2022 • 26 • 3

Print ISSN 2500-0462
Online ISSN 2500-3259

Научный рецензируемый журнал

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

Основан в 1997 г.
Периодичность 8 выпусков в год
DOI 10.18699/VJGB-22-27

Учредители

Сибирское отделение Российской академии наук

Федеральное государственное бюджетное научное учреждение «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук»

Межрегиональная общественная организация Вавиловское общество генетиков и селекционеров

2022 • 26 • 3

Print ISSN 2500-0462
Online ISSN 2500-3259

Scientific Peer Reviewed Journal

VAVILOV JOURNAL OF GENETICS AND BREEDING

VAVILOVSKII ZHURNAL GENETIKI I SELEKTSII

Founded in 1997
Published 8 times annually
DOI 10.18699/VJGB-22-27

Founders

Siberian Branch of the Russian Academy of Sciences

Federal Research Center Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences

The Vavilov Society of Geneticists and Breeders

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

СОДЕРЖАНИЕ • 2022 • 26 • 3

Молекулярная и клеточная биология

Генетика и селекция растений

Генетика и селекция животных

Медицинская генетика

Тест-системы и вакцинопрофилактика

VAVILOV JOURNAL OF GENETICS AND BREEDING

CONTENTS • 2022 • 26 • 3

Вавиловский журнал генетики и селекции. 2022;26(3):227-233
DOI 10.18699/VJGB-22-29

Original Russian text www.bionet.nsc.ru/vogis/

An experimental study of the effects of SNPs in the TATA boxes of the *GRIN1*, *ASCL3* and *NOS1* genes on interactions with the TATA-binding protein

E.B. Sharypova, I.A. Drachkova, I.V. Chadaeva, M.P. Ponomarenko, L.K. Savinkova ✉

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
✉ savinkl@mail.ru

Abstract. The *GRIN1*, *ASCL3*, and *NOS1* genes are associated with various phenotypes of neuropsychiatric disorders. For instance, these genes contribute to the development of schizophrenia, Alzheimer's and Parkinson's diseases, and epilepsy. These genes are also associated with various cancers. For example, *ASCL3* is overexpressed in breast cancer, and *NOS1*, in ovarian cancer cell lines. Based on our findings and literature data, we had previously obtained results suggesting that the single-nucleotide polymorphisms (SNPs) that disrupt erythropoiesis are highly likely to be associated with cognitive and neuropsychiatric disorders in humans. In the present work, using SNP_TATA_Z-tester, we investigated the influence of unannotated SNPs in the TATA boxes of the promoters of the *GRIN1*, *ASCL3*, and *NOS1* genes (which are involved in neuropsychiatric disorders and cancers) on the interaction of the TATA boxes with the TATA-binding protein (TBP). Double-stranded oligodeoxyribonucleotides identical to the TATA-containing promoter regions of the *GRIN1*, *ASCL3*, and *NOS1* genes (reference and minor alleles) and recombinant human TBP were employed to study *in vitro* (by an electrophoretic mobility shift assay) kinetic characteristics of the formation of TBP–TATA complexes and their affinity. It was found, for example, that allele A of rs1402667001 in the *GRIN1* promoter increases TBP–TATA affinity 1.4-fold, whereas allele C in the TATA box of the *ASCL3* promoter decreases the affinity 1.4-fold. The lifetime of the complexes in both cases decreased by ~20 % due to changes in the rates of association and dissociation of the complexes (k_a and k_d, respectively). Our experimental results are consistent with the literature showing *GRIN1* underexpression in schizophrenic disorders as well as an increased risk of cervical, bladder, and kidney cancers and lymphoma during *ASCL3* underexpression. The effect of allele A of the –27G>A SNP (rs1195040887) in the *NOS1* promoter is suggestive of an increased risk of ischemic damage to the brain in carriers. A comparison of experimental TBP–TATA affinity values (K_D) of wild-type and minor alleles with predicted ones showed that the data correlate well (linear correlation coefficient $r = 0.94$, $p < 0.01$).
Key words: *GRIN1*; *ASCL3*; *NOS1*; TATA-binding protein; affinity; TBP/TATA interaction.

For citation: Sharypova E.B., Drachkova I.A., Chadaeva I.V., Ponomarenko M.P., Savinkova L.K. An experimental study of the effects of SNPs in the TATA boxes of the *GRIN1*, *ASCL3* and *NOS1* genes on interactions with the TATA-binding protein. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2022;26(3):227-233. DOI 10.18699/VJGB-22-29

Экспериментальное изучение влияния SNP TATA-боксов генов *GRIN1*, *ASCL3* и *NOS1* на взаимодействие с TATA-связывающим белком

Е.Б. Шарыпова, И.А. Драчкова, И.В. Чадаева, М.П. Пономаренко, Л.К. Савинкова ✉

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ savinkl@mail.ru

Аннотация. Гены *GRIN1*, *ASCL3* и *NOS1* связаны с различными фенотипами нервно-психических расстройств. Эти гены делают вклад в развитие шизофрении, болезней Альцгеймера и Паркинсона, эпилепсии и др. и ассоциируются также с различными онкологическими заболеваниями. Например, повышенная экспрессия *ASCL3* наблюдается при раке молочной железы, *NOS1* – в клеточных линиях рака яичников. Ранее на основе наших и литературных данных мы получили результаты, свидетельствующие в пользу того, что SNP, нарушающие эритропоэз, с большой вероятностью могут быть связаны с когнитивными и нервно-психическими расстройствами у человека. В настоящей работе исследовано влияние выявленных с помощью SNP_TATA_Z-tester не-

аннотированных SNP TATA-боксов промоторов генов *GRIN1*, *ASCL3* и *NOS1*, участвующих в нервно-психических расстройствах и онкологических заболеваниях, на взаимодействие TATA-связывающего белка (TBP). Для изучения *in vitro* кинетических характеристик образования комплексов TBP/TATA и аффинности с помощью метода задержки ДНК в геле использованы двуцепочечные олигодезоксирибонуклеотиды, идентичные TATA-содержащим участкам промоторов генов *GRIN1*, *ASCL3* и *NOS1* (референсным и минорным аллелям), и рекомбинантный TBP человека. Показано, например, что аллель «A» rs1402667001 промотора гена *GRIN1* повышает аффинность TBP/TATA в 1.4 раза, а аллель «C» TATA-бокса промотора гена *ASCL3* снижает аффинность в 1.4 раза; при этом время жизни комплексов в обоих случаях уменьшается примерно на 20 % за счет изменения скоростей образования и диссоциации комплексов (k_a и k_d соответственно). Наши экспериментальные результаты согласуются с литературными данными, показывающими низкую экспрессию гена *GRIN1* при шизофренических расстройствах и повышенный риск возникновения рака шейки матки, мочевого пузыря, почек и лимфомы при пониженной экспрессии гена *ASCL3*. Влияние аллеля «A» SNP –27G>A (rs1195040887) промотора гена *NOS1* гипотетически может свидетельствовать о повышенном риске возникновения ишемического повреждения мозга у носителей. Сравнение экспериментальных значений аффинности (K_D) TBP/TATA «диких» (WT) и минорных аллелей с прогнозируемыми показало, что данные хорошо коррелируют друг с другом: коэффициент линейной корреляции $r = 0.94$ ($p < 0.01$).
Ключевые слова: *GRIN1*; *ASCL3*; *NOS1*; TATA-связывающий белок; аффинность; TBP/TATA взаимодействие.

Introduction

Previously, using Web service SNP_TATA_Comparator (Ponomarenko M. et al., 2015) and *in vitro* experiments, we studied effects of SNPs in TATA boxes within core promoters of human genes to predict potential SNP markers, for example, markers of obesity (Arkova et al., 2015), autoimmune diseases (Ponomarenko M. et al., 2016), Alzheimer's disease (Ponomarenko P. et al., 2017), aggressiveness (Chadaeva et al., 2016), circadian rhythm disorders (Ponomarenko P. et al., 2016), anomalies of female reproductive potential (Chadaeva et al., 2018), erythropoiesis disorders (Sharypova et al., 2018), and resistance to anticancer therapy (Turnaev et al., 2016). Then, on the basis of our results and literature data, we made findings suggesting that SNPs that disrupt erythropoiesis are likely to be associated with cognitive and neuropsychiatric disorders in humans (Ponomarenko M. et al., 2020).

The aim of the current work was to search for and to experimentally verify *in vitro* the effects of unannotated SNPs in TATA boxes within promoters of genes *NOS1*, *GRIN1*, and *ASCL3* (which are involved in neuropsychiatric disorders and cancers) on the affinity and kinetic characteristics of TBP–TATA complexes. Identification of causal regulatory mechanisms of diseases is becoming a common practice, but experimental annotation of variants in target genes, especially in regulatory regions, is still a major bottleneck for the use of such genetic data in personalized medicine. Therefore, experimental quantitative methods for annotating SNPs in regulatory regions of specific genes remain important and relevant. This paper presents predictions of effects of unannotated SNPs in TATA boxes of genes *GRIN1*, *ASCL3*, and *NOS1* on the thermodynamic and kinetic characteristics of TBP–TATA complexes as well as the results of their experimental verification *in vitro*.

The *GRIN1* gene, located in chromosomal region 9q34.3, codes for the GluN1 (NR1) subunit of the *N*-methyl-D-aspartate receptor (NMDAR) and plays a key role in synaptic functions (Sin et al., 2002). The protein encoded by *GRIN1* is a critical subunit of *N*-methyl-D-aspartate receptors, which are members of the superfamily of glutamate receptor channels. The latter are heteromeric protein complexes with multiple subunits arranged to form a ligand-gated ion channel. These subunits play a key part in synaptic plasticity, which is thought to underlie memory and learning.

The first meta-analysis and convergence analysis (Forero, 2020) of available genome-wide expression studies regarding epileptogenesis in humans and model animals made it possible to identify several major candidate genes, including *GRIN1*. Animal models and postmortem studies on patients' brains have shown that transcription and expression levels of the gene of the GluN1 protein in schizophrenia differ from those in controls (conditionally healthy volunteers), although the changes varied among different regions of the brain (Ding et al., 2017).

The achaete-scute complex-like (*ASCL*) gene family consists of five members, namely *ASCL1*, *ASCL2*, *ASCL3*, *ASCL4*, and *ASCL5*. The *ASCL3* gene (*SGN1*) is located on chromosome 11, and its product was originally characterized as a transcription factor specifically localized to cells of salivary gland ducts (Park et al., 2017). Dysregulation of *ASCL* family genes has been reported to play a key role in psychiatric and neurological disorders (Hanahan, Weinberg, 2011). All *ASCL* genes encode the basic helix-loop-helix transcription factors that control nervous system development (Rugel-Stahl et al., 2012); thus, they are called proneural genes. Expression of *ASCL* family genes and the impact of their products on cells are not limited to the nervous system. For example, ASCL family members have been shown to be expressed in progenitor cells during muscle and intestinal-cell differentiation (Fox, 1998).

Using bioinformatic analyses, researchers have determined potential involvement of several ASCL family members in the initiation and progression of tumors in various types of cancer. *ASCL3* is overexpressed in breast cancer (Hanahan, Weinberg, 2011) but is underexpressed (relative to normal controls) in kidney, cervical, and bladder cancers as well as lymphoma and melanoma. Analysis of different subtypes of kidney tumors has revealed that *ASCL3* is downregulated in renal oncocytoma. As has already been mentioned, in lymphoma and cancers of the cervix, bladder, kidney, and epithelium, a decrease in the expression of *ASCL3* has been documented (Hanahan, Weinberg, 2011), suggesting that this gene is a suitable research object not only in psychiatric and neurological disorders but also in cancers.

The *NOS1* gene codes for the major isoform of nitric oxide synthase and is widely expressed in all tissues, and NOS1 produces approximately 90 % of nitric oxide in the central nervous system (Akyol et al., 2004). The gene is mapped to chromosomal region 12q24. Several studies indicate that *NOS1* variants are associated with such disorders as Alzheimer's disease (Mishizen-Eberz et al., 2004), schizophrenia (Shinkai et al., 2002; Saadat, 2010), and Parkinson's disease (Hancock et al., 2008; Yu et al., 2018). Using reverse-transcription polymerase chain reaction (RT-PCR), some authors (Freudenberg et al., 2015) demonstrated that protein expression of NOS1 is constitutively high in ovarian cancer cell lines and that mRNA expression of *NOS1* varies among such cell lines. The results of that study mean that NOS1 promotes malignant characteristics of ovarian cancer cells, including proliferation, invasion, and chemoresistance, thereby constituting a potential therapeutic target.

Materials and methods

DNA sequences. Unannotated SNPs (from the *GRIN1* gene: rs1402667001, from *ASCL3*: rs1049743008:c, and from *NOS1*: rs1195040887) were retrieved from the dbSNP database (Sherry et al., 2001). Promoter sequences within the [–100; –1] region relative to a transcription start site were retrieved from the Eukaryotic Promotor Database (EPD) (Praz et al., 2002), and the presence of TATA boxes in these regions was determined in the same database.

Analysis of DNA sequences *in silico.* DNA sequences of human genes *GRIN1*, *ASCL3*, and *NOS1* between nucleotide positions –100 and –1 upstream of the transcription start site, which were taken from the reference genome, were analyzed using our Web service SNP_TATA_Z-tester, which is a modified version of SNP_TATA_Comparator (Ponomarenko M. et al., 2015).

Synthetic double-stranded oligodeoxyribonucleotides (ODNs). For experimental verification, we used 26 bp ODNs identical to the reference and minor alleles of genes *GRIN1*, *ASCL3*, and *NOS1*; the ODNs were synthesized and then purified by polyacrylamide gel electrophoresis by Biosan (Novosibirsk, Russia).

Sequences of these double-stranded ODNs – identical to the promoter regions of genes *GRIN1*, *ASCL3*, and *NOS1* containing TATA-like elements – were as follows (reference [wild type; WT] alleles and minor alleles):

GRIN1 (WT) – 5′-tggagg***g***ggACAAAGACAgggtggtg-3′
GRIN1 (–34g>a) – 5′-tggagg***a***ggACAAAGACAgggtggtg-3′
ASCL3 (WT) – 5′-tcgaaaaaTAAAA**T**AAAAtaaaacat-3′
ASCL3 (–45T>c) – 5′-tcgaaaaaTAAAA***c***AAAAtaaaacat-3′
NOS1 (WT) – 5′-tgtttcctGATA**G**AAAaaaaaaatgg-3′
NOS1 (–27G>a) – 5′-tgtttcctGATA***a***AAAaaaaaaatgg-3′

Labeling of the ODNs at 5′ ends with ^{32}P-γATP. To prepare labeled double-stranded ODNs, both their strands were labeled with ^{32}P-γATP (Biosan, Novosibirsk, Russia) by means of T4 polynucleotide kinase (SibEnzyme, Novosibirsk, Russia), annealed at 95 °C (at an equimolar ratio), and slowly cooled to room temperature. The annealed duplexes were purified and analyzed by electrophoresis in a nondenaturing 15 % polyacrylamide gel followed by autoradiography on a phosphorimager, Molecular Imager PharosFX Plus (Bio-Rad, Hercules, CA, USA). Unlabeled duplexes were prepared in the same way and used without further purification by polyacrylamide gel electrophoresis.

Isolation and purification of recombinant TATA-binding (TBP) protein. We used recombinant human TBP expressed in *Escherichia coli* BL21(DE3) cells via the pAR3038-hTBP plasmid (a gift from Prof. B. Puhg, Center for Gene Regulation, Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, Pennsylvania, USA). Expression and purification of TBP were performed according to ref. (Pugh, 1995), except for the concentration of isopropyl β-D-1-thiogalactopyranoside (1.0 instead of 0.1 mM) and induction time (3 instead of 1.5 h).

Determination of kinetic parameters (rates of formation and dissociation of complexes, k_a and k_d) and equilibrium dissociation constants, K_D, of TBP–ODN complexes. Association rate constants (k_a) and dissociation rate constants (k_d), which characterize the rates of formation and disintegration of complexes, were determined by measuring the kinetics of TBP binding to the TATA-containing double-stranded ODNs identical to either a WT TATA box (reference allele) or a TATA box carrying an SNP. The experiments were conducted at several concentrations of the ^{32}P-labeled ODNs and a fixed concentration of TBP (0.4 nM unless stated otherwise). TBP–ODN binding experiments were performed at 25 °C in a buffer composed of 20 mM HEPES-KOH (pH 7.6), 5 mM $MgCl_2$, 70 mM KCl, 1 mM dithiothreitol, 100 μg/ml bovine serum albumin, 0.01 % of NP-40, and 5 % of XYZ, as described in detail before (Drachkova et al., 2014). The electrophoretic mobility shift assay was carried out in a native 5 % polyacrylamide gel in Tris-glycine buffer (pH 8.3) for 40 min at 10 °C. The gels were dried, and Imaging Screen-K (Kodak) for Molecular Imager PharosFX Plus (Bio-Rad) was exposed to the gels. Each screen was scanned on the phosphorimager, and the autoradiographs were quantified in Quantity One v.4.5.0 software (Bio-Rad).

Statistical analysis. The comparison of the predicted and experimental values of TBP–TATA complexes' affinity for the "normal" and minor alleles was performed using the Statistica software package (Statsoft™, Tulsa, OK, USA).

Results and discussion

Transcription, as a rule, is the gene expression stage most sensitive to internal signals and external signals entering the cell and is the main mechanism that controls gene expression. Here, we analyzed its initial stage, i. e., the interaction of TBP with a promoter: the process that triggers the assembly of the transcription complex.

The Table presents the results of *in vitro* verification of our predictions (made by the SNP_TATA_Z-tester Web service) regarding the effect of substitutions in the TATA boxes of genes *GRIN1*, *ASCL3*, and *NOS1* on TBP–TATA affinity. These data include experimental affinity (K_D) values for WT and minor alleles and association and dissociation constants (k_a and k_d, respectively), reflecting the rates of formation and disintegration of TBP–TATA complexes.

The results given in the Table were obtained by the electrophoretic mobility shift assay. Figure 1 shows electropherograms for the *GRIN1* gene as an example.

As readers can see in the Table, the TBP–TATA affinity for WT *GRIN1* can be described as low-specificity binding

Experimental *in vitro* verification of our predictions of the effect of rs1402667001, rs1049743008, and rs1195040887 (from genes *GRIN1*, *ASCL3*, and *NOS1*, respectively) on affinity and kinetic characteristics of TBP–TATA complexes

Gene	SNP	Prediction		Experiment					
		$-\ln(K_D)$	Δ, ln units	$-\ln(K_D)$	K_D, nM	Δ, ln units	k_a, $(M^{-1} \cdot s^{-1}) \cdot 10^3$	k_d, $s^{-1} \cdot 10^{-4}$	$t_{1/2}$, min
GRIN1	WT	17.59 ± 0.08	0.00	14.95 ± 0.05	320 ± 50	0.00	1.4 ± 0.3	4.5 ± 0.4	25.7 ± 3
	–34G>a rs1402667001	17.74 ± 0.08	0.15	15.42 ± 0.08	200 ± 50	2.32	2.8 ± 0.4	5.7 ± 0.8	20.3 ± 5
ASCL3	WT	19.04 ± 0.08	0.00	18.24 ± 0.09	12 ± 5	0.00	7.0 ± 1.0	8.0 ± 3	14.4 ± 6
	–45T>c rs1049743008	18.75 ± 0.08	0.29	17.90 ± 0.08	17 ± 3	0.85	6.2 ± 0.5	10.0 ± 1	11.5 ± 1
NOS1	WT	18.56 ± 0.08	0.00	18.77 ± 0.08	70 ± 20	0.00	2.8 ± 0.3	2.0 ± 0.4	58.0 ± 6
	–27G>a rs1195040887	18.96 ± 0.08	0.34	18.95 ± 0.08	59 ± 9	0.18	4.7 ± 0.4	2.8 ± 0.3	38.0 ± 5

Note. The data are presented as mean ± standard deviation. $K_D = k_d/k_a$; Δ, a TBP affinity difference (between ODNs containing and not containing an SNP) expressed in logarithmic units: $\Delta = -\ln[K_{D, TATA(Mut)}] - (-\ln [K_{D, TATA}])$; $t_{1/2} = \ln 2/k_d$.

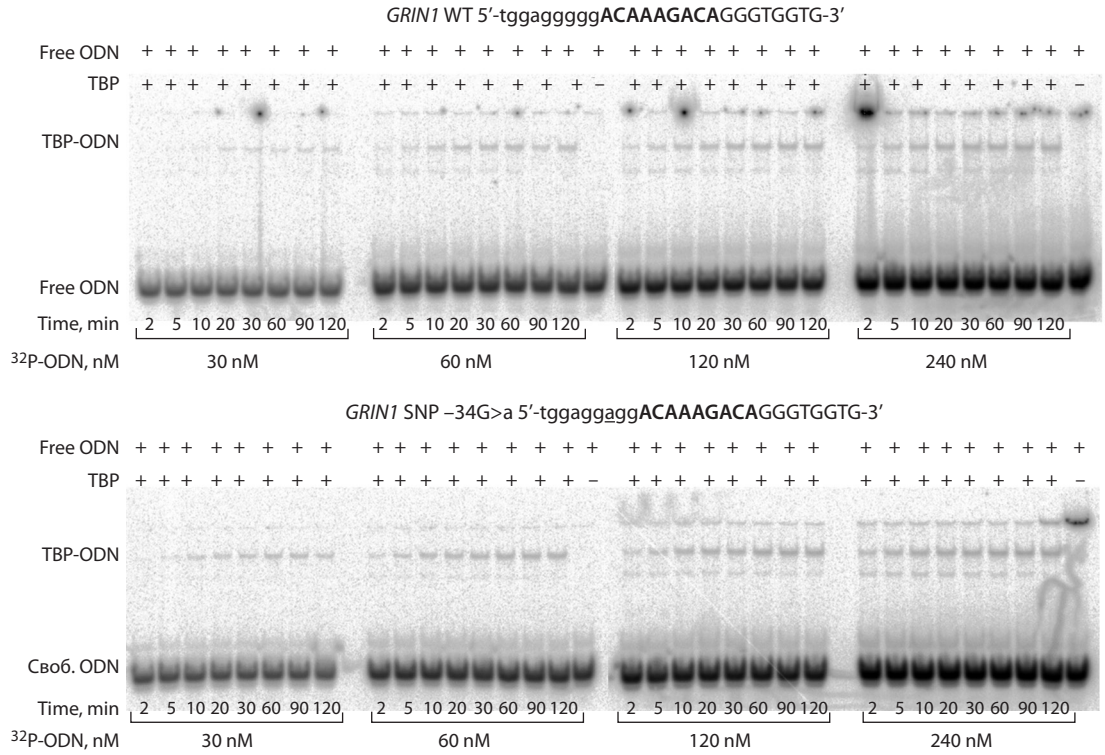


Fig. 1. The electropherograms obtained to determine kinetic isotherms of TBP binding to the ODNs identical to the TATA-like element of the *GRIN1* gene promoter: either the WT allele or minor allele "a" (SNP: –34G>a).

(K_D = 320 nM). Although the TATA box sequence is AT-rich (ACAAAGACA), the third T, which has the greatest weight in the TATA box, is missing. In addition, conformational flexibility of the three A*s*, which is clearly not enough to generate the conformation suitable for TBP binding, is blocked on both sides by high-melting-point base pairs containing C and G. The substitution (–34G>a) in the sequence flanking the TATA-like element did not significantly increase the flexibility of this DNA region, but the affinity strengthened almost 1.4-fold after introduction of the substitution (minor allele) as compared to the WT allele (K_D = 200 nM). The rate of formation of the TBP–ODN complex increased twofold: 2.8 versus 1.4 $M^{-1} \cdot s^{-1}$ (Fig. 2). The rate of complex dissociation was also slightly higher, by 20 %, in the case of the minor allele. Overall, these changes somewhat shortened the complex's half-life (from 25 to 20 min), that is, made it less stable. Judging from the sequence of the ODN in question, which is identical to the region of the *GRIN1* promoter, this promoter does not contain the TATA box consensus sequence but contains a G-rich box, which can bind to transcription

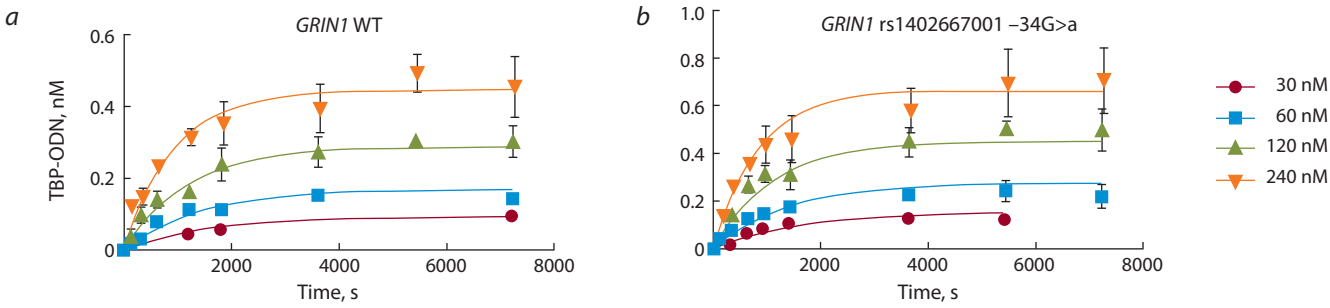


Fig. 2. Kinetic isotherms of the TBP binding to the ODNs identical to the TATA box of the *GRIN1* gene promoter, for either the WT allele (panel *a*) or minor allele "a" (panel *b*).

The binding isotherms and k_a and k_d values were determined by means of the electropherograms (see Fig. 1) in GraphPad Prism 5 software.

factor SP3, which activates transcription of genes in chicken embryonic cortical neurons and represses it in undifferentiated cells (Chaudhary et al., 2017), as demonstrated in a study on the chicken *GRIN1* gene.

The observed (in our study) weak affinity of TBP for the TATA-like element of the *GRIN1* promoter, which implies gene underexpression (Liu et al., 2019), is in agreement with findings of low *GRIN1* expression in animal models of schizophrenic disorders. On the basis of our data on enhanced TBP–TATA affinity for minor allele "a" (rs1402667001), we can theorize that carriers of this allele are at a reduced risk of schizophrenia as compared to carriers of WT allele G. It should be noted that despite active research, the role of *GRIN1* in the etiology of schizophrenia remains uncertain. For instance, in ref. (Zhao et al., 2006), using Sanger DNA sequencing, researchers conducted a case-control study to investigate the association between *GRIN1* and the risk of schizophrenia in a population of northern China. Distributions of both a genotype and allele of rs117783907 (–1945G/t) significantly differed between the case group and control group ($p < 0.0083$). In that article, genotype frequencies of rs138961287 and rs11146020 are statistically significantly different too ($p < 0.05$), indicating that rs138961287, rs117783907, and rs11146020 are associated with schizophrenia. In another association study conducted in a northern Chinese Han population, allele "c" of rs11146020 was reported to reduce the risk of schizophrenia (Begni et al., 2003), although Saadat (2010) found that this allele is a risk factor of schizophrenia in an Italian population. Furthermore, a meta-analysis (Zwicker et al., 2018) suggests that allele "c" of rs11146020 is associated with an increased risk of schizophrenia, i. e., the results are inconsistent among studies. It is likely that among different ethnic groups, the influence of environmental and genetic factors and their interactions may differ in their impact on the risk of mental disorders (Zwicker et al., 2018). The risk of psychosis increases with the accumulation of multiple variants carrying a genetic risk and with exposure to multiple adverse environmental factors (Gray et al., 2015).

Some authors (Ding et al., 2017) investigated the expression of a large group of genes in the brains of patients with major depressive disorder and controls (postmortem analysis). The results showed elevated expression of most of glutamatergic genes (e. g., *GRIN1*, *GRIN2A–D*, *GRIA2–4*, *GRIK1* and *-2*, and *GRM1*) tested in the dorsolateral prefrontal cortex (mainly in women). Based on these findings, it can be assumed that rs1402667001 (studied by us), which enhances the affinity of TBP for the TATA-like element, may be a candidate SNP marker of an increased risk of schizophrenia. Despite conflicting results, some researchers (Zou et al., 2020) believe that the association of *GRIN1* with schizophrenia and other psychotic disorders is undeniable and that subunit NR1 encoded by this gene may be a promising therapeutic target in schizophrenia.

As readers can see in the sequence of the *ASCL3* promoter region, which is identical to the region where a TATA box is usually located (positions –20 to –70 relative to the transcription start site), it is enriched in A nucleotides and contains a T. The latter has the greatest weight in the TATA box sequence (for the binding to TBP) and can take the third position in our case. Accordingly, we observed strong affinity of the TBP–TATA complex: K_D = 12 nM. SNP –45T>C (rs1049743008), replacing T with high-melting-point C, led to a 1.4-fold weakening of the affinity (K_D = 17 nM), although the rate of formation of the TBP–TATA complex increased slightly (12 %), while the rate of dissociation increased a little more: by 20 %. As a consequence, the complex's half-life with the minor allele is also slightly shorter (11.5 versus 14.4 min), i. e., stability decreased. Because changes in the affinity of the TBP–TATA interaction correlate with alterations of gene expression (Mogno et al., 2010), it can be hypothesized that carriers of the C allele (with weakened TBP–TATA affinity and *ASCL3* expression) are at a higher risk of a malignant tumor: lymphoma and cancers of the cervix, bladder, epithelium, and kidneys. This notion is confirmed by the results in ref. (Hanahan, Weinberg, 2011), where a database analysis revealed that out of 21 analyzed tumor types, five correlate with the *ASCL3* expression that is diminished to various degrees.

The *NOS1* promoter contains a TATA-like element with sequence GATA**G**AAA, to which TBP binds with 70 nM affinity. When high-melting-point G was replaced in our study by A, the affinity strengthened, albeit slightly: by 14 % (K_D = 59 ± 9 nM). In this context, the rate of TBP–TATA complex formation (k_a) increased by a factor of 1.7, while the dissociation rate of the complex (k_d) accelerated by a factor of 1.4, and the lifetime of the complex diminished ~1.4-fold.

Based on the results in ref. (Zou et al., 2020) indicating that NOS1 inhibitors can effectively reduce the severity of ischemic brain damage, it can be theorized that the A allele (SNP –27G>A, rs1195040887) – with enhanced TBP–TATA

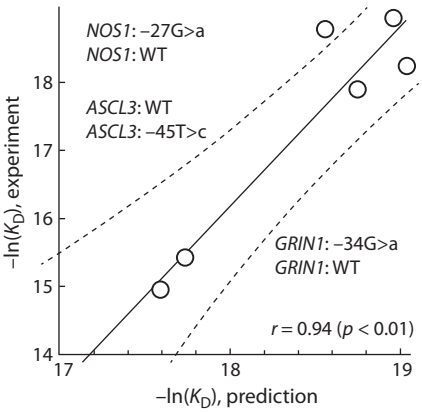


Fig. 3. The significant correlation of the experimentally measured TBP–DNA affinity values with those predicted *in silico* by means of Web service SNP_TATA_Z-tester.

The dashed curves: a 95 % confidence interval for the regression line. The estimates were made using the Statistica package (Statsoft™, USA).

affinity and gene expression – may be a candidate marker of an elevated risk of the ischemic brain injury associated with cerebral palsy. The association of *NOS1* with various diseases points to a pleiotropic role of NOS1 in many physiological processes and potentially to a pathogenesis that is shared among these diseases.

Our comparison of the experimental affinity values (K_D) of TBP–TATA complexes of reference (WT) alleles and minor alleles with those predicted by the SNP_TATA_Z-tester Web service (Ponomarenko M. et al., 2015) indicates that the data correlate well (linear correlation coefficient $r = 0.94$, $p < 0.01$) (Fig. 3).

Thus, we determined the affinity and kinetic characteristics of the interaction of TBP with TATA boxes containing unannotated SNPs. We found that these SNPs may be functionally significant and correlate with an increased risk of such neuropsychiatric diseases as schizophrenia and ischemic brain damage (associated with cerebral palsy) as well as the risk of malignant tumors: lymphoma and cancers of the cervix, epithelium, bladder, and kidneys.

Conclusion

The results show effects of the analyzed SNPs (rs1402667001, rs1049743008, and rs1195040887) in the TATA boxes within promoters of genes *GRIN1*, *ASCL3*, and *NOS1* on affinity and the rates of formation and disintegration of TBP–TATA complexes (k_a and k_d, respectively). Our experimental data suggest that the identified candidate SNP markers in neuronal genes can contribute to the development of not only neuropsychiatric but also oncological diseases, in agreement with the results obtained by other authors. Our findings about the influence of SNPs on TBP–TATA affinity and therefore on the expression of the genes in question point to their possible contribution to a higher risk of the diseases associated with these genes. Furthermore, our results have the potential to improve human health and to facilitate the development of new diagnostic markers.

References

Akyol O., Zoroglu S.S., Armutcu F., Sahin S., Gurel A. Nitric oxide as a physiopathological factor in neuropsychiatric disorders. *In Vivo*. 2004;18:377-390.

Arkova O.V., Ponomarenko M.P., Rasskazov D.A., Drachkova I.A., Arshinova T.V., Ponomarenko P.M., Savinkova L.K., Kolchanov N.A. Obesity-related known and candidate SNP markers can significantly change affinity of TATA-binding protein for human gene promoters. *BMC Genom*. 2015;16(Suppl.13):S5. DOI 10.1186/1471-2164-16-S13-S5.

Begni S., Moraschi S., Bignotti S., Fumagalli F., Rillosi L., Perez J., Gennarelli M. Association between the G1001C polymorphism in the *GRIN1* gene promoter region and schizophrenia. *Biol. Psychiatry*. 2003;53(7):617-619. DOI 10.1016/s0006-3223(02)01783-3.

Chadaeva I.V., Ponomarenko M.P., Rasskazov D.A., Sharypova E.B., Kashina E.V., Matveeva M.Y., Arshinova T.V., Ponomarenko P.M., Arkova O.V., Bondar N.P., Savinkova L.K., Kolchanov N.A. Candidate SNP markers of aggressiveness-related complications and comorbidities of genetic diseases are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters. *BMC Genom*. 2016;17(Suppl.14):995. DOI 10.1186/s12864-016-3353-3.

Chadaeva I.V., Ponomarenko P.M., Rasskazov D.A., Sharypova E.B., Kashina E.V., Zhechev D.A., Drachkova I.A., Arkova O.V., Savinkova L.K., Ponomarenko M.P., Kolchanov N.A., Osadchuk L.V., Osadchuk A.V. Candidate SNP markers of reproductive potential are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters. *BMC Genom*. 2018;19(Suppl.3):0. DOI 10.1186/s12864-018-4478-3.

Chaudhary S., Kaushik M., Kukreti R., Kukreti S. Structural switch from a multistranded G-quadruplex to single strands as a consequence of point mutation in the promoter of the human *GRIN1* gene. *Mol. Biosyst*. 2017;13(9):1805-1816. DOI 10.1039/c7mb00360a.

Ding J., Zhou H.-H., Ma Q.-R., He Z.-Y., Ma J.-B., Liu Y.-M., Zhang Y.-W., He Y.-Q., Liu J. Expression of NR1 and apoptosis levels in the hippocampal cells of mice treated with MK-801. *Mol. Med. Rep*. 2017;16(6):8359-8364. DOI 10.3892/mmr.2017.7674.

Drachkova I., Savinkova L., Arshinova T., Ponomarenko M., Peltek S., Kolchanov N. The mechanism by which TATA-box polymorphisms associated with human hereditary diseases influence interactions with the TATA-binding protein. *Hum. Mutat*. 2014;35(5):601-608. DOI 10.1002/humu.22535.

Forero D.A. Functional genomics of epileptogenesis in animal models and humans. *Cell Mol. Neurobiol*. Publ. online 28 July 2020. Publ. 2021;41:1579-1587. DOI 10.1007/s10571-020-00927-x.

Fox P.C. Acquired salivary dysfunction. Drugs and radiation. *Ann. N.Y. Acad. Sci*. 1998;842:132-137. DOI 10.1111/j.1749-632.1998.tb09641.x.

Freudenberg F., Alttoa A., Reif A. Neuronal nitric oxide synthase (*NOS1*) and its adaptor, *NOS1AP*, as a genetic risk factors for psychiatric disorders. *Genes Brain Behav*. 2015;14(1):46-63. DOI 10.1111/gbb.12193.

Gray A.L., Hyde T.M., Deep-Soboslay A., Kleinman J.E., Sodhi M.S. Sex differences in glutamate receptor gene expression in major depression and suicide. *Mol. Psychiatry*. 2015;20(9):1057-1068. DOI 10.1038/mp.2015.91.

Hanahan D., Weinberg R.A. Hallmarks of cancer: the next generation. *Cell*. 2011;144:646-674. DOI 10.1016/j.cell.2011.02.013.

Hancock D.B., Martin E.R., Vance J.M., Scott W.K. Nitric oxide synthase genes and their interactions with environmental factors in Parkinson's disease. *Neurogenetics.* 2008;9(4):249-262. DOI 10.1007/s10048-008-0137-1.

Liu Y.-P., Ding M., Zhang X.-C., Liu Y., Xuan J.-F., Xing J.-K., Xia X., Yao J., Wang B.-J. Association between polymorphisms in the *GRIN1* gene 5′ regulatory region and schizophrenia in a northern Han Chinese population and haplotype effect on protein expression *in vitro*. *BMC Med. Genet*. 2019;20(1):26. DOI 10.1186/s12881-019-0757-3.

Mishizen-Eberz A.J., Rissman R.A., Carter T.L., Ikonomovic M.D., Wolfe B.B., Armstrong D.M. Biochemical and molecular studies of NMDA receptor subunits NR1/2A/2B in hippocampal subregions throughout progression of Alzheimer's disease pathology. *Neurobiol. Dis*. 2004;15(1):80-92. DOI 10.1016/j.nbd.2003.09.016.

Mogno I., Vallania F., Mitra R.D., Cohen B.A. TATA is a modular component of synthetic promoters. *Genome Res*. 2010;20(10):1391-1397. DOI 10.1101/gr.106732.110.

Park Y.-J., Koh J., Kwon J.T., Park Y.-S., Yang L., Cha S. Uncovering stem cell differentiation factors for salivary gland regeneration by quantitative analysis of differential proteomes. *PLoS One*. 2017; 12(2):e0169677. DOI 10.1371/journal.pone.0169677.

Ponomarenko M.P., Arkova O., Rasskazov D., Ponomarenko P., Savinkova L., Kolchanov N. Candidate SNP markers of gender-biased autoimmune complications of monogenic diseases are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters. *Front Immunol*. 2016;7:130. DOI 10.3389/fimmu.2016.00130.

Ponomarenko M., Rasskazov D., Arkova O., Ponomarenko P., Suslov V., Savinkova L., Kolchanov N. How to use SNP_TATA_Comparator to find a significant change in gene expression caused by the regulatory SNP of this gene's promoter via a change in affinity of the TATA-binding protein for this promoter. *Biomed. Res. Int*. 2015; 2015:359835. DOI 10.1155/2015/359835.

Ponomarenko M., Sharypova E., Drachkova I., Chadaeva I., Arkova O., Podkolodnaya O., Ponomarenko P., Kolchanov N., Savinkova L. Unannotated single nucleotide polymorphisms in the TATA box of erythropoiesis genes show in vitro positive involvements in cognitive and mental disorders. *BMC Med. Genet*. 2020;21(Suppl.1):165. DOI 10.1186/s12881-020-01106-x.

Ponomarenko P., Chadaeva I., Rasskazov D.A., Sharypova E., Kashina E.V., Drachkova I., Zhechev D., Ponomarenko M.P., Savinkova L.K., Kolchanov N. Candidate SNP markers of familial and sporadic Alzheimer's diseases are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters. *Front. Aging Neurosci*. 2017;20(9):231. DOI 10.3389/fnagi.2017.00231.

Ponomarenko P., Rasskazov D., Suslov V., Sharypova E., Savinkova L., Podkolodnaya O., Podkolodny N.L., Tverdokhleb N.N., Chadaeva I., Ponomarenko M., Kolchanov N. Candidate SNP markers of chronopathologies are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters. *Biomed. Res. Int*. 2016;2016:8642703. DOI 10.1155/2016/8642703.

Praz V., Périer R.C., Bonnard C., Bucher P. The Eukaryotic Promoter Database, EPD: new entry types and links to gene expression data. *Nucleic Acids Res*. 2002;30:322-324. DOI 10.1093/nar/30.1.322.

Pugh B.F. Purification of the human TATA-binding protein, TBP. In: Tymms M.J. (Ed.) In Vitro Transcription and Translation Protocols. (Ser. Methods in Molecular Biology, Vol. 37). Totowa, NJ: Humana Press Inc., 1995.

Rugel-Stahl A., Elliott M.E., Ovitt C.E. Ascl3 marks adult progenitor cells of the mouse salivary gland. *Stem Cell Res*. 2012;8(3):379-387. DOI 10.1016/j.scr.2012.01.002.

Saadat M. N-methyl-D-aspartate receptor NR1 subunit gene (*GRIN1*) G1001C polymorphism and susceptibility to schizophrenia: a meta-analysis. *EXCLI J*. 2010;9:11-6.

Sharypova E.B., Drachkova I.A., Kashina E.V., Rasskazov D.A., Ponomarenko P.M., Ponomarenko M.P., Kolchanov N.A., Savinkova L.K. An experimental study of the effect of rare polymorphisms of human *HBB*, *HBD* and *F9* promoter TATA boxes on the kinetics of interaction with the TATA-binding protein. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding.* 2018;22(1):145-152. DOI 10.18699/VJ18.342. (in Russian)

Sherry S., Ward M., Kholodov M., Baker J., Phan L., Smigielski E., Sirotkin K. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 2001;29:308-311. DOI 10.1093/nar/29.1.308.

Shinkai T., Ohmori O., Hori H., Nakamura J. Allelic association of the neuronal nitric oxide synthase (NOS1) gene with schizophrenia. *Mol. Psychiatry.* 2002;7(6):560-563. DOI 10.1038/sj.mp.4001041.

Sin W.C., Haas K., Ruthazer E.S., Cline H.T. Dendrite growth increased by visual activity requires NMDA receptor and Rho GTPases. *Nature.* 2002;419(6906):475-480. DOI 10.1038/nature00987.

Turnaev I.I., Rasskazov D.A, Arkova O.V., Ponomarenko M.P., Ponomarenko P.M., Savinkova L.K., Kolchanov N.A. Hypothetical SNP markers that significantly affect the affinity of the TATA-binding protein to *VEGFA*, *ERBB2*, *IGF1R*, *FLT1*, *KDR*, and *MET* oncogene promoters as chemotherapy targets. *Molecular Biology.* 2016;50(1):141-152. DOI 10.1134/S0026893316010209.

Yu T., Xia L., Bi D., Wang Y., Shang Q., Zhu D., Song J., Wang J., Wang X., Zhu C., Xing Q. Association of *NOS1* gene polymorphisms with cerebral palsy in a Han Chinese population: a case-control study. *BMC Med. Genomics*. 2018;11(1):56. DOI 10.1186/s12920-018-0374-6.

Zhao X., Li H., Shi Y., Tang R., Chen W., Liu J., Feng G., Shi J., Yan L., Liu H., He L. Significant association between the genetic variations in the 5′ end of the N-methyl-D-aspartate receptor subunit gene *GRIN1* and schizophrenia. *Biol. Psychiatry*. 2006;59:747-753. DOI 10.1016/j.biopsych.2005.10.023.

Zou Z., Li X., Sun Y., Li L., Zhang Q., Zhu L., Zhong Z., Wang M., Wang Q., Liu Z., Wang Y., Ping Y., Yao K., Hao B., Liu Q. NOS1 expression promotes proliferation and invasion and enhances chemoresistance in ovarian cancer. *Oncol. Lett.* 2020;19(4):2989-2995. DOI 10.3892/ol.2020.11355.

Zwicker A., Denovan-Wright E.M., Uher R. Gene-environment interplay in the etiology of psychosis. *Psychol. Med*. 2018;48(12):1925-1936. DOI 10.1017/S003329171700383X.

ORCID ID

E.B. Sharypova orcid.org/0000-0002-1467-9312
I.A. Drachkova orcid.org/0000-0002-2522-1657
I.V. Chadaeva orcid.org/0000-0002-2724-5441
M.P. Ponomarenko orcid.org/0000-0003-1663-318X
L.K. Savinkova orcid.org/0000-0003-4543-4104

Acknowledgements. The work was supported by publicly funded project No. FWNR-2022-0016 of the Federal Research Center ICG SB RAS.

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

Received October 21, 2021. Revised December 27, 2021. Accepted December 30, 2021.

Вавиловский журнал генетики и селекции. 2022;26(3):234-239
DOI 10.18699/VJGB-22-30

Original Russian text www.bionet.nsc.ru/vogis/

The role of highly dispersed silica nanoparticles in the realization of the effects of granulosa on the maturation and fertilization competence of *Sus scrofa domesticus* oocytes

T.I. Kuzmina[1] ✉, I.V. Chistyakova[1], A.O. Prituzhalova[1] ✉, D.N. Tatarskaya[2]

[1] Russian Research Institute of Farm Animal Genetics and Breeding – Branch of the L.K. Ernst Federal Research Center for Animal Husbandry, Pushkin, St. Petersburg, Russia
[2] Pushkin Leningrad State University, Pushkin, St. Petersburg, Russia
✉ prof.kouzmina@mail.ru; aklevakina14@mail.ru

Abstract. Reproductive technologies are some of the key directions in the context of the need to preserve and select highly productive farmed animals in terms of economically useful traits. Improvements of the existing models of the *in vitro* oocyte maturation system help to solve the problem of low yield of porcine embryos at the final stages of preimplantation development. In the present study, a model of culture medium for gametes (NCSU-23 with 10 % homologous follicular fluid, 10 IU hCG and 10 IU eCG) modernized by the addition of $1 \cdot 10^6$ granulosa cells (GCs) per ml and 0.001 % of highly dispersed silica nanoparticles (HDSn) is proposed for use in the IVM and IVF technology of donor porcine oocytes. Analysis of the oocyte chromatin status by the Tarkowsky method and assessment of the level of destructive changes in chromatin (apoptosis, pyknosis) revealed a significant percentage increase in matured oocytes and a decrease in the proportion of granulosa cells with degenerated chromatin when using the original culture system. The positive effects of a joint addition of GCs and HDSn to the maturation system have made it possible to increase the indicators of the meiotic maturation and fertilization of oocytes. Optimal results of developmental competence of oocytes were achieved with the joint use of GCs and HDSn in the maturation system (the proportion of matured cells reached 89 %, the level of oocytes with chromosome degeneration was 12 %, 39 % of embryos reached the final stage of preimplantation development). The positive effect of HDSn on oocyte fertilization was accompanied by an abrupt decrease in destructive processes in GCs during culture in the presence of HDSn. The level of somatic cells with pyknotic nuclei was 32 % and the level of apoptosis (TUNEL-test), 21 %, compared with the control (43 and 31 %, $p < 0.01$, respectively). Thus, a high efficiency of the porcine oocyte maturation system in the joint culture of gametes with GCs and HDSn was revealed. It makes it possible to recommend a model of this culture medium at the IVM of female gametes of *Sus scrofa domesticus* for improving the quality of donor oocytes used in cell and genetic engineering.
Key words: porcine oocytes; maturation *in vitro*; highly dispersed silica nanoparticles; apoptosis; granulosa.

For citation: Kuzmina T.I., Chistyakova I.V., Prituzhalova A.O., Tatarskaya D.N. The role of highly dispersed silica nanoparticles in the realization of the effects of granulosa on the maturation and fertilization competence of *Sus scrofa domesticus* oocytes. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2022;26(3):234-239. DOI 10.18699/VJGB-22-30

Роль наночастиц высокодисперсного кремнезема в реализации эффектов гранулезы на компетентность к созреванию и оплодотворению ооцитов *Sus scrofa domesticus*

Т.И. Кузьмина[1] ✉, И.В. Чистякова[1], А.О. Притужалова[1] ✉, Д.Н. Татарская[2]

[1] Всероссийский научно-исследовательский институт генетики и разведения сельскохозяйственных животных – филиал Федерального научного центра животноводства – ВИЖ им. академика Л.К. Эрнста, Пушкин, Санкт-Петербург, Россия
[2] Ленинградский государственный университет им. А.С. Пушкина, Пушкин, Санкт-Петербург, Россия
✉ prof.kouzmina@mail.ru; aklevakina14@mail.ru

Аннотация. Репродуктивные технологии являются одним из ключевых направлений в условиях необходимости сохранения и отбора выдающихся по хозяйственно полезным признакам особей сельскохозяйственных животных. Совершенствование имеющихся моделей созревания ооцитов *in vitro* в различных вариациях способствует решению проблемы низкого выхода эмбрионов свиней на завершающих стадиях доимплантационного развития. В настоящем исследовании с использованием технологии созревания и оплодотворения донорских ооцитов свиней *in vitro* предложена модель среды для культивирования гамет (NCSU-23 с 10% гомологичной фолликулярной жидкостью, 10 МЕ ХГЧ и 10 МЕ ХГ лошади), модернизированная введением

$1 \cdot 10^6$ клеток гранулезы (КГ) на 1 мл среды и 0.001 % наночастиц высокодисперсного кремнезема (нВДК). Анализ статуса хроматина ооцитов по методу Тарковского и оценка уровня деструктивных изменений хроматина соматических клеток овариальных фолликулов (апоптоз, пикноз) выявили значительное повышение показателей ядерного созревания гамет и снижение доли клеток гранулезы с дегенерированным хроматином при применении разработанной системы культивирования. Обнаружено позитивное влияние совместного введения КГ и нВДК в систему дозревания, позволившего увеличить показатели мейотического созревания и оплодотворяемости ооцитов. Оптимальные показатели фертильности ооцитов достигнуты при сочетанном использовании в системе дозревания КГ и нВДК (доля созревших клеток достигла 89 %, уровень ооцитов с дегенерацией хромосом составил 12 %, 39 % эмбрионов достигли завершающей стадии доимплантационного развития). Положительный эффект нВДК на показатели оплодотворяемости ооцитов сопровождался резким снижением деструктивных процессов в КГ при их культивировании в присутствии нВДК. Уровень пикнозов составил 32 %, а уровень апоптозов (TUNEL-test) – 21 % по сравнению с контролем (43 и 31 % соответственно, $p < 0.01$). Таким образом, выявлена высокая эффективность системы созревания ооцитов свиней в условиях совместного кокультивирования гамет с КГ и нВДК, что позволяет рекомендовать модель разработанной среды в технологии экстракорпорального созревания женских гамет *Sus scrofa domesticus* для повышения качества донорских яйцеклеток, используемых в клеточной и генетической инженерии.
Ключевые слова: ооциты свиней; созревание *in vitro*; наночастицы высокодисперсного кремнезема; апоптоз; гранулеза.

Introduction

Cell reproductive and DNA biotechnologies play an important role in the intensification of breeding process in animal husbandry because they are a tool for increasing the number of individuals outstanding in economic traits (Romar et al., 2019). Biotechnological interest in the species *Sus scrofa domesticus* has increased because it can be used in biomedicine, due to the features of its physiology (proximity to the species *Homo sapiens*), for organ xenotransplantation. *In vitro* production of viable native and reconstructed (cloned, transgenic) porcine embryos on a mass scale is possible; however, now, certain steps in the technology of *in vitro* maturation of *S. scrofa domesticus* eggs and their fertilization require improvement (Fowler et al., 2018). Development of standardized protocols for methodology of obtaining porcine embryos *in vitro* is necessary to take full advantage of the possibilities of innovative cellular reproductive technologies in porcine breeding and biomedicine, including production of genetically modified pigs.

Effectiveness on various stages of extracorporeal production of porcine embryos is ambiguous. Improvement of oocyte maturation systems, low percentage of monospermic zygotes and zygotes that develop to the final stage of preimplantation development (blastocyst) require solutions (Martinez et al., 2019). Nowadays, there are many works on the development of a unified maturation system of donor porcine oocytes *in vitro*, but yield of embryos at the final stages of preimplantation development still does not exceed 45–50 % (Soriano-Úbeda et al., 2017).

The abovementioned allows us to define the task of modeling the culture media composition for completion of porcine oocyte meiotic maturation *in vitro* as highly relevant. *In vivo*, an egg is forming in close relationship with the somatic cells of the ovarian follicle (cumulus, granulosa), which produce a number of bioactive molecules involved in the growth and maturation of oocytes. The pioneering works of L.R. Abeydeera showed effectiveness of using follicle walls and follicular fluid as part of oocyte maturation systems (Abeydeera et al., 1998). However, procedures to dissect follicle, objectivity of its quality evaluation by embryotechnologist prolong the duration of the first stage of embryo production technology. The use of innovative materials, including their nanoscale particles, in *in vitro* maturation system of animal gametes is a rapidly developing branch of bionanotechnology (Remião et al., 2018). Many researchers have evaluated the cyto- and gene-toxicity of nanoparticles of different origin on mammalian germ cells (Roy et al., 2020).

Our previous studies revealed positive effects of HDSn on cell compartments functioning of native and devitrified female gametes of farm animals, destructive chromatin processes in the nuclei of germinal and somatic cells of ovarian follicles (Kuzmina et al., 2017, 2020). Based on these considerations, it seems logical to add granulosa cells into the basic culture media as a potential supplier of natural origin biologically active substances, primarily steroids, and nanoparticles of different origin.

The aim of this study was to evaluate the role of highly dispersed silica nanoparticles in realizing the effects of the addition of granulosa cells into the system of extracorporeal maturation of porcine oocytes on gamete fertility indices.

Materials and methods

All reagents used in the experiments, except as indicated in the text, were produced by Sigma-Aldrich (USA). Plastic laboratory glassware was from BD Falcon™ (USA).

In experiments we used cumulus oocyte complexes (COC) isolated from the antral follicles of *post mortem* ovaries of *S. scrofa domesticus* landrace breed at the age of 6–8 months. The ovaries after ovariectomy of animals at a local slaughterhouse were delivered to laboratory in 0.9 % NaCl solution at 30–35 °C containing 100 IU/ml penicillin, 100 μg/ml streptomycin, and 0.25 ng/ml amphotericin. COCs were aspirated from antral follicles (with high turgor, from 3 to 6 mm diameter, and a high degree of vascularization). Oocytes with homogeneous ooplasm, zone pellucid uniform in width, surrounded by a compact layer of cumulus cells (at least 5–6 layers) were used in experiments.

After morphological evaluation, 40–50 COCs were placed in droplets (500 μl volume) of culture media with the following composition: Group I – synthetic culture medium North

Carolina State University-23 (NCSU-23) + 10 IU human chorionic gonadotropin + 10 IU equine chorionic gonadotropin + 10 % follicular fluid (follicle diameter 3–6 mm); Group II – synthetic culture medium NCSU-23 + 10 IU human chorionic gonadotropin + 10 IU equine chorionic gonadotropin + 10 % follicular fluid (follicle diameter 3–6 mm) + 0.001 % HDSn; Group III – synthetic culture medium NCSU-23 + 10 IU human chorionic gonadotropin + 10 IU equine chorionic gonadotropin + 10 % follicular fluid (follicle diameter 3–6 mm) + $1 \cdot 10^6$ granulosa cells (GC) per ml of medium; Group IV – synthetic culture medium NCSU-23 + 10 IU human chorionic gonadotropin + 10 IU equine chorionic gonadotropin + 10 % follicular fluid (follicle diameter 3–6 mm) + $1 \cdot 10^6$ GC per ml of medium + 0.001 % HDSn. HDSn were synthesized in the Chuiko Institute of Surface Chemistry, National Academy of Sciences of Ukraine. Concentration was chosen according to the recommendations of developers (Zyuzyn et al., 2015). COCs were cultured for 22 hours at 38.5 °C in the atmosphere of 5 % CO_2 in aforementioned media, then the media were changed, hormones were excluded in all studied groups and they were cultured for the next 22 hours.

The chromatin status of oocytes at meiotic maturation and the level of pyknosis in granulosa cells were tested with cytology method (Kuzmina et al., 2008). Oocytes were placed for 5–10 min in a warm (37 °C) 0.9 % hypotonic solution of 3-substituted sodium citrate and purified from cumulus. Then cells were transferred on dry non-fat glass and fixed with a mixture of methanol and acetic acid (3:1). Dried samples of oocyte and granulosa cells were stained with 4 % Romanovsky–Giemsa solution (azure-eosin) for 3–4 min.

The level of apoptosis in GCs after culture for 22 hours in NCSU-23 medium with 10 IU human chorionic gonadotropin, 10 IU equine chorionic gonadotropin, 10 % follicular fluid (follicle diameter 3–6 mm) and 22 hours later (total culture time 44 hours) after culture medium change (exclusion of hormonal supplements) was assessed by TUNEL (Janowski et al., 2012). The experimental group was supplemented with 0.001 % HDSn at all stages of culture.

We used modified mTBM medium containing 113.1 mM NaCl, 3.0 mM KCl, 7.5 mM $CaCl_2 \cdot 2H_2O$, 20.0 mM Tris, 11.0 mM glucose, 5.0 mM sodium pyruvate, 1 mM caffeine and 0.1 % BSA for *in vitro* fertilization, after 44 hours of culture, the oocytes were mechanically (by pipetting) released from the cumulus cells. Then in amount of 10 pcs. placed in drops of mTBM medium (volume 90 µl under paraffin oil) in 35 mm culture dishes for 30 min in CO_2 incubator for equilibration. The oocytes were fertilized with native sperm (initial concentration in diluent $3 \cdot 10^9$ spermatozoa per ml). After 3-fold centrifugation (80 g for 3 min at room temperature), 10 ml of sperm suspension was resuspended in 10 ml of DPBS with 0.1 % BSA and sperm concentration was adjusted to $2 \cdot 10^6$ cells per ml. 16 µl of sperm suspension was added to 90 µl droplets with oocyte and cultured in a CO_2 incubator at 38.5 °C in an atmosphere of 5 % CO_2 and 90 % humidity. After 6 hours of incubation with spermatozoa, the oocytes were transferred to 500 µl of NCSU-23 medium with 0.4 % BSA for culturing in a CO_2 incubator for 7 days at 38.5 °C in an atmosphere of 5 % O_2, 5 % CO_2 and 90 % N_2 with medium changes every 48 hours of culture (Egerszegi et al., 2010).

To determine the level of apoptosis in GCs, its suspension was placed on poly-L-lysine-coated slides and dried. Next, apoptosis levels were tested according to the manufacturer's instructions and the method adapted for granulosa cells presented by Janowski et al. (2012). For this purpose, GCs were fixed in 4 % (v/v) paraformaldehyde solution for 30 min, incubated for 2 min in 10 % Triton X-100 solution on 0.1 % sodium citrate. Then GCs were incubated with TUNEL reagent (Roche Diagnostics, GmbH, Mannheim, Germany) for 60 min at 37 °C in the dark. After incubation cells were washed in DPBS solution, stained in 0.1 % (w/v) propidium iodide solution (20 min exposure), washed again in DPBS, and exposed for 1 hour in the dark at room temperature. Samples were stored in the refrigerator at +3 to +5 °C. Samples were analyzed using a ZEISS AxioLab. fluorescence microscope A1 (Carl Zeiss, Germany).

The results were processed using the SigmaStat statistical software package (Jandel Scientific Software, USA). Pearson's χ^2 test was used to assess the reliability of frequency variables. Significance of differences between the compared values was assessed at the following levels: $p < 0.05$, $p < 0.01$, and $p < 0.001$ for 3–5 independent experiments.

Results and discussion

Granulosa and cumulus cells produce a great number of growth and other factors determining oocyte formation and subsequent embryo development (Canipari, 2000). Nanoparticles of various chemical compounds, including HDSn, can synchronize the nuclear and cytoplasmic maturation of animal oocytes and protect intracellular components from factors detrimental to their functioning, including reactive oxygen species (ROS) (Kuzmina et al., 2017, 2020). Data of chromatin status analysis in porcine oocytes were cultured with granulosa cells and HDSn are presented in Figure 1.

Addition of HDSn to culture medium promoted re-initiation and completion of meiosis (Fig. 2) in oocytes cultured without GC compared to cells in the control group (79 and 75 % versus 89 and 84 %, $p < 0.05$). Moreover, the stimulating effect of HDSn's on oocyte maturation was also observed in co-cultured gametes with GC (85 and 79 % versus 93 and 89 %, $p < 0.05$). It is important to note that adding HDSn resulted in a decrease in the percentage of degenerated oocytes in experimental groups compared to control groups cultured with or without somatic cells (15 and 12 % versus 25 and 18 %, $p < 0.01$).

In the second series of experiments, we evaluated the effect of HDSn on destructive processes in GCs during *in vitro* culture (Fig. 3). The inhibitory effect of HDSn on destructive processes in chromatin (apoptosis, pyknosis) of granulosa cells during prolonged culture was shown. Thus, after 22 hours of culture, proportion of cells with pyknotic nuclei was lower by 7 % in the group cultured with HDSn compared to the control (21 and 28 %, $p < 0.01$), and proportion of apoptotic cells – by 6 % (13 and 19 %, $p < 0.05$). After 44 hours of culture, proportion of cells in pyknotic state reached 43 % ($p < 0.01$) in the control group and the level of apoptotic cells reached 31 % ($p < 0.01$). In the experimental group, these indices were significantly lower (32 and 21 %, $p < 0.01$).

The results of analysis fertility parameters of oocytes matured in various systems are shown in Fig. 4. Fertilizable

addition of HDSn to maturation medium provided an increase in the level of fertilized oocytes, which was expressed by an increase of 12 % in the cleavage level (51 %, $p < 0.05$) and 11 % in the yield of blastocysts (23 %, $p < 0.01$) and an increase in the yield of preimplantation embryos at blastocyst stage compared with the control group (39 and 12 %, respectively). At the same time, maximum fertilization rates were observed in the group of oocytes co-cultured with granulosa and HDSn cells (61 and 39 %, respectively, $p < 0.01$ versus control groups).

Oxidative stress is one of the main factors reducing development competence of oocytes at culture (Wei et al., 2016). Positive effect of nanoparticles on oocyte maturation can probably be explained by the ability of HDSn to level the damaging effect of free-radical processes at cell culture by reducing the formation of oxidative modification products of proteins (Savchenko, 2013). In addition, as a result of oxidative stress on cells in the endoplasmic reticulum (ER), occur synthesis and assembly of lipid droplets (LD), which act as a protective mechanism when reactive oxygen species act on the membrane structures of organelles, as well as supply mitochondria with fatty acids for ATP production (Lee et al., 2012; Zhang X., Zhang K., 2012). It is known that the intracellular form of LDs in the shape of "granules" and their diffuse arrangement provide fatty acid mobilization, determining normal maturation of cumulus oocyte complexes (Bradley et al., 2019). It was shown that the addition of HDSn to culture medium provides an increase in the level of oocytes with the LDs with diffuse localization, which, as indicated earlier, ensures normal gamete development (Novichkova, Kuzmina, 2019).

It is known that communication of GCs, as well as interaction of cumulus cells with the oocyte, determines the growth and formation of female gamete. Successful oocyte maturation and further embryonic development depend on the action of certain hormones, in particular progesterone and estradiol secreted by granulosa cells. In turn, the effect of these hormones on oocytes is mediated by the radial crown cells expressing FSHR (follicle-stimulating hormone receptor), which is necessary for cumulus cell proliferation and normal gamete development (Okazaki et al., 2003). It was shown that the addition of progesterone and β-estradiol in culture medium increases the level of FSHR expression, cumulus cell survival, and reduces the level of apoptosis (Okamoto et al., 2016). HDSn prevent apoptosis in somatic cells and male gametes of animals by stimulating antioxidant system through interacting with receptors on the cell surface (Boytseva et al., 2017; Kuzmina et al., 2017).

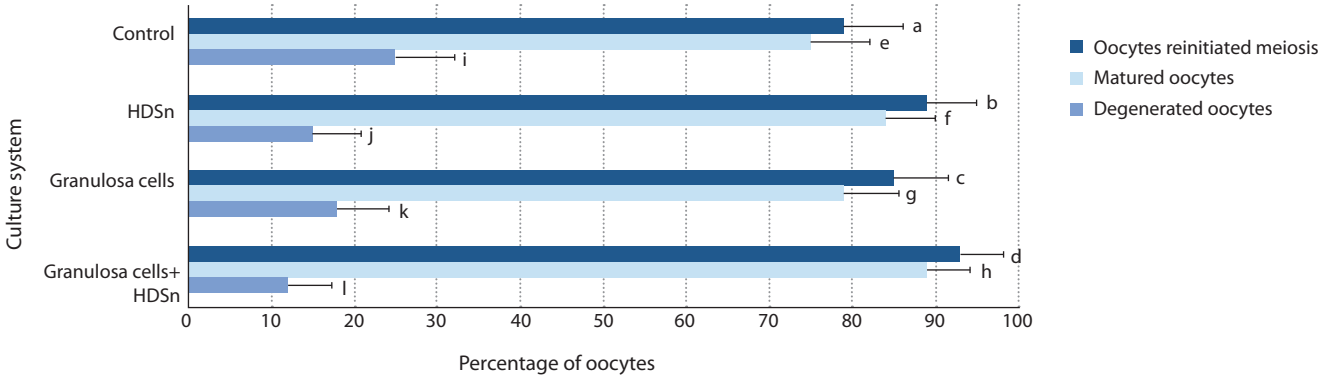


Fig. 1. Indicators of porcine oocytes chromatin status after culture with granulosa cells and HDSn (time of culture – 44 hours, number of oocytes – 600).

* Differences are statistically significant (χ^2 test): a : b; c : d; e : f; g : h; i : j – $p < 0.05$; a : d; e : h; i : l – $p < 0.1$.

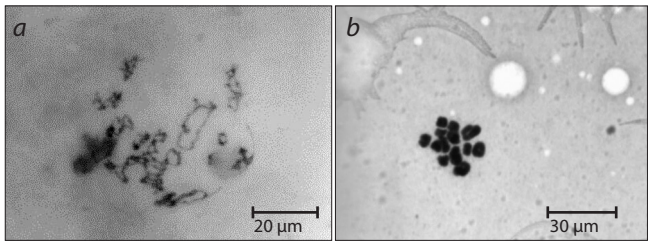


Fig. 2. Representative image of *S. scrofa domesticus* oocytes chromatin at diplotene (*a*) and metaphase II (*b*) stages.

Cytological sample, staining with azure-eosin according to Romanovsky–Giemsa, microscope ZEISS AxioLab. A1, Carl Zeiss.

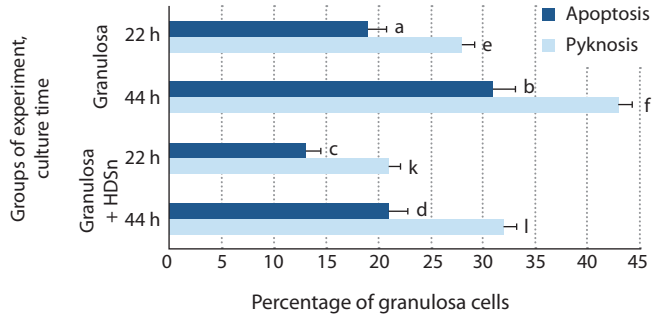


Fig. 3. Destructive processes of chromatin in granulosa cells of porcine ovarian follicles (number of cells –7539).

* Differences are statistically significant (χ^2 test): a : c; a : d; e : l – $p < 0.05$; a : b; b : c; b : d; c : d; e : f; e : k; f : k; f : l; k : l – $p < 0.01$.

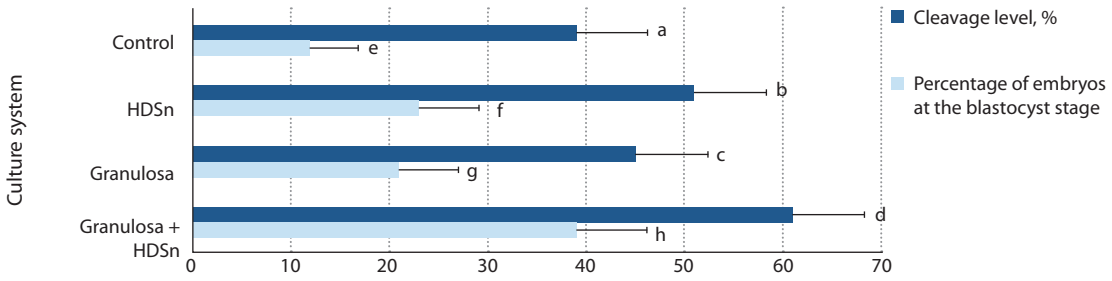


Fig. 4. Analysis of fertility indicators of *S. scrofa domesticus* oocytes matured in different culture systems (number of oocytes – 736).
* Differences are statistically significant (χ^2 test): a : b, e : g – $p < 0.05$; a : d, c : d, e : f, e : h, f : h, g : h – $p < 0.01$.

Conclusion

The development of an effective protocol for obtaining native and reconstructed *S. scrofa domesticus* embryos *in vitro* will significantly intensify stages of innovative cellular reproductive technologies used in animal husbandry, veterinary medicine and biomedicine. The aims of the present study are to improve extracorporeal maturation system of donor porcine oocytes to obtain oocytes competent for fertilization and embryo development. Considering the importance of somatic cells of ovarian follicles in the formation of a mature oocyte, coculture of cumulus-oocyte complexes with granulosa cells was used in the experiments. Maturation system was upgraded by the addition of HDSn into culture medium.

Experiments revealed a positive effect of the developed system on indicators of fertility of oocytes (yield of matured oocytes, cleavage, and level of embryos that reached the final stage of preimplantation development). The most positive effect was observed when HDSn and granulosa cells were used in the culture system together. High fertility rates of oocytes matured in medium with HDSn are probably explained by a reduced level of destructive changes in surrounding cumulus cells (subpopulation of granulosa cells).

The study found that the addition of HDSn into the culture medium causes the levels of apoptosis and pyknosis in granulosa cells to decrease, which indicates an increase in the number of viable cells, their hormonesynthetic activity and provides physiological processes involved in formation of oocytes with high fertility. The most significant indicator in evaluation of the effectiveness of any culture system for maturation of oocytes is the embryo yield. In our studies, the yield of embryos at the final stage of preimplantation development was the highest (39 %) in the case of the joint use of HDSn and granulosa cells in oocyte maturation system. Results of the study allow to recommend the developed culture system for extracorporeal maturation of donor oocytes of *S. scrofa domesticus*.

References

Abeydeera L.R., Wang W.H., Cantley T.C., Rieke A., Day B.N. Coculture with follicular shell pieces can enhance the developmental competence of pig oocytes after in vitro fertilization: relevance to intracellular glutathione. *Biol. Reprod*. 1998;58(1):213-218. DOI 10.1095/biolreprod58.1.213.

Boytseva E.N., Bychkova N.V., Kuzmina T.I. Effects of highly dispersed silica nanoparticles on the apoptosis of *Bos taurus* spermatozoa. *Tsitologiya = Cytology*. 2017;5(59):375-380. (in Russian)

Bradley J., Pope I., Wang Y., Langbein W., Borri P., Swann K. Dynamic label-free imaging of lipid droplets and their link to fatty acid and pyruvate oxidation in mouse eggs. *J. Cell Sci*. 2019;132(13):jcs228999. DOI 10.1242/jcs.228999.

Canipari R. Oocyte-granulosa cell interactions. *Hum. Reprod. Update*. 2000;6(3):279-289. DOI 10.1093/humupd/6.3.279.

Egerszegi I., Alm H., Rátky J., Heleil B., Brüssow K.P., Torner H. Meiotic progression, mitochondrial features and fertilisation characteristics of porcine oocytes with different G6PDH activities. *Reprod. Fertil. Dev*. 2010;22(5):830-838. DOI 10.1071/RD09140.

Fowler K.E., Mandawala A.A., Griffin D.K., Walling G.E., Harvey S.C. The production of pig preimplantation embryos *in vitro*: current progress and future prospects. *Reprod. Biol*. 2018;18(3):203-211. DOI 10.1016/j.repbio.2018.07.001.

Janowski D., Salilew-Wondim D., Torner H., Tesfaye D., Ghanem N., Tomek W., El-Sayed A., Schellander K., Holker M. Incidence of apoptosis and transcript abundance in bovine follicular cells is associated with the quality of the enclosed oocyte. *Theriogenology*. 2012;78(3):656-669. DOI 10.1016/j.theriogenology.2012.03.012.

Kuzmina T.I., Alm H., Torner H. Methods of Porcine Embryos Production *in vitro*. St. Petersburg; Puskin, 2008. (in Russian)

Kuzmina T.I., Chistyakova I.V., Tatarskaya D.N. The influence of highly dispersed silica nanoparticles on the functional activity of mitochondria and chromatin state in native and devitrified *Bos taurus* oocytes. *Sel'skokhozyaystvennaya Biologiya = Agricultural Biology*. 2020;55(4):784-793. DOI 10.15389/agrobiology.2020.4.784eng.

Kuzmina T.I., Novichkova D.A., Chistyakova I.V., Epishko O.A. Effects of highly dispersed silica nanoparticles on the chromatin in somatic cells of porcine follicles. *Veterinariya = Veterinary Medicine Journal*. 2017;2:43-45. (in Russian)

Lee J.S., Mendez R., Heng H.H., Yang Z.Q., Zhang K. Pharmacological ER stress promotes hepatic lipogenesis and lipid droplet formation. *Am. J. Transl. Res.* 2012;4(1):102-113.

Martinez E.A., Martinez C.A., Cambra J.M., Maside C., Lucas X., Vazquez J.L., Vazquez J.M., Roca J., Rodriguez-Martinez H., Gil M.A., Parrilla I., Cuello C. Achievements and future perspectives of embryo transfer technology in pigs. *Reprod. Domest. Anim*. 2019;54(4):4-13. DOI 10.1111/rda.13465.

Novichkova D.A., Kuzmina T.I. Effect of highly dispersed silica nanoparticles on the functioning of lipidome in *Sus scrofa domesticus* oocytes. *Meditsina Ekstremal'nykh Situatsiy = Medicine of Extreme Situations*. 2019;S1:30-34. (in Russian)

Okamoto A., Ikeda M., Kaneko A., Kishida C., Shimada M., Yamashita Y. The novel pig in vitro maturation system to improve developmental competence of oocytes derived from atretic non-vascularized follicle. *Biol. Reprod*. 2016;95(4):7. DOI 10.1095/biolreprod.116.138982.

Okazaki T., Nishibori M., Yamashita Y., Shimada M. LH reduces proliferative activity of cumulus cells and accelerates GVBD of porcine oocytes. *Mol. Cell. Endocrinol.* 2003;209:43-50. DOI 10.1016/j.mce.2003.08.002.

Remião M.H., Segatto N.V., Pohlmann A., Guterres S.S., Seixas F.K., Collares T. The potential of nanotechnology in medically assisted reproduction. *Front. Pharmacol*. 2018;8:994. DOI 10.3389/fphar.2017.00994.

Romar R., Canovas S., Matas C., Gadea J., Coy P. Pig in vitro fertilization: where are we and where do we go? *Theriogenology*. 2019;137:113-121. DOI 10.1016/j.theriogenology.2019.05.045.

Roy P.K., Qamar A.Y., Fang X., Kim G., Bang S., Zoysa M., Shin S.T., Cho J. Chitosan nanoparticles enhance developmental competence of in vitro-matured porcine oocytes. *Reprod. Domest. Anim.* 2020;56(2):342-350. DOI 10.1111/rda.13871.

Savchenko D.S. Studyng of antioxidatic properties of nanocomposite of highly dispersive silicon dioxide with silver nanoparticles. *Meditsina i Obrazovanie v Sibiri = Medicine and Education in Siberia*. 2013;6:23-30. (in Russian)

Soriano-Úbeda C., García-Vazquez F.A., Romero-Aguirregomezcorta J., Matas C. Improving porcine *in vitro* fertilization output by simulating the oviductal environment. *Sci. Rep*. 2017;7:43616. DOI 10.1038/srep43616.

