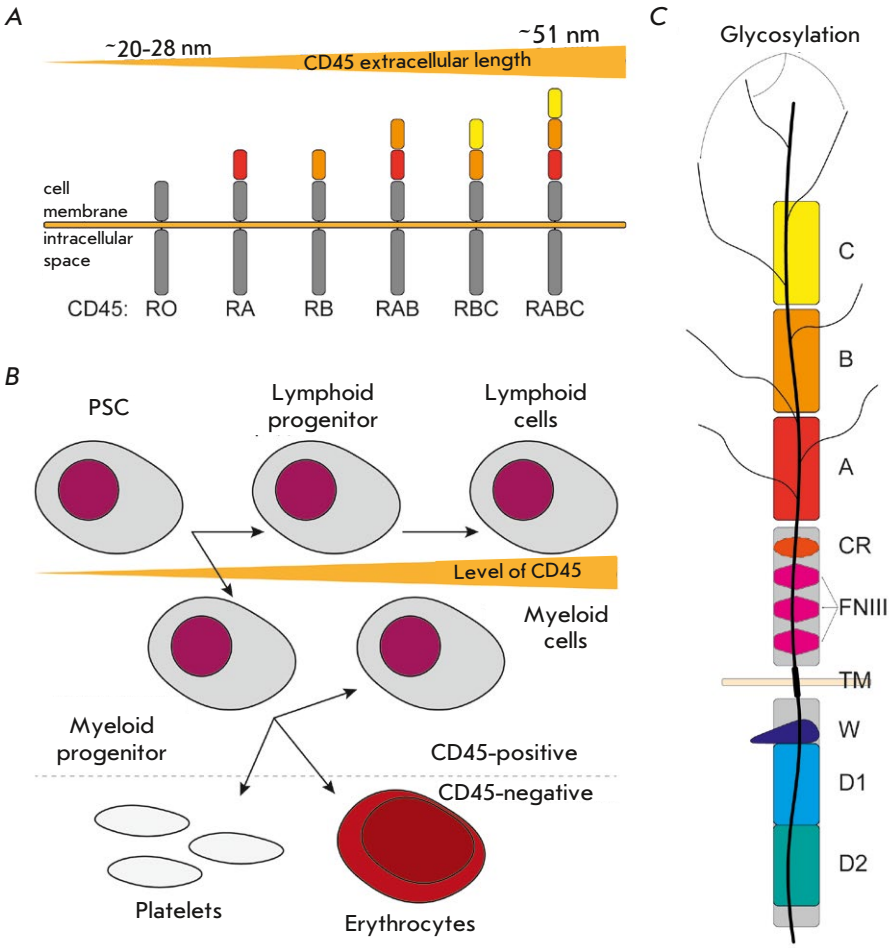


Fig. 1. The structure and prevalence of CD45 isoforms in blood cells. (*A*) Six main CD45 isoforms found in humans differ in the composition of their extracellular part owing to the alternative splicing of pre-mRNA of the *PTPRC* gene; (*B*) CD45 resides on the membrane of all cells of hematopoietic origin except for platelets and erythrocytes. The amount of CD45 increases during cell differentiation; (*C*) the structure of CD45RABC. PSC – pluripotent stem cell; A, B, C – the extracellular domains of CD45 responsible for a particular isoform; CR – cysteine-rich region; FNIII – fibronectin type III domains; TM – transmembrane domain; W – wedge domain; D1 – domain with phosphatase activity; D2 – domain required for CD45 to function in the cell

and the CD3ζ subunit of the CD3/TCR receptor. Lck is the main substrate of CD45 phosphatase, which is capable of dephosphorylating both the inhibitory tyrosine residue (Y505) at the C-terminus of kinase and activating the tyrosine residue (Y394) [9, 10]. During dephosphorylation of the inhibitory tyrosine residue Y505, CD45 competes with tyrosine kinase Csk, which inhibits Lck [11]. Courtney and colleagues studied the dual function of phosphatase and inferred that CD45 regulates the intensity and frequency of the signal transduced via TCR by exerting an impact on different substrates. By varying the activity of CD45, they revealed that the phosphatase maintains a significant amount of Lck active but prevents the activation of CD3ζ. A detailed study of the dynamics of immune synapse formation showed that before activation, the TCR complex has a non-active conformation and does not interact with the major histocompatibility complex (MHC) class I or class II. Meanwhile, CD45 inhibits the recruitment of kinase Csk [12] and dephosphorylates CD3ζ and Lck (*Fig. 2A*). During cell interaction, CD45 and Lck molecules are first recruited to the central supramolecular activation cluster (cSMAC) via TCR. However, during IS formation, CD45 is expelled to the distal supramolecular activation cluster (dSMAC) [13–17] (*Fig. 2B*). The expulsion of CD45 from the IS is seemingly related to molecule size and the high level of glycosylation and sialation [18–20] (size reduction of the CD45 ectodomain increases colocalization of phosphatase and TCR, as well as reduces the activity of TCR [5, 20–23]). Furthermore, CD45 needs to be removed from the IS center to shift the equilibrium in the central part of the synapse toward kinases. Changes in the balance enable the phosphorylation of CD3ζ, thus ensuring transduction of the signal for TCR activation. For the activation cycle to be completed, Csk molecules and the CD45RO isoform need to begin accumulating within the immune synapse [2]; the CD45RO isoform gradually penetrates into the central portion of the immune synapse and shifts the kinase–phosphatase equilibrium toward phosphatases, as well as dephosphorylates CD3ζ and

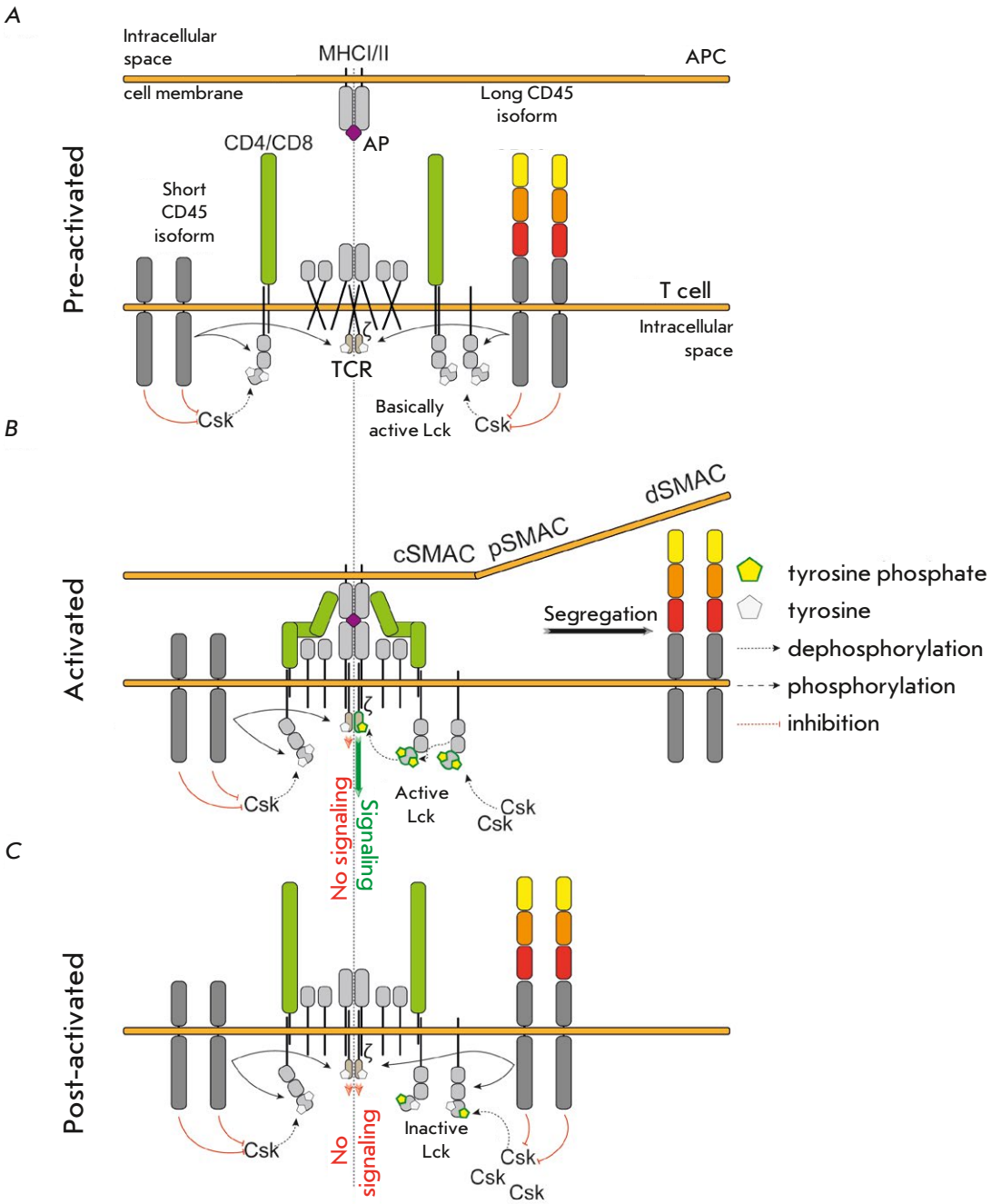


Fig. 2. The role of CD45 in transmitting T-cell receptor activation signal. T cell activation stages are shown: the pre-active (*A*), active (*B*), and post-active (*C*) ones. During the activation cycle, the composition and phosphorylation of IS participants changes: first (*A*), Lck kinase exists in the state of basal activity due to dephosphorylation by CD45 phosphatase prevailing over phosphorylation by Csk kinase; in the active state (*B*), CD45 is "segregated" in dSMAC due to its rigid and bulky structure (long isoforms) and Lck becomes active due to autophosphorylation and prevalence over Csk and phosphorylates CD3ζ, which enables further signaling; short isoforms of CD45 gradually penetrating cSMAC are synthesized at this time, and Csk accumulates there, leading to a transition of Lck to its inactive form and the end of signaling (*C*). MHCI/II I – major histocompatibility complex class I or II; APC – antigen-presenting cell; PA – presenting antigen; ζ – CD3ζ; TCR – T-cell receptor; Lck, Csk – protein kinases; cSMAC, pSMAC, dSMAC – the central, peripheral, and distal supramolecular activation clusters, respectively

activates the tyrosine residue in Lck (*Fig. 2C*). The signal intensity drops, and the composition of CD45 isoforms changes (toward increasing length and volume), while Lck returns to its basal activity state.

Presumably, through this mechanism for TCR regulation, CD45 impedes spontaneous T-cell activation, thus preventing hyperactivation and its negative sequelae [24] induced by low-affinity antigens or in the absence of an antigen. Zikherman and colleagues changed the levels of CD45 and Csk expression and showed that the balance between these molecules plays a crucial role in T-cell development. During the maturation of T cells in the thymus, the basal and inducible TCR signaling are regulated at the positive and negative selection stages. CD45 plays a positive and negative role simultaneously in antigen recognition. Variation of the Csk level regulates only the basal signal transduction. Meanwhile, an identical reduction in the Csk and CD45 levels leads to opposite changes in basal signaling by the same value. Therefore, a fluctuating CD45 level is needed for the following two processes to properly unfold: regulation of inducible signaling during positive and negative selection and compensation for Csk upon maintenance of the basal activity of T cells [25, 26]. In CD45-deficient mice with deleted exons 6 [27], 9 [28], or 12 [29], CD45 was shown to play a pivotal role in the transition of double negative (DN, $CD4^-CD8^-$) thymocytes to double positive (DP, $CD4^+CD8^+$) ones [30]. The initiation and gradual changes in the synthesis of CD4 and CD8 coreceptors during thymocyte differentiation depend on signals generated by pre-TCR and TCR. The involvement of CD45 is needed for these phenotypic and functional changes to occur. Deficiency of Lck, an important participant in TCR signaling controlled by CD45, manifested itself in a way similar to the lack of CD45 [31]: the transition of double-negative T cells to double positive ones in mice was disturbed [30], while the number of mature peripheral T cells was $\leqslant$ 5–10% of that for wild-type animals [27, 28].

The role and functions of CD45 in B cells

In B cells, CD45 also plays a crucial role in the modulation of the signal transduced via the B-cell receptor and is required to ensure normal B-cell development and an adequate response to an antigen. CD45-deficient mice were found to have defects in B-cell maturation [27, 28, 32]. Interestingly, the number of peripheral B cells does not decrease but their phenotype changes noticeably. The population of mature B cells in the spleen, as well as the population carrying the CD23 marker and MHC class II molecules, significantly decreases (IgD^{high} IgM^{low}) [32]. In mice carrying a mutation within CD45 exon 9, the number of immature peripheral B cells is increased (IgM^{high}) [28]. Importantly, CD45-deficient B cells did not proliferate in response to the stimulation of B-cell receptors (with polyclonal anti-IgD/anti-IgM antibodies); however, when other activation pathways were stimulated (by lipopolysaccharide (LPS), interleukin 4 (IL-4) and monoclonal anti-CD4 antibody), the proliferation of CD45-negative B cells was identical to that in control cells.

Cyster et al. [33] found that in response to antigen stimulation, naïve B cells isolated from CD45-deficient mice employed the ERK/RSK/EGR1 kinase pathways to a lesser extent and exhibited a low level of intracellular calcium mobilization. A difference in the markers of CD86 and CD54 activation was also observed: their level was lower in CD45-negative B cells compared to that in the CD45-positive ones. However, it was demonstrated by stimulating B cells with phorbol ester in combination with ionomycin that CD45-negative and CD45-positive B cells are activated identically.

It is essential to mention the role of CD45 in germinal centers and upon autoimmune diseases. It has been demonstrated that high-affinity autoreactive B cells fail selection in the bone marrow of native and CD45-deficient mice. However, the loss of CD45 expression allowed low-affinity autoreactive B cells to pass positive selection. Because of the lack of CD45, these B cells during selection did not induce the ERK/RSK/EGR1 pathway; intracellular calcium mobilization in response to the antigen in them was significantly lower compared to that in high-affinity cells, which provided protection to autoreactive cells against elimination. Therefore, CD45 regulates BCR and TCR activation in different ways. Unlike for TCR, the higher level of CD45 favorably affects BCR signaling, as well as it enhances the activation of the ERK/RSK/EGR1 and PI3K/AKT/mTOR kinase pathways and intracellular calcium mobilization [34]. In the case of increased CD45 expression, the Src family kinases (SFKs) involved in the BCR pathway are dephosphorylated at inhibitory tyrosine (Y507) more actively, whereas the level of phosphorylated activating tyrosine (Y416) in them remains unchanged [35]. Decreased CD45 expression has no effect on the Ca^{2+} level, since B cells contain CD148 phosphatase, which partially duplicates the functions of CD45 [35].

The role and functions of CD45 in macrophages

Leukocyte adhesion to the extracellular matrix and other cells is regulated by proteins belonging to the family of integrins [36, 37]. The targets of CD45 phosphatase activity, SFKs, are involved in the regulation

of integrin-dependent phagocytosis, as well as macrophage differentiation and activation caused by adhesion [38–40]. Roach et al. have demonstrated that in the absence of CD45, the regulation of integrin-dependent adhesion is disturbed, while the activity of PTKs Hck and Lyn (SFKs exhibiting activity in myeloid cells) is increased [41]. The CD45-mediated regulation of Hck and Lyn kinases in macrophages differs from the regulation of SFKs in T and B cells, where CD45 activity is needed to a greater extent for the dephosphorylation of the inhibitory tyrosine residues at the C-termini of Lck and Fyn and the enhancement of their activity [42–44]. The simultaneously increased phosphorylation of C-terminal tyrosine residues and SFK activity in CD45-negative macrophages indicate that the phosphatase inhibits SFK. This possibly occurs due to the dephosphorylation of autocatalytic tyrosine.

The role and function of CD45 in neutrophils

Experiments on mice deficient in kinases Hck, Fgr, and Lyn [45] have demonstrated that a loss of SFK reduces neutrophil adhesion and the level of post-translational modification of proteins. In those experiments, the Rab27a-dependent mobilization of neutrophil elastase and vesicles containing integrins α3β1 and α6β1 was out of balance. This also led to a disruption of the neutrophil migration through the vascular basement membrane and extravasation upon inflammation [45]. Hck and Fgr are involved in chemoattractant-dependent oxidative stress and F-actin polymerization [46]. SFKs also have something to do with the regulation of the mRNA transcription of many important cytokines and chemokines, which are synthesized by neutrophils either constitutively or upon stimulation (with interleukins-1, -6, -8, -10, -12; tumor necrosis factor-α (TNF-α); granulocyte–macrophage colony-stimulating factor, etc.) [47, 48]. Liles et al. [49] showed that CD45 activation by antibodies increases the level of oxidative stress induced by neutrophil activators. In turn, Harvath et al. [50] demonstrated that CD45 interacts with the molecules that are coupled to receptors for leukotriene B4 and the complement component C5a and that it regulates neutrophil chemotaxis in response to stimulation with the respective ligands. Gao et al. [51] revealed that colocalization of CD45 receptors and FcγRIIa reduces the antibody-dependent cytotoxicity of neutrophils, while simultaneously increasing IL-6 production upon FcγRIIa-mediated activation. Zhu et al. [52] showed that CD45 in neutrophils increases signaling of G-protein-coupled receptors and enhances Ca^{2+} mobilization and the activity of the PI3K and ERK kinases.

The role and function of CD45 in dendritic cells

Dendritic cells (DCs) play a crucial role in sustaining the relationship between the innate and adaptive immunity. CD45 is involved in the formation of these functional distinctions, since its specific isoforms (CD45RB) mark different DC populations. Foreign molecules bind to pattern recognition receptors (TLRs being among them), thus initiating the program of dendritic cell maturation. This process determines the further DC-mediated activation of naïve T cells together with their response to the presented antigen [53]. Early stages of TLR signaling pathways are believed to regulate SFKs [54]. The signal from TLRs increases translation of the co-stimulatory molecules that are required for activating naïve T cells and secretion of proinflammatory cytokines such as IL-12, IL-6, and TNF-α, which affect the type of generated effector T cells [55]. TLR is one of the key components of DC activation. Although the main elements of the signaling pathways of these receptors are known, the contribution of SFK has not been thoroughly described. A comparative analysis of dendritic cells isolated from native and CD45-deficient mice [56] showed that CD45 is not required for dendritic cell development but affects their maturation induced by TLR agonists. CD45 has an impact on the phosphorylation of Lyn, Hck, and Fyn and reduces LPS-induced activation of Lyn. CD45 had a favorable effect on TLR4-induced secretion of proinflammatory cytokines and interferon-β (IFN-β). Moreover, CD45 exhibited different effects on TLR activation: negative (TLR2 and TLR9 or MyD88-dependent cytokine production) and positive (TLR3 and TLR4 or MyD88-independent IFN-β secretion) [56].

The role and function of CD45 in NK cells

Many of the known NK cell receptors (CD16, NK1.1, NKG2D, NKp44, etc.) are associated with the intracellular proteins FcεRIγ, DAP10 or DAP12, which contain immunoreceptor tyrosine-based activation motifs (ITAMs) [57]. CD45 regulates the activity of SFKs, which phosphorylate tyrosine residues in ITAM and trigger the activation signaling pathways [58]. Hesslein and colleagues demonstrated that during the stimulation of the Ly49H and NKG2D receptors, CD45-negative NK cells are only 20% less toxic than control CD45-positive NK cells, while no differences are observed during the stimulation of the CD16 receptor. However, cytokine and chemokine secretion was much lower in CD45-negative NK cells in [59]. Therefore, CD45 can have different effects on the same activation signaling pathway. Signal intensity and/or duration are responsible for CD45 recruitment. Cytotoxic granules are released within several minutes after

receptor activation, in direct proximity from many components of the cellular signaling pathways involved in activation. In turn, cytokine secretion [60] is a longer process that involves the transduction of the transcriptional activation signal, mRNA synthesis and maturation, as well as translation and secretion. Therefore, persistent signaling is needed for cytokine release, whereas short-term stimulation is sufficient for eliciting a cytotoxic effect in DCs [61].

The role of CD45 in cancer

In hematopoietic cancers, CD45 expression depends on the cancer type. Thus, Feuillard and colleagues found that in patients with chronic lymphocytic leukemia (CLL), atypical tumor cells and the low level of CD45 on their surface are positive markers of patient survival [62]. Loss of CD45 was detected in patients with Hodgkin lymphoma [63] and childhood acute lymphoblastic leukemia (ALL) [64]. The higher CD45 expression in patients with ALL is associated with a higher risk of tumor recurrence [65]. It still remains unclear how CD45 is involved in the pathogenesis of multiple myeloma (MM) [66]. Patients with MM simultaneously had both CD45-positive and CD45-negative tumor cells [67]. The increased CD45 expression enhances the sensitivity of MM cells to 17-dimethylaminoethylamino-17-demethoxygeldanamycin, an inhibitor of HSP90 chaperone [67], and various apoptotic stimuli (e.g., oxidative stress and endoplasmic reticulum stress) [68]. Regardless of whether MM cells contained CD45, stimulation with IL-6 was able to unlock the JAK/STAT signaling pathways; however, only CD45-positive cells can proliferate after activation [69]. The overall survival chances of patients with the predominance of CD45-positive MM cells was lower compared to that for patients with a predominance of CD45-negative cells [70]. On the other hand, the role of CD45 has been characterized much better in patients with diffuse large B cell lymphoma (DLBCL). Phosphatase CD45 acts as a galectin-3 receptor [71]; transcription of the gene encoding it is upregulated in DLBCL cells [72]. Galectin-3 exhibits antiapoptotic activity [73]. When binding to CD45, galectin-3 remains anchored to the cell membrane. Its removal was shown to increase the number of apoptotic tumor cells [71].

The immunosuppressive tumor microenvironment (TME) is a determinant factor of the resistance of solid tumors to immunity. Several layers comprising various types of cells can be differentiated within the tumor microenvironment; an appreciably significant portion of these cells are myeloid-derived immune cells that become immunosuppressive under tumor signaling (myeloid-derived suppressor cells, MDSC) [74, 75]. MDSCs typically positively express the surface markers CD11b and Gr-1. The key function of these cells is suppression of the effector functions of NK and T cells [76, 77]. MDSCs were also shown to potentiate the immunosuppressive activity of APCs [78], which may inhibit the activity of T cells. van Vliet et al. demonstrated that macrophages and DCs inhibit effector T cells through MGL (macrophage galactose-type lectin), one of the C-type lectin receptors [79]. The interaction between MGL and CD45 of effector T cells reduced their proliferation and caused apoptosis. Schuette et al. [80] also revealed that the mannose receptor on DCs interacts with CD45 on cytotoxic T cells, thus resulting in their inhibition, reprogramming, and development of immunological tolerance.

THE EFFECT OF CD45 ON THE ACTIVITY OF CARS

Chimeric antigen receptors (CARs) are recombinant receptors that allow targeting immune cells to surface tumor-associated antigens (TAAs) [81]. CAR is a transmembrane molecule comprising an antigen recognition domain (it typically is a single-chain variable antibody fragment), the transmembrane domain, intracellular costimulatory domains (CD28, 4-1BB, and OX40 being most common), and the signaling domain (typically, CD3ζ) [82]. Several generations of CARs are currently known; they differ in the number of costimulatory domains or a set of auxiliary intracellular domains [83]. Despite the similar functionality, the effects of any activation of CARs and TCRs on cell proliferation and the cytotoxic response are different [84].

Structural and functional distinctions between TCRs and CARs

Unlike TCRs, which activate T cells after the recognition of 1–10 MHC molecules, several thousand surface TAA molecules are needed for CAR activation [84, 85]. There are many distinctions between CARs and TCRs, which are responsible for the increased threshold of the antigen content required for efficient activation of T cells. First, it is the receptor/ligand affinity: TCRs ensure MHC–antigen binding with micromolar affinity [85], while CARs bind their ligands with nanomolar affinity [86]. The increased affinity of CAR binding alters the kinetics that turn off the receptor, as well as its ability to be repeatedly activated and its mechanoreceptor function; these properties are believed to contribute to the ability of TCRs to perceive low ligand levels [87, 88]. After interaction with the antigen, the TCR and CD3δ, CD3ε, CD3γ, and CD3ζ bound to it assemble multicomponent signaling complexes [89, 90]. CARs interact with some signaling proteins of TCR; however, the quantitative and quali-

tative changes in the assembly of signaling complexes and the structure of IS alters the sensitivity to antigens [84, 91]. Imaging of the CAR and TCR synapses revealed that CAR synapses depend on the interaction between intercellular adhesion molecule-1 and integrin αLβ2 to a lesser extent, and that they are characterized by altered Lck localization compared to that for TCR [92–94]. Actin rings in the CAR synapse are much smaller than those in the TCR synapse, thus causing faster transduction of mechanical signals and dissociation of a CAR T cell from the target cell. Signal in the CAR synapse is initiated faster and is more intense, while signal duration is shorter than in the TCR synapse. This fact accelerates the involvement of the CAR T cell in the interaction with the target cell, causes fast release of cytotoxic granules in the IS, and rapid cytolysis of tumor cells [94].

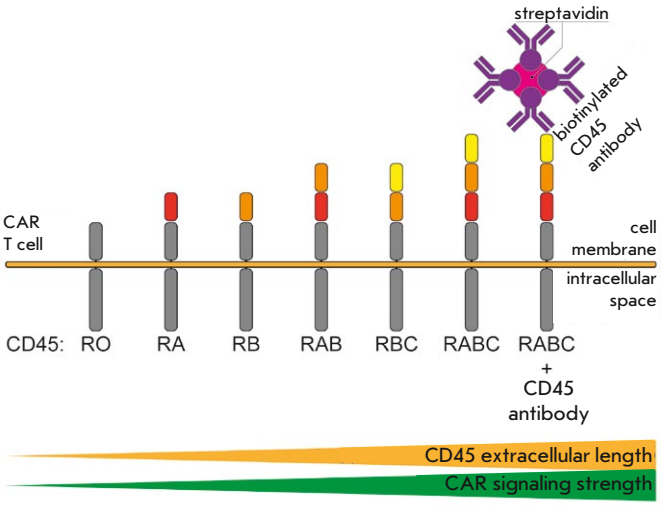


Fig. 3. The length of the extracellular part of CD45 influences CAR signaling. CD45 segregation from the immune synapse and CAR signaling strength as a result of an increase in the length of the extracellular portion of phosphatase. CAR T cell – T cell modified by the chimeric antigen receptor

The effect of CD45 on signal transduction upon activation of CARs

The release of CD45 molecules into the distal region of the immune synapse facilitates the phosphorylation of Lck and CD3ζ and it ensures tight contact in the receptor–antigen complex [5]. Karlsson et al. demonstrated that CD45 segregation is also required for activation of CAR19 to proceed, which is similar to TCR activation [95]. It is logical that activation of both CAR and TCR depends on CD45 segregation from the domain where the immune synapse takes shape. In both cases, CD45 substrates (SFKs) are expected to play a crucial role in signal transduction. The effects of CAR size, the distance to the TAA epitope to be recognized, and the CD45 length on the transduction of signals from CAR and activation of CAR19 T cells have been identified [96]: an increased size of the extracellular domain of CAR reduces CD45 segregation from the immune synapse area, phosphorylation of the signaling participants, and release of proinflammatory cytokines. The same dependence is also observed upon varying the distance to the TAA epitope to be recognized. An increase in CD45 length has an opposite effect regardless of the reason why it happens: this is true both for different CD45 isoforms and for the case when the molecule volume is increased by using specific antibodies (*Fig. 3*). These findings support the kinetic segregation model for CAR T cells [97] proposed by Karlsson et al. in their experiments [95].

CONCLUSIONS

Disruption of phosphorylation is one of the many causes of cancer. CD45 and other phosphatases play a positive role in oncogenesis by regulating the pro-oncogenic mechanisms; therefore, they are potential candidates for a targeted elimination of tumors or for increasing their sensitivity to chemo- or radiation therapy. Progress continues to be made in the research into CD45 using novel methods for designing and screening CD45 inhibitors, as well as technologies that enable synthesis inhibition and allow one to change the composition of CD45 isoforms in human hemocytoblast (CRISPR/Cas9). The known CD45 ligands (pUL11, E3/49K) and their analogs are also considered promising targets for cancer therapy. The importance of CD45 synthesis by tumor cells in predicting the clinical outcome of patients with CLL, ALL, MM, and DLBCL has been demonstrated. The course of many oncological diseases may also depend on the activity of this phosphatase. An important feature of CD45 is the typical composition of isoforms, which depends on cell differentiation. T cells with a naïve phenotype express CD45RA, while central memory and effector memory T cells express the CD45RO isoforms only. This segregation allows one to easily isolate the T-cell population of interest and then obtain CAR T cells with tailored properties. For example, memory T cells not expressing the CD45RA marker can be used to reduce the risk of graft-versus-host disease [98, 99]. The abundance of phosphatase among lymphoid and myeloid cells, as well as the high receptor level on the membrane [4], makes CD45 an extremely attractive target for CAR-T therapy both for hematopoietic tumors and upon the

conditioning of recipient's hematopoiesis preceding bone marrow transplantation. Along with the other methods used to control the activity of CAR T cells, the idea of regulating CAR activation by changing the CD45 length is very promising [100, 101]. Recent developments indicate that when designing a novel CAR, one needs to take into account the ratio between the sizes of the chimeric antigen receptor, the targeted antigen, and CD45 [102]. CD45 knockout is another performance potential associated with the enhancement of the CAR T cell. CD45 is essential for T-cell development and maturation; however, the absence of CD45 expression on CAR T cells can increase the safety of adoptive immunotherapy by reducing the risk of adverse reactions related to TCR signaling. ●

This work was supported by the Russian Science Foundation (grant No. 17-74-30019).

REFERENCES

1. Lynch K.W., Weiss A. // J. Biol. Chem. 2001. V. 276. № 26. P. 24341–24347.
2. Tong A., Nguyen J., Lynch K.W. // J. Biol. Chem. 2005. V. 280. № 46. P. 38297–38304.
3. Dahlke M.H., Larsen S.R., Rasko J.E.J., Schlitt H.J. // Leuk. Lymphoma. 2004. V. 45. № 2. P. 229–236.
4. Hermiston M.L., Xu Z., Weiss A. // Annu. Rev. Immunol. 2003. V. 21. № 1. P. 107–137.
5. Chang V.T., Fernandes R.A., Ganzinger K.A., Lee S.F., Siebold C., McColl J., Jönsson P., Palayret M., Harlos K., Coles C.H., et al. // Nat. Immunol. 2016. V. 17. № 5. P. 574–582.
6. Nam H.-J., Poy F., Saito H., Frederick C.A. // J. Exp. Med. 2005. V. 201. № 3. P. 441–452.
7. Kashio N., Matsumoto W., Parker S., Rothstein D.M. // J. Biol. Chem. 1998. V. 273. № 50. P. 33856–33863.
8. Trowbridge I.S., Thomas M.L. // Annu. Rev. Immunol. 1994. V. 12. № 1. P. 85–116.
9. D'Oro U., Sakaguchi K., Appella E., Ashwell J.D. // Mol. Cell. Biol. 1996. V. 16. № 9. P. 4996–5003.
10. Mustelin T., Taskén K. // Biochem. J. 2003. V. 371. № 1. P. 15–27.
11. Chow L.M.L., Fournel M., Davidson D., Veillette A. // Nature. 1993. V. 365. № 6442. P. 156–160.
12. Castro-Sanchez P., Teagle A.R., Prade S., Zamoyska R. // Front. Cell Dev. Biol. 2020. V. 8. P. 1–31.
13. Dustin M.L. // Cancer Immunol. Res. 2014. V. 2. № 11. P. 1023–1033.
14. Freiberg B.A., Kupfer H., Maslanik W., Delli J., Kappler J., Zaller D.M., Kupfer A. // Nat. Immunol. 2002. V. 3. № 10. P. 911–917.
15. Leupin O., Zaru R., Laroche T., Müller S., Valitutti S. // Curr. Biol. 2000. V. 10. № 5. P. 277–280.
16. Johnson K.G., Bromley S.K., Dustin M.L., Thomas M.L. // Proc. Natl. Acad. Sci. USA. 2000. V. 97. № 18. P. 10138–10143.
17. Davis S.J., Shaw A.S., Dustin M.L. // Semin. Immunol. 2000. V. 12. № 1. P. 5–21.
18. Irles C., Symons A., Michel F., Bakker T.R., van der Merwe P.A., Acuto O. // Nat. Immunol. 2003. V. 4. № 2. P. 189–197.
19. Burroughs N.J., Wülfing C. // Biophys. J. 2002. V. 83. № 4. P. 1784–1796.
20. Cordoba S.-P., Choudhuri K., Zhang H., Bridge M., Basat A.B., Dustin M.L., van der Merwe P.A. // Blood. 2013. V. 121. № 21. P. 4295–4302.
21. Carbone C.B., Kern N., Fernandes R.A., Hui E., Su X., Garcia K.C., Vale R.D. // Proc. Natl. Acad. Sci. USA. 2017. V. 114. № 44. P. 9338–9345.
22. Burroughs N.J., Lazic Z., van der Merwe P.A. // Biophys. J. 2006. V. 91. № 5. P. 1619–1629.
23. Choudhuri K., Wiseman D., Brown M.H., Gould K., van der Merwe P.A. // Nature. 2005. V. 436. № 7050. P. 578–582.
24. Courtney A.H., Shvets A.A., Lu W., Griffante G., Mollenauer M., Horkova V., Lo W.-L., Yu S., Stepanek O., Chakraborty A.K., et al. // Sci. Signal. 2019. V. 12. № 604. P. 1–14.
25. McNeill L., Salmond R.J., Cooper J.C., Carret C.K., Cassady-Cain R.L., Roche-Molina M., Tandon P., Holmes N., Alexander D.R. // Immunity. 2007. V. 27. № 3. P. 425–437.
26. Zikherman J., Jenne C., Watson S., Doan K., Raschke W., Goodnow C.C., Weiss A. // Immunity. 2010. V. 32. № 3. P. 342–354.
27. Kishihara K., Penninger J., Wallace V.A., Kündig T.M., Kawal K., Wakeham A., Timms E., Pfeffer K., Ohashi P.S., Thomas M.L., et al. // Cell. 1993. V. 74. № 1. P. 143–156.
28. Byth K.F., Conroy L.A., Howlett S., Smith A.J., May J., Alexander D.R., Holmes N. // J. Exp. Med. 1996. V. 183. № 4. P. 1707–1718.
29. Mee P.J., Turner M., Basson M.A., Costello P.S., Zamoyska R., Tybulewicz V.L.J. // Eur. J. Immunol. 1999. V. 29. № 9. P. 2923–2933.
30. Pingel S., Baker M., Turner M., Holmes N., Alexander D.R. // Eur. J. Immunol. 1999. V. 29. № 8. P. 2376–2384.
31. Molina T.J., Kishihara K., Siderovskid D.P., van Ewijk W., Narendran A., Timms E., Wakeham A., Paige C.J., Hartmann K.-U., Veillette A., et al. // Nature. 1992. V. 357. № 6374. P. 161–164.
32. Benatar T., Carsetti R., Furlonger C., Kamalia N., Mak T., Paige C.J. // J. Exp. Med. 1996. V. 183. № 1. P. 329–334.
33. Cyster J.G., Healy J.I., Kishihara K., Mak T.W., Thomas M.L., Goodnow C.C. // Nature. 1996. V. 381. № 6580. P. 325–328.
34. Zikherman J., Doan K., Parameswaran R., Raschke W., Weiss A. // Proc. Natl. Acad. Sci. USA. 2012. V. 109. № 1. P. 3–12.
35. Zhu J.W., Brdicka T., Katsumoto T.R., Lin J., Weiss A. // Immunity. 2008. V. 28. № 2. P. 183–196.
36. Springer T.A. // Nature. 1990. V. 346. № 6283. P. 425–434.
37. Hynes R.O. // Cell. 1992. V. 69. № 1. P. 11–25.
38. English B.K., Ihle J.N., Myracle A., Yi T. // J. Exp. Med. 1993. V. 178. № 3. P. 1017–1022.
39. Zaffran Y., Escallier J.C., Ruta S., Capo C., Mege J.L. // J. Immunol. 1995. V. 154. № 7. P. 3488–3497.
40. Lichtenberg U., Quintrell N., Bishop J.M. // Oncogene. 1992. V. 7. № 5. P. 849–858.

41. Roach T., Slater S., Koval M., White L., McFarland E.C., Okumura M., Thomas M., Brown E. // Curr. Biol. 1997. V. 7. № 6. P. 408–417.
42. Cahir McFarland E.D., Hurley T.R., Pingel J.T., Sefton B.M., Shaw A., Thomas M.L. // Proc. Natl. Acad. Sci. USA. 1993. V. 90. № 4. P. 1402–1406.
43. Shiroo M., Goff L., Biffen M., Shivnan E., Alexander D. // EMBO J. 1992. V. 11. № 13. P. 4887–4897.
44. Sieh M., Bolen J.B., Weiss A. // EMBO J. 1993. V. 12. № 1. P. 315–321.
45. Rohwedder I., Kurz A.R.M., Pruenster M., Immler R., Pick R., Eggersmann T., Klapproth S., Johnson J.L., Alsina S.M., Lowell C.A., et al. // Haematologica. 2020. V. 105. № 7. P. 1845–1856.
46. Fumagalli L., Zhang H., Baruzzi A., Lowell C.A., Berton G. // J. Immunol. 2007. V. 178. № 6. P. 3874–3885.
47. Cassatella M.A. // Immunol. Today. 1995. V. 16. № 1. P. 21–26.
48. Lloyd A.R., Oppenheim J.J. // Immunol. Today. 1992. V. 13. № 5. P. 169–172.
49. Liles W.C., Ledbetter J.A., Waltersdorph A.W., Klebanoff S.J. // J. Immunol. 1995. V. 155. № 4. P. 2175–2184.
50. Harvath L., Balke J.A., Christiansen N.P., Russell A.A., Skubitz K.M. // J. Immunol. 1991. V. 146. № 3. P. 949–957.
51. Gao H., Henderson A., Flynn D.C., Landreth K.S., Ericson S.G. // Exp. Hematol. 2000. V. 28. № 9. P. 1062–1070.
52. Zhu J.W., Doan K., Park J., Chau A.H., Zhang H., Lowell C.A., Weiss A. // Immunity. 2011. V. 35. № 5. P. 757–769.
53. Janeway C.A., Medzhitov R. // Annu. Rev. Immunol. 2002. V. 20. № 1. P. 197–216.
54. Byeon S.E., Yi Y.-S., Oh J., Yoo B.C., Hong S., Cho J.Y. // Mediators Inflamm. 2012. V. 2012. P. 1–18.
55. Medzhitov R. // Nat. Rev. Immunol. 2001. V. 1. № 2. P. 135–145.
56. Cross J.L., Kott K., Miletić T., Johnson P. // J. Immunol. 2008. V. 180. № 12. P. 8020–8029.
57. Lanier L.L. // Curr. Opin. Immunol. 2003. V. 15. № 3. P. 308–314.
58. Samelson L.E. // Annu. Rev. Immunol. 2002. V. 20. № 1. P. 371–394.
59. Hesslein D.G.T., Takaki R., Hermiston M.L., Weiss A., Lanier L.L. // Proc. Natl. Acad. Sci. USA. 2006. V. 103. № 18. P. 7012–7017.
60. Valitutti S., Dessing M., Aktories K., Gallati H., Lanzavecchia A. // J. Exp. Med. 1995. V. 181. № 2. P. 577–584.
61. Valitutti S., Müller S., Dessing M., Lanzavecchia A. // J. Exp. Med. 1996. V. 183. № 4. P. 1917–1921.
62. Rizzo D., Lotay A., Gachard N., Marfak I., Faucher J.-L., Trimoreau F., Guérin E., Bordessoule D., Jaccard A., Feuillard J. // Am. J. Hematol. 2013. V. 88. № 9. P. 747–753.
63. Ozdemirli M., Mankin H.J., Aisenberg A.C., Harris N.L. // Cancer. 1996. V. 77. № 1. P. 79–88.
64. Ratei R., Sperling C., Karawajew L., Schott G., Schrappe M., Harbott J., Riehm H., Ludwig W.-D. // Ann. Hematol. 1998. V. 77. № 3. P. 107–114.
65. Cario G., Rhein P., Mitlöhner R., Zimmermann M., Bandapalli O.R., Romey R., Moericke A., Ludwig W.-D., Ratei R., Muckenthaler M.U. // Haematologica. 2014. V. 99. № 1. P. 103–110.
66. Ishikawa H., Mahmoud M.S., Fujii R., Abroun S., Kawano M.M. // Leuk. Lymphoma. 2000. V. 39. № 1–2. P. 51–55.
67. Lin H., Kolosenko I., Björklund A.-C., Protsyuk D., Österborg A., Grandér D., Tamm K.P. // Exp. Cell Res. 2013. V. 319. № 5. P. 600–611.
68. Liu S., Ishikawa H., Tsuyama N., Li F.-J., Abroun S., Otsuyama K., Zheng X., Ma Z., Maki Y., Iqbal M.S., et al. // Oncogene. 2006. V. 25. № 3. P. 419–429.
69. Ishikawa H., Tsuyama N., Kawano M.M. // Int. J. Hematol. 2003. V. 78. № 2. P. 95–105.
70. Gonsalves W.I., Timm M.M., Rajkumar S.V., Morice W.G., Dispenzieri A., Buadi F.K., Lacy M.Q., Dingli D., Leung N., Kapoor P., et al. // Leuk. Res. 2016. V. 44. P. 32–39.
71. Clark M.C., Pang M., Hsu D.K., Liu F.-T., de Vos S., Gascoyne R.D., Said J., Baum L.G. // Blood. 2012. V. 120. № 23. P. 4635–4644.
72. Hoyer K.K., Pang M., Gui D., Shintaku I.P., Kuwabara I., Liu F.-T., Said J.W., Baum L.G., Teitell M.A. // Am. J. Pathol. 2004. V. 164. № 3. P. 893–902.
73. Nangia-Makker P., Nakahara S., Hogan V., Raz A. // J. Bioenerg. Biomembr. 2007. V. 39. № 1. P. 79–84.
74. Schiavoni G., Gabriele L., Mattei F. // Front. Oncol. 2013. V. 3. P. 1–15.
75. Laplane L., Duluc D., Larmonier N., Pradeu T., Bikfalvi A. // Trends Cancer. 2018. V. 4. № 12. P. 802–809.
76. Gabrilovich D.I., Nagaraj S. // Nat. Rev. Immunol. 2009. V. 9. № 3. P. 162–174.
77. Serafini P., De Santo C., Marigo I., Cingarlini S., Dolcetti L., Gallina G., Zanovello P., Bronte V. // Cancer Immunol. Immunother. 2004. V. 53. № 2. P. 64–72.
78. Ostrand-Rosenberg S., Sinha P., Beury D.W., Clements V.K. // Semin. Cancer Biol. 2012. V. 22. № 4. P. 275–281.
79. van Vliet S.J., Gringhuis S.I., Geijtenbeek T.B.H., van Kooyk Y. // Nat. Immunol. 2006. V. 7. № 11. P. 1200–1208.
80. Schuette V., Embgenbroich M., Ulas T., Welz M., Schulte-Schrepping J., Draffehn A.M., Quast T., Koch K., Nehring M., König J., et al. // Proc. Natl. Acad. Sci. USA. 2016. V. 113. № 38. P. 10649–10654.
81. Gross G., Waks T., Eshhar Z. // Proc. Natl. Acad. Sci. USA. 1989. V. 86. № 24. P. 10024–10028.
82. Jayaraman J., Mellody M.P., Hou A.J., Desai R.P., Fung A.W., Pham A.H.T., Chen Y.Y., Zhao W. // EBioMedicine. 2020. V. 58. P. 1–12.
83. Fu Z., Zhou J., Chen R., Jin Y., Ni T., Qian L., Xiao C. // Oncol. Lett. 2020. V. 20. № 4. P. 1–9.
84. Salter A.I., Ivey R.G., Kennedy J.J., Voillet V., Rajan A., Alderman E.J., Voytovich U.J., Lin C., Sommermeyer D., Liu L. // Sci. Signal. 2018. V. 11. № 544. P. 1–35.
85. Watanabe K., Terakura S., Martens A.C., van Meerten T., Uchiyama S., Imai M., Sakemura R., Goto T., Hanajiri R., Imahashi N. // J. Immunol. 2015. V. 194. № 3. P. 911–920.
86. Hudecek M., Lupo-Stanghellini M.-T., Kosasih P.L., Sommermeyer D., Jensen M.C., Rader C., Riddell S.R. // Clin. Cancer Res. 2013. V. 19. № 12. P. 3153–3164.
87. Feng Y., Brazin K.N., Kobayashi E., Mallis R.J., Reinherz E.L., Lang M.J. // Proc. Natl. Acad. Sci. USA. 2017. V. 114. № 39. P. 8204–8213.
88. Valitutti S., Müller S., Cella M., Padovan E., Lanzavecchia A. // Nature. 1995. V. 375. № 6527. P. 148–151.
89. Voisinne G., Kersse K., Chaoui K., Lu L., Chaix J., Zhang L., Goncalves Menoita M., Girard L., Ounoughene Y., Wang H. // Nat. Immunol. 2019. V. 20. № 11. P. 1530–1541.
90. Chakraborty A.K., Weiss A. // Nat. Immunol. 2014. V. 15. № 9. P. 798–807.
91. Ramello M.C., Benzaïd I., Kuenzi B.M., Lienlaf-Moreno M., Kandell W.M., Santiago D.N., Pabón-Saldaña M., Darville L., Fang B., Rix U. // Sci. Signal. 2019. V. 12. № 568. P. 1–31.

92. Gudipati V., Rydzek J., Doel-Perez I., Gonçalves V.D.R., Scharf L., Königsberger S., Lobner E., Kunert R., Einsele H., Stockinger H. // Nat. Immunol. 2020. V. 21. № 8. P. 848–856.
93. James J.R., Vale R.D. // Nature. 2012. V. 487. № 7405. P. 64–69.
94. Davenport A.J., Cross R.S., Watson K.A., Liao Y., Shi W., Prince H.M., Beavis P.A., Trapani J.A., Kershaw M.H., Ritchie D.S. // Proc. Natl. Acad. Sci. USA. 2018. V. 115. № 9. P. 2068–2076.
95. Karlsson H., Svensson E., Gigg C., Jarvius M., Olsson-Strömberg U., Savoldo B., Dotti G., Loskog A. // PLoS One. 2015. V. 10. № 12. P. 1–20.
96. Xiao Q., Zhang X., Tu L., Cao J., Hinrichs C.S., Su X. // Sci. Immunol. 2023. V. 7. № 74. P. 1–30.
97. Davis S.J., van der Merwe P.A. // Nat. Immunol. 2006. V. 7. № 8. P. 803–809.
98. Chan W.K., Suwannasaen D., Throm R.E., Li Y., Eldridge P.W., Houston J., Gray J.T., Pui C.-H., Leung W. // Leukemia. 2015. V. 29. № 2. P. 387–395.
99. Ukrainskaya V., Molostova O., Shelikhova L., Pershin D., Kulakovskaya E., Volkov D., Rakhteenko A., Muzalevskii Y., Kazachenok A., Brilliantova V., et al. // Blood Adv. 2022. V. 6. № 19. P. 5582–5588.
100. Stepanov A.V., Kalinin R.S., Shipunova V.O., Zhang D., Xie J., Rubtsov Y.P., Ukrainskaya V.M., Schulga A., Konovalova E.V., Volkov D.V., et al. // Proc. Natl. Acad. Sci. USA. 2022. V. 119. № 46. P. 1–8.
101. Kalinin R.S., Petukhov A.V., Knorre V.D., Maschan M.A., Stepanov A.V., Gabibov A.G. // Acta Naturae. 2018. V. 10. № 2. P. 16–23.
102. Kulemzin S.V., Kuznetsova V.V., Mamonkin M., Taranin A.V., Gorchakov A.A. // Acta Naturae. 2017. V. 9. № 1. P. 6–14.

Chromosomal Aberrations As a Biological Phenomenon in Human Embryonic Development

A. D. Ivanova*, M. L. Semenova
Lomonosov Moscow State University, Biological Faculty, Moscow, 119991 Russian Federation
*E-mail: ankivanova2@gmail.com
Received: July 20, 2023; in final form, August 09, 2023
DOI: 10.32607/actanaturae.25255
Copyright © 2023 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **Frequent chromosomal abnormalities are a distinctive feature of early embryonic development in mammals, especially humans. Aneuploidy is considered as a contributing factor to failed embryo implantation and spontaneous abortions. In the case of chromosomal mosaicism, its effect on the potency of embryos to normally develop has not been sufficiently studied. Although, a significant percentage of chromosomal defects in early human embryos are currently believed to be associated with the features of clinical and laboratory protocols, in this review, we focus on the biological mechanisms associated with chromosomal abnormalities. In particular, we address the main events in oocyte meiosis that affects not only the genetic status of an unfertilized oocyte, but also further embryo viability, and analyze the features of first cleavage divisions and the causes of frequent chromosomal errors in early embryonic development. In addition, we discuss current data on self-correction of the chromosomal status in early embryos.**
KEYWORDS **chromosomal mosaicism, aneuploidy, preimplantation development.**
ABBREVIATIONS **ART – assisted reproductive technology; PGT – preimplantation genetic testing; IVF – *in vitro* fertilization; ICM – inner cell mass; TE – trophectoderm; MI – meiosis I; MII – meiosis II; SAC – spindle assembly checkpoint; ICSI – intracytoplasmic sperm injection; FISH – fluorescence *in situ* hybridization.**

INTRODUCTION

In the 2000s, preimplantation genetic testing (PGT) became widely used in assisted reproductive technology (ART) clinics. Using PGT techniques, approximately half of early human embryos were found to carry chromosomal abnormalities, whereas this rate was only 1% in early mouse embryos [1]. Apparently, embryonic chromosomal abnormalities are an inherent part of *Homo sapiens* evolution and control the reproduction process throughout life [2]. Chromosomal abnormalities span a wide range of genomic imbalances of varying severity, from whole-chromosome polyploidy and large structural aneuploidies to submicroscopic deletions and duplications. Aneuploid embryos contain cells with the same karyotype abnormalities. Mosaic embryos contain at least two cell lineages with different karyotypes.

At the preimplantation development stages, chromosomal abnormalities cannot be accurately diagnosed based on the morphological features of embryos [3], but later, they affect the ability to develop and are of great importance in ART clinical practice (*Fig. 1*). Any chromosomal abnormalities cause a genetic imbalance that adversely affects development processes driven by the embryo's own genome. In humans, massive activation of the embryonic genome (day 3 of development) coincides in time with a usually observed delay and arrest of embryo development, which may be related to the genetic imbalance caused by chromosomal abnormalities [4, 5]. However, even completely aneuploid embryos are capable of forming morphologically normal blastocysts. Later, aneuploidy prevents implantation and further development of the embryo, leading to spontaneous abortions in the early stages or to postnatal abnormalities [6, 7]. Therefore, in modern clinical practice, transfer is not performed if blastocyst aneuploidy is detected by PGT techniques. An increase in the embryonic an-

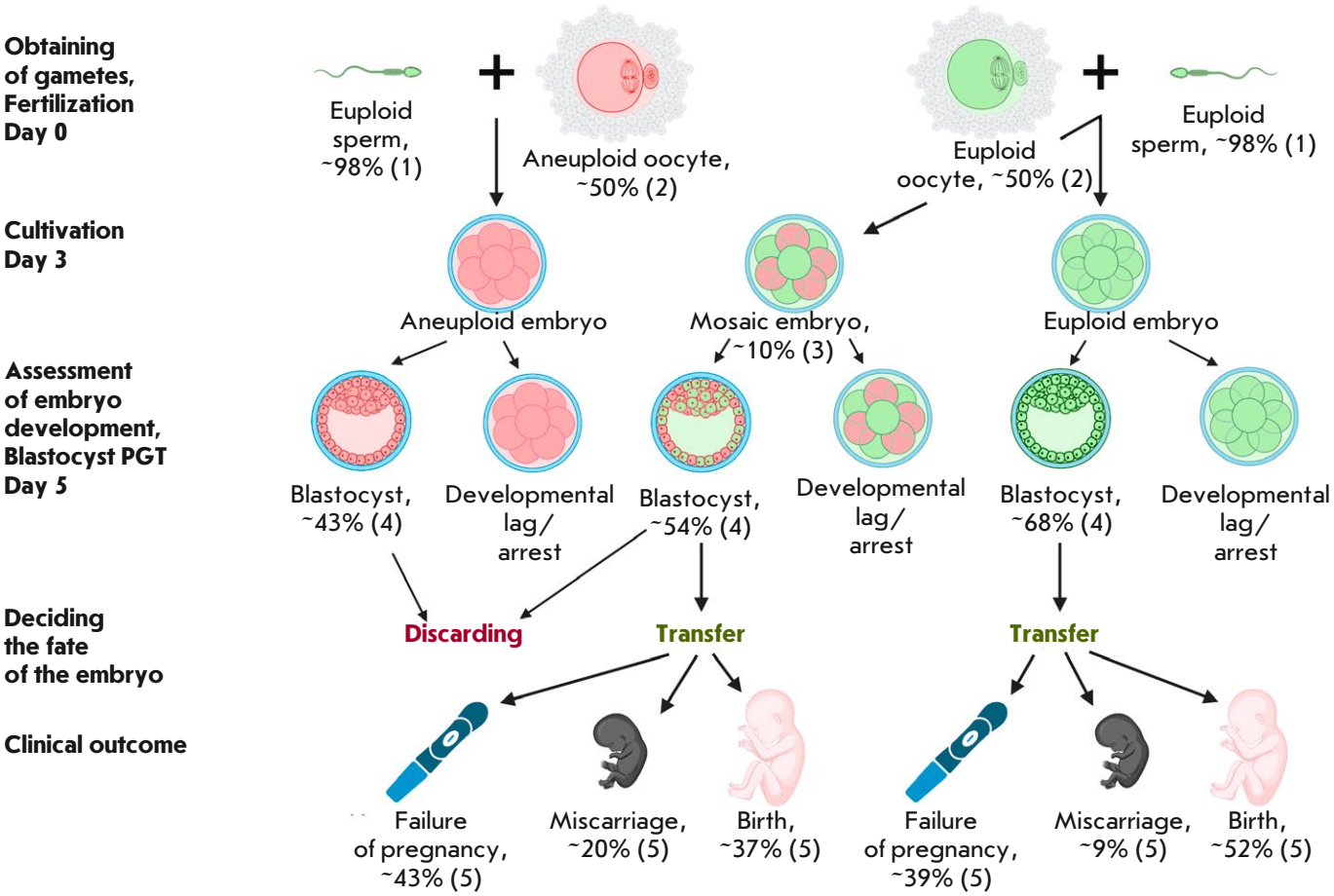


Fig. 1. Efficiency of IVF cycles, depending on the chromosomal status of gametes and embryos. Euploid cells are indicated in green, aneuploid cells are indicated in pink. 1) According to literature data, human spermatozoa in the vast majority of cases do not carry chromosomal abnormalities [23, 24]. 2) Mean rate of chromosomal abnormalities in human oocytes. The proportion of aneuploid oocytes varies from 20% to 80–90%, depending on maternal age (see [80]). 3) Mean rate of embryo mosaicism, based on experimental data [10, 15, 16]. 4) Blastocyst rate in embryos with different chromosomal statuses, according to experimental data [81]. 5) Clinical outcomes after transfer of euploid and mosaic embryos, according to experimental data [19]

euploidy rate with increasing age of a female is believed to be the main factor behind the gradual decrease in fertility [8].

According to current data, the rate of chromosomal mosaicism is not associated with maternal age [9–11]. Chromosomal mosaicism of human embryos is a phenomenon that is being actively studied by both researchers at scientific laboratories and embryologists at IVF clinics. Although chromosomal mosaicism in preimplantation embryos is increasingly recognized as a natural biological phenomenon [12], there is still a chance that the overall mosaicism rate is artificially increased by clinical factors [13, 14]. Recent multicenter studies have reported the mosaicism rate in PGT-screened embryos to be approximately 17% [15, 16], whereas another study reports a mosaicism rate of only 2.6% [10], and these differences are likely due to laboratory protocols.

Mosaic embryos containing an euploid cell line (euploid–aneuploid mosaics) are considered the most common [17] and, in some cases, have potencies to normal development. In clinical practice, births of healthy babies with normal karyotypes have been reported by females that had undergone mosaic embryo transfer [6, 18, 19]. If chromosomal mosaicism is detected, the decision to transfer or discard the

blastocyst depends on the mosaicism type, aneuploid cell percentage, and the chromosomes involved in the aneuploidy. Unfortunately, there is still no definitive data on the involvement of inner cell mass (ICM) cells, which would produce the fetus, in chromosomal mosaicism. There is evidence of a different probability of fetal involvement in chromosomal mosaicism, depending on the chromosome: the highest risk is associated with mosaicism of the autosomes 13, 18, and 21 and sex chromosomes [20].

Good clinical outcomes in mosaic embryo transfer may be associated both with certain biological mechanisms that promote the restoration of euploidy in cell lines and with an initially erroneous diagnosis of mosaicism. First, during PGT, the chromosomal status is determined in a limited area of the trophectoderm (TE). According to studies analyzing several biopsies from each embryo, euploidy and whole-chromosome aneuploidy are fairly reliable diagnoses, whereas a single analysis of a TE biopsy in mosaicism and segmental aneuploidy often does not reflect the chromosomal status of the entire embryo [21, 22]. Second, there is a widespread belief that the high rate of mosaic embryos in some clinics may be due not to biological reasons but to laboratory manipulations or technical factors [13, 14].

Despite the ongoing discussion about the technical aspects of mosaic embryo diagnosis, this review addresses in detail only the truly biological aspects of the formation of mosaic and aneuploid embryos and possible self-correction of their chromosomal status.

