VAVILOV JOURNAL OF GENETICS AND BREEDING

VAVILOVSKII ZHURNAL GENETIKI I SELEKTSII

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ
СОДЕРЖАНИЕ • 2021 • 25 • 5

VAVILOV JOURNAL OF GENETICS AND BREEDING

CONTENTS • 2021 • 25 • 5

The transcription factor dFOXO controls the expression of insulin pathway genes and lipids content under heat stress in *Drosophila melanogaster*

M.A. Eremina[1] ✉, P.N. Menshanov[1, 2], O.D. Shishkina[1], N.E. Gruntenko[1]

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

Abstract. The insulin/insulin-like growth factor signaling (IIS) pathway is one of the key elements in an organism's response to unfavourable conditions. The deep homology of this pathway and its evolutionary conservative role in controlling the carbohydrate and lipid metabolism make it possible to use *Drosophila melanogaster* for studying its functioning. To identify the properties of interaction of two key IIS pathway components under heat stress in *D. melanogaster* (the forkhead box O transcription factor (dFOXO) and insulin-like peptide 6 (DILP6), which intermediates the dFOXO signal sent from the fat body to the insulin-producing cells of the brain where DILPs1–5 are synthesized), we analysed the expression of the genes *dilp6*, *dfoxo* and insulin-like receptor gene (*dInR*) in females of strains carrying the hypomorphic mutation *dilp6⁴¹* and hypofunctional mutation *foxoᴮᴳ⁰¹⁰¹⁸*. We found that neither mutation influenced *dfoxo* expression and its uprise under short-term heat stress, but both of them disrupted the stress response of the *dilp6* and *dInR* genes. To reveal the role of identified disruptions in metabolism control and feeding behaviour, we analysed the effect of the *dilp6⁴¹* and *foxoᴮᴳ⁰¹⁰¹⁸* mutations on total lipids content and capillary feeding intensity in imago under normal conditions and under short-term heat stress. Both mutations caused an increase in these parameters under normal conditions and prevented decrease in total lipids content following heat stress observed in the control strain. In mutants, feeding intensity was increased under normal conditions; and decreased following short-term heat stress in all studied strains for the first 24 h of observation, and in *dilp6⁴¹* strain, for 48 h. Thus, we may conclude that dFOXO takes part in regulating the IIS pathway response to heat stress as well as the changes in lipids content caused by heat stress, and this regulation is mediated by DILP6. At the same time, the feeding behaviour of imago might be controlled by dFOXO and DILP6 under normal conditions, but not under heat stress.

Key words: *Drosophila melanogaster*; insulin/insulin-like growth factors signaling pathway; *dInR*; *dilp6*; *dfoxo*; gene expression; feeding behaviour; total lipids.

Транскрипционный фактор dFOXO регулирует экспрессию генов инсулинового сигнального каскада и содержание липидов при тепловом стрессе у *Drosophila melanogaster*

М.А. Еремина[1] ✉, П.Н. Меньшанов[1, 2], О.Д. Шишкина[1], Н.Е. Грунтенко[1]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук,
  Новосибирск, Россия
[2] Новосибирский государственный технический университет, Новосибирск, Россия
✉ eremina@bionet.nsc.ru

Аннотация. Одним из основных элементов ответа организма на неблагоприятные условия является сигнальный каскад инсулина/инсулиноподобных факторов роста (И/ИФР). Благодаря глубокой гомологии этого каскада и эволюционной консервативности его роли в регуляции углеводно-жирового метаболизма, возможно использование модельного объекта *Drosophila melanogaster* для изучения механизмов его функционирования. Для определения особенностей взаимодействия двух ключевых компонентов каскада И/ИФР у *D. melanogaster* – транскрипционного фактора dFOXO и инсулиноподобного пептида DILP6, «посредника» в передаче сигнала от dFOXO в жировом теле к инсулин-продуцирующим клеткам мозга (месту синтеза DILPs1–5), в условиях теплового стресса мы провели анализ экспрессии генов *dilp6*, *dfoxo* и гена инсулинопо-

добного рецептора (*dlnR*) у самок линий, несущих гипоморфную мутацию *dilp6⁴¹* и гипофункциональную мутацию *foxo^BG01018*. Обнаружено, что обе мутации не оказывали влияния на экспрессию *dfoxo* и ее повышение при кратковременном тепловом стрессе, однако нарушали ответ на стресс генов *dilp6* и *dlnR*. Для выявления роли обнаруженных нарушений в контроле метаболизма и метаболического поведения мы проанализировали влияние мутаций *dilp6⁴¹* и *foxo^BG01018* на содержание общих липидов и интенсивность капиллярного питания имаго в нормальных условиях и при кратковременном тепловом стрессе. Обе мутации приводили к усилению данных признаков в нормальных условиях и препятствовали снижению содержания общих липидов после стресса, наблюдаемому у контрольной линии. Интенсивность питания была повышена у мутантов в нормальных условиях и снижалась после кратковременного теплового стресса у всех изученных линий в течение первых суток наблюдения, а у линии *dilp6⁴¹* – в течение двух суток. Таким образом, можно заключить, что dFOXO принимает участие в регуляции как ответа сигнального каскада И/ИФР на тепловой стресс, так и вызываемых тепловым стрессом изменений в содержании липидов, причем эта регуляция опосредуется DILP6. В то же время метаболическое поведение имаго, по-видимому, регулируется dFOXO и DILP6 в нормальных условиях, но не при тепловом стрессе.

Ключевые слова: *Drosophila melanogaster*; сигнальный каскад инсулина/инсулиноподобных факторов роста; *dlnR*; *dilp6*; *dfoxo*; экспрессия генов; пищевое поведение; общие липиды.

Introduction

Nowadays, as living beings often encounter unfavourable environmental conditions such as pollution and global warming, the study of deeply conservative mechanisms that contribute to adaptation is of current interest. It is known that such influences launch the development of nonspecific adaptive defensive responses on molecular (Garbuz, Evgen'ev, 2017), behavioral (Kaluev, 1999), biochemical and physiological (Gruntenko, 2008; Even et al., 2012; Miyashita, Adamo, 2020) levels. The ability to respond to stress in an integrated manner, which comprises behavioral, metabolic and molecular reactions, is key for survival and adaptation of animals including insects (Koyama et al., 2020). It was proven that besides its role as crucial modulator of growth and metabolism, in insects, the IIS pathway is an essential component of the neuroendocrine stress reaction (Gruntenko, Rauschenbach, 2018; Lubawy et al., 2020). Due to the deep homology of this pathway in animals of different taxa including humans and flies, it is possible to use the latter as an object for investigating evolutionary-conservative mechanisms underlying molecular-genetic regulation of the IIS pathway, and carbohydrate and lipid metabolism it controls. As in most animals, in insects, carbohydrates and lipids serve as the main energy supply (Arrese, Soulages, 2010). The processes of producing and storing energy undergo complex modulation by many inner factors including heritage, lifestyle, hormones, metabolites, as well as various outside influences (Mattila, Hietakangas, 2017).

Drosophila's applicability to the research of metabolism is defined by the simplicity of its IIS pathway regulation (Fig. 1), which involves homologues of insulin (DILPs1–5) and insulin-like growth factors (DILP6) of mammals connecting to a single insulin-like receptor (dInR), which activates the pathway (Gruntenko, Rauschenbach, 2018), and two homologues of relaxin (DILPs7,8) (Gontijo, Garelli, 2018). The dInR signal being transduced directly or *via* its substrate CHICO (the homologue of insulin receptor substrates of mammals, IRS1–4) causes dAkt/PKB (protein kinase B homologue) to activate, which in turn modulates the activity of a number of proteins, in particular, it phosphorylates transcriptional factor of Forkhead box class O family, dFOXO (homologue of mammalian FOXO), which is synthesized in the fat body and controls the transcription of more than a thousand genes (Bai et al., 2012), and inhibits its translocation into the nucleus (Puig et al., 2003; Slack et al., 2011; Álvarez-Rendón et al., 2018). Under stress, dFOXO is translocated to the nucleus (Jünger et al., 2003; Hwangbo et al., 2004; Gruntenko et al., 2016) activating the expression of a number of genes including *dInR via* a feedback loop (Gruntenko, Rauschenbach, 2018). It was also previously shown that the expression of *dilp6* in the fat body inhibits the expression of *dilp2* and *dilp5* in imago's brain as well as the secretion of DILP2 into the hemolymph, and that dFOXO influence on the expression of DILPs produced in median neurosecretory cells is mediated by DILP6 synthesized in the fat body (Slaidina et al., 2009; Bai et al., 2012). Thus, DILPs seems to connect dFOXO, adipose tissue and endocrine function of the brain, creating a feedback loop back to dInR.

Stress reaction causes the mobilization of organism's energy reserves along with a variety of metabolic changes. In a changing environment, feeding behaviour plays an important role in adaptation (Rabasa, Dickson, 2016). It is known that in mammals, acute stress is usually accompanied by feeding suppression and a decrease in weight gain; chronic stress can result in excessive food intake, weight gain and obesity (Rabasa, Dickson, 2016).

This study aimed to analyse the expression of *dInR*, *dilp6* and *dfoxo* genes of three key components of the IIS pathway, which is involved in neuroendocrine stress reaction, in *D. melanogaster* strains carrying *dilp6⁴¹* и *foxo^BG01018* mutations under heat stress, and to evaluate the latter's influence on feeding behaviour and total lipids content in these strains.

Materials and methods

Drosophila melanogaster strains and stress conditions. Three *D. melanogaster* strains were used in this study: strain *dilp6⁴¹* with the deletion covering the 3′ region of *phl* gene and 5′ upstream region of *dilp6* including the first exon and part of the first intron (Rauschenbach et al., 2017); strain

М.А. Еремина, П.Н. Меньшанов
О.Д. Шишкина, Н.Е. Грунтенко

dFOXO регулирует экспрессию генов инсулинового
сигнального каскада и содержание липидов при стрессе

2021
25•5
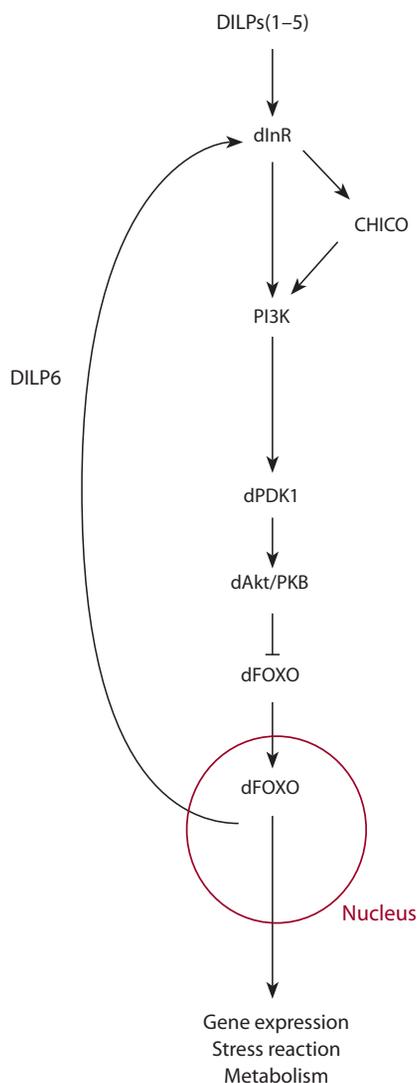
Fig. 1. The scheme of insulin/insulin-like growth factors signaling pathway in *Drosophila*.

DILPs are *Drosophila* insulin-like growth peptides, dInR – *Drosophila* insulin-like receptor, CHICO – homologue of mammalian insulin receptor substrate, PI3K – phosphoinositide 3 kinase, dPDK1 – *Drosophila* phosphoinositide-dependent kinase-1, dAkt/PKB – a homolog of mammalian protein kinase B, dFOXO – *Drosophila* forkhead box O transcription factor.

foxo^{BG01018}, which carries a P[GT1] element transposon in the 5′ upstream region of the *dfoxo* gene, resulting in a mild loss of function (Dionne et al., 2006); and their progenitor strain *w*^{1118} as a control. The stocks were obtained from the Bloomington Drosophila Stock Center (Bloomington, IN, USA).

The cultures were raised on standard medium (agar-agar, 7 g/l; corn grits, 50 g/l; dry yeast, 18 g/l; sugar, 40 g/l) and kept at 25 °C, 12:12 h photoperiod, relative humidity 50 %. Imagoes were synchronised at eclosion (flies were collected every 3–4 hours). Females were exposed to heat stress by transferring vials with flies from a 25 °C incubator to a 38 °C incubator for 60 or 90 min. After 60 min of stressing flies were returned to 25 °C, after 90 min they were subsequently frozen in liquid nitrogen and stored at –80 °C.

Quantitative real-time polymerase chain reaction (qRT-PCR). mRNA quantity of *dilp6*, *dfoxo* and *dInR* genes was evaluated in whole body homogenates of Canton-S females (15 flies/sample) using TRI reagent Lot #BCBT8883 (Sigma-Aldrich, USA) for total RNA extraction, Revert Aid First Strand cDNA Synthesis Kit #K1621 (Thermo Fisher Scientific, USA) with oligo (dT)18 primer for synthesis of cDNA, M-427 Kit with SYBR-Green I (Syntol, Russia) and CFX96 Touch qPCR System (Bio-Rad, USA) for performing qRT-PCR. Each reaction was performed in triplicates with three biological replicates. Data were normalized against *Act5C*. High stability of *Act5C* expression under heat stress was shown by Ponton et al. (2011). The primers used in the study are shown in the Table.

Total lipid quantification. Quantification of total lipids was performed using Van Handel's method (1985) modified for *D. melanogaster* (Eremina, Gruntenko, 2020) under normal conditions or in 24 h after 60 min under 38 °C. Flies (1 fly per sample, 10–20 samples per each studied group) were decapitated to avoid the influence of eye pigment on the measurement results, homogenised on ice in 100 μl of chloroform-methanol (1:1) and shacked for 10 min. 50 μl of supernatant were transferred to new tubes and placed in microthermostat M-208 (Bis-N, Russia) at 90 °C till the solvent completely evaporated. Then 10 μl of 95 % H_2SO_4 were added to each sample and they were again kept at 90 °C for 2 min. After

Primers used in RT-PCR

Gene	Amplicon, bp	Forward/Reverse	Sequence (5′–3′)	Tm, °C	Reference
dfoxo	196	F	GCCTAGATCACTTTCCCGAG	53	Gruntenko et al., 2016
		R	GTCAGCTCATCCGCCATTGT	55	
dilp6	149	F	CACGGAATACGAACAGAGACG	55	Eremina et al., 2019
		R	TCGGTTACGTTCTGCAAGTC	55	
dInR	123	F	TGAGCATGTGGAGCACATCAAGATG	59	Okamoto et al., 2013
		R	CGTAGGAGATTTTCTCGTTTGGCTG	58	
Act5C	90	F	GCGCCCTTACTCTTTCACCA	58	Guio et al., 2014
		R	ATGTCACGGACGATTTCACG	55	

that the samples were cooled on ice and the phosphovanillin color reagent (85 % H_3PO_4 + 6 % vanillin solution (4:1)) was added up to 1 ml of volume. The samples were incubated for 15 min at room temperature till pink colouration appeared and was stable for 1 h. Then the samples were measured by Smart Spec Plus spectrophotometer (Bio-Rad, USA) at 525 nm.

Feeding behaviour analysis (CAFE). Ingestion was measured using the Capillary Feeder (CAFÉ) method of Ja et al. (2007), modified by Williams et al. (2014). To provide flies with a humid environment, flat-bottomed glass vials (20 × 100 mm) with 1 % agarose (5 cm high) were placed into microcentrifuged 50 ml tubes filled with 7 ml of water. Each glass capillary (10 × 90 mm, Narishige, Japan) was filled with 20 µl of liquid food containing 5 % sugar and 5 % east extract (Biospringer, France). Five females were placed into each vial (4–9 vials per group), which was plugged with a foam plug. A capillary was inserted into it through 10 µl and 200 µl pipette tips and was held in place by them. The vials with flies were kept in an incubator (Sanyo, Japan) at 25 °C, 50 % relative humidity, 12:12 h photoperiod for 24 or 48 h. Before that the experimental group was subjected to short-term heat stress (38 °C, 60 min). Initial and final food levels in capillaries were marked to determine total food consumption per day. To minimize food evaporation, capillaries were topped with a 0.1 µl oil layer. To adjust for food evaporation, a vial without flies was used.

Statistical analysis. Data on gene expression were analyzed by the $2^{-\Delta\Delta CT}$ method (Livak, Schmittgen, 2001). All data are presented as means ± SEM and analysed by ANOVA. The results were considered significant at $p < 0.05$.

Results and discussion

To discover whether disruption of the feedback loop of the IIS pathway regulation affects its stress response, we studied the expression of three key genes of the pathway, *dilp6*, *dfoxo* and *dInR*, in *D. melanogaster* females carrying hypomorphic mutation *dilp6⁴¹* and hypofunctional mutation *foxo^{BG01018}* under normal conditions or heat stress (38 °C, 90 min). There were no quantitative changes in mRNA expression level of *dilp6* and *dInR* genes in *dilp6⁴¹* and *foxo^{BG01018}* strains under heat stress, whereas in their progenitor strain *w^{1118}* the expression of *dilp6* decreased, and the expression of *dInR* increased under heat stress (Fig. 2, $p < 0.05$ for both genes). At the same, *dfoxo* expression level increased or had a tendency to increase under heat stress in all strains under study (see Fig. 2, STRAIN – $F_{(2, 12)}$ = 3.14, $p < 0.081$; STRESS – $F_{(1, 12)}$ = 12.80, $p < 0.0038$). Notably, *dilp6⁴¹* mutants are characterised by a lower *dilp6* expression ($p < 0.001$); however, *dfoxo* expression in *foxo^{BG01018}* mutants does not differ from the control strain *w^{1118}* (see Fig. 2). This allows us to assume that the previously described loss of dFOXO function in *foxo^{BG01018}* strain (Dionne et al., 2006) is connected not with a lowered expression level of the corresponding gene but with a defect in its structure.

The results of qualitative measurement of total lipids in *D. melanogaster* females with *dilp6⁴¹* and *foxo^{BG01018}* mutations under normal conditions or following heat stress (38 °C, 60 min) signify that both mutations cause an increase in lipid content in comparison with the control strain *w^{1118}*, and lipid content in *dilp6⁴¹* and *foxo^{BG01018}* strains, unlike in their progenitor strain, does not decrease in 24 h after heat stress
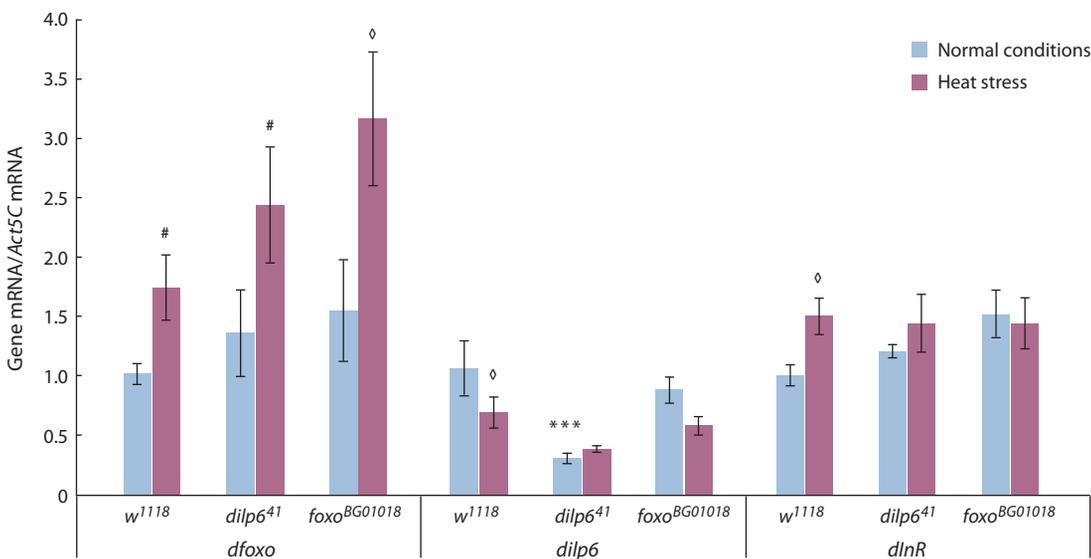
Fig. 2. *dilp6*, *dfoxo* and *dInR* mRNA levels in *D. melanogaster* females of *w^{1118}*, *dilp6⁴¹* and *foxo^{BG01018}* strains under normal conditions and after short-term heat stress (38 °C, 90 min).

Each value is a mean of three biological replicates. Error bars show standard error of the mean. Asterisks indicate significant differences between females with mutation of Drosophila insulin-like peptide 6 gene (*dilp6⁴¹*) and females of the control strain *w^{1118}* ($p < 0.001$). Diamond indicates significant differences between stressed and control groups of the same genotype ($p < 0.05$). Hash indicates a tendency for such differences ($p < 0.07$).

М.А. Еремина, П.Н. Меньшанов
О.Д. Шишкина, Н.Е. Грунтенко

dFOXO регулирует экспрессию генов инсулинового
сигнального каскада и содержание липидов при стрессе

2021
25·5
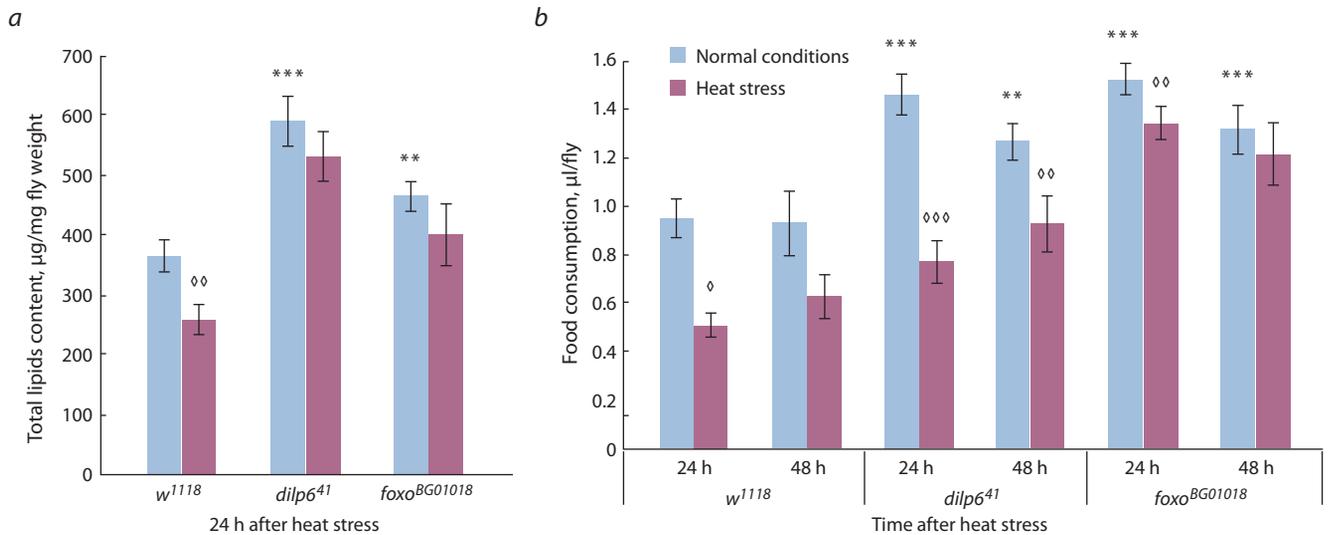
Fig. 3. Total lipids level (*a*) and capillary feeding intensity (*b*) in females of *D. melanogaster* strains *w¹¹¹⁸*, *dilp6⁴¹* and *foxo^BG01018* under normal conditions and following short-term heat stress (38 °C, 60 min).

Each value is a mean of 10–20 (*a*) and 9–11 (*b*) measurements. Error bars indicate s.e.m. Asterisk indicates significant differences between control females of *w¹¹¹⁸* strain and females with *dilp6⁴¹* and *foxo^BG01018* mutations (** $p < 0.01$, *** $p < 0.001$). Diamond indicates significant differences between stressed and control groups of the same genotype (◊ – $p < 0.05$, ◊◊ – $p < 0.01$, ◊◊◊ – $p < 0.001$).

(Fig. 3, *a*, STRAIN – $F_{(1, 96)} = 26.78$, $p \ll 0.0001$; STRESS – $F_{(1, 96)} = 141.56$, $p < 0.012$; STRAIN*STRESS – $F_{(2, 96)} = 0.25$, $p = 0.777$).

The increased lipid content in females of the mutant strains could be explained by their discovered increased food consumption in comparison with control females of *w¹¹¹⁸* strain throughout the entire experiment (see Fig. 3, *b*, STRAIN – $F_{(2, 59)} = 44.40$, $p \ll 0.0001$; TIME – $F_{(1, 59)} = 5.12$, $p < 0.028$; STRAIN*TIME – $F_{(2, 59)} = 1.41$, $p = 0.252$). However, in the first 24 h after heat stress feeding intensity decreases in comparison with normal conditions in both females of the control strain *w¹¹¹⁸* and the mutant strains; in *dilp6⁴¹* strain, this effect is maintained for 48 h (see Fig. 3, *b*, STRESS – $F_{(1, 59)} = 36.09$, $p \ll 0.0001$; STRAIN*STRESS – $F_{(2, 59)} = 6.28$, $p < 0.0034$; STRAIN*STRESS*TIME – $F_{(2, 59)} = 1.26$, $p = 0.291$).

It was previously shown that the IIS pathway can interact with gonadotropins and biogenic amines in *Drosophila* modulating their dynamics under stress and thus participating in the control of organism's stress response (Gruntenko, Rauschenbach, 2018). However, it remained unclear (1) which links of the IIS pathway were involved in stress response and (2) what effect does the participation of the IIS pathway in stress response have on its ability to control the carbohydrate and lipid metabolism.

It was demonstrated by us earlier that in *D. melanogaster* females dFOXO translocates to the nucleus under heat stress (Gruntenko et al., 2016), and here we showed that this translocation is accompanied by a tendency to an increase of *dfoxo* expression (see Fig. 2). Our data also allow us to assume that dFOXO activation under stress results in *dilp6* being inhibited as the decrease of *dilp6* expression found in the control strain *w¹¹¹⁸* is not observed in *foxo^BG01018* mutants (see Fig. 2).

dilp6 expression in the fat body was previously shown to suppress *dilp2* and *dilp5* expression in imago's brain and the secretion of DILP2 into hemolymph; dFOXO's influence on the expression of DILPs produced in median neurosecretory cells is inhibited by a simultaneous repression of DILP6 in the fat body *via* RNA interference (Bai et al., 2012). This allow us to suppose that a decrease in DILP6 activity under heat stress leads to an increase in level of DILPs expressed in median neurosecretory cells of the brain. Indeed, we were able to demonstrate earlier that DILP3 synthesis in these cells is increased in response to heat stress in wild type flies (Andreenkova et al., 2018), and in *dilp6⁴¹* larvae – under normal conditions (Andreenkova et al., 2017), which corresponds well with our assumption about a signal being transmitted from dFOXO to DILP3 through DILP6 under heat stress. Then, DILP3 appears to activate dInR, inhibiting the IIS pathway, which is confirmed by our data on the lack of a shift in *dInR* expression level under heat stress in flies with mutations of *dilp6* and *dfoxo* genes as opposed to the shift in laboratory strain *w¹¹¹⁸*, in which a decrease in *dilp6* and an increase in *dInR* expression is shown to occur in response to heat stress (see Fig. 2).

System defects in the IIS pathway cause *D. melanogaster* to manifest a number of different phenotypes including those connected to metabolism, which usually involves an increase in organism's carbohydrates and lipids reserves (Mattila, Hietakangas, 2017). Murillo-Maldonado et al. (2011) demonstrated almost all viable combinations of mutations with partial loss of function or hypomorphism of IIS genes to have changes in carbohydrates and lipids levels. Slaidina et al. (2009) showed *dilp6* knockdown to cause an increased level of triglycerides and glycogen in *Drosophila* larvae.

These results correspond well with our data on the increased content of total lipids in females of *dilp6⁴¹* and *foxo^BG01018* strains (see Fig. 3, *a*), as well as with increased glucose and trehalose levels in *dilp6⁴¹* and *foxo^BG01018* mutants we previously demonstrated (Eremina et al., 2019).

Regarding regulation of feeding behaviour under heat stress it seems to occur independently from *dilp6* and *dfoxo* genes as their mutations do not inhibit loss of appetite following stress (see Fig. 3, *b*).

Conclusion

Thus, we have shown that the disruption of *dilp6* and *dfoxo* gene functions in *Drosophila melanogaster* (1) results in the feedback loop of the IIS pathway being disrupted under heat stress, (2) leads to an increase in total lipids content under normal conditions and impedes their decrease following heat stress, and (3) causes an increase in feeding intensity under normal conditions but does not impede its decrease following heat stress.

References

Álvarez-Rendón J.P., Salceda R., Riesgo-Escovar J.R. *Drosophila melanogaster* as a model for diabetes type 2 progression. *Biomed. Res. Int.* 2018;2018:1417528. DOI 10.1155/2018/1417528.

Andreenkova O.V., Eremina M.A., Gruntenko N.E., Rauschenbach I.Y. Effect of heat stress on expression of DILP2 and DILP3 insulin-like peptide genes in *Drosophila melanogaster* adults. *Russ. J. Genet.* 2018;54(3):363-365. DOI 10.1134/S102279541803002X.

Andreenkova O.V., Rauschenbach I.Y., Gruntenko N.E. Hypomorphic mutation of the *dilp6* gene increases DILP3 expression in insulin-producing cells of *Drosophila melanogaster*. *Russ. J. Genet.* 2017; 53(10):1159-1161. DOI 10.1134/S1022795417080026.

Arrese E.L., Soulages J.L. Insect fat body: energy, metabolism, and regulation. *Annu. Rev. Entomol.* 2010;55:207-225. DOI 10.1146/annurev-ento-112408-085356.

Bai H., Kang P., Tatar M. *Drosophila* insulin-like peptide-6 (*dilp6*) expression from fat body extends lifespan and represses secretion of *Drosophila* insulin-like peptide-2 from the brain. *Aging Cell.* 2012;11(6):978-985. DOI 10.1111/acel.12000.

Dionne M.S., Pham L.N., Shirasu-Hiza M., Schneider D.S. *Akt* and *foxo* dysregulation contribute to infection-induced wasting in *Drosophila*. *Curr. Biol.* 2006;16(20):1977-1985. DOI 10.1016/j.cub.2006.08.052.

Eremina M.A., Gruntenko N.E. Adaptation of the sulfophosphovanillin method of analysis of total lipids for various biological objects as exemplified by *Drosophila melanogaster*. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2020; 24(4):441-445. DOI 10.18699/VJ20.636. (in Russian)

Eremina M.A., Karpova E.K., Rauschenbach I.Yu., Pirozhkova D.S., Andreenkova O.V., Gruntenko N.E. Mutations in the insulin signaling pathway genes affect carbohydrate level under heat stress in *Drosophila melanogaster* females. *Russ. J. Genet.* 2019; 55(4):519-521. DOI 10.1134/S1022795419030050.

Even N., Devaud J.M., Barron A.B. General stress responses in the honey bee insects. *Insects.* 2012;3(4):1271-1298. DOI 10.3390/insects3041271.

Garbuz D.G., Evgen'ev M.B. The evolution of heat shock genes and expression patterns of heat shock proteins in the species from temperature contrasting habitats. *Russ. J. Genet.* 2017;53(1):21-38. DOI 10.1134/S1022795417010069.

Gontijo A.M., Garelli A. The biology and evolution of the Dilp8-Lgr3 pathway: a relaxin-like pathway coupling tissue growth and developmental timing control. *Mech. Dev.* 2018;154:44-50. DOI 10.1016/j.mod.2018.04.005.

Gruntenko N.E. Stress and Reproduction of Insects: Hormonal Control. Novosibirsk; Moscow: KMK Publ., 2008. (in Russian)

Gruntenko N.E., Adonyeva N.V., Burdina E.V., Karpova E.K., Andreenkova O.V., Gladkikh D.V., Ilinsky Y.Y., Rauschenbach I.Yu. The impact of FOXO on dopamine and octopamine metabolism in *Drosophila* under normal and heat stress conditions. *Biol. Open.* 2016; 5(11):1706-1711. DOI 10.1242/bio.022038.

Gruntenko N.E., Rauschenbach I.Yu. The role of insulin signalling in the endocrine stress response in *Drosophila melanogaster*: a mini-review. *Gen. Comp. Endocrinol.* 2018;258:134-139. DOI 10.1016/j.ygcen.2017.05.019.

Guio L., Barron M.G., Gonzalez J. The transposable element *Bari-Jheh* mediates oxidative stress response in *Drosophila. Mol. Ecol.* 2014;23:2020-2030.

Hwangbo D.S., Gershman B., Tu M.P., Palmer M., Tatar M. *Drosophila* dFOXO controls lifespan and regulates insulin signalling in brain and fat body. *Nature.* 2004;429(6991):562-566. DOI 10.1038/nature02549.

Ja W.W., Carvalho G.B., Mak E.M., de la Rosa N.N., Fang A.Y., Liong J.C., Brummel T., Benzer S. Prandiology of *Drosophila* and the CAFE assay. *Proc. Natl. Acad. Sci. USA.* 2007;104(20):8253-8256.10.1073/pnas.0702726104.

Jünger M.A., Rintelen F., Stocker H., Wasserman J.D., Végh M., Radimerski T., Greenberg M.E., Hafen E. The *Drosophila* forkhead transcription factor FOXO mediates the reduction in cell number associated with reduced insulin signaling. *J. Biol. Chem.* 2003;2(3):20. DOI 10.1186/1475-4924-2-20.

Kaluev A.V. Problems of Studying Stressful Behavior, Kiev: KSF Publ., 1999. (in Russian)

Koyama T., Texada M.J., Halberg K.A., Rewitz K. Metabolism and growth adaptation to environmental conditions in *Drosophila*. *Cell. Mol. Life Sci.* 2020;77:4523-4551.

Livak K.J., Schmittgen T.D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods.* 2001;25(4):402-408. DOI 10.1006/meth.2001.1262.

Lubawy J., Urbanski A., Colinet H., Pfluger H.-J., Marciniak P. Role of the insect neuroendocrine system in the response to cold stress. *Front. Physiol.* 2020;11:376. DOI 0.3389/fphys.2020.00376.

Mattila J., Hietakangas V. Regulation of carbohydrate energy metabolism in *Drosophila melanogaster*. *Genetics.* 2017;207(4):1231-1253. DOI 10.1534/genetics.117.199885.

Miyashita A., Adamo S.A. Stayin' alive: Endocrinological stress responses in insects. In: Saleuddin S., Lange A., Orchard I. (Eds.). Advances in Invertebrate Endocrinology. Toronto: Apple Acad. Press, 2020;283-325.

Murillo-Maldonado J.M., Sánchez-Chávez G., Salgado L.M., Salceda R., Riesgo-Escovar J.R. *Drosophila* insulin pathway mutants affect visual physiology and brain function besides growth, lipid, and carbohydrate metabolism. *Diabetes.* 2011;60(5):1632-1636. DOI 10.2337/db10-1288.

Okamoto N., Nakamori R., Murai T., Yamauchi Y., Masuda A., Nishimura T. A secreted decoy of InR antagonizes insulin/IGF signaling to restrict body growth in *Drosophila*. *Genes Dev.* 2013;27(1):87-97. DOI 10.1101/gad.204479.112.

Ponton F., Chapuis M.-P., Pernice M., Sword G.A., Simpson S.J. Evaluation of potential reference genes for reverse transcription-qPCR studies of physiological responses in *Drosophila melanogaster*. *J. Insect Physiol.* 2011;57:840-850. DOI 10.1016/j.jinsphys.2011.03.014.

Puig O., Marr M.T., Ruhf M.L., Tjian R. Control of cell number by *Drosophila* FOXO: downstream and feedback regulation of the insulin receptor pathway. *Genes Dev.* 2003;17(16):2006-2020. DOI 10.1101/gad.1098703.

М.А. Еремина, П.Н. Меньшанов
О.Д. Шишкина, Н.Е. Грунтенко

dFOXO регулирует экспрессию генов инсулинового
сигнального каскада и содержание липидов при стрессе

2021
25•5

Rabasa C., Dickson S.L. Impact of stress on metabolism and energy balance. *Curr. Opin. Behav. Sci.* 2016;9:71-77. DOI 10.1016/j.cobeha.2016.01.011.

Rauschenbach I.Y., Karpova E.K., Burdina E.V., Adonyeva N.V., Bykov R.A., Ilinsky Y.Y., Menshanov P.N., Gruntenko N.E. Insulin-like peptide DILP6 regulates juvenile hormone and dopamine metabolism in *Drosophila* females. *Gen. Comp. Endocrinol.* 2017; 243:1-9. DOI 10.1016/j.ygcen.2016.11.004.

Slack C., Giannakou M.E., Foley A., Goss M., Partridge L. dFOXO-independent effects of reduced insulin-like signaling in *Drosophila*.

Aging Cell. 2011;10(5):735-748. DOI 10.1111/j.1474-9726.2011.00707.x.

Slaidina M., Delanoue R., Grönke S., Partridge L., Leopold P. A *Drosophila* insulin-like peptide promotes growth during nonfeeding states. *Dev. Cell.* 2009;17(6):874-884. DOI 10.1016/j.devcel.2009.10.009.

Van Handel E. Rapid determination of total lipids in mosquitoes. *J. Am. Mosq. Control Assoc.* 1985;1:302-304.

Williams M.J., Wang Yi., Klockars A., Lind P.M., Fredriksson R., Schiöth H.B. Exposure to bisphenol A affects lipid metabolism in *Drosophila melanogaster. Basic Clin. Pharmacol. Toxicol.* 2014; 114(5):414-420. DOI 10.1111/bcpt.12170.

ORCID ID
M.A. Eremina orcid.org/0000-0001-6136-6928
N.E. Gruntenko orcid.org/0000-0003-3272-1518

Parent-of-origin effects on nuclear chromatin organization and behavior in a Drosophila model for Williams–Beuren Syndrome

A.V. Medvedeva[1], E.V. Tokmatcheva[1], A.N. Kaminskaya[2], S.A. Vasileva[1], E.A. Nikitina[1, 3], A.V. Zhuravlev[1], G.A. Zakharov[1], O.G. Zatsepina[4], E.V. Savvateeva-Popova[1] ✉

[1] Pavlov Institute of Physiology of the Russian Academy of Sciences, St. Petersburg, Russia
[2] Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Moscow, Russia
[3] Herzen State Pedagogical University of Russia, St. Petersburg, Russia
[4] Engelhardt Institute of Molecular Biology of the Russian Academy of Sciences, Moscow, Russia
✉ esavvateeva@mail.ru

Abstract. Prognosis of neuropsychiatric disorders in progeny requires consideration of individual (1) parent-of-origin effects (POEs) relying on (2) the nerve cell nuclear 3D chromatin architecture and (3) impact of parent-specific miRNAs. Additionally, the shaping of cognitive phenotypes in parents depends on both learning acquisition and forgetting, or memory erasure. These processes are independent and controlled by different signal cascades: the first is cAMP-dependent, the second relies on actin remodeling by small GTPase Rac1 – LIMK1 (LIM-kinase 1). Simple experimental model systems such as Drosophila help probe the causes and consequences leading to human neurocognitive pathologies. Recently, we have developed a Drosophila model for Williams–Beuren Syndrome (WBS): a mutant *agn^ts3* of the *agnostic* locus (X:11AB) harboring the *dlimk1* gene. The *agn^ts3* mutation drastically increases the frequency of ectopic contacts (FEC) in specific regions of intercalary heterochromatin, suppresses learning/memory and affects locomotion. As is shown in this study, the polytene X chromosome bands in reciprocal hybrids between *agn^ts3* and the wild type strain *Berlin* are heterogeneous in modes of FEC regulation depending either on maternal or paternal gene origin. Bioinformatic analysis reveals that FEC between X:11AB and the other X chromosome bands correlates with the occurrence of short (~30 bp) identical DNA fragments partly homologous to Drosophila 372-bp satellite DNA repeat. Although learning acquisition in a conditioned courtship suppression paradigm is similar in hybrids, the middle-term memory formation shows patroclinic inheritance. Seemingly, this depends on changes in miR-974 expression. Several parameters of locomotion demonstrate heterosis. Our data indicate that the *agn^ts3* locus is capable of trans-regulating gene activity via POEs on the chromatin nuclear organization, thereby affecting behavior.

Key words: POE (parent-of-origin effects); 3D nuclear architecture; chromatin ectopic contacts; LIM-kinase 1 (LIMK1); actin; mir-RNA; learning acquisition; memory formation; locomotion.

Влияние родительского происхождения аллелей на пространственную организацию хромосом и поведение на модели синдрома Вильямса–Бойерна на дрозофиле

А.В. Медведева[1], Е.В. Токмачева[1], А.Н. Каминская[2], С.А. Васильева[1], Е.А. Никитина[1, 3], А.В. Журавлев[1], Г.А. Захаров[1], О.Г. Зацепина[4], Е.В. Савватеева-Попова[1] ✉

[1] Институт физиологии им. И.П. Павлова Российской академии наук, Санкт-Петербург, Россия
[2] Институт биоорганической химии Российской академии наук, Москва, Россия
[3] Российский государственный педагогический университет им. А.И. Герцена, Санкт-Петербург, Россия
[4] Институт молекулярной биологии им. В.А. Энгельгардта Российской академии наук, Москва, Россия
✉ esavvateeva@mail.ru

Аннотация. Прогноз развития нейропсихиатрических заболеваний требует учитывать родительское происхождение аллелей как существенный фактор предрасположенности у потомства. Родительское наследование определяет 3D организацию хромосом в ядре нервных клеток, в том числе за счет эпигенетического влияния микро-РНК генеративных клеток родителей. Кроме того, когнитивные нейропатологии у родителей зависят от двух процессов – обучения и забывания, или стирания памяти. Эти процессы независимы и контролируются разными сигнальными каскадами: обучение – цАМФ-зависимым, забывание – каскадом ремоделирования актина: малая ГТФаза Rac1 – LIMK1 (LIM-kinase 1). Для понимания становления нейропатологии человека необходимо

А.В. Медведева, Е.В. Токмачева, А.Н. Каминская …
Г.А. Захаров, О.Г. Зацепина, Е.В. Савватеева-Попова

Роль родительского происхождения аллелей
в организации хромосом

2021
25•5

привлечение простых модельных объектов. Нами создана модель синдрома Вильямса–Бойерна на дрозофиле с мутационным повреждением гена *dlimk1 – agnostic* (*agn^ts3*), кодирующего ключевой фермент ремоделирования актина LIMK1. У *agn^ts3* повышена частота формирования негомологичных контактов в специфических районах интеркалярного гетерохроматина, резко нарушены способность к обучению, формированию памяти и локомоция. У реципрокных гибридов между *agn^ts3* и линией дикого типа *Berlin* частота эктопических контактов, сформированных дисками политенных хромосом, зависит от направления скрещивания, воспроизводя либо отцовские, либо материнские свойства. Биоинформационный анализ показывает, что частота эктопических контактов между Х:11АВ и другими районами Х хромосомы обусловлена присутствием короткого (~30 п.н.) фрагмента ДНК, частично гомологичного участку 372 п.н. сателлитной ДНК. Гибриды, имея одинаковую способность к обучению в парадигме условно-рефлекторного подавления ухаживания, проявляют патроклинный характер наследования среднесрочной памяти. Это может быть связано с уровнем экспрессии миР-794. Параметры локомоторной активности проявляют гетерозис. По-видимому, локус *agn^ts3* осуществляет трансрегуляцию пространственной организации ядра, тем самым влияя на количественные признаки (поведение).

Ключевые слова: эффект родительского происхождения аллелей; 3D организация ядра; эктопические контакты хромосом; LIM-киназа 1 (LIMK1); актин; микро-РНК; обучение; формирование памяти; локомоция.

Introduction

Genome plasticity is ensured by the architecture of specific nuclear loci and nuclear localization of transcriptional machinery (Medrano-Fernández, Barco, 2016; Iourov et al., 2019) where the chromatin organization is the priority-driven factor (Ito et al., 2014; Medrano-Fernández, Barco, 2016; Li et al., 2018). The outcome of recent achievements in systems biology is the notion that the plasticity of 3D chromatin architecture of nervous cell nuclei plays the leading role in cognition and neuropsychiatric disorders (Medrano-Fernández, Barco, 2016; Kim et al., 2018; Iourov et al., 2019). The epigenetic component is still an underestimated source of psychomotor disturbances and neuronal diversity (Savvateeva-Popova et al., 2017). Therefore, a new field of human biomedical research named molecular cytogenetics and cytogenomics (Iourov et al., 2008), or chromosomics (Liehr, 2019) has evolved. The main goal of chromosomics is the study of chromosomes, their 3D architecture in the interphase nucleus, the outcomes of chromosomal sub-region plasticity and gene interactions for shaping interindividual and intercellular genomic variations in normal behavior and disease.

Recently, this topical problem has turned out to be the pursuit of understanding the concomitant role of active forgetting, since the antithesis to learning acquisition is the forgetting or memory erasure (Davis, Zhong, 2017). Both processes are independent and controlled by different signal cascades: learning acquisition and memory consolidation occurs via cAMP cascade, its components being CREB and C/EBP. Active forgetting relies on actin remodeling cascade responsible for structural alterations of neurons and synapses: small GTPase Rac1 – LIMK1 (the key enzyme of actin remodeling LIM-kinase 1) and its phosphorylation substrate cofilin. The absence of Rac1-dependent forgetting causes the autistic spectrum disorders. Expression changes (active or non-active state) of LIMK1 and cofilin lead to different neuro pathologies. The most studied example embracing all the aforementioned facets of manifestations is Microdeletion (Deletion) Williams–Beuren Syndrome, or WBS in 7q11.23. WBS deletion leads to cardiovascular pathology, cognitive deficit in visuospatial construction and hypersociability (Kaiser-Rogers, Rao, 2005; Nikitina et al., 2014c). This is because long-term synaptic plasticity and depression determine successfulness of learning and memory, as well as of locomotor behavior, and depend on epigenetic regulation of LIM-kinase 1 (LIMK1) gene, one of approximately 28 genes uncovered by WBS deletion. Epigenetic regulation of LIMK1 activity involves DNA methylation, chromatin remodeling and the noncoding RNA-mediated process (Smrt, Zhao, 2010).

LIMK1, a member of serine/threonine (Ser/Thr) family kinases regulated by the Rho-GTPase pathway, is the key enzyme of actin remodeling cascade. Dendritic spines are actin-rich structures, and spine dynamics is driven mainly by actin remodeling, thus sharing several molecular pathways with dendrite growth (Smrt, Zhao, 2010). Increasing sets of evidence suggest that nuclear actin also plays a pivotal role in transcriptional regulation and DNA repair. Interestingly, monomeric actin is a stoichiometric subunit of a variety of chromatin remodeling complexes. A shift between monomeric and polymeric states modifies activity of histone deacetylases (Klages-Mundt et al., 2018).

Recently, we have developed a simple and appropriate Drosophila model for chromosomics (Savvateeva-Popova et al., 2017) using the properties of the *agnostic* locus harboring LIMK1 gene (X:11AB). This region possesses the properties of intercalary heterochromatin. Being a hotspot of chromosome breaks, ectopic contacts, underreplication and recombination, the region attains strain-specific architecture marked by single base changes and small insertion/deletions. The EMS-induced temperature-sensitive (ts) mutation *agn^ts3* carries the insertion of transposable element (TE) from Tc1/mariner superfamily ~460 bp downstream 3′UTR of Drosophila LIMK1 gene (*dlimk1*), as well as A/T-rich 28 bp insertion within intron 1 of *dlimk1* capable of pairing with 5′ TIR of the TE.

When maintained at 29 °C, *agn^ts3* shows a temperature-sensitive lethality at all stages of development except for the imaginal stage. At normal temperature, the adult flies show drastic learning acquisition and memory retention defects, as well as locomotor impairments and amyloid-like inclusions (Nikitina et al., 2014b; Kaminskaya et al., 2015). Stress exposure (heat shock for 30 min at 37 °C) suppresses these manifestations (Nikitina et al., 2012, 2014a). Also, *agn^ts3* mutation leads to: (1) LIMK1 and p-cofilin increase in the adult brain and salivary glands of 3rd instar larvae at 22–25 °C and a fall down to the level of the wild type strain *Canton S* at 29–37 °C; (2) high level of ts-induced recombination within *agn^ts3* region; (3) 3-fold increase in frequency of non-allelic ectopic contacts (FEC) within 2L arm of the chromosome 2

and in the 11B X chromosome region (Medvedeva et al., 2010). Additionally, miRNAs expression including the bio-markers for human neuropathologies is drastically reduced in *agn^{ts3}* relative to the wild type strains (Savvateeva-Popova et al., 2017).

agn^{ts3}-specific nuclear organization is shaped in early em-bryogenesis alongside with formation of chromosomal hetero-chromatin regions. Intrinsic *agn^{ts3}* FEC is maternally inherited (Medvedeva et al., 2010). Therefore, *agn^{ts3}* is a promising model for studies on parent-of-origin effects (POEs) on prog-eny considered as significant causative factors of psychiatric disorders (Zayats et al., 2015). For instance, a 1.5 Mb WBS deletion recurrently arises *de novo* and depends on POEs: ma-ternal origin leads to more severe developmental abnormali-ties and microcephaly (Pérez Jurado et al., 1996). Moreover, when WBS deletion has a paternal origin, expression levels of a number of genes within the WBS deletion decrease. Among these genes crucial for the brain development is a gene for ge-neral transcription factor II-I (GTF2I). It regulates transcrip-tion by binding to DNA and histone deacetylase (HDAC) (Col-lette et al., 2009). The main goal of the study is the analysis of POEs role in shaping quantitative traits, namely learning acquisition, memory retention, locomotion, and miRNAs ex-pression while using the advantages of the Drosophila model for POEs in progeny from reciprocal crosses between *agn^{ts3}* and the wild type strain *Berlin* (Fig. 1).

To meet the requirements of chromosomics, FECs between the region X:11AB and the other bands of the X chromosome may be estimated as an indicator of chromosomal spatial organization. This approach is justified by the existence of late-replicating genomic territories including the underrepli-cated regions of polytene chromosomes. These regions over-lap with late-replicating regions of mitotically dividing cells (Belyakin et al., 2005). Suppression of SUUR gene responsible for underreplication of intercalary heterochromatin and, as a consequence, for the ectopic pairing leads to death in early em-bryogenesis (Belyaeva et al., 1998). Additionally, to elucidate the contribution of DNA sequence homology as components of epigenetic regulation in the ectopic chromatin pairing we have developed the special software package Homology Seg-ment Analysis.

Materials and methods

Drosophila stocks. The fly stocks used belong to Biocollec-tion of Pavlov Institute of Physiology of the Russian Acade-my of Sciences:
- *Berlin*, a wild type strain;
- *Canton S*, a wild type strain;
- *agn^{ts3}*, a temperature-sensitive mutation on *Canton S* genetic background within *agnostic* locus (X:11AB) af-fecting *dlimk1* activity.

The reciprocal hybrids between *agn^{ts3}* and *Berlin* were used because *Berlin dlimk1* sequence is closer to FlyBase reference sequence (Savvateeva-Popova et al., 2017). At the same time, the reciprocal hybrids *agn^{ts3}* × *Canton S* (the genetic back-ground for *agn^{ts3}*) and *Canton S* × *agn^{ts3}* demonstrate exactly the same cognitive behavior as reciprocal hybrids with *Berlin* (Vasiljeva et al., 2019) approving the usage of *Berlin*.

Figure 1 shows the X chromosome and autosomes archi-tecture in *Berlin* and *agn^{ts3}* female and male parents and
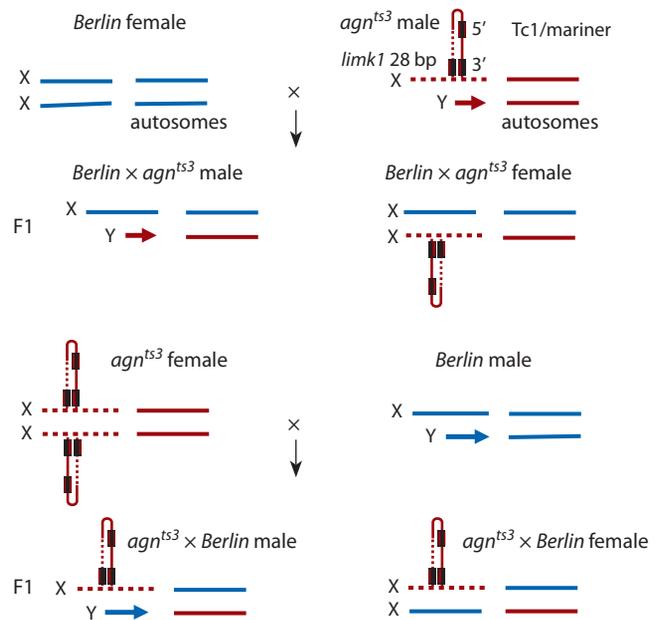
Fig. 1. Architecture of Drosophila chromosomes in the studied strains.

F1 female and male progeny from reciprocal crosses with the accent on the putative hairpin formed in the X chromosome by 28 bp A/T rich insertion within intron 1 of *dlimk1* gene and 3′ end of Tc1/mariner element.

Flies were maintained on standard Drosophila yeast-raisin medium at +22±0.5 °C under a 12-h light/dark cycle. For the memory and locomotion tests, males were collected upon eclosion without narcotization and kept individually in culture vials till the behavioral experiments on the 5th day.

Estimation of frequency of ectopic contacts (FECs). The aceto-orcein squash preparations were prepared from salivary glands of III instar *D. melanogaster* female larvae. 20 to 30 animals were examined, therefore the number of analyzed chromosomes varied from 300 to 500. The examples of ectopic contacts are presented on Fig. 2. The number of non-homologous contacts between the region X:11AB and different bands of the X chromosome was calculated and ex-pressed as per cent of the total number of the examined nuclei. FECs in parents and F1 reciprocal hybrids were compared using Student's *t*-test. Identification of genes localized in the X chromosome bands forming contacts with the 11AB region was performed using NCBI Genome Data Viewer database (https://www.ncbi.nlm.nih.gov/genome/gdv/) and molecular function of identified genes was derived from FlyBase (https://flybase.org).

Bioinformatics analysis of DNA segments homology. *D. melanogaster* genome sequence (release 6) was taken from (Zerbino et al., 2018). Special software package Homology Segment Analysis searching the matches of short single-stranded DNA fragments within the chromosome areas involved in the ectopic pairing in Drosophila has been deve-loped. The software written in Python 3 can be freely down-loaded from (Zhuravlev, 2019a). Software version from git (commit 41719cddc6283edbd79c5bf2aee237cde48d4b7d) was used. The algorithm of the program is described in brief in (Zhuravlev, 2019b). The exact run parameters for 11AB region

А.В. Медведева, Е.В. Токмачева, А.Н. Каминская …
Г.А. Захаров, О.Г. Зацепина, Е.В. Савватеева-Попова

Роль родительского происхождения аллелей
в организации хромосом

2021
25 • 5
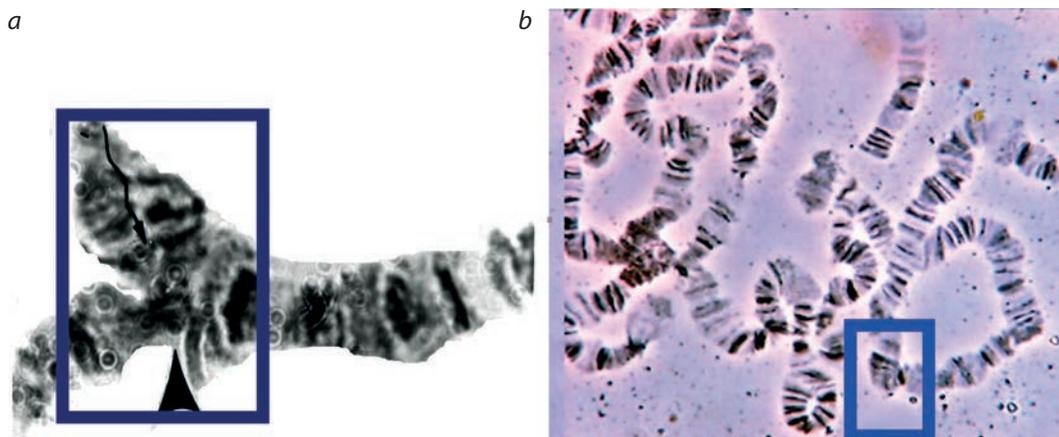
Fig. 2. Examples of ectopic contacts in the 2L (*a*) and in the X (*b*) chromosomes.

are following: segmentanalysis.py -v -s 30 dm6.nounmapped.
fa.gz :X:11982050:12772070 DmelMapTable.160615c.bed

Here, dm6.nounmapped.fa.gz is Drosophila genome canonical sequence (Ensembl release 6) and DmelMapTable.160615c.bed is chromosome bands location from Ensembl database converted to BED format. Both files are included in software repository.

For other tested regions, all parameters were the same except for the regions location.

Preparation of miRNAs libraries and bioinformatic analysis. The detailed description of the procedure is given in (Savvateeva-Popova et al., 2017). Extract RNA reagent (Evrogen, Russia) was used for total RNA extraction from adult 5 days old males. To obtain the fraction of small RNA, 25 µg of total RNA were separated using 15 % polyacrylamide gel electrophoresis in the presence of Urea (8 M) following excision of small RNA fraction corresponding to 21–29 nts. Illumina TruSeq Small RNA prep kit (Illumina, USA) was used for small RNA libraries preparation. Sequencing was performed on an Illumina HiSeq 2000 platform.

The amount of mapped miRNAs reads was counted by BEDTools (v. 2.22) and mirbase annotation (r. 19) (Quinlan, Hall, 2010). Analysis of differentially expressed miRNAs was performed using edgeR (v. 3.10.2) package in R environment (v. 3.2.2) (Robinson et al., 2010). miRNA functions were derived from miRBase (Kozomara et al., 2019). Small RNA libraries were deposited in NCBI SRA under the number PRJNA633483.

Locomotor activity. Computer-aided automatic device for simultaneous registration of 20 animals is described in (Zakharov et al., 2012). The experiment lasted for 1 h. Spontaneous locomotor activity of flies was detected in a plate with eight chambers and transparent cover using high-resolution video camera. Software used for locomotor activity analysis is freely available at (Zakharov, 2017).

The following parameters of locomotion were assessed: activity index (%); run frequency (the number of run bouts in 100 seconds); running speed (mm/s). The full record was divided into 1 s quanta, and the mean speed of fly movement in each quantum was calculated. If the result was less than the threshold value (5 mm/s), the fly was considered to be resting during this time quantum; otherwise, it was considered

moving. Neighboring quanta with similar movement pattern were merged in intervals of moving and resting. Activity index is determined as a time spent in movement. Running speed is an average fly speed, determined using only intervals of movement. The Kruskal–Wallis analysis of variance with the multiple comparison of mean ranks was used to compare all the experimental groups.

Learning acquisition and middle-term memory formation in Drosophila males. Detailed description of learning/memory assessments in conditioned courtship suppression paradigm (CCSP) and specially designed software for observation and statistical analysis (randomization test) of learning indices based on courtship indices is given in (Kamyshev et al., 1999). CCSP employs the natural stimuli of Drosophila courtship. Both virgin and fertilized females emit an aphrodisiac pheromone, attracting a naïve male without courtship experience. However, a fertilized female rejects a male at the courtship stage of attempted copulation via emitting an aversive pheromone. Repetitive rejections during 30 min training provoke a kind of learned helplessness when a male stops courting another female. This courtship suppression might last for one hour when test female is virgin and for eight hours when fertilized. Males with defective memory formation continue to court after such a training as vigorously, as naïve males.

The courtship index (CI, percentage of time spent in courtship) was calculated for each male. The learning index (LI) was computed according to the formula:

$$LI = [(CI_{NA} - CI_{TR})/CI_{NA}] \cdot 100 = (1 - CI_{TR}/CI_{NA}) \cdot 100,$$

where CI_{NA} and CI_{TR} are the mean courtship indices for independent samples of naïve and trained males, respectively.

Results

Analysis of spatial nuclear organization delimited to FECs formed by the X chromosome region 11AB

Spatial nuclear organization was analyzed using microscopic images of polytene chromosomes in larvae salivary glands. The results of comparative analysis of FECs between the 11AB region and the other X chromosome regions in

Table 1. Comparative analysis of X – X:11AB FECs in *Berlin*, *agn*^ts3, and their reciprocal hybrids

X chromosome region		1	2	3	4	1 *vs* 2	3 *vs* 4	2 *vs* 3	2 *vs* 4	1 *vs* 3
		agn^ts3	*agn*^ts3 × *Berlin*	*Berlin* × *agn*^ts3	*Berlin*					
1	A	0.18	0.54	0	0.91	#				
	B	0.18	0.27	0	0.18	*			*	
	C	0	0.54	0	0	#				
	D	0	0.27	0	0	#				
	E	0	0	0	0					
	F	0	0	0	0					
2	A	0	0	0	0					
	B	0	0.54	0	0	#				
	C	0	0.27	0	0	#				
	D	0	0	0	0					
	E	0	0	0	0					
	F	0	0	0	0					
3	A	0	0.54	0	0	#				
	B	0.18	0	0	0					
	C	0.18	0.27	0	0	*				
	D	0	0	0	0.18					
	E	0	0	0	0					
	F	0	0	0	0.18					
4	A	0.18	0.27	0	0	*				
	B	0	0.27	0	0	#				
	C	0	0	0	0					
	D	0.18	0	0.34	0.18		#			*
	E	0	0.27	0	0.18	#			*	
	F	0	0.27	0.34	0	#	#	*		
5	A	0	0	0.67	0		#			
	B	0	0	0	0					
	C	0.18	0.27	0	0.18	*			*	
	D	0.18	0.27	0.34	0	*	#	*		*
	E	0	0	0	0					
	F	0	0	0	0					
6	A	0.72	0.27	0.34	0.18		#	*	*	
	B	0	0.27	0	0	#				
	C	0	0	0	0.18					
	D	0	0	0	0.18					
	E	0	0	0	0					
	F	0.18	0	0.34	0		#			*
7	A	0.18	0.27	0.67	0.73	*	*			
	B	0.36	0	0.34	0.55		*			*
	C	0.36	0.81	0	0.36	#				
	D	0	1.08	0	0	#				
	E	0.18	0.27	0	0	*				
	F	0.18	0.27	0	0.18	*			*	
8	A	0	0.27	0	0.36	#			*	
	B	0	0	0	1.28					
	C	0	0	0.34	0.91					
	D	0.72	0.81	1.35	0.18	*	#	*		
	E	2.15	1.62	1.35	2.19	*		*	*	*
	F	0.36	0	0	0.36					
9	A	2.69	0.27	3.70	4.19		*			*
	B	0.18	1.89	1.01	0.18	#	#			
	C	0	0	0.34	0.55		*			
	D	0	0	0.67	0.73		*			
	E	0.18	0.81	0.34	0.18	#	#			
	F	0.18	1.35	0	0.36	#				

А.В. Медведева, Е.В. Токмачева, А.Н. Каминская …
Г.А. Захаров, О.Г. Зацепина, Е.В. Савватеева-Попова

Роль родительского происхождения аллелей
в организации хромосом

2021
25•5

Table 1 (continued)

X chromosome region		1	2	3	4	1 *vs* 2	3 *vs* 4	2 *vs* 3	2 *vs* 4	1 *vs* 3
		agn^{ts3}	$agn^{ts3} \times Berlin$	$Berlin \times agn^{ts3}$	$Berlin$					
10	A	2.87	5.14	4.71	2.91	#	*	*		*
	B	11.29	5.95	8.08	5.65		*	*	*	*
	C	1.08	2.16	2.02	0.73	#	#	*		
	D	0.36	3.24	3.37	0.73	#	#	*		
	E	1.25	2.97	4.71	0.55	#	#	*		
	F	0.36	2.97	0.67	0	#	#			
11	A	NA	NA	NA	NA					
	B	NA	NA	NA	NA					
	C	9.14	12.43	12.12	6.92	*	#	*		
	D	0	12.43	12.12	6.92	#	#	*		
	E	0	0.27	0	0.36	#			*	
	F	0	0.27	0	0	#				
12	A	1.79	2.43	2.36	4.02	*		*		*
	B	0.36	1.08	0	1.82	#				
	C	0.72	2.16	0	0.91	#				
	D	1.61	3.24	1.68	1.28	#	*			*
	E	0.36	1.89	6.06	0.18		#			
	F	8.24	1.62	1.01	10.56		*			
13	A	1.79	0.81	0.67	0.91		*	*		
	B	0.36	0.27	1.01	0.73	*	*			
	C	0.36	0.81	0.67	0.55		*	*	*	
	D	0.72	0.81	0	0.36	*				
	E	0.36	0	0	0					
	F	0	0	0	0.18					
14	A	0	0.54	1.01	0.18	#	#	*		
	B	0	0.27	0.34	0.18	#	#	*	*	
	C	0	0	0	0					
	D	0.18	0	0.34	0		#			
	E	0	0	0	0					
	F	0	0	0	0					
15	A	0	0	0	0					
	B	0	0	0	0.18					
	C	0	0	0	0					
	D	0	0	0	0.18					
	E	0	0	0	0					
	F	0	0	0	0					
16	A	0.18	0	0.34	0.18		#			*
	B	0.18	0	0	0					
	C	0	0	0	0.18					
	D	0	0	0.34	0.18		#			
	E	0	0	0	0					
	F	0.18	0	0	0					
17	A	0.18	0	0.34	0.36		#			*
	B	0.18	0.27	0	0	*				
	C	0	0.27	0	0	#				
	D	0	0	0	0					
	E	0	0	0	0					
	F	0	0	0	0					
18	A	0.18	0	0	0.18					
	B	0.18	0	0	0.36					
	C	0	0	0	0.36					
	D	0	0	0	0					
	E	0	0	0	0					
	F	0	0	0	0					

Table 1 (end)

X chromosome region		1	2	3	4	1 vs 2	3 vs 4	2 vs 3	2 vs 4	1 vs 3
		agn^ts3	*agn^ts3* × *Berlin*	*Berlin* × *agn^ts3*	*Berlin*					
19	A	0	0.27	0	0.18	#			*	
	B	0	0	0	0					
	C	0.18	0.27	0	0	*				
	D	0.36	0.54	1.01	0	*	#	*		
	E	0	0.81	1.01	0.36	#	#	*		
	F	0.18	0	0	0					
20	A	0.18	0	0	0					
	B	0	0	0	0					
	C	0	0	0	0					
	D	0	0.27	0	0.36	#			*	
	E	0	0	0	0					
	F	0	0	0	0					
Chromocenter		1.25	1.62	0.34	0.73	*				

Note. NA (not applicable) for 11AB region frequencies indicates that region cannot pair with itself.

* No difference, two-sample Student t-test for difference of means (t_{dif}) <1 reflecting a mode of FEC inheritance: maternal – n columns 1 *vs* 2 and 3 *vs* 4, paternal – in columns 2 *vs* 4 and 1 *vs* 3, hybrid-specific – columns 2 vs 3.

Difference $t_{dif} \geq 1$, in columns 1 *vs* 2 and 3 *vs* 4 reflecting the region with FEC higher in hybrids than in parents (heterosis) or regions present only in the hybrid strains.

No symbol represents the absence of contacts, therefore statistical significance cannot be estimated.

Berlin, *agn^ts3* and their hybrids are presented in Table 1. The columns 1–4 show the pattern of FECs between the 11AB region and other X chromosome regions pertinent to listed strains. The absence of significant differences of FECs between columns, i. e. 1 *vs* 2 and 3 *vs* 4 indicates matroclinic inheritance, 2 *vs* 3 indicates hybrid-specific frequencies, 2 *vs* 4 and 1 *vs* 3 pinpoints the patroclinic inheritance. The polytene chromosome bands in hybrids demonstrate differences in FECs, a part of them showing either matroclinic properties or properties of the father strain. In certain bands, FECs are similar in reciprocal hybrids (hybrids-specific, i. e. FEC is equal for hybrids, being different from at least one parent) or depend on the direction of a cross, but significantly differ from that of parents.

NCBI Genome Data Viewer software helped to reveal the genes located in the X chromosome bands forming the ectopic contacts. Based on the assumption that shared location of genes determines their functional features (Liu et al., 2019), it was worth elucidating what biological processes are under the influence of epigenetic factors related to allelic parent-of-origin. Therefore, the grouping of genes implied their involvement in control of a certain biologic process (Fig. 3).

Figure 3, *a* presents provisional definition of the biological processes controlled by genes within bands forming contacts with 11AB region either in matroclinic or in reciprocal hybrid-specific manner. Among the functional groups of genes having increased number in *agn^ts3* × *Berlin* relative to the reciprocal hybrid are genes for motor proteins (8, 4.5-fold) and sensory receptors (2, 3.5-fold). Genes involved in neurodevelopment, oxidative-reduction process and proliferation do not present in *Berlin* × *agn^ts3* progeny. Heterosis, when the number of genes involved in hybrid-specific contacts is more than 2-fold higher than in both parents, is manifested for oxidative-reduction process (3), signal transduction (4), proliferation (9) and non-coding RNAs (13).

The biological processes and genes involved in ectopic contacts with 11AB in a mode of father strain are shown in Fig. 3, *b*. Among them are the genes responsible for chromatin remodeling (5, 5-fold increase in *agn^ts3* × *Berlin* relative to the reciprocal hybrid), reactive oxygen species metabolic process (3, 3-fold), and metabolism (6, 3-fold). Sensory perception and signal transduction groups do not present in *Berlin* × *agn^ts3*.

Data on biological processes and the number of genes involved in contacts with 11AB region exclusively in hybrids or with FECs prevailing those of parents (FEC heterosis), are shown on Fig. 3, *c*. The ratio of gene numbers for *agn^ts3* × *Berlin* and *Berlin* × *agn^ts3* reveals functional significance of other gene groups. The magnitudes of differences may be ranged as follows: lipid metabolism and neurotransmitter secretion (20, 8-fold); metabolism (6, 2.3-fold); cell proliferation (9, 6-fold; as in previously published evidence on *agn^ts3* (Tokmacheva, 1995)); chromatine remodeling (5, 3.5-fold). Reparation/recombination group does not present in *Berlin* × *agn^ts3*.

Profiles of the localized fragment frequencies (LFFs)
Using our specially designed software Homology Segment Analysis, we analyzed correlation between ectopic pairing and distribution of small identical fragments within contacting regions.

The region X:11AB (~790 kb) involved in Drosophila X – X:11AB ectopic pairing was selected as the source of small 30 nt fragments, and the profile of X:11AB localized fragment frequencies (LFFs) was constructed for the X chromosome (Fig. 4). The region X:14B–15B was selected as a control region having almost equal length (~790 kb).

Both LFF and the FEC (X – X:11AB) distributions significantly differ from normal. Therefore, we calculated Spearman's rank-order correlation coefficients (r_S) for LFF and the strain-specific FEC (X – X:11AB) (Table 2). A positive correlation between LFF and FEC is observed within the re-

А.В. Медведева, Е.В. Токмачева, А.Н. Каминская ...
Г.А. Захаров, О.Г. Зацепина, Е.В. Савватеева-Попова

Роль родительского происхождения аллелей
в организации хромосом

2021
25•5

1. Neurodevelopment
2. Sensory receptors
3. Oxidative-reduction process/mitochondria
4. Signal transduction
5. Chromatin remodeling
6. Metabolism
7. Transcription factors
8. Actin and microtubules-binding proteins, motor function
9. Proliferation
10. Chitin
11. Morphogenesis
12. Membranes and lipid metabolism
13. Non-coding RNAs
14. Sensory perception
15. Transcription apparatus
16. Muscle system, actin-binding, microtubules
17. Meiosis, reparation, recombination
18. Translation apparatus
19. Learning, memory, courtship behavior, locomotion
20. Neurotransmitters, secretion, lipid metabolism
21. Ubiquitination
22. Morphogenesis, oogenesis

Fig. 3. Processes controlled by genes within bands involved in ectopic contacts: *a* – matroclinic- and reciprocal hybrid-specific contacts; *b* – patroclinic-specific contacts; *c* – hybrids-specific contacts.

X axis – the biological processes, Y axis – the number of genes.
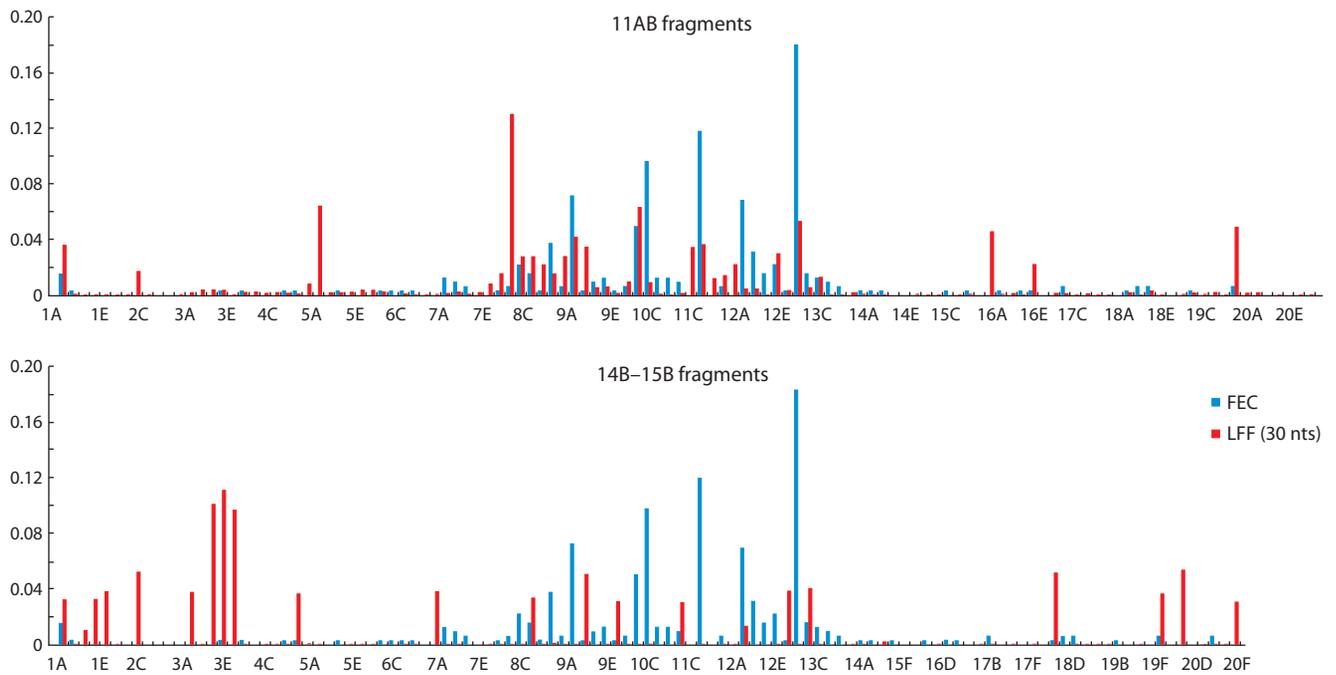
Fig. 4. *Berlin* ectopic contacts frequencies (FECs) and localized fragment frequencies (LFFs) for fragments of 11AB and 14B–15B regions. The X chromosome bands are shown.

gion of interest X:11AB. The influence of the fragment length on r_S value has been tested. For 30 nt fragments, r_S is highly significant for all Drosophila strains ($p < 0.001$). As to LFF (X:14B–15B), there is no significant correlation with FEC (X – X:11AB) ($p ≥ 0.01$) pointing to specificity of analysis of DNA homology within the region of ectopic pairing. For

20 nt fragments, there is a false-positive correlation between LFF (X:14B–15B) and *agn^{ts3}×Berlin* FEC (X:11AB). Probably, 20 nt fragments are too short to display DNA homology specifically associated with ectopic pairing. The increase in fragments length from 20 to 50 nt leads to decrease in LFF (X:14B–15B) – FEC (X – X:11AB) unspecific correlation

Table 2. Spearman's rank correlation (r_S) between LFF and strain-specific X – X:11AB FEC

X chromosome region	Fragments length, nt					
	20		30		50	
	r_S	p	r_S	p	r_S	p
11AB						
Berlin	0.294	0.001	0.354	0.0001	0.393	0.0000
agn^{ts3}	0.287	0.002	0.355	0.0001	**0.207**	**0.0239**
Berlin × agn^{ts3}	0.334	0.0002	0.354	0.0001	0.347	0.0001
agn^{ts3} × Berlin	0.319	0.0004	0.318	0.0004	0.278	0.0022
14B–15B						
Berlin	**0.186**	**0.043**	**0.113**	**0.219**	**0.041**	**0.634**
agn^{ts3}	**0.206**	**0.045**	**0.138**	**0.134**	**0.030**	**0.746**
Berlin × agn^{ts3}	**0.214**	**0.020**	**0.131**	**0.154**	**0.085**	**0.372**
agn^{ts3} × Berlin	0.236	0.009	**0.206**	**0.025**	**0.171**	**0.063**

Note. Bold: non-significant correlation ($p \geq 0.01$).

Table 3. Spearman's rank correlation (r_S) between *Berlin* X – X:11AB FEC and localization frequency values for specific DNA sequences

Sequence	NO	X area	r_S	p
aattttcgcatttttttgtaaggggtaacatcatcaaaatttgcaaaaaat[#]	69	5E–11F	**0.2180**	**0.0172**
aaaaaacaaactcaacgaggtatgacattc[*]	39	7D–12F	0.303	0.0008
atgatgttaccccttacaaaaaatgcgaaa[*]	110	3B–13B	0.300	0.0009
atttttttgtaaggggtaacatcatcaaaat[*]	116	4F–13E	**0.206**	**0.0241**
(-at-)$_{15}$	694	2B–20C	**0.195**	**0.033**
(-ca-)$_{15}$	311	1C–20F	**0.027**	**0.770**
(-gt-)$_{15}$	289	3A–20F	**–0.002**	**0.985**
(-tc-)$_{15}$	75	2A–19C	**–0.001**	**0.989**

Note. Bold: non-significant correlation ($p \geq 0.01$). NO – the number of occurrence of fragment within the X chromosome excluding area 11AB; X area – the area of the X chromosome where the fragment has been localized.
[#] 50 bp fragment; [*] fragments with a partial homology to 372 bp repeat (see in text).

and simultaneous increase in LFF – FEC (X:11AB) specific correlation. However, there is a lack of LFF (50 nt) – ECF correlation for *agn^{ts3}*. Hence, 30 nt fragment length seems to be optimal in search for the identical fragments within the candidate chromosomal regions involved in ectopic pairing.

To find out what DNA sequences impact the ectopic pairing the most, we calculated r_S between *Berlin* FEC and the localization frequencies for the several specific X:11AB 30 nt fragments having the highest number of occurrences (NO) in the X chromosome (Table 3). The maximal correlation is evident in the part of 372 bp middle repetitive DNA sequence (marked with asterisk (Waring, Pollack, 1987)). The majority of fragments with significant positive r_S values appear to be the parts of a ~50 bp repeat (marked with #) having significant self-complementarity. This sequence also shows an almost complete identity to another part of 372 bp repeat. Such repeats with slight sequence variations occur in both DNA strands of the X chromosome. r_S is much lower for (-at-)$_{15}$ repeat and is nearly absent for (-gt-)$_{15}$, (-tc-)$_{15}$ and (-ca-)$_{15}$ repeats, although their NO can be higher. Noteworthy, all r_S values for 372 bp repeat fragments are lower compared to r_S for the whole set

of the localized fragments (see Table 2), hence all of them seemingly impact ectopic pairing.

miRNAs expression profile

Analysis of miRNAs expression in *Berlin*, *agn^{ts3}* and their reciprocal hybrids demonstrates significant differences between males of parent strains and hybrids in content of 44 miRNAs. For the reciprocal hybrids, the heat maps of miRNAs expression are presented in Fig. 5. Among 44 miRNAs 10 miRNAs belong to the same cluster of testis-specific miRNAs (Mohammed et al., 2014). However, only miR-980 and miR-974 are involved in memory processes. Expression of miR-980 suppresses Drosophila memory (Guven-Ozkan et al., 2016), while lowered expression of miR-974 impairs memory formation (Busto et al., 2015).

The heat map represents the RPM-normalized and log$_2$-transformed counts of miRNAs reads with z-scale normalization of the rows. Thirty percent of low-expressed miRNAs were removed from further analysis. Only the miRNAs with altered content in a hybrid compared to at least one parent are shown.

А.В. Медведева, Е.В. Токмачева, А.Н. Каминская …
Г.А. Захаров, О.Г. Зацепина, Е.В. Савватеева-Попова

Роль родительского происхождения аллелей
в организации хромосом

2021
25•5
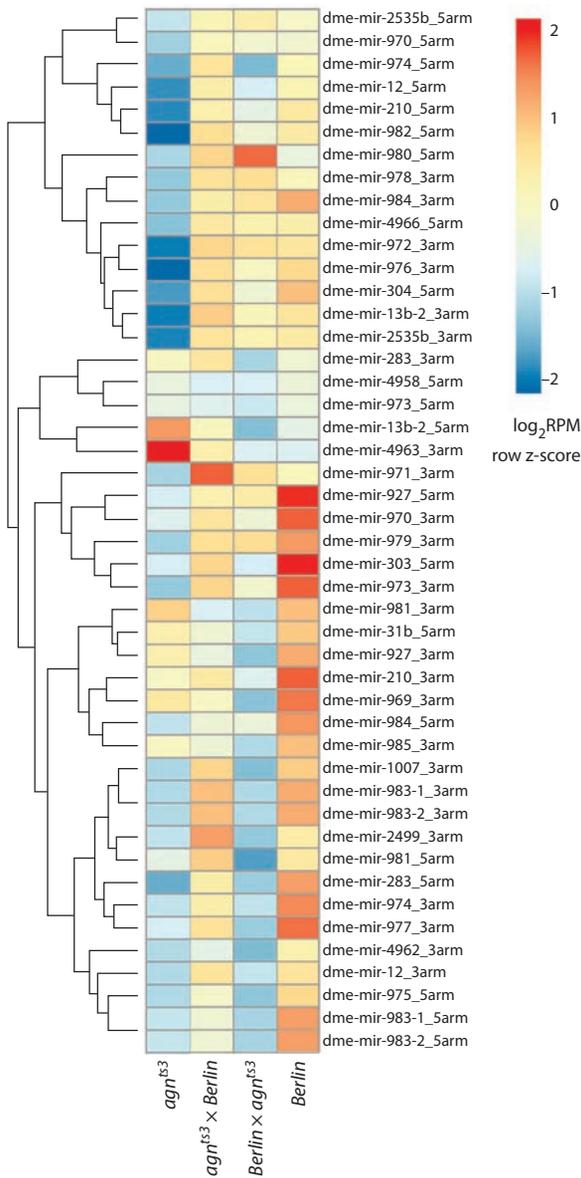
Fig. 5. The relative content of miRNAs in reciprocal hybrids compared to parents.

Behavioral analysis of parent strains and their reciprocal hybrids

Locomotor behavior. Figure 6 presents the parameters of locomotion. As to activity indices, they are significantly different in agn^{ts3} and *Berlin*, both hybrids differ from agn^{ts3} and *Berlin* × agn^{ts3} differs from *Berlin*. Running speed and run frequency in agn^{ts3} and *Berlin* are similar. Both hybrids differ from agn^{ts3} and *Berlin* × agn^{ts3} from *Berlin*. However, the hybrid running speed exceeds that of parents, demonstrating heterosis. Comparatively to parents, run frequency in hybrids is intermediate. At the same time, any alterations in locomotor parameters in hybrids are similar and unidirectional.

Learning acquisition and memory formation. In all strains, CIs of males decrease after training with fertilized females compared to naïve flies (Fig. 7, *a*). Learning/memory scores in reciprocal hybrids show that memory formation (3 hours after training), but not learning acquisition (0 hours after training), demonstrates patroclinic inheritance (see Fig. 7, *b*).
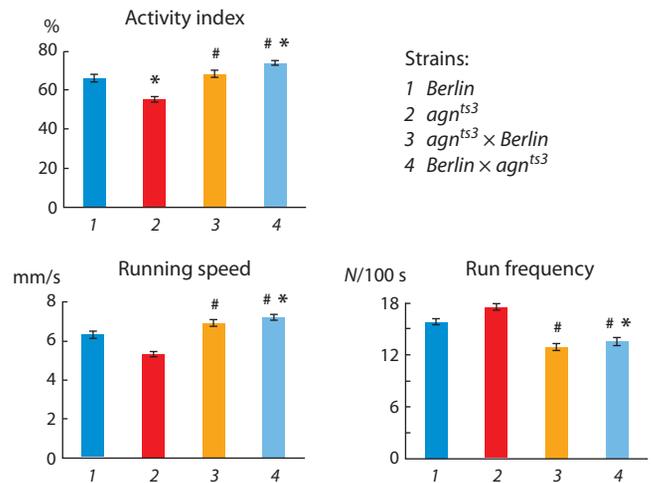
Fig. 6. Locomotor activity in agn^{ts3}, *Berlin* and their reciprocal hybrids.

N – the number of run bouts; # the difference from agn^{ts3}; * the difference from *Berlin* (Kruskal–Wallis analysis, the multiple comparison of ranges, $n = 20$, $p \leq 0.05$).
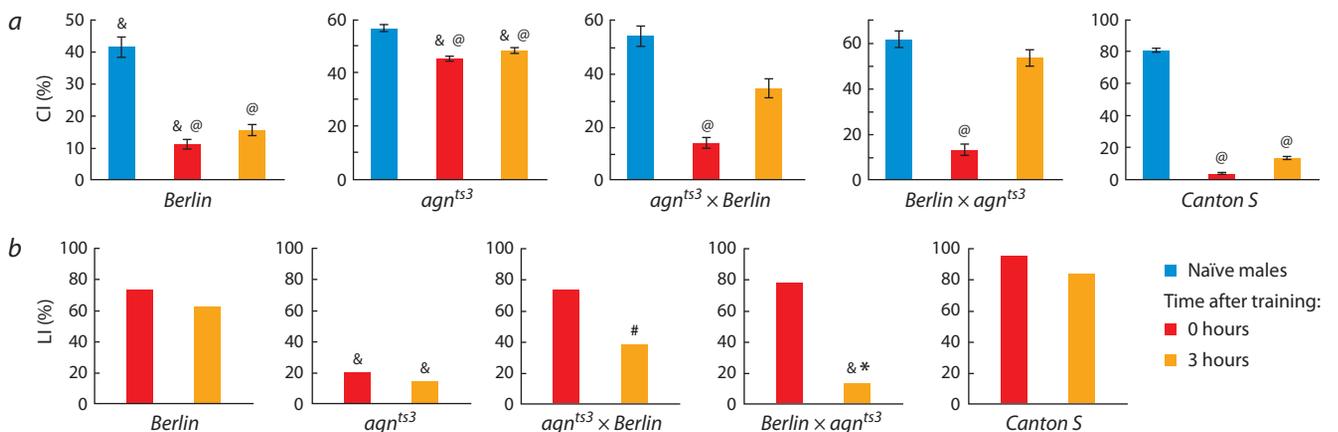
Fig. 7. Learning and memory in parents agn^{ts3}, *Berlin* and their reciprocal hybrids immediately and 3 hrs after training: *a* – courtship indices; *b* – learning indices.

@ difference from naïve males; # difference from agn^{ts3}; * difference from *Berlin*; & difference from *Canton S* (two-way randomization test, $p < 0.05$).

When male parent originates from *Berlin* strain in crosses *agn^{ts3}* × *Berlin*, learning and 3-hour memory do not differ from LIs of *Berlin*, but do differ from LIs of *agn^{ts3}*. Patroclinic inheritance is evident in the reciprocal hybrid *Berlin* × *agn^{ts3}*. In this case, 3-hour LIs in *Berlin* × *agn^{ts3}* and *agn^{ts3}* are similar. The patroclinic inheritance cannot be associated with the Y chromosome, as it is likely to be similar in *Berlin*, *agn^{ts3}* and *Canton S* (see Materials and methods, Drosophila stocks description). Also, it cannot be attributed to the X chromosomes, since they are different in *Berlin* and *agn^{ts3}* × *Berlin*. Seemingly, this may be caused by some paternal epigenetic factors, such as miRNAs in cytoplasm of male sperm.

Discussion
Studies of genome-wide associations between DNA polymorphisms and phenotypic traits have revealed genetic variants predisposing to different mental diseases. Findings pinpointing the role of POEs in genetic risk for neuropathology open new possibilities for therapy and preventive medicine (Zayats et al., 2015). In this study, we estimated FECs in reciprocal hybrids considering an impact both of genetic variants of *agn^{ts3}* gene and epigenetic factors (POEs) in spatial nuclear architecture, learning/memory formation and spontaneous locomotor activity.

To exploit the advantages given by the model, we delimited the analysis of spatial nuclear organization to FECs formed by the X chromosome region 11AB harboring *dlimk1* gene. Our assumption was that FECs partly reflect the restricted homology of short DNA sequences in different, seemingly "non-homologous" regions.

As shown in this study for the short identic fragments within the contacting regions, FECs correlate with LFFs. Although the correlation is rather moderate (~0.35), it is highly specific for 30 nt fragments. Many factors affect ectopic pairing, such as DNA homology, the distance between the interacting regions and epigenetic factors causing the interstrain FECs differences (Zykova et al., 2018). Our computational algorithm concerns only fragments aligned with the X chromosome without gaps. This reveals the partial homology of interacting bands. Thus, the interacting chromosomal areas may be significantly larger than 30 or 50 nts. Although different mechanisms are involved in ectopic pairing, including POEs, our data indicate the significant role of DNA sequence itself. However, as FEC–LFF correlation is mainly observed for specific DNA fragments, their pairing mediated by some proteins or non-coding RNAs cannot be ruled out.

Most of the found 30–50 nt fragments are similar to the *D. melanogaster* dispersed 372 bp A/T-rich noncoding repeat (Waring, Pollack, 1987). This moderately repeated sequence is located in the euchromatin of the X chromosome between the regions 4 and 14A in ~300–400 copies per haploid genome. The 372 bp repeat is a part of 1.688 g/cm³ class of satellite DNA (1.688X repeats) (Jagannathan et al., 2017). siRNA from the 1.688X repeats is involved in dosage compensation in recognition of the X and autosomal chromatin, thereby delimiting activities of male-specific lethal (MSL) complex to sex chromosomes through up-regulation of the X chromosome (Menon et al., 2014).

The data obtained consider the common mechanisms of ectopic contacts formation and dosage compensations. Seem-ingly, POEs might influence the spatial chromatin organization, thereby affecting behavioral performances.

This indicates that each Drosophila strain possesses its own pattern of ectopic contacts with the region 11AB. The polytene chromosome bands are heterogeneous in their modes of regulation of ectopic pairing. A part of them is regulated by genes of either maternal, or paternal origin. A separate class is comprised of regions manifesting only hybrid properties. Similar POEs were observed for the pattern of methylation and nucleosome distribution within the imprinted loci in humans and plants (Dong et al., 2018; Zink et al., 2018).

In both reciprocal crosses, bands 7A, 9A and 13B display the maternal properties. When mother is *agn^{ts3}*, FECs in these bands significantly decrease. In this case maternal control of spatial localization and therefore, of gene expression is genetically determined. These are genes controlling membrane receptor regulation (*PPYR1*) and signal transduction (*gce*), chromatin remodeling (*Top1*, *HDAC6*), axon guidance and chemosensory jumping behavior (*acj6*).

In the cross *agn^{ts3}* × *Berlin*, the number of genes with known functions contacting with 11AB with maternal-specific frequency is 2-fold higher than in reciprocal cross. Possibly, this is due to the *agn^{ts3}*-specific miRNAs pattern of expression. The role of miRNAs in maternal inheritance and expression in embryogenesis is sparsely studied. As we have shown earlier, the expression level of miR-9, miR-34 and miR-124 differs in *agn^{ts3}* from that in *Berlin*, *Canton S* and *Oregon-R* (Savvateeva-Popova et al., 2017). miR-9 and miR-124 are also expressed in early development (0–12 hrs) (Sempere et al., 2003), miR-34 is detected in embryos till zygotic reduction (Soni et al., 2013). As known, the switch from maternal to zygotic development program occurs between the second and the third hours of embryonic stage, hence miRNA found in early development have maternal origin (Schier, 2007). These miRNAs targets are Swi/Snf-like complex, neural-progenitor-specific npBAF, repressor-element-1-silencing transcription factor (REST belonging to 1 class of histone deacetylases (HDAC1/2) and silent information regulator 1 (SIRT1) – 3 class of NAD⁺-dependent histone deacetylases involved in heterochromatin formation, Bourassa, Ratan, 2014). They are involved in neurogenesis, dendrite morphogenesis and axon guidance which depend on global chromatin remodeling.

Therefore, it is not surprising that in cross *agn^{ts3}* × *Berlin* ectopic contacts between 11AB and regions containing genes involved in chromosome remodeling are formed with frequency characteristic for the maternal genome. The products of these genes are: Tip60 – histone acetyltransferase, HDAC6 – histone deacetylase; mxc – regulator of histone synthesis of Polycomb group; Top1 – DNA topoisomerase. However, only two regions containing genes *HDAC6* and *Top1* are present in cross *Berlin* × *agn^{ts3}*.

Noteworthy, new knowledge about topologically associating domains (TADs) indicates that polytene, diploid, and embryonic TADs condensation along the chromosome axis is just the same everywhere (Eagen et al., 2015). Moreover, comparison of TADs with 3D chromatin organization revealed by the Hi-C method confirms that the interphase nucleus spatial organization into TADs is directly represented by banding pattern of polytene chromosomes (Kolesnikova, 2018). Therefore, this allows to bridge the ratio of genes forming

А.В. Медведева, Е.В. Токмачева, А.Н. Каминская …
Г.А. Захаров, О.Г. Зацепина, Е.В. Савватеева-Попова

Роль родительского происхождения аллелей
в организации хромосом

2021
25•5

ectopic contacts in a mode of either maternal, or paternal strain in reciprocal crosses and their physiologic manifestations. The later might result from alterations in the 11AB region architecture. As shown in Fig. 3, *a*, the *agn^ts3*-like matroclinic mode of inheritance is pertinent to genes responsible for actin and microtubules-binding proteins with motor function and neurodevelopment. Noteworthy, the state of actin remodeling determining neurologic manifestations is a diagnostic feature of *agn^ts3* (Savvateeva-Popova et al., 2017).

The regions with FECs similar in reciprocal hybrids, but differing from parents, i. e. manifesting hybrid properties, contain a large set of genes responsible for motor functions.

Figure 3, *c* shows genes and biological processes for chromosomal regions forming ectopic contacts with X:11AB only in hybrids or mainly in hybrids compared to parents. These processes are pertinent to main manifestations of *agn^ts3*: meiosis, reparation, recombination; transcription factors; metabolism; proliferation; actin-binding proteins, microtubule-associated proteins.

Interestingly, in the cross *Berlin × agn^ts3* the chromosomal bands 8D, 12E, and 19D demonstrate FECs significantly exceeding these of parents. These bands contain genes involved in taste and odor perception and neurodevelopment, in particular of the mushroom bodies of the brain. The other examples of father strain manifestations might result from activities of trans-acting factors, such as miRNAs (Wittkopp et al., 2006).

miR-974 is involved in memory processes: its lowered expression impairs memory formation (Busto et al., 2015). Decrease in its content in olfactory neurons and the mushroom body V2 neurons promotes 3-hour memory. Noteworthy, the content of miR-974 is decreased both in *agn^ts3* and in progeny of *Berlin × agn^ts3* (impaired 3-hour memory) and is similar to wild type in the *agn^ts3 × Berlin* cross (normal memory). Likely, miR-974 might act as trans-acting factor presumed to regulate genes in patroclinic mode. The prevailing role of the paternal genome in memory formation is evident in *Canton S* and *agn^ts3* reciprocal hybrids (Vasiljeva et al., 2019).

Taken together, our data indicate that the *agnostic* locus might belong to the class of quantitative trait loci (QTL) (Qin et al., 2019).

Conclusion

One of the requirements of predictive and personalized medicine is consideration of POEs for prognosis of clinical phenotype of many multifactorial neuropsychiatric disorders. These different and individual manifestations of cognitive abilities and motor functions in patients with the same disease, i. e. behavioral plasticity, results from genome plasticity provoked by 3D chromatin architecture of the nerve cells nuclei. The evolutionary gene conservation approves the usage of simple low cost, fast and efficient models as Drosophila to probe the causes, consequences and mechanisms of pathology leading to human disease (Peffer et al., 2015). The Drosophila *agnostic* LIMK1 gene is a good candidate for linking the neuronal activity (spine remodeling, neurite outgrowth, trafficking of intracellular components, postsynaptic density functioning) and genetic apparatus (transcription machinery, chromatin-remodeling factors). Additionally, quite recent and unexpected findings (Davis, Zhong, 2017) reveal a new target of intellectual disabilities: learning acquisition and memory

erasure (forgetting) are governed by different signal cascades, correspondingly cAMP-dependent and actin remodeling cascade small GTPase Rac1 – LIMK1 (the key enzyme of actin remodeling LIM-kinase 1) and its phosphorylation substrate cofilin. The absence of Rac1-dependent forgetting causes the autistic spectrum disorders. Expression changes (active or non-active state) of LIMK1 and cofilin lead to different neurological disorders. Therefore, in the tradition of Russian genetic school (Lobashev et al., 1973), the *agnostic* gene might be a functional link between genetic and cytogenetic processes within the nervous system and serve as a model for elucidating both the maternal and paternal modes of transgenerational inheritance.

References

Belyaeva E.S., Zhimulev I.F., Volkova E.I., Alekseyenko A.A., Moshkin Y.M., Koryakov D.E. Su(UR)ES: a gene suppressing DNA underreplication in intercalary and pericentric heterochromatin of *Drosophila melanogaster* polytene chromosomes. *Proc. Natl. Acad. Sci. USA*. 1998;95(13):7532-7537. DOI 10.1073/pnas.95.13.7532.

Belyakin S.N., Christophides G.K., Alekseyenko A.A., Kriventseva E.V., Belyaeva E.S., Nanayev R.A., Makunin I.V., Kafatos F.C., Zhimulev I.F. Genomic analysis of *Drosophila* chromosome underreplication reveals a link between replication control and transcriptional territories. *Proc. Natl. Acad. Sci. USA*. 2005;102(23):8269-8274. DOI 10.1073/pnas.0502702102.