Wei J.H., Yuan X.Y., Zhang J.M., Wei J.Q. Caspase activity and oxidative stress of granulosa cells are associated with the viability and developmental potential of vitrified immature oocytes. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 2016;198:22-26. DOI 10.1016/j.ejogrb.2015.12.010.

Zhang X., Zhang K. Endoplasmic reticulum stress-associated lipid droplet formation and type II diabetes. *Biochem. Res. Int*. 2012;2012:247275. DOI 10.1155/2012/247275.

Zyuzyn A.B., Shcherbak O.V., Osypchuk O.S., Kovtun S.I., Dzitsyuk V.V. Using of nanomaterials in embryogenetic system for receiving pig's embryos *in vitro*. *Faktory Eksperimentalnoi Evolutsii Organizmov = Factors in Experimental Evolution of Organisms*. 2015;17:164-168. (in Ukrainian)

ORCID ID

T.I. Kuzmina orcid.org/0000-0002-4218-6080
I.V. Chistyakova orcid.org/0000-0001-7229-5766
A.O. Prituzhalova orcid.org/0000-0002-2865-9582
D.N. Tatarskaya orcid.org/0000-0002-8834-1912

Благодарности. This work was supported by the Ministry of Science and Higher Education of the Russian Federation (Project No. 0445-2021-0005 and 121052600350-9).

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

Received November 8, 2021. Revised January 9, 2022. Accepted January 11, 2022.

Вавиловский журнал генетики и селекции. 2022;26(3):240-249
DOI 10.18699/VJGB-22-31

Biochemical and technological properties of moose (*Alces alces*) recombinant chymosin

D.V. Balabova[1], A.P. Rudometov[2], S.V. Belenkaya[1, 2, 3], A.N. Belov[4], A.D. Koval[4], A.A. Bondar[5], A.Yu. Bakulina[2, 3], E.A. Rukhlova[2], V.V. Elchaninov[4], D.N. Shcherbakov[1, 2]

[1] Altai State University, Barnaul, Russia
[2] State Research Center of Virology and Biotechnology "Vector", Koltsovo, Novosibirsk Region, Russia
[3] Novosibirsk State University, Novosibirsk, Russia
[4] Federal Altai Scientific Center for Agrobiotechnology, Siberian Research Institute of Cheese-Making, Barnaul, Russia
[5] Institute of Chemical Biology and Fundamental Medicine of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
dnshcherbakov@gmail.com

Abstract. Recombinant chymosins (rChns) of the cow and the camel are currently considered as standard milk coagulants for cheese-making. The search for a new type of milk-clotting enzymes that may exist in nature and can surpass the existing "cheese-making" standards is an urgent biotechnological task. Within this study, we for the first time constructed an expression vector allowing production of a recombinant analog of moose chymosin in the expression system of *Escherichia coli* (strain SHuffle express). We built a model of the spatial structure of moose chymosin and compared the topography of positive and negative surface charges with the correspondent structures of cow and camel chymosins. We found that the distribution of charges on the surface of moose chymosin has common features with that of cow and camel chymosins. However, the moose enzyme carries a unique positively charged patch, which is likely to affect its interaction with the substrate. Biochemical and technological properties of the moose rChn were studied. Commercial rChns of cow and camel were used as comparison enzymes. In some technological parameters, the moose rChn proved to be superior to the reference enzymes. Compared with the cow and camel rChns, the moose chymosin specific activity is less dependent on the changes in $CaCl_2$ concentration in the range of 1–5 mM and pH in the range of 6–7, which is an attractive technological property. The total proteolytic activity of the moose rChn occupies an intermediate position between the rChns of cow and camel. The combination of biochemical and technological properties of the moose rChn argues for further study of this enzyme.
Key words: moose; recombinant chymosin; milk-clotting activity; biochemical properties; cheese-making; *Alces alces*.

For citation: Balabova D.V., Rudometov A.P., Belenkaya S.V., Belov A.N., Koval A.D., Bondar A.A., Bakulina A.Yu., Rukhlova E.A., Elchaninov V.V., Shcherbakov D.N. Biochemical and technological properties of moose (*Alces alces*) recombinant chymosin. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2022;26(3):240-249. DOI 10.18699/VJGB-22-31

Биохимические и технологические свойства рекомбинантного химозина лося (*Alces alces*)

Д.В. Балабова[1], А.П. Рудометов[2], С.В. Беленькая[1, 2, 3], А.Н. Белов[4], А.Д. Коваль[4], А.А. Бондарь[5], А.Ю. Бакулина[2, 3], Е.А. Рухлова[2], В.В. Ельчанинов[4], Д.Н. Щербаков[1, 2]

[1] Алтайский государственный университет, Барнаул, Россия
[2] Государственный научный центр вирусологии и биотехнологии «Вектор» Роспотребнадзора, р. п. Кольцово, Новосибирская область, Россия
[3] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
[4] Федеральный Алтайский научный центр агробиотехнологий, Сибирский научно-исследовательский институт сыроделия, Барнаул, Россия
[5] Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук, Новосибирск, Россия
dnshcherbakov@gmail.com

Аннотация. Эталонными коагулянтами молока для сыроделия в настоящее время считаются рекомбинантные химозины (рХн) коровы и верблюда. Нахождение молокосвертывающих ферментов, способных превзойти эталонные коагулянты молока, является актуальной биотехнологической задачей. Нами сконструирован экспрессирующий вектор, который позволил впервые получить рекомбинантный аналог химозина лося в системе экспрессии *Escherichia coli* (штамм SHuffle Express). Построена модель пространственной структуры химозина лося, и проведено сравнение топографии положительных и отрицательных поверхностных зарядов с соответствующими структурами химозинов коровы и верблюда. Обнаружено, что распределение зарядов на поверхности химозина лося имеет общие черты с распределением зарядов химозинов коровы и верблюда. Отличительная особенность химозина лося – наличие положительно заряженного поверхностного участка, который, вероятно, влияет на его взаимодействие с субстратом. Исследованы основные биохимические и технологические свойства рХн лося. В качестве ферментов сравнения использовали коммерческие рХн

коровы и верблюда. Установлено, что по некоторым технологическим показателям рХн лося превосходил ферменты сравнения. По сравнению с рХн коровы и верблюда специфическая активность рХн лося в меньшей степени зависит от изменения концентрации $CaCl_2$ в диапазоне 1–5 мМ и pH в диапазоне 6–7, что является привлекательным технологическим свойством. По общей протеолитической активности рХн лося занимает промежуточное положение между ферментами коровы и верблюда. Совокупность биохимических и технологических свойств рХн лося свидетельствует о необходимости дальнейшего изучения этого фермента.
Ключевые слова: лось; рекомбинантный химозин; молокосвертывающая активность; биохимические свойства; сыроделие; *Alces alces*.

Introduction

The segment of recombinant enzymes occupies a significant part of the modern biotechnology market (Trono, 2019). One of the first industrial enzymes obtained using genetic engineering technologies was cow's recombinant chymosin (Flamm, 1991), which has been considered the standard of a milk-clotting enzyme (ME) in cheese-making for a long time (Belov et al., 2009; Jacob et al., 2011). The rapid development of molecular biology methods (primarily next-generation sequencing and genetic engineering) has intensified the search for enzymes with superior biochemical and technological properties compared to the milk coagulants traditionally used in the industry.

The main goal of such a search is to find the enzymes possessing approximately the same sensitivity to pH and concentration of calcium ions in milk compared with a bovine (*Bos taurus*) chymosin but would outperform it in a milk-clotting activity (MA) and, at the same time, would demonstrate a lower overall proteolytic activity (PA) and thermal stability (TS).

Previously, researchers have obtained and studied rChns of sheep (*Ovis aries*) (Rogelj et al., 2001), goat (*Capra hircus*) (Vega-Hernández et al., 2004; Vallejo et al., 2012; Tyagi et al., 2016), water buffalo (*Bubalus arnee bubalis*) (Vallejo et al., 2012; Tyagi et al., 2017), and camel (*Camelus dromedarius*) (Kappeler et al., 2006). It was shown that the rChns of goat, buffalo, and sheep are ordinary ME and cannot compete with the bovine Chn. The camel rChn showed a higher affinity toward bovine κ-casein (κ-Cs) and had a better MA/PA ratio than the cow rChn, but was inferior to the bovine enzyme in TS (Bansal et al., 2009). Nevertheless, after a comprehensive study of its biochemical and technological properties, the camel's rChn is widely used in the practice of cheese-making (Bansal et al., 2009; Moynihan et al., 2014; Gumus, Hayaloglu, 2019) and is now considered a reference ME along with the bovine rChn.

Later, the rChns of yak (*Bos grunniens*) (Luo et al., 2016; Ersöz, İnan, 2019), alpaca (*Vicugna pacos*) (Belenkaya et al., 2018), and Altai maral (*Cervus elaphus sibiricus*) (Belenkaya et al., 2020a, b) have been obtained and studied. The complete biochemical and technological characteristics of the yak rChn have not yet been established. The available literature data indicate, on the one hand, that the yak rChn has a low technologically significant threshold for TS, and on the other hand, a higher total PA, compared to that of the bovine rChn (Belenkaya et al., 2020c). According to our data, the rChn of maral has an excessively high total PA and TS, limiting its potential use only to the production of cheeses with short ripening and storage periods (Belenkaya et al., 2020a). The genetically engineered Chn of alpaca surpasses the bovine rChn in the MA/PA ratio but is inferior in this parameter to the rChn of the camel. In addition, similar to the rChn of camel, the rChn of alpaca has a higher TS than the bovine rChn. Taken together, the facts mentioned above show that despite the presence of some interesting characteristics, the complex of biochemical and technological properties of the rChns of yak, alpaca, and maral hardly allows these enzymes to be considered as an alternative to the rChns of cow and camel.

Here we present a new milk-clotting enzyme – recombinant chymosin of moose (rChn-Alc) in a prokaryotic expression system and investigate some of its biochemical properties in comparison with standard milk-clotting enzymes. The recombinant prochymosin (rProChn) of moose developed in the prokaryotic expression system was activated by a stepwise pH change method, and an active rChn-Alc preparation capable of effectively coagulating cow's milk was obtained. It was shown that compared to the reference milk coagulants, the specific enzymatic activity of rChn-Alc was less sensitive to changes in the H^+ concentration in the pH range of 6.0–7.0. Concerning an important technological indicator, the total PA, rChn-Alc was found to occupy an intermediate position between the cow rChn and the camel rChn. The specific MA of the moose rChn was lower than that of the cow and camel rChns, possibly due to incomplete refolding of the enzyme obtained in the *E. coli* expression system. The results obtained expand the understanding of the biochemical and technological properties of Chns of various species and create a basis for further search for technological coagulants of cow's milk that would surpass the existing reference milk-converting enzymes in their properties.

Materials and methods

Work organization. The optimization of the structure of the moose prochymosin gene and the construction of a producer strain were carried out at the State Research Center of Virology and Biotechnology "Vector". Works on obtaining a preparation of recombinant moose prochymosin and determining its biochemical and technological properties were carried out at Altai State University. All work was carried out in 2019.

Strains and media. *Escherichia coli* strain NEB Stable used to construct and propagate all plasmids was purchased from New England Biolabs (NEB, Ipswich, USA). *E. coli* strain SHuffle express was purchased from New England Biolabs (NEB, Ipswich, USA) and used as a heterologous host to produce the rProChn of moose (GenBank MT542132). The medium Lysogeny broth (LB) (1.0 % bacto-peptone, 0.5 % yeast extract, and 1.0 % NaCl) in liquid or solid (1.5 % agar) form was used to culture NEB stable cells at 37 °C. *E. coli*

SHuffle express cells were cultured at 30 °C in LB medium (AppliChem, USA) with the addition of isopropyl-β-D-1-thiogalactopyranoside (IPTG) for induction (final concentration 1.0 mM).

Subcloning of prochymosin gene into pET21a expression vector. Codon optimization of the moose prochymosin sequence (accession number MT542132) for the selected expression system was performed by the online service Integrated DNA Technologies (https://eu.idtdna.com/CodonOpt), followed by synthesis and integration into pGH cloning plasmid. Synthetic gene sequence containing *BamHI* and *HindIII* restriction sites at the 5′- and 3′-ends, respectively, was digested and subcloned into the expression vector pET21a (Novagen, Germany). The structure of the recombinant plasmid was verified by Sanger sequencing. As a result, the expression vector pET21-CYM-Alc was obtained.

***E. coli* transformation and recombinant protein production.** For obtaining the target protein, the chemical transformation of *E. coli* strain SHuffle express was carried out with the resulting construct. Individual *E. coli* colonies containing recombinant plasmids were cultured overnight on an orbital shaker (Biosan, Latvia) in LB medium containing 100 μg/ml ampicillin at 37 °C and 180 rpm. The inoculum in a ratio of 1/100 was transferred to an Erlenmeyer flask containing LB medium and grown at 37 °C and 180 rpm. After the optical density (OD600) reached a value of 0.8, IPTG (Anatrace Products, USA) was added to the mixture to a final concentration of 1.0 mM. The culture was additionally incubated on a shaker for 12 h at 25 °C and 180 rpm. The biomass was centrifuged for 20 min at 5000 g and 4 °C to precipitate the inclusion bodies. *E. coli* cells were then resuspended in STET buffer (AppliChem, USA) (8.0 % sucrose; 50 mM Tris-HCl; 20 mM EDTA; 5.0 % (w/v) Triton X-100, pH 8.0) in proportion of 20 ml per 1 gram of biomass and incubated overnight at 4 °C. Thereafter cells were destroyed using a Soniprep 150 Plus ultrasonic homogenizer (MSE, PRC). Inclusion bodies were precipitated by centrifugation at 20,000 g for 20 min at 4 °C (Wei et al., 1999). The sedimented inclusion bodies were solubilized in buffer A (50 mM KH_2PO_4, 150 mM NaCl, pH 10.7) containing 8.0 M urea, incubated for 24 h at 15 °C and centrifuged at 20,000 g for 20 min.

Further work was carried out with a supernatant containing recombinant ProChn (rProChn). The target protein was renaturated according to the method of Wei et al. (1999). The supernatant was diluted 3× with buffer A and incubated for 12 h at 15 °C. Following the incubation, the supernatant diluted with alkaline buffer was adjusted to pH 8.0 with 1.0 M HCl, kept at 15 °C for 1 h, and dialyzed against buffer B (50 mM Tris, 150 mM NaCl, pH 8.0) overnight at 4 °C (Wei et al., 1999). As a result, an experimental preparation of moose rProChn was obtained.

The recombinant protein production in *E. coli* cells was analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) according to the Laemmli method (Laemmli, 1970). To analyze the electrophoretic mobility of the protein and determine molecular weights, the molecular weight markers PageRuler Unstained Protein Ladder (Thermo Scientific, USA) were used. Protein concentration was measured by the Bradford method (Bradford, 1976).

Activation of moose recombinant prochymosin. Activation of the moose rProChn was carried out by a stepwise changing of pH (Belenkaya et al., 2020b). To avoid autocatalytic conversion (Pedersen et al., 1979) of zymogen to an active enzyme after the isolation from inclusion bodies and partial purification, the preparations of moose rProChn were stored in weakly alkaline conditions (buffer B) prior to activation. For activation, HCl (2.0 M) was added to the rProChn sample, adjusting pH to 3.0 with continuous stirring. Then stirring was stopped, and the mixture was incubated at pH 3.0 for 2 h. After incubation, pH of the sample was adjusted to 5.8 using 0.5 M NaOH. As a result, the rChn of moose (rChn-Alc) was obtained.

Comparison preparations of commercial reference coagulants. Biochemical properties of rChn-Alc were compared with the properties of commercial reference coagulants: bovine rChn (rChn-Bos) (granular dry form with declared MA – 2201 IMCU/g) and commercial camel rChn (rChn-Cam) (liquid form with declared MA – 1000 IMCU/ml), produced by "Chr. Hansen" (Denmark).

Three-dimensional structure modeling and imaging. The homology model of the moose Chn three-dimensional structure was built on Swiss-model server (Waterhouse et al., 2018). The bovine Chn structure was used as a template for the modeling (Jensen et al., 2013). The images of surface charges were built with Coulombic Surface Coloring function of Chimera 1.14 software package.

Milk-clotting activity assays. A 10.0 % solution of standardized skimmed milk powder (MZSF OJSC, Russia) in 5 mM $CaCl_2$, pH 6.5, was used as a substrate. A 0.5 % aqueous solution of a dry bovine rChn with a certified MA value was used as a control. Prior to determining the MA, the control sample and the liquid preparation of rChn were kept in a water bath at 35 °C for 15 min and cooled to room temperature. The procedure for determining MA was carried out in a water bath at 35 °C. Substrate solution (2.5 ml) was placed into a glass tube and heated at 35 °C for 5 min. An aliquot (0.25 ml) of an enzyme was added to the substrate, a stopwatch was turned on, and the resulting reaction mixture was immediately thoroughly mixed. The time when the first flakes of the coagulated substrate were observed in the drops of reaction mixture applied onto the tube wall was considered to be the clotting time. The milk-clotting activity was expressed in arbitrary units (AU) per 1 ml (AU/ml) and calculated using the equation:

$$MA = 0.005 \cdot A \cdot T_1/T_2,$$

where A – certified MA value of the control rChn sample in AU per 1 gram, 0.005 – the dilution factor, T_1 – coagulation time for the control rChn sample of chymosin, T_2 – coagulation time for the test rChn sample.

Determination of total MA in each sample was performed in triplicate ($n = 3$). The relative MA of rChn-Alc was calculated after determining the total MA and protein concentration. For determining the relative MA of commercial reference chymosins, a 1.0 % aqueous solution of rChn-Bos was prepared, and the liquid rChn-Cam was diluted 10 times with distilled water. Protein concentration was determined in the resulting solutions with the Bradford assay. The MA values declared by the manufacturer were used to calculate the relative MA of the commercial reference enzymes. The relative MA was expressed in AU per milligram (AU/mg). To convert IMCU

(International Milk Clotting Units) values into AU, a multiplication coefficient of 125 was used.

General proteolytic activity assays. A 1.0 % solution of Hammerstein-grade casein in a 20 mM Na-phosphate buffer (pH 5.6) was used as a substrate. The investigated MEs were introduced into the substrate solution in a 1:4 ratio and incubated at 35 °C for 0 ('zero' point), 30, 90, and 180 min. The reaction was stopped by adding trichloroacetic acid. The precipitates were filtered, and the OD of the filtrate was measured at 280 nm (OD280) with a 'zero' point as a control. To assess the specificity of the rChn preparations, the OD280 values of the samples incubated for 180 min were designated as the PA values. The specificity was defined as the ratio of MA to general PA (MA/PA). When calculating the specificity of rChn-Bos and rChn-Cam, the MA values stated by the manufacturer were used. The enzymes studied were normalized by MA.

Thermal stability assays. Aliquots of ME were heated in the temperature range of 30–60 °C for 30 min and then assessed for residual MA. The MA values obtained in the samples heated at 30 °C were assigned as 100 %. The enzymes studied were normalized by MA.

Dependence of rennet coagulation time on pH. Solutions (10.0 %) of standardized skimmed milk (SSM) were adjusted to pH levels of 6.0, 6.2, 6.4, 6.6, 6.8 and 7.0, and the rennet coagulation time (RCT) of the studied preparations of rChns was then determined. The RCT values at a pH of 6.0 was assigned as 100 %. The enzymes studied were normalized by MA.

Dependence of rennet coagulation time on the calcium chloride concentration. Dry powder of $CaCl_2$ was added to the SSM to a final concentration of 1–5 mM, and the clot formation time was measured therein. The RCT values obtained in $CaCl_2$-free samples of SSM were taken as 1.0. The enzymes studied were normalized by MA.

Results

Expression of recombinant moose prochymosin. We used the *E. coli* strain SHuffle express to produce moose prochymosin in the laboratory to study its biochemical properties since *E. coli* is the most studied system for the expression of heterologous genes; functionally active chymosins of a number of mammals have already been obtained in this system (Rogelj et al., 2001; Belenkaya et al., 2020a, b), and despite the presence of drawbacks it allows obtaining samples of recombinant proteins in quantities sufficient for primary biochemical analysis in a short time. To obtain a producer strain, the designed nucleotide sequence of the moose ProChn gene, 1095 bp in size, was synthesized and cloned as part of the pET21a plasmid vector. The production and purification of the target protein were carried out as described previously (Belenkaya et al., 2018, 2020a). In order to evaluate the efficiency of the synthesis of rProChn of moose, as well as to determine its localization in the *E. coli* cell, an electrophoretic analysis of protein preparations obtained from the cells of the producer strain was carried out (Fig. 1).

Analysis of *E. coli* cells containing the pET21-CYM-Alc plasmid after induction with IPTG showed a high protein content, which coincides with the calculated one for rProChn of moose in terms of electrophoretic mobility (41 kDa). Its content was ≥30 % (see Fig. 1, lane *1*) of the total amount of cell proteins. It can be seen that the soluble fraction of *E. coli* biomass after treatment with STET buffer and centrifugation (lane *3*) contains almost no target protein, while the fraction of inclusion bodies is nearly completely represented by rProChn of moose (lane *4*).

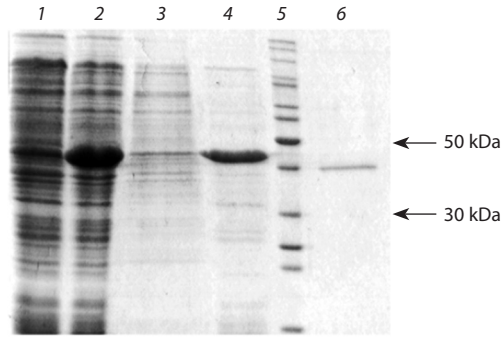


Fig. 1. SDS-PAGE analysis of protein samples obtained from the producer strain cells:

1 – producer cell biomass before IPTG adding; *2* – producer cell biomass after 5 h IPTG adding; *3* – soluble biomass fraction after treatment with STET buffer; *4* – insoluble fraction (inclusion bodies) after treatment with lysis buffer; *5* – molecular-weight markers (200, 150, 120, 100, 85, 70, 60, 50, 40, 30, 25, 20 and 15 kDa); *6* – moose rChn obtained as a result of zymogen activation.

Activation of rProChn and obtaining of rChn-Alc. The initial MA of the rProChn was <1.0 AU/ml. After activation, MA was equal to 843 AU/ml. Thus, as a result of activation, the total MA of the preparation increased more than 840 times, indicating the efficiency of the conversion of rProChn into active rChn of moose. In this case, a propetide is cleaved from the N-terminus of prochymosin molecule, resulting in a change in the length of the protein in the polyacrylamide gel, which is recorded using SDS-PAGE analysis (see Fig. 1, lane *6*).

Three-dimensional structure and surface charges of chymosin. Analyses of Chn sequences from different mammals demonstrated that the moose Chn is close to the other ones, especially the bovine Chn. Amino acid sequences of the bovine and moose Chn share 93.5 % identity, differing in 21 out of 323 positions. Three-dimensional structures of proteins with such a high similarity level are expected to be very close. In comparison, the camel and bovine Chn have 83.3 % identity, and their structures are similar. Therefore, we built a homology model of the moose Chn and used it for analyzing the surface charges.

Previous studies of camel and bovine chymosins revealed three positively charged patches on their surfaces that can contribute to the enzyme-substrate interaction (Jensen et al., 2013). Patch 1 and patch 3 are identical in the bovine and moose Chn, and patch 2 in the moose Chn has the same total charge as patch 2 in the bovine Chn, but charge distributions are different. Also, an additional charged patch in the moose Chn can be seen, designated as patch 4 (Fig. 2).

Technological properties

Specific MA. Milk-clotting activity is a basic technological characteristic of any new rChn since it indicates its ability to hydrolyze the Chn-sensitive peptide bond in the kappa-casein molecule and cause milk coagulation. Specialists in cheesemaking are aware of the paradox "cow Chn – camel milk and

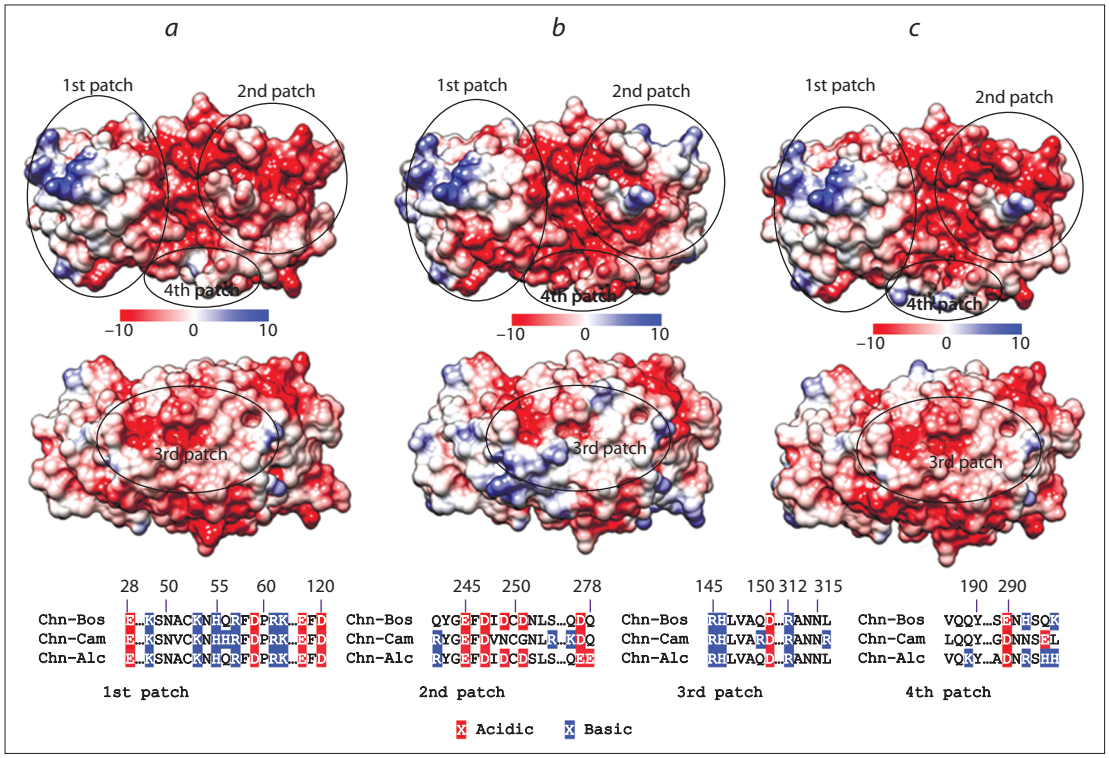


Fig. 2. Surface charged patches on chymosin: bovine (*a*), camel (*b*), moose (*c*).

Molecular surfaces are colored with Chimera software by the potential values in kcal/mole at 298 K. All proteins oriented with the C-terminal domain to the left and the N-terminal domain to the right, looking into the binding cleft (top) and rotated 180 degrees around the horizontal direction (bottom). The sequences of the chymosin charged patches are aligned, charged residues are highlighted.

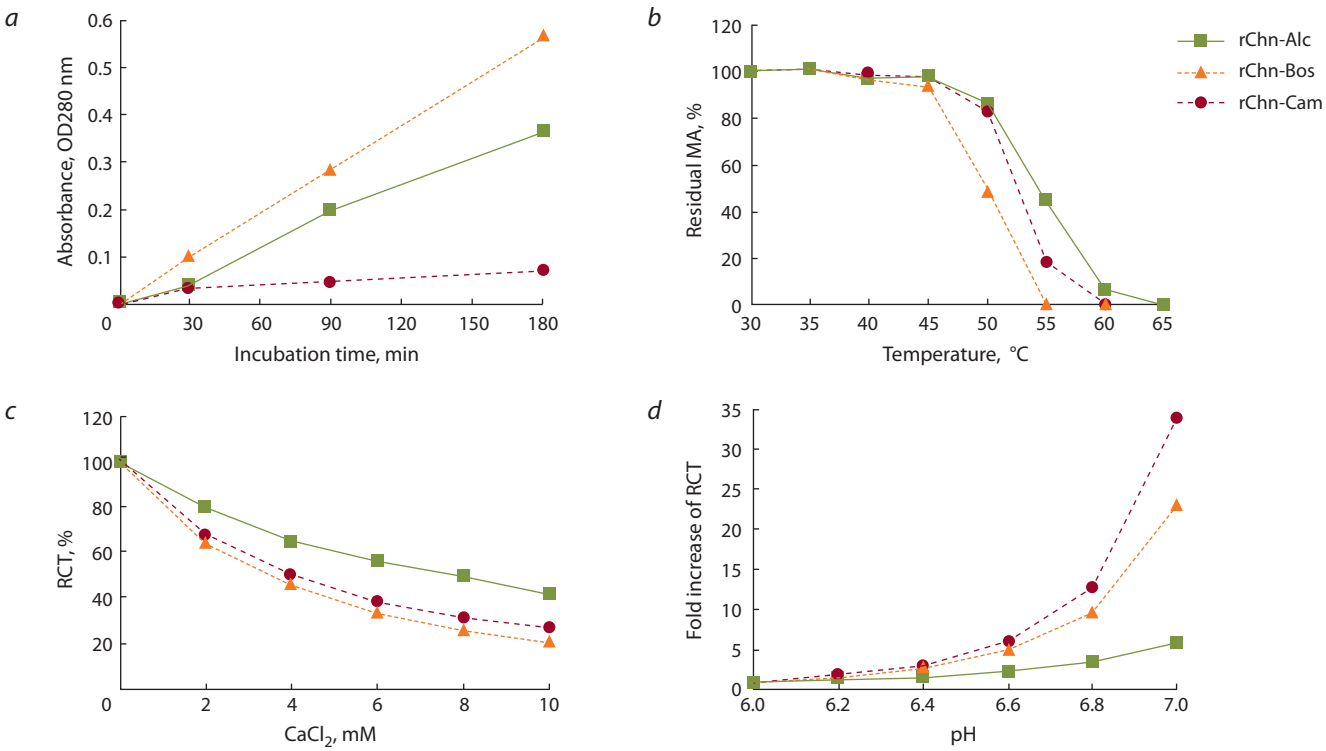


Fig. 3. Results of a comparative study of the following dependency patterns: general proteolytic activity (OD280) on the incubation time (*a*), residual MA (%) on the heating temperature (*b*), RCT on the calcium chloride concentration (*c*) and pH (*d*).

Table 1. Total and specific milk-clotting activity of recombinant chymosins

Sample	Total MA, AU/ml	Protein concentration, mg/ml	Specific MA, AU/mg	Specific MA, %
rChn-Alc	843 ± 14	0.027 ± 0.002	31 197 ± 526	37
rChn-Bos	2751	0.033 ± 0.005	85 323 ± 12 928	100
rChn-Cam	125 000	0.928 ± 0.029	136 944 ± 2099	161

camel Chn – cow milk". The paradox lies in the inability of the cow Chn to coagulate camel milk, while the camel enzyme effectively coagulates cow milk (Kappeler et al., 2006). Therefore, the study of any new ME for cheese-making begins with determining its MA in relation to cow's milk as the main raw material for cheese production. Only when one is sure that the new enzyme is capable of coagulating cow's milk is it reasonable to start investigating its other technological properties. Since the cow rChn and the camel rChn can be considered reference MEs for cheese-making, it is advisable to compare the biochemical properties of the new milk coagulant with them in order to assess its technological prospects.

In terms of specific MA, the moose rChn was inferior to the reference MEs – the cow and camel rChns – by 2.7 and 4.4 times, respectively (Table 1). The specific MA of rChn-Cam is 1.61 times higher than the specific coagulation activity of rChn-Bos. This is in good agreement with the data of (Kappeler et al., 2006; Belenkaya et al., 2020c), where it was shown that the ratio of the specific MA of rChn-Bos to the rChn-Cam is 1:1.7.

General proteolytic activity and specificity. To predict the technological prospects of any new ME, it is necessary to study its general PA. Excessive PA of milk coagulant is considered a negative factor in cheese production since it leads to a decrease in the yield and deterioration of the organoleptic properties of the cheeses produced (Singh et al., 2003; Harboe et al., 2010).

Conventionally, the PA of milk coagulants can be divided into specific and non-specific. The specific or milk-clotting activity of ME provides the hydrolysis of the F105-M106 bond in the κ-casein molecule, causing the destabilization of casein micelles and leading to the formation of a milk clot. Non-specific or general PA characterizes the ability of ME to hydrolyze any peptide bonds, with the exception of the F105-M106 bond of κ-casein. The ideal milk coagulant for cheese-making should exhibit the maximal MA with the minimal general PA (Harboe et al., 2010). The ratio of MA to general PA (MA/PA) is called specificity. The higher the value, the more versatile the ME, and the wider the range of cheeses to be produced.

The dynamics of accumulation of milk substrate proteolysis products under the action of rChn-Alc is similar to rChn-Bos and differs markedly from rChn-Cam (Fig. 3, *a*). These differences are most clearly observed after 90 min of incubation. After 180 min of incubation of the enzyme-substrate mixture, the general PA values (expressed in OD280 units) of the rChns of moose, cow, and camel were 0.362 ± 0.023, 0.565 ± 0.020, and 0.072 ± 0.012, respectively. As expected, rChn-Cam showed an exceptionally low level of non-specific proteolysis, which is in good agreement with the data (Bansal et al., 2009), according to which the PA of rChn-Cam was 4 times lower than that of rChn-Bos. Apparently, low values of the general PA are typical for the rChns of representatives of the Camelidae family. According to our data, the general PA of another member of this family, the alpaca, is about 3 times lower than that of the cow (Belenkaya et al., 2018).

Table 2. Specific MA, general PA, and specificity of recombinant moose, cow, and camel chymosins

Sample	Specific MA, %	General PA, %	Specificity, MA/PA
rChn-Alc	37	64	0.6
rChn-Bos	100	100	1.0
rChn-Cam	65	13	5.0

If we take the general PA of rChn-Bos as 100 %, then the PA of rChn-Alc and rChn-Cam will be 64 and 13 %, respectively. Using data on specific MA and general PA, specificity can be calculated (Table 2).

By specificity, and hence by the degree of cheese-making universality, the studied enzymes were arranged in the following order: rChn-Cam > rChn-Bos > rChn-Alc. The ratio of MA/PA calculated for the moose rChn was 1.6 and 8.3 times lower than for the cow and camel rChns. It is possible that the low specificity of rChn-Alc is a consequence of its low specific MA, which, as we have already noted, may be due to incomplete refolding of its zymogen. On average, the efficiency of restoring the correct folding (in terms of MA) of genetically engineered chymosins obtained after solubilization of inclusion bodies rarely exceeds 30 % (Wei et al., 1999, 2000; Chen et al., 2000; Eskandari et al., 2012).

Thermostability. Milk coagulants with a high threshold of thermal inactivation may show undesirable PA at the stages of cheese production associated with an increase in the heating temperature of the clot, as well as during prolonged maturation and storage of finished products. Therefore, the TS is an important technological characteristic of any new ME that claims to be used in cheese-making.

The proteolytic activity of MEs is registered in various types of cheeses (Masotti et al., 2010; Sforza et al., 2012; Gumus, Hayaloglu, 2019; Lamichhane et al., 2019; D'Incecco et al., 2020; Mane, McSweeney, 2020) and makes a significant contribution to the "proteolytic maturation" of the product. Information about the TS of the milk coagulant used allows one to regulate the degree of proteolysis and influence the maturation time of cheeses by varying the processing temperature of the cheese grain or by using ME with different thermal activation thresholds (Lamichhane et al., 2019).

It was found that the rChn of a camel is more thermally stable than that of a cow (Kappeler et al., 2006; Jensen et

al., 2013; Belenkaya et al., 2020c). It is also known that the general PA of these enzymes increases with increasing temperature – rChn-Bos shows the maximum PA at 55.0 °C, and rChn-Bos at 52.5 °C (Kappeler et al., 2006). An increase in the heating temperature of the clot from 50 to 56 °C when producing very hard, granular, cows' milk cheese using the rChn of a cow or a camel leads to a significant decrease in the concentration of products of proteolysis of αS1-casein. But even after processing the clot at 56 °C, the concentration of markers of proteolysis of αS1-casein was higher in maturing and stored cheeses produced using a more thermally stable rChn-Cam than when using rChn-Bos (Costabel et al., 2015). This is despite the fact that the general PA of a camel enzyme is 3.5–4.0 times lower than that of a cow (Kappeler et al., 2006).

The ranges of TS of the rChns of the same species obtained in different expression systems may differ. The thresholds of total temperature inactivation of the camel rChn expressed in higher mold fungi (*Aspergillus niger*) and yeast (*Komagataella* (*Pichia*) *pastoris*) differed by 10 °C (Belenkaya et al., 2020c). The experimental rChn-Bos synthesized in the *E. coli* BL21(DE3) system exceeded the commercial rChn-Bos expressed in *A. niger* by 15 °C (Belenkaya et al., 2018). These data indicate a possible role of posttranslational modifications as a factor influencing the temperature stability of rChns.

The threshold of thermal inactivation was considered the T (°C) at which the studied rChn lost >20 % of the initial coagulation activity. According to this criterion, the TS threshold for rChn-Bos was 50 °C, and for rChn-Alc and rChn-Cam – 55 °C (see Fig. 3, *b*). After 30 min of heating up at 55 °C, rChn-Bos was completely inactivated. Despite the same TS threshold, rChn-Alc and rChn-Cam differed in the dynamics of thermal inactivation in the temperature range of 50–65 °C. After heating up to 55 °C, the residual coagulation activity of the moose rChn was almost 2.5 times higher than that of rChn-Cam and amounted to 44.9 and 18.2 %, respectively. The recombinant camel rChn was completely inactivated after heating at 60 °C, while rChn-Alc still retained 6.5 % of the original MA at this temperature, suggesting higher temperature stability of rChn-Alc compared to rChn-Cam.

Thus, taking into account the same threshold of thermal inactivation of rChn-Alc and rChn-Cam, according to the TS criterion, the studied enzymes are arranged as follows: rChn-Alc > rChn-Cam > rChn-Bos. Increased, in comparison with reference enzymes, TS limits the scope of application of the moose rChn assumes its use, first of all, in the production of cheeses with short maturation and storage periods.

Dependence of rennet coagulation time on the calcium chloride concentration. Most rennet cheeses are made from pasteurized milk. It is known that during pasteurization, denatured β-lactoglobulin binds to micellar κ-casein, which leads to an increase in the duration of RCT (Fox et al., 2017). In addition, during high-temperature processing of raw milk, part of the salts and calcium ions present in it precipitates in the form of insoluble calcium phosphate. As a result, the concentration of Ca^{2+} in milk decreases, which also increases the RCT. In order to avoid increasing the dose of introduced ME and improve the coagulation ability of pasteurized milk, $CaCl_2$ is added to it in an amount of 0.1–0.4 g/l (~1–4 mM). However, an increase of the $CaCl_2$ concentration in the milk substrate causes not only an increase in the coagulation activity but also in the general PA of the enzyme, especially at the stage of milk coagulation (Wang et al., 2015). Therefore, the use of ME with high sensitivity to Ca^{2+} concentration is associated with the risk of negative consequences of increasing its general PA. Based on this, it is necessary that the new milk coagulant, in comparison with modern reference technological enzymes, has a comparable or lower sensitivity to changes in the concentration of $CaCl_2$ in the milk substrate.

Just as in the case of other MEs (Fox et al., 2017), the duration of RCT under the action of the studied rChns decreased in response to an increase in the concentration of calcium chloride. In the range of 0–10 mM of $CaCl_2$ clot formation time is reduced by 0–58 % for rChn-Alc, 0–79 % for rChn-Bos, and 0–73 % for rChn-Cam. The dynamics of changes in the dependence of RCT on the concentration of calcium chloride for the cow and camel rChns is almost the same (see Fig. 3, *c*). Recombinant Chn of moose differs from reference enzymes – its coagulation activity is less sensitive to changes in the concentration of $CaCl_2$ in the milk substrate. At 4 mM $CaCl_2$, the RCT of the milk substrate decreases by 2.0 and 2.2 times, respectively, under the action of rChn-Bos and rChn-Cam, and only by 1.5 times for rChn-Alc. This, in particular, means that the risk of an increase in the general PA when using the moose rChn to curdle pasteurized milk with added $CaCl_2$ is much less than that of reference coagulants, which is a positive factor from the point of view of cheese production.

Thus, we suppose that the sensitivity of coagulation activity of rChn-Alc to an increase in the concentration of $CaCl_2$ in milk fully meets the requirements of cheese production.

Dependence of rennet coagulation time on pH. The optimums of the specific activity of various types of chymosins lie in the pH range of 4.6–6.0 (Belenkaya et al., 2020c). However, in the production of most types of rennet cheeses, ME is added to the milk mixture at a pH of 6.5–6.6. Therefore, one of the technological requirements for any new coagulant is its ability to effectively curdle milk in a slightly acidic pH range that is far from the pH optimum.

The duration of RCT depends on the electrostatic and hydrophobic properties of casein micelles, which are related to the H^+ concentration. When milk is acidified, the total negative charge of caseins decreases due to the pH approaching the pI values. This reduces the forces of electrostatic repulsion between the micelles and simultaneously increases the casein-casein hydrophobic interactions, which accelerates the formation of milk clot. If the pH increases, the casein-casein hydrophobic interactions weaken as the total negative charges of caseins increase. The growing forces of electrostatic repulsion prevent the convergence of similarly charged casein micelles and slow down the formation of milk clot (Lucey, 2002; Harboe et al., 2010; Fox et al., 2017).

By the nature of the dependence of the coagulation ability on pH, the most promising for cheese-making are MEs, which slowly lose activity when moving away from the pH-optimum to the alkaline region and can exhibit high MA at weakly acidic and neutral pH values.

Compared to the reference rChns, the coagulation activity of rChn-Alc is much less dependent on changes in milk pH from 6.0 to 7.0 (see Fig. 3, *d*). At a pH of 6.4–6.6, the RCT for rChn-Alc increases 1.6–2.3 times, and for rChn-Bos and rChn-Cam, this parameter increases 2.5–5.0 and 2.9–6.0 times,

respectively. At pH 7.0, the rChns of cow and camel showed extremely low (we can say, trace) coagulation activity, and the differences between them and the rChn of moose were most clearly manifested. Based on the data obtained, it can be argued that in the working "cheese-making" pH range of 6.5–6.6, the consumption of the moose rChn will be lower than that of the reference MEs, which is an important technological characteristic. However, the ability of rChn-Alc to show significant coagulation activity at neutral pH values is not unique. Previously, similar properties were found in the rChns of the goat (Vallejo et al., 2012) and yak (Ersöz, İnan, 2019).

Thus, the moose rChn is able to effectively curdle cow's milk at a pH of 6.5–6.6, and is not inferior in this indicator to commercial genetically engineered chymosins.

Discussion

For the first time, recombinant moose chymosin was obtained, and its characteristics, important for the production of rennet cheeses, were also investigated. We have chosen a prokaryotic expression system for preliminary characterization of the enzyme, because it is easier to work with and since it was known that other recombinant chymosins obtained in prokaryotes retain their activity (Eskandari et al., 2012; Belenkaya et al., 2020a, c). The conditions used for the expression of the moose chymosin gene in the *E. coli* system lead to a highly efficient synthesis of the target protein, with almost all of it accumulating in an insoluble form in inclusion bodies. As expected, MA rProChn was very low, and after activation MA rChn-Alc it was 843 AU/ml. According to the total MA, the moose rChn obtained by us was 2.4–2.8 times inferior to other genetically engineered rChns (2014 AU/ml for alpaca and 2330 AU/ml for maral) obtained in the *E. coli* expression system (Belenkaya et al., 2018, 2020a, b).

It is possible that the low specific MA of rChn-Alc, compared with the reference milk coagulants, is due to the insufficient efficiency of its zymogen refolding after isolation from the inclusion bodies. It is known that the stage of restoring the correct three-dimensional structure is a "bottleneck" in obtaining rChn in *E. coli* expression systems and leads to a decrease in the yield and specific activity of the target product (Wei et al., 1999, 2000; Chen et al., 2000; Eskandari et al., 2012).

Also, we cannot exclude the possibility that under our conditions the moose chymosin was not activated quite correctly, with the N-end cut off in a different position, thus affecting its activity. For example, loss of the first three residues of camel chymosin significantly decreased its activity (Jensen et al., 2013). We have so far characterized only the enzymatic properties and determined the approximate molecular weight using SDS-PAGE analysis, but we do not know the exact amino acid sequence of rChn-Alc.

Previously, it was suggested that the technological properties of the camel Chn depend on its surface charge distribution (Jensen et al., 2013). The total charge of κ-casein C-terminal part is negative, as is the total charge of all known chymosins. Positively charged patches on the chymosin surface can play a role in properly positioning and binding the enzyme to the substrate (Jensen et al., 2013). Most positively charged patches in the moose Chn are similar to those in the bovine Chn, but patch 2 has intermediate characteristics between the corresponding patches in the bovine and camel Chn. An additional patch 4 in the moose Chn is located close to the substrate-binding cleft (see Fig. 2). It is challenging to conclude whether the differences in positively charged patches in chymosins are stochastic or whether they result from adaptation to some conditions, such as species-specific variations in κ-caseins charge distributions. Further studies of the chymosins from different mammals may clarify this question.

The resulting preparation of rChn-Alc is able to coagulate cow's milk. In terms of specific MA, however, it is inferior to the reference commercial rChns of cow and camel. It means that in cheese-making the consumption of the rChn of moose, obtained in the *E. coli* expression system, will be higher than that of rChn-Bos and rChn-Cam. In order to compete with reference enzymes, the specific MA of the moose rChn should be increased 3–4 times. However, in a number of technological parameters, the moose rChn is superior to the reference commercial enzymes. Thus, in comparison with the rChn of a cow and a camel, the specific activity of the rChn of moose is less dependent on changes in the concentration of $CaCl_2$ in the range of 1–5 mM and pH in the range of 6–7, which is an attractive technological property.

In general, though obtained in the prokaryotic system, the moose chymosin meets the basic requirements for enzymes for cheese-making, encouraging us to study this protein. The main problem of yeast expression systems is a strong ability to glycosylate proteins. *Pichia* may have an advantage in the glycosylation of secreted proteins over *Saccharomyces cerevisiae* because the former does create proteins with long carbohydrate chains via hyperglycosylation (Akishev et al., 2021). In an experiment to obtain recombinant camel chymosin, the prochymosin gene was successfully cloned and expressed in *P. pastoris* under the control of the GAP promoter and purified from culture via a combination of cation and anion exchange chromatography. *Camelus bactrianus* recombinant chymosin manifested high milk-clotting activity (9605 U/mg) (Akishev et al., 2021). One of the priority tasks is to obtain the moose rChn in the eukaryotic expression system and to compare its technological properties (primarily specific MA) with the properties of the enzyme produced in the *E. coli* expression system.

Conclusion

The nucleotide sequence encoding moose (*Alces alces*) prochymosin was optimized for its efficient expression in *E. coli* cells of the SHuffle express strain. The synthesized prochymosin gene was cloned into the pET21a vector, resulting in the pET21-CYM-Alc expression vector. The constructed model of the spatial structure of the moose Chn showed that the ionic charges on the surface of the protein molecule are distributed similarly to those for the cow and camel Chn, but the moose enzyme has a unique charged site, which probably affects its MA.

A sample of moose rProChn was developed and its biochemical and technological properties were studied. In some of the technological parameters, it surpasses the reference commercial enzymes. Thus, the specific activity of the moose rChn is less dependent on changes in $CaCl_2$ concentration in the range of 1–5 mM and substrate pH in the range of 6–7, compared to the cow and camel rChn. The total proteolytic activity of the moose rChn occupies an intermediate position between

the cow and camel rChn. In terms of such indicators as specific milk-clotting activity, specificity and thermal stability, the mooserChn is inferior to reference commercial chymosins.

References

Akishev Z., Kiribayeva A., Mussakhmetov A., Baltin K., Ramankulov Y., Khassenov B. Constitutive expression of *Camelus bactrianus* prochymosin B in *Pichia pastoris*. *Heliyon.* 2021;7(5):e07137. DOI 10.1016/j.heliyon.2021.e07137.

Bansal N., Drake M.A., Piraino P., Broe M.L., Harboe M., Fox P.F., McSweeney P.L.H. Suitability of recombinant camel (*Camelus dromedarius*) chymosin as a coagulant for Cheddar cheese. *Int. Dairy J.* 2009;19(9):510-517. DOI 10.1016/j.idairyj.2009.03.010.

Belenkaya S.V., Balabova D.V., Belov A.N., Koval A.D., Shcherbakov D.N., Elchaninov V.V. Basic biochemical properties of recombinant chymosins (review). *Appl. Biochem. Microbiol.* 2020a;56(4):363-372. DOI 10.1134/S0003683820040031.

Belenkaya S.V., Bondar A.A., Kurgina T.A., Elchaninov V.V., Bakulina A.Yu., Rukhlova E.A., Lavrik O.I., Ilyichev A.A., Shcherbakov D.N. Characterization of the Altai maral chymosin gene, production of a chymosin recombinant analog in the prokaryotic expression system, and analysis of its several biochemical properties. *Biochemistry* (*Mosc*.). 2020b;85(7):781-791. DOI 10.1134/S0006297920070068.

Belenkaya S.V., Rudometov A.P., Shcherbakov D.N., Balabova D.V., Kriger A.V., Belov A.N., Koval A.D., Elchaninov V.V. Biochemical properties of recombinant chymosin in alpaca (*Vicugna pacos* L.). *Appl. Biochem. Microbiol.* 2018;54(6):569-576. DOI 10.1134/S0003683818060054.

Belenkaya S.V., Shcherbakov D.N., Balabova D.V., Belov A.N., Koval A.D., Elchaninov V.V. Production of maral (*Cervus elaphus sibiricus* Severtzov) recombinant chymosin in the prokaryotic expression system and the study of the aggregate of its biochemical properties relevant for the cheese-making industry. *Appl. Biochem. Microbiol.* 2020c;56(6):647-656. DOI 10.1134/S0003683820060034.

Belov A.N., Koval A.D., Avdanina E.A., Elchaninov V.V. Effects of milk clotting preparation on cheese organoleptic indices. *Syrodelie i Maslodelie = Cheese Making and Butter Making*. 2009;1:22-24. (in Russian)

Bradford M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem*. 1976;72(1-2):248-254. DOI 10.1016/0003-2697(76)90527-3.

Chen H., Zhang G., Zhang Y., Dong Y., Yang K. Functional implications of disulfide bond, Cys206−Cys210, in recombinant prochymosin (chymosin). *Biochemistry.* 2000;39(40):12140-12148. DOI 10.1021/bi000976o.

Costabel L.M., Bergamini C.V., Pozza L., Cuffia F., Candioti M.C., Hynes E. Influence of chymosin type and curd scalding temperature on proteolysis of hard cooked cheeses. *J. Dairy Res.* 2015;82(3):375-384. DOI 10.1017/S0022029915000175.

D'Incecco P., Limbo S., Hogenboom J., Rosi V., Gobbi S., Pellegrino L. Impact of extending hard-cheese ripening: a multiparameter characterization of Parmigiano Reggiano cheese ripened up to 50 months. *Foods*. 2020;9(3):268. DOI 10.3390/foods9030268.

Ersöz F., İnan M. Large-scale production of yak (*Bos grunniens*) chymosin A in *Pichia pastoris*. *Protein Expr. Purif*. 2019;154:126-133. DOI 10.1016/j.pep.2018.10.007.

Eskandari M.H., Hosseini A., Zarasvand S.A., Aminlari M. Cloning, expression, purification and refolding of caprine prochymosin. *Food Biotechnol.* 2012;26(2):143-153. DOI 10.1080/08905436.2012.670829.

Flamm E.L. How FDA approved chymosin: a case history. *Nat. Biotechnol*. 1991;9(4):349-351. DOI 10.1038/nbt0491-349.

Fox P.F., Guinee T.P., Cogan T.M., McSweeney P.L.H. Enzymatic coagulation of milk. In: Fundamentals of Cheese Science. Boston: Springer, 2017;185-229. DOI 10.1007/978-1-4899-7681-9_7.

Gumus P., Hayaloglu A.A. Effects of blends of camel and calf chymosin on proteolysis, residual coagulant activity, microstructure, and sensory characteristics of Beyaz peynir. *J. Dairy Sci.* 2019;102(7):5945-5956. DOI 10.3168/jds.2018-15671.

Harboe M., Broe M.L., Qvist K.B. The production, action and application of rennet and coagulants. In: Law B.A., Tamime A.Y. (Eds.) Technology of Cheesemaking. Chichester: Wiley-Blackwell, 2010;98-129. DOI 10.1002/9781444323740.ch3.

Jacob M., Jaros D., Rohm H. Recent advances in milk clotting enzymes. *Int. J. Dairy Tech.* 2011;64(1):14-33. DOI 10.1111/j.1471-0307.2010.00633.x.

Jensen J.L., Molgaard A., Poulsen J.-C.N., Harboe M.K., Simonsen J.B., Lorentzen A.M., Hjernø K., Brink J.M., Qvist K.B., Larsen S. Camel and bovine chymosin: the relationship between their structures and cheese-making properties. *Acta Crystallogr. D. Biol. Crystallogr*. 2013;69(Pt.5):901-913. DOI 10.1107/S0907444913003260.

Kappeler S.R., van den Brink H.J.M., Rahbek-Nielsen H., Farah Z., Puhan Z., Hansen E.B., Johansen E. Characterization of recombinant camel chymosin reveals superior properties for the coagulation of bovine and camel milk. *Biochem. Biophys. Res. Commun.* 2006;342(2):647-654. DOI 10.1016/j.bbrc.2006.02.014.

Laemmli U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* 1970;227(5259):680-685. DOI 10.1038/227680a0.

Lamichhane P., Sharma P., Kennedy D., Kelly A.L., Sheehan J.J. Microstructure and fracture properties of semi-hard cheese: differentiating the effects of primary proteolysis and calcium solubilization. *Food Res. Int.* 2019;125:108525. DOI 10.1016/j.foodres.2019.108525.

Lucey J.A. Formation and physical properties of milk protein gels. *J. Dairy Sci.* 2002;85(2):281-294. DOI 10.3168/JDS.S0022-0302(02)74078-2.

Luo F., Jiang W.H., Yang Y.X., Li J., Jiang M.F. Cloning and expression of yak active chymosin in *Pichia pastoris*. *Asian-Australas. J. Anim. Sci.* 2016;29(9):1363-1370. DOI 10.5713/AJAS.16.0038.

Mane A., McSweeney P.L.H. Proteolysis in Irish farmhouse Camembert cheese during ripening. *J. Food Biochem.* 2020;44(1):1-14. DOI 10.1111/jfbc.13101.

Masotti F., Hogenboom J.A., Rosi V., de Noni I., Pellegrino L. Proteolysis indices related to cheese ripening and typicalness in PDO Grana Padano cheese. *Int. Dairy J.* 2010;20(5):352-359. DOI 10.1016/j.idairyj.2009.11.020.

Moynihan A.C., Govindasamy-Lucey S., Jaeggi J.J., Johnson M.E., Lucey J.A., McSweeney P.L.H. Effect of camel chymosin on the texture, functionality, and sensory properties of low-moisture, part-skim Mozzarella cheese. *J. Dairy Sci*. 2014;97(1):85-96. DOI 10.3168/jds.2013-7081.

Pedersen V.B., Christensen K.A., Foltmann B. Investigations on the activation of bovine prochymosin. *Eur. J. Biochem.* 1979;94(2):573-580. DOI 10.1111/j.1432-1033.1979.tb12927.x.

Rogelj I., Perko B., Francky A., Penca V., Pungerčar J. Recombinant lamb chymosin as an alternative coagulating enzyme in cheese production. *J. Dairy Sci.* 2001;84(5):1020-1026. DOI 10.3168/jds.S0022-0302(01)74561-4.

Sforza S., Cavatorta V., Lambertini F., Galaverna G., Dossena A., Marchelli R. Cheese peptidomics: a detailed study on the evolution of the oligopeptide fraction in Parmigiano-Reggiano cheese from curd to 24 months of aging. *J. Dairy Sci.* 2012;95(7):3514-3526. DOI 10.3168/jds.2011-5046.

Singh T.K., Drake M.A., Cadwallader K.R. Flavor of Cheddar cheese: a chemical and sensory perspective. *Compr. Rev. Food Sci. Food Saf.* 2003;2(4):166-189. DOI 10.1111/j.1541-4337.2003.tb00021.x.

Trono D. Recombinant enzymes in the food and pharmaceutical industries. In: Singh R.S., Singhania R.R., Pandey A., Larroche C. (Eds.) Biomass, Biofuels, Biochemicals: Advances in Enzyme Technology. Elsevier, 2019;349-387. DOI 10.1016/B978-0-444-64114-4.00013-3.

Tyagi A., Kumar A., Mohanty A.K., Kaushik J.K., Grover S., Batish V.K. Expression of buffalo chymosin in *Pichia pastoris* for application in mozzarella cheese. *LWT – Food Sci. Technol.* 2017;84:733-739. DOI 10.1016/j.lwt.2017.06.033.

Tyagi A., Kumar A., Yadav A.K., Chandola Saklani A., Grover S., Batish V.K. Functional expression of recombinant goat chymosin in *Pichia pastoris* bioreactor cultures: a commercially viable alternate. *LWT – Food Sci. Technol.* 2016;69:217-224. DOI 10.1016/j.lwt.2016.01.015.

Vallejo J.A., Ageitos J.M., Poza M., Villa T.G. Short communication: a comparative analysis of recombinant chymosins. *J. Dairy Sci.* 2012;95(2):609-613. DOI 10.3168/jds.2011-4445.

Vega-Hernández M.C., Gómez-Coello A., Villar J., Claverie-Martín F. Molecular cloning and expression in yeast of caprine prochymosin. *J. Biotechnol*. 2004;114(1-2):69-79. DOI 10.1016/j.jbiotec.2004.06.002.

Wang N., Wang K.Y., Li G., Guo W., Liu D. Expression and characterization of camel chymosin in *Pichia pastoris*. *Protein Expr. Purif.* 2015;111:75-81. DOI 10.1016/j.pep.2015.03.012.

Waterhouse A., Bertoni M., Bienert S., Studer G., Tauriello G., Gumienny R., Heer F.T., de Beer T.A.P., Rempfer C., Bordoli L., Lepore R., Schwede T. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 2018;46(W1):W296-W303. DOI 10.1093/nar/gky427.

Wei C., Tang B., Zhang Y., Yang K. Oxidative refolding of recombinant prochymosin. *Biochem. J.* 1999;340(1):345-351. DOI 10.1042/0264-6021:3400345.

Wei C., Zhang Y., Yang K. Chaperone-mediated refolding of recombinant prochymosin. *J. Protein Chem.* 2000;19(6):449-456. DOI 10.1023/A:1026593113633.

ORCID ID

D.V. Balabova orcid.org/0000-0002-3925-0862
A.P. Rudometov orcid.org/0000-0003-2808-4309
A.A. Bondar orcid.org/0000-0001-9181-0487
A.Yu. Bakulina orcid.org/0000-0002-4723-9225
V.V. Elchaninov orcid.org/0000-0002-4355-6055
D.N. Shcherbakov orcid.org/0000-0001-8023-4453

Acknowledgements. Sanger DNA sequencing was performed in Genomics Core Facility (ICBFM SB RAS, Novosibirsk).

This study was funded by the State Task of the Ministry of Science and Education of the Russian Federation (topic No. FZMW-2020-0002, "Design of recombinant enzyme producers for the cheese-making industry") and The Russian Foundation for Basic Research (grant number 19-44-220010).

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

Received November 17, 2021. Revised February 19, 2022. Accepted March 3, 2022.

Вавиловский журнал генетики и селекции. 2022;26(3):250-263
DOI 10.18699/VJGB-22-32

Original Russian text www.bionet.nsc.ru/vogis/

Starch metabolism in potato *Solanum tuberosum* L.

E.M. Sergeeva ✉, K.T. Larichev, E.A. Salina, A.V. Kochetov

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
✉ sergeeva@bionet.nsc.ru

Abstract. Starch is a major storage carbohydrate in plants. It is an important source of calories in the human and animal diet. Also, it is widely used in various industries. Native starch consists of water-insoluble semicrystalline granules formed by natural glucose polymers amylose and amylopectin. The physicochemical properties of starch are determined by the amylose:amylopectin ratio in the granule and degrees of their polymerization and phosphorylation. Potato *Solanum tuberosum* L. is one of the main starch-producing crops. Growing industrial needs necessitate the breeding of plant varieties with increased starch content and specified starch properties. This task demands detailed information on starch metabolism in the producing plant. It is a complex process, requiring the orchestrated work of many enzymes, transporter and targeting proteins, transcription factors, and other regulators. Two types of starch are recognized with regard to their biological functions. Transitory starch is synthesized in chloroplasts of photosynthetic organs and degraded in the absence of light, providing carbohydrates for cell needs. Storage starch is synthesized and stored in amyloplasts of storage organs: grains and tubers. The main enzymatic reactions of starch biosynthesis and degradation, as well as carbohydrate transport and metabolism, are well known in the case of transitory starch of the model plant *Arabidopsis thaliana*. Less is known about features of starch metabolism in storage organs, in particular, potato tubers. Several issues remain obscure: the roles of enzyme isoforms and different regulatory factors in tissues at various plant developmental stages and under different environmental conditions; alternative enzymatic processes; targeting and transport proteins. In this review, the key enzymatic reactions of plant carbohydrate metabolism, transitory and storage starch biosynthesis, and starch degradation are discussed, and features specific for potato are outlined. Attention is also paid to the known regulatory factors affecting starch metabolism.
Key words: potato; *Solanum tuberosum*; starch; amylose; amylopectin; synthesis; degradation.

For citation: Sergeeva E.M., Larichev K.T., Salina E.A., Kochetov A.V. Starch metabolism in potato *Solanum tuberosum* L. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2022;26(3):250-263. DOI 10.18699/VJGB-22-32

Метаболизм крахмала у картофеля *Solanum tuberosum* L.

Е.М. Сергеева ✉, К.Т. Ларичев, Е.А. Салина, А.В. Кочетов

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ sergeeva@bionet.nsc.ru

Аннотация. Крахмал – основной запасной углевод растений и важный источник калорий в рационе питания человека и животных, широко применяется также для ряда промышленных нужд. Нативный крахмал представляет собой не растворимые в воде полукристаллические гранулы, сформированные молекулами амилозы и амилопектина, которые являются природными полимерами глюкозы. Физико-химические свойства крахмала определяются соотношением амилозы и амилопектина в грануле, степенью их полимеризации и фосфорилирования. Одно из основных растений-продуцентов крахмала – картофель *Solanum tuberosum* L. Растущие потребности рынка диктуют необходимость получения сортов с повышенным содержанием и заданными свойствами крахмала, необходимым условием чего является получение детальной информации о процессе его метаболизма в организме растения. Процесс метаболизма крахмала сложен и представляет собой согласованную работу множества ферментов, транспортных и направляющих белков, транскрипционных и других регуляторных факторов. По принципу выполняемой биологической функции крахмал делится на два типа: транзиторный, который синтезируется в хлоропластах фотосинтезирующих органов и распадается при отсутствии освещения, обеспечивая клетку углеводами; и запасной, который синтезируется и хранится в амилопластах запасающих органов (зерен, клубней). Основные ферментативные реакции биосинтеза и деградации крахмала, а также транспорта и метаболизма углеводов хорошо изучены на транзиторном крахмале модельного объекта *Arabidopsis thaliana*; об особенностях метаболизма крахмала в запасающих органах, в частности клубнях картофеля, известно несколько меньше. При этом ряд вопросов остается открытым: недостаточно изучены участие различных изоформ ферментов и влияние на них

регуляторных факторов в зависимости от ткани и стадии развития растения, а также условий внешней среды; реализация альтернативных путей ферментативных превращений; участие направляющих и транспортных белков. В данном обзоре рассмотрены ключевые ферментативные реакции углеводного обмена, биосинтеза и деградации транзиторного и запасного крахмала, общие для растений, и отмечены особенности, характерные для картофеля; уделено внимание известным регуляторным факторам, влияющим на метаболизм крахмала.
Ключевые слова: картофель; *Solanum tuberosum*; крахмал; амилоза; амилопектин; синтез; деградация.

Introduction

Starch is the main storage carbohydrate in plants. It constitutes up to 85 % of the dry matter of their edible parts: cereal grains (maize *Zea mays* L., rice *Oryza sativa* L., wheat *Triticum* spp., barley *Hordeum vulgare* L., etc.), potato tubers *Solanum tuberosum* L., edible roots (cassava *Manihot esculenta* Crantz, sweet potato *Ipomoea batatas* (L.) Lam., and yam *Dioscorea alata* L.), sago palm stems *Metroxylon sagu* Rottb., plantain fruit *Musa* spp. (Zeeman et al., 2010; Santana, Meireles, 2014). Starch provides a great portion of calories for human and animal nutrition. In addition, it is a natural reproducible and biodegradable material used in non-food industry, e. g., in the production of fabric, paper, drugs, and plastics.

Chemically, starch is a mixture of amylose and amylopectin. These natural glucose polymers form water-insoluble semicrystalline granules. Amylopectin consists of highly branched glucan molecules, the linear regions of which are formed by α-1,4-glycosidic bonds, whereas the branching points are formed by α-1,6-bonds. Amylose is a practically linear polymer with few branches. Amylopectin molecules constitute about 75–80 % of starch weight. They form the structural framework of the granule, consisting of repetitive amorphous and semicrystalline lamellae. Amylose molecules are dispersed in the semicrystalline amylopectin matrix (Zeeman et al., 2010; Tetlow, Bertoft, 2020).

The amylose:amylopectin ratio determines the dietetic, physicochemical, and functional starch properties essential for particular industries. The starch present in food is classified into glycemic and resistant. Being readily digestible in the small intestine, amylopectin increases the glycemic potential of starch. In contrast, higher amylose contents make starch more resistant (Li et al., 2008). Resistant starch is less degradable by amylases in the small intestine. It serves as substrate for microbes in the large intestine to produce short-chain fatty acids, which exert local antiinflammatory and antitumor effects (Birt et al., 2013). Also, physicochemical (gelatinization and retrogradation) and functional (swelling and viscosity) properties are taken into consideration in certain applications. These properties are determined genetically: by size and morphology of starch granules, amylose:amylopectin ratio, glucan branching, and glucan phosphorylation (Visser et al., 1991; Schwall et al., 2000; Hofvander et al., 2004; Khlestkin et al., 2017).

Potato (*Solanum tuberosum* L.) ranks fourth among starch-producing crops in the world, next to maize, cassava, and wheat. Potato starch differs from cereal starches in a variety of important features. Potato amylose and amylopectin have higher degrees of polymerization and phosphorylation; therefore, potato starch is more suitable for bioplastic production (Hofvander et al., 2004; Reyniers et al., 2020). In response to the increasing commercial demand, the global production of potato starch steadily increases: 3.7 million tons in 2018 and 3.9 million tons in 2020 (https://www.researchandmarkets.com/reports/5330932/potato-starch-market-global-industry-trends). To obtain native starches with specified properties and to increase the overall amount of starch per plant are topical tasks in potato breeding.

The key enzymes in starch biosynthesis (see the Table) and their genes have been studied in detail in model plants (*Arabidopsis thaliana* L.) and in crops, including potato (Streb, Zeeman, 2012; Van Harsselaar et al., 2017; Slugina, Kochieva, 2018).

Starch is produced by the polymerization of ADP-glucose, catalyzed by granule-bound (GBSS) and soluble (SS) starch synthases. Other enzymes involved are the starch branching enzyme (SBE) and the debranching enzyme (DBE).

Starch metabolism is a component of general carbohydrate metabolism. It is essential for plant functions: growth, development, and stress response. The enzymes and genes associated with the metabolism of starch and other plant carbohydrates have been extensively studied for three decades. A considerable body of data on the location of these genes in plant genomes and their expression has been accumulated. Seventy-five *S. tuberosum* genes have been mapped on the reference genome, and the expression patterns of 64 genes in leaves and tubers have been studied (Van Harsselaar et al., 2017; Slugina, Kochieva, 2018). It has been shown that starch metabolism genes experienced numerous duplications and produced paralogs by sub- and neofunctionaliza-

Enzymes involved in starch metabolism

Enzyme	EC no.	Izozyme	Function
		Carbohydrate metabolism	
Sucrose synthase	EC:2.4.1.13	SuSy1, SuSy2, SuSy3, SuSy4, SuSy5, SuSy6, SuSy7	Reversible cleavage of sucrose to fructose and UDP-glucose (Stein, Granot, 2019))
UDP-glucose pyrophosphorylase	EC:2.7.7.9	UGPase1, UGPase2	Reversible conversion of UDP-glucose and pyrophosphate to glucose-1-phosphate and UDP (Kleczkowski et al., 2010)
Phosphoglucoisomerase	EC:5.3.1.9	PGI, PGI-like1, PGI-like2	Reversible conversion of fructose-6-phosphate to glucose-6-phosphate (Yu et al., 2000)
Phosphoglucomutase	EC:5.4.2.2	PGM1, PGM2.1, PGM2.2	Reversible conversion of glucose-6-phosphate to glucose-1-phosphate (Yu et al., 2000)
ADP-glucose pyrophosphorylase: large and small subunits	EC:2.7.7.27	AGPL1, AGPL2, AGPL3	Reversible synthesis of ADP-glucose and pyrophosphate from glucose-1-phosphate and ATP (Geigenberger et al., 1999)
	EC:2.7.7.27	AGPS1.1, AGPS1.2, AGPS2	
Inorganic pyrophosphatase	EC:3.6.1.1	PPase, PPase-like	Pyrophosphate cleavage to orthophosphate (George et al., 2010)
		Starch granule synthesis	
Granule-bound starch synthase	EC:2.4.1.242	GBSS1	Amylose synthesis by forming α-1,4-glycosidic bonds (Pfister, Zeeman, 2016)
Soluble starch synthases	EC:2.4.1.21	SS1, SS2, SS3, SS4, SS5, SS6	SS1, SS2, SS3: amylopectin synthesis by forming α-1,4-glycosidic bonds (Pfister, Zeeman, 2016). SS4, SS5: granule initiation (Helle et al., 2018; Tetlow, Bertoft, 2020). SS6: associated with starch granules in potato tubers, function unknown (Helle et al., 2018)
Starch-branching enzymes	EC:2.4.1.18	SBE1.1, SBE1.2, SBE2, SBE3	SBE2, SBE3: branching of amylose and amylopectin molecules by forming α-1,6-glycosidic bonds (Tetlow, Bertoft, 2020). The SBE1 function is unknown (Van Harsselaar et al., 2017)
Debranching enzymes α- (1→6) glucan hydrolase starch-debranching enzyme (DBE). Isoamylase	EC:3.2.1.68	ISA1.1, ISA1.2, ISA2, ISA3	Hydrolyse α-1,6-glycosidic bonds of amylopectin. ISA1 and ISA2 are involved in amylopectin synthesis and the formation of its semicrystalline structure (Hennen-Bierwagen et al., 2012). The ISA1/ISA2 complex participates in the regulation of starch granule formation (Bustos et al., 2004). ISA3 participates in starch degradation (Streb, Zeeman, 2012; Pfister, Zeeman, 2016)
		Starch degradation	
Glucan water dikinase	EC:2.7.9.4	GWD	Glucan phosphorylation at position C6 (Ritte et al., 2006)
Phosphoglucan water dikinase	EC:2.7.9.5	PWD	Glucan phosphorylation at position C3 (Ritte et al., 2006)

End of the Table

Enzyme	EC no.	Izozyme	Function
Phosphoglucan phosphatases Like SEX Four; Starch Excess 4	EC:3.1.3.48	LSF1, LSF2 SEX4, SEX4-like	Glucan dephosphorylation. SEX4 dephosphorylates positions C3 and C6 of glycoside residues. LSF2 dephosphorylates position C3 (Hejazi et al., 2010; Santelia et al., 2011)
α-Amylase	EC:3.2.1.1	AMY1.1, AMY1.2, AMY2 (AMY23), AMY3, AMY3-like	AMY3: hydrolysis of α-1,4-glycosidic bonds of amylose and amylopectin (Yu et al., 2005). AMY2 (AMY23): may be involved in the cold-induced sweetening of starch (Hou et al., 2017)
β-Amylase	EC:3.2.1.2	BAM1, BAM2, BAM3.1, BAM3.2, BAM4, BAM6.1, BAM6.2, BAM6.3, BAM7, BAM9	BAM1 and BAM3: cleavage of α-1,4-glycosidic bonds in amylose and amylopectin; maltose formation (Fulton et al., 2008). BAM1, BAM9 may be involved in the cold-induced sweetening of starch (Hou et al., 2017)
α-Glucan phosphorylase	EC:2.4.1.1	PHO1a, PHO1b, PHO2a, PHO2b	Glucose-1-phosphate production by maltooligosaccharide degradation; participation in starch granule growth (Pfister, Zeeman, 2016)
4-α-Glucanotransferase, disproportionating enzyme	EC:2.4.1.25	DPE1, DPE2	Maltooligosaccharide degradation to glucose (Critchley et al., 2001). Involved in granule initiation (Tetlow, Bertoft, 2020)

tion during evolution. The paralogous genes, which encode different enzyme isoforms, show tissue- and/or stage-specific expression patterns (Van Harsselaar et al., 2017; Qu et al., 2018; López-González et al., 2019). Various factors affect the expression of starch metabolism genes: circadian rhythms, photoperiod, levels of plant hormones and sugars, and stressing factors (drought and cold) (López-González et al., 2019). Although the key enzymatic reactions in starch biosynthesis and degradation are well known, there are many unclear points concerning alternative enzymatic processes and their localization inside the cell, roles of particular isozymes in starch metabolism in different organs at different developmental stages, and the influence of regulatory factors.

By now, functions of many proteins involved in starch metabolism in potato have been identified (see the Table). Most enzymes have isoforms with partially overlapping functions (Van Harsselaar et al., 2017). Translocator proteins essential for transporting metabolites through plastid membranes are important for starch metabolism as well. They also have multiple forms: adenylate translocators NTT1 and NTT2 (Tjaden et al., 2001), glucose transporter pGlcT1 (Cho et al., 2011), glucose 6-phosphate translocator (isoforms GPT1.1, GPT1.2, GPT2.1, and GPT2.2) (Kammerer et al., 1998), maltose transporter MEX1 (Cho et al., 2011), and triose phosphate translocator (TPT, TPT-like) (Flügge et al., 1989).

The amylose:amylopectin ratio in plants can be modified by raising lines carrying certain alleles of genes involved in starch synthesis. Such accessions were obtained in cereals; for instance, in maize *Z. mays*. The *amylose extender* (*ae*$^-$) mutation is associated with the loss of the activity of the starch-branching enzyme SBEIIb. Starch in plants with the *ae*$^-$ phenotype is enriched with amylose, and its amylopectin chains are longer (Stinard et al., 1993). The maize phenotype whose starch has practically no amylose is named *waxy*, and it is determined by a mutation in the gene for granule-bound starch synthase *GBSSI*. Its endosperm is gluey (Hossain et al., 2019).

However, the breeding of potato varieties with specified starch properties is complicated by its autotetraploid genome. The only amylose-free potato cultivar Eliane was obtained by mutation-assisted breeding (Muth et al., 2008). Genetic engineering and genome editing aimed at the modification of key starch biosynthesis genes produced plants with expected phenotypes: amylose-free starch (knockout and knockdown of the *GBSS1* gene), amylose-enriched starch (knockdown of

both *SBE1* and *SBE2*), and starch with modified amylopectin properties (editing of *SBE1* and/or *SBE2*) (Visser et al., 1991; Schwall et al., 2000; Hofvander et al., 2004; Andersson et al., 2006, 2017; Tuncel et al., 2019).

In some cases, plants with a desired phenotype acquired additional traits. For example, the increase in amylose content obtained by antisense suppression of the genes for starch-branching enzymes *SBE1* and *SBE2* was accompanied by a decrease in starch content, formation of smaller granules, and larger tubers (Hofvander et al., 2004). Apparently, changes in certain starch metabolism steps may affect the overall carbohydrate metabolism in the plant. An association study revealed genetic loci associated with starch content and productivity (tuber weight), whereas the functions of some detected genes were unknown at all, and some other genes were involved in signaling and regulation: transcriptional and posttranscriptional (Schönhals et al., 2017). Part of the detected SNPs exerted antagonistic effects on potato productivity and starch content (Schönhals et al., 2017). Thus, the investigation of regulation pathways of starch metabolism genes is important for improving potato quality and productivity.

