

Письма

в

ВАВИЛОВСКИЙ ЖУРНАЛ
ГЕНЕТИКИ И СЕЛЕКЦИИ

Letters to Vavilov Journal of Genetics and Breeding

2025
М А Р Т

Генетика животных • Популяционная генетика • Биоразнообразие •
Молекулярная биология • Математическая биология

Том 11
№1

PISMAVAVILOV@BIONET.NSC.RU

ПИСЬМА В ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

Сетевое издание
Листья



ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

Основано в 2015 году
Периодичность четыре выпуска в год
doi 10.18699/letvjgb-2025-11-01

Учредитель

Федеральное государственное бюджетное научное учреждение «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук» (ИЦиГ СО РАН), Новосибирск, Россия

Главный редактор

А.В. Кочетов – академик РАН, д-р биол. наук (ИЦиГ СО РАН, Новосибирск, Россия)

Заместители главного редактора

Н.П. Гончаров – академик РАН, д-р биол. наук (ИЦиГ СО РАН, Новосибирск, Россия)

Е.А. Салина – д-р биол. наук, профессор (ИЦиГ СО РАН, Новосибирск, Россия)

Ответственный секретарь

О.Ю. Шоева – канд. биол. наук (ИЦиГ СО РАН, Новосибирск, Россия)

Редакционная коллегия

З. Акрам – профессор, директор (Кафедра селекции растений и генетики Университета аридного земледелия им. Пира Мехра Али Шаха, Равалпинди, Пакистан)

О.С. Афанасенко – академик РАН, д-р биол. наук, профессор (Всероссийский НИИ защиты растений, Санкт-Петербург, Пушкин, Россия)

О.В. Ваулин – канд. биол. наук (ИЦиГ СО РАН, Новосибирск, Россия)

М.А. Вишнякова – д-р биол. наук, профессор (Федеральный исследовательский центр Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова (ВИР), Санкт-Петербург, Россия)

Т.А. Гавриленко – д-р биол. наук, профессор (Федеральный исследовательский центр Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова (ВИР), Санкт-Петербург, Россия)

Ю.Э. Гербек – канд. биол. наук (Еврейский университет в Иерусалиме, Реховот, Израиль)

И.М. Горобей – д-р с.-х. наук, профессор РАН (СО РАН, Новосибирск, Россия)

Е.И. Гулятьева – д-р биол. наук (Всероссийский НИИ защиты растений, Санкт-Петербург, Пушкин, Россия)

Ш. Демберел – академик Монгольской академии наук, профессор (Институт ветеринарной медицины, Монгольский университет естественных наук, Улан-Батор, Монголия)

Н.И. Дубовец – чл.-кор. НАН Беларуси, д-р биол. наук, доцент (Институт генетики и цитологии НАН Беларуси, Минск, Беларусь)

И.К. Захаров – д-р биол. наук, профессор (ИЦиГ СО РАН, Новосибирск, Россия)

А.П. Карташова – канд. с.-х. наук (Мурманская ГСХОС – филиал ВИР, Мурманск, Россия)

К.В. Крутовский – канд. биол. наук, профессор (Гёттингенский университет им. Георга-Августа, Гёттинген, Германия)

А.М. Кудрявцев – чл.-кор. РАН, д-р биол. наук (Институт общей генетики им. Н.И. Вавилова РАН, Москва, Россия)

С.А. Лашин – д-р биол. наук (ИЦиГ СО РАН, Новосибирск, Россия)

А.Ю. Летягин – д-р мед. наук, профессор (ИЦиГ СО РАН, Новосибирск, Россия)

П.Н. Мальчиков – д-р с.-х. наук (Самарский НИИ сельского хозяйства им. Н.М. Тулайкова, пос. Безенчук, Россия)

Е.А. Орлова – канд. с.-х. наук (СибНИИРС – филиал ИЦиГ СО РАН, р. п. Краснообск, Новосибирская обл., Россия)

А.С. Пилипенко – канд. биол. наук (ИЦиГ СО РАН, Новосибирск, Россия)

Ю.И. Рагино – чл.-кор. РАН, д-р мед. наук, профессор (НИИТПМ – филиал ИЦиГ СО РАН, Новосибирск, Россия)

Т. Ражабов – профессор, директор (Институт агробиотехнологий и продовольственной безопасности, Самаркандский государственный университет им. Шарофа Рашидова, Самарканд, Узбекистан)

И.Д. Рашаль – академик АН Латвии, д-р биол. наук, профессор (Институт биологии Латвийского университета, Саласпилс, Латвия)

Р.Р. Садоян – д-р биол. наук, профессор (Армянский государственный педагогический университет им. Х. Абовяна, Ереван, Армения)

А.А. Соловьев – д-р биол. наук, профессор РАН, профессор (Всероссийский центр карантина растений, Москва, Россия)

Н.А. Сурин – академик РАН, д-р с.-х. наук, профессор (Федеральный исследовательский центр КНЦ СО РАН – обособленное подразделение Красноярский НИИ сельского хозяйства, Красноярск, Россия)

В.А. Трифонов – д-р биол. наук, профессор (Институт молекулярной и клеточной биологии СО РАН, Новосибирск, Россия)

В.С. Фишман – д-р биол. наук (ИЦиГ СО РАН, Новосибирск, Россия)

С.В. Шеховцов – канд. биол. наук (ИЦиГ СО РАН, Новосибирск, Россия)

Online edition
Letters

to **VAVILOV JOURNAL
OF GENETICS AND BREEDING**

Founded in 2015

Four issues annually

doi 10.18699/letvjgb-2025-11-01

Founder

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences (ICG SB RAS), Novosibirsk, Russia

Editor-in-Chief

A.V. Kochetov – Full Member of the Russian Academy of Sciences, Dr. Sci. (Biology) (ICG SB RAS, Novosibirsk, Russia)

Deputy Editors-in-Chief

N.P. Goncharov – Full Member of the Russian Academy of Sciences, Dr. Sci. (Biology) (ICG SB RAS, Novosibirsk, Russia)

E.A. Salina – Professor, Dr. Sci. (Biology) (ICG SB RAS, Novosibirsk, Russia)

Executive Secretary

O.Yu. Shoeva – Cand. Sci. (Biology) (ICG SB RAS, Novosibirsk, Russia)

Editorial board

O.S. Afanasenko – Full Member of the RAS, Professor, Dr. Sci. (Biology) (All-Russia Research Institute for Plant Protection, St. Petersburg, Russia)

Z. Akram – Professor, Director (Department of Plant Breeding and Genetics at the Pir Mehr Ali Shah Arid Agriculture University, Rawalpindi, Pakistan)

S. Demberel – Full Member of the Mongolian Academy of Sciences, Professor (Institute of Veterinary Medicine, Mongolian University of Life Sciences, Ulaanbaatar, Mongolia)

N.I. Dubovets – Corr. Member of the NAS of Belarus, Associate Professor, Dr. Sci. (Biology) (Institute of Genetics and Cytology, NASB, Minsk, Belarus)

V.S. Fishman – Dr. Sci. (Biology) (ICG SB RAS, Novosibirsk, Russia)

T.A. Gavrilenko – Professor, Dr. Sci. (Biology) (N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), St. Petersburg, Russia)

I.M. Gorobei – Dr. Sci. (Biology) (Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia)

E.I. Gulyaeva – Dr. Sci. (Biology) (All-Russia Research Institute for Plant Protection, Saint Petersburg, Pushkin, Russia)

Yu.E. Herbeck – Cand. Sci. (Biology) (The Hebrew University of Jerusalem, Rehovot, Israel)

A.P. Kartashova – Cand. Sci. (Agricul.) (Murmansk State Agricultural Experimental Station – Branch of the VIR, Murmansk, Russia)

K.V. Krutovsky – Professor, Cand. Sci. (Biology) (Georg-August University of Göttingen, Göttingen, Germany)

A.M. Kudryavtsev – Corr. Member of the RAS, Dr. Sci. (Biology) (Vavilov Institute of General Genetics, RAS, Moscow, Russia)

S.A. Lashin – Dr. Sci. (Biology) (ICG SB RAS, Novosibirsk, Russia)

A.Y. Letyagin – Professor, Dr. Sci. (Medicine) (ICG SB RAS, Novosibirsk, Russia)

P.N. Malchikov – Dr. Sci. (Agricul.) (Tulaikov Research Institute of Agriculture, Russian Agricultural Academy, Bezenchuk, Samara oblast, Russia)

E.A. Orlova – Cand. Sci. (Agricul.) (Siberian Research Institute of Plant Production and Breeding – Branch of the ICG SB RAS, Krasnoobsk, Novosibirsk oblast, Russia)

A.S. Pilipenko – Cand. Sci. (Biology) (ICG SB RAS, Novosibirsk, Russia)

Yu.I. Ragino – Corr. Member of the RAS, Professor, Dr. Sci. (Medicine) (Research Institute of Internal and Preventive Medicine – Branch of the ICG SB RAS, Novosibirsk, Russia)

T. Rajabov – Professor, Dean (Institute of Agrobiotechnologies and Food Safety, Samarkand State University named after Sharof Rashidov, Samarkand, Uzbekistan)

I.D. Rashal – Full Member of the LAS, Professor, Dr. Sci. (Biology) (University of Latvia, Salaspils, Latvia)

R.R. Sadoyan – Professor, Dr. Sci. (Biology), Dean of the Faculty of Biology, Chemistry and Geography (Kh. Abovyan Armenian State Pedagogical University, Yerevan, Armenia)

S.V. Shekhovtsov – Cand. Sci. (Biology), Head of the Genogeography Sector of the Palearctic (Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia)

A.A. Soloviev – Professor, Dr. Sci. (Biology), Deputy Director (All-Russian Plant Quarantine Center, Moscow, Russia)

N.A. Surin – Full Member of the Russian Academy of Sciences, Dr. Sci. (Agricul.), Professor, Head of Scientific Direction

V.A. Trifonov – Professor, Dr. Sci. (Biology) (Institute of Molecular and Cellular Biology, SB RAS, Novosibirsk, Russia)

O.V. Vaulin – Cand. Sci. (Biology) (ICG SB RAS, Novosibirsk, Russia)

M.A. Vishnyakova – Professor, Dr. Sci. (Biology) (N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), St. Petersburg, Russia)

I.K. Zakharov – Professor, Dr. Sci. (Biology) (ICG SB RAS, Novosibirsk, Russia)

СОДЕРЖАНИЕ • 2025 • 11 • 1

- Генетика животных**
- 5 **Оригинальное исследование**
Гены-кандидаты, ассоциированные с весом тела, у животных молочных пород крупного рогатого скота
Н.С. Юдин, А.В. Игошин, Д.М. Ларкин
- Популяционная генетика**
- 12 **Оригинальное исследование**
Экспериментальный анализ полиморфных микросателлитных локусов у кладоцер: стратегия смешанного размножения и динамика популяций
А.А. Порошина, Д.Ю. Щербаков
- 17 **Обзор**
Полвека популяционной генетики народов Сибири: исследования Людмилы Павловны Осиповой
И.К. Захаров, О.Л. Посух
- Биоразнообразие**
- 29 **Оригинальное исследование**
Половой полиморфизм гинодиэцичного растения *Geranium sylvaticum* (Geraniaceae)
Н.И. Гордеева
- Молекулярная биология**
- 36 **Оригинальное исследование**
Определение *in vitro* средства ТВР к ТАТА-боксу промоторов генов *Arabidopsis thaliana*
Е.Б. Шарыпова, М.П. Пономаренко, Н.А. Колчанов, Л.К. Савинкова
- Математическая биология**
- 43 **Дискуссия**
Как считать критерий Стьюдента статистической значимости разности средних двух выборок.
I. Проверка нормальности
В.М. Ефимов

CONTENTS • 2025 • 11 • 1

- 5 **Animal genetics**
Original article
Weight-associated candidate genes
in dairy cattle
N.S. Yudin, A.V. Igoshin, D.M. Larkin
- 12 **Population genetics**
Original article
Experimental analysis of polymorphic
microsatellite loci in Cladocera: mixed breeding
strategy and population dynamics
A.A. Poroshina, D.Yu. Sherbakov
- 17 **Review**
Half a century of population genetics of the peoples
of Siberia: research by Lyudmila P. Osipova
I.K. Zakharov, O.L. Posukh
- 29 **Biodiversity**
Original article
Sexual polymorphism of a gynodioecious
plant *Geranium sylvaticum* (Geraniaceae)
N.I. Gordeeva
- 36 **Molecular biology**
Original article
In vitro determination of TBP affinity
to the TATA-box of promoters
of *Arabidopsis thaliana* genes
*E.B. Sharypova, M.P. Ponomarenko,
N.A. Kolchanov, L.K. Savinkova*
- 43 **Mathematical biology**
Discussion
How to calculate the Student's test
for the statistical significance of the difference
between the means of two samples.
I. Testing for normality
V.M. Efimov

 pismavavilov.ru

doi 10.18699/letvjgb-2025-11-02

Оригинальное исследование

Гены-кандидаты, ассоциированные с весом тела, у животных молочных пород крупного рогатого скота

Н.С. Юдин ¹, А.В. Игошин ¹, Д.М. Ларкин ²

Аннотация: Признаки, связанные с весом тела в разном возрасте, хорошо коррелируют с продуктивностью, здоровьем и продолжительностью хозяйственного использования крупного рогатого скота. В то же время у крупного рогатого скота вес тела – признак с высокой наследуемостью. Полногеномный анализ ассоциаций позволяет найти генетические варианты и потенциальные гены-кандидаты, контролируемые интересующий признак. Однако обзоры работ по изучению веса тела с помощью полногеномного анализа ассоциаций у животных молочных пород крупного рогатого скота практически отсутствуют. Целью этой работы было выявление потенциальных генов-кандидатов, ассоциированных с признаками веса тела, у животных молочных пород крупного рогатого скота на основании исследований по полногеномному анализу ассоциаций. Поиск литературы проведен с использованием поискового запроса “Cattle AND Dairy AND Weight AND Genome-wide AND Association AND Study”. Критерием для включения публикации в нашу работу служила идентификация хотя бы одного гена-кандидата, достоверно ассоциированного с признаком, связанным с весом тела у животных молочной породы крупного рогатого скота. Количество пересечений между списками генов-кандидатов подсчитывали с помощью программы Venn. Анализ взаимодействий между генами/белками проводили с использованием веб-ресурса GeneMANIA; перепредставленные термины геной онтологии были найдены с помощью веб-инструмента DAVID. Было обнаружено 20 генов, ассоциированных в одной публикации с разными признаками веса либо с одинаковым признаком, но в нескольких публикациях. Среди них наибольшее количество пересечений имели гены *TNNT3*, *DYRK4*, *AKAP3*, *GALNT8*, *NDUFA9* и *KCNA1*. Общие биохимические пути и белок-белковые взаимодействия были найдены для пар генов *TNNT3–TNNI2* и *FGF6–FGF23*. Общие белковые домены были обнаружены у пар *TNNT3–TNNI2*, *FGF6–FGF23* и *KCNA1–KCNA6*. При функциональной аннотации вышеупомянутого списка из 20 генов были выявлены перепредставленные термины геной онтологии, указывающие на обогащение списка генами, кодирующими белки комплекса тропонина (*TNNT3* и *TNNI2*), белки калиевых каналов внутреннего выпрямления (*KCNA1* и *KCNA6*), а также белки-рецепторы фактора роста фибробластов (*FGF6* и *FGF23*).

Ключевые слова: крупный рогатый скот; молочная порода; вес тела; полногеномный анализ ассоциаций; ген-кандидат

Для цитирования: Юдин Н.С., Игошин А.В., Ларкин Д.М. Гены-кандидаты, ассоциированные с весом тела, у животных молочных пород крупного рогатого скота. *Письма в Вавиловский журнал генетики и селекции*. 2025;11(1):5-11. doi 10.18699/letvjgb-2025-11-02

Финансирование: Работа выполнена при поддержке бюджетного проекта № FWNR-2022-0039.

Original article

Weight-associated candidate genes in dairy cattle

N.S. Yudin ¹, A.V. Igoshin ¹, D.M. Larkin ²

Abstract: Traits associated with body weight at different ages correlate strongly with the productivity, health, and longevity of cattle. At the same time, body weight is a trait with high heritability in cattle. Genome-wide association study can identify genetic variants and potential candidate genes that control the trait of interest. However, there are practically no reviews of genome-wide association studies of body weight traits in dairy cattle breeds. The aim of the review was to identify potential candidate genes associated with body weight traits in dairy cattle based on genome-wide association studies. Literature search was conducted using the query “Cattle AND Dairy AND Weight AND Genome-wide AND Association AND Study”. The criterion for inclusion in the review was the identification of at least one candidate gene significantly associated with a trait associated with body weight in dairy cattle. A web tool Venn was used to calculate the number of intersections between lists of candidate genes. Gene/protein interactions were analyzed using the GeneMANIA web resource. Overrepresented gene ontology terms were identified using the DAVID web tool. Twenty genes were identified that were associated in one publication with different weight traits, or with the same trait, but in several publications. We identified 20 genes that were associated in one publication with different weight traits, or with the same trait but in several publications. The genes *TNNT3*,

¹ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

² Королевский ветеринарный колледж, Университет Лондона, Лондон, Великобритания
Royal Veterinary College, University of London, London, United Kingdom

 yudin@bionet.nsc.ru

© Юдин Н.С., Игошин А.В., Ларкин Д.М., 2025

DYRK4, *AKAP3*, *GALNT8*, *NDUFA9*, and *KCNA1* had the highest number of intersections among them. Shared biochemical pathways and protein-protein interactions have been identified between *TNNT3–TNNI2* and *FGF6–FGF23* gene pairs. Shared protein domains were found in the pairs *TNNT3–TNNI2*, *FGF6–FGF23* and *KCNA1–KCNA6*. Functional annotation of the above-mentioned list of 20 genes revealed overrepresented gene ontology terms, indicating the enrichment of the list with genes encoding troponin complex proteins (*TNNT3* and *TNNI2*), proteins of delayed rectifier potassium channels (*KCNA1* and *KCNA6*), as well as fibroblast growth factor receptor proteins (*FGF6* and *FGF23*).

Key words: cattle; dairy breed; body weight; genome-wide association study; candidate gene

For citation: Yudin N.S., Igoshin A.V., Larkin D.M. Weight-associated candidate genes in dairy cattle. *Pisma v Vavilovskii Zhurnal Genetiki i Seleksii = Letters to Vavilov Journal of Genetics and Breeding*. 2025;11(1):5-11. doi 10.18699/letvjgb-2025-11-02 (in Russian)

Funding: This study was supported by a budget project No. FWNR-2022-0039.

Введение

Вес тела включен в селекционные индексы племенной ценности быков молочных пород крупного рогатого скота (КРС) во многих развитых странах в силу его влияния на производственные характеристики, здоровье и продолжительность хозяйственного использования животных (Tempelman, Lu, 2020). У КРС голштинской породы показана положительная корреляция веса телочек при рождении с удоем (Berry et al., 2003), потреблением сухого вещества и энергетическим балансом (Veerkamp et al., 2000), а также с репродуктивными показателями и здоровьем (Yin, König, 2018). Выживаемость новорожденных телят и легкость отела умеренно коррелируют с массой тела коров (Johanson et al., 2011). В исследовании (Han et al., 2021) показано, что более тяжелые телки давали больше молока в первую лактацию, но при этом они сильнее снижали массу тела, сталкивались с повышенным риском выбраковки и не производили больше молока в долгосрочной перспективе.

Полногеномный анализ ассоциаций (ПГАА) позволяет выявить значимые генетические варианты и потенциальные гены-кандидаты хозяйственно важных признаков КРС (Sahana et al., 2023). Принцип ПГАА основан на тестировании связи между генотипами и фенотипом, выполняемом для множества локусов исследуемого генома. Как правило, для ПГАА используют биочипы со множеством равномерно распределенных однонуклеотидных полиморфизмов, которые расположены близко к функциональным мутациям и находятся в неравновесии по сцеплению с причинными локусами, которые непосредственно влияют на признак. Известно, что вес тела – признак с высокой наследуемостью у КРС (Coffey et al., 2006). Это обстоятельство способствовало многочисленным исследованиям ПГАА признаков, связанных с весом тела у животных мясных пород КРС (Raza et al., 2019, 2020; Naserkheil et al., 2022). Однако в литературе практически отсутствуют обзоры работ по изучению этих признаков с помощью ПГАА у животных молочных пород КРС.

Цель настоящей работы – выявление на основании исследований ПГАА потенциальных генов-кандидатов, ассоциированных с весом тела у животных молочных пород КРС для планирования дальнейших экспериментов и проведения маркер-ориентированной и геномной селекции.

Материалы и методы

Поиск и отбор литературы были проведены в соответствии с общепринятыми критериями, предъявляемыми к систематическим обзорам (Page et al., 2021). Мы выполнили систематический поиск литературы, опубликованной в базах

данных PubMed, Scopus, Web of Science и Google Scholar до 1 ноября 2024 г., с использованием поискового запроса “Cattle AND Dairy AND Weight AND Genome-wide AND Association AND Study”. Критерием для включения публикации в нашу работу служила идентификация хотя бы одного гена-кандидата, достоверно ассоциированного с признаком, связанным с весом тела у животных молочной породы КРС. Для расчета количества пересечений между списками генов-кандидатов использовали программу Venn (<http://bioinformatics.psb.ugent.be/webtools/Venn/>).

Анализ взаимодействий между генами/белками проводили с помощью веб-ресурса GeneMANIA (Franz et al., 2018) посредством исследования совместной экспрессии и локализации, физических и генетических взаимодействий, а также общих белковых доменов и биохимических путей. Гены считали связанными совместной экспрессией, если их уровни экспрессии были схожими в разных условиях эксперимента. Гены считали связанными совместной локализацией, если они экспрессируются в одной и той же ткани или если их белковые продукты идентифицированы в одном месте клетки. Гены считали связанными физическим взаимодействием, если было обнаружено, что они контактируют при исследовании белок-белковых взаимосвязей. Гены считали связанными генетическим взаимодействием, если было обнаружено, что эффекты нарушения одного гена модифицируются нарушениями во втором гене.

Перепредставленные термины генной онтологии были выявлены с помощью веб-инструмента DAVID (Sherman et al., 2022). DAVID позволяет находить термины генной онтологии, которыми насыщены аннотации генов из исследуемого списка. При анализе использовали словари клеточных компонентов и молекулярных функций. В качестве контрольного (референсного) набора был список генов вида *Bos taurus*. Термины считали достоверными при пороге $p < 0.05$ и кратности обогащения > 70 .

Результаты и обсуждение

Гены-кандидаты, ассоциированные с весом тела у животных молочных пород

В результате проведенного поиска литературы для детального анализа нами было отобрано восемь публикаций (табл. 1). Из них семь статей посвящено исследованию голштинской и голштино-фризской породы, а в одной был выполнен анализ, помимо голштинской, красных молочных пород. Минимальный размер выборки составлял 763 животных (Doran et al., 2014), максимальный – 53 644 животных (Cole et al., 2014). Фенотипические признаки включают

Таблица 1. Характеристика исследований по полногеномному анализу ассоциаций признаков веса у животных молочных пород КРС**Table 1.** Characteristics of genome-wide association studies of weight traits in dairy cattle genome-wide association study

Порода	Размер выборки	Число SNP	Признак (число генов-кандидатов)	Литературный источник
Голштинская (Германия)	15 921	45 613	Вес при рождении (51), вес в 2–3 мес. (12), вес в 13–14 мес. (38)	Yin, König, 2019
Голштинская (Австралия и Новая Зеландия)	1782	624 930	Вес в возрасте 250 дней (3)	Pryce et al., 2012
Голштинская (Германия)	53 644	43 188	Вес при рождении (20)	Cole et al., 2014
Голштинская (КНР)	971	114 658	Ежемесячные показатели веса с рождения до 12 мес. (187)	Ma et al., 2023
Голштино-фризская (Ирландия)	763	42 477	Убойный вес (184)	Doran et al., 2014
Голштинская (Израиль)	1749	40 498	Ежегодный привес (1)	Gershoni et al., 2021
Голштино-фризская (Ирландия, Великобритания, Нидерланды, Швеция)	1629	37 590	Вес в возрасте 44 недели (4)	Veerkamp et al., 2012
Голштинская (Дания, Финляндия, Швеция)	5519	13 396 556*	Суточные привесы в возрасте менее и более 550 дней (17)	Mao et al., 2016
Красные молочные породы (Danish Red, Swedish Red, Finnish Ayrshire)	4410	13 396 556*	Суточные привесы в возрасте менее и более 550 дней (8)	Mao et al., 2016

* Импутация по данным полногеномного секвенирования данных генотипирования на биочипе Illumina 50k Bovine BeadChip.

вес тела в разном возрасте, а также ежегодные или суточные привесы. Большинство работ было проведено с использованием стандартного биочипа Illumina BovineSNP50 BeadChip, по одной работе – с использованием биочипов Illumina 150K Bovine Beadchip (Ma et al., 2023) и Illumina High Density Bovine SNP chip (Pryce et al., 2012). Еще в одной работе использована импутация по данным полногеномного секвенирования (Mao et al., 2016).

В вышеперечисленных статьях было экстрагировано всего 525 генов-кандидатов, связанных с признаками веса/привеса (см. табл. 1). Путем пересечения списков нами выявлено 20 генов, ассоциированных либо с разными признаками веса/привеса в одной статье, либо с одинаковым признаком, но более чем в одной работе (табл. 2). Из них ген *TNNT3* имел четыре таких пересечения, а гены *DYRK4*, *AKAP3*, *GALNT8*, *NDUFA9* и *KCNA1* – три. Следует отметить, что 9 из 20 генов, *AKAP3*, *CCND2*, *DYRK4*, *FGF23*, *FGF6*, *GALNT8*, *KCNA1*, *KCNA6*, *TIGAR*, которые мы нашли на пересечении ранее с помощью ПГАА, были ассоциированы с метаболической массой тела (масса тела в степени 0.75) у голштинов (Hardie et al., 2017). Этот показатель оценивает количество метаболически активной ткани в организме и хорошо коррелирует с убойным весом у животных молочных пород КРС (Mehtio et al., 2021).

Гены *TNNT3* и *TNNI2*

Гены *TNNT3* и *TNNI2* кодируют субъединицы тропонинового комплекса, являющегося необходимым компонентом аппарата мышечного сокращения (Mullen, Barton, 2000). Вероятно, эти гены вовлечены в контроль массы тела/мышечной массы также и у животных мясных пород КРС. В. Santiago с коллегами (2023) провели сравнительное протеомное исследование образцов скелетной мышечной ткани (длиннейшей мышцы груди) кастрированных и некастрированных бычков, полученных от скрещивания пород ангус и нелоре. Показано, что некастрированные бычки, которые при этом заметно превосходят кастрированных по среднесуточному

привесу, массе туши и живой массе, имеют значимо большую экспрессию продуктов генов *TNNT3* и *TNNI2* (Santiago et al., 2023). В другой работе проведен сравнительный транскриптомный анализ тканей длиннейшей мышцы спины у яков и у гибридов яка и КРС. Было обнаружено, что у гибридов значимо повышена экспрессия шести сплайс-изоформ гена *TNNI2*, чем могут быть объяснены лучшие показатели роста и развития скелетной мускулатуры у гибридов по сравнению с чистокровными яками (Cao et al., 2025). Примечательны данные R.J. Schilder с коллегами (2011), полученные на модели крыс. Они показали, что при повышении массы тела животных *TNNT3* меняет профиль альтернативного сплайсинга. Причем сходный эффект наблюдается как при действительном повышении массы тела крыс, так и при искусственном – с помощью жилетов с утяжелителями. Ученые полагают, что ген *TNNT3* выполняет роль молекулярного сенсора массы тела, который позволяет изменять адаптацию мышечной ткани в ответ на возросшие нагрузки. Вероятно, нарушение этой функции гена *TNNT3* приводит к ожирению (Schilder et al., 2011).

Гены в районе ~0.5 млн п. н. на 5-й хромосоме

Десять из двадцати генов нашего списка (*KCNA1*, *KCNA6*, *GALNT8*, *NDUFA9*, *AKAP3*, *DYRK4*, *FGF6*, *FGF23*, *TIGAR*, *CCND2*) попадают в район на 5-й хромосоме длиной ~0.5 млн п. н. (BTA5:105.26–105.79 м. п. н.). Ранее для этого же района нами была показана ассоциация с живым весом у животных герефордской мясной породы КРС (Igoshin et al., 2019). Однако в нашей работе сигналы ассоциации были в основном сконцентрированы около гена *CCND2*, кодирующего белок циклин D2, участвующий в контроле клеточного цикла и являющийся стимулятором клеточной пролиферации (Zhao et al., 2021). Также у герефордской породы, по данным С.М. Seabury с коллегами (2017), *CCND2* ассоциирован со среднесуточным привесом. Кроме того, имеются данные о связи этого гена с отложением жира у свиней (Li et al., 2022; Oliveira et al., 2022), набором мышечной массы у кур (Willson

Таблица 2. Описание и хромосомная локализация генов, ассоциированных с весом тела, у животных молочных пород крупного рогатого скота**Table 2.** Description and chromosomal localization of weight-associated genes in dairy cattle breeds

Символ гена	Полное название	Хромосома	Локализация гена (п. н.) ^а	Литературный источник
<i>TNNT3</i>	Тропонин Т3, быстрый скелетный мышечный	29	49579057–49594681	Doran et al., 2014; Mao et al., 2016; Yin, König, 2019 ^б
<i>DYRK4</i>	Регулируемая фосфорилированием тирозина киназа с двойной специфичностью 4	5	105528478–105577743	Doran et al., 2014; Yin, König, 2019 ^б
<i>AKAP3</i>	Якорный белок А-киназы 3	5	105495981–105523869	
<i>GALNT8</i>	Полипептид N-ацетилгалактосаминилтрансфераза 8	5	105400923–105441611	
<i>NDUFA9</i>	Субъединица А9 НАДН: убихинон оксидоредуктазы	5	105474046–105495864	
<i>KCNA1</i>	Член подсемейства А потенциал-зависимых калиевых каналов	5	105263614–105271979	
<i>KCNA6</i>	Член подсемейства А потенциал-зависимых калиевых каналов	5	105327366–105370526	Yin, König, 2019 ^б
<i>TSPAN32</i>	Тетраспанин 32	29	49265270–49282239	Doran et al., 2014; Yin, König, 2019 ^б
<i>CTSD</i>	Катепсин D	29	49712746–49722179	
<i>LSP1</i>	Лимфоцит-специфический белок 1	29	49598905–49636903	
<i>FGF6</i>	Фактор роста фибробластов 6	5	105667368–105679381	
<i>TNNI2</i>	Тропонин I, быстрый скелетный мышечный	29	49645742–49648340	
<i>CCND2</i>	Циклин D2	5	105763348–105786522	
<i>TIGAR</i>	Индукцированная гликолизом регуляторная фосфатаза TP53	5	105728270–105747498	
<i>FGF23</i>	Фактор роста фибробластов 23	5	105716618–105728073	
<i>PLAG1</i>	Фактор транскрипции PLAG1	14	23323709–23375679	Pryce et al., 2012; Mao et al., 2016
<i>RFX6</i>	Регуляторный фактор X6	9	33719317–33842656	Cole et al., 2014; Doran et al., 2014
<i>SCO2</i>	Фактор сборки цитохром-С-оксидазы 2	5	119822432–119824250	Gershoni et al., 2021; Ma et al., 2023
<i>CETN3</i>	Центрин 3	7	89936060–89953545	Veerkamp et al., 2012; Ma et al., 2023
<i>CYP19A1</i>	Цитохром P450 19A1	10	59102240–59156621	Mao et al., 2016 ^г

Примечание. ^а Согласно референсной сборке генома KPC ARS-UCD2.0; ^б ассоциирован с весом при рождении и весом в 2–3 месяца; ^в ассоциирован с весом при рождении; ^г ассоциирован с суточными привесами в возрасте менее и более 550 дней у голштинской и красных молочных пород. Гены ассоциированы либо с разными признаками веса/привеса в одной статье, либо с одинаковым признаком, но более чем в одной работе.