Bourassa M.W., Ratan R.R. The interplay between microRNAs and histone deacetylases in neurological diseases. *Neurochem. Int*. 2014; 77:33-39. DOI 10.1016/j.neuint.2014.03.012.

Busto G.U., Guven-Ozkan T., Fulga T.A., Van Vactor D., Davis R.L. microRNAs that promote or inhibit memory formation in *Drosophila melanogaster*. *Genetics*. 2015;200(2):569-580. DOI 10.1534/genetics.114.169623.

Collette J.C., Chen X.N., Mills D.L., Galaburda A.M., Reiss A.L., Bellugi U., Korenberg J.R. William's syndrome: gene expression is related to parental origin and regional coordinate control. *J. Hum. Genet*. 2009;54(4):193-198. DOI 10.1038/jhg.2009.5.

Davis R.L., Zhong Y. The biology of forgetting – a perspective. *Neuron*. 2017;95(3):490-503. DOI 10.1016/j.neuron.2017.05.039.

Dong X., Chen J., Li T., Zhang X., Zhang M., Song W., Zhao H., Lai J. Parent-of-origin-dependent nucleosome organization correlates with genomic imprinting in maize. *Genome Res*. 2018;28(7):1020-1028. DOI 10.1101/gr.230201.117.

Eagen K.P., Hartl T.A., Kornberg R.D. Stable chromosome condensation revealed by chromosome conformation capture. *Cell*. 2015; 163(4):934-946.

Guven-Ozkan T., Busto G.U., Schutte S.S., Cervantes-Sandoval I., O'Dowd D.K., Davis R.L. *MiR-980* is a memory suppressor microRNA that regulates the autism-susceptibility gene *A2bp1*. *Cell Rep*. 2016;14(7):1698-1709. DOI 10.1016/j.celrep.2016.01.040.

Iourov I.Y., Vorsanova S.G., Yurov Y.B. Molecular cytogenetics and cytogenomics of brain diseases. *Curr. Genomics*. 2008;9(7):452-465. DOI 10.2174/138920208786241216.

Iourov I.Y., Vorsanova S.G., Yurov Y.B. Pathway-based classification of genetic diseases. *Mol. Cytogenet*. 2019;12:4. DOI 10.1186/s13039-019-0418-4.

Ito S., Magalska A., Alcaraz-Iborra M., Lopez-Atalaya J.P., Rovira V., Contreras-Moreira B., Lipinski M., Olivares R., Martinez-Hernandez J., Ruszczycki B., Lujan R., Geijo-Barrientos E., Wilczynski G.M., Barco A. Loss of neuronal 3D chromatin organization causes transcriptional and behavioural deficits related to serotonergic dysfunction. *Nat. Commun*. 2014;5:4450. DOI 10.1038/ncomms 5450.

Jagannathan M., Warsinger-Pepe N., Watase G.J., Yamashita Y.M. Comparative analysis of satellite DNA in the *Drosophila melano-*

gaster species complex. *G3* (*Bethesda*). 2017;7(2):693-704. DOI 10.1534/g3.116.035352.

Kaiser-Rogers K., Rao K. Structural chromosome rearrangements. In: Gersen S.L., Keagle M.B. (Eds.). The Principles of Clinical Cytogenetics. Humana Press, Totowa, NJ., 2005;165-206. DOI 10.1385/1-59259-833-1:165.

Kaminskaya A.N., Nikitina E.A., Medvedeva A.V., Gerasimenko M.S., Chernikova D.A., Savvateeva-Popova E.V. The influence of the *limk1* gene polymorphism on learning acquisition and memory formation, pCREB distribution and aggregate formation in neuromuscular junctions in *Drosophila melanogaster*. *Russ. J. Genet*. 2015; 51(6):582-590. DOI 10.1134/S1022795415060071.

Kamyshev N.G., Iliadi K.G., Bragina J.V. *Drosophila* conditioned courtship: two ways of testing memory. *Learn. Mem*. 1999;6:1-20. DOI 10.1101/lm.6.1.1.

Kim S., Yu N.K., Shim K.W., Kim J.I., Kim H., Han D.H., Choi J.E., Lee S.W., Choi D.I., Kim M.W., Lee D.S., Lee K., Galjart N., Lee Y.S., Lee J.H., Kaang B.K. Remote memory and cortical synaptic plasticity require neuronal CCCTC-binding factor (CTCF). *J. Neurosci*. 2018;38(22):5042-5052. DOI 10.1523/JNEUROSCI.2738-17.2018.

Klages-Mundt N.L., Kumar A., Zhang Y., Kapoor P., Shen X. The nature of actin-family proteins in chromatin-modifying complexes. *Front. Genet*. 2018;9:398. DOI 10.3389/fgene.2018.00398.

Kolesnikova T.D. Banding pattern of polytene chromosomes as a representation of universal principles of chromatin organization into topological domains. *Biochemistry*. (*Mosc*). 2018;83(4):338-349.

Kozomara A., Birgaoanu M., Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res*. 2019;47(D1): D155-D162. DOI 10.1093/nar/gky1141.

Li K.L., Zhang L., Yang X.M., Fang Q., Yin X.F., Wei H.M., Zhou T., Li Y.B., Chen X.L., Tang F., Li Y.H., Chang J.F., Li W., Sun F. Histone acetyltransferase CBP-related H3K23 acetylation contributes to courtship learning in *Drosophila*. *BMC Dev. Biol*. 2018; 18(1):20. DOI 10.1186/s12861-018-0179-z.

Liehr T. From human cytogenetics to human chromosomics. *Int. J. Mol. Sci*. 2019;20(4):826. DOI 10.3390/ijms20040826.

Liu L., Li Q.Z., Jin W., Lv H., Lin H. Revealing gene function and transcription relationship by reconstructing gene-level chromatin interaction. *Comput. Struct. Biotechnol. J*. 2019;17:195-205. DOI 10.1016/j.csbj.2019.01.011.

Lobashev M.E., Ponomarenko B.B., Polyanskaya G.G., Tsapygina R.I. The role of nervous system in the regulation of genetic and cytogenetic processes. *Zh. Evol. Biokhim. Phiziol*. 1973;9:398-405. (in Russian)

Medrano-Fernández A., Barco A. Nuclear organization and 3D chromatin architecture in cognition and neuropsychiatric disorders. *Mol. Brain*. 2016;9(1):83. DOI 10.1186/s13041-016-0263-x.

Medvedeva A., Zhuravlev A., Savvateeva-Popova E. LIMK1, the key enzyme of actin remodeling bridges spatial organization of nucleus and neural transmission: from heterochromatin via non-coding RNAs to complex behavior. In: Cytoskeleton: Cell Movement, Cytokinesis and Organelles Organization. Nova Science Publishers, Inc., 2010;1: 37-67.

Menon D.U., Coarfa C., Xiao W., Gunaratne P.H., Meller V.H. siRNAs from an X-linked satellite repeat promote X-chromosome recognition in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*. 2014; 111(46):16460-16465. DOI 10.1073/pnas.1410534111.

Mohammed J., Bortolamiol-Becet D., Flynt A.S., Gronau I., Siepel A., Lai E.C. Adaptive evolution of testis-specific, recently evolved, clustered miRNAs in *Drosophila*. *RNA*. 2014;20(8):1195-1209. DOI 10.1261/rna.044644.114.

Nikitina E.A., Kaminskaya A.N., Molotkov D.A., Popov A.V., Savvateeva-Popova E.V. Effect of heat shock on courtship behavior, sound production, and learning in comparison with the brain content of LIMK1 in *Drosophila melanogaster* males with altered structure of the LIMK1 gene. *J. Evol. Biochem. Physiol*. 2014a;50(2):154-166. DOI 10.1134/S0022093014020082.

Nikitina E., Medvedeva A., Zakharov G., Savvateeva-Popova E. The *Drosophila agnostic* locus: involvement in the formation of cognitive defects in Williams Syndrome. *Acta Naturae*. 2014b;6(2): 53-61.

Nikitina E., Medvedeva A., Zakharov G., Savvateeva-Popova E. Williams syndrome as a model for elucidation of the pathway genes – the brain – cognitive functions: genetics and epigenetics. *Acta Naturae*. 2014c;6(1):9-22.

Nikitina E.A., Medvedeva A.V., Dolgaya Yu.F., Korochkin L.I., Pavlova G.V., Savvateeva-Popova E.V. Involvement of GDNF and LIMK1 and heat shock proteins in Drosophila learning and memory formation. *J. Evol. Biochem. Physiol*. 2012;48:529-539. DOI 10.1134/S0022093012050076.

Peffer S., Cope K., Morano K. Unraveling protein misfolding diseases using model systems. *Future Sci. OA*. 2015;1(2). DOI 10.4155/fso.15.41.

Pérez Jurado L.A., Peoples R., Kaplan P., Hamel B.C., Francke U. Molecular definition of the chromosome 7 deletion in Williams syndrome and parent-of-origin effects on growth. *Am. J. Hum. Genet*. 1996;59(4):781-792.

Qin H., Niu T., Zhao J. Identifying multi-omics causers and causal pathways for complex traits. *Front. Genet*. 2019;10:110. DOI 10.3389/fgene.2019.00110. PMID: 30847004. PMCID: PMC6393387.

Quinlan A.R., Hall I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26(6):841-842. DOI 10.1093/bioinformatics/btq033.

Robinson M.D., McCarthy D.J., Smyth G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139-140. DOI 10.1093/bioinformatics/btp616.

Savvateeva-Popova E.V., Zhuravlev A.V., Brázda V., Zakharov G.A., Kaminskaya A.N., Medvedeva A.V., Nikitina E.A., Tokmatcheva E.V., Dolgaya J.F., Kulikova D.A., Zatsepina O.G., Funikov S.Y., Ryazansky S.S., Evgen'ev M.B. *Drosophila* model for the analysis of genesis of LIM-kinase 1-dependent Williams–Beuren Syndrome cognitive phenotypes: INDELs, transposable elements of the Tc1/*mariner* superfamily and microRNAs. *Front. Genet*. 2017; 8:123. DOI 10.3389/fgene.2017.00123.

Schier A.F. The maternal-zygotic transition: death and birth of RNAs. *Science*. 2007;316(5823):406-407. DOI 10.1126/science.1140693. PMID: 17446392.

Sempere L.F., Sokol N.S., Dubrovsky E.B., Berger E.M., Ambros V. Temporal regulation of microRNA expression in *Drosophila melanogaster* mediated by hormonal signals and Broad-Complex gene activity. *Dev. Biol*. 2003;259(1):9-18. DOI 10.1016/s0012-1606(03)00208-2. PMID: 12812784.

Smrt R.D., Zhao X. Epigenetic regulation of neuronal dendrite and dendritic spine development. *Front. Biol*. (*Beijing*). 2010;5(4):304-323. DOI 10.1007/s11515-010-0650-0.

Soni K., Choudhary A., Patowary A., Singh A.R., Bhatia S., Sivasubbu S., Chandrasekaran S., Pillai B. miR-34 is maternally inherited in *Drosophila melanogaster* and *Danio rerio*. *Nucleic Acids Res*. 2013;41(8):4470-4480. DOI 10.1093/nar/gkt139.

Tokmacheva E.V. The mitotic activity of the cells in the head neural ganglion in larvae of the Drosophila ts mutant with an altered capacity for learning and augmented calmodulin activation properties. *Zh. Vyssh. Nerv. Deiat. Im. I.P. Pavlova*. 1995;45(3):565-571. (in Russian)

Vasiljeva S., Tokmacheva E., Medvedeva A., Ermilova A., Nikitina E., Shchegolev B., Surma S.V., Savvateeva-Popova E.V. Parent-of-origin effect in genetic instability of somatic brain's cells of Drosophila and memory formation under normal and stress conditions. *Tsitologiia*. 2019;61(12):951-963. DOI 10.1134/S0041377119120071. (in Russian)

Waring G.L., Pollack J.C. Cloning and characterization of a dispersed, multicopy, X chromosome sequence in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*. 1987;84(9):2843-2847. DOI 10.1073/pnas.84.9.2843.

А.В. Медведева, Е.В. Токмачева, А.Н. Каминская …
Г.А. Захаров, О.Г. Зацепина, Е.В. Савватеева-Попова

Роль родительского происхождения аллелей
в организации хромосом

2021
25•5

Wittkopp P.J., Haerum B.K., Clark A.G. Parent-of-origin effects on mRNA expression in *Drosophila melanogaster* not caused by genomic imprinting. *Genetics*. 2006;173(3):1817-1821. DOI 10.1534/genetics.105.054684.

Zakharov G. Locotrack. https://github.com/GennadiyZakharov/locotrack. 2017. (Accessed April 3, 2020).

Zakharov G.A., Zhuravlev A.V., Payalina T.L., Kamyshov N.G., Savvateeva-Popova E.V. The effect of mutations of the kynurenine pathway of tryptophan metabolism on locomotor behavior and gene expression in glutamatergic and cholinergic systems of *D. melanogaster*. *Russ. J. Genet. Appl. Res.* 2012;2:197-204. DOI 10.1134/S2079059712020141.

Zayats T., Johansson S., Haavik J. Expanding the toolbox of ADHD genetics. How can we make sense of parent of origin effects in ADHD and related behavioral phenotypes? *Behav. Brain Funct.* 2015;11(1):33. DOI 10.1186/s12993-015-0078-4.

Zerbino D.R., Achuthan P., Akanni W., Amode M.R., Barrell D., Bhai J., Flicek P. Ensembl 2018. *Nucleic Acids Res.* 2018;46:D754-D761. DOI 10.1093/nar/gkx1098.

Zhuravlev A. Homology Segment Analysis. Available at: https://bitbucket.org/beneor/homology-segment-analysis/src/master/ 2019a. (Accessed September 3, 2019).

Zhuravlev A. Homology Segment Analysis protocol. protocols.io. 2019b. DOI 10.17504/protocols.io.bakyicxw.

Zink F., Magnusdottir D.N., Magnusson O.T., Walker N.J., Morris T.J., Sigurdsson A., Halldorsson G.H., Gudjonsson S.A., Melsted P., Ingimundardottir H., Kristmundsdottir S., Alexandersson K.F., Helgadottir A., Gudmundsson J., Rafnar T., Jonsdottir I., Holm H., Eyjolfsson G.I., Sigurdardottir O., Olafsson I., Masson G., Gudbjartsson D.F., Thorsteinsdottir U., Halldorsson B.V., Stacey S.N., Stefansson K. Insights into imprinting from parent-of-origin phased methylomes and transcriptomes. *Nat. Genet.* 2018;50(11):1542-1552. DOI 10.1038/s41588-018-0232-7.

Zykova T.Y., Levitsky V.G., Belyaeva E.S., Zhimulev I.F. Polytene chromosomes – a portrait of functional organization of the *Drosophila* genome. *Curr. Genomics*. 2018;9(3):179-191. DOI 10.2174/1389202918666171016123830.

ORCID ID

A.V. Medvedeva orcid.org/0000-0001-7989-8746
E.V. Tokmatcheva orcid.org/0000-0002-2864-9894
S.V. Vasileva orcid.org/0000-0002-7785-7091
E.A. Nikitina orcid.org/0000-0003-1897-8392

A.V. Zhuravlev orcid.org/0000-0003-2673-4283
G.A. Zakharov orcid.org/0000-0003-4850-3777
O.G. Zatsepina orcid.org/0000-0001-5982-941X
E.V. Savvateeva-Popova orcid.org/0000-0002-6925-4370

Variability of mitochondrial DNA D-loop sequences in Zabaikalskaya horse breed

L.A. Khrabrova[1] ✉, N.V. Blohina[1], B.Z. Bazaron[2], T.N. Khamiruev[2]

[1] All-Russian Research Institute for Horse Breeding, Divovo, Ryazan Region, Russia
[2] Scientific Research Institute of Veterinary Medicine of Eastern Siberia – Branch of the Siberian Federal Scientific Centre of Agro-BioTechnologies of the Russian Academy of Sciences, Chita, Russia
✉ l.khrabrova@yandex.ru

Abstract. The Zabaikalskaya horse is an indigenous breed of horses from Siberia with diverse use. It is characterized by endurance and good adaptability to year-round herd maintenance in the harsh conditions of the Baikal steppes. To determine the genetic characteristics of the maternal lineage of the Zabaikalskaya horse breed based on mitochondrial DNA polymorphisms, we collected hair samples from 31 horses belonging to breeding farms in the Trans-Baikal Territory. Analysis of the 530 bp sequence of the mtDNA D-loop was performed using the maximum composite likelihood (MCL) model in combination with bootstrap analysis. When studying the polymorphism of the hypervariable region of the mtDNA D-loop in Zabaikalskaya horses, we identified 31 haplotypes representing 8 haplogroups: B, C, G, H, L, M, Q and R according to modern classification. The sequenced fragment of the D-loop from nucleotide position 15471 to 16000 contained 17 polymorphic sites, mainly represented by the A→G, G→A and T→C transitions. The haplogroups Q (25.81 %), B (19.35 %), G (16.13 %) and H (12.90 %) were prevailing in the mtDNA structure of this breed. Genetic analysis of the mitochondrial genome of the Zabaikalskaya horse revealed a high level of diversity of haplotypes and haplogroups, which are typical for the horse populations of Eurasia.
Key words: genetic diversity; haplogroups mtDNA; horse; phylogenetic analysis; Zabaikalskaya breed.

Вариабельность последовательности D-петли митохондриальной ДНК у лошадей забайкальской породы

Л.А. Храброва[1] ✉, Н.В. Блохина[1], Б.З. Базарон[2], Т.Н. Хамируев[2]

[1] Всероссийский научно-исследовательский институт коневодства, пос. Дивово, Рыбновский район, Рязанская область, Россия
[2] Научно-исследовательский институт ветеринарии Восточной Сибири – филиал Сибирского федерального научного центра агробиотехнологий Российской академии наук, Чита, Россия
✉ l.khrabrova@yandex.ru

Аннотация. Забайкальская лошадь – аборигенная порода лошадей Сибири универсального использования. Она характеризуется выносливостью и хорошей приспособленностью к круглогодичному табунному содержанию в суровых условиях байкальских степей. Для определения генетических особенностей материнских линий забайкальской породы на основе полиморфизмов митохондриальной ДНК мы собрали образцы волос от 31 лошади, принадлежащей племенным хозяйствам Забайкальского края. Анализ последовательности D-петли мтДНК размером 530 п.н. проводили с использованием модели максимального правдоподобия (MCL) в сочетании с бутстрэп-анализом. При изучении полиморфизма гипервариабельного региона D-петли мтДНК у забайкальских лошадей был выявлен 31 гаплотип, представляющий восемь гаплогрупп: B, C, G, H, L, M, Q и R, согласно современной классификации. Секвенированный фрагмент D-петли (нуклеотидная позиция 15471–16000) содержал 17 полиморфных сайтов, в основном представленных транзициями A→G, G→A и T→C. В структуре мтДНК породы преобладали гаплогруппы Q (25.81 %), B (19.35 %), G (16.3 %) и H (12.90 %). Генетический анализ митохондриального генома забайкальской лошади выявил высокий уровень разнообразия гаплотипов и гаплогрупп, которые типичны для популяции лошадей Евразии.
Ключевые слова: генетическое разнообразие; гаплогруппы мтДНК; лошадь; филогенетический анализ; забайкальская порода.

Introduction

Zabaikalskaya horse is one of the native horse breeds from Siberia and it is distributed in the territories located to the south and southeast of Lake Baikal. The breed was formed through process of long interbreeding of Mongolian horses and later representatives of cultural breeds brought by Russian settlers in the 17th and 18th centuries (Karaush, 1952; Khakimov et al., 2002). As a result, the local population of horses in the Trans-Baikal Territory was significantly improved, resulting in a relatively large array of horses, versatile in their use and well

Л.А. Храброва, Н.В. Блохина
Б.З. Базарон, Т.Н. Хамируев

Вариабельность последовательности D-петли
митохондриальной ДНК у лошадей забайкальской породы

2021
25•5

adapted to year-round herd maintenance. The Zabaikalskaya horse is characterized by low growth rate but with a massive and bony appearance. In 1993, these horses were given the status of the Zabaikalskaya breed, and it was included in the State Register of Breeding Achievements of the Russian Federation (Khamiruev et al., 2014).

Today, modern genetics technology has allowed researchers to unravel some of the mysteries surrounding the development of Russian native breeds. The study of features of nuclear and mitochondrial DNA of horses of different breeds and areas, including the found remains of ancient horses, allowed clarifying of many important questions of Equids evolution. In particular, high variability of mitochondrial DNA was revealed, indicating the presence of multiple wild ancestors in domestic horses and the existence of different regions of domestication (Bowling, Ruvinski, 2000; Jansen et al., 2002). Due to the study of the mitochondrial genome, it was finally established that the species *Equus caballus* is not a direct descendant of the wild horse *E. ferus przewalski*.

The study of the mitochondrial genome of the horse began with the work of X. Xu and U. Arnason (1994), which led to the sequencing of equine mtDNA and demonstrated that its configuration varies due to the different number of GTGCACCT repeats in the control region. Further studies have shown that mtDNA polymorphism can be studied using various technologies, among which the method of direct sequencing of the hyper variable region of the D-loop is most often practiced. This approach allows to study matrilineal diversity within horse breeds and their phylogenetic relationships (Bowling et al., 2000; Hill et al., 2002; Lopes et al., 2006; McGahern et al., 2006; Glazewska et al., 2007; Moridi et al., 2012; Vilstrup et al., 2013).

Comparative analysis of mitochondrial DNA in different populations and breeds of horses in Europe and Asia and phylogenetic reconstruction made on its basis showed the presence of complex variability of mitochondrial haplogroups; this was not observed in other domesticated species (McGahern et al., 2006). Comparison of European and Asian horse breeds revealed differentiated distribution of mtDNA haplogroup variants and evidence of a biogeographic wedge in Asian populations, including association of "Eastern" mtDNA types with haplogroup F. A number of previously unknown additional mtDNA sequences were obtained from horses of Akhal-Teke, Vyatskaya, Mezenskaya, Orlov Trotter and Yakutskaya breeds from Russia, among which the greatest similarity with the European populations was observed in the Mezenskaya horse (McGahern et al., 2006).

A new stage of studying the mitochondrial genome of the horse began in 2012, when a team of researchers led by A. Achilli conducted a complete sequencing of 83 mitochondrial genomes of modern horses in Europe, Asia, the Middle East and America. The phylogenetic analysis with high molecular resolution revealed 18 main haplogroups (A–R) with their diagnostic mutational motifs that arose during the Neolithic period. The researchers concluded that the proposed classification of the encoded and control regions of the mitochondrial genome could be used in the study of the remains of ancient horses, phylogenetic relations of modern breeds, intra-breed diversity and evaluation of the possible connection of mtDNA with racing performance.

Subsequent studies of mitochondrial DNA polymorphism in horses of different populations showed that the number of haplogroups in breeds can vary from 4 to 14 (Bigi et al., 2014; Sorokin, 2015; Cardinali et al., 2016; Khrabrova et al., 2020). The study of mitochondrial sequence variation in 251 Arabian horses from different countries showed that they belong to 13 mtDNA haplogroups based on Achilli (Khanshour, Cothran, 2013). The greatest number of mtDNA haplotypes and haplogroups was revealed in Arab horses of the Syrian population, that is, in the region of creation of this breed. Arabian horses had a much wider range of mitochondrial haplotypes compared to Thoroughbred, with a high frequency of haplogroup L in all populations.

When studying the mitochondrial genome of horses, researchers found a clear association of dam lines with certain mtDNA haplogroups and haplotypes (Hill et al., 2002; Lopes et al., 2006; Sorokin, 2015; Khrabrova et al., 2019). This allows, if necessary, to control the female line of horses using the method of mtDNA sequencing (Bowling et al., 2000; Khrabrova et al., 2020).

The task of our research was to study the variability of the D-loop sequence of mtDNA and maternal lineage in the native Zabaikalskaya horse breed, bred in the steppes of Transbaikalia.

Material and methods
To analyze the 530 bp hypervariable region of mtDNA D-loop we sequenced 31 Zabaikalskaya horses from Chita State Stable (*n* = 13) and Kalinin stud (*n* = 18) located in the Trans-Baikal Region. DNA was isolated from hair follicles using a set of ExtraGene DNA Prep 200, produced by laboratory "Isogen" (Moscow), according to the manufacturer's instructions. The original primers for amplification of the studied mtDNA D-loop site were selected by S. Sorokin (2015) taking into account the reference sequence of the fossil Swedish horse X79547 (Xu, Arnason, 1994).

For PCR amplification we used the composition of the reaction mixture included 0.2 mM dNTP, 0.5 µM of each primer, 2.5 mM MgCl$_2$, 1× PCR buffer, 1 unit Taq polymerase (PE Applied Biosystems, USA) and 1 unit AmpliTaq Gold polymerase (PE Applied Biosystems), 50 ng DNA. The procedure involved heating the reaction mixture at 95 °C (5 min), followed by 30 cycle's denaturation at 94 °C (30 s), annealing at 55 °C (30 s) and elongation at 72 °C (45 s). Final elongation was carried out at 72 °C for 2 min. The sequencing of the PCR fragment from agarose gel was performed using the Bigdye Terminator Cycle Sequencing Kit (PE Applied Biosystems) on the ABI 3130xl genetic analyzer (PE Applied Biosystems) using manufacturer's protocol.

All mtDNA sequences were prepared with BioEdit 7.2.1. Additional GenBank data set of 18 known haplogroups (A–R) under the number JN398377–JN398457 was used to identify the obtained nucleotide sequences. The phylogenetic analysis of the mtDNA D-loop, including a 530 bp fragment (position 15471–16000) was performed using the Neighbor-Joining (NJ) method and program MEGA 7 (www.megasoftware.net). The statistical confidence of each node was estimated by 1000 random bootstrap runs. When constructing a phylogenetic tree, bootstrap values of more than 50 % were taken into account. For comparison of mtDNA haplogroups

distribution in Zabaikalskaya horse, data from 156 Kabardinian, 64 Mongolian (Khaudov et al., 2018), 74 Thoroughbred (Khrabrova et al., 2019) and 22 Vyatskaya horses (Khrabrova et al., 2020) were used.

Results

Sequence analysis of the 530 bp region of the mitochondrial D-loop of Zabaikalskaya horses showed the presence of 31 haplotypes corresponding to 8 haplogroups: B, C, G, H, L, M, Q and R according to Achilli's classification (2012). The sequenced fragment of the D-loop from 15471 to 16000 nucleotide position contained 17 polymorphic sites, mainly represented by the A-G, G-A and T-C transitions (Table 1).

The average number of substitutions per site was 0.03 ± 0.022, indicating a relatively high level of nucleotide diversity.

Sequences with haplogroups B, G, M and Q were found in both subpopulations of Zabaikalskaya horses. In addition, mtDNA haplogroups C, L and R were found in horses of the Chita State Stable, while haplogroup H was found only in the Kalinin stud. The genetic structure of Zabaikalskaya horse was dominated by haplogroups Q (25.81 %), B (19.35 %), G (16.3 %) and H (12.90 %), which is clearly demonstrated

Table 1. Variability of nucleotides in a 530 bp fragment of the mtDNA D-loop of Zabaikalskaya horse haplotypes compared with reference sequence X79547

ID	15494	15585	15597	15602	15603	15604	15615	15616	15617	15649	15650	15720	15956	17703	17720	17771	17777
X79547	T	G	A	C	T	G	A	A	T	A	A	G	A	T	G	C	A
126ZAB	C	A		T	C						G	A			A	T	
127ZAB	C	A		T	C						G	A			A	T	
128ZAB	C	A		T	C						G	A			A	T	
129ZAB	C	A		T	C						G	A			A	T	
130ZAB	C			T					C			A	G		A	T	
131ZAB	C										G	A			A		
132ZAB	C										G	A			A		
133ZAB	C										G	A			A		
134ZAB	C										G	A			A		
135ZAB	C			T		A						A	G	C	A	T	G
136ZAB	C	A		T		A						A	G	C	A	T	G
137ZAB	C	A	G	T							G	A		C	A		
138ZAB	C			T		A						A	G	C	A	T	G
139ZAB	C	A	G	T							G	A		C	A		
140ZAB	C	A	G	T							G	A		C	A		
141ZAB	C										G	A			A		
142ZAB	C	A		T		A						A	G	C	A	T	G
143ZAB	C	A		T		A						A	G	C	A	T	G
Chardash	C			T			G	G				A	G	C	A		
Praga	C			T					C			A	G		A	T	
Zabayka	C	A		T		A						A	G	C	A	T	G
Arbalet	C			T			G	G				A	G	C	A		
Chegrash	C		G	T							G	A	G		A		
149ZAB	C	A	G	T							G	A		C	A		
150ZAB	C			T					C			A	G		A	T	
151ZAB	C										G	A			A		
152ZAB	C		G	T							G	A		C	A		
153ZAB	C	A		T	C	A				G		A	G		A	T	
154ZAB	C	A		T		A						A	G	C	A	T	G
155ZAB	C			T		A						A	G	C	A	T	G
156ZAB	C		G	T							G	A	G		A		

Л.А. Храброва, Н.В. Блохина
Б.З. Базарон, Т.Н. Хамируев

Вариабельность последовательности D-петли
митохондриальной ДНК у лошадей забайкальской породы

2021
25•5
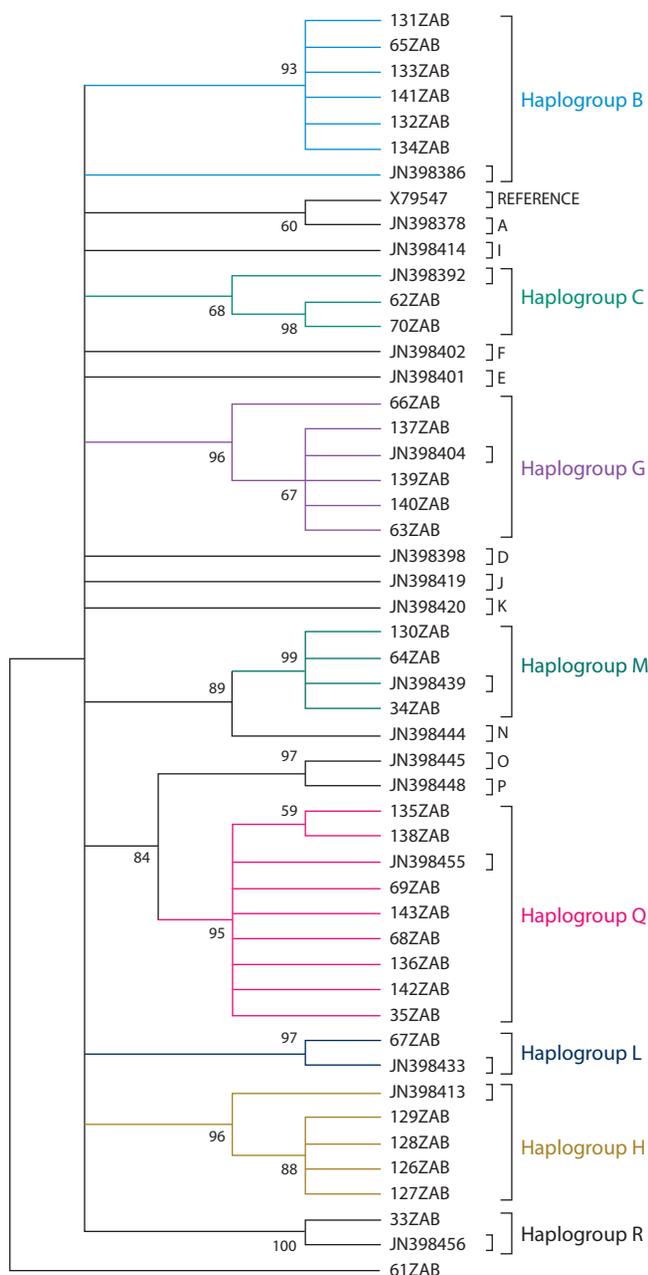
Fig. 1. Phylogenetic tree of mtDNA D-loop sequences from haplotypes of Zabaikalskaya horses constructed by use of the Neighbor-Joining method in combination with bootstrap analysis (bootstrap value > 50 are shown at nodes).

For identification of haplogroups were used GenBank data (JN398377–JN398457) and their classification according to (Achilli et al., 2012).

Fig. 2. Scheme of evolutionary relationships between Zabaikalskaya horses from different subpopulations (Chita State Stable – blue, Kalinin stud – red) derived from phylogenetic analysis using Neighbor-Joining method and MEGA7 program.

by the dendrograms in Fig. 1 and 2. Haplotype Q, which combines horses from different farms, is the most common in Asia and the Middle East (Achilli et al., 2012; Khanshour, Cothran, 2013).

The statistical analysis of phylogenetic relationships calculated by the MCL distances shows a fairly high level of bootstrap value (94–100) matching of haplotypes within haplogroups (see Fig. 1). The genetic divergence of the mitochondrial genome of Zabaikalskaya horses from different subpopulations is shown in Figure 2.

Most haplotypes of Zabaikalskaya horse (74.19 %) are found in the haplogroups B, G, H and Q; haplogroup L was rarely found, which is typical of many cultural and local breeds (Table 2).

Comparative analysis of mtDNA in horses of different breeds bred in Russia shows that, as a rule, they have a multilinear origin on the maternal line. At the same time, haplogroups B and L were represented quite significantly in all breeds (5.41–36.36 and 3.23–50.0 %, respectively). The genetic analysis of mtDNA of Zabaikalskaya horse breed indicates a high level of diversity of haplotypes and a clear differentiation of I maternal line in the subpopulation of this breed.

Discussion

The high level of diversity of the genetic structure of mitochondrial DNA and maternal inheritance make it a unique object for the study of evolutionary processes, phylogenetic analysis and evaluation of population diversity (Lukashov, 2009). We found a high mitochondrial polymorphism represented by 31 D-loop haplotypes from 8 haplogroups: B, C, G, H, L, M, Q and R based upon Achilli et al. (2012). In addition, haplogroups A (12.5 %) and D (8.3 %) were identified in the sequences of 24 Zabaikalskaya horse presented in GenBank (Khaudov et al., 2018).

The genetic structure of Zabaikalskaya horse breed shown by mtDNA D-loop haplogroups is generally typical for horse populations in Asia, in which haplogroups B, C, G, and Q were most common (Achilli et al., 2012; Khanshour, Cothran, 2013; Khaudov et al., 2018). It is obvious that the Mongolian horse had a great influence on the formation of Zabaikalskaya horse. The haplogroup variants obtained from the Zabaikalskaya horse were more similar to other local breeds from Siberia (Voronkova, Stolpovskiy, 2018), which indicates their common matrilineal genealogy.

Table 2. Distribution of the haplogpoups mtDNA (in %) in different horse breeds

Haplogroup mtDNA	Zabaikalskaya (n = 31)	Kabardian* (n = 156)	Vyatskaya (n = 22)	Thoroughbred (n = 74)	Mongolian* (n = 64)
A	0.00	6.50	4.55	4.05	6.20
B	19.35	11.00	36.36	5.41	11.00
C	6.45	5.20	0.00	0.00	6.20
D	0.00	2.60	0.00	0.00	3.10
E	0.00	6.50	0.00	0.00	6.20
F	0.00	0.00	0.00	0.00	0.00
G	16.13	19.50	0.00	9.46	3.10
H	12.90	0.00	0.00	2.70	1.60
I	0.00	7.20	0.00	16.21	0.00
J	0.00	3.30	0.00	0.00	0.00
K	0.00	0.00	0.00	0.00	0.00
L	3.23	12.30	31.82	50.00	17.20
M	9.68	5.20	4.55	5.41	11.00
N	0.00	1.90	4.55	6.76	0.00
O–P	0.00	5.80	4.55	0.00	12.50
Q	25.81	11.70	4.55	0.00	18.80
R	6.45	1.30	0.00	0.00	3.10

* Date from A.D. Khaudov et al. (2018).

A comparative analysis of polymorphism of microsatellite loci in local horse breeds in Russia showed that Zabaikalskaya horse is characterized by a relatively high level of genetic diversity and has a high level of similarity with Buryat horse (Khrabrova, 2015). Mitochondrial analysis confirms the close relationship of these neighboring breeds, which share 6 common haplogroups: B, G, L, M, Q and R.

Comparison of European and Asian horse breeds revealed differentiated distribution of mtDNA haplogroup variants and evidence of a biogeographic wedge in Asian populations, including association of "Eastern" mtDNA types with haplogroups Q and R. Interestingly, these haplogroups were identified in horses of the Kabardian and Vyatskaya horse breeds bred in the European part of the Russian Federation (Khaudov et al., 2018; Khrabrova et al., 2020), but were absent in the Thoroughbred and Cleveland Bay horses created in England (Khrabrova et al., 2019; Dell et al., 2020).

The study of the sequence of the hypervariable fragment of the mtDNA D-loop makes it possible to assess the interbreed diversity of horses along maternal lines. The data of many researchers confirm the evidence for biogeographic patterning of mtDNA sequences in Eastern horse populations (McGahern et al., 2006; Khanshour, Cothran, 2013; Khaudov et al., 2018).

Conclusion
The mtDNA analysis of Zabaikalskaya horse identified 31 different haplotypes clustered in 8 haplogroups (B, C, G, H, L, M, Q and R), which indicates the genetic diversity of the maternal ancestry lines in this breed. Zabaikalskaya horses show high frequencies in haplogroups Q, B, G and H; these haplogroups are also the most common in native horse breeds of Siberia and Mongolian horses. The high variability level

of the mitochondrial genome in horses of local breeds may determine good adaptive qualities. The obtained data allow us to supply important information about the genetic features of the existing maternal structure of the Zabaikalskaya horse breed.

References
Achilli A., Olivieri A., Soares P., Landoni H., Kashani B.H., Perego U.A., Nergadze S.G., Carossa V., Santagostino M., Capomaccio S., Felicetti M., Al-Achkar W., Penedo M.C.T., Verini-Supplizi A., Houshmand M., Woodward S.R., Semino O., Silvestrelli M., Giulotto E., Pereira L., Bandelt H.-J., Torroni A. Mitochondrial genomes from modern horses reveal the major haplogroups that underwent domestication. *Proc. Natl. Acad. Sci. USA.* 2012;109:7:2449-2454. DOI 10.1073/pnas.1111637109.

Bigi D., Perrota G., Zambonelli P. Genetic analysis of seven Italian horse breeds on mitochondrial DNA D-loop variation. *Anim. Genet.* 2014;45(4):593-595. DOI 10.1111/age.12156.

Bowling A.T., Del Valle A., Bowling M. A pedigree-based study of mitochondrial D-loop DNA sequence variation among Arabian. *Anim. Genet.* 2000;31(1):1-7.

Bowling A.T., Ruvinski A. The Genetics of the Horse. Wallingford: CABI Publishing, 2000.

Cardinali I., Lancioni H., Giontella A., Capodiferro M.R., Capomaccio S., Buttazzoni L., Biggio G.P., Cherchi R., Albertini E., Olivieri A., Cappelli K., Achilli A., Silvestrelli M. An overview of ten Italian horse breeds through mitochondrial DNA. *PLoS One.* 2016; 11(4):e0153004. DOI 10.1371/journal.pone.0153004.

Dell A.C., Curry M.C., Yarnell K.M., Starbuck G.R., Wilson Ph.B. Mitochondrial D-loop sequence variation and maternal lineage in the endangered Cleveland Bay horse. *PLoS One.* 2020;15(12): e0243247. DOI 10.1371/journal.pone.02443247.

Glazewska I., Wysocka A., Gralak B., Sell J. A new view on dam lines in Polish Arabian horses based on mtDNA analysis. *Genet. Sel. Evol.* 2007;39(5):609-619. DOI 10.1051/gse:2007025.

Л.А. Храброва, Н.В. Блохина
Б.З. Базарон, Т.Н. Хамируев

Вариабельность последовательности D-петли
митохондриальной ДНК у лошадей забайкальской породы

2021
25•5

Hill E.W., Bradley M., Al-Barody M., Ertugrul O., Splan R.K., Zakharov I., Cunningham E.P. History and integrity of Thoroughbred dam lines revealed in equine mtDNA variation. *Anim. Genet.* 2002; 33(4):287-294.

Jansen T., Foster P., Levine M.A., Oelke H., Hurles M., Renfrew C., Weber C., Olek K. Mitochondrial DNA and the origin of the domestic horse. *Proc. Natl. Acad. Sci. USA.* 2002;99(16):10905-10910. DOI 10.1073/pnas.152330099. PMCID: PMC125071.

Karaush O.N. Zabaikalskaya horse. In: The Book about Horse. Moscow: Selkhozgiz Publ., 1952;1:564-568. (in Russian)

Khakimov N.V., Pankova T.Yu., Vinogradov I.I. To the question of Zabaikalskaya horse. In: Scientific Support for Sustainable Development of the Agro-industrial Complex of East Transbaikalia: Proceedings of the International Scientific-Practical Conference. Chita, 2002;II:102-104. (in Russian)

Khamiruev T.N., Bazaron B.Z., Kalashnikov R.V. Some biological features of the Transbaikal horse. *Konevodstvo i Konny Sport = Horse Breeding and Equestrian Sports.* 2014;4:20-22. (in Russian)

Khanshour A.M., Cothran E.G. Maternal phylogenetic relationships and genetic variation among Arabian horse populations using whole mitochondrial DNA D-loop sequencing. *BMC Genet.* 2013;14:83. DOI 10.1186/1471-2156-14-83. PMCID: PMC3847362.

Khaudov A.D., Duduev A.S., Kokov Z.A., Amshokov K.K., Zhekamukhov M.Kh., Zaitsev A.M., Reissmann M. Genetic analysis of maternal and paternal lineages in Kabardian horses by uniparental molecular markers. *Open Vet. J.* 2018;8(1):40-46. DOI 10.4314/ovj. v8iI.7.

Khrabrova L.A. Characterization of Genetic Horse Breeding Resources in Russia. Lap Lambert Acad. Publ., 2015.

Khrabrova L.A., Zaitsev A.M., Kalashnikov V.V., Blohina N.V., Belousova N.F., Sorokin S.I. Structure of Vyatskaya horse breed by haplogroups mtDNA. *Konevodstvo i Konny Sport = Horse Breeding and Equestrian Sports.* 2020;4:4-6. DOI 10.25727/HS.2020.4.62190. (in Russian)

Khrabrova L.A., Zaitsev A.M., Vikulova L.L., Adamkovskaya M.V., Blohina N.V., Sorokin S.I. MtDNA haplotype analysis in dam families of the Thoroughbred riding horses. In: Modern Trends in Agri-cultural Production in the World Economy: Proceedings of XVIII International Scientific and Practical Conference. Kemerovo, Russia, 2019;34-42. DOI 10.32743/kuz.agri.2020.34-42.

Lopes M.S., Mendonca D., Cymbron T., Valera M., da Costa-Ferreira J., da Câmara Machado A. The Lusitano horse maternal lineage based on mitochondrial D-loop sequence variation. *Anim. Genet.* 2006; 36(3):196-202. DOI 10.1111/j.1365-2052.2005.01279.x.

Lukashov V.V. Molecular Evolution and Phylogenetic Analysis. Moscow: Binom. Laboratoriya Znanij Publ., 2009. (in Russian)

McGahern A., Bower M.A., Edwards C.J., Brophy P.O., Sulimova G., Zakharov I., Vizuete-Forster M., Levine M., Li S., MacHugh D.E., Hil E.W. Evidence for biogeographic patterning of mitochondrial DNA in Eastern horse populations. *Anim. Genet.* 2006;37(5):494-497. DOI 10.1111/j.1365-2052.2006.01495.x.

Moridi M., Masoudi A.A., Vaez Torshizi R., Hill E.W. Mitochondrial DNA D-loop sequence variations in maternal lineages of Iranian native horses. *Anim. Genet.* 2012;44(2):209-213. DOI 10.1111/j.1365-2052.2012.02389.x.

Sorokin S.I. Molecular and genetic analysis of D-loop mitochondrial DNA of blood families of Vladimir breed. *Konevodstvo i Konny Sport = Horse Breeding and Equestrian Sports.* 2015;6:27-29. (in Russian)

Vilstrup J.T., Seguin-Orlando A., Stiller M., Ginolhac A., Raghavan M., Nielsen S.C.A., Weinstock J., Froese D., Vasiliev S.K., Ovodov N.D., Clary J., Helgen K.M., Fleischer R.C., Cooper A., Shapiro B., Orlando L. Mitochondrial phylogenomics of modern and ancient equids. *PLoS One.* 2013;8(2):e55950. DOI 10.1371/journal.pone.0055950. PMCID: PMCPMC3577844.

Voronkova V.N., Stolpovskiy Yu.A. Assessment of the genetic diversity of native breeds of the Sayano-Altai region using nuclear and mitochondrial DNA markers. In: Aboriginal Horse Breeding in Russia: History, Present, Prospects: Proceedings of II All-Russian Res. Pract. Conference. Arkhangelsk, 2018;60-69. (in Russian)

Xu X., Arnason U. The complete mitochondrial DNA sequence of the horse. *Equus caballus*: extensive heteroplasmy of control region. *Gene.* 1994;1480:357-362.

ORCID ID

L.A. Khrabrova orcid.org/0000-0003-2590-8472
N.V. Blohina orcid.org/0000-0001-7406-6385
B.Z. Bazaron orcid.org/0000-0001-7525-5815
T.N. Khamiruev orcid.org/0000-0002-0147-2929

Analysis of the structure and function of the tomato *Solanum lycopersicum* L. MADS-box gene *SlMADS5*

A.V. Nezhdanova, M.A. Slugina, E.A. Dyachenko, A.M. Kamionskaya, E.Z. Kochieva, A.V. Shchennikova ✉

Institute of Bioengineering, Federal Research Centre "Fundamentals of Biotechnology" of the Russian Academy of Sciences, Moscow, Russia
✉ shchennikova@yandex.ru

Abstract. At all stages of flowering, a decisive role is played by the family of MADS-domain transcription factors, the combinatorial action of which is described by the ABCDE-model of flower development. The current volume of data suggests a high conservatism of ABCDE genes in angiosperms. The E-proteins SEPALLATA are the central hub of the MADS-complexes, which determine the identity of the floral organs. The only representative of the *SEPALLATA3* clade in tomato *Solanum lycopersicum* L., *SlMADS5*, is involved in determining the identity of petals, stamens, and carpels; however, data on the functions of the gene are limited. The study was focused on the *SlMADS5* functional characterization. Structural and phylogenetic analyses of SlMADS5 confirmed its belonging to the SEP3 clade. An *in silico* expression analysis revealed the absence of gene transcripts in roots, leaves, and shoot apical meristem, and their presence in flowers, fruits, and seeds at different stages of development. Two-hybrid analysis showed the ability of SlMADS5 to activate transcription of the target gene and interact with TAGL1. Transgenic plants *Nicotiana tabacum* L. with constitutive overexpression of *SlMADS5* cDNA flowered 2.2 times later than the control; plants formed thickened leaves, 2.5–3.0 times thicker stems, 1.5–2.7 times shortened internodes, and 1.9 times fewer flowers and capsules than non-transgenic plants. The flower structure did not differ from the control; however, the corolla petals changed color from light pink to magenta. Analysis of the expression of *SlMADS5* and the tobacco genes *NtLFY*, *NtAP1*, *NtWUS*, *NtAG*, *NtPLE*, *NtSEP1*, *NtSEP2*, and *NtSEP3* in leaves and apexes of transgenic and control plants showed that *SlMADS5* mRNA is present only in tissues of transgenic lines. The other genes analyzed were highly expressed in the reproductive meristem of control plants. Gene transcripts were absent or were imperceptibly present in the leaves and vegetative apex of the control, as well as in the leaves and apexes of transgenic lines. The results obtained indicate the possible involvement of *SlMADS5* in the regulation of flower meristem development and the pathway of anthocyanin biosynthesis in petals.

Key words: *Solanum lycopersicum*; *Nicotiana tabacum*; heterologous gene expression; MADS-domain transcription factors; SEPALLATA; SlMADS5.

Структурно-функциональный анализ MADS-box гена *SlMADS5* томата *Solanum lycopersicum* L.

А.В. Нежданова, М.А. Слугина, Е.А. Дьяченко, А.М. Камионская, Е.З. Кочиева, А.В. Щенникова ✉

Институт биоинженерии, Федеральный исследовательский центр «Фундаментальные основы биотехнологии» Российской академии наук, Москва, Россия
✉ shchennikova@yandex.ru

Аннотация. На всех этапах цветения решающую роль играет семейство MADS-доменных транскрипционных факторов, комбинаторное действие которых описывается ABCDE-моделью развития цветка. Современный объем данных позволяет говорить о высоком консерватизме ABCDE-генов у покрытосеменных. Е-белки SEPALLATA являются центральным звеном MADS-комплексов, определяющих идентичность цветковых органов. Единственный представитель клады *SEPALLATA3* томата *Solanum lycopersicum* L., *SlMADS*, участвует в определении идентичности лепестков, тычинок и плодолистиков, однако данные о функциях гена ограничены. Целью работы стала функциональная характеристика *SlMADS5*. Структурно-филогенетический анализ *SlMADS5* подтвердил его принадлежность к SEP3-кладе. Экспрессионный анализ *in silico* выявил отсутствие транскриптов гена в корнях, листьях и апикальной меристеме побега и их присутствие в цветках, плодах и семенах на разных стадиях развития. Двугибридный анализ показал способность белка SlMADS5 активировать транскрипцию гена-мишени и взаимодействовать с белком TAGL1. Трансгенные растения *Nicotiana tabacum* L. с конститутивной сверхэкспрессией кДНК *SlMADS5* цвели в 2.2 раза позже и формировали утолщенные листья, имели в 2.5–3.0 раза более толстый стебель, в 1.5–2.7 раза укороченные междоузлия и в 1.9 раза меньше

А.В. Нежданова, М.А. Слугина, Е.А. Дьяченко
А.М. Камионская, Е.З. Кочиева, А.В. Щенникова

Структурно-функциональный анализ MADS-box
гена *SlMADS5* томата *Solanum lycopersicum* L.

2021
25•5

цветков и коробочек, чем нетрансгенные растения. Строение цветков от контроля не отличалось, однако лепестки венчика сменили окраску со светло-розовой на маджент. Анализ экспрессии *SlMADS5* и генов табака *NtLFY*, *NtAP1*, *NtWUS*, *NtAG*, *NtPLE*, *NtSEP1*, *NtSEP2* и *NtSEP3* в тканях листьев и апексов трансгенных и контрольных растений показал, что мРНК *SlMADS5* присутствует только в тканях трансгенных линий. Остальные анализируемые гены высоко экспрессировались в репродуктивной меристеме контрольных растений. Транскрипты данных генов отсутствовали или присутствовали в следовых количествах в листьях и вегетативном апексе контроля, а также в листьях и апексах трансгенных линий. Полученные результаты свидетельствуют о возможном участии гена *SlMADS5* в регуляции развития меристемы цветка и пути биосинтеза антоцианов в лепестках.

Ключевые слова: *Solanum lycopersicum*; *Nicotiana tabacum*; гетерологичная экспрессия гена; MADS-транскрипционные факторы; SEPALLATA; SlMADS5.

Introduction

Throughout the plant's life cycle, its root and shoot apical meristems maintain a pool of pluripotent stem cells, which give rise to new organs: roots and leaves respectively, during vegetative development and flowers during reproduction stage. At the reproductive stage, the shoot apical meristem of the angiosperms turns into the inflorescence meristem, which forms determined flower meristems (Hugouvieux et al., 2018). In all aspects of flowering, the MADS-domain family of transcription factors (TFs) plays a key role according to the well-known ABCDE flower development model (Smaczniak et al., 2012).

The ABCDE model is based on genetic and molecular studies, primarily of model species *Arabidopsis thaliana* (L.) Heynh., *Antirrhinum majus* L., and *Petunia×hybrida* hort. ex E. Vilm. (Coen, Meyerowitz, 1991; Angenent et al., 1995; Pelaz et al., 2000; Theissen, 2001; Ditta et al., 2004). According to the model, the identity of flower organs is determined by five classes of genetic activities: A and E – sepals; A, B and E – petals; B, C and E – stamens; C and E – carpels; C, E and D – ovules. At the molecular level, the ABCDE-model is explained by the so-called "quartet" model, according to which MADS-TFs of ABCDE classes in various combinations form tetramers: for example, C/C/E/E – to determine carpel identity, or A/B1/B2/E – to specify petal identity (Honma, Goto, 2001; Theissen, Saedler, 2001). These tetramers activate or suppress transcription of target genes (Melzer et al., 2009; Smaczniak et al., 2012). The current data suggest a high structural and functional conservatism of A, B, C, D, and E genes in flowering plants (Smaczniak et al., 2012).

The genes of the E-class, *A. thaliana SEPALLATA* (*SEP1*, *SEP2*, *SEP3*, and *SEP4*), which are involved in determining the identity of all floral organs, deserve special attention (Pelaz et al., 2000; Smaczniak et al., 2012). The knockout of only one of the *SEP* genes does not have a significant effect on the *A. thaliana* flower, while the *sep1 sep2 sep3* triple mutation transforms all the flower organs into sepals; a new flower with the same development pattern is formed instead of the pistil (Pelaz et al., 2000). The quadruple *sep1 sep2 sep3 sep4* mutation leads to the replacement of all flower organs with leaf-like organs (Ditta et al., 2004).

SEP proteins are the central hub in the formation of MADS-TF quartets (Immink et al., 2009). Among SEPs, SEP3 is the most functionally pleiotropic and interacts with almost all MADS-TFs responsible for the identity of flower organs (Alhindi et al., 2017). *SEP3* gene simultaneous ectopic expression with the A-, B-, or C-class genes transforms leaves into flower organs (Honma, Goto, 2001; Pelaz et al., 2001b).

During plant evolution, *SEP* genes are believed to have arisen later than other flower-related MADS-box genes, but at the same time they became key players in the origin of flowering plants, as well as in the domestication and breeding of crops (Theissen, 2001; Schilling et al., 2018). Therefore, their study in cultivated plants can expand the understanding of the role of these genes in determining economically valuable traits.

The tomato *Solanum lycopersicum* L. is one of the most important vegetables and, at the same time, a model for studying the fleshy fruit development and ripening. The tomato genome has been sequenced and annotated (https://www.solgenomics.net/), and contains several *SEP* genes: *TAGL2* (Solyc05g015750.2.1), *SlMADS6/TM29/LeSEP1* (Solyc02g089200.2.1), *RIPENING INHIBITOR* (*MADS-RIN*) (Solyc05g012020.2.1), *SlMADS98/SlCMB1* (Solyc04g005320.2.1), *SlMADS1/ENHANCER-OF-JOINTLESS-2* (Solyc03g114840.2.1), *SlMBP21/JOINTLESS-2* (*J2*) (Solyc12g038510.1.1) and *SlMADS5/TM5/TDR5/LeSEP3* (Solyc05g015750.3.1) (Wang Y. et al., 2019).

In addition to determining the flower organ identity, SEP proteins, together with MADS-TFs of the FRUITFULL (FUL) and AGAMOUS (AG) subfamilies, are actively involved in the regulation of fruit ripening. This is clearly demonstrated in tomato, the fruit ripening of which is controlled by FUL1/FUL2, TOMATO AGAMOUS 1 (TAG1)/TOMATO AGAMOUS-LIKE 1 (TAGL1) and MADS-RIN (Karlova et al., 2014; Shima et al., 2014; Wang R. et al., 2019). At the same time, FUL2 and TAGL1 have been shown to play an additional role in pistil initiation and early fruit development (Vrebalov et al., 2009; Wang R. et al., 2019), which is likely to be performed in combination with the tomato SEP3 homolog, SlMADS5 (Leseberg et al., 2008).

SEP1-like gene *TAGL2* was shown to be expressed at stages I (anthesis) and II of the tomato fruit development (Busi et al., 2003). Suppression of *SEP1*-like *TM29* causes the development of parthenocarpic fruits and the flower reversion (Ampomah-Dwamena et al., 2002). Tomato *SEP4*-like *SlCMB1* regulates ethylene biosynthesis and the accumulation of carotenoids during fruit ripening; suppression of *SlCMB1* leads to a change in the inflorescence architecture and an increase in the sepal size (Zhang et al., 2018a, b). *SEP4*-like *SlMADS1* acts as a negative regulator of fruit ripening (Dong et al., 2013). *SEP4*-like *SlMBP21* specifies the sepal size me-

diated by ethylene and auxin signaling, as well as the abscission zone formation (Li et al., 2017; Roldan et al., 2017). *SEP4*-like *MADS-RIN* is the main regulator of fruit ripening: gene knockout leads to the formation of an unripe fruit, including the absence of carotenoid accumulation (Vrebalov et al., 2002; Leseberg et al., 2008).

The only representative of the tomato clade SEP3, TF SlMADS5, is involved in determining the identity of the organs of the three inner flower whorls (Pnueli et al., 1994), interacting with MADS-TFs of the SEP and AG subfamilies (Leseberg et al., 2008). Despite the *SEP3* significance, this gene variability has been characterized in cultivated and wild tomato species, and the *SlMADS5* expression was observed in some organs and tissues (Pnueli et al., 1994; Slugina et al., 2020).

The aim of the present study was to characterize the function of *S. lycopersicum SlMADS5*. *SlMADS5* structural, phylogenetic and expression analysis confirmed its belonging to the SEP3-clade. Analysis in the yeast two-hybrid GAL4-system showed the SlMADS5 TF activator properties and its interaction with C-class MADS-TF. Transgenic *Nicotiana tabacum* L. plants with *SlMADS5* constitutive overexpression exhibited a pronounced phenotype of reproductive development suppression.

Materials and methods

Tomato *S. lycopersicum* cv. Silvestre recordo and tobacco *N. tabacum* cv. Samsun plants were used in the study. Tomato accessions were grown under controlled greenhouse conditions (day/night: +21/23 °C, 16 h/8 h; 300–400 µmol/m^{-2}/s^{-1}) until flowering. Roots, leaves, flowers and ripe fruits were collected separately. Tissues were grounded in liquid nitrogen and stored at –70 °C. Tobacco accessions were grown *in vitro* on a sterile MS medium in a climatic chamber (day/night: +21/23 °C, 16 h/8 h; 300 µmol/m^{-2}/s^{-1}) until the formation of 4–6 leaves.

Total RNA was isolated from tomato (roots, leaves, flowers, and ripe fruits) and tobacco (leaves, vegetative apex, and reproductive apex) tissues using the RNeasy Plant Mini Kit (QIAGEN, USA), and used for cDNA synthesis (the Reverse Transcription System, Promega, USA). Genomic DNA was isolated from leaf tissues by the standard potassium acetate method (Dyachenko et al., 2018) and used for PCR tests for the presence of a transgene in the plant genome.

Primers for gene amplification, sequencing, and expression analysis were generated based on the MADS-box transcripts of *S. lycopersicum* cv. Heinz and tobacco *N. tabacum* genes available in the NCBI (http://www.ncbi.nlm.nih.gov/) (*NtAPETALA1* (*NtAP1*; JQ686939.1, AF068724.2, XM_016635359.1, AF009127.1, U63162.1); *NtLEAFY* (*NtLFY*; JQ686928.1, XM_016593842.1); *NtWUSCHEL* (*NtWUS*; XM_016637596.1, MG843891.1, XM_016619508.1, JQ686923.1); *NtAG* (NM_001325900, XM_016638054.1, XM_016580096.1, XM_016580095.1, XM_016580097.1); *NtPLENA* (*NtPLE*; XM_016631079.1, XM_016631071.1, XM_016615571.1, XM_016615578.1, U63163.1): *NtSEP1* (XM_016653813.1, XM_016645589.1, XM_016620650.1, XM_016596552.1, XM_016611481.1, XM_016645132.1, NM_001324748.1, XM_016620651.1, XM_016647424.1,

XM_016644825.1); *NtSEP3* (NM_001325160.1, XM_016582910.1); *NtSEP2* (XM_016645132.1, NM_001324748.1, XM_016645589.1)) so that forward and reverse primers are separated by at least one intron and match all possible transcripts for each of the analyzed genes (Table 1). The primer sequences were additionally verified using Primer 3 and BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). Primers for CDS in-frame cloning into plasmid vectors (GAL4 system) contained *EcoR*I (forward, F) and *Sal*I (reverse, R) restriction sites at the 5′ end.

Full-length *SlMADS5*, *TAG1*, and *FUL2* cDNAs were amplified using the cDNA, isolated from *S. lycopersicum* cv. Silvestre recordo flowers; PCR conditions: initial denaturation at 95 °C for 5 min; 30 cycles of denaturation (94 °C for 30 s),

Table 1. The list of primers used in the study

Gene	Primer name	Primer sequence (5′→3′)
		CDS amplification
SlMADS5	SlMADS5 F/R	TAATCAGAATTCATGGGAAGGGGTAGGGTTGA TTGCATGTCGACTCAAGGCAACCAGCCAGCCA
TAGL1	TAGL1 F/R	TAATCAGAATTCATGGTTTTTCCTA TTGCATTGTCGACTCAGACAAGCTGGAGAGGAG
FUL2	FUL2 F/R	TAATCAGAATTCATGGGTAGAGGAAGAGTACA TTGGATGTCGACTTAACCGTTGAGATGGCGAA
		qRT-PCR
NtAP1	NtAP1 F/R	AGGACCTGCAAAACTTGGAA TGATTTTGCTGATGCCATTC
NtLFY	NtLFY F/R	TAATGCCCTTGACGCTCTCT TCGACACCACCTTCTTCCTC
NtWUS	NtWUS F/R	CTTCTCATGGTGTACTGGCC CAGTTCCTCATAATCGTCTACTAG
NtSEP1	NtSEP1 F/R	AATAATGGCGGAACAGATGG TGGATCAGGTTCACATTCCA
NtSEP3	NtSEP3 F/R	TCACTTGAGAGGCAGCTTGA CATCGCCCTGAGTTTGAGTT
NtSEP2	NtSEP2-2 F/R	GCAACATGCTCAATCTCAGG TTGGGCATTTGTTACTGCTG
NtAG	NtAG F/R	ATGAGCTGCTGTTTGCTGAA TGAACTCCCTGGCATCAAGT
NtPLE	NtPLE F/R	GCCATTGGTAGAGTCCGTTC AGCTGGAGAGCAGTTTGGTC
SlMADS5	SlMADS5 F/R	GCCAAATGCACAAGATGTGGG CCAGCCATGTAGTTATTCACAC
actin-7	Actin-7 F/R	CTACGAGCAGGAGCTTGACA TAATCTTCATGCTGCTGGGA
		PCR for the presence of a transgene in the plant genome*
NOS-T	NOS-R	CGAATTCCCGGGATCTAGTAACATAGATGAC

* SlMADS5-F primer was used as a direct primer for PCR analysis of plants for the presence of a transgene in the genome.

А.В. Нежданова, М.А. Слугина, Е.А. Дьяченко
А.М. Камионская, Е.З. Кочиева, А.В. Щенникова

Структурно-функциональный анализ MADS-box
гена *SlMADS5* томата *Solanum lycopersicum* L.

2021
25•5

annealing (55 °C – 30 s) and synthesis (72 °C – 1 min); final synthesis (72 °C – 7 min). The PCR fragments of the expected length were purified using the MinElute Gel Extraction Kit (QIAGEN, USA), cloned into the pGEM®-T Easy plasmid vector (Promega, Madison, WI, USA) at *Eco*RI and *Sal*I sites and sequenced (Core Facility "Bioengineering"). Further, the *SlMADS5*, *FUL2*, and *TAGL1* CDSs were cloned into hybrid vectors pAD-GAL4 and pBD-GAL4cam (Aglient Technologies, USA): each gene was ligated in frame with the activator domain (pAD) and DNA-binding domain (pBD) of the yeast TF GAL4. Recombinant pJ69-4a strains carrying each pAD-gene and pBD-gene construct separately, as well as in pairs pAD-gene + pBD-gene, were obtained. For plant transformation, *SlMADS5* cDNA was cloned in a sense orientation into a binary vector based on pBin19, under the control of the enhanced cauliflower mosaic virus promoter 35S and nopaline synthase (NOS) terminator. With this construct, a recombinant agrobacterial strain AGLØ was obtained.

For sequence structural analysis, the NCBI-CDD (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi), MEGA 7.0 (Kumar et al., 2016) and Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2/) were used. Sequence phylogeny was assessed in the MEGA7, using Maximum Likelihood method based on the JTT model.

Gene expression analysis was performed *in silico* (using TomExpress database; http://tomexpress.toulouse.inra.fr/select-data), as well as by quantitative (q) real-time (RT) PCR in two biological and three technical replicates. The kit "Reaction mixture for carrying out qRT-PCR in the presence of SYBR Green I and ROX" (JSC Syntol, RF) and the CFX96 Real-Time PCR Detection System (Bio-Rad Laboratories, USA) were applied. The qRT-PCR conditions were as follows: 95 °C – 5 min; 40 cycles (95 °C – 15 s, 60 °C – 50 s). The reference gene *actin-7* (XM_016658880.1) (Schmidt, Delaney, 2010) was used for normalizing the expression of tobacco genes. Statistical processing of the results was carried out using the GraphPad Prism v. 7.02 (https://www.graphpad.com).

The analysis of SlMADS5 interactions with TAGL1 and FUL2 proteins was carried out *in vivo* in a two-hybrid GAL4-yeast system using the *Saccharomyces cerevisiae* Pj69-4a strain, according to the HybriZAP-2.1-Hybrid cDNA Two-Hybrid Synthesis Kit protocol (Stratagene).

Leaf explants of tobacco (*N. tabacum* cv. Samsun) were transformed using *Agrobacterium tumefaciens* strain AGLØ. To select transgenic regenerants, an MS medium containing kanamycin (Km, 100 mg/L) for selection and carbenicillin (500 mg/L), which suppresses agrobacteria growth, was used. The rooted regenerants were adapted to the soil in greenhouse conditions and then tested for the presence of a transgene in the genome by PCR with primers specific to the sequences of the 5′ end of the transgene and the NOS-terminator (see Table 1).

Results

To confirm the conservatism of the SlMADS5 function in tomato (cv. Silvestre recordo), an analysis of its interactions with MADS-TFs TAGL1 and FUL2, the interaction with which was and was not, respectively, shown earlier (Leseberg et al., 2008), was carried out.

Structural analysis of the SlMADS5 protein was carried out in comparison with the known tomato, tobacco, and *A. thaliana* SEP homologs. The presence of the main domains characteristic of MIKC^c type MADS-TFs was confirmed, namely the highly conserved MEF2-like MADS-domain (1–76 aa), an I-region (77–92 aa), a conserved keratin (K)-like domain (93–173 aa), and a variable C-region (174–241 aa) (Fig. 1, *a*). The performed phylogenetic analysis testified the belonging of SlMADS5 to the SEP3 clade (see Fig. 1, *b*).
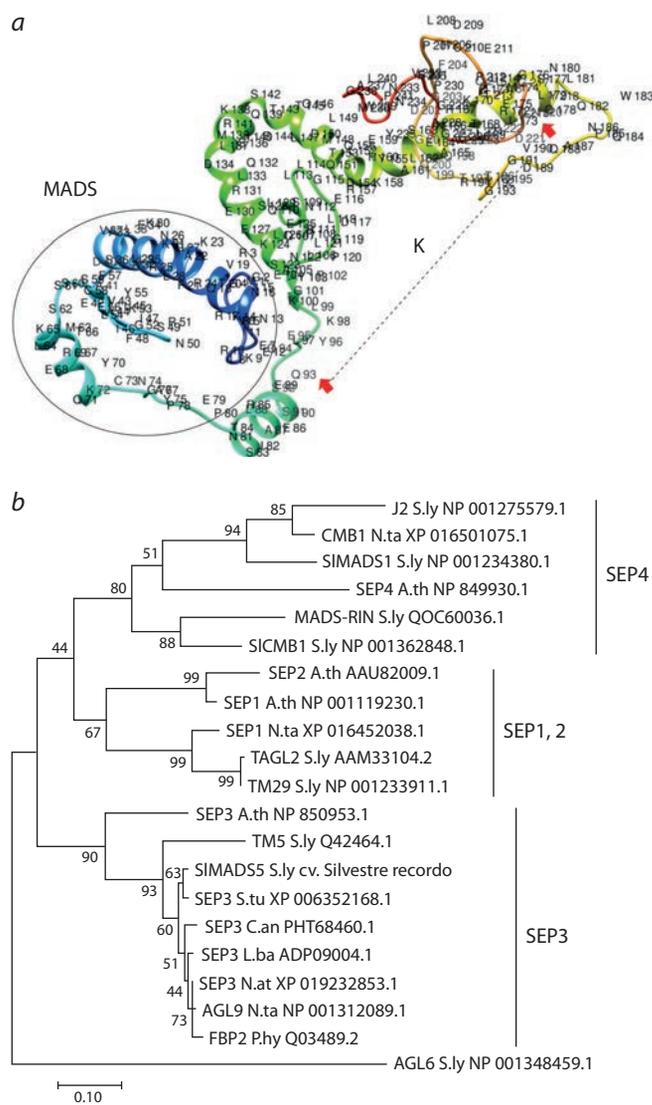
Fig. 1. Structure and phylogenetic analysis of the SlMADS5.

a – three-dimensional structure of TF SlMADS5 according to Phyre2. The MADS-domain is indicated with a circle; the beginning and end of the K-domain are indicated with red arrows; *b* – dendrogram based on the alignment of 19 MADS-TF sequences from the SEP clade of tomato, other Solanaceae species, and the model species *Arabidopsis thaliana*.

The analysis was carried out in MEGA 7.0, using the Maximum Likelihood method based on the JTT model. The tree is rooted with *S. lycopersicum* AGL6 homolog. The significant bootstrap values for 1000 replicates are shown at the base of the branches. The NCBI accession numbers are shown opposite the protein names. S.ly – *S. lycopersicum*, N.ta – *N. tabacum*, A.th – *A. thaliana*, S.tu – *S. tuberosum* L., C.an – *Capsicum annuum* L., L.ba – *Lycium barbarum* L., N.at – *N. attenuate* Torr. ex S. Watson, P.hy – *Petunia × hybrida*.
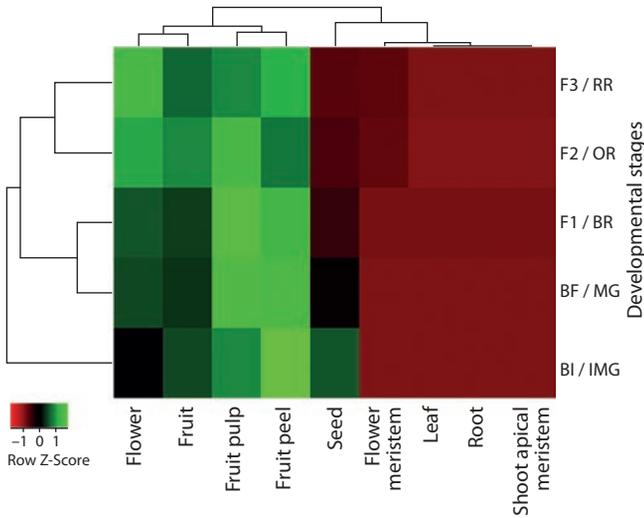
Fig. 2. Heatmap of *SlMADS5* gene expression in roots, shoot apical meristems, leaves, flower meristems, and flowers at the stages of bud initiation (BI), bud formation (BF), flower opening (F_1–F_3), as well as in whole fruits, fruit peels, fruit pulps, and seeds at the stages IMG, MG, BR, OR, and RR.

Expression of *SlMADS5* in roots and reproductive tissues is shown for *S. lycopersicum* cv. MicroTom; in leaves and shoot apical meristems, for cv. M82.

Table 2. Results of the analysis of SlMADS5 protein-protein interactions

pAD_GAL4	pBD_GAL4cam	–LH +10 mM 3AT	–LTH +10 mM 3AT	–LTA	X-gal test
		Autoactivation test			
	SlMADS5	+		+	+
	CDM44 (positive control)*	+		+	+
	CDM37 (negative control)*	–		–	–
		Protein-protein interaction test			
SlMADS5	TAGL1		+	+	+
SlMADS5	FUL2		–	–	–
CDM44 (+)*	CDM37 (+)*		+	+	+
CDM37 (–)*	CDM111 (–)*		–	–	–

* According to (Shchennikova et al., 2004). The experiment was carried out in parallel at room temperature and 30 °C (the same results were obtained for both temperatures). L – L-Leucine; H – L-Histidine; T – L-Tryptophan; A – L-Adenine hemisulfate salt; 3AT – 3-amino-1,2,4-triazole; –LH, –LTH and –LTA – nutritional medium without leucine/histidine, leucine/histidine/tryptophan, and leucine/tryptophan/adenine, respectively; X-gal – 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside. X-gal test: yeast colonies, where the analyzed proteins interact and, as a result, activate the expression of the β-galactosidase (lacZ) gene, acquire a blue color due to the cleavage of the X-gal substrate added to the medium by the β-galactosidase enzyme.

To characterize TF SlMADS5 functionally, we analyzed the expression of the *SlMADS5* gene in various tomato organs and the ability of SlMADS5 protein to activate gene transcription and interact with MADS proteins of the C and A classes. Also, transgenic *N. tabacum* model plants with constitutive overexpression of *SlMADS5* cDNA were obtained.

In silico analysis of the *SlMADS5* expression pattern was carried out in roots, leaves, vegetative shoot meristem, flower meristem, flower (from bud to fully open and anthesis stage), fruits (4–8 days after anthesis), fruit skin and pulp (stages: Immature Green (IMG); Mature Green (MG); Breaker (BR), color change; Orange (OR); Red Ripe (RR)), and in seeds (IMG, MG, BR, RR) (Fig. 2). *SlMADS5* transcripts were not found in roots, leaves, and the vegetative apical meristem. At the same time, *SlMADS5* expression was shown in flowers (maximum – at the anthesis stage), fruits, fruit peel (maximum at MG and BR stages), fruit pulp (maximum at IMG, MG, and BR stages), and seeds (maximum at IMG stage) (see Fig. 2).

In vivo analysis in the yeast two-hybrid GAL4 system showed that TF SlMADS5 has the property of activating the transcription of target genes, interacts with the C-class MADS protein TAGL1, but does not interact with the A-class MADS protein FUL2 (Table 2).

The characterization of transgenic tobacco plants with *SlMADS5* constitutive overexpression was performed. Independent regenerants T_0 35S::*SlMADS5* (18 plants) were adapted to the greenhouse, tested by PCR for the presence of a transgene expression cassette in the genome, and compared with the control (non-transgenic tobacco plants) during development. In comparison with the control, 35S::*SlMADS5* plants (Fig. 3) bloomed much later (on average, 138 days *vs.* 62 in the control). Also, 35S::*SlMADS5* phenotype was characterized by a 2.5–3.0 times thicker stem, 2.0 times shortened internodes, thickened and darker leaves, and 2.5 times fewer flowers and capsules. The 35S::*SlMADS5* flower structure did not differ from the control.

Seeds of two transgenic T_0 lines (S5-16 and S5-17) with a pronounced phenotype were planted in a greenhouse. T_1 plants, which gave a positive PCR signal for the presence of a transgene in the genome, bloomed 1.3–1.5 times later than the control, had a 35S::*SlMADS5* phenotype, and formed flowers with magenta-colored corolla petals, in contrast to light pink petals in the control.

Seeds of lines T_1 S5-16-6, S5-16-7, S5-17-1 and S5-17-4 were planted on MS medium (Km 50 mg/l); the 3:1 ratio of the number of Km-resistant to Km-sensitive seedlings indicated a heterozygous state of the transgene and one copy of it in the genome of transgenic lines. In seedlings, internodes were near absent, and only T_2 plants of the S5-16-7 line (14 accessions) formed a noticeable stem and were adapted to the greenhouse (the rest of the plants died after transfer to the soil). Plants T_2 S5-16-7 demonstrated the 35S::*SlMADS5* phenotype: they bloomed 2.4 times later than the control; formed thickened stems and leaves, shortened internodes, and 2.3 times less seed capsules.

In T_1 lines S5-16-7 and S5-17-1, in comparison with the control, we analyzed the *SlMADS5* expression, as well as the expression of tobacco genes associated with reproductive development: *NtLFY*, *NtAP1* (plant transition to flowering),

А.В. Нежданова, М.А. Слугина, Е.А. Дьяченко
А.М. Камионская, Е.З. Кочиева, А.В. Щенникова

Структурно-функциональный анализ MADS-box
гена *SlMADS5* томата *Solanum lycopersicum* L.

2021
25•5

Fig. 3. Transgenic tobacco plants T$_0$ (*SlMADS5*) (*b–d, f*) in comparison with the control non-transgenic *N. tabacum* plant (WT) (*a, e*) at the stages of bud formation (*c, d*), flowering (*a, e, f*), and seed formation (*b*).

(*c*) and (*d*) – the top of the same plant 35S::*SlMADS5*. The photos were taken one and a half weeks apart. Scale bar 1 cm.

NtWUS (central regulator of stem cells in the meristem), *NtAG*, *NtPLE*, *NtSEP1*, *NtSEP2*, *NtSEP3* (key genes for the identity of the flower meristem and flower organs). For the analysis, we used tissues of leaves and apical meristems (vegetative and reproductive in the control, and shoot meristem in lines S5-16-7 and S5-17-1) of transgenic and control plants.

Expression of the *SlMADS5* transgene was present only in the tissues of S5-16-7 and S5-17-1 plants. The expression pattern of the remaining analyzed genes was similar: their mRNA was absent or was minimal in the leaves of the control and transgenic lines, as well as in the S5-16-7 and S5-17-1 apexes of undefined status. At the same time, these genes were highly transcribed in the reproductive meristems of control plants (Fig. 4).

Discussion

In this study, a functional analysis of the *SlMADS5* gene, the *SEP3* homolog in tomato, was carried out. Structural analysis (see Fig. 1) confirmed that SlMADS5 belongs to the SEP3 clade, which may indicate the conservatism of its role in the

reproductive development of tomato, namely, its participation in determining the identity of petals, stamens, carpels, and ovules.

It is known that *SlMADS5* is not expressed in tomato leaves and roots and is expressed in flowers and fruits (Slugina et al., 2020). Also, *SlMADS5* mRNA is present in the meristem domains that correspond to the future three inner whorls of the tomato flower, as well as during organogenesis and in the corresponding mature organs (Pnueli et al., 1991, 1994). A detailed *in silico* analysis of the *SlMADS5* expression pattern carried out in this study revealed that *SlMADS5* mRNA is absent not only in roots and leaves, but also in the shoot apical meristems and flower meristems at early stages of development (see Fig. 2). Gene transcription is activated late in the development of the flower meristem, and reaches a peak in an open flower and in the peel of an immature fruit (see Fig. 2). This corresponds not only to the well-known role of SEP3 homologs in determining the differentiation of flower meristem cells corresponding to the three inner whorls of organs (Pnueli et al., 1991, 1994), but also suggests the active
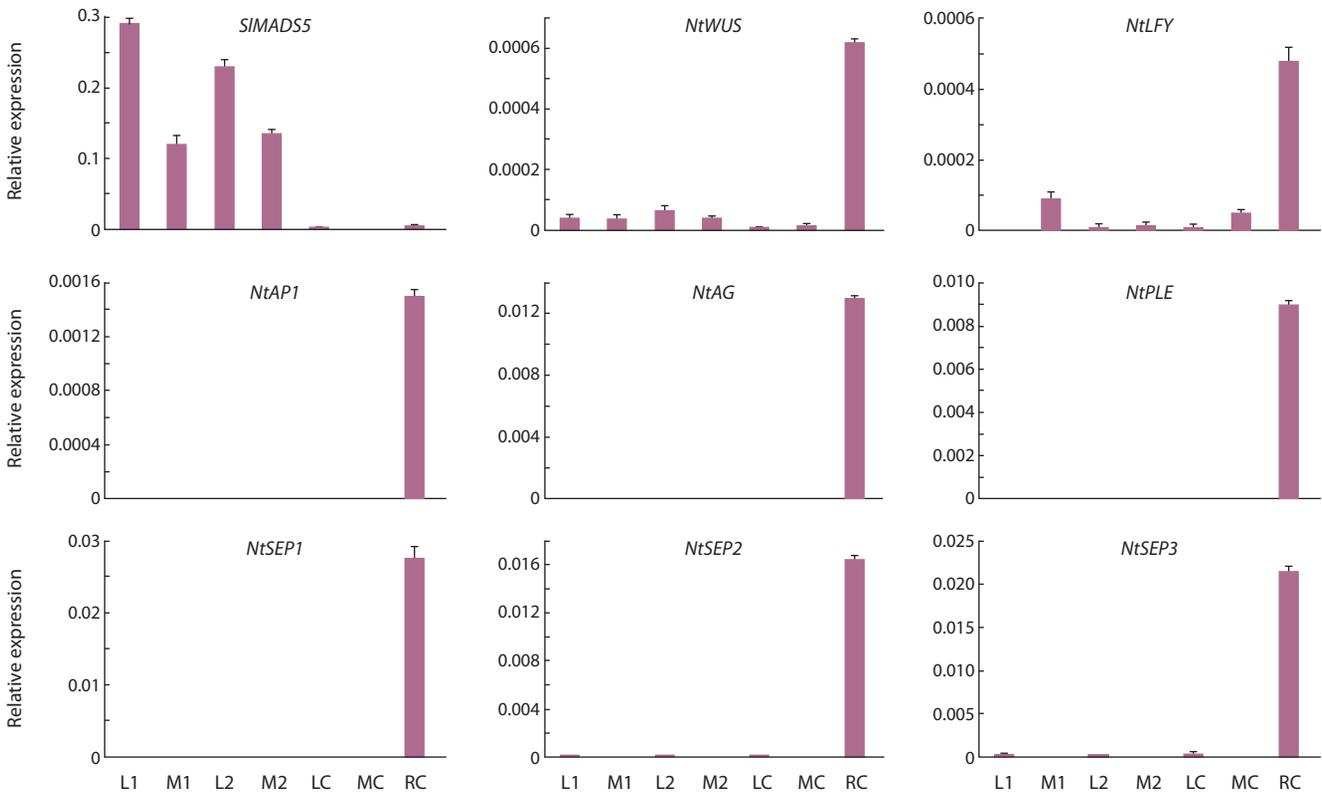
Fig. 4. Expression of *SlMADS5* and *N. tabacum* genes *NtWUS*, *NtLFY*, *NtAP1*, *NtAG*, *NtPLE*, *NtSEP1*, *NtSEP2* and *NtSEP3* in control (C) and transgenic lines S5-16-7 (1) and S5-17-1 (2).

L – leaf; M – shoot apical meristem, R – reproductive meristem.

participation of SlMADS5 in the aspects of development and ripening of tomato fruits and seeds.

To characterize the *SlMADS5* function, transgenic tobacco plants with constitutive overexpression of *SlMADS5* cDNA were obtained. The phenotype of transgene overexpression does not determine its function; however, it may indicate a similarity with the already characterized homologs. Earlier, the effect of heterologous overexpression of *SEP3* homologs of different plant species was studied mainly using transgenic *A. thaliana* plants, but there are works with the use of *Nicotiana* spp. plants. Tobacco, like tomato, belongs to the Solanaceae family and has the same flower structure; therefore, in this study, a heterologous expression system in tobacco was selected.

Various effects of overexpression of *SEP3* homologs have been described. Thus, *SEP3* constitutive expression in *A. thaliana* significantly accelerates flowering (Pelaz et al., 2001a). In these plants, the *APETALA3* (B-class) and *AG* (C-class) genes are transcribed ectopically (Castillejo et al., 2005). Overexpression of the *P. × hybrida SEP3*-like gene *FBP2* leads to early flowering of the *A. thaliana* plants (Ferrario et al., 2003). Early flowering is caused by overexpression of tobacco *SEP3*-like gene *NsMADS3* in *N. sylvestris* Speg. & Comes (Jang et al., 1999) and chrysanthemum *SEP3*-like gene *CDM44* in *N. tabacum* (Goloveshkina et al., 2012).

At the same time, no influence of overexpression of *SEP3*-homologous genes on the flowering time was also observed. Thus, homologous overexpression of *FBP2* in *P. × hybrida*

has no effect on plant vegetation period (Ferrario et al., 2006). Heterologous overexpression of *Platanus acerifolia SEP3*-like genes in *A. thaliana* causes early flowering only in the case of the *PlacSEP3.2* gene, while overexpression of the second gene, *PlacSEP3.1*, causes early flowering only in transgenic tobacco plants (Zhang et al., 2017).

In the case of *SlMADS5* constitutive overexpression, a significant delay in flowering was observed, most likely associated with the incorrect development of the shoot apical meristem (see Fig. 3). Different effects of heterologous ectopic expression of *SEP3* homologs in transgenic plants may be associated with structural differences in encoded protein sequences responsible for binding to promoters of target genes or to partner proteins.

Normally, traces of the *A. thaliana SEP3* transcripts are found in the inflorescence meristem, and gene expression is noticeably activated only in the flower meristem parts, from which petals, stamens, and carpels are subsequently formed (Ferrario et al., 2003; Urbanus et al., 2009). Therefore, the presence of the TF SlMADS5 in tissues, where there should be no tobacco SEP3 homologs, can lead to nonspecific protein-protein and DNA-binding interactions of SlMADS5, which can disrupt the pattern of meristem development.

To clarify the status of transgenic meristems S5-16-7 and S5-17-1, visually ready for flowering, we analyzed the expression of genes whose activity is associated with the identity of the reproductive inflorescence and flower meristems (*NtLFY* and *NtAP1*) (Weigel et al., 1992). Considering the results

А.В. Нежданова, М.А. Слугина, Е.А. Дьяченко
А.М. Камионская, Е.З. Кочиева, А.В. Щенникова

Структурно-функциональный анализ MADS-box
гена *SlMADS5* томата *Solanum lycopersicum* L.

2021
25·5

obtained (see Fig. 4), only the inflorescence meristem of the control plant has reproductive status. The presence of a low level of *LFY* expression in the vegetative apex of the control and in the S5-16-7 apex (see Fig. 4) suggests the initial stages of the meristem transition to the reproductive state, since it has been shown that in *A. thaliana LFY* begins to be expressed in the flower meristem primordia at the periphery of the inflorescence meristem (Weigel et al., 1992).

It is known that SEP3 is the central hub of the MADS-complexes in *A. thaliana* (Immink et al., 2009). TF SlMADS5 also shows an exceptional ability to assemble tetrameric complexes of MADS TFs (Leseberg et al., 2008). The interaction of SlMADS5 with FUL2 and TAGL1 shown in this work (see Table 2), as well as the role of FUL2 and TAGL1 in pistil initiation and early fruit development (Vrebalov et al., 2009; Wang R. et al., 2019), indicate the possible involvement of SlMADS5 in determining the identity of the tomato pistil in complex with FUL2 and TAGL1.

One of the complexes, SEP3/SEP3/AG/AG, is required for flower determination and completion of its development (Hugouvieux et al., 2018). This is due to a decrease in the number of stem cells because of the *WUS* gene suppression with the key participation of TF AG (Lenhard et al., 2001). Accordingly, in transgenic petunia plants with simultaneous overexpression of *SEP3*-like *FBP2* and D-class gene *FBP11*, where developmental arrest is observed at the cotyledon stage, transcription of *AG*-like *FBP6* is activated and mRNA of *WUS*-like *TERMINATOR* is absent (Ferrario et al., 2006). This suggests the joint participation of *SEP3*, *AG*, and D-class genes in the suppression of stem cells in the meristem.

Taking into account the activation of *AG* expression in *A. thaliana* with *SEP3* overexpression (Castillejo et al., 2005), as well as the participation of SEP3 and AG in the suppression of *WUS* transcription (Lenhard et al., 2001; Ferrario et al., 2006) and the interaction of TF SlMADS5 with the AG homolog TAGL1 (see Table 2), it can be assumed that the ectopically synthesized TF SlMADS5 is able to activate transcription of the tobacco *AG*-like genes *NtAG* and *NtPLE* in transgenic shoot meristem. Subsequent formation of complexes SlMADS5/SlMADS5/NtAG/NtAG or SlMADS5/SlMADS5/NtPLE/NtPLE can lead to inhibition of meristem development due to the tobacco *WUS*-like gene *NtWUS* suppression, since WUS plays a key role in determining the stem cell identity, the population of which is not supported in plants with loss of WUS function (Ferrario et al., 2006; Jha et al., 2020).

To test this possibility, we analyzed the expression of *SlMADS5*, *NtWUS*, *AG*-like genes *NtAG* and *NtPLE*, as well as *SEP*-like genes *NtSEP1*, *NtSEP2*, and *NtSEP3*. However, the presence of *SlMADS5* ectopic expression did not lead to the activation of *AG*-like genes, and the expression of *NtWUS* was significantly higher in the tissues of transgenic lines in comparison with the control (excluding the control inflorescence meristem) (see Fig. 4). The latter can be a probable reason for the formation of significantly thickened, in comparison with the control, stem and leaves of transgenic plants of all 11 lines with the 35S::*SlMADS5* phenotype (see Fig. 3) as a result of the increased number of stem cells and the meristem overgrowth.

It should also be noted that in transgenic plants, the anthocyanin color of the flower corolla changed from pale pink (control) to magenta (35S::*SlMADS5*) (see Fig. 3). Previously, it was shown that the expression of the *SEP*-like gene *MrMADS01* in *Myrica rubra* berries significantly increases at the last stage of ripening, which allowed the authors to suggest the involvement of this gene in the biosynthesis of anthocyanins (Zhao et al., 2019). Silencing the *SEP*-like gene *PaMADS7* in sweet cherry (*Prunus avium*) leads to a change in the content of anthocyanins in fruits (Qi et al., 2020). It can be assumed that SlMADS5 is also involved in the regulation of anthocyanin biosynthesis in transgenic tobacco petals.

Silencing of *SlMADS5* gene leads to a change in the number of flower whorls and the number of organs in whorls, as well as the formation of green petals with signs of sepals, and sterile anthers and carpels with signs of sepals and petals, respectively (Pnueli et al., 1994), which may indicate the participation of the gene in determining the identity of tomato flower organs. Nevertheless, no complete homeotic transformation of certain flower organs was observed when *SlMADS5* was suppressed (Pnueli et al., 1994).

Conclusion

The data on the effect of *SlMADS5* overexpression on the development of transgenic tobacco plants obtained in this study also do not confirm the involvement of the gene in determining the floral organ identity. Also, the data obtained may indicate that the ectopic expression of this single gene in a heterologous system (*N. tabacum*) is insufficient to activate transcription of the MADS-box tobacco genes associated with flowering, but it is sufficient for a long delay in the reproductive development of the plant.

References

Alhindi T., Zhang Z., Ruelens P., Coenen H., Degroote H., Iraci N., Geuten K. Protein interaction evolution from promiscuity to specificity with reduced flexibility in an increasingly complex network. *Sci. Rep.* 2017;7:44948. DOI 10.1038/srep44948.

Ampomah-Dwamena C., Morris B.A., Sutherland P., Veit B., Yao J.L. Down-regulation of TM29, a tomato SEPALLATA homolog, causes parthenocarpic fruit development and floral reversion. *Plant Physiol.* 2002;130(2):605-617. DOI 10.1104/pp.005223.

Angenent G.C., Franken J., Busscher M., van Dijken A., van Went J.L., Dons H.J., van Tunen A.J. A novel class of MADS box genes is involved in ovule development in petunia. *Plant Cell.* 1995;7(10): 1569-1582. DOI 10.1105/tpc.7.10.1569.

Busi M.V., Bustamante C., D'Angelo C., Hidalgo-Cuevas M., Boggio S.B., Valle E.M., Zabaleta E. MADS-box genes expressed during tomato seed and fruit development. *Plant Mol. Biol.* 2003;52(4): 801-815. DOI 10.1023/a:1025001402838.

Castillejo C., Romera-Branchat M., Pelaz S. A new role of the Arabidopsis *SEPALLATA3* gene revealed by its constitutive expression. *Plant J.* 2005;43(4):586-596. DOI 10.1111/j.1365-313X.2005. 02476.x.

Coen E.S., Meyerowitz E.M. The war of the whorls: genetic interactions controlling flower development. *Nature.* 1991;353(6339):31-37. DOI 10.1038/353031a0.

Ditta G., Pinyopich A., Robles P., Pelaz S., Yanofsky M.F. The *SEP4* gene of *Arabidopsis thaliana* functions in floral organ and meristem identity. *Curr. Biol.* 2004;14(21):1935-1940. DOI 10.1016/ j.cub.2004.

Dong T., Hu Z., Deng L., Wang Y., Zhu M., Zhang J., Chen G. A tomato MADS-box transcription factor, SlMADS1, acts as a negative regu-

lator of fruit ripening. *Plant Physiol.* 2013;163(2):1026-1036. DOI 10.1104/pp.113.224436.

Dyachenko E.A., Slugina M.A. Intraspecific variability of the *Sus1* sucrose synthase gene in *Pisum sativum* accessions. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding.* 2018;22(1):108-114. DOI 10.18699/VJ18.338. (in Russian)

Ferrario S., Immink R.G., Shchennikova A., Busscher-Lange J., Angenent G.C. The MADS box gene *FBP2* is required for SEPALLATA function in petunia. *Plant Cell.* 2003;15(4):914-925. DOI 10.1105/tpc.010280.

Ferrario S., Shchennikova A.V., Franken J., Immink R.G., Angenent G.C. Control of floral meristem determinacy in petunia by MADS-box transcription factors. *Plant Physiol.* 2006;140(3):890-898. DOI 10.1104/pp.105.072660.

Goloveshkina E.N., Shulga O.A., Shchennikova A.V., Kamionskaya A.M., Skryabin K.G. Functional characterization of chrysanthemum SEPALLATA3 homologs CDM77 and CDM44 in transgenic tobacco plants. *Dokl. Biol. Sci.* 2012;443:87-90. DOI 10.1134/S0012496612020020.

Honma T., Goto K. Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. *Nature.* 2001;409(6819):525-529. DOI 10.1038/35054083.

Hugouvieux V., Silva C.S., Jourdain A., Stigliani A., Charras Q., Conn V., Conn S.J., Carles C.C., Parcy F., Zubieta C. Tetramerization of MADS family transcription factors SEPALLATA3 and AGAMOUS is required for floral meristem determinacy in *Arabidopsis. Nucleic Acids Res.* 2018;46(10):4966-4977. DOI 10.1093/nar/gky205.

Immink R.G., Tonaco I.A., de Folter S., Shchennikova A., van Dijk A.D., Busscher-Lange J., Borst J.W., Angenent G.C. SEPALLATA3: the 'glue' for MADS box transcription factor complex formation. *Genome Biol.* 2009;10(2):R24. DOI 10.1186/gb-2009-10-2-r24.

Jang S., Hong M.Y., Chung Y.Y., An G. Ectopic expression of tobacco MADS genes modulates flowering time and plant architecture. *Mol. Cells.* 1999;9(6):576-586. http://www.molcells.org/journal/view.html?year=1999&volume=9&number=6&spage=576.

Jha P., Ochatt S.J., Kumar V. WUSCHEL: a master regulator in plant growth signaling. *Plant Cell. Rep.* 2020;39(4):431-444. DOI 10.1007/s00299-020-02511-5.

Karlova R., Chapman N., David K., Angenent G.C., Seymour G.B., de Maagd R.A. Transcriptional control of fleshy fruit development and ripening. *J. Exp. Bot.* 2014;65(16):4527-4541. DOI 10.1093/jxb/eru316.

Kumar S., Stecher G., Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0. molecular biology and evolution. *Mol. Biol. Evol.* 2016;33:1870-1874. DOI 10.1093/molbev/msw054.

Lenhard M., Bohnert A., Jürgens G., Laux T. Termination of stem cell maintenance in *Arabidopsis* floral meristems by interactions between *WUSCHEL* and *AGAMOUS. Cell.* 2001;105(6):805-814. DOI 10.1016/s0092-8674(01)00390-7.

Leseberg C.H., Eissler C.L., Wang X., Johns M.A., Duvall M.R., Mao L. Interaction study of MADS-domain proteins in tomato. *J. Exp. Bot.* 2008;59:2253-2265. DOI 10.1093/jxb/ern094.

Li N., Huang B., Tang N., Jian W., Zou J., Chen J., Cao H., Habib S., Dong X., Wei W., Gao Y., Li Z. The MADS-box gene *SlMBP21* regulates sepal size mediated by ethylene and auxin in tomato. *Plant Cell Physiol.* 2017;58(12):2241-2256. DOI 10.1093/pcp/pcx158.

Melzer R., Verelst W., Theissen G. The class E floral homeotic protein SEPALLATA3 is sufficient to loop DNA in 'floral quartet'-like complexes *in vitro. Nucleic Acids Res.* 2009;37(1):144-157. DOI 10.1093/nar/gkn900.

Pelaz S., Ditta G.S., Baumann E., Wisman E., Yanofsky M.F. B and C floral organ identity functions require *SEPALLATA* MADS-box genes. *Nature.* 2000;405(6783):200-203. DOI 10.1038/35012103.

Pelaz S., Gustafson-Brown C., Kohlami S.E., Crosby W.L., Yanofsky M.F. *APETALA1* and *SEPALLATA3* interact to promote flower development. *Plant J.* 2001a;26:385-394. DOI 10.1046/j.1365-313x.2001.2641042.x.

Pelaz S., Tapia-López R., Alvarez-Buylla E.R., Yanofsky M.F. Conversion of leaves into petals in *Arabidopsis. Curr. Biol.* 2001b;11(3):182-184. DOI 10.1016/s0960-9822(01)00024-0.

Pnueli L., Abu-Abeid M., Zamir D., Nacken W., Schwarz-Sommer Z., Lifschitz E. The MADS box gene family in tomato: temporal expression during floral development, conserved secondary structures and homology with homeotic genes from *Antirrhinum* and *Arabidopsis. Plant J.* 1991;1(2):255-266. DOI 10.1111/j.1365-313X.1991.00255.x.

Pnueli L., Hareven D., Broday L., Hurwitz C., Lifschitz E. The *TM5* MADS box gene mediates organ differentiation in the three inner whorls of tomato flowers. *Plant Cell.* 1994;6(2):175-186. DOI 10.1105/tpc.6.2.175.

Qi X., Liu C., Song L., Li M. PaMADS7, a MADS-box transcription factor, regulates sweet cherry fruit ripening and softening. *Plant Sci.* 2020;301:110634. DOI 10.1016/j.plantsci.2020.110634.