Carbohydrate metabolism in potato plants

Two starch forms are recognized with regard to biologic function: transitory and storage. Transitory starch is synthesized and accumulated in chloroplasts of photosynthetic organs (leaves) in the daytime and degraded in darkness to provide nutrients for the cell. Storage starch is synthesized in amyloplasts (nonphotosynthetic plastids) of storage organs (e. g., potato tubers) and stored there for a long time to be utilized in the preparation for sprouting (Zeeman et al., 2010; Streb, Zeeman, 2012).

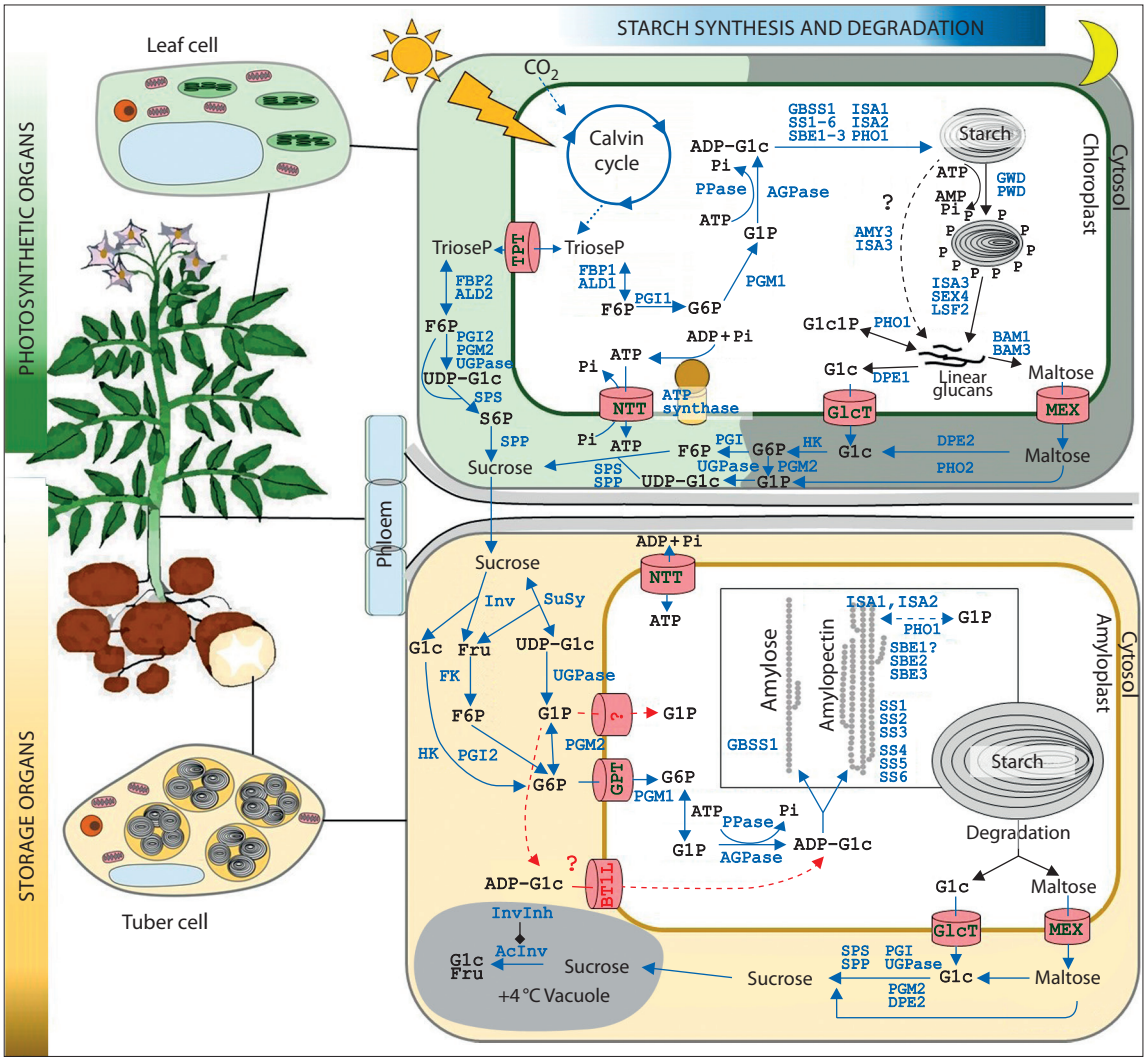
The main difference in carbohydrate metabolism between cells of leaves and storage organs is in the sources of carbohydrates and ATP required for starch synthesis and enzymatic reactions. In leaves, they can form in the same cells that produce transitory starch, and in storage organs, they are imported from photosynthetic ones.

During photosynthesis, chloroplasts produce ATP and fix atmospheric carbon dioxide by the Calvin–Benson cycle, in which triose phosphate is produced as an intermediate (Streb, Zeeman, 2012). Part of triose phosphate molecules remain in the chloroplast stroma to serve as carbohydrate material for transitory starch synthesis. The rest is transported to cytosol by triose phosphate translocator TPT (Flügge et al., 1989). In chloroplasts, a series of enzymatic reactions converts triose phosphate to glucose-1-phosphate (G1P). Then ADP-glucose pyrophosphorylase (AGPase) converts G1P to ADP-glucose, the main substrate for starch synthesis (see the Figure).

As triose phosphate molecules are exported to cytosol, they are converted to sucrose, which is then delivered to storage organs through phloem and apoplast to be a carbohydrate material for storage starch synthesis. Sucrose is transported into cells of storage organs either by sucrose transporter proteins or, after being hydrolyzed by invertase to glucose and fructose, by hexose transporters (Ruan, 2014). There are two pathways to cleave sucrose in cell cytoplasm: saccharolytic, catalyzed by sucrose synthase SuSy, or hydrolytic, catalyzed by invertase Inv. Invertase irreversibly cleaves sucrose to glucose and fructose, and SuSy catalyzes the reversible cleavage to fructose and UDP-glucose (Stein, Granot, 2019). The predominant pathway depends on the tuber development stage. At the beginning of growth, at a high cell division rate, the hydrolytic pathway prevails, and the saccharolytic pathway steps forward at the starch accumulation stage (Appeldoorn et al., 1997). The SuSy-catalyzed pathway is important for the rate of starch accumulation in potato. It has been shown that a decrease in SuSy activity reduces starch content in mature tubers (Zrenner et al., 1995; Baroja-Fernández et al., 2009). Seven *SuSy* isoforms have been predicted in potato, and the *SuSy4* gene had tissue-specific expression in growing tubers (Van Harsselaar et al., 2017). It is likely that some sucrose synthase isoforms are involved in sucrose transport through phloem, as observed in *A. thaliana* (Yao et al., 2020).

The subsequent steps, catalyzed by fructokinase FK, hexokinase HK, and UDP-glucose pyrophosphorylase, produce a pool of phosphorylated hexoses in cytosol. They can be reversibly interconverted by cytosolic isoforms of phosphoglucoisomerase PGI2 and phosphoglucomutase PGM2 (Yu et al., 2000; Kleczkowski et al., 2010). One of the phosphorylated hexoses, glucose-6-phosphate, is transported to amyloplasts by membrane glucose-6-phosphate/phosphate-translocator GPT (Kammerer et al., 1998). In amyloplasts, G6P is utilized for ADP-glucose production. The rate of starch biosynthesis in potato tubers directly depends on G6P transport (Tauberger et al., 2000; Fernie et al., 2002).

ATP is also imported to storage tissues from photosynthetic ones. It is delivered inside amyloplasts by plastid adenylate translocator NTT. It is known that even a slight decrease in NTT activity reduces the overall starch content in potato tubers (Tjaden et al., 2001), whereas the combination of NTT and GPT overexpression increases it (Zhang L. et al., 2008). There may be an alternative route of hexose transport to amyloplasts: direct import of glucose-1-phosphate (G1P) and its utilization in ADP-glucose synthesis. There is evidence that this route acts in potato, although the corresponding transporter



Starch metabolism in potato photosynthesizing (leaves) and storage (tubers) organs.

Blue arrows indicate the starch biosynthesis and sugar metabolism pathways; black arrows, the starch degradation pathway. Starch degradation in leaves and starch biosynthesis in tubers are shown in more detail. The red arrow indicates the alternative pathway of ADP-glucose synthesis in cytosol and subsequent transport to amyloplasts performed by BT1 (Brittle1-like transporter). Carbohydrates are shown in black: Fru, fructose; F6P, fructose-6-phosphate; Glc, glucose; G1P, glucose-1-phosphate; G6P, glucose-6-phosphate; S6P, sucrose-6-phosphate; TrioseP, triose phosphate; ADP-Glc, ADP-glucose; UDP-Glc, UDP-glucose. Enzymes and invertase inhibitor (in blue): AcInv, vacuolar acid invertase; AGPase, ADP-glucose pyrophosphorylase; ALD, aldolase; AMY, α-amylase; BAM, β-amylase; FK, fructokinase; DPE, 4-α-glucanotransferase; FBP, fructose-1,6-bisphosphatase; GBSS, granule-bound starch synthase; GWD, α-glucan water dikinase; HK, hexokinase; ISA, isoamylase; InvInh, invertase inhibitor; Inv, invertase; PGI, phosphoglucoisomerase; PGM, phosphoglucomutase; PHO, α-glucan phosphorylase; PPase, inorganic pyrophosphatase; PWD, phosphoglucan water dikinase; SBE, starch-branching enzyme; SEX, LSF, phosphoglucan phosphatases; SPP, sucrose phosphate phosphatase; SPS, sucrose phosphate synthase; SS, starch synthase; SuSy, sucrose synthase; UGPase, UDP-glucose pyrophosphorylase. Transporter proteins (in green): GPT, glucose-6-phosphate translocator; NTT, ATP-ADP antiporter; MEX, maltose transporter; TPT, triose phosphate/phosphate translocator; GlcT, glucose transporter; BT1L, Brittle1-like transporter.

protein is not known yet (Fettke et al., 2010). Two candidate G1P transporters through plasma membrane were recently found in *A. thaliana* (Malinova et al., 2020).

ADP-glucose acts as the substrate for starch biosynthesis. It is produced in a reversible reaction catalyzed by ADP-glucose pyrophosphorylase (AGPase) in the stroma of chloroplasts and amyloplasts. AGPase synthesizes ADP-glucose and pyrophosphate (PPi) from G1P and ATP. AGPase is a heterotetramer consisting of two large and two small subunits, AGPL and AGPS. Its activity is essential for starch synthesis in potato tubers (Geigenberger et al., 1999). Inorganic pyrophosphatase (PPase) degrades pyrophosphate to orthophosphate (George et al., 2010). The plastid PPase isoform con-

tributes much to starch accumulation in potato tubers. Lines knocked down for the *StpsPPase* gene had lower contents of starch, in particular, amylose, and smaller granules. That study also recorded elevated amounts of starch biosynthesis intermediates: pyrophosphate, glucose, fructose, hexose phosphates, and, unexpectedly, ADP-glucose. The increase in ADP-glucose content indicates that pyrophosphate does not affect the direction of the AGPase-catalyzed reaction in potato. Thus, the mechanism by which PPase participates in starch synthesis in potato tubers is still to be understood (Andersson et al., 2018).

An alternative cytosolic pathway of ADP-glucose synthesis, catalyzed by SuSy and UGPase, acts in cereals (monocots). It is essential for grain growth. ADP-glucose is transported to amyloplast by the Brittle1-like transporter protein (BT1) (Bowsher et al., 2007). The homolog protein of BT1 (StBT1) has been found in *S. tuberosum* plant. However, there is no evidence for ADP-glucose transport through the amyloplast membrane. The StBT1 protein performs unidirectional transport of AMP, ADP, and ATP (Leroch et al., 2005).

Starch granule synthesis

Transitory starch synthesis in leaf chloroplasts and storage starch synthesis in tuber amyloplasts follow basically the same route. Amylose and amylopectin synthesis is performed by 16 key enzymes belonging to the following groups: starch synthases, starch-branching enzymes, and starch-debranching enzymes (see the Table). Most enzymes exist as isoforms, the functions of which may partly overlap (Pfister, Zeeman, 2016; Van Harsselaar et al., 2017).

Starch synthases catalyze the formation of glycosidic bonds by transferring the glucose residue of ADP-glucose to the nonreducing end of the glucose polymer. They are subdivided into granule-bound (GBSS) and soluble (SS) starch synthases. The former synthesize long chains, mainly in amylose, and long chain fragments in amylopectin. The latter include a series of isoforms: SS1, SS2, SS3, SS4, SS5, and SS6. Of them, SS1, SS2, and SS3 synthesize chains of various lengths in amylopectin (Pfister, Zeeman, 2016). The SS4 isoform performs a special function among starch synthases, as it initiates starch granule formation (Tetlow, Bertoft, 2020). One or two large starch granules instead of five to seven wild-type small ones were found in *A. thaliana* plants with the knocked out *ss4* gene (Roldán et al., 2007). The functions of SS5 and SS6 are still vague. The C end of the SS5 protein lacks the conservative fragment characteristic of starch synthases, which has catalytic domain GT1, although the protein has the conservative glucan-binding site. Probably, SS5 is involved in starch granule initiation, as it has been shown that the loss of SS5 from *A. thaliana* reduces the granule number in leaves (Abt et al., 2020). The SS6 isoform and its gene were found in potato in recent years (Van Harsselaar et al., 2017), and the role of this enzyme is unknown. It may participate in granule growth, as it is directly bound to it; in addition, it bears conservative motifs XXGGL and KXGGL, characteristic of glycosyl transferase domains of starch synthases GT1 and GT5, respectively (Helle et al., 2018).

Starch granule initiation was an obscure issue for a long time. The studies reported by now concern transitory starch initiation in the model plant *A. thaliana*, but it seems that our notion of some key steps in granule formation may be extended to other plant species (Mérida, Fettke, 2021). As mentioned above, SS4 is the main granule-initiating enzyme, and SS5 and SS6 also take part in the process. Maltooligosaccharides, probably forming in the degradation of starch polyglucans by amylases, are the substrate (Mérida, Fettke, 2021). The steric interaction of starch synthases, substrate molecules, and the growing granule is driven by the PTST2 and PTST1 proteins, targeting to starch. They are associated with starch synthases SS4 and GBSS1, respectively. Also, they contain a carbohydrate-binding domain (Seung et al., 2015, 2017). Note that PTST2 is not found in potato tubers, and this fact indicates that the starch granule initiation processes in *A. thaliana* and potato differ (Helle et al., 2018). A heteromultimeric complex of isoamylases ISA1 and ISA2 has been shown to influence starch granule initiation in potato tubers. By all appearances, isoamylases suppress the formation of new starch granules by disrupting the formation of soluble glucan molecules in chloroplast stroma (Bustos et al., 2004).

As the chains of amylose and amylopectin molecules are elongated by starch synthases SS1, SS2, and SS3, starch-branching enzymes SBE attach side branches to them (Pfister, Zeeman, 2016). Starch-branching enzymes cleave α-1,4-glycosidic bonds of polyglucans, synthesized by starch synthases, and attach short chains to the so-called acceptor chain by forming α-1,6-glycosidic bonds. The starch-branching enzymes of potato have three isoforms: SBE1.1, SBE1.2, SBE2, and SBE3 (formerly designated as SBE1) (see the Table) (Van Harsselaar et al., 2017). Thus, forms referred to in other papers as SBE1 and SBE2 are designated as SBE3 and SBE2 according to the notation of Van Harsselaar et al. SBE3 produces mainly long side chains, and SBE2 produces short amylopectin chains (Tetlow, Bertoft, 2020). The roles of SBE1.1 and SBE1.2 in starch production are unknown, but studies of *A. thaliana* demonstrate a pleiotropic effect of SBE1 on plant growth and development. Transformants overexpressing SBE1 were

white-colored and low. They had a longer life cycle and produced fewer seeds than control plants (Wang X. et al., 2010). The joint action of different isoforms affects amylopectin structure. Experiments with potato plants with the knocked out genes *sbe3* and/or *sbe2* (designated by the experimenters as *sbe1* and *sbe2*) demonstrate that *sbe3* inactivity results in the formation of starch with longer amylopectin chains and lower branching level. The knockout of *sbe2* with active *sbe3* did not affect the amylopectin structure much, but the number of starch granules in potato tubers increased and size decreased (Tuncel et al., 2019).

Debranching enzymes (DBE) are another group of enzymes involved in the formation of the amylopectin structure (see the Table). They reconstruct branched glucans into easier crystallizable forms, which is essential for granule formation (Pfister, Zeeman, 2016). Debranching enzymes include isoamylases (ISA), which catalyze the hydrolysis of α-1,6-glycosidic amylopectin bonds and remove excessive branching. Potato isoamylases include isoforms ISA1, ISA2, and ISA3. The ISA1 and ISA2 proteins can form heteromultimers, capable of more efficient removal of long outer chains of amylopectin (Hussain et al., 2003). The ISA3 isoenzyme is important for starch degradation, as it cleaves short outer chains of glucans (Streb et al., 2008). Transgenic potato plants with lower expression of the *isa1*, *isa2*, and *isa3* genes had significantly less starch in developing tubers, whereas the starch contents in leaves did not change. The plants also had fewer and larger granules and higher sucrose contents, probably resulting from the increase in the overall granule surface and easier access for degrading enzymes (Ferreira et al., 2017).

In addition to starch synthases, branching and debranching enzymes, the synthesis of starch granules involves α-glucan phosphorylases. Their plastid (PHO1) and cytoplasmic (PHO2) isoforms catalyze the reversible transfer of the glycosylic group of glucose-1-phosphate to the nonreducing end of the chain of an α-1,4-bound glucan (Pfister, Zeeman, 2016). The PHO2 enzyme is involved in carbohydrate metabolism in cytoplasm, and PHO1 contributes to starch synthesis and degradation in plastids. It has been shown that at lower temperatures starch synthesis in potato tubers can also follow the phosphorylase pathway with G1P as the substrate (Fettke et al., 2012).

Starch granule degradation

Degradation is an intrinsic part of the metabolism of starch and carbohydrates in general, although it has been studied much poorer than starch biosynthesis. The degradation pathways of transitory starch have been investigated in most detail in leaves of the model plant *A. thaliana*. The knowledge of starch degradation in potato tubers is limited to cold-induced sweetening and sprouting. The main steps of starch degradation are the release of soluble glucan from starch granules, glucan conversion to linear forms (maltooligosaccharides), maltooligosaccharide hydrolysis to maltose, and subsequent maltose metabolism in the cell. Starch degradation is performed by a broad range of enzymes: α- and β-amylases, isoamylase, α-glucan water dikinase (GWD), phosphoglucan water dikinase (PWD), α-glucan phosphorylase, phosphoglucan phosphatase, and 4-α-glucanotransferase (see the Table).

Starch granule degradation is initiated by GWD and PWD. They phosphorylate glucans at positions C6 and C3 of glucose residues, making them more hydrophilic and allow α-, β-, and isoamylases access to them (see the Table) (Ritte et al., 2006; Streb, Zeeman, 2012). The phosphorylation by GWD seems to play the key role in starch degradation in potato tubers and leaves (Claassen et al., 1993; Orzechowski et al., 2021). Tubers of transgenic potato plants with lower expression of the *StGWD* gene were less prone to starch degradation at low temperatures (Lorberth et al., 1998).

The next step of starch granule degradation is glucan hydrolysis by amylases. Potato α- and β-amylases include many isoforms, and functions of some of them are not known in detail (see the Table) (Van Harsselaar et al., 2017). By extrapolating data on *A. thaliana*, we suppose that β-amylases BAM1 and BAM3 hydrolyze linear fragments of amylose and amylopectin, and the degradation of branched fragments demands the debranching enzyme DBE (ISA3 in potato) (Hussain et al., 2003; Fulton et al., 2008; Pfister, Zeeman, 2016). With knocked down *StBAM3*, starch content in potato leaves was higher than in the wild genotype (Scheidig et al., 2002). Cold-induced sweetening in potato tubers is also affected by some amylase species: α-amylase AMY2 (AMY23) and β-amylases BAM1 and BAM9. The supposed function of BAM1 and BAM9 is starch degradation in plastids, and AMY2 is likely to degrade phytoglycogen in cytosol (Hou et al., 2017). An alternative pathway of starch degradation is observed in *A. thaliana*. It is initiated by α-amylase AMY3, which releases linear and branched glucans from starch granules, and these glucans are then hydrolyzed by β- and isoamylases (see the Figure) (Streb et al., 2008).

Alongside starch glucan hydrolysis by amylases, the glucans are dephosphorylated by phosphoglucan phosphatases SEX4 (Starch Excess) and LSF2 (LIKE SEX FOUR2), first described in *A. thaliana*. These processes are interrelated: phosphorylation by dikinases increases granule solubility and makes them accessible for amylases, whereas phosphate moieties may hamper

hydrolysis (Hejazi et al., 2010; Santelia et al., 2011). Reduction of SEX4 or LSF2 activities in potato inhibited starch degradation in leaves. Starch content in tubers remained unchanged, and granules were smaller and less phosphorylated (Samodien et al., 2018).

The cooperation of dikinases, amylases, and phosphatases produces a pool of soluble maltooligosaccharides (linear glucans). Maltooligosaccharides are degraded by two pathways: hydrolytic, by β-amylases, or phosphorolytic, by α-glucan phosphorylase PHO1 (Weise et al., 2006; Fulton et al., 2008). The end product of the phosphorolytic pathway is G1P, which can be utilized in metabolism inside the plastid. Also, glucose can be produced by 4-α-glucanotransferase DPE1 (DisProportionating Enzyme) and exported to cytosol by glucose transporter pGlcT1 (Critchley et al., 2001; Cho et al., 2011). Knockdown of the chloroplast enzyme DPE slows down starch degradation in potato leaves in the cold and induces maltooligosaccharide accumulation, although these effects are not observed in tubers (Lloyd et al., 2004). Cold-induced sweetening in potato tubers is accompanied by increasing β-amylase activity and higher maltose content (Nielsen et al., 1997).

Maltose, which is the predominant product of hydrolytic starch degradation, is exported to cytosol by the transmembrane transporter MEX1 (Cho et al., 2011). In cytosol, maltose is processed by 4-α-glucanotransferase DPE2 or phosphorylase PHO2 to glucose or G1P, which are then converted to sucrose by the joint action of PGI2, PGM2, HK, UGPase, SPS (sucrose phosphate synthase), and SPP (sucrose phosphate phosphatase) (see the Figure) (López-González et al., 2019).

Sucrose is exported from leaf cells to storage organs; also, it is used in cell metabolism. In potato tubers, sucrose is used as a source of nutrients in sprouting, and its level controls dormancy release (Sonnewald S., Sonnewald U., 2014).

To delay sprouting, potato tubers are stored at low temperatures, 2–5 °C, and these conditions initiate cold-induced sweetening. This process involves sucrose hydrolysis by vacuolar acid invertase AcInv, encoded by the *Pain-1* gene, and the accumulation of reducing sugars (glucose and fructose) in tubers (see the Figure) (Sowokinos et al., 2018). Knockout of *Pain-1* resulted in lower contents of reducing sugars (Clasen et al., 2016). One of the key regulators of cold-induced sweetening is invertase inhibitor SbAI, which inhibits AcInv (McKenzie et al., 2013). It has been shown that SbAI can also inhibit α- and β-amylases (StAmy23, StBAM1, and StBAM9), in potato tubers, thereby influencing the rate of starch degradation in cold-induced sweetening (Zhang H. et al., 2014).

Mechanisms controlling starch metabolism

Starch metabolism requires orchestrated work of many enzymes, transporters, and targeting proteins, which implies many regulation levels: gene expression, post-transcriptional regulation, and the posttranslational regulation of enzymatic activity. The expression patterns of genes for key enzymes involved in starch metabolism are well known in various plant species, but much less is known about expression-regulating factors (López-González et al., 2019). The difficulty is that the starch-metabolizing enzymes exist as numerous isoforms, which are encoded by the corresponding number of paralogous genes. The expression patterns of these genes depend on tissue (leaves, developing seeds, or growing tubers) and developmental stage, as shown in *A. thaliana* and maize (Tsai et al., 2009; Chen et al., 2014). In potato, the tissue-specific mode of expression has been shown for *SuSy4*, *SS5*, *SBE3*, *APL3*, *PHO1a*, *PHO1b*, *GPT1.1*, *GPT2.1*, *SEX4*, and *NTT2* in tubers and for *AMY1.1*, *APL1*, and *BAM3.1* in leaves (Van Harsselaar et al., 2017). A number of external and internal factors affect the expression of starch biosynthesis genes: circadian rhythms, photoperiod, and sugar content (Tiessen et al., 2002; Kötting et al., 2010). It is known that the expression rates of *GBSSI*, *LSF1*, *LSF2*, *SEX4*, and *BAM3* in *A. thaliana* leaves are governed by transcription factors depending on circadian rhythms and photoperiod, so that the demand for energy is rapidly met in response to ambient changes (Tenorio et al., 2003; Flis et al., 2016). The expression rates of the genes *GBSS*, *SuSy*, and *AGPase* respond to photoperiod in growing potato tubers, being highest in the end of the light time and lowest in the beginning. This variation is determined by the influx of photoassimilates from leaves (Geigenberger, Stitt, 2000; Ferreira et al., 2010).

The formation of the storage organ, tuber, from the stolon is an important step in potato plant development. It includes intense starch production, the formation of starch granules, and increase in metabolite flux. Tuber formation is a complex process, influenced by environmental factors (photoperiod) and a variety of signals: biochemical, hormonal, and molecular, mediated by microRNAs and transcription factors (Hannapel et al., 2017; Kondhare et al., 2021). The investigation of tuber formation contributed much to the understanding of mechanisms that regulate starch metabolism in potato tubers.

Plant hormones are an important factor influencing the expression of genes involved in starch metabolism, and their effect on tuber formation has been studied in sufficient detail. The level of abscisic acid correlates with starch accumulation in potato tubers (Borzenkova, Borovkova, 2003). Treatment of stolons with indole ace-

tic acid increased starch content in growing tubers, but a twofold increase in concentration caused the opposite effect (Wang D. et al., 2018). A correlation between the transcription rates of the genes for auxin, on the one hand, and starch biosynthesis (*PGM*, *AGPase*, *GBSS*, *SS*, and *BE*), on the other hand, was observed in the formation of cassava storage roots (Rüscher et al., 2021).

Sugars (hexoses, sucrose, and trehalose) are another group of signaling molecules influencing the expression of starch metabolism genes. Sucrose increases the expression of the *SuSy* and *AGPase* genes in potato (Salanoubat, Belliard, 1989; Müller-Röber et al., 1990). The rates of *SuSy* and *AGPase* expression are high in growing tubers, but they decrease rapidly after the separation of the tuber from the plant and, correspondingly, cease of sucrose import from photosynthesizing organs (Ferreira et al., 2010).

The differential expression of starch biosynthesis genes was detected at various tuber development stages (Ferreira et al., 2010; Van Harsselaar et al., 2017). The expression rate of the *SS4* gene was elevated at the stolon stage, and it lowered with tuber growth, confirming the role of this starch synthase in granule initiation (Ferreira et al., 2010). Also, tuber growth was accompanied by an increase in the expression rate of sucrose synthase *SuSy4* and decrease in the expression of cell wall invertase *cw-Inv*. These changes point to transition to the sucrose synthase-mediated pathway of sucrose degradation. Genes for glucose-6-phosphate translocator GPT, adenylate translocator NTT, ADP-glucose pyrophosphorylase (AGPase), starch synthases, and starch-branching enzymes increased their expression with tuber growth. Of this group, the isogenes *SuSy4*, *SBE3*, and *GPT2.1* demonstrated just tuber-specific expression (Ferreira et al., 2010; Van Harsselaar et al., 2017). Coexpression analysis was employed to investigate the mechanisms of molecular regulation of gene activity, and transcription factors LOB, TIFY5a, and WRKY4 were found to be associated with the expression of the *SuSy4* and *GPT2.1* genes (Van Harsselaar et al., 2017). Analysis of coexpression networks for starch biosynthesis genes of seven plant species (*Arabidopsis thaliana*, cassava *Manihot esculenta*, millet *Panicum virgatum*, maize *Zea mays*, rice *Oryza sativa*, barley *Hordeum vulgare*, and sweet potato *Ipomoea batatas*) revealed the involvement of 24 transcription factors (López-González et al., 2019).

Little is known about mechanisms regulating starch metabolism genes at the posttranscriptional step. Posttranscriptional regulation involves a variety of factors, including RNA-binding proteins (RBPs), microRNAs, and alternative splicing, so that plants can rapidly reprogram their transcriptomes in response to external and internal factors. Photoperiod significantly influences microRNA expression patterns in the growth and development of potato tubers. It has been shown that differentially expressing microRNAs are targeted to genes coding for transcription factors and RNA-binding regulatory proteins *StGRAS*, *StTCP2/4*, and *StPTB6* (Kondhare et al., 2018).

Posttranslational regulation is the next step of protein activity control. It is mediated by allosteric regulation, in which an effector molecule is bound to a noncatalytic site of the enzyme, altering its conformation, catalytic properties, and, thereby, its specificity and interaction with other proteins (Zeeman et al., 2010). Allosteric regulation involves protein phosphorylation and the formation of multimeric complexes and disulfide bridges (Kötting et al., 2010; Zeeman et al., 2010). Many starch-metabolizing enzymes assume the phosphorylated state: PGI, PGM1, AGPase, SS3, GWD1, GWD2, DPE2, AMY3, BAM1, BAM3, LDA, pGlcT, and MEX1 (Kötting et al., 2010). ADP-glucose pyrophosphorylase (AGPase) is a clear example of allosterically regulated potato enzyme. It is activated by 3-phosphoglyceric acid and inhibited by inorganic phosphate (Sowokinos, Preiss, 1982). Depending on the redox state in the cell, AGPase can be reversibly inactivated by the formation of disulfide bridges between small subunits of the heterotetramer (Ballicora et al., 2000).

Enzymes can aggregate into complexes known as metabolons (Sweetlove, Fernie, 2013). Complexes formed by starch biosynthesis enzymes were found in the endosperm of growing cereal seeds; in particular, SSIII, SSIIa, SBEIIa, and SBEIIb form a protein complex (Tetlow et al., 2008). Protein complexes of PTST2 and SS4 form in the initiation of starch granules in *A. thaliana* leaves (Seung et al., 2015, 2017). Potato isoamylases ISA1 and ISA2 form a heterotetrameric complex, which controls starch granule formation (Bustos et al., 2004).

Conclusions

The investigation of starch metabolism in potato plants, particularly, starch biosynthesis and degradation in tubers, is topical in connection with the growing demand for potato starch in industry. A large body of information on key enzymes for starch and carbohydrate metabolism in various crops and the model species *A. thaliana* has been accumulated in the past three decades. Although the starch biosynthesis scheme is basically the same in different species, there are significant variations associated with different sets of isozymes, features of their functions, metabolite transport pathways (e. g., ADP-glucose transport through plastid membranes in cereals), and the existence of intricate and multileveled regulation, governed by external (photoperiod and temperature) and internal (plant hormones, metabolites,

microRNA, and regulatory proteins) factors. Isogenes encoding six starch synthase isoforms, seven sucrose synthases, nine β-amylases, and three to five for each of the starch-branching and other enzymes were identified in the potato genome (Van Harsselaar et al., 2017).

The functions of many isoforms, including the majority of α- and β-amylases, are still unknown. Some isogenes (*SuSy4*, *SS5*, *SBE3*, *APL3*, *PHO1a*, *PHO1b*, *GPT1.1*, *GPT2.1*, *SEX4*, and *NTT2*) demonstrate tuber-specific expression and activity variation at various stages of tuber formation (Ferreira et al., 2010; Van Harsselaar et al., 2017). Isoenzymes AMY23, BAM1, BAM9 are specifically involved in starch degradation and carbohydrate metabolism in cold-induced sweetening (Hou et al., 2017). Also, the action of various factors on starch accumulation during tuber development has been shown: transcription factors LOB, TIFY5a, and WRKY4; plant hormones (auxin and abscisic acid); sugars; and microRNAs, the contents of which may mediate the effect of photoperiod. However, the functions of many isoenzymes and proteins involved in the regulatory and directing functions in starch metabolism in potato plants are poorly explored. To resolve this issue, modern methods are proposed: combined analysis of the metabolome and transcriptome inside a single cell or tissue (López-González et al., 2019). Bottom-up proteomics also seems promising in search for new components (Helle et al., 2018). For example, the analysis of 36 proteins associated with potato starch granules revealed, in addition to already known starch metabolism enzymes, targeting and regulatory proteins described in *A. thaliana*: PTST1 (Protein Targeting to Starch), ESV1 (Early StarVation1), and LESV (Like ESV). Also, Kunitz-type proteinase inhibitor and enzymes involved in redox regulation (thioredoxin TRX and glutathione peroxidase GPX) were found (Helle et al., 2018). Detailed information on all components involved in starch metabolism and on their interactions, including their behavior under varying ambient conditions, is essential for raising potato varieties with high performance and specified starch properties.

References

Abt M., Pfister B., Sharma M., Eicke S., Bürgy L., Neale I., Seung D., Zeeman S.C. STARH SYNTHASE5, a noncanonical starch synthase-like protein, promotes starch granule initiation in Arabidopsis. *Plant Cell.* 2020;32(8):2543-2565. DOI 10.1105/tpc.19.00946.

Andersson M., Melander M., Pojmark P., Larsson H., Bülow L., Hofvander P. Targeted gene suppression by RNA interference: an efficient method for production of high-amylose potato lines. *J. Biotechnol.* 2006;123:137-148. DOI 10.1016/j.jbiotec.2005.11.001.

Andersson M., Turesson H., Arrivault S., Zhang Y., Fält A.-S., Fernie A.R., Hofvander P. Inhibition of plastid PPase and NTT leads to major changes in starch and tuber formation in potato. *J. Exp. Bot.* 2018;69(8):1913-1924. DOI 10.1093/jxb/ery051.

Andersson M., Turesson H., Nicolia A., Fält A.S., Samuelsson M., Hofvander P. Efficient targeted multiallelic mutagenesis in tetraploid potato (*Solanum tuberosum*) by transient CRISPR-Cas9 expression in protoplasts. *Plant Cell Rep*. 2017;36:117-128. DOI 10.1007/s00299-016-2062-3.

Appeldoorn N.J.G., de Bruijn S.M., Koot-Gronsveld E.A.M., Visser R.G.F., Vreugdenhil D., van der Plas L.H.W. Developmental changes of enzymes involved in conversion of sucrose to hexose-phosphate during early tuberisation of potato. *Planta.* 1997;202:220-226. DOI 10.1007/s004250050122.

Ballicora M.A., Frueauf J.B., Fu Y., Schurmann P., Preiss J. Activation of the potato tuber ADP-glucose pyrophosphorylase by thioredoxin. *J. Biol. Chem.* 2000;275(2):1315-1320. DOI 10.1074/jbc.275.2.1315.

Baroja-Fernández E., Muñoz F.J., Montero M., Etxeberria E., Sesma M.T., Ovecka M., Bahaji A., Ezquer I., Li J., Prat S., Pozueta-Romero J. Enhancing sucrose synthase activity in transgenic potato (*Solanum tuberosum* L.) tubers results in increased levels of starch, ADPglucose and UDPglucose and total yield. *Plant Cell Physiol*. 2009;50(9):1651-1662. DOI 10.1093/pcp/pcp108.

Birt D.F., Boylston T., Hendrich S., Jane J.L., Hollis J., Li L., McClelland J., Moore S., Phillips G.J., Rowling M., Schalinske K., Scott M.P., Whitley E.M. Resistant starch: promise for improving human health. *Adv. Nutr*. 2013;4(6):587-601. DOI 10.3945/an.113.004325.

Borzenkova R.A., Borovkova M.P. Developmental patterns of phytohormone content in the cortex and pith of potato tubers as related to their growth and starch content. *Russ. J. Plant Physiol.* 2003;50:119-124. DOI 10.1023/A:1021957022595.

Bowsher C.G., Scrase-Field E.F.A.L., Esposito S., Emes M.J., Tetlow I.J. Characterization of ADP-glucose transport across the cereal endosperm amyloplast envelope. *J. Exp. Bot.* 2007;58(6):1321-1332. DOI 10.1093/jxb/erl297.

Bustos R., Fahy B., Hylton C.M., Seale R., Nebane N.M., Edwards A., Martin C., Smith A.M. Starch granule initiation is controlled by a heteromultimeric isoamylase in potato tubers. *Proc. Natl. Acad. Sci. USA.* 2004;101(7):2215-2220. DOI 10.1073/pnas.0305920101.

Chen J., Zeng B., Zhang M., Xie S., Wang G., Hauck A., Lai J. Dynamic transcriptome landscape of maize embryo and endosperm development. *Plant Physiol*. 2014;166(1):252-264. DOI 10.1104/pp.114.240689.

Cho M.H., Lim H., Shin D.H., Jeon J.S., Bhoo S.H., Park Y.I., Hahn T.R. Role of the plastidic glucose translocator in the export of starch degradation products from the chloroplasts in *Arabidopsis thaliana*. *New Phytologist*. 2011;190(1):101-112. DOI 10.1111/j.1469-8137.2010.03580.x.

Claassen P.A.M., Budde M.A.W., Van Calker M.H. Increase in phosphorylase activity during cold-induced sugar accumulation in potato tubers. *Potato Res*. 1993;36:205-217. DOI 10.1007/BF02360529.

Clasen B.M., Stoddard T.J., Luo S., Demorest Z.L., Li J., Cedrone F., Tibebu R., Davison S., Ray E.E., Daulhac A., Coffman A., Yabandith A., Retterath A., Haun W., Baltes N.J., Mathis L., Voytas D.F., Zhang F. Improving cold storage and processing traits in potato through targeted gene knockout. *Plant Biotechnol. J.* 2016;14(1):169-176. DOI 10.1111/pbi.12370.

Critchley J.H., Zeeman S.C., Takaha T., Smith A.M., Smith S.M. A critical role for disproportionating enzyme in starch breakdown is revealed by a knock-out mutation in *Arabidopsis*. *Plant J.* 2001;26(1):89-100. DOI 10.1046/j.1365-313x.2001.01012.x.

Fernie A.R., Swiedrych A., Tauberger E., Lytovchenko A., Trethewey R.N., Willmitzer L. Potato plants exhibiting combined antisense repression of cytosolic and plastidial isoforms of phosphoglucomutase surprisingly approximate wild type with respect to the rate of starch synthesis. *Plant Physiol. Biochem*. 2002;40(11):921-927. DOI 10.1016/S0981-9428(02)01457-2.

Ferreira S.J., Senning M., Fischer-Stettler M., Streb S., Ast M., Neuhaus H.E., Zeeman S.C., Sonnewald S., Sonnewald U. Simultaneous silencing of isoamylases ISA1, ISA2 and ISA3 by multi-target

RNAi in potato tubers leads to decreased starch content and an early sprouting phenotype. *PLoS One*. 2017;12(7):e0181444. DOI 10.1371/journal.pone.0181444.

Ferreira S.J., Senning M., Sonnewald S., Kessling P.M., Goldstein R., Sonnewald U. Comparative transcriptome analysis coupled to X-ray CT reveals sucrose supply and growth velocity as major determinants of potato tuber starch biosynthesis. *BMC Genomics.* 2010;11:93. DOI 10.1186/1471-2164-11-93.

Fettke J., Albrecht T., Hejazi M., Mahlow S., Nakamura Y., Steup M. Glucose 1-phosphate is efficiently taken up by potato (*Solanum tuberosum*) tuber parenchyma cells and converted to reserve starch granules. *New Phytol.* 2010;185(3):663-675. DOI 10.1111/j.1469-8137.2009.03126.x.

Fettke J., Leifels L., Brust H., Herbst K., Steup M. Two carbon fluxes to reserve starch in potato (*Solanum tuberosum* L.) tuber cells are closely interconnected but differently modulated by temperature. *J. Exp. Bot.* 2012;63(8):3011-3029. DOI 10.1093/jxb/ers014.

Flis A., Sulpice R., Seaton D.D., Ivakov A.A., Liput M., Abel C., Millar A.J., Stitt M. Photoperiod-dependent changes in the phase of core clock transcripts and global transcriptional outputs at dawn and dusk in *Arabidopsis*. *Plant Cell Environ*. 2016;39(9):1955-1981. DOI 10.1111/pce.12754.

Flügge U.I., Fischer K., Gross A., Sebald W., Lottspeich F., Eckerskorn C. The triose phosphate-3-phosphoglycerate-phosphate translocator from spinach chloroplasts: nucleotide sequence of a full-length cDNA clone and import of the *in vitro* synthesized precursor protein into chloroplasts. *EMBO J*. 1989;8(1):39-46.

Fulton D.C., Stettler M., Mettler T., Vaughan C.K., Li J., Francisco P., Gil M., Reinhold H., Eicke S., Messerli G., Dorken G., Halliday K., Smith A.M., Smith S.M., Zeeman S.C. β-AMYLASE4, a non-catalytic protein required for starch breakdown, acts upstream of three active β-amylases in *Arabidopsis* chloroplasts. *Plant Cell*. 2008;20(4):1040-1058. DOI 10.1105/tpc.107.056507.

Geigenberger P., Muller-Rober B., Stitt M. Contribution of adenosine 5′-diphosphoglucose pyrophosphorylase to the control of starch synthesis is decreased by water stress in growing potato tubers. *Planta.* 1999;209(3):338-345. DOI 10.1007/s004250050641.

Geigenberger P., Stitt M. Diurnal changes in sucrose, nucleotides, starch synthesis and AGPS transcript in growing potato tubers that are suppressed by decreased expression of sucrose phosphate synthase. *Plant J.* 2000;23:795-806. DOI 10.1046/j.1365-313x.2000.00848.x.

George G.M., van der Merwe M.J., Nunes-Nesi A., Bauer R., Fernie A.R., Kossmann J., Lloyd J.R. Virus-induced gene silencing of plastidial soluble inorganic pyrophosphatase impairs essential leaf anabolic pathways and reduces drought stress tolerance in *Nicotiana benthamiana*. *Plant Physiol*. 2010;154(1):55-66. DOI 10.1104/pp.110.157776.

Hannapel D.J., Sharma P., Lin T., Banerjee A.K. The multiple signals that control tuber formation. *Plant Physiol.* 2017;174(2):845-856. DOI 10.1104/pp.17.00272.

Hejazi M., Fettke J., Kötting O., Zeeman S.C., Steup M. The Laforin-like dual-specificity phosphatase SEX4 from Arabidopsis hydrolyzes both C6- and C3-phosphate esters introduced by starch-related dikinases and thereby affects phase transition of α-glucans. *Plant Physiol.* 2010;152(2):711-722. DOI 10.1104/pp.109.149914.

Helle S., Bray F., Verbeke J., Devassine S., Courseaux A., Facon M., Tokarski C., Rolando C., Szydlowski N. Proteome analysis of potato starch reveals the presence of new starch metabolic proteins as well as multiple protease inhibitors. *Front. Plant Sci*. 2018;9:746. DOI 10.3389/fpls.2018.00746.

Hennen-Bierwagen T.A., James M.G., Myers A.M. Involvement of debranching enzymes in starch biosynthesis. In: Tetlow I.J. (Ed.) Essential Reviews in Experimental Biology. Vol. 5. The Synthesis and Breakdown of Starch. Society for Experimental Biology, London, 2012;179-215.

Hofvander P., Andersson M., Larsson C.T., Larsson H. Field performance and starch characteristics of high-amylose potatoes obtained by antisense gene targeting of two branching enzymes. *Plant Biotechnol. J.* 2004;2:311-320. DOI 10.1111/j.1467-7652.2004.00073.x.

Hossain F., Chhabra R., Devi E.L., Zunjare R.U., Jaiswal S.K., Muthusamy V. Molecular analysis of mutant *granule-bound starch synthase-I* (*waxy1*) gene in diverse waxy maize inbreds. *3 Biotech*. 2019;9(1):3. DOI 10.1007/s13205-018-1530-6.

Hou J., Zhang H., Liu J., Reid S., Liu T., Xu S., Tian Z., Sonnewald U., Song B., Xie C. Amylases StAmy23, StBAM1 and StBAM9 regulate cold-induced sweetening of potato tubers in distinct ways. *J. Exp. Bot*. 2017;68(9):2317-2331. DOI 10.1093/jxb/erx076.

Hussain H., Mant A., Seale R., Zeeman S., Hinchliffe E., Edwards A., Hylton C., Bornemann S., Smith A.M., Martin C., Bustos R. Three isoforms of isoamylase contribute different catalytic properties for the debranching of potato glucans. *Plant Cell.* 2003;15(1):133-149. DOI 10.1105/tpc.006635.

Kammerer B., Fischer K., Hilpert B., Schubert S., Gutensohn M., Weber A., Flügge U.I. Molecular characterization of a carbon transporter in plastids from heterotrophic tissues: the glucose 6-phosphate/phosphate antiporter. *Plant Cell.* 1998;10(1):105-117. DOI 10.1105/tpc.10.1.105.

Khlestkin V.K., Peltek S.E., Kolchanov N.A. Target genes for development of potato (*Solanum tuberosum* L.) cultivars with desired starch properties. *Selskokhozyaystvennaya Biologiya = Agricultural Biology.* 2017;52(1):25-36. DOI 10.15389/agrobiology.2017.1.25eng.

Kleczkowski L.A., Kunz S., Wilczynska M. Mechanisms of UDP-glucose synthesis in plants. *Crit. Rev. Plant Sci.* 2010;29(4):191-203. DOI 10.1080/07352689.2010.483578.

Kondhare K.R., Malankar N.N., Devani R.S., Banerjee A.K. Genome-wide transcriptome analysis reveals small RNA profiles involved in early stages of stolon-to-tuber transitions in potato under photoperiodic conditions. *BMC Plant Biol.* 2018;18:284. DOI 10.1186/s12870-018-1501-4.

Kondhare K.R., Patil A.B., Giri A.P. Auxin: an emerging regulator of tuber and storage root development. *Plant Sci.* 2021;306:110854. DOI 10.1016/j.plantsci.2021.110854.

Kötting O., Kossmann J., Zeeman S.C., Lloyd J.R. Regulation of starch metabolism: the age of enlightenment? *Curr. Opin. Plant Biol.* 2010;13(3):321-329. DOI 10.1016/j.pbi.2010.01.003.

Leroch M., Kirchberger S., Haferkamp I., Wahl M., Neuhaus H.E., Tjaden J. Identification and characterization of a novel plastidic adenine nucleotide uniporter from *Solanum tuberosum*. *J. Biol. Chem*. 2005;280(18):17992-8000. DOI 10.1074/jbc.M412462200.

Li L., Jiang H.X., Campbell M., Blanco M., Jane J.L. Characterization of maize amylose-extender (*ae*) mutant starches. Part I: Relationship between resistant starch contents and molecular structures. *Carbohydr. Polym*. 2008;74(3):396-404. DOI 10.1016/j.carbpol.2008.03.012.

Lloyd J.R., Blennow A., Burhenne K., Kossmann J. Repression of a novel isoform of disproportionating enzyme (stDPE2) in potato leads to inhibition of starch degradation in leaves but not tubers stored at low temperature. *Plant Physiol.* 2004;134(4):1347-1354. DOI 10.1104/pp.103.038026.

López-González C., Juárez-Colunga S., Morales-Elías N.C., Tiessen A. Exploring regulatory networks in plants: transcription factors of starch metabolism. *Peer. J.* 2019;7:e6841. DOI 10.7717/peerj.6841.

Lorberth R., Ritte G., Willmitzer L., Kossmann J. Inhibition of a starch-granule-bound protein leads to modified starch and repression of cold sweetening. *Nat. Biotechnol*. 1998;16(5):473-477. DOI 10.1038/nbt0598-473.

Malinova I., Kössler S., Orawetz T., Matthes U., Orzechowski S., Koch A., Fettke J. Identification of two *Arabidopsis thaliana* plasma membrane transporters able to transport glucose 1-phosphate. *Plant Cell Physiol.* 2020;61(2):381-392. DOI 10.1093/pcp/pcz206.

McKenzie M.J., Chen R.K., Harris J.C., Ashworth M.J., Brummell D.A. Post-translational regulation of acid invertase activity by vacuolar invertase inhibitor affects resistance to cold-induced sweetening of potato tubers. *Plant Cell Environ.* 2013;36(1):176-185. DOI 10.1111/j.1365-3040.2012.02565.x.

Mérida A., Fettke J. Starch granule initiation in *Arabidopsis thaliana* chloroplasts. *Plant J.* 2021;107(3):688-697. DOI 10.1111/tpj.15359.

Müller-Röber B.T., Kossmann J., Hannah L.C., Willmitzer L., Sonnewald U. One of two different ADP-glucose pyrophosphorylase genes from potato respond strongly to elevated levels of sucrose. *Mol. Gen. Genet.* 1990;224:136-146. DOI 10.1007/BF00259460.

Muth J., Hartje S., Twyman R.M., Hofferbert H.R., Tacke E., Prüfer D. Precision breeding for novel starch variants in potato. *Plant Biotechnol. J.* 2008;6(6):576-584. DOI 10.1111/j.1467-7652.2008.00340.x.

Nielsen T.H., Deiting U., Stitt M. A β-amylase in potato tubers is induced by storage at low temperature. *Plant Physiol*. 1997;113(2):503-510. DOI 10.1104/pp.113.2.503.

Orzechowski S., Sitnicka D., Grabowska A., Compart J., Fettke J., Zdunek-Zastocka E. Effect of short-term cold treatment on carbohydrate metabolism in potato leaves. *Int. J. Mol. Sci*. 2021; 22(13):7203. DOI 10.3390/ijms22137203.

Pfister B., Zeeman S.C. Formation of starch in plant cells. *Cell Mol. Life Sci*. 2016;73(14):2781-2807. DOI 10.1007/s00018-016-2250-x.

Qu J., Xu S., Zhang Z., Chen G., Zhong Y., Liu L., Zhang R., Xue J., Guo D. Evolutionary, structural and expression analysis of core genes involved in starch synthesis. *Sci. Rep.* 2018;8:12736. DOI 10.1038/s41598-018-30411-y.

Reyniers S., Ooms N., Gomand S.V., Delcour J.A. What makes starch from potato (*Solanum tuberosum* L.) tubers unique: a review. *Compr. Rev. Food Sci. Food Saf.* 2020;19(5):2588-2612. DOI 10.1111/1541-4337.12596.

Ritte G., Heydenreich M., Mahlow S., Haebel S., Kötting O., Steup M. Phosphorylation of C6- and C3-positions of glucosyl residues in starch is catalysed by distinct dikinases. *FEBS Lett*. 2006;580(20):4872-4876. DOI 10.1016/j.febslet.2006.07.085.

Roldán I., Wattebled F., Lucas M.M., Delvallé D., Planchot V., Jiménez S., Pérez R., Ball S., D'Hulst C., Mérida A. The phenotype of soluble starch synthase IV defective mutants of *Arabidopsis thaliana* suggests a novel function of elongation enzymes in the control of starch granule formation. *Plant J.* 2007;49:492-504. DOI 10.1111/j.1365-313X.2006.02968.x.

Ruan Y.L. Sucrose metabolism: gateway to diverse carbon use and sugar signaling. *Annu. Rev. Plant Biol.* 2014;65:33-67. DOI 10.1146/annurev-arplant-050213-040251.

Rüscher D., Corral J.M., Carluccio A.V., Klemens P.A.W., Gisel A., Stavolone L., Neuhaus H.E., Ludewig F., Sonnewald U., Zierer W. Auxin signaling and vascular cambium formation enable storage metabolism in cassava tuberous roots. *J. Exp. Bot.* 2021;72(10):3688-3703. DOI 10.1093/jxb/erab106.

Salanoubat M., Belliard G. The steady-state level of potato sucrose synthase mRNA is dependent on wounding, anaerobiosis and sucrose concentration. *Gene.* 1989;84:181-185. DOI 10.1016/0378-1119(89)90153-4.

Samodien E., Jewell J.F., Loedolff B., Oberlander K., George G.M., Zeeman S.C., Damberger F.F., van der Vyver C., Kossmann J., Lloyd J.R. Repression of *Sex4* and *Like Sex Four2* orthologs in potato increases tuber starch bound phosphate with concomitant alterations in starch physical properties. *Front. Plant Sci*. 2018;9:1044. DOI 10.3389/fpls.2018.01044.

Santana A.L., Meireles M.A.A. New starches are the trend for industry applications: a review. *Food Public Health*. 2014;4(5):229-241. DOI 10.5923/j.fph.20140405.04.

Santelia D., Kötting O., Seung D., Schubert M., Thalmann M., Bischof S., Meekins D.A., Lutz A., Patron N., Gentry M.S., Allain F.H., Zeeman S.C. The phosphoglucan phosphatase Like Sex Four2 dephosphorylates starch at the C3-position in *Arabidopsis*. *Plant Cell*. 2011;23(11):4096-4111. DOI 10.1105/tpc.111.092155.

Scheidig A., Fröhlich A., Schulze S., Lloyd J.R., Kossmann J. Downregulation of a chloroplast-targeted β-amylase leads to a starch-excess phenotype in leaves. *Plant J.* 2002;30(5):581-591. DOI 10.1046/j.1365-313x.2002.01317.x.

Schönhals E.M., Ding J., Ritter E., Paulo J.M., Cara N., Tacke E., Hofferbert H.R., Lübeck J., Strahwald J., Gebhardt C. Physical mapping of QTL for tuber yield, starch content and starch yield in tetraploid potato (*Solanum tuberosum* L.) by means of genome wide genotyping by sequencing and the 8.3 K SolCAP SNP array. *BMC Genomics*. 2017;18:642. DOI 10.1186/s12864-017-3979-9.

Schwall G.P., Safford R., Westcott R.J., Jeffcoat R., Tayal A., Shi Y.C., Gidley M.J., Jobling S.A. Production of very-high-amylose potato starch by inhibition of SBE A and B. *Nat. Biotechnol*. 2000;18(5):551-554. DOI 10.1038/75427.

Seung D., Boudet J., Monroe J., Schreier T.B., David L.C., Abt M., Lu K.J., Zanella M., Zeeman S.C. Homologs of PROTEIN TARGETING TO STARCH control starch granule initiation in Arabidopsis leaves. *Plant Cell*. 2017;29(7):1657-1677. DOI 10.1105/tpc.17.00222.

Seung D., Soyk S., Coiro M., Maier B.A., Eicke S., Zeeman S.C. PROTEIN TARGETING TO STARCH is required for localising GRANULE-BOUND STARCH SYNTHASE to starch granules and for normal amylose synthesis in Arabidopsis. *PLoS Biol*. 2015;13(2):e1002080. DOI 10.1371/journal.pbio.1002080.

Slugina M.A., Kochieva E.Z. The use of carbohydrate metabolism genes for potato (*Solanum tuberosum* L.). *Selskokhozyaystvennaya Biologiya = Agricultural Biology.* 2018;53(3):450-463. DOI 10.15389/agrobiology.2018.3.450eng.

Sonnewald S., Sonnewald U. Regulation of potato tuber sprouting. *Planta*. 2014;239(1):27-38. DOI 10.1007/s00425-013-1968-z.

Sowokinos J.R., Hayes R.J., Thill C.A. Coordinated regulation of cold induced sweetening in tetraploid potato families by isozymes of UDP-glucose pyrophosphorylase and vacuolar acid invertase. *Am. J. Potato Res.* 2018;95:487-494. DOI 10.1007/s12230-018-9653-1.

Sowokinos J.R., Preiss J. Pyrophosphorylases in *Solanum tuberosum*: III. Purification, physical, and catalytic properties of ADPglucose pyrophosphorylase in potatoes. *Plant Physiol*. 1982;69(6):1459-1466. DOI 10.1104/pp.69.6.1459.

Stein O., Granot D. An overview of sucrose synthases in plants. *Front. Plant Sci.* 2019;10:95. DOI 10.3389/fpls.2019.00095.

Stinard P.S., Robertson D.S., Schnable P.S. Genetic isolation, cloning, and analysis of a mutator-induced, dominant antimorph of the maize *amylose extender1* locus. *Plant Cell*. 1993;5(11):1555-1566. DOI 10.1105/tpc.5.11.1555.

Streb S., Delatte T., Umhang M., Eicke S., Schorderet M., Reinhardt D., Zeeman S.C. Starch granule biosynthesis in *Arabidopsis* is abolished by removal of all debranching enzymes but restored by the subsequent removal of an endoamylase. *Plant Cell*. 2008;20(12):3448-3466. DOI 10.1105/tpc.108.063487.

Streb S., Zeeman S.C. Starch metabolism in Arabidopsis. *Arabidopsis Book*. 2012;10:e0160. DOI 10.1199/tab.0160.

Sweetlove L.J., Fernie A.R. The spatial organization of metabolism within the plant cell. *Annu. Rev. Plant Biol*. 2013;64:723-746. DOI 10.1146/annurev-arplant-050312-120233.

Tauberger E., Fernie A.R., Emmermann M., Renz A., Kossmann J., Willmitzer L., Trethewey R.N. Antisense inhibition of plastidial phosphoglucomutase provides compelling evidence that potato tuber amyloplasts import carbon from the cytosol in the form of glucose-6-phosphate. *Plant J.* 2000;23(1):43-53. DOI 10.1046/j.1365-313x.2000.00783.x.

Tenorio G., Orea A., Romero J.M., Mérida A. Oscillation of mRNA level and activity of granule-bound starch synthase I in *Arabidopsis* leaves during the day/night cycle. *Plant Mol. Biol*. 2003;51(6):949-958. DOI 10.1023/a:1023053420632.

Tetlow I.J., Beisel K.G., Cameron S., Makhmoudova A., Liu F., Bresolin N.S., Wait R., Morell M.K., Emes M.J. Analysis of protein complexes in wheat amyloplasts reveals functional interactions among starch biosynthetic enzymes. *Plant Physiol*. 2008;146(4):1878-1891. DOI 10.1104/pp.108.116244.

Tetlow I.J., Bertoft E. A review of starch biosynthesis in relation to the building block-backbone model. *Int. J. Mol. Sci.* 2020;21:7011. DOI 10.3390/ijms21197011.

Tiessen A., Hendriks J.H.M., Stitt M., Branscheid A., Gibon Y., Farré E.M., Geigenberger P. Starch synthesis in potato tubers is regulated by post-translational redox modification of ADP-glucose pyrophosphorylase: a novel regulatory mechanism linking starch

synthesis to the sucrose supply. *Plant Cell.* 2002;14:2191-2213. DOI 10.1105/tpc.003640.

Tjaden J., Möhlmann T., Kampfenkel K., Neuhaus E. Altered plastidic ATP/ADP-transporter activity influences potato (*Solanum tuberosum* L.) tuber morphology, yield and composition of tuber starch. *Plant J*. 2001;16:531-540. DOI 10.1046/j.1365-313x.1998.00317.x.

Tsai H.-L., Lue W.-L., Lu K.-J., Hsieh M.-H., Wang S.-M., Chen J. Starch synthesis in Arabidopsis is achieved by spatial cotranscription of core starch metabolism genes. *Plant Physiol*. 2009;151(3):1582-1595. DOI 10.1104/pp.109.144196.

Tuncel A., Corbin K.R., Ahn-Jarvis J., Harris S., Hawkins E., Smedley M.A., Harwood W., Warren F.J., Patron N.J., Smith A.M. Cas9-mediated mutagenesis of potato starch-branching enzymes generates a range of tuber starch phenotypes. *Plant Biotechnol. J.* 2019;17(12):2259-2271. DOI 10.1111/pbi.13137.

Van Harsselaar J.K., Lorenz J., Senning M., Sonnewald U., Sonnewald S. Genome-wide analysis of starch metabolism genes in potato (*Solanum tuberosum* L.). *BMC Genomics.* 2017;18(1):37. DOI 10.1186/s12864-016-3381-z.

Visser R.G., Somhorst I., Kuipers G.J., Ruys N.J., Feenstra W.J., Jacobsen E. Inhibition of the expression of the gene for granule-bound starch synthase in potato by antisense constructs. *Mol. Gen. Genet*. 1991;225(2):289-296. DOI 10.1007/BF00269861.

Wang D., Cheng L., Wang Y., Zhang F. Comparative proteomic analysis of potato (*Solanum tuberosum* L.) tuberization *in vitro* regulated by IAA. *Am. J. Potato Res*. 2018;95(4):395-412. DOI 10.1007/s12230-018-9640-6.

Wang X., Xue L., Sun J., Zuo J. The *Arabidopsis BE1* gene, encoding a putative glycoside hydrolase localized in plastids, plays crucial roles during embryogenesis and carbohydrate metabolism. *J. Integr. Plant Biol.* 2010;52:273-288. DOI 10.1111/j.1744-7909.2010.00930.x.

Weise S.E., Schrader S.M., Kleinbeck K.R., Sharkey T.D. Carbon balance and circadian regulation of hydrolytic and phosphorolytic breakdown of transitory starch. *Plant Physiol*. 2006;141(3):879-886. DOI 10.1104/pp.106.081174.

Yao D., Gonzales-Vigil E., Mansfield S.D. Arabidopsis sucrose synthase localization indicates a primary role in sucrose translocation in phloem. *J. Exp. Bot*. 2020;71(6):1858-1869. DOI 10.1093/jxb/erz539.

Yu T.S., Lue W.L., Wang S.M., Chen J. Mutation of *Arabidopsis* plastid phosphoglucose isomerase affects leaf starch synthesis and floral initiation. *Plant Physiol.* 2000;123(1):319-326. DOI 10.1104/pp.123.1.319.

Yu T.S., Zeeman S.C., Thorneycroft D., Fulton D.C., Dunstan H., Lue W.L., Hegemann B., Tung S.Y., Umemoto T., Chapple A., Tsai D.L., Wang S.M., Smith A.M., Chen J., Smith S.M. α-Amylase is not required for breakdown of transitory starch in *Arabidopsis* leaves. *J. Biol. Chem*. 2005;280(11):9773-9779. DOI 10.1074/jbc.M413638200.

Zeeman S.C., Kossmann J., Smith A.M. Starch: its metabolism, evolution, and biotechnological modification in plants. *Annu. Rev. Plant Biol*. 2010;61:209-234. DOI 10.1146/annurev-arplant-042809-112301.

Zhang H., Liu J., Hou J., Yao Y., Lin Y., Ou Y., Song B., Xie C. The potato amylase inhibitor gene *SbAI* regulates cold-induced sweetening in potato tubers by modulating amylase activity. *Plant Biotechnol. J.* 2014;12(7):984-993. DOI 10.1111/pbi.12221.

Zhang L., Häusler R.E., Greiten C., Hajirezaei M.R., Haferkamp I., Neuhaus H.E., Flügge U.I., Ludewig F. Overriding the co-limiting import of carbon and energy into tuber amyloplasts increases the starch content and yield of transgenic potato plants. *Plant Biotechnol. J.* 2008;6(5):453-464. DOI 10.1111/j.1467-7652.2008.00332.x.

Zrenner R., Salanoubat M., Willmitzer L., Sonnewald U. Evidence of the crucial role of sucrose synthase for sink strength using transgenic potato plants (*Solanum tuberosum* L.). *Plant J*. 1995;7(1):97-107. DOI 10.1046/j.1365-313x.1995.07010097.x.

ORCID ID
E.M. Sergeeva orcid.org/0000-0001-8234-3063
E.A. Salina orcid.org/0000-0001-8590-847X
A.V. Kochetov orcid.org/0000-0003-3151-5181

Acknowledgements. This work was supported by State Budgeted Project 0259-2021-0012.
Conflict of interest. The authors declare no conflict of interest.
Received October 13, 2021. Revised December 20, 2021. Accepted December 20, 2021.

Вавиловский журнал генетики и селекции. 2022;26(3):264-271
DOI 10.18699/VJGB-22-33

Original Russian text www.bionet.nsc.ru/vogis/

Forecast for the zone of viticulture in European Russia under climate change

L.Yu. Novikova ✉, P.V. Ozerski

Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), St. Petersburg, Russia
✉ l.novikova@vir.nw.ru

Abstract. Climate warming has turned out to be a significant factor in viticulture and winemaking in all grape-growing areas of the world. Many countries consider the advance of viticulture to the north and to mountainous areas as a possible way to adapt to warming. The factors limiting the zone of viticulture in Russia have been identified by Soviet scientist F.F. Davitaya in 1948, and they are still relevant. They are the sum of active temperatures above 10 °C ($\Sigma T_{10} > 2500$ °C), mean of absolute minimum temperatures ($T_{min} > -35$ °C), length of the frost-free period ($L_{ff} < 150$ days), and hydrothermal coefficient ($0.5 < HTC < 2.5$). The values of these limiting factors in the present-day zone of commercial viticulture (ZCV) correspond to the ranges defined by F.F. Davitaya, with the exception of T_{min}, which in the modern ZCV in European Russia is above –26 °C everywhere. The objective of this work was to assess the possibility of moving the boundaries of the ZCV to the north under the existing and predicted climate conditions in European Russia. The 1980–2019 daily data from 150 weather stations of the Federal Service for Hydrometeorology and Environmental Monitoring were used to calculate mean long-term values, trends and forecasts for 2050 for the ZCV limiting factors and locate the points lying in the range acceptable for viticulture. The QGIS program was applied to plot the points on the European Russia map and mark the terminal latitude. Versions with $T_{min} > -26$ °C and $T_{min} > -35$ °C were considered. On average for European Russia, in 1980–2019, there was an increase in ΣT_{10}, T_{min}, and L_{ff} and a decrease in HTC. However, in the same period, T_{min} showed a tendency toward decreasing at a number of points at latitudes lower than 55° N. The increase in heat supply during the growing season in European Russia implies a possibility of expanding the ZCV northward, beyond the present-day terminal latitude of 46.6° N, to 51.8° N under the existing conditions, and up to 60.7° N by 2050. In addition, even under the current conditions viticulture is possible in the area of Kaliningrad (54° N, 20° E). Using extra protective measures in winters not colder than –35 °C would make it possible to grow grapes at up to 53.3° N under the current conditions and at up to 60.7° N under the prognosticated ones. At the same time, a possible decrease in the minimum winter temperature at the south of European Russia will require additional protective measures in winter, while an increase in the aridity of the climate on the northwest coast of the Caspian Sea will reduce the area under non-irrigated vineyards.

Key words: viticulture; climatic limiting factors; climate change; trends; forecast; GIS.

For citation: Novikova L.Yu., Ozerski P.V. Forecast for the zone of viticulture in European Russia under climate change. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2022;26(3):264-271. DOI 10.18699/VJGB-22-33

Прогноз зоны возделывания винограда на европейской территории России в условиях изменения климата

Л.Ю. Новикова ✉, П.В. Озерский

Федеральный исследовательский центр Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова (ВИР), Санкт-Петербург, Россия
✉ l.novikova@vir.nw.ru

Аннотация. Потепление климата оказалось существенным фактором для виноградарства и виноделия всех виноградарских районов мира. Многие страны рассматривают продвижение виноградарства на север и в горные районы как возможный путь адаптации к потеплению. Факторы, лимитирующие зону виноградарства в России, определены советским ученым Ф.Ф. Давитая в 1948 г. и актуальны до сих пор. Это сумма активных температур выше 10 °C ($\Sigma T_{10} > 2500$ °C), средний из абсолютных минимумов температуры ($T_{min} > -35$ °C), продолжительность безморозного периода ($L_{ff} < 150$ сут) и гидротермический коэффициент ($0.5 < ГТК < 2.5$). Значения лимитирующих факторов современной зоны промышленного виноградарства (ЗПВ) соответствуют определенным Ф.Ф. Давитая диапазонам, за исключением T_{min}, которая в современной ЗПВ на европейской территории России везде выше –26 °C. Целью исследования было определение возможности продвижения на север границ зоны промышленного виноградарства в современных и прогнозируемых климатических условиях европейской территории России. По суточным данным 1980–2019 гг. для 150 метеостанций Росгидромета рассчитали среднемноголетние значения, тренды и прогнозы к 2050 г. значений лимитирующих факто-

ров ЗПВ, определили точки, лежащие в допустимом для виноградарства диапазоне. В программе QGIS нанесли точки на карту европейской территории России, определили предельную широту. Были рассмотрены варианты с $T_{min} > -26$ °C и $T_{min} > -35$ °C. В 1980–2019 гг. в среднем на европейской территории России наблюдался рост ΣT_{10}, T_{min}, L_{ff} и снижение ГТК. Однако южнее 55° N в ряде точек прослеживалась тенденция к снижению T_{min}. Рост теплообеспеченности вегетационного периода на европейской территории России создает предпосылки продвижения промышленного виноградарства к северу от современной предельной широты 46.6° до 51.8° в текущих условиях, а к 2050 г. – до 60.7° N. Кроме того, уже сейчас виноградарство возможно в районе Калининграда (54° N, 20° E). При дополнительных мерах по укрытию на зиму до –35 °C виноградарство возможно до 53.3° N в текущих условиях и до 60.7° N – в прогнозируемых. Возможное снижение минимальной температуры зимы на юге европейской территории России потребует дополнительных мер защиты зимой, а повышение засушливости климата на северо-западном побережье Каспийского моря будет уменьшать площади под неорошаемыми виноградниками.

Ключевые слова: виноградарство; лимитирующие климатические факторы; изменения климата; тренды; прогнозы; ГИС.

Introduction

Climate zones suited for high-quality viticulture and winemaking are narrow and greatly depend on the effect of climate change (Hannah et al., 2013; Mozell, Thach, 2014; Santos et al., 2020). A more than 1 °C temperature rise during the small climatic optimum in the 8th–13th centuries led to the advance of the viticulture border in Western and Central Europe to the north by 3–4° N, but from the 15th century the grape-growing area shifted southward (Barash, 1989; Khromov, Petrosyants, 2012). At present, the area under viticulture in the Northern hemisphere is confined between 30° and 50° N, corresponding to the limits of the mean April-to-October temperature of 12–22 °C, or 13–21 °C for high-quality wine production (Schultz, Jones, 2010; Jones, 2012).

According to forecasts, future warming, on the one hand, will produce a beneficial effect on viticulture as a result of the inclusion of new areas, but, on the other hand, will generate serious problems in the areas of traditional viticulture (Roy et al., 2017; Hewer, Brunette, 2020; Vyshkvarkova, Rybalko, 2021). By 2050, the area suitable for viticulture in main winemaking regions is expected to decrease by 19–62 %, as predicted by the RCP 4.5 global climate change scenario, or by 25–73 % according to the RCP 8.5 scenario (Hannah et al., 2013). Contemporary climate change initiates the shift of the zone of commercial viticulture to the north and to mountainous regions (Jones, 2012; Vršič, Vodovnik, 2012; Hannah et al., 2013; Mozell, Thach, 2014; Quénol et al., 2014). Russia is among the countries that may face expressly significant consequences of climate warming (Houtan et al., 2021).

To predict the impact of climate change on the efficiency of viticulture in grape-growing regions, their climate resources are assessed using various indicators, such as the sums of active and effective temperatures, biologically active sums of effective temperatures, Winkler index, mean April-to-October temperatures, spring frost risk index, aridity index, cold index, Huglin and Branas heliothermal indices, Selyaninov's hydrothermal coefficient, etc. (Lorenzo et al., 2013; Blanco-Ward et al., 2019; Rybalko, 2020; Pipan et al., 2021; Vyshkvarkova, Rybalko, 2021). The limiting factors in viticulture are not the same under different environmental and geographic conditions. The most important requirements for grapevine cultivation are temperature and light during the active growing season. In arid regions, rainfall becomes a limiting factor, so irrigation is introduced. Close to the northern boundaries of commercial viticulture, it is limited by winter conditions (Likhovskoi et al., 2016; Roy et al., 2017).

F.F. Davitaya (Davitaya, 1948) analyzed the world's viticulture zones in the first third of the 20th century and made a comprehensive assessment of the range of climatic requirements for grapes, highlighting the characteristics that were relevant for the USSR: the temperature in the beginning and end of the growing season was 10 °C; the sum of temperatures above 10 °C during the growing season (ΣT_{10}) was higher than 2500 °C; the inhibitory high temperature was 35–40 °C; the required minimum mean temperature of the warmest month was 16–18 °C, or 17–19 °C for high-quality wine production; the length of the frost-free period (L_{ff}) was at least 150 days; the mean of absolute minimum temperatures (T_{min}) for uncovered grapevine cultivation was not lower than –15 °C or, with conventional ways of protection from the cold, –35 °C; and Selyaninov's hydrothermal coefficient (HTC) was within the range from 0.5 to 1.5–2.5 (Davitaya, 1948, p. 172–174). This system of indicators is still valid (Mishchenko, 2009; Roy et al., 2017; Hewer, Brunette, 2020).

Geographic areas with a climate that is currently suitable for growing certain crops or may become so in the future are visualized using GIS techniques (Hannah et al., 2013; Nesbitt et al., 2018). Such approach also makes it possible to identify and adjust the parameters of the climatic niche for a species, i. e., the range of agroclimatic parameters under which its development is possible. For this purpose, data from definite geographic points where this species occurs are analyzed (Soberon, Nakamura, 2009; Peterson et al., 2015; Wójtowicz M., Wójtowicz A., 2020).

In Russia, the zone of commercial viticulture (ZCV) is located between the Black, Azov and Caspian seas and in the Crimea at the latitudes of 41.6–46.6° N and the longitudes of 32.5–48.5° E (AgroAtlas, 2008) (Fig. 1). The range of climate characteristics for the ZCV in European Russia (ER) in the early 21st century has mostly remained within the limits outlined by F.F. Davitaya (1948), significantly deviating from them in only one indicator – the minimum winter temperature, which in the present-day ZCV zone does not fall below –26 °C

(Chistyakov, Novikova, 2020). The main limitation for the advance of grapevine cultivation to the north is T_{min}, and it is low HTC that limits its shift to the northwestern coast of the Caspian Sea. If we accept the possibility of cultivation at $T_{min} > -35$ °C, then ΣT_{10} and L_{ff} would become limiting factors in the north.

Our subject of interest is the advance of grapevine to the north, so the limitations imposed by temperatures above 35–40 °C are not discussed (Leewen et al., 2013). July temperatures above 16 °C are observed in ER to the south of 60–63° N; this factor is also not limiting and is not considered below. Thus, ΣT_{10}, L_{ff}, T_{min}, and HTC are the limiting factors for the ZCV in ER when viticulture is moving northward. A deficit in moisture supply (the requirement is HTC > 0.5) limits non-irrigated viticulture on the northeastern coast of the Caspian Sea.

The objective of this work was to assess the possibility of moving the boundaries of the ZCV to the north under the existing and predicted climate conditions in ER.

Material and method

Conventionally, the ER territory is regarded here as limited by 63° N and 60° E. The software product applied was QGIS 3.22.0[1]. The analysis of the climate in ER was performed pointwise, according to the data of 150 weather stations of the Federal Service for Hydrometeorology and Environmental Monitoring, with more than 20 years of observations in the period of 1980–2019. We used daily data from an open web source (RIHMI – World Data Center, 2020)[2]. The VITIS TIME SERIES program (Novikova, Lebedeva, 2019) was used to calculate the values of ΣT_{10}, T_{min}, L_{ff} and HTC for each point for each year and their trends for the period of 1980–2019.

Values average for 1980–2019 were attributed to 2000, individual ΣT_{10}, L_{ff}, T_{min} and HTC forecasts for 2050 were calculated for each point, and points where grapevine cultivation is possible were identified according to the set of requirements proposed by F.F. Davitaya (1948) for protected viticulture and taking into account specific features of the modern Russian viticulture with $T_{min} < -26$ °C. The study adopted a significance level of 5 %.