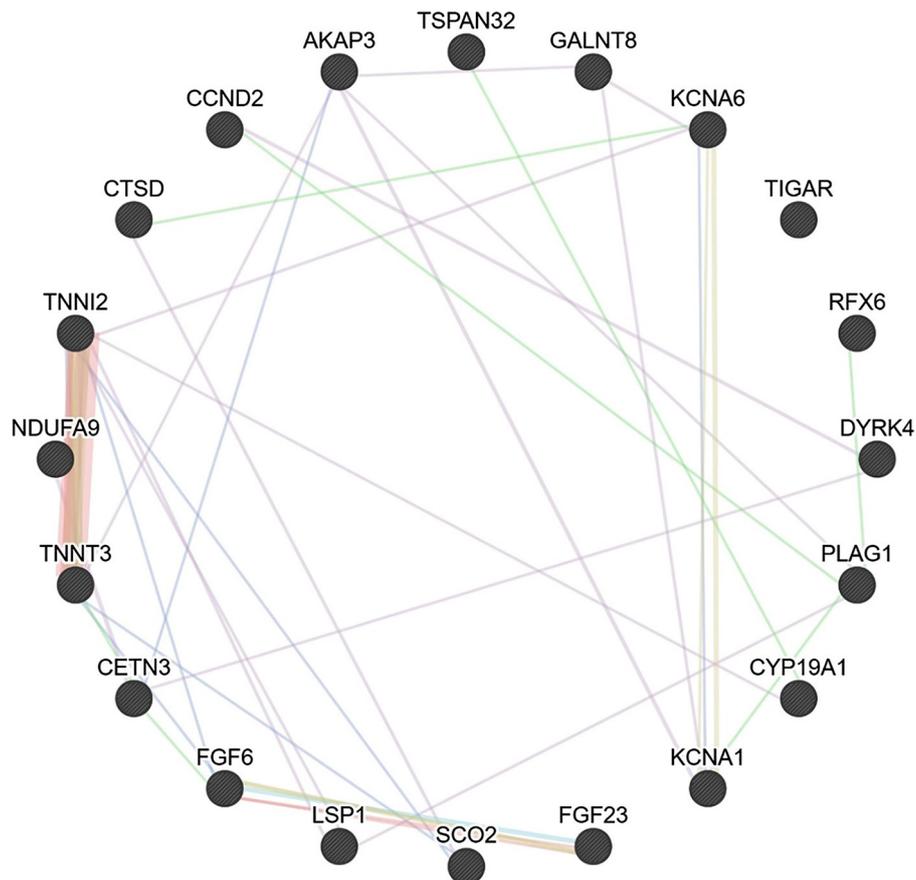
et al., 2018), весом тела у мышей (Brown, Moore, 2012) и ростом и индексом массы тела у человека (Steinthorsdottir et al., 2014). Исходя из вышесказанного, мы предложили *CCND2* в качестве основного гена-кандидата, вовлеченного в контроль массы тела у животных исследуемой породы. Таким образом, можно предположить, что сигналы ассоциации в других генах – это результат эффекта неравновесия по сцеплению. Это возможно, с учетом, что в обоих исследованиях они были выявлены в голштинской породе, для которой вследствие инбридинга и интенсивной селекции характерны достаточно протяженные гаплотипы (Kim, Kirkpatrick, 2009).

Нельзя также исключать и роль других генов из вышеуказанного района в контроле массы тела животных. Например, В. Ху с коллегами (2021) на модели мышей обнаружили гиперметилирование промотора и снижение экспрессии мРНК гена *FGF6* у животных с ожирением. Исследователи искусственно индуцировали сверхэкспрессию *FGF6* в скелетной мускулатуре мышей, благодаря чему у них произошло увеличение мышечной и снижение жировой массы в условиях высокожировой диеты по сравнению с контролем. При этом общий вес тела был ниже в тестовой группе (Ху et

al., 2021). R. Marsell с коллегами (2009) измеряли сывороточный уровень белка *FGF23*, также относящегося к семейству факторов роста фибробластов, в выборке пожилых людей. Было показано, что уровень *FGF23* в сыворотке крови положительно коррелирует с массой тела.

Прочие гены

Ген *TIGAR* также может быть ассоциирован с весом у КРС. По данным исследования на модели мышей, сверхэкспрессия *TIGAR* у животных вызывает ожирение (Zhang et al., 2024). Белковый продукт гена *CTSD* – катепсин D – участвует в деградации полипептидных гормонов и факторов роста (Benes et al., 2008). Однонуклеотидный полиморфизм в гене *PLAG1* ранее был ассоциирован с весом туши у животных мясной японской черной породы (Nishimura et al., 2012). Ген *CYP19A1* кодирует цитохром P450-зависимую монооксигеназу, участвующую в синтезе холестерина и стероидов, в том числе эстрадиола. Мутации в этом гене приводят к нарушению роста и дифференцировки клеток (Glubb et al., 2017). Информация об участии других генов в контроле признаков, связанных с весом тела, в доступной литературе отсутствует.



Сеть взаимодействий между 20 генами/белками, ассоциированными с весом тела, у животных молочных пород крупного рогатого скота.

Сиреневым цветом показаны сходные паттерны экспрессии (41.32 %), красным – физические взаимодействия (26.39 %), желтым – общие белковые домены (14.90 %), синим – биохимические пути (9.77 %), фиолетовым – совместная локализация (5.56 %), зеленым – генетические взаимодействия (2.06 %)

A network of interactions between 20 genes and the proteins they encode associated with body weight in dairy cattle

Lilac shows similar expression patterns (41.32 %), red shows physical interactions (26.39%), yellow shows shared protein domains (14.90 %), blue shows biochemical pathway (9.77 %), violet shows co-localization (5.56 %), and green shows genetic interactions (2.06 %)

Анализ взаимодействий и функциональная аннотация генов

Результаты анализа взаимодействий между генами/белками, ассоциированными с весом, с помощью веб-ресурса GeneMANIA приведены на рисунке. Различные виды взаимодействий охватывают все гены, за исключением *TIGAR*. Физические взаимодействия и общие биохимические пути были обнаружены между парами генов *TNNT3–TNNI2*, а также *FGF6–FGF23*. Общие белковые домены были найдены у пар белков *TNNT3–TNNI2*, *FGF6–FGF23* и *KCNA1–KCNA6*.

В нашей работе был использован веб-инструмент DAVID для идентификации клеточных компонентов и молекулярных функций, которые могут участвовать в генетическом контроле признаков веса у животных молочных пород КРС (табл. 3). Мы выявили перепредставленные термины, связанные с белками комплекса тропонина *TNNT3* и *TNNI2*, ко-

торые входят в состав миофиламентов и тонких филаментов поперечно-полосатых мышц (Pepper, Galkin, 2022).

Гены *KCNA1* и *KCNA6* кодируют калиевые каналы внутреннего выпрямления. Через эти каналы ионы калия проходят только внутрь клетки, а не наружу, как в случае «типичных» калиевых каналов. Считается, что этот ток ионов внутрь клетки может играть важную роль в стабилизации мембранного потенциала покоя возбудимых клеток (Hibino et al., 2010). У человека мутации в этих генах приводят к наследственным заболеваниям опорно-двигательного аппарата (Maggi et al., 2021).

Белок *FGF6* экспрессируется главным образом в миоцитах и их предшественниках. Он играет ключевую роль в процессах миогенеза и ангиогенеза, регенерации мышц, а также транспорте железа и липидном метаболизме (Smith, Jerome-Majewska, 2024). Белок *FGF23* выделяется остеоцитами в костной ткани и регулирует гомеостаз и транспорт фосфатов в почках. Нарушение его секреции приводит к

Таблица 3. Перепредставленные термины генной онтологии, выявленные при анализе 20 генов, контролирующих вес тела у животных молочных пород КРС**Table 3.** Overrepresented gene ontology terms identified in the analysis of the 20 genes controlling body weight in dairy cattle breeds

Термин	Ген	Кратность обогащения	P
GO:0005861~ комплекс тропонина	<i>TNNT3, TNNI2</i>	252.73	0.007
GO:0005865~ тонкий филамент поперечно-полосатой мышцы	<i>TNNT3, TNNI2</i>	113.73	0.02
GO:0036379~ миофиламент	<i>TNNT3, TNNI2</i>	103.39	0.02
GO:0005251~ калиевые каналы внутреннего выпрямления	<i>KCNA1, KCNA6</i>	83.85	0.03
GO:0005104~ связывание рецептора фактора роста фибробластов	<i>FGF6, FGF23</i>	74.86	0.03

хронической почечной недостаточности, которая сопровождается атрофией скелетной мускулатуры (Elsurer Afsar et al., 2023).

Заключение

Использование ПГАА позволило идентифицировать ряд генов-кандидатов, ассоциированных с признаками веса тела у животных молочных пород КРС. В настоящей работе обобщена информация об этих генах и проведена их функциональная аннотация. Нами выявлено 20 генов, которые были ассоциированы либо с разными признаками веса в одной статье, либо с одинаковым признаком, но в нескольких публикациях. Среди них наибольшее количество совпадений имел ген *TNNT3*, далее по рангу шли гены *DYRK4*, *AKAP3*, *GALNT8*, *NDUFA9* и *KCNA1*. Физические взаимодействия и общие биохимические пути были обнаружены между парами генов *TNNT3-TNNI2*, а также *FGF6-FGF23*. Общие белковые домены были найдены у пар *TNNT3-TNNI2*, *FGF6-FGF23* и *KCNA1-KCNA6*. Список, включающий 20 генов, характеризовался значимо перепредставленным количеством связей с терминами генной онтологии, относящимися к белкам комплекса тропонина *TNNT3* и *TNNI2*, которые входят в состав миофиламентов и тонких филаментов поперечно-полосатых мышц, а также с белками калиевых каналов внутреннего выпрямления *KCNA1* и *KCNA6*. Представленные нами результаты позволяют лучше понять генетическую архитектуру и биологические механизмы, регулирующие признаки веса тела у КРС. Они также могут быть использованы для проведения маркер-ориентированной и геномной селекции на улучшение этих признаков у животных молочных пород.

Список литературы / References

Benes P., Vetvicka V., Fusek M. Cathepsin D – many functions of one aspartic protease. *Crit Rev Oncol Hematol*. 2008;68(1):12-28. doi 10.1016/j.critrevonc.2008.02.008

Berry D.P., Buckley F., Dillon P., Evans R.D., Rath M., Veerkamp R.F. Genetic relationships among body condition score, body weight, milk yield, and fertility in dairy cows. *J Dairy Sci*. 2003;86(6):2193-2204. doi 10.3168/jds.S0022-0302(03)73809-0

Brown S.D.M., Moore M.W. The International Mouse Phenotyping Consortium: past and future perspectives on mouse phenotyping. *Mamm Genome*. 2012;23(9-10):632-640. doi 10.1007/s00335-012-9427-x

Cao M., Ding Z., Wang X., Guo S., Kang Y., Hu L., Zhang B., Pei J., Ma Y., Guo X. Full-length transcriptome sequencing of the longissimus dorsi muscle of yak and cattle-yak using nanopore technology. *Int J Biol Macromol*. 2025;284(1):138071. doi 10.1016/j.jbiomac.2024.138071

Coffey M.P., Hickey J., Brotherstone S. Genetic aspects of growth of Holstein-Friesian dairy cows from birth to maturity. *J Dairy Sci*. 2006;89(1):322-329. doi 10.3168/jds.S0022-0302(06)72097-5

Cole J.B., Waurich B., Wensch-Dorendorf M., Bickhart D.M., Swalve H.H. A genome-wide association study of calf birth weight in Holstein cattle using single nucleotide polymorphisms and phenotypes predicted from auxiliary traits. *J Dairy Sci*. 2014;97(5):3156-3172. doi 10.3168/jds.2013-7409

Doran A.G., Berry D.P., Creevey C.J. Whole genome association study identifies regions of the bovine genome and biological pathways involved in carcass trait performance in Holstein-Friesian cattle. *BMC Genomics*. 2014;15(1):837. doi 10.1186/1471-2164-15-837

Elsurer Afsar R., Afsar B., Ikizler T.A. Fibroblast growth factor 23 and muscle wasting: a metabolic point of view. *Kidney Int Rep*. 2023;8(7):1301-1314. doi 10.1016/j.ekir.2023.04.027

Franz M., Rodriguez H., Lopes C., Zuberi K., Montojo J., Bader G.D., Morris Q. GeneMANIA update 2018. *Nucleic Acids Res*. 2018;46(W1):W60-W64. doi 10.1093/nar/gky311

Gershoni M., Weller J.I., Ezra E. Genetic and genome-wide association analysis of yearling weight gain in Israel Holstein dairy calves. *Genes (Basel)*. 2021;12(5):708. doi 10.3390/genes12050708

Glubb D.M., O'Mara T.A., Shamsani J., Spurdle A.B. The association of CYP19A1 variation with circulating estradiol and aromatase inhibitor outcome: can CYP19A1 variants be used to predict treatment efficacy? *Front Pharmacol*. 2017;8:218. doi 10.3389/fphar.2017.00218

Han L., Heinrichs A.J., De Vries A., Dechow C.D. Relationship of body weight at first calving with milk yield and herd life. *J Dairy Sci*. 2021;104(1):397-404. doi 10.3168/jds.2020-19214

Hardie L.C., VandeHaar M.J., Tempelman R.J., Weigel K.A., Armentano L.E., Wiggans G.R., Veerkamp R.F., de Haas Y., Coffey M.P., Connor E.E., Hanigan M.D., Staples C., Wang Z., Dekkers J.C.M., Spurlock D.M. The genetic and biological basis of feed efficiency in mid-lactation Holstein dairy cows. *J Dairy Sci*. 2017;100(11):9061-9075. doi 10.3168/jds.2017-12604

Hibino H., Inanobe A., Furutani K., Murakami S., Findlay I., Kurachi Y. Inwardly rectifying potassium channels: their structure, function, and physiological roles. *Physiol Rev*. 2010;90(1):291-366. doi 10.1152/physrev.00021.2009

Igoshin A.V., Yudin N.S., Belonogova N.M., Larkin D.M. Genome-wide association study for body weight in cattle populations from Siberia. *Anim Genet*. 2019;50(3):250-253. doi 10.1111/age.12786

Johanson J.M., Berger P.J., Tsuruta S., Misztal I. A Bayesian threshold-linear model evaluation of perinatal mortality, dystocia, birth weight, and gestation length in a Holstein herd. *J Dairy Sci*. 2011;94(1):450-460. doi 10.3168/jds.2009-2992

Kim E.S., Kirkpatrick B.W. Linkage disequilibrium in the North American Holstein population. *Anim Genet*. 2009;40(3):279-288. doi 10.1111/j.1365-2052.2008.01831.x

Li J., Wu J., Jian Y., Zhuang Z., Qiu Y., Huang R., Lu P., Guan X., Huang X., Li S., Min L., Ye Y. Genome-wide association studies revealed significant QTLs and candidate genes associated with backfat and loin muscle area in pigs using imputation-based whole genome sequencing data. *Animals (Basel)*. 2022;12(21):2911. doi 10.3390/ani12212911

Ma Z., Chang Y., Brito L.F., Li Y., Yang T., Wang Y., Yang N. Multitrait meta-analyses identify potential candidate genes for growth-related traits

- in Holstein heifers. *J Dairy Sci.* 2023;106(12):9055-9070. doi 10.3168/jds.2023-23462
- Maggi L., Bonanno S., Altamura C., Desaphy J.F. Ion channel gene mutations causing skeletal muscle disorders: pathomechanisms and opportunities for therapy. *Cells.* 2021;10(6):1521. doi 10.3390/cells10061521
- Mao X., Sahana G., De Koning D.J., Gulbrandsen B. Genome-wide association studies of growth traits in three dairy cattle breeds using whole-genome sequence data. *J Anim Sci.* 2016;94(4):1426-1437. doi 10.2527/jas.2015-9838
- Marsell R., Mirza M.A., Mallmin H., Karlsson M., Mellström D., Orwoll E., Ohlsson C., Jonsson K.B., Ljunggren O., Larsson T.E. Relation between fibroblast growth factor-23, body weight and bone mineral density in elderly men. *Osteoporos Int.* 2009;20(7):1167-1173. doi 10.1007/s00198-008-0780-2
- Mehtio T., Pitkanen T., Leino A.M., Mantysaari E.A., Kempe R., Negussie E., Lidauer M.H. Genetic analyses of metabolic body weight, carcass weight and body conformation traits in Nordic dairy cattle. *Animal.* 2021;15(12):100398. doi 10.1016/j.animal.2021.100398
- Mullen A.J., Barton P.J. Structural characterization of the human fast skeletal muscle troponin I gene (*TNNI2*). *Gene.* 2000;242(1-2):313-320. doi 10.1016/S0378-1119(99)00519-3
- Naserkheil M., Manzari Z., Dang C.G., Lee S.S., Park M.N. Exploring and identifying candidate genes and genomic regions related to economically important traits in Hanwoo cattle. *Curr Issues Mol Biol.* 2022;44(12):6075-6092. doi 10.3390/cimb44120414
- Nishimura S., Watanabe T., Mizoshita K., Tatsuda K., Fujita T., Watanabe N., Sugimoto Y., Takasuga A. Genome-wide association study identified three major QTL for carcass weight including the PLAG1-CHCHD7 QTN for stature in Japanese Black cattle. *BMC Genet.* 2012;13:40. doi 10.1186/1471-2156-13-40
- Oliveira H.C., Derks M.F.L., Lopes M.S., Madsen O., Harlizius B., van Son M., Grindflek E.H., Godia M., Gjuvsland A.B., Otto P.I., Groenen M.A.M., Guimaraes S.E.F. Fine mapping of a major backfat QTL reveals a causal regulatory variant affecting the *CCND2* gene. *Front Genet.* 2022;13:871516. doi 10.3389/fgene.2022.871516
- Page M.J., McKenzie J.E., Bossuyt P.M., Boutron I., Hoffmann T.C., Mulrow C.D., Shamseer L., ... Thomas J., Tricco A.C., Welch V.A., Whiting P., Moher D. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021;372:n160. doi 10.1136/bmj.n71
- Pepper I., Galkin V.E. Actomyosin Complex. In: Harris J.R., Marles-Wright J. (Eds) *Macromolecular Protein Complexes IV. Subcellular Biochemistry.* Vol. 99. Springer, 2022;421-470. doi 10.1007/978-3-031-00793-4_14
- Pryce J.E., Arias J., Bowman P.J., Davis S.R., Macdonald K.A., Waghorn G.C., Wales W.J., Williams Y.J., Spelman R.J., Hayes B.J. Accuracy of genomic predictions of residual feed intake and 250-day body weight in growing heifers using 625,000 single nucleotide polymorphism markers. *J Dairy Sci.* 2012;95(4):2108-2119. doi 10.3168/jds.2011-4628
- Raza S.H.A., Khan R., Abdelnour S.A., Abd El-Hack M.E., Khafaga A.F., Taha A., Ohran H., Mei C., Schreurs N.M., Zan L. Advances of molecular markers and their application for body variables and carcass traits in Qinchuan cattle. *Genes (Basel).* 2019;10(9):717. doi 10.3390/genes10090717
- Raza S.H.A., Khan S., Amjadi M., Abdelnour S.A., Ohran H., Alanazi K.M., Abd El-Hack M.E., Taha A.E., Khan R., Gong C., Schreurs N.M., Zhao C., Wei D., Zan L. Genome-wide association studies reveal novel loci associated with carcass and body measures in beef cattle. *Arch Biochem Biophys.* 2020;694:108543. doi 10.1016/j.abb.2020.108543
- Sahana G., Cai Z., Sanchez M.P., Bouwman A.C., Boichard D. Invited review: good practices in genome-wide association studies to identify candidate sequence variants in dairy cattle. *J Dairy Sci.* 2023;106(8):5218-5241. doi 10.3168/jds.2022-22694
- Santiago B., Baldassini W., Neto O.M., Chardulo L.A., Torres R., Pereira G., Curi R., Chiaratti M.R., Padilha P., Alessandroni L., Gagaoua M. Post-mortem muscle proteome of crossbred bulls and steers: relationships with carcass and meat quality. *J Proteomics.* 2023;278:104871. doi 10.1016/j.jpro.2023.104871
- Schilder R.J., Kimball S.R., Marden J.H., Jefferson L.S. Body weight-dependent troponin T alternative splicing is evolutionarily conserved from insects to mammals and is partially impaired in skeletal muscle of obese rats. *J Exp Biol.* 2011;214(9):1523-1532. doi 10.1242/jeb.051763
- Seabury C.M., Oldeschulte D.L., Saatchi M., Beever J.E., Decker J.E., Halley Y.A., Bhattarai E.K., Molaei M., Freetly H.C., Hansen S.L., Yampara-Iquise H., Johnson K.A., Kerley M.S., Kim J., Loy D.D., Marques E., Neibergs H.L., Schnabel R.D., Shike D.W., Spangler M.L., Weaber R.L., Garrick D.J., Taylor J.F. Genome-wide association study for feed efficiency and growth traits in U.S. beef cattle. *BMC Genomics.* 2017;18(1):386. doi 10.1186/s12864-017-3754-y
- Sherman B.T., Hao M., Qiu J., Jiao X., Baseler M.W., Lane H.C., Imamichi T., Chang W. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* 2022;50(W1):W216-W221. doi 10.1093/nar/gkac194
- Smith J., Jerome-Majewska L.A. Reprint of: fibroblast growth factor 6. *Differentiation.* 2024;139:100805. doi 10.1016/j.diff.2024.100805
- Steinthorsdottir V., Thorleifsson G., Sulem P., Helgason H., Grarup N., Sigurdsson A., Helgadóttir H.T., ... Kong A., Gudbjartsson D.F., Pedersen O., Thorsteinsdóttir U., Stefansson K. Identification of low-frequency and rare sequence variants associated with elevated or reduced risk of type 2 diabetes. *Nat Genet.* 2014;46(3):294-298. doi 10.1038/ng.2882
- Tempelman R.J., Lu Y. Symposium review: genetic relationships between different measures of feed efficiency and the implications for dairy cattle selection indexes. *J Dairy Sci.* 2020;103(6):5327-5345. doi 10.3168/jds.2019-17781
- Veerkamp R.F., Oldenbroek J.K., Van der Gaast H.J., Van der Werf J.H. Genetic correlation between days until start of luteal activity and milk yield, energy balance, and live weights. *J Dairy Sci.* 2000;83(3):577-583. doi 10.3168/jds.s0022-0302(00)74917-4
- Veerkamp R.F., Coffey M., Berry D., de Haas Y., Strandberg E., Bovenhuis H., Calus M., Wall E. Genome-wide associations for feed utilization complex in primiparous Holstein-Friesian dairy cows from experimental research herds in four European countries. *Animal.* 2012;6(11):1738-1749. doi 10.1017/S1751731112001152
- Willson N.L., Forder R.E.A., Tearle R., Williams J.L., Hughes R.J., Natrass G.S., Hynd P.I. Transcriptional analysis of liver from chickens with fast (meat bird), moderate (F1 layer x meat bird cross) and low (layer bird) growth potential. *BMC Genomics.* 2018;19(1):309. doi 10.1186/s12864-018-4723-9
- Xu B., Liu C., Zhang H., Zhang R., Tang M., Huang Y., Jin L., Xu L., Hu C., Jia W. Skeletal muscle-targeted delivery of Fgf6 protects mice from diet-induced obesity and insulin resistance. *JCI Insight.* 2021;6(19):e149969. doi 10.1172/jci.insight.149969
- Yin T., König S. Genetic parameters for body weight from birth to calving and associations between weights with test-day, health, and female fertility traits. *J Dairy Sci.* 2018;101(3):2158-2170. doi 10.3168/jds.2017-13835
- Yin T., König S. Genome-wide associations and detection of potential candidate genes for direct genetic and maternal genetic effects influencing dairy cattle body weight at different ages. *Genet Sel Evol.* 2019;51(1):4. doi 10.1186/s12711-018-0444-4
- Zhang T., Linghu K.G., Tan J., Wang M., Chen D., Shen Y., Wu J., Shi M., Zhou Y., Tang L., Liu L., Qin Z.H., Guo B. TIGAR exacerbates obesity by triggering LRRK2-mediated defects in macroautophagy and chaperone-mediated autophagy in adipocytes. *Autophagy.* 2024;20(8):1741-1761. doi 10.1080/15548627.2024.2338576
- Zhao M., Nakada Y., Wei Y., Bian W., Chu Y., Borovjagin A.V., Xie M., Zhu W., Nguyen T., Zhou Y., Serpooshan V., Walcott G.P., Zhang J. Cyclin D2 overexpression enhances the efficacy of human induced pluripotent stem cell-derived cardiomyocytes for myocardial repair in a swine model of myocardial infarction. *Circulation.* 2021;144(3):210-228. doi 10.1161/CIRCULATIONAHA.120.049497

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 29.11.2024. После доработки 24.12.2024. Принята к публикации 25.12.2024.

 pismavavilov.ru

doi 10.18699/letvjgb-2025-11-03

Original article

Experimental analysis of polymorphic microsatellite loci in Cladocera: mixed breeding strategy and population dynamics

A.A. Poroshina  , D.Yu. Sherbakov 

Abstract: The aim of this work is to identify polymorphic loci in *Daphnia* genomes using genome sequence analysis and experimental analysis of real accessions of samples collected in nature. We use primers predicted by an algorithm for DNA amplification from *Daphnia* samples and capillary electrophoresis to detect polymorphic loci. This was done to detect polymorphic loci in a single experiment, when the presence of polymorphic loci with a frequency sufficient for genetic studies is detected as the appearance of multiple peaks on the electrophoregram. Thus, we evaluate the effectiveness of the capillary electrophoresis method and refine population parameters, including the proportion of sexual reproduction and the effective population size. Studying polymorphic loci becomes especially relevant in the context of climate change and anthropogenic impacts on ecosystems, as these factors can significantly affect genetic diversity and adaptation in populations. Polymorphic loci can serve as indicators of genetic stability and the ability of species to adapt to changing environmental conditions. In the face of global climate change and increasing anthropogenic impacts, such as pollution, urbanization, and landscape changes, populations face new challenges that require rapid adaptation. Polymorphic loci, due to their high variability, can provide valuable information about the genetic structure of populations and their potential for adaptation. In the context of our research, studying polymorphic loci in *Daphnia* becomes particularly important, given their ecological role and sensitivity to environmental changes. *Daphnia*, as key components of freshwater ecosystems, can serve as model organisms for studying genetic adaptation to climate change and anthropogenic impacts. Understanding polymorphic loci in *Daphnia* can provide valuable data for assessing genetic stability and adaptive potential in these populations, which is particularly important in the face of rapidly changing environmental conditions.

Key words: polymorphic loci; capillary electrophoresis; agent-based modeling; microsatellites; *Daphnia*; genetic diversity; adaptation; climate change; anthropogenic impacts

For citation: Poroshina A.A., Sherbakov D.Yu. Experimental analysis of polymorphic microsatellite loci in Cladocera: mixed breeding strategy and population dynamics. *Pisma v Vavilovskii Zhurnal Genetiki i Seleksii* = *Letters to Vavilov Journal of Genetics and Breeding*. 2025;11(1):12-16. doi 10.18699/letvjgb-2025-11-03

Funding: The study was carried out within the framework of the state budget theme No. 0279-2021-0010 "Genetics of Baikal organism communities: genofond structure, conservation strategies".

Оригинальное исследование

Экспериментальный анализ полиморфных микросателлитных локусов у кладоцер: стратегия смешанного размножения и динамика популяций

А.А. Порошина , Д.Ю. Щербаков 

Аннотация: Целью данной работы была идентификация полиморфных локусов в геномах *Daphnia* с помощью анализа последовательности генома и экспериментального анализа реальных образцов, собранных в естественных условиях. Мы использовали праймеры, предсказанные алгоритмом амплификации ДНК из образцов дафний, и капиллярный электрофорез для выявления полиморфных локусов. Это сделано для детекции полиморфных локусов в результате единственного эксперимента, когда наличие полиморфизмов с частотой, достаточной для генетических исследований, детектируется как появление множественных пиков на электрофореграмме. Таким образом, мы оценивали эффективность метода капиллярного электрофореза и уточняли

Limnological Institute of the Siberian Branch of the Russian Academy of Sciences, Irkutsk, Russia
Лимнологический институт Сибирского отделения Российской академии наук, Иркутск, Россия

 a.poroshina@lin.irk.ru

 Poroshina A.A., Sherbakov D.Yu., 2025

параметры популяции, включая долю полового размножения и эффективную численность популяции. Изучение полиморфных локусов становится особенно актуальным в контексте изменения климата и усиления антропогенного воздействия на экосистемы, поскольку эти факторы могут существенно влиять на генетическое разнообразие и адаптацию популяций. Полиморфные локусы, содержащие микросателлитные повторы, могут служить индикаторами генетической стабильности и способности видов адаптироваться к изменяющимся условиям окружающей среды. В условиях глобального изменения климата и усиливающегося антропогенного воздействия, такого как загрязнение окружающей среды, урбанизация и изменения ландшафта, гидробионты сталкиваются с новыми вызовами, требующими быстрой адаптации. Полиморфные локусы, благодаря своей высокой изменчивости, могут предоставить ценную информацию о генетической структуре популяций и их потенциале к адаптации. В контексте наших исследований изучение полиморфных локусов у *Daphnia* приобретает особую важность, учитывая их экологическую роль и чувствительность к изменениям окружающей среды. Представители рода *Daphnia* как ключевые компоненты пресноводных экосистем могут служить модельными организмами для изучения генетической адаптации к изменению климата и антропогенным воздействиям. Понимание полиморфных локусов у *Daphnia* может дать ценные данные для оценки генетической стабильности и адаптивного потенциала этих популяций, что особенно важно в условиях быстро меняющихся условий окружающей среды.