Roldan M.V.G., Périlleux C., Morin H., Huerga-Fernandez S., Latrasse D., Benhamed M., Bendahmane A. Natural and induced loss of function mutations in *SlMBP21* MADS-box gene led to *jointless-2* phenotype in tomato. *Sci. Rep.* 2017;7(1):4402. DOI 10.1038/s41598-017-04556-1.

Schilling S., Pan S., Kennedy A., Melzer R. MADS-box genes and crop domestication: the jack of all traits. *J. Exp. Bot.* 2018;69(7):1447-1469. DOI 10.1093/jxb/erx479.

Schmidt G.W., Delaney S.K. Stable internal reference genes for normalization of real-time RT-PCR in tobacco (*Nicotiana tabacum*) during development and abiotic stress. *Mol. Genet. Genom.* 2010;283(3):233-241. DOI 10.1007/s00438-010-0511-1.

Shchennikova A.V., Shulga O.A., Immink R., Skryabin K.G., Angenent G.C. Identification and characterization of four chrysanthemum MADS-box genes, belonging to the *APETALA1/FRUITFULL* and *SEPALLATA3* subfamilies. *Plant Physiol.* 2004;134(4):1632-1641. DOI 10.1104/pp.103.036665.

Shima Y., Fujisawa M., Kitagawa M., Nakano T., Kimbara J., Nakamura N., Shiina T., Sugiyama J., Nakamura T., Kasumi T., Ito Y. Tomato *FRUITFULL* homologs regulate fruit ripening via ethylene biosynthesis. *Biosci. Biotechnol. Biochem.* 2014;78(2):231-237. DOI 10.1080/09168451.2014.878221.

Slugina M.A., Dyachenko E.A., Kochieva E.Z., Shchennikova A.V. Structural and functional diversification of *SEPALLATA* genes *TM5* and *RIN* in tomato species (section Lycopersicon). *Dokl. Biochem. Biophys.* 2020;492(1):152-158. DOI 10.1134/S1607672920030102.

Smaczniak C., Immink R.G., Angenent G.C., Kaufmann K. Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies. *Development.* 2012;139(17):3081-3098. DOI 10.1242/dev.074674.

Theissen G. Development of floral organ identity: stories from the MADS house. *Curr. Opin. Plant Biol.* 2001;4(1):75-85. DOI 10.1016/s1369-5266(00)00139-4.

Theissen G., Saedler H. Plant biology. Floral quartets. *Nature.* 2001;409(6819):469-471. DOI 10.1038/35054172.

Urbanus S.L., de Folter S., Shchennikova A.V., Kaufmann K., Immink R.G., Angenent G.C. In planta localisation patterns of MADS domain proteins during floral development in *Arabidopsis thaliana. BMC Plant Biol.* 2009;9:5. DOI 10.1186/1471-2229-9-5.

Vrebalov J., Pan I.L., Arroyo A.J., McQuinn R., Chung M., Poole M., Rose J., Seymour G., Grandillo S., Giovannoni J., Irish V.F. Fleshy fruit expansion and ripening are regulated by the tomato *SHATTERPROOF* gene *TAGL1. Plant Cell.* 2009;21(10):3041-3062. DOI 10.1105/tpc.109.066936.

Vrebalov J., Ruezinsky D., Padmanabhan V., White R., Medrano D., Drake R., Schuch W., Giovannoni J. A MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (rin) locus. *Science.* 2002;296:343-346. DOI 10.1126/science.1068181.

А.В. Нежданова, М.А. Слугина, Е.А. Дьяченко
А.М. Камионская, Е.З. Кочиева, А.В. Щенникова

Структурно-функциональный анализ MADS-box
гена *SlMADS5* томата *Solanum lycopersicum* L.

2021
25•5

Wang R., Tavano E.C.D.R., Lammers M., Martinelli A.P., Angenent G.C., de Maagd R.A. Re-evaluation of transcription factor function in tomato fruit development and ripening with CRISPR/Cas9-mutagenesis. *Sci. Rep.* 2019;9(1):1696. DOI 10.1038/s41598-018-38170-6.

Wang Y., Zhang J., Hu Z., Guo X., Tian S., Chen G. Genome-wide analysis of the MADS-box transcription factor family in *Solanum lycopersicum*. *Int. J. Mol. Sci.* 2019;20(12):2961. DOI 10.3390/ijms20122961.

Weigel D., Alvarez J., Smyth D.R., Yanofsky M.F., Meyerowitz E.M. *LEAFY* controls floral meristem identity in *Arabidopsis*. *Cell*. 1992; 69(5):843-859. DOI 10.1016/0092-8674(92)90295-n.

Zhang J., Hu Z., Wang Y., Yu X., Liao C., Zhu M., Chen G. Suppression of a tomato SEPALLATA MADS-box gene, *SlCMB1*, generates altered inflorescence architecture and enlarged sepals. *Plant Sci.* 2018a;272:75-87. DOI 10.1016/j.plantsci.2018.03.031.

Zhang J., Hu Z., Yao Q., Guo X., Nguyen V., Li F., Chen G. A tomato MADS-box protein, SlCMB1, regulates ethylene biosynthesis and carotenoid accumulation during fruit ripening. *Sci. Rep.* 2018b;8(1): 3413. DOI 10.1038/s41598-018-21672-8.

Zhang S., Lu S., Yi S., Han H., Liu L., Zhang J., Bao M., Liu G. Functional conservation and divergence of five SEPALLATA-like genes from a basal eudicot tree, *Platanus acerifolia*. *Planta*. 2017; 245(2):439-457. DOI 10.1007/s00425-016-2617-0.

Zhao H.B., Jia H.M., Wang Y., Wang G.Y., Zhou C.C., Jia H.J., Gao Z.S. Genome-wide identification and analysis of the MADS-box gene family and its potential role in fruit development and ripening in red bayberry (*Morella rubra*). *Gene*. 2019;717:144045. DOI 10.1016/j.gene.2019.144045.

ORCID ID

A.V. Nezhdanova orcid.org/0000-0003-4175-3175
M.A. Slugina orcid.org/0000-0003-1281-3837
E.A. Dyachenko orcid.org/0000-0002-0570-9751
A.M. Kamionskaya orcid.org/0000-0001-9815-9578
E.Z. Kochieva orcid.org/0000-0002-6091-0765
A.V. Shchennikova orcid.org/0000-0003-4692-3727

Phylogenetic and structural analysis of annexins in pea (*Pisum sativum* L.) and their role in legume-rhizobial symbiosis development

O.A. Pavlova, I.V. Leppyanen, D.V. Kustova, A.D. Bovin, E.A. Dolgikh ✉

All-Russia Research Institute for Agricultural Microbiology, Pushkin, St. Petersburg, Russia
✉ dol2helen@yahoo.com

Abstract. Annexins as Ca^{2+}/phospholipid-binding proteins are involved in the control of many biological processes essential for plant growth and development. In a previous study, we had shown, using a proteomic approach, that the synthesis of two annexins is induced in pea roots in response to rhizobial inoculation. In this study, phylogenetic analysis identified these annexins as PsAnn4 and PsAnn8 based on their homology with annexins from other legumes. The modeling approach allowed us to estimate the structural features of these annexins that might influence their functional activity. To verify the functions of these annexins, we performed comparative proteomic analysis, experiments with calcium influx inhibitors, and localization of labeled proteins. Essential down-regulation of PsAnn4 synthesis in a non-nodulating pea mutant P56 (*sym10*) suggests an involvement of this annexin in the rhizobial symbiosis. Quantitative RT-PCR analysis showed that PsAnn4 was upregulated at the early stages of symbiosis development, starting from 1–3 days after inoculation to up to 5 days after inoculation, while experiments with the Ca^{2+} channel blocker $LaCl_3$ revealed its negative influence on this expression. To follow the PsAnn4 protein localization in plant cells, it was fused to the fluorophores such as red fluorescent protein (RFP) and yellow fluorescent protein (YFP) and expressed under the transcriptional regulation of the 35S promoter in *Nicotiana benthamiana* leaves by infiltration with *Agrobacterium tumefaciens*. The localization of PsAnn4 in the cell wall or plasma membrane of plant cells may indicate its participation in membrane modification or ion transport. Our results suggest that PsAnn4 may play an important role during the early stages of pea-rhizobial symbiosis development.
Key words: legume-rhizobial symbiosis; pea annexins; three-dimensional modeling; proteomics; calcium inhibitors; localization.

Филогенетический и структурный анализ аннексинов у гороха (*Pisum sativum* L.) и их роль в развитии бобово-ризобиального симбиоза

О.А. Павлова, И.В. Леппянен, Д.В. Кустова, А.Д. Бовин, Е.А. Долгих ✉

Всероссийский научно-исследовательский институт сельскохозяйственной микробиологии, Пушкин, Санкт-Петербург, Россия
✉ dol2helen@yahoo.com

Аннотация. Аннексины являются Ca^{2+}/фосфолипид-связывающими белками, которые вовлечены в контроль многих биологических процессов, необходимых для роста и развития растений. Ранее выполненный протеомный анализ позволил нам выявить два аннексина, синтез которых усиливается в ответ на ризобиальную инокуляцию. В этой работе с помощью филогенетического анализа два аннексина были классифицированы как PsAnn4 и PsAnn8 на основании их гомологии с аннексинами других бобовых растений. С помощью молекулярного моделирования мы изучили структурные особенности этих аннексинов, которые могут влиять на их функциональную активность. Для анализа функции PsAnn4 и PsAnn8 были проведены сравнительный протеомный анализ, эксперименты с ингибиторами поступления кальция в клетку и локализация в тканях растений. Отсутствие активации синтеза PsAnn4 у мутанта гороха P56 (*sym10*), не способного формировать клубеньки, предполагает участие этого аннексина в бобово-ризобиальном симбиозе. Количественная ПЦР, совмещенная с обратной транскрипцией, показала, что экспрессия гена *PsAnn4* увеличивается на ранних стадиях развития симбиоза начиная с 1–3-го дня после инокуляции до 5-го дня, тогда как блокатор Ca^{2+} канала LaCl3 подавляет эту экспрессию. Для изучения локализации PsAnn4 в клетках растений были получены конструкции для синтеза этого белка, слитого с такими флуорофорами, как красный флуоресцентный белок (RFP) и желтый флуоресцентный белок (YFP) при транскрипционной регуляции под промотором 35S в листьях *Nicotiana benthamiana* при инфильтрации *Agrobacterium tumefaciens*. Локализация PsAnn4 в клеточной

О.А. Павлова, И.В. Леппянен
Д.В. Кустова, А.Д. Бовин, Е.А. Долгих

Филогенетический и структурный анализ
аннексинов у гороха

2021
25•5

стенке или плазматической мембране клеток растений указывает на возможность участия этого аннексина в ионном транспорте или модификации мембраны. Обсуждается возможная роль аннексина PsAnn4 в регуляции ранних стадий развития симбиоза у гороха.

Ключевые слова: бобово-ризобиальный симбиоз; аннексины гороха; 3D-моделирование; протеомика; ингибиторы кальция; локализация.

Introduction

Annexins are of particular research interest due to their ability to regulate various aspects of plant growth and development. Annexins belong to the evolutionarily conserved superfamily of proteins that are involved in Ca^{2+}-dependent or Ca^{2+}-independent binding with membrane phospholipids (Laohavisit, Davies, 2011; Davies, 2014). Most annexins have four putative annexin repeats of around 70 amino acids, with the conservative repeat GxGT-(38 residues)-D/E, which confers Ca^{2+}/phospholipid-binding activity to these proteins (Gerke, Moss, 2002; Laohavisit, Davies, 2011). In addition, some plant annexins have motifs demonstrating F-actin binding and peroxidase and ATPase/GTPase activities (Mortimer et al., 2008; Konopka-Postupolska et al., 2011).

Despite the general structural similarity of these proteins, the functions of annexins are diverse, and individual annexins may have specific activities. Annexins are involved in a wide variety of essential cellular processes, including the regulation of membrane organization, vesicle trafficking, cytoskeletal dynamics, exocytosis, cell cycle control, ion transport, and signal transduction (Laohavisit, Davies, 2011; Clark et al., 2012; Davies, 2014). Annexins as phospholipid-binding proteins are being implicated in the fusion of membrane vesicles, as was shown for annexins from bell pepper and cotton (Clark et al., 2012; Lizarbe et al., 2013). They are also involved in the regulation of exocytosis, e. g., annexins in *Zea mays* root cap cells (Carroll et al., 1998). Moreover, annexins can function as cationic channels activated by various stimuli in cells. Annexins can influence the Ca^{2+} influx in plant cells, as was demonstrated for a *Capsicum annuum* annexin, which has Ca^{2+}-channel activity (Hofmann et al., 2000). The *Arabidopsis thaliana* annexin AtAnn1, which is expressed in root cells, exhibits pH-dependent cation-channel activity, while *Z. mays* annexins cause active conductivity of Ca^{2+} in lipid bilayers at slightly acidic pH (Gorecka et al., 2005; Laohavisit et al., 2009). Since annexins can be Ca^{2+} sensors, these proteins are likely to be involved in signal transduction; for example, the annexin from *Triticum aestivum* was suggested to be engaged in low-temperature signaling (Breton et al., 2000).

Participation of annexins in the responses to cold, oxidative, and saline stresses is well-studied in plants (Mortimer et al., 2008; Clark et al., 2012; Espinoza et al., 2017). The annexin AtAnn1 from *A. thaliana* is involved in plant protection against oxidative stress (Konopka-Postupolska et al., 2009). The overexpression of AtAnn has been found to confer tolerance to drought and salt stresses and fungal attack in transgenic plants (Konopka-Postupolska et al., 2009). Similarly, the overexpression of the wild tomato (*Solanum pennellii*) annexin SpAnn2 in cultivated tomato *Solanum lycopersicum* enhances drought and salt tolerance through the elimination of reactive oxygen species (ROS) (Ijaz et al., 2017).

Some annexins are also known to be activated in plants during interaction with plant-growth promoting bacteria

(Kwon et al., 2016) and the development of mutualistic symbioses (De Carvalho-Niebel et al., 1998, 2002; Wienkoop, Saalbach, 2003; Manthey et al., 2004; Talukdar et al., 2009; Limpens et al., 2013; Breakspear et al., 2014; Carrasco-Castilla et al., 2018). During legume-rhizobial symbiosis, physiological changes occur, which are necessary for rhizobial infection and nodule organogenesis, such as the stimulation of ion fluxes, membrane depolarization, ROS production, cytoplasm alkalinization, perinuclear calcium oscillations, and cytoskeletal rearrangements. In *Medicago truncatula*, the transcription of *MtAnn1* is activated directly by Nod factors or inoculation with rhizobia in epidermal cells and later in cortical cells (De Carvalho-Niebel et al., 1998, 2002; Breakspear et al., 2014). Studies using confocal microscopy showed GFP-labeled MtAnn1 to be localized in the cytoplasm, but protein accumulation in response to inoculation occurred at the periphery of the nucleus. MtAnn1 has been shown to be able to bind to the membrane phospholipid phosphatidylserine. Therefore, MtAnn1 is probably related to the events occurring at the early stages of symbiosis, leading to bacterial infection or nodule organogenesis (De Carvalho-Niebel et al., 2002).

Transcriptome profiling of roots inoculated with rhizobia revealed enhanced expression of *MtAnn2*, as well as *MtAnn1* (Manthey et al., 2004). The expression of the *MtAnn2* gene is associated with cell division in the nodule primordium (Manthey et al., 2004). Proteomic analysis revealed the MtAnn2 protein presence in lipid rafts from root plasma membrane preparations (Lefebvre et al., 2007). Another annexin MtAnn3 was found to be important for root hair deformations in *M. truncatula* (Gong et al., 2012). The increased expression of *MtAnn1* and *MtAnn2* is also associated with the early stages of AM fungal symbiosis, which corresponds to the stages of pre-infection and infection in this type of symbiosis (Manthey et al., 2004). This may indicate the general role of these annexins in the regulation of signaling pathways that lead to the development of two types of symbiosis.

A protein homologous to MtAnn1 – PvAnn1 from *Phaseolus vulgaris* – is activated at the early stages of symbiosis development (Jáuregui-Zúñiga et al., 2016; Carrasco-Castilla et al., 2018). The stimulation of Ca^{2+} ion transfer through the plasma membrane and ROS production caused by Nod factors constitute an early response in the signal transduction pathway. Analysis of *PvAnn1*-RNAi transgenic roots inoculated with rhizobia showed a decrease in ROS production and Ca^{2+} influx into the cells, which resulted in impaired progression and decreased numbers of infection threads and nodules (Carrasco-Castilla et al., 2018). Taken together, these findings point to the involvement of PvAnn1 in the regulation of signal transduction at early stages.

Previously performed proteomic analysis in pea (*Pisum sativum* L.) allowed us to reveal two annexins, the synthesis of which was increased in response to inoculation with

Rhizobium leguminosarum bv. *viciae* RCAM1026 in 24 h (Leppyanen et al., 2018). In this work, searching in the recently released pea genome database using available coding sequences for annexin genes from *M. truncatula* and *P. vulgaris* revealed 15 annexins in pea. Phylogenetic analysis showed the relationship among members of the annexin superfamily in other legumes and allowed the identification of two previously revealed pea annexins responsive to rhizobial inoculation as PsAnn4 and PsAnn8 based on their homology with the *M. truncatula* and *P. vulgaris* proteins. To verify the function of these annexins, we performed comparative proteomic analysis using pea mutant P56 (*sym10*) unable to form symbiosis and wild type cv. Frisson. The approaches employed included quantitative RT-PCR, experiments with calcium channel inhibitors, and localization of labeled proteins.

Materials and methods

Plant material and bacterial strain. Pea *Pisum sativum* L. seeds cv. Frisson were sterilized with sulphuric acid for 5 min, washed with water 3 times, transferred on 1 % water agar plates and germinated at room temperature in the dark. 4–5 days-old seedlings were transferred into pots with vermiculite saturated with Jensen medium (van Brussel et al., 1982), grown in a growth chamber at 21 °C at 16 h light/8 h dark cycles, 60 % humidity. For experiments with inhibitor, the Ca²⁺ channels blocker LaCl₃, the plants were grown in pots saturated with Jensen medium with 100 μM CaCl₂ × 2 H₂O. The *Rhizobium leguminosarum* bv. *viciae* strain RCAM 1026 (WDCM 966) was cultivated at 28 °C on TY (Orosz et al., 1973) agar medium with 0.5 mg/ml of streptomycin. Fresh liquid bacterial culture was grown in B⁻ medium (Van Brussel et al., 1977) and the optical density of the suspension at 600 nm (OD$_{600}$) was adjusted to 0.5. Pea seedlings were inoculated with 2 ml of *R. leguminosarum* bv. *viciae* per plant. Pea roots (segments of main roots susceptible for rhizobial infection without lateral roots) were harvested 1 day after inoculation (dai).

Nicotiana benthamiana seeds were surface sterilized with 10 % hypochlorite for 10 min, washed with water 5 times and left for imbibition on a plate with sterile filter paper at 4 °C. All seeds were germinated in a large plastic box with soil for seven days, and then transferred into individual pots with soil. Plants were grown at 23 °C with 16 h light/8 h dark cycles, 60 % humidity.

Phylogenetic analysis. Multiple sequence alignments were performed using ClustalΩ http://www.clustal.org/omega/ (Sievers et al., 2011). The phylogenetic tree was generated with the Maximum Likelihood method using MEGA X https://www.megasoftware.net/ with 1000 bootstrap replicates. The domain composition of the corresponding encoded proteins was assessed using PFAM https://www.sanger.ac.uk/science/tools/pfam (Bateman et al., 2004).

Protein homology modeling was performed in Modeller 9.20 https://salilab.org/modeller/9.20/release.html (Webb, Sali, 2016). Visualization of the three-dimensional structure was obtained using the PyMol program https://pymol.org/2/ (Ordog, 2008). The three-dimensional crystal structure of the GhAnn1 *G. hirsutum* protein (Hu et al., 2008) was used as a template for building the model. To refine the model, the energy was minimized twice by the conjugate gradient method (VTFM) and the method of molecular dynamics in vacuum. The reliability of the model was calculated by the formula

$$P = (1 - F(Z)) \cdot 100 \text{ \%},$$

where Z is the estimation of discretely optimized protein energy, F is the Gaussian function with $\mu = 0$ and $\sigma^2 = 1$.

Isolation of total protein from pea roots. A modified method was used to isolate proteins from pea roots (Dam et al., 2014). 100 mg of the roots were ground in liquid nitrogen, then extraction buffer (0.1 M tris-HCl (pH 8.0), 30 % sucrose, 10 mM dithiothreitol (DTT), 2 % sodium dodecyl sulfate (SDS), a mixture of protease inhibitors (Sigma-Aldrich, USA) was added to the material and extraction was performed at +4 °C. After centrifugation at 12 000 g for 15 min, the supernatant was mixed in a 1:1 ratio with phenol (pH 8.0) (Thermo Fisher Scientific, USA), centrifuged at 12 000 g for 5 min. The upper phase was taken for precipitation of proteins. Five volumes of cold 100 mM ammonium acetate in methanol were added and incubated for 30 min at –20 °C. After centrifugation at 12 000 g for 5 min, the pellet was washed twice with 100 mM ammonium acetate in methanol and twice with 80 % acetone. The precipitate was dried in air and dissolved in the buffer for isoelectric focusing (25 mM tris-HCl (pH 8.0), 9 M urea, 4 % CHAPS, 50 mm DTT, 0.2 % ampholytes (Bio-Rad Laboratories, USA)). Protein concentration was measured using Bradford assay (Bradford, 1976).

Two-dimensional differential gel electrophoresis. Two-dimensional differential gel electrophoresis (DIGE) of proteins was performed using staining of samples with various fluorescent dyes (Voss, Haberl, 2000). The samples were conjugated for 30 min on ice with fluorescent dyes Cyanine 2 or Cyanine 5 (Cy2 or Cy5) in various combinations. The incubation solution contained 400 pM of each dye dissolved in dimethylformamide for 30 min on ice. The reaction was stopped by adding 10 mM L-lysine (Sigma-Aldrich), followed by incubation on ice for 10 min. After that, the control and experimental samples were mixed, DTT and ampholytes (50 mM DTT, 0.2 % ampholytes (Bio-Rad Laboratories) were added. Passive in-gel rehydration with immobilized pH gradient (Bio-Rad Laboratories) was performed overnight at room temperature. The total amount of sample applied to 7 cm gel (pH 3–10, Bio-Rad Laboratories) was up to 100 μg. Isoelectric focusing (IEF) was performed in a Protean IEF system (Bio-Rad Laboratories) at a temperature of 20 °C, the samples were desalted at 250 V for 15 min, after which the voltage was linearly increased to 4,000 V for 2 hours, then IEF was carried out with increasing voltage up to 10000 V. Before electrophoresis in polyacrylamide gel (PAGE), protein recovery was carried out in buffer with DTT (6 M urea, 0.375 M tris, pH 8.8, 2 % SDS, 20 % glycerol, 2 % DDT) for 10 min followed by alkylation in iodoacetamide buffer (6 M urea, 0.375 M tris, pH 8.8, 2 % SDS, 20 % glycerol, 2.5 % iodoacetamide) for 15 min. The second direction of two-dimensional electrophoresis was carried out in tris-glycine buffer (25 mM Tris-HCl, 192 mM glycine, 0.1 % SDS, pH 8.3) in 15 % polyacrylamide gel using a 4 % stacking gel. After separation of proteins the gels were visualized using a laser scanner Typhoon FLA 9500 (GE Healthcare, Germany).

О.А. Павлова, И.В. Леппянен
Д.В. Кустова, А.Д. Бовин, Е.А. Долгих

Филогенетический и структурный анализ
аннексинов у гороха

2021
25·5

Table 1. List of primers used in this study

Gene name	Forward primer	Reverse primer
Ubiquitin	5′-ATGCAGATC/TTTTGTGAAGAC-3′	5′-ACCACCACGG/AAGACGGAG-3′
PsAnn4	5′-CATCTTTGGGCACTTGAATCC-3′	5′-TATCTTTGCCTCCGCTTTTGCTAT-3′
PsAnn8	5′-GAACATGGCGTCTCCGTCAGTAA-3′	5′-CTTCTCGGCCCTCGTAACAATCA-3′
PsEnod5	5′-CGATACTATCGATGTAGTGG-3′	5′-GACTGTAATTGACCTTCACC-3′
PsNIN	5′-CCGCAAAGAGCATCGGTGTATG-3′	5′-GCATAGAAAGATCCAATCTGTATAGC-3′

Mass spectrometry. The proteins were rehydrated in trypsin solution (20 ng/µl trypsin, 30 mM tris, pH 8.2) on ice for 1 h and then incubated for 1 h at 56 °C. The peptides were extracted from the gel with 50 % acetonitrile, 0.1 % formic acid. This solution was evaporated in vacuum concentrator CentriVap (Labconco) at 4 °C and dissolved in phase A (5 % acetonitrile, 0.1 % formic acid). Mass spectrometry was performed using Agilent ESI-Q-TOF 6538 UHD (Agilent Technologies) combined with high performance liquid chromatograph Agilent 1260 (Agilent Technologies). Chromatography was performed in system water – acetonitrile in the presence of 0.1 % formic acid (phase A – 5 % acetonitrile with 0.1 % formic acid, phase B – 90 % acetonitrile with 0.1 % formic acid) in the gradient of acetonitrile (from 5 to 60 % phase B for 25 min and to 100 % phase B for 5 min) on Zorbax 300SB-C18 column 3.5 µm, 150 mm length (Agilent Technologies) with flow rate 15 µl/min.

RNA extraction and quantitative reverse transcription PCR (RT-PCR). RNA extraction and RT-PCR were performed as described previously (Kirienko et al., 2018). The quantitative RT-PCR analysis was performed on a CFX-96 real-time PCR detection system with C1000 thermal cycler (Bio-Rad Laboratories). All primer pairs (Table 1) were designed using the Vector NTI program and produced by the Evrogen company (www.evrogen.com). PCR amplification specificity was verified using a dissociation curve (55–95 °C). mRNA levels were normalized against Ubiquitin and values were calculated as ratios relative to non-inoculated root expression levels. The data of two-three independent biological experiments were analysed. Statistical analysis was conducted by Student's test ($p < 0.05$) to assess the differences between variants.

Genetic constructs for plant transformation. To obtain the pBIN19 vector for plant transformation, carrying the gene of interest, the coding sequence of PsAnn4 gene without stop-codon has been amplified using cDNA as a template with corresponding primers (see Table 1). Total RNA was isolated from 2 dai pea roots of cv. Frisson. Amplification was done using Phusion Flash High-Fidelity PCR Master Mix (Thermo Scientific). The amplified products were restricted with XbaI and EcoRI and subcloned in the pMON vector under 35S promoter in the frame with the sequences encoding RFP or YFP and nopaline synthase terminator (Tnos). The inserts were verified by sequencing. The cassette composed of the 35S promoter, gene of interest fused with RFP or YFP and Tnos was excised from pMON using HindIII, SmaI and cloned in the pBIN19. All verified constructs were transferred into the Agrobacterium tumefaciens LBA4404.

Transient protein expression in N. benthamiana leaves. A. tumefaciens strain LBA4404 was used for infiltration in N. benthamiana leaves. Bacterial culture was grown at 28 °C overnight, then centrifuged at 3000 g and resuspended in 10 mM MES-KOH, 10 mM MgCl$_2$ and 0.5 mM acetosyringone up to culture density OD$_{600}$ = 0.5. Bacterial cells were infiltrated into the leaves of 3-week-old N. benthamiana. Plants were analyzed 48–96 h after infiltration.

Results

Phylogenetic analysis of annexins in pea and other legumes

The search of the sequences presumably coding for annexins in legumes was performed using BlastX with 8 previously revealed M. truncatula and 13 P. vulgaris nucleotide sequences encoding these proteins (Kodavali et al., 2013; Carrasco-Castilla et al., 2018) as queries against different plant sequence databases: https://phytozome.jgi.doe.gov/pz/portal.html for M. truncatula and P. vulgaris, http://www.kazusa.or.jp/lotus/ for L. japonicus, and the URGI database v. 1 https://urgi.versailles.inra.fr/blast for P. sativum L. (Clark et al., 2001; Carrasco-Castilla et al., 2018; Kreplak et al., 2019). As a result, we were able to identify 18 coding sequences (CDSs) for annexins in M. truncatula, 15 in P. sativum L., and 13 in L. japonicus (Table 2). Twenty-three genes had been previously found to encode annexins in soybean (Feng et al., 2013). The coding sequences for annexins from P. sativum were named based on their phylogenetic relationships with the corresponding homologous sequences from M. truncatula and P. vulgaris (see Table 2) (Clark et al., 2012; Kodavali et al., 2013; Carrasco-Castilla et al., 2018).

The phylogenetic analysis (Fig. 1) was performed using the deduced amino acid sequences of annexins found and annotated for P. sativum along with those of other legumes (M. truncatula, P. vulgaris, Lotus japonicus, and Glycine max) and non-legumes (A. thaliana, G. raimondii), which were available in the Phytozome database v. 12.1 and other databases.

Based on our analysis, the previously found MtAnn1 (Medtr8g038210) and PvAnn1 (Phvul.011g209300) clustered in the subclade with proteins corresponding to P. sativum Psat4g147120 and Psat4g191080, named PsAnn1a and PsAnn1b (see Table 2). Revealed in M. truncatula MtAnn2 (Medtr8g038220) and P. vulgaris PvAnn2 (Phvul.011g209200) clustered in the subclade with Psat4g191040, named PsAnn2.

Two previously described pea annexins induced in roots in response to rhizobial inoculation (Leppyanen et al., 2018)

Table 2. Accession numbers and annotations of annexin sequences in *P. sativum*, *M. truncatula*, *P. vulgaris*, and *L. japonicus*

Gene accession number *P. vulgaris*	Protein	Gene accession number *M. truncatula*	Protein	Gene accession number *P. sativum*	Protein	Gene accession number *L. japonicus*
Phvul.011G209300.1	PvAnn1	Medtr8g038210.1	MtAnn1	Psat4g147120.1, Psat4g191080.1	PsAnn1a, PsAnn1b	Lj0g3v0203419.1
Phvul.011G209200.1	PvAnn2	Medtr8g038220.2	MtAnn2	Psat4g191040.1	PsAnn2	Lj0g3v0363079.1
Phvul.011G209500.1	PvAnn3	Medtr8g038150.1	MtAnn3	Psat4g146920.1	PsAnn3	Lj0g3v0203449.1
Phvul.005G030100.1	PvAnn4	Medtr3g018780.1	MtAnn4	Psat5g217440.1	PsAnn4	Lj0g3v0261959.1
Phvul.004G146900.1	PvAnn5	Medtr6g071595.2	MtAnn5	Psat1g028960.1	PsAnn5	Lj2g3v0636730.1, Lj4g3v2858470.1
Phvul.005G030200.1	PvAnn6	Medtr3g018790.1	MtAnn6	Psat5g217920.1	PsAnn6	Lj0g3v0261939.1
Phvul.002G332200.1	PvAnn7	Medtr8g107640.1	MtAnn7	Psat7g000680.1	PsAnn7	Lj4g3v3117410.1
Phvul.008G173100.1	PvAnn8	Medtr5g063670.1	MtAnn8	Psat2g074960.1	PsAnn8	Lj0g3v0166899.1
Phvul.006G123400.1	PvAnn9	Medtr2g031980.1	MtAnn9	Psat1g164360.1	PsAnn9	–
Phvul.003G013700.1	PvAnn10	Medtr1g033560.1	MtAnn10	Psat6g095440.1	PsAnn10	Lj2g3v0062280.1, Lj5g3v0768290.1
Phvul.011G209400.1	PvAnn11	Medtr8g038170.1	MtAnn11	Psat4g146960.1	PsAnn11	Lj0g3v0203439.1
Phvul.002G255700.1	PvAnn12	Medtr0276s0050.1	MtAnn12	Psat7g054960.1	PsAnn12	Lj4g3v2823370.1
Phvul.004G052200.1	PvAnn13	Medtr6g028030.1	MtAnn13	Psat1g094800.1	PsAnn13	–
		Medtr8g038180.1	MtAnn14	Psat4g147000.1	PsAnn14	–
		Medtr3g018920.1	MtAnn15	–	–	–
		Medtr1g112520.1	MtAnn16	–	–	–
		Medtr6g071605.1	MtAnn17	–	–	–
		Medtr6g071615.1	MtAnn18	–	–	–

were identified as proteins corresponding to Psat5g217440 and Psat2g074960 coding sequences using a new database https://urgi.versailles.inra.fr/blast for *P. sativum* (see Table 2) (Kreplak et al., 2019). The phylogenetic analysis depicted an additional branch in the phylogenetic group with MtAnn1/ PvAnn1 and MtAnn2/PvAnn2, comprising MtAnn4 (Medtr3g018780), PvAnn4 (Phvul.005g030100), and their homolog Psat5g217440, named PsAnn4 (identified by proteomic screening) (see Table 2). Another previously found pea annexin, Psat2g074960, might be closely related to Medtr5g063670 and Phvul.008G173100.1, defined as *MtAnn8* and *PvAnn8* based on phylogenetic analysis (see Table 2).

Analysis of the domain composition of pea annexins and modeling of three-dimensional structure of PsAnn4 and PsAnn8

Analysis of the domain composition of the corresponding proteins in pea showed the presence of four typical domains of plant annexins (Fig. 2). This suggests that the annexin gene family indeed comprises several members in pea. Although plant annexins have four putative annexin repeats, not all Ca^{2+}-binding motifs in these repeats seem to be functional. In plant annexins, the Ca^{2+}-binding site is highly conservative in the first (I) repeat but is not conservative in the second (II) and third (III) repeats, while in the fourth (IV) repeat moderate conservatism is preserved (see Fig. 2).

The crystal structure of the *Gossypium hirsutum* annexin GhAnn1 bound to calcium was obtained in an earlier study (Hu et al., 2008). Since PsAnn4 and PsAnn8 may be involved in regulation of pea-rhizobial symbiosis, we modeled the three-dimensional (3D) structure of these two annexins using GhAnn1, with 50 % sequence identity for PsAnn4 and 78 % sequence identity for PsAnn8 as a template (Fig. 3, *a*, *b*). The resulting 3D structures of PsAnn4 and PsAnn8 proteins indicated the coordination of calcium ions in the first and fourth annexin repeats. In the first repeat of both proteins, the calcium-binding site of the type II was coordinated by three carbonyl oxygen atoms of the residues Phe-23, Gly-25, and Gly-27, and carboxylate of Glu-67 in PsAnn4 and PsAnn8 (see Fig. 2 and 3, *c*, *d*), as was shown earlier for GhAnn1 (Hu et al., 2008).

We suppose that the second calcium ion is bound in the loop of the fourth annexin repeat of PsAnn4 and PsAnn8 proteins. It is coordinated in the binding site of type II by Ile-254, Lys-256, and Gly-258 in pea annexins (see Fig. 2, 3, *e*, *f*). The third calcium ion (in the binding site of type III) is coordinated by two oxygen atoms of the residues Val-296 and Thr-299 and carboxylate of Glu-304 in this protein (similarly, Val, Thr, and Glu are involved in Ca^{2+} binding in the fourth repeat of GhAnn1) (see Fig. 2, 3, *g*) (Hu et al., 2008). However, in the fourth repeat of PsAnn4 protein, the Val-296 is replaced by Ser and Glu-304 by Lys (see Fig. 2). This might potentially

О.А. Павлова, И.В. Леппянен
Д.В. Кустова, А.Д. Бовин, Е.А. Долгих

Филогенетический и структурный анализ
аннексинов у гороха

2021
25•5

obstruct the binding of the calcium ion, as was shown in our modeling (see Fig. 3, *g*). Although we cannot rule out that this might be due to low homology between PsAnn4 and GhAnn1, which was used as a template in the modeling, the results suggest the potential difference in Ca^{2+} binding between PsAnn4 and PsAnn8 proteins.

Comparative analysis of protein patterns in wild-type and non-nodulating pea mutant

To verify whether the stimulation of synthesis of PsAnn4 and PsAnn8 proteins depends on Nod factor perception, the protein patterns were analyzed in wild-type pea cv. Frisson and a P56 mutant with a defective *sym10* gene (which encodes a putative Nod factor receptor) (Madsen et al., 2003).

Two-dimensional differential in-gel electrophoresis-based proteomics was used to characterize the pattern of protein distribution (Fig. 4). Two spots corresponding to the location of the previously characterized annexins (Leppyanen et al., 2018) were excised from the gel. Mass spectrometric analysis confirmed their identity to annexins Psat5g217440 (PsAnn4) and Psat2g074960 (PsAnn8). Enhanced level of PsAnn4 was found in the inoculated roots of wild type pea plants (cv. Frisson) compared to the inoculated P56 mutant roots.

The amount of PsAnn8 protein was also slightly higher in response to inoculation in the wild type than in the P56 mutant, but not as essential as for PsAnn4. In accordance with this, low amounts of PsAnn4 and PsAnn8 proteins were found in the roots of the P56 mutant and didn't change in response to inoculation. This suggests that the up-regulation of both annexins may depend on Nod factor recognition in pea plants and may be connected with the functioning of these annexin during symbiotic interaction of plants with rhizobia at early stages. Since the increase in the amount of PsAnn4 protein was more significant in response to inoculation, we focused on this annexin in our next experiments.

PsAnn4 expression pattern in response to rhizobial inoculation and treatment with Ca^{2+} inhibitors

The *PsAnn4* expression pattern in response to rhizobial inoculation was analyzed in our experiments (Fig. 5, *a*). A quantitative RT-PCR analysis revealed that *Rhizobium* infection enhanced the *PsAnn4* gene expression at the early stages of nodulation, starting from 1–3 days after inoculation up to 5 days after inoculation, but thereafter their transcript levels did not significantly change upon nodule development (see Fig. 5, *a*). In our experiments the expression of another annexin gene, *PsAnn1a*, the closest homolog of *MtAnn1* gene was also analyzed (see Fig. 5, *b*). As it was expected, the *PsAnn1a* gene expression was primarily enhanced at the early stages of symbiosis development and reached the highest levels in the nodules. Similar pattern had been previously found for *MtAnn1* (De Carvalho-Niebel et al., 1998, 2002). Therefore, up-regulation of *PsAnn4* expression may be related to the early stages of nodulation. The upregulation of the *PsAnn4* transcription level was not as significant as it was at the protein level, which implies that the regulation of this annexin can be mainly achieved at the post-transcriptional and translational level.

To verify the influence of calcium inhibitors on the regulation of *PsAnn4* gene, its expression level was estimated after

Fig. 1. Phylogenetic tree of annexin sequences from legumes (*P. sativum*, *G. max*, *M. truncatula*, and *P. vulgaris*) and non-legumes (*A. thaliana*, *G. barbadense*, and *G. hirsutum*).

The phylogenetic tree was generated with the maximum-likelihood method using MEGAX with 1,000 bootstrap replicates. PsAnn4 and PsAnn8 are indicated in boxes. The annexin sequences from *P. sativum* are indicated with blue circles.

```
           1                                                                                                 103
MtAnn1   (1)   --MATLVVQDSTA-SVQDAELLHKSFEGWGTDEKTVITILGHRNSNQIQQIRKAYEGIYNEDLIKRLESEIKG----------DFEKAVYRWILEPAERDAVL
MtAnn2   (1)   --MATLIAPINHS-PVADAEALHGAFKGWGTDEKSVITILGHRNVYQRQQIRKSYQEQIYQEDILKRLESELSG----------DFERAVYRWMLEPADRDAVL
PvAnn1   (1)   --MATLIAPSNHS-PQEDAEALRKSFEGWGTDENTIIAILGHRTVHQRQQIRRAYEEIYQEDLVKRLSSEIKG----------DFEKAVYRWILEPADRDAIL
PvAnn2   (1)   --MATLIAPITFS-PGADAEDLHKAFEGWGTDEKTVIAILGHRNVYQRQQIKIYEEIYQELLKRLESELTG----------DFERSVYRWMLEPADRDAVL
PsAnn1a  (1)   --MATLIAPSNHS-PQEDAESLRKAFEGWGTDERTVITILGHRNSHQIQQIRRAYEEIYQEDLIKRLESELKG----------DFEKAVYRWILEPADRDAVL
PsAnn1b  (1)   --MATLIAPSNHS-PQEDAESLRKAFEGWGTDERTVITILGHRSHQIQQIRRAYEEIYQEDLIKRLESELKG----------DFEKAVYRWILEPADRDAVL
PsAnn2   (1)   --MATRIAPTNHS-PLADAEALHKAFEGWGTDEKSVITILGHRSVYQRQQIRKTYQETYQEDLIKRLESEISG----------DFERAVYRWMLEPADRDAVL
PsAnn3   (1)   --MASLKLPEVVFSPTQDSERLRNAFQGIGTDEKELILVLGHRNAQQRKEIRETYHQIYKESLVDSLQSELSG----------DFRNAIVLWTCDSSERDAKL
PsAnn4   (1)   --MATIVAPSQIS-PVEDAEALRLAFKGWGADNKAIIAILGHRNVHQRQQIRKAYDELFEEDLLKRLESEISG----------DFERAVYRWILEPADRDAVL
PsAnn5   (1)   --MATLILPPIPPSPRDDAMQLYRAFKGFGCDTTAVINILAHRDAAQRAYLQHEYKATYSEELSKRLVSELSG----------KLETAVLLWLHDPAGRDAEI
PsAnn6   (1)   --MASLMAPSKHS-PVEDAEALQRAVKGWGADNKAIIAILGHRNGTQRTQIRQAYYDLYQEDLIKRLESELSG----------DFERAMYRWILEPAEREALL
PsAnn7   (1)   --MSTLVLPPVPPSPKDDATELHRAFKGFGCDTGAVINILAHRDAAQRAYLQQEYRAIYYEDLIKRLSSELSG----------KLENAVLLWMHDPAGRDAVI
PsAnn8   (1)   --MSTLTVPHPLPPVSDDVEQLRKAFFGWGTNEDLIISILGHRNGSQRKAIREGYAQTYGEDLLKALDKELTS----------DFERLVHLWALESAERDAFL
PsAnn9   (1)   --MASLIDAKDFS-PIEDAETIMNACKGFGTNETVLISILANRNAAQRKLVRLAYQEVYHQDLIELLKSELSG----------NFERAICLWTMDPAEREATL
PsAnn10  (1)   --MTTLTIPPTIFSPREDAKKLHKAFKGLGCDTTRVIQILANRNSEQRSLIQGEYETTYSEPLSKRLSSEIRG----------HLKKALLLWLHDPPTRDAKI
PsAnn11  (1)   -----------MAFNQELEAITQAFSGHGVDEKSLATIILGKWDPVERETYRKNTSNFFVEDHERKFQKWNDHHVRLLKHEFVRFKNAVVLWSMHPWERDARL
PsAnn12  (1)   --MATLKIPSHVPHPSEDSEQLRGAFQGWGTNEGLIISILAHRNASQRKAIRETYAQTHGEDLLKDLDKELSS----------DFEKAVLLWTLDPAERDAFL
PsAnn13  (1)   MATTNILTMMTNNNFELDCKKIHDSLSGL--SLSQLIPSLSSLTILERKQLRETYKEVYGEDLINQLQKYDEDDF-----FSSMKFSTLSLWMLCPHDRDAFI
PsAnn14  (1)   --MATLVAPEHFT-PEDDAIVLYKAVKGWGTDESAIIAIMGHRNATQRQQIRQIYEDMYEEDLIKRLESELSG----------NFEKAMYRWILEPSDCYAVL
                                        Repeat I
```

```
           104                                                                                                    206
MtAnn1   (91)  ANVAIKSG-KN------------------------------------------YNVIVEISAVLSPEELLNVRRAYVKRYKHSLEEDLAAHTS-GHLRQLLVGL
MtAnn2   (91)  ANVAIKDGSKS------------------------------------------YHVIIEIVSVLSPEEVLAMRRAYHNRYKHSLEEDLAAHTT-GHLRQLLVGL
PvAnn1   (91)  ANVAIKSG-KN------------------------------------------YNVIVEIATINSPEELLAVRRAYLHRYKHSLEEDLAAHTT-GPLRQLLVGL
PvAnn2   (91)  ANVAIKSGSKG------------------------------------------YHVIVEIACVLSSEELLAARRAYHNRYKRSLEEDVATHTS-GDLRQLLVGL
PsAnn1a  (91)  ANVALKSG-KN------------------------------------------YNVVVEIASILSPEELFNVRRAYIKRYKHSLEEDLATHTS-GHLRQLLVGL
PsAnn1b  (91)  AHVALKSE-KN------------------------------------------YNVVVEIASILSPEELFNVRRAYIKRYKHSLEEDLATHTS-GHLRQLLVGL
PsAnn2   (91)  ANVALKDGSKG------------------------------------------YPVIIEIVSVLSPEEVLALRRAYHNRYKHSLEEDLAAHTT-GHLRQLLVGL
PsAnn3   (92)  ARDALKAKRKG------------------------------------------IKQLQILVEIACATSPNHLMAVRQTYCALYDCSLEEDIIASVS-PPLTKILVGL
PsAnn4   (91)  INVAIKHGGKD------------------------------------------YHVVAEIASVLSAEELLAVRRAYRNRYKRSIEEDVSANTT-GHLRQLLVGL
PsAnn5   (92)  IRQSLIVD---------------------------------------------RSLVAATEVICSRTPSQLQHLKQLYHSKFGVYLEHEIEANTS-GDLQKILLKY
PsAnn6   (91)  ANIALKNANMN------------------------------------------YHVIVEISCVSSPDELFIVRRAYHNRYKRSMEEDVATNTT-GHLRQLLVGL
PsAnn7   (92)  IRQSLIVT---------------------------------------------KILDAATEVICSRTPAQLQYLRQIYHAKFGDFVDHDIERNTS-GDHKKILLAY
PsAnn8   (92)  ANEATKKWTSS------------------------------------------NQVLVAEIACTRSSDQLFSAKKAYHVLYKKSLEEDVAHHTT-GDYRKLLLPL
PsAnn9   (91)  INEALKKATPD------------------------------------------YKVIIEIACTKNSEELLAVKRSYQSLYKHSLEEDVASQTI-GDIRKLLVAV
PsAnn10  (92)  TRSALTSSV---------------------------------------------VDNQATTEIICSRTPSQLRRLKEVYLSKYNHSPLERDIENQTS-GDHKKLLLGY
PsAnn11  (92)  AKEALKKGPSS------------------------------------------YNVLIEIACTRSSEELLGARKAYHSLFDHSIEEDVASHIH-GIDRKLLVAL
PsAnn12  (92)  ANQATKMLTSNNAIIMEIASTRSPLELLKETLDPAERDAFLANQATKMLTSNNAIIMEIASTRSPLELLKAKSAYQSRFKKSLEEDVAYHTS-GDIRKLLVPL
PsAnn13  (97)  AREALLQDETN------------------------------------------FKALVEVFVGRKSSHVGLITQAYLKMFRRQLDQDIMNLDPPHPFQKILMAL
PsAnn14  (91)  ANVALKNVSRD------------------------------------------YHVIVEIASVLKPEELVNVRRAYQYRFHHSLEEHVAAHTS-SGGYFRQFLVGL
                                        Repeat II
```

```
           207                                                                                                    309
MtAnn1   (151) VTAFRYVGDEINPKLAQTEAGILHESVK---EKKGSHEE-AIRILTTRSKTQLIATFNRYRETHGTSITKKLLDEGSDEFQKALYTTIRSFNDHVKYYEKVVR
MtAnn2   (152) VTSFRYGGAEINPKLAKTEADILHESIK--EKKGNHEE-AIRILTTRSKTQLLATFNRYRDDHGISITKKLLDNASDDFHKALHTTIRCINDHKKYYEKILR
PvAnn1   (151) VTSFRYVGDEINPKLAQNEAEILHDAVK---EKKSSYEE-AIRVLTTRSKTQLVATFNRYREIHGGSISKKLVDEGSDDFEKALHTTIRSINDHFKYYEKVVR
PvAnn2   (152) VTSYRYGGDEINARLAKTEADILHSSIK---EKKGNHEE-AIRILTTRSKTQLLATFNRYRDDHGTSITKKLLDDASDDFRKALHTAIRCINDHKKYYEKVLR
PsAnn1a  (151) VTAFRYVGDDVNARLAQSEADILHEAVK---EKKGSQEE-AVRILTTRSKAQLIATFNRYRETHGTSITKKLLEEGSDDFQKALYTTIRSFNDHVKYYEKVVR
PsAnn1b  (151) VTAFRYVGDDVNARLAQSEAHIIHESLK---EKKGSQEE-AVRILTTRSKAQLIATFNRYRDEHGVSVSKKLLDQTSDDFHKALHTAIRSINDHKKYYEKVLR
PsAnn2   (152) VSSFRYGGDEINPKLAKTEADILHESVK---EKKGNYEE-AIRILTTRSKTQLLATFNRYRDDHGTSITKKLLDNASDDFHKALHTAIRCISDHKKYYEKVLR
PsAnn3   (156) VSSYRHDKVTVNSEVAKSEAEKLHEAIK---NKQLEDDH-VVWILCTRNFFQLRETFTCYKQLYNNTFEEDIKVCGKGDLASLLNVVVWCIDRPEKHFAKVIR
PsAnn4   (152) VSSFRYEGDDEINPKLAQSEAHIIHESLK---EKKGNNEEEVIVITTRSKTQLVATFNRYRDEHGVSVSKKLLDQTSDDFHKALHTAIRSINDHKKYYEKVLR
PsAnn5   (152) ISTPRHEGPEVNREVAQKDAKVLFKAGE---KKLGTDEKTFVQIFSERSSVHLAAVSSYYHDMYGHSLKKAVKNEASGNFGLALLTITECANNPAKYFAKVLY
PsAnn6   (152) VSSFRYGGSEVNARLAESEADILHEAIK---NKNHNHEE-VIRILTTRSKTQLVATFNCYRHDHGIAITKKLLDEGSDDFHKALRIAISCINDHKKYYEKVLR
PsAnn7   (153) LNTPRHEGPETNRDMAENDAVALYRAGE---KKLGTDEKIFVQIFSGHSAAQLAAINQCYNTKYGHSLKKAIKNETSGDFFAHALSTIVQCAENPARYFAKVLR
PsAnn8   (153) VSCHRYEGDEVNLTIAKAEAKILHEKIS---KKAYNDDD-LIRILATRSKAQVNATLNHYKDAFGKDINKDLKEDPKDEFLSLLRSTVKCLTRPEKYFAKIIR
PsAnn9   (152) ISTYKYDGEEFDENVAHSEANILHQLIE---RKAFNDDE-MIRILSTRSKKQLSVTFNIFKDLFGTTINKGLLASTTDEYIGALRTIVRCINDPQRYLAKVLC
PsAnn10  (153) ITTPRYEGPEFDHLMVEEDAKQLYKSGE---KKIGTDEKTFIRIFTERSTTHLAAVSSAYTASFGNSLDKAIKSETSGSFMRGLLTILRCATDSNMYFAKLLR
PsAnn11  (153) VSAYRYEGSKVKDDTAKSEAKTLANAIKNDQKKPIIEDDEVIRILATRSKLHLQAISKHYKEISGKNLEEDLNDLR----FKET---VQCLCTPQVYFSKVLD
PsAnn12  (194) VGTFRYEGDEVNMTLAKSEAKLLHDKIA---DKAYNHED-LIRIVTTRSKAQLNATLNHYNNAFGNVIDKDLETGSEDEYLKLLKAAIKCLTYPEKYFEELLR
PsAnn13  (159) SASHKAHQVDISHHISKCDARRLYESGER---SLGAIDEAVVLEILSKRSIQQLKLTFLSYKHIYGHDYTKSIKRGNYGQFGKALMVVVKCICNQAHYYAKGLY
PsAnn14  (153) VSSFRYDGDYIHPSLAKYEAEILHEAIQ---NKNGNLDE-VIRILTTRSKKQLKATFNRYRDDHGYSISKRLLNEESDDFLKAAHVAIRCIDDHKKYYEKVLR
                                        Repeat III
```

```
           310                                                                  384
MtAnn1   (250) DAIKKVGTD--EDALTRVIVSRAQHDLKVISDVYYKRNSVL-LEHVVAKETSGDYKKFLLTLLGKEE--------
MtAnn2   (251) GALKRVGTD--EDALTRVVVTRAEKDLKDIKELYYKRNSVH-LEDAVAKEISGDYKKFILTLLGKQD--------
PvAnn1   (250) NAIKKVGTD--EDALTRVVVSRAEKDLKTVSEVYYKRNSVL-LEHAIAKEISGDYKKFLLTLLGKED--------
PvAnn2   (251) NAMKKVGTD--EDALTRVVVSRAEKDLRNIAEIYYKRNSVH-LEDAVAKEISGDYKKFILTLLGKDV--------
PsAnn1a  (250) DAVKKVGTD--EDALTRVIVSRAQHDLKAISDVYHKRNSVP-FEHVVAKETSGDYKKLLLTLLGQED--------
PsAnn1b  (250) DAVKRIGTD--EDALTRVIVSRAQHDLKAISDVYHKRNSVP-LEHVVAKETSGDYKKLLLTLLGQED--------
PsAnn2   (251) SALKRIGTD--EDGLTRVVVTRAEKDLKDIKEVYYKRNSVH-LEDAIAKEISGDYKKFILTLLGK----------
PsAnn3   (255) DSIVGPGTD--EDSLNRGIVTRAEIDLLKVRFEYANMFKTN-LDDDVIGDTSGDYKEFLLTLLGKGPKGD------
PsAnn4   (252) NAIKKGTD--EDALSRVIVTRAEKDLKAISDVYKKFILALLGKQE--------
PsAnn5   (252) KAMKGMGTN--DDTLIRVIVSRTEIDMQYIKAEYSKKYKKT-LNDAVHSETSGNYRAFLLALLGPNH--------
PsAnn6   (251) NAMETVGTD--EDALTRVIVTRAEKDLEDIKKVYYKRNSVH-LEHAVAKKTSGDYKKFLLTLMAKEE--------
PsAnn7   (252) KSMKGLGTD--DTKLIRVIVTRCEIDLHYIKAELYKKYKKT-LNDAVKSETSGGYKFTLLLGSNQ--------
PsAnn8   (252) LSIMKRGTD--EGALTRVVATRAEKDLKIVGDEYQRRDSVP-LDRAIVKDTTGDYEKMLLAILGHDDA-------
PsAnn9   (251) NALMDLVNE--DDALNRVIVTRAEKDLKEIKDHFLKRNNTS-IHDSVRTKTWGNYKTFLLHLLGKE---------
PsAnn10  (253) KSMKGIGTD--DKRLIRVIVTRSTEIDMVHIKAEYRYSDTSGHYKDFIIHLLGTDY-------
PsAnn11  (249) ASLKIDVDKNVKKSLTRVIITRADVDMKDIKTEYQKLYGVS-LPQKIEETAKGNYKDFLLTLIARGG--------
PsAnn12  (293) LAINKMGTD--ELALTRVVTTRAEVDLQRIAEEYQRRNSVP-LDRAIDNDTSGDYQKILLALLGRED--------
PsAnn13  (260) TSIKKG---RRDIGTLARVLVSRADIDMDEIRSVFKEKYEKELGDVICESIPCGDYRDFLVALATKS--SYIFTN
PsAnn14  (252) GALKRIGTD--EDELTRVVVTRAEKDLKDLKELYYHRNSVR-LEDAVASEISGDYKRFLVTLIGREY--------
                                        Repeat IV
```

Fig. 2. Multiple sequence alignment of the amino acid sequences of 15 presumable *P. sativum* annexins, 2 *M. truncatula* annexins (MtAnn1, MtAnn2), and 2 *P. vulgaris* annexins (PvAnn1, PvAnn2) by ClustalΩ.

Four annexin repeats are underlined. Yellow highlights potential calcium-binding motifs. In the calcium-binding motif of the first annexin repeat, the conservative tryptophan (W) necessary for binding to the membrane is indicated in gray. Important for calcium binding amino acid residues in the calcium-binding site of the type II (repeat I, Phe-23, Gly-25, Gly-27, and Glu-67) as well as in the calcium-binding site of the type III (repeat IV, Ile-254, Lys-256, Gly-258, and Val-296, Thr-299, Glu-304) are indicated in bold and underlined. *P. sativum* annexins PsAnn4 and PsAnn8 are marked in red.

О.А. Павлова, И.В. Леппянен
Д.В. Кустова, А.Д. Бовин, Е.А. Долгих

Филогенетический и структурный анализ
аннексинов у гороха

2021
25•5
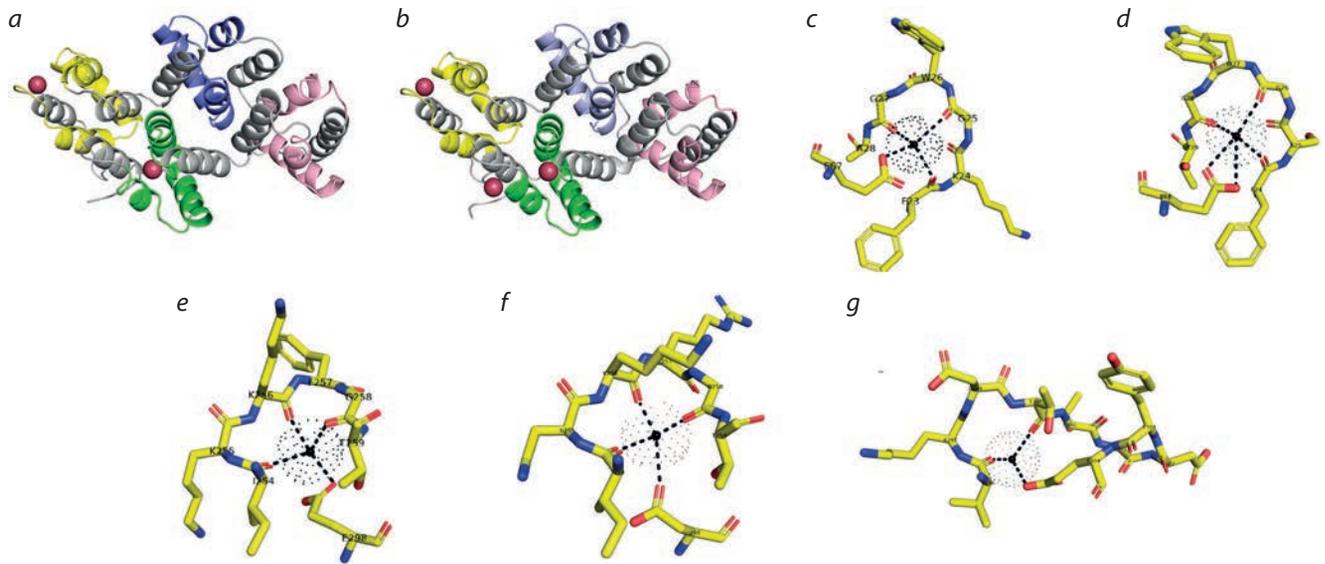
Fig. 3. Modeling of the three-dimensional structures of PsAnn4 (*a*) and PsAnn8 (*b*) using the crystal structure of *G. hirsutum* annexin (GhAnn1, PDB code 3BRX) as a template and their binding with calcium ions in the first (*c*, *d*) and fourth repeats (*e*, *f*, *g*).

The 3D structures of PsAnn4 and PsAnn8 proteins indicated the coordination of calcium ions in the first (*c*, *d*) and fourth (*e*, *f*, *g*) annexin repeats.
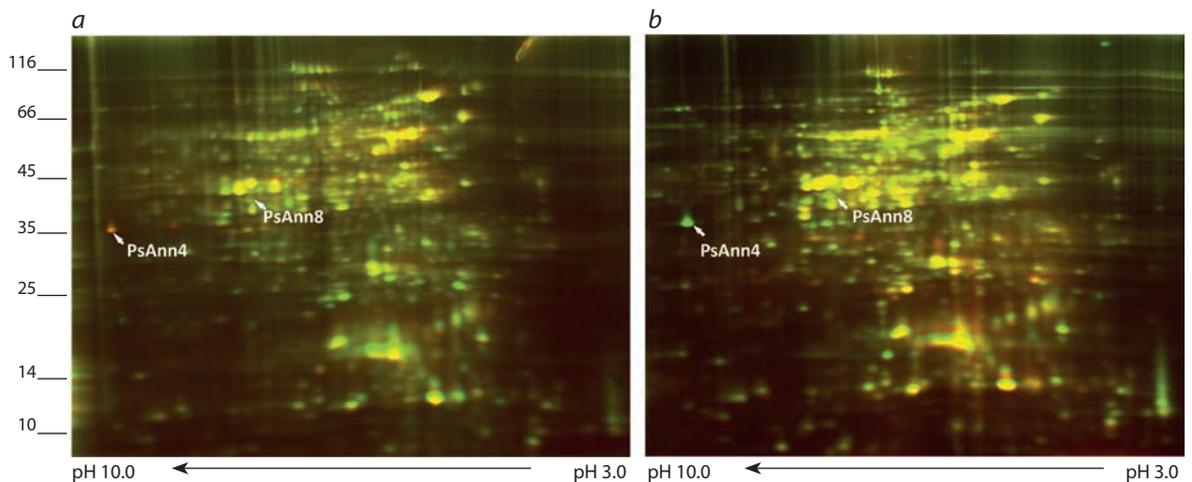
Fig. 4. Comparative analysis of protein patterns in wild-type pea plant and P56 mutant with an impaired *sym10* gene using two-dimensional differential gel electrophoresis 1 day after inoculation (1 dai).

The protein extract from wild type pea roots inoculated with *R. leguminosarum* bv. *viciae* RCAM1026 was labelled with Cy2 (red) and protein extract from inoculated roots of P56 mutant was labelled with Cy5 (green) (*a*) and conversely the extract from inoculated wild type roots was labelled with Cy5 (green) and protein extract from inoculated roots of P56 mutant was labelled with Cy2 (red) (*b*).
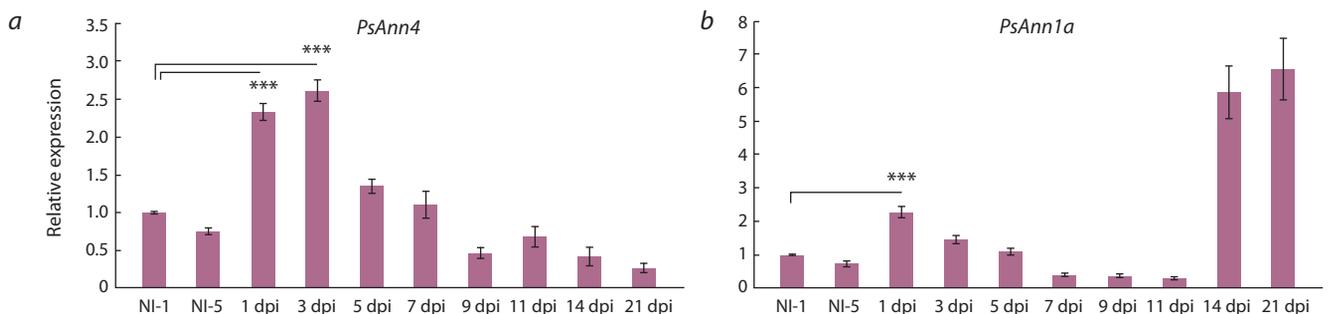
Fig. 5. Quantitative RT-PCR analysis of *PsAnn4* (*a*) and *PsAnn1b* (*b*) expression in pea roots upon nodulation. mRNA levels were normalized against *Ubiquitin* and values were calculated as ratios relative to non-inoculated root (NI) expression levels.

The data of three independent biological experiments were analyzed. Bars represent the mean ± SEM of two biological replicates. Asterisks indicate significant differences compared to non-inoculated roots, based on Student's *t*-test and *p*-value less than 0.001 is flagged with three asterisks (***).
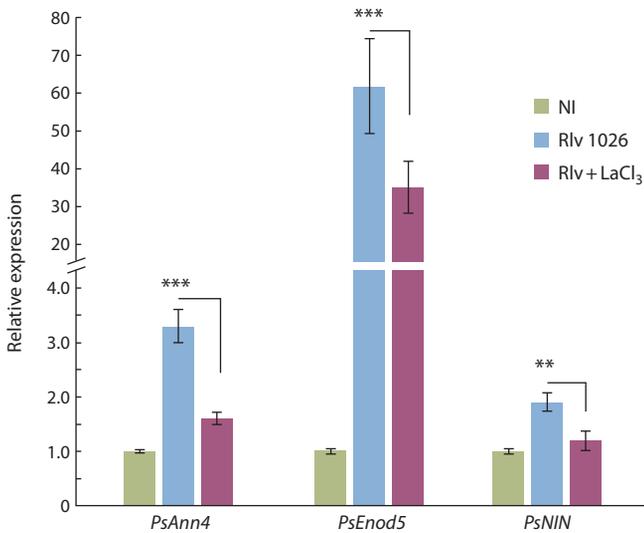
Fig. 6. *PsAnn4*, *PsEnod5*, and *PsNIN* expression levels in pea roots after inoculation (1 dai) with *R. leguminosarum* bv. *viciae* RCAM1026 (Rlv) and after treatment with the Ca^{2+} channel blocker $LaCl_3$ (Rlv + $LaCl_3$). mRNA levels were normalized against *Ubiquitin* and values were calculated as ratios relative to non-inoculated root expression levels (NI).

The data of three independent biological experiments were analyzed. Bars represent the mean ± SEM. Asterisks indicate significant differences between treated (Rlv + $LaCl_3$) and non-treated (Rlv) roots, based on Student's *t*-test and *p*-values less than 0.001 and 0.01 are flagged with three (***) and two (**) asterisks, respectively.

Fig. 7. Localization of PsAnn4 fused to red fluorescent protein (RFP) (*a*) and yellow fluorescent protein (YFP) (*b*) at the C-terminus under the transcriptional regulation of the 35S promoter in *N. benthamiana* leaves by infiltration with *A. tumefaciens* LBA4404. Scale bars are 200 µm. Immunoblot analysis of different cell fractions obtained from the *N. benthamiana* leaves after infiltration of PsAnn4-YFP with *A. tumefaciens* LBA4404 (*c*).

IS-1 – insoluble fraction was pelleted at 36 000 g; IS-2 – insoluble fraction was pelleted at 100 000 g; S – soluble fraction at 100 000 g; MW – molecular weight marker.

plant treatment with the Ca^{2+} channel blocker $LaCl_3$ (Fig. 6). Two previously described as symbiosis-specific genes *PsNIN* and *PsEnod5* were also used in our experiments as a control for effective inoculation. In pea roots, the upregulation of *PsAnn4* expression in response to inoculation was revealed in 1 dai, corresponding with experiments on the dynamics of this gene expression upon nodulation. The significant decrease in the expression of *PsAnn4* was found in our experiments in the presence of $LaCl_3$. Down-regulation of symbiosis-specific genes *PsEnod5* and *PsNIN* was also observed, which indicated the importance of Ca^{2+} influx for their regulation. Therefore, the influx of calcium ions into the cell, which is observed at the early stages of symbiosis development, may affect the expression level of *PsAnn4* in pea roots (see Fig. 6).

Subcellular localization of pea PsAnn4 annexin

To follow the PsAnn4 protein localization in plant cells, it was fused to the fluorophores such as red fluorescent protein (RFP) and yellow fluorescent protein (YFP) at the C-terminus and expressed under the transcriptional regulation of the 35S promoter in *N. benthamiana* leaves by infiltration with *A. tumefaciens* (Fig. 7, *a*, *b*). The infiltration of constructs for the synthesis of proteins fused with RFP and YFP allowed us to visualize the protein in leaf tissues after transformation. In the cells of *N. benthamiana* leaves, PsAnn4 protein was localized in the plasma membrane or in the cell wall. In addition, we also estimated the presence of PsAnn4 in different cell fractions by Western-blot hybridization using anti-YFP or anti-RFP antibodies. PsAnn4-YFP was found in insoluble fraction of leaf tissue pelleted at 36 000 g (see Fig. 7, *c*). It suggests that PsAnn4 may be involved in cell wall or membrane modification as well as in ion transport.

Discussion

Available pea genome information (Kreplak et al., 2019) allowed us to determine the composition of the annexin gene family in this legume. Database searches revealed 15 annexin genes in *P. sativum* L., 18 in *M. truncatula* as well as 13 in both *P. vulgaris* and *L. japonicus*. Based on the phylogenetic analysis of these annexins, close homologs can be identified among these legume species (see Fig. 1).

At present, only one pea annexin, p35, has been functionally characterized (Clark et al., 1992). The localization of this annexin in root cells involved in active secretion suggests its function in exocytosis. Subsequently, the use of antibodies against this protein revealed its localization in epidermal cells of the leaf and stem (Clark et al., 1998, 2000). However, annexins involved in nodulation have not been characterized in *P. sativum*. In contrast, in *M. truncatula*, two annexins, MtAnn1 (Medtr8g038210) and MtAnn2 (Medtr8g038220),

О.А. Павлова, И.В. Леппянен
Д.В. Кустова, А.Д. Бовин, Е.А. Долгих

Филогенетический и структурный анализ
аннексинов у гороха

2021
25•5

demonstrated a high level of expression during nodulation and were found to be involved in controlling bacterial infection and nodule organogenesis (De Carvalho-Niebel et al., 1998, 2002; Manthey et al., 2004; Breakspear et al., 2014). Another annexin, MtAnn3 (Medtr4g097180), was found to be important for root hair deformations in *M. truncatula* (Gong et al., 2012). At the same time, close homologs of MtAnn1 – PvAnn1 (Phvul.011g209300) and LjAnn1 (Lj0g3v0203419), which belong to the same phylogenetic group as MtAnn1, play important roles in the symbiotic process in *P. vulgaris* and *L. japonicus* (Wienkoop, Saalbach, 2003; Jáuregui-Zúñiga et al., 2016; Carrasco-Castilla et al., 2018).

In our earlier work, two annexins activated at the early stages of symbiosis development in pea were found using the proteomics approach (Leppyanen et al., 2018). This approach might be helpful for the identification of new regulators of signal transduction pathways at the initial stages of nodulation in pea. Our present analysis revealed that these two identified annexins of pea belong to different phylogenetic groups, defined as homologs of MtAnn4, PvAnn4 and MtAnn8, PvAnn8, respectively. Although PsAnn4, and MtAnn4 and PvAnn4 have high levels of homology with MtAnn1 and PvAnn1, they belong to another group of annexins based on phylogenetic analysis. PsAnn8 belongs to a less studied phylogenetic group. Therefore, two previously unknown annexins were identified in our study. In addition to stimulation during rhizobial inoculation, the dependence of PsAnn4 and PsAnn8 activation on the LysM-receptor-like kinase SYM10, encoding a putative Nod factor receptor, was revealed in the present study (see Fig. 4), which suggested that rhizobial signaling molecules Nod factors may be important for their activation. It also suggests the participation of these two annexins in the development of the symbiotic interaction of plants with rhizobia.

Phylogenetic analysis and prediction of the overall 3D structure of PsAnn4 and PsAnn8 proteins showed differences in the Ca^{2+}-binding motif in the fourth annexin repeat of these proteins, and therefore, in the potential ability to bind calcium ions. This can potentially influence the binding of these annexins to phospholipids by means of a calcium bridge mechanism. It was predicted that three calcium ions were coordinated in the first and fourth repeats, which is consistent with the data of the canonical binding of the *G. hirsutum* annexin GhAnn1 and animal annexins to the phospholipids of membranes using the mechanism of calcium bridges (Hu et al., 2008). In the predicted structures of *Arabidopsis* annexins (AtAnn1, AtAnn3, and AtAnn4), the canonicity of the Ca^{2+}-binding motif in the first repeat and the presence of modified motifs in the fourth repeats of AtAnn1 and AtAnn3 were also shown, while AtAnn4 had no recognizable Ca^{2+} – or phospholipid-binding motifs (Konopka-Postupolska, Clark, 2017).

Since the level of PsAnn4 synthesis in response to inoculation was more significant in the roots of wild type pea plants compared with mutant defective in symbiosis, we carried out the analysis of this annexin in more detail. It was shown that the regulation of PsAnn4 annexin in pea could be achieved at the transcriptional level as well as post-transcriptional and translational levels, probably. Significant activation of *MtAnn1* and *MtAnn2* gene expression level was found in the roots of *M. truncatula* treated with Nod factors or inoculated

with rhizobia (De Carvalho-Niebel et al., 1998, 2002; Manthey et al., 2004; Breakspear et al., 2014). Meanwhile, the expression of *PvAnn1* in *P. vulgaris* was slightly upregulated in developing nodules (Carrasco-Castilla et al., 2018). However, a phosphoproteomic approach revealed that PvAnn1 was a phosphorylated protein with enhanced levels of synthesis during nodulation (Jáuregui-Zúñiga et al., 2016). Hence, the regulation of annexins involved in nodulation might be different and is probably connected with different functions that annexins fulfil in this process.

Localization of annexins might differ depending on their function. Some annexins show cytoplasmic and nuclear localization, while other annexins are associated with various plant membranes, including the plasma membrane, endoplasmic reticulum, and nuclear membrane (Laohavisit, Davies, 2011; Clark et al., 2012; Davies, 2014). Some annexins may be embedded in the membrane in the form of monomers or oligomers. One of the distinctive characteristics of annexins is their ability to change their cellular localization in response to various stimuli. In our experiments, the localization of annexin 4 (PsAnn4) in the cell wall or plasma membrane was shown, suggesting the participation of this annexin in processes associated either with membrane modification or ion transport at the early stages of symbiosis establishment in pea. Similarly, the localization of the other annexin, MtAnn2, involved in nodulation in *M. truncatula*, was revealed to be associated with the plasma membrane, particularly with lipid rafts from root plasma membrane preparations (Lefebvre et al., 2007). In addition, the annexin PvAnn1 is essential for ROS-dependent regulation of Ca^{2+} influx into the cells of *P. vulgaris*, which strongly suggests the localization of this protein in the plasma membrane. Therefore, specific subcellular localization of annexins might be associated with their function signal transduction at the early stages of symbiosis.

Conclusion

In this study, phylogenetic analysis of the pea annexins PsAnn4 and PsAnn8 was performed based on their homology with annexins from other legumes. The modeling approach allowed us to estimate the structural features of these annexins that might influence their functional activity. To verify the functions of these annexins, we performed comparative proteomic analysis, experiments with calcium influx inhibitors, and localization of labeled proteins. Essential down-regulation of PsAnn4 synthesis in a non-nodulating pea mutant P56 (sym10) suggests an involvement of this annexin in the rhizobial symbiosis. The localization of PsAnn4 in the cell wall or plasma membrane of plant cells may indicate its participation in membrane modification or ion transport.

References

Bateman A., Coin L., Durbin R., Finn R.D., Hollich V., Griffiths-Jones S., Khanna A., Marshall M., Moxon S., Sonnhammer E.L.L., Studholme D.J., Yeats C., Eddy S.R. The Pfam protein families database. *Nucleic Acids Res.* 2004;32:D138-D141. DOI 10.1093/nar/gkh121.

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.

Breakspear A., Liu C., Roy S., Stacey N., Rogers C., Trick M., Morieri G., Mysore K.S., Wen J., Oldroyd G.E.D., Downie J.A., Murray J.D. The root hair "Infectome" of *Medicago truncatula* uncovers changes in cell cycle genes and reveals a requirement for auxin signaling in rhizobial infection. *Plant Cell Online.* 2014;26(12):4680-4701. DOI 10.1105/tpc.114.133496.

Breton G., Vazquez-Tello A., Danyluk J., Sarhan F. Two novel intrinsic annexins accumulate in wheat membranes in response to low temperature. *Plant Cell Physiol.* 2000;41(2):177-184. DOI 10.1093/pcp/41.2.177.

Carrasco-Castilla J., Ortega-Ortega Y., Jáuregui-Zúñiga D., Juárez-Verdayes M.A., Arthikala M.K., Monroy-Morales E., Nava N., Santana O., Sánchez-López R., Quinto C. Down-regulation of a *Phaseolus vulgaris* annexin impairs rhizobial infection and nodulation. *Environ. Exp. Bot.* 2018;153:108-119. DOI 10.1016/j.envexpbot.2018.05.016.

Carroll A.D., Moyen C., Van Kesteren P., Tooke F., Battey N.H., Brownlee C. Ca²⁺, annexins, and GTP modulate exocytosis from maize root cap protoplasts. *Plant Cell.* 1998;10(8):1267-1276. DOI 10.1105/tpc.10.8.1267.

Clark G.B., Dauwalder M., Roux S.J. Purification and immunolocalization of an annexin-like protein in pea seedlings. *Planta.* 1992; 187(1):1-9. DOI 10.1007/BF00201617.

Clark G.B., Dauwalder M., Roux S.J. Immunological and biochemical evidence for nuclear localization of annexin in peas. *Plant Physiol. Biochem.* 1998;36(9):621-627. DOI 10.1016/S0981-9428(98)80010-7.

Clark G.B., Morgan R.O., Fernandez M.P., Roux S.J. Evolutionary adaptation of plant annexins has diversified their molecular structures, interactions and functional roles. *New Phytol.* 2012;196(3):695-712. DOI 10.1111/j.1469-8137.2012.04308.x.

Clark G.B., Rafati D.S., Bolton R.J., Dauwalder M., Roux S.J. Redistribution of annexin in gravistimulated pea plumules. *Plant Physiol. Biochem.* 2000;38(12):937-947. DOI 10.1016/S0981-9428(00)01206-7.

Clark G.B., Sessions A., Eastburn D.J., Roux S.J. Differential expression of members of the annexin multigene family in *Arabidopsis. Plant Physiol.* 2001;126(3):1072-1084. DOI 10.1104/pp.126.3.1072.

Dam S., Dyrlund T.F., Ussatjuk A., Jochimsen B., Nielsen K., Goffard N., Ventosa M., Lorentzen A., Gupta V., Andersen S.U., Enghild J.J., Ronson C.W., Roepstorff P., Stougaard J. Proteome reference maps of the *Lotus japonicus* nodule and root. *Proteomics.* 2014;14(2-3):230-240. DOI 10.1002/pmic.201300353.

Davies J.M. Annexin-mediated calcium signalling in plants. *Plants.* 2014;3(1):128-140. DOI 10.3390/plants3010128.

De Carvalho Niebel F., Lescure N., Cullimore J.V., Gamas P. The *Medicago truncatula MtAnn1* gene encoding an annexin is induced by Nod factors and during the symbiotic interaction with *Rhizobium meliloti. Mol. Plant Microbe Interact.* 1998;11(6):504-513. DOI 10.1094/MPMI.1998.11.6.504.

De Carvalho-Niebel F., Timmers A.C.J., Chabaud M., Defaux-Petras A., Barker D.G. The Nod factor-elicited annexin MtAnn1 is preferentially localised at the nuclear periphery in symbiotically activated root tissues of *Medicago truncatula. Plant J.* 2002;32(3):343-352. DOI 10.1046/j.1365-313X.2002.01429.x.

Espinoza C., Liang Y., Stacey G. Chitin receptor CERK1 links salt stress and chitin-triggered innate immunity in *Arabidopsis. Plant J.* 2017;89(5):984-995. DOI 10.1111/tpj.13437.

Feng Y.M., Wei X.K., Liao W.X., Huang L.H., Zhang H., Liang S.C., Peng H. Molecular analysis of the annexin gene family in soybean. *Biol. Plant.* 2013;57(4):655-662. DOI 10.1007/s10535-013-0334-0.

Gerke V., Moss S.E. Annexins: from structure to function. *Physiol. Rev.* 2002;82(2):331-371. DOI 10.1152/physrev.00030.2001.

Gong Z.Y., Song X., Chen G.Y., Zhu J.B., Yu G.Q., Zou H.S. Molecular studies of the *Medicago truncatula MtAnn3* gene involved in root hair deformation. *Chinese Sci. Bull.* 2012;57(15):1803-1809. DOI 10.1007/s11434-011-4937-6.

Gorecka K.M., Konopka-Postupolska D., Hennig J., Buchet R., Pikula S. Peroxidase activity of annexin 1 from *Arabidopsis thaliana. Biochem. Biophys. Res. Commun.* 2005;336(3):868-875. DOI 10.1016/j.bbrc.2005.08.181.

Hofmann A., Proust J., Dorowski A., Schantz R., Huber R. Annexin 24 from *Capsicum annuum.* X-ray structure and biochemical characterization. *J. Biol. Chem.* 2000;275(11):8072-8082. DOI 10.1074/jbc.275.11.8072.

Hu N.J., Yusof A.M., Winter A., Osman A., Reeve A.K., Hofmann A. The crystal structure of calcium-bound annexin Gh1 from *Gossypium hirsutum* and its implications for membrane binding mechanisms of plant annexins. *J. Biol. Chem.* 2008;283(26):18314-18322. DOI 10.1074/jbc.M801051200.

Ijaz R., Ejaz J., Gao S., Liu T., Imtiaz M., Ye Z., Wang T. Overexpression of annexin gene *AnnSp2*, enhances drought and salt tolerance through modulation of ABA synthesis and scavenging ROS in tomato. *Sci. Rep.* 2017;7(1):1-14. DOI 10.1038/s41598-017-11168-2.

Jáuregui-Zúñiga D., Ortega-Ortega Y., Pedraza-Escalona M., Reyes-Grajeda J.P., Ruiz M.I., Quinto C. Phosphoproteomic analysis in *Phaseolus vulgaris* roots treated with *Rhizobium* etli nodulation factors. *Plant Mol. Biol. Report.* 2016;34(5):961-969. DOI 10.1007/s11105-016-0978-y.

Kirienko A.N., Porozov Y.B., Malkov N.V., Akhtemova G.A., Le Signor C., Thompson R., Saffray C., Dalmais M., Bendahmane A., Tikhonovich I.A., Dolgikh E.A. Role of a receptor-like kinase K1 in pea *Rhizobium* symbiosis development. *Planta.* 2018;248(5):1101-1120. DOI 10.1007/s00425-018-2944-4.

Kodavali P.K., Skowronek K., Koszela-Piotrowska I., Strzelecka-Kiliszek A., Pawlowski K., Pikula S. Structural and functional characterization of annexin 1 from *Medicago truncatula. Plant Physiol. Biochem.* 2013;73:56-62. DOI 10.1016/j.plaphy.2013.08.010.

Konopka-Postupolska D., Clark G. Annexins as overlooked regulators of membrane trafficking in plant cells. *Int. J. Mol. Sci.* 2017;18(4):1-34. DOI 10.3390/ijms18040863.

Konopka-Postupolska D., Clark G., Goch G., Debski J., Floras K., Cantero A., Fijolek B., Roux S., Hennig J. The role of annexin 1 in drought stress in *Arabidopsis. Plant Physiol.* 2009;150(3):1394-1410. DOI 10.1104/pp.109.135228.

Konopka-Postupolska D., Clark G., Hofmann A. Structure, function and membrane interactions of plant annexins: An update. *Plant Sci.* 2011;181(3):230-241. DOI 10.1016/j.plantsci.2011.05.013.

Kreplak J., Madoui M.-A., Cápal P., Novák P., Labadie K., Aubert G., Bayer P.E., Gali K.K., Syme R.A., Main D., Klein A., Bérard A., Vrbová I., Fournier C., D'Agata L., Belser C., Berrabah W., Toegelová H., Milec Z., Vrána J., Lee H., Kougbeadjo A., Térézol M., Huneau C., Turo C.J., Mohellibi N., Neumann P., Falque M., Gallardo K., McGee R., Tar'an B., Bendahmane A., Aury J.-M., Batley J., Le Paslier M.-C., Ellis N., Warkentin T.D., Coyne C.J., Salse J., Edwards D., Lichtenzveig J., Macas J., Doležel J., Wincker P., Burstin J. A reference genome for pea provides insight into legume genome evolution. *Nat. Genet.* 2019;51(9):1411-1422. DOI 10.1038/s41588-019-0480-1.

Kwon Y.S., Lee D.Y., Rakwal R., Baek S.B., Lee J.H., Kwak Y.S., Seo J.S., Chung W.S., Bae D.W., Kim S.G. Proteomic analyses of the interaction between the plant-growth promoting rhizobacterium *Paenibacillus polymyxa* E681 and *Arabidopsis thaliana. Proteomics.* 2016;16(1):122-135. DOI 10.1002/pmic.201500196.

Laohavisit A., Davies J.M. Annexins. *New Phytol.* 2011;189(1):40-53. DOI 10.1111/j.1469-8137.2010.03533.x.

Laohavisit A., Mortimer J.C., Demidchik V., Coxon K.M., Stancombe M.A., Macpherson N., Brownlee C., Hofmann A., Webb A.A.R., Miedema H., Battey N.H., Davies J.M. *Zea mays* annexins modulate cytosolic free Ca²⁺ and generate a Ca²⁺-permeable conductance. *Plant Cell.* 2009;21(2):479-493. DOI 10.1105/tpc.108.059550.

Lefebvre B., Furt F., Hartmann M.-A., Michaelson L.V., Carde J.-P., Sargueil-Boiron F., Rossignol M., Napier J.A., Cullimore J., Bessoule J.-J., Mongrand S. Characterization of lipid rafts from *Medi-*

О.А. Павлова, И.В. Леппянен
Д.В. Кустова, А.Д. Бовин, Е.А. Долгих

Филогенетический и структурный анализ
аннексинов у гороха

2021
25·5

cago truncatula root plasma membranes: A proteomic study reveals the presence of a raft-associated redox system. *Plant Physiol.* 2007; 144(1):402-418. DOI 10.18362/bjta.v4.i1-2.59.

Leppyanen I.V., Kirienko A.N., Lobov A.A., Dolgikh E.A. Differential proteome analysis of pea roots at the early stages of symbiosis with nodule bacteria. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2018;22(2):196-204. DOI 10.18699/VJ18.34.7.

Limpens E., Moling S., Hooiveld G., Pereira P.A., Bisseling T., Becker J.D., Küster H. Cell- and tissue-specific transcriptome analyses of *Medicago truncatula* root nodules. *PLoS One.* 2013;8(5):e64377. DOI 10.1371/journal.pone.0064377.

Lizarbe M.A., Barrasa J.I., Olmo N., Gavilanes F., Turnay J. Annexin-phospholipid interactions. Functional implications. *Int. J. Mol. Sci.* 2013;14:2652-2683. DOI 10.3390/ijms14022652.

Madsen E.B., Madsen L.H., Radutoiu S., Olbryt M., Rakwalska M., Szczyglowski K., Sato S., Kaneko T., Tabata S., Sandal N., Stougaard J. A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature.* 2003;425(6958):637-640. DOI 10.1038/nature02045.

Manthey K., Krajinski F., Hohnjec N., Firnhaber C., Pünler A., Perlick A.M., Küster H. Transcriptome profiling in root nodules and arbuscular mycorrhiza identifies a collection of novel genes induced during *Medicago truncatula* root endosymbioses. *Mol. Plant-Microbe Interact.* 2004;17(10):1063-1077. DOI 10.1094/MPMI. 2004.17.10.1063.

Mortimer J.C., Laohavisit A., Macpherson N., Webb A., Brownlee C., Battey N.H., Davies J.M. Annexins: multifunctional components of growth and adaptation. *J. Exp. Bot.* 2008;59(3):533-544. DOI 10.1093/jxb/erm344.

Ordog R. PyDeT, a PyMOL plug-in for visualizing geometric concepts around proteins. *Bioinformation.* 2008;2(8):346-347. DOI 10.6026/97320630002346.

Orosz L., Sváb Z., Kondorosi A., Sik T. Genetic studies on rhizobiophage 16-3. I. Genes and functions on the chromosome. *Mol. Gen. Genet.* 1973;125(4):341-350. DOI 10.1007/BF00276589.

Sievers F., Wilm A., Dineen D., Gibson T.J., Karplus K., Li W., Lopez R., McWilliam H., Remmert M., Söding J., Thompson J.D., Higgins D.G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 2011; 7(1):539. DOI 10.1038/msb.2011.75.

Talukdar T., Gorecka K.M., de Carvalho-Niebel F., Downie J.A., Cullimore J., Pikula S. Annexins – calcium- and membrane-binding proteins in the plant kingdom: potential role in nodulation and mycorrhization in *Medicago truncatula*. *Acta Biochim. Pol.* 2009;56(2): 199-210. DOI 20091709.

Van Brussel A.A.N., Planque K., Quispel A. The wall of *Rhizobium leguminosarum* in bacteroid and free-living forms. *J. Gen. Microbiol.* 1977;101(1):51-56. DOI 10.1099/00221287-101-1-51.

Van Brussel A.A.N., Tak T., Wetselaar A., Pees E., Wijffelman C. Small leguminosae as test plants for nodulation of *Rhizobium leguminosarum* and other rhizobia and agrobacteria harbouring a *leguminosarum* sym plasmid. *Plant Sci. Lett.* 1982;27(3):317-325. DOI 10.1016/0304-4211(82)90134-1.

Voss T., Haberl P. Observations on the reproducibility and matching efficiency of two-dimensional electrophoresis gels: consequences for comprehensive data analysis. *Electrophoresis.* 2000;21(16): 3345-3350. DOI 10.1002/1522-2683(20001001)21:16<3345::AID-ELPS3345>3.0.CO;2-Z.

Webb B., Sali A. Comparative protein structure modeling using Modeller. *Curr. Protoc. Bioinform.* 2016;54(1):5.6.1-5.6.37. DOI 10.1002/cpbi.3.

Wienkoop S., Saalbach G. Proteome analysis. Novel proteins identified at the peribacteroid membrane from *Lotus japonicus* root nodules. *Plant Physiol.* 2003;131(3):1080-1090. DOI 10.1104/pp.102. 015362.

ORCID ID
O.A. Pavlova orcid.org/0000-0003-0528-5618
I.V. Leppyanen orcid.org/0000-0002-2158-0855
A.D. Bovin orcid.org/0000-0003-4061-435X
E.A. Dolgikh orcid.org/0000-0002-5375-0943

Biochemical composition of tomato fruits of various colors

A.B. Kurina ✉, A.E. Solovieva, I.A. Khrapalova, A.M. Artemyeva

Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), St. Petersburg, Russia
✉ nastya_n11@mail.ru

Abstract. Tomato (*Lycopersicon esculentum* Mill.) is an economically important and widely cultivated vegetable crop that is consumed both fresh and processed. The nutritional value of tomato fruits is related to the content of carotenoids, polyphenols, sugars, organic acids, minerals and vitamins. Currently, there is a growing interest in the qualitative and quantitative increase in the content of health-promoting compounds in tomato fruits. VIR *Lycopersicon* (Tourn.) Mill. genetic resources collection includes 7678 accessions of one cultivated and nine wild species, which in turn provides ample opportunities for searching for information on the variability of the content of biologically active substances and searching for sources with a high content of them in the gene pool. Our work presents the results of the study of 70 accessions of cultivated and wild tomato on the main biochemical characteristics: the content of dry matter, ascorbic acid, sugars, carotenoids, chlorophylls and anthocyanins. As the basis for the selection of accessions for the study, accessions with various colors of fruits, including new accessions with varying content of anthocyanin, were taken. As a result of this study, the amplitude of variability in the content of dry matter (3.72–8.88 and 9.62–11.33 %), sugars (1.50–5.65 and 2.20–2.70 %), ascorbic acid (12.40–35.56 and 23.62–28.14 mg/100 g), titratable acidity (0.14–0.46 and 0.33–0.48 %), chlorophylls (0.14–5.11 and 2.95–4.57 mg/100 g), carotenoids (0.97–99.86 and 1.03–10.06 mg/100 g) and anthocyanins (3.00–588.86 and 84.31–152.71 mg/100 g) in the fruits of cultivated and wild tomatoes, respectively, was determined. We have determined correlations between the content of dry matter and monosaccharides ($r = 0.40$, $p \le 0.05$), total sugars ($r = 0.37$, $p \le 0.05$) and ascorbic acid ($r = 0.32$, $p \le 0.05$); the content of ascorbic acid and carotenoids ($r = 0.25$, $p \le 0.05$). A high dependence of the content of chlorophyll *a* and *b* among themselves ($r = 0.89$, $p \le 0.05$), as well as between the content of chlorophyll *b* and anthocyanins ($r = 0.47$, $p \le 0.05$), the content of β-carotene ($r = 0.26$, $p \le 0.05$) and the content of monosaccharides ($r = -0.29$, $p \le 0.05$) has been noted. We have identified tomato accessions with a high content of individual chemical substances, as well as with a complex of traits that can be used as sources in breeding for a high content of dry matter, sugars, ascorbic acid, pigments and anthocyanins.
Key words: tomato; fruit color; biochemical compounds; pigments; anthocyanins.

Биохимический состав плодов томата различной окраски

А.Б. Курина ✉, А.Е. Соловьева, И.А. Храпалова, А.М. Артемьева

Федеральный исследовательский центр Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова (ВИР), Санкт-Петербург, Россия
✉ nastya_n11@mail.ru

Аннотация. Томат (*Lycopersicon esculentum* Mill.) – экономически важная и широко возделываемая овощная культура, потребляется как в свежем, так и в переработанном виде. Пищевая ценность плодов томата связана с содержанием в них каротиноидов, полифенолов, растворимых сахаров, органических кислот, минералов и витаминов. В настоящее время растет интерес к качественному и количественному увеличению содержания полезных для здоровья соединений в плодах томата. Коллекция генетических ресурсов *Lycopersicon* (Tourn.) Mill. Всероссийского института генетических ресурсов растений им. Н.И. Вавилова (ВИР) включает 7678 образцов одного культурного и девяти диких видов, что представляет широкие возможности для поиска сведений об изменчивости содержания биологически активных веществ и отбора в генофонде источников с высоким их содержанием. В нашей работе приведены результаты изучения 70 образцов культурного и дикорастущего томата по основным биохимическим признакам – содержанию сухого вещества, аскорбиновой кислоты, сахаров, каротинов, хлорофиллов и антоцианов. Для изучения взяты образцы с разнообразной окраской плодов, включая новые образцы с различным содержанием антоциана. В результате исследования определена амплитуда изменчивости содержания сухих веществ (3.72–8.88 и 9.62–11.33 %), сахаров (1.50–5.65 и 2.20–2.70 %), аскорбиновой кислоты (12.40–35.56 и 23.62–28.14 мг/100 г), титруемой кислотности (0.14–0.46 и 0.33–0.48 %), хлорофиллов (0.14–5.11 и 2.95–4.57 мг/100 г), общих каротиноидов (0.97–99.86 и 1.03–10.06 мг/100 г) и антоцианов (3.00–588.86 и 84.31–152.71 мг/100 г) в плодах культурного и дикорастущего томата соответственно. Определены корреляционные связи между содержанием сухих веществ и моно-

А.Б. Курина, А.Е. Соловьева
И.А. Храпалова, А.М. Артемьева

Биохимический состав плодов томата
различной окраски

2021
25•5

сахаридов ($r = 0.40$, $p \leq 0.05$), суммы сахаров ($r = 0.37$, $p \leq 0.05$) и аскорбиновой кислоты ($r = 0.32$, $p \leq 0.05$), содержанием аскорбиновой кислоты и каротиноидов ($r = 0.25$, $p \leq 0.05$). Выявлена высокая зависимость содержания хлорофиллов *a* и *b* между собой ($r = 0.89$, $p \leq 0.05$), а также зависимость средней степени между содержанием хлорофилла *b* и антоцианов ($r = 0.47$, $p \leq 0.05$), содержанием β-каротина ($r = 0.26$, $p \leq 0.05$) и содержанием моносахаридов ($r = -0.29$, $p \leq 0.05$). Выделены образцы томата с высоким содержанием отдельных химических веществ, а также по комплексу признаков, которые могут быть использованы в качестве источников в селекции на повышенное содержание сухого вещества, сахаров, аскорбиновой кислоты, пигментов и антоцианов.

Ключевые слова: томат; окраска плодов; биохимические соединения; пигменты; антоцианы.

Introduction

The beneficial properties of vegetables are associated with the presence of various compounds in them – phytochemicals beneficial to human health. Vegetable crops are the main source of natural antioxidants that have chemoprotective and anti-cancer effects (Zanfini et al., 2010; Chandra, Ramalingam, 2011).

Tomato (*Lycopersicon esculentum* Mill. = syn. *Solanum lycopersicum* L.) is an economically important crop that ranks first among vegetable crops in terms of cultivation area in the world, and is consumed both fresh and processed. About 182.3 million tons of tomato fruits are grown on 4.85 million hectares in the world annually (FAOSTAT, 2019). Asia accounts for 61.1 % of world production, Europe, America and Africa – 13.5, 13.4 and 11.8 % of the total harvest, respectively. Tomato consumption is concentrated in China, India, North Africa, the Middle East, the USA and Brazil, where it ranges from 61.9 to 198.9 kg per capita (FAOSTAT, 2019).

Tomato was introduced to Europe from Central and Southwest America in the 16th century. It was originally used as an ornamental plant, and then gradually became an important crop in human nutrition (Peralta, Spooner, 2007). As a result of domestication in the world, several different groups of tomato cultivars have been bred, differing in the size, shape and color of the fruits (Bhattarai et al., 2018). Much of the genetic variation was lost during domestication (Bai, Lindhout, 2007), and selection for new productivity traits had a negative impact on several other important traits, such as stress tolerance and fruit quality (Tanksley, 2004; Gascuel et al., 2017).

The nutraceutical value of tomato fruits is explained by the content of carotenoids, polyphenols, soluble sugars, organic acids, minerals and vitamins, especially vitamin C and E (Leiva-Brondo et al., 2012; Raiola et al., 2015; Martí et al., 2016), as well as volatile compounds (Wang, Seymour, 2017). Their antioxidant capacity depends on both lipophilic (carotenoids and vitamin E) and hydrophilic (vitamin C and phenolic compounds) fractions (Ilić et al., 2009). Tomato carotenoids are the main source of lycopene in the human diet (Viuda-Martos et al., 2014). Anthocyanins are not usually found in tomatoes, but flavonols (mainly quercetin, myricetin, and kaempferol) and flavanones (naregenin) have been found (Scarano et al., 2018). Bioactive compounds of tomato fruits have a wide range of physiological properties, including anti-inflammatory, antiallergenic, antimicrobial, vasodilating, antithrombotic, cardioprotective and antioxidant effects (Martí et al., 2016; Mozos et al., 2018). Epidemiological evidence suggests that consumption of tomatoes and tomato products is associated with a reduced risk of prostate cancer and other chronic diseases (Campbell et al., 2004; Zanfini et al., 2010; Wei, Giovannucci, 2012; Friedman, 2013).