Results

Changes of climatic factors important for viticulture in ER during 1980–2019

In 1980–2019, on average, ER showed an increase in ΣT_{10}, T_{min} and L_{ff}, and a decrease in HTC. The average trend for 150 stations was: $\Delta\Sigma T_{10} = 11.52$ °C/year, $\Delta T_{min} = 0.02$ °C/year, $\Delta L_{ff} = 0.31$ days/year, and ΔHTC = –0.01 units/year (see the Table).

ΣT_{10} increased at all examined points: in 144 out of 150 locations it was significant. On average, the intensity of summer warming decreased with the latitude: the correlation between the growth rate ΣT_{10} and latitude was $r = -0.51$. The L_{ff}, T_{min} and HTC trends had both positive and negative values (see the Table, Fig. 1).

With the average tendency for the studied stations to increase T_{min}, the trend was zero or negative at 48 points. There were only 8 statistically significant T_{min} trends, and two of them were negative. The T_{min} trend increased with the latitude ($r = 0.52$), i. e., warming in winter was more intense in the north of ER; out of 61 points located to the north of 55° N, a negative trend was observed only at three points (see Fig. 1, *b*).

L_{ff} increased on average; however, negative trends were observed at 25 points out of 150 studied. There were only 40 significant trends, and positive values were registered in 39 of them.

With the average tendency toward a decrease in HTC, induced by an active increase in temperatures and, on average, the absence of a tendency toward changes in rainfall, the HTC values increased at 8 points. The HTC trends were significant at 20 points, all of them being negative.

Potential zone of commercial viticulture

Under the present-day climate conditions, defined as the average values of the limiting factors for the period of 1980–2019, grapevine cultivation without irrigation and with usual protective measures for the winter season is possible at 36 points out of 150 studied (Fig. 2, *a*), including one point in Kaliningrad Province (Baltiysk, in 1980–2019: $\Sigma T_{10} = 2567$ °C, $L_{ff} = 229$ days, $T_{min} = -15$ °C, HTC = 1.2). If we add the points where insufficient moisture (HTC < 0.5) can be compensated by irrigation, then their number will reach 41 (see Fig. 2, *b*). If we add areas with winter temperature minima reaching –35 °C, then the number of points will increase to 58 (see Fig. 2, *c*). The heat supply during the growing season in ER implies a possibility to move grapevine cultivation to the north of the current terminal latitude of the ZCV (46.6° N) up to 51.8° N even now. With additional sheltering measures for the winter season with temperatures down to –35 °C, it can be moved to 53.3° N.

By 2050, wintering conditions are expected to worsen in the south of ER, i. e., in a number of places the minimum winter temperature will drop below –26 °C. However, due to the movement of heat to the north, the number of points suitable for commercial grapevine cultivation will increase to 43. Additional viticultural practice measures will increase this number even more – up to 56, with the inclusion of irrigated vineyards (see Fig. 2, *e*), and up to 95, with the inclusion of points with temperatures in winter above –35 °C (see Fig. 2, *f*). The increase of heat supply during the growing season creates the prerequisites for the advance of viticulture to 60.7° N by 2050. St. Petersburg may enter the zone of viticulture: by 2050, the forecasts for the area are: $\Sigma T_{10} = 2772$ °C, $T_{min} = -18$ °C, $L_{ff} = 191$ days, and HTC = 1.4. Without the Baltic Sea coast, the northernmost point among the forecasted ones is located on the latitude of 58.1° N.

Discussion

The pace of climate change and anomalies in the harvests of the world's staple crops are on the rise (Jägermeyr et al., 2021). However, the changes in agroclimatic indicators have regional specificities (Sirotenko et al., 2013; Hewer, Brunette, 2020), which was confirmed in the process of this study. Since the 1970s, the sums of active temperatures have been increasing in all ER regions, and more actively in the south; however, the January temperature increases faster in northern

[1] QGIS. A Free and Open Source Geographic Information System. https://qgis.org/ru/site (Accessed August 20, 2021).

[2] RIHMI – World Data Center. http://www.meteo.ru (Accessed June 7, 2020).

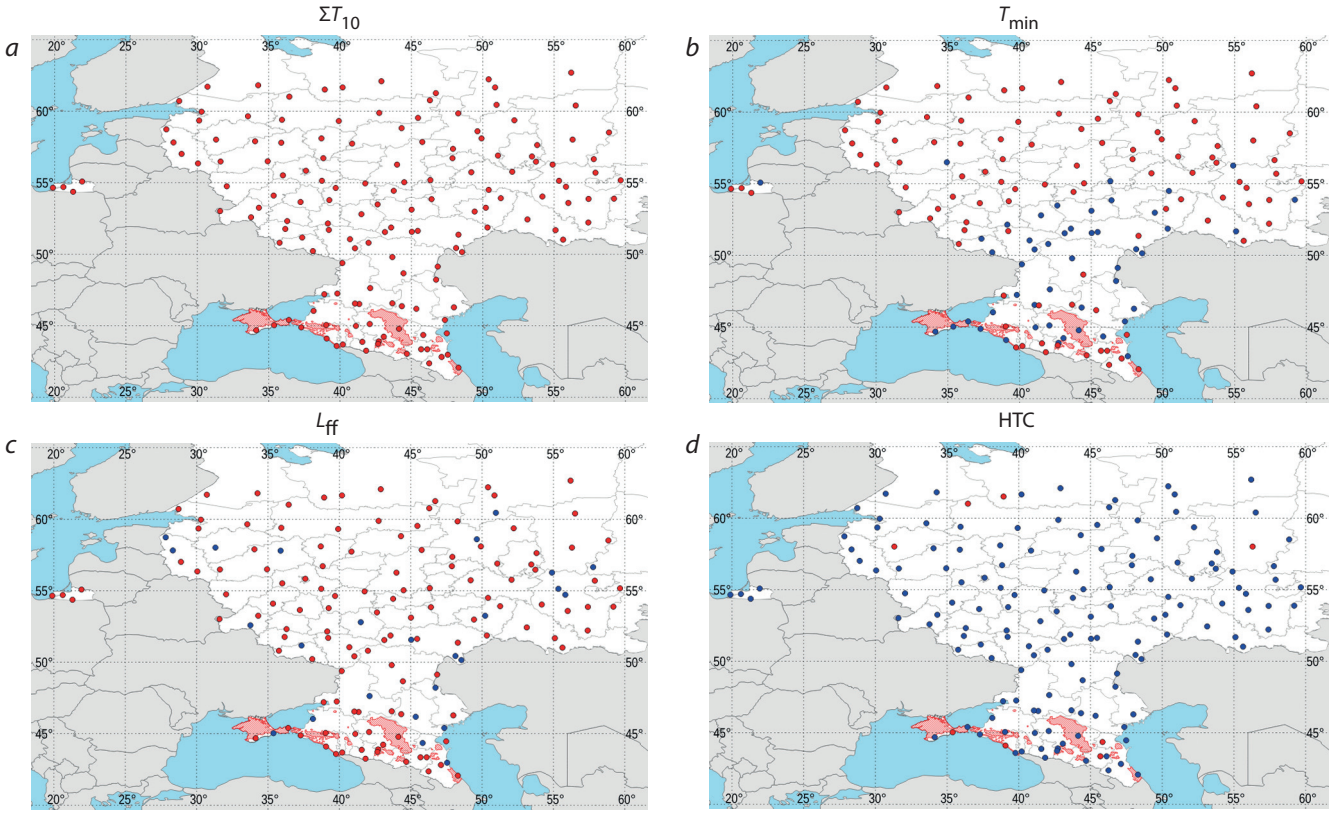


Fig. 1. Trends of the limiting factors for viticulture recorded at 150 weather stations in European Russia: *a*, the sum of temperatures above 10 °C; *b*, the absolute minimum of temperature; *c*, the length of the frost-free period; *d*, HTC.

Red dots are positive trends, and blue dots are zero or negative trends. The zone of commercial viticulture in the early 21st century is shaded. The ZCV map was taken from the AgroAtlas resource (AgroAtlas, 2008) and modified.

Trends of the limiting factors in the zone of commercial viticulture at 150 weather stations in ER in 1980–2019

Indicator	Average			Trend, units/year		
	Mean	Min	Max	Mean	Min	Max
Sum of active temperatures above 10 °C	2533.8	365.1	4333.1	11.52	0.62	25.01
Absolute minimum of temperature, °C	–27.4	–40.2	–4.6	0.02	–0.16	0.19
Length of the frost-free period, days	156.7	79.7	295.3	0.31	–0.57	1.19
HTC	1.2	0.3	3.7	–0.01	–0.02	0.01

latitudes. Total rainfalls have both positive and negative tendencies (Sirotenko et al., 2013). Changes accelerated in the early 21st century: for example, from 1975 to 2004 the HTC reduced in most of ER, with the exception of several regions, and in 1980–2019 we observed that the forecasts concerning the growth of aridity throughout ER came true (Sirotenko, Pavlova, 2009). Warming is accompanied by intensification of climate instability. In 2007, a decrease in the number of winters with threatening drops in air temperature was predicted (The Economics of Climate Change, 2006), but according to the data of 1980–2019 the minimum temperature and the number of dangerous frosts in winter did not decrease despite the rise of the mean winter temperature. The tendency toward an increase in the frequency of extreme events and a decrease in the absolute minimum of annual temperature was observed in many regions of the world under global warming (Bucur et al., 2019).

The study of climate change near the northern border of ER's zone of commercial viticulture – at the Don Ampelographic Collection named after Y.P. Potapenko – also showed that in 1981–2017 the sum of temperatures above 10 °C increased (by 170 °C/10 years), the total rainfall during the period of active growth decreased (by 21 mm/10 years), the mean temperature of the winter dormancy period grew (by

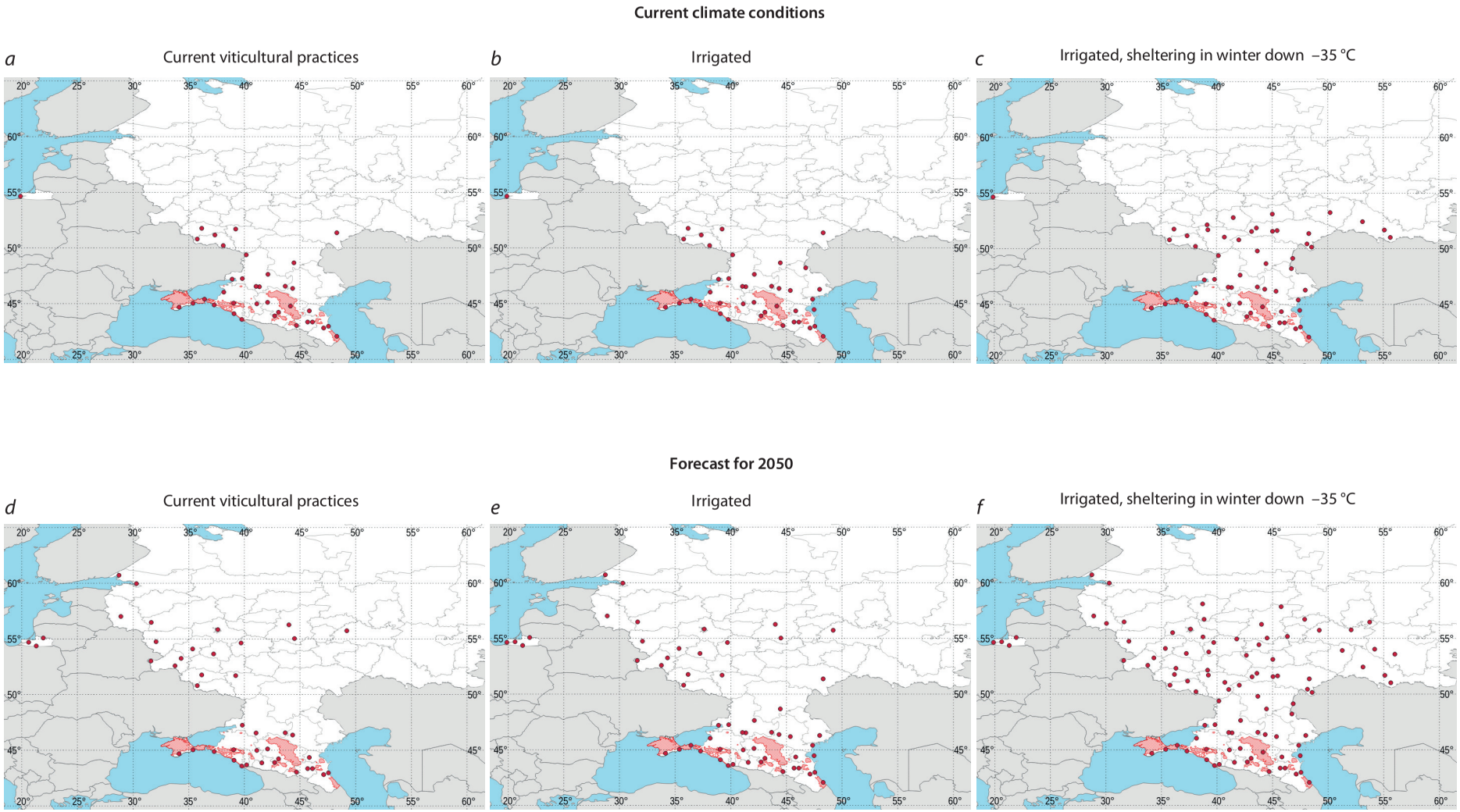


Fig. 2. Points in European Russia where climate conditions are suitable for grapevine cultivation with existing viticultural practices (*a*, *d*), irrigation (*b*, *e*), sheltering in winters down to –35 °C (*c*, *f*) under the present-day conditions and by 2050.

The zone of commercial viticulture in the early 21st century is shaded.

0.5 °C/10 years), the length of the frost-free period extended (by 0.7 days/year), while the number of winter days with temperatures below –20 °C and the minimum winter temperature did not change (Novikova, Naumova, 2018, 2019).

Assessments of the climatic requirements of grapes, made by F.F. Davitaya in the 1930s on the basis of the world's experience in viticulture, turned out to be relevant for ER in the early 21st century. Despite the fact that there are grapevine varieties with temperature requirements of 2100 °C and lower (Mishchenko, 2009; Naumova, Novikova, 2015), the ZCV in ER is limited to temperature sums of 2500 °C, which is explained by the need to have 80–90 % of years with the sums of temperatures necessary for profitable commercial cultivation of a variety (Losev, Zhurina, 2004). There is also concordance between the ranges of other climatic factors, with the exception of the minimum winter temperature: in the contemporary ZCV in ER it is –26 °C versus –35 °C reported by F.F. Davitaya. Values of the limiting factors for the zone of viticulture in the Canadian province of Quebec are also close to F.F. Davitaya's estimates: $L_{ff} > 150$ days, the sum of effective April-to-October temperatures above 10 °C are $DD_{10} >$ 900 °C, $T_{min} > -34$ °C, and the annual number of very cold days ($T < -22$ °C) is less than 30 (Roy et al., 2017). Canadian researchers (Hewer, Brunette, 2020) rank the territories according to the minimum winter temperatures reflecting the degree of suitability for viticulture: –34…–30 °C means poorly suitable conditions; –30…–27 °C medium; –27…–22 °C good, and >–22 °C very good. Thus, the considered options with temperature limits $T_{min} > -35$ °C and $T_{min} > -26$ °C correspond to different degrees of risk and economic efficiency of viticulture. The boundary of $T_{min} > -35$ °C, reported by F.F. Davitaya, possibly corresponds to amateur viticulture.

The ongoing climate change affects all grapevine traits (Vršič, Vodovnik, 2012; Novikova, Naumova, 2019, 2020) and requires adaptation of viticulture and winemaking in all viticultural regions of the world (White et al., 2006; Schultz, Jones, 2010; Jones, 2012; Hannah et al., 2013; Quénol et al., 2014; Bardaji, Iraizoz, 2015). Many countries regard moving northward and into mountainous areas as a possible way for viticulture to adapt to warming (White et al., 2006; Hannah et al., 2013; Schultze et al., 2016; Tóth, Végrári, 2016; Roy et al., 2017; Vyshkvarkova, Rybalko, 2021). Our calculations have shown that a significant advance of viticulture in ER to the north from the current latitude of 46.6° N is possible: even under the existing climate conditions grapevine cultivation could be extended to Kaliningrad and by 2050 to Leningrad Province. During the maximum warming in the 12–13th centuries, well-developed viticulture was underway on the Baltic coast as well as in England (Khromov, Petrosyants, 2012).

However, the trends of 1980–2019 show a decrease in the minimum winter temperature in the southern regions of ER, which may make viticulture less profitable there due to the need for additional winter sheltering measures.

For the main viticulture regions of the world, a decrease in rainfall and an increase in high temperatures (Biasi et al., 2019; Santos et al., 2020) become risk factors and enhance the need for irrigation (Hall et al., 2016; Chrysargyris et al., 2020). The HTC decrease throughout ER is limiting non-irrigated viticulture to the north of the Caspian Sea, where climate aridity will grow. For the rest of ER, moisture conditions remain favorable. A study conducted by Crimean colleagues confirms that viticulture in the vicinity of Sevastopol in the 21st century will be possible without irrigation, but grapevines may experience moisture deficiency (Vyshkvarkova et al., 2021).

Short-term adaptation measures should focus on specific threats, mainly changes in crop management practices (e. g., irrigation, sunscreens to protect leaves, etc.). Further, the change in the composition and taste of grapes and wine will cause regional changes in the assortment of varieties and style of winemaking (Mira de Orduña, 2010; Fraga, Santos, 2017) and stimulate the advance of viticulture to the northern and mountainous regions.

We did not consider the negative effect of rising high temperatures, since the southern boundary of the ZCV in Russia lies at the latitude of 41.6° N, higher than the southern boundary of the world's viticulture (30° N). However, the threat of excessively high temperatures in the south of the ZCV will remain in the future. In addition, the most important issue of matching the quality of soils in ER to the needs of viticulture has not been considered. These aspects require further research.

Conclusion

The increase of heat supply during the growing season in ER implies a possibility to expand the zone of commercial viticulture northward from the current terminal latitude (46.6° N) up to 51.8° N and by 2050 to 60.7° N. Besides, even under the existing conditions it is possible to develop viticulture in the area of Kaliningrad (54° N, 20° E). Additional winter sheltering measures at temperatures down to –35 °C would make it possible to cultivate grapevine up to 53.3° N under the current conditions and up to 60.7° N under the predicted conditions. The increasing aridity of the climate on the northwestern coast of the Caspian Sea will reduce the area under non-irrigated vineyards. A possible drop of the minimum winter temperature in the south of ER will require additional protective measures during the winter season.

References

AgroAtlas. Interactive Agricultural Ecological Atlas of Russia and Neighboring Countries. Economic Plants and their Diseases, Pests and Weeds. 2008. [Electronic resource]. http://www.agroatlas.ru/en/index.html (Accessed August 20, 2021).

Barash S.I. History of Bad Harvests and Weather in Europe. Leningrad: Hydrometeoizdat Publ., 1989. (in Russian)

Bardaji I., Iraizoz B. Uneven responses to climate and market influencing the geography of high-quality wine production in Europe. *Reg. Environ. Change.* 2015;15:79-92. DOI 10.1007/s10113-014-0623-y.

Biasi R., Brunori E., Ferrara C., Salvati L. Assessing impacts of climate change on phenology and quality traits of *Vitis vinifera* L.: the contribution of local knowledge. *Plants.* 2019;8:121. DOI 10.3390/plants8050121.

Blanco-Ward D., Ribeiro A.C., Barreales D., Castro J., Verdial J., Feliciano M., Viceto C., Rocha A., Carlos C., Silveira C., Miranda A. Climate change potential effects on grapevine bioclimatic indices: a case study for the Portuguese demarcated Douro Region (Portugal). *BIO Web of Conf.* 2019;12:01013. DOI 10.1051/bioconf/20191201013.

Bucur G.M., Cojocaru G., Antoce A.O. The climate change influences and trends on the grapevine growing in Southern Romania: a long-

term study, 42nd World Congress of Vine and Wine. *BIO Web Conf.* 2019;15:01008. DOI 10.1051/bioconf/20191501008.

Chistyakov P.N., Novikova L.Yu. Evaluation of the possibility of moving northward of the zone of grape cultivation in the ETR. In: Book of abstracts of the All-Russian sci. conf. with international participation "Contribution of Agrophysics to Solving Fundamental Problems of Agricultural Science", St. Petersburg, October 1–2, 2020. St. Petersburg: FGBNU AFI, 2020;275-281. (in Russian)

Chrysargyris A., Xylia P., Litskas V., Stavrinides M., Heyman L., Demeestere K., Höfte M., Tzortzakis N. Assessing the impact of drought stress and soil cultivation in Chardonnay and Xynisteri grape cultivars. *Agronomy.* 2020;10:670. DOI 10.3390/agronomy10050670.

Davitaya F.F. Climatic Zones of Grapes in the USSR. Moscow: Pishchepromizdat Publ., 1948. (in Russian)

Fraga H., Santos J.A. Daily prediction of seasonal grapevine production in the Douro wine region based on favourable meteorological conditions. *Aust. J. Grape Wine Res*. 2017;23:296-304. DOI 10.1111/ajgw.12278.

Hall A., Mathews A.J., Holzapfel B. Potential effect of atmospheric warming on grapevine phenology and post-harvest heat accumulation across a range of climates. *Int. J. Biometeorol.* 2016;60(9):1405-1422. DOI 10.1007/s00484-016-1133-z.

Hannah L., Roehrdanz P.R., Ikegami M., Shepard A.V., Shaw M.R., Tabor G., Zhi L., Marquet P.A., Hijmans R.J. Climate change, wine, and conservation. *Proc. Natl. Acad. Sci. USA*. 2013;110(17):6907-6912. DOI 10.1073/pnas.1210127110.

Hewer M., Brunette M. Climate change impact assessment on grape and wine for Ontario, Canada's appellations of origin. *Reg. Environ. Change.* 2020;20(3):86. DOI 10.1007/s10113-020-01673-y.

Houtan K.S., Tanaka K.R., Gagné T.O., Becker S.L. The geographic disparity of historical greenhouse emissions and projected climate change. *Sci. Adv*. 2021;7:eabe4342. DOI 10.1126/sciadv.abe4342.

Jägermeyr J., Müller C., Ruane A.C., Elliott J., Balkovic J., Castillo O., Faye B., Foster I., Folberth C., Franke J.A., Fuchs K., Guarin J.R., Heinke J., Hoogenboom G., Iizumi T., Jain A.K., Kelly D., Khabarov N., Lange S., Lin T.-S., Liu W., Mialyk O., Minoli S., Moyer E.J., Okada M., Phillips M., Porter C., Rabin S.S., Scheer C., Schneider J.M., Schyns J.F., Skalsky R., Smerald A., Stella T., Stephens H., Webber H., Zabel F., Rosenzweig C. Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. *Nat. Food.* 2021;2:873-885. DOI 10.1038/s43016-021-00400-y.

Jones G. Climate, grapes, and wine: structure and suitability in a changing climate. *Acta Hort.* 2012;931:19-28. DOI 10.17660/ActaHortic.2012.931.1.

Khromov S.P., Petrosyants M.A. Meteorology and Climatology. Moscow: Moscow State University Publ., 2012. (in Russian)

Leewen C., Schultz H., de Cortazar-Atauri I.G., Duchêne E., Ollat N., Pieri P., Bois B., Goutouly J.-P., Quénol H., Touzard J.-M., Malheiro A.C., Bavaresco L., Delrot S. Why climate change will not dramatically decrease viticultural suitability in main wine-producing areas by 2050. *Proc. Natl. Acad. Sci. USA*. 2013;110(33):3051-3052. DOI 10.1073/pnas.1307927110.

Likhovskoi V.V., Zlenko V.A., Volinkin V.A., Oleinikov N.P., Polylyax A.A., Vasylyk I.A., Troshin L.P. Frost resistance of Crimean indigenous grape varieties and their hybrids. *Nauchnyy Zhurnal KubGAU = Scientific Journal of KubSAU.* 2016;117(03):681-694. (in Russian)

Lorenzo M.N., Taboada J.J., Lorenzo J.F., Ramos A.M. Influence of climate on grape production and wine quality in the Rías Baixas, north-western Spain. *Reg. Environ. Change.* 2013;13:887-896. DOI 10.1007/s10113-012-0387-1.

Losev A.P., Zhurina L.L. Agrometeorology. Moscow: KolosS Publ., 2004. (in Russian)

Mira de Orduña R. Climate change associated effects on grape and wine quality and production. *Food Res. Int.* 2010;43:1844-1855. DOI 10.1016/j.foodres.2010.05.001.

Mishchenko Z.A. Agro-climatology. Kiev: KNT Publ., 2009. (in Russian)

Mozell M.R., Thach L. The impact of climate change on the global wine industry: challenges & solutions. *Wine Econ. Policy.* 2014;3(2):81-89. DOI 10.1016/j.wep.2014.08.001.

Naumova L.G., Novikova L.Yu. Temperature analysis of interphase periods of grape varieties of the collection of the All-Russian Scientific Research Institute of Viticulture and Winemaking named after Ya.I. Potapenko. *Vinodelie i Vinogradarstvo = Wine-making and Viticulture.* 2015;5:46-50. (in Russian)

Nesbitt A., Dorling S., Lovett A. A suitability model for viticulture in England and Wales: opportunities for investment, sector growth and increased climate resilience. *J. Land Use Sci*. 2018;13(4):414-438. DOI 10.1080/1747423X.2018.1537312.

Novikova L.Yu., Lebedeva E.G. Certificate of state registration of the computer program 'Program for predicting the reaction of grape varieties to climate change VITIS TIME SERIES' No. 2019664805 dated November 13, 2019. (in Russian)

Novikova L.Yu., Naumova L.G. Regression analysis of winter hardiness of grape cultivars from Ya.I. Potapenko Don ampelographic collection. *Magarach. Vinogradarstvo i Vinodelie = Magarach. Viticulture and Winemaking*. 2018;4:59-61. (in Russian)

Novikova L.Yu., Naumova L.G. Structuring ampelographic collections by phenotypic characteristics and comparing the reaction of grape varieties to climate change. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2019;23(6):772-779. DOI 10.18699/VJ19.551.

Novikova L.Y., Naumona L.G. Dependence of fresh grapes and wine taste scores on the origin of varieties and weather conditions of the harvest year in the northern zone of industrial viticulture in Russia. *Agronomy*. 2020;10(10):1613. DOI 10.3390/agronomy10101613.

Peterson A.T., Papeş M., Soberón J. Mechanistic and correlative models of ecological niches. *Eur. J. Ecol*. 2015;1(2):28-38. DOI 10.1515/eje-2015-0014.

Pipan P., Hall A., Rogiers S.Y., Holzapfel B.P. Accuracy of interpolated versus in-vineyard sensor climate data for heat accumulation modelling of phenology. *Front. Plant Sci.* 2021;12:635299. DOI 10.3389/fpls.2021.635299.

Quénol H., Grosset M., Barbeau G., van Leeuwen C., Hofmann M., Foss C., Irimia L., Rochard J., Boulanger J.-P., Tissot C., Miranda C. Adaptation of viticulture to climate change: high resolution observation of adaptation scenario for viticulture: the ADVICLIM European project. *Bull. de l'OIV*. 2014;87(1001-1002-1003):395-406.

Roy P., Grenier P., Barriault E., Logan T., Blondlot A., Bourgeois G., Chaumont D. Probabilistic climate change scenarios for viticultural potential in Québec. *Clim. Change*. 2017;143(1):43-58. DOI 10.1007/s10584-017-1960-x.

Rybalko E.A. Climatic indices in viticulture. *Magarach. Vinogradarstvo i Vinodelie = Magarach. Viticulture and Winemaking*. 2020;22(1):26-28. DOI 10.35547/iM.2020.22.1.005. (in Russian)

Santos J.A., Fraga H., Malheiro A.C., Moutinho-Pereira J., Dinis L.-T., Correia C., Moriondo M., Leolini L., Dibari C., Costafreda-Aumedes S., Kartschall T., Menz C., Molitor D., Junk J., Beyer M., Schultz H.R. A review of the potential climate change impacts and adaptation options for European viticulture. *Appl. Sci.* 2020;10(9):3092. DOI 10.3390/app10093092.

Schultz H.R., Jones G.V. Climate induced historic and future changes in viticulture. *J. Wine Res*. 2010;21:137-145. DOI 10.1080/09571264.2010.530098.

Schultze S.R., Sabbatini P., Luo L. Effects of a warming trend on cool climate viticulture in Michigan, USA. *SpringerPlus*. 2016;5(1):1119. DOI 10.1186/s40064-016-2777-1.

Sirotenko O.D., Abashina E.V., Pavlova V.N. Dynamics of climate-conditioned changes in heat supply, moisture content and productivity of the agricultural zone of Russia. *Trudy FGBU VNIISHM = Proceedings of the FSBI VNIISHM.* 2013;38:41-53. (in Russian)

Sirotenko O.D., Pavlova V.N. The impact of climate change on agriculture. In: Development of Agricultural Meteorology in Russia. Obninsk, 2009;168-175. (in Russian)

Soberon J., Nakamura M. Niches and distributional areas: concepts, methods and assumptions. *Proc. Natl. Acad. Sci. USA*. 2009;106:19644-19650. DOI 10.1073/pnas.0901637106.

The Economics of Climate Change. The Stern Review. Nicholas Stern. Cabinet Office – HM Treasury, UK, 2006.

Tóth J.P., Végvári Z. Future of wine grape growing regions in Europe. *Aust. J. Grape Wine Res*. 2016;22:64-72. DOI 10.1111/ajgw.12168.

Vršič S., Vodovnik T. Reactions of grape varieties to climate changes in North East Slovenia. *Plant Soil Environ*. 2012;58(1):34-41. DOI 10.17221/352/2011-PSE.

Vyshkvarkova E., Rybalko E. Forecast of changes in air temperatures and heat indices in the Sevastopol region in the 21st century and their impacts on viticulture. *Agronomy*. 2021;11:954. DOI 10.3390/agronomy11050954.

Vyshkvarkova E., Rybalko E., Marchukova O., Baranova N. Assessment of the current and projected conditions of water availability in the Sevastopol region for grape growing. *Agronomy*. 2021;11(8):1665. DOI 10.3390/agronomy11081665.

White M.A., Diffenbaugh N.S., Jones G.V., Pal J.S., Giorgi F. Extreme heat reduces and shifts United States premium wine production in the 21st century. *Proc. Natl. Acad. Sci. USA*. 2006;103:11217-11222. DOI 10.1073/pnas.0603230103.

Wójtowicz M., Wójtowicz A. the effect of climate change on linolenic fatty acid in oilseed rape. *Agronomy.* 2020;10(12):2003. DOI 10.3390/agronomy10122003.

ORCID ID

L.Yu. Novikova orcid.org/0000-0003-4051-3671
P.V. Ozerski orcid.org/0000-0001-7627-5412

Acknowledgements. The work was carried out within the framework of the state task according to the VIR thematic plan on the topic No. 0481-2022-0004 "Improvement of approaches and methods of *ex situ* conservation of the identified gene pool of vegetatively propagated crops and their wild relatives, development of technologies for their effective use in breeding".

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

Received November 16, 2021. Revised January 14, 2022. Accepted January 16, 2022.

Вавиловский журнал генетики и селекции. 2022;26(3):272-280
DOI 10.18699/VJGB-22-34

Transmission of potato spindle tuber viroid between *Phytophthora infestans* and host plants

O.S. Afanasenko ✉, A.V. Khiutti, N.V. Mironenko, N.M. Lashina

All-Russian Research Institute of Plant Protection, Pushkin, St. Petersburg, Russia
✉ olga.s.afan@gmail.com

Abstract. Potato spindle tuber viroid (PSTVd) is a naked, circular, single-stranded RNA (356–363 nucleotides in length) which lacks any protein-coding sequences. It is an economically important pathogen and is classified as a high-risk plant quarantine disease. Moreover, it is known that PSTVd is mechanically transmitted by vegetative plant propagation through infected pollen, and by aphids. The aim of this study is to determine the possibility of viroid transmission by potato pathogen *Phytophthora infestans* (Mont.) de Bary. PSTVd-infected (strain VP87) potato cultivars Gala, Colomba, and Riviera were inoculated with *P. infestans* isolate PiVZR18, and in 7 days, after the appearance of symptoms, re-isolation of *P. infestans* on rye agar was conducted. RT-PCR diagnostics of PSTVd in a mixture of mycelia and sporangia were positive after 14 days of cultivation on rye agar. The PSTVd-infected *P. infestans* isolate PiVZR18v+ was used to inoculate the healthy, viroid-free plants of potato cv. Gala and tomato cv. Zagadka. After 60 days, an amplification fragment of PSTVd was detected in the tissues of one plant of tomato cv. Zagadka by RT-PCR with the primer set P3/P4, indicating successful transmission of PSTVd by *P. infestans* isolate PiVZR18v+. This result was confirmed by sequencing of the RT-PCR amplicon with primers P3/P4. The partial sequence of this amplicon was identical (99.5 %) to PSTVd strain VP87. RT-PCR showed the possibility of viroid stability in a pure culture of *P. infestans* isolate PiVZR18v+ after three consecutive passages on rye agar. PSTVd was not detected after the eighth passage on rye agar in *P. infestans* subculture. These results are initial evidence of potato viroid PSTVd being bidirectionally transferred between *P. infestans* and host plants.
Key words: potato; tomato; PSTVd strains; transmission; *Phytophthora infestans*; RT-PCR detection.

For citation: Afanasenko O.S., Khiutti A.V., Mironenko N.V., Lashina N.M. Transmission of potato spindle tuber viroid between *Phytophthora infestans* and host plants. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2022;26(3):272-280. DOI 10.18699/VJGB-22-34

Трансмиссия вироида веретеновидности клубней картофеля между *Phytophthora infestans* и растениями-хозяевами

О.С. Афанасенко ✉, А.В. Хютти, Н.В. Мироненко, Н.М. Лашина

Всероссийский научно-исследовательский институт защиты растений, Пушкин, Санкт-Петербург, Россия
✉ olga.s.afan@gmail.com

Аннотация. Вироид веретеновидных клубней картофеля (ВВКК) представляет собой кольцевую одноцепочечную РНК длиной 356–363 нуклеотида, в которой отсутствуют какие-либо последовательности, кодирующие белок. ВВКК является экономически значимым заболеванием картофеля, имеющим статус карантинного. Известно, что ВВКК передается механически при вегетативном размножении растений, через инфицированную пыльцу и с помощью тлей. Целью данного исследования было определение возможности передачи ВВКК (штамм VP87) от растений картофеля и томата, зараженных вироидом, патогену *Phytophthora infestans* (Mont.) de Bary и от *P. infestans* – растениям картофеля и томата. Сорта картофеля Гала, Коломба и Ривьера, инфицированные ВВКК, инокулировали изолятом *P. infestans* PiVZR18; через 7 дней после появления симптомов фитофтороза провели повторное выделение *P. infestans* в чистую культуру на ржаной агар. Через 14 дней культивирования *P. infestans* на ржаном агаре в смеси мицелия и спорангиев методом ОТ-ПЦР был обнаружен ВВКК. ВВКК-инфицированным изолятом *P. infestans* (PiPSTVdv+) провели инокуляцию растений томата сорта Загадка и растений картофеля сорта Гала. Через 60 дней в листьях томата сорта Загадка методом ОТ-ПЦР с праймерами P3/P4 был выявлен диагностический продукт амплификации 360 п.о., свидетельствующий об успешной трансмиссии ВВКК изолятом PiVZR18v+. Результаты были подтверждены секвенированием продукта амплификации ВВКК. Нуклеотидная последовательность вироида в растении томата, зараженном изолятом PiVZR18v+, оказалась на 99.5 % идентичной использованному в эксперименте штамму VP87. Для доказательства возможности сохранения вироида в чистой культуре *P. infestans* изолят PiVZR18v+ пассировали на ржаном агаре с интервалом 30 дней. После трех последовательных пассажей на ржаном агаре ВВКК был диагностирован в культуре изолята, что подтверждено секвенированием продукта

амплификации с вироид-специфичными праймерами. В субкультуре *P. infestans* после восьмого пассажа на ржаном агаре ВВКК не обнаруживался. Полученные данные свидетельствуют о двунаправленной передаче ВВКК в патосистеме *P. infestans* – растение-хозяин.
Ключевые слова: картофель; томаты; штаммы ВВКК; трансмиссия; *Phytophthora infestans*; ОТ-ПЦР-диагностика.

Introduction

Potato spindle tuber viroid (PSTVd) is an economically important pathogen, classified as a high-risk plant quarantine disease. According to the European Plant Protection Organization (EPPO), the disease has been reported in 37 countries on all continents (https://gd.eppo.int/taxon/PSTVD0/distribution). In Russia, and other former Soviet Union regions, PSTVd was detected in about 50–70 % of *in vitro* potato plants (Kastalyeva et al., 1992).

Natural infections of PSTVd have been found in the field, mainly in solanaceous crops, such as pepino (Puchta et al., 1990), potato (Diener, Raymer, 1969), and tomato (Puchta et al., 1990). Potato plants infected with PSTVd become smaller and show leaf yellowing, and infected tubers become smaller and cracked. The reduction in tuber weight depends on the viroid strain, potato cultivar, and length of time they have been infected with PSTVd (Pfannenstiel, Slack, 1980). Furthermore, a reduction in tuber yield of up to 24 % has been reported in cultivar Saco infected with mild strains of PSTVd, however, severe strain reduced the yield by up to 64 % (Singh R.P., 1970). In addition to direct losses, it is important to take into account indirect losses that can be significant due to the quarantine status.

PSTVd belongs to Family Pospiviroidae (1POSPF), Genus *Pospiviroid* (1POSPG) and consists of a naked, circular, single-stranded RNA (356–363 nucleotides in length) – the smallest among plant pathogens lacking a protein-coding ability – therefore, it is a parasite of the host transcription mechanism (Yanagisawa et al., 2019).

PSTVd has a wide host range of at least 138 species across 10 families (Singh R.P., 1973). The main hosts are from the Solanaceae family (Owens et al., 1992; Mertelik et al., 2010; Mackie et al., 2016). PSTVd is transmitted mechanically (Verhoeven, Roenhorst, 2010), by aphids (Syller et al., 1997). Moreover, it was found to be vertically transmitted through pollen to progeny seeds on potato (*Solanum tuberosum*) and tomato (*S. lycopersicum*) and horizontally transmitted through infected pollen to other potato and tomato plants (Kryczyński et al., 1988; Singh R.P. et al., 1992; Matsushita, Yanagisawa, 2018).

Viruses are common in fungi and oomycetes and some of these viruses share sequence identities with plant viruses belonging to different families and genera (Mascia et al., 2019). There are several examples of plant–virus transmission by phytopathogenic fungi and oomycetes. It was shown that soil-inhabiting fungi *Olpidium brassicae* and *O. radicale* belonging to Chytridiales and *Polymyxa graminis*, *Spongospora subterranean*, and *Synchytrium endobioticum* – belonging to the order Plasmodiophorales – transmit plant viruses (Bhat, Rao, 2020). Replications of the tobacco mosaic virus were demonstrated in the phytopathogenic fungi *Colletotrichum acutatum*, *C. clavatum*, and *C. theobromicola* (Mascia et al., 2019), cucumber mosaic virus was reported in *Rhizoctonia solani* (Andika et al., 2017), artichoke Italian latent virus, artichoke mottled crinkle virus, potato virus X, potato virus Y, tobacco mosaic virus and cucumber mosaic virus plus its satellite RNA can replicate and persist in *Phytophthora infestans* at least through the first subculture (Mascia et al., 2019).

Wei et al. (2019) obtained preliminary data on the possibilities of replicating hop stunt viroid (HSVd), iresine 1 viroid belonging to the Pospiviroidae and avocado sunblotch viroid (Avsunviroidae) in at least one of phytopathogenic ascomycete fungi *Cryphonectria parasitica*, *Valsa mali*, and *Fusarium graminearum*.

Oomycete *P. infestans* causes significant losses to potato and tomato crops on a global scale. Despite intensive use of fungicides, the pathogen is constantly and ubiquitously present in potato crops. *P. infestans* is also the most harmful and widespread tomato pathogen, both in field and greenhouse conditions. This oomycete has a high adaptive potential to the host plants, which may indicate the formation of competitive relationships with other potato and tomato pathogens. In this regard, it is of interest to identify a possible role of *P. infestans* in transmission of PSTVd to potato and tomato plants.

Materials and methods

Plant materials. Potato cultivars that, according to our data, were susceptible to both PSTVd strain VP87 and *P. infestans* – Gala, Riviera, and Colomba and tomato cultivars Zagadka, Moskvich, and Damskiy Palchik were included in the study.

These potato and tomato cultivars were registered in the Russian State Register of Breeding Achievements.

PSTVd strains. Two intermediate PSTVd strains, VP35 (GenBank accession no. LC523658) and VP87 (LC523667), and severe strain FP10-13 (LC523676) deposited in the international information database DDBJ (DNA Data Bank of Japan), Data set "Viruses" (http://blast.ddbj.nig.ac.jp/) were used in the study. These strains were isolated from infected potato leaves from the Volga Federal District (VP87 and VP35) and tubers from the Far Eastern Federal District (FP10-13) in 2019 (Matsushita et al., 2021).

The PSTVd strains were supported on living tomato plants of Russian cultivars Zagadka, Moskvich, and Damskiy Palchik.

Isolate of *P. infestans*. Isolate PiVZR18 of *P. infestans* was used in the experiments on viroid transmission. PiVZR18 was isolated from the natural population of *P. infestans* in the Leningrad Region (northwest of the European part of Russia) in 2018. Eight virulence genes (1, 2, 3, 4, 6, 7, 10, and 11) were identified in this isolate on a set of Black's differentials (Black et al., 1953).

Viroid inoculation of plants. Potato and tomato plants were grown in a growth room at a temperature of 25 °C with a photoperiod of 16h/8h (day/night) in 2l pots filled with "Terra vita" soil. Seven-day germination potato plants and 14-day tomato plants were used for inoculation by PSTVd.

To prepare the inoculum, 0.1 g of fresh tomato leaf tissue – 60 days post inoculation (dpi) with PSTVd strain VP87 – was

ground in 1 ml sodium phosphate buffer (pH 7.0) and filtered through cheesecloth.

For mechanical inoculation, the cotyledons of tomato were dusted with carborundum and gently rubbed over the surface of the leaves with a plastic pestle. Ten microliters of inoculum was placed on the injured leaf surface and rubbed several times with a sterile plastic pestle. The inoculated plants were incubated for two months at 25 °C with light intensity (fluorescent, 40 W, ×4).

At 60 dpi, the presence of viroid in the inoculated tomato plants was determined by RT-PCR.

To inoculate 7-day potato plants of the cultivars Gala, Riviera, and Colomba, a 0.5–1.0 cm longitudinal stem incision was performed with a sterile razor on a stem apex (Suppl. Fig. 1)[1], and 10 µl of the PSTVd VP87 strain suspension – obtained as described above – was applied. Three plants of each potato cultivar were inoculated and the assay was repeated three times. In 60 dpi, the presence of PSTVd in the inoculated plants of potato cultivars was determined by RT-PCR.

***P. infestans* inoculation of plants.** Isolate PiVZR18 of *P. infestans* was cultured on rye agar medium (1.0 Li ddH_2O, 60.0 g rye organic berries (grind in blender), 20.0 g sucrose, 15.0 g agar) for 30 days in the dark at 15 °C for propagation and morphological observation (Medina, Platt, 1999).

Before inoculating the plants, the suspension was incubated at 12 °C for 2.5–3 h to release zoospores. Upon RT-PCR detection for viroid infection, both healthy and viroid-infected tomato and potato plants were inoculated with a suspension of *P. infestans* at a concentration of 50.000 zoosporangia in 1 ml. After inoculation, the plants were placed in humid chambers with a 14 h light period, at 23 °C during the day and 15 °C at night for a period of 13 dpi. To study PSTVd transmission from *P. infestans* to host plants, when typical symptoms of late blight appeared, the humid chamber was removed and the development of *P. infestans* slowed down. The affected leaves were removed and the plants continued to grow. For the purposes of PSTVd diagnostics, the upper leaves of the plants without late blight symptoms were cut off.

Viroid inoculation of *P. infestans*. Inoculum of PSTVd was obtained from the infected tomato plants as described above and applied to the 14-day *P. infestans* culture by transferring 10 µl per Petri dish (in the dish center). After inoculation, the culture was left to grow for 15 days at 10 °C. Then, mycelia from the periphery and from the center of the colony were transferred separately to a fresh medium. The culture was left in the same conditions for 30 days, after which RT-PCR analysis was conducted.

Isolation of *P. infestans* from infected potato and tomato plants. Seven days post inoculation (dpi) after the symptoms of late blight appeared, *P. infestans* was isolated from the plants. Sections of the infected leaves were placed between tuber slices of the healthy cv. Colomba, and at 6 days, mycelium was transferred with a needle to the surface of rye agar. The isolates were cultured for 30 days at 15 °C in the dark and then transferred to a fresh medium.

Viroid detection and sequencing. We collected the uppermost leaves from the inoculated potato or tomato plants at 60 dpi. PSTVd detection in pure culture of *P. infestans* was carried out after 30 days of growing on rye agar. Approximately 0.1 g of tissue from leaves or mycelium was used for RNA extraction. Total RNA was extracted using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) as per the manufacturer's instructions (http://www.genome.duke.edu/cores/microarray/services/rna-qc/documents/RNeasy_Mini_Handbook.pdf) and subsequently used for one-step RT-PCR. Primer sets P3/P4 (Behjatnia et al., 1996) and P1/P2 (Gross et al., 1978) or 68PV-R+87PV-F (Yanagisawa et al., 2019) were used to detect PSTVd. RT-PCR was prepared with the PrimerScript One-Step RT-PCR Kit ver 2 reagents in 10 ml (Takara Bio Inc., Shiga, Japan) following the manufacturer's instructions. The primer set ITS4/ITS5 (White et al., 1990; Ristaino et al., 1998) was used to detect the ITS region of *P. infestans* as an internal control.

RT-PCR was carried out on a MyCycler Thermal Cycler (Bio-Rad, California, USA) at 50 °C for 30 min, 94 °C for 2 min, followed by 35 cycles of 94 °C for 30 sec, 60 °C for 30 sec and 72 °C for 30 sec. An additional elongation step was performed at 72 °C for 5 min followed by storage at 12 °C. The sizes of the diagnostic fragments of PSTVd and the ITS region were 360 and 946 bp, respectively.

The PSTVd amplicons were sequenced at the Beagle Company (St. Petersburg, Russia). Alignment and manual editing of nucleotide sequences were performed using Vector NTI Advance 10 software (Thermo Fisher Scientific). The obtained nucleotide sequences were tested for similarity with PSTVd strain VP87 (LC523667), used in this study, deposited in the international information database DDBJ (DNA Data Bank of Japan), data set "Viruses" (http://blast.ddbj.nig.ac.jp/).

Results

There is an absence of data on the possibility of replicating PSTVd in *P. infestans* and bidirectionally transferring it between host plants and *P. infestans*. In this study, we investigated the possibility of PSTVd transmission (1) from host plants to *P. infestans*, (2) from *P. infestans* to host plants, and (3) the possibility of PSTVd stability in pure cultures of *P. infestans*.

Transmission of PSTVd from host plants to *P. infestans*

From potato plants. Upon confirming PSTVd infection in tomato plants of cv. Zagadka by RT-PCR, inoculation of three potato cultivars (Gala, Colomba, and Riviera), using as an inoculum source extracted-sap of tomato infected with PSTVd strain VP87, was conducted (see Suppl. Fig. 1). After 60 days, detection of PSTVd presence in plants of these cultivars was carried out by RT-PCR with the P3/P4 primer set (Fig. 1). The brightest amplicons indicating a high accumulation of the viroid were found for cv. Gala and cv. Colomba. The viroid accumulation was lower in three plants of cv. Riviera (see Fig. 1).

PSTVd-infected and uninfected (control) potato plants were inoculated with *P. infestans* isolate PiVZR18. Seven days after the appearance of symptoms (Fig. 2, *a*) caused by *P. infestans*, the pathogen was first re-isolated from these cultivars on tuber slices of healthy cv. Colomba and then transferred to pure culture on rye agar (see Fig. 2).

[1] Supplementary Figures 1–3 are available in the online version of the paper: http://www.bionet.nsc.ru/vogis/download/pict-2022-26/appx5.pdf

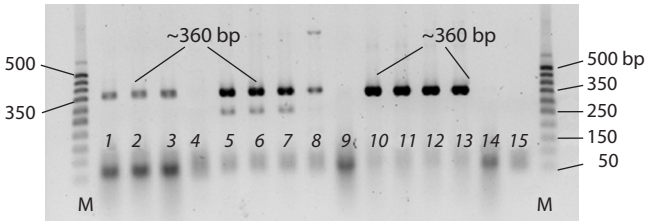


Fig. 1. RT-PCR detection of PSTVd with the primer set P3/P4 in leaves of infected potato plants by PSTVd strain VP87 before inoculation with *P. infestans*.

1–3, Individual plants of cv. Riviera; *4*, mock-inoculated cv. Riviera; *5–8*, individual plants of cv. Colomba; *9*, mock-inoculated cv. Colomba; *10–13*, individual plants of cv. Gala; *14*, mock-inoculated cv. Gala; *15*, negative control (distilled water); M, molecular weight marker 50 bp DNA ladder (Primetech DNA Ladder).

After culturing of *P. infestans* isolates on rye agar, PSTVd detection by RT-PCR was conducted with the primer set P3/P4 (Fig. 3, *a*). Amplicons of ~360 bp – indicating the presence of PSTVd – were detected in a mixture of mycelium and sporangia of pure culture of *P. infestans* isolates after colonization on the cultivars Gala, Riviera, and Colomba (see Fig. 3, *a*). To control for the negative results of viroid detection that are not due to the quality of RNA extracted from *P. infestans* samples, we used PCR with universal primers ITS4/ITS5 on rDNAs, which are species specific for *P. infestans* and displayed an amplicon of 946 bp (Ristaino et al., 1998) (see Fig. 3, *b*).

From tomato plants. The 15 plants of the two tomato cultivars (Moskvich and Damskiy Palchik) infected with three PSTVd strains (VP87, FP10-13, and VP35) were inoculated with the *P. infestans* isolate PiVZR18 to confirm the results obtained. Seven days after the appearance of the symptoms caused by *P. infestans*, 15 cultures of the pathogen were re-isolated from these plants, first on tuber slices of healthy cv. Colomba. After mycelial overgrowth on the surface of tuber slices, the first detection of viroid presence in the mycelium was performed. Out of 15 isolates, the brightest fragment indicating viroid infection of mycelium was detected with the primer set P3/P4 in the *P. infestans* isolate from cv. Moskvich infected with PSTVd strain VP87 (Fig. 4, line *1*). Positive PSTVd detection was also obtained for cv. Damskiy Palchik, infected by PSTVd strain VP87 (line *4*) and for the isolates from cv. Moskvich infected by PSTVd strain FP10-13 (lines *6–8*) and by strain VP35 (lines *10–13*). Weak amplicons were obtained in lines *3*, *5*, *9*, *14*, which indicates a low initial concentration of PSTVd in *P. infestans* mycelium obtained from the host plant (see Fig. 4).

PSTVd transmission from *P. infestans* to host plants

The PSTVd-infected *P. infestans* isolate PiVZR18v+ was used to inoculate the healthy, viroid-free plants of potato cv. Gala and tomato cv. Zagadka.

After 60 days, an amplification fragment of PSTVd was detected in the tissues of one plant of tomato cv. Zagadka by RT-PCR with the primer set P3/P4, indicating successful transmission of PSTVd by *P. infestans* isolate PiVZR18v+ (Fig. 5).

The detection of PSTVd in the RNA of cv. Zagadka inoculated with the *P. infestans* isolate PiVZR18v+ was

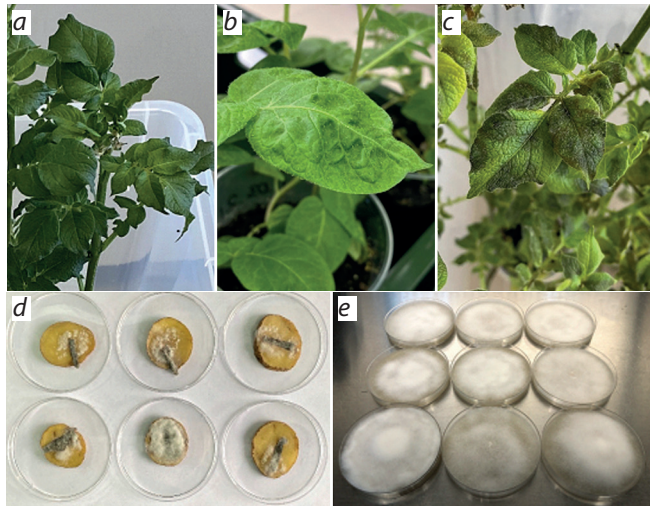


Fig. 2. PSTVd transmission from potato to *P. infestans*.

a, Mock-inoculated (distilled water) potato cv. Colomba; *b*, PSTVd-infected (strain VP87) potato cv. Colomba; *c*, symptoms of *P. infestans* on PSTVd-infected potato cv. Colomba (strain VP87); *d*, tuber slices of healthy cv. Colomba inoculated with *P. infestans* from the infected plants; *e*, *P. infestans* isolates from the PSTVd-infected plants of cv. Colomba on rye agar.

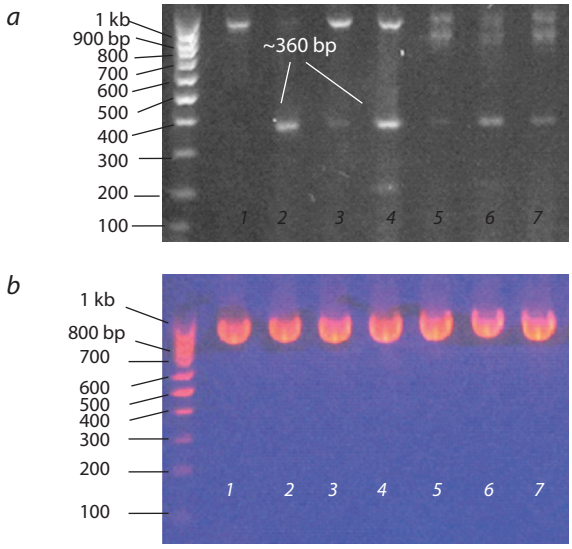


Fig. 3. RT-PCR detection of PSTVd in cultures of *P. infestans* isolated from potato plants infected with the PSTVd strain VP87 (*a*) and control of *P. infestans* RNA presence (*b*).

a, RT-PCR detection with the primer set P3/P4; *b*, RT-PCR amplification of the same RNA samples with the primer set ITS4/ITS5. Lines: *1*, PSTVd-uninfected *P. infestans* isolate PiVZR18; *2*, *P. infestans* isolates from PSTVd-infected potato cv. Gala; *3*, *4*, from cv. Riviera; *5–7*, from cv. Colomba. Left line: molecular weight marker 100 bp (Gene Ruller, Fermentas).

confirmed by sequencing of the RT-PCR amplicon with primers P3/P4 (see Fig. 5, line *3*). The partial sequence (204 bp) of this amplicon was identical (99.5 %) to PSTVd strain VP87 (LC523667) (Suppl. Fig. 2).

Stability of PSTVd in pure culture of *P. infestans*

The stability of strain VP87 in pure culture of *P. infestans* isolate PiVZR18v+ after consecutive passages on rye agar was studied. RT-PCR with primer sets P3/P4 (Fig. 6, *a*) and 68PV/87PV (Fig. 6, *b*) revealed amplification products

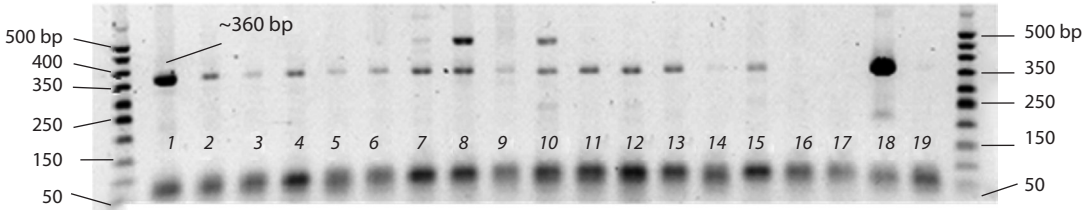


Fig. 4. RT-PCR detection of PSTVd with the primer set P3/P4 in *P. infestans* isolates grown on slices of potato cv. Colomba tubers after their isolation from infected with viroid tomato plants.

1, *2*, Isolates from tomato cv. Moskvich infected by PSTVd strain VP87; *3*, *4*, isolates from tomato cv. Damskiy Palchik, infected by PSTVd strain VP87; *5–9*, isolates from tomato cv. Moskvich, infected by PSTVd strain FP10-13; *10–13*, isolates from tomato cv. Moskvich, infected by PSTVd strain VP35; *14*, *15*, isolates from tomato cv. Damskiy Palchik, infected by PSTVd strain VP35; *16*, isolate PiVZR18 of *P. infestans* uninfected by PSTVd (negative control); *17*, cv. Colomba potato tuber (negative control); *18*, tomato cv. Moskvich infected by PSTVd strain VP87 (positive control); *19*, distilled water (negative control). On the right and left sides of the gel, 50 bp DNA ladders (Primetech DNA Ladder) are shown.

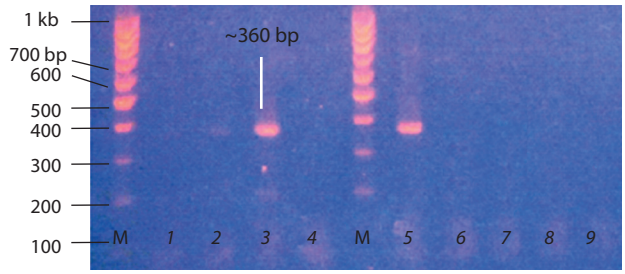


Fig. 5. RT-PCR detection of PSTVd with the primer set P3/P4 in potato and tomato plants 60 dpi with viroid-infected *P. infestans* isolate PiVZR18v+.

1, Potato cv. Gala; *2–4*, tomato cv. Zagadka; *5*, tomato cv. Zagadka infected by PSTVd strain VP87 (positive control); *6*, potato cv. Gala uninfected by PSTVd (negative control); *7*, tomato cv. Zagadka uninfected by PSTVd (negative control); *8*, PSTVd-uninfected *P. infestans* isolate PiVZR18 (negative control); *9*, distilled water (negative control); M, molecular weight marker 100 bp (Gene Ruller, Fermentas).

indicating the presence of PSTVd after the second and third passages on rye agar of *P. infestans* isolates from viroid-infected cv. Colomba (see the Table, Fig. 6). PSTVd stability in *P. infestans* isolates after three passages on rye agar was shown by sequencing of the RT-PCR amplicon with primers P3/P4.

The partial sequence of RT-PCR amplicon (near 232 bp) of viroid in *P. infestans* isolate PiVZRv+ after the third passage on rye agar received with the primer set P3/P4 is identical (98.3 %) to PSTVd strain VP87 (LC523667). Another partial sequence of RT-PCR amplicon (270 bp) received with the primer set 68PV/87PV of the same RNA sample is identical to PSTVd strain VP87 (LC523667) – 99.3 % (Suppl. Fig. 3).

On the other hand, PSTVd was not detected after the eighth passage on rye agar in *P. infestans* subculture (see the Table).

P. infestans isolates infected with viroid strain VP87 were characterized by more abundant, but also more compact mycelium, forming an almost felt-like colony (Fig. 7).

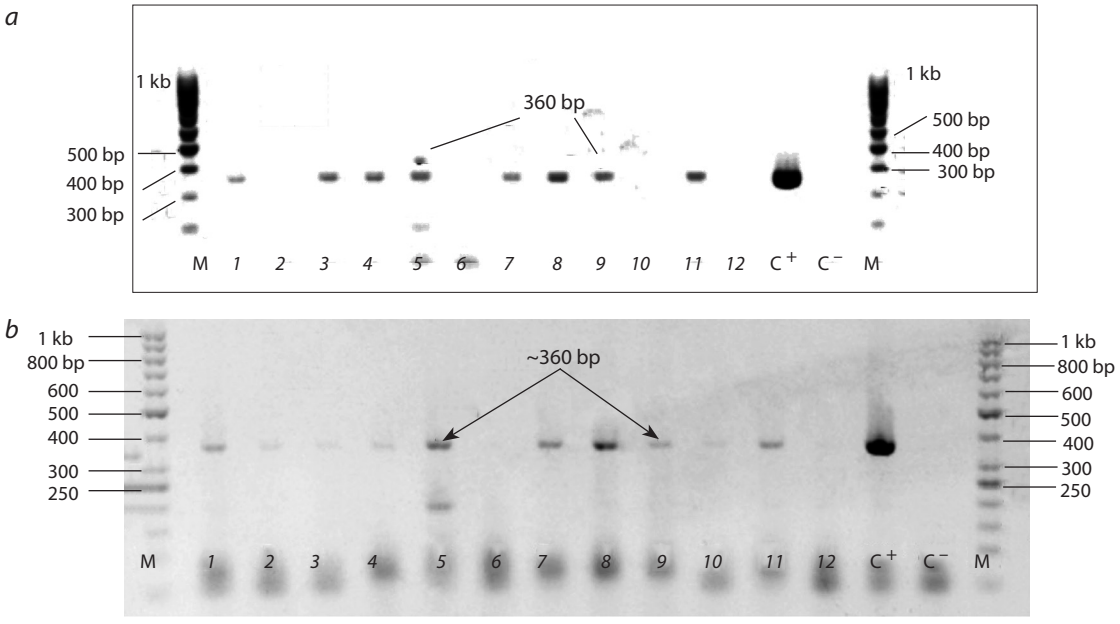


Fig. 6. RT-PCR detection of PSTVd with primer sets P3/P4 (*a*) and 68PV/87PV (*b*) in cultures of *P. infestans* isolates after colonization on PSTVd infected potato cv. Colomba.

1–4, Second passage on rye agar; *5–11*, third passage on rye agar; *12*, PSTVd-uninfected isolate PiVZR18 (negative control); C^+, cv. Colomba infected by PSTVd (positive control); C^-, distilled water; M, molecular weight marker 100 bp (*a*) and 50 bp (*b*) (Gene Ruller, Fermentas).

Virulence testing on a set of Black's potato differentials (11 lines with different resistance genes) (Black et al., 1953) of the PSTVd-infected PiVZR18v+ and the initial uninfected PiVZR18 isolates showed the same types of reactions on detached leaves of 11 lines after 7 dpi (Fig. 8, 9).

Presence of amplification product 360 bp in *P. infestans* isolates after colonization on viroid-infected plants of potato cv. Colomba (each passage is 30 days)

Number of passages on rye agar	Number of *P. infestans* isolates	
	Positive PSTVd detection	Negative PSTVd detection
One passage 30 days after re-isolation on rye agar	5	2
Two passages	3 (see Fig. 6)	1
Three passages	5 (see Fig. 6)	2
Eight passages	0	7

For both isolates, eight virulence genes (1, 2, 3, 4, 6, 7, 10, and 11) were identified. The PSTVd-infected *P. infestans* isolate seems to be less aggressive in comparison to the uninfected isolate; however, this observation requires further study.

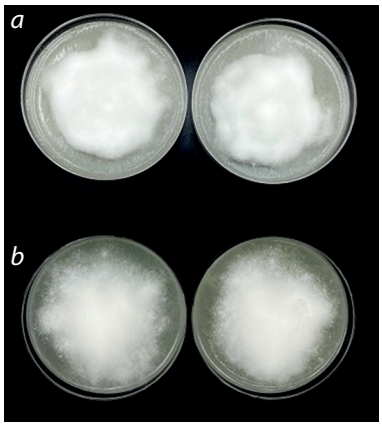


Fig. 7. Cultures of *P. infestans* PiVZR18: *a*, infected with viroid after three passages on rye agar: *b*, healthy.

Fig. 8. Virulence of the PSTVd-infected PiVZR18v+ to the lines of a set of Black's differentials (11 resistance genes). Eight virulence genes (1, 2, 3, 4, 6, 7, 10, 11) were determined.

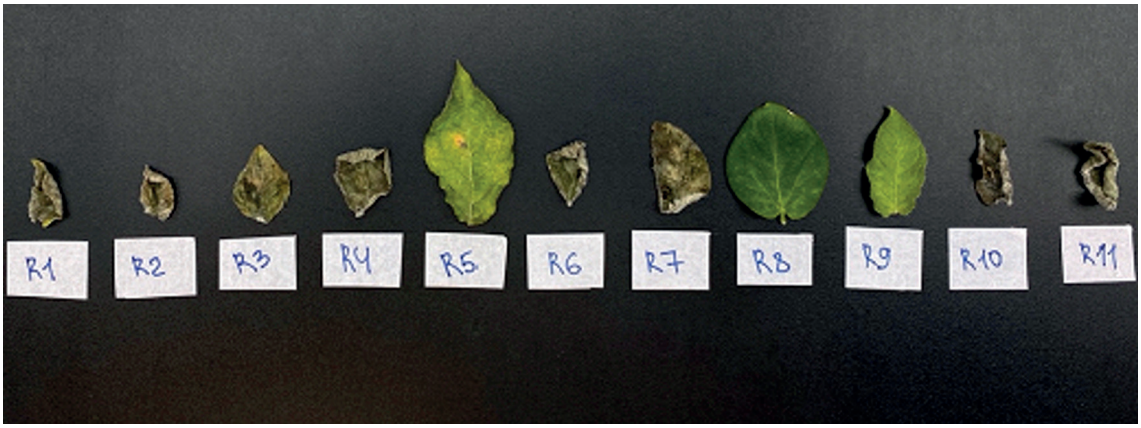


Fig. 9. Virulence of the initial PSTVd-uninfected *P. infestans* PiVZR18 isolate to the lines of a set of Black's differentials (11 resistance genes). Eight virulence genes (1, 2, 3, 4, 6, 7, 10, 11) were determined.

Discussion

Potato (*Solanum tuberosum*) is one of the most important staple crops worldwide. According to the FAO, over 390 million tons are produced on over 19 million ha of farmland worldwide (http://www.fao.org/faostat/en/#data/QCL). The quarantine status of potato tuber spindle viroids and possible significant losses of potato yield determine the importance of studying various aspects of pathogen epidemiology.

PSTVd replication is accompanied by the accumulation of viroid-derived small RNAs (vd-siRNAs) suggested to play a central role in disease symptom development.

Potatoes were domesticated in the Andes in Southern Peru around 10,000 years ago. Nevertheless, the introduction of the potato to Europe, along with all its associated diseases, dates back to the 16–17th centuries (Khavkin, 2015). Late blight of potato and tomato is a devastating disease caused by fungus-like oomycete *P. infestans.* Despite many efforts, the severity of this disease has increased dramatically in recent years.

It is known that the mode and speed of spreading plant pathogens are the major factors in the development of epiphytotics. Early studies showed that PSTVd is spread primarily through the use of infected plant material produced vegetatively or as botanical seeds (Fernow et al., 1970), through mechanical spreading across the growing crop, particularly between plants of different species of the Solanaceae family (Manzer, Merriam, 1961; Verhoeven, Roenhorst, 2010). The success of mechanical transmission depends on infected host plant species or cultivars, as well as the frequency and severity of the disease and the temperature (Bulletin OEPP/EPPO Bulletin, 2011). Importantly, transmission to potato and other test plants by aphids (*Myzus persicae*) was successful only if PSTVd RNA was encapsidated by potato leafroll virus (PLRV) particles (Salazar et al., 1995; Querci et al., 1997).

Some viruses are spread by vectors, which can include pathogenic fungi (Andika et al., 2017; Sutela et al., 2019), oomycetes (Mascia et al., 2019), and nematodes (Brown et al., 1989; Singh S. et al., 2020). Bidirectional transfer between *Fusarium graminearum* and tobacco plants of hop stunt viroid (HSVd) during infection was shown by Wei et al. (2019). However, *Nicotiana benthamiana* is not a natural host for either HSVd or *F. graminearum*. Given this fact (Serra et al., 2020), more evidence is needed to validate the possibility of viroid transmission by phytopathogenic fungi. We showed the presence of PSTVd in *P. infestans* isolates after colonization on plants of three potato cultivars infected with viroid. After three passages on rye agar (30 days each), RT-PCR analysis showed the presence of viroid in pure cultures of *P. infestans.* The partial RT-PCR amplicon sequence of viroid in *P. infestans* isolate PiVZRv+ after the third passage on rye agar is identical (98.3–99.3 %) to PSTVd strain VP87 that was used for the initial inoculation.

Sixty days after inoculation of healthy tomato plants with *P. infestans* isolate carrying PSTVd, RT-PCR revealed a 360 bp amplification product, indicating successful infection of the plants. This is the first report of horizontal transfer of potato viroid PSTVd between *P. infestans* and host plants.

Moreover, there is evidence that small RNAs (sRNAs, approximately 20–30 nt) can horizontally transfer from microbes to plants and spread silencing information toward the targeted genes (Han, Luan, 2015). Small RNAs were also found in fungi (Wang, Dean, 2020) and fungal-like *Oomycota* (Jahan et al., 2015). In addition, sRNAs of 19–40 nt were found from *P. infestans* (Vetukuri et al., 2012). There are numerous reports of sRNA cross-transfer between plants and pathogens (Zeng et al., 2019; Wang, Dean, 2020). sRNAs can be transported within an organism through the inner side of the plasma membrane (symplast), or cell wall (apoplast) (Wang, Dean, 2020). It is suggested that sRNAs are translocated by extracellular vesicles (EVs) from *Arabidopsis* to *P. capsica* (Hou et al., 2019) and *B. cinerea* (Cai et al., 2018). Furthermore, it is possible that interaction between oomycete and potato involves not only sRNA exchange but also the movement of larger viroid RNA molecules from mycelium into a plant and *vice versa*.

PSTVd replicates in the nucleus, traffics long distances in the phloem, and moves cell-to-cell via plasmodesmata in plants (Takeda, Ding, 2009). After the third subculture, PSTVd was detected from *P. infestans*, suggesting that PSTVd can replicate in the nucleus and locate to non-septate hyphae of *P. infestans* (see the Table). On the other hand, after the eighth subculture, PSTVd accumulation in *P. infestans* was not detected by RT-PCR. The same results were obtained by Wei et al. (2019), in which PSTVd was eliminated from *Cryphonectria parasitica*, *Valsa mali*, and *Fusarium graminearum* after eight subcultures. This disappearance could be caused by a defense mechanism against viroid, namely, the RNA silencing system. Viroids are the target of the RNA silencing system and become elicitors of the host defense system via RNA silencing (Cottilli et al., 2019; Wei et al., 2020). Thus, PSTVd could have been degraded by the silencing system, resulting in the elimination of PSTVd from *P. infestans*.

Phytophthora infestans produce sporangia on the surface of potato leaves, and then zoospores, released from sporangia, form walled cysts on the plant surface (Mazumdar et al., 2021). The cysts germinate and extend a germ tube into the leaves and stems of the host plants. PSTVd transferred from *P. infestans* to plants, suggesting that PSTVd was present not only in mycelium but also in sporangia and zoospores. Mature sporangia were dispersed by wind or water (Leesutthiphonchai et al., 2018). Thus, there is a possibility that PSTVd can spread long distances via infected sporangia. In contrast, there is still no evidence of viroid infection in isolates of *P. infestans* from field populations and the possibility of viroid stability in the mycelium of *P. infestans* in tubers is unclear.

Concerning mycoviruses, there are two hypotheses of their origin: the first states that they are of an unknown but ancient origin and have coevolved along with their hosts, the second one suggests they have relatively recently moved from a fungal plant host into fungus (Pearson et al., 2009). Both hypotheses are also applicable to PSTVd. Prolonged co-existence of viroid–*P. infestans*–host plants can lead to viroid transition from a host plant to an oomycete.

Conclusion

Potato spindle tuber viroid is known as autonomously replicating pathogen only of plants and mainly of solanaceous crops, that lacks any protein-coding sequences. Herein, we demonstrate the possibility of viroid transmission from host plants (potato and tomato) to *Phytophthora* infestans, from *P. infestans* to host plants, and the possibility of PSTVd

stability in pure cultures of *P. infestans* after three consecutive passages on rye agar. These results are initial evidence of bidirectionally transferred potato viroid PSTVd between *P. infestans* and host plants.

References

Andika I.B., Weia Sh., Cao Ch., Salaipetha L., Kondo H., Suna L. Phytopathogenic fungus hosts a plant virus: a naturally occurring cross-kingdom viral infection. *Proc. Natl. Acad. Sci. USA.* 2017;114(46):12267-12272. DOI 10.1073/pnas.1714916114.

Behjatnia A., Dry I., Krake L., Condé B.D., Connelly M.I., Randles J., Rezaian M.A. New potato spindle tuber viroid and tomato leaf curl geminivirus strains from a wild *Solanum* sp. *Phytopathology*. 1996;86:880-886. DOI 10.1094/Phyto-86-880.

Bhat Al.I., Rao G.P. Characterization of Plant Viruses. Methods and Protocols. Humana, New York: Springer Science+Business Media, 2020. DOI 10.1007/978-1-0716-0334-5_1.

Black W., Mastenbroek C., Mills W.R., Peterson L.C. A proposal for an international nomenclature of races of *Phytophthora infestans* and of genes controlling immunity in *Solanum demissum* derivatives. *Euphytica*. 1953;2:173-178. DOI 10.1007/BF00053724.

Brown D.J.F., Ploeg A.T., Robinson D.J. A review of reported associations between *Trichodorus* and *Paratrichodorus* species (Nematoda : Trichodoridae) and tobraviruses with a description of laboratory methods for examining virus transmission by trichodorids. *Rev. Nématol.* 1989;12(3):235-241.

Bulletin OEPP/EPPO Bulletin. 2011. *Potato spindle tuber viroid* on potato. 2011;41(3):394-399.

Cai Q., Qiao L., Wang M., He B., Lin F.M., Palmquist J., Huang S.D., Jin H. Plants send small RNAs in extracellular vesicles to fungal pathogen to silence virulence genes. *Science*. 2018;360:1126-1129. DOI 10.1126/science.aar4142.

Cottilli P., Belda-Palazón B., Adkar-Purushothama C.R., Perreault J.-P., Schleiff E., Rodrigo I., Ferrando A., Lisón P. Citrus exocortis viroid causes ribosomal stress in tomato plants. *Nucleic Acids Res*. 2019;47(16):8649-8661. DOI 10.1093/nar/gkz679.

Diener T.O., Raymer W.B. Potato spindle tuber virus: a plant virus with properties of a free nucleic acid. II. Characterization and partial purification. *Virology*. 1969;37(3):351-366. DOI 10.1016/0042-6822(69)90219-0.

Fernow K.H., Peterson L.C., Plaisted R.L. Spindle tuber virus in seeds and pollen of infected plants. *Am. Potato J*. 1970;47:75-80. DOI 10.1007/BF02864807.

Gross H.J., Domdey H., Lossow C., Jank P.R.M., Alberty H., Sänger H.L. Nucleotide sequence and secondary structure of potato spindle tuber viroid. *Nature*. 1978;273:203-208. DOI 10.1038/273203a0.

Han L., Luan Y.-S. Horizontal transfer of small RNAs to and from plants. *Front. Plant Sci.* 2015;6:1113. DOI 10.3389/fpls.2015.01113.

Hou Y., Zhai Y., Feng L., Karimi H.Z., Rutter B.D., Zeng L., Choi D.S., Zhang B., Gu W., Chen X., Ye W., Innes R.W., Zhai J., Ma W. A *Phytophthora* effector suppresses trans-kingdom RNAi to promote disease susceptibility. *Cell Host Microbe*. 2019;25(1):153-165.e5. DOI 10.1016/j.chom.2018.11.007.

Jahan S.N., Åsman A.K.M., Corcoran P., Fogelqvist J., Vetukuri R.R., Dixelius C. Plant-mediated gene silencing restricts growth of the potato late blight pathogen *Phytophthora infestans*. *J. Exp. Bot*. 2015;66:2785-2794. DOI 10.1093/jxb/erv094.