Ключевые слова: полиморфные локусы; капиллярный электрофорез; агентное моделирование; микросателлиты; дафнии; генетическое разнообразие; адаптация; изменение климата; антропогенные воздействия

Финансирование: Исследование выполнено в рамках темы госзадания № 0279-2021-0010 «Генетика сообществ байкальских организмов: структура генофонда, стратегии консервации».

Introduction and motivation

Polymorphic loci play a crucial role in studying genetic diversity and evolutionary processes in populations. For cladocerans of the genus *Daphnia* O.F. Müller, 1785, which exhibit a mixed reproductive strategy involving both sexual and asexual reproduction, understanding the polymorphism of loci helps estimate the proportion of sexual reproduction and other population parameters. A previously developed agent-based model allowed us to assess genetic diversity and the disequilibrium parameter (D) on simulated data, opening up the possibility of using disequilibrium to estimate the proportion of sexual reproduction in a population (Poroshina, Sherbakov, 2023). This work focuses on identifying polymorphic loci that we are developing to monitor changes in the reproduction strategy, utilizing the previously established agent-based model.

The aim of this work is to identify polymorphic loci in daphniids genomes using genome sequence analysis and experimental analysis of real samples. We use primers predicted by an algorithm for DNA amplification from *Daphnia* samples and capillary electrophoresis to detect polymorphic loci (Thielsch et al., 2012). Thus, we evaluate the effectiveness of the method and refine population parameters, including the proportion of sexual reproduction and the effective population size.

Studying polymorphic loci becomes especially relevant in the context of climate change and anthropogenic impacts on ecosystems, as these factors can significantly affect genetic diversity and adaptation in populations. Polymorphic loci can serve as indicators of genetic stability and the ability of species to adapt to changing environmental conditions. In the face of global climate change and increasing anthropogenic impacts, such as pollution, urbanization, and landscape changes, populations face new challenges that require rapid adaptation. Polymorphic loci, due to their high variability, can provide valuable information about the genetic structure of populations and their potential for adaptation.

Climate change affects ecosystems worldwide, altering temperature regimes, precipitation, and seasonal cycles. These changes can lead to shifts in species ranges, changes in phenology, and even the extinction of some populations. Polymorphic

loci play a key role in the adaptation of species to new climatic conditions (Hoffmann, Sgrò, 2011).

Polymorphism can contribute to the adaptation of species to climate change. Examples where polymorphic loci played a key role in adaptation, such as changes in phenology and physiology in response to climatic changes, highlight the importance of polymorphic loci in evolutionary adaptation and their potential for use in monitoring and managing populations (Franks, Hoffmann, 2012).

The mechanisms of genetic adaptation to environmental changes, including the role of polymorphic loci, are discussed in various studies. Polymorphic loci can influence adaptation through changes in allele frequencies and genetic structure of populations. This emphasizes the importance of studying polymorphic loci for understanding evolutionary processes and developing conservation strategies (Hansen et al., 2012; Yuan et al., 2012; Zhang et al., 2023).

The loss of genetic diversity can affect the survival of species in the face of climate change and anthropogenic impacts. Preserving genetic diversity, including polymorphic loci, is critical for the resilience of species to climate change. This highlights the importance of preserving genetic diversity for the resilience of species to climate change and includes a discussion of polymorphic loci (Kim et al., 2011; Pauls et al., 2013; Engelhardt et al., 2014).

In the context of our research, studying polymorphic loci in *Daphnia* becomes particularly important, given their ecological role and sensitivity to environmental changes. *Daphnia*, as key components of freshwater ecosystems, can serve as model organisms for studying genetic adaptation to climate change and anthropogenic impacts. Understanding polymorphic loci in *Daphnia* can provide valuable data for assessing genetic stability and adaptive potential in these populations, which is particularly important in the face of rapidly changing environmental conditions.

Materials and methods

As model organisms, we used two species of cladocerans: *Daphnia longispina* O.F. Müller, 1785 and *Simocephalus vetulus*

O.F. Müller, 1786. Previously developed primers from the article “Microsatellite markers for European *Daphnia*” (Brede et al., 2006) were tested only on representatives of the genus *Daphnia*. Our task was to test these primers on their closest relatives from the same family, namely *S. vetulus*.

We previously developed an object-oriented model to obtain theoretical data which was published in the article “A procedure for modeling genetic diversity distortions in populations of organisms with mixed reproductive strategies” (Poroshina, Sherbakov, 2023). The theoretical data were verified using various population metrics such as *D*-linkage disequilibrium parameters and *I_a*-index association. The R package poppr was used for this analysis (Kamvar et al., 2014). We compared the analysis of microsatellite state tables with real data from the article “Discrimination of hybrid classes using cross-species amplification of microsatellite loci: methodological challenges and solutions in *Daphnia*” (Thielsch et al., 2012). This comparison was necessary to confirm the correctness of our model. The above work analyzed the species *D. cucullata*, *D. galeata*, and *D. longispina*, while we chose *S. vetulus* in addition to *D. longispina*.

Daphnia longispina employs cyclical parthenogenesis for reproduction, which means that for the majority of the year, they reproduce asexually without male involvement. During this phase, females generate diploid eggs that develop into females without fertilization. However, under stressful conditions such as temperature fluctuations, food scarcity, or other adverse factors, females start producing haploid eggs that require fertilization by males. This process results in the formation of diapausing eggs, which can endure unfavorable conditions and initiate a new generation when conditions improve.

Simocephalus vetulus, on the other hand, also reproduces through cyclical parthenogenesis, similar to *D. longispina*. This means that under favorable conditions, females produce diploid eggs that develop into females without fertilization. However, under stressful conditions, *S. vetulus* can also switch to sexual reproduction, producing haploid eggs that require fertilization by males. Consequently, both *D. longispina* and *S. vetulus* demonstrate cyclical parthenogenesis. These species are ideal for our analysis due to their distinct reproductive strategies. By utilizing the previously developed model and method for assessing the proportion of sexual reproduction based on the states of microsatellite loci (whether homozygous or heterozygous), we can determine the reproductive method of a given population.

Simocephalus vetulus specimens were collected in August 2024 in a lake near the village of Bolshye Koty (coordinates 52.92067, 105.07231), which is characterized by a rocky-muddy bottom, an abundance of *Elodea* sp., *Lemna* sp., *Potamogeton lucens* and a diverse fish population. In the same month, samples of *D. longispina* were collected from a reservoir on the Angara River in Irkutsk (coordinates 52.26504, 104.27950), which has a depth of more than 1 m, a pelagic zone and flooded grasses. The samples were collected using a plankton net with a mesh size of 100 µm to ensure the capture of both adult and juvenile individuals.

Approximately 100 individuals per each species were collected for the study. DNA extraction was performed using a classical phenol-chloroform extraction protocol (Abdulla, 2014;

Ye, Lei, 2023). For each species, 10 individuals were pooled to create a mixed sample, resulting in 5 mixed samples per species. The extracted DNA was amplified using previously designed microsatellite primers labeled with four different fluorescent dyes: TAMRA, R6G, ROX, and FAM. The PCR amplification was carried out in a 20 µL reaction volume containing 10 ng of genomic DNA, 1 × PCR buffer, 2 mM MgCl₂, 0.2 mM dNTPs, 0.2 µM of each primer, and 1 unit of Taq DNA polymerase. The PCR conditions included an initial denaturation step at 94 °C for 5 minutes, followed by 35 cycles of denaturation at 94 °C for 30 seconds, annealing at temperatures ranging from 53 to 57 °C for 30 seconds (as specified in the original article by Brede N. et al., 2006), and extension at 72 °C for 45 seconds, with a final extension step at 72 °C for 10 minutes.

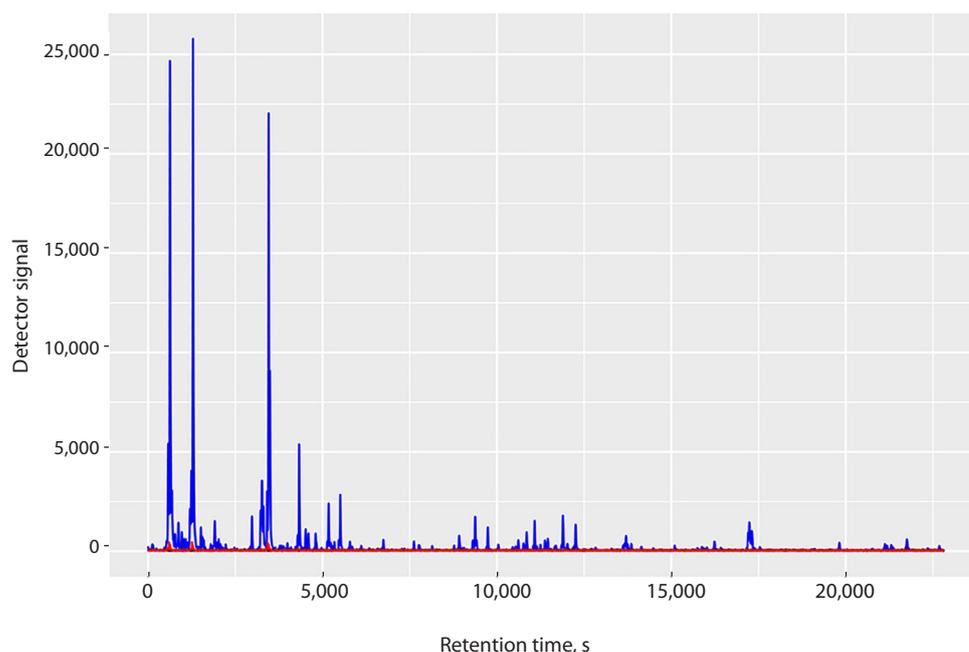
The PCR products were analyzed using the NANOFOR-05 genetic analyzer to identify polymorphic primers (Shewale et al., 2012). The primary objective of this study was to identify primers that exhibit polymorphism in both *D. longispina* and *S. vetulus*. The analysis yielded a set of polymorphic primers, which are listed in the Table below. These primers will be utilized in subsequent experiments to further investigate the genetic diversity and population dynamics of these species.

Below is a Table listing 5 pairs of polymorphic primers for the two species that we will use subsequently. The table presents the results of the polymorphism analysis of 19 primers for two species of cladocerans: *Simocephalus vetulus* (denoted as s1–s5) and *Daphnia longispina* (denoted as d1–d4).

Polymorphic primers for *Daphnia longispina* and *Simocephalus vetulus*

Primer	s1	s2	s3	s4	s5	d1	d2	d3	d4
DaB17/16(R6G)	-	-	-	+	-	-	+	-	-
DaB10/14(TAMRA)	-	-	+	-	-	+	-	-	-
Dgm101(ROX)	-	+	+	+	+	+	-	+	-
Dp281(FAM)	+	-	+	+	-	-	-	-	+
SwiD15(R6G)	+	+	+	?	+	-	-	+	+
SwiD18(TAMRA)	+	+	-	-	-	-	+	-	+
DaB10/15(ROX)	+	-	+	-	-	+	-	+	+
DaB17/17(FAM)	-	?	+	+	-	-	+	+	-
SwiD6(R6G)	-	-	+	-	-	+	-	-	-
SwiD2(TAMRA)	-	+	-	-	+	-	-	-	-
Dgm105(ROX)	+	-	-	+	-	-	-	-	-
SwiD18(FAM)	-	-	-	-	-	-	-	-	-
SwiD14(TMARA)	-	-	-	-	-	+	+	-	-
SwiD12(R6G)	-	-	-	-	-	-	+	-	-
SwiD6(ROX)	+	+	-	-	-	-	-	-	-
Dgm112(R6G)	-	-	-	-	-	-	+	+	-
Dgm109(TAMRA)	-	-	-	-	-	-	+	+	+
Dp519(FAM)	-	-	-	-	-	-	-	-	-
Dp196(ROX)	-	-	-	-	-	-	-	-	-

Note. The “+” sign indicates the presence of polymorphism, the “-” sign indicates its absence, and “?” denotes uncertain results. Primers that are suitable for both species are highlighted in bold.



Electropherogram of primer DaB17/17 obtained using capillary electrophoresis.

The graph shows the dependence of the detector signal on the retention time for the primer DaB17/17. The blue peaks indicate the presence of polymorphic loci, demonstrating differences in the lengths of DNA fragments amplified by this primer. The red line represents the baseline signal level.

Results

As part of this work, we conducted a comprehensive study to search for polymorphic loci in the genomes of *Daphnia*. As a result of the study, out of 19 previously developed microsatellite markers, 8 were identified as polymorphic in the *Daphnia* genomes. Of these, 5 are polymorphic for both *D. longispina* and *S. vetulus*. These polymorphic markers will be used in the future to determine the breeding strategy within *Daphnia* populations. We analyzed capillary electrophoresis graphs constructed from .fas files using the SeqinR package (Charif, Lobry, 2007) (see the Figure).

Discussion

Polymorphic markers are crucial indicators of genetic stability and adaptation potential in species, especially under the pressures of global climate change and increasing anthropogenic impacts like pollution and urbanization. These markers, due to their high variability, offer valuable insights into population genetics and adaptive capabilities.

Global climate change is altering ecosystems by modifying temperature patterns, rainfall, and seasonal cycles. Such changes can shift species' geographic ranges, affect life cycle events, and potentially lead to population declines. Polymorphic loci significantly influence a species' ability to adapt to these new conditions (Hoffmann, Sgrò, 2011).

Polymorphic loci play a key role in facilitating phenological and physiological adaptations, as demonstrated in studies showing how genetic variations drive adaptation through allele frequency modifications (Franks, Hoffmann, 2012; Hansen et al.,

2012). Preserving genetic diversity is essential for species resilience against climate change and human activities (Kim et al., 2011; Pauls et al., 2013).

Our study focuses on polymorphic loci in *Daphnia*, given their ecological importance and sensitivity to environmental shifts. As key components of freshwater ecosystems, *Daphnia* serve as model organisms for investigating genetic responses to climate change and anthropogenic influences. Understanding these loci provides critical data on genetic stability and adaptive potential, vital for assessing how *Daphnia* populations will respond to rapidly changing environments.

A limitation of our study is the potential influence of seasonal changes on the genetic diversity of *Daphnia* populations.

Conclusion

In conclusion, we have identified polymorphic markers suitable for estimation of genetic status and adaptive potential in *Daphnia* populations. These markers provide essential insights into genetic structure and diversity, are crucial for understanding adaptation to changing conditions. Future research using these markers can monitor population dynamics and assess responses in various ecosystems, such as lakes and rivers. This emphasizes the importance of preserving genetic diversity to ensure species resilience in the face of climate change and anthropogenic pressures.

References

Abdulla A.A. Optimization of DNA extraction of *Lactobacillus* spp for identification by *tufB* gene-based polymerase chain reaction. *J Biol Agric Healthcare*. 2014;4(8):122-127

- Brede N., Thielsch A., Sandrock C., Spaak P., Keller B., Streit B., Schwenk K. Microsatellite markers for European *Daphnia*. *Mol Ecol Notes*. 2006;6(2):536-539. doi 10.1111/j.1471-8286.2005.01218.x
- Charif D., Lobry J.R. SeqinR 1.0-2: a contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In: Structural Approaches to Sequence Evolution: Molecules, Networks, Populations. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007;207-232
- Engelhardt K.A.M., Lloyd M.W., Neel M.C. Effects of genetic diversity on conservation and restoration potential at individual, population, and regional scales. *Biol Conserv*. 2014;179:6-16. doi 10.1016/j.biocon.2014.08.011
- Franks S.J., Hoffmann A.A. Genetics of climate change adaptation. *Ann Rev Genet*. 2012;46(1):185-208. doi 10.1146/annurev-genet-110711-155511
- Hansen M.M., Olivieri I., Waller D.M., Nielsen E.E. Monitoring adaptive genetic responses to environmental change. *Mol Ecol*. 2012;21(6):1311-1329. doi 10.1111/j.1365-294X.2011.05463.x
- Hoffmann A.A., Sgrò C.M. Climate change and evolutionary adaptation. *Nature*. 2011;470(7335):479-485. doi 10.1038/nature09670
- Kamvar Z.N., Tabima J.F., Grünwald N.J. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*. 2014;2:e281. doi 10.7717/peerj.281
- Kim M.S., Richardson B.A., McDonald G.L., Klopfenstein N.B. Genetic diversity and structure of western white pine (*Pinus monticola*) in North America: a baseline study for conservation, restoration, and addressing impacts of climate change. *Tree Genet Genomes*. 2011;7:11-21. doi 10.1007/s11295-010-0311-0
- Pauls S.U., Nowak C., Bálint M., Pfenninger M. The impact of global climate change on genetic diversity within populations and species. *Mol Ecol*. 2013;22(4):925-946. doi 10.1111/mec.12152
- Poroshina A., Sherbakov D. A procedure for modeling genetic diversity distortions in populations of organisms with mixed reproductive strategies. *Mathematics*. 2023;11(13):2985. doi 10.3390/math11132985
- Shewale J.G., Qi L., Calandro L.M. Principles, practice, and evolution of capillary electrophoresis as a tool for forensic DNA analysis. *Forensic Sci Rev*. 2012;24(2):79-100
- Thielsch A., Völker E., Kraus R.H.S., Schwenk K. Discrimination of hybrid classes using cross-species amplification of microsatellite loci: methodological challenges and solutions in *Daphnia*. *Mol Ecol Resour*. 2012;12(4):697-705. doi 10.1111/j.1755-0998.2012.03142.x
- Ye X., Lei B. The current status and trends of DNA extraction. *Bioessays*. 2023;45(8):2200242. doi 10.1002/bies.202200242
- Yuan X., Miller D.J., Zhang J., Herrington D., Wang Y. An overview of population genetic data simulation. *J Comput Biol*. 2012;19(1):42-54. doi 10.1089/cmb.2010.0188
- Zhang K., Wu J., Deng D., Zhao H., Liu Q., Peng S., Zhang Y., Zhou Z. Population genetic differentiation of *Daphnia sinensis* in a lasting high-phosphorus Chinese lake, Lake Chaohu. *Front Ecol Evol*. 2023;10:913738. doi 10.3389/fevo.2022.913738

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

Received November 2, 2024. Revised December 2, 2024. Accepted December 16, 2024.

 pismavavilov.ru

doi 10.18699/letvjgb-2025-11-04

Обзор

Полвека популяционной генетики народов Сибири: исследования Людмилы Павловны Осиповой

И.К. Захаров¹, О.Л. Посух^{1,2} 

Аннотация: Полувековая жизнь в науке Л.П. Осиповой была тесно связана со становлением, организацией и развитием комплексных и широкомасштабных исследований генетической структуры популяций человека Северной Евразии. Были инициированы и проводились исследования в различных направлениях популяционной и эволюционной генетики человека, молекулярной генетики и иммуногенетики, демографии и биомедицины коренных народов; изучались механизмы адаптации человека к факторам окружающей среды, а также влияние техногенных факторов на окружающую среду и на здоровье человека. Совокупность результатов масштабных геномных исследований международных консорциумов с участием коллектива Л.П. Осиповой позволила реконструировать предковые геномы и воссоздать пути миграции человека в историческом прошлом.

Ключевые слова: популяционная генетика; генетическая демография; генетическая изменчивость; адаптация; коренные народы; Северная Евразия; этногенетика; Л.П. Осипова

Для цитирования: Захаров И.К., Посух О.Л. Полвека популяционной генетики народов Сибири: исследования Людмилы Павловны Осиповой. *Письма в Вавиловский журнал генетики и селекции*. 2025;11(1):17-28. doi 10.18699/letvjgb-2025-11-04

Review

Half a century of population genetics of the peoples of Siberia: research by Lyudmila P. Osipova

I.K. Zakharov¹, O.L. Posukh^{1,2} 

Abstract: L.P. Osipova's life in science lasted half a century and was closely connected with the establishment, organization and development of complex and large-scale studies of the genetic structure of human populations in Northern Eurasia. Her research was initiated and conducted in various areas of human population and evolutionary genetics, molecular genetics and immunogenetics, demography and biomedicine of indigenous peoples; the mechanisms of human adaptation to environmental factors were studied, as well as the impact of technogenic factors on the environment and human health. The totality of the results of large-scale genomic studies of international consortiums with the participation of L.P. Osipova's team made it possible to reconstruct ancestral genomes and recreate human migration routes in the historical past.

Key words: population genetics; genetic demography; genetic variation; adaptation; indigenous peoples; Northern Eurasia; ethnogenetics; L.P. Osipova

For citation: Zakharov I.K., Posukh O.L. Half a century of population genetics of the peoples of Siberia: research by Lyudmila P. Osipova. *Pisma v Vavilovskii Zhurnal Genetiki i Seleksii = Letters to Vavilov Journal of Genetics and Breeding*. 2025;11(1):17-28. doi 10.18699/letvjgb-2025-11-04 (in Russian)

5 сентября 2024 г. ушла из жизни заведующая лабораторией популяционной этногенетики ИЦиГ СО РАН кандидат биологических наук Людмила Павловна Осипова – известный специалист в многочисленных областях исследований, посвященных человеку: популяционная и эволюционная генетика человека, генетическая демография, молекулярная биология, иммуногенетика и биомедицина, изучение механизмов адаптации человека, оценка влияния техногенных факторов на окружающую среду и на здоровье человека.

Изучение генетической структуры коренных малочисленных популяций человека Сибири и Крайнего Севера стало делом всей ее жизни.

Людмила Павловна Осипова (в девичестве Ехаева) родилась 17 сентября 1946 г. в поселке Чалдай Цюрупинского района Павлодарской области Казахской ССР. Ее детские годы были чередой смены местожительства и, соответственно, школ, в которых она обучалась. В юности формировались многие черты характера Людмилы: ответственность,

¹ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

² Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
Novosibirsk State University, Novosibirsk, Russia

 posukh@bionet.nsc.ru

 Захаров И.К., Посух О.Л., 2025



Людмила Павловна Осипова

целестремленность, тяга к получению новых знаний, активная жизненная позиция. Интересы не ограничивались школьными рамками: занятия спортивной гимнастикой, туризмом, а в старших классах она была избрана в состав Экибастузского горкома ВЛКСМ. Она много читала – книги и свежая пресса сопровождали ее всегда. В частности, Людмилу увлекали научно-популярные издания по физиологии и медицине. Ей также были интересны все грани многогранной жизни людей – социально-бытовые, общественно-политические, культурные, биологические.

Вся дальнейшая судьба Людмилы Павловны будет связана с Новосибирском, с Сибирью. В 1964 г., после окончания средней школы № 1 города Экибастуза с золотой медалью она поступила на физический факультет Новосибирского государственного университета. К концу первого курса у нее созрело твердое убеждение: «физика – не мое призвание...» И в 1965 г. Людмила вновь сдала вступительные экзамены в НГУ, но в этот раз выбрала медико-биологическое отделение факультета естественных наук. В 1971 г. Людмила Павловна окончила НГУ по специальности «биология» с присвоением квалификации «генетик».

Дипломную работу она выполнила под руководством Веры Георгиевны Матвеевой в Институте цитологии и генетики СО АН СССР в лаборатории радиационной генетики, возглавляемой д. б. н. проф. Юлием Яковлевичем Керкисом. Первой опубликованной статьей Л.П. Осиповой стала «Мутагенный эффект пестицида ДДБ и хлорида в костном мозге и половых клетках мышей» (Матвеева и др., 1973). Было ясно, что ее жизнь будет связана с наукой. Казалось,

что место работы определено – в лаборатории ИЦиГ, где она выполняла курсовую и дипломную работу, были готовы принять увлеченного наукой, умного и пытливого молодого работника. Однако при распределении возникла непреодолимая преграда. После четвертого курса Людмила трудилась в составе студенческого стройотряда на Чукотке. Студенческие стройотряды – это не только возможность освоить рабочие профессии, не только юношеская и трудовая романтика, но и способ заработать, существенная материальная поддержка для «бедного» студента. Отряд задержался с возвращением к началу учебного года в Новосибирск. Этот факт возмутил директора ИЦиГ и зав. кафедрой цитологии и генетики НГУ Д.К. Беляева – возможность принять Л.П. Осипову в институт исключалась...

В 1970 г. в Новосибирске создается Сибирский филиал Академии медицинских наук СССР, призванный скоординировать и интегрировать медико-биологические и медицинские исследования, проводимые в Сибири. Здесь получают свое развитие исследования по генетике человека. В Институте клинической и экспериментальной медицины (ИКЭМ) СФ АМН СССР была открыта лаборатория генетических основ адаптации человека (впоследствии лаборатория популяционной генетики человека), организатором и заведующим которой стал врач по образованию, позднее д.б.н. проф. Рем Израилевич Сукерник (1934–2025). В дальнейшем лаборатория приобрела статус межведомственной (ИКЭМ СО АМН СССР, куратор академик АМН В.П. Казначеев, и ИЦиГ СО АН СССР, куратор академик Д.К. Беляев). По инициативе Р.И. Сукерника начинают разворачиваться медико-биологические и генетические исследования популяций малочисленных коренных народов Сибири. Наряду с поиском научных подходов, освоением и разработкой методов исследований, Сукернику приходилось решать и непростые кадровые вопросы при организации и становлении лаборатории. В августе 1971 г. Людмила Павловна была принята в лабораторию, сначала на должность старшего лаборанта, а с декабря 1973 г. – младшего научного сотрудника.

Научная тематика лаборатории была направлена на изучение механизмов адаптации и анализа генетической изменчивости популяций коренных народов Сибири. Исследования на тот момент базировались на анализе биохимических эритроцитарных и сывороточных маркеров (полиморфные системы групп крови, изоферменты, сывороточные белки, аллотипы иммуноглобулинов). Кроме того, в лаборатории работали генетики-демографы (Т.В. Гольцова и Т.А. Абанина), и генетические исследования сопровождались сбором демографической информации: анализировались родственные связи, состав семей, рождаемость, смертность, другие демографические параметры, составлялись детальные родословные. Анализ всех этих данных позволял описать генетико-демографическую структуру популяций народов Севера.

В 1980–1990-е годы благодаря Р.И. Сукернику у лаборатории налаживались контакты с зарубежными специалистами в области популяционной генетики и иммуногенетики человека: из США – М. Кроуфорд (M.H. Crawford) и М. Шенфилд (M.S. Schanfield), Японии – Х. Матсумото (H. Matsumoto), Франции – М.П. Лефранк (M.-P. Lefranc) и Ж. Лефранк



Студентка физфака НГУ
Людмила Павловна Ехаева



Л.П. Осипова в лаборатории генетических основ адаптации человека



Слева направо: зав. лабораторией Р.И. Сукерник, сотрудники лаборатории Л.П. Осипова, В.П. Вибе, Т.М. Карафет
и зарубежный гость проф. М. Шенфилд (M.S. Schanfield). Академгородок, около Дома ученых. Май 1984 г.

(G. Lefranc). Некоторые из них приезжали в Новосибирск для знакомства и обсуждения совместных исследований, результаты которых были опубликованы в серии научных статей (см. в Приложении «Список избранных публикаций Л.П. Осиповой»)¹.

Популяционно-генетические исследования лаборатории были связаны с организацией и проведением ежегодных экспедиций в отдаленные, труднодоступные и малонаселенные районы Сибири и Крайнего Севера с суровыми климатическими условиями. По словам Людмилы Павловны, при приеме ее на работу Р.И. Сукерника больше всего интересовал вопрос, сможет ли она ездить в экспедиции. Ответ был предсказуем. Успех экспедиций зависел от налаживания контактов с органами местного самоуправления для помощи в организации работы, с медицинским персоналом на местах для содействия в работе с приглашенными на обследование людьми, с сотрудниками транспортных организаций (авиация, автомобильный и речной транспорт) для перемещений экспедиционного отряда.

В 1987 г. приказом Министерства здравоохранения СССР лаборатория популяционной генетики человека формально была переведена в красноярский Научно-исследовательский институт медицинских проблем Севера СО АМН СССР.

В феврале 1988 г. на диссертационном совете при ИЦиГ СО АН СССР Людмила Павловна успешно защитила диссертацию «Полиморфизм аллотипов иммуноглобулинов (система Gm) в антропологических изолятах Северной Сибири» на соискание ученой степени кандидата биологических наук, в которой суммированы данные, полученные в популяциях лесных ненцев Ямала и нгансан Таймыра.

В 1989 г. Л.П. Осипову переводят на должность и. о. старшего научного сотрудника. За время работы в лаборатории популяционной генетики человека она стала авторитетным, признанным специалистом-генетиком по изучению популяций коренных сибирских народов и являлась соавтором около двух десятков статей. Определилась и сформировалась область научных интересов Л.П. Осиповой: генетика человека; популяционная генетика; демография; коренные народы Сибири.

В феврале 1990 г. лаборатория популяционной генетики человека была переведена в Институт цитологии и генетики СО АН СССР, а в сентябре 1991 г. претерпела коренную реорганизацию – произошло ее разделение. На базе лаборатории популяционной генетики человека был организован сектор молекулярной эволюции человека, возглавляемый Р.И. Сукерником, а группа сотрудников в составе Л.П. Осиповой, Т.М. Карафет, О.Л. Посух, В.П. Вибе, С.А. Пьянкова и М.А. Казаковцевой перешла в лабораторию теории молекулярной эволюции (заведующий С.Н. Родин). В ноябре 1991 г. лабораторию теории молекулярной эволюции переименовали в лабораторию молекулярной эволюции, в ее составе был создан сектор молекулярной и эволюционной генетики человека. Руководителем сектора, а в последующем лаборатории была назначена Людмила Павловна, и с тех пор она оставалась бессменным руководителем подразделения. Это назначение стало временем ответственности за судьбы

других людей, периодом становления и раскрытия таланта организатора научного поиска. Позднее, в 2010 г., лабораторию молекулярной и эволюционной генетики человека переименовали в лабораторию популяционной этногенетики. В 2017 г. Л.П. Осипова была переведена на должность ведущего научного сотрудника с возложением на нее обязанностей заведующей лабораторией популяционной этногенетики.

