

Научный рецензируемый журнал

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

Основан в 1997 г.

Периодичность 8 выпусков в год

Учредители

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

Межрегиональная общественная организация Вавиловское общество генетиков и селекционеров

Сибирское отделение Российской академии наук

Главный редактор

В.К. Шумный – академик РАН, д-р биол. наук, профессор (Россия)

Заместители главного редактора

Н.А. Колчанов – академик РАН, д-р биол. наук, профессор (Россия)

Н.Б. Рубцов – д-р биол. наук, профессор (Россия)

Е.К. Хлесткина – д-р биол. наук, профессор (Россия)

Ответственный секретарь

Г.В. Орлова – канд. биол. наук (Россия)

Редакционный совет

В.С. Баранов – чл.-кор. РАН, д-р мед. наук (Россия)
Л.А. Беспалова – академик РАН, д-р с.-х. наук (Россия)
А. Бёрнер – д-р наук (Германия)
В.М. Говорун – академик РАН, д-р биол. наук (Россия)
И. Гроссе – д-р наук, проф. (Германия)
Г.Л. Дианов – д-р биол. наук, проф. (Великобритания)
Ю.Е. Дуброва – д-р биол. наук, проф. (Великобритания)
И.К. Захаров – д-р биол. наук, проф. (Россия)
И.А. Захаров-Гезехус – чл.-кор. РАН, д-р биол. наук (Россия)
С.Г. Инге-Вечтомов – академик РАН, д-р биол. наук (Россия)
И.Е. Керкис – д-р наук (Бразилия)
А.В. Кильчевский – чл.-кор. НАНБ, д-р биол. наук (Беларусь)
С.В. Костров – чл.-кор. РАН, д-р хим. наук (Россия)
Ж. Ле Гуи – д-р наук (Франция)
Б. Люгтенберг – д-р наук, проф. (Нидерланды)
В.И. Молодин – академик РАН, д-р ист. наук (Россия)
В.П. Пузырев – академик РАН, д-р мед. наук (Россия)
А.Ю. Ржецкий – канд. биол. наук, проф. (США)
И.Б. Рогозин – канд. биол. наук (США)
А.О. Рувинский – д-р биол. наук, проф. (Австралия)
К.Г. Скрябин – академик РАН, д-р биол. наук (Россия)
К.В. Славин – д-р наук, проф. (США)
И.А. Тихонович – академик РАН, д-р биол. наук (Россия)
Л.В. Хотылева – академик НАНБ, д-р биол. наук (Беларусь)
Э.К. Хуснутдинова – д-р биол. наук, проф. (Россия)
М.Ф. Чернов – д-р мед. наук (Япония)
С.В. Шестаков – академик РАН, д-р биол. наук (Россия)
Н.К. Янковский – академик РАН, д-р биол. наук (Россия)

Редакционная коллегия

Т.Г. Амстиславская – д-р биол. наук, доцент (Россия)
Е.Е. Андронов – канд. биол. наук (Россия)
Ю.С. Аульченко – д-р биол. наук (Россия)
Д.А. Афонников – канд. биол. наук, доцент (Россия)
Л.И. Афтанас – академик РАН, д-р мед. наук (Россия)
Е.В. Березиков – канд. биол. наук, проф. (Россия, Нидерланды)
С.А. Боринская – д-р биол. наук (Россия)
П.М. Бородин – д-р биол. наук, проф. (Россия)
М.И. Воевода – академик РАН, д-р мед. наук (Россия)
Т.А. Гавриленко – д-р биол. наук, доцент (Россия)
В.Н. Даниленко – д-р биол. наук, проф. (Россия)
С.А. Демаков – д-р биол. наук (Россия)
Е.А. Долгих – канд. биол. наук (Россия)
Н.Н. Дыгало – чл.-кор. РАН, д-р биол. наук (Россия)
С.Л. Киселев – д-р биол. наук, проф. (Россия)
В.А. Козлов – академик РАН, д-р мед. наук (Россия)
Ю.М. Константинов – д-р биол. наук, проф. (Россия)
А.В. Кочетов – чл.-кор. РАН, д-р биол. наук (Россия)
О. Кребс – д-р биол. наук, проф. (Германия)
И.Н. Лаврик – канд. хим. наук (Германия)
Л.А. Лутова – д-р биол. наук, проф. (Россия)
В.Ю. Макеев – чл.-кор. РАН, д-р физ.-мат. наук (Россия)
М.П. Мошкин – д-р биол. наук, проф. (Россия)
Н.А. Проворов – д-р биол. наук (Россия)
Д.В. Пышный – чл.-кор. РАН, д-р хим. наук (Россия)
А.В. Ратушный – канд. биол. наук (США)
Е.А. Салина – д-р биол. наук, проф. (Россия)
М.Г. Самсонова – д-р биол. наук (Россия)
В.А. Степанов – чл.-кор. РАН, д-р биол. наук (Россия)

Scientific Peer Reviewed Journal

VAVILOV JOURNAL OF GENETICS AND BREEDING

Vavilovskii Zhurnal Genetiki i Seleksii

Founded in 1997

Published 8 times annually

Founders

Federal State Budget Scientific Institution "The Federal Research Center Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences"

The Vavilov Society of Geneticists and Breeders

Siberian Branch of the Russian Academy of Sciences

Editor-in-Chief

V.K. Shumny, Full Member of the Russian Academy of Sciences, Dr. Sci. (Biology), Russia

Deputy Editor-in-Chief

N.A. Kolchanov, Full Member of the Russian Academy of Sciences, Dr. Sci. (Biology), Russia

N.B. Rubtsov, Professor, Dr. Sci. (Biology), Russia

E.K. Khlestkina, Professor, Dr. Sci. (Biology), Russia

Executive Secretary

G.V. Orlova, Cand. Sci. (Biology), Russia

Editorial council

V.S. Baranov, Corr. Member of the RAS, Dr. Sci. (Medicine), Russia

L.A. Bespalova, Full Member of the RAS, Dr. Sci. (Agricul.), Russia

A. Börner, Dr. Sci., Germany

M.F. Chernov, Dr. Sci. (Medicine), Japan

G.L. Dianov, Professor, Dr. Sci. (Biology), Great Britain

Yu.E. Dubrova, Professor, Dr. Sci. (Biology), Great Britain

V.M. Govorun, Full Member of the RAS, Dr. Sci. (Biology), Russia

J. Le Gouis, Dr. Sci., France

I. Grosse, Professor, Dr. Sci., Germany

S.G. Inge-Vechtomov, Full Member of the RAS, Dr. Sci. (Biology), Russia

I.E. Kerkis, Dr. Sci., Brazil

L.V. Khotyleva, Full Member of the NAS of Belarus, Dr. Sci. (Biology), Belarus

E.K. Khusnutdinova, Professor, Dr. Sci. (Biology), Russia

A.V. Kilchevsky, Corr. Member of the NAS of Belarus, Dr. Sci. (Biology), Belarus

S.V. Kostrov, Corr. Member of the RAS, Dr. Sci. (Chemistry), Russia

B. Lugtenberg, Professor, Dr. Sci., Netherlands

V.I. Molodin, Full Member of the RAS, Dr. Sci. (History), Russia

V.P. Puzyrev, Full Member of the RAS, Dr. Sci. (Medicine), Russia

I.B. Rogozin, Cand. Sci. (Biology), United States

A.O. Ruvinsky, Professor, Dr. Sci. (Biology), Australia

A.Yu. Rzhetsky, Professor, Cand. Sci. (Biology), United States

S.V. Shestakov, Full Member of the RAS, Dr. Sci. (Biology), Russia

K.G. Skryabin, Full Member of the RAS, Dr. Sci. (Biology), Russia

K.V. Slavin, Professor, Dr. Sci., United States

I.A. Tikhonovich, Full Member of the RAS, Dr. Sci. (Biology), Russia

N.K. Yankovsky, Full Member of the RAS, Dr. Sci. (Biology), Russia

I.K. Zakharov, Professor, Dr. Sci. (Biology), Russia

I.A. Zakharov-Gezekhus, Corr. Member of the RAS, Dr. Sci. (Biology), Russia

Editorial board

D.A. Afonnikov, Associate Professor, Cand. Sci. (Biology), Russia

L.I. Aftanas, Full Member of the RAS, Dr. Sci. (Medicine), Russia

T.G. Amstislavskaya, Associate Professor, Dr. Sci. (Biology), Russia

E.E. Andronov, Cand. Sci. (Biology), Russia

Yu.S. Aulchenko, Dr. Sci. (Biology), Russia

E.V. Berezikov, Professor, Cand. Sci. (Biology), Russia, Netherlands

S.A. Borinskaya, Dr. Sci. (Biology), Russia

P.M. Borodin, Professor, Dr. Sci. (Biology), Russia

V.N. Danilenko, Professor, Dr. Sci. (Biology), Russia

S.A. Demakov, Dr. Sci. (Biology), Russia

E.A. Dolgikh, Cand. Sci. (Biology), Russia

N.N. Dygalo, Corr. Member of the RAS, Dr. Sci. (Biology), Russia

T.A. Gavrilenko, Associate Professor, Dr. Sci. (Biology), Russia

S.L. Kiselev, Professor, Dr. Sci. (Biology), Russia

A.V. Kochetov, Corr. Member of the RAS, Dr. Sci. (Biology), Russia

Yu.M. Konstantinov, Professor, Dr. Sci. (Biology), Russia

V.A. Kozlov, Full Member of the RAS, Dr. Sci. (Medicine), Russia

O. Krebs, Professor, Dr. Sci. (Biology), Germany

I.N. Lavrik, Cand. Sci. (Chemistry), Germany

L.A. Lutova, Professor, Dr. Sci. (Biology), Russia

V.Yu. Makeev, Corr. Member of the RAS, Dr. Sci. (Physics and Mathem.), Russia

M.P. Moshkin, Professor, Dr. Sci. (Biology), Russia

N.A. Provorov, Dr. Sci. (Biology), Russia

D.V. Pyshnyi, Corr. Member of the RAS, Dr. Sci. (Chemistry), Russia

A.V. Ratushny, Cand. Sci. (Biology), United States

E.A. Salina, Professor, Dr. Sci. (Biology), Russia

M.G. Samsonova, Dr. Sci. (Biology), Russia

V.A. Stepanov, Corr. Member of the RAS, Dr. Sci. (Biology), Russia

M.I. Voevoda, Full Member of the RAS, Dr. Sci. (Medicine), Russia

877

ОТ РЕДАКТОРА

Структура и взаимодействие макромолекул

878

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Кватернионное моделирование траектории спирали для анализа формы молекулы ДНК
А.Ф. Мутерко

887

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Взаимодействия белок-ДНК: статистический анализ межатомных контактов в большой и малой бороздках
А.А. Анашкина, Е.Н. Кузнецов, А.В. Батяновский, Л.А. Урошлев, В.Г. Туманян, Н.Г. Есипова

Системная биология человека и животных

895

ОБЗОР

Интернет-доступные информационные ресурсы по геномным сетям, включающие данные по человеку и животным
Е.В. Игнатьева, Д.А. Афонников, Н.А. Колчанов

903

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Анализ циркадного ритма биологических процессов в печени и почках мыши
Н.Л. Подколотный, Н.Н. Твердохлеб, О.А. Подколотная

911

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Роль генов апоптоза в контроле агрессивного поведения, выявленная с помощью комбинированного анализа ассоциативных геномных сетей, экспрессионных и геномных данных по серым крысам с агрессивным поведением
А.О. Брагин, О.В. Сайк, И.В. Чадаева, П.С. Деменков, А.Л. Маркель, Ю.Л. Орлов, Е.И. Рогов, И.Н. Лаврик, В.А. Иванисенко

Системная биология и биомедицина

920

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Перспективные маркеры СІМР+ опухолей толстой кишки, выявленные на основе анализа данных ресурса TCGA
Г.С. Краснов, А.Д. Бениаминов, Р.А. Тычко, Г.А. Пузанов, Р.О. Новаковский, А.В. Кудрявцева, А.А. Дмитриев

925

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Особенности экспрессионного профиля больных злокачественной меланомой в зависимости от ответа на иммунотерапию
Е.Н. Лукьянова, М.С. Федорова, Е.А. Пудова, Т.В. Наседкина, Е.В. Степанова, К.М. Ньюшко, А.Ю. Попов, Н.В. Коробан, А.А. Дмитриев, М.В. Киселева, А.В. Липатова, А.С. Заседателев, А.В. Кудрявцева

932

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Ингибирование гексокиназы 2 приводит к снижению экспрессии ферментов гликолиза PFKF, BPGM и GPI в клеточной линии RKO
М.С. Федорова, И.Ю. Карпова, А.В. Липатова, Е.А. Пудова, З.Г. Гуватова, Д.В. Кочетков, А.В. Чайка, Б.Я. Алексеев, М.В. Киселева, А.Д. Каприн, А.В. Кудрявцева, А.В. Снежкина

937

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Отсроченные эффекты хирургического воздействия, перенесенного матерью в период ранней беременности, на артериальное давление и поведение потомков у крыс линии OXYS
И.Н. Рожкова, Т.Н. Игонина, Д.С. Рагаева, О.М. Петрова, Е.Ю. Брусенцев, В.А. Напримеров, С.Я. Амстиславский

Экологическая и популяционная генетика

943

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Паразитические организмы родов *Nosema*, *Crithidia* и *Lotmaria* в популяциях пчел и шмелей: исследование в Индии
В.Ю. Вавилова, И.Д. Конопацкая, С.Л. Лузянин, М. Войцеховский, А.Г. Блинов

952

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Генетическая и биохимическая характеристика стафилококков, встречающихся в Новосибирске
Ю.Н. Козлова, Н.В. Фоменко, В.В. Морозова, И.В. Саранина, А.Ю. Тикунов, Д.А. Ганичев, А.Г. Самохин, В.В. Павлов, О.М. Рожнова, И.А. Бондарь, Е.В. Зенкова, В.В. Нимаев, В.В. Климонтов, Н.В. Тикунова

- 959 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Филогенетический анализ
микробного мата в горячем
источнике Гарга (Байкальская
рифтовая зона) и разнообразие
природных пептидаз
*Е.В. Лаврентьева, А.А. Раднагуруева,
Т.Г. Банзаракцаева, С.М. Базаров, Д.Д. Бархутова,
И.Д. Ульзетуева, М.К. Чернявский, М.Р. Кабилов,
В.В. Хахинов*
- 964 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Полиморфизм гена трегалазы (*TREH*)
у коренного населения Сибири
Б.А. Малярчук, М.В. Деренко
- Актуальные технологии генетики
и клеточной биологии**
- 969 **ОБЗОР**
Методы геномного редактирования
дрожжей: история и современное
состояние
*А.С. Розанов, В.Н. Шляхтун, Л.А. Текутьева,
О.М. Сон, С.В. Сизова, С.Е. Пельтек*
- 979 **ОБЗОР**
Растительные системы экспрессии
в качестве продуцентов
рекомбинантных фармацевтически
ценных белков
Е.В. Дейнеко, А.А. Загорская
- 986 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Создание кандидатной вакцины
против клещевого энцефалита на
основе гибридного рекомбинантного
flagG-protE-белка
*П.А. Белавин, Д.А. Куньк, Е.В. Протопопова,
В.Б. Локтев, Е.В. Дейнеко*
- 993 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Создание стабильного штамма-
продуцента полноразмерного
антитела человека на примере
антитела против вируса экстремелии
*А.Л. Матвеев, Я.А. Хлусевич, И.К. Байков,
И.В. Бабкин, Е.П. Гончарова, В.В. Морозова,
Н.В. Тикунова*
- 1001 Алфавитный указатель авторов
статей, опубликованных в журнале
в 2017 г.

877

FROM THE EDITOR

Structure and interaction of macromolecules

878

ORIGINAL ARTICLE

Quaternion modeling of the helical path for analysis of the shape of the DNA molecule

A.F. Muterko

887

ORIGINAL ARTICLE

Protein-DNA interactions: statistical analysis of interatomic contacts in major and minor grooves

A.A. Anashkina, E.N. Kuznetsov, A.V. Batianovskii, L.A. Uroshlev, V.G. Tumanyan, N.G. Esipova

Human and animal systems biology

895

REVIEW

Online resources on gene networks containing human and animal data

E.V. Ignatieva, D.A. Afonnikov, N.A. Kolchanov

903

ORIGINAL ARTICLE

Analysis of the circadian rhythm of biological processes in mouse liver and kidney

N.L. Podkolodnyy, N.N. Tverdokhleby, O.A. Podkolodnaya

911

ORIGINAL ARTICLE

Role of apoptosis genes in aggression revealed using combined analysis of ANDSystem gene networks, expression and genomic data in grey rats with aggressive behavior

A.O. Bragin, O.V. Saik, I.V. Chadaeva, P.S. Demenkov, A.L. Markel, Yu.L. Orlov, E.I. Rogayev, I.N. Lavrik, V.A. Ivanisenko

Systems biology and biomedicine

920

ORIGINAL ARTICLE

Promising markers of CIMP+ colon tumors identified on the basis of TCGA data analysis

G.S. Krasnov, A.D. Beniaminov, R.A. Tychko, G.A. Puzanov, R.O. Novakovskiy, A.V. Kudryavtseva, A.A. Dmitriev

925

ORIGINAL ARTICLE

Differences in expression profiles in malignant melanoma patients according to immunotherapy response

E.N. Lukyanova, M.S. Fedorova, E.A. Pudova, T.V. Nasedkina, E.V. Stepanova, K.M. Nyushko, A.Y. Popov, N.V. Koroban, A.A. Dmitriev, M.V. Kiseleva, A.V. Lipatova, A.S. Zasedatelev, A.V. Kudryavtseva

932

ORIGINAL ARTICLE

Knockdown of hexokinase 2 results in a decreased expression level of the glycolytic enzymes PFKF, BPGM, and GPI in RKO cell line

M.S. Fedorova, I.Y. Karpova, A.V. Lipatova, E.A. Pudova, Z.G. Guvatova, D.V. Kochetkov, A.V. Chaika, B.Y. Alekseev, M.V. Kiseleva, A.D. Kaprin, A.V. Kudryavtseva, A.V. Snezhkina

937

ORIGINAL ARTICLE

Long-term effects of maternal exposure to surgical stress at the earliest stage of pregnancy on blood pressure and behavior in offspring of OXYS rats

I.N. Rozhkova, T.N. Igonina, D.S. Ragaeva, O.M. Petrova, E.Yu. Brusentsev, V.A. Naprimerov, S.Ya. Amstislavsky

Ecological and population genetics

943

ORIGINAL ARTICLE

Parasites of the genus *Nosema*, *Crithidia* and *Lotmaria* in the honeybee and bumblebee populations: a case study in India

V.Y. Vavilova, I. Konopatskaia, S.L. Luzyanin, M. Woyciechowski, A.G. Blinov

952

ORIGINAL ARTICLE

Genetic and biochemical characterization of staphylococci occurring in Novosibirsk, Russia

Y.N. Kozlova, N.V. Fomenko, V.V. Morozova, I.V. Saranina, A.Yu. Tikunov, D.A. Ganichev, A.G. Samokhin, V.V. Pavlov, O.M. Rozhnova, I.A. Bondar, E.V. Zenkova, V.V. Nimaev, V.V. Klimontov, N.V. Tikunova

- 959 **ORIGINAL ARTICLE**
Phylogenetic analysis of the microbial mat in the hot spring Garga (Baikal rift zone) and the diversity of natural peptidases
E.V. Lavrentyeva, A.A. Radnagurueva, T.G. Banzaraktsaeva, S.M. Bazarov, D.D. Barkhutova, I.D. Ulzetueva, M.K. Chernyavsky, M.R. Kabilov, V.V. Khakhinov
- 964 **ORIGINAL ARTICLE**
Polymorphism of the trehalase gene (*TREH*) in native populations of Siberia
B.A. Malyarchuk, M.V. Derenko
- Mainstream technologies in genetics and cell biology**
- 969 **REVIEW**
Methods of yeast genome editing
A.S. Rozanov, V.N. Shlyahatun, L.A. Tekutieva, O.M. Son, S.V. Sizova, S.E. Peltek
- 979 **REVIEW**
Plant expression systems for production of recombinant pharmaceutically important proteins
E.V. Deineko, A.A. Zagorskaya
- 986 **ORIGINAL ARTICLE**
Candidate vaccine construction against tick-born encephalitis based on hybrid recombinant flagG-protE-protein
P.A. Belavin, D.A. Kunyk, E.V. Protopopova, V.B. Loktev, E.V. Deineko
- 993 **ORIGINAL ARTICLE**
Development of a stable eukaryotic strain producing fully human monoclonal antibody on the basis of the human antibody against ectromelia virus
A.L. Matveev, Ya.A. Khlusevich, I.K. Baykov, I.V. Babkin, E.P. Goncharova, V.V. Morozova, N.V. Tikunova
- 1001 Alphabetical author index for the list of papers published in the journal in 2017

Предлагаемый вниманию читателя выпуск журнала посвящен ряду актуальных направлений биоинформатики, системной биологии, биомедицины, экологической и популяционной генетики, технологий генетики и клеточной биологии.

Раздел «Структура и взаимодействие макромолекул» состоит из двух статей. В первой из них описаны алгоритмы моделирования траектории молекулы ДНК с использованием кватернионов – системы гиперкомплексных чисел, удобной для представления операций вращения и смещения объектов в 3D пространстве. В сравнении с матрицами они обладают большей вычислительной устойчивостью и эффективностью. Во второй статье рассмотрено взаимодействие белок–ДНК на основе статистического анализа межатомных контактов в большой и малой бороздках ДНК.

Обзор информационных ресурсов по генным сетям человека и животных, которые важны для реконструкции моделей генных сетей, открывает рубрику «Системная биология человека и животных». В следующей статье представлены результаты оригинального исследования тканеспецифичности циркадных фазовых характеристик биологических процессов, полученные на основе сравнительного анализа экспериментальных данных по суточной динамике уровня трансляции генов мыши. Для двенадцати временных точек суток в каждой ткани (печень, почки) выделены группы генов, находящиеся в фазе с повышенным уровнем трансляции, для которых проведен анализ обогащения терминами GO. Благодаря этому выявлены циркадные паттерны биологических процессов и проведено сравнение временной активности биологических процессов в течение суток в различных органах. Последняя статья раздела посвящена взаимосвязи нейронального апоптоза и агрессивного поведения крыс. Комплексный анализ экспрессионных и геномных данных у серых крыс с агрессивным и дружелюбным поведением, а также ассоциативных генных сетей позволил выявить набор генов,

потенциально участвующих в регуляции нейронального апоптоза посредством молекулярно-генетических взаимодействий с генами/белками, напрямую вовлеченными в нейрональный апоптоз.

Серия статей по выявлению перспективных SNP-маркеров канцерогенеза, анализу экспрессионных профилей больных злокачественной меланомой в зависимости от лечения, исследованию влияния ингибирования гексокиназы 2 на фосфорилирование глюкозы, влиянию стресса на ранней стадии беременности у крыс линии OXYS на их потомство представлена в разделе «Системная биология и биомедицина». Эти работы связаны с развитием концепции персонализированной медицины, которая во многом опирается на современные методы системной биологии и молекулярной генетики.

Раздел «Экологическая и популяционная генетика» открывает статья, в которой суммированы результаты исследований распространенности паразитов в различных популяциях пчел и шмелей. Другое интересное исследование, посвященное анализу генетических и биохимических характеристик стафилококков, распространенных в Новосибирске, имеет не только теоретическое, но и прикладное значение для принятия решений в лечении стафилококковых инфекций. Авторы следующей работы изучают разнообразие природных пептидаз на основе филогенетического анализа микробного сообщества, живущего в экстремальных температурных условиях, что может получить применение в современных биотехнологиях. В последней статье раздела представлены исследования полиморфизмов в популяциях коренного населения Сибири.

И наконец, завершает выпуск журнала рубрика, посвященная актуальным технологиям генетики и клеточной биологии. В двух обзорных статьях рассмотрен вопрос применения и перспектив дальнейшего развития методов геномного редактирования дрожжей, а также описаны методы генетической инженерии растений, в частности использование растительных систем экспрессии в качестве продуцентов рекомбинантных фармацевтически ценных белков. Оригинальные исследования знакомят с технологией создания стабильного штамма-продуцента полноразмерного антитела человека на примере антитела против вируса экстремелии и с технологией создания кандидатной вакцины против клещевого энцефалита на основе гибридного рекомбинантного flagG-protE-белка.

*Научные редакторы:
академик Н.А. Колчанов,
заведующий ЦКП «Биоинформатика» Н.Л. Подколотный,
заместитель директора по научной работе С.Е. Пельтек*

Quaternion modeling of the helical path for analysis of the shape of the DNA molecule

A.F. Muterko

Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

The three-dimensional shape of a DNA molecule is a key property influencing its functional specificity and the nature of its molecular interactions. The characteristic shape into which a DNA molecule folds under certain conditions is a manifestation of its micromechanical and structural features, which are sequence-dependent. DNA shape-related properties can therefore be determined in a predictable manner. A number of models have been designed to describe intrinsic DNA curvature, incorporating a set of helical parameters which can be applied to operative three-dimensional reconstruction of the DNA structures. Alternatively, desired base pair parameters can be computed based on publicly available information about atomic DNA structures. Further, taking the base pairs as rigid bodies, their relative location in space can be estimated based on these parameters. Matrices are a common method to implement any rigid body transformations and are widely used in the modeling of DNA structures. Quaternions are the more straightforward and robust alternative for matrices. Unit quaternions can represent only a rotation, whereas dual quaternions combine rotation and translation into a single state. In the present guide, the algebra of unit and dual quaternions is applied for the first time to modeling of the DNA helical path, based on conformational parameters of the base pair steps. Although dual quaternions are preferable for modeling of DNA structure in detail, the use of unit quaternions is sufficient to predict the DNA trajectory and all calculations of DNA shape features. In order to analyze DNA shape and chain statistics, and predict the micromechanical properties of DNA molecules based on coordinates of the helical path, the widely used as well as original algorithms for computing DNA curvature, radius of gyration, persistence length and phasing of DNA bends are described. Taken together, these algorithms will be useful both in the *in silico* analysis of relatively short DNA fragments as well as in topological mapping of whole genomes.

Key words: quaternion modeling; DNA structure; curvature; radius of gyration; persistence length; Fourier transform; phasing; bending.

HOW TO CITE THIS ARTICLE:

Muterko A.F. Quaternion modeling of the helical path for analysis of the shape of the DNA molecule. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2017;21(8):878-886. DOI 10.18699/VJ17.308

Received 12.09.2017
Accepted for publication 29.11.2017
© AUTHOR, 2017

e-mail: muterko@gmail.com

Кватернионное моделирование траектории спирали для анализа формы молекулы ДНК

А.Ф. Мутерко

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

Пространственная организация, форма молекулы ДНК являются ключевой характеристикой, определяющей ее функциональную специфичность и природу межмолекулярных взаимодействий. Специфическая форма, которую молекула ДНК принимает при определенных условиях, обусловлена ее микромеханическими и структурными особенностями, зависящими от последовательности нуклеотидов. Следовательно, отдельные характеристики формы ДНК могут быть прогнозированы. Предложен ряд моделей для описания внутренней кривизны ДНК, включающий набор геометрических параметров двойной спирали, применяемых при компьютерной реконструкции пространственных структур. С другой стороны, необходимые параметры пар оснований можно рассчитать исходя из общедоступной информации атомной структуры ДНК. Принимая пары оснований как твердые тела, их относительное расположение в пространстве можно оценить по полученным параметрам. Матрицы являются наиболее распространенным способом реализации преобразований твердого тела и широко используются в моделировании формы ДНК. Более простая и надежная альтернатива матрицам – кватернионы. Единичные кватернионы представляют только поворот, тогда как двойные кватернионы объединяют в себе и поворот, и смещение. В настоящем руководстве алгебра единичных и двойных кватернионов впервые применена для моделирования траектории молекулы ДНК исходя из конформационных параметров динуклеотидных шагов. Хотя использование двойных кватернионов оптимально для детального моделирования структуры, единичные кватернионы достаточны для прогнозирования траектории двойной спирали и последующих расчетов ее пространственных характеристик. Обсуждаются широко используемые, а также оригинальные алгоритмы вычисления кривизны, радиуса гирации, персистентной длины и фазирования статических изгибов для анализа формы молекулы, вычисления статистики полимерной цепи и прогнозирования микромеханических свойств на основе координат траектории ДНК-спирали. Приведенные алгоритмы будут полезны как в ходе *in silico* анализа относительно коротких фрагментов ДНК, так и в топологическом картировании полных геномов.

Ключевые слова: кватернионное моделирование; структура ДНК; кривизна; радиус гирации; персистентная длина; преобразование Фурье; фазирование; изгибаемость.

The shape of DNA is sequence-dependent and determined by the intrinsic curvature and flexibility of the molecule, which is driven by external forces. DNA bending and curving play a crucial role in many important biological processes, including recombination, replication and the excision of damaged nucleotides. They are also important for transcriptional regulation of numerous prokaryotic and eukaryotic genes, as well as DNA-protein recognition and interactions. By now it is definitively known that intrinsic DNA curvature is a major determinant of nucleosome organization and positioning, facilitating chromosome folding and DNA packaging in the nucleoid, promoting the appropriate mode of supercoiling and protecting the prokaryotic genome from phage integration (see References in the Supplementary material)¹.

Predictions of sequence-dependent mechanical properties of DNA are important for understanding many biological processes associated with the various DNA-protein interactions, including the phenomenon of “indirect readout”. The micromechanical characteristics of the DNA molecule can be described by a variety of models, including the remarkably effective worm-like chain (WLC) model for semiflexible polymers (Kratky, Porod, 1949). In this model, the bending stiffness of a molecule is described by its persistence length. However, the WLC model is valid only for isotropic, intrinsically straight, homogeneous polymers and cannot be applied in its original form to DNA molecules, which contain non-Gaussian deformations, such as the intrinsic bends (Schellman, Harvey, 1995; Rivetti et al., 1998).

The intrinsic curvature of DNA is mainly determined by the length and localization of adenine tracts (A-tract, A_nT_m , $n+m \geq 4$). The helical phasing of A-tracts and other bent DNA-related sequences for a long time formed the basis of the central concept of intrinsic DNA curvature. More than 30 years ago it was demonstrated how localization of local bends affects the global curvature of a DNA molecule (Wu, Crothers, 1984; Hagerman, 1985; Koo et al., 1986). In particular, it has been shown that macroscopic DNA curvature is strongly affected by the phasing of local bends. Since the bends produced by A-tracts have directional preference, their systematic alternation in phase with the helix screw add coherently and significantly increases the macroscopic DNA curvature, whilst for “straighter” DNA molecules the systematic bends are nearly exactly out of phase (Koo et al., 1986). The Discrete Fourier Transform (DFT) is used to detect the periodicity of any property along a sequence, through calculation of the Fourier transform power spectrum. However, this method is valid only for planar systematic bends, which are oriented in the same direction or, often, the bends, which are determined by repeats of a sequence motif. It is important therefore to predict the macroscopic DNA curvature based on the phase of local bends, regardless of their origin.

Since DNA conformation has been shown to be sequence dependent, this has led to the development of series of models for *in silico* prediction of DNA shape based on sequence-dependent parameters of base pair steps in di-, three- and tetranucleotide contexts. These base pair steps (or often wedge) parameters were chosen through many different methods, such as circularization, gel-mobility, DNase I digestion and

nucleosome positioning data, X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, conformational energy minimization and computer simulation (see References in Supplementary material). Nonetheless, there is no complete consensus on sequence-dependent parameters that are optimal for accurate prediction of DNA shape.

Base pair orientation in the DNA helical structures is described by a set of intra and inters parameters, each including of three rotational and three translational parameters. Internal parameters determine the conformation of bases in the local coordinate frame of a base pair, hence, they do not affect the trajectory of the DNA helical path. The set of inter base pair geometrical parameters consists of the three rotations by *tilt* (τ), *roll* (ρ) and *twist* (Ω), and three translations *shift* (Dx), *slide* (Dy) and *rise* (Dz) about X , Y and Z axes, respectively (Fig. 1), in the standard coordinate frame approved by the Cambridge convention (Dickerson, 1989). In this manner, the inter base pair parameters are used to construct low resolution (coarse-grained) three-dimensional DNA structures. Such translational parameters as *shift* and *slide* define the local displacement between adjacent base pairs ($\sqrt{\text{slide}^2 + \text{shift}^2}$). The magnitude of *twist* characterizes the degree of torsional twisting of the double helix. Only the parameters of *roll* and *tilt* make the main contribution to bending of the DNA molecule, herewith the total local bend of each dinucleotide step can be represented as $\sqrt{\text{tilt}^2 + \text{roll}^2}$. Thus, the DNA trajectory, modeled in this way is a spatial curve, the shape of which is mainly determined by the accumulation of local bends (of specific magnitude and direction), local displacements, translation by *rise* and rotation by *twist*. In a recent study, the genome-scale computational analysis of DNA curvature based on three-dimensional trajectories of DNA molecules calculated from a set of inter base pair conformational parameters was carried out for *Arabidopsis* and rice (Masoudi-Nejad et al., 2011).

The rigid transform includes six degrees of freedom, consisting of three translational and three rotational components. The most popular method of storing and combining these transforms are matrices. Use of the matrix transformation technique is a standard method for calculation of the DNA path (De Santis et al., 1990; Bolshoy et al., 1991; Babcock et al., 1994; Liu, Beveridge, 2001). Quaternions are the more straightforward and robust alternative for matrices. They are more compact, reduce the volume of algebra and minimize computation. However, the unit quaternion can represent only a rotation without translation. Dual quaternions allow us to unify the translation and rotation into a single state.

In the present guide to DNA shape analysis the use of unit and dual quaternions to coarse-grained modeling of three-dimensional DNA structures is described in detail. Proven algorithms for efficient prediction and analysis of the shape of DNA molecules are described as follows. The discussed methods are fast and robust making them appropriate for operative prediction of curvature- and bendability-dependent DNA shape at any scale.

1. Obtaining base pair step parameters

The calculation of the DNA helix trajectory is performed based on the average values of local helical parameters. These parameters can be obtained from various sources, such as scientific papers, the base pair parameters database (DiProDB, <http://www.bionet.nsc.ru/vogis/download/pict-2017-21/appx11.pdf>)

¹ Supplementary materials are available in the online version of the paper: <http://www.bionet.nsc.ru/vogis/download/pict-2017-21/appx11.pdf>

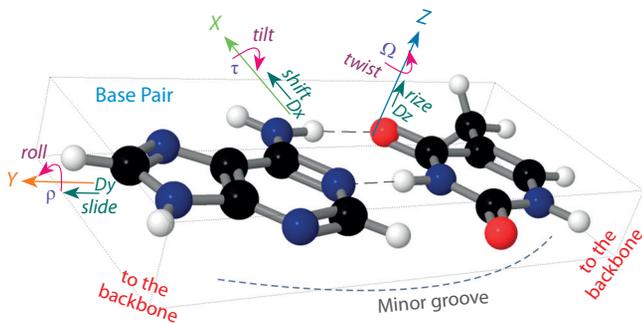


Fig. 1. Base pair orientation parameters, following the Cambridge convention. Three rotations (*twist*, *roll*, *tilt*) and three translations (*rise*, *slide*, *shift*) of a base pair about the Z, Y and X axes are indicated.

diprod.b.leibniz-fl.de) or alternatively, they can be deduced from three-dimensional DNA structures. This last method is optimal, since it allows us to choose the most suitable or desirable to resolve specific problems the DNA structures (with specific nucleotide content or conformation, obtained in different conditions etc.) to deduce base pairs parameters. This significantly increases the flexibility of DNA molecular modeling and expands the possibilities for further investigations. In particular, specific sets of base pair parameters can be used for modeling and analysis of unusual DNA structures that are different from canonical B-DNA.

As the Protein Data Bank (PDB) is the largest current source of information about the three-dimensional structures of biological molecules (Berman et al., 2003), including nucleic acids, its will be enough to obtain primary data for further estimation of DNA base pair geometry. Currently, more than 1 700 DNA structures are publicly available in the PDB database. They are represented in different conformations obtained using a range of methods, such as X-ray crystallography (of various resolutions), solution NMR spectroscopy, neutron diffraction and others. However, many DNA molecules are chemically modified or represented in complex with ligands. It seems quite clear that DNA conformation is substantially affected by crystal packing, bases modification and interactions of the DNA with the other molecules or ions. Thus, the most suitable base pair conformational parameters for modeling of the DNA structures are provided by NMR in water only solutions. Previously it was shown, that distinct from the X-ray crystallographic analysis, helical parameters derived from NMR structures can correctly predict the curvature of DNA molecules (Gabrielian, Pongor, 1996). However, it should be noted that until the late 1990s the accuracy of the NMR method was not sufficient to evaluate small intrinsic bends of the DNA axis (Vermeulen et al., 2000). A full set of helical parameters from the selected DNA structures can be deduced using the Curves+ (Lavery et al., 2009) or 3DNA (Lu, Olson, 2003) programs.

As a rule, each DNA structure from PDB is represented by 5–10 conformers possessing the lowest energy. For this reason, base pair parameters deduced from individual structures should be averaged over all conformers, and the standard deviation for each parameter should be evaluated (further, the standard deviations can be used for estimation of the dynamic persistence length). To eliminate errors associated with “end

effects”, only the central base pairs should be used in computation. Since structural properties of dinucleotides in the DNA molecule depend upon the flanking base pairs, the parameters of base pair steps should wherever possible be applied in tetra-nucleotide context. (The base pair step parameters used here for reconstruction of the DNA trajectories are provided in the Supplementary material.)

2. DNA path modeling

According to quaternion algebra, the vector function of the DNA sequence describing the position of the n -th base pair in the DNA path in the Cartesian coordinate system can be represented as:

$$P_{(n)} = \sum_{i=2}^n \left[\prod_{k=2}^i \left[Q_{\Omega_k \rho_k \tau_k} \right] \circ \vec{V}_{(x_i + Dx_i, y_i + Dy_i, z_i + Dz_i)} \circ \left\{ \prod_{k=2}^i \left[Q_{\Omega_k \rho_k \tau_k} \right] \right\}^{-1} \right], \quad (1)$$

where $\prod_{k=2}^i \left[Q_{\Omega_k \rho_k \tau_k} \right]$ is the consistent quaternion multiplication of the quaternion $Q_{\Omega_k \rho_k \tau_k}$ realizing rotation about the Z, Y and X axes consistently by angles, defined by the *twist* (Ω), *roll* (ρ) and *tilt* (τ) parameters of the neighboring base pairs ($k-1, k$):

$$Q_{\Omega_k \rho_k \tau_k} = \begin{bmatrix} \cos \frac{\Omega_k}{2} \cos \frac{\rho_k}{2} \cos \frac{\tau_k}{2} + \sin \frac{\Omega_k}{2} \sin \frac{\rho_k}{2} \sin \frac{\tau_k}{2} \\ \cos \frac{\Omega_k}{2} \cos \frac{\rho_k}{2} \sin \frac{\tau_k}{2} - \sin \frac{\Omega_k}{2} \sin \frac{\rho_k}{2} \cos \frac{\tau_k}{2} \\ \cos \frac{\Omega_k}{2} \sin \frac{\rho_k}{2} \cos \frac{\tau_k}{2} + \sin \frac{\Omega_k}{2} \cos \frac{\rho_k}{2} \sin \frac{\tau_k}{2} \\ -\cos \frac{\Omega_k}{2} \sin \frac{\rho_k}{2} \sin \frac{\tau_k}{2} + \sin \frac{\Omega_k}{2} \cos \frac{\rho_k}{2} \cos \frac{\tau_k}{2} \end{bmatrix}. \quad (2)$$

Vector $\vec{V}_{(x_i + Dx_i, y_i + Dy_i, z_i + Dz_i)}$ contains the coordinates of the i -th element (atom, molecule, point or any structure) in the intrinsic coordinate frame of the dinucleotide step (for example, the coordinates of the idealized phosphate of the first strand are contained in the vector $\vec{V}_{(Dx_i, 8.91 + Dy_i, 2.08 + Dz_i)}$). The quaternion of rotation and translational vector together calculate the rigid-body transformation between two successive base pair steps.

The initial displacement on *shift* (Dx), *slide* (Dy) and *rise* (Dz) for each transformation can be simplistically represented using dual quaternions:

$$\begin{cases} P_{(n)} = \sum_{i=2}^n \left[\hat{Q}_{TR_i} \circ \hat{Q}_{e_i} \circ \hat{Q}_{TR_i}^* \right] \\ \hat{Q}_{TR_i} = Q_{r_i} + \varepsilon \frac{1}{2} Q_{r_i} \circ Q_{d_i} \\ Q_{r_i} = \prod_{k=2}^i \left[Q_{\Omega_k \rho_k \tau_k} \right] \\ Q_{d_i} = [0, Dx_i, Dy_i, Dz_i] \\ \hat{Q}_{TR_i}^* = Q_{r_i}^* - \left[\varepsilon \frac{1}{2} Q_{r_i} \circ Q_{d_i} \right]^* \end{cases}. \quad (3)$$

The dual quaternion \hat{Q}_{e_i} contains the coordinate of any element, which is associated with the i -th base pair (for example, the coordinates of the base pair center are contained in the dual quaternion $\hat{Q}_{e_i} = [1, 0, 0, 0][0, 0, 0, 0]$; coordinates of the idealized phosphate of the first strand in $\hat{Q}_{e_i} = [1, 0, 0, 0][0, 0, 8.91, 2.08]$).

The use of dual quaternions is a fast, simple and robust way of molecular modeling. For example, to model base pair conformation, determined by such local (inter base pair) parameters as *shear* (Sx), *stretch* (Sy), *stagger* (Sz), *buckle* (κ),

propeller (π) and *opening* (σ) the dual quaternion \hat{Q}_{e_i} for each base in the i -th pair can be represented as:

$$\begin{cases} \hat{Q}_{e_i} = Q_{\sigma_i \pi_i} + \varepsilon \frac{1}{2} Q_{\sigma_i \pi_i} \circ Q_{S_i} \\ Q_{S_i} = [0, Sx_i, Sy_i, Sz_i] \end{cases} \quad (4)$$

Similar to dinucleotide step parameters, the local geometric parameters of base pairs can be obtained from articles or deduced from published three-dimensional DNA structures using Curves+ or 3DNA programs. The atomic coordinates for each base and sugar phosphate backbone can be obtained from relevant publications (Clowney et al., 1996; Gelbin et al., 1996; Parkinson et al., 1996; Olson et al., 2001) or structure databases.

Although displacements by *slide* and *shift* do not contribute to DNA bending, they make the main contribution to noise in further measurements. For this reason, these parameters can be ignored in subsequent *in silico* estimations of the DNA curvature. If for subsequent measurements only the DNA trajectory, excluding displacement by *shift* and *slide* is required, the use of dual quaternions is redundant and the equations above for unit quaternions can be simplified to:

$$\begin{cases} P_{(n)} = \sum_{i=2}^n [x_i, y_i, z_i] \\ x_i = \frac{Dz_i}{\|Q^i\|} (2Q_3^i Q_1^i + 2Q_2^i Q_0^i) \\ y_i = \frac{Dz_i}{\|Q^i\|} (2Q_3^i Q_2^i - 2Q_1^i Q_0^i) \\ z_i = \frac{Dz_i}{\|Q^i\|} ([Q_3^i]^2 - [Q_2^i]^2 - [Q_1^i]^2 + [Q_0^i]^2) \\ Q^i = \prod_{k=2}^i [Q_{\Omega_k \rho_k \tau_k}] \end{cases} \quad (5)$$

Calculation of the three-dimensional DNA structures was carried out in the local Cartesian coordinate system in accordance with the transformations defined by the Cambridge convention on the definitions and nomenclature of nucleic acid structure components (Dickerson, 1989). The center of the first base pair is located at the origin of the coordinate frame, at point 0.

3. Features of DNA molecule estimation from helical patch

Curvature

Overall, the curvature vector of a spatial line is defined as the derivative of the tangent unit vector along this line. Its modulus is the inverse of the curvature radius and its direction is the direction of the main normal to the curve (Landau, Lifshitz, 1970). In terms of a DNA molecule this means the angular deviation between the local helical axes of the successive base pairs. The curvature forms the basis for further estimations of the shape of the DNA molecule. For this reason, there are many ways in which it can be measured. As a rule, the curvature is estimated in DNA curvature units, where one curvature unit is defined as the average curvature of DNA in the nucleosome core particle, 1/42.8 Å (Trifonov, Ulanovsky, 1987).

The DNA curvature of a segment can be calculated from the inverse of the radius of a circumscribed circle of a triangle

with vertices on helix axis coordinates at the center and both ends of this segment. If the sliding window has a length of $2hw$ bp, the DNA curvature C in the i -th position is given by:

$$\begin{cases} C_i = 4\sqrt{(P(P-a)(P-b)(P-c))/abc} \\ P = (a+b+c)/2 \end{cases} \quad (6)$$

where a , b and c are the distances between the points $(i-hw, i)$, $(i, i+hw)$ and $(i-hw, i+hw)$, respectively.

Curvature estimation during the least square circle (LSC) fit is based on the fitting of a circle to the coordinates of the DNA molecule curved in a plane. The radius of this circle is taken as a measure of the curvature of the DNA fragment (Kanhare, Bansal, 2003). In order to evaluate the planarity of the bend and further projection, the best fit (least squares) plane (LSP) is calculated to a set of base pair centers of the analyzed segment. The possibility of applying this method should be grounded by low values of the root-mean-square deviation (RMSD), estimated for LSC and LSP (RMSD values of distances from the reference and fitted feature (circle or plane)).

The original interpretation of DNA curvature was proposed by De Santis (De Santis et al., 1988). Since DNA curvature $C(n)$ is a local property of the DNA axis and represents its directional change along the sequence, the distortions of the B-DNA axis along the chain gives the value of curvature per turn in modulus and phase, calculated for recurrent turns along the sequence (De Santis et al., 1988):

$$\overline{C}(n) = v(n_2 - n_1)^{-1} \sum_{s=n_1}^{n_2} [\rho_s - i\tau_s] e^{\frac{2\pi is}{v}} \quad (7)$$

where $C(n)$ is the average curvature vector, characterizing the orientation deviation of the helical axis between n_1 , and n_2 sequence numbers, per turn of DNA with helical periodicity v ($v = 10.4$), assigned to the position of the average sequence number of the tract ($n = (n_2 + n_1)/2$), the local deviation of the s -th base pair plane from the canonical B-DNA represented as a complex vector in terms of the *roll* and *tilt* angles. This is the most popular method for estimation of DNA curvature from the nucleotide sequence directly, without modeling of the helical path in advance.

The local bend angle characterizes the deflection of the helix axis and is calculated as the angle between the tangent vectors in the direction along the contour of the DNA molecule (or as arccosine of a scalar product if these tangents are unit vectors). The tangent vectors can be represented as the unit vectors, indicating orientation of the base pair centers, or as the vectors, connecting n and $n+s$ bp ($s \in [3, 15]$), or alternatively, as the unit vectors (normals) averaged over 11–15 bp (centroids of helical turns). Some authors distinguish the bend angles estimated using of tangent vectors obtained in the latter two ways as successive and cumulative (Kanhare, Bansal, 2003). The cumulative bending angle between the averaged normal vectors >15 bp apart is often used in the estimation of DNA curvature (Goodsell, Dickerson, 1994; Gabrielian, Pongor, 1996).

DNA curving can be estimated based on radius of gyration (see Supplementary materials for detail). When comparing DNA molecules of identical sizes the radius of gyration will be smaller for curved and larger for extended molecules (Dlagic, Harrington, 1998). Furthermore, information about the different shape properties such as symmetry, anisotropy,

asphericity, acylindricity and more can be extracted from the various combinations of the orthogonal components of the radius of gyration tensor (Jernigan et al., 1987; Olson et al., 1993; Dlakic, Harrington, 1998; Kanhere, Bansal, 2003; Rawat, Biswas, 2009).

The *d-max* (Tung, Burks, 1987) is the maximum orthogonal distance from the base pair center to the straight line connecting the ends of the analyzed segment, evaluated over all base pair centers of this segment. This value is estimated as the maximal perpendicular distance (*p-dist*) over all base pair centers. In its turn, the perpendicular distance is the minimal distance from the *i*-th base pair to the straight line connecting the ends of the segment.

Lastly, the curvature of the DNA molecule can be simply characterized by the ratio of the curvilinear distance (contour length) to the linear distance between the ends of the segment (*SD* value) (Eckdahl, Anderson, 1987; Tan, Harvey, 1987) and vice versa (*Qc*, (Dlakic, Harrington, 1998; Kanhere, Bansal, 2003; Matyašek et al., 2013)). Large values of *C*, *d-max* and *SD* indicate a more curved DNA fragment.

The sliding window technique is a common way for analyzing this distribution. Reducing the size of the sliding window increases the resolution and magnitude of the curvature measurements, but also increases the noise that is determined by local bends (by *roll* and *tilt*) and displacements (by *shift* and *slide*) of adjacent base pairs. In order to reduce the noise, the size of the sliding window as a multiple of a helical turn (10 or 10.4 bp), from 20 bp is recommended. Furthermore, since local bends during a helical turn, as a rule, are mutually compensated, when the curvature is estimated for a position in the center of sliding window, the size of a half window should be a multiple of 10 (for example, see the analysis of the “Hagerman paradox”, below).

Persistence length

Persistence length is a key parameter for quantitative interpretation of the conformational properties of DNA. Theoretically, the persistence length can be formulated in terms of the magnitude of the projection vector and the tangent-tangent correlation function. In previous studies, both methods were successfully applied to estimation of the persistence length of DNA molecules from the reconstructed 3D trajectories of helical paths (Shpigelman et al., 1993; Bednar et al., 1995; Schellman, Harvey, 1995; Vologodskaja, Vologodskii, 2002).

In terms of projection vector, the persistence length (*P*) can be defined as the average projection of the end-to-end vector (I_N) onto the unit vector of the first segment (I_1) in the limit of infinite chain length (Flory, 1969). It is likely that for a long heterogeneous chain with average link angle the more statistically significant is the average projection of end-to-end vector on the direction (unit vector) of each segment along the chain. Thus the persistence length is averaged over position and direction is given by:

$$P = \lim_{N \rightarrow \infty} \frac{1}{N} \left\langle \sum_{j=1}^N \frac{\overline{I_N \cdot I_j}}{|I_j|} \right\rangle. \quad (8)$$

Persistence length, measured in this way characterizes the chain length through which the memory of the initial orientation persists. So that its will be strongly dependent on the direction, length, homogeneity and shape of molecule.

Nevertheless, for the cases when $L \gg P$ this method can be effectively used for the operative estimation of the DNA persistence length.

According to the tangent-tangent correlation function (Landau, Lifshitz, 1958):

$$\langle \cos(\theta) \rangle = \left\langle \overline{I(s_i) \cdot I(s_{i+1})} \right\rangle = e^{-\frac{s_{i+1}-s_i}{P}}, \quad (9)$$

where $\langle \cos(\theta) \rangle$ is the average cosine calculated from the complete set of local bend angles (θ) measured between the unit vectors (I) tangent to the chain at the points s_i and s_{i+1} (spaced of contour length l ($s_{i+1}-s_i$)) collected over the entire chain. Tangent vectors can be calculated in the various ways described above keeping in mind that the method employed will dramatically affect the deflection angle and, hence, persistence length estimation. In particular the persistence length strongly depends on the scale of measurements, i. e. when unit segments of the DNA chain are bonds between base pairs, or when they are centroids of helical turns (see Supplementary materials).

The average cosine θ is a multiplicative function of segment length l , assuming that the average directional correlation between two segments decays exponentially along the chain. The distribution of $-\ln\langle \cos(\theta) \rangle$ is evaluated for the segments of different length l , where l is the average curvilinear distance between base pair centers, separated by the n bp window, moving along the sequence with a step of ≥ 1 bp. Since correlation in the orientation of segments decays exponentially to their distance, the $-\ln\langle \cos(\theta) \rangle$ will be linear. The persistent length can be calculated from the inverse of the regression slope of the $-\ln\langle \cos(\theta) \rangle$ distribution plot. Application of this method should be justified by a good linear fit of the $-\ln\langle \cos(\theta) \rangle$ distribution (coefficient of determination, $R^2 \geq 0.9$).

The equations above are suitable only for intrinsically straight, homogeneous polymers. However, in most cases the analyzed DNA is not straight, but contains numerous planar and coplanar static bends as well as segments with different flexibilities, located at any position along the chain. For such heterogeneous polymers, the persistence length strongly depends on the starting position, the direction, and the length of the fragment. One of possible solutions is the division of the DNA path into a set of large fragments between the bends and a set of small fragments overlapping these bends, then estimation of P can be conducted for each of these fragments separately (Supplementary material contains examples for different cases). Following that, the average persistence length of a chain can be derived from summation of P over all fragments (N_j) with length l (Rivetti et al., 1998):

$$\frac{L}{P} = \sum_{i=1}^{N_j} \frac{l_i}{P_i}. \quad (10)$$

Thus, the resultant persistence length (averaged over all fragments of molecule) will be mainly determined by the minimal P among all fragments. Furthermore, it seems that the contribution of intrinsic bends to persistence length calculated in this way will be decreased.

The apparent persistence length (P_a) of the DNA molecule includes contributions from both static (P_s) and dynamic (P_d) persistence lengths, which are related as follows (Trifonov et al., 1988; Schellman, Harvey, 1995):

$$\frac{1}{P_a} = \frac{1}{P_s} + \frac{1}{P_d}. \quad (11)$$

Static persistence length is determined by the intrinsic DNA curvature (shape) at the minimum energy conformation, without accounting for thermal fluctuations. The equations above, applied to the predicted DNA trajectories, estimate the static persistence length. Conversely, the dynamic persistence length characterizes the rigidity of the DNA molecule in thermal fluctuations of the angles between adjacent base pairs. Assuming crystal packing, interactions with proteins and thermal fluctuations as external perturbing force fields, eliminating correlations in the bend directions, the dynamic persistence length can be estimated from the DNA sequence using a set of dispersions of *roll* (ρ) and *tilt* (τ) angles (which can be obtained from the same sources as for the base pair parameters or during averaging of base pair parameters, see Section 1) for all dinucleotide steps (Vologodskaya, Vologodskii, 2002):

$$\frac{1}{P_d} = \frac{1}{2l} \sum_{i=1}^{16} p_i \theta_i^2, \quad (12)$$

where p_i is the probability of dinucleotide i , the length of segment $l \ll P$ is an average *rise* ($\sim 3.38 \text{ \AA}$) and θ_i^2 is the variance of the local bend angle (in radians) of i -th dinucleotide ($\theta_i^2 = \Delta\rho_i^2 + \Delta\tau_i^2$, ρ and τ in radians, $\theta_i \ll \pi/2$).

However it should be noted that whereas flexibility of base pair steps in overall correlates well with local bends, as determined by *roll* and *tilt* (Olson et al., 1998; Packer et al., 2000; McConnell, Beveridge, 2001), the bends produced by A-tracts are highly stable with a midpoint of structural transition near 30 °C (Chan et al., 1993) and a melting temperature above 37 °C (Chan et al., 1990; Jerkovic, Bolton, 2000). Thus, although the static persistence length of A-tracts-produced bends is very small, the dynamic persistence length of these bends will be very large. Since P_a is always less than either P_s or P_d the correlation between apparent persistence length, calculated from the summation of P_s^{-1} and P_d^{-1} , and the stiffness of DNA containing A-tract induced bends will be lost. The extremely low static (P_s) and exceptionally high dynamic (P_d) persistence lengths at relatively large apparent persistence length (P_a) were noted in a recent study of DNA molecules containing A-tract related large intrinsic bends (Mitchell et al., 2017).

Helical phasing. DFT analysis

Discrete Fourier transform is extremely useful for macroscopic DNA curvature analysis. This is due to its ability to reveal periodicity in analyzed DNA properties, such as local curvature in sliding window, as well as the relative strength of any periodic components. Subsequent DFT frequency spectrum analysis allows us to detect and quality evaluate the regular alternation of DNA bends along the molecule. The discrete Fourier transform is defined as:

$$F_k = \sum_{j=0}^{N-1} x_j e^{2\pi ijk/N}, \quad k \in [0, N-1], \quad (13)$$

where x_j is a property value in the j -th position, represented as a complex number with a zero-valued imaginary part, and k is the frequency domain (for positive frequencies $k \in [0, N/2-1]$).

In the case of DNA analysis, the sampling frequency (f_s , sampling rate) is the ratio of the data set size (N) to the step of the sliding window. The minimal frequency (frequency resolution) is given by $\Delta f = f_s/N$ and maximal frequency (Nyquist frequency) is $f_s/2$. Since this DFT spectrum is N -periodic, each

frequency domain characterizes the distribution of values of the analyzed features with frequencies of f per N bp and a maximum $1/2 \text{ bp}^{-1}$ ($f = \Delta f x = x$, $x \in [0, N/2]$).

The magnitude of the frequency domain for the real input data is calculated as:

$$|F_k| = \sqrt{\left[\sum_{j=0}^{N/2-1} x_{2j} e^{2\pi ijk/(N/2)} \right]^2 + \left[e^{2\pi ik/N} \sum_{j=0}^{N/2-1} x_{2j+1} e^{2\pi ijk/(N/2)} \right]^2}. \quad (14)$$

The phasing of systematic bends can be evaluated from the data of distribution of the local curvature in the course of frequency analysis of the DFT power spectrum. The frequency of helix turns (fh) in the DFT spectrum is $N/10$ (where 10 is the floor of average *twist*, 10.4 bp) and the initial phase is 0° . The local bends repeated with a frequency multiple of fh or with variations around this are phased with the helix screw: they have the same direction and progressively increase the macroscopic curvature of the DNA molecule. The fractional part of the ratio of fh to the frequency of local bends repeated in antiphase is near to 0.5: these bends are oriented oppositely (phase shift is 180° , $2\pi \cdot 0.5$) and mutually compensate each other.

In order to demonstrate how phasing of DNA bends affects the shape of molecules, the trajectories of the $(A_5N_5)_n$ (curved DNA, local bends in phase) and $(A_5N_{10})_n$ (straight DNA, local bends out of phase) sequences were modeled and the DFT power spectrum of curvature distribution was analyzed (Fig. 2). If sampling is 128, the $fh = 13$ ($128/10$). The $(A_5N_{10})_n$ DNA fragment contains systematic bends, repeated with a frequency between 8 and 9 that corresponds to $\{fh/8.5\} \approx 0.5$. Hence, the phases of neighboring bends will be opposite and the molecule is almost straight (see Fig. 2). The systematic bends of the $(A_5N_5)_n$ DNA fragment repeated with a frequency of 13 corresponds to fh , hence, the local bends alternate in phase with the helix periodicity, and the resulting global curvature of DNA molecule will be large (see Fig. 2).

For operative estimation of the phasing of DNA bends, the DFT power spectrum can be evaluated only for frequencies near to the frequency of the helix screw (ω , average twist angle, 34.6°) (Gabrielian, Pongor, 1996). In this case, the period is 2π (360° , one helical turn), hence frequency (k/N) is $34.6^\circ/360^\circ$ and the average twist corresponds to angular frequency $\omega = 2\pi 34.6^\circ \pi / 2\pi 180^\circ$ (in radians). The discrete Hartley transform (DHT) is an alternative for the real DFT. DHT for angular frequency ω is given by:

$$H_\omega = \frac{1}{N} \sum_{j=0}^{N-1} x_j \text{cas}(j\omega), \quad (15)$$

where $\text{cas}(j\omega) = \cos(j\omega) + \sin(j\omega)$.

The magnitude of this frequency bin in terms of DHT can be expressed as:

$$|H_\omega| = \sqrt{\left[\frac{1}{N} \sum_{j=0}^{N-1} x_j \cos(j\omega) \right]^2 + \left[\frac{1}{N} \sum_{j=0}^{N-1} x_j \sin(j\omega) \right]^2}. \quad (16)$$

The DFT power spectrum should be calculated for frequencies in range of minimal and maximal values of *twist* angle (as a rule from 27° to 42° for B-DNA). The supplementary material contains some recommendations on the use of DFT analysis in studying of the DNA curvature distribution.

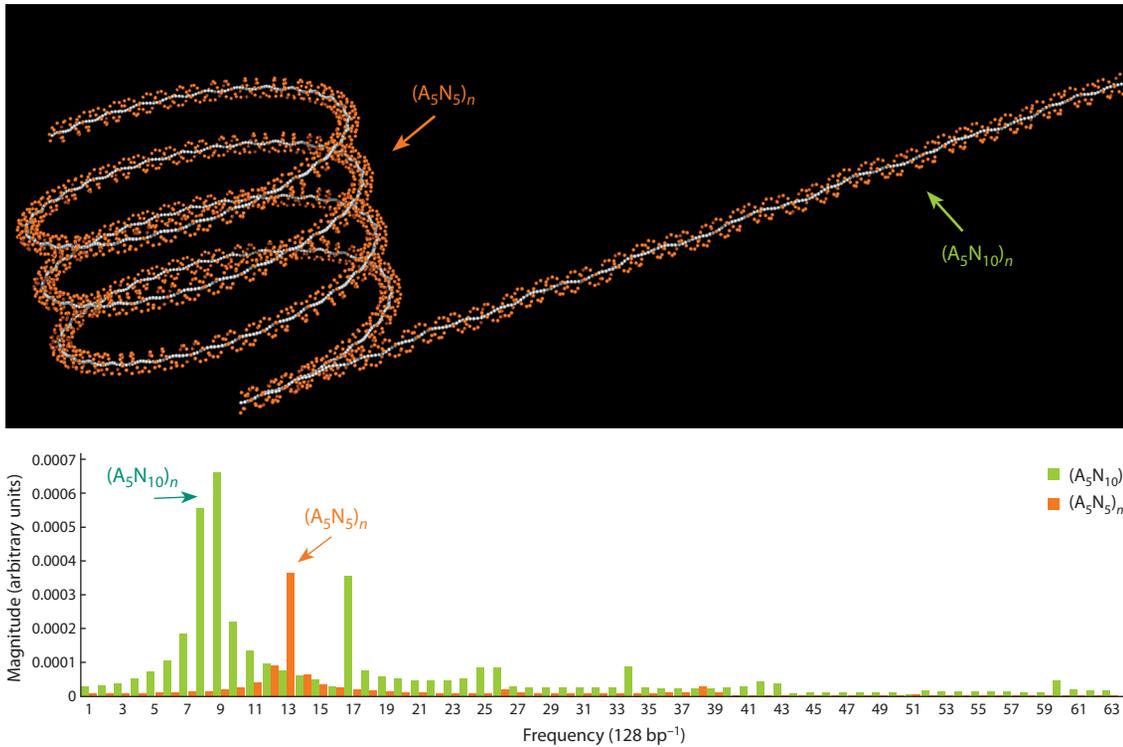


Fig. 2. Reconstructed DNA paths of the experimental molecules, containing repeated A-tract in phase $((A_5N_5)_n)$ and out of phase $((A_5N_{10})_n)$ with the helix screw. DFT power spectral for both molecules were calculated using values of DNA curvature as input data, estimated in a sliding window of 40 bp and step of 1 bp, using equation (6). The dominant frequencies for each spectrum are indicated.

Phasing of coplanar and non-periodic bends

It seems obvious that non-periodically bends in phase as well as bends in opposite directions in antiphase will increase the macroscopic DNA curvature. Furthermore, the orientation of bends is not distinguished by the DFT method, so that coplanar bends will be taken as planar, leading to distortion of the results. For this reason, the definition of phase relative to the direction of the DNA trajectory will be useful. In order to define the relative phase (φ_{vb}) of the vector position of the center of bend (“bend vector”, for example this can be the center of an analyzed segment: $vb_{n+s} = v_{n+s} - v_n$) the “directional vector” $vd_n = v_{n+2s} - v_n$ was aligned with the X axis and then the coordinate frame was rotated around the Y axis by 90° to align vd with the Z axis (the direction of this rotation is determined by the sign of x coordinate of vd). The quaternion for realizing these rotations is given by Q^φ :

$$\begin{cases} Q^\varphi = \left[0.7, 0, -0.7 \frac{x_{vd}}{|x_{vd}|}, 0 \right] \cdot Q^{vx} \\ Q^{vx} = \left[\sqrt{\frac{1+qR_0}{2}}, \frac{qR_1}{\langle qR \rangle} \sqrt{\frac{1-qR_0}{2}}, \frac{qR_2}{\langle qR \rangle} \sqrt{\frac{1-qR_0}{2}}, \frac{qR_3}{\langle qR \rangle} \sqrt{\frac{1-qR_0}{2}} \right], \quad (17) \\ qR = [0, x_{vd}, 0, 0] \cdot \frac{\overline{vd}}{|vd|} \end{cases}$$

where $\sqrt{\frac{1+qR_0}{2}}$ equivalent to the $\cos(\text{acos}(qR_0)/2)$, $\sqrt{\frac{1-qR_0}{2}}$ is the $\sin(\text{acos}(qR_0)/2)$, and $\{qR_1, qR_2, qR_3\}/\langle qR \rangle$ are the direction cosines at X, Y and Z axes respectively.