Kastalyeva T.B., Mozhaeva K.A., Pisetskaya N.F., Romanova S.A., Trofimets L.N. The potato spindle tuber viroid and bringing potato into a healthy state. *Vestnik RASKHN = Herald of the Russian Academy of Agricultural Sciences*. 1992;3:22-24. (in Russian)

Khavkin E.E. Potato late blight as a model of pathogen-host plant coevolution. *Russ. J. Plant Physiol*. 2015;62(3):408-419. DOI 10.1134/S1021443715030103.

Kryczyński S., Paduch-Cichal E.L., Skrzeczkowski J. Transmission of three viroids through seed and pollen of tomato plants. *J. Phytopathol*. 1988;121(1):51-57. DOI 10.1111/j.1439-0434.1988.tb00952.x.

Leesutthiphonchai W., Vu A.L., Ah-Fong A.M.V., Judelson H.S. How does *Phytophthora infestans* evade control efforts? Modern insight into the late blight disease. *Phytopathology*. 2018;108(8):916-924. DOI 10.1094/PHYTO-04-18-0130-IA.

Mackie A.E., Rodoni B.C., Barbetti M.J., McKirdy S.J., Jones R.A.C. Potato spindle tuber viroid: alternative host reservoirs and strain found in a remote subtropical irrigation area. *Eur. J. Plant Pathol*. 2016;145(2):433-446. DOI 10.1007/s10658-016-0857-2.

Manzer F.E., Merriam D. Field transmission of potato spindle tuber virus and virus X by cultivating and hilling equipment. *Am. Potato J*. 1961;38:346-352.

Mascia T., Vučurović A., Minutillo S.A., Nigro F., Labarile R., Savoia M.A., Palukaitis P., Gallitelli D. Infection of *Colletotrichum acutatum* and *Phytophthora infestans* by taxonomically different plant viruses. *Eur. J. Plant Pathol*. 2019;153(4):1001-1017. DOI 10.1007/s10658-018-01615-9.

Matsushita Y., Yanagisawa H. Distribution of *Tomato planta macho viroid* in germinating pollen and transmitting tract. *Virus Genes*. 2018;54:124-129. DOI 10.1007/s11262-017-1510-7.

Matsushita Y., Yanagisawa H., Khiutti A., Mironenko N., Ohto Y., Afanasenko O. Genetic diversity and pathogenicity of potato spindle tuber viroid and chrysanthemum stunt viroid isolates in Russia. *Eur. J. Plant Pathol.* 2021;161:529-542. DOI 10.1007/s10658-021-02339-z.

Mazumdar P., Singh P., Kethiravan D., Ramathani I., Ramakrishnan N. Late blight in tomato: insights into the pathogenesis of the aggressive pathogen *Phytophthora infestans* and future research priorities. *Planta*. 2021;253(6):119. DOI 10.1007/s00425-021-03636-x.

Medina M.V., Platt H.W. Comparison of different culture media on the mycelial growth, sporangia and oospore production of *Phytophthora infestans*. *Am. J. Potato Res.* 1999;76:121-125. DOI 10.1007/BF02853576.

Mertelik J., Kloudova K., Cervena G., Necekalova J., Mikulkova H., Levkanicova Z., Dedic P., Ptacek J. First report of *Potato spindle tuber viroid* (PSTVd) in *Brugmansia* spp., *Solanum jasminoides*, *Solanum muricatum* and *Petunia* spp. in the Czech Republic. *Plant Pathol.* 2010;59(2):392. DOI 10.1111/j.1365-3059.2009.02115.x.

Owens R.A., Khurana S.M.P., Smith D.R., Singh M.N., Garg I.D. A new mild strain of potato spindle tuber viroid isolated from wild *Solanum* spp. in India. *Plant Dis*. 1992;76(5):527-529. DOI 10.1094/PD-76-0527.

Pearson M.N., Beever R.E., Boine B., Arthur K. Mycoviruses of filamentous fungi and their relevance to plant pathology. *Mol. Plant Pathol.* 2009;10(1):115-128. DOI 10.1111/j.1364-3703.2008.00503.x.

Pfannenstiel M.A., Slack S.A. Response of potato cultivars to infection by the potato spindle tuber viroid. *Phytopathology*. 1980;70(9):922-926. DOI 10.1094/Phyto-70-922.

Puchta H., Herold T., Verhoeven K., Roenhorst A., Ramm K., Schmidt-Puchta W., Sänger H.L. A new strain of potato spindle tuber viroid (PSTVd-N) exhibits major sequence differences as compared to all other PSTVd strains sequenced so far. *Plant Mol. Biol.* 1990;15(3):509-511. DOI 10.1007/BF00019169.

Querci M., Owens R.A., Bartolini I., Lazarte V., Salazar L.F. Evidence for heterologous encapsidation of potato spindle tuber viroid in particles of potato leafroll virus. *J. Gen. Virol*. 1997;78(Pt.6):1207-1211. DOI 10.1099/0022-1317-78-6-1207.

Ristaino J.B., Madritch M., Trout C.L., Parra G. PCR amplification of ribosomal DNA for species identification in the plant pathogen genus *Phytophthora*. *Appl. Environ. Microbiol.* 1998;64(3):948-954. DOI 10.1128/AEM.64.3.948–954.

Salazar L.F., Querci M., Bartolini I., Lazarte V. Aphid transmission of potato spindle tuber viroid assisted by potato leafroll virus. *Fitopatologia*. 1995;30(1):56-58.

Serra P., Carbonell A., Navarro B., Gago-Zachert S., Li S., Di Serio F., Flores R. Symptomatic plant viroid infections in phytopathogenic fungi: a request for a critical reassessment. *Proc. Natl. Acad. Sci. USA.* 2020;117(19):10126-10128. DOI 10.1073/pnas.1922249117.

Singh R.P. Seed transmission of potato spindle tuber virus in tomato and potato. *Am. Potato J*. 1970;47:225-227. DOI 10.1007/BF02872303.

Singh R.P. Experimental host range of the potato spindle tuber 'virus'. *Am. Potato J*. 1973;50:111-123. DOI 10.1007/BF02857207.

Singh R.P., Boucher A., Somerville T.H. Detection of *Potato spindle tuber viroid* in the pollen and various parts of potato plant pollinated with viroid-infected pollen. *Plant Dis*. 1992;76:951-953.

Singh S., Awasthi L.P., Jangre A., Nirmalkar V.K. Chapter 22 – Transmission of plant viruses through soil-inhabiting nematode vectors. In: Awasthi L.P. (Ed.) Applied Plant Virology. Acad. Press, 2020;291-300. DOI 10.1016/B978-0-12-818654-1.00022-0.

Sutela S., Poimala A., Vainio E.J. Viruses of fungi and oomycetes in the soil environment. *FEMS Microbiol. Ecol.* 2019;95(9):fiz119. DOI 10.1093/femsec/fiz119.

Syller J., Marczewski W., Pawłowicz J. Transmission by aphids of potato spindle tuber viroid encapsidated by potato leafroll luteovirus particles. *Eur. J. Plant Pathol.* 1997;103:285-289. DOI 10.1023/A:1008648822190.

Takeda R., Ding B. Viroid intercellular trafficking: RNA motifs, cellular factors and broad impacts. *Viruses*. 2009;1(2):210-221. DOI 10.3390/v1020210.

Verhoeven J.T.J., Roenhorst J.W. High stability of original predominant pospiviroid genotypes upon mechanical inoculation from ornamentals to potato and tomato. *Arch. Virol*. 2010;155:269-274. DOI 10.1007/s00705-009-0572-9.

Vetukuri R.R., Åsman A.K.M., Tellgren-Roth C., Jahan S.N., Reimegård J., Fogelqvist J., Savenkov E., Söderbom F., Avrova A.O., Whisson S.C., Dixelius C. Evidence for small RNAs homologous to effector-encoding genes and transposable elements in the oomycete *Phytophthora infestans*. *PLoS One*. 2012;7(12):e51399. DOI 10.1371/journal.pone.0051399.

Wang M., Dean R.A. Movement of small RNAs in and between plants and fungi. *Mol. Plant Pathol*. 2020;21:589-601. DOI 10.1111/mpp.12911.

Wei Sh., Bian R., Andika I.B., Niu E., Liu Q., Kondo H., Yang L., Zhou H., Pang T., Lian Z., Liu X., Wu Y., Sun L. Symptomatic plant viroid infections in phytopathogenic fungi. *Proc. Natl. Acad. Sci. USA.* 2019;116(26):13042-13050. DOI 10.1073/pnas.1900762116.

Wei S., Bian R., Andika I.B., Niu E., Liu Q., Kondo H., Yang L., Zhou H., Pang T., Lian Z., Liu X., Wu Y., Sun L. Nucleotide substitutions in plant viroid genomes that multiply in phytopathogenic fungi. *Proc. Natl. Acad. Sci. USA.* 2020;117(19):10129-10130. DOI 10.1073/pnas.2001670117.

White T.J., Bruns T., Lee S., Taylor J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis M.A., Gelfand D.H., Sninsky J.J., White T.J. (Eds.) PCR Protocols, a Guide to Methods and Applications. Acad. Press, 1990;315-322. DOI 10.1016/B978-0-12-372180-8.50042-1.

Yanagisawa H., Sano T., Hase S., Matsushita Y. Influence of the terminal left domain on horizontal and vertical transmissions of tomato planta macho viroid and potato spindle tuber viroid through pollen. *Virology*. 2019;526(2):22-31. DOI 10.1016/j.virol.2018.09.021.

Zeng J., Gupta V.K., Jiang Y., Yang B., Gong L., Zhu H. Cross-kingdom small RNAs among animals, plants and microbes. *Cells*. 2019;8(4):371. DOI 10.3390/cells8040371.

ORCID ID

O.S. Afanasenko orcid.org/0000-0001-7368-0797
A.V. Khiutti orcid.org/0000-0003-1479-7746
N.V. Mironenko orcid.org/0000-0003-3383-2973
N.M. Lashina orcid.org/0000-0003-1419-7134

Acknowledgements. This study was supported by the Russian Science Foundation, project No. 20-46-07001.

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

Received December 27, 2021. Revised February 3, 2022. Accepted February 3, 2022.

Вавиловский журнал генетики и селекции. 2022;26(3):281-289
DOI 10.18699/VJGB-22-35

Social behavior and spatial orientation in rat strains with genetic predisposition to catatonia (GC) and stereotypes (PM)

V.S. Plekanchuk[1, 2] ✉, O.I. Prokudina[1], M.A. Ryazanova[1]

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Novosibirsk State University, Novosibirsk, Russia
✉ lada9604@mail.ru

Abstract. Various psychopathologies, including schizophrenia, bipolar disorder and major depression, are associated with abnormalities in social behavior and learning. One of the syndromes that may also take place in these disorders is catatonia. Catatonia is a psychomotor syndrome in which motor excitement, stereotypy, stuporous state, including the phenomenon of "waxy flexibility" (catalepsy), can be observed. Rats with genetic catatonia (GC) and pendulum-like movements (PM) of the anterior half of the body have physiological and behavioral changes similar to those observed in schizophrenia and depression in humans and can be considered as incomplete experimental models of these pathologies. The social behavior of the GC and PM rats has not been previously studied, and the cognitive abilities of animals of these strains are also insufficiently studied. To determine whether the GC and PM rats have changes in social behavior and spatial learning, behavioral phenotyping was performed in the resident-intruder test, three-chamber test, Barnes maze test. Some deviations in social behavior, such as increased offensive aggression in PM rats in the resident-intruder test, increased or decreased social interactions depending on the environment in different tests in GC, were shown. In addition, principal component analysis revealed a negative association between catatonic freezing and the socialization index in the three-chamber test. Decreased locomotor activity of GC rats can adversely affect the performance of tasks on spatial memory. It has been shown that PM rats do not use a spatial strategy in the Barnes maze, which may indicate impairment of learning and spatial memory.
Key words: catatonia; GC rat strain; PM rat strain; epilepsy; stereotypes; learning; social interaction.

For citation: Plekanchuk V.S., Prokudina O.I., Ryazanova M.A. Social behavior and spatial orientation in rat strains with genetic predisposition to catatonia (GC) and stereotypes (PM). *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2022;26(3):281-289. DOI 10.18699/VJGB-22-35

Социальное поведение и пространственная ориентация у линий крыс с генетической предрасположенностью к кататонии (ГК) и стереотипиям (МД)

В.С. Плеканчук[1, 2] ✉, О.И. Прокудина[1], М.А. Рязанова[1]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
✉ lada9604@mail.ru

Аннотация. Различные психопатологии, включая шизофрению, биполярное расстройство и большое депрессивное расстройство, ассоциированы с отклонениями в социальном поведении и обучении. Одним из синдромов, который также может иметь место при этих расстройствах, является кататония. Кататония – психомоторное расстройство, при котором могут наблюдаться двигательное возбуждение, стереотипии, ступор, в том числе с явлением «восковой гибкости» (каталепсии). Крысы с генетической кататонией (ГК) и маятникообразными стереотипными движениями (МД) имеют физиологические и поведенческие изменения, сходные с наблюдаемыми при шизофрении и депрессии у человека, и могут рассматриваться как неполные экспериментальные модели этих патологий. Социальное поведение крыс линий ГК и МД ранее не было исследовано, также недостаточно изучены когнитивные способности животных данных линий. Чтобы определить, имеются ли изменения в социальном поведении и пространственном обучении у крыс линий ГК и МД, было проведено их поведенческое фенотипирование в тесте «резидент-интрудер», трехкамерном тесте и лабиринте Барнс. В нашей работе показаны некоторые отклонения в социальном поведении, такие как усиление оборонительной агрессии у крыс МД, увеличение или уменьшение уровня социальных взаимодействий в зависимости от условий тестирования у крыс ГК. Кроме того, анализ главных компонент

выявил отрицательную связь между кататоническим застыванием и индексом социальности в трехкамерном тесте. Снижение двигательной активности крыс ГК может негативно влиять на выполнение заданий по оценке пространственной памяти. Показано, что крысы МД не используют пространственную стратегию в лабиринте Барнс, что может указывать на нарушение обучения и пространственной памяти.
Ключевые слова: кататония; линия крыс ГК; линия крыс МД; эпилепсия; стереотипии; обучение; социальное взаимодействие.

Introduction

In psychiatric classification, there is an acute issue of division and diagnosis of individual nosological units. A lot of evidence pointing to the generally continuous nature of psychopathological variation versus discrete has been accumulated (Krueger et al., 2018). In DSM-5 (Diagnostic and Statistical Manual of mental disorders), there are many "spectra" and groups of disorders (Autism Spectrum Disorder, Schizophrenia Spectrum and Other Psychotic Disorders, Bipolar and related disorders), the symptoms of which overlap very strongly. The comorbidity observed between major depression and schizophrenia (Samsom, Wong, 2015), bipolar disorder, attention deficit hyperactivity disorder, and autism (Kiser et al., 2015) implies that the same pathophysiological processes occur in these diseases. In this regard, new concepts are being created that try to explain the pathogenesis of distinct psychiatric symptoms and emphasize the exploration of endophenotypes but not of complex diseases (Anderzhanova et al., 2017). This approach solves the "problem of comorbidity" by explicitly modeling patterns of co-occurrence among signs and symptoms (Krueger et al., 2018).

One of the syndromes that can be used as a "specifier" in DSM-5 for the characterization of several clinical phenotypes including schizophrenia spectrum disorders, affective, and neurodevelopmental disorders is catatonia (Wilson et al., 2015). Catatonia is a psychomotor syndrome characterized by various signs: stupor, catalepsy (posturing, waxy flexibility), stereotypy, mutism. This motor and behavioral alteration may occur in many psychiatric conditions but predominantly in schizophrenia, affective psychosis, autism (Fink, Taylor, 2001). While many aspects of human psychopathologies cannot be simulated in animals, some symptoms of catatonia can. Different animal models can help characterize the nature of specific psychopathology symptoms, and there are special behavioral parameters of potential relevance to signs and symptoms of schizophrenia. Excessive catatonic reactions in animals can also correspond to catatonia in humans and include presence of bizarre motor activity, decrease in motor activity, or catatonic excitement (intense bursts of agitated stereotypy).

For genetically based modeling of schizophrenia-relevant and catatonia-relevant symptoms, the GC (genetic catatonia) and the PM (pendulum-like movements) rat strains were offered (Timofeeva, 1985). The strains were obtained by selection for intensification of such catatonic reactions as freezing or catalepsy (GC strain) and stereotyped pendulum movements (PM strain). The GC rats demonstrate occasional freezing or, instead, hyperkinetic behavioral reactions that resemble the manifestations of the catatonic syndrome (Ryazanova et al., 2012). These reactions can be spontaneous, as well as in response to a weak stressful stimulus, such as in a special test for catatonic freezing (Fig. 1, *a*). In addition, rats of this strain are characterized by increased stress reactivity (Alekhina et al., 2015), increased shock-induced aggression (Nikulina et al., 1987), impaired filtration of sensorimotor information (manifested by a deficiency of PPI) (Ryazanova et al., 2017). PM rats are characterized by rhythmic side-to-side rocking of the head and forebody in the absence of locomotion (see Fig. 1, *b*). More than that, rats selected for an increased amplitude of pendulum-like movements after the 40th generation started generating seizures to audiogenic stimulation (Alekhina et al., 2007).

Despite some parameters supporting face validity of this model, phenotype of these strains is not yet well explored. For example, aspects of behavior and cognitive activity such as social interactions and learning are also of interest. A variety of neuropsychiatric disorders are characterized by disruptions in social behavior and social recognition, including depression, autism spectrum disorders, bipolar disorders, obsessive-compulsive disorders, and schizophrenia. In animals, altered social interaction responses in a variety of situations are considered as analogs related to negative – social withdrawal – symptoms of schizophrenia (Powell, Miyakawa, 2006), hyperactivity and aggressive behavior directly related to positive symptoms of schizophrenia (Volavka, Citrome, 2008).

To determine whether selection for predisposition to the catatonic freezing and the amplitude of pendulum-like movements influenced social interactions and learning in the GC and PM rat strains, behavioral phenotyping of rats in the resident intruder test, three-chamber test, Barnes maze test was carried out.

Methods

The study was carried out on male rats of the GC (genetic catatonia), PM (pendulum-like movements), Wistar and WAG (Wistar Albino Glaxo) strains. Since the PM rat strain is outbred, rats of the outbred Wistar strain were used as a control, while for the inbred GC, the inbred WAG were used.

Experiment 1 included the catatonic freezing test (22 males at the age of 2 months from each strain); the three-chamber paradigm test (the same 15 males at the age of 5 months from each strain) and the resident-intruder test (5 days after the three-chamber test). In the Experiment 2, another 60 rats (15 males at the age of 4 months from each strain) were tested in the Barnes maze.

Rats were kept under standard vivarium conditions with a free access to food and water. All experimental procedures complied with the rules and regulations formulated in the EU Council Directive 1986 (86/609/EEC) and the Declaration of Helsinki on the protection of vertebrate animals used in experimental research and approved by the ICG SB RAS Bioethics Committee (protocol No. 43, 28.09.2018).

Experiment 1. Social behavior and catatonia

The catatonic freezing test is a selection criterion for the GC rats and was carried out according to the protocol (Timofeeva, 1985). To determine the presence or absence of freezing

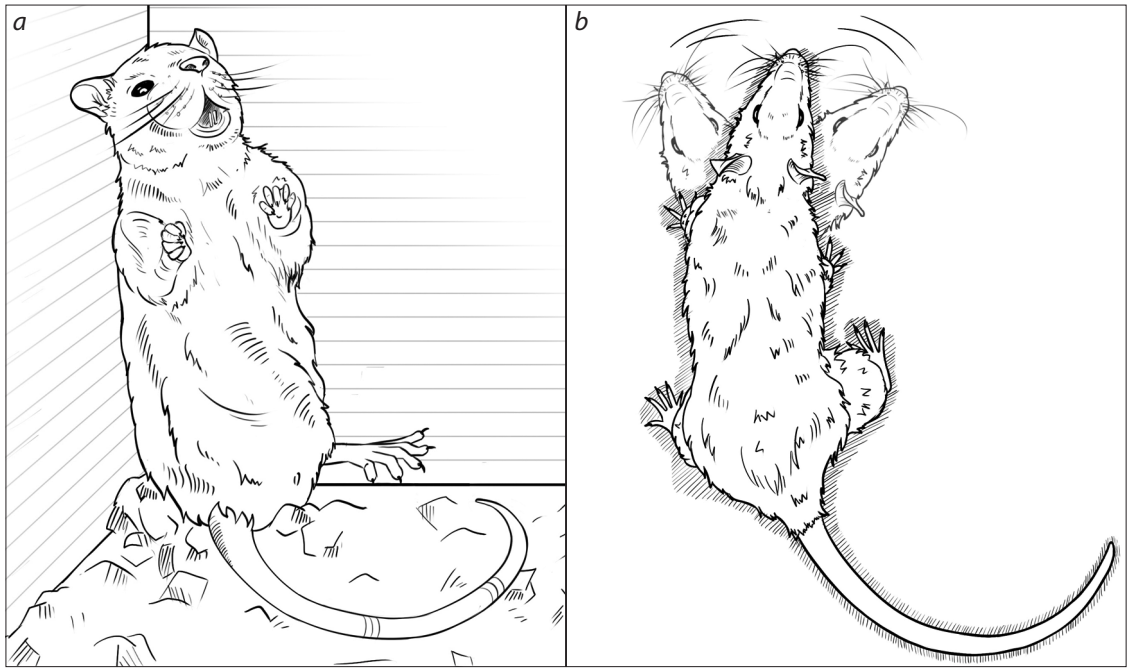


Fig. 1. Catatonic reactions in rats: *a*, catalepsy in GC; *b*, pendulum-like movements of the PM rats.

reactions and their duration, the rat was uplifted in the corner of the cage by the forelegs using a test stick. The freezing time was estimated as a time during which the animal retained the induced posture or freezing position on 4 paws after the stick was removed. Rats were tested two times on different days.

The three-chamber paradigm test. The three-chamber social interaction assay was performed to assess social deficits according to the protocol (Kaidanovich-Beilin et al., 2011). Testing was carried out in a test arena manufactured by OpenScience, Russia, model TS1701-R. The apparatus for the test is comprised of a rectangular, three-chamber box. Each chamber is 40×85 cm and the dividing walls are made from clear Plexiglas, with an open middle section, which allows free access to each chamber. For habituation, the test rat was placed into a Plexiglas's arena containing two empty cylindrical containers in two side chambers for 10 minutes.

Session I. Wistar males of the same weight without any prior contact (not littermates) with the subject were used as control animals (Stranger 1 and Stranger 2). One of the control rats (Stranger 1) was placed in one of the containers located in one of the side chambers. The placement of Stranger 1 on the left or right side of the chamber was systematically altered between trials. After removing the walls between the compartments, the following parameters were monitored and recorded: duration of direct contacts of the subject rat with Stranger 1; duration of contacts with empty enclosure. The duration of session I was 10 minutes. Then the Session II began. The second control rat was placed in the empty cylinder in the opposite side chamber. Duration of direct contacts of the subject rat with Stranger 1 and Stranger 2 were monitored and recorded within 10 minutes. The socialization index was calculated by the formula $(T_1-T_0)/(T_1+T_0)\times 100$ %, where T_1 is the time of contact with the containment cup housing Stranger 1 rat; T_0 – time of contact with the empty enclosure. The social novelty index is calculated by the formula $(T_2-T_1)/(T_2+T_1)\times 100$ %, where T_1 is the time of contact with familiar rat (Stranger 1), T_2 is the time of contact with the container housing Stranger 2 rat. The freezing time in this test was also recorded at each session (Freezing 1 and Freezing 2, respectively).

The resident-intruder test. To measure offensive aggression, the resident-intruder test was performed according to the standard protocol (Koolhaas et al., 2013). To assess the defensive behavior of resident males of the studied strain, they were placed in cages for 7 days before the test. To preserve olfactory signals, the cage was not cleaned before the test. The intruder (the Wistar male of the same size) was placed in the resident's cage through the partition, then the partition was removed. Testing was carried out for 10 minutes. Durations of the behavioral parameters were registered: (1) total offense: sum of lateral threat, upright, clinch and keep down; (2) social exploration: sum of social explore, ano-genital sniffing and move towards; (3) non-social activity: non-social explore, rearing, grooming; (4) inactivity, including rest and freezing (freezing in the RI). Also the numbers of mounts and attack latencies were analyzed.

The analysis of the main factors determining the variability of behavior characteristics in Experiment 1 was investigated by the Principal Component Analysis.

Experiment 2. Spatial learning

The Barnes maze was used to test the acquisition of spatial memory. Testing was carried out in a setup manufactured by RPC OpenScience, Russia, model TS1101-R (field diameter 122 cm, 18 holes are located around the perimeter). Testing in the Barnes maze included 3-minute training sessions once a day for 5 days (Stansley, Yamamoto, 2015). Probe trial was administered 24 hours after the acquisition session (Day 7). The following parameters were to be calculated: (1) primary latency, (2) primary errors, (3) distance moved (in cm), and (4) velocity (cm/s) (Gawel et al., 2019).

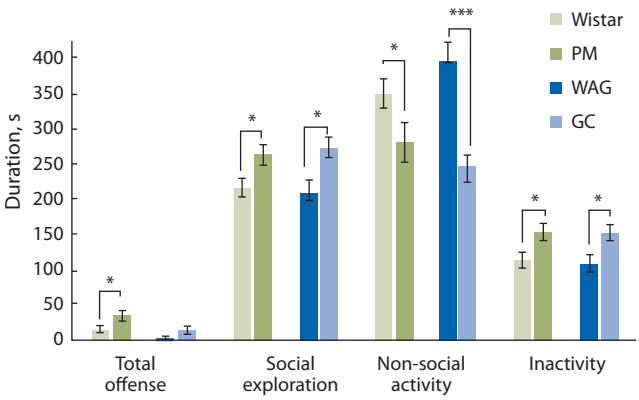


Fig. 2. Behavioral profile of resident males during a ten minutes' resident-intruder test.

Differences marked are shown for PM versus Wistar rats, and GC versus WAG. * $p < 0.05$; *** $p < 0.001$, Student's *t*-test.

Component patterns for Experiment 1

Test	Variables	C1	C2	C3
The catatonic freezing test	Catatonic freezing	–0.69	–	–
The resident-intruder test	Freezing in the RI	–	–	0.57
	Mount	–	0.74	–
	Offense	–	–	0.71
	Social exploration	–	0.66	–
The three-chamber test	Sociability index	0.59	–	–
	Social novelty index	–	–	–
	Freezing 1	–0.68	–	–
	Freezing 2	–0.65	–	–

Note. Catatonic freezing – duration of stupor in the catatonic freezing test; Freezing in RI – duration of immobility in the resident-intruder test; Mount – number of mounts in the resident-intruder test; Offense in RI – total duration of aggressive behavior in the resident-intruder test; Social exploration in RI – total duration of non-aggressive social behavior in the resident-intruder test; Sociability index – in the three-chamber paradigm test; Social novelty index – in the three-chamber paradigm test; Freezing 1 and Freezing 2 in the three-chamber test – duration of immobility in the session I and in the session II, respectively, of the three-chamber paradigm test. Only component patterns above 0.55 were recorded.

Video tracking and registration of behavioral parameters were carried out using the program EthoVision XT 15 (Noldus, Wageningen, Netherlands). In addition, on 4, 5 and 7 day trial was classified into 1 of 3 categories of search strategy (Yassine et al., 2013) reflecting the use of either a direct spatial strategy (defined as direct visit to the target, sometimes preceded by at most 1 adjacent hole visit), a serial strategy (minimum of 2 adjacent hole visits in a serial manner before reaching the target) or a mixed (i. e., random) strategy (remaining trials). The data were subsequently analyzed in terms of percentage of trials with a direct spatial strategy.

Statistics

The obtained data were processed using STATISTICA 10.0. In the paper, data are presented as mean ± SEM. Behavioral scores from Experiment 1 were analyzed by Student *t*-tests (except for parameters: lateral threating, clinch attack, attack latencies, which was analyzed by Mann–Whitney U test). When comparing a rate, Fisher's exact test was used. The analysis of the main factors that determine the variability of behavior characteristics in Experiment 1 was investigated by the principal component analysis. In Experiment 2, comparisons of components were made using the Mann–Whitney U test. Data analysis from the training sessions of Barnes maze was carried out using repeated measures ANOVA, followed by Fisher LSD post hoc analyses to analyze group differences. Statistical evaluation of the probe trial data was performed using one-way ANOVA, Fisher LSD post hoc analyses.

Results

Experiment 1. Social behavior and catatonia

The catatonic freezing test revealed a mean duration of freezing in the GC and PM rats is by far longer than in the control rats (35.6 ± 3.4 s in GC vs 18.4±3.0 s in WAG, $p<0.001$; 23.7±3.7 s in PM vs 16.6±4.5 s in Wistar, $p<0.05$). In addition, a rate of rats in populations that freeze for longer than 10 seconds was estimated. In the GC (95.5 %) and PM (77.3 %) strains, the rate is significantly higher than in the control strains (63.6 and 39.1 %, respectively; $p < 0.01$, F=0.0351 for GC; F = 0.0155 for PM).

A study of behavior in the three-chamber paradigm test showed a decrease in the sociability index in the GC rats (18.6 ± 10.2) compared to WAG (56.0 ± 10.1) ($p < 0.05$). The sociability index in the PM rats (35.7 ± 15.4 vs 44.6 ± 14.6), as well as the social novelty index in both groups (–9.4±12.1 in GC, –0.7±15.8 in PM) did not differ from the control (–15.8±12.9 in WAG, –29.6 ± 12.5 in Wistar).

In the resident-intruder test the parameters of resident's behavior in the home cage when adding an intruder were registered and combined in categories (see Methods). The analysis revealed an increased level of social exploration of PM versus Wistar, as well as GC compared to WAG ($p<0.05$) (Fig. 2). Moreover, unlike the GC, PM rats exhibited more aggressive behavior both in total duration ($p < 0.05$) and short attack latencies (90.3± 16.9) compared to Wistar (273.7±65.8, $p<0.05$). In addition, the GC and PM strains showed significantly increased sexual behavior ($p < 0.01$), which was estimated in the number of mounts (3.5 ± 0.9 in GC vs 0.0±0 in WAG; 2.9 ± 0.7 in PM vs 0.6 ± 0.4 in Wistar). Non-social activity of the GC and PM rats was significantly lower compared to control ($p < 0.05$ and $p < 0.001$), while the time of inactivity was higher ($p < 0.05$). Thus, the behavior of the GC and PM rats in the home cage when the intruder is placed shifts towards an increase in social interactions with a decrease in exploratory activity.

The principal component analysis of Experiment 1 parameters produced three factors with eigenvalues greater than 1. These three factors explain 57 % of the variance in the correlation matrix. The factor patterns are presented in the Table.

Component 1 (24.1 % of variance) was explained by stupor in the catatonic freezing test (–0.69) and in the three-chamber paradigm test (–0.68 for session I; –0.65 for session II) and sociability index value (0.59).

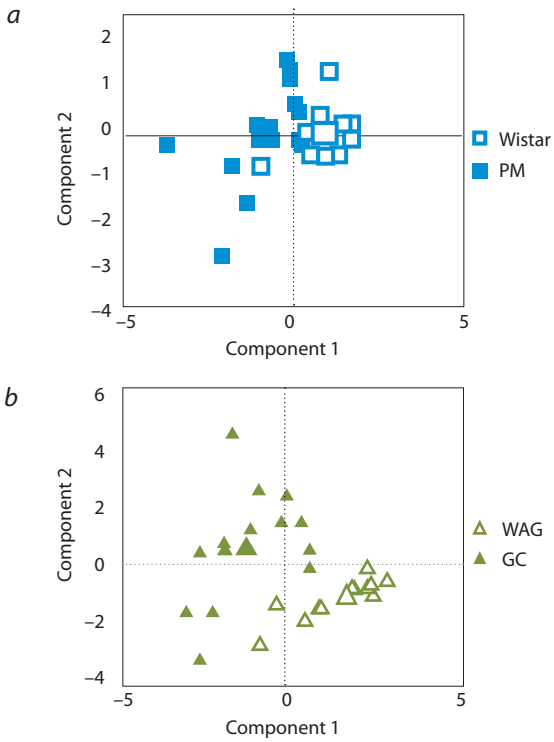


Fig. 3. Principal component scores plot: *a*, the PM compared with Wistar. The mean scores of two principal components are indicated by larger squares; *b*, the GC compared with WAG. The mean scores of two principal components are indicated by larger triangles.

Component 2 (18.7 % of variance) was mainly loaded by number of mounts (0.74) and total duration of non-aggressive social behavior (0.66) in the resident-intruder test.

Component 3 (14.2 % of variance) was loaded by the total duration of aggressive behavior (0.71) and duration of immobility (0.57) in the resident-intruder test.

Mann–Whitney U test procedures showed a strain effect for Component 1 in PM and Wistar rats ($p < 0.001$) (Fig. 3, *a*). For WAG and GC, a significant difference was shown in Component 1 ($p < 0.001$) and Component 2 ($p < 0.01$) (Fig. 3, *b*).

Experiment 2. Barnes maze task

The data analysis revealed that latency time for the GC group was significantly increased in the probe trial when compared to the WAG group ($F[1.26] = 5.9$, $p < 0.05$) (Fig. 5, *a*). No difference was found between PM and Wistar. The average velocity of movement across the maze field did not differ for Wistar and PM. Comparison of GC and WAG rats speed revealed a significant effect of the test day ($F[4, 108] = 13.7$, $p < 0.0001$) and the interaction of factors of the genotype and the test day on the speed ($F[4, 108] = 3.95$, $p < 0.001$) was found. The average velocity of movement across the maze field was significantly lower for GC in day 3 (effect of the genotype factor, $p < 0.001$), in day 4 ($p < 0.05$), in day 5 ($p < 0.05$) (Fig. 4, *b*) and in the probe trial ($F[1.28] = 12$, $p < 0.01$) (see Fig. 5, *b*). Total distance moved did not differ between groups. The use of spatial strategy increased with the training during the acquisition phase, except for the PM group: in the 7th day

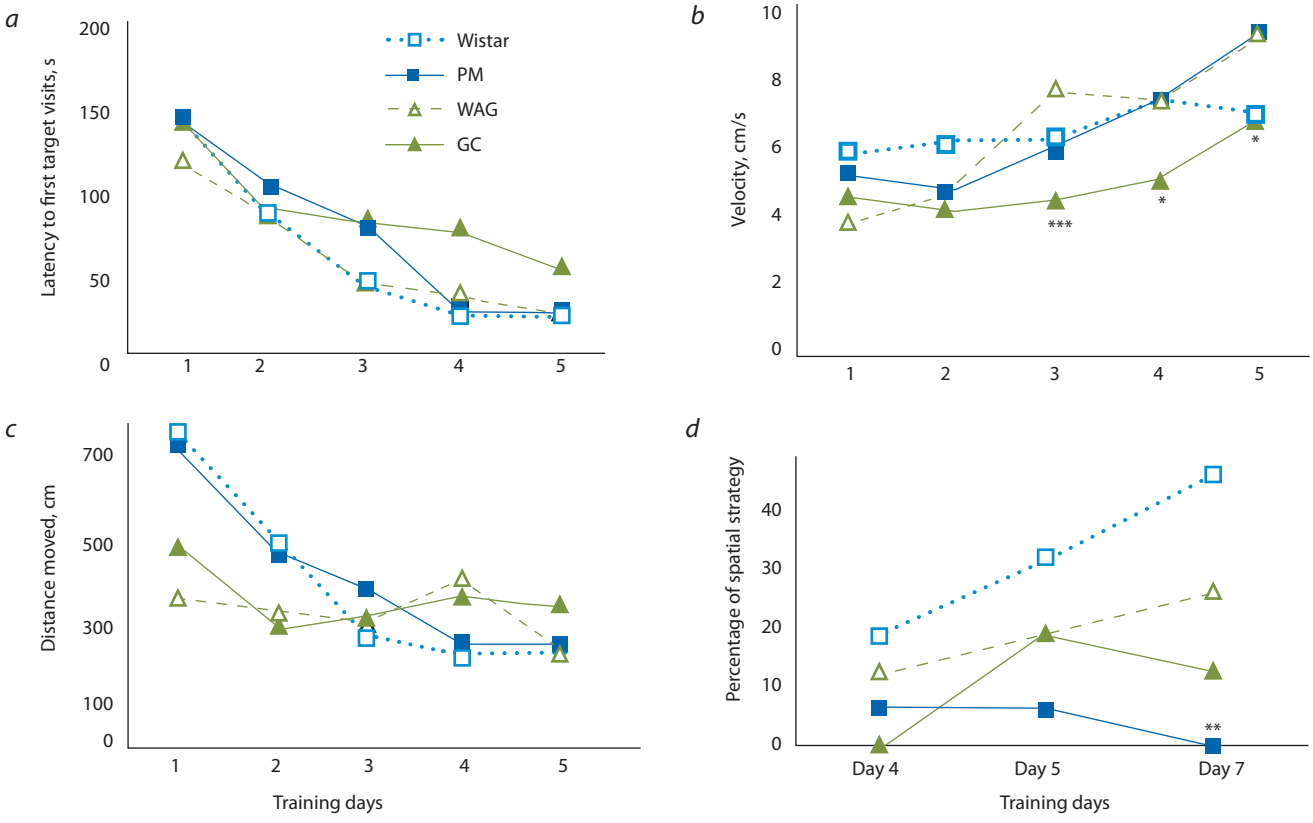


Fig. 4. Spatial learning of the PM and GC rats during the acquisition session in the Barnes maze compared to Wistar and WAG, respectively: *a*, mean latencies to enter the escape hole; *b*, the average velocity of movement across the maze field; *c*, mean distance traveled; *d*, the incidence of spatial strategy in groups.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

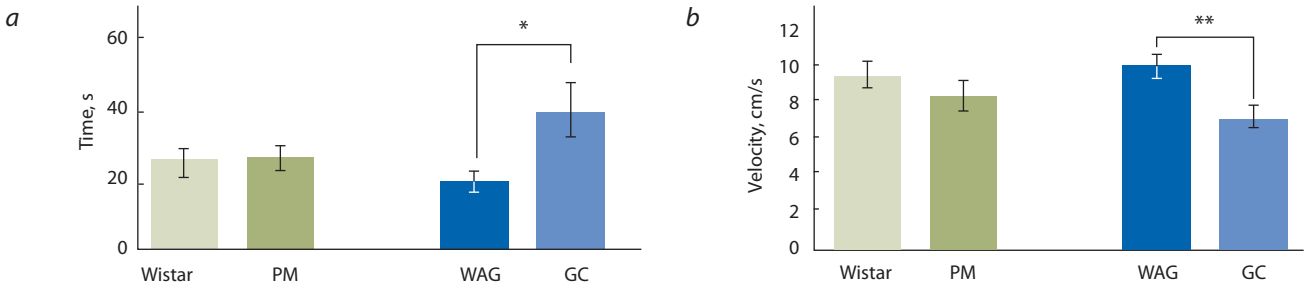


Fig. 5. Probe trial in the Barnes maze of the PM and GC rats compared to Wistar and WAG, respectively: *a*, mean latencies to enter the escape hole; *b*, the average velocity of movement across the maze field.

* $p < 0.05$, ** $p < 0.01$.

of trials, the incidence of spatial strategy in the PM rats was 0 % (0/15) compared to 46.7 % (7/15) in the Wistar rats ($p < 0.01$, F = 0.0063, Fisher's exact test) (see Fig. 4, *d*).

There was no significant effect of genetic group on the mean number of errors per trial made during the probe trial.

Discussion

Experiment 1. Social behavior and catatonia

Decreased sociability in the three-chamber test shown by GC rats in this work is consistent with literature data about social abnormalities in different animal models of psychopathologies. Most of the animal models of schizophrenia have decreased or normal social interaction (Jones et al., 2011; Nani et al., 2019). In particular, in the model of negative symptoms of schizophrenia in animals induced by NMDA-receptor antagonists, social interaction deficits have been shown (Neill et al., 2010). DISC-1 mutations known to cause schizophrenia-like abnormalities in rodents can impairs cognitive and social behaviors in some transgenic mice (Shevelkin et al., 2017; Sultana, Lee, 2020), but not in rats (Li, Zhang, 2017; Glenn et al., 2021). Research of knockouts of the Neuregulin-1 (NRG1) gene which has been identified as a candidate susceptibility gene for schizophrenia, revealed a selective impairment in response to social novelty in NRG1 mutants, but not in sociability (O'Tuathaigh et al., 2007). Developmental models of schizophrenia, such as using neonatal lesions of the rat ventral hippocampus or prenatal administration of methylazomethanol into pregnant rats, result in deficits in social behavior, as well as impaired memory, and increased anxiety (Sams-Dodd et al., 1997; Winship et al., 2019). Another selective breeding model of psychopathology that exhibits increased freezing to context (but unlike rats of the GC strain only onto acute prior stress) is the Wistar Kyoto (WKY) rats (Nosek et al., 2008). WKY is a depression model characterized by elevated anxiety- and depression-like behavior. In the social interaction assessment, the WKY rats avoided contact with another rats (Nam et al., 2014).

However, in the resident-intruder test, the total time that the GC and PM rats spent on direct contact with a social object (intruder) was significantly higher than that of control. This discrepancy in social activity in the two tests may be explained by different environmental conditions that affect emotional state and motivation. Stress level of the three-chamber test is mostly caused by placing the experimental animal into a novel environment by the experimenter. Earlier it was shown that the GC rats react more strongly even to the handling required to place the animal in the experimental setup: corticosterone concentrations were increased during handling, but reduced at rest (Alekhina et al., 2016). Such an increased stress reactivity in the GC rats may explain the decrease in sociability index in the three-chamber test due to the passive-defensive reaction in response to handling stress in the three-chamber test, but not in the resident-intruder test, which does not require handling. Previously, it was shown that passive-defensive reflex expressed in the form of catatonic stupor is of dominant character and significantly prevails over cognitive and alimentary reflexes (Petrova, 1990). The results of this work suggest that the predisposition to catatonic stupor also negatively affects social motivation during testing in the three-chamber paradigm.

The data on the increase in social contacts of both GC rats and PM rats in the resident-intruder test shown in this work are of interest. It is known that an increase in social interaction in rodents can be achieved in certain ways, such as medial prefrontal cortex lesions (Gonzalez et al., 2000) or low doses of ethanol (Varlinskaya et al., 2001). At the neurochemical level, a wide variety of systems have been examined for their role in the normal expression of social behavior (Crowley et al., 1989). Oxytocin, vasopressin, endogenous opioids and catecholamines appear to participate in a wide variety of affiliative behaviors (Nelson, Panksepp, 1998). Acute administration of opiate drugs, low dose morphine and naltrexone produced a more robust attenuation of social investigation than non-social exploratory activity in rats. Amphetamine increased both forms of investigation and haloperidol had the opposite effect (Deak et al., 2009). More than that, there is evidence of the involvement of the glutamate system in the formation of social deviations. D-Cycloserine, a partial agonist at the glycine recognition site of the glutamatergic NMDA receptor, can increase social investigation and sexual behavior and decrease aggressiveness in mice (McAllister, 1994). There are results supporting a role of glutamate receptors subunits in the modulation of social behavior (Vekovischeva et al., 2004; Adamczyk et al., 2012), however the study of the glutamate receptors genes mRNA in the hippocampus and frontal cortex of the GC rats did not reveal any changes (Plekanchuk, Ryazanova, 2021).

The increased mounting in the GC and PM rats shown in this paper may be indicative of aggressiveness between rats of the same sex. It has previously been shown that the

GC rats demonstrate a high level of shock-induced aggression (Nikulina et al., 1987), but not aggression towards male rats or interspecies aggression towards mice (Alekhina et al., 1987). In addition, both PM and GC rats have an increased aggressive response towards humans (Alekhina et al., 2016; Alekhina, Kozhemyakina, 2019).

Considering the fact that PM rats, in addition to catatonic symptoms, have a predisposition to audiogenic seizures, the connection between epilepsy and psychopathology in humans should be mentioned. Many symptoms of neurologic or psychiatric illnesses – such as cognitive impairment, depression, anxiety, attention deficits – occur more frequently in people with epilepsy than in the general population (Brooks-Kayal et al., 2013). The rat lines selectively bred for differences in amygdala excitability, manifested by "fast" or "slow" kindling epileptogenesis, display several comorbid features related to anxiety and learning. Seizure-prone genetic background provides poorer original learning and easier disruption of new learning, as well as increased anxiety and impulsivity (McLntyre et al., 2004). Rats in the chronic phase of the lithium-pilocarpine model of epilepsy showed disturbed communicative behavior, with impaired social behavioral patterns, increased motor activity and impaired memory function (Smolensky et al., 2019). Aggression is one of several psychiatric disorders that is observed, among others, in epileptic patients (Deb et al., 2020). This association has been reliably replicated in several animal models including those using pilocarpine (Desjardins et al., 2001) and domoic acid (Fuquay et al., 2012), in which aggression develops either in parallel to spontaneous seizures or precedes the development of recurrent seizures. The increased offensive behavior of the PM rats in the resident-intruder test shown in this work may confirm the likely relationship between seizure predisposition and aggressiveness.

Experiment 2. The Barnes maze task

Rodent basal cognitive abilities include, along with elementary logic tasks solutions and generalization capacity of a low level, spatial behavior and memory. This type of cognitive ability requires the formation of mental representations of spatial environmental characteristics (Poletaeva, Zorina, 2014). To test the acquisition of spatial memory in PM and GC, the Barnes maze was used. The increased time required to search for the target hole in the GC rats may indicate impaired spatial learning. However, a decreased locomotor activity has earlier been shown in rats of this line (Petrova, 1990), and to assess whether potential disturbances are in fact memory impairments it is necessary to take into account such parameters as primary errors and search strategy. No differences were shown for these parameters in GC compared to WAG. The reduced GC rats velocity of movement across the maze field for 3–7 days confirms the effect of motor activity on latency to first target visits. The GC rats appear to have no learning impairment in this test. The fact that the velocity of movement of the GC rats in the field does not differ in the first two days of testing, but is less than in the control on the following days, may indicate a slower adaptation to new conditions.

Estimation of the search strategy showed differences in PM in comparison with Wistar. After a few days of training, non-cognitively impaired animals frequently use the spatial strategy to resolve the BM task. The fact that after a few days of learning trials the PM rats still use mixed (i. e., random) and serial strategies instead of spatial to resolve the maze means that they are cognitively impaired and do not employ spatial clues to reach the target hole (Yassine et al., 2013). It has previously been shown that the PM rats exhibit longer latency and lower rate of successful trials in the Morris water test, at the same time, the GC rats did not differ from the control in these parameters (Barykina et al., 2009). The Morris water maze is more stressful for animals than the Barnes maze, because there is water immersion (Gawel et al., 2019). Water-maze training induced greater increases in plasma corticosterone which may affect the performance of animals (Harrison et al., 2009). In addition, the GC rats are inclined to passive drift and longer floating episodes in the Morris water test (Barykina et al., 2009). Low movement speed in Barnes's maze, high time of inactivity and low exploratory activity in the resident-intruder test in the GC rats are caused by catatonic freezing. The data shown in this work confirm the manifestation of catatonic inhibition by the GC rats in different stressful situations (Barykina et al., 2009).

Conclusion

Selection for the duration of catatonic freezing and the amplitude of pendulum-like movements influenced social interactions and learning in the GC and PM rat strains. In particular, the GC rats have increased or decreased social interactions depending on the environment, and a negative relationship between catatonic freezing and sociality were shown in this work. The PM rats show increased social activity and offensive aggression in the resident-intruder test. Except for the reduced velocity of movement across the maze field, the GC rats appear to have no difficulty in solving the Barnes maze, whereas the PM rats do not use a spatial strategy in the maze, which may indicate impairment of learning and spatial memory.

References

Adamczyk A., Mejias R., Takamiya K., Yocum J., Krasnova I.N., Caldero J., Wang T. GluA3-deficiency in mice is associated with increased social and aggressive behavior and elevated dopamine in striatum. *Behav. Brain Res*. 2012;229(1):265-272. DOI 10.1016/j.bbr.2012.01.007.

Alekhina T.A., Kozhemyakina R.V. Modeling of focal seizures with automatisms in rats with pendulum movements. *Bull. Exp. Biol. Med.* 2019;168(2):300-303. DOI 10.1007/s10517-019-04695-7.

Alekhina T.A., Palchikova N.A., Igonina T.N., Kuznetsova N.V. Comparative analysis of imipramine intake reactions in catatonic and wistar rats. *Rossiiskii Fiziologicheskii Zhurnal im. I.M. Sechenova = Russian Journal of Physiology*. 2015;101(3):249-257. (in Russian)

Alekhina T.A., Palchikova N.A., Kozhemyakina R.V., Prokudina O.I. The signs of destabilization in behavioral and somatovegetative parameters of rats selected for catatonia. *Russ. J. Genet. Appl. Res.* 2016;6(8):798-803. DOI 10.1134/S2079059716080025.

Alekhina T.A., Prokudina O.I., Ryazanova M.A., Ukolova T.N., Barykina N.N., Kolpakov V.G. Typological characteristics of behavior in strains of rats bred for enhancement and absence of pendulum movements. Association with brain monoamines. *Zhurnal Vysshey Nervnoy Deyatel'nosti im. I.P. Pavlova = I.P. Pavlov Journal of Higher Nervous Activity*. 2007;57(3):336-343. (in Russian)

Alekhina T.A., Shtilman N.I., Nikulina E.M., Pavlov I.F., Barykina N.N. Aggression and learning in a strain of rats predisposed to catalepsy. *Zhurnal Vysshey Nervnoy Deyatel'nosti im. I.P. Pavlova = I.P. Pavlov Journal of Higher Nervous Activity*. 1987;37(3):537-541. (in Russian)

Anderzhanova E., Kirmeier T., Wotjak C.T. Animal models in psychiatric research: the RDoC system as a new framework for endophenotype-oriented translational neuroscience. *Neurobiol. Stress.* 2017;7:47-56. DOI 10.1016/j.ynstr.2017.03.003.

Barykina N.N., Chugui V.F., Alekhina T.A., Ryazanova M.A., Ukolova T.N., Sakharov D.G., Kolpakov V.G. Learning of rats predisposed to catalepsy in Morris water test. *Zhurnal Vysshey Nervnoy Deyatel'nosti im. I.P. Pavlova = I.P. Pavlov Journal of Higher Nervous Activity*. 2009;59(6):728-735. (in Russian)

Brooks-Kayal A.R., Bath K.G., Berg A.T., Galanopoulou A.S., Holmes G.L., Jensen F.E., Scharfman H.E. Issues related to symptomatic and disease-modifying treatments affecting cognitive and neuropsychiatric comorbidities of epilepsy. *Epilepsia.* 2013;54 (Suppl.4):44-60. DOI 10.1111/epi.12298.

Crowley W.R., O'Connor L.H., Feder H.H. Neurotransmitter systems and social behavior. In: Balthazart J. (Ed.) Molecular and Cellular Basis of Social Behavior in Vertebrates. Advances in Comparative and Environmental Physiology. Vol. 3. Berlin; Heidelberg; Springer, 1989;161-208. DOI 10.1007/978-3-642-73827-2_4.

Deak T., Arakawa H., Bekkedal M.Y., Panksepp J. Validation of a novel social investigation task that may dissociate social motivation from exploratory activity. *Behav. Brain Res.* 2009;199(2):326-333. DOI 10.1016/j.bbr.2008.12.011.

Deb S., Brizard B.A., Limbu B. Association between epilepsy and challenging behaviour in adults with intellectual disabilities: systematic review and meta-analysis. *BJPsych Open.* 2020;6(5):e114. DOI 10.1192/bjo.2020.96.

Desjardins D., Parker G., Cook L.L., Persinger M.A. Agonistic behavior in groups of limbic epileptic male rats: pattern of brain damage and moderating effects from normal rats. *Brain Res.* 2001;905(1-2): 26-33. DOI 10.1016/S0006-8993(01)02454-4.

Fink M., Taylor M.A. The many varieties of catatonia. *Eur. Arch. Psychiatry Clin. Neurosci.* 2001;251(Suppl.1):I/8-I/13. DOI 10.1007/pl00014200.

Fuquay J.M., Muha N., Pennington P.L., Ramsdell J.S. Domoic acid induced status epilepticus promotes aggressive behavior in rats. *Physiol. Behav.* 2012;105(2):315-320. DOI 10.1016/j.physbeh.2011.08.013.

Gawel K., Gibula E., Marszalek-Grabska M., Filarowska J., Kotlinska J.H. Assessment of spatial learning and memory in the Barnes maze task in rodents – methodological consideration. *Naunyn-Schmiedeberg's Arch. Pharmacol.* 2019;392(1):1-18. DOI 10.1007/s00210-018-1589-y.

Glenn M.J., Batallán Burrowes A.A., Yu W., Blackmer-Raynolds L., Norchi A., Doak A.L. Progression of behavioral deficits during periadolescent development differs in female and male DISC1 knockout rats. *Genes Brain Behav.* 2021;e12741. DOI 10.1111/gbb.12741.

Gonzalez L.E., Rujano M., Tucci S., Paredes D., Silva E., Alba G., Hernandez L. Medial prefrontal transection enhances social interaction: I: Behavioral studies. *Brain Res.* 2000;887(1):7-15. DOI 10.1016/S0006-8993(00)02931-0.

Harrison F.E., Hosseini A.H., McDonald M.P. Endogenous anxiety and stress responses in water maze and Barnes maze spatial memory tasks. *Behav. Brain Res.* 2009;198(1):247-251. DOI 10.1016/j.bbr.2008.10.015.

Jones C.A., Watson D.J.G., Fone K.C.F. Animal models of schizophrenia. *Br. J. Pharmacol*. 2011;164(4):1162-1194. DOI 10.1111/j.1476-5381.2011.01386.x.

Kaidanovich-Beilin O., Lipina T., Vukobradovic I., Roder J., Woodgett J.R. Assessment of social interaction behaviors. *J. Vis. Exp.* 2011;48:e2473. DOI 10.3791/2473.

Kiser D.P., Rivero O., Lesch K.P. Annual research review: the (epi)genetics of neurodevelopmental disorders in the era of whole-genome sequencing – unveiling the dark matter. *J. Child Psychol. Psychiatry.* 2015;56(3):278-295. DOI 10.1111/jcpp.12392.

Koolhaas J.M., Coppens C.M., de Boer S.F., Buwalda B., Meerlo P., Timmermans P.J. The resident-intruder paradigm: a standardized test for aggression, violence and social stress. *J. Vis. Exp.* 2013;77:e4367. DOI 10.3791/4367.

Krueger R.F., Kotov R., Watson D., Forbes M.K., Eaton N.R., Ruggero C.J., Zimmermann J. Progress in achieving quantitative classification of psychopathology. *World Psychiatry.* 2018;17(3):282-293. DOI 10.1002/wps.20566.

Li M., Zhang M. SU10. Behavioral characteristics of a DISC1 knockout rat model. *Schizophr. Bull.* 2017;43(Suppl.1):S164. DOI 10.1093/schbul/sbx024.009.

McAllister K.H. D-cycloserine enhances social behaviour in individually-housed mice in the resident-intruder test. *Psychopharmacology.* 1994;116(3):317-325. DOI 10.1016/0031-9384(86)90007-7.

McLntyre D.C., McLeod W.S., Anisman H. Working and reference memory in seizure-prone and seizure-resistant rats: impact of amygdala kindling. *Behav. Neurosci.* 2004;118(2):314-323. DOI 10.1037/0735-7044.118.2.314.

Nam H., Clinton S.M., Jackson N.L., Kerman I.A. Learned helplessness and social avoidance in the Wistar-Kyoto rat. *Front. Behav. Neurosci*. 2014;8:109. DOI 10.3389/fnbeh.2014.00109.

Nani J.V., Rodríguez B., Cruz F.C., Hayashi M.A.F. Animal models in psychiatric disorder studies. In: Tvrdá E., Yenisetti S.C. (Eds.) Animal Models in Medicine and Biology. IntechOpen, 2019. DOI 10.5772/intechopen.89034.

Neill J.C., Barnes S., Cook S., Grayson B., Idris N.F., McLean S.L., Harte M.K. Animal models of cognitive dysfunction and negative symptoms of schizophrenia: focus on NMDA receptor antagonism. *Pharmacol. Ther.* 2010;128(3):419-432. DOI 10.1016/j.pharmthera.2010.07.004.

Nelson E.E., Panksepp J. Brain substrates of infant–mother attachment: contributions of opioids, oxytocin, and norepinephrine. *Neurosci. Biobehav. Rev.* 1998;22(3):437-452. DOI 10.1016/S0149-7634(97)00052-3.

Nikulina E.M., Popova N.K., Kolpakov V.G., Alekhina T.A. Brain dopaminergic system in rats with a genetic predisposition to catalepsy. *Biog. Amines.* 1987;4(4-6):399-406.

Nosek K., Dennis K., Andrus B.M., Ahmadiyeh N., Baum A.E., Woods L.C.S., Redei E.E. Context and strain-dependent behavioral response to stress. *Behav. Brain Funct.* 2008;4(1):1-8. DOI 10.1186/1744-9081-4-23.

O'Tuathaigh C.M.P., Babovic D., O'Sullivan G.J., Clifford J.J., Tighe O., Croke D.T., Waddington J.L. Phenotypic characterization of spatial cognition and social behavior in mice with 'knockout' of the schizophrenia risk gene neuregulin 1. *Neuroscience.* 2007;147(1): 18-27. DOI 10.1016/j.neuroscience.2007.03.051.

Petrova E.V. Features of changes in congenital and acquired forms of behavior in rats with genetic catalepsy. *Zhurnal Vysshey Nervnoy Deyatel'nosti im. I.P. Pavlova = I.P. Pavlov Journal of Higher Nervous Activity.* 1990;40(3):475-480. (in Russian)

Plekanchuk V.S., Ryazanova M.A. Expression of glutamate receptor genes in the hippocampus and frontal cortex in GC rat strain with genetic catatonia. *J. Evol. Biochem. Physiol*. 2021;57(1):156-163. DOI 10.1134/s0022093021010154.

Poletaeva I.I., Zorina Z.A. A genetic approach to the study of simple cognitive abilities in animals. *Rossiyskiy Zhurnal Kognitivnoy Nauki = Russian Journal of Cognitive Science.* 2014;1(3):31-55.

Powell C.M., Miyakawa T. Schizophrenia-relevant behavioral testing in rodent models: a uniquely human disorder? *Biol. Psychiatry.* 2006;59(12):1198-1207. DOI 10.1016/j.biopsych.2006.05.008.

Ryazanova M.A., Igonina T.N., Alekhina T.A., Prokudina O.I. The increase in the proportion of nervous animals bred for catatonia: the participation of central adrenoreceptors in catatonic reactions. *Russ. J. Genet*. 2012;48:1141-1147. DOI 10.1134/S1022795412100092.

Ryazanova M.A., Prokudina O.I., Plekanchuk V.S., Alekhina T.A. Expression of catecholaminergic genes in the midbrain and prepulse inhibition in rats with a genetic catatonia. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding.* 2017;21(7):798-803. DOI 10.18699/VJ17.296. (in Russian)

Sams-Dodd F., Lipska B.K., Weinberger D.R. Neonatal lesions of the rat ventral hippocampus result in hyperlocomotion and deficits in social behaviour in adulthood. *Psychopharmacology*. 1997;132(3): 303-310. DOI 10.1007/s002130050349.

Samsom J.N., Wong A.H.C. Schizophrenia and depression co-morbidity: what we have learned from animal models. *Front. Psychiatry*. 2015;6:13. DOI 10.3389/fpsyt.2015.00013.

Shevelkin A.V., Terrillion C.E., Abazyan B.N., Kajstura T.J., Jouroukhin Y.A., Rudow G.L., Pletnikov M.V. Expression of mutant DISC1 in Purkinje cells increases their spontaneous activity and impairs cognitive and social behaviors in mice. *Neurobiol. Dis.* 2017;103:144-153. DOI 10.1016/j.nbd.2017.04.008.

Smolensky I.V., Zubareva O.E., Kalemenev S.V., Lavrentyeva V.V., Dyomina A.V., Karepanov A.A., Zaitsev A.V. Impairments in cognitive functions and emotional and social behaviors in a rat lithium-pilocarpine model of temporal lobe epilepsy. *Behav. Brain Res.* 2019;372:112044. DOI 10.1016/j.bbr.2019.112044.

Stansley B.J., Yamamoto B.K. Behavioral impairments and serotonin reductions in rats after chronic L-dopa. *Psychopharmacology.* 2015;232(17):3203-3213. DOI 10.1007/s00213-015-3980-4.

Sultana R., Lee C.C. Expression of behavioral phenotypes in genetic and environmental mouse models of schizophrenia. *Front. Behav. Neurosci.* 2020;14:29. DOI 10.3389/fnbeh.2020.00029.

Timofeeva A.S. (Ed.) Genetic and Evolutionary Problems in Psychiatry. Novosibirsk: Nauka Publ., 1985. (in Russian)

Varlinskaya E.I., Spear L.P., Spear N.E. Acute effects of ethanol on behavior of adolescent rats: role of social context. *Alcohol. Clin. Exp. Res*. 2001;25(3):377-385. DOI 10.1111/j.1530-0277.2001.tb02224.x.

Vekovischeva O.Y., Aitta-aho T., Echenko O., Kankaanpää A., Seppälä T., Honkanen A., Korpi E.R. Reduced aggression in AMPA-type glutamate receptor GluR-A subunit-deficient mice. *Genes Brain Behav.* 2004;3(5):253-265. DOI 10.1111/j.1601-1848.2004.00075.x.

Volavka J., Citrome L. Heterogeneity of violence in schizophrenia and implications for long-term treatment. *Int. J. Clin. Pract.* 2008;62(8):1237-1245. DOI 10.1111/j.1742-1241.2008.01797.x.

Wilson J.E., Niu K., Nicolson S.E., Levine S.Z., Heckers S. The diagnostic criteria and structure of catatonia. *Schizophr. Res.* 2015;164(1-3): 256-262. DOI 10.1016/j.schres.2014.12.036.

Winship I.R., Dursun S.M., Baker G.B., Balista P.A., Kandratavicius L., Maia-de-Oliveira J.P., Howland J.G. An overview of animal models related to schizophrenia. *Can. J. Psychiatry*. 2019;64(1): 5-17. DOI 10.1177/0706743718773728.

Yassine N., Lazaris A., Dorner-Ciossek C., Després O., Meyer L., Maitre M., Mathis C. Detecting spatial memory deficits beyond blindness in tg2576 Alzheimer mice. *Neurobiol. Aging.* 2013;34(3):716-730. DOI 10.1016/j.neurobiolaging.2012.06.016.

ORCID ID

V.S. Plekanchuk orcid.org/0000-0002-3492-0513

Acknowledgements. The authors thank R.V. Kozhemyakina, Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia, for help with planning the experiment. This work was supported by the Ministry of Education and Science of Russian Federation, grant No. 2019-0546 (FSUS-2020-0040) and budget project No. FWNR-2022-0019.

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

Received September 17, 2021. Revised November 13, 2021. Accepted December 1, 2021.

Вавиловский журнал генетики и селекции. 2022;26(3):290-297
DOI 10.18699/VJGB-22-36

Original Russian text www.bionet.nsc.ru/vogis/

Molecular-genetic approaches to species identification of platyhelminthes of the genus *Ligophorus* (Monogenea) parasitising flathead mullet

E.A. Vodiasova[1] ✉, E.S. Chelebieva[1], O.V. Shikhat[1], D.M. Atopkin[2], E.V. Dmitrieva[1]

[1] A.O. Kovalevsky Institute of Biology of the Southern Seas of the Russian Academy of Sciences, Sevastopol, Russia
[2] Federal Scientific Center of the East Asia Terrestrial Biodiversity, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, Russia
✉ eavodiasova@gmail.com

Abstract. *Mugil cephalus* L., 1758 (flathead mullet) is a valuable commercial fish and a promising object of artificial breeding in the Black Sea and the Sea of Azov, and the study of its parasite fauna is important for fishery and mariculture. Monogeneans of the genus *Ligophorus* are common ectoparasites dwelling on the gills of mullets. Two representatives of this genus parasitise flathead mullet in the Azov-Black Sea region, namely *Ligophorus mediterraneus* Sarabeev, Balbuena et Euzet, 2005 and *Ligophorus cephali* Rubtsova, Balbuena, Sarabeev, Blasco-Costa et Euzet, 2006. Morphological identification of these species requires spending much time and a high level of experience in monogenean taxonomy. For quick and correct species identification of these parasites, we have developed a genotyping approach based on the polymerase chain reaction of allele-specific gene sites for various Monogenea species. A fragment of the 28S ribosomal gene, which includes conserved and variable sites, was chosen as a genetic marker. Three approaches were used as follows: amplified fragment length analysis, allele-specific PCR with endpoint detection and allele-specific real-time PCR using SYBR Green intercalating dye. The first approach was by obtaining PCR products of different lengths that were specific either to *L. mediterraneus* or to *L. cephali*. This approach was implemented due to the presence of several variable sites located at a distance from each other. The PCR mixture contained three primers: one forward and two reverse. The forward primer was complementary to the conserved site, which did not differ between species. Reverse primers were species-specific and, for each species, they were complementary to different DNA regions located 100 bp apart. As a result, *L. mediterraneus* was characterized by shorter amplicons than *L. cephali*. For the second and third approaches, a pair of primers was designed according to the following principle: the forward primer was complementary to both species, since it was selected for the conserved gene region. Reverse primers were species-specific and were designed for the 28S variable region. The two parasite species were distinguished by three-point mutations. Thus, one pair of primers was complementary to *L. mediterraneus*, the other, to *L. cephali*. The amplified fragment length analysis and the allele-specific real-time PCR demonstrated 100 % coincidence of genotyping results compared with Sanger sequencing. The developed genotyping protocols can be used not only to distinguish two species of *Ligophorus* from flathead mullet in ecological studies and veterinary practice but also for further development of similar approaches for other monogeneans, among which there are many pathogenic species.
Key words: genotyping; allele-specific PCR; Monogenea; *Ligophorus*; *Mugil cephalus*.

For citation: Vodiasova E.A., Chelebieva E.S., Shikhat O.V., Atopkin D.M., Dmitrieva E.V. Molecular-genetic approaches to species identification of platyhelminthes of the genus *Ligophorus* (Monogenea) parasitising flathead mullet. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2022;26(3):290-297. DOI 10.18699/VJGB-22-36

Молекулярно-генетические подходы к видовой идентификации паразитических плоских червей рода *Ligophorus* (Monogenea), обитающих на лобане

Е.А. Водясова[1] ✉, Э.С. Челебиева[1], О.В. Шихат[1], Д.М. Атопкин[2], Е.В. Дмитриева[1]

[1] Федеральный исследовательский центр «Институт биологии южных морей им. А.О. Ковалевского Российской академии наук», Севастополь, Россия
[2] Федеральный научный центр биоразнообразия наземной биоты Восточной Азии, Дальневосточное отделение Российской академии наук, Владивосток, Россия
✉ eavodiasova@gmail.com

Аннотация. *Mugil cephalus* L., 1758 (лобан) – ценная промысловая рыба и перспективный объект разведения в бассейнах Черного и Азовского морей. Изучение его паразитофауны крайне важно для рыбного промысла и марикультуры. Одними из массовых эктопаразитов, обитающих на жабрах кефалевых, являются моногенеи рода *Ligophorus*. На лобане в Азово-Черноморском регионе паразитируют два представителя этого рода: *Ligophorus mediterraneus* Sarabeev, Balbuena et Euzet, 2005 и *Ligophorus cephali* Rubtsova, Balbuena, Sarabeev, Blasco-Costa et Euzet, 2006. Морфологическое определение этих видов весьма трудоемко и требует высокого

уровня квалификации. Для быстрого и точного различия двух названных видов паразитов нами разработан метод, основанный на ПЦР вариабельных участков рибосомного гена 28S. Данный ген широко используется для молекулярной таксономии рода *Ligophorus.* В настоящей работе применялись три подхода: анализ длины амплифицированных фрагментов, аллель-специфичная ПЦР с детекцией в конечной точке и аллель-специфичная ПЦР в реальном времени с использованием интеркалирующего красителя SYBR Green. Первый подход заключался в подборе праймеров для получения ПЦР-продуктов различной длины, которые были характерны для *L. mediterraneus* или *L. cephali.* Этот подход был осуществлен благодаря наличию нескольких вариабельных локусов, которые находятся на расстоянии друг от друга. ПЦР-смесь содержала три праймера: один прямой и два обратных. Прямой праймер был комплементарен консервативному участку, который не различался между видами. Обратные праймеры были видоспецифичны, при этом для каждого вида они были комплементарны различным участкам ДНК, которые расположены на удалении 100 п. н. друг от друга. В результате *L. mediterraneus* характеризовался более короткими ампликонами, чем *L. cephali.* Для второго и третьего подхода конструировалась пара праймеров по следующему принципу: прямой праймер был комплементарен обоим видам, так как подбирался к консервативному участку гена. Обратные праймеры были видоспецифичными и разрабатывались к вариабельному участку 28S. На этом участке два вида паразита различались тремя точечными мутациями. Таким образом, одна пара праймеров была комплементарна *L. mediterraneus*, вторая – *L. cephali.* Анализ длины амплифицированных фрагментов и аллель-специфичная ПЦР в реальном времени продемонстрировали 100 % совпадение результатов генотипирования при сравнении с морфологической идентификацией и секвенированием по Сэнгеру. Разработанные протоколы генотипирования могут быть использованы не только для различия обитающих на лобане двух видов *Ligophorus* при экологических исследованиях и в ветеринарной практике, но и для последующей разработки подобных методов для других моногеней, среди которых много патогенных видов.
Ключевые слова: генотипирование; аллель-специфическая ПЦР; Monogenea; *Ligophorus*; *Mugil cephalus.*

Introduction

Monogenea (Platyhelminthes: Monogenea) are parasites, mainly of fish, with a direct life cycle. Dozens of new taxa of these parasites are described each year. Their diversity has reached 5000 species (Vanhove, 2013), and many of them are of epizootic importance (Cribb et al., 2002; Bakke et al., 2007; Rubio-Godoy, 2007). The boundaries of most species are established based on morphological criteria, and for species identification, the shape and size of attachment disc structures are mainly used (Yamaguti, 1963; Gusev et al., 1985; Pugachev et al., 2009; Vignon, 2011; Strona et al., 2014; Kalafi et al., 2016). However, these structures have high intraspecific variability (Ergens, Gelnar, 1985; Caltran et al., 1995; Dmitrieva, Dimitrov, 2002; Olstad et al., 2009; Mladineo et al., 2013). The latter makes it very difficult to determine the species identity of monogeneans and raises the question of defining the framework of their intra- and interspecific variability. Appealing to real collection specimens to confirm the determination is often difficult due to accessing collections with type specimens. Comparison of organisms with many "similar" species from different areas based on brief descriptions and often inaccurate drawings does not always allow reliable species identification. As a result, an increase in the number of "false" and underestimation of "hidden" species taxa can lead to misunderstanding of the phylogeny, diversity and distribution of representatives of individual monogenean groups (Poisot et al., 2011), and sometimes to problems in determining the status of pathogenic species, as in the case of *Gyrodactylus salaris* and *G. thymalli* (Fromm et al., 2014; Mieszkowska et al., 2018).

Given the above, the development of approaches and methods allowing for the most accurate identification of monogenean species remains an urgent task, both in theoretical and practical terms. One of the promising directions in molecular genetic studies of parasites is the development of methods for genotyping species and local intraspecific groupings, both for biodiversity studies of individual taxa and for rapid diagnosis of species and their populations (Tokarev et al., 2015). Such works with relation to monogeneans are rare (Fromm et al., 2013, 2014; Mieszkowska et al., 2018). A few papers address the problems of DNA barcoding of monogenean species (Littlewood, 2008; Vanhove, 2013). Molecular studies on the genus *Ligophorus* Euzet et Suriano, 1977 are limited to a few studies, with 135 ribosomal nuclear DNA sequences deposited in the NCBI GenBank (as of 27.11.2021). The 18S, ITS1, 5.8S and 28S fragments were obtained for 12 species from the Mediterranean Sea and 2 species from the Black Sea (Blasco-Costa et al., 2012; Rodríguez-González et al., 2015). For two species off the coast of Brazil, 18S, ITS1, 5.8S and 28S were sequenced (Marchiori et al., 2015), and 18S, 28S and ITS1 fragments were obtained for 14 species from the Indian Ocean (Soo et al., 2015; Khang et al., 2016; Pakdee et al., 2019). Several studies (Blasco-Costa et al., 2012; Rodríguez-González et al., 2015; Khang et al., 2016) have compared morphological and genetic variability, showing a greater degree of congruence between phylogenetic reconstructions based on these data, suggesting that the use of ribosomal cluster sequences for genotyping species of this genus is promising.

The flathead mullet *Mugil cephalus* L., 1758 is a commercial fish of the Black and Azov Seas and a promising object of mariculture in the region; therefore, the study of its parasitofauna is critical not only from the scientific but also from the practical point of view. Monogeneans of the genus *Ligophorus*, which parasitise on the gills of mullets, are one of the ectoparasites for the flathead mullet. In the Azov-

Black Sea region, *L. mediterraneus* Sarabeev, Balbuena et Euzet, 2005 and *L. cephali* Rubtsova, Balbuena, Sarabeev, Blasco-Costa et Euzet, 2006 have morphologically similar attachment structures (Dmitrieva et al., 2009a, b), which makes their identification difficult. At the same time, these species have a good level of genetic divergence based on the variability of 28S and ITS1 (Blasco-Costa et al., 2012). This divergence is due to the single nucleotide substitutions characteristic of *L. cephali* and *L. mediterraneus*. When assessing the infestation of these species in large samples of fish, e. g. in ecological or veterinary surveys, the use of morphological characters is problematic, and sequencing followed by molecular taxonomy is costly and time-consuming. In addition, up to eight *Ligophorus* species may parasite on one individual mullet (Dmitrieva et al., 2012; Soo et al., 2015). This situation is not unique and occurs for species of the same genus in many members of the family Dactylogyridae, which includes *Ligophorus*.

With the appearance of real-time PCR, alternative approaches for genotyping based on allele-specific PCR that allows rapid and reliable species identification have begun to develop. However, for members of the family Dactylogyridae, such approaches have not been used. Thus, this work aimed to develop an express methodology to distinguish two monogenic species *L. cephali* and *L. mediterraneus* parasitising on the proboscis in the Azov-Black Sea region based on 28S gene variability. Considering that there are many representatives of epizootic importance among Dactylogyridae, the development of inexpensive and straightforward methods for rapid genotyping of species of this taxon to distinguish between pathogenic and nonpathogenic species, including at the larval stage, is also relevant in a practical sense.

Materials and methods

Sampling. The material for this work was 20 specimens of monogeneans of the genus *Ligophorus* collected from the gills of 3 individuals of *Mugil cephalus* in autumn 2019 in the Black Sea off the coast of Crimea, in Balaklava Bay. The worms were collected alive, a glycerol-gelatin preparation (Gusev, 1983) was prepared from part of an individual for species identification by the morphology of the attachment disc structures, and another part of the same monogenean was fixed in 96 % ethanol for molecular genetic studies.

Taxonomy identification. Species identification based on the shape and size of the haptoral structures and male copulatory organ of monogeneans using an Olympus CX41 microscope and phase-contrast optics at ×800–2,000 magnification according to the descriptions of *Ligophorus* species from Black Sea mullet (Dmitrieva et al., 2009a, b). Measurements and photographs were taken using CellSense digital image processing software.

DNA isolation and genetic analyses. The isolation was performed using a DNA-EXTRAN kit (Sintol Ltd., Russia). Each individual was incubated in 100 μL of lysis buffer (Sintol Ltd.) with 5 μL of proteinase K (Sintol Ltd.) and 1 μL of 2-mercaptoethanol at 56 °C for 3 hours. After lysis, samples were shaken for 20 s and further DNA extraction was performed according to the manufacturer's recommendations. DNA elution was performed in 30 μL. The isolated DNA was stored at –20 °C.