Currently, there is a growing interest in the qualitative and quantitative increase in the content of healthy compounds in tomato fruits in order to further increase the nutraceutical potential of the crop. Modern biochemical research is aimed at identifying and quantifying the components of plant materials, as well as determining their biological activity. Such data are needed, among other things, for the development of beneficial nutritional and nutraceutical supplements.

VIR *Lycopersicon* (Tourn.) Mill. worldwide collection includes 7678 accessions of one cultivated and nine wild species (according to C.M. Rick (1959)). Tomato (*L. esculentum* Mill.) has 6536 varietal and 1505 hybrid populations (F$_3$–F$_5$). The first tomato accessions entered in the collection in 1922 as a result of the expedition of N.I. Vavilov and S.M. Bukasov to the USA and Canada. These were stem indeterminate forms with different colors of the fruit, relevant for use in breeding until now. Then the collection was expanded with accessions of various types of growth and development and morphological features from 95 countries all over the world, that is, the widest crop diversity for various uses and sources for breeding is concentrated in it. The collection continues to grow. Currently, much attention is paid to the involvement in the collection of competitive accessions with unconventional fruit color: yellow, orange, pink, crimson, green, brown, purple, "black", characterized by a high content of biologically active flavonoids and pigments.

The modern structure of the VIR tomato collection: accessions of wild species – 196; primitive forms – 371; landraces – 551; breeding cultivars – 4188; hybrids – 1511; mutant forms – 49; self-pollinated lines – 118; genetic sources with identified genes – 278; donors – 17 accessions.

The purpose of this work is to conduct a comparative assessment of tomato accessions with different fruit colors from the VIR collection in terms of biochemical composition. The main task was to determine the content of the main chemical compounds – dry matter, ascorbic acid, sugars, chlorophylls, carotenoids and anthocyanins – in various tomato accessions from the VIR collection.

Materials and methods

The material for the study included breeding cultivars and hybrids of different time of creation of cultivated and wild tomatoes from the VIR collection (70 accessions in total), differing in many phenological and morphological characteris-

tics: the duration of the vegetative period, the type of growth, height of inflorescences formation, the number and type of inflorescences, flower features, fruit formation, shape, size, surface and internal structure of the fruit. Accessions with various colors of fruits were taken for research, including new accessions with different contents of anthocyanin. The studied tomato accessions represented a wide range of colors of fruits in biological ripeness: green-purple, green-yellow, yellow, yellow-orange, yellow-purple, orange, orange-red, red, pink, crimson, red-brown and purple-red, both with a uniform color and with the presence of yellow or green stripes (Table 1).

Tomato accessions were grown at the VIR Pushkin and Pavlovsk Laboratories (St. Petersburg, Northwest Russia) in a glass greenhouse in March-October of 2019–2020. The temperature was 22–30 °C during the day and 16–22 °C at night. The plants were grown only under natural light. Daylight duration varied from 11.5 h in March and September to 18.6 h in the third decade of June. Illumination varied from 4 to 10 thousand lx/m^2, depending on the growing season. The plants were grown in AgroBalt peat substrate, completely filled with mineral fertilizers. 3–9 plants were placed per 1 m^2, depending on the plant habit. Generally accepted in the Northwest region agricultural technologies included the garter and the formation of plants, taking into account the specific conditions of the winter shelving greenhouse.

Phenological and morphological descriptions were carried out according to the "International CMEA Classifier of the Genus *Lycopersicon* Tourn." (1986), "Descriptors Tomato (*Lycopersicon* spp.) IPGRI" (1996) and "Tomato – UPOV (*Solanum lycopersicum* L.)" (2012).

Biochemical analysis was carried out in the VIR Laboratory of Biochemistry and Molecular Biology in the Biological Ripeness of the Fruits. The study took 1/2 part of at least five fruits of each accession, in two replications. The analysis and processing of the material were carried out according to the VIR methods (Ermakov et al., 1987): the dry matter content was measured by a gravimetric method; sugars – by the Bertrand's method; total (titratable) acidity – by titrating with 0.1 n of alkali, calculated as malic acid; ascorbic acid – by the method of direct extraction from plants with 1 % hydrochloric acid, followed by titration with 2,6-dichloroindophinol (Tillman's reagent); carotenoids and chlorophylls were isolated with 100 % acetone and their absorption was measured on an Ultrospec II spectrophotometer at different wavelengths (nm): 645, 662 – for chlorophylls *a* and *b*, 440 – for carotenoids, 454 – for carotenes (total carotenes determined by paper chromatography), 454 – for β-carotene, 503 – for lycopene. Anthocyanins were extracted by 1% hydrochloric acid, then measured by spectrophotometry at 510 nm wavelength, in terms of cyanidin-3,5-diglycoside (453 nm), with a correction for the content of the green pigment at 657 nm. All data are presented on raw material.

Statistical analysis. Descriptive statistics (mean, median, standard error, standard deviation, range of variability) were calculated for all biochemical parameters to assess the genetic diversity of tomatoes. Data analysis was performed using the STATISTICA v.12.0 software (StatSoft Inc., USA). Data testing for normality of distribution was performed using the Shapiro–Wilk test and the quantile-quantile plot (QQ Plot). The mean values of the data with normal distribution were compared using one-way analysis of variance (ANOVA); Pearson's correlation coefficient was used for correlation analysis. Data with a distribution other than normal were compared using the Kruskal–Wallis test, and correlation analysis was compared using the Spearman's rank correlation coefficient. Cluster analysis was performed using the UPGMA method in the PAST program (Hammer et al., 2001).

Results and discussion
As a result of studying the most important indicators of the biochemical composition of tomato fruits, the large differences between the studied accessions were established.

Dry matter content
One of the most important indicators of the quality of tomato fruits and their technological properties is the dry matter content. The dry matter content in the fruits of cultivated tomato was in the range of 3.72–8.88 % (Cv = 14.7 %), in the fruits of wild species – 9.62–11.33 % (Cv = 6.2 %) (Table 2). Fruits with a high concentration of dry substances taste good, give a higher yield during processing, and have better transportability and keeping quality during storage. On average, the red-brown accessions accumulated more dry matter (6.46 %) than the rest. A high content of dry matter (more than 7.00 %) was noted in the accessions Slivka krasnaya, Ampel'nyj F_1 and Patrikeevna. Among the accessions of wild species, the largest amount of dry matter in fruits was accumulated by accessions of *L. peruvianum*: 10.25–11.33 %.

In our study, we found weak positive correlations in cultivated tomato accessions between the content of dry matter and monosaccharides ($r = 0.40$, $p \leq 0.05$), the amount of sugars ($r = 0.37$, $p \leq 0.05$) and ascorbic acid ($r = 0.32$, $p \leq 0.05$). Thus, an increase in the amount of these biochemical characteristics in fruits will have little effect on an increase in the other three indicators.

Our results on the dry matter content are consistent with the results of other studies (Gupta et al., 2011; Nour et al., 2013; Kondratyeva, Engalychev, 2019; Ignatova et al., 2020), which reported on the dry matter content in tomato fruits within 5.55–8.80 %.

Sugar content
Most of the dry matter in tomato fruits is carbohydrates, the main of which are soluble sugars. In our study, the total sugar content in cultivated forms was 1.50–5.65 % (Cv = 26.2 %), in accessions of wild species – 2.20–2.70 % (Cv = 6.4 %) (see Table 2). The high variability of the sugar content in cultivated tomato is associated with both genetic characteristics and growing conditions. Soluble monosaccharides are represented by glucose and fructose in tomato fruits. The average content of monosaccharides was 2.84 % in cultivated and 1.98 % in wild tomatoes. The accessions Superklusha and Patrikeevna showed a high content of total sugars, including monosaccharides – 5.35 and 5.65 %, respectively.

Oligosaccharides in tomato fruits are mainly represented by the disaccharide sucrose. The content of disaccharides

А.Б. Курина, А.Е. Соловьева
И.А. Храпалова, А.М. Артемьева

Биохимический состав плодов томата
различной окраски

2021
25•5

Table 1. List of the studied tomato accessions

No.	VIR catalog number	Accession name	Origin/firm	Fruit color
		Accessions of cultivated edible tomato *Lycopersicon esculentum* Mill. = *Solanum lycopersicum* L.		
1	vr.k-15325	Citrusovyj sad	Crimea, Feodosia	Lemon yellow
2	vr.k-15343	Patrikeevna	Semena Altaya, Yu.V. Fotev	Yellow
3	vr.k-15344	Zolotce	Semena Altaya, Yu.V. Fotev	
4	k-3766	Yantarnyj	Northwest Research Institute of Agriculture "Belogorka"	
5	vr.k-11243	Mestnyj	Madagascar	
6	vr.k-15368	Zheltyj delikates	Poisk	
7	vr.k-15303	Yellow Ruffles	USA	
8	g.k-01040	Cypa F_1	VIR	
9	vr.k-15361	Indigo Gold Berries	USA, Wisconsin	Yellow-purple
10	vr.k-15306	Stripes of Yorc	USA	
11	vr.k-15333	Utenok	All-Russian Research Institute of Horticulture, V.I. Kozak	Yellow-orange
12	vr.k-14426	Hurma	SSF Gisok	
13	vr.k-15338	Dina	Institute of General Genetics	
14	vr.k-15365	Gold Medal	USA, Wisconsin	
15	vr.k-15309	Yaponskij tryufel' oranzhevyj	Sibirskiy sad	Orange
16	vr.k-15315	Il'ya Muromec	Ogorodnoye izobiliye	
17	vr.k-15319	Chemal'skij 2	Altai region	
18	vr.k-15328	Mestnyj	Altai region, Rubtsovsk, market	
19	vr.k-15369	Maksi-karotin F_1	VIR	
20	k-4085	Slivka krasnaya	Botanical Garden, Semena Altaya	Orange-red
21	k-6582	Karlik kartofel'nyj	Volkhov, Leningrad region	
22	vr.k-15357	Podarok Kubani	FSBSI FSVC, Poisk	
23	k-4482	Novichok	VIR Volgograd experimental station	
24	k-4895	Assorti	VIR	
25	vr.k-14430	Valyuta	VIR	
26	vr.k-15321	Beduin	Sevastopol	
27	vr.k-15323	Hybrid Budenovka × Chernyj princ	Sevastopol	Orange-red with purple
28	vr.k-15310	Majkl Pollan	Nash sad	Orange-red with green stripes
29	vr.k-15326	Percevidnyj mestnyj	Crimea	Orange-red with yellow stripes
30	vr.k-15345	Cherry rozovyj	Botanical Garden	Pink
31	vr.k-15352	Nepas 12	Sedek	
32	vr.k-15356	Superklusha	O.V. Postnikova	
33	k-6881	Dikaya roza	Pridnestrovskiy Research Institute of Agriculture, Aelita	
34	vr.k-15308	Yaponskij tryufel' rozovyj	Sibirskiy sad	
35	vr.k-15312	Persik	Sedek	
36	vr.k-15317	Bych'e serdce rozovoe	Altai region	
37	vr.k-15320	Amurskij tigr	Crimea	Pink with yellow stripes
38	vr.k-15322	Percevidnyj rozovyj	Sevastopol	Pink
39	k-6938	Zyryanka	Botanical Garden	Red

Table 1 (end)

No.	VIR catalog number	Accession name	Origin/firm	Fruit color
40	vr.k-15347	Hardins Miniature	USA	Red
41	k-4571	Ion	FSBSI FSVC	
42	k-2497	Mongol'skij karlik	Partner	
43	k-3043	Nevskij	Northwest Research Institute of Agriculture "Belogorka"	
44	vr.k-15350	Kamennyj cvetok	V.I. Blokin-Mechtalin	
45	vr.k-15351	Krasnaya rossyp'	Poisk	
46	k-5647	Kraynij sever	Biotekhnika, V.I. Kozak	
47	vr.k-15335	Severnaya malyutka	VSTISP, V.I. Kozak	
48	k-7035	Budenovka	Tomagros	
49	vr.k-15367	Speckled Roman	USA, Wisconsin	Red with yellow stripes
50	k-6128	Chernyj princ	Scientific and Production Corporation "NK.LTD"	Red-brown
51	vr.k-15354	Ampel'nyj F_1	Seed-breeding agrofirm Il'inichna	
52	k-7256	Chernyj mavr	Scientific and Production Corporation "NK.LTD"	
53	k-5169	Cherokee Purple	USA, Tennessee	
54	vr.k-15314	Viagra	Gavrish	
55	vr.k-13341	Indigo Blue Berries	USA, Wisconsin	Purple-red
56	vr.k-15362	Indigo Helsing Junction Blue	USA, Wisconsin	
57	vr.k-15363	Indigo Clackamas Blue Berry	USA, Wisconsin	
58	vr.k-15364	Indigo Apple	USA, Wisconsin	
59	vr.k-15366	Ananas Noire	USA, Wisconsin	
60	vr.k-15302	OSU Blue	USA	
61	vr.k-15304	Blue Berry	USA	
62	vr.k-15305	Amethyst Jewel	USA	
63	vr.k-15307	Lyagushka-tsarevna	Gavrish	Green-yellow
64	k-5407	Rin	All-Russian Research Institute of Vegetable Growing	
		Accessions of wild inedible species *Lycopersicon peruvianum* Mill. = *Solanum peruvianum* L. *Lycopersicon glandulosum* C.H. Muell = *Solanum corneliomuelleri* J.F. Macbr.		
65	k-3924	*L. peruvianum* Mill. PL 129152	Japan	Green-purple
66	k-3960	*L. peruvianum* Mill. var. *dentatum* FPI 128650/1236	Chile	
67	k-3962	*L. peruvianum* Mill. var. *dentatum* FPI 128652/1238		
68	k-2904	*L. glandulosum* C.F. Mull. ES 495	India	
69	k-3944	*L. glandulosum* C.F. Mull.	Peru	
70	k-2099	*L. peruvianum* Mill.	USA	

in the studied accessions was low and averaged 0.2–0.4 % in both cultivated and wild forms. Accessions Kamennyj cvetok, Zolotce and Dikaya roza contained more than 1.2 % disaccharides in fruits.

Several studies have shown that green fruits of wild tomato species accumulate mainly sucrose, while fruits of cultivated tomato accumulate glucose and fructose (Stommel, 1992; Beckles et al., 2012). In our study, wild tomato fruits accumulated more monosaccharides than disaccharides, which is possibly related to the growing conditions. The Leningrad region is characterized by low insolation, which is possibly the reason for the low accumulation of disaccharides.

А.Б. Курина, А.Е. Соловьева
И.А. Храпалова, А.М. Артемьева

Биохимический состав плодов томата
различной окраски

2021
25•5

Table 2. Parameters of descriptive statistics of cultivated and wild tomato accessions by biochemical characteristics

Trait	Type	Mean ± SE	Median	Min–max	Std. dev.
Dry matter, %	Cultivated	5.65 ± 0.10	5.65	3.72–8.88	0.83
	Wild	10.28 ± 0.25	10.28	9.62–11.33	0.61
Total sugars, %	Cultivated	3.06 ± 0.10	3.06	1.50–5.65	0.79
	Wild	2.42 ± 0.07	2.44	2.20–2.70	0.17
Monosaccharides, %	Cultivated	2.84 ± 0.10	2.83	1.47–5.65	0.82
	Wild	1.98 ± 0.10	1.94	1.65–2.33	0.25
Total acidity, %	Cultivated	0.28 ± 0.01	0.26	0.14–0.46	0.08
	Wild	0.40 ± 0.02	0.39	0.33–0.48	0.06
Ascorbic acid, mg/100g	Cultivated	20.78 ± 0.64	19.22	12.40–35.56	5.12
	Wild	26.22 ± 0.75	26.29	23.62–28.14	1.84
Chlorophylls*, mg/100g	Cultivated	1.40 ± 0.13	1.12	0.14–5.11	1.17
	Wild	3.83 ± 0.24	3.94	2.95–4.57	0.60
Chlorophyll a*, mg/100g	Cultivated	0.64 ± 0.07	0.45	0.05–2.81	0.58
	Wild	2.56 ± 0.16	2.59	2.02–3.03	0.38
Chlorophyll b*, mg/100g	Cultivated	0.76 ± 0.07	0.69	0.08–2.61	0.59
	Wild	1.27 ± 0.09	1.34	0.93–1.54	0.22
Total carotenoids*, mg/100g	Cultivated	21.86 ± 2.83	12.18	0.97–99.86	22.61
	Wild	2.68 ± 1.48	1.27	1.03–10.06	3.62
Carotenoids*, mg/100g	Cultivated	5.74 ± 0.44	5.98	0.68–15.91	3.51
	Wild	1.19 ± 0.07	1.13	1.03–1.40	0.16
Lycopene*, mg/100g	Cultivated	16.20 ± 2.63	7.44	0.00–89.39	21.02
	Wild	–	–	–	–
Carotenes*, mg/100g	Cultivated	2.31 ± 0.19	2.10	0.27–6.24	1.50
	Wild	0.62 ± 0.07	0.58	0.50–0.90	0.15
β-carotene, mg/100g	Cultivated	0.68 ± 0.05	0.68	0.08–1.62	0.39
	Wild	0.20 ± 0.01	0.21	0.17–0.24	0.03
Anthocyanins*, mg/100g	Cultivated	45.20 ± 12.81	10.63	3.00–588.86	102.48
	Wild	125.30 ± 9.68	125.85	84.31–152.71	23.70

* The data have abnormal distribution.

The sugar content in tomato fruits varied in the range of 2.81–4.41 % in the study of A.V. Kuzyomensky (2004), the sugar content was in the range of 2.12–6.00 % in F_1 hybrids (Ignatova et al., 2020).

Titratable acidity and sugar-acid index
The titratable acidity in fruits of cultivated tomato accessions varies within 0.14–0.46 % (Cv = 28.4 %) with an average content of 0.28 %, in wild tomatoes – 0.33–0.48 % (Cv = 15.0 %) with an average content of 0.40 %. A low content of titratable acids (less than 0.19 %) was observed in tomato accessions: Karlik kartofel'nyj, Utenok, Yantarnyj, Gold Medal and Yellow Ruffles. A high content (more than 0.40 %) was noted in

tomato accessions with pink, orange, orange-red and yellow-purple color of fruits: Amurskij tigr, Bych'e serdce rozovoe, Stripes of Yorc, Yaponskij tryufel' oranzhevyj and rozovyj, Valyuta. High acidity in wild tomato species (0.48 %) was noted in two accessions: *L. glandulosum* (k-3944, Peru) and *L. peruvianum* (k-2099, USA).

Similar results on the level of titratable acidity were obtained in other studies. In R.V. Nour et al. (2013) and J. Owusu et al. (2012) studies titratable acidity varied from 0.10 to 0.41 %.

The taste of the fruit is determined by the index of sugar to acid. It has been proven that this indicator changes depending on soil and climatic conditions, cultivation techniques and
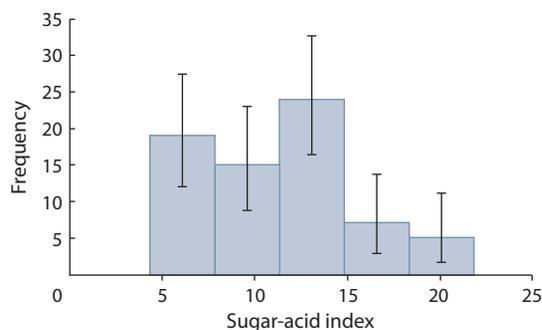
Fig. 1. Distribution of tomato accessions by sugar-acid index.

varietal characteristics of the crop, as well as the degree of fruit ripeness (Kondratyeva, Pavlov, 2009). The index of sugar to acid is an indicator of the quality of the fruit: the higher it is, the tastier the product.

It was found that all accessions had a different sugar-acid index. The level of sugar-acid index in cultivated tomatoes ranged from 4.41–21.80 (Cv = 34.1 %), in wild ones – 4.63–7.01 % (Cv = 14.2 %). Tomato accessions were divided into six statistically significant groups of the sugar-acid index (Fig. 1). The first group included 16 tomato accessions, which were characterized by a low index – 4.4–7.3. This group included all accessions of wild tomato, as well as accessions of cultivated tomato with yellow (Mestnyj), yellow-purple (Stripes of Yorc), orange-red (Valyuta), orange (Yaponskij tryufel' oranzhevyj), pink (Yaponskij tryufel' rozovyj, Amur-skij tigr), red (Severnaya malyutka, Krajnij sever) and green-yellow (Rin) color of fruits. The second (7.3–10.2), third (10.2–13.1) and fourth (13.1–16.0) groups included 14–17 accessions with different fruit colors. The fifth and sixth groups included accessions with a high index: 16.0–21.8. These groups were represented by accessions with yellow (Patrikeevna), yellow-orange (Dina, Gold Medal), pink (Di-kaya roza), red (Zyryanka), orange-red (Karlik kartofel'nyj), purple-red (Indigo Helsing Junction Blue, Amethyst Jewel) and red-brown (Chernyj princ) fruit color.

Ascorbic acid content

The nutritional value of tomato fruits is determined, first of all, by the high content of vitamins, among which ascorbic acid (vitamin C) occupies one of the first places. The content of ascorbic acid in the analyzed fruits of cultivated tomato varied from 12.40 to 35.56 mg/100 g (Cv = 24.6 %) with an average content of 20.78 mg/100 g, in wild ones – from 23.62 to 28.14 mg/100 g (Cv = 6.0 %) with an average content of 26.22 mg/100 g. A high content of ascorbic acid (more than 30 mg/100 g) was found in the accessions Utenok, Amethyst Jewel, Yaponskij tryufel' rozovyj and oranzhevyj.

A weak correlation between the content of ascorbic acid with a dry matter content ($r = 0.32$, $p \leq 0.05$) and carotenoids ($r = 0.25$, $p \leq 0.05$) was found in the studied accessions.

R.V. Nour et al. (2013) found significant differences in the content of ascorbic acid in different tomato cultivars: 91.9–329.7 mg·kg^{-1}. R.A. Dar and J.P. Sharma (2011) found ascorbic acid content in the range of 197.7 to 378 mg·kg^{-1} FW, Harish et al. (2012) – within 20.23–29.32 mg/100 g.

Thus, our results are partially consistent with the already available results and also expand the range of variability of the content of ascorbic acid in tomato fruits.

Chlorophyll content

The amount of pigments and their ratio significantly affect the metabolism of plants and can differ depending on the species or cultivar of the plant, as well as on the phase of its ontogenesis (Belova et al., 2012).

Chlorophyll is found in large quantities in unripe tomato fruits; in the process of ripening, it is destroyed. During the ripening of tomatoes, the degradation of chlorophyll happens with the biosynthesis and accumulation of carotenoids at the same time; both processes are responsible for the change of fruit color.

In our study, the total chlorophyll content in cultivated tomatoes was in the range of 0.14–5.11 mg/100 g, in wild ones – 2.95–4.57 mg/100 g (see Table 2). Tomato accessions with different fruit colors differed in the content of chloro-phyll *a* and *b* (Fig. 2).

In addition to the total chlorophyll content, the adaptabil-ity of plants to a certain lighting regime is also manifested in the qualitative composition of pigments. In our study, as expected, the largest amount of chlorophyll *a* in fruits was accumulated by accessions of wild tomato. A high content was found in accessions of cultivated forms with green-yellow (1.45–1.69 mg/100 g), red-brown (1.25–2.32 mg/100 g) fruit color and in several accessions with purple-red fruit color: Blue Berry (vr.k-15304) – 1.53 mg/100 g and Indigo Clackamas Blue Berry (vr.k-15363) – 2.81 mg/100 g. The other accessions contained not more than 1.00 mg/100 g of chlorophyll *a*.

Tomato accessions with yellow, yellow-orange, red, orange, orange-red, pink and purple-red fruit color accumulated more chlorophyll *b* in the fruit. The highest content (more than 1.00 mg/100 g) was noted in the fruits of wild tomato and in most accessions of cultivated tomato with purple-red color of fruits, as well as some accessions with red: Mongol'skij kar-lik (1.20 mg/100 g), Nevskij (1.76 mg/100 g), Krajnij sever (2.43 mg/100 g); orange-red: Beduin (1.14 mg/100 g), Podarok Kubani (1.89 mg/100 g), Valyuta (2.61 mg/100 g); and pink: Cherry rozovyj (1.46 mg/100 g), Nepas 12 (1.58 mg/100 g), fruit color.

One of the informative indicators characterizing the poten-tial photochemical activity of fruits is the ratio of chlorophyll *a* to chlorophyll *b* (*a/b*). The possible effect of fruit ripening on the rate of destruction of pigments was reflected in the value of the ratio of the chlorophyll content – *a* to *b*. In our study, chlorophyll *b* prevails in the total chlorophyll pool of culti-vated tomatoes. The chlorophyll *a/b* ratio was in the range of 0.45–2.17 in most cultivated tomato accessions, and in wild ones – 1.85–2.53. All accessions with yellow-green and red-brown fruit color had a chlorophyll *a/b* ratio more than 1, as well as some accessions with yellow (Zheltyj delikates, Zolotce), red (Zyryanka), orange-red (Novichok, Hybrid Bu-denovka × Chernyj princ) and purple-red (Blue Berry, Indigo Clackamas Blue Berry) color of the fruit.

In general, accessions with green-purple, green-yellow, red-brown and purple-red coloration in total accumulated

А.Б. Курина, А.Е. Соловьева
И.А. Храпалова, А.М. Артемьева

Биохимический состав плодов томата
различной окраски

2021
25•5
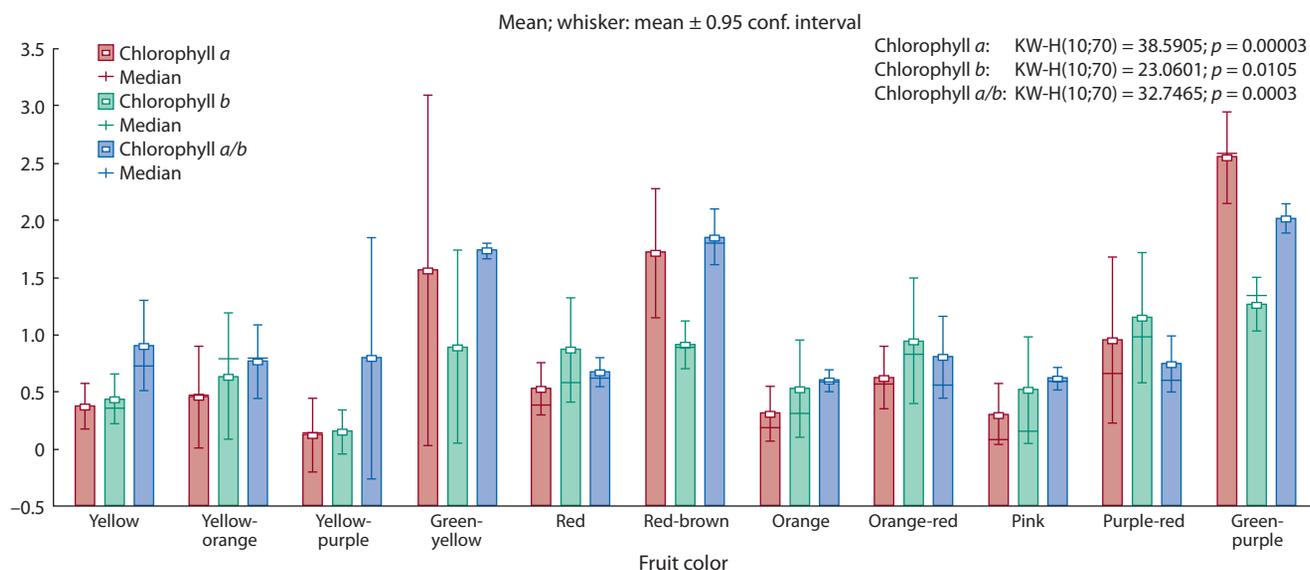
Fig. 2. Variability of tomato accessions with different fruit colors in terms of chlorophyll content.

more chlorophylls in their fruits – more than 2.40 mg/100 g, accessions with orange-red, red and yellow-orange coloration – within the range of 1.10–1.58 mg/100 g, the total content of chlorophylls in fruits with a different color did not exceed 0.85 mg/100 g.

Thus, it can be assumed that the presence of chlorophylls in tomato fruits with yellow, yellow-orange, orange, orange-red, red and pink coloration is due to the fact that the chlorophyll degradation process was not yet completed and proceeded in parallel with the synthesis of carotenoids, and the prevalence of chlorophyll *b* indicates that the rate of photosynthesis has already been reduced. The high accumulation of chlorophylls in the fruits of some tomato accessions with red, orange-red and yellow-orange color of the fruit may be associated with the presence of a green spot.

Carotenoid content

The diversity of tomato fruit color is the result of mutations in the genes of the carotenoid pathway, which arose as a result of domestication and improvement of varieties, such as *yellow-flesh* (*r*), *tangerine* (*t*), *green-flesh* (*gf*), *green ripe* (*gr*), *apricot* (*at*), *beta carotene* (*B*), *high pigment* (*hp*), *old gold* (*og*), and *y* (*yellow*) (Roohanitaziani et al., 2020). Changing the classic red color of tomato, such genes primarily affect its biochemical composition, and especially the content of carotenoids, allowing the creation of varieties with a changed content of these substances (Kuzyomensky, 2004).

The most common carotenoids of red tomato varieties are lycopene (red pigment) and β-carotene (yellow-orange pigment), while lutein, ζ-carotene, neurosporin, and others may also be present in orange and yellow fruits (Khachik et al., 2002). Other identified tomato carotenoids – γ-carotene, phytoene, phytofluen, are found in small amounts (Golubkina et al., 2017).

The pigment complex of the fruits of the studied tomato accessions was characterized by a high content of carotenoids. The total content of carotenoids in the fruits of cultivated

tomato was in the range of 0.97–99.86 mg/100 g, with an average content of 21.86 mg/100 g, in wild tomato – 1.03–10.06 mg/100 g, with average content – 2.68 mg/100 g.

The variability of the carotene content in tomato accessions was high (Cv = 64.9 %). The average carotene content of cultivated tomato accessions was 2.31 mg/100 g, of which β-carotene was 0.68 mg/100 g (Fig. 3). High carotene content was found in tomato accessions with red-brown (average 3.25 mg/100 g) and orange (4.03 mg/100 g) fruit color, low (less than 0.80 mg/100 g) was found in accessions with yellow, green-yellow, yellow-purple and green-purple color of fruits, the rest contained on average 1.66–2.85 mg/100 g. At the same time, a high content of β-carotene was found in tomato accessions with pink (average 0.89 mg/100 g) and orange-red (0.95 mg/100 g) fruit color, slightly less (0.81–0.82 mg/100 g) in accessions with red-brown and orange colored fruits. Accessions Valyuta (vr.k-14430), Kraynij sever (k-5647), and Novichok (k-4482) contained more than 1.40 mg/100 g of β-carotene.

Lycopene is a non-cyclic β-carotene isomer. The content of lycopene in fruits of cultivated tomato varied from 0.00 to 89.39 mg/100 g; lycopene was not found in fruits of wild tomato (see Table 2). The differences between the accessions in terms of lycopene content are very large, including the differences within the fruit color groups. Accessions with pink and orange-red color of fruits were characterized by a high lycopene content (on average 26.32–32.52 mg/100 g), accessions with green-yellow, yellow and yellow-purple color of the fruit accumulated significantly less (less than 6.5 mg/100 g). Accessions with red and orange color of fruits in our study had similar values for lycopene content – 8.80 and 8.37 mg/100 g, as well as accessions with red-brown, yellow-orange and purple-red color of fruits: 16.12, 17.04 and 18.04 mg/100 g, respectively (see Fig. 3).

Accessions with yellow-orange (Dina and Gold Medal), red-brown (Viagra) and purple-red (OSU Blue) color of the fruit showed a high content of lycopene (for each group: 21.62,
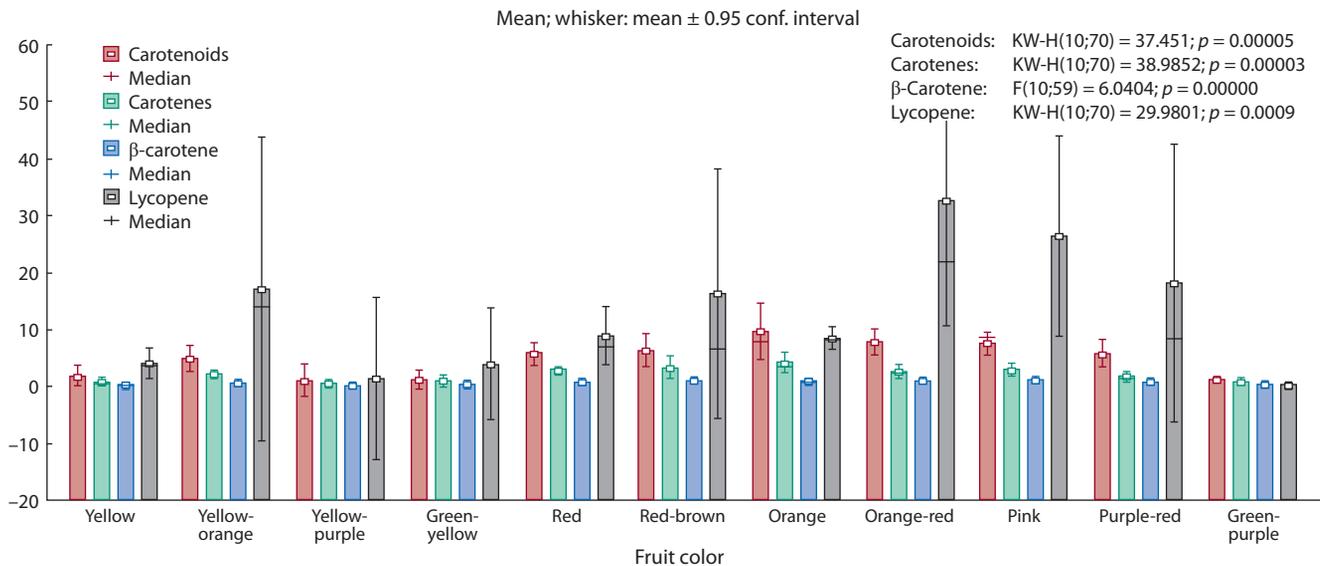
Fig. 3. Variability of tomato accessions of various fruit colors in terms of the content of carotenoids.

38.71, 45.67 and 89.39 mg/100 g, respectively). Accession Zheltyy delikates did not accumulate lycopene in the fruit.

The data on the content of lycopene and β-carotene is very different in the works of other authors. As a result of studying 10 red tomatoes, R.V. Nour et al. (2013) found that the content of lycopene was in the range of 19.7–49.0 mg · kg^{-1}, and β-carotene – 6.4–12.8 mg · kg^{-1}. After studying 185 tomato accessions, S. Anjum et al. (2020) determined that the content of lycopene was 1.57–23.24 mg · 100 g^{-1}, β-carotene – 1.32–7.61 mg · 100 g^{-1}. R.S. Pal et al. (2018) reported the content of lycopene in the studied 22 tomato lines in the range of 3.05–9.83 mg/100 g and β-carotene – 4.32–7.31 mg/100 g. In a study by I.Yu. Kondratyeva and N.A. Golubkina (2016), the content of lycopene in tomato accessions with yellow and orange color of the fruit was in the range of 0.0–2.6 mg/100 g, in fruits with red and pink color – 3.3–11.5 mg/100 g, β-carotene – 0.8–6.2 and 0.8–3.1 mg/100 g, respectively.

In our study, the proportion of β-carotene from the total content of carotenes is 25.7–28.4 % in accessions with yellow, yellow-orange and yellow-purple color of fruits, and the proportion of carotenes from carotenoids is 41.5–42.8 %. Thus, we can assume that the remaining carotenoid pigments in these accessions are xanthophylls, including lutein. At the same time, accessions with a yellow-orange color of the fruit accumulated a significant amount of lycopene (on average 17.0 mg/100 g). Accessions with a green-yellow color of the fruit were characterized by a high proportion of carotenes in the carotenoid complex – 71.7 %, but β-carotene averaged only 20.0 %. In accessions with red, red-brown and orange fruit color, the proportion of carotenes was in the range of 42.5–52.0 %, and β-carotene – 20.1–26.3 %, while the accessions with these fruit colors accumulated the greatest amount of carotenes (on average 2.7–4.0 mg/100 g) compared to the rest of the accessions. Accessions with orange-red and pink color of fruits accumulated the greatest amount of lycopene – on average 26.3–32.5 mg/100 g, while the proportion of carotenes was small – 34.5–38.2 %, with a proportion of β-carotene within 31.3–36.1 %. In accessions with a purple-

red color of the fruit, the proportion of carotenes was 29.0 % with a prevalence of β-carotene.

Thus, we can assume that tomato fruits also contain other carotenoid pigments that were not identified by us – lutein, ζ-carotene, γ-carotene, neurosporin, phytoene, phytofluen and others, which is consistent with the studies of other authors (Khachik et al., 2002; Golubkina et al., 2017).

Anthocyanin content

Normally, cultivated tomato plants do not synthesize anthocyanins in fruits. Three loci, *Anthocyanin fruit* (*Aft*), *atroviolacium* (*atv*), and *Aubergine* (*Abg*), enhance the accumulation of anthocyanins in fruits when they introgress from wild species into cultivated tomatoes (Kendrick et al., 1997; Jones et al., 2003). The *atv*, *Aft*, and *Abg* loci in wild tomato species can contribute to the pigmentation of anthocyanins in fruits, and the *atv* locus can dramatically increase the amount of anthocyanins in cultivated tomato fruits when it is combined with the *Aft* or *Abg* locus (Mes et al., 2008). Most of the anthocyanins present in the fruits of such tomatoes are concentrated in the skin, and almost complete absent in the seeds and pulp (Ooe et al., 2016).

In our study, a significant amount of anthocyanins was observed in accessions of cultivated tomato with purple-red (32.89–588.86 mg/100 g) and yellow-purple (87.91–161.22 mg/100 g) fruit color, as well as in wild tomato fruits (84.31–152.71 mg/100 g) (Fig. 4).

In accessions with other fruit colors, anthocyanins were also found, but in much smaller quantities. Fruits with red coloration accumulated anthocyanins on average 14.09 mg/100 g, with yellow, yellow-orange, green-yellow, orange and orange-red – within 10.62–11.77 mg/100 g, and accessions with red-brown and pink fruit color – less than 9.0 mg/100 g. Anthocyanin content of 53.3 mg/100 g was found in the Speckled Roman accession with red and yellow stripes.

The correlation analysis revealed a high dependence of the content of chlorophyll *a* and *b* among themselves ($r = 0.89$, $p \leq 0.05$), as well as an average positive correlation between

А.Б. Курина, А.Е. Соловьева
И.А. Храпалова, А.М. Артемьева

Биохимический состав плодов томата
различной окраски

2021
25•5
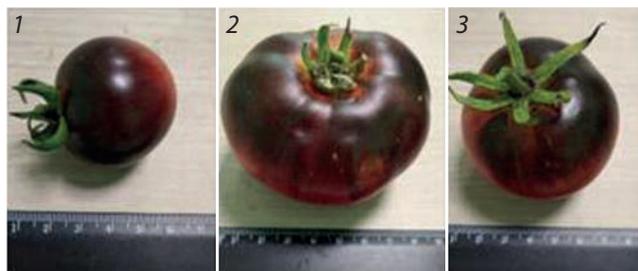
Fig. 4. Accessions of tomato high in anthocyanins: Indigo Clackamas Blue Berry (*1*), Indigo Apple (*2*), Ananas Noire (*3*).

the content of chlorophyll *b* and anthocyanins ($r = 0.47$, $p \leq 0.05$), weak – with the content of β-carotene ($r = 0.26$, $p \leq 0.05$) and weak negative – with the content of monosaccharides ($r = -0.29$, $p \leq 0.05$). Between chlorophyll *a* and anthocyanin there is also a positive correlation of average degree ($r = 0.37$, $p \leq 0.05$).

C.M. Jones et al. (2003) reported that the amount of anthocyanins in the peel of "blue" tomatoes ranged from 20.6 to 66.5 mg/100 g, in another study the amount of anthocyanins in the peel ranged from 7.79 to 110.79 mg/100 g (Peter et al., 2008). In studies of "purple" tomatoes obtained by the method of genetic engineering, *Del/Ros1*, *Del/Ros1 × AtMYB12*, the anthocyanin content is reported to be 5.1 ± 0.5 g·kg^{-1} DW and 1.154 ± 0.011 mg·g^{-1} FW (Lim et al., 2014; Zhang et al., 2015), and in tomato accessions obtained by breeding *Aft/Aft × atv/atv* – 116.11 mg·100 g^{-1} FW (Mes et al., 2008), V118 – 50.18 mg·100 g^{-1} DW (Li et al., 2011), Blue Ja-

pan Indigo tomato – 17 mg·g^{-1} DW (Ooe et al., 2016) and *Aft/Aft × atv/atv × hp2/hp2* – 90.91 mg·100 g^{-1} FW (Da Silva-Souza et al., 2020). E. Ooe et al. (2016) also reported that "blue" tomato accessions accumulate significant amounts of lycopene.

Thus, our studies of the anthocyanin content in tomato fruits are consistent with previous studies. As a result of the biochemical analysis, we identified tomato accessions by a set of traits that can be used as sources in breeding for a high content of sugars and biologically active substances (Table 3).

Cluster analysis

A dendrogram was constructed based on the results of cluster analysis of the studied biochemical parameters of tomato accessions (in accordance with Table 2) (Fig. 5). Tomato accessions were divided into two groups, small and large; within the second group, five clusters were identified.

The first cluster included two tomato accessions with a high content of anthocyanins and chlorophylls in fruits: Ananas Noire (430.3 and 2.63 mg/100 g) and Indigo Clackamas Blue Berry (588.9 and 5.11 mg/100 g).

The second cluster is divided into three sub-clusters. The first subcluster is represented by one accession from the USA – OSU Blue, the second – by accessions of wild tomato *L. glandulosum* (k-2904, k-3944) and *L. peruvianum* (k-2099, k-3924, k-3962), as well as accessions of cultivated tomato Stripes of Yorc with a yellow-purple color of the fruit and Indigo Apple (vr.k-15364) and Blue Berry (vr.k-15304) with a purple-red color of the fruit. The third subcluster included two accessions with a purple-red color of the fruit – Amethyst Jewel and Indigo Helsing Junction Blue. This group of acces-

Table 3. Distinguished tomato accessions by a complex of biochemical characteristics

No.	Accessions name	Total of sugars, %	Ascorbic acid, mg/100 g	Lycopene, mg/100 g	β-carotene, mg/100 g	Anthocyanins, mg/100 g
1	Superklusha	5.35 ± 0.32	16.12 ± 1.32	15.22 ± 2.34	1.07 ± 0.11	4.25 ± 0.25
2	Yaponskij tryufel' rozovyj	2.91 ± 0.97	34.58 ± 9.49	26.22 ± 9.29	0.86 ± 0.02	21.29 ± 8.18
3	Yaponskij tryufel' oranzhevyj	2.94 ± 0.81	35.56 ± 6.35	6.48 ± 2.94	1.14 ± 0.01	13.43 ± 2.35
4	Beduin	3.10 ± 0.07	24.88 ± 2.47	86.57 ± 15.32	1.10 ± 0.36	10.78 ± 1.88
5	OSU Blue	3.12 ± 0.41	24.05 ± 4.19	89.39 ± 14.28	1.29 ± 0.11	169.28 ± 94.94
6	Kraynij sever	1.85 ± 0.20	19.84 ± 2.81	29.01 ± 5.41	1.53 ± 0.59	11.39 ± 1.57
7	Novichok	4.36 ± 0.49	17.36 ± 2.59	2.82 ± 0.92	1.62 ± 0.14	7.65 ± 0.93
8	Ananas Noire	3.10 ± 0.78	23.56 ± 3.77	7.01 ± 1.04	0.35 ± 0.12	430.30 ± 98.35
9	Indigo Clackamas Blue Berry	2.48 ± 0.2	24.80 ± 1.75	9.61 ± 1.87	1.03 ± 0.18	588.86 ± 171.89
10	Chernyj mavr	2.43 ± 1.07	16.12 ± 2.56	19.15 ± 4.7	1.02 ± 0.13	4.17 ± 2.34
11	Viagra	4.42 ± 0.23	17.30 ± 1.84	45.67 ± 6.50	1.12 ± 0.11	7.80 ± 1.21
12	Percevidnyj rozovyj	2.98 ± 0.56	29.70 ± 9.37	69.24 ± 15.84	1.01 ± 0.64	3.37 ± 1.24
13	Bych'e serdce rozovoe	3.24 ± 0.94	16.24 ± 1.92	43.24 ± 10.75	1.06 ± 0.63	4.72 ± 1.32
14	Cherry rozovyj	3.75 ± 0.74	26.04 ± 3.23	34.15 ± 3.59	1.17 ± 0.13	3.79 ± 1.12
	Means	3.06 ± 0.10	20.78 ± 0.64	15.95 ± 2.60	0.68 ± 0.05	50.35 ± 11.94
	LSD	1.02	4.32	–	0.15	–

Fig. 5. Dendrogram of tomato accessions by basic biochemical parameters. UPGMA method.

The numbers indicate the size of the bootstrap; the names of the accessions are given in accordance with Table 1.

А.Б. Курина, А.Е. Соловьева
И.А. Храпалова, А.М. Артемьева

Биохимический состав плодов томата
различной окраски

2021
25•5

sions was also characterized by a high content of anthocyanins from 120.4 to 281.3 mg/100 g, and the accessions of the first and third subclusters had a high content of lycopene: 89.4, 16.4 and 11.6 mg/100 g, respectively.

The third cluster included tomato accessions with pink (Dikaya roza, Bych'e serdce rozovoe, Cherry rozovyj, Yaponskij tryufel' rozovyj), orange-red (Slivka krasnaya, Hybrid Budenovka × Chernyj princ), yellow-orange (Gold Medal), red (Krajnij sever) and red-brown (Viagra) fruit color. These accessions were characterized by a high content of total carotenoids – 44.96 ± 5.97 mg/100 g, of which there were 36.57 ± 6.45 mg/100 g of lycopene, 3.01 ± 1.37 mg/100 g of carotenes, and were low in anthocyanins.

The fourth cluster was the largest; it combined 41 accessions with different fruit colors and was divided into six subclusters. The first subcluster is represented by 16 accessions, mainly with red and orange color of the fruit, which were characterized by the content of carotenes – on average 3.03 ± 1.31 mg/100 g and lycopene – 6.35 ± 1.92 mg/100 g; this group also included several accessions with a high content of chlorophylls. The second subcluster also combined 16 accessions, but mainly with yellow and green-yellow fruit coloration and several with red and pink. These accessions were characterized by a low content of total carotenoids (6.28 ± 2.36 mg/100 g), including lycopene – an average of 3.78 ± 2.50 mg/100 g. The third subcluster is represented by three accessions with orange and orange-red fruit coloration. They were characterized by a high content of chlorophylls – 1.19–2.88 mg/100 g, anthocyanins – 18.41 ± 5.04 mg/100 g and total carotenoids – 17.04 ± 2.22 mg/100 g, of which the content of lycopene was 10.54 ± 0.15 mg/100 g, carotene – 3.56 ± 1.33 mg/100 g. The fourth subcluster is formed by two accessions – Indigo Blue Berries with a purple-red color and Hurma with a yellow-orange color of the fruit. The fifth subcluster is represented by three accessions: Dina, Chernyj mavr and Superklusha with a high content of total carotenoids (on average 25.71 ± 1.59 mg/100 g), of which the content of lycopene was 18.66 ± 3.23 mg/100 g and the content of carotenes was 3.28 ± 1.75 mg/100 g, of which β-carotene content was an average of 0.83 ± 0.38 mg/100 g, and a total sugar content of 3.91 % on average. The sixth subcluster included one accession of the Yaponskij tryufel' oranzhevyj, which is characterized by a high content of all carotenoids and low chlorophylls, as well as a high content of ascorbic acid and titratable acidity.

The fifth cluster included accession of wild tomato *L. peruvianum* (k-3960) and two accessions of cultivated tomato with yellow-purple (Indigo Gold Berries) and red with yellow stripes (Speckled Roman) fruit color. These accessions had an average content of anthocyanins in fruits – in the range of 53.3–87.9 mg/100 g.

The sixth cluster is represented by four accessions with orange-red (Beduin, Assorti, Majkl Pollan) and pink (Percevidnyj rozovyj) color of the fruit, which were characterized by a high content of lycopene – on average 71.90 ± ± 9.91 mg/100 g.

Thus, the accessions of the first two clusters were characterized by a high content of anthocyanins and chlorophylls in fruits, as well as ascorbic acid. The accessions of the second and fifth clusters were distinguished by a high dry matter content, while the accessions of the third and sixth clusters had a high content of total sugars, total carotenoids, with a predominance of lycopene and β-carotene. The fourth cluster united tomato accessions, on average, with a low content of carotenoids and anthocyanins, but a high content of carotenes. The accessions of the fifth cluster were characterized by an average content of anthocyanins and a low content of carotenoids.

Conclusion

As a result of this study, it was revealed that tomato accessions from the VIR collection with different fruit colors greatly differ in biochemical composition. We have determined the amplitude of variability of the main biochemical characteristics: dry matter, sugars, ascorbic acid, titratable acidity, pigments and anthocyanins. Correlations were revealed between the content of dry matter and monosaccharides ($r = 0.40$, $p \leq 0.05$), the total sugars ($r = 0.37$, $p \leq 0.05$) and ascorbic acid ($r = 0.32$, $p \leq 0.05$); the content of ascorbic acid and carotenoids ($r = 0.25$, $p \leq 0.05$). A high dependence of the content of chlorophyll *a* and *b* among themselves ($r = 0.89$, $p \leq 0.05$), as well as an average positive relationship between the content of chlorophyll *b* and anthocyanins ($r = 0.47$, $p \leq 0.05$), weak with the content of β-carotene ($r = 0.26$, $p \leq 0.05$) and weak negative with the content of monosaccharides ($r = -0.29$, $p \leq 0.05$) was demonstrated. There was also a moderate positive correlation between the content of chlorophyll *a* and anthocyanin ($r = 0.37$, $p \leq 0.05$).

It was revealed that accessions with red-brown color of fruits accumulate more dry matter. Accessions with green-purple, green-yellow, red-brown and purple-red coloration in total accumulate more chlorophylls in fruits – more than 2.40 mg/100 g, accessions with orange-red, red and yellow-orange coloration – within 1.10–1.58 mg/100 g. Tomato accessions characterized by a high content of carotene are those with red-brown (average 3.25 mg/100 g) and orange (4.03 mg/100 g) fruit color, whereas accessions with yellow, green-yellow, yellow-purple and green-purple color of fruits – by a low carotene content (less than 0.80 mg/100 g). A high content of β-carotene was found in tomato accessions with pink (average 0.89 mg/100 g) and orange-red (0.95 mg/100 g) fruit color, a lower content (0.81–0.82 mg/100 g) – in accessions with red-brown and orange fruit color.

It was determined that the differences in the content of lycopene between the accessions are very large, including the differences within the fruit color groups. A high content of lycopene was found in accessions with pink and orange-red color of fruits (on average 26.32–32.52 mg/100 g), accessions with green-yellow, yellow and yellow-purple color of the fruit accumulated it much lower – less than 6.5 mg/100 g. A large amount of anthocyanins was contained in tomato accessions with purple-red (32.89–588.86 mg/100 g) and yellow-purple (87.91–161.22 mg/100 g) fruit color, as well as in accessions of wild tomato (84.31–152.71 mg/100 g). Anthocyanins were also found in accessions with different color of fruits, but in much smaller quantities.

We have identified tomato accessions with a high content of both individual chemicals and a complex of traits that can be used as sources in breeding for a high content of dry matter, sugars, ascorbic acid, pigments and anthocyanins.

References

Anjum S., Hamid A., Ghafoor A., Tahira R., Shah S., Awan S.I., Ahmad Kh.Sh. Evaluation of biochemical potential in tomato (*Solanum lycopersicum*) germplasms. *Pak. J. Agric. Sci.* 2020;57(1):177-187. DOI 10.21162/PAKJAS/20.8140.

Bai Y., Lindhout P. Domestication and breeding of tomatoes: What have we gained and what can we gain in the future? *Ann. Bot.* 2007; 100(5):1085-1094. DOI 10.1093/aob/mcm150.

Beckles D.M., Hong N., Stamova L., Luengwilai K. Biochemical factors contributing to tomato fruit sugar content: a review. *Fruits.* 2012;67(1):49-64. DOI 10.1051/fruits/2011066.

Belova A.Yu., Murashev S.V., Verzhuk V.G. Influence of pigments in plants leaves on formation and properties of fruit production. *Nauchnyy Zhurnal NIU ITMO. Seriya Protsessy i Apparaty Pishchevykh Proizvodstv = Scientific Journal NRU ITMO. Series Processes and Food Production Equipment.* 2012;1(13):13. (in Russian)

Bhattarai K., Sharma S., Panthee D.R. Diversity among modern tomato genotypes at different levels in fresh-market breeding. *Int. J. Agron.* 2018;2018:1-15. DOI 10.1155/2018/4170432.

Campbell J.K., Canene-Adams K., Lindshield B.L., Boileau T.W.-M., Clinton S.K., Erdman J.W. Jr. Tomato phytochemicals and prostate cancer risk. *J. Nutr.* 2004;134(12):3486-3492. DOI 10.1093/jn/134. 12.3486S.

Chandra H.M., Ramalingam S. Antioxidant potentials of skin, pulp, and seed fractions of commercially important tomato cultivars. *Food Sci. Biotechnol.* 2011;20(1):15-21. DOI 10.1007/s10068-011-0003-z.

Da Silva-Souza M.A., Peres L.E.P., Freschi J.R., Purgatto E., Lajolo F.M., Hassimotto N.M.A. Changes in flavonoid and carotenoid profiles alter volatile organic compounds in purple and orange cherry tomatoes obtained by allele introgression. *J. Sci. Food Agric.* 2020;100(4):1662-1670. DOI 10.1002/jsfa.10180.

Dar R.A., Sharma J.P. Genetic variability studies of yield and quality traits in tomato (*Solanum lycopersicum* L.). *Int. J. Plant Breed. Genet.* 2011;5(2):168-174. DOI 10.3923/ijpbg.2011.168.174.

Descriptors Tomato (Lycopersicon spp.) IPGRI (International Plant Genetic Resources Institute). Rome. Italy, 1996.

Ermakov A.I., Arasimovich V.V., Yarosh N.P. Biochemical Methods in Plant Studies. Leningrad: Agropromizdat Publ., 1987. (in Russian)

FAOSTAT, 2019. Available at: http://www.fao.org/faostat/en/#home.

Friedman M. Anticarcinogenic, cardioprotective, and other health benefits of tomato compounds lycopene, α-tomatine, and tomatidine in pure form and in fresh and processed tomatoes. *J. Agric. Food Chem.* 2013;61(40):9534-9550. DOI 10.1021/jf402654e.

Gascuel Q., Diretto G., Monforte A.J., Fortes A.M., Granell A. Use of natural diversity and biotechnology to increase the quality and nutritional content of tomato and grape. *Front. Plant Sci.* 2017;8:652. DOI 10.3389/fpls.2017.00652.

Golubkina N.A., Molchanova A.V., Tareeva M.M., Baback O.G., Nekrashevich N.A., Kondratyeva I.Yu. Quantitative thing layer chromatography for evaluation of carotenoid composition of tomatoes *Solanum licopersicum. Ovoshchi Rossii = Vegetable Crops of Russia.* 2017;5:96-99. DOI 10.18619/2072-9146-2017-5-96-99. (in Russian)

Gupta A., Kawatra A., Sehgal S. Physical-chemical properties and nutritional evaluation of newly developed tomato genotypes. *Afr. J. Food Sci. Technol.* 2011;2(7):167-172.

Hammer Ø., Harper D.A.T., Ryan P.D. PAST: paleontological statistics software package for education and data analysis. *Palaeontol. Electron.* 2001;4(1):1-9.

Harish M.Ch., Shanmugaraj B.M., Balamurugan S., Sathishkumar R. Influence of genotypic variations on antioxidant properties in diffe-

rent fractions of tomato. *J. Food Sci.* 2012;77(11):1174-1178. DOI 10.1111/j.1750-3841.2012.02962.x.

Ignatova S.I., Babak O.G., Bagirova S.F. Development of high-lycopene tomato hybrids using conventional breeding techniques and molecular markers. *Ovoshchi Rossii = Vegetable Crops of Russia.* 2020;5:22-28. DOI 10.18619/2072-9146-2020-5-22-28. (in Russian)

Ilić Z., Aharon Z., Perzelan Y., Alkalai-Tuvia S., Fallik E. Lipophilic and hydrophilic antioxidant activity of tomato fruit during postharvest storage on different temperatures. *Acta Hortic.* 2009;830(91):627-634. DOI 10.17660/ActaHortic.2009.830.91.

International CMEA Classifier of the Genus *Lycopersicon* Tourn. Leningrad, 1986. (in Russian)

Jones C.M., Mes P., Myers J.R. Characterization and inheritance of the *Anthocyanin fruit* (*Aft*) tomato. *J. Hered.* 2003;94(6):449-456. DOI 10.1093/jhered/esg093.

Kendrick R.E., Kerckhoffs L.H.J., Van Tuinen A., Koornneef M. Photomorphogenic mutants of tomato. *Plant Cell Environ.* 1997;20: 746-751.

Khachik F., Carvalho L., Bernstein P.S., Muir G.J., Zhao D.-Y., Katz N.B. Chemistry, distribution, and metabolism of tomato carotenoids and their impact on human health. *Exp. Biol. Med.* 2002; 227(10):845-851. DOI 10.1177/153537020222701002.

Kondratyeva I.Yu., Engalychev M.R. Orange-fruited tomato varieties with high taste and preventive-therapeutic properties. *Izvestiya Federalnogo Nauchnogo Tsentra Ovoshchevodstva = News of FSVC.* 2019;2:71-78. DOI 10.18619/2658-4832-2019-2-71-78. (in Russian)

Kondratyeva I.Yu., Golubkina N.A. Lycopene and β-carotene in tomato. *Ovoshchi Rossii = Vegetable Crops of Russia.* 2016;4:80-83. DOI 10.18619/2072-9146-2016-4-80-83. (in Russian)

Kondratyeva I.Yu., Pavlov L.V. Dry components concentration in tomato fruits in dependence of qualitative and quantitative parameters. *Kartofel i Ovoshchi = Potato and Vegetables.* 2009;5:21. (in Russian)

Kuzyomensky A.V. Breeding and Genetic Studies of Mutant Forms of Tomato. Kharkov, 2004. (in Russian)

Leiva-Brondo M., Valcárcel M., Cortés-Olmos C., Roselló S., Cebolla-Cornejo J., Nuez F. Exploring alternative germplasm for the development of stable high vitamin C content in tomato varieties. *Sci. Hortic.* 2012;133:84-88. DOI 10.1016/J.SCIENTA.2011.10.013.

Li H., Deng Z., Liu R., Young J.C., Zhu H., Loewen S., Tsao R. Characterization of phytochemicals and antioxidant activities of a purple tomato (*Solanum lycopersicum* L.). *J. Agric. Food Chem.* 2011; 59(21):11803-11811. DOI 10.1021/jf202364v.

Lim W., Miller R., Park J., Park S. Consumer sensory analysis of high flavonoid transgenic tomatoes. *J. Food Sci.* 2014;79(6):1212-1217. DOI 10.1111/1750-3841.12478.

Martí R., Roselló S., Cebolla-Cornejo J. Tomato as a source of carotenoids and polyphenols targeted to cancer prevention. *Cancers (Basel).* 2016;8(6):58. DOI 10.3390/cancers8060058.

Mes P.J., Boches P., Myers J.R., Durst R. Characterization of tomatoes expressing anthocyanin in the fruit. *J. Am. Soc. Hortic. Sci.* 2008; 133(2):262-269. DOI 10.21273/JASHS.133.2.262.

Mozos I., Stoian D., Caraba A., Malainer C., Horbanczuk J.O., Atanasov A.G. Lycopene and vascular health. *Front. Pharmacol.* 2018;9: 521. DOI 10.3389/fphar.2018.00521.

Nour R.V., Trandafir I., Ionica M.E. Antioxidant compounds, mineral content and antioxidant activity of several tomato cultivars grown in Southwestern. *Not. Bot. Horti Agrobot. Cluj-Napoca.* 2013;41(1): 136-142. DOI 10.15835/nbha4119026.

Ooe E., Ogawa K., Horiuchi T., Tada H., Murase H., Tsuruma K., Shimazawa M., Hara H. Analysis and characterization of anthocyanins and carotenoids in Japanese blue tomato. *Biosci. Biotechnol. Biochem.* 2016;80(2):341-349. DOI 10.1080/09168451.2015.1091715.

Owusu J., Ma H., Wang Z., Amissah A. Effect of drying methods on physicochemical properties of pretreated tomato (*Lycopersicon esculentum* Mill.) slices. *Croat. J. Food Technol. Biotechnol. Nutr.* 2012;7(1-2):106-111.

А.Б. Курина, А.Е. Соловьева
И.А. Храпалова, А.М. Артемьева

Биохимический состав плодов томата
различной окраски

2021
25•5

Pal R.S., Hedau N.K., Lakshmi Kant, Pattanayak A. Functional quality and antioxidant properties of tomato genotypes for breeding better quality varieties. *Electron. J. Plant Breed.* 2018;9(1):1-8. DOI 10.5958/0975-928X.2018.00001.7.

Peralta I.E., Spooner D.M. History, origin and early cultivation of tomato (Solanaceae). In: Razdan M.K., Mattoo A.K. (Eds.) Genetic Improvement of Solanaceous. Enfield, USA: Science Publ., 2007. DOI 10.1201/b10744-2.

Peter J.M., Peter B., James R.M. Characterization of tomatoes expressing anthocyanin in the fruit. *J. Amer. Soc. Hort. Sci.* 2008;133:262-269.

Raiola A., Tenore G.C., Barone A., Frusciante L., Rigano M.M. Vitamin E content and composition in tomato fruits: Beneficial roles and bio-fortification. *Int. J. Mol. Sci.* 2015;16(12):29250-29264. DOI 10.3390/ijms161226163.

Rick C.M. Tomato. *Genet. Coop. Rept.* 1959;9:41-42.

Roohanitaziani R., de Maagd R.A., Lammers M., Molthoff J., Meijer-Dekens R., van Kaauwen M.P.W., Finkers H.T., Tikunov Yu., Visser R.G.F., Bovy A.G. Exploration of a resequenced tomato core collection for phenotypic and genotypic variation in plant growth and fruit quality traits. *Genes.* 2020;11(11):1278. DOI 10.3390/genes11111278.

Scarano A., Butelli E., de Santis S., Cavalcanti E., Hill L., de Angelis M., Giovinazzo G., Chieppa M., Martin C., Santino A. Combined dietary anthocyanins, flavonols, and stilbenoids alleviate inflammatory bowel disease symptoms in mice. *Front. Nutr.* 2018;4:75. DOI 10.3389/fnut.2017.00075.

Stommel J.R. Enzymatic components of sucrose accumulation in the wild tomato species *Lycopersicon peruvianum*. *Plant Physiol*. 1992; 99(1):324-328. DOI 10.1104/pp.99.1.324.

Tanksley S.D. The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. *Plant Cell*. 2004;16(Suppl.): S181-S189. DOI 10.1105/tpc.018119.

Tomato – UPOV (*Solanum lycopersicum* L.). 0007 TG/44/11 Rev. Geneva, 2012.

Viuda-Martos M., Sanchez-Zapata E., Sayas-Barberá E., Sendra E., Pérez-Álvarez J.A., Fernández-López J. Tomato and tomato by-products. Human health benefits of lycopene and its application to meat products: a review. *Crit. Rev. Food Sci. Nutr.* 2014;54(8):1032-1049. DOI 10.1080/10408398.2011.623799.

Wang D., Seymour G.B. Tomato flavor: Lost and found? *Mol. Plant*. 2017;10(6):782-784. DOI 10.1016/j.molp.2017.04.010.

Wei M.Y., Giovannucci E.L. Lycopene, tomato products, and prostate cancer incidence: a review and reassessment in the PSA screening era. *J. Oncol*. 2012;2012:271063. DOI 10.1155/2012/271063.

Zanfini A., Corbini G., Rosa C.L., Dreassi E. Antioxidant activity of tomato lipophilic extracts and interactions between carotenoids and α-tocopherol in synthetic mixtures. *Food Sci. Technol*. 2010;1(43): 67-72. DOI 10.1016/j.lwt.2009.06.011.

Zhang Y., Butelli E., Alseekh S., Tohge T., Rallapalli G., Luo J., Kawar P.G., Hill L., Santino A., Fernie A.R., Martin C. Multi-level engineering facilitates the production of phenylpropanoid compounds in tomato. *Nat. Commun*. 2015;6:8635. DOI 10.1038/ncomms9635.

ORCID ID
A.B. Kurina orcid.org/0000-0002-3197-4751
A.E. Solovieva orcid.org/0000-0002-6201-4294
A.M. Artemyeva orcid.org/0000-0002-6551-5203

Powdery mildew resistance of barley accessions from Dagestan

R.A. Abdullaev[1], T.V. Lebedeva[1], N.V. Alpatieva[1], B.A. Batasheva[2], I.N. Anisimova[1] ✉, E.E. Radchenko[1]

[1] Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), St. Petersburg, Russia
[2] Dagestan Experiment Station of VIR, N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), Vavilovo Village, Derbent, Dagestan, Russia
✉ irina_anisimova@inbox.ru

Abstract. Powdery mildew caused by the parasitic fungus *Blumeria graminis* (DC.) Golovin ex Speer f. sp. *hordei* Marchal is one of the most common diseases of barley. Growing resistant varieties can significantly minimize harmful effects of the pathogen. The specificity in the interaction between the fungus and its host plant requires a continuous search for new donors of the resistance trait. The powdery mildew resistance of 264 barley accessions from Dagestan and genetic control of the trait in resistant forms were studied under field and laboratory conditions. Forty-seven barley lines carrying previously identified powdery mildew resistance genes were also examined. During three years, the experimental material was evaluated under severe infection pressure at the Dagestan Experiment Station of VIR (North Caucasus, Derbent). Juvenile resistance against the Northwest (St. Petersburg, Pushkin) pathogen population was evaluated in a climatic chamber. The genetic control of *B. graminis* resistance in the selected accessions was studied with the application of hybridological and molecular analyses. The level of genetic diversity of Dagestan barley for effective resistance to powdery mildew is very low. Only two accessions, VIR-23787 and VIR-28212, are resistant against *B. graminis* at both seedling and adult plant stages. The high-level resistance of breeding line VIR-28212 originating from barley landrace VIR-17554 (Ep-80 Abyssinien) from Ethiopia is controlled by the recessive gene *mlo11*. Accession VIR-17554 is heterogeneous for the studied trait, with the powdery mildew resistant genotypes belonging to two varieties, *dupliatrum* (an awnless phenotype) and *nigrinudum* (an awned phenotype). In accession VIR-23787, a recessive resistance gene distinct from the *mlo11* allele was identified. This accession is supposed to be protected by a new, effective pathogen resistance gene.
Key words: barley landraces; *Blumeria graminis* f. sp. *hordei*; durable resistance; genes.

Устойчивость образцов ячменя из Дагестана к мучнистой росе

Р.А. Абдуллаев[1], Т.В. Лебедева[1], Н.В. Алпатьева[1], Б.А. Баташева[2], И.Н. Анисимова[1] ✉, Е.Е. Радченко[1]

[1] Федеральный исследовательский центр Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова (ВИР),
Санкт-Петербург, Россия
[2] Федеральный исследовательский центр Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова (ВИР),
Дагестанская опытная станция – филиал ВИР, с. Вавилово, Дербент, Республика Дагестан, Россия
✉ irina_anisimova@inbox.ru

Аннотация. Мучнистая роса (возбудитель *Blumeria graminis* (DC.) Golovin ex Speer f. sp. *hordei* Marchal) – одно из самых распространенных заболеваний ячменя. Возделывание устойчивых сортов может существенно ограничить вредоносность патогена. Специфичность взаимодействия гриба с растением-хозяином обусловливает необходимость постоянного поиска новых доноров устойчивости. Носителями эффективных генов устойчивости к патогену часто являются образцы местного ячменя. В полевых и лабораторных экспериментах оценили потенциал изменчивости 264 образцов ячменя из Дагестана по устойчивости к *B. graminis* и изучили генетический контроль признака у выделившихся форм. Исследовали также 47 линий ячменя, несущих ранее идентифицированные гены устойчивости к мучнистой росе. В течение трех лет с использованием балловой шкалы экспериментальный материал оценивали на жестком инфекционном фоне в условиях Дагестанской опытной станции Всероссийского института генетических ресурсов растений им. Н.И. Вавилова (Северный Кавказ, г. Дербент), а в климатической камере исследовали ювенильную устойчивость растений к северо-западной (Санкт-Петербург, г. Пушкин) популяции гриба. Генетический контроль устойчивости к *B. graminis* у выделенных образцов изучали с помощью гибридологического и молекулярного анализов. Установили, что генетическое разнообразие ячменей Дагестана по эффективной устойчивости к мучнистой росе весьма невелико. Лишь образцы к-23787 и к-28212 устойчивы к дагестанской популяции *B. graminis* в фазе цветения и к петербургской популяции патогена – в фазе двух листьев. Две географически очень разобщенные популяции гриба оказались идентичны по вирулентности к линиям с ранее идентифицированными генами устойчивости к мучнистой росе. Высокая устойчивость селек-

Р.А. Абдуллаев, Т.В. Лебедева, Н.В. Алпатьева
Б.А. Баташева, И.Н. Анисимова, Е.Е. Радченко

Устойчивость образцов ячменя из Дагестана
к мучнистой росе

2021
25•5

ционной линии к-28212, полученной с использованием образца местного ячменя из Эфиопии к-17554 (Ep-80 Abyssinien), контролируется рецессивным геном *mlo11*. Образец к-17554 гетерогенен по изученному признаку, при этом устойчивые к мучнистой росе генотипы относятся к двум разновидностям: *dupliatrum* (остистый фенотип) и *nigrinudum* (безостый фенотип). У образца к-23787 обнаружен рецессивный ген устойчивости, отличающийся от аллеля *mlo11*. Предполагается, что этот образец защищен новым эффективным геном устойчивости к патогену.

Ключевые слова: образцы местного ячменя; *Blumeria graminis* f. sp. *hordei*; длительная устойчивость; гены.

Introduction

Powdery mildew caused by the parasitic fungus *Blumeria graminis* (DC.) Golovin ex Speer f. sp. *hordei* Marchal) is one of the most common and harmful diseases of barley. The pathogen predominantly affects leaves, leaf sheath, and stems throughout the growing season. In the infected plants, photosynthetic activity of leaves is being decreased while water loss and respiration intensity are being increased, which results in growth retardation, reduced tillering ability, and decreasing seed mass and grain number per spike. Yield reduction caused by powdery mildew can reach 30 %, with an average of 5–10 % across all regions (Balkema-Boomstra, Masterbroek, 1995; Gong et al., 2013).

Selection of resistant plant genotypes is a radical and environment friendly way to combat the disease. Unfortunately, the pathogen is characterized by differential interaction with the host plant genotype. This means that the ubiquitously observed genetic uniformity of cultivated varieties creates conditions for adaptive microevolution of the fungus.

In barley, numerous powdery mildew resistance genes designated by various symbols have been identified, most of them are alleles at the loci *Mla* and *Mlo*. Thus, 39 alleles at the *Mla* (*Mildew resistance locus a*) (chromosome 1H) and 44 alleles at the *Mlo* (*Mildew resistance locus o*) locus (chromosome 4H) are known (Jørgensen, 1994; Seeholzer, 2009; Reinstädler et al., 2010; Kusch, Panstruga, 2017). However, most alleles are ineffective against the causative agent. The allele *mlo11* is practically the only effective gene that confers durable resistance to the pathogen. Barley landraces often possess effective genes for resistance against phytopathogens. For example, a study of 925 Ethiopian barley accessions has revealed 15 accessions harboring the *mlo11* allele, and 59 forms whose resistance to *B. graminis* was controlled by effective genes distinct from the *mlo11* (Abdullaev et al., 2019).

Since recently, the attention of researchers is drawn to Dagestan, a region of ancient agriculture. In a small area, very contrasting soil, climatic and landscape conditions are present: from the low-lying Caspian basins to high mountains with eternal snows, from the semi-desert and desert sharply continental northern dry-steppe zones to the regions of the subtropical type. The results of the study of US Barley Collection (USDA National Small Grains Collection) accessions for resistance against several pathogens and insects have shown that the Caucasus could be considered as the "center of concentration" of yellow rust resistant barley forms (the causative agent *Puccinia striiformis* Westend. f. sp. *hordei*). In these studies, three accessions from Dagestan were highly resistant to yellow rust and net blotch (*Pyrenophora teres* (Died.) Drechsl.), as well as moderately resistant to spot blotch (*Cochliobolus sativus* (Ito et Kurib.) Drechsler ex Dastur.) and viral diseases (Bonman

et al., 2005). Powdery mildew resistance of Dagestan barleys has not yet been studied.

This study is aimed at evaluating variability potential of barley accessions from Dagestan for resistance to *B. graminis*, and at elucidating genetic control of the trait in the selected resistant forms.

Material and methods

The material used in the study included 264 barley accessions (187 spring, 76 winter, and one facultative) from Dagestan (Supplementary 1)[1], among them the landraces (228 accessions) prevailed whereas only 36 accessions represented cultivars and breeding lines. The studied forms belong to the two subspecies: six-row barley (subsp. *vulgare*) and two-row barley (subsp. *distichon*) and represent 29 botanical varieties. Some accessions were registered in the collection of the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR) as populations involving up to five varieties. Forty-seven barley lines carrying previously identified powdery mildew resistance genes were also studied (Supplementary 2).

Field trials were carried out at the Dagestan Experiment Station of VIR (DES VIR, Derbent; latitude 41°59′03.3″N, longitude 48°19′47.7″E) in 2012–2014. Accessions were sown in the third decade of October in the field plots of 1 m² area, with a row to row spacing of 15 cm and row length of 1 m. Spring barley cultivar Temp (VIR-22055, Krasnodar region) sown after each 20 accessions was used as a control. Resistance to disease was evaluated at the heading stage and at the milk-ripe stage and expressed as infection type (IT) in the following scale (Loskutov et al., 2012):

IT 1 – resistance is very low – pustules cover all leaves and internodes in abundance, including the upper ones; a lesion can capture an ear;

IT 3 – low resistance – pustules in bulk develop mainly on the lower leaves and internodes; individual scattered pustules are observed on the upper leaves;

IT 5 – medium resistance – a moderate number of pustules on the lower leaves and internodes;

IT 7 – high resistance – single small pustules on the lower leaves and internodes, pustules can be more numerous, but small;

IT 9 – very high resistance – no pustules are visible.

To exclude the presence of known powdery mildew resistance genes in barley accessions from Dagestan, the seedling resistance test was performed. Fungal inoculum was propagated on plants grown in a Barnstead 845-2 climatic chamber at a 12-hour photoperiod and a temperature of 16 °C (day), 13 °C (night).

---

[1] Supplementary materials 1–2 are available at
http://www.bionet.nsc.ru/vogis/download/pict-2021-25/appx10.pdf

The Northwest (St. Petersburg, Pushkin) population of the fungus was used for inoculation. The population was collected from a susceptible barley cultivar Belogorsky (VIR-22089, Leningrad region). Twenty seeds of each of resistant accessions and 47 barley lines carrying previously identified powdery mildew resistance genes were sown on water-moistened cotton in plastic trays, placed in a climatic chamber, and after a week the seedlings were inoculated through shaking conidia on them from plants strongly affected by powdery mildew. Infection types were scored using 0–4 point scale (Mains, Dietz, 1930) as follows:

IT 0 – highly resistant, no mycelium evident. Chlorotic or necrotic spots may be developed by some varieties;

IT 1 – very resistant, slight to moderate mycelial development, but with little or no sporulation. Chlorotic or necrotic spots may develop in some varieties;

IT 2 – moderately resistant, moderate mycelial development, accompanied by limited sporulation. Chlorotic or necrotic areas may be formed;

IT 3 – moderately susceptible. Moderate to abundant mycelial development, accompanied by moderate sporulation;

IT 4 – very susceptible. Abundant mycelial development, accompanied by abundant sporulation.

To specify resistance genes, we have estimated segregation ratios in the F_2 hybrid populations obtained from crossing resistant accessions with a susceptible variety (VIR-15033). For allelic testing, the resistant accessions were crossed with each other, as well as with near-isogenic line Ingrid *mlo11*. The resistance tests were conducted in a climatic chamber.

The germinated seeds were sown on water-moistened cotton in plastic trays. Each tray contained one row of each parental form (P_1, P_2) and F_1 hybrid, and 7–8 rows of F_2 plants. At the two-leaf stage, the seedlings were inoculated with the Northwest pathogen population (collected near St. Petersburg). Hybrids from the crossings of resistant accessions and a susceptible tester were evaluated for resistance at the time of death of susceptible parental form. The assessment of allelic relations among powdery mildew resistance genes was carried out at the time of death of susceptible control (cultivar Belogorsky) which was sown along with F_2 plants in the same tray. Plants exhibiting infection type similar to that of either the susceptible parent or the control (IT scores of 3 or 4 according to the scale of E.B. Mains and S.M. Dietz (1930) were classified as homozygous susceptible (S). The resistant class (R) comprised plants similar in infection type to the resistant parental form (IT scores of 0–1).

For identifying the *mlo11* gene, the PCR markers (Table 1) developed by P. Piffanelli et al. (2004) were used. Total DNA was isolated from 7-day-old seedlings according to the method of D.B. Dorokhov and E. Kloke (1997). Amplification was carried out in a 25 µl volume reaction mixture containing 50–100 ng of genomic DNA, 1×reaction buffer, 2 mM $MgCl_2$, 0.25 mM dNTPs, 0.25 µM of each primer, 1 U *Taq* DNA polymerase (Dialat Ltd). PCR was conducted on a MyCycler thermal cycler (Bio-Rad, USA). The protocol consisted of an initial cycle of denaturation at 94 °C for 5 min, followed by 35 cycles (94 °C for 30 s, 60 °C for 30 s and 72 °C for 1.5 min). Final extension was done at 72 °C for 10 min. Amplification products were analyzed by electrophoresis on 1.5 %

Table 1. Primers used in the study (Piffanelli et al., 2004)

Primer	Sequence (5′–3′)
ADUP7	CTCAAGCTTGCCACCATGTCGGACAAAAAGGG
Mlo6	CATCTACTACTAGCATGTACC
Mlo10	GTCCTGCCACCTAAGTAGCAG

Table 2. Barley accessions exhibiting differences in powdery mildew resistance (DES VIR, 2012–2014)

VIR catalogue number	Accession	Infection type*		
		2012	2013	2014
23787	Landrace	7	7	9
25615	Dagestan × VIR-17348	7	7	5
28211	Mutant KNIISH × PF-24 × Ketzoraz Beta	7	7	5
28212	KNIISH 11/1 × Krasnyj Dar × VIR-17554	7	7	7
30781	Dagestanskij zolotistyj	7	7	5
22055	Temp (susceptible control)	3	3	3

* IT according to the 0–9 point scale (Loskutov et al., 2012).

agarose gels and visualized under ultraviolet light. Fragment size was estimated using FastRuler™ SM1113 DNA-marker (Fermentas).

Results

In 2012–2014, the epiphytotic disease development was observed at the DES VIR: the susceptible control cultivar Temp exhibited the IT score of 3 according to the 0–9 point scale (Loskutov et al., 2012). At a severe infection pressure, five accessions whose IT scores did not exceed 7 points have been initially isolated; in 2014 only two accessions (VIR-23787 and VIR-28212) were resistant (Table 2).

Forty-seven barley lines carrying previously identified powdery mildew resistance genes were evaluated under laboratory conditions and in the field experiments at the DES VIR. Twelve lines have been shown to be resistant to the pathogen in the field (Table 3). The same lines were resistant to the Northwest population of *B. graminis*, i. e. there were no significant differences in virulence between the two pathogen populations.

Two selected accessions, VIR-23787 and VIR-28212, exhibited resistance to the pathogen in the field and laboratory experiments. According to the information documented at the VIR Department of Oat, Rye and Barley Genetic Resources, the landrace accession VIR-23787 is of unknown pedigree, and line VIR-28212 originated from Ethiopian spring barley accession VIR-17554.

Thus, a comparative analysis of infection types of powdery mildew resistant accessions VIR-23787 and VIR-28212, and lines resistant to the Dagestan and St. Petersburg populations of *B. graminis* (see Table 3) has shown that these accessions might be protected either by the recessive gene(s) *mlo* or dominant *Mla*.

Р.А. Абдуллаев, Т.В. Лебедева, Н.В. Алпатьева
Б.А. Баташева, И.Н. Анисимова, Е.Е. Радченко

Устойчивость образцов ячменя из Дагестана
к мучнистой росе

2021
25•5

Table 3. Barley accessions carrying effective genes for resistance against *B. graminis*

VIR catalogue number	Accession	Origin	Infection type	
			Field trials**	Laboratory experiments***
30218	IL* Ingrid *mlo1*	Germany	7	1
30219	IL Ingrid *mlo3*		7	0
30220	IL Ingrid *mlo4*		7	0
30221	IL Ingrid *mlo5*		7	0
30222	IL Ingrid *mlo8*		7	0
30223	IL Ingrid *mlo9*		7	0
30224	IL Ingrid *mlo10*		7	0
30225	IL Ingrid *mlo11*		7	0
30237	Diamant × IB-54B *Mla16*		7	0
30238	Diamant × IB-86B *Mla19*		7	1
31007	Line *Mla18*	Israel	7	1
31011	Line *Mli*		7	1
23787	Landrace	Dagestan	9	1
28212	KNIISKH11/1 × Krasny Dar × VIR-17554		7	1
22055	Temp (control)	Russia	3	–
22089	Belogorsky (control)		–	4

* IL – isogenic line; ** IT according to the 0–9 point scale (Loskutov et al., 2012); *** IT according to the 0–4 point scale (Mains, Dietz, 1930).

Table 4. Segregation for resistance to *B. graminis* of the F$_2$ hybrids obtained from crossing resistant barley accessions and susceptible testers

Cross combination	Total number of plants	The ratio of phenotypes (R:S)		χ^2	*p*
		observed	expected		
VIR-15033 × VIR-23787	244	51:193	1:3	2.48	0.10–0.20
VIR-15033 × VIR-28212	935	244:691	1:3	0.60	0.25–0.50
VIR-28212 × VIR-15033	223	61:162	1:3	0.65	0.25–0.50
VIR-17554 × VIR-22089	433	93:340	1:3	2.87	0.05–0.10

Note. $\chi^2_{0.05} = 3.84$.

To determine the number of genes conferring resistance, the F$_1$ and F$_2$ hybrids obtained from crossing resistant accessions VIR-23787 and VIR-28212 with susceptible tester VIR-15033 were evaluated for resistance at the seedling stage. The parental genotypes VIR-23787 and VIR-28212 were resistant to the Northwest pathogen population with the IT score of 1, whereas F$_1$ hybrid plants were susceptible to the pathogen (IT score of 4). Phenotypic segregation in the F$_2$ populations fitted the expected ratio 1R: 3S (Table 4). We suggest that accessions VIR-23787 and VIR-28212 possess recessive genes conferring powdery mildew resistance at the seedling stage. Analysis of F$_2$ hybrid progenies derived from reciprocal crosses between the resistant (VIR-28212) and susceptible (VIR-15033) accessions did not reveal an effect of the maternal or paternal genotype on the segregation pattern ($\chi^2_{1:3} = 0.60$ and $\chi^2_{1:3} = 0.65$).

We have also analyzed segregation for pathogen resistance of the F$_2$ hybrids obtained from crossing accession VIR-17554 from Ethiopia (a putative donor of disease resis-

tance in the line VIR-28212) and susceptible cultivar Belogorsky (VIR-22089). The results indicated that accession VIR-17554 also carries a single recessive gene for *B. graminis* resistance (see Table 4).

The segregation pattern in the F$_2$ generation allows us to suggest that *B. graminis* resistance of accessions VIR-28212 and VIR-17554 is controlled by the gene *mlo11*. There was no segregation in the F$_2$ hybrid progenies from the crossings of these accessions with the *mlo11* allele carrier VIR-30225 (IL Ingrid *mlo11*) (Table 5). This means that accessions VIR-28212, VIR-17554 and VIR-30225 are protected by the *mlo11* gene.

Segregation in the F$_2$ hybrid population from a cross between IL Ingrid *mlo11* and accession VIR-23787 fitted the expected ratio (7R: 9S; $\chi^2_{7:9} = 0.00$) for two recessive resistance genes.

Accessions VIR-28212, VIR-17554 and VIR-23787 were analyzed with the use of the *mlo11* specific molecular markers (see Table 1). Amplification of specific 1200 and 440 bp

Table 5. Segregation for resistance to *B. graminis* of the F_2 hybrids
obtained from crossing resistant barley accessions with IL Ingrid *mlo11* (VIR-30225)

Cross combination	Total number of plants	The ratio of phenotypes (R:S)		χ^2	*p*
		observed	expected		
VIR-30225 × VIR-28212	363	363:0	–	–	–
VIR-30225 × VIR-17554	203	203:0	–	–	–
VIR-30225 × VIR-23787	582	254:328	7:9	0.00	0.995
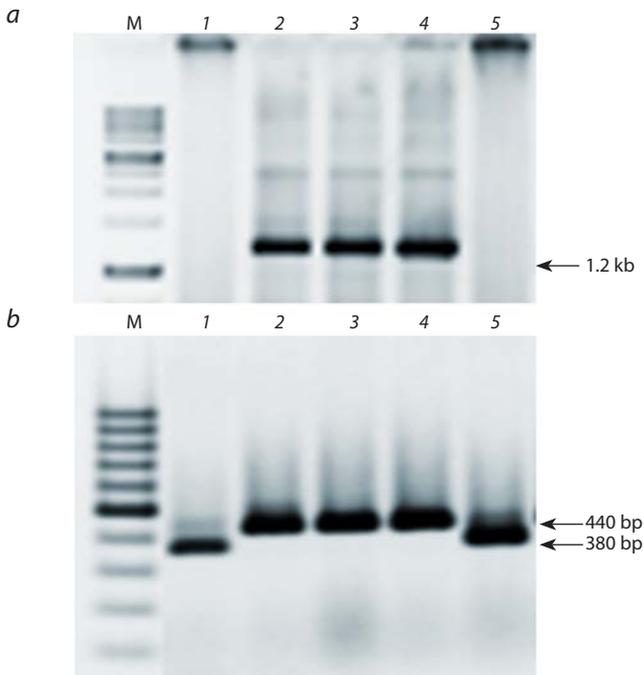
Molecular marker-based assessment of powdery mildew resistant barley plants with the use of primer pairs ADUP7-Mlo6 (*a*) and Mlo6-Mlo10 (*b*).

a – the 1.2 kb fragment is specific for the *mlo11* allele; *b* – the 380 bp fragment is amplified in the genotypes with the wild-type *Mlo* allele and a fragment of approx. 440 bp size is amplified in the genotypes carrying the *mlo11* allele. *1* – susceptible control VIR-22089 (cultivar Belogorsky); *2* – line VIR-30225 with the *mlo11* allele; *3* – VIR-28212; *4* – VIR-17554; *5* – VIR-23787. M – molecular weight marker.

fragments with primer pairs ADUP7-Mlo6 and Mlo6-Mlo10 respectively is usually an indication of the presence of the *mlo11* gene. Genotypes carrying the *mlo11* allele were found in accessions VIR-17554 and VIR-28212 (see the Figure). No carriers of the *mlo11* allele were detected among the 20 plants analyzed in accession VIR-23787, therefore this form is protected by another resistance gene.

Discussion

The results of the study indicate a rather low level of genetic diversity for powdery mildew resistance within the studied set of Dagestan barley accessions. Five accessions (VIR-23787, VIR-25615, VIR-28211, VIR-28212, VIR-30781) were selected for powdery mildew resistance in 2012 and 2013. However, in the past years their susceptibility is enhanced, which may be due to changes of the pathogen population structure. Only two accessions, VIR-23787 and VIR-28212, were resistant to

the Dagestan population of *B. graminis* at the adult stage and to the St. Petersburg population of the pathogen at the seedling stage respectively. It is interesting that two geographically very distant populations of the fungus turned out to be identical in virulence to the tester lines (see Table 3).

It was somewhat surprising, that accession VIR-28212 was protected by the effective resistance gene *mlo11*, which was introgressed from VIR-17554 (Ep-80 Abyssinien). This accession, which has entered the VIR collection from the German Gene Bank in 1949, turned out to be heterogeneous for resistance (eight resistant plants out of the ten studied). Plants belonging to the variety *nigrinudum* are most likely a result of genetic contamination in consequence of the cross-pollination, which is quite typical for Ethiopian barley.

Durable non-specific resistance of barley to *B. graminis* is associated with mutations at the *Mlo* locus at the long arm of chromosome 4 (Jørgensen, 1992). The resistance of *mlo* mutants is associated with physiological processes which prevent successful penetration of the pathogen into epidermal cells of the host plant (Ge et al., 2016). Carriers of the recessive *mlo* allele are characterized by leaf damage which is considered as a manifestation of premature cell death symptoms after cell wall appositions (callose deposits on adult leaves), observed even in the absence of the pathogen (Skou et al., 1984). Despite a number of limitations associated with negative pleiotropic effects leading to yield decrease, the use of *mlo* alleles (mostly *mlo11* and, in part, *mlo9*) in barley breeding programs provided durable protection against *B. graminis* in the regions with temperate humid climate. Currently, 75 % of modern spring barley varieties in Europe are protected by the *mlo* genes (Dreiseitl, 2017).

The recessive *B. graminis* resistance gene in accession VIR-23787 is distinct from the *mlo11* and not associated with negative pleiotropic effects which are characteristic of the *mlo* alleles induced by chemical mutagenesis. In addition, unlike the accessions carrying recessive *mlo* alleles, no symptoms of fungal damage were detected on plants of accession VIR-23787. Therefore, we assume that this accession is protected by a new pathogen resistance gene.

Recently, a novel recessive gene conferring broad-spectrum resistance against *Blumeria graminis* f. sp. *hordei* was found in a spring barley line selected from a Moroccan landrace at the Polish Plant Breeding and Acclimatization Institute (Piechota et al., 2020). The gene designated *mlmr* was mapped at the long arm of chromosome 6H. In the other study performed with the use of phytopathological testing, a new resistance allele *MlLu* was identified among 16 winter barley accessions originating from four gene banks (Dreiseitl, 2019). Thus, identification of novel genes in barley landraces can facilitate

Р.А. Абдуллаев, Т.В. Лебедева, Н.В. Алпатьева
Б.А. Баташева, И.Н. Анисимова, Е.Е. Радченко

Устойчивость образцов ячменя из Дагестана
к мучнистой росе

2021
25•5

the broadening of the available powdery mildew resistance germplasm. Moreover, knowledge of resistance phenotypes can assist in determining accessions authenticity and their genotype purity in gene banks (Dreiseitl, Zavřelová, 2018).

Conclusion

The cultivation of barley varieties carrying effective genes for resistance against *B. graminis* f. sp. *hordei* can significantly limit the harmfulness of the pathogen. The specificity in the interaction between the fungus and its host plant requires a continuous search for new resistance donors. Barley landraces are traditionally considered as sources of novel genes for pathogen resistance. The present study performed under field and laboratory conditions has revealed rather low genetic diversity for effective resistance against powdery mildew within the examined set of 264 barley accessions from Dagestan. Only two accessions, VIR-23787 and VIR-28212, were resistant to the Dagestan population of *B. graminis* at the adult stage and to the St. Petersburg population of the pathogen at the seedling stage. Accession VIR-28212 is protected by the effective resistance gene *mlo11*, which was probably introgressed from accession VIR-17554 (Ethiopia). Accession VIR-23787 has another recessive *B. graminis* resistance gene which is distinct from *mlo11* and does not have negative pleiotropic effects typical for other *mlo* alleles obtained by chemical mutagenesis.

References

Abdullaev R.A., Lebedeva T.V., Alpatieva N.V., Yakovleva O.V., Kovaleva O.N., Radchenko E.E., Anisimova I.N., Batasheva B.A., Karabitsina Yu.I., Kuznetsova E.B. Genetic diversity of barley accessions from Ethiopia for powdery mildew resistance. *Russ. Agric. Sci.* 2019;45(2):232-235. DOI 10.3103/S1068367419030029.

Balkema-Boomstra A.G., Masterbroek H.D. Effect of powdery mildew (*Erysiphe graminis* f. sp. *hordei*) on photosynthesis and grain yield of partially resistant genotypes of spring barley (*Hordeum vulgare* L.). *Plant Breed.* 1995;114(2):126-130. DOI 10.1111/j.1439-0523.1995.tb00776.x.

Bonman J.M., Bockelman H.E., Jackson L.F., Steffenson B.J. Disease and insect resistance in cultivated barley accessions from the USDA National Small Grains Collection. *Crop Sci.* 2005;45(4):1271-1280. DOI 10.2135/cropsci2004.0546.

Dorokhov D.B., Kloke E. Rapid and economical technology of RAPD analysis of plant genomes. *Genetika = Genetics* (*Moscow*). 1997; 33(4):443-450. (in Russian)

Dreiseitl A. Genes for resistance to powdery mildew in European barley cultivars registered in the Czech Republic from 2011 to 2015. *Plant Breed.* 2017;136(3):351-356. DOI 10.1111/pbr.12471.

Dreiseitl A. A novel resistance against powdery mildew found in winter barley cultivars. *Plant Breed.* 2019;138(6):840-845. DOI 10.1111/pbr.12730.

Dreiseitl A., Zavřelová M. Identification of barley powdery mildew resistances in gene bank accessions and the use of gene diversity for verifying seed purity and authenticity. *PLoS One.* 2018;13(12): e0208719. DOI 10.1371/journal.pone.0208719.

Ge X.T., Deng W.W., Lee Z.Z., Lopez-Ruiz F.J., Schweizer P., Ellwood S.R. Tempered *mlo* broad-spectrum resistance to barley powdery mildew in an Ethiopian landrace. *Sci. Rep.* 2016;6:29558. DOI 10.1038/srep29558.

Gong X., Li C., Zhang G., Yan G., Lance R., Sun D. Novel genes from wild barley *Hordeum spontaneum* for barley improvement. In: Advance in Barley Sciences. Proc. 11th Int. Barley Genetic Symp. Zhang G., Li C., Liu X. (Ed.). Dordrecht Heidelberg; New York; London: Zhejiang University Press, Springer, 2013;69-89. DOI 10.1007/2F978-94-007-4682-4_6.pdf.

Jørgensen J.H. Discovery, characterization and exploitation of *Mlo* powdery mildew resistance in barley. *Euphytica.* 1992;63(1-2):141-152. DOI 10.1007/BF00023919.

Jørgensen J.H. Genetics of powdery mildew resistance in barley. *Crit. Rev. Plant Sci.* 1994;13(1):97-119. DOI 10.1080/713608055.

Kusch S., Panstruga R. *mlo*-based resistance: an apparently universal "weapon" to defeat powdery mildew disease. *Mol. Plant Microbe Interact.* 2017;30(3):179-189. DOI 10.1094/MPMI-12-16-0255-CR.

Loskutov I.G., Kovaleva O.N., Blinova E.V. Methodological Guidelines for Studying and Maintaining the Global Collection of Barley and Oat. St. Petersburg: VIR, 2012. (in Russian)

Mains E.B., Dietz S.M. Physiologic forms of barley mildew *Erysiphe graminis* hordei Marchal. *Phytopathology.* 1930;20:229-239.

Piechota U., Słowacki P., Czembor P.C. Identification of a novel recessive gene for resistance to powdery mildew (*Blumeria graminis* f. sp. *hordei*) in barley (*Hordeum vulgare*). *Plant Breed.* 2020;139: 730-742. DOI 10.1111/pbr.12819.

Piffanelli P., Ramsay L., Waugh R., Benabdelmouna A., D'Hont A., Hollricher K., Jørgensen J.H., Schulze-Lefert P., Panstruga R. A barley cultivation-associated polymorphism conveys resistance to powdery mildew. *Nature.* 2004;430(7002):887-891. DOI 10.1038/nature02781.

Reinstädler A., Müller J., Jerzy H., Czembor J.H., Piffanelli P., Panstruga R. Novel induced *mlo* mutant alleles in combination with site-directed mutagenesis reveal functionally important domains in the heptahelical barley *Mlo* protein. *BMC Plant Biol.* 2010;10:31. DOI 10.1186/1471-2229-10-31.

Seeholzer S. Isolation and Characterization of New R-protein Variants Encoded at the Barley *Mla* Locus that Specify Resistance against the Fungus Powdery Mildew. University of Zurich, Faculty of Science, 2009;131. DOI 10.5167/uzh-31283.

Skou J.P., Jørgensen J.H., Lilholt U. Comparative studies on callose formation in powdery mildew compatible and incompatible barley. *Phytopathol. Z.* 1984;109(2):147-168. DOI 10.1111/j.1439-0434.1984.tb00702.x.

ORCID ID

R.A. Abdullaev orcid.org/0000-0003-1021-7951
T.V. Lebedeva orcid.org/0000-0003-2344-9233
N.V. Alpatieva orcid.org/0000-0002-5531-2728
B.A. Batasheva orcid.org/0000-0002-2266-281X
I.N. Anisimova orcid.org/0000-0003-0474-8860
E.E. Radchenko orcid.org/0000-0002-3019-0306

The implication of leaf anatomical structure for the selective breeding of lilacs

L.M. Pshennikova

Botanical Garden-Institute of the Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, Russia
✉ pshennikova1@yandex.ru

Abstract. The cultivars of the common lilac (*Syringa vulgaris*) grown in the south of the Russian Far East are not always winter-hardy and are often damaged by fungal diseases due to a very humid climate. A promising trend in the selective breeding of lilacs in Russia is the creation of new breeding material based on the gene pool of the broadleaf lilac (*S. oblata*) and its hybrids in order to introduce valuable adaptive traits into cultivars. The present work aimed to identify the traits of leaf anatomy in species and cultivars of *Syringa* resistant and susceptible to *Pseudocercospora lilacis*, the causative agent of brown leaf spot disease. The study was carried out on the living collection of the Botanical Garden-Institute, Far Eastern Branch, Russian Academy of Sciences (Vladivostok). The leaf anatomical structure of two *Syringa* species showing different degrees of resistance to *P. lilacis* in the monsoon climate of the Far East (resistant *S. oblata* and weakly resistant *S. vulgaris*, and also their hybrid cultivars) has been analyzed. The differences between species, subspecies, and cultivars are quantitative: they differ in the number of spongy mesophyll layers, the cell height in the first layer of palisade mesophyll, the cell height in the upper and lower epidermises, and the thickness of both mesophylls. The interspecific hybrids resistant or weakly resistant to *P. lilacis* (brown leaf spot disease) mainly retain the leaf anatomy structure of the maternal plant. One of the traits determining the resistance of hybrid lilac cultivars is an increased number of spongy mesophyll layers in the leaf blade. The study of leaf anatomy has shown that the four-layered spongy mesophyll leaf parenchyma correlates with the resistance of lilacs from the subsection *Euvulgaris* to *P. lilacis*. In *S. oblata*, this trait is inherited down the maternal line. To establish lilac cultivars resistant to fungal diseases, it is advisable to cross the two species (*S. oblata* and *S. vulgaris*) or their cultivars using one of *S. oblata* subspecies as a maternal plant.