To align the unit vector with the X axis we can simplify the calculation of qR to:

$$qR = \left[|x_{vd}|, 0, z_{vd} \frac{x_{vd}}{|x_{vd}|}, -y_{vd} \frac{x_{vd}}{|x_{vd}|} \right], \quad (18)$$

where the $\frac{x_{vd}}{|x_{vd}|}$ ratio is necessary to get a sign. The phase is estimated by all four quadrants as the cosine of the angle between the projection of curvature vector on XY plane and the direction of the X axis $(1, 0, 0)$:

$$\begin{cases} vb = Q^\varphi \cdot vb \cdot (Q^\varphi)^{-1} \\ \varphi_{vb} = [1, 0, 0] \cdot \frac{vb}{|vb|} \end{cases} \quad (19)$$

For example, Hagerman has shown that DNA sequences containing repeating runs of A_4T_4 in phase were significantly bent, whereas those with T_4A_4 (A-tract in opposite polarity) were almost straight (Hagerman, 1986). Thus, there is a fundamental difference in the structure of A_nT_n and T_nA_n DNA segments, despite their identical nucleotide composition and phased bends. In fact, the distribution of curvature in these DNA molecules have the same DFT frequency spectrum with a peak of magnitude for 10.04 bp period, which is multiplied to the period of helix axis $(\{10.4/10.04\} = 0.036 \ll 5)$ that is also consistent with experimental data (Price, Tullius, 1993). Thus, at first glance in both cases the phased bends should increase the macroscopic curvature of DNA molecules significantly. However analysis of the phase spectrum makes it clear (Fig. 3) why in the first case $((CA_4T_4G)_n)$ DNA fragments

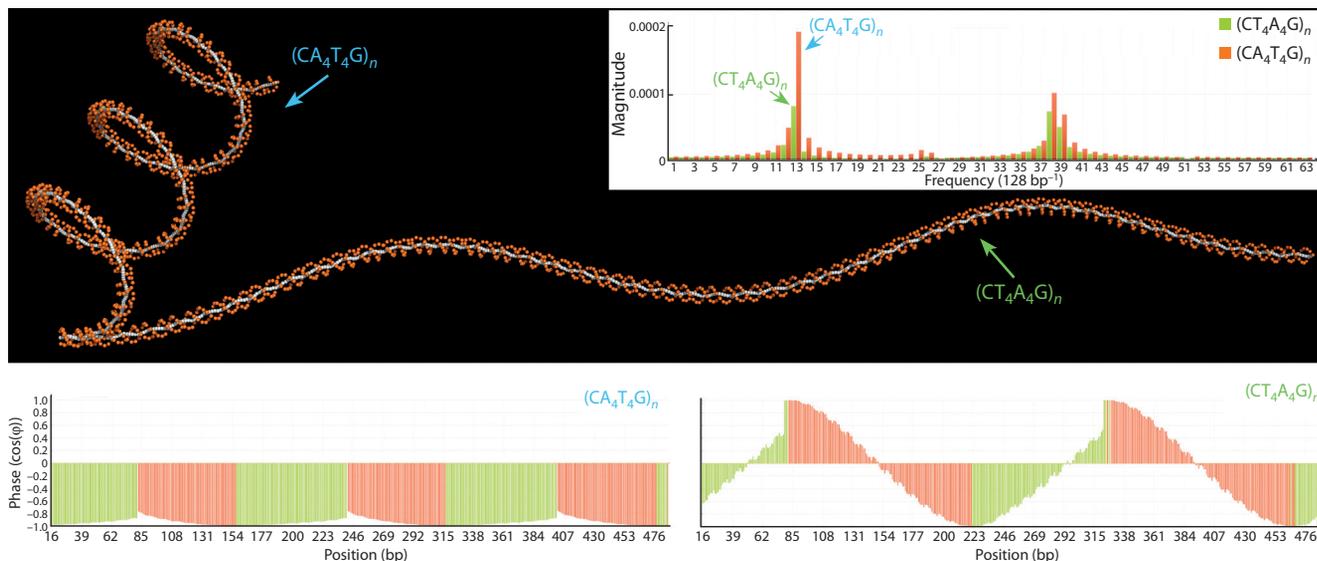


Fig. 3. The “Hagerman paradox” analysis using discrete Fourier transform and the phase spectrum calculated by equation (17). Curved $((CA_4T_4G)_n)$ and “straight” $((CT_4A_4G)_n)$ DNA molecules have similar frequency spectral, but differ in the phase spectrum of curvature, calculated relative to direction of the DNA path. The macroscopic bends of the “straight” fragment orient in opposite directions and compensate each other. Curvature was analyzed in a sliding window of 30 bp and 1 bp step.

are exclusively curved, but in the other case $((CT_4A_4G)_n)$ the fragments are almost straight (“zig-zag” structure). In particular it shows that the bends produced by $((CT_4A_4G)_n)$ are oriented in opposite directions and mutually compensate each other (see Fig. 3).

However, it should be noted that magnitudes of DFT power spectrum of the $((CT_4A_4G)_n)$ straight fragment will be relative high only if the size of the sliding window wherein DNA curvature is estimated is in multiples of an odd number of turns. If the size of the half window is a multiple of 10 (one helical turn) the amplitude of curvature of the $((CT_4A_4G)_n)$ fragment is near to 0. This is consistent with the fact that oppositely oriented bends, in this motif, are mutually compensated in one turn (Steffl et al., 2004). The present example shows how a change in the size of the sliding window can be adopted for specific research tasks.

Conclusion

More than 30 years ago it was shown that computer modeling of DNA sequences is a viable approach to the study of the biological implications of DNA structure. A lot of research has been devoted to this problem. The estimation of the intrinsic curvature of relatively short DNA fragments (< 10 kb) can be useful in investigating various features, such as analysis of promoter and regulatory regions of specific genes, or the parts of the genome associated with recombination suppression and heterochromatin packing, nucleosome positioning, design of shape-related DNA markers and many others. Furthermore, the actual aim is the monitoring of DNA curvature and bendability over hundreds and thousands base pairs. Similar studies, previously performed for many prokaryotic and eukaryotic genomes, provided important information about their spatial organization and its influence on various biological processes. The algorithms discussed in the present guide formed the basis for such investigations and will be useful in analysis

of relatively short DNA fragments as well as for topological mapping of whole genomes.

Acknowledgments

This work was supported by the Siberian Branch of the Russian Academy of Sciences, integrated program 0324-2016-0025, project II.2P/VI.53-17.

Conflict of interest

The author declares no conflict of interest.

References

Babcock M.S., Pednault E.P., Olson W.K. Nucleic acid structure analysis. Mathematics for local Cartesian and helical structure parameters that are truly comparable between structures. *J. Mol. Biol.* 1994; 237(1):125-156.

Bednar J., Furrer P., Katritch V., Stasiak A.Z., Dubochet J., Stasiak A. Determination of DNA persistence length by cryo-electron microscopy. Separation of the static and dynamic contributions to the apparent persistence length of DNA. *J. Mol. Biol.* 1995;254(4): 579-594.

Berman H., Henrick K., Nakamura H. Announcing the worldwide Protein Data Bank. *Nat. Struct. Biol.* 2003;10(12):980.

Bolshoy A., McNamara P., Harrington R.E., Trifonov E.N. Curved DNA without A-A: experimental estimation of all 16 DNA wedge angles. *Proc. Natl. Acad. Sci. USA.* 1991;88(6):2312-2316.

Chan S.S., Breslauer K.J., Austin R.H., Hogan M.E. Thermodynamics and premelting conformational changes of phased (dA)₅ tracts. *Biochemistry.* 1993;32(44):11776-11784.

Chan S.S., Breslauer K.J., Hogan M.E., Kessler D.J., Austin R.H., Ojemann J., Passner J.M., Wiles N.C. Physical studies of DNA premelting equilibria in duplexes with and without homo dA.dT tracts: correlations with DNA bending. *Biochemistry.* 1990;29(26):6161-6171.

Clowney L., Jain S.C., Srinivasan A.R., Westbrook J., Olson W.K., Berman H.M. Geometric parameters in nucleic acids: Nitrogenous bases. *J. Am. Chem. Soc.* 1996;118(3):509-518.

- De Santis P., Palleschi A., Savino M., Scipioni A. A theoretical model of DNA curvature. *Biophys. Chem.* 1988;32(2-3):305-317.
- De Santis P., Palleschi A., Savino M., Scipioni A. Validity of the nearest-neighbor approximation in the evaluation of the electrophoretic manifestations of DNA curvature. *Biochemistry.* 1990;29(39):9269-9273.
- Dickerson R.E. Definitions and nomenclature of nucleic acid structure components. *Nucleic Acids Res.* 1989;17(5):1797-1803.
- Dlakic M., Harrington R.E. Unconventional helical phasing of repetitive DNA motifs reveals their relative bending contributions. *Nucleic Acids Res.* 1998;26(18):4274-4279.
- Eckdahl T.T., Anderson J.N. Computer modelling of DNA structures involved in chromosome maintenance. *Nucleic Acids Res.* 1987;15(20):8531-8545.
- Flory P. *Statistical Mechanics of Chain Molecules.* N.Y.: Inter-science, 1969.
- Gabrielian A., Pongor S. Correlation of intrinsic DNA curvature with DNA property periodicity. *FEBS Lett.* 1996;393(1):65-68.
- Gelbin A., Schneider B., Clowney L., Hsieh S.H., Olson W.K., Berman H.M. Geometric parameters in nucleic acids: Sugar and phosphate constituents. *J. Am. Chem. Soc.* 1996;118(3):519-529.
- Goodsell D.S., Dickerson R.E. Bending and curvature calculations in B-DNA. *Nucleic Acids Res.* 1994;22(24):5497-5503.
- Hagerman P.J. Sequence dependence of the curvature of DNA: a test of the phasing hypothesis. *Biochemistry.* 1985;24(25):7033-7037.
- Hagerman P.J. Sequence-directed curvature of DNA. *Nature.* 1986;321(6068):449-450.
- Jerkovic B., Bolton P.H. The curvature of dA tracts is temperature dependent. *Biochemistry.* 2000;39(40):12121-12127.
- Jernigan R.L., Sarai A., Shapiro B., Nussinov R. Relationship between curved DNA conformations and slow gel migration. *J. Biomol. Struct. Dyn.* 1987;4(4):561-567.
- Kanhere A., Bansal M. An assessment of three dinucleotide parameters to predict DNA curvature by quantitative comparison with experimental data. *Nucleic Acids Res.* 2003;31(10):2647-2658.
- Koo H.S., Wu H.M., Crothers D.M. DNA bending at adenine thymine tracts. *Nature.* 1986;320(6062):501-506.
- Kratky O., Porod G. Röntgenuntersuchung gelöster fadenmoleküle. *Recueil des Travaux Chimiques des Pays-Bas.* 1949;68:1106-1123.
- Landau L.D., Lifshitz E.M. *Statistical Physics, Course of Theoretical Physics.* London, UK: Pergamon, 1958.
- Landau L.D., Lifshitz E.M. *Theory of Elasticity.* Oxford, NY: Pergamon Press, 1970.
- Lavery R., Moakher M., Maddocks J.H., Petkeviciute D., Zakrzewska K. Conformational analysis of nucleic acids revisited: curves+. *Nucleic Acids Res.* 2009;37(17):5917-5929.
- Liu Y., Beveridge D.L. A refined prediction method for gel retardation of DNA oligonucleotides from dinucleotide step parameters: reconciliation of DNA bending models with crystal structure data. *J. Biomol. Struct. Dyn.* 2001;18(4):505-526.
- Lu X.J., Olson W.K. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* 2003;31(17):5108-5121.
- Masoudi-Nejad A., Movahedi S., Jáuregui R. Genome-scale computational analysis of DNA curvature and repeats in Arabidopsis and rice uncovers plant-specific genomic properties. *BMC Genomics.* 2011;12:214.
- Matyašek R., Fulneček J., Kovařík A. Evaluation of DNA bending models in their capacity to predict electrophoretic migration anomalies of satellite DNA sequences. *Electrophoresis.* 2013;34(17):2511-2521.
- McConnell K.J., Beveridge D.L. Molecular dynamics simulations of B'-DNA: sequence effects on A-tract-induced bending and flexibility. *J. Mol. Biol.* 2001;314:23-40.
- Mitchell J.S., Glowacki J., Grandchamp A.E., Manning R.S., Maddocks J.H. Sequence-dependent persistence lengths of DNA. *J. Chem. Theory Comput.* 2017;13(4):1539-1555.
- Olson W.K., Bansal M., Burley S.K., Dickerson R.E., Gerstein M., Harvey S.C., Heinemann U., Lu X.J., Neidle S., Shakked Z., Sklenar H., Suzuki M., Tung C.S., Westhof E., Wolberger C., Berman H.M. A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.* 2001;313(1):229-237.
- Olson W.K., Gorin A.A., Lu X.J., Hock L.M., Zhurkin V.B. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. USA.* 1998;95(19):11163-11168.
- Olson W.K., Marky N.L., Jernigan R.L., Zhurkin V.B. Influence of fluctuations on DNA curvature. A comparison of flexible and static wedge models of intrinsically bent DNA. *J. Mol. Biol.* 1993;232(2):530-554.
- Packer M.J., Dauncey M.P., Hunter C.A. Sequence-dependent DNA structure: tetranucleotide conformational maps. *J. Mol. Biol.* 2000;295(1):85-103.
- Parkinson G., Vojtechovsky J., Clowney L., Brünger A.T., Berman H.M. New parameters for the refinement of nucleic acid-containing structures. *Acta Crystallogr. Sect. D.* 1996;52(1):57-64.
- Price M.A., Tullius T.D. How the structure of an adenine tract depends on sequence context: a new model for the structure of TnAn DNA sequences. *Biochemistry.* 1993;32(1):127-136.
- Rawat N., Biswas P. Size, shape, and flexibility of proteins and DNA. *J. Chem. Phys.* 2009;131(16):165104.
- Rivetti C., Walker C., Bustamante C. Polymer chain statistics and conformational analysis of DNA molecules with bends or sections of different flexibility. *J. Mol. Biol.* 1998;280(1):41-59.
- Schellman J.A., Harvey S.C. Static contributions to the persistence length of DNA and dynamic contributions to DNA curvature. *Biophys. Chem.* 1995;55(1-2):95-114.
- Shpigelman E.S., Trifonov E.N., Bolshoy A. CURVATURE: software for the analysis of curved DNA. *Comput. Appl. Biosci.* 1993;9(4):435-440.
- Steff R., Wu H., Ravindranathan S., Sklenar V., Feigon J. DNA A-tract bending in three dimensions: solving the dA4T4 vs dT4A4 conundrum. *Proc. Natl. Acad. Sci. USA.* 2004;101(5):1177-1182.
- Tan R.K., Harvey S.C. A comparison of six DNA bending models. *J. Biomol. Struct. Dyn.* 1987;5(3):497-512.
- Trifonov E.N., Tan R.K.Z., Harvey S.C. Static persistence length of DNA. In: W.K. Olson, M.H. Sarma, R.H. Sarma, M. Sundaralingam (Eds.). *DNA Bending and Curvature.* N.Y.: Adenine Press, 1988; 243-253.
- Trifonov E.N., Ulanovsky L.E. Inherently curved DNA and its structural elements. In: R.D. Wells, S.C. Harvey (Eds.). *Unusual DNA Structures.* Berlin, Germany: Springer-Verlag, 1987;173-187.
- Tung C.S., Burks C. A quantitative measure of DNA curvature enabling the comparison of predicted structures. *J. Biomol. Struct. Dyn.* 1987;4(4):553-559.
- Vermeulen A., Zhou H., Pardi A. Determination DNA global structure and DNA bending by application of NMR residual dipolar coupling. *J. Am. Chem. Soc.* 2000;122:9638-9647.
- Vologodskaya M., Vologodskii A. Contribution of the intrinsic curvature to measured DNA persistence length. *J. Mol. Biol.* 2002;317(2):205-213.
- Wu H.M., Crothers D.M. The locus of sequence-directed and protein-induced DNA bending. *Nature.* 1984;308(5959):509-513.

Взаимодействия белок–ДНК: статистический анализ межатомных контактов в большой и малой бороздках

А.А. Анашкина¹✉, Е.Н. Кузнецов², А.В. Батяновский³, Л.А. Урошлев⁴, В.Г. Туманян¹, Н.Г. Есипова¹

¹ Институт молекулярной биологии им. В.А. Энгельгардта Российской академии наук, Москва, Россия

² Институт проблем управления им. В.А. Трапезникова Российской академии наук, Москва, Россия

³ Институт биофизики и клеточной инженерии НАН Беларуси, Минск, Республика Беларусь

⁴ Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, Россия

Взаимодействия между белком и ДНК по существу лежат в основе всех процессов, происходящих в живой клетке. Понимание принципов специфического распознавания сайтов ДНК позволит понять, как управляются эти процессы, и даст возможность сознательно вмешиваться в управление ими. В работе изучены методом Вороного–Делоне контакты белок–ДНК на атомном уровне в структурах 3518 комплексов из PDB (все имеющиеся на май 2017 г.), содержащих как белковые цепи, так и ДНК. Метод не содержит параметров и позволяет однозначно выявлять непосредственные контакты между атомами и характеризовать каждый контакт, помимо расстояния между атомами, площадью контакта, определяемой соответствующей гранью полиэдра Вороного. Показано, что большая часть контактов образуется между атомами белка и атомами сахарофосфатного остова ДНК (72.9%). На контакты с атомами нуклеиновых оснований, выходящих в бороздки ДНК, приходится для большой бороздки 17.0% и для малой бороздки 10.1% от всех атомных контактов. Суммарно на взаимодействия между атомами белка и атомами нуклеиновых оснований приходится 27.1% всех атомных контактов. Анализ площади доступной поверхности атомов большой и малой бороздок показал, что она коррелирует с числом контактов (коэффициенты линейной корреляции 0.94 и 0.93 соответственно), однако атомы нуклеиновых оснований, образующие водородные связи, контактируют чаще, чем этого можно было ожидать из статистических соображений. Показано, что конформационно-стабильные пептиды достаточно часто встречаются в областях связывания с ДНК. Анализ остатков в предопределенной конформации в 3518 комплексах белок–ДНК выявил 159 аминокислотных остатков в предопределенной конформации β-изгиба типа I, 15 остатков в конформации β-изгиба типа I', 6 остатков в конформации β-изгиба типа II. Остатков в конформации β-изгиба типа II' найдено не было. Анализ контактов показал, что такие остатки практически не образуют контактов с ДНК. Контакты с атомами нуклеиновых оснований найдены только в двух гомологичных структурах 3qea и 3qe9, где атомы треонина образуют контакты с атомами нуклеиновых оснований AT-пары.

Ключевые слова: комплексы белок–ДНК; разбиение Вороного–Делоне; взаимодействия белок–ДНК; специфичность узнавания.

Protein-DNA interactions: statistical analysis of interatomic contacts in major and minor grooves

А.А. Анашкина¹✉, Е.Н. Кузнецов²,
А.В. Батяновский³, Л.А. Урошлев⁴,
В.Г. Туманян¹, Н.Г. Есипова¹

¹ Engelhardt Institute of Molecular Biology RAS, Moscow, Russia

² V.A. Trapeznikov Institute of Control Sciences RAS,
Moscow, Russia

³ Institute of Biophysics and Cell Engineering of NAS of Belarus,
Minsk, Belarus

⁴ Vavilov Institute of General Genetics RAS, Moscow, Russia

The interactions between protein and DNA in essence underlie all processes in a living cell. Understanding the principles of specific recognition of DNA sites will open the way to understand how these processes are controlled and to interfere in their operation. In the paper we studied contacts between the protein and DNA at the atomic level in the structures of all the 3518 protein-DNA complexes available in PDB by the Voronoi–Delaunay tessellation method. The method unambiguously defines contacts between atoms without any parameters, and characterizes each contact by the distance between atoms and the contact area, which is determined by the corresponding face of the Voronoi polyhedron. It was shown that most contacts are formed between the protein atoms and the sugar-phosphate backbone of the DNA (72.9%). The contact with the atoms of the nucleic bases emerging into the grooves of DNA is 17.0% for a major groove and 10.1% for all atomic contacts for a minor groove. Totally, the interaction between protein atoms and nucleic base atoms accounts for 27.1% of all contacts. Analysis of the accessible surface area of atoms in the major and the minor grooves showed a correlation with the number of contacts (coefficient of linear correlation 0.94 and 0.93, respectively), however, nucleic acid atoms forming hydrogen bonds make contacts more often than may be expected from statistical considerations. It was shown that conformationally stable peptides occur sometimes in the binding regions with DNA. Analysis of the residues in a predefined conformation in 3518 protein-DNA complexes revealed 159 amino acid residues in a predefined β-bend type I conformation, 15 residues in the conformation of β-bend type I', and 6 residues in the conformation of β-bend type II. No residues in the conformation of β-bend type II' were

found. Analysis of contacts showed that such residues virtually do not form contacts with DNA. Contacts with nucleic base atoms are found only in the two homologous structures 3qea and 3qe9, where threonine atoms form contacts with atoms of nucleotide bases of the AT-pair.

Key words: protein-DNA complexes; Voronoi–Delaunay tessellation; protein-DNA interaction; specificity of recognition.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Анашкина А.А., Кузнецов Е.Н., Батяновский А.В., Урошлев Л.А., Туманян В.Г., Есипова Н.Г. Взаимодействия белок-ДНК: статистический анализ межатомных контактов в большой и малой бороздках. Вавиловский журнал генетики и селекции. 2017;21(8):887-894. DOI 10.18699/VJ17.309

HOW TO CITE THIS ARTICLE:

Anashkina A.A., Kuznetsov E.N., Batianovskii A.V., Uroshlev L.A., Tumanyan V.G., Esipova N.G. Protein-DNA interactions: statistical analysis of interatomic contacts in major and minor grooves. Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding. 2017;21(8):887-894. DOI 10.18699/VJ17.309 (in Russian)

Специфические взаимодействия между белками и соответствующими сайтами ДНК лежат в основе фундаментальных биологических процессов, происходящих в живой клетке. Изменение специфичности связывания может привести к серьезным нарушениям биологической функции и, как следствие, к заболеваниям организма. Так, изменение специфичности связывания мутированных транскрипционных факторов с ДНК наблюдается при разных типах рака (Filipova et al., 2002). В отличие от процессов с участием молекулы-адаптера, как в случае генетического кода, взаимоотношения между мономерными единицами (аминокислотами, нуклеотидами) в интерфейсах белок-ДНК не могут быть произвольными, а подчиняются определенным физико-химическим ограничениям. Гипотеза о том, что эти взаимодействия могут быть описаны неким псевдокодом, пусть гораздо менее определенным, чем генетический код, до сих пор не подтверждена и не опровергнута (Choo, Klug, 1997; Pabo, Nekludova, 2000; Venos et al., 2002), но для отдельных классов белков оказалось возможным выявить некий шаблон взаимодействий белков и ДНК и высказать соображения о причинах специфичности и влияния мутаций на специфичность (Wolfe et al., 2000).

Рассмотрим кратко историю анализа белково-нуклеиновых взаимодействий. В начале в связи с малыми объемами информации по комплексам белок-ДНК все данные суммировались, так как только таким образом можно было сформировать достаточно объемный для статистического анализа материал. Своеобразный итог этого этапа подвела работа Торнтон (Luscombe et al., 2001), в которой изучались 129 комплексов белок-ДНК. Были рассмотрены взаимодействия разных типов: вандерваальсовы контакты, ионные мостики, водородные связи. Уже тогда авторы различали неспецифические, «универсальные» взаимодействия, ответственные за связывание с ДНК в целом, и взаимодействия, «зависящие от контекста», отвечающие за распознавание последовательности нуклеиновых оснований ДНК. В дальнейших исследованиях структура комплексов белок-ДНК анализировалась для определенных типов белков, связывающихся с ДНК. Во-первых, уже были накоплены достаточные объемы данных по белковым семействам. Во-вторых, исследователи исходили из того, что взаимодействие является специфическим для данного класса связывающихся с ДНК белков (например,

цинковых пальцев (Wolfe et al., 2000)). Между тем вопрос об общих закономерностях взаимодействий белок-ДНК оставался открытым.

В нашей работе мы поставили задачу анализа пространственных контактов между атомами в структурах комплексов белок-ДНК и поиска структурных детерминант специфичности. Для начала необходимо было решить, какие данные суммируются (что дает очевидные преимущества в смысле статистики), а какие, напротив, должны быть рассмотрены отдельно в силу их специфической природы. Некоторые авторы полагают, что следует отдельно рассматривать комплексы белок-ДНК с разной специфичностью. Так, выделяют высоко-, мульти- и неспецифические комплексы (Cologna, Guo, 2016). При этом под высокоспецифическими понимают белки, которые распознают единственный сайт на ДНК. Мультиспецифические белки распознают несколько сходных сайтов. Неспецифические белки связывают ДНК независимо от последовательности. Соответственно, комплексы разделены на группы по типу распознающего белка (Cologna, Guo, 2016). Мы полагаем, что такое деление является достаточно условным. Если же оно имеет под собой реальные физико-химические основы, то анализ интерфейсов должен объективным образом выявить такое деление. Во избежание перепредставленности тех или иных взаимодействий, некоторые авторы составляют также репрезентативную выборку, исключая из рассмотрения комплексы с гомологичными белковыми цепями. Ранее в работе (Anashkina et al., 2007) мы показали, что минимальный для изучения набор должен содержать около 460 интерфейсов для оценки аминокислотного состава интерфейсов с точностью 2 % при достоверности 95 %. Поэтому мы полагаем, принимая во внимание объем существующих данных по белок-ДНК комплексам, что надо рассматривать все имеющиеся в распоряжении исследователей экспериментальные структуры комплексов. Это необходимо для повышения достоверности исследования. Таким образом, мы отбирали комплексы из PDB по наличию в структурах комплексов как белковых цепей, так и ДНК. Однако принципиальным моментом является то, что эти данные должны быть позиционированы относительно большой и малой бороздок и сахарофосфатного остова ДНК, поскольку большая и малая бороздки разделены в пространстве атомами сахарофосфатного остова,

и смешивать данные, полученные для разных бороздок, нелогично.

Материалы и методы

Заметим, что ДНК имеет сольватную оболочку, кроме того, в бороздках ДНК находится «структурированная» вода (Привалов, 1958) и ионы. Далеко не во всех комплексах положения атомов кислорода воды и ионов локализованы. По этой причине опосредованные контакты (например, мостик через воду) не рассматривались.

Разбиение Вороного–Делоне. Контакты между белками и ДНК были определены на основе анализа координат атомов пространственной структуры 3518 комплексов белок-ДНК методом разбиения Вороного–Делоне (Medvedev, 1986; Анашкина и др., 2008; Anashkina et al., 2013), аналогично тому, как это было ранее сделано для белок-белковых взаимодействий (Anashkina et al., 2007). Соответственно, термином «контакт» мы будем называть далее общую грань полиэдра Вороного, полученного в результате разбиения Вороного–Делоне. Каждый контакт характеризуется типом пары атомов, расстоянием между этими атомами и площадью общей грани полиэдра Вороного.

Результатом проведенного разбиения являются таблицы контактов между атомами аминокислот и нуклеотидов, а также между более крупными пространственными единицами – аминокислотными остатками и нуклеотидами. Список использованных в исследовании комплексов белок-ДНК и полные таблицы контактов между атомами белка и ДНК приведены в Приложении¹. Программа для построения разбиения написана на языке C++, ее исходный код доступен по запросу авторам статьи через электронную почту.

Конформационно-стабильные пептиды – короткие аминокислотные последовательности, которые включены в различные белковые глобулы преимущественно в одном конформационном состоянии. Методология выявления таких последовательностей изложена в работах (Батяновский, Власов, 2008; Батяновский и др., 2009) и сводится к сравнению конформационных состояний по совокупностям двугранных углов. При наличии преимущественной (более 60 % всех случаев) группы близких конформаций утверждается наличие конформационной стабильности у данной короткой последовательности.

Предопределенные конформации. Конформационно-предопределенные конформации – это локальные области полипептидной цепи белка, конформация которых определяется контекстом, т. е. участками полипептидной цепи справа и слева от этой области (Урошлев и др., 2015; Торшин и др., 2016). Типичным примером такого рода конформаций являются β-изгибы полипептидной цепи. Фиксация остатков β-шпильки, примыкающих к изгибу, резко ограничивает число конформационных вариантов вследствие образования псевдоцикла. Механизм «навязывания» изгибу, соединяющему два антипараллельных β-тяжа, определенной конформации объясняет образование β-изгибов I, I' и II типов, а также β-изгиба II' типа с напряженным в стереохимическом смысле элементом.

¹ См. Приложение по адресу:
<http://www.bionet.nsc.ru/vogis/download/pict-2017-21/appx12.pdf>

Результаты

В настоящей работе мы поставили цель подробно проанализировать весь имеющийся экспериментальный материал по белок-ДНК комплексам в поисках закономерностей связывания белков для каждой из бороздок отдельно, выделяя общие и уникальные черты в обоих случаях. На момент проведения исследования в банке данных PDB было найдено 3518 комплексов белок-ДНК. На каждой стадии анализа были выявлены межатомные взаимодействия, ответственные за специфическое распознавание (распознавание последовательности нуклеотидов), и сделано заключение в пользу или против существования подобия кода узнавания на уровне взаимодействия атомов или мономеров нуклеиновых кислот и белков.

На рис. 1 показано разделение атомов нуклеиновых оснований ДНК на большую и малую бороздки. В большую бороздку выходят атомы N9, C8, N7, C5, C6, N6 аденина, N1, C6, C5, C8, C4, O4 тимина, N9, C8, N7, C5, C6, O6 гуанина, N1, C6, C5, C4, N4 цитозина. В малую бороздку выходят атомы N9, C4, N3, C2, N1 аденина, N1, C2, O2, N3 тимина, N9, C4, N3, C2, N2 гуанина и N1, C2, O2 цитозина. Заметим, что атомы N9 аденина и гуанина и атомы N1 тимина и цитозина могут выходить как в малую, так и в большую бороздку (см. рис. 1), однако трехмерное строение молекулы ДНК таково, что эти атомы не доступны для контакта со стороны малой бороздки. Таким образом,

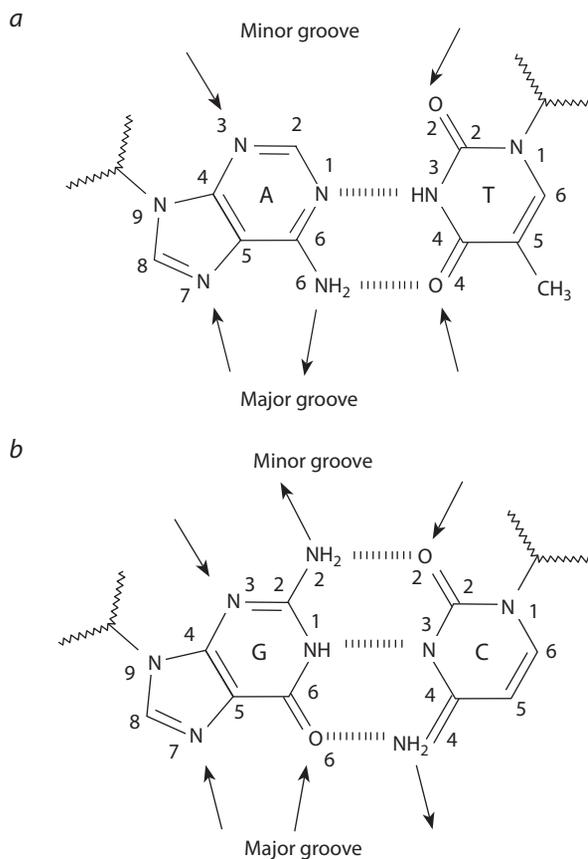


Fig. 1. DNA atoms of the major and minor grooves: a, AT pair; b, GC pair. Arrows indicate atoms that are donors and acceptors of possible hydrogen bonds.

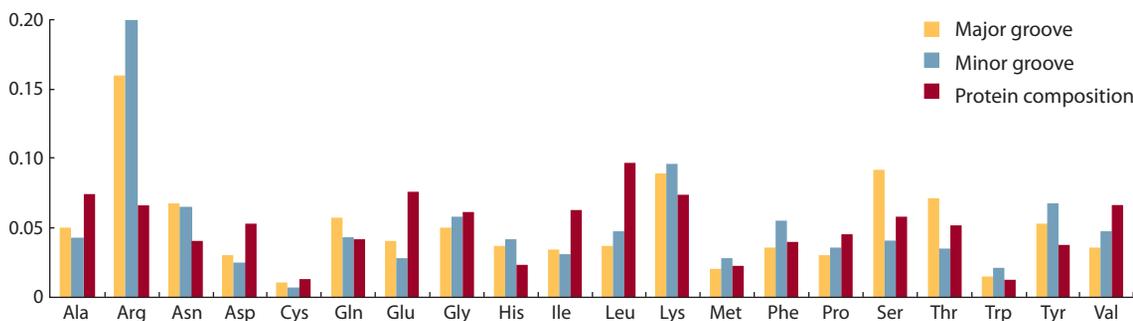


Fig. 2. Proportions of amino acids in the protein–DNA interfaces along the major (yellow) and minor (blue) grooves and in the total protein composition.

Table 1. Contacts of sidechain and backbone atoms of proteins with major and minor groove atoms, deoxyribose, and phosphate groups of DNA*

Protein atoms	DNA atoms			
	major groove	minor groove	deoxyribose	phosphate
Backbone	33380 (2.6)**	18818 (1.8)	145773 (11.2)	108173 (8.6)
Sidechain	181801 (14.4)	85026 (8.3)	444666 (33.6)	246468 (19.5)

* Based on 3518 protein–DNA complexes.

** Parenthesized are the percentages of the total number of contacts.

Table 2. The number of contacts of arginine atoms with minor groove DNA atoms calculated on a sample of 3518 protein–DNA complexes

Protein atoms	DNA base atoms													Total
	DA_C4	DA_N3	DA_C2	DA_N1	DT_N3	DT_O2	DT_C2	DG_C4	DG_N3	DG_C2	DG_N2	DC_O2	DC_C2	
ARG_C	1	14	19	8	3	24	0	2	6	0	19	27	0	123
ARG_CA	8	88	21	11	12	49	2	65	31	26	50	80	9	452
ARG_CB	22	205	172	32	45	246	52	27	42	7	110	151	24	1135
ARG_CD	65	463	278	61	64	424	45	44	157	81	266	265	79	2292
ARG_CG	28	245	173	36	54	268	44	65	97	52	133	141	30	1366
ARG_CZ	36	277	229	62	87	310	41	64	157	35	300	193	28	1819
ARG_N	7	163	140	12	10	127	5	7	26	25	64	67	12	665
ARG_NE	79	353	241	75	94	353	85	101	209	77	290	272	99	2328
ARG_NH1	253	1263	949	120	167	1187	184	190	606	134	799	731	133	6716
ARG_NH2	207	1132	872	97	231	1415	161	188	645	125	988	958	183	7202
ARG_O	25	75	70	29	28	104	21	39	77	11	57	110	5	651

Gray filling indicates contacts of guanidine nitrogen atoms.

контакты этих атомов могут быть отнесены только к большой бороздке. Атомы дезоксирибозы и фосфатной группы полагаем одинаковыми для всех нуклеотидов.

Для начала следует упомянуть, что области связывания белок–ДНК обогащены положительно заряженными аргинином, лизином и гистидином и образующими водородные связи аспарагином, серином, тирозином, в то время как содержание отрицательно заряженных и гидрофобных аминокислот понижено (рис. 2). Также можно видеть, что наблюдаются различия в аминокислотном составе областей взаимодействия белка с большой и малой бороздкой. Наиболее ярко это различие наблюдается для серинов и треонинов, которых много в большой бороздке и суще-

ственно меньше в малой. Тирозин и аргинин, напротив, чаще встречаются в малой бороздке, чем в большой.

Мы разделили весь массив контактов между атомами белков и ДНК на три категории: контакты с атомами нуклеиновых оснований по большой бороздке, по малой бороздке и контакты с атомами сахарофосфатного остова ДНК. Согласно полученным данным, большая часть контактов (53.1 %) образуется между атомами боковых радикалов аминокислот белка и сахарофосфатным остовом ДНК. Контакты боковых радикалов аминокислот белка с атомами нуклеиновых оснований, выходящих в бороздки ДНК, составляют 14.4 % в большой бороздке и 8.3 % в малой (от всех атомных контактов). Атомы основной

Table 3. The most frequent contacts of protein atoms with atoms of the major and minor grooves and of the sugar–phosphate backbone

Major groove			Minor groove			Sugar–phosphate backbone		
Protein atom type	Number of contacts with DNA		Protein atom type	Number of contacts with DNA		Protein atom type	Number of contacts with DNA	
	number	%		number	%		number	%
ARG_NH2	16469	7.4	ARG_NH2	7202	6.7	ARG_NH2	43071	4.5
ARG_NH1	13251	13.3	ARG_NH1	6716	13.0	ARG_NH1	37252	8.4
LYS_NZ	6857	16.4	LYS_NZ	3583	16.3	LYS_NZ	33617	11.9
ARG_NE	6361	19.3	PHE_CZ	2366	18.5	ARG_CD	24686	14.5
ARG_CZ	6329	22.1	ASN_ND2	2342	20.7	ARG_NE	23775	17.0
SER_OG	5508	24.6	ARG_NE	2328	22.8	SER_OG	22962	19.4
ASN_ND2	5057	26.9	ARG_CD	2292	25.0	SER_CB	20718	21.6
ARG_CD	4961	29.1	LYS_CE	2239	27.0	THR_CG2	20579	23.7
GLN_NE2	4378	31.1	PHE_CE1	2183	29.1	GLY_CA	20486	25.9
GLN_OE1	4266	33.0	GLY_CA	1919	30.9	LYS_CE	20382	28.0
ASN_OD1	4131	34.8	TYR_OH	1918	32.7	THR_OG1	20331	30.2
THR_OG1	4100	36.7	ARG_CZ	1819	34.3	LYS_CD	19627	32.2
THR_CG2	3694	38.3	LEU_CD1	1720	35.9	ARG_CZ	17349	34.0
LYS_CE	3661	40.0	PHE_CE2	1640	37.5	ARG_CG	17285	35.8
ALA_CB	3542	41.6	VAL_CG1	1486	38.9	LYS_CB	15656	37.5
SER_CB	3426	43.1	LEU_CD2	1480	40.2	ALA_CB	15401	39.1
TYR_OH	3076	44.5	VAL_CG2	1391	41.5	ARG_CB	15005	40.7
GLY_CA	2891	45.8	ARG_CG	1366	42.8	LYS_CG	14285	42.2
ARG_CG	2826	47.1	ALA_CB	1353	44.1	ASN_ND2	14250	43.7
LYS_CD	2756	48.3	LYS_CD	1314	45.3	GLY_N	13934	45.1
GLU_OE2	2201	49.3	ILE_CD1	1280	46.5	TYR_OH	13928	46.6
GLU_OE1	2188	50.3	GLN_NE2	1259	47.6	GLY_O	13516	48.0
			TYR_CE2	1256	48.8	LYS_O	10699	49.1
			PRO_CG	1251	50.0	GLN_NE2	10250	50.2

For each type of atoms, contacts totaling over 50 % of the contacts of the type are shown.

цепи также образуют контакты с ДНК, однако в меньшем количестве (табл. 1).

Частота контактов атомов в большой, малой бороздках и с сахарофосфатным остовом

Мы полагаем, что все контакты можно разделить на два класса: специфические контакты, обеспечивающие ключевые взаимодействия, и неспецифические, образующиеся вследствие пространственного сближения атомов, расположенных вблизи от атомов, участвующих в специфических контактах. Например, представим себе водородную (или ионную) связь между атомом белка и атомом ДНК в сайте взаимодействия. Разбиение Вороного–Делоне «увидит» в такой ситуации контакт между этими двумя атомами. В то же время разбиение Вороного–Делоне выявит много других контактов, образованных соседними по структуре атомами белка и ДНК вследствие их сближения в пространстве. Пусть эта водородная (ионная) связь будет взаимодействием «специфическим», участвующим именно в распознавании сайта ДНК, тогда разбиением Вороного–Делоне мы получаем смесь «специфических» и «неспецифических» контактов. В случае применения метода определения контактов между атомами по критерию расстояния ситуация еще сильнее ухудшается, поскольку

дополнительно возникает неопределенность, вызванная неточностью и сильной зависимостью результатов от выбора величины критического расстояния.

Рассмотрим контакты атомов аминокислоты аргинина, а именно контакты этой аминокислоты с атомами нуклеиновых оснований, выходящих в малую бороздку (табл. 2). Из табл. 2 видно, что большая часть контактов (65.6 %) образуется с участием атомов азота гуанидиновой группы.

Таким образом, можно считать контакты атомов азота NH1 и NH2 гуанидиновой группы наиболее значимыми, а контакты других атомов аргинина вторичными, образующимися вследствие пространственного сближения атомов азота NH1 и NH2 гуанидиновой группы со своими атомами-мишенями.

Рассматривая контакты каждой аминокислоты, можно выделить для нее наиболее важные атомы, образующие много контактов. Таким же образом, выделяя наиболее часто контактирующие атомы, можно рассматривать атомы ДНК. Результат такого двойного отбора приведен в Приложении.

Кроме того, мы отобрали наиболее часто контактирующие атомы белка в каждой из трех рассматриваемых категорий контактов (табл. 3). Для всех категорий получилось, что наиболее часто контактируют атомы азота

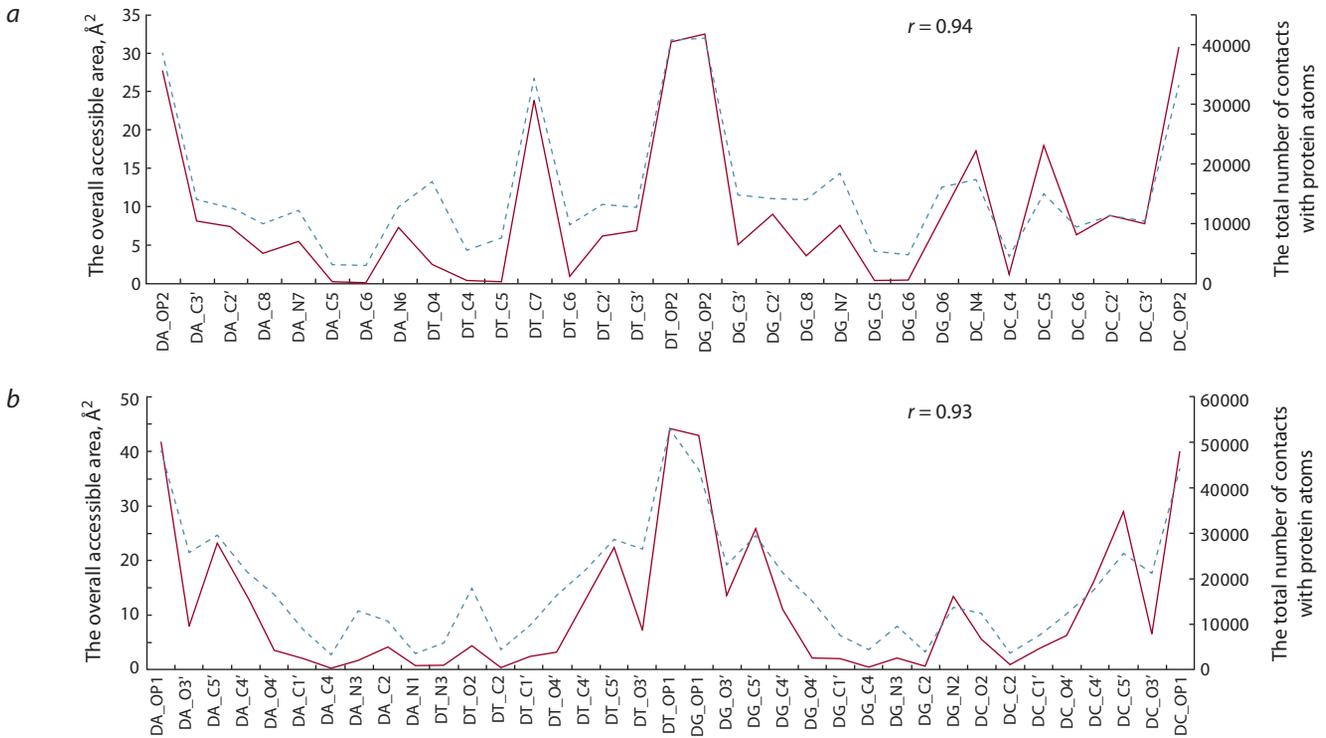


Fig. 3. The total number of contacts and the accessible surface area of the (a) major and (b) minor groove atoms. The solid line shows the area of the accessible surface, and the dotted line shows the number of contacts formed by the atoms. The DNA bases facing the major groove are arranged in accordance with their order in the spatial structure.

гуанидиновой группы аргинина и аминогруппы лизина, также следует отметить атомы амидной группы аспарагина и глутамин. В большой бороздке часто контактируют гидроксильные группы серина, треонина, тирозина и атомы карбоксильной группы глутаминовой кислоты.

В малой бороздке часто контактирует гидроксильная группа тирозина, наблюдается также большое число контактов с гидрофобными атомами фенилаланина, лейцина, валина, изолейцина, аланина, глицина. С сахарофосфатным остовом часто контактируют атомы аргинина, лизина, серина, треонина, глицина, глутамин, причем как атомы полярных групп, так и гидрофобные (см. табл. 3).

Сравнение частоты контактов с площадью доступной поверхности для атомов ДНК большой и малой бороздок

Мы проанализировали, насколько соответствует частота контактов с атомами ДНК, выходящими в большую и малую бороздку, доступности этих атомов растворителю (рис. 3).

Сравнение показало, что, несмотря на очень хорошее соответствие количества контактов и площади доступной поверхности (коэффициенты корреляции 0.94 и 0.93 для атомов большой и малой бороздок соответственно), некоторые атомы образуют больше контактов, чем можно было бы ожидать из площади доступной поверхности. Это атомы кислорода и азота DA_N7, DT_O4, DG_N7, DG_O6, выходящие в большую бороздку, и атомы DA_N3, DT_O2, DG_N3, DC_O2, выходящие в малую бороздку. Для таких атомов превышение может быть более чем двукратным

(см. рис. 3). Атомы сахарофосфатного остова и углероды краев оснований, выходящих в бороздки, не проявляют подобной тенденции. Естественно предположить, что наблюдаемая картина связана с возможностью или невозможностью образования водородной связи.

Конформационно-стабильные пептиды в комплексах белок-ДНК

Анализ расположения конформационно-стабильных пептидов в структуре ДНК-связывающих белков показал, что такие пептиды встречаются в разных местах белка, а также в участках взаимодействия между белковыми субъединицами в составе белкового комплекса и в области взаимодействия с ДНК. На рис. 4 приведен пример разметки конформационно-стабильных пептидов в структуре комплекса белок-ДНК (1a02).

Вопрос расположения конформационно-стабильных пептидов в области взаимодействия с ДНК и их роли в распознавании ДНК требует более подробного дальнейшего анализа.

Взаимодействия остатков в предопределенной конформации с ДНК

В изученных комплексах белок-ДНК было выявлено 159 аминокислотных остатков в предопределенной конформации β-изгиба типа I, 15 остатков в конформации β-изгиба типа I', 6 остатков в конформации β-изгиба типа II. Остатков в конформации β-изгиба типа II' найдено не было. Всего был обнаружен 91 межатомный контакт между аминокислотами в предопределенной конформа-



Fig. 4. Structure of the complex of DNA-binding domains of proteins NFAT (blue), FOS (pink) and JUN (red) with DNA (green and dark-green) (Protein Data Bank accession 1a02). Conformationally stable peptides are shown in yellow.



Fig. 5. The T120 residue in a predefined conformation of the type I β turn forms contacts with the nucleotide bases of the AT pair in the 3qea structure.

ции и атомами ДНК: 72 типа I и 19 типа II. Почти все эти контакты образованы с атомами сахарофосфатного остова ДНК. Контакты с атомами нуклеиновых оснований найдены только в двух гомологичных структурах 3qea и 3qe9 (34 контакта). В этих структурах остаток треонина в предопределенной конформации I типа взаимодействует с атомами нуклеиновых оснований АТ пары (рис. 5).

Таким образом, аминокислотные остатки в предопределенной конформации основной цепи могут принимать участие во взаимодействии с ДНК, однако частота таких событий достаточно мала.

Обсуждение

Можно надеяться, что выяснение принципов специфического распознавания сайтов ДНК в перспективе откроет путь к управлению важнейшими процессами в живых системах. Очевидным методическим подходом для исследований в рамках этой проблемы является анализ экспериментально определенных структур белково-нуклеиновых комплексов. Ранее нами была создана программа для построения пространственного разбиения Вороного–Делоне на системе точек, соответствующих расположению атомов в комплексах белок-ДНК (Anashkina et al., 2007). В рамках этого однозначного математического построения контактирующие атомы определяются как соседние вершины разбиения Делоне или как атомы, имеющие общие грани полиэдров Вороного ненулевой площади. Во всех работах других исследователей определение контактирующих атомов проводится на основе критерия расстояния. Однако, как показано нами ранее, это ведет к систематическим ошибкам: недоопределению числа контактов при выборе недостаточно большого критического расстояния и переопределению числа контактов при выборе слишком большого критического расстояния. В последнем случае фактически между выделенными контактирующими атомами может находиться один атом или более. Особенно трудно контролировать вхождение молекулы воды в такой искусственно увеличенный контакт.

В последнее время основное внимание исследователей обращено на изучение отдельных семейств распознающих ДНК белков. Нашей задачей было включить в рассмотрение всю совокупность имеющихся данных с целью получения характеристик, общих для разных распознающих схем, причем суммирование данных сопровождалось дифференциацией по бороздкам. Аналогичная работа была проделана в свое время Торнтон с соавторами (Luscombe et al., 2001), но статистика взаимодействий отдельно по бороздкам ими не рассматривалась. В цитируемой работе было исследовано 129 комплексов белок-ДНК, в то время как в нашей работе – 3518 комплексов. Для изучения были отобраны все имеющиеся в PDB на весну 2017 г. комплексы белок-ДНК. Мы проанализировали контакты между атомами белка и ДНК в структурах белок-ДНК комплексов методом Вороного–Делоне, что позволило исключить не прямые взаимодействия, которые обычно попадают в поле зрения при использовании критерия расстояния, тем самым «зашумляя» картину. В предположении, что существуют специфические контакты, управляющие распознаванием, и неспецифические, образующиеся вследствие сближения атомов в пространстве, мы выделили наиболее часто контактирующие атомы белка и ДНК. Так, например, контакты всего лишь трех атомов азота гуанидиновой группы аргинина образуют с сахарофосфатным остовом 10.9 %, с малой бороздкой 15.1 % и с большой бороздкой 16.2 % контактов от общего числа контактов в каждой категории соответственно (см. табл. 3).

Наши вычисления показали, что три десятка типов атомов из 13 аминокислот образуют более 50 % контактов во всех трех категориях контактов: по большой бороздке, по малой бороздке и с сахарофосфатным остовом (см. табл. 3). Мы пришли к выводу, что число контактов атомов нуклеотидов с атомами белка хорошо коррелирует с площадью доступной поверхности (коэффициент линейной корреляции по большой бороздке 0.94, по малой – 0.93). Однако для атомов кислорода и азота DA_N7, DT_O4, DG_N7, DG_O6, выходящих в большую бороздку, и

атомов DA_N3, DT_O2, DG_N3, DC_O2, выходящих в малую бороздку, показано, что они контактируют чаще, чем можно ожидать исходя из площади доступной поверхности. Для таких атомов превышение может быть более чем двукратным (см. рис. 3). Это атомы кислорода и азота нуклеиновых оснований, выходящие в малую и большую бороздку и способные к формированию водородных связей, будучи донорами и/или акцепторами таких связей.

Что касается роли конформационно-стабильных пептидов в структуре ДНК-связывающих белков, то по нашим наблюдениям такие пептиды достаточно часто встречаются в разных местах белка, а также в участках взаимодействия между белковыми субъединицами в составе белкового комплекса и в области взаимодействия с ДНК, что требует более подробного анализа.

Анализ аминокислотных остатков в предопределенной конформации основной цепи показал, что эти остатки могут принимать участие во взаимодействии с ДНК, однако частота таких событий достаточно мала. Полученные в работе данные могут оказаться полезными, в частности, для проектирования рекомбинантных белков, распознающих последовательность ДНК.

Acknowledgments

This study was supported by the Russian Foundation for Basic Research, projects 15-04-99605a, 17-04-02105a, and 16-54-00219Bel-a, and the Belarus Republican Foundation for Basic Research, project B16R178.

Conflict of interest

The authors declare no conflict of interest.

References

- Anashkina A.A., Kuznetsov E.N., Batyanovskii A.V., Gnuchev N.V., Tumanyan V.G., Esipova N.G. Classification of amino acids based on comparative analysis of contacts in DNA-protein complexes and specific DNA-protein interactions. *Biophysics*. 2013;58(6):766-770.
- Anashkina A.A., Kuznetsov E.N., Esipova N.G., Tumanyan V.G. Comprehensive statistical analysis of residues interaction specificity at protein-protein interfaces. *Proteins: Struct. Funct. Bioinf*. 2007; 67(4):1060-1077. DOI 10.1002/prot.21363.
- Anashkina A.A., Tumanyan V.G., Kuznetsov E.N., Galkin A.V., Esipova N.G. Relative occurrence of amino acid-nucleotide

- contacts assessed by Voronoi-Delaunay tessellation of protein-DNA interfaces. *Biophysics*. 2008;53(3):199-201. DOI 10.1134/S0006350908030032.
- Batyanovskii A.V., Esipova N.G., Shnoll S.E. Mutual disposition of short conformationally stanch oligopeptides in the 3D structure of globular proteins. *Biophysics*. 2009;54(6):748-752. DOI 10.1134/S0006350909060153.
- Batyanovskii A.V., Vlasov P.K. Short protein segments with prevalent conformation. *Biophysics*. 2008;53(4):264-267. DOI 10.1134/S0006350908040040.
- Benos P.V., Lapedes A.S., Stormo G.D. Is there a code for protein-DNA recognition? *Probab(istical)ly... BioEssays*. 2002;24(5):466-475. DOI 10.1002/bies.10073.
- Choo Y., Klug A. Physical basis of a protein-DNA recognition code. *Curr. Opin. Struct. Biol*. 1997;7(1):117-125. DOI 10.1016/S0959-440X(97)80015-2.
- Corona R.I., Guo J.-T. Statistical analysis of structural determinants for protein-DNA-binding specificity. *Proteins*. 2016;84(8):1147-1161. DOI 10.1002/prot.25061.
- Filippova G.N., Qi C.-F., Ulmer J.E., Moore J.M., Ward M.D., Hu Y.J., Loukinov D.I., Pugacheva E.M., Klenova E.M., Grundy P.E., Feinberg A.P., Cleton-Jansen A.M., Moerland E.W., Cornelisse C.J., Suzuki H., Komiya A., Lindblom A., Dorion-Bonnet F., Neiman P.E., Morse H.C., Collins S.J., Lobanenko V.V. Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alter its DNA-binding specificity. *Cancer Res*. 2002;62(1):48-52.
- Luscombe N.M., Laskowski R.A., Thornton J.M. Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res*. 2001;29(13):2860-2874. DOI 10.1093/nar/29.13.2860.
- Medvedev N.N. The algorithm for three-dimensional Voronoi polyhedra. *J. Comput. Physics*. 1986;67(1):223-229. DOI 10.1016/0021-9991(86)90123-3.
- Pabo C.O., Nekludova L. Geometric analysis and comparison of protein-DNA interfaces: Why is there no simple code for recognition? *J. Mol. Biol*. 2000;301(3):597-624. DOI 10.1006/jmbi.2000.3918.
- Privalov P.L. The state and role of water in biological systems. *Biofizika = Biophysics*. 1958;3(6):738-743. (in Russian)
- Torshin I.Yu., Uroshlev L.A., Esipova N.G., Tumanyan V.G. Descriptive statistics of disallowed regions and various protein secondary structures in the context of studying twisted β -hairpins. *Biophysics*. 2016;61(1):6-12. DOI 10.1134/S0006350916010243.
- Uroshlev L.A., Torshin I.Yu., Batyanovskii A.V., Esipova N.G., Tumanyan V.G. Disallowed conformations of a polypeptide chain as exemplified by the β -turn of the β -hairpin in the α -spectrin SH3 domain. *Biophysics*. 2015;60(1):1-9. DOI 10.1134/S0006350915010236.
- Wolfe S.A., Nekludova L., Pabo C.O. DNA recognition by Cys₂His₂ zinc finger proteins. *Annu. Rev. Biophys. Biomol. Struct*. 2000;29:183-212. DOI 10.1146/annurev.biophys.29.1.183.



Интернет-доступные информационные ресурсы по геномным сетям, включающие данные по человеку и животным

Е.В. Игнатьева^{1, 2}✉, Д.А. Афонников^{1, 2}, Н.А. Колчанов^{1, 2}

¹ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия

Геномные сети – это молекулярно-генетические системы, обеспечивающие формирование фенотипических характеристик организмов (молекулярных, биохимических, структурных, морфологических, поведенческих и т.д.) на основе информации, закодированной в их геномах. Реконструкция геномных сетей обеспечивает методическую основу современной системной биологии. Большую ценность представляет информация о структурно-функциональной организации геномных сетей, накопленная в современных базах данных. В настоящем обзоре представлена характеристика интернет-доступных информационных ресурсов, ориентированных на человека и животных и содержащих данные по геномным сетям и их функциональным модулям. Не претендуя на полноту охвата абсолютно всех информационных ресурсов, содержащих данные, относящиеся к человеку и животным по этой тематике, предложен обзор создан для того, чтобы оценить современное состояние проблемы, а также представить критерии, согласно которым целесообразно оценивать полезность информационных ресурсов для конкретных исследовательских задач. Исходя из этого нами, во-первых, была сформирована и охарактеризована подборка баз данных, содержащих сведения о метаболических и сигнальных путях, а также о путях регуляции биологических процессов на клеточном и организменном уровнях. Во-вторых, в качестве примера описаны несколько известных баз данных по межмолекулярным взаимодействиям различных типов. В обзоре рассматриваются следующие характеристики баз данных: 1) типы накопленной информации; 2) способы представления информации; 3) способы наполнения баз данных; 4) основные источники информации; 5) программные средства, позволяющие осуществлять поиск и анализ данных. Сопоставление перечисленных характеристик показало, что рассмотренные базы данных очень гетерогенны по тематике, источникам, типам и способу представления информации, а также по возможностям формировать запросы и анализировать данные. Делается вывод о том, что до начала реконструкции геномной сети определенного биологического процесса очень важно иметь представление о максимально полном наборе информационных источников, из которых может быть взята информация. Приведены примеры веб-порталов, аккумулирующих сведения о базах данных и информационных ресурсах, которые могут быть полезны для реконструкции и анализа геномных сетей.

Ключевые слова: системная биология; базы данных; геномные сети; Интернет.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Игнатьева Е.В., Афонников Д.А., Колчанов Н.А. Интернет-доступные информационные ресурсы по геномным сетям, включающие данные по человеку и животным. Вавиловский журнал генетики и селекции. 2017;21(8):895-902. DOI 10.18699/VJ17.310

HOW TO CITE THIS ARTICLE:

Ignatieva E.V., Afonnikov D.A., Kolchanov N.A. Online resources on gene networks containing human and animal data. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2017;21(8):895-902. DOI 10.18699/VJ17.310 (in Russian)

Received 15.09.2017

Accepted for publication 20.11.2017

© AUTHORS, 2017

✉ e-mail: eignat@bionet.nsc.ru

Online resources on gene networks containing human and animal data

E.V. Ignatieva^{1, 2}✉, D.A. Afonnikov^{1, 2},
N.A. Kolchanov^{1, 2}

¹ Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

² Novosibirsk State University, Novosibirsk, Russia

Gene networks are molecular genetic systems that ensure the formation of phenotypic characteristics of organisms (molecular, biochemical, structural, morphological, behavioral, etc.) based on information encoded in their genomes. Reconstruction of gene networks provides a methodological basis for modern systems biology. In this regard, the information on the structural and functional organization of gene networks accumulated in modern databases is extremely valuable. This review characterizes a number of Internet-accessible information resources oriented to humans and animals and containing data on gene networks and their functional modules. Without pretending to fully cover all information resources containing data related to humans and animals on the subject, the current review was created to report the current status of the problem and to present the criteria according to which we propose to evaluate the utility of web-resources for specific research tasks. On this basis, we compiled and characterized a collection of databases containing information on metabolic and signaling pathways, as well as pathways of regulation of biological processes at the cellular and organismal levels. In addition, we observed the characteristics of several well-known databases containing data on interactions between biomolecules of various types. The following characteristics of databases were considered: (1) the types of information accumulated in the databases; (2) methods of data presentation; (3) methods of data collection; (4) data sources; (5) special search tools and options for data analysis. A comparison of the above characteristics showed that the databases are very heterogeneous according to their scopes, sources and types of data, interfaces, as well as according to their search options and data analysis tools. It was concluded that at the first step of the gene network reconstruction it is important to form a full set of information resources from which the data can be obtained. The web portals accumulating information about the databases that may be useful for the reconstruction and analysis of gene networks are specified.

Key words: systems biology; databases; gene networks; Internet.

При изучении генетических механизмов регуляции молекулярно-биологических процессов в современной биологической науке широко используется понятие «генные сети» (в англоязычной литературе – gene regulatory networks). Современное определение понятия «генные сети» базируется на представлениях о молекулярно-генетических системах управления (МГСУ), сформулированных профессором В.А. Ратнером. Термин МГСУ применялся для информационно-кибернетического описания и моделирования совокупности универсальных (репликация, транскрипция, трансляция, репарация, рекомбинация, сегрегация) и неуниверсальных молекулярно-биологических процессов и реакций, протекающих в клетках живых организмов (Ратнер, 1966). Позднее С. Кауфманом был предложен сходный термин «генные сети» (первоначально звучавший как genetic control networks), который и закрепился в мировой литературе (Kauffman, 1969).

В дальнейшем под генными сетями мы будем подразумевать молекулярно-генетические системы, обеспечивающие формирование фенотипических характеристик организмов (молекулярных, биохимических, структурных, морфологических, поведенческих и т. д.) на основе информации, закодированной в их геномах. Генные сети включают группы координированно функционирующих генов, которые взаимодействуют друг с другом как через свои первичные продукты (РНК и белки), так и через разнообразные метаболиты и другие вторичные продукты функционирования генных сетей (Колчанов и др., 2000, 2013).

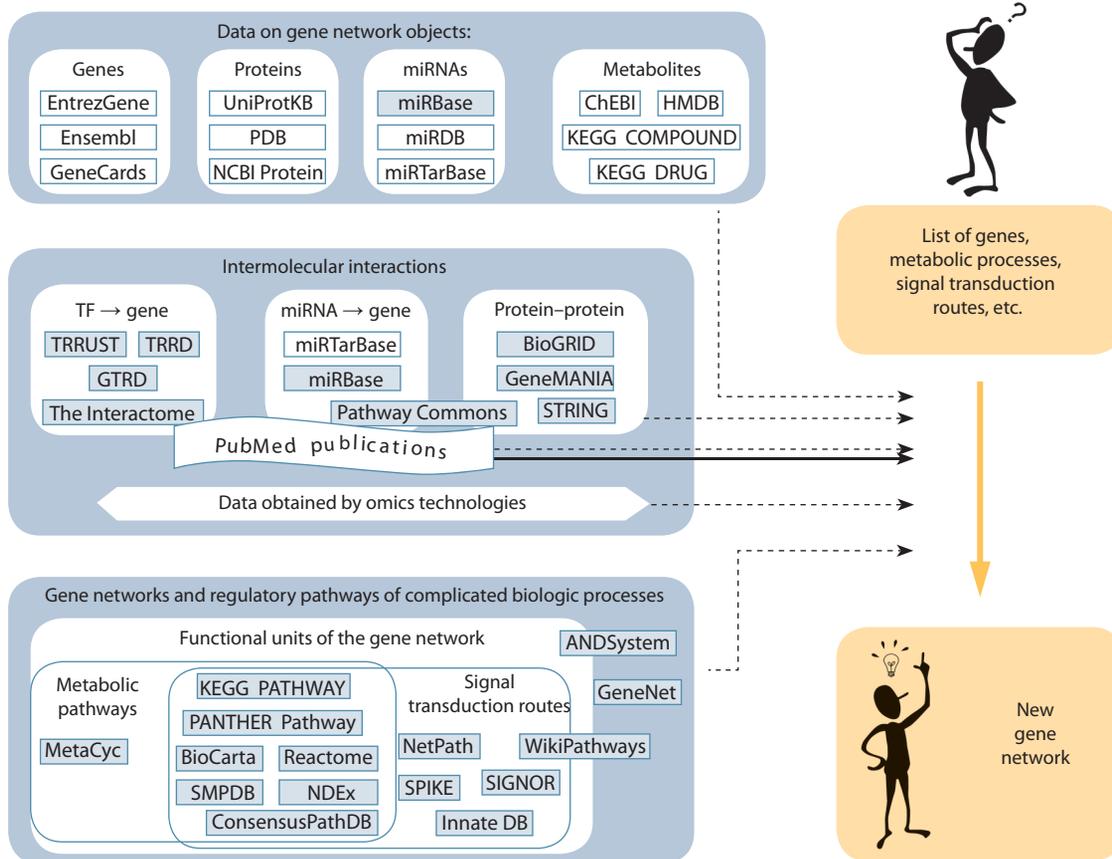
Концепция генных сетей постоянно развивается, обеспечивая тем самым методическую основу для решения все более широкого круга задач. Высокопроизводительные экспериментальные методы, применяемые в современных исследованиях, позволяют получать и анализировать все взаимодействия между определенными типами молекулярно-генетических объектов, идентифицированных в конкретном типе клеток (например, ДНК-белковые и белок-белковые взаимодействия). Массовое накопление подобных данных стало предпосылкой для выделения в рамках теории генных сетей таких сущностей, как: 1) сети транскрипционной регуляции (transcriptional regulatory networks) (Reese-Hoyes et al., 2005; Neph et al., 2012); 2) сети взаимодействий между миРНК и соответствующими генами-мишенями (miRNA regulatory networks) (Plaisier et al., 2012; Obermayer, Levine, 2014); а также 3) сети белок-белковых взаимодействий (protein-protein interaction (PPI) networks) (Lei et al., 2013; Ignatieva et al., 2016).