For the molecular taxonomy of the species, the 28S ribosomal gene, which is used in the analysis of this genus, was chosen as a genetic marker (Blasco-Costa et al., 2012; Soo et al., 2015; Pakdee et al., 2018). The 28S gene fragment was amplified using primers U178 (5′-GCACCCGCT GAAYTTAAG-3′) and LSU1200R (5′-GCATAGTTCAC CATCTTTCGG-3′) (Littlewood et al., 2000; Lockyer et al., 2003) according to the following protocol: pre-denaturation at 95 °C for 3 min followed by 38 cycles (denaturation at 94 °C for 40 s, annealing at 56 °C for 30 s, and elongation at 72 °C for 45 s). Each PCR reaction was performed in 25 μL of reaction mixture, containing 1–10 ng of matrix DNA, 0.4 μM of each primer, 5x ScreenMix PCR mix with Taq polymerase (Eurogen Ltd., Russia). Amplification products were detected by electrophoresis in 1 % agarose gel, staining with ethidium bromide and visualization under UV light. PCR products were sequenced in both directions using a standard BigDye Terminator Cycle Sequencing Ready Reaction Kit on an ABI PRISM 3130 analyzer (Applied Biosystems Inc.). The obtained 28S fragments were aligned in BIOEDIT software (Hall, 1999), *L. mediterraneus* (JN996829, JN996828, JN996827) and *L. cephali* (JN996830) were used as reference sequences. All nucleotide sequences obtained in this study are deposited in the GenBank: *L. mediterraneus* (MZ413895–MZ413898) and *L. cephali* (MZ413887–MZ413893).

Selection of primers for genotyping *L. mediterraneus* and *L. cephali*. Variability analysis of the 28S ribosomal gene fragment showed no intraspecific variability for this genetic marker. All nucleotide sequences for each monogenean species parasitising on flathead mullet from both the Mediterranean Sea (JN996829, JN996828, JN996827, JN996830) and the Black Sea (this work) were identical. Seven sequences of the 28S fragment for *L. mediterraneus* and eight for *L. cephali* were analysed. At the same time, several sites with mutations typical for *L. mediterraneus* and *L. cephali* in the region of 450–480, 540–570 and 680–705 bp were found (Fig. 1).

DNA regions that differed by at least 3 nucleotide substitutions between the two species were selected for genotyping. Primers flanking the polymorphic regions were designed using the internet resource https://benchling.com/. All developed primers are presented in Table 1.

All reverse primers were tested for their level of identity to other species using the blastn algorithm against the NCBI genetic database. Only the reverse primers were tested, as they are responsible for identifying the species. The primers CR450 and CR550 showed 100 % identity with 100 % coverage only to the species *L. cephali*. The MR450 primer, apart from 100 % identity to *L. mediterraneus*, also showed the same identity to *L. saladensis* (GenBank number KF442629). This species occurs only off the coast of Brazil and inhabits a different host, *Mugil liza*. The situation with primer CR650 is similar: in addition to 100 % identity with *L. cephali*, there is also 100 % identity with *L. heteronchus* (GenBank number JN996812). This parasite also inhabits another host, *Planiliza saliens*. Thus, among all known flathead mullet parasites, the developed primers allow identifying two species of *L. mediterraneus* and *L. cephali*, which makes it possible to use them not only in the Azov-Black Sea basin.

Analysis of amplified fragment lengths. Two versions of the primer mixture were selected for genotyping based

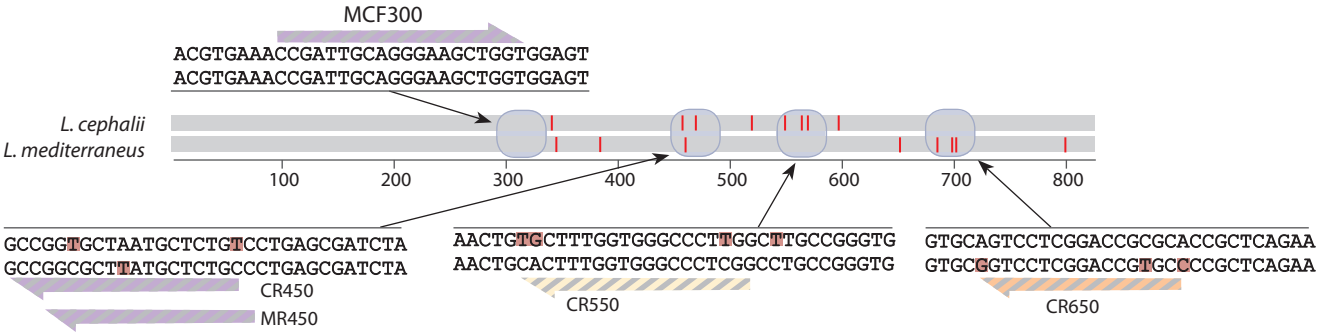


Fig. 1. Conserved and polymorphic regions of the 28S ribosomal gene for *L. mediterraneus* and *L. cephali*.

Table 1. Sequences of developed primers used for genotyping

Primer	Nucleotide sequence, 5′–3′	Annealing temperature, °C
MCF300	AAACCGATTGCAGGGAAGCTGG	59.8
CR450	GGACAGAGCATTAGCACCGGC	60.0
MR450	GGGCAGAGCATAAGCGCCG	60.7
CR550	AGCCAAGGGCCCACCAAAGCA	63.4
CR650	GTGCGCGGTCCGAGGACT	61.5

Table 2. Primers for different genotyping methods

Metod	Forward primer	Reverse primer	Amplicon length, bp	Species
Analysis of amplified fragment lengths	U178	MR450	630	*L. mediterraneus*
		CR550	750	*L. cephali*
	U178	MR450	630	*L. mediterraneus*
		CR650	880	*L. cephali*
Allele-specific end-point PCR	U178	MR450	630	*L. mediterraneus*
		CR450	630	*L. cephali*
Allele-specific real-time PCR	MCF300	MR450	170	*L. mediterraneus*
		CR450	170	*L. cephali*

on the analysis of different amplicon lengths. In the first case, the amplicon lengths specific to *L. mediterraneus* and *L. cephali* differed by 120 nucleotides, and in the second, by 250 nucleotides (Table 2). The essence of the approach we developed is as follows. Three primers are added to the PCR mix instead of two primers (as in traditional PCR). One primer (forward, U178) is complementary to the conserved region of 28S and will be annealed in both species accordingly. The second primer (reverse primer) was designed for a site that differs between the two species by several mutations. In this mixture it is the MR450 primer, which is complementary to the sequence specific to the *L. mediterraneus*. A third primer (reverse, CR550 or CR650) was also developed for a site that differs between the two species by several mutations, but it is complementary to *L. cephali*.

Thus, depending on the DNA matrix, only one of the two reverse primers will be annealed and the product will be produced. The reverse primers are chosen so that the product will be 630 bp long when MR450 is annealed, but the amplicon will be longer when the other reverse primer is annealed. So, with

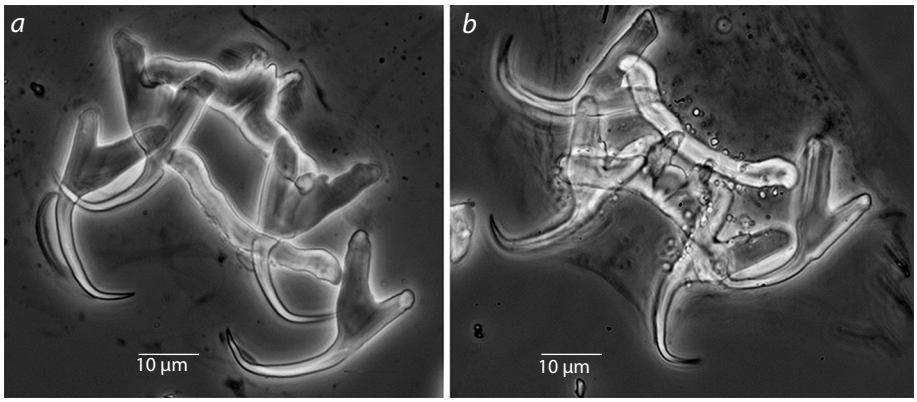


Fig. 2. Structures of the attachment discs of *L. cephali* (*a*) and *L. mediterraneus* (*b*) collected from the gills of *Mugil cephalus* in the Black Sea off the coast of Crimea.

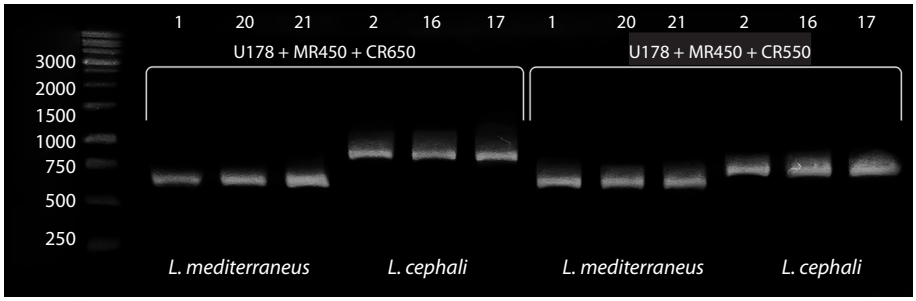


Fig. 3. Typing of DNA samples using amplified fragment length analysis.

Here and in Fig. 4: the numbers of test specimens and the primer mixture used are shown at the top, while the species that were identified based on morphology and molecular taxonomy are shown at the bottom.

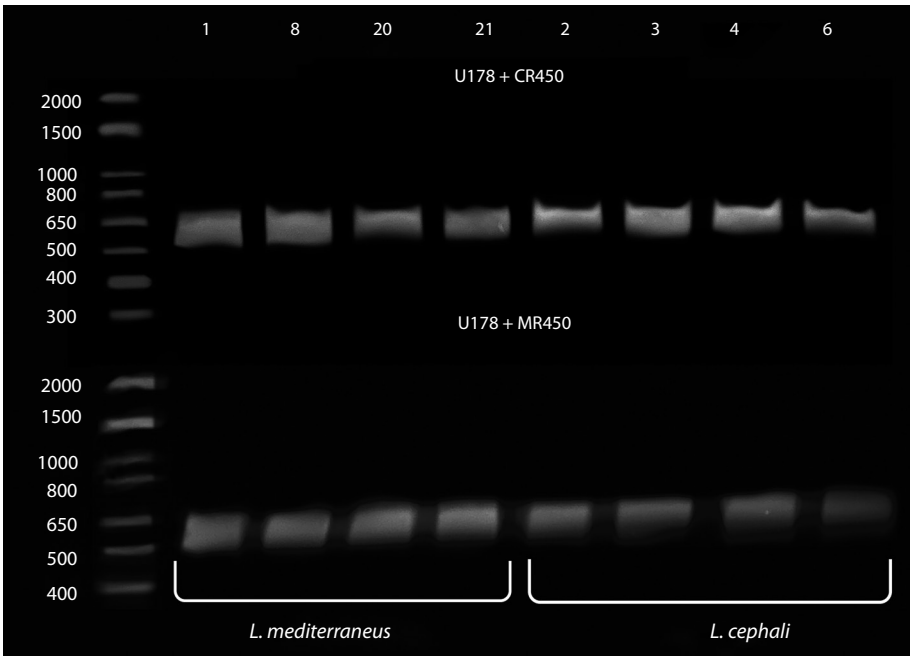


Fig. 4. Typing of DNA samples using allele-specific PCR with end-point detection.

the CR550 primer the length will be 750 bp and with the CR650 the length will be 880 bp. By performing a PCR reaction with the three primers, two *Ligophorus* species can be distinguished based on the length of the amplicons.

The PCR mixture and the amplification conditions were the same in both variants. The volume of the reaction mixture was 20 μL, and the final concentration of each primer (Eurogen, Russia) was 0.25 μM. The amplification was carried out according to the following protocol: pre-denaturation at 95 °C – 3 min followed by 38 cycles (denaturation at 94 °C – 40 s, annealing at 56 °C – 30 s, elongation at 72 °C – 45 s). The amplification products were detected by electrophoresis in 1 % agarose gel, staining with ethidium bromide and visualization in UV light. Monogenea species were characterized by their amplicon length, shown in Table 1.

Allele-specific end-point and real-time PCR. Genotyping based on allele-specific PCR with detection at the endpoint, as in real-time, was performed in a 20 μL reaction mixture. The final concentration of each primer (Eurogen, Russia) was 0.2 μM. The primer pairs used for each approach are listed in Table 2. Amplification with detection of PCR product at the end-point was performed according to the following protocol: pre-denaturation at 95 °C – 3 min followed by 38 cycles (denaturation at 94 °C – 40 s, annealing at 60 °C – 30 s, elongation at 72 °C – 30 s).

Each sample was analysed in three replicates when testing the genotyping method by allele-specific real-time PCR. The volume and composition of the reaction mixture were not changed, whereas the amplification conditions were changed: pre-denaturation at 95 °C for 3 min followed by 40 cycles (denaturation at 95 °C for 10 s, annealing at 60 °C for 10 s and synthesis at 72 °C for 30 s). In the end, a melting curve analysis was performed to evaluate the formation of primer dimers.

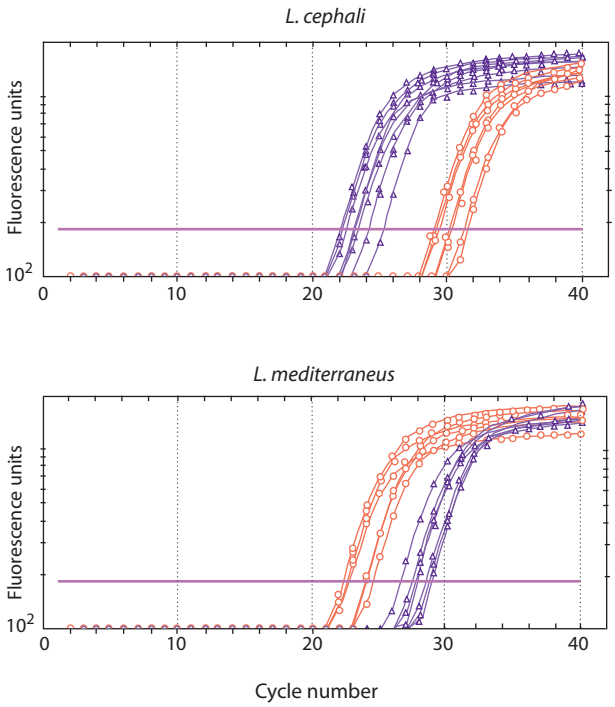


Fig. 5. Typing of DNA samples by allele-specific real-time PCR.

On the ordinate axis, the values are given in logarithmic scale. Purple indicates amplification curves for a mixture of primers MCF300+CR450, yellow for MCF300+MR450.

Results and discussion

Morphological species identification

Among the 20 collected specimens, 2 species were identified by morphology (Fig. 2): 9 specimens of *L. cephali*, sample numbers 2, 3, 4, 6, 7, 10, 16, 17, 19, and 11 specimens of *L. mediterraneus*, sample numbers 1, 5, 8, 11, 12, 13, 14, 15, 18, 20, 21.

Species identification using different genotyping methods

Morphological analysis was performed for all 20 individuals, based on which the monogenic species were identified. Analysis of the nucleotide sequences of the 28S ribosomal gene fragment obtained by Sanger sequencing confirmed the morphological identification of 11 individuals and allowed us to distinguish between the two species (see Fig. 1). All 20 *Ligophorus* individuals were then subjected to the methods described above for separating the two species by allele-specific PCR to assess their performance.

The method of genotyping based on PCR product length analysis is based on using two polymorphic regions of the 28S ribosomal gene and has been described in detail above. This approach separated *L. cephali* and *L. mediterraneus* species (Fig. 3). When both primer mixture variants (U178+MR450+CR650 and U178+MR450+CR550) were used, amplification of PCR products with only one reverse primer, which had complete complementarity to the 28S region, was observed for all individuals.

For genotyping based on allele-specific PCR with end-point detection, two amplification reactions with different primer compositions were performed for each sample. In one version, the reverse primer was complementary to the 28S gene region characteristic of *L. mediterraneus* (MR450); in the other, it was complementary to the same 28S gene region specific to *L. cephali* (CR450). The primers differed by 3 nucleotides. Using this approach, it was not possible to select amplification conditions that would not result in annealing of primers that are not fully complementary. As a result, when detected in an agarose gel at the end-point, PCR products were always detected, although with different intensities (Fig. 4).

At the same time, using this approach, but with real-time detection, allows the two species to be distinguished (Fig. 5). It is due to the different accumulation rates of PCR products when using fully and partially complementary primers. In this approach, the direct primer has been replaced to obtain shorter amplification products, which is recommended for real-time PCR. Two amplification reactions are also performed for each individual, and then the species is determined by the lower Ct value (number of the cycle in which the fluorescence signal crosses the threshold line). Accumulation of the products is faster when the primer and the matrix of the tested DNA are entirely complementary. 100 % concordance in identification by allele-specific PCR with real-time detection with morphological analysis and sequencing data was shown for all individuals.

Conclusion

This work developed a molecular genetic approach to rapidly distinguish between *L. mediterraneus* and *L. cephali* inhabiting *Mugil cephalus* in the Azov-Black Sea basin. Of the three approaches tested, two (amplified fragment length analysis method and allele-specific real-time PCR) allowed a reliable distinction between these two monogenean species. The use of allele-specific PCR with end-point detection of amplification products is inefficient because annealing and product accumulation occur for both primers complementary to *L. mediterraneus* and *L. cephali*. The approach using a PCR mixture containing three primers proposed in this work is the most cost-effective. The allele-specific real-time PCR method can be considered as the fastest and most efficient, the disadvantage of which is only its relatively high cost. Nevertheless, the developed approach is much faster and more cost-effective than sequencing the nucleotide sequences of the 28S ribosomal gene fragment.

The proposed genotyping methods can be used to rapidly separate two flatworms of the genus *Ligophorus* when assessing the degree of infestation of the flathead mullet with these parasites in the Azov-Black Sea region. It should also be noted that based on the data on 28S nucleotide sequences for other parasites of this genus, our developed primers have 100 % identity only with these two species of all that inhabit the flathead mullet. It allows them to be used in other parts of the world's oceans as well. The developed approach is vital when carrying out various works to study these species, such as studying the distribution of these species, changes in the ratio of two species on one host individual, competition of these species, the influence of various factors on their abundance, etc. In addition, the results obtained

demonstrate the promise of developing such approaches to estimate the abundance of other monogenic species, including pathogens.

References

Bakke T.A., Cable J., Harris P.D. The biology of gyrodactylid monogeneans: the "Russian-Doll Killers". *Adv. Parasitol*. 2007;64:161-376. DOI 10.1016/S0065-308X(06)64003-7.

Blasco-Costa I., Mìguez-Lozano R., Sarabeev V., Balbuena J.A. Molecular phylogeny of species of *Ligophorus* (Monogenea: Dactylogyridae) and their affinities within the Dactylogyridae. *Parasitol. Int*. 2012;61(6):619-627. DOI 10.1016/j.parint.2012.06.004.

Caltran H., Silan P., Roux M. *Ligophorus imitans* (Monogenea) Ectoparasite de *Liza ramada* (Teleostei). I. Populations naturelles et variabilité morphologique. *Ecologie*. 1995;26(2):95-104.

Cribb T.H., Chisholm L.A., Bray R.A. Diversity in the Monogenea and Digenea: does lifestyle matter? *Int. J. Parasitol*. 2002;32(3):321-328. DOI 10.1016/S0020-7519(01)00333-2.

Dmitrieva E., Dimitrov G. Variability in the taxonomic characters of Black Sea gyrodactylids (Monogenea). *Syst. Parasitol*. 2002;51(3):199-206. DOI 10.1023/A:1014594614921.

Dmitrieva E.V., Gerasev P.I., Gibson D.I., Pronkina N.V., Galli P. Descriptions of eight new species of *Ligophorus* Euzet & Suriano, 1977 (Monogenea: Ancyrocephalidae) from Red Sea mullets. *Syst. Parasitol*. 2012;81(3):203-237. DOI 10.1007/s11230-011-9341-8.

Dmitrieva E.V., Gerasev P.I., Merella P., Pugachev O.N. Redescription of *Ligophorus mediterraneus* Sarabeev, Balbuena & Euzet, 2005 (Monogenea: Ancyrocephalidae) with some methodological notes. *Syst. Parasitol*. 2009a;73(2):95-105. DOI 10.1007/s11230-009-9177-7.

Dmitrieva E.V., Gerasev P.I., Merella P., Pugachev O.N. Redescriptions of *Ligophorus cephali* Rubtsova, Balbuena, Sarabeev, Blasco-Costa & Euzet, 2006 and *L. chabaudi* Euzet & Suriano, 1977 (Monogenea: Ancyrocephalidae), with notes on the functional morphology of the copulatory organ. *Syst. Parasitol*. 2009b;73(3):175-191. DOI 10.1007/s11230-009-9192-8.

Ergens R., Gelnar M. Experimental verification of the effect of temperature on the size of hard parts of opisthaptor of *Gyrodactylus katharineri* Malmberg, 1964 (Monogenea). *Folia Parasitologica*. 1985;32(4):377-380.

Fromm B., Burow S., Hahn C., Bachmann L. MicroRNA loci support conspecificity of *Gyrodactylus salaris* and *Gyrodactylus thymalli* (Platyhelminthes: Monogenea). *Int. J. Parasitol*. 2014;44(11):787-793. DOI 10.1016/j.ijpara.2014.05.010.

Fromm B., Worren M.M., Hahn C., Hovig E., Bachmann L. Substantial loss of conserved and gain of novel microRNA families in flatworms. *Mol. Biol. Evol*. 2013;30(12):2619-2628. DOI 10.1093/molbev/mst155.

Gusev A.V. Method for Collecting and Processing Materials on Monogeneans Parasitizing Fish. Leningrad: Nauka Publ., 1983. (in Russian)

Gusev A.V., Pugachev O.N., Ergens R.R., Khotenovskiy I.A. Class Monogenea. In: Bauer O.N., Gusev A.V. (Eds.) Key to the Parasites of Freshwater Fish of the USSR Fauna. Vol. 2. Leningrad: Nauka Publ., 1985;10-387. (in Russian)

Hall T.A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser. 41*. 1999;95-98.

Kalafi E.Y., Tan W.B., Town C., Dhillon S.K. Automated identification of Monogeneans using digital image processing and K-nearest neighbour approaches. *BMC Bioinform*. 2016;17(19):511. DOI 10.1186/s12859-016-1376-z.

Khang T.F., Soo O.Y.M., Tan W.B., Lim L.H.S. Monogenean anchor morphometry: systematic value, phylogenetic signal, and evolution. *PeerJ*. 2016;4:e1668. DOI 10.7717/peerj.1668.

Littlewood D.T.J. Platyhelminth systematics and the emergence of new characters. *Parasite*. 2008;15:333-341. DOI 10.1051/parasite/2008153333.

Littlewood D.T.J., Curini-Galletti M., Herniou E.A. The interrelationships of Proseriata (Platyhelminthes: Seriata) tested with molecules and morphology. *Mol. Phylogenet. Evol*. 2000;16(3):449-466. DOI 10.1006/mpev.2000.0802.

Lockyer A.E., Olson P.D., Littlewood D.T.J. Utility of complete large and small subunit rRNA genes in resolving the phylogeny of the Neodermata (Platyhelminthes): implications and a review of the cercomer theory. *Biol. J. Linn. Soc*. 2003;78(2):155-171. DOI 10.1046/j.1095-8312.2003.00141.x.

Marchiori N.C., Pariselle A., Pereira J.J., Agnèse J.-F., Durand J.-D., Vanhove M.P.M. A comparative study of *Ligophorus uruguayense* and *L. saladensis* (Monogenea: Ancyrocephalidae) from *Mugil liza* (Teleostei: Mugilidae) in southern Brazil. *Folia Parasitol*. 2015;62:024. DOI 10.14411/fp.2015.024.

Mieszkowska A., Górniak M., Jurczak-Kurek A., Ziętara M.S. Revision of *Gyrodactylus salaris* phylogeny inspired by new evidence for Eemian crossing between lineages living on grayling in Baltic and White sea basins. *PeerJ*. 2018;6:e5167. DOI 10.7717/peerj.5167.

Mladineo I., Šegvić-Bubić T., Stanić R., Desdevises Y. Morphological plasticity and phylogeny in a monogenean parasite transferring between wild and reared fish populations. *PLoS One*. 2013;8(4):e62011. DOI 10.1371/journal.pone.0062011.

Olstad K., Bachmann L., Bakke T.A. Phenotypic plasticity of taxonomic and diagnostic structures in gyrodactylosis-causing flatworms (Monogenea, Platyhelminthes). *Parasitology*. 2009;136:1305-1315. DOI 10.1017/S0031182009990680.

Pakdee W., Ogawa K., Pornruseetriratn S., Thaenkham U., Yeemin T. The first record of *Ligophorus* Euzet & Suriano, 1977 (Monogenea: Dactylogyridae) on *Crenimugil buchanani* (Teleostei: Muglidae) from Thailand based on morphological and molecular analyses. *J. Helminthol*. 2019;93(6):752-762. DOI 10.1017/S0022149X1800072X.

Poisot T., Verneau O., Desdevises Y. Morphological and molecular evolution are not linked in *Lamellodiscus* (Plathyhelminthes, Monogenea). *PLoS One*. 2011;6(10):e26252. DOI 10.1371/journal.pone.0026252.

Pugachev O.N., Gerasev P.I., Gussev A.V., Ergens R., Khotenowsky I. Guide to Monogenoidea of freshwater fish of Palaearctic and Amur regions. Milan: Ledizione-Ledi Publ., 2009.

Rodríguez-González A., Miguez-Lozano R., Llopis-Belenguer C., Balbuena J.A. Phenotypic plasticity in haptoral structures of *Ligophorus cephali* (Monogenea: Dactylogyridae) on the flathead mullet (*Mugil cephalus*): a geometric morphometric approach. *Int. J. Parasitol*. 2015;45(5):295-303. DOI 10.1016/j.ijpara.2015.01.005.

Rubio-Godoy M. Fish host-monogenean parasite interactions, with special reference to Polyopisthocotylea. In: Terrazas L.I. (Ed.) Advances in the Immunobiology of Parasitic Diseases. Thiruvananthapuram: Research Signpost, 2007;91-109.

Soo O.Y.M., Tan W.B., Lim L.H.S. Three new species of *Ligophorus* Euzet & Suriano, 1977 (Monogenea: Ancyrocephalidae) from *Moolgarda buchanani* (Bleeker) off Johor, Malaysia based on morphological, morphometric and molecular data. *Raffles Bull. Zool*. 2015;63:49-65.

Strona G., Montano S., Seveso D., Paolo G., Fattorini S. Identification of Monogenea made easier: a new statistical procedure for an auto-

matic selection of diagnostic linear measurements in closely related species. *J. Zoolog. Syst. Evol. Res*. 2014;52(2):95-99. DOI 10.1111/jzs.12050.

Tokarev Yu.S., Vasilieva A.A., Grushevaya I.V., Malysh Yu.M. Multilocus genotyping as a modern approach to the diagnosis of microsporidia, obligate intracellular parasites of animals. *Problemy Sovremennoi Nauki i Obrazovaniya = Problems of Modern Science and Education*. 2015;12:51-55. (in Russian)

Vanhove M.P.M., Tessens B., Schoelinck C., Jondelius U., Littlewood D.T.J., Artois T., Huyse T. Problematic barcoding in flatworms: a case-study on monogeneans and rhabdocoels (Platyhelminthes). *ZooKeys*. 2013;365:355-379. DOI 10.3897/zookeys.365.5776.

Vignon M. Putting in shape – towards a unified approach for the taxonomic description of monogenean haptoral hard parts. *Syst. Parasitol.* 2011;79(3):161-174. DOI 10.1007/s11230-011-9303-1.

Yamaguti S. Systema helminthum. Vol. IV. Monogenea and Aspidocotylea. New York; London: Intersci. Publ., 1963.

ORCID ID

E.A. Vodiasova orcid.org/0000-0003-3886-2880
E.S. Chelebieva orcid.org/0000-0002-7662-2573
O.V. Shikhat orcid.org/0000-0003-1909-3836
D.M. Atopkin orcid.org/0000-0001-8417-3424
E.V. Dmitrieva orcid.org/0000-0001-6300-2458

Acknowledgements. The work was supported financially by RFBR under the scientific project No. 20-44-920004 and according to the State Assignment No. 0556-2021-0002, state registration number 121030100028-0.

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

Received July 13, 2021. Revised January 6, 2022. Accepted January 18, 2022.

Вавиловский журнал генетики и селекции. 2022;26(3):298-307
DOI 10.18699/VJGB-22-28

Original Russian text www.bionet.nsc.ru/vogis/

Comparative analysis of allele frequencies for DNA polymorphisms associated with disease and economically important traits in the genomes of Russian and foreign cattle breeds

A.V. Igoshin[1], G.A. Romashov[1], E.N. Chernyaeva[2], N.P. Elatkin[2], N.S. Yudin[1], D.M. Larkin[3]

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] LLC "Miratorg-Genetika", Moscow, Russia
[3] Royal Veterinary College, London, United Kingdom
dmlarkin@gmail.com

Abstract. The genetic makeup of a breed including its genetic differences from other breeds determines its appearance and characteristics, including economically important traits and resistance to pathologies. To date, many loci controlling significant phenotypes have been identified, which is successfully used in the world practice of marker-assisted selection to improve breed properties. The aim of this study was a comparative analysis of frequencies for known causative nucleotide substitutions, insertions and deletions associated with disease and economically important traits in Russian and foreign cattle breeds. As a result, we identified frequencies of these DNA polymorphisms in the populations of Russian cattle breeds, compared them with those of foreign populations of the same breed, as well as other foreign breeds. Our results indicate similarities in frequencies for most of such alleles within breeds (populations of Russian and foreign breeding), as well as the relationship between the causative allele prevalence and the presence of phenotypic traits under the effect. We also found an excess of some undesirable alleles in the Russian cattle populations, which should be paid attention to when designing breeding programs. We found that the alleles increasing fertility in the Hereford breed have a higher frequency in the Russian Hereford population compared to the foreign counterpart. Interestingly, unlike for the European breeds, for Asian Turano-Mongolian Wagyu and Yakut cattle, there was a less clear link between phenotypic traits and frequencies of known causative alleles. Our work points to specific genetic variants that could be used to improve and/or maintain the performance of certain cattle breeds bred in the Russian Federation.
Key words: cattle; selection; breed; Russian Federation; genetic variants; SNP, insertion; deletion.

For citation: Igoshin A.V., Romashov G.A., Chernyaeva E.N., Elatkin N.P., Yudin N.S., Larkin D.M. Comparative analysis of allele frequencies for DNA polymorphisms associated with disease and economically important traits in the genomes of Russian and foreign cattle breeds. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2022;26(3):298-307. DOI 10.18699/VJGB-22-28

Сравнительный анализ частот ДНК-полиморфизмов, ассоциированных с заболеваниями и хозяйственно важными признаками, в геномах российских и зарубежных пород крупного рогатого скота

А.В. Игошин[1], Г.А. Ромашов[1], Е.Н. Черняева[2], Н.П. Елаткин[2], Н.С. Юдин[1], Д.М. Ларкин[3]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] ООО «Мираторг-Генетика», Москва, Россия
[3] Королевский ветеринарный колледж, Лондон, Великобритания
dmlarkin@gmail.com

Аннотация. Генетический состав породы и ее генетические отличия от других пород определяют ее облик и характерные особенности, включая экономически важные признаки и встречаемость патологий. К настоящему времени выявлено множество локусов, контролирующих наиболее значимые фенотипы, что успешно используется в мировой практике для маркер-ассоциированной селекции в целях улучшения свойств пород. В настоящей работе проведен сравнительный анализ частот известных каузативных нуклеотидных замен, вставок и делеций, связанных с заболеваниями и хозяйственно ценными признаками, в российских и зарубежных породах крупного рогатого скота. Выявлены частоты вышеуказанных ДНК-полиморфизмов в популяциях российских пород крупного рогатого скота, выполнено их сравнение с частотами в зарубежных популяциях для пород, разводимых в Российской Федерации, а также с другими зарубежными породами. Наши

результаты показывают схожесть частот большинства аллелей внутри пород (российского или зарубежного разведения), а также связь между представленностью аллелей исследуемых полиморфизмов и наличием определяемых ими фенотипических признаков. Были найдены и превышения по частотам ряда нежелательных аллелей в российских популяциях крупного рогатого скота, на которые стоит обратить внимание при селекционной работе с породами. Обнаружено, что аллели, отвечающие за повышенную фертильность породы герефорд, имеют повышенную частоту в популяциях российского разведения по сравнению с зарубежными популяциями. Интересно, что для азиатских турано-монгольских вагю и якутского скота наблюдалась меньшая связь между фенотипическими признаками и частотами известных каузативных аллелей по сравнению с европейскими породами. Наша работа указывает на конкретные генетические варианты, которые могут быть использованы для улучшения и/или поддержания качеств ряда пород крупного рогатого скота, разводимых в Российской Федерации.
Ключевые слова: крупный рогатый скот; селекция; порода; Российская Федерация; генетические варианты; SNP; инсерция; делеция.

Introduction

Common types of genetic variations, such as single nucleotide polymorphisms, nucleotide insertions and deletions, among others, can have "beneficial" or "harmful" effects on animal health and productivity (Liu, Bickhart, 2012; Bourque et al., 2018). That is why the sequencing of the *Bos taurus* genome caused a surge in research on the genetic diversity of cattle breeds and its relationship with economically important traits, adaptations and diseases, which opened up opportunities to use the knowledge gained for creating breeds with the necessary qualities and improving existing breeds (Larkin, Yudin, 2016; Yudin, Larkin, 2019). Now, according to the OMIA database (www.omia.org; Lenffer et al., 2006), 272 bovine traits are known to be genetically controlled, including a number of diseases. For 175 of them, causative mutations in the coding and non-coding regions of DNA have already been identified, the effect of which is related to various mechanisms, including changes in the protein sequence, in the stability, expression or processing of RNA (Ibeagha-Awemu et al., 2008; Yudin, Voevoda, 2015; Ciepłoch et al., 2017). Using this information, tests were developed for genotyping pathological mutations and removing carrier animals from the breeding herd (Romanenkova et al., 2015; Fornara et al., 2019; Sabetova et al., 2021). With this approach, it is possible to identify mutations at an early age for the timely culling of animals or embryos (Terletskiy et al., 2016). At the same time, it is worth considering that a "harmful" mutation may be "useful" for another economically important trait (Fasquelle et al., 2009). Identification of gene alleles associated with economically important traits allowed using them for marker-assisted selection (Pighetti, Elliott, 2011; Abd El-Hack et al., 2018). Marker-assisted selection is particularly important for traits that become evident with age or only in animals of the same sex, such as productivity or fertility (Zinovieva, 2016; Raina et al., 2020).

So far, Russian cattle breeds have been investigated for the presence of only a few, the most common mutations associated with economically important traits and health (Romanenkova et al., 2016, 2018; Usova et al., 2017; Surzhikova et al., 2019). The purpose of our work was to analyze the spectrum and frequencies of known causative DNA polymorphisms in nine Russian cattle breeds using genome sequencing data and to compare the frequencies of these polymorphisms with those in worldwide breeds or foreign populations of the same breeds to determine the options for which the selection in Russian cattle could be conducted.

Materials and methods

The list of single-nucleotide polymorphisms (SNPs), insertions and deletions, clinically and economically important for cattle, was compiled based on the information from the OMIA database (www.omia.org; Lenffer et al., 2006) and practical guidance of the Irish Cattle Breeding Federation (McClure M., McClure J., 2016). The genomic positions of polymorphisms specified in the *Bos taurus* UMD3.1 assembly coordinates were converted to the ARS-UCD1.2 assembly coordinates using liftOver (Kuhn et al., 2013). For polymorphisms present in the sample of Russian breeds, reference and alternative alleles were verified for matching those specified in the publications. For four substitutions of the twelve possible (T↔A and G↔C), such a verification is complicated, since: (1) there may be a change in the reference allele during the transition to a new genome assembly; (2) in the publication, the allele can be specified for a chain which is complementary to reference sequence. In such cases, we verified the alleles of polymorphisms in the context of codons (for substitutions in the coding sequence) or proximate sequences. For example, according to Hirano and colleagues (2013) and the OMIA database, replacing the nucleotide G with C at the BTA8:83909754 position, leading to the replacement of valine with leucine, results in perinatal weak calf syndrome. However, apparently, this replacement was indicated by the authors for the messenger RNA sequence – since in the assembly ARS-UCD1.2 C stands for the reference nucleotide, being a part of the "AAC" triplet, which, in turn, corresponds to "GUU" mRNA codon, encoding valine. Thus, in the reference assembly of ARS-UCD1.2, the G allele will be "harmful".

In this paper, we used data on SNPs, insertions and deletions in the worldwide breeds from the "1000 Bull Genomes" Project (Hayes, Daetwyler, 2019), including the resequencing data of eight Russian breeds obtained earlier, as well as the resequencing data (".fastq"-files) for the Russian population of the Aberdeen Angus breed (hereinafter simply Angus), provided by LLC "Miratorg-Genetika". Of note, some of these animals were imported from the USA and Australia

Table 1. Breed analyzed

Breed	Geographic origin	Sample size
Altai	Russia	20
Buryat	Russia	19
Kalmyk	Russia	13
Kholmogory	Russia	32
Yakut	Russia	30
Yaroslavl	Russia	22
Aberdeen Angus (foreign)	Australia, Canada, New Zealand, USA etc.	401
Aberdeen Angus (Russian)	Russia (partially imported from USA and Australia)	46
Hereford (foreign)	Australia, Canada, New Zealand, USA etc.	123
Hereford (Russian)	Russia	18
Wagyu (foreign)	Australia	9
Wagyu (Russian)	Russia	20
Northern Finncattle	Finland	34
Western Finncattle	Finland	25
Eastern Finncattle	Finland	25
The rest (>180 populations/breeds)	–	4409

(Table 1). Additionally, we also used data on three native Finnish breeds provided by the Natural Resources Institute Finland (Luke). Finland borders with Russia and has a largely similar (although milder) climate, so the inclusion of Finnish breeds in the study could shed light on features of the selective breeding manifesting in the close natural conditions of the two countries.

Removal of adapter sequences from raw paired reads was performed using Trimmomatic-0.39. Clean reads were aligned to the ARS-UCD1.2 reference sequence using BWA-MEM v.0.7.17 (Li, Durbin, 2009). Files containing aligned sequences (".sam"-files) were then converted to the ".bam" format and sorted using the SAMtools v.1.8 software (Li et al., 2009). Further, libraries belonging to the same animal were pooled using the 'MergeSamFiles' module of the Picard v.2.18.2 package (http://broadinstitute.github.io/picard). Duplicates were marked using the 'MarkDuplicates' module of the above-mentioned software. The OPTICAL_DUPLICATE_PIXEL_DISTANCE parameter equaling 2500 was chosen according to the recommendations of the "1000 Bull Genomes" protocol. Base quality score recalibration was performed using the 'BaseRecalibrator' and 'PrintReads' modules of the GATK v.3.8 package (McKenna et al., 2010) using data provided by the "1000 Bull Genomes" Project (Hayes, Daetwyler, 2019). The variant calling and the merging of the resulting gVCF files were performed using the 'HaplotypeCaller' and 'GenotypeGVCFs' modules of the GATK v.3.8 program, respectively.

Extraction of SNPs, insertions, and deletions from genome-wide VCF files was performed with the Tabix utility (Li, 2011), using the coordinates of polymorphisms from a previously generated list. The resulting VCF files containing the selected polymorphisms were used to calculate the frequencies of alternative alleles in the samples using the PLINK 2.0 program (Purcell et al., 2007) with the following parameters: --vcf --chr-set 30 --freq --pheno --loop-cats. The count has been carried out for (1) breeds bred in Russia (Kholmogory, Yaroslavl, Altai, Yakut, Buryat, Kalmyk, Angus, Wagyu and Hereford), (2) foreign populations of those breeds (if present), (3) three Finnish breeds (Northern Finncattle, Western Finncattle and Eastern Finncattle), and (4) a combined sample of all other worldwide cattle breeds (see Table 1).

The presence of allele frequency differences between the abovementioned samples was tested using Fisher's exact test implemented in the 'fisher.test()' R function. Contingency tables 2×2 were composed by counting the number of reference and alternative alleles in the chromosomal pool of each of the two groups studied. Three types of comparisons were made: (1) between a breed bred in Russia (or a foreign population of the same breed, if present) and a combined sample of other world breeds; (2) between a breed bred in Finland and the combined sample of the world's breeds using the polymorphisms identified in the first type of comparisons; (3) only between the Russian population and the foreign population of the same breed. To correct for multiple comparisons, we used the Storey and Tibshirani method

(Storey, Tibshirani, 2003) implemented in the 'qvalue()' R function (Storey et al., 2020).

Results

Our list of clinically and economically important polymorphisms contained 193 SNPs and 63 insertions/deletions. A search in the VCF files revealed in Russian breeds the presence of 38 SNPs and one insertion from the above-mentioned list (Supplementary Table 1)[1], which corresponded to at least 21 phenotypic traits.

When comparing 15 populations for 39 polymorphisms (585 comparisons in total) with a global sample, in 229 cases statistically significant ($q < 0.05$) differences in allele frequencies were found (see the Figure). The most significant differences with the total sample of worldwide breeds were observed for foreign populations of Angus and Hereford breeds (29 and 27 loci, respectively). Of the Russian populations, the Yakut breed had the largest number (16 loci) of differences from the worldwide sample. Of the Finnish breeds, the Northern Finncattle had the largest number (20 loci) of such differences.

The most significant (q = 4.24E–286) allele frequency difference from the global sample was observed for the foreign Angus population for SNP rs109688013 in the melanocortin-1 receptor gene *MC1R*, carriers of the alternative allele C of which have a black coat color (Klungland et al., 1995). The difference from the worldwide sample for this locus was also statistically significant for most of other populations as well, with the exception of the Northern Finncattle, as well as Russian and foreign Wagyu populations. In particular, the difference at this SNP was the highest among 39 loci for the Russian population of Angus (q = 6.01E–35), both populations of Herefords (q = 6.22E–37 for foreign and 7.34E–07 for Russian), for Altai (q = 1.99E–06), Kholmogory (q = 9.27E–12) and Yaroslavl (q = 2.76E–06) breeds. In foreign and Russian Angus populations, the frequency of the C allele coding for black color reaches 0.973 and 0.989, while in other worldwide breeds it is 0.339. In the populations of Altai, Kholmogory, Yaroslavl breeds, Russian and foreign Herefords, it has a frequency of 0.026, 0.828, 0.772, 0 and 0.019, respectively. In Finnish breeds, the frequency of the C allele varies from zero in Western Finncattle to 0.052 in Eastern Finncattle and 0.258 in Northern Finncattle.

Of the remaining loci, the greatest difference in the studied breeds from the global cattle population was observed for polymorphisms associated with milk traits, coat color and bleeding disorders. Thus, the Russian Wagyu population had the most significant (q = 6.44E–21) allele frequency difference from the worldwide sample for 15 bp insertion located at BTA27:16305660, which disrupts the *F11* gene function and, as a result, leads to a deficiency of blood coagulation factor XI, encoded by this gene (Kunieda et al., 2005). In Russian Wagyu population, the frequency of this insertion reaches 0.25, while in the global cattle population it is close to zero. The most significant differences from the worldwide sample for the foreign Wagyu population (q = 2.60E–05) and the Yakut breed (q = 2.21E–18) were observed for SNP rs210634530 in the gene of microphthalmia-associated transcription factor MITF, which is associated with the 'white spotting' phenotype (Fontanesi et al., 2012). The frequencies of the 'white spotting'-associated allele T in the Yakut breed and foreign Wagyu population are 0.083 and 0.111, respectively, while in the worldwide sample it reaches 0.65. In the Buryat and Kalmyk breeds, the most significant difference (q = 6.81E–10 and 2.33E–06, respectively) had SNP rs109191047 in the growth hormone gene *GH1*, associated with the composition of milk (Mullen et al., 2010). The frequency of G allele increasing the milk fat and protein content is 0.100 in the worldwide population, while in the above-mentioned breeds it reaches 0.526 and 0.500, respectively.

Comparisons between the Russian and corresponding foreign populations, made for the Angus, Hereford and Wagyu breeds, revealed four loci, statistically significantly ($q < 0.05$) differing in allele frequencies. Of these, three SNPs (rs43703017, rs43703015 and rs110014544) had differing frequencies in the Russian and foreign Angus populations and specified the alleles of the kappa-casein gene *CSN3*. One SNP located in the *CAPN1* gene (rs17872050) differed between the Hereford populations and was associated with meat tenderness. Taking into account the frequency differences at the nominal significance level ($p < 0.05$), eight additional loci can be noted (Table 2), among which the V311A missense substitution (BTA26:34340886T>C) in the *NHLRC2* gene differing between the Angus populations and in homozygotes leading to notomelia, a type of polymelia in which the additional limb is located along or near the midline of the back (Beever et al., 2014).

Discussion

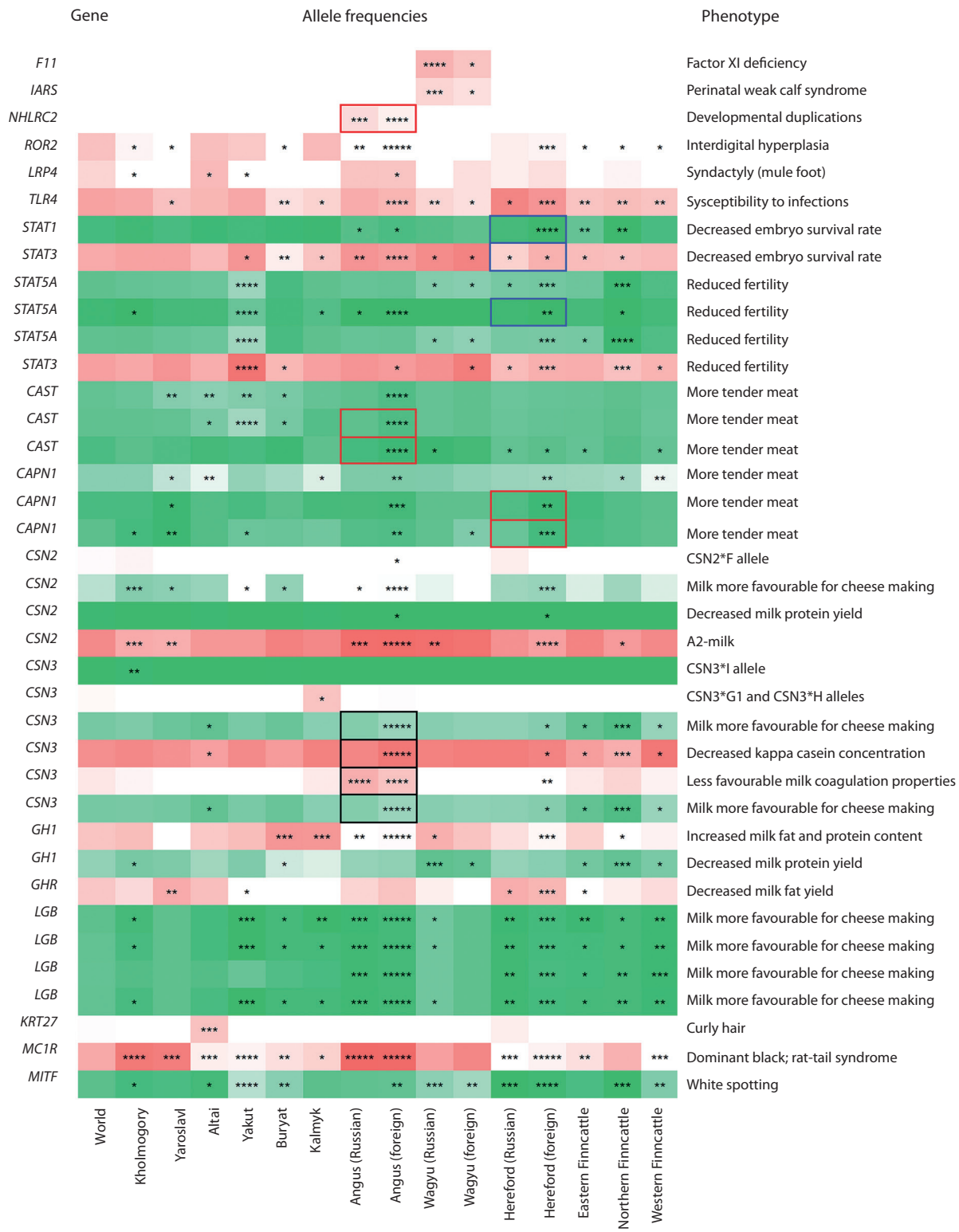
Breed-specific genetic features

The gene pool of farm animals is formed under the influence of factors such as selection for productive traits, adaptation to environmental conditions, hybridization, *de novo* mutations, the founder effect and genetic drift (Notter, 1999; Xu et al., 2015).

As we showed above, a significant part of the polymorphisms taken into the analysis in the studied breeds differs in frequencies from the "worldwide average", reflecting the gene pool features of particular populations. For example, the Yakut cattle shows the highest divergence in allele frequencies among Russian breeds, expressed both in a greater number of differing loci and in a greater significance of these differences, which is consistent with the data on phylogeny of this breed and the analysis of its population structure (Yurchenko et al., 2018; Buggiotti et al., 2021).

Some of the polymorphisms studied make a definitive contribution to characteristic features of the breeds. For example, the content of the *MC1R* gene allele rs109688013-C in the breeds coincides well with the typical color of their representatives. Thus, in Angus having a black coat color, the frequency of this allele is close to one. In Yaroslavl and Kholmogory cattle, rs109688013-C also predominates, apparently defining black and black-mottled coats, mainly characteristic of these animals. At the same time, in Herefords, which are not characterized by a black color, the frequency of the C allele is close to zero. Similarly, there is a link between color and the frequency of the C allele in populations of Finnish breeds. In breeds that have mainly lighter coats (fawn,

[1] Supplementary Tables 1 and 2 are available in the online version of the paper: http://vavilov.elpub.ru/jour/manager/files/Suppl_Igoshin_engl.pdf



Frequencies of clinically and economically significant polymorphisms in Russian and foreign cattle populations.

On the left are the genes containing the polymorphic variants under study. A darker tone corresponds to a higher frequency of a reference (green) or alternative (red) allele. On the right is the phenotype associated with this allele. The polymorphism designations and their frequencies are given in Suppl. Table 2.
The asterisks indicate loci that have significant frequency differences between the specified breed and the global cattle population: ***** q < 1.0E–25, **** q < 1.0E–10, *** q < 1.0E–5, ** q < 1.0E–3, * q < 0.05. The frame indicates loci that differ ($p < 0.05$) between the Russian and foreign populations of the same breed: red color means that the "harmful" allele has a large content in the Russian population, blue – in the foreign, black – the significance of the allele for beef breeds is not established.

Table 2. Differences between Russian and foreign populations within the same breed

Locus	Allele		Frequency of allele associated with the phenotype specified		Breed	Phenotype	Statistical significance	
	Reference	Alternative	Russian population	Foreign population			*p*-value	q-value
rs43705173	G*	A	0.853	0.971	Hereford	Decreased embryo survival rate	0.0086	0.1465
rs43703015	T*	C	0.304	0.157	Angus	Decreased rennet coagulation time, decreased lactose concentration	0.0011	0.0353
rs43703016	C	A*	0.696	0.833	Angus	Decreased kappa casein concentration	0.0024	0.0563
rs43703017	A	G*	0.283	0.113	Angus	Less favourable coagulation properties and increased milk fat content	4.88E–05	0.0057
rs110014544	G*	A	0.304	0.159	Angus	Decreased rennet coagulation time	0.0012	0.0353
rs41255587	G*	A	0.620	0.725	Angus	More tender meat	0.0385	0.3753
rs109221039	A*	G	0.793	0.894	Angus	More tender meat	0.0088	0.1465
rs208753173	G*	A	0.917	0.988	Hereford	Reduced fertility	0.0298	0.349
rs110942700	T	C*	0.083	0.267	Hereford	Decreased embryo survival rate	0.0201	0.262
BTA26:34340886	T	C*	0.065	0.024	Angus	Developmental duplications	0.037	0.3753
rs17871051	G*	A	0.722	0.894	Hereford	More tender meat	0.0127	0.1854
rs17872050	C*	T	0.500	0.799	Hereford	More tender meat	0.0003	0.0154

* The allele associated with the phenotype specified.

light brown and red, often white muzzle, belly and back), it is low (0.053 in Eastern Finncattle) or zero (in Western Finncattle). In Northern Finncattle, which has a predominantly white coat color (some individuals are black-mottled), the frequency of rs109688013-C is 0.258. Breeds for which the red (Kalmyk) or brown (Altai, Buryat) colors are typical have rs109688013-C in low frequency (0.03–0.08). However, in Wagyu populations, which are usually characterized by black color, the frequency of this allele is far from one and has values of 0.42 in Russian population and 0.67 in foreign population, probably reflecting the genetic characteristics of Turano-Mongolian breeds. This discrepancy is also observed in Yakut cattle, in which a black-and-white color is common, but the frequency of rs109688013-C is vanishingly small. Given the genetic divergence of Turano-Mongolian breeds from other breeds, it can be assumed that other loci are involved in the control of body coloration.

Also, coat color is associated with the SNP rs210634530 in the *MITF* gene, the T allele of which defines the 'white spotting' phenotype. The highest frequency of rs210634530-T is observed in populations of Hereford cattle (fixed in the Russian sample and 0.92 in the foreign population), which is characterized by a white head and belly. In addition, this allele prevails in the populations of Kholmogory, Yaroslavl, Altai and Kalmyk breeds, which have white spotting in color, as well as in Angus. In other populations, the frequency of the T allele varies from low (Yakut breed) to moderate (Buryat, Wagyu). The link between the content of rs210634530-T and coat color can be demonstrated for Finnish breeds. As mentioned above, many individuals of Western Finncattle and Eastern Finncattle have a white muzzle, back, and belly. Northern Finncattle has either a white or, less often, black-and-white coat. It should be noted that in addition to SNP rs210634530, additional loci appear to be involved in the control of the 'white spotting' phenotype (Fontanesi et al., 2012), therefore, the link between the frequency of rs210634530-T and coat color may not be so straightforward.

Some of the genetic features of the breeds are not quite obvious at first glance. For example, both Russian and foreign populations of Angus and Wagyu have a high (0.89–0.95) frequency of rs43703011-G allele of the beta-casein gene *CSN2*. Variations of the *CSN2* gene at several non-synonymous

positions determine its alleles – A1, A2, A3, B, C, etc. The above-mentioned allele G of rs43703011 is shared by several alleles of the *CNS2* gene, the most common of which is A2. The so-called A2-milk is considered more preferable for consumption, due to better absorption and fewer undesirable effects from the human digestive system (Jianqin et al., 2016). In recent years, breeding programs in many countries have aimed to increase the frequency of the A2 allele in dairy cattle (Sebastiani et al., 2020). Given that Angus and Wagyu are beef breeds and are not used for milk production, the increased G allele content they have can hardly be explained by selection to improve milk quality. The most plausible explanation is selection for meat productivity. Thus, according to Hohmann et al. (2020), the carriage of the A2 allele increases average daily weight gain and weaning weight in German Angus and Simmentals. Therefore, increasing the frequency of the rs43703011-G allele, and consequently the A2 allele of the *CSN2* gene, can be useful for improving not only dairy but also beef breeds.

Some of the variants found are specific to one breed and virtually absent in others. The most breed-specific are clinically important polymorphisms in the *F11*, *IARS* and *NHLRC2* genes. The previously mentioned insertion in the *F11* gene, leading to a deficiency of blood coagulation factor XI, is almost exclusively observed in foreign and Russian Wagyu populations. At the same time, among more than 5 thousand other animals from the "1000 Bull Genomes" Project, this mutation is harbored by only two animals. Association of factor XI activity with ATATGTGCAGAATAT insertion has been initially demonstrated for Wagyu (Kunieda et al., 2005). The homozygous genotype for this mutation is associated with a blood clotting disorder and an increase in the duration of bleeding. In the Russian population of Wagyu, its frequency equals 0.25, which is consistent with the data of early publications on its prevalence in the Japanese black breed (Watanabe et al., 2006; Ohba et al., 2008). At the same time, in the foreign Wagyu population, here represented by a sample from Australia, this insertion has a frequency of 0.11.

Other examples of breed-specific variants are single nucleotide substitutions in the *IARS* (BTA8:83909754C>G) and *NHLRC2* (BTA26:34340886T>C) genes. BTA8:83909754C>G mutation in the *IARS* gene in homozygote leads to perinatal weak calf syndrome and increased prenatal mortality (Hirano et al., 2013, 2016). This variant is specific to Wagyu, and besides, it was found only in one animal from the "1000 Bull Genomes" Project. In the Russian and Australian samples of this breed, its frequencies are 0.075 and 0.056, respectively. The BTA26:34340886T>C mutation in the *NHLRC2* gene mentioned earlier, in homozygote leading to notomelia, is breed-specific for Angus, and was first discovered in this breed (Beever et al., 2014). Besides Angus, in the sample of "1000 Bull Genomes", the mutant allele is found only in one animal of an unknown ('crossbreed') breed. In the Russian and foreign populations of this breed, it has frequencies of 0.065 and 0.024, respectively.

Differences between Russian and foreign populations of the same breed

The presence in our analysis of foreign Angus, Herefords and Wagyu populations can shed light on the features of the selection and adaptation of Russian populations of these breeds. Overall, Russian and foreign samples of the same breed demonstrate similar allele frequency profiles, with statistically confirmed differences present only in a small number of loci. The differences observed can be explained by many factors or their combinations. For example, an almost threefold excess of the BTA26:34340886-C allele (leading to the appearance of additional limbs) content in the Russian Angus population compared to the foreign one (see Table 2) may be a consequence of the founder effect or genetic drift in general, as well as less intensive efforts for elimination of this variant in the Russian herd.

Interpopulation differences in the loci associated with reproduction may result from an adaptation to environmental conditions. In the Russian population of Herefords, alleles of several polymorphisms that negatively affect the survival of embryos (rs43705173-G and rs110942700-C) and fertility (rs208753173-G) have a lower frequency compared to the foreign sample of this breed. It can be assumed that the Russian sample of Herefords, in this work represented by a population bred in Western Siberia since the 1960s (Vsyakikh, Kurinsky, 1976), was subject to selection for reproductive performance. This assumption is supported by the data of Afanasyeva and co-authors, according to which in the conditions of the Altai Territory, the population of Herefords of Siberian selection shows a much lower stillbirth rate (1.4 %) compared to animals of Finnish selection (6.6 %) imported in 2011 (Afanasyeva et al., 2015). Low temperatures are known to negatively affect the reproduction of cattle, reducing fertility and increasing perinatal mortality (Gwazdauskas, 1985; Mee, 2020). Therefore, population differences at these loci may reflect the process of genetic adaptation aimed at compensating a decrease in reproductive functions caused by cold.

Of particular interest are single nucleotide polymorphisms associated with meat traits and differing in samples of Angus (rs41255587 and rs109221039 in the *CAST* gene) and Herefords (rs17871051 and rs17872050 in the *CAPN1* gene). For all four SNPs, the foreign populations of these breeds demonstrate a higher content of alleles increasing meat tenderness. This is an important gastronomic feature and its improvement is included in the breeding programs of foreign beef breeds (Tatum, 2006). At the same time, we are not aware of extensive breeding attempts of this kind in Russia, which is probably the reason for the observed differences between the samples. Therefore, the Russian populations of Angus and Herefords have the potential for the improvement of meat quality by selection for *CAST* and *CAPN1* alleles.

Of the studied loci differing between populations of the same breed, four SNPs (rs43703015, rs43703016, rs43703017 and rs110014544) determining the kappa-casein gene *CSN3* alleles deserve to be noticed. Their allele frequencies differ between Russian and foreign Angus populations. These polymorphisms are associated with milk traits, in particular, with the concentration of kappa-casein in milk and milk coagulation properties, which is important for cheesemaking. At the same time, the effect of *CSN3* alleles on the productivity of beef cattle is poorly understood. Investigations of Tambasco et al. (2003) и Curi et al. (2005) found no association between *CSN3* alleles and meat traits. Thus, the observed differences can be attributed to the founder effect, or selection for

economically important traits whose associations with *CSN3* polymorphisms have not yet been identified.

Polymorphisms of clinical significance present in Russian breeds

In Russian breeds, there is a number of polymorphic variants, in homozygous state causing hereditary diseases, some of which (mutations in the genes *F11*, *IARS* and *NHLRC2*) have already been discussed above due to their breed specificity. Also, the variants in the *ROR2* and *LRP4* genes should be mentioned that are associated with the manifestation of interdigital hyperplasia (proliferation of tissue between the hooves) and syndactyly (fusion of the fingers, also called 'mule foot'), respectively. Unlike the *F11*, *IARS* and *NHLRC2* genes, the "harmful alleles" in *ROR2* (rs377953295-A) and *LRP4* (rs453049317-T) are not breed-specific, and are widespread both in Russian breeds and in the rest of the worldwide cattle population. Of the Russian populations, the Kalmyk (0.192) and Altai (0.15) breeds have the highest content of the rs37795322-A allele of the *ROR2* gene. In the worldwide sample, its frequency reaches 0.13. The rs453049317-T variant in the *LRP4* gene has the highest frequency in the Altai breed (0.2) and in the Russian Angus (0.12), while in the rest of the worldwide population it is 0.076.

Currently, testing for genetic defects is widely used in the practice of animal husbandry in many countries (Terletskiy et al., 2016). For example, testing for mutations in the *F11* and *IARS* genes is included in the genetic screening programs recommended by the Australian Wagyu Association (https://www.wagyu.org.au/content/uploads/2020/08/Genetic-Conditions-in-Wagyu-FactSheet-2020.pdf). At the same time, the elimination of undesirable alleles should be approached with caution. For example, there is an assumption that the carriage of mutations associated with syndactyly improves the milk productivity of cows, which can partially explain the spread of this pathology in cattle (Johnson et al., 2006).

Conclusion

Our analysis showed the allele frequency distribution for the most clinically and economically important DNA polymorphisms present in Russian cattle breeds. A number of variants leading to common hereditary disorders in cattle have significant representation in Russian populations, and probably need to be eliminated. Also, the differences between Russian and foreign cattle populations at several loci are presumably of adaptive importance. The data of this study may be useful in cattle breeding programs aimed at improving the existing cattle breeds, and creating new ones.

References

Abd El-Hack M.E., Abdelnour S.A., Swelum A.A., Arif M. The application of gene marker-assisted selection and proteomics for the best meat quality criteria and body measurements in Qinchuan cattle breed. *Mol. Biol. Rep*. 2018;45(5):1445-1456. DOI 10.1007/s11033-018-4211-y.

Afanasyeva A.I., Knyazev S.S., Lotz K.N. Reproductive capacity of Hereford beef cattle of Siberian and Finnish breeding under the conditions of the Altai region. *Vestnik Altayskogo Gosudarstvennogo Agrarnogo Universiteta = Bulletin of the Altai State Agricultural University*. 2015;8(130):86-89. (in Russian)

Beever J.E., Marron B.M., Parnell P.F., Teseling C.F., Steffen D.J., Denholm L.J. Developmental Duplications (DD): 1. Elucidation of the underlying molecular genetic basis of polymelia phenotypes in Angus cattle. In: Proc. XXVIII World Buiatrics Congress. Cairns, 2014.

Bourque G., Burns K.H., Gehring M., Gorbunova V., Seluanov A., Hammell M., Imbeault M., Izsvák Z., Levin H.L., Macfarlan T.S., Mager D.L., Feschotte C. Ten things you should know about transposable elements. *Genome Biol*. 2018;19:199. DOI 10.1186/s13059-018-1577-z.

Buggiotti L., Yurchenko A.A., Yudin N.S., Vander Jagt C.J., Vorobieva N.V., Kusliy M.A., Vasiliev S.K., Rodionov A.N., Boronetskaya O.I., Zinovieva N.A., Graphodatsky A.S., Daetwyler H.D., Larkin D.M. Demographic history, adaptation, and NRAP convergent evolution at amino acid residue 100 in the world northernmost cattle from Siberia. *Mol. Biol. Evol*. 2021;38(8):3093-3110. DOI 10.1093/molbev/msab078.

Ciepłoch A., Rutkowska K., Oprzkadek J., Poławska E. Genetic disorders in beef cattle: a review. *Genes Genomics*. 2017;39(5):461-471. DOI 10.1007/s13258-017-0525-8.

Curi R.A., de Oliveira H.N., Gimenes M.A., Silveira A.C., Lopes C.R. Effects of CSN3 and LGB gene polymorphisms on production traits in beef cattle. *Genet. Mol. Biol*. 2005;28(2):262-266. DOI 10.1590/S1415-47572005000200015.

Fasquelle C., Sartelet A., Li W., Dive M., Tamma N., Michaux C., Druet T., Huijbers I.J., Isacke C.M., Coppieters W., Georges M., Charlier C. Balancing selection of a frame-shift mutation in the *MRC2* gene accounts for the outbreak of the Crooked Tail Syndrome in Belgian Blue Cattle. *PLoS Genet*. 2009;5(9):e1000666. DOI 10.1371/journal.pgen.1000666.

Fontanesi L., Scotti E., Russo V. Haplotype variability in the bovine *MITF* gene and association with piebaldism in Holstein and Simmental cattle breeds. *Anim. Genet*. 2012;43(3):250-256. DOI 10.1111/j.1365-2052.2011.02242.x.

Fornara M.S., Kostyunina O.V., Filipchenko A.A., Sermyagin A.A., Zinovyeva N.A. Polymorphism determination system of gene SUGT1 associated with Fleckvieh fertility haplotype FH4. *Veterinariya, Zootekhniya i Biotekhnologiya = Veterinary Medicine, Zootechnics and Biotechnology*. 2019;3:92-97. DOI 10.26155/vet.zoo.bio.201903015. (in Russian)

Gwazdauskas F.C. Effects of climate on reproduction in cattle. *J. Dairy Sci*. 1985;68(6):1568-1578. DOI 10.3168/jds.S0022-0302(85)80995-4.

Hayes B.J., Daetwyler H.D. 1000 Bull Genomes Project to map simple and complex genetic traits in cattle: applications and outcomes. *Annu. Rev. Anim. Biosci*. 2019;7:89-102. DOI 10.1146/annurev-animal-020518-115024.

Hirano T., Kobayashi N., Matsuhashi T., Watanabe D., Watanabe T., Takasuga A., Sugimoto M., Sugimoto Y. Mapping and exome sequencing identifies a mutation in the *IARS* gene as the cause of hereditary perinatal weak calf syndrome. *PLoS One*. 2013;8(5):e64036. DOI 10.1371/journal.pone.0064036.

Hirano T., Matsuhashi T., Takeda K., Hara H., Kobayashi N., Kita K., Sugimoto Y., Hanzawa K. *IARS* mutation causes prenatal death in Japanese Black cattle. *Anim. Sci. J*. 2016;87(9):1178-1181. DOI 10.1111/asj.12639.

Hohmann L.G., Weimann C., Scheper C., Erhardt G., König S. Associations between maternal milk protein genotypes with preweaning calf growth traits in beef cattle. *J. Anim. Sci*. 2020;98(10):skaa280. DOI 10.1093/jas/skaa280.

Ibeagha-Awemu E.M., Kgwatalala P., Ibeagha A.E., Zhao X. A critical analysis of disease-associated DNA polymorphisms in the genes of cattle, goat, sheep, and pig. *Mamm. Genome.* 2008;19(4):226-245. DOI 10.1007/s00335-008-9101-5.

Jianqin S., Leiming X., Lu X., Yelland G.W., Ni J., Clarke A.J. Effects of milk containing only A2 beta casein versus milk containing both A1 and A2 beta casein proteins on gastrointestinal physiology, symptoms of discomfort, and cognitive behavior of people with self-

reported intolerance to traditional cows' milk. *Nutr. J*. 2016;15:35. DOI 10.1186/s12937-016-0147-z.

Johnson E.B., Steffen D.J., Lynch K.W., Herz J. Defective splicing of *Megf7/Lrp4*, a regulator of distal limb development, in autosomal recessive mulefoot disease. *Genomics*. 2006;88(5):600-609. DOI 10.1016/j.ygeno.2006.08.005.

Klungland H., Våge D.I., Gomez-Raya L., Adalsteinsson S., Lien S. The role of melanocyte-stimulating hormone (MSH) receptor in bovine coat color determination. *Mamm. Genome*. 1995;6(9):636-639. DOI 10.1007/BF00352371.

Kuhn R.M., Haussler D., Kent W.J. The UCSC genome browser and associated tools. *Brief. Bioinform*. 2013;14(2):144-161. DOI 10.1093/bib/bbs038.

Kunieda M., Tsuji T., Abbasi A.R., Khalaj M., Ikeda M., Miyadera K., Ogawa H., Kunieda T. An insertion mutation of the bovine *F11* gene is responsible for factor XI deficiency in Japanese black cattle. *Mamm. Genome*. 2005;16(5):383-389. DOI 10.1007/s00335-004-2462-5.

Larkin D.M., Yudin N.S. The genomes and history of domestic animals. *Molecular Genetics, Microbiology and Virology*. 2016;31(4):197-202. DOI 10.3103/S0891416816040054.

Lenffer J., Nicholas F.W., Castle K., Rao A., Gregory S., Poidinger M., Mailman M.D., Ranganathan S. OMIA (Online Mendelian Inheritance in Animals): an enhanced platform and integration into the Entrez search interface at NCBI. *Nucleic Acids Res*. 2006;34(1):D599-D601. DOI 10.1093/nar/gkj152.

Li H. Tabix: fast retrieval of sequence features from generic TAB-delimited files. *Bioinformatics*. 2011;27(5):718-719. DOI 10.1093/bioinformatics/btq671.

Li H., Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760. DOI 10.1093/bioinformatics/btp324.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079. DOI 10.1093/bioinformatics/btp352.

Liu G.E., Bickhart D.M. Copy number variation in the cattle genome. *Funct. Integr. Genomics*. 2012;12(4):609-624. DOI 10.1007/s10142-012-0289-9.

McClure M., McClure J. Genetic Disease and Trait Information for IDB Genotyped Animals in Ireland. Bandon: Irish Cattle Breeding Federation, 2016.

McKenna A., Hanna M., Banks E., Sivachenko A., Cibulskis K., Kernytsky A., Garimella K., Altshuler D., Gabriel S., Daly M., DePristo M.A. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-1303. DOI 10.1101/gr.107524.110.

Mee J.F. Investigation of bovine abortion and stillbirth/perinatal mortality – similar diagnostic challenges, different approaches. *Ir. Vet. J.* 2020;73:20. DOI 10.1186/s13620-020-00172-0.

Mullen M.P., Berry D.P., Howard D.J., Diskin M.G., Lynch C.O., Berkowicz E.W., Magee D.A., MacHugh D.E., Waters S.M. Associations between novel single nucleotide polymorphisms in the *Bos taurus* growth hormone gene and performance traits in Holstein-Friesian dairy cattle. *J. Dairy Sci*. 2010;93(12):5959-5969. DOI 10.3168/jds.2010-3385.

Notter D.R. The importance of genetic diversity in livestock populations of the future. *J. Anim. Sci*. 1999;77(1):61-69. DOI 10.2527/1999.77161x.

Ohba Y., Takasu M., Nishii N., Takeda E., Maeda S., Kunieda T., Kitagawa H. Pedigree analysis of factor XI deficiency in Japanese black cattle. *J. Vet. Med. Sci.* 2008;70(3):297-299. DOI 10.1292/jvms.70.297.

Pighetti G.M., Elliott A.A. Gene polymorphisms: the keys for marker assisted selection and unraveling core regulatory pathways for mastitis resistance. *J. Mammary Gland Biol. Neoplasia*. 2011;16(4):421-432. DOI 10.1007/s10911-011-9238-9.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., Maller J., Sklar P., de Bakker P.I.W., Daly M.J., Sham P.C. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet*. 2007;81(3):559-575. DOI 10.1086/519795.

Raina V.S., Kour A., Chakravarty A.K., Vohra V. Marker-assisted selection vis-à-vis bull fertility: coming full circle – a review. *Mol. Biol. Rep*. 2020;47(11):9123-9133. DOI 10.1007/s11033-020-05919-0.

Romanenkova O.V., Gladyr E.A., Kostyunina O.V., Zinovieva N.A. Development of test system for diagnostics of cattle fertility haplotype HH3 associated with early embryonic mortality. *Dostizheniya Nauki i Tekhniki APK = Achievements of Science and Technology of AIC.* 2015;29(11):91-94. (in Russian)

Romanenkova O.V., Gladyr E.A., Kostyunina O.V., Zinovieva N.A. Screening of cattle for the presence of mutation in *APAF1* gene, which is associated with fertility haplotype HH1. *Dostizheniya Nauki i Tekhniki APK = Achievements of Science and Technology of AIC.* 2016;30(2):94-97. (in Russian)

Romanenkova O.S., Volkova V.V., Kostyunina O.V., Zinovieva N.A. Diagnostics of HH5 haplotype for Russian Holstein and Black-and-White cattle population. *Molochnoe i Myasnoe Skotovodstvo = Dairy and Beef Cattle Breeding*. 2018;6:13-15. DOI 10.25632/MMS.2018.2018.20295. (in Russian)

Sabetova K.D., Podrechneva I.Yu., Belokurov S.G., Schegolev P.O., Kofiadi I.A. Test system for *BLAD* mutation diagnosis in cattle populations. *Russ. J. Genet.* 2021;57(8):936-941. DOI 10.1134/S1022795421080135.

Sebastiani C., Arcangeli C., Ciullo M., Torricelli M., Cinti G., Fisichella S., Biagetti M. Frequencies evaluation of β-casein gene polymorphisms in dairy cows reared in Central Italy. *Animals* (*Basel*). 2020;10(2):252. DOI 10.3390/ani10020252.

Storey J.D., Bass A.J., Dabney A., Robinson D. qvalue: Q-value estimation for false discovery rate control. R Packag. version 2.24.0. 2020. DOI 10.18129/B9.bioc.qvalue.

Storey J.D., Tibshirani R. Statistical significance for genome-wide experiments. *Proc. Natl. Acad. Sci. USA*. 2003;100(16):9440-9445. DOI 10.1073/pnas.1530509100.

Surzhikova E.S., Sharko G.N., Mikhailenko T.N. Allelic spectrum of CSN3, PIT-1, PRL genes in horned cattle of Black-and-White breed. *Novosti Nauki v APK = Science News of AIC.* 2019;3:136-139. DOI 10.25930/2218-855X/032.3.12.2019. (in Russian)

Tambasco D.D., Paz C.C.P., Tambasco-Studart M., Pereira A.P., Alencar M.M., Freitas A.R., Coutinho L.L., Packer I.U., Regitano L.C.A. Candidate genes for growth traits in beef cattle crosses *Bos taurus* × *Bos indicus*. *J. Anim. Breed. Genet*. 2003;120(1):51-56. DOI 10.1046/j.1439-0388.2003.00371.x.

Tatum J.D. Pre-Harvest Cattle Management Practices for Enhancing Beef Tenderness. Colorado State Univ., 2006.

Terletskiy V.P., Buralkhiyev B.A., Usenbekov Y.S., Yelubayeva M., Tyshchenko V.I., Beyshova I.S. Screening for mutations that determine the development of hereditary diseases in breeding cattle. *Aktual'nye Voprosy Veterinarnoi Biologii = Actual Questions of Veterinary Biology*. 2016;3:3-7. (in Russian)

Usova T.P., Usmanova N.N., Litvina N.I., Usov N.V. The spread of BLAD-syndrome of breeding bulls of Holstein breed of Russian and import selection. *Vestnik Rossiyskogo Gosudarstvennogo Agrarnogo Zaochnogo Universiteta = Bulletin of the Russian State Agricultural Correspondence University*. 2017;25:20-24. (in Russian)

Vsyakikh A.S., Kurinsky M.S. Imported Cattle in the USSR. Moscow: Kolos Publ., 1976. (in Russian)

Watanabe D., Hirano T., Sugimoto Y., Ogata Y., Abe S., Ando T., Ohtsuka H., Kunieda T., Kawamura S. Carrier rate of Factor XI deficiency in stunted Japanese black cattle. *J. Vet. Med. Sci*. 2006;68(12):1251-1255. DOI 10.1292/jvms.68.1251.