Первоначально тематика научных исследований возглавляемой Осиповой лаборатории оставалась в русле направлений, разработанных под руководством Р.И. Сукерника в лаборатории популяционной генетики человека. Здесь уместно подчеркнуть, что Людмила Павловна всегда с большим уважением и почтением относилась к своему научному учителю – Рему Израилевичу Сукернику. Но за три десятилетия работы лаборатории под руководством Л.П. Осиповой тематика существенно расширилась и обогатилась исследованиями в области медицины, физиологии, экологии за счет привлечения новых подходов и методов, расширения сотрудничества и кооперации с отечественными и зарубежными коллегами и коллективами ученых.

В тяжелые для российской науки 90-е годы прошлого века (недостаток финансирования, массовая эмиграция российских ученых) Людмиле Павловне, несмотря на многие сложности, удавалось поддерживать работу лаборатории. «Утечка мозгов» не обошла стороной и ее коллектив – часть сотрудников уехала в США, Германию, Канаду. Однако важно подчеркнуть, что все годы с ними сохранялась связь и продолжалось научное сотрудничество.

В России исследования лаборатории, наряду с бюджетным финансированием, поддерживались многочисленными грантами РФФИ, Интеграционными проектами СО РАН, экспедиционными грантами СО РАН. Финансовую поддержку в разные годы оказывали Федеральные целевые научно-технические программы, Комплексные программы фундаментальных исследований СО РАН, РФФИ. Специфика исследований определяла и возможность финансовой поддержки со стороны администраций северных регионов (хоздоговорные работы). Одни направления исследований коллектива, возглавляемого Людмилой Павловной, проводились при наличии целевого финансирования непрерывно, они оставались «стержневыми, сквозными», другие – в определенные периоды времени, либо в рамках финансирования совместных научных программ, либо как результат личной инициативы и заинтересованности различных специалистов.

Людмила Павловна обладала талантом налаживать и поддерживать долгое время научную кооперацию с широким кругом ученых и специалистов разных научных направлений: молекулярными биологами и генетиками, врачами и физиологами, антропологами и археологами, этнографами и демографами, геологами и экологами, физиками и биоинформатиками. Она притягивала людей. Ею были установлены и сохранялись многолетние творческие связи с к.б.н. В.Г. Матвеевой, Т.В. Гольцовой, академиком РАН Л.Е. Паниным, д.б.н. В.Г. Селятицкой, д.б.н. Н.Б. Пиковской, д.х.н. Г.А. Невинским, д.б.н. М.Л. Филипенко, к.б.н. Е.Н. Ворониной, Т.Н. Ткаченко, д.б.н. Э.К. Хуснутдиновой, к.б.н. В.Л. Ко-

¹ Приложение см. по адресу:

<https://pismavavilov.ru/wp-content/uploads/2025/03/appx1.pdf>



Л.П. Осипова с проф. Х. Матсумото. Осака, Япония



Работа в лаборатории М.-П. Лефранк. Монпелье, Франция

рогодиной, академиком РАН д.б.н. М.И. Воеводой, д.и.н. И.В. Октябрьской, к.б.н. В.А. Мануйловым, к.б.н. И.Г. Нетесовой, академиком РАН С.В. Нетесовым, д.м.н. С.А. Петровым, д.ф.-м.н. К.П. Куценогим, к.г.-м.н. Ф.В. Сухоруковым, к.г.-м.н. Б.Л. Щербовым, д.г.-м.н. В.Д. Страховенко, д.б.н. Т.И. Аксенович, д.м.н. проф. Н.К. Поповой, д.б.н. А.В. Куликовым – вот далеко не полный перечень соавторов и соисполнителей/соруководителей совместных, как правило комплексных, исследований.

Совместные работы проводились не только с отечественными учеными. Финансирование российской науки в последнее десятилетие прошлого века и в нулевых двухтысячных ставило ее на грань выживания. Но Л.П. Осиповой удавалось поддерживать свои исследования благодаря налаживанию научных связей с зарубежными коллегами и иностранными фондами (Франция, Япония, США, Канада, Эстония, Германия). В результате плодотворного международного сотрудничества в разные годы была достигнута финансовая поддержка за счет различных грантов и соглашений. Это позволяло проводить эксперименты как в лаборатории, так и во время зарубежных командировок, как самой Людмиле Павловне, так и ее сотрудникам.

Так, например, сотрудничество с проф. М.П. Лефранк (M.-P. Lefranc, лаборатория молекулярной иммуногенетики, Университет Монпелье II, Франция) и проф. Х. Матsumото (H. Matsumoto, Медицинский колледж Осаки, Япония) обеспечило выполнение многих исследований в области изучения генов иммуноглобулинов человека (Engström et al., 1996; Osipova et al., 1999; Dard et al., 2001). Совместные популяционно-генетические исследования проводились с проф. М. Кроуфордом (M.H. Crawford, Университет Канзаса, г. Лоуренс, США) (McComb et al., 1996).

Многолетнее плодотворное сотрудничество с проф. М.Ф. Хаммером (M.F. Hammer) и к.б.н. Т.М. Карафет (лаборатория молекулярной систематики и эволюции, Университет Аризоны, г. Тусон, США) по изучению генетического разнообразия сибирских этнических групп по маркерам Y-хромосомы привело к появлению высокоцитируемых статей о формировании генетического ландшафта коренных народов на территории Сибири и об актуальных аспектах раннего заселения Нового Света мигрантами с территории Сибири (Karafet et al., 1997, 1999, 2002). В сотрудничестве с проф. Р. Виллемсом (R. Villems, Эстонский биоцентр, г. Тарту, Эстония) и проф. Т. Шурром (T.G. Schurr, кафедра антропологии, Пенсильванский университет, г. Филадельфия, США) осуществлялись исследования по анализу маркеров митохондриальной ДНК и Y-хромосомы в сибирских популяциях (Reidla et al., 2003; Tamm et al., 2007; Dulik et al., 2012). С д-ром В.Р. Леонардом (W.R. Leonard, кафедра антропологии, Северо-Западный университет, г. Эванстон, США) были организованы совместные экспедиции и изучались особенности основного метаболизма коренного населения Сибири при адаптации к северному климату (Leonard et al., 1999, 2002).

Экспедиции и биологические коллекции

Научным подвигом Л.П. Осиповой стали сбор и создание в результате экспедиций лаборатории уникальной коллекции образцов биологического материала, полученных от пред-

ставителей разных народов из географически удаленных районов Сибири и Крайнего Севера. Первая экспедиция Людмилы Павловны состоялась еще в далеком 1974 г. на Алтай, к северным алтайцам, в поселок Курмач-Байгол. С тех пор она – организатор (начальник отряда) и участник около ста экспедиционных выездов в различные регионы Сибири и Крайнего Севера. Исследованные популяции коренных народов: лесные и тундровые ненцы, селькупы, нганасаны, долганы, коми, ханты, кеты, алтайцы (северные и южные), телеуты, якуты, юкагиры, эвены, эвенки, буряты, тувинцы, а также казахи Алтая, русские староверы и русские Сибири (из разных локаций).

Обеспечение успешной работы в нелегких экспедиционных условиях – непростая задача, которую Людмила Павловна успешно решала благодаря своему таланту налаживать контакты с местными администрациями и органами здравоохранения, помогавшими в транспортировке экспедиционного отряда и организации работы в поселковых медицинских учреждениях. Кроме того, работа с представителями местного населения (детальный опрос, получение генеалогических и других сведений, забор биоматериала для генетических исследований) требует отдельного подхода, человеколюбия, такта, умения расположить к себе – всех тех качеств, которыми Людмила Павловна обладала как никто другой. Со многими людьми «на местах» у Людмилы Павловны установились хорошие теплые отношения, она помогала советами, помогала организовать врачебную консультацию и лечение в новосибирских клиниках.

Фотоаппарат был неотъемлемым атрибутом в жизни Людмилы Павловны, сопровождал ее во всех путешествиях. Фотографиями она радовала местных жителей при возвращении в места экспедиционных работ, многим отсылала фотографии по почте. Особенно это было актуально в первые десятилетия ее путешествий – до появления фотоаппаратов «мыльниц» и современного повального увлечения «селфи» – когда собственные фотографии для многих коренных жителей часто были первыми в их жизни.

Важным является не только сбор, но и организация долговременной сохранности банка биологических образцов человека (образцов крови и другого биологического материала, образцов ДНК) в Институте цитологии и генетики СО РАН. В настоящее время этот банк насчитывает более 15 тыс. образцов, хранящихся в низкотемпературных условиях. Следует отметить, что все образцы сопровождаются детальной информацией об этнической принадлежности, о географической локализации, детальными демографическими данными. Для доноров подавляющего большинства образцов биологического материала известна генеалогия и составлены родословные, имеются сведения о состоянии здоровья. Для Людмилы Павловны и ее сотрудников скрупулезность при описании собираемого материала была непреложным законом. Все это обеспечивает уникальность информационной базы и банка биологического материала человека в лаборатории популяционной этногенетики ИЦиГ СО РАН, не только в пределах России, но и во всем мире. Образцы использовались в экспериментальной работе для изучения по различным научным направлениям. Часть образцов анализировалась на первых этапах на уровне био-



Сбор информации и душевные беседы



Л.П. Осипова с семьей тундровиков. Село Гыда, Тазовский район ЯНАО



На празднике. Село Самбург, Пуровский район ЯНАО

химических маркеров, однако с появлением молекулярных методов анализа в лаборатории активно развивались исследования и на уровне ДНК. Ценность и востребованность представленных в коллекции образцов со временем только возрастает, ведь существенная часть коллекции в настоящее время уже невозможна.

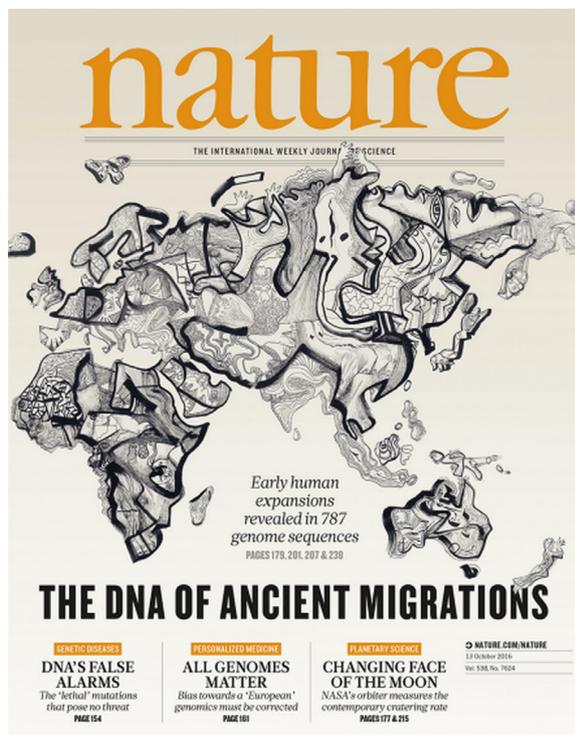
Публикации Л.П. Осиповой

Крайне сложно охарактеризовать в деталях научные интересы Людмилы Павловны. Она – автор и соавтор около 350 научных публикаций. Среди них более 200 статей в отечественных и международных научных журналах. Уже по названиям этих журналов (а их около 50) можно судить о широчайшем спектре научных направлений, по которым велись исследования. Приведем здесь только некоторые журналы из этого обширного списка: отечественные – «Генетика», «Молекулярная биология», «Медицинская генетика», «Сибирский экологический журнал», «Информационный вестник ВОГиС»/«Вавиловский журнал генетики и селекции», «Физиология человека», «Терапевтический архив», «Оптика атмосферы и океана», «Молекулярная генетика, микробиология и вирусология»; зарубежные – «Nature», «Science», «Human Genetics», «American Journal of Human Genetics», «American Journal of Physical Anthropology», «Pharmacogenetics», «International Journal of Immunogenetics», «Human Biology», «Genome Research», «PLoS Genetics» и др.

В Приложении приведен список избранных научных работ Л.П. Осиповой, сгруппированных по разделам в соответствии с научными направлениями исследований (при ее участии или под ее руководством).

В раздел 1 «Популяционная генетика (группы крови и другие биохимические маркеры). Генетическая демография» вошли избранные статьи Л.П. Осиповой и ее коллектива, подготовленные по результатам изучения популяционно-генетической структуры коренных народов Северной Азии на уровне биохимических маркеров (группы крови, изоферменты, сывороточные белки, иммуноглобулины). Важно отметить, что особое внимание всегда уделялось тщательному сбору демографических данных, что позволяло анализировать динамику изучаемых популяций и другие генетико-демографические параметры (половозрастная структура, уровень метисации, репродукция).

Раздел 2 «Митохондриальная ДНК, Y-хромосома и геномные данные в свете эволюционно-исторических исследований происхождения и миграций популяций человека» включает в себя статьи об изучении маркеров митохондриальной ДНК, Y-хромосомы и полногеномных данных в популяциях человека, в которых Л.П. Осипова была соавтором (включая пять статей в журнале «Nature» и одну статью в «Science»). Развитие методов секвенирования как современных, так и древних геномов человека и объединенные усилия ученых со всего мира, в том числе российских, в области эволюционно-исторических исследований происхождения и миграций популяций человека привели к появлению целого ряда публикаций в высокорейтинговых журналах. Так, например, важность новых данных о происхождении популяций Евразии подчеркивает тот факт, что художественные иллюстрации к этим статьям были размещены на обложках журнала «Nature» (2014 и 2016 гг.) (Lazaridis et al., 2014; Pagani et al., 2016).



Обложка журнала "Nature" (2014 г.) с иллюстрацией к статье Lazaridis et al. "Ancient human genomes suggest three ancestral populations for present-day Europeans"



Обложка журнала "Nature" (2016 г.) с иллюстрацией к статье Pagani et al. "Genomic analyses inform on migration events during the peopling of Eurasia"



Тематические выпуски журналов, сборники научных трудов, главы в монографиях с участием Л.П. Осиповой

В разделе 3 «Влияние природных и техногенных факторов среды на здоровье и генофонд популяций человека» представлены избранные публикации Л.П. Осиповой о результатах изучения влияния природных и техногенных факторов среды на здоровье и генофонд популяций человека. В конце 1990-х годов Людмила Павловна была координатором крупного интеграционного проекта СО РАН «Влияние техногенной нагрузки на биогеохимические циклы, биоценозы, биологическое здоровье и генофонд коренных жителей Ямало-Ненецкого автономного округа», в котором участвовало восемь институтов СО РАН (ИЦиГ, НИБХ, ОИГГМ, ИХКГ, ИНХ, НИОХ, ИВМиМГ, ИЯФ). В рамках этого проекта было проведено множество экспедиций. При широкой меж-

дисциплинарной кооперации исследовались химический состав атмосферных аэрозолей, осадков и поверхностных вод, различные характеристики (содержание радионуклидов и тяжелых металлов и др.) биогеохимических объектов, включая пищевые цепочки, показатели здоровья населения (медицинские данные, анализы крови, в том числе цитогенетические и иммунологические исследования).

Людмилу Павловну всегда интересовала научная область, связанная с изучением адаптационных особенностей и возможностей коренных жителей Сибири, постоянно живущих в суровых климатических условиях, и в разделе 4 «Изучение адаптационных возможностей северных народов» сгруппированы публикации, посвященные этой тематике.

В разделе 5 «Биомедицина. Медицинская генетика. Физиология. Вирусология» собраны многочисленные публикации, касающиеся так или иначе широкого спектра различных аспектов здоровья, медицинской генетики, физиологии у коренных народов Сибири. В сотрудничестве с различными специалистами Л.П. Осипова и ее коллектив анализировали распространенность гельминтозов, гормональный статус, особенности липидного состава сыворотки крови, полиморфизм факторов системы свертывания крови и ряда генов серотониновой системы у представителей коренного и пришлого населения. Многие статьи связаны с изучением полиморфизма в генах системы биотрансформации ксенобиотиков, в генах, участвующих в процессах свертывания крови, метаболизма липидов и глюкозы, иммунитета у представителей коренных народов Сибири. Часть работ подготовлена в рамках проекта РНФ «Исследование генетических детерминант мультифакторных заболеваний у коренных народов Сибири» (2019–2021 гг.) под руководством Л.П. Осиповой.

Есть целый цикл работ (раздел 6), выполненных совместно с исследователями в области изучения генетического разнообразия вируса гепатита В на территории Сибири.

Результаты работ Л.П. Осиповой опубликованы также в многочисленных трудах отечественных и зарубежных конференций. Она принимала участие в подготовке и редактировании различных тематических выпусков журналов и сборников научных трудов, глав в книгах, монографиях, в том числе зарубежных, в которых обобщался и систематизировался материал, полученный в лаборатории и/или в сотрудничестве с другими учреждениями в рамках выполнения программ, договоров, личной инициативы (раздел 7).

Популяризатор науки

Людмила Павловна обладала незаурядным талантом просветителя и популяризатора генетических исследований, проводимых в ее лаборатории. Подтверждением этому служат научно-популярные статьи в средствах массовой информации, выступления и интервью на радио и по телевидению. Телевизионный фильм «Иду на зов» (сентябрь 2007 г., КМУ ТК «Альянс», главный редактор Т.И. Чернышова, научный консультант Л.П. Осипова) получил Гран-при на окружном фестивале 2007 г. в Салехарде.

Из экспедиционных путешествий она привозила множество фотографий, которыми украшала свои публикации и доклады на конференциях и тем самым знакомила читателей и слушателей с жизнью и бытом коренных жителей Севера.

Образовательная деятельность

В 1992–2009 гг. Л.П. Осипова читала авторский спецкурс «Генетика человека» объемом 20 ч для студентов четвертого курса, специализирующихся на кафедре цитологии и генетики факультета естественных наук НГУ. Под ее руководством выполнено и защищено около десятка дипломных (студенты-биологи НГУ) и несколько аспирантских работ, защитили кандидатские диссертации А.В. Пономарева, Ю.О. Кашинская (совм. руководство с О.Л. Посух), М.А. Казаковцева (Губина) (совм. руководство с М.И. Воеводой), А.В. Кириченко (совм. руководство с Т.И. Аксенович), Р.П. Корчагина (Тийс).



Людмила Павловна во время одного из своих выступлений

Награды

В связи с 275-летием Российской академии наук за многолетнюю и плодотворную работу на благо отечественной науки Л.П. Осиповой была выражена благодарность в Грамоте от Президента Российской академии наук (1999 г.). Среди ее наград: Памятные знаки мэрии г. Новосибирска в честь 110-летия (2003) и «За труд и благо города» в честь 115-летия со дня основания г. Новосибирска (2008); диплом лауреата Городского дня науки мэрии г. Новосибирска, номинация «Лучший результат в области фундаментальных исследований здоровья человека» (2014); Почетная грамота Министерства образования, науки и инновационной политики Новосибирской области (2015). В 2010 г. Л.П. Осиповой было присвоено почетное звание «Заслуженный ветеран Сибирского отделения РАН».

Стоит отметить, что неоднократно Людмила Павловна была награждена и в регионах экспедиционных исследований. Ее вклад в развитие науки в Ямало-Ненецком автономном округе отмечен Памятным знаком в честь 75-летия ЯНАО (2005 г.) от губернатора ЯНАО; благодарностью «За профессионализм и высокие достижения в области молекулярной, популяционной и эволюционной генетики человека, генетической демографии, иммуногенетики и биомедицины, за большой вклад в научные исследования по изучению генофондов коренных этносов ЯНАО» (2012) и Почетными грамотами «За многолетний добросовестный труд, большой вклад в развитие науки в ЯНАО и в связи с профессиональным праздником Днем российской науки» (2017 и 2018 гг.) от Департамента по науке и инновациям

ЯНАО. В 2017 г. Л.П. Осиповой была вручена памятная медаль «90 лет Институту» в связи с юбилеем Научно-исследовательского клинического института педиатрии и детской хирургии им. академика Ю.Е. Вельтищева (г. Москва).

Труд ученого – результат сложного и уникального жизненного пути. Годы учебы, выбора и вхождения в профессию, становления, формирования вокруг себя коллег и учеников – результат сочетания множества факторов. И не всегда благоприятных... Полвека жизни в науке Л.П. Осиповой оказались созвучны ее человеческим качествам – высоким требованиям, предъявляемым не только к другим, но прежде всего к себе, трудолюбию, целеустремленности, таланту организатора научного поиска.

Участие в развитии нового научного направления – изучении генетического разнообразия народов, населяющих Северную Азию, не было широкой и прямой дорогой. Исследования носили комплексный характер, популяции изучались всесторонне – вопросы адаптации к условиям среды, проблемы состояния здоровья, демографии, влияния средовых факторов на генофонд. Длительные наблюдения за географически удаленными популяциями со временем приобрели свою особую ценность. Основой успешности Л.П. Осиповой в профессии были любовь и уважение к объекту своего исследования – человеку. И хотя конечным результатом этих работ были популяционные характеристики, для Людмилы Павловны важным всегда оставался каждый человек с его неповторимой историей жизни.

Судьба не баловала Людмилу Павловну, ее жизнь не была беззаботной и легкой. Она пережила утраты и болезни родных и близких ей людей. Перенесла вместе со всеми тяжелые времена безденежья в финансировании науки и ученых конца прошлого века. Преодолевала разочарование в связи с уходом из лаборатории и отъездом сотрудников. Разгонила черные тучи, сгущавшиеся над ее коллективом и обширным банком биологического материала, собранного в тяжелых экспедициях ее собственным трудом и стараниями многих сотрудников и соратников. Присущие ей природные трудолюбие и жизнелюбие, оптимизм и любовь к людям, забота о благополучии семьи, радость, счастье и гордость при общении с детьми и внуками не давали унынию захватить ее.

Век Людмилы Павловны был отмерен двумя семерками – 77 лет насыщенной и в целом-то счастливой жизни.

Список литературы / References

Матвеева В.Г., Керкис Ю.Я., Осипова Л.П. Мутагенный эффект пестицида ДДБ и хлоридина в костном мозге и половых клетках мышей. *Генетика*. 1973;9(2):67-76
[Matveeva V.G., Kerkis Y.Y., Osipova L.P. Mutagenic effect of the pesticide DDB and of chloridine in the bone marrow and germ cells of mice. *Soviet Genetics*. 1974;9(2):186-192]

Dard P., Lefranc M.P., Osipova L., Sanchez-Mazas A. DNA sequence variability of IGHG3 alleles associated to the main G3m haplotypes in human populations. *Eur J Hum Genet*. 2001;9(10):765-772. doi 10.1038/sj.ejhg.5200700

Dulik M.C., Zhadanov S.I., Osipova L.P., Askapuli A., Gau L., Gokcumen O., Rubinstein S., Schurr T.G. Mitochondrial DNA and Y chromosome variation provides evidence for a recent common ancestry between Native Americans and Indigenous Altaians. *Am J Hum Genet*. 2012;90(2):229-246. doi 10.1016/j.ajhg.2011.12.014

Engström P.E., Norhagen G., Osipova L., Helal A., Wiebe V., Brusco A., Carbonara A.O., Lefranc G., Lefranc M.P. Salivary IgG subclasses in individuals with and without homozygous IGHG gene deletions. *Immunology*. 1996;89(2):178-182. doi 10.1046/j.1365-2567.1996.d01-738.x

Karafet T., Zegura S.L., Vuturo-Brady J., Posukh O., Osipova L., Wiebe V., Romero F., Long J.C., Harihara S., Jin F., Dashnyam B., Gerelsaikhan T., Omoto K., Hammer M.F. Y chromosome markers and Trans-Bering Strait dispersals. *Am J Phys Anthropol*. 1997;102(3):301-314. doi 10.1002/(SICI)1096-8644(199703)102:3<301::AID-AJPA1>3.0.CO;2-Y

Karafet T.M., Zegura S.L., Posukh O., Osipova L., Bergen A., Long J., Goldman D., Klitz W., Harihara S., de Knijff P., Wiebe V., Griffiths R.C., Templeton A.R., Hammer M.F. Ancestral Asian source(s) of new world Y-chromosome founder haplotypes. *Am J Hum Genet*. 1999;64(3):817-831. doi 10.1086/302282

Karafet T.M., Osipova L.P., Gubina M.A., Posukh O.L., Zegura S.L., Hammer M.F. High levels of Y-chromosome differentiation among native Siberian populations and the genetic signature of a boreal hunter-gatherer way of life. *Hum Biol*. 2002;74(6):761-789. doi 10.1353/hub.2003.0006

Lazaridis I., Patterson N., Mittnik A., Renaud G., Mallick S., Kirsanow K., Sudmant P.H., ... Osipova L., ... Slatkin M., Pääbo S., Kelso J., Reich D., Krause J. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*. 2014;513(7518):409-413. doi 10.1038/nature13673

Leonard W.R., Galloway V.A., Ivakine E., Osipova L., Kazakovtseva M. Nutrition, thyroid function and basal metabolism of the Evenki of central Siberia. *Int J Circumpolar Health*. 1999;58(4):281-295

Leonard W.R., Sorensen M.V., Galloway V.A., Spencer G.J., Mosher M.J., Osipova L., Spitsyn V.A. Climatic influences on basal metabolic rates among circumpolar populations. *Am J Hum Biol*. 2002;14(5):609-620. doi 10.1002/ajhb.10072

McComb J., Crawford M.H., Osipova L., Karaphet T., Posukh O., Schanfield M.S. DNA interpopulational variation in Siberian indigenous populations: The Mountain Altai. *Am J Hum Biol*. 1996;8(5):599-607. doi 10.1002/(SICI)1520-6300(1996)8:5<599::AID-AJHB5>3.0.CO;2-R

Osipova L.P., Posukh O.L., Wiebe V.P., Miyazaki T., Matsumoto H., Lefranc G., Lefranc M.P. BamHI-SacI RFLP and Gm analysis of the immunoglobulin IGHG genes in the Northern Selkups (west Siberia): new haplotypes with deletion, duplication and triplication. *Hum Genet*. 1999;105(6):530-541. doi 10.1007/s004399900187

Pagani L., Lawson D.J., Jagoda E., Mörseburg A., Eriksson A., Mitt M., Clemente F., ... Osipova L.P., ... Nielsen R., Villemes R., Willerslev E., Kivisild T., Metspalu M. Genomic analyses inform on migration events during the peopling of Eurasia. *Nature*. 2016;538(7624):238-242. doi 10.1038/nature19792

Reidla M., Kivisild T., Metspalu E., Kaldma K., Tambets K., Tolk H.V., Parik J., ... Osipova L., ... Geberhiwot T., Herrnstadt C., Howell N., Torroni A., Villemes R. Origin and diffusion of mtDNA haplogroup X. *Am J Hum Genet*. 2003;73(5):1178-1190. doi 10.1086/379380

Tamm E., Kivisild T., Reidla M., Metspalu M., Smith D.G., Mulligan C.J., Bravi C.M., ... Ossipova L.P., Damba L., Voevoda M.I., Dipierri J.E., Villemes R., Malhi R.S. Beringian standstill and spread of Native American founders. *PLoS One*. 2007;2(9):e829. doi 10.1371/journal.pone.0000829

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 22.02.2025. После доработки 07.03.2025. Принята к публикации 10.03.2025.

 pismavavilov.ru

doi 10.18699/letvjgb-2025-11-05

Оригинальное исследование

Половой полиморфизм гинодиэцичного растения *Geranium sylvaticum* (Geraniaceae)

Н.И. Гордеева 

Аннотация: Исследован половой полиморфизм *Geranium sylvaticum* L. в двух популяциях в луговом и лесном фитоценозах подтаежного правобережья Оби (окрестности г. Новосибирска). В зависимости от комбинации обоеполых, пестичных и частично андростерильных цветков обнаружены три половых фенотипа: гермафродитный, женский и гиномоноэцичный. У ряда гермафродитных фенотипов формируются частично андростерильные цветки, среди которых преобладают цветки с 5 стаминодиями и 5 нормальными тычинками. Частично андростерильные цветки малочисленны и образуются в верхней и/или нижней части соцветия в начале и/или конце цветения растений. Гиномоноэцичные фенотипы представлены в двух вариантах: женский фенотип с единичными частично андростерильными цветками и гермафродитный фенотип с несколькими пестичными цветками. Женские фенотипы образуют приблизительно в 2 раза меньше цветков и в 1.6 раза меньше плодов, чем гермафродитные (различия статистически значимы), однако образование плодов в расчете на цветок у женских фенотипов в 1.4 раза выше, чем у гермафродитных. В условиях низкой освещенности гермафродитные фенотипы с частично андростерильными цветками встречаются в 2.2 раза чаще, чем при высоком уровне освещения. Соотношение двух форм гермафродитных фенотипов (с частично андростерильными цветками и без них) в луговом и лесном фитоценозах различается статистически значимо ($p = 0.0247 < 0.05$). При высоком уровне освещения гермафродитные фенотипы продуцируют в 1.6 раза больше плодов, чем при низком; различия статистически значимы ($p < 0.05$). В исследуемых популяциях *G. sylvaticum* выявлена низкая доля женских фенотипов как в луговом, так и в лесном фитоценозах: 7.1 и 4.7 % соответственно. Гиномоноэцичные фенотипы в популяциях отмечаются единично – 0.8–1.6 %. Для *G. sylvaticum* характерна сравнительно низкая встречаемость женских и гиномоноэцичных фенотипов и заметная доля гермафродитных с частично андростерильными цветками в европейской и азиатской частях ареала вида.

Ключевые слова: *Geranium sylvaticum* L.; гинодиэция; гиномоноэция; фенотип; соотношение половых фенотипов

Для цитирования: Гордеева Н.И. Половой полиморфизм гинодиэцичного растения *Geranium sylvaticum* (Geraniaceae). *Письма в Вавиловский журнал генетики и селекции*. 2025;11(1):29-35. doi 10.18699/letvjgb-2025-11-05

Финансирование: Работа выполнена в рамках государственного задания Центрального сибирского ботанического сада СО РАН № АААА-А21-121011290026-9.