Рассмотрим в качестве примера сети транскрипционной регуляции (СТР). Эти сети отображают регуляторные взаимодействия типа «транскрипционный фактор → регулируемый ген-мишень». Каждая вершина СТР представляет сразу две сущности: ген и кодируемый им белок. Данные о регуляторных взаимодействиях в СТР могут быть получены на основе различных экспериментально-теоретических подходов, например: 1) компьютерного анализа регуляторных областей генов, выявленных методикой ДНКазы I футпринтинга (Neph et al., 2012); 2) сопоставления экспрессионных данных и данных, полученных методикой иммунопреципитации

хроматина с последующим секвенированием (ChIP-seq) (McMullen et al., 2014); 3) исследования эффектов выключения генов, кодирующих транскрипционные факторы, с помощью siRNA (Tomaru et al., 2014; Li et al., 2016). Как и все высокопроизводительные методы анализа, иммунопреципитация хроматина и ДНКазы I футпринтинг не позволяют получать данные со стопроцентным уровнем точности. Кроме того, компьютерные методы анализа, которые, как правило, применяют в таких исследованиях, характеризуются определенной долей ложноположительных и ложноотрицательных результатов (Kolchanov et al., 2007; Levitsky et al., 2007). Таким образом, регуляторные взаимодействия, выявленные с помощью описанных выше методов (а также ряда других высокопроизводительных методов анализа), должны быть использованы при реконструкции генных сетей с определенной осторожностью.

В настоящее время активно исследуются также ассоциативные генные сети, содержащие расширенный набор типов вершин (включая биологические процессы, заболевания, фенотипические характеристики, мутации, терапевтические воздействия). Такой подход позволяет объединять разрозненные локальные генные сети, не пересекающиеся между собой по молекулярно-генетическим объектам, и выявлять добавочные объекты (фенотипический признак, заболевание и т. д.), связывающие локальные генные сети (Glotov et al., 2015; Ivanisenko et al., 2015).

Реконструкция генных сетей – неотъемлемый этап исследования сложных молекулярно-генетических систем и процессов. При этом сначала конкретизируется объект исследования (фенотипический признак либо патологический процесс) и формируются списки элементарных биологических процессов, а также объектов (генов, белков, миРНК, метаболитов), вовлеченных в генную сеть. Данные об объектах генной сети доступны в базах данных EntrezGene, Ensembl, UniProtKB, miRBase, KEGG COMPOUND и др. (см. рисунок). Далее необходимо собрать информацию о взаимодействиях между объектами. Для этого можно обращаться как к разрозненным публикациям из PubMed, так и к специализированным информационным ресурсам, содержащим данные по различным типам межмолекулярных взаимодействий (BioGRID, GeneMANIA, miRBase, TRRUST, Pathway Commons и т. д.). Кроме того, данные о взаимодействиях между объектами (регуляторные, образование комплексов и т. д.) могут быть предсказаны на основе компьютерного анализа данных, полученных с помощью омиксных технологий (транскриптомных, протеомных, метаболомных и т. д.). При реконструкции генных сетей очень полезно использовать также сведения о функциональных модулях генных сетей (метаболических путях (МП), путях передачи сигналов (ППС) и др.), а также о путях регуляции сложных биологических процессов на клеточном и организменном уровне (ПРБП), которые представлены в современных базах данных. Наиболее крупными базами подобного рода являются KEGG PATHWAY, Reactome, MetaCyc, WikiPathways. Объем, а также разнообразие типов данных, доступных в интернет-ресурсах, чрезвычайно велики, в связи с чем планирование работ по реконструкции генных сетей обязательно должно включать этап поиска и систематизации сведений об источниках данных.



Employment of sources of information in gene network reconstruction.

Databases are indicated in bars. Shaded bars indicate databases reviewed in this paper. The solid arrow denotes manual annotation of scientific publications, and dashed arrows, automated annotation or extraction from databases.

Сведения об интернет-доступных информационных ресурсах по тематике «геномные сети» можно почерпнуть из современных веб-порталов. Например, PathGuide (<http://www.pathguide.org/>) отображает список из 690 баз, расклассифицированных по нескольким категориям, включая: 1) белок-белковые взаимодействия; 2) метаболические пути; 3) сигнальные пути; 4) сети транскрипционных взаимодействий. Другим очень полезным ресурсом является веб-коллекция баз данных по молекулярной биологии, размещенная на сайте журнала Nucleic Acids Research (NAR Molecular Biology Database Collection, <http://www.oxfordjournals.org/nar/database/c/>). Согласно этому ресурсу, в 2017 г. общий список баз данных по всем видам организмов, имеющих отношение к тематике «геномные сети», включал 143 наименования (36 из категории «метаболические пути», 93 из категории «белок-белковые взаимодействия» и 19 из категории «сигнальные пути») (Galperin et al., 2017).

В нашем обзоре охарактеризованы базы данных из вышеперечисленных тематических разделов веб-коллекции журнала NAR, ориентированные (в числе прочих видов) на человека и животных и находящиеся в открытом доступе в сети Интернет в 2017 г. Нами сформирована и охарактеризована подборка баз данных, содержащих сведения о метаболических и сигнальных путях, а также

о путях регуляции биологических процессов на клеточном и организменном уровнях (табл. 1 и Приложение 2¹). Чтобы познакомить читателя с предметной областью, мы представили характеристики ряда наиболее известных и значимых баз данных по межмолекулярным взаимодействиям. Таким образом, не претендуя на полноту охвата абсолютно всех баз данных, включающих сведения по геномным сетям человека и животных, представленный обзор знакомит читателя с современным состоянием проблемы, а также рассматривает критерии, согласно которым целесообразно оценивать полезность информационных ресурсов для конкретных исследовательских задач.

Типы информации и способы ее представления в базах данных

Содержащаяся в базах данных информация может быть отнесена к следующим иерархическим уровням: 1) пути регуляции сложных биологических процессов (включая заболевания и патологии), протекающих на уровне клетки, ткани, органа, а также целого организма; 2) функциональные модули геномных сетей – метаболические или сигнальные пути; 3) межмолекулярные взаимодействия. Информация, соответствующая первым двум иерархиче-

¹ Приложения 1–4 см. по адресу: <http://www.bionet.nsc.ru/vogis/download/pict-2017-21/appx13.pdf>

Table 1. Online databases on gene networks, metabolic and signaling pathways, pathways regulating other biologic processes, and molecular interactions in the alphabetical order

Base name	Filling method or data source	Data type	Reference
ANDSystem	DB, TM	PPI, RI, coexpression, involvement of proteins in biologic processes according to UniProt-GOA, use of the substance in disease treatment, gene–disease associations	http://www-bionet.sccc.ru/andvisio/
BiGG Models	M	MP	http://bigg.ucsd.edu/
BioCarta	M	MP, STR, BPRP, diseases	https://cgap.nci.nih.gov/Pathways/BioCarta_Pathways
BioCyc	M, DB, CP	MP	https://biocyc.org/
BioGRID	M	PPI, GI	https://thebiogrid.org/
ConsensusPathDB	M, DB	MP, STR, BPRP, PPI, RI, GI	http://cpdb.molgen.mpg.de/
GeneMANIA	TM, DB, CP	PPI, GI, coexpression, relationships between genes and proteins predicted by computerized methods	http://genemania.org/
GeneNet	M	Structure–functional organization of gene networks, STR, MP	http://wwwmgs.bionet.nsc.ru/mgs/gnw/genenet/
GTRD	CAD	DNA–protein interactions	http://gtrd.biouml.org
InnateDB	M, DB	STR	http://www.innatedb.ca/
The Interactome	CAD	Tissue–specific TRN	http://www.regulatorynetworks.org/
KEGG PATHWAY	M	MP, STR, BPRP, diseases, medicines	http://www.genome.ad.jp/kegg
miRBase	CAD, CP, M	RI	http://www.mirbase.org/
MetaCyc	M	MP	http://metacyc.org/
NDEX	DB, M	MP, STR, BPRP, PPI, RI	http://www.ndexbio.org/#/
NetPath	M	Human STR	http://www.netpath.org/
PANTHER Pathway	M, DB	MP, STR	http://pantherdb.org/
Pathway Commons	DB	PPI, RI	http://www.pathwaycommons.org/
Reactome	M	MP, STR, routes of molecule transport in the cell and DNA replication	http://www.reactome.org/
SIGNOR	M	STR and regulatory interactions between their components, BPRP, diseases	http://signor.uniroma2.it/
SMPDB	M	MP, STR, BPRP, diseases	http://smpdb.ca/
SPIKE	M, DB	Human STR	http://www.cs.tau.ac.il/~spike/
STRING	M, TM, DB, CP	PPI, coexpression, computer-predicted relationships between genes and proteins	string-db.org/
TRED	M, DB, CP	TRN, RI	http://rulai.cshl.edu/TRED
TRRD	M	RI, DNA–protein interactions	http://wwwmgs.bionet.nsc.ru/mgs/gnw/trrd/
TRRUST	TM, M	TRN, RI	http://www.grnpedia.org/trrust
WikiPathways	M	MP, STR, BPRP, diseases	https://www.wikipathways.org/index.php/WikiPathways

Filling methods: M, manual annotation of scientific publications; TM, text mining; DB, import from other databases; CP, computerized predictions; CAD, computerized analysis of data obtained by high-throughput experimental methods. Data types: PPI, protein–protein interactions; GI, genetic interactions; MP, metabolic pathways; STR, signal transduction routes; BPRP, biologic process regulation pathways at the cellular and organismal levels; RI, regulatory interactions (transcription factor–regulated gene or miRNA–regulated gene); TRN, transcription regulation networks. Appendix 2 provides more comprehensive database descriptions with references to recent papers.

ским уровням, как правило, представлена в базах данных в виде диаграмм. Данные, соответствующие третьему иерархическому уровню (межмолекулярные взаимодействия), отображаются в виде сетей взаимодействий между объектами.

Диаграммы. В таких базах, как KEGG PATHWAY, MetaCyc, WikiPathways, Reactome, SIGNOR, GeneNet и др., основными единицами информации являются диаграммы – статические изображения, содержащие фиксированный набор объектов (генов, белков, метаболитов и т. д.) и отображающие метаболические либо сигнальные пути, а также схемы регуляции сложных биологических процессов. Объекты диаграмм, а также узлы реакций и регуляторных процессов, как правило, интерактивны, обеспечивая переход к их текстовым описаниям.

В каждой базе данных используется специально разработанный графический язык представления информации. В качестве примера описан графический язык базы KEGG PATHWAY (Приложение 1).

Известны базы данных, диаграммы которых отображают только сигнальные пути (SIGNOR, SPIKE, InnateDB) (см. аббревиатуру ППС в табл. 1). Другие базы (KEGG PATHWAY, WikiPathways, GeneNet, BioCarta, PANTHER Pathway) представляют на диаграммах как метаболические, так и сигнальные пути, а также пути регуляции других процессов (МП, ППС, ПРБП в табл. 1). Базы KEGG PATHWAY, BioCarta, WikiPathways содержат также диаграммы, отображающие механизмы развития патологических процессов.

Помимо того, что диаграммы являются единицами информации в базах данных, они выполняют еще одну очень важную функцию. Диаграммы создаются аннотаторами-биологами и отражают их представления о регуляции биологических процессов. Эти представления формировались в ходе многолетних исследований. Таким образом, наряду с функциональной аннотацией генов терминами из словаря Gene Ontology, диаграммы биологических процессов служат еще одним курируемым источником данных о функциональной роли гена.

Сети взаимодействий. Данные о парных либо множественных межмолекулярных взаимодействиях (регуляторных, белок-белковых, формировании комплексов, биохимических реакциях и т. д.) очень важны для реконструкции генных сетей. Известен целый ряд баз данных (STRING, GeneMANIA, Pathway Commons, The Interactome, TRRUST и др.), накапливающих сведения о межмолекулярных взаимодействиях (табл. 2). Базы этой категории не содержат готовых статических диаграмм. Сети взаимодействий между заданным набором объектов (генов или белков) генерируются в этих базах на основании запроса (списка генов/белков, заданного пользователем). Результат запроса отображается в графическом виде, в виде интерактивной сети, включающей объекты (гены, белки и т. д.) и связи различных типов – ассоциации. В каждом случае имеется возможность получить данные об ассоциациях в текстовом виде. В качестве запроса может быть задан список генов/белков (STRING, GeneMANIA, Interactome, ConsensusPathDB и т. д.) или только один ген/белок (BioGRID, TRRUST). Основные типы межмолекулярных взаимодействий/ассоциаций

между объектами генных сетей, накопленные в соответствующих информационных ресурсах, представлены в табл. 2. Большинство из рассмотренных ресурсов содержат данные об ассоциациях различных типов. Отдельную группу составляют базы The Interactome, TRRUST, TRRD, TRED и GTRD, содержащие только один тип ассоциаций – регуляторные взаимодействия типа «транскрипционный фактор → регулируемый ген».

Способы наполнения баз данных и источники данных

Данные могут быть внесены в базу путем ручной аннотации научных публикаций, автоматической аннотации различных источников данных (текстов статей, результатов компьютерного анализа, данных экспериментов и т. д.), а также экстракции данных из других ресурсов и их интеграции в едином формате.

Ручная аннотация научных публикаций использовалась при создании баз данных KEGG PATHWAY, MetaCyc, WikiPathways, Reactome, SIGNOR, GeneNet, TRRD, а также некоторых других (см. табл. 1, базы с пометкой «Р»). Ручная аннотация – трудоемкий процесс, однако он обеспечивает более высокое качество информации по сравнению с данными, внесенными с помощью автоматической аннотации. Именно поэтому такие базы данных очень востребованы. Источниками информации могут быть как научные публикации, описывающие результаты экспериментов с отдельными генами и белками (основная масса данных в базах KEGG PATHWAY, MetaCyc, WikiPathways, SIGNOR), так и публикации, представляющие результаты современных высокопроизводительных экспериментов (подобные сведения накапливаются в BioGRID и ConsensusPathDB) и их анализа (накоплены в BiGG Models). Команды аннотаторов многих баз достаточно многочисленны, например, в наполнении базы WikiPathways участвовало более 400 экспертов из разных стран (Kutmon et al., 2016).

Автоматическая аннотация была использована (наряду с другими методами) при создании таких информационных ресурсов, как STRING, GeneMANIA, The Interactome, TRED и др. Этот способ позволяет получать данные о парных взаимодействиях (ассоциациях) между объектами (генами, белками и т. д.). Источниками данных при таком способе наполнения могут быть: 1) сведения, извлеченные из рефератов и текстов научных публикаций методами текст-майнинга (см. табл. 1, базы AAT); 2) ассоциации, выявленные (предсказанные) на основе компьютерного анализа геномов и протеомов (см. табл. 1, базы КП); 3) результаты компьютерного анализа данных широкомасштабных геномных, протеомных и транскриптомных исследований (см. табл. 1, базы КАОД).

Интеграция информации из интернет-доступных информационных источников является еще одним возможным способом организации новых информационных ресурсов. Таким путем создана база Pathway Commons (см. табл. 1, обозначение «БД»).

Кроме того, при создании информационных ресурсов широко применяется комбинация трех вышеперечисленных подходов (PANTHER Pathway, STRING, GeneMANIA, ConsensusPathDB, TRED, TRRUST, ANDSysystem).

Table 2. Online resources storing data on intermolecular interactions between gene network objects/associations and their interaction types

Resource	Type of intermolecular interactions									
	Protein-protein	Cooccurrence in publications	Protein similarity	Coexpression	Location in the genome	Regulatory TF → target genes	miRNA → target genes	Genetic interactions*	Gene-disease association	Drug-target molecule or drug-disease association
ANDSystem	✓	✓	×	✓	×	✓	✓	×	✓	✓
BioGRID	✓	×	×	×	×	×	✓	✓	×	×
ConsensusPathDB	✓	×	×	×	×	✓	×	✓	×	✓
GeneMANIA	✓	×	✓	✓	×	×	×	✓	×	×
GTRD	×	×	×	×	×	✓	×	×	×	×
miRBase	×	×	×	×	×	×	✓	×	×	×
The Interactome	×	×	×	×	×	✓	×	×	×	×
Pathway Commons	✓	×	×	×	×	✓	×	×	×	✓
STRING	✓	✓	✓	✓	✓	×	×	×	×	×
TRED	×	×	×	×	×	✓	×	×	×	×
TRRD	×	×	×	×	×	✓	×	×	×	×
TRUST	×	×	×	×	×	✓	×	×	×	×

* By genetic interactions is implied that disturbances in one gene affect the function of the other gene.

** According to UniProt-GOA.

Программные инструменты баз данных

Программные средства баз данных обеспечивают поиск данных, их анализ, а также экспорт в различных форматах.

Практически все базы снабжены специальными поисковыми системами, позволяющими проводить поиск по названию объекта (гена, белка, метаболита и др.). В базах KEGG PATHWAY, Reactome, WikiPathways и др. результатом поиска являются диаграммы биологических процессов. В базах по межмолекулярным взаимодействиям (STRING, GeneMANIA, BioGRID, The Interactome, ANDSysSystem) на основе такого запроса (как уже отмечено выше) выполняется реконструкция сети взаимодействий, включающая заданный объект (или объекты, если их было несколько в запросе). В информационных системах STRING, GeneMANIA и ANDSysSystem есть возможность включить в сеть добавочные объекты (гены/белки), имеющие максимальное количество связей с объектами из сети. Таким образом, можно выявить наиболее короткий путь, связывающий два объекта, если изначальный результат запроса не содержал прямой связи между ними. В системе ConsensusPathDB имеется специальная опция «найти самый короткий путь».

Экспорт данных из баз можно осуществить: 1) в виде графических файлов, содержащих изображения диаграмм либо сетей взаимодействий между объектами (форматы png, pdf и др.); 2) в виде текстовых файлов (таблиц), содержащих списки объектов или бинарных взаимодействий между ними; 3) в специальных форматах (BioPAX, PSI-MITAB, SBML, SBGN, gpml, owl, pwl и т. п.), представляющих описание объектов и связей на диаграмме. Более подробно возможности экспорта данных описаны в Приложении 2. У некоторых баз (KEGG PATHWAY, Reactome, ConsensusPathDB, PANTHER Pathway, NDEx, BiGG Models) имеются программные интерфейсы APIs (application programming interfaces), позволяющие формировать и выполнять запросы ко всей информации, содержащейся в базах.

Наиболее распространенный и востребованный вариант анализа данных, предоставляемый веб-сайтами информационных ресурсов, – это анализ, проводимый с целью функциональной аннотации списка генов. В англоязычной литературе он именуется как «pathway enrichment analysis», а также «over-representation analysis, ORA» (Khatgi et al., 2012). Входными данными является список генов. В результате анализа пользователь получает сведения о диаграммах базы, в которых присутствуют гены из исходного списка, а также статистическую оценку повышенной (пониженной) представленности генов из исходного набора в диаграммах. Этот вид анализа широко используется для интерпретации функций групп дифференциально экспрессирующихся генов, полученных на основе микрочиповых технологий, а также для интерпретации данных транскриптомных и протеомных исследований. Программные средства для функциональной аннотации наборов генов имеются у баз PANTHER Pathway, ConsensusPathDB, Reactome, SPIKE, InnateDB и некоторых других.

Известны и другие программные инструменты, предоставляемые отдельными базами. Например, в базе KEGG PATHWAY есть возможность выделить на диаграмме объ-

екты из списка, введенного пользователем. В базах WikiPathways, SPIKE, NDEx имеются программные средства, позволяющие создавать диаграммы. Анализ информации из баз ConsensusPathDB, GeneMANIA, NDEx, InnateDB (и ряда других) можно осуществлять через специальные приложения (плагины) системы Cytoscape.

Заключение

Реконструкция генных сетей обеспечивает методическую основу для исследования закономерностей организации и механизмов генетического контроля функционирования живых систем (Neph et al., 2012; Stergachis et al., 2014), их эволюции (Mustafin et al., 2017), роли генов в развитии патологий (Ivanisenko et al., 2015; Ignatieva et al., 2016; Saik et al., 2016), а также для разработки математических моделей биологических процессов (Подколотный и др., 2016; Podkolodnaya et al., 2017).

В настоящем обзоре представлены основные сведения о ряде информационных ресурсов, содержащих данные по генным сетям и их функциональным модулям, которые находятся в открытом доступе в сети Интернет в 2017 г. и ориентированы на человека и животных. Известны и коммерческие базы данных по этой тематике (см. Приложение 4), которые могут быть полезны для реконструкции молекулярно-генетических сетей. Однако они доступны через Интернет на платной основе (через логин и пароль). Сведения об этих коммерческих системах предоставляются фирмами-разработчиками.

Сопоставление основных характеристик интернет-ресурсов (см. табл. 1 и 2, а также Приложение 2 и 3) выявило большое разнообразие предлагаемых ими возможностей. Кроме того, очевидно, что каждый ресурс содержит уникальный набор данных и использует оригинальные подходы к представлению информации. В связи с этим на этапе планирования работы по реконструкции генной сети определенного биологического процесса очень важно иметь представление о максимально полном наборе информационных источников, из которых может быть взята информация. Стремительный рост объема новых экспериментальных данных по молекулярно-генетическим взаимодействиям побуждает исследователей к созданию новых информационных ресурсов и выпуску более современных версий (релизов) уже известных баз. Поэтому для получения наиболее свежих сведений о базах по тематике «генные сети» целесообразно обращаться к ежегодным выпускам журнала *Nucleic Acids Research*, посвященным базам данных (первый номер каждого года) (Galperin et al., 2017), к другим веб-порталам, а также обзорным публикациям по этой тематике из специальных научных журналов (*Database: The Journal of Biological Databases and Curation*, *BMC Bioinformatics* и т. п.).

Acknowledgments

Analysis of the potential of databases including graphic representations of signal transduction, metabolism, and regulation of other biologic processes was supported by the Government of the Russian Federation, project 14.B25.31.0033. Analysis of functional features of online resources storing data on intermolecular interactions between gene network objects was supported by State Budgeted Project 0324-2016-0008.

Conflict of interest

The authors declare no conflict of interest.

References

- Galperin M.Y., Fernández-Suárez X.M., Rigden D.J. The 24th annual Nucleic Acids Research database issue: a look back and upcoming changes. *Nucleic Acids Res.* 2017;45(D1):D1-D11. DOI 10.1093/nar/gkw1188.
- Glotov A.S., Tiys E.S., Vashukova E.S., Pakin V.S., Demenkov P.S., Saik O.V., Ivanisenko T.V., Arzhanova O.N., Mozgovaya E.V., Zainulina M.S., Kolchanov N.A., Baranov V.S., Ivanisenko V.A. Molecular association of pathogenetic contributors to pre-eclampsia (pre-eclampsia associome). *BMC Syst. Biol.* 2015;9(Suppl.2):S4. DOI 10.1186/1752-0509-9-S2-S4.
- Ignatieva E.V., Afonnikov D.A., Saik O.V., Rogaev E.I., Kolchanov N.A. A compendium of human genes regulating feeding behavior and body weight, its functional characterization and identification of GWAS genes involved in brain-specific PPI network. *BMC Genet.* 2016;17(Suppl.3):158. DOI 10.1186/s12863-016-0466-2.
- Ivanisenko V.A., Saik O.V., Ivanisenko N.V., Tiys E.S., Ivanisenko T.V., Demenkov P.S., Kolchanov N.A. ANDSystem: an Associative Network Discovery System for automated literature mining in the field of biology. *BMC Syst. Biol.* 2015;9(Suppl.2):S2. DOI 10.1186/1752-0509-9-S2-S2.
- Kauffman S. Homeostasis and differentiation in random genetic control networks. *Nature.* 1969;224(5215):177-178.
- Khatri P., Sirota M., Butte A.J. Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Comput. Biol.* 2012;8(2):e1002375. DOI 10.1371/journal.pcbi.1002375.
- Kolchanov N.A., Anan'ko E.A., Kolpakov F.A., Podkolodnaya O.A., Ignatieva E.V., Goryachkovskaya T.N., Stepanenko E.L. Gene networks. *Molekulyarnaya biologiya = Molecular Biology (Moscow).* 2000;34(4):533-544. (in Russian)
- Kolchanov N.A., Ignatieva E.V., Podkolodnaya O.A., Likhoshvai V.A., Matushkin Yu.G. Gene networks. *Vavilovskii Zhurnal Genetiki i Selektcii = Vavilov Journal of Genetics and Breeding.* 2013;17(4/2): 833-850. (in Russian)
- Kolchanov N.A., Merkulova T.I., Ignatieva E.V., Ananko E.A., Oshchepkov D.Y., Levitsky V.G., Vasiliev G.V., Klimova N.V., Merkulov V.M., Hodgman T.C. Combined experimental and computational approaches to study the regulatory elements in eukaryotic genes. *Brief. Bioinform.* 2007;8(4):266-274.
- Kutmon M., Riutta A., Nunes N., Hanspers K., Willighagen E.L., Bohler A., Mélius J., Waagmeester A., Sinha S.R., Miller R., Coort S.L., Cirillo E., Smeets B., Evelo C.T., Pico A.R. WikiPathways: capturing the full diversity of pathway knowledge. *Nucleic Acids Res.* 2016;44(D1):D488-494. DOI 10.1093/nar/gkv1024.
- Lei X., Wu S., Ge L., Zhang A. Clustering and overlapping modules detection in PPI network based on IBFO. *Proteomics.* 2013;13(2):278-290. DOI 10.1002/pmic.201200309.
- Levitsky V.G., Ignatieva E.V., Ananko E.A., Turnaev I.I., Merkulova T.I., Kolchanov N.A., Hodgman T.C. Effective transcription factor binding site prediction using a combination of optimization, a genetic algorithm and discriminant analysis to capture distant interactions. *BMC Bioinformatics.* 2007;8:481. DOI 10.1186/1471-2105-8-481.
- Li J.R., Suzuki T., Nishimura H., Kishima M., Maeda S., Suzuki H. Asymmetric regulation of peripheral genes by two transcriptional regulatory networks. *PLoS ONE.* 2016;11(8):e0160459. DOI 10.1371/journal.pone.0160459.
- McMullen P.D., Bhattacharya S., Woods C.G., Sun B., Yarborough K., Ross S.M., Miller M.E., McBride M.T., LeCluyse E.L., Clewell R.A., Andersen M.E. A map of the PPAR α transcription regulatory network for primary human hepatocytes. *Chem. Biol. Interact.* 2014;209:14-24. DOI 10.1016/j.cbi.2013.11.006.
- Mustafin Z.S., Lashin S.A., Matushkin Y.G., Gunbin K.V., Afonnikov D.A. Orthoscape: a cytoscape application for grouping and visualization KEGG based gene networks by taxonomy and homology principles. *BMC Bioinformatics.* 2017;18(Suppl.1):1427. DOI 10.1186/s12859-016-1427-5.
- Neph S., Stergachis A.B., Reynolds A., Sandstrom R., Borenstein E., Stamatoyannopoulos J.A. Circuitry and dynamics of human transcription factor regulatory networks. *Cell.* 2012;150(6):1274-1286. DOI 10.1016/j.cell.2012.04.040.
- Obermayer B., Levine E. Exploring the miRNA regulatory network using evolutionary correlations. *PLoS Comput. Biol.* 2014;10(10): e1003860. DOI 10.1371/journal.pcbi.1003860.
- Plaisier C.L., Pan M., Baliga N.S. A miRNA-regulatory network explains how dysregulated miRNAs perturb oncogenic processes across diverse cancers. *Genome Res.* 2012;22(11):2302-2314. DOI 10.1101/gr.133991.111.
- Podkolodnaya O.A., Tverdokhlebs N.N., Podkolodnyy N.L. Computational modeling of the cell-autonomous mammalian circadian oscillator. *BMC Syst. Biol.* 2017;11(Suppl.1):379. DOI 10.1186/s12918-016-0379-8.
- Podkolodnyy N.L., Tverdokhlebs N.N., Podkolodnaya O.A. Computational model for mammalian circadian oscillator: interacting with NAD $^{+}$ /SIRT1 pathway and age-related changes in gene expression of circadian oscillator. *Vavilovskii Zhurnal Genetiki i Selektcii = Vavilov Journal of Genetics and Breeding.* 2016;20(6):848-856. DOI 10.18699/VJ16.201. (in Russian)
- Ratner V.A. *Geneticheskie upravlyayushchie sistemy [Genetic Control Systems].* Novosibirsk: Nauka Publ., 1966. (in Russian)
- Reece-Hoyes J.S., Deplancke B., Shingles J., Grove C.A., Hope I.A., Walhout A.J. A compendium of *Caenorhabditis elegans* regulatory transcription factors: a resource for mapping transcription regulatory networks. *Genome Biol.* 2005;6(13):R110.
- Saik O.V., Ivanisenko T.V., Demenkov P.S., Ivanisenko V.A. Interactome of the hepatitis C virus: Literature mining with ANDSystem. *Virus Res.* 2016;218:40-48. DOI 10.1016/j.virusres.2015.12.003.
- Stergachis A.B., Neph S., Sandstrom R., Haugen E., Reynolds A.P., Zhang M., Byron R., Canfield T., Stelting-Sun S., Lee K., Thurman R.E., Vong S., Bates D., Neri F., Diegel M., Giste E., Dunn D., Vierstra J., Hansen R.S., Johnson A.K., Sabo P.J., Wilken M.S., Reh T.A., Treuting P.M., Kaul R., Groudine M., Bender M.A., Borenstein E., Stamatoyannopoulos J.A. Conservation of trans-acting circuitry during mammalian regulatory evolution. *Nature.* 2014; 515(7527):365-370. DOI 10.1038/nature13972.
- Tomaru Y., Hasegawa R., Suzuki T., Sato T., Kubosaki A., Suzuki M., Kawaji H., Forrest A.R., Hayashizaki Y., FANTOM Consortium, Shin J.W., Suzuki H. A transient disruption of fibroblastic transcriptional regulatory network facilitates *trans*-differentiation. *Nucleic Acids Res.* 2014;42(14):8905-8913. DOI 10.1093/nar/gku567.

Анализ циркадного ритма биологических процессов в печени и почках мыши

Н.Л. Подколотный^{1,2}✉, Н.Н. Твердохлеб^{1,3}, О.А. Подколотная¹

¹ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Институт вычислительной математики и математической геофизики Сибирского отделения Российской академии наук, Новосибирск, Россия

³ Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия

В статье представлены результаты исследования тканеспецифичности циркадных фазовых характеристик биологических процессов в печени и почках мыши. Основываясь на экспериментальных данных по суточной динамике уровня трансляции генов мыши из базы данных GEO (GSE67305 и GSE81283), полученных методом профилирования рибосом в печени и почках, мы провели сравнительный анализ транслятомов в этих двух органах. Были выявлены гены, демонстрирующие выраженную суточную динамику трансляции (3358 генов в печени и 2938 в почках). Далее для двенадцати временных точек (ZT0–ZT22) в каждой ткани (печень, почки) были выделены группы генов, находящиеся в фазе с повышенным уровнем трансляции. В работе было принято, что ген находится в фазе с повышенным уровнем трансляции, если в данной временной точке его показатель профилирования рибосом для обеих реплик превышал среднесуточное значение показателя для этого гена. Наибольшее количество ритмичных генов в печени имеет повышенный уровень трансляции в начале темной фазы суток, соответствующей повышенной активности животных. В почках различия в распределении по времени суток числа генов, находящихся в фазе повышенного уровня трансляции, были менее выражены, а максимальное число таких генов наблюдалось с середины светлой фазы суток до середины темной. Был проведен анализ обогащения терминами GO категории Biological Process этих двенадцати групп генов в печени и почках. Среди процессов, ритмичность которых характерна как для печени, так и для почек, выявлены процессы, циркадные фазовые характеристики которых в этих тканях совпадают, и процессы, имеющие существенно различные временные фазовые паттерны. Также выявлены процессы со строгой тканеспецифичностью ритмической трансляции. Подход, использованный в нашей работе, позволяет проводить анализ органа/тканеспецифичности фазовых характеристик биологических процессов, а полученные результаты подчеркивают необходимость учитывать фазовые циркадные характеристики при сравнении особенностей протекания биологических процессов в различных органах.

Ключевые слова: циркадный ритм; трансляция; анализ обогащения терминами GO; тканеспецифичность; биологические процессы; фазовые характеристики.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Подколотный Н.Л., Твердохлеб Н.Н., Подколотная О.А. Анализ циркадного ритма биологических процессов в печени и почках мыши. Вавиловский журнал генетики и селекции. 2017;21(8):903-910. DOI 10.18699/VJ17.311

HOW TO CITE THIS ARTICLE:

Podkolodny N.L., Tverdokhle N.N., Podkolodnaya O.A. Analysis of the circadian rhythm of biological processes in mouse liver and kidney. Vavilovskii Zhurnal Genetiki i Selektzii = Vavilov Journal of Genetics and Breeding. 2017;21(8):903-910. DOI 10.18699/VJ17.311 (in Russian)

Received 17.10.2017
Accepted for publication 09.11.2017
© AUTHORS, 2017

✉ e-mail: pnl@bionet.nsc.ru

Analysis of the circadian rhythm of biological processes in mouse liver and kidney

N.L. Podkolodny^{1,2}✉, N.N. Tverdokhle^{1,3},
O.A. Podkolodnaya¹

¹ Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

² Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk

³ Novosibirsk State University, Novosibirsk, Russia

The paper presents the results of a study of the tissue-specificity of the circadian phase characteristics of biological processes in the mouse liver and kidneys. We performed a comparative analysis of the translatoemes in these two organs based on experimental data on the daily dynamics of the level of translation of mouse genes from the GEO database (GSE67305 and GSE81283) obtained by ribosome profiling. Genes with a pronounced daily dynamics of translation were revealed (3358 genes in the liver and 2938 in the kidneys). Further, for each of the 12 time points (ZT0–ZT22), for each tissue (liver, kidneys), groups of genes that were in a phase with an increased level of translation were identified. It was assumed that the gene is in a phase with an increased level of translation if at a given time point its ribosome profiling rate for both replicas exceeded the daily average value for this gene. The greatest number of rhythmic genes in the liver has an increased level of translation at the beginning of the dark phase of the day corresponding to increased animal activity. In the kidneys, the differences in the distribution of the number of genes in the phase of an elevated translation level by the time of day were less pronounced, and the maximum number of such genes was observed from the middle of the light phase of the day to the middle of the dark one. A statistical analysis of enrichment of Gene Ontology terms in these twelve gene groups in the liver and kidneys was performed. Analyzing the processes, the rhythmicity of which is typical of both liver and kidneys, we have identified the processes, the circadian phase characteristics of which in these tissues coincide and the processes having essentially different temporal phase patterns for these tissues. Processes with strict tissue-specific rhythmic translation have also been identified. The approach used in our work allows us to analyze the organo/tissue-specificity of the phase characteristics of biological processes, and the results emphasize the need to take into account the phase circadian characteristics when comparing the features of the course of biological processes in various organs.

Key words: circadian rhythm; translation; GO enrichment analysis; tissue specificity; biological processes; phase characteristics.

Циркадные часы являются одним из фундаментальных эволюционно консервативных механизмов, который координирует большую часть процессов, протекающих в живых организмах. В основе их лежит ритмичная экспрессия генов, формирование которой осуществляется как на транскрипционном, так и на посттранскрипционном уровне. Автономные клеточные циркадные осцилляторы обеспечивают ритмичность биологических процессов в каждой клетке. Синхронизация ритмических процессов на уровне организма осуществляется главным водителем циркадного ритма, локализованным в супрахиазматических ядрах (СХЯ) гипоталамуса. Предполагается, что автономный клеточный механизм циркадных часов, основанный на транскрипционно-трансляционных взаимосвязанных петлях обратных связей, синхронизирует ритмическую экспрессию мРНК широкого спектра генов в клетке (Takahashi, 2017). Это находит отражение в ритмических изменениях трансляции белков, их уровней активности и, в конечном счете, синхронизирует ритмику разнообразных клеточных функций и биологических процессов. Нарушение нормальных циркадных ритмов может привести к развитию серьезных патологий, включая рак, сердечно-сосудистые, нейродегенеративные, воспалительные заболевания и др. (Oishi et al., 2005; Cao et al., 2009; McCarthy, Welsh, 2012; Kettner et al., 2014; Fang et al., 2015; Labrecque, Cermakian, 2015; Sundar et al., 2015).

Известно, что суммарно не менее сорока процентов от всех белок-кодирующих генов организма млекопитающих проявляют циркадные осцилляции по меньшей мере в одном органе на уровне мРНК (Zhang et al., 2014). Отмечают также наличие органо/тканеспецифичности циркадной ритмики экспрессии генов (Zhang et al., 2014; Laing et al., 2015). В то же время показано координирование фаз ритмов экспрессии генов, участвующих в одном биологическом пути в различных органах. Примером могут служить гены, экспрессирующиеся в печени и мышцах, продукты которых участвуют в регуляции уровня глюкозы в крови. Фазы экспрессии этих генов координируются таким образом, чтобы облегчить регуляцию уровня глюкозы в крови (Harfmann et al., 2015).

Циркадная синхронизация функционирования генетических и метаболических сетей, путей передачи сигналов и т. д. может осуществляться не только на уровне транскрипции, но и на уровне трансляции, посттранскрипционной и посттрансляционной модификации (Kojima et al., 2012; Menet et al., 2012; Morf et al., 2012; Beckwith, Yanovsky, 2014; Lück et al., 2014; Nolte, Staiger, 2015). Оказалось, что только у относительно небольшой части генов ритмический характер экспрессии наблюдается на уровне транскрипции и не всегда ему соответствует ритмический характер экспрессии кодируемых ими белков (Koike et al., 2012; Menet et al., 2012), и наоборот. Например, в печени, по разным оценкам, от 20 до 50 % белков, демонстрирующих суточный ритм, кодируются неритмичными мРНК (Reddy et al., 2006; Mauvoisin et al., 2014; Robles et al., 2014). Появление методов исследования транслятома (всей совокупности фрагментов мРНК, транслируемых в определенный момент) позволяет не только лучше понять роль трансляционной и посттрансляционной регуляции

в функционировании циркадных часов организма, но и выявить их связь с суточной динамикой биологических процессов, физиологических характеристик и поведения. Поскольку корреляция данных протеома и транслятома выше, чем протеома и транскриптома, можно рассматривать транслятом как более точную оценку уровня экспрессии генов (Smircich et al., 2015). Это кажется тем более верным, что именно белки являются функциональными единицами в клетке, в отличие от мРНК, которые в данном случае можно рассматривать как промежуточные звенья (Robles, Mann, 2013).

В настоящей работе мы провели сравнительный анализ транслятома печени и почек мыши, представленных в исследованиях (Janich et al., 2015; Castelo-Szekely et al., 2017), выявили гены, имеющие выраженную суточную динамику трансляции, определили периоды времени, когда она достигает максимального уровня. Это позволило провести исследование тканеспецифичности фазовых характеристик биологических процессов в течение суток на основе функционального анализа ГО.

Материалы и методы

В исследовании мы использовали данные, полученные методом профилирования рибосом при изучении транслятома печени и почек мыши, представленные в работах группы David Gatfield (Janich et al., 2015; Castelo-Szekely et al., 2017) (GSE67305 и GSE81283 в базе данных GEO). В настоящее время метод профилирования рибосом находит все более широкое применение для решения широкого круга задач, связанных с изучением механизмов регуляции трансляции в живой клетке. В частности, этот подход обеспечивает информацию о транслируемых мРНК и их трансляционной активности в определенный момент времени *in vivo* (Ingolia, 2014). Мыши (самцы линии C57BL/6 возраста 12 недель) содержались в течение двух недель в условиях свето-темнового режима (12 ч свет/12 ч темнота). Затем мышей забивали через каждые 2 ч в течение 24-часового суточного цикла (12 временных точек (ZT0–ZT22), где ZT0 – включение света, ZT12 – выключение света). Для дальнейших исследований у животных забивали печень и почки с целью проведения экспериментов по профилированию рибосом в каждой ткани (Janich et al., 2015; Castelo-Szekely et al., 2017). Таким образом, для каждой временной точки в двух независимых репликах эксперимента были получены значения показателя профилирования рибосом анализируемых генов, которые мы использовали в нашем исследовании.

Идентификация генов, демонстрирующих циркадную ритмику трансляции, была осуществлена методом, основанным на комбинации критерия, используемого в программе eJTK_CYCLE (empP < 0.05) (Hughes et al., 2010; Hutchison et al., 2015), с последующим анализом разброса измерений в каждой временной точке суточного ритма для выявления устойчивых паттернов экспрессии, воспроизводимых в каждом суточном цикле наблюдения для нескольких реплик (в нашем случае две реплики). Показатель empP в программе eJTK_CYCLE соответствует эмпирической оценке уровня значимости (*p-value*) с учетом множественности выбора, полученной путем имитационного моделирования (Hutchison et al., 2015).

Ритмичный ген должен также удовлетворять дополнительным условиям: 1) амплитуда колебаний уровня показателя профилирования рибосом гена должна быть не менее 20 % от его среднего значения за сутки; 2) хотя бы для одной временной точки показатель профилирования рибосом в обеих репликах должен превышать его среднесуточное значение. В нашем исследовании принимается, что ген находится в фазе с повышенным уровнем трансляции, если в данной временной точке показатель профилирования рибосом этого гена для обеих реплик превышал его среднесуточное значение для этого гена.

Анализ обогащения терминами GO групп генов проведен с использованием биоинформационного ресурса DAVID (Huang et al., 2009). Для оценки достоверности использовали критерий Бенджамини с поправкой на множественность выбора ($p < 0.05$).

Результаты и обсуждение

Выделение групп ритмичных генов

Анализ экспериментальных данных по экспрессии генов (уровень трансляции), полученных методом профилирования рибосом, показал, что из общего числа генов, экспрессирующихся в почках и печени, 10573 оказались общими для двух органов, 407 уникальны для печени, 2049 – для почек (рис. 1).

На первом этапе работы мы выявили гены, демонстрирующие циркадный ритм трансляции, оцениваемый по величине показателя профилирования рибосом, в печени и почках. В дальнейшем будем называть такие гены ритмичными. Количество выявленных генов с ритмичным характером трансляции составляет: 3358 из 10980, экспрессирующихся в печени, и 2938 из 12622 в почках (см. рис. 1, a). Так же, как и в исследовании (Castelo-Szekely et al., 2017), в печени количество ритмичных генов оказалось выше, чем в почках.

Согласно нашим результатам, процентное содержание ритмичных генов составляло 30 % в печени и 23 % в почках и оказалось выше, чем в исследованиях (Janich et al., 2015; Castelo-Szekely et al., 2017) – 17 и 8 % соответственно. Это можно объяснить различием подходов и критериев, использованных для выявления ритмичных генов, что определяется особенностями задач, поставленных в наших исследованиях. Отметим, кроме того, что в литературе упоминаются существенные различия числа ритмичных генов, выявленных разными экспериментальными и математическими методами. Например, с использованием микрочипов одними исследователями выявлено 335 ритмичных генов в печени (Panda et al., 2002), а другими – 2828 (Eckel-Mahan et al., 2012). Методом RNA-seq в этом же органе выявлено 1160 (Vollmers et al., 2009) и 4544 (Atger et al., 2015) ритмичных гена.

Исходной посылкой для нашего исследования является то, что ритмичные гены с высокой вариабельностью уровня экспрессии в определенной ткани в течение суток можно рассматривать для этой ткани как время-специфические гены, экспрессия которых в определенные периоды суток увеличивается для выполнения специфических функций в конкретных биологических процессах и динамика активности которых может иметь циркадный характер.

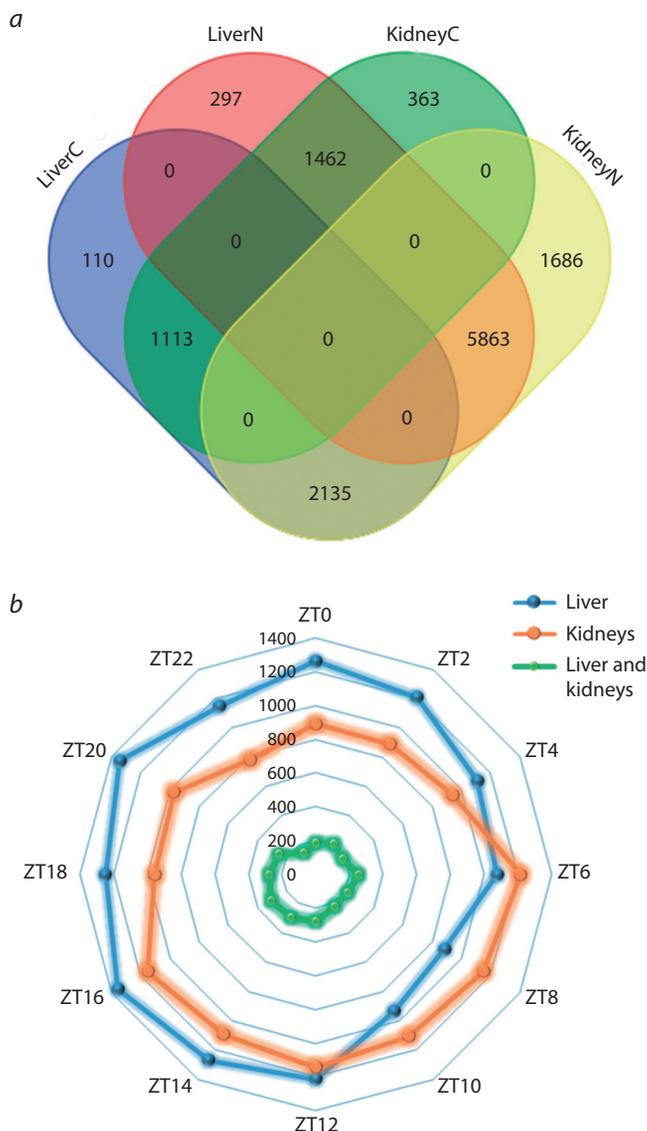


Fig. 1. (a) A Venn diagram illustrating overlaps of the sets of cyclic and noncyclic genes expressed in the liver (LiverC and LiverN, respectively) and in kidneys (KidneyC and KidneyN). (b) A circle diagram of the distribution of genes in the elevated translation state at certain timepoints of the circadian cycle in the liver, kidneys, and both.

Для наших целей важно выделить группу генов, которые воспроизводят цикличность экспрессии (трансляции) с периодом 24 ч. При отборе ритмичных генов мы использовали не только критерии циркадности, но и дополнительные критерии отбора, которые и позволили выявить гены с высокой относительной амплитудой уровня трансляции (см. раздел Материалы и методы).

На диаграмме Венна (см. рис. 1, a) представлены пересечения множеств выделенных нами ритмичных и неритмичных генов в печени и почках. Общее число анализируемых генов в печени ($LiverC \cup LiverN$) – 10980, в почках ($KidneyC \cup KidneyN$) – 12622. Из них в печени ритмичны 3358 генов (LiverC), неритмичны – 7622 гена (LiverN); в почках ритмичны 2938 генов (KidneyC), не-

ритмичны – 9684 гена (KidneyN); 1113 генов ритмичны в обоих органах (LiverC \cap KidneyN).

Для конкретных ритмичных генов были выявлены временные интервалы, соответствующие их фазам повышенной экспрессии (трансляции). Это позволило сформировать группы генов, находящихся в фазе повышенной экспрессии, для каждой из 12 исследованных временных точек суточного цикла.

На следующем этапе работы мы задались вопросом, как распределено по времени суток число генов, входящих в эти группы в печени и почках. Фазовые гистограммы, описывающие это распределение, представлены на рис. 1, б. В печени максимальное количество генов находится в фазе повышенной экспрессии (трансляции) в период, соответствующий ZT14–ZT20, минимальное – в период времени ZT8–ZT10.

В целом этот результат не противоречит данным (Janich et al., 2015), показавшим при более жестких условиях отбора ритмичных генов, что акрофаза (максимальное значение) ритма трансляции в печени приходится на ZT15–ZT16. Таким образом, наибольшее количество ритмичных генов в фазе повышенной экспрессии наблюдается в печени в начале темной фазы суток, соответствующей повышенной активности животных, а минимальное – в конце светлой, в период покоя, предшествующий переходу к активной фазе. В почках различия на протяжении суток были менее выражены, а максимум числа генов в фазе повышенной трансляции наблюдался с середины светлой фазы суток до середины темной (ZT6–ZT16), что соответствует результатам работы (Castelo-Szekely et al., 2017), в которой исследовались циркадные ритмы эффективности трансляции.

Максимальные числа генов в фазе с повышенной экспрессией (трансляцией) одновременно и в почках, и в печени наблюдаются в период ZT14–ZT16, что совпадает с пересечением временных периодов с максимальным количеством таких генов для печени и почек. Таким образом, в этот период в печени и почках задействовано большое число одних и тех же генов, которые, по-видимому, участвуют в регуляции сходных биологических процессов в этих тканях.

Функциональный анализ генов с ритмичным типом трансляции

Анализ обогащения терминами GO категории Biological Process (BP_DIRECT) был проведен в группах генов с ритмичной трансляцией в печени и почках. Каждая из групп демонстрирует обогащение как специфическими терминами GO, так и терминами, характерными для обеих групп. Среди общих можно отметить ожидаемые термины *rhythmic process*, *circadian regulation of gene expression*, *circadian rhythm*, что может служить подтверждением правильности формирования групп ритмичных генов (см. таблицу). Остальные термины из этого списка также не имеют органной специфичности и касаются процессов транспорта, транскрипции, процессинга мРНК, ответа на стресс эндоплазматического ретикулума и др.

Проведенный анализ позволил выявить общие термины GO, которыми обогащены группы ритмичных генов печени и почек. Однако остается неясным, насколько

такое функциональное сходство объясняется совпадением множеств генов в печени и почках, дающих вклад в обогащение термином GO.

Функциональный GO анализ обогащения терминами категории Biological Process показал, что термины, являющиеся «общими» для ритмичных генов печени и почек, также в значительной степени и с высокой достоверностью представлены в группе генов, соответствующих пересечению множеств ритмичных генов печени и почек (см. таблицу). Данное наблюдение может свидетельствовать о высоком вкладе в уровень обогащения указанными терминами GO группы генов, имеющих повышенную трансляцию как в печени, так и в почках.

Таким образом, можно отметить, что биологические процессы из рассмотренного выше списка не являются тканеспецифичными ни для печени, ни для почек и, вероятно, ритмичны в обоих органах. В то же время орган/тканеспецифичность их фазовых характеристик требует дальнейшей проверки. На это же указывают результаты работы (Castelo-Szekely et al., 2017), свидетельствующие о существенном различии фазовых характеристик эффективности трансляции в печени и почках.

Фазовые характеристики обогащения терминами GO категории Biological Process в группах ритмичных генов печени и почек

На следующем этапе работы мы задались вопросом о сходстве или различии фазовых характеристик обогащения терминами категории Biological Process групп ритмичных генов, выявленных в печени и почках.

Для каждой из 12 временных точек (ZT0–ZT22) были выявлены группы генов, находящихся в фазе, характеризующейся повышенным уровнем экспрессии (трансляции). Формирование групп генов было осуществлено так, что для каждого гена, входящего в конкретную временную группу, его показатель профилирования рибосом для обеих реплик в указанной временной точке превышал среднесуточное значение. GO анализ обогащения этих 12 групп генов терминами категории Biological Process был проведен для каждой ткани (печень и почки).

Сначала мы сосредоточились на биологических процессах, которые оказались «общими» для печени и почек (см. таблицу). При этом мы предполагали, что наличие циркадной динамики обогащения терминами определенных биологических процессов можно рассматривать как свидетельство ритмичности этих процессов.

Оказалось, что фазы обогащения некоторыми терминами этой категории совпадают или частично перекрываются для генов печени и почек. Примером таких совпадений могут служить термины *rhythmic process*, *response to unfolded protein*, *lipid metabolic process* (рис. 2). Фазовые характеристики других могут существенно различаться, как в случае термина *fatty acid metabolic process*. Также можно отметить термины, обогащение которыми в группах генов одного органа проявляет циркадную динамику, а другого – нет. Примером таких терминов является *cell-cell adhesion*, *cellular response to insulin stimulus* (см. рис. 2).

Дальнейший анализ касался терминов категории Biological Process, обогащение которыми не является «об-

GO terms of the Biological Process category overrepresented in groups of cyclic genes in the liver and kidneys

Terms	Liver, <i>p</i> -value	Kidneys, <i>p</i> -value
*GO:0048511~rhythmic process	8.07E-05	11.04E-08
*GO:0032922~circadian regulation of gene expression	9.90E-04	1.35E-07
*GO:0006457~protein folding	1.10E-05	2.71E-07
*GO:0006629~lipid metabolic process	1.34E-09	3.13E-08
*GO:0007623~circadian rhythm	7.41E-04	8.68E-07
*GO:0034976~response to endoplasmic reticulum stress	0.00258	3.74E-05
*GO:0006397~mRNA processing	1.22E-05	4.59E-09
*GO:0016310~phosphorylation	3.69E-06	2.54E-05
*GO:0008380~RNA splicing	1.23E-04	9.21E-07
*GO:0006986~response to unfolded protein	5.66E-05	0.0127
*GO:0006631~fatty acid metabolic process	3.73E-06	1.03E-05
*GO:0006810~transport	4.09E-12	1.05E-15
*GO:0055114~oxidation-reduction process	2.94E-13	0.0288
GO:0098609~cell-cell adhesion	5.63E-07	0.0019
GO:0015031~protein transport	1.65E-05	4.79E-05
GO:0006914~autophagy	1.30E-04	0.0178
GO:0006915~apoptotic process	4.32E-04	0.018
GO:0045944~positive regulation of transcription from RNA polymerase II promoter	0.01428	0.0188
GO:0032869~cellular response to insulin stimulus	0.00129	0.0296

p-value calculated with the Benjamini correction for multiple choice.

* GO terms found to be overrepresented in the group of cyclic genes shared by both organs.

цим» для ритмичных генов печени и почек, а также не выявлено в группах генов печени и почек по отдельности, но обнаружено при анализе циркадной динамики обогащения в одной или нескольких временных точках. Так, известно, что один из процессов – метаболизм фруктозы, осуществляется преимущественно в печени. Обогащение термином *fructose metabolic process* не наблюдается в группе ритмичных генов почек и группе общих генов, но выявлено в группе ритмичных генов печени в период ZT14–ZT16 при анализе циркадной динамики обогащения этим термином (см. рис. 2). Для терминов *cholesterol metabolic process* и *cholesterol biosynthetic process* обогащение выявлено только в группах ритмичных генов печени в период ZT20–ZT22 (см. рис. 2).

Можно предполагать, что сходство фазовых характеристик биологических процессов в двух органах является отражением существования некоторого единого координирующего воздействия. Ранее было отмечено, что для большинства генов именно ритмическая транскрипция преимущественно управляет ритмическим накоплением мРНК и их трансляцией, во всяком случае в печени (Atger et al., 2015), тогда как эффективность трансляции скорее модулирует экспрессию в направлении большего ее межорганного сходства на уровне биосинтеза белка (Castelo-Szekely et al., 2017).

С этой точки зрения в качестве источника координирующего воздействия можно рассматривать синхронизирующие сигналы центральных циркадных часов организма

(супрахиазматических ядер гипоталамуса (СХЯ)), которые через различного рода посредников передаются к периферическим органам. Показано также координирующее воздействие циркадных часов на процессы трансляции в печени за счет их влияния на процессы биогенеза рибосом, максимум которого приходится на начало темной фазы суток (Jouffe et al., 2013).

В то же время различие фаз биологических процессов в различных органах может определяться специфическими для данной ткани или органа синхронизирующими сигналами, важными для выполнения основных функций данного органа, исходящими из других источников. Влияние таких сигналов может проявляться на различных уровнях регуляции экспрессии гена. Это находит подтверждение в исследовании (Atger et al., 2015), показавшем, что режим питания является синхронизирующим сигналом, определяющим ритм трансляции генов в печени, уровень мРНК которых не демонстрирует циркадного ритма. Отмечено также, что, хотя механизмы автономных клеточных циркадных осцилляторов в различных органах имеют практически одинаковый молекулярный состав, наличие органоспецифических особенностей регуляции эффективности трансляции их компонентов может позволить осуществлять настройку специфических для данного органа процессов в соответствии с его основными функциями (Castelo-Szekely et al., 2017).

В настоящее время помимо циркадных часов СХЯ, датчиком времени (*Zeitgeber*) для которых служит свет,

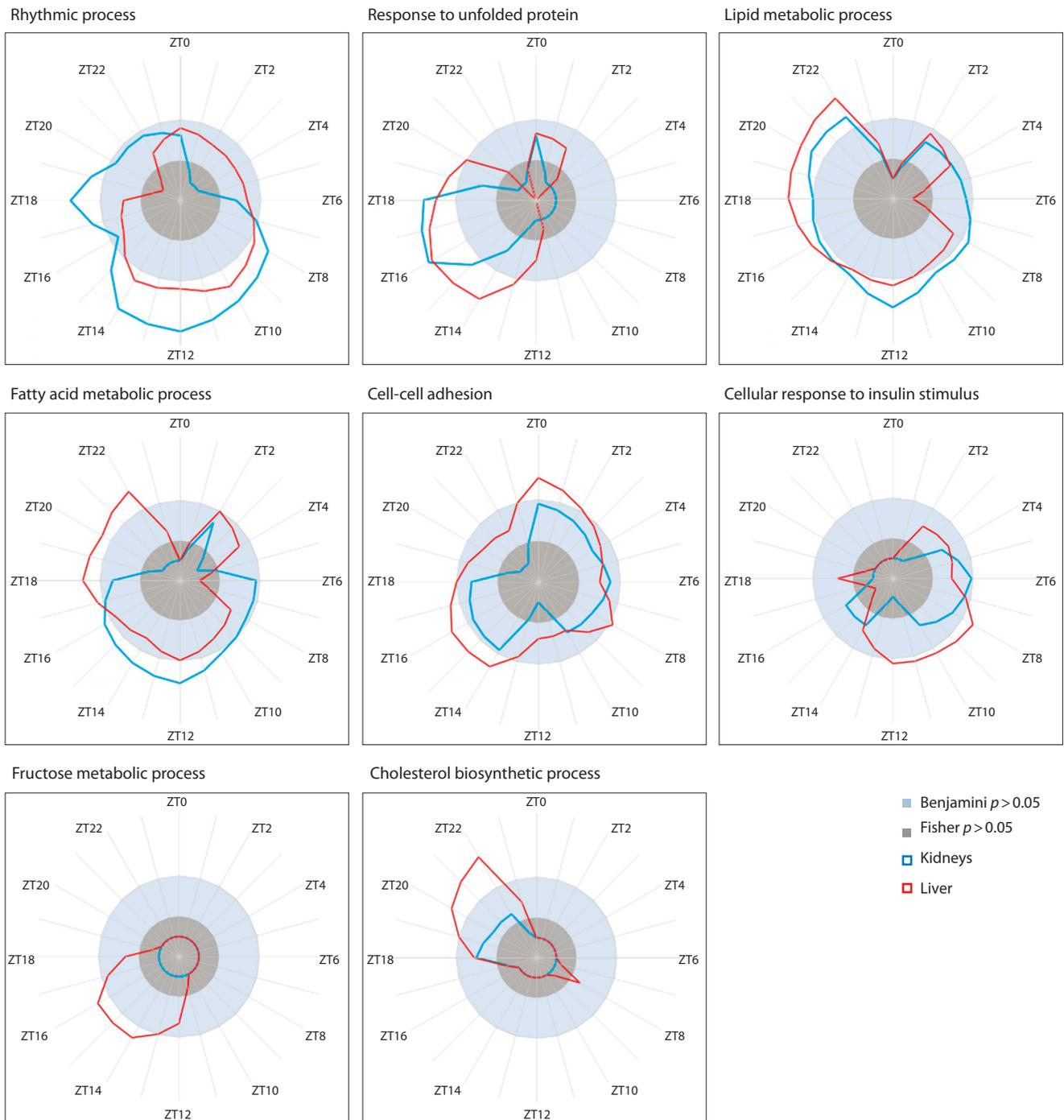


Fig. 2. Circadian variation of the enrichment of the Biological Process category with GO terms in groups of cyclic genes in the liver and kidneys.

описаны и другие, хотя их изучение затруднительно и находится в самом начале пути. К ним, например, относятся осцилляторы, обусловленные пищей (food entrainable oscillator – FEO), физической активностью и другими сигналами-синхронизаторами (Blum et al., 2012; Pendergast, Yamazaki, 2014; Flóres et al., 2016). Эти неканонические осцилляторы могут участвовать в регуляции циркадных ритмов, конкурируя в определенных ситуациях за влияние на синхронизацию определенных процессов с центральным циркадным осциллятором СХЯ.

Заключение

Изучение тканеспецифических особенностей циркадных ритмов экспрессии генов – одна из важных проблем современной хронобиологии. Такого рода знания позволяют по-новому взглянуть на некоторые задачи в области физиологии, медицины, хронотерапии, хронофармакологии и т. д. Использование данных профилирования рибосом предоставляет дополнительные возможности в исследовании циркадной регуляции трансляции и ее роли в формировании тканеспецифической циркадной ритмики

биологических процессов. В настоящей работе мы провели сравнительный анализ циркадных транскриптомов печени и почек мыши, представленных в исследованиях (Janich et al., 2015; Castelo-Szekely et al., 2017), выявили группы генов, имеющих выраженную суточную динамику трансляции, определили периоды времени, когда эти гены находятся в фазе с повышенным уровнем экспрессии (трансляции) в каждом из органов.

В результате были получены распределения ритмично транскрибуемых генов по времени их повышенной экспрессии в течение 24-часового цикла. В печени максимальное число генов транскрибируется на высоком уровне в начале темной фазы суток. В почках эти различия выражены меньше, и максимум приходится на период с середины светлой до середины темной фазы суток.

Анализ фазовых характеристик обогащения терминами GO категории Biological Process в группах ритмичных генов печени и почек показал, что некоторые процессы, например *response to unfolded protein* и *lipid metabolic process*, не только ритмичны в обоих органах, но и имеют близкие фазовые характеристики. В то время как ряд других биологических процессов (такие как *cell-cell adhesion* и *cellular response to insulin stimulus*) проявляет строгую тканеспецифичность ритмичности, т. е. процессы, проявляющие ритмичность в одном органе, оказываются неритмичными в другом.

Таким образом, ритмичные гены с высокой амплитудой экспрессии можно рассматривать для определенной ткани как времяспецифические гены, которые имеют повышенную экспрессию в определенные периоды суток для выполнения специфических функций в конкретных биологических процессах. Этот подход позволяет проводить анализ органа/тканеспецифичности фазовых характеристик биологических процессов, а полученные результаты подчеркивают необходимость учитывать фазовые циркадные характеристики при сравнении особенностей протекания биологических процессов в различных органах.

Acknowledgments

This work was supported by State Budgeted Project 0324-2016-0008.

Conflict of interest

The authors declare no conflict of interest.

References

Atger F., Gobet C., Marquis J., Martin E., Wang J., Weger B., Lefebvre G., Descombes P., Naef F., Gachon F. Circadian and feeding rhythms differentially affect rhythmic mRNA transcription and translation in mouse liver. *Proc. Natl. Acad. Sci. USA.* 2015;112(47):E6579-E6588. Epub 2015; Nov 9. DOI 10.1073/pnas.1515308112.

Beckwith E.J., Yanovsky M.J. Circadian regulation of gene expression: at the crossroads of transcriptional and post-transcriptional regulatory networks. *Curr. Opin. Gene Dev.* 2014;27:35-42. DOI 10.1016/j.gde.2014.03.007.

Blum I.D., Lamont E.W., Abizaid A. Competing clocks: metabolic status moderates signals from the master circadian pacemaker. *Neurosci. Biobehav. Rev.* 2012;36(1):254-270. Epub 2011; Jun 12. DOI 10.1016/j.neubiorev.2011.06.003.

Cao Q., Gery S., Dashti A., Yin D., Zhou Y., Gu J., Koeffler H.P. A role for the clock gene *Per1* in prostate cancer. *Cancer Res.* 2009;69(19):7619-7625.

Castelo-Szekely V., Arpat A.B., Janich P., Gatfield D. Translational contributions to tissue specificity in rhythmic and constitutive gene expression. *Genome Biol.* 2017;18(1):116. PMID: 28622766Ko.

Eckel-Mahan K.L., Patel V.R., Mohney R.P., Vignola K.S., Baldi P., Sassone-Corsi P. Coordination of the transcriptome and metabolome by the circadian clock. *Proc. Natl. Acad. Sci. USA.* 2012;109(14):5541-5546. Epub 2012; Mar 19. DOI 10.1073/pnas.1118726109.

Fang L., Yang Z., Zhou J., Tung J.Y., Hsiao C.D., Wang L., Deng Y., Wang P., Wang J., Lee M.H. Circadian clock gene *CRY2* degradation is involved in chemoresistance of colorectal cancer. *Mol. Cancer Ther.* 2015;14(6):1476-1487.

Flóres D.E., Bettilyon C.N., Yamazaki S. Period-independent novel circadian oscillators revealed by timed exercise and palatable meals. *Sci. Rep.* 2016;6:21945. DOI 10.1038/srep21945.

Harfmann B.D., Schroder E.A., Esser K.A. Circadian rhythms, the molecular clock, and skeletal muscle. *J. Biol. Rhythms.* 2015;30(2):84-94. Epub 2014; Dec 15. DOI 10.1177/0748730414561638.

Huang D.W., Sherman B.T., Lempicki R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 2009;4:44-57.

Hughes M.E., Hogenesch J.B., Kornacker K. JTK_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *J. Biol. Rhythms.* 2010;25(5):372-380. DOI 10.1177/0748730410379711.

Hutchison A.L., Maienschein-Cline M., Chiang A.H., Ali Tabei S.M., Gudjonson H., Bahroos N., Allada R., Dinner A.R. Improved statistical methods enable greater sensitivity in rhythm detection for genome-wide data. *PLoS Comput. Biol.* 2015;11(3):e1004094. DOI 10.1371/journal.pcbi.1004094.

Ingolia N.T. Ribosome profiling: new views of translation, from single codons to genome scale. *Nat. Rev. Genet.* 2014;15(3):205-213. Epub 2014; Jan 28. DOI 10.1038/nrg3645.

Janich P., Arpat A.B., Castelo-Szekely V., Lopes M., Gatfield D. Ribosome profiling reveals the rhythmic liver transcriptome and circadian clock regulation by upstream open reading frames. *Genome Res.* 2015;25(12):1848-1859. PMID: 26486724.

Jouffe C., Cretenet G., Symul L., Martin E., Atger F., Naef F., Gachon F. The circadian clock coordinates ribosome biogenesis. *PLoS Biol.* 2013;11(1):e1001455. Epub 2013; Jan 3. DOI 10.1371/journal.pbio.1001455.