MECHANISMS OF ANEUPLOIDY INDUCTION IN EARLY HUMAN EMBRYOS

Embryonic chromosomal abnormalities may result from meiotic errors in oogenesis and spermatogenesis or mitotic errors in early development. Complete aneuploidy is of meiotic origin in 90% of cases. Sperm is believed to account for only 1–2% of embryonic aneuploidies, mainly segmental ones [23, 24]. For example, genotyping of 967 embryo biopsies revealed that about 70% of segmental aneuploidies were of paternal origin, whereas whole-chromosome aneuploidies were, mainly, related to maternal errors. About 70% and 30% of meiotic trisomies occur during the first and second meiotic divisions, respectively, in oogenesis [25].

In mammalian oocytes, centrioles are destroyed after the pachytene stage [26]. In some species, their function in meiosis is performed by acentriolar microtubule organizing centers [27]. After germinal vesicle breakdown in mouse oocytes, microtubules of the meiotic spindle are assembled and stabilized around chromatin, forming a few vesicular structures, followed by their orientation and the establishment of spindle poles and bipolarity; i.e., the meiotic spindle is assembled “inside out” by means of multiple acentriolar microtubule organizing centers [26, 27]. Unlike mouse oocytes, human oocytes lack not only centrosomes but also prominent acentriolar microtubule organizing centers. A few hours after germinal vesicle breakdown, microtubules form a small aster within the chromosome aggregate, and several more hours are required for the early spindle to form [28]. Spindle assembly in human oocytes relies on a gradient of the Ran-GTP complex around each chromosome. In addition to microtubule assembly, Ran-GTP also regulates the activity of motor proteins, such as HSET, a motor protein responsible for spindle pole focusing, and Kid, a motor protein that promotes chromosome alignment on the metaphase plate [29]. The meiosis I (MI) spindle poles of the human oocyte are initially poorly defined; chromosomes often change their position on a spindle that can temporarily become multipolar. In this case, kinetochores are often attached to more than one pole, which can further lead to errors in chromosome segregation [28]. Chromosomes are aligned on the metaphase plate 16 h, the anaphase begins 18 h, and the first polar body is abscised approximately 20 h after germinal vesicle breakdown. Meiosis II (MII) spindle assembly occurs more rapidly. The MII metaphase plate forms in the oocyte approximately 24 h after the onset of maturation, and the oocyte becomes ready for fertilization [28]. In contrast to MI, the multipolar spindle stage is rare in MII [30], which may explain the more frequent chromosomal errors in MI.

Paradoxically, meiosis in the absence of centrosomes may be a mechanism meant to protect against additional increases in the rate of maternal aneuploidy. For example, an artificially increased HSET level in mouse oocytes was shown to accelerate spindle bipolarization and promote the formation of more focused poles, similarly to mitotic ones. This change in meiotic spindle morphogenesis was sufficient for total disruption of chromosome segregation [31].

Aneuploidies are an order of magnitude more common in early human embryos than in the embryos of other mammalian species [1]. One of the causes may be the insufficient levels of KIFC1, which stabilizes the meiotic spindle, in human oocytes. Thus, administration of exogenous KIFC1 to human oocytes reduces the rate of meiotic chromosomal errors; on the contrary, a decrease in the KIFC1 level in cattle and mouse oocytes leads to spindle instability and increased chromosome segregation errors [32].

The aneuploidy rate in early human embryos is known to increase with maternal age. Oocytes are arrested in the MI prophase, from the embryonic period until ovulation. During this long period, chromatid cohesion is weakened due to the depletion of cohesion molecules, which is a major factor contributing to an increase in the rate of chromosomal errors as females age [33]. In MI, both homologous chromosomes and sister chromatids in the bivalent are held together by a ring-like cohesin structure. Cohesin which holds together homologous chromosomes is cleaved at the MI anaphase, whereas cohesin which holds together sister chromatids needs to remain in place longer to ensure sister chromatid cohesion until the MII anaphase. More than 90% of meiotic chromosomal errors arise due to premature separation of sister chromatids [34]. In MI, there can be reverse chromosome segregation when sister chromatids, rather than homologous chromosomes, separate at the anaphase. The rate of this phenomenon in human oocytes soars with maternal age [35]. Reverse chromosome segregation in MI results in normal DNA copy numbers in daughter cells; but in MII, the chromatids are unbound by cohesin, which contributes to segregation errors [36]. Finally, as maternal age increases, spindle assembly checkpoint (SAC) efficiency decreases; the SAC is the spindle assembly control point that inhibits the onset of anaphase until all chromosome kinetochores are correctly attached to the spindle [37, 38].

Mammalian oocyte meiosis is a complex multistep process that is subject to frequent chromosomal malfunction. Furthermore, additional species-specific features interfere with a correct progression of meiosis in human oocytes. The lack of centrosomes and acentriolar microtubule organizing centers, spindle pole instability, multipolar spindle stages, insufficient expression of the genes whose products stabilize the spindle and control meiosis stages, and depletion of cohesion molecules – all these factors together contribute a great deal to the emergence of chromosomal aberrations in the meiosis of human oocytes.

MECHANISMS OF CHROMOSOMAL MOSAICISM OCCURRENCE IN EARLY HUMAN EMBRYOS

The most common cause of chromosomal mosaicism in early embryos is postzygotic (mitotic) errors in chromosome segregation. Unlike aneuploidy, no significant relationship between chromosomal mosaicism and maternal age has been found [9, 10]. The first cell divisions are at the highest risk of mitotic errors [17, 39]. Mosaicism has recently been shown to occur in most cases as early as at the two-cell stage [40], although it was previously thought that mitotic errors most often occur in the second or third division, probably due to the gradual depletion of the maternal transcripts involved in mitosis [41]. Insufficient or absent expression of cell cycle checkpoint genes potentially increases the rate of mitotic errors. Recently, it has been found that the first transcriptional processes in the human embryo occur as early as at the pronuclei stage [42], but massive activation of the genome occurs only after the second or third cell division [43, 44]. Cell cycle drivers are intensively activated only at the morula stage [45]. In addition, the SAC efficiency is suggested to become sufficiently reliable only when the nuclear–cytoplasmic ratio in embryonic cells is restored [46].

Sperm centrosome destruction may also be the cause of mosaicism in early human embryos [47]. The sperm centrosome forms the spindle of the first cleavage division (the egg does not carry its own centrosome), and its integrity is required for mitotic divisions after fertilization [48]. Otherwise, the spindle is not constructed correctly, which leads to errors in the distribution of chromosomes between daughter cells. This was confirmed by a clinical study that revealed that fertilization of oocytes using ICSI by physically separated sperm segments increased the rate of chromosomal mosaicism in embryos [47].

Mitotic errors associated with mosaicism in an originally euploid embryo include anaphase lag, mitotic nondisjunction, endoreplication, formation of tripolar spindles, premature division of cells before DNA replication, and chromosome breakage [39, 49].

Anaphase lag and mitotic nondisjunction are considered the most common causes of mosaicism in cleavage embryos. Anaphase lag results in chromosome loss in one cell line without a corresponding increase in the number of chromosomes in another cell line. This phenomenon implies the retention of one or more chromosomes at the mitotic spindle equator after most sister chromatids of other chromosomes have separated and begun segregation towards the poles. The most common cause of anaphase lag is the attachment of kinetochores to microtubules emanating from both poles of the spindle (merotelic attachments [50]). In addition, lagging chromosomes may be insufficiently replicated, entangled, or not captured by the spindle at all. Later, the lagging chromosomes can be included in micronuclei [51].

Mitotic nondisjunction implies an uneven distribution of chromatids between two daughter cells, without loss of chromosomal material, which results in an increase in the number of DNA copies in one cell line and a decrease in another. Apparently, this is also associated with abnormalities in kinetochore orientation (i.e., their attachment to the spindle poles via microtubules). A single-cell FISH analysis of 138 mosaic

cleavage-stage embryos revealed that 78% of mosaic chromosomal abnormalities in chromosomes 5–8 had to do with mitotic nondisjunction (monosomic and trisomic abnormal cell lines in the embryo), and that only 20% of abnormalities were associated with anaphase lag (only monosomic abnormal cell lines in the embryo) [52]. Opposite results were obtained in a recent study using 24-chromosomal FISH: a total of 35.21% of the chromosomes were characterized by monosomy, and only 5.64% were characterized by trisomy (tested chromosomes, n = 5,547; tested cells, n = 250; tested blastocysts, n = 17); i.e., the predominant mechanism of mosaicism could be presumed to be associated with anaphase lag and chromosome loss. Analysis of mosaicism using chromosome copy numbers revealed that trisomy occurs more often than monosomy only in sex chromosomes [53].

Less common is mosaicism in preimplantation embryos associated with other mitotic errors. Endoreplication (the cause of mosaicism in 1.4% of cases [52]), which implies repeated replication of chromosomes without cell division, leads to the formation of tetraploid cells. Then, the chromosomes of tetraploid cells can be redistributed in subsequent divisions in various ways, but the number of chromosome copies in most daughter cells exceeds the norm. Chromosome breakage and premature cell division before DNA replication lead to the opposite situation when the chromosome copy number is decreased. In addition, abnormal tripolar spindles formed due to disturbances in the centrosomal regulator PLK4 lead to massive chromosome loss in nascent cell lines [54].

Therefore, the occurrence of chromosomal mosaicism in early embryos may theoretically be associated with many different mechanisms. However, there is still no reliable data that allows us to draw clear conclusions about the predominance of one mechanism over the other. For example, studies comparing the rates of anaphase lag and mitotic nondisjunction have a number of limitations. The rate of cell division in different cell lines may vary. Upon an initially equal number of monosomic and trisomic cells, one of the cell lines under study may be more noticeable due to a high rate of cell division [55], or one of the cell lines may be more actively eliminated during embryo development.

HYPOTHESIS OF SELF-CORRECTION OF ABNORMAL EMBRYOS AT EARLY DEVELOPMENTAL STAGES

In clinical practice, there have been reported cases of mosaic embryo transfer to patients who had not produced euploid embryos in IVF cycles. Although, the risk of negative clinical outcomes upon mosaic embryo transfer is higher than that upon euploid embryo transfer [56], in some cases, mosaic embryo transfer results in births of children with normal karyotypes. The first evidence-based study on this issue was published in 2015. Mosaic embryos were transferred to 18 female patients; there were 8 clinical pregnancies which led to the birth of 6 healthy children. All pregnancies that got to term were confirmed, by means of sampling of the chorionic villi, to have a normal karyotype [57]. These results, as well as the Preimplantation Genetic Diagnosis International Society (PGDIS) recommendations stating the possibility of mosaic embryo transfer in the absence of euploid ones [58], enabled large sample size studies. One of the latest large studies provides data on the outcomes of 137 mosaic embryo transfers. For 8 of the 37 registered live births, prenatal genetic testing was performed and normal chromosomal complement was detected [18]. Another publication reported 29 transfers of low-level mosaic blastocysts, which resulted in clinical pregnancy; prenatal testing revealed a 100% euploidy rate [6]. Positive clinical outcomes were also obtained in 36 pregnancies after the transfer of embryos with various levels and types of mosaicism: amniocentesis revealed a normal karyotype in each of these cases, and the pregnancies led to the birth of healthy children [59]. In addition, there were cases of mosaicism detected at prenatal testing which resulted in healthy live births with normal karyotypes [60]. Another interesting clinical case is the birth of a child after transfer of an embryo with 35% mosaicism of monosomy 2. A peripheral blood chromosome analysis of this newborn revealed only 2% mosaic monosomy 2 [61].

Definitely, positive clinical outcomes of mosaic embryo transfer may be partly explained by a low level of true biological mosaicism; i.e., by a false-positive diagnosis of mosaicism at preimplantation stages. However, an alternative explanation may be the elimination of the genetic aberrations detected at the blastocyst stage at later stages of development [12, 13, 62]. Probably, self-correction processes are activated in order to prevent the consequences of associated gene imbalance [13].

There are three hypothetical models of self-correction: predominant distribution, clonal depletion, and correction through a second mitotic error (*Fig. 2)* [63]. The predominant distribution model suggests an uneven allocation of aneuploid cells to the ICM and TE as the early embryo divides into these two cell lineages. If most abnormal cells are allocated into the TE, the effect of mosaicism on fetal development is not that significant. The clonal depletion model presumes a higher division rate of euploid cells compared with

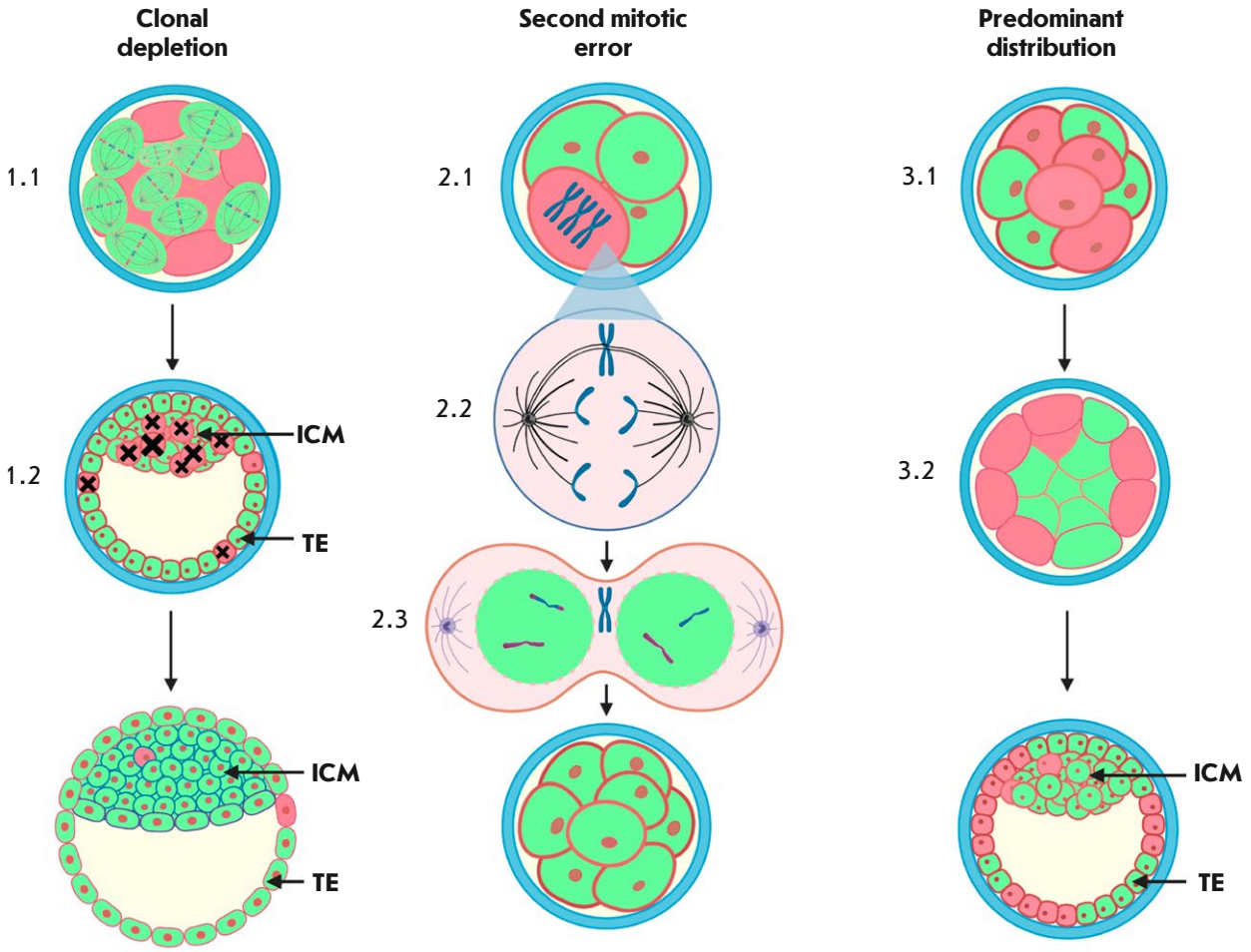


Fig. 2. Models of self-correction of the chromosomal status in mosaic embryos. Euploid cells are indicated in green, aneuploid cells are indicated in pink. Spindles (1.1) reflect an increase in the proliferative activity of euploid cell lines in mosaic embryos. Black crosses indicate apoptotic processes in aneuploid cells (1.2). A trisomal aneuploid embryonic cell (2.1) can undergo corrective mitotic division. One of the chromosomes remains at the mitotic spindle equator due to merotelic attachment of microtubules to kinetochores (2.2) and is not further included in the nuclei of daughter cells (2.3). (3.1, 3.2) Displacement of aneuploid cells to the embryo periphery, to the area of the nascent trophectoderm

that of aneuploid cells, as well as apoptotic death and elimination of abnormal cells. According to the third self-correction model, a second mitotic error can correct the chromosome set in abnormal cells to a normal configuration.

Data that the rate of fetal mosaicism (~0.2% according to amniocentesis results) is an order of magnitude less than that of placental mosaicism (~2% according to chorionic villi karyotyping results) [20, 64] may indicate a predominant distribution of abnormal cells in the TE. On the other hand, the initial ratio of euploid and abnormal cells in the TE and ICM may be similar, but in the ICM, the mechanisms for eliminating aneuploid cells work more efficiently. Even during normal development, a surge in programmed cell death is observed in the ICM of euploid embryos, which is associated with choosing the future by ICM cells and their division into the hypoblast and the epiblast [65]. Numerous studies comparing TE and ICM samples from human mosaic blastocysts have revealed no evi-

dence of predominant distribution of aneuploid cells in blastocyst TEs [66–69]. Time-lapse recording of embryo development in a mouse model of artificially induced chromosomal mosaicism did also not detect a predominant distribution of abnormal cells in the TE [12].

However, this study [12] revealed severe proliferative defects in the abnormal cell line in the TE and frequent apoptotic death of aneuploid ICM cells. The mechanisms of cell elimination in mammalian embryos are activated at the late stages of preimplantation development. Apoptotic cell death is first observed at the blastocyst stage, with these processes being more marked in ICM cells than in TE cells [70]. Probably, this fact may explain the higher activity of self-correction mechanisms through clonal depletion in fetal tissues. Experiments with chimeric embryos showed that some mosaic embryos have full developmental potential, provided that they contain a sufficient percentage of euploid cells [12]. A similar study clarified that the elimination of aneuploid cells is based on p53-dependent processes involving both autophagy and apoptosis before, during, and after implantation; on the other hand, euploid cells undertake compensatory proliferation during the implantation period [71]. In human embryos, proliferation and cell death levels are also increased in mosaic blastocysts compared with those in euploid blastocysts [67, 69]. A study conducted on rhesus monkey embryos demonstrated that self-correction of mosaicism may involve cellular fragmentation of abnormal blastomeres [51]. Studies in the laboratory of I.N. Lebedev (Research Institute of Medical Genetics, Tomsk) have revealed that dead cells are present in the cavity of mosaic blastocysts, and that karyotype abnormalities in them are much more common than in ICM and TE cells of the same blastocysts [72, 73]. Similar results were reported in a recent study that compared the chromosomal status of TE biopsies and samples consisting of cells left in the zona pellucida after blastocyst hatching (cellular debris). An abnormal karyotype was detected in 85.7% of cellular debris samples (n = 18); in this case, aneuploid and euploid statuses in the corresponding TE biopsies were detected in an equal ratio (9 : 9) [74]. Thus, the results of many studies argue for self-correction through clonal depletion of abnormal cells, and the mechanisms of action of this model may be different.

The model of self-correction through a second mitotic error is poorly supported by recent studies, at least in the case of whole-chromosome mosaic aneuploidies. Trisomic cell populations are theoretically able to self-correct by losing an additional chromosome [62], but in this case, the percentage of uniparental disomies should be quite high, whereas at the blastocyst stage, uniparental disomies are extremely rare (0.06%) [75]. However, the rate of uniparental disomies increases at the later development stages. A frequency of uniparental disomies of 2.1% was reported in fetuses with a normal karyotype, for which preliminary karyotyping of chorionic villi showed the presence of mosaicism [64]. Thus, the possibility of self-correction of mosaic embryos through a second mitotic error cannot be completely excluded. In the case of segmental abnormalities, this pathway seems more likely. Acentric chromosome fragments are unable to attach to the mitotic spindle; therefore, they can be lost [76].

Interestingly, fetal mosaicism usually involves sex chromosomal abnormalities or trisomy of chromosomes 21, 18, and 16 [20, 54], whereas individuals with complete aneuploidy of these chromosomes are viable. This observation suggests that self-correction mechanisms are more effective in the case of mosaicism of chromosomes whose aneuploid set more often leads to lethal outcomes. Another interesting fact is that transfer of mosaic embryos derived from the oocytes of young female patients provides better clinical outcomes compared with transfer of mosaic embryos from patients of late reproductive age (⩾ 34 years of age); i.e., self-correction mechanisms may be more effective in the embryonic cells of young female patients.

CONCLUSION

The topicality of studying chromosomal aberrations in early embryos and their impact on normal development has sharply increased as ART clinics have spread. The mechanisms of induction of complete embryonic aneuploidy are quite well studied, and aneuploidy has long been recognized as a factor that negatively affects the normal development of the embryo. The mechanisms of induction of chromosomal mosaicism have been less studied than those of complete aneuploidy. In addition, mitotic errors, unlike meiotic errors, can occur at different stages of embryo development. Decisions about the fate of mosaic embryos identified at IVF clinics are still made “doubtfully” due to the lack of sufficient fundamental knowledge. From a biological point of view, the developmental potential of mosaic embryos may depend both on the proportion and location of abnormal cells and on the numbers of chromosomes involved in mosaicism [6, 20, 76–78]. However, data from different research groups vary, probably, due to the effect of laboratory and technical factors on the actual biological events associated with chromosomal mosaicism. Most diagnoses of mosaic embryos may be false-positives [68, 79], which means that

most of the accumulated data on clinical outcomes after mosaic embryo transfer are no longer relevant. To date, regarding available scarce data, one may unequivocally say that, in some cases, mosaic embryo transfer results in the birth of a healthy child. Some data, discussed in this review, on self-correction of mosaic embryos inspire confidence and give hope to patients who have failed euploid embryos [12, 71]. On the other hand, potential risks should be taken into account. All patients who planned to undergo mosaic embryo transfer should receive thorough genetic counseling.

In this review, we have focused on the biological mechanisms of induction of chromosomal defects and combined data on the possible mechanisms of self-correction of abnormalities in embryo development. However, it should be borne in mind that the array of studies reviewed has a number of limitations, in particular embryo cultivation *in vitro* and differences in the techniques used for the diagnosis of the chromosomal status. Therefore, the data here address one aspect of the issue and are insufficient to understand the full picture. This mainly concerns such a controversial phenomenon as chromosomal mosaicism. Primarily, further research should focus more on a clear differentiation between "true" and "apparent in the PGT results" chromosomal mosaicism. ●

REFERENCES

1. Bond D.J., Chandley A.C., Chandlev A. // Aneuploidy. 1983. P. 27–54.
2. Viotti M. // Genes. 2020. V. 11. № 6. P. 602.
3. Lee C.I., Chen C.H., Huang C.C., Cheng E.H., Chen H.H., Ho S.T., Lin P., Lee M.S., Lee T.H. // Reprod. Biomed. Online. 2019. V. 39. № 4. P. 569–579.
4. Munné S., Grifo J., Cohen J., Weier H.U.G. // Am. J. Hum. Genet. 1994. V. 55. № 1. P. 150.
5. Campbell A., Fishel S., Bowman N., Duffy S., Sedler M., Hickman C.F. // Reprod. Biomed. Online. 2013. V. 26. № 5. P. 477–485.
6. Munné S., Spinella F., Grifo J., Zhang J., Beltran M.P., Fragouli E., Fiorentino F. // Eur. J. Med. Genet. 2020. V. 63. № 2. P. 103741.
7. Tiegs A.W., Tao X., Zhan Y., Whitehead C., Kim J., Hanson B., Osman E., Kim T.J., Patounakis G., Gutmann J., et al. // Fertility, Sterility. 2021. V. 115. № 3. P. 627–637.
8. Liu K., Case A. // J. Obstetrics Gynecology Canada. 2011. V. 33. № 11. P. 1165–1175.
9. Sachdev N.M., Ribustello L., Liu E., McCulloh D.H., Grifo J., Munne S. // Fertility, Sterility. 2016. V. 106. № 3. P. e156–e157.
10. Katz-Jaffe M., McReynolds S., De Klerk K., Henry L.N., Schweitz M., Swain J., Schoolcraft W.B. // Fertility, Sterility. 2017. V. 108. № 3. P. e87–e88.
11. Gao J., Wei N., Zhu X., Li R., Yan L., Qiao J. // J. Assisted Reproduction Genet. 2023. V. 40. № 5. P. 1089–1098.
12. Bolton H., Graham S.J.L., van der Aa N., Kumar P., Theunis K., Fernandez Gallardo E., Voet T., Zernicka-Goetz M. // Nat. Commun. 2016. V. 7. № 1. P. 1–12.
13. Capalbo A., Ubaldi F.M., Rienzi L., Scott R., Treff N. // Hum. Reprod. 2017. V. 32. № 3. P. 492–498.
14. Practice Committee and Genetic Counseling Professional Group (GCPG) of the American Society for Reproductive Medicine // Fertility, Sterility. 2020. V. 114. № 2. P. 246–254.
15. Munné S., Kaplan B., Frattarelli J.L., Child T., Nakhuda G., Shamma F.N., Silverberg K., Kalista T., Handyside A.H., Katz-Jaffe M., et al. // Fertility, Sterility. 2019. V. 112. № 6. P. 1071–1079. e7.
16. Capalbo A., Poli M., Rienzi L., Girardi L., Patassini C., Fabiani M., Cimadomo D., Benini F., Farcomeni A., Cuzzi J., et al. // Am. J. Hum. Genet. 2021. V. 108. № 12. P. 2238–2247.
17. van Echten-Arends J., Mastenbroek S., Sikkema-Raddatz B., Korevaar J.C., Heineman M.J., van der Veen F., Repping S. // Hum. Reprod. Update. 2011. V. 17. № 5. P. 620–627.
18. Zhang Y.X., Chen J.J., Nabu S., Yeung Q.S.Y., Li Y., Tan J.H., Suksalak W., Chanchamroen S., Quangkananurug W., Wong P.S., et al. // Genes. 2020. V. 11. № 9. P. 973.
19. Viotti M., Victor A.R., Barnes F.L., Zouves C.G., Besser A.G., Grifo J.A., Cheng E.H., Lee M.S., Horcajadas J.A., Corti L., et al. // Fertility, Sterility. 2021. V. 115. № 5. P. 1212–1224.
20. Grati F.R., Gallazzi G., Branca L., Maggi F., Simoni G., Yaron Y. // Reprod. Biomed. Online. 2018. V. 36. № 4. P. 442–449.
21. Navratil R., Horak J., Hornak M., Kubicek D., Balcova M., Tauwinklova G., Travnik P., Vesela K. // Mol. Hum. Reprod. 2020. V. 26. № 4. P. 269–276.
22. Marin D., Xu J., Treff N.R. // Prenatal Diagnosis. 2021. V. 41. № 5. P. 545–553.
23. McCoy R.C., Demko Z.P., Ryan A., Banjevic M., Hill M., Sigurjonsson S., Rabinowitz M., Petrov D.A. // PLoS Genet. 2015. V. 11. № 10. P. e1005601.
24. Bell A.D., Mello C.J., Nemesh J., Brumbaugh S.A., Wysoker A., McCarroll S.A. // Nature. 2020. V. 583. № 7815. P. 259–264.
25. Kubicek D., Hornak M., Horak J., Navratil R., Tauwinklova G., Rubes J., Vesela K. // Reprod. Biomed. Online. 2019. V. 38. № 3. P. 330–339.
26. Szollosi D., Calarco P., Donahue R.P. // J. Cell Sci. 1972. V. 11. № 2. P. 521–541.
27. Schuh M., Ellenberg J. // Cell. 2007. V. 130. № 3. P. 484–498.
28. Holubcová Z., Blayney M., Elder K., Schuh M. // Science. 2015. V. 348. № 6239. P. 1143–1147.
29. Ems-McClung S.C., Emch M., Zhang S., Mahnoor S., Weaver L.N., Walczak C.E. // J. Cell Biol. 2020. V. 219. № 2. e201906045.
30. Roeles J., Tsiavaliaris G. // Nat. Commun. 2019. V. 10. № 1. P. 1–10.

31. Bennabi I., Quéguiner I., Kolano A., Boudier T., Mailly P., Verlhac M.H., Terret M.E. // EMBO Repts. 2018. V. 19. № 2. P. 368–381.
32. So C., Menelaou K., Uraji J., Harasimov K., Steyer A.M., Seres K.B., Bucevicius J., Lukinavicius G., Mobius W., Sibold C., et al. // Science. 2022. V. 375. № 6581. P. eabj3944.
33. Lee H.L., McCulloh D.H., Hodes-Wertz B., Adler A., McCaffrey C., Grifo J.A. // J. Assist. Reprod. Genet. 2015. V. 32. № 3. P. 435–444.
34. Capalbo A., Bono S., Spizzichino L., Biricik A., Baldi M., Colamaria S., Ubaldi F.M., Rienzi L., Fiorentino F. // Hum. Reprod. 2013. V. 28. № 2. P. 509–518.
35. Patel J., Tan S.L., Hartshorne G.M., McAinsh A.D. // Biology Open. 2016. V. 5. № 2. P. 178–184.
36. Zielinska A.P., Holubcova Z., Blayney M., Elder K., Schuh M. // Elife. 2015. V. 4. P. e11389.
37. Musacchio A., Salmon E.D. // Nat. Rev. Mol. Cell Biol. 2007. V. 8. № 5. P. 379–393.
38. Lagirand-Cantaloube J., Ciabrini C., Charrasse S., Ferrieres A., Castro A., Anahory T., Lorca T. // Sci. Rept. 2017. V. 7. № 1. P. 1–14.
39. Taylor T.H., Gitlin S.A., Patrick J.L., Crain J.L., Wilson J.M., Griffin D.K. // Hum. Reprod. Update. 2014. V. 20. № 4. P. 571–581.
40. Cavazza T., Takeda Y., Politi A.Z., Aushev M., Aldag P., Baker C., Choudhary M., Bucevicius J., Lukinavicius G., Elder K., et al. // Cell. 2021. V. 184. № 11. P. 2860–2877. e22.
41. Baart E.B., Martini E., van den Berg I., Macklon N.S., Galjaard R.H., Fauser B.C.J.M., van Opstal D. // Hum. Reprod. 2006. V. 21. № 1. P. 223–233.
42. Asami M., Lam B.Y.H., Ma M.K., Rainbow K., Braun S., VerMilyea M.D., Yeo G.S.H., Perry A.C.F. // Cell Stem Cell. 2022. V. 29. № 2. P. 209–216. e4.
43. Leng L., Sun J., Huang J., Gong F., Yang L., Zhang S., Yuan X., Fang F., Xu X., Luo Y., et al. // Cell Stem Cell. 2019. V. 25. № 5. P. 697–712. e6.
44. Wells D., Bermudez M., Steuerwald N., Thornhill A., Walker D., Malter H., Delhanty J., Cohen J. // Hum. Reprod. 2005. V. 20. № 5. P. 1339–1348.
45. Kiessling A. A. // Nat. Biotechnol. 2010. V. 28. № 10. P. 1025–1026.
46. Kyogoku H., Kitajima T.S. // Developmental Cell. 2017. V. 41. № 3. P. 287–298. e4.
47. Palermo G.D., Colombero L.T., Rosenwaks Z. // Rev. Reprod. 1997. V. 2. P. 19–27.
48. Silber S., Escudero T., Lenahan K., Abdelhadi I., Kilani Z., Munne S. // Fertility, Sterility. 2003. V. 79. № 1. P. 30–38.
49. Mantikou E., Wong K.M., Repping S., Mastenbroek S. // Biochim. Biophys. Acta (BBA)-Mol. Basis Dis. 2012. V. 1822. № 12. P. 1921–1930.
50. Cimini D., Howell B., Maddox P., Khodjakov A., Degrassi F., Salmon E. // J. Cell Biol. 2001. V. 153. № 3. P. 517–528.
51. Daughtry B.L., Rosenkrantz J.L., Lazar N.H., Fei S.S., Redmayne N., Torkenczy K.A., Adey A., Yan M., Gao L., Park B., et al. // Genome Res. 2019. V. 29. № 3. P. 367–382.
52. Munné S., Sandalinas M., Escudero T., Márquez C., Cohen J. // Reprod. Biomed. Online. 2002. V. 4. № 3. P. 223–232.
53. Ioannou D., Fonseka K.G.L., Meershoek E.J., Thornhill A.R., Abogrein A., Ellis M., Griffin D.K. // Chromosome Res. 2012. V. 20. № 4. P. 447–460.
54. Levy B., Hoffmann E.R., McCoy R.C., Grati F.R. // Prenatal Diagnosis. 2021. V. 41. № 5. P. 631–641.
55. Munné S., Wells D. // Fertility, Sterility. 2017. V. 107. № 5. P. 1085–1091.
56. Viotti M., Greco E., Grifo J.A., Madjunkov M., Librach C., Cetinkaya M., Kahraman S., Yakovlev P., Kornilov N., Corti L., et al. // Fertility, Sterility. 2023. S0015-0282(23)00716-1.
57. Greco E., Minasi M.G., Fiorentino F. // N. Engl. J. Med. 2015. V. 373. № 21. P. 2089–2090.
58. Leigh D., Cram D.S., Rechitsky S., Handyside A., Wells D., Munne S., Kahraman S., Grifo J., Katz-Jaffe M., Rubio C. // Reprod. Biomed. Online. 2022. V. 45. № 1. P. 19–25.
59. Lee C.I., Cheng E.H., Lee M.S., Lin P.Y., Chen Y.C., Chen C.H., Huang L.S., Huang C.C., Lee T.H. // J. Assisted Reprod. Genet. 2020. V. 37. № 9. P. 2305–2313.
60. Chen C.P., Lin Y.H., Chern S.R., Wu P.S., Chen S.W., Wu F.T., Lee M.S., Chen Y.Y., Wang W. // Taiwanese J. Obstetrics Gynecol. 2020. V. 59. № 1. P. 146–149.
61. Kahraman S., Cetinkaya M., Yuksel B., Yesil M., Pirkevi Cetinkaya C. // Hum. Reprod. 2020. V. 35. № 3. P. 727–733.
62. Bazrgar M., Gourabi H., Valojerdi M.R., Yazdi P.E., Baharvand H. // Stem Cells Devel. 2013. V. 22. № 17. P. 2449–2456.
63. Delhanty J.D.A. // Hum. Fertility. 2013. V. 16. № 4. P. 241–245.
64. Malvestiti F., Agrati C., Grimi B., Pompilii E., Izzi C., Martinoni L., Gaetani E., Liuti M.R., Trotta A., Maggi F., et al. // Prenatal Diagnosis. 2015. V. 35. № 11. P. 1117–1127.
65. Xenopoulos P., Kang M., Puliafito A., Di Talia S., Hadjantonakis A.K. // Cell Rept. 2015. V. 10. № 9. P. 1508–1520.
66. Fragouli E., Lenzi M., Ross R., Katz-Jaffe M., Schoolcraft W.B., Wells D. // Hum. Reprod. 2008. V. 23. № 11. P. 2596–2608.
67. Capalbo A., Wright G., Elliott T., Ubaldi F.M., Rienzi L., Nagy Z.P. // Hum. Reprod. 2013. V. 28. № 8. P. 2298–2307.
68. Popovic M., Dheedene A., Christodoulou C., Taelman J., Dhaenens L., van Nieuwerburgh F., Deforce D., van den Abbeel E., De Sutter P., Menten B., et al. // Hum. Reprod. 2018. V. 33. № 7. P. 1342–1354.
69. Victor A.R., Tyndall J.C., Brake A.J., Lepkowsky L.T., Murphy A.E., Griffin D.K., McCoy R.C., Barnes F.L., Zouves C.G., Viotti M. // Fertility, Sterility. 2019. V. 111. № 2. P. 280–293.
70. Winiarczyk D., Piliszek A., Sampino S., Lukaszewicz M., Modliński J.A. // Reproduction, Fertility, Devel. 2021. V. 33. № 12. P. 725–735.
71. Singla S., Iwamoto-Stohl L.K., Zhu M., Zernicka-Goetz M. // Nat. Commun. 2020. V. 11. № 1. P. 1–15.
72. Zhigalina D.I., Skryabin N.A., Artyukhova V.G., Svetlakov A.V., Lebedev I.N. // Tsitologiya. 2016. V. 58. № 6. P. 488–492.
73. Tšuiko O., Zhigalina D.I., Jatsenko T., Skryabin N.A., Kanbekova O.R., Artyukhova V.G., Svetlakov A.V., Teearu K., Trosin A., Salumets A., et al. // Fertility, Sterility. 2018. V. 109. № 6. P. 1127–1134. e1.
74. Wang X., Zhao J., Yao Z., Xia Q., Chang T., Zeng J., Liu J., Li Y., Zhu H. // Reproductive Sci. 2023. P. 1–11.
75. Gueye N.A., Devkota B., Taylor D., Pfundt R., Scott Jr. R.T., Treff N.R. // Fertility, Sterility. 2014. V. 101. № 1. P. 232–236.
76. Fragouli E., Alfarawati S., Spath K., Babariya D., Tarozzi N., Borini A., Wells D. // Hum. Genet. 2017. V. 136. № 7. P. 805–819.

77. Spinella F., Fiorentino F., Biricik A., Bono S., Rubert, A., Cotroneo E., Baldi M., Cursio E., Minasi M.G., Greco E. // Fertility, Sterility. 2018. V. 109. № 1. P. 77–83.
78. Martin A., Mercader A., Dominguez F., Quiñonero A., Perez M., Gonzalez-Martin R., Delgado A., Mifsud A., Pellicer A., De Los Santos M.J. // Front. Mol. Biosci. 2023. V. 10. P. 264.
79. Fragouli E., Munné S., Wells D. // Hum. Reprod. Update. 2019. V. 25. № 1. P. 15–33.
80. Franasiak J.M., Forman E.J., Hong K.H., Werner M.D., Upham K.M., Treff N.R., Scott R.T. // J. Assist. Reprod. Genet. 2014. V. 31. № 11. P. 1501–1509.
81. Rubio C., Rodrigo L., Mercader A., Mateu E., Buendia P., Pehlivan T., Viloria T., Santos D.L., Simon C., Remohi J., et al. // Prenatal Diagnosis: Published in Affiliation With the International Society for Prenatal Diagnosis. 2007. V. 27. № 8. P. 748–756.

The Role of Autophagy in the Development of Pathological Conditions of the Body

U. S. Kench[1,2], S. S. Sologova[2], V.S. Prassolov[1*], P. V. Spirin[1**]
[1]Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russian Federation
[2]Department of Pharmacology, Nelyubin Institute of Pharmacy, I.M. Sechenov First Moscow State Medical University (Sechenov University), Moscow, 119991 Russian Federation
*E-mail: prassolov45@mail.ru
**E-mail: spirin.pvl@gmail.com
Received: June 28, 2023; in final form, July 23, 2023
DOI: 10.32607/actanaturae.23838
Copyright © 2023 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License,which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **Autophagy is the process of lysosomal elimination of the cell organelles, cytoplasmic sites, and pathogenic microorganisms that enter the cell. This process is associated with both cell death regulation and an increase in cell survival chances. Autophagy is involved in the development of various diseases (Crohn disease, cancer, atherosclerosis, etc.). For these reasons, it is of significant interest to establish the molecular targets involved in autophagy regulation and the factors that mediate its participation in pathogenesis. The review describes the potential molecular mechanisms involved in the regulation of autophagy, its contribution to the vital cell activity in a healthy organism, and pathologies.**
KEYWORDS **autophagy, apoptosis, cell death, lysosomes.**

INTRODUCTION

Autophagy is the mechanism of removal of non-required and damaged organelles and cell cytosol regions. It is considered a compensatory response that is a result of the lack of nutrients in a cell, as well as a response to stress. In some cases, activation of autophagy leads to cell death. Thus, autophagy, on the one hand, protects cells from unfavorable external and internal factors, and, on the other hand, leads to cell death if it is impossible to save the cell and in case of viral or bacterial infection.

1. Mechanisms of autophagy regulation

During autophagy, an autophagosome is formed around the target to be degraded and the target then undergoes lysis. The following autophagy stages are usually distinguished: initiation, elongation, autophagosome formation, and formation of an autophagolysosome, followed by its degradation (*Fig. 1*).

Stage I. The initiation of autophagy begins with the extension of a section of the rough endoplasmic reticulum (ER) membrane, followed by its detachment. During initiation, the ULK complex is recruited to the outer ER membrane, leading to a change in the membrane structure. The ULK complex, which consists of the ULK1, Atg13, FIP200, and Atg101 proteins, is formed through dephosphorylation of the Atg13 and ULK1 proteins and simultaneous drop in the kinase activity of the mTORc1 complex. Dephosphorylation of Atg13 and ULK1 triggers the assembly of an active ULK complex [1]. Dephosphorylated Atg13, a ULK complex member, binds to Atg14, a member of the PI3KC3 complex (Vps34), while ULK1 phosphorylates the proteins Beclin1 (Atg6) and Vps34 and, thus, activates them (*Fig. 1*).

Beclin1 is the key protein for the PI3KC3 complex formation, while Vps34 is involved in the production of phosphoinositol triphosphate from the phosphoinositol diphosphate (PI2P) on the ER membrane surface. PI3P is required for the recruitment of the other proteins involved in phagophore formation and its subsequent transition to an autophagosome.

Stage II. The PI3KC3 complex (phosphatidylinositol-3-kinase class 3), together with the ULK complex, promotes the extension of the ER membrane fragment and its subsequent detachment with the formation of a phagophore [1].

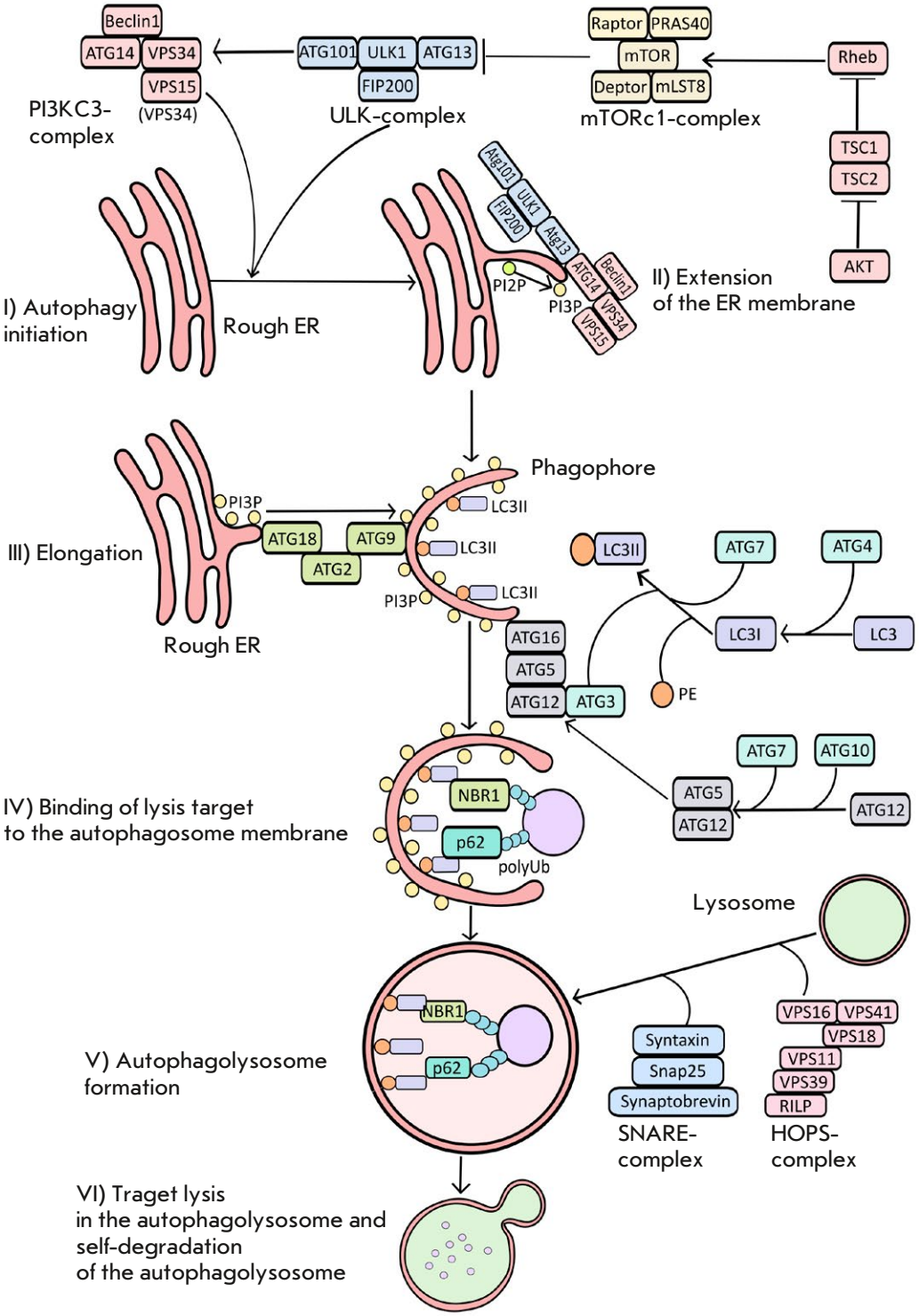


Fig. 1. Schematic representation of autophagy. Autophagy is initiated by the inhibition of the mTORc1 complex, which prevents the ULK complex assembly (mTORc1 phosphorylates Atg13, which inhibits the assembly of the active ULK complex). Stage I – autophagy initiation. Stage II – endoplasmic reticulum membrane extension. Stage III – elongation. Stage IV – recruitment of the degradation target to the autophagosome. Stages V–VI – autophagolysosome formation and target lysis

Stage III. Phagophore elongation includes modifications of its structure (enrichment of the phagophore membrane with PI3P, recruitment of LC3II), which are required for the binding of the target to be degraded to the autophagosome membrane. At this stage, the major conjugate complex Atg12/Atg5/Atg16 plays a key role. Conjugate formation begins with the processing of the ubiquitin-like protein Atg12, which is carried out by the ubiquitin-E1-like activating enzyme Atg7 [2] and the ubiquitin-E2-like enzyme Atg10 [3]. Atg5 and Atg16 then join the activated Atg12 (*Fig. 1*). The conjugate is also necessary to recruit the other proteins involved in elongation and ensure phagophore membrane extension [1–3].

Stage IV. At this stage, the target to be degraded is bound and positioned inside the autophagosome with the use of LC3II. LC3II is produced as a result of proteolytic cleavage of LC3 by the cysteine protease Atg4 with formation of the intermediate product LC3I. With the involvement of Atg7 and Atg3, LC3I interacts with phosphatidylethanolamine (PE) to form LC3II and anchor it on the phagophore membrane [2, 4, 5]. The Atg8 and GABARAP proteins possess functions that are similar to those of LC3II [5].

Simultaneously with LC3II processing, additional enrichment of the PI3P phagophore is carried out by the Atg9/Atg2/Atg18 conjugate, which transfers PI3P from the ER to the phagophore (*Fig. 1*, stage III) [6]. Atg13 initiates formation of the Atg9/Atg2/Atg18 conjugate.

For further autophagosome formation, the proteins that have already completed their function must detach from the phagophore. LC3II is one of the few proteins that remains on the phagophore membrane. Adaptor proteins are required to bind the target to be degraded to LC3II. Protein p62 (SQSTM1) is one of the adaptor proteins. It is involved in the regulation of various signaling pathways, since it can bind to polyubiquitinated proteins, which are components of a number of signaling pathways, and induce their degradation in the autophagosome [7].

Stage V. Autophagolysosomes are formed as a result of autophagosome and lysosome fusion with the participation of a complex of proteins. HOPS is the main complex. It is composed of the proteins VPS16, VPS41, VPS18, VPS11, VPS39, RILP, and Rab7. This complex is responsible for the fusion of autophagosome and the lysosome membranes [8]. The SNARE/SNAP25 complex, which consists of the proteins Syntaxin, SNAP25 (SNAP27), and Synaptobrevin, is also required for membrane fusion.

Stage VI. At this stage, the target is degraded by the lysosomal enzymes inside the autophagolysosome. The same enzymes that performed substrate degradation eventually degrade the autophagolysosome.

1.1. Role of mTORc1

The mTORc1 complex, which consists of the proteins mTOR, PRAS40, Deptor, Raptor, and mLST8, is the main regulator of autophagy (*Fig. 1*). Multiple signaling pathways, with PI3K/AKT/mTOR being the main one, regulate mTORc1 activity. Positive regulation of mTORc1 involves an active protein Rheb, which is repressed by the TSC1/2 complex. AKT acts as a negative regulator of TSC1/2 and thus functions as one of the main kinases responsible for autophagy regulation [9].

1.2. Role of calcium

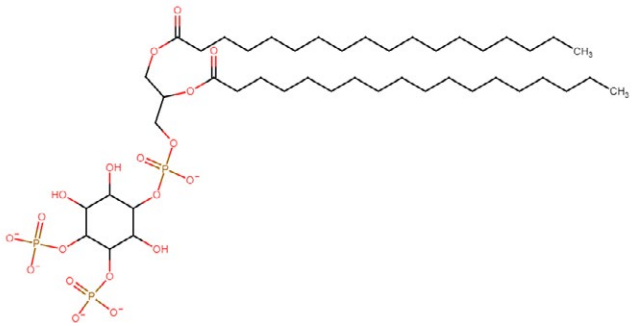
Calcium can act as both an inducer and a repressor in autophagy regulation.

The inhibitory effect of Ca2+ ions is implemented through their ability to activate calpain, a calcium-dependent cysteine protease that degrades autophagy-initiating proteins (Atg5, Beclin1, and PTEN).

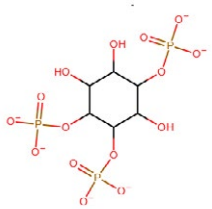
Activation of the PHLPP1β phosphatase due to an increased calpain level leads to the suppression of the ERK1/2 and AKT activities, which disrupts lysosome function. The intracellular protein calpastatin is a natural calpain inhibitor.

The action of Ca2+ ions on autophagy is implemented through CaMKKβ kinase-mediated activation of AMPK. AMPK inhibits the mTORc1 complex and activates the TSC1/2 and ULK1 proteins. Calmodulin-dependent protein calcineurin is another calcium target. This protein dephosphorylates the transcription factor TFEB, which results in its activation (*Fig. 2*). Activated TFEB regulates the expression of genes encoding such autophagy proteins as LC3, Beclin1, and p62. Calmodulin activates Vps34 and calmodulin-dependent kinase DAPK, which is a direct inducer of Beclin1 [10].

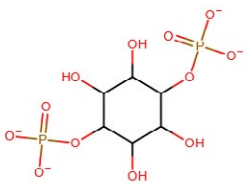
The calcium ion level in the cytoplasm is regulated by the activity of calcium channels, including IP3R. IP3R is an ER calcium channel; its function directly depends on the level of IP3, which opens the channel. Channel opening leads to calcium release from the ER into the cytosol. IP3R has an anti-autophagic effect, since it can inhibit the dissociation of the Beclin1/Bcl-2 complex, thereby reducing the level of active unbound Beclin1 in the cell. The PLC and IMPase proteins are involved in the regulation of the intracellular concentration of IP3; they convert phosphoinositol diphosphate (PI2P) to inositol triphosphate (IP3) and inositol monophosphate (IP1) to inositol, respectively. Inositol can be converted to its original state, PI2P [10].



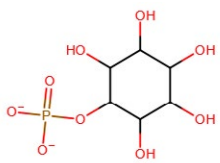
Phosphatidylinositol 4,5-bisphosphate (PIP2)



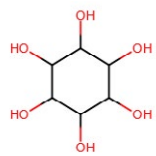
Inositol 1,4,5-trisphosphate (IP3)



Inositol 1,4-bisphosphate (IP2)



Inositol phosphate (IP)



Inositol (Ins)

The tumor suppressor p53 plays an important role in autophagy regulation. Depending on the molecular target it interacts with, p53 can act either as an autophagy activator or an inhibitor. The direct interaction of p53 with anti-apoptotic Bcl-2 family proteins (Bcl-2, Bcl-Xl, Mcl-1) inhibits their activity and induces apoptosis. In particular, the interaction of p53 with Bcl-2 causes dissociation of the Bcl-2/Beclin1 complex, with subsequent Beclin1 release and autophagy initiation [11]. An example of p53-mediated autophagy induction is the activation of TSC1/2 and Beclin1 by direct interaction between p53 and DAPK, a Beclin1 activator. The p53 protein can also interfere with autophagy initiation by disrupting ULK complex assembly through binding to the FIP200 protein. The p53 protein can also inhibit AMPK, one of the important autophagy activators (*Fig. 2*) [12].

Autophagy, despite its adaptive function, can be the cause of autophagy-dependent cell death (lethal autophagy), which is characterized by the appearance of a significant number of vacuoles in cells [13].

One of the possible mechanisms of lethal autophagy is the activation of ceramide synthase 1 (CerS1), which forms ceramide on the outer mitochondrial membrane. This causes mitochondrial degradation, since ceramide interacts with the LC3II receptor anchored on the autophagosomal membrane. Excessive accumulation of ceramide on the mitochondrial membrane was found to significantly increase the risk of lethal autophagy induction [13].

Target degradation can also occur without the formation of an autophagosome and other specific vesicles; this pathway is called chaperone-associated autophagy [14]. It begins with the formation of a transmembrane channel by the oligomeric lysosomal protein LAMP2A (CD107). This channel materiliazes upon the occurrence of a so-called misfolded protein, which has an abnormal conformation and contains a unique KFERQ motif, in the cytosol. To form the channel, the KFERQ motif of the misfolded protein must recruit a complex of proteins including HSC70, which acts as a chaperone [10, 14]. The misfolded protein then passes through the LAMP2A channel into the lysosome, where it is degraded.

2. ROLE OF AUTOPHAGY IN DISEASE DEVELOPMENT

Autophagy is involved in the development of a number of human diseases (atherosclerosis, diabetes mellitus, ischemia of different localization, hepatocirrhosis, chronic obstructive pulmonary disease, etc.), including both disease onset and response to it.

2.1. Autophagy and neurodegenerative diseases

Neurodegenerative diseases form an extensive group of pathologies caused by neuronal cell death. The mechanisms underlying the development of these diseases are not fully clear. However, it is known that they are usually associated with the production and accumulation of agglomerates of misfolded proteins with an abnormal structure both in the intercellular

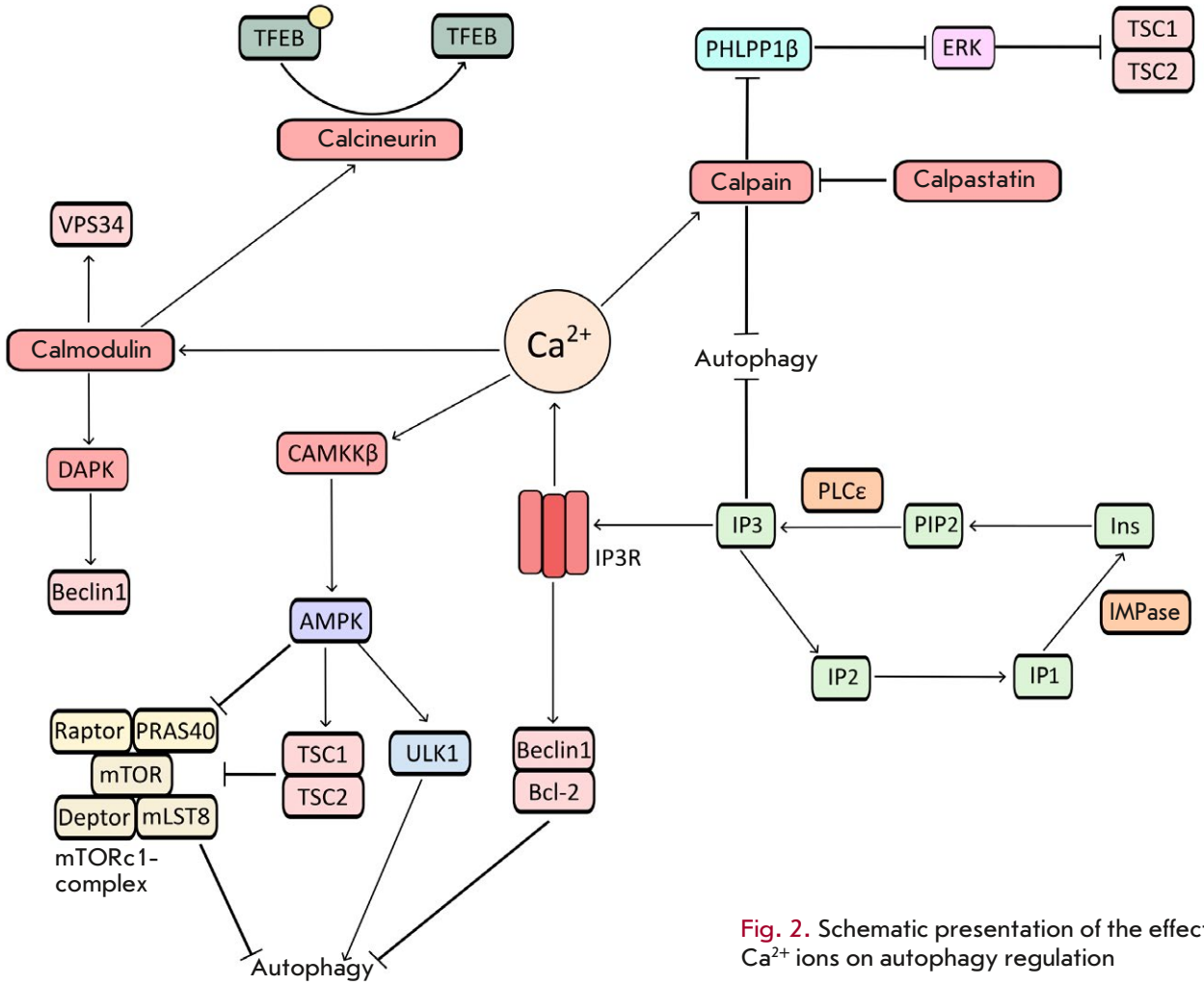


Fig. 2. Schematic presentation of the effect of Ca^{2+} ions on autophagy regulation

space and in cells. Both central and peripheral nervous system cells can be involved in neurodegeneration, which causes gradual impairment of motor, psychological, and cognitive functions.