Key words: *Syringa oblata*; *Syringa vulgaris*; *Pseudocercospora lilacis*; leaf anatomical structure; adaptation; interspecific hybridization.

Значение анатомического строения листа в селекции сиреней

Л.М. Пшенникова

Ботанический сад-институт Дальневосточного отделения Российской академии наук, Владивосток, Россия
✉ pshennikova1@yandex.ru

Аннотация. Сорта сирени обыкновенной (*Syringa vulgaris*) на юге российского Дальнего Востока не всегда зимостойки, повреждаются грибными болезнями, что связано с избыточной влажностью климата. Перспективным направлением отечественной селекции сиреней является создание нового селекционного материала с использованием генофонда *Syringa oblata* и ее гибридов для введения в культурные сорта ценных адаптивных признаков. Целью работы было выявление анатомических особенностей листьев видов и сортов рода *Syringa*, устойчивых и восприимчивых к возбудителю бурой листовой пятнистости *Pseudocercospora lilacis*. Исследования проведены на коллекционном материале открытого грунта Ботанического сада-института ДВО РАН. Проанализировано анатомическое строение листа двух видов *Syringa*, различающихся по устойчивости к *P. lilacis* в муссонном климате Дальнего Востока: устойчивого *S. oblata* и слабоустойчивого *S. vulgaris*, а также гибридных сортов, полученных на их основе. Различия между видами, подвидами и сортами носят количественный характер: число рядов губчатой паренхимы, высота клеток первого ряда палисадной паренхимы, высота клеток верхней и нижней эпидермы, толщина палисадного и губчатого слоя. Устойчивые и малоустойчивые к *P. lilacis* межвидовые гибриды преимущественно сохраняют анатомическое строение листа материнского растения. Одним из определяющих признаков устойчивости гибридных сортов сиреней является повышенное количество рядов губчатой паренхимы листовой пластинки – признак, передающийся от *S. oblata* по материнской линии. Создание устойчивых к грибным заболеваниям сортов сиреней целесообразно проводить путем скрещивания видов *S. oblata* и *S. vulgaris* или их сортов, используя в качестве материнского растения один из подвидов *S. oblata*. В ходе исследования анатомии листа установлено, что губчатая ткань из четырех рядов коррелирует с устойчивостью сиреней из подсекции *Euvulgaris* Schneid. к *P. lilacis* на юге Приморского края. Этот признак листа *S. oblata* передается по наследству по материнской линии.

Ключевые слова: *Syringa oblata*; *Syringa vulgaris*; *Pseudocercospora lilacis*; анатомическое строение листа; адаптация; межвидовая гибридизация.

Introduction

The use of new cultivars that are resistant to pathogenic biota is a solution to problems of not only economic, but also environmental significance. Lilac species and cultivars have long been recognized as valuable ornamental plants. However, the habitat and climatic conditions of Vladivostok and its suburbs are very specific (Agroclimatic Resources…, 1973), which becomes a serious obstacle to the introduction of many exotic, alien trees and shrubs in southern Primorsky Krai, Russia, and their outdoor cultivation. The present study was conducted to extend our knowledge about the mechanisms of plant adaptation to the specific climate in the south of the Russian Far East such as, in particular, the mechanisms of protection against adverse biotic factors of the environment shown by some cultivars from the subsection *Euvulgaris* Schneid. of the genus *Syringa* L. There are a number of works published by various botanical institutions that elucidate the species composition of the pathogenic biota associated with the genus *Syringa* (Khomyakov, Tereschenko, 2000; Tomoshevich, Vorobjova, 2010; Chervyakova, Keldish, 2018; Pavlenkova, 2018; Polyakova, 2018, etc.). We could not find any studies that consider the factors of lilacs' resistance or susceptibility to fungal diseases, and there is also a lack of recent data on the mechanism of lilacs' resistance to fungal diseases.

According to our observations (Pshennikova, 2007, 2018), the most resistant cultivars in southern Primorsky Krai are those from the garden group *Hyacinthiflora*, which have been obtained through interspecific hybridization of *S. oblata* and *S. vulgaris*, freely interbreeding with each other. However, this group also includes cultivars that differ in the degree of susceptibility to pathogenic fungi.

The broadleaf lilac, *Syringa oblata* Lindl., is an introduced plant in Primorsky Krai, brought by S.I. Elovitsky from China in the early 20th century (Vasilyuk et al., 1987). In nature, it is found in the northern part of Northeast China (Saakov, 1960; Mei-chen et al., 1996). Currently it is often used for decorating the landscape of the city of Vladivostok and other populated areas of Primorsky Krai. This species is resistant not only to the winter conditions of the region, but also to pests and fungal diseases, and, apparently, has immunity acquired during the evolution in similar climatic conditions of China. Some interspecific hybrid cultivars possess this resistance (Pshennikova, 2007, 2018).

The common lilac, *Syringa vulgaris* L., a species close to *S. oblata*, is winter-hardy in the conditions of Primorsky Krai. It was introduced into Primorsky Krai in the mid-20th century, probably, from Chernigov Oblast, Ukraine (Vasilyuk et al., 1987). Under the continental climate of its natural habitats in highlands (cretaceous slopes on the Balkan Peninsula), this plant has developed a high drought resistance and tolerance to sudden temperature variations. However, for the same reason, it has not formed mechanisms of protection from high humidity characteristic of the climate in southern Primorsky Krai, which is one of the factors responsible for the high prevalence of pathogenic fungi damaging the species and its cultivars. The lilac brown leaf spot disease (Bunkina et al., 1971) causes especially serious damage to this lilac, resulting in loss of decorative appearance of bushes and premature leaf fall.

The features of leaf anatomical structure in pathogen-resistant species and cultivars are considered to be the primary barriers or passive immunity factors (Vavilov, 1964; Shkalikov et al., 2005; Plotnikova, 2007; Shestakova, 2010, 2013). Leaf traits (pubescence, the thick cuticular layer, the thick epidermis, and also the anatomical specifics of mesophyll) have a significant effect on plant immunity (Furst, 1968; Pautov et al., 2002; Sokolova, 2010; Motyleva, Dzhigadlo, 2012). The first data on specifics of the leaf apparatus structure in hybrid cultivars of ornamental woody plants in the literature date back to the 1970s–1980s (Eremin, Novikova, 1976; Novikova, 1976, 1982; Pham van Nang, 1976; Turovsky et al., 1978; Bykova, 1979).

The present study aimed to identify the features of leaf blade anatomical structure in the *Syringa* species and cultivars, bred on the basis of *S. oblata* and *S. vulgaris*, which differ in the degree of resistance to *Pseudocercospora lilacis* (Desm.) Deighton.

Materials and methods

A total of 22 representatives of subsection *Euvulgaris* Schneid. of the genus *Syringa* L. (Table 1) from the live collection grown on open-air plots of the Botanical Garden-Institute, Far Eastern Branch, Russian Academy of Sciences (BGI FEB RAS), located in the coastal zone of southern Primorsky Krai, were used as objects of the study. The material was collected from 2016 to 2019.

The degree of resistance of lilac species and cultivars to *P. lilacis* was scored on a 5-point scale for ornamental cultivated plants (Tamberg, Ulyanova, 1969), which we adapted for the genus *Syringa* (see Table 2): (1) no disease or up to 10 % of leaf surface of the plant damaged; (2) up to 25 % of leaf surface damaged; (3) up to 50 % of leaf surface damaged; (4) up to 75 % of leaf surface damaged; (5) over 75 % of leaf surface damaged.

To analyze mesophyll, five leaf blades were used. The third leaf from the base of a vegetative shoot, completely grown, was sampled from the southern aspect of crown. Leaves were fixed in 70 % ethyl alcohol. Cross-sections through the middle part of a leaf blade between the midrib and the leaf edge were cut on a freezing microtome, stained with a safranin solution, and embedded in glycerol/jelly. The sections were examined under a Zeiss Axioplan 2 Imaging microscope (Carl Zeiss, Germany) using the AxioVision 4 software. The data obtained were processed in the MS Excel package.

The following anatomical characters of a leaf were considered: leaf thickness, height of upper and lower epidermis, thickness of palisade mesophyll, thickness of spongy mesophyll, number of layers of spongy and palisade mesophylls, and size of cells in the 1st and 2nd layers of palisade mesophyll in hybrid cultivars (Fig. 1). The study was conducted at the Center for Collective Use "Microtechnical Laboratory", BGI FEB RAS, and "Bioresourse Collection".

Statistical analysis was carried out using the STATISTICA 6.0 software package. The data were tested for normality of distribution using the Shapiro–Wilk W-test (Shapiro, Wilk, 1965). To search for a statistical relationship between the variables, a correlation analysis was performed using

Table 1. Objects of study

No.	Species/subspecies/cultivar, Author	Cultivar's parents	Year of cultivar's creation
1	*S. vulgaris* L.	–	–
2	*S. oblata* subsp. *oblata* Lindl.	–	–
3	*S. oblata* subsp. *dilatata* (Nakai) P.S. Green et M.C. Chang	–	–
4	*S. oblata* var. *alba* Rehder	–	–
5	*S. oblata* 'Wan Hua Zi', Zang et Fan	*S. oblata* × ?	1984
6	'Xiang Xue', Zang et Fan	*S. oblata* × *S. vulgaris* 'Alba plena'	1984
7	'Luo Lan Zi', Zang et Fan	*S. oblata* × *S. vulgaris* 'Alba plena'	1962
8	'Maiden's Blush', Skinner	*S. oblata* subsp. *dilatata* × *S. vulgaris*	1966
9	'Olimpiada Kolesnikova', Kolesnikov	'Tamara Kolesnikova' × 'Berryer'	1941
10	'Vechernii Vladivostok' ('Evening Vladivostok'), Pshennikova	*S. oblata* subsp. *oblata* × *S. vulgaris*	2007
11	'Neizvestnyi Soldat' ('Unknown Soldier'), Pshennikova	*S. vulgaris* 'Bogdan Khmelnitsky' × *S. oblata* subsp. *oblata*	2017
12	'Neznakomka' ('Stranger Woman'), Pshennikova	*S. oblata* subsp. *oblata* × *S. vulgaris*	2008
13	'Bogdan Khmelnitsky', Rubtsov, Zhogoleva, Lyapunova	*S. vulgaris* × ?	1954
14	'Capitaine Baltet', Lemoine	*S. vulgaris* × ?	1919
15	'Charles Joly', Lemoine	*S. vulgaris* × ?	1896
16	'Mme Florent Stepman', Stepman-Demessemaeker	*S. vulgaris* × ?	1908
17	'Romance', Havemeyer et Eaton	*S. vulgaris* × ?	1954
18	'Buffon', Lemoine	*S.* × *hyacinthiflora*	1921
19	'Esther Staley', Clarke	*S.* × *hyacinthiflora*	1948
20	'Pocahontas', Skinner	*S.* × *hyacinthiflora*	1935
21	'Tom Tayler', Skinner	*S.* × *hyacinthiflora*	1962
22	'Dal'nevostochnitsa' ('Far Eastern Woman'), Pshennikova	'Olimpiada Kolesnikova' × ?	2018
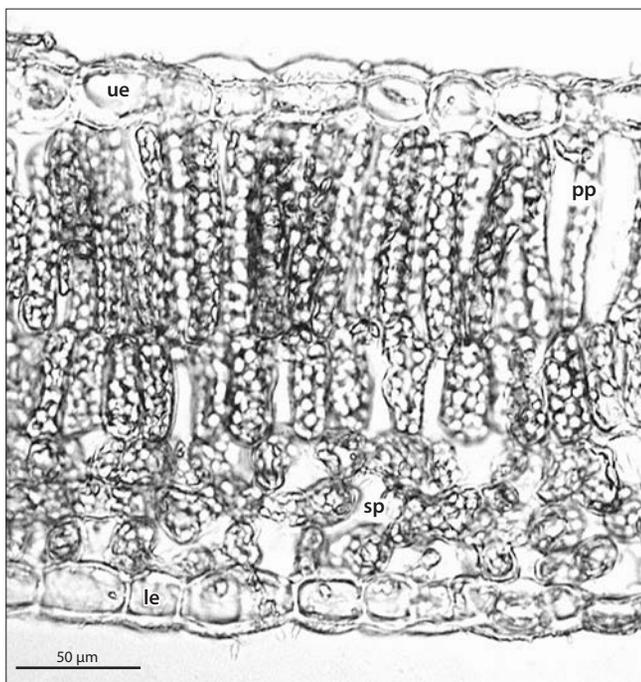
Fig. 1. A leaf cross-section from the lilac cultivar 'Neizvestnyi Soldat'.

The letter designations are as follows: ue, upper epidermis; pp, palisade mesophyll parenchyma; sp, spongy mesophyll parenchyma; le, lower epidermis.

the Spearman Rank Order Correlations (Fieller et al., 1957). Linear measurements of leaf blade tissues were assumed to be independent variables. Degree of resistance of a species/cultivar to *P. lilacis* was assumed to be a dependent variable. According to the *W*-test, the distribution of analyzed data differed from the normal one (p-value < 0.05). Measurements for each representative were made in 20 to 30 replicates; the total number of observations was 454.

Results and discussion

Long-term observations on the species and cultivars of the genus *Syringa* made it possible to arrange them in the order of increasing degree of their leaves' resistance to *P. lilacis* (Table 2).

The first group included two subspecies of *S. oblata* and their hybrid cultivars. Subspecies of *S. oblata* differ in the shape and size of leaf blade. The study of the leaf anatomy of *S. oblata* and *S. vulgaris* allowed identification of the distinguishing traits of these species (Fig. 2, Table 3).

According to the data obtained, the species *S. oblata* resistant to *P. lilacis* differs from the non-resistant species *S. vulgaris* by an increased number of spongy mesophyll layers. The subspecies of *S. oblata* differ from one another in the thickness of leaf blade and the height of palisade mesophyll cells. *S. oblata* subsp. *oblata* has cells of the upper layer of

Table 2. The degree of resistance of the *Syringa* varieties and species to the fungal diseases *Pseudocercospora lilacis*

Degree of stability	Taxa and cultivars
1 No disease or up to 10 % of leaf surface of the plant damaged	*S. oblata* subsp. *oblata*, *S. oblata* subsp. *dilatata*, *S. oblata* var. *alba*, 'Wan Hua Zi', 'Xiang Xue', 'Luo Lan Zi', 'Tom Tayler', 'Pocahontas', 'Maiden's Blush', 'Vechernii Vladivostok', 'Dal'nevostochnitsa', 'Neznakomka', 'Neizvestnyi Soldat'
2 Up to 25 % of leaf surface damaged	'Olimpiada Kolesnikova', 'Buffon'
3 Up to 50 % of leaf surface damaged	'Bogdan Khmelnitsky'
4 Up to 75 % of leaf surface damaged	'Romance'
5 Over 75 % of leaf surface damaged	*S. vulgaris*, 'Capitaine Baltet', 'Charles Joly', 'Mme Florent Stepman', 'Esther Staley'
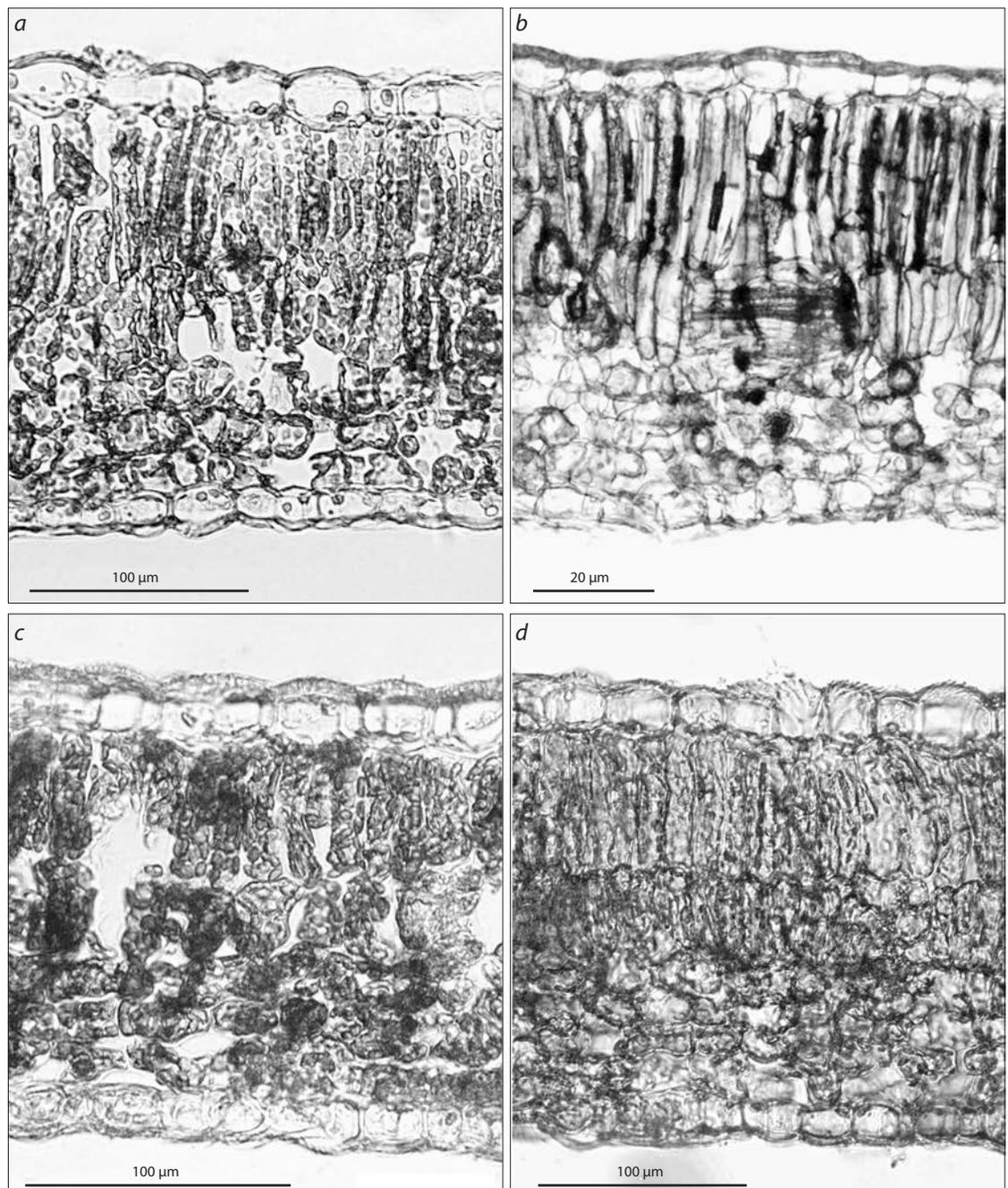
Fig. 2. Leaf cross-sections from the following lilacs:

a, *S. vulgaris*; *b*, *S. oblata* subsp. *oblata*; *c*, *S. oblata* subsp. *dilatata*; *d*, *S. oblata* var. *alba*.

Table 3. Characteristics of leaf cross-sections from species of the genus *Syringa*

Taxon	Epidermis cell length, μm		Palisade mesophyll				Spongy mesophyll		Leaf thickness, μm
	upper	lower	Thickness, μm	Number of layers	Cell length, μm		Number of layers	Thickness, μm	
					1st layer	2nd layer			
S. oblata subsp. *oblata*	24.9±0.7	19.8±0.6	209.9±5.0	2 (3)	140.5±1.0	64.7±1.3	4 (5)	88.6±2.4	351.8±3.0
S. oblata subsp. *dilatata*	26.1±0.6	17.9±0.5	80.4±2.2	2	46.6±1.3	38.4±1.5	3–4	50.9±1.4	183.7±2.8
S. oblata var. *alba*	26.3±0.7	18.3±0.5	103.6±1.2	2	63.5±0.8	41.7±0.7	4 (5)	61.9±0.9	212.8±1.3
S. vulgaris	26.8±0.9	18.6±0.7	110.1±5.8	2 (3)	69.5±1.9	44.3±1.0	3	55.4±1.5	219.6±1.9

Table 4. Characteristics of leaf cross-sections from the *S. oblata* cultivars

Cultivar	Epidermis cell length, μm		Palisade mesophyll				Spongy mesophyll		Spongy mesophyll
	upper	lower	Thickness, μm	Number of layers	Cell length, μm		Number of layers	Thickness, μm	
					1st layer	2nd layer			
'Vechernii Vladivostok'	24.9±0.9	18.2±0.5	233.6±3.3	2 (3)	142.6±9.1	70.2±2.4	4 (5)	102.2±4.1	391.7±3.9
'Neznakomka'	19.2±0.6	16.7±0.4	144.8±2.5	2 (3)	92.7±1.8	51.3±1.4	4	79.6±2.1	262.0±3.5
'Luo Lan Zi'	34.8±1.0	20.7±0.5	141.6±3.6	2	88.8±4.8	44.9±0.9	4 (3)	49.5±1.6	245.9±1.9
'Maiden's Blush'	29.6±0.9	19.5±0.5	120.3±1.5	2	77.4±1.1	42.8±0.8	4 (3)	58.5±1.1	225.8±1.8
'Xiang Xue'	22.9±0.5	20.7±0.6	88.8±1.2	2	50.8±0.9	39.4±0.8	4 (3)	61.2±1.3	192.4±1.7
'Wan Hua Zi'	23.9±0.7	19.9±0.7	188.8±3.1	2 (3)	115.4±2.0	55.1±0.9	4	84.9±1.9	312.8±3.2

palisade mesophyll almost twice as large as the second layer cells. The revealed difference may be a systematic trait of *S. oblata* subsp. *oblata* on the anatomical level. In *S. oblata* subsp. *dilatata*, the heights of the layers of palisade mesophyll cells are either equivalent, or the first layer is slightly, almost 1.2-fold, larger than the second one (see Fig. 2, *c*). This subspecies has upper epidermis cells as large as those in *S. vulgaris.*

We found that the thickness of leaf epidermis and mesophyll in the species under study are not related to resistance against *P. lilacis* (see Table 3). Thus, the height of the upper epidermis cells in the weakly resistant species *S. vulgaris* is greater than that in the fungus-immune subspecies *S. oblata* subsp. *oblata*, or the ratio of the 1st and 2nd layers of palisade mesophyll in *S. vulgaris* is the same as in the subspecies *S. oblata* subsp. *dilatata* resistant to the fungal pathogen. The number of palisade mesophyll layers is not a constant trait in these species and can vary from 2 to 3. The subspecies *S. oblata* subsp. *dilatata* has the leaf anatomical traits close to those of *S. vulgaris* and *S. oblata* subsp. *oblata.*

The following group of lilacs combines cultivars obtained through hybridization of two species, *S. oblata* (maternal species) and *S. vulgaris*. The cultivars 'Vechernii Vladivostok', 'Neznakomka', 'Luo Lan Zi', and 'Wan Hua Zi' repeat the structural features of *S. oblata* subsp. *oblata*; they do not differ in the number of layers of palisade and spongy mesophylls (Table 4, Fig. 3, *a–d*); these cultivars are resistant to *P. lilacis*. It has been found that the *S. oblata* cultivars are distinguished by the height of the upper and lower epidermises, the thick-

ness of palisade and spongy mesophyll parenchyma, and the leaf thickness. The resistant cultivars are characterized by the increased number of spongy mesophyll layers, 4 or more.

The cultivars 'Xiang Xue' and 'Maiden's Blush' are also resistant to *P. lilacis*. Their leaf anatomical structure is characteristic of *S. oblata* subsp. *dilatata*. Their spongy mesophyll is mostly 4-layered, but 3 layers are also observed sometimes.

Another group under study combined the cultivars with both parents being *S. vulgaris*: 'Capitaine Baltet', 'Charles Joly', 'Mme Florent Stepman', 'Romance', and 'Bogdan Khmelnitsky', which proved to be non-resistant to *P. lilacis* to varying degrees. Their leaves are medium in thickness, the palisade mesophyll is 2-layered, and the spongy mesophyll is 3-layered. In 'Bogdan Khmelnitsky', the spongy mesophyll consists sometimes of 4 layers of cells (see Fig. 3, *f*). This can apparently be explained by the fact that the species *S. oblata* was used for breeding the parental cultivars of 'Bogdan Khmelnitsky' at some stage in the past. Compared to the other *S. vulgaris* cultivars, this one is more resistant to the fungal pathogen. The cultivars differ in the height of cells of upper and lower epidermises and in the thickness of palisade and spongy mesophylls. The leaf structure is typical of *S. vulgaris* (Table 5).

The following group of lilacs combined hybrid cultivars referred to as *Hyacinthiflora*, or hyacinth lilacs. These are the cultivars 'Olimpiada Kolesnikova', 'Dal'nevostochnitsa', 'Neizvestnyi Soldat', 'Buffon', 'Esther Staley', etc. (see Table 1). Plants from this group are distinguished both by
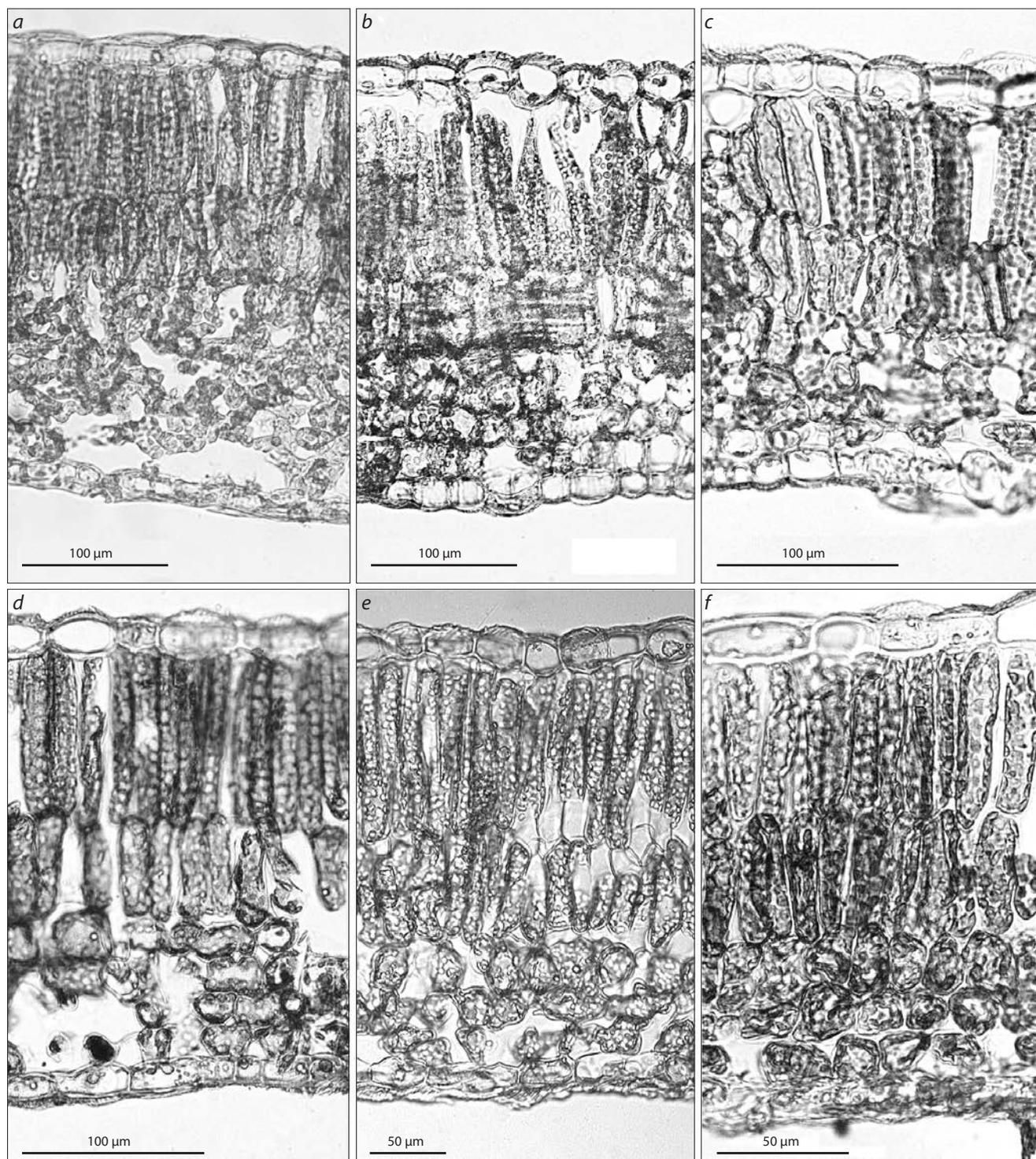
Fig. 3. Leaf cross-sections from lilacs of the following cultivars:
a, 'Vechernii Vladivostok'; *b*, 'Wan Hua Zi'; *c*, 'Maiden's Blush'; *d*, 'Neznakomka'; *e*, 'Olimpiada Kolesnikova'; *f*, 'Bogdan Khmelnitsky'.

leaf anatomical structure (Table 6) and by their resistance to *P. lilacis*. In Russian literature, the cultivar 'Olimpiada Kolesnikova' is attributed to the group of *S. vulgaris* cultivars (Rubtzov et al., 1980; Okuneva et al., 2008). According to the literature data, the cultivar 'Berryer' (*S. oblata* × *S. vulgaris*) was involved in breeding this cultivar. The mesophyll is

5-layered, consisting of 2 layers of palisade and 3(4) layers of spongy mesophylls. By its resistance to *P. lilacis*, the cultivar is placed in group 2 (see Table 2). The cultivar 'Esther Staley' bred by American lilac breeders (parents are unknown) is non-resistant to *P. lilacis*. The cultivar 'Ester Staley' was probably obtained on the basis of *S. vulgaris*. The mesophyll

Table 5. Characteristics of leaf cross-sections from the *S. vulgaris* cultivars

Cultivar	Epidermis cell length, μm		Palisade mesophyll		Spongy mesophyll		Leaf thickness, μm
	upper	lower	Thickness, μm	Number of layers	Thickness, μm	Number of layers	
'Capitaine Baltet'	29.4±0.9	20.6±0.5	150.4±3.5	2	66.7±2.7	3	257.9±2.8
'Charles Joly'	25.0±0.9	20.8±0.9	133.7±3.1	2	74.9±1.8	3	255.3±3.5
'Mme Florent Stepman'	28.9±0.9	20.5±0.5	108.5±1.3	2	68.5±0.9	3	221.9±1.0
'Romance'	19.9±0.6	13.1±0.6	105.4±5.9	2	52.6±1.4	3	200.9±4.1
'Bogdan Khmelnitsky'	21.1±0.5	17.8±0.5	113.6±1.2	2	55.7±1.6	3 (4)	204.7±2.1

Table 6. Characteristics of leaf cross-sections from lilacs of the *Hyacinthiflora* group

Cultivar	Epidermis cell length, μm		Palisade mesophyll		Spongy mesophyll		Leaf thickness, μm
	upper	lower	Thickness, μm	Number of layers	Thickness, μm	Number of layers	
'Buffon'	21.8±0.6	16.4±0.6	126.5±2.7	2 (3)	58.9±2.13	3 (4)	226.5±1.84
'Esther Staley'	25.5±0.8	18.9±0.7	148.3±2.7	2 (3)	69.4±1.5	3 (4)	248.9±5.3
'Pocahontas'	25.3±0.5	16.7±0.4	89.8±1.5	2	51.2±1.1	3–4	192.02±1.8
'Tom Tayler'	28.3±0.8	22.2±0.5	133.1±1.9	2 (3)	57.8±1.8	4 (3)	235.2±1.7
'Olimpiada Kolesnikova'	31.4±1.4	22.6±0.6	163.2±1.9	2	93.8±2.5	3 (4)	301.2±3.8
'Dal'nevostochnitsa'	24.9±1.0	18.1±0.8	154.1±3.4	2	67.9±1.8	4	270.7±5.9
'Neizvestnii Soldat'	28.3±0.8	22.2±0.5	133.1±1.9	2 (3)	57.8±1.8	3 (4)	235.2±1.7

Table 7. Spearman's coefficient of correlation between the leaf characters and the resistance to *Pseudocercospora lilacis*

Leaf character	N	R	t	p-value
Upper epidermis *vs*. Degree of resistance	454	0.013074	0.2780	0.781154
Lower epidermis *vs*. Degree of resistance	454	−0.019263	−0.4096	0.682290
Palisade mesophyll thickness *vs*. Degree of resistance	454	−0.091657	−1.9569	0.050975
Number of palisade mesophyll layers *vs*. Degree of resistance	454	−0.029622	−0.6301	0.528974
Number of spongy mesophyll layers *vs*. Degree of resistance	454	−0.555699	−14.2104	0.000001
Spongy mesophyll thickness *vs*. Degree of resistance	454	−0.028454	−0.6052	0.545361
Leaf thickness *vs*. Degree of resistance	454	−0.076463	−1.6304	0.103712

N o t e. *N*, number of observations; *R*, Spearman's correlation coefficient; *t*, value of the Student's *t*-test for the number of degrees of freedom of $n-2$; *p*-value, probability of error for the null hypothesis that there is no relationship between the characters.

structure in the cultivar 'Neizvestnyi Soldat' is similar to that described above, but it proved to be resistant to *P. lilacis*. Some cultivars of this group ('Dal'nevostochnitsa', 'Tom Tayler', and 'Pocahontas') are resistant to *P. lilacis* and have mainly 4-layered spongy mesophyll. It is likely that the resistance of the cultivars from the *Hyacinthiflora* group is related to the amount of genetic material obtained from parents. The 4-layered spongy mesophyll is not a trait of passive immunity, but rather serves as a marker of the presence of genetic material from *S. oblata* in the hybrid.

A statistical analysis (Table 7) showed the relationship between the number of spongy mesophyll layers and the degree of plants' resistance to *P. lilacis* (*p*-value < 0.001).

We have found that the more layers the spongy mesophyll includes, the higher the resistance to *P. lilacis* (see Table 2). We have not observed any correlation relationship between the other leaf characters and the degree of resistance to *P. lilacis*.

The cultivars that are non-resistant to *P. lilacis* have the leaf anatomical structure similar to that of *S. vulgaris*. Thus, the number of spongy mesophyll layers can be used to predict

the resistance of new lilac cultivars in the climatic conditions of southern Primorsky Krai.

According to our data, a characteristic structural feature of *S. oblata* and all cultivars resistant to *P. lilacis* is the 4-layered spongy mesophyll parenchyma. The cultivars of *S. oblata* subsp. *oblata* are similar in the ratio of heights of the first and second layers of palisade mesophyll parenchyma. The increased number of spongy mesophyll layers (4 or more) correlates (see Table 7) with the resistance to fungal disease. On the other hand, hybrids *S. vulgaris*×*S. oblata*, in which 4-layered spongy mesophyll is sometimes found, can be both resistant and non-resistant to *P. lilacis*. Our conclusions drawn from the examination of the leaf blade anatomical structure in lilacs are consistent with those published by Chinese breeders (Zang et al., 1983; Shuying et al., 1995). These authors reported that the inheritance of maternal traits dominates in hybrid offspring of *S. oblata*×*S. vulgaris*.

We have also found that the examination of leaf anatomy allows identification of the parents of *Syringa* hybrid offspring on the species level, because lilac cultivars retain the structural plan of the maternal plant. In members of the genus *Cerasus* Mill. (Motyleva, Dzhigadlo, 2012; Shestakova, 2013), there is a clear relationship between the primary leaf barriers and resistance to another fungal disease, coccomycosis. Contradictory data were obtained by Turovsky with co-authors (Turovsky et al., 1978), who explain the resistance of cherry tree to a fungal pathogen by the functional features of the host plant. We have established that the structural elements of the leaf anatomical structure (such as the thickness of epidermis and mesophyll) in the studied taxa of the genus *Syringa* are not the primary defense against infection.

Conclusion

Certain traits of the leaf blade anatomical structure in lilacs from the subsection *Euvulgaris* Schneid. of the genus *Syringa*, such as, in particular, the spongy mesophyll parenchyma consisting of 4 layers, can be considered an indicator of the degree of their resistance to *Pseudocercospora lilacis* in southern Primorsky Krai. For creating lilac cultivars resistant to fungal diseases, it is expedient to cross two species (*S. oblata* and *S. vulgaris*) or their cultivars, using one of the subspecies of *S. oblata* as a maternal plant. With free pollination, only seeds from resistant cultivars should be taken. The difference in the mesophyll anatomical structure observed in *S. oblata* subsp. *oblata* and *S. oblata* subsp. *dilatata* can be used as an additional diagnostic trait.

References

Agroclimatic Resources of the Primorskiy Territory. Leningrad: Gidrometeoizdat Publ., 1973. (in Russian)

Bunkina I.A., Koval E.Z., Nelen E.S. Mycoflora and Fungal Diseases in Community Landscapes of the Far East. Vladivostok, 1971. (in Russian)

Bykova N.B. Anatomy of the leaf and annual stem of some almond species and hybrids. In: Biological and Structural Features of Useful Plants in Uzbekistan (Hazelnuts, Cloves, Legumes). Tashkent: Fan Publ., 1979;68-74. (in Russian)

Chervyakova O.N., Keldish M.A. Peculiarities of *Syringa* L. protection from noxious organisms at introduction. In: Procedures of the International Lilac conference "International Syringa 2018". May 21–27, 2018. Moscow; St. Petersburg, 2018;224-228. (in Russian)

Eremin G.V., Novikova L.N. Anatomical and morphological features of hybrids between plum species with different ploidies. *Doklady VASKHNIL = Reports of the Academy of Agricultural Sciences.* 1976;7:17-19. (in Russian)

Fieller E.C., Hartley H.O., Pearson E.S. Tests for rank correlation coefficients. I. *Biometrika.* 1957;44(3-4):470-481. DOI 10.1093/biomet/44.3-4.470.

Furst G.G. The structure of the seed peel in different types and varieties of onions. *Bulleten Glavnogo Botanicheskogo Sada = Bulletin of the Main Botanical Garden.* 1968;69:55-60. (in Russian)

Khomyakov M.T., Tereschenko S.I. The resistance of lilac to diseases. *Zashchita i Karantin Rasteniy = Plant Protection and Quarantine.* 2000;7:31-32. (in Russian)

Mei-chen C., Lian-qing Q., Green P.S. Oleaceae. In: Flora of China. 1996;15:272-319. Available at: http://flora.huh.harvard.edu/china/mss/volume15/index.htm

Motyleva S.M., Dzhigadlo E.N. Morphoanatomical characteristics and elemental composition of cherry leaves in connection with resistance to coccomycosis. *Plodovodstvo i Yagovodstvo Rossii = Pomiculture and Small Fruits Culture in Russia.* 2012;30:253-261. (in Russian)

Novikova L.N. Some anatomical features of plum species and varieties. *Byulleten VIR = Bulletin of the Institute of Plant Industry* (*Leningrad*). 1976;60:73-76. (in Russian)

Novikova L.N. Anatomical and morphological characteristics of the leaf of distant hybrids of stone fruits. *Nauchno-Tekhnicheskiy Byulleten VIR = Bulletin of the Institute of Plant Industry* (*Leningrad*). 1982;123:54-57. (in Russian)

Okuneva I.B., Mikhailov N.I., Demidov A.S. Lilac. Collection of the Main Botanical Garden of the Russian Academy of Sciences: History and the Current State. Moscow: Nauka Publ., 2008. (in Russian)

Pautov A.A., Yakovleva O.V., Kolodjagnii S.F. Leafe epidermis microrelief in *Populus* (Salicaceae). *Botanicheskiy Zhurnal = Botanical Journal.* 2002;87(1):63-71. (in Russian)

Pavlenkova G.A. Resistange of species of the genus *Syringa* L. to damaging abiotic and biotic factors of the environment in conditions of Orel region. In: Procedures of the International Lilac conference "International Syringa 2018". May 21–27, 2018. Moscow; St. Petersburg, 2018;219-223. (in Russian)

Pham van Nang. Comparative anatomical study of some alleged hybrids in the genus *Crataegus* L. *Uzbekskiy Botanicheskiy Zhurnal = Uzbek Botanical Journal.* 1976;6:48-52. (in Russian)

Plotnikova L.Ya. Plant Immunity and Breeding for Resistance to Diseases and Pests. Moscow: KolosS Publ., 2007;69. (in Russian)

Polyakova N.V. Diseases and pests of the *Syringa* L. collection of the South Ural Botanical Garden-Institute. *Agrarnaya Rossiya = Agricultural Russia.* 2018;12:17-19. (in Russian)

Pshennikova L.M. Lilac Species Cultivated in the Botanical Garden-Institute, Far Eastern Branch of the Russian Academy of Sciences. Vladivostok: Dalnauka Publ., 2007. (in Russian)

Pshennikova L.M. Perspectives of lilac introduction in the south of the Russian Far East. In: Procedures of the International Lilac conference "International Syringa 2018". May 21–27, 2018. Moscow; St. Petersburg, 2018;152-156. (in Russian)

Rubtsov L.I., Mikhaylov N.L., Zhogoleva V.G. Species and Varieties of Lilac Cultivated in the USSR. Repertory Catalog. Kiev: Naukova Dumka Publ., 1980. (in Russian)

Saakov S.G. Genus Lilac (*Syringa* L.). In: Trees and Shrubs of the USSR. Moscow–Leningrad, 1960;5:435-458. (in Russian)

Shapiro S.S., Wilk M.B. An analysis of variance test for normality (complete samples). *Biometrika.* 1965;52(3-4):591-611. DOI 10.1093/biomet/52.3-4.591.

Shestakova V.V. Assessment of the resistance of cherry and cherry varieties to coccomycosis by anatomical and morphological features. *Plodovodstvo i Vinogradarstvo Yuga Rossii = Fruit Growing and Viticulture of Southern Russia.* 2010;20(2):76-82. (in Russian)

Shestakova V.V. The use of biochemical and anatomical and morphological parameters to study the resistance of *Cerasus* Mill. represen-

tatives to *Coccomyces hiemalis* Higgins. *Nauka Kubani = Science in Kuban.* 2013;1:16-20. (in Russian)

Shkalikov V.A., Dyakov Yu.T., Smirnov A.N. et al. Plant Immunity. Moscow: KolosS Publ., 2005. (in Russian)

Shuying Z., Yinghan F., Ronghui L. Breeding of new cultivarisin the genus *Syringa* (Oleaceae). *Acta Hortic.* 1995;404:63-67. DOI 10.17660/actahortic.1995.404.9.

Sokolova E.A. The significance of anatomical traits for the taxonomy of the Prunoideae (Rosaceae) subfamily. Synopsis of Dr. Biol. Sci. Diss. St. Petersburg, 2000. (in Russian)

Tamberg T.G., Ulyanova T.N. Guidelines for Studying the Collection of Ornamental Crops. Leningrad: Vavilov Institute of Plant Industry, 1969. (in Russian)

Tomoshevich M.A., Vorobjova I.G. Diseases of lilac in Siberian urban plantings. *Zashchita i Karantin Rasteniy = Plant Protection and Quarantine.* 2010;5:51. (in Russian)

Turovsky I.I., Zhukov O.S., Shcheglova L.A. Anatomical and ultra-structural features of mesophyll cells of cherry forms immune and susceptible to coccomycosis. *Bulleten TsGL imeni I.V. Michurina = Bulletin of the Michurin Central Genetic Laboratory.* 1978;31:33-36. (in Russian)

Vasilyuk V.K., Vrisch D.L., Zhuravkov A.F., Kostenko K.A., Lobano-va I.I., Mironova L.N., Petukhova I.P., Rout A.N., Seledets V.P., Smirnova O.A., Urusov V.M., Filatova L.D., Khmelnitsky K.A., Hrapko O.V., Centalovich V.T., Chipizubova M.N., Bityukov S.A., Pozdnyakov D.L., Voronkova N.M., Prilutsky A.N. Gardening of Cities of the Primorskiy Territory. Vladivostok: Far Eastern Branch of the USSR Academy of Sciences, 1987. (in Russian)

Vavilov N.I. Immunity Problems of Cultivated Plants. Vol. 4. Moscow–Leningrad, 1964. (in Russian)

Zang S.Y., Fan Y.H., Li R.H. Hybridization and breeding of *Syringa* plants. *Collection of Papers on Transplanting and Domestication.* 1983;3:117-121.

ORCID ID

L.M. Pshennikova orcid.org/0000-0002-8243-3752

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

Genotypic and ecological variability of zinc content in the grain of spring bread wheat varieties in the international nursery KASIB

V.P. Shamanin[1] ✉, P. Flis[2], T.V. Savin[3], S.S. Shepelev[1], O.G. Kuzmin[1], A.S. Chursin[1], I.V. Pototskaya[1], I.E. Likhenko[4],
I.Yu. Kushnirenko[5], A.A. Kazak[6], V.A. Chudinov[7], T.V. Shelaeva[8], A.I. Morgounov[9]

[1] Omsk State Agrarian University named after P.A. Stolypin, Omsk, Russia
[2] University of Nottingham, Nottingham, United Kingdom
[3] Kazakh Research Institute of Agriculture and Plant Growing, Almalybak, Almaty region, Kazakhstan
[4] Siberian Research Institute of Plant Production and Breeding – Branch of the Institute of Cytology and Genetics of the Siberian Branch
  of the Russian Academy of Sciences, Novosibirsk, Russia
[5] Chelyabinsk Agricultural Research Institute, Chelyabinsk, Russia
[6] Northern Trans-Ural State Agricultural University, Tyumen, Russia
[7] Karabalyk Experimental Agricultural Research Station, Kostanai region, Kazakhstan
[8] Research and Production Center for Grain and Farming named after A.I. Baraev, Shortandy, Akmola region, Kazakhstan
[9] Ministry of Environment, Water and Agriculture, Riyadh, Kingdom of Saudi Arabia
✉ vp.shamanin@omgau.org

Abstract. Spring bread wheat is the staple crop in Western Siberia and Kazakhstan, a significant portion of which goes for export. Wheat breeding with a high level of zinc in wheat grain is the most cost-effective and environmentally friendly way to address zinc deficiency in the diet. The purpose of this work was to evaluate the contribution of the factors 'location' and 'genotype' in the variability of zinc content in wheat grain, and to identify the best varieties as sources of this trait for breeding. The research on screening zinc content in the wheat grain of 49 spring bread wheat varieties from the Kazakhstan-Siberia Spring Wheat Trial (KASIB) nursery was carried out at 4 sites in Russia (Chelyabinsk, Omsk, Tyumen, Novosibirsk) and 2 sites in Kazakhstan (Karabalyk and Shortandy) in 2017–2018. The content of zinc in wheat grain was evaluated at the Ionomic Facility of University of Nottingham in the framework of the EU project European Plant Phenotyping Network-2020. The analysis of variance showed that the main contribution into the general phenotypic variation of the studied trait, 38.7 %, was made by the factor 'location' due to different contents of zinc and moisture in the soil of trial sites; the effect of the factor 'year' was 13.5 %, and the effect of the factor 'genotype' was 8.0 %. The most favorable environmental conditions for accumulation of zinc in wheat grain were observed in the Omsk region. In Omsk, the average zinc content in all studied varieties was 50.4 mg/kg, with 63.7 mg/kg in the best variety 'OmGAU 100'. These values are higher than the target values of the international program Harvest Plus. 'Novosibirskaya 16' (49.4 mg/kg), 'Silach' (48.4 mg/kg), 'Line 4-10-16' (47.2 mg/kg), 'Element 22' (46.3 mg/kg) and 'Lutescens 248/01' (46.0 mg/kg) were identified as being the best varieties. Significant possibilities for the production of wheat grain with high zinc content, which is in demand for the production of bread and pastry products with functional properties, were identified in the Western Siberian region.
Key words: variety; grain of wheat; zinc; protein; ecology.

Генотипическая и экологическая изменчивость содержания цинка в зерне сортов яровой мягкой пшеницы международного питомника КАСИБ

В.П. Шаманин[1] ✉, П. Флис[2], Т.В. Савин[3], С.С. Шепелев[1], О.Г. Кузьмин[1], А.С. Чурсин[1], И.В. Потоцкая[1], И.Е. Лихенко[4],
И.Ю. Кушниренко[5], А.А. Казак[6], В.А. Чудинов[7], Т.В. Шелаева[8], А.И. Моргунов[9]

[1] Омский государственный аграрный университет им. П.А. Столыпина, Омск, Россия
[2] Ноттингемский университет, Ноттингем, Великобритания
[3] Казахский научно-исследовательский институт земледелия и растениеводства, Алмалыбак, Алматинская область, Казахстан
[4] Сибирский научно-исследовательский институт растениеводства и селекции – филиал Федерального научного центра Институт цитологии
  и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[5] Челябинский научно-исследовательский институт сельского хозяйства, Челябинск, Россия
[6] Государственный аграрный университет Северного Зауралья, Тюмень, Россия
[7] Карабалыкская сельскохозяйственная опытная станция, пос. Карабалык, Костанайская область, Казахстан
[8] Научно-производственный центр зернового хозяйства им. А.И. Бараева, пос. Шортанды, Акмолинская область, Казахстан
[9] Продовольственная и сельскохозяйственная организация ООН, Рияд, Саудовская Аравия
✉ vp.shamanin@omgau.org

Аннотация. Яровая мягкая пшеница является основной культурой в Западной Сибири и Казахстане, где значительная доля производимого зерна идет на экспорт. Селекция пшеницы на повышенное содержание цинка в зерне – наиболее рентабельный и экологичный способ решения проблемы дефицита цинка в рационе питания. Цель настоящей работы – установить вклад факторов «пункт» и «генотип» в изменчивость содержания цинка в зерне пшеницы и выделить лучшие сорта в качестве источников данного признака для селекции. Исследования по скринингу накопления цинка в зерне пшеницы 49 сортов яровой мягкой пшеницы из питомника КАСИБ-18 проведены в четырех пунктах России (Челябинск, Омск, Тюмень, Новосибирск) и двух пунктах Казахстана (Карабалык и Шортанды) в течение 2017–2018 гг. Содержание цинка в зерне определяли на факультете иономики Университета г. Ноттингем в рамках проекта EPPN-2020. Результаты дисперсионного анализа показали, что основной вклад в общее фенотипическое варьирование признака вносил фактор «пункт» (38.7 %) вследствие разного содержания цинка в почве и влагообеспеченности в пунктах испытания; влияние факторов «год» и «генотип» составило 13.5 и 8.0 % соответственно. Наиболее благоприятные экологические условия для получения зерна пшеницы с повышенным содержанием цинка сложились в Омской области, где в среднем по всем сортам содержание цинка было равно 50.4 мг/кг, а у лучшего сорта ОмГАУ 100 – 63.7 мг/кг. Эти показатели выше целевых значений международной программы Harvest Plus. Выделены лучшие сорта – Новосибирская 16 (49.4 мг/кг), Силач (48.4 мг/кг), Линия 4-10-16 (47.2 мг/кг), Элемент 22 (46.3 мг/кг) и Лютесценс 248/01 (46.0 мг/кг). В Западно-Сибирском регионе выявлены значительные потенциальные возможности производства зерна пшеницы с повышенным содержанием цинка, востребованного для получения хлеба и кондитерских продуктов с функциональными свойствами.

Ключевые слова: сорт; зерно пшеницы; цинк; белок; экология.

Introduction

Wheat remains one of three crop commodities (along with maize and rice) contributing to global food security. Global wheat production has been increasing at a steady annual rate of 1–2 % to meet the growing population demand. According to FAO (http://www.fao.org/faostat/en), in Russian Federation, area under wheat has grown from 23.9 mln ha in 2014 to 26.5 mln ha in 2018 (+10.9 %), grain yield – from 2.50 to 2.72 t/ha (+8.8 %) and the total production – from 59.7 to 72.1 mln t (+20.7 %). The grain exports have increased more than two times and exceeded 35 mln t in 2019 and over 38.5 mln t in 2020. At present, wheat production in the world satisfies the demand and more attention shall be paid to wheat grain quality.

One of the pioneering works on this subject is European Union Project HEALTH GRAIN. The project was implemented in 2005–2010 and laid out the foundation for the studies for improving grain wheat nutritional value: protein content and composition, carbohydrates, vitamins, micronutrients and phytochemicals (Björck et al., 2012). Unfortunately, in Russia, the work on functional properties of wheat grain is limited to the study of purple wheat and its products in the Institute of Cytology and Genetics (Khlestkina et al., 2019; Gordeeva et al., 2020). The enhancement of functional properties and nutritional value of wheat grain products will have beneficial effect on human health and immune status, especially in connection with threats similar to coronavirus pandemic.

Wheat biofortification was started in mid-2000s by Harvest Plus consortium (https://www.harvestplus.org/what-we-do/crops) and made tremendous progress. The grain zinc concentration of new biofortified wheat varieties increased by 40 % (+12 mg/kg) compared to commercial varieties (Velu et al., 2011; Singh R., Velu, 2017).

Recent results obtained from Harvest Plus and Harvest Zinc projects in China, India, Mexico, Pakistan, South Africa, and Turkey indicate positive effects of foliar-applied Zn (zinc) alone, and a micronutrient cocktail solution containing I (iodine), Zn, Se (selenium), and Fe (iron) that significantly improve grain accumulation of micronutrients, particularly in new biofortified wheat varieties. Grain-Zn was increased from 28.6 to 46.0 mg/kg with Zn-spray and 47.1 mg/kg with micronutrient cocktail spray (Zou et al., 2019).

Grain Zn contents of wheat varied among different countries from 25.10 mg/kg in Europe to 33.91 mg/kg in North America depending on: (1) the amount of Zn available in the soil; (2) genotypic characteristics of cultivated varieties; (3) cultivation types, environments, climates (Wang et al., 2020). Modern wheat varieties have limited grain Zn concentration: on the average – 14–42 mg/kg (Bouis, 1995; Morgunov et al., 2007; Velu et al., 2011; Guttieri et al., 2015). In this connection a large-scale screening of wheat genetic resources at the germplasm bank of the International Maize and Wheat Improvement Center (CIMMYT) was initiated to explore variation for Zn amongst the wheat wild relatives *T. monococcum*, *T. dicoccoides*, *Ae. tauschii*, *T. boeticum, T. spelta*, *T. polonicum*, landraces, and wheat hexaploid synthetics, which detected the most promising sources for development of varieties with high grain Zn concentration (Cakmak et al., 2004; Velu et al., 2014; Verma et al., 2016; Savin et al., 2018; Bhatta et al., 2019).

A field evaluation of a set of core-collection of landraces of CIMMYT screened under Zn-enriched soil conditions at Cd. Obregon (Mexico) showed that there was high variation for grain Zn concentration – from 40 to 96 mg/kg. *T. dicoccoides* introgression lines with bread wheat background showed up to 88 mg/kg grain Zn concentration. The first high zinc wheat variety Zincol 2016, having *T. spelta* in its pedigree, was released in Pakistan. Zn-enriched wheat varieties such as Zinc Shakti, WB 02, and HPBW 01 were adapted by more than 500,000 farmers in India. These varieties were developed using synthetic hexaploid wheat with the genome of *Ae. tauschii* (Velu et al., 2019).

V. Govindan et al. (2018) reported a moderate level of broad-sense heritability for grain Zn concentration, and a significant Genotype × Environment interaction effect on this trait. The search and introgression of genes controlling high zinc content into the initial material for grain quality breeding

В.П. Шаманин, П. Флис, Т.В. Савин …
В.А. Чудинов, Т.В. Шелаева, А.И. Моргунов

Изменчивость содержания цинка в зерне
сортов пшеницы питомника КАСИБ

2021
25•5

through marker assisted selection has been conducted. One study identified QTLs associated with grain Zn concentration in wheat. These were located on chromosomes 2A, 5A, 7A (Peleg et al., 2009; Xu et al., 2012; Krishnappa et al., 2017). According to the research results of Y. Genc et al. (2009), the combination of four loci located on chromosomes 7A, 4B, 6B, and 3D increased the grain Zn by 23 %. The gene *GPC-B1* (*NAM-B1*) was transferred to bread wheat genome from *T. dicoccoides*. Current tetraploid and hexaploid wheat varieties have non-active allele *GPC-B1*, except for some landraces and old varieties of *T. dicoccum*, *T. durum*, *T. spelta*, and *T. aestivum* (Mitrofanova, Khakimova, 2016). The active allele of this gene can be effective in improving high protein content, and remobilization of micronutrients from flag leaf to grains, which increases the concentration of Fe and Zn by 18 and 12 %, respectively (Uauy et al., 2006; Waters et al., 2009).

Grain Zn concentration is negatively correlated with yield in spring wheat varieties in several studies (Welch, Graham, 2002; Morgunov et al., 2007; Murphy et al., 2008). Some experiments, on the contrary, indicate that this correlation does not necessarily occur, and illuminate the possibility of combining high grain Zn with high grain yield and protein content in new varieties (Chen et al., 2017; Krishnappa et al., 2017; Abugalieva, Savin, 2018). The minerals' bioavailability to humans, including Zn, depends on the phytic acid, which binds them. In this connection, the current wheat varieties should combine high yield with low phytic acid/Zn ratio (<5) (Qi et al., 2013; Liu et al., 2014).

Omsk State Agrarian University (Omsk SAU) coordinates Kazakhstan-Siberia network on spring wheat improvement (KASIB), which combines 20 breeding and scientific research institutions from Kazakhstan and Russia. In earlier studies, more than 150 genotypes were evaluated at 4–8 sites of KASIB network in Kazakhstan and Western Siberia in search for genetic resources of high zinc content. The relationship of Zn grain concentration with protein content and effects of Genotype × Environment interaction for this trait were studied (Morgounov et al., 2006; Gomez-Becerra et al., 2007).

In 2016, Kazakh Research Institute of Farming and Crop Production won a grant of the project EPPN-2020 (European Plant Phenotyping Network) to conduct ionomics analysis of spring wheat grain from Kazakhstan and Russia using ionomics phenotyping platform at the University of Nottingham (UK). This platform couples high throughput elemental analysis based on automated data capture and their processing with bioinformatic methods (https://www.ionomicshub.org/home/PiiMS). Ionomics analysis of 23 elements including heavy and rare metals, micronutrients in 49 spring bread wheat varieties at 6 sites in Kazakhstan and Russia was conducted in 2017–2018. This study identified that the variability of elemental analysis of grain wheat depends on the Genotype × Environment factors, and their interaction. The highest Zn and Fe concentrations in grain wheat were detected in Omsk oblast. Some varieties and breeding lines with high Zn and Fe content were identified (Abugalieva et al., 2020).

The objective of this research was to determine the contribution of the 'Location' and 'Genotype' factors on variability of wheat grain Zn content, and to select the best varieties as sources of this trait for breeding.

Materials and methods

The study of 49 varieties of spring bread wheat of the nursery KASIB-18 (Kazakh-Siberian nursery of spring bread wheat) was carried out at four sites in Western Siberia, Southern Urals, and in two sites in Kazakhstan (Fig. 1). Geographic coordinates of Russian experimental sites are: Chelyabinsk Research Institute of Agriculture (Chelyabinsk) – 54°93′ N, 60°73′ E; Omsk SAU (Omsk) – 55°01′ N, 73°18′ E; Northern Trans-Ural SAU (Tyumen) – 57°09′ N, 65°25′ E; Siberian Research Institute of Plant Cultivation and Breeding (Novosibirsk) – 54°89′ N, 82°97′ E; Kazakhstan experimental sites: Karabalyk Experimental Agricultural Research Station (Karabalyk) – 53°51′ N, 62°06′ E; Research and Production Center for Grain and Farming (Shortandy) – 51°63′ N, 71°04′ E.

Variety trial of the nursery KASIB-18 was carried out in 2017–2018. The weather conditions differed significantly in geographic experimental sites (Table 1).

At all experimental sites of the KASIB network, the sum of active temperatures above 10 °C in 2017–2018 was higher than the values necessary for normal growth and development of wheat plants: the lowest at Tyumen – 2118–2124 °C and the highest at Karabalyk – 2553–2637 °C. According to the hydrothermic coefficient (HTC), calculated by G.T. Selyaninov method (1958), the most favorable conditions for moisture availability were in the sites Tyumen and Novosibirsk (HTC = 1.38–1.53) in both years of research, as well as in Chelyabinsk in 2018 (HTC = 1.42), which had a positive effect on the formation of higher grain yield in these sites. In general, 2017 was characterized by drier conditions in Omsk (HTC = 0.72) and Shortandy (HTC = 0.46) compared to 2018 (HTC = 1.15 and 1.24, respectively). In Karabalyk, in both years, dry conditions were observed during the plant growing season (HTC = 0.80–0.83).

There are no significant differences in the soil morphological characters of the experimental sites, with the exception of higher humus content in Tyumen and Novosibirsk (7.0–7.5 %). Based on the literature sources, zinc content in the humus layer of meadow chernozem soils of the Omsk region is 20.1–69.4 mg/kg (Azarenko et al., 2019). According to the data from JSC "Kazakhstan Agrarian Expertise" branch (www.kazagrex.kz) in Akmola region, zinc content in low-humus soil of Shortandy is 3.3 mg/kg. The data of zinc content in the soil of the remaining sites are not available.

Fig. 1. Map of experimental sites of nursery KASIB-18 in the regions of Russia and Kazakhstan in 2017 and 2018.

Table 1. Characteristics of soil and weather conditions at experimental sights
during the growing season (May–September), 2017–2018

Trait	Chelyabinsk	Omsk	Tyumen	Novosibirsk	Karabalyk	Shortandy
Sum of active temperatures, °C						
2017	2349	2364	2118	2292	2637	2508
2018	2307	2169	2124	2157	2553	2130
Rainfall, mm						
2017	273.7	171.0	324.0	315.9	218.5	114.6
2018	328.1	249.0	322.0	296.9	204.5	264.5
Hydrothermic coefficient						
2017	1.16	0.72	1.53	1.38	0.83	0.46
2018	1.42	1.15	1.52	1.38	0.80	1.24
Soil type	Chernozem, medium loamy	Meadow- chernozem, heavy loamy	Chernozem, medium loamy	Chernozem, medium loamy	Chernozem, heavy loamy	Chernozem, heavy loamy
Soil acidity, units	5.3	6.8	6.7	6.7	6.8	7.7
Humus, %	5.2	5.2	7.0	7.5	4.7	3.3

Sowing, selection assessments, and observations in the nursery were carried out in accordance with the Methodology of State Variety Trial of Agricultural Cultures (1989) and the program of the Kazakhstan-Siberia network on spring wheat improvement. The plot area was 3 m² with a sowing rate of 500 seeds per 1 m². The sowing date was May 20–30, the sowing depth – 4–5 cm. Field trials utilized a systematic complete block design with three replicates. The preceding crop was black fallow.

Grain samples from each experimental site were analyzed in the Kazakh Research Institute of Agriculture and Plant Growing (Almalybak), for protein content in grain and its fractions determined by Kjeldahl method (State standard No. 10846-91) using Infratec FOSS 1841 on the basis of previously created calibration equations. Zinc content in grain was determined at the Ionomics Faculty of the University of Nottingham. Zinc concentration was calculated in mg/kg of dry weight. Statistical data processing was reconstructed by variational, correlation, and ANOVA analysis using Microsoft Excel and Statistica application software packages.

Results

Analysis of zinc accumulation in wheat grain of 49 varieties of KASIB-18 indicates significant differences in the grain zinc content, depending on the experimental site (Table 2).

In Omsk, the yield was low, but the grain Zn content, on the contrary, was the highest (50.4 mg/kg). In Shortandy, the average grain yield was almost at the level of the varieties yield in Omsk, but the grain Zn content was 1.9 times lower, which indicates a significant influence of soil and climatic properties of the region on the accumulation of this microelement in wheat grain. The highest grain Zn content was found in Omsk and in other sites of Russia – Tyumen, Novosibirsk, Chelyabinsk (44.1–44.8 mg/kg), and significantly less in Ka-

zakhstan – Karabalyk and Shortandy (37.3 and 26.8 mg/kg, respectively).

No correlation was found between grain Zn content and yield. One site was an exception – under dry conditions of 2017, an average negative relationship was observed in Karabalyk ($r = -0.35$) with the lowest yield (1786 kg/ha) compared to the rest experimental sight.

On average, for two years of research, the highest grain yield was in Novosibirsk – 3985 kg/ha, in Tyumen, and Chelyabinsk – 3210 and 3780 kg/ha, respectively. The yield obtained in Karabalyk, Omsk, and Shortandy was less than 3000 kg/ha. Significant differences in the grain average protein content of different experimental sites were found. The highest protein content on average for two years of research was observed in Omsk (14.9 %), in Shortandy (14.0 %), in Novosibirsk (13.1 %), in Chelyabinsk (12.2 %), in Karabalyk (12.1 %), and the lowest – in Tyumen (10.7 %). In Tyumen, in 2017–2018, an average positive correlation was observed between grain Zn and protein content – 0.3 and 0.4, respectively.

On the basis of the experiment results of 49 varieties for two years in 6 ecological sites, a three-factor ANOVA analysis was carried out, and the contribution of the main factors to variability of wheat grain Zn was determined (Fig. 2). The main contribution to the variability of the studied trait was made by the 'Location' factor – 38.7 %. ANOVA analysis revealed significant influence of the following factors: 'Year' – 13.5 %, 'Genotype – 8.0 %, 'Genotype × Location' – 14.3 %, and 'Location × Year' – 7.8 %. The combined effect of three factors interaction was significant – 15.1 %.

Figure 3 presents the limits of average indicators of grain Zn content of the studied varieties for two years of research. The maximum grain Zn content was observed in one variety – 49.4 mg/kg, in six varieties Zn content varied from 45.7 to 49 mg/kg, in 16 varieties – from 39.1 to 42.4 mg/kg, in

В.П. Шаманин, П. Флис, Т.В. Савин …
В.А. Чудинов, Т.В. Шелаева, А.И. Моргунов

Изменчивость содержания цинка в зерне
сортов пшеницы питомника КАСИБ

2021
25•5

Table 2. Grain yield, protein content (PC), Zn content, and coefficients of correlation between (*r*) Zn content, yield and PC in the varieties from experimental sights of nursery KASIB-18 in 2017–2018

Year	Grain yield, kg/ha	PC, %	Zn, mg/kg	(*r*) Zn with yield	(*r*) Zn with protein
		Omsk			
2017	3032	16.3	47.6	0.08	0.01
2018	2019	13.6	53.2	–0.05	0.26*
Average	2526	14.9	50.4		
		Chelyabinsk			
2017	4098	10.7	36.4	0.14	0.23
2018	3461	13.7	51.8	0.06	0.25
Average	3780	12.2	44.1		
		Novosibirsk			
2017	3113	13.2	46.5	–0.06	0.14
2018	4857	12.9	43.1	–0.03	0.07
Average	3985	13.1	44.8		
		Tyumen			
2017	2382	11.1	40.8	–0.09	0.30*
2018	4037	10.3	48.9	–0.02	0.40*
Average	3210	10.7	44.8		
		Karabalyk			
2017	1786	12.5	31.7	–0.35*	0.03
2018	4003	11.7	42.9	–0.01	0.12
Average	2894	12.1	37.3		
		Shortandy			
2017	2580	14.8	18.5	–0.18	–0.04
2018	2413	13.2	35.0	0.08	0.29*
Average	2496	14.0	26.8		
LSD$_{05}$	663	1.1	6.2		

* Significant at *p* < 0.05 probability level.

Fig. 2. Contribution of various factors for Zn grain content of wheat varieties, %.
Significant at * *p* < 0.05 and ** *p* < 0.001 probability level.
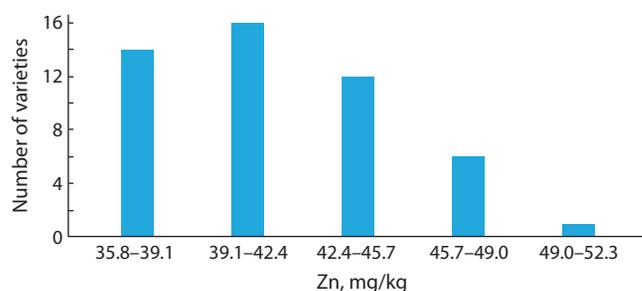
Fig. 3. Number of varieties by intervals of Zn grain content from experimental sites, on average of 2017–2018.

14 varieties – from 35.8 to 39.1 mg/kg, in 12 varieties – from 42.4 to 45.7 mg/kg. The varieties with a high grain Zn content identified by the experiment results on average for all sites for two years, and the variability limits of the trait, depending on year and site, are presented in Table 3. The differences between the presented varieties on the average zinc content in grain correspond to error of the experiment.

Variety Novosibirskaya 16 on average had Zn content of 49.4 mg/kg, but this indicator was not stable, the trait value varied from 12.9 to 74.0 mg/kg, the variation coefficient was 35.5 %. Varieties Silach, Line 4-10-16, and Element 22 were more stable than Novosibirskaya 16. On average, the grain Zn content of studied varieties was 41.5 mg/kg.

Figure 4, *a* presents the distribution of varieties in Omsk by the grain Zn content. The highest content of this microelement for all varieties was noticed in this site in all years of research. According to the ranging of varieties by the grain Zn content three groups were distinguished: 21 varieties (51.6–63.7 mg/kg) were assigned to the first, 18 varieties (44.7–51.3 mg/kg) were assigned to the second, and 10 varieties with a relatively low microelement content (39.4–44.4 mg/kg) were assigned to the third. The significant differences on Zn content were noticed for varieties of the first and third groups as the contribution of the genotypic factor to expression of studied trait was 8 %. The highest zinc content of 49 varieties was observed in variety OmGAU 100 – 63.7 mg/kg on average for two years, and the lowest – in variety Lutescens 30 – 39.4 mg/kg.

In Shortandy, there were no significant differences among varieties on grain Zn content: the first group included 42 varieties (21.9–39.7 mg/kg), the second and third – seven varieties (18.2–21.5 mg/kg). Variety Lutescens 443 had the highest grain Zn content – 39.7 mg/kg, which is almost at the level of the lowest indicator of the trait in Omsk (see Fig. 4, *b*). Variety Lutescens 857 was characterized by low grain Zn content (only 18.2 mg/kg) in Shortandy.

Thus, varieties OmGAU 100 and Lutescens 443 selected in two experimental sites should also be included in the hybridization program for the improvement of wheat grain Zn content.

Table 3. Varieties with the highest Zn grain content (mg/kg) from experimental sites, on average of 2017–2018

Variety	Origin	X_{av}	Max	Min	Cv, %
Novosibirskaya 16	Novosibirsk (Rus)	49.4	74.0	12.9	35.5
Silach	Chelyabinsk (Rus)	48.4	70.3	22.2	25.4
Line 4-10-16	Karabalyk (Kaz)	47.2	61.4	20.6	23.4
Element 22	Omsk (Rus)	46.3	60.7	29.7	23.9
Lutescens 248/01	Shortandy (Kaz)	46.0	65.0	9.2	32.5
Average in experiment		41.5			
LSD_{05}		6.2			

Discussion

According to the World Health Organization, two billions of people worldwide are at risk of suffering from Zn deficiency (WHO, 2017). The regions where zinc deficiency is most common are Southeast Asia, southern Africa, and other developing countries. The main factors leading to Zn deficiency should first of all be attributed to insufficient consumption of

Zn with food in the regions due to low content of food with low Zn content in soil, drought climate and lack of moisture (http://cgon.rospotrebnadzor.ru/content/62/2683).

The health benefits of wheat grain and products for consumers are a strategic priority worldwide (Saleh et al., 2019). Biofortified wheat with higher Zn concentration has proven its positive effect on human health in India and Pakistan (List-
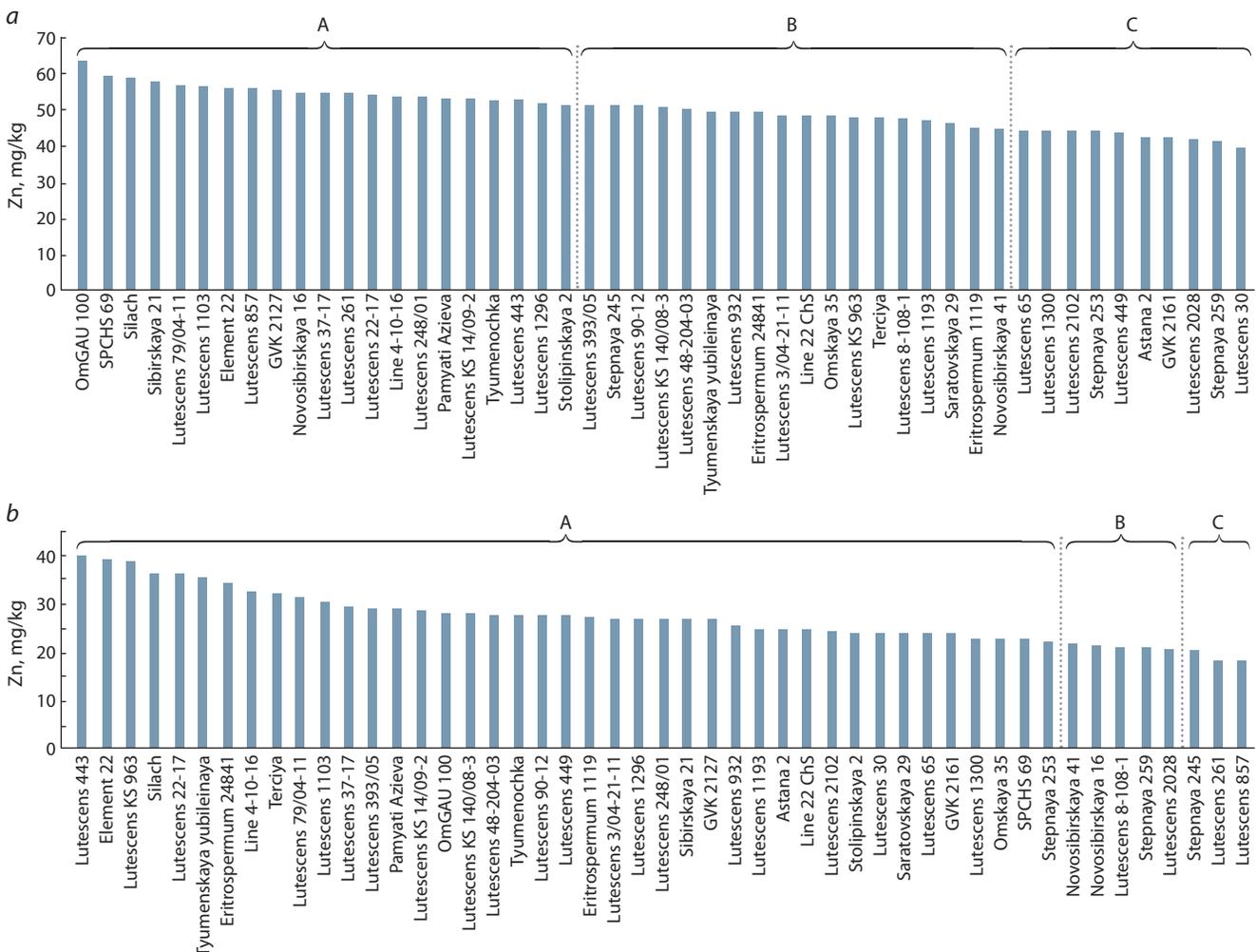
Fig. 4. Ranging of varieties of nursery KASIB-18 depending on Zn grain content, on average for 2017–2018: *a*, Omsk; *b*, Shortandy.

Different letters indicate homogeneous groups.

В.П. Шаманин, П. Флис, Т.В. Савин …
В.А. Чудинов, Т.В. Шелаева, А.И. Моргунов

Изменчивость содержания цинка в зерне
сортов пшеницы питомника КАСИБ

2021
25•5

man et al., 2019). Russia is one of the world's leading wheat exporters, exporting mainly to the Middle East and Africa, where there is a significant Zn shortage in the products of poor population eating mainly bread. Breeding wheat with enhanced levels of grain Zn provides a cost-effective, sustainable solution to the malnutrition problems in the developing world (Bouis et al., 2011).

Corresponding to our research, wheat grain grown in the experimental sites of Western Siberian and Ural regions has an increased Zn content on average from 44.1 to 50.4 mg/kg (see Table 2). In the sites Shortandy and Karabalyk, the grain Zn content was on average 26.8–37.3 mg/kg for two years of research. These data are consistent with the publication by J. Wang et al. (2020), according to which the Zn content in grain grown in Kazakhstan is on average 28.4 mg/kg. This is probably explained by dry climate in Kazakhstan. The plants need a mass flow of soil solution and ions for effective mineral nutrition, which depends primarily on the presence of moisture in the soil (Singh B. et al., 2005).

The ANOVA analysis revealed the influence of genotypic factor, soil, climatic, and weather conditions on the grain Zn accumulation in studied wheat varieties. The main contribution to the trait phenotypic variation was made by the 'Location' factor due to different Zn content in the soil and moisture availability in the experimental sites – 38.7 %. The influence of the 'Year' factor was 13.5 % and the 'Genotype' factor – 8.0 %.

In most sites of research, 2018 was more favorable on the available moisture in the soil compared to dry 2017. In 2018, grain Zn content varied from 35.0 mg/kg (Shortandy) to 53.2 mg/kg (Omsk). In 2017, under drought conditions, lower grain Zn concentration was noticed in Karabalyk (31.7 mg/kg), 18.5 mg/kg – in Shortandy, 36.4 mg/kg – in Chelyabinsk, and 47.6 mg/kg – in Omsk (see Table 2).

According to the two years research results of the nursery KASIB-18 in all experimental sites, sources of high Zn content were identified: Novosibirskaya 16 (49.4 mg/kg), Silach (48.4 mg/kg), Line 4-10-16 (47.2 mg/kg), Element 22 (46.3 mg/kg), and Lutescens 248/01 (46.0 mg/kg). These indicators are higher than the target ones of the international Harvest Plus program. This program allowed to increase the grain Zn content in the wheat varieties cultivated in India and Pakistan from 25 mg/kg on average to 37 mg/kg, approximately by 30–40 % (Singh R., Velu, 2017).

There was high variability of the trait in selected varieties (Cv = 23.4–35.5 %), which indicates a strong influence of environmental conditions on this trait (see Table 3). The study of the same varieties set under contrasting environmental conditions revealed differences among varieties and lines in grain Zn accumulation. Varieties OmGAU 100 (63.7 mg/kg) and Lutescens 443 (39.4 mg/kg) were characterized by the maximum grain Zn accumulation in Omsk and Shortandy, while varieties Lutescens 30 (39.4 mg/kg) and Lutescens 857 (18.2 mg/kg) were characterized by the minimum, respectively (see Fig. 4, *a*, *b*).

A.V. Volkov (2015) reported the positive effect of zinc fertilizers on yield, protein content, gluten, technological, and bread making qualities in agrochemical experiments on the spring wheat variety Zlata. In our experiments, the average positive correlation between the grain Zn and protein con-

tent (0.3 and 0.4) was revealed in Tyumen for both years of research, as protein content was low (on average 10.7 %). In the other five sites, the correlation was not found. There is no correlation between yield and grain Zn content (see Table 2). Probably, the grain Zn content is not an indicator of the necessity to use Zn-fertilizers. This issue requires additional research under the conditions of a specific region and different soil types.

Varieties Novosibirskaya 16, Silach, Element 22, OmGAU 100 included in the State Register of Breeding Achievements and cultivated on large areas under conditions of Western Siberian region, are able to form Zn content in grain exceeding 60–70 mg/kg (see Table 3, Fig. 4, *a*). The analysis of wheat grain from farmer's crops in the Omsk region also revealed a high grain Zn content. It is feasible to use it both in the domestic market for the production of bread and bakery products with functional properties, and for the export of wheat grain (Abugalieva et al., 2020). Our research confirms the high potential of Omsk, Novosibirsk, Tyumen, and Chelyabinsk regions for the production of grain with a high Zn content. It is advisable to use the grain of the best varieties on the Zn content to form separate batches for the production of "healthy" bread and for export. These are unused reserves that will contribute to the improvement of human health, especially for people with low incomes, as well as increase the export potential of the country.

Conclusion

According to the research results, the most favorable soil and climatic conditions to form grain wheat with a high Zn content were in the Omsk region. On average, during two years of research of the nursery KASIB-18 in Omsk, the grain Zn content of the studied varieties was 50.4 mg/kg, which is more than in other Russian experimental sites – Tyumen, Novosibirsk, and Chelyabinsk (44.1–44.8 mg/kg). In Kazakhstan, the average grain Zn content of wheat varieties was 37.3 mg/kg in Karabalyk and 26.8 mg/kg in Shortandy. Significant differences on the grain Zn content among the varieties indicate the possibility of breeding improvement of wheat on this trait. On average, for all sites, for two years of research, the highest grain Zn content was revealed in varieties Novosibirskaya 16 (49.4 mg/kg), Silach (48.4), Line 4-10-16 (47.2), Element 22 (46.3), and Lutescens 248/01 (46.0 mg/kg). In the conditions of Omsk, variety OmGAU 100 (63.7 mg/kg), and in Shortandy, under less favorable conditions, variety Lutescens 443 (39.7 mg/kg) were distinguished. These varieties should be included in the hybridization program for the improvement of wheat grain Zn content.

The main contribution (38.7 %) to the variability of the grain Zn content was made by soil and climate conditions of the region (the 'Location' factor). Significant influences of the 'Genotype' factor – 8.0 %, and the 'Year' factor – 13.5 % were revealed. There was no correlation between grain Zn content and yield, the correlation between grain Zn and protein content was revealed in Tyumen (–0.3 and –0.4). The potential opportunities for production of wheat grain with a high Zn content, which will be in demand for production of bread and pastry with functional properties, were identified in Western Siberian region.

References

Abugalieva A., Flis P., Shamanin V., Savin T., Morgounov A. Ionomic analysis of spring wheat grain produced in Kazakhstan and Russia. *Commun. Soil Sci. Plant. Anal.* 2020. Available at: https://www.tandfonline.com. DOI 10.1080/00103624.2020.1865398.

Abugaliyeva A.I., Savin T.V. The wheat introgressive form evaluation by grain biochemical and technological properties. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding.* 2018;22(3):353-362. DOI 10.18699/VJ18.371. (in Russian)

Azarenko Yu.A., Ermohin Yu.I., Aksenova Yu.V. Zinc in soils of agrocenosis of Omsk Region and efficiency of zinc fertilizers application. *Zemledelije = Agriculture.* 2019;2:13-17. DOI 10.24411/0044-3913-2019-10203. (in Russian)

Bhatta M., Shamanin V., Shepelev S., Baenziger S., Pozherukova V., Pototskaya I., Morgounov A. Marker-trait associations for enhancing agronomic performance, disease resistance, and grain quality in synthetic and bread wheat accessions in Western Siberia. *G3: Genes Genom. Genet.* 2019;9(12):4209-4222. DOI 10.1534/g3.119.400811.

Björck I., Östman E., Kristensen M., Anson N.M., Price R.K., Haenen G.R.M., Havenaar R., Knudsen K.E.B., Frid A., Mykkanen H., Welch R.W., Riccardi G. Cereal grains for nutrition and health benefits: overview of results from *in vitro*, animal and human studies in the HEALTHGRAIN project. *Trends Food Sci. Technol.* 2012;25(2):87-100.

Bouis H. Enrichment of food staples through plant breeding: a new strategy for fighting micronutrient malnutrition. *SCN News.* 1995;12:15-19. PMID: 12346314.

Bouis H.E., Hotz C., McClafferty B., Meenakshi J.V., Pfeiffer W.H. Biofortification: a new tool to reduce micronutrient malnutrition. *Food Nutr. Bull.* 2011;32:31S-40S.

Cakmak I., Torun A., Millet E., Feldman M., Fahim T., Korol A., Nevo E., Braun H.J., Ozkan H. *Triticum dicoccoides*: an important genetic resource for increasing zinc and iron concentration in modern cultivated wheat. *Soil Sci. Plant Nutr.* 2004;50:1047-1054.

Chen X.-P., Zhang Y.-Q., Tong Y.-P., Xue Y.-F., Liu D.-Y., Zhang W., Deng Y., Meng Q.-F., Yue S.-C., Yan P., Cui Z.-L., Shi X.-J., Guo S.-W., Sun Y.-X., Ye Y.-L., Wang Z.-H., Jia L.-L., Ma W.-Q., He M.-R., Zhang X.-Y., Kou C.-L., Li Y.-T., Tan D.-S., Cakmak I., Zhang F.-S., Zou C.-Q. Harvesting more grain zinc of wheat for human health. *Sci. Rep.* 2017;7:7016. DOI 10.1038/s41598-017-07484-2.

Genc Y., Verbyla A.P., Torun A.A., Cakmak I., Willsmore K., Wallwork H., McDonald G.K. Quantitative trait loci analysis of zinc efficiency and grain zinc concentration in wheat using whole genome average interval mapping. *Plant Soil.* 2009;314:49-66. DOI 10.1007/s11104-008-9704-3.

Gomez-Becerra H.F., Morgunov A.I., Abugalieva A.I. Evaluation of the germplasm through the Kazakhstan-Siberian network of spring wheat improvement: I. Genotype × environment interactions and site classification for grain yield and grain protein content. *Austr. J. Agric. Res.* 2007;4:649-660.

Gordeeva E., Shamanin V., Shoeva O., Kukoeva T., Morgounov A., Khlestkina E. The strategy for marker-assisted breeding of anthocyanin-rich spring bread wheat (*Triticum aestivum* L.) cultivars in Western Siberia. *Agronomy.* 2020;10:1603. DOI 10.3390/agronomy10101603.

Govindan V., Singh R.P., Crespo-Herrera L., Juliana Ph., Dreisigacker S., Valluru R., Stangoulis J., Sohu V.S., Mavi G.S., Mishra V.K., Balasubramaniam A., Chatrath R., Gupta V., Singh G.P., Joshi A.K. Genetic dissection of grain zinc concentration in spring wheat for mainstreaming biofortification in CIMMYT wheat breeding. *Sci. Rep.* 2018;8:13526. DOI 10.1038/s41598-018-31951-z.

Guttieri M.J., Seabourn B.W., Liu C., Baenziger P.S., Waters B.M. Distribution of cadmium, iron, and zinc in millstreams of hard winter wheat (*Triticum aestivum* L.). *J. Agric. Food Chem.* 2015;63:10681-10688. DOI 10.1021/acs.jafc.5b04337.

Khlestkina E.K., Shoeva O.Y., Gordeeva E.I., Otmakhova Y.S., Usenko N.I., Tikhonova M.A., Tenditnik M.V., Amstislavskaya T.G. Anthocyanins in wheat grain: genetic control, health benefit and breadmaking quality. In: Current Challenges in Plant Genetics, Genomics, Bioinformatics, and Biotechnology: Proceed. of the Fifth Int. Sci. Conf. PlantGen2019 (June 24–29, 2019, Novosibirsk, Russia). Novosibirsk, 2019;5-18. DOI 10.18699/ICG-PlantGen2019-02.

Krishnappa G., Singh A.M., Chaudhary S., Ahlawat A.K., Singh S.K., Shukla R.B., Jaiswa J.P., Singh G.P., Solanki I.S. Molecular mapping of the grain iron and zinc concentration, protein content and thousand kernel weight in wheat (*Triticum aestivum* L.). *PLoS One.* 2017;12(4):e0174972. DOI 10.1371/journal.pone.0174972.

Listman M., Guzman C., Palacios-Rojas N., Pfeiffer W.H., San Vicente F., Govindan V. Improving nutrition through biofortification: preharvest and postharvest technologies. *Cereal Foods World.* 2019;64(3). DOI 10.1094/CFW-64-3-0025.

Liu H., Wang Z.H., Li F., Li K., Yang N., Yang Y., Huang D., Liang D., Zhao H., Mao H., Liu J., Qiu W. Grain iron and zinc concentrations of wheat and their relationships to yield in major wheat production areas in China. *Field Crops Res.* 2014;156:151-160. DOI 10.1016/j.fcr.2013.11.011.

Methods of State Crop Variety Trial. Iss. 2: Cereals, legumes, maize, and fodder crops. Moscow, 1989. (in Russian)

Mitrofanova O.P., Khakimova A.G. New genetic resources in wheat breeding for an increased grain protein content. *Russ. J. Genet. Appl. Res.* 2017;7:477-487. DOI 10.1134/S2079059717040062.

Morgounov A., Gomez-Becerra H.F., Abugalieva A. Iron and zinc concentration in grain of spring bread wheat from Kazakhstan and Siberia. *Agromeridian.* 2006;1(2):5-16.

Morgunov A.I., Gomez-Becerra H.F., Abugalieva A.I., Dzhunusova M., Yessimbekova M.A., Muminjanov H., Zelenskiy Y., Ozturk L., Cakmak Y. Iron and zinc grain density in common wheat grown in Central Asia. *Euphytica.* 2007;155:193-203. DOI 10.1007/s10681-006-9321-2.

Murphy K.M., Reeves P.G., Jones S.S. Relationship between yield and mineral nutrient concentrations in historical and modern spring wheat cultivars. *Euphytica.* 2008;163:381-390. DOI 10.1007/s10681-008-9681-x.

Peleg Z., Cakmack I., Ozturk L., Yazici A., Jun Y., Budak H., Korol A.B., Fahima T., Saranga Y. Quantitative trait loci conferring grain mineral nutrient concentrations in durum wheat × wild emmer wheat RIL population. *Theor. Appl. Genet.* 2009;119:353-369. DOI 10.1007/s00122-009-1044-z.

Qi Y.T., Zhou S.N., Zhang Q., Yang L.X. The effect of foliar Zn application at grain filling stage on Zn bioavailability in grain fractions of modern winter wheat cultivars. *J. Agro Environ. Sci.* 2013;32:1085-1091.

Saleh A.S.M., Wang P., Wang N., Yang S., Xiao Z. Technologies for enhancement of bioactive components and potential health benefits of cereal and cereal-based foods: research advances and application challenges. *Crit. Rev. Food Sci. Nutr.* 2019;59(2):207-227. DOI 10.1080/10408398.2017.1363711.

Savin T.V., Abugaliyeva A.I., Cakmak I., Kozhakhmetov K. Mineral composition of wild relatives and introgressive forms in wheat selection. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding.* 2018;22(1):88-96. DOI 10.18699/VJ18.335. (in Russian)

Selyaninov G.T. Fundamentals of agroclimatic zoning of the USSR. In: Issues of Agroclimatic Zoning in the USSR. Moscow, 1958. (in Russian)

Singh B., Natesan S.K.A., Singh B.K., Usha K. Improving zinc efficiency of cereals under zinc deficiency. *Curr. Sci.* 2005;88:36-44.

Singh R.P., Velu G. Zinc-biofortified wheat: harnessing genetic diversity for improved nutritional quality. *Sci. Brief: Biofort.* 2017. Available at: http://www.harvestplus.org/sites/default/files/publications/ScienceBrief-Biofortification-1_ZincWheat_May2017.pdf

Uauy C., Distelfeld A., Fahima T., Blechl A., Dubcovsky J. A NAC gene regulating senescence improves grain protein, zinc, and iron

В.П. Шаманин, П. Флис, Т.В. Савин …
В.А. Чудинов, Т.В. Шелаева, А.И. Моргунов

Изменчивость содержания цинка в зерне
сортов пшеницы питомника КАСИБ

2021
25•5

content in wheat. *Science*. 2006;314(5803):1298-1301. DOI 10.1126/science.1133649.

Velu G., Crespo-Herrera L., Huert J., Payne T., Guzman C., Singh R.P. Assessing genetic diversity to breed competitive biofortified wheat with increased grain Zn and Fe concentrations. *Front. Plant Sci*. 2019;9:1971. DOI 10.3389/fpls.2018.01971.

Velu G., Ortiz-Monasterio I., Cakmak I., Hao Y., Singh R.P. Biofortification strategies to increase grain zinc and iron concentrations in wheat. *J. Cereal Sci*. 2014;59:365-372.

Velu G., Singh R.P., Huerta-Espino J., Peña R.J. Breeding for enhanced zinc and iron concentration in CIMMYT spring wheat germplasm. *Czech J. Genet. Plant Breed*. 2011;47:S174-S177. DOI 10.17221/3275-CJGPB.

Verma S.K., Kumar S., Sheikh I., Malik S., Mathpal P., Chugh V., Kumar S., Prasad R., Dhaliwal H.S. Transfer of useful variability of high grain iron and zinc from *Aegilops kotschyi* into wheat through seed irradiation approach. *Int. J. Radiat. Biol*. 2016;92(3):132-139. DOI 10.3109/09553002.2016.1135263.

Volkov A.V. The efficiency of various application methods, forms, and dozes of zinc fertilizers for spring wheat grown on sod-podzolic soils. Dr. Sci. Diss. (Biol.). Moscow, 2015. (in Russian)

Wang J.W., Kong F., Liu R., Fan Q., Zhang X. Zinc in wheat grain, processing, and food. Front. Nutr. 2020;7:124. DOI 10.3389/fnut.2020.00124.

Waters B.M., Uauy C., Dubcovsky J., Grusak M.A. Wheat (*Triticum aestivum*) NAM proteins regulate the translocation of iron, zinc, and nitrogen compounds from vegetative tissues to grain. *J. Exp. Bot*. 2009;60:4263-4274.

Welch R.M., Graham R.D. Breeding crops for enhanced micronutrient content. *Plant Soil*. 2002;245:205-214.

WHO. The World Health Report. Geneva: World Health Organization. Accessed June 1, 2017.

Xu Y., Diaoguo A., Dongcheng L., Aimin Z., Hongxing X., Bin L. Molecular mapping of QTLs for grain zinc, iron and protein concentration of wheat across two environments. *Field Crops Res*. 2012;138:57-62.

Zou C., Du Y., Rashid A., Ram H., Savasli E., Pieterse P.J., Ortiz-Monasterio I., Yazici A., Kaur C., Mahmood K., Singh S., Le Roux M.R., Kuang W., Onder O., Kalayci M., Cakmak I. Simultaneous biofortification of wheat with zinc, iodine, selenium, and iron through foliar treatment of a micronutrient cocktail in six countries. *J. Agric. Food Chem*. 2019;67(29):8096-8106. DOI 10.1021/acs.jafc.9b01829.

ORCID ID

V.P. Shamanin orcid.org/0000-0003-4767-9957
P. Flis orcid.org/0000-0002-5529-7599
T.V. Savin orcid.org/0000-0002-3550-647x
S.S. Shepelev orcid.org/0000-0002-4282-8725
O.G. Kuzmin orcid.org/0000-0003-4918-0025
A.S. Chursin orcid.org/0000-0001-6797-6145

I.V. Pototskaya orcid.org/0000-0003-3574-2875
I.E. Likhenko orcid.org/0000-0002-0305-1036
I.Yu. Kushnirenko orcid.org/0000-0001-5463-9932
A.A. Kazak orcid.org/0000-0002-0563-3806
V.A. Chudinov orcid.org/0000-0002-2639-9000
T.V. Shelaeva orcid.org/0000-0001-7328-572X
A.I. Morgounov orcid.org/0000-0001-7082-5655

Mechanosensitive molecular interactions in atherogenic regions of the arteries: development of atherosclerosis

E.L. Mishchenko[1] ✉, A.M. Mishchenko[2], V.A. Ivanisenko[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
✉ elmish@bionet.nsc.ru

Abstract. A terrible disease of the cardiovascular system, atherosclerosis, develops in the areas of bends and branches of arteries, where the direction and modulus of the blood flow velocity vector change, and consequently so does the mechanical effect on endothelial cells in contact with the blood flow. The review focuses on topical research studies on the development of atherosclerosis – mechanobiochemical events that transform the proatherogenic mechanical stimulus of blood flow – low and low/oscillatory arterial wall shear stress in the chains of biochemical reactions in endothelial cells, leading to the expression of specific proteins that cause the progression of the pathological process. The stages of atherogenesis, systemic risk factors for atherogenesis and its important hemodynamic factor, low and low/oscillatory wall shear stress exerted by blood flow on the endothelial cells lining the arterial walls, have been described. The interactions of cell adhesion molecules responsible for the development of atherosclerosis under low and low/oscillating shear stress conditions have been demonstrated. The activation of the regulator of the expression of cell adhesion molecules, the transcription factor NF-κB, and the factors regulating its activation under these conditions have been described. Mechanosensitive signaling pathways leading to the expression of NF-κB in endothelial cells have been described. Studies of the mechanobiochemical signaling pathways and interactions involved in the progression of atherosclerosis provide valuable information for the development of approaches that delay or block the development of this disease.

Key words: atherogenesis; shear stress; transcription factor NF-κB; RelA expression; mechanosensitive receptors; cell adhesion molecules; signaling pathways; mechanotransduction.

Механочувствительные молекулярные взаимодействия в атерогенных районах артерий: развитие атеросклероза

Е.Л. Мищенко[1] ✉, А.М. Мищенко[2], В.А. Иванисенко[1]

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

Аннотация. Атеросклероз, грозное заболевание сердечно-сосудистой системы, развивается в местах изгибов и разветвлений артерий, где меняются направление и модуль вектора скорости тока крови, а следовательно, механическое воздействие на контактирующие с током крови эндотелиальные клетки. Обзор посвящен актуальным исследованиям развития атеросклероза: механобиохимическим событиям, преобразующим проатерогенный механический стимул тока крови – низкое и низкое/осциллирующее напряжение сдвига, оказываемое на стенки артерий, – в цепи биохимических реакций в эндотелиальных клетках, приводящих к экспрессии специфичных белков, вызывающих прогрессирование патологического процесса. Описаны стадии, системные факторы риска, а также важный гемодинамический фактор атерогенеза: низкое и низкое/осциллирующее напряжение сдвига, оказываемое током крови на эндотелиальные клетки, выстилающие стенки артерий. Показаны взаимодействия молекул клеточной адгезии, ответственные за развитие атеросклероза в условиях низкого и низкого/осциллирующего напряжения сдвига. Описаны активация регулятора экспрессии молекул клеточной адгезии – транскрипционного фактора NF-κB – и факторы, контролирующие его активацию в этих условиях. Описаны механочувствительные сигнальные пути, приводящие к экспрессии NF-κB в эндотелиальных клетках. Исследования механобиохимических сигнальных путей и взаимодействий, вовлеченных в прогрессирование атеросклероза, необходимы для разработки подходов, задерживающих или блокирующих развитие заболевания.