Xu L., Bickhart D.M., Cole J.B., Schroeder S.G., Song J., Tassell C.P., Sonstegard T.S., Liu G.E. Genomic signatures reveal new evidences

for selection of important traits in domestic cattle. *Mol. Biol. Evol*. 2015;32(3):711-725. DOI 10.1093/molbev/msu333.

Yudin N.S., Larkin D.M. Whole genome studies of origin, selection and adaptation of the Russian cattle breeds. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2019;23(5):559-568. DOI 10.18699/VJ19.525. (in Russian)

Yudin N.S., Voevoda M.I. Molecular genetic markers of economically important traits in dairy cattle. *Russ. J. Genet.* 2015;51(5):506-517. DOI 10.1134/S1022795415050087.

Yurchenko A., Yudin N., Aitnazarov R., Plyusnina A., Brukhin V., Soloshenko V., Lhasaranov B., Popov R., Paronyan I.A., Plemyashov K.V., Larkin D.M. Genome-wide genotyping uncovers genetic profiles and history of the Russian cattle breeds. *Heredity* (*Edinb.*). 2018;120(2):125-137. DOI 10.1038/s41437-017-0024-3.

Zinovieva N.A. Haplotypes affecting fertility in Holstein cattle. *Sel'skokhozyaistvennaya Biologiya = Agricultural Biology.* 2016;51(4):423-435. DOI 10.15389/agrobiology.2016.4.423eng.

ORCID ID

N.S. Yudin orcid.org/0000-0002-1947-5554
D.M. Larkin orcid.org/0000-0001-7859-6201

Acknowledgements. The work was supported by the Russian Ministry of Education and Science grant No. 075-15-2021-1004.
Conflict of interest. The authors declare no conflict of interest.

Received November 30, 2021. Revised December 20, 2021. Accepted December 30, 2021.

Вавиловский журнал генетики и селекции. 2022;26(3):308-318
DOI 10.18699/VJGB-22-37

Original Russian text www.bionet.nsc.ru/vogis/

Molecular genetics of idiopathic pulmonary fibrosis

R.N. Mustafin

Bashkir State Medical University, Ufa, Russia
ruji79@mail.ru

Abstract. Idiopathic pulmonary fibrosis (IPF) is a severe progressive interstitial lung disease with a prevalence of 2 to 29 per 100,000 of the world's population. Aging is a significant risk factor for IPF, and the mechanisms of aging (telomere depletion, genomic instability, mitochondrial dysfunction, loss of proteostasis) are involved in the pathogenesis of IPF. The pathogenesis of IPF consists of TGF-β activation, epithelial-mesenchymal transition, and SIRT7 expression decrease. Genetic studies have shown a role of mutations and polymorphisms in mucin genes (*MUC5B*), in the genes responsible for the integrity of telomeres (*TERC, TERC, TINF2, DKC1, RTEL1, PARN*), in surfactant-related genes (*SFTPC, SFTPCA, SFTPA2, ABCA3, SP-A2*), immune system genes (*IL1RN, TOLLIP*), and haplotypes of HLA genes (*DRB1*15:01, DQB1*06:02*) in IPF pathogenesis. The investigation of the influence of reversible epigenetic factors on the development of the disease, which can be corrected by targeted therapy, shows promise. Among them, an association of a number of specific microRNAs and long noncoding RNAs was revealed with IPF. Therefore, dysregulation of transposons, which serve as key sources of noncoding RNA and affect mechanisms of aging, may serve as a driver for IPF development. This is due to the fact that pathological activation of transposons leads to violation of the regulation of genes, in the epigenetic control of which microRNA originating from these transposons are involved (due to the complementarity of nucleotide sequences). Analysis of the MDTE database (miRNAs derived from Transposable Elements) allowed the detection of 12 different miRNAs derived in evolution from transposons and associated with IPF (miR-31, miR-302, miR-326, miR-335, miR-340, miR-374, miR-487, miR-493, miR-495, miR-630, miR-708, miR-1343). We described the relationship of transposons with TGF-β, sirtuins and telomeres, dysfunction of which is involved in the pathogenesis of IPF. New data on IPF epigenetic mechanisms can become the basis for improving results of targeted therapy of the disease using noncoding RNAs.
Key words: idiopathic pulmonary fibrosis; immune system; microRNA; telomeres; transposons; epigenetic factors.

For citation: Mustafin R.N. Molecular genetics of idiopathic pulmonary fibrosis. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2022;26(3):308-318. DOI 10.18699/VJGB-22-37

Молекулярно-генетические особенности патогенеза идиопатического легочного фиброза

Р.Н. Мустафин

Башкирский государственный медицинский университет, Уфа, Россия
ruji79@mail.ru

Аннотация. Идиопатический легочный фиброз – тяжелая прогрессирующая интерстициальная болезнь легких с распространенностью 2–29 случаев на 100 000 человек населения в мире. Значимым фактором риска заболевания является старение, механизмы развития которого задействованы в патогенезе идиопатического легочного фиброза. К ним относятся истощение теломер, геномная нестабильность, дисфункция митохондрий и потеря протеостаза. Важную роль в развитии идиопатического легочного фиброза играют также эпителиально-мезенхимальный переход, активация TGF-β и снижение экспрессии сиртуина SIRT7. Молекулярно-генетические исследования показали, что в патогенезе идиопатического легочного фиброза имеют значение мутации и полиморфизмы в генах муцина (*MUC5B*), в генах, ответственных за целостность теломер (*TERC*, *TERT*, *TINF2*, *DKC1*, *RTEL1*, *PARN*), генов сурфактанта (*SFTPC*, *SFTPCA*, *SFTPA2*, *ABCA3*, *SP-A2*) и иммунной системы (*IL1RN*, *TOLLIP*), а также гаплотипы генов HLA (*DRB1*15:01*, *DQB1*06:02*). Перспективно изучение влияния на развитие болезни обратимых эпигенетических факторов, которые могут быть скорректированы таргетной терапией. Среди них с идиопатическим легочным фиброзом ассоциированы специфические микроРНК и длинные некодирующие РНК. Сделано предположение, что драйверным событием для идиопатического легочного фиброза служит дисрегуляция транспозонов, которые являются ключевыми источниками некодирующих РНК и влияют на механизмы старения. Это обусловлено тем, что при патологической активации транспозонов происходит нарушение регуляции генов, в эпигенетическом управлении которых участвуют происходящие от этих транспозонов микроРНК (в связи с комплементарностью нуклеотидных последовательностей). Анализ базы данных MDTE (miRNAs derived from Transposable Elements) позволил выявить 12 различных микроРНК, гены которых в эволюции возникли от транспозонов и ассоциированы

с идиопатическим легочным фиброзом (miR-31, miR-302, miR-326, miR-335, miR-340, miR-374, miR-487, miR-493, miR-495, miR-630, miR-708, miR-1343). Описаны взаимосвязи мобильных элементов с TGF-β, сиртуинами и теломерами, дисфункция которых вовлечена в патогенез идиопатического легочного фиброза. Новые данные об эпигенетических механизмах развития патологии могут стать основой для улучшения результатов таргетной терапии болезни с использованием в качестве мишени некодирующих РНК.
Ключевые слова: идиопатический легочный фиброз; иммунная система; микроРНК; теломеры; транспозоны; эпигенетические факторы.

Introduction

Idiopathic pulmonary fibrosis (IPF) is a progressive severe interstitial lung disease. The annual incidence of IPF is up to 17.4 per 100,000 people in the world (Chioma, Drake, 2017). The prevalence of IPF in different countries ranges from 2 to 29 per 100,000 people (Zhao et al., 2017) (for example, in Finland – 16–18 per 100,000 (Hodgson et al., 2002); in the USA – 14–42,7 per 100,000 people). IPF is associated with aging. Therefore, for people over 75 years of age, the prevalence of the disease is 227.2 per 100,000, while for people aged 18 to 34 years, the prevalence of IPF is 4 per 100,000. The average age of patients with IPF is 66 years (Raghu et al., 2006). Survival for IPF is about 3 years after diagnosis, and available drugs only slow the decline in lung function with little to no effect on mortality (Wyman et al., 2017).

IPF pathogenesis involves environmental influences and microorganisms (Sgalla et al., 2018). Viral (Epstein–Barr, cytomegalovirus, herpesvirus-1,-7,-8, Kaposi's sarcoma and hepatitis C), bacterial and fungal infections play a potential role in the development of IPF (Sheng et al., 2020). Smoking and metal dust inhalation are also associated with the risk of IPF (Chioma, Drake, 2017; Sgalla et al., 2018). The development of IPF is affected by occupational hazards, such as contact with silicon, beryllium, coal dust, asbestos, and radiation. In addition, IPF is associated with anti-inflammatory drugs (sulfasalazine, rituximab), chemotherapy drugs (bleomycin, methotrexate), heart drugs (amiodarone, propranolol), and antibiotics (nitrofurantoin, ethambutol) (Chioma, Drake, 2017). In 2019, a meta-analysis including 3206 patients and 9368 healthy individuals showed the role of gastroesophageal reflux disease in the development of IPF (Methot et al., 2019).

According to the generally accepted hypothesis, IPF develops as a result of immune reactions to restore the structure of lung tissue in case of repeated damage to the alveolar epithelium or endothelium. In this mechanism, the inflammatory mediator profibrotic cytokine – transforming growth factor β (TGF-β) activates angiogenesis and the production of extracellular matrix components (collagen and fibronectin). Failure to inactivate the fibrotic trigger leads to an exacerbation of the inflammatory response with excessive deposition of matrix components and lung scarring (Chioma, Drake, 2017). Molecular mediators of IPF include cell surface proteins, intracellular proteins, and soluble molecules (cytokines). The development of IPF is associated with sirtuins, a family of histone deacetylases that require NAD+ for their catalytic activity. The expression of sirtuins in fibroblasts of patients with IPF is significantly reduced. Similarly, a decrease in the concentration of SIRT7 in lung tissues was found in experimental mouse models with IPF induced by bleomycin. Inhibition of SIRT7 in fibroblast cultures by siRNA caused an increase in collagen synthesis. Overexpression of SIRT7 in lung fibroblasts leads to lower levels of COL1A1, COL1A2, COL3A1, exerting an antifibrotic effect (Wyman et al., 2017).

In the pathogenesis of IPF, an important role is played by the epithelial-mesenchymal transition, during which the expression of adhesion molecules (E-cadherin) is suppressed, and the cytokeratin cytoskeleton is transformed into a vimentin one. Accordingly, epithelial cells acquire a mesenchymal morphology (Li J. et al., 2021). However, there is still no complete theory that would fully explain the mechanism of IPF development. The most accurate data on the pathogenesis of IPF can be obtained using molecular genetic studies, which are promising for identifying the individual risk of the disease and developing its effective targeted therapy (Spagnolo, Cottin, 2017).

Genetic factors in idiopathic pulmonary fibrosis

Familial IPF involving two or more family members averages 10 to 15 % of all IPF cases (Chioma, Drake, 2017). There are sporadic, familial and syndromal forms of IPF (Lawson et al., 2004; Gochuico et al., 2012). Sporadic cases of the disease are multifactorial diseases, that is, their development is influenced by environmental factors. These forms comprise the majority of IPF cases and are associated with polymorphic variants of various genes (Table 1). Risk factors for sporadic IPFs are male gender, smoking, inhalation of metal and wood dust, or use of certain medications such as methotrexate and bleomycin (Fernandez et al., 2012). Familial IPFs are similar to sporadic, but are characterized by an earlier manifestation. They are caused by mutations in certain genes (see Table 1) (Lawson et al., 2004).

Familial IPFs were first described in 1958 by McKusick and Fisher as an autosomal dominant disorder with variable penetrance (McKusick, Fisher, 1958). Up to 18 % of all familial IPFs are caused by mutations in the genes of telomerase components: *TERT* (c.97C>T, c.430G>A, c.1456C>T, c.2240delT, c.2593C>T, c.2594G>A, c.3346_3522del) и *TERC* (r.37a>g) (Tsakiri et al., 2007). Exome sequencing also made it possible to identify rarer forms of familial IPF caused by mutations in the helicase gene that regulates telomere elongation (*RTEL1*: c.602delG, c.1451C>T, c.1940C>T, c.2005C>T, c.3371A>C) and in the deadenylation nuclease gene (*PARN*: IVS4-2a>g, c.529C>T, c.563_564insT, c.751delA, IVS16+1g>a, c.1262A>G) (Stuart et al., 2015). Cases of familial IPF caused by a mutation in exon 5 (+128T>A) in the SFTPC surfactant protein gene are also described (Thomas et al., 2002).

Syndromal IPF develops in autosomal recessive Hermansky–Pudlak syndrome, which is caused by an *AP3B1* gene

Table 1. Genetics of various forms of idiopathic pulmonary fibrosis

Gene/mutation (polymorphism)	Protein (RNA) product	Reference
Hereditary form		
RTEL1 / c.602delG, c.1451C > T, c.1940C > T, c.2005C > T, c.3371A > C	Telomere elongation regulating helicase	Stuart et al., 2015
PARN / IVS4-2a > g, c.529C > T, c.563_564insT, c.751delA, IVS16+1g > a, c.1262A > G	Deadenylating nuclease	
MUC5B / (rs35705950)	Mucin 5B	Seibold et al., 2011
TERT / c.97C > T, c.430G > A, c.1456C > T, c.2240delT, c. 2593C > T, c.2594G > A, c.3346_3522del	Telomeres reverse transcriptase	Tsakiri et al., 2007
TERC / r.37a > g	Telomeres RNA component	
TERT / c.1892G > A, c.2594G > A, c.2648T > G	Telomeres reverse transcriptase	Fernandez et al., 2012
SFTPC / экзон 5 (+128T > A)	Surfactant protein C	Thomas et al., 2002
Syndromal form		
AP3B1 / c.1525C > T (p.R509X), c.1739T > G (p.L580R), IVS10+5G > A, IVS11–1G > C	Intracellular traffic protein	Gochuico et al., 2012
Sporadic form		
AKAP13 / (rs62023891)	Lymphoblastic oncogene	Allen et al., 2020
ATP11A / (rs9577395)	ATPase phospholipid transporting 11A	
DPP9 / (rs12610495)	Serine protease	
DSP / (rs2076295)	Desmoplakin for intercellular contacts	
IVD / (rs59424629)	Isovaleryl-CoA-dehydrogenase	
IL1RN / (VNTR*2)	Interleukin	Korthagen et al., 2012
FAM13A / (rs2013701)	Protein involved in receptor signaling	Allen et al., 2020
MUC5B / (rs35705950)	Mucin 5B	Seibold et al., 2011; Noth et al., 2013; Lee M.G., Lee Y.H., 2015; Allen et al., 2020
SFTPC / (G4702C, C4859G, G4877A, G5089A, C5210A, G5236A, G5574A, A5786C, T6108C, C6699T)	Surfactant protein C	Lawson et al., 2004
SPPL2C / (rs17690703)	Lysosomal membrane protein	Noth et al., 2013
TERC / (rs12696304)	Telomeres RNA component	Allen et al., 2020
TERT / (rs7725218)	Telomeres reverse transcriptase	
TOLLIP / (rs111521887, rs5743894, rs5743890)	Innate immune system Toll-interacting protein	Noth et al., 2013

mutation (encodes an intracellular traffic protein). In this case, the specific mutations in the *AP3B1* gene are the following: c.1525C>T (p.R509X), c.1739T>G (p.L580R), IVS10+5G>A, IVS11–1G>C (Gochuico et al., 2012).

The promoter region of the mucin gene (*MUC5B*) contains a highly conserved polymorphic variant rs35705950 for primates, which is associated with sporadic and familial forms of IPF (Seibold et al., 2011). *SFTPC* gene polymorphisms (G4702C, C4859G, G4877A, G5089A, C5210A, G5236A, G5574A, A5786C, T6108C, C6699T) are associated with sporadic IPF (Lawson et al., 2004). In this form of IPF, shortening of the telomeres of circulating lymphocytes was revealed, which indicates the role of changes in the *TERT* and *TERC* genes (Fernandez et al., 2012). According to epidemiological data, familial forms with autosomal dominant inheritance range from 0.5–2 % (in the USA) (Allam, Limper, 2006) to 3.3–3.7 % (in Finland) (Hodgson et al., 2002) of all cases of IPF.

The most reliable data on the genes involved in the pathogenesis of IPF can be obtained from large-scale studies using genome-wide association studies (GWAS). A meta-analysis of five studies of IPF patients compared with healthy controls (the numbers of IPF patients in samples are 88, 61, 54, 22 and 77 from different countries) revealed the haploblock VNTR*2 of the *IL1RN* gene (encodes an interleukin-1 receptor antagonist), associated with susceptibility to the development of sporadic IPF (Korthagen et al., 2012). In a study of 544 patients with IPF, associations with various alleles of the *TOLLIP* gene (rs111521887, rs5743894, rs5743890), the *SPPL2C* gene allele (rs17690703), and the *MUC5B* gene allele (rs35705950) were found. The *TOLLIP* gene encodes a Toll-interacting protein involved in the innate immune system; the *SPPL2C* gene encodes a lysosomal membrane protein with a conserved transmembrane domain (Noth et al., 2013). The role of the *MUC5B* allelic variant (rs35705950) in the predisposition to IPF was confirmed in a meta-analysis of 2859 patients with IPF (control group consisted of 6901 people) (Lee M.G., Lee Y.H., 2015). The Tollip protein plays an important role in modulating the transport and degradation of TGF-β (Zhu L. et al., 2012). These results are consistent with the role of TGF-β in the pathogenesis of IPF (Chioma, Drake, 2017).

A GWAS conducted in 2016 on 1616 patients (control – 4683 people) showed the association of two haplotypes of the genes of the major histocompatibility complex (HLA): *DRB1*15:01* and *DQB1*06:02* with the development of IPF. It allowed researchers to suggest the role of autoimmune processes in the development of IPF (Fingerlin et al., 2016). A GWAS conducted in 2020 on DNA samples from 2668 patients showed an association of sporadic IPF with alleles of genes *MUC5B* (rs35705950), *TERC* (rs12696304), *TERT* (rs7725218), *DSP* (encodes desmoplakin for intercellular contacts, allele rs2076295), *ATP11A* (encodes a membrane ATPase that regulates calcium ions transport, rs9577395 variant), *IVD* (encodes isovaleryl-CoA dehydrogenase, rs59424629 polymorphism), *AKAP13* (encodes lymphoblastic oncogene, rs62023891 allele), *FAM13A* (hypoxia inducible gene associated with lung cancer, rs2013701 variant), *DPP9* (encodes a serine protease, polymorphism rs12610495) (Allen et al., 2020).

Thus, according to most genetic studies, IPF is associated with allelic variants of the genes responsible for the production of mucin, the functioning of telomeres and the immune system, which indicates a complex pathogenesis of the disease. In addition, IPF is associated with aging. At the molecular level, IPF development involves processes characteristic of aging, including telomere depletion, genomic instability, mitochondrial dysfunction, cellular senescence, and loss of proteostasis (Gulati, Thannickal, 2019). One of the causes of aging is the dysfunction of the immune system and telomeres caused by impaired transposon expression (Mustafin, 2019). This is due to the fact that in evolution, transposons became sources of the nucleotide sequence of both telomeres (Arkhipova et al., 2017) and telomerase encoding genes (Garavis et al., 2013). In Drosophila, the role of telomerase is performed directly by retrotransposons: TAHRE (Telomere Associated and HeT-A Related), TART (Telomere Associated Retrotransposon) и HeT-A (Healing Transposon) (Casacuberta, 2017). In humans, the ability of LINE1 retrotransposons to participate in alternative telomere elongation was revealed (Bondarev, Khavinson, 2016). Transposons likely play a role in the IPF pathogenesis, since familial IPF is most often caused by mutations in the genes maintaining telomeres (the *TERC* and *TERT* genes) (Tsakiri et al., 2007; Fernandez et al., 2012), while sporadic forms of IPF are associated with polymorphic variants of these genes (Allen et al., 2020).

Transposons serve as the basis for the epigenetic regulation of ontogenesis (Mustafin, Khusnutdinova, 2019). Transposons are specific genome structures capable of moving to a new locus and occupy 45 % of human DNA. They are classified into DNA transposons (movement by the "cut and paste" mechanism) and retrotransposons (movement with reverse transcription of mRNA and insertion of cDNA) (Wei G. et al., 2016).

Role of miRNAs in the pathogenesis of idiopathic pulmonary fibrosis

Epigenetic factors include DNA methylation, histone modifications and chromatin remodeling, as well as RNA interference via non-coding RNAs. Transposons are the most important sources of miRNA genes during evolution, in connection with which the MDTE (miRNAs derived from Transposable Elements) database was created in 2016 (Wei G. et al., 2016). Data from this database are taken from the results of the work of various authors (Piriyapongsa et al., 2007; Gu et al., 2009; Filshtein et al., 2012; Tempel et al., 2012; Qin et al., 2015). Investigation of miRNAs can provide information about IPF pathogenesis, as well as become the basis for the development of effective disease therapy. Lung fibroblasts play an important role in the initiation and progression of IPF. Investigation of microRNA expression in these cells revealed a decrease in miR-101 levels in human patients with IPF and in experimental models (bleomycin-induced pulmonary fibrosis) (Huang C. et al., 2017). In the development of IPF, dysregulation of various miRNAs that affect the TGF-β signaling pathways, which induce cell differentiation, migration, invasion, and hyperplastic changes, was revealed. These microRNAs include miR-21, miR-424 (profibrotic); miR-9-5p, miR-18a-5p, miR-26a, miR-27b, miR-101, miR-153, miR-326, miR-489, miR-1343 (antifibrotic) (Kang, 2017).

A pronounced imbalance in the expression of microRNA families miR-29, miR-21-5p, miR-92a-3p, miR-26a-5p, let-7d-5p in IPF was found, and therefore these molecules are considered as potential therapeutic targets for treatment of the disease (Bagnato et al., 2017). In human lung epithelium with IPF and mice with bleomycin-induced lung fibrosis, a decrease in the level of miR-323a was found, which attenuates TGF-α and TGF-β signaling (Ge et al., 2016). MiR-21 also influences these signaling pathways. The expression of miR-21 is increased in lung tissues of IPF patients and experimental mice. MiR-21 is produced by fibroblasts and regulates Smad7 expression by influencing TGF-β1, promoting extracellular matrix hyperproduction (Liu G. et al., 2010). Low expression of miR-184 in IPF patients correlates with high levels of p63 oncosuppressive protein, knockdown of which reduces TGF-β1-induced lung fibrosis. It was found that miR-184 binds complementarily to the 3′-UTR of the mRNA of the *TP63* gene, suppressing its expression (Li J. et al., 2021).

Among the microRNAs listed above associated with IPF (Huang C. et al., 2017), miR-326 (source – *hAT-Tip100* DNA transposon) and miR-1343 (source – LINE2 retrotransposon) originated from transposons, according to MDTE and data of various authors (Piriyapongsa et al., 2007; Gu et al., 2009; Filshtein et al., 2012; Tempel et al., 2012; Qin et al., 2015; Wei G. et al., 2016). In 2015, Yang et al. identified significant changes in the levels of 47 different miRNAs in the blood plasma of IPF patients compared with healthy controls (Yang et al., 2015). Of these 47 microRNAs, 4 originated from transposons: miR-31 (from LINE2), miR-302 (from the nonautonomous retroelement SINE/MIR), miR-335 (from SINE/MIR), miR-374 (from LINE2) (Wei G. et al., 2016). These 47 microRNAs are involved in the signaling pathways of TGF-β, mitogen-activated protein kinase (MAPK), PI3K-Akt, Wnt, HIF-1, Jak-STAT, Notch, actin cytoskeleton regulation (Yang et al., 2015). Reduced expression of miR-630 (Li R. et al., 2018) (derived from SINE/MIR (Wei G. et al., 2016)), miR-708-3p (Liu B. et al., 2018) (from LINE2 (Wei G. et al., 2016)) was detected in the blood plasma of patients with IPF. Elevated levels of transposon-derived miRNAs were shown for miR-487b (from SINE/MIR), miR-493 (from LINE2), miR-495 (from the LTR-containing retroelement ERVL-MaLT) (Zhang et al., 2021). MiR-340-5p, which promotes fibroblast proliferation in IPF by affecting the ATF and MAPK/p38 pathways (Wei Y.Q. et al., 2020), originated from the *TcMar-Mariner* DNA transposon (Wei G. et al., 2016).

Table 2 presents data on changes in the expression of miRNAs that originated in evolution from transposons (as well as long non-coding RNAs (lncRNA)) in IPF with a comparative analysis of scientific literature data on these miRNAs in bronchial asthma and chronic obstructive pulmonary disease. As can be seen from Table 2, among 24 miRNAs, 13 of them are unique in the changes in expression in patients with IPF: miR-9-5p, miR-27b, miR-153, miR-184, miR-326, miR-340, miR-374, miR-424, miR-487b, miR-489, miR-493, miR-630, miR-1343. Of these, 8 microRNAs (miR-153, miR-326, miR-340, miR-374, miR-487b, miR-493, miR-630, miR-1343) are evolutionarily derived from TE (Piriyapongsa et al., 2007; Gu et al., 2009; Filshtein et al., 2012; Tempel et al., 2012; Qin et al., 2015; Wei G. et al., 2016).

Investigation of the role of epigenetic factors in the development of IPF serves as the basis for the development of new methods of targeted therapy for the disease. Potential agents for the treatment of IPF may be non-coding RNAs. It was found that lncRNA PCAT29 (prostate cancer-associated transcript 29), which activates miRNA-221 and suppresses TGF-β, can be used to treat patients with IPF (Liu X. et al., 2018). It was discovered that expression of miR-506, which is complementary to the 3′-UTR of the p65 NF-κB subunit, is downregulated during IPF. Accordingly, the use of miR-506 as a target for targeted therapy may have an impact on apoptosis and inflammation in IPF (Zhu M. et al., 2019). Administration of antisense miR-21 reduced the severity of pathology in mice with bleomycin-induced lung fibrosis, suggesting the potential use of this miRNA in the treatment of IPF (Liu G. et al., 2010). Similar data were obtained for miR-708-3p (Liu B. et al., 2018). Overexpression of miR-184 suppresses TGF-β-induced fibrotic processes in the lung, therefore miR-184 can be considered for targeted therapy of IPF (Li J. et al., 2021). In animal experiments and in clinical studies on patients with IPF, the effectiveness of the interfering sequence for the long non-coding RNA lncITPF (sh-lncITPF), which reduces the index of fibrosis, collagen and vimentin, was also revealed. In patients with IPF, an increased expression of lncRNA-ITPF was revealed, which affects the acetylation of histones H3 and H4 in the promoter region of the *ITGBL1* gene, thus stimulating fibrosis. Transcription of lncITPF is under the control of TGF-β1/Smad2/3 (Song et al., 2019). For IPF treatment, the DR8 peptide (DHNNPQIR-NH2), which has a powerful antioxidant activity, was proposed. In an animal experiment with bleomycin-induced IPF, it was shown that after the use of DR8, fibrosis indicators, including profibrogenic and pro-inflammatory cytokines and marker proteins, were significantly reduced. DR8 reduced pathological changes caused by bleomycin, as well as collagen deposits (especially COL1). *In vivo* experiments showed that DR8 is able to suppress the proliferation and generation of reactive oxygen species stimulated by TGF-β1 (Wang et al., 2019).

Long non-coding RNAs (lncRNAs) are epigenetic factors, since they have transcriptional, post-transcriptional, and translational regulatory effects on the functioning of the genome. This effect is realized both due to the secondary structure of RNA, which provides interaction with proteins, and through hybridization with DNA and RNA due to the complementarity of nucleotides. Many lncRNA genes evolved from transposons (Johnson, Guigo, 2014). According to the NONCODEv4 database (http://www.noncode.org), more than 96,000 lncRNA genes have been annotated in humans, many of which contain TE sequences, which indicates the role of TEs in the origin of lncRNA genes (Johnson, Guigo, 2014). In addition, lncRNA can be formed during the processing of transcripts of LTR-containing retroelements (Lu et al., 2014) or LINE retrotransposons (Honson, Macfarlan, 2018). Analysis of GENOCODE and expressed RNA sequences showed that the majority of lncRNAs originated from transposons, since at least 83 % of them contain one or more retroelement fragments. On average, about 41 % of all lncRNA nucleotide sequences are identical to transposons (Kelley, Rinn, 2012). Thus, changes in lncRNA expression during IPF could indicate the role of transposons in the pathogenesis of the disease. Indeed, in 2017, Hao et al.

Table 2. Comparative analysis of the role of microRNAs in the development of idiopathic pulmonary fibrosis and other lung diseases

MicroRNA locus)/origin from TE	Direction of expression change (tissue)	Mechanism of influence		
		on IPF	on asthma (author)	on COPD (author)
let-7d (9q22.32)/–	↓ Bronchioles epithelium, alveoli	Targeted effect on mRNA of *EDA*, *LIX1L*, *MAPK11*, *NME4* genes	–*	Positively correlated (Tasena et al., 2018)
miR-9-5p (5q14.3)/–	↓ Bronchioles epithelium, alveoli	Antifibrotic	–	–
miR-18a (13q31.3)/–	↓ Bronchioles epithelium, alveoli	Antifibrotic, targeted effect on mRNA of *TGF-β*, *IL-6*, *IL-8* genes	Decreased expression (Martinez-Nunez et al., 2014)	–
miR-21 (17q23.1)/–	↑ Lung fibroblasts	Profibotic (regulates Smad7 expression by enhancing TGF-α and TGF-β signaling)	Increased expression in severe asthma (Liu J. et al., 2020)	Increased expression (He et al., 2021)
miR-26a (3p22.2)/–	↓ Bronchioles epithelium, alveoli	Antifibrotic, pro-inflammatory (increases levels of IL-5,-8,-12, TNF-α)	Increased expression (Shi et al., 2019)	–
miR-27b (9q22.32)/–	↓ Bronchioles epithelium, alveoli	Antifibrotic	–	–
miR-29 (7q32.3)/–	↓ Lung fibroblasts	Antifibrotic, extracellular matrix synthesis regulation	–	Increased expression (Kara et al., 2016)
miR-31 (9p21.3) /LINE2	↓ Bronchioles epithelium, alveoli (Yang et al., 2015)	Antifibrotic, pro-inflammatory (increases levels of IL-5,-8,-12, TNF-α)	Increased expression (Shi et al., 2019)	–
miR-92a-3p (13q31.3)/–	↓ Lung fibroblasts	Inhibits matrix metalloproteinase (MMP-1) synthesis	–	Decreased expression (Kara et al., 2016)
miR-101 (1p31.3)/–	↓ Bronchioles epithelium, alveoli	Antifibrotic	–	Increased expression (Hassan et al., 2012)
miR-153 (2q35)/–	↓ Lung fibroblasts	Antifibrotic (TGF-βRII regulation)	–	–
miR-184 (15q25.1)/–	↓ Bronchioles epithelium, alveoli	Onhibits the p63 protein, reducing TGF-β1 signaling; suppresses expression of *TP53*	–	–
miR-302 (4q25) /SINE/MIR	↑ Bronchioles epithelium, alveoli (Yang et al., 2015)	Regulator of allergic inflammation in mast cells, increases the production of IL-1β, IL-6, TNF-α	Increased expression (Xiao et al., 2018)	–
miR-323 (14q32.31)/–	↓ Bronchioles epithelium, alveoli	Attenuates TGF-α and TGF-β signaling, regulates T-lymphocyte differentiation	Increased expression (Karner et al., 2017)	–
miR-326 (11q13.4) /DNA-TE *hAT-Tip100*	↓ Bronchioles epithelium, alveoli (Huang C. et al., 2017)	Antifibrotic	–	–

End of Table 2

MicroRNA locus)/origin from TE	Direction of expression change (tissue)	Mechanism of influence		
		on IPF	on asthma (author)	on COPD (author)
miR-335 (7q32.2) /SINE/MIR	↓ Lung fibroblasts (Yang et al., 2015)	Suppresses the proliferation, migration and differentiation of fibroblasts, expression of *RB1*, *CARF*, *SGK3* genes	–	Reduced expression in smokers (Ong et al., 2019)
miR-340 (5q35.3) /DNA-TE *TcMar-Mariner*	↑ Lung fibroblasts (Wei Y.Q. et al., 2020)	Acts on the ATF and MAPK/p38 pathways, enhancing fibroblast proliferation	–	–
miR-374 (Xq13.2) /LINE2	↓ Lung fibroblasts (Yang et al., 2015)	Suppress the expression of MID1 ubiquitin ligase, inhibit mTOR signaling pathways (Unterbruner et al., 2018	–	–
miR-424 (Xq26.3)/–	↑ Lung fibroblasts	Profibrotic	–	–
miR-487b (14q32.31) /SINE/MIR	↑ Lung fibroblasts (Zhang et al., 2021)	Suppresses IL-33 expression, reducing Ig-E levels (Liu B. et al., 2018)	–	–
miR-489 (7q21.3)/–	↓ Bronchioles epithelium, alveoli	Antifibrotic	–	–
miR-493 (14q32.2) /LINE2	↑ Lung fibroblasts (Zhang et al., 2021)	Inhibits Wnt/B-catenin, Wnt/PCP, MEK/ERK, PI3K/AKT pathways (Huang L. et al., 2019)	–	–
miR-495 (14q32.31) /ERVL-MaLT	↑ Lung fibroblasts (Zhang et al., 2021)	Inhibits TNF-α, IL-1β, IL-6 synthesis	Decreased expression (Li W. et al., 2021)	Positive correlation (Li R. et al., 2020)
miR-630 (15q24.1) /SINE/MIR	↓ Lung fibroblasts (Li R. et al., 2018)	Regulates expression of *CDH2*, *VIM*, *EZH2*, *SOCS2*, *TFG*, *TLR4*, *Smad9*, *EP300* genes	–	–
miR-708 (11q14.1) /LINE2	↓ Lung fibroblasts (Liu B. et al., 2018)	Suppresses the expression of the metalloproteinase gene (*ADAM17*), inhibits CD44, RARRES2, ADAM33	Decreased expression (Dileepan et al., 2016)	–
miR-1343 (11p13) /LINE2	↓ Bronchioles epithelium, alveoli (Huang C. et al., 2017)	Antifibrotic (regulates the expression of TGF-β receptors)	–	–
lncRNA AP003419.16	↑ Lung tissue (Hao et al., 2017)	Regulates TGF-β1 signaling pathways	–	–
lncRNA ITPF	↑ Lung tissue (Song et al., 2019)	Regulates *ITGBL1* gene expression, stimulating lung fibrosis	–	–

Note. TE – transposable elements, IPF – idiopathic pulmonary fibrosis, COPD – chronic obstructive pulmonary disease, "–" – no association or correlation data.

determined a decrease in the levels of 1,376 different lncRNAs and an increase in the levels of 440 lncRNAs in the blood plasma of patients with ILF compared with healthy controls. The highest level was observed for lncRNA AP003419.16, which is involved in TGF-β1 signaling pathways and can be used as a marker of disease (Hao et al., 2017).

Influence of transposons on pulmonary fibrosis pathogenesis

The above data indicate the role of transposons in the emergence of noncoding RNAs that are involved in the pathogenesis of IPF and many other human diseases. The obtained results of molecular genetic studies of IPF are consistent with this assumption. It refers to the influence of transposons on the

aging processes that are involved in the pathogenesis of IPF and other multifactorial diseases (Gulati, Thannickal, 2019). In aging, retrotransposons containing long terminal repeats (Navalainen et al., 2018) and LINE1 (Mahmood et al., 2020) are activated. Moreover, their overexpression during aging enhances the production of interferon, contributing to aseptic inflammation in tissues (De Cecco et al., 2013).

Transposons (due to the relationship with microRNAs derived from them) are involved in the functioning of the immune system, the changes in which are associated with IPF (Korthagen et al., 2012; Noth et al., 2013; Fingerlin et al., 2016). For example, the miR-31 microRNA derived from LINE2 has a pro-inflammatory effect, enhancing the synthesis of IL-5,-8,-12, TNF-α (Shi et al., 2019); miR-302 (evolved from SINE/MIR) increases production of IL-1β, IL-6, TNF-α (Xiao et al., 2018). SINE/MIR are also a source of miR-487b, which represses IL-33 expression, reducing Ig-E levels (Liu H.C. et al., 2018). MiR-495 derived from ERVL-MaLT inhibits the synthesis of TNF-α, IL-1β, IL-6 (Li W. et al., 2021). In mammalian evolution, *RAG* genes were domesticated from ancient DNA transposons for V(D)J recombination in the immune system. Vertebrate antigen-specific immunity has two main features of DNA transposons. The components of immunity consist of recombinase (encoded by the *RAG1* and *RAG2* genes) and mobile DNA (limited to specific sites that the recombinase recognizes). RAG proteins are homologous to Tc1-element transposase (Lescale, Deriano, 2016). LTR-containing retroelements are involved in the regulation of the human immune system, as they are enhancers for the HLA-G gene (Chuong, 2018).

Transposons also affect the sirtuins (Wyman et al., 2017) and TGF-β (Liu G. et al., 2010; Chioma, Drake, 2017; Kang, 2017) involved in the pathogenesis of IPF. SIRT7 epigenetically represses LINE1 expression throughout the genome. An important role in this process is played by the interaction of SIRT7 with lamins A/C, since SIRT7 ensures the deacetylation of histone H3K18, facilitating the interaction of LINE1 with the nuclear lamina (Vazquez et al., 2019). Derived from an LTR-containing retroelement, the *PEG10* gene encodes a PEG10-RF1 protein that interacts with members of the TGF-β type I and II superfamily (Lux et al., 2005). The role of evolutionarily young retroelements in the regulation of TGF-β pathways, along with PDGF, EGFR and p38 signaling, was revealed (Nikitin et al., 2018). The role of retroelements in the epithelial-mesenchymal transition important for the development of IPF was shown (Sgalla et al., 2018; Li J. et al., 2021), which is induced by the non-autonomous retrotransposon Alu due to the modulating of miR-566 expression (Ruocco et al., 2018). Telomere dysfunction leading to the development of IPF (Mathai et al., 2015; Chioma, Drake, 2017; Allen et al., 2020) and other diseases is likely associated with changes in the activity of transposons, which are the evolutionary sources of genes involved in the functioning of telomeres (Arkhipova, 2017) and the telomerase gene (Garavis et al., 2013).

Conclusion

The investigation of epigenetic factors in the development of IPF is a promising direction in revealing the pathogenesis of the disease and developing more effective methods of its therapy. Through the study of miRNAs, it was shown that IPF is associated with an imbalance in the epigenetic regulation of the genome. Therefore, the reason for the development of IPF may be an imbalance in the control of the work of the genome by dynamic structures that play a role in age-associated pathology and aging of the body. The most appropriate control elements are transposons, since they affect the functioning of the immune system and are closely related to it evolutionarily. It has been suggested that the study of the role of transposons in the pathogenesis of IPF can reveal the pathways of the molecular cascade of the disease. Evidence for the role of transposons in the pathogenesis of IPF is the evolutionary emergence of long noncoding RNAs and miRNAs from transposons. Analysis of the MDTE database and scientific literature revealed 12 specific IPF-associated miRNAs that originated from transposons. Eight of these 12 microRNAs (miR-153, miR-326, miR-340, miR-374, miR-487b, miR-493, miR-630, miR-1343) are unique, since the change in their expression is specific for IPF and has not been described with other diseases of the bronchopulmonary system.

References

Allam J.S., Limper A.H. Idiopathic pulmonary fibrosis: is it a familial disease. *Curr. Opin. Pulm. Med.* 2006;12:312-317. DOI 10.1097/01.mcp.0000239546.24831.61.

Allen R.J., Guillen-Guio B., Oldham J.M., Ma S.F., Dressen A., Paynton M.L., Kraven L.M., Obeidat M., Li X., Ng M., Braybrooke R., Molina M., Hobbs B.D., Putman R.K., Flores C., Noth I., Jenkins R.G., Wain L.V. Genome-wide association study of susceptibility to idiopathic pulmonary fibrosis. *Am. J. Respir. Crit. Care. Med.* 2020;201(5):564-574. DOI 10.1164/rccm.201905-1017OC.

Arkhipova I.R., Yushenova I.A., Rodriguez F. Giant reverse transcriptase-encoding transposable elements at telomeres. *Mol. Biol. Evol.* 2017;34(9):2245-2257. DOI 10.1093/molbev/msx159.

Bagnato G., Roberts W.N., Roman J., Gangemi S. A systematic review of overlapping microRNA patterns in systemic sclerosis and idiopathic pulmonary fibrosis. *Eur. Respir. Rev.* 2017;26:160125. DOI 10.1183/16000617.0125-2016.

Bondarev I.E., Khavinson V.Kh. Suppression of alternative telomere lengthening in cancer cells with reverse transcriptase inhibitors. *Adv. Gerontol.* 2016;6(4):272-274. DOI 10.1134/S2079057016040020.

Casacuberta E. Drosophila: retrotransposons making up telomeres. *Viruses*. 2017;9(7):192. DOI 10.3390/v9070192.

Chioma O.S., Drake W.P. Role of microbial agents in pulmonary fibrosis. *Yale J. Biol. Med.* 2017;90(2):219-227.

Chuong E.B. The placenta goes viral: retroviruses control gene expression in pregnancy. *PLoS Biol.* 2018;16(10):e3000028. DOI 10.1371/journal.pbio.3000028.

De Cecco M., Criscione S.W., Peterson A.L., Neretti N., Sediby J.M., Kreiling J.A. Transposable elements become active and mobile in the genomes of aging mammalian somatic tissues. *Aging* (*Albany NY*). 2013;5:867-883. DOI 10.18632/aging.100621.

Dileepan M., Sarver A.E., Rao S.P., Panettieri R.A., Jr., Subramanian S., Kannan M.S. MicroRNA mediated chemokine responses in human airway smooth muscle cells. *PLoS One.* 2016;11(3):e0150842. DOI 10.1371/journal.pone.0150842.

Fernandez B.A., Fox G., Bhatia R., Sala E., Noble B., Nash D., Fernandez D., Duguid N., Dohey A., Kamel F., Edwards L., Mahoney K., Stuckless S., Parfrey P.S., Woods M.O. A Newfoundland cohort of familial and sporadic idiopathic pulmonary fibrosis patients: clinical and genetic features. *Respir. Res.* 2012;13:64. DOI 10.1186/1465-9921-13-64.

Filshtein T.J., Mackenzie C.O., Dale M.D., Dela-Cruz P.S., Ernst D.M., Frankenberger E.A., He C., Heath K.L., Jones A.S., Jones D.K., King E.R., Maher M.B., Mitchell T.J., Morgan R.R., Sirobhushanam S., Halkyard S.D., Tiwari K.B., Rubin D.A., Borchert G.M.,

Larson E.D. Orbid: Origin-based identification of microRNA targets. *Mob. Genet. Elements.* 2012;2(4):184-192. DOI 10.4161/mge.21617.

Fingerlin T.E., Zhang W., Yang I.V., Ainsworth H.C., Russell P.H., Blumhagen R.Z., Schwarz M.I., Brown K.K., Steele M.P., Loyd J.E., Cosgrove G.P., Lynch D.A., Growhong S., Markin C.R., Beckman K.B., Langefeld C.D., Schwartz D.A. Genome-wide imputation study identifies novel HLA locus for pulmonary fibrosis and potential role for auto-immunity in fibrotic idiopathic interstitial pneumonia. *BMC Genet.* 2016;17(1):74. DOI 10.1186/s12863-016-0377-2.

Garavis M., Gonzalez C., Villasante A. On the origin of the eukaryotic chromosome: the role of noncanonical DNA structures in telomere evolution. *Genome Biol. Evol.* 2013;5:1142-1150. DOI 10.1093/gbe/evt079.

Ge L., Habiel D.M., Hansbro P.M., Kim R.Y., Gharib S.A., Edelman J.D., Konigshoff M., Parimon T., Brauer R., Huang Y., Allen J., Jiang D., Kurkciyan A.A., Mizuno T., Stripp B.R., Noble P.W., Hogaboam C.M., Chen P. miR-323a-3p regulates lung fibrosis by targeting multiple profibrotic pathways. *JCI Insight.* 2016;1(20):e90301. DOI 10.1172/jci.insight.90301.

Gochuico B.R., Huizing M., Golas G.A., Scher C.D., Tsokos M., Denver S.D., Frei-Jones M.J., Gahl W.A. Interstitial lung disease and pulmonary fibrosis in Hermansky–Pudlak syndrome type 2, an adaptor protein-3 complex disease. *Mol. Med.* 2012;18(1):56-64. DOI 10.2119/molmed.2011.00198.

Gu T.J., Yi X., Zhao X.W., Zhao Y., Yin J.Q. Alu-directed transcriptional regulation of some novel miRNAs. *BMC Genomics.* 2009;10:563.

Gulati S., Thannickal V.J. The aging lung and idiopathic pulmonary fibrosis. *Am. J. Med. Sci.* 2019;357:384-389. DOI 10.1016/j.amjms.2019.02.008.

Hao X., Du Y., Qian L., Li D., Liu X. Upregulation of long noncoding RNA AP003419.16 predicts high risk of aging-associated idiopathic pulmonary fibrosis. *Mol. Med. Rep.* 2017;16(6):8085-8091. DOI 10.3892/mmr.2017.7607.

Hassan F., Nuovo G.J., Crawford M., Boyaka P.N., Kirkby S., Nana-Sinkam S.P., Cormet-Boyaka E. MiR-101 and miR-144 regulate the expression of the CFTR chloride channel in the lung. *PLoS One*. 2012;7(11):e50837. DOI 10.1371/journal.pone.0050837.

He S., Sun S., Lu J., Chen L., Mei X., Li L., Zeng Z., Zhong M., Xie L. The effects of the miR-21/SMAD7/TGF-β pathway on Th17 cell differentiation in COPD. *Sci. Rep.* 2021;11:6338. DOI 10.1038/s41598-021-85637-0.

Hodgson U., Laitinen T., Tukiainen P. Nationwide prevalence of sporadic and familial idiopathic pulmonary fibrosis: evidence of founder effect among multiplex families in Finland. *Thorax.* 2002;57(4):338-342. DOI 10.1136/thorax.57.4.338.

Honson D.D., Macfarlan T.S. A lncRNA-like role for LINE1s in development. *Dev. Cell.* 2018;46:132-134. DOI 10.1016/j.devcel.2018.06.022.

Huang C., Xiao X., Yang Y., Mishra A., Liang Y., Zeng X., Yang X., Xu D., Blackburn M.R., Henke C.A., Liu L. MicroRNA-101 attenuates pulmonary fibrosis by inhibiting fibroblast proliferation and activation. *J. Biol. Chem.* 2017;292:16420-16439. DOI 10.1074/jbc.M117.805747.

Huang L., Huang L., Li Z., Wei Q. Molecular mechanisms and therapeutic potential of miR-493 in cancer. *Crit. Rev. Eukaryot. Gene Expr*. 2019;29(6):521-528. DOI 10.1615/CritRevEukaryotGeneExpr.2019030056.

Johnson R., Guigo R. The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs. *RNA.* 2014;20:959-976. DOI 10.1261/rna.044560.114.

Kang H. Role of microRNAs in TGF-β signaling pathway-mediated pulmonary fibrosis. *Int. J. Mol. Sci.* 2017;18(12):2527. DOI 10.3390/ijms18122527.

Kara M., Kirkil G., Kalemci S. Differential expression of microRNAs in chronic obstructive pulmonary disease. *Adv. Clin. Exp. Med.* 2016;25(1):21-26. DOI 10.17219/acem/28343.

Karner J., Wawrzyniak M., Tankov S., Runnel T., Aints A., Kisand K., Altraja A., Kingo K., Akdis C.A., Akdis M., Rebane A. Increased microRNA-323-3p in IL-22/IL-17-producing T cells and asthma: a role in the regulation of the TGF-β pathway and IL-22 production. *Allergy.* 2017;72(1):55-65. DOI 10.1111/all.12907.

Kelley D., Rinn J. Transposable elements reveal a stem cell-specific class of long noncoding RNAs. *Genome Biol.* 2012;13(11):R107. DOI 10.1186/gb-2012-13-11-r107.

Korthagen N.M., van Moorsel C.H., Kazemier K.M., Ruven H.J., Grutters J.C. IL1RN genetic variations and risk of IPF: a meta-analysis and mRNA expression study. *Immunogenetics.* 2012;64:371-377. DOI 10.1007/s00251-012-0604-6.

Lawson W.E., Grant S.W., Ambrosini V., Womble K.E., Dawson E.P., Lane K.B., Markin C., Renzoni E., Lympany P., Thomas A.Q., Roldan J., Scott T.A., Blackwell T.S., Phillips J.A., Loyd J.E., du Bois R.M. Genetic mutations in surfactant protein C are a rare cause of sporadic cases of IPF. *Thorax.* 2004;59(11):977-980. DOI 10.1136/thx.2004.026336.

Lee M.G., Lee Y.H. A meta-analysis examining the association between the MUC5B rs35705950 T/G polymorphism and susceptibility to idiopathic pulmonary fibrosis. *Inflamm. Res.* 2015;64(6):463-470. DOI 10.1007/s00011-015-0829-6.

Lescale C., Deriano L. The RAG recombinase: beyond breaking. *Mech. Ageing Dev.* 2016;16:30263-30269. DOI 10.1016/j.mad.2016.11.003.

Li J., Pan C., Tang C., Tan W., Zhang W., Guan J. miR-184 targets TP63 to block idiopathic pulmonary fibrosis by inhibiting proliferation and epithelial-mesenchymal transition of airway epithelial cells. *Lab. Invest.* 2021;101(2):142-154. DOI 10.1038/s41374-020-00487-0.

Li R., Wang Y., Song X., Sun W., Zhang J., Liu Y., Li H., Meng C., Zhang J., Zheng Q., Lv C. Potential regulatory role of circular RNA in idiopathic pulmonary fibrosis. *Int. J. Mol. Med.* 2018;42(6):3256-3268. DOI 10.3892/ijmm.2018.3892.

Li R., Xu F., Wu X., Ji S., Xia R. CUL1-mediated organelle fission pathway inhibits the development of chronic obstructive pulmonary disease. *Comput. Math. Methods Med.* 2020;2020:5390107. DOI 10.1155/2020/5390107.

Li W., Wang X., Sun S., An H. Long non-coding RNA colorectal neoplasia differentially expressed correlates negatively with miR-33a and miR-495 and positively with inflammatory cytokines in asthmatic children. *Clin. Respir. J.* 2021;15(11):1175-1184. DOI 10.1111/crj.13424.

Liu B., Li R., Zhang J., Meng C., Zhang J., Song X., Lv C. MicroRNA-708-3p as a potential therapeutic target via the ADAM17-GATA/STAT3 axis in idiopathic pulmonary fibrosis. *Exp. Mol. Med.* 2018;50(3):e465. DOI 10.1038/emm.2017.311.

Liu G., Friggeri A., Yang Y., Milosevic J., Ding Q., Thannickal V.J., Kaminski N., Abraham E. miR-21 mediates fibrogenic activation of pulmonary fibroblasts and lung fibrosis. *J. Exp. Med.* 2010;207(8):1589-1597. DOI 10.1084/jem.20100035.

Liu H.C., Liao Y., Liu C.Q. miR-487b mitigates allergic rhinitis through inhibition of the IL-33/ST2 signaling pathway. *Eur. Rev. Med. Pharmacol. Sci.* 2018;22(23):8076-8083. DOI 10.26355/eurrev_201812_16497.

Liu J., Li C., Zhang C., Zhang Z. LncRNA-CASC7 enhances corticosteroid sensitivity via inhibiting the PI3K/AKT signaling pathway by targeting miR-21 in severe asthma. *Pulmonology.* 2020;26(1):18-26. DOI 10.1016/j.pulmoe.2019.07.001.

Liu X., Gao S., Xu H. LncRNACAT29 inhibits pulmonary fibrosis via the TGF-β1-regulated RASAL1/ERK1/2 signal pathway. *Mol. Med. Rep*. 2018;17(6):7781-7788. DOI 10.3892/mmr.2018.8807.

Lu X., Sachs F., Ramsay L., Jacques P.-E., Göke J., Bourque G., Ng H.-H. The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. *Nat. Struct. Mol. Biol.* 2014;21:423-425. DOI 10.1038/nsmb.2799.

Lux A., Beil C., Majety M., Barron S., Gallione C.J., Kuhn H.M., Gerg J.N., Kioschis P., Marchuk D.A., Hafner M. Human retroviral *gag*- and *gag-pol*-like proteins interact with the transforming growth factor-beta receptor activin receptor-like kinase 1. *J. Biol. Chem.* 2005;280(9):8482-8493. DOI 10.1074/jbc.M409197200.

Mahmood W., Erichsen L., Ott P., Schulz W., Fischer J.C., Arauzo-Bravo M.J., Bendhack M.L., Hassan M., Santourlidis S. Aging-associated distinctive DNA methylation changes of LINE-1 retrotransposons in pure cell-free DNA from human blood. *Sci. Rep.* 2020;10(1):22127. DOI 10.1038/s41598-020-79126-z.

Martinez-Nunez R., Bondanese V.P., Louafi F., Francisco-Garcia A.S., Rupani H., Bedke N., Holgate S., Howerth P.H., Davies D.E., Sanchez-Elsner T. A microRNA network dysregulated in asthma controls IL-6 production in bronchial epithelial cells. *PLoS One.* 2014;9(10):e111659. DOI 10.1371/journal.pone.0111659.

Mathai S.K., Yang I.V., Schwarz M.I., Schwartz D.A. Incorporating genetics into the identification and treatment of idiopathic pulmonary fibrosis. *BMC Med.* 2015;13:191. DOI 10.1186/s12916-015-0434-0.

McKusick V.A., Fisher A.M. Congential cystic disease of the lung with progressive pulmonary fibrosis and carcinomatosis. *Ann. Intern. Med.* 1958;48:774-790. DOI 10.7326/0003-4819-48-4-774.

Methot D.B., Leblanc E., Lacasse Y. Meta-analysis of gastroesophageal reflux disease and idiopathic pulmonary fibrosis. *Chest.* 2019;155(1):33-43. DOI 10.1016/j.chest.2018.07.038.

Mustafin R.N. Aging and interrelation of telomeres with transposable elements. *Adv. Gerontol.* 2019;32(5):693-701. (in Russian).

Mustafin R.N., Khusnutdinova E.K. The role of transposable elements in the ecological morphogenesis under the influence of stress. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding.* 2019;23(4):380-389. DOI 10.18699/VJ19.506.

Nevalainen T., Autio A., Mishra B.H., Marttila S., Jyha M., Hurme M. Aging-associated pattern in the expression of human endogenous retroviruses. *PLoS One.* 2018;13(12):e0207407. DOI 10.1371/journal.pone.0207407.

Nikitin D., Penzar D., Garazha A., Sorokin M., Tkachev V., Borisov N., Pltorak V., Buzdin A.A. Profiling of human molecular pathways affected by retrotransposons at the level of regulation by transcription factor proteins. *Front. Immunol.* 2018;9:30. DOI 10.3389/fimmu.2018.00030.

Noth I., Zhang Y., Ma S.F., Flores C., Barbes M., Huang Y., Broderick S.M., Wade M.S., Kaminski N., Garcia J.G.N. Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. *Lancet Respir Med.* 2013;1(4):309-317. DOI 10.1016/S2213-2600(13)70045-6.

Ong J., van den Berg A., Faiz A., Boudewijn I.M., Timens W., Vermeulen C.J., Oliver B.G., Kok K., Terpstra M.M., van den Berge M., Brandsma C.A., Kluiver J. Current smoking is associated with decreased expression of miR-335-5p in parenchymal lung fibroblasts. *Int. J. Mol. Sci.* 2019;20(20):5176. DOI 10.3390/ijms20205176.

Piriyapongsa J., Marino-Ramirez L., Jordan I.K. Origin and evolution of human microRNAs from transposable elements. *Genetics.* 2007;176(2):1323-1337. DOI 10.1534/genetics.107.072553.

Qin S., Jin P., Zhou X., Chen L., Ma F. The role of transposable elements in the origin and evolution of microRNAs in human. *PLoS One.* 2015;10(6):e0131365. DOI 10.1371/journal.pone.0131365.

Raghu G., Weycker D., Edelsberg J., Bradford W.Z., Oster G. Incidence and prevalence of idiopathic pulmonary fibrosis. *Am. J. Respir. Crit. Care. Med.* 2006;174:810-816. DOI 10.1164/rccm.200602-163OC.

Ruocco F.D., Basso V., Rivoire M., Mehlen P., Ambati J., De Falco S., Tarallo V. *Alu* RNA accumulation induces epithelial-to-mesenchymal transition by modulating miR-566 and is associated with cancer progression. *Oncogene.* 2018;37(5):627-637. DOI 10.1038/onc.2017.369.

Seibold M.A., Wise A., Speer M., Steele M., Brown K., Lloyd J.E., Fingerlin T.E., Garantziotis S., Herron A., Slifer S.H., Schwartaz D.A. A common MUC5B promoter polymorphism and pulmonary fibrosis. *N. Engl. J. Med.* 2011;364:1503-1512. DOI 10.1056/NEJMoa1013660.

Sgalla G., Iovene B., Clavello M., Ori M., Varone F., Richeldi L. Idiopathic pulmonary fibrosis: pathogenesis and management. *Respir. Res.* 2018;19(1):32. DOI 10.1186/s12931-018-0730-2.

Sheng G., Chen P., Wei Y., Yue H., Chu J., Zhao J., Wang Y., Zhang W., Zhang H.L. Viral infection increases the risk of idiopathic pulmonary fibrosis: a meta-analysis. *Chest.* 2020;157(5):1175-1187. DOI 10.1016/j.chest.2019.10.032.

Shi Z., Sun Y., Wang K., Jia J., Yang J., Li Y. Effects of miR-26a/miR-146a/miR-31 on airway inflammation of asthma mice and asthma children. *Eur. Rev. Med. Pharmacol. Sci.* 2019;23(12):5432-5440. DOI 10.26355/eurrev_201906_18212.

Song X., Xu P., Meng C., Song C., Blackwell T.S., Li R., Li H., Zhang J., Lv C. LncITPF promotes pulmonary fibrosis by targeting hnRNP-L depending on its host gene ITGBL1. *Mol. Ther.* 2019;27(2):380-393. DOI 10.1016/j.ymthe.2018.08.026.

Spagnolo P., Cottin V. Genetics of idiopathic pulmonary fibrosis: from mechanistic pathways to personalized medicine. *J. Med. Genet.* 2017;54:93-99. DOI 10.1136/jmedgenet-2016-103973.

Stuart B.D., Choi J., Zaidi S., Xing C., Holohan B., Chen R., Choi M., Dharwadkar P., Torres F., Girod C.E., Weissler J., Lifton R.P., Garcia C.K. Exome sequencing links mutations in PARN and RTEL1 with familial pulmonary fibrosis and telomere shortening. *Nat. Genet.* 2015;47:512-517. DOI 10.1038/ng.3278.

Tasena H., Faiz A., Timens W., Noordhoek J., Hylkema M.N., Gosens R., Hiemstra P.S., Spira A., Postma D.S., Tew G.W., Grimbaldeston M.A., van den Berge M., Heijink I.H., Brandsma C. MicroRNA-mRNA regulatory networks underlying chronic mucus hypersecretion in COPD. *Eur. Respir. J.* 2018;52(3):1701556. DOI 10.1183/13993003.01556-2017.

Tempel S., Pollet N., Tahi F. NcRNAclassifier: a tool for detection and classification of transposable element sequences in RNA hairpins. *BMC Bioinformatics.* 2012;13:246. DOI 10.1186/1471-2105-13-246.

Thomas A.Q., Lane K., Phillips J., Prince M., Markin C., Speer M., Schwartz D.A., Gaddipati R., Marney A., Johnson J., Roberts R., Haines J., Stahlman M., Loyd J.E. Heterozygosity for a surfactant protein C gene mutation associated with usual interstitial pneumonitis and cellular nonspecific interstitial pneumonitis in one kindred. *Am. J. Respir. Crit. Care. Med.* 2002;165(9):1322-1328. DOI 10.1164/rccm.200112-123OC.

Tsakiri K.D., Cronkhite J.T., Kuan P.J., Xing C., Raghu G., Weissler J.C., Rosenblatt R.L., Shay J.W., Gracia C.K. Adult-onset pulmonary fibrosis caused by mutations in telomerase. *Proc. Natl. Acad. Sci. USA.* 2007;104(18):7552-7557. DOI 10.1073/pnas.0701009104.

Unterbruner K., Matthes F., Schilling J., Nalavade R., Weber S., Winter J., Kraub S. MicroRNAs miR-19, miR-340, miR-374 and miR-542 regulate MID1 protein expression. *PLoS One.* 2018;13(1):e0190437. DOI 10.1371/journal.pone.0190437.

Vazquez B.N., Thackray J.K., Simonet N.G., An W., Vaquero A., Tischfield J.A., Serrano L. SIRT7 mediates L1 elements transcriptional repression and their association with the nuclear lamina. *Nucleic Acids Res.* 2019;47:7870-7885. DOI 10.1093/nar/gkz519.

Wang D., Yan Z., Bu L., An C., Deng B., Zang J., Rao J., Cheng L., Zhang J., Zhang B., Xie J. Protective effect of peptide DR8 on bleomycin-induced pulmonary fibrosis by regulating the TGF-β/MAPK signaling pathway and oxidative stress. *Toxicol. Appl. Pharmacol.* 2019;382:114703. DOI 10.1016/j.taap.2019.114703.

Wei G., Qin S., Li W., Chen L., Ma F. MDTE DB: a database for microRNAs derived from transposable element. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 2016;13(6):1155-1160. DOI 10.1109/TCBB.2015.2511767.

Wei Y.Q., Guo Y.F., Yang S.M., Ma H.H., Li J. MiR-340-5p mitigates the proliferation and activation of fibroblast in lung fibrosis by targeting TGF-β/p38/ATF1 signaling pathway. *Eur. Rev. Med. Pharmacol. Sci.* 2020;24(11):6252-6261. DOI 10.26355/eurrev_202006_21523.

Wyman A.E., Noor Z., Fishelevich R., Lockatell V., Shah N.G., Todd N.W., Atamas S.P. Sirtuin 7 is decreased in pulmonary fibrosis and regulates the fibrotic phenotype of lung fibroblasts. *Am. J. Physiol. Lung. Cell. Mol. Physiol.* 2017;312:L945-L958. DOI 10.1152/ajplung.00473.2016.

Xiao L., Jiang L., Hu Q., Li Y. MiR-302e attenuates allergic inflammation *in vitro* model by targeting RelA. *Biosci. Rep.* 2018;38(3):BSR20180025. DOI 10.1042/BSR20180025.

Yang G., Yang L., Wang W., Wang J., Wang J., Xu Z. Discovery and validation of extracellular/circulating microRNAs during idiopathic pulmonary fibrosis disease progression. *Gene.* 2015;562:138-144. DOI 10.1016/j.gene.2015.02.065.

Zhang Y.F., Gu L.N., Qi J., Xia Q.Q., Tian L.J., Jiang W.L., Cao M.S. Construction of potential idiopathic pulmonary fibrosis related microRNA and messenger RNA regulatory network. *Chin. Med. J.* (*Engl*). 2021;134(5):584-586. DOI 10.1097/CM9.0000000000001276.

Zhao J., Ren Y., Qu Y., Jiang W., Lv C. Pharmacodynamic and pharmacokinetic assessment of pulmonary rehabilitation mixture for the treatment of pulmonary fibrosis. *Sci. Rep.* 2017;7:3458. DOI 10.1038/s41598-017-02774-1.

Zhu L., Wang L., Luo X., Zhang Y., Ding Q., Jiang X., Wang X., Pan Y., Chen Y. Tollip, an intracellular trafficking protein, is a novel modulator of the transforming growth factor-β signaling pathway. *J. Biol. Chem.* 2012;287(47):39653-39663. DOI 10.1074/jbc.M112.388009.

Zhu M., An Y., Zhang X., Wang Z., Duan H. Experimental pulmonary fibrosis was suppressed by microRNA-506 through NF-kappa-mediated apoptosis and inflammation. *Cell. Tissue Res.* 2019;378:255-265. DOI 10.1007/s00441-019-03054-2.

ORCID ID

R.N. Mustafin orcid.org/0000-0002-4091-382X

Conflict of interest. The author declares no conflict of interest.

Received August 5, 2021. Revised September 14, 2021. Accepted January 13, 2022.

Вавиловский журнал генетики и селекции. 2022;26(3):319-326
DOI 10.18699/VJGB-22-38

Original Russian text www.bionet.nsc.ru/vogis/

Analysis of the low density lipoprotein receptor gene (*LDLR*) mutation spectrum in Russian familial hypercholesterolemia

V.B. Vasilyev ✉, F.M. Zakharova, T.Yu. Bogoslovskaya, M.Yu. Mandelshtam

Institute of Experimental Medicine, St. Petersburg, Russia
✉ vadim@biokemis.ru

Abstract. Familial hypercholesterolemia (FH) is a very common human hereditary disease in Russia and in the whole world with most of mutations localized in the gene coding for the low density lipoprotein receptor (*LDLR*). The object of this review is to systematize the knowledge about *LDLR* mutations in Russia. With this aim we analyzed all available literature on the subject and tabulated the data. More than 1/3 (80 out of 203, i. e. 39.4 %) of all mutations reported from Russia were not described in other populations. To date, most *LDLR* gene mutations have been characterized in large cities: Moscow (130 entries), Saint Petersburg (50 entries), Novosibirsk (34 mutations) and Petrozavodsk (19 mutations). Other regions are poorly studied. The majority of pathogenic mutations (142 out of 203 reported here or 70 %) were revealed in single pedigrees; 61 variants of mutations were described in two or more genealogies; only 5 mutations were found in 10 or more families. As everywhere, missense mutations prevail among all types of nucleotide substitutions in *LDLR*, but the highest national specificity is imparted by frameshift mutations: out of 27 variants reported, 19 (or 70 %) are specific for Russia. The most abundant in mutations are exons 4 and 9 of the gene due to their largest size and higher occurrence of mutations in them. Poland, the Czech Republic, Italy and the Netherlands share the highest number of mutations with the Russian population. Target sequencing significantly accelerates the characterization of mutation spectra in FH, but due to the absence of systematic investigations in the regions, one may suggest that most of *LDLR* mutations in the Russian population have not been described yet.
Key words: familial hypercholesterolemia; low density lipoprotein receptor gene; mutations.

For citation: Vasilyev V.B., Zakharova F.M., Bogoslovskaya T.Yu., Mandelshtam M.Yu. Analysis of the low density lipoprotein receptor gene (*LDLR*) mutation spectrum in Russian familial hypercholesterolemia. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2022;26(3):319-326. DOI 10.18699/VJGB-22-38

Анализ спектра мутаций гена рецептора низкой плотности (*LDLR*) при семейной гиперхолестеринемии в России

В.Б. Васильев ✉, Ф.М. Захарова, Т.Ю. Богословская, М.Ю. Мандельштам

Институт экспериментальной медицины, Санкт-Петербург, Россия
✉ vadim@biokemis.ru

Аннотация. Семейная гиперхолестеринемия – распространенное во всем мире наследственное заболевание человека, при котором чаще всего дефекты обнаруживаются в гене рецептора липопротеинов низкой плотности (*LDLR*). Цель работы – систематизировать знания о мутациях гена *LDLR* в России. Проведен анализ литературы по предмету исследования, составлены сводные таблицы, показывающие встречаемость мутаций в отдельных регионах, и определены часто встречающиеся мутации. Более трети (80 из 203, т. е. 39.4 %) патогенных или вероятно патогенных мутаций представлены вариантами, специфичными для России и не встречающимися в других странах. Наибольшее количество вариантов охарактеризовано в крупных городах: Москве (130 патогенных мутаций), Санкт-Петербурге (50), Новосибирске (34) и Петрозаводске (19), тогда как регионы охарактеризованы гораздо хуже. Подавляющее число патогенных мутаций (142 из 203, или 70 %) найдено в единичных семьях, и только 61 вид мутаций встречался в двух или в нескольких родословных. Лишь 5 видов мутаций были найдены не менее чем в 10 семьях. Как и везде в мире, в России в гене *LDLR* преобладают миссенс-мутации, но особенным национальным своеобразием характеризуются мутации типа сдвига рамки считывания: из 27 найденных вариантов 19 (70 %) специфичны для России. Наивысшее число мутаций в гене *LDLR* в российской популяции обнаружено в четвертом и девятом экзонах. Это определяется тем, что четвертый и девятый экзоны являются самыми протяженными в гене и кодируют функционально важные участки белка, что обусловливает повышенную плотность патогенных мутаций в расчете на один нуклеотид длины именно в этих экзонах. Российская популяция имеет наибольшее число совпадающих му-

таций с популяциями Польши, Чешской Республики, Нидерландов и Италии. Внедрение методов таргетного секвенирования существенно ускорило характеристику мутационного спектра при семейной гиперхолестеринемии, но из-за отсутствия систематических исследований в регионах большинство видов мутаций в России, вероятнее всего, еще не описано.
Ключевые слова: семейная гиперхолестеринемия; рецептор липопротеинов низкой плотности; мутации.

Introduction

The term 'familial hypercholesterolemia' (FH) is generally used to refer to monogenic diseases caused by mutations in the low-density lipoprotein (LDL) receptor (*LDLR*) gene (OMIM 606945), in the apolipoprotein B (*APOB*) gene (OMIM 107730), in the *PCSK9* gene (OMIM 607786), in the adapter protein gene for the LDL receptor *LDLRAP1* (OMIM 605747) and some minor genes, such as *STAP1*, *APOE*, *LIPA*, or in the sterol transporter genes, sterolins *ABCG5/ABCG8* (Defesche et al., 2017; Berberich, Hegele, 2019). At the same time, 80–85 % of FH cases are caused by mutations in the LDL receptor gene. Mutations in the apolipoprotein B gene are responsible for 5–10 % of FH cases. Mutations in the *PCSK9* gene and in the LDL receptor adapter protein gene are the rarest, occurring in no more than 1 % of patients with FH.

It was previously believed that heterozygous FH occurs in 1 out of 500 people examined in the population, but the current data allow us to conclude that it is much more frequent. A study of 69,106 patients in Denmark who were diagnosed with FH based on the recommendations of the Dutch Lipid Clinic Network (DLCN) demonstrated that the prevalence of the disease is 1:219 (Benn et al., 2012). It may be even higher in Russia, i. e. 1:148 (Ershova et al., 2017). However, in this instance, cases of not only definite, but also probable FH were taken into account. Such frequency allows attributing FH to the most common monogenic human diseases.

Already in 2018, the ClinVar database (Landrum et al., 2016) included 4973 variants of the *LDLR* gene (Iacocca et al., 2018) associated with FH, of which 2351 variants were classified as pathogenic, and 1525 as probably pathogenic, the rest considered as benign variants or variants of uncertain clinical significance. The history of FH research in Russia has recently been reviewed (Vasilyev et al., 2020; Meshkov et al., 2021a). Most of the mutations leading to FH, as expected, were found in the *LDLR* gene, 187 pathogenic or likely pathogenic variants of which were identified in Russia (Meshkov et al., 2021a); 67 out of 187 were not described in other populations of the world. An important article on the genetics of FH in St. Petersburg was later published based on targeted sequencing of genes involved in the origin of the disease (Miroshnikova et al., 2021). As a result, 23 variants of the *LDLR* gene sequence were found in the St. Petersburg population, most of which had not been described in that area (Mandelshtam et al., 1993; Tatishcheva et al., 2001; Zakharova et al., 2005, 2007; Vasilyev et al., 2020). The results of studying mutations in the regions of Russia appeared only recently (Meshkov et al., 2021b). Continuously replenished data on the subject indicate the necessity for regular revisions of the tables of the *LDLR* gene mutations in Russia (Meshkov et al., 2021a). In our view, such a notion supports the relevance of the present review: it already mentions 203 pathogenic or likely pathogenic variants in the gene discussed.

Methods

All available literature concerning *LDLR* gene mutations in the Russian population was analyzed. As a result, a summarizing table was compiled that significantly expands our knowledge about the spectrum of mutations in Russia, as compared to previously published data (Mandelshtam et al., 1993, 1998a, b; Chakir et al., 1998a, b; Krapivner et al., 2001; Mandelshtam, Maslennikov, 2001; Tatishcheva et al., 2001; Zakharova et al., 2001, 2005, 2007; Meshkov et al., 2004, 2009, 2021a, b; Voevoda et al., 2008; Komarova et al., 2013a–c; Korneva et al., 2013–2016, 2017a, b; Shakhtshneider et al., 2017, 2019a, b, 2021; Averkova et al., 2018; Semenova et al., 2020) (Supplementary Table)[1].

The term 'mutation', used throughout the text, implies all rare variants of a gene (widespread polymorphisms excluded) that are potentially capable of causing a disease, including the variants with proven or highly probable pathogenicity. Synonymous substitutions are not considered in this review (their list is presented in Vasilyev et al., 2020).

Results and discussion

There are more than 4900 variants of the *LDLR* gene described in the world, as was already mentioned (Iacocca et al., 2018). This review reports 203 pathogenic or likely pathogenic mutations of this gene in Russia (see Suppl. Table). However, this diversity is not likely to represent the variability of the receptor gene in the Russian population, since GWS was introduced somewhat recently, and a systematic study of the FH genetics in quite a few regions of Russia has not been conducted. The studies were carried out mainly in large cities (Fig. 1). Currently most of the mutations found are specific for each of these cities, and a significantly smaller proportion is common with other regions. Thus, the largest number of pathogenic or probably pathogenic variants in Russia (101) was found only in Moscow (Krapivner et al., 2001; Meshkov et al., 2004, 2009, 2021a; Averkova et al., 2018; Semenova et al., 2020), 35 variants were found only in St. Petersburg (Mandelshtam et al., 1993; Zakharova et al., 2001, 2005, 2007; Tatishcheva et al., 2001), 23 – only in Novosibirsk (Voevoda et al., 2008; Shakhtshneider et al., 2017, 2019a, b), 11 – only in Petrozavodsk (Komarova et al., 2013a–c; Korneva et al., 2013, 2014, 2017a, b), 33 – in other regions, sometimes in several regions simultaneously (Meshkov et al., 2021b).

[1] Supplementary Table is available in the online version of the paper: http://vavilov.elpub.ru/jour/manager/files/Suppl_Vasilyev_engl.pdf

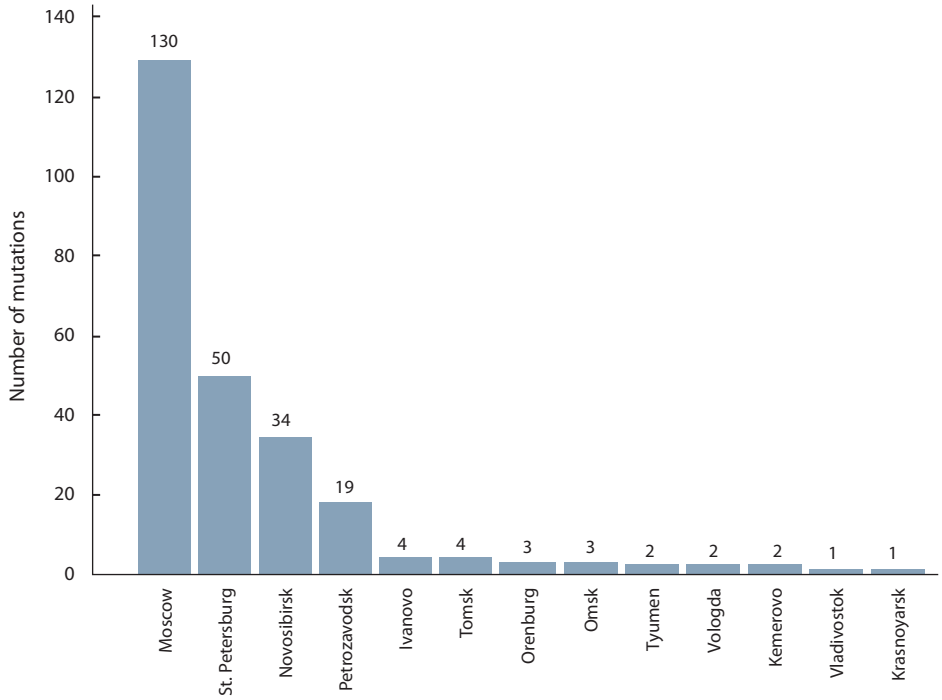


Fig. 1. The number of pathogenic and likely pathogenic variants in the *LDLR* gene found in Russian cities with examined populations (excluding common polymorphisms and benign variants).