There are many mechanisms that aim to eliminate misfolded proteins in the cell, including autophagy. Autophagy can be either triggered by ER stress induction, in particular, the PERK/eIF2A/ATF4-signaling pathway, as a response to the production of misfolded proteins (*Fig. 3*) or mediated by chaperones [15]. Autophagy can also participate in the elimination of misfolded proteins characteristic of a specific neurodegenerative disease [1]. Impaired removal of misfolded proteins leads to their accumulation and further aggregation in bodies and plaques. Lewy bodies (α-synuclein) are formed in cells in Parkinson disease, while senile (β-amyloid) and neurofibrillary (tau protein) plaques are produced in Alzheimer disease.

Autophagy impairment in neurodegenerative diseases is accompanied by the accumulation of lysosomes and immature autophagosomes in neurons. This phenomenon is associated with impaired inactivation of the TFEB transcription factor, which regulates the expression of many genes encoding autophagy proteins (LC3, Beclin1, p62, etc.) and the proteins involved in lysosome biogenesis.

In normal conditions, the mTORc1 complex plays a significant role in the regulation of TFEB activity; it inhibits TFEB translocation to the nucleus by phosphorylating it, which results in the formation of the 14-3-3σ/TFEB(P) complex in the cytoplasm. In neurodegenerative diseases, the activity of the mTORc1 complex drops, which results in TFEB release from the 14-3-3σ/TFEB(P) complex and its translocation to the nucleus (*Fig. 3*). Misfolded proteins were found to prevent TFEB inactivation, thus significantly increasing the expression of the genes they regulate. In neu-

rodegenerative diseases, this can be considered a cell compensatory response to a decrease in autophagy efficiency [16].

TFEB is also involved in the regulation of PINK1, a serine/threonine kinase responsible for the localization of ubiquitin ligase Parkin on the mitochondrial outer membrane (MOM) due to a decrease in the membrane potential of damaged mitochondria. Parkin polyubiquitinates MOM proteins, leading to the formation of the OMM/polyUb/p62-LC3II protein complex. This complex is required to recruit autophagosomes to mitochondra for the degradation of the latter in the autophagolysosome. Mitochondrial degradation in the autophagolysosome is called mitophagy. In normal conditions, PINK1 is transported to the mitochondrial matrix by TOM/TIM translocases, where it is degraded by the proteases PARL and MRR (*Fig. 3*) [16].

2.1.1. Huntington disease. Huntington disease is an autosomal-dominant disease. Its early stage is characterized by neurodegeneration of basal brain structures (striatum), while disease progression leads to complete atrophy of the cerebral cortex. The first symptoms appear at the age of 35–45 years. At early disease stages, motor functions are impaired, and cognitive and mental abnormalities can also be observed. Mental disorders such as aggression, depression, panic attacks, etc. develop during disease progression. Memory impairment and motor disorders become pronounced; bradykinesia, ataxia, and decreased reflexes are observed. Death occurs 15–20 years after the diagnosis. To date, there are no drugs to treat Huntington disease.

The pathogenesis of Huntington disease is associated with the expansion of CAG trinucleotide repeats in the *HTT* gene, which codes for the huntingtin (Htt) protein. In normal conditions, the number of such repeats does not exceed three. The expansion of repeats has a cumulative nature; i.e., the more there are repeats in *HTT*, the higher the risk of disease development is; however, ⩾40 CAG repeats are considered a critical threshold [17, 18]. In normal conditions, Htt participates in the axonal transport and acts as an adaptor protein to kinesin. The mutant Htt protein (mHtt) lacks the ability to recruit kinesin to the vesicle, resulting in impaired vesicular transport through the axon [19]. The mutant protein mHtt can interact with transcription factors such as CREB, CBP, TFIID, p53, and SP1 [19], disrupting their DNA-binding activity. This leads to reduced production of vital proteins. Unlike wild-type Htt, mHtt can induce autophagy gene expression through the ER stress activation (the PERK/eIF2A/ATF4 signaling pathway) (*Fig. 3*) [15]. In addition, mHtt can enhance TFEB dephosphorylation, leading to an increased transport of the latter into the nucleus and expression of genes encoding autophagy proteins (*Fig. 3*).

It was found that mHtt can directly bind to Beclin1 thus impairing the PI3KC3 complex assembly and autophagy initiation [18]. Thus, mHtt can affect autophagy activity through different pathways. Disruption of mHtt degradation due to impaired autophagy leads to its accumulation in the cell cytoplasm with the formation of protein aggregates, which ultimately results in a more aggressive disease course [15, 19].

Accumulation of mHtt in the cytoplasm can result in its association with p62 and disruption of LC3II function on the autophagasome membrane. This affects autophagosome formation around the target to be degraded, resulting in the formation of empty autophagosomes, with a possibility to induce cell necroptosis [19, 20].

2.1.2. Alzheimer disease and the role of autophagy in its development. Alzheimer disease (AD) is the most common type of senile dementia. Its early stages are characterized by impaired short-term memory and cognitive decline. As the disease progresses, a loss of communication functions, self-care ability, and speech impairment, up to complete aphasia, are observed.

The mechanisms of AD pathogenesis are not fully understood. There are several hypotheses describing the mechanisms underlying AD, including the tau hypothesis and the hypothesis of the accumulation of senile plaques. The role of mitochondria in AD has also been actively investigated.

The role of β-amyloid in AD pathogenesis has not been established yet; recent studies have questioned the theory of the leading role of senile plaques in neurodegeneration [21]. Beta-amyloid (βA) is a polypeptide consisting of 42 amino acid residues; stacks of these polypeptides form senile plaques. Beta-amyloid is formed by amyloidogenic proteolytic cleavage of the β-amyloid precursor protein (APP), whose gene is located on chromosome 21. APP participates in cell adhesion and contributes to cell survival. APP cleavage resulting in βA formation is mediated by β-secretase (BACE1) and γ-secretase. The C-terminus of APP is cleaved by γ-secretase, while its N-terminus is cleaved by β-secretase. Proteolytic cleavage of APP yields βA monomers, which form extracellular protein conglomerates: so-called senile plaques [22]. These structures spatially interfere with the formation of synaptic connections and initiate local inflammation due to the release of numerous pro-inflammatory factors by microglia cells. This inflammation is caused by the interaction of the transmembrane receptor TREM2

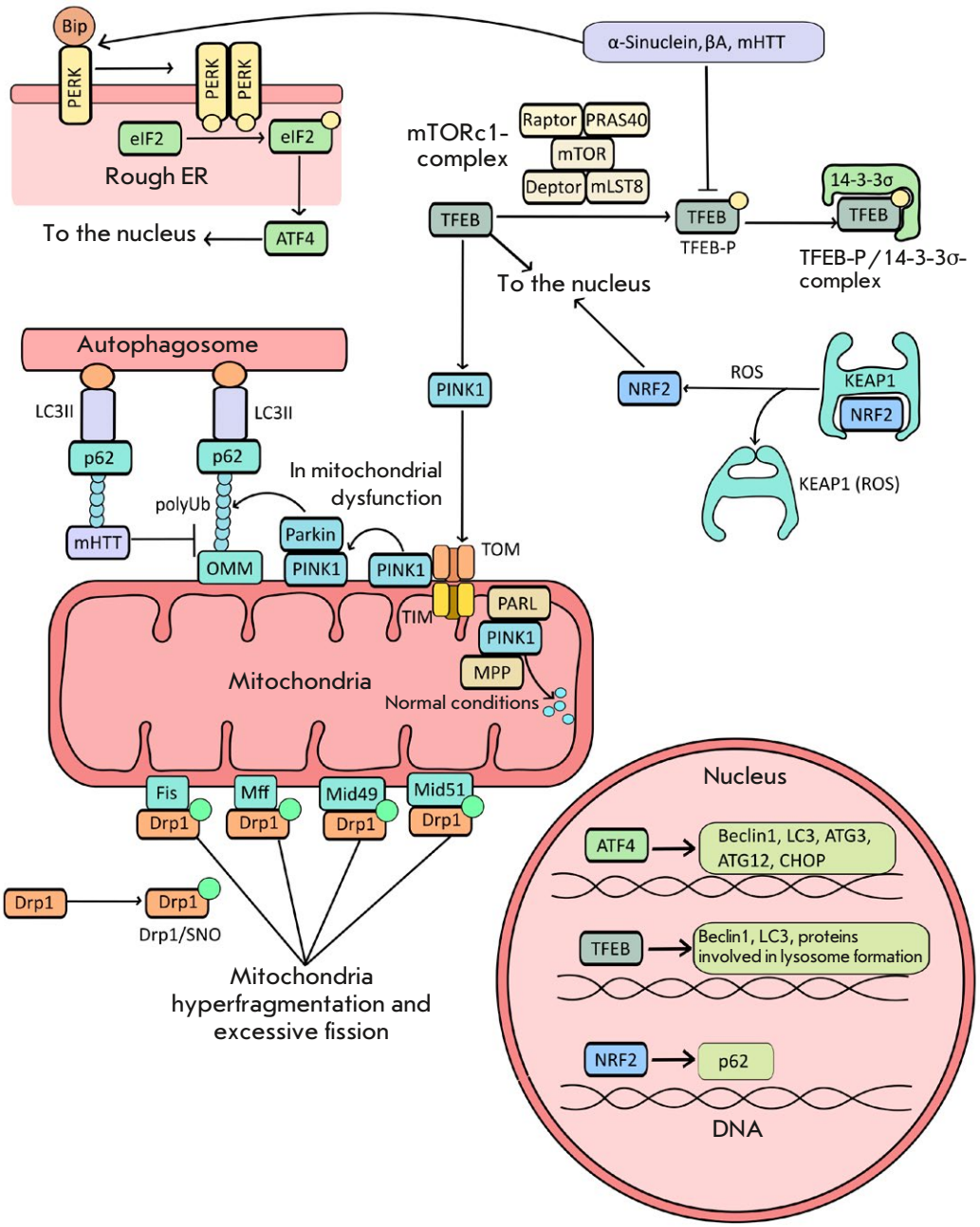


Fig. 3. Effect of ER stress-induced misfolded proteins on the regulation of TFEB activity and the mTORc1 complex in neurodegenerative diseases (α-synuclein, β-amyloid, and mHTT in Parkinson, Alzheimer, and Huntington diseases, respectively). The PERK/Bip heterodimeric complex is located on the ER membrane. The complex of chaperone Bip and PERK prevents ER stress activation. The Bip protein recognizes misfolded proteins and delivers them to proteasomes for degradation. Bip binds to misfolded proteins, leading to PERK release followed by its dimerization and activation. The PERK dimer phosphorylates the translation initiation factor eIF2, resulting in inactivation of the latter and inhibition of the translation of many proteins, except for the ATF4 transcription factor, which migrates to the nucleus and activates autophagy gene expression

with amyloid plaques, resulting in the activation of the NF-κB and Syk-kinase proteins, which are involved in the activation of cytokines and other inflammatory factors (IL-2 and NO synthases), and neuronal death [23].

Tau protein, which is a member of the microtubule-associated protein (MAP) family, is also involved in AD pathogenesis. MAP proteins provide microtubule rigidity and stiffness. This is achieved through the ability of tau to bind to tubulin and form "stiffening ribs" along microtubules. The efficiency of this binding depends on the tau phosphorylation level. The higher the phosphorylation level of the tau protein is, the lower its affinity for tubulin. In normal conditions, two to three amino acid residues are phosphorylated in the tau protein. Some tau protein mutations increase its phosphorylation level. For example, four missense mutations (G272V, P301L, V337M, and R406W) cause tau overphosphorylation and its detachment from microtubules. Tau detachment results in its accumulation in the cytoplasm, followed by its export into the intercellular space and aggregation in neurofibrillary tangles [24]. The condition characterized by such aggregation is called tauopathy [25].

There has been a growing body of evidence of the involvement of mitochondria in AD. APP accumulates in mitochondria as a result of its transfer from the cytoplasm to the mitochondrial intermembrane space by translocase TOMM40, where APP inhibits the cytochrome oxidase complex (complex IV of the electron transport chain), which reduces ATP production [26]. Beta-amyloid also binds to cyclophilin D, which is involved in the regulation of calcium levels in mitochondria and mitochondrial gene expression. As a consequence, disruption of cyclophilin D function decreases mitochondrial gene transcription and impairs the mitochondrial functioning [27].

Beta-amyloid is known to affect mitochondrial fission. During mitochondrial fission, Drp1 proteins form a ring-like structure around the organelle. Accumulation of βA triggers the production of inducible NO synthase (iNOS) in the cell, which is involved in the S-nitrosylation of Drp1 (Drp1/SNO) (*Fig. 3*). This modification disrupts the regulation of Drp1 oligomerization on the mitochondrial wall, resulting in abnormal fragmentation and increased mitochondrial number [28, 29]. Drp1 is recruited to the mitochondrial membrane through association with the adaptor transmembrane proteins Fis1, Mff, and Mid49/51 (*Fig. 3*) [29].

Autophagy is involved in APP degradation in vesicles. Autophagy impairment in AD is associated with the accumulation of a large number of immature autophagosomes in neurons due to a dysfunction of the ESCRT-III complex (cytosolic protein complex) involved in the formation of multivesicular bodies (MVBs). This complex is responsible for the transport of ubiquitinated membrane proteins to MVB. Inhibition of MVB formation makes their fusion with the autophagosome impossible. It also abrogates the formation of the late endosome and further destruction of APP in lysosomes [30, 31]. The direct effect of autophagy on AD pathogenesis is associated with Atg7, which is involved in βA transport to MVB. Atg7 participates in the accumulation of amyloid agglomerates in exosome vesicles and transportation of these agglomerates to the intercellular space. It was experimentally shown that suppression of the Atg7 activity by small interfering RNAs decreases βA production in neurons. Atg7 deficiency was shown to lead to a significant accumulation of hyperphosphorylated tau. Thus, Atg7 is actively involved in tau protein degradation and, therefore, can be directly involved in its turnover [31]. The evidence suggests that the proteins associated with autophagy regulation can be involved in AD development.

Another AD trait is the accumulation of reactive oxygen species (ROS) in neurons. ROS production is mediated by NADPH-oxidase 4 (NOX4), which is activated by the interaction of the transmembrane protein RAGE with βA molecules. It is important to note that ROS can both activate and inhibit autophagy.

An example of a positive effect of ROS on autophagy is the activation of the ROS-KEAP1-NRF2-p62 signaling pathway. ROS oxidize cysteine residues in KEAP1, which forms the heterodimeric KEAP1/NRF2 complex. Oxidation of KEAP1 residues leads to the release of the NRF2 transcription factor, which enhances the expression of the p62-encoding gene (*Fig. 3*) [32].

The negative effect of ROS on autophagy is associated with a decrease in HIF-1α activity. In the active state, this transcription factor enhances the transcription of the LC3, BNIP3/NIX, and REDD genes by interacting with their enhancers. Upon ROS accumulation in cells, the proline residues in HIF-1α are oxidized, resulting in HIF-1α polyubiquitination and its further proteolysis [33].

2.2. Autophagy and autoimmune diseases

There are numerous causes behind autoimmune diseases. Their development is associated, on the one hand, with the formation of a pool of mature B cells (plasmocytes) producing autoreactive antibodies, and, on the other, with a decrease in either the activity or number of regulatory T cells [34]. The occurrence of a pool of autoreactive lymphocytes can be associated with a disrupted selection of the entire pool of lymphocytes in the major immune organs. The subse-

quent defense response to autoreactive lymphocytes involves their elimination through interaction with the epithelial cells of the medullary region of the thymus stroma. These cells produce tissue-specific antibodies that interact with autoreactive lymphocytes, which ultimately leads to their death. This process is called autoresistance. Impaired autoresistance in autoimmune diseases is believed to maintain the pool of autoreactive lymphocytes.

In addition to maintaining a pool of autoreactive lymphocytes that are abnormally aggressive towards normal cells, the pathological immune response can be associated with impaired degradation of damaged or dead cell fragments, components of pathogenic microorganisms, and other antigens, followed by their accumulation due to impaired autophagy [35].

Autophagy also promotes the assembly of the MHC complexes involved in antigen presentation on the cell membrane. These complexes act as immune response activation signals. Autophagy disruption leads to impaired MHC II assembly, since no pathogen fragmentation in the autophagolysosome or interaction of the resulting fragments with MHC II takes place [35].

There are also mechanisms in which autophagy acts as a negative regulator of autoimmune processes. In particular, autophagy affects immune cell survival and differentiation. This is evidenced by the fact that Atg5 dysfunction in B lymphocytes leads to impaired differentiation of pro-B cells into pre-B cells. B cells with mutant inactive Atg5 are less viable than cells with wild-type Atg5. Autophagy can also affect the BCR signaling pathway required for B cell activation. Apoptotic B cells with an activated BCR signaling pathway are characterized by an abnormal increase in autophagosome formation, which leads to cell death. This evidence indicates that autophagy can be involved in the inhibition of autoreactive B cells [36, 37].

Autophagy is one of the processes mediating T cell viability. Inhibited formation of components of the autophagy initiator complex PI3KC3-C1 in T cells is known to result in impaired removal of damaged organelles, impaired differentiation, and all-out death [38]. In addition, T cell survival is reduced in Atg7, Atg5, and Atg3 deficiency.

2.2.1. Crohn disease and the role of autophagy in its pathogenesis. Similar to many autoimmune diseases, the pathogenesis of Crohn disease, which is manifested by chronic inflammation of the large intestine, has not been fully elucidated. The intestinal mucosa in Crohn disease resembles a cobblestone sidewalk and has characteristic thickened areas. Symptoms are similar to dyspeptic disorders; they include abdominal pain, diarrhea, anorexia, nausea, vomiting, and weight loss. The inflammation area can spread to the entire gastrointestinal tract, up to the oral mucosa. The absorption of nutrients in the intestine is impaired in disease. The molecular mechanism of Crohn disease has not been revealed; therefore, there are no effective ways to treat it [39].

To date, there are several hypotheses on the mechanism of Crohn disease onset and progression. According to one of them, mutations in the gene encoding the NOD2 receptor play a key role.

NOD2 is a cytosolic receptor protein located on the inner side of the cytoplasmic membrane. It is involved in the antibacterial immune response. NOD2 contains three distinct domains: NOD, LRR, and CARD. Muramyl dipeptide (MDP), a component of bacterial cell wall peptidoglycan, is a NOD2-activaing ligand. Receptor activation leads to simultaneous interaction of MDP with the LRR domains of two NOD2 molecules, resulting in their dimerization (*Fig. 4*). This causes NOD2 activation and recruitment of two RIP2 molecules to the CARD domain. An E3 ubiquitin ligase complex containing cIAP1/2 and XIAP associates with RIP2, leading to complex activation and polyubiquitin formation on RIP2. A complex consisting of the TAB1/2 and TAK proteins is formed on polyubiquitin, initiating the assembly of the IKKα/β/γ complex, which participates in the phosphorylation of Iκβ, which, in turn, forms a complex with NF-κβ. This leads to NF-κβ release and activation, followed by its migration to the nucleus [40]. NOD2 also regulates the activity of α- and β-defensins, which form "holes" on the bacterial membrane, eventually leading to cell death.

Mutations in the LRR domain have been found to impair the immune response and increase the chances of survival of intracellular pathogenic bacteria. It ultimately results in increased production of the cytokine IL-23, leading to enhanced chemotaxis of Th17 cells to the intestinal mucosa [41, 42].

Association of Atg16L with NOD2 yields the Agt12/Atg5/Atg16L complex, which is necessary for the formation of an autophagosome surrounding the bacterium and further bacterial lysis. This subtype of autophagy is called xenophagy. If the LRR domain of NOD2 carries a mutation, Atg16L is not recruited to the membrane. This leads to impaired autophagosome formation and promotes the survival of pathogenic bacteria inside the cell [41, 42].

Another protein involved in Crohn disease pathogenesis is the IRGM (immunity-related GTPase family M) protein, which possesses GTPase activity. This protein binds to the MOM by interacting with the cardiolipin on its surface. IFN-γ synthesis in the cell, as well as cell infection with Gram-negative bacte-

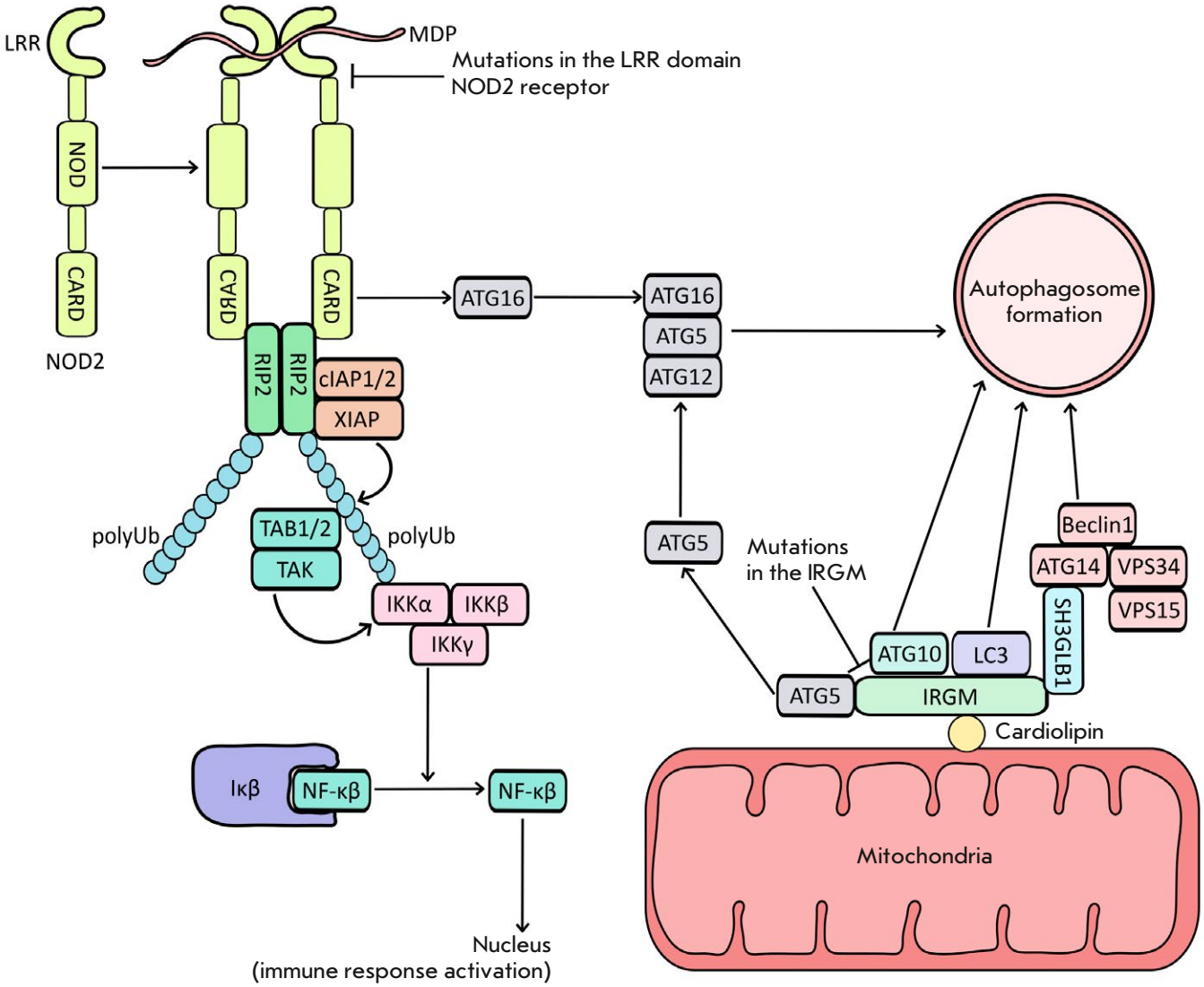


Fig. 4. Crohn disease and autophagy. Schematic presentation of intracellular receptor NOD2 activation and signaling pathways affecting autophagosome assembly and interaction with mitochondria (details can be found in the article)

ria, enhances IRGM activity. IRGM is involved in the regulation of antibacterial immune mechanisms in the cell. Active IRGM initiates autophagy through its interaction with the following autophagosome assembly proteins: Atg5, Atg10, Bif-1, LC3, SH3GLB1, UVRAG, Beclin1, and Vps34 (*Fig. 4*) [43].

Inactivating *IRGM* mutations are known to increase the risk of Crohn disease. Introduction of a deletion in the *IRGM* promoter region and an increase in the amount of microRNA-196 targeting IRGM mRNA were shown to reduce autophagy activity [44].

2.3. Autophagy and cancer

Similar to other diseases, autophagy has a dual effect on cancer development. On the one hand, autophagy serves as one of the sources of nutrients for rapidly dividing cancer cells. On the other hand, it can inhibit cell division and even cause cancer cell death [45].

Hypoxia develops in the tumor due to a lack of adequate blood supply resulting from aggressive, uncontrolled growth of cancer cells. The metabolism of cancer cells is altered; glycolysis and subsequent anaerobic catabolism are activated. HIF-1α plays a key

role in cell adaptation to tissue hypoxia; the HIF-1α transcription factor enhances angiogenesis in tumor, triggers glycolysis, and activates cellular adaptation processes. Disruption of oxygen-dependent proteolysis of HIF-1α in cancer cells impedes its degradation, which results in its accumulation in the cytosol. This, in turn, leads to increased expression of the genes encoding autophagy proteins (Beclin1 and BNIP3) [33, 45].

Depletion of energy reserves in cancer cells leads to the activation of the AMP kinase AMPK, which is induced by ATP deficiency. AMPK is a sensor of the lack of cell energy resources. Activated AMPK phosphorylates Beclin1 and ULK1 at S93, S96, and T388 and at S467, S555, T574, and S637, respectively, leading to their activation. AMPK is also involved in the phosphorylation of mTORc1 complex proteins, causing complex inactivation (*Fig. 2*). These processes are an adaptation to nutrient deficiency; they lead to an increase in autophagy activity and nutrient acquisition by eliminating cancer cell components [46].

The adaptor protein p62 is involved in the autophagosome-mediated degradation of the toxic substances formed during metabolism in cancer cells (*Fig. 1*). A decrease in the p62 level impedes cancer growth. A high p62 level is detected in pancreatic, lung, and liver cancer cells [47].

In addition to the positive effects of autophagy on cancer cell survival, there are also examples of its negative effect. For instance, autophagy can inhibit cancer cell growth and cause cancer cell death through the interaction of Beclin1 with a mutant EGFR tyrosine kinase which is involved in carcinogenesis. The Beclin1–mEGFR interaction inhibits the mitotic activity of the mutant receptor, resulting in suppressed cancer cell growth. In addition, introduction of an inactivating mutation in *BECN1*, which codes for Beclin1, and experimental reduction of its expression lead to enhanced cancer cell growth [48]. A monoallelic deletion in *BECN1* is often found in breast, prostate, and ovarian cancers. This mutation is found in 40–75% of all the above pathologies and is most common in breast cancer. Beclin1 protein deficiency has also been observed in kidney cancer, non-small cell lung cancer, and cholangiocarcinoma. Hyperexpression of the gene encoding Bcl-2, which can form a complex with Beclin1 (Beclin1/Bcl-2) inhibiting autophagy, has been observed in various lymphomas [49].

2.3.1. Follicular lymphoma. Lymphoma is a lymphatic system disease; its development is caused by uncontrolled growth of lymphocytes in the major immune organs and lymph nodes. One variant of follicular lymphoma is non-Hodgkin lymphoma [50]. This disease is characterized by slow progression. Symptoms are observed at late disease stages and include enlargement of lymph nodes in the groin region, neck, and armpits, as well as back pain and intoxication. Displacement of immunocompetent cells and development of immunodeficiency take place during disease progression.

The pathogenesis of follicular lymphoma is associated with the t(14;18)(q32;q21) chromosomal translocation, which is characterized by a rearrangement between the chromosome 18 region encoding the anti-apoptotic protein Bcl-2 and the chromosome 14 region encoding the enhancer region of the immunoglobulin heavy chain. This translocation results in a fusion gene expressing Bcl-2 at an abnormally high rate. As noted above, Bcl-2 accumulation leads to excessive recruitment of Bcl-2 to Beclin1, BNIP3, and other autophagy-associated proteins. This results in autophagy inhibition in cells carrying the mutation [49]. One of the main characteristics of the Bcl-2 protein is its anti-apoptotic activity. Accumulation of this protein leads to a decrease in the apoptosis of transformed immature B cells and an enlargement of their pool. Another mutation, which is often detected in lymphomas, is associated with the *Bcl-6* gene: t(3;14)(q27;q32). In this mutation, a rearrangement of fragments between chromosomes 3 and 14 takes place, which results in a gene sequence encoding mutant Bcl-6 that fails to function properly; i.e., it performs normal differentiation of B cells [51].

In addition, the p62 and LC3II levels are decreased in follicular lymphoma cells, resulting in autophagy inhibition and autophagy-associated cell death [52].

CONCLUSION

Autophagy plays an important role in cells. Autophagy impairment is associated with the development of various diseases, while autophagy activity can affect the course of various diseases in different ways. It should be noted that, despite the active study of the role of autophagy in various cell processes and diseases, the contribution of the individual signaling pathways related to autophagy remains poorly understood and is of great interest. ●

Data collection and writing of the manuscript was supported by the Russian Science Foundation (grant No. 21-14-00355). The work on preparing the manuscript for publication was supported by the Russian Foundation for Basic Research (grant No. 20-54-76005).

REFERENCES

1. Cao W., Li J., Yang K., Cao D. // Bull. Cancer. 2021. V. 108. № 3. P. 304–322. https://doi.org/10.1016/j.bulcan.2020.11.004.
2. Xiong J. // Protein & Cell. 2015. V. 6. № 10. P. 722–734. https://doi.org/10.1007/s13238-015-0195-8.
3. Xie K., Liang C., Li Q., Yan C., Wang C., Gu Y., Zhu M., Du F., Wang H., et al. // Cancer. 2016. V. 139. № 7. P. 1564–1573. https://doi.org/10.1002/ijc.30205.
4. Kuma A., Matsui M., Mizushima N. // Autophagy. 2007. V. 3. № 4. P. 323–328. https://doi.org/10.4161/auto.4012.
5. Schaaf M.B.E., Keulers T.G., Vooijs M.A., Rouschop K.M.A. // FASEB. J. 2016. V. 30. № 12. P. 3961–3978. https://doi.org/10.1096/fj.201600698R.
6. Gómez-Sánchez R., Rose J., Guimarães R., Mari M., Papinski D., Rieter E., Geerts W.J., Hardenberg R., Kraft C., Ungermann C., et al. // J. Cell. Biol. 2018. V. 217. № 8. P. 2743–2763. https://doi.org/10.1083/jcb.201710116.
7. Katsuragi Y., Ichimura Y., Komatsu M. // FEBS J. 2015. V. 282. № 24. P. 4672–4678. https://doi.org/10.1111/febs.13540.
8. McEwan D.G., Popovic D., Gubas A., Terawaki S., Suzuki H., Stadel D., Coxon F.P., Miranda de Stegmann D., Bhogaraju S., et al. // Mol Cell. 2015. V. 57. № 1. P. 39–54. doi: 10.1016/j.molcel.2014.11.006.
9. Dibble C.C., Cantley L.C. // Trends Cell Biol. 2015. V. 25. № 9. P. 545–555. doi: 10.1016/j.tcb.2015.06.002.
10. Decuypere J.P., Bultynck G., Parys J.B. // Cell Calcium. 2011. V. 50. № 3. P. 242–250. https://doi.org/10.1016/j.ceca.2011.04.001.
11. Tasdemir E., Maiuri M.C., Orhon I., Kepp O., Morselli E., Criollo A., Kroemer G. // Cell Cycle. 2008. V. 7. № 19. P. 3006–3011. https://doi.org/10.4161/cc.7.19.6702.
12. Mrakovcic M., Fröhlich L.F. // Biomolecules. 2018. V. 8. № 2. P. 14. https://doi.org/10.3390/biom8020014.
13. Bialik S., Dasari S.K., Kimchi A. // J. Cell Sci. 2018. V. 131. № 18. https://doi.org/10.1242/jcs.215152.
14. Majeski A.E., Dice J.F. // Internat. J. Biochem. Cell Biol. 2004. V. 36. № 12. P. 2435–2444. https://doi.org/10.1016/j.biocel.2004.02.013.
15. Banerjee R., Beal M.F., Thomas B. // Trends Neurosci. 2010. V. 33. № 12. P. 541–549. https://doi.org/10.1016/j.tins.2010.09.001.
16. Dou C., Zhang Y., Zhang L., Qin C. // Animal Models Exp. Med. 2022. V. 6. № 1. P. 10–17. https://doi.org/10.1002/ame2.12229.
17. Moumné L., Betuing S., Caboche J. // Front. Neurol. 2013. V. 4. P. 127. https://doi.org/10.3389/fneur.2013.00127.
18. Saudou F., Humbert S. // Neuron. 2016. V. 89. № 5. P. 910–926. https://doi.org/10.1016/j.neuron.2016.02.003.
19. Croce K.R., Yamamoto A. // Neurobiol. Disease. 2019. V. 122. P. 16–22. https://doi.org/10.1016/j.nbd.2018.08.010.
20. Martin D.D.O., Ladha S., Ehrnhoefer D.E., Hayden M.R. // Trends Neurosci. 2015. V. 38. № 1. P. 26–35. https://doi.org/10.1016/j.tins.2014.09.003.
21. Piller C. // Science. 2022. V. 377. № 6604. P. 358–363. https://doi.org/10.1126/science.add9993.
22. Vetrivel K.S., Thinakaran G. // BBA - Mol. Cell. Biol. Lipids. 2010. V. 1801. № 8. P. 860–867. https://doi.org/10.1016/j.bbalip.2010.03.007.
23. Gratuze M., Leyns C.E.G., Holtzman D.M. // Mol. Neurodegeneration. 2018. V. 13. № 1. P. 66. https://doi.org/10.1186/s13024-018-0298-9.
24. Iqbal K., Liu F., Gong C.-X., Grundke-Iqbal I. // Curr. Alzheimer Res. 2010. V. 7. № 8. P. 656–664. https://doi.org/10.2174/156720510793611592.
25. Chi H., Sang T.-K., Chang H.-Y. // Cognitive Disorders, IntechOpen. 2019. V. 8. http://dx.doi.org/10.5772/intechopen.73198.
26. Monzio Compagnoni G., Di Fonzo A., Corti S., Comi G.P., Bresolin N., Masliah E. // Mol. Neurobiol. 2020. V. 57. № 7. P. 2959–2980. https://doi.org/10.1007/s12035-020-01926-1.
27. Swerdlow R.H. // J. Alzheimer's Dis. 2018. V. 62. № 3. P. 1403–1416. https://doi.org/10.3233/JAD-170585.
28. Nakamura T., Cieplak P., Cho D.-H., Godzik A., Lipton S.A. // Mitochondrion. 2010. V. 10. № 5. P. 573–578. https://doi.org/10.1016/j.mito.2010.04.007.
29. Losón O.C., Song Z., Chen H., Chan D.C. // Mol. Biol. Cell. 2013. V. 24. № 5. P. 659–667. https://doi.org/10.1091/mbc.E12-10-0721.
30. Babst M., Katzmann D.J., Estepa-Sabal E.J., Meerloo T., Emr S.D. // Dev. Cell. 2002. V. 3. № 2. P. 271–282. https://doi.org/10.1016/s1534-5807(02)00220-4.
31. Uddin M.S., Stachowiak A., Mamun A.A., Tzvetkov N.T., Takeda S., Atanasov A.G., Bergantin L.B., Abdel-Daim M.M., Stankiewicz A.M. // Front. Aging Neurosci. 2018. V. 10. № 4. https://doi.org/10.3389/fnagi.2018.00004.
32. Wang Y., Mandal A.K., Son Y.O., Pratheeshkumar P., Wise J.T.F., Wang L., Zhang Z., Shi X., Chen Z. // Toxicol. Appl. Pharmacol. 2018. V. 353. P. 23–30. https://doi.org/10.1016/j.taap.2018.06.003.
33. Choi H., Merceron C., Mangiavini L., Seifert E.L., Schipani E., Shapiro I.M. // Autophagy. 2016. V. 12. № 9. P. 1631–1646. https://doi.org/10.1080/15548627.2016.1192753.
34. Dominguez-Villar M., Hafler D.A. // Nat. Immunol. 2018. V. 19. P. 665–673. https://doi.org/10.1038/s41590-018-0120-4.
35. Yang Z., Goronzy J.J., Weyand C.M. // J. Mol. Med. 2015. V. 93. № 7. P. 707–717. https://doi.org/10.1007/s00109-015-1297-8.
36. Pierdominici M., Vomero M., Barbati C., Colasanti T., Maselli A., Vacirca D., Giovannetti A., Malorni W., Ortona E. // FASEB J. 2012. V. 26. № 4. P. 1400–1412. https://doi.org/10.1096/fj.11-194175.
37. Miller B.C., Zhao Z., Stephenson L.M., Cadwell K., Pua H.H., Lee H.K., Mizushima N., Iwasaki A., He Y.-W., Swat W., et al. // Autophagy. 2008. V. 4. P. 309–314. https://doi.org/10.4161/auto.5474.
38. Yang G., van Kaer L. // Cell Death Dis. 2020. V. 11. № 5. P. 334. https://doi.org/10.1038/s41419-020-2568-z.
39. Torres J., Mehandru S., Colombel J.F., Peyrin-Biroulet L. // Lancet. 2017. V. 389. № 10080. P. 1741–1755. https://doi.org/10.1016/S0140-6736(16)31711-1.
40. Al Nabhani Z., Dietrich G., Hugot J.P., Barreau F. // PLoS Pathogens. 2017. V. 13. № 3. P. e1006177. https://doi.org/10.1371/journal.ppat.1006177.
41. Boyapati R., Satsangi J., Ho G.T. // F1000Prime Rep. 2015. V. 7. P. 44. https://doi.org/10.12703/P7-44.
42. Travassos L., Carneiro L., Ramjeet M. Hussey S., Kim Y.G., Magalhães J.G., Yuan L., Soares F., et al. // Nat. Immunol. 2010. V. 11. P. 55–62. https://doi.org/10.1038/ni.1823.
43. Petkova D.S., Viret C., Faure M. // Front. Immunol. 2013. V. 3. P. 426. https://doi.org/10.3389/fimmu.2012.00426.
44. Carrière J., Darfeuille-Michaud A., Nguyen H.T. // World J. Gastroenterol. 2014. V. 20. № 34. P. 12102–12117. https://doi.org/10.3748/wjg.v20.i34.12102.
45. Yun C.W., Lee S.H. // Int. J. Mol. Sci. 2018. V. 19. № 11. P. 3466. https://doi.org/10.3390/ijms19113466.
46. Vega-Rubín-de-Celis S. // Biology. 2020. V. 9. № 1. P. 4.

https://doi.org/10.3390/biology9010004.
47. Amaravadi R.K., Kimmelman A.C., Debnath J. // Cancer Discov. 2019. V. 9. № 9. P. 1167–1181. https://doi.org/10.1158/2159-8290.CD-19-0292.
48. Rohatgi R.A., Janusis J., Leonard D., Bellve K.D., Fogarty K.E., Baehrecke E.H., Corvera S., Shaw L.M. // Oncogene. 2015. V. 34. № 42. P. 5352–5362. https://doi.org/10.1038/onc.2014.454.
49. Pattingre S., Tassa A., Qu X., Garuti R., Liang X.H., Mizushima N., Packer M., Schneider M.D., Levine B. // Cell. 2005. V. 122. № 6. P. 927–939. https://doi.org/10.1016/j.cell.2005.07.002.
50. Carbone A., Roulland S., Gloghini A., Younes A., von Keudell G., Lopez-Guillermo A., Fitzgibbon J. // Nat. Rev. Dis. Primers. 2019. V. 5. № 1. P. 83. https://doi.org/10.1038/s41572-019-0132-x.
51. Ott G., Rosenwald A. // Haematologica. 2008. V. 93. № 12. P. 1773–1776. https://doi.org/10.3324/haematol.2008.001495.
52. McCarthy A., Marzec J., Clear A., Petty R.D., Coutinho R., Matthews J., Wilson A., Iqbal S., Calaminici M., Gribben J.G., Jia L. // Oncotarget. 2014. V. 5. №. 22. P. 11653–11668. https://doi.org/10.18632/oncotarget.2605.

The Rurikids: The First Experience of Reconstructing the Genetic Portrait of the Ruling Family of Medieval Rus' Based on Paleogenomic Data

K. V. Zhur[1], F. S. Sharko[1], Vl. V. Sedov[2], M. V. Dobrovolskaya[2], V. G. Volkov[3], N. G. Maksimov[4], A. N. Seslavine[5], N. A. Makarov[2], E. B. Prokhortchouk[1*]

[1]Federal Research Centre "Fundamentals of Biotechnology" of the Russian Academy of Sciences, Moscow, 119071 Russian Federation

[2]Institute of Archeology, Russian Academy of Sciences, Moscow, 117292 Russian Federation

[3]Regional State Autonomous Institution "Center of Tatar Culture", Tomsk, 634050 Russian Federation

[4]ANO "Runiverse", Moscow, 119071 Russian Federation

[5]Russian Public Organisation "RDS", Moscow, 109028 Russian Federation

*Email: prokhortchouk@gmail.com

Received: June 20, 2023; in final form, August 17, 2023

DOI: 10.32607/actanaturae.23425

Copyright © 2023 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT The Rurikids were the reigning house of Rus', its principalities and, ultimately the Tsardom of Russia, for seven centuries: from the IX to the end of the XVI century. According to the Primary Chronicle (the Tale of Bygone Years), the main chronicle of Rus', the Rurik dynasty was founded by the Varangian prince Rurik, invited to reign in Novgorod in 862, but still there is no direct genetic evidence of the origin of the early Rurikids. This research, for the first time, provides a genome-wide paleogenetic analysis of bone remains belonging to one of the Rurikids, Prince Dmitry Alexandrovich (?–1294), the son of the Grand Prince of Vladimir Alexander Yaroslavich Nevsky (1221–1263). It has been established that his Y chromosome belongs to the N1a haplogroup. Most of the modern Rurikids, according to their genealogies, belonging to the N1a haplogroup, have the most similar variants of Y chromosomes to each other, as well as to the Y chromosome of Prince Dmitry Alexandrovich. Genome-wide data of the medieval and modern Rurikids unequivocally indicates that they belong to the N1a haplogroup of the Y chromosome, starting at least from the XI century (since the time of Prince Yaroslav the Wise). All the other alleged Rurikids, both ancient and modern, being carriers of other haplogroups (R1a, I2a), possess high heterogeneity of the sequence of Y chromosomes, meaning that we cannot confirm their common ancestry. The most probable ancestors of Prince Dmitry Alexandrovich in the male line were the men who left the burial ground Bolshoy Oleny Island on the coast of the Kola Peninsula about 3,600 years ago. The reconstruction of the genome of Prince Dmitry Alexandrovich indicates the contribution of three ancestral components to his origin: (1) the early medieval population of the east of Scandinavia from the island of Oland, (2) representatives of the steppe nomadic peoples of the Eurasian steppes of the Iron Age or the early medieval population of central Europe (steppe nomads from the territory of Hungary), and (3) the ancient East-Eurasian component. Reliable statistics were also obtained when the Scandinavians were replaced with the Medieval Russian Slavic populations of the XI century. Thus, for the first time, we have shown the complex nature of interethnic interactions in the formation of the nobility of medieval Rus' on the example of the ancient Rurikid.

KEYWORDS the Rurikids, Prince Dmitry Alexandrovich, whole genome sequencing, N1a-haplogroup.

ABBREVIATIONS aDNA – ancient DNA; SNP – single-nucleotide polymorphism; PCA – principal component analysis.

INTRODUCTION

The application of paleogenetic methods when studying the genetic identity and origin of the medieval Russian nobility is one of the most productive among many modern scientific approaches capable of expanding and verifying the existing knowledge about the Medieval Russian society, its ethnic composition, and political organization. Meanwhile, the remains of the Rurikids, the most ancient reigning family whose members were the major actors in the history of Russia in the IX–XVI centuries, remain almost untouched by paleogenetic researches. The XI–XII centuries Rurikids' haplogroups were reconstructed based on genetic materials of modern individuals whose genealogy, according to historical data, ascends to Rurik with different degrees of reliability [1]. The accuracy in the selection of these genetic materials and the possibility of verification of the historical and genealogical information on the basis of which the selection was made remain debatable and are discussed by the authors – experts of absolutely different qualifications and fields [2, 3].

The existence of a "blind spot" in the study of the genomes of the Medieval Russian nobility in many aspects is due to the complexity of personally identifying the remains of the Rurikids and other aristocratic families in the necropolises of the X–XIV centuries. It is well known that the names of the buried were not indicated in any way on funerary structures, sarcophagi, or tombstones until the beginning of the XV century. The location of princely burial grounds is established by annalistic messages, synodics of the XVI–XVII centuries, taking into account the later tradition of church veneration of many representatives of the princely family. Archaeological research of burial grounds in Medieval Russian churches and the anthropological study of bones are the main ways of identifying noble burials, but the conditions of necropolises do not always allow for such identification. The long use of necropolises, the practice of placing new burials over old ones, moving the revered remains during their examination in the XV–XIX centuries, and, finally, the removal of relics in the course of the anti-religion campaigns in Soviet times led to part of the princely remains of the XI–XIV centuries from the burial places in the Medieval Russian churches getting lost, or the remains not being reliably matched with certain historical persons whose graves were disposed in these necropolises. One example of the use of bone remains in order to conduct the genetic analysis of the Rurikids whose membership in the princely ruling family cannot be verified by archaeological data is the study of the presumed remains of Prince Gleb Svyatoslavich from the Chernigov Transfiguration Cathedral – a skull found during repair works at the temple without archaeological documentation [4].

Thus, the few burial places of the Rurikids with bone remains that can be reliably attributed to the princely family, based on the archaeological data, anthropological definitions, and a set of historical evidence acquire special significance. To such trustworthy burial places belongs the burial site of prince Dmitry Alexandrovich, found in the southern altar apse in the south-eastern part of the Transfiguration cathedral in Pereslavl-Zalessky (Supplementary 1).

EXPERIMENTAL PART

DNA isolation and genomic library preparation

All experiments with aDNA were carried out in "a clean room" – a room specially equipped for these purposes at the Federal Research Center "Fundamentals of Biotechnology" of the Russian Academy of Sciences (Skryabin Institute of Bioengineering).

DNA was isolated from the bone remains found in the ruined sarcophagus of the Transfiguration cathedral in Pereslavl-Zalessky. According to the historical information on the burial place, archeological data and anthropological definitions, these remains belong to the son of Prince Alexander Yaroslavich Nevsky – Prince Dmitry Alexandrovich (Supplementary 1). The remains are characterized by good preservation of bone tissue, which is typical for remains found long after burial out of contact with the ground, suggesting a rather late episode of destruction of the sarcophagus. To isolate aDNA from the samples provided for genetic analysis, we obtained three portions of bone powder weighing 20, 50, and 80 mg from the metacarpal bone of the hand, rotula, and navicular bone of the foot, respectively. DNA was isolated by magnetic separation using buffer D – that of the Dabney method (5 M guanidine hydrochloride, 40% (v/v), isopropanol, 0.12 M sodium acetate, and 0.05% (v/v) Tween 20) and silica-coated magnetic beads [5].

The resulting DNA was used to prepare libraries of single-stranded DNA fragments using the ACCEL-NGS 1S Plus DNA Library Kit (Swift Biosciences, USA) according to the original protocol but with minor modifications: for the steps providing strand elongation and sample indexing, uracil-tolerant polymerase (KAPA HiFi HS Uracil+RM, USA) was used. To assess the content of endogenous DNA, test sequencing of the constructed libraries of low-coverage DNA fragments was carried out, approximately 3–4 million single reads per sample (50 bp

long). For the sample with the best preservation of the genetic material (high endogeneity and the presence of C > T substitutions at the 5' ends of DNA fragments), an additional library was prepared from the same DNA extract and pre-treated with a mixture of uracil-DNA glycosylase (UDG) and endonuclease VIII [6]. The mixture of enzymes made it possible to remove uracil from the aDNA strands and turn the resulting abasic sites into single nucleotide breaks, while some of the uracils at the ends of the fragments were preserved, which is associated with the low efficiency of enzymes in these regions. The removal of uracils improved the quality of mapping and prevented a distortion of the results of the subsequent statistical processing [7].

The MyBaits Expert Human Affinities Prime Plus Kit (Daicel Arbor Biosciences) was used for subsequent enrichment of the genome regions of interest. Biotinylated single-stranded DNA probes from the kit cover single nucleotide polymorphisms (SNPs) from the panel "1240K capture" [8], 46,000 additional unique SNPs of the Y chromosome of known haplogroups according to the classifier of the International Society of Genetic Genealogy (ISOGG) [9], and a set of MitoTrio probes for three different mitochondrial genomes: the Revised Cambridge Reference Sequence (rCRS), the Reconstructed Sapiens Reference Sequence (RSRS), and the Vindija Neanderthal sequence (Genbank NC_011137) [10]. Libraries were sequenced on a HiSeq 1500 instrument (Illumina, USA) in paired read mode 2 × 150 bp for genome-wide sequencing and in the mode of single readings 50 bp long for test libraries.

Bioinformatics analysis

To remove contaminating DNA reads from the sequencing data, we used the BBDuk software [11] included in the BBMap package, and bacteria, fungi, plants, viruses, and other organism databases. The output of the BBDuk tool was analyzed using the PALEOMIX pipe-line (version 1.2.14) [12]. Sequencing adapters were trimmed using the Cutadapt v3.4 tool [13]. Sequences were aligned to the reference human genome sequence (hg19/GRCh37) using BWA (version 0.7.17) [14].

Aligned reads were filtered to ensure a minimum display quality of 20 using samtools view (version 1.9) [15]. Indexing, sorting, and removal of duplicates (rmdup) were performed using the samtools tool (version 1.9) [15]. To call genotypes from aligned reads, a PileupCaller (https://github.com/stschiff/sequenceTools) with the "–randomHaploid" mode was used, which calls haploid genotypes by randomly selecting one high-quality base (phred base quality score ⩾ 30) on the 1240K SNP panel (https://reich.hms.harvard.edu/).

Postmortem DNA damage patterns were analyzed using the MapDamage2 software [16], which offers a series of tools for imaging and modeling postmortem damage patterns observed in ancient samples. MapDamage2.0 also makes it possible to recalculate base quality scores in order to mitigate the impact of postmortem damage on further analysis.

To determine the genetic clustering of the NEV2.3 sample among the ancient samples known at the time of the study presented in the Allen Ancient DNA Resource (AADR) panel [17], the ADMIXTURE v.1.3.0 software [18] was used. SNPs were trimmed for sites with linkage disequilibrium using PLINK v1.9 [19]. The sliding window was 50 SNPs; the step was 5 SNPs; the r2 threshold was 0.2 (–in-dep-pairwise 50 5 0.2). There were 10 runs with random starting values for a number of clusters (K) in the range of 4–12; the run with the lowest cross-validation error was selected to plot the graph of population admixture.

For principal component analysis (PCA), the smartpca tool from the EIGENSOFT package was used. Ancient samples were projected onto the first two components of the modern samples. A list of samples is presented in *Table 1* of Supplementary 2. The following parameters were set by default: lsqproject: YES, numoutlieriter; 0, shrinkmode; and YES for the smartpca analysis. Mitochondrial haplotypes were determined using the HaploGrep program [20]. Determination of Y chromosome haplogroups was carried out by comparing alleles on the phylogenetic tree ISOGG version 15.73. F4-statistics were calculated using the qpDstat program from the ADMIXTOOLS software package with default parameters. All constructions were based on available data obtained from whole genome sequencing of the samples. To model the genome from the components of ancestral populations, we used the qpWave and qpAdmix programs with the "allsnps: YES" parameter and "Russia_Yana_UP," "Russia_Sunghir," "Bichon_LP," "Zagros_EN," "Russia_DevilsCave_N," "Alaska_LP," "Russia_Ust_Ishim.DG," "Papuan.DG," "Han.DG," "Chukchi.DG," "Russia_Kostenki14," "ONG," "Yoruba.SDG," "Mbuti.SDG," and "Karitiana.SDG" were chosen as the right populations.

The HIrisPlex-S online tool [21–23] was used to predict eye, hair, and skin color.

DNA isolation and genomic library preparation from a modern human blood sample

To avoid the bias caused by the method of library preparation, we sequenced a sample of a modern

Rurikid (sample with an identifier Olgovich3), according to his genealogy. Informed consent to participate in the genetic research was obtained from the modern Rurikid before the start of any research procedures.

Genomic DNA was isolated from 200 μl of blood using a Magen DNA blood mini kit according to the protocol. Overall, 1 μg of genomic DNA was used to fragment the sample to an average DNA fragment size of 200 nucleotides on a Covaris platform. The library for subsequent whole genome sequencing was prepared according to the instructions in the NEBNEXT DNA UltraII (NEB) kit. Sequencing was carried out on a HiSeq1500 platform (Illuimina, USA). In total, 115,564,028 reads of 150 nucleotides were generated. Mapping was performed using BWA (v0.7.17) on the hg19/GRCh37 reference human genome, followed by the removal of PCR duplicates. For further identification of SNPs, we used 104,109,318 reads generated on the bcftools software.

Genetic testing of the alleged modern Rurikids was carried out in the commercial laboratory FamilyTreeDNA in Houston (USA), in the laboratory of human population genetics at the Research Centre of Medical Genetics named after the Academician N.P. Bochkov (Moscow) and in the laboratory of evolutionary genetics at the Research institute of medical genetics (Tomsk). The results of the genetic testing were provided by Seslavin A.N., Volkov V.G. and Maksimov N.G., the leaders of the international research project "Rurikovichi. The genome of Russian princes." Some of the results were presented as bam files (Olgovich1, Yurievich1, Mstislavich1, Mstislavich2, Yurievich2, Olgovich4, Mstislavich3, Mstislavich4, Yurievich3, Mstislavich5), part as a list of SNPs of the Y chromosome (Olgovich2 and Olgovich5). All participants donated their genetic data to be used in this project.

RESULTS AND DISCUSSION

Discovery of the remains of Prince Dmitry Alexandrovich in the Transfiguration Cathedral in Pereslavl-Zalessky

The architectural and archaeological team of the Institute of Archeology of the Russian Academy of Sciences, led by V.V. Sedov, examined the alleged burial site of Prince Dmitry Alexandrovich in Pereslavl-Zalessky. Prince Dmitry Alexandrovich (?–1294) was the second son of Prince Alexander Yaroslavich, who inherited the Principality of Pereslavl after his father's death (1263). At different times Prince Dmitry possessed the Novgorod and the great Vladimir principalities. He died on Volok Lamsky on his way to Pereslavl from Tver, and he was buried in Pereslavl (Complete collection of Russian chronicles, vol. 1, 282 p.; Complete collection of Russian chronicles, vol. 3, 328 p.). At the same time, a number of chronicles (including the 4th Novgorod chronicle, the Moscow chronicle of the late XV century, Voskresenskaya and Nikonovskaya chronicles) contain direct references to the fact that he was buried in the Transfiguration Cathedral of Pereyaslavl-Zalessky (Complete collection of Russian chronicles, vol. IV, 249 p.; Complete collection of Russian chronicles, vol. XXV, 157 p.; Complete collection of Russian chronicles, vol. VII, 181 p.; Complete collection of Russian chronicles, vol. XIII, 170 p.) [24–28]. The identification of the remains from the sarcophagus in the southwestern part of the cathedral as the burial place of Prince Dmitry Alexandrovich is based on a combination of historical information on the burial, archaeological data, and anthropological definitions (Supplementary 1).

Paleogenetic analysis of bone remains from the sarcophagus

To isolate aDNA, bone samples (metacarpal, patella, and navicular bone of the foot) of an adult individual, presumably Prince Dmitry Alexandrovich (identification number Nev2), were collected. Bone powder weights were obtained from the samples with the corresponding identification numbers Nev2.1, Nev2.2, and Nev2.3, from which DNA was isolated and libraries of single-stranded fragments were prepared for the initial shotgun sequencing in order to assess endogeneity. Sequencing results are presented in Supplementary 3.

The Nev2.3 sample was characterized by the highest proportion of endogenous DNA and the frequency of cytosine to thymine substitutions at the 5' ends of aDNA fragments and was selected for further in-solution enrichment and sequencing. The frequency of C to T substitutions around the 5′ ends of the sequences are presented in the figure in Supplementary 4.

As a result of genome-wide sequencing of the library of Nev2.3 aDNA fragments, more than 15 million reads were generated and 532,154 single nucleotide polymorphisms (SNPs) were identified. The data suggested that the genome belonged to a man, his mitochondrial haplogroup was F1b, and the Y chromosome haplogroup was N1a (*Table 1*). As a result of assessing the level of contamination of the sample according to such parameters as the degree of heterozygosity of mtDNA and X chromosome (*Tables 1–2* of Supplementary 5), no contamination of the sample was detected.