Kettner N.M., Katchy C.A., Fu L. Circadian gene variants in cancer. *Ann. Med.* 2014;46(4):208-220.

Koike N., Yoo S.H., Huang H.C., Kumar V., Lee C., Kim T.K., Takahashi J.S. Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science.* 2012;338(6105):349-354.

Kojima S., Sher-Chen E.L., Green C.B. Circadian control of mRNA polyadenylation dynamics regulates rhythmic protein expression. *Genes Dev.* 2012;26:2724-2736.

Labrecque N., Cermakian N. Circadian clocks in the immune system. *J. Biol. Rhythms.* 2015;30(4):277-290.

Laing E.E., Johnston J.D., Möller-Levet C.S., Bucca G., Smith C.P., Dijk D.J., Archer S.N. Exploiting human and mouse transcriptomic data: Identification of circadian genes and pathways influencing health. *Bioessays.* 2015;37(5):544-556. Epub 2015; Mar 14. DOI 10.1002/bies.201400193.

Lück S., Thurley K., Thaben P.F., Westermark P.O. Rhythmic degradation explains and unifies circadian transcriptome and proteome data. *Cell. Rep.* 2014;9(2):741-751.

Mauvoisin D., Wang J., Jouffe C., Martin E., Atger F., Waridel P., Quadroni M., Gachon F., Naef F. Circadian clock-dependent and -independent rhythmic proteomes implement distinct diurnal functions in mouse liver. *Proc. Natl. Acad. Sci. USA.* 2014;111(1):167-172.

McCarthy M.J., Welsh D.K. Cellular circadian clocks in mood disorders. *J. Biol. Rhythms.* 2012;27(5):339-352.

Menet J.S., Rodriguez J., Abruzzi K.C., Rosbash M. Nascent-Seq reveals novel features of mouse circadian transcriptional regulation. *eLIFE.* 2012;1:e00011. DOI 10.7554/eLife.00011.

- Morf J., Rey G., Schneider K., Stratmann M., Fujita J., Naef F., Schibler U. Cold-inducible RNA-binding protein modulates circadian gene expression posttranscriptionally. *Science*. 2012;338(6105):379-383.
- Nolte C., Staiger D. RNA around the clock – regulation at the RNA level in biological timing. *Front. Plant Sci.* 2015;6:311. DOI 10.3389/fpls.2015.00311.
- Oishi K., Ohkura N., Amagai N., Ishida N. Involvement of circadian clock gene *Clock* in diabetes-induced circadian augmentation of plasminogen activator inhibitor-1 (*PAI-1*) expression in the mouse heart. *FEBS Lett.* 2005;579(17):3555-3559. DOI 10.1016/j.febslet.2005.05.027.
- Panda S., Antoch M.P., Miller B.H., Su A.I., Schook A.B., Straume M., Schultz P.G., Kay S.A., Takahashi J.S., Hogenesch J.B. Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell*. 2002;109(3):307-320.
- Pendergast J.S., Yamazaki S. Effects of light, food, and methamphetamine on the circadian activity rhythm in mice. *Physiol. Behav.* 2014;128:92-98. Epub 2014; Feb 11. DOI 10.1016/j.physbeh.2014.01.021.
- Reddy A.B., Karp N.A., Maywood E.S., Sage E.A., Deery M., O'Neill J.S., Wong G.K., Chesham J., Odell M., Lilley K.S., Kyriacou C.P., Hastings M.H. Circadian orchestration of the hepatic proteome. *Curr. Biol.* 2006;16(11):1107-1115.
- Robles M.S., Cox J., Mann M. In-vivo quantitative proteomics reveals a key contribution of post-transcriptional mechanisms to the circadian regulation of liver metabolism. *PLoS Genet.* 2014;10(1):e1004047. DOI 10.1371/journal.pgen.1004047.
- Robles M.S., Mann M. Proteomic approaches in circadian biology. In: *Circadian Clocks: Handb. Exp. Pharmacol.* A. Kramer, M. Meroz (Eds.) Springer, 2013;(217):389-407. DOI 10.1007/978-3-642-25950-0_17.
- Smircich P., Eastman G., Bispo S., Duhagon M.A., Guerra-Slompo E.P., Garat B., Goldenberg S., Munroe D.J., Dallagiovanna B., Holecz F., Sotelo-Silveira J.R. Ribosome profiling reveals translation control as a key mechanism generating differential gene expression in *Trypanosoma cruzi*. *BMC Genomics.* 2015;16:443. DOI 10.1186/s12864-015-1563-8.
- Sundar I.K., Yao H., Sellix M.T., Rahman I. Circadian molecular clock in lung pathophysiology. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 2015;309(10):L1056-L1075. DOI 10.1152/ajplung.00152.2015.
- Takahashi J.S. Transcriptional architecture of the mammalian circadian clock. *Nat. Rev. Genet.* 2017;18(3):164-179. Epub 2016; Dec 19. DOI 10.1038/nrg.2016.150.
- Vollmers C., Gill S., DiTacchio L., Pulivarthy S.R., Le H.D., Panda S. Time of feeding and the intrinsic circadian clock drive rhythms in hepatic gene expression. *Proc. Natl. Acad. Sci. USA.* 2009;106(50):21453-21458. Epub 2009; Nov 25. DOI 10.1073/pnas.0909591106.
- Zhang R., Lahens N.F., Balance H.I., Hughes M.E., Hogenesch J.B. A circadian gene expression atlas in mammals: implications for biology and medicine. *Proc. Natl. Acad. Sci. USA.* 2014;111(45):16219-16224. PMID: 25349387. DOI 10.1073/pnas.1408886111.

Роль генов апоптоза в контроле агрессивного поведения, выявленная с помощью комбинированного анализа ассоциативных генных сетей, экспрессионных и геномных данных по серым крысам с агрессивным поведением

А.О. Брагин¹✉, О.В. Сайк¹, И.В. Чадаева^{1,2}, П.С. Деменков¹, А.Л. Маркель^{1,2}, Ю.Л. Орлов^{1,2}, Е.И. Погаев^{1,3,4,5}, И.Н. Лаврик^{1,6}, В.А. Иванисенко¹

¹ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия

³ Медицинская школа Массачусетского университета в Вустере, Вустер, Массачусетс, США

⁴ Институт общей генетики Российской Академии наук, Москва, Россия

⁵ Центр генетики и генетических технологий, биологический факультет, факультет биоинженерии и биоинформатики, Московский государственный университет им. М.В. Ломоносова, Москва, Россия

⁶ Магдебургский университет имени Отто фон Гуерике, Магдебург, Германия

Агрессивное поведение животных играет важную роль при защите территории, потомства, установлении социально-иерархических отношений и т.д. При ряде заболеваний (шизофрения, маниакально-депрессивный психоз, нейродегенеративные заболевания) наблюдается повышенная агрессия. В нейрогенезе большое значение в поддержании клеточного гомеостаза имеет нейрональный апоптоз. Нарушения нейронального апоптоза отмечаются при старении и различных нейропатологиях (эпилепсия, болезнь Альцгеймера, нейротравмы), сопровождающихся изменениями психоэмоционального состояния. Известно, что в мозге высокоагрессивных крыс значительно изменяется уровень экспрессии ключевых генов нейронального апоптоза (*Casp3*, *Bax* и *Bcl-xl*). В связи с этим актуальным является изучение связей нейронального апоптоза и агрессивного поведения. Целью данной работы был анализ ассоциативных сетей, описывающих молекулярно-генетические взаимодействия между генами/белками, вовлеченными в нейрональный апоптоз, дифференциально экспрессированными генами и генами, имеющими полиморфизмы, у серых крыс с агрессивным поведением. Выявлено 819 дифференциально экспрессирующихся генов в гипоталамусе, вентральной тегментальной области и в сером веществе периакведуктума у крыс с агрессивным и дружелюбным поведением. Анализ ассоциативной сети дифференциально экспрессирующихся генов позволил выявить три вершины с максимальной центральностью, которые соответствовали генам *Stx1a*, *Mbp* и *Th*. При анализе генома было обнаружено 137 полиморфных генов, три из которых (*Lig4*, *Parp1* и *Pigt*) вовлечены в нейрональный апоптоз. Показано, что среди генов, взаимодействующих в ассоциативной сети с генами, вовлеченными в нейрональный апоптоз, статистически значимо представлены полиморфные и дифференциально экспрессированные гены (p -value < 0.01). Реконструированы три молекулярно-генетические цепочки, описывающие связи между полиморфными генами и генами/белками нейронального апоптоза, опосредованные через дифференциально экспрессированные гены. Цепочки включали полиморфные гены *Tsc1*, *Adamts4* и *Lgals3*, дифференциально экспрессированные гены *Ezr*, *Acan*, *Th* и 19 генов нейронального апоптоза. Было показано, что процесс нейронального апоптоза тесно связан с агрессивным поведением у животных.

Ключевые слова: нейрональный апоптоз; агрессивное поведение; ANDSystem; ассоциативные генные сети; дифференциально экспрессирующиеся гены; полиморфизмы.

Role of apoptosis genes in aggression revealed using combined analysis of ANDSystem gene networks, expression and genomic data in grey rats with aggressive behavior

A.O. Bragin¹✉, O.V. Saik¹, I.V. Chadaeva^{1,2}, P.S. Demenkov¹, A.L. Marke^{1,2}, Yu.L. Orlov^{1,2}, E.I. Pogayev^{1,3,4,5}, I.N. Lavrik^{1,6}, V.A. Ivanisenko¹

¹ Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

² Novosibirsk State University, Novosibirsk, Russia

³ University of Massachusetts Medical School, Worcester, USA

⁴ Institute of General Genetics RAS, Moscow, Russia

⁵ Center for Genetics and Genetic Technologies, Faculty of Biology, Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Moscow, Russia

⁶ Otto von Guericke University, Magdeburg, Germany

Aggressive behavior in animals plays an important role in protecting the territory, offspring, establishing social hierarchical relations, etc. Increased aggression is observed in a number of diseases (schizophrenia, bipolar disorder, brain degenerative disorders). Neuronal apoptosis is crucial in the maintenance of developmental processes during neurogenesis. Alterations in neuronal apoptosis are observed in aging and neuropathologies accompanied by changes in psycho-emotional state (epilepsy, Alzheimer's disease, neurotrauma). The expression of key neuronal apoptosis genes (*Casp3*, *Bax* and *Bcl-xl*) in the brain of highly aggressive rats is significantly altered. The aim of this work was to analyze associative networks that describe genetic interactions between genes/proteins involved in neuronal apoptosis, differentially expressed genes and genes with polymorphisms in grey rats with aggressive behavior. Analysis revealed 819 differentially expressed genes in the hypothalamus, ventral tegmental region and periaqueductus Sylvii grey matter in grey rats with aggressive and tame behavior. The *Stx1a*, *Mbp*

and *Th* genes have the highest index of betweenness centrality in the associative network of differentially expressed genes. Genome analysis revealed 137 polymorphic genes. Three of them (*Lig4*, *Parp1* and *Pigt*) were involved in neuronal apoptosis. It was shown that polymorphic and differentially expressed genes were statistically significantly overrepresented among genes interacting with neuronal apoptosis genes (*p*-value < 0.01). Three molecular-genetic chains describing connections between polymorphic and neuronal apoptosis genes mediated by differentially expressed genes were reconstructed. Chains included the polymorphic genes *Tsc1*, *Adamts4* and *Lgals3*, differentially expressed genes *Ezr*, *Acan*, *Th* and 19 neuronal apoptosis genes. It was shown that neuronal apoptosis is closely related to aggressive behavior in animals.

Key words: neuronal apoptosis; aggressive behavior; ANDSystem; associative gene networks; differentially expressed genes; polymorphisms.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Брагин А.О., Сайк О.В., Чадаева И.В., Деменков П.С., Маркель А.Л., Орлов Ю.Л., Робаев Е.И., Лаврик И.Н., Иванисенко В.А. Роль генов апоптоза в контроле агрессивного поведения, выявленная с помощью комбинированного анализа ассоциативных генных сетей, экспрессионных и геномных данных по серым крысам с агрессивным поведением. Вавиловский журнал генетики и селекции. 2017;21(8):911-919. DOI 10.18699/VJ17.312

HOW TO CITE THIS ARTICLE:

Bragin A.O., Saik O.V., Chadaeva I.V., Demenkov P.S., Markel A.L., Orlov Yu.L., Rogaev E.I., Lavrik I.N., Ivanisenko V.A. Role of apoptosis genes in aggression revealed using combined analysis of ANDSystem gene networks, expression and genomic data in grey rats with aggressive behavior. Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding. 2017;21(8):911-919. DOI 10.18699/VJ17.312 (in Russian)

Агрессивное поведение животных, как определил его один из основоположников этологии (науки о поведении) австрийский зоолог К. Лоренц (Lorenz, 2002), или инстинкт агрессии, направленной против собратьев по виду, у животных изучается уже давно и занимает ведущее место в научных исследованиях основ поведения. В настоящее время агрессию определяют как врожденную форму поведения и считают ее частью агонистического поведения (от греч. *agonistikos* – способный к борьбе, воинственный) (Хайнд, 1975). Выявление генетических детерминант агрессивного поведения весьма актуально, поскольку в человеческом обществе повышенная агрессивность – одна из основных проблем социальных отношений, здравоохранения и других сфер деятельности. Большинство исследователей генетики агрессивного поведения сходятся на том, что этот тип поведения имеет полигенную детерминацию (Craig, Halton, 2009). И кроме генов, непосредственно участвующих в поведенческих реакциях, есть генетические ансамбли, которые детерминируют агрессивность опосредованно, через различные метаболические каскады биохимических реакций в нейронах (Benus et al., 1991).

Многофакторная система формирования агрессивного поведения предполагает поиск и использование удобных и надежных генетических моделей, например модельных животных с повышенной и пониженной агрессивностью, у которых влияние внешних факторов сведено к минимуму. Логично предположить, что разница в уровне экспрессии генов у животных с повышенной агрессивностью по сравнению с неагрессивными (обе группы получены путем генетической селекции из одной и той же исходной популяции серых крыс-пасюков – *Rattus norvegicus*) позволит определить, какие гены участвуют в реализации агрессивных реакций. Такая модель была выведена в

Институте цитологии и генетики СО РАН путем селекции серых крыс-пасюков в двух противоположных направлениях – на повышение агрессивности и на толерантное поведение по отношению к человеку (Plyusnina, Oskina, 1997). Отбор проводился при одинаковых условиях содержания и единообразной методике оценки поведенческой реакции в отношении человека на протяжении более чем 80 поколений крыс. С помощью современных технологий высокопроизводительного секвенирования ДНК агрессивное поведение в этой модели изучается не только на фенотипическом уровне, который включает наблюдение и анализ конкретных поведенческих и нейроэндокринных реакций (Кожемякина и др., 2016), но и на геномном уровне, путем исследования конкретных генов, влияющих на повышенную или пониженную агрессивность особи. В сочетании с компьютерными технологиями реконструкции генных сетей, поиском информации в базах данных по экспрессии генов в отделах мозга результаты таких комплексных исследований становятся полноценными и достоверными для понимания феномена агрессивного поведения.

Нейрональный апоптоз выполняет важнейшую физиологическую роль в нервной системе при нейрогенезе. При этом происходит апоптотическая гибель большого количества избыточных нейронов и, таким образом, поддерживается баланс клеток в развивающейся нервной системе (Franklin, 2011). В работе (Ilchibaeva et al., 2016) показано, что у крыс с высокой степенью агрессии обнаруживается значительное увеличение экспрессии гена каспазы 3 в гипоталамусе, снижение уровня экспрессии проапоптотического гена *Bax* в гиппокампе и увеличение уровня мРНК антиапоптотического гена *Bcl-xl* в области ядер шва среднего мозга. Эти данные говорят в пользу того, что генетически обусловленная агрессия у крыс

связана со значительными изменениями регуляции нейронального апоптоза.

Целью работы был анализ ассоциативных сетей, описывающих молекулярно-генетические взаимодействия между белками и генами, вовлеченными в нейрональный апоптоз, дифференциально экспрессированными генами и генами, несущими полиморфизмы у серых крыс с агрессивным поведением. Выявление генов, вовлеченных в такие взаимодействия, может быть основой для разработки лекарственных препаратов, корректирующих агрессивное поведение.

Материалы и методы

Для построения списка дифференциально экспрессирующихся (ДЭ) и полиморфных генов у серых крыс с агрессивным поведением были использованы данные РНК-секвенирования трех отделов мозга (гипоталамус, вентральная тегментальная область среднего мозга и периакведуктум). Библиотеки чтений транскриптов от анализируемых крыс получены в результате секвенирования на платформе Illumina согласно протоколу NEBNext mRNA Library Prep Reagent Set for Illumina (NEB, США) и доступны на сайте The European Nucleotide Archive (<http://www.ebi.ac.uk/ena>) под индексом ERP011250. Всего проанализировано 12 библиотек, из них по две были получены для каждого из трех рассматриваемых отделов мозга крыс с агрессивным поведением, по две – для каждого изучаемого отдела мозга крыс с дружелюбным поведением. Картирование чтений на референсный геном крысы *Rattus norvegicus* (Rnor_6.0) осуществляли в программе STAR (Dobin et al., 2013), входные параметры использовали заданные по умолчанию. Референсный геном *Rattus norvegicus* был взят с сайта Ensembl (Yates et al., 2015). Дифференциально экспрессирующиеся гены между крысами с агрессивным и дружелюбным поведением выявляли при помощи пакета DESeq2 (Love et al., 2014), реализованного на языке программирования R, с учетом поправки на множественное сравнение ($p\text{-value} < 0.05$). Анализ сверхпредставленности Gene Ontology (GO) биологических процессов проводили веб-сервисом DAVID 6.8 (Huang et al., 2009; <https://david.ncifcrf.gov>), все настройки были взяты по умолчанию.

Поиск полиморфизмов в генах у серых крыс с агрессивным поведением проводили в пакете Samtools (Li, 2011). В качестве референсного использовали геном крысы *Rattus norvegicus* версия Rnor_6.0 из базы данных Ensembl (Yates et al., 2015). Программой VCFtools (Danecsek et al., 2011) были отобраны полиморфизмы, которые встречаются во всех шести образцах тканей мозга агрессивных крыс, но отсутствуют у дружелюбных особей. Среди выявленных полиморфизмов с помощью программы VEP (McLaren et al., 2016) были найдены миссенс-мутации.

Реконструкцию ассоциативных генных сетей, описывающих молекулярно-генетические взаимодействия между генами и белками, осуществляли с помощью системы ANDSystem (Demenkov et al., 2012; Ivanisenko et al., 2015). Система ANDSystem предназначена для автоматического анализа текстов научных публикаций с целью извлечения информации о взаимодействиях между молекулярно-генетическими объектами (гены, белки, метаболиты,

микро-РНК), а также их ассоциативных и регуляторных связей с заболеваниями, биологическими процессами и фенотипическими признаками различных организмов. Всего рассматривали 24 типа взаимодействий, включая физические взаимодействия (белок–белок, белок–ДНК, белок–лиганд), каталитические реакции, регуляторные связи (регуляция активности, экспрессии, транспорта, деградации), коэкспрессию, ассоциацию (интегральный тип связей, характеризующий наличие любого типа взаимодействий). Следует отметить, что тип связи «ассоциация» в ANDSystem используется для описания связей между парами объектов любого типа. Центральность вершин (betweenness centrality) в генной сети оценивалась с помощью функции `nx.betweenness_centrality` пакета Networkx, реализованного на языке программирования Python (Hagberg et al., 2008). Список генов, вовлеченных в нейрональный апоптоз (GO:0051402), был получен с помощью базы данных AmiGO 2 (Ashburner et al., 2000). Статистическую значимость сверхпредставленности набора генов в анализируемой выборке вычисляли с помощью программы GeneProf (<https://www.geneprof.org>), использующей гипергеометрическое распределение.

Поиск цепочек в глобальной ассоциативной генной сети ANDSystem по заданному шаблону проводился инструментом Pathway Wizard системы ANDSystem. Шаблон для поиска цепочек, связывающих полиморфные гены, ДЭ гены и нейрональный апоптоз, включал три вершины и задавался следующим образом: 1) первый участник цепочки – полиморфный ген; 2) второй участник – ДЭ ген; 3) третий участник – ген/белок, вовлеченный в нейрональный апоптоз; 4) тип связей между первым и вторым участником, а также между вторым и третьим участником цепочки может быть любой из числа представленных в системе ANDSystem.

Результаты и обсуждение

Анализ дифференциальной экспрессии и полиморфизмов генов серых крыс с агрессивным поведением

Было выявлено 819 генов, дифференциально экспрессирующихся у серых крыс с агрессивным поведением по сравнению с экспрессией генов у серых крыс с дружелюбным поведением. В том числе выявлено 45, 675 и 153 ДЭ гена в гипоталамусе, вентральной тегментальной области и периакведуктуме соответственно (Приложение 1)¹. Экспрессия 8, 242 и 90 генов у крыс с агрессивным поведением была повышена в гипоталамусе, вентральной тегментальной области и периакведуктуме соответственно по сравнению с таковой у дружелюбных крыс, а экспрессия 37, 433 и 63 – понижена. Три гена (*Ifi27*, *Mcm7* и *AY172581.17*) были дифференциально экспрессированы в трех анализируемых отделах мозга у серых крыс с агрессивным поведением. Одновременно в вентральной тегментальной области и периакведуктуме были дифференциально экспрессированы 32 гена, в гипоталамусе и вентральной тегментальной области – 17 генов, а в гипоталамусе и периакведуктуме – 8 генов.

¹ Приложения 1–4 см. по адресу: <http://www.bionet.nsc.ru/vogis/download/pict-2017-21/appx14.pdf>

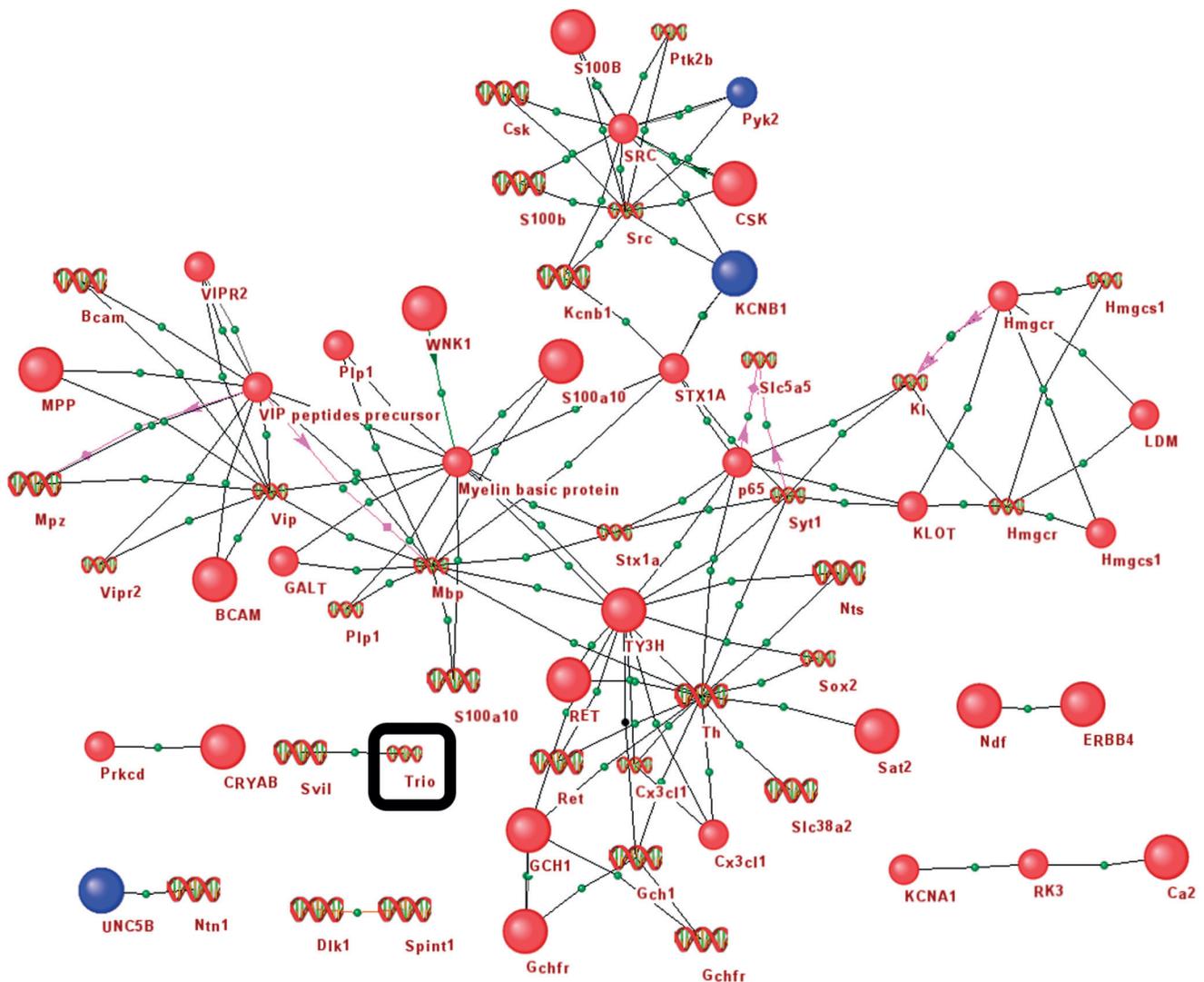


Fig. 1. Associative gene network of DE genes found to be directly interconnected according to ANDSystem. Large icons denote DE genes/proteins with upregulated expression in laboratory rats with aggressive behavior; small icons, DE genes/proteins with downregulated expression. Genes/proteins involved in neuronal apoptosis are highlighted in blue. The polymorphic *Trio* gene is enclosed in a bar. Genes are shown as DNA helices and proteins, as balls. Gray links indicate associations and protein–protein interactions; green links, catalytic interactions; and pink links, expression regulation.

Для ряда ДЭ генов в литературе присутствует информация по их связи с агрессивным поведением. Например, для гена тирозингидроксилазы (*Th*), экспрессия которого повышена в вентральной тегментальной области крыс с агрессивным поведением, в работе (Bondar et al., 2009) показано, что у самцов мышей, проявляющих агрессивное поведение, уровень экспрессии этого гена повышается. Экспрессия гена *Kirrel3* повышена в периакведуктуме у серых крыс с агрессивным поведением. Делеция *Kirrel3* приводит к снижению межсамцовой агрессии у мышей (Prince et al., 2013). Экспрессия гена вазоактивного кишечного пептида (*Vip*) была понижена в вентральной тегментальной области у серых крыс с агрессивным поведением. У певчих птиц введение пептида VIP непосредственно в боковую перегородку влияет на агрессивное поведение (Maney et al., 2005). Экспрессия гена нейротензина (*Nts*) была повышена в вентральной тегментальной области у серых крыс с агрессивным поведением. В работе (Gammie

et al., 2009) показано, что у самок мышей, которые проявляли высокие уровни материнской агрессии, уровень *Nts* был снижен, а интрацеребровентрикулярное введение *Nts* значительно снижало материнскую агрессию. Экспрессия гена синтазы оксида азота 1 (*Nos1*) была повышена в периакведуктуме у серых крыс с агрессивным поведением. Имеются данные о том, что ингибирование гена нейрональной NO-синтазы (*nNOS*) приводит к повышению агрессивности самцов мышей (Hotchkiss et al., 2005), в то время как в исследовании (Demas et al., 1999) показано, что делеция гена эндотелиальной NO-синтазы (*eNOS*) у самцов мышей снижает агрессию. Оказалось, что 819 дифференциально экспрессирующихся генов вовлечены в 14 сверхпредставленных Gene Ontology биологических процессов, в том числе в регуляцию трансмембранного транспорта ионов (GO:0034765), развитие нервной системы (GO:0007399), развитие мозга (GO:0007420), регуляцию секреции дофамина

(GO:0014059), положительную регуляцию развития проекции нейронов (GO:0010976) и др. (Приложение 2).

С помощью пакета Samtools (Li, 2011) выявлены 1136 полиморфизмов в 464 генах, которые были обнаружены во всех шести образцах тканей мозга агрессивных крыс. Анализ программой VEP (McLaren et al., 2016) показал, что среди этих 1136 полиморфизмов миссенс-мутациями являются 156 полиморфизмов в 141 гене. Аналогично для серых крыс с дружелюбным поведением выявлено 795 полиморфизмов в 338 генах, среди которых было 106 миссенс-мутаций в 97 генах (Приложение 3). Для дальнейшего анализа были отобраны 137 генов, несущих миссенс-мутации, которые встречаются во всех шести образцах тканей мозга агрессивных крыс, но отсутствуют у дружелюбных особей (полиморфные гены). Среди этих полиморфных генов двенадцать оказались ДЭ генами (*Frrs11*, *Gpr137b*, *Hmgcll1*, *Hps4*, *Mcm7*, *Nkd2*, *Npas2*, *Pdprn*, *Plod1*, *Sp110*, *Tnks1bp1*, *Trio*).

Для анализа взаимодействий ДЭ и полиморфных генов была построена ассоциативная генная сеть с помощью системы ANDSystem. Рассматривали только прямые связи между анализируемыми генами, без учета посредников. Ассоциативная генная сеть состояла из кластера 34 генов (рис. 1), имеющих связи хотя бы с одним из рассмотренных генов, и 910 ни с кем не связанных генов. В ассоциативной генной сети присутствовали 119 взаимодействий следующих типов: ассоциации, каталитические взаимодействия, регуляция экспрессии и белок-белковые взаимодействия (см. рис. 1). Оказалось, что с ДЭ генами имеет связи только один полиморфный ген (*Trio*), экспрессия которого была понижена у крыс с агрессивным поведением.

Анализ центральности вершин (betweenness centrality) показал, что наиболее центральными в генной сети, представленной на рис. 1, являются вершины *Stx1a*, *Mbp* и *Th*. Показатели их центральности составили 0.22, 0.18 и 0.17 соответственно. Синтаксин 1А (*Stx1a*), экспрессия которого была понижена в вентральной тегментальной области у крыс с агрессивным поведением, кодирует белок, участвующий в регуляции серотонинергических и ГАМК-ергических систем. Известно, что его экспрессия изменяется при аутизме, а ряд мутаций ассоциирован с синдромом Аспергера (Durdiaková et al., 2014). Обе эти патологии характеризуются эмоциональными нарушениями, трудностями в социальном взаимодействии и стереотипным, ограниченным, повторяющимся репертуаром интересов и действий. Экспрессия гена основного белка миелина (*Mbp*) была понижена в периакведуктуме у крыс с агрессивным поведением. По данным (Wang et al., 2004), значительное снижение уровня *Mbp* наблюдается у пациентов с болезнью Альцгеймера и нарушением когнитивных функций. Экспрессия гена тирозингидроксилазы (*Th*), связанного с агрессивным поведением у самцов мышей (Bondar et al., 2009), была повышена в вентральной тегментальной области серых крыс с агрессивным поведением. Участники генной сети, обладающие высоким показателем центральности, могут быть ключевыми регуляторами ДЭ генов и являются перспективными для дальнейшего изучения при поиске потенциальных мишеней для фармпрепаратов, направленных на снижение агрессивного поведения.

Анализ генов нейронального апоптоза

Согласно базе данных AmiGO 2 в Gene Ontology (GO), в биологический процесс нейронального апоптоза (Neuron Apoptotic Process, GO:0051402) у крыс вовлечены 276 генов. Среди 819 ДЭ генов в нейрональном апоптозе участвуют 23 гена (*Adarb1*, *Cdk5r1*, *Cit*, *Coro1a*, *Crlf1*, *Dlx1*, *Egln3*, *Egr1*, *En1*, *Epha7*, *Gabra5*, *Gclc*, *Kcnb1*, *Mef2c*, *Ngfr*, *Nrbp2*, *Nsmf*, *Pcp4*, *Prkcg*, *Ptk2b*, *Syngap1*, *Unc5b*, *Vegfb*). Например, на рис. 1 видно, что экспрессия гена *Unc5b* повышена. Согласно (Tang et al., 2008), *Unc5b* обладает проапоптотическим действием, и в отсутствие сигнальной молекулы нетрин-1 способен активировать нейрональный апоптоз. Для гена *Egln3*, экспрессия которого повышена в вентральной тегментальной области у серых крыс с агрессивным поведением, показано, что увеличение его активности приводит к нейрональному апоптозу (Schlisio, 2009). Экспрессия гена *Egr1* была повышена в периакведуктуме у серых крыс с агрессивным поведением. Известно, что повышение уровня экспрессии гена *Egr1* способно активировать нейрональный апоптоз (Xie et al., 2011). Экспрессия гена *Mef2c* была снижена в вентральной тегментальной области и повышена в периакведуктуме у серых крыс с агрессивным поведением. В работе (Okamoto et al., 2002) показано, что активация *Mef2c* вызывает снижение нейронального апоптоза. В статье (Mao et al., 1999) говорится о том, что, хотя во время развития нервной системы активация гена *Mef2c* приводит к снижению апоптоза, в зрелых нейронах при стрессах активация этого гена способствует повышению уровня апоптоза.

Среди полиморфных генов три (*Lig4*, *Parp1*, *Pigt*) участвуют в нейрональном апоптозе. Известно, что дефицит *Lig4* приводит к повышению уровня нейронального апоптоза (Sekiguchi et al., 2001). Расщепление *Parp1* каспазой 9 вызывает апоптоз, в то время как гиперактивация этого гена приводит к некрозу (Sairanen et al., 2009).

Интересной задачей был поиск прямых межмолекулярных взаимодействий ДЭ и полиморфных генов с генами, вовлеченными в нейрональный апоптоз. С помощью системы ANDSystem было реконструировано 478 взаимодействий между ДЭ генами и генами/белками, участвующими в нейрональном апоптозе (Приложение 4). Рассматривались такие типы взаимодействий, как регуляция активности, каталитические взаимодействия, регуляция деградации, регуляция экспрессии, белок-белковые взаимодействия, регуляция транспорта и ассоциации. Наибольшее число взаимодействий наблюдалось для ассоциаций и белок-белковых взаимодействий: 401 и 34 связи соответственно. Так, нокаут гена *Sox2*, экспрессия которого была снижена в вентральной тегментальной области у крыс с агрессивным поведением, приводит к снижению уровня экспрессии гена *ErbB3* (Koike et al., 2015). В свою очередь, повышенный нейрональный апоптоз наблюдался у мышей с нокаутом гена *ErbB3* (Narkis et al., 2007). Следует отметить, что количество генов/белков нейронального апоптоза, участвующих во взаимодействиях с ДЭ генами, составило 55 из 276, что существенно превышало число, ожидаемое по случайным причинам ($p\text{-value} < 10^{-8}$).

Для полиморфных генов с помощью ANDSystem было найдено 14 ассоциативных связей и 3 белок-белковых взаимодействия с генами/белками, вовлеченными в ней-

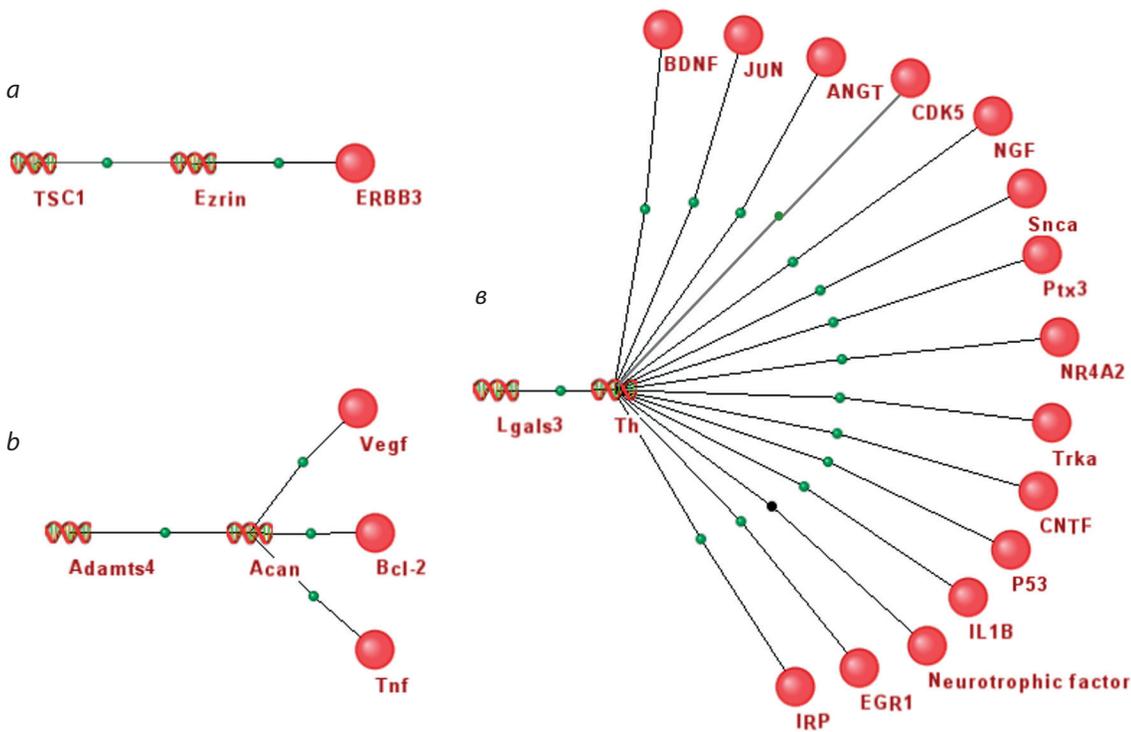


Fig. 2. Associative gene network describing the potential interactions of polymorphic genes, DE genes, and genes/proteins involved in neuronal apoptosis through molecular pathways reconstructed with the Pathway Wizard function of ANDSystem.

(a) The molecular pathway comprising the polymorphic *Tsc1* gene, the DE gene *Ezr*, and the *ErbB3* gene, involved in neuronal apoptosis. (b) The molecular pathway comprising the polymorphic gene *Adamts4*, the DE gene *Acan*, and three genes (*Vegf*, *Bcl-2*, *Tnf*) involved in neuronal apoptosis. (c) The molecular pathway comprising the polymorphic gene *Lgals3*, the DE gene *Th*, and 15 genes involved in neuronal apoptosis. Genes are shown as DNA helices, and proteins, as balls. Gray links indicate associative interactions.

рональный апоптоз (см. Приложение 4). Например, ингибирование полиморфного гена *Parp1* может приводить к снижению активации гена *GSK-3* (Songin et al., 2007). Интересно отметить, что ген *GSK-3* обладает как про-, так и антиапоптотическим действием, так как по-разному регулирует два основных пути апоптоза – внутренний и внешний. *GSK-3* активирует внутренний митохондриальный путь апоптоза, но ингибирует внешний путь апоптоза, опосредованный через рецепторы клеточной смерти (Beugle, Jore, 2006). Для полиморфных генов, так же как и в случае ДЭ генов, оказалось, что количество генов/белков нейронального апоптоза, участвующих во взаимодействиях с ними, превышало число, ожидаемое по случайным причинам, и составило 8 из 276. Согласно гипергеометрическому распределению, перепредставленность участников взаимодействий с полиморфными генами среди генов/белков нейронального апоптоза имела статистическую значимость ($p\text{-value} < 0.01$).

Участники генных сетей могут оказывать влияние на функцию других участников посредством прямых взаимодействий, а также путем передачи сигнала по цепочкам вершин сети, имеющих между собой соответствующие связи. В настоящей работе проведен поиск всех возможных цепочек из трех вершин, описывающих связи между полиморфными генами и генами/белками нейронального апоптоза, опосредованные через ДЭ гены, инструментом Pathway Wizard, который реализован в

системе ANDSystem. Таким образом были выявлены три цепочки, связывающие, согласно заданным требованиям, три полиморфных гена (*Adamts4*, *Lgals3*, *Tsc1*), три ДЭ гена (*Acan*, *Ezr*, *Th*) и 19 генов/белков, вовлеченных в нейрональный апоптоз (рис. 2).

На рис. 2, а представлена цепочка, включающая полиморфный ген *Tsc1*, ДЭ ген *Ezr* и ген *ErbB3*, вовлеченный в нейрональный апоптоз. Продукт гена *Tsc1* – гамартин, взаимодействует с белком – продуктом гена *Ezr*, экспрессия которого снижена в вентральной тегментальной области у крыс с агрессивным поведением (Narayanan, 2003). Эти белки участвуют в организации плазматической мембраны и актинового цитоскелета, а также адгезии и миграции клеток. Экспрессия гена *Ezr* регулируется белком ErbB3. У мышей с нокаутом гена *ErbB3* наблюдался повышенный нейрональный апоптоз (Narkis et al., 2007). Кроме того, было показано снижение уровня экспрессии гена *ErbB3* при нокауте гена *Sox2*, экспрессия которого была снижена в вентральной тегментальной области у крыс с агрессивным поведением (Koike et al., 2015).

Цепочка, включающая полиморфный ген *Adamts4*, ДЭ ген *Acan* и три гена (*Vegf*, *Bcl-2*, *Tnf*), вовлеченных в нейрональный апоптоз, представлена на рис. 2, б. Фермент агреканазы *Adamts4* расщепляет агрекан, продукт гена *Acan*, экспрессия которого была повышена в вентральной тегментальной области у крыс с агрессивным поведением (Kim et al., 2009). Известно, что *Vegf*, *Tnf* и антисенс

Bcl-2 способны снижать уровень экспрессии гена *Acan* (Klooster, Bernier, 2004), а повышение уровня экспрессии гена *Bcl-2* приводит к увеличению экспрессии гена *Acan* (Feng et al., 1999). Белки *Vegf* и *Bcl-2* способны снижать уровень нейронального апоптоза (Rong et al., 2012), а *Tnf* – активировать нейрональный апоптоз (Ryan et al., 2004).

Третья цепочка (см. рис. 2, в) включает полиморфный ген *Lgals3*, ДЭ ген *Th* и 15 генов, вовлеченных в нейрональный апоптоз (*Bdnf*, *Jun*, *Angt*, *Cdk5*, *Ngf*, *Snca*, *Ptx3*, *Nr4A2*, *Trka*, *Cntf*, *P53*, *IL1B*, *Gdnf*, *Egr1*, *Irp*). Ген *Lgals3* кодирует β-галактозид-связывающий белок галектин-3, который имеет коллагеноподобный домен, обеспечивающий взаимодействие данного белка с внеклеточным матриксом. Галектин-3 участвует во многих биологических процессах, включая врожденный иммунитет, клеточную адгезию, регуляцию функции Т-клеток и микроглии (Chip et al., 2017). Показано, что галектин-3 регулирует белок CREB, который является фактором транскрипции гена *Th*, и повышенная экспрессия гена *Lgals3* приводит к увеличению экспрессии гена *Th* (Wu et al., 2010). Известно, что *Jun*, *Angt*, *Cdk5*, *Ngf*, *Ptx3*, *Nr4A2*, *Trka*, *Cntf*, *P53*, *IL1B*, *Egr1* могут положительно регулировать экспрессию гена *Th* или активность его продукта (Kansu et al., 2004; Dorofeeva et al., 2013). *Gdnf* приводит к снижению уровня экспрессии гена *Th* (Georgievska et al., 2004). В работе (Fukuchi et al., 2010) показано, что *Bdnf* может активировать экспрессию гена *Th*, тогда как в статье (Berhow et al., 1995) говорится о том, что *Bdnf* снижает уровень *Th* в вентральной тегментальной области. Гены *Snca*, *Irp* могут коэкспрессироваться с *Th* (Wang et al., 2013).

Заключение

Анализ геномных и транскриптомных данных по лабораторным крысам с агрессивным и дружелюбным поведением с использованием реконструкции ассоциативных генных сетей позволил выявить набор генов, потенциально участвующих в регуляции нейронального апоптоза, посредством молекулярно-генетических взаимодействий с генами/белками, напрямую вовлеченными в нейрональный апоптоз. Число таких генов значительно превышало ожидаемое по случайным причинам (p -value < 0.01). Анализ молекулярно-генетических путей в ассоциативной генной сети, описывающих связи полиморфных и ДЭ генов с генами нейронального апоптоза, показал, что гены, представленные в данных путях, могут играть существенную роль в функционировании ассоциативной генной сети. В частности, участник одного из этих путей – ген *Th*, который, согласно литературным данным (Bondar et al., 2009), ассоциирован с агрессивным поведением, обладает высокой центральностью (betweenness centrality).

Таким образом, полученные результаты могут служить дополнительным свидетельством того, что процесс нейронального апоптоза играет важную роль в агрессивном поведении у животных.

Acknowledgments

The reconstruction and analysis of associative networks of neuronal apoptosis were supported by the Russian Science Foundation, project 14-44-00011 “Programmed cell death induced via death receptors: identification of apoptosis-

triggering molecular mechanisms by molecular simulation. The search for gene polymorphisms and differentially expressed genes in rats with aggressive behavior was supported by the Government of the Russian Federation, project 14.B25.31.0033 (decree 220 of April 9, 2010). Works at the SPF Vivarium Shared Access Center were supported by project RFMEFI62117X0015. The development of scripts and algorithms for the calculation of graph node centralities was supported by State Budgeted Project 0324-2016-0008.

Conflict of interest

The authors declare no conflict of interest.

References

- Ashburner M., Ball C.A., Blake J.A., Botstein D., Butler H., Cherry J.M., Davis A.P., Dolinski K., Dwight S.S., Eppig J.T., Harris M.A., Hill D.P., Issel-Tarver L., Kasarskis A., Lewis S., Matese J.C., Richardson J.E., Ringwald M., Rubin G.M., Sherlock G. Gene Ontology: tool for the unification of biology. *Nature Genetics*. 2000;25(1):25-29. DOI 10.1038/75556.
- Benus R.F., Bohus B., Koolhaas J.M., van Oortmerssen G.A. Behavioural differences between artificially selected aggressive and non-aggressive mice: response to apomorphine. *Behav. Brain Res.* 1991; 43(2):203-208. DOI 10.1016/S0166-4328(05)80072-5.
- Berhow M.T., Russell D.S., Terwilliger R.Z., Beitner-Johnson D., Self D.W., Lindsay R.M., Nestler E.J. Influence of neurotrophic factors on morphine- and cocaine-induced biochemical changes in the mesolimbic dopamine system. *Neuroscience*. 1995; 68(4):969-979.
- Beurel E., Jope R.S. The paradoxical pro- and anti-apoptotic actions of GSK3 in the intrinsic and extrinsic apoptosis signaling pathways. *Prog. Neurobiol.* 2006;79(4):173-189. DOI 10.1016/j.pneurobio.2006.07.006.
- Bondar N.P., Ul'yana A.B., Kovalenko I.L., Filipenko M.L., Kudryavtseva N.N. Molecular implications of repeated aggression: *Th*, *Dat1*, *Snca* and *Bdnf* gene expression in the VTA of victorious male mice. *PLoS ONE*. 2009; 4(1):e4190. DOI 10.1371/journal.pone.0004190.
- Chip S., Fernandez-Lopez D., Li F., Faustino J., Derugin N., Vexler Z.S. Genetic deletion of galectin-3 enhances neuroinflammation, affects microglial activation and contributes to sub-chronic injury in experimental neonatal focal stroke. *Brain Behav. Immun.* 2017;60:270-281. DOI 10.1016/j.bbi.2016.11.005.
- Craig I.W., Halton K.E. Genetics of human aggressive behaviour. *Human Genetics*. 2009;126(1):101-113. DOI 10.1007/s00439-009-0695-9.
- Danecek P., Auton A., Abecasis G., Albers C.A., Banks E., DePristo M.A., Handsaker R.E., Lunter G., Marth G.T., Sherry S.T., McVean G. The variant call format and VCFtools. *Bioinformatics*. 2011;27(15):2156-2158. DOI 10.1093/bioinformatics/btr330.
- Demas G.E., Kriegsfeld L.J., Blackshaw S., Huang P., Gammie S.C., Nelson R.J., Snyder S.H. Elimination of aggressive behavior in male mice lacking endothelial nitric oxide synthase. *J. Neurosci.* 1999; 19:RC30. DOI 0270-6474/99/190001.
- Demenkov P.S., Ivanisenko T.V., Kolchanov N.A., Ivanisenko V.A. ANDVisio: a new tool for graphic visualization and analysis of literature mined associative gene networks in the ANDSystem. *In Silico Biology*. 2012;11(3,4):149-161. DOI 10.3233/ISB-2012-0449.
- Dobin A., Davis C.A., Schlesinger F., Drenkow J., Zaleski C., Jha S., Batut P., Chaisson M., Gingeras T.R. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21. DOI 10.1093/bioinformatics/bts635.
- Dorofeeva N.A., Nikitina L.S., Glazova M.V., Kirillova O.D., Chernigovskaia E.V. Inactivation of p53 leads to enhancement of tyrosine hydroxylase biosynthesis in the dopaminergic brain neurons. *Zhurnal Evoliutsionnoi Biokhimii i Fiziologii*. 2013;49(2):137-143.
- Durdiaková J., Warrior V., Banerjee-Basu S., Baron-Cohen S., Chakrabarti B. *STX1A* and Asperger syndrome: a replication study. *Mol.*

- Autism. 2014;5:14. DOI 10.1186/2040-2392-5-14.
- Feng L., Balakir R., Precht P., Horton W.E. Bcl-2 regulates chondrocyte morphology and aggrecan gene expression independent of caspase activation and full apoptosis. *J. Cell. Biochem.* 1999;74(4):576-586. DOI 10.1002/(SICI)1097-4644(19990915)74:4<576::AID-JCB7>3.0.CO;2-N.
- Franklin J.L. Redox regulation of the intrinsic pathway in neuronal apoptosis. *Antioxid. Redox Signal.* 2011;14(8):1437-1448. DOI 10.1089/ars.2010.3596.
- Fukuchi M., Fujii H., Takachi H., Ichinose H., Kuwana Y., Tabuchi A., Tsuda M. Activation of tyrosine hydroxylase (TH) gene transcription induced by brain-derived neurotrophic factor (BDNF) and its selective inhibition through Ca²⁺ signals evoked via the N-methyl-d-aspartate (NMDA) receptor. *Brain Res.* 2010; 1366:18-26. DOI 10.1016/j.brainres.2010.10.034.
- Gammie S.C., D'Anna K.L., Gerstein H., Stevenson S.A. Neurotensin inversely modulates maternal aggression. *Neuroscience.* 2009; 158(4):1215-1223. DOI 10.1016/j.neuroscience.2008.11.045.
- Georgievska B., Carlsson T., Lacar B., Winkler C., Kirik D. Dissociation between short-term increased graft survival and long-term functional improvements in Parkinsonian rats overexpressing glial cell line-derived neurotrophic factor. *Eur. J. Neurosci.* 2004;20(11):3121-3130. DOI 10.1111/j.1460-9568.2004.03770.x.
- Hagberg A., Swart P., Chult D.S. Exploring network structure, dynamics, and function using NetworkX: Proc. of the 7th Python in Sci. Conf. (SciPy 2008). USA: Los Alamos National Laboratory (LANL), 2008;11-15.
- Hinde R. *Animal Behaviour: A Synthesis of Ethology and Comparative Psychology.* 2nd ed. N. Y., St. Louis, San Francisco, Dusseldorf, London, Mexico, Panama, Sydney, Toronto: McGraw-Hill, 1970. [Russ. ed.: Khaynd R. Povedenie zhivotnykh. Sintez etologii i sravnitelnoy psikhologii. Moscow: Mir Publ., 1975. (in Russian)]
- Hotchkiss A.K., Pyter L.M., Gation M.L., Wen J.C., Milman H.A., Nelson R.J. Aggressive behavior increases after termination of chronic sildenafil treatment in mice. *Physiol. Behav.* 2005;83(5):683-688. DOI 10.1016/j.physbeh.2004.09.005.
- Huang D.W., Sherman B.T., Lempicki R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 2009;4(1):44-57. DOI 10.1038/nprot.2008.211.
- Ilchibaeva T.V., Tsybko A.S., Kozhemyakina R.V., Naumenko V.S. Expression of apoptosis genes in the brain of rats with genetically defined fear-induced aggression. *Mol. Biol.* 2016;50(5):719-724. DOI 10.1134/S0026893316030079.
- Ivanisenko V.A., Saik O.V., Ivanisenko N.V., Tiys E.S., Ivanisenko T.V., Demenkov P.S., Kolchanov N.A. ANDSystem: an Associative Network Discovery System for automated literature mining in the field of biology. *BMC Syst. Biol.* 2015;9(Suppl.2):S2. DOI 10.1186/1752-0509-9-S2-S2.
- Kansy J.W., Daubner S.C., Nishi A., Sotogaku N., Lloyd M.D., Nguyen C., Lu L., Haycock J.W., Hope B.T., Fitzpatrick P.F., Bibb J.A. Identification of tyrosine hydroxylase as a physiological substrate for Cdk5. *J. Neurochem.* 2004;91(2):374-384. DOI 10.1111/j.1471-4159.2004.02723.x.
- Kim M.S., Kim J.H., Lee M.R., Kang J.H., Kim H.J., Ko H.M., Choi C.H., Jung J.Y., Koh J.T., Kim B.K., Oh H.K. Effects of alendronate on a disintegrin and metalloproteinase with thrombospondin motifs expression in the developing epiphyseal cartilage in rats. *Anat. Histol. Embryol.* 2009;38(2):154-160. DOI 10.1111/j.1439-0264.2008.00920.x.
- Klooster A.R., Bernier S.M. Tumor necrosis factor alpha and epidermal growth factor act additively to inhibit matrix gene expression by chondrocyte. *Arthritis. Res. Ther.* 2004;7(1):R127. DOI 10.1186/ar1464.
- Koike T., Wakabayashi T., Mori T., Hirahara Y., Yamada H. Sox2 promotes survival of satellite glial cells *in vitro*. *Biochem. Biophys. Res. Commun.* 2015;464(1):269-274. DOI 10.1016/j.bbrc.2015.06.141.
- Kozhemyakina R.V., Konoshenko M.Yu., Sakharov D.G., Smaigin D.A., Markel A.L. Comparative analysis of the behavior of wild Norway rats (*Rattus norvegicus*) and Norway rats selected for tolerant and aggressive behavior in the open field test. *Zhurnal vysshey nervnoy deyatelnosti im. I.P. Pavlova = I.P. Pavlov Journal of Higher Nervous Activity.* 2016;66(1):92-102. DOI 10.7868/S0044467716010093. (in Russian)
- Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics.* 2011;27(21):2987-2993. DOI 10.1093/bioinformatics/btr509.
- Lorenz K. *On Aggression.* Psychology Press, 2002.
- Love M.I., Huber W., Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014; 15(12):550. DOI 10.1186/s13059-014-0550-8.
- Maney D.L., Erwin K.L., Goode C.T. Neuroendocrine correlates of behavioral polymorphism in white-throated sparrows. *Horm. Behav.* 2005;48(2):196-206. DOI 10.1016/j.yhbeh.2005.03.004.
- Mao Z., Bonni A., Xia F., Nadal-Vicens M., Greenberg M.E. Neuronal activity-dependent cell survival mediated by transcription factor MEF2. *Science.* 1999;286(5440):785-790. DOI 10.1126/science.286.5440.785.
- McLaren W., Gil L., Hunt S.E., Riat H.S., Ritchie G.R., Thormann A., Flicek P., Cunningham F. The ensembl variant effect predictor. *Genome Biol.* 2016;17(1):122. DOI 10.1186/s13059-016-0974-4.
- Narayanan V. Tuberous sclerosis complex: genetics to pathogenesis. *Pediatr. Neurol.* 2003;29(5):404-409. DOI 10.1016/j.pediatrneurol.2003.09.002.
- Narkis G., Ofir R., Manor E., Landau D., Elbedour K., Birk O.S. Lethal congenital contractural syndrome type 2 (LCCS2) is caused by a mutation in *ERBB3* (*Her3*), a modulator of the phosphatidylinositol-3-kinase/Akt pathway. *Am. J. Hum. Genet.* 2007;81(3):589-595. DOI 10.1086/520770.
- Okamoto S.I., Li Z., Ju C., Schölzke M.N., Mathews E., Cui J., Salvesen G.S., Bossy-Wetzel E., Lipton S.A. Dominant-interfering forms of MEF2 generated by caspase cleavage contribute to NMDA-induced neuronal apoptosis. *Proc. Natl. Acad. Sci. USA.* 2002;99(6):3974-3979. DOI 10.1073/pnas.022036399.
- Plyusnina I., Oskina I. Behavioral and adrenocortical responses to open-field test in rats selected for reduced aggressiveness toward humans. *Physiol. Behav.* 1997;61(3):381-385. DOI 10.1016/S0031-9384(96)00445-3.
- Prince J.E., Brignall A.C., Cutforth T., Shen K., Cloutier J.F. Kirrel3 is required for the coalescence of vomeronasal sensory neuron axons into glomeruli and for male-male aggression. *Development.* 2013; 140(11):2398-2408. DOI 10.1242/dev.087262.
- Rong W., Wang J., Liu X., Jiang L., Wei F., Hu X., Han X., Liu Z. Naringin treatment improves functional recovery by increasing BDNF and VEGF expression, inhibiting neuronal apoptosis after spinal cord injury. *Neurochem. Res.* 2012;37(8):1615-1623. DOI 10.1007/s11064-012-0756-7.
- Ryan L.A., Peng H., Erichsen D.A., Huang Y., Persidsky Y., Zhou Y., Gendelman H.E., Zheng J. TNF-related apoptosis-inducing ligand mediates human neuronal apoptosis: links to HIV-1-associated dementia. *J. Neuroimmunol.* 2004;148(1):127-139. DOI 10.1016/j.jneuroim.2003.11.019.
- Sairanen T., Szepesi R., Karjalainen-Lindsberg M.L., Saksi J., Paetau A., Lindsberg P.J. Neuronal caspase-3 and PARP-1 correlate differentially with apoptosis and necrosis in ischemic human stroke. *Acta Neuropathol.* 2009;118(4):541-552. DOI 10.1007/s00401-009-0559-3.
- Schlisio S. Neuronal apoptosis by prolyl hydroxylation: implication in nervous system tumours and the Warburg conundrum. *J. Cell. Mol. Med.* 2009;13(10):4104-4112. DOI 10.1111/j.1582-4934.2009.00881.x.
- Sekiguchi J., Ferguson D.O., Chen H.T., Yang E.M., Earle J., Frank K., Whitlow S., Gu Y., Xu Y., Nussenzweig A., Alt F.W. Genetic interactions between ATM and the nonhomologous end-joining factors in genomic stability and development. *Proc. Natl. Acad. Sci. USA.* 2001;98(6):3243-3248. DOI 10.1073/pnas.051632098.

- Songin M., Jęško H., Czapski G., Adamczyk A., Strosznajder R.P. GSK-3 β and oxidative stress in aged brain. Role of poly (ADP). *Folia Neuropathol.* 2007;45(4):220-229.
- Tang X., Jang S.W., Okada M., Chan C.B., Feng Y., Liu Y., Luo S.W., Hong Y., Rama N., Xiong W.C., Mehlen P. Netrin-1 mediates neuronal survival through PIKE-L interaction with the dependence receptor UNC5B. *Nat. Cell Biol.* 2008;10(6):698. DOI 10.1038/ncb1732.
- Wang D.S., Bennett D.A., Mufson E.J., Mattila P., Cochran E., Dickson D.W. Contribution of changes in ubiquitin and myelin basic protein to age-related cognitive decline. *Neurosci. Res.* 2004;48(1):93-100. DOI 10.1016/j.neures.2003.10.002.
- Wang Y.C., He F., Ma J., Zhou D., Liang Y., Gong Y.X. Impacts of electroacupuncture on ubiquitin-proteasome system in rats with Parkinson's disease. *Zhongguo Zhen Jiu = Chinese Acupuncture & Moxibustion.* 2013;33(8):725-729.
- Wu L., Zhao Q., Zhu X., Peng M., Jia C., Wu W., Zheng J., Wu X.Z. A novel function of microRNA let-7d in regulation of galectin-3 expression in attention deficit hyperactivity disorder rat brain. *Brain Pathol.* 2010;20(6):1042-1054. DOI 10.1111/j.1750-3639.2010.00410.x.
- Xie B., Wang C., Zheng Z., Song B., Ma C., Thiel G., Li M. Egr-1 transactivates *Bim* gene expression to promote neuronal apoptosis. *J. Neurosci.* 2011;31(13):5032-5044. DOI 10.1523/JNEUROSCI.5504-10.2011.
- Yates A., Akanni W., Amode M.R., Barrell D., Billis K., Carvalho-Silva D., Cummins C., Clapham P., Fitzgerald S., Gil L., Girón C.G., Gordon L., Hourlier T., Hunt S.E., Janacek S.H., Johnson N., Juettemann T., Keenan S., Lavidas I., Martin F.J., Maurel T., McLaren W., Murphy D.N., Nag R., Nuhn M., Parker A., Patricio M., Pignatelli M., Rahtz M., Riat H.S., Sheppard D., Taylor K., Thormann A., Vullo A., Wilder S.P., Zadissa A., Birney E., Harrow J., Muffato M., Perry E., Ruffier M., Spudich G., Trevanion S.J., Cunningham F., Aken B.L., Zerbino D.R., Flicek P. Ensembl 2016. *Nucleic Acids Res.* 2015;44(D1):D710-D716. DOI 10.1093/nar/gkv1157.

Перспективные маркеры CIMP+ опухолей толстой кишки, выявленные на основе анализа данных ресурса TCGA

Г.С. Краснов, А.Д. Бениаминов, Р.А. Тычко, Г.А. Пузанов, Р.О. Новаковский, А.В. Кудрявцева, А.А. Дмитриев

Институт молекулярной биологии им. В.А. Энгельгардта Российской академии наук, Москва, Россия

CIMP+ (CpG-Island Methylator Phenotype) опухоли характеризуются плотным метилированием промоторных CpG-островков одновременно многих генов и представляют отдельную группу злокачественных новообразований толстой кишки. Несмотря на то что диагностика CIMP+ опухолей имеет значительную прогностическую ценность, до сих пор не разработано эффективного набора маркеров для их выявления. Для определения CpG-сайтов, уровень метилирования которых может быть использован для идентификации CIMP+ опухолей, с помощью созданного ранее приложения CrossHub нами проведен анализ профилей экспрессии и метилирования 297 образцов первичных опухолей и 38 парных к ним «условных норм» толстой кишки, представленных в базе проекта TCGA (The Cancer Genome Atlas). Разработан скоринг, учитывающий уровень метилирования CpG-сайтов, их расположение, а также уровень экспрессии соответствующих генов. Определено, что статус метилирования CpG-сайтов, относящихся к генам *AMOTL1*, *ZNF43*, *ZNF134* и *CHFR*, является перспективным маркером CIMP+ опухолей. Более того, идентифицированы конкретные районы промоторных областей этих генов, уровень метилирования которых ассоциирован с исследуемым фенотипом. Для валидации полученных данных на независимой выборке, сначала оценили относительный уровень мРНК генов *AMOTL1*, *ZNF43*, *ZNF134* и *CHFR* в 30 парных (опухоль/«условная норма») образцах толстой кишки методом количественной ПЦР. Для всех генов выявлено частое (50–60 % случаев) и значительное (2–30 раз) снижение экспрессии. Затем методом бисульфитной конверсии ДНК с последующим клонированием и секвенированием исследовали статус метилирования CpG-сайтов, отобранных в результате биоинформатического анализа, и обнаружили высокий уровень метилирования (β -value = 0.3–0.9) в образцах с одновременно сниженным уровнем экспрессии всех четырех генов и низкий уровень метилирования (β -value = 0.0–0.2) в образцах с неизменным уровнем экспрессии четырех генов и в «условных нормах». Таким образом, статус метилирования CpG-сайтов промоторных областей генов *AMOTL1*, *ZNF43*, *ZNF134* и *CHFR* является перспективным потенциальным маркером CIMP+ опухолей толстой кишки.

Ключевые слова: рак толстой кишки; CIMP+; эпигенетика; метилирование ДНК; TCGA; количественная ПЦР; молекулярные маркеры.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Краснов Г.С., Бениаминов А.Д., Тычко Р.А., Пузанов Г.А., Новаковский Р.О., Кудрявцева А.В., Дмитриев А.А. Перспективные маркеры CIMP+ опухолей толстой кишки, выявленные на основе анализа данных ресурса TCGA. Вавиловский журнал генетики и селекции. 2017;21(8):920-924. DOI 10.18699/VJ17.313

HOW TO CITE THIS ARTICLE:

Krasnov G.S., Beniaminov A.D., Tychko R.A., Puzanov G.A., Novakovskiy R.O., Kudryavtseva A.V., Dmitriev A.A. Promising markers of CIMP+ colon tumors identified on the basis of TCGA data analysis. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2017;21(8):920-924. DOI 10.18699/VJ17.313 (in Russian)

Received 11.09.2017
Accepted for publication 23.11.2017
© AUTHORS, 2017

e-mail: Alex_245@mail.ru

Promising markers of CIMP+ colon tumors identified on the basis of TCGA data analysis

G.S. Krasnov, A.D. Beniaminov, R.A. Tychko, G.A. Puzanov, R.O. Novakovskiy, A.V. Kudryavtseva, A.A. Dmitriev

Engelhardt Institute of Molecular Biology RAS, Moscow, Russia

CIMP+ (CpG-Island Methylator Phenotype) tumors are characterized by dense methylation of promoter CpG islands of many genes at once and represent a separate group of malignant neoplasms of the colon. Despite the fact that the diagnostics of CIMP+ tumors has a significant prognostic value, an effective set of markers has not been developed yet. For the identification of CpG sites, the methylation level of which could be used to detect CIMP+ tumors, an analysis of expression and methylation profiles of 297 primary colon tumors and 38 histologically normal tissues paired to them, which are presented in the TCGA (The Cancer Genome Atlas) project database, was performed by us using the CrossHub tool created previously. We developed the scoring, which takes into account the methylation level of CpG sites, their location, and the expression level of the corresponding genes. It was revealed that the methylation status of CpG sites of the *AMOTL1*, *ZNF43*, *ZNF134*, and *CHFR* genes is a promising marker of CIMP+ tumors. Moreover, specific regions of promoters of these genes, the methylation level of which was associated with the examined phenotype, were identified. To verify the obtained data in independent sampling, first, the quantitative PCR was used to assess the relative mRNA level of the *AMOTL1*, *ZNF43*, *ZNF134*, and *CHFR* genes in 30 paired (tumor/histologically normal tissue) colon samples. For all the genes, a frequent (50–60 % of cases) and significant (2–30-fold) expression decrease was revealed. Then, the bisulfite conversion of DNA followed by cloning and sequencing was applied to examine the methylation status of CpG sites that were selected as the result of bioinformatics analysis. We observed a high methylation level (β -value = 0.3–0.9) of the CpG sites in the samples with simultaneous downregulation of all 4 genes and a low methylation level (β -value = 0.0–0.2) in the samples with the unchanged expression level of 4 genes and in histologically normal tissues. Thus, the methylation status of the CpG sites of promoter regions of the *AMOTL1*, *ZNF43*, *ZNF134*, and *CHFR* genes is a promising potential marker of CIMP+ colon tumors.