Ключевые слова: атерогенез; напряжение сдвига; транскрипционный фактор NF-κB; экспрессия RelA; механочувствительные рецепторы; молекулы клеточной адгезии; сигнальные пути; механотрансдукция.

Е.Л. Мищенко, А.М. Мищенко
В.А. Иванисенко

Механочувствительные молекулярные взаимодействия
в атерогенных районах артерий: развитие атеросклероза

2021
25•5

Risk factors and stages of atherogenesis. Shear stress is an important haemodynamic factor in atherogenesis

Nowadays, cardiovascular disease is a major public health issue. Moreover, atherosclerosis is one of the most common pathologies of the cardiovascular system. The systemic risk factors for the development of atherosclerosis include age, hypertension, diabetes mellitus, smoking, low physical activity, fatty diet, renal failure, increased level of fibrinogen, low-density lipoproteins, cholesterol and blood plasma C-reactive protein (Virani et al., 2020). The level of low-density lipoproteins (LDLs) is classified into a separate group of factors that account for the atherogenicity of the subfraction profile of apo-B-containing lipoproteins (Chang et al., 2017; Ozerova et al., 2018). The penetration of blood plasma LDLs through the endothelium in athero-susceptible areas of the arteries and their retention and accumulation in the extracellular matrix (ECM) of the subendothelial space initiates atherogenesis. LDLs are retained in the intima (mainly due to interaction with proteoglycans), undergo oxidation (formation of oxLDLs) and cause an inflammatory response – the infiltration of circulating blood monocytes into the intima. In the intima, monocytes differentiate into macrophages, uptake oxLDLs and become foam cells (Libby et al., 2019).

The development of atherosclerosis occurs in the following stages: (*i*) adaptive intimal thickening, (*ii*) formation of fatty streaks, (*iii*) pathological intimal thickening (PIT), (*iv*) early fibroatheroma and (*v*) late fibroatheroma. In stage (*i*), smooth muscle cells (SMCs) of the media migrate to the intima and secrete proteoglycans. The formation of fatty streaks in stage (*ii*) is accompanied by the accumulation of foamy cells (macrophages loaded with lipids) in the intima. Lipid-loaded SMCs are less represented. The PIT process (*iii*) occurs with and without the infiltration of macrophages. In both cases, SMCs and extracellular lipid pools are present in the intima. The accumulation of SMCs occurs towards the lumen of the artery, and lipid pools accumulate close to the media. The formation of a fibrous cap that covers the necrotic core occurs in the later stages of development of the atherosclerotic lesions, including early and late fibroatheroma (*iv*, *v*). The cap includes SMCs, infiltrated macrophages, T-lymphocytes, as well as collagens and proteoglycans of the extracellular matrix. Programmed cell death, via apoptosis and necroptosis, plays an essential role in early fibroatheroma (*iv*) with the formation of foci of necrosis and cholesterol crystals. In late fibroatheroma (*v*), an extensive necrotic core that consists of cellular debris and a large number of crystals of free cholesterol and its esters is formed (Otsuka et al., 2015).

The molecular and genetic processes of atherogenesis remain unclear. Numerous haemodynamic studies have shown that, based on systemic risk factors, atherosclerosis develops mainly in the bends and branching of the arteries, where there is a change in the nature of the blood flow (Cecchi et al., 2011; Morbiducci et al., 2016; Zou et al., 2016). The haemodynamic characteristics of the effect of blood flow on the vessel walls are wall shear stress (WSS), hydrostatic pressure and cyclic deformation. WSS is the friction force that occurs when flowing blood comes into contact with the inner wall of the

artery. WSS on the arterial wall is described by (Ku et al., 1985):

$$\vec{\tau}_w = \mu \frac{d\vec{V}(t)}{dr},$$

where μ is the blood viscosity index, $\vec{V}(t)$ is the blood flow rate parallel to the vessel wall at time t and r is the radial coordinate.

Studies conducted on animal models and observing patients revealed a regular maximum thickening of the intima and the formation of atherosclerotic plaques in areas with low WSS (< 10 dyn/cm^2 in humans) and low/oscillatory WSS (with a deviation of the instantaneous WSS vector from its average direction). Such damage was minimal in areas of high WSS (> 25 dyn/cm^2 in humans). High values of WSS were realized in the rectilinear sections of the arteries with laminar blood flow. In the areas of branching and bending of the arteries near the walls, vortex flows were formed, leading to mechanical stress on the walls, which was accompanied by pathological effects. These flows were characterised by low and low/oscillatory WSS (Cecchi et al., 2011; Morbiducci et al., 2016). A study conducted on isolated segments of blood vessels through which LDLs flowed demonstrated that the transport of LDLs into the vascular wall increased with a decrease in WSS and, on the contrary, decreased with an increase in WSS (Colic et al., 2015). Patient-specific modelling of the subendothelial accumulation of LDLs in the stenotic right coronary artery also showed an inverse relationship between the distribution of WSS and the accumulation of LDLs (Sakellarios et al., 2013). The zone of recirculating flow and low WSS corresponded to the maximum accumulation of LDLs, and in areas of high WSS, the accumulation of LDLs was low.

To localise the segments of arteries with low and low/oscillatory WSS and monitor the transformation of atherosclerotic plaques into a stable or unstable phenotype, computer modelling of blood flow in the vessels is being developed. This will facilitate the identification of patient-specific fields and gradients of blood flow rates depending on the geometry of the vessels (Soulis et al., 2006; Timmins et al., 2015, 2017; Hung et al., 2016). To solve these issues, the Navier–Stokes equations for an incompressible viscous fluid are used. To reconstruct the geometric shape of the vessels, intravascular ultrasound methods are used as well as X-ray microcomputer tomography with the use of contrast agents (Nebuloni et al., 2013; Xing et al., 2016). The ANSYS Fluent, OpenFOAM, FLUENT 6.0 and other software packages are widely used to conduct calculations via mathematical models of stationary and unsteady blood flows in various areas of the arteries. Computer modelling of the distribution of WSS, which accounts for patient-specific data on the geometry of blood vessels, is of high value for clinical practice.

The molecules of cell adhesion and their interactions in the early stage of atherogenesis

OxLDLs in the subendothelial space as well as low and low/oscillatory WSS cause pro-inflammatory activation of endothelial cells (ECs), which leads to the rolling of leukocytes in the circulating blood flow to the endothelium, their adhesion

and transendothelial migration (TEM). The mediators of these critical events of the early stage of atherogenesis are cell adhesion molecules. These molecules are expressed on the surface of ECs and circulating blood cells (monocytes or leukocytes and platelets) and include platelet endothelial cell adhesion molecule-1 (PECAM-1); intercellular adhesion molecule-1 and -2 (ICAM-1, ICAM-2); vascular cell adhesion molecule-1 (VCAM-1); E-, L- and P-selectins; vascular endothelial (VE) cadherin; β1 and β2 integrins; proline-rich glycoprotein CD99 and junctional adhesion molecule-A (JAM-A).

Selectins (transmembrane glycoproteins) are expressed on the surface of ECs (E-selectin, ELAM-1, P-selectin), leukocytes (L-selectin) and platelets (P-selectin) (Carlos, Harlan, 1994). Early experiments conducted in a flow chamber with a laminar flow of monocytes on an EC monolayer (physiologically low WSS, the use of functionally blocking monoclonal antibodies to L-, P-, E-selectin, ICAM-1, VCAM-1, β1 and β2 integrins) showed that the rolling of the monocytes to the ECs, weak, reversible contact of the monocytes with ECs (initial adhesion) and slowing down of the rate along the endothelium determine the interactions of the L-selectin of the monocytes with glycoprotein ligands of the ECs and, to a lesser extent, the P-selectin of the ECs with the glycoprotein ligand PSGL-1 of the monocytes. E-selectin is not involved in the process (Luscinskas et al., 1994, 1996). Firm, irreversible adhesion of the leukocytes to the ECs occurs during the interaction of the leukocyte α4β1 (VLA-4) integrin with the endothelial immunoglobulin VCAM-1 (Luscinskas et al., 1994; Huo, Ley, 2001) and the interaction of the leukocyte αLβ2 (LFA-1, CD11a/CD18; Mac-1, CD11b/CD18) integrins with the endothelial immunoglobulin ICAM-1 (Luscinskas et al., 1994; Sigal et al., 2000; Huo, Ley, 2001).

OxLDLs and lysophosphatidylcholine (a component of oxLDLs) induce the expression of ICAM-1 and VCAM-1 on the surface of cultured ECs and stimulate monocyte adhesion (Kume et al., 1992; Amberger et al., 1997). A physiologically low WSS, created by the flow of leukocytes (monocytes, neutrophils, lymphoblasts, lymphocytes) on the EC monolayer, generates upward docking structures on the endothelium that contain ICAM-1 and VCAM-1 clusters within 1–2 minutes. These clusters surround the leukocytes and function as an anchor for them (Barreiro et al., 2002; Carman et al., 2003). In turn, the structures of VCAM-1 and ICAM-1 that surround the leukocytes stimulate the formation of lateral linear tracks of leukocyte β1 (VLA-4) and β2 (LFA-1) integrins that are oriented parallel to the ICAM-1 and VCAM-1 clusters (Carman, Springer, 2004). Moreover, 90 % of the leukocytes that are surrounded by the VCAM-1 and ICAM-1 clusters transmigrate to the subendothelial space, and the suppression of VCAM-1 and ICAM-1 via inhibitors (BARTA-AM, colchicine, toxin-B) significantly suppresses TEM. Regardless of the TEM pathway (the paracellular pathway between the ECs, the transcellular pathway through the ECs), the TEM process is associated with the formation of a cupped traction structure by the VCAM-1/VLA-4 and ICAM-1/LFA-1 interactions of endothelial and leukocyte cells, which guide and facilitate TEM (Carman et al., 2003).

VE-cadherin plays an important role in the TEM of leukocytes. VE-cadherin is only expressed in ECs, localised mainly in the intercellular contacts and plays an important role in the intercellular adhesion and barrier functions of the ECs (Garrett et al., 2017). The adhesion of leukocytes to the endothelium in the lateral intercellular contacts, preceding TEM, induces the formation of gaps in the intercellular distribution of VE-cadherin and the components of the VE-cadherin complex (α-, β-, Υ-catenin, p120-catenin) at the sites of adhesion or transmigration of the leukocytes. The gaps are formed as a result of lateral displacement of VE-cadherin in the membrane and facilitate the TEM of the leukocytes. Following the completion of TEM, VE-cadherin moves in the opposite direction and closes the gaps (curtain opening and closing effect) (Allport et al., 2000; Shaw et al., 2001). The lateral displacement of VE-cadherin in the membrane most likely occurs due to the destabilisation of the VE-cadherin bond with the actin cytoskeleton by the following mechanism: the ICAM-1 and VCAM-1 clusters in the sites of intercellular adhesion of the leukocytes to the ECs induce the intracellular activation of the Src and Pyk2 tyrosine kinases and the phosphorylation of Tyr658 and Tyr731 of the cytoplasmic domain of VE-cadherin, which are involved in the low-affinity binding of VE-cadherin to p120- and β-catenin, respectively. The weakening of these interactions disrupts the VE-cadherin bond with the actin cytoskeleton, destabilises the VE-cadherin or VE-cadherin cell-cell interactions and facilitates the lateral movement of the phosphorylated VE-cadherin in the membrane (Allingham et al., 2007). p120-catenin regulates the phosphorylation of VE-cadherin and the paracellular TEM of leukocytes via a competition mechanism with the activated Src and Pyk2 tyrosine kinases: the overexpression of p120-catenin in the ECs leads to the absence of gaps in the distribution of VE-cadherin and the blocking of the TEM of the leukocytes (Alcaide et al., 2008).

The glycoprotein PECAM-1 (CD31) plays an important role in the TEM of leukocytes. In ECs, PECAM-1 is mainly localised in the intercellular contacts. The homophilic interactions of this protein with adjacent ECs occur through the extracellular Ig-like domains IgD1 and IgD2 (Paddock et al., 2016). PECAM-1 is also present in a distinct membrane-vesicular recycling compartment adjacent to the lateral border membrane of the ECs (Mamdouh et al., 2009). In the resting endothelium (in the absence of adhesive leukocytes), there is constitutive membrane traffic between the lateral cell border and the membrane-vesicular compartment (known as the lateral border recycling compartment, LBRC) (Mamdouh et al., 2003, 2008). In the presence of adhesive leukocytes on the endothelium, directed kinesin-dependent migration of the PECAM-1-bearing LBRC membrane along the microtubules of the ECs to the sites of para- and transmigration of the leukocytes occurs, in addition to the surrounding of the leukocytes by the LBRC membrane. The LBRC provides non-ligated PECAM-1 and CD99 in the ECs to interact with the homophilic ligands (PECAM-1 and CD99, respectively) in the leukocytes and initiates the signals for further recruitment of the LBRC as the leukocytes move through the endothelial layer. The antibodies to PECAM-1 and CD99 block the TEM of the leukocytes (Mamdouh et al., 2008, 2009). The recruitment of the LBRC to paracellular TEM sites precedes the formation of gaps in the intercellular distribution of VE-cadherin and is necessary for the formation of these gaps (Gonzalez et al., 2016). In ApoE−/− PECAM-1−/− mice, the load of plaques in the areas of the carotid artery with a low and low/oscillatory

Е.Л. Мищенко, А.М. Мищенко
В.А. Иванисенко

Механочувствительные молекулярные взаимодействия
в атерогенных районах артерий: развитие атеросклероза

2021
25•5

WSS was significantly less than in the control ApoE[−/−] mice (Harrison et al., 2013). A study on the relationship of single nucleotide polymorphisms (Val125Leu, exon 3; Asn563Ser, exon 8; Arg670Gly, exon 12) in functionally important domains of PECAM-1 in patients who were at risk of developing coronary heart disease and myocardial infarction showed that Arg670Gly substitution can be a homozygous protector for the development of myocardial infarction. This substitution is localised close to Tyr663, the phosphorylation of which, under low WSS conditions, initiates the signaling pathway of activation of the key transcription factor NF-κB for the expression of cell adhesion molecules (Sahebkar et al., 2013). Val125Leu and Asn563Ser substitutions are not associated with the risk of coronary heart disease (Xia et al., 2015).

Integrins also play an important role in cell adhesion. The integrins are a large family of receptors that are localised in the plasma membrane and consist of 18 α and 8 β subunits that form 24 different heterodimers. The extracellular domains of the integrins interact with ECM proteins (collagens, CL; fibronectin, FN; laminins, LN; vitronectin, VN; etc.) and ligands (for example, VCAM-1) on the surface of other types of cells, causing cell-substratum or cell-cell adhesion. The integrin-ligand interactions induce the activation of a variety of signaling pathways that modulate cellular behaviour, including proliferation, shape, motility, survival or apoptosis, differentiation, protein phosphorylation, cytoskeleton organisation and gene expression. Many integrins are expressed in an inactive state on the cell surface since the membrane-proximal highly conserved sequences of the cytoplasmic domains of the α and β subunits form a structural constraint that locks the conformation of integrins in an inactive, low-affinity state. The activation of integrins is often induced by intracellular signals and regulatory factors that act on the cytoplasmic domains, as well as phosphorylation. This alters the affinity of integrins for ligands through conformational changes in their extracellular domains, as well as clustering (Hynes, 2002).

The adhesion of leukocytes to the endothelium and their infiltration into the subendothelial space is enhanced by cytokines, chemokines and other factors. Thus, monocytic chemotactic protein-1 (MCP-1) (produced by macrophages and vascular wall cells) and interleukin-8 (IL-8) (produced by macrophages) are involved in the delay of peripheral circulation monocytes and their adhesion and migration into the arterial intima via interaction with monocyte receptors belonging to the CCR2 and CXCR2 types, respectively (Peters, Charo, 2001; Charo, Taubman, 2004). IL-9, secreted mainly by CD45[+]CD3[−]CD19[−] leukocytes in АпоE[−/−] mice, stimulates the expression of VCAM-1 in the ECs of the aorta of mice through interaction with the IL-9 receptor (IL-9R) and activation (phosphorylation) of the signaling protein and transcription activator STAT3 (Zhang et al., 2015). In the atherosclerotic aortas of АпоE[−/−] mice, a high level of IL-17A expression was observed, as well as numerous IL-17A-producing CD4[+] T helper 17 (Th17) and γδ[+] T cells. IL-17A initiates the production of several cytokines and chemokines by aortic cells, in particular, the pro-inflammatory chemokine CXCL1, which activates peripheral circulation monocytes and stimulates their adhesion and migration to the aortic wall (Erbel et al., 2014). Pro-atherogenic IL-17C, expressed mainly by aortic SMCs in АпоE[−/−] mice, is involved in the recruitment of T cells

and macrophages into the aortic wall (Butcher et al., 2016). TGF-β-(H$_2$O$_2$-) inducible clone 5 (His-5), expressed on the surface of ECs and SMCs, is involved in the formation of structures that are similar to the microvilli on the surface of ECs, which enhance the adhesion of monocytes to the ECs (Arita-Okubo et al., 2015).

Transcription factor NF-κB is a key regulator of the gene expression of cell adhesion molecules under conditions of physiologically low and low/oscillatory wall shear stress

The transcription factor NF-κB positively regulates the expression of cell adhesion molecules with the participation of other transcription factors and coactivators. The *VCAM-1* gene promoter has two NF-κB sites that are required for transcription activation (Neish et al., 1992). The C/EBP and NF-κB sites were identified in the *ICAM-1* gene promoter, and mutations in the latter completely suppress the activation of the *ICAM-1* promoter (Ledebur et al., 1995). The *ELAM-1* gene promoter includes a CRE/ATF site, three NF-κB sites and three HMG I(Y) sites; two of the HMG I(Y) sites are located within the NF-κB sites (the interaction of HMG I(Y) and NF-κB with small and large DNA grooves, respectively). All three NF-κB sites are required for promoter activation and enhance the affinity of NF-κB and ATF-2 for the promoter (Whitley et al., 1994). The promoter of the *MCP-1* gene (monocyte chemoattractant protein-1) contains the NF-κB site and the AP-1 site, which are necessary for maximum induction of the promoter (Martin et al., 1997). The core element GAGACC (SSRE) was identified in the pro-atherogenic platelet growth factor (*PDGF*) promoter that stimulates the proliferation and migration of SMCs (Resnick et al., 1993); it was found to interact with NF-κB (Davis et al., 2003).

In ECs, the most common p50/p65 (RelA) heterodimer is NF-κB. It is well known that, in the cytoplasm, latent NF-κB is associated with an inhibitor of IκB (mainly IκBα) and is inactive. When ECs are stimulated by cytokines TNFα, IL-1 or bacterial lipopolysaccharide (LPS) that interact with specific EC receptors, signaling pathways are activated, leading to the activation of the IκB kinase (IKK complex), which specifically phosphorylates IκB. After the ubiquitination of IκB and its proteasomal degradation, free NF-κB is translocated into cell nuclei (factor activation), interacts with elements of the DNA major groove and activates the expression of genes involved in the immune response, cell survival, carcinogenesis and inflammation (with the participation of other transcription factors and coactivators) (Oeckinghaus et al., 2011; Yu et al., 2015). Early studies demonstrated that active NF-κB and ICAM-1 were present in ECs, macrophages and SMCs in atherosclerotic plaques of arteries of deceased patients but were absent in the intima or media of the healthy arteries of those patients (Brand et al., 1996). OxLDL triggers in the ECs *in vitro* and *in vivo* signaling pathways for IKK complex activation; this includes the activation of focal adhesion kinase (FAK) and ribosomal S6 kinase (RSK) and leads to NF-κB activation, VCAM-1 expression and monocyte adhesion to ECs (Yurdagul et al., 2016). Using the model flow channel system (Frangos et al., 1985) that is created by a laminar pulsating flow of the culture medium containing monocytes, in which an average physiologically low and uniformly distributed WSS acts on

the EC monolayer, it was shown that a mechanical stimulus (physiological low WSS) activates the IKK complex and NF-κB and stimulates VCAM-1 expression and monocyte adhesion to the ECs (Mohan et al., 1999). A model spatial gradient of averaged low/oscillatory WSS, close to the gradient of WSS in arterial branchings and bends, resulted in a more efficient activation of NF-κB compared to uniformly distributed low WSS (Nagel et al., 1999).

High activation of NF-κB was observed in ECs that were cultivated based on a model of the calculated WSS profile in a site of the human carotid sinus (atherogenic region, averaged low/oscillatory WSS) compared with a model of the calculated WSS profile in the distal segment of the carotid artery bifurcation (atheroprotective region, high average WSS) or cells at rest (Dai et al., 2004). A study of the transcriptional expression of the VCAM-1 endothelium from the nature of haemodynamics and calculated profiles of WSS in the segments of the left and right coronary arteries showed that low/oscillatory WSS at the outer wall of the bifurcation of the left anterior descending artery (atherogenic region) induces a significantly higher expression of VCAM-1 compared to a straight region of the right artery (laminar flow, high WSS), in which atherosclerotic plaques are largely unformed (O'Keeffe et al., 2009). The mapping of the endothelial expression of NF-κB/IκB and the activation of NF-κB in areas of the aortic arch of mice with high (high probability, HP; internal arch of the aortic arch) and low (low probability, LP; external arch of the aortic arch) probability of atherosclerotic plaque formation demonstrated that in the HP region with low/oscillatory WSS (Suo et al., 2007), the expression of NF-κB/IκB and the activation of NF-κB significantly exceeded those in the LP region (Hajra et al., 2000). S. Cuhlmann et al. (2011) also detected increased expression of the RelA subunit NF-κB and a higher nuclear localisation compared with the LP region in the HP region of the aortic arch of mice. Moreover, this was correlated with increased expression of VCAM-1 and the accumulation of CD68[+] macrophages in the HP region.

The gap-junction protein connexin 40 (Cx40), expressed in ECs under low WSS conditions, whose cytoplasmic C-terminus interacts with IκBα and inhibits its phosphorylation, is a negative regulator of NF-κB activation in areas with low WSS (Denis et al., 2017). Conversely, mechanosensitive phosphatidic acid phosphatase (PPAP2B) regulates NF-κB activity, expression of adhesion proteins, monocyte adhesion and TEM under high WSS. PPAP2B shows increased expression in ECs under atheroprotective flow characteristics (high WSS, laminar flow) and shows decreased expression under atherogenic characteristics (low WSS, disturbed flow). PPAP2B hydrolyses lysophosphatidic acid (LPA), a pro-atherogenic and thrombogenic glycerophospholipid in the blood and, thereby, blocks signaling pathways activated by the interaction of LPA with cellular LPA receptors (LPAR). LPA-LPAR1 interaction activates the Rho kinase-NF-κB signaling pathway and the subsequent transcriptional expression of ICAM-1, VCAM-1 and E-selectin in ECs (Shimada et al., 2010; Wu et al., 2015).

LPA-LPAR1/2 interaction leads to an increase in the contractility of ECs and the permeability of the endothelial monolayer (Wu et al., 2015). LPA is an extracellular signaling molecule that is capable of interacting with at least six G pro-tein-coupled cellular receptors, initiating intracellular signaling cascades (Yung et al., 2014) and exerting multiple effects on blood cells (platelets, monocytes) and vascular wall cells (ECs, SMCs). Therefore, in platelets, LPA induces a change in the shape, aggregation, and formation of platelet-monocyte aggregates. In ECs, LPA induces cell migration, expression of VCAM-1, ICAM-1, E-selectin, chemokines (CXCL1), formation of actin stress fibres and cell contraction. In SMCs, LPA induces cell contraction, migration and proliferation. Additionally, LPA accumulates in the lipid-rich core of atherosclerotic plaques, and, when ruptured, enters the bloodstream and activates platelets, leading to the formation of blood clots (Schober, Siess, 2012). Endothelial NO synthase (eNOS) is also expressed under physiologically high WSS. Under these conditions, the eNOS promoter is activated through the interaction of NF-κB with the GAGACC element (Davis et al., 2003). eNOS produces NO, an inducer of IκBα expression. NO stabilises the NF-κB·IκBα heterotrimer in the cytoplasm and induces IκBα translocation into cell nuclei, leading to the inactivation of NF-κB and the termination of NF-κB-mediated transcription (Spiecker et al., 1997). Moreover, NO inhibits IκB kinase (IKK complex), a positive regulator of NF-κB activation (Yurdagul et al., 2013).

Mechanosensitive signaling pathways that control the transcriptional expression of the RelA subunit of NF-κB in endothelial cells

The WSS, created by the blood flow to the walls of the arteries, activates mechanosensitive signaling pathways in the ECs. A mechanosensor localised in the cell membrane perceives a mechanical stimulus (WSS) and triggers intracellular signaling that activates specific transcription factors, which regulate the transcriptional expression of proteins. The internal bend of the aortic arch and the areas of arterial bifurcation, where low and low/oscillatory WSS are realized, are associated with the localisation of atheroma. In the same areas, increased expression of active c-Jun N-terminal kinase 1 (JNK1), which belongs to mitogen-activated protein kinases (MAPK), was revealed (Zakkar et al., 2008), as well as increased expression and nuclear localisation of the RelA subunit of NF-κB (Cuhlmann et al., 2011). The hypothesis that physiologically low WSS regulates the expression of RelA and NF-κB target genes (*VCAM-1* and others) through the JNK1-dependent pathway was proved by modelling the haemodynamics of the carotid artery in mice. The implantation of a tapered cuff around the mouse carotid artery generates a laminar blood flow with a low velocity and low WSS upstream of the cuff and disturbed, low velocity and low/oscillatory WSS downstream of the cuff (Cheng et al., 2006; Xing et al., 2016). This approach demonstrated that increased expression of RelA and activated JNK1 is realised in regions with low and low/oscillatory WSS. The transcription factor ATF2 is activated by JNK1 (phosphorylation) and interacts with the RelA promoter sites to activate the promoter. Thus, in areas with low/oscillatory WSS, the JNK-ATF2-RelA signaling pathway is implemented, which stimulates the expression of RelA and NF-κB target genes (*VCAM-1*). The pathway also stimulates the accumulation of CD68[+] macrophages, which are an indicator of the development of arterial inflammation (Cuhlmann et al., 2011).

Е.Л. Мищенко, А.М. Мищенко
В.А. Иванисенко

Механочувствительные молекулярные взаимодействия
в атерогенных районах артерий: развитие атеросклероза

2021
25 • 5

The signaling pathway for the transcriptional expression of RelA upstream of JNK1 includes the activation of integrins, which is initiated by the stimulation of the cell surface mechanosensory complex and an integrin-dependent signaling cascade, leading to the activation of JNK1. The cell surface mechanosensory complex includes the PECAM-1 and VE-cadherin receptors and the vascular endothelial growth factor 2 receptor (VEGFR2), which belongs to the receptor tyrosine kinase subfamily (Tzima et al., 2005). In this complex, PECAM-1 is a key mechanosensitive signaling molecule that perceives a mechanical signal and converts it into a chain of intracellular biochemical reactions. WSS, acting on the extracellular domain of PECAM-1, affects the conformation of the cytoplasmic domain of PECAM-1 and the availability of Tyr663 and Tyr686 of this domain for phosphorylation. Phosphorylation is conducted by membrane-bound Fyn tyrosine kinase (family of Src tyrosine kinases), which is localised in intercellular contacts near PECAM-1 (Chiu et al., 2008). PECAM-1/PECAM-1 intercellular EC interactions through the extracellular domains of PECAM-1 (Paddock et al., 2016) are required for efficient phosphorylation and the triggering of intracellular mechanosensitive signaling (Chiu et al., 2008; Snyder et al., 2017). Activated PECAM-1 and VE-cadherin (which functions as an adapter and is associated with VEGFR2) facilitate the phosphorylation of VEGFR2 at Tyr801 and Tyr1175 by the Src tyrosine kinase. In turn, the phosphorylated VEGFR2, through direct interaction with the regulatory p85 subunit of the phosphatidylinositol-3-OH kinase (PI(3)K), phosphorylates PI(3)K (Tzima et al., 2005).

The activated PI(3)K stimulates the conformational activation of integrins through the conservative pathway of the association of phosphatidylinositol-3,4,5 triphosphate (PI(3)K product) with pleckstrin homology (PH) domains of cytohesin-1 or cytohesin-like proteins. This is followed by the translocation of these proteins to the plasma membrane and their association with the cytoplasmic domains of the β subunits of integrins, which is realized in various types of cells. This interaction leads to conformational changes in the extracellular domains of the α and β subunits of integrins and an increase in their affinity for specific ECM proteins, as well as clustering (Hughes, Pfaff, 1998; Hynes, 2002). In athero-resistant areas of arteries, the ECM is rich in CL (IV) and LN. However, in atherogenic regions (low/oscillatory WSS conditions), the ECM is enriched in the pro-inflammatory proteins FN and FG (Orr et al., 2005; Feaver et al., 2010; Collins et al., 2014). Antibodies specific to ligated β1 and β3 integrins were used to show that, under conditions of low/oscillatory WSS, there is an increase in the binding of α5β1 integrin to FN (ligand of α5β1, αvβ3, αvβ5 integrins) and of αvβ3 integrin to VN. Thus, the conformational activation of integrins and the dynamic formation of new bonds of integrins with specific ligands occur – ECM proteins are realized (Jalali et al., 2001; Tzima et al., 2001).

The composition of the ECM and the ligation of integrins by the ECM proteins activate many intracellular signaling cascades – in this case, the Shc-Grb·Sos-Ras-MAPK signaling pathway. The cellular adapter protein Shc, activated under conditions of low/oscillatory WSS by tyrosine kinases Src and VEGFR2 close to the EC contacts (phosphorylation of Tyr239/240), forms an early unstable complex Shc·VEGFR2·VE-cadherin (Liu et al., 2008). Next, a stable complex Shc with an αvβ3 integrin that is ligated with FN or VN is formed (Chen et al., 1999; Jalali et al., 2001; Liu et al., 2008). Additionally, Shc associates with the β1 and β5 integrins that are ligated with FN and VN (Chen et al., 1999; Jalali et al., 2001). In this manner, Shc coordinates the intercellular contact proteins (VE-cadherin) and integrin-ECM interactions under low/oscillatory WSS conditions. Membrane-associated small G proteins Ras (small GTPases Ras) function cyclically between active Ras·GTP and inactive Ras·GDP forms, which are a molecular switch of an intracellular signal in response to an extracellular stimulus (Johnson, Chen, 2012). The complexes of phosphorylated Shc with the αvβ3, β1 and β5 integrins formed on the cytoplasmic side of the plasma membrane are accompanied by the association of Shc with the cytoplasmic Grb2·Sos complex of the growth factor receptor-bound protein 2 (Grb2) and the guanine nucleotide exchange factor Sos. This complex stimulates the rate of exchange of the GDP associated with Ras, on the GTP (the interaction of Shc with Grb2·Sos recruits Sos to the cytoplasmic side of the plasma membrane, providing the activation of membrane-localised Ras·GDP) (McCormick et al., 1993). In turn, activated Ras stimulates cytoplasmic Raf kinase (kinase kinase MAPK, MAPKKK, MEKK) and recruits Raf to the inner surface of the plasma membrane through direct interaction with its regulatory domains and the subsequent phosphorylation of four sites of the kinase domain (Dumaz, Marais, 2005). Raf, activated via a MAP kinase cascade, activates MAPK (JNK) (Davis, 2000).

However, the activation of Shc and membrane-associated heterotrimeric G proteins (that consist of α, β, and ϒ subunits) leads to the stimulation of Ras under the conditions of the action of WSS on the cells. In the absence of WSS, G proteins are activated via association with ligand-activated receptors: the α subunit exchanges the GDP bound to it for GTP and dissociates from the βϒ dimer. The α·GTP and βϒ complexes become mediators of cellular event signaling until the α subunit restores its inactive GDP-bound state (since it has GTPase activity.) The reassociation of α·GDP with the βϒ dimer provides an inactive $G_{\alpha\beta\Upsilon}$ heterotrimer that is capable of entering a new activation cycle (Simon et al., 1991). Alterations in the physical properties of membranes under WSS conditions (including the ordered configuration of the phospholipid bilayer [lipid order], fluidity and cholesterol content) affect the conformation and functions of membrane-associated proteins and, as a result, the signaling pathways activated by these proteins (Yamamoto, Ando, 2018).

Purified heterotrimeric G proteins in phospholipid liposomes loaded with [Υ-^{32}P]GTP were shown to be activated by the action of physiological levels of WSS on liposomes (Gudi et al., 1998). Within 1 second, a physiologically low WSS activated the $G\alpha_{q/11}$ and $G\alpha_{i3/0}$ subunits of G proteins in ECs (Gudi et al., 1996). The activated $G\alpha_q$ and $G_{\beta\Upsilon}$ subunits dissociated from them under conditions of a WSS gradient in a physiologically low range and initiated Ras activation and downstream MAPK signaling (Gudi et al., 2003). Moreover, two primary basic mechanosensors, $G\alpha_{q/11}$ and PECAM-1, establish a mechanosensitive $G\alpha_{q/11}$·PECAM-1 complex, which is formed under conditions of laminar flow (physiological high WSS) *in vivo* (atheroprotective straight

region of the descending aorta of mice) and *in vitro*. Under oscillatory flow conditions (low/oscillatory WSS), leading to the activation of G proteins, this complex is rapidly (within 30 seconds) destroyed (Otte et al., 2009). The formation of the $G\alpha_{q/11}$·PECAM-1 complex involves PECAM-1 extracellular Ig-like domains 2 and 3, as well as a $G\alpha_{q/11}$ interacting receptor that associates with $G\alpha_{q/11}$ and PECAM-1 Ig-like domains 2 and 3 and serves as a bridge in the formation of the complex (Yeh et al., 2008). The PECAM-1·$G\alpha_{q/11}$ complex also includes heparan sulfate proteoglycan, which associates with the Ig-like domain 3 of PECAM-1 and mediates the formation of the PECAM-1·$G\alpha_{q/11}$ complex (Paz et al., 2014).

Conclusion

Computer modelling of the blood flow in the arteries makes it possible to determine the most atherogenic areas of the arteries, which are characterised by low and low/oscillatory WSS. The transcription factor NF-κB and the cell adhesion molecules ICAM-1, VCAM-1 and E-selectin are the earliest markers of atherogenesis. Therefore, the processes involved in the expression of these proteins under the conditions induced by a mechanical stimulus (low and low/oscillatory WSS) on endothelial cells are of great interest. This review presents an analysis of numerous studies that demonstrated how the activation of membrane-bound proteins that perceive a mechanical stimulus (low and low/oscillatory WSS) triggers a cascade of biochemical reactions that lead to the transcriptional expression of NF-κB, a key regulator of the expression of cell adhesion molecules. This review also describes in detail the mechanisms of interaction between the endothelial cell adhesion molecules and blood leukocytes that are responsible for adhesion and the subsequent TEM of leukocytes during the initial stage of atherogenesis. Studying the molecular processes involved in the initiation and development of atherosclerosis is extremely important for the development of an effective defence against this disease.

References

Alcaide P., Newton G., Auerbach S., Sehrawat S., Mayadas T.N., Golan D.E., Yacono P., Vincent P., Kowalczyk A., Luscinskas F.W. p120-Catenin regulates leukocyte transmigration through an effect on VE-cadherin phosphorylation. *Blood.* 2008;112(7):2770-2779. DOI 10.1182/blood-2008-03-147181.

Allingham M.J., van Buul J.D., Burridge K. ICAM-1-mediated, Src- and Pyk2-dependent vascular endothelial cadherin tyrosine phosphorylation is required for leukocyte transendothelial migration. *J. Immunol.* 2007;179(6):4053-4064. DOI 10.4049/jimmunol.179.6.4053.

Allport J.R., Muller W.A., Luscinskas F.W. Monocytes induce reversible focal changes in vascular endothelial cadherin complex during transendothelial migration under flow. *J. Cell. Biol.* 2000;148(1):203-216. DOI 10.1083/jcb.148.1.203.

Amberger A., Maczek C., Jürgens G., Michaelis D., Schett G., Trieb K., Eberl T., Jindal S., Xu Q., Wick G. Co-expression of ICAM-1, VCAM-1, ELAM-1 and Hsp60 in human arterial and venous endothelial cells in response to cytokines and oxidized low-density lipoproteins. *Cell Stress Chaperones.* 1997;2(2):94-103. DOI 10.1379/1466-1268(1997)002<0094:ceoive>2.3.co;2.

Arita-Okubo S., Kim-Kaneyama J.R., Lei X.F., Fu W.G., Ohnishi K., Takeya M., Miyauchi A., Honda H., Itabe H., Miyazaki T., Miyazaki A. Role of Hic-5 in the formation of microvilli-like structures and the monocyte-endothelial interaction that accelerates athero-

sclerosis. *Cardiovasc. Res.* 2015;105(3):361-371. DOI 10.1093/cvr/cvv003.

Barreiro O., Yanez-Mo M., Serrador J.M., Montoya M.C., Vicente-Manzanares M., Tejedor R., Furthmayr H., Sanchez-Madrid F. Dynamic interaction of VCAM-1 and ICAM-1 with moesin and ezrin in a novel endothelial docking structure for adherent leukocytes. *J. Cell Biol.* 2002;157(7):1233-1245. DOI 10.1083/jcb.200112126.

Brand K., Page S., Rogler G., Bartsch A., Brandl R., Knuechel R., Page M., Kaltschmidt C., Baeuerle P.A., Neumeier D. Activated transcription factor nuclear factor-kappa B is present in the atherosclerotic lesion. *J. Clin. Invest.* 1996;97(7):1715-1722. DOI 10.1172/JCI118598.

Butcher M.J., Waseem T.C., Galkina E.V. Smooth muscle cell-derived interleukin-17C plays an atherogenic role via the recruitment of proinflammatory interleukin-17A+ T cells to the aorta. *Arterioscler. Thromb. Vasc. Biol.* 2016;36(8):1496-1506. DOI 10.1161/ATVBAHA.116.307892.

Carlos T.M., Harlan J.M. Leukocyte-endothelial adhesion molecules. *Blood.* 1994;84(7):2068-2101. DOI 10.1111/j.1582-4934.2009.00811.x.

Carman C.V., Jun C.D., Salas A., Springer T.A. Endothelial cells proactively form microvilli-like membrane projections upon intercellular adhesion molecule 1 engagement of leukocyte LFA-1. *J. Immunol.* 2003;171(11):6135-6144. DOI 10.4049/jimmunol.171.11.6135.

Carman C.V., Springer T.A. A transmigratory cup in leukocyte diapedesis both through individual vascular endothelial cells and between them. *J. Cell Biol.* 2004;167(2):377-388. DOI 10.1083/jcb.200404129.

Cecchi E., Giglioli C., Valente S., Lazzeri C., Gensini G.F., Abbate R., Mannini L. Role of hemodynamic shear stress in cardiovascular disease. *Atherosclerosis.* 2011;214(2):249-256. DOI 10.1016/j.atherosclerosis.2010.09.008.

Chang C.T., Shen M.Y., Lee A.S., Wang C.C., Chen W.Y., Chang C.M., Chang K.C., Stancel N., Chen C.H. Electronegative low-density lipoprotein increases the risk of ischemic lower-extremity peripheral artery disease in uremia patients on maintenance hemodialysis. *Sci. Rep.* 2017;7(1):4654-4662. DOI 10.1038/s41598-017-04063-3.

Charo I.F., Taubman M.B. Chemokines in the pathogenesis of vascular disease. *Circ. Res.* 2004;95(9):858-866. DOI 10.1161/01.RES.0000146672.10582.17.

Chen K.D., Li Y.S., Kim M., Li S., Yuan S., Chien S., Shyy J.Y. Mechanotransduction in response to shear stress. Roles of receptor tyrosine kinases, integrins, and Shc. *J. Biol. Chem.* 1999;274(26):18393-18400. DOI 10.1074/jbc.274.26.18393.

Cheng C., Tempel D., van Haperen R., van der Baan A., Grosveld F., Daemen M.J., Krams R., de Crom R. Atherosclerotic lesion size and vulnerability are determined by patterns of fluid shear stress. *Circulation.* 2006;113(23):2744-2753. DOI 10.1161/CIRCULATIONAHA.105.590018.

Chiu Y.J., McBeath E., Fujiwara K. Mechanotransduction in an extracted cell model: Fyn drives stretch- and flow-elicited PECAM-1 phosphorylation. *J. Cell Biol.* 2008;182(4):753-763. DOI 10.1083/jcb.200801062.

Colic M., Pantovic S., Jeremic M., Jokovic V., Obradovic Z., Rosic M. Transport of low-density lipoprotein into the blood vessel wall during atherogenic diet in the isolated rabbit carotid artery. *Circ. J.* 2015;79(8):1846-1852. DOI 10.1253/circj.CJ-14-1316.

Collins C., Osborne L.D., Guilluy C., Chen Z., O'Brien E.T. 3rd, Reader J.S., Burridge K., Superfine R., Tzima E. Haemodynamic and extracellular matrix cues regulate the mechanical phenotype and stiffness of aortic endothelial cells. *Nat. Commun.* 2014;5:3984. DOI 10.1038/ncomms4984.

Cuhlmann S., Van der Heiden K., Saliba D., Tremoleda J.L., Khalil M., Zakkar M., Chaudhury H., Luong le A., Mason J.C., Udalova I., Gsell W., Jones H., Haskard D.O., Krams R., Evans P.C. Disturbed blood flow induces RelA expression via c-Jun N-terminal kinase 1: a novel mode of NF-κB regulation that promotes arterial inflamma-

Е.Л. Мищенко, А.М. Мищенко
В.А. Иванисенко

Механочувствительные молекулярные взаимодействия
в атерогенных районах артерий: развитие атеросклероза

2021
25·5

tion. *Circ. Res.* 2011;108(8):950-959. DOI 10.1161/CIRCRESAHA. 110.233841.

Dai G., Kaazempur-Mofrad M.R., Natarajan S., Zhang Y., Vaughn S., Blackman B.R., Kamm R.D., García-Cardeña G., Gimbrone M.A. Jr. Distinct endothelial phenotypes evoked by arterial waveforms derived from atherosclerosis-susceptible and -resistant regions of human vasculature. *Proc. Natl. Acad. Sci. USA.* 2004;101(41):14871-14876. DOI 10.1073/pnas.0406073101.

Davis M.E., Grumbach I.M., Fukai T., Cutchins A., Harrison D.G. Shear stress regulates endothelial nitric-oxide synthase promoter activity through nuclear factor kappaB binding. *J. Biol. Chem.* 2003; 279(1):163-168. DOI 10.1074/jbc.M307528200.

Davis R.J. Signal transduction by the JNK group of MAP kinases. *Cell.* 2000;103:239-252. DOI 10.1016/S0092-8674(00)00116-1.

Denis J.F., Scheckenbach K.E.L., Pfenniger A., Meens M.J., Krams R., Miquerol L., Taffet S., Chanson M., Delmar M., Kwak B.R. Connexin40 controls endothelial activation by dampening NFκB activation. *Oncotarget.* 2017;8(31):50972-50986. DOI 10.18632/onco target.16438.

Dumaz N., Marais R. Integrating signals between cAMP and the RAS/RAF/MEK/ERK signalling pathways. *FEBS J.* 2005;272(14):3491-3504. DOI 10.1111/j.1742-4658.2005.04763.x.

Erbel C., Akhavanpoor M., Okuyucu D., Wangler S., Dietz A., Zhao L., Stellos K., Little K.M., Lasitschka F., Doesch A., Hakimi M., Dengler T.J., Giese T., Blessing E., Katus H.A., Gleissner C.A. IL-17A influences essential functions of the monocyte/macrophage lineage and is involved in advanced murine and human atherosclerosis. *J. Immunol.* 2014;193(9):4344-4355. DOI 10.4049/jimmunol.1400 181.

Feaver R.E., Gelfand B.D., Wang C., Schwartz M.A., Blackman B.R. Atheroprone hemodynamics regulate fibronectin deposition to create positive feedback that sustains endothelial inflammation. *Circ. Res.* 2010;106(11):1703-1711. DOI 10.1161/CIRCRESAHA. 109.216283.

Frangos J.A., Eskin S.G., McIntire L.V., Ives C.L. Flow effects on prostacyclin production by cultured human endothelial cells. *Science.* 1985;227(4693):1477-1479. DOI 10.1126/science.3883488.

Garrett J.P., Lowery A.M., Adam A.P., Kowalczyk A.P., Vincent P.A. Regulation of endothelial barrier function by p120-catenin·VE-cadherin interaction. *Mol. Biol. Cell.* 2017;28(1):85-97. DOI 10.1091/mbc.E16-08-0616.

Gonzalez A.M., Cyrus B.F., Muller W.A. Targeted recycling of the lateral border recycling compartment precedes adherens junction dissociation during transendothelial migration. *Am. J. Pathol.* 2016; 186(5):1387-1402. DOI 10.1016/j.ajpath.2016.01.010.

Gudi S.R., Clark C.B., Frangos J.A. Fluid flow rapidly activates G proteins in human endothelial cells. Involvement of G proteins in mechanochemical signal transduction. *Circ. Res.* 1996;79(4):834-839. DOI 10.1161/01.RES.79.4.834.

Gudi S., Huvar I., White C.R., McKnight N.L., Dusserre N., Boss G.R., Frangos J.A. Rapid activation of Ras by fluid flow is mediated by Galpha(q) and Gbetagamma subunits of heterotrimeric G proteins in human endothelial cells. *Arterioscler. Thromb. Vasc. Biol.* 2003; 23(6):994-1000. DOI 10.1161/01.ATV.0000073314.51987.84.

Gudi S., Nolan J.P., Frangos J.A. Modulation of GTPase activity of G proteins by fluid shear stress and phospholipid composition. *Proc. Natl. Acad. Sci. USA.* 1998;95(5):2515-2519. DOI 10.1073/PNAS. 95.5.2515.

Hajra L., Evans A.I., Chen M., Hyduk S.J., Collins T., Cybulsky M.I. The NF-kappa B signal transduction pathway in aortic endothelial cells is primed for activation in regions predisposed to atherosclerotic lesion formation. *Proc. Natl. Acad. Sci. USA.* 2000;97(16):9052-9057. DOI 10.1073/pnas.97.16.9052.

Harrison M., Smith E., Ross E., Krams R., Segers D., Buckley C.D., Nash G.B., Rainger G.E. The role of platelet-endothelial cell adhesion molecule-1 in atheroma formation varies depending on the site-specific hemodynamic environment. *Arterioscler. Thromb. Vasc. Biol.* 2013;33(4):694-701. DOI 10.1161/ATVBAHA.112.300379.

Hughes P.E., Pfaff M. Integrin affinity modulation. *Trends. Cell Biol.* 1998;8(9):359-364. DOI 10.1016/s0962-8924(98)01339-7.

Hung O.Y., Molony D., Corban M.T., Rasoul-Arzrumly E., Maynard C., Eshtehardi P., Dhawan S., Timmins L.H., Piccinelli M., Ahn S.G., Gogas B.D., McDaniel M.C., Quyyumi A.A., Giddens D.P., Samady H. Comprehensive assessment of coronary plaque progression with advanced intravascular imaging, physiological measures, and wall shear stress: a pilot double-blinded randomized controlled clinical trial of nebivolol versus atenolol in nonobstructive coronary artery disease. *J. Am. Heart. Assoc.* 2016;5(1):e002764. DOI 10.1161/JAHA.115.002764.

Huo Y., Ley K. Adhesion molecules and atherogenesis. *Acta Physiol. Scand.* 2001;173(1):35-43. DOI 10.1046/j.1365-201X.2001.00882.x.

Hynes R.O. Integrins: bidirectional, allosteric signaling machines. *Cell.* 2002;110(6):673-87. DOI 10.1016/s0092-8674(02)00971-6.

Jalali S., del Pozo M.A., Chen K., Miao H., Li Y., Schwartz M.A., Shyy J.Y., Chien S. Integrin-mediated mechanotransduction requires its dynamic interaction with specific extracellular matrix (ECM) ligands. *Proc. Natl. Acad. Sci. USA.* 2001;98(3):1042-1046. DOI 10.1073/pnas.98.3.1042.

Johnson D.S., Chen Y.H. Ras family of small GTPases in immunity and inflammation. *Curr. Opin. Pharmacol.* 2012;12(4):458-463. DOI 10.1016/j.coph.2012.02.003.

Ku D.N., Giddens D.P., Zarins C.K., Glagov S. Pulsatile flow and atherosclerosis in the human carotid bifurcation. Positive correlation between plaque location and low oscillating shear stress. *Arteriosclerosis.* 1985;5(3):293-302. DOI 10.1161/01.ATV.5.3.293.

Kume N., Cybulsky M.I., Gimbrone M.A.Jr. Lysophosphatidylcholine, a component of atherogenic lipoproteins, induces mononuclear leukocyte adhesion molecules in cultured human and rabbit arterial endothelial cells. *J. Clin. Invest.* 1992;90(3):1138-1144. DOI 10.1172/JCI115932.

Ledebur H.C., Parks T.P. Transcriptional regulation of the intercellular adhesion molecule-1 gene by inflammatory cytokines in human endothelial cells. Essential roles of a variant NF-kappa B site and p65 homodimers. *J. Biol. Chem.* 1995;270(2):933-943. DOI 10.1074/jbc.270.2.933.

Libby P., Buring J.E., Badimon L., Hansson G.K., Deanfield J., Bittencourt M.S., Tokgözoğlu L., Lewis E.F. Atherosclerosis. *Nat. Rev. Dis. Primers.* 2019;5(1):56. DOI 10.1038/s41572-019-0106-z.

Liu Y., Sweet D.T., Irani-Tehrani M., Maeda N., Tzima E. Shc coordinates signals from intercellular junctions and integrins to regulate flow-induced inflammation. *J. Cell Biol.* 2008;182(1):185-196. DOI 10.1083/jcb.200709176.

Luscinskas F.W., Ding H., Tan P., Cumming D., Tedder T.F., Gerritsen M.E. L- and P-selectins, but not CD49d (VLA-4) integrins, mediate monocyte initial attachment to TNF-alpha-activated vascular endothelium under flow *in vitro*. *J. Immunol.* 1996;157(1):326-335.

Luscinskas F.W., Kansas G.S., Ding H., Pizcueta P., Schleiffenbaum B.E., Tedder T.F., Gimbrone M.A. Jr. Monocyte rolling, arrest and spreading on IL-4-activated vascular endothelium under flow is mediated via sequential action of L-selectin, beta 1-integrins, and beta 2-integrins. *J. Cell Biol.* 1994;125(6):1417-1427. DOI 10.1083/jcb.125.6.1417.

Mamdouh Z., Chen X., Pierini L.M., Maxfield F.R., Muller W.A. Targeted recycling of PECAM from endothelial surface-connected compartments during diapedesis. *Nature.* 2003;421(6924):748-753. DOI 10.1038/nature01300.

Mamdouh Z., Kreitzer G.E., Muller W.A. Leukocyte transmigration requires kinesin-mediated microtubule-dependent membrane trafficking from the lateral border recycling compartment. *J. Exp. Med.* 2008;205(4):951-966. DOI 10.1084/jem.20072328.

Mamdouh Z., Mikhailov A., Muller W.A. Transcellular migration of leukocytes is mediated by the endothelial lateral border recycling compartment. *J. Exp. Med.* 2009;206(12):2795-2808. DOI 10.1084/jem.20082745.

Martin T., Cardarelli P.M., Parry G.C., Felts K.A., Cobb R.R. Cytokine induction of monocyte chemoattractant protein-1 gene expression in

human endothelial cells depends on the cooperative action of NF-kappa B and AP-1. *Eur. J. Immunol.* 1997;27(5):1091-1097. DOI 10.1002/eji.1830270508.

McCormick F. Signal transduction. How receptors turn Ras on. *Nature.* 1993;363(6424):15-16. DOI 10.1038/363015a0.

Mohan S., Mohan N., Valente A.J., Sprague E.A. Regulation of low shear flow-induced HAEC VCAM-1 expression and monocyte adhesion. *Am. J. Physiol.* 1999;276(5):C1100-1107. DOI 10.1152/ajpcell.1999.276.5.C1100.

Morbiducci U., Kok A.M., Kwak B.R., Stone P.H., Steinman D.A., Wentzel J.J. Atherosclerosis at arterial bifurcations: evidence for the role of haemodynamics and geometry. *Thromb. Haemost.* 2016; 115(3):484-492. DOI 10.1160/TH15-07-0597.

Nagel T., Resnick N., Dewey C.F. Jr., Gimbrone M.A. Jr. Vascular endothelial cells respond to spatial gradients in fluid shear stress by enhanced activation of transcription factors. *Arterioscler. Thromb. Vasc. Biol.* 1999;19(8):1825-1834. DOI 10.1161/01.ATV.19.8.1825.

Nebuloni L., Kuhn G.A., Müller R. A comparative analysis of water-soluble and blood-pool contrast agents for *in vivo* vascular imaging with micro-CT. *Acad. Radiol.* 2013;20(10):1247-1255. DOI 10.1016/j.acra.2013.06.003.

Neish A.S., Williams A.J., Palmer H.J., Whitley M.Z., Collins T. Functional analysis of the human vascular cell adhesion molecule 1 promoter. *J. Exp. Med.* 1992;176(6):1583-1593. DOI 10.1084/jem.176.6.1583.

Oeckinghaus A., Hayden M.S., Ghosh S. Crosstalk in NF-κB signaling pathways. *Nat. Immunol.* 2011;12(8):695-708. DOI 10.1038/ni.2065.

O'Keeffe L.M., Muir G., Piterina A.V., McGloughlin T. Vascular cell adhesion molecule-1 expression in endothelial cells exposed to physiological coronary wall shear stresses. *J. Biomech. Eng.* 2009; 131(8):081003. DOI 10.1115/1.3148191.

Orr A.W., Sanders J.M., Bevard M., Coleman E., Sarembock I.J., Schwartz M.A. The subendothelial extracellular matrix modulates NF-kappaB activation by flow: a potential role in atherosclerosis. *J. Cell Biol.* 2005;169(1):191-202. DOI 10.1083/jcb.200410073.

Otsuka F., Kramer M.C., Woudstra P., Yahagi K., Ladich E., Finn A.V., de Winter R.J., Kolodgie F.D., Wight T.N., Davis H.R., Joner M., Virmani R. Natural progression of atherosclerosis from pathologic intimal thickening to late fibroatheroma in human coronary arteries: a pathology study. *Atherosclerosis.* 2015;241(2):772-782. DOI 10.1016/j.atherosclerosis.2015.05.011.

Otte L.A., Bell K.S., Loufrani L., Yeh J.-C., Melchior B., Dao D.N., Stevens H.Y., White C.R., Frangos J.A. Rapid changes in shear stress induce dissociation of a G alpha(q/11)-platelet endothelial cell adhesion molecule-1 complex. *J. Physiol.* 2009;587(Pt.10):2365-2373. DOI 10.1113/jphysiol.2009.172643.

Ozerova I.N., Metelskaya V.A., Gavrilova N.E. Atherogenic normolipidemia in men with coronary atherosclerosis: some peculiarities of subfractional distribution of apo B-containing lipoproteins. *Ateroscleroz = Atherosclerosis.* 2018;14(3):5-11. DOI 10.15372/ATER 20180301. (in Russian)

Paddock C., Zhou D., Lertkiatmongkol P., Newman P.J., Zhu J. Structural basis for PECAM-1 homophilic binding. *Blood.* 2016;127(8): 1052-1061. DOI 10.1182/blood-2015-07-660092.

Paz dela N.G., Melchior B., Shayo F.Y., Frangos J.A. Heparan sulfates mediate the interaction between platelet endothelial cell adhesion molecule-1 (PECAM-1) and the Gαq/11 subunits of heterotrimeric G proteins. *J. Biol. Chem.* 2014;289(11):7413-7424. DOI 10.1074/jbc.M113.542514.

Peters W., Charo I.F. Involvement of chemokine receptor 2 and its ligand, monocyte chemoattractant protein-1, in the development of atherosclerosis: lessons from knockout mice. *Curr. Opin. Lipidol.* 2001;12(2):175-180. DOI 10.1097/00041433-200104000-00011.

Resnick N., Collins T., Atkinson W., Bonthron D.T., Dewey C.F. Jr., Gimbrone M.A. Jr. Platelet-derived growth factor B chain promoter contains a cis-acting fluid shear-stress-responsive element.

Proc. Natl. Acad. Sci. USA. 1993;90(10):4591-4595. DOI 10.1073/pnas.90.10.4591.

Sahebkar A., Morris D.R., Biros E., Golledge J. Association of single nucleotide polymorphisms in the gene encoding platelet endothelial cell adhesion molecule-1 with the risk of myocardial infarction: a systematic review and meta-analysis. *Thromb. Res.* 2013;132(2):227-233. DOI 10.1016/j.thromres.2013.07.007.

Sakellarios A.I, Papafaklis M.I., Siogkas P., Athanasiou L.S., Exarchos T.P., Stefanou K., Bourantas C.V., Naka K.K., Michalis L.K., Parodi O., Fotiadis D.I. Patient-specific computational modeling of subendothelial LDL accumulation in a stenosed right coronary artery: effect of hemodynamic and biological factors. *Am. J. Physiol. Heart Circ. Physiol.* 2013;304(11):H1455-70. DOI 10.1152/ajpheart.00539.2012.

Schober A., Siess W. Lysophosphatidic acid in atherosclerotic diseases. *Br. J. Pharmacol.* 2012;167(3):465-482. DOI 10.1111/j.1476-5381.2012.02021.x.

Shaw S.K., Bamba P.S., Perkins B.N., Luscinskas F.W. Real-time imaging of vascular endothelial-cadherin during leukocyte transmigration across endothelium. *J. Immunol.* 2001;167(4):2323-2330. DOI 10.4049/jimmunol.167.4.2323.

Shimada H., Rajagopalan L.E. Rho kinase-2 activation in human endothelial cells drives lysophosphatidic acid-mediated expression of cell adhesion molecules via NF-kappaB p65. *J. Biol. Chem.* 2010; 285(17):12536-12542. DOI 10.1074/jbc.M109.099630.

Sigal A., Bleijs D.A., Grabovsky V., van Vliet S.J., Dwir O., Figdor C.G., van Kooyk Y., Alon R. The LFA-1 integrin supports rolling adhesions on ICAM-1 under physiological shear flow in a permissive cellular environment. *J. Immunol.* 2000;165(1):442. DOI 10.4049/jimmunol.165.1.442.

Simon M.I., Strathmann M.P., Gautam N. Diversity of G proteins in signal transduction. *Science.* 1991;252(5007):802-808. DOI 10.1126/science.1902986.

Snyder J.L., McBeath E., Thomas T.N., Chiu Y.J., Clark R.L., Fujiwara K. Mechanotransduction properties of the cytoplasmic tail of PECAM-1. *Biol. Cell.* 2017;109(8):312-321. DOI 10.1111/boc.201600079.

Soulis J.V., Farmakis T.M., Giannoglou G.D., Louridas G.E. Wall shear stress in normal left coronary artery tree. *J. Biomech.* 2006;39(4): 742-749. DOI 10.1016/j.jbiomech.2004.12.026.

Spiecker M., Peng H.B., Liao J.K. Inhibition of endothelial vascular cell adhesion molecule-1 expression by nitric oxide involves the induction and nuclear translocation of IkappaBalpha. *J. Biol. Chem.* 1997;272(49):30969-30974. DOI 10.1074/jbc.272.49.30969.

Suo J., Ferrara D.E., Sorescu D., Guldberg R.E., Taylor W.R., Giddens D.P. Hemodynamic shear stresses in mouse aortas – implications for atherogenesis. *Arterioscler. Thromb. Vasc. Biol.* 2007;27: 346-351. DOI 10.1161/01.ATV.0000253492.45717.46.

Timmins L.H., Molony D.S., Eshtehardi P., McDaniel M.C., Oshinski J.N., Giddens D.P., Samady H. Oscillatory wall shear stress is a dominant flow characteristic affecting lesion progression patterns and plaque vulnerability in patients with coronary artery disease. *J. R. Soc. Interface.* 2017;14(127):20160972. DOI 10.1098/rsif.2016.0972.

Timmins L.H., Molony D.S., Eshtehardi P., McDaniel M.C., Oshinski J.N., Samady H., Giddens D.P. Focal association between wall shear stress and clinical coronary artery disease progression. *Ann. Biomed. Eng.* 2015;43(1):94-106. DOI 10.1007/s10439-014-1155-9.

Tzima E., del Pozo M.A., Shattil S.J., Chien S., Schwartz M.A. Activation of integrins in endothelial cells by fluid shear stress mediates Rho-dependent cytoskeletal alignment. *EMBO J.* 2001;20(17): 4639-4647. DOI 10.1093/emboj/20.17.4639.

Tzima E., Irani-Tehrani M., Kiosses W.B., Dejana E., Schultz D.A., Engelhardt B., Cao G., DeLisser H., Schwartz M.A. A mechanosensory complex that mediates the endothelial cell response to fluid shear stress. *Nature.* 2005;437(7057):426-431. DOI 10.1038/nature03952.

Е.Л. Мищенко, А.М. Мищенко
В.А. Иванисенко

Механочувствительные молекулярные взаимодействия
в атерогенных районах артерий: развитие атеросклероза

2021
25•5

Virani S.S., Alonso A., Benjamin E.J., Bittencourt M.S., Callaway C.W., Carson A.P., Chamberlain A.M., Chang A.R., Cheng S., Delling F.N., ..., Stokes A., Tirschwell D.L., VanWagner L.B., Tsao C.W., American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics-2020 update: a report from the American Heart Association. *Circulation*. 2020;141(9):e139-e596. DOI 10.1161/CIR.0000000000000757.

Whitley M.Z., Thanos D., Read M.A., Maniatis T., Collins T. A striking similarity in the organization of the E-selectin and beta interferon gene promoters. *Mol. Cell Biol.* 1994;14(10):6464-6475. DOI 10.1128/mcb.14.10.6464.

Wu C., Huang R.T., Kuo C.H., Kumar S., Kim C.W., Lin Y.C., Chen Y.J., Birukova A., Birukov K.G., Dulin N.O., Civelek M., Lusis A.J., Loyer X., Tedgui A., Dai G., Jo H., Fang Y. Mechanosensitive PPAP2B regulates endothelial responses to atherorelevant hemodynamic forces. *Circ. Res.* 2015;117(4):e41-e53. DOI 10.1161/CIRCRESAHA.117.306457.

Xia T., Liu X., Du C.J., Jin X., Kong X.Q., Li G. Association of Leu-125Val polymorphisms in the PECAM-1 gene with the risk of coronary heart disease: a meta-analysis. *Int. J. Clin. Exp. Med.* 2015; 8(2):2219-2225.

Xing R., De Wilde D., McCann G., Ridwan Y., Schrauwen J.T., van der Steen A.F., Gijsen F.J., Van der Heiden K. Contrast-enhanced micro-CT imaging in murine carotid arteries: a new protocol for computing wall shear stress. *Biomed. Eng. Online.* 2016;15(Suppl.2):156. DOI 10.1186/s12938-016-0270-2.

Yamamoto K., Ando J. Emerging role of plasma membranes in vascular endothelial mechanosensing. *Circ. J.* 2018;82(11):2691-2698. DOI 10.1253/circj.CJ-18-0052.

Yeh J.-C., Otte L.A., Frangos J.A. Regulation of G protein-coupled receptor activities by the platelet-endothelial cell adhesion molecule, PECAM-1. *Biochemistry.* 2008;47(34):9029-9039. DOI 10.1021/bi8003846.

Yu X.H., Zheng X.L., Tang C.K. Nuclear factor-κB activation as a pathological mechanism of lipid metabolism and atherosclerosis. *Adv. Clin. Chem.* 2015;70:1-30. DOI 10.1016/bs.acc.2015.03.004.

Yung Y.C., Stoddard N.C., Chun J. LPA receptor signaling: pharmacology, physiology, and pathophysiology. *J. Lipid Res.* 2014;55(7): 1192-1214. DOI 10.1194/jlr.R046458.

Yurdagul A.Jr., Chen J., Funk S.D., Albert P., Kevil C.G., Orr A.W. Altered nitric oxide production mediates matrix-specific PAK2 and NF-κB activation by flow. *Mol. Biol. Cell.* 2013;24(3):398-408. DOI 10.1091/mbc.E12-07-0513.

Yurdagul A.Jr., Sulzmaier F.J., Chen X.L., Pattillo C.B., Schlaepfer D.D., Orr A.W. Oxidized LDL induces FAK-dependent RSK signaling to drive NF-κB activation and VCAM-1 expression. *J. Cell Sci.* 2016;129(8):1580-1591. DOI 10.1242/jcs.182097.

Zakkar M., Chaudhury H., Sandvik G., Enesa K., Luong le A., Cuhlmann S., Mason J.C., Krams R., Clark A.R., Haskard D.O., Evans P.C. Increased endothelial mitogen-activated protein kinase phosphatase-1 expression suppresses proinflammatory activation at sites that are resistant to atherosclerosis. *Circ. Res.* 2008;103(7): 726-732. DOI 10.1161/CIRCRESAHA.108.183913.

Zhang W., Tang T., Nie D., Wen S., Jia C., Zhu Z., Xia N., Nie S., Zhou S., Jiao J., Dong W., Lu B., Xu T., Sun B., Lu Y., Li Y., Cheng L., Liao Y., Cheng X. IL-9 aggravates the development of atherosclerosis in ApoE-/- mice. *Cardiovasc. Res.* 2015;106(3):453-464. DOI 10.1093/cvr/cvv110.

Zou Y., Huang X., Feng L., Hou J., Xing L., Yu B. Localization of in-stent neoatherosclerosis in relation to curvatures and bifurcations after stenting. *J. Thorac. Dis.* 2016;8(12):3530-3536. DOI 10.21037/jtd.2016.11.108.

Phage display as a tool for identifying HIV-1 broadly neutralizing antibodies

A.N. Chikaev[1] ✉, A.P. Rudometov[2], Yu.A. Merkulyeva[2], L.I. Karpenko[2]

[1] Institute of Molecular and Cellular Biology of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] State Research Center of Virology and Biotechnology "Vector", Rospotrebnadzor, Koltsovo, Novosibirsk region, Russia
✉ chikaev@mcb.nsc.ru

Abstract. Combinatorial biology methods offer a good solution for targeting interactions of specific molecules by a high-throughput screening and are widely used for drug development, diagnostics, identification of novel monoclonal antibodies, search for linear peptide mimetics of discontinuous epitopes for the development of immunogens or vaccine components. Among all currently available techniques, phage display remains one of the most popular approaches. Despite being a fairly old method, phage display is still widely used for studying protein-protein, peptide-protein and DNA-protein interactions due to its relative simplicity and versatility. Phage display allows highly representative libraries of peptides, proteins or their fragments to be created. Each phage particle in a library displays peptides or proteins fused to its coat protein and simultaneously carries the DNA sequence encoding the displayed peptide/protein in its genome. The biopanning procedure allows isolation of specific clones for almost any target, and due to the physical link between the genotype and the phenotype of recombinant phage particles it is possible to determine the structure of selected molecules. Phage display technology continues to play an important role in HIV research. A major obstacle to the development of an effective HIV vaccine is an extensive genetic and antigenic variability of the virus. According to recent data, in order to provide protection against HIV infection, the so-called broadly neutralizing antibodies that are cross-reactive against multiple viral strains of HIV must be induced, which makes the identification of such antibodies a key area of HIV vaccinology. In this review, we discuss the use of phage display as a tool for identification of HIV-specific antibodies with broad neutralizing activity. We provide an outline of phage display technology, briefly describe the design of antibody phage libraries and the affinity selection procedure, and discuss the biology of HIV-1-specific broadly neutralizing antibodies. Finally, we summarize the studies aimed at identification of broadly neutralizing antibodies using various types of phage libraries.
Key words: phage display; antibody libraries; HIV-1; broadly neutralizing antibodies (bnAbs).

Применение фагового дисплея для поиска ВИЧ-1-нейтрализующих антител

А.Н. Чикаев[1] ✉, А.П. Рудометов[2], Ю.А. Меркульева[2], Л.И. Карпенко[2]

[1] Институт молекулярной и клеточной биологии Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Государственный научный центр вирусологии и биотехнологии «Вектор» Роспотребнадзора, р. п. Кольцово, Новосибирская область, Россия
✉ chikaev@mcb.nsc.ru

Аннотация. Комбинаторная белковая инженерия – востребованный инструмент для решения задач, связанных со скринингом большого разнообразия взаимодействующих молекул: разработки лекарств, средств диагностики, идентификации антител, поиска конформационных имитаторов антигенных детерминант для создания иммуногенов или компонентов вакцин. Среди всех подобных методик одна из наиболее популярных – технология фагового дисплея, появившаяся во второй половине 1980-х гг., однако в силу относительной простоты и универсальности по-прежнему активно применяющаяся для изучения белок-белковых, пептид-белковых и ДНК-белковых взаимодействий. Фаговый дисплей позволяет создавать высокопредставительные библиотеки пептидов, белков или их фрагментов, в которых каждая фаговая частица экспонирует на своей поверхности исследуемые пептиды или белки и одновременно несет в своем геноме последовательность ДНК, кодирующую экспонируемый пептид/белок. Процедура аффинной селекции позволяет находить специфические фаговые клоны практически к любой мишени, а за счет наличия физической связи между генотипом и фенотипом можно эффективно определить структуру отобранных молекул. Значительную роль технология фагового дисплея сыграла в исследованиях, направленных

А.Н. Чикаев, А.П. Рудометов
Ю.А. Меркульева, Л.И. Карпенко

Применение фагового дисплея
для поиска ВИЧ-1-нейтрализующих антител

2021
25·5

на изучение антигенной структуры вируса иммунодефицита человека (ВИЧ-1) и разработку средств борьбы с этим заболеванием. Серьезная проблема, из-за которой до сих пор не удается создать эффективную анти-ВИЧ-вакцину, – сильная антигенная изменчивость вируса. Согласно современным представлениям, для обеспечения защиты от инфицирования необходимо стимулировать индукцию в организме вирус-нейтрализующих антител, активных в отношении большого числа различных штаммов ВИЧ-1. Соответственно, идентификация подобных антител является важной исследовательской задачей. Тема настоящего обзора – применение фагового дисплея в качестве инструмента для поиска ВИЧ-1-нейтрализующих антител широкого спектра действия. Представлены ключевые характеристики технологии фагового дисплея, кратко описан процесс получения библиотек антител, проведения процедуры аффинной селекции, а также обсуждается феномен ВИЧ-1-нейтрализующих антител широкого спектра действия. Приводится обзор исследований, посвященных поиску кросс-нейтрализующих антител с использованием различных типов фаговых библиотек.

Ключевые слова: фаговый дисплей; библиотеки антител; ВИЧ-1; нейтрализующие антитела широкого спектра действия (bnAbs).

Introduction

Phage display was first described in 1985 by George Smith and Gregory Winter, who were awarded the 2018 Nobel Prize in Chemistry for this discovery. They reported that foreign peptides could be successfully expressed on the surface of bacteriophage particles by integrating a gene of interest into a phage genome upstream of its coat protein open reading frame (Smith, 1985). It is noteworthy that a conceptually similar study was independently conducted by a Russian scientific group led by A.A. Ilyichev, who incorporated a peptide-coding sequence into the pVIII protein gene of M13 phage (Ilyichev et al., 1992; Minenkova et al., 1993). Later, G. Smith and colleagues proposed a selection strategy for the enrichment of population of recombinant phage clones that specifically bind to the target ligand, using affinity enrichment process (Smith, 1985). Since there is a direct physical link between the genotype of the recombinant phage particle and the phenotype of the fusion protein, this method allows the identification of DNA sequences encoding selected molecules.

Subsequently, G. Smith and colleagues described the creation of combinatorial phage libraries that contain a large number of phage particles, each carrying a unique protein or peptide on its surface. Currently, one the most common types of phage libraries used for studying various protein-to-protein, receptor-ligand interactions or protein engineering are antibody phage libraries displaying single-stranded (scFv) and antigen-binding (Fab) fragments of IgG molecules (McCafferty et al., 1990; Winter et al., 1994). There are also alternative antibody formats used for the construction of antibody phage libraries, such as variable domains of antibodies from the heavy chains of camelids (VHH, or nanobodies) and sharks (vNAR) (Davies, Riechmann, 1995; Greenberg et al., 1995).

In order to create phage antibody library, antibody fragments to be exposed are usually fused to the N-terminus of the pIII phage coat protein. Despite the fact that all the phage coat proteins can be used for phage display, only pIII is suited to expose large peptides or proteins without loss of infectivity and functional activity of phage particles (Kay et al., 1993; Kishchenko et al., 1994; Mullen et al., 2006; Tikunova, Morozova, 2009). In early phage display systems,

gene sequences encoding for antibody fragments were inserted directly into phage genome (McCafferty et al., 1990; Scott, Smith, 1990). Currently, a separate plasmid vector, also known as phagemid, is commonly used to introduce target DNA inserts into the phage genome. Phagemid carries recombinant pIII fusion gene, as well as phage and bacterial replication origins (thus can be replicated independently of phage production), but it lacks phage genes necessary for infecting, replicating, assembling and budding phage particles. In order to produce recombinant phages, phagemid-transformed *Escherichia coli* cells should be coinfected with a helper phage that carries wild-type phage genome including all the remaining phage genes required for the phage life cycle (Ledsgaard et al., 2018). The phage origin of replication in the phagemid enables its packaging into the forming virions as a single-strand DNA. Thus, the resulting phage particles contain both recombinant and wild-type forms of pIII from the helper-phage, so the infectivity is not compromised (Felici et al., 1991).

Immune libraries are usually generated from B-cell derived antibody repertoire of immunized animals or reconvalescent donors. Phage immune libraries contain about 10^7–10^8 unique phage clones displaying antigen-specific antibodies on their surface (Kennedy et al., 2018). In some cases, "naive" libraries, based on lymphocyte mRNA of unvaccinated/healthy donors, or intact animals, as well as "synthetic" libraries, based on *de novo* synthesized oligonucleotides, may be used in order to enhance diversity of the antibody repertoire (Griffiths, Duncan, 1998; Tikunova, Morozova, 2009). The representativeness of the phage libraries can reach up to 10^9–10^{10} for "naive" and 10^{10}–10^{11} for "synthetic" libraries (Zhao et al., 2016; Kennedy et al., 2018; Muyldermans, 2021).

Construction of antibody phage display libraries

The library's construction begins with RNA isolation from hybridoma cell lines, spleen cells from immunized animals, or B-lymphocytes from human peripheral blood, and subsequent cDNA synthesis (Clackson et al., 1991). Then, using isotype-specific primers, the variable regions of immunoglobulin light (VL) and heavy (HV) chain genes (or solely VH in the case of VHH) are amplified and cloned

into a phagemid vector between the pIII-encoding gene and N-terminal signal sequence, which directs fusion protein to periplasmic translocation. These phagemids encoding diverse VH/VL gene combinations are then used for the transformation of *E. coli* cells that are co-infected with a helper phage, leading to the production of a set of phage particles exposing different antibody fragments (Skerra, Pluckthun, 1988; Tikunova, Morozova, 2009; Hammers, Stanley, 2014).

Following that, a biopanning procedure (a process of selecting phage clones that carry antigen-specific variants of antibody fragments) is carried out. To do this, the library is incubated with the target antigen that is immobilized on an immune plate, magnetic beads, or immunosorbent. Unbound particles are then washed, and a fraction of target-specific phages can be eluted using buffer with low or high pH or by adding a competing protein or peptide that strongly interferes with binding of the target molecule to selected phages (Smith, Petrenko, 1997). Another frequently used method of biopanning, probably the most effective, involves affinity selection using target molecules labeled with SS-biotin (Chames, Baty, 2010). The target molecule must be immobilized via a streptavidin-coated template. After that, the procedure involves standard incubation of the phage library and washing the unbound phage clones. SS-biotin contains a disulfide bond that can be cleaved by treatment with sulfhydryl, which enables the separation of a complex target (specific phage) from the substrate by adding a reducing agent such as dithiothreitol or 2-mercaptoethanol. This method provides a significant increase in the percentage of target-specific clones during biopanning, since the eluate obtained this way doesn't contain any phages that are non-specifically bound to the substrate.

After each round of biopanning, eluted phage clones are used to infect *E. coli*, which are then cultured and super-infected with helper phage. The produced progeny phage particles are used for subsequent rounds of affinity selection. Commonly, one or two rounds of panning are enough to enrich the library with antigen-specific phage clones, though in the case of synthetic libraries (which are more representative but less specific) the number of rounds might be increased up to five.

Phage titering is done after every affinity screening to assess the amount of target-specific clones. The specificity of each phage clone can be assessed using enzyme-linked immunosorbent assay (ELISA), immunoblotting or flow cytometry. After the final round of biopanning, phages with the highest affinity are picked up, amplified, sequenced and used for phage DNA extraction and amplification of VHH, Fab- or scFv-coding sequences, which are then subcloned into an expression vector in order to express the soluble forms of corresponding proteins. Such Fab/scFvs can also be converted into full-length monoclonal antibodies (mAbs) by in-frame cloning of VH and VL genes into the cassette vector that harbors appropriate heavy IgG constant region genes.

In the final step, identified antibodies should be validated for their avidity and affinity against the target antigen via ELISA, western blotting or another immunological assay
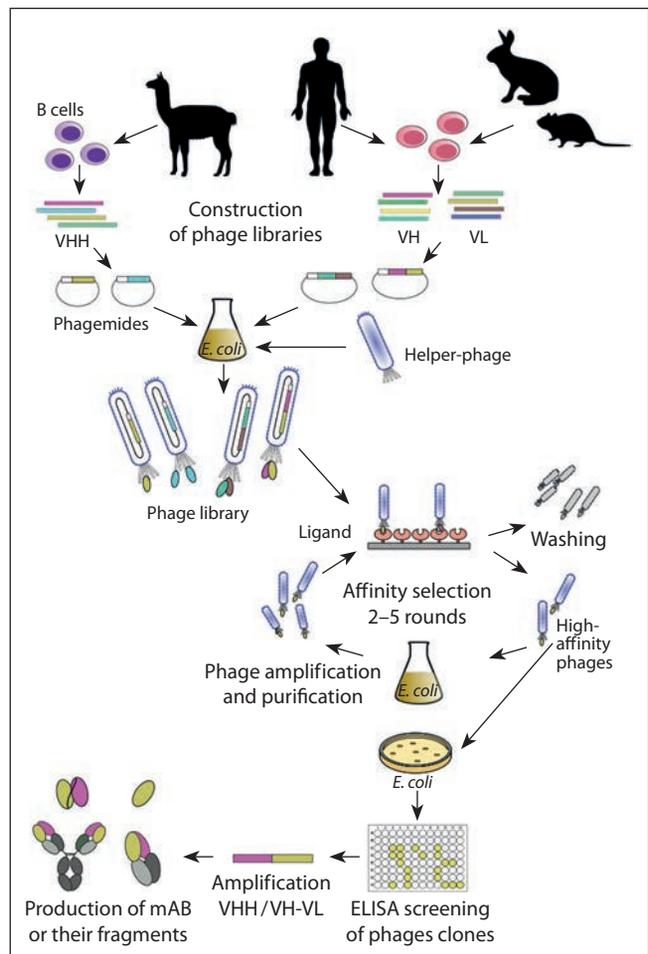
Fig. 1. Scheme representing generation of phage display antibody libraries and biopanning procedure.

(Alfaleh et al., 2020). The general scheme of the method is shown in Fig. 1.

The number of panning rounds can vary depending on whether a greater variety of clones or a greater specificity of fusing to an antigen is required. Additionally, double recognition panning against two antigens can be carried out for the selection of bispecific antibodies (Hammers, Stanley, 2014).

Phage display has plenty of applications: it is used for the development of antibacterial therapeutic agents (Christensen et al., 2001; Huang et al., 2012; Ashby et al., 2017), biosensors (Moon et al., 2019; Sozhamannan, Hofmann, 2020), identification of mAbs for treatment of dermatological, autoimmune diseases or cancers (Chan et al., 2014; Hammers, Stanley, 2014; Nixon et al., 2014; Alfaleh et al., 2020), as a platform for targeted drug and vaccine delivery (Clark, March, 2004; Petrenko, Jayanna, 2014; Nemudraya et al., 2016), as a tool for diagnostics and treatment of viral infections (Castel et al., 2011; Hess, Jewell, 2020). Phage display also has broad applications in the field of HIV-1 research: mapping epitopes recognized by HIV-neutralizing antibodies; searching for HIV-derived peptide mimics, which could be used a fusion inhibitors, components of vaccines

А.Н. Чикаев, А.П. Рудометов
Ю.А. Меркульева, Л.И. Карпенко

Применение фагового дисплея
для поиска ВИЧ-1-нейтрализующих антител

2021
25·5

and diagnostics; identifying HIV-neutralizing antibodies with broad neutralizing activity.

Below we review some examples of identification of HIV-1 broadly neutralizing antibodies using phage display technology.

Broadly neutralizing antibodies

One of the prominent features of human immunodeficiency virus is its phenomenal ability to evade the humoral immune response by rapidly mutating due to the low fidelity of HIV-1 reverse transcriptase which markedly enhances the genetic variation of the virus and makes it mutate very quickly – at the highest rate for any biological entity (Cuevas et al., 2015). Mutations occurring after each cycle of viral replication often cause structural changes in HIV immunodominant regions. As a result, the majority of antibodies elicited against HIV infection are strain-specific, and either are non-neutralizing or lose the ability to neutralize the virus after several replication cycles due to antigen escape. In this regard, it was believed that HIV-1 neutralizing antibodies could not be induced or it occurs extremely rarely (Mccoy, Burton, 2017). Nevertheless, these antibodies were lately found in so-called HIV long-term non-progressors – HIV-infected patients who do not develop immunodeficiency in the absence of antiretroviral therapy. Sera of non-progressors exhibited HIV-neutralizing activity not only against host strains, but also against a panel of different HIV-1 isolates (Dhillon et al., 2007; Walker et al., 2010; Sok, Burton, 2018; Dashti et al., 2019).

It was originally thought that induction of antibodies with broad HIV-1 neutralizing activity is a unique feature of non-progressors which provide them with the ability to control viremia for a long time (Montefiori et al., 1996). Later, it was shown that bnAbs is elicited in 20–50 % of all HIV-1 infected patients, but it takes a very long time before mature neutralizing antibodies can be arisen: the affinity maturation process may last up to several years from the moment of infection (Doria-Rose et al., 2009; Hraber et al., 2014; Rusert et al., 2016). It was shown that passive administration of a single bnAbs or its combinations to non-human primates completely protected animals against SHIV infection (Hessell et al., 2009; Moldt et al., 2012; Shingai et al., 2014). Moreover, passive transfer of bnAbs to HIV-infected individuals correlated with a long-term viral load reduction to undetectable levels (Lynch et al., 2015; Scheid et al., 2016), and in some cases, a host-protective humoral immune response was formed (Schoofs et al., 2016).

Today, the majority of HIV researchers acknowledge that an immunogen capable of eliciting broadly neutralizing antibodies may provide protection against HIV. Thus, the search for bnAbs and the development of immunogens aimed at elicitation of broadly neutralizing antibodies are among the most important tasks of modern vaccinology.

Phage display as a tool for identification of HIV-1 broadly neutralizing antibodies

The first studies devoted to the identification of bnAbs using phage display were published in the early 1990s. At that time, there were practically no data on broadly neutralizing antibodies. However, detailed information about the antigenic structure of HIV-1 had already been obtained, which resulted in the understanding that in order to provide effective protection against the virus, the humoral immune response must be targeted to the conserved viral epitopes, which are less susceptible to mutagenesis (Kowalski et al., 1987; Habeshaw et al., 1990; Putney, 1992). These fragments have been considered as the main targets for neutralizing antibodies. Subsequently, other HIV-1 antigenic determinants were discovered, which are critical for the HIV entry into the host cells, also known as sites of vulnerability (Shcherbakov et al., 2015; Kwong, Mascola, 2018). Up to date, at least seven sites on HIV Env that are vulnerable to antibody-mediated protection have been identified (Fig. 2).

bnAbs recognizing conserved regions of the gp120 glycoprotein

IgG1b12 was the first HIV-1 broadly neutralizing antibody derived using phage display (and also one of the first described bnAbs). In 1991, D.R. Burton et al. obtained an immune Fab library of bacteriophages based on B-cells from the bone marrow of an HIV-positive non-progressor (Barbas et al., 1991; Burton et al., 1991). After the panning of the resulting library against the HIV-1 IIIB gp120 glycoprotein, they selected phage clones that specifically bound with gp120. As a result, specific combinations of VH: VL genes were identified and expressed in the Fab format. It was shown that these Fabs were able to compete with the soluble CD4 molecule (sCD4) for binding to gp120 in ELISA (Burton et al., 1991). In their next work, the authors demonstrated the ability of selected Fabs to neutralize IIIB, MN and RF HIV-1 strains (Barbas et al., 1992).

Expanded screening of the phage library and more detailed analysis revealed a clone displaying the Fab fragment numbered b12, which binds with high affinity to the mature form of gp120 at the CD4bs (Roben et al., 1994). Later, a full-length recombinant IgG1b12 antibody was obtained, which became one of the first HIV-1 broadly neutralizing antibodies discovered (Burton et al., 1994).

Neutralization breadth of IgG1b12 has been repeatedly evaluated using different viral strains and panels of primary HIV-1 isolates. Depending on the panel used, b12 neutralized 30–63 % of the pseudovirus/primary isolates panels used in the experiment (at a concentration of IC50 < 50 μg/ml), whereas the highest neutralization potency was demonstrated against HIV-1 subtype B (clade B HIV-1) (Burton et al., 1994; Walker et al., 2009; Corti et al., 2010; Wu et al., 2010; Zhang et al., 2012; Gach et al., 2013). Before the advent of the second generation bnAbs, obtained by sorting affinity B-cells memory (Sok, Burton, 2018), among all detected at that time cross-neutralizing antibodies IgG1b12 was one of the leaders in the number of neutralized isolates HIV-1.

bNAbs specific to CD4 binding site

M.Y. Zhang et al. were the first researchers who identified two CD4bs-specific broadly neutralizing antibodies. They constructed Fab phage immune libraries derived from the bone marrow B-cells of an HIV-infected non-progressor
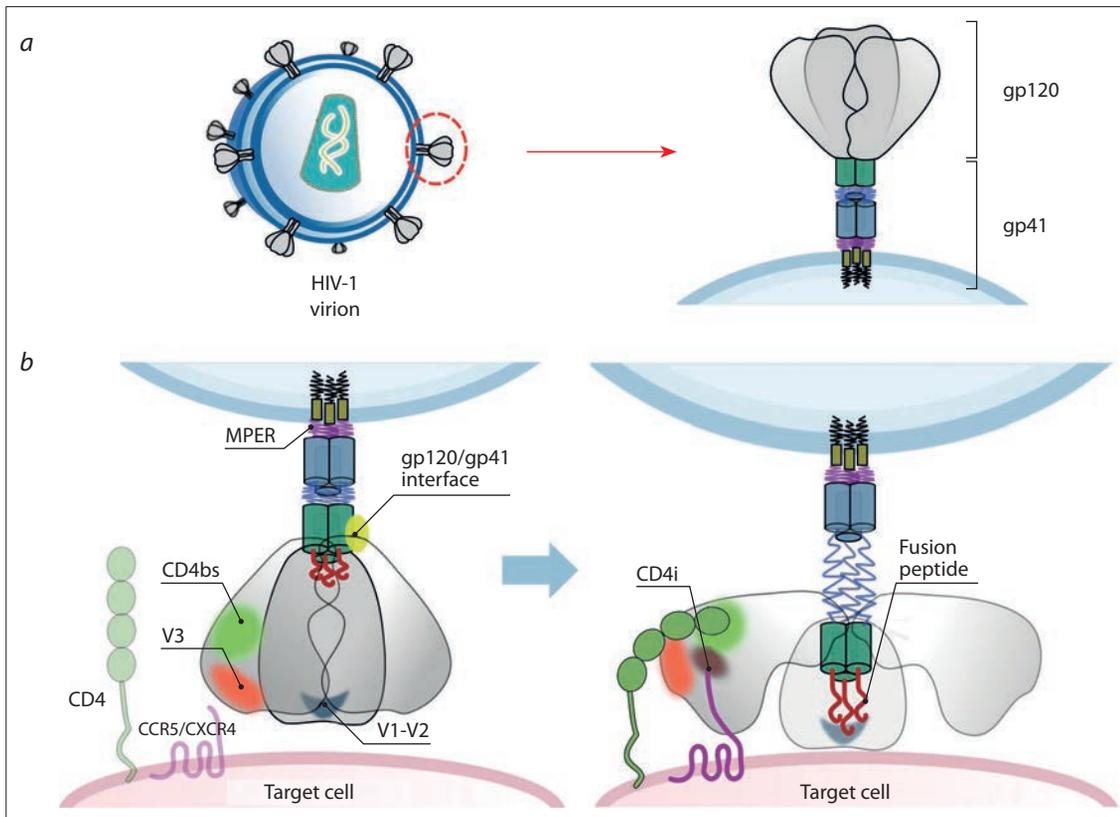
Fig. 2. Sites of HIV-1 vulnerability to bnAbs.

a – schematic representation of HIV-1 envelope glycoprotein (Env). Left – HIV-1 virion structure, right – mature HIV-1 envelope (Env) protein anchored to the viral membrane. Env is a trimer of heterodimers composed of gp41 (transmembrane subunit) and gp120 (surface subunit) molecules; *b* – conformational changes of HIV-1 Env trimers occurring during infection of host cell. Known HIV-1 sites of vulnerabilities to broadly neutralizing antibodies are indicated: gp120 binding site (CD4bs); V1/V2 glycan apex, loops V3 glycan loop of gp120 (V1/V2, V3); CD4-induced CCR5/CXCR4 binding site which is exposed following CD4 binding (CD4i); gp120/gp41 interface; membrane-proximal external region of gp41 (MPER); gp41 N-terminal fusion peptide, which anchors to the host cell membrane.

who had high titers of HIV-1-specific broadly neutralizing antibodies.

In order to identify clones that bind to the conserved antigenic determinants of the virus, biopanning was performed using two antigens. The first round of selection was carried out against sCD4 in complex with recombinant gp140$_{89.6}$ – a truncation form of gp160 Env of HIV-1 89.6 strain with removed transmembrane and cytoplasmic domains. In the second round, library was panned against HIV-1 IIIB gp140 and sCD4 complex. Subsequent selection rounds were carried out against gp140$_{89.6}$ and gp140IIIB molecules, respectively, with a gradual decrease in antigen concentration at each round. The binding affinity of the selected clones was evaluated by ELISA using single gp140$_{89.6}$/gp140IIIB molecules or in complex with sCD4.

The selected m18 clone showed the highest binding affinity to all of the antigens and was capable of neutralizing 11 out of 15 pseudoviruses bearing HIV-1 Envs from different strains (Zhang et al., 2003). A year later, the selected clones were re-screened by ELISA using the additional JR-FLgp120 antigen, resulting in selection of m14 Fab clone with enhanced affinity and neutralization breadth compared to m18 (Zhang et al., 2004a). Next, these clones were tested for neu-

tralization activity against extended panel of 30 HIV isolates. It was shown that m14 and m18 were able to neutralize about 21–23 % and 13–21 % of a panel, respectively, thereby demonstrating lower neutralization breadth compared to bnAb IgG1b12 (Zhang et al., 2012).

Using a similar approach, the same research group screened a Fab phage library derived from B-cells of R2 donor with high titers of cross-neutralizing antibodies. Two of the identified clones, m22 and m24, were expressed as Fab fragments of CD4bs-specific antibodies, which demonstrated neutralization potency and breadth similar to m14 and m18 (Zhang et al., 2006).

bNAbs specific to CCR5/CXCR4 coreceptor-binding sites of gp120

M. Moulard et al. screened a Fab antibody phage library (IgG1κ) derived from an HIV-infected individual against the gp120-CD4-CCR5 complex (Moulard et al., 2002). After five rounds of affinity selection, a Fab clone X5 with a unique CDR3 heavy-chain was selected. It was shown that the binding affinity of X5 to the CD4-gp120/CD4-gp140 complexes was significantly higher than that to the single gp120 and gp140 molecules, respectively. Addition of denatured

А.Н. Чикаев, А.П. Рудометов
Ю.А. Меркульева, Л.И. Карпенко

Применение фагового дисплея
для поиска ВИЧ-1-нейтрализующих антител

2021
25•5

CCR5 to the CD4-gp120 complex increased the X5 affinity, indicating that its epitope is formed by a CD4-dependent conformational change of gp120. X5 partially competed for binding to gp120 with other CD4i-specific antibodies, and with CD4bs-recognizing bnAb IgG1b12. Furthermore, X5 Fab neutralized 11 out of 12 primary HIV-1 isolates, thus demonstrating affinity, breadth, and potency comparable to the full-length IgG1b12 (Moulard et al., 2002).

However, the hypothesis that the full-length bivalent variant of IgG X5 would have even greater neutralizing activity was not confirmed. Apparently, the availability of the epitope recognized by X5 is limited by steric hindrance, which may cause the lack of binding efficacy of the larger molecules (Labrijn et al., 2003; Choudhry et al., 2006). It was thus concluded that the single-chain fragment of this X5 antibody possesses the highest activity and neutralization breadth, compared to its Fab and IgG variants (Choudhry et al., 2006).

Later, non-specific mutagenesis of the X5 scFv-encoding sequences was carried out, the resulting "mutant" phage sublibrary was screened against the oligomeric form of gp140$_{89.6}$ (which was non-homologous to gp120JR-FL) in complex with sCD4. Two scFvs, m6 and m9, capable of neutralizing 96 and 100 % of primary isolates from a panel comprising 33 different strains, were identified. Moreover, X5 neutralized only 45 % of the isolates from this panel (Zhang et al., 2004b). Neutralization assay performed using another panel of 30 HIV-1 different strains revealed that m9 neutralized 76 % of the primary isolates (Zhang et al., 2012).

bNabs specific to the MPER region of the gp41 glycoprotein

Another site of vulnerability of HIV-1, the membrane-proximal outer region of the gp41 glycoprotein (MPER), which is located between the transmembrane region and the gp41 C-terminal α-helical fragment, became another target of HIV bnAbs. Since MPER plays a crucial role in the process of viral fusion to the target cell, it is highly conservative and thus considered as one of the most promising targets for the development of antiviral drugs (Burton, Hangartner, 2016). The study performed by M. Zwick and his colleagues (2001) should be mentioned as one of the first attempts to search for MPER-specific bnAbs using phage display. They developed an immune Fab phage library based on cDNA of VH/VL genes isolated from the B cells of the bone marrow of an HIV-1 non-progressor who had a high titer of broadly neutralizing antibodies.

Two biopanning strategies were applied: in the first case, the HIV-1MN gp41 peptide MN 2031 comprising the MPER sequence was used as an antigen. In the second, the selection was carried out against a whole HIV-1MN virion. After the panning, several MPER-specific clones were identified, including Fab Z13 clone which had the highest affinity and neutralization breadth. Next, the authors created a "mutant" phage library displaying Fab Z13 with random mutations in the LCDR3 and screened it against the gp41 glycoprotein, in order to identify the Z13 clones with enhanced affinity. Among the selected phages, clone exposing Fab Z13e1 bound

to gp41 MPER with the highest affinity. Subsequently, a full-length Z13e1 IgG molecule was obtained, which provided a more-than-100-fold enhanced affinity for binding to MPER, and a significant increase in HIV-1-neutralizing activity compared to the initial IgG Z13 variant. The amount of neutralized isolates increased from 35 to 50 % (Zwick et al., 2001; Nelson et al., 2007).

bNabs recognizing the gp120–gp41 interface

Phage display was also used for identification of bnAbs that bind to the N-terminal domain of gp41, the so-called fusion peptide (see Fig. 2). Antibodies are able to bind to gp41 after conformational changes occurring during the last stages of HIV cell entry.

Such anti-gp41 antibodies were first described by M. Miller's research group (Miller et al., 2005). They used B-cells from the bone marrow of an HIV-negative patient for naive scFv phage library construction and subsequent identification of human monoclonal antibodies specific to the gp41 N-terminal region (NHR). The library was subsequently panned against the polymer that mimics gp41 6HB peptide complex, and then against the IZN36 peptide mimetic of the N-terminal heptad repeats (NHR). As a result, both antigen-binding phage clones were selected and used for reconstruction of corresponding soluble scFvs and full-length IgGs. Virus neutralization assay led to the detection of H/I1-BMV-D5 antibody capable of neutralizing 9 of 19 tested HIV-1 isolates (Miller et al., 2005).

Searching for bnAbs using phage display of single domain antibody fragments

In addition to the "classical" scFv/Fab phage display, phage libraries based on camelid single domain antibodies (nanobodies) were used to search for HIV bnAbs (see Fig. 1). Compared to IgG, nanobodies are more stable and have a smaller size, which can be beneficial for binding to sterically restricted antigenic determinants. Besides, the absence of light chains in the VHH structure facilitates gene manipulations and cloning procedures during library construction.

One of the first immune VHH phage libraries was obtained from llamas immunized with trimeric form of HIV-1 gp140 derived from a clade C (CN54). Total RNA of lymphocytes was isolated and used for cDNA synthesis; the repertoire of VHH genes was amplified and cloned into a phagemid vector, obtaining a library of phages displaying HIV-specific VHHs as pIII fusion proteins. Biopanning of the library led to the isolation of VHH A12, C8 and D7 that were able to neutralize 24 and 26 of 65 HIV-1 Env-pseudotyped virions from tier 1, 2, and 3 of various isolates (Forsman et al., 2008). Similarly, a phagemid immune library based on a VHH from llama immunized with clades A and B/C HIV-1 gp140 was obtained. After the biopanning, J3 and 3E3 clones were identified, which bound specifically to CD4bs and were able to neutralize 96 and 95 % of the pseudovirus panel (Mccoy et al., 2012, 2014; Strokappe et al., 2012).