Table 1. Analysis of the spectrum of pathogenic and likely pathogenic mutations of the *LDLR* gene in Russia and in the world

Mutation type	Total number of pathogenic variants found in Russia (%)	Number and proportion of variants specific for Russia (%)	Number and proportion of variants shared with other populations in the world (%)	Proportion of variants of this type in the world	
				Defesche et al., 2017	Chora et al., 2018
Large deletions	10 (5)	6 (60)	4 (40)	8–10	9
Splice site mutations	18 (9)	8 (44)	10 (56)	8–10	9
In-frame deletions and insertions	7 (3.5)	2 (29)	5 (71)	15–20	4
Frameshift mutations	27 (13.5)	19 (70)	8 (30)		18
Nonsense mutations	20 (10)	2 (10)	18 (90)	12–15	9
Missense mutations	121 (59)	43 (35)	78 (65)	40–50	46
Regulatory mutations	Not found	–	–	No data	2
Synonymous substitutions	Not considered	–	–	No data	1
Intron mutations	Not considered	–	–	No data	2
Total	203 (100)	80 (39.4)	123 (60.6)		100

Table 2. Pathogenic variants in the *LDLR* gene that were found in patients with FH from the Russian population in 10 or more families

Nucleotide change according to reference sequence NM_000527.5 (*LDLR*)	Predicted change in protein	Number of families	Nucleotide change according to dbSNP	Populations in Russia	References for Russia	Other countries in the world
c.478T > G	p.(Cys160Gly)	10	rs879254540	St. Petersburg, Novosibirsk, Moscow	Chakir et al., 1998a; Mandelshtam, Maslennikov, 2001; Meshkov et al., 2004, 2009, 2021a	None
c.654_656delTGG	p.(Gly219del)	14	rs121908027	St. Petersburg, Novosibirsk, Moscow	Mandelshtam et al., 1998; Mandelshtam, Maslennikov, 2001; Zakharova et al., 2005; 2007; Meshkov et al., 2021a	The Czech Republic, Germany, UK, Israel, The Netherlands, Poland, RSA, USA, mainly in Ashkenazi Jews
c.986G > A	p.(Cys329Tyr)	13	rs761954844	St. Petersburg, Novosibirsk, Moscow, Petrozavodsk	Zakharova et al., 2005, 2007; Shakhtshneider et al., 2019b; Semenova et al., 2020b; Meshkov et al., 2021a, b; Miroshnikova et al., 2021	Canada, China, The Czech Republic, Philippines, Poland, Taiwan, The Netherlands
c.1202T > A	p.(Leu401His)	33	rs121908038	St. Petersburg, Novosibirsk, Moscow, Krasnoyarsk, Petrozavodsk	Zakharova et al., 2005; Shakhtshneider et al., 2019b; Meshkov et al., 2021a, b; Miroshnikova et al., 2021	Finland, The Netherlands, Brazil, Mexico, Norway
c.1775G > A	p.(Gly592Glu)	43	rs137929307	St. Petersburg, Novosibirsk, Moscow	Zakharova et al., 2001, 2007; Voevoda et al., 2008; Semenova et al., 2020; Meshkov et al., 2021a, b	Austria, Belgium, Brazil, Greece, Canada, Germany, The Czech Republic, Italy, The Netherlands, Norway, Poland, Portugal, Spain, USA, etc.

Table 3. Occurrence of Russian *LDLR* gene mutations in other countries

Country	Total number of mutations described	Reference	Number of mutations shared with Russia
Poland	99	Chmara et al., 2010	19
The Czech Republic	129	Tichý et al., 2012; Chora et al., 2018	23
The Netherlands	306	Fouchier et al., 2005; Chora et al., 2018	36
Spain	205	Mozas et al., 2004; Chora et al., 2018	12
Italy	251	Bertolini et al., 2013; Pirillo et al., 2017; Chora et al., 2018	28

At the moment, we can state a wide variety of mutations in Russia, of which more than one third (39.4 %) are specific for the local population and have not been found anywhere else in the world so far (Table 1). The distribution of mutations by type in Russia is very similar to that in the world (see Table 1).

From the analysis of Table 1 it follows that the underestimated percentage of large deletions is due to the fact that targeted sequencing was introduced only recently, and a targeted large-scale search for large deletions was not carried out: the researchers focused on exon screening, which determined a slightly higher percentage of missense mutations than in the world as a whole. The systematic search for large deletions in the *LDLR* gene began very recently (Shakhtshneider et al., 2021). It included patients with FH in whom high-throughput targeted sequencing did not reveal significant mutations in a panel of 43 lipid metabolism genes using multiplex ligase-dependent PCR (MLPA), which revealed two deletions of the *LDLR* gene in a studied sample of 80 patients with FH.

Only a few variants of the *LDLR* gene occur in several families, but unique mutations predominate. The majority of pathogenic mutations (142 out of 203 or 70 %) in Russia were also found in singular families, and only 61 types of mutations were found in two families or in more pedigrees. Around the world, the largest number of mutations is described in the 4th exon. Firstly, it is the largest of all exons in the *LDLR* gene. Secondly, it has the highest density of mutations, amounting to 0.882 variants per nucleotide (Chora et al., 2018). It is in this exon that the largest number of functionally characterized mutations was found, and almost all of those have a pathogenic effect. Our study showed that the largest number of mutations in the *LDLR* gene in the Russian population is localized in the largest exons, i. e. the 4th and 9th (Fig. 2). Considering all the information available, we conclude

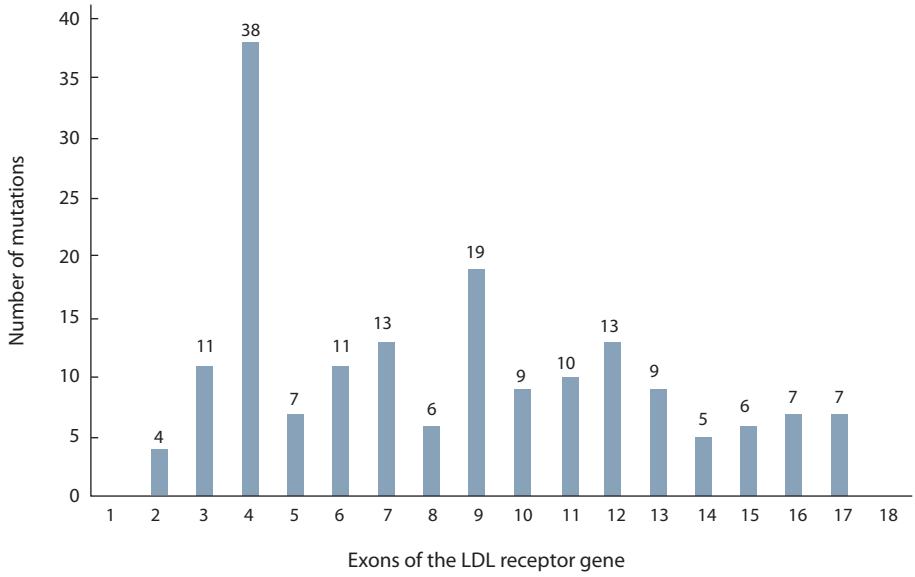


Fig. 2. Distribution of pathogenic and probably pathogenic variants by exons in the LDL receptor gene in patients with FH in Russia.

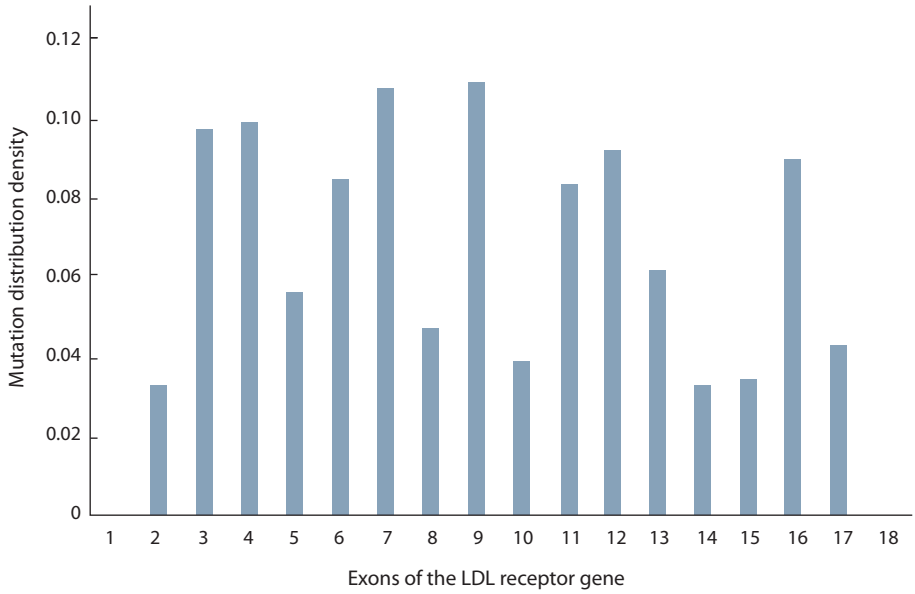


Fig. 3. Distribution density of pathogenic and likely pathogenic mutations in the exons of the LDL receptor gene per nucleotide in patients with FH in Russia.

that the highest mutation density per nucleotide (Fig. 3) is determined in the 9th, but not in the 4th exon, in contrast to other databases reported worldwide (Chora et al., 2018).

Only five pathogenic variants of the *LDLR* gene in Russia can be classified as major, found in 10 or more families (Table 2). Of these, only one is specific for Russia, while the rest are widespread in the world.

The greatest similarity in the spectrum of mutations in the *LDLR* gene in Russia is observed with Poland, the Czech Republic, the Netherlands, Spain and Italy, which is

partly determined by the fact that these populations are the best characterized (Table 3). This similarity probably results from the presence of widespread Caucasian race mutations in the world, but is not due to migration or the founder effect.

Conclusion

Thus, the success of further study of the mutation spectrum of the *LDLR* gene will depend on several factors, one of which is the formation of a complete nation-wide register of patients with FH, and the other is the introduction of targeted sequencing into routine practice.

References

Averkova A.O., Brazhnik V.A., Speshilov G.I., Rogozhina A.A., Koroleva O.S., Zubova E.A., Galyavich A.S., Tereshenko S.N., Boyeva O.I., Zateyshchikov D.A. Targeted sequencing in patients with clinically diagnosed hereditary lipid metabolism disorder and acute coronary syndrome. *Bulletin of Russian State Medical University.* 2018;5:80-86. DOI 10.24075/brsmu.2018.061.

Benn M., Watts C.F., Tybjaerg-Hansen A., Nordestgaard B.G. Familial hypercholesterolemia in the Danish general population: prevalence, coronary artery disease and cholesterol-lowering medication. *J. Clin. Endocrinol. Metab.* 2012;97(11):3956-3964. DOI 10.1210/jc.2012-1563.

Berberich A.J., Hegele R.A. The complex molecular genetics of familial hypercholesterolaemia. *Nat. Rev. Cardiol.* 2019;16(1):9-20. DOI 10.1038/s41569-018-0052-6.

Bertolini S., Pisciotta L., Rabacchi C., Cefalù A.B., Noto D., Fasano T., Signori A., Fresa R., Averna M., Calandra S. Spectrum of mutations and phenotypic expression in patients with autosomal dominant hypercholesterolemia identified in Italy. *Atherosclerosis.* 2013;227(2):342-348. DOI 10.1016/j.atherosclerosis.2013.01.007.

Chakir Kh., Skobeleva N.A., Shevtsov S.P., Konstantinov V.O., Denisenko A.D., Schwartz E.I. Two novel slavic point mutations in the low-density lipoprotein receptor gene in patients with familial hypercholesterolemia from St. Petersburg, Russia. *Mol. Genet. Metab.* 1998a;63(1):31-34. DOI 10.1006/mgme.1997.2614.

Chakir K., Mandelshtam M.Ju., Shevtsov S.P., Golubkov V.I., Skobeleva N.A., Shur Yu.A., Zakharova F.M., Lipovetskyi B.M., Konstantinov V.O., Denisenko A.D., Gaitskhoki V.S., Schwartz E.I. Two novel low-density lipoprotein receptor gene mutations (E397X and 347delGCC) in St. Petersburg familial hypercholesterolemia. *Mol. Genet. Metab.* 1998b;65(4):311-314. DOI 10.1006/mgme.1998.2762.

Chmara M., Wasag B., Zuk M., Kubalska J., Wegrzyn A., Bednarska-Makaruk M., Pronicka E., Wehr H., Defesche J.C., Rynkiewicz A., Limon J. Molecular characterization of Polish patients with familial hypercholesterolemia: novel and recurrent *LDLR* mutations. *J. Appl. Genet.* 2010;51(1):95-106. DOI 10.1007/BF03195716.

Chora J.R., Medeiros A.M., Alves A.C., Bourbon M. Analysis of publicly available *LDLR*, *APOB*, and *PCSK9* variants associated with familial hypercholesterolemia: application of ACMG guidelines and implications for familial hypercholesterolemia diagnosis. *Genet. Med.* 2018;20(6):591-598. DOI 10.1038/gim.2017.151.

Defesche J.C., Gidding S.S., Harada-Shiba M., Hegele R.A., Santos R.D., Wierzbicki A.S. Familial hypercholesterolaemia. *Nat. Rev. Dis. Prim.* 2017;3:17093. DOI 10.1038/nrdp.2017.93.

Ershova A.I., Meshkov A.N., Bazhan S.S., Storozhok M.A., Efanov A.Y., Medvedeva I.V., Indukaeva E.V., Danilchenko Y.V., Kuzmina O.K., Barbarash O.L., Deev A.D., Shalnova S.A., Boytsov S.A. The prevalence of familial hypercholesterolemia in the West Siberian region of the Russian Federation: a substudy of the ESSE-RF. *PLoS One.* 2017;12(7):e0181148. DOI 10.1371/journal.pone.0181148.

Fouchier S.W., Kastelein J.J., Defesche J.C. Update of the molecular basis of familial hypercholesterolemia in the Netherlands. *Hum. Mutat.* 2005;26(6):550-556. DOI 10.1002/humu.20256.

Iacocca M.A., Chora J.R., Carrié A., Freiberger T., Leigh S.E., Defesche J.C., Kurtz C.L., DiStefano M.T., Santos R.D., Humphries S.E., Mata P., Jannes C.E., Hooper A.J., Wilemon K.A., Benlian P., O'Connor R., Garcia J., Wand H., Tichy L., Sijbrands E.J., Hegele R.A., Bourbon M., Knowles J.W., ClinGen FH Variant Curation Expert Panel. ClinVar database of global familial hypercholesterolemia-associated DNA variants. *Hum. Mutat.* 2018;39(11):1631-1640. DOI 10.1002/humu.23634.

Komarova T.Yu., Golovina A.S., Grudinina N.A., Zakharova F.M., Korneva V.A., Lipovetsky B.M., Serebrenitskaya M.P., Konstantinov V.O., Vasilyev V.B., Mandelshtam M.Yu. New mutations in low-density lipoprotein receptor gene in familial hypercholesterolemia patients from Petrozavodsk. *Russ. J. Genet.* 2013a;49(6):673-676. DOI 10.1134/S1022795413040066.

Komarova T.Yu., Golovina A.S., Grudinina N.A., Zakharova F.M., Korneva V.A., Lipovetskii B.M., Serebrenitskaya M.P., Konstantinov V.O., Vasilyev V.B., Mandelshtam M.Yu. "Finnish" mutations in LDL receptor gene: a rare cause of familial hypercholesterolemia in St. Petersburg and Petrozavodsk. *Bull. Exp. Biol. Med.* 2013b;155(3):380-383. DOI 10.1007/s10517-013-2159-4.

Komarova T.Yu., Korneva V.A., Kuznetsova T.Yu., Golovina A.S., Vasilyev V.B., Mandelshtam M.Yu. Familial hypercholesterolemia mutations in Petrozavodsk: no similarity to St. Petersburg mutation spectrum. *BMC Med. Genet.* 2013c;14(1):128. DOI 10.1186/1471-2350-14-128.

Korneva V.A., Bogoslovskaya T.Yu., Kuznetsova T.Yu., Mandelshtam M.Yu., Vasiliev V.B. Familial hypercholesterolemia due to a new mutation in the low density lipoprotein receptor gene. *Klinicheskaia Meditsina = Clinical Medicine.* 2014;92(7):49-53. (in Russian)

Korneva V.A., Kuznetsova T.Yu., Bogoslovskaya T.Yu., Murtazina R.Z., Didio A.V., Serebrenitskaya M.P., Konstantinov V.O., Mandelshtam M.Yu., Vasilyev V.B. Clinical case of a patient with family hypercholesterolemia caused by mutation c.1859G>C (p.W620S) in the gene of human low density lipoproteins receptor. *Kardiovaskulyarnaya Terapiya i Profilaktika = Cardiovascular Therapy and Prevention.* 2015;14(6):52-53. DOI 10.15829/1728-8800-2015-6-52-53. (in Russian)

Korneva V.A., Kuznetsova T.Yu., Bogoslovskaya T.Yu., Murtazina R.Z., Didio A.V., Serebrenitskaya M.P., Konstantinov V.O., Mandelshtam M.Yu., Vasilyev V.B. Modern diagnostic abilities of familial hypercholesterolemia before clinical features of atherosclerosis appearance. *Ateroscleros i Dislipidemii = Journal of Atherosclerosis and Dyslipidemias.* 2016;1:22-32. (in Russian)

Korneva V.A., Kuznetsova T.Yu., Bogoslovskaya T.Yu., Polyakov D.S., Vasilyev V.B., Orlov A.V., Mandelshtam M.Yu. Cholesterol levels in genetically determined familial hypercholesterolaemia in Russian Karelia. *Cholesterol.* 2017a;2017:9375818. DOI 10.1155/2017/9375818.

Korneva V.A., Kuznetsova T.Yu., Komarova T.Yu., Golovina A.S., Mandelshtam M.Yu., Konstantinov V.O., Vasilyev V.B. A case of familial hypercholesterolemia caused by a novel mutation p. FsS65:D129X of human low density lipoprotein receptor gene. *Kardiologiia = Cardiology.* 2013;53(5):50-54. (in Russian)

Korneva V.A., Kuznetsova T.Yu., Murtazina R.Z., Didio A.V., Bogoslovskaya T.Yu., Mandelshtam M.Yu., Vasilyev V.B. The familial hypercholesterolemia caused by a novel human low density lipoprotein receptor gene mutation c.1327 T>C (p.W433R).

Kardiologiia = Cardiology. 2017b;57(2):12-16. DOI 10.18565/cardio.2017.2.12-16. (in Russian)

Krapivner S.R., Malyshev P.P., Poltaraus A.B., Kukharchuk V.V., Bochkov V.N. A case of familial hypercholesterolemia caused by a novel mutation D461Y in the low density lipoprotein receptor gene. *Kardiologiia = Cardiology.* 2001;41(1):92-94. (in Russian)

Landrum M.J., Lee J.M., Benson M., Brown G., Chao C., Chitipiralla S., Gu B., Hart J., Hoffman D., Hoover J., Jang W., Katz K., Ovetsky M., Riley G., Sethi A., Tully R., Villamarin-Salomon R., Rubinstein W., Maglott D.R. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* 2016;44(D1):D862-D868. DOI 10.1093/nar/gkv1222.

Mandelshtam M., Chakir K., Shevtsov S., Golubkov V., Skobeleva N., Lipovetsky B., Konstantinov V., Denisenko A., Gaitskhoki V., Schwartz E. Prevalence of Lithuanian mutation among St. Petersburg Jews with familial hypercholesterolemia. *Hum. Mutat.* 1998a;12(4):255-258. DOI 10.1002/(SICI)1098-1004(1998)12:4<255::AID-HUMU6>3.0.CO;2-E.

Mandelshtam M.Ju., Golubkov V.I., Schur Yu.A., Lipovetsky B.M., Gaitskhoki V.S. A novel mutation 347delGCC in the human low density lipoprotein receptor gene. *Russ. J. Bioorg. Chem*. 1998b;24(10):710-711.

Mandelshtam M.Ju., Lipovetskyi B.M., Schwartzman A.L., Gaitskhoki V.S. A novel deletion in the low density lipoprotein receptor gene in a patient with familial hypercholesterolemia from Petersburg. *Hum. Mutat.* 1993;2(4):256-260. DOI 10.1002/humu.1380020404.

Mandelshtam M.Yu., Maslennikov A.B. Study of familial hypercholesterolemia molecular genetics in Russia. In: Maslennikov A.B. (Ed.) Molecular Technologies in Practical Medicine. Novosibirsk: Publ. House “Manuskript”, 2001;58-64. (in Russian)

Meshkov A., Ershova A., Kiseleva A., Zotova E., Sotnikova E., Petukhova A., Zharikova A., Malyshev P., Rozhkova T., Blokhina A., Limonova A., Ramensky V., Divashuk M., Khasanova Z., Bukaeva A., Kurilova O., Skirko O., Pokrovskaya M., Mikova V., Snigir E., Akinshina A., Mitrofanov S., Kashtanova D., Makarov V., Kukharchuk V., Boytsov S., Yudin S., Drapkina O. The *LDLR*, *APOB*, and *PCSK9* variants of index patients with familial hypercholesterolemia in Russia. *Genes.* 2021a;12(1):66. DOI 10.3990/genes12010066.

Meshkov A.N., Ershova A.I., Kiseleva A.V., Shalnova S.A., Drapkina O.M., Boytsov S.A. On behalf of the FH-ESSE-RF Investigators. The prevalence of heterozygous familial hypercholesterolemia in selected regions of the Russian Federation: the FH-ESSE-RF study. *J. Pers. Med.* 2021b;11(6):464. DOI 10.3390/jpm11060464.

Meshkov A.N., Malyshev P.P., Kukharchuk V.V. Familial hypercholesterolemia in Russia: genetic and phenotypic characteristics. *Terapevticheskiy Arkhiv = Therapeutic Archive*. 2009;81(9):23-28. (in Russian)

Meshkov A.N., Stambolsky D.V., Krapivner S.R., Bochkov V.N., Kukharchuk V.V., Malyshev P.P. Low density lipoprotein receptor gene mutations in patients with clinical diagnosis of familial hypercholesterolemia. *Kardiologiia = Cardiology.* 2004;44(9):58-61. (in Russian)

Miroshnikova V.V., Romanova O.V., Ivanova O.N., Fedyakov M.A., Panteleeva A.A., Barbitoff Y.A., Muzalevskaya M.M., Urazgildeeva S.A., Gurevich V.S., Urazov S.P., Scherbak S.G., Sarana A.M., Semenova N.A., Anisimova I.V., Guseva D.M., Pchelina S.N., Glotov A.S., Zakharova E.Y., Glotov O.S. Identification of novel variants in the *LDLR* gene in Russian patients with familial hypercholesterolemia using targeted sequencing. *Biomed. Rep.* 2021;14(1):15. DOI 10.3892/br.2020.1391.

Mozas P., Castillo S., Tejedor D., Reyes G., Alonso R., Franco M., Saenz P., Fuentes F., Almagro F., Mata P., Pocoví M. Molecular characterization of familial hypercholesterolemia in Spain: identification of 39 novel and 77 recurrent mutations in *LDLR*. *Hum. Mutat.* 2004;24(2):187. DOI 10.1002/humu.9264.

Pirillo A., Garlaschelli K., Arca M., Averna M., Bertolini S., Calandra S., Tarugi P., Catapano A.L., LIPIGEN Group. Spectrum of mutations in Italian patients with familial hypercholesterolemia: new results from the LIPIGEN study. *Atheroscler. Suppl.* 2017;29:17-24. DOI 10.1016/j.atherosclerosissup.2017.07.002.

Semenova A.E., Sergienko I.V., García-Giustiniani D., Monserrat L., Popova A.B., Nozadze D.N., Ezhov M. Verification of underlying genetic cause in a cohort of Russian patients with familial hypercholesterolemia using targeted next generation sequencing. *J. Cardiovasc. Dev. Dis.* 2020;7(2):16. DOI 10.3390/jcdd7020016.

Shakhtshneider E.V., Ivanoshchuk D.E., Fishman V.S., Belokopytova P.S., Orlov P.S., Timoshchenko O.V., Voevoda M.I. Target sequencing of genes *LDLR*, *APOB*, *PCSK9*, *LDLRAP1* and gene cluster *APOA5-A4-C3-A1* in patients with familial hypercholesterolemia. *Kompleksnye Problemy Serdechno-sosydistykh Zabolevaniy = Complex Issues of Cardiovascular Diseases.* 2019a;8(S3-2):71. (in Russian)

Shakhtshneider E.V., Ivanoshchuk D.E., Makarenkova K.V., Orlov P.S., Timoshchenko O.V., Bazhan S.S., Nikitin Y.P., Voevoda M.I. Cascade genetic screening in diagnostics of heterozygous familial hypercholesterolemia: clinical case. *Russian Journal of Cardiology*. 2017;6(146):178-179. DOI 10.15829/1560-4071-2017-6-178-179. (in Russian)

Shakhtshneider E., Ivanoshchuk D., Orlov P., Timoshchenko O., Voevoda M. Analysis of the *LDLR*, *APOB*, *PCSK9* and *LDLRAP1* genes variability in patients with familial hypercholesterolemia in West Siberia using targeted high throughput resequencing. *Atherosclerosis*. 2019b;287:e285. DOI 10.1016/j.atherosclerosis.2019.06.883.

Shakhtschneider E., Ivanoshchuk D., Timoshenko O., Orlov P., Semayev S., Valeev E., Goonko A., Ladygina N., Voevoda M. Analysis of rare variants in genes related to lipid metabolism in patients with familial hypercholesterolemia in Western Siberia (Russia). *J. Pers. Med.* 2021;11(11):1232. DOI 10.3990/jpm11111232.

Tatishcheva Yu.A., Mandelshtam M.Yu., Golubkov V.I., Lipovetsky B.M., Gaitskhoki V.S. Four new mutations and two polymorphic variants of the low-density lipoprotein receptor gene in familial hypercholesterolemia patients from St. Petersburg. *Russ. J. Genet.* 2001;37(9):1082-1087. DOI 10.1023/A:1011973817437.

Tichý L., Freiberger T., Zapletalová P., Soška V., Ravčuková B., Fajkusová L. The molecular basis of familial hypercholesterolemia in the Czech Republic: spectrum of *LDLR* mutations and genotype-phenotype correlations. *Atherosclerosis.* 2012;223(2):401-408. DOI 10.1016/j.atherosclerosis.2012.05.014.

Vasilyev V., Zakharova F., Bogoslovskaya T., Mandelshtam M. Familial hypercholesterolemia in Russia: three decades of genetic studies. *Front. Genet.* 2020;11:550591. DOI 10.3389/fgene.2020.550591.

Voevoda M.I., Kulikov I.V., Shakhtshneider E.V., Maksimov V.N., Pilipenko I.V., Tereschenko I.P., Kobzev E.F., Romaschenko A.G., Nikitin Yu.P. The spectrum of mutations in the low-density lipoprotein receptor gene in the Russian population. *Russ. J. Genet.* 2008;44(10):1191-1194. DOI 10.1134/S1022795408100074.

Zakharova F.M., Damgaard D., Mandelshtam M.Y., Golubkov V.I., Nissen P.H., Nilsen G.G., Stenderup A., Lipovetsky B.M., Konstantinov V.O., Denisenko A.D., Vasilyev V.B., Faergeman O. Familial

hypercholesterolemia in St.-Petersburg: the known and novel mutations found in the low density lipoprotein receptor gene in Russia. *BMC Med. Genet.* 2005;6:6. DOI 10.1186/1471-2350-6-6.

Zakharova F.M., Golubkov V.I., Mandelshtam M.Ju., Lipovetskii B.M., Gaitskhoki V.S. Identification of novel missense mutation G571E, novel silent mutation H229H, nonsense mutation C74X, and four single nucleotide polymorphisms in the low-density lipoprotein receptor gene in patients with familial hypercholesterolemia from St. Petersburg. *Russ. J. Bioorg. Chem.* 2001;27:349-351. DOI 10.1023/a:1012300632671.

Zakharova F.M., Tatishcheva Yu.A., Golubkov V.I., Lipovetsky B.M., Konstantinov V.O., Denisenko A.D., Faergeman O., Vasilyev V.B., Mandelshtam M.Yu. Familial hypercholesterolemia in St. Petersburg: diversity of mutations argues against a strong founder effect. *Russ. J. Genet*. 2007;43(9):1046-1052. DOI 10.1134/S1022795407090116.

ORCID ID

V.B. Vasilyev orcid.org/0000-0002-9707-262X
F.M. Zakharova orcid.org/0000-0002-9558-3979
T.Yu. Bogoslovskaya orcid.org/0000-0002-9480-1073
M.Yu. Mandelshtam orcid.org/0000-0002-7135-3239

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

Received November 9, 2021. Revised December 14, 2022. Accepted February 10, 2022.

Вавиловский журнал генетики и селекции. 2022;26(3):327-335
DOI 10.18699/VJGB-22-39

Original Russian text www.bionet.nsc.ru/vogis/

Design and assembly of plant-based COVID-19 candidate vaccines: recent development and future prospects

E.A. Uvarova[1#] ✉, P.A. Belavin[1#], E.V. Deineko[1, 2]

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Tomsk State University, Tomsk, Russia
✉ uvarova@bionet.nsc.ru

Abstract. An outbreak of a new variant of the coronavirus infection, known as COVID-19, occurred at the end of 2019 in China, in the city of Wuhan. It was caused by the SARS-CoV-2 virus. This variant of the virus is characterized by a high degree of variability and, as the current situation with its spread across different regions of the globe shows, it can lead to a progressive spread of infection among the human population and become the cause of a pandemic. The world scientific community is making tremendous efforts to develop means of protection, prevention and treatment of this disease based on modern advances in molecular biology, immunology and vaccinology. This review provides information on the current state of research in the field of vaccine development against COVID-19 with an emphasis on the role of plants in solving this complex problem. Although plants have long been used by mankind as sources of various medicinal substances, in a pandemic, plant expression systems become attractive as biofactories or bioreactors for the production of artificially created protein molecules that include protective antigens against viral infection. The design and creation of such artificial molecules underlies the development of recombinant subunit vaccines aimed at a rapid response against the spread of infections with a high degree of variability. The review presents the state of research covering a period of just over two years, i.e. since the emergence of the new outbreak of coronavirus infection. The authors tried to emphasize the importance of rapid response of research groups from various scientific fields towards the use of existing developments to create means of protection against various pathogens. With two plant expression systems – stable and transient – as examples, the development of work on the creation of recombinant subunit vaccines against COVID-19 in various laboratories and commercial companies is shown. The authors emphasize that plant expression systems have promise for the development of not only protective means under conditions of rapid response (subunit vaccines), but also therapeutic agents in the form of monoclonal antibodies against COVID-19 synthesized in plant cells.
Key words: plant-based vaccines; plant expression systems; virus-like particles; transient expression; stable expression; recombinant proteins.

For citation: Uvarova E.A., Belavin P.A., Deineko E.V. Design and assembly of plant-based COVID-19 candidate vaccines: recent development and future prospects. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2022;26(3):327-335. DOI 10.18699/VJGB-22-39

Разработка и создание кандидатных вакцин против COVID-19 на основе растительных систем экспрессии: состояние исследований и перспективы

Е.А. Уварова[1#] ✉, П.А. Белавин[1#], Е.В. Дейнеко[1, 2]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Национальный исследовательский Томский государственный университет, Томск, Россия
✉ uvarova@bionet.nsc.ru

Аннотация. Вспышка нового варианта коронавирусной инфекции, известной как COVID-19, произошла в конце 2019 г. в Китае, в городе Ухань, и была вызвана вирусом SARS-CoV-2. Данный вариант вируса характеризуется высокой степенью изменчивости и, как показывает сложившаяся ситуация с его распространением по различным регионам земного шара, способен приводить к прогрессирующему распространению инфекции среди человеческой популяции и становиться причиной возникновения пандемии. Мировое сообщество исследователей прилагает огромные усилия для разработки средств защиты, профилактики и

* E.A. Uvarova and P.A. Belavin contibuted equally to the preparation of the manuscript

лечения этого заболевания, основываясь на современных достижениях в области молекулярной биологии, иммунологии и вакцинологии. В предлагаемом обзоре суммирована информация о современном состоянии исследований в области создания вакцин против COVID-19 с акцентом на роль растений в решении этой сложной проблемы. Хотя растения издавна использовались человечеством в качестве источников различных лекарственных субстанций, в условиях пандемии растительные системы экспрессии становятся привлекательными в качестве биофабрик или биореакторов для наработки искусственно созданных белковых молекул, включающих протективные антигены для вакцины против вирусной инфекции. Дизайн и конструирование таких искусственных молекул лежат в основе создания рекомбинантных субъединичных вакцин, нацеленных на быстрое реагирование против распространения инфекций с высокой степенью изменчивости. В обзоре представлено состояние исследований, охватывающее период немногим более двух лет, т. е. с момента появления новой коронавирусной инфекции. Обсуждается важность быстрого реагирования исследовательских групп из разных научных областей в направлении использования уже существующих разработок для создания средств защиты против различных патогенов. На примере двух растительных систем экспрессии – стабильной и транзиентной – показано развитие работ по конструированию рекомбинантных субъединичных вакцин против COVID-19 в различных лабораториях и коммерческих компаниях. Подчеркнута перспективность использования растительных систем экспрессии для разработки не только средств защиты в условиях быстрого реагирования (субъединичные вакцины), но и терапевтических средств в виде моноклональных антител против COVID-19, синтезируемых в растительных клетках.
Ключевые слова: растительные вакцины; системы экспрессии растений; вирусоподобные частицы; транзиентная экспрессия; стабильная экспрессия; рекомбинантные белки.

Introduction

Vaccination is one of the most effective methods of combating infectious diseases. A vaccine is a preparation that stimulates the body to form a protective reaction against an infectious agent. Vaccination is based on the programming of specific immunological mechanisms for protection against pathogens of various infections. Although humanity has managed to avoid outbreaks of many dangerous infections precisely thanks to vaccination, the vaccines available in the arsenal are still far from "ideal". The use of traditional vaccines, the production of which is based on attenuated or inactivated pathogens, is sometimes accompanied by sensitization of the body, a large load on the immune system, reactogenicity, toxicity, etc. (Francis, 2018).

The methods and approaches developed to date in the field of molecular biology, immunology, vaccinology, cellular and synthetic biology, as well as bioinformatics, allow us to take a fresh look at the opening opportunities for creating more advanced means of protection against pathogens of viral and bacterial origin, devoid of the above disadvantages. The use of modern biology methods makes it possible to identify and isolate biological macromolecules or their fragments that could be used as immunogenic components to activate the immune system in response to a pathogen. Such components can be proteins of pathogens (for example, envelope proteins of infectious agents), which are immunogens. With the use of genetic engineering technologies, the direction of creating recombinant subunit vaccines is successfully developing – artificially created protein molecules that include protective antigens in combination with adjuvants synthesized in various expression systems (Salazar-González et al., 2015; Demurtas et al., 2016; Fischer, Buyel, 2020; McNulty et al., 2020; Rybicki, 2020).

The creation of recombinant subunit vaccines is most relevant for pathogens characterized by a high level of variability. These pathogens include viral pathogens that cause acute respiratory infections and influenza (Shoji et al., 2011; Ward et al., 2020). These pathogens can lead to a progressive spread of infection in the human population and cause epidemics and pandemics. These pathogens include a new type of coronavirus – SARS-CoV-2.

The availability of data on the genome structure of a new virus strain isolated using classical methods of virology, electron microscopy, and molecular analysis at the end of 2019 (Zhu et al., 2020) opened up wide opportunities for applying new approaches to designing vaccines. The first statements about clinical trials (https://clinicaltrials.gov/ct2/show/NCT04283461) appeared two months after the publication of the primary structure of the genome of this virus (Zhu et al., 2020). This fact indicates that the existing developments and understanding of the molecular mechanisms of the formation of protective reactions on the part of the body's immune system make it possible to respond quickly enough to the emergence of new variants of a viral infection (Pogrebnyak et al., 2005). However, the use of new vaccines for the prevention of the population is dictated by the need for a deep assessment of their effectiveness and impact on the human body, as well as the possibilities for their industrial development (Jiang, 2020).

In this review, the authors made an attempt to analyze the possibilities of using plant expression systems aimed at creating antiviral subunit recombinant vaccines, in particular, candidate vaccines against COVID-19.

Plant-based expression systems

When developing new generation vaccines, including recombinant vaccines, the question of finding highly effective and cost-effective systems for their expression remains topical. Currently, *Escherichia coli*, several species of *Saccharomyces*, and mammalian cells are most commonly used for these purposes. New prospects for the production of recombinant proteins are opening up with the use of recombinant plants (biopharming) that could act as plant (edible) vaccines (Sala-

zar-González et al., 2015). Plants in whose tissues recombinant immunogens are synthesized and accumulated are attractive for obtaining substances for veterinary and medical purposes, including for obtaining subunit recombinant antiviral vaccines. Many leading biotechnology laboratories and commercial firms use plant cells as an alternative expression system for the production of recombinant proteins for medical purposes (Fischer, Buyel, 2020; Rybicki, 2020).

Plant-made expression platforms used for the synthesis of recombinant proteins, including vaccinogenic ones, are based on stable expression of the target gene when it is delivered to the nuclear or chloroplast genomes of the plant, as well as on its transient expression. Figure 1 shows two main platforms being developed in leading biotechnological centers for the production of recombinant proteins, including medical ones, using the synthetic capabilities of the transcription-translation apparatus of plants. The general principle underlying these platforms is as follows: by genetic engineering, an artificial matrix with a target gene is created, according to which the corresponding protein is synthesized and accumulated in plant tissues. Plant tissues can be freeze-dried and encapsulated, or the recombinant protein can be isolated and purified directly from the tissues. As a rule, genes of envelope proteins of pathogens of infectious diseases, which are immunogens, are used as a target gene. As part of expression cassettes, target genes can be integrated into the plant genome (nuclear or chloroplast), which will ensure stable expression of the target gene and accumulation of the target product in plant tissues (see Fig. 1, *a*). However, the use of the chloroplast genome for these purposes, although it seems very promising, is still far from practical application due to the existence of a large number of unsolved problems (Waheed et al., 2015; Yu Y. et al., 2020).

To deliver target genes to plant tissues in a transient expression system, viral vectors specifically designed for this purpose are used (Sainsbury et al., 2010), as well as the *Nicotiana benthamiana* plant species, the structural features of the leaf parenchyma of which are optimal for successful agroinfiltration (see Fig. 1, *b*). IconGenetics (Germany) has developed and patented a "magnification" system in which the yield of a recombinant protein in a transient expression system can reach up to 80 % of the total soluble protein (TSP) (Gleba et al., 2005). Despite the relatively low (slightly more than 1 % ORP) yield of the recombinant protein in plants with a stable expression system in the case of nuclear transformation, the use of already available agricultural technologies for growing transgenic plants provides them with unlimited scalability at minimal cost (Kermode, 2018). Thus, a plant platform with stable expression of the target gene is promising for the production of high-volume products, such as vaccines for disease prevention, especially in developing countries.

Transgenic or transplastomic plants with stable expression of the target gene are used for large-scale production of recombinant proteins over a long period, while the characteristics of transient expression make it possible to obtain the required amounts of recombinant protein in short periods of time, which seems to be extremely important when an emergency response to the spread of a pathogen is required. For example, the US Food and Drug Administration (FDA) approved an emergency

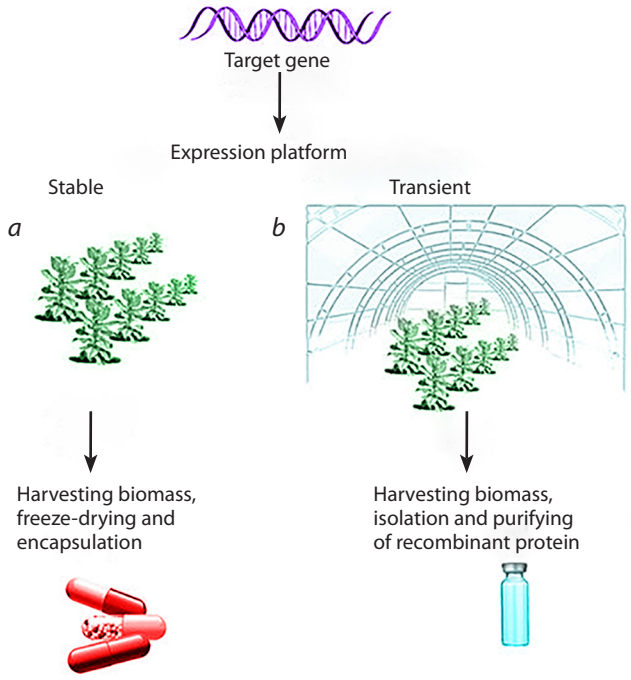


Fig. 1. Plant-made expression platforms used for the synthesis of recombinant proteins.

cocktail against Ebola virus called ZMapp™, consisting of three monoclonal antibodies transiently synthesized in tobacco plants (Phoolcharoen et al., 2011).

The transient expression system is promising for the development of small-scale accumulation of personalized drugs, such as anti-idiotypic scFv antibodies for non-Hodgkin's lymphoma, as well as in case of a need for mass vaccination of the population in case of outbreaks of seasonal viral diseases caused by rapidly mutating viruses. In *N. benthamiana* plants, after three weeks from the moment the viral sequence was already isolated, sufficiently large amounts of antigens were synthesized from influenza virus strains H5N1 (bird flu) and H1N1 (swine flu) (Hodgins et al., 2019; Makarkov et al., 2019). These recombinant proteins, synthesized in a transient plant expression system, are being considered as candidate influenza vaccines and have completed phase II human trials (Pillet et al., 2019).

DowAgroSciences (USA) has developed the Concert™ plant cell culture system as an advanced platform for the production of a recombinant antigen against Newcastle disease virus (pseudoplague) in poultry. Although the company did not launch commercial production of this recombinant vaccine, this technology has served as the basis for other commercial products. The reality and effectiveness of this approach has been repeatedly confirmed by researchers from the world's leading biotechnological laboratories, as well as by the activities of numerous companies and firms specializing in the production of one or more closely related products based on their own expression platform (Margolin et al., 2018; Rybicki, 2018).

Given the dramatic impact of the COVID-19 pandemic, it is critical to consider all the technologies at the disposal of researchers that could be applied to combat the causative agent of this infectious disease, the SARS-CoV-2 virus. Since the technology for the production of plant biopharmaceuticals has already been generally developed, it seems very attractive in the context of a pandemic in terms of producing not only inexpensive vaccines, but also antibodies used for therapy, prevention and diagnosis. The production of antibodies, such as anti-COVID-19, seems even more promising than vaccinogenic proteins, since recombinant plant-derived antibodies can be produced and approved for human use in a timely manner compared to vaccine development (Hiatt et al., 1989; Tian et al., 2020). The promise of plant expression systems for use in the fight against COVID-19 is discussed in reviews (Rosales-Mendoza, 2020; Shanmugaraj et al., 2020b).

General idea of the immune response to viral infection

The causative agents of respiratory diseases, which include various types of coronaviruses, enter the human body through the mucous membranes of the upper respiratory tract. Virus particles attach to cell receptors, fuse with the cell membrane, and enter the cell. Using the replicative apparatus of the cell, the virus multiplies and the viral particles come out, affecting the cells adjacent to it. In the case of the SARS-CoV-2 virus, penetration into the cell is provided by the S-2 protein, which is one of the two parts of the surface viral S protein (spike protein). The second part of this protein – S-1 provides binding to the ACE2 receptor of the lung epithelium. Having penetrated inside the cell, viruses become intracellular parasites, and the fight against them by the host's immune system becomes a difficult task.

Evolutionary, two systems of protecting the body from the penetration of pathogens have been formed – innate, immediately responding to danger, aimed at identifying the pathogen as a whole (innate immunity) and adaptive, aimed at identifying a huge number of specificities (antigens) in various pathogens (acquired immunity). The molecular mechanism of pathogen recognition is based on the detection of some standard "molecular marks" or pathogen-associated molecular patterns (PAMPs). Figure 2 shows the general scheme of the development of the body's immune responses to the penetration of the virus.

Innate immune responses are triggered at the first stage of interaction between the organism and the pathogen. Pathogen structures are recognized by the receptors of phagocytic cells and natural killers, upon interaction with which T-cell immune response cascades are launched and the elimination of pathogens and infected cells is coordinated. Toll-like and NOD-like receptors constitute the main group of receptors during the development of nonspecific protection (Takeuchi, Akira, 2010; Channappanavar et al., 2014).

It should be emphasized that in defending itself against viruses, the cell uses both antibodies (the humoral link of immunity) and the strategy of destroying cells infected with the virus (the cellular link of immunity). Membrane proteins of most viruses are "identification marks" or targets for the

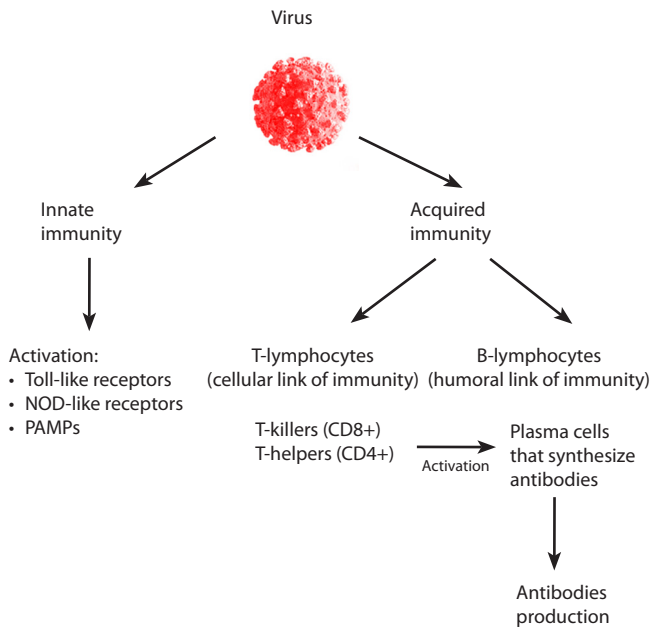


Fig. 2. Scheme of the immune response of the body to the invasion of a viral infection.

cell, on which B-lymphocytes activated by T-helpers (CD4+) differentiate into plasma cells that synthesize antibodies (see Fig. 2), which prevent the attachment and penetration of the virus into the cell. Such a defense strategy is effective in the early stages of infection, until the virus has entered the cell. After a cell is infected with a virus, another strategy is activated to destroy them, which is carried out by natural killers and cytotoxic T-lymphocytes (CD8+) (see Fig. 2). The importance of the formation of cytotoxic reactions in the fight against coronaviruses was emphasized earlier (Channappanavar et al., 2014). The scheme of immune responses of the body to the invasion of a viral infection, shown in Figure 2, is extremely simplified in order to draw attention to the key points that are important when choosing a strategy for developing a vaccine.

Principles for COVID-19 vaccine development

Modern knowledge in the field of molecular biology, immunology and vaccinology provides researchers with a wide range of methods and approaches for designing new generation vaccines based not only on data on the antigenic structure of the pathogen, but also on the mechanisms of the immune response to the pathogen and its components.

Nucleotide sequences of the SARS-CoV-2 virus genome are available on the websites of the National Library for Medicine and the Gene Bank (https://www.ncbi.nlm.nih.gov/sars-cov-2/). As of mid-July 2021, information on more than 377,000 fully read genomes of this virus, as well as more than 526,000 partially read genomes, is freely available. Since the outbreak of a new coronavirus infection COVID-19 occurred at the end of 2019 in China, in the city of Wuhan, the nucleotide sequence of this strain of the virus was condition-

ally chosen as a reference. Reference sequence data and all sequenced genomes are available from the gene bank at https://www.ncbi.nlm.nih.gov/nuccore/NC_045512.2.

At this point in time, on the WHO website, you can find information on the status of completed developments for the creation of vaccines against SARS-CoV-2 (https://cdn.who.int/media/docs/default-source/blue-print/15april2022-novel-covid-19-vaccine-tracker.zip?sfvrsn=225505e5_3&download=true). It should be emphasized that to date, 196 candidate vaccines at the stage of preclinical studies and 153 preparations at the stage of clinical trials have been registered in the world. On the WHO website via the same link, you can also find information about specific vaccine preparations and their manufacturers that are at the stage of clinical trials and receiving approval as vaccines from WHO (see the Table).

Analyzing the state of research in the development of vaccines against COVID-19, it should be noted that almost all creators use the immunogenic protein S of the coronavirus as a basis, which is presented to the immune system in different ways. It is this protein of the SARS-CoV-2 virus that binds to the ACE2 receptor of mucosal epithelial cells and ensures its penetration into the cells of the human body. As can be seen from the results of the analysis of the status of vaccines that have been registered and prepared by developers for use (see the Table), the current market for COVID-19 vaccines includes both classical vaccines based on the presentation of antigens of inactivated viruses to the immune system (Gao et al., 2020) and mRNA vaccines, in which the mRNA encoding the S protein is packaged in a lipid envelope. Such mRNA, when it enters human cells, is a template for the synthesis of the S protein, which is recognized by the cells of the immune system as a danger signal (Pardi et al., 2018). RNA vaccines have been shown to induce neutralizing antibodies with high titers (Jackson et al., 2020). Based on the same mechanism of antigen presentation, DNA vaccines are being developed that include a DNA fragment encoding the S protein into vectors, for example, into plasmids or adenoviruses (see the Table). In studies in rhesus monkeys, such vaccines stimulated the production of high antibody titers as well as the production of cytotoxic lymphocytes (Yu J. et al., 2020). The disadvantage of vector vaccines is the immunogenicity of the vectors themselves.

Vaccines based on recombinant proteins or peptides are considered promising. In the case of SARS-CoV-2, full-length or domains of S, M, and N proteins are considered as candidates for antigens, to increase the immunogenicity of which epitopes recognized by T- and B-cells of the immune system are additionally used (Marian, 2021). Such artificially created recombinant proteins, when they enter the body, activate the cells of the immune defense systems, which trigger the formation of the corresponding subcellular populations, the biosynthesis of antibodies, and the formation of "memory cells". The most complete strategy for creating vaccines against SARS-CoV-2 and the current state of research in this area are presented in the review (Bakhiet, Taurin, 2021).

It should be emphasized that in the development of antiviral vaccines, including against COVID-19, two important stages can be conditionally distinguished, the first of which is directly related to the creation of the vaccine itself, which is presented to the immune system either in the form of a large number of antigens (inactivated virus), or in the form of the dominant antigen(s) in the form of mRNA, DNA, recombinant protein or peptide. The importance of the second stage is determined by reliable systems for the production of either the virus itself or its antigens. Analyzing the current state of research in the development of vaccines and, in particular, subunit vaccines of a new generation, it should be noted that, along with well-established platforms, for example, Chinese hamster cells (CHO), used in the development of the recombinant vaccine 'Recombinant Novel Coronavirus Vaccine' by the Chinese company Zhifei Longcom (see the Table), plant expression systems are attracting increasing attention of the global research community (Fischer, Buyel, 2020; Kannan et al., 2020).

The state of research in the development of plant-based vaccines against COVID-19

Despite the fact that the first work on the attractiveness of plant expression systems for the biotechnological production of vaccinogenic proteins against COVID-19 appeared relatively recently (Rosales-Mendoza, 2020; Shanmugaraj et al., 2020a), by now the number of such works has increased markedly (Capell et al., 2020; Dhama et al., 2020; Ma et al., 2020; Prasad et al., 2020; Shanmugaraj et al., 2020a). The authors comprehensively discuss the possibilities of applying existing biotechnological developments to create subunit vaccines based on plant expression systems (Capell et al., 2020; Ma et al., 2020), as well as the features of creating this type of vaccine in case of a need for rapid response to the spread of a pathogen (Shanmugaraj et al., 2020a).

The researchers' developments on the creation of plant-based influenza vaccines based on virus-like particles also formed the basis for the development of vaccines against COVID-19 (Hodgins et al., 2019; Makarkov et al., 2019).

Moreover, the Canadian company 'Medicago', which uses plant expression systems for the production of recombinant proteins for medical purposes, used a transient expression system in *N. benthamiana* plants to develop a vaccinogenic protein including S-1 protein of the SARS-CoV-2 virus (https://www.medicago.com/en/covid-19-programs/). The company's developers used the fusion of a sequence encoding the viral protein S-1 with a sequence providing conformational transformations of a protein molecule that mimic the surface of a viral particle. The prospects of creating recombinant proteins conformationally folded in the form of virus-like particles, on the surface of which antigens are presented in the form of recombinant polypeptides, were noted earlier (Bai et al., 2008). Folding a recombinant protein in the form of a virus-like particle significantly increases the efficiency of antigen presentation to cells of the body's immune system (Rybicki, 2020). Despite the fact that the vaccine prepared on the basis of virus-like particles, although it mimics a virus, such an "artificial virus" lacks a genetic apparatus (RNA or DNA) and, accordingly, the ability to replicate when it enters a cell. Previously, numerous studies have shown that post-translational protein transformations in plant expression systems ensure its folding into a virus-like particle (D'Aoust

Status of COVID-19 vaccines within WHO EUL/PQ evaluation process 02.04.2022

Platform	Name of Vaccine	Manufacturer / WHO EUL holder	Status of assessment
Human Adenovirus Vector-based Covid-19 vaccine	AZD1222	AstraZeneca, UK	*
	Ad26.COV2.S	Janssen Infectious Diseases & Vaccinece (J&J), USA	*
	Sputnik-V	The FGB Institution "National Research Center for Epidemiology and Microbiology named after Honorary Academician N.F. Gamaleya of the Ministry of Health of the Russian Federation"	**
	Ad5-nCoV	SanSinoBIO, China	**
	Covishield (ChAdOx1_nCoV-19)	Serum Institute of India, India	*
Inactivated	SARS-CoV-2 Vaccine, Inactivated (Vero Cell) CoronavacTM	IMBCAMS, China	*
	Inactivated SARS-CoV-2 Vaccine (Vero Cell)	Sinopharm/WBIBP, China	**
	COVAXIN	Bharat Biotech, India	**
	SARS-CoV-2 Vaccine, Inactivated (Vero Cell)	IMBCAMS, China	**
	CovIran® vaccine	Iran	***
	SARS-CoV-2 Vaccine, Inactivated (Vero Cell)	IMBCAMS, China	***
Nucleoside modified mNRA	BNT162b2/COMIRNATY	Pfizer, USA	*
	mRNA-1273	Moderna, USA	*
	Zorecimeran	CureVac, Germany	***
Recombinant protein subunit	Recombinant Novel Coronavirus Vaccine (CHO Cell)	Zhifei Longcom, China	***
	NVX-CoV2373/Covovax	Novavax, USA	**
	CoV2 preS dTM-AS03 vaccine	Sanofi, France	***
	SCB-2019	Clover Biopharmaceuticals, China	***
	Soberana 01, Soberana 02 Soberana Plus Abdala	BioCubaFarma, Cuba	***
	Corbevax	Biological E, India	***
	GBP510	SK Bioscience, South Korea	***
Peptide antigen	EpiVacCorona	SRC VB "Vector", Russia	***
Plant-based expression systems	COVIFENZ®	Medicago Inc., Canada	***

Note. WHO website was used: https://extranet.who.int/pqweb/sites/default/files/documents/Status_COVID_VAX_02April2022.pdf
Status of assessment: * registered, ** end of registration, *** preparatory procedures.

et al., 2010; Lua et al., 2014). Moreover, it has been demonstrated that plant expression systems support the synthesis of functional recombinant proteins, including such complex ones as antibodies (Diamos et al., 2020).

The website of the biotech company Medicago Inc. published information on the approval by the Canadian regulator of a vaccine, which is a recombinant S protein of the SARS-CoV-2 virus, in the form of virus-like particles synthesized in tobacco plants (*N. benthamiana*) (https://medicago.com/app/uploads/2022/02 /Covifenz-PM-en.pdf). The vaccinogenic protein isolated and purified from plant biomass is used in tests on volunteers (about 30 thousand people participated in the experiment). It should be emphasized that the candidate vaccine obtained on the basis of the plant expression system has successfully passed three phases of clinical trials on volunteers (Pillet et al., 2019; Ward et al., 2021) and is now approved under the commercial name COVIFENZ® in Canada. The company announced the formation of high antibody titers in the subjects. Medicago Inc. has evaluated its candidate vaccine with GSK pandemic adjuvant. Currently, companies such as GlaxoSmithKlein (GSK, UK), Seqirus (UK) and Dynavax (USA) are developing licensed adjuvants (AS03, MF59 and CpG 1018, respectively) for their use with COVID-19 vaccines. The use of an adjuvant may be of particular importance in a pandemic situation, as it can reduce the amount of vaccine protein required per dose, which allows more doses of the vaccine to be produced and therefore contributes to the protection of more people. Although the exact dosage of the vaccine for humans has not yet been determined, the company estimates potential production volumes starting with 2021 to 80 million doses per year with an increase in productivity from 2023 to more than 1 billion doses of the COVID-19 vaccine per year.

Conclusion

As the experience of leading biotechnology companies and laboratories in optimizing expression systems for the production of recombinant proteins shows, plant expression systems are very attractive for these purposes and are already in demand by some large and medium-sized biotech companies. The prospect of using plant cells for the production of recombinant proteins intended for vaccine prophylaxis is also based on the possibility of their oral and intranasal administration and the activation of mucosal responses.

Oral delivery of pharmaceutical proteins appears to be a desirable target for the biopharmaceutical industry, as it provides more convenient drug administration than intravenous, intramuscular, and subcutaneous injections. Oral delivery will lead to better patient outcomes along with improved quality of life. Moreover, the attractiveness of plant expression systems is based on the ability to quickly respond to pathogens with a high degree of variability. The examples of successful testing of a plant vaccine against COVID-19 and the provision of large production volumes for the production of a vaccinogenic recombinant protein given in this review confirm the promise of plant expression systems for obtaining recombinant subunit vaccines.

References

Bai B., Hu Q., Hu H., Zhou P., Shi Z., Meng J., Lu B., Huang Y., Mao P., Wang H. Virus-like particles of SARS-like coronavirus formed by membrane proteins from different origins demonstrate stimulating activity in human dendritic cells. *PLoS One*. 2008;3:e2685. DOI 10.1371/journal.pone.0002685.

Bakhiet M., Taurin S. SARS-CoV-2: targeted managements and vaccine development. *Cytokine Growth Factor Rev*. 2021;58:16-29. DOI 10.1016/j.cytogfr.2020.11.001.

Capell T., Twyman R.M., Armario-Najera V., Ma J.K., Schillberg S., Christou P. Potential applications of plant biotechnology against SARS-CoV-2. *Trends Plant Sci.* 2020;25(7):635-643. DOI 10.1016/j.tplants.2020.04.009.

Channappanavar R., Fett C., Zhao J., Meyerholz D.K., Perlman S. Virus-specific memory CD8 T cells provide substantial protection from lethal severe acute respiratory syndrome coronavirus infection. *J. Virol.* 2014;88:11034-11044. DOI 10.1128/jvi.01505-14.

D'Aoust M.A., Couture M.M., Charland N., Trepanier S., Landry N., Ors F., Vézina L.-P. The production of hemagglutinin-based virus-like particles in plants: a rapid, efficient and safe response to pandemic influenza. *Plant Biotechnol. J.* 2010;8(5):607-619. DOI 10.1111/j.1467-7652.2009.00496.x.

Demurtas O.C., Massa S., Illiano E., De Martinis D., Chan P.K., Di Bonito P., Franconi R. Antigen production in plant to tackle infectious diseases flare up: the case of SARS. *Front. Plant Sci.* 2016;7:1-12. DOI 10.3389/fpls.2016.00054.

Dhama K., Natesan S., Yatoo M.I., Patel S.K., Tiwari R., Saxena S.K., Harapan H. Plant-based vaccines and antibodies to combat COVID-19: current status and prospects. *Hum. Vaccin. Immunother.* 2020;16(12):2913-2920. DOI 10.1080/21645515.2020.1842034.

Diamos A.G., Hunter J.G.L., Pardhe M.D., Rosenthal S.H., Sun H., Foster B.C., DiPalma M.P., Chen Q., Mason H.S. High level production of monoclonal antibodies using an optimized plant expression system. *Front. Bioeng. Biotechnol.* 2020;7:472. DOI 10.3389/fbioe.2019.00472.

Fischer R., Buyel J.F. Molecular farming – the slope of enlightenment. *Biotechnol. Adv*. 2020;40:107519. DOI 107519 10.1016/j.biotechadv.2020.107519.

Francis M.J. Recent advances in vaccine technologies. *Vet. Clin. Small Anim*. 2018;48(2):231-241. DOI 10.1016/j.cvsm.2017.10.002.

Gao Q., Bao L., Mao H., Wang L., Xu K., Yang M. Development of an inactivated vaccine candidate for SARS-CoV-2. *Science.* 2020;369(6499):77-81. DOI 10.1126/science.abc1932.

Gleba Y., Klimyuk V., Marillonnet S. Magnifection: a new platform for expressing recombinant vaccines in plants. *Vaccine.* 2005;23:2042-2048. DOI 10.1016/j.vaccine.2005.01.006.

Hiatt A., Cafferkey R., Bowdish K. Production of antibodies in transgenic plants. *Nature*. 1989;342(6245):76-78. DOI 10.1038/342076a0.

Hodgins B., Pillet S., Landry N., Ward B.J. Prime-pull vaccination with a plant-derived virus-like particle influenza vaccine elicits a broad immune response and protects aged mice from death and frailty after challenge. *Immun. Ageing*. 2019;16:27. DOI 10.1186/s12979-019-0167-6.

Jackson L.A., Anderson E.J., Rouphael N.G., Roberts P.C., Makhene M., Coler R.N., McCullough M.P., Chappell J.D., Denison M.R., Stevens L.J., Pruijssers A.J., McDermott A., Flach B., Doria-Rose N.A., Corbett K.S., Morabito K.M., O'Dell S., Schmidt S.D., Swanson P.A., Padilla M., Mascola J.R., Neuzil K.M., Bennett H., Sun W., Peters E., Makowski M., Albert J., Cross K., Buchanan W., Pikaart-Tautges R., Ledgerwood J.E., Graham B.S., Beigel J.H. An

mRNA vaccine against SARS-CoV-2 – preliminary report. *N. Engl. J. Med.* 2020;383:1920-1931. DOI 10.1056/nejmoa2022483.

Jiang S. Don't rush to deploy COVID-19 vaccines and drugs without sufficient safety guarantees. *Nature*. 2020;579:321. DOI 10.1038/d41586-020-00751-9.

Kannan S., Shaik Syed Ali P., Sheeza A., Hemalatha K. COVID-19 (Novel Coronavirus 2019) – recent trends. *European Review for Medical and Pharmacological Sciences*. 2020;24:2006-2011. DOI 10.26355/eurrev_202002_20378.

Kermode A.R. Current Status and Perspectives of the Molecular Farming Landscape. In: Kermode A.R. (Ed.) Molecular Pharming, First Edition. John Wiley & Sons, 2018;3-25. DOI 10.1002/9781118801512.ch10.

Lua L.H.L, Connors N.K., Sainsbury F., Chuan Y.P., Wibowo N., Middelberg A.P.J. Bioengineering virus-like particles as vaccines. *Biotechnol. Bioeng*. 2014;111(3):425-440. DOI 10.1002/bit.25159.79.

Ma C., Su S., Wang J., Wei L., Du L., Jiang S. From SARS-CoV to SARS-CoV-2: safety and broad-spectrum are important for coronavirus vaccine development. *Microbes Infect.* 2020;22(6-7):245-253. DOI 10.1016/j.micinf.2020.05.004.

Makarkov A.I., Golizeh M., Ruiz-Lancheros E., Gopal A.A., Costas-Cancelas I.N., Chierzi S., Pillet S., Charland N., Landry N., Rouiller I., Wiseman P.W., Ndao M., Ward B.J. Plant-derived virus-like particle vaccines drive cross-presentation of influenza A hemagglutinin peptides by human monocyte-derived macrophages. *NPJ Vaccines.* 2019;4:1-13. DOI 10.1038/s41541-019-0111-y.

Margolin E., Chapman R., Williamson A.L., Rybicki E.P., Meyers A.E. Production of complex viral glycoproteins in plants as vaccine immunogens. *Plant Biotechnol. J.* 2018;16(9):1531-1545. DOI 10.1111/pbi.12963.

Marian A.J. Current state of vaccine development and targeted therapies for COVID-19: impact of basic science discoveries. *Cardiovasc. Pathol*. 2021;50:107278. DOI 10.1016/j.carpath.2020.107278.

McNulty M.J., Gleba Y., Tusé D., Hahn-Löbmann S., Giritch A., Nandi S., McDonald K.A. Techno-economic analysis of a plant-based platform for manufacturing antimicrobial proteins for food safety. *Biotechnol. Prog*. 2020;36(1):e2896. DOI 10.1002/btpr.2896.

Pardi N., Hogan M.J., Porter F.W., Weissman D. mRNA vaccines – a new era in vaccinology. *Nat. Rev. Drug Discov.* 2018;17(4):261-279. DOI 10.1038/nrd.2017.243.

Phoolcharoen W., Bhoo S.H., Lai H., Ma J., Arntzen C.J., Chen Q., Mason H.S. Expression of an immunogenic Ebola immune complex in *Nicotiana benthamiana*. *Plant Biotechnol. J.* 2011;9(7):807-816. DOI 10.1111/j.1467-7652.2011.00593.x.

Pillet S., Couillard J., Trépanier S., Poulin J.F., Yassine-Diab B., Guy B., Ward B.J., Landry N. Immunogenicity and safety of a quadrivalent plant-derived virus like particle influenza vaccine candidate – two randomized Phase II clinical trials in 18 to 49 and ≥50 years old adults. *PLoS One.* 2019;14(6):e0216533. DOI 10.1371/journal.pone.0216533.

Pogrebnyak N., Golovkin M., Andrianov V., Spitsin S., Smirnov Y., Egolf R., Koprowski H. Severe acute respiratory syndrome (SARS) S protein production in plants: development of recombinant vaccine. *Proc. Natl. Acad. Sci. USA*. 2005;102(25):9062-9067. DOI 10.1073/pnas.0503760102.

Prasad A., Muthamilarasan M., Prasad M. Synergistic antiviral effects against SARS-CoV-2 by plant-based molecules. *Plant Cell Rep*. 2020;39(9):1109-1114. DOI 10.1007/s00299-020-02560-w.

Rosales-Mendoza S. Will plant-made biopharmaceuticals play a role in the fight against COVID-19? *Expert Opin. Biol. Ther.* 2020;20(6):545-548. DOI 10.1080/14712598.2020.1752177.

Rybicki E. History and promise of plant-made vaccines for animals. In: MacDonald J. (Ed.) Prospects of Plant-Based Vaccines in Veterinary Medicine. Cham: Springer, 2018;1-22. DOI 10.1007/978-3-319-90137-4_1.

Rybicki E.P. Plant molecular farming of virus-like nanoparticles as vaccines and reagents. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* 2020;12:1-22. DOI 10.1002/wnan.1587.

Sainsbury F., Sack M., Stadlmann J., Quendler H., Fischer R., Lomonossoff G.P. Rapid transient production in plants by replicating and non-replicating vectors yields high quality functional anti-HIVantibody. *PLoS One.* 2010;5(11):e13976. DOI 10.1371/journal.pone.0013976.

Salazar-González J.A., Bañuelos-Hernández B., Rosales-Mendoza S. Current status of viral expression systems in plants and perspectives for oral vaccines development. *Plant Mol. Biol*. 2015;87(3):203-217. DOI 10.1007/s11103-014-0279-5.

Shanmugaraj B., Malla A., Phoolcharoen W. Emergence of Novel Coronavirus 2019-nCoV: need for rapid vaccine and biologics development. *Pathogens*. 2020a;9(2):148. DOI 10.3390/pathogens9020148.

Shanmugaraj B., Siriwattananon K., Wangkanont K., Phoolcharoen W. Perspectives on monoclonal antibody therapy as potential therapeutic intervention for Coronavirus disease-19 (COVID-19). *Asian Pac. J. Allergy Immunol.* 2020b;38(1):10-18. DOI 10.12932/AP-200220-0773.

Shoji Y., Chichester J.A., Jones M., Manceva S.D., Damon E., Mett V., Musiychuk K., Bi H., Farrance C., Shamloul M., Kushnir N., Sharma S., Yusibov V. Plant-based rapid production of recombinant subunit hemagglutinin vaccines targeting H1N1 and H5N1 influenza. *Hum. Vaccine*. 2011;7(Suppl.):41. DOI 10.4161/hv.7.0.14561.

Takeuchi O., Akira S. Pattern recognition receptors and inflammation. *Cell.* 2010;140(6):805-820. DOI 10.1016/j.cell.2010.01.022.

Tian X., Li C., Huang A., Xia S., Lu S., Shi Z., Lu L., Jiang S., Yang Z., Wu Y., Ying T. Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human mono-clonal antibody. *Emerg. Microbes Infect.* 2020;9(1):382-385. DOI 10.1080/22221751.2020.1729069.

Waheed M.T., Ismail H., Gottschamel J., Mirza B., Lössl A.G. Plastids: the green frontiers for vaccine production. *Front. Plant Sci.* 2015;6:1-11. DOI 10.3389/fpls.2015.01005.

Ward B.J., Gobeil P., Séguin A., Atkins J., Boulay I., Charbonneau P.Y., Couture M., D'Aoust M.A., Dhaliwall J., Finkle C., Hager K., Mahmood A., Makarkov A., Cheng M.P., Pillet S., Schimke P., St-Martin S., Trépanier S., Landry N. Phase 1 randomized trial of a plant-derived virus-like particle vaccine for COVID-19. *Nat. Med.* 2021;27:1071-1078. DOI 10.1038/s41591-021-01370-1.

Ward B.J., Makarkov A., Séguin A., Pillet S., Trépanier S., Dhaliwall J., Libman M.D., Vesikari T., Landry N. Efficacy, immunogenicity, and safety of a plant-derived, quadrivalent, virus-like particle influenza vaccine in adults (18–64 years) and older adults (≥65 years): two multicentre, randomised phase 3 trials. *Lancet.* 2020;396(10261):1491-1503. DOI 10.1016/S0140-6736(20)32014-6.

Yu J., Tostanosk L.H., Peter L., Mercad N.B., McMahan K., Mahrokhia S.H., Nkolol J.P., Liu J., Li Z., Chandrashekar A., Martine D.R., Loos C., Atyeo C., Fischinger S., Burk J.S., Slei M.D., Chen Y., Zuiani A., Lelis F.J.N., Travers M., Habibi S., Pessaint L.,

Van Ry A., Blade K., Brown R., Cook A., Finneyfrock B., Dodson A., Teow E., Velasco J., Zahn R., Wegmann F., Bondzi E.A., Dagotto G., Gebr M.S., He X., Jacob-Dolan C., Kirilova M., Kordana N., Lin Z., Maxfiel L.F., Nampanya F., Nityanandam R., Ventur J.D., Wan H., Cai Y., Chen B., Schmid A.G., Weseman D.R., Bari R.S., Alter G., Andersen H., Lewi M.G., Barou D.H. DNA vaccine protection against SARS-CoV-2 in rhesus macaques. *Science.* 2020;369(6505):806-811. DOI 10.1126/science.abc6284.

Yu Y., Yu P.C., Chang W.J., Yu K., Lin C.S. Plastid transformation: how does it work? Can it be applied to crops? What can it offer? *Int. J. Mol. Sci.* 2020;21(14):4854. DOI 10.3390/ijms21144854.

Zhu N., Zhang D., Wang W., Li X., Yang B., Song J., Zhao X., Huang B., Shi W., Lu R., Niu P., Zhan F., Ma X., Wang D., Xu W., Wu G., Gao G.F., Tan W. A novel coronavirus from patients with pneumonia in China. *N. Engl. J. Med.* 2020;382:727-733. DOI 10.1056/NEJMoa2001017.

ORCID ID
E. Uvarova orcid.org/0000-0002-1490-4196

Acknowledgements. This research was funded by the Ministry of Science and Higher Education of the Russian Federation, grant number FWNR-2022-0022.
Conflict of interest. The authors declare no conflict of interest.
Received July 22, 2021. Revised November 8, 2021. Accepted November 10, 2021.

Прием статей через электронную редакцию на сайте http://vavilov.elpub.ru/index.php/jour
Предварительно нужно зарегистрироваться как автору, затем в правом верхнем углу страницы выбрать «Отправить рукопись». После завершения загрузки материалов обязательно выбрать опцию «Отправить письмо», в этом случае редакция автоматически будет уведомлена о получении новой рукописи.

«Вавиловский журнал генетики и селекции»/"Vavilov Journal of Genetics and Breeding"
до 2011 г. выходил под названием «Информационный вестник ВОГиС»/
"The Herald of Vavilov Society for Geneticists and Breeding Scientists".

Регистрационное свидетельство ПИ № ФС77-45870 выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций 20 июля 2011 г.

«Вавиловский журнал генетики и селекции» включен ВАК Минобрнауки России в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук, Российский индекс научного цитирования, ВИНИТИ, базы данных Emerging Sources Citation Index (Web of Science), Zoological Record (Web of Science), Scopus, PubMed Central, Ebsco, DOAJ, Ulrich's Periodicals Directory, Google Scholar, Russian Science Citation Index на платформе Web of Science, каталог научных ресурсов открытого доступа ROAD.

Открытый доступ к полным текстам:
на сайте ИЦиГ СО РАН – bionet.nsc.ru/vogis/
платформе Elpub – vavilov.elpub.ru/index.php/jour
платформе Научной электронной библиотеки – elibrary.ru/title_about.asp?id=32440
PubMed Central, https://www.ncbi.nlm.nih.gov/pmc/journals/3805/

Подписку на «Вавиловский журнал генетики и селекции» можно оформить в любом почтовом отделении России. Индекс издания 42153 по каталогу «Пресса России».

✉ e-mail: vavilov_journal@bionet.nsc.ru

Издатель: Федеральное государственное бюджетное научное учреждение
«Федеральный исследовательский центр Институт цитологии и генетики
Сибирского отделения Российской академии наук»,
проспект Академика Лаврентьева, 10, Новосибирск, 630090.
Адрес редакции: проспект Академика Лаврентьева, 10, Новосибирск, 630090.
Секретарь по организационным вопросам С.В. Зубова. Тел.: (383)3634977.
Издание подготовлено информационно-издательским отделом ИЦиГ СО РАН. Тел.: (383)3634963*5218.
Начальник отдела: Т.Ф. Чалкова. Редакторы: В.Д. Ахметова, И.Ю. Ануфриева. Дизайн: А.В. Харкевич.
Компьютерная графика и верстка: Т.Б. Коняхина, О.Н. Савватеева.

Подписано в печать 22.03.2022. Выход в свет 31.03.2022. Формат 60 × 84 $^{1}/_{8}$. Усл. печ. л. 12.32.
Уч.-изд. л. 14.6. Тираж 150 экз. (1-й завод 1–45 экз.) Заказ № 143. Цена свободная.

Отпечатано в Сибирском отделении РАН, Морской проспект, 2, Новосибирск, 630090.