Key words: colon cancer; CIMP+; epigenetics; DNA methylation; TCGA; quantitative PCR; molecular markers.

Фенотип CIMP+ (CpG island methylator phenotype) отражает особый тип онкогенеза, который проявляется в плотном гиперметиловании одновременно множества CpG-островков в регуляторных (чаще всего промоторных) областях генома (Rhee et al., 2017). Метилирование и последующая супрессия распространяются на такие гены, как *CDKN2A (p16)*, *MLH1* и др., которые в нормальных тканях метилированию не подвергаются (Barzily-Rokni et al., 2011). Впервые фенотип CIMP+ был обнаружен при раке толстой кишки (Toyota et al., 1999), но в последние годы данный термин все чаще используют для описания интенсивного гиперметилования промоторов генов и в других видах рака, включая рак легкого, молочной железы, мочевого пузыря, матки, желудка, яичника, поджелудочной железы, почек и предстательной железы (Suzuki et al., 2014). Проведенные нами ранее исследования нарушений паттернов метилирования генов хромосомы 3 с применением NotI-микрочипов позволили также предположить существование опухолей CIMP+ (20–30 % опухолей) при раке легкого, яичника, шейки матки и почки (Dmitriev et al., 2012, 2014; Kashuba et al., 2012; Senchenko et al., 2013). Молекулярные и клинико-патологические особенности этого типа опухолей позволяют считать, что возникновение различных CIMP+ опухолей обусловлено единым молекулярным механизмом (Suzuki et al., 2014). Однако причины, ответственные за образование данного фенотипа, пока не установлены.

Ранняя диагностика CIMP+ опухолей имеет значительную прогностическую ценность (Hughes et al., 2013). Кроме того, для некоторых видов рака показано, что CIMP+ опухоли проявляют чувствительность к деметилирующим препаратам и ингибиторам EZH2 (Suzuki et al., 2014). Необходимо персонализированное назначение препаратов на основе деметилирующих агентов, так как их нецелевое использование крайне опасно и может стимулировать прогрессию заболевания (Mani, Herceg, 2010). Основной целью большинства исследований является попытка установить наличие CIMP+ опухолей в конкретных видах рака. При этом наборы маркерных генов для определения данного фенотипа различаются в разных работах, что приводит к серьезным затруднениям при анализе результатов (Suzuki et al., 2014).

Целью нашего исследования было определение перспективных маркеров для диагностики CIMP+ опухолей толстой кишки на основе комплексного подхода, включающего анализ современных транскриптомных и метиломных баз данных (таких как TCGA – The Cancer Genome Atlas) и последующую экспериментальную валидацию на выборке первичных опухолей.

Материалы и методы

Биоинформатический анализ. С помощью разработанного ранее приложения CrossHub (Krasnov et al., 2016) был проведен анализ материалов проекта TCGA (<https://cancergenome.nih.gov/>) для 297 образцов первичных опухолей и 38 парных к ним гистологически нормальных тканей («условных норм») толстой кишки (Cancer Genome Atlas, 2012). Использовали данные о статусе метилирования ДНК, полученные с помощью микрочипов Infinium HumanMethylation450 (Illumina, США), и данные об уров-

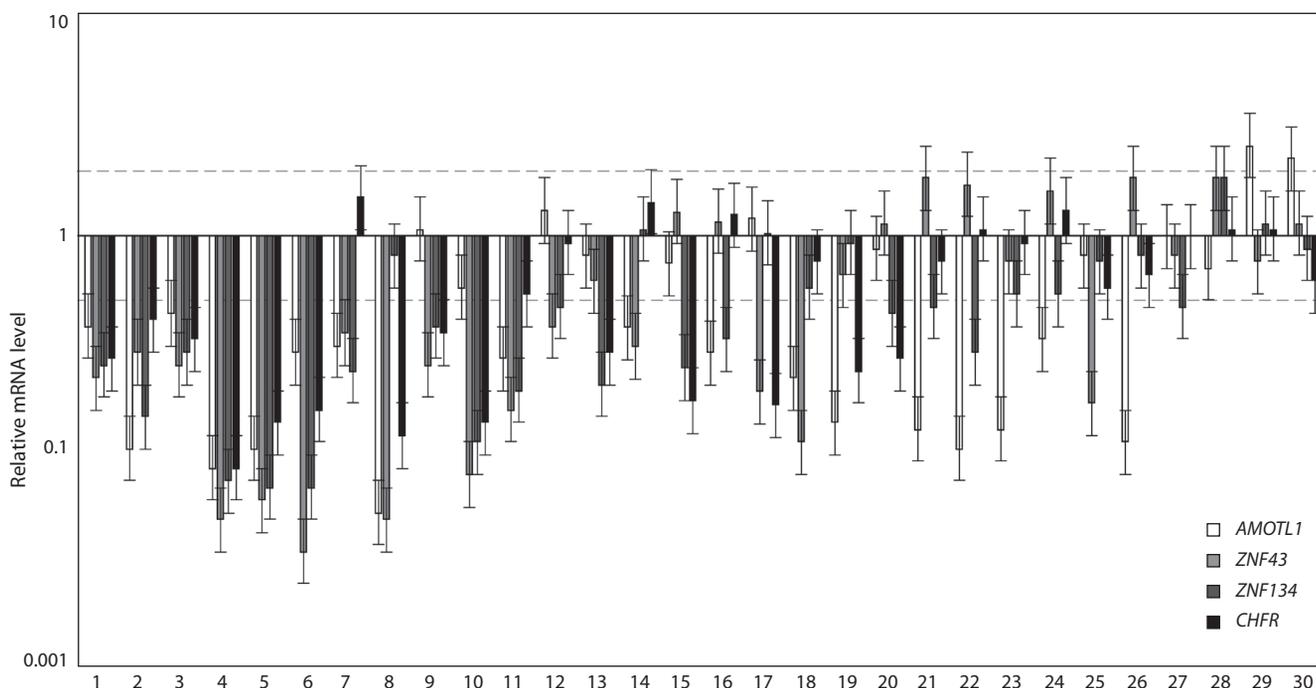
не экспрессии генов, полученные с помощью высокопроизводительных секвенаторов серии HiSeq (Illumina). Для установления ассоциации CpG-сайтов с промоторными областями генов воспользовались материалами проекта ENCODE для шести клеточных линий.

Образцы тканей. Операционные образцы опухолевых тканей толстой кишки и тканей «условной нормы» собраны и охарактеризованы в Московском научно-исследовательском онкологическом институте (МНИОИ) им. П.А. Герцена. Цели и осуществление данной работы одобрены этическим комитетом МНИОИ им. П.А. Герцена. Опухолевые клетки в каждом образце рака толстой кишки составляли не менее 70 %. Образцы тканей замораживали в жидком азоте и хранили при температуре –70 °С.

Выделение РНК, ДНК, получение кДНК. Замороженные в жидком азоте образцы опухолевых и парных к ним нормальных тканей толстой кишки подвергались механической гомогенизации на дисмембраторе Mikro-Dismembrator S (Sartorius, Германия). Для выделения РНК и ДНК использовали наборы RNeasy Mini Kit и QIAamp DNA Mini Kit (фирма Qiagen, Германия). Оценка количества РНК проводилась на флуориметре Qubit 2.0 (Invitrogen, США), оценка качества – на биоанализаторе Agilent 2100 Bioanalyzer (Agilent Technologies, США). Параметр RIN (RNA integrity number – показатель качества РНК) для образцов, включенных в исследования, составлял не менее 7. Концентрацию ДНК определяли на спектрофотометре NanoDrop 1000 (Nanodrop, США). Качество выделенной ДНК проверяли электрофорезом в 2 % агарозном геле в присутствии бромистого этидия, а также по значению отношения 260/280 нм, которое составляло 1.8–2.0. Выделенную РНК обрабатывали ДНКазой I (Thermo Fisher Scientific, США). Для проведения реакции обратной транскрипции применяли обратную транскриптазу M-MuLV (Thermo Fisher Scientific) и набор случайных гексануклеотидов.

Количественная ПЦР (кПЦР). Для количественной оценки уровня мРНК целевых и контрольных генов использовали наборы праймер-зонд TaqMan Gene Expression Assays (Thermo Fisher Scientific, США). В качестве контрольных выбраны гены *RPNI* и *PUM1* (Krasnov et al., 2011). С учетом биологической вариабельности уровня мРНК контрольных генов, при оценке экспрессии целевых генов значимыми считали изменения в 2 раза и более. Количественную ПЦР проводили на приборе Applied Biosystems 7500 Real-Time PCR System (Thermo Fisher Scientific) в трех технических повторностях по протоколу, описанному нами ранее (Dmitriev et al., 2016). Данные кПЦР обрабатывали при помощи оригинального программного приложения АТГ (Melnikova et al., 2016). Относительный уровень мРНК оценивали методом относительных измерений ($\Delta\Delta Ct$ -метод) с учетом эффективности реакций, как описано ранее (Dmitriev et al., 2016).

Анализ метилирования промоторных областей целевых генов проводили методом бисульфитной конверсии ДНК с последующим клонированием и секвенированием по Сэнгеру. ДНК обрабатывали бисульфитом при помощи набора EZ DNA Methylation Kit (Zymo Research, США). Продукты амплификации вырезали из 2 % агарозного геля и очищали с помощью набора DNA Clean & Concentrator-5



Relative levels of *AMOTL1*, *ZNF43*, *ZNF134*, and *CHFR* mRNAs in 30 colon cancer samples.

Quantitative PCR data are presented in the logarithmic scale. Dashed lines indicate the range of twofold mRNA level changes, and values within the range are considered the preservation of expression.

Methylation level of the selected CpG sites of the *AMOTL1*, *ZNF43*, *ZNF134*, and *CHFR* promoter regions in colon cancer samples

Gene	Number of CpG sites	Methylation level of CpG sites (β -value)		
		Tumor, samples 1–6	Tumor, samples 27–30	Conventional norm, samples 1–6, 27–30
<i>AMOTL1</i>	8	0.3–0.9	0.0–0.2	0.0–0.2
<i>ZNF43</i>	4	0.3–0.9	0.0–0.2	0.0–0.2
<i>ZNF134</i>	7	0.3–0.9	0.0–0.2	0.0–0.2
<i>CHFR</i>	6	0.3–0.9	0.0–0.2	0.0–0.2

лирования сайтов, статус метилирования которых может быть использован в качестве маркера CIMP+ опухолей толстой кишки (β -value = 0.3–0.9).

Диагностика CIMP+ опухолей толстой кишки необходима в силу ее высокой прогностической значимости (Lee et al., 2017), однако до сих пор нет унифицированного надежного набора маркеров, что вызывает существенные сложности при интерпретации результатов из различных исследований (Berg et al., 2014). В представленной работе проведен комплексный анализ данных ресурса TCGA, позволивший идентифицировать CpG-сайты, гиперметилирование которых функционально и ассоциировано с CIMP+ фенотипом опухолей толстой кишки. Полученный результат валидирован на независимой выборке первичных опухолей толстой кишки и парных к ним «условных норм». Таким образом, показано, что уровень метилирования CpG-сайтов промоторных областей генов *AMOTL1*, *ZNF43*, *ZNF134* и *CHFR* является перспективным маркером CIMP+ опухолей толстой кишки. Для определения клинической значимости предложенных маркеров

необходимы дальнейшие масштабные исследования на первичных опухолях.

Acknowledgments

This study was supported by the Russian Foundation for Basic Research and the Moscow government, project 15-34-70055 mol_a_Mos. Experiments were performed at the Genome Shared Access Center, Institute of Molecular Biology, Russian Academy of Sciences (http://www.eimb.ru/rus/ckp/ccu_genome_c.php).

Conflict of interest

The authors declare no conflict of interest.

References

Barzily-Rokni M., Friedman N., Ron-Bigger S., Isaac S., Michlin D., Eden A. Synergism between DNA methylation and macroH2A1 occupancy in epigenetic silencing of the tumor suppressor gene p16(CDKN2A). *Nucleic Acids Res.* 2011;39(4):1326-1335. DOI 10.1093/nar/gkq994.

- Berg M., Hagland H.R., Soreide K. Comparison of CpG island methylator phenotype (CIMP) frequency in colon cancer using different probe- and gene-specific scoring alternatives on recommended multi-gene panels. *PLoS ONE*. 2014;9(1):e86657. DOI 10.1371/journal.pone.0086657.
- Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012;487(7407):330-337. DOI 10.1038/nature11252.
- Couderec C., Boin A., Fuhrmann L., Vincent-Salomon A., Mandati V., Kieffer Y., Mehta-Grigoriou F., Del Maestro L., Chavrier P., Valleraud D., Brito I., Dubois T., De Koning L., Bouvard D., Louvard D., Gautreau A., Lallemand D. AMOTL1 promotes breast cancer progression and is antagonized by Merlin. *Neoplasia*. 2016;18(1):10-24. DOI 10.1016/j.neo.2015.11.010.
- Derks S., Cleven A.H., Melotte V., Smits K.M., Brandes J.C., Azad N., van Criekinge W., de Bruine A.P., Herman J.G., van Engeland M. Emerging evidence for CHFR as a cancer biomarker: from tumor biology to precision medicine. *Cancer Metastasis Rev*. 2014;33(1):161-171. DOI 10.1007/s10555-013-9462-4.
- Dmitriev A.A., Kashuba V.I., Haraldson K., Senchenko V.N., Pavlova T.V., Kudryavtseva A.V., Anedchenko E.A., Krasnov G.S., Pronina I.V., Loginov V.I., Kondratieva T.T., Kazubskaya T.P., Braga E.A., Yenamandra S.P., Ignatjev I., Ernberg I., Klein G., Lerman M.I., Zabarovsky E.R. Genetic and epigenetic analysis of non-small cell lung cancer with NotI-microarrays. *Epigenetics*. 2012;7(5):502-513. DOI 10.4161/epi.19801.
- Dmitriev A.A., Krasnov G.S., Rozhmina T.A., Kishlyan N.V., Zyablitsin A.V., Sadritdinova A.F., Snezhkina A.V., Fedorova M.S., Yurkevich O.Y., Muravenko O.V., Bolsheva N.L., Kudryavtseva A.V., Melnikova N.V. Glutathione S-transferases and UDP-glycosyltransferases are involved in response to aluminum stress in flax. *Front. Plant Sci*. 2016;7:1920. DOI 10.3389/fpls.2016.01920.
- Dmitriev A.A., Rudenko E.E., Kudryavtseva A.V., Krasnov G.S., Gordiyuk V.V., Melnikova N.V., Stakhovskiy E.O., Kononenko O.A., Pavlova L.S., Kondratieva T.T., Alekseev B.Y., Braga E.A., Senchenko V.N., Kashuba V.I. Epigenetic alterations of chromosome 3 revealed by NotI-microarrays in clear cell renal cell carcinoma. *BioMed Res. Int*. 2014;2014:735292. DOI 10.1155/2014/735292.
- Hughes L.A., Melotte V., de Schrijver J., de Maat M., Smit V.T., Bovee J.V., French P.J., van den Brandt P.A., Schouten L.J., de Meyer T., van Criekinge W., Ahuja N., Herman J.G., Weijenberg M.P., van Engeland M. The CpG island methylator phenotype: what's in a name? *Cancer Res*. 2013;73(19):5858-5868. DOI 10.1158/0008-5472.CAN-12-4306.
- Jen J., Wang Y.C. Zinc finger proteins in cancer progression. *J. Biomed. Sci*. 2016;23(1):53. DOI 10.1186/s12929-016-0269-9.
- Kashuba V., Dmitriev A.A., Krasnov G.S., Pavlova T., Ignatjev I., Gordiyuk V.V., Gerashchenko A.V., Braga E.A., Yenamandra S.P., Lerman M., Senchenko V.N., Zabarovsky E. NotI microarrays: novel epigenetic markers for early detection and prognosis of high grade serous ovarian cancer. *Int. J. Mol. Sci*. 2012;13(10):13352-13377. DOI 10.3390/ijms131013352.
- Krasnov G.S., Dmitriev A.A., Melnikova N.V., Zaretsky A.R., Nasedkina T.V., Zasedatelev A.S., Senchenko V.N., Kudryavtseva A.V. CrossHub: a tool for multi-way analysis of The Cancer Genome Atlas (TCGA) in the context of gene expression regulation mechanisms. *Nucleic Acids Res*. 2016;44(7):e62. DOI 10.1093/nar/gkv1478.
- Krasnov G.S., Oparina N.Y., Dmitriev A.A., Kudryavtseva A.V., Anedchenko E.A., Kondrat'eva T.T., Zabarovsky E.R., Senchenko V.N. *RPNI*, a new reference gene for quantitative data normalization in lung and kidney cancer. *Mol. Biol*. 2011;45(2):211-220. DOI 10.1134/S0026893311020129.
- Lee M.S., Menter D.G., Kopetz S. Right versus left colon cancer biology: integrating the consensus molecular subtypes. *J. Natl. Compr. Canc. Netw*. 2017;15(3):411-419.
- Mani S., Herceg Z. DNA demethylating agents and epigenetic therapy of cancer. *Adv. Genet*. 2010;70:327-340. DOI 10.1016/B978-0-12-380866-0.60012-5.
- Melnikova N.V., Dmitriev A.A., Belenikin M.S., Koroban N.V., Spersanskaya A.S., Krinitsina A.A., Krasnov G.S., Lakunina V.A., Snezhkina A.V., Sadritdinova A.F., Kishlyan N.V., Rozhmina T.A., Klimina K.M., Amosova A.V., Zelenin A.V., Muravenko O.V., Bolsheva N.L., Kudryavtseva A.V. Identification, expression analysis, and target prediction of flax genotroph microRNAs under normal and nutrient stress conditions. *Front. Plant Sci*. 2016;7:399. DOI 10.3389/fpls.2016.00399.
- Rhee Y.Y., Kim K.J., Kang G.H. CpG island methylator phenotype-high colorectal cancers and their prognostic implications and relationships with the serrated neoplasia pathway. *Gut Liver*. 2017;11(1):38-46. DOI 10.5009/gnl15535.
- Sanbhnani S., Yeong F.M. CHFR: a key checkpoint component implicated in a wide range of cancers. *Cell. Mol. Life Sci*. 2012;69(10):1669-1687. DOI 10.1007/s00018-011-0892-2.
- Senchenko V.N., Kisseljova N.P., Ivanova T.A., Dmitriev A.A., Krasnov G.S., Kudryavtseva A.V., Panasenko G.V., Tsitrin E.B., Lerman M.I., Kissel'ov F.L., Kashuba V.I., Zabarovsky E.R. Novel tumor suppressor candidates on chromosome 3 revealed by NotI-microarrays in cervical cancer. *Epigenetics*. 2013;8(4):409-420. DOI 10.4161/epi.24233.
- Suzuki H., Yamamoto E., Maruyama R., Niinuma T., Kai M. Biological significance of the CpG island methylator phenotype. *Biochem. Biophys. Res. Commun*. 2014;455(1-2):35-42. DOI 10.1016/j.bbrc.2014.07.007.
- Toyota M., Ahuja N., Ohe-Toyota M., Herman J.G., Baylin S.B., Issa J.P. CpG island methylator phenotype in colorectal cancer. *Proc. Natl. Acad. Sci. USA*. 1999;96(15):8681-8686.
- Zheng Y., Vertuani S., Nystrom S., Audebert S., Meijer I., Tegnebratt T., Borg J.P., Uhlen P., Majumdar A., Holmgren L. Angiomotin-like protein 1 controls endothelial polarity and junction stability during sprouting angiogenesis. *Circ. Res*. 2009;105(3):260-270. DOI 10.1161/CIRCRESAHA.109.195156.



Особенности экспрессионного профиля больных злокачественной меланомой в зависимости от ответа на иммунотерапию

Е.Н. Лукьянова¹, М.С. Федорова¹, Е.А. Пудова¹, Т.В. Наседкина^{1, 2}, Е.В. Степанова², К.М. Ньюшко³, А.Ю. Попов⁴, Н.В. Коробан³, А.А. Дмитриев¹, М.В. Киселева³, А.В. Липатова¹, А.С. Заседателев^{1, 2}, А.В. Кудрявцева^{1, 3}

¹ Институт молекулярной биологии им. В.А. Энгельгардта Российской академии наук, Москва, Россия

² Национальный медицинский исследовательский центр онкологии им. Н.Н. Блохина Министерства здравоохранения Российской Федерации, Москва, Россия

³ Национальный медицинский исследовательский центр радиологии Министерства здравоохранения Российской Федерации, Москва, Россия

⁴ Городская клиническая больница им. Д.Д. Плетнева Департамента здравоохранения города Москвы, Москва, Россия

Одним из важнейших направлений современной молекулярной генетики и биомедицины является поиск предиктивных маркеров, помогающих выбрать наиболее эффективный способ лечения и препарат, а также индивидуально подобрать его дозировку. Особенно перспективными считаются те маркеры, которые могут обеспечить возможность применения неинвазивной, так называемой жидкостной биопсии. Этот метод позволяет оценить состояние опухоли по анализу естественных жидкостей организма, таких как кровь, моча или слюна. Подобные исследования наиболее удобны в случаях, когда необходимо проводить мониторинг эффективности терапии, чтобы вовремя зафиксировать момент возникновения резистентности опухолевых клеток, начало рецидива и перейти на следующую линию терапии. При лечении агрессивных и быстро метастазирующих злокачественных новообразований, таких как меланома, наличие достоверных маркеров, позволяющих быстро и точно определить тактику лечения, особенно важно. В последнее время появляется все больше исследований, посвященных поиску предиктивных маркеров эффективности иммунотерапии. Меланома характеризуется высокой иммуногенностью, в связи с чем она стала модельным объектом для исследования и внедрения новых подходов иммунотерапии. В настоящей работе проведено сравнение двух групп больных метастатической меланомой кожи с различным ответом на иммунотерапию блокаторами иммунных контрольных точек с целью идентифицировать новые предиктивные экспрессионные биомаркеры среди микро- и матричных РНК, а также определить гены, ответственные за возникновение объективного ответа на терапию у пациентов. В исследовании выявлено несколько микроРНК со значительным изменением содержания в опухолевой ткани пациентов, по-разному отвечающих на иммунотерапию. Обнаружены также различия на уровне экспрессии их генов-мишеней, которые позволят более детально проанализировать молекулярные механизмы, определяющие чувствительность или резистентность клеток злокачественной меланомы к действию иммунотерапии. На основе полученных результатов предложены экспрессионные маркеры (мРНК и микроРНК), которые после дальнейшей апробации могут быть использованы как предикторы ответа опухолей злокачественной меланомы к иммунотерапии.

Ключевые слова: меланома; иммунотерапия; предиктивные маркеры; профилирование микроРНК; высокопроизводительное секвенирование; транскриптомный анализ; регуляция экспрессии генов.

Differences in expression profiles in malignant melanoma patients according to immunotherapy response

E.N. Lukyanova¹, M.S. Fedorova¹, E.A. Pudova¹, T.V. Nasedkina^{1, 2}, E.V. Stepanova², K.M. Nyushko³, A.Y. Popov⁴, N.V. Koroban³, A.A. Dmitriev¹, M.V. Kiseleva³, A.V. Lipatova¹, A.S. Zasedatelev^{1, 2}, A.V. Kudryavtseva^{1, 3}

¹ Engelhardt Institute of Molecular Biology RAS, Moscow, Russia

² Blokhin National Cancer Research Center, Ministry of Health of the Russian Federation, Moscow, Russia

³ National Medical Research Radiological Center, Ministry of Health of the Russian Federation, Moscow, Russia

⁴ Pletnev State Hospital, Department of Health of the Moscow, Moscow, Russia

One of the most important branch of modern molecular genetics and biomedicine is the search for predictive markers that help choose the most effective way of treatment, drug and also determine its individual dosage. Among the markers, those that can provide the possibility of using a non-invasive, so-called "liquid biopsy" are considered particularly promising. This method allows the condition of the tumor to be assessed by analyzing the body's natural fluids, such as blood, urine or saliva. Such studies are most convenient in those cases when it is necessary to monitor the effectiveness of therapy in order to record the time of the onset of resistance of tumor cells, the onset of relapse and to move on to the next line of therapy. In the treatment of aggressive and rapidly became metastatic malignant tumors, such as melanoma, the presence of reliable markers that allow quick and accurate determination of treatment tactics is especially important. Nowadays, there is an increasing number of studies devoted to the search for predictive markers of the effectiveness of immunotherapy. Melanoma is one of the most immunogenic tumors and, as a result, has become a model object for research into and introduction of new approaches to immunotherapy. In this study, we compared two groups of patients with metastatic skin melanoma, with different responses to immunotherapy with blockers of immune control points, to identify new predictive expression biomarkers among microRNAs and mRNAs, and to identify the

Received 13.09.2017

Accepted for publication 08.11.2017

© AUTHORS, 2017

E.N. Lukyanova and M.S. Fedorova equal contribution to manuscript preparation.

e-mail: rhizamoeba@mail.ru

genes responsible for the occurrence of an objective response to therapy. As a result, the study detected several microRNAs with a significant change in expression level within the tumor tissue of patients responding differently to immunotherapy. Differences in the level of expression of their target genes have also been found, that will allow a more detailed analysis of the molecular mechanisms that determine the sensitivity or resistance of malignant melanoma cells to the immunotherapy. Based on the obtained data, we have proposed expression markers (mRNAs and microRNAs) that can be used as predictors of malignant melanoma tumors to immunotherapy.

Key words: melanoma; immunotherapy; predictive biomarkers; NGS; miRNA profiling; transcriptomic analysis; regulation of gene expression.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Лукьянова Е.Н., Федорова М.С., Пудова Е.А., Наседкина Т.В., Степанова Е.В., Ньюшко К.М., Попов А.Ю., Коробан Н.В., Дмитриев А.А., Киселева М.В., Липатова А.В., Заседателев А.С., Кудрявцева А.В. Особенности экспрессионного профиля больных злокачественной меланомой в зависимости от ответа на иммунотерапию. Вавиловский журнал генетики и селекции. 2017; 21(8):925-931. DOI 10.18699/VJ17.314

HOW TO CITE THIS ARTICLE:

Lukyaynova E.N., Fedorova M.S., Pudova E.A., Nasedkina T.V., Stepanova E.V., Nyushko K.M., Popov A.Y., Koroban N.V., Dmitriev A.A., Kiseleva M.V., Lipatova A.V., Zasedatelev A.S., Kudryavtseva A.V. Differences in expression profiles in malignant melanoma patients according to immunotherapy response. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2017; 21(8):925-931. DOI 10.18699/VJ17.314 (in Russian)

Несмотря на значительные достижения в области расшифровки механизмов канцерогенеза, разработки методов диагностики и новых лекарственных средств, заболеваемость злокачественными новообразованиями и смертность в России и во всем мире неуклонно растут (Петрова и др., 2014).

Стремительное течение онкологических заболеваний, а также высокая токсичность лекарственной и лучевой терапии обуславливают необходимость первоначально правильного определения тактики лечения. Поэтому одним из важнейших направлений современной молекулярной генетики и биомедицины является поиск предиктивных маркеров, помогающих выбрать наиболее эффективный способ лечения и препарат, а также индивидуально подобрать дозу (Szczerpaniak Sloane et al., 2017). Предиктивными маркерами могут быть однонуклеотидные замены и небольшие инсерции и делеции (Martin-Liberal, Larkin, 2015; Krawczyk et al., 2017), амплификации (Bouchalova et al., 2009), хромосомные перестройки (Sokolenko, Imyaninov, 2017), слитые транскрипты (Rix et al., 2013), а также экспрессия генов (на уровне мРНК и белка) и некодирующих РНК (Shindo et al., 2017). Среди маркеров особенно перспективными считаются те, которые могут обеспечить возможность применения так называемой жидкостной биопсии. Этот метод позволяет оценить состояние опухоли не по фрагменту ткани биоптата, а по анализу естественных жидкостей организма (кровь, моча, слюна). Такие исследования особенно удобны и важны, когда необходимо проводить мониторинг эффективности терапии, чтобы вовремя зафиксировать момент возникновения резистентности опухолевых клеток, начало рецидива и перейти на следующую линию терапии. Особого внимания заслуживают микроРНК, так как в крови они главным образом находятся в составе экзосом или связаны с белковыми комплексами и поэтому более стабильны (Huang, Hoop, 2016).

Необходимое условие развития новых способов лечения онкологических заболеваний – детальное понимание

молекулярных механизмов, лежащих в основе ответа злокачественных клеток на препараты. Это позволяет правильно выбирать целевую группу пациентов, контролировать лечение и корректировать тактику терапии на основе понимания механизмов возникновения резистентности (Milik et al., 2017).

Подобные данные приобретают особую значимость при терапии агрессивных и быстро метастазирующих злокачественных новообразований, таких как меланома. Удовлетворительные результаты лечения меланомы, как правило, могут быть получены только в случае диагностики заболевания на самых ранних стадиях и при своевременном хирургическом вмешательстве (Leong et al., 2012). Меланома – одна из наиболее иммуногенных опухолей, в связи с чем она стала модельным объектом для исследования и внедрения новых подходов иммунотерапии. Высокая чувствительность к онкоиммунопрепаратам во многом объясняется высокой мутационной нагрузкой при меланоме, обуславливающей продукцию большого числа неоантигенов, что, в свою очередь, усиливает иммунный ответ (Alexandrov et al., 2013).

В настоящее время наиболее эффективный способ иммунотерапии – ингибирование так называемых иммунных контрольных точек для последующей активации Т-клеточного ответа и опосредованной им смерти опухолевых клеток. К основным мишеням блокады контрольных точек относятся CTLA4 и PD-1, два ключевых рецептора иммунной толерантности.

До сих пор нет общепринятого подхода к отбору пациентов, для которых будет эффективна терапия ингибиторами иммунных контрольных точек, и к оптимальному комбинированию различных взаимодополняющих подходов к иммунотерапии друг с другом и с традиционными методами лечения рака (Hodi et al., 2010; Roh et al., 2017). Основным методом оценки эффективности иммунотерапии считается иммуногистохимическое окрашивание, например антителами к PD-1. Однако предиктивная точ-

ность остается относительно низкой, более того, существуют значительные сложности с унификацией интерпретации данных.

Появляется все больше исследований, посвященных поиску предиктивных маркеров эффективности иммунотерапии. Показано, что наличие анеуплоидии в опухоли коррелирует с маркерами иммунного «ускользания» и с пониженным ответом на иммунотерапию (Davoli et al., 2017). Повышенная мутационная нагрузка (МН) может свидетельствовать о чувствительности, в то время как накопление делеций генов – о резистентности к препаратам-блокаторам иммунных контрольных точек (Roh et al., 2017). В настоящее время высокая МН – наиболее точный генетический маркер. При меланоме этот показатель в целом выше, чем при других злокачественных новообразованиях (Alexandrov et al., 2013). Однако тестирование МН имеет существенный недостаток – высокую стоимость исследования, что, несомненно, ограничивает возможности применения этого подхода в рутинной лабораторной диагностике.

Высокая МН, как правило, связана с нарушением функционирования системы репарации. При меланоме она вызвана действием ультрафиолетового излучения и сопровождается специфической «мутационной подписью» (Alexandrov et al., 2013). На опухолях толстой кишки показано, что пациенты с мутациями в генах системы репарации неспаренных оснований, таких как *MLH1*, *MSH2*, *MSH6* и других, характеризуются повышенной МН и наиболее чувствительны к иммунотерапии. Выявлены также эпигенетические предикторы. Амплификации и активирующие соматические мутации в гене *EZH2* коррелируют с уровнем метилирования ДНК, обеспечивая эпигенетическую инактивацию ряда генов, вовлеченных в опухолевую супрессию, и иммунный ответ при меланоме (Tiffen et al., 2016).

В работе проведено сравнение двух групп больных метастатической меланомой кожи с различным ответом

на иммунотерапию блокаторами контрольных точек, для того чтобы идентифицировать новые предиктивные экспрессионные биомаркеры, а также выявить гены, ответственные за возникновение объективного ответа на терапию у пациентов. В дальнейшем это позволит разработать способы оптимального выбора терапии в рамках персонализированного подхода к лечению, а также модулировать иммунный ответ в целях улучшения результатов лечения ингибиторами контрольных точек.

Материалы и методы

В исследовании использованы результаты высокопроизводительного секвенирования на платформе Illumina, депонированные в базу данных проекта «The Cancer Genome Atlas» (TCGA-project). Общая выборка больных метастатической меланомой кожи составила 13 пациентов, получавших после хирургического лечения препараты ипилимумаб или ниволумаб. В аннотации образцов присутствовали сведения об эффективности терапии: частичный ответ (PR), полный ответ (CR) и прогрессия заболевания (PD). Выборка была разделена на две группы: 1) пациенты, имевшие частичный или полный ответ на терапию (при частичном уменьшении размеров опухоли или ее полном видимом исчезновении); 2) пациенты, у которых наблюдалась быстрая прогрессия заболевания. Эффективность ответа на терапию оценивалась в соответствии с критериями RECIST (Response Evaluation Criteria in Solid Tumors), при этом учитывалась возможность временного увеличения объема опухоли вследствие повышения содержания клеток иммунной системы. Основные характеристики пациентов, включенных в исследуемую выборку, представлены в табл. 1.

Анализ дифференциальной экспрессии микроРНК и мРНК проведен в статистической среде R с использованием пакета DESeq2. Получены списки дифференциально экспрессируемых микроРНК и мРНК, удовлетворяющих следующим условиям: 1) уровень достоверности

Table 1. Clinicopathological characterization of samples

Patient	Response to treatment	Sex	Age	Treatment	Clinical stage	Response
TCGA-DA-A3F2	PR	M	55	aCTLA4	IIIb	+
TCGA-EE-A2G5	PD	W	28	aCTLA4	Ib	-
TCGA-EE-A3JI	PD	M	48	aCTLA4	I	-
TCGA-FR-A3YN	CR	M	44	aCTLA4	Ib	+
TCGA-GF-A3OT	PD	W	58	aCTLA4	IIIc	-
TCGA-GN-A4U4	PR	M	73	aCTLA4	IIa	+
TCGA-GN-A4U9	PR	M	71	aCTLA4	IIIc	+
TCGA-GN-A8LK	PD	M	70	aCTLA4	Ib	-
TCGA-GN-A8LN	PR	M	68	aCTLA4	IIc	+
TCGA-QB-AA90	PD	M	73	aCTLA4	IIIc	-
TCGA-WE-A8K5	PD	M	65	aCTLA4	Iv	-
TCGA-WE-A8ZN	PD	M	57	aPD1	IIb	-
TCGA-WE-AAA0	PD	M	47	aCTLA4	Ia	-

(р-значение) с коррекцией средней доли ложных отклонений (FDR-corrected p-value) менее 0.05; 2) параметр порога дифференциальной экспрессии ($\log_2\text{FoldChange}$) по модулю более 1.0.

Далее для каждой микроРНК из полученного списка идентифицировали мРНК-мишени. С этой целью проведен анализ баз данных: miRDB (<http://mirdb.org/download.html>) и microRNA.org (<http://www.microRNA.org/microRNA/getDownloads.do>). Полученный список мРНК-потенциальных мишеней сопоставили со списком дифференциально экспрессируемых мРНК в сравниваемых группах больных. Были выбраны только те мРНК, которые присутствовали в обоих списках и изменение экспрессии которых имело разнонаправленный характер по отношению к соответствующей микроРНК. Использованный подход позволил выбрать для дальнейшего анализа среди всех предсказанных мРНК-мишеней наиболее достоверно связанные с молекулярными механизмами, ответственными за чувствительность или резистентность клеток меланомы к иммунотерапии. Следует отметить, что при небольшом размере выборки необходимо рассматривать только изменения, демонстрирующие высочайшие значения достоверности. Проанализированы функции выявленных генов для оценки их роли в механизмах, обеспечивающих чувствительность или устойчивость к иммунотерапии.

Результаты

Проведен биоинформатический анализ, направленный на поиск в образцах злокачественной меланомы дифференциально экспрессируемых микроРНК и их генов-мишеней, которые отличают пациентов, демонстрирующих объективный лечебный эффект иммунотерапии, от пациентов без значимого ответа. Идентифицирован ряд микроРНК, удовлетворяющих условиям отбора, и выявлены потенциальные мРНК-мишени (табл. 2).

Биоинформатический поиск генов-мишеней, регулируемых выявленными нами маркерными микроРНК, проведенный с использованием двух разных баз данных, miRDB и microRNA.org, позволил идентифицировать девять генов (табл. 3).

Перечисленные в табл. 3 гены вовлечены в различные клеточные процессы. Функции некоторых из этих генов до сих пор детально не известны.

Ген *GPR22* (G белок-сцепленный рецептор 22) расположен на длинном плече хромосомы 7 в регионе 22.3 (7q22.3), его белковый продукт принадлежит к семейству рецепторов 1 и связан с G-белком-рецептором; функции *GPR22* не известны (O'Dowd et al., 1997).

Фосфолипаза C, кодируемая геном *PLCXD3*, локализованным на коротком плече хромосомы 5 в области 13.1 (5p13.1), относится к суперсемейству фосфоинозитидспецифических фосфолипаз C, регулируемых рецептором фосфодиэстеразы. Фосфолипазы C принимают участие в регуляции уровня цитозольного кальция и активности протеинкиназы (Gellatly et al., 2012). Детально функции гена *PLCXD3* не изучены, однако кодируемый им фермент, предположительно, влияет на активность многих белков клетки (Katan, 2005).

Ген *MYT1L* (миелиновый транскрипционный фактор 1) расположен на коротком плече хромосомы 2 в области 25.3

(2p25.3), кодирует белок, содержащий мотив цинкового пальца. Этот белок функционирует в развивающейся центральной нервной системе млекопитающих. В экспериментальных условиях экспрессия *MYT1L* в сочетании с основным факторами транскрипции типа спираль-петля-спираль может преобразовывать человеческие фибробласты в индуцированные нейроны, которые способны генерировать потенциалы действия (Pang et al., 2011; Yoo et al., 2011). Данные о вовлеченности *MYT1L* в иммунные реакции или в процесс канцерогенеза в литературе отсутствуют.

Белок A4 семейства переносчиков растворенных веществ 24 кодируется геном *SLC24A4*, локализованным на длинном плече хромосомы 14 в регионе 32 (14q32). Он принадлежит к семейству переносчиков растворенных веществ. В частности, *SLC24A4* связан с транспортом глюкозы и других сахаров, органических кислот, ионов металлов и аминов. Показано, что три полиморфизма в гене *SLC24A4*, а именно: rs28777, rs35391 и rs16891982, ассоциированы с повышенным риском развития злокачественной меланомы (Duffy et al., 2010).

В области импринтированных генов на коротком плече хромосомы 11 в регионе 15.4 (11p15.4) располагается ген *PHLDA2* (*TSSC3*). Этот регион содержит кластер генов-супрессоров опухолевого роста, изменения в котором могут быть связаны со многими видами рака, такими как карцинома надпочечника, рак легких, яичников и молочной железы (Schwienbacher et al., 1998; Müller et al., 2000).

Ген *LIN28A* локализуется на коротком плече хромосомы 1 (1p36.11), кодирует субъединицу РНК-связывающего белка LIN28A/LIN28B, регулирующего, в частности, биогенез микроРНК семейства let-7, образующих регуляторную петлю. При злокачественных новообразованиях часто наблюдается повышенная экспрессия *LIN28A*, что может быть ассоциировано с дополнительной активацией онкогенов *RAS*, *MYC* и *HMGA2*, а также с изменениями в клеточном метаболизме, пролиферацией, ангиогенезом, процессами программируемой клеточной гибели, геномной нестабильностью, воспалением в зоне опухоли, метастазированием. В совокупности эти изменения приводят к иммортализации злокачественных клеток и «ускользанию» от иммунной системы. Показано также, что повышенная экспрессия *LIN28A* коррелирует с неблагоприятным прогнозом при ряде онкологических заболеваний (Wang et al., 2015).

Меланосомальный трансмембранный белок, кодируемый геном *OCA2*, располагающимся на длинном плече хромосомы 15 (15q12-q13.1), является высоко полиморфным и функционально связан с процессами пигментации. Ряд полиморфных вариантов гена *OCA2* у человека играет роль в определении нормальных фенотипических вариаций цвета глаза, а некоторые из них ассоциированы с риском возникновения меланомы (Hawkes et al., 2013). Показано, что сниженная экспрессия гена *OCA2* коррелирует с прогрессией меланомы (Ryu et al., 2007).

Идентифицированные в данной работе гены *BANK1* и *RYR2* вовлечены в процессы внутриклеточной передачи кальция. Поддержание баланса Ca^{2+} внутриклеточного свободного кальция (Ca^{2+}), в том числе работа кальциевых каналов, играет большую роль в реализации многих жизненно важных функций клетки и необходимо для проли-

Table 2. miRNAs differentially expressing in skin melanoma tissues and their target genes

miRNAs	log2FoldChange	FDR-corrected p-value	Target genes
miR-187	-7.504300342	1.26E-06	<i>RYR2</i>
miR-214	3.185228481	0.005	<i>SLC24A4, PHLDA2</i>
miR-221	-1.571716577	0.019	<i>PLCXD3, MYT1L, LIN28A</i>
miR-30E	1.219245476	0.015	<i>OCA2</i>
miR-3662	-2.588305151	0.036	<i>MYT1L, LIN28A, PLCXD3, BANK1</i>
miR-373	4.550029362	0.014	<i>SLC24A4</i>
miR-539	-3.186501375	0.009	<i>MYT1L, RYR2, GPR22, PLCXD3, BANK1</i>

ферации и миграции. Нарушение этих процессов связано с рядом заболеваний, в первую очередь онкологических.

Ген *BANK1* (В-клеточный каркасный белок с анкириновыми повторностями 1) локализован на длинном плече хромосомы 4 (4q24). Кодированный им белковый продукт функционирует в качестве важной адаптерной молекулы, связывается с В-клеточным рецептором, после чего в В-клетках образуются внутриклеточные вторичные мессенджеры, что приводит к мобилизации кальция из внутриклеточных хранилищ (Aiba et al., 2006). Установлена ассоциация гена *BANK1* с патогенезом аутоиммунных заболеваний, таких как системная красная волчанка (Kozutsev et al., 2008) и аутоиммунный тиреоидит (Muhali et al., 2013).

Кодирующий рецептор риадолина 2 ген *RYR2* локализуется на длинном плече хромосомы 1 в области 43 (1q43). Белковый продукт данного гена – один из компонентов кальциевого канала (Benkuskuy et al., 2004).

Обсуждение

Наиболее эффективным способом поиска предикторов ответа на лекарственную терапию считается сопоставление групп пациентов, различающихся по эффективности лечения. Даже в случае применения таргетных препаратов с заведомо известной молекулярной мишенью могут быть выявлены дополнительные предикторы ответа. Например, при разработке препарата кризотиниб, ингибитора *MET*, оказалось, что предиктором ответа на таргетную терапию аденокарциномы легкого является наличие транслокации гена *ALK*.

При сравнении двух групп больных злокачественной меланомой, различающихся по чувствительности к иммунотерапии, нами определены дифференциально экспрессируемые микроРНК, а дальнейший анализ позволил идентифицировать их гены-мишени, также характеризующиеся дифференциальной экспрессией: *GPR22, PLCXD3, MYT1L, BANK1, SLC24A4, PHLDA2, LIN28A, OCA2* и *RYR2*.

По литературным данным, идентифицированная нами группа генов (*GPR22, PLCXD3* и *MYT1L*) изучена крайне слабо. Для некоторых из перечисленных генов отсутствуют сведения о функциях белкового продукта или уровне экспрессии при различных патологических состояниях. Ограниченная информация может объясняться, например, крайней тканеспецифичностью экспрессии гена либо тем, что он экспрессируется на очень коротком эта-

Table 3. Change of the expression of genes regulated by miRNAs under study in the group of patients responding to immune therapy

Target gene	Log2FoldChange	FDR-corrected p-value
<i>OCA2</i>	-6.088176199	2.81E-05
<i>SLC24A4</i>	-4.443078093	3.11E-05
<i>PHLDA2</i>	-3.337098901	1.61E-07
<i>BANK1</i>	2.673397565	8.97E-06
<i>RYR2</i>	4.229018476	6.99E-05
<i>GPR22</i>	5.11669104	7.35E-06
<i>PLCXD3</i>	6.487593377	1.39E-05
<i>LIN28A</i>	8.508599155	4.77E-07
<i>MYT1L</i>	9.447990872	2.71E-09

пе онтогенеза и на низком уровне. Возможно также, что ген произошел в результате дупликации, у него есть ряд родственных генов, справляющихся с необходимыми организму и исходно возложенными на него функциями. Это позволяет гену постепенно дивергировать и приобрести дополнительные свойства, обеспечивающие связь с иммунной системой и опосредующие степень ответа клетки на иммунотерапию.

Идентифицированные в результате биоинформатического анализа гены *SLC24A4, PHLDA2, LIN28A* и *OCA2* продемонстрировали вовлеченность в процессы онкогенеза, но для многих из них нет достаточного количества литературных данных для проведения полноценного анализа. Тем не менее даже по имеющейся информации видна функциональная связь идентифицированных генов с канцерогенезом при меланоме. Например, ген *LIN28A* участвует и в онкогенезе, и в процессах иммунного ответа, в то время как определенные наследственные варианты гена *OCA2* связаны с повышенным риском возникновения меланомы. В работе, посвященной иммунотерапии онкологических заболеваний (Yang et al., 2013), проведено сравнение различных комбинаций цитокинов с целью получения генно-инженерных CD8+Т-клеток. Исследователи попытались создать менее дифференцированные противоопухолевые CD8+Т-клетки *ex vivo*, закладывая основу для дальнейшего развития онкоиммунотерапии.

Обнаружено, что под воздействием смеси цитокинов активировалась экспрессия гена *LIN28A*.

Изменения в экспрессии белков, участвующих в процессе продвижения кальция через плазматическую мембрану и клеточные органеллы, идентифицированы при различных злокачественных новообразованиях (Monteith et al., 2012). В некоторых случаях кальциевые каналы рассматриваются в качестве потенциальных терапевтических мишеней для определенных нозологических форм, а уровень их экспрессии часто коррелирует с прогнозом заболевания. Накопленные данные свидетельствуют о том, что кальциевая сигнализация играет важную роль во многих клеточных процессах, в том числе злокачественной трансформации и поддержании иммунитета (Hanahan, Weinberg, 2011). Полученные нами результаты демонстрируют значительные различия в уровнях экспрессии генов *BANK1* и *RYR2*, в зависимости от степени ответа больных меланомой на иммунотерапию. Этот факт открывает возможность их использования в качестве предиктивных маркеров, а также подтверждает роль кальциевого баланса при нормальных и патологических состояниях организма.

Установлены различия на уровне экспрессии генов, которые позволят более детально проанализировать молекулярные механизмы, определяющие чувствительность или резистентность клеток злокачественной меланомы к действию иммунотерапии. Предложены экспрессионные маркеры на основе мРНК и микроРНК, которые после дальнейшей апробации могут быть использованы как предикторы ответа опухолей злокачественной меланомы на иммунотерапию ингибиторами контрольных точек.

Среди маркеров мРНК выделено пять генов, экспрессия которых увеличивается более чем в четыре раза. Применение в качестве маркера повышения уровня экспрессии (а не понижения) традиционно считается более удобным в лабораторной практике. Это позволяет использовать для анализа меньшее количество исходного материала и точнее интерпретировать полученные результаты тестов.

Выявление нескольких микроРНК со значительным изменением содержания в опухолевой ткани пациентов, по-разному отвечающих на иммунотерапию, позволяет надеяться в будущем на возможность перехода на неинвазивную диагностику.

Acknowledgments

This work was supported by the Russian Science Foundation, project 14-35-00107.

The authors are grateful to the staff of the Genome Shared Access Center, Institute of Molecular Biology, Russian Academy of Sciences (http://www.eimb.ru/rus/ckp/ccu_genome_c.php) and Initium Pharm Company for granting access to computational facilities.

Conflict of interest

The authors declare no conflict of interest.

References

Aiba Y., Yamazaki T., Okada T., Gotoh K., Sanjo H., Ogata M., Kurosaki T. BANK negatively regulates Akt activation and subsequent B cell responses. *Immunity*. 2006;24(3):259-268. DOI 10.1016/j.immuni.2006.01.002.

Alexandrov L.B., Nik-Zainal S., Wedge D.C., Aparicio S.A., Behjati S., Biankin A.V., Bignell G.R., Bolli N., Borg A., Børresen-Dale A.L., Boyault S., Burkhardt B., Butler A.P., Caldas C., Davies H.R., Desmedt C., Eils R., Eyfjörd J.E., Foekens J.A., Greaves M., Hosoda F., Hutter B., Ilcic T., Imbeaud S., Imielinski M., Jäger N., Jones D.T., Jones D., Knappskog S., Kool M., Lakhani S.R., López-Otin C., Martin S., Munshi N.C., Nakamura H., Northcott P.A., Pajic M., Paemmanuil E., Paradiso A., Pearson J.V., Puente X.S., Raine K., Ramakrishna M., Richardson A.L., Richter J., Rosenstiel P., Schlessner M., Schumacher T.N., Span P.N., Teague J.W., Totoki Y., Tutt A.N., Valdés-Mas R., van Buuren M.M., van 't Veer L., Vincent-Salomon A., Waddell N., Yates L.R. Australian Pancreatic Cancer Genome Initiative; ICGC Breast Cancer Consortium; ICGC MMML-Seq Consortium; ICGC PedBrain; Zucman-Rossi J., Futreal P.A., McDermott U., Lichten P., Meyerson M., Grimmond S.M., Siebert R., Campo E., Shibata T., Pfister S.M., Campbell P.J., Stratton M.R. Signatures of mutational processes in human cancer. *Nature*. 2013;22;500(7463):415-421. DOI 10.1038/nature12477.

Benkusky N.A., Farrell E.F., Valdivia H.H. Ryanodine receptor channelopathies. *Biochem. Biophys. Res. Commun.* 2004;322(4):1280-1285. DOI 10.1016/j.bbrc.2004.08.033.

Bouchalova K., Cizkova M., Cwiertka K., Trojanec R., Hajduch M. Triple negative breast cancer – current status and prospective targeted treatment based on HER1 (EGFR), TOP2A and C-MYC gene assessment. *Biomed. Pap. Med. Fac. Univ. Palacky Olomouc Czech. Repub.* 2009;153(1):13-17.

Davoli T., Uno H., Wooten E.C., Elledge S.J. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science*. 2017;355(6322). DOI 10.1126/science.aaf8399.

Duffy D.L., Zhao Z.Z., Sturm R.A., Hayward N.K., Martin N.G., Montgomery G.W. Multiple pigmentation gene polymorphisms account for a substantial proportion of risk of cutaneous malignant melanoma. *J. Invest. Dermatol.* 2010;130(2):520-528. Epub 2009; Aug 27. DOI 10.1038/jid.2009.258.

Gellatly S.A., Kalujnaia S., Cramb G. Cloning, tissue distribution and sub-cellular localisation of phospholipase C X-domain containing protein (PLCXD) isoforms. *Biochem. Biophys. Res. Commun.* 2012;424(4):651-656. Epub 2012; Jun 22. DOI 10.1016/j.bbrc.2012.06.079.

Hanahan D., Weinberg R.A. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-674. DOI 10.1016/j.cell.2011.02.013.

Hawkes J.E., Cassidy P.B., Manga P., Boissy R.E., Goldgar D., Cannon-Albright L., Florell S.R., Leachman S.A. Report of a novel OCA2 gene mutation and an investigation of OCA2 variants on melanoma risk in a familial melanoma pedigree. *J. Dermatol. Sci.* 2013;69(1):30-7. Epub 2012; Oct 13. DOI 10.1016/j.jdermsci.2012.09.016.

Hodi F.S., O'Day S.J., McDermott D.F., Weber R.W., Sosman J.A., Haanen J.B., Gonzalez R., Robert C., Schadendorf D., Hassel J.C., Akerley W., van den Eertwegh A.J., Lutzky J., Lorigan P., Vaubel J.M., Linette G.P., Hogg D., Ottensmeier C.H., Lebbé C., Peschel C., Quirt I., Clark J.I., Wolchok J.D., Weber J.S., Tian J., Yellin M.J., Nichol G.M., Hoos A., Urba W.J. Improved survival with ipilimumab in patients with metastatic melanoma. *N. Engl. J. Med.* 2010;363(8):711-723. Epub 2010; Jun 5. DOI 10.1056/NEJMoa1003466.

Huang S.K., Hoon D.S. Liquid biopsy utility for the surveillance of cutaneous malignant melanoma patients. *Mol. Oncol.* 2016;10(3):450-463. Epub 2015; Dec 17. DOI 10.1016/j.molonc.2015.12.008.

Katan M. New insights into the families of PLC enzymes: looking back and going forward. *Biochem. J.* 2005;391(3):e7-9. DOI 10.1042/BJ20051506.

Kozyrev S.V., Abelson A.K., Wojcik J., Zaghlool A., Linga Reddy M.V., Sanchez E., Gunnarsson I., Svenungsson E., Sturfelt G., Jönsen A., Truedsson L., Pons-Estel B.A., Witte T., D'Alfonso S., Barizzone N., Danieli M.G., Gutierrez C., Suarez A., Junker P., Lastrup H., González-Escribano M.F., Martin J., Abderrahim H., Alarcón-Riquelme M.E. Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. *Nat. Genet.* 2008;40(2):211-216. DOI 10.1038/ng.79.

- Krawczyk P., Kowalski D.M., Ramlau R., Kalinka-Warzocha E., Winiarczyk K., Stencel K., Powrózek T., Reszka K., Wojas-Krawczyk K., Bryl M., Wójcik-Superczyńska M., Głogowski M., Barinow-Wojewódzki A., Milanowski J., Krzakowski M. Comparison of the effectiveness of erlotinib, gefitinib, and afatinib for treatment of non-small cell lung cancer in patients with common and rare EGFR gene mutations. *Oncol. Lett.* 2017;(6):4433-4444. Epub 2017; Apr 3. DOI 10.3892/ol.2017.5980.
- Leong S.P., Mihm M.C. Jr., Murphy G.F., Hoon D.S., Kashani-Sabet M., Agarwala S.S., Zager J.S., Hauschild A., Sondak V.K., Guild V., Kirkwood J.M. Progression of cutaneous melanoma: implications for treatment. *Clin. Exp. Metastasis.* 2012;29(7):775-796. DOI 10.1007/s10585-012-9521-1.
- Martin-Liberal J., Larkin J. Vemurafenib for the treatment of BRAF mutant metastatic melanoma. *Future Oncol.* 2015;11(4):579-589. DOI 10.2217/fon.14.252.
- Milik S.N., Lasheen D.S., Serya R.A.T., Abouzid K.A.M. How to train your inhibitor: Design strategies to overcome resistance to Epidermal Growth Factor Receptor inhibitors. *Eur. J. Med. Chem.* 2017;142:131-151. DOI 10.1016/j.ejmech.2017.07.023.
- Monteith G.R., Davis F.M., Roberts-Thomson S.J. Calcium channels and pumps in cancer: changes and consequences. *J. Biol. Chem.* 2012;287(38):31666-31673. Epub 2012; Jul 20. DOI 10.1074/jbc.R112.343061.
- Muhali F.S., Song R.H., Wang X., Shi X.H., Jiang W.J., Xiao L., Li D.F., He S.T., Xu J., Zhang J.A. Genetic variants of BANK1 gene in autoimmune thyroid diseases: a case-control association study. *Exp. Clin. Endocrinol. Diabetes.* 2013;121(9):556-560. Epub 2013; Oct 14. DOI 10.1055/s-0033-1348220.
- Müller S., van den Boom D., Zirkel D., Köster H., Berthold F., Schwab M., Westphal M., Zunkeller W. Retention of imprinting of the human apoptosis-related gene TSSC3 in human brain tumors. *Hum. Mol. Genet.* 2000;22(9):757-763. PMID: 10749982.
- O'Dowd B.F., Nguyen T., Jung B.P., Marchese A., Cheng R., Heng H.H., Kolakowski L.F. Jr., Lynch K.R., George S.R. Cloning and chromosomal mapping of four putative novel human G-protein-coupled receptor genes. *Gene.* 1997;187(1):75-81.
- Pang Z.P., Yang N., Vierbuchen T., Ostermeier A., Fuentes D.R., Yang T.Q., Citri A., Sebastiano V., Marro S., Südhof T.C., Wernig M. Induction of human neuronal cells by defined transcription factors. *Nature.* 2011;26;476(7359):220-223. DOI 10.1038/nature10202.
- Petrova G.V., Kaprin A.D., Starinskij V.V., Gretsova O.P. The incidence of malignant neoplasms in the Russian population. *Onkologiya. Zhurnal imeni P.A. Gertsena = P.A. Herzen Journal of Oncology.* 2014;2(5):5-10. (in Russian)
- Rix U., Colinge J., Blatt K., Gridling M., Remsing Rix L.L., Parapatics K., Cerny-Reiterer S., Burkard T.R., Jäger U., Melo J.V., Bennett K.L., Valent P., Superti-Furga G. A target-disease network model of second-generation BCR-ABL inhibitor action in Ph+ ALL. *PLoS One.* 2013;8(10):e77155. DOI 10.1371/journal.pone.0077155.
- Roh W., Chen P.L., Reuben A., Spencer C.N., Prieto P.A., Miller J.P., Gopalakrishnan V., Wang F., Cooper Z.A., Reddy S.M., Gumbs C., Little L., Chang Q., Chen W.S., Wani K., De Mace-
- do M.P., Chen E., Austin-Breneman J.L., Jiang H., Roszik J., Tetzlaff M.T., Davies M.A., Gershenwald J.E., Tawbi H., Lazar A.J., Hwu P., Hwu W.J., Diab A., Glitza I.C., Patel S.P., Woodman S.E., Amaria R.N., Prieto V.G., Hu J., Sharma P., Allison J.P., Chin L., Zhang J., Wargo J.A., Futreal P.A. Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Sci. Transl. Med.* 2017;9(379). DOI 10.1126/scitranslmed.aah3560. Erratum: *Sci. Transl. Med.* 2017. Apr 12;9(385).
- Ryu B., Kim D.S., Deluca A.M., Alani R.M. Comprehensive expression profiling of tumor cell lines identifies molecular signatures of melanoma progression. *PLoS One.* 2007;2(7):e594. DOI 10.1371/journal.pone.0000594.
- Schwienbacher C., Sabbioni S., Campi M., Veronese A., Bernardi G., Menegatti A., Hatada I., Mukai T., Ohashi H., Barbanti-Brodano G., Croce C.M., Negrini M. Transcriptional map of 170-kb region at chromosome 11p15.5: identification and mutational analysis of the BWR1A gene reveals the presence of mutations in tumor samples. *Proc. Natl. Acad. Sci USA.* 1998;95(7):3873-3878.
- Shindo Y., Hazama S., Nakamura Y., Inoue Y., Kanekiyo S., Suzuki N., Takenouchi H., Tsunedomi R., Nakajima M., Ueno T., Takeda S., Yoshino S., Okuno K., Fujita Y., Hamamoto Y., Kawakami Y., Oka M., Nagano H. miR-196b, miR-378a and miR-486 are predictive biomarkers for the efficacy of vaccine treatment in colorectal cancer. *Oncol. Lett.* 2017;14(2):1355-1362. Epub 2017; Jun 2. DOI 10.3892/ol.2017.6303.
- Sokolenko A.P., Imyanitov E.N. Molecular tests for the choice of cancer therapy. *Curr. Pharm. Des.* 2017; DOI 10.2174/1381612823666170719110125.
- Szczepaniak Sloane R.A., Gopalakrishnan V., Reddy S.M., Zhang X., Reuben A., Wargo J.A. Interaction of molecular alterations with immune response in melanoma. *Cancer.* 2017;123(S11):2130-2142. DOI 10.1002/cncr.30681. 28543700.
- Tiffen J., Wilson S., Gallagher S.J., Hersey P., Filipp F.V. Somatic copy number amplification and hyperactivating somatic mutations of EZH2 correlate with DNA methylation and drive epigenetic silencing of genes involved in tumor suppression and immune responses in melanoma. *Neoplasia.* 2016;18(2):121-132. DOI 10.1016/j.neo.2016.01.003.
- Wang T., Wang G., Hao D., Liu X., Wang D., Ning N., Li X. Aberrant regulation of the LIN28A/LIN28B and let-7 loop in human malignant tumors and its effects on the hallmarks of cancer. *Mol. Cancer.* 2015;14:125. DOI 10.1186/s12943-015-0402-5.
- Yang S., Ji Y., Gattinoni L., Zhang L., Yu Z., Restifo N.P., Rosenberg S.A., Morgan R.A. Modulating the differentiation status of ex vivo-cultured anti-tumor T cells using cytokine cocktails. *Cancer Immunol. Immunother.* 2013;62(4):727-736. Epub 2012; Dec 4. DOI 10.1007/s00262-012-1378-2.
- Yoo A.S., Sun A.X., Li L., Shcheglovitov A., Portmann T., Li Y., Lee-Messer C., Dolmetsch R.E., Tsien R.W., Crabtree G.R. MicroRNA-mediated conversion of human fibroblasts to neurons. *Nature.* 2011; 476(7359):228-231. DOI 10.1038/nature10323.

Ингибирование гексокиназы 2 приводит к снижению экспрессии ферментов гликолиза ПФКР, ВРГМ и GPI в клеточной линии RKO

М.С. Федорова¹, И.Ю. Карпова¹, А.В. Липатова¹, Е.А. Пудова¹, З.Г. Гуватова¹, Д.В. Кочетков¹, А.В. Чайка², Б.Я. Алексеев², М.В. Киселева², А.Д. Каприн², А.В. Кудрявцева^{1,2}, А.В. Снежкина¹

¹ Институт молекулярной биологии им. В.А. Энгельгардта Российской академии наук, Москва, Россия

² Национальный медицинский исследовательский центр радиологии Министерства здравоохранения Российской Федерации, Москва, Россия

Колоректальный рак (КРР) – одно из наиболее распространенных злокачественных новообразований в мире, характеризующееся высоким уровнем смертности. Изучение ключевых аспектов формирования и прогрессии КРР необходимо для разработки новых подходов к его терапии, а также поиска новых диагностических, прогностических и предиктивных биомаркеров. Для многих видов опухолей важным изменением метаболизма является активация гликолиза, ассоциированная с нарушением экспрессии основных ферментов, принимающих участие в этом процессе, и регуляторных молекул. Чаще всего в опухолевых клетках наблюдается повышение экспрессии гексокиназы 2 (HK2), что делает ее многообещающей мишенью для таргетной терапии. В модифицированной нами клеточной линии RKO, постоянно экспрессирующей короткие шпилечные РНК для ингибирования гексокиназы 2, проведен количественный анализ экспрессии 15 генов (*GAPDH*, *ADPGK*, *ALDOA*, *ENO3*, *PFKL*, *PGK1*, *PGAM1*, *PKM2*, *ENO1*, *PDK1*, *PDK3*, *PFKP*, *ENO2*, *GPI* и *BPGM*), кодирующих ключевые ферменты гликолиза, а также гена *HIF1A*. Выявлено значительное снижение экспрессии генов *PFKP*, *BPGM* и *GPI* на уровне мРНК (в 5, 86 и 93 раза соответственно) и белка (в 2.5, 3.5 и 19 раз соответственно). Вероятно, снижение экспрессии *GPI* и *PFKP* связано с уменьшением количества их субстратов, глюкозо-6-фосфата и фруктозо-6-фосфата, при подавлении гексокиназы 2. Однако вопрос о причинах снижения уровня мРНК этих трех ферментов при одновременном сохранении уровня экспрессии других участников гликолиза требует дальнейшего изучения.

Ключевые слова: колоректальный рак; shRNA; кПЦР; Вестерн-блот; гликолиз; эффект Варбурга; HK2.

Knockdown of hexokinase 2 results in a decreased expression level of the glycolytic enzymes PFKP, BPGM, and GPI in RKO cell line

M.S. Fedorova¹, I.Y. Karpova¹, A.V. Lipatova¹, E.A. Pudova¹, Z.G. Guvatova¹, D.V. Kochetkov¹, A.V. Chaika², B.Y. Alekseev², M.V. Kiseleva², A.D. Kaprin², A.V. Kudryavtseva^{1,2}, A.V. Snezhkina¹

¹ Engelhardt Institute of Molecular Biology RAS, Moscow, Russia

² National Medical Research Radiological Center, Ministry of Health of the Russian Federation, Moscow, Russia

Colorectal cancer (CRC) is one of the most common malignant neoplasms in the world, and is characterized by a high mortality rate. The study of the key aspects of colorectal cancer formation and progression is necessary to develop new approaches to its therapy, as well as to search for new diagnostic, prognostic and predictive biomarkers of CRC. In many types of tumors, one of the key changes in metabolism is the activation of glycolysis, which is associated with alterations in the expression of the main glycolytic enzymes and regulatory molecules. There is often an increase in hexokinase 2 (HK2) exogenous expression in tumor cells, which makes it a promising target for anticancer therapy. Quantitative expression analysis of 15 genes (*GAPDH*, *ADPGK*, *ALDOA*, *ENO3*, *PFKL*, *PGK1*, *PGAM1*, *PKM2*, *ENO1*, *PDK1*, *PDK3*, *PFKP*, *ENO2*, *GPI*, and *BPGM*), encoding the key glycolysis enzymes, as well as *HIF1A* gene was carried out in a modified RKO cell line, which constantly expresses the short hairpin RNA (shRNA) for the inhibition of hexokinase 2. A significant decrease in the expression of *PFKP*, *BPGM*, and *GPI* genes both at the mRNA (5-, 86-, and 93-fold, respectively) and protein (2.5-, 3.5-, and 19-fold, respectively) levels was revealed. Probably, the downregulation of *GPI* and *PFKP* is associated with a decrease in the amount of their substrates, glucose-6-phosphate and fructose-6-phosphate, under the inhibition of hexokinase 2. Nevertheless, the cause of a decreased mRNA level of these three enzymes, while the expression level of other glycolytic participants is constant, requires further investigation.