Later, the same library was used for selection of VHH clones that were specific to the other HIV-1 glycoproteins:

1F10, which binds to the V3 loop of gp120; 1B5, which recognizes the CCR5 binding region, gp41-specific 2H10 and 2E7 clones. These nanobodies were capable of neutralizing from 45 to 80 % of Env pseudoviruses from the panels used (Lutje Hulsik et al., 2013; Strokappe et al., 2019). Bivalent nanobodies carrying VHHs with the highest neutralizing activity were also designed. The neutralization potency of these bispecific nanobodies increased approximately 1400-fold compared to the mixture of the individual VHHs; the highest efficiency of the nanobodies was observed against clade C HIV-1 viral strains (Lutje Hulsik et al., 2013; Strokappe et al., 2019).

K. Koch and colleagues (2017) separately prepared a phage immune library of HIV-specific VHHs using lymphocytes of camel immunized with soluble stabilized HIV-1 clade C gp140 Env trimer (SOSIP gp140). After affinity selection of the library, several CD4bs-specific nanobodies were identified, the best of which (VHH-9, VHH-28, VHH-A6) were capable of neutralizing 53, 65, and 77 % of a 21-isolate HIV-1 Env pseudovirus panel.

The authors of the above-mentioned studies emphasize that display of nanobody immune libraries via phage display is a convenient and effective alternative to the "traditional" scFv/Fab libraries for searching for high-affinity HIV-1 broadly-neutralizing antibodies. Small size and chemical stability of VHH facilitate various genetic manipulations directed to obtain clones with improved characteristics, such as site-targeted mutagenesis or creation of humanized and multivalent nanobodies specific to different regions of viral antigens. Finally, VHH phage libraries may be considered as a cheaper-to-manufacture alternative to full-sized human MAbs for HIV treatment (Weiss, Verrips, 2019).

Conclusion

Phage display technology played an essential role as a tool for searching, studying and epitope mapping of HIV-neutralizing antibodies. Phage display yielded the first HIV-1 bnAb, thereby leading to the extensive development of this research area. Soon, broadly neutralizing antibodies became a major focus of HIV vaccine design. Current methods for isolating HIV-specific bnAbs include the sorting of antigen-specific B cells with one memory on virus-like particles or variable loop removed recombinant viral proteins (Wu et al., 2010). These techniques, along with high-throughput screening of selected antibody clones (Walker et al., 2009), allowed the identification of second-generation HIV-1 bnAbs with markedly increased potency and breadth. The discovery of such antibodies capable of neutralizing more than 90 % of viral isolates has reinvigorated interest in the use of bnAbs in HIV-1 therapy.

Since 2010, more than 30 clinical trials of broadly neutralizing antibodies have been registered (Mahomed et al., 2021). Among them, 12 studies successfully passed Phase I, demonstrating safety of bnAbs and their combinations; four Phase II bnAb trials are currently underway, first data are expected to be made publicly available in 2021 (Julg, Barouch, 2019; Karuna, Corey, 2020; Mahomed et al., 2020;

Stephenson et al., 2020). Lastly, relying on the progress achieved in generating recombinant viral antigens (Jardine et al., 2013; Medina-Ramirez et al., 2017; Stamatatos et al., 2017; Duan et al., 2018), together with a vast amount of data accumulated on broadly neutralizing antibodies (Mascola, Haynes, 2013; Mouquet, Nussenzweig, 2013), novel strategies for design of HIV vaccines aimed for induction of 2nd generation bnAbs have been proposed (Del Moral-Sanchez, Sliepen, 2019).

Hence, bNAbs represent a promising novel approach for effective HIV-1 immunotherapy and prevention. Thus, today bnAbs are one of the most important objects in the study of HIV infection. It is likely that in the foreseeable future they will become a worthy alternative to existing antiretroviral therapy, and in the longer term, one can expect the emergence of preventive vaccines that induce their production.

References

Alfaleh M.A., Alsaab H.O., Mahmoud A.B., Alkayyal A.A., Jones M.L., Mahler S.M., Hashem A.M. Phage display derived monoclonal antibodies: from bench to bedside. *Front. Immunol.* 2020;11:1986. DOI 10.3389/fimmu.2020.01986.

Ashby M., Petkova A., Gani J., Mikut R., Hilpert K. Use of peptide libraries for identification and optimization of novel antimicrobial peptides. *Curr. Top. Med. Chem.* 2017;17(5):537-553. DOI 10.2174/1568026616666160713125555.

Barbas C.F., 3rd, Crowe J.E., Jr., Cababa D., Jones T.M., Zebedee S.L., Murphy B.R., Chanock R.M., Burton D.R. Human monoclonal Fab fragments derived from a combinatorial library bind to respiratory syncytial virus F glycoprotein and neutralize infectivity. *Proc. Natl. Acad. Sci. USA.* 1992;89(21):10164-10168. DOI 10.1073/pnas.89.21.10164.

Barbas C.F., 3rd, Kang A.S., Lerner R.A., Benkovic S.J. Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proc. Natl. Acad. Sci. USA.* 1991;88(18):7978-7982. DOI 10.1073/pnas.88.18.7978.

Burton D.R., Barbas C.F., 3rd, Persson M.A., Koenig S., Chanock R.M., Lerner R.A. A large array of human monoclonal antibodies to type 1 human immunodeficiency virus from combinatorial libraries of asymptomatic seropositive individuals. *Proc. Natl. Acad. Sci. USA.* 1991;88(22):10134-10137. DOI 10.1073/pnas.88.22.10134.

Burton D.R., Hangartner L. Broadly neutralizing antibodies to HIV and their role in vaccine design. *Annu. Rev. Immunol.* 2016;34:635-659. DOI 10.1146/annurev-immunol-041015-055515.

Burton D.R., Pyati J., Koduri R., Sharp S.J., Thornton G.B., Parren P.W., Sawyer L.S., Hendry R.M., Dunlop N., Nara P.L. Efficient neutralization of primary isolates of HIV-1 by a recombinant human monoclonal antibody. *Science.* 1994;266(5187):1024-1027. DOI 10.1126/science.7973652.

Castel G., Chteoui M., Heyd B., Tordo N. Phage display of combinatorial peptide libraries: application to antiviral research. *Molecules.* 2011;16(5):3499-3518. DOI 10.3390/molecules16053499.

Chames P., Baty D. Phage display and selections on biotinylated antigens. In: Kotermann R., Dübel S. (Eds.). Antibody Engineering. Humana Press, 2010;151-164. DOI 10.1007/978-3-642-01144-3_11.

Chan C.E., Lim A.P., Macary P.A., Hanson B.J. The role of phage display in therapeutic antibody discovery. *Int. Immunol.* 2014;26(12):649-657. DOI 10.1093/intimm/dxu082.

Choudhry V., Zhang M.Y., Dimitrova D., Prabakaran P., Dimitrov A.S., Fouts T.R., Dimitrov D.S. Antibody-based inhibitors of HIV infection. *Expert Opin. Biol. Ther.* 2006;6(5):523-531. DOI 10.1517/14712598.6.5.523.

А.Н. Чикаев, А.П. Рудометов
Ю.А. Меркульева, Л.И. Карпенко

Применение фагового дисплея
для поиска ВИЧ-1-нейтрализующих антител

2021
25•5

Christensen D.J., Gottlin E.B., Benson R.E., Hamilton P.T. Phage display for target-based antibacterial drug discovery. *Drug Discov. Today.* 2001;6(14):721-727. DOI 10.1016/s1359-6446(01)01853-0.

Clackson T., Hoogenboom H.R., Griffiths A.D., Winter G. Making antibody fragments using phage display libraries. *Nature.* 1991; 352(6336):624-628. DOI 10.1038/352624a0.

Clark J.R., March J.B. Bacteriophage-mediated nucleic acid immunisation. *FEMS Immunol. Med. Microbiol.* 2004;40(1):21-26. DOI 10.1016/S0928-8244(03)00344-4.

Corti D., Langedijk J.P., Hinz A., Seaman M.S., Vanzetta F., Fernandez-Rodriguez B.M., Silacci C., Pinna D., Jarrossay D., Balla-Jhagjhoor-singh S., Willems B., Zekveld M.J., Dreja H., O'sullivan E., Pade C., Orkin C., Jeffs S.A., Montefiori D.C., Davis D., Weissenhorn W., Mcknight A., Heeney J.L., Sallusto F., Sattentau Q.J., Weiss R.A., Lanzavecchia A. Analysis of memory B cell responses and isolation of novel monoclonal antibodies with neutralizing breadth from HIV-1-infected individuals. *PLoS One.* 2010;5(1):e8805. DOI 10.1371/journal.pone.0008805.

Cuevas J.M., Geller R., Garijo R., Lopez-Aldeguer J., Sanjuan R. Extremely high mutation rate of HIV-1 *in vivo. PLoS Biol.* 2015;13(9): e1002251. DOI 10.1371/journal.pbio.1002251.

Dashti A., Devico A.L., Lewis G.K., Sajadi M.M. Broadly neutralizing antibodies against HIV: back to blood. *Trends Mol. Med.* 2019; 25(3):228-240. DOI 10.1016/j.molmed.2019.01.007.

Davies J., Riechmann L. Antibody VH domains as small recognition units. *Biotechnology* (*NY*)*.* 1995;13(5):475-479. DOI 10.1038/nbt 0595-475.

Del Moral-Sanchez I., Sliepen K. Strategies for inducing effective neutralizing antibody responses against HIV-1. *Expert Rev. Vaccines.* 2019;18(11):1127-1143. DOI 10.1080/14760584.2019.1690458.

Dhillon A.K., Donners H., Pantophlet R., Johnson W.E., Decker J.M., Shaw G.M., Lee F.H., Richman D.D., Doms R.W., Vanham G., Burton D.R. Dissecting the neutralizing antibody specificities of broadly neutralizing sera from human immunodeficiency virus type 1-infected donors. *J. Virol.* 2007;81(12):6548-6562. DOI 10.1128/JVI. 02749-06.

Doria-Rose N.A., Klein R.M., Manion M.M., O'dell S., Phogat A., Chakrabarti B., Hallahan C.W., Migueles S.A., Wrammert J., Ahmed R., Nason M., Wyatt R.T., Mascola J.R., Connors M. Frequency and phenotype of human immunodeficiency virus envelope-specific B cells from patients with broadly cross-neutralizing antibodies. *J. Virol.* 2009;83(1):188-199. DOI 10.1128/JVI. 01583-08.

Duan H., Chen X., Boyington J.C., Cheng C., Zhang Y., Jafari A.J., Stephens T., Tsybovsky Y., Kalyuzhniy O., Zhao P., Menis S., Nason M.C., Normandin E., Mukhamedova M., Dekosky B.J., Wells L., Schief W.R., Tian M., Alt F.W., Kwong P.D., Mascola J.R. Glycan masking focuses immune responses to the HIV-1 CD4-binding site and enhances elicitation of VRC01-class precursor antibodies. *Immunity.* 2018;49(2):301-311 e305. DOI 10.1016/j.immuni.2018. 07.005.

Felici F., Castagnoli L., Musacchio A., Jappelli R., Cesareni G. Selection of antibody ligands from a large library of oligopeptides expressed on a multivalent exposition vector. *J. Mol. Biol.* 1991;222(2):301-310. DOI 10.1016/0022-2836(91)90213-p.

Forsman A., Beirnaert E., Aasa-Chapman M.M., Hoorelbeke B., Hijazi K., Koh W., Tack V., Szynol A., Kelly C., Mcknight A., Verrips T., De Haard H., Weiss R.A. Llama antibody fragments with cross-subtype human immunodeficiency virus type 1 (HIV-1)-neutralizing properties and high affinity for HIV-1 gp120. *J. Virol.* 2008;82(24): 12069-12081. DOI 10.1128/JVI.01379-08.

Gach J.S., Quendler H., Tong T., Narayan K.M., Du S.X., Whalen R.G., Binley J.M., Forthal D.N., Poignard P., Zwick M.B. A human antibody to the CD4 binding site of gp120 capable of highly potent but sporadic cross clade neutralization of primary HIV-1. *PLoS One.* 2013;8(8):e72054. DOI 10.1371/journal.pone.0072054.

Greenberg A.S., Avila D., Hughes M., Hughes A., McKinney E.C., Flajnik M.F. A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. *Nature.* 1995;374(6518):168-173. DOI 10.1038/374168a0.

Griffiths A.D., Duncan A.R. Strategies for selection of antibodies by phage display. *Curr. Opin. Biotechnol.* 1998;9(1):102-108. DOI 10.1016/s0958-1669(98)80092-x.

Habeshaw J.A., Dalgleish A.G., Bountiff L., Newell A.L., Wilks D., Walker L.C., Manca F. AIDS pathogenesis: HIV envelope and its interaction with cell proteins. *Immunol. Today.* 1990;11(11):418-425. DOI 10.1016/0167-5699(90)90162-3.

Hammers C.M., Stanley J.R. Antibody phage display: technique and applications. *J. Invest. Dermatol.* 2014;134(2):1-5. DOI 10.1038/jid. 2013.521.

Hess K.L., Jewell C.M. Phage display as a tool for vaccine and immunotherapy development. *Bioeng. Transl. Med.* 2020;5(1):e10142. DOI 10.1002/btm2.10142.

Hessell A.J., Rakasz E.G., Poignard P., Hangartner L., Landucci G., Forthal D.N., Koff W.C., Watkins D.I., Burton D.R. Broadly neutralizing human anti-HIV antibody 2G12 is effective in protection against mucosal SHIV challenge even at low serum neutralizing titers. *PLoS Pathog.* 2009;5(5):e1000433. DOI 10.1371/journal.ppat. 1000433.

Hraber P., Korber B.T., Lapedes A.S., Bailer R.T., Seaman M.S., Gao H., Greene K.M., Mccutchan F., Williamson C., Kim J.H., Tovanabutra S., Hahn B.H., Swanstrom R., Thomson M.M., Gao F., Harris L., Giorgi E., Hengartner N., Bhattacharya T., Mascola J.R., Montefiori D.C. Impact of clade geography and age of the epidemic on HIV-1 neutralization by antibodies. *J. Virol.* 2014;88(21):12623-12643. DOI 10.1128/JVI.01705-14.

Huang J.X., Bishop-Hurley S.L., Cooper M.A. Development of anti-infectives using phage display: biological agents against bacteria viruses and parasites. *Antimicrob. Agents Chemother.* 2012;56(9): 4569-4582. DOI 10.1128/AAC.00567-12.

Ilyichev A.A., Minenkova O.O., Kishchenko G.P., Tat'kov S.I., Karpishev N.N., Eroshkin A.M., Ofitzerov V.I., Akimenko Z.A., Petrenko V.A., Sandakhchiev L.S. Inserting foreign peptides into the major coat protein of bacteriophage M13. *FEBS Lett.* 1992;301(3):322-324. DOI 10.1016/0014-5793(92)80267-k.

Jardine J., Julien J.P., Menis S., Ota T., Kalyuzhniy O., Mcguire A., Sok D., Huang P.S., Macpherson S., Jones M., Nieusma T., Mathison J., Baker D., Ward A.B., Burton D.R., Stamatatos L., Nemazee D., Wilson I.A., Schief W.R. Rational HIV immunogen design to target specific germline B cell receptors. *Science.* 2013;340(6133):711-716. DOI 10.1126/science.1234150.

Julg B., Barouch D.H. Neutralizing antibodies for HIV-1 prevention. *Curr. Opin. HIV AIDS.* 2019;14(4):318-324. DOI 10.1097/COH. 0000000000000556.

Karuna S.T., Corey L. Broadly neutralizing antibodies for HIV prevention. *Annu. Rev. Med.* 2020;71(1):329-346. DOI 10.1146/annurev-med-110118-045506.

Kay B.K., Adey N.B., He Y.S., Manfredi J.P., Mataragnon A.H., Fowlkes D.M. An M13 phage library displaying random 38-amino-acid peptides as a source of novel sequences with affinity to selected targets. *Gene.* 1993;128(1):59-65. DOI 10.1016/0378-1119(93)90153-t.

Kennedy P.J., Oliveira C., Granja P.L., Sarmento B. Monoclonal antibodies: technologies for early discovery and engineering. *Crit. Rev. Biotechnol.* 2018;38(3):394-408. DOI 10.1080/07388551.2017.135 7002.

Kishchenko G., Batliwala H., Makowski L. Structure of a foreign peptide displayed on the surface of bacteriophage M13. *J. Mol. Biol.* 1994;241(2):208-213. DOI 10.1006/jmbi.1994.1489.

Koch K., Kalusche S., Torres J.L., Stanfield R.L., Danquah W., Khazanehdari K., Von Briesen H., Geertsma E.R., Wilson I.A., Wernery U., Koch-Nolte F., Ward A.B., Dietrich U. Selection of nanobodies with broad neutralizing potential against primary HIV-1 strains using soluble subtype C gp140 envelope trimers. *Sci. Rep.* 2017;7(1):8390. DOI 10.1038/s41598-017-08273-7.

Kowalski M., Potz J., Basiripour L., Dorfman T., Goh W.C., Terwilliger E., Dayton A., Rosen C., Haseltine W., Sodroski J. Functional regions of the envelope glycoprotein of human immunodeficiency virus type 1. *Science.* 1987;237(4820):1351-1355. DOI 10.1126/science.3629244.

Kwong P.D., Mascola J.R. HIV-1 vaccines based on antibody identification B cell ontogeny and epitope structure. *Immunity.* 2018;48(5): 855-871. DOI 10.1016/j.immuni.2018.04.029.

Labrijn A.F., Poignard P., Raja A., Zwick M.B., Delgado K., Franti M., Binley J., Vivona V., Grundner C., Huang C.C., Venturi M., Petropoulos C.J., Wrin T., Dimitrov D.S., Robinson J., Kwong P.D., Wyatt R.T., Sodroski J., Burton D.R. Access of antibody molecules to the conserved coreceptor binding site on glycoprotein gp120 is sterically restricted on primary human immunodeficiency virus type 1. *J. Virol.* 2003;77(19):10557-10565. DOI 10.1128/jvi.77.19.10557-10565.2003.

Ledsgaard L., Kilstrup M., Karatt-Vellatt A., McCafferty J., Laustsen A.H. Basics of antibody phage display technology. *Toxins* (*Basel*). 2018;10(6). DOI 10.3390/toxins10060236.

Lutje Hulsik D., Liu Y.Y., Strokappe N.M., Battella S., El Khattabi M., Mccoy L.E., Sabin C., Hinz A., Hock M., Macheboeuf P., Bonvin A.M., Langedijk J.P., Davis D., Forsman Quigley A., Aasa-Chapman M.M., Seaman M.S., Ramos A., Poignard P., Favier A., Simorre J.P., Weiss R.A., Verrips C.T., Weissenhorn W., Rutten L. A gp41 MPER-specific llama VHH requires a hydrophobic CDR3 for neutralization but not for antigen recognition. *PLoS Pathog.* 2013;9(3):e1003202. DOI 10.1371/journal.ppat.1003202.

Lynch R.M., Boritz E., Coates E.E., Dezure A., Madden P., Costner P., Enama M.E., Plummer S., Holman L., Hendel C.S., Gordon I., Casazza J., Conan-Cibotti M., Migueles S.A., Tressler R., Bailer R.T., Mcdermott A., Narpala S., O'dell S., Wolf G., Lifson J.D., Freemire B.A., Gorelick R.J., Pandey J.P., Mohan S., Chomont N., Fromentin R., Chun T.W., Fauci A.S., Schwartz R.M., Koup R.A., Douek D.C., Hu Z., Capparelli E., Graham B.S., Mascola J.R., Ledgerwood J.E., VRC 601 Study Team. Virologic effects of broadly neutralizing antibody VRC01 administration during chronic HIV-1 infection. *Sci. Transl. Med.* 2015;7(319):319ra206. DOI 10.1126/scitranslmed.aad5752.

Mahomed S., Garrett N., Baxter C., Abdool Karim Q., Abdool Karim S.S. Clinical trials of broadly neutralizing monoclonal antibodies for Human Immunodeficiency Virus prevention: a review. *J. Infect. Dis.* 2021;223(3):370-380. DOI 10.1093/infdis/jiaa377.

Mahomed S., Garrett N., Karim Q.A., Zuma N.Y., Capparelli E., Baxter C., Gengiah T., Archary D., Samsunder N., Rose N.D., Moore P., Williamson C., Barouch D.H., Fast P.E., Pozzetto B., Hankins C., Carlton K., Ledgerwood J., Morris L., Mascola J., Abdool Karim S. Assessing the safety and pharmacokinetics of the anti-HIV monoclonal antibody CAP256V2LS alone and in combination with VRC07-523LS and PGT121 in South African women: study protocol for the first-in-human CAPRISA 012B phase I clinical trial. *BMJ Open.* 2020;10(11):e042247. DOI 10.1136/bmjopen-2020-042247.

Mascola J.R., Haynes B.F. HIV-1 neutralizing antibodies: understanding nature's pathways. *Immunol. Rev.* 2013;254(1):225-244. DOI 10.1111/imr.12075.

McCafferty J., Griffiths A.D., Winter G., Chiswell D.J. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature.* 1990;348(6301):552-554. DOI 10.1038/348552a0.

Mccoy L.E., Burton D.R. Identification and specificity of broadly neutralizing antibodies against HIV. *Immunol. Rev.* 2017;275(1):11-20. DOI 10.1111/imr.12484.

Mccoy L.E., Quigley A.F., Strokappe N.M., Bulmer-Thomas B., Seaman M.S., Mortier D., Rutten L., Chander N., Edwards C.J., Ketteler R., Davis D., Verrips T., Weiss R.A. Potent and broad neutralization of HIV-1 by a llama antibody elicited by immunization. *J. Exp. Med.* 2012;209(6):1091-1103. DOI 10.1084/jem.20112655.

Mccoy L.E., Rutten L., Frampton D., Anderson I., Granger L., Bashford-Rogers R., Dekkers G., Strokappe N.M., Seaman M.S., Koh W., Grippo V., Kliche A., Verrips T., Kellam P., Fassati A., Weiss R.A. Molecular evolution of broadly neutralizing llama antibodies to the CD4-binding site of HIV-1. *PLoS Pathog.* 2014;10(12):e1004552. DOI 10.1371/journal.ppat.1004552.

Medina-Ramirez M., Garces F., Escolano A., Skog P., De Taeye S.W., Del Moral-Sanchez I., Mcguire A.T., Yasmeen A., Behrens A.J., Ozorowski G., Van Den Kerkhof T., Freund N.T., Dosenovic P., Hua Y., Gitlin A.D., Cupo A., Van Der Woude P., Golabek M., Sliepen K., Blane T., Kootstra N., Van Breemen M.J., Pritchard L.K., Stanfield R.L., Crispin M., Ward A.B., Stamatatos L., Klasse P.J., Moore J.P., Nemazee D., Nussenzweig M.C., Wilson I.A., Sanders R.W. Design and crystal structure of a native-like HIV-1 envelope trimer that engages multiple broadly neutralizing antibody precursors *in vivo*. *J. Exp. Med.* 2017;214(9):2573-2590. DOI 10.1084/jem.20161160.

Miller M.D., Geleziunas R., Bianchi E., Lennard S., Hrin R., Zhang H., Lu M., An Z., Ingallinella P., Finotto M., Mattu M., Finnefrock A.C., Bramhill D., Cook J., Eckert D.M., Hampton R., Patel M., Jarantow S., Joyce J., Ciliberto G., Cortese R., Lu P., Strohl W., Schleif W., Mcelhaugh M., Lane S., Lloyd C., Lowe D., Osbourn J., Vaughan T., Emini E., Barbato G., Kim P.S., Hazuda D.J., Shiver J.W., Pessi A. A human monoclonal antibody neutralizes diverse HIV-1 isolates by binding a critical gp41 epitope. *Proc. Natl. Acad. Sci. USA.* 2005; 102(41):14759-14764. DOI 10.1073/pnas.0506927102.

Minenkova O.O., Ilyichev A.A., Kishchenko G.P., Petrenko V.A. Design of specific immunogens using filamentous phage as the carrier. *Gene.* 1993;128(1):85-88. DOI 10.1016/0378-1119(93) 90157-x.

Moldt B., Rakasz E.G., Schultz N., Chan-Hui P.Y., Swiderek K., Weisgrau K.L., Piaskowski S.M., Bergman Z., Watkins D.I., Poignard P., Burton D.R. Highly potent HIV-specific antibody neutralization *in vitro* translates into effective protection against mucosal SHIV challenge *in vivo*. *Proc. Natl. Acad. Sci. USA.* 2012;109(46):18921-18925. DOI 10.1073/pnas.1214785109.

Montefiori D.C., Baba T.W., Li A., Bilska M., Ruprecht R.M. Neutralizing and infection-enhancing antibody responses do not correlate with the differential pathogenicity of SIVmac239delta3 in adult and infant rhesus monkeys. *J. Immunol.* 1996;157(12):5528-5535.

Moon J.S., Choi E.J., Jeong N.N., Sohn J.R., Han D.W., Oh J.W. Research progress of M13 bacteriophage-based biosensors. *Nanomaterials* (*Basel*). 2019;9(10). DOI 10.3390/nano9101448.

Moulard M., Phogat S.K., Shu Y., Labrijn A.F., Xiao X., Binley J.M., Zhang M.Y., Sidorov I.A., Broder C.C., Robinson J., Parren P.W., Burton D.R., Dimitrov D.S. Broadly cross-reactive HIV-1-neutralizing human monoclonal Fab selected for binding to gp120-CD4-CCR5 complexes. *Proc. Natl. Acad. Sci. USA.* 2002;99(10):6913-6918. DOI 10.1073/pnas.102562599.

Mouquet H., Nussenzweig M.C. HIV: Roadmaps to a vaccine. *Nature.* 2013;496(7446):441-442. DOI 10.1038/nature12091.

Mullen L.M., Nair S.P., Ward J.M., Rycroft A.N., Henderson B. Phage display in the study of infectious diseases. *Trends Microbiol.* 2006; 14(3):141-147. DOI 10.1016/j.tim.2006.01.006.

Muyldermans S. A guide to: generation and design of nanobodies. *FEBS J.* 2021;288(7):2084-2102. DOI 10.1111/febs.15515.

Nelson J.D., Brunel F.M., Jensen R., Crooks E.T., Cardoso R.M., Wang M., Hessell A., Wilson I.A., Binley J.M., Dawson P.E., Burton D.R., Zwick M.B. An affinity-enhanced neutralizing antibody against the membrane-proximal external region of human immunodeficiency virus type 1 gp41 recognizes an epitope between those

А.Н. Чикаев, А.П. Рудометов
Ю.А. Меркульева, Л.И. Карпенко

Применение фагового дисплея
для поиска ВИЧ-1-нейтрализующих антител

2021
25•5

of 2F5 and 4E10. *J. Virol.* 2007;81(8):4033-4043. DOI 10.1128/JVI.02588-06.

Nemudraya A.A., Richter V.A., Kuligina E.V. Phage peptide libraries as a source of targeted ligands. *Acta Naturae.* 2016;8(1):48-57. DOI 10.32607/20758251-2016-8-1-48-57.

Nixon A.E., Sexton D.J., Ladner R.C. Drugs derived from phage display: from candidate identification to clinical practice. *MAbs.* 2014;6(1):73-85. DOI 10.4161/mabs.27240.

Petrenko V.A., Jayanna P.K. Phage protein-targeted cancer nanomedicines. *FEBS Lett.* 2014;588(2):341-349. DOI 10.1016/j.febslet.2013.11.011.

Putney S. How antibodies block HIV infection: paths to an AIDS vaccine. *Trends Biochem. Sci.* 1992;17(5):191-196. DOI 10.1016/0968-0004(92)90265-b.

Roben P., Moore J.P., Thali M., Sodroski J., Barbas C.F., 3rd, Burton D.R. Recognition properties of a panel of human recombinant Fab fragments to the CD4 binding site of gp120 that show differing abilities to neutralize human immunodeficiency virus type 1. *J. Virol.* 1994;68(8):4821-4828. DOI 10.1128/JVI.68.8.4821-4828.1994.

Rusert P., Kouyos R.D., Kadelka C., Ebner H., Schanz M., Huber M., Braun D.L., Hoze N., Scherrer A., Magnus C., Weber J., Uhr T., Cippa V., Thorball C.W., Kuster H., Cavassini M., Bernasconi E., Hoffmann M., Calmy A., Battegay M., Rauch A., Yerly S., Aubert V., Klimkait T., Boni J., Fellay J., Regoes R.R., Gunthard H.F., Trkola A., Swiss Hivcs T. Determinants of HIV-1 broadly neutralizing antibody induction. *Nat. Med.* 2016;22(11):1260-1267. DOI 10.1038/nm.4187.

Scheid J.F., Horwitz J.A., Bar-On Y., Kreider E.F., Lu C.L., Lorenzi J.C., Feldmann A., Braunschweig M., Nogueira L., Oliveira T., Shimeliovich I., Patel R., Burke L., Cohen Y.Z., Hadrigan S., Settler A., Witmer-Pack M., West A.P., Jr., Juelg B., Keler T., Hawthorne T., Zingman B., Gulick R.M., Pfeifer N., Learn G.H., Seaman M.S., Bjorkman P.J., Klein F., Schlesinger S.J., Walker B.D., Hahn B.H., Nussenzweig M.C., Caskey M. HIV-1 antibody 3BNC117 suppresses viral rebound in humans during treatment interruption. *Nature.* 2016;535(7613):556-560. DOI 10.1038/nature18929.

Schoofs T., Klein F., Braunschweig M., Kreider E.F., Feldmann A., Nogueira L., Oliveira T., Lorenzi J.C., Parrish E.H., Learn G.H., West A.P., Jr., Bjorkman P.J., Schlesinger S.J., Seaman M.S., Czartoski J., Mcelrath M.J., Pfeifer N., Hahn B.H., Caskey M., Nussenzweig M.C. HIV-1 therapy with monoclonal antibody 3BNC117 elicits host immune responses against HIV-1. *Science.* 2016;352(6288):997-1001. DOI 10.1126/science.aaf0972.

Scott J.K., Smith G.P. Searching for peptide ligands with an epitope library. *Science.* 1990;249(4967):386-390. DOI 10.1126/science.1696028.

Shcherbakov D.N., Bakulina A.Y., Karpenko L.I., Ilyichev A.A. Broadly neutralizing antibodies against HIV-1 as a novel aspect of the immune response. *Acta Naturae.* 2015;7(4):11-21.

Shingai M., Donau O.K., Plishka R.J., Buckler-White A., Mascola J.R., Nabel G.J., Nason M.C., Montefiori D., Moldt B., Poignard P., Diskin R., Bjorkman P.J., Eckhaus M.A., Klein F., Mouquet H., Cetrulo Lorenzi J.C., Gazumyan A., Burton D.R., Nussenzweig M.C., Martin M.A., Nishimura Y. Passive transfer of modest titers of potent and broadly neutralizing anti-HIV monoclonal antibodies block SHIV infection in macaques. *J. Exp. Med.* 2014;211(10):2061-2074. DOI 10.1084/jem.20132494.

Skerra A., Pluckthun A. Assembly of a functional immunoglobulin Fv fragment in *Escherichia coli. Science.* 1988;240(4855):1038-1041. DOI 10.1126/science.3285470.

Smith G.P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science.* 1985;228(4705):1315-1317. DOI 10.1126/science.4001944.

Smith G.P., Petrenko V.A. Phage display. *Chem. Rev.* 1997;97(2):391-410. DOI 10.1021/cr960065d.

Sok D., Burton D.R. Recent progress in broadly neutralizing antibodies to HIV. *Nat. Immunol.* 2018;19(11):1179-1188. DOI 10.1038/s41590-018-0235-7.

Sozhamannan S., Hofmann E.R. The state of the art in biodefense related bacterial pathogen detection using bacteriophages: how it started and how it's going. *Viruses.* 2020;12(12). DOI 10.3390/v12121393.

Stamatatos L., Pancera M., Mcguire A.T. Germline-targeting immunogens. *Immunol. Rev.* 2017;275(1):203-216. DOI 10.1111/imr.12483.

Stephenson K.E., Wagh K., Korber B., Barouch D.H. Vaccines and broadly neutralizing antibodies for HIV-1 prevention. *Annu. Rev. Immunol.* 2020;38(673-703). DOI 10.1146/annurev-immunol-080219-023629.

Strokappe N.M., Hock M., Rutten L., Mccoy L.E., Back J.W., Caillat C., Haffke M., Weiss R.A., Weissenhorn W., Verrips T. Super potent bispecific llama VHH antibodies neutralize HIV via a combination of gp41 and gp120 epitopes. *Antibodies* (*Basel*). 2019;8(2). DOI 10.3390/antib8020038.

Strokappe N., Szynol A., Aasa-Chapman M., Gorlani A., Forsman Quigley A., Hulsik D.L., Chen L., Weiss R., De Haard H., Verrips T. Llama antibody fragments recognizing various epitopes of the CD4bs neutralize a broad range of HIV-1 subtypes A, B and C. *PLoS One.* 2012;7(3):e33298. DOI 10.1371/journal.pone.0033298.

Tikunova N.V., Morozova V.V. Phage display on the base of filamentous bacteriophages: application for recombinant antibodies selection. *Acta Naturae.* 2009;1(3):20-28.

Walker L.M., Phogat S.K., Chan-Hui P.Y., Wagner D., Phung P., Goss J.L., Wrin T., Simek M.D., Fling S., Mitcham J.L., Lehrman J.K., Priddy F.H., Olsen O.A., Frey S.M., Hammond P.W., Investigators P.G.P., Kaminsky S., Zamb T., Moyle M., Koff W.C., Poignard P., Burton D.R. Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science.* 2009;326(5950):285-289. DOI 10.1126/science.1178746.

Walker L.M., Simek M.D., Priddy F., Gach J.S., Wagner D., Zwick M.B., Phogat S.K., Poignard P., Burton D.R. A limited number of antibody specificities mediate broad and potent serum neutralization in selected HIV-1 infected individuals. *PLoS Pathog.* 2010;6(8):e1001028. DOI 10.1371/journal.ppat.1001028.

Weiss R.A., Verrips C.T. Nanobodies that neutralize HIV. *Vaccines* (*Basel*). 2019;7(3). DOI 10.3390/vaccines7030077.

Winter G., Griffiths A.D., Hawkins R.E., Hoogenboom H.R. Making antibodies by phage display technology. *Annu. Rev. Immunol.* 1994;12(433-455). DOI 10.1146/annurev.iy.12.040194.002245.

Wu X., Yang Z.Y., Li Y., Hogerkorp C.M., Schief W.R., Seaman M.S., Zhou T., Schmidt S.D., Wu L., Xu L., Longo N.S., Mckee K., O'dell S., Louder M.K., Wycuff D.L., Feng Y., Nason M., Doria-Rose N., Connors M., Kwong P.D., Roederer M., Wyatt R.T., Nabel G.J., Mascola J.R. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science.* 2010;329(5993):856-861. DOI 10.1126/science.1187659.

Zhang M.Y., Choudhry V., Sidorov I.A., Tenev V., Vu B.K., Choudhary A., Lu H., Stiegler G.M., Katinger H.W., Jiang S., Broder C.C., Dimitrov D.S. Selection of a novel gp41-specific HIV-1 neutralizing human antibody by competitive antigen panning. *J. Immunol. Methods.* 2006;317(1-2):21-30. DOI 10.1016/j.jim.2006.09.016.

Zhang M.Y., Shu Y., Phogat S., Xiao X., Cham F., Bouma P., Choudhary A., Feng Y.R., Sanz I., Rybak S., Broder C.C., Quinnan G.V., Evans T., Dimitrov D.S. Broadly cross-reactive HIV neutralizing human monoclonal antibody Fab selected by sequential antigen panning of a phage display library. *J. Immunol. Methods.* 2003;283(1-2):17-25. http://dx.doi.org/10.1016/j.jim.2003.07.003.

Zhang M.Y., Shu Y., Rudolph D., Prabakaran P., Labrijn A.F., Zwick M.B., Lal R.B., Dimitrov D.S. Improved breadth and potency of an HIV-1-neutralizing human single-chain antibody by random mutagenesis and sequential antigen panning. *J. Mol. Biol.* 2004a; 335(1):209-219. DOI 10.1016/j.jmb.2003.09.055.

Zhang M.Y., Xiao X., Sidorov I.A., Choudhry V., Cham F., Zhang P.F., Bouma P., Zwick M., Choudhary A., Montefiori D.C., Broder C.C., Burton D.R., Quinnan G.V., Jr., Dimitrov D.S. Identification and characterization of a new cross-reactive human immunodeficiency virus type 1-neutralizing human monoclonal antibody. *J. Virol.* 2004b;78(17):9233-9242. DOI 10.1128/JVI.78.17.9233-9242.2004.

Zhang M.Y., Yuan T., Li J., Rosa Borges A., Watkins J.D., Guenaga J., Yang Z., Wang Y., Wilson R., Li Y., Polonis V.R., Pincus S.H., Ru-precht R.M., Dimitrov D.S. Identification and characterization of a broadly cross-reactive HIV-1 human monoclonal antibody that binds to both gp120 and gp41. *PLoS One.* 2012;7(9):e44241. DOI 10.1371/journal.pone.0044241.

Zhao A., Tohidkia M.R., Siegel D.L., Coukos G., Omidi Y. Phage antibody display libraries: a powerful antibody discovery platform for immunotherapy. *Crit. Rev. Biotechnol.* 2016;36(2):276-289. DOI 10.3109/07388551.2014.958978.

Zwick M.B., Bonnycastle L.L., Menendez A., Irving M.B., Barbas C.F., 3rd, Parren P.W., Burton D.R., Scott J.K. Identification and characterization of a peptide that specifically binds the human broadly neutralizing anti-human immunodeficiency virus type 1 antibody b12. *J. Virol.* 2001;75(14):6692-6699. DOI 10.1128/JVI.75.14.6692-6699.2001.

ORCID ID

A.N. Chikaev orcid.org/0000-0001-5423-3457
A.P. Rudometov orcid.org/0000-0003-2808-4309
Iu.A. Merkuleva orcid.org/0000-0002-6974-0686
L.I. Karpenko orcid.org/0000-0003-4365-8809

Chromatin remodeling in oligodendrogenesis

E.V. Antontseva ✉, N.P. Bondar

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

Abstract. Oligodendrocytes are one type of glial cells responsible for myelination and providing trophic support for axons in the central nervous system of vertebrates. Thanks to myelin, the speed of electrical-signal conduction increases several hundred-fold because myelin serves as a kind of electrical insulator of nerve fibers and allows for quick saltatory conduction of action potentials through Ranvier nodes, which are devoid of myelin. Given that different parts of the central nervous system are myelinated at different stages of development and most regions contain both myelinated and unmyelinated axons, it is obvious that very precise mechanisms must exist to control the myelination of individual axons. As they go through the stages of specification and differentiation – from multipotent neuronal cells in the ventricular zone of the neural tube to mature myelinating oligodendrocytes as well as during migration along blood vessels to their destination – cells undergo dramatic changes in the pattern of gene expression. These changes require precisely spatially and temporally coordinated interactions of various transcription factors and epigenetic events that determine the regulatory landscape of chromatin. Chromatin remodeling substantially affects transcriptional activity of genes. The main component of chromatin is the nucleosome, which, in addition to the structural function, performs a regulatory one and serves as a general repressor of genes. Changes in the type, position, and local density of nucleosomes require the action of specialized ATP-dependent chromatin-remodeling complexes, which use the energy of ATP hydrolysis for their activity. Mutations in the genes encoding proteins of the remodeling complexes are often accompanied by serious disorders at early stages of embryogenesis and are frequently identified in various cancers. According to the domain arrangement of the ATP-hydrolyzing subunit, most of the identified ATP-dependent chromatin-remodeling complexes are classified into four subfamilies: SWI/SNF, CHD, INO80/SWR, and ISWI. In this review, we discuss the roles of these subunits of the different subfamilies at different stages of oligodendrogenesis.
Key words: oligodendrocyte; myelination; epigenetic regulation; gene expression.

Ремоделирование хроматина в олигодендрогенезе

Е.В. Антонцева ✉, Н.П. Бондарь

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

Аннотация. Олигодендроциты – это один из типов глиальных клеток, ответственных за миелинизацию и обеспечивающих трофическую поддержку аксонов в центральной нервной системе позвоночных. Благодаря миелину скорость проведения электрического сигнала увеличивается в сотни раз, так как он служит своего рода электроизолятором нервных волокон и позволяет осуществлять скачкообразную передачу потенциала действия через лишенные миелина перехваты Ранвье. Поскольку разные части ЦНС миелинизируются на различных стадиях развития и большинство регионов содержит как миелинизированные, так и немиелинизированные аксоны, очевидно, что должны существовать очень точные механизмы для контроля миелинизации отдельных аксонов. При прохождении через стадии спецификации и дифференцировки – от мультипотентных нейрональных клеток вентрикулярной зоны нервной трубки до зрелых миелинизирующих олигодендроцитов, а также во время миграции вдоль кровеносных сосудов к пункту назначения, клетки претерпевают кардинальные изменения в паттерне экспрессии генов. Эти изменения требуют тщательно скоординированного в пространстве и времени взаимодействия различных транскрипционных факторов (ТФ) и эпигенетических событий, определяющих регуляторный ландшафт хроматина. Ремоделирование хроматина существенно влияет на транскрипционную активность генов. Основной компонент хроматина – это нуклеосома, которая, помимо структурной, выполняет регуляторную функцию и служит общим репрессором генов. Для изменения типа, положения и локальной плотности нуклеосом необходимо действие специализированных АТФ-зависимых комплексов ремоделирования хроматина, которые используют для своей работы энергию гидролиза АТФ. Мутации в генах, кодирующих белки комплексов ремоделирования, часто сопровождаются серьезными нарушениями на ранних стадиях эмбриогенеза и с высокой частотой

идентифицируются при различных раковых заболеваниях. Большинство идентифицированных АТФ-зависимых комплексов ремоделирования хроматина классифицируется на четыре подсемейства: SWI/SNF, CHD, INO80/SWR и ISWI, согласно доменной организации их АТФ-гидролизующей субъединицы. В настоящем обзоре мы подробно остановимся на роли этих субъединиц разных подсемейств на различных этапах олигодендрогенеза.

Ключевые слова: олигодендроцит; миелинизация; эпигенитическая регуляция; экспрессия генов.

Introduction

Until recently, in the research on the workings of the brain, the central role in the functioning of the central nervous system has been assigned to neurons, and various pathological conditions have been regarded as a result of impaired functioning of neurons. Glial cells have been assigned the function of a filler for the space among neurons, as reflected in the name: glia is translated as "glue". Nonetheless, at present, there is no doubt about their necessity for the maintenance of axonal functions, for synaptic plasticity, and for the formation of neural networks.

Oligodendrocytes (OLs) are a type of glial cells responsible for myelination of axons in the central nervous system of vertebrates. As a result of an extremely specialized process of intercellular interaction, each OL forms several processes, each of which repeatedly wraps a part of an axon like insulating tape. One OL can simultaneously myelinate up to 50 axonal segments (Nave, Werner, 2014). The cytoplasm is almost absent in these processes, and, accordingly, the myelin sheath is actually many layers of the cell membrane with dielectric properties; this arrangement prevents the scattering of an electrical signal traveling along the axon. In addition, between the layered wrappings (sheaths) of myelin, there are ~1 micron-wide gaps, called Ranvier nodes, which enable saltatory transmission of a nerve impulse.

Thus, due to the myelin sheath, the speed of the electrical signal along the axon increases manyfold. It has been shown that the most active axons in the brain receive heavier myelin insulation, which allows them to work even more efficiently. It should be noted that myelination of the central nervous system proceeds over a long period and is the final stage in the development of the nervous system. According to magnetic resonance imaging data from humans, the bulk of myelination of most brain structures occurs in adolescence and reaches 90 % by the age of 20–25, whereas the late-maturing prefrontal cortex is the last to be myelinated (Lebel et al., 2008). In mice, most of myelination takes place in the first 4–7 weeks after birth. It is worth mentioning that myelin formation continues throughout the lifespan in the form of either remyelination of damaged nerve fibers or myelination of previously unmyelinated ones (Zhu et al., 2011; Bartzokis et al., 2012; Young et al., 2013).

Given that different parts of the central nervous system are myelinated at different stages of development and most of central nervous system regions contain both myelinated and unmyelinated axons, it is clear that very precise mechanisms must exist to control the myelination of individual axons. In addition, it is becoming increasingly obvious that there is some level of plasticity (remodeling) based on life experience in the myelination process (Mitew et al., 2014).

Specification, migration, and differentiation of OLs

Oligodendrocytes come into being as a result of gradual differentiation of oligodendrocyte precursor cells (OPCs) that arise during specification of multipotent neuronal progenitor cells located in the ventral zone of the neural tube. In the specification of OPCs, one of the key factors that determines this localization is the sonic hedgehog (SHH) protein secreted by the cells of the notochord and floor plate. The dorsal-ventral SHH concentration gradient promotes the formation of OPCs mainly from the ventral neuroepithelium by inducing the expression of a number of transcription factors (NKX2.2, PAX6, SOX10, OLIG1, and OLIG2). Oligodendrocyte precursor cells are characterized by the expression of proteoglycan NG2 and platelet growth factor receptor alpha (PDGFRα). In the absence of SHH in the anterior or spinal cord, OPCs of ventral origin do not form (Orentas, Miller, 1996; Pringle et al., 1996; Tekki-Kessaris et al., 2001). After specification, OPCs proliferate and migrate along blood vessels, ensuring an even distribution of white matter in the central nervous system (Dejana, Betsholtz, 2016; Tsai et al., 2016). After the migration of OPCs, some of them remain in the precursor state, while others – through the stage of premyelinating OLs – differentiate within at least 2–3 days into mature myelinating cells, which interact with axons and give rise to myelin internodes (Zhu et al., 2011; Mitew et al., 2014).

To avoid a shortage of OLs during axon myelination, an excess of these cells is generated, and unnecessary cells are subsequently eliminated by apoptosis. One of the mechanisms that determine the final OL number is competition for a limited amount of specific growth and survival factors, such as platelet growth factor (PDGF)-A, fibroblast growth factor (FGF)-2, insulinlike growth factor (IGF)-1, neurotrophin (NT)-3, and ciliary neurotrophic factor (CNTF) (Barres, Raff, 1994; Miller, 2002).

As expected, the differentiation into OLs should be consistent with a neuronal development program, and there are neuronal signals that control the conversion of OPCs into myelin-forming OLs. For instance, in the peripheral nervous system, on the axonal membrane, there is a protein called neuregulin 1, which controls myelination by Schwann cells; however, a similar regulatory protein triggering the myelination by OLs in the central nervous system has not yet been identified. Moreover, there is evidence that there are inhibitory neuronal signals that keep OPCs in a "suppressed" state (Emery, 2010). These inhibitory signals (for example, Jagged, PSA-NCAM, and LINGO-1) coming from axons in turn activate various regulators of transcription, such as SOX5/6, HES5, and ID2/4, which actively prevent OPCs from entering the stage of terminal differentiation (Piaton et al., 2010; Taveggia et al., 2010).

Е.В. Антонцева
Н.П. Бондарь

Ремоделирование хроматина
в олигодендрогенезе

2021
25·5

More and more data indicate that different areas of the central nervous system correspond to different OPC populations that are controlled by some local signaling mechanisms. Thus, in different areas of the brain and spinal cord, OPCs are under the influence of different signaling molecules. In particular, dissimilarities in the temporal expression of these factors and signals in a developing central nervous system ensure that myelination begins earlier in the spinal cord and later in cortical regions.

When going through the stages of specification, migration, proliferation, and differentiation, OLs undergo dramatic changes in the pattern of gene expression, which require a precisely (temporally and spatially) coordinated interaction of various transcription factors and epigenetic events that determine the regulatory landscape of chromatin (Copray et al., 2009). Chromatin remodeling considerably affects transcriptional activity of genes and occurs mainly due to nucleosome repositioning (under the action of ATP-dependent chromatin-remodeling complexes), chemical modifications of histones, DNA methylation, and interactions with noncoding RNAs (Gregath, Lu, 2018; Koreman et al., 2018).

In the present review, we will examine in detail the functions of ATP-dependent chromatin-remodeling complexes at various stages of oligodendrogenesis.

ATP-dependent chromatin-remodeling complexes
The regulation of gene expression is possible for the most part due to changes in the structure of chromatin: its remodeling. The main component of chromatin is the nucleosome, which consists of eight protein subunits (histones), wrapped by a 146 bp DNA region for each nucleosome. In addition to the structural function (there is growing evidence that it is less important), the nucleosome performs a regulatory function and serves as a general repressor of genes. It interferes with virtually all DNA-related processes, including transcription, replication, and DNA repair (Kornberg, Lorch, 2020). Changes in the type, position, and local density of nucleosomes require the action of specialized ATP-dependent chromatin-remodeling complexes (so-called remodelers), which use the energy of ATP hydrolysis for their functioning.

All ATP-dependent chromatin-remodeling complexes contain an ATP-hydrolyzing subunit belonging to the SNF2 DNA helicase/translocase family and one or more subunits that are bound to it (Hota, Bruneau, 2016). It should be noted that these ATPases lack helicase activity for strand unwinding and therefore act only as DNA translocases (Dürr et al., 2005) that move DNA along the surface of the nucleosome and disrupt DNA–histone interactions. ATPases of all chromatin remodelers contain a highly conserved core ATPase domain, which is sufficient for the catalytic activity of the complex; however, ATPase activity regulation is carried out by the domains flanking the ATPase domain and/or by proteins physically associated with it (Clapier et al., 2017). It must be pointed out that tissue specificity of the subunits of a chromatin-remodeling complex gives it unique properties, which determine its landing on tissue-specific regulatory loci of the genome (Hota, Bruneau, 2016).

According to sequence homology of the central ATP-hydrolyzing subunit, most of the ATP-dependent chromatin-

remodeling complexes identified so far are classified into four subfamilies: SWI/SNF (switch/sucrose nonfermentable), CHD (chromodomain helicase DNA-binding), INO80/SWR (Inositol-requiring 80/SWi2/snf2-related 1), and ISWI subfamily (imitation switch). Aside from differences in the number of subunits in the complex, these subfamilies are distinguished by exclusive domains adjacent to the core ATPase region. For instance, SWI/SNF is characterized by the presence of a bromodomain, CHD – by tandem chromodomains, ISWI remodelers contain HSS helicase domains, and INO80/SWR members have an HSA domain (SANT helicase) (Bartholomew, 2014; Clapier et al., 2017).

The structural differences between the subfamilies determine their functionality. For example, members of the SWI/SNF subfamily are mainly responsible for the removal and repositioning of nucleosomes, thereby providing access to DNA for transcription factors (in order to regulate the level of gene expression) and for DNA repair and recombination factors. The assembly and positioning of nucleosomes is mainly conducted by remodelers of subfamilies ISWI and CHD. The editing of nucleosomes, namely the replacement of canonical histones with specialized types and vice versa, is carried out mainly by members of the INO80 subfamily. Changing the composition of nucleosomes allows for the creation of specialized regions of chromatin in a replication-independent manner (Venkatesh, Workman, 2015; Clapier et al., 2017). For example, histone H2A.Z is a part of the nucleosomes located in promoter regions of most genes; the proximity of such a nucleosome to a transcription start site directly correlates with the expression level of the gene in question (Bargaje et al., 2012), H2A.X is required for the repair of double-strand breaks (Elsesser et al., 2019).

Mutations in the genes encoding subunits of ATP-dependent chromatin-remodeling complexes are often accompanied by severe abnormalities in the early stages of embryogenesis and are frequently identified in various cancers (Ho, Crabtree, 2010; Wilson, Roberts, 2011).

Remodelers' actions during oligodendrogenesis

Subfamily SWI/SNF
Mammalian members of SWI/SNF are large complexes of ~1.5 MDa and consist of at least 15 different subunits (Hota, Bruneau, 2016). They are characterized by the presence of enzyme Brm (encoded by the *Brahma* gene, also known as *Smarca2*) or Brg1 (*Brahma-related gene 1*, aka *Smarca4*) as core translocases. These remodelers are mainly involved in the regulation of the cell cycle and differentiation of several cell types (Matsumoto et al., 2016).

A comparative analysis of genome-wide chromatin immunoprecipitation-high-throughput sequencing (ChIP-seq) data on the localization of RNA polymerase RNAPII on the genomic DNA of OLs at different stages of differentiation (rat primary OPCs in culture, immature OLs, and mature OLs) has revealed significant enrichment in genes coding for proteins of the SWI/SNF remodeling complex during the transition from OPCs to immature OLs (Yu et al., 2013). In particular, in response to differentiation signals, a significant increase in the RNAPII amount was registered on exons of the *Brg1*

gene (encoding the core ATPase of the SWI/SNF complex) but was not observed on the gene of its only homolog, *Brm* (Yu et al., 2013).

During early embryonic development (embryonic day 12; E12), *Brg1* transcripts are virtually undetectable in proliferating neuronal stem cells (NSCs) of the ventricular zone (VZ) in the mouse forebrain, in contrast to the already differentiated postmitotic cells of the mantle zone (Randazzo et al., 1994). This pattern of *Brg1* expression changes after E13, in such a way that most VZ cells become Brg1-immunoreactive (Matsumoto et al., 2006), in good agreement with the data on the beginning of the first wave of OPC specification in mice on E12.5 during this period (Cai et al., 2005). A Cre-mediated knockout involving a conditional deletion of Brg1 under the control of nestin (an NSC marker) in mice leads to a decrease (in the ventricular zone, E13.5) in the number of proliferating NSCs capable of subsequently undergoing gliogenesis. It is assumed that this phenomenon can be mediated by stem cell differentiation into postmitotic cells (neurons) and/or by apoptosis. Those authors believe that Brg1 keeps NSCs in an undifferentiated state until they respond to some gliogenic signals during development in mammals (Matsumoto et al., 2006).

An increase in the transcription of *Brg1* has been noted exclusively during the transition from precursors to immature OLs and is not observed during differentiation of other cell types, suggesting that this is a unique event initiating oligodendrocytic differentiation. Western blotting and immunohistochemical results also indicate that Brg1 expression is mainly limited to differentiating oligodendrocytic cells (Yu et al., 2013). By contrast, in another study, according to immunohistochemical findings, Brg1 expression was detectable at all stages of OL development, and there was no obvious difference in staining intensity between OPCs and maturing OLs (Bischof et al., 2015).

In mice with a conditional knockout of *Brg1* (with Cre under the control of Olig1, the expression of which starts with the formation of OPCs), a myelination-deficient phenotype manifests itself, although the number of OPCs and their proliferation rate are comparable to those in control mice. It turns out that these mice feature increased expression levels of differentiation inhibitors (ID2/4, Nfia/b, Sox5, and β-catenin) and downregulation of genes associated with both the synthesis of lipids and myelin sheath proteins and with differentiation regulation (Gm98/MRF and Sox10) (Yu et al., 2013). In the aforementioned work of M. Bischof and coworkers (2015), it was also found that early deletion of *Brg1* (with Cre under the control of Brn4 expressed in NSCs) throughout the ventricular zone of the spinal cord also prevents Sox10 expression in OLPs until the end of embryogenesis.

ChIP-seq with antibodies to Brg1 revealed an order of magnitude more peaks in immature OLs than in OPCs, while almost all the peaks in OPCs overlapped with the peaks of immature OLs but had significantly lower intensity. Gene analysis showed that Brg1 target genes are mainly associated with myelination, oligodendrocytic differentiation, and cell cycle arrest (e. g., *Cnp*, *Cldn11*, *Klf9*, and *Zfp191*). Those authors hypothesized that Brg1 mainly targets intergenic enhancer regions because ~80 % of peaks of Brg1 binding in immature OLs were located at a considerable distance from a transcription start site (Yu et al., 2013).

Transcription factor Olig2 is thought to play an important role in the regulation of Brg1 functions in oligodendrogenesis. It is reported that Olig2 not only regulates *Brg1* expression but also recruits the SWI/SNF complex with the Brg1 ATPase to OL-specific enhancers during the critical transition from OPCs to immature OLs; as a consequence, directed chromatin remodeling ensues, which is necessary for activation of the oligodendrocytic differentiation program (Yu et al., 2013).

Moreover, it has been revealed that at early stages of development (E14.5), Brg1 within the SWI/SNF complex interacts with the proximal promoter of the *Olig2* gene in cortical NSCs and suppresses its expression, whereas in cells of the ventricular zone, where oligodendrocytic differentiation is seen first, no suppressive effect of Brg1 on *Olig2* was found (Matsumoto et al., 2016).

Therefore, all these data indicate that Brg1, which is a part of the SWI/SNF complex, is required for both OPC specification and oligodendrocytic differentiation.

Subfamily CHD
ATPases of this subfamily of remodelers are represented by nine proteins, CHD1–9, which differ in their domain organization. The complexes formed by them can combine 1 to 10 subunits. Some of them shift or push nucleosomes to facilitate transcription, whereas others play a repressive part (Clapier, Cairns, 2009), and thus are important regulators of cell differentiation (Martin, 2010). It has been demonstrated that chromodomain ATPases Chd7 and Chd8 are important for the initiation of the processes of nerve-fiber myelination as well as remyelination in the pathogenesis of some disorders (He et al., 2016; Doi et al., 2017; Marie et al., 2018).

Chd7 expression is present in most of oligodendrocytic cells in a number of structures within a developing brain (postnatal day 14, PND14) and in the spinal cord. A substantial proportion of these cells are differentiated OLs with Chd7 overexpression, whereas in OPCs, Chd7 expression is reported to be lower (He et al., 2016). A conditional knockout of the *Chd7* gene using the *Olig1*–Cre construct reduces the number of OLs expressing MBP both at embryonic and early postnatal stages of development and decreases the expression of basic myelin proteins Mbp and Plp1 at the mRNA level. As a result, in the mutant animals, there is a decrease of white-matter volume in the brain and impaired myelination of nerve fibers (a lower g-ratio) in comparison with control animals (PND14). In the cerebral cortex, the *Chd7* deletion also leads to a decrease in the number of mature OLs at the early postnatal stages of development but does not have a significant impact on the number of OPCs and their proliferation. Nevertheless, with age, mutant mice showed a gradually rising number of mature OLs, and by PND60, the degree of myelination in the spinal cord approached normal. Accordingly, Chd7 is necessary for the initiation of oligodendrocytic differentiation, and Chd7 loss causes a noticeable delay in the myelination of nerve fibers. A comparative analysis of gene expression profiles in the spinal cord between control mice and mice with the conditional

Е.В. Антонцева
Н.П. Бондарь

Ремоделирование хроматина
в олигодендрогенезе

2021
25•5

Chd7 knockout (PND8) confirmed the importance of Chd7 for the regulation of genes responsible for the differentiation into OLs and for myelination; furthermore, in most cases (~84 %), it functions as a transcriptional activator (He et al., 2016).

Normally, in adult mice, Chd7 expression in the white matter of the spinal cord is practically undetectable; however, lysolecithin-induced demyelination causes local re-expression of Chd7 against the background of myelin regeneration via OPC recruitment. In the case of a *Chd7* deletion in the affected tissue area, there was a significant reduction not only in the OL number but also in the expression levels of *Mbp* and *Plp1* as compared to the control, and morphometric characteristics of myelinated fibers deteriorated. For these reasons, those authors believe that Chd7 is crucial for remyelination in case of white-matter damage (He et al., 2016). Chd7 has been reported to be a key regulator of OPC activation and proliferation after spinal cord injury (Doi et al., 2017).

In the expression regulation of the *Chd7* gene itself, an important role is played by the Brg1 remodeler, the conditional knockout of which significantly suppresses *Chd7* expression in mutant mice. Furthermore, within the *Chd7* gene, multiple sites of cooperative binding of factors Brg1 and Olig2 have been identified, and these sites are functionally significant at the stage of OPCs and immature OLs (He et al., 2016).

A genome-wide search for Chd7-binding sites by ChIP-seq in differentiating OLs has shown that they are predominantly located in the region +5 kbp relative to a transcription start site of target genes and also serve as binding sites for transcription factor Olig2, which is known to regulate enhancers that are functionally important for oligodendrogenesis, in particular, by recruiting Brg1 to them (Yu et al., 2013; He et al., 2016).

A comparative analysis of distributions of ChIP-seq peaks for the remodelers Chd7 and Brg1 revealed that the peaks do not overlap in ~76 % of cases, and therefore Chd7 has unique molecular functions that control OL maturation (He et al., 2016).

CHD7 mutations cause human CHARGE syndrome, which is characterized by multiple pathologies including craniofacial anomalies, neurological dysfunction, and growth retardation. Most CHARGE patients show some degree of intellectual disability, and many have structural aberrations of the corpus callosum and cerebellar vermis (Martin, 2010).

Subfamily INO80/SWR

The remodeling complexes belonging to this subfamily can contain more than 10 subunits and perform a variety of functions, including facilitation of transcription activation and DNA repair (Clapier, Cairns, 2009). In one study (Elsesser et al., 2019), it was demonstrated that a member of this family, chromatin-remodeling complex TIP60/EP400, is an important component of the oligodendrocytic differentiation program. It turned out that a Cre-mediated knockout of Ep400 (core ATPase) at various oligodendrogenesis stages in transgenic mice does change the OPC number but is the reason for a sharp drop in the number of OLs that started terminal differentiation and initiated myelin gene expression.

Analysis of gene expression in the OLs from mice with the *Ep400* knockout confirmed downregulation of not only termi-

nal differentiation and myelination genes, such as *Plp1*, *Mbp*, *Mog*, and *Nfasc*, but also key components of the transcriptional network, e. g., *Sox10*, *Nkx2.2*, *Olig2*, *Olig1*, *Myrf*, and *Fyn*, the products of which regulate the processes of oligodendrocytic differentiation and myelination. In particular, in immature OLs, EP400 specifically binds to the promoter and enhancer ECR9 in the first intron of the *Myrf* gene and i) catalyzes the replacement of histone H2A by the specialized H2A.Z histone in many nucleosomes as well as ii) recruits the Sox10 transcription factor, with which EP400 physically interacts to activate *Myrf* expression during oligodendrocytic differentiation. After the differentiation, EP400 induces changes in the pattern of H2A.Z localization on the promoter.

Additionally, in cultured primary OLs, it was found that the absence of EP400 impairs repair processes, as evidenced by an increase in the amount of histone γH2A.X, which is a marker of a double-stranded DNA break and emerges as a result of phosphorylation of the specialized H2A.X histone. Because such breaks are a signal for apoptosis, it is possible to define EP400 as a factor that promotes OL survival and protects DNA from damage or helps with its repair (Elsesser et al., 2019). Given that upon completion of the differentiation into OLs and at initial stages of myelination, there is an increase in heterochromatinization and nuclear condensation (Mori, Leblond, 1970), it can be theorized that such critical changes in the structure of chromatin make it especially vulnerable during this period of oligodendrogenesis and require more attention from a "surveillance" system (Elsesser et al., 2019).

Thus, EP400 is not required for OPC specification or for early development of clones in embryogenesis but is necessary during terminal oligodendrocytic differentiation and the active phase of myelination.

Conclusion

Each stage of oligodendrogenesis is accompanied by a change in the expression of a large number of genes. The activation of some genes and the repression of others are implemented by precisely spatially and temporally coordinated interactions of various transcription factors with the promoters and enhancers of these genes. Such fine regulation is possible due to the coordinated fine-tuned work of transcription factors and ATP-dependent chromatin-remodeling complexes, which determine the regulatory landscape of chromatin and contribute to its epigenetic modifications (Copray et al., 2009). In addition to many common properties among all chromatin remodelers, they have distinct features that explain their functional specialization in the cell (Hota, Bruneau, 2016), in particular, at different stages of differentiation in oligodendrogenesis. It has been shown that OP specification requires SWI/SNF complexes, which keep NSCs in an undifferentiated state until they respond to certain gliogenic signals. Besides, members of this subfamily are curators of transcriptional activity of genes at all stages of oligodendrocytic differentiation. By contrast, CHD7 is required for OPC proliferation, for the onset of oligodendrocytic differentiation, and for the activation of genes responsible for myelination and remyelination, whereas EP400 is necessary only for terminal oligodendrocytic differentiation and the active phase of myelination (Matsumoto

et al., 2006, 2016; Yu et al., 2013; He et al., 2016; Doi et al., 2017; Elsesser et al., 2019).

The roles of ISWI subfamily members in oligodendrogenesis have not yet been investigated; however, given their functional feature – the assembly and positioning of nucleosomes (Hota, Bruneau, 2016; Clapier et al., 2017) – it is possible that they are important for heterochromatinization processes, which intensify with the gradual transition from OPCs to mature myelinating OLs, and for nuclear condensation (Mori, Leblond, 1970).

References

Bargaje R., Alam M.P., Patowary A., Sarkar M., Ali T., Gupta S., Garg M., Singh M., Purkanti R., Scaria V., Sivasubbu S., Brahmachari V., Pillai B. Proximity of H2A.Z containing nucleosome to the transcription start site influences gene expression levels in the mammalian liver and brain. *Nucleic Acids Res*. 2012;40:8965-8978. DOI 10.1093/nar/gks665.

Barres B.A., Raff M.C. Control of oligodendrocyte number in the developing rat optic nerve. *Neuron*. 1994;12:935-942. DOI 10.1016/0896-6273(94)90305-0.

Bartholomew B. Regulating the chromatin landscape: structural and mechanistic perspectives. *Annu. Rev. Biochem*. 2014;83:671-696. DOI 10.1146/annurev-biochem-051810-093157.

Bartzokis G., Lu P.H., Heydari P., Couvrette A., Lee G.J., Kalashyan G., Freeman F., Grinstead J.W., Villablanca P., Finn J.P., Mintz J., Alger J.R., Altshuler L.L. Multimodal magnetic resonance imaging assessment of white matter aging trajectories over the lifespan of healthy individuals. *Biol. Psychiatry*. 2012;72:1026-1034. DOI 10.1016/j.biopsych.2012.07.010.

Bischof M., Weider M., Kuspert M., Nave K.-A., Wegner M. Brg1-dependent chromatin remodelling is not essentially required during oligodendroglial differentiation. *J. Neurosci*. 2015;35:21-35. DOI 10.1523/JNEUROSCI.1468-14.2015.

Cai J., Qi Y., Hu X., Tan M., Liu Z., Zhang J., Li Q., Sander M., Qiu M. Generation of oligodendrocyte precursor cells from mouse dorsal spinal cord independent of Nkx6 regulation and Shh signaling. *Neuron*. 2005;45:41-53. DOI 10.1016/j.neuron.2004.12.028.

Clapier C.R., Cairns B.R. The biology of chromatin remodeling complexes. *Annu. Rev. Biochem*. 2009;78:273-304. DOI 10.1146/annurev.biochem.77.062706.153223.

Clapier C.R., Iwasa J., Cairns B.R., Peterson C.L. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell Biol*. 2017;18:407-422. DOI 10.1038/nrm.2017.26.

Copray S., Huynh J.L., Sher F., Casaccia-Bonnefil P., Boddeke E. Epigenetic mechanisms facilitating oligodendrocyte development, maturation, and aging. *Glia*. 2009;57:1579-1587. DOI 10.1002/glia.20881.

Dejana E., Betsholtz C. Oligodendrocytes follow blood vessel trails in the brain. *Science*. 2016;351:341-342. DOI 10.1126/science.aaf1139.

Doi T., Ogata T., Yamauchi J., Sawada Y., Tanaka S., Nagao M. Chd7 collaborates with Sox2 to regulate activation of oligodendrocyte precursor cells after spinal cord injury. *J. Neurosci*. 2017;37:10290-10309. DOI 10.1523/JNEUROSCI.1109-17.2017.

Dürr H., Körner C., Müller M., Hickmann V., Hopfner K.-P. X-ray structures of the *Sulfolobus solfataricus* SWI2/SNF2 ATPase core and its complex with DNA. *Cell*. 2005;121:363-373. DOI 10.1016/j.cell.2005.03.026.

Elsesser O., Fröb F., Küspert M., Tamm E.R., Fujii T., Fukunaga R., Wegner M. Chromatin remodeler Ep400 ensures oligodendrocyte survival and is required for myelination in the vertebrate central nervous system. *Nucleic Acids Res*. 2019;47:6208-6224. DOI 10.1093/nar/gkz376.

Emery B. Regulation of oligodendrocyte differentiation and myelination. *Science*. 2010;330:779-782. DOI 10.1126/science.1190927.

Gregath A., Lu Q.R. Epigenetic modifications – insight into oligodendrocyte lineage progression, regeneration, and disease. *FEBS Lett*. 2018;592:1063-1078. DOI 10.1002/1873-3468.12999.

He D., Marie C., Zhao C., Kim B., Wang J., Deng Y., Clavairoly A., Frah M., Wang H., He X., Hmidan H., Jones B.V., Witte D., Zalc B., Zhou X., Choo D.I., Martin D.M., Parras C., Lu Q.R. Chd7 cooperates with Sox10 and regulates the onset of CNS myelination and remyelination. *Nat. Neurosci*. 2016;19:678-689. DOI 10.1038/nn.4258.

Ho L., Crabtree G.R. Chromatin remodelling during development. *Nature*. 2010;463:474-484. DOI 10.1038/nature08911.

Hota S.K., Bruneau B.G. ATP-dependent chromatin remodeling during mammalian development. *Development*. 2016;143:2882-2897. DOI 10.1242/dev.128892.

Koreman E., Sun X., Lu Q.R. Chromatin remodeling and epigenetic regulation of oligodendrocyte myelination and myelin repair. *Mol. Cell Neurosci*. 2018;87:18-26. DOI 10.1016/j.mcn.2017.11.010.

Kornberg R.D., Lorch Y. Primary role of the nucleosome. *Mol. Cell*. 2020;79:371-375. DOI 10.1016/j.molcel.2020.07.020.

Lebel C., Walker L., Leemans A., Phillips L., Beaulieu C. Microstructural maturation of the human brain from childhood to adulthood. *Neuroimage*. 2008;40:1044-1055. DOI 10.1016/j.neuroimage.2007.12.053.

Marie C., Clavairoly A., Frah M., Hmidan H., Yan J., Zhao C., Van Steenwinckel J., Daveau R., Zalc B., Hassan B., Thomas J.-L., Gressens P., Ravassard P., Moszer I., Martin D.M., Lu Q.R., Parras C. Oligodendrocyte precursor survival and differentiation requires chromatin remodeling by Chd7 and Chd8. *Proc. Natl. Acad. Sci*. 2018;115:E8246-E8255. DOI 10.1073/pnas.1802620115.

Martin D.M. Chromatin remodeling in development and disease: focus on CHD7. *PLoS Genet*. 2010;6:e1001010. DOI 10.1371/journal.pgen.1001010.

Matsumoto S., Banine F., Feistel K., Foster S., Xing R., Struve J., Sherman L.S. Brg1 directly regulates *Olig2* transcription and is required for oligodendrocyte progenitor cell specification. *Dev. Biol*. 2016;413:173-187. DOI 10.1016/j.ydbio.2016.04.003.

Matsumoto S., Banine F., Struve J., Xing R., Adams C., Liu Y., Metzger D., Chambon P., Rao M.S., Sherman L.S. Brg1 is required for murine neural stem cell maintenance and gliogenesis. *Dev. Biol*. 2006;289:372-383. DOI 10.1016/j.ydbio.2005.10.044.

Miller R.H. Regulation of oligodendrocyte development in the vertebrate CNS. *Prog. Neurobiol*. 2002;67:451-467. DOI 10.1016/S0301-0082(02)00058-8.

Mitew S., Hay C.M., Peckham H., Xiao J., Koenning M., Emery B. Mechanisms regulating the development of oligodendrocytes and central nervous system myelin. *Neuroscience*. 2014;276:29-47. DOI 10.1016/j.neuroscience.2013.11.029.

Mori S., Leblond C.P. Electron microscopic identification of three classes of oligodendrocytes and a preliminary study of their proliferative activity in the corpus callosum of young rats. *J. Comp. Neurol*. 1970;139:1-29. DOI 10.1002/cne.901390102.

Nave K.-A., Werner H.B. Myelination of the nervous system: mechanisms and functions. *Annu. Rev. Cell Dev. Biol*. 2014;30:503-533. DOI 10.1146/annurev-cellbio-100913-013101.

Orentas D.M., Miller R.H. The origin of spinal cord oligodendrocytes is dependent on local influences from the notochord. *Dev. Biol*. 1996;177:43-53. DOI 10.1006/dbio.1996.0143.

Piaton G., Gould R.M., Lubetzki C. Axon-oligodendrocyte interactions during developmental myelination, demyelination and repair. *J. Neurochem*. 2010;114(5):1243-1260. DOI 10.1111/j.1471-4159.2010.06831.x.

Pringle N.P., Yu W.-P., Guthrie S., Roelink H., Lumsden A., Peterson A.C., Richardson W.D. Determination of neuroepithelial cell fate: induction of the oligodendrocyte lineage by ventral midline

Е.В. Антонцева
Н.П. Бондарь

Ремоделирование хроматина
в олигодендрогенезе

2021
25•5

cells and sonic hedgehog. *Dev. Biol.* 1996;177:30-42. DOI 10.1006/dbio.1996.0142.

Randazzo F.M., Khavari P., Crabtree G., Tamkun J., Rossant J. *brg1:* a putative murine homologue of the *Drosophila brahma* gene, a homeotic gene regulator. *Dev. Biol*. 1994;161:229-242. DOI 10.1006/dbio.1994.1023.

Taveggia C., Feltri M.L., Wrabetz L. Signals to promote myelin formation and repair. *Nat. Rev. Neurol*. 2010;6:276-287. DOI 10.1038/nrneurol.2010.37.

Tekki-Kessaris N., Woodruff R., Hall A.C., Gaffield W., Kimura S., Stiles C.D., Rowitch D.H., Richardson W.D. Hedgehog-dependent oligodendrocyte lineage specification in the telencephalon. *Development*. 2001;128:2545-2554.

Tsai H.-H., Niu J., Munji R., Davalos D., Chang J., Zhang H., Tien A.-C., Kuo C.J., Chan J.R., Daneman R., Fancy S.P.J. Oligodendrocyte precursors migrate along vasculature in the developing nervous system. *Science*. 2016;351:379-384. DOI 10.1126/science.aad3839.

Venkatesh S., Workman J.L. Histone exchange, chromatin structure and the regulation of transcription. *Nat. Rev. Mol. Cell Biol*. 2015;16: 178-189. DOI 10.1038/nrm3941.

Wilson B.G., Roberts C.W.M. SWI/SNF nucleosome remodellers and cancer. *Nat. Rev. Cancer*. 2011;11:481-492. DOI 10.1038/nrc3068.

Young K.M., Psachoulia K., Tripathi R.B., Dunn S.-J., Cossell L., Attwell D., Tohyama K., Richardson W.D. Oligodendrocyte dynamics in the healthy adult CNS: evidence for myelin remodeling. *Neuron*. 2013;77:873-885. DOI 10.1016/j.neuron.2013.01.006.

Yu Y., Chen Y., Kim B., Wang H., Zhao C., He X., Liu L., Liu W., Wu L.M.N., Mao M., Chan J.R., Wu J., Lu Q.R. Olig2 targets chromatin remodelers to enhancers to initiate oligodendrocyte differentiation. *Cell*. 2013;152:248-261. DOI 10.1016/j.cell.2012.12.006.

Zhu X., Hill R.A., Dietrich D., Komitova M., Suzuki R., Nishiyama A. Age-dependent fate and lineage restriction of single NG2 cells. *Development*. 2011;138:745-753. DOI 10.1242/dev.047951.

ORCID ID
E.V. Antontseva orcid.org/0000-0002-4214-7153
N.P. Bondar orcid.org/0000-0002-5602-5149

Prioritization of biological processes based on the reconstruction and analysis of associative gene networks describing the response of plants to adverse environmental factors

P.S. Demenkov[1, 2], E.A. Oshchepkova[1], T.V. Ivanisenko[1], V.A. Ivanisenko[2, 3] ✉

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Novosibirsk State University, Novosibirsk, Russia
[3] Kurchatov Genomic Center of ICG SB RAS, Novosibirsk, Russia
✉ salix@bionet.nsc.ru

Abstract. Methods for prioritizing or ranking candidate genes according to their importance based on specific criteria via the analysis of gene networks are widely used in biomedicine to search for genes associated with diseases and to predict biomarkers, pharmacological targets and other clinically relevant molecules. These methods have also been used in other fields, particularly in crop production. This is largely due to the development of technologies to solve problems in marker-oriented and genomic selection, which requires knowledge of the molecular genetic mechanisms underlying the formation of agriculturally valuable traits. A new direction for the study of molecular genetic mechanisms is the prioritization of biological processes based on the analysis of associative gene networks. Associative gene networks are heterogeneous networks whose vertices can depict both molecular genetic objects (genes, proteins, metabolites, etc.) and the higher-level factors (biological processes, diseases, external environmental factors, etc.) related to regulatory, physicochemical or associative interactions. Using a previously developed method, biological processes involved in plant responses to increased cadmium content, saline stress and drought conditions were prioritized according to their degree of connection with the gene networks in the SOLANUM TUBEROSUM knowledge base. The prioritization results indicate that fundamental processes, such as gene expression, post-translational modifications, protein degradation, programmed cell death, photosynthesis, signal transmission and stress response play important roles in the common molecular genetic mechanisms for plant response to various adverse factors. On the other hand, a group of processes related to the development of seeds ("seeding development") was revealed to be drought specific, while processes associated with ion transport ("ion transport") were included in the list of responses specific to salt stress and processes associated with the metabolism of lipids were found to be involved specifically in the response to cadmium.

Key words: knowledge base SOLANUM TUBEROSUM; Gene Ontology; *Arabidopsis thaliana*; text mining methods; associative gene networks; centrality of vertices; network-based prioritization methods.

Приоритизация биологических процессов на основе реконструкции и анализа ассоциативных генных сетей, описывающих ответ растений на неблагоприятные факторы внешней среды

П.С. Деменков[1, 2], Е.А. Ощепкова[1], Т.В. Иванисенко[1], В.А. Иванисенко[2, 3] ✉

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
[3] Курчатовский геномный центр ИЦиГ СО РАН, Новосибирск, Россия
✉ salix@bionet.nsc.ru

Аннотация. Методы приоритизации или ранжирование кандидатных генов по их важности в соответствии с заданными критериями, основанными на анализе генных сетей, широко применяются в биомедицине для поиска ассоциаций генов с заболеваниями, предсказания биомаркеров, фармакологических мишеней и т.д. При этом наблюдается тенденция их использования и в других областях знаний, в частности в растениеводстве. В значительной степени это обусловлено развитием технологий для решения задач маркер-ориентированной и геномной селекции, требующих знаний о молекулярно-генетических механизмах, лежащих в основе формирования хозяйственно ценных признаков. Новым направлением для изучения молекулярно-генетических механизмов является приоритизация биологических процессов с применением анализа ассоциативных генных

П.С. Деменков, Е.А. Ощепкова
Т.В. Иванисенко, В.А. Иванисенко

Приоритизация биологических процессов, описывающих
ответ растений на неблагоприятные факторы внешней среды

2021
25 • 5

сетей. Ассоциативная генная сеть – это гетерогенная сеть, в качестве вершин которой наряду с молекулярно-генетическими объектами (гены, белки, метаболиты и т.д.) могут быть представлены сущности более высокого уровня (биологические процессы, заболевания, факторы внешней среды и т.д.), связанные между собой регуляторными, физико-химическими или ассоциативными взаимодействиями. С использованием разработанного нами ранее метода осуществлена приоритизация биологических процессов по степени их связи с генными сетями, представленными в базе знаний SOLANUM TUBEROSUM и описывающими ответ растений на повышенное содержание кадмия, солевой стресс и условия засухи. Результаты приоритизации свидетельствуют о том, что фундаментальные процессы, такие как экспрессия генов, посттрансляционная модификация, деградация белков, программируемая клеточная смерть, фотосинтез, передача сигналов, ответ на стресс, играют важную роль в общих молекулярно-генетических механизмах ответа растений на различные неблагоприятные факторы. С другой стороны, среди специфичных для устойчивости к засухе была выявлена группа процессов, связанных с развитием семян (seeding development). Процессы, связанные с ионным транспортом (ion transport), вошли в список специфичных для ответа на солевой стресс, а связанные с метаболизмом липидов (phospholipid degradation – деградация фосфолипидов) – для ответа на кадмий.

Ключевые слова: база знаний SOLANUM TUBEROSUM; Gene Ontology; *Arabidopsis thaliana*; методы text mining; ассоциативные генные сети; центральность вершин; сетевые методы приоритизации.

Introduction

The rapid development of high-performance experimental methods has significantly expanded the ability to generate large sets of genomic, transcriptomic and proteomic data in scientific research. This, in turn, has increased the relevance of bioinformatics methods that allow researchers to interpret omic data, both at the level of key genes and at the level of molecular genetic mechanisms. Among the widely used approaches in computational analysis of gene sets identified in experiments are prioritization methods (Raj, Sreeja, 2018), which rank the studied genes (or other objects, such as diseases) by characterising their proximity to a set from a given learning sample. Depending on the prioritization problem, the training set may consist of genes associated with diseases or phenotypic traits or sets of differentially expressed genes, for example. The higher the proximity in relative units, the greater the priority of the analysed object as a candidate possessing the same properties as objects in the training set. Such methods are used in biomedicine to detect candidate genes associated with diseases (Tranchevent et al., 2016), disease biomarkers (Jha et al., 2020), potential pharmacological targets (Cesur et al., 2020) and drug republic (Pushpakom et al., 2019). In animal husbandry and crop production, prioritization methods have been applied to analyse genomic data related to marker-oriented and genomic selection (Arruda et al., 2016; Crossa et al., 2017; Kochetov et al., 2017; Kolchanov et al., 2017; Cai et al., 2019; Voss-Fels et al., 2019; Sun et al., 2020), as well as Raspanic loci analysis (Bargsten et al., 2014; Schaefer et al., 2018; Lin et al., 2019).

A special place among the prioritization methods is occupied by approaches based on the analysis of genetic network graphs, including network protein interactions, metabolic networks, signal transmission networks, and networks of genes associated with disease. In such network prioritization methods, the proximity of the studied genes to the training sample is estimated using various topological characteristics of the genetic network graph. Methods for analysing the structure of genetic network graphs to solve prioritization tasks can be divided into three large groups (Shim et al., 2017; Raj, Sreeja, 2018): (1) methods based on identifying hubs using centrality indicators of vertices in the column (Cho et al., 2016); (2) methods based on network diffusion, including random wandering (Chen et al., 2011; Shim, Lee, 2015; Le, Pham, 2017; Lysenko et al., 2017); and (3) methods based on identifying functional modules (clusters or subnets) (Jia et al., 2011; Leung et al., 2014). All of these methods are aimed at identifying genes (or other entities) that are important for the phenotype or process being studied. The importance of information about the topology of networks in assessing the functional significance of genes was demonstrated by the example of the centrality of the vertices in a *Saccharomyces cerevisiae* protein–protein interactions network (Jeong et al., 2001). The authors showed that the deletion of vertices with a large number of connections in the network of protein–protein interactions was fatal more often than the deletion of other vertices.

Among the frequently used characteristics to determine the importance of the vertex in the network structure are centrality indicators, including Degree Centrality, which is the number of links this vertex has with other vertices in the network (Freeman, 1978); centrality proximity to the centre of the graph (Closeness Centrality), which is the reverse amount of the sum of the lengths of all the shortest pathways passing through the top (Sabidussi, 1966); and centrality (Betweenness Centrality), which is the number of the shortest pathways passing through the vertex (Freeman, 1977).

Previously, we proposed the concept of an associative gene network (Ivanisenko V.A. et al., 2015), an extended gene network whose vertices can represent not only molecular genetic objects (genes, proteins, metabolites, etc.), but also higher-level factors (biological processes, phenotypic traits, diseases, factors of the external environment, etc.) related to regulatory, physical, chemical or associative interactions (Ivanisenko V.A. et al., 2015). Automatic analysis of texts of scientific publications and factual databases was used to create knowledge bases: ANDSystem, which contains associative gene networks for animals and humans (Ivanisenko V.A. et al., 2015, 2019; Ivanisenko T.V. et al., 2020), and SOLANUM TUBEROSUM (Saik et al., 2017; Ivanisenko T.V. et al., 2018), which contains associative gene networks of plants. Analysis of associative gene networks from the knowledge bases ANDSystem and SOLANUM TUBEROSUM was used to develop methods for

prioritizing human genes associated with diseases (Saik et al., 2018, 2019) and the potato genes involved in breeding and the development of agriculturally meaningful traits (Demenkov et al., 2019), respectively. The prioritization method was based on the assessment of the centrality indicator according to the degree of vertices corresponding to genes in the analysed associative gene network.

In this work, an approach is proposed to prioritize biological processes by calculating the vertex centrality indicator in the associative gene network. This work uses the example of the model plant *Arabidopsis thaliana* and the information about the biological processes associated with the response of plants to the unfavourable factors cadmium content, drought conditions and salt stress in the knowledge base SOLANUM TUBEROSUM.

Materials and methods

SOLANUM TUBEROSUM (Saik et al., 2017; Ivanisenko T.V. et al., 2018) consists of three main blocks.

The first (block 1) is an automatic analysis of texts of scientific publications and factual databases that is designed to extract information about the relationship between objects using semantic–linguistic templates. To extract knowledge, ANDSystem software tools configured to the subject area under study are used (Ivanisenko V.A. et al., 2015). The setting of the subject area includes the creation of new semantic–linguistic templates that take into account the specifics of the texts of scientific publications in the field of biology of plants and crop production.

The second block is the knowledge base containing object dictionaries and information about the relationships between objects extracted in block 1 in the form of an integrated associative gene network (a graph in which the vertices correspond to objects and the ribs indicate the specified types of connections). The SOLANUM TUBEROSUM knowledge base contains dictionaries of molecular genetic objects (genes, proteins, metabolites, microRNAs), Gene Ontology biological processes (Gene Ontology Consortium, 2019), phenotypic traits, diseases and pathogens of potatoes and model organisms, among other information. To describe relationships between these objects in associative gene networks, more than 25 different types of interactions are used. These correspond to, for example, physical interactions, catalytic reactions, regulation, participation and associations.

In block 3, ANDVisio (Demenkov et al., 2012), designed to reconstruct and analyse associative gene networks related to the problem under study, is applied based on the knowledge base of SOLANUM TUBEROSUM. The SOLANUM TUBEROSUM knowledge base can be accessed via the Internet at https://www-bionet.sscc.ru/and/plant/.

The number of interactions in the SOLANUM TUBEROSUM knowledge base for three plant species is shown in Table 1.

The schematic diagram of all stages of the prioritization is shown in Fig. 1. A detailed description of the development of the SOLANUM TUBEROSUM knowledge base (see Fig. 1, *a*) can be found in previous studies (Saik et al., 2017;

Ivanisenko T.V. et al., 2018). The prioritization algorithm consists of several steps (see Fig. 1, *b*), which are fully automated in the ANDVisio program. The ANDVisio program provides the user interface for access to the SOLANUM TUBEROSUM knowledge base, and has a wide range of tools for reconstruction, graphic visualisation and analysis of associative gene networks.

In the first step, the associative gene network for the problem being studied is automatically reconstructed as specified by the user, incorporating biological processes from the Gene Ontology database. To achieve this, the "Reconstruction of the Network" procedure is implemented in ANDVisio. In the interface window for this procedure, the name of the Gene Ontology biological process is selected from the list provided (which contains all of the Gene Ontology biological processes, updated annually). Upon the execution of this procedure, an associative gene network related to a given biological process will be automatically reconstructed and visualised in graphical form. The reconstruction algorithm includes an appeal using MySQL requests to the SOLANUM TUBEROSUM knowledge base, which generates a list of genes with a connection to the specified biological process. The user can choose whether to use interactions identified by analysing the texts of scientific publications or those extracted from factual databases. Genes from this list become vertices in the desired associative gene network. The formal description of the algorithm is shown in Fig. 1, *b*. Using this algorithm, networks were reconstructed for three biological processes: "Response to Cadmium Ion", "Drought Tolerance" and "Response to Salt Stress".

In the second step, for each reconstructed associative gene network, biological processes were prioritized according to the algorithm shown in Fig. 1, *b*. Prioritization is an automated iterative process performed by the ANDVisio program. Briefly, all Gene Ontology biological processes in the SOLANUM TUBEROSUM knowledge base dictionaries are sequentially analysed. The analysis included the calculation of the CTC (cross-talk centrality) indicator. The biological processes were ranked according to their CTC values thus, biological processes with the highest indicator received maximum priority.

The CTC indicator for the vertex *i* was calculated using the formula:

$$CTC_i = N_i/M, \quad M \neq 0,$$

where N_i is the number of vertices to which the *i* vertex is connected by an edge in an analysed column of an associative gene network, *M* is the total number of vertices of the associative gene network, and *i* takes values from *1* to *M*. The values of the CTC indicator range from 0 to 1, with the value of 0 occurring when the peak under consideration on the network is not related to other vertices. The characteristic N_i is known as the indicator of the Degree centrality of the vertex (Freeman, 1978). This indicator is used because it is assumed that the larger the vertices associated with this vertex, the greater its influence in the associative gene network and the molecular genetic mechanism as a whole. To assess the significance of the centrality of the vertex (biological process), an approach to calculate the overrepresentation of biological processes was used (Subramanian et al., 2005). Using this approach,

П.С. Деменков, Е.А. Ощепкова
Т.В. Иванисенко, В.А. Иванисенко

Приоритизация биологических процессов, описывающих
ответ растений на неблагоприятные факторы внешней среды

2021
25•5

Table 1. Number of basic types of interactions, including molecular genetic interactions and associations, presented in the knowledge base of SOLANUM TUBEROSUM for three plant species

Interaction type	*Solanum tuberosum*/potato	*Arabidopsis thaliana*/arabidopsis	*Zea mays*/maize
Association	4783	294 520	15 280
Catalyze	1140	26 842	1203
Regulation*	346	20 213	1120
Molecular interaction	69	16 589	178
Involvement	3857	81 067	5906

* The total number of links is shown for the following types of interactions: upregulation, downregulation, regulation, activity upregulation, activity downregulation, activity regulation, expression upregulation, expression downregulation, expression regulation, transport upregulation, transport downregulation and transport regulation.

a

Dictionaries of objects

Gene/proteins – set **Vg**
Metabolites – **Vm**
GO biological processes – **Vgo**
Phenotypic traits – **Vf**
Diseases and pathogens – **Vd**

Automatic knowledge extraction from the texts of scientific publications

ANDSystem text mining tools:
text normalization,
entity name recognition,
interaction extraction.
Interaction types:
Protein-protein interactions – **Eint**
Associations – **Ea**
Regulations – **Er**
…

The SOLANUM TUBEROSUM knowledge base

The knowledge representation model is a semantic network.
Let the knowledge graph is **AGN(V, E)**, where **V** = {Vgo, Vg, Vm, …} is a set of vertices (objects),
E = {Eint, Ea, Er, …} – a set of edges (interactions)

ANDVisio program

Reconstruction and analysis of associative gene networks, based on access to the SOLANUM TUBEROSUM knowledge base

b

The ANDVisio program

The user selects the Gene Ontology process of interest **vgo** ∈ **Vgo** and starts the procedure for reconstructing the associative gene network according to the selected process, provided by the ANDVisio program

Reconstruction of the associative gene network vgo

MySQL query to the SOLANUM TUBEROSUM knowledge base.
The result is a subgraph AGN_vgo(v, **e** ∈ **E**).
v = {**vgo**, vg ∈ Vg}

Prioritization of biological processes Gene Ontology

$\forall\, m \in \text{Vgo}, m \neq \text{vgo}$

Extending AGN_vgo by adding a vertex from the Vgo set:
AGN_vgo(v, **e** ∈ **E**) $\Rightarrow$ AGN_vgo({v, *m*}, **e** ∈ **E**)

CTC calculation for vertices M = {*m*}

$i < N$

Sorting M = {*m*}
CTC(m_i) ≥ CTC(m_j), $i < j$

Fig. 1. Schematic diagram of all steps performed during prioritization: *a*, creating the SOLANUM TUBEROSUM knowledge base; *b*, prioritization of Gene Ontology biological processes (Gene Ontology Consortium, 2019).

the probability (*p*-value) that this vertex will have N_i or more links by random chance in the associative gene network was evaluated using a hypergeometric distribution adjusted to the Benjamini–Yekutieli multiple comparison procedure (Benjamini, Yekutieli, 2001).

For the characterisation of the identified processes, a cluster analysis was carried out by the Markov Cluster algorithm (Van Dongen, Abreu-Goodger, 2012), using the semantic metric for biological processes by Wang and colleagues (Wang J.Z. et al., 2007).

Results and discussion

Using the SOLANUM TUBEROSUM knowledge base, we reconstructed the associative gene networks of *A. thaliana* that described the interaction of genes with biological processes related to plant responses to adverse environmental factors, including drought, salt stress and increased cadmium content. To study the potential molecular genetic mechanisms underlying the reconstructed gene networks, biological processes were prioritized based on the centrality of their interactions with the network's genes/proteins.

Plant responses to drought

The associative gene network for the biological process "drought tolerance", reconstructed in the SOLANUM TUBEROSUM knowledge base, included 292 vertices (proteins) interconnected by 440 ribs (Fig. 2). To simplify the representation of the interactions, the genes corresponding to the proteins are not illustrated. The relatively large number of proteins associated with the term "drought tolerance" can be explained by the fact that plant resistance to drought is conditioned by numerous aspects of plant physiology – a search query in PubMed for the keywords "drought" and "plants" yields more than 18,000 publications.

It should be noted that the analysis based on automatic processing of scientific literature data depends on how deeply the subject area has been studied. Thus, our conclusions can indicate the presence of an interaction, according to the published data, but cannot assert the absence of an interaction simply because the interaction is not discussed in the literature.

Among the biological processes associated with proteins in the associative gene network "drought tolerance", there were 208 processes with a Q-value below the significance level of 0.05. Their centrality scores ranged from 0.0067 to 0.58. The identified biological processes were divided into 22 clusters according to semantic proximity (Fig. 3).

A list of 12 biological processes with the largest number of links with other objects in the drought tolerance gene network is shown in Table 2.

The prioritization enables the identification of such fundamental processes as "transcription", "signaling" and "gene expression", which fell into the clusters "post-translational modifications", "cellular processes" and "cellular metabolism", respectively. The processes associated with transcription, translation and gene expression are widely discussed in the context of transcription factors and their participation in plant responses to adverse environmental conditions (Leng, Zhao, 2020).

Signaling processes induced by external and endogenous factors also play an important role. In particular, the process "ABA signaling" (abscisic acid signaling pathway) was highly rated (see Table 2). The CTC level for this process was found to be 0.34. This means that when the analysed associative gene network "drought tolerance" is expanded by adding the "ABA signaling" vertex, 34 % of all vertices will be associated with it. Many studies have demonstrated the high importance of ABA in plant responses to drought (approximately 3000 publications in PubMed). It has been shown that ABA is synthesised in leaves in response to a water deficiency signal transmitted from roots to shoots and is involved in mechanisms of drought resistance, such as stomata closure and production of osmotic defence proteins (Takahashi et al., 2018). ABA is also involved in the regulation of flowering time ("flowering" is on the list of priority processes, see Table 2). For example, it was shown that under drought conditions, ABA-responsive element binding factor 3 (ABF3) and ABF4, along with nuclear transcription factor Y subunit gamma (NF-YC), enhance the expression of suppressor of overexpression of constans 1 (SOC1), thus promoting acceleration of plant flowering, allowing the plant to complete its life cycle at an earlier date (Hwang et al., 2019).

Another process associated with the response of plants to drought, located in the upper lines of the prioritization results (see Table 2), is phosphorylation (one of the main ways to transmit regulatory signals in a cell). For example, it was shown in *A. thaliana* that the transcriptional activation of genes sensitive to drought requires phosphorylation of the transcription factor RD26 by the kinase brassinosteroid in-

Table 2. Ranking of biological processes according to their potential relationship with the "drought tolerance" process

No.	Biological process	Cluster name	Number of links*	CTC	*p*-value	*Q*-value**
1	Transcription	post-translational modifications	173	0.59	7E–64	3E–61
2	Signaling	cellular processes	121	0.41	1E–81	7E–79
3	Gene expression	cellular metabolism	121	0.41	2E–73	1E–70
4	Signaling pathways	signaling pathways	107	0.37	1E–68	5E–66
5	Transcriptional control	transcriptional control	106	0.36	1E–27	1E–25
6	ABA signaling	signaling pathways	98	0.34	1E–86	1E–83
7	Flowering	regulation of leaf senescence	84	0.29	8E–50	2E–47
8	Phosphorylation	cellular metabolism	74	0.25	4E–20	4E–18
9	Degradation	cellular metabolism	69	0.24	5E–33	7E–31
10	Cellular processes	cellular processes	59	0.20	2E–61	5E–59
11	Translation	post-translational modifications	59	0.20	6E–24	6E–22
12	Response to stress	response to stimulus	58	0.20	5E–33	7E–31

* The number of connections between the biological process and proteins/genes in the associative network.
** Significance adjusted for multiple Benjamini–Yekutieli comparisons.