Original article

Sexual polymorphism of a gynodioecious plant *Geranium sylvaticum* (Geraniaceae)

N.I. Gordeeva 

Abstract: Sexual polymorphism of a gynodioecious plant *Geranium sylvaticum* L. were investigated in meadow and forest phytocoenoses in subtaiga of the Ob right riverside (surroundings of Novosibirsk). In the populations of *G. sylvaticum* depending on the combination of bisexual, pistillate and intermediate flowers of one individual, three sexual phenotypes are distinguished: hermaphrodite, female and gynomonoeocious. In hermaphrodite phenotypes, flowers with 5 staminodes and 5 normal stamens prevailed among intermediate flowers. Gynomonoeocious phenotypes are presented in two variants: initially a female individual with single intermediate flowers and initially hermaphrodite individuals with several pistillate flowers. Intermediate flowers are few in number and are formed in the upper and/or lower part of the inflorescence, at the beginning and/or end of flowering plants. Females produce approximately 2 times fewer flowers and 1.6 times fewer fruits than hermaphrodite phenotypes; the differences are statistically significant ($p < 0.05$); however, fruit formation per flower in females is 1.4 times higher than in hermaphrodites. In low-light conditions, hermaphrodites with intermediate flowers occur 2.2 times more than in high-light conditions; the ratio of two forms of hermaphrodites (with and without intermediate flowers) between meadow and forest phytocenoses differs statistically significantly ($p = 0.0247 < 0.05$). At a high level of light, hermaphrodites produce 1.6 times more fruits than at a low level of light; the differences are statistically significant

Центральный сибирский ботанический сад Сибирского отделения Российской академии наук, Новосибирск, Россия
Central Siberian Botanical Garden of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

 nataly.gordeeva@gmail.com

 © Гордеева Н.И., 2025

($p < 0.05$). A low proportion of *G. sylvaticum* females was detected in both meadow and forest communities: 7.1 and 4.7 %, respectively. Gynomonoeious phenotypes in cenopulations are rarely observed only once: 0.8–1.6 %. *G. sylvaticum* is characterized by a relatively low frequency of females and gynomonoeious phenotypes and a noticeable participation of hermaphrodites with intermediate flowers in the European and Asian parts of the species' range.

Key words: *Geranium sylvaticum* L.; gynodioecy; gynomonoeicy; phenotype; sex ratio

For citation: Gordeeva N.I. Sexual polymorphism of a gynodioecious plant *Geranium sylvaticum* (Geraniaceae). *Pisma v Vavilovskii Zhurnal Genetiki i Seleksii = Letters to Vavilov Journal of Genetics and Breeding*. 2025;11(1):29-35. doi 10.18699/letvjgb-2025-11-05 (in Russian)

Funding: The work was carried out within the framework of the State task of the Central Siberian Botanical Garden, SB RAS No. AAAA-A21-121011290026-9.

Введение

Гинодиэция – тип полового полиморфизма у растений, при котором в природных популяциях совместно произрастают растения с обоеполюми цветками (гермафродитный фенотип) и растения с пестичными цветками (женский фенотип) (Darwin, 1897; Charlesworth, 2006). Гинодиэция описана примерно у 7 % цветковых растений и встречается во многих семействах покрытосеменных (Richards, 1997). Морфофизиологические различия у половых фенотипов растений являются наследственно устойчивыми (Меликян, 2000). Гинодиэция обеспечивает аутбридинг особей в популяциях, влияет на популяционную изменчивость видов и способствует адаптации растений к условиям окружающей среды. По современным представлениям, в основе гинодиэции выявлены две главные генетические системы, которые связаны с генами, вызывающими мужскую стерильность: 1) мутации ядерных генов; 2) мутации митохондриальных и хлоропластных генов (генов цитоплазматической мужской стерильности – ЦМС) (Couvet et al., 1990, 1998; Charlesworth, Laporte, 1998; Chase, 2007; Dufay, Billard, 2012). Ядерные или цитоплазматические гены мужской стерильности нарушают процесс образования фертильной пыльцы, в результате чего формируются пестичные цветки с недоразвитыми тычинками. Гены ЦМС могут взаимодействовать с ядерными генами-восстановителями фертильности, что приводит к восстановлению фертильности пыльцы (Couvet et al., 1990, 1998; Chase, 2007).

Для многих видов рода *Geranium* L. (семейство Geraniaceae Juss.) свойствен половой полиморфизм в форме гинодиэции (Knuth, 1898; Демьянова, 1985; Годин, 2020). Гинодиэцичные виды рода характеризуются следующими признаками полового дифференцирования растений: обоеполые цветки имеют больший размер лепестков по сравнению с пестичными цветками и фертильные тычинки; у пестичных цветков андроцей представлен стаминодиями со стерильными пыльниками или вовсе без пыльников (Knuth, 1898; Asikainen, Mutikainen, 2003; Chang, 2006; Демьянова, 2013; Varga, Kytöviita, 2016; Abdusalam et al., 2017; Гордеева, 2020; Гордеева, Комаревцева, 2023). Среди представителей рода *Geranium* обнаружены также виды с гиномоноэцичными фенотипами (растения с обоеполюми и пестичными цветками на одном растении): *G. sylvaticum* L., *G. pratense* L. (Демьянова, Аксенова, 2009; Varga, Kytöviita, 2016) и андромоноэцичными фенотипами (растения с обоеполюми и тычиночными цветками на одном растении): *G. linearilobum* DC. (= *G. transversale* (Kar. et Kir) Vved. ex Pavlov) (Abdusalam et al., 2017).

Geranium sylvaticum L. относится к растениям, система размножения которых хорошо изучена. В популяциях вида,

кроме гермафродитных и женских фенотипов, встречаются гиномоноэцичные, а также гермафродитные с частично андростерильными цветками (Asikainen, Mutikainen, 2003; Демьянова, Аксенова, 2009). Частично андростерильные цветки имеют одинаковый с обоеполюми цветками размер, но помимо фертильных пигментированных тычинок в них имеются от 1 до 9 стаминодиев с белыми недоразвитыми пыльниками (Asikainen, Mutikainen, 2003; Varga, Kytöviita, 2016). У женских фенотипов установлена либо немного более высокая семенная продуктивность по сравнению с гермафродитными (Asikainen, Mutikainen, 2003), либо почти одинаковая с ними (Varga, Kytöviita, 2016). Доля женских фенотипов может варьировать в разных местообитаниях от 0.4 до 27.2 % (Asikainen, Mutikainen, 2003; Демьянова, 2013).

S. Varga, M.M. Kytöviita (2016) исследовали половую экспрессию растений *G. sylvaticum*, которую определяли в зависимости от числа фертильных тычинок в цветках: 0 тычинок – пестичные цветки, ≥ 1 тычинки – обоеполые цветки (в том числе от 1 до 9 тычинок – частично андростерильные цветки). Показатель рассчитывали как среднее значение фертильных тычинок на цветок в конце каждого периода цветения растений. Установлено, что половая экспрессия фенотипов *G. sylvaticum* нестабильна по годам и находится в зависимости от условий освещения. Доля растений, изменяющих свою половую экспрессию при пересадке в условия с высокой освещенностью, была выше, чем доля растений, пересаженных в места с низкой освещенностью (Varga, Kytöviita, 2016). На частоту встречаемости женских фенотипов в популяциях *G. sylvaticum* могут влиять условия окружающей среды: при низком содержании минеральных веществ в почве женские фенотипы получают преимущество (Asikainen, Mutikainen, 2005). Детальные исследования опыления обоеполых и пестичных цветков *G. sylvaticum* позволяют предположить, что диморфизм размеров цветков вида приспособлен для оптимального опыления и, возможно, способствует эволюционному развитию *G. sylvaticum* по направлению к диэции (двудомности) (Soininen, Kytöviita, 2022).

В настоящее время отсутствуют данные об особенностях полового полиморфизма *G. sylvaticum* в азиатской части ареала вида. Цель настоящей работы – исследование гинодиэции *G. sylvaticum* L. в луговом и лесном сообществах подтаежного правобережья Оби (окрестности г. Новосибирска). Для этого проводили наблюдение за образованием типов цветков и плодов у растений в течение одного сезона и определяли соотношение половых фенотипов в двух популяциях с разными условиями освещенности.

Материалы и методы

Geranium sylvaticum L. относится к короткостебельным травянистым поликарпикам с евразийским ареалом (общее распространение: Атлантическая, Северная, Центральная, Южная, Восточная Европа; Кавказ, Арктика, Западная и Восточная Сибирь, Средняя Азия). Произрастает на суходольных и лесных лугах, в осиново-березовых, осиновых и смешанных, темнохвойно-осиновых лесах, на просеках, в зарослях кустарников; вид заходит в субальпийский пояс (Пешкова, 1996; Трошкина, 2019). Генеративная сфера (синфлуоресценция) *G. sylvaticum* представляет собой закрытый тирс из монохазиев, паракладии немногочисленные (Кузнецова и др., 1992).

Исследования проводились в двух популяциях *G. sylvaticum* в луговом и лесном сообществах в 2023 г. в окрестности г. Новосибирска (подтаежное правобережье Оби).

1. Лесной злаково-разнотравный луг на вырубке березово-соснового леса (54°49'22.64" с. ш., 83°06'27.85" в. д.). Общее проективное покрытие (ОПП) – 100 %, проективное покрытие (ПП) вида *G. sylvaticum* – 2 %. Доминанты: *Aegopodium podagraria* L., *Dactylis glomerata* L., *Calamagrostis arundinacea* (L.) Roth., *Brachypodium pinnatum* (L.) P. Beauv., *Anthriscus sylvestris* (L.) Hoffm., *Heracleum dissectum* Ledeb.
2. Разнотравный сосново-березовый лес (54°50'00.83" с. ш., 83°06'08.50" в. д.). ОПП – 90 %, ПП *G. sylvaticum* – 1 %. Доминанты: *Aegopodium podagraria*, *Pteridium aquilinum* (L.) Kuhn., *Dactylis glomerata*, *Urtica dioica* L., *Carex pediformis* var. *macroura* (Meinsh.) Kük.

Для учета фертильных тычинок vs стаминодиев (видоизмененных, недоразвитых, лишенных пыльника тычинок, утративших способность производить пыльцу) в цветках разных фенотипов *G. sylvaticum* в каждой популяции в начале фазы цветения отмечали 20 растений с обоеполюми цветками и 6–9 растений с пестичными цветками (все обнаруженные женские фенотипы на начало наблюдений). Несколько отмеченных растений были повреждены в процессе вегетации и потому не включены в окончательную обработку данных. В результате в популяции № 1 выборка состояла из 19 растений с первоначально обоеполюми цветками и 9 растений с пестичными цветками. В популяции № 2 выборка состояла из 20 растений с первоначально обоеполюми цветками и 2 растений с пестичными цветками. Во время цветения через каждые сутки у отмеченных особей просматривали раскрытые цветки и подсчитывали в них число тычинок и стаминодиев. Фертильными считали тычинки, пыльники которых были пурпурно окрашены и содержали пигментированную пыльцу (Asikainen, Mutikainen, 2003). Стаминодии либо отличались стерильными пыльниками белого цвета, либо были короткими и без пыльников. Наблюдения за особями проводили в течение 20 дней до фазы «конец цветения–начало плодоношения».

Число цветков и плодов у особей *G. sylvaticum* определяли в каждой популяции в начале периода плодоношения растений. Большая часть плодов находилась на ранней стадии формирования, поэтому семена не подсчитывали. Ранее при изучении *G. sylvaticum* выявлено, что различия между фенотипами по показателю числа плодов коррели-

руют с количеством семян, т. е. с семенной продуктивностью (Asikainen, Mutikainen, 2003). Сравнение гермафродитных и женских фенотипов по показателям числа цветков и плодов в расчете на одну особь проводили с помощью *t*-критерия Уэлча для выборок с неодинаковой дисперсией ($p < 0.05$) только в популяции № 1, так как в популяции № 2 фазы плодоношения достигли лишь две женские формы из всех первоначально отмеченных.

Исследуемые популяции *G. sylvaticum* лугового (№ 1) и лесного (№ 2) фитоценозов различались условиями освещения. Освещенность местообитаний определяли во время массового цветения *G. sylvaticum* в середине июня с помощью прибора Люксметр Ю-116 в безоблачный день в 12:00 ч на двадцати произвольно выбранных точках на уровне травянистого яруса, затем данные усреднялись. В луговом сообществе уровень освещенности составлял около 70 000 ± 4400 лк, в лесном сообществе под древесным пологом – в 4.4 раза меньше, около 16 000 ± 3400 лк. Влияние условий освещения на образование частично андростерильных цветков исследовали только у гермафродитных фенотипов каждой выборки, поскольку женские фенотипы были малочисленны. Для этого оценивали различия в распределении двух гермафродитных фенотипов в двух популяциях с помощью точного критерия Фишера для малых выборок ($p < 0.05$) (Зайцев, 1991). Влияние условий освещения на образование плодов у гермафродитных особей в популяциях анализировали с помощью критерия Стьюдента, $p < 0.05$ (Зайцев, 1991).

Соотношение разных половых фенотипов в популяциях *G. sylvaticum* определяли в период массового цветения растений по общепринятым методикам, путем подсчета всех генеративных особей в популяциях на площадях 200–300 м² (Asikainen, Mutikainen, 2003). Для удобства подсчета растений исследуемую площадь разбивали на трансекты 2–5 м шириной. Выборка в популяции № 1 составляла 139 особей, в популяции № 2 – 128 особей.

Результаты

Типы цветков и половые фенотипы

В результате исследования *G. sylvaticum* выявлено три типа цветков: обоеполые, пестичные и частично андростерильные (табл. 1, рис. 1). Частично андростерильные цветки, кроме фертильных тычинок, имели разное число стаминодиев. В большинстве цветков в андростерильных цветках отмечалось 5 стаминодиев (92 % от всех частично андростерильных цветков), в единичных случаях – 8 или 9 стаминодиев (8 %). В зависимости от сочетания разных типов цветков у одной особи могут формироваться: 1 – только обоеполые цветки (гермафродитные фенотипы); 2 – только пестичные цветки (женские фенотипы); 3 – совместно обоеполые и частично андростерильные цветки (гермафродитные фенотипы); 4 – совместно обоеполые и пестичные цветки (гиномоноэцичные фенотипы); 5 – совместно пестичные и частично андростерильные цветки (гиномоноэцичные фенотипы) (см. табл. 1).

В исследуемых выборках гермафродитные фенотипы с постоянным числом фертильных тычинок (10 шт.) и женские без фертильных тычинок составляют значительную долю:

Таблица 1. Половые фенотипы и типы цветков *Geranium sylvaticum*
Table 1. Sexual phenotypes and flower types of *Geranium sylvaticum*

Фенотип	Число особей	Число цветков на одно растение, $\frac{M \pm m}{\%}$		
		обоеполюх	пестичных	частично андростерильных
Популяция № 1				
1	11	$\frac{34.0 \pm 5.9}{100}$	$\frac{0}{0}$	$\frac{0}{0}$
2	8	$\frac{0}{0}$	$\frac{15.4 \pm 2.4}{100}$	$\frac{0}{0}$
3	8	$\frac{32.3 \pm 6.2}{94.4}$	$\frac{0}{0}$	$\frac{1.9 \pm 0.3}{5.6}$
4	1	$\frac{0}{0}$	$\frac{11.0}{91.7}$	$\frac{1.0}{8.3}$
Популяция № 2				
1	4	$\frac{30.8 \pm 8.4}{100}$	$\frac{0}{0}$	$\frac{0}{0}$
2	2	$\frac{0}{0}$	$\frac{11.5 \pm 1.5}{100}$	$\frac{0}{0}$
3	14	$\frac{21.5 \pm 3.4}{94.1}$	$\frac{0}{0}$	$\frac{1.4 \pm 0.2}{5.9}$
5	5	$\frac{22.0 \pm 11.0}{88.0}$	$\frac{2.5 \pm 1.5}{10.0}$	$\frac{1.0}{2.0}$

Примечание. Фенотипы: 1 – особи с обоеполюми цветками (гермафродитный фенотип); 2 – особи с пестичными цветками (женский фенотип); 3 – особи с обоеполюми и частично андростерильными цветками (гермафродитный фенотип); 4 – особи с пестичными и частично андростерильными цветками (гиномоноэцичный фенотип); 5 – особи с обоеполюми, частично андростерильными и пестичными цветками (гиномоноэцичный фенотип). $\frac{M \pm m}{\%}$ – над чертой среднее значение и ошибка среднего, под чертой – доля от общего числа цветков.



Рис. 1. Типы цветков *Geranium sylvaticum*: a – обоеполюый; b – частично андростерильный; c – пестичный
Fig. 1. Type of *Geranium sylvaticum* flowers: a, bisexual; b, intermediate; c, pistillate

68 и 27 % растений для первой и второй популяций соответственно. Обнаружено большое количество гермафродитных фенотипов с частично андростерильными цветками: 29 и 64 % растений для первой и второй популяций соответственно. Однако число частично андростерильных цветков на отдельных гермафродитных фенотипах очень мало: 1–3 шт. на растение (5.6 и 5.9 % от всех цветков). Эти цветки формируются в начале и/или в конце цветения растений в верхней и нижней частях соцветия.

Гиномоноэцичные фенотипы малочисленны и обнаружены в двух вариантах: первоначально женский фенотип с единичными частично андростерильными цветками (одно растение в популяции № 1) и первоначально гермафродитные фенотипы с несколькими пестичными цветками (два растения в популяции № 2) (см. табл. 1). По нашим наблюдениям, у гиномоноэцичных фенотипов частично андростерильные и пестичные цветки образуются единично в верхней части соцветия в конце цветения растений.

Таблица 2. Число цветков и плодов гермафродитных фенотипов *Geranium sylvaticum*
Table 2. Number of flowers and fruits of hermaphrodite of *Geranium sylvaticum*

Фенотип	Число цветков на одно растение			Число плодов на одно растение		
	Min–max	$M \pm m$	$C_v, \%$	Min–max	$M \pm m$	$C_v, \%$
Популяция № 1						
A	10–69	34.0 ± 5.95	58.1	9–39	19.5 ± 3.44	58.6
B	16–67	34.1 ± 6.39	53.0	9–40	22.1 ± 4.06	52.0
Популяция № 2						
A	6–44	30.8 ± 8.44	54.9	3–23	14.8 ± 4.33	58.7
B	7–44	24.5 ± 3.30	48.6	1–27	11.9 ± 2.12	66.6

Примечание. А – фенотипы с обоеполыми цветками, В – фенотипы с обоеполыми и частично андростерильными цветками; Min–max – минимальное и максимальное значения; $M \pm m$ – среднее значение и ошибка среднего; C_v – коэффициент вариации.

Плоды

В популяции № 1 между гермафродитными и женскими фенотипами *G. sylvaticum* установлены статистически значимые различия по показателю числа цветков и плодов ($p < 0.05$): $t_{\text{крит}2,14} = 3.88$ и $t_{\text{крит}2,2} = 2.54$ соответственно. Женские фенотипы образуют приблизительно в 2 раза меньше цветков и в 1.6 раза меньше плодов, чем гермафродитные (рис. 2). Однако образование плодов в расчете на один цветок у женских фенотипов в 1.4 раза выше, чем у гермафродитных. Коэффициенты вариации показателей у разных половых фенотипов сравнительно близкие и составляют 43.8 % (женские) и 54.4 % (гермафродитные).

Влияние условий освещения

Анализ, проведенный с помощью критерия Фишера, показал, что в условиях низкой освещенности гермафродитные фенотипы с частично андростерильными цветками встречаются в 2.2 раза чаще, чем в фитоценозе с высокой освещенностью; различия статистически значимы (см. табл. 1). В разных условиях освещения не установлено статистически значимых различий по числу цветков у гермафродитных фенотипов (табл. 2). Однако в условиях высокой освещенности у гермафродитных фенотипов образуется в 1.6 раза больше плодов (в среднем на одно растение), чем при низком уровне освещенности; различия статистически значимы ($t = 2.87, p < 0.05$) (см. табл. 2).

Соотношение половых фенотипов

В популяциях как лугового, так и лесного сообществ преобладают гермафродитные фенотипы: 92.1 и 93.7 % соответственно. Встречаемость женских фенотипов намного ниже: 7.1 и 4.7 % для лугового и лесного сообществ. Кроме того, в популяциях отмечаются единично гиномоноэцичные фенотипы: 0.8 и 1.6 % для лугового и лесного сообществ соответственно.

Обсуждение

При исследовании *G. sylvaticum* в зависимости от комбинации обоеполых, пестичных и частично андростерильных цветков обнаружены три половых фенотипа: гермафродитные, женские и гиномоноэцичные. Выявленные фенотипы характерны для вида также в других частях его ареала (Asikainen, Mutikainen, 2003; Varga, Kytöviita, 2016).

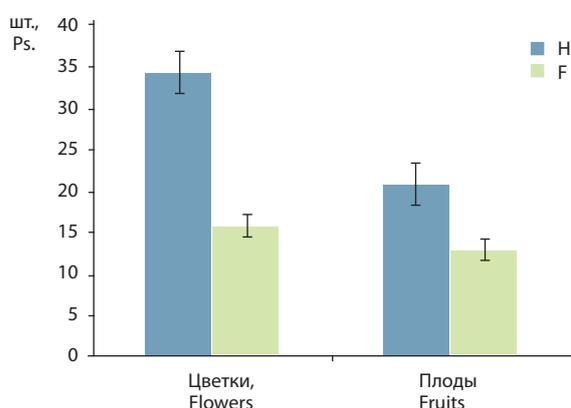


Рис. 2. Число цветков и плодов у гермафродитных (H) и женских (F) фенотипов *Geranium sylvaticum* в луговом сообществе

Fig. 2. Number of flowers and fruits in hermaphrodite (H) and female (F) phenotypes of *Geranium sylvaticum* in the meadow community

В литературе имеются противоречивые сведения в отношении семенной продуктивности гермафродитных и женских фенотипов *G. sylvaticum*. Одни авторы отмечают небольшое превышение семенной продуктивности для женских фенотипов: число семян на цветок у женских особей в 1.2–1.7 раза выше, чем у гермафродитных (Asikainen, Mutikainen, 2003). Другие указывают, что число семян не различается (Varga, Kytöviita, 2016). Наши исследования показали, что женские фенотипы *G. sylvaticum* отличаются меньшим числом плодов, чем гермафродитные, однако в расчете на один цветок у женских образуется в 1.4 раза больше плодов. Эти результаты подтверждают предположение J.O. Soininen, M.M. Kytöviita (2022) о том, что пестичные цветки *G. sylvaticum* опыляются успешнее, чем обоеполые, и, возможно, успешности опыления способствует диморфизм по размеру цветков. В работе (Varga, Kytöviita, 2016) не обнаружено связи между частичной андростерильностью цветков и семенной продуктивностью у гермафродитных фенотипов *G. sylvaticum*. Мы также не установили статистически значимых различий для *G. sylvaticum* по числу плодов между двумя формами гермафродитных фенотипов (с обоеполыми цветками или наличием частично андростерильных цветков), что позволяет предположить отсутствие связи между образованием частично андростерильных цветков и наличием плодов.

При исследовании соотношения половых фенотипов в популяциях *G. sylvaticum* следует учитывать, что образование пестичных цветков у гиномоноэичных фенотипов происходит в начале или в конце цветения растений (Демьянова, Аксенова, 2009; наша работа). В лесных и лесостепных местообитаниях европейской части ареала встречаемость женских фенотипов невысока и составляет в среднем около 11.5 % (от 0.4 до 27.4 %) от числа всех генеративных особей (Asikainen, Mutikainen, 2003; Демьянова, 2013). Доля гиномоноэичных фенотипов в популяциях еще меньше – до 2 % (Демьянова, Аксенова, 2009). По нашим данным, в популяциях наблюдается низкая встречаемость женских фенотипов (4.7–7.1 %) от всех генеративных особей, гиномоноэичные фенотипы отмечаются еще реже (1.6 %). Согласно (Демьянова, Аксенова, 2009), в популяциях гермафродитные фенотипы с частично андростерильными цветками составляют около 26 %; в наших исследованиях этот показатель не превышает 11 %. Таким образом, *G. sylvaticum* характеризуются сравнительно низкой долей женских и гиномоноэичных фенотипов и заметным участием гермафродитных фенотипов с частично андростерильными цветками.

Причины возникновения половых различий у видов рода *Geranium* обсуждаются. По мнению ряда авторов, одна из основных причин недоразвития гинецея и андроеца при морфогенезе цветка – ядерно-цитоплазматический контроль в системе размножения организмов, который влияет на мужскую стерильность, т. е. нарушает образование фертильной пыльцы (Asikainen, Mutikainen, 2005; Varga, Kutöviita, 2016). В этом случае у пестичных цветков формируются вместо тычинок стаминодии, что определяет женский фенотип растения. А. Abdusalam с коллегами (2017) при изучении *G. linearilobum* обнаружили, помимо обоеполюх и пестичных цветков, тычиночные цветки, которые отличаются от обоеполюх только недоразвитием столбиков в пестиках, хотя в завязи имеются семяпочки нормального размера, как у обоеполюх цветков. Они предположили, что недоразвитие гинецея в тычиночных цветках *G. linearilobum* может быть связано с влиянием неблагоприятных условий окружающей среды, т. е. с экологическими факторами, а не детерминировано генетическими факторами ядерно-цитоплазматического контроля (Abdusalam et al., 2017). В наших исследованиях в условиях с низкой освещенностью отмечается увеличение числа особей с частично андростерильными цветками в популяции. Кроме того, обнаружено, что частично андростерильные цветки образуются в начале периода цветения растений при нестабильной весенней температуре и в конце цветения при ухудшении освещенности из-за развития растений более высоких ярусов. Мы предполагаем, что частичная андростерильность *G. sylvaticum* связана с влиянием неблагоприятных условий на экспрессию генов репродуктивной системы и не зависит от ядерно-цитоплазматического контроля мужской стерильности.

Заключение

В зависимости от сочетания обоеполюх, пестичных и частично андростерильных цветков обнаружены три половых фенотипа: гермафродитные, женские и гиномоноэичные. Частично андростерильные цветки малочисленны (1–3 шт.,

в среднем 2.3 % от всех цветков) и образуются в верхней или в нижней частях соцветия в начале и/или конце цветения растений. Гиномоноэичные фенотипы встречаются в двух вариантах: первоначально женские с единичными частично андростерильными цветками и первоначально гермафродитные с несколькими пестичными цветками. В условиях низкой освещенности гермафродитные фенотипы с частично андростерильными цветками встречаются в 2.2 раза чаще, чем в условиях с высокой освещенностью (различия статистически значимы, $p=0.0247 < 0.05$). Женские фенотипы образуют приблизительно в 2 раза меньше цветков и в 1.6 раза меньше плодов, чем гермафродитные (различия статистически значимы, $p < 0.05$), однако образование плодов в расчете на цветок у женских фенотипов в 1.4 раза выше, чем у гермафродитных. В условиях высокой освещенности гермафродитные фенотипы продуцируют в 1.6 раза больше плодов, чем при низком уровне освещенности (различия статистически значимы, $p < 0.05$).

Выявлена низкая доля женских фенотипов *G. sylvaticum* в луговом и лесном сообществах: 7.1 и 4.7 % соответственно. Гиномоноэичные фенотипы в исследованных популяциях отмечаются единично – 0.8–1.6 %. Для *G. sylvaticum* характерна сравнительно низкая встречаемость женских и гиномоноэичных фенотипов и заметное участие гермафродитных фенотипов с частично андростерильными цветками в разных частях ареала вида.