Key words: colorectal cancer; shRNA; qPCR; Western blot; glycolysis; Warburg effect; HK2.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Федорова М.С., Карпова И.Ю., Липатова А.В., Пудова Е.А., Гуватова З.Г., Кочетков Д.В., Чайка А.В., Алексеев Б.Я., Киселева М.В., Каприн А.Д., Кудрявцева А.В., Снежкина А.В. Ингибирование гексокиназы 2 приводит к снижению экспрессии ферментов гликолиза ПФКР, ВРГМ и GPI в клеточной линии RKO. Вавиловский журнал генетики и селекции. 2017;21(8):932-936. DOI 10.18699/VJ17.315

HOW TO CITE THIS ARTICLE:

Fedorova M.S., Karpova I.Y., Lipatova A.V., Pudova E.A., Guvatova Z.G., Kochetkov D.V., Chaika A.V., Alekseev B.Y., Kiseleva M.V., Kaprin A.D., Kudryavtseva A.V., Snezhkina A.V. Knockdown of hexokinase 2 results in a decreased expression level of the glycolytic enzymes PFKP, BPGM, and GPI in RKO cell line. Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding. 2017;21(8):932-936. DOI 10.18699/VJ17.315 (in Russian)

Received 10.09.2017

Accepted for publication 27.11.2017

© AUTHORS, 2017

M.S. Fedorova and I.Y. Karpova equal contribution to manuscript preparation.

e-mail: rhizamoeba@mail.ru

Колоректальный рак (КРР) – наиболее распространенное новообразование и вторая ведущая причина смертности от рака во всем мире (Siegel et al., 2017). Это заболевание может долгое время протекать бессимптомно – на момент постановки диагноза около 25 % больных имеют отдаленные метастазы. При этом пятилетняя выживаемость пациентов с метастатическим КРР составляет менее 10 %. Необходимо развитие новых методов терапии данного заболевания, а также поиск новых биомаркеров для ранней диагностики, прогнозирования течения заболевания и персонализированного подбора лекарственных средств.

Для КРР характерны множественные генетические и эпигенетические нарушения, возникающие в ходе образования и развития опухоли (Haraldson et al., 2012; Kudryavtseva et al., 2016c). Одним из ключевых событий при развитии КРР является перестройка энергетического обмена клеток (Snezhkina et al., 2016; Graziano et al., 2017). Активная пролиферация и быстрый рост опухолевых клеток, находящихся в условиях гипоксии, требуют большого запаса энергии и строительных ресурсов, которые образуются на различных этапах гликолиза. При этом усиливается транспорт глюкозы в клетки и увеличивается скорость образования лактата. Однако даже при последующей васкуляризации опухоли и появлении кислорода в достаточном количестве клетки продолжают использовать гликолиз в качестве основного источника энергии (эффект Варбурга) (Warburg, 1956; Krasnov et al., 2013b; Kudryavtseva et al., 2016b).

Активация гликолиза происходит в злокачественных клетках настолько часто, что многие исследователи рассматривают ферменты, принимающие участие в этом процессе, как многообещающие мишени для таргетной терапии. Однако при выборе такой мишени и предварительной оценке возможности ее использования необходимо выявить все изменения в клетке, которые могут быть вызваны воздействием на нее. Прежде всего, следует определить нарушения экспрессии генов при экспериментальном подавлении на моделях *in vitro*.

В данной работе проведен количественный анализ экспрессии 15 генов (*GAPDH*, *ADPGK*, *ALDOA*, *ENO3*, *PFKL*, *PGK1*, *PGAM1*, *PKM2*, *ENO1*, *PDK1*, *PDK3*, *PFKP*, *ENO2*, *GPI* и *BPGM*), продукты которых являются ключевыми ферментами гликолиза, и гена *HIF1A*, кодирующего индуцируемый гипоксией фактор 1A, в условиях ингибирования экспрессии гена *HK2* в клеточной линии карциномы толстой кишки RKO. Для генов с изменением экспрессии на уровне мРНК проведена также оценка уровня экспрессии белка методом Вестерн-блот.

Материалы и методы

Клеточные линии и условия культивирования. Использована клеточная линия карциномы толстой кишки RKO, полученная из лаборатории пролиферации клеток Института молекулярной биологии РАН под руководством П.М. Чумакова. Клеточную линию RKO культивировали в ростовой среде ДМЕМ (ПанЭко, Россия) с добавлением 10 % фетальной бычьей сыворотки (FBS) (HyClone, США), 100 ед/мл пенициллина (ПанЭко) и 100 мкг/мл стрептомицина (ПанЭко) при влажности 95 %, темпера-

туре 37 °С, а также в условиях содержания углекислого газа на уровне 5 %.

Создание лентивирусных конструкций и получение стабильной линии клеток. Для получения стабильной клеточной линии RKO с подавленной экспрессией гена *HK2* созданы конструкции лентивирусного вектора pLSLP, экспрессирующие короткие шпилечные РНК (shRNA) (Kudryavtseva et al., 2016a). Трансдукцию проводили в четырех повторях. После лентивирусной трансдукции фракции клеток с интегрированной конструкцией выделяли путем селекции на среде с пуромицином (2 мкг/мл) (Thermo Fisher Scientific) в течение пяти дней. Степень подавления *HK2* определяли с помощью методов количественной ПЦР (кПЦР) и Вестерн-блота.

Выделение РНК и белка. Выделение тотальной РНК и белка из клеточной линии RKO проводили с использованием набора AllPrep DNA/RNA/Protein Mini Kit (Qiagen, Германия) согласно инструкции производителя. Количество выделенных РНК и белка оценивали с помощью флуориметра Qubit 2.0 (Invitrogen, США). Очищенные препараты РНК хранили при температуре –80 °С, препараты белка – при температуре –20 °С.

Получение кДНК. Выделенную РНК (1 мкг) обрабатывали ДНКазой I (Thermo Fisher Scientific) согласно протоколу производителя. Реакцию обратной транскрипции проводили с использованием ревертазы M-MuLV (Thermo Fisher Scientific) и набора случайных гексануклеотидов.

Метод количественной ПЦР (кПЦР). Для количественной оценки уровня мРНК целевых генов методом кПЦР использованы коммерческие наборы праймеров и проб TaqMan Gene Expression Assay (*HK2*: Hs00606086_m1; *HIF1A*: Hs00153153_m1; *ADPGK*: Hs00229849_m1; *PFKP*: Hs00242993_m1; *GAPDH*: Hs02758991_g1; *PDK1*: Hs01561850_m1; *PDK3*: Hs00178440_m1; *ENO1*: Hs00361415_m1; *GPI*: Hs00976711_m1; *ENO3*: Hs01093275_m1; *PKM2*: Hs00762869_s1; *PFKL*: Hs00160027_m1; *PGAM1*: Hs01652468_g1; *BPGM*: Hs00156139_m1; *PGK1*: Hs00943178_g1; *ENO2*: Hs00157360_m1; *ALDOA*: Hs00605108_g1) (Thermo Fisher Scientific). В качестве эндогенного контроля выбраны гены *RPN1* и *GUSB*. Ранее нами показана возможность их использования для транскриптомных исследований колоректального рака (Krasnov et al., 2015). Каждую реакцию ПЦР проводили в трех повторях на приборе Applied Biosystems 7500 Real-Time PCR System (Thermo Fisher Scientific) по протоколу, описанному в работе (Snezhkina et al., 2016).

Данные кПЦР обрабатывали при помощи оригинального программного приложения АТГ (Melnikova et al., 2016). Относительный уровень мРНК оценивали методом относительных измерений ($\Delta\Delta C_t$ -метод) с учетом эффективности реакций, как описано ранее (Dmitriev et al., 2016).

Разделение белков в полиакриламидном геле (ПААГ) и Вестерн-блот. Электрофорез белков осуществляли в ПААГ в денатурирующих условиях по методу Лэммли. Для отмывки геля от лаурилсульфата натрия (SDS) использовали трис-глицериновый буфер, содержащий 10 % метанола. Затем гель перенесли на мембрану из поливинилиденфторида (PVDF) (Millipore, США). Мембрану промывали буферным раствором (TBS), содержащим 0.1 % Твин 20, и блокировали в 3 % бычьего

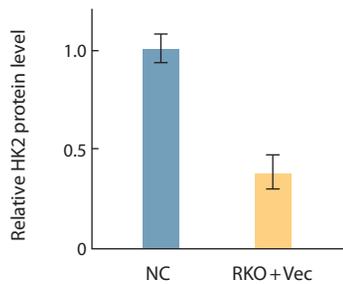


Fig. 1. Expression of *HK2* in RKO cells after the application of the lentiviral vector carrying shRNA.

NC, negative control, Vec, lentiviral vector pLSLP carrying shRNA; $p < 0.05$.

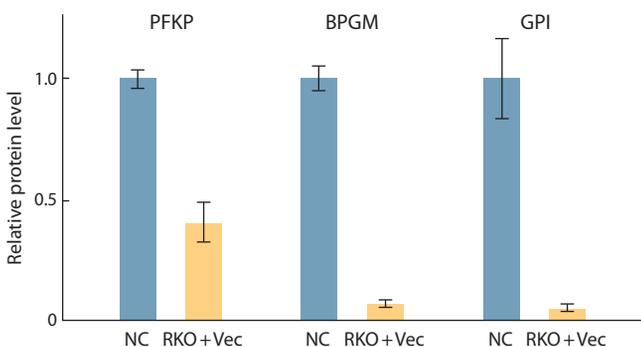


Fig. 2. Expression of PFKP, GPI, and BPGM in RKO cells with repressed *HK2* expression.

Designations follow fig. 1.

сывороточного альбумина (BSA) в TBS, содержащем 0.1 % Твин 20 в течение 1 часа при комнатной температуре. Далее мембрану инкубировали с первичными антителами в течение 10–16 ч при 4 °С, затем ее три раза отмывали в TBS, содержащем 0.1 % Твин-20. После отмывки мембрану инкубировали в растворе вторичных антител в течение 1 часа при комнатной температуре и затем отмывали пять раз, как описано ранее. Нормализацию проводили по β-актину.

В работе использовали первичные антитела, специфичные к *HK2* (MA5-14849), β-actin (PA1-183), PFKP (PA5-28673), *ENO2* (PA5-27452), GPI (PA5-29665), BPGM (PA5-21821), и вторичные антитела, конъюгированные с флуорохромом Alexa Fluor 488 (A-11008) (Thermo Fisher Scientific). Детекцию флуоресценции проводили с использованием Турбоон FLA 9500 (General Electric, США).

Статистический анализ. В качестве критерия достоверности использован *t*-критерий Стьюдента. Различия считали статистически значимыми при $p < 0.05$.

Результаты

Подавление экспрессии гена *HK2* в клеточной линии RKO. С помощью shRNA, встроенной в лентивирусный вектор, получена стабильная линия клеток RKO со сниженной экспрессией *HK2*. Методом кПЦР показано снижение экспрессии мРНК гена *HK2* в линии RKO в четыре раза. Уровень экспрессии белкового продукта –

гексокиназы 2, определенный методом Вестерн-блот, понизился в три раза (рис. 1).

Снижение экспрессии генов *PFKP*, *GPI* и *BPGM* в клеточной линии RKO в условиях подавления экспрессии *HK2*. С помощью метода кПЦР проведен анализ экспрессии 15 генов, кодирующих ключевые ферменты гликолиза (*GAPDH*, *ADPGK*, *ALDOA*, *ENO3*, *PFKL*, *PGK1*, *PGAM1*, *PKM2*, *ENO1*, *PDK1*, *PDK3*, *PFKP*, *ENO2*, *GPI* и *BPGM*), а также гена *HIF1A* в клеточной линии RKO со сниженной экспрессией *HK2*. Выявлено, что уровень мРНК большинства исследуемых генов (*GAPDH*, *ADPGK*, *ALDOA*, *ENO3*, *PFKL*, *PGK1*, *PGAM1*, *PKM2*, *ENO1*, *PDK1*, *PDK3* и *HIF1A*) не изменялся. Уровень мРНК гена *PFKP* снижался в пять раз ($p < 0.05$), гена *ENO2* – в три раза ($p < 0.05$) в клеточной линии RKO со сниженной экспрессией *HK2*. Еще для двух генов, *BPGM* и *GPI*, было характерно значительное подавление экспрессии в 86 и 93 раза ($p < 0.05$) соответственно. Методом Вестерн-блот подтверждено снижение экспрессии белков PFKP, BPGM и GPI от 2.5 до 19 раз (рис. 2). Экспрессия белка *ENO2* не изменялась.

Обсуждение

Первая ферментативная реакция гликолиза – АТФ-зависимое фосфорилирование глюкозы до глюкозо-6-фосфата – катализируется ферментами гексокиназами (Oragina et al., 2013). В тканях животных описано четыре изофермента гексокиназ, кодируемых соответствующими генами – *HK1*, *HK2*, *HK3* и *HK4* (*GCK*) (Wilson, 2003). Гексокиназа 2 играет важную роль на начальных этапах развития опухоли и характеризуется повышенной экспрессией во многих видах рака, что связано главным образом с увеличением транспорта глюкозы в клетку (Mathupala et al., 2001). Это способствует повышению скорости гликолиза, адаптирует энергетический метаболизм для поддержания бесконтрольного роста и деления опухолевых клеток, которые даже в присутствии кислорода продолжают использовать гликолиз в качестве основного источника аденозинтрифосфата (АТФ) (эффект Варбурга). Происходит быстрое преобразование глюкозы в лактат, что сопровождается закислением внеклеточной среды и способствует метастазированию (Hsu, Sabatini, 2008; Wang et al., 2017). При онко-ассоциированной активации *HK2* связывается с потенциал-зависимым анионным каналом (VDAC) на наружной мембране митохондрий, препятствуя пермеабилзации митохондриальной мембраны и, как следствие, гибели опухолевых клеток (Pastorino et al., 2002; Krasnov et al., 2013a). Связывание между комплексом VDAC и *HK2* позволяет также избежать ингибирования гексокиназы 2 глюкозо-6-фосфатом и обеспечивает прямой доступ к АТФ, генерируемому суперкомплексом АТФ-синтасомой в митохондриях, поддерживая, таким образом, высокую скорость метаболизма в опухолевых клетках (Kim, Dang, 2005).

В ряде исследований показана вовлеченность *HK2* в развитие КРР, при котором наблюдается как повышение экспрессии гена *HK2* (Izuishi et al., 2012), так и ее снижение (Krasnov et al., 2015). С помощью иммуногистохимического анализа выявлена связь между экспрессией *HK2* и размером опухоли, стадией заболевания, степенью

инвазии, а также наличием отдаленных метастазов. Подавление HK2 в клеточных линиях KPP (HCT8 и HT29) приводит как к снижению образования лактата, так и к пониженной пролиферативной и миграционной активности опухолевых клеток (Katagiri et al., 2017). Одновременное подавление HK1 и HK2, как показано нами ранее, также ассоциировано со снижением жизнеспособности клеток линий KPP HT-29, SW480, HCT-15, RKO и HCT116 (Kudryavtseva et al., 2016a).

На модели клеточной линии RKO нами проведен анализ экспрессии генов, кодирующих ключевые ферменты гликолиза, а также гена *HIF1A* в условиях ингибирования экспрессии *HK2*. Уровень экспрессии большинства генов не изменялся, для гена *ENO2* обнаружено снижение экспрессии только на уровне мРНК, для генов *PFKP*, *BPGM* и *GPI* – на уровне мРНК и белка.

Ген *GPI* кодирует фермент глюкозо-6-фосфат изомеразу, которая катализирует второй этап гликолиза. Недавние исследования показали участие белка GPI в регуляции роста опухолевых клеток и процесса метастазирования. Обнаружено (Funasaka et al., 2005), что в условиях гипоксии наблюдается повышение экспрессии гена *GPI* в клеточных линиях рака молочной железы (BT-549), а ингибирование его экспрессии приводит к снижению подвижности опухолевых клеток. При стимуляции гликолиза D-(+)-глюкозой в клеточных линиях KPP (SW480 и SW620) отмечено повышение экспрессии гена *GPI* и, наоборот, при обработке клеток ингибитором гликолиза 2-дезоксид-D-глюкозой (2-DG) экспрессия *GPI* снижалась (Yeh et al., 2008). Нами выявлено значительное подавление экспрессии гена *GPI* в условиях ингибирования *HK2*. По-видимому, снижение экспрессии *GPI* связано с уменьшением производства его субстрата – глюкозо-6-фосфата, который является продуктом первой реакции гликолиза, катализируемой гексокиназой 2.

При ингибировании *HK2* обнаружено также снижение экспрессии гена *PFKP*, кодирующего фосфофруктокиназу. Этот фермент катализирует фосфорилирование фруктозо-6-фосфата в процессе гликолиза. Следует отметить, что экспрессия гена *PFKP* снижалась в меньшей степени, чем экспрессия *GPI*. Этот факт может объясняться тем, что в условиях ингибирования экспрессии *HK2* и, соответственно, снижения содержания продукта катализируемой ей реакции, фруктозо-6-фосфата, происходит постепенное уменьшение количества субстрата для последующих реакций гликолиза. Известно, что PFKP часто вовлечен в канцерогенез и играет важную роль в нарушении энергетического обмена опухолевых клеток (Wang et al., 2016). Показано, что в клеточных линиях светлоклеточной почечно-клеточной карциномы (786-O, 769-P и Saki-1) подавление экспрессии фосфофруктокиназы приводит к ингибированию роста клеток, а также вызывает остановку клеточного цикла и индуцирует апоптоз (Wang et al., 2016). Согласно полученным нами данным, ингибирование экспрессии *HK2* приводит к снижению уровня мРНК и белка PFKP, что, в свою очередь, может индуцировать гибель опухолевых клеток, особенно в условиях *in vivo*. Таким образом, ингибирование гена *PFKP* в дополнение к *HK2* может быть рассмотрено как вариант усиления противоопухолевого эффекта.

В норме ген *BPGM* экспрессируется в различных тканях организма, преобладая в костном мозге, плаценте, сердце и эндотелиальных клетках, и кодирует фермент бисфосфоглицеромутазу, катализирующую образование 2,3-бисфосфоглицерата. Эта реакция называется шунтом Рапопорта–Люберинга (Cho et al., 2008). Известно, что в метаболических путях при проксимальных нарушениях (в случае шунта Рапопорта–Люберинга это, например, недостаточность гексокиназы, фосфоглюкоизомеразы или альдолазы А) уровни фермента, катализирующего определенный этап, и его продукта будут снижены, так как уменьшается интенсивность синтеза предшественников. Если же нарушение расположено дистально, то, наоборот, концентрация увеличивается. По-видимому, экспрессия гена *BPGM* значительно подавляется при ингибировании *HK2* именно на фоне общего сокращения количества промежуточных продуктов гликолиза на предшествующих этапах. Вероятно, не во всех случаях количество продуктов гликолиза напрямую связано с изменениями экспрессии ферментов, катализирующих их образование, нами выявлено подавление только трех из них.

Таким образом, получены данные об изменении экспрессии генов при экспериментальном подавлении наиболее часто активирующегося в опухолях различных локализаций фермента гликолиза – гексокиназы 2 – на модели карциномы толстой кишки. Полученные данные важны для понимания результатов и последствий воздействия на ту или иную молекулу, рассматриваемую в качестве мишени при разработке способов таргетной терапии онкологических заболеваний.

Acknowledgments

The study was supported by the Russian Science Foundation, project 14-15-01083.

The authors are grateful to the staff of the National Medical Research Radiological Center, Ministry of Health of the Russian Federation, for providing access to its facilities, where the RKO cell line was cultured and experiments on lentivirus transduction followed by selection of stable clones were conducted. The authors also acknowledge the access to equipment of the Genome Shared Access Center, Institute of Molecular Biology, Russian Academy of Sciences (http://www.eimb.ru/rus/ckp/ccu_genome_c.php) and to computing facilities of Initium-Farm Company.

Conflict of interest

The authors declare no conflict of interest.

References

- Cho J., King J.S., Qian X., Harwood A.J., Shears S.B. Dephosphorylation of 2,3-bisphosphoglycerate by MIPK expands the regulatory capacity of the Rapoport-Luebering glycolytic shunt. *Proc. Natl. Acad. Sci. USA.* 2008;105(16):5998-6003. DOI 10.1073/pnas.0710980105.
- Dmitriev A.A., Kudryavtseva A.V., Krasnov G.S., Koroban N.V., Speranskaya A.S., Krintsina A.A., Belenikin M.S., Snehkina A.V., Sadriddinova A.F., Kishlyan N.V., Rozhmina T.A., Yurkevich O.Y., Muravenko O.V., Bolsheva N.L., Melnikova N.V. Gene expression profiling of flax (*Linum usitatissimum* L.) under edaphic stress. *BMC Plant Biol.* 2016;16(Suppl. 3):237. DOI 10.1186/s12870-016-0927-9.
- Funasaka T., Yanagawa T., Hogan V., Raz A. Regulation of phosphoglucose isomerase/autocrine motility factor expression by hypoxia. *FASEB J.* 2005;19(11):1422-1430. DOI 10.1096/fj.05-3699com.

- Graziano F., Ruzzo A., Giacomini E., Ricciardi T., Aprile G., Loupakis F., Lorenzini P., Ongaro E., Zoratto F., Catalano V., Sarti D., Rulli E., Cremolini C., De Nicolis M., De Maglio G., Falcone A., Fiorentini G., Magnani M. Glycolysis gene expression analysis and selective metabolic advantage in the clinical progression of colorectal cancer. *Pharmacogenomics J.* 2017;17(3):258-264. DOI 10.1038/tpj.2016.13.
- Haraldson K., Kashuba V.I., Dmitriev A.A., Senchenko V.N., Kudryavtseva A.V., Pavlova T.V., Braga E.A., Pronina I.V., Kondratov A.G., Rynditch A.V., Lerman M.I., Zabarovsky E.R. LRRC3B gene is frequently epigenetically inactivated in several epithelial malignancies and inhibit cell growth and replication. *Biochimie.* 2012;94:1151-1157. DOI 10.1016/j.biochi.2012.01.019.
- Hsu P.P., Sabatini D.M. Cancer cell metabolism: Warburg and beyond. *Cell.* 2008;134(5):703-707. DOI 10.1016/j.cell.2008.08.021.
- Izuishi K., Yamamoto Y., Sano T., Takebayashi R., Nishiyama Y., Mori H., Masaki T., Morishita A., Suzuki Y. Molecular mechanism underlying the detection of colorectal cancer by 18F-2-fluoro-2-deoxy-D-glucose positron emission tomography. *J. Gastrointest. Surg.* 2012;16(2):394-400. DOI 10.1007/s11605-011-1727-z.
- Katagiri M., Karasawa H., Takagi K., Nakayama S., Yabuuchi S., Fujishima F., Naitoh T., Watanabe M., Suzuki T., Unno M. Hexokinase 2 in colorectal cancer: a potent prognostic factor associated with glycolysis, proliferation and migration. *Histol. Histopathol.* 2017;32(4):351-360. DOI 10.14670/HH-11-799.
- Kim J.W., Dang C.V. Multifaceted roles of glycolytic enzymes. *Trends Biochem. Sci.* 2005;30(3):142-150. DOI 10.1016/j.tibs.2005.01.005.
- Krasnov G.S., Dmitriev A.A., Lakunina V.A., Kirpiy A.A., Kudryavtseva A.V. Targeting VDAC-bound hexokinase II: a promising approach for concomitant anti-cancer therapy. *Expert Opin. Ther. Targets.* 2013a;17(10):1221-1233. DOI 10.1517/14728222.2013.833607.
- Krasnov G.S., Dmitriev A.A., Sadritdinova A.F., Fedorova M.S., Snezhkina A.V., Melnikova N.V., Poteryakhina A.V., Nyushko K.M., Belyakov M.M., Kaprin A.D., Zaretsky A.R., Kudryavtseva A.V. Evaluation of gene expression of hexokinases in colorectal cancer with the use of bioinformatics methods. *Biofizika.* 2015;60(6):1050-1056. DOI 10.1134/S0006350915060172.
- Krasnov G.S., Dmitriev A.A., Snezhkina A.V., Kudryavtseva A.V. Deregulation of glycolysis in cancer: glyceraldehyde-3-phosphate dehydrogenase as a therapeutic target. *Expert Opin. Ther. Targets.* 2013b;17(6):681-693. DOI 10.1517/14728222.2013.775253.
- Kudryavtseva A.V., Fedorova M.S., Zhavoronkov A., Moskalev A.A., Zasedatelev A.S., Dmitriev A.A., Sadritdinova A.F., Karpova I.Y., Nyushko K.M., Kalinin D.V., Volchenko N.N., Melnikova N.V., Klimina K.M., Sidorov D.V., Popov A.Y., Nasedkina T.V., Kaprin A.D., Alekseev B.Y., Krasnov G.S., Snezhkina A.V. Effect of lentivirus-mediated shRNA inactivation of HK1, HK2, and HK3 genes in colorectal cancer and melanoma cells. *BMC Genet.* 2016a;17(Suppl. 3):156. DOI 10.1186/s12863-016-0459-1.
- Kudryavtseva A.V., Krasnov G.S., Dmitriev A.A., Alekseev B.Y., Kardymon O.L., Sadritdinova A.F., Fedorova M.S., Pokrovsky A.V., Melnikova N.V., Kaprin A.D., Moskalev A.A., Snezhkina A.V. Mitochondrial dysfunction and oxidative stress in aging and cancer. *Oncotarget.* 2016b;7(29):44879-44905. DOI 10.18632/oncotarget.9821.
- Kudryavtseva A.V., Lipatova A.V., Zaretsky A.R., Moskalev A.A., Fedorova M.S., Rasskazova A.S., Shibukhova G.A., Snezhkina A.V., Kaprin A.D., Alekseev B.Y., Dmitriev A.A., Krasnov G.S. Important molecular genetic markers of colorectal cancer. *Oncotarget.* 2016c;7(33):53959-53983. DOI 10.18632/oncotarget.9796.
- Mathupala S.P., Rempel A., Pedersen P.L. Glucose catabolism in cancer cells: identification and characterization of a marked activation response of the type II hexokinase gene to hypoxic conditions. *J. Biol. Chem.* 2001;276(46):43407-43412. DOI 10.1074/jbc.M108181200.
- Melnikova N.V., Dmitriev A.A., Belenikin M.S., Koroban N.V., Speranskaya A.S., Krinitsina A.A., Krasnov G.S., Lakunina V.A., Snezhkina A.V., Sadritdinova A.F., Kishlyan N.V., Rozhmina T.A., Klimina K.M., Amosova A.V., Zelenin A.V., Muravenko O.V., Bolshova N.L., Kudryavtseva A.V. Identification, expression analysis, and target prediction of flax genotroph microRNAs under normal and nutrient stress conditions. *Front. Plant Sci.* 2016;7:399. DOI 10.3389/fpls.2016.00399.
- Oparina N.Y., Snezhkina A.V., Sadritdinova A.F., Veselovskii V.A., Dmitriev A.A., Senchenko V.N., Melnikova N.V., Speranskaya A.S., Darii M.V., Stepanov O.A., Barkhatov I.M., Kudryavtseva A.V. Differential expression of genes that encode glycolysis enzymes in kidney and lung cancer in humans. *Russ. J. Genet.* 2013;49(7):707-716. DOI 10.1134/S1022795413050104.
- Pastorino J.G., Shulga N., Hoek J.B. Mitochondrial binding of hexokinase II inhibits Bax-induced cytochrome c release and apoptosis. *J. Biol. Chem.* 2002;277(9):7610-7618. DOI 10.1074/jbc.M109950200.
- Siegel R.L., Miller K.D., Fedewa S.A., Ahnen D.J., Meester R.G.S., Barzi A., Jemal A. Colorectal cancer statistics, 2017. *CA: Cancer J. Clin.* 2017;67(3):177-193. DOI 10.3322/caac.21220.
- Snezhkina A.V., Krasnov G.S., Zaretsky A.R., Zhavoronkov A., Nyushko K.M., Moskalev A.A., Karpova I.Y., Afremova A.I., Lipatova A.V., Kochetkov D.V., Fedorova M.S., Volchenko N.N., Sadritdinova A.F., Melnikova N.V., Sidorov D.V., Popov A.Y., Kalinin D.V., Kaprin A.D., Alekseev B.Y., Dmitriev A.A., Kudryavtseva A.V. Differential expression of alternatively spliced transcripts related to energy metabolism in colorectal cancer. *BMC Genomics.* 2016;17(Suppl. 14):1011. DOI 10.1186/s12864-016-3351-5.
- Wang W., Liu Z., Zhao L., Sun J., He Q., Yan W., Lu Z., Wang A. Hexokinase 2 enhances the metastatic potential of tongue squamous cell carcinoma via the SOD2-H2O2 pathway. *Oncotarget.* 2017;8(2):3344-3354. DOI 10.18632/oncotarget.13763.
- Wang J., Zhang P., Zhong J., Tan M., Ge J., Tao L., Li Y., Zhu Y., Wu L., Qiu J., Tong X. The platelet isoform of phosphofructokinase contributes to metabolic reprogramming and maintains cell proliferation in clear cell renal cell carcinoma. *Oncotarget.* 2016;7(19):27142-27157. DOI 10.18632/oncotarget.8382.
- Warburg O. Origin of cancer cells. *Oncologia.* 1956;9(2):75-83. DOI 10.1126/science.123.3191.309.
- Wilson J.E. Isozymes of mammalian hexokinase: structure, subcellular localization and metabolic function. *J. Exp. Biol.* 2003;206(12):2049-2057. DOI 10.1242/jeb.00241.
- Yeh C.S., Wang J.Y., Chung F.Y., Lee S.C., Huang M.Y., Kuo C.W., Yang M.J., Lin S.R. Significance of the glycolytic pathway and glycolysis related-genes in tumorigenesis of human colorectal cancers. *Oncol. Rep.* 2008;19(1):81-91. DOI 10.3892/or.19.1.81.



Отсроченные эффекты хирургического воздействия, перенесенного матерью в период ранней беременности, на артериальное давление и поведение потомков у крыс линии OXYS

И.Н. Рожкова¹, Т.Н. Игонина¹, Д.С. Рагаева¹, О.М. Петрова^{1,2}, Е.Ю. Брусенцев¹, В.А. Напримеров^{1,3}, С.Я. Амстиславский¹

¹ Федеральное исследовательское учреждение Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия

³ Новосибирский государственный аграрный университет, Новосибирск, Россия

Применение некоторых вспомогательных репродуктивных технологий, в частности трансплантации эмбрионов, может вызывать различные физиологические и поведенческие изменения у потомства. Целью нашего исследования являлась проверка влияния хирургического вмешательства, используемого при трансплантации эмбрионов, на вес, артериальное давление и поведение в тестах «открытое поле» (ТОП) и «приподнятый крестообразный лабиринт» (ПКЛ) у взрослых потомков. В работе исследованы отсроченные эффекты хирургического воздействия на четвертые сутки беременности на потомков крыс линии OXYS. Самок линии OXYS спаривали с фертильными самцами той же линии, через 96 ч после обнаружения сперматозоидов во влагалищных мазках им проводили хирургическую операцию, имитирующую трансплантацию эмбрионов. У потомков самок, подвергавшихся хирургическому воздействию во время беременности (группа OXYS-PS), в возрасте 3 мес. измеряли вес тела, систолическое (САД) и диастолическое (ДАД) артериальное давление, а также исследовали поведение в тестах ТОП и ПКЛ. Контролем служили потомки интактных крыс OXYS. Средний вес у крыс линии OXYS не отличался от такового у крыс OXYS-PS. Крысы линий OXYS и OXYS-PS имели высокое САД, превышающее гипертензивный порог (150 мм рт. ст.), и высокое ДАД. При этом данные показатели достоверно выше у крыс OXYS-PS, чем у животных контрольной линии. У крыс OXYS-PS в тесте ТОП было снижено время, проведенное в центре арены, меньше исследованная область, а также снижено число стоек и их продолжительность по сравнению с крысами OXYS. Как показано в тесте ПКЛ, число выглядываний из закрытых рукавов и их длительность меньше у крыс OXYS-PS, чем в контроле. Таким образом, хирургическое воздействие, испытанное самкой OXYS на ранних стадиях беременности, приводит к следующим эффектам у потомков: повышение САД и ДАД, снижение исследовательской активности и повышение тревожности.

Ключевые слова: крысы OXYS; беременность; хирургическое воздействие; отсроченные эффекты.

Long-term effects of maternal exposure to surgical stress at the earliest stage of pregnancy on blood pressure and behavior in offspring of OXYS rats

I.N. Rozhkova¹, T.N. Igonina¹, D.S. Ragaeva¹, O.M. Petrova^{1,2}, E.Yu. Brusentsev¹, V.A. Naprimerov^{1,3}, S.Ya. Amstislavsky¹

¹ Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

² Novosibirsk State University, Novosibirsk, Russia

³ Novosibirsk State Agrarian University, Novosibirsk, Russia

The use of some assisted reproductive technologies, in particular, embryo transfer, may cause various physiological and behavioral changes in the offspring. The purpose of our study was to study the effects of surgery (which is used for embryo transfer) done with pregnant dams on the weight, blood pressure and behavior in the open field and elevated plus-maze tests in adult offspring. Thus, long-term effects on the offspring after maternal exposure to surgical stress given to dams at the 4th day of pregnancy were studied in OXYS rats. OXYS females were mated in estrus with fertile males of the same strain. 96 hours after spermatozoa were found in vaginal smears the surgery (sham operation, imitating embryo transfer) was performed. Body weight (BW), systolic (SAP) and diastolic (DAP) arterial pressure as well as behavior in open field (OF) and elevated plus maze (EPM) tests were studied in the offspring of females exposed to surgical treatment during pregnancy (OXYS-PS) at the age of 3 mo. Untreated offspring of OXYS rats were used as controls. BW in naturally born OXYS rats did not differ from those of the OXYS-PS group. OXYS and OXYS-PS rats exhibited higher SAP (more than 150 mm Hg) and DAP; it is noteworthy that both SAP and DAP were higher in the OXYS-PS group than in the control group. The time spent in the center of arena, the area studied, the time and number of rearing were decreased in OXYS-PS rats in the OF test as compared to the OXYS controls. Moreover, OXYS-PS rats were characterized by the absence of grooming in the OF test. As was demonstrated by the EPM test, the duration and numbers of peeking out from closed arms were decreased

Received 03.08.2017

Accepted for publication 05.10.2017

© AUTHORS, 2017

in the OXYS-PS rats as compared to the OXYS controls. Thus, OXYS dams' exposure to surgical stress at their early pregnancy led to such effects in the offspring as elevated SAP and DAP, decreased overall activity and increased anxiety.

Key words: OXYS rats; pregnancy; surgical treatment; long-term effects.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Рожкова И.Н., Игонина Т.Н., Рагаева Д.С., Петрова О.М., Брусенцев Е.Ю., Напримеров В.А., Амстиславский С.Я. Отсроченные эффекты хирургического воздействия, перенесенного матерью в период ранней беременности, на артериальное давление и поведение потомков у крыс линии OXYS. Вавиловский журнал генетики и селекции. 2017;21(8):937-942. DOI 10.18699/VJ17.316

HOW TO CITE THIS ARTICLE:

Rozhkova I.N., Igonina T.N., Ragaeva D.S., Petrova O.M., Brusentsev E.Yu., Naprimerov V.A., Amstislavsky S.Ya. Long-term effects of maternal exposure to surgical stress at the earliest stage of pregnancy on blood pressure and behavior in offspring of OXYS rats. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2017;21(8):937-942. DOI 10.18699/VJ17.316 (in Russian)

Взаимодействие эмбриона с материнским организмом во время беременности играет важную роль в формировании здоровья будущего ребенка. В соответствии с гипотезой DOHaD (Developmental Origin of Health and Disease) стресс и другие негативные воздействия на мать во время беременности вызывают долгосрочные последствия, такие как задержка роста, гипертония и инсулин-независимый диабет в постнатальный период (Barker, 2000). Исследование воздействия пренатального стресса на поведение будущих потомков, в частности повышение у них уровня тревожности, показано в работах (Wilson et al., 2013; Said et al., 2015; Badache et al., 2017).

Экспериментально подтверждено, что гипотеза DOHaD применима и к преимплантационной стадии пренатального онтогенеза (Kwong et al., 2000; Брусенцев и др., 2014; Игонина и др., 2016). Хирургическое воздействие является одним из видов стресса, который возникает при трансплантации эмбрионов. Он может оказывать влияние на развивающийся организм в пренатальный период и иметь отсроченные последствия, что требует изучения. Ранее было показано, что применение некоторых вспомогательных репродуктивных технологий (ВРТ), в частности трансплантации эмбрионов, может привести к изменениям артериального давления у потомства линий крыс с гипертонией (Di Nicolantonio et al., 2006; Lee, Azar, 2010; Рагаева и др., 2014).

Влияние трансплантации эмбрионов на изменение АД у крыс гипертонивных линий подробно описано в обзоре (Рагаева и др., 2014). В работах (Dene, Rapp, 1985; Gray, 1991) исследователи не наблюдали никаких эффектов по влиянию трансплантации эмбрионов крысам реципиентам-нормотоникам на АД крыс гипертонивных линий, однако в более поздних исследованиях обнаружено смягчение и/или задержка развития гипертонии у крыс линий SHR (Di Nicolantonio et al., 2006; Lee, Azar, 2010) и Dahl (Kubisch, Gomez-Sanchez, 1999). В экспериментах на крысах линии НИСАГ данные были противоречивы: при трансплантации эмбрионов крысам-реципиентам нормотонивных линий наблюдалось не смягчение, а, напротив, усиление проявления гипертонии (Amstislavsky et al., 1996; Рагаева и др., 2015), при трансплантации же в матку крыс того же гипертонивного генотипа обнаружено смягчение гипертонии (Amstislavsky et al., 1996). Поскольку трансплантация у крыс производится хирургически, при объяснении результатов следует учитывать

возможные последствия этого стрессирующего воздействия. Необходимо отметить, что влияние хирургического воздействия, испытываемого матерью на самых ранних стадиях беременности, на поведение потомков практически не изучалось.

Линия крыс OXYS, созданная в Институте цитологии и генетики СО РАН, является уникальной генетической моделью преждевременного старения и связанных с ним заболеваний. В настоящее время линия насчитывает уже свыше 100 поколений селекции, хорошо изучена и характеризуется комплексом различных нарушений (Kolosova et al., 2014), в том числе артериальной гипертонией (Shteklina et al., 2005). Влияние трансплантации эмбрионов на АД хорошо известно, но какой именно фактор влияет на изменение показателей АД: материнская внутриутробная среда, культивирование эмбрионов *in vitro* или же сама процедура трансплантации, не установлено.

Цель нашего исследования – выяснение эффектов одного из вышеперечисленных факторов, а именно процедуры трансплантации, т. е. хирургического воздействия, оказываемого на материнский организм на наиболее ранних стадиях беременности; массу тела рожденного потомства, артериальное давление и поведение этих потомков во взрослом состоянии в тестах «открытое поле» (ТОП) и «приподнятый крестообразный лабиринт» (ПКЛ) у крыс линии OXYS.

Материалы и методы

Экспериментальные животные. Использовали крыс линии OXYS в возрасте трех месяцев. Животных содержали в SPF-виварии Института цитологии и генетики СО РАН в индивидуально вентилируемых клетках (OptiRAT, Animal Care, США) при температуре 22–24 °С и влажности 40–50 %; в качестве подстилки использовали опилки. Суточный цикл состоял из 12 ч дня и 12 ч ночи, с восходом в 4 ч утра и закатом в 16 ч вечера. Животные имели постоянный доступ к автоклавируемому стандартизированному корму (V1534-300, Sniff, Soest, Германия) и очищенной воде, обогащенной минеральными добавками (Экопроект, Россия). Взрослых самцов крыс в возрасте 12–14 нед содержали в группах по две-три особи.

Все экспериментальные процедуры одобрены Комиссией по биоэтике Института цитологии и генетики СО РАН (протокол № 12 от 6 декабря 2012 г.) и согласуются с Европейской конвенцией о защите позвоночных жи-

вотных, используемых для экспериментальных и других научных целей. Исследования проведены в Центре генетических ресурсов лабораторных животных Института цитологии и генетики СО РАН (RFMEFI61914X0005, RFMEFI62114X0010).

Самцы крыс линии OXYS были разделены на две группы: OXYS – интактные животные ($n = 13$), группу формировали из четырех пометов и OXYS-PS – потомки самок, подвергавшихся хирургическому воздействию во время беременности ($n = 10$), – из пяти пометов. Исследовано поведение крыс в тестах ТОП и ПКЛ, измерены масса тела и артериальное давление.

Хирургическое воздействие. У самок линии OXYS ($n = 9$) определяли стадию эструса по влагалищным мазкам. По достижении данной стадии их спаривали с фертильными самцами этой же линии на ночь. Наличие вагинальных пробок свидетельствовало о фертильном спаривании. Именно таких самок брали в эксперимент. Через 96 ч после спаривания с фертильными самцами самкам линии OXYS экспериментальной группы (OXYS-PS, $n = 5$) проводили полостную операцию, аналогичную той, которую проводят при трансплантации эмбрионов. Для анестезии самке-реципиенту вводили медитин (0.02 мл на крысу) внутримышечно, через 10 мин – кетамин (0.2 мл на крысу) внутривенно. Затем подкожно вводили 0.02 мл амоксициллина (тригидрат амоксициллина, 150 мг/мл). Волосы в области операционного поля (правая сторона спины, 1 см от нижнего ребра в вентральном направлении) выбривали, а кожу обрабатывали 70 % спиртом. Кожу и лежащий под ней мышечный слой в области матки разрезали дорсовентрально. Висцеральный жир вместе с яичником, яйцеводом и верхней частью матки захватывали пинцетом и доставали наружу. В правый рог матки с помощью стеклянного капилляра вводили 5 мкл среды FertiCult Flushing medium (FertiPro, Бельгия). Самки OXYS интактной группы ($n = 4$) не подвергались хирургической операции во время беременности и служили контролем.

Измерение артериального давления и веса тела. Систолическое (САД) и диастолическое (ДАД) артериальное давление измеряли с помощью двухканальной неинвазивной системы измерения кровяного давления с высокой пропускной способностью (CODA™) (Kent Scientific, Torrington, США). Непосредственно перед измерением артериального давления неанестезированных животных осторожно помещали в рестриктор и взвешивали с помощью весов (ScoutPro SPS2001 F, Ohaus Corporation, США). Для привыкания к экспериментальным условиям крысам было дано минимум 5 мин. Во время измерений крыс выдерживали на нагретой платформе (35 °С). Для каждого животного выполнено не менее шести повторов измерений; при статистическом анализе использовано усредненное значение.

Поведенческие тесты. Поведение взрослых крыс изучали с помощью тестов ТОП и ПКЛ. Процедура габитуации к рукам оператора выполнена за 24 ч до тестирования. Каждое животное помещали в чистую индивидуальную вентилируемую клетку и транспортировали в затемненную комнату для наблюдений с красным светом (20 Вт) за 30–60 мин до тестирования. Все тесты проводили в

темный световой период, начиная с 16 ч. После каждого теста животное возвращали в свою клетку. Следующий тест повторяли через 24 ч. Все устройства, используемые для поведенческих тестов, очищали после животного с помощью 0.5 % Combidez. Поведение крыс в каждом тесте регистрировали датчиком изображения Kinect 3-D, расположенным над аппаратом, и оригинальным программным обеспечением EthoStudio software (Kulikov et al., 2008).

Тест ТОП традиционно используется для оценки поведенческих профилей у грызунов (Belzung, Griebel, 2001). Он выполнен в модификации, которая в настоящее время нами используется (Amstislavsky et al., 2001; Kulikov et al., 2008). Устройство для тестирования представляло собой круглую арену с отверстиями в полу ($d = 2$ см), изготовленную из белого поливинилхлорида ($d = 97$ см), разделенную на секторы ($S = 13$ см²) (OpenScience, Россия). Арена была ограничена бортом высотой 42 см и освещена лампой накаливания красного света, расположенной на расстоянии 130 см над поверхностью. Продолжительность теста составляла 5 мин. Фиксировали общепринятые параметры ТОП: время, проведенное в центре арены; пройденное расстояние; исследованная область; число и длительность грумингов; латентный период первого груминга; число и длительность вертикальных стоек; латентный период первой стойки. Все параметры были записаны на видео, а затем проанализированы с помощью программного обеспечения EthoStudio.

Тест ПКЛ используется для оценки тревожности и общей активности, часто в сочетании с ТОП (Carobrez, Bertoglio, 2005). Устройство, используемое в данном исследовании, состояло из двух открытых ($50 \times 14 \times 1$ см) и двух закрытых ($50 \times 14 \times 30$ см) рукавов, изготовленных из поливинилхлорида и приподнятых на 55 см над поверхностью пола. Лабиринт был освещен лампой накаливания (40 Вт), расположенной сверху на высоте 130 см. Продолжительность теста составляла 5 мин. Регистрировали стандартные параметры: время, проведенное в открытых и закрытых рукавах, количество заходов в открытые и закрытые рукава, общее количество переходов (Rodgers, Cole, 1994). Оценивали также некоторые дополнительные поведенческие параметры: пройденное расстояние, частоту, латентность и продолжительность свешиваний вниз головой из открытых рукавов и частоту, латентность и продолжительность выглядываний из закрытых рукавов (Cruz et al., 1994; Carobrez, Bertoglio, 2005). Индекс тревожности рассчитан по (Lee et al., 2016).

Статистический анализ. Различия в весе и уровне артериального давления и поведенческие данные у потомков, рожденных интактными матерями и самками, подвергавшимися хирургическому воздействию, проанализированы с помощью t -критерия Стьюдента при использовании стандартного пакета программного обеспечения STATISTICA V 8.0 (StatSoft, Inc). Данные в табл. 1–3 представлены как среднее \pm SEM, значения при $p < 0.05$ считались статистически значимыми.

Результаты

Получено потомство от девяти самок крыс OXYS, участвующих в эксперименте, при этом плодовитость самок

Table 1. Arterial blood pressure and body weight in the OXYS and OXYS-PS rat groups

Group	SBP (mmHg)	DBP (mmHg)	Body weight (g)
OXYS (<i>n</i> = 13)	147.3 ± 3.58	103.3 ± 3.30	296.3 ± 4.29
OXYS-PS (<i>n</i> = 10)	164.55 ± 5.27*	119.61 ± 4.07**	295.92 ± 4.37

* *p* < 0.05; ** *p* < 0.01 as compared to OXYS.

Table 2. Behavioral parameters of OXYS rats in the OF test

Parameters	Group (number of rats per group)	
	OXYS (<i>n</i> = 13)	OXYS-PS (<i>n</i> = 10)
Time in the center (% of the test time)	19.5 ± 4.0	4.7 ± 1.21**
Explored area (% of the whole area)	76.3 ± 2.94	59.4 ± 4.95**
Distance travelled (m)	7.1 ± 4.40	7.3 ± 0.48
Grooming bouts (<i>n</i>)	0.7 ± 0.36	0
Grooming duration (s)	0.6 ± 0.34	0
Grooming latency (s)	263.9 ± 18.25	300 ± 0.0
The number of rearings (<i>n</i>)	18.4 ± 1.62	12.1 ± 1.55*
Rearing duration (s)	25.0 ± 3.61	14.5 ± 2.31*
Rearing latency (s)	10.9 ± 5.64	17.3 ± 3.35

* *p* < 0.05; ** *p* < 0.01 as compared to control.

Table 3. Behavioral parameters of OXYS rats in the EPM test

Parameters	Group (number of rats per group)	
	OXYS (<i>n</i> = 13)	OXYS-PS (<i>n</i> = 10)
Time in open arms (% of the test time)	21.79 ± 7.67	17.39 ± 4.95
Time in closed arms (% of the test time)	69.60 ± 7.67	72.70 ± 6.76
Distance travelled (m)	3.50 ± 0.42	3.95 ± 0.44
Head dips (<i>n</i>)	5.50 ± 1.42	5.00 ± 1.44
Head dip duration (s)	7.02 ± 1.86	5.25 ± 1.66
Head dip latency (s)	15.34 ± 7.25	20.13 ± 7.71
Peeking out from closed arms (<i>n</i>)	6.00 ± 1.79	2.40 ± 0.60
Duration of peeking (s)	32.16 ± 10.89	6.09 ± 1.51*
Latency of peeking (s)	49.66 ± 27.11	80.09 ± 23.89
Anxiety index	0.66 ± 0.07	0.70 ± 0.04

* *p* < 0.05 as compared to control.

крыс, перенесших стресс на ранней стадии беременности, не отличалась от таковой у интактных крыс (10.8 ± 0.95 в сравнении с 9.2 ± 0.8 соответственно).

Сравнение артериального давления и веса тела крыс линии OXYS и группы OXYS-PS приведено в табл. 1. В возрасте трех месяцев средний вес у взрослых крыс линии OXYS не отличается от среднего веса крыс OXYS-PS. В этом же возрасте крысы линий OXYS и OXYS-PS имеют высокое САД, превышающее гипертензивный порог 150 мм рт. ст. Данный показатель был достоверно выше у крыс OXYS-PS, чем у крыс OXYS (*p* < 0.05). Аналогичные результаты получены по ДАД для крыс линий OXYS и

OXYS-PS. У крыс OXYS-PS этот показатель также достоверно выше, чем у крыс OXYS (*p* < 0.01) (см. табл. 1).

Данные по поведению крыс OXYS опытной и контрольной групп в ТОП представлены в табл. 2. Самцы крыс из группы OXYS-PS проводили в центре арены меньше времени, чем самцы линии OXYS (*p* < 0.01). Также исследованная область была меньше (*p* < 0.01) у крыс OXYS-PS по сравнению с контролем. Продолжительность и число стоек было меньше (*p* < 0.05) у крыс OXYS-PS по сравнению с контролем.

В табл. 3 приведены показатели по поведению крыс OXYS опытной и контрольной групп в ПКЛ. Длитель-

ность выглядываний из закрытых рукавов была меньше ($p < 0.05$) у крыс OXYS-PS по сравнению с контролем.

Обсуждение

В настоящей работе показано, что вес потомства не зависит от хирургического воздействия на преимплантационной стадии беременности. Имеются данные о снижении веса тела новорожденных детей, полученные в результате применения ВРТ у человека, даже в тех случаях, когда трансплантировали только один эмбрион (Schieve et al., 2002), хотя в некоторых работах различий по весу тела новорожденных, полученных после применения ВРТ на человеке (Scherrer et al., 2012) и крысах (Игониная и др., 2016; Di Nicolantonio et al., 2006), не обнаружено. В исследовании, проведенном на мышах (Sjoblom et al., 2005), показано, что изменение веса тела потомства вызвано не самими процедурами ВРТ, а неоптимальными условиями их культивирования *in vitro*. На гипертензивных крысах линии SHR продемонстрировано (Lee, Azar, 2010), что на вес тела крысят в помете влияет линия вскармливающей их самки, а не применение ВРТ. В другом эксперименте на этой же линии было показано, что трансплантация эмбрионов на вес потомства не влияет (Di Nicolantonio et al., 2006). Эти данные хорошо согласуются с нашими выводами об отсутствии эффекта по весу тела у крыс линии OXYS, рожденных после хирургического воздействия на преимплантационной стадии беременности. Совокупность всех результатов показывает, что у крыс такая физиологическая характеристика, как вес тела, достаточно стабильна и не зависит от стресса, вызванного хирургическим воздействием на преимплантационной стадии беременности.

В нашем исследовании крысы экспериментальной группы OXYS-PS имели более высокое САД и ДАД, чем крысы контрольной группы, что, по всей видимости, может быть связано с хирургическим воздействием, схожим с тем, что испытывает самка-реципиент при трансплантации эмбрионов. Противоречивый характер имеют данные по изменению АД у потомков, полученных в результате исследований на разных гипертензивных линиях крыс. Так, ранее сообщалось об отсутствии влияния трансплантации эмбрионов на АД у новорожденных в постнатальном онтогенезе (Dene, Rapp, 1985; Gray, 1991). Однако более поздние исследования показали, что, если эмбрионы гипертензивных крыс развиваются в матке нормотензивной самки-реципиента, то у полученного потомства происходит либо смягчение гипертензии, либо задержка ее наступления (Di Nicolantonio et al., 2006; Lee, Azar, 2010). Хотя в ряде исследований не обнаружено изменений АД у потомков нормотензивных крыс, пересаженных на стадии эмбриона гипертензивным самкам-реципиентам (Рагаева и др., 2014), в некоторых работах, напротив, показано, что АД повышается у потомков крыс без предрасположенности к гипертензии, если они рождаются после трансплантации эмбрионов от самки-реципиента с гипертензией (Lee, Azar, 2010). Это несоответствие в результатах исследований может быть объяснено различиями в выборе линии донора и реципиента, а также используемых методов.

Результаты нашего исследования на крысах OXYS, как и данные более ранних работ (Wilson et al., 2013; Badache

et al., 2017), свидетельствуют о том, что животные, рожденные после стресса матери во время беременности, имеют сниженную исследовательскую активность по сравнению с интактными. Действительно, тест ТОП выявил, что крысы экспериментальной группы достоверно меньше времени проводили в центре платформы. В данном тесте у них были меньше исследованная область и меньшее число вертикальных стоек, а также время, проведенное в стойках.

Кроме того, в нашей работе показано, что крысы, рожденные после хирургического воздействия, отличались повышенной тревожностью по сравнению с контрольной группой, о чем свидетельствует снижение длительности выглядываний из закрытых рукавов в тесте ПКЛ. Данные ранее проведенных исследований находятся в полном соответствии с нашими результатами. Потомки стрессированных во время беременности матерей проявляли повышенную тревожность в ТОП и/или ПКЛ (Said et al., 2015; Badache et al., 2017).

Вероятно, хирургическое воздействие на преимплантационной стадии беременности негативно влияет на развивающийся эмбрион и, согласно гипотезе Д. Баркера (Barker, 2000), может приводить к изменению некоторых физиологических и поведенческих характеристик у потомства в постнатальном онтогенезе. Хирургическое воздействие, произведенное на самых ранних этапах беременности, имеет пролонгированный эффект, так как восстановление после операции занимает не менее недели. Таким образом, последствия этого стресса проявляются не только на преимплантационной стадии, но и на более поздних этапах беременности. Именно с этими воздействиями в раннем пренатальном онтогенезе связаны обнаруженные нами эффекты у потомков – повышение САД и ДАД, снижение исследовательской активности и повышение тревожности.

Acknowledgments

This study was supported by the Russian Foundation for Basic Research, project 15-04-05509, and State Budgeted Project 0324-2016-0002 for the Human and Animal Genetics Department, Institute of Cytology and Genetics, Novosibirsk. Experiments were conducted at the SPF vivarium Shared Access Center, Institute of Cytology and Genetics, projects RFMEFI61914X0005 and RFMEFI62114X0010.

Conflict of interest

The authors declare no conflict of interest.

References

- Amstislavsky S.Y., Alekhina T.A., Barykina N.N., Chuguy V.F., Petrenko O.I., Kolpakov V.G. Effects of change of maternal environment during early postnatal development on behaviour in cataleptic rats. *Behav. Proc.* 2001;56(1):41-47. PMID: 11566236.
- Amstislavsky S., Amstislavskaya T., Stein M., Maksimovsky L., Markel A., Ternovskaya Yu., Ternovsky D., Amstislavsky S. Embryo cryobanking for conserving laboratory and wild animal species. *Scand. J. Lab. Anim. Sci.* 1996;23:269-277.
- Badache S., Bouslama S., Brahmia O., Baïri A.M., Tahraoui A.K., Ladjama A. Prenatal noise and restraint stress interact to alter exploratory behavior and balance in juvenile rats, and mixed stress reverses these effects. *Stress.* 2017;20:320-328. DOI 10.1080/10253890.2017.1307962.

- Barker D.J. *In utero* programming of cardiovascular disease. *Theriogenology*. 2000;53:555-574.
- Belzung C., Griebel G. Measuring normal and pathological anxiety-like behaviour in mice: a review. *Behav. Brain Res.* 2001;125(1-2):141-149. PMID: 11682105.
- Brusentsev E.Yu., Igonina T.N., Amstislavsky S.Ya. Traditional and modern approaches to the culture of preimplantation mammalian embryos *in vitro*. *Ontogenez = Ontogenesis (Moscow)*. 2014;45(2):73-88. DOI 10.7868/S0475145014020037. (in Russian)
- Carobrez A.P., Bertoglio L.J. Ethological and temporal analyses of anxiety-like behavior: the elevated plus-maze model 20 years on. *Neurosci. Biobehav.* 2005;29:1193-1205. DOI 10.1016/j.neubiorev.2005.04.017.
- Cruz A.P., Frei F., Graeff F.G. Ethopharmacological analysis of rat behavior on the elevated plus-maze. *Pharmacol. Biochem. Behav.* 1994;49(1):171-176. PMID: 7816869.
- Dene H., Rapp J.P. Lack of effects of maternal salt intake on blood pressure of offspring in Dahl salt-sensitive rats. *Clin. Exp. Hypertens.* 1985;7:1121-1133.
- Di Nicolantonio R., Koutsis K., Westcott K.T., Wlodek M.E. Relative contribution of the prenatal versus postnatal period on development of hypertension and growth rate of the spontaneously hypertensive rat. *Clin. Exp. Pharm. Phys.* 2006;33:9-16.
- Gray S.D. Reciprocal embryo transfer between SHR and WKY. II. Effect on cardiovascular development. *Clin. Exp. Hypertens. Part A. Theory Pract.* 1991;13:963-969.
- Igonina T.N., Ragaeva D.S., Prokudina O.I., Brusentsev E.Yu., Rozhkova I.N., Abramova T.O., Amstislavsky S.Ya. Effects of *in vitro* culture and embryo transfer on arterial blood pressure, body weight and behaviour in ISIAH rats. *Rossiyskiy fiziologicheskiy zhurnal im. I.M. Sechenova = I.M. Sechenov Physiological Journal*. 2016;102(9):1089-1098. (in Russian)
- Kolosova N.G., Stefanova N.A., Korbolina E.E., Fursova A.Zh., Kozhevnikova O.S. Senescence-accelerated OXYS rats: A genetic model of premature aging and age-related diseases. *Adv. Gerontol.* 2014;4:294-298.
- Kubisch H.M., Gomez-Sanchez E.P. Embryo transfer in the rat as a tool to determine genetic components of the gestational environment. *Lab. Anim. Sci.* 1999;49(1):90-94. PMID: 10090101.
- Kulikov A.V., Tikhonova M.A., Kulikov V.A. Automated measurement of spatial preference in the open field test with transmitted lighting. *J. Neurosci. Meth.* 2008;170(2):345-351. DOI 10.1016/j.jneumeth.2008.01.024.
- Kwong W.Y., Wild A.E., Roberts P., Willis A.C., Fleming T.P. Maternal undernutrition during the preimplantation period of rat development causes blastocyst abnormalities and programming of postnatal hypertension. *Development*. 2000;127:4195-4202.
- Lee B., Sur B., Cho S.G., Yeom M., Shim I., Lee H., Hahm D.H. Ginsenoside Rb1 rescues anxiety-like responses in a rat model of post-traumatic stress disorder. *J. Nat. Med.* 2016;70(2):133-144. DOI 10.1007/s11418-015-0943-3.
- Lee J.Y., Azar S.H. Wistar-Kyoto and spontaneously hypertensive rat blood pressure after embryo transfer into different wombs and cross-suckling. *Exp. Biol. Med. (Maywood)*. 2010;235:1375-1384.
- Ragaeva D.S., Abramova T.O., Rozhkova I.N., Brusentsev E.Yu., Kalinichenko E.V., Igonina T.N., Amstislavsky S.Ya. Effects of reproductive technologies and the SPF status on some physiological and behavioral characteristics in rats with arterial hypertension (ISIAH strain). *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2015;19(4):383-387. DOI 10.18699/VJ15.048. (in Russian)
- Ragaeva D.S., Brusentsev E.Yu., Amstislavsky S.Ya. Assisted reproductive technologies and arterial hypertension. *Ontogenez = Ontogenesis (Moscow)*. 2014;45(5):299-313 DOI 10.7868/S0475145014050085. (in Russian)
- Rodgers R.J., Cole J.C. Anxiolytic-like effect of (S)-WAY 100135, a 5-HT1A receptor antagonist, in the murine elevated plus-maze test. *Eur. J. Pharmacol.* 1994;261(3):321-325. PMID: 7813555.
- Said N., Lakehayli S., Battas O., Hakkou F., Tazi A. Effects of prenatal stress on anxiety-like behavior and nociceptive response in rats. *J. Integr. Neurosci.* 2015;14:223-234. DOI 10.1142/S0219635215500107.
- Scherrer U., Rimold S.F., Rexhaj E., Stuber T., Duplain H., Garcin S., de Marchi S.F., Nicod P., Germond M., Allemann Y., Sartori C. Systemic and pulmonary vascular dysfunction in children conceived by assisted reproductive technologies. *Circulation*. 2012;125:1890-1896.
- Schieve L.A., Meikle S.F., Ferre C., Peterson H.B., Jeng G., Wilcox L.S. Low and very low birth weight in infants conceived with use of assisted reproductive technology. *N. Engl. J. Med.* 2002;346:731-737.
- Shteklina T.A., Kolosova N.G., Terner A.Ia. Hypertensive syndrome and changes of water-salt metabolism in the OXYS rats with early ageing. *Russ. Fiziol. Zh.* 2005;91:1213-1220.
- Sjoblom C., Roberts C.T., Wikland M., Robertson S.A. Granulocyte-macrophage colony-stimulating factor alleviates adverse consequences of embryo culture on fetal growth trajectory and placental morphogenesis. *Endocrinology*. 2005;146:2142-2153.
- Wilson C.A., Vazdarjanova A., Terry A.V. Jr. Exposure to variable prenatal stress in rats: effects on anxiety-related behaviors, innate and contextual fear, and fear extinction. *Behav. Brain Res.* 2013;238:279-288. DOI 10.1016/j.bbr.2012.10.003.



Parasites of the genus *Nosema*, *Crithidia* and *Lotmaria* in the honeybee and bumblebee populations: a case study in India

V.Y. Vavilova¹✉, I. Konopatskaia^{1, 2}, S.L. Luzyanin³, M. Woyciechowski⁴, A.G. Blinov^{1, 5}

¹ Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

² Novosibirsk State University, Novosibirsk, Russia

³ Institute of Biology, Ecology and Natural Resources, Kemerovo State University, Kemerovo, Russia

⁴ Institute of Environmental Sciences, Jagiellonian University, Krakow, Poland

⁵ Institute of Systematics and Ecology SB RAS, Novosibirsk, Russia

The populations of honeybees and bumblebees have been decreasing around the world in the recent decades. A variety of pathogens and parasites, including bacteria, fungi, protozoa, nematodes, mites and insects play significant role in honeybee and bumblebee colonies loss. Parasites of the genus *Nosema* (Microsporidia: Nosematidae) and the genera *Crithidia* and *Lotmaria* (Kinetoplastida: Trypanosomatidae) have a significant negative impact on honeybee and bumblebee colonies. Recent studies of nuclear DNA markers of these parasites allowed to describe new species and genetic variants. The aim of this study was to investigate the Microsporidia (*Nosema* spp.) and Trypanosomatidae (*Crithidia* spp. and *Lotmaria passim*) prevalence and genetic diversity in honeybee and bumblebee populations of Indian territories that haven't been studied before. In total 119 specimens of 4 honeybee and 5 bumblebee species were analyzed in this study. The prevalence of parasites in honeybee and bumblebee populations of the two Indian states (Jammu and Kashmir, Karnataka) were identified using PCR with primers specific for the ribosomal RNA genes cluster of *Nosema*, *Crithidia* and *Lotmaria* species. Co-infection by microsporidian and trypanosomatid parasites was detected in several honeybee and bumblebee specimens from Jammu and Kashmir state. Comparative analysis of ribosomal RNA genes sequences showed that honeybee samples from India studied were infected by *N. bombi*, *N. ceranae* and *L. passim*. Bumblebee populations were infected by *Nosema D*, *Crithidia bombi* and *Crithidia expoeki*. No honeybee's specimen with trypanosomatid infection was found in Karnataka state. For the first time *N. bombi* infection was detected in the honeybee population. The studies of distribution of microsporidia and trypanosomatid parasites among the honeybee and bumblebee populations all over the World were summarized and supplemented.

Key words: honeybees; bumblebees; infection; ribosomal RNA genes; *Nosema* spp.; *Crithidia* spp.; *Lotmaria passim*.