Table 1. Characteristics of the Nev2.3 sample sequencing

ID of the library	Original number of reads	Number of reads after filtering	Number of reads mapped on hg19	After removal of PCR duplicates	Coverage	Endogenous of DNA, %	SNP (for analysis)	Genetic gender	mtDNA haplogroup	Y chromosome haplogroup
NEV_2.3	15001647	14976811	14299210	3025176	0.06	20.2	532154	M	F1b1	N1a1a1a1a1a7a~

Table 2. Results of the prediction of the phenotypic traits of sample NEV2.3

Trait	Probability (*P*)
Brown eye color	0.962
Dark hair color	0.810
Intermediate skin tone	0.635
Brown hair color	0.555
Black hair color	0.355
Fair skin	0.306
Light hair color	0.190
Light blond hair color	0.090
Dark skin	0.053
Intermediate eye color	0.035
Very light skin	0.005
Blue eye color	0.003
Very dark or black skin tone	0.001
Red hair color	0.000

Prediction of the phenotypic traits of Prince Dmitri Alexandrovich based on genetic data

We have investigated the sample NEV2.3, presumably belonging to Prince Dmitry Alexandrovich, son of Prince Alexander Yaroslavich Nevsky, and managed to identify the single-nucleotide polymorphisms that allow us to predict his phenotype with a reasonable probability: hair color, skin color, and eye color. Eye color was most likely to be brown (P = 0.962), hair color was dark (P = 0.810) or brown (P = 0.555), and skin tone was intermediate (0.635), that is, neither light nor dark. Prediction results for the phenotypic properties of sample NEV2.3 are shown in *Table 2*.

Analysis of the Y chromosome sequences of Prince Dmitry Alexandrovich and other alleged later representatives of the Rurik family

There are three possible variants of the Y chromosome haplogroup of Rurik and his descendants – N1a, R1a, and I2a. The hypotheses are based on the results of genetic studies of the alleged 43 modern Rurikids (who are representatives of 32 genera from different branches of the alleged direct descendants of Prince Yaroslav the Wise and the Polotsk Rurikids) and three ancient descendants of Rurik [1, 2, 29–32].

For the phylogenetic positioning of the Y chromosome, we used all the samples available in the Allen Ancient DNA Resource (AADR) database [17] carrying the haplogroup N1a, as well as the results of genotyping of the Y chromosomes of modern Rurikids with a similar haplogroup. The analysis did not aim at establishing a high-resolution haplogroup, but we concentrated on the analysis of all established polymorphisms of the Y chromosome of the sample (51017 SNP, Supplementary 6).

As a result of phylogenetic positioning, the Y chromosome of Prince Dmitry Alexandrovich was clustered together with the Y chromosomes of alleged modern Rurikids (*Fig. 1*) originating from various noble families: Mstislavich2 (M), Mstislavich3 (M), Mstislavich4 (M), Mstislavich5 (M), Yurievich1 (Yu), Yurievich2(Yu), Yurievich3 (Yu), Olgovich2 (O), Olgovich3 (O), and Olgovich5 (O). M stands for the Mstislaviches, heirs of the branch of the princely Monomakh family descended from Prince Mstislav Vladimirovich (1076–1132); Yu, for the Yurieviches, a branch of the Rurikids derived from the great prince of Kiev Yurii Dolgorukii († 1157); and O, for the Olgoviches, the heirs of the middle line of Chernigov princes, the descendants of Oleg "Gorislavich" Svyatoslavich († 1115). Hereinafter we will use the symbols M, O, and Yu (these branches are reproduced in Supplementary 7) in order to attribute the samples to the genealogic branch of Rurikids. The detailed genealogy of the Rurikids is provided in Supplementary 8.

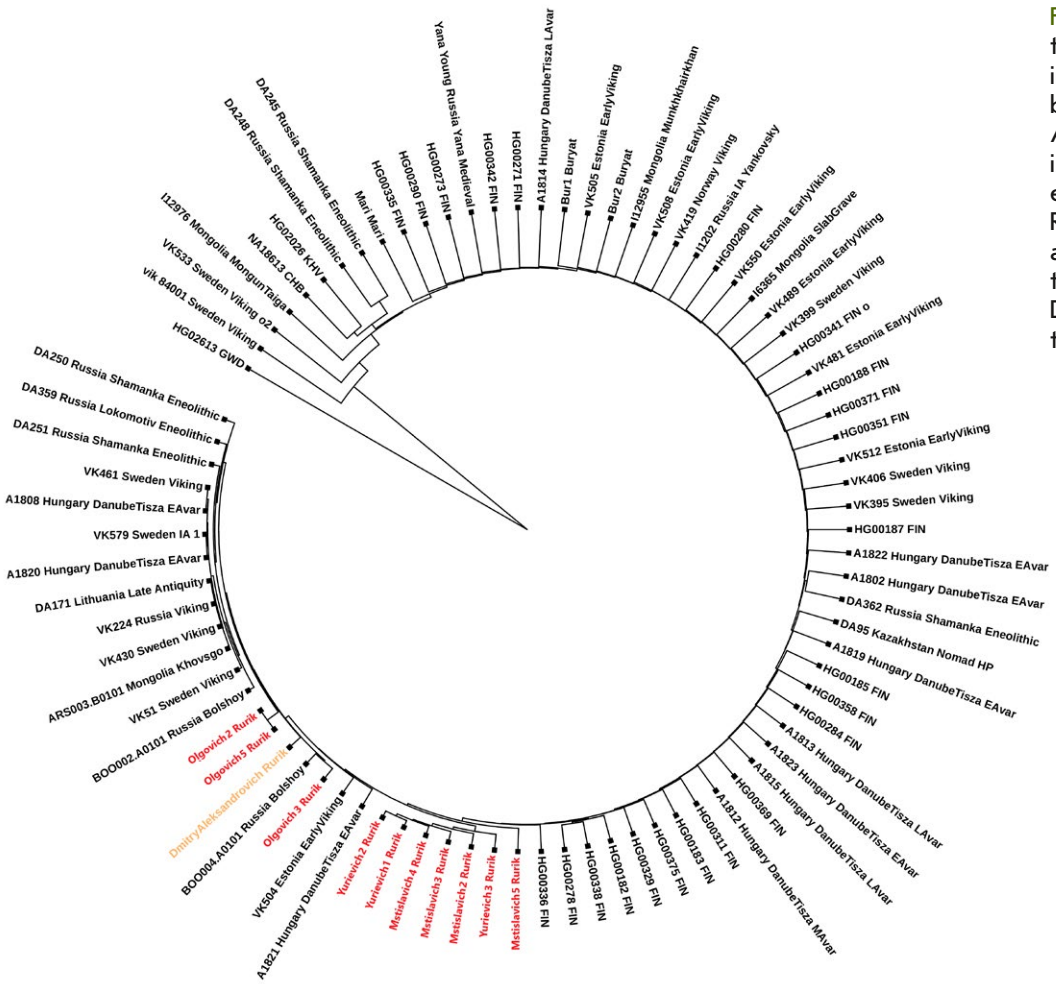


Fig. 1. Phylogenetic tree of the Y chromosome, reflecting the family relationships between Prince Dmitry Alexandrovich (the sample is marked in orange), modern representatives of the Rurikids (marked in red), and samples from the database "The Allen Ancient DNA Resource," belonging to the haplogroup N1a

The three ancient alleged Rurikids, whose Y chromosome haplogroups were previously determined by other scientific groups, include a sample allegedly belonging to Prince Gleb Svyatoslavich of Chernigov (O), published under the identification number VK542 [4], a sample presumably belonging to Prince Izyaslav Ingvarevich Lutsky (M) with the identification number VK541 [4], and a sample belonging to Bela Rostislavovich (O), a large Hungarian feudal lord, a representative of the Chernigov line of the princely family of the Rurikids [33]. The Y chromosomal haplogroups established for these samples are as follows: Prince Gleb – I2a (whole genome sequence); Prince Izyaslav – R1a (whole genome sequence); and Prince Bela – N1a1a1a1a1a1a (according to STR markers). It is important to note that the belonging of the Chernigov and Lutsk burial places to the Rurikids cannot be substantiated by archaeological data, which calls into question the hypotheses that follow from the genetic analysis of these samples.

For carriers of the haplogroup R1a of the modern representatives of the Rurikids and samples from the AADR database, an analysis was carried out using the same algorithm as for carriers of the haplogroup N1a (*Fig. 2*). It turned out that the Y chromosome of the alleged prince Izyaslav Ingvarevich Lutsky, although it belongs to the haplogroup R1a, does not cluster together with the samples of contemporary representatives of the Rurikids: the samples of Mstislavich1 (M), Olgovich1 (O), and Olgovich4 (O). Moreover, Mstislavich1 is clustered separately from Olgovich1 and Olgovich4. We should mention that these samples do not cluster with other "Vikings" with haplogroup R1a whose remains were found in Gnezdovo (VK466), Staraya Ladoga (VK408, VK18), and Kurevanikha (VK160) [4].

Thus, all modern descendants of the legendary Prince Rurik (according to their pedigrees) belonging to the N1a haplogroup and Prince Dmitry Alexandrovich have highly similar Y chromosomes.

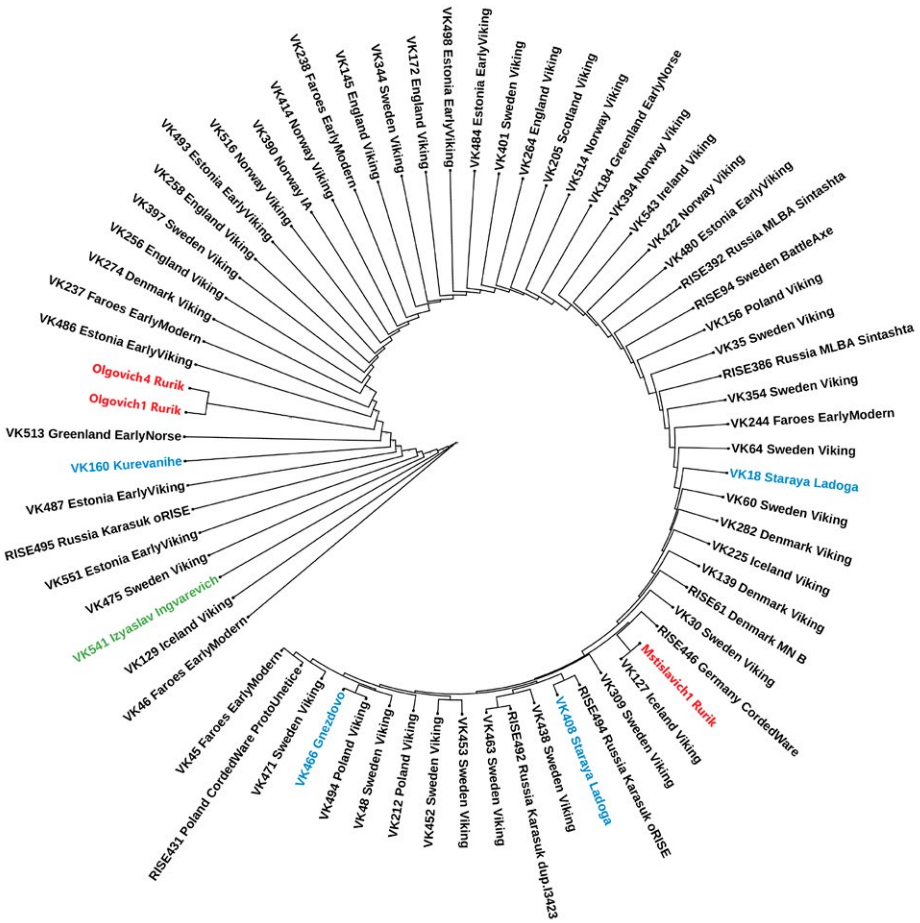


Fig. 2. Phylogenetic tree of the Y chromosome, reflecting the relationship between the alleged modern Rurikids (marked in red) and samples from the database "The Allen Ancient DNA Resource," belonging to the haplogroup R1a. The green color marks the sample of the alleged Prince Izyaslav Ingvarevich Lutsky, the blue color marks the samples of "Vikings" from the territory of modern Russia

The aggregate of genome-wide data on the medieval and modern Rurikids unequivocally indicates that they belong to the N1a haplogroup of the Y chromosome, starting at least from the XIth century (since the time of Prince Yaroslav the Wise). All the other prospective Rurikids, both ancient and modern, being carriers of other haplogroups (R1a, I2a), possess high heterogeneity of the sequence of Y chromosomes; we cannot, therefore, confirm their common ancestry.

Search for archaeological samples with the Y chromosome sequences closest to Prince Dmitry Alexandrovich

The Y chromosome of Prince Dmitry Alexandrovich, in addition to the modern Rurikids, is clustered in the same branch with the ancient people from Bolshoy Oleny Island (Russia_Bolshoy), a burial ground dating back to the middle of the 2nd millennium BC, located in the Kola district of the Murmansk region (*Fig. 1*). Previously, using these samples as an example, the gene flow of the peoples of Siberia (East Eurasian component) to the North and East of Europe was shown [34]. A high degree of homology in the Y chromosome of a representative of the Russian noble family and people of the early metal era led us to the hypothesis of the possible contribution of the East Eurasian gene pool to the formation of the northern European population of the early Middle Ages.

We studied the contribution of the genome of people from Bolshoy Oleny Island to the formation of the medieval population living in the Baltic territories of modern Finland, Denmark, Sweden, and Norway. For this purpose, we used the genomes of the "Vikings" published in the 2020 Margaryan article [4] (we use this term not for historical purposes but for brevity of the reference to the population under study). All of these samples had a VK identifier and a digital code. F4-statistics of the form (VK, Test; Bolshoy Oleny Island, Yoruba) showed negative values ($z > 10$) only when the population of Finns (Finland_Levanluhta) or Saami were used as a test (Finland_Saami_IA.SG), but positive values were encountered when the population of southern Europeans was used as a test: for example, sample Italy_Medieval_EarlyModern.SG ($z > 13$). The results are presented in *Fig. 3* and in *Table 1* of Supplementary 9.

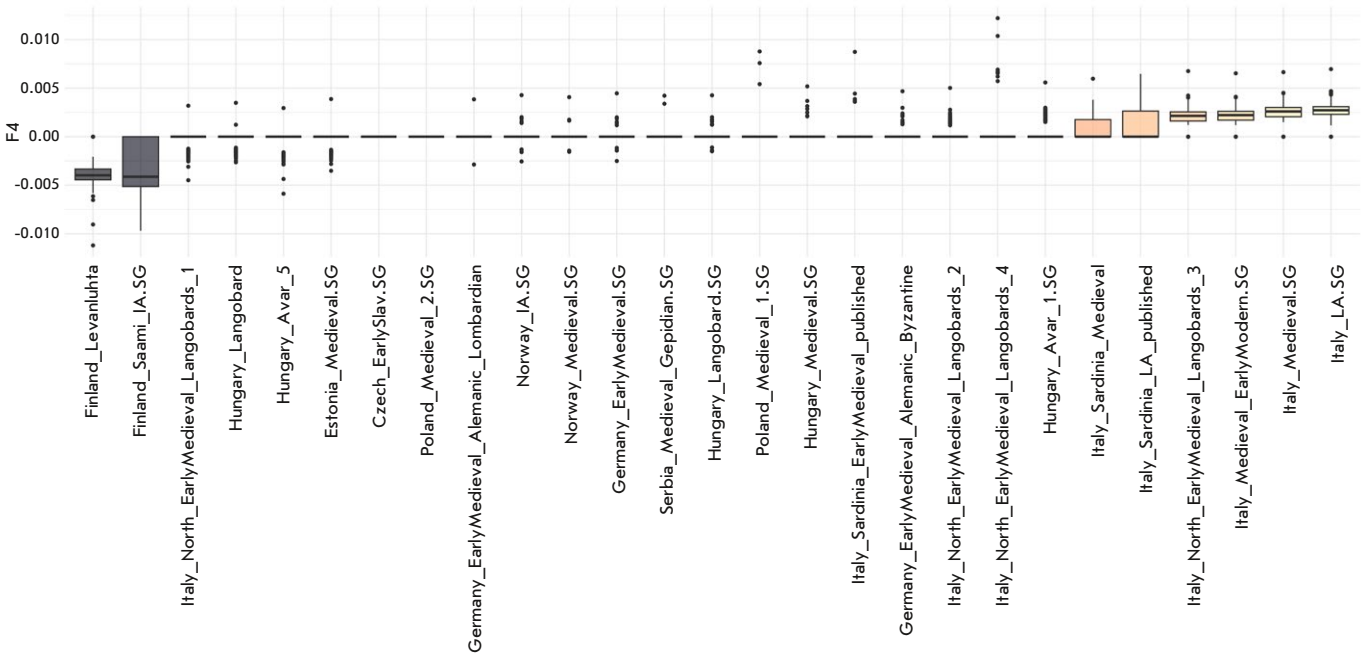


Fig. 3. Values of F4-statistics in the form ("Vikings," Test population; Bolshoy Oleny Island, Yoruba2). Test populations are plotted horizontally. The boxplot displays the F4-statistic values for the "Vikings" group. Statistically significant values are marked in gray for negative values of the F4-statistic and in yellow for positive values

Table 3. Values of F4-statistics in the form (Vikings_R1a, Vikings_N1a; Big Deer Island, Yoruba2)

H1	H2	X	O	D	Z
Viking_R1a	Viking_N1a	BolOlen	Yoruba2	-0.000315	-3.463
Viking_R1a	Viking_N1a	BolOlen	Yoruba	-0.000247	-2.737
Viking_R1a	Viking_N1a	BolOlen	Mbuti	-0.00019	-1.988

Thus, when comparing the VK–Saami pair, genes flow from Bolshoy Oleni Ostrov individuals to the Saami. But when comparing the VK–Southern Europe pair, a significant contribution of Bolshoy-related ancestry in the populations of the "Viking" was detected. Most likely, this gene flow occurred through the contacts of the "Vikings" with the Finno-Ugric population of the Baltic region.

The unexpected similarity of the Y chromosomes of Prince Dmitry Alexandrovich and the ancient people from Bolshoy Oleny Island made it possible to hypothesize that the contribution of the East Eurasian genes might have been significantly higher in the "Vikings" with the N1a haplogroup compared to the "Vikings" with the R1a haplogroup. Indeed, F4-statistics in the form (Vikings_R1a, Vikings_N1a; Big Deer Island, Yoruba2) indicated a significant flow of East Eurasian genes into the "Vikings" with the N1a haplogroup (F4 = -0.00032, Z = -3.46). The results are presented in *Table 3*. At the same time, the genome of Prince Dmitry Alexandrovich did not show a significant difference in terms of the East Eurasian genetic component compared to other "Vikings" with haplogroup N1a.

The hypothesis that the people from Bolshoy Oleny Island are one of the optimal proxy populations for four-ways admixture was tested by repeating the modeling that was performed by Margaryan et al [4]. Some Viking populations, such as Ladoga and Estonia_IA, could not be modeled as a mixture of three ancestral populations: European hunter-gatherers, Neolithic farmers, and steppe pastoralists (*Table 1* of Supplementary 10). To achieve a successful modeling, a fourth source was added, which was represented by the eastern samples of the Xiongnu Iron Age (about 100 BC–50 AD) or samples of the Bolshoy

Oleny Island. It turned out that the Scandinavian populations were modeled equally effectively ($p > 0.05$) using both the Xiongnu [4] and Bolshoy Oleny Island samples (*Table 2* Supplementary 10). Their genetic contribution to these populations was as follows: to Ladoga – 4.7% Xiongnu and 4.7% Bolshoy Oleny Island; and to EstoniaIA – 6.5% Xiongnu and 8.4% Bolshoy Oleny Island.

Thus, it is clear that the gene pool of medieval "Vikings," representing a significant part of Northern Europe (island and mainland), came into being partly through a flow of genes from Siberia, and the male ancestors of Prince Dmitry Alexandrovich were, with a high probability, men who left the Bolshoy Oleny burial ground island on the coast of the Kola Peninsula about 3,600 years ago.

Analysis of mtDNA of Prince Dmitry Alexandrovich

The mitochondrial haplogroup of Prince Dmitry Alexandrovich was identified as F1b1. This haplogroup refers to the East Eurasian cluster, and its representation at different frequencies in the gene pools of most of the previously studied ancient and modern populations of the Baikal region and adjacent territories of Central Asia is noted [35–38]. Also, the mitochondrial haplogroup F was found in three Avars of the VII century in the Danube-Tiss interfluve (F1b1b and two samples with F1b1f). The genomic profiles of these individuals of the middle Avar period correspond to the genomes of other members of the elite of the early Avar period in this region and consist of 90–98% of the ancestral component AR_Xianbei_P_2c, which has an eastern steppe origin and acts as a genetic component of the ancient northeast Asians (ANA). Two of the three burial places (male burial places) were characterized by a rather rich inventory of gold and gilded objects, which indicates their belonging to the nobility [39].

It is rather difficult to interpret the origin of the mitochondrial haplogroup of Prince Dmitry Alexandrovich, since for almost all historical epochs there is an increased variability and "diversity" of the mitochondrial composition of the female part of the groups of the ancient population. This is due to the fact that marriages of an official and unofficial nature concentrated representatives of completely different genetic lines in one geographical locus. When examining the history of dynasties, it is important to keep in mind that the attraction of women of various backgrounds as a beneficial or forced political move is a widespread phenomenon. Thus, the F1b mitochondrial group of Prince Dmitry Alexandrovich can be associated both with the ancient northern flow from the territory of Siberia (East Eurasian component) [34] and with migration of early medieval nomads [39], while the source of this group can probably be the same.

Results of PCA analysis

The principal component analysis (PCA) was used to assess the genetic affinities of the genome of Prince Dmitry Alexandrovich to other known ancient and modern populations. The results of the PCA-analysis for 740 samples are shown in the figure in Supplementary 11 (the list of samples is presented in *Table 1* of Supplementary 2). A simplified version of these results is reproduced in *Fig. 4* (only 116 samples are displayed, the list of samples is presented in *Table 2* of Supplementary 2). The location of ancient and modern genomes on the PCA map correlates with the geographical coordinates of the corresponding archaeological sites (Pearson correlation 0.76). The PC1 axis mainly coincides with the West–East direction; and the PC2, with the North–South. The genome of Prince Dmitry Alexandrovich (sample coordinates PC1: -0.0071, PC2: 0.0062) holds an intermediate position between the European and Central Asian clusters. The ancient samples closest in time to Prince Dmitry Alexandrovich belong to an early medieval population of Central Europe, the Avars steppe nomads of the late period; for example, the Hungary_LateAvar (ID I16741) [40] and Hungary_Transtisza_LAvar (ID ARK-11) [41].

The Avars were a nomadic people originating from Central Asia who moved to Central Europe in the 6th century and created the state of the Avar Khaganate there (VI–IX centuries). Archaeologists often define the Avars as Caucasoids, suggesting that only a small ruling stratum, the elite, retained a pronounced Mongoloid feature. The recently published genomes of ancient individuals of the Avar period demonstrate their genetic heterogeneity; on the principal component plot, the studied samples are scattered over the entire wedge from the populations of Western Eurasia to the populations of Northeast Asia [41]. Despite this heterogeneity, some patterns were identified: representatives of the early Avar elite form a dense cluster with a high content of the "Ancient Northeast Asians" (ANA) component, while the samples of the Late Avar period are shifted towards Western Eurasia. In turn, representatives of the Avars who are not associated with the elite are quite diverse and have a significantly smaller component of the "ancient northeast Asians" or it is completely absent. Hungary_Late Avar (I16741), an individual of the late Avar period with a mixed genomic profile consisting of ~20% of the Eastern Steppe component and ~80% of the component most pronounced in the previous local inhabit-

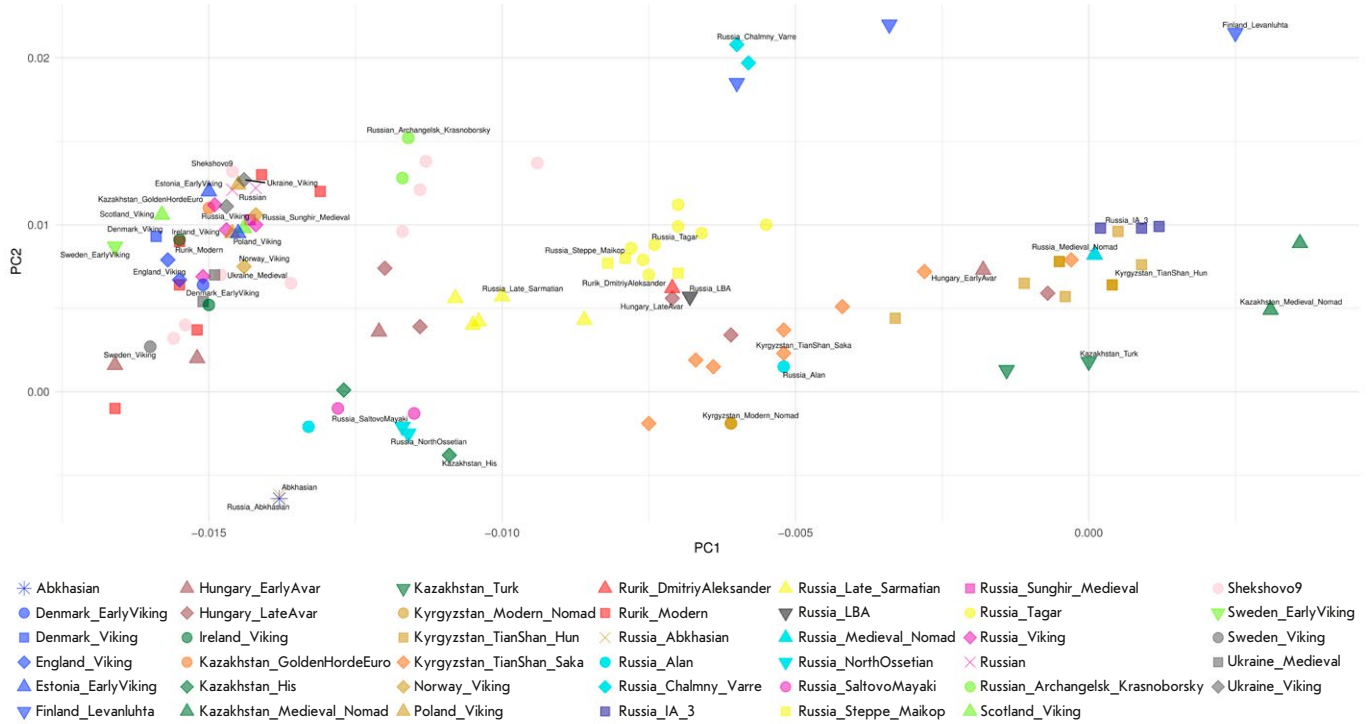


Fig. 4. Principal component analysis. The genome of Prince Dmitry Alexandrovich was projected onto the ancient and modern populations of Eurasia

ants of the Carpathian Basin clustered next to Prince Dmitry Alexandrovich, belongs to this group of samples [41].

Admixture analysis

Analysis of the genetic origin of Prince Dmitry Alexandrovich was carried out using the Admixture method. Supplementary 12 presents the results of an Admixture analysis with parameters K from 6 to 12. The results of the Admixture analysis in simplified form with the number of ancestral populations equal to six ($K = 6$) are shown in *Fig. 5*. Representatives of ancient populations are shown whose genomes were used for modeling the genetic origin of the prince.

When decomposing the genome of Prince Dmitry Alexandrovich into ancestral components, it should be noted its genetic similarity to representatives of the early medieval population of the east of Scandinavia, the "Vikings," which may militate in favor of the version of the "Vikings" (Scandinavian) origin of Rurik, the ancestor of the princely family called Rus', which the Chronicle directly indicates. Here and below, we use the term "Vikings" in quotation marks to show that this is a heterogeneous and complex European population in its historical formation, united only by the way of life and habitat.

Comparison of the genome of Prince Dmitry Alexandrovich with the genomes of the Scandinavian populations of the "Viking" Age, including those from the territory of modern Russia [4], indicates the presence of an additional East Eurasian component in a significant amount (indicated in dark blue color). The indicated component is maximally expressed among the Nganasans, an indigenous people in Siberia, the Finno-Ugric Mansi people, representatives of the indigenous Han people in China (East Asia), among the Avars elite from the Danube-Tisza Interfluve (Hungary_DanubeTisza_MLAvar) [41], as well as among the Early Neolithic of Lake Baikal (Russia_Shamanka_Eneolithic.SG) [42] and Mongolia (Mongolia_North_N). To a lesser extent, this component is present in samples of the early Middle Ages from the territory of modern Finland (Finland_Levanluhta), in a sample synchronous with Prince Dmitry Alexandrovich from the Caspian steppe (Russia_Medieval_Nomad), as well as in more ancient Iranian-speaking steppe nomads of the Iron Age from the territories of modern Kazakhstan and Kyrgyzstan (Kazakhstan_TianShan_Saka, Kyrgyzstan_TianShan_Hun). Due to the fact that the steppe and Finno-Ugric populations share a common origin, this type of analysis does not allow

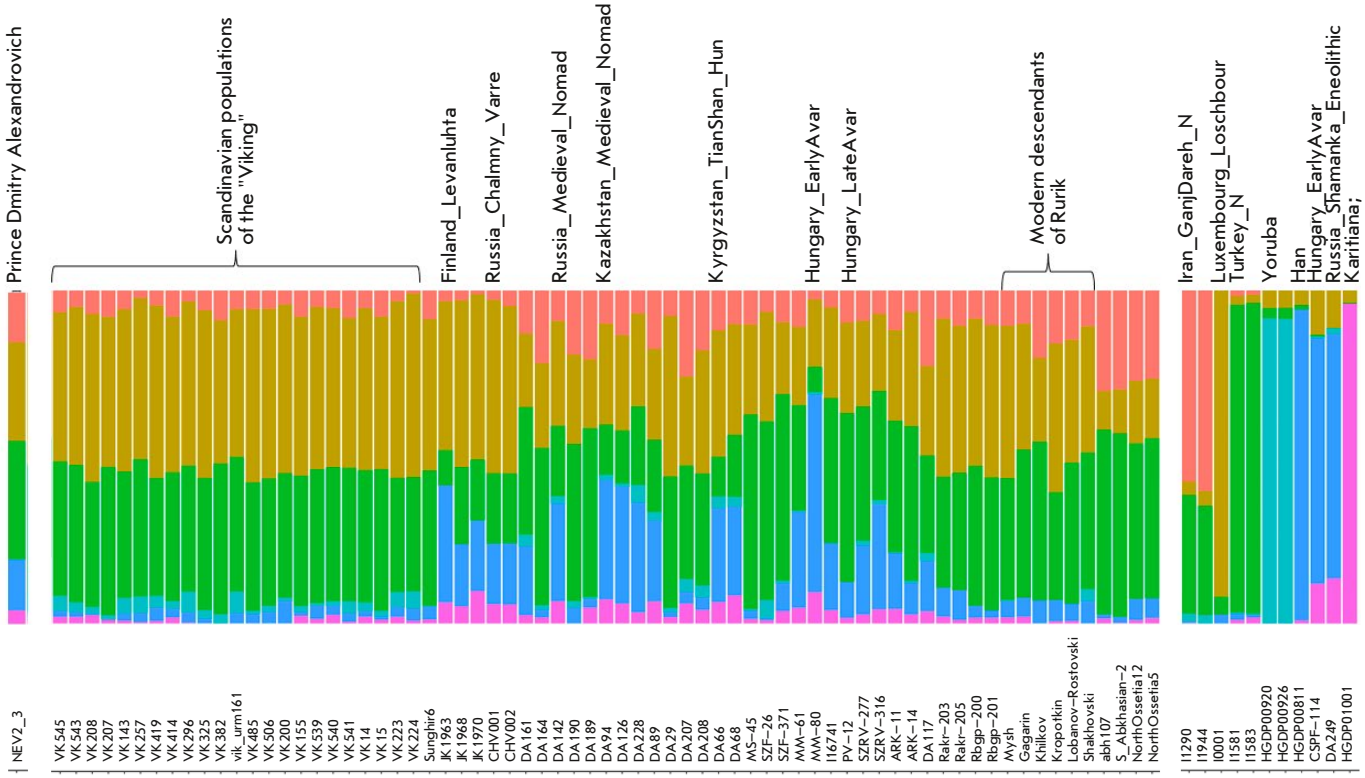


Fig. 5. Results of the Admixture analysis for representatives of ancient populations whose genomes were used to assess the genetic origin of Prince Dmitry Alexandrovich (the number of ancestral populations is six ($K = 6$))

us to specifically attribute this component to one of these groups, with all used K equal to 6–12.

Thus, based on PCA data, Admixture analysis, and information on mitochondrial DNA, it can be argued that Prince Dmitry Alexandrovich had a significant eastern component in his genome. This distinguishes him from the early medieval population of the east of Scandinavia, the "Vikings," and the medieval Slavic sample from Vladimir (Sungir6), but it makes him closer to the ancient population of Finland, the Kola Peninsula, and the early medieval population of central Europe, which includes a well-known component of steppe nomads. Probably, this contribution came from both the male and female lines, which corresponds to the routes of ancient migration from Siberia to the north of Europe and migrations from Siberia in the first millennium BC – the first millennium AD along the Eurasian steppe corridor.

Modeling the genome of Prince Dmitry Alexandrovich from the genomes of ancestral populations

After analyzing the results of the PCA and Admixture analysis, as well as available historical information, we selected populations that could participate in the formation of the genome of Prince Dmitry Alexandrovich: the genomes of the early medieval population of the east of Scandinavia; representatives of the Iranian-speaking nomads of the Eurasian steppes of the Iron Age, and the population of the early Middle Ages of Central Europe, which includes a well-known component of the steppe nomads and samples of individuals representing the ancient East Eurasian component. To assess the contribution of the Slavic component to the genome of Prince Dmitry Alexandrovich, samples of the Medieval Russian population of the XI century from the rural necropolis of the Shekshovo settlement in Suzdal Opolye and an individual of the XII century from the territory of modern Vladimir (Sungir6) were used [43, 44]. Several models were tested:

Modeling of the genome of Prince Dmitry Alexandrovich using the genomes of the "Vikings". QpWave-analysis determined that a minimum of three ancestry sources were sufficient to model the genome of Prince Dmitry Alexandrovich (Supplementary 13). Exploring all possible combinations of three sourc-

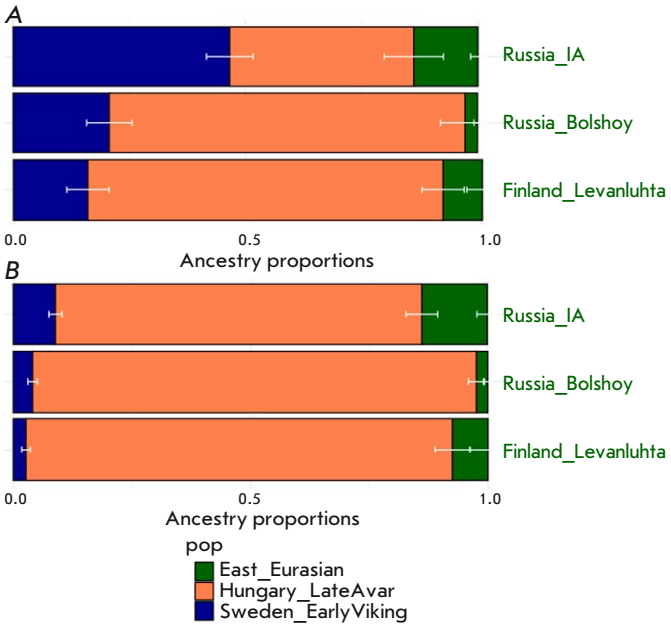


Fig. 6. The results of modeling the genome of Prince Dmitry Alexandrovich as a three-way admixture of "Vikings" (dark blue color), the early medieval population of Central Europe (orange), which includes a well-known component of the steppe nomads, and the ancient East Eurasian component (green). Early "Vikings" from the territory of Sweden (*A*) and early "Vikings" from the territory of Estonia (*B*) are shown as representatives of the "Vikings"

es from the selected list was accomplished using the qpAdm tool. Results of the modeling of the genome of Prince Dmitry Alexandrovich (p greater than 0.05) are presented in Supplementary 14 (*Table 1*). The genome of Prince Dmitry Alexandrovich can be successfully modeled as a three-way admixture of the "Vikings," steppe nomads, and ancient East Eurasian components. For example, one of such models includes 46.6% of the early medieval population of the east of Scandinavia (Sweden_EarlyViking); 39.6% of the early medieval population of central Europe, which includes the known component of the steppe nomads (Hungary_LateAvar); and 13.8% of the Russia_IA source as the third component (Iron Age sample from the territory of the Republic Altai). Substitution of Russia_IA with Iron Age samples from the territory of Finland (Finland_Levanluhta) or the Bolshoy Oleny Island in the Kola district of the Murmansk region produced a reliable three-way admixture model. These third components (Russia_IA, Finland Levanluhta and Bolshoy Oleny) have the ancient East Eurasian component in common. According to recent studies, people of that ancestry arrived in the Kola Peninsula more than 3,500 years ago from Siberia [34] and mixed with local populations thus forming people who speak the Finno-Ugric language today. The results are presented in *Fig. 6A* and in Supplementary 14 (*Table 1*).

Among "Vikings" Sweden_EarlyViking was the largest contributing population to the genome of Prince Dmitry Aleksandrovich (46.6%). For other "Vikings" this proportion did not exceed 9%. The minimum contribution is made by the Estonia_EarlyViking population: 2.7% (*Fig. 6B*). The Sweden_EarlyViking population is represented by three samples from the village of Bode (Böde) on the island of Öland and dated around the period between the VII and VIII centuries AD (with ID's VK379, VK382, VK359). The analysis of strontium isotopes for these samples [4] attributed them to the category of migrants to the town of Bode, although the question of whether they were the original inhabitants of the island of Oland remains open.

Our qpAdm analysis showed a significant genetic difference between the Oland_Sweden_EarlyViking group and other samples from Oland Island, designated as Oland_Sweden_Viking: the first one can be successfully (p = 0.64) modeled as a three-way admixture between European hunter-gatherers, Neolithic farmers, and steppe pastoralists (Supplementary 14 *Table 2*), while the Oland_Sweden_Viking group could not be modeled as a mixture of these ancestral populations (p = 0.01). There are also significant differences in the modeling of these two groups of "Vikings" using a single source group, which are Iron Age populations from the territory of Europe or a sample of the Medieval Russian population of the XI century from the territory of modern Vladimir, Russia (Supplementary 14 *Table 3*). In an extended analysis, all "Viking" population groups were used to test the single-source model for the Sungir6 sample (*Fig. 7*, Supplementary 14, *Table 4*). Those populations of "Vikings" that provide reliable values of F4-statistics for this model are concentrated in the northern part of Europe, Ireland and Iceland, while none of the southern populations fits into such a model. These results raised the question of the relationship of the Scandinavian population groups to the Slavs in the period from the VI to the XI centuries.

Modeling the genome of Prince Dmitry Alexandrovich using the genomes of Iranian-speaking nomads of the Eurasian steppes in the Iron Age

A statistically reliable genetic model of the genome of Prince Dmitry Alexandrovich was also obtained

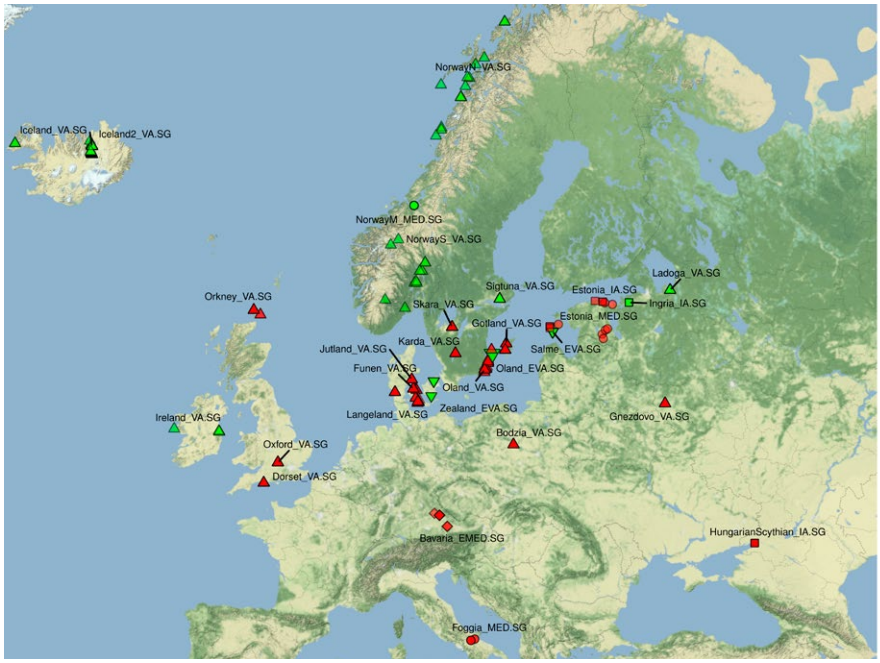


Fig. 7. The map shows the burial sites of "Vikings" who confirmed (green) or did not confirm (red) the model of one source for the Sungir6 sample. The shape of the label reflects the temporal classification of images; the decoding is given in the legend to the right of the figure

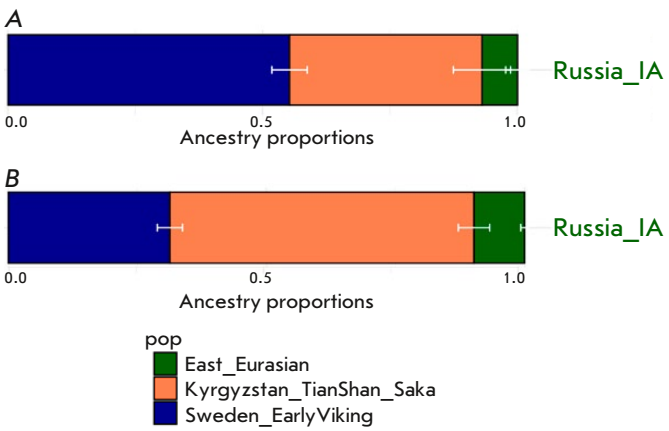


Fig. 8. The results of modeling the genome of Prince Dmitry Alexandrovich as a three-way admixture of "Vikings" (dark blue color), ancient East Eurasian component (green), Early Iron Age nomads from the territory of Kyrgyzstan (*A*) and Kazakhstan (*B*) (orange)

by replacing the early medieval population of Central Europe (including a component of steppe nomads) with Iranian-speaking nomads of The Eurasian steppes in The Iron Age (Kazakhstan_TianShan_Saka, Kyrgyzstan_TianShan_Hun). The results are shown in *Fig. 8* and in Supplementary 14 (*Table 5*). The model provides a fits only with early Iron Age nomads from the Tien Shan region, perhaps, due to their genetic profile: high proportion (70%) of the Late Bronze Age steppe pastoralists, 25% of the South Siberian hunter-gatherer component, and 5% of the component associated with the Neolithic population of Iran [45]. This group of nomads, according to the results of f3-outgroup-statistics, is genetically closer to northern European populations compared to other nomads of the early Iron Age from the Asian cluster.

Modeling the genome of Prince Dmitry Alexandrovich using the genomes of the Medieval Russian Slavic population

We hypothesize that alternative models – replacing the population of "Vikings" with Medieval Russian Slavic populations – will likely also provide a fit. Since the pagan Slavic tradition practiced cremation until the end of the X century, we used samples of the medieval Russian population of the XI century from the rural necropolis of the settlement of Shekshovo9 in the Suzdal Opolye and an individual of the XI century from the territory of modern Vladimir, Russia (Sungir6) [42, 43]. The genomes of individuals from Shekshovo9 are the result of a mix of the Central European (Slavic) and local (Finnish) genetic components, while Sungir6 is considered a pure Medieval Russian Slavic population. These features of the samples are confirmed in the PCA plot: Sungir is located in a cluster of European (Danish, Polish, Norwegian, Ukrainian, etc.) medieval samples, and Shekshovo9 is shifted to the "East" along the PC1 axis (*Fig. 4*).

We successfully modeled Prince Dmitry Alexandrovich's ancestry as being derived from pop-

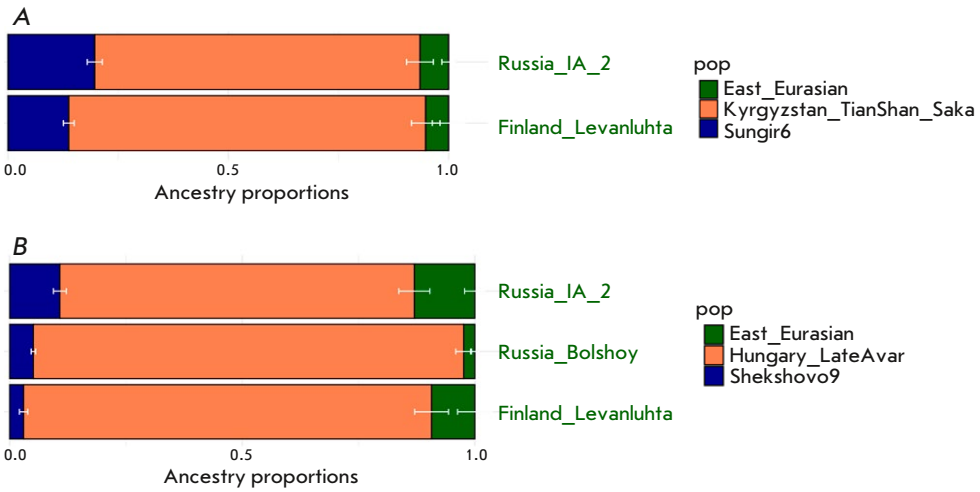


Fig. 9. The results of modeling the genome of Prince Dmitry Alexandrovich as a three-way admixture of the Medieval Russian Slavic population (dark blue color), the ancient East Eurasian component (green), and (*A*) representatives of the steppe nomads (orange) of the Early Iron Age or (*B*) the early medieval population of Central Europe (steppe nomads from the territory of Hungary)

ulations related to the Sungir6 sample – 19.7%; Iron Age nomadic steppe peoples – 73.8% for Kyrgyzstan_TianShan_Saka; and the Iron Age from the territory of the present-day Altai Republic – 6.5% for Russia_IA (*Fig. 8* and in Supplementary 14, *Table 6*). Replacing the Sungir6 with Shekshovo9 as a source, the proportion of the Medieval Russian Slavic population amounts to 18.7%; the nomadic Iron Age steppe peoples (Kyrgyzstan_TianShan_Saka) – 78.2%; and Russia_IA_2 – about 3% (*Fig. 9A*). In the slightly different three-source modeling with the Hungarian Avars, Slavic Medieval Russian and Russia IA – the proportion reached 76,2%, 10,8%, and 13,1%, correspondingly (*Fig. 9B*). The decrease in the Slavic proportion can be explained by the partial compensation of Slavic origin by the Avars individuals who were used for modeling. Their origin suggests 80% of the local East European source and only 20% of the Central Asian one (ID I16741) [40].

Thus, modeling of the genome of Prince Dmitry Alexandrovich indicates the contribution of three ancestral sources to its origin: (1) the early medieval population of the east of Scandinavia from the island of Oland, (2) the steppe nomadic peoples of the Eurasian steppes of the Iron Age or the early medieval population of central Europe (steppe nomads from the territory of Hungary), and (3) the East Eurasian component. Alternative models, replacing the population of "Vikings" with the Medieval Russian Slavic populations (Shekshovo9 and Sungir6) also provide a fit.

CONCLUSION

Paleogenetics and mathematical statistics provided an opportunity to discuss the origin of the Rurikids and design a reliable tool to attribute remains with disrupted documentation to this ruling family.

An analysis of the genealogical tree of the Rurikids showed that the modern individuals of this family, who have a Y chromosome clustered with Prince Dmitry Alexandrovich's Y chromosome, belong to three different branches – the Olgoviches, Mstislaviches, and Yuryeviches. Thus, the N1a haplogroup of the Y chromosome characterizes all three branches of the tree, suggesting that their first common ancestor, Prince Yaroslav the Wise, was a carrier of N1a haplogroup also.

The mitochondrial haplogroup of Prince Dmitry Alexandrovich was determined as F1b1, which may point to the contribution of eastern populations to his genome. This hypothesis is also supported by autosomal data (PCA and Admixture). Although the main genetic makeup of Dmitry Aleksandrovich can be attributed to the Scandinavians/Slavic/European populations our results provide clear evidence of the input of the Eastern genetic component. This is in line with historical data: marriages of Russian princes with the daughters of the Polovtsian khans from the end of the XI century were a common practice that cemented allied relations and political interaction [46, 47]. Dmitry's mother, the wife of Alexander Nevsky, Alexandra Bryachislavna, came from the Polotsk Izyaslaviches family. Information about the wives of these princes is scarce, and the name and origin of Alexandra's mother is unknown. However, men of the Polotsk branch of the Rurikids did not avoid marriage alliances with Polovtsy women. From the Polovtsian family came the second wife of the Polotsk Prince Svyatopolk Izyaslavich (1050–1113), Elena, the daughter of Khan Tugorkan (Complete collection of Russian chronicles,

1997, vol. I, p. 231–232). The circulation of eastern mitochondrial groups in this situation is quite expected. Alternatively, the origins of the eastern component in the genome of Dmitry Alexandrovich might be associated with the marriages of the Rurikids with representatives of the dynasties of Central and Southern Europe (Serbian Vukanovichi, Hungarian Arpads). Eastern genes brought by the migration of the first millennium AD [48] could be much better preserved within elites than within plebs.

Our results raise some questions on Rurikids' genetic history. The most obvious of them are (i) how to explain the presence of the "Eastern component" in the genome of Dmitry Aleksandrovich and (ii) were there a genetic connection between Scandinavians and Slavs in the pre Rurik era? The answers may come after a systematic paleogenomic study of new, reliably documented paleoanthropological materials from the territory of Russia. ●

This work was financially supported by the Ministry of Education and Science of the Russian Federation, project No. 075-10-2020-116 (grant No. 13.1902.21.0023).

Supplementarys are available on the website https://doi.org/10.32607/actanaturae.23425.

REFERENCES

1.Volkov V.G. // Genealogy of the pre-Petrine time: source study, methodology, research. 2012. P. 11–40.
2.Balanovsky O.P. // Ethnographic review. 2021. № 4. P. 119–131.
3.Pchelov E.V. How the genetics of the Rurikids is studied and what comes out of it // Genofond.rf. 2016. http://xn--c1acc6aafa1c.xn--p1ai/?page_id=7053 (accessed 18.06.2023).
4. Margaryan A., Lawson D.J., Sikora M., Racimo F., Rasmussen S., Moltke I, Cassidy L.M., Jørsboe E., Ingason A., Pedersen M.W., et al. // Nature. 2020. V. 585. P. 390–396.
5. Rohland N., Glocke I., Aximu-Petri A., Meyer M. // Nat. Protoc. 2018. V. 13. P. 2447–2461.
6. Gansauge M.T., Meyer M. // Nat. Protoc. 2013. V. 8. P. 737–748.
7. Sharko F.S., Zhur K.V., Trifonov V.A., Prokhortchouk E.B. // Acta Naturae. 2023. V. 15. P. 87–96.
8. Mathieson I., Lazaridis I., Rohland N., Mallick S., Patterson N., Roodenberg S.A., Harney E., Stewardson K., Fernandes D., Novak M., et al. // Nature. 2015. V. 528. P. 499–503.
9. Y-DNA Haplogroup Tree 2019–2020. https://isogg.org/tree (accessed 18.06.2023).
10. Green R.E., Malaspinas A.S., Krause J., Briggs A.W., Johnson PL.F., Uhler C., Meyer M., Good J.M., Maricic T., Stenzel U., et al. // Cell. 2008. V. 134. P. 416–426.
11. Bushnell B., Rood J., Singer E. // PLoS One. 2017. V. 12. P. e0185056.
12. Schubert M., Ermini L., Der Sarkissian C., Jónsson H., Ginolhac A., Schaefer R., Martin M., Fernandez R., Kircher M., McCue M., et al. // Nat. Protoc. 2014. V. 9. P. 1056–1082.
13. Martin M. // EMB net.journal. 2011. V. 17. P. 10–12.
14. Li H., Durbin R. // Bioinformatics. 2009. V. 25. P. 1754–1760.
15. Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. // Bioinformatics. 2009. V. 25. P. 2078–2079.
16. Jónsson H., Ginolhac A., Schubert M., Johnson P.L.F., Orlando L. // Bioinformatics. 2013. V. 29. P. 1682–1684.
17. Mallick S., Micco A., Mah M., Ringbauer H., Lazaridis I., Olalde I., Patterson N., Reich D. // bioRxiv. 2023. doi:10.1101/2023.04.06.535797
18. Alexander D.H., Novembre J., Lange K. // Genome Res. 2009. V. 19. P. 1655–1664.
19. Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R, Bender D., Maller J., Sklar P., de Bakker P.I., Daly M.J. et al // Am. J. Hum. Genet. 2007. V. 81. P. 559–575.
20. Weissensteiner H., Pacher D., Kloss-Brandstätter A., Forer L., Specht G., Bandelt H.J., Kronenberg F., Salas A., Schönherr S. // Nucl. Acids Res. 2016. V. 44. P. 58–63.
21. Walsh S., Chaitanya L., Clarisse L., Wirken L., Draus-Barini J., Kovatsi L., Maeda H., Ishikawa T., Sijen T., de Knijff P., et al. // Forensic Sci. Int. Genet. 2014. V. 9. P. 150–161.
22. Walsh S., Chaitanya L., Breslin K., Muralidharan C., Bronikowska A., Pospiech E., Koller J., Kovatsi L., Wollstein A., Branicki W., et al. // Hum. Genet. 2017. V. 136. P. 847–863.
23. Chaitanya L., Breslin K., Zuñiga S., Wirken L., Pośpiech E., Kukla-Bartoszek M. Sijen T., Knijff P., Liu F., Branicki W., et al. // Forensic Sci. Int. Genet. 2018. V. 35. P. 123–135.
24. Complete collection of Russian chronicles. V. 3. Novgorod First Chronicle of the Senior and Junior Recensions. M., 2000. 692 p.
25. Complete collection of Russian chronicles. V. 3, part 1. Novgorod fourth chronicle. M., 2000. 690 p.
26. Complete collection of Russian chronicles. V. 7. Chronicle according to the Resurrection List. M., 2001. 345 p.
27. Complete collection of Russian chronicles. V. 1. Laurentian Chronicle. L., 1926–1928. 496 p.
28. Complete collection of Russian chronicles. V. 8. Chronicle collection, called the Patriarchal or Nikon chronicle. M., 2000. 532 p.
29. Volkov V.G., Seslavin N.A. Genetic study of the Rurik Dynasty // Centenary of human population genetics. Conference proceedings. 29–31 May 2019. P. 84–85.
30. Barabanov O.N. // Space and time. 2015. № 3 (21). P. 198–207.
31. Maksimov N.G. // Russian Newsweek. 2006. № 50 (128). P. 68–69.
32. Maksimov N.G. // Russian Newsweek. 2007. № 52 (176). P. 50–58.
33. Borbely N., Mende B.G., Pinhasi R., Cheronet O., Reich D. Takacs A., Hajdu T., Szecsenyi-Nagy A. // 28th EAA Annu. Meet. Budapest, Hungary, 31 August – 3 September 2022. P. 193.
34. Lamnidis T.C., Majander K., Jeong C., Salmela E., Wessman A., Moiseyev V., Khartanovich V., Balanovsky O., Ongyerth M., Weihmann A., et al. // Nat. Commun. 2018. V. 9. P. 5018.

35. Yao Y.G., Kong Q.P., Bandelt H.J., Kivisild T., Zhang Y.P. // Am. J. Hum. Genet. 2002. V. 70. P. 635–651.
36. Derenko M.V., Grzybowski T., Malyarchuk B.A., Dambueva I.K., Denisova G.A., Czarny J., Dorzhu C.M., Kakpakov V.T., Miścicka-Sliwka D., Woźniak M., et al. // Ann. Hum. Genet. 2003. V. 67. P. 391–411.
37. Keyser-Tracqui C., Crubézy E., Ludes B. // Am. J. Hum. Genet. 2003. № 73 (2). P. 247–260.
38. Trapezov R.O., Pilipenko A.S., Molodin V.I. // Vavilov Journal of Genetics and Breeding. 2014. V. 18. № 3. P. 469–477.
39. Csáky V., Gerber D., Koncz I., Csiky G., Mende B.G., Szeifert B., Egyed B., Pamjav H., Marcsik A., Molnár, et al. // Sci. Rep. 2020. V. 10. P. 948.
40. Maróti Z., Neparáczki E., Schütz O., Maár K., Varga G.I.B., Kovács B., Kalmár T., Nyerki E., Nagy I., Latinovics D., et al. // Curr. Biol. 2022. V. 32. P. 2858–2870.e7.
41. Gnecchi-Ruscone G.A., Szécsényi-Nagy A., Koncz I., Csiky G., Rácz Z., Rohrlach A.B., Brandt G., Rohland N., Csáky V., Cheronet O., et al. // Cell. 2022. V. 185. P. 1402–1413.e21.
42. de Barros Damgaard P., Martiniano R., Kamm J., Moreno-Mayar J.V., Kroonen G., Peyrot M., Barjamovic G., Rasmussen S., Zacho C., Baimukhanov N. // Science. 2018. V. 360. doi: 10.1126/science.aar7711.
43. Peltola S., Majander K., Makarov N., Dobrovolskaya M., Nordqvist K., Salmela E., Onkamo P. // Curr. Biol. 2023. V. 33. P. 174–182.e10.
44. Sikora M., Seguin-Orlando A., Sousa V.C., Albrechtsen A., Korneliussen T., Ko A., Rasmussen S., Dupanloup I., Nigst P.R., Bosch M.D., et al. // Science. 2017. V. 358. P. 659–662.
45. Damgaard P.B., Marchi N., Rasmussen S., Peyrot M., Renaud G., Korneliussen T., Moreno-Mayar J.V., Pedersen M.W., Goldberg A., Usmanova E., et al. // Nature. 2018. V. 557. P. 369–374.
46. Pletneva S.A. Polovtsy. M.: Nauka, 1990. 208 p.
47. Litvina A.F., Uspensky F.B., 2013. Russian names of the Polovtsian princes: Interdynastic contacts through the prism of anthroponymy. M.: Polimedia, 2013. 280 p.
48. Veeramah K.R., Rott A., Groß M., van Dorp L., López S., Kirsanow K., Sell C., Blöcher J., Wegmann D., Link V., et al. // Proc. Natl. Acad. Sci. USA. 2018. V. 115. P. 3494–3499.