П.С. Деменков, Е.А. Ощепкова
Т.В. Иванисенко, В.А. Иванисенко

Приоритизация биологических процессов, описывающих
ответ растений на неблагоприятные факторы внешней среды

2021
25 • 5
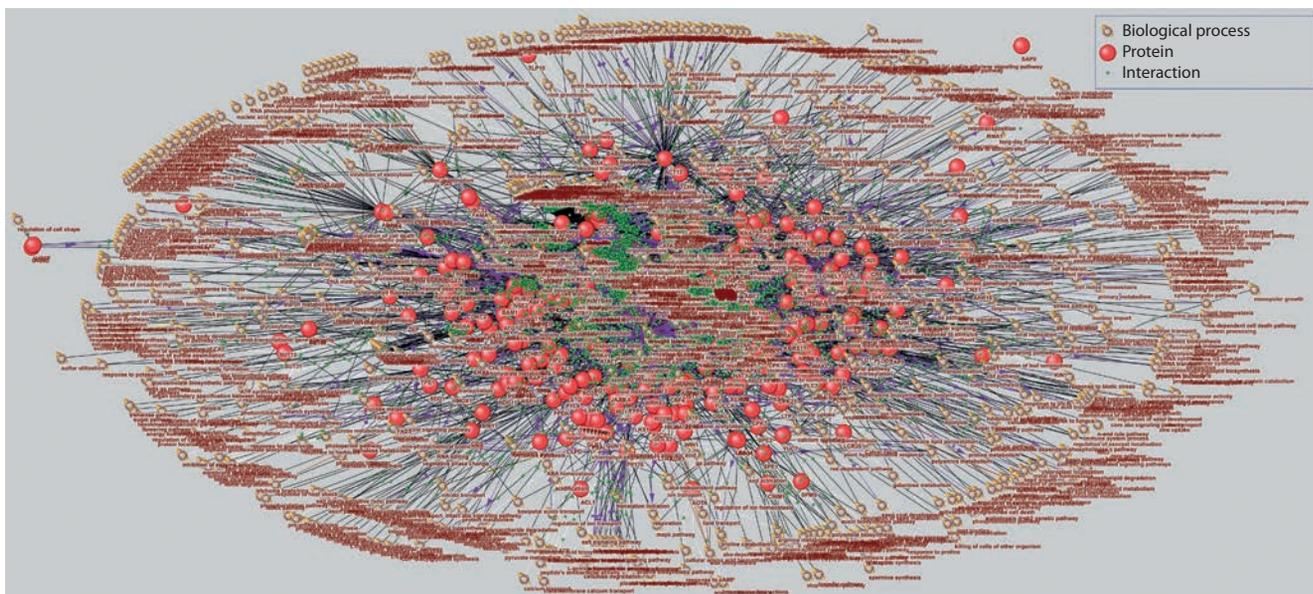
Fig. 2. Associative gene network "drought tolerance" for *Arabidopsis thaliana*, including proteins and biological processes as vertices.
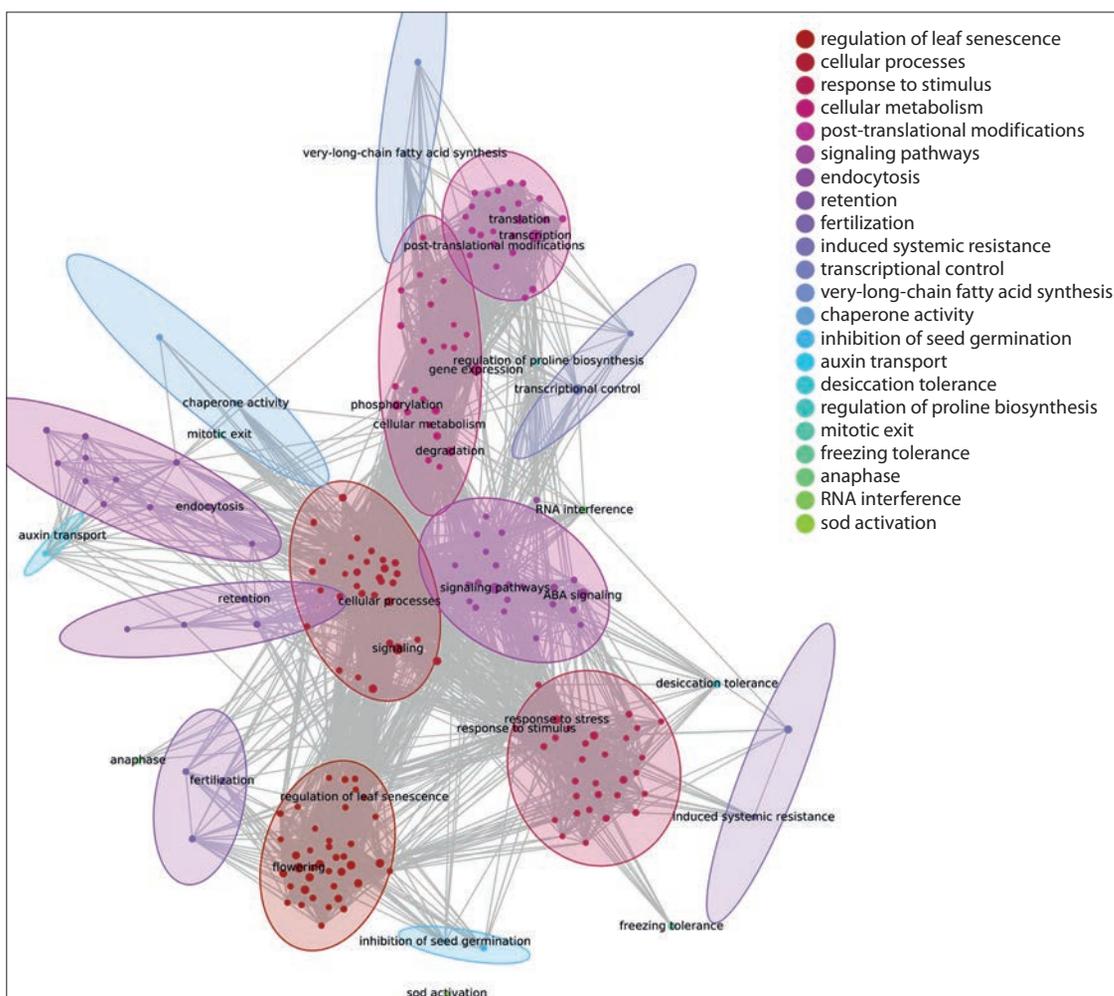
Fig. 3. Semantic proximity clustering of biological processes significantly overrepresented in the "drought tolerance" associative gene network.

Ovals outline vertices from one cluster. The colour of the vertices and ovals matches the colour of the cluster in the legend.

sensitive 2 (BIN2) (Jiang et al., 2019). In a study of a drought-resistant rapeseed cultivar (*Brassica napus* L.), an important role was shown for beta carbonic anhydrase 1 (βCA1) phosphorylation in photosynthesis regulation under drought conditions (Wang L. et al., 2016).

The position of biological processes such as plant resistance to cold and salt stress in the list of prioritized processes in the analysis of plant response to drought conditions deserves special discussion. Although these processes were not at the top of the list of processes with the highest priority for the associative gene network "drought tolerance" (see Table 2), they had a rather high rating (CTC = 0.1). Indeed, plant resistance to various adverse environmental factors is often mediated by the same molecular genetic mechanisms. For example, it was shown in corn (*Zea mays*) that the expression of the transcription factor MYB (MYB3R) is induced during both drought and salt stress, contributing to plant resistance to these environmental factors (Wu et al., 2019). In soybean (*Glycine max*), it was shown that another transcription factor in the MYB family, MYB118, also plays an important role in plant resistance to drought (Du et al., 2018). Its expression, like that of MYB3R, is induced under drought and salt stress, and the newly synthesised MYB118 enhances the expression of stress-associated genes that mediate the plant response to these stresses. In cotton (*Gossypium* spp.), an important role was shown for proteins in the cyclin-dependent kinase family in plant response to drought and salt stress (Magwanga et al., 2018). The mechanisms of plant response to drought overlap with the mechanisms of response to cold stress: in both cases, the synthesis and mobilisation of abscisic acid is induced in the vascular tissue of the leaves, which is necessary for the regulation of stomatal closure (reviewed in Agurla et al., 2018).

Plant response to salinity

The "response to salt stress" associative gene network contained 81 vertices and 102 links (Fig. 4). Twelve of the highest-priority processes with a centrality index of at least 0.2 were identified (Table 3). As in the case of the drought resistance associative gene network, all of the processes included in the list of the highest priority processes are statistically significant.

Among the biological processes associated with proteins in the "response to salt stress" associative gene network, 85 had a *Q*-value of <0.05. Their centrality scores ranged from 0.02 to 0.34. The identified biological processes were divided into 14 clusters according to semantic proximity (Fig. 5).

A list of the 12 biological processes with the largest number of connections with other objects in the "response to salt stress" gene network is shown in Table 3.

Salt stress tolerance is widely represented in the scientific literature – a PubMed search for the keywords "salt" and "plants" yields more than 24,000 publications. The analysis yielded a list of top-rated processes that included "seed germination", which fell into the "respiration" cluster (see Table 3). For example, the authors showed in *Stylosanthes humilis* that increased salt content suppresses seed germination and results in decreased ethylene production and increased abscisic acid production (Silva et al., 2018). Leaf senescence during normal plant development and caused by salt stress occurs via common hydrogen peroxide-mediated signaling pathways, as shown in *A. thaliana* (Allu et al., 2014).

Plant response to the presence of cadmium

Cadmium is toxic to most plants and animals (Genchi et al., 2020). It can be present in soil, causing stress responses in plants, suppressing the growth of roots and shoots and re-

Table 3. Ranking of biological processes according to their potential relationship with the "response to salt stress" process

No.	Biological process	Cluster name	Number of links*	CTC	*p*-value	*Q*-value**
1	Transcription	RNA synthesis	30	0.34	1.8E–5	8.0E–4
2	Signaling	cellular processes	28	0.32	4.5E–16	1.3E–13
3	Salt tolerance	response to stress	28	0.32	6.5E–32	3.1E–29
4	Gene expression	cellular lipid metabolism	24	0.27	6.4E–11	5.9E–9
5	Phosphorylation	cellular processes	22	0.25	5.9E–7	3.1E–5
6	Signaling pathways	signaling pathways	20	0.23	1.7E–9	1.2E–7
7	Response to salt	response to stress	18	0.20	5.3E–25	1.9E–22
8	Flowering	respiration	17	0.19	3.6E–8	2.2E–6
9	Seed germination	respiration	16	0.18	4.2E–12	5.2E–10
10	Drought tolerance	response to stress	15	0.17	1.2E–9	9.2E–8
11	Response to stress	response to stress	15	0.17	1.6E–8	5.2E–10
12	Degradation	cellular processes	14	0.16	1.4E–5	6.0E–4

* The number of connections between the biological process and proteins/genes in the associative network.
** Significance adjusted for multiple Benjamini–Yekutieli comparisons.

П.С. Деменков, Е.А. Ощепкова
Т.В. Иванисенко, В.А. Иванисенко

Приоритизация биологических процессов, описывающих
ответ растений на неблагоприятные факторы внешней среды

2021
25•5
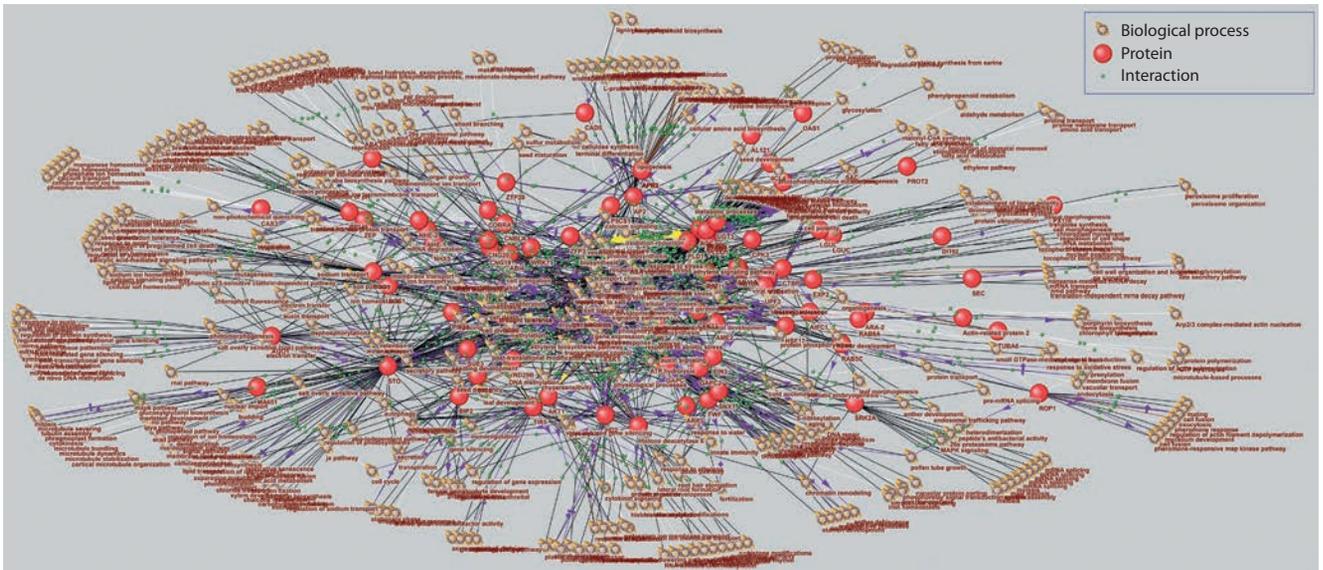
Fig. 4. Associative gene network "response to salt stress" for *Arabidopsis thaliana*, including proteins and biological processes as vertices.
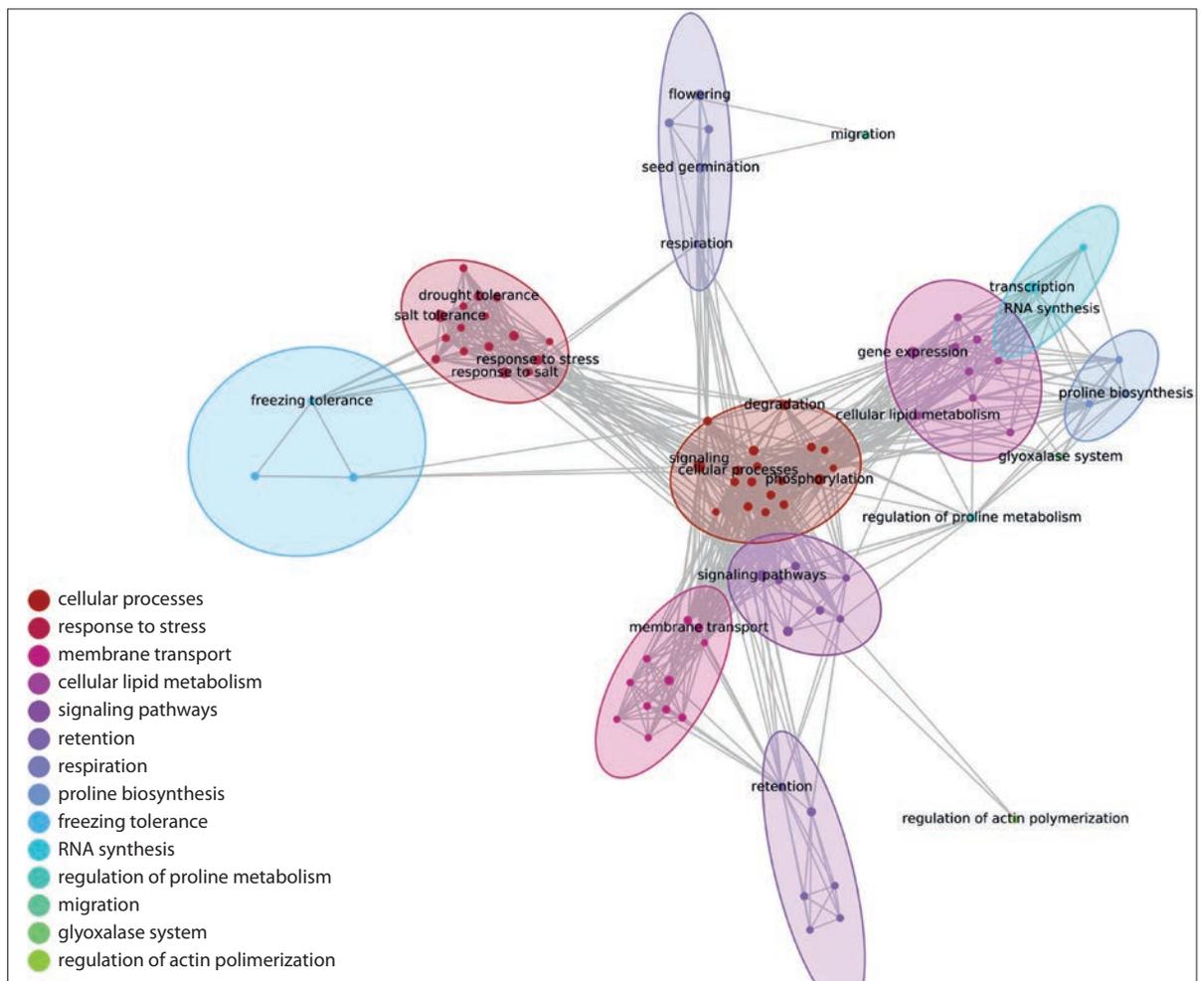
Fig. 5. Semantic proximity clustering of biological processes significantly overrepresented in the "response to salt stress" associative gene network.

Ovals outline vertices from one cluster. The colour of the vertices and ovals matches the colour of the cluster in the legend.

Fig. 6. Associative gene network "response to cadmium ion" for *Arabidopsis thaliana*, including proteins and biological processes as vertices.
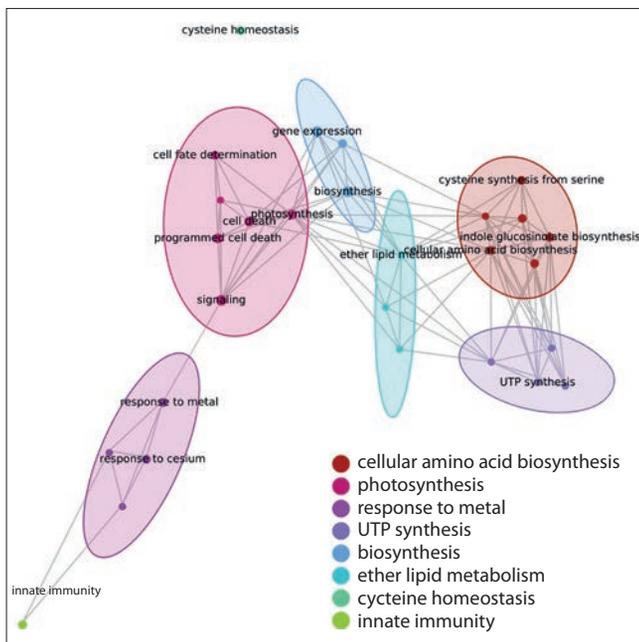
Fig. 7. Semantic proximity clustering of biological processes significantly overrepresented in the "response to cadmium ion" associative gene network.

Ovals outline vertices from one cluster. The colour of the vertices and ovals matches the colour of the cluster in the legend.

ducing the rate of photosynthesis and nutrient consumption (Genchi et al., 2020; Kaya et al., 2020). The increased interest in the problem is due to some plants being prone to cadmium hyperaccumulation, which raises prospects for their use for both soil cleaning and other industrial purposes (Küpper, Leitenmaier, 2013). However, the mechanisms underlying

plant response to cadmium remain poorly understood. Nine genes were associated with the biological process "response to cadmium ion" in the SOLANUM TUBEROSUM knowledge base for *A. thaliana* (Fig. 6).

Among the biological processes related to proteins in the "response to cadmium ion" associative gene network, 28 had a Q-value below 0.05. Their centrality scores ranged from 0.11 to 0.44. The identified biological processes were divided into 8 clusters according to semantic proximity (Fig. 7).

A list of the 12 biological processes with the largest number of connections with other objects in the "response to cadmium ion" gene network is shown in Table 4. Among the biological processes with a high rating in the cadmium response gene network, there were two processes associated with cell death ("programmed cell death" and "cell death"), which were in the "photosynthesis" cluster. The importance of programmed cell death in response to cadmium has been confirmed in dozens of publications. For example, it was shown that cadmium causes morphophysiological changes and programmed cell death in *Genipa americana* L. (Souza et al., 2011); moreover, in tomato cell culture, it was demonstrated that cadmium induces programmed cell death using caspase-like proteases (Iakimova et al., 2008).

In studies of the molecular mechanisms of plant responses to unfavourable factors, it may be of great interest to identify both general and specific biological processes. Below we consider examples of general and specific processes.

Comparative analysis of reconstructed associative gene networks

Comparison of the "response to salt stress" and "drought tolerance" associative networks revealed 41 common biological processes, 40 processes specific to the response to salt stress

П.С. Деменков, Е.А. Ощепкова
Т.В. Иванисенко, В.А. Иванисенко

Приоритизация биологических процессов, описывающих
ответ растений на неблагоприятные факторы внешней среды

2021
25•5

Table 4. Ranking of biological processes according to their potential relationship with the "response to cadmium ion" process

No.	Biological process	Cluster name	Number of links*	CTC	*p*-value	*Q*-value**
1	Signaling	photosynthesis	4	0.44	5.0E–4	1.0E–2
2	Biosynthesis	biosynthesis	4	0.44	3.1E–4	8.4E–3
3	Gene expression	biosynthesis	4	0.44	1.0E–3	1.7E–2
4	Programmed cell death	photosynthesis	3	0.33	7.8E–5	2.6E–3
5	Cell death	photosynthesis	3	0.33	5.1E–4	1.3E–2
6	Photosynthesis	photosynthesis	3	0.33	7.5E–4	1.3E–2
7	Indole glucosinolate biosynthesis	cellular amino acid biosynthesis	2	0.22	1.9E–5	9.7E–4
8	Cysteine synthesis from serine	cellular amino acid biosynthesis	2	0.22	3.8E–5	1.4E–2
9	Cysteine biosynthesis	cellular amino acid biosynthesis	2	0.22	1.6E–4	4.6E–3
10	Camalexin biosynthesis	cellular amino acid biosynthesis	2	0.22	2.9E–3	4.0E–2
11	Cellular amino acid biosynthesis	cellular amino acid biosynthesis	2	0.22	3.7E–3	4.4E–2
12	Cell fate determination	photosynthesis	2	0.22	2.8E–5	1.1E–3

* The number of connections between the biological process and proteins/genes in the associative network.
** Significance adjusted for multiple Benjamini–Yekutieli comparisons.

network and 156 processes specific to the drought tolerance network. Examination of all three associative gene networks revealed five common biological processes: signaling, cell death, programmed cell death, gene expression and photosynthesis. Notably, the tables of biological processes with the greatest number of connections with other objects (see Tables 2–4) mainly include general processes that have a statistically significant relationship with two or all three of the gene networks of plant responses to unfavourable factors.

Identification of biological processes specific to the plant response to drought compared with the responses to salinity and cadmium has revealed a group of processes associated with seed development: "seeding development", "seed development", "seed dormancy", "seed maturation", "regulation of seed size" and "inhibition of seed germination". Indeed, many articles demonstrating the impact of drought on these processes have been published. It was shown that a lack of water leads to a decrease in the rate and duration of maturation of lentil seeds, leading to a decrease in their size (Sehgal et al., 2019). Similar studies on soybean lines clearly showed that drought conditions affect the quality of seed preparation and subsequent germination (Wijewardana et al., 2019).

When identifying biological processes specific to the response of plants to salinity, a group of processes associated with ion transport was distinguished: "ion transport", "membrane transport", "sodium transport", "cation transport", "transmembrane proton transport", "sodium ion transmembrane transport" and "potassium ion homeostasis". Indeed, the adaptation of plants to conditions of salt excess depends on their ability to remove Na+ and Cl– ions or to increase resistance to osmotic stress and ion accumulation in tissues (Munns, Tester, 2008).

Comparing the biological processes specific to the plant response to the presence of cadmium with those involved in

responses to drought and salinity highlights several processes associated with lipid metabolism: "ether lipid metabolism", "phospholipid degradation" and "glycerophospholipid metabolism". It is known that lipid peroxidation is one of the manifestations of the toxicity of cadmium. For example, in tomatoes, it has been shown that cadmium induces substantial changes in lipid composition, causing premature aging of leaves (Djebali et al., 2005).

Conclusion

The molecular genetic mechanisms of potato resistance to unfavourable conditions remain rather poorly understood. The accumulated knowledge about the *A. thaliana* model plant can shed light on the molecular interactions in the gene networks involved in the response of potatoes to stress conditions. This approach is also used in studies by other researchers. For example, Ž. Ramšak and colleagues (2018) built a model of immune signaling pathways in *A. thaliana*, then superimposed the data obtained onto *S. tuberosum* to gain new knowledge about the immune signaling pathways in potatoes.

The Gene Ontology database (Gene Ontology Consortium, 2019) contains information on biological processes and the genes involved in their functioning. Using Gene Ontology terms, molecular genetic mechanisms involved in the breeding and important traits of plants, knowledge of which is necessary for the development of modern approaches for marker-oriented and genomic selection, can be described. However, the analysis of gene networks built by analysing literature sources containing data on the relationship between genes and the studied processes is also of great interest.

Analysis of the prioritization of biological processes that may be involved in the response of plants to unfavourable factors (cadmium, salinity and drought conditions), performed

using the reconstruction of associative genes, showed good agreement with the well-known literature data. Among the biological processes that received high priority were fundamental processes related to the expression of genes, post-translational modification of proteins and degradation, as well as processes associated with cell death. The various ways to transmit signals occupies an important role in the mechanisms of plants' responses to adverse factors.

When comparing the biological processes identified as playing a role in responses to unfavourable conditions, in addition to finding common processes associated with all three of the unfavourable factors in the external environment, we also identified processes that were statistically significantly connected with only one of the associative networks studied. Among the biological processes specifically involved in response to drought is a rather large group associated with the development of seeds (for example, "seeding development" and "seed dormancy"). Among the processes involved specifically in the response to saline stress are a group of processes associated with ion transport (for example, "ion transport", "transmembrane proton transport" and "sodium ion transmembrane transport"). Finally, processes involved in lipid metabolism (such as "phospholipid degradation") were specific to the plant response to cadmium.

References

Agurla S., Gahir S., Munemasa S., Murata Y., Raghavendra A.S. Mechanism of stomatal closure in plants exposed to drought and cold stress. In: Iwaya-Inoue M., Sakurai M., Uemura M. (Eds.). Survival Strategies in Extreme Cold and Desiccation. Advances in Experimental Medicine and Biology. Vol. 1081. Singapore: Springer, 2018;215-232. DOI 10.1007/978-981-13-1244-1_12.

Allu A.D., Soja A.M., Wu A., Szymanski J., Balazadeh S. Salt stress and senescence: identification of cross-talk regulatory components. *J. Exp. Bot.* 2014;65(14):3993-4008. DOI 10.1093/jxb/eru173.

Arruda M.P., Lipka A.E., Brown P.J., Krill A.M., Thurber C., Brown-Guedira G., Dong Y., Foresman B.J., Kolb F.L. Comparing genomic selection and marker-assisted selection for Fusarium head blight resistance in wheat (*Triticum aestivum* L.). *Mol. Breed.* 2016;36(7):84. DOI 10.1007/s11032-016-0508-5.

Bargsten J.W., Nap J.P., Sanchez-Perez G.F., van Dijk A.D. Prioritization of candidate genes in QTL regions based on associations between traits and biological processes. *BMC Plant Biol.* 2014;14:330. DOI 10.1186/s12870-014-0330-3.

Benjamini Y., Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann. Statist.* 2001;29(4):1165-1188. DOI 10.1214/aos/1013699998.

Cai Z., Guldbrandtsen B., Lund M.S., Sahana G. Prioritizing candidate genes for fertility in dairy cows using gene-based analysis, functional annotation and differential gene expression. *BMC Genomics.* 2019;20(1):255. DOI 10.1186/s12864-019-5638-9.

Cesur M.F., Siraj B., Uddin R., Durmuş S., Çakır T. Network-based metabolism-centered screening of potential drug targets in *Klebsiella pneumoniae* at genome scale. *Front. Cell. Infect. Microbiol.* 2020;9:447. DOI 10.3389/fcimb.2019.00447.

Chen Y., Jiang T., Jiang R. Uncover disease genes by maximizing information flow in the phenome–interactome network. *Bioinformatics.* 2011;27(13):i167-i176. DOI 10.1093/bioinformatics/btr213.

Cho A., Shim J.E., Kim E., Supek F., Lehner B., Lee I. MUFFINN: cancer gene discovery via network analysis of somatic mutation data. *Genome Biol.* 2016;17(1):1-16. DOI 10.1186/s13059-016-0989-x.

Crossa J., Pérez-Rodríguez P., Cuevas J., Montesinos-López O., Jarquín D., de Los Campos G., Burgueño J., González-Camacho J.M.,

Pérez-Elizalde S., Beyene Y., Dreisigacker S., Singh R., Zhang X., Gowda M., Roorkiwal M., Rutkoski J., Varshney R.K. Genomic selection in plant breeding: methods, models, and perspectives. *Trends Plant Sci.* 2017;22(11):961-975. DOI 10.1016/j.tplants.2017.08.011.

Demenkov P.S., Ivanisenko T.V., Kolchanov N.A., Ivanisenko V.A. ANDVisio: a new tool for graphic visualization and analysis of literature mined associative gene networks in the ANDSystem. *In Silico Biol.* 2012;11(3):149-161. DOI 10.3233/ISB-2012-0449.

Demenkov P.S., Saik O.V., Ivanisenko T.V., Kolchanov N.A., Kochetov A.V., Ivanisenko V.A. Prioritization of potato genes involved in the formation of agronomically valuable traits using the SOLANUM TUBEROSUM knowledge base. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding.* 2019; 23(3):312-319. DOI 10.18699/VJ19.501.

Djebali W., Zarrouk M., Brouquisse R., El Kahoui S., Limam F., Ghorbel M.H., Chaïbi W. Ultrastructure and lipid alterations induced by cadmium in tomato (*Lycopersicon esculentum*) chloroplast membranes. *Plant Biol. (Stuttg).* 2005;7(4):358-368. DOI 10.1055/s-2005-837696.

Du Y.T., Zhao M.J., Wang C.T., Gao Y., Wang Y.X., Liu Y.W., Chen M., Chen J., Zhou Y.B., Xu Z.S., Ma Y.Z. Identification and characterization of *GmMYB118* responses to drought and salt stress. *BMC Plant Biol.* 2018;18(1):320. DOI 10.1186/s12870-018-1551-7.

Freeman L.C. A set of measures of centrality based on betweenness. *Sociometry.* 1977;40:35-41. DOI 10.2307/3033543.

Freeman L.C. Centrality in social networks conceptual clarification. *Social Networks.* 1978;1(3):215-239. DOI 10.1016/0378-8733(78)90021-7.

Genchi G., Sinicropi M.S., Lauria G., Carocci A., Catalano A. The effects of cadmium toxicity. *Int. J. Environ. Res. Public Health.* 2020; 17(11):3782. DOI 10.3390/ijerph17113782.

Gene Ontology Consortium. The gene ontology resource: 20 years and still GOing strong. *Nucleic Acids Res.* 2019;47(D1):D330-D338. DOI 10.1093/nar/gky1055.

Hwang K., Susila H., Nasim Z., Jung J.Y., Ahn J.H. *Arabidopsis* ABF3 and ABF4 transcription factors act with the NF-YC complex to regulate *SOC1* expression and mediate drought-accelerated flowering. *Mol. Plant.* 2019;12(4):489-505. DOI 10.1016/j.molp.2019. 01.002.

Iakimova E.T., Woltering E.J., Kapchina-Toteva V.M., Harren F.J., Cristescu S.M. Cadmium toxicity in cultured tomato cells – role of ethylene, proteases and oxidative stress in cell death signaling. *Cell Biol. Int.* 2008;32(12):1521-1529. DOI 10.1016/j.cellbi.2008.08.021.

Ivanisenko T.V., Saik O.V., Demenkov P.S., Ivanisenko N.V., Savostianov A.N., Ivanisenko V.A. ANDDigest: a new web-based module of ANDSystem for the search of knowledge in the scientific literature. *BMC Bioinformatics.* 2020;21(Suppl.11):228. DOI 10.1186/s12859-020-03557-8.

Ivanisenko T.V., Saik O.V., Demenkov P.S., Khlestkin V.K., Khlestkina E.K., Kolchanov N.A., Ivanisenko V.A. The SOLANUM TUBEROSUM knowledge base: the section on molecular-genetic regulation of metabolic pathways. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding.* 2018;22(1): 8-17. DOI 10.18699/VJ18.325. (in Russian)

Ivanisenko V.A., Demenkov P.S., Ivanisenko T.V., Mishchenko E.L., Saik O.V. A new version of the ANDSystem tool for automatic extraction of knowledge from scientific publications with expanded functionality for reconstruction of associative gene networks by considering tissue-specific gene expression. *BMC Bioinformatics.* 2019; 20(Suppl.1):34. DOI 10.1186/s12859-018-2567-6.

Ivanisenko V.A., Saik O.V., Ivanisenko N.V., Tiys E.S., Ivanisenko T.V., Demenkov P.S., Kolchanov N.A. ANDSystem: an Associative Network Discovery System for automated literature mining in the field of biology. *BMC Syst. Biol.* 2015;9(Suppl.2):S2. DOI 10.1186/1752-0509-9-S2-S2.

П.С. Деменков, Е.А. Ощепкова
Т.В. Иванисенко, В.А. Иванисенко

Приоритизация биологических процессов, описывающих
ответ растений на неблагоприятные факторы внешней среды

2021
25·5

Jeong H., Mason S.P., Barabási A.L., Oltvai Z.N. Lethality and centrality in protein networks. *Nature*. 2001;411(6833):41-42. DOI 10.1038/35075138.

Jha M., Roy S., Kalita J.K. Prioritizing disease biomarkers using functional module based network analysis: a multilayer consensus driven scheme. *Comput. Biol. Med*. 2020;126:104023. DOI 10.1016/j.compbiomed.2020.104023.

Jia P.L., Zheng S.Y., Long J.R., Zheng W., Zhao Z.M. dmGWAS: dense module searching for genome-wide association studies in protein–protein interaction networks. *Bioinformatics*. 2011;27:95-102. DOI 10.1093/bioinformatics/btq615.

Jiang H., Tang B., Xie Z., Nolan T., Ye H., Song G.Y., Walley J., Yin Y. GSK3-like kinase BIN2 phosphorylates RD26 to potentiate drought signaling in *Arabidopsis*. *Plant J*. 2019;100(5):923-937. DOI 10.1111/tpj.14484.

Kaya C., Ashraf M., Alyemeni M.N., Ahmad P. Responses of nitric oxide and hydrogen sulfide in regulating oxidative defence system in wheat plants grown under cadmium stress. *Physiol. Plant*. 2020;168(2):345-360. DOI 10.1111/ppl.13012.

Kochetov A.V., Glagoleva A.Y., Strygina K.V., Khlestkina E.K., Gerasimova S.V., Ibragimova S.M., Shatskaya N.V., Vasilyev G.V., Afonnikov D.A., Shmakov N.A., Antonova O.Y., Gavrilenko T.A., Alpatyeva N.V., Khiutti A., Afanasenko O.S. Differential expression of NBS-LRR-encoding genes in the root transcriptomes of two *Solanum phureja* genotypes with contrasting resistance to *Globodera rostochiensis*. *BMC Plant Biol*. 2017;17(Suppl.2):251. DOI 10.1186/s12870-017-1193-1.

Kolchanov N.A., Kochetov A.V., Salina E.A., Pershina L.A., Khlestkina E.K., Shumny V.K. Status and prospects of marker-assisted and genomic plant breeding. *Herald of the Russian Academy of Sciences*. 2017;87(2):125-131. DOI 10.1134/S1019331617020113.

Küpper H., Leitenmaier B. Cadmium-accumulating plants. In: Sigel A., Sigel H., Sigel R. (Eds.). Cadmium: From Toxicity to Essentiality. Metal Ions in Life Sciences. Vol. 11. Dordrecht: Springer, 2013;373-393. DOI 10.1007/978-94-007-5179-8_12.

Le D.H., Pham V.H. HGPEC: a Cytoscape app for prediction of novel disease-gene and disease-disease associations and evidence collection based on a random walk on heterogeneous network. *BMC Syst. Biol*. 2017;11(1):61. DOI 10.1186/s12918-017-0437-x.

Leng P., Zhao J. Transcription factors as molecular switches to regulate drought adaptation in maize. *Theor. Appl. Genet*. 2020;133(5):1455-1465. DOI 10.1007/s00122-019-03494-y.

Leung A., Bader G.D., Reimand J. Hypermodules: identifying clinically and phenotypically significant network modules with disease mutations for biomarker discovery. *Bioinformatics*. 2014;30:2230-2232. DOI 10.1093/bioinformatics/btu172.

Lin F., Fan J., Rhee S.Y. QTG-Finder: a machine-learning based algorithm to prioritize causal genes of quantitative trait loci in Arabidopsis and rice. *G3: Genes, Genomes, Genetics*. (*Bethesda*). 2019; 9(10):3129-3138. DOI 10.1534/g3.119.400319.

Lysenko A., Boroevich K.A., Tsunoda T. Arete – candidate gene prioritization using biological network topology with additional evidence types. *BioData Min*. 2017;10:22. DOI 10.1186/s13040-017-0141-9.

Magwanga R.O., Lu P., Kirungu J.N., Cai X., Zhou Z., Wang X., Diouf L., Xu Y., Hou Y., Hu Y., Dong Q., Wang K., Liu F. Whole genome analysis of cyclin dependent kinase (*CDK*) gene family in cotton and functional evaluation of the role of *CDKF4* gene in drought and salt stress tolerance in plants. *Int. J. Mol. Sci*. 2018;19(9):2625. DOI 10.3390/ijms19092625.

Munns R., Tester M. Mechanisms of salinity tolerance. *Annu. Rev. Plant Biol*. 2008;59:651-681. DOI 10.1146/annurev.arplant.59.032607.092911.

Pushpakom S., Iorio F., Eyers P.A., Escott K.J., Hopper S., Wells A., Doig A., Guilliams T., Latimer J., McNamee C., Norris A., Sanseau P., Cavalla D., Pirmohamed M. Drug repurposing: progress, challenges and recommendations. *Nat. Rev. Drug Discov*. 2019; 18(1):41-58. DOI 10.1038/nrd.2018.168.

Raj M.R., Sreeja A. Analysis of computational gene prioritization approaches. *Procedia Comput. Sci*. 2018;143:395-410. DOI 10.1016/j.procs.2018.10.411.

Ramšak Ž., Coll A., Stare T., Tzfadia O., Baebler Š., Van de Peer Y., Gruden K. Network modeling unravels mechanisms of crosstalk between ethylene and salicylate signaling in potato. *Plant Physiol*. 2018;178(1):488-499. DOI 10.1104/pp.18.00450.

Sabidussi G. The centrality index of a graph. *Psychometrika*. 1966;31: 581-603. DOI 10.1007/BF02289527.

Saik O.V., Demenkov P.S., Ivanisenko T.V., Bragina E.Y., Freidin M.B., Goncharova I.A., Dosenko V.E., Zolotareva O.I., Hofestaedt R., Lavrik I.N., Rogaev E.I., Ivanisenko V.A. Novel candidate genes important for asthma and hypertension comorbidity revealed from associative gene networks. *BMC Med. Genomics*. 2018;11(Suppl.1): 15. DOI 10.1186/s12920-018-0331-4.

Saik O.V., Demenkov P.S., Ivanisenko T.V., Kolchanov N.A., Ivanisenko V.A. Development of methods for automatic extraction of knowledge from texts of scientific publications for the creation of a knowledge base SOLANUM TUBEROSUM. *Selskokhozyaystvennaya Biologiya = Agricultural Biology*. 2017;52(1):63-74. DOI 10.15389/agrobiology.2017.1.63eng.

Saik O.V., Nimaev V.V., Usmonov D.B., Demenkov P.S., Ivanisenko T.V., Lavrik I.N., Ivanisenko V.A. Prioritization of genes involved in endothelial cell apoptosis by their implication in lymphedema using an analysis of associative gene networks with ANDSystem. *BMC Med. Genomics*. 2019;12(Suppl.2):47. DOI 10.1186/s12920-019-0492-9.

Schaefer R.J., Michno J.M., Jeffers J., Hoekenga O., Dilkes B., Baxter I., Myers C.L. Integrating coexpression networks with GWAS to prioritize causal genes in maize. *Plant Cell*. 2018;30(12):2922-2942. DOI 10.1105/tpc.18.00299.

Sehgal A., Sita K., Bhandari K., Kumar S., Kumar J., Vara Prasad P.V., Siddique K.H.M., Nayyar H. Influence of drought and heat stress, applied independently or in combination during seed development, on qualitative and quantitative aspects of seeds of lentil (*Lens culinaris* Medikus) genotypes, differing in drought sensitivity. *Plant Cell Environ*. 2019;42(1):198-211. DOI 10.1111/pce.13328.

Shim J.E., Lee I. Network-assisted approaches for human disease research. *Animal Cells Syst*. 2015;19:231-235.

Shim J.E., Lee T., Lee I. From sequencing data to gene functions: co-functional network approaches. *Animal Cells Syst*. 2017;21(2):77-83. DOI 10.1080/19768354.2017.1284156.

Silva N.C.Q., de Souza G.A., Pimenta T.M., Brito F.A.L., Picoli E.A.T., Zsögön A., Ribeiro D.M. Salt stress inhibits germination of *Stylosanthes humilis* seeds through abscisic acid accumulation and associated changes in ethylene production. *Plant Physiol. Biochem*. 2018;130:399-407. DOI 10.1016/j.plaphy.2018.07.025.

Souza V.L., de Almeida A.A., Lima S.G., de M. Cascardo J.C., da C. Silva D., Mangabeira P.A., Gomes F.P. Morphophysiological responses and programmed cell death induced by cadmium in *Genipa americana* L. (Rubiaceae). *Biometals*. 2011;24(1):59-71. DOI 10.1007/s10534-010-9374-5.

Subramanian A., Tamayo P., Mootha V.K., Mukherjee S., Ebert B.L., Gillette M.A., Paulovich A., Pomeroy S.L., Golub T.R., Lander E.S., Mesirov J.P. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA*. 2005;102(43):15545-15550. DOI 10.1073/pnas.0506580102.

Sun C., Dong Z., Zhao L., Ren Y., Zhang N., Chen F. The Wheat 660K SNP array demonstrates great potential for marker-assisted selection in polyploid wheat. *Plant Biotechnol. J*. 2020;18(6):1354-1360. DOI 10.1111/pbi.13361.

Takahashi F., Kuromori T., Sato H., Shinozaki K. Regulatory gene networks in drought stress responses and resistance in plants. *Adv. Exp. Med. Biol*. 2018;1081:189-214. DOI 10.1007/978-981-13-1244-1_11.

Tranchevent L.C., Ardeshirdavani A., ElShal S., Alcaide D., Aerts J., Auboeuf D., Moreau Y. Candidate gene prioritization with Endeavour. *Nucleic Acids Res*. 2016;44(W1):W117-W121. DOI 10.1093/nar/gkw365.

van Dongen S., Abreu-Goodger C. Using MCL to extract clusters from networks. *Methods Mol. Biol*. 2012;804:281-295. DOI 10.1007/978-1-61779-361-5_15.

Voss-Fels K.P., Cooper M., Hayes B.J. Accelerating crop genetic gains with genomic selection. *Theor. Appl. Genet*. 2019;132(3):669-686. DOI 10.1007/s00122-018-3270-8.

Wang J.Z., Du Z., Payattakool R., Yu P.S., Chen C.F. A new method to measure the semantic similarity of GO terms. *Bioinformatics*. 2007; 23(10):1274-1281. DOI 10.1093/bioinformatics/btm087.

Wang L., Jin X., Li Q., Wang X., Li Z., Wu X. Comparative proteomics reveals that phosphorylation of β carbonic anhydrase 1 might be important for adaptation to drought stress in *Brassica napus*. *Sci. Rep*. 2016;6:39024. DOI 10.1038/srep39024.

Wijewardana C., Reddy K.R., Krutz L.J., Gao W., Bellaloui N. Drought stress has transgenerational effects on soybean seed germination and seedling vigor. *PLoS One*. 2019;14(9):e0214977. DOI 10.1371/journal.pone.0214977.

Wu J., Jiang Y., Liang Y., Chen L., Chen W., Cheng B. Expression of the maize MYB transcription factor ZmMYB3R enhances drought and salt stress tolerance in transgenic plants. *Plant Physiol. Biochem*. 2019;137:179-188. DOI 10.1016/j.plaphy.2019.02.010.

ORCID ID
P.S. Demenkov orcid.org/0000-0001-9433-8341

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

Comprehensive analysis of the 5-HTTLPR allelic polymorphism effect on behavioral and neurophysiological indicators of executive control in people from different ethnic groups in Siberia

A.N. Savostyanov[1, 2, 5] ✉, D.V. Bazovkina[1], S.A. Lashin[1, 5], S.S. Tamozhnikov[2], A.E. Saprygin[1, 2], T.N. Astakhova[5],
U.N. Kavai-ool[4], N.V. Borisova[3], A.G. Karpova[3]

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Scientific-Research Institute of Neuroscience and Medicine, Novosibirsk, Russia
[3] M.K. Ammosov North-Eastern Federal University, Yakutsk, Sakha Republic, Russia
[4] Tuvan Scientific Center, Kyzyl, Tyva Republic, Russia
[5] Novosibirsk State University, Novosibirsk, Russia
✉ a-sav@mail.ru

Abstract. The allelic polymorphism of the serotonin transporter's gene 5-HTTLPR is considered as one of the factors determining an individual genetic predisposition to the development of a wide range of affective disorders, including depression. Many studies have shown that the climatic and social conditions of people's life can have a significant impact on the connections of 5-HTTLPR with the risk of depression. The stop-signal paradigm (SSP) is an experimental method allowing evaluating an individual ability to the self-control of behavior in a changing environment. In the SSP experiment, a subject should either press one of several buttons quickly after the appearance of the target stimuli or suppress the already started movement if an inhibitory signal follows the target stimulus. The aim of this study is a research of associations between the allelic the 5-HTTLPR polymorphism and the individual scores of the personal anxiety level, as well as the behavioral and neurophysiological indicators of the ability to self-control over motor reactions in the SSP. The study was conducted among people from three ethno-regional groups: healthy Caucasoids from Novosibirsk, the Mongoloid groups of the indigenous population of the Tuva Republic and Sakha Republic (Yakutia). Genetic, ethnographic, and psychological influences on an individual's ability to control motor responses were compared. The amplitude of the premotor peak of the evoked brain potential was used as a neurophysiological marker of the person's readiness to the execution of target-directed activity. It was revealed that the frequency of the *S*-allele polymorphism 5-HTTLPR was significantly higher for both mongoloid groups compared to the Caucasoids. The *S/S* genotype was associated with an increased level of personal anxiety and at the same time with a better ability to the self-control of behavior in the SSP experiment. Anxiety level, participants' sex, ethnicity, and allelic polymorphism 5-HTTLPR had a statistically significant effect on the amplitude of the premotor readiness potential recorded under the SSP conditions in the frontal and parietal-occipital cortical regions. Our data support the hypothesis that the *S/S* genotype of the 5-HTTLPR polymorphism may be associated with more success in adapting to the climatic conditions connected with high life risk in comparison to *L/L* and *L/S* genotypes.
Key words: serotonin transporter; 5-HTTLPR polymorphism; personal anxiety; stop-signal paradigm; premotor evoked potential.

Комплексный анализ влияния аллельного полиморфизма 5-HTTLPR на поведенческие и нейрофизиологические показатели исполнительного контроля у людей из разных этнических групп в Сибири

А.Н. Савостьянов[1, 2, 5] ✉, Д.В. Базовкина[1], С.А. Лашин[1, 5], С.С. Таможников[2], А.Е. Сапрыгин[1, 2], Т.Н. Астахова[5],
У.Н. Кавай-оол[4], Н.В. Борисова[3], А.Г. Карпова[3]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Научно-исследовательский институт нейронаук и медицины, Новосибирск, Россия
[3] Северо-Восточный федеральный университет им. М.К. Аммосова, Якутск, Россия
[4] Тувинский научный центр, Кызыл, Республика Тыва, Россия
[5] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
✉ a-sav@mail.ru

Аннотация. Аллельный полиморфизм гена транспортера серотонина 5-HTTLPR рассматривается как один из факторов, определяющих генетическую предрасположенность человека к развитию широкого ряда аффективных нарушений, включая депрессию. Во многих исследованиях было показано, что климатические и социальные условия жизни людей могут оказать существенное влияние на взаимосвязь 5-HTTLPR с риском депрессии. Стоп-сигнал парадигма (ССП) – это экспериментальный метод, позволяющий оценить способности человека к контролю своего поведения в условиях изменяющейся внешней среды. В ССП эксперименте испытуемый должен либо быстро нажимать на одну из нескольких кнопок после появления целевых стимулов, либо подавлять уже начатое движение, если после целевого стимула следует запрещающий сигнал. В представленной работе исследуются ассоциации полиморфизма 5-HTTLPR с оценками уровня личностной тревожности, а также поведенческими и нейрофизиологическими (электроэнцефалограмма, ЭЭГ) показателями способности к контролю двигательных реакций в ССП у представителей трех этно-региональных групп – здоровых европеоидов из г. Новосибирска, монголоидных групп коренного населения Республики Тыва и Республики Саха (Якутия). Целью исследования было сопоставление генетических, этнографических и психологических эффектов на способность индивида к контролю моторных ответов. Амплитуда премоторного пика вызванного потенциала использована в качестве нейрофизиологического маркера готовности человека к выполнению целенаправленных действий. Выявлено, что частота встречаемости аллеля *S* полиморфизма 5-HTTLPR достоверно выше для обеих монголоидных групп в сравнении с европеоидной. Генотип *S/S* ассоциирован с повышенным уровнем личностной тревожности и одновременно с лучшей способностью к контролю движений в условиях ССП эксперимента. Уровень тревожности, пол испытуемых, этническая принадлежность и аллельный полиморфизм 5-HTTLPR оказывают статистически достоверное влияние на амплитуду премоторного потенциала готовности, регистрируемого в условиях ССП в лобных и теменно-затылочных областях коры. Наши данные подтверждают гипотезу, что генотип *S/S* полиморфизма 5-HTTLPR может быть ассоциирован с большей успешностью адаптации к климатическим условиям, связанным с высоким риском для жизни, в сравнении с генотипами *L/L* и *L/S*.

Ключевые слова: транспортер серотонина; полиморфизм 5-HTTLPR; личностная тревожность; парадигма стоп-сигнал; премоторный вызванный потенциал.

Introduction

The subject of psychological genetics is the identification of molecular markers associated with psychological characteristics of healthy people and predisposition to the onset of psychiatric and neurological diseases (Eysenck, 1990; Miller, Lynam, 2003). Serotonin neurotransmitter transporter (5-HTT) polymorphism is one of the most extensively studied molecular markers of predisposition to a wide range of mental disorders (Lesch et al., 1997; Arango et al., 2003). In humans, the serotonin transporter is encoded by the *SLC6A4* gene located on chromosome 17 (Gelernter et al., 1995). The promoter region of the serotonin transporter gene (5-HTTLPR) contains 16 tandem repeats about 20 bp units. Polymorphism 5-HTTLPR is presented by two allelic variants: long variant contains 16 repeats (*L* allele) and short variant contains 14 repeats (*S* allele). It is known, the *S* allele is associated with a reduced efficiency of the transport function of this protein (Lesch et al., 1996). In addition, the long allele contains an A/G single nucleotide polymorphism, with the *Lg* allele functionally similar to the *S* allele (Hu et al., 2005).

It was revealed, that the *S* allele increased the risk of depression in people who have experienced life stress (Caspi et al., 2003). However, the relationship of this allele with stress and depression is still a topic of active discussion (Munafò et al., 2009; Risch et al., 2009; Knyazev et al., 2017). In some works, this connection was confirmed, while in others, on the contrary, it was rejected. There are also studies in which the *S* allele is associated not only with negative qualities. A number of studies have shown that people with the *S* allele performed better than carriers of the *L/L* genotype when solving a wide range of cognitive tasks (Homberg, Lesch, 2011). *S* allele carriers showed greater success on the divergent thoughts test (Volf et al., 2009), demonstrated better visual planning abilities (Roiser et al., 2006) and better attention to

differences in the probability of winning, showed more intensive error handling (Althaus et al., 2009), high performance in card sorting tests in the Wisconsin problem (Borg et al., 2009), and higher IQ scores (Volf et al., 2015). All of these results can be explained by the "differential susceptibility" hypothesis, according to which the so-called "risk alleles" may have a higher sensitivity to environmental challenges. What is harmful in some circumstances may be beneficial in other life situations.

The association of 5-HTTLPR polymorphism with personality traits in healthy individuals is also a topic of intense debates (Hariri et al., 2005; Dannlowski et al., 2008; van der Meer et al., 2016). We have previously shown that associations between different 5-HTTLPR alleles and personality traits in healthy subjects were significantly modulated by ethnic and cultural affiliation of people (Savostyanov et al., 2015). In general, from a review of the scientific literature, we can conclude that when searching for markers of mental illness or personality traits, it is impossible to limit ourselves only to the molecular-genetic level of the description of the nervous system. It is also necessary to take into account the characteristics of human behavior in changing environmental conditions, including their behavior in society.

When studying complex multifactorial associations between genotype and behavior, it is paramount to choose a proper experimental model that will allow to conduct the study. One of these models is the stop-signal paradigm (SSP, Band et al., 2003). SSP is designed as a method to assess a person's ability to control their own actions in a changing environment with an acute shortage of time for decision-making. The essence of the method is that a person in a random order either performs quick targeted actions in response to the appearance of target events, or suppresses the already started activity if the target event is followed by a prohibitory signal. Previously,

А.Н. Савостьянов, Д.В. Базовкина, С.А. Лашин …
У.Н. Кавай-оол, Н.В. Борисова, А.Г. Карпова

Взаимосвязь полиморфизма 5-HTTLPR с нейрофизиологи-
ческими реакциями в условиях парадигмы стоп-сигнал

2021
25•5

we showed that 5-HTTLPR polymorphism is associated with indicators of motor control under ERP conditions (Karpova et al., 2017). Carriers of the *S* allele showed significantly better scores on control over movements compared to people with the *L/L* genotype. In another our study, it was shown that SSP can be used to study endophenotypic differences between subjects (Savostyanov et al., 2009). Comparison of EEG reactions under ERP conditions revealed significant differences between people with different levels of anxiety regarding the dynamics of neurophysiological processes associated with control over movements. Thus, on the basis of preliminary studies, it can be concluded that the analysis of behavioral and neurophysiological indicators recorded in the SSP experiments makes it possible to reveal their dependence simultaneously on genetic differences of people, the level of their personal anxiety and other indicators including gender and age of participants.

The current article presents the multivariate analysis of the associations between the 5-HTTLPR allelic polymorphism, the level of personal anxiety, and behavioral and neurophysiological indicators reflecting brain activity under motor control conditions. One of the well-studied neurophysiological markers of motor control detected in ERP is the premotor readiness potential (the so-called Bereitschaftspotential or readiness potential). This electrographic brain response is detected on the EEG in time intervals immediately preceding the execution of the planned actions. The amplitude and cortical topography of the premotor potential reflects a person's readiness to perform an action.

We carried out a comparative study in three groups of healthy young people, differing in ethnicity and region of residence. A large group of Caucasians (mainly Russians), permanently residing in a large industrial city (Novosibirsk), was examined. In addition, two independent groups of Siberian Mongoloids were examined – Tuvans living in the Republic of Tuva, Kyzyl, and a group of Mongoloids, consisting mainly of Yakuts and Evenks, living in the Republic of Sakha (Yakutia). Our hypothesis assumed that the influence of the level of anxiety and 5-HTTLPR polymorphism affects the amplitude of the cerebral premotor potential under conditions of movement activation in the ERP.

Materials and methods
Subjects. A total of 294 young, healthy subjects, in average 23.4 ± 3.2 y. o., 117 men, 177 women, mainly students of various universities participated in the survey. 121 of them were permanently living in Novosibirsk and considered themselves to be one of the Caucasian ethnic groups (approximately 72 % – Russians, 10 % – Ukrainians, 7 % – Tatars, 4 % – Jews, 7 % – the rest). 94 people were examined at the Tuva State University, Kyzyl, the Republic of Tuva. All subjects from this group referred to themselves as ethnic Tuvans. Another 79 people were examined at the North-Eastern Federal University in Yakutsk. In this group, approximately 80 % identified themselves as Yakuts (Sakha), 15 % – as Evenks and about 5 % – as Yukagirs.

Before the examination, all subjects signed an informed consent to participate in the examination and filled out a questionnaire in which they noted the presence of various

diseases. The exclusion criteria were the presence of mental or neurophysiological diseases, as well as brain injuries within three years before the examination. In addition, participants who indicated in the questionnaire that they use drugs or psychoactive substances were excluded from the experimental sample. Also, pregnant women and women in the first four days of the menstrual cycle were excluded from the study. The examination protocol met the requirements of the Declaration of Helsinki on Biomedical Ethics and was approved by the Ethics Committee of the Research Institute of Physiology and Fundamental Medicine.

Before the start of the examination, all subjects completed the Russian version of the psychological questionnaire of C. Spielberger to assess the level of situational and personal anxiety. In addition, buccal epithelium samples were taken from all participants to determine the 5-HTTLPR allelic polymorphism.

Determination of 5-HTTLPR polymorphism. Genomic DNA was isolated from buccal epithelium samples by using a DNA isolation kit (Biosilica, Russia). The genotypes of the subjects (*L/L*, *L/S*, *S/S*) were determined by PCR using specific primers F 5′-ggcgtgcgctgtgaattgc-3′ and R 5′-gagga ctgagctgacaaccac-3′ (Lesch et., 1996) and the genomic DNA of the subjects as a template. PCR products were separated by electrophoresis in 3 % agarose gel stained with ethidium bromide. The *S* and *L* allele sizes for 5-HTTLPR were 489 and 529 bp, respectively. For determination of polymorphism of *La*/*Lg* the products of amplification split during 3 h by MspI endonuclease. The resulting fragments were separated and visualized on 3 % agarose gel stained with ethidium bromide. Cleavage product sizes for the *La* allele were 340, 127 and 62 bp, while for the *Lg* allele were 174, 166, 127 and 62 bp. *Lg* allele was included in the *S* allele group because the two alleles are functionally similar (Hu et al., 2005).

Experimental method "Stop-Signal Paradigm". In the experiment, the technique proposed by Band et al. (2003) and adapted for EEG recording by Savostyanov and co-authors (2009) was used. The experiment was designed in the form of a computer game "Hunt". The participant was randomly presented with 130 visual stimuli (50 % tanks and 50 % deers). Stimuli with a size of 5×7 cm appeared in the center of the computer screen, located about 1 meter from the subject's head. The stimulus was shown on the screen for 0.75 seconds; the interval between stimuli varied randomly within 3–5 seconds. The subject's task was to press the left button as quickly as possible after the appearance of a deer (which corresponded to a shot from a crossbow) and the right button after the appearance of a tank (which corresponded to a shot from an anti-tank weapon). If the participant managed to press the button correctly before the image disappeared from the screen, they were awarded game points. If the participant chose the button incorrectly or pressed it after 0.75 seconds, their game score was decreased. In 35 % of cases, after the appearance of the target stimulus, a stop signal was presented (a red square in the center of the figure with the inscription "Stop"). The interval between the appearance of the target light and the stop light ranged from 0.25 to 0.75 seconds. In the event of a stop light, the participant had to interrupt the movement that had already started. If the participant stopped moving, their

score did not change. If the subject pressed the button after the stop-light appeared, their score would decrease. Accordingly, all tasks of the SSP were subdivided into the "Go" condition, when it was necessary to press the button (100 tasks out of 130), and the "Stop" condition (30 tasks out of 100), when it was necessary to suppress the movement. The sequence of tasks from both conditions was randomized for all subjects.

EEG registration. EEG was recorded using an actiChamp biopotential amplifier from Brain Products, Germany, with a bandwidth of 0.3–100 Hz, a signal sampling rate of 1000 Hz. The electrodes were placed according to the international scheme 10–5 % with grounding at AFz and reference at Cz. For participants from Novosibirsk, EEG was recorded using 128 channels, for participants from Kyzyl and Yakutsk, EEG was recorded using 64 channels. Additionally, an electrooculogram (VEOG, HEOG) and an ECG were recorded for all participants.

EEG preprocessing and calculation of event-related potentials. To assess changes in signal amplitude associated with the appearance of a target stimulus, event-related potentials (ERPs) were calculated using the ERPLAB software package (https://erpinfo.org/erplab). EEG fragments containing muscle artifacts that could not be corrected were excluded from the analysis. For each participant, 80–90 EEG fragments containing the mark of the target event occurrence in the "Go" condition were selected. The time interval from –1.5 to +3.0 seconds before and after the appearance of the target signal was selected for analysis. The time interval from –1.5 to –0.5 seconds before the appearance of the target signal was used for baseline correction.

The EEG was preliminarily filtered in the range of 1–40 Hz using eleptic filters. Following the recommendations of Delorme and Makeig (2004), the re-reference procedures for the averaged referent and subtraction of the baseline were performed during data pre-processing. Independent components analysis (ICA) was performed to exclude oculomotor and blink artifacts. Initially, component weights were calculated individually for each participant. Components that corresponded to ocular artifacts were identified by visual inspection in conjunction with EOG and ECG. Components with artifacts were removed during EEG pre-processing.

Event-related potentials (ERPs), which were calculated in the ERPLAB software package, were used to assess changes in brain activity associated with motor tasks. After removing the artifacts, we calculated the ERP values using the ERPLAB program independently for each EEG channel and each participant. The results were filtered with a 15 Hz cutoff filter. After that, the average amplitude for the premotor readiness peak was calculated in a time window from 350 to 600 ms after the appearance of the target stimulus.

Statistical analysis of results. One Way analysis of variance (ANOVA) with one target and three fixed variables was used to statistically assess the significance of the obtained behavioral results. As a target variable, one of three indicators was chosen independently of each other – the level of personal anxiety, time (in milliseconds) and quality (percentage of correct decisions) of the task in the "Go" condition. The fixed variables were simultaneously stated: "gender" (men or women), "ethno-regional group" (Novosibirsk Caucasians, Tuvinians, Yakut groups), "5-HTTLPR polymorphism" (*L/L*, *L/S* or *S/S* genotypes).

To process the amplitude of the event-related potentials, the data obtained for each EEG channel were averaged over 11 groups of electrodes corresponding to the left frontal, medial frontal, right frontal, left temporal, right temporal, left central, medial central, right central, left parieto-occipital, medial parieto-occipital and the right parieto-occipital regions of the cerebral cortex. After that, multivariate analysis of variance ANOVA was applied with repeated measurements and Greenhouse–Geisser sphericity correction with the factors "cortical regions" (11 sections of the cortex), "sex" (men or women), "ethno-regional group" (Novosibirsk Caucasians, Tuvinians, Yakut group), "5-HTTLPR polymorphism" (*L/L*, *L/S* or *S/S* genotypes), "anxiety level" (people with relatively low or relatively high anxiety, the sample was divided according to the median value of anxiety scores).

Results

Prevalence of 5-HTTLPR alleles in regional groups
The prevalence of various alleles of 5-HTTLPR polymorphism among subjects from the cities of Novosibirsk, Kyzyl and Yakutsk is presented in the Table. Among the representatives of the Caucasian sample from Novosibirsk, the frequency of occurrence of the *L* allele (56.6 %) exceeded the frequency of the *S* allele (43.4 %), while in both Mongoloid samples, on the contrary, the *S* allele was found significantly more often than the *L* allele. When comparing the three groups, there were four degrees of freedom, the boundary value of χ^2 for three degrees of freedom is more than 9.50; between the examined groups $\chi^2 = 23.55$, the significance of intergroup differences when comparing the Novosibirsk Caucasoid and two Mongoloid samples was $p < 0.01$ (see the Table).

It was important to note that in the sample of Yakuts and Evenks, the frequency of occurrence of the *S* allele (84.8 %) was higher than in the sample of Tuvans (73.4 %). When comparing the Yakut and Tuvan groups, there were two degrees of freedom, the boundary value of χ^2 for two degrees of freedom was more than 5.99; between the two examined groups $\chi^2 = 8.30$, the significance of intergroup differences when comparing two Mongoloid (i. e. Yakut and Tuvan) samples among themselves was $p < 0.03$. Thus, the two Mongoloid samples differed from each other in the prevalence of alleles of the 5-HTTLPR polymorphism, although not as contrastingly as they both differed from the Caucasian sample.

Association of effects of group, gender and 5-HTTLPR polymorphism with the level of personal anxiety
The level of personal anxiety assessed using the Spielberger questionnaire significantly differed between participants of different sex. Anxiety was significantly higher for women (mean 26.8 ± 0.8) than for men (mean 23.8 ± 0.9), $F_{(1, 294)} = 9.05$; $p = 0.003$; $\eta^2 = 0.050$ (Fig. 1, *a*). Also, the main effect of the "group" factor was revealed when comparing the assessments of anxiety for representatives of the subjects from different regions, $F_{(2, 294)} = 3.91$; $p = 0.021$; $\eta^2 = 0.053$ (see Fig. 1, *b*). The minimum anxiety was found in the Tuvin group

А.Н. Савостьянов, Д.В. Базовкина, С.А. Лашин …
У.Н. Кавай-оол, Н.В. Борисова, А.Г. Карпова

Взаимосвязь полиморфизма 5-HTTLPR с нейрофизиологи-
ческими реакциями в условиях парадигмы стоп-сигнал

2021
25•5

Frequency of occurrence of *L* and *S* alleles of 5-HTTPLR polymorphism in the regional groups of the subjects

Experiment region	Number of people with genotype			Total number of subjects	Probability of occurrence of allele in the group, %	
	L/L	*L/S*	*S/S*		*L*	*S*
Novosibirsk	41	55	25	121	56.6	43.4
The Republic of Tuva	6	38	50	94	26.6	73.4
The Republic of Sakha (Yakutia)	4	16	59	79	15.2	84.8
For all regions	51	109	134	294	35.8	64.1

N o t e. For comparison of the Caucasian and two Mongoloid (Yakut and Tuvan) groups $\chi^2 = 23.55$, $p < 0.01$.
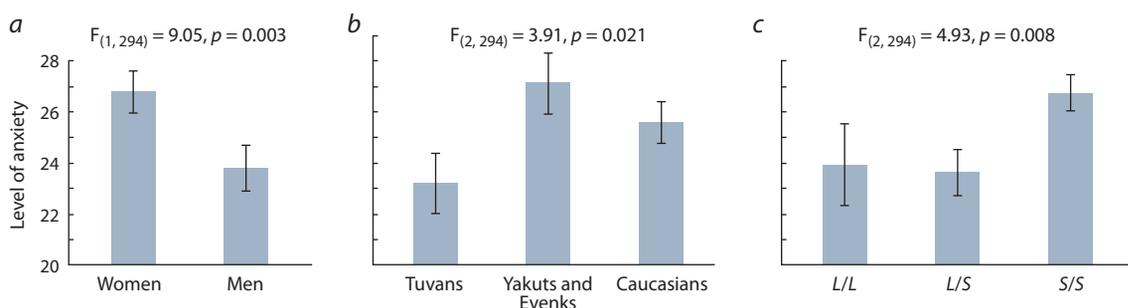
Fig. 1. Relationship between the level of personal anxiety and gender (*a*), ethnic and regional affiliation of the participants (*b*), and 5-HTTLPR polymorphism (*c*).

(mean 23.2 ± 1.2), and the maximum in the Yakut group (mean 27.1 ± 1.2), while in the Caucasian group, the average level of anxiety was intermediate between the two Mongoloid groups (25.6 ± 0.8).

The main effect of 5-HTTLPR polymorphism on the level of personal anxiety was statistically significant, $F_{(2, 294)} = 4.93$; $p = 0.008$; $\eta^2 = 0.032$ (see Fig. 1, *c*). The level of anxiety for carriers of genotypes *L/L* (mean 23.9 ± 1.6) and *L/S* (23.6 ± 0.9) was significantly lower than for people with the genotype *S/S* (26.7 ± 0.7). Post-hoc comparisons did not reveal significant differences in the level of anxiety between people with the *L/L* and *L/S* genotypes ($p > 0.5$), but revealed a significant difference between people with the *S/S* genotype and carriers of two other genotypes ($p < 0.03$).

The factors of gender, group and 5-HTTLPR polymorphism on the level of anxiety did not significantly interact with each other. Calculation of the effect of each of these factors under the control of other factors did not lead to the disappearance of the significance of the effects, although the significance of the 5-HTTLPR factor in this case slightly decreased ($p = 0.008$ excluding group and sex *vs* $p = 0.035$ under the control of group and sex). Thus, the influence of these three factors on the level of anxiety can be considered as independent of each other.

Association of the effects of group, gender, personality anxiety, and 5-HTTLPR polymorphism with behavioral indicators in the stop-signal paradigm

For indicators assessing a person's ability to suppress movements after stop-signal onset, no significant effects of gender, group or genotype, as well as their interactions, were identified.

A significant effect of the group was revealed for the reaction time in the "Go" condition, $F_{(2, 276)} = 3.66$; $p = 0.052$; $\eta^2 = 0.013$. Tuvans (mean time 561 ± 3 ms) and Yakuts (562 ± 4 ms) showed faster reaction time in comparison with Caucasians (569 ± 3 ms). Post-hoc comparisons did not reveal significant differences in reaction time between the Tuvan and Yakut groups ($p > 0.7$), but revealed differences between Caucasians from both other groups ($p < 0.05$). Also, for the response time, a significant effect of gender was revealed, $F_{(1, 276)} = 3.72$; $p = 0.055$; $\eta^2 = 0.013$. The average reaction time was lower for men (561 ± 3 ms) than for women (568 ± 3 ms). The effect of 5-HTTLPR polymorphism or its interaction with other effects for the response time was statistically not significant.

For the indicator of the quality of performance of tasks in the condition "Go", a significant main effect of the factor "gender" was revealed, $F_{(1, 276)} = 3.81$; $p = 0.052$; $\eta^2 = 0.014$ (Fig. 2, *a*). Men performed this task with better average quality (86.3 ± 0.8 %) than women (84.2 ± 0.8 %). Also, for the quality indicator, the main effect of the "group" factor was significant, $F_{(2, 276)} = 4.55$; $p = 0.011$; $\eta^2 = 0.032$ (see Fig. 2, *b*). Tuvans (84.3 ± 0.9 %) and Caucasians (84.0 ± 0.9 %) showed the same average quality of task performance, while the average quality of task performance in the Yakut group (87.7 ± 1.0 %) was significantly higher than in both other groups. The main effect of 5-HTTLPR polymorphism, calculated without control for other factors, was statistically marginal ($p = 0.090$). However, when calculating the effect of polymorphism under the control of the level of personal anxiety, it became significant, $F_{(2, 276)} = 3.03$; $p = 0.050$; $\eta^2 = 0.019$ (see Fig. 2, *c*). People with the *S/S* genotype showed the best average quality of solving motor tasks (86.7 ± 0.8 %), in comparison with carriers
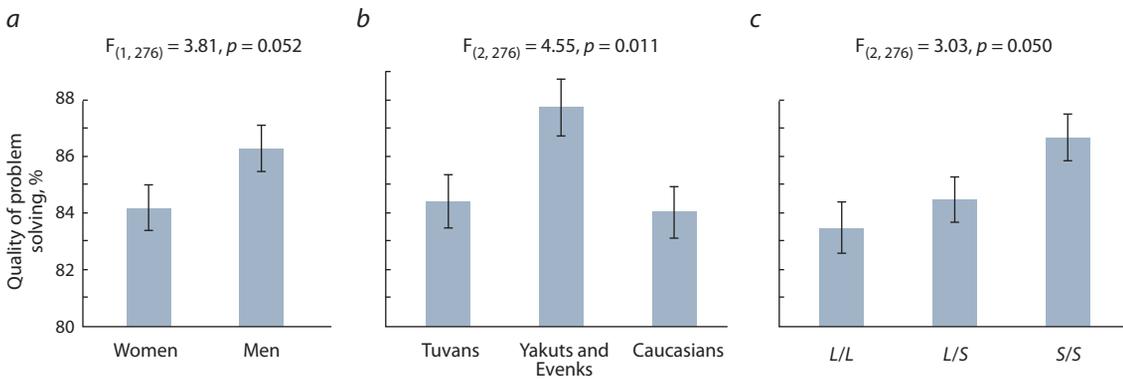
Fig. 2. The relationship between the quality of problem solving in the "Go" condition of the stop-signal paradigm with gender (*a*),
ethnic and regional affiliation of participants (*b*), and 5-HTTLPR polymorphism (*c*).

of genotypes *L/S* (84.5±0.8 %) and *L/L* (83.5±0.9 %). The effect of the level of anxiety on indicators of speed or quality of solving motor tasks was not revealed. The factors of gender, group and polymorphism of the serotonin transporter did not significantly interact with each other.

In general, based on the results of the analysis of the effects of factors of gender, regional-ethnic group and 5-HTTLPR polymorphism on psychological and behavioral indicators, it can be concluded that all three selected factors affect both the level of anxiety and the indicators of motor control. However, their effects did not significantly affect each other. It can also be noted that, although 5-HTTLPR simultaneously affected both the level of anxiety and the quality of motor control, none of the indicators of motor control were directly dependent on the level of personal anxiety.

Association of the effects of group, gender, personality anxiety and 5-HTTLPR polymorphism with the amplitude of the premotor event-related potentials in the stop-signal paradigm

The amplitude-time graph of the event-related potentials in the left motor area and the topographic amplitude distribution of the premotor potential are shown in Fig. 3. Initially, the influence of various factors on the amplitude of the premotor event-related potentials was simultaneously assessed for all 11 cortical regions. With this method of assessment, only significant effects of the region were revealed, $F_{(10, 2920)} = 300.05$; $p < 0.0001$, and group $F_{(2, 294)} = 4.30$; $p = 0.014$. The premotor potential had a negative amplitude in the frontal and temporal regions of the cortex and a positive amplitude in the central and parieto-occipital regions. The amplitude in the left ($r = -0.18$; $p = 0.003$) and right ($r = -0.15$; $p = 0.011$) frontal lobes negatively correlated with the quality of task solution under "Go" conditions, and in the medial central ($r = 0.17$; $p = 0.005$) and the medial parieto-occipital ($r = 0.14$; $p = 0.024$) areas, these correlations were positive. If we take into account that the selected peak had a negative polarity for the frontal areas and positive for the central and parieto-occipital areas, we can conclude that its large amplitude in magnitude corresponded to the best quality of tasks in all areas of the cortex.

Interactions of factors "region" to "group", $F_{(20, 2940)} = 21.28$; $p < 0.0001$; $\eta^2 = 0.195$, "region" to "5-HTTLPR polymorphism", $F_{(20, 2940)} = 5.81$; $p < 0.0001$; $\eta^2 = 0.046$, "region" to "level of anxiety", $F_{(10, 2950)} = 2.38$; $p = 0.008$; $\eta^2 = 0.049$, and "region" to "sex", $F_{(10, 2920)} = 3.99$; $p = 0.011$; $\eta^2 = 0.061$ were statistically highly significant. The effects of all factors appeared only in the frontal and occipital-parietal regions of the cortex and did not affect other regions. In addition, since the directivity of the peak amplitude in the anterior (negative) and posterior (positive) regions was different, the statistical analysis was performed separately for the frontal and occipital-parietal regions.

In the frontal parts of the cortex, a significant main effect of the group was revealed, $F_{(2, 294)} = 8.91$; $p < 0.0001$; $\eta^2 = 0.023$. The negative peak amplitude was the maximum modulus in the Yakut group ($-2.0±0.2$ μkV), less in the maximum modulus in the Tuvan group ($-1.7±0.1$ μkV) and the lowest in the maximum modulus in the Caucasian group ($-1.2±0.1$ μkV). Post-hoc comparisons revealed pairwise significant differences in the frontal negative amplitude of the premotor peak between all three groups ($p < 0.01$). The effect of the gender factor for this indicator was also significant, $F_{(1, 292)} = 8.30$; $p = 0.004$; $\eta^2 = 0.030$. The negative amplitude was higher in absolute value for women ($-1.8±0.1$ μkV) than for men ($-1.3±0.1$ μkV). The main effect of 5-HTTLPR polymorphism was significant, $F_{(2, 294)} = 3.90$; $p = 0.021$; $\eta^2 = 0.026$. The amplitude of the negative peak was greater in modulus for carriers of the *S/S* genotype ($-1.8±0.1$ μkV) than for people with the *L/S* ($-1.4±0.1$ μkV) and *L/L* ($-1.3±0.2$ μkV) genotypes. No significant interactions were found for all selected factors. The effect of personality anxiety and its interaction with other effects in the frontal cortex was also not significant.

In the parieto-occipital parts of the cortex, significant effects of sex were revealed, $F_{(1, 292)} = 5.00$; $p = 0.026$; $\eta^2 = 0.033$, group, $F_{(2, 294)} = 40.71$; $p < 0.0001$; $\eta^2 = 0.218$, and 5-HTTLPR polymorphism, $F_{(2, 294)} = 10.29$; $p < 0.0001$; $\eta^2 = 0.065$ (Fig. 4). The positive amplitude in the posterior parts of the cortex was maximum in the Yakut group ($3.1±0.2$ μkV), less in the Tuvan group ($2.8±0.2$ μkV) and the lowest in the Caucasian

А.Н. Савостьянов, Д.В. Базовкина, С.А. Лашин …
У.Н. Кавай-оол, Н.В. Борисова, А.Г. Карпова

Взаимосвязь полиморфизма 5-HTTLPR с нейрофизиологи-
ческими реакциями в условиях парадигмы стоп-сигнал

2021
25•5

a Time-amplitude plot for ERP under C3 electrode (left motor cortical area)

b Cortical topography of pre-motor potential in time-window from 350 till 600 ms afrer task onset
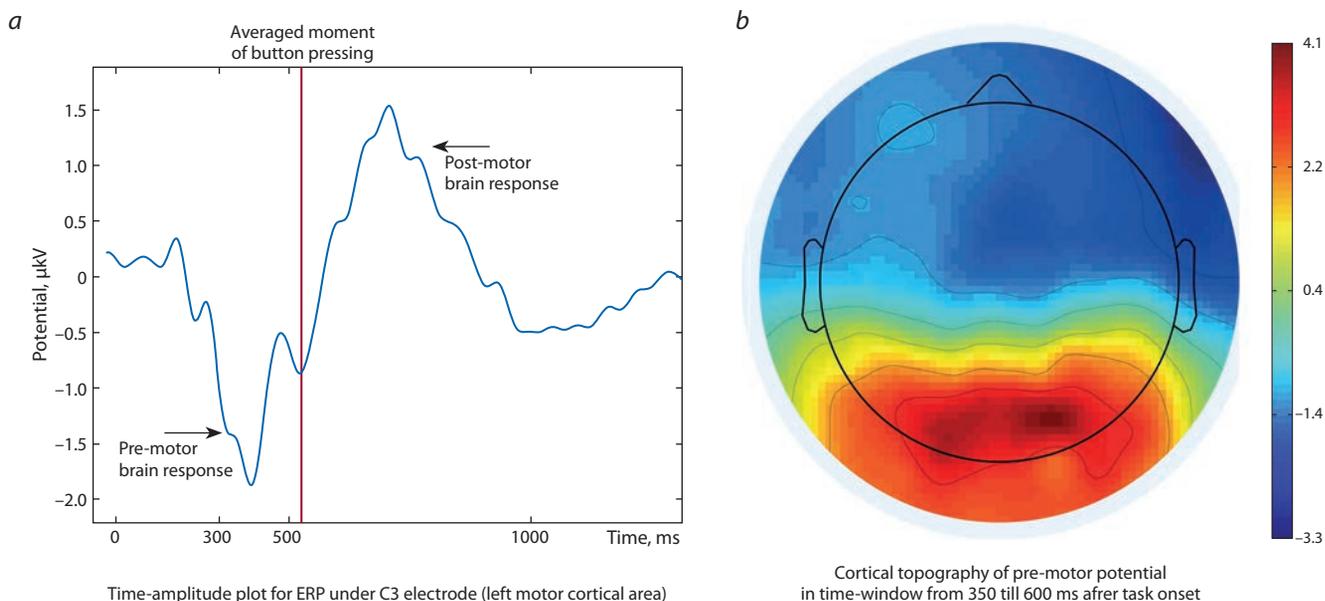
Fig. 3. Amplitude-time graph (*a*) and cortical topography (*b*) for cerebral ERP responses in the "Go" condition of SSP.
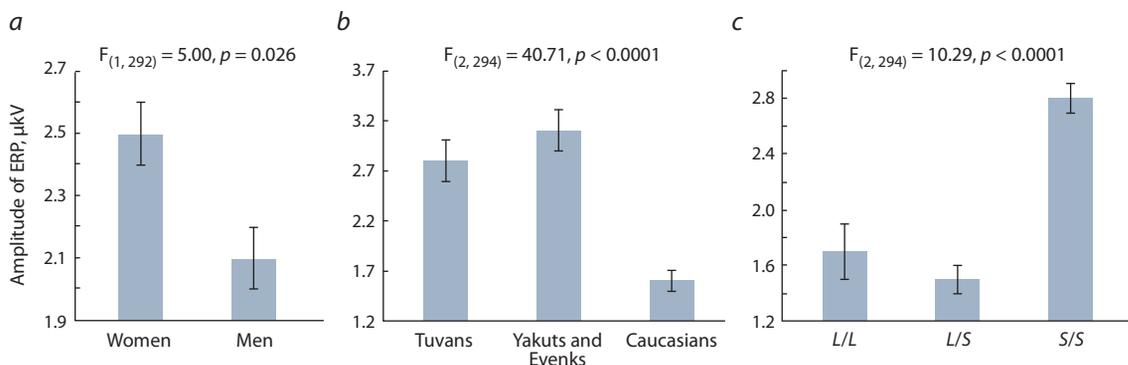
Fig. 4. Correlation of the amplitude of the premotor peak of the event related potential in the "Go" condition of ERP in the parieto-occipital cortex with sex (*a*), ethno-regional affiliation of participants (*b*), and 5-HTTLPR polymorphism (*c*).

group (1.6 ± 0.1 μkV). Post-hoc comparisons revealed pairwise significant differences in the parieto-occipital amplitude of the premotor peak between all three groups ($p < 0.01$). The positive amplitude was higher for women (2.5 ± 0.1 μkV) than for men (2.1 ± 0.1 μkV). The amplitude of the positive peak was greater for carriers of the *S/S* genotype (2.8 ± 0.1 μkV) than for people with genotypes *L/S* (1.5 ± 0.1 μkV) and *L/L* (1.7 ± 0.2 μkV). As for the frontal cortex, in the parieto-occipital cortex, no significant interactions were found for all selected factors.

When calculating the effect of personal anxiety simultaneously for three areas of the parieto-occipital cortex, this effect was insignificant. However, it turned out to be significant separately for the left ($F_{(1, 296)} = 3.93$; $p = 0.048$; $\eta^2 = 0.013$, the amplitude for low-anxious people was lower (1.1 ± 0.1 μkV) than for highly anxious (2.5 ± 0.2 μkV)) and for the right one ($F_{(1, 296)} = 6.19$; $p = 0.013$; $\eta^2 = 0.021$, the amplitude for low-anxious people was lower (1.9 ± 0.2 μkV) than for highly

anxious (2.5 ± 0.2 μkV)) parieto-occipital areas and was not significant for the medial parieto-occipital cortex ($p > 0.5$).

Thus, based on the analysis of the amplitude of the premotor event-related potential, the effects of group (the strongest brain responses in Yakuts, the weakest in Caucasians), gender (the amplitude of responses was greater in women than in men), and 5-HTTLPR polymorphism (the highest responses in carriers of the *S/S* genotype) and anxiety (stronger in high- than in low-anxious) were identified. The amplitude of the premotor potentials correlated with the behavioral indicator of the quality of task solving. However, the effects for all the factors we selected on the premotor potential amplitude did not interact with each other.

Discussion

The frequency of occurrence of *S* and *L* alleles of 5-HTTLPR polymorphism in Caucasoid and Mongoloid samples revealed is consistent with well-known patterns obtained by comparing

different ethnic groups (see Esau et al., 2008; Noskova et al., 2008; Ivanov et al., 2019). It is well known that the *L* allele is more common in Caucasians, while the *S* allele is more common in Mongoloids. Our data are broadly consistent with these results. It can be noted that Caucasians from Novosibirsk are more likely to carry the *S* allele than Caucasians from Europe, the United States and even from the European part of Russia. Among Siberian Mongoloids, the frequency of occurrence of this allele is higher in Yakuts and Evenks in comparison with Tuvans. This indirectly indicates the relationship of the *S* allele with increased adaptability to extreme or sub-extreme climate conditions. Indeed, in the series Western Europe – Western Siberia – Southeast Siberia and Northeast Siberia, extreme climatic conditions for human life are escalating. The frequency of the *S* allele also increases from West to North-East. Although at present we do not have direct data indicating a relationship between 5-HTTLPR polymorphism and mechanisms of adaptation to extreme climates, we can assume the existence of such an association as a working hypothesis.

This hypothesis is supported by the relationship of 5-HTTLPR with behavioral indicators of motor control and the level of anxiety. In the modern psychological literature, anxiety is usually viewed as a negative marker associated with an increased risk of a number of diseases, such as depression, or psychosomatic disorders. However, in conditions accompanied by an increased danger to life, anxiety should serve as an adaptive factor that reduces the risk of human death. It can be noted that the genetic marker of high anxiety (*S* allele) is most common in groups of people living in subpolar or polar climates. The same allele is a marker associated with higher rates of motor control in an experimental model assessing the ability to self-regulate behavior under time pressure. The facts above allowed to formulate the assumption that the *S* allele, which is "bad" from the point of view of the urban environment, may turn out to be a marker of increased ability to adapt in conditions associated with high danger to life.

As mentioned above, the *S* allele is associated with a reduced efficiency of the transport function of this protein. Biochemical studies showed that animals with the *S/S* genotype were characterized by a reduced level of serotonin in the synaptic cleft and a reduced level of functional activity of serotonergic neurons (Lesch et al., 1996). It is also known that the serotonergic system in the regulation of behavior is responsible for the performance of inhibitory control (Munafò et al., 2009). According to literature data obtained in psychiatric patients, it is known that the *S/S* genotype should be associated with the lower ability to delay irrelevant behavioral responses (Malloy-Diniz et al., 2011). However, our data under the "Stop" condition in the SSP did not reveal differences between carriers of different 5-HTTLPR alleles in either behavioral or ERP parameters. It can be assumed that in healthy people with the *S/S* genotype, a decrease in the activity of 5-OHT neurons is associated not with a deterioration in inhibitory control, but with an improvement in the parameters of activation control due to a lower suppression of motor neurons. In this case, a decrease in the concentration of serotonin in the brain due

to a decrease in the transport function of the carrier protein under some external conditions can be considered as a marker of an increased tendency to impulsive-anxious behavior, and under other living conditions – as a mechanism of adaptation to high danger.

It can also be noted that, in addition to the allelic polymorphism chosen, the behavioral indicators of motor control and the level of anxiety are influenced by several other factors independent of each other. Women are, on average, more anxious than men. Caucasians are more anxious than Tuvans, but less anxious in comparison with Yakuts and Evenks. Men are better at motor control tasks than women. Mongoloids perform these tasks on average faster and better than Caucasians. At the same time, no statistical interaction of the factors we selected was found. The occurrence of the *S* allele associated with high anxiety was higher in both Mongoloid groups in comparison with the Caucasoids, but at the same time, Tuvinians are less anxious, and Yakuts and Evenks are more anxious than Caucasians. Thus, both the level of anxiety and the ability to control movements are determined not by one, but by a wide range of factors which interact unclearly.

The attempt made in the framework of this study to find a mechanism for integrating the effects of genetic and environmental factors using the analysis of cerebral event-related potentials has not yet given a completely satisfactory result. We have confirmed the previously established fact that the frontal and parieto-occipital amplitude of the premotor readiness potential correlates with the success of solving motor tasks. We have also shown that the amplitude of this potential depends on the 5-HTTLPR allelic polymorphism. People with the *S/S* genotype show both increased abilities for movement control of behavior and an increased amplitude of premotor cerebral responses to EEG in the frontal and parieto-occipital regions of the cortex. This allows us to conclude that the relationship between 5-HTTLPR polymorphism and the ability for behavioral control is mediated by the electrophysiological activity of the corresponding parts of the cortex. It can also be noted that the effect of anxiety was detected only for the left and right, but not the medial part of the parieto-occipital cortex, while the 5-HTTLPR effect was reliably revealed for six selected parts of the cortex, including the medial parieto-occipital and all frontal regions. On this basis, it can be argued that although both the anxiety effect and the allelic polymorphism of the serotonin transporter are equally manifested in the amplitude of the premotor potential, they have different topography in the cortical areas – the anxiety effect affects a significantly narrower area of the cortex than the allelic polymorphism effect. As for the effects of gender and ethnic-regional affiliation of subjects on the amplitude of brain responses, we have so far failed to separate them from the effect of allelic polymorphism, or to describe the mechanism of their interaction. All three factors affect the amplitude of the premotor potential in the same areas of the cortex and at the same time intervals of the brain reaction. Therefore, at the present stage of the study, we can only conclude that the statistical model we have chosen for the pairwise assessment of the effects of various factors on the neurophysiological

А.Н. Савостьянов, Д.В. Базовкина, С.А. Лашин …
У.Н. Кавай-оол, Н.В. Борисова, А.Г. Карпова

Взаимосвязь полиморфизма 5-HTTLPR с нейрофизиологи-
ческими реакциями в условиях парадигмы стоп-сигнал

2021
25•5

processes that underlie the voluntary control of movements in the stop-signal paradigm did not allow us to achieve the study aim and identify the brain mechanism of their interaction.

Conclusion

Allelic polymorphism 5-HTTLPR is associated simultaneously with an increased level of personal anxiety and with a better ability to control movements in experimental conditions associated with the need to make decisions with a lack of time. It can be hypothesized that the *S* allele of the serotonin transporter is associated with better adaptability to living conditions under conditions of increased danger, which is indirectly confirmed by the frequency of occurrence of this allele in various ethno-regional groups. Analysis of neurophysiological processes recorded by EEG assessment recorded under the stop-signal paradigm showed that both the level of anxiety and 5-HTTLPR polymorphism affect the amplitude of the premotor readiness, but the topography of the effects of anxiety and polymorphism is significantly different.

References

Althaus M., Groen Y., Wijers A.A., Mulder L.J.M., Minderaa R.B., Kema I.P., Dijck J.D.A., Hartman C.A., Hoekstra P.J. Differential effects of 5-HTTLPR and DRD2/ANKK1 polymorphisms on electrocortical measures of error and feedback processing in children. *Clin. Neurophysiol*. 2009;120:93-107. DOI 10.1016/j.clinph.2008.10.012.

Arango V., Huang Y.Y., Underwood M.D., Mann J.J. Genetics of the serotonergic system in suicidal behavior. *J. Psychiatr. Res*. 2003;37: 375-386. DOI 10.1016/s0022-3956(03)00048-7.

Band G.P.H., van der Molen M.W., Logan G.D. Horse-race model simulations of the stop-signal procedure. *Acta Psychol*. 2003;112: 105-142. DOI 10.1016/s0001-6918(02)00079-3.

Borg J., Henningsson S., Saijo T., Inoue M., Bah J., Westberg L., Lundberg J., Jovanovic H., Andree B., Nordstrom A.L., Halldin C., Eriksson E., Farde L. Serotonin transporter genotype is associated with cognitive performance but not regional 5-HT1A receptor binding in humans. *Int. J. Neuropsychopharmacol*. 2009;12:783-792. DOI 10.1017/S1461145708009759.

Caspi A., Sugden K., Moffitt T.E., Taylor A., Craig I.W., Harrington H., McClay J., Mill J., Martin J., Braithwaite A., Poulton R. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science*. 2003;301:386-389. DOI 10.1126/science.1083968.

Dannlowski U., Ohrmann P., Bauer J., Deckert J., Hohoff C., Kugel H., Arolt V., Heindel W., Kersting A., Baune B.T., Suslow T. 5-HTTLPR biases amygdala activity in response to masked facial expressions in major depression. *Neuropsychopharmacology*. 2008;33:418-424. DOI 10.1038/sj.npp.1301411.

Delorme A., Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Meth*. 2004;134(1):9-21. DOI 10.1016/j.jneumeth.2003.10.009.

Esau L., Kaur M., Adonis L., Arieff Z. The 5-HTTLPR polymorphism in South African healthy populations: a global comparison. *J. Neural Transm*. 2008;115(5):755-760. DOI 10.1007/s00702-007-0012-5.

Eysenck H. Biological dimensions of personality. In: Pervin L.A. (Ed.). Handbook of Personality: Theory and Research. New York, Guilford, 1990;244-276.

Gelernter J., Pakstis A.J., Kidd K.K. Linkage mapping of serotonin transporter protein gene SLC6A4 on chromosome 17. *Hum. Genet*. 1995;95:677-680. DOI 10.1007/BF00209486.

Hariri A.R., Drabant E.M., Munoz K.E., Kolachana B.S., Mattay V.S., Egan M.F., Weinberger D.R. A susceptibility gene for affective disorders and the response of the human amygdala. *Arch. Gen. Psychiatry*. 2005;62:146-152. DOI 10.1001/archpsyc.62.2.146.

Homberg J.R., Lesch K.P. Looking on the bright side of serotonin transporter gene variation. *Biol. Psychiatr*. 2011;69:513-519. DOI 10.1016/j.biopsych.2010.09.024.

Hu X.Z., Oroszi G., Chun J., Smith T.L., Goldman D., Schuckit M.A. An expanded evaluation of the relationship of four alleles to the level of response to alcohol and the alcoholism risk. *Alcohol: Clin. Exp. Res*. 2005;29:8-16. DOI 10.1097/01.alc.0000150008.68473.62.

Ivanov R., Zamyatin V., Klimenko A., Matushkin Y., Savostyanov A., Lashin A. Reconstruction and analysis of gene networks of human neurotransmitter systems reveal genes with contentious manifestation for anxiety, depression, and intellectual disabilities. *Genes (Basel)*. 2019;10(9):699. DOI 10.3390/genes10090699.

Karpova A.G., Savostyanov A.N., Bazovkina D.V., Tamozhnikov S.S., Saprygin A.E., Proshina E.A., Borisova N.V., Aftanas L.I. Allelic polymorphism of the serotonin transporter as a formation factor of the behavioral cognitive control in the Yakuts. *Yakutskij Meditsynskyi Zhurnal = Yakut Medical Journal*. 2017;3(59):21-24. (in Russian)

Knyazev G.G., Bazovkina D.V., Savostyanov A.N., Naumenko V.S., Kuznetsova V.B., Proshina E.A. Suppression mediates the effect of 5-HTTLPR by stress interaction on depression. *Scand. J. Psychol*. 2017;58(5):373-378. DOI 10.1111/sjop.12389.

Lesch K.P., Bengel D., Heils A., Sabol S.Z., Greenberg B.D., Petri S., Benjamin J., Muller C.R., Hamer D.H., Murphy D.L. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science*. 1996;274(5292):1527-1531. DOI 10.1126/science.274.5292.1527.

Lesch K.P., Meyer J., Glatz K., Flügge G., Hinney A., Hebebrand J., Klauck S.M., Poustka A., Poustka F., Bengel D., Mössner R., Riederer P., Heils A. The 5-HT transporter gene-linked polymorphic region (5-HTTLPR) in evolutionary perspective: alternative biallelic variation in rhesus monkeys. *J. Neural Transm*. 1997;104(11-12): 1259-1266. DOI 10.1007/BF01294726.

Malloy-Diniz L.F., Neves F.S., Paiva de Moraes P.H., de Marco L.A., Romano-Silva M.A., Krebs M.-O., Correa H. The 5-HTTLPR polymorphism, impulsivity and suicide behavior in euthymic bipolar patients. *J. Affect. Disord*. 2011;133(1-2):221-226. DOI 10.1016/j.jad.2011.03.051.

Miller J.D., Lynam D.R. Psychopathy and the five-factor model of personality: a replication and extension. *J. Pers. Assess*. 2003;81(2): 168-178. DOI 10.1207/S15327752JPA8102_08.

Munafò M.R., Durrant C., Lewis G., Flint J. Gene X environment interactions at the serotonin transporter locus. *Biol. Psychiatry*. 2009; 65:211-219. DOI 10.1016/j.biopsych.2008.06.009.

Noskova T., Pivac N., Nedic G., Kazantseva A., Gaysina D., Faskhutdinova G., Gareeva A., Khalilova Z., Khusnutdinova E., Kovacic D.K., Kovacic Z., Jokic M., Seler D.M. Ethnic differences in the serotonin transporter polymorphism (5-HTTLPR) in several European populations. *Prog. Neuropsychopharmacol. Biol. Psychiatry*. 2008;32(7):1735-1739. DOI 10.1016/j.pnpbp.2008.07.012.

Risch N., Herrell R., Lehner T., Liang K.Y., Eaves L., Hoh J., Griem A., Kovacs M., Ott J., Merikangas K.R. Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression. A meta-analysis. *J. Am. Med. Assoc*. 2009;301:2462-2471. DOI 10.1001/jama.2009.878.

Roiser J.P., Rogers R.D., Cook L.J., Sahakian B.J. The effect of polymorphism at the serotonin transporter gene on decision-making, memory and executive function in ecstasy users and controls. *Psychopharmacology*. 2006;188:213-227. DOI 10.1007/s00213-006-0495-z.

Savostyanov A.N., Naumenko V.S., Sinyakova N.A., L'vova M.N., Levin E.A., Zaleshin M.S., Kavay-ool U.N., Mordvinov V.A., Kolchanov N.A., Aftanas L.I. Association of anxiety level with polymorphic variants of serotonin transporter gene in Russians and Tuvinians. *Russ. J. Genet. Appl. Res.* 2015;5:656-665. DOI 10.1134/S2079059715060155.

Savostyanov A.N., Tsai A.C., Liou M., Levin E.A., Lee J.D., Yurganov A.V., Knyazev G.G. EEG-correlates of trait anxiety in the stop-signal paradigm. *Neurosci. Lett.* 2009;449(2):112-116. DOI 10.1016/j.neulet.2008.10.084.

van der Meer D., Hoekstra P.J., Bralten J., van Donkelaar M., Heslenfeld D.J., Oosterlaan J., Faraone S.V., Franke B., Buitelaar J.K., Hartman C.A. Interplay between stress response genes associated with attention-deficit hyperactivity disorder and brain volume. *Genes Brain Behav.* 2016;15(7):627-636. DOI 10.1111/gbb.12307.

Volf N.V., Belousova L.V., Kulikov A.V. Association between the 5-HTTLPR polymorphism of serotonin transporter gene and EEG in young and postmenopausal women. *Zhurnal Vysshei Nervnoi Deyatelnosti im. I.P. Pavlova = I.P. Pavlov Journal of Higher Nervous Activity.* 2015;65(3):324-332. DOI 10.7868/S0044467715030132. (in Russian)

Volf N.V., Kulikov A.V., Bortsov C.U., Popova N.K. Association of verbal and figural creative achievement with polymorphism in the human serotonin transporter gene. *Neurosci. Lett.* 2009;463:154-157. DOI 10.1016/j.neulet.2009.07.070.

ORCID ID
A.N. Savostyanov orcid.org/0000-0002-3514-2901
S.A. Lashin orcid.org/0000-0003-3138-381X