Список литературы / References

- Годин В.Н. Распространение гинодиэции у цветковых растений. *Ботанический журнал*. 2020;105(3):236-252. doi 10.31857/S0006813620030023
- [Godin V.N. Distribution of gynodioecy in flowering plants. *Botanicheskii Zhurnal*. 2020;105(3):236-252. doi 10.31857/S0006813620030023 (in Russian)]
- Гордеева Н.И. Гинодиэция *Geranium bifolium* (Geraniaceae). *Ботанический журнал*. 2020;105(3):293-299. doi 10.31857/S0006813620030035
- [Gordeeva N.I. Gynodioecia of *Geranium bifolium* (Geraniaceae). *Botanicheskii Zhurnal*. 2020;105(3):293-299. doi 10.31857/S0006813620030035 (in Russian)]
- Гордеева Н.И., Комаревцева Е.К. Гинодиэция *Geranium pseudosibiricum* (Geraniaceae). *Ботанический журнал*. 2023;108(12):1093-1099. doi 10.31857/S0006813623120049
- [Gordeeva N.I., Komarevtseva E.K. Gynodioecia of *Geranium pseudosibiricum* (Geraniaceae). *Botanicheskii Zhurnal*. 2023;108(12):1093-1099. doi 10.31857/S0006813623120049 (in Russian)]
- Демьянова Е.И. Распространение гинодиэции у цветковых растений. *Ботанический журнал*. 1985;70(10):1289-1301
- [Dem'yanova E.I. Distribution of gynodioecy in flowering plants. *Botanicheskii Zhurnal*. 1985;70(10):1289-1301 (in Russian)]
- Демьянова Е.И. Половой полиморфизм некоторых степных растений Центрально-Черноземного заповедника. *Вестник Пермского университета. Серия: Биология*. 2013;2:11-18
- [Dem'yanova E.I. Sexual polymorphism of some steppe plants of Central Chernozomnii reserve. *Bulletin of Perm University. Biology*. 2013;2:11-18 (in Russian)]
- Демьянова Е.И., Аксенова Л.В. Половая структура популяций *Geranium sylvaticum* L. и *G. pratense* L. в Приуралье. В: Ботанические исследования на Урале. Пермь, 2009;101-102
- [Dem'yanova E.I., Aksenoval L.V. The sexual structure of the populations of *Geranium sylvaticum* L. and *G. pratense* L. in the Ural. In: Botanical Research in the Urals. Perm, 2009;101-102 (in Russian)]

- Зайцев Г.Н. Математический анализ биологических данных. М., 1991 [Zaitsev G.N. Mathematical Analysis of the Biological Data. Moscow, 1991 (in Russian)]
- Кузнецова Т.В., Прякина Н.И., Яковлев Г.П. Соцветия. Морфологическая классификация. СПб., 1992 [Kuznetsova T.V., Pryakhina N.I., Yakovlev G.P. Inflorescences. The Morphological Classification. St. Petersburg, 1992 (in Russian)]
- Меликян А.П. Половой полиморфизм. В: Эмбриология цветковых растений. Терминология и концепции. Системы репродукции. Т. 3. СПб., 2000;73-75 [Melikyan A.P. Sex polymorphism. In: Embryology of Flowering Plants. Terminology and Concepts. Reproductive Systems. Vol. 3. St. Petersburg, 2000;73-75 (in Russian)]
- Пешкова Г.А. Семейство Geraniaceae. Флора Сибири. Т. 10. Новосибирск, 1996;8-22 [Peshkova G.A. Family Geraniaceae. In: Flora of Siberia. Vol. 10. Novosibirsk, 1996;8-22 (in Russian)]
- Трошкина В.И. Конспект видов рода *Geranium* (Geraniaceae) Алтайской горной страны. *Растительный мир Азиатской России*. 2019;3:13-28. doi 10.21782/RMAR1995-2449-2019-3(13-28) [Troshkina V.I. The synopsis of the genus *Geranium* (Geraniaceae) of the Altai mountain country. *Rastitel'nyj Mir Aziatskoj Rossii = Flora and Vegetation of Asian Russia*. 2019;3:13-28. doi 10.21782/RMAR1995-2449-2019-3(13-28) (in Russian)]
- Abdusalam A., Tan D., Chang S.M. Sexual expression and reproductive output in the ephemeral *Geranium transversale* are correlated with environmental conditions. *Am J Bot.* 2017;104(12):1920-1929. doi 10.3732/ajb.1700258
- Asikainen E., Mutikainen P. Female frequency and relative fitness of females and hermaphrodites in gynodioecious *Geranium sylvaticum* (Geraniaceae). *Am J Bot.* 2003;90(2):226-234. doi 10.3732/ajb.90.2.226
- Asikainen E., Mutikainen P. Pollen and resource limitation in a gynodioecious species. *Am J Bot.* 2005;92(3):487-494. doi 10.3732/ajb.92.3.487
- Chang S.M. Female compensation through the quantity and quality of progeny in a gynodioecious plant, *Geranium maculatum* (Geraniaceae). *Am J Bot.* 2006;93(2):263-270. doi 10.3732/ajb.93.2.263
- Charlesworth D. Evolution of plant breeding systems. *Curr Biol.* 2005;16(17):R726-R735. doi 10.1016/j.cub.2006.07.068
- Charlesworth D., Laporte V. The male-sterility polymorphism of *Silene vulgaris*: analysis of genetic data from two populations and comparison with *Thymus vulgaris*. *Genetics*. 1998;150(3):1267-1282. doi 10.1093/genetics/150.3.1267
- Chase C.D. Cytoplasmic male sterility: a window to the world of plant mitochondrial-nuclear interactions. *Trends Genet.* 2007;23(2):81-90. doi 10.1016/j.tig.2006.12.004
- Couvet D., Atlan A., Belhassen E., Gliddon C., Gouyon P.H., Kjellberg F. Co-evolution between two symbionts: the case of cytoplasmic male-sterility in higher plants. In: Futuyma D.J., Antonovics J. (Eds) *Oxford Surveys in Evolutionary Biology*. Vol. 7. Oxford University Press, 1990;225-249
- Couvet D., Ronce O., Gliddon C. The maintenance of nucleocytoplasmic polymorphism in a metapopulation: the case of gynodioecy. *Am Nat.* 1998;152(1):59-70. doi 10.1086/286149
- Darwin C. *The Different Forms of Flowers on Plants of the Same Species*. New York: D. Appleton, 1897
- Dufay M., Billard E. How much better are females? The occurrence of female advantage, its proximal causes and its variation within and among gynodioecious species. *Ann Bot.* 2012;109(3):505-519. doi 10.1093/aob/mcr062
- Knuth P. *Handbuch der Blütenbiologie*. Bd. II. Leipzig: W. Engelmann, 1898. doi 10.5962/bhl.title.23080
- Richards A.J. *Plant Breeding Systems*. Garland Science, 1997
- Soininen J.O., Kytöviita M.M. *Geranium sylvaticum* increases pollination probability by sexually dimorphic flowers. *Ecol Evol.* 2022;12(12):e9670. doi 10.1002/ece3.9670
- Varga S., Kytöviita M.M. Light availability affects sex lability in a gynodioecious plant. *Am J Bot.* 2016;103(11):1928-1936. doi 10.3732/ajb.1600158

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Поступила в редакцию 12.11.2024. После доработки 14.01.2025. Принята к публикации 30.01.2025.

 pismavavilov.ru

doi 10.18699/letvjgb-2025-11-06

Оригинальное исследование

Определение *in vitro* сродства ТВР к ТАТА-боксу промоторов генов *Arabidopsis thaliana*

Е.Б. Шарыпова ¹, М.П. Пономаренко ¹, Н.А. Колчанов ^{1,2}, Л.К. Савинкова ¹✉

Аннотация: Для инициации транскрипции РНК-полимеразой II с ТАТА-содержащих промоторов необходима сборка транскрипционного комплекса, начинающаяся со взаимодействия ТВР (TATA binding protein) или TFIIID, в состав которого он входит, с ТАТА-боксом, и присоединения к ним других базальных факторов транскрипции. Изучению взаимодействия ТВР с ТАТА-боксами и ТАТА-подобными элементами млекопитающих посвящено большое количество работ, увеличивающееся с обнаружением влияния однонуклеотидных замен (SNP) в ТАТА-последовательностях на фенотипические проявления, в том числе на связь с различными заболеваниями человека. Взаимодействие ТВР с ТАТА-боксами растений, в частности *Arabidopsis thaliana*, практически не изучалось, за исключением единичных ранних работ. Широкомасштабное предсказание промоторных последовательностей и входящих в них сайтов связывания транскрипционных факторов с помощью различных инструментов *in silico* успешно используется благодаря достижениям в транскриптомных технологиях и секвенировании геномов довольно многих видов растений (*A. thaliana*, рис, томат, свекла, картофель, пшеница, хлопчатник и др.). Однако предсказанные регуляторные последовательности могут быть функциональными (а могут и не быть), и необходима экспериментальная верификация вклада потенциального сайта связывания фактора транскрипции в активность промотора для подтверждения его функциональности. В данной работе применяется количественный метод задержки ДНК в геле для определения сродства ТВР к ТАТА-боксам промоторов генов *A. thaliana* и его сравнения с предсказанным с помощью веб-сервиса Plant_SNP_TATA_Z-tester значением. Для изучения мы использовали гены *A. thaliana* *PUR7*, *TFL1*, *MLH1*, *FT*, *AOP2*, транскрибируемые в разных органах растения. Выявлено, что олигодезоксирибонуклеотиды, идентичные последовательностям ТАТА-боксов с фланкирующими их нуклеотидами генов *A. thaliana*, имеют сродство к ТВР, различающееся до 66 раз. Фланкирующие последовательности ТАТА-боксов промоторов рассмотренных генов *A. thaliana* имеют более низкое содержание GC-нуклеотидов во фланках (на 23 %) по сравнению с генами человека. Эксперименты показали, что для предсказанных с помощью Plant_SNP_TATA_Z-tester равновесных констант диссоциации K_D и их экспериментальных значений, выраженных в нМ, коэффициент линейной корреляции (r) = 0.97 при $p < 0.05$. Полученные результаты указывают на надежность *in silico* оценки сродства ТВР *Arabidopsis* к промоторам генов этого растения с помощью сервиса Plant_SNP_TATA_Z-tester.

Ключевые слова: *Arabidopsis thaliana*; ТАТА-боксы; промоторы генов *A. thaliana*; ТВР/ТАТА сродство

Для цитирования: Шарыпова Е.Б., Пономаренко М.П., Колчанов Н.А., Савинкова Л.К. Определение *in vitro* сродства ТВР к ТАТА-боксам промоторов генов *Arabidopsis thaliana*. Письма в Вавиловский журнал генетики и селекции. 2025;11(1):36-42. doi 10.18699/letvjgb-2025-11-06

Финансирование: Работа выполнена при поддержке бюджетного финансирования, № FWNR-2022-0016.

Original article

In vitro determination of TBP affinity to the TATA-box of promoters of *Arabidopsis thaliana* genes

Е.Б. Sharypova ¹, М.П. Ponomarenko ¹, Н.А. Kolchanov ^{1,2}, Л.К. Savinkova ¹✉

Abstract: Initiation of transcription by RNA polymerase II from TATA-containing promoters requires the assembly of a transcription complex, which begins with the interaction of TBP (TATA binding protein) or TFIIID, of which it is a part, with the TATA box and the addition of other basal transcription factors to them. A large number of studies have been devoted to studying the interaction of TBP (TATA binding protein) with TATA boxes and TATA-like elements in mammals, which began to increase with the discovery of the effect of single nucleotide polymorphisms (SNPs) in TATA sequences on phenotype, in particular, on association with various human diseases.

¹ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

² Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
Novosibirsk State University, Novosibirsk, Russia

 savinkl@mail.ru

 Шарыпова Е.Б., Пономаренко М.П., Колчанов Н.А., Савинкова Л.К., 2025

The interaction of TBP with the TATA boxes in plants, in particular, in *Arabidopsis thaliana*, has not been practically studied, with the exception of a few early studies. Large-scale prediction of promoter sequences and their constituent transcription factor binding sites using various *in silico* tools is successfully used due to advances in transcriptomic technologies and genome sequencing of many plant species (model plant *A. thaliana*, rice, tomato, beet, potato, wheat, cotton, etc.). However, predicted regulatory sequences may or may not be functional, and experimental verification of the contribution of a potential transcription factor binding site to promoter activity is required to confirm its functionality. This work uses a quantitative DNA gel retention method to determine the affinity of TBP for TATA boxes of *A. thaliana* gene promoters predicted using Plant_SNP_TATA_Z-tester, which correlates with the activity of transcription and gene expression. For the study, we used the genes of *A. thaliana* *PUR7*, *TFL1*, *MLH1*, *FT*, *AOP2* transcribed in different tissues of the plant. We showed that oligodeoxyribonucleotides identical to the sequences of TATA boxes with flanking nucleotides of *A. thaliana* genes have an affinity for TBP that differs up to 66 times. Flanking sequences of TATA boxes of the promoters of the considered *A. thaliana* genes have a lower content of GC-nucleotides in the flanks (by 23 %) compared to human genes. Experiments have shown that for K_D predicted by Plant_SNP_TATA_Z-tester and their experimental values expressed in nM, the linear correlation coefficient (r) = 0.97 at $p < 0.05$, for predicted and experimental values expressed in ln units ($-\ln K_D$), $r = 0.76$ ($p < 0.05$). The results obtained indicate the reliability of *in silico* assessment of the affinity of Arabidopsis TBP for promoters of the genes of this plant using the Plant_SNP_TATA_Z-tester service.

Key words: *Arabidopsis thaliana*; TATA-box; *A. thaliana* genes promoters; TBP/TATA affinity

For citation: Sharypova E.B., Ponomarenko M.P., Kolchanov N.A., Savinkova L.K. *In vitro* determination of TBP affinity to the TATA-box of promoters of *Arabidopsis thaliana* gene. *Pisma v Vavilovskii Zhurnal Genetiki i Selekcii = Letters to Vavilov Journal of Genetics and Breeding*. 2025;11(1):36-42. doi 10.18699/letvjgb-2025-11-06 (in Russian)

Funding: The work was carried out with the support of budget funding, No. FWNR-2022-0016.

Введение

Для инициации транскрипции с помощью РНК-полимеразы II всем эукариотам требуется сборка базального аппарата транскрипции на кор-промоторе – области, расположенной в районе примерно (–50, +50) пар нуклеотидов выше и ниже сайта инициации транскрипции (transcription start site, TSS) (Hampsey, 1998; Sandelin et al., 2007; Juven-Gershona, Kadonaga, 2010; Kadonaga, 2012). Кор-промоторы часто содержат консервативные элементы ДНК, распознаваемые компонентами базального транскрипционного аппарата, – общими (general, basal) факторами транскрипции: TFIID, TFIIB, TFIIA, TFIIF, TFIIE и TFIIH (Sandelin et al., 2007). Наиболее хорошо описан элемент ДНК кор-промотора – TATA-бокс, который распознается TATA-связывающим белком (TATA binding protein, TBP). TATA-бокс представляет собой АТ-богатую последовательность, расположенную примерно на 25–35 пар оснований выше TSS почти у всех эукариот (Burley, Roeder, 1996). Рекрутирование TBP и ассоциированных с ним факторов, являющихся частью основного фактора транскрипции TFIID, направляет сборку преинициаторного комплекса (preinitiation complex, PIC) на кор-промоторе. Это строго регулируемый процесс, обеспечивающий точную инициацию транскрипции.

У растений, как показано для *Arabidopsis thaliana* (*Arabidopsis*) и *Oryza sativa* (рис), достаточно строгое позиционирование некоторых элементов и кор-промотора относительно старта транскрипции. Выявлено, что TATA-бокс чаще расположен в области (–30, –25) вверх от старта транскрипции (Patikoglou et al., 1999). TBP у растений связывается со специфическими последовательностями кор-промоторов генов, TATA-боксами, имеющими консенсус TATAWAWR (Rhee, Pugh, 2012) или TATAWA (где W – А или Т) (Dikstein, 2011), и с TATA-элементами, содержащими допустимые замены внутри канонической последовательности. TBP связывает и изгибает ДНК, затем происходит присоединение основных факторов транскрипции TFIIA и TFIIB, стабилизирующих комплекс, что необходимо для дальнейшей сборки PIC, заключающейся в присоедине-

нии к этому субкомплексу РНК-полимеразы II и базальных факторов IIF, IIH, IIE. Следует отметить, что универсального состава кор-промоторных элементов не существует. Вариации последовательности внутри кор-промотора могут направлять образование альтернативных транскрипционных комплексов, необходимых в разных физиологических условиях.

Взаимодействие TBP с TATA-боксами у растений, в частности у *Arabidopsis*, практически не изучалось, за исключением единичных ранних работ, в которых исследовали роль TATA-боксов в системах транскрипции растений. Следует отметить работу (Mukumoto et al., 1993), где подробно исследовано влияние замен в TATA-боксах *Arabidopsis* на транскрипцию с использованием *in vitro* химерной системы *HeLa*, в которой TATA-связывающий белок человека был заменен очищенным TBP *Arabidopsis*. В работе было показано, что при замене третьего Т на А в $T_1A_2T_3A_4T_5A_6T_7A_8$ сохранялось менее 15 % транскрипционной активности от дикого типа, тогда как другие замены А на Т или Т на А позволяли сохранить более 36 % от уровня транскрипции дикого типа. Три последовательности TATA-боксов, TATATATA, TATAAATA и TATATAAA, эффективные в транскрипции эукариотических генов, также продуктивны были в транскрипции *Arabidopsis*. Замены А или Т на С или Г в положениях 2, 3, 4 или 5 сильно снижали эффективность транскрипции, тогда как влияние замен в положениях 1, 6 или 7 было менее значительным. Среди мутантных TATA-элементов последовательности TAGAGATA и GAGAGAGA были полностью неактивны в системе транскрипции.

Полногеномное предсказание промоторных последовательностей и входящих в них сайтов связывания транскрипционных факторов с помощью различных инструментов *in silico* успешно используется благодаря достижениям в транскриптомных технологиях и секвенировании геномов довольно многих видов растений (Marks et al., 2021). Однако предсказанные регуляторные последовательности могут быть функциональными (но могут и не быть), и необходима экспериментальная верификация вклада потенциального

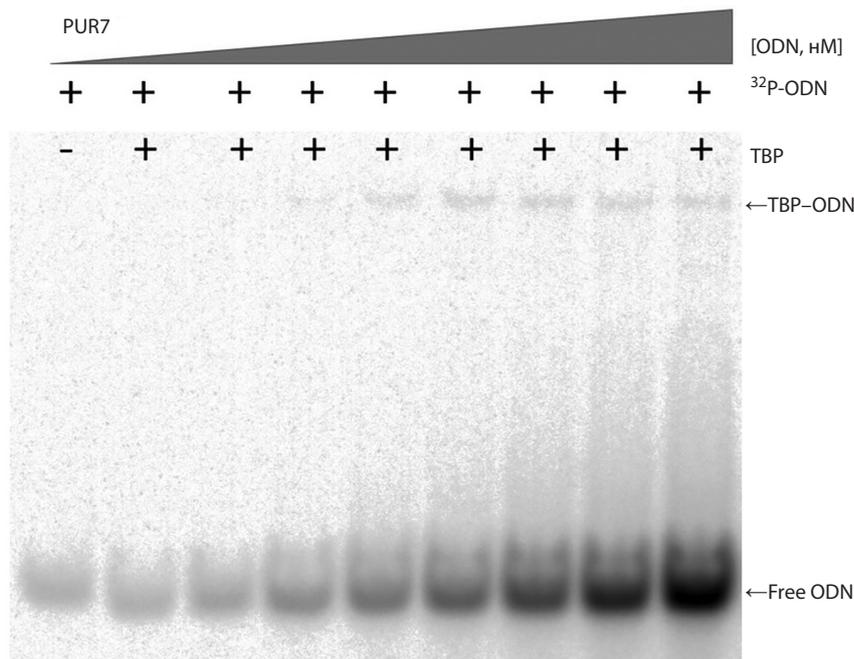


Рис. 1. Электрофореграмма, полученная при измерении равновесной константы диссоциации комплексов ТВР с ТАТА-боксами на примере промотора гена *PUR7* *Arabidopsis*

Fig. 1. Electropherogram obtained by measuring the equilibrium constant dissociation of TBP complexes with TATA boxes using the example of the *Arabidopsis PUR7* gene promoter

сайта связывания фактора транскрипции в активность промотора для подтверждения его функциональности.

В настоящей работе для экспериментального определения сродства ТВР/ТАТА мы использовали олигодезоксирибонуклеотиды (oligodeoxyribonucleotides, ODN), идентичные ТАТА-боксам промоторов генов *Arabidopsis*, экспрессирующихся в разных органах растения (корнях, стебле, листьях, цветках), и рекомбинантный ТВР человека. ТАТА-связывающий белок человека в работе применяли на основании высокой гомологии ТВР эукариот и, в частности, > 80 % гомологии карбокситерминального конца ТВР *Arabidopsis* и человека (животных и дрожжей), участвующего в связывании ТАТА-боксов, и их взаимозаменяемости в системах транскрипции *in vitro* и *in vivo* (Holdsworth et al., 1992; Mukumoto et al., 1993; Iwataki et al., 1997; Yamaguchi et al., 1998). С помощью разработанного нами веб-сервиса *Plant_SNP_TATA_Z-tester* (Rasskazov et al., 2022) сделаны прогнозы *in silico* зависимости сродства ТАТА-боксов к ТВР от последовательности ТАТА-боксов и его окружения и проведено их сравнение с экспериментальными данными.

Материалы и методы

Последовательности ДНК. Последовательности ДНК ТАТА-боксов генов *A. thaliana* (табл. 1) взяты из базы данных TAIR (Garcia-Hernandez et al., 2002).

Анализ последовательности ДНК *in silico*. Мы проанализировали последовательности ДНК перед стартом транскрипции ТАТА-содержащих промоторов генов *Arabidopsis* с использованием разработанного нами веб-сервиса *Plant_SNP_TATA_Z-tester* ([http://www.mgs.bionet.nsc.ru/cgi-](http://www.mgs.bionet.nsc.ru/cgi-bin/mgs/tatascan_plant/start.pl)

[bin/mgs/tatascan_plant/start.pl](http://www.mgs/tatascan_plant/start.pl)) (Rasskazov et al., 2022). Ранее нами был создан веб-сервис *Human_SNP_TATA_Z-tester* для оценки влияния SNP кор-промоторов на связывание ТВР с ТАТА-боксами и, соответственно, на транскрипцию генов и наследственные заболевания (Ponomarenko et al., 2015). Веб-сервис *Plant_SNP_TATA_Z-tester* создан для оценки влияния мутаций проксимальных промоторов генов растений на экспрессию и фенотип; прогнозируемые значения сродства ТВР/ТАТА выражаются в молях на литр.

Синтетические двуцепочечные ODN. ODN, идентичные ТАТА-боксам промоторов генов *Arabidopsis*, синтезированы и очищены в компании «Биосан» (Новосибирск, Россия) (<https://biosan-nsk.ru/>).

Мечение ODN по 5'-концам ³²P-АТФ. В работе использовали ODN длиной 26 пар оснований, синтезированные и дополнительно очищенные электрофорезом в ПААГ. Для получения меченых двуцепочечных ODN обе цепи метили ³²P-АТФ (ИХБФМ СО РАН, Новосибирск, Россия) с помощью Т4-полинуклеотидкиназы (SibEnzyme, Новосибирск, Россия), смешивали в эквимолярном соотношении и отжигали (нагревали до 95 °С и медленно охлаждали до комнатной температуры). Полученные дуплексы очищали и анализировали электрофорезом в 15 % ПААГ в неденатурирующих условиях с последующей радиоавтографией и визуализацией с помощью *Molecular Imager PharoFX Plus*. Получение меченых двуцепочечных олигонуклеотидов подробно описано в статье (Drachkova et al., 2011).

Выделение и очистка рекомбинантного ТАТА-связывающего белка. В работе использовали рекомбинантный ТВР человека, экспрессированный в клетках *Escherichia coli*

Таблица 1. Краткое описание генов, экспрессирующихся в разных органах *Arabidopsis*, промоторы которых использовались для изучения взаимодействия TATA-бокса с TBP

Table 1. Summary of genes expressed in different organs of the *Arabidopsis*, whose promoters were used to study the interaction of the TATA box with TBP

Ген	ID (база данных TAIR)	Описание гена	Литературный источник
<i>PUR7</i>	AT3G21110	Ген <i>PUR7</i> кодирует 5'-фосфорибозил-4-(N-сукцинокарбоксамид)-5-аминоимидазолсинтетазу, участвующую в синтезе пуринов. С использованием трансгенной репортерной системы показано, что наиболее высокий уровень биосинтеза пуринов происходит в митотически активных тканях растения. Наибольшее содержание мРНК <i>PUR7</i> обнаружено в цветках <i>Arabidopsis</i>	Senecoff et al., 1996
<i>TFL1</i>	AT5G03840	Ген <i>TERMINAL FLOWER 1 (TFL1)</i> кодирует белок-гомолог гена <i>FT</i> (флориген) с противоположным действием (ингибирует переход к цветению в результате замены одной аминокислоты). Играет важную роль в поддержании вегетативного состояния апикальной меристемы побега, задерживая цветение, в том числе и у <i>Arabidopsis</i>	Shannon, Meeks-Wagner, 1993; Goretti et al., 2020; Cerise et al., 2023
<i>MLH1</i>	AT4G09140	Ген <i>MUTL-homologue 1 (MLH1)</i> у <i>Arabidopsis</i> кодирует одноименный белок, являющийся гомологом белков MLH1 репарационной системы других эукариот. Исправляет ошибки репликации ДНК. Потеря функциональности гена <i>MLH1</i> приводит к значительному снижению фертильности как у гомозигот, так и у гетерозигот	Dion et al., 2007
<i>FT</i>	AT1G10880	Ген <i>FLOWERING LOCUS (FT)</i> кодирует белок флориген, связывающий фосфатидилэтаноламин, мощный стимулятор цветения	Corbesier et al., 2007; Song et al., 2014; Lv et al., 2021
<i>AOP2</i>	AT4G03060	Ген <i>ALKENYL HYDROXALKYL PRODUCING 2 (AOP2)</i> кодирует одну из диоксигеназ, зависимых от 2-оксиглутарата, катализирует широкий спектр окислительных реакций. Ген является частью регуляторной петли обратной связи, связывающей биосинтез глюкозинолатов и передачу сигналов жасмоновой кислоты	Burow et al., 2015; Das, 2021

BL21(DE3), несущих плазмиду pAR3038-hTBP (любезно предоставлена проф. В. Pugh, Центр регуляции гена, Отделение биохимии и молекулярной биологии, Университет Пенсильвании, США). Экспрессию и очистку TBP проводили согласно (Pugh, 1995), за исключением концентрации изопропилтио-галактопиранозиды (1 вместо 0.1 mM) и времени индукции (3 вместо 1.5 ч).

Определение равновесных констант диссоциации (K_D). Равновесные константы диссоциации комплексов TBP с TATA-содержащими двуцепочечными ODN, характеризующие сродство TBP к ODN, идентичным TATA-боксам с фланкирующими нуклеотидами реальных промоторов, измеряли с применением традиционного подхода, который включал титрование фиксированного количества активного TBP (0.5–1 нМ, в зависимости от нуклеотидного состава ODN), с увеличением концентрации TATA-содержащего ODN до достижения равновесия. Время достижения равновесия определяли предварительно для каждого ODN. Каждое значение K_D определяли в результате не менее трех опытов.

Эксперименты по связыванию TBP/ODN выполняли при 25 °C в буфере (20 mM HEPES-KOH (pH 7.6), 5 mM MgCl₂, 70 mM KCl, 1 mM DTT, 100 мкг/мл BSA, 0.01 % NP-40, 5 % глицерин) до уравнивания. Комплексы TBP/ODN отделяли от несвязанного ODN, применяя метод задержки ДНК в геле (EMSA). Электрофорез проводили с использованием 5 % ПААГ в трис-глициновом буфере (pH = 8.3) в течение 40 мин при температуре 10 °C и напряженности поля 25 В/см. Гели

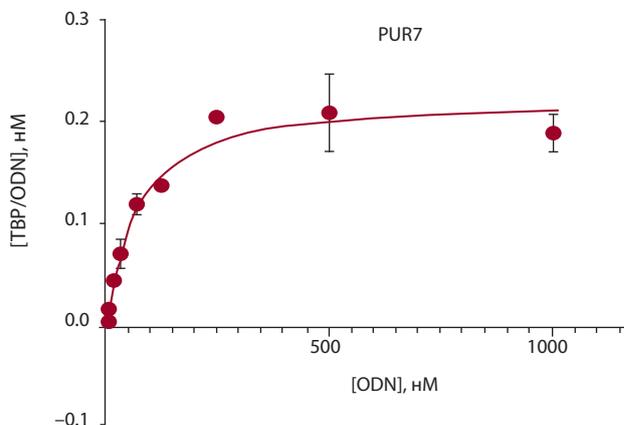


Рис. 2. Изотерма связывания, полученная на основании электрофореграммы (см. рис. 1)

Fig. 2. Binding isotherm obtained from the electropherogram (Fig. 1)

высушивали и экспонировали с экраном Imaging Screen-K (Kodak, США) для сканера Molecular Imager PharoFX Plus (Bio-Rad, США). Экран сканировали на фосфоимиджере и осуществляли количественный анализ радиоавтографов с помощью программы Quantity One – v. 4.5.0 (Bio-Rad). Расчет равновесных констант диссоциации комплексов TBP/ODN проводили с использованием программы GraphPad Prism 5 (<http://graphpad-prism.software.informer.com/5.01>).

Статистический анализ данных. Сравнение наших прогнозов с экспериментальными значениями сродства

Таблица 2. Равновесные константы диссоциации (K_D) комплексов TBP/TATA *A. thaliana*, предсказанные с помощью веб-сервиса Plant_SNP_TATA_Z-tester (Rasskazov et al., 2022) и определенные *in vitro* с применением количественного метода EMSA

Table 2. Equilibrium dissociation constants (K_D) of *A. thaliana* TBP/TATA complexes, predicted by the Plant_SNP_TATA_Z-tester Web service (Rasskazov et al., 2022) and determined *in vitro* using the quantitative EMSA method

Ген	ID (база данных TAIR)	Последовательность ODN 5' – 3', п. о.	Plant_SNP_TATA Z tester, нМ	Эксперимент <i>in vitro</i> K_D , нМ
<i>PUR7</i>	AT3G21110	AATTCGTC CTACACTT CCCCCACC	8.28	60 ± 20
<i>TFL1</i>	AT5G03840	GGAAAACCC CTATAAAT AGATGTCTC	2.10	0.9 ± 0.2
<i>MLH1</i>	AT4G09140	GATGACAA ATTATA ACCCTAAATAA	3.99	10 ± 4
<i>FT</i>	AT1G10880	ATAAGTGTATGC ATTTATA GTGCATA	2.10	9 ± 4
<i>AOP2</i>	AT4G03060	GAAGGGTCC CCATTTTATA GAGAAAT	2.17	5.4 ± 3
Коэффициент линейной корреляции Пирсона			$r = 0.97$	
Уровень статистической значимости			$p < 0.05$	

Примечание. Курсивом выделена последовательность TATA-бокса.

TBP/TATA выполняли с помощью пакета программ Statistica (Statsoft™, США).

Результаты и обсуждение

С помощью веб-сервиса Plant_SNP_TATA_Z-tester (Rasskazov et al., 2022) мы сделали прогнозы *in silico* значений равновесных констант диссоциации, характеризующих сродство TBP к TATA-боксам, и затем экспериментально определили значения K_D с применением количественного метода EMSA с использованием ODN длиной 26 пар оснований, идентичных TATA-боксам с фланкирующими областями промоторов реальных генов *A. thaliana*. На рис. 1 в качестве примера приведена электрофореграмма, полученная при измерении K_D . На рис. 2 показана изотерма связывания, определенная из электрофореграммы с помощью программы GraphPad Prism 5. Для гена *PUR7* $K_D = 60 \pm 20$ нМ.

В табл. 2 представлены прогнозы сродства TBP/ODN в нМ (10^{-9} М) и их экспериментальная верификация. Для изучения мы выбрали промоторы с последовательностями TATA-боксов, содержащими разное количество Т-нуклеотидов в TATA-боксах и разное количество ГС-нуклеотидов во фланкирующих последовательностях.

Как видно из табл. 2, почти все последовательности содержат похожие TATA-элементы (за исключением гена *PUR7*), в той или иной степени гомологичные консенсусу TATA-бокса. Однако сродство между ними различается в 1.2–66 раз. Если сравнивать влияние замен в TATA-боксе *Arabidopsis* и человека – $T_1A_2T_3A_4A_5A_6A_7$, то самыми консервативными, с наибольшим влиянием на транскрипцию и экспрессию являются замены во 2, 3 и 4-й позициях А и Т на G или С, что их объединяет. Что касается T_3 в последовательности TATA-бокса, то он имеет максимальный вес среди нуклеотидов TATA-бокса (Bucher, 1990), поэтому его замена очень часто оказывает существенное влияние на сродство TBP-TATA и транскрипцию. Нуклеотид A_2 имеет более низкий вес (Bucher, 1990), а A_4 , согласно весовой матрице Шахмурадова для TATA-боксов растений (Shahmuradov et al., 2003), – даже чуть выше, чем T_3 .