CNA Landscape of HER2-Negative Breast Cancer in Anthracycline-Based Neoadjuvant Chemotherapy Regimens

M. K. Ibragimova[1,2,3*], E. A. Kravtsova[1,2], M. M. Tsyganov[1,3], N. V. Litviakov[1,2,3]
[1]Cancer Research Institute, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, 634009 Russian Federation
[2]National Research Tomsk State University, Tomsk, 634050 Russian Federation
[3]Siberian State Medical University, Tomsk, 634050 Russian Federation
*E-mail: imk1805@yandex.ru
Received: May 12, 2023; in final form, July 28, 2023
DOI: 10.32607/actanaturae.20377
Copyright © 2023 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **Critical evaluation of how and when to include anthracyclines in preoperative chemotherapy is becoming more relevant in an era when the molecular genetic approach not only allows for the development of biologically targeted therapeutics, but also implies the ability to select the patients likely to benefit from certain cytotoxic agents. Changes in the copy number aberration (CNA) landscape of luminal B HER2-negative (HER2−) breast cancer (BC) during anthracycline-based neoadjuvant chemotherapy (NAC) regimens were studied in order to identify groups of potential CNA markers of objective response and CNA markers for predicting the development of hematogenous metastasis. Comparison of CNA frequencies depending on the response to NAC showed that objective response was observed in a larger number of deletions in the 11q22.3 and 11q23.1 loci (p = 0.004). Comparison of CNA frequencies in groups of patients after treatment showed that hematogenous metastasis was observed with a greater number of amplifications in the 9p22.2 locus (p = 0.003) and with a greater number of deletions in the 9p21.3 locus (p = 0.03). Potential predictive CNA markers of objective response and prognostic CNA markers of hematogenous metastasis in anthracycline-based NAC regimens have been identified.**
KEYWORDS **breast cancer, CNA landscape, anthracycline-based regimens, neoadjuvant chemotherapy, prognosis.**
ABBREVIATIONS **CNA – copy number aberration; BC – breast cancer; NAC – neoadjuvant chemotherapy; pCR – pathological complete response; SNV – single nucleotide variation; VAF – variant allele frequency; HER2 – human epidermal growth factor receptor 2; MFS – metastasis-free survival.**

INTRODUCTION

Neoadjuvant chemotherapy (NAC) is considered the standard for the combination treatment of HER2-positive (HER2+) breast cancer (BC). At the same time, the treatment of localized ER+/HER2-negative (HER2−) BC, which is characterized by poorer chemosensitivity compared to other clinical BC subtypes, can be challenging [1]. Pathological complete response (pCR) rates in HER2-negative (HER2−) BC are low, while the presence of residual disease does not have the same prognostic value as in other clinical subtypes [2]. However, the rate of survival of patients who display complete or partial regression after NAC differs significantly from that of patients for whom the disease has stabilized or continues to progress [3, 4]. In this regard, in the case of HER2− BC, it is important that we search for predictive markers of complete and partial regression, in contrast to pCR markers in the triple-negative (TN) and HER2+ BC subtypes.

Several approaches to the treatment of HER2− and metastatic BC patients exist to date. However, no gold standard has been established for first-line treatment so far. Anthracycline- and taxane-based regimens are considered traditional systemic approaches to first-line chemotherapy and neoadjuvant therapy in this disease subtype [5].

The presence of toxic side effects for anthracycline-based NAC regimens (cardiotoxicity, leukemogenic effects, and secondary malignancies) [6–8], along with the central medical ethics principle of “not to do harm”, makes it extremely arduous to not only

identify patients with the highest positive response to chemotherapy, but also find the systemic approach with the highest possible therapeutic index and minimal risk of significant long-term treatment-related toxicity.

In 2021, the National Comprehensive Cancer Network Guidelines removed the anthracycline-based therapy both from the list of preferred regimens for the treatment of early-stage HER2+ BC and from the "regimen of choice" category [9]. However, the results of meta-analyses available to date show that anthracycline use is justified in luminal B ER+/HER2− BC [10].

Thus, there are problems involved in selecting a NAC approach in HER2− BC. The question of when and in which HER2− BC patients a certain NAC regimen should be used requires further discussion.

Anthracycline-based regimens are an important component of BC treatment, especially in TN BC with a high risk of recurrence (regardless of axillary lymph node involvement) and HER2−/ER+ BC with axillary lymph node involvement. For this reason, it is necessary to search for biomarkers that predict the response to anthracyclines during NAC in BC, including both pCR and partial regression, which are associated with a favorable outcome.

In this work, we studied the changes in the copy number aberration (CNA) landscape of luminal B HER2− BC in the presence of anthracycline-based NAC regimens to identify groups of predictive CNA markers of objective (pCR + partial > 50% regression) response to treatment and potential CNA markers for predicting hematogenous metastasis.

EXPERIMENTAL

Material and methods

The study included 35 patients aged 25–68 years (mean age, 49.3 ± 0.1 years (Mean ± SE)) with morphologically verified luminal B HER2− BC of the stages $T_{1-4}N_{0-3}M_0$ (IIA–IIIB). Luminal B HER2− subtype was defined as ER +, PR + or −, Ki67 > 30%.

According to the Consensus Conference on Neoadjuvant Chemotherapy in Carcinoma of the Breast (April 26–28, 2003, Philadelphia, Pennsylvania), patients received 4–8 NAC courses using the following regimens: FAC (fluorouracil, doxorubicin, and cyclophosphamide) AC (doxorubicin, cyclophosphamide), and CAX (cyclophosphamide, doxorubicin, and xeloda). The effectiveness of preoperative chemotherapy was evaluated based on the WHO and International Union against Cancer criteria using ultrasound and/or mammography, which were performed before treatment, after two NAC courses, and before surgery. Complete regression (100% tumor reduction), partial regression (> 50% decrease in tumor volume), stabilization (< 50% decrease or > 25% increase in tumor volume), and progression (> 25% increase in tumor volume) were recorded. All cases of complete regression were confirmed by morphological analysis. According to international recommendations, BC patients with disease course stabilization or progression are included in the group with no response to NAC, and patients with partial regression form the group with an objective response during preoperative chemotherapy. It is impossible to obtain tumor samples if the tumor went into complete regression after NAC.

Table 1 presents the main clinical and morphological characteristics of the patients included in the study.

Tumor biopsy samples obtained before treatment under ultrasound guidance and surgical samples resected after NAC were used in the study. DNA was isolated from 35 paired BC tissue samples obtained from each woman before and after NAC.

Microarray analysis was carried out using the high-density Affymetrix CytoScan HD Array (USA). Sample preparation, hybridization, and scanning were performed using an Affymetrix GeneChip® Scanner 3000 7G (Affymetrix). The results were analyzed using Chromosome Analysis Suite 4.0 (Affymetrix).

A statistical analysis of the data was performed using the Statistica 8.0 package (StatSoft Inc., USA). The $\chi 2$ test was used to assess the differences between frequencies (http://vassarstats.net/index.html). The survival rate was analyzed using the Kaplan–Meier method and logrank test.

Compliance with patients' rights and bioethics principles. The study was conducted in accordance with the 1964 Declaration of Helsinki (amended in 2013). The study protocol was approved by the Biomedical Ethics Committee of the Cancer Research Institute of the Tomsk National Research Medical Center of the Russian Academy of Sciences (protocol No. 1, 01/14/2013). All patients provided an informed consent to participate in the study.

RESULTS

At the first stage of the study, in order to evaluate the changes in the CNA landscape in the presence of anthracycline-based NAC, we described the tumor CNA landscape before and after treatment (*Fig. 1*) and assessed the changes in the CNA frequency in tumor.

The highest amplification number (68.6%) was found in the tumor loci 1q32.1–32.2, 1q42.12–42.13, and 1q42.2 in the patients before treatment. The highest deletion number (68.6%) was observed in the

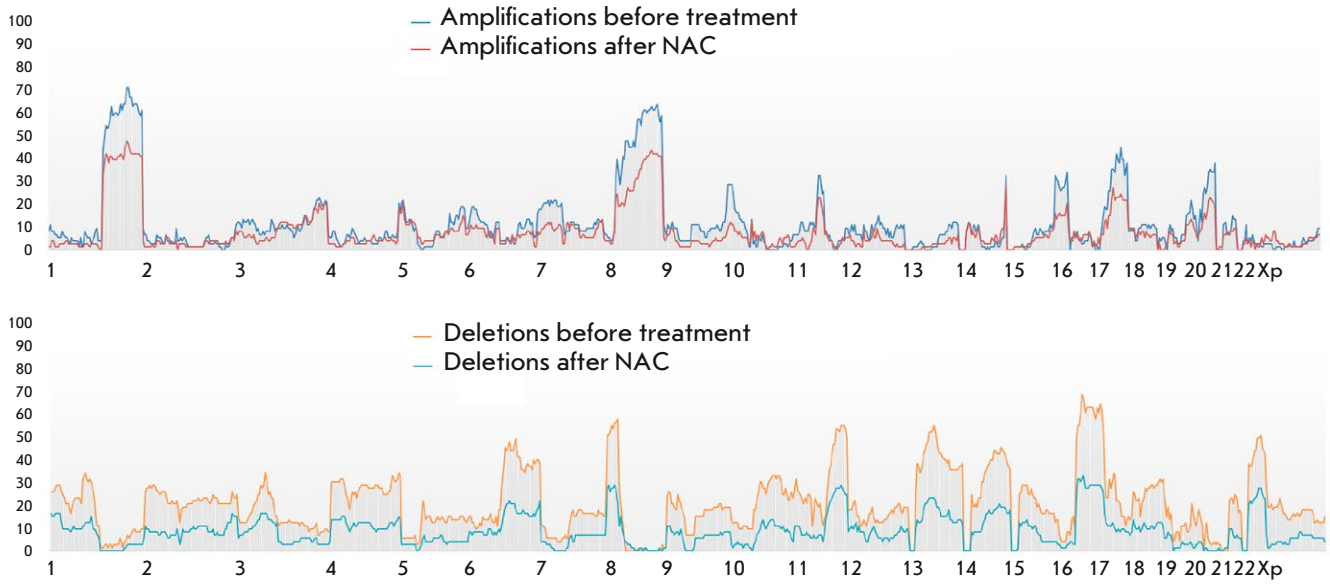


Fig. 1. Frequency of amplifications and deletions in each chromosome in patients receiving anthracycline-based regimens before and after NAC as part of preoperative chemotherapy

loci 17p13.3 and 17p13.1 (in the complete absence of amplifications). A total of 62.9% of amplifications were detected in the loci 8q21.3, 8q22.1–22.2, 8q23.3, 8q24.11–24.12m, and 8q24.21 on an extended region of the long arm of chromosome 8 in a complete absence of deleted regions.

The highest amplification number (48.6%) in tumor after treatment was noted in the loci 1q21.3, 1q32.1-1q32.3, 1q41, 1q42.11–1q42.13, 1q42.2–42.3, 8q22.3, and 8q23.3; however, no deleted regions were found in the loci 1q21.3, 8q22.3, and 8q23.3. The highest rate of deletions (37.1%) was observed in the loci 16q21 and 16q22.1. A total of 34.3% of the deletions were detected in the tumor loci 11q23.3 and 17p13.3 after NAC, in a complete absence of amplified regions.

Analysis of the number of CNAs resulting from the use of anthracycline-based NAC regimens revealed a statistically significant decrease in the frequency of deletions in the loci 17p13.3 and 17p13.1: from 68.6 to 34.3%, which is 24/35 events before treatment and 12/35 events after NAC, respectively ($p = 0.002$).

Figure 1 presents data on the amplification and deletion frequencies in each chromosome in BC patients before and after anthracycline-based NAC regimens.

Next, in order to search for potential CNA markers that could help predict an objective response to NAC during anthracycline-based regimens, we analyzed the distribution of CNA frequencies in tumor before treatment, depending on the response to preoperative chemotherapy.

Partial tumor regression was observed in 23 out of 35 patients (group 1) after treatment. Disease stabilization and progression were noted in 12 out of 35 patients (group 2) after therapy (*Table 1*).

Among group 1 patients, the highest amplification number (82.6%) was found in the loci 1q32.1–32.2, 1q42.12–42.13, and 1q42.2 in the absence of deletions. The highest deletion number (78.3%) was observed in the loci 11q23.1, 11q23.3, and 17p13.1 in the absence of amplifications. Among group 2 patients, the highest amplification number (58.3%) was detected in the loci 1q23.3, 8q21.11–21.13, 8q21.2, 8q21.3, 8q22.1–22.3, 8q23.1–23.3, 8q24.11–24.13, 8q24.21–24.23, and 8q24.3 in the absence of deletions. The highest number of deletions (59.0%) was recorded in 16q21 and 16q22.1 in the absence of amplifications.

Comparison of the CNA frequencies in these groups of patients demonstrated that an objective response to NAC was observed in the higher deletion number (18/23 events, 78.3%) in the loci 11q22.3 and 11q23.1 in group 1 compared to group 2 (3/12 events, 25.0%) ($p = 0.004$). *These loci could serve as predictive markers of objective response to anthracycline-based regimens in preoperative chemotherapy.*

In order to illustrate the complete picture of the tumor CNA landscape during treatment, we analyzed

Table 1. Clinical and morphological characteristics of the breast cancer patients included in the study

Clinical and morphological characteristic		Number of patients (abs. number, %)
Age (years)	⩽ 45	10 (28.6)
	> 45	25 (71.4)
Menstrual status	Regular	22 (62.9)
	Perimenopause	4 (11.4)
	Menopause	5 (14.3)
	Postmenopause	4 (11.4)
Histological type	Invasive ductal carcinoma	20 (57.1)
	Invasive lobular carcinoma	3 (8.6)
	Invasive unspecified carcinoma	5 (14.3)
	Other type	7 (20.0)
Tumor size	T_1	8 (22.8)
	T_2	25 (71.4)
	T_3	1 (2.9)
	T_4	1 (2.9)
Lymphatic metastasis	N_0	16 (45.7)
	N_1	14 (40.0)
	N_2	1 (2.9)
	N_3	4 (11.4)
Histological type	Monofocal	23 (65.7)
	Multifocal	12 (34.3)
Response to NAC	Progression	1 (2.9)
	Stabilization	11 (31.4)
	Partial regression	23 (65.0)
Median follow-up, months (M ± SE)		80.5 ± 1.1 (min–max: 24–148)
Metastasis rate		13 (37.1)
Median onset of metastasis, months (M ± SE)		45.7 ± 0.4 (min–max: 4–130)
Recurrence rate		4 (11.4)
Median recurrence rate, months (M ± SE)		72.5 ± 1.5 (min–max: 52–107)

the distribution of CNA frequencies depending on the response to preoperative chemotherapy in patients after treatment.

Groups with partial regression (group 3, after NAC) and stabilization/progression (group 4, after NAC) were also formed.

In group 3 patients, the highest amplification number (47.8%) was observed in the loci 1q32.1–32.3, 1q41, 1q42.11–42.13, 1q42.2, 1q42.3, 8q22.3, and 8q23.3 in the absence of deletions; the highest number of deletions (34.8%) was found in the loci 8p23.2, 11q21, 11q22.1–22.3, 11q23.1-23.3, 11q24.1, and 11q24.2.

Comparison of CNA frequencies in the group with partial tumor regression after anthracycline-based NAC regimens showed a statistically significant decrease in the amplification frequency, from 82.6% before treatment to 41.7% after treatment, in the loci 1q32.1 and 1q32.2 ($p = 0.0001$). Furthermore, the frequency of deletions in the loci 17p13.3 and 17p13.1 decreased after NAC (78.3 and 30.4% before and after treatment, respectively; $p = 0.0002$).

In group 4 patients, the highest amplification rate (75.0%) was found in the loci 1q21.3, 1q22, and 1q23.1–23.3 in the absence of deletions; the highest num-

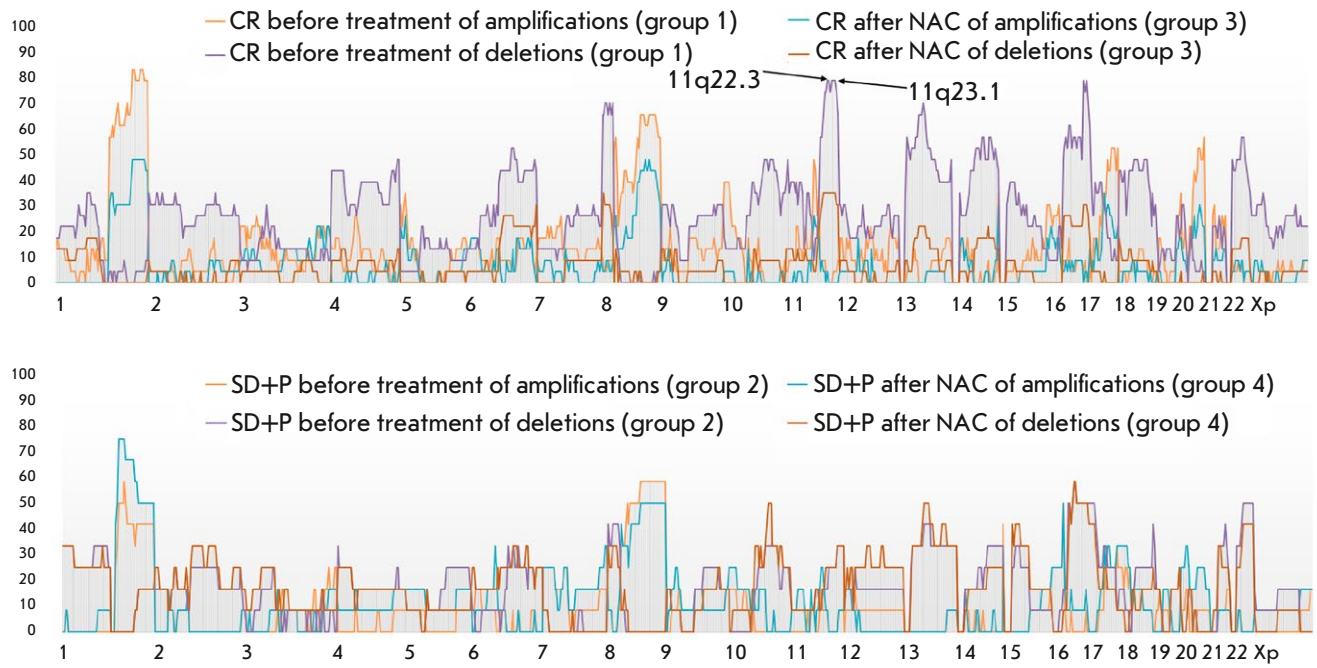


Fig. 2. CNA frequency in breast cancer patients receiving anthracycline-based regimens before and after treatment as part of preoperative chemotherapy depending on the NAC effect. For group 1 and 3 patients with objective response to NAC (partial cancer regression), the CNA frequency before and after NAC is presented as amplification/deletion 1 and amplification/deletion 3, respectively. For group 2 and 4 patients with an absence of an objective response to NAC (cancer stabilization/progression, SD+P), the CNA frequency before and after NAC is presented as amplification/deletion 2 and amplification/deletion 4, respectively

ber of deletions (58.3%) was observed in 16q21 and 16q22.1.

Comparison of CNA frequencies in the group with tumor stabilization and progression after anthracycline-based NAC regimens demonstrated an increase in the amplification rate in 6p12.2, from 0% to 25.0% before and after NAC, respectively (p = 0.001). The frequency of deletions in the loci 6p11.1 increased after NAC (0 and 25.0% before and after treatment, respectively; p = 0.001).

Figure 2 presents the summarized data on CNA frequencies in BC patients before and after therapy, depending on the NAC effect.

Next, in order to identify potential prognostic CNA markers of hematogenous metastasis during anthracycline-based NAC regimens, we analyzed the distribution of CNA frequencies in tumor before treatment, depending on the status of hematogenous metastasis.

Hematogenous metastasis was registered in 13 patients in the studied group. The following groups were formed: groups 5 and 6, which included patients without hematogenous metastasis before and after NAC, respectively, and groups 7 and 8, consisting of individuals with hematogenous metastasis before and after NAC, respectively. *Table 1* presents data on the rate of hematogenous metastasis and the median onset of metastasis.

In group 5, the highest number of amplifications (59.1%) was registered in the loci 8q21.3, 8q22.1–22.3, 8q23.1–23.3, 8q24.11–12, and 8q24.21 in the absence of deletions and in 1q32.1–32.2, 1q42.12–42.13, and 1q42.2 with a deletion rate of 9.1%. The highest deletions rate (77.3%) was found in 17p13.3–13.1 in the absence of amplifications.

In group 7, the highest amplification rate (84.6%) was detected in 1q23.2–23.3, 1q24.1–24.3, and 1q25.1–25.3 in the absence of deletions. The highest deletion rate (76.9%) was observed in 11q23.3, 11q24.1, and 11q24.2 in the absence of amplifications.

Comparison of CNA frequencies in these groups showed that hematogenous metastasis took place in the highest amplification rate in the loci 18q11.2, 18q12.1, and 18q12.2. In particular, 23.0% of patients with diagnosed hematogenous metastasis had amplifications in these loci, while patients without hematog-

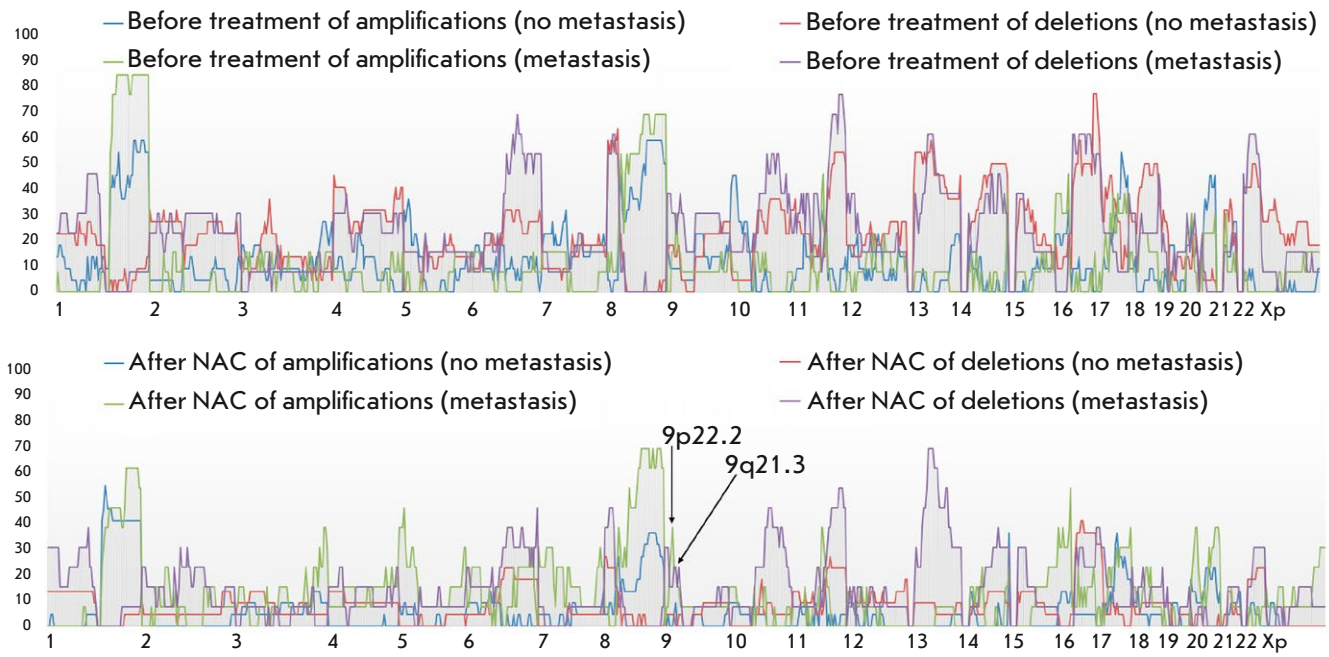


Fig. 3. CNA frequency in breast cancer patients receiving anthracycline-based regimens before and after treatment as part of preoperative chemotherapy, depending on the presence of hematogenous metastasis

enous metastasis demonstrated no amplifications in these loci ($p = 0.035$).

In group 6 patients, the highest amplifications rate (54.6%) was found in the 1q21.3 locus, while no deletions were observed. The highest deletions rate (40.9%) was found in loci 16q21 and 16q22.1.

In group 8 patients, the highest amplifications rate (69.2%) was detected in 8q21.13, 8q21.2, 8q21.3, 8q22.1–22.3, 8q23.3, 8q24.13, 8q24.21, and 8q24.22; the highest deletion rate (69.1%) was observed in 13q14.11–14.13 and 13q14.2, in the absence of amplifications.

Comparison of CNA frequencies in these groups after treatment showed that hematogenous metastasis is associated with a high number of amplifications in the 9p22.2 locus: 0/22 events (0%) and 5/13 events (38.5%) in the absence/presence of hematogenous metastasis, respectively ($p = 0.002$). Hematogenous metastasis was also developed in a greater number of deletions in 9p21.3 (0/22 events (0%) and 3/13 events (23.1%) in the absence/presence of hematogenous metastasis, respectively) ($p = 0.053$). *These loci could serve as prognostic markers of hematogenous metastasis in patients with luminal B HER2– BC receiving anthracycline-based NAC regimens.*

Figure 3 shows CNA frequencies in BC patients receiving anthracycline-based regimens as part of preoperative chemotherapy before and after treatment, depending on the presence/absence of hematogenous metastasis.

To assess the metastasis-free survival (MFS) rate depending on the identified potential prognostic CNA markers of hematogenous metastasis after the use of anthracycline-based NAC regimens, we plotted survival curves using the Kaplan–Meier method.

Figure 4 presents the curves of the BC patients included in the study, depending on the presence of 9p22.2 amplifications ($p = 0.003$) and 9p21.3 deletions ($p = 0.03$) in tumor.

DISCUSSION

It is important to note that results of numerous studies on the search for prognostic and predictive markers in various adjuvant and neoadjuvant chemotherapy regimens against known molecular subtypes of BC have been published to date. The data of these studies are rather contradictory and aimed at a reevaluation of the use of available agents, including anthracyclines, while the attention is focused on the search for pCR markers.

In particular, centrosome duplication on chromosome 17 (*CEP17* duplication) was studied as a marker of sensitivity to anthracyclines. An increased *CEP17* copy number is often found in BC [11, 12]. Analysis of BC samples for the presence of the *CEP17* duplication

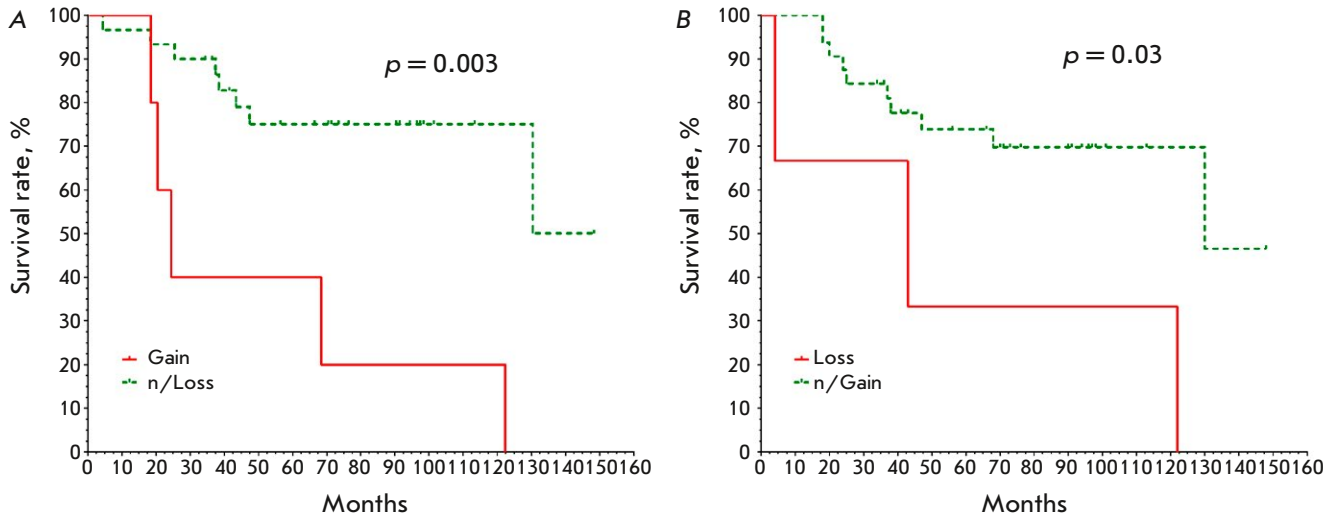


Fig. 4. Non-metastatic survival rates of breast cancer patients depending on the presence of amplifications in the 9p22.2 locus (*A*) and deletions in the 9p21.3 locus (*B*) in tumor

at various rates (e.g. > 1.86 *CEP17*/cell, HER2/*CEP17* ⩾ 2.0) was performed by several research groups and yielded contradictory results; either the presence or an absence of a linear relationship between the HER2/CEP17 ratio and pCR [11, 13, 14].

The dynamic changes in the HER2− BC genetic landscape were also studied. In particular, a series of prospective molecular profiling of HER2− BC tumors was conducted during chemotherapy. Tumor biopsies were obtained before and after 2 weeks of chemotherapy (doxorubicin/cyclophosphamide, ddAC). The tumor samples were obtained during surgery, after 8 weeks of combination therapy. To assess the single nucleotide variants (SNVs) and CNAs of 440 tumor-associated genes (ACTOnco®), NGS sequencing of the DNA of each patient (n = 34) was carried out at three time points. New mutations that developed due to therapy were found in 13% of cases (after one treatment cycle). A total of 72% of patients exhibited changes in variant allele frequencies (VAFs) of pathogenic SNVs: 51% of these changes developed early (after 2 weeks of therapy) and persisted for 8 weeks. The changes in SNV VAF were mostly associated with the PI3K/mTOR/AKT pathway. Tumors with a poor response to treatment (50% events, [7/14]) were less likely to develop SNV VAF compared to tumors with a good response (15% events, [4/24], p = 0.029). No significant difference in the CNA was noted between patients with a good and poor response after 2 weeks of therapy (22 [0–100] vs. 35 [0–106] events, respectively, p = 0.605). However, after 8 weeks, patients with a good response had a lower CNA load compared to patients with a poor response (12 [3–26] vs. 32 [15–73] events, respectively, p = 0.042) [15].

The results of an integrated multi-omics profiling of high-grade HER2− BC were presented. Identification of metastatic candidate driver events in stage III ER+HER2− tumors based on primary (n = 270) and metastatic diseases after treatment (n = 243) revealed amplification of 8q24.13 and 8q24.21 in 44.5% of metastatic cases [16].

We performed a search for potential predictive CNA markers of objective response to NAC and markers of hematogenous metastasis. In particular, we found deletions in the **11q22.3** and **11q23.1** loci as potential predictive markers of objective response to anthracycline-based NAC regimens as part of preoperative chemotherapy in patients with luminal B HER2− BC. Patients with deletions in these loci are statistically significantly more likely to develop an objective response to anthracycline-based NAC than those without deletions (p = 0.004).

In a study by Elin Barnekow et al., the **11q22.3** locus is considered one of the eight most important BC susceptibility loci [17]. This locus was identified as a new risk locus (the most significant SNP is rs228595, $p = 7 \times 10^{-6}$) in *BRCA1* mutation carriers [18]. The 11q22.3 locus contains several genes, including *ACAT1*, *NPAT*, and *ATM* (according to the data obtained using the genecards.org database).

It is important to note that recent studies indicate new and surprising functions for *ACAT1*, which encodes acetyl-CoA acetyltransferase. ACAT1 has lysine acetyltransferase activity, and it acetylates pyruvate

dehydrogenase, which contributes to the Warburg effect and tumor cell proliferation [19]. According to recent data, ACAT1-mediated acetylation of METTL3 inhibits cell migration and invasion in TN BC [20]. It was also shown that inhibition of NPAT (nuclear protein, transcription coactivator) and p-NPAT prevents BC from entering the S phase of the cell cycle due to reduced DNA synthesis [21]. The common c.7271T>G mutation in *ATM* increases the risk of BC fourfold [22]. The role of *ATM* in BC has been studied in detail. *ATM* mutations have been found to correlate with certain clinical characteristics of BC such as lymph node involvement and HER2+ phenotype. *ATM* mutations are generally associated with a poor BC prognosis. In addition, since mutations in the *ATM*-encoding protein are involved in DNA repair mechanisms, *ATM* aberrations may also enhance the sensitivity of BC cells to platinum-based drugs and *PARP* inhibitors. Some data point to an association between *ATM* mutations and resistance of luminal positive BC cells to CDK4/6 inhibitors [23]. In our study, deletion of the *ATM* gene locus resulted in increased sensitivity of HER2− BC cells to anthracyclines.

We also showed that **9p22.2** and **9p21.3** could be considered prognostic makers of hematogenous metastasis in luminal B HER2− BC patients receiving anthracycline-based NAC regimens.

Both the amplifications in the **9p22.2** locus and deletions in the **9p21** locus (in contrast to its normal and amplified states) are considered unfavorable prognostic markers ($p = 0.03$).

Deletions of the short arm of chromosome 9 were shown to be associated with such aggressive BC signs as a highly malignant phenotype and a shortened survival time. The 9p deletions usually involve large fragments or even the entire chromosome arm [24]. The 9p21 deletions are associated with an unfavorable BC phenotype. In particular, 9p21 was found in 15.3% of 1,089 analyzed cases and associated with an unfavorable disease course, including a highly malignant phenotype ($p < 0.0001$), lymph node metastasis ($p = 0.0063$), and a high Ki67 index ($p < 0.0001$). The presence of the 9p21 deletion was associated with a poor disease outcome ($p = 0.0720$) [25].

According to published data, a homozygous deletion of the 9p21.3 locus is found in 15% of all human cancer diseases [26]. Recently, Han et al. analyzed large-scale genomic data presented in the Cancer Genome Atlas (TCGA) and showed that the 9p21.3 deletion is a marker of a poor prognosis in several cancers, including BC. The study demonstrated a clear association between homozygous 9p21.3 deletion and shorter overall survival time [27].

CONCLUSION

In the present study, we analyzed changes in the CNA landscape in breast cancer patients receiving anthracycline-based neoadjuvant chemotherapy regimens. We found potential predictive CNA markers of objective response (frequencies of deletions in loci **11q22.3** and **11q23.1**) and prognostic CNA markers of hematogenous metastasis (amplifications in the 9p22.2 locus and deletions in the 9p21.3 locus) in luminal B HER2− BC patients receiving anthracycline-based NAC regimens as part of preoperative chemotherapy. The obtained results are partially confirmed by the literature. However, validation of the obtained results is required in order to use the identified predictive and prognostic markers. ●

This research was supported by the Russian Science Foundation grant No. № 22-25-00499.

REFERENCES

1. Spring L.M., Bar Y., Isakoff S.J. // J. Natl. Compr. Canc. Netw. 2022. V. 20. № 6. P. 723–734.
2. Spring L.M., Fell G., Arfe A., Sharma C., Greenup R., Reynolds K.L., Smith B.L., Alexander B., Moy B., Isakoff S.J., et al. // Clin. Cancer Res. 2020. V. 26. P. 2838–2848.
3. LeVasseur N., Sun J., Gondara L., Diocee R., Speers C., Lohrisch C., Chia S. // J. Cancer Res. Clin. Oncol. 2020. V. 146. P. 529–536.
4. Yee D., De Michele A.M., Yau C., Symmans W.F., Albain K.S., Yi Chen Y., Krings G., Wei S., Harada S., et al. // JAMA Oncol. 2020. V. 6. № 9. P. 1355–1362.
5. Cardoso F., Senkus E., Costa A., Papadopoulos E., Aapro M., André F., Harbeck N., Aguilar Lopez B., Barrios C.H., Bergh J., et al. // Ann. Oncol. 2018. V. 29. P. 1634–1657.
6. Melendez G.C., Hundley W.G. // Circ. Cardiovasc. Imaging. 2016. V. 9. №12. e005797
7. Armenian S.H., Lacchetti C., Barac A. // J. Clin. Oncol.: Off. J. Am. Soc. Clin. Oncol. 2017. V. 35. № 8. P. 893–911.
8. Wolff A.C., Blackford A.L., Visvanathan K., Rugo H.S., Moy B., Goldstein L.J., Stockerl-Goldstein K., Neumayer L., Langbaum T.S., et al. // J. Clin. Oncol.: Off. J. Am. Soc. Clin. Oncol. 2015. V. 33. № 4. P. 340–348.
9. National Comprehensive Cancer Network Clinical Practice Guidelines in Oncology: Breast Cancer. v3.2021. https://www.nccn.org/professionals/physician_gls/pdf/breast.pdf.
10. Vuger A.T., Tiscoski K., Apolinario T., Cardoso F. // The Breast. 2022. V. 65. P. 67–76.
11. Rakha E.A., Miligy I.M., Quinn C.M., Provenzano E., Shaaban A.M., Marchiò C., Toss M.S., Gallagy G., Murray C., Walshe J., et al. // Br. J. Cancer. 2021. V. 124. P. 1836–1842.
12. Hoda R.S., Brogi E., Xu J., Ventura K., Ross D.S., Dang

MC., Robson M., Norton L., Morrow M., Wen H.Y. // Arch. Pathol. Lab. Med. 2020. V. 144. № 5. P. 597–601.
13. Bartlett J.M., McConkey C.C., Munro A.F., Desmedt C., Dunn J.A., Larsimont D.P., O'Malley F.P., Cameron D.A., Earl H.M., Poole C.J., et al. // J. Clin. Oncol. 2015. V. 33. P. 1680–1687.
14. Greenwell K., Hussain L., Lee D. Bramlage M., Bills G., Mehta A., Jackson A., Wexelman B. // Breast Cancer Res. Treat. 2020. V. 181. P. 249–254.
15. Choo J.R.E., Jan YH., Ow S.G.W., Wong A., Lee M.X., Ngoi N., Yadav K., Lim J.S.J., Lim S.E., Chan C.W., et al. // Targ. Oncol. 2022. V. 17. P. 355–368.
16. Wang K., Li L., Franch-Expósito S., Le X., Tang J., Li Q., Wu Q., Bassaganyas L., Camps J., Zhang X., et al. // Mol. Oncol. 2022. V. 16. № 12. P. 2413–2431.
17. Barnekow E., Hasslow J., Liu W., Bryant P., Thutkawkorapin J., Wendt C., Czene K., Hall P., Margolin S., Lindblom A. // Internat. J. Mol. Sci. 2023. V. 24. № 5. P. 44–68.
18. Hamdi Y., Soucy P., Kuchenbaeker K.B., Pastinen T., Droit A., Lemaçon A., Adlard J., Aittomäki K., Andrulis I.L., Arason A., et al. // Breast Cancer Res. Treat. 2017. V. 161. P. 117–134.
19. Goudarzi A. // Life Sci. 2019. V. 232. P. 116592.
20. Zhang G., Huang R., Zhao H., Xia Y., Huang H., Qian M., Fu Y., Cui Y. // Genes Immunity. 2023. V. 24. P. 99–107.
21. Ismail A., El-Mahdy H.A., Abulsoud A.I., Sallam A.-A.M., Gomaa Eldeib M., Elsakka E.G.E., Bakr Zaki M., Doghish A.S. // Internat. J. Biol. Macromolecules. 2023. V. 224. P. 1541–1565.
22. Hall M.J., Bernhisel R., Hughes E., Larson K., Rosenthal E.T., Singh N.A., Lancaster J.M., Kurian A.W. // Cancer Prev. Res. 2021. V. 14. P. 433–440.
23. Stucci L.S., Internò V., Tucci M., Perrone M., Mannavola F., Palmirotta R., Porta C. // Genes. 2021. V. 12. P. 727.
24. Jong Y.J., Li L.H., Tsou M.H., Chen Y.J., Cheng S.H., Wang-Wuu S., Tsai S.F., Chen C.M., Huang A.T., Hsu M.T., et al. // Cancer Genet. Cytogenet. 2004. V. 148. P. 55–65.
25. Lebok P., Roming M., Kluth M., Koop C., Özden C., Taskin B., Hussein K., Lebeau A., Witzel I., Wölber L., et al. // Oncotarget. 2016. V. 7. P. 81322–81331.
26. R. Beroukhim, Mermel C.H., Porter D., Wei G., Raychaudhuri S., Donovan J., Barretina J., Boehm J.S., Dobson J., Urashima M., et al. // Nature. 2010. V. 463. P. 899–905.
27. Han G., Yang G., Hao D., Lu Y., Thein K., Simpson B.S., Chen J., Sun R., Alhalabi O., Wang R., et al. // Nat. Commun. 2021. V. 12. P. 5606.

Analysis of the Association between the *Tgfb1* Gene Haplotype and Liver Diseases in Children

R. M. Kurabekova[1*], O. E. Gichkun[1,2], O. M. Tsirulnikova[1,2], I. E. Pashkova[1], V. A. Fomina[2], O. P. Shevchenko[1,2], S. V. Gautier[1,2]

[1]V.I. Shumakov National Medical Research Center of Transplantology and Artificial Organs, Moscow, 123182 Russian Federation

[2]I.M. Sechenov First Moscow State Medical University (Sechenov University), Moscow, 119991 Russian Federation

*E-mail: kourabr@yandex.ru

Received: May 19, 2023; in final form, July 28, 2023

DOI: 10.32607/actanaturae.19425

Copyright © 2023 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **Transforming growth factor-β1 (TGF-β1), a cytokine with immunosuppressive and pro-fibrogenic activity, is a potential marker of infection, liver transplant rejection, and fibrosis. Its levels in the blood and tissues depend on many factors; however, the role of gene polymorphism is still unclear. In this work, the distribution frequency of three single nucleotide polymorphism (SNP) variants of the *Tgfb1* gene, namely rs1800469, rs1800470, and rs1800471, was studied in children with end-stage liver disease (ESLD). The study included 225 pediatric liver recipients aged 1 month to 16 years (median, 8 months), including 100 boys and 125 girls, and 198 healthy individuals aged 32.7 ± 9.6 years, including 78 men and 120 women. The indication for liver transplantation in children was ESLD, which was mostly caused by congenital and inherited liver diseases. SNPs were detected by real-time polymerase chain reaction using TaqMan probes and DNA isolated from peripheral blood. SNP frequency distribution was in Hardy–Weinberg equilibrium and did not differ between children with liver diseases and the healthy ones. Analysis of the SNPs frequency based on allelic interaction models did not reveal any differences between patients and the healthy individuals. Evaluation of linkage disequilibrium for *Tgfb1* polymorphic variant pairs revealed a statistically significant linkage between all studied variants. Seven haplotypes, which are variants of SNP combinations, were observed in the studied groups of patients and healthy individuals. A total of 80% of the group had three haplotypes, whose frequencies did not differ between patients and the healthy individuals. Significant differences were found in the frequency of the haplotypes A-A-C, G-G-C, and G-A-G (at rs1800469, rs1800470, and rs1800471, respectively), which were observed up to 11 times more often in recipients compared to the healthy individuals. It is possible that these haplotypes are ESLD-predisposing variants, which may also contribute to the development of complications after liver transplantation in children.**

KEYWORDS **congenital and inherited liver diseases, biliary atresia and hypoplasia, pediatric liver recipients, liver transplantation.**

ABBREVIATIONS **TGF-β1 – transforming growth factor β1; *Tgfb1* – TGF-β1 gene; SNP – single nucleotide polymorphism; ESLD – end-stage liver disease; OR – odds ratio; BA – biliary atresia; BH – biliary hypoplasia; LD – linkage disequilibrium; HLA – human leukocyte antigen.**

INTRODUCTION

Liver transplantation is the only effective treatment for children with end-stage liver disease (ESLD) [1]. Despite advances in organ transplantation and the high survival rates of recipients, the period after transplantation can be accompanied by such complications as infections, transplant rejection, and fibrosis. Prevention of complications requires the use of accurate approaches to their prediction and diagnosis using molecular and genetic markers.

Transforming growth factor-β1 (TGF-β1), an important component of the immune system with im-

munosuppressive and pro-fibrogenic activities, plays an essential role in the development of complications after organ transplantation [2, 3].

A series of studies, including our work, showed that the TGF-β1 cytokine level in pediatric liver recipients correlates with the graft function and has a prognostic and diagnostic significance [4–6]. TGF-β1 levels in recipient's blood and tissues are determined by numerous factors, including genetic ones. Considering that ESLDs in children are mainly congenital and hereditary, this genetic factor can play some role in both the progression of the disease and emergence of post-transplant complications.

Tgfb1 contains a series of single nucleotide polymorphisms (SNPs), which can be associated with various pathologies [7–9]. The greatest interest of researchers specializing in the field of solid organ transplantation has been drawn to the following three SNPs in *Tgfb1*: rs1800469 with a C > T substitution in the promoter region (C(–509)T); rs1800470 with a T > C substitution in codon 10 of the first exon (T(+869)C), resulting in Leu-to-Pro substitution in the protein; and rs1800471 with a C > G substitution in codon 25 of the first exon (C(+915)G), leading to a Arg-to-Pro substitution in the protein. The rs1800469 polymorphism is located in the promoter region and affects the recruitment of transcription factors, thus disrupting transcription regulation. SNPs rs1800470 and rs1800471 are located in the first exon and affect protein expression. These SNPs are considered to be the cause of the differences in the TGF-β1 activity level in tissues and can be associated with the development of post-transplant complications [10–12]. The role of the *Tgfb1* polymorphism in pediatric liver diseases is still unknown.

The aim of this study is to evaluate the frequencies of three *Tgfb1* SNPs in young children with ESLD.

Establishment of the role of the polymorphism of the genes determining the activity of pro- and anti-inflammatory cytokines, including TGF-β1, in the pathogenesis of various diseases in solid organ recipients will make it possible to both predict the risk of developing the disease and its severity and select a therapeutic approach for an individual patient. An example of polymorphism analysis in clinical practice is the genotyping of human major histocompatibility complex genes and further selection of a recipient-compatible donor organ for transplantation. Another example is the polymorphism of *CYP3A5*, which encodes a member of the cytochrome P450 superfamily that can disrupt functional protein synthesis and exert a significant effect on the clearance of the immunosuppressive drug tacrolimus. Selection of a tacrolimus daily dose that takes into account the *CYP3A5* genotype allows one to attain the desired drug concentration.

EXPERIMENTAL

The study protocol was approved by the local ethics committee of the V.I. Shumakov National Medical Research Center of Transplantology and Artificial Organs. Either patients or their guardians signed a written informed consent to participate in the study. The consent is stored in the patient's medical records.

The study included 225 pediatric liver recipients (100 boys and 125 girls) aged 1–192 months (16 years; median, 8 months) and 198 healthy individuals aged 32.7 ± 9.6 years (78 males and 120 females). This sample was considered an open Russian population, since the ethnicity of the study participants was not identified.

Liver diseases in patients included the following pathologies: biliary atresia (BA), Caroli disease, biliary hypoplasia (BH), Alagille syndrome, Byler disease, and other rare liver diseases like Crigler–Najjar syndrome, von Gierke disease, alpha-1 antitrypsin deficiency, tyrosinemia, fulminant autoimmune hepatitis, cryptogenic cirrhosis, etc. The demographic and clinical characteristics of the pediatric liver recipients included in the study are presented in *Table 1*.

The patients included in the study underwent transplantation of a liver fragment from a living related donor. Recipients received two- and three-component immunosuppressive therapy, which included tacrolimus, corticosteroids, and mycophenolate drugs. Routine examination and treatment of patients were conducted in accordance with the clinical recommendations of the Russian Transplant Society and the protocols of the V.I. Shumakov National Medical Research Center of Transplantology and Artificial Organs.

Table 1. Patients included in the study

Characteristic	Value
Number of patients, *n*	225
Age, months	8 (1–192)
Male/female ratio (%)	100(44)/125(56)
Disease, number of cases (%)	
BA	107(48)
BH	24(11)
Caroli disease	11(5)
Alagille syndrome	12(5)
Byler disease	10(4)
Others	61(27)

Note: BA – biliary atresia; BH – biliary hypoplasia.

Genomic DNA was isolated from peripheral blood using a QIAamp DNA Blood Mini Kit (Qiagen, Germany) on an automated QIAcube™ system (Qiagen, Germany) according to the manufacturer's instructions. The *Tgfb1* polymorphic variants rs1800469, rs1800470, and rs1800471 were analyzed by real-time polymerase chain reaction using TaqMan probes (Applied Biosystems, USA) on a CFX96™ real-time PCR detection system (Bio-Rad, USA) according to the manufacturer's instructions.

The statistical analysis was performed using the Microsoft Excel software. The distribution of the studied SNP genotype frequencies, haplotype structure, and pairwise linkage disequilibrium were analyzed using the SNPstats software [13]. In order to confirm the independent distribution of the studied polymorphic alleles, we tested them for compliance with the Hardy–Weinberg law. Allele frequency was calculated using the following formula: allele frequency = ((2 × number of homozygotes) + number of heterozygotes))/2 × total number of individuals. The frequencies of the genotypes and individual alleles were compared between different groups using the Pearson $\chi 2$ criterion. To quantitatively represent the impact of a possible genotype on a certain characteristic, odds ratios (ORs) and their 95% confidence intervals (CI) were calculated. To assess the linkage disequilibrium, *D*-statistics and the correlation coefficient *r* were calculated. The critical significance level was considered 0.05.

RESULTS

Three *Tgfb1* polymorphic variants (rs1800469, rs1800470, and rs1800471) were genotyped in the DNA of the studied patients. The frequencies of the various genotypes and alleles were calculated. *Figure 1* presents the distribution of the genotypes and alleles in children with liver diseases and healthy individuals.

A comparative analysis of the frequencies of the studied genotypes and alleles in the children with liver diseases and healthy donors did not reveal any statistically significant differences.

No gender-related statistically significant differences in the distribution of the studied SNPs were found between the patients and healthy individuals. At the same time, significant differences were observed in the frequency of the rare genotype C/G rs1800471 between young female patients and healthy women; this genotype was 3.96 times more common in girls than in healthy women (OR, 3.96, 95% CI 1.09–14.43, $p = 0.01$).

The frequency distribution of the studied genotypes of all three polymorphic variants was in agreement with the Hardy–Weinberg equilibrium in both the patients and healthy individuals (*Table 2*).

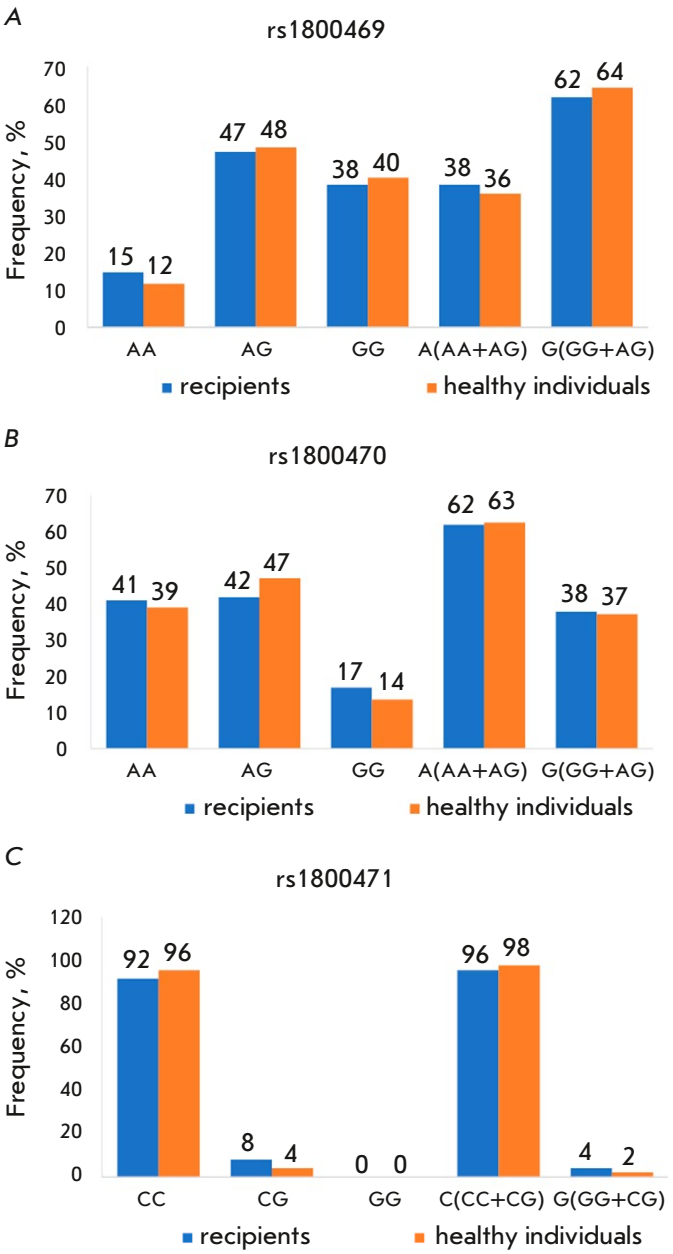


Fig. 1. Frequency distributions of *Tgfb1* SNPs rs1800469 (*A*), rs1800470 (*B*), and rs1800471 (*C*) in children with liver diseases and healthy individuals

A comparative analysis of the distribution of the genotype and allele frequencies of three *Tgfb1* SNPs using different models of allelic interactions (codominant, dominant, recessive, and superdominant) was conducted in healthy individuals and pediatric liver recipients. For each model, OR and the error rate

Table 2. Compliance of *Tgfb1* polymorphism distributions with the Hardy–Weinberg law

Group	rs1800469		rs1800470		rs1800471	
	χ^2	p	χ^2	p	χ^2	p
Healthy individuals	0.4779	0.54	0.0166	0.898	0.0842	0.772
Patients	0.0013	0.971	2.6579	0.120	0.3925	0.531

Note: $p < 0.05$ do not meet the Hardy–Weinberg equilibrium.

Table 3. Comparison of *Tgfb1* polymorphism distributions in children with liver diseases and healthy individuals using different inheritance models

Model	Genotype	Frequency, recipients	Frequency, healthy individuals	OR (95% CI)	p
rs1800469					
Codominant	G/G	85 (38%)	74 (39.8%)	1.00	0.71
	A/G	106 (47.3%)	90 (48.4%)	1.02 (0.67–1.55)	
	A/A	33 (14.7%)	22 (11.8%)	1.29 (0.69–2.41)	
Dominant	G/G	85 (38%)	74 (39.8%)	1.00	0.73
	A/G-A/A	139 (62%)	112 (60.2%)	1.07 (0.72–1.60)	
Recessive	G/G-A/G	191 (85.3%)	164 (88.2%)	1.00	0.41
	A/A	33 (14.7%)	22 (11.8%)	1.27 (0.71–2.28)	
Superdominant	G/G-A/A	118 (52.7%)	96 (51.6%)	1.00	0.83
	A/G	106 (47.3%)	90 (48.4%)	0.96 (0.65–1.41)	
rs1800470					
Codominant	A/A	91 (40.8%)	74 (39.8%)	1.00	0.49
	A/G	94 (42.1%)	87 (46.8%)	0.87 (0.57–1.33)	
	G/G	38 (17%)	25 (13.4%)	1.23 (0.68–2.22)	
Dominant	A/A	91 (40.8%)	74 (39.8%)	1.00	0.81
	A/G-G/G	132 (59.2%)	112 (60.2%)	0.95 (0.64–1.42)	
Recessive	A/A-A/G	185 (83%)	161 (86.6%)	1.00	0.32
	G/G	38 (17%)	25 (13.4%)	1.32 (0.76–2.28)	
Superdominant	A/A-G/G	129 (57.9%)	99 (53.2%)	1.00	0.33
	A/G	94 (42.1%)	87 (46.8%)	0.82 (0.56–1.22)	
rs1800471					
Codominant	C/C	205 (91.9%)	180 (96.3%)	1.00	0.063
	C/G	18 (8.1%)	7 (3.7%)	2.26 (0.92–5.53)	

were calculated by comparing children with liver diseases and healthy individuals (*Table 3*).

No statistically significant differences were found in the distribution of the studied *Tgfb1* variant frequencies between patients and healthy individuals using different allelic interaction models. In addition, no significant gender-related differences in genotype distribution were noted.

Since the analyzed loci are located on the same chromosome, linkage disequilibrium (LD), i.e., linked locus inheritance and haplotype formation, can be observed. *Table 4* presents the results of the statistical analysis of pairwise linkage of the studied *Tgfb1* variants as *D*, *D*', and *r*-statistics, including the error rate.

A statistically significant linkage was found between all studied variants. The strongest linkage

Table 4. Statistical assessment of the linkage disequilibrium for pairs of polymorphic variants of the *Tgfb1* gene

SNP pair	*D*	*D'*	*r*	*p*
rs1800469–rs1800470	0.1447	0.6259	0.6184	0
rs1800469–rs1800471	-0.0113	0.9934	-0.136	0.0001
rs1800470–rs1800471	0.0089	0.4628	0.1062	0.0021

was observed for the first pair of loci (rs1800469–rs1800470); the weakest linkage was found between the other two pairs.

Seven combinations of the studied SNPs were found in the studied groups. *Table 5* presents the identified haplotypes in order of decreasing frequency, frequencies of various groups, the OR between healthy individuals and recipients, and the OR error rate.