Паразитические организмы родов *Nosema*, *Crithidia* и *Lotmaria* в популяциях пчел и шмелей: исследование в Индии

В.Ю. Вавилова¹✉, И.Д. Конопацкая^{1, 2}, С.Л. Лузянин³, М. Войцеховский⁴, А.Г. Блинов^{1, 5}

¹ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия

³ Институт биологии, экологии и природных ресурсов, Кемеровский государственный университет, Кемерово, Россия

⁴ Институт наук об окружающей среде, Ягеллонский университет, Краков, Польша

⁵ Институт систематики и экологии животных Сибирского отделения Российской академии наук, Новосибирск, Россия

В последние десятилетия наблюдается резкое снижение численности популяций медоносных пчел и шмелей на территории большинства стран мира. Вклад в снижение численности данных опылителей вносят различные паразитические организмы (бактерии, грибы, простейшие, нематоды, клещи и насекомые). Паразиты рода *Nosema* (Microsporidia: Nosematidae) и родов *Crithidia* и *Lotmaria* (Kinetoplastida: Trypanosomatidae) оказывают значительное негативное влияние на численность медоносных пчел и шмелей. В недавних исследованиях, проведенных с использованием ядерных ДНК-маркеров, были описаны новые виды и генетические варианты данных паразитов. Целью настоящей работы являлось установление уровня зараженности медоносных пчел и шмелей микроспоридиями (*Nosema* spp.) и трипаносоматидами (*Crithidia* spp. и *Lotmaria passim*), а также изучение генетической вариабельности этих паразитов на ранее не исследованной территории Индии. В работе проанализировано 119 образцов из четырех видов медоносных пчел и пяти видов шмелей. Уровни зараженности популяций пчел и шмелей паразитическими организмами на территории двух штатов (Джамму и Кашмир, Карнатака) были определены с помощью полимеразной цепной реакции с праймерами, специфичными к кластеру генов рибосомной РНК *Nosema*, *Crithidia* и *Lotmaria*. Совместное заражение популяций медоносных пчел и шмелей микроспоридиями и трипаносоматидами было зафиксировано на территории штата Джамму и Кашмир. В результате сравнительного анализа нуклеотидных последовательностей кластера генов рибосомной РНК установлено, что в популяциях медоносных пчел на территории Индии были представлены *N. bombi*, *N. ceranae* и *L. passim*. Популяции шмелей были поражены микроспоридией *Nosema D* и трипаносоматидами *Crithidia bombi* и *Crithidia*

expoeiki. В образцах медоносных пчел, собранных на территории штата Карнатака, паразиты родов *Crithidia* и *Lotmaria* не выявлены. В популяции медоносных пчел впервые выявлена микроспоридия *N. bombi*. Данные о распространении микроспоридий и трипаносоматид в популяциях медоносных пчел и шмелей по всему миру были обобщены и дополнены.

Ключевые слова: медоносные пчелы; шмели; заражение; гены рибосомной PHK; *Nosema* spp.; *Crithidia* spp.; *Lotmaria passim*.

HOW TO CITE THIS ARTICLE:

Vavilova V.Y., Konopatskaia I., Luzyanin S.L., Woyciechowski M., Blinov A.G. Parasites of the genus *Nosema*, *Crithidia* and *Lotmaria* in the honeybee and bumblebee populations: a case study in India. *Vavilovskii Zhurnal Genetiki i Selekcii* = Vavilov Journal of Genetics and Breeding. 2017;21(8):943-951. DOI 10.18699/VJ17.317

Microsporidia of the genus *Nosema* (Microsporidia: Nosematidae) and trypanosomatid parasites of the genera *Crithidia* and *Lotmaria* (Kinetoplastida: Trypanosomatidae) have a negative impact on the honeybees and bumblebees colonies fitness (Schmid-Hempel, 2001; Brown et al., 2003; Higes et al., 2008; Hornitzky, 2008; Yourth et al., 2008; Ravoet et al., 2013).

Nosema and *Tubulosema* species represent the obligate intracellular spore forming organisms that are related to the Fungi (Han, Weiss, 2017). Two microsporidium species, *Nosema ceranae* (Fries et al., 1996) and *Nosema apis* (Zander, 1909), are known to infect honeybees. *Nosema bombi* is another parasite belonging to the phylum Microsporidia which is widespread in the bumblebee populations (Fantham, Porter, 1914). Analysis of standard nuclear DNA markers of *N. bombi* in bumblebee colonies from USA, Russia, China and several European countries revealed new genetic variants of the parasite (Fries et al., 2001; Tay et al., 2005; Szentgyörgyi et al., 2011; Cordes et al., 2012; Li et al., 2012; Vavilova et al., 2015). Three new *Nosema* variants (A, B, and C) isolated from bumblebees in China were suggested to be genetic variants of *N. ceranae* (Li et al., 2012; Vavilova et al., 2015). All detected genetic variants of *Nosema* species did not receive the status of separate species.

Another microsporidium species *Tubulosema pampeana* (Microsporidia: Tubulinosematidae) was described for the first time in *Bombus araratus* individuals from Argentine (Plischuk et al., 2015). Currently, there are no cases of *T. pampeana* infections in other regions.

Crithidia and *Lotmaria* species are the protozoan flagellated trypanosomatid parasites of honeybees and bumblebees. For a long time *Crithidia mellifica* (Langridge, McGhee, 1967) was the only Trypanosomatidae species described for the *Apis mellifera* and it was poorly investigated. Recent identification of several DNA markers of the American honeybees parasites revealed their high genetic diversity (Cox-Foster et al., 2007; vanEngelsdorp et al., 2009; Runckel et al., 2011; Cornman et al., 2012). After detailed analysis of trypanosome stains the *C. mellifica* SF was redesignated as *Lotmaria passim*. Therefore at the moment two stains of *C. mellifica* (designated as ATCC 30254 and ATCC 30862) and two stains of *Lotmaria passim* (designated as BRL and SF) for honeybee populations are described (Schwarz et al., 2015). Trypanosomatid *Crithidia bombi* (Lipa, Triggiani, 1988) infecting bumblebees is highly researched (Schmid-Hempel, Reber Funk, 2004; Meeus et al., 2010; Schmid-Hempel, Tognazzo, 2010). Microsatellite data showed that several *C. bombi* genotypes circulates in bumblebee populations from Switzerland, Argentina and Chile (Schmid-Hempel, Reber Funk, 2004; Schmid-Hempel

et al., 2011, 2014). Recently two new *Crithidia* species, *Crithidia expoeiki* and *Crithidia mexicana*, have been identified in bumblebees from North America and Mexico, respectively (Schmid-Hempel, Tognazzo, 2010; Gallot-Lavallée et al., 2016).

Microsporidian and trypanosomatid parasites described above have a negative impact on the honeybees and bumblebees fitness. The parasites cause the rapid honeybees and bumblebees loss at both individual and colony levels (Schmid-Hempel, 2001; Brown et al., 2003; Higes et al., 2008; Hornitzky, 2008; Yourth et al., 2008; Ravoet et al., 2013).

Thus, investigation of these parasites in host populations from new geographical regions allows to characterize new genetic variants and describe its specific distribution. Despite the high importance of honeybees and bumblebees for the economy of India no studies of their parasites have been performed so far. In this study the diversity of *Nosema*, *Crithidia*, and *Lotmaria* parasites from honeybees and bumblebees in an unexplored regions of India, states Jammu and Kashmir and Karnataka were analyzed.

Materials and methods

Sample collection, DNA extraction, PCR amplification, and sequencing. 80 samples of honeybee species (*A. cerana*, *A. dorsata*, *A. florea* and *A. mellifera*) were collected in two Indian states (Jammu and Kashmir, Karnataka) in May and June, 2007, respectively (Fig. 1, Table). 39 bumblebee specimens of *B. asiaticus*, *B. lucorum*, *B. rufofasciatus*, *B. simillimus* and *B. trifasciatus* species were collected in Jammu and Kashmir state in May, 2007 (see Fig. 1, Table).

All samples were obtained by entomological sweep nets, identified to the species level in the field and preserved in 70 % ethanol. Total DNA were extracted from abdomens of the specimens fixed in ethanol using DNeasy Blood and Tissue Kit (QIAGEN) according to the manufacturer's protocol.

The standard sets of primers and PCR conditions were used for PCR amplification of small subunit ribosomal RNA (SSU rRNA), internal transcribed spacer (ITS) and large subunit ribosomal RNA (LSU rRNA) genes of *Nosema* spp. and *Tubulosema* spp. (Tay et al., 2005; Szentgyörgyi et al., 2011) and 18S rRNA genes of *Crithidia* spp. and *L. passim* (Meeus et al., 2010; Arismendi et al., 2016). Polymerase chain reactions (PCR) were performed in 20 µl volume containing 0.1 µg of genomic DNA, 10 mM Tris-HCl (pH 8.9), 1 mM (NH₄)₂SO₄, 1.5 mM MgCl₂, 200 µM of each of four dNTPs, 0.5 µM primers, and 2.5 units of Taq DNA polymerase. The PCR products were analyzed in 1.2 % agarose gel electrophoresis and extracted from gel with a QIAquick Gel Extraction Kit (QIAGEN). From 40 ng to 200 ng of each PCR product

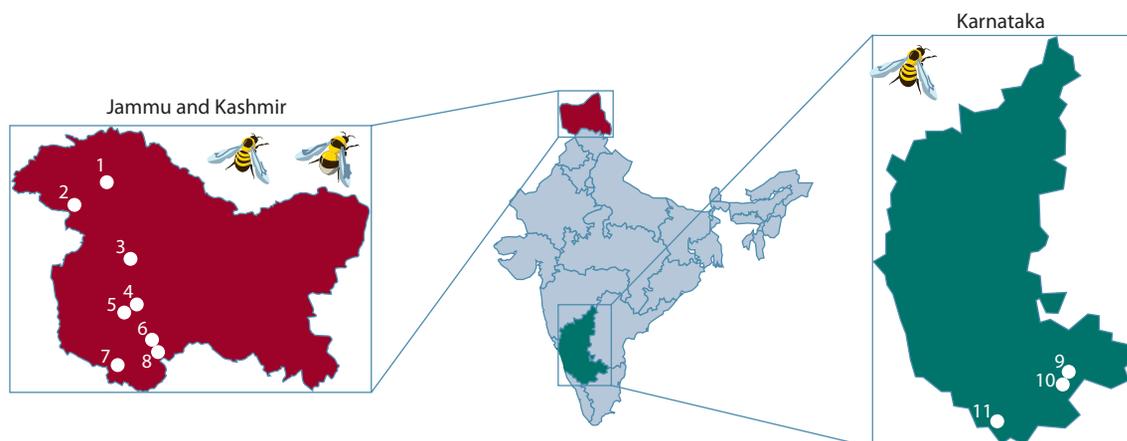


Fig. 1. Map of Indian states with collection sites designation. Collection sites are correlated with the Table.

Collection sites and list of honeybee and bumblebee species used in this study. The site areas are designated according to the Fig. 1

Indian state	Site	Coordinates of sites	Species of genus <i>Apis</i> and <i>Bombus</i>	Specimen number	Number of infected specimens (positive results of PCR)	
					<i>Nosema</i> spp./ <i>Tubulinosema</i> spp.	<i>Crithidia</i> spp./ <i>Lotmaria passim</i>
Jammu and Kashmir	1	34°08'50.2" N, 74°53'01.2" E	<i>Apis cerana</i> Fabricius, 1793 <i>Apis mellifera</i> Linnaeus, 1758	4 13	2/0 7/0	0/2 0/5
	2	34°01'57.2" N, 74°21'50.5" E	<i>A. cerana</i>	2	0/0	0/0
	3	33°32'09.2" N, 75°14'57.8" E	<i>A. cerana</i>	1	0/0	0/0
			<i>A. mellifera</i>	5	0/0	0/1
	4	33°07'38.1" N, 75°22'19.2" E	<i>A. mellifera</i>	9	3/0	0/1
	6	32°59'38.0" N, 75°42'10.9" E	<i>A. cerana</i>	7	0/0	0/0
	7	32°43'07.4" N, 74°51'59.1" E	<i>A. cerana</i>	2	0/0	0/0
<i>Apis dorsata</i> Fabricius, 1793			19	4/0	0/1	
<i>Apis florea</i> Fabricius, 1787			1	0/0	0/0	
Karnataka	9	13°01'09.0" N, 77°34'07.0" E	<i>A. cerana</i>	6	1/0	0/0
			<i>A. dorsata</i>	2	1/0	0/0
			<i>A. florea</i>	2	1/0	0/0
10	12°53'05.8" N, 77°28'21.5" E	<i>A. mellifera</i>	5	0/0	0/0	
11	12°01'07.4" N, 76°06'06.6" E	<i>A. dorsata</i>	2	1/0	0/0	
Total specimen number				80	20/0	0/10
Jammu and Kashmir	1	34°08'50.2" N, 74°53'01.2" E	<i>Bombus simillimus</i> Smith, 1852	8	0/0	2/0
	2	34°01'57.2" N, 74°21'50.5" E	<i>Bombus asiaticus</i> Morawitz, 1875	2	0/0	1/0
			<i>Bombus lucorum</i> Linnaeus, 1761	1	0/0	0/0
			<i>Bombus rufofasciatus</i> Smith, 1852	1	0/0	0/0
	4	33°07'38.4" N, 75°22'19.5" E	<i>Bombus trifasciatus</i> Smith, 1852	3	2/0	0/0
	5	33°05'09.9" N, 75°19'49.2" E	<i>B. trifasciatus</i>	5	2/0	2/0
8	32°57'11.1" N, 75°43'31.5" E	<i>B. asiaticus</i>	16	0/0	7/0	
		<i>B. trifasciatus</i>	3	0/0	0/0	
Total specimen number				39	4/0	12/0

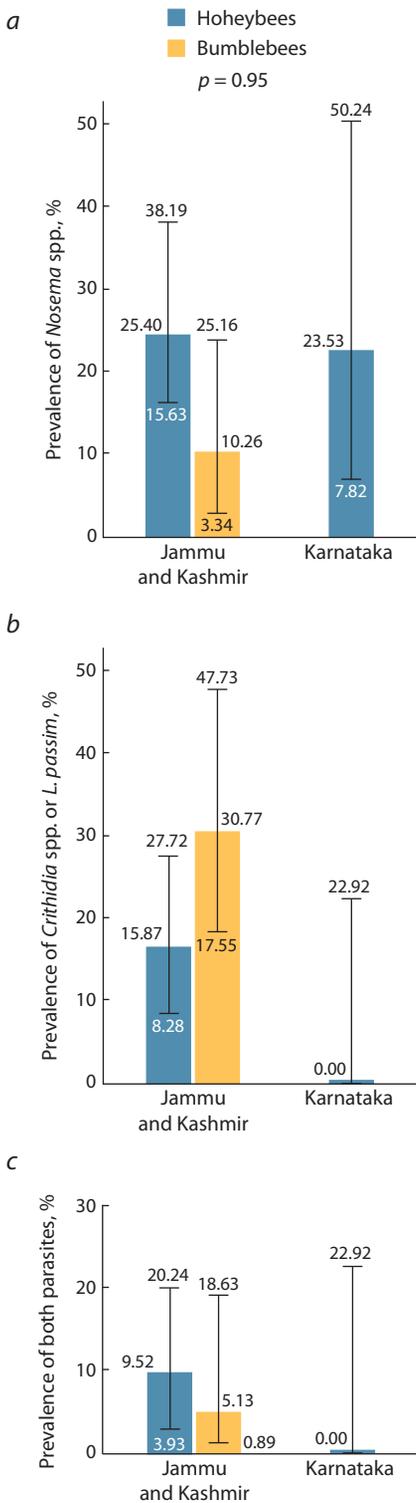


Fig. 2. Prevalence of *Nosema* spp. (a), *Crithidia* spp. or *L. passim* (b) infection and co-infection of all parasites (c) in honeybee and bumblebee populations from Jammu and Kashmir and Karnataka states.

Bars represent confidence intervals defined by chi-square test ($p = 0.95$) in STATISTICA. The studied parasites were characterized by low host-specificity; thus, their prevalence was calculated as percentage of the infected specimens to the total number of honeybee or bumblebee specimens in each sampling location.

was used in a 10- μ l cycle sequencing reaction with the BigDye® Terminator Kit (Applied Biosystems) and analyzed on an ABI 3130XL Genetic Analyser (Applied Biosystems) at the SB RAS Genomics Core Facility (Novosibirsk, Russia, <http://sequest.niboch.nsc.ru>). The partial SSU rRNA sequences of *N. bombi* and *N. ceranae* were deposited in the European Nucleotide Archive (ENA) under accession numbers HG321408–HG321411 and LT548980–LT548999. The obtained SSU rRNA, ITS2, and LSU rRNA gene sequences of *Nosema* spp. were deposited to GenBank under accession numbers KF188752–KF188755. The partial sequences of 18S rRNA gene from *Crithidia* spp. and *L. passim* were placed in GenBank under accession numbers KX151671–KX151702.

Parasite prevalence. The prevalence of microsporidia and trypanosomatid parasites was calculated as percentage of the infected specimens to the total number of honeybee or bumblebee specimens in each sampling sites. Confidence intervals for parasite prevalence were defined using STATISTICA by chi-square test.

Comparative and phylogenetic analyses. Logical search was performed to identify rRNA genes sequences from insects parasites of the phylum Microsporidia and family Trypanosomatidae available from the United States National Center for Biotechnology Information. Nucleotide sequences alignments were performed by ClustalW (Larkin et al., 2007) and improved by MUSCLE algorithm (Edgar, 2004) in Unipro UGENE software (Okonechnikov et al., 2012) (<http://ugene.unipro.ru>) for each parasite group. Phylogenetic analyses were performed using the Neighbor-Joining (NJ) method in MEGA 6.0 and Maximum Likelihood (ML) method in PhyML 3.0 (Guindon, Gascuel, 2003; Tamura et al., 2013). Statistical support for the NJ and ML trees was evaluated by bootstrapping, 100 replications for the ML method and 1,000 replications for the NJ method (Felsenstein, 1985).

Results and discussion

Prevalence of parasite infection in honeybee and bumblebee populations

We studied 80 honeybee and 39 bumblebee specimens collected in two Indian states. *Nosema* spp. were detected by PCR amplification with primers specific to SSU rRNA sequence. *Nosema* spp. were discovered in 20 honeybee specimens from all investigated honeybee species and 4 bumblebee specimens of *B. trifasciatus*. The prevalence of *Nosema* spp. in honeybee populations was 25 and 24 % in Jammu and Kashmir, and Karnataka states, respectively (Fig. 2, a). In Jammu and Kashmir state the prevalence of *Nosema* spp. in bumblebee population was 10 %.

The trypanosomatid parasites were identified by PCR amplification with primers specific to 18S rRNA sequence in 10 honeybee and 12 bumblebee specimens. *A. cerana*, *A. dorsata*, *A. mellifera*, *B. asiaticus*, *B. simillimus* and *B. trifasciatus* species were infected. In Jammu and Kashmir state the prevalence of trypanosomatid parasites was 16 and 31 % in honeybee and bumblebee populations, respectively (Fig. 2, b). No infected honeybee specimens were found in Karnataka state.

Co-infection by *Nosema* spp. and one of the trypanosomatid parasites (*Crithidia* spp. or *L. passim*) was detected in 6 honeybee (*A. cerana* and *A. mellifera*) and 2 bumblebee (*B. trifasciatus*) specimens in Jammu and Kashmir state (Fig. 2, c).

Genetic diversity of *Nosema* spp. in honeybee and bumblebee populations

Comparative analyses of the SSU rRNA sequences of *Nosema* spp. Totally, we obtained 24 nucleotide sequences of *Nosema* spp. SSU rRNA gene from honeybees and bumblebees (see Table). The results of the comparative analysis showed that 13 out of 20 sequences from honeybee specimens were identical to *N. bombi* SSU rRNA sequences (KF002566, HG321391, KF188769, JN872234, and JN872233). *N. bombi* was described for bumblebees only. The remaining 7 sequences were identical to SSU rRNA of *N. ceranae* (KF640602, JX205150).

Previously *Nosema* species were considered to be the host specific parasites. *N. apis* infected only the European honeybee *A. mellifera*, while *N. ceranae* was a specific parasite for the Asian honeybee *A. cerana* (Smith, 2012). In the recent years *N. ceranae* was identified in *A. mellifera* and some bumblebee species in the different parts of the world. Moreover, some stains of *N. ceranae* are predicted to replace *N. apis* in populations of *A. mellifera* honeybees (Chen, Huang, 2010;

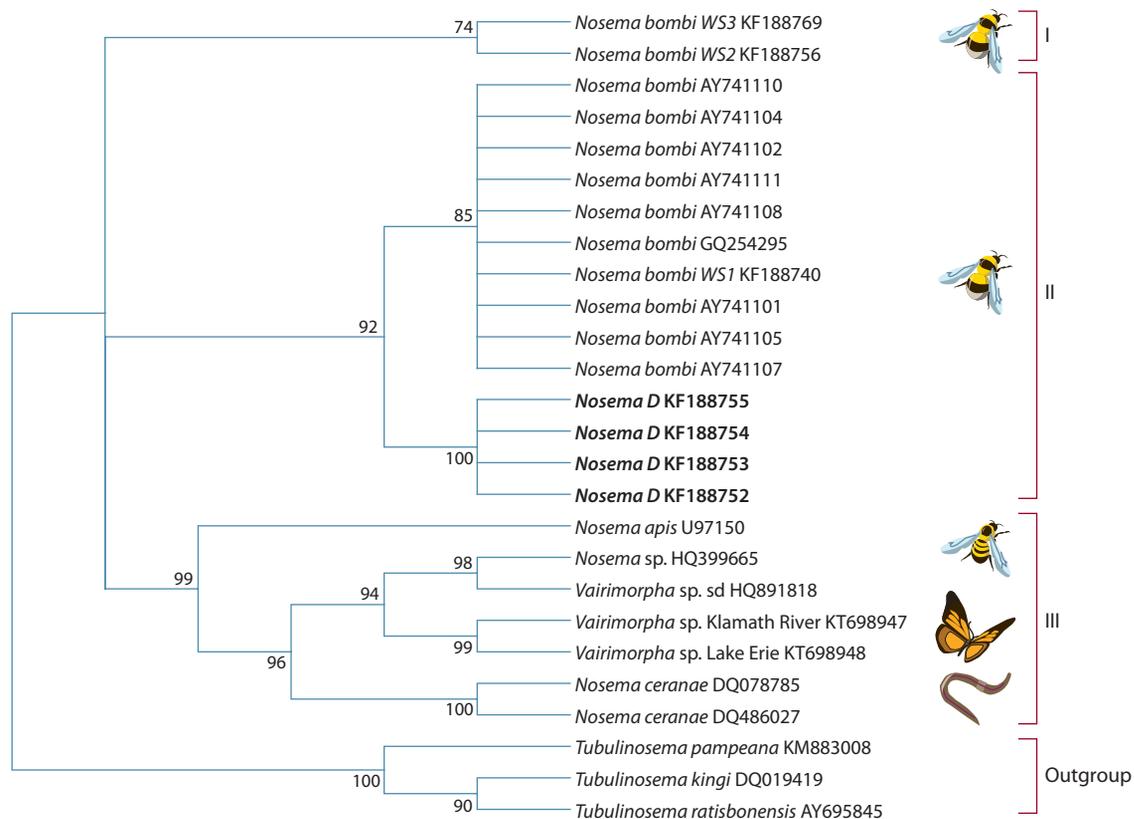


Fig. 3. Neighbor-joining (NJ) phylogenetic tree was based on SSU rRNA, ITS2, and the partial LSU rRNA nucleotide sequences of several *N. bombi* genetic variants, *Nosema* spp. and *Vairimorpha* spp. from various insects and annelids available from GenBank. *Tubulinoosema* spp. were used as outgroup.

Statistical support was evaluated by bootstrapping (1000 replications). The nodes with bootstrap values over 50 % are indicated.

Martín-Hernández et al., 2012). In this study a bumblebee parasite *N. bombi* was identified in honeybees from Jammu and Kashmir and Karnataka states, thus additionally refuting the assumption of the host specificity of *Nosema* species.

All *Nosema* spp. SSU rRNA sequences from infected bumblebee specimens were almost identical to *Nosema D* sequences (JN872219–JN872229) (Li et al., 2012). A distinguish is only a single nucleotide substitution. No other *Nosema* as well as *Tubulinoosema* species were found in bumblebees in this study (see Table).

Comparative and phylogenetic analyses of SSU rRNA, ITS2 and partial LSU rRNA sequences from *Nosema* spp. To expand information about *Nosema D* we obtained SSU rRNA, ITS2, and partial LSU rRNA sequences for four *B. trifasciatus* specimens. Sequences of *Vairimorpha* spp., *N. bombi*, *N. ceranae*, *N. apis*, and *Nosema* sp. from *Pieris rapae* from GenBank, as well as the obtained sequences were used for phylogenetic analysis. Sequences of several *Tubulinoosema* species were taken as an outgroup. Phylogenetic tree built by the NJ method is presented in Fig. 3.

The phylogenetic tree was divided on outgroup and three clusters (see Fig. 3). The outgroup was presented by sequences of *T. pampeana*, which were described as parasite of *B. araratus* from South America and two parasites of *Drosophila* spp. (*T. ratisbonensis* and *T. kingi*). The first cluster (I) consists of sequences of *N. bombi* WS2 and *N. bombi* WS3 that were

previously described in populations of bumblebee from West Siberia (Vavilova et al., 2015). The second cluster (II) includes two clades. The sequences of *N. bombi* previously identified from the Europe, USA and West Siberia formed the first clade (Tay et al., 2005; Sokolova et al., 2010; Szentgyörgyi et al., 2011). The second clade consists of the newly identified *Nosema D* sequences. The third cluster (III) is also split into two clades. Sequence of *N. apis*, obtained from *A. mellifera* apiary specimens in New Zealand (Gatehouse, Malone, 1998), is in the first clade of the third cluster. The second clade of this cluster consists of sequences of several *Vairimorpha* spp. from *Bombyx mori* and *Manayunkia speciosa* (Liu et al., 2012; Malakauskas et al., 2015); sequences of *N. ceranae* from Taiwan honeybees (Huang et al., 2007); and sequence of unspecified *Nosema* from *Pieris rapae* (Chen et al., 2012). Thus, the analysis of complete SSU rRNA, ITS2 and partial LSU rRNA gene sequences confirmed that *Nosema D* is a genetic variant of *N. bombi* and distributed in the bumblebee populations at least in China and India.

Genetic diversity of *Crithidia* spp. and *Lotmaria passim* in honeybee and bumblebee populations

Comparative analyses of the 18S rRNA sequences of *Crithidia* spp. and *Lotmaria passim*. We obtained 10 and 12 nucleotide sequences of 18S rRNA gene of trypanosomatid parasites in honeybees and bumblebees, respectively

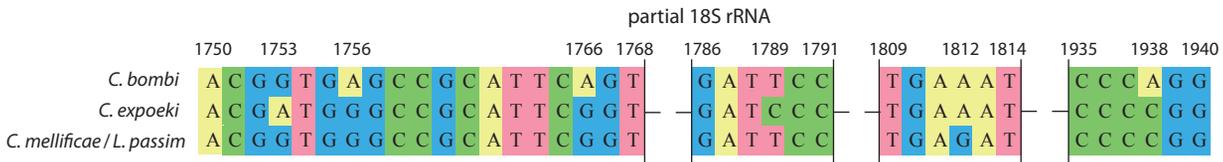


Fig. 4. Alignment of partial 18S rRNA gene sequences of *C. bombi*, *C. expoeki* and *C. mellifcae/L. passim* specified from honeybee and bumblebee specimens in this study.

Nucleotide positions are indicated according to *Crithidia fasciculata* sequence (Y00055) of full-length rRNA gene cluster.

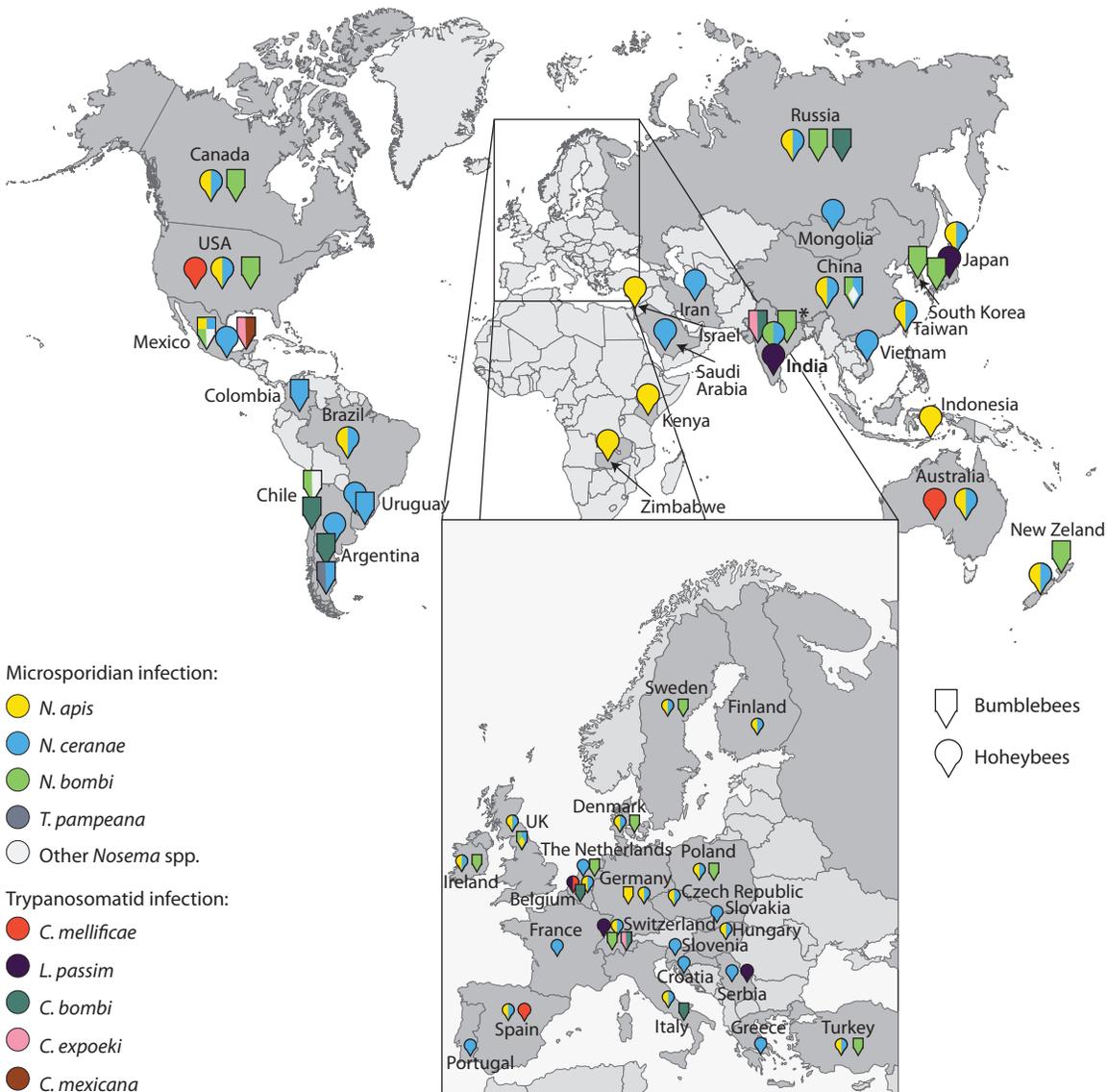


Fig. 5. World map of microsporidian and trypanosomatid distributions across the honeybee and bumblebee populations.

* *Nosema D*, which was determined in bumblebee populations from China (Li et al., 2012) and from India (the present study).

(see Table). The results of comparative analysis showed that nine of infected bumblebee specimens refer to *C. expoeki* (KM980187) and three others refer to *C. bombi* (FN546181, KM980184, KM980185). The distinguish between *C. bombi* and *C. expoeki* sequences amounted five nucleotide substitutions (Fig. 4).

All ten 18S rRNA gene sequences were identical and they could belong to either *C. mellifcae* or *L. passim* parasites. Sequences of *C. mellifcae/L. passim* differ in 4 and 3 nucleotide substitutions from *C. bombi* and *C. expoeki*, respectively (see Fig. 4). Using the primers for 18S rRNA specific to *L. passim* and to *C. mellifcae* (Arismendi et al., 2016) on the next

step, we proved that all the obtained sequences belonged to *L. passim* (KJ713378, KM980188, KT252553, KX953206).

Summarizing the data about microsporidian and trypanosomatid parasites in honeybee and bumblebee population from India, we identified that two *A. ceranae* and one *A. mellifera* specimens were co-infected by *N. ceranae* and *L. passim*; three specimens of *A. mellifera* were infected by both *N. bombi* and *L. passim*. Co-infection by *Nosema D* and *C. expoeki* in bumblebee populations was established in two *B. trifasciatus* specimens. No cases of *Nosema D* and *C. bombi* co-infection were found in this study.

Co-infection by *N. ceranae* and *L. passim* was also previously established in honeybee samples from Switzerland (Tritschler et al., 2017). Infection of both *N. ceranae* and *C. mellificaе* parasites was described for honeybees from Belgian apiaries (Ravoet et al., 2013) (see Fig. 4). Nevertheless, sequences of *C. mellificaе* 18S rRNA, identified by Ravoet et al. (2013), were identical for *C. mellificaе* and *L. passim*. Thus, these data should be clarified. Gallot-Lavallée et al. (2016) investigated co-infection by *Nosema* spp. and *Crithidia* spp. in the bumblebee populations from Mexico and established the cases of shared parasite infection (Fig. 5). However, species of *Nosema* and *Crithidia* genera found in infected bumblebee samples were not specified. Our data about co-infection of honeybee and bumblebee specimens by microsporidia and trypanosomatid parasites coincide with previously described studies. For the first time *N. bombi/L. passim* and *N. bombi/C. expoeki* co-infection were detected.

Geographic distribution of *Nosema* spp., *Crithidia* spp. and *L. passim* in honeybee and bumblebee populations

The results of this study supplement the knowledge of the distribution of microsporidian parasites among the honeybee and bumblebee populations all over the world (see Fig. 5).

Apian parasites of genus *Nosema* (*N. apis* and *N. ceranae*) are widely distributed in honeybee populations. Joint presence of these parasites was described in numerous studies (Table S1 in the Supplementary material)¹. Nevertheless, there are cases in the several countries such as Indonesia, Israel, Kenya and Zimbabwe of honeybee infections by *N. apis* only (see Table S1). Presence of *N. ceranae* only was established in the honeybee's population from countries of Latin America (except Brazil), several European countries, Iran, Mongolia, Saudi Arabia and Vietnam (see Table S1). In this study, we discovered that honeybee populations were infected by *N. ceranae*. There were no cases of *N. apis* presence. Presence of *N. bombi* parasite in honeybee specimens was detected for the first time (see Fig. 5).

N. bombi is widespread in the natural and commercial bumblebee populations of North and South America, Eurasia and New Zealand (Gallot-Lavallée et al., 2016; Brown, 2017). Several *N. bombi* genetic variants (*WS1*, *WS2* and *WS3*) were described in Siberian bumblebee populations (Vavilova et al., 2015). Four new *Nosema* variants (A, B, C and D) were isolated from bumblebees in China (Li et al., 2012). The microsporidian parasite, *T. pampeana*, was described in bumblebee populations from Argentina (Plischuk et al., 2015). In the recent decades the cases of bumblebees infection by apian

parasites *N. ceranae* (Argentina, China, Colombia, Mexico, UK and Uruguay), *N. apis* (Mexico and UK) and other *Nosema* species (Chile, China and Mexico) have been described (Gallot-Lavallée et al., 2016; Brown, 2017). We established the presence of *Nosema D* in bumblebee population from Jammu and Kashmir state (India). *Nosema D* was previously described by (Li et al., 2012) (see Fig. 5).

Presence of two trypanosomatid parasites, *C. mellificaе* and *L. passim*, was indicated in honeybee populations globally. *C. mellificaе* was found in honeybee specimens from Australia, Belgium, USA and Spain (Table S2). *L. passim* infections were described for honeybees from Belgium, Japan, Serbia and Switzerland (see Table S2). In this study distribution of *L. passim* in Indian honeybee populations were established (see Fig. 5). There were no cases of honeybee infection by *C. mellificaе*.

The cases of trypanosomatid infections were determined in commercial and native populations of bumblebees on the territories of North and South America and Eurasia. *C. bombi* is the most common trypanosomatid parasite that infects bumblebees from Argentina, Belgium, Chile, Germany, Italy, Russia, Switzerland and UK (Table S3). The second species *C. expoeki* is presented in Mexican and Swiss bumblebee populations (Schmid-Hempel, Tognazzo, 2010; Gallot-Lavallée et al., 2016). Bumblebee infection by *C. mexicana* was indicated in Mexico (Gallot-Lavallée et al., 2016). Both *C. bombi* and *C. expoeki* are distributed among bumblebee populations from India (see Fig. 5).

Thus, in this study the prevalence of *Nosema*, *Crithidia* and *Lotmaria* parasites in honeybee and bumblebee populations of Jammu and Kashmir and Karnataka states were identified. In addition, co-infection by Microsporidia and Trypanosomatidae parasites was identified in several honeybee and bumblebee specimens from Jammu and Kashmir state. Honeybee and bumblebee specimens from India studied were infected by several microsporidian parasites (*N. bombi*, *N. ceranae* and *Nosema D*). Trypanosomatid parasites of *C. bombi*, *C. expoeki* and *L. passim* species were detected in honeybee and bumblebee populations. Moreover, for the first time *N. bombi* infection was detected in the honeybee population. Thus, further investigations are required to determine distribution of microsporidia and trypanosomatid parasites among the honeybee and bumblebee populations all over the World.

Acknowledgments

The study was supported by the project 0324-2016-0008 from the Russian State Budget.

Conflict of interest

The authors declare no conflict of interest.

References

- Arismendi N., Bruna A., Zapata N., Vargas M. PCR-specific detection of recently described *Lotmaria passim* (Trypanosomatidae) in Chilean apiaries. *J. Invertebr. Pathol.* 2016;134:1-5.
- Brown M.J.F. Microsporidia: an emerging threat to bumblebees? *Trends Parasitol.* 2017;S1471-4922.
- Brown M.J.F., Schmid-Hempel R., Schmid-Hempel P. Strong context-dependent virulence in a host-parasite system: reconciling genetic evidence with theory. *J. Anim. Ecol.* 2003;72:994-1002.

¹ Supplementary materials are available in the online version of the paper: <http://www.bionet.nsc.ru/vogis/download/pict-2017-21/appx15.pdf>

- Chen D., Shen Z., Zhu F., Guan R., Hou J., Zhang J., Xu X., Tang X., Xu L. Phylogenetic characterization of a microsporidium (*Nosema* sp. MPr) isolated from the *Pieris rapae*. *Parasitol. Res.* 2012;111:263-269.
- Chen Y.P., Huang Z.Y. *Nosema ceranae*, a newly identified pathogen of *Apis mellifera* in the USA and Asia. *Apidologie.* 2010;41:364-374.
- Cordes N., Huang W.F., Strange J.P., Cameron S.A., Griswold T.L., Lozier J.D., Solter L.F. Interspecific geographic distribution and variation of the pathogens *Nosema bombi* and *Crithidia* species in United States bumble bee populations. *J. Invertebr. Pathol.* 2012;109:209-216.
- Corman R.S., Tarpy D.R., Chen Y., Jeffreys L., Lopez D., Pettis J.S., vanEngelsdorp D., Evans J.D. Pathogen webs in collapsing honey bee colonies. *PLoS ONE.* 2012;7:e43562.
- Cox-Foster D.L., Conlan S., Holmes E.C., Palacios G., Evans J.D., Moran N.A., Quan P.L., Briese T., Hornig M., Geiser D.M., Martinson V., vanEngelsdorp D., Kalkstein A.L., Drysdale A., Hui J., Zhai J., Cui L., Hutchison S.K., Simons J.F., Egholm M., Pettis J.S., Lipkin W.I. A metagenomic survey of microbes in honey bee colony collapse disorder. *Science.* 2007;318:283-287.
- Edgar R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32:1792-1797.
- Fantham H.B., Porter A. The morphology, biology and economic importance of *Nosema bombi*, N. sp., parasitic in various humble bees (*Bombus* spp.). *Ann. Trop. Med. Parasit.* 1914;8:623-638.
- Felsenstein J. Phylogenies and the comparative method. *Am. Nat.* 1985;125:1-15.
- Fries I., De Ruijter A., Paxton R.J., Da Silva A.J., Slemenda S.B., Pieniazek N.J. Molecular characterization of *Nosema bombi* (Microsporidia: Nosematidae) and a note on its sites of infection in *Bombus terrestris* (Hymenoptera: Apoidea). *J. Apic. Res.* 2001;40:91-96.
- Fries I., Feng F., da Silva A., Slemenda S.B., Pieniazek N.J. *Nosema ceranae* n. sp. (Microspora, Nosematidae), morphological and molecular characterization of a microsporidian parasite of the Asian honey bee *Apis cerana* (Hymenoptera, Apidae). *Eur. J. Protistol.* 1996;32:356-365.
- Gallot-Lavallée M., Schmid-Hempel R., Vandame R., Vergara C.H., Schmid-Hempel P. Large scale patterns of abundance and distribution of parasites in Mexican bumblebees. *J. Invertebr. Pathol.* 2016;133:73-82.
- Gatehouse H.S., Malone L.A. The ribosomal RNA gene region of *Nosema apis* (Microspora): DNA sequence for small and large subunit rRNA genes and evidence of a large tandem repeat unit size. *J. Invertebr. Pathol.* 1998;71:97-105.
- Guindon S., Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 2003;52:696-704.
- Han B., Weiss L.M. Microsporidia: obligate intracellular pathogens within the Fungal Kingdom. *Microbiol. Spectr.* 2017;5(2).
- Higes M., Martín-Hernández R., Botías C., Bailón E.G., González-Porto A.V., Barrios L., Del Nozal M.J., Bernal J.L., Jiménez J.J., Palencia P.G., Meana A. How natural infection by *Nosema ceranae* causes honeybee colony collapse. *Environ. Microbiol.* 2008;10:2659-2669.
- Hornitzky M. *Nosema* disease: Literature review and three year survey of beekeepers. Pt. 2. Rural Industries Research and Development Corporation, 2008.
- Huang W.F., Jiang J.H., Chen Y.W., Wang C.H. A *Nosema ceranae* isolate from the honeybee *Apis mellifera*. *Apidologie.* 2007;38:30-37.
- Langridge D.F., McGhee R.B. *Crithidia mellifica* n. sp.: An acidophilic trypanosomatid of honey bee *Apis mellifera*. *J. Protozool.* 1967;14:485-487.
- Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., Thompson J.D., Gibson T.J., Higgins D.G. Clustal W and Clustal X version 2.0. *Bioinformatics.* 2007;23:2947-2948.
- Li J., Chen W., Wu J., Peng W., An J., Schmid-Hempel P., Schmid-Hempel R. Diversity of *Nosema* associated with bumblebees (*Bombus* spp.) from China. *Int. J. Parasitol.* 2012;42:49-61.
- Lipa J.J., Triggiani O. *Crithidia bombi* sp. n., a new flagellated parasite of a bumble bee *Bombus terrestris* L. (Hymenoptera, Apidae). *Acta Protozool.* 1988;27:287-290.
- Liu H., Pan G., Li T., Huang W., Luo B., Zhou Z. Ultrastructure, chromosomal karyotype, and molecular phylogeny of a new isolate of microsporidian *Vairimorpha* sp. BM (Microsporidia, Nosematidae) from *Bombyx mori* in China. *Parasitol. Res.* 2012;110:205-210.
- Malakauskas D.M., Altman E.C., Malakauskas S.J., Thiem S.M., Schloesser D.W. Ribosomal DNA identification of *Nosema/Vairimorpha* in freshwater polychaete, *Manayunkia speciosa*, from Oregon/California and the Laurentian Great Lakes. *J. Invertebr. Pathol.* 2015;132:101-104.
- Martín-Hernández R., Botías C., Bailón E.G., Martínez-Salvador A., Prieto L., Meana A., Higes M. Microsporidia infecting *Apis mellifera*: Coexistence or competition. Is *Nosema ceranae* replacing *Nosema apis*? *Environ. Microbiol.* 2012;14:2127-2138.
- Meeus I., De Graaf D.C., Jans K., Smagge G. Multiplex PCR detection of slowly-evolving trypanosomatids and neogregarines in bumblebees using broad-range primers. *J. Appl. Microbiol.* 2010;109:107-115.
- Okonechnikov K., Golosova O., Fursov M., Varlamov A., Vaskin Y., Efreinov I., Grehov G., Kandrov O.G., Rasputin D., Syabro K., Tleukenov M. Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics.* 2012;28:1166-1167.
- Plischuk S., Sanscrainte N.D., Becnel J.J., Estep A.S., Lange C.E. *Tubulinosema pampeana* sp. n. (Microsporidia, Tubulinosematidae), a pathogen of the South American bumble bee *Bombus atratus*. *J. Invertebr. Pathol.* 2015;126:31-42.
- Ravoet J., Maharramov J., Meeus I., De Smet L., Wenseleers T., Smagge G., de Graaf D.C. Comprehensive bee pathogen screening in Belgium reveals *Crithidia mellifica* as a new contributory factor to winter mortality. *PLoS ONE.* 2013;8:e72443.
- Runckel C., Flenniken M.L., Engel J.C., Ruby J.G., Ganem D., Andino R., DeRisi J.L. Temporal analysis of the honey bee microbiome reveals four novel viruses and seasonal prevalence of known viruses, *Nosema*, and *Crithidia*. *PLoS ONE.* 2011;6:e20656.
- Schmid-Hempel P. On the evolutionary ecology of host-parasite interactions: Addressing the question with regard to bumblebees and their parasites. *Naturwissenschaften.* 2001;88:147-158.
- Schmid-Hempel P., Reber Funk C. The distribution of genotypes of the trypanosome parasite, *Crithidia bombi*, in populations of its host, *Bombus terrestris*. *Parasitology.* 2004;129:147-158.
- Schmid-Hempel R., Eckhardt M., Goulson D., Heinzmann D., Lange C., Plischuk S., Escudero L.R., Salathé R., Scriven J.J., Schmid-Hempel P. The invasion of southern South America by imported bumblebees and associated parasites. *J. Anim. Ecol.* 2014;83:823-837.
- Schmid-Hempel R., Salathe R., Tognazzo M., Schmid-Hempel P. Genetic exchange and emergence of novel strains in directly transmitted trypanosomatids. *Infect. Genet. Evol.* 2011;11:564-571.
- Schmid-Hempel R., Tognazzo M. Molecular divergence defines two distinct lineages of *Crithidia bombi* (Trypanosomatidae), parasites of bumblebees. *J. Eukaryot. Microbiol.* 2010;57:337-345.
- Schwarz R.S., Bauchan G.R., Murphy C.A., Ravoet J., De Graaf D.C., Evans J.D. Characterization of two species of trypanosomatidae from the honey bee *Apis mellifera*: *Crithidia mellifica* Langridge and McGhee, and *Lotmaria passim* n. gen., n. sp. *J. Eukaryot. Microbiol.* 2015;62:567-583.
- Smith M.L. The honey bee parasite *Nosema ceranae*: transmissible via food exchange? *PLoS ONE.* 2012;7:e43319.
- Sokolova Y.Y., Sokolov I.M., Carlton C.E. Identification of *Nosema bombi* Fantham and Porter 1914 (Microsporidia) in *Bombus impatiens* and *Bombus sandersoni* from Great Smoky Mountains National Park (USA). *J. Invertebr. Pathol.* 2010;103:71-73.
- Szentgyörgyi H., Blinov A., Ereemeeva N., Luzyanin S., Grześ I.M., Woyciechowski M. Bumblebees (Bombidae) along pollution gradient – heavy metal accumulation, species diversity, and *Nosema bombi* infection level. *Polish J. Ecol.* 2011;59:599-610.

- Tamura K., Stecher G., Peterson D., Filipiński A., Kumar S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 2013;30:2725-2729.
- Tay W.T., O'Mahony E.M., Paxton R.J. Complete rRNA gene sequences reveal that the microsporidium *Nosema bombi* infects diverse bumblebee (*Bombus* spp.) hosts and contains multiple polymorphic sites. *J. Eukaryot. Microbiol.* 2005;52:505-513.
- Tritschler M., Retschnig G., Yañez O., Williams G.R., Neumann P. Host sharing by the honey bee parasites *Lotmaria passim* and *Nosema ceranae*. *Ecol. Evol.* 2017;7:1850-1857.
- vanEngelsdorp D., Evans J.D., Saegerman C., Mullin C., Haubruge E., Nguyen B.K., Frazier M., Frazier J., Cox-Foster D., Chen Y., Underwood R., Tarry D.R., Pettis J.S. Colony collapse disorder: A descriptive study. *PLoS ONE*. 2009;4:e6481.
- Vavilova V., Sormacheva I., Woyciechowski M., Eremeeva N., Fet V., Strachecka A., Bayborodin S.I., Blinov A. Distribution and diversity of *Nosema bombi* (Microsporidia: Nosematidae) in the natural populations of bumblebees (*Bombus* spp.) from West Siberia. *Parasitol. Res.* 2015;114:3373-3383.
- Yourth C.P., Brown M.J.F., Schmid-Hempel P. Effects of natal and novel *Crithidia bombi* (Trypanosomatidae) infections on *Bombus terrestris* hosts. *Insectes Soc.* 2008;55:86-90.
- Zander E. Tierische Parasiten als Krankheitserreger bei der Biene. *Münch. Bienenztg.* 1909;31:196-204.

Генетическая и биохимическая характеристика стафилококков, встречающихся в Новосибирске

Ю.Н. Козлова¹, Н.В. Фоменко², В.В. Морозова¹, И.В. Саранина¹, А.Ю. Тикун¹, Д.А. Ганичев³, А.Г. Самохин⁴,
В.В. Павлов⁴, О.М. Рожнова⁴, И.А. Бондарь⁵, Е.В. Зенкова⁵, В.В. Нимаев⁶, В.В. Климонтов⁶, Н.В. Тикун¹✉

¹ Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук, Новосибирск, Россия

² АО Вектор-Бест, Новосибирск, Россия

³ Негосударственное учреждение здравоохранения Дорожная клиническая больница на ст. Новосибирск-Главный ОАО «Российские железные дороги», Новосибирск, Россия

⁴ Новосибирский научно-исследовательский институт травматологии и ортопедии им. Я.Л. Цивьяна Министерства здравоохранения Российской Федерации, Новосибирск, Россия

⁵ Новосибирский медицинский университет Министерства здравоохранения Российской Федерации, Новосибирск, Россия

⁶ Научно-исследовательский институт клинической и экспериментальной лимфологии – филиал Федерального исследовательского центра Института цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

Стафилококки способны поражать практически любые органы и ткани организма человека, вызывая поверхностные и глубокие гнойные инфекции, поражения дыхательных и мочевыводящих путей, а также провоцировать пищевые отравления и интоксикации. В последние годы участились случаи заболеваний, возбудителями которых оказываются коагулазонегативные стафилококки. Такие виды, как *Staphylococcus epidermidis*, *Staphylococcus haemolyticus* и *Staphylococcus hominis* значительно чаще выявляются при внутрибольничных инфекциях. Особая опасность этих патогенов заключается в повышенной вирулентности и патогенности штаммов, а также в их частой устойчивости к различным антибиотикам. Особенно трудно поддаются лечению заболевания, вызванные метициллин-резистентными стафилококками. Правильная идентификация стафилококков и их чувствительности к антибиотикам важна для постановки клинического диагноза и назначения адекватной лекарственной терапии. В связи с этим востребованы методы быстрого и точного определения видов стафилококков для оценки их патогенных свойств и наличия метициллин-резистентности. Цель данного исследования – характеристика стафилококков, в том числе и по устойчивости к антибиотикам, изолированных в г. Новосибирске из образцов, полученных от людей, животных и из окружающей среды. Проанализирована коллекция из 100 штаммов стафилококков. Видовая идентификация стафилококков проведена по результатам секвенирования гена 16S рРНК. Выявлено 11 видов стафилококков. Среди штаммов, полученных от больных из стационаров, доминировали *Staphylococcus aureus* (79.1 %), *Staphylococcus epidermidis* составили около 12.5 %. Среди внебольничных штаммов *S. aureus* и *S. epidermidis* выделены примерно в равных соотношениях. Обнаружение коагулазоположительных видов проводили стандартным биохимическим методом и выявлением гена *coa* методом ПЦР в реальном времени. Для штаммов *S. aureus* показано 100 % совпадение результатов между наличием гена *coa* и присутствием коагулазной активности, что свидетельствует о том, что выявление гена *coa* может служить точным методом идентификации *S. aureus*. Получен высокий уровень совпадения (99 %) между наличием в клетках стафилококков гена *tesA* и фенотипической устойчивостью к оксациллину. Исследование стафилококков на присутствие гена *tesA* может рассматриваться как альтернатива фенотипическому методу обнаружения метициллин-резистентных штаммов стафилококков.

Ключевые слова: *Staphylococcus*; коагулазонегативные стафилококки; метициллин-резистентные стафилококки (MRS); *tesA*; *coa*; диагностика стафилококковой инфекции.

Genetic and biochemical characterization of staphylococci occurring in Novosibirsk, Russia

Y.N. Kozlova¹, N.V. Fomenko², V.V. Morozova¹,
I.V. Saranina¹, A.Yu. Tikunov¹, D.A. Ganichev³,
A.G. Samokhin⁴, V.V. Pavlov⁴, O.M. Rozhnova⁴,
I.A. Bondar⁵, E.V. Zenkova⁵, V.V. Nimaev⁶,
V.V. Klimontov⁶, N.V. Tikunov¹✉

¹ Institute of Chemical Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia

² Joint-Stock Company Vector-Best, Novosibirsk, Russia

³ Railway Clinical Hospital on the station Novosibirsk-Glavnyy, Novosibirsk, Russia

⁴ Novosibirsk Research Institute of Traumatology and Orthopaedics n. a. L.A. Tsivyan of Ministry of Health of the Russian Federation, Novosibirsk, Russia

⁵ Novosibirsk State Medical University of Ministry of Health of the Russian Federation, Novosibirsk, Russia

⁶ Scientific Institute of Clinical and Experimental Lymphology – Branch of Federal Scientific Center Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

Staphylococci are capable of penetrating many human tissues and organs, causing superficial and deep purulent infections, respiratory and urinary tract infections, food poisoning and intoxication. Last years, coagulase-negative staphylococci were the cause of infection in many cases. Infectious agents, namely *Staphylococcus epidermidis*, *Staphylococcus haemolyticus*, and *Staphylococcus hominis*, were detected more often as nosocomial infections. A particular danger of these infections is a high virulence and pathogenicity of bacterial strains and their resistance to various antibiotics. Methicillin-resistant staphylococci are especially difficult to treat. The correct identification of staphylococci and their sensitivity to antibiotics are important for clinical diagnosis and appointment of adequate drug therapy. Rapid and accurate identification of *Staphylococcus* species and detection of their sensitivity to antibiotics is quite important. The aim of this study was to study staphylococci isolated in Novosibirsk from human, animal and environmental samples. A collection of 100 staphylococcus strains was analyzed. *Staphylococcus* species were identified by sequencing the 16S rRNA gene. Eleven staphylococcus species were identified. Among the strains obtained from hospitalized patients, *Staphylococcus aureus* do-

minated (79.1 %), *Staphylococcus epidermidis* amounted to about 12.5 %. However, *S. aureus* and *S. epidermidis* strains were isolated in an approximately equal proportion from community-associated samples. Identification of coagulase positive strains was performed using a standard biochemical method and by real-time PCR of the *coa* gene. 100 % coincidence between the presence of the gene and coagulase activity for *S. aureus* strains was recorded, which suggests that detection of the *coa* gene can be used as a correct method for *S. aureus* identification. A high coincidence rate (99 %) was revealed between the phenotypic resistance to oxacillin and the presence of the staphylococcal *mecA* gene. The study of staphylococci for the presence of the *mecA* gene can be considered as an alternative to the phenotypic method for identification of methicillin-resistant strains of staphylococci.

Key words: *Staphylococcus*; coagulase-negative staphylococci; methicillin-resistant staphylococcus strains (MRS); *mecA*; *coa*; diagnostics of staphylococcal infection.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Козлова Ю.Н., Фоменко Н.В., Морозова В.В., Саранина И.В., Тикунов А.Ю., Ганичев Д.А., Самохин А.Г., Павлов В.В., Рожнова О.М., Бондарь И.А., Зенкова Е.В., Нимаев В.В., Климонтов В.В., Тикунова Н.В. Генетическая и биохимическая характеристика стафилококков, встречающихся в Новосибирске. Вавиловский журнал генетики и селекции. 2017;21(8):952-958. DOI 10.18699/VJ17.318

HOW TO CITE THIS ARTICLE:

Kozlova Y.N., Fomenko N.V., Morozova V.V., Saranina I.V., Tikunov A.Yu., Ganichev D.A., Samokhin A.G., Pavlov V.V., Rozhnova O.M., Bondar' I.A., Zenkova E.V., Nimaev V.V., Klimontov V.V., Tikunova N.V. Genetic and biochemical characterization of staphylococci occurring in Novosibirsk, Russia. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2017;21(8):952-958. DOI 10.18699/VJ17.318 (in Russian)

Стафилококки – грамположительные факультативно-анаэробные малоподвижные бактерии, не образующие спор. К настоящему времени описано более 40 видов, входящих в род *Staphylococcus* (Ghebremedhin et al., 2008). Большинство бактерий из этого рода в норме обитают на коже и слизистых человека и животных, а также являются частью почвенных микробных сообществ (Madigan, Martinko, 2005). В зависимости от способности к продукции коагулазы, отвечающей за превращение фибриногена в фибрин и вызывающей свертывание крови, стафилококки подразделяются на коагулазопозитивные и коагулазонегативные. К первым относятся пять видов стафилококков, из которых только *Staphylococcus aureus* играет важную роль в патологии человека. Коагулазопозитивные стафилококки – *Staphylococcus pseudintermedius* и *Staphylococcus intermedius* – возбудители самых распространенных и опасных стафилококковых инфекций в ветеринарии, при тесном контакте с животными способны также передаваться человеку (Bond, Loeffler, 2012). Большинство представителей рода *Staphylococcus* являются коагулазонегативными и считаются менее вирулентными (Götz et al., 2006).

Стафилококки способны поражать практически любые органы и ткани организма человека, вызывая поверхностные и глубокие гнойные абсцессы, инфекции дыхательных и мочевыводящих путей, гнойно-некротические процессы в послеоперационных ранах, а также провоцировать пищевые отравления и интоксикации. Считается, что более 85 % всех стафилококковых инфекций человека вызваны *S. aureus* (Карпов, Качанко, 2006; Sievert et al., 2013). В последние годы участились случаи гнойно-септических инфекций у людей, возбудителями которых являются коагулазонегативные стафилококки *Staphylococcus*

epidermidis, *Staphylococcus saprophyticus*, *Staphylococcus haemolyticus* и *Staphylococcus hominis*. Они значительно чаще стали регистрироваться при нозокомиальных инфекциях, особенно у больших в отделениях интенсивной терапии и хирургических отделениях. Какое-то время эти инфекции могут протекать в виде бессимптомной бактериемии, не переходя в классические манифестные формы гнойной инфекции, в то же время резко истощая компенсаторные механизмы больного и повышая риск неблагоприятных исходов (Черный и др., 2010). Особая опасность внутрибольничных стафилококковых инфекций заключается в повышенной вирулентности и патогенности таких штаммов, а также в их частой устойчивости к различным антибиотикам. Особенно трудно поддаются лечению заболевания, вызванные метициллин-резистентными стафилококками (MRS), обладающими устойчивостью к бета-лактамам – группе антибиотиков, включающей метициллин, оксациллин, цефалоспорины и пр. Правильная идентификация стафилококков важна для постановки клинического диагноза и назначения адекватной лекарственной терапии, однако принятые процедуры лабораторного тестирования клинических изолятов (исследование способности к ферментации маннита, коагулазной и гемолитической активностей и др.) в ряде случаев сопровождаются ошибками и нуждаются в перепроверке другими методами. В связи с этим востребована быстрая и точная идентификация видов стафилококков для оценки их патогенных свойств и наличия метициллин-резистентности.

Цель данного исследования – характеристика стафилококков, в том числе и по устойчивости к антибиотикам, изолированных в г. Новосибирске из образцов, полученных от людей, животных и из окружающей среды.

Table 1. Synthetic oligonucleotides for analysis of the *mecA* and *coa* genes

Gene	Oligonucleotide	Sequence
<i>mecA</i>	Fmec	5'-AGGGACTCGAAAACTTTACGAT-3'
	Rmec	5'-ATTAATGTATGTGCGATTGTATTGC-3'
	Zmec	5'-(HEX)-GCTCCAACATGAAGATGGCTATCGTGTACAATCGT-(BHQ1)-3'
<i>coa</i>	Fco	5'-ACTCAAGGAGAATCAAGTGATATTGA-3'
	Rco	5'-GAATCTGGTCTCGCTTCATATC-3'
	Zco	5'-(HEX)-TCGTGTATTACGGATACCTGTACCAGCATCTC(T-BHQ1)-3'

Материалы и методы

Бактериальные штаммы. Исследовали 100 штаммов бактерий рода *Staphylococcus*, выделенных из образцов, взятых от людей, животных и из окружающей среды. Забор материала производили стерильным ватным тампоном. Бактериальные клетки смывали с тампона стерильным 0.9 % раствором NaCl, затем 10-кратные разведения полученной клеточной суспензии высевали на маннит-солевой агар и агар CLED (Laboratorios Conda S.A., Испания). Чистые культуры бактерий получали путем трехкратного пересева; принадлежность к роду *Staphylococcus* определяли по культурально-морфологическим признакам. Чистые культуры стафилококков были депонированы в коллекции экстремофильных микроорганизмов и типовых культур (КЭМТК) Института химической биологии и фундаментальной медицины СО РАН (ИХБФМ СО РАН).

Определение вида бактерий. Видовую принадлежность стафилококков определяли путем секвенирования последовательностей гена 16S рРНК. В качестве матрицы для амплификации гена 16S рРНК использовали бактериальные лизаты. ПЦР проводили с универсальными олигонуклеотидов 16s-8-f-b 5'-AGRGTTTGATCCTGGCTCA-3' и 16s-1350-r-b 5'-ACGGGCGGTGTGTACAAG-3', описанных ранее (Wang, Qian, 2009). В результате ПЦР получали фрагменты ДНК длиной 1342 н.п. Секвенирование полученных ПЦР-фрагментов осуществляли в обоих направлениях с применением тех же праймеров и набора BigDye™ Terminator v.3.1 (Applied Biosystems, США) по протоколам производителя. Капиллярный электрофорез продуктов реакции проводили на анализаторе ABI 3500 Genetic Analyzer (Applied Biosystems, США). Нуклеотидные последовательности верифицировали с помощью программы Sequencher v.4.0.5. и анализировали с использованием базы данных GenBank Database. Для определения вида сопоставляли полученную последовательность с наиболее близкими последовательностями рода *Staphylococcus*, присутствующими в базе данных GenBank Database (при уровне сходства последовательностей не менее 99 %). Видовую идентификацию в ряде случаев подтверждали на биохимическом анализаторе Gene III OmniLog System (BioLog, США).

Анализ биохимических свойств. Биохимические свойства выделенных бактерий проверяли стандартными методами: а) выявлением гемолитической активности на агаризованной среде с добавлением до 10 % крови кролика; б) выявлением наличия коагулазы пробирочным методом с плазмой кролика (Лебедева, 1973). Чувствительность к антибиотикам (бензилпенициллин, окса-

циллин, гентамицин, эритромицин, ципрофлоксацин, левофлоксацин, клиндамицин, ванкомицин) выделенных штаммов *Staphylococcus* spp. определяли диско-диффузионным методом, согласно МУК 4.2.1890-04 «Определение чувствительности к антибактериальным препаратам», с использованием агара Мюллера–Хинтона. Штаммы MRS идентифицировали на основании чувствительности к оксациллину.

Генетический анализ. Наличие генов *coa* и *mecA* в геномах исследуемых штаммов стафилококков подтверждали методом ПЦР в реальном времени (ПЦР-рв). Для этого из 100 мкл бактериальной суспензии выделяли ДНК с использованием набора реагентов «РеалБест ДНК-экспресс» (АО «Вектор-Бест», Россия). В качестве положительного контроля ПЦР использовали ДНК внутреннего контрольного образца, входящего в состав набора для выделения. ПЦР-рв проводили с использованием соответствующих праймеров и зондов, представленных в табл. 1, на амплификаторе с детекцией в режиме реального времени CFX96 Touch™ (Bio-Rad, США). Количественную оценку ДНК осуществляли по калибровочным кривым, построенным по результатам анализа серии разведений стандартного образца. Стандартным образцом являлась очищенная ДНК *S. aureus*, штамм КЭМТК 707, концентрация которой определена спектрофотометрически. В качестве контрольных штаммов для фенотипических тестов и генотипирования были взяты чувствительные к метициллину штаммы *S. aureus* ATCC 25923 и *S. aureus* ATCC 29213 и резистентный штамм *S. aureus* ATCC 43300. В ряде случаев полученные ПЦР-фрагменты секвенировали с применением тех же праймеров (см. табл. 1). Секвенирование и анализ последовательностей осуществляли аналогично описанному выше.

Результаты и обсуждение

Анализ видового разнообразия исследуемых штаммов стафилококков. В 2010–2014 гг. из различных образцов, взятых от людей, животных и из окружающей среды выделены 100 штаммов стафилококков, которые были депонированы в КЭМТК ИХБФМ СО РАН. Из 100 штаммов 85 изолированы от здоровых и больных людей, 48 штаммов – от пациентов, находившихся на стационарном лечении с торпидными трофическими язвами, парапротезной инфекцией и инфекцией мягких тканей, в том числе на фоне сахарного диабета (образцы из стационаров получены из Дорожной клинической больницы на станции Новосибирск-Главный ОАО «РЖД»; Новосибирского научно-исследовательского института травматологии и

Table 2. Identification of *Staphylococcus* strains

No	Species	Number of strains				
		overall	from indoor patients	from the community	from healthy persons and outpatients	from animals and the environment
1	<i>S. aureus</i>	56	38	18	15	3
2	<i>S. capitis</i>	2	0	2	2	0
3	<i>S. epidermidis</i>	24	6	18	13	5
4	<i>S. equorum</i>	1	0	1	1	0
5	<i>S. haemolyticus</i>	2	1	1	1	0
6	<i>S. hominis</i>	4	1	3	3	0
7	<i>S. intermedius</i>	1	0	1	0	1
8	<i>S. lentus</i>	1	0	1	0	1
9	<i>S. pseudintermedius</i>	4	0	4	0	4
10	<i>S. simulans</i>	2	1	1	1	0
11	<i>S. warneri</i>	3	1	2	1	1
Total		100	48	52	37	15

ортопедии им. Я.Л. Цивьяна Минздрава России; Научно-исследовательского института клинической и экспериментальной лимфологии; Государственной Новосибирской областной клинической больницы Минздрава РФ), 37 – от условно здоровых лиц и амбулаторных больных. От животных и из окружающей среды изолировано 15 штаммов (табл. 2).