Менее значительное влияние оказывают замены в 1, 6 и 7-й позициях. Следует отметить, что влияние замены

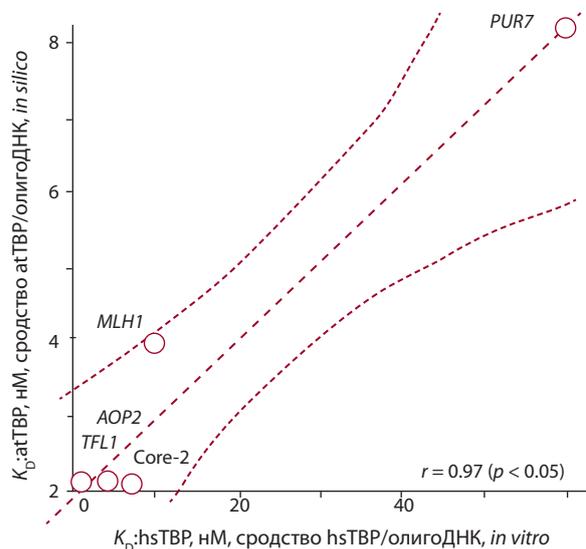


Рис. 3. Корреляция между предсказанными *in silico* и измеренными *in vitro* значениями сродства TBP к TATA-боксам.

AtTBP/олигоДНК – сродство TBP *Arabidopsis* к ODN; hsTBP/олигоДНК – сродство TBP человека к ODN

Fig. 3. Correlation between predicted *in silico* and measured *in vitro* values of TBP affinity for TATA boxes.

AtTBP/oligo DNA – affinity of *Arabidopsis* TBP to ODN; hsTBP/oligoDNA – affinity human TBP to ODN

первого нуклеотида Т на С в TATA-боксе имеет противоречивые результаты. Так, в работе (Stewart et al., 2006) было показано, что эта замена практически не влияет на связывание. J.J. Stewart и L.A. Stargell (2001) ранее выявили, что промотор, содержащий последовательность CATAAAA, в 40 раз менее эффективен для индукции транскрипции *in vivo*, чем классическая последовательность TATAAAA. В работе (Wobbe, Struhl, 1990) продемонстрировано, что транскрипция в клетках *HeLa* снижается в 3 раза при замене первого Т на С. Вероятно, различия обусловлены разным контекстом, в котором находятся последовательности TATA-боксов, и/или разным составом белков ядерных экстрактов.

Самое низкое сродство TBP/TATA (60 нМ) определено для гена *PUR7*, последовательность TATA-бокса которого осно-

вательно изменена: вместо консервативного А во второй позиции наблюдается С, вместо консервативных ТА в 7-й и 8-й позициях находится СТ.

Содержание G- и C-нуклеотидов во фланках всех использованных TATA-боксов колеблется от 6 до 10.6 % в генах *MLH1* и *COR2E*, у *PUR7* – 12 GC. Сравнение среднего содержания GC-нуклеотидов во фланкирующих последовательностях TATA-боксов генов человека (Savinkova et al., 2013) и *Arabidopsis* показывает, что в генах человека оно выше на 23 %: у человека оно равно в среднем 12, а у *Arabidopsis* – 9. Как известно, взаимодействие TBP с TATA-элементом в значительной степени зависит от фланкирующих последовательностей, мотивы которых уникальны для определенных TATA-боксов (Faiger et al., 2006). В ряде работ изучали содержание GC-нуклеотидов в кодирующих и некодирующих областях *Arabidopsis*; обнаружена обратная зависимость экспрессии генов от их процентного содержания (Morey et al., 2011; Das, Bansal, 2019). Но содержание GC во фланкирующих последовательностях TATA-боксов и их влияние на сродство TBP-TATA не рассматривались.

Как видно из рис. 3, наши предсказанные и экспериментальные значения для сродства (K_D) TBP/ODN хорошо коррелируют друг с другом. Коэффициент линейной корреляции r для предсказанных K_D и их экспериментальных значений, выраженных в нМ, составил 0.97 при $p < 0.05$.

Заключение

Мы представили результаты первого экспериментального систематического изучения с количественным определением сродства TBP с TATA-боксами промоторов реальных генов *A. thaliana*. Из полученных данных по взаимодействию TBP-TATA можно видеть, что ODN, идентичные последовательностям TATA-боксов с фланкирующими нуклеотидами промоторов генов *Arabidopsis*, имеют сродство к TBP человека, различающееся до 66 раз и, соответственно, различающиеся уровни транскрипции и экспрессии генов (Mogno et al., 2010) в разных тканях растения. Фланкирующие последовательности TATA-боксов промоторов рассмотренных генов *Arabidopsis* имеют более низкое содержание GC-нуклеотидов во фланках (на 23 %) по сравнению с генами человека. Наши данные при сравнении с аналогичными результатами, полученными с TBP *Arabidopsis*, позволят точно выяснить, есть ли разница и насколько она значительна во взаимодействии TBP-TATA при замене TBP человека на TBP *Arabidopsis*.

В свете широко применяемой технологии редактирования геномов сельскохозяйственных растений данные расширенной и более точной картины взаимодействия TBP с TATA-боксами и их вариантами с использованием TBP *Arabidopsis* могут быть применены для регулирования урожайности важных продовольственных культур. Полученные результаты открывают возможность оценки *in silico* сродства TBP *Arabidopsis* к промоторам генов этого растения с помощью разработанного нами веб-сервиса Plant_SNP_TATA_Z-tester (Rasskazov et al., 2022). С учетом консервативности TBP высших растений (Holdsworth et al., 1992; Mukumoto et al., 1993; Iwataki et al., 1997; Yamaguchi et al., 1998; Porto et al., 2014) предложенный метод *in silico* оцен-

ки сродства TBP к промоторам может применяться также и к анализу промоторов генов других растений в пределах этого таксона.

Список литературы / References

- Bucher P. Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *J Mol Biol.* 1990;212(4):563-578. doi 10.1016/0022-2836(90)90223-9
- Burley S.K., Roeder R.G. Biochemistry and structural biology of transcription factor IID (TFIID). *Annu Rev Biochem.* 1996;65:769-799. doi 10.1146/annurev.bi.65.070196.004005
- Burrow M., Atwell S., Francisco M., Kerwin R.E., Halkier B.A., Kliebenstein D.J. The glucosinolate biosynthetic gene *AOP2* mediates feedback regulation of jasmonic acid signaling in *Arabidopsis*. *Mol Plant.* 2015;8(8):1201-1212. doi 10.1016/j.molp.2015.03.001
- Corbese M., da Silveira Falavigna V., Rodríguez-Maroto G., Signal A., Severing E., Gao H., van Driel A., Vincent C., Wilkens S., Iacobini F.R., Formosa-Jordan P., Pajoro A., Coupland G. Two modes of gene regulation by TFL1 mediate its dual function in flowering time and shoot determinacy of *Arabidopsis*. *Development.* 2023;150(23):dev202089. doi 10.1242/dev.202089
- Corbesier L., Vincent C., Jang S., Fornara F., Fan Q., Searle I., Giakountis A., Farrona S., Gissot L., Turnbull C., Coupland G. FT protein movement contributes to long-distance signaling in floral induction of *Arabidopsis*. *Science.* 2007;316(5827):1030-1033. doi 10.1126/science.1141752
- Das B. Glucosinolate biosynthesis: role of MAM synthase and its perspectives. *Biosci Rep.* 2021;41(10):BSR20211634. doi 10.1042/BSR20211634
- Das S., Bansal M. Variation of gene expression in plants is influenced by gene architecture and structural properties of promoters. *PLoS One.* 2019;14(3):e0212678. doi 10.1371/journal.pone.0212678
- Dikstein R. The unexpected traits associated with core promoter elements. *Transcription.* 2011;2(5):201-206. doi 10.4161/trns.2.5.17271
- Dion E., Li L., Jean M., Belzile F. An *Arabidopsis* MLH1 mutant exhibits reproductive defects and reveals a dual role for this gene in mitotic recombination. *Plant J.* 2007;51(3):431-440. doi 10.1111/j.1365-3113.2007.03145.x
- Drachkova I.A., Ponomarenko P.M., Arshinova T.V., Ponomarenko M.P., Suslov V.V., Savinkova L.K., Kolchanov N.A. *In vitro* examining the existing prognoses how TBP binds to TATA with SNP associated with human diseases. *Health.* 2011;3(9):577-583. doi 10.4236/health.2011.39099
- Faiger H., Ivanchenko M., Cohen I., Haran T.E. TBP flanking sequences: asymmetry of binding, long-range effects and consensus sequences. *Nucleic Acids Res.* 2006;34:104-119. doi 10.1093/nar/gkj414
- García-Hernández M., Berardini T.Z., Chen G., Crist D., Doyle A., Huala E., Knee E., Lambrecht M., Miller N., Mueller L.A., Mundodi S., Reiser L., Rhee S.Y., Scholl R., Tacklind J., Weems D.C., Wu Y., Xu I., Yoo D., Yoon J., Zhang P. TAIR: a resource for integrated *Arabidopsis* data. *Funct Integr Genomics.* 2002;2(6):239-253. doi 10.1007/s10142-002-0077-z
- Goretti D., Silvestre M., Collani S., Langenecker T., Méndez C., Madaño F., Schmid M. TERMINAL FLOWER1 functions as a mobile transcriptional cofactor in the shoot apical meristem. *Plant Physiol.* 2020;182(4):2081-2095. doi 10.1104/pp.19.00867
- Hampsey M. Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiol Mol Biol Rev.* 1998;62:465-503. doi 10.1128/MMBR.62.2.465-503.1998
- Holdsworth M.J., Grierson C., Schuch W., Bevan M. DNA-binding properties of cloned TATA-binding protein from potato tubers. *Plant Mol Biol.* 1992;19:455-464. doi 10.1007/BF00023393
- Iwataki N., Hoya A., Yamazaki K. Restoration of TATA-dependent transcription in a heat-inactivated extract of tobacco nuclei by recombinant TATA-binding protein (TBP) tobacco. *Plant Mol Biol.* 1997;34:69-79. doi 10.1023/A:1005759521285
- Juven-Gershona T., Kadonaga J.T. Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev Biol.* 2010;339(2):225-229. doi 10.1016/j.ydbio.2009.08.009
- Kadonaga J.T. Perspectives on the RNA polymerase II core promoter. *Wiley Interdiscip Rev Dev Biol.* 2012;1(1):40-51. doi 10.1002/wdev.21

- Lv X., Zeng X., Hu H., Chen L., Zhang F., Liu R., Liu Y., Zhou X., Wang C., Wu Z. Structural insights into the multivalent binding of the *Arabidopsis FLOWERING LOCUS T* promoter by the CO-NF-Y master transcription factor complex. *Plant Cell*. 2021;33(4):1182-1195. doi 10.1093/plcell/koab016
- Marks R.A., Hotaling S., Frandsen P.B., VanBuren R. Representation and participation across 20 years of plant genome sequencing. *Nat Plants*. 2021;7(12):1571-1578. doi 0.1038/s41477-021-01031-8
- Mogno I., Vallania F., Mitra R.D., Cohen B.A. TATA is a modular component of synthetic promoters. *Genome Res*. 2010;20(10):1391-1397. doi 10.1101/gr.106732.110
- Morey C., Mookherjee S., Rajasekaran G., Bansal M. DNA free energy-based promoter prediction and comparative analysis of *Arabidopsis* and rice genomes. *Plant Physiol*. 2011;156(3):1300-1315. doi 10.1104/pp.110.167809
- Mukumoto F., Hirose S., Imaseki H., Yamazaki K.-I. DNA sequence requirement of a TATA element-binding protein from *Arabidopsis* for transcription *in vitro*. *Plant Mol Biol*. 1993;23:995-1003. doi 10.1007/BF00021814
- Patikoglou G.A., Kim J.L., Sun L., Yang S.H., Kodadek T., Burley S.K. TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. *Genes Dev*. 1999;13(24):3217-3230. doi 10.1101/gad.13.24.3217
- Ponomarenko M., Rasskazov D., Arkova O., Ponomarenko P., Suslov V., Savinkova L., Kolchanov N. How to use SNP_TATA_Comparator to find a significant change in gene expression caused by the regulatory SNP of this gene's promoter via a change in affinity of the TATA-binding protein for this promoter. *Biomed Res Int*. 2015;2015:359835. doi 10.1155/2015/359835
- Porto M.S., Pinheiro M.P., Batista V.G., dos Santos R.C., Filho Pde A., de Lima L.M. Plant promoters: an approach of structure and function. *Mol Biotechnol*. 2014;56(1):38-49. doi 10.1007/s12033-013-9713-1
- Pugh B.F. Purification of the human TATA-binding protein, TBP. In: Tymms M.J. (Ed.) *In Vitro Transcription and Translation Protocols*. (Ser. Methods in Molecular Biology. Vol. 37). Totowa, NJ: Humana Press Inc, 1995;359-367. doi 10.1385/0-89603-288-4:359
- Rasskazov D., Chadaeva I., Sharypova E., Zolotareva K., Khandaev B., Ponomarenko P., Podkolodnyy N., Tverdokhlebs N., Vishnevsky O., Bogomolov A., Podkolodnaya O., Savinkova L., Zemlyanskaya E., Golubyatnikov V., Kolchanov N., Ponomarenko M. Plant_SNP_TATA_Z-tester: a Web service that unequivocally estimates the impact of proximal promoter mutations on plant gene expressions. *Int J Mol Sci*. 2022;23(15):8684. doi 10.3390/ijms23158684
- Rhee H.S., Pugh B.F. Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature*. 2012;483(7389):295-301. doi 10.1038/nature10799
- Sandelin A., Carninci P., Lenhard B., Ponjavic J., Hayashizaki Y., Hume D.A. Mammalian RNA polymerase II core promoters: insights from genome-wide studies. *Nat Rev Genet*. 2007;8:424-436. doi 10.1038/nrg2026
- Savinkova L., Drachkova I., Arshinova T., Ponomarenko P., Ponomarenko M., Kolchanov N. An experimental verification of the predicted effects of promoter TATA-box polymorphisms associated with human diseases on interactions between the TATA boxes and TATA-binding protein. *PLoS One*. 2013;8(2):e54626. doi 10.1371/journal.pone.0054626
- Senecoff J.F., McKinney E.C., Meagher R.B. *De novo* purine synthesis in *Arabidopsis thaliana*. II. The *PUR7* gene encoding 5'-phosphoribosyl-4-(N-succinocarboxamide)-5-aminoimidazole synthetase is expressed in rapidly dividing tissues. *Plant Physiol*. 1996;112(3):905-917. doi 10.1104/pp.112.3.905
- Shahmuradov I.A., Gammerman A.J., Hancock J.M., Bramley P.M., Solov'yev V.V. PlantProm: a database of plant promoter sequences. *Nucleic Acids Res*. 2003;31(1):114-117. doi 10.1093/nar/gkg041
- Shannon S., Meeks-Wagner D.R. Genetic interactions that regulate inflorescence development in *Arabidopsis*. *Plant Cell*. 1993;5(6):639-655. doi 10.1105/tpc.5.6.639
- Song Y.H., Shim J.S., Kinmonth-Schultz H.A., Imaizumi T. Photoperiodic flowering: time measurement mechanisms in leaves. *Annu Rev Plant Biol*. 2014;66:441-464. doi 10.1146/annurev-arplant-043014-115555
- Stewart J.J., Stargell L.A. The stability of the TFIIA-TBP-DNA complex dependent on the sequence of the TATAAA element. *Biol Chem*. 2001;276:30078-30084. doi 10.1074/jbc.M105276200
- Stewart J.J., Fischback J.A., Chen X., Stargell L.A. Non optimal TATA element exhibits diverse mechanistic consequences. *J Biol Chem*. 2006;281:22665-22673. doi 10.1074/jbc.M603237200
- Wobbe C.R., Struhl K. Yeast and human TATA-binding proteins have nearly identical DNA sequence requirements for transcription *in vitro*. *Mol Cell Biol*. 1990;10:3859-3867. doi 10.1128/MCB.10.8.3859
- Yamaguchi Y., Itoh Y., Takeda Y., Yamazaki K. TATA sequence requirements for the initiation of transcription for an RNA polymerase II *in vitro* transcription system from *Nicotiana tabacum*. *Plant Mol Biol*. 1998;38(6):1247-1252. doi 10.1023/a:1006056128129

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 13.11.2024. После доработки 11.01.2025. Принята к публикации 27.01.2025.

 pismavavilov.ru

doi 10.18699/letvjgb-2025-11-07

Дискуссия

Как считать критерий Стьюдента статистической значимости разности средних двух выборок. I. Проверка нормальности

В.М. Ефимов ^{1, 2, 3, 4}✉

Аннотация: Критерий Стьюдента статистической значимости разности средних двух выборок, предложенный Р. Фишером в 1925 г., до сих пор является одним из самых востребованных методов статистического анализа. За почти столетие его использования сложилась устойчивая система рекомендаций и условий его применения, которая изложена в учебниках и руководствах по статистической обработке данных. Как правило, настоятельно требуется предварительная проверка нормальности распределения исходных выборок и равенства их дисперсий. В случае ненормальности рекомендуется использовать непараметрические методы, например критерий Манна–Уитни. В работе представлена более современная точка зрения на эту проблему, обусловленная несколькими взаимосвязанными причинами. Во-первых, за прошедшее столетие накоплен значительный практический опыт применения t -критерия Стьюдента, который заставляет сильно сомневаться в обязательности проверки нормальности и равенства дисперсий, а также применения ранговых критериев в случае отсутствия нормальности. Во-вторых, теория тоже не стояла на месте. Появились расчет критерия Стьюдента через точно-бисериальный коэффициент корреляции и альтернативы методам «нормальной теории» в виде свободных от распределения процедур. В-третьих, кардинально выросли вычислительные возможности, позволяющие без дополнительных предположений моделировать в компьютере генеральные распределения исходных выборок и по ним оценивать требуемые p -value.

Ключевые слова: нормальное распределение; точно-бисериальный коэффициент корреляции; свободные от распределения процедуры; p -value; бутстреп

Для цитирования: Ефимов В.М. Как считать критерий Стьюдента статистической значимости разности средних двух выборок. I. Проверка нормальности. *Письма в Вавиловский журнал генетики и селекции*. 2025;11(1):43-50. doi 10.18699/letvjgb-2025-11-07

Финансирование: Работа поддержана бюджетным проектом № FWNR-2022-0019.

Discussion

How to calculate the Student's test for the statistical significance of the difference between the means of two samples. I. Testing for normality

V.M. Efimov ^{1, 2, 3, 4}✉

Abstract: The Student's t -test for the statistical significance of the difference in the means of two samples, proposed by R. Fisher in 1925, is still one of the most popular methods of statistical analysis. Over almost a century of its use, a fairly stable system of recommendations and conditions for its application has developed, which is set out in textbooks and manuals on statistical data processing. As a rule, a preliminary check of the normality of the original sample distributions and the equality of their variances is urgently required. In case of abnormality, it is recommended to use nonparametric methods, for example, the Mann–Whitney test. The paper presents a more modern point of view on this problem, caused by several interrelated reasons. Firstly, a century of practical experience has accumulated in using the Student's t -test, which makes one strongly doubt the necessity of checking normality and equality of variances, as well as the use of rank criteria in the absence of normality. Secondly, the theory has not stood still either. An alternative to the methods of

¹ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

² Институт систематики и экологии животных Сибирского отделения Российской академии наук, Новосибирск, Россия
Institute of Systematics and Ecology of Animals of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

³ Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
Novosibirsk State University, Novosibirsk, Russia

⁴ Национальный исследовательский Томский государственный университет, Томск, Россия
Tomsk State University, Tomsk, Russia

 vmefimov@gmail.com

© Ефимов В.М., 2025

the “normal theory” has appeared in the form of distribution-free procedures. Thirdly, over the past century, computing capabilities have increased dramatically, making it possible to model general distributions of initial samples on a computer without additional assumptions and use them to estimate the required p -values.

Key words: normal distribution; point-biserial correlation coefficient; free-distribution procedures; p -value; bootstrap

For citation: Efimov V.M. How to calculate the Student's test for the statistical significance of the difference between the means of two samples. I. Testing for normality. *Pisma v Vavilovskii Zhurnal Genetiki i Seleksii = Letters to Vavilov Journal of Genetics and Breeding*. 2025;11(1):43-50. doi 10.18699/letvjgb-2025-11-07 (in Russian)

Funding: The work was supported by BP #FWNR-2022-0019.

Введение

Критериев Стьюдента много. Самым востребованным является критерий Стьюдента статистической значимости разности средних двух выборок (далее t -критерий Стьюдента для двух средних), и именно его мы будем рассматривать. Задача заключается в следующем. Даны две совокупности объектов, у которых измерен некоторый количественный признак. Для каждой совокупности вычислены средние значения этого признака. Как правило, они несколько отличаются друг от друга. Требуется принять решение: считаем ли мы разницу между этими средними случайной или нет. Алгоритм должен быть пригодным для использования на практике, т. е. простым и надежным.

В рамках статистической науки проблема формулируется так: имеется нулевая гипотеза (H_0) (обычно отсутствия различий), а наблюдаемые отклонения от нее возникли по случайным причинам. Необходимо принять или отклонить нулевую гипотезу. Для этого надо оценить суммарную вероятность (p -value) наблюдаемых и всех менее вероятных отклонений от H_0 . Если p -value меньше некоторого заранее установленного порога – уровня значимости (α), то нулевая гипотеза отвергается (Fisher, 1925b). На сегодняшний день приняты и повсеместно используются три стандартных уровня статистической значимости: 0.05, 0.01, 0.001.

Трехсотлетняя история возникновения и развития идеи “ $p < 0.05$ ” обстоятельно изложена в работе (Kennedy-Shaffer, 2019). Однако в наше время этот подход вызывает очень сильный протест прикладных статистиков из-за преимущественного использования только p -value, и цитируемая работа является частью их мощного движения против мира “ $p < 0.05$ ”. Протестующие не призывают отменить p -value, но предлагают наряду с ним в обязательном порядке вычислять размер эффекта и доверительные интервалы, а также приводить научные и практические обоснования, например экономическую значимость, принимаемого решения (Wasserstein et al., 2019). Некоторые просто предлагают заменить $p < 0.05$ на $p < 0.005$ (Benjamin et al., 2018). Но редакторы и рецензенты большинства научных журналов еще не в курсе предстоящих перемен и требуют от авторов работать по привычным для всех шаблонам, следуя заветам Фишера столетней давности.

Для средних двух выборок нулевая гипотеза формулируется просто: ($\bar{x}_1 = \bar{x}_2$), или $\bar{x}_1 - \bar{x}_2 = 0$. Для каждой совокупности ($k = 1, 2$) достаточно знать число объектов (N_k), среднее (\bar{x}_k) и сумму квадратов отклонения от среднего (SS_k). Тогда t -критерий Стьюдента для двух средних вычисляется по формуле, предложенной Фишером (Fisher, 1925a):

$$t = \frac{(\bar{x}_1 - \bar{x}_2) \cdot \sqrt{N_1 + N_2 - 2}}{\sqrt{SS_1 + SS_2}} \cdot \sqrt{\frac{N_1 N_2}{N_1 + N_2}}, \quad (1)$$

$$\bar{x}_k = \frac{\sum \bar{x}_i}{N_k}, \quad SS_k = \sum (x_i - \bar{x}_k)^2.$$

Статистика t подчиняется t -распределению Стьюдента с $df = N_1 + N_2 - 2$ степенями свободы. Само распределение и первый t -критерий Стьюдента были найдены ранее при решении более простой задачи Уильямом Сили Госсетом. ‘Student’ – псевдоним Госсета. Задача заключалась в оценке статистической значимости отклонения средней одной выборки от константы (Student, 1908). Позже оказалось, что этому распределению подчиняется множество критериев, полученных другими авторами (в частности, t -критерии статистической значимости линейного коэффициента корреляции Пирсона между двумя количественными признаками и точно-бисериального коэффициента корреляции между количественным признаком и двоичным) (Кендалл, Стюарт, 1973), и многие из них тоже стали называться t -критериями Стьюдента. (Возможно, во имя справедливости и во избежание путаницы стоило бы называть критерий (1) t -критерием Фишера или критерием Стьюдента–Фишера.)

Проверка нормальности и равенства дисперсий

До появления компьютеров исследователи были вынуждены пользоваться печатными таблицами, в которых приводились значения нормального, χ^2 , t и некоторых других распределений для разных уровней значимости α и степеней свободы df . Госсет лично вычислил и напечатал первые таблицы своего распределения, назвав его z -распределением и придавая из практических соображений особое значение малым степеням свободы, начиная с $N = 4$ (Student, 1917). Его таблицы не пользовались популярностью. Фишер переформатировал таблицу Госсета под свою концепцию оценки значимости, которая на практике почти полностью свелась к использованию “ $p < 0.05$ ”, переименовал z - в t -распределение (поскольку у него было свое z -распределение) и включил в свой знаменитый учебник, выдержавший около полутора десятков изданий (Fisher, 1925b). Свой выбор $\alpha = 0.05$ он обосновывал тем, что $t(0, N) = 1.96 \approx 2$ при больших N , и тогда 95 % доверительный интервал для среднего \bar{x} практически равен $\pm 2\sigma_{\bar{x}}$. Фишер считал это разумным и весьма удобным для исследователей. Исследователи действительно приняли это на ура, так же как и идею “ $p < 0.05$ ”. Кроме того, Фишер придавал большое значение проверке нормальности выборок, а также равенства их дисперсий, поскольку все теоретические результаты были получены им именно при этих

предположениях, а практики применения критерия, естественно, еще не было и не могло быть. К сожалению, несмотря на уже столетие практического применения t -критерия Стьюдента, этого и сегодня все еще строго требуют авторы многих статистических учебников и руководств, а также редакторы и рецензенты большинства научных журналов, добавляя, что в случае ненормальности надо использовать непараметрические критерии.

Устаревшая рекомендация проверять нормальность обеих выборок – чисто математическая (Лойко и др., 2019), поскольку t -критерий Стьюдента для двух средних выведен Фишером именно при этом предположении. Причем это предположение нужно только для того, чтобы получить нормальность распределения разности средних. В формуле (1) вообще не используются значения самих выборок, есть только параметры, полученные из этих значений: объемы, средние и суммы квадратов отклонений от средней. Поэтому не очень-то и важно, как распределены сами значения. Но Фишер всегда стремился к математической строгости, в отличие от Госсета, которому было достаточно, чтобы формула хорошо работала на практике, даже если она просто угадана. И предложение перейти в случае ненормальности выборок к непараметрическим критериям (т. е. заменить значения их рангами) – тоже чисто математическое, распределение рангов всегда известно заранее, поэтому вопрос всегда решается до конца (Copover, 2012). Математически очень удобно. А то, что при этом происходит подмена понятий на предметном уровне и изучается уже не совсем та задача, которая на самом деле интересует практиков, математики просто не замечают, поскольку это не их проблема. Хотя разность средних рангов, вообще говоря, совсем не то же самое, что разность самих средних, например, в случае, когда в выборке много малых значений, близких друг к другу, и мало больших.

Вопрос о том, надо ли все-таки проверять нормальность, возникал много раз. Достаточно указать на 31-ю главу второго тома знаменитого трехтомника Кендалла–Стюарта, полностью посвященную этому вопросу. Авторы приходят к четкому мнению: критерии типа критерия Стьюдента, касающиеся генеральных средних, довольно нечувствительны к отклонениям от нормальности. Особенно выделяются два случая: «...если объемы выборок равны, то даже асимметрия исходного распределения вызывает малое отклонение от нормальной теории. Если исходное распределение симметрично, то критерий будет устойчив даже при неодинаковых объемах выборок» (Кендалл, Стюарт, 1973).

На практике критерий Стьюдента достаточно хорошо работает для любых реально встречающихся непрерывных распределений независимо от того, нормальны они или нет. А также для распределений, которые по определению не могут быть нормальными, например ранговых или двоичных, но, заметим, их средние являются количественными признаками. Причина проста: распределение любых средних всегда приближенно нормально в силу центральной предельной теоремы. Соответственно, разность средних тоже будет всегда распределена приближенно нормально. Поэтому даже в случае отклонения от нормальности распределения исходных выборок больших погрешностей не

ожидается. Более того, еще сам Госсет указывал, что в этом случае (формула (1)) гораздо более вероятно занижение статистической значимости, а не ее завышение, так как большие отклонения в данных, увеличивая числитель, еще больше увеличивают знаменатель в силу его квадратичности. Некоторое завышение статистической значимости происходит только в случае, когда более пологие хвосты асимметричных распределений («треугольники») ориентированы навстречу друг другу, но такая ситуация крайне редко встречается в реальных данных. А отношение дисперсий в F -тесте Фишера, наоборот, очень неустойчиво по той же причине, и его вообще не стоит применять для проверки равенства дисперсий. Любопытно, что для того, чтобы прийти к этим выводам без строгой математики, Госсет использовал численное моделирование распределений на основе случайного выбора чисел, написанных на отдельных листочках бумаги, – по сути метод Монте-Карло, причем задолго до появления компьютеров (Lehmann, 1999). Современная наука приходит к тем же выводам (Кендалл, Стюарт, 1973; Lumley et al., 2002; Лемешко Б.Ю., Лемешко С.Б., 2008; Орлов, 2020).

Как практик, Госсет считал, что гораздо важнее оценивать размер эффекта, чем его статистическую значимость. А насчет проверки нормальности писал, обращаясь к Фишеру (через *Nature*): «...вопрос о применимости нормальной теории к ненормальному материалу имеет большое значение и заслуживает внимания как со стороны математики, так и со стороны тех из нас, в чьей компетенции лежит применение результатов его трудов в практической работе. Лично я всегда считал, возможно, без каких-либо вполне определенных оснований для этой веры, что на самом деле на распределение Стьюдента будут очень мало влиять те небольшие отклонения от нормальности, которые имеют место в большинстве биологических и экспериментальных исследований, и недавние работы с небольшими выборками подтверждают меня в этом убеждении» (Student, 1929).

Да и точка зрения Фишера со временем стала более реалистичной: «В основе статистической оценки средних лежит следующее основное положение: если некоторая величина распределена нормально с дисперсией σ^2 , то средняя случайной выборки, состоящей из n наблюдений, распределена нормально с дисперсией σ^2/n . Практическое значение этого положения отчасти усиливается еще тем фактом, что даже если исходное распределение не будет точно нормальным, все же распределение средней стремится к нормальной форме при возрастании выборки. Таким образом, это положение имеет широкую область применения и остается правомерным и в тех случаях, когда мы не имеем достаточных оснований считать, что исходное распределение нормально, лишь бы при этом у нас была уверенность в том, что оно не принадлежит к тому исключительному классу распределений, при которых распределение средней не стремится к нормальному» (Фишер, 1958).