Table 5 shows that the combination of G-A-C alleles is the most prevalent (about 50% of cases in patients and 60% of cases in healthy individuals), whose distribution does not differ significantly between the groups. The second most common haplotype, A-G-C, is present in approximately 30% of the studied groups; its frequency also did not differ between the patients and healthy individuals. The frequencies of the fifth most common haplotype, G-G-G, also did not differ between the groups: there were about 2% of cases in both groups. In general, about 80% of the studied individuals had three out of seven haplotype variants with the same frequencies in recipients and healthy individuals.

Statistically significant differences in the frequency of the least common haplotypes were found. These haplotypes are more prevalent in recipients compared to healthy individuals. Haplotypes No. 3 (A-A-C), No. 4 (G-G-C), and No. 6 (G-A-G) are 3.12 (p = 0.0002), 2.88 (p = 0.0008), and 11.18 (p = 0.025) times more prevalent in patients than in healthy individuals, respectively. In general, the less common haplotypes No. 3, 4, and 6 were found in 26.84% of the patients and 7.71% of the healthy individuals. The least common haplotype A-G-G (No. 7) was practically absent in healthy individuals, while its frequency in recipients was <1%, which made it impossible to compare the groups for this parameter.

DISCUSSION

The development of approaches to the non-invasive diagnosis of post-transplant complications is important in relation to pediatric liver recipients due to the high risk of invasive procedures. Gene diagnosis has such important advantages over other methods as independence from the physiological state, immutability, and the possibility to perform only a single test. The results of genetic tests can provide information on a patient's predispositions and allow for the use of personalized therapy by selecting drugs based on an individual patient's characteristics.

In this work, we analyzed the frequencies of the three most prevalent *Tgfb1* SNPs in children with ESLD and healthy individuals in an open Russian population. We showed that the distribution of these SNPs does not differ between patients and healthy individuals and meets the Hardy–Weinberg equilibrium.

The data we obtained did not reveal an association between different *Tgfb1* SNPs and pediatric liver diseases. We have not found publications on the study of the *Tgfb1* genetic polymorphism in young children with congenital and hereditary liver diseases in Russian and other populations. The role of *Tgfb1* SNPs in the development of various liver pathologies has been studied in adult patients; however, the results of these studies are inconclusive [7–12].

Table 5. *Tgfb1* haplotype frequencies in children with liver diseases and healthy individuals

No.	Nucleotide			Frequency			Odds ratio (95% CI)	*p*
	rs1800469	rs1800470	rs1800471	total	patients	healthy individuals		
1	G	A	C	0.5236	0.4680	0.5864	1.00	
2	A	G	C	0.2841	0.2410	0.3190	1.05 (0.75–1.47)	0.76
3	A	A	C	0.0862	0.1340	0.0374	3.12 (1.72–5.67)	0.0002*
4	G	G	C	0.0754	0.1170	0.0371	2.88 (1.56–5.32)	0.0008*
5	G	G	G	0.0180	0.0154	0.0176	1.54 (0.51–4.65)	0.44
6	G	A	G	0.0127	0.0174	0.0026	11.18 (1.37–91.18)	0.025*
7	A	G	G	0.0038	0.0073	0	–	–

*$p < 0.05$.

On the one hand, there is data indicating an association between these SNPs and transplant rejection and chronic dysfunction [10–12]. On the other hand, there are also studies that did not uncover any association between the *Tgfb1* polymorphism and both transplant rejection and donor liver fibrosis in adult patients [14–16].

The frequencies of SNPs rs1800469, rs1800470, and rs1800471 in healthy individuals identified in our study are consistent with the data of other domestic authors [17, 18]. A comparison of the distribution of the allele frequencies studied in our work with the data deposited into the U.S. National Center for Biotechnology Information (NCBI) also did not reveal significant differences from that of *Tgfb1* in the European population: rs1800469 – A(37%)/G(63%); rs1800470 – A(56%)/G(44%); and rs1800471 – C(94%)/G(6%).

The TGF-β1 cytokine is a vital protein involved in the regulation of the key cellular processes; therefore, a significant impairment of its function can be fatal [19]. It is possible that individual single nucleotide substitutions have a limited effect on the protein function, while a combination of several substitutions can have a pronounced effect. Therefore, analysis of the haplotypes of several SNPs can be more informative than a study of a single nucleotide substitution.

We noticed a linkage disequilibrium of SNPs rs1800469, rs1800470, and rs1800471, which form seven haplotype variants, in patients and healthy individuals. The prevalence of the three most common haplotypes did not vary significantly between patients and healthy individuals. The analysis of the same *Tgfb1* haplotypes conducted by other authors showed a similar prevalence of the most common haplotype, G-A-C, in healthy individuals, which was about 50–60% [17, 20].

The least common haplotypes identified in our study were more prevalent in ESLD patients compared to the healthy individuals. This suggests involvement of these haplotypes in the predisposition to liver diseases. A significant number of diseases in the studied patients are congenital and hereditary, while the genetic nature of the majority has not been studied in detail. Therefore, a search for disease-associated haplotypes can be of scientific and practical value in transplantation. It is possible that the identified haplotypes not only predispose children to the liver disease, but also contribute to the complications that emerge after liver transplantation. However, further studies are required to unambiguously establish such a causal relationship.

The study design is based on the case–control method, which imposes certain limitations on the legitimacy of establishing a causal relationship between the identified associations. It should be noted that it is not always possible to unambiguously determine the haplotype based on the genotype using PCR. Only sequencing allows for accurate haplotype identification.

The limitation in the conclusion on a possible association between the studied haplotypes and predisposition to ESLD is also due to the fact that some pathologies can be determined by numerous genetic factors/polymorphisms, each of which makes only a small contribution to the overall risk of developing the disease, while their significance is difficult to evaluate when analyzing a small patient sample. For instance, the functional activity of TGF-β1 can be determined by not only the gene polymorphism, but also by the genetic variants of the other factors participating in the cytokine pathway, such as binding proteins and receptors. The risk of developing hepatitis C after transplantation in patients was shown to be associated with the frequency of the SNP rs868 located in the non-coding 3′-UTR region of the TGF-β1 receptor gene (*Tgfbr1*) [21]. In addition, the interaction of different genes can have a clinical significance. An association between *HLA* genes and the genes of various cytokines, including TGF-β1, was found in patients with such a multifactorial autoimmune disease as type 1 diabetes mellitus, in which polymorphisms of the human major histocompatibility complex genes may play an important role [22]. Some combinations of polymorphic variants of the cytokine TNF-α, IFN-γ, IL-6, and TGF-β1 genes were shown to be less common in patients with type 1 diabetes mellitus compared to the control, which suggested a protective role for these haplotypes [22]. Linkage disequilibrium of the TNF-α variant characteristic of the protective haplotype with two polymorphic HLA variants was also noted.

Predisposition to various polygenic diseases can also be determined by an individual's ethnicity, which points to the need to study ethnically homogeneous groups. However, we did not determine the ethnicity of the individuals in our study. Therefore, the obtained results can be considered valid for an open Russian population.

CONCLUSION

The level of the multifunctional cytokine TGF-β1 is a potential biomarker of infection, transplant rejection, and fibrosis. In this work, we studied the distribution of the three most significant *Tgfb1* polymorphisms (rs1800469, rs1800470, and rs1800471) in pediatric patients with congenital and hereditary liver diseases. We demonstrated that the frequency of each individual polymorphism does not differ significantly from

that of healthy individuals and meets the Hardy–Weinberg equilibrium.

However, the frequency of haplotypes of the three studied *Tgfb1* polymorphisms differs statistically significantly between patients and healthy individuals. Seven different haplotypes were found in the studied group. Of them, three were observed 3 to 11 times more often in recipient children compared to healthy ones. These haplotypes, namely A-A-C, G-G-C, and G-A-G, which correspond to rs1800469, rs1800470, and rs1800471, respectively, can be associated with predisposition to end-stage liver disease in children. Additional studies are warranted in order to further elucidate the role of these haplotypes in post-transplant complications. ●

REFERENCES

1. Gautier S.V. // Russian Journal of Transplantology and Artificial Organs. 2017. V. 19. № 3. P. 10–32.
2. Dudek K., Koziak K., Placha G., Kornaiewicz O., Zieniewicz K., Zurakowski J., Krawczyk M. // Transplant. Proc. 2009. V. 41. № 1. P. 240–245.
3. Zhang Y., Wang Y.L., Liu Y.W., Li Q., Yuan Y.H., Niu W.Y., Sun L.Y., Zhu Z.J., Shen Z.Y., Han R.F. // Transplant. Proc. 2009. V. 41. № 5. P. 1767–1769.
4. Briem-Richter A., Leuschner A., Krieger T., Grabhorn E., Fischer L., Nashan B., Haag F., Ganschow R. // Pediatr. Transplant. 2013. V. 17. № 8. P. 757–764.
5. Hussein M.H., Hashimoto T., Abdel-Hamid Daoud G., Kato T., Hibi M., Tomishige H., Hara F., Suzuki T., Nakajima Y., Goto T., et al. // Pediatr. Surgery Int. 2011. V. 27. № 3. P. 263–268.
6. Kurabekova R.M., Tsirulnikova O.M., Gichkun O.E., Pashkova I.E., Olefirenko G.A., Shevchenko O.P. // Russian Journal of Transplantology and Artificial Organs. 2018. V. 20. № 4. P. 38–43.
7. Dhaouadi T., Sfar I., Bardi R., Jendoubi-Ayed S., Abdallah T.B., Ayed K., Gorgi Y. // Transplant. Proc. 2013. V. 45. № 6. P. 2152–2157.
8. Kim Y.H., Kim T.H., Kang S.W., Kim H.J., Park S.J., Jeong K.H., Kim S.K., Lee S.H., Ihm C.G., Lee T.W., et al. // Immunol. Invest. 2013. V. 42. № 4. P. 285–295.
9. Paladino N., Flores A.C., Fainboim H., Schroder T., Cuarterolo M., Lezama C., Ballerga E.G., Levi D., Tanno H., Costanzo G., et al. // Clin. Immunol. 2010. V. 134. № 3. P. 305–312.
10. Zhang X.X., Bian R.J., Wang J., Zhang Q.Y. // Genet. Mol. Res. 2016. V. 15. № 2. P. 15027599.
11. Arrieta-Bolanos E., Mayor N.P., Marsh S.G., Madrigal J.A., Apperley J.F., Kirkland K., Mackinnon S., Marks D.I., McQuaker G., Perry J., et al. // Haematologica. 2016. V. 101. № 3. P. 382–390.
12. Benza R.L., Coffey C.S., Pekarek D.M., Barchue J.P., Tallaj J.A., Passineau M.J., Grenett H.E. // J. Heart Lung Transplant. 2009. V. 28. № 10. P. 1057–1062.
13. Solé X., Guinó E., Valls J., Iniesta R., Moreno V. // Bioinformatics. 2006. V. 22. № 15. P. 1928–1929.
14. Liu K., Liu X., Gu S., Sun Q., Wang Y., Meng J., Xu Z. // Oncotarget. 2017. V. 8. № 37. P. 62463–62469.
15. Eurich D., Bahra M., Boas-Knoop S., Lock J.F., Golembus J., Neuhaus R., Neuhaus P., Neumann U.P. // Liver Transplant. 2011. V. 17. № 3. P. 279–288.
16. Ferrarese A., Sartori G., Orrù G., Frigo A.C., Pelizzaro F., Burra P., Senzolo M. // Transplant. Int. 2021. V. 34. № 3. P. 398–411.
17. Barsova R., Titov B., Matveeva N., Favorov A., Rybalkin I., Vlasik T., Tararak E., Sukhinina T., Shakhnovich R., Ruda M. // Acta Naturae (in Russian). 2012. V. 4. № 2. P. 76–82.
18. Razvodovskaya A.V., Cherkashina I.I., Nikulina S.Yu., Shestovitskiy V.A., Voyevoda M.I., Maksimov V.N., Aver'yanov A.B. // Bulletin of physiology and pathology of respiration. 2014. № 54. P. 23–29.
19. Li M.O., Wan Y.Y., Sanjabi S., Robertson A.K., Flavell R.A. // Annu. Rev. Immunol. 2006. V. 24. P. 99–146.
20. Iriyoda T.M.V., Flauzino T., Costa N.T., Lozovoy M.A.B., Reiche E.M.V., Simão A.N.C. // Clin. Exp. Med. 2022. V. 22. № 1. P. 37–45.
21. Sajjad E.A., Radkowski M., Perkowska-Ptasińska A., Pacholczyk M., Durlik M., Fedorowicz M., Pietrzak R., Ziarkiewicz-Wróblewska B., Włodarski P., Malejczyk J. // Annals of Transplantation. 2017. V. 22. P. 638-645.
22. Kumar R., Goswami R., Agarwal S., Israni N., Singh S.K., Rani R. // Tissue Antigens. 2007. V. 69. № 6. P. 557–567.

Effect of the *ati* Gene Deletion on the Pathogenicity and Immunogenicity of the Vaccinia Virus

S. N. Yakubitskiy, A. A. Sergeev, K. A. Titova, I. S. Shulgina, E. V. Starostina, M. B. Borgoyakova, L. I. Karpenko, S. N. Shchelkunov*
State Research Center of Virology and Biotechnology VECTOR, Rospotrebnadzor, Koltsovo, Novosibirsk region, 630559 Russian Federation
*E-mail: snshchel@rambler.ru
Received April 14, 2023; in final form, July 13, 2023
DOI: 10.32607/actanaturae.17872
Copyright © 2023 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **Among the nonvirion proteins of the vaccinia virus (VACV), a 94-kDa long protein is most abundantly present; the protein is a truncated form of the 150-kDa A-type inclusion (ATI) protein of the cowpox virus encoded by the *ati* gene. This VACV protein does not form intracellular ATIs, being as it is a major immunogen upon infection/immunization of humans or animals with the VACV. Antibodies specific to this protein are not virus-neutralizing. The present study focused on the effect of the production of this nonstructural major immunogenic VACV protein on the manifestation of pathogenicity and immunogenicity of the virus in the BALB/c mouse model of infection. In order to introduce a targeted deletion into the VACV LIVP genome, the recombinant integration/deletion plasmid pΔati was constructed and further used to generate the recombinant virus LIVPΔati. The pathogenicity of the VACV LIVP and LIVPΔati strains was studied in 3-week-old mice. The mice were intranasally infected with the viruses at a dose of 10^7 pfu; 50% of the animals infected with the parent LIVP strain died, while infection with the LIVPΔati strain led to the death of only 20% of the mice. Intradermal vaccination of mice aged 6–7 weeks with the LIVPΔati virus statistically significantly increased the production of VACV-specific IgG, compared to that after intradermal vaccination with VACV LIVP. Meanwhile, no differences were noted in the cell-mediated immune response to the vaccination of mice with VACV LIVP or LIVPΔati, which was assessed by ELISpot according to the number of splenocytes producing IFN-γ in response to stimulation with virus-specific peptides. Intranasal infection of mice with lethal doses of the cowpox virus or the ectromelia virus on day 60 post-immunization with the studied VACV variants demonstrated that the mutant LIVPΔati elicits a stronger protective response compared to the parent LIVP.**

KEYWORDS **orthopoxviruses, vaccinia virus, *ati* gene, intradermal injection, immunogenicity, protectivity.**
ABBREVIATIONS **CPXV – cowpox virus; ECTV – ectromelia virus; VACV – vaccinia virus; pfu – plaque-forming units; dpi – day post immunization; i.d. – intradermal; i.n. – intranasal.**

INTRODUCTION

The vaccinia virus (VACV) belongs to the genus *Orthopoxvirus* (family Poxviridae), which also comprises such virus species as the variola virus (VARV), monkeypox virus (MPXV), cowpox virus (CPXV), camelpox virus (CMLV), and others [1, 2]. Orthopoxviruses are the largest DNA-containing mammalian viruses whose entire life cycle occurs in the cytoplasm of infected cells. The members of this genus are morphologically indiscernible in terms of virion structure and antigenically close to each other; therefore, infection with one orthopoxvirus species affords immunity against other members of this genus [3]. For this very reason, smallpox has been eradicated using the live vaccine based on different VACV stains [1, 4].

Like for other orthopoxviruses, there exist two infectious forms of VACV. Intracellular mature virions (IMVs) make up the overwhelming majority of virus progeny, while a much smaller portion is represented by extracellular enveloped virions (EEVs) [5, 6].

The so-called LS antigen (an antigen complex consisting of the thermolabile (L) and thermostable (S) antigenic components), against which antibodies are intensely produced, was discovered in early studies focusing on the immune response elicited by intra-

dermally (i.d.) infecting rabbits with VACV [7]. The highly immunogenic LS antigen, which is a nonvirion protein and is abundantly present, is isolated from the extracts of infected animal tissue. The antibodies produced after the animals are immunized with LS antigen exhibit no virus-neutralizing activity but react to clinical samples collected from patients with smallpox and monkeypox in the complement binding and gel immunoprecipitation assays [2].

It was demonstrated later that the 94-kDa LS antigen of VACV is the truncated form of the protein that forms intracellular A-type inclusion bodies (ATIs) in CPXV infected cell. The ATI protein of CPXV is 150 kDa in size [8], is encountered in large quantities in infected cells during the later stage of the virus replication cycle (up to 4% of total cell protein), and is aggregated to yield gel-like bodies that may incorporate mature virions in their cytoplasm [9, 10]. C-terminally truncated forms of this protein 92–96 kDa in size are synthesized in large quantities by VARV, MPXV, CMLV, and VACV, without forming ATIs (*Fig. 1*).

It was demonstrated that immunization of laboratory animals or human volunteers with VACV results in the production of antibodies against a broad range of virion proteins, including the highly immunogenic nonvirion ATI-like protein [11]. The T-cell-mediated immune response to a VACV infection is mainly induced against early nonvirion proteins [6]. The only exception is the truncated ATI protein synthesized during the late stage of the virus replication cycle; nevertheless, a strong immune response is elicited to it [12]. A plausible reason is that the truncated *ati* gene (*A29L* in the case of VACV-LIVP, see *Fig. 1*) is one of the most intensely expressed VACV genes, and that the protein encoded by it is produced in the largest quantities among nonvirion proteins [9].

The *ati* gene is not among the essential genes for the VACV, since natural variants of the virus with this gene deleted have been discovered [13, 14]. It was also demonstrated for CPXV that *ati* gene deletion affects viral replication ability neither *in vitro* nor *in vivo* [15, 16]. However, the effect of *ati* gene deletion on the immunogenic properties of VACV still remains unclear.

This study focused on the impact of the deletion of the *ati* gene, which encodes the major immunogenic protein of VACV (antibodies specific to it exhibit no virus-neutralizing activity), on the pathogenicity and immunogenicity of the virus.

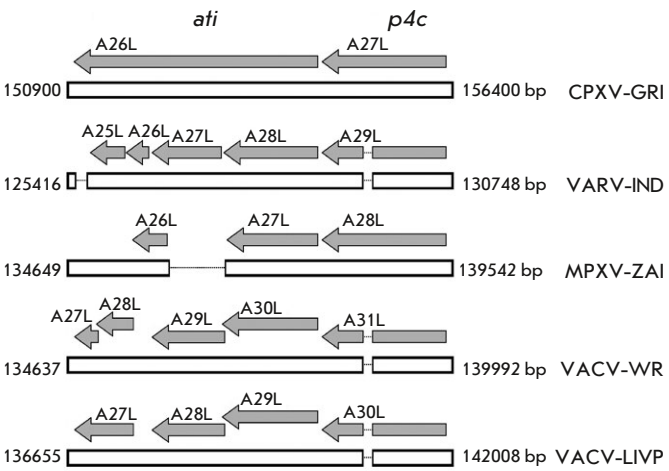


Fig. 1. Comparison of the potential orthopoxvirus open reading frames (ORFs) within the *ati* (ORF A26L) and *p4c* ((A27L) CPXV-GRI) genes. Gray arrows indicate the size and direction of the respective ORFs; their names are provided above these arrows for the CPXV-GRI, VARV-IND, MPXV-ZAI, VACV-WR [2], and VACV-LIVP viruses. Numbers to the right and left of white boxes denote the positions of the respective regions of the viral genomes. Thin lines indicate deletions in the viral genomes with respect to the CPXV-GRI DNA sequence. The truncated form of the ATI protein is encoded by ORF A29L VACV-LIVP and the respective ORFs of other viruses

MATERIALS AND METHODS

Bacteria, viruses, and cell culture

The *Escherichia coli* XL2-Blue strain and LIVP VACV cl. 14 obtained previously by finite dilution in agarose gel [17], as well as the cowpox virus (CPXV) strain GRI-90 and the ectromelia virus (ECTV) strain K-1, obtained from the virus collection, were used in this study. The passaged African green monkey kidney cell culture CV-1 was procured from the Cell Culture Collection of the State Research Center of Virology and Biotechnology VECTOR, Rospotrebnadzor. The viruses were grown and titrated using a monolayer CV-1 culture according to the procedure described in [18].

The Animals

BALB/c inbred mice procured from the husbandry of the Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences (Novosibirsk, Russia) were used for the study. The experimental animals were fed the standard diet with a sufficient amount of water, in compliance with the veterinary legislation and requirements for humane care and use of laboratory animals (State Standard GOST 33216-2014 “Guidelines for Accommodation and Care of Animals. Species-specific Provisions for Laboratory Rodents and Rabbits”). The studies

and manipulations involving animals were approved by the Bioethics Committee of the State Research Center of Virology and Biotechnology VECTOR, Rospotrebnadzor (Protocol No. 02-06.2022 dated June 28, 2022).

Pathogenicity assessment of the VACV strains

Three-week-old BALB/c mice (10 animals per group) were used in the study focusing on the pathogenicity of the VACV LIVP and LIVPΔati viruses (for intranasal (i.n.) infection). After inhalation anesthesia with diethyl ether, the mice received either a virus-containing fluid (50 μL, 10^7 plaque-forming units (pfu)) or saline into the nasal cavity. The animals were followed up for 22 days; clinical manifestations of infection and animal death were documented.

The following scoring system for disease symptoms was used: 0 – no disease signs; 1 – slightly unkempt hair coat; 3 – significantly unkempt hair coat, as well as back-arching or conjunctivitis; 4 – hard breathing or immobility; and 5 – death.

Each mouse was weighed individually every two days. The arithmetic mean body weights of the mice in each group at a given time point were calculated and expressed as a percentage of the initial weight.

Immunization of mice and sample collection for analyses

BALB/c mice aged 6–7 weeks were immunized with the LIVP and LIVPΔati VACV strains by making an intradermal (i.d.) injection into the dorsal side of the tail (~ 1 cm from the tail base), according to the procedure described earlier [19] using a virus dose of 10^5 pfu/20 μL/mouse. The mice that had received saline were used as negative controls.

The humoral immune response in the mice was analyzed on days 7, 14, 21, 28, 42, and 56 post immunization (dpi). At each time point specified above, six mice per group were taken for analysis. Mouse blood samples were intravitally collected from the retro-orbital venous sinus using a 23G needle. Blood sampling from the retro-orbital sinus does not cause visual organ injury. Blood sampling is a short-lasting but painful procedure; however, no analgesia was used, as it has been demonstrated that the known analgesic or anesthetic techniques may affect the immunological parameters of the animal's blood.

Sera were obtained from individual animal blood samples by precipitation of blood cells via centrifugation at a relative centrifugal force of 1,000 *g* during 10 min. The sera were exposed at 56°C during 30 min and stored at –20°C.

On 7, 14, and 21 dpi, after blood collection, the mice were euthanized by cervical dislocation; spleens were isolated from each out of six mice in the study groups under sterile conditions.

Splenocyte separation

Splenocytes were separated by passing individual spleens through a cell strainer (BD Falcon™, USA) using a syringe plunger. After red blood cells had been removed using a ACK lysing buffer (Thermo Fisher Scientific, USA), the splenocytes were washed and resuspended in a RPMI-1640 medium supplemented with 10% fetal bovine serum and gentamicin (50 μg/mL).

Quantification of IFN-γ-producing cells by ELISpot

The intensity of the T-cell-mediated immunity in the vaccinated mice was quantified by ELISpot according to the number of splenocytes producing IFN-γ. The assay was carried out using MABTECH kits (Sweden) in accordance with the manufacturer's instruction. Cells were stimulated using a mixture of VACV-specific peptides SPYAAGYDL, SPGAAGYDL, VGPSNSPTF, KYGRLFNEI, GFIRSLQTI, and KYMWCYSQV immunodominant for BALB/c mice (20 μg/mL of each peptide) [20, 21]. The counts of IFN-γ-producing cells were determined using an ELISpot reader (Carl Zeiss, Germany).

Enzyme-linked immunosorbent assay of murine sera

Enzyme-linked immunosorbent assay (ELISA) of mouse sera was carried out according to the procedure described in [18]. A preparation of LIVP VACV virions purified by sucrose cushion centrifugation was used as an antigen. All the analyzed mouse serum samples were titrated using a series of twofold dilutions (from 1 : 100 to 1 : 12,800). Titration was repeated the next day when conducting ELISA. The IgM and IgG titers were determined using solutions containing peroxidase conjugates of anti-mouse IgM and anti-mouse IgG (Sigma, USA), respectively. The IgM and IgG titers were determined in each analyzed serum sample (individually for each repeat and then averaged). The geometric mean logarithms of inverse titer of VACV-specific IgG or IgM were calculated for the study groups, and 95% confidence intervals were determined.

Assessment of protective efficacy in immunized mice

On 60 dpi, the groups of animals immunized with VACV LIVP or LIVPΔati or the ones in the control group were i.n. infected with CPXV GRI-90 at a dose of 2.0×10^6 pfu/50 μL/mouse (six animals per group) or ECTV K-1 at a dose of 2.2×10^3 pfu/50 μL/mouse (six animals per group). The animals were followed up for 14 days; clinical manifestations of infection and

deaths were documented. Each mouse was weighed every 2 days. The arithmetic mean body weights of the mice in each group at each time point were calculated and expressed as a percentage of the initial weight.

The data were obtained for the animal groups i.d. immunized with VACV LIVP or LIVPΔati, as well as the groups of non-immunized and non-infected mice (negative control) or mice infected with CPXV GRI-90 or ECTV K-1 (positive control).

Statistical analysis

Statistical analysis and data comparison were performed using the standard methods, employing the Statistica 13.0 software package (StatSoft Inc. 1984–2001). $P < 0.05$ were considered statistically significant.

RESULTS

Construction of the LIVPΔati virus

Targeted deletion in the VACV LIVP genome (GenBank: KX781953.1) was performed in accordance with the scheme shown in *Fig. 2*. At the first stage of constructing the recombinant integrative/deletion plasmid pΔati, we calculated and synthesized oligonucleotide primers for PCR and the amplification of flanking VACV LIVP DNA sequences adjacent to the left or right border of the viral genome region to be deleted (ORF A28L–A29L, position on the genome, 137618–140470 bp) using the Oligo software (version 3.3) (Borland International, USA) (*Figs. 1, 2*).

The left-border flanking fragment (L) was synthesized using the primer pair 5'-**AAGCTT**GTTTGGTAGTAGATACATATCAATATCATC-3' (**HindIII**) and 5'-**CTGCAG**GCTGACTCAATTGCATGAAGAT-3' (**PstI**); the right-border flanking fragment (R) was synthesized using the primers 5'-**CTGCAG**GGGTAATTATAAGATCGTAGATCTCC-3' (***PstI***) and 5'-**CCCGGG**ATGGCGAACATTATAAATTTATGG-3' (**XmaI**) and Platinum *Taq* DNA high-fidelity proofreading polymerase (Invitrogen, USA); DNA of LIVP VACV cl. 14 was used as the DNA template. The resulting target fragments L and R were purified using a QIAquick Gel Extraction Kit (QIAGEN, Netherlands) inserted into the vector HindIII–XmaI fragment pMGC20-gpt (*Fig. 2*) and cloned employing the transformation of competent *E. coli* cells (XL2-Blue strain) and ampicillin as a selective marker. Sequence correctness of the recombinant plasmid pΔati was confirmed by sequencing.

At the next stage of the study, the monolayer culture of CV-1 cells was infected with VACV LIVP and transfected with the recombinant plasmid pΔati, with

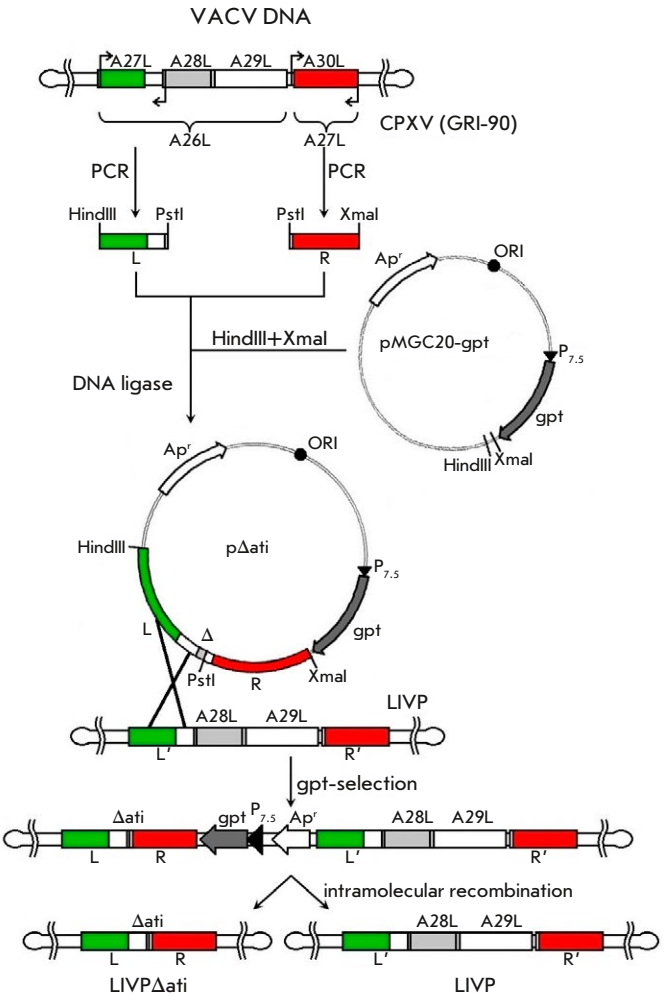


Fig. 2. Scheme for constructing the plasmid pΔati and VACV LIVPΔati (see explanation in the text). L and R – left and right flanking *ati* gene regions

simultaneous gpt selection of VACV recombinant according to the procedure described earlier [17]. Single crossover of the integrative plasmid and viral DNA gave rise to a recombinant viral genome carrying both the selective *gpt* gene and sequences represented by a viral genome segment carrying the target deletion and the same segment without the deletion (*Fig. 2*). This genetic construct carrying the long forward repeats R, R' and L, L' is unstable and can exist only under selective pressure [15, 16]. Withdrawal of selective pressure on the *gpt* gene and intramolecular recombination at R-R' led to the formation of the recombinant virus LIVPΔati (*Fig. 2*). Clones of this viral variant were identified by PCR followed by sequencing of viral DNA.

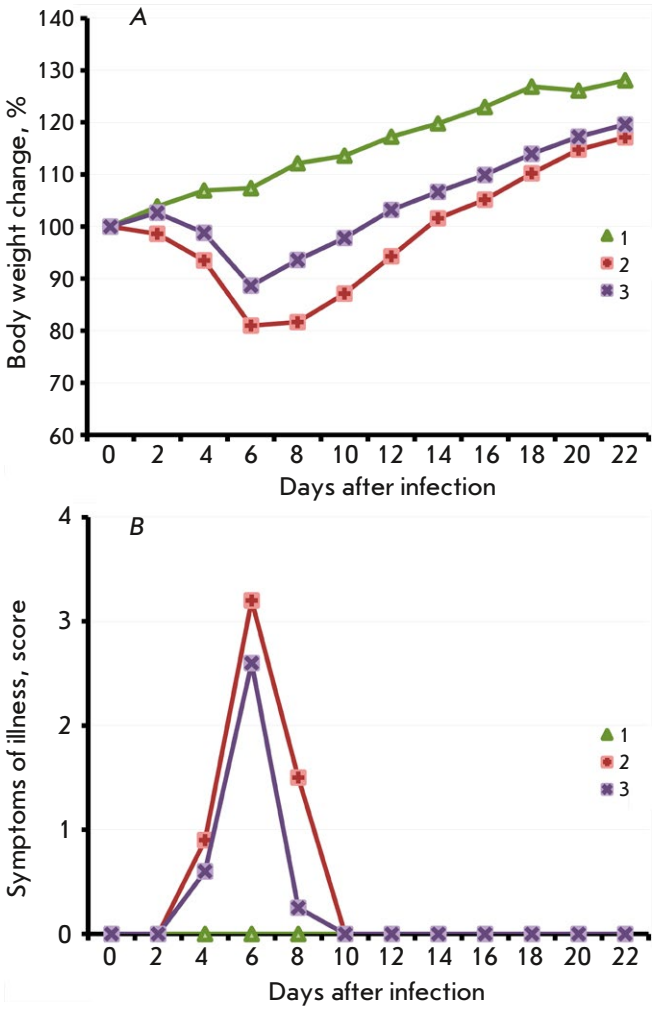


Fig. 3. The dynamics of changes in mouse body weight (*A*) and clinical manifestations of the infection (*B*) after i.n. inoculation of the LIVP (*2*) or LIVPΔati (*3*) viruses at a dose of 10^7 pfu. Data for groups consisting of 10 animals infected with the respective viruses, as well as the control group (*1*), are presented

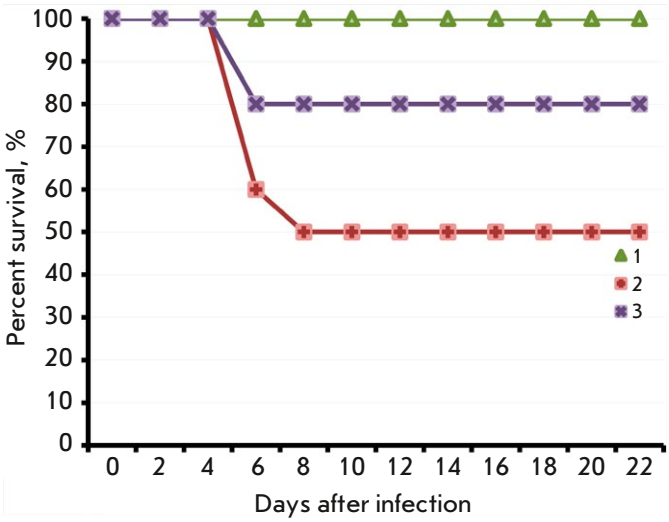


Fig. 4. The dynamics of death of mice i.n. infected with the LIVP (*2*) or LIVPΔati (*3*) viruses at a dose of 10^7 pfu. Control group – non-infected animals (*1*)

Assessment of the pathogenicity of the LIVP and LIVPΔati viruses upon intranasal infection of mice

The pathogenicities of the VACV LIVP and LIVPΔati strains were studied using 3-week-old BALB/c mice. The mice (10 animals per group) were i.n. infected with the viruses at a dose of 10^7 pfu. The animals were followed up during 22 days; each mouse was weighed; clinical manifestations of infection and deaths were documented. In mice infected with VACV LIVP, profound clinical manifestations of the infection were visible starting on day 4; the maximum was attained on day 6; and the animals had recovered after day 10 (*Fig. 3B*). Illness was accompanied by significant body weight reduction (*Fig. 3A*). Under the same conditions, the LIVPΔati virus caused less profound clinical manifestations of the infection (*Fig. 3B*) and less significant body weight loss compared to mice infected with LIVP (*Fig. 3A*), although the differences were statistically insignificant. Infection of mice with the LIVP strain led to the death of 50% of the animals, while only 20% of the mice infected with the LIVPΔati strain died (*Fig. 4*).

These findings are indicative of a reduced pathogenicity of the VACV LIVP with *ati* gene deletion (*Fig. 2*).

Analysis of the development of a cell-mediated immune response to vaccination of mice with VACV variants

The intensity of the cell-mediated immune response in mice i.d. immunized with LIVP or LIVPΔati was determined on 7, 14, and 21 dpi by IFN-γ ELISpot, according to the number of splenocytes producing IFN-γ in response to stimulation with virus-specific peptides. Six animals per group were analyzed at each time point. The results shown in *Fig. 5* demonstrate that a high level of cell-mediated immune response was observed already on 7 dpi, peaking on 14 dpi and significantly dropping by 21 dpi. The dynamics and level of cell-mediated immune response coincided for both VACV strains.

Comparison of the dynamics of emergence of a humoral immune response to the vaccination of mice with the LIVP and LIVPΔati viruses

The levels of VACV-specific IgM and IgG in the sera of mice i.d. immunized with LIVP or LIVPΔati were determined by ELISA on 7, 14, 21, 28, 42, and 56 dpi.

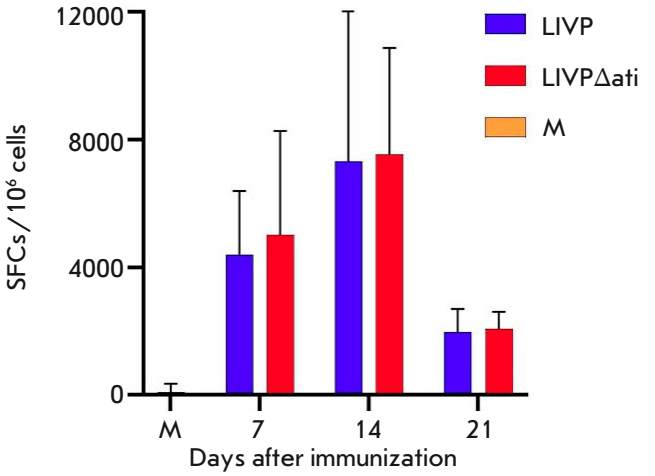


Fig. 5. The results of ELISpot assay of the VACV-specific cell-mediated response in immunized BALB/c mice. SFCs – interferon-γ-producing cells; M – non-immunized mice

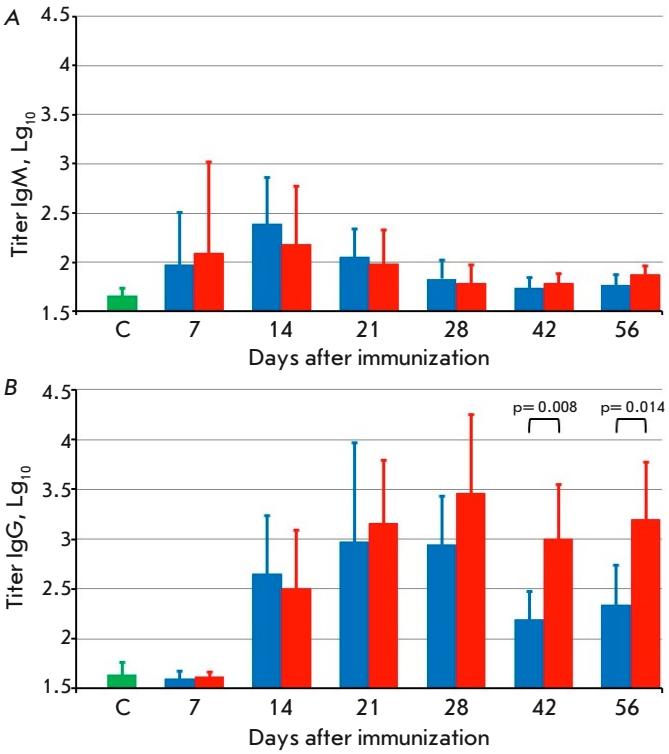


Fig. 6. The titers of VACV-specific IgM (*A*) and IgG (*B*) in the sera of mice immunized with the LIVP (blue bars) or LIVPΔati (red bars) viruses. C – sera of mice that received saline

A relatively high IgM level was observed on 7 dpi; it reached its maximum on 14 dpi and then dropped. Both VACV strains did not differ in terms of the dynamics and level of production of virion-specific IgM (*Fig. 6A*).

An intense synthesis of VACV-specific IgG was observed on 14 dpi; the antibody level subsequently increased on 21–28 dpi and remained high throughout the entire follow-up period (up to 56 dpi, *Fig. 6B*) in mice immunized with LIVPΔati but declined in mice immunized with LIVP. In terms of the geometric mean of inverse IgG titers on 28, 42, and 56 dpi, LIVPΔati was noticeably superior to LIVP; these differences were statistically significant on 42 and 56 dpi (*Fig. 6B*).

Assessment of the protective effect of immunization of mice with the VACV variants against challenge with lethal doses of heterologous orthopoxviruses

On 60 dpi, mice i.d. immunized with VACV LIVP or LIVPΔati, as well as the control (non-immunized) mice, were i.n. infected with either CPXV (at a dose of 2.0×10^6 pfu/mouse) or ECTV (at a dose of 2.2×10^3 pfu/mouse) (six animals per group). During the 14-day follow-up period, we monitored clinical manifestations and deaths of mice. The criterion of VACV infection development according to body weight loss was used.

All the mice in the control group infected with CPXV died after 6 days, while the animals immunized with both VACV variants survived. Weight loss and clinical manifestations were less acute in mice vaccinated with LIVPΔati compared to those vaccinated with LIVP (*Fig. 7*). Differences in body weight loss between the groups of vaccinated mice were statistically significant on days 10–14 after infection with CPXV (*Fig. 7A*).

Intergroup differences were more significant after the infection of immunized mice with the ECTV virus. All control mice died 12 days after infection with ECTV; 83% of the animals vaccinated with LIVP survived, while all the animals survived in the LIVPΔati group (*Fig. 8*). In the LIVPΔati group, manifestations of infection were very mild and were almost never accompanied by body weight loss (*Fig. 9*). Meanwhile, in mice vaccinated with LIVP, clinical signs of infection were observed on days 6–12 after infection with ECTV. The body weight of the animals had substantially dropped; statistically significant differences in this parameter were observed on days 8–14 after infection with ECTV compared to the group of mice vaccinated with LIVPΔati (*Fig. 9A*).

DISCUSSION

The truncated form of the ATI protein (the 94-kDa protein that does not form intracellular ATIs but is a major immunogen upon infection/immunization of

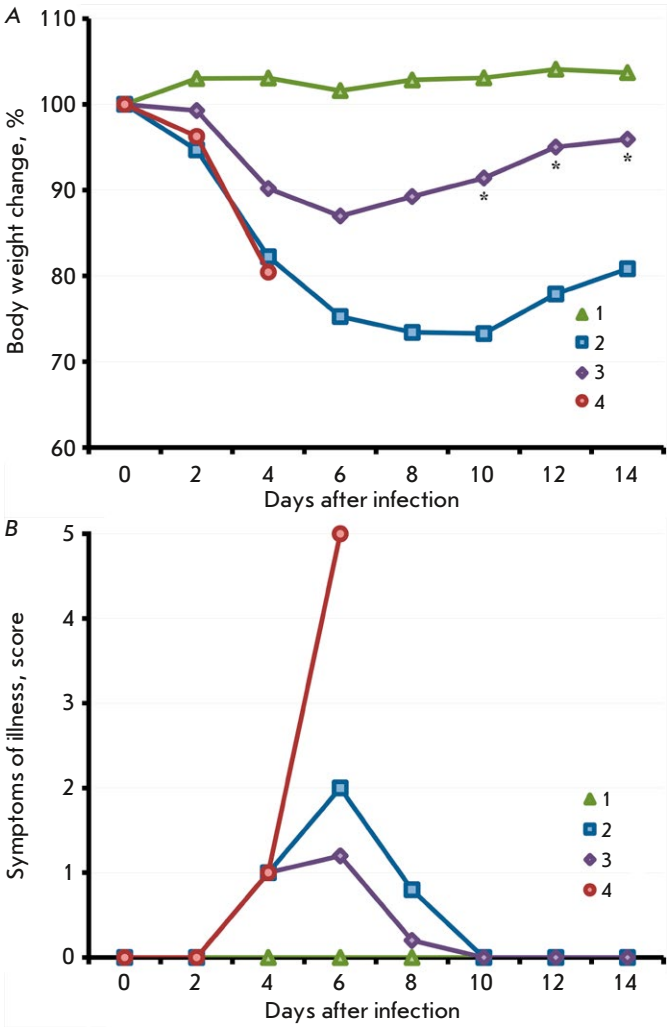


Fig. 7. The dynamics of changes in body weight (*A*) and clinical manifestations of the infection (*B*) after i.n. infection of mice vaccinated with LIVP (*2*) or LIVPΔati (*3*) with CPXV-GRI on day 60 post-immunization. The data are presented for groups consisting of six animals. The controls were groups of non-vaccinated mice, both non-infected (*1*) and infected with CPXV-GRI (*4*). An asterisk shows the time points at which the mean body weights (expressed as a percentage of the initial weight) in the group of mice immunized with LIVPΔati differ statistically significantly from those in the group of mice immunized with LIVP. Comparison was performed using the Student's t-test for independent samples

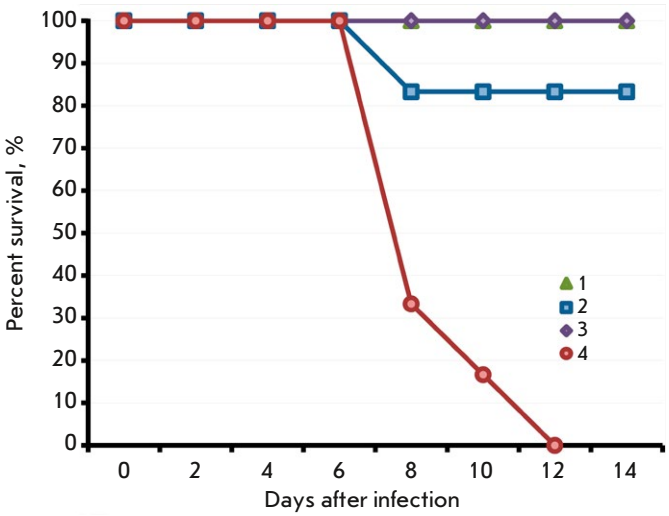


Fig. 8. The dynamics of deaths of mice vaccinated with LIVP (*2*) or LIVPΔati (*3*) after i.n. infection with ECTV-K1 on day 60 post immunization. The data are presented for groups consisting of six animals. The controls were groups of non-vaccinated mice, both non-infected (*1*) and infected with CPXV-ECTV-K1 (*4*)

humans or animals with VACV) is synthesized most abundantly among nonvirion VACV proteins [11]. Meanwhile, antibodies specific to this protein do not exhibit any virus-neutralizing activity. The effect of the production of this nonstructural major immunogenic protein on the manifestation of pathogenicity and immunogenicity of VACV still needs research. Therefore, this work aimed to produce the VACV carrying a targeted deletion of the *ati* gene and investigate the properties of this virus in mice.

The previously characterized clonal variant of LIVP VACV [17] was used as the study object. In LIVP, the *ati* gene (*A26L* in CPXV-GRI) falls into three short potential open reading frames (ORFs); of those, the *A29L* gene encodes the major immunogenic protein A29 (94 kDa) (*Fig. 1*).

The integrative/deletion plasmid pΔati, and subsequently the LIVPΔati strain, were constructed according to the scheme shown in *Fig. 2*.

The pathogenicities of the LIVP and LIVPΔati strains were compared at the first stage. The sensitivity of mice to orthopoxviruses depends largely on their age [2]; therefore, young (3-week-old) BALB/c mice were used. The animals (10 mice per group) were i.n. infected with the viruses, since this method mimics the natural route of infection transmission and is responsible for the fact that mice are most sensitive to this very infection [22, 23].

It turned out that upon i.n. infection of young mice, the LIVP strain at a dose of 10^7 pfu induced the development of a clinically apparent infection (*Fig. 3*) and death of 50% of the animals (*Fig. 4*), while the LIVPΔati strain led to the appearance of less apparent illness signs in mice (*Fig. 3*) and caused the death of 20% of the animals (*Fig. 4*). Hence, the *ati* gene deletion in VACV LIVP caused its attenuation compared to the original viral strain. That is consistent with the

previously held assumption that the reduced pathogenicity of some natural VACV strains can be caused by a spontaneous deletion of the *ati* gene in them [13], although experimental evidence to support this has not been provided.

The immunogenicity of VACV LIVP and LIVPΔati was studied in adult mice (aged 6–7 weeks) with a mature immune system. The development of a VACV-specific cell-mediated immune response to i.d. vaccination of mice was assessed by ELISpot, according to the number of splenocytes producing IFN-γ in response to stimulation with peptides. A strong cell-mediated immune response was observed already on 7 dpi, reaching a maximum on 14 dpi and significantly declining by 21 dpi (*Fig. 5*). Meanwhile, deletion of the *ati* gene in VACV LIVP affected neither the dynamics nor the strenght of the cell-mediated immune response to vaccination in mice.

The antibody response is known to make the greatest contribution to the eliciting of the adaptive immune response to vaccination with VACV [3, 24]. Therefore, the dynamics of synthesis of IgM and IgG specific to the virion proteins of the VACV after i.d. vaccination of mice with the LIVP or LIVPΔati strain at doses of 10^5 pfu were assessed by ELISA.

A relatively high IgM level was detected on 7 dpi, attaining its maximum on 14 dpi and then declining. Both VACV strains did not differ in terms of the dynamics and the level of production of virion-specific IgM (*Fig. 6A*).

Intense synthesis of VACV-specific IgG was observed starting on 14 dpi; the antibody level subsequently increased on 21–28 dpi. LIVPΔati was noticeably superior to LIVP in terms of the geometric means of inverse IgG titers on 28, 42, and 56 dpi; this superiority was statistically significant on 42 and 56 dpi (*Fig. 6B*). One of the plausible reasons for this is that the absence of synthesis of the major, late nonstructural protein A29 in LIVPΔati does not distract the immune system from synthesizing IgG specific to this protein and ensures a more intense synthesis of antibodies specific to the virion proteins of VACV.

A number of studies have demonstrated that the humoral immune response makes the greatest contribution to the protection against a challenge with orthopoxviruses [3, 6, 24, 25]; therefore, it was important to assess the protective immunity that had developed in response to i.d. vaccination of mice with VACV LIVP and LIVPΔati. For this purpose, six mice in each group were i.n. infected with lethal doses of CPXV GRI-90 or ECTV K-1 on 60 dpi. In both cases, the protective effect of vaccination with LIVPΔati was stronger than that for vaccination with the parent LIVP strain (*Figs. 7–9*), which supports the earlier

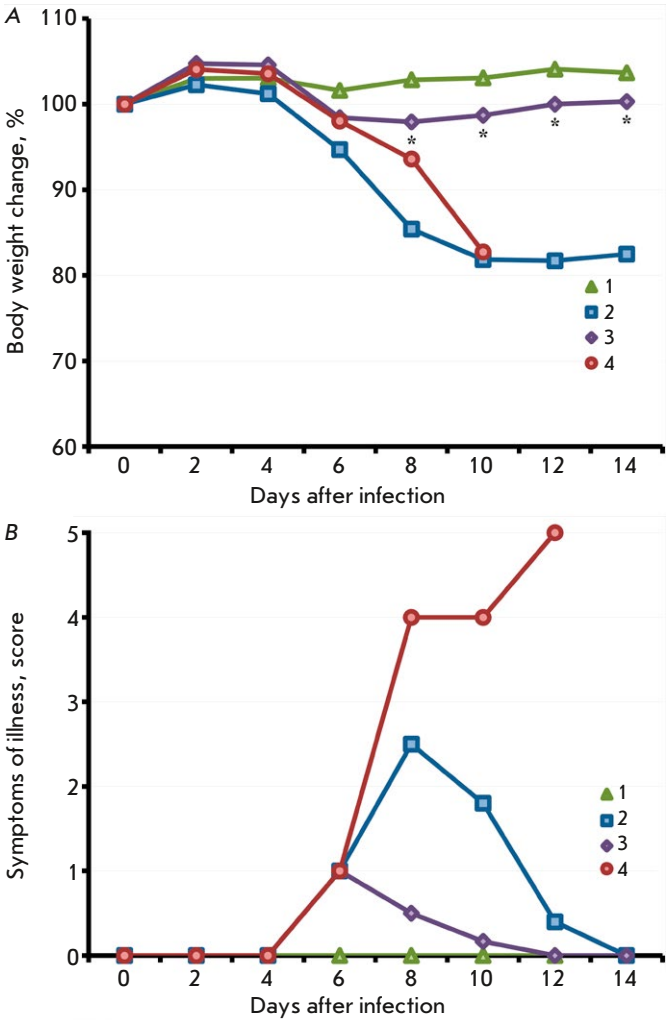


Fig. 9. The dynamics of changes in body weight (*A*) and clinical manifestations of infection (*B*) after i.n. infection of mice vaccinated with LIVP (*2*) or LIVPΔati (*3*) with ECTV-K1 on day 60 post immunization. The data are presented for groups consisting of six animals. The controls were groups of non-vaccinated mice, both non-infected (*1*) and infected with ECTV-K1 (*4*). An asterisk shows time points at which the mean body weight (expressed as a percentage of the initial weight) in the group of mice immunized with LIVPΔati differs statistically significantly from that in the group of mice immunized with LIVP. Comparison was performed using the Student's t-test for independent samples

conclusions about the crucial role of the antibody response in the development of a body's defense against a orthopoxvirus infection.

A conclusion can be drawn that deletion of the genomic region of VACV LIVP comprising the *ati* gene weakens the pathogenic properties of the LIVPΔati virus upon i.n. infection of BALB/c mice and increases the production of virion-specific IgG in response to i.d. vaccination of mice with this mutant virus, thus

ensuring stronger protection for mice (compared to the parent LIVP) against the subsequently induced lethal infection with the heterologous orthopoxviruses CPXV and ECTV. Therefore, the LIVPΔati strain can be considered a promising vector for constructing polyvalent recombinant vaccines against various infectious diseases.

This work was supported by the Russian Science Foundation (grant No. 19-14-00006-P).

The authors would like to thank I.V. Kolosova for providing the VACV LIVP preparations as well as D.N. Kisakov and S.A. Pyankov for statistical analysis of the experimental data.

REFERENCES

1. Fenner F., Henderson D.A., Arita I., Jezek Z., Ladnyi I.D. Smallpox and Its Eradication. World Health Organization. Geneva, Switzerland, 1988. 1460 p.
2. Shchelkunov S.N., Marennikova S.S., Moyer R.W. Orthopoxviruses Pathogenic for Humans. New York, USA: Springer, 2005. 425 p.
3. Moss B. // Immunol. Rev. 2011. V. 239. № 1. P. 8–26.
4. Shchelkunova G.A., Shchelkunov S.N. // Acta Naturae. 2017. V. 9. № 4. P. 4–12.
5. Payne L.G. // J. Gen. Virol. 1980. V. 50. P. 89–100.
6. Shchelkunov S.N., Shchelkunova G.A. // Acta Naturae. 2020. V. 12. № 1. P. 33–41.
7. Shedlovsky T., Smadel J.E. // J. Exp. Med. 1942. V. 75. № 2. P. 165–178.
8. Funahashi S., Sato T., Shida H. // J. Gen. Virol. 1988. V. 69. № 1. P. 35–47.
9. de Carlos A., Paez E. // Virology. 1991. V. 185. № 2. P. 768–778.
10. Katsafanas G.C., Moss B. // J. Virol. 2020. V. 94. № 4. P. e01671–19.
11. Jones-Trower A., Garcia A., Meseda C.A., He Y., Weiss C., Kumar A., Weir J.P., Merchlinsky M. // Virology. 2005. V. 343. № 1. P. 128–140.
12. Moutaftsi M., Tscharke D.C., Vaughan K., Koelle D.M., Stern L., Calvo-Calle M., Ennis F., Terajima M., Sutter G., Crotty S., et al. // Future Microbiol. 2010. V. 5. № 2. P. 221–239.
13. Marques J.T., Trindade G.D., Da Fonseca F.G., Dos Santos J.R., Bonjardim C.A., Ferreira P.C., Kroon E.G. // Virus Genes. 2001. V. 23. № 3. P. 291–301.
14. Leite J.A., Drumond B.P., de Souza Trindade G., Bonjardim C.A., Ferreira P.C.P., Kroon E.G. // Virus Genes. 2007. V. 35. № 3. P. 531–539.
15. Leite J.A., da Fonseca F.G., de Souza Trindade G., Abrahao J.S., Arantes R.M., de Almeida-Leite C.M., dos Santos J.R., Guedes M.I., Ribeiro B.M., Bonjardim C.A., et al. // Arch. Virol. 2011. V. 156. № 4. P. 617–628.
16. Kastenmayer R.J., Maruri-Avidal L., Americo J.L., Earl P.L., Weisberg A.S., Moss B. // Virology. 2014. V. 452–453. P. 59–66.
17. Yakubitskiy S.N., Kolosova I.V., Maksyutov R.A., Shchelkunov S.N. // Acta Naturae. 2015. V. 7. № 4. P. 113–121.
18. Shchelkunov S.N., Yakubitskiy S.N., Bauer T.V., Sergeev A.A., Kabanov A.S., Bulichev L.E., Yurganova I.A., Odnoshevskiy D.A., Kolosova I.V., Pyankov S.A., et al. // Acta Naturae. 2020. V. 12. № 4. P. 120–132.
19. Shchelkunov S.N., Yakubitskiy S.N., Sergeev A.A., Starostina E.V., Titova K.A., Pyankov S.A., Shchelkunova G.A., Borgoyakova M.B., Zadorozhny A.M., Orlova L.A., et al. // Viruses. 2022. V. 14. № 7. P. 1453.
20. Oseroff C., Peters B., Pasquetto V., Moutaftsi M., Sidney J., Panchanathan V., Tscharke D.C., Maillere B., Grey H., Sette A. // J. Immunol. 2008. V. 180. № 11. P. 7193–7202.
21. Russell T.A., Tscharke D.C. // Immunol. Cell Biol. 2014. V. 92. № 5. P. 466–469.
22. Hughes L.J., Townsend M.B., Gallardo-Romero N., Hutson C.L., Patel N., Dotty J.B., Salzer J.S., Damon I.K., Carroll D.S., Satheshkumar P.S., et al. // Virology. 2020. V. 544. P. 55–63.
23. Shchelkunov S.N., Yakubitskiy S.N., Sergeev A.A., Kabanov A.S., Bauer T.V., Bulichev L.E., Pyankov S.A. // Viruses. 2020. V. 12. № 8. P. 795.
24. Belyakov I.M., Earl P., Dzutsev A., Kuznetsov V.A., Lemon M., Wyatt L.S., Snyder J.T., Ahlers J.D., Franchini G., Moss B., et al. // Proc. Natl. Acad. Sci. USA. 2003. V. 100. P. 9458–9463.
25. Lambert P.H., Laurent P.E. // Vaccine. 2008. V. 26. P. 3197–3208.

Attenuating Neuronal Autophagy Alleviates Inflammatory Injury in OGD-Deprived Co-culture of HT22 with BV2

Z. W. Huang[1], Y. Y. Liu[1], X. M. Chen[1], C. L. Yu[2*], H. Y. He[2**], Y. H. Deng[1***]
[1]Department of basic medicine, Medical School, Kunming University of Science and Technology, Kunming, 650500 China
[2]Anning First People's Hospital Affiliated to Kunming University of Science and Technology, Kunming, 650300 China
E-mail: * – yuchunlei1828820@163.com; ** – 511869321@qq.com; *** – 827821533@qq.com
Received October 10, 2022; in final form, July 10, 2023
DOI: 10.32607/actanaturae.11830
Copyright © 2023 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License,which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **Neuronal CX3CL1 suppressed microglial inflammation by binding to its receptor CX3CR1 expressed on microglia. Neuronal autophagy was prominently activated by cerebral ischemia, whereas CX3CL1 expression in autophagic neurons was conversely down-regulated to exacerbate microglial inflammation. Accordingly, this study was meant to investigate whether ischemia-activated microglial inflammation could be repressed by promoting CX3CL1 expression via the attenuation of neuronal autophagy. Immunofluorescence showed that autophagy predominantly occurred in neurons but barely in microglia. Western blot and immunofluorescence demonstrated that attenuating HT22 autophagy significantly increased its CX3CL1 expression and subsequently mitigated the BV2-mediated inflammatory responses, as indicated by decreased inflammatory factors of NF-κB-p65, IL-6, IL-1β, TNF-α, and PGE2. Meanwhile, CCK-8, Nissl staining, and FJC staining showed that an OGD (Oxygen-glycogen deprivation)-created neuronal injury was greatly alleviated by CX3CL1-suppressed microglial inflammation. Contrarily, elevating HT22 autophagy markedly decreased its CX3CL1 expression, which consequently worsened microglial inflammation and the neuronal injury. Our data suggests that attenuating neuronal autophagy may be an effective method to alleviate a microglial inflammatory injury after an ischemic stroke.**
KEYWORDS **Ischemic stroke, neuronal autophagy, CX3CL1 expression, microglial inflammation, neuroprotection.**

INTRODUCTION

Cerebral stroke, a serious cerebrovascular disease, remains the main cause of disability and the second leading cause of death worldwide. Approximately 87% of patients suffer from an ischemic stroke [1]. The pathogenesis of cerebral ischemia has been investigated for decades, but means to alleviate post-stroke neurological injury remain troublingly few. A cascade of pathological processes causes neuronal death after an ischemic stroke, such as nutrient and energy depletion, release of reactive oxygen species, intracellular calcium overload, neuro-excitotoxicity, etc. [2]. Cerebral ischemia simultaneously activates a microglial inflammation and autophagic signaling. Microglia maintain cellular homeostasis by monitoring the microenvironment for responding to an injurious stimulus, such as ischemia. However, microglia-medicated inflammatory responses have been confirmed to be excessively amplified and, thereby, to accelerate the pathological aftershocks of an ischemic stroke [3]. A growing body of evidence demonstrates that autophagy remains prominently activated at the acute phase of a stroke. Yet this activated autophagy is predominantly displayed in neurons but seldom in microglia at the penumbra [4]. Recent studies have shown that there are close interactions between neuronal autophagy and a microglial inflammation [3]. Thus, understanding the mutual regulations between them might offer more clues for stroke treatment.

The chemokine fractalkine/CX3CL1 is a unique member of the CX3C family of chemokines. It is crucial in mediating the inflammatory response in the central nervous system [5]. Studies have revealed that communication between neurons and microglia is established via CX3CL1–CX3CR1 signaling [6]. CX3XL1 is only expressed on the membranes of neurons, while its receptor CX3CR1 is for the most part located on microglia [7]. Microglia are native inflammatory cells in the brain and are kept quiescent through conjugation with neurons by the CX3CL1–CX3CR1 contact under physiological conditions [8]. Thus, microglial activity is kept at an appropriate level to avoid trig-

gering an excessive inflammatory response that can lead to neurological injury [9]. Thus, the CX3CL1–CX3CR1-mediated interaction between neurons and microglia was critical in maintaining normal brain function [10]. However, the inhibitory effects of neurons on the microglial inflammation are likely disrupted if the CX3CL1 and/or CX3CLR expressions are altered under a pathological state, such as cerebral ischemia [11]. Therefore, this study is meant to investigate what and how the CX3CL1-repressed microglial inflammation is disturbed, using an ischemia model of a co-culture of HT22 neurons with BV2 microglia.

Autophagy is a metabolic process by which damaged organelles, old proteins, superfluous cytoplasmic ingredients, and waste substrates to lysosomes are delivered for degradation [12]. At the same time, excessive autophagy accelerates cell death due to the uncontrolled autophagy initiation [13]. Autophagic/lysosomal signaling is prominently activated by cerebral ischemia [14]. Meanwhile, both reported studies and our previous investigations demonstrated that autophagy in neurons is excessively elevated by ischemic ischemia, leading to a massive accumulation of autophagic cargo within cells. Ultimately, the neurons at the penumbra suffer from autophagic cell death [15]. Intriguingly, more evidence shows that autophagy predominantly occurs in neurons, but seldom in microglia after an acute ischemic stroke [1]. Based on the CX3CL1–CX3CR1-mediated crosstalk mechanism, we asked ourselves whether the CX3CL1 expression could be changed in autophagic neurons, thereby subsequently weakening the suppressive effect of neurons on a microglial inflammation [5]. Consequently, the microglia-triggered inflammation response can be amplified to increase neuronal death [16]. To verify this hypothesis, a rat model of ischemic stroke was prepared to attempt to better understand the correlation between neuronal autophagy and microglial inflammation in our previous study [16]. The results showed that ischemia-induced neuronal autophagy leads to a reduction of the CX3CL1 expression. Moreover, further autophagy decreases the CX3CL1 expression and, consequently, aggravates the microglial inflammation and neurological injury. Conversely, attenuating autophagy significantly elevates the CX3CL1 expression of neurons, which in turn alleviates the microglial inflammatory injury and brain damage. These data support the contention that neuronal autophagy aggravates the microglial inflammatory injury by down-regulating the CX3CL1 expression on neurons. However, the study failed to elucidate the direct regulative mechanism of neuronal autophagy on a microglial inflammation after cerebral ischemia.

To investigate the direct crosstalk mechanism between neuronal autophagy and a microglial inflammation, an OGD co-culture of HT22 neurons with BV2 microglia was first prepared in this study. Thereafter, the culture condition was made to meet the requirement that autophagy is mostly induced in HT22, but rarely in BV2. Based on this understanding, the autophagy level in the neurons is pharmacologically altered to reveal the exact effect of neuronal autophagy on the microglial inflammatory response. The effect of a autophagy-regulated microglial inflammation on a neuronal injury is correspondingly explored. Through our study, the correlative regulation between neuronal autophagy and a microglial inflammation after cerebral ischemia ought to be fully elucidated.

1. MATERIAL AND METHODS

1.1. Cell culture

Mouse hippocampal neuron (HT22) and mouse microglial cells (BV2) were purchased from Wuhan Procell Life Technology Co., Ltd (Wuhan, China). HT22 and BV2 cells were firstly cultured in a high glucose DMEM medium (Hyclone, UT, USA) containing 10% fetal bovine serum (Biological Industries, CT, USA), respectively. After 2 days of separated culture, the HT22 and BV2 cells were collected and counted, and they were seeded into T25 culture flasks at a ratio of 9 : 1 for co-culture. After 24 h of co-culture, the model of oxygen-glucose deprivation/reoxygenation (OGD/R) was ready.

1.2. Oxygen-glycogen deprivation/reoxygenation (OGD/R)

To prepare the model of cell ischemia *in vitro*, the complete co-culture medium of HT22 with BV2 was replaced with a serum-free, sugar-free medium (glucose deprivation). The culture plates were placed in 95% N_2 and 5% CO_2 chambers (oxygen deprivation). After 1.5 h of OGD, the culture medium was replaced with a complete DMEM medium (resupply of glucose) and the plates were moved into the incubator with 5% CO_2 (reintroduction of oxygen). In this way, a OGD cell model mimicking the microenvironment in the ischemic brain tissues was created. The main objective of our study was to establish the correlation between neuronal autophagy and a microglial inflammation. For this reason, the culture conditions had to be adjusted to account for the fact that autophagy is mostly induced in neurons but little in microglia. Following this, the autophagy level in the HT22 and BV2 cells, respectively, was measured by double immunofluorescence.

We investigated how neuronal autophagy affects its CX3CL1 expression, which subsequently regulates the microglial inflammatory response by altering CX3CL1–CX3CR1 signaling. The autophagy inducer

Tat-Beclin1 and inhibitor 3-methyladenine (3-MA, 15 μM) were additionally added into the co-culture medium. Our preliminary study had confirmed that a dose of 15 μM Tat-Beclin1 could further promote autophagy in HT22 cells upon OGD, whereas the same dose of Tat-Beclin1 had little effect on the autophagy level in BV2 in the co-culture.

1.3. Western blot

The total proteins of the co-cultured cells were extracted using a protein extraction kit (Beyotime Biotechnology, Shanghai, China). After quantification by the BCA method, the proteins were separated by molecular weight using a polyacrylamide gel before being transferred onto PVDF membranes (Millipore Corporation, Ma, USA). Nonspecific proteins were blocked with 10% nonfat milk for 2 h at room temperature. After washing with TBST, the PVDF membranes were incubated with rabbit primary antibodies against mouse LC3 (1 : 10000, Sigma, MO, USA), beclin1 (1 : 1000, ABclonal, Wuhan, China), CX3CL1 (1 : 2000, GeneTex, CA, USA), NF-κB-p65 (1 : 1000, GeneTex, CA, USA), and β-actin (1 : 10000, Abclonal, Wuhan, China) overnight at 4°C. After the washing step, the secondary antibodies were labeled for 2 h at room temperature. The fluorescence signal intensity was analyzed by Image J, and the band density values were normalized to β-actin.

1.4. Immunofluorescence

The co-cultured cells were seeded onto six-well plates coated with polylysine. Thereafter, the cells were permeabilized with 0.2% Triton-X100 for 5 min and washed with PBS. After blocking with 10% BSA (Beyotime Biotechnology, Shanghai, China) for 1 h at room temperature, the rabbit primary antibodies against rat LC3 (1 : 400, Sigma, MO, USA), NeuN (1 : 400, Abcam, Cambs, UK), Iba-1 (1 : 400, Abcam, Cambs, UK), CX3CL1 (1 : 400, GeneTex, Texas, USA), and NF-κB-p65 (1 : 400, GeneTex, TX, USA) were incubated overnight at 4°C. After washing, they were labeled with Alexa Fluor-coupled secondary antibodies (1 : 1000, Jackson ImmunoResearch Laboratories, INC. PA, USA) in the dark. Finally, the cells were washed and fixed with DAPI (1 : 2000, Sigma, MO, USA). The results were represented in the form of fluorescence intensity. Under high magnification (×200), the fluorescence density was calculated by Image J and the fluorescence density values required background removal.

1.5. Enzyme-linked immunosorbent assay (ELISA)

The HT22 and BV2 co-cultured cells were collected by centrifugation, and the culture medium was also obtained for the measurement of inflammatory factors. The concentrations of TNF-α, IL-6, IL-1β, and PGE2 in the co-culture medium were measured by an ELISA kit (Biotech & Jingmeibio, Beijing, China), according to the instructions provided by the manufacturer. The detected values were quantified according to the standard curve.

1.6. Nissl staining

The co-cultured cells seeded on the six-well plates were firstly fixed with 4% paraformaldehyde and then dehydrated by 70% ethanol for 1 min. Following that, they were immersed in a Cresyl violet Stain solution (Leagene, Beijing, China) at 56°C for 1 h and then rinsed with deionized water. Thereafter, the Nissl Differentiation solution (Leagene, Beijing, China) was added to incubate for 2 min. Finally, rapid dehydration using ethanol, as well as verification using xylene, was conducted. The staining was observed and photographed with a stereoscopic microscope (Nikon Instruments Co., Ltd., Tokyo, Japan). The result was expressed as the number of Nissl bodies in 10 randomly selected non-overlapping fields under high magnification (×200). Five plates had to be counted for each group.

1.7. Fluoro-Jade C (FJC) staining

A Fluoro-Jade C (FJC) Staining Kit (Thermofisher, MA, USA) was used to detect the necrotic neurons, according to the instructions provided by the manufacturer. The co-cultured cells were processed with sodium hydroxide from Solution A for 5 min and incubated with 70% ethanol for 2 min, then washed in distilled water for 2 min. The cells were further incubated with potassium permanganate from the working solution B for 10 min. After washing with distilled water, the cells were stained with Fluoro-Jade C from the working solution C for 10 min. After the washing, the cells were permeabilizated with xylene and examined with a fluorescent microscope (Nikon Instruments Co., Ltd., Tokyo, Japan). The result was expressed in the number of FJC-stained cells in 6 randomly selected non-overlapping fields under high magnification (×200). Five plates had to be counted for each group.

1.8. Cell Counting Kit-8 (CCK-8) kit

The cell viability was measured by a CCK-8 kit (Beyotime Biotechnology, Shanghai, China), according to the instructions provided by the manufacturer. Absorbance at a 450-nm wavelength was detected using a microplate reader. Cell viability was calculated using the formula of the survival rate: [(As – Ab)/(Ac – Ab)] × 100%. Meanwhile, the inhibition

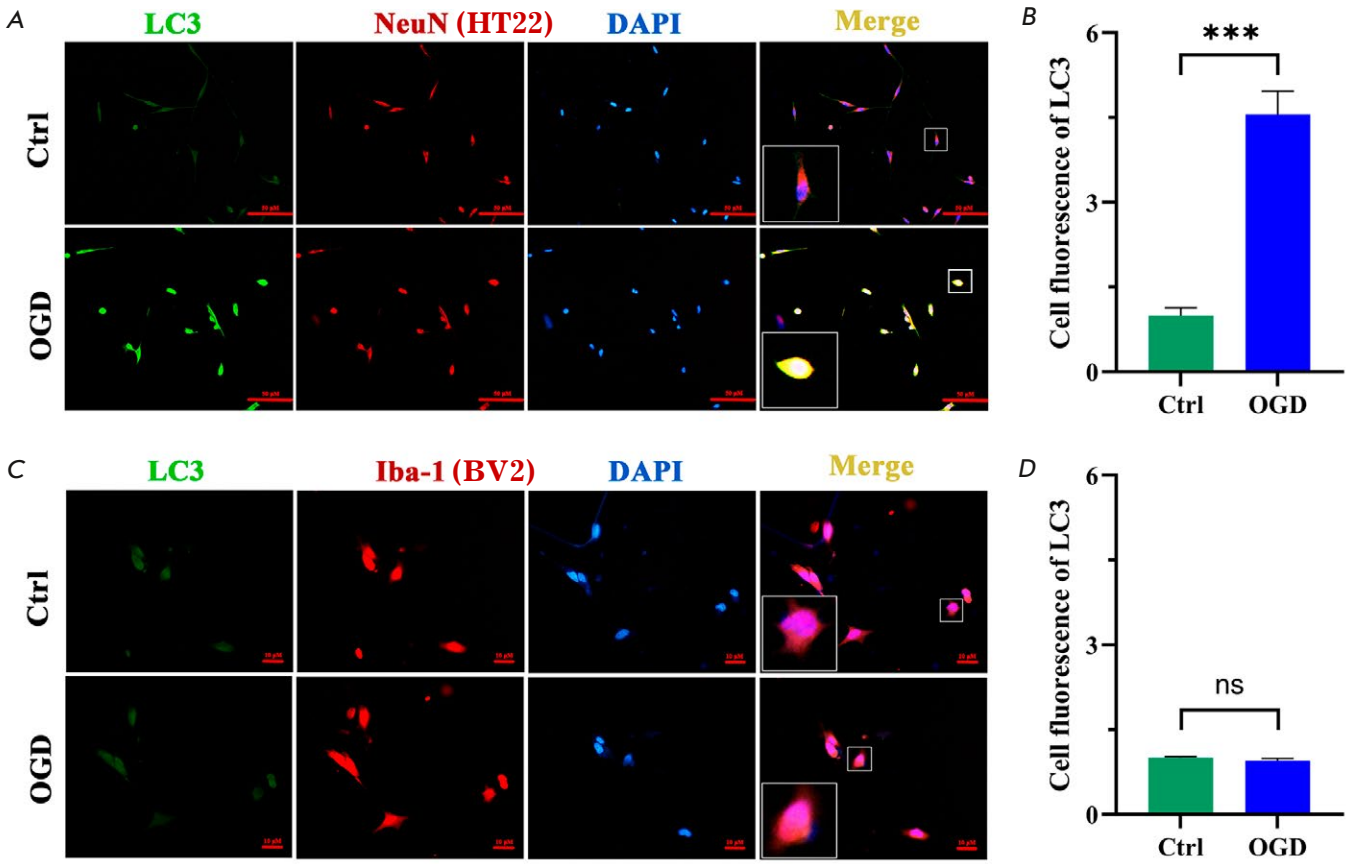


Fig. 1. Immunofluorescence was performed to identify the co-culture condition g that meets the requirement that autophagy is mostly induced in HT22 cells but barely in BV2 upon OGD. By Screening, 1.5 h of OGD followed by 2 h of reoxygenation was recognized to meet the requirement that autophagy mainly occurs in neurons but seldom in microglia. (*A*, *C*) Immunofluorescence images of LC3 (green) / NeuN (red), and Iba-1 (red) / DAPI (blue) co-localization. (*B*, *D*) Cell fluorescence of LC3. Bar: 50 μm, $n = 6$. $^{***}p < 0.001$, ns

rate was calculated using the formula [(Ac – As)/(Ac – Ab)] × 100%. As: Autophagy intervention group, Ac: OGD group, Ab: Sham group.

1.9. Statistical analysis

All the data in this study were subjected to one-way ANOVA or the t-test for statistical differences by SPSS 24.0, and values of $P < 0.05$ were considered statistically significant. The statistical analyses were expressed as a mean ± SEM. Western blot strips were analyzed using Image J, bar statistical histograms were drawn using GraphPad Prism 9, and immunofluorescence graphs were processed using Adobe Photoshop CC 2022.

2. RESULT

2.1. The condition of autophagy predominantly occurs in neurons but barely so in microglia

An OGD co-culture model of HT22 neurons with BV2 microglia was prepared. In order to establish the correlation between neuronal autophagy and a microglial inflammation, a culture condition was created firstly so as to meet the requirement that autophagy is mostly induced in HT22, while scarcely so in BV2 cells. Double immunofluorescence demonstrated that 1.5 h of OGD, followed by 2 h of reoxygenation, was the appropriate culture condition in which autophagy is predominantly induced in HT22 cells but barely so in BV2 (*Fig. 1A–D*).

2.2. OGD-induced neuronal autophagy decreased its CX3CL1 expression

Autophagy was for the most part induced in HT22 neurons 2 h after OGD, as mentioned above. Thus, the variation in CX3CL1 expression could be directly observed in the HT22 neurons that had suffered from autophagy. Western blot showed that the ratio of LC3-II/LC3-I and the Beclin1 expression were prominently high in the co-cultured cells of HT22 with BV2 (*Fig. 2C,D*), whereas the CX3CL1 expression was conversely low (*Fig. 2E*) in the OGD group compared

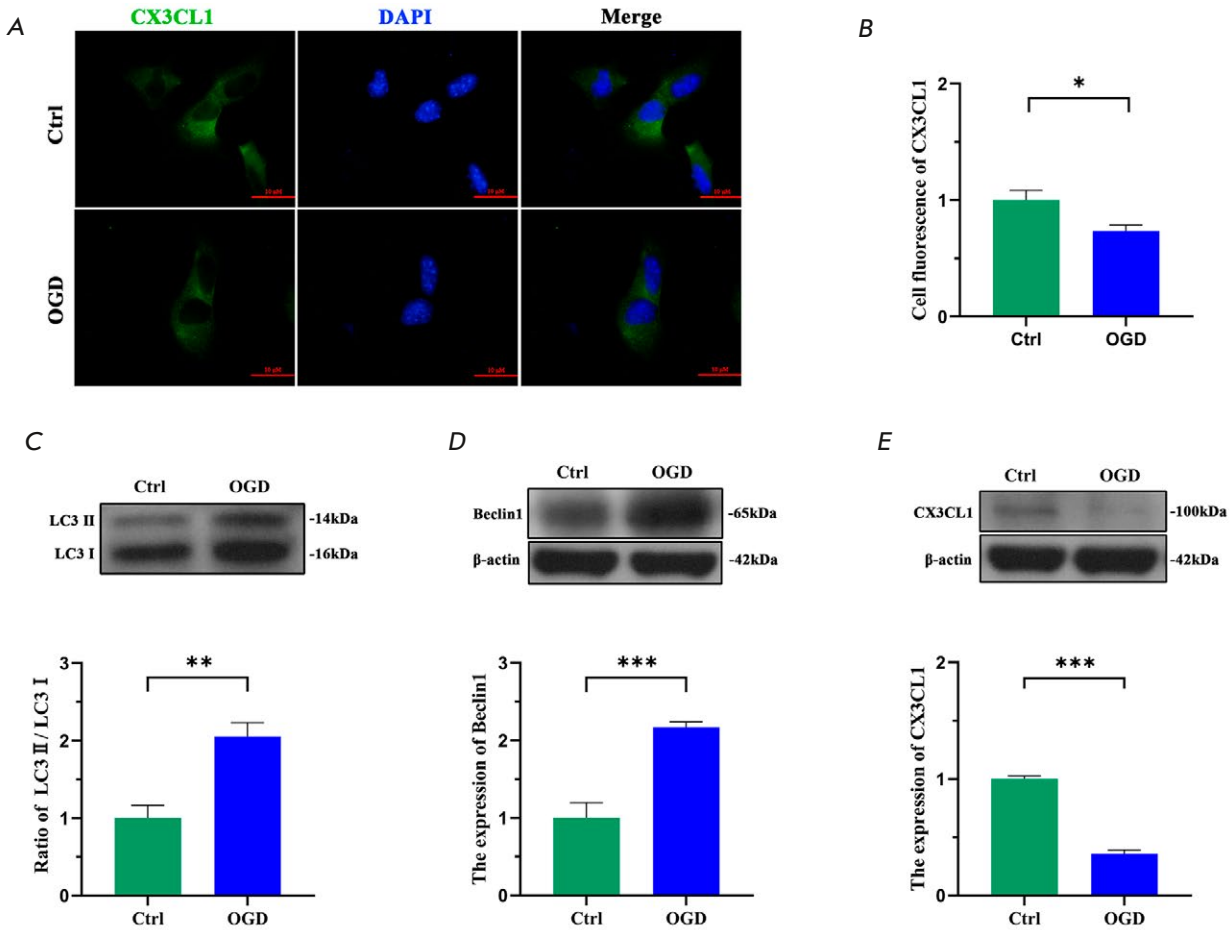


Fig. 2. OGD-induced autophagy in HT22 led to decreased CX3CL1 expression. (*A*) Immunofluorescence images of CX3CL1 (green) / DAPI (blue) localization. (*B*) Cell fluorescence of CX3CL1 (green) from image. (*C–E*) Western blot images of LC3, Beclin1, CX3CL1, and β-actin expression. Quantitative analysis of the immunoblotted proteins by Image J. Bar: 50 μm, $n = 6$. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

with those in the control group. Furthermore, double immunofluorescence demonstrated that OGD significantly elevated the autophagy level in the HT22 cells. However, the CX3CL1 expression was contrarily decreased in the OGD HT22 neurons (*Fig. 2A,B*). These results seem to indicate that neuronal autophagy down regulates CX3CL1 expression.

2.3. Attenuating HT22 autophagy suppresses BV2 inflammatory activation

The OGD-elevated autophagy in HT22 cells resulted in decreased CX3CL1 expression, which likely weakened the efficacy of their suppressive action on the microglial inflammatory response, due to a disruption of the CX3CL1–CX3CR1 cross-talk. We, therefore, inquired whether the BV2-mediated inflammation could be abated by increasing the CX3CL1 expression via the attenuation of autophagy in HT22 cells. To alter the autophagy level, the autophagy inducer Tat-Beclin1 and its inhibitor 3-MA, respectively, were added into the co-culture medium upon OGD. Western blot showed that both the Beclin1 expression and the ratio LC3-II/LC3-I could be effectively altered by the autophagic agents (*Fig. 3A,B*). The promoted autophagy further decreased CX3CL1 expression in the OGD+Tat group, compared with that in the OGD group (*Fig. 3C*). By contrast, decreasing autophagy increased the CX3CL1 expression in our example (*Fig. 3C*). Meanwhile, western blot demonstrated that 15 μM of Tat-Beclin1 greatly boosts autophagy in HT22 neurons (*Fig. 3A,B*). Moreover, Tat-Beclin1-promoted HT22 autophagy markedly reduced its CX3CL1 expression (*Fig. 4C*). Conversely, 3-MA-inhibited autophagy greatly increased its CX3CL1 expression in the OGD+3-MA group (*Fig. 3C*), compared with that in the OGD group. We further investigated the effect of the altered CX3CL1 expression in the HT22 neurons on microglial inflammatory signaling. The result indicates (*Fig. 4A–D*) that the inflammatory signaling of NF-κB-p65 was prominently reinforced

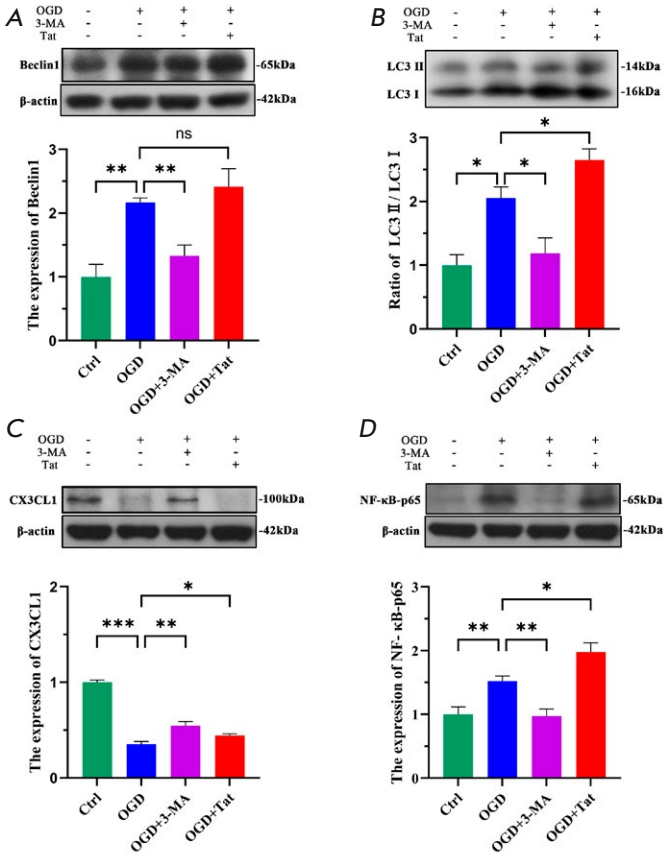


Fig. 3. Inhibiting HT22 autophagy suppressed BV2 inflammatory activation. (*A–D*) Western blot images of LC3, Beclin1, CX3CL1, NF-κB-p65, and β-actin expression. Quantitative analysis of the immunoblotted proteins by Image J. $n = 6$. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, ns

by the autophagy-decreased CX3CL1 (*Fig. 4C,D*). Contrarily, the increased CX3CL1 expression through autophagy inhibition proved effective in mitigating the microglial activation.

2.4. Inhibiting neuronal autophagy repressed the microglial inflammatory response

To explore whether the microglial inflammatory response could be suppressed by inhibiting neuronal autophagy, the presence of the inflammatory factors IL-6, IL-1β, TNF-α, and PGE2 was assessed by measuring their concentrations in the OGD co-culture medium. The results (*Fig. 5A–D*) showed that Tat-Beclin1-elevated autophagy in HT22 cells noticeably aggravated the OGD-induced microglial inflammatory response, as was reflected by increased concentrations of IL-6, IL-1β, TNF-α, and PGE2. By contrast, inhibiting autophagy in HT22 cells proved effective in suppressing the inflammation in the OGD+3-MA group, compared with that in the OGD+Tat group or OGD group.

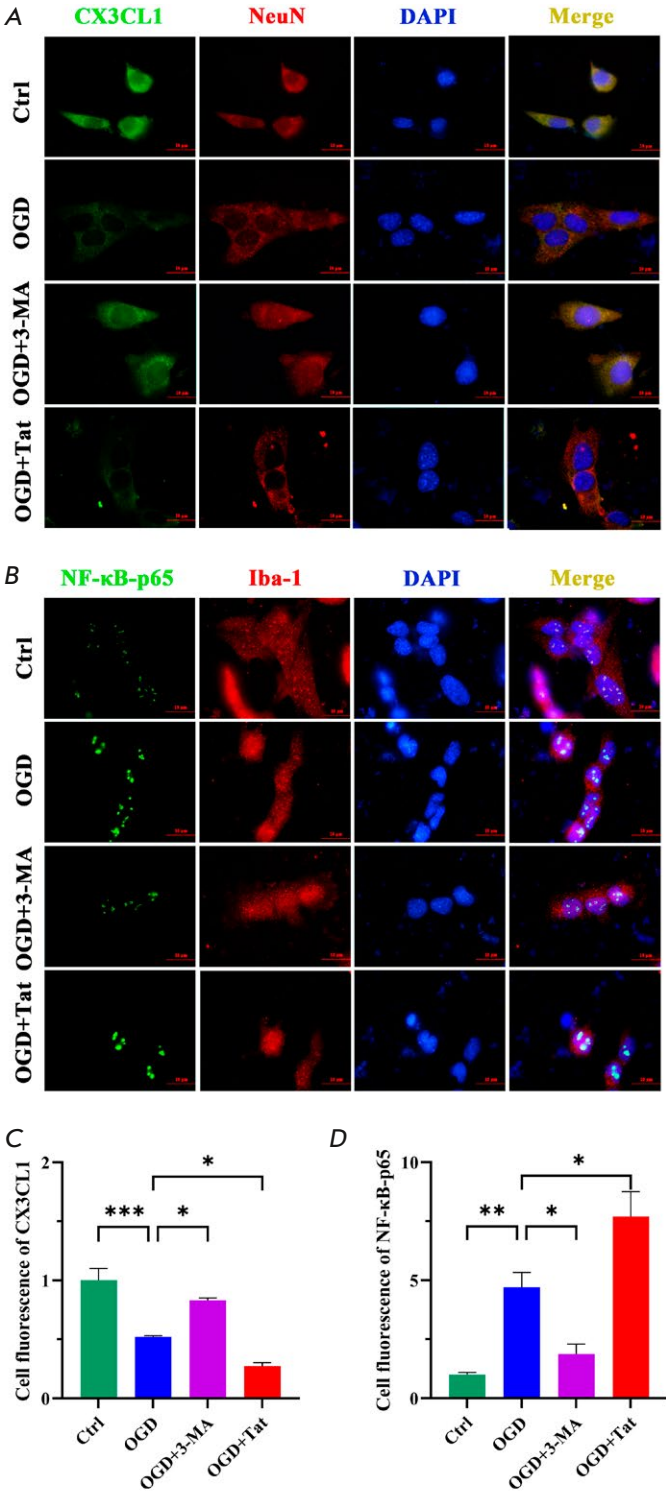


Fig. 4. Reducing HT22 autophagy suppressed BV2 inflammatory activation. (*A*, *B*) Immunofluorescence images of CX3CL1 (green) / NeuN (red), NF-κB-p65 (green) / Iba-1 (red), and DAPI (blue). (*C*, *D*) Cell fluorescence of CX3CL1 and NF-κB-p65 (green). Bar: 50 μm, $n = 6$. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

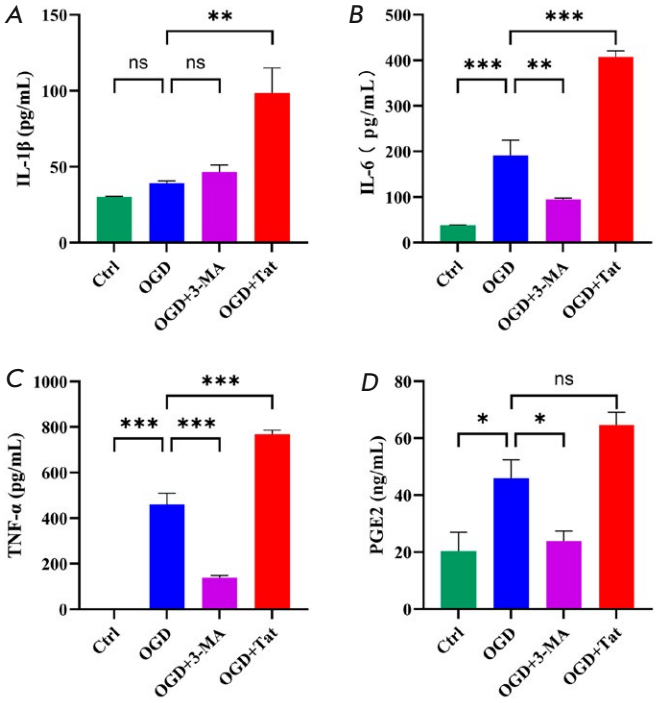


Fig. 5. Attenuation of autophagy in HT22 cells repressed the BV2-mediated inflammatory response. ELISA analysis of inflammatory factors including IL-1β (*A*), IL-6 (*B*), TNF-α (*C*), and PGE2 (*D*) converted from the standard curve by measurement of the OD_{450nm} value. $n = 6$. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, ns

2.5. Autophagy inhibition-suppressed microglial inflammation benefited neuron survival

To determine whether the CX3CL1-suppressed microglial inflammation could alleviate the OGD-induced neuronal injury, with the CCK-8 kit, Nissl staining and FJC staining were performed to evaluate cell viability, neuron survival, and cell death, respectively. The results showed that cell viability (*Fig. 6A,C*) and neuron survival (*Fig. 6E,F*) were significantly improved, while the cell death rate was correspondingly decreased in the OGD+3-MA group, compared with those in the OGD+Tat group, or in the OGD group. Conversely, the OGD-induced neuronal injury was further aggravated in the OGD+Tat group (*Fig. 6B,D*), compared with that in the OGD group.

3. DISCUSSION

Ischemic stroke caused by cerebrovascular occlusion is a fatal disease that threatens human beings [17]. The pathological mechanism underlying ischemic stroke has been extensively investigated in recent years, and yet the recombinant tissue plasminogen activator (rtPA) remains the only approved agent for stroke treatment [18]. Its clinical application can reduce the likelihood of disability by 25%. However, the therapeutic efficacy of rtPA rapidly drops past the 4.5 h that follow an ischemic stroke [19]. Besides, rtPA administration frequently increases the risk of a hemorrhage, which, as we know, aggravates brain injury [19]. Thrombectomy is another efficacious way to remove an infarction, but it may lead to more serious damage to the brain than the cerebral ischemia itself, because of the ischemia/reperfusion injury induced by the instantaneous complete blood resupply [20]. The neurons at the ischemic core rapidly die within several minutes after an ischemic stroke, but the cell death at the penumbra (the peripheral area around the core) lags, due to the milder ischemia by the blood supply with the arterial collateral anastomoses [21]. Mounting evidence points to the fact that neurons that have suffered from autophagy can be rescued back to life by modulation targeting at autophagic/lysosomal signaling [22]. However, neurons and glial cells coexist within penumbra tissues. Thus, this study particularly concerned itself with whether the fate of autophagic neurons is regulated by microglia, using a co-culture of neurons with microglia *in vitro*. This study might provide more clues as to how to improve stroke treatment.

Microglia are native immune cells that are responsible for neuroinflammation and are prominently activated by cerebral ischemia to maintain cellular homeostasis [23]. A modest microglial inflammation benefits neuroprotection, while an amplified immune response leads to neurological injury. CX3CL1, a chemokine anchored to the membranes on neurons, is efficacious in suppressing a microglial inflammation by binding to its receptor CX3CR1 expressed on microglia [6]. Under normal conditions, the microglial inflammatory response can be limited by neurons through the CX3CL1–CX3CR1 signaling pathway [24]. Studies have shown that neuronal autophagy at the penumbra is excessively activated, resulting in aggravated ischemic brain damage [25]. Furthermore, our previous study [16] established that the CX3CL1 expression was significantly reduced on autophagic neurons. We, therefore, looked into whether this reduced CX3CL1 on neurons weakened its suppressive effects on the microglial inflammatory response, resulting in the worsened neurological injury after an ischemic stroke. Based on the CX3CL1-CX3CR1 regulative mechanism, the correlation between neuronal autophagy and the microglial inflammation was investigated using an OGD co-culture of HT22 neurons with BV2 microglia.

Our study demonstrated that 1.5 h of OGD followed by 2 h of reoxygenation was the ideal culture condition under which autophagy was mostly induced

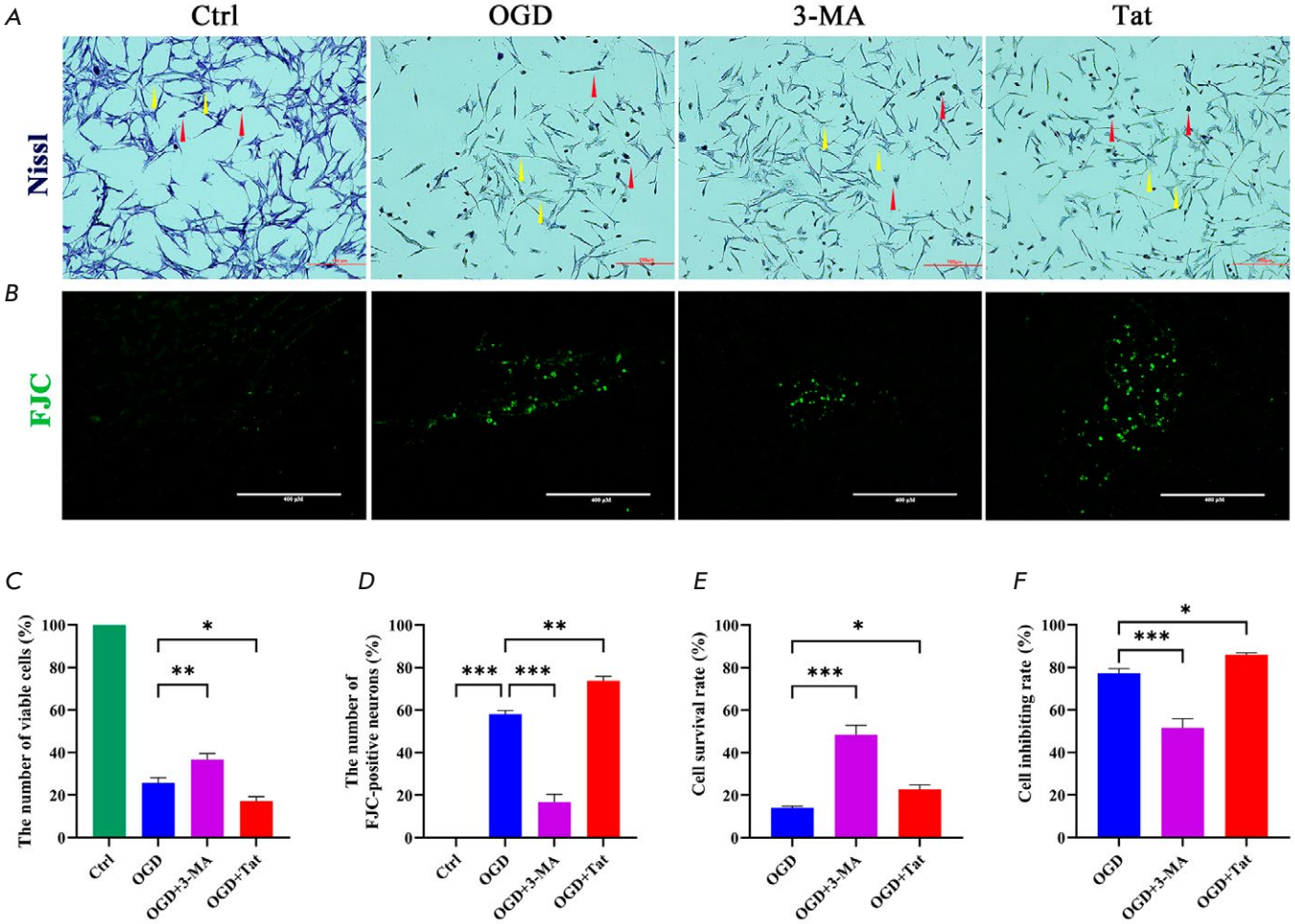


Fig. 6. Autophagy inhibition-suppressed microglial inflammation alleviated the neuronal injury after OGD. (*A*) Nissl staining images of pyknotic neurons (red arrows) and viable neurons (yellow arrows). (*B*) FJC staining images of degenerative neurons. (*C*) The number of FJC-positive cells was statistically analyzed. (*D*) The number of viable cells was statistically analyzed. (*E*) Statistical analysis of the cell survival rate detected by the CCK-8 kit. $n = 6$. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

in HT22 cells but barely in BV2 (*Fig. 1A,B*). The same study also indicated that neurons were more susceptible to ischemia than microglia, similarly to what was reported in [16]. CX3CL1 was uniquely expressed on neurons. We wondered whether its expression was altered in autophagic neurons. The results demonstrated that the significantly elevated autophagic activity was coupled with a markedly reduced CX3CL1 expression, suggesting that CX3CL1 expression is negatively regulated by autophagy in OGD HT22 neurons. Studies have indicated that a microglial inflammation could be suppressed by neurons through CX3CL1–CX3CR1 signaling [25]. Therefore, we discussed whether the microglial inflammatory injury was aggravated by the down-regulated CX3CL1 expression in OGD HT22 cells. Under OGD condition, the co-cultured cells were treated with the autophagy inhibitor 3-MA and the inducer Tat-Beclin1, respectively. The results showed that attenuation of HT22 autophagy significantly restored CX3CL1 expression (*Fig. 3C*). Consequently, the microglial inflammatory signaling of NF-κB-p65 was greatly suppressed (*Fig. 3D*). Meanwhile, the inflammatory factors of IL-6, IL-1β, TNF-α, and PGE2 were also attenuated (*Fig. 5*). By contrast, promoting HT22 autophagy further reduced its CX3CL1 expression and, in turn, exacerbated the inflammatory response. Moreover, the neuronal autophagy-worsened microglial inflammation led to increased death amongst HT22 cells (*Fig. 6B*). Conversely, down-regulation of autophagy alleviated the inflammatory injury and subsequently promoted neuron survival in OGD HTT cells (*Fig. 6A*). Our data collectively suggest that

neuronal autophagy aggravated the microglial inflammatory injury by reducing its CX3CL1, due to the disruption in CX3CL1–CX3CR1 communication that took place after ischemia. Contrarily, promotion of CX3CL1 on neurons by attenuating autophagy could have enhanced its suppressive effects on the microglial inflammatory response and, thereby, alleviate the ischemic injury in the neurons.

In summary, the main purpose of our study was to investigate the correlation between neuronal autophagy and microglial inflammation in a co-culture of neurons with microglia, based on the suppressive impact of neurons on the microglial inflammatory response through CX3CL1–CX3CR1 signaling. The results showed that the OGD-induced neuronal autophagy significantly decreases its CX3CL1 expression, which consequently exacerbates the microglial inflammatory response and neurological injury. Furthermore, promoting neuronal autophagy upon OGD further lessens its CX3CL1 expression and, in turn, worsens the microglial inflammation. Conversely, inhibiting autophagy effectively alleviates the microglial inflammatory injury by up-regulating CX3CL1 expression and, thereby, improving neuronal survival. Our data suggest that inhibiting neuronal autophagy might be a reliable way to alleviate the microglial inflammatory injury after an ischemic stroke. ●

Huang Z.W. designed and performed the experiments. Liu Y.Y., and Chen X.M. performed part of the experiments, collected the data, and did the statistical analysis. He H.Y. and Deng Y.H. provided the ideal, supervised and corrected the research strategies, wrote and revised the manuscript. All authors have read and agreed to the published version of this manuscript.

The authors declare that they have no conflict of interest.

This work was supported by grants from the National Natural Science Foundation of China (Nos. 81960418, 82160240, 82160241), and Yunnan Ten Thousand Talents Plan Young & Elite Talents Project (No. YNWR-QNBJ-2018-034), and Yunnan Applied Basic Research Projects Fund of Yunnan Provincial Department of Science & Technology (No. 202001AT070049).

REFERENCE

1. Wang P., Shao B.Z., Deng Z., Chen S., Yue Z., Miao C.Y. // Prog. Neurobiol. 2018. V. 163. P. 98–117.
2. Iadecola C., Anrather J. // Nat. Med. 2011. V. 17. № 7. P. 796–808.
3. Su P., Zhang J., Wang D., Zhao F., Cao Z., Aschner M., Luo W. // Neuroscience. 2016. V. 319. P. 155–167.
4. Berglund R., Guerreiro-Cacais A.O., Adzemovic M.Z., Zeitelhofer M., Lund H., Ewing E., Ruhrmann S., Nutma E., Parsa R., Thessen-Hedreul M., et al. // Sci. Immunol. 2020. V. 5. № 52. P. 5077.
5. Lu M., Zhao W., Han S., Lin X., Xu T., Tan Q., Wang M., Yi C., Chu X., Yang W., et al. // Sci. Adv. 2022. V. 8. № 26. P. 8048.
6. Mendiola A.S., Church K.A., Cardona S.M., Vanegas D., Garcia S.A., Macklin W., Lira S.A., Ransohoff R.M., Kokovay E., Lin C-HA., et al. // J. Neurochem. 2022. V. 162. № 5. P. 430–443.
7. Chen X., Jiang M., Li H., Wang Y., Shen H., Li X., Zhang Y., Wu J., Yu Z., Chen G. // J. Neuroinflammation. 2020. V. 17. № 1. P. 209–224.
8. Lauro C., Chece G., Monaco L., Antonangeli F., Peruzzi G., Rinaldo S., Paone A., Cutruzzolà F., Limatola C. // Front. Cell Neurosci. 2019. V. 13. P. 414.
9. Xu S., Lu J., Shao A., Zhang J.H., Zhang J. // Front. Immunol. 2020. V. 11. P. 294.
10. Subbarayan M.S., Joly-Amado A., Bickford P.C., Nash K.R. // Pharmacol. Ther. 2022. V. 231 P. 107989.
11. Liu M., Xu Z., Wang L., Zhang L., Liu Y., Cao J., Fu Q., Liu Y., Li H., Lou J., et al. // J. Neuroinflammation. 2020. V. 17. № 1. P. 270.
12. Campbell B.C.V., Khatri P. // Lancet. 2020. V. 396. № 10244. P. 129–142.
13. Zhang Y., Wu Z., Huang Z., Liu Y., Chen X., Zhao X., He H., Deng Y. // Brain Res. 2022. V. 1778. P. 147768.
14. Camargos Q.M., Silva B.C., Silva D.G., Toscano E.Cd.B., Oliveira B.d.S., Bellozi P.M.Q., Jardim B.L.d.O., Vieira É.L.M., de Oliveira A.C.P., Sousa L.P., et al. // Brain Res. Bull. 2020. V. 155. P. 1–10.
15. Ermine C.M., Bivard A., Parsons M.W., Baron J.-C. // Int. J. Stroke. 2021. V. 16. № 5. P. 497509.
16. He H.Y., Ren L., Guo T., Deng Y.H. // Neural. Regen. Res. 2019. V. 14. № 2. P. 280–288.
17. Wu Z., Zhang Y., Liu Y., Chen X., Huang Z., Zhao X., He H., Deng Y. // Life (Basel). 2021. V. 11. № 9. P. 948.
18. Ajoolabady A., Wang S., Kroemer G., Penninger J.M., Uversky V.N., Pratico D., Henninger N., Reiter R.J., Bruno A., Joshipura K., et al. // Pharmacol. Ther. 2021. V. 225. P. 107848.
19. Barthels D., Das H. // Biochim. Biophys. Acta Mol. Basis Dis. 2020. V. 1866. № 4. P. 165260.
20. Al-Mufti F., Amuluru K., Roth W., Nuoman R., El-Ghanem M., Meyers P.M. // Neurosurgery. 2018. V. 82. № 6. P. 781–789.
21. Yu S., Yu M., Bu Z., He P., Feng J. // Front. Cell Neurosci. 2020. V. 14. P. 193.
22. Lingling D., Miaomiao Q., Yili L., Hongyun H., Yihao D. // Brain Res. Bull. 2022. V. 184. P. 24–33.
23. Lian L., Zhang Y., Liu L., Yang L., Cai Y., Zhang J., Xu S. // Front. Mol. Neurosci. 2020. V. 13. P. 612439.
24. Cipriani R., Villa P., Chece G., Lauro C., Paladini A., Micotti E., Perego C., De Simoni M.G., Fredholm B.B., Eusebi F., et al. // J. Neurosci. 2011. V. 31. № 45. P. 16327–16335.
25. Mao M., Xu Y., Zhang X.Y., Yang L., An X.B., Qu Y., Chai Y.N., Wang Y.R., Li T.T., Ai J. // J. Neuroinflammation. 2020. V. 17. № 1. P. 244.

A DNA Replication Stress-Based Prognostic Model for Lung Adenocarcinoma

S. Shi[1], G. Wen[1], C. Lei[1], J. Chang[1], X. Yin[1], X. Liu[1], S. Huang[2*]
[1]Department of Cardiothoracic Surgery, The People's Hospital of Dazu District, Chongqing, 402360 China
[2]Department of Orthopedics, The People's Hospital of Dazu District, Chongqing, 402360 China
*E-mail: cqslhuang@163.com
Received: July 08, 2023; in final form, September 25, 2023
DOI: 10.32607/actanaturae.25112
Copyright © 2023 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License,which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **Tumor cells endure continuous DNA replication stress, which opens the way to cancer development. Despite previous research, the prognostic implications of DNA replication stress on lung adenocarcinoma (LUAD) have yet to be investigated. Here, we aimed to investigate the potential of DNA replication stress-related genes (DNARSs) in predicting the prognosis of individuals with LUAD. Differentially expressed genes (DEGs) originated from the TCGA-LUAD dataset, and we constructed a 10-gene LUAD prognostic model based on DNARSs-related DEGs (DRSDs) using Cox regression analysis. The receiver operating characteristic (ROC) curve demonstrated excellent predictive capability for the LUAD prognostic model, while the Kaplan-Meier survival curve indicated a poorer prognosis in a high-risk (HR) group. Combined with clinical data, the Riskscore was found to be an independent predictor of LUAD prognosis. By incorporating Riskscore and clinical data, we developed a nomogram that demonstrated a capacity to predict overall survival and exhibited clinical utility, which was validated through the calibration curve, ROC curve, and decision curve analysis curve tests, confirming its effectiveness in prognostic evaluation. Immune analysis revealed that individuals belonging to the low-risk (LR) group exhibited a greater abundance of immune cell infiltration and higher levels of immune function. We calculated the immunopheno score and TIDE scores and tested them on the IMvigor210 and GSE78220 cohorts and found that individuals categorized in the LR group exhibited a higher likelihood of deriving therapeutic benefits from immunotherapy intervention. Additionally, we predicted that patients classified in the HR group would demonstrate enhanced sensitivity to Docetaxel using anti-tumor drugs. To summarize, we successfully developed and validated a prognostic model for LUAD by incorporating DNA replication stress as a key factor.**

KEYWORDS **DNA replication stress, lung adenocarcinoma, prognostic model, immunotherapy response, anti-tumor drug prediction.**

INTRODUCTION

Lung cancer (LC) is a highly heterogeneous and lethal malignancy, representing a significant contributor to cancer incidence and mortality rates [1]. Lung adenocarcinoma (LUAD) stands as the predominant subtype of LC [2]. Surgery and radiation therapy offer hope for curing LUAD patients, while chemotherapy, targeted therapy, and immunotherapy can maximize the improvement of tumor prognosis. However, the prognosis for patients with LUAD still poses a significant challenge, with a relatively low long-term survival rate [3]. Parameters such as tumor size, TNM staging, and tumor grading cannot meet the demands of prognosis prediction and more precise treatment guidance, and finding new evaluation methods is a pressing need for precision medicine. The establishment of robust prognostic risk models holds the potential to significantly enhance our ability to forecast the prognosis of individuals diagnosed with LUAD.

The preservation of genome integrity heavily relies on the integrity and accuracy of DNA replication. However, the DNA replication process constantly faces challenges from various intrinsic and extrinsic stresses, including DNA damage and other factors, which can pose threats to overall genomic stability [4]. Various obstacles that delay, prevent, or terminate DNA replication are defined as DNA replication stress [5]. DNA replication stress activated by oncogene abnormalities is an important factor affecting cancer progression. On the one hand, it abets genomic instability, advancing cancer development. On the

other hand, it retards cell proliferation and triggers anti-cancer defense mechanisms to induce cell apoptosis or senescence [6]. Tumor cells frequently exhibit a prominent characteristic of chronic replication stress, which arises from the persistent presence of replication stress sources due to impaired replication stress responses, diminished repair protein activity, and ongoing proliferation signal transduction. This chronic replication stress contributes significantly to the genomic instability and aberrant cell proliferation observed in tumor cells [7]. Previous studies have found that the DNA replication stress-related genes POLQ, PLK51, RAD6, CLASPIN, and CDC14 can predict the prognosis of early and mid-stage non-small cell LC (NSCLC) patients [8]. Additionally, DNA replication stress is an important mechanism for the chemotherapy and targeted therapy of LC. The integration of immunotherapy with these therapies represented a compelling strategy to augment the efficacy of LC treatment [9]. Therefore, the value of DNA replication stress-related genes (DNARSs) lies in their potential to be valuable prognostic markers and aid in predicting drug efficacy in the context of LUAD.

The proportion of immune cell infiltration in the tumor microenvironment (TME) affects cancer patient survival and the immunotherapy response [10, 11]. The expression levels of immune checkpoint inhibitors (ICIs) like cytotoxic T lymphocyte-associated protein 4 (CTLA4) and programmed cell death protein 1 (PD1)/programmed cell death ligand 1 (PD-L1) are usually significantly increased in hypoxic malignant tumors, and ICIs are more effective for a small proportion of LC patients [12]. However, there are currently no tools available for forecasting the efficacy of immunotherapy in LUAD individuals.

We hereby used bioinformatics analysis to assess LUAD feature genes related to DNA replication stress and analyzed their roles in predicting the prognosis and drug efficacy for LUAD individuals.

MATERIALS AND METHODS

Data collection

Gene expression datasets of LUAD with complete clinical data, including age, gender, tumor grade, and TNM staging, were provided by The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/) and Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/) databases. The TCGA-LUAD dataset (539 cancer tissue samples and 59 normal tissue samples) was utilized as the training set, while the GSE26939 dataset (116 LUAD cancer tissue samples, platform number GPL9053) was used as the validation set.

Twenty-one DNA replication stress features were obtained from references, including 982 DNARSs (*Table 1*) [13, 14].

We collected the gene sequencing data of 119 tumor samples from individuals with urothelial cancer treated with atezolizumab (anti-PD-L1) from the IMvigor210 immune therapy cohort [15]. The GSE78220 dataset (platform number GPL11154) contained tumor samples from melanoma patients treated with anti-PD-1 therapy and was supplied by the GEO database [16].

Differential analysis

The R package "edgeR" [17] was used to conduct a differential analysis on LUAD tissue specimens and normal tissue specimens in the training set, and the differentially expressed genes (DEGs) of LUAD were selected along the criteria of standard FDR < 0.05 and $|\log(FC)| > 1$. The intersection of DEGs and DNARSs was used to obtain the LUAD differential genes associated with DNA replication stress (DRSDs).

Prognostic model construction and evaluation

We first screened LUAD tumor patient specimens with a survival time greater than 30 days from the training set based on clinical data. Then, the univariate Cox regression analysis was tapped utilizing the R package "survival" (https://CRAN.R-project.org/package=survival) to select the genes in DRSDs significantly associated with the overall survival (OS) of LUAD individuals. To mitigate the risk of overfitting in the statistical model, we employed the LASSO Cox analysis to identify a subset of feature genes from the larger pool of identified genes, utilizing the R packages "glmnet" [18] and "survival." Feature genes were subjected to a multivariate Cox regression analysis to establish the LUAD prognostic model, using R packages "survival" and "survminer" (https://rdocumentation.org/packages/survminer/versions/0.4.9). The formula for calculating the Riskscores was

$$\text{Riskscore} = \Sigma\ \text{Coefficient (gene)} \times \text{Expressionvalue (gene)}.$$

Coefficient is the coefficient of the gene. Expressionvalue is the relative expression level of gene standardized by Z-score.

Riskscore was calculated for each LUAD patient sample in both the training and validation sets, and the samples were separated as high-risk (HR) and low-risk (LR) groups as per the median value. The distribution of Riskscore scores, patient survival status, and expression levels of feature factors in the two risk groups of LUAD patient specimens in the train-

ing set were analyzed. Kaplan-Meier survival curves were constructed utilizing the R package "survival" to compare the difference in the survival rates between the patients in the two groups. Receiver operating characteristic (ROC) curves were constructed using the R packages "timeROC" [19] and "survival" to calculate the area under the curve (AUC) and test the prognostic performance of the model.