Современная Интернет-энциклопедия уже не считает, что должны быть распределены нормально сами выборки: «Для применения данного критерия необходимо, чтобы выборочные средние имели нормальное распределение». Но дальше все по-старому: «При маленьких выборках это

означает требование нормальности исходных значений. В случае применения двухвыборочного критерия для независимых выборок также необходимо соблюдение условия равенства дисперсий. Существуют, однако, альтернативы критерию Стьюдента для ситуации с неравными дисперсиями. Также не вполне корректно применять *t*-критерий Стьюдента при наличии в данных значительного числа выбросов. При несоблюдении этих условий при сравнении выборочных средних должны использоваться аналогичные методы непараметрической статистики, среди которых наиболее известными являются U-критерий Манна-Уитни (в качестве двухвыборочного критерия для независимых выборок...)» (https://руни.рф/Т-критерий_Стьюдента).

Как видно из этого текста, несмотря на некоторые подвижки, многие догматические требования, к сожалению, остались.

Дихотомия (разделение на две части)

Вычислять *t*-критерий Стьюдента для двух средних можно по-разному. Во втором томе трехтомника Кендалла-Стюарта приведена формула 26.80 (исправлена опечатка):

$$\frac{r_{pb}^2}{1-r_{pb}^2} = \frac{(\bar{x}_1 - \bar{x}_2)^2}{SS_1 + SS_2} \cdot \frac{N_1 N_2}{N_1 + N_2} = \frac{t^2}{N_1 + N_2 - 2}, \quad (2a)$$

где r_{pb} – точно-бисериальный коэффициент корреляции между количественным признаком и двоичным(!). Она же (в словесной формулировке) приведена в Википедии (https://en.wikipedia.org/wiki/Point-biserial_correlation_coefficient).

Из формулы (2a) следует

$$t^2 = \frac{(N_1 + N_2 - 2) \cdot r_{pb}^2}{1 - r_{pb}^2}. \quad (2b)$$

Это означает, что вычислять *t*-критерий Стьюдента для двух средних можно и через точно-бисериальный коэффициент корреляции признаков x и b . Количественный признак (x) получается, если объединить значения обеих выборок в одну. Двоичный (b) принимает значение 1, если значение x_i объединенной выборки относится к первой из выборок, и 0, если ко второй. Между количественным признаком (x) и двоичным (b) можно вычислить точно-бисериальный коэффициент корреляции r_{pb} либо по формуле, выведенной исходя из биномиального распределения двоичного признака b (Кендалл, Стюарт, 1973):

$$r_{pb} = \frac{(\bar{x}_1 - \bar{x}_2) \sqrt{b(1-b)}}{\sqrt{\frac{\sum(x_i - \bar{x})^2}{N}}} = \frac{(\bar{x}_1 - \bar{x}_2) \sqrt{\left(\frac{N_1 N_2}{N}\right)}}{\sqrt{\sum x_i^2 - N \bar{x}^2}}, \quad (3a)$$

либо по обычной формуле Пирсона для двух количественных признаков:

$$r_{pb} = \frac{\sum(x_i - \bar{x})(b_i - \bar{b})}{\sqrt{\sum(x_i - \bar{x})^2} \cdot \sqrt{\sum(b_i - \bar{b})^2}} = \frac{\sum x_i b_i - N \bar{x} \bar{b}}{\sqrt{\sum x_i^2 - N \bar{x}^2} \cdot \sqrt{\sum b_i^2 - N \bar{b}^2}}, \quad (3b)$$

которые эквивалентны друг другу.

Тогда, используя

$$N = N_1 + N_2; \quad \bar{x} = \frac{N_1 \bar{x}_1 + N_2 \bar{x}_2}{N}; \quad \bar{b} = \frac{N_1 \bar{b}_1 + N_2 \bar{b}_2}{N} = \frac{N_1}{N},$$

$$\sum x_i b_i = N_1 \bar{x}_1; \quad \sum x_i^2 = SS_1 + N_1 \bar{x}_1^2 + SS_2 + N_2 \bar{x}_2^2,$$

получим

$$\begin{aligned} \frac{r_{pb}^2}{1-r_{pb}^2} &= \frac{(\sum x_i b_i - N \bar{x} \bar{b})^2}{(\sum x_i^2 - N \bar{x}^2) \cdot (\sum b_i^2 - N \bar{b}^2) - (\sum x_i b_i - N \bar{x} \bar{b})^2} = \\ &= \frac{(N_1 \bar{x}_1 - N \left(\frac{N_1 \bar{x}_1 + N_2 \bar{x}_2}{N}\right) \left(\frac{N_1}{N}\right))^2}{(SS_1 + N_1 \bar{x}_1^2 + SS_2 + N_2 \bar{x}_2^2 - N \bar{x}^2) \left(N_1 - N \left(\frac{N_1}{N}\right)^2\right) - \left[\frac{N_1 N_2}{N} (\bar{x}_1 - \bar{x}_2)\right]^2} = \\ &= \frac{\left[\frac{N_1 N_2}{N} (\bar{x}_1 - \bar{x}_2)\right]^2}{\left(SS_1 + SS_2 + \frac{N_1 N_2}{N} (\bar{x}_1 - \bar{x}_2)^2\right) \frac{N_1 N_2}{N} - \left[\frac{N_1 N_2}{N} (\bar{x}_1 - \bar{x}_2)\right]^2} = \\ &= \frac{\frac{N_1 N_2}{N} (\bar{x}_1 - \bar{x}_2)^2}{SS_1 + SS_2} = \frac{(\bar{x}_1 - \bar{x}_2)^2}{SS_1 + SS_2} \cdot \frac{N_1 N_2}{N}. \quad (4) \end{aligned}$$

Из формулы (4) следует, что формулы (1) и (2a) математически эквивалентны друг другу. Следовательно, оценивать достоверность разности средних двух выборок можно по любой из них. Все это отмечено также в трехтомнике Кендалла, Стюарта (разделы 26.34, 26.35). Более того, формула (2b) статистической значимости точно-бисериального коэффициента корреляции выведена исходя из биномиального распределения двоичного признака b . Но заметим, что именно по этой формуле вычисляется *t*-критерий Стьюдента статистической значимости линейного коэффициента корреляции Пирсона между двумя количественными признаками, подчиняющимися двумерному нормальному распределению (Крамер, 1975, разд. 29.7).

Таким образом, и для двух количественных признаков, подчиняющихся двумерному нормальному распределению, и для ситуации, когда один из признаков является количественным и разбит на две части, по отдельности подчиняющиеся нормальному распределению, а второй – двоичным, который подчиняется биномиальному, а не нормальному распределению, профессионалы-статистики утверждают, что применим один и тот же *t*-критерий Стьюдента статистической значимости коэффициента корреляции. Заметим, что в случае действительного различия средних распределение объединенного количественного признака становится бимодальным (двугорбым), т. е. явно не нормальным. А поскольку второй признак двоичный, он тоже никак не может иметь нормальное распределение. Следовательно, *t*-критерий Стьюдента статистической значимости линейного коэффициента корреляции Пирсона между двумя количественными признаками вполне закономерно применим даже для двух заведомо ненормальных признаков (один двоичный, другой количественный, но бимодальный). Применять, действительно, можно и в теории, и на практике, но почему тогда от нас при расчете корреляций всегда требуют проверить нормальность исходных признаков?

И этот же t -критерий математически эквивалентен t -критерию Стьюдента статистической значимости разности средних. Стоит ли после этого удивляться, что такой универсальный критерий хорошо работает и во многих других, отклоняющихся от нормальности ситуациях?

Таким образом, нормальность проверять, как правило, незачем. Применяя t -критерий Стьюдента даже для явно ненормально распределенных данных, мы практически никогда не зависим достоверность наших результатов. По крайней мере, не зависим сильно.

С равенством дисперсий дело обстоит хуже. Хорошо бы его все-таки проверять, но все критерии проверки равенства дисперсий сами по себе крайне неустойчивы, в частности обычно рекомендуемый F -критерий отношения дисперсий Фишера, и требуют большого числа наблюдений (Орлов, 2020). По этой причине многие авторы рекомендуют вместо t -критерия Стьюдента для двух средних использовать t -критерий Уэлча (Welch, 1938), который тоже выведен из предположения нормальности распределения выборок (и тоже в нем не нуждается по той же причине):

$$t_w = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{SEM_1^2 + SEM_2^2}}, \quad (5)$$

$$s_k^2 = \frac{SS_k}{N_k - 1}, \quad SEM_k^2 = \frac{s_k^2}{N_k} = \frac{SS_k}{N_k(N_k - 1)},$$

где s_k^2 – выборочная дисперсия, SEM (standard error of mean) – стандартная ошибка среднего.

Но t -критерий Уэлча удобен тем, что изначально не требует равенства дисперсий. Несколько раньше Уэлча этот же критерий предложил аспирант Фишера Смит (Smith, 1936a; см. Davenport, Webster, 1975). (Возможно, селекционерам интересно, что этот же Смит является автором знаменитого селекционного индекса Смита–Хазеля (Smith, 1936b; Hazel, 1943).)

Некоторое неудобство доставляет расчет степеней свободы для t -критерия Уэлча:

$$df = \frac{(SEM_1^2 + SEM_2^2)^2}{\frac{(SEM_1^2)^2}{N_1 - 1} + \frac{(SEM_2^2)^2}{N_2 - 1}}. \quad (6)$$

К счастью, при $N_1 = N_2$ оба критерия совпадают, так что в наиболее важном для практики случае – контроль и опыт, когда размеры выборок, как правило, выбираются равными, – можно спокойно использовать обычный t -критерий Стьюдента, не обращая внимания на любую разницу дисперсий. А также, как мы уже знаем, когда распределения исходных выборок симметричны.

Любопытно, что советские биологи, когда им разрешили использовать «буржуазную статистику» (см. предисловие к (Фишер, 1958)), применяли при счете вручную именно t -критерий Уэлча (формула (5)), но с $df = N_1 + N_2 - 2$ степенями свободы, искренне считая, что это и есть «настоящий» t -критерий Стьюдента. Причина кроется в отечественных учебниках биометрии того времени. В них задача сравнения

средних вполне логично сводилась к отношению разности средних и стандартного отклонения этой разности. В случае нормального распределения разности это отношение, конечно, подчиняется t -распределению и, следовательно, может служить статистическим t -критерием. Упускался маленький математический нюанс: для того чтобы быть t -критерием Стьюдента именно с $df = N_1 + N_2 - 2$ степенями свободы, нужно было дополнительно предположить равенство дисперсий двух выборок, как это сделал Фишер, и ввести его в формулу (1). А без этого предположения такое отношение как раз является t -критерием Уэлча (5), но тогда число степеней свободы должно определяться по формуле (6). Считать вручную формулу (6) было довольно затруднительно, да ее и не знали. Большой беды от этого не было, всего лишь чуть-чуть завышалось число степеней свободы, а следовательно, и статистическая значимость разности средних. При равенстве объемов выборок все просто совпадало.

Когда появились большие компьютеры (типа БЭСМ-6), ситуация изменилась не сильно. Ввиду полного отсутствия готового программного обеспечения каждый исследователь писал себе программы сам. А это еще надо было уметь. Да и вычислительный ресурс был крайне ограничен. Например, три БЭСМ-6 с объемом ОЗУ каждой менее 800 килобайт(!) обслуживали почти весь Академгородок, включая все институты биологического профиля. И только с переходом в начале 1990-х на персональные компьютеры из Юго-Восточной Азии и западные статистические пакеты возникла возможность более-менее адекватно обрабатывать реальные биологические данные. К сожалению, своих (отечественных) персональных компьютеров и статистических пакетов с того времени и до сих пор практически так и не появилось, поскольку в этом вроде бы и не было необходимости. Но сейчас, в изменившихся международных условиях, это может стать большой проблемой.

Статпакеты за редким исключением пишут не статистики, а программисты, иногда биологи (например, PAST (Hammer et al., 2001)), добросовестно копируя статистические учебники. А учебники всегда отстают от современного уровня, отраженного в статьях, на одно, а то и два поколения ученых. И конечно, в этих статпакетах запрограммирован более старый t -критерий Стьюдента (формула (1)), а критерий Уэлча (формула (5)) либо не назван прямо (а назван, например, t -критерием с неравными дисперсиями), либо не запрограммирован вообще. Вместе с пакетами появились и руководства к ним, в которых, конечно, требуются ненужные проверки нормальности и равенства дисперсий. И конечно, во всех пакетах, если у вас вычислен выборочный коэффициент корреляции r между двумя признаками вместе с его достоверностью p -value, то можете не сомневаться, что он вычислен с использованием формулы (2b):

$$t = \frac{r\sqrt{N-2}}{\sqrt{1-r^2}}, \quad (2c)$$

хотя еще тот же Фишер более века назад нашел более точное нормализующее z -преобразование для коэффициента корреляции (Fisher, 1915):

$$z = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right); \quad s_z = \frac{1}{\sqrt{N-3}}, \quad (7)$$

где s_z – стандартное отклонение z , а z распределено нормально. Соответственно, отношение $t = z/s_z$ подчиняется t -распределению с $N-3$ степенями свободы. z -преобразование Фишера тоже есть во всех пакетах, но понятно, что если r и p -value для него уже посчитаны компьютером по формуле (2с), то исследователь-практик вряд ли будет прилагать дополнительные усилия, чтобы вычислить достоверность чуточку точнее. По-хорошему, это должно делаться прямо в пакете при вычислении корреляций.

Что действительно надо проверять, так это наличие выбросов. Статистический анализ всегда проводится в предположении некоторого единства или контролируемой исследователем неоднородности данных, например разбиения на группы по их биологическому смыслу. Но даже одна точка, по каким-то случайным или неслучайным причинам далеко отлетевшая от своей группы, может сильно исказить ситуацию. Эту возможность надо проверять, лучше всего визуально, через построение графиков (например, бокс-плота), и до статистической обработки удалять такие точки на том основании, что они не принадлежат основной совокупности. Причем это необязательно будут ошибки измерения или учета, это могут быть и реальные объекты (например, экзотические животные, сбежавшие от своих хозяев, при изучении диких животных в природе). Можно и даже желательно описать их отдельно. Правда, со временем такие объекты могут стать частью местной фауны, как, например, верблюды в Австралии, когда-то завезенные для перевозки грузов (https://ru.wikipedia.org/wiki/Дикие_верблюды_Австралии).

Еще пример из экспедиционного опыта автора: в зоологических отловах обнаружилась трехногая водяная полевка, самка, судя по плацентарным пятнам в матке, рожавшая. Полностью отсутствовала передняя лапа. Никаких повреждений, даже их следов. Нормальная ровная шкурка на том месте, где должна быть лапа. Но описания нового таксона в систематике не случилось. При вскрытии оказалось, что четвертая лапа все же была, но в виде бледной прозрачной полоски, протянувшейся под шкуркой вдоль тела. Однако трехногая жизнь заметно отразилась на параметрах внутренних органов, особенно их симметрии, и поэтому весь набор ее показателей пришлось исключить из дальнейшего статистического анализа.

Для выявления выбросов можно применять и статистические методы. Для этого тоже существуют аналоги t -критериев, например: «Таблицы 4.8. Критерии исключения резко выделяющихся наблюдений» (Большев, Смирнов, 1983). Современная литература на эту тему весьма обширна (Wang et al., 2019; Boukerche et al., 2020).

При ручном счете, который применялся в докомпьютерную эпоху, если все-таки для двух средних считать именно t -критерий Стьюдента, формула (1) была предпочтительнее формулы (2b) из-за удобства вычислений. Появление персональных компьютеров резко изменило ситуацию с расчетами. Во-первых, они стали доступны всем исследователям и сложность формул перестала быть серьезным препятстви-

ем. Соответственно, отпала необходимость пользоваться статистическими таблицами, все необходимые расчеты можно делать в электронных таблицах типа Excel или в специализированных статистических пакетах. Во-вторых, формула (2b) оказалась более удобной при машинном представлении данных в виде таблиц, особенно если измерено несколько признаков у одних и тех же объектов, разбитых на две группы. Тогда один двоичный признак отражает это разбиение сразу для всех признаков и рассчитать статистические значимости разностей сразу для всех признаков можно и удобнее через корреляции (не забывая про необходимость учета множественных сравнений).

Свободные от распределения процедуры

Почему мы всегда исходим из гипотезы нормальности? Почему мы так за нее держимся? В реальности мы, как правило, не знаем не только параметров, но даже формы распределения, которому подчиняется наша исходная выборка. Можем только более или менее уверенно предполагать, что это гипотетическое распределение непрерывно. Можем ли мы подобрать критерий для решения нашей задачи, который не зависит от формы этого распределения? Оказывается, можем, по крайней мере для некоторых задач. Например, для задачи независимости двух количественных признаков. В случае независимости генеральный коэффициент корреляции равен нулю. Поэтому в качестве критерия отклонения от независимости можно взять выборочный коэффициент корреляции r . Можно даже вычислить его распределение, причем не делая никаких предположений относительно формы распределения, которому подчиняется исходная выборка. Профессионалам-статистикам достаточно его непрерывности. И удивительнейшим образом оказывается, что наилучшим распределением выборочного коэффициента корреляции r в общем случае является так называемое перестановочное распределение r , которое с очень хорошим приближением даже для малых N совпадает с t -распределением Стьюдента с $N-2$ степенями свободы для коэффициента корреляции r в «нормальной теории». Первые два момента распределений совпадают полностью (Кендалл, Стюарт, 1973, п. 31.19).

«Но если фактическое совпадение перестановочного распределения с распределением нормальной теории не является полной неожиданностью, то оно во всяком случае очень удобно и приятно, так как мы можем по-прежнему пользоваться таблицами нормальной теории (в данном случае t -распределения Стьюдента) для свободного от распределения критерия непараметрической гипотезы независимости» (Кендалл, Стюарт, 1973, п. 31.20).

Теперь становится понятным, почему и в случае двумерного нормального распределения (формула (3b)), и в случае пары бинарного и двумодального распределений (формула (3a)) t -распределение для выборочного коэффициента корреляции r (и одновременно для t -критерия разности двух средних) вычисляется по одной и той же формуле. А оно и должно всегда по ней вычисляться, если мы не знаем настоящего генерального распределения. Так говорит теория. Разница только в том, что если мы исходим из гипотезы нормального распределения, то мы должны ее

проверять и только в случае удачи использовать t -критерий Стьюдента, а иначе уходить на ранги и критерии типа Манна–Уитни. Если же мы исходим из гипотезы неизвестного непрерывного распределения, то ничего проверять не должны, а должны использовать тот же самый t -критерий Стьюдента, причем в любой удобной для нас форме. Исследователь-практик вправе сам выбирать, из какой гипотезы ему исходить.

Бутстреп для расчета статистической значимости t -критерия Стьюдента разности средних двух выборок

А можно ли проверить теорию? То есть можно ли что-то сказать относительно статистической значимости, например, различия средних, имея только сами данные и совсем не выдвигая никаких теоретических гипотез насчет того, как именно они должны себя вести и какому распределению обязаны подчиняться? У нас же теперь компьютеры есть!

Американский статистик Б. Эфрон нашел такой путь и назвал его бутстрепом (Efron, 1979). Он предложил размножить исходную выборку прямо в компьютере. Пусть она состоит из N элементов. Новую выборку получим следующим образом. С помощью датчика случайных чисел выберем с равными вероятностями любой элемент исходной выборки и включим его копию в новую таблицу. Повторим процесс N раз. Бутстреп-копия исходной выборки сформирована.

Повторим весь процесс много раз (K) и получим K бутстреп-копий исходной выборки. Поскольку все элементы в них равновероятно выбраны из исходной выборки, они подчиняются ее эмпирическому распределению. А с ростом K оно сходится к генеральному. То, что оно нам неизвестно, не имеет значения. Главное – что оно одно и то же для исходной выборки и для всех ее бутстреп-копий. По сути, мы получили некую искусственную модель генерального распределения, которому заведомо подчиняется наша выборка, поскольку именно ее эмпирическое распределение и взято за основу (Rochowicz, 2010; Hesterberg, 2015; Rousselet et al., 2023). Поскольку это все же «не настоящее» генеральное распределение, то у него есть и свои преимущества, и неизбежные недостатки. Начнем с преимуществ: 1. Проверка предположений о нормальности и равенстве дисперсий не требуется. И вообще никаких предположений о распределениях выборок не требуется. 2. Выводы делаются на более солидном, хотя и насчитанном искусственно материале. Недостатки: 1. Необходимы более мощные компьютеры, специализированное программное обеспечение и много машинного времени. 2. Не для всех задач его можно применять. Например, бутстреп в принципе не позволяет оценить среднее «настоящего» генерального распределения, а следовательно, и его доверительные интервалы. Чем больше бутстреп-копий выборки мы сделаем, тем ближе среднее их средних приблизится к среднему самой выборки, а не к среднему генерального распределения.

В случае средних двух выборок бутстреп можно применять по-разному, хотя нулевая гипотеза всегда одинакова: средние обеих выборок равны. Прямой способ изложен в (Hesterberg, 2015). Вычисляем для выборок стандартный t -критерий разности средних, для которого нам и нужно

получить p -value. Центрируем обе выборки, каждую своим средним. Теперь для них нулевая гипотеза заведомо выполняется, средние значения каждой выборки равны нулю. Остальные важные параметры каждой выборки (объем, дисперсия, асимметрия, эксцесс) при центрировании не меняются. Делаем для каждой центрированной выборки бутстреп-копию. Естественно, все параметры обеих выборок по случайным причинам отклоняются от исходных. Вычисляем t -критерий разности между ними. Повторяем много раз (10^3 , 10^4 , 10^5 – чем больше, тем лучше). Получаем бутстреп-распределение t -критерия при гарантированном выполнении нулевой гипотезы. Заметим, что это бутстреп-распределение совсем не обязано быть табличным t -распределением Стьюдента. Оно может быть сильно искажено особенностями обеих выборок – неравенством дисперсий, объемов, асимметрией и т. д. Но именно это нам и надо – учесть особенности самих выборок. А для этого достаточно посмотреть, какова доля насчитанных t -критериев в бутстреп-распределении, которые превысили (по модулю) t -критерий для исходных нецентрированных выборок. Это и будет бутстреп-оценкой p -value. Никаких дополнительных предположений и проверок не потребовалось.

Список литературы / References

- Болшев Л.Н., Смирнов Н.В. Таблицы математической статистики. М.: Наука, 1983
[Bolshev L.N., Smirnov N.V. Tables of Mathematical Statistics. Moscow: Nauka Publ., 1983 (in Russian)]
- Кендалл М., Стюарт А. Статистические выводы и связи. Т. 2. М.: Наука, 1973
[Kendall M., Stewart A. The Advanced Theory of Statistics. Vol. 2. Moscow: Nauka Publ., 1973 (in Russian)]
- Крамер Г. Математические методы статистики. М.: Мир, 1975
[Cramer H. Mathematical Methods of Statistics. Moscow: Mir Publ., 1975 (in Russian)]
- Лемешко Б.Ю., Лемешко С.Б. Об устойчивости и мощности критериев проверки однородности средних. *Измерительная техника*. 2008;9:23-28
[Lemeshko B.Y., Lemeshko S.B. Power and robustness of criteria used to verify the homogeneity of means. *Meas Tech*. 2008;51(9):950-959. doi 10.1007/s11018-008-9157-3]
- Лойко В.И., Луценко Е.В., Орлов А.И. Высокие статистические технологии и системно-когнитивное моделирование в экологии. Краснодар: КубГАУ, 2019
[Loiko V.I., Lutsenko E.V., Orlov A.I. High Statistical Technologies and System-cognitive Modeling in Ecology. Krasnodar: KubSAU Publ., 2019 (in Russian)]
- Орлов А.И. О методах проверки однородности двух независимых выборок. *Заводская лаборатория. Диагностика материалов*. 2020;86(3):67-76. doi 10.26896/1028-6861-2020-86-3-67-76
[Orlov A.I. On methods of testing the homogeneity of two independent samples. *Zavodskaya Laboratoriya. Diagnostika Materialov = Industrial Laboratory. Diagnostics of Materials*. 2020;86(3):67-76. doi 10.26896/1028-6861-2020-86-3-67-76 (in Russian)]
- Фишер Р.А. Статистические методы для исследователей. М.: Госстатиздат, 1958
[Fisher R.A. Statistical Methods for Research Workers. Moscow: Gosstatizdat Publ., 1958 (in Russian)]
- Benjamin D.J., Berger J.O., Johannesson M., Nosek B., Wagenmakers E.J., Berk R., Bollen K.A., ... Wolpert R., Xie Y., Young C., Zinman J., Johnson V.E. Redefine statistical significance. *Nat Hum Behav*. 2018;2(1):6-10. doi 10.1038/s41562-017-0189-z
- Boukerche A., Zheng L., Alfandi O. Outlier detection: methods, models, and classification. *ACM Computing Surveys (CSUR)*. 2020;53(3):1-37. doi 10.1145/3381028
- Conover W.J. The rank transformation – an easy and intuitive way to connect many nonparametric methods to their parametric counter-

- parts for seamless teaching introductory statistics courses. *Wiley Interdiscip Rev Comput Stat*. 2012;4(5):432-438. doi 10.1002/wics.1216
- Davenport J.M., Webster J.T. The Behrens–Fisher problem, an old solution revisited. *Metrika*. 1975;22(1):47-54
- Efron B. Computers and the theory of statistics: thinking the unthinkable. *SIAM Rev*. 1979;21(4):460-480. doi 10.1137/1021092
- Fisher R.A. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*. 1915;10(4):507-521
- Fisher R.A. Applications of “Student’s” distribution. *Metron*. 1925a;5:90-104
- Fisher R.A. *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd, 1925b
- Hammer Ø, Harper D.A.T., Ryan P.D. PAST: PAleontological STATistics software package for education and data analysis. *Palaeontol Electronica*. 2001;4(1):9
- Hazel L.N. The genetic basis for constructing selection indexes. *Genetics*. 1943;28(6):476-490
- Hesterberg T.C. What teachers should know about the bootstrap: resampling in the undergraduate statistics curriculum. *Am Stat*. 2015;69(4):371-386. doi 10.1080/00031305.2015.1089789
- Kennedy-Shaffer L. Before $p < 0.05$ to beyond $p < 0.05$: using history to contextualize p -values and significance testing. *Am Stat*. 2019;73(sup1):82-90. doi 10.1080/00031305.2018.1537891
- Lehmann E.L. “Student” and small-sample theory. *Stat Sci*. 1999;14(4):418-426
- Lumley T., Diehr P., Emerson S., Chen L. The importance of the normality assumption in large public health data sets. *Annu Rev Public Health*. 2002;23(1):151-169. doi 10.1146/annurev.publhealth.23.100901.140546
- Rochowicz J.A., Jr. Bootstrapping analysis, inferential statistics and EXCEL. *Spreadsheets in Education (SiE)*. 2010;4(3). <http://epublications.bond.edu.au/ejsie/vol4/iss3/4>
- Rousselet G., Pernet C.R., Wilcox R.R. An introduction to the bootstrap: a versatile method to make inferences by using data-driven simulations. *Meta-Psychology*. 2023;7:2058. doi 10.15626/MP.2019.2058
- Smith H.F. The problem of comparing the results of two experiments with unequal errors. *J Council Sci Industrial Res*. 1936a;9:211-212
- Smith H.F. A discriminant function for plant selection. *Annals of Eugenics*. 1936b;7(3):240-250.
- Student. The probable error of a mean. *Biometrika*. 1908;6:1-25. doi 10.2307/2331554
- Student. Tables for estimating the probability that the mean of a unique sample of observations lies between $-\infty$ and any given distance of the mean of the population from which the sample is drawn. *Biometrika*. 1917;XV:414-417
- Student. Statistics in biological research. *Nature*. 1929;124:93. doi 10.1038/124093b0
- Student’s t -test (t -критерий Стьюдента) https://en.wikipedia.org/wiki/Student%27s_t-test [обновлено 10 января 2023; процитировано 30 июня 2024. доступно: <https://руни.рф>]
- Wang H., Bah M.J., Hammad M. Progress in outlier detection techniques: a survey. *IEEE Access*. 2019;7:107964-108000. doi 10.1109/ACCESS.2019.2932769
- Wasserstein R.L., Schirm A.L., Lazar N.A. Moving to a world beyond “ $p < 0.05$ ”. *The American Statistician*. 2019;73(sup.1):1-19. <https://www.jstor.org/stable/48783683>
- Welch B.L. The significance of the difference between two means when the population variances are unequal. *Biometrika*. 1938;29(3/4):350-362

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Поступила в редакцию 01.09.2024. После доработки 13.11.2024. Принята к публикации 09.12.2024.

Сетевое издание «Письма в Вавиловский журнал генетики и селекции» – реестровая запись СМИ Эл № ФС77-75536, зарегистрировано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций 08 мая 2019 г.

Основано в 2015 году (до 2019 года выходило под названием «Письма в Вавиловский журнал»). На страницах издания публикуются результаты экспериментальных, методических и теоретических исследований, аналитические обзоры по всем разделам генетики и селекции, а также по смежным областям биологических и сельскохозяйственных наук; материалы и документы по истории генетики и селекции; описания сортов растений и пород животных; рецензии; письма, адресованные редактору; персоналии и мемориальные статьи; хроника и информация из региональных отделений Вавиловского общества генетиков и селекционеров.

Цель издания – донести новейшие результаты фундаментальных и прикладных исследований в области генетики растений, животных, человека, микроорганизмов, описание новых методов и селекционных достижений до наибольшего числа ученых, включая специалистов из смежных областей науки и техники, а также до преподавателей вузов, читающих курсы лекций по генетике и селекции.

Сетевое издание «Письма в Вавиловский журнал генетики и селекции» с 15.06.2023 включено в [Перечень рецензируемых научных изданий](#), в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук, по специальностям и отраслям науки:

- 1.5.7. Генетика (биологические науки)
- 1.5.22. Клеточная биология (биологические науки)
- 4.1.2. Селекция, семеноводство и биотехнология растений (биологические науки)
- 4.1.2. Селекция, семеноводство и биотехнология растений (сельскохозяйственные науки)

Индексируется в РИНЦ, включено в DOAJ.

Прием статей осуществляется через электронную почту редакции: pismavavilov@bionet.nsc.ru

Адрес издания в сети интернет: <https://pismavavilov.ru/>

При перепечатке материалов ссылка на издание обязательна.

Учредитель и издатель: Федеральное государственное бюджетное научное учреждение «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук» (ИЦиГ СО РАН)
Почтовый адрес учредителя и издателя: проспект Академика Лаврентьева, 10, Новосибирск, 630090
Почтовый адрес редакции: проспект Академика Лаврентьева, 10, Новосибирск, 630090
Телефон редакции: (383) 363 4963, доб. 5316
✉ Электронный адрес редакции: pismavavilov@bionet.nsc.ru

Выпуск подготовлен информационно-издательским отделом ИЦиГ СО РАН.

Дата публикации: 26.03.2025