Independent prognostic analysis, nomogram construction, and evaluation

The Riskscore from the training set was used as the single feature and combined with clinical data to perform univariate Cox and multivariate Cox regression analyses, evaluating the independent ability of the model to predict the patient survival chances. A LUAD prognostic nomogram was constructed using clinical factors and Riskscore, and a calibration curve was utilized to evaluate the disparity between the predicted event rate and the actual event rate. The R packages "rms" [20] and "survival" were used for this analysis. The ROC curves were depicted utilizing the R packages "timeROC" [19] and "survival" to evaluate the performance of the model in forecasting the prognosis of LUAD patients based on nomogram, Riskscore, age, gender, tumor grade, and TNM staging. The standardized net benefit of the nomogram was analyzed using the decision curve analysis (DCA).

Tumor immune analysis

Immune infiltration analysis was done utilizing the R packages "GSVA" [21] and "estimate" (https://R-Forge.R-project.org/projects/estimate/). The ssGSEA method was used to analyze immune cell infiltration and function in the HR and LR groups, and the expression of human leukocyte antigen (HLA)-related genes was evaluated. The differences between different risk groups were compared using the Wilcoxon test.

Prediction of immunotherapy response

To forecast the response of the HR and LR groups to immunotherapy, a series of studies were conducted. Immune checkpoints expression was analyzed in the two groups. The immunophenoscore (IPS) demonstrates high accuracy in predicting the response to anti-CTLA-4 and anti-PD-1 therapies, making it a valuable tool for determining the tumor's likelihood of responding to ICI therapy. The IPS score of each patient was obtained from The Cancer Immunome Atlas (TCIA, https://tcia.at), and the differences in IPS scores between the two groups were compared. Tumor Immune Dysfunction and Exclusion (TIDE) can forecast the response to immunotherapy by simulating the main mechanisms of tumor immune escape. We employed TIDE score to predict the response of the two groups to ICI immunotherapy.

Furthermore, we used the Imvigor210 immune therapy cohort of individuals with urothelial cancer treated with the anti-PD-L1 inhibitor atezolizumab and the GSE78220 transcriptome dataset of melanoma individuals treated with anti-PD1 to test the effectiveness of the model in predicting the response to immunotherapy, including treatment efficacy and survival.

Anti-tumor drug screening

To identify potential targets and effective drugs, we used the CellMiner database (https://discover.nci.nih.gov/cellminer/) and R package "pRRophetic" (https://github.com/paulgeeleher/pRRophetic/) to screen for anti-tumor drugs related to the IC_{50} of feature genes. Different drug IC_{50} values were predicted in the two groups, with lower IC_{50} values indicating a more effective cancer treatment [22].

RESULTS

Identification of DRSDs

This study's training set included expression data from 539 LUAD cancer tissue specimens and 59 normal tissue specimens. DEGs of the LUAD differential gene sets were obtained through a differential analysis, including 6,005 genes. Among the analyzed genes, we observed differential upregulation in 4,217 genes and differential downregulation in 1,788 genes (*Fig. 1A, Table 2*). Intersection of the DNARSs with 982 genes and DEGs was taken to obtain the Venn diagram of DRSDs, which contained 279 genes (*Fig. 1B*).

Establishment of a prognostic model

To develop robust risk features for clinical use, a series of Cox regression analyses were conducted. First, 163 genes that may affect OS were identified from the 279 genes in DRSDs through univariate Cox analysis. Then, 10 candidate genes were determined using LASSO regression (*Fig. 2A,B*). Multivariate Cox analysis showed that the coefficients of 10 feature genes were non-zero, with NT5E being a prognostic risk factor and GTF2H4 being a protective factor. The model was established ground on 10 genes (*Fig. 2C*). The 10-gene LUAD prognostic risk model based on DNA repair stress is shown below:

$$\text{Riskscore} = 0.05 \times \text{HMMR} + 0.03 \times \text{TEX15} + 0.04 \times \text{PLK1} + 0.10 \times \text{EX01} + 0.09 \times \text{H2BC4} + 0.21 \times \text{H2AX} - 0.08 \times \text{GTF2H4} + 0.19 \times \text{NME4} + 0.09 \times \text{UCK2} + 0.16 \times \text{NT5E}$$

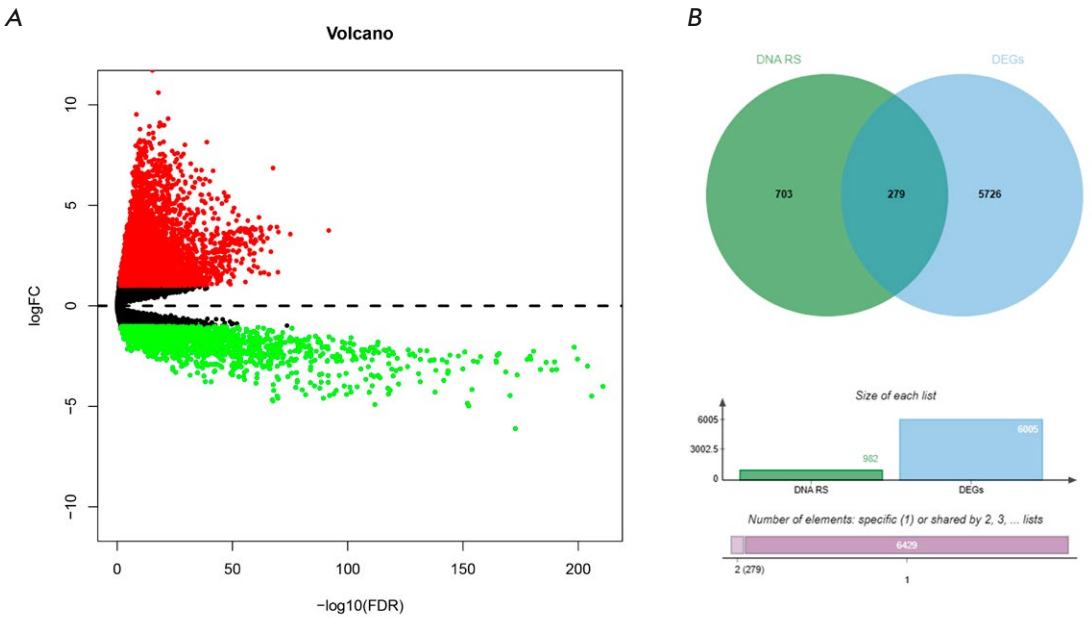


Fig. 1. Screening of DRSDs. (*A*) Volcano plot of DEGs related to LUAD. (*B*) Venn diagram of the intersection between DEGs and DNARSs, corresponding to DRSDs

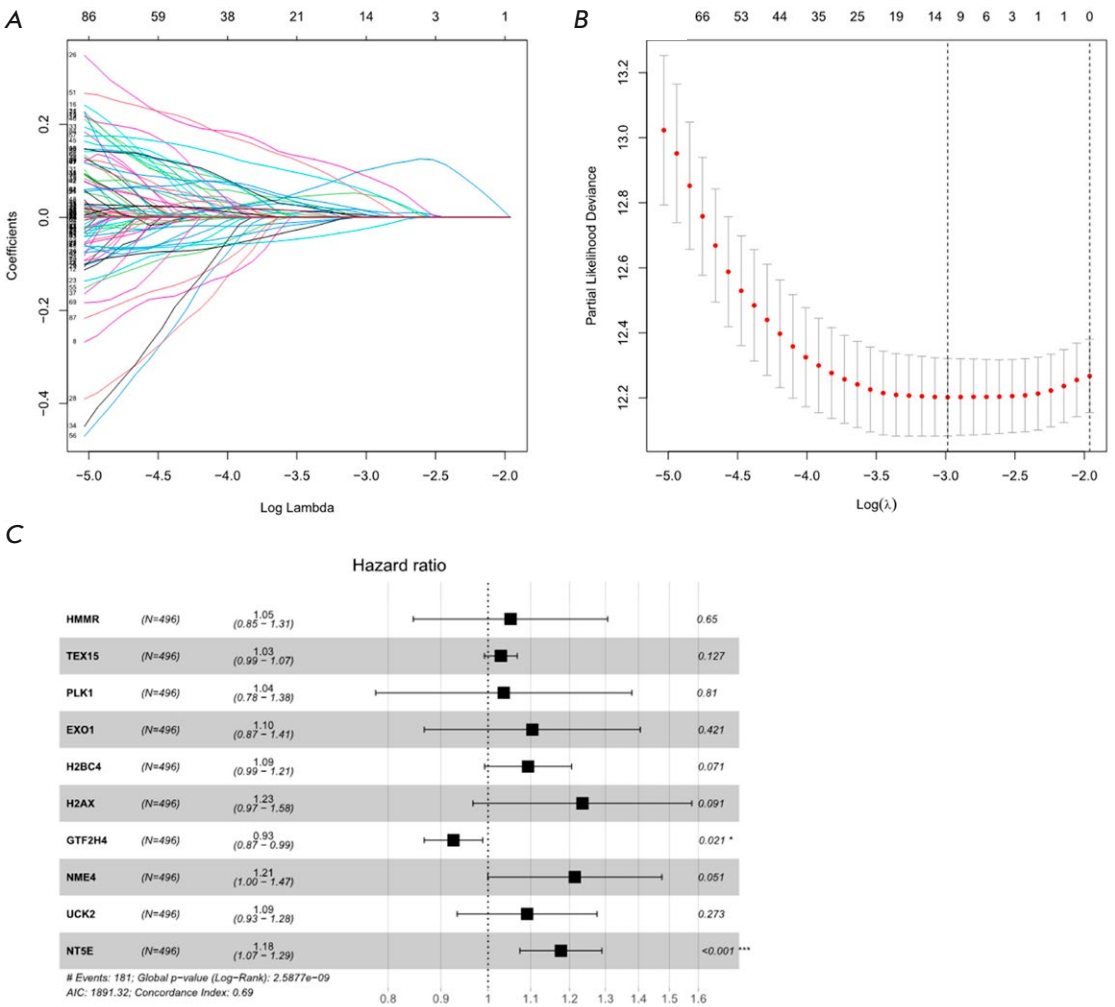


Fig. 2. Construction of a LUAD prognostic model using DRSDs. (*A*) Cross-validation plot of the logarithmic (λ) sequence in the LASSO model, with the selection of the best parameter (lambda) indicated by the first black dotted line. (*B*) LASSO coefficient spectrum of 10 OS-related genes. (*C*) Forest plot of the multivariate Cox regression analysis based on the 10 feature genes in DRSDs

Evaluation of the prognostic model

The Riskscore of LUAD samples in both the training and validation sets were computed by utilizing the LUAD prognostic risk model, and specimens were divided into HR and LR groups accordingly. The distribution of Riskscore values and survival status within the training set revealed that patients in the HR group exhibited a higher mortality rate (*Fig. 3A,B*). The heatmap of feature gene expression in the training set samples showed that all genes except GTF2H4 were highly expressed in the HR group (*Fig. 3C*). From the training set, we found that the survival rate of HR patients was lower ($P < 0.05$), indicating better overall prognosis for LR individuals (*Fig. 3D*). The ROC curve of the training set showed that the AUC values for 1-, 3-, and 5-year were between 0.67 and 0.74, indicating good sensitivity and specificity of the risk model (*Fig. 3E*). External validation of the validation set showed that patients in the HR group had a lower survival rate than those in the LR group ($P < 0.05$) (*Fig. 3F*). The ROC curve of the validation set showed the AUC values for 1-, 3-, and 5-year were between 0.69 and 0.73, proving that the risk model also did well in the validation set (*Fig. 3G*). In summary, the LUAD prognostic model based on DRSDs exhibits high accuracy and reliability in predicting patient likelihood of survival.

Independent prognostic analysis

To examine the independent impact of Riskscore on the survival of LUAD patients, we conducted both univariate and multivariate Cox analyses. These analyses involved incorporating the patients' Riskscore along with other relevant clinical-pathological indicators. The findings revealed that Riskscore independently served as a prognostic factor for LUAD patients' OS (*Fig. 4A,B*). Then, we combined Riskscore with prognostic clinical features to construct a nomogram for a more comprehensive prediction of patient chances of survival (*Fig. 4C*). According to the calibration curve, the nomogram predicted the OS of LUAD individuals at 1-, 3-, and 5-year with little difference from the ideal model (Fig. 4*D–F*). The ROC curve illustrated that the AUC values of Riskscore and the nomogram were 0.7 and 0.73, respectively, higher than those of other clinical factors, indicating good prognostic predictive ability (*Fig. 4G*). We analyzed the clinical net benefit of the nomogram via DCA curve analysis, which showed that the nomogram was of clinical utility in forecasting the prognosis of LUAD individuals (*Fig. 4H*). Therefore, the nomogram established here helped predict the survival probability of LUAD patients.

Tumor immune cell infiltration

Tumor immune cell infiltration is tightly linked to tumor progression [23]. By analyzing the immune cell infiltration and immune-related functional pathways between the two groups, we probed the disparities in the immune activity status between the two groups (*Fig. 5A,B*). The proportions of immune cell infiltration of dendritic cells (aDCs, iDCs), B_cells, Mast_cells, Neutrophils, T_helper_cells, and TIL were tellingly downregulated in the HR group ($P < 0.05$) (*Fig. 5A*). The immune-related pathway APC_co-inhibition was notably upregulated, while HLA and Type_II_IFN_Response were significantly downregulated in the same group ($P < 0.05$) (*Fig. 5B*). In addition, most HLA genes were significantly downregulated in the same group ($P < 0.05$) (*Fig. 5C*). In summary, the proportion of immune cell infiltration in HR LUAD patients was lower compared to that in the LR group.

Prediction of immunotherapy response

The Riskscore of LUAD individuals is tightly linked to their immune function, suggesting that the HR and LR groups may have different responses to immunotherapy. Therefore, we further explored the ability of the prognostic model to predict the immunotherapy response of cancer individuals. Expression of most immune checkpoints was notably higher in the LR group, with significant differences ($P < 0.05$) (*Fig. 6A*). The IPS score indicated that individuals in the LR group exhibited a better response to CTLA-4 and anti-PD-1 treatment, denoting that LR LUAD individuals had stronger immunogenicity and were more likely to benefit from immune therapy ($P < 0.05$) (*Fig. 6B*). LR LUAD individuals with lower TIDE scores indicated a weaker inclination to evade the immune system and a stronger inclination to benefit from immune therapy, with significant differences ($P < 0.05$) (*Fig. 6C*). Since there is currently no transcriptome data on the response of LUAD individuals to ICI treatment, we used other cancer data to ascertain the performance of the model in predicting the immunotherapy response. Using the IMvigor210 and GSE78220 datasets to verify the response of the HR and LR groups, we found that the samples responsive to immunotherapy in the LR group were higher than those in the HR group (*Fig. 6D–E*), and that OS of the LR group was tellingly better than that of the HR group, showing a better survival trend (*Fig. 6F–G*). In summary, LR LUAD patients displayed a greater likelihood of responding to immunotherapy than HR patients and had a better prognosis.

Prediction of potential anti-cancer drugs

To mine the response of LUAD patients to anti-cancer drug treatment, we dissected the linkage between

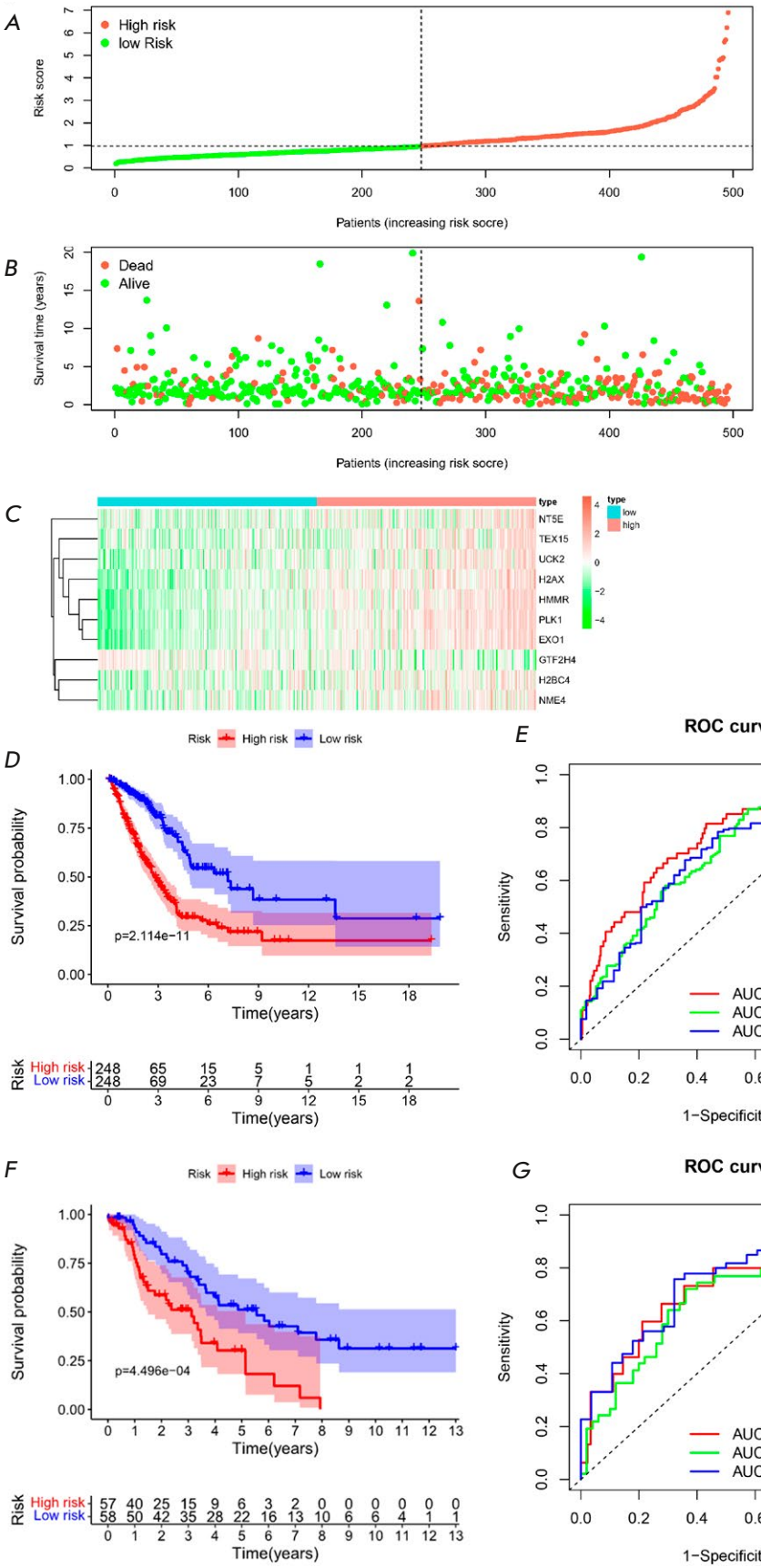


Fig. 3. Performance evaluation of the prognostic model in predicting the prognosis risk of LUAD patients. (*A*) Distribution of Riskscore values in the TCGA training set, with the dotted line indicating the optimal threshold between the LR and HR groups. (*B*) Distribution of survival status in the TCGA training set, with the dotted line indicating the optimal threshold between the LR and HR groups. (*C*) Heatmap of the expression levels of the 10 feature genes in the TCGA training set. (*D*) Kaplan-Meier survival curve in the TCGA training set. (*E*) ROC curve in the TCGA training set. (*F*) Kaplan-Meier survival curve in the GEO validation set. (*G*) ROC curve in the GEO validation set

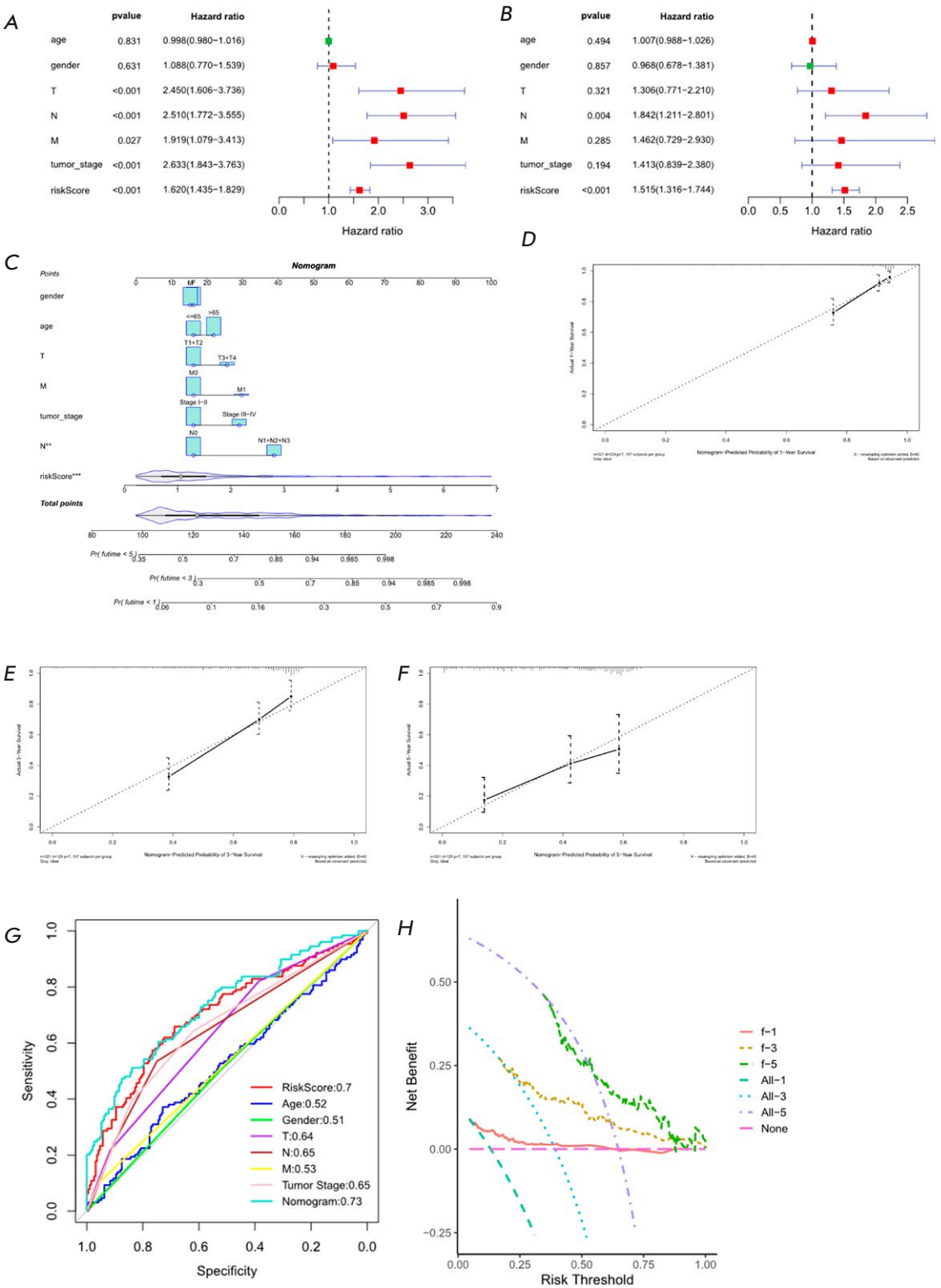


Fig. 4. Independent prognosis analysis of Riskscore in LUAD patients in the TCGA training set. (*A*) Forest plot of the univariate Cox regression analysis combining Riskscore with clinical information. (*B*) Forest plot of the multivariate Cox regression analysis combining Riskscore and clinical information on LUAD patients. (*C*) Nomogram constructed by combining Riskscore and clinical information. (*D*), (*E*), and (*F*) Calibration curves for predicting the risk of 1-, 3-, and 5-year death, respectively. (*G*) Clinical features, Riskscore, and ROC curve used to diagnose Nomograms. (*H*) DCA curve for diagnosing Nomograms

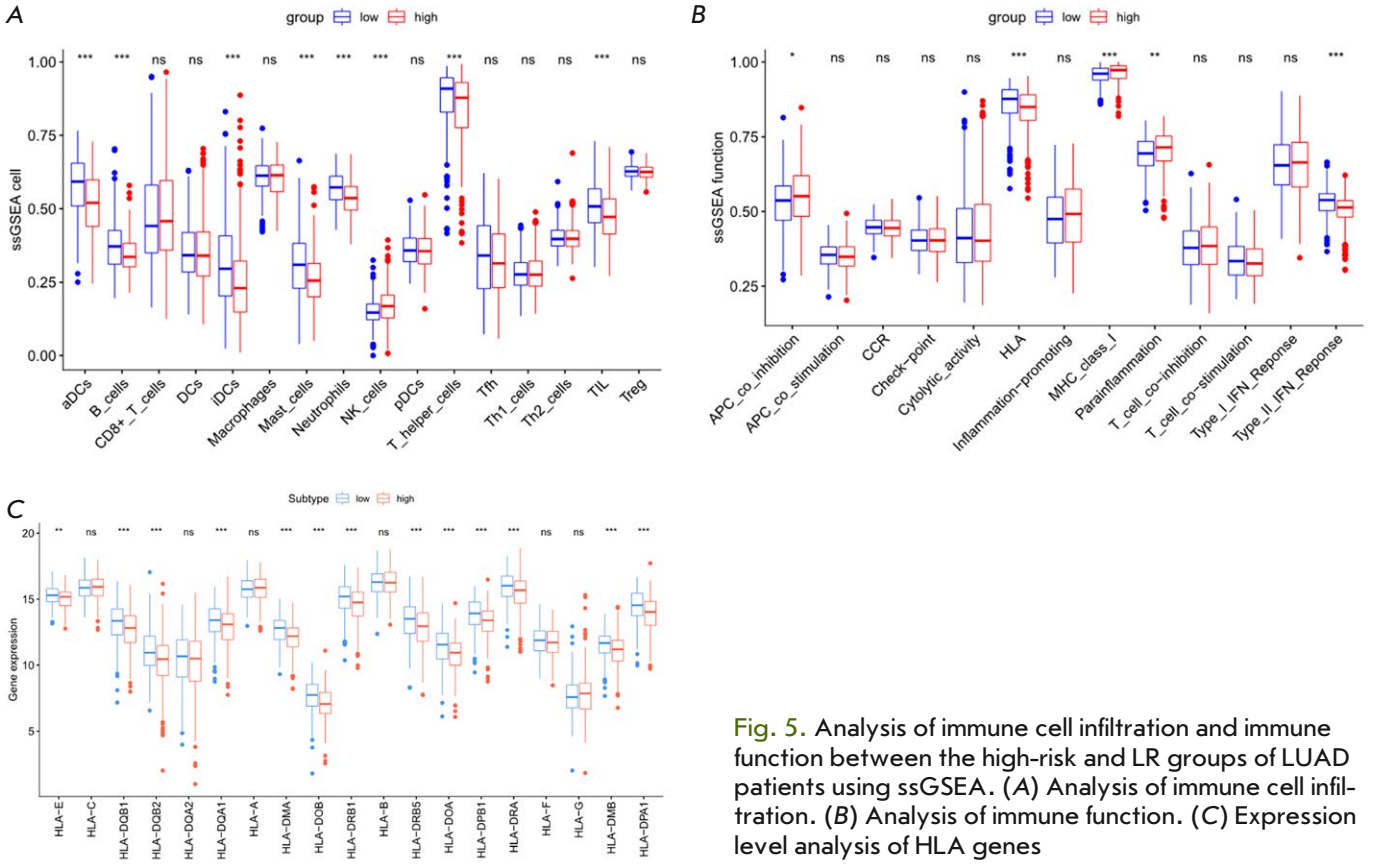


Fig. 5. Analysis of immune cell infiltration and immune function between the high-risk and LR groups of LUAD patients using ssGSEA. (*A*) Analysis of immune cell infiltration. (*B*) Analysis of immune function. (*C*) Expression level analysis of HLA genes

the expression of prognostic feature genes and the IC_{50} values of drug antagonists, with results displaying a significant positive linkage between the expression of the PLK1 and IC_{50} value of 5-Fluoro deoxy uridine 10mer (cor = 0.510), while the expression level of NT5E showed a significant negative linkage with the IC_{50} values of Idarubicin (cor = -0.510), XR-5944 (cor = -0.501), and Fluorouracil (cor = -0.499) (*Fig. 7A*). Furthermore, we investigated the association between the prognostic risk and drug sensitivity. The findings revealed that the HR group, characterized by a poor OS, exhibited heightened sensitivity to the drugs FTI-277, JNK Inhibitor VIII, CCT018159, and Docetaxel ($P < 0.001$) (*Fig. 7B*).

CONCLUSION

Despite the availability of various treatments like surgery, radiotherapy, chemotherapy, targeted therapy, and immunotherapy, the mortality rate of LUAD remains high. DNA replication abnormalities are the main cause of genomic instability leading to tumor initiation and progression [24]. DNA replication stress not only affects the autonomous cell response of cancer patients, but also alters the cellular microenvironment, activates innate immune responses, and helps the organism to protect itself against proliferating damaged cells [25]. Here, we developed a LUAD prognosis model grounded in DNARSs. In the training and validation cohorts, our novel LUAD prognosis model showed a reliable prognostic prediction performance and can serve as an independent prognostic tool for LUAD patients. The nomogram grounded in the Riskscore and clinical factors exhibits reliability and accuracy in forecasting the survival probability of LUAD individuals. The LR group of LUAD patients is characterized by high anti-tumor immune cell infiltration and high immune activity status.

Based on the Cox regression analysis, we obtained ten DNA replication stress biomarkers that impact the prognosis for LUAD individuals, including HMMR, TEX15, PLK1, EXO1, H2BC4, H2AX, NME4, UCK2, NT5E, and GTF2H4. The expression levels of HMMR, TEX15, PLK1, EXO1, H2BC4, H2AX, NME4, UCK2, and NT5E increased with increase in Riskscore. High expression of HMMR fosters malignant behaviors in LUAD individuals [26]. PLK1 mediates the phosphorylation of SKA3 and enhances the stability of the SKA3 protein, thereby promoting the malignant progression of LC [27]. The high expression of the EXO1 gene is an independent risk factor for a poor

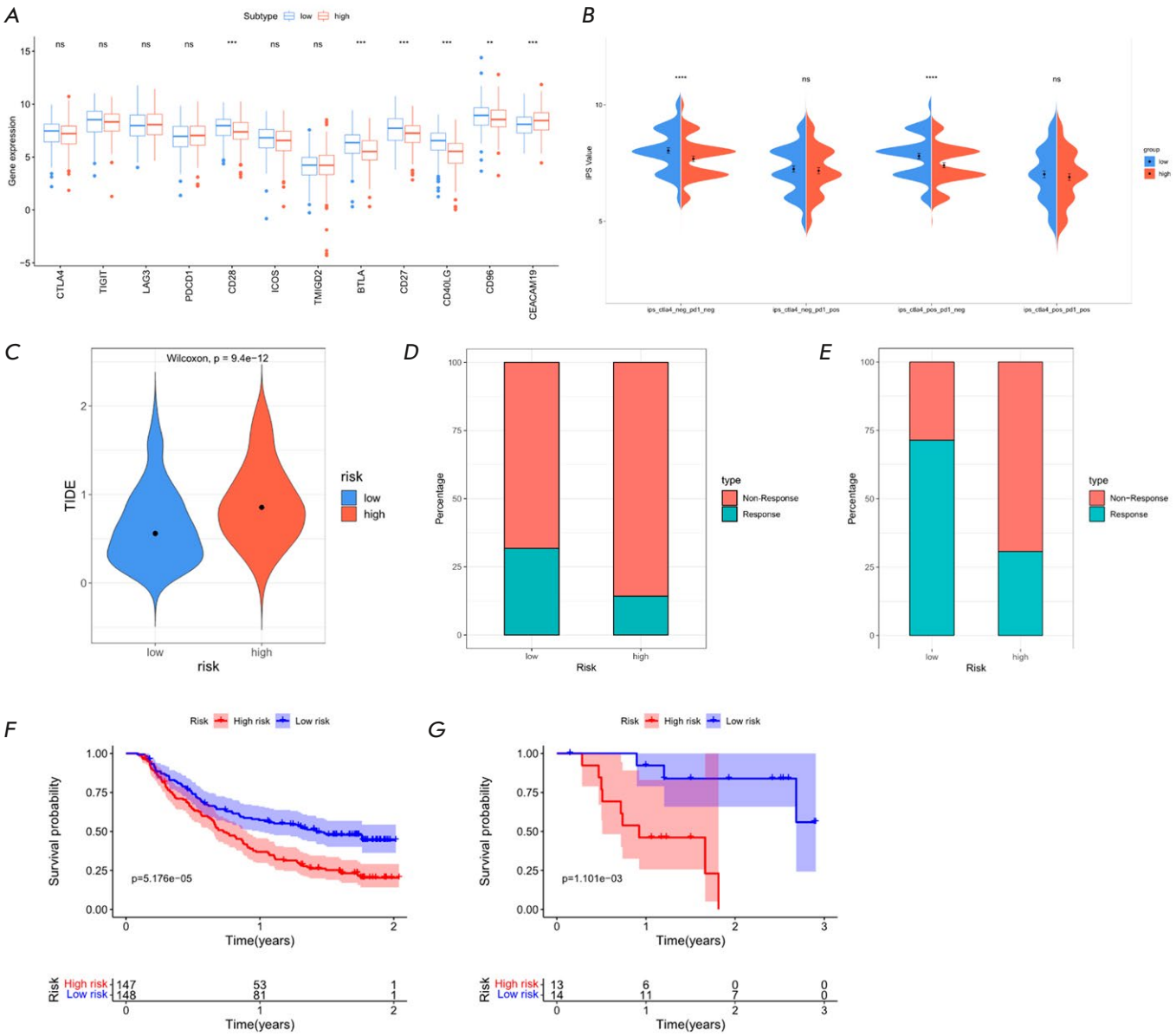


Fig. 6. Analysis of the immunotherapy response in the HR and LR groups of LUAD patients. (*A*) Boxplot of immune checkpoint expression levels in the HR and LR groups of LUAD patients. (*B*) Violin plot of IPS scores in the HR and LR groups of LUAD patients. (*C*) Violin plot of TIDE scores in the HR and LR groups of LUAD patients. (*D*) ICI treatment response of the HR and LR groups of LUAD patients in the iMvigor210 cohort. (*E*) ICI treatment response of the HR and LR groups of LUAD patients in the GSE78220 cohort. (*F–G*) Kaplan-Meier survival curve of the HR and LR groups of LUAD patients in the iMvigor210 (*F*) and GSE78220 cohorts (*G*), respectively

prognosis of LUAD, and EXO1 can also predict the response to chemotherapy [28–30]. Phosphorylated γH2AX at Ser-139 is a cellular response to DNA double-strand breaks and DNA damage, which features in tumor cell apoptosis. Studies have reported that the expression of γH2AX can predict the efficacy of ICI treatment in LUAD [31, 32]. NME4 affects NSCLC by overcoming cell cycle arrest and enhancing cell proliferation [33]. UCK2 is a rate-limiting enzyme in the pyrimidine salvage synthesis pathway, which promotes LC cell proliferation and migration [34, 35]. The NT5E gene encodes CD73, which promotes LUAD proliferation and metastasis via the EGFR/AKT/mTOR axis [36, 37]. Additionally, an upregulation in the expression of GTF2H4 results in a corresponding decrease in Riskscore. As research has revealed, a decreased expression of GTF2H4 is associated with a decreased DNA repair capacity. Genetic variations in

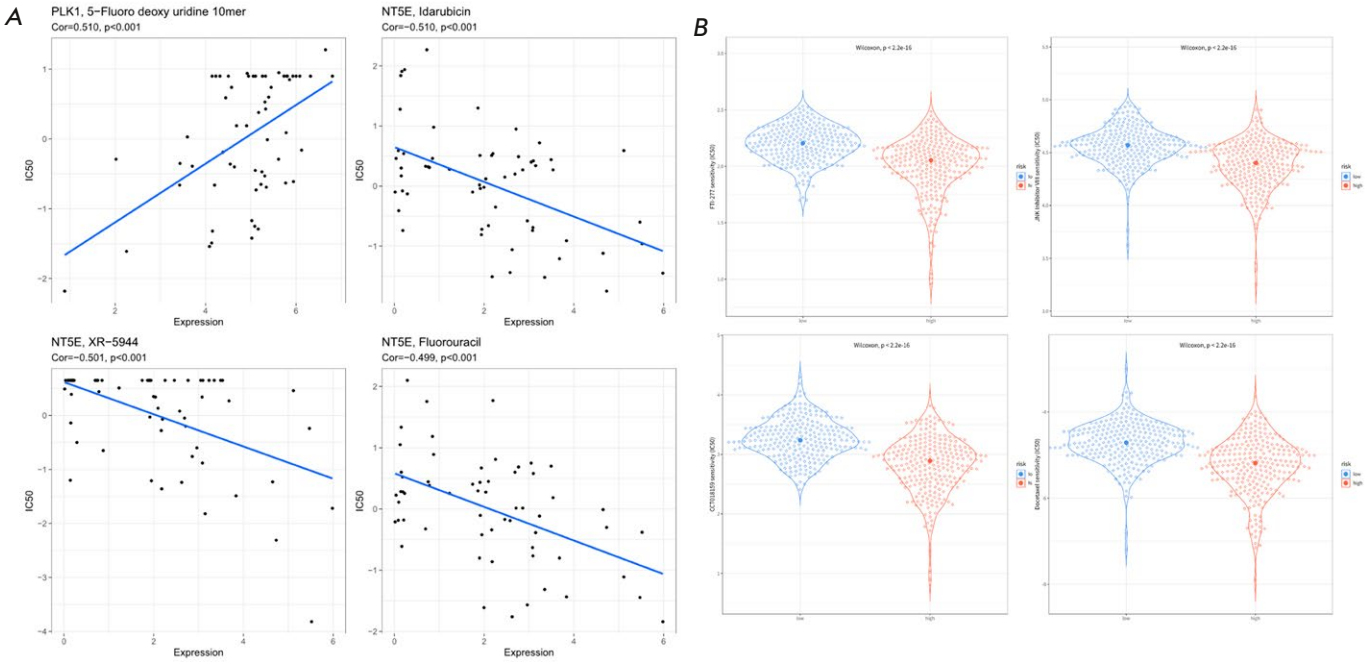


Fig. 7. Prediction of the response of LUAD patients to anticancer drug inhibitors. (*A*) Correlation between the expression levels of feature genes in LUAD and the IC_{50} values of patients to drug inhibitors. (*B*) Prediction of the treatment response of LUAD patients to FTI-277, JNK Inhibitor VIII, CCT018159, and Docetaxel in the HR and LR groups of LUAD patients

GTF2H4 raise the risk of LC, and GTF2H4 is a potential predictor of clinical outcomes of platinum-based chemotherapy in NSCLC patients [38, 39]. Although the effects of TEX15 and H2BC4 on LUAD are unknown, the effects of other DNA replication stress biomarkers on the risk of LUAD patient prognosis echo the findings of this study.

ICI therapy has greatly improved the dilemma of cancer treatment, but the probability of a response to ICI therapy in LUAD individuals remains comparatively low, while the majority of cancer patients may not derive substantial benefits from immunotherapy drugs [40]. Compared with HR LUAD patients, LR individuals have higher IPS and significantly lower TIDE scores, indicating that LR LUAD individuals display a greater likelihood of benefiting from immunotherapy. In addition, based on prognostic genes and prognostic risk grouping, it is helpful to highlight the efficacy of chemotherapy drugs widely used in the clinical treatment of LUAD. Idarubicin is an anthracycline chemotherapy drug commonly used to treat malignant tumors like LC and leukemia [41]. Our results showed that LUAD patients with high expression of NT5E were more sensitive to Idarubicin. Docetaxel belongs to the taxane class of chemotherapy drugs and is utilized to treat non-small cell lung cancer. They stabilize microtubules by preventing depolymerization and cause cell death [42]. Research has shown that LUAD individuals with a high Riskscore are more sensitive to Docetaxel. In addition, research found that the DNA-targeted drugs XR5944 [43], HSP90, and DDX39B inhibitor CCT018159 [44], farnesyl transferase inhibitor FTI-277 [45], and the JNK inhibitor VIII [46] with potential cancer therapeutic effects are related to the risk score of LUAD individuals. In summary, the LUAD prognostic risk score calculated using DNA replication stress biomarkers had the potential to predict the drug treatment response.

In conclusion, we have established a new DRSDs feature with the potential to forecast the immunotherapy response of LUAD individuals. Undeniably, limitations exist. Although the prognostic value of the DRSDs feature we established has been fully validated in the TCGA and GEO cohorts, the retrospective and potential biases of this study still need attention. Secondly, this study only conducted analyses based on public databases, and it is necessary to attempt more *in vitro* and *in vivo* experiments to study the molecular mechanisms of DNARSs affecting LUAD. In addition, external clinical studies are needed to determine the potential estimation accuracy of the DRSDs feature for the prognosis of LUAD individuals who have not received or have received immunotherapy. ●

Huang S. and Shi S. conceived and designed the study. Wen G. collected the data. Lei C. and Liu X. analyzed and interpreted the data. Chang J. and Yin X. contributed to the manuscript writing and editing.

All authors read and approved the final manuscript.

The authors declare that they have no conflicts of interest with the contents of this article.

Tables are available on the website https://doi.org/10.32607/actanaturae.25112.

REFERENCES

1. Sung H., Ferlay J., Siegel R.L., Laversanne M., Soerjomataram I., Jemal A., Bray F. // CA Cancer J. Clin. 2021. V. 71. № 3. P. 209–249.
2. Chen J.W., Dhahbi J. // Sci. Rep. 2021. V. 11. № 1. P. 13323.
3. Hao C.C., Xu C.Y., Zhao X.Y., Luo J.N., Wang G., Zhao L.H., Ge X., Ge X.F. // J. Exp. Clin. Cancer Res. 2020. V. 39. № 1. P. 256.
4. Saxena S., Zou L. // Mol. Cell. 2022. V. 82. № 12. P. 2298–2314.
5. Maiorano D., El Etri J., Franchet C., Hoffmann J.S. // Int. J. Mol. Sci. 2021. V. 22. № 8. P. 3924.
6. Bianco J.N., Bergoglio V., Lin Y.L., Pillaire M.J., Schmitz A.L., Gilhodes J., Lusque A., Mazieres J., Lacroix-Triki M., Roumeliotis T.I., et al. // Nat. Commun. 2019. V. 10. № 1. P. 910.
7. Ubhi T., Brown G.W. // Cancer Res. 2019. V. 79. № 8. P. 1730–1739.
8. Allera-Moreau C., Rouquette I., Lepage B., Oumouhou N., Walschaerts M., Leconte E., Schilling V., Gordien K., Brouchet L., Delisle M.B., et al. // Oncogenesis. 2012. V. 1. № 10. P. e30.
9. Baillie K.E., Stirling P.C. // Trends Cancer. 2021. V. 7. № 5. P. 430–446.
10. Zuo S., Wei M., Wang S., Dong J., Wei J. // Front Immunol. 2020. V. 11. P. 1218.
11. Sun S., Guo W., Wang Z., Wang X., Zhang G., Zhang H., Li R., Gao Y., Qiu B., Tan F., et al. // Cancer Med. 2020. V. 9. № 16. P. 5960–5975.
12. Hu M., Li Y., Lu Y., Wang M., Li Y., Wang C., Li Q., Zhao H.// PeerJ. 2021. V. 9. P. e11306.
13. Huang R.H., Hong Y.K., Du H., Ke W.Q., Lin B.B., Li Y.L. // J. Transl. Med. 2023. V. 21. № 1. P. 20.
14 Dreyer S.B., Upstill-Goddard R., Paulus-Hock V., Paris C., Lampraki E.M., Dray E., Serrels B., Caligiuri G., Rebus S., Plenker D., et al. // Gastroenterology. 2021. V. 160. № 1. P. 362–377.
15. Bala A.V., Galsky M.D., Rosenberg J.E., Powles T., Petrylak D.P., Bellmunt J., Loriot Y., Necchi A., Hoffman-Censits J., Perez-Gracia J.L., et al. // Lancet. 2017. V. 389. № 10064. P. 67–76.
16. Hugo W., Zaretsky J.M., Sun L., Song C., Moreno B.H., Hu-Lieskovan S., Berent-Maoz B., Pang J., Chmielowski B., Cherry G., et al. // Cell. 2016. V. 165. № 1. P. 35–44.
17. Robinson M.D., McCarthy D.J., Smyth G.K. // Bioinformatics. 2010. V. 26. № 1. P. 139–140.
18. Friedman J., Hastie T., Tibshirani R. // J. Stat. Softw. 2010. V. 33. № 1. P. 1–22.
19. Blanche P., Dartigues J.F., Jacqmin-Gadda H. // Stat. Med. 2013. V. 32. № 30. P. 5381–5397.
20. Huang C., Liu Z., Xiao L., Xia Y., Huang J., Luo H., Zong Z., Zhu Z. // Front. Oncol. 2019. V. 9. P. 1159.
21. Hanzelmann S., Castelo R., Guinney J. // Bioinformatics. 2013. V. 14. P. 7.
22. Qu X., Zhao X., Lin K., Wang N., Li X., Li S., Zhang L., Shi Y. // Front. Immunol. 2022. V. 13. P. 994019.
23. Lv W., Tan Y., Zhou X., Zhang Q., Zhang J., Wu Y. // Front. Immunol. 2022. V. 13. P. 989928.
24. Sun Y., Cheng Z., Liu S. // Mol. Med. 2022. V. 28. № 1. P. 128.
25. Ragu S., Matos-Rodrigues G., Lopez B.S. // Genes (Basel). 2020. V. 11. № 4. P. 409.
26. Wang Q., Wu G., Fu L., Li Z., Wu Y., Zhu T., Yu G. // Mutation Research. 2023. V. 826. P. 111811.
27. Ning G., Lu C., Chen Y., Jiang M., Si P., Zhang R. //Anticancer Drugs. 2023. V. 34. № 7. P. 866–876.
28. Wang S., Cai W., Li J., An W., Zheng H., Liao M. // Biochem. Genet. 2022. V. 60. № 6. P. 1934–1945.
29. He J., Wang Z., Wang Y., Zou T., Li X.P., Chen J. // Dis. Markers. V. 2022. https://doi.org/10.1155/2022/3306189.
30. Hong M.J., Park J.E., Lee S.Y., Lee J., Choi J.E., Kang H.G., Do S.K., Jeong J.Y., Shin K.M., Lee W.K., et al. // J. Cancer. 2022. V. 13. № 15. 3701–3709.
31. Panneerselvam J., Srivastava A., Mehta M., Chen A., Zhao Y.D., Munshi A., Ramesh R. // Cancers (Basel). 2019. V. 11. № 12. P. 1879.
32. Sakurai E., Ishizawa H., Kiriyama Y., Michiba A., Hoshikawa Y., Tsukamoto T. // Int. J. Mol. Sci. 2022. V. 23. № 12. P. 6679.
33. Wang W., Dong M., Cui J., Xu F., Yan C., Ma C., Yi L., Tang W., Dong J., Wei Y. // Mol. Med. Rep. 2019. V. 20. № 2. P. 1629–1636.
34. Wu Y., Jamal M., Xie T., Sun J., Song T., Yin Q., Li J., Pan S., Zeng X., Xie S., et al. // Cancer Sci. 2019. V. 110. № 9. P. 2734–2747.
35. Fu Y., Wei X.D., Guo L., Wu K., Le J., Ma Y., Kong X., Tong Y., Wu H. // Front. Oncol. 2022. V. 12. P. 904887.
36. Kowash R.R., Akbay E.A. // Front. Immunol. 2023. V. 14. P. 1130358.
37. Zhang H., Cao Y., Tang J., Wang R. // Biomed. Res. Int. 2022. V. 2022. P. 9944847.
38. Song X., Wang S., Hong X., Li X., Zhao X., Huai C., Chen H., Gao Z., Qian J., Wang J., et al. // Sci. Rep. 2017. V. 7. № 1. P. 11785.
39. Wang M., Liu H., Liu Z., Yi X., Bickeboller H., Hung R.J., Brennan P., Landi M.T., Caporaso N., Christiani D.C., et al. // Carcinogenesis. 2016. V. 37. № 9. P. 888–896.
40. Attili I., Tarantino P., Passaro A., Stati V., Curigliano G., de Marinis F. // Lung Cancer. 2021. V. 154. P. 151–160.
41. Capeloa T., Benyahia Z., Zampieri L.X., Blackman M., Sonveaux P. // Semin. Cell Dev. Biol. 2020. V. 98. P. 181–191.
42. Borghaei H., Gettinger S., Vokes E.E., Chow L.Q.M., Burgio M.A., de Castro Carpeno J., Pluzanski A., Arrieta O., Frontera O.A., Chiari R., et al. // J. Clin. Oncol. 2021. V. 39. № 7. P. 723–733.
43. Buric A.J., Dickerhoff J., Yang D. // Molecules. 2021. V. 26. № 14. P. 4132.
44. Tu Q., Liu X., Yao X., Li R., Liu G., Jiang H., Li,K., Chen Q., Huang X., Chang Q., et al. // J. Exp. Clin. Cancer Res. 2022. V. 41. № 1. P. 274.
45. Tateishi K., Tsubaki M., Takeda T., Yamatomo Y., Imano M., Satou T., Nishida S. // J. BUON. 2021. V. 26. № 2. P. 606–612.
46. Heslop K.A., Rovini A., Hunt E.G., Fang D., Morris M.E., Christie C.F., Gooz M.B., DeHart D.N., Dang Y., Lemasters J.J., et al. // Biochem. Pharmacol. 2020. V. 171. P. 113728.

GENERAL RULES

Acta Naturae publishes experimental articles and reviews, as well as articles on topical issues, short reviews, and reports on the subjects of basic and applied life sciences and biotechnology.

The journal *Acta Naturae* is on the list of the leading periodicals of the Higher Attestation Commission of the Russian Ministry of Education and Science. The journal Acta Naturae is indexed in PubMed, Web of Science, Scopus and RCSI databases.

The editors of *Acta Naturae* ask of the authors that they follow certain guidelines listed below. Articles which fail to conform to these guidelines will be rejected without review. The editors will not consider articles whose results have already been published or are being considered by other publications.

The maximum length of a review, together with tables and references, cannot exceed 50,000 characters with spaces (approximately 30 pages, A4 format, 1.5 spacing, Times New Roman font, size 12) and cannot contain more than 16 figures.

Experimental articles should not exceed 30,000 symbols (approximately 15 pages in A4 format, including tables and references). They should contain no more than ten figures.

A short report must include the study's rationale, experimental material, and conclusions. A short report should not exceed 12,000 symbols (5–6 pages in A4 format including no more than 12 references). It should contain no more than three figures.

The manuscript and all necessary files should be uploaded to www.actanaturae.ru:

1) text in Word 2003 for Windows format;
2) the figures in TIFF format;
3) the text of the article and figures in one pdf file;
4) the article's title, the names and initials of the authors, the full name of the organizations, the abstract, keywords, abbreviations, figure captions, and Russian references should be translated to English;
5) the cover letter stating that the submitted manuscript has not been published elsewhere and is not under consideration for publication;
6) the license agreement (the agreement form can be downloaded from the website www.actanaturae.ru).

MANUSCRIPT FORMATTING

The manuscript should be formatted in the following manner:

- Article title. Bold font. The title should not be too long or too short and must be informative. The title should not exceed 100 characters. It should reflect the major result, the essence, and uniqueness of the work, names and initials of the authors.
- The corresponding author, who will also be working with the proofs, should be marked with a footnote *.
- Full name of the scientific organization and its departmental affiliation. If there are two or more scientific organizations involved, they should be linked by digital superscripts with the authors' names. Abstract. The structure of the abstract should be very clear and must reflect the following: it should introduce the reader to the main issue and describe the experimental approach, the possibility of practical use, and the possibility of further research in the field. The average length of an abstract is 20 lines (1,500 characters).
- Keywords (3 – 6). These should include the field of research, methods, experimental subject, and the specifics of the work. List of abbreviations.
- INTRODUCTION
- EXPERIMENTAL PROCEDURES
- RESULTS AND DISCUSSION
- CONCLUSION

  The organizations that funded the work should be listed at the end of this section with grant numbers in parenthesis.
- REFERENCES

  The in-text references should be in brackets, such as [1].

RECOMMENDATIONS ON THE TYPING AND FORMATTING OF THE TEXT

- We recommend the use of Microsoft Word 2003 for Windows text editing software.
- The Times New Roman font should be used. Standard font size is 12.
- The space between the lines is 1.5.
- Using more than one whole space between words is not recommended.
- We do not accept articles with automatic referencing; automatic word hyphenation; or automatic prohibition of hyphenation, listing, automatic indentation, etc.
- We recommend that tables be created using Word software options (Table → Insert Table) or MS Excel. Tables that were created manually (using lots of spaces without boxes) cannot be accepted.
- Initials and last names should always be separated by a whole space; for example, A. A. Ivanov.
- Throughout the text, all dates should appear in the "day.month.year" format, for example 02.05.1991, 26.12.1874, etc.
- There should be no periods after the title of the article, the authors' names, headings and subheadings, figure captions, units (s – second, g – gram, min – minute, h – hour, d – day, deg – degree).
- Periods should be used after footnotes (including those in tables), table comments, abstracts, and abbreviations (mon. – months, y. – years, m. temp. – melting temperature); however, they should not be used in subscripted indexes (T_m – melting temperature; $T_{p.t}$ – temperature of phase transition). One exception is mln – million, which should be used without a period.
- Decimal numbers should always contain a period and not a comma (0.25 and not 0,25).
- The hyphen ("-") is surrounded by two whole spaces, while the "minus," "interval," or "chemical bond" symbols do not require a space.
- The only symbol used for multiplication is "×"; the "×" symbol can only be used if it has a number to its

right. The "·" symbol is used for denoting complex compounds in chemical formulas and also noncovalent complexes (such as DNA·RNA, etc.).
- Formulas must use the letter of the Latin and Greek alphabets.
- Latin genera and species' names should be in italics, while the taxa of higher orders should be in regular font.
- Gene names (except for yeast genes) should be italicized, while names of proteins should be in regular font.
- Names of nucleotides (A, T, G, C, U), amino acids (Arg, Ile, Val, etc.), and phosphonucleotides (ATP, AMP, etc.) should be written with Latin letters in regular font.
- Numeration of bases in nucleic acids and amino acid residues should not be hyphenated (T34, Ala89).
- When choosing units of measurement, SI units are to be used.
- Molecular mass should be in Daltons (Da, KDa, MDa).
- The number of nucleotide pairs should be abbreviated (bp, kbp).
- The number of amino acids should be abbreviated to aa.
- Biochemical terms, such as the names of enzymes, should conform to IUPAC standards.
- The number of term and name abbreviations in the text should be kept to a minimum.
- Repeating the same data in the text, tables, and graphs is not allowed.

GUIDENESS FOR ILLUSTRATIONS

- Figures should be supplied in separate files. Only TIFF is accepted.
- Figures should have a resolution of no less than 300 dpi for color and half-tone images and no less than 600 dpi.
- Files should not have any additional layers.

REVIEW AND PREPARATION OF THE MANUSCRIPT FOR PRINT AND PUBLICATION

Articles are published on a first-come, first-served basis. The members of the editorial board have the right to recommend the expedited publishing of articles which are deemed to be a priority and have received good reviews.

Articles which have been received by the editorial board are assessed by the board members and then sent for external review, if needed. The choice of reviewers is up to the editorial board. The manuscript is sent on to reviewers who are experts in this field of research, and the editorial board makes its decisions based on the reviews of these experts. The article may be accepted as is, sent back for improvements, or rejected.

The editorial board can decide to reject an article if it does not conform to the guidelines set above.

The return of an article to the authors for improvement does not mean that the article has been accepted for publication. After the revised text has been received, a decision is made by the editorial board. The author must return the improved text, together with the responses to all comments. The date of acceptance is the day on which the final version of the article was received by the publisher.

A revised manuscript must be sent back to the publisher a week after the authors have received the comments; if not, the article is considered a resubmission.

E-mail is used at all the stages of communication between the author, editors, publishers, and reviewers, so it is of vital importance that the authors monitor the address that they list in the article and inform the publisher of any changes in due time.

After the layout for the relevant issue of the journal is ready, the publisher sends out PDF files to the authors for a final review.

Changes other than simple corrections in the text, figures, or tables are not allowed at the final review stage. If this is necessary, the issue is resolved by the editorial board.

FORMAT OF REFERENCES

The journal uses a numeric reference system, which means that references are denoted as numbers in the text (in brackets) which refer to the number in the reference list.

For books: the last name and initials of the author, full title of the book, location of publisher, publisher, year in which the work was published, and the volume or issue and the number of pages in the book.

For periodicals: the last name and initials of the author, title of the journal, year in which the work was published, volume, issue, first and last page of the article. Must specify the name of the first 10 authors. Ross M.T., Grafham D.V., Coffey A.J., Scherer S., McLay K., Muzny D., Platzer M., Howell G.R., Burrows C., Bird C.P., et al. // Nature. 2005. V. 434. № 7031. P. 325–337.

References to books which have Russian translations should be accompanied with references to the original material listing the required data.

References to doctoral thesis abstracts must include the last name and initials of the author, the title of the thesis, the location in which the work was performed, and the year of completion.

References to patents must include the last names and initials of the authors, the type of the patent document (the author's rights or patent), the patent number, the name of the country that issued the document, the international invention classification index, and the year of patent issue.

The list of references should be on a separate page. The tables should be on a separate page, and figure captions should also be on a separate page.

The following e-mail addresses can be used to contact the editorial staff: actanaturae@gmail.com, tel.: (495) 727-38-60.