По результатам секвенирования гена 16S рПНК определена видовая принадлежность выделенных штаммов стафилококков (см. табл. 2). Более половины штаммов составили штаммы *S. aureus*, около 1/3 – *S. epidermidis* (см. табл. 2). Видовое разнообразие внебольничных штаммов отличалось от такового у штаммов, выделенных из стационаров (см. табл. 2). Так, среди штаммов, полученных от больных из стационаров, доминировали *S. aureus* (79.1 %), *S. epidermidis* составили около 12.5 %, а *S. haemolyticus*, *S. simulans*, *S. hominis* и *S. warneri* встретились в единичных случаях. Среди внебольничных штаммов *S. aureus* и *S. epidermidis* выделены в примерно равных соотношениях, кроме того, у условно здоровых и амбулаторных больных спектр выявленных видов стафилококков шире, и, помимо вышперечисленных видов, обнаружены *S. capitis* и *S. equorum* (см. табл. 2). *S. lentus*, *S. intermedius* и *S. pseudintermedius* были изолированы только из образцов от животных (см. табл. 2).

Идентификация стафилококков с использованием биохимических маркеров. Наличие у стафилококков ряда биохимических свойств, в частности гемолитической и коагулазной активности, во многом обуславливает их патогенность и клиническую значимость. Исследуемые штаммы стафилококков протестированы на наличие некоторых биохимических свойств методами классической микробиологии. Анализ выявленных штаммов на наличие гемолитической активности показал, что 72 % проанализированных штаммов стафилококков обладали этой активностью, причем гемолитические свойства присутствовали как у штаммов *S. aureus*, так и у других видов стафилококков (табл. 3).

Считается, что одна из основных биохимических характеристик *S. aureus* – способность разлагать маннит до кислоты в маннит-солевом агаре, поэтому тест на ферментацию маннита используется для идентификации штаммов *S. aureus* (Нетрусов и др., 2005). В ходе исследования выявлено, что расщеплять маннит могут не только *S. aureus*, но и бактерии других видов стафилококков: *S. haemolyticus*, *S. warneri*, *S. capitis*, *S. lentus*, *S. equorum*, *S. hominis* и *S. intermedius* (см. табл. 3). Полученные результаты подтвердили, что тест на способность к ферментации маннита не позволяет точно выявлять *S. aureus* и требуется дополнительная идентификация.

Тесты на наличие коагулазы также используются для межвидовой диагностики стафилококков и выявления *S. aureus* (Нетрусов и др., 2005). Сравнение результатов выявления этого маркера у стафилококков фенотипическим и генетическим методами продемонстрировало 100 % совпадение только для штаммов *S. aureus*, включая тест-штаммы (см. табл. 3). У других коагулазопозитивных стафилококков (*S. intermedius* и *S. pseudintermedius*) соответствующий ПЦР-фрагмент не обнаружен (см. табл. 3). Таким образом, разработанная система детекции гена *coa* является высокоспецифичной в отношении штаммов *S. aureus* и может быть использована для идентификации бактерий этого вида.

Детекция резистентных к антибиотикам стафилококков, включая MRS. Исследование чувствительности штаммов стафилококков к антибиотикам показало, что большинство исследованных штаммов устойчиво к эритромицину (более 75 % штаммов) и бензилпенициллину (86 % резистентных штаммов среди *S. aureus*, 96 % – среди *S. epidermidis*). Вместе с тем все протестированные штаммы чувствительны к ванкомицину и более 65 % – к оксациллину и гентамицину (табл. 4).

Маркером чувствительности стафилококков к β-лактамам антибиотикам считается оксациллин. Если стафилококк обладает резистентностью к оксациллину, то и другие β-лактамы антибиотиками (метициллин, пеницил-

Table 3. Analysis of *Staphylococcus* strain features by biochemical and genetic methods

Species	Overall number of strains	Number of strains with activities			Detection of the <i>coa</i> gene ²
		hemolytic ¹	mannitol-metabolizing	coagulase	
<i>S. aureus</i>	56	44	56	56	56
<i>S. capitis</i>	2	1	2	0	0
<i>S. epidermidis</i>	24	14	0	0	0
<i>S. equorum</i>	1	1	1	0	0
<i>S. haemolyticus</i>	2	2	2	0	0
<i>S. hominis</i>	4	1	1	0	0
<i>S. intermedius</i>	1	1	1	1	0
<i>S. lentus</i>	1	1	1	0	0
<i>S. pseudintermedius</i>	4	3	0	3	0
<i>S. simulans</i>	2	2	0	0	0
<i>S. warneri</i>	3	2	3	0	0
Total	100	72	67	60	56

¹ The hemolytic activity was assayed after 20 h incubation of strains on agar medium with up to 10 % rabbit blood; ² The coagulase activity and the *coa* gene were also found in control *S. aureus* test strains ATCC 25923, ATCC 29213 and ATCC 43300.

Table 4. Antibiotic sensitivity in *Staphylococcus* strains

Antibiotic Class	Concentration	Number of strains, % of the overall number of strains								
		<i>S. aureus</i>			<i>S. epidermidis</i>			Other species		
		Susceptible	Intermediate	Resistant	Susceptible	Intermediate	Resistant	Susceptible	Intermediate	Resistant
Penicillins	Benzylpenicillin, 10 U	8 (14)	0	48 (86)	1 (4)	0	23 (96)	10 (50)	0	10 (50)
Semisynthetic penicillins	Oxacillin, 1 µg	45 (81)	0	11 (19)	17 (71)	1 (4)	6 (25)	19 (95)	0	1 (5)
Generation II aminoglycoside	Gentamicin, 10 µg	37 (67)	7 (12)	12 (21)	19 (79)	1 (4)	4 (17)	18 (90)	0	2 (10)
Macrolide	Erythromycin, 15 µg	8 (14)	35 (63)	13 (23)	5 (21)	9 (38)	10 (42)	3 (15)	10 (50)	7 (35)
Generation I fluoroquinolone	Ciprofloxacin, 5 µg	13 (22)	21 (39)	22 (39)	14 (58)	2 (8)	8 (34)	16 (80)	1 (5)	3 (15)
Generation III fluoroquinolone	Levofloxacin, 5 µg	39 (70)	7 (12)	10 (18)	14 (58)	7 (29)	3 (13)	14 (70)	2 (10)	4 (20)
Semisynthetic lincosamide	Clindamycin, 2 µg	28 (49)	12 (23)	16 (28)	13 (54)	2 (8)	9 (38)	14 (70)	4 (20)	2 (10)
Glycopeptide	Vancomycin, 30 µg	56 (100)	0	0	24 (100)	0	0	18 (90)	0	0

Note. Resistant to two or more antibiotic classes: *S. aureus*, 33 strains; *S. epidermidis*, 13 strains; and other *Staphylococcus* species, 6 strains.

лины, ингибитор-защищенные аминопенициллины, цефалоспорины всех поколений, карбапенемы) будут неэффективны (Сидоренко, Тишков, 2004; Карпов, Качанко, 2006; Черный и др., 2010). Чувствительность стафилококков к β-лактамам антибиотикам обусловлена способностью β-лактамов необратимо связываться с серин-содержащим активным центром так называемых пенициллин-связывающих белков (ПСБ), необходимых для построения клеточной стенки стафилококка, что приводит к гибели бактерии.

Различают, по крайней мере, пять различных ПСБ, из которых ПСБ2а, кодируемый геном *mecA*, обладает почти в 1000 раз меньшим аффинитетом к β-лактамам. Поэтому стафилококк, носитель гена *mecA*, способен выдержать воздействие этих антибиотиков (Сидоренко, Тишков, 2004). Ген *mecA* располагается в крупном мобильном генетическом элементе *SCCmec*. Метициллин-резистентные штаммы *S. aureus* (MRSA) часто обладают устойчивостью и ко многим другим антибиотикам, поскольку *SCCmec*

Table 5. Test of MRS strains for antibiotic resistance

No	Strain	Detection of the <i>mecA</i> gene by qPCR	Resistance detected by the disk-diffusion test		Strain origin
			to oxacillin ¹	to other antibiotics ²	
<i>S. aureus</i>					
1	KEMTK 675	+	R	PEN, GEN, ERY, CLI, CIP, LEV	Discharge from a postoperative fistula
2	KEMTK 707	+	R	PEN, ERY, CLI, CIP, LEV	Biopsy sample from the tissue necrosis area after surgery
3	KEMTK 1639	+	R	PEN, GEN, ERY, CLI, CIP, LEV	Discharge from a septic wound in diabetic foot infections
4	KEMTK 1695	+	R	PEN, GEN, CIP, LEV	»
5	KEMTK 1732	+	R	PEN, GEN, CIP	»
6	KEMTK 1733	+	R	PEN, CIP, LEV	Discharge from a postoperative fistula
7	KEMTK 1850	+	R	PEN, GEN, ERY, CLI, CIP, LEV	Discharge from a septic wound in diabetic foot infections
8	KEMTK 2077	+	R	PEN, GEN, ERY, CLI, CIP, LEV	»
9	KEMTK 2078	+	R	PEN, GEN, ERY, CLI, CIP, LEV	»
10	KEMTK 2105	+	R	PEN, GEN, ERY, CLI, CIP	»
11	KEMTK 2125	+	R	PEN, GEN, CIP	Biopsy sample and discharge from a surgical site
12	ATSS 43300	+	R	PEN, GEN, ERY, CLI	Test strain
13	ATSS 25923	-	S	S	»
14	ATSS29213	-	S	PEN	»
<i>S. epidermidis</i>					
1	KEMTK 1547	+	R	PEN, ERY, CIP, LEV	Discharge from a septic wound in diabetic foot infections
2	KEMTK 1548	+	R	PEN, ERY	»
3	KEMTK 1734	+	R	PEN, GEN	Discharge from a postoperative fistula
4	KEMTK 1827	+	R	PEN, GEN, ERY, CLI	Discharge from a septic wound in diabetic foot infections
5	KEMTK 1833	+	R	PEN, GEN, ERY, CLI, CIP	Discharge from a postoperative fistula
6	KEMTK 2034	+	I	PEN, CIP	Nasal smear in chronic sinusitis
7	KEMTK 2079	+	R	PEN, ERY	Discharge from a septic wound in diabetic foot infections
<i>S. haemolyticus</i>					
1	KEMTK 1700	+	R	PEN, GEN, ERY, CIP, LEV	Discharge from a septic wound in diabetic foot infections

Designations: ¹ R, resistant; S, susceptible; I, intermediate. ² PEN, benzylpenicillin; GEN, gentamicin; ERY, erythromycin; CLI, clindamycin; CIP, ciprofloxacin; LEV, levofloxacin. All strains except for KEMTK 2034 were obtained from indoor patients.

может содержать множественные локусы устойчивости (Сидоренко, Тишков, 2004), что существенно осложняет лечение. В связи с этим необходима идентификация MRSA и других MRS.

Фенотипическое тестирование штаммов стафилококков из Новосибирска на устойчивость к оксациллину выявило 18 штаммов MRS, 11 из которых – *S. aureus*, 6 – *S. epidermidis*, 1 – *S. haemolyticus* (табл. 5). Доля MRSA составила 19.6 % (11 из 56), метициллин-резистентных *S. epidermidis* (MRSE) – 25.0 % (6 из 24), а доля MRS среди остальных стафилококков – 5.0 % (1 из 20).

Штаммы MRS составили 38 % (18 из 48) из всех штаммов стафилококков, выделенных в условиях стационара, при этом половина штаммов MRS была устойчива к пяти и более антибиотикам (см. табл. 5). В то же время среди 52 внебольничных штаммов выявлен лишь один штамм *S. epidermidis*, КЭМТК 2034, обладающий промежуточной

устойчивостью к метициллину из образца от пациента, находящегося на амбулаторном лечении.

Одним из методов выявления MRS, включая MRSA, может быть генетический анализ на присутствие гена *mecA*. Все штаммы стафилококков были протестированы на присутствие гена *mecA* методом ПЦР-рв. Данные по встречаемости гена *mecA* у фенотипически выявленных штаммов MRS представлены в табл. 5. Оба вида анализа показали высокую степень соответствия (99 %). Только один из штаммов *S. epidermidis* КЭМТК 2034 обладал промежуточной устойчивостью к оксациллину, при этом выявление гена *mecA* подтвердило его принадлежность к MRSE. Среди штаммов, фенотипически чувствительных к оксациллину, ген *mecA* не выявлен.

Таким образом, детекция гена *mecA* методом ПЦР в реальном времени является специфичным методом, превосходит фенотипический метод по чувствительности и

быстроте анализа. Выявление гена *tecA* у стафилококков может рассматриваться как альтернатива фенотипическому выявлению штаммов MRS.

Распространение штаммов MRS в Российской Федерации существенно различается в зависимости от года, профиля учреждения и региона, при этом данные собраны преимущественно в европейской части страны (Сидоренко и др., 1998; Мурашкин и др., 2012; Бо и др., 2014). Опубликованы результаты двух многоцентровых исследований по эпидемиологии нозокомиальных штаммов *S. aureus* в Российской Федерации, включивших в том числе и штаммы, выделенные в стационарах Новосибирска и других городов Сибири (Дехнич и др., 2002, 2008). Результаты исследования 2002 г. показали, что штаммы MRSA составили 33.5 % от общего количества *S. aureus*, выявленных в стационарах России. При этом доля MRSA варьировала в различных отделениях стационаров от 0 до 89.5 % и в основном зависела от профиля отделения (Дехнич и др., 2002). В работе (Дехнич и др., 2008) исследована антибиотикорезистентность штаммов *S. aureus* в отделениях реанимации и интенсивной терапии: в среднем по России встречаемость MRSA составила 49.9 %, в Новосибирске – 72.1 %. В нашем исследовании доля MRSA составила 29 % от общего количества штаммов *S. aureus*, выделенных от госпитализированных пациентов, что приблизительно соответствует ранее полученным данным для стационаров России (Дехнич и др., 2002).

Заключение

Определена видовая принадлежность 100 штаммов стафилококков, выделенных в Новосибирске от больных и здоровых людей, от животных и из окружающей среды, и показано, что среди штаммов, полученных от больных из стационаров, доминировали *S. aureus* (79.1 %), *S. epidermidis* составили около 12.5 %, стафилококки других видов обнаружены в единичных случаях. Среди внебольничных штаммов спектр видов шире, а *S. aureus* и *S. epidermidis* встречаются примерно в равных соотношениях. Анализ наличия гена *coa* в отличие от биохимического теста на коагулазную активность позволил точно идентифицировать *S. aureus*. Выявлен высокий уровень совпадения между наличием в клетках стафилококков гена *tecA* и фенотипической устойчивостью к оксациллину. Исследование стафилококков на присутствие гена *tecA* можно использовать как альтернативу фенотипическому методу идентификации метициллин-резистентных штаммов стафилококков.

Acknowledgments

This study was supported by the Program of academic studies of the RAS Presidium, project 2014-155. Strains examined were obtained from the Collection of Extremophilic Microorganisms and Type Cultures of the Institute of Chemical Biology and Fundamental Medicine, Novosibirsk, supported by the Russian Federal Agency for Scientific Organizations, Program for Bioresource Collections, project 0309-2017-0008.

Conflict of interest

The authors declare no conflict of interest.

References

- Bo Lju, Tikhilov R.M., Shubnyakov I.I., Bozhkova S.A., Artyukh V.A., Denisov A.O. Analysis of the effectiveness of insulating operations in paraprosthesis infection. *Travmatologiya i ortopediya Rossii = Traumatology and Orthopedics of Russia*. 2014;2(72):22-29. (in Russian)
- Bond R., Loeffler A. What's happened to *Staphylococcus intermedius*? Taxonomic revision and emergence of multi-drug resistance. *J. Small Anim. Pract.* 2012;53(3): 147-154. DOI 10.1111/j.1748-5827.2011.01165.x.
- Cherniy V.I., Kolesnikov A.N., Kuznetsova I.V. Antibiotic therapy in emergency medicine. Donetsk: Zaslavsky Publ., 2010. (in Russian)
- Dekhnich A.V., Edelstain I.A., Narezkina A.D., Afinogenov G.E., Akhmetova L.I., Boronina L.G., Gugutcidze E.N., Gudkova L.V., Zdzitovetcki D.E., Ilyina V.N., Kretchikova O.I., Marusina N.E., Multih I.G., Pylaeva S.I., Smirnov I.V., Suborova T.N., Taraban V.K., Furletova N.M., Hasanova S.G., Schetinina E.V., Stratchounski L.S. Epidemiology of antimicrobial resistance of nosocomial strains of *Staphylococcus aureus* in Russia: Results of prospective study. *Klinicheskaya mikrobiologiya i antimikrobnaya khimioterapiya = Clinical Microbiology and Antimicrobial Chemotherapy*. 2002;4(4):325-337. (in Russian)
- Dekhnich A.V., Nikulin A.A., Ryabkova E.L., Krechikova O.I., Sukhorukova M.V., Kozlov R.S., and the ROSNET Study Group. Epidemiology of antimicrobial resistance of *S. aureus* isolated from ICU patients in Russia: Results of prospective multicenter study. *Klinicheskaya mikrobiologiya i antimikrobnaya khimioterapiya = Clinical Microbiology and Antimicrobial Chemotherapy*. 2008;10(4):333-345. (in Russian)
- Ghebremedhin B., Layer F., König W., König B. Genetic classification and distinguishing of *Staphylococcus* species based on different partial gap, 16S rRNA, hsp60, rpoB, sodA, and tuf gene sequences. *J. Clin. Microbiol.* 2008;46(3):1019-1025.
- Götz F., Bannerman T., Schleifer K-H. The Genera *Staphylococcus* and *Micrococcus*. The Prokaryotes. N. Y.: Springer, USA, 2006;5-75.
- Karpov I.A., Kachanko E.F. Community-acquired infections caused by methicillin-resistant *Staphylococcus aureus* (MRSA): approaches to antimicrobial therapy. *Meditinskije novosti = Medical News*. 2006; 10:28-32. (in Russian)
- Lebedeva M.N. Rukovodstvo k prakticheskim zanyatiyam po meditsinskoj mikrobiologii [A Guide to Practical Classes in Medical Microbiology]. Moscow: Meditsina Publ., 1973. (in Russian)
- Madigan M., Martinko J. Brock Biology of Microorganisms. N. Y.: Prentice Hall, 2005.
- Murashkin N.N., Gluzmin M.I., Skoblikov N.Je., Bakulev A.L., Materikin A.I., Gluzmina M.M., Khotko A.A. Role of MRSA strains in the pathogenesis of severe atopic dermatitis in childhood. The ways of remission achievement. *Vestnik dermatologii i venerologii = Bulletin of Dermatology and Venereology*. 2012;1:66-74. (in Russian)
- Netrusov A.I., Egorova M.A., Zakharchuk L.M., Kolotilova N.N. Praktikum po mikrobiologii [Practical Course in Microbiology]. Moscow: Akademiya Publ., 2005. (in Russian)
- Sidorenko S.V., Rezvan S.P., Grudinina S.A., Krotova L.A., Sterkhova G.V. Results of a multicenter study of the sensitivity of staphylococci to antibiotics in Moscow and St. Petersburg. *Antibiotiki i khimioterapiya = Antibiotics and Chemotherapy*. 1998;7:15-25. (in Russian)
- Sidorenko S.V., Tishkov V.I. Molecular basis of antibiotic resistance. *Uspekhi biologicheskoi khimii = Success of Biological Chemistry*. 2004;44:263-306. (in Russian)
- Sievert D.M., Ricks P., Edwards J.R., Schneider A., Patel J., Srinivasan A., Kallen A., Limbago B., Fridkin S. Antimicrobial-resistant pathogens associated with healthcare-associated infections: summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2009-2010. *Infect. Control Hosp. Epidemiol.* 2013;34(1):1-14. DOI 10.1086/668770.
- Wang Y., Qian P.Y. Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS One*. 2009;4(10):e7401. DOI 10.1371/journal.pone.0007401.

Филогенетический анализ микробного мата в горячем источнике Гарга (Байкальская рифтовая зона) и разнообразии природных пептидаз

Е.В. Лаврентьева^{1, 2}, А.А. Раднагуруева¹, Т.Г. Банзаракцаева¹, С.М. Базаров¹, Д.Д. Бархутова¹, И.Д. Ульзетуева³, М.К. Чернявский⁴, М.Р. Кабилов⁵, В.В. Хахинов²

¹ Институт общей и экспериментальной биологии Сибирского отделения Российской академии наук, Улан-Удэ, Россия

² Бурятский государственный университет, Улан-Удэ, Россия

³ Байкальский институт природопользования Сибирского отделения Российской академии наук, Улан-Удэ, Россия

⁴ Геологический институт Сибирского отделения Российской академии наук, Улан-Удэ, Россия

⁵ Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук, Новосибирск, Россия

В настоящее время в микробной экологии активно применяются методы метагеномного анализа, которые позволяют охарактеризовать таксономический состав и разнообразие микробных сообществ. Гидролитические бактерии, в частности протеолитики, в горячих источниках занимают нишу первичных деструкторов, благодаря способности секретировать ферменты, активные в широких диапазонах значений pH и температур. Целью данной работы было определение таксономического состава, структуры бактериального микробного мата и выявление пептидаз в термофильном микробном сообществе Гарга. Гидрохимический анализ воды показал высокое содержание сульфатов – 390 мг/дм³. В микроэлементном составе воды отмечены повышенные концентрации В, Rb, Li, Ba и Sr. Проведен анализ таксономического разнообразия микробного мата в горячем источнике Гарга, в температурной зоне 54 °С. Структура микробного мата представлена разнообразными филогенетическими группами мезофильных и термофильных бактерий с различными метаболическими и экологическими функциями. Наибольшую долю в сообществе составил филум Firmicutes (64 %). Анализ собранных метагеномных последовательностей микробного мата позволил впервые систематизировать и дать характеристику выявленных пептидаз в микробном мате горячего источника Гарга. Сравнение метагеномных последовательностей репрезентативных данных выявило доминирование ферментов класса сериновых пептидаз. Природные пептидазы в исследуемом микробном сообществе обеспечивают гидролиз биополимеров на первых этапах деструкции органического вещества и могут представлять биотехнологический интерес.

Ключевые слова: филогенетический анализ; микробное сообщество; метагеномный анализ; пептидазы; Байкальская рифтовая зона.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Лаврентьева Е.В., Раднагуруева А.А., Банзаракцаева Т.Г., Базаров С.М., Бархутова Д.Д., Ульзетуева И.Д., Чернявский М.К., Кабилов М.Р., Хахинов В.В. Филогенетический анализ микробного мата в горячем источнике Гарга (Байкальская рифтовая зона) и разнообразии природных пептидаз. Вавиловский журнал генетики и селекции. 2017;21(8):959-963. DOI 10.18699/VJ17.319

HOW TO CITE THIS ARTICLE:

Lavrentyeva E.V., Radnagurueva A.A., Banzaraktsaeva T.G., Bazarov S.M., Barkhutova D.D., Ulzetueva I.D., Chernyavsky M.K., Kabilov M.R., Khakhinov V.V. Phylogenetic analysis of the microbial mat in the hot spring Garga (Baikal rift zone) and the diversity of natural peptidases. Vavilovskii Zhurnal Genetiki i Selektii = Vavilov Journal of Genetics and Breeding. 2017;21(8):959-963. DOI 10.18699/VJ17.319 (in Russian)

Received 17.07.2017

Accepted for publication 05.10.2017

© AUTHORS, 2017

✉ e-mail: lena_l@mail.ru

Phylogenetic analysis of the microbial mat in the hot spring Garga (Baikal rift zone) and the diversity of natural peptidases

E.V. Lavrentyeva^{1, 2}, A.A. Radnagurueva¹, T.G. Banzaraktsaeva¹, S.M. Bazarov¹, D.D. Barkhutova¹, I.D. Ulzetueva³, M.K. Chernyavsky⁴, M.R. Kabilov⁵, V.V. Khakhinov²

¹ Institute of General and Experimental Biology SB RAS, Ulan-Ude, Russia

² Buryat State University, Ulan-Ude, Russia

³ Baikal Institute of Nature Management SB RAS, Ulan-Ude, Russia

⁴ Geological Institute SB RAS, Ulan-Ude, Russia

⁵ Institute of Chemical Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia

Hydrolytic bacteria (in particular, proteolytics) are the primary destructors in hot springs. The proteolytic bacteria are able to secrete enzymes that are active in wide ranges of pH and temperature. The aim of this work was to study the taxonomic composition, the structure of the bacterial microbial mat, and to study the distribution of peptidases in the thermophilic microbial Garga community. For the study, we sampled the microbial mat at a water temperature of 54.2 °C and a pH of 8.3. Hydrochemical analysis of water showed a high content of sulfates, 390 mg/dm³. The microelement composition of water showed that the Garga water had increased concentrations of B, Rb, Li, Ba, Sr. We analyzed the taxonomic diversity of the microbial community in the hot spring Garga at a temperature zone of 54 °C. The structure of the microbial mat is represented by various phylogenetic groups of mesophilic and thermophilic bacteria, with various metabolic and ecological functions. The dominant group in this community was the phylum Firmicutes (64 %). The analysis of the collected metagenomic sequences of the microbial community allowed the detected peptidases in the microbial community in the hot spring Garga to be for the first time systematized and characterized. Comparisons of metagenomic sequences of representative data showed a dominance of serine peptidase class enzymes. Natural peptidases in the investigated microbial community ensure the hydrolysis of biopolymers at the first stages of the destruction of organic matter and may have biotechnological relevance.

Key words: phylogenetic analysis; microbial community; metagenomic analysis; peptidases; Baikal rift zone.

Термофильные и алкалофильные гидролитические бактерии представляют одну из наиболее обширных и активно изучаемых групп микроорганизмов и являются источником новых ферментов и метаболитов для промышленности и медицины. Множество термо-/алкалофильных микроорганизмов осуществляет гидролиз полимерных субстратов – это амилолитики (Krishnan, Chandra, 1983; Gupta et al., 2003), целлюлозолитики (Bergquist et al., 1999; Uhl, Daniel, 1999; Andrade et al., 2001; Kozina et al., 2010; и др.) и протеолитики. Основная часть протеолитиков в качестве источника углерода и энергии использует пептиды благодаря наличию внеклеточных пептидаз, активных в широких диапазонах значений pH и температур (Ward et al., 2002). В целом в природных местообитаниях гидролитики занимают нишу первичных деструкторов, благодаря способности гетеротрофно расти на биополимерах различной природы (Кубланов, Подсокорская, 2011).

Использование современных методов молекулярной биологии позволило значительно расширить знания о разнообразии культивируемых и некультивируемых представителей различных местообитаний, в том числе горячих источников. Высокопроизводительное пиросеквенирование дает возможность охарактеризовать состав и разнообразие микробных сообществ и выявляет как доминирующие микроорганизмы, так и минорные компоненты (Бонч-Осмоловская, 2011).

В настоящей работе проведена оценка состава и структуры микробного мата с использованием комплексного подхода и изучено распространение пептидаз в термофильном микробном сообществе Гарга.

Материалы и методы

Высокотемпературный источник Гарга находится в долине р. Гарги, в отрогах Икатского хребта, окаймляющего котловину с восточной стороны ($54^{\circ}19'203''$ с. ш. и $110^{\circ}59'646''$ в. д., высота над уровнем моря 674 м). Участок выхода сложен водноледниковыми верхнечетвертичными отложениями, которые перекрывают палеозойские граниты и приурочены к Гаргинскому разлому, протягивающемуся в северо-восточном направлении на 30–40 км.

Микроэлементный состав воды определен в Геологическом институте СО РАН (г. Улан-Удэ) на масс-спектрометре с индуктивной связанной плазмой высокого разрешения ICP-MS Element XR (Thermo Scientific Fisher, Германия).

Изучение филогенетического разнообразия и поиск пептидаз в микробном сообществе проведены в ООО «Биоспарк» (г. Москва). Поиск пептидаз в микробном сообществе осуществляли в несколько этапов.

- Выделение тотальной ДНК из микробного сообщества. Непосредственно само выделение ДНК выполняли в три этапа: разрушение клеток, промывка и осаждение.
- Подготовка на основе полученных препаратов ДНК библиотеки для секвенирования на платформе Illumina MiSeq проводилась с использованием коммерческого набора реактивов New England Biolabs NEBNext® Ultra™ DNA Library Prep Kit for Illumina® в строгом соответствии с рекомендациями производителя (<https://www.neb.com/products/e7370-nebnext-ultra-dna-library-prep-kit-for-illumina>).

- Секвенирование полученной библиотеки на платформе Illumina MiSeq.
- Биоинформатический анализ полученных данных с целью выявления всего разнообразия пептидаз в микробном сообществе Гарга. Для сборки контигов использовали пакет программного обеспечения SPAdes v3.9.1 (<http://bioinf.spbau.ru/spades>). На основании обработанной аннотации получен перечень всех выявленных типов пептидаз.

Результаты и обсуждение

Физико-химические условия горячего источника Гарга.

Термальная сульфатно-натриевая вода изливается с дебитом 5 л/с в небольшой грот, расположенный на правом берегу р. Гарга, на высоте около 100 м от уреза воды (674 м над уровнем моря). Температура воды на выходе достигает 74°C . В газовом составе преобладает азот. По температурному режиму воды источника гипертермальные, с температурой выше 41°C . Вода источника, по классификации И.С. Ломоносова (1974), относится к горячинскому типу сульфатно-натриевых кремнистых терм. Отличительная особенность вод этого типа – абсолютное преобладание сульфатов натрия и небольшая минерализация, изменяющаяся от 0.5 до 1.0 г/л. Вода, стекая по ручью, образует травертин длиной до 50 м и шириной до 25 м, который считается одним из самых крупных в Байкальском регионе.

Высота уступов террас колеблется от 0.3 до 1.5 м, их образование зависит от водообильности источника. Состав травертинов практически полностью карбонатно-кальциевый, с высоким содержанием SiO_2 .

Пробы микробного мата отобраны при температуре воды 54.2°C и pH 8.3. Минерализация составила 0.74 г/дм^3 . Химический состав воды представлен в табл. 1.

Доминирующий катион в воде – натрий; среди анионов обнаружено максимальное количество ионов сульфата и гидрокарбоната. Характерной чертой для вод источника Гарга является высокое содержание радона, до 43 эман.

Анализ результатов микроэлементного состава воды показал, что в Гаргинских водах отмечена повышенная концентрация В, Rb, Li, Ba и Sr (табл. 2). Преобладание данных элементов в целом характерно для термальных источников, разгружающихся в пределах Икатского хребта. Наличие Sr, Li, Ba и некоторых других элементов в воде горячего источника связано с их высоким содержанием в гранитоидных породах Баргузинского комплекса.

Установлено, что содержание элементов и их соединений колеблется по разрезу травертинов, что может свидетельствовать о непостоянстве гидрохимического режима в период осадкообразования (Замана и др., 2011).

Филогенетический анализ микробного мата. Структура микробного мата горячего источника Гарга представлена различными филогенетическими группами мезофильных и термофильных бактерий с различными метаболическими и экологическими функциями (рисунок).

В микробном мате наибольшую долю в сообществе составил филум Firmicutes (30925 нуклеотидных последовательностей, 64 %). Подавляющее большинство последовательностей отнесено к порядку Bacillales и является представителями рода *Paenibacillus* (57 %). В про-

Table 1. Chemical composition of water in the Garga hot spring at the site of microbial mat sampling

Ion concentration	Na ⁺	K ⁺	Ca ²⁺	Mg ²⁺	CO ₃ ²⁻	HCO ₃ ⁻	NO ₃ ⁻	SO ₄ ²⁻	H ₂ SiO ₃	F ⁻	PO ₄ ³⁻	Cl ⁻
mg/L	312	11.3	23	0.08	6	109.8	7.4	390	89	10.5	2.02	51.4

Table 2. Concentrations of microelements in the Garga hot spring

Element	µg/L								
Be	0.6023	Hf	0.0011	La	0.0217	Mn	13.30	Gd	< DL
B	260.82	Tl	0.1946	Ce	0.0277	Fe	16.67	Tb	0.0005
Rb	118.59	Bi	0.0071	Pr	0.0026	Ni	0.4905	Dy	< DL
Y	0.0161	Th	0.0038	Nd	0.0095	Pb	0.4948	Ho	0.0005
Zr	0.0525	U	0.0056	Sm	0.0009	Al	33.70	Er	0.0013
Nb	0.0017	Li	1736	P	92.40	Ge	10.61	Tm	< DL
Mo	14.35	Co	< DL*	Sc	0.0015	As	0.0877	Yb	< DL
Ag	0.0080	Zn	6.144	Ti	0.8597	Se	< DL	Lu	< DL
Sn	0.1240	Ga	3.861	V	0.0405	Sr	3794	Ta	0.0025
Sb	0.0641	Cd	0.0287	Cr	0.2082	Eu	0.0010	Cu	1.306
Cs	42.85	Ba	46.25	W	80.82				

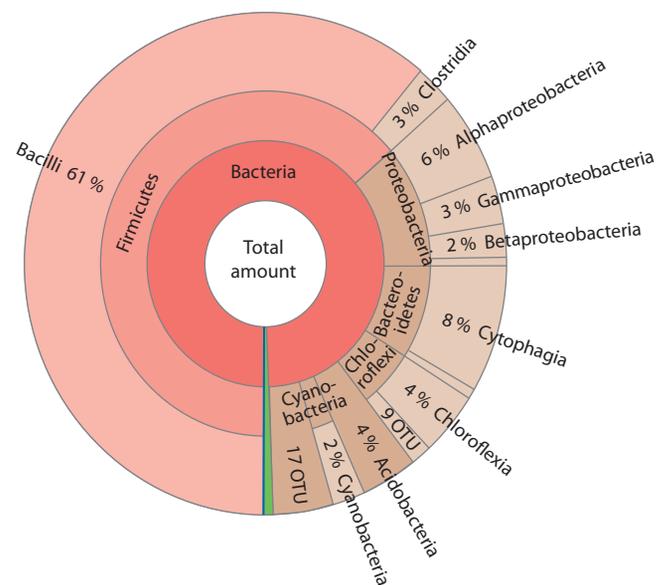
* DL, detection limit.

веденных ранее исследованиях нам удалось выделить алкало-термофильную протеолитическую бактерию в виде чистой культуры Га-35, которая была отнесена к *Paenibacillus dendritiformis*. Определение пептидазной активности показало, что штамм Га-35 обладает высокой активностью (3.7 ед/мг белка) на специфичном для субтилизин-подобных пептидаз субстрате GlpAALpNa (Раднагуруева и др., 2016). Известно, что представители рода *Paenibacillus* являются активными продуцентами гидролитических ферментов, в частности пептидаз, и наиболее перспективными объектами для использования в промышленной биотехнологии (Grady et al., 2016).

Филум Proteobacteria представлен четырьмя основными классами: Alpha (6 %), Beta (2 %), Gamma (3 %) и Delta (0.6 %). Для этого филума характерны большое разнообразие и высокая скорость аэробных и анаэробных метаболических путей (Madigan, 2000). Активность протеобактерий способствует деградации и быстрому повторному использованию органических веществ внутри микробного мата (Thibault et al., 2010).

В микробном мате горячего источника Гарга представители Gammaproteobacteria включают порядок Xanthomonadales (2 %), к которому относится семейство Xanthomonadaceae. Характерно, что многие бактерии этого семейства способны осуществлять разложение органического вещества. В классе Alphaproteobacteria доминирует порядок Rhodospirillales. Большая доля последовательностей в филуме Proteobacteria (18.0 %) принадлежит неклассифицированным в настоящее время видам бактерий.

Филум Vacteroidetes (8 %) представлен в сообществе микробного мата классом Cytophagia. Следует отметить, что все последовательности этого филума отнесены к некультивируемым бактериям. Филум Vacteroidetes включа-



Species composition in the microbial mat of the Garga hot spring. OTU, operational taxonomic unit, ≤ 1% in the community.

ет разнообразные гетеротрофные бактерии с чрезвычайно разнообразными метаболическими функциями. В работе (Portillo et al., 2009) показано, что филум Vacteroidetes составляет важную группу прокариот в горячих источниках Таиланда. Они потенциально важные потребители органического вещества, производимого цианобактериями и другими автотрофными организмами (Mackenzie et al., 2013).

Доля филума Chloroflexi в микробном сообществе составляет 5%. Термофильные бактерии Chloroflexus являются доминирующей группой в самом многочисленном

классе Chloroflexia (4 %). Метаболическая гибкость различных представителей филума Chloroflexi определяет эти бактерии как матообразующий компонент во многих исследованных горячих источниках и способствует в борьбе за питательные вещества и/или физическое пространство.

Около 4 % микроорганизмов определены как Acidobacteria. Полученные последовательности филума Acidobacteria отнесены к неклассифицированным бактериям на уровне класса.

Вклад Cyanobacteria в структуру микробного мата составил 2 % и представлен родами *Leptolyngbya*, *Chlorogloeopsis* и *Synechococcus*.

Минорными компонентами сообщества были бактерии филумов Armatimonadetes и Deinococcus-Thermus (0.9 %).

В проведенных ранее исследованиях в горячем источнике Гарга показано, что основа микробного мата в температурной зоне (52–57 °C) – цианобактерии. Культуральными и микроскопическими методами установлено доминирование цианобактерий родов *Phormidium*, *Oscillatoria* и *Synechococcus* (Бархутова и др., 2011; Намсараев и др., 2011). Характерно, что русло ручья в источнике периодически меняется, и, вероятно, в связи с этим произошла смена доминирующих видов в микробном мате в период наших исследований.

Обнаружено, что в составе микробного мата большое количество последовательностей имеет низкий уровень гомологии с известными культивируемыми представителями в мировой базе данных. Это, возможно, указывает на потенциально новые виды бактерий, не описанные ранее.

Метагеномный анализ пептидаз в микробном мате Гарга. Особенностью микробных сообществ термальных источников Байкальской рифтовой зоны являются его термофилия и алкалитолерантность, что предполагает термо- и pH стабильность соответствующих ферментов.

Ферменты, функционирующие в экстремальных условиях, определяют метаболические процессы и специфические биологические функции этих микроорганизмов в местах обитания. Среди секретируемых внеклеточных ферментов важная роль принадлежит пептидазам, принимающим активное участие в использовании микроорганизмами органических субстратов.

Результаты метагеномного анализа показали разнообразие пептидаз в природном образце микробного мата горячего источника Гарга. Основные данные биоинформатического анализа для изученного образца представлены ниже.

Кол-во контигов	467 457
Совокупная длина, п. н.	398 557 590
Медианная длина, п. н.	320
Средняя длина, п. н.	853
Максимальная длина, п. н.	994 249
N50	2330
N90	274
Кол-во аннотированных генов	332 830
Кодирующие	327 004
тРНК	3 974
тмРНК	55
Повторы	1 797
Прочие	–
С названием	121 875
С кодом фермента EC	82 125

Анализ полученных результатов позволил впервые систематизировать и дать характеристику установленных пептидаз в микробном мате горячего источника Гарга.

Сравнение метагеномных последовательностей репрезентативных данных показало доминирование ферментов класса сериновых пептидаз. Из литературных данных известно, что сериновые пептидазы обычно активны при нейтральных и щелочных значениях pH и имеют оптимум pH между 7 и 11 (Rao et al., 1998). На основании аминокислотной последовательности пептидазы делят на семейства, в настоящее время выделено более 20 семейств (Rao et al., 1998).

Метагеномный анализ выявил распространение пептидаз класса металлопептидаз в микробном мате Гарга. Обнаружено 12 последовательностей металлопептидаз различной номенклатуры (Germination protease precursor, ATP-dependent zinc metalloprotease и др.). В классе металлопептидазы выделяют 25 семейств, относящихся к разным систематическим группам. Металлопептидазы, как правило, синтезируются в неактивной форме и катализируются в присутствии ионов металлов (магния, марганца, кобальта, цинка). Оптимальные значения pH для металлопептидаз также находятся в диапазоне 7–9. Известны металлопептидазы, обладающие свойствами термостабильности (Oliveira et al., 2003).

Нами выявлены лишь единичные последовательности генов класса цистеиновых пептидаз в микробном сообществе Гарга. Возможно, это связано с тем, что цистеиновые пептидазы эффективны при нейтральных значениях pH; реже зона оптимума лежит в слабощелочной или слабощелочной среде (pH 4–9) в зависимости от природы гидролизующего белка (Oliveira et al., 2003). У изученного нами микробного сообщества диапазон развития находится в высокотемпературной зоне и щелочной области pH.

Заключение

Таким образом, впервые методом пиросеквенирования удалось подробно описать таксономический состав и количественное соотношение основных групп микроорганизмов в микробном мате горячего источника Гарга в температурной зоне 54 °C. Характерно, что изученное микробное сообщество представлено весомой долей бактерий, участвующих на разных этапах продукции и деструкции органического вещества в микробных матах. Большое количество филогенетически разнообразных, метаболически разнонаправленных групп свидетельствует о сбалансированном сложном сообществе, где каждая группа занимает свою экологическую нишу. Доминирование бактерий филума Firmicutes позволяет предположить, что в сообществе изученного микробного мата активно протекают процессы деструкции органического вещества.

Использование метагеномного метода позволило выявить распространение пептидаз в изученном микробном мате. Показано доминирование ферментов класса сериновых пептидаз. Изучение пептидаз микроорганизмов, функционирующих в экстремальных местах обитания, предоставляет уникальную возможность исследовать стратегии отдельных регуляторных процессов микробной клетки и границы функционирования микроорганизмов, так как известно, что усиленный синтез гидролитических

ферментов является одним из способов адаптации микроорганизмов к экстремальным условиям окружающей среды.

Acknowledgments

This study was supported by the Federal Agency for Scientific Organization, projects 037-2017-0003; Russian Federation Ministry of Education and Science 6.9754. 2017/BCh 15-04-01275; and the Russian Foundation for Basic Research, projects 15-04-01275; 15-44-04335-r_Siberia_a.

The authors are grateful to Cand. Sci. (Biol.) A. Yu. Merkel, Director General of the Biospark Company (Moscow, Russia), and Cand. Sci. (Geol.-Min.) V.B. Khubanov, Researcher, Head of the Group of Inductively Coupled Plasma Mass Spectrometry, Geological Institute of the Siberian Branch of the RAS (Ulan-Ude, Buryatia, Russia).

Conflict of interest

The authors declare no conflict of interest.

References

- Andrade C.M., Aguiar W.B., Antranikian G. Physiological aspects involved in production of xylanolytic enzymes by deep-sea hyperthermophilic archaeon *Pyrodicticum abyssi*. *Appl. Biochem. Biotechnol.* 2001;91-93:655-669.
- Barkhutova D.D., Tsyrenova D.D., Bryanskaya A.V., Danilova E.V., Zaitseva S.V., Namsaraev Z.B. Mikrobnye maty. Geokhimicheskaya deyatel'nost' mikroorganizmov gidroterm Baykal'skoy riftovoy zony [Microbial mats. Geochemical activity of microorganisms in thermal springs of the Baikal Rift zone]. Novosibirsk: Acad. Publ. House "Geo", 2011. (in Russian)
- Bergquist P.L., Gibbs M.D., Morris D.D., Te'o V.S.J., Saul D.J., Morgan H.W. Molecular diversity of thermophilic cellulolytic and hemicellulolytic bacteria. *FEMS Microbiol. Ecol.* 1999;28:99-110.
- Bonch-Osmolovskaya E.A. Thermophilic microorganisms: a general overview. *Trudy Instituta mikrobiologii im. S.N. Vinogradskogo. [Gal'chenko V.F. (Ed.) Proceeding of the Winogradsky Institute of Microbiology]*. Moscow: Nauka Publ., 2011;5-14. (in Russian)
- Grady E.N., MacDonald J., Liu L., Richman A., Yuan Z.-C. Current knowledge and perspectives of *Paenibacillus*: a review. *Microb. Cell Factories.* 2016;15:203. DOI 10.1186/s12934-016-0603-7.
- Gupta R., Beg Q., Lorenz P. Bacterial alkaline proteases: molecular approaches and industrial applications. *Appl. Microbiol. Biotechnol.* 2003;59(1):15-32.
- Kozina I.V., Kublanov I.V., Kolganova T.V., Chernyh N.V., Bonch-Osmolovskaya E.A. *Caldanaerobacter uzonensis* sp. nov., an anaerobic, thermophilic, heterotrophic bacterium isolated from a hot spring. *Int. J. Syst. Evol. Microbiol.* 2010;60:1372-1375.
- Krishnan T., Chandra A.K. Purification and characterization of α -amylase from *Bacillus licheniformis* CUMC305. *Appl. Environ. Microbiol.* 1983;46:430-437.
- Kublanov I.V., Podosokorska O.A. Thermophilic microorganisms decomposing biopolymers. *Trudy Instituta mikrobiologii im. S.N. Vinogradskogo. [Gal'chenko V.F. (Ed.) Proceeding of the Winogradsky Institute of Microbiology]*. Moscow: Nauka Publ., 2011;315-342. (in Russian)
- Lomonosov I.S. Geokhimiya i formirovanie sovremennykh gidroterm Baykalskoy riftovoy zony [Geochemistry and Formation of Modern Thermal Springs in the Baikal Rift Zone]. Novosibirsk: Nauka Publ., 1974. (in Russian)
- Mackenzie R., Pedrós-Alió C., Díez B. Bacterial composition of microbial mats in hot springs in Northern Patagonia: variations with seasons and temperature. *Extremophiles.* 2013;17:123-136. DOI 10.1007/s00792-012-0499-z.
- Madigan M.T. Bacterial habitats in extreme environments. *Journey to Diverse Microbial Worlds.* 2000; 2:61-72.
- Namsaraev Z.B., Zaitseva S.V., Dmitrieva O.M., Barkhutova D.D. Structure and functional activity of microbial mats of the Garga thermal spring (Barguzin valley). *Vestnik Buryatskogo gosudarstvennogo universiteta = Bulletin of the Buryat State University.* 2011;14a: 231-239. (in Russian)
- Oliveira A.S., Xavier-Filho J., Sales M.P. Cysteine proteinases and cystatins. *Brazil. Arch. Biol. Technol.* 2003;46(1):91-104.
- Portillo M.C., Sririn V., Kanoksilapatham W., Gonzalez J.M. Differential microbial communities in hot spring mats from Western Thailand. *Extremophiles.* 2009; 13(2):321-331. DOI 10.1007/s00792-008-0219-x.
- Radnagurueva A.A., Lavrentieva E.V., Budagaeva V.G., Barkhutova D.D., Namsaraev B.B., Dunaevsky Y.E. Organotrophic bacteria of the Baikal Rift Zone hot springs. *Microbiology (Mikrobiologiya).* 2016;85(3):367-378.
- Rao M.B., Tanksale A.M., Ghatge M.S., Deshpande V.V. Molecular and biotechnological aspects of microbial proteases. *Microbiol. Mol. Biol. Rev.* 1998;62(3):597-635.
- Thibault V., Lovejoy C., Jungblut A.D., Vincent W.F., Corbeil J. Metagenomic profiling of Arctic microbial mat communities as nutrient scavenging and recycling systems. *Limnol. Oceanogr.* 2010; 55(5):1901-1911. DOI 10.4319/lo.2010.55.5.190.
- Uhl A.M., Daniel R.M. The first description of an archaeal hemicellulase: the xylanase from *Thermococcus zilligii* AN1. *Extremophiles.* 1999;3:263-267.
- Ward D.E., Shockley K.R., Chang L.S., Levy R.D., Michel J.K., Connors S.B., Kelly R.M. Proteolysis in hyperthermophilic microorganisms. *Archaea.* 2002;1:63-74.
- Zamana L.V., Khakhinov V.V., Danilova E.V., Barkhutova D.D. Hydrochemistry of mineral waters. *Geokhimicheskaya deyatel'nost' mikroorganizmov gidroterm Baykalskoy riftovoy zony [Geochemical activity of microorganisms in thermal springs of the Baikal Rift zone]*. Novosibirsk: Acad. Publ. House "Geo", 2011;62-101. (in Russian)

Полиморфизм гена трегалазы (*TREH*) у коренного населения Сибири

Б.А. Малярчук , М.В. Деренко

Институт биологических проблем Севера Дальневосточного отделения Российской академии наук, Магадан, Россия

Дефицит некоторых углеводов в традиционной диете коренного населения Крайнего Севера способствовал высокой распространенности в популяциях неактивных вариантов генов, кодирующих, например, амилазу (ген *AMY2A*) и сахаразу-изомальтазу (ген *SI*). Одним из малоусваиваемых коренными жителями Крайнего Севера дисахаридов является трегалоза, которая содержится в водорослях, высших грибах, лишайниках и некоторых высших растениях. В настоящей работе исследован полиморфизм гена *TREH* в популяциях Сибири. Этот ген кодирует трегалазу – фермент, расщепляющий трегалозу. Анализ экзомного полиморфизма показал наличие в популяциях семи гаплотипов гена *TREH*. Три из них определяются вариантом rs2276064-A, который ассоциируется с самой низкой активностью трегалазы. Максимальная частота гаплотипов этой группы наблюдается в выборках коренного населения Северо-Восточной Азии (около 60%), в популяциях других регионов Сибири ее частота составляет 30–40%. Таким образом, высокая частота варианта rs2276064-A, ассоциирующегося с малоактивной трегалазой, объясняет, почему коренные северяне избегают пищи, содержащей трегалозу. Предполагается, что увеличению частоты этого варианта в популяциях Северо-Востока Азии мог способствовать дрейф генов, действующий в популяциях малой эффективной численности. Однако не исключено, что искусственно вызванный дефицит трегалозы в пище коренных народов Крайнего Севера (вследствие традиции отказа от грибов) также мог стать причиной увеличения частоты малоактивной трегалазы при условии, что эта традиция существует у жителей Крайнего Севера на протяжении многих поколений.

Ключевые слова: ген *TREH*; экзом; однонуклеотидные полиморфизмы; популяции человека; Сибирь; адаптивная эволюция.

Polymorphism of the trehalase gene (*TREH*) in native populations of Siberia

B.A. Malyarchuk , M.V. Derenko

Institute of Biological Problems of the North FEB RAS,
Magadan, Russia

Deficiency of some carbohydrates in the traditional diet of native populations of the Far North contributed to a high population prevalence of inactive variants of genes encoding, for example, amylase (*AMY2A* gene) and sucrase-isomaltase (*SI* gene). Trehalose, which is found in algae, higher fungi, lichens and some higher plants, is another of the disaccharides, poorly digested by native people of the Far North. Here, in native populations of Siberia, we studied the polymorphism of the *TREH* gene, which encodes trehalase, an enzyme that cleaves trehalose. The analysis of exome polymorphism showed the presence of seven haplotypes of the *TREH* gene in the population. Three of them are determined by the variant rs2276064-A, which is associated with the lowest activity of trehalase. The maximum frequency of this group of haplotypes is observed in the samples of native populations of Northeast Asia (about 60%), in the remaining samples of the Siberian populations, its frequency is 30–40%. Thus, the high frequency of variant rs2276064-A, associated with a low-activity trehalase, explains why Northern aboriginal peoples avoid food containing trehalose. It is assumed that the increase in the frequency of this variant in the populations of Northeast Asia can have been facilitated by gene drift acting in populations of a small effective size. However, it is possible that artificially induced trehalose deficiency in the traditional diet of the indigenous peoples of the Far North (due to the tradition of rejecting fungi) could also cause an increase in the frequency of low-activity trehalase, provided that this tradition has been existing among Northern aboriginal peoples for many generations.

Key words: gene *TREH*; exome; single nucleotide polymorphisms; human populations; Siberia; adaptive evolution.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Малярчук Б.А., Деренко М.В. Полиморфизм гена трегалазы (*TREH*) у коренного населения Сибири. Вавиловский журнал генетики и селекции. 2017;21(8):964-968. DOI 10.18699/VJ17.320

HOW TO CITE THIS ARTICLE:

Malyarchuk B.A., Derenko M.V. Polymorphism of the trehalase gene (*TREH*) in native populations of Siberia. Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding. 2017;21(8):964-968. DOI 10.18699/VJ17.320 (in Russian)

Received 31.07.2017

Accepted for publication 31.10.2017

© AUTHORS, 2017

Издвеле диета играла большую роль в жизни людей и направляла эволюцию метаболических путей. Усвоение простых и сложных полисахаридов невозможно без целой системы ферментов и транспортных белков, и любое изменение в диете находит свое отражение в этой системе (Karasov et al., 2011). Один из наиболее известных тому примеров – распространение в популяциях человека генетических вариантов лактазы, метаболизирующих лактозу у взрослых, в связи с распространением молочного животноводства в посленеолитическое время (Tishkoff et al., 2007). Известна также связь копияности гена амилазы слюны *AMY1* с уровнем потребления крахмала: копияность гена *AMY1* тем выше, чем больше крахмала содержится в диете (Perry et al., 2007). Однако наиболее яркие примеры взаимосвязи изменений диеты и частот аллелей генов преподносят исследования коренного населения Крайнего Севера. Так, у эскимосов, чукчей и коряков обнаружена максимальная частота (52 %) делеции гена панкреатической амилазы *AMY2A*, необходимой для расщепления крахмала, а также минимальная копияность этого гена в сочетании с пониженной копияностью гена амилазы слюны *AMY1*. Примерно у 30 % представителей коренного населения Северо-Востока Азии вообще отсутствует ген панкреатической амилазы *AMY2A* (Inchley et al., 2016).

Дефицитом олигосахаридов можно объяснить и высокую распространенность среди коренного населения крайнего севера Азии и Америки делеции динуклеотида AG в локусе rs781470490 гена сахаразы-изомальтазы (*SI*). Так, у канадских эскимосов частота этой делеции составляет 17 % (Marcadier et al., 2015), у эскимосов Гренландии – 39 % (Pedersen et al., 2017). По нашим данным, частота делеции AG в локусе rs781470490 у чукчей составляет 14.3 %, у коряков – 7.3 % и у охотских эвенов – 3.5 % (Малярчук и др., 2017). Делеция в локусе rs781470490 гена *SI* приводит к преждевременной терминации синтеза сахаразы-изомальтазы и, как следствие, к полному отсутствию этого фермента.

Высокая распространенность указанных выше вариантов полиморфизма генов *AMY2A* и *SI* на Северо-Востоке Азии может объясняться дефицитом крахмала и дисахаридов в рационе питания коренного населения в прошлом. Традиционные диеты аборигенов Крайнего Севера включают очень малые количества природных моно- и олигосахаридов, получаемых в основном при потреблении ягод и почек растений (Боринская и др., 2009).

Еще одним из малоусваиваемых коренным населением Арктики дисахаридов является трегалоза. По данным (Kozlov et al., 2005), 10.5 % коренного населения Арктики не усваивают трегалозу, в то время как среди европейцев неусваиваемость этого дисахаридов составляет 0.25–2 %. Трегалоза содержится в водорослях, высших грибах, лишайниках и некоторых высших растениях. Известно, что коренные народы Севера традиционно не едят грибов, объясняя это тем, что грибы – пища северного оленя, а не человека (Боринская и др., 2009). Однако основной причиной отказа от грибов, по всей видимости, является дефицит фермента трегалазы, расщепляющей трегалозу, у многих аборигенов Севера; из-за этого у них возникает боль в животе после употребления грибов в пищу (Arola

et al., 1999; Боринская и др., 2009). Традиция отказа от грибов могла привести к повышению частоты низкоактивной формы трегалазы у коренных жителей Крайнего Севера. Однако дефицит ферментов, метаболизирующих дисахариды, становится проблемой для здоровья коренного населения лишь в наше время, по мере расширения рациона питания аборигенов Севера за счет продуктов, содержащих трегалозу, сахарозу, лактозу и т. д.

О генетических причинах дефицита трегалазы стало известно относительно недавно. В 2013 г. опубликованы результаты исследования полиморфизма гена *TREH*, кодирующего трегалазу, в выборках индейцев пима, различающихся по активности трегалазы в плазме крови (Muller et al., 2013). Оказалось, что из проанализированных авторами вариантов полиморфизма локусов rs2276064, rs117619140, rs558907 и rs10790256 только аллель rs2276064-A достоверно ассоциируется с самой низкой активностью трегалазы: гомозиготам AA соответствовала активность 10.8 единиц фермента, гетерозиготам AG – 20.5, гомозиготам GG – 29.3 единицы активности (Muller et al., 2013). Поскольку популяции коренного населения Сибири ранее не были охвачены исследованиями полиморфизма гена *TREH*, то основной целью настоящей работы является изучение этой проблемы.

Материалы и методы

Секвенирование экзонов и прилегающих нетранслируемых участков генов проводили с помощью системы Agilent SureSelectXT Human All Exon V5+UTRs (Agilent Technologies) на платформе HiSeq 1500 (Illumina, США) для шести образцов ДНК от представителей коренного населения Сибири (3 коряка, 2 эвена и 1 эвенк). В качестве референтной последовательности использовали геномную сборку GRCh37.p13. Аннотацию вариантов полиморфизма производили с помощью сервера SeattleSeq Annotation. В работе анализировали участок хромосомы 11 (между позициями 118528026–118550399) с геном *TREH*, 3'-область которого перекрывается с таковой для гена *PHLDB1*.

В сравнительном анализе использованы также данные о полноэкзомном полиморфизме у представителей коренного населения Северо-Восточной Азии (4 эскимоса, 5 чукчей и 16 коряков) (Clemente et al., 2014). Проанализированы также полные экзоны коренного населения других регионов Сибири: Центральной Сибири (8 эвенов, 13 эвенков, 8 якутов), Южной Сибири (3 тувинца, 2 шорца, 6 алтайцев, 17 бурят, 6 монголов) и Западной Сибири (3 кета, 3 ханта, 3 манси, 3 селькупа, 6 ненцев, 2 нганасана) (Pagani et al., 2016).

Частоту аллелей и соответствие распределения генотипов равновесию Харди–Вайнберга рассчитывали с помощью пакета программ Arlequin 3.01 (Excoffier et al., 2007). Для выявления гаплотипов из генотипов с неизвестной гаметной фазой применяли алгоритм ELB (Excoffier et al., 2003) пакета программ Arlequin 3.01. При получении медианных сетей гаплотипов гена *TREH* использовали программу Network 4.6 (www.fluxus-engineering.com).

Результаты и обсуждение

Замена G → A в локусе rs2276064 гена *TREH*, которая, согласно работе (Muller et al., 2013), ассоциируется с самой

Table 1. Prevalence of allele rs2276064-A of the *TREH* gene in human populations

Population	rs2276064-A frequency	Source
Eastern Asia (<i>n</i> = 4300)	42.4	ExAC database
Northeastern Siberia (<i>n</i> = 28)	58.9	Clemente et al., 2014; this work
Central Siberia (<i>n</i> = 32)	31.3	Pagani et al., 2016; this work
Southern Siberia (<i>n</i> = 34)	32.4	Pagani et al., 2016
West Siberia (<i>n</i> = 20)	40.0	»
Latin America (<i>n</i> = 5784)	29.8	ExAC database
Southern Asia (<i>n</i> = 8251)	4.4	»
Europe (<i>n</i> = 33 196)	1.9	»
Africa (<i>n</i> = 4860)	0.6	»

n, sample size.

Table 2. Polymorphic variants of the *TREH* and *PHLDB1* genes in the indigenous population of Siberia

Nucleotide position on chromosome 11; nucleotide substitution	Polymorphism ID (0, newly found)	Gene	Site location; type of amino acid substitution
118528424, C→T	rs17748	<i>PHLDB1</i>	3'-UTR
118528466, T→G	rs7389	<i>PHLDB1-TREH</i>	3'-UTR
118529127, G→A	rs7928371	<i>TREH</i>	Gly541Gly
118529388, C→T	0	<i>TREH</i>	Интрон
118529703, G→A	rs2276064	<i>TREH</i>	Arg486Trp
118530611, T→C	rs2276065	<i>TREH</i>	Thr389Ala
118531335, T→C	rs140687140	<i>TREH</i>	Ile339Val
118532376, A→G	rs147473566	<i>TREH</i>	Met196Thr
118534082, C→T	rs10790256	<i>TREH</i>	Lys52Lys

низкой активностью трегалазы, приводит к аминокислотной замене Arg486Trp. Результаты анализа с помощью программы PolyPhen-2 (Adzhubei et al., 2013) показали, что эта замена сопровождается существенными изменениями физико-химических свойств фермента (индекс PolyPhen = 0.998). Следовательно, можно предположить, что замена Arg486Trp является патологической. Однако в некоторых популяциях мира распространенность аллеля rs2276064-A довольно высока (табл. 1). Так, по данным базы ExAC (<http://exac.broadinstitute.org>), частота этого варианта снижается от 42 % в Восточной Азии до 1.9 % в Европе и 0.6 % в Африке. Полученные нами результаты экзомного генотипирования популяций коренного населения Сибири показали, что в сибирских популяциях частота варианта rs2276064-A составляет 30–40 %, однако на северо-востоке Сибири наблюдается повышение частоты почти до 60 % (см. табл. 1). Таким образом, низкоактивная форма трегалазы с максимальной частотой распространена в популяциях коренного населения Северо-Восточной Азии.

Генотипирование экзонов и прилегающих участков гена *TREH*, включая 3'-область гена *PHLDB1*, у коренного населения Сибири позволило выявить девять полиморфных сайтов, обнаруженных у исследованных индивидумов более одного раза (табл. 2). Анализ генотипов, сформированных полиморфными вариантами в этих сайтах, показал наличие семи гаплотипов (табл. 3). Из них пять гап-

лотипов распространены в различных популяциях Сибири довольно равномерно (гаплотипы SQ1-SQ5), два редких гаплотипа обнаружены только в Южной Сибири (SQ6) и Северо-Восточной Сибири (SQ7).

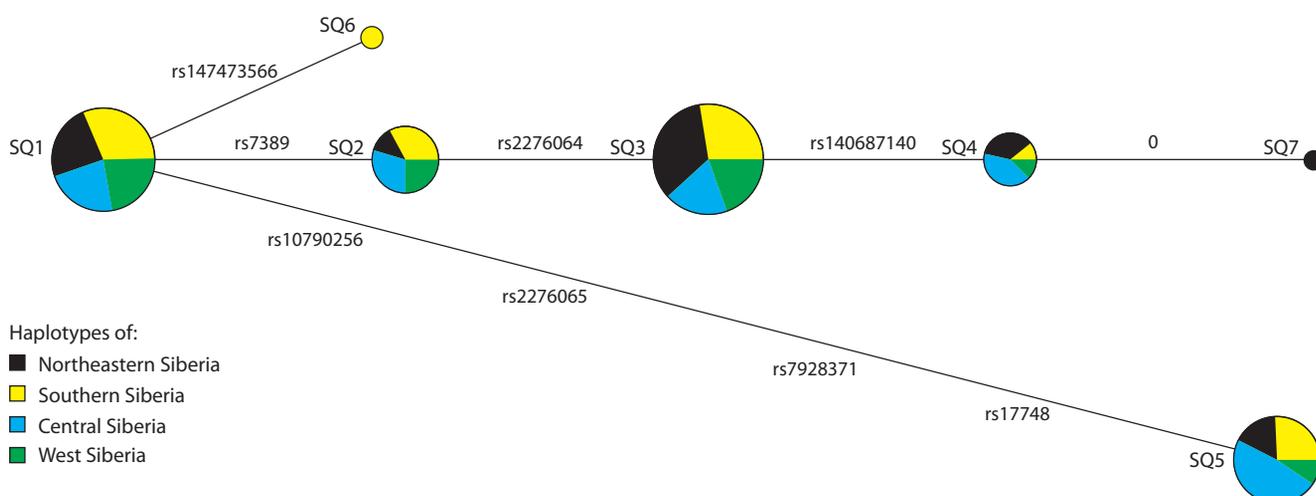
Анализ медианных сетей гаплотипов гена *TREH* в популяциях Сибири выявил, что центральным гаплотипом в сети является SQ1 (рисунок). От него отходит ветвь, сформированная гаплотипами SQ2, SQ3, SQ4 и SQ7. Единственное отличие гаплотипа SQ2 от SQ1 – замена в локусе rs7389, который расположен в нетранслируемой области между генами *PHLDB1* и *TREH*. Предполагается, что эта мутация повреждает сайт связывания с микроРНК, ингибируя тем самым трансляцию белковых продуктов расположенных рядом генов (White et al., 2015). Можно предположить, что уже на данном этапе начинается процесс снижения активности трегалазы. Кроме того, ранее была обнаружена ассоциация между полиморфизмом в локусе rs7389 и ингибированием активации плазминогена типа 1, что может повышать риск сосудистых осложнений и различных тромбозов (White et al., 2015). Следующая мутация появляется в гаплотипе SQ3. Это замена G→A в упомянутом уже локусе rs2276064, которая приводит к снижению активности трегалазы.

Гаплотип SQ4 характеризуется отсутствием варианта rs140687140-C, а гаплотип SQ7, выявленный только у коряков, определяется новым вариантом полиморфизма

Table 3. Prevalence of *TREH* haplotypes in Siberian populations, %

No.	Haplotype	Northeastern Siberia (n = 28)	Central Siberia (n = 32)	Southern Siberia (n = 34)	West Siberia (n = 20)
SQ1	CTGCGTTAC	26.8	23.4	30.9	35.0
SQ2	CGGCGTTAC	5.4	10.9	13.2	15.0
SQ3	CGGCATTAC	44.6	21.9	29.4	35.0
SQ4	CGGCATCAC	10.7	10.9	4.4	5.0
SQ5	TTACGCTAT	12.5	32.8	17.6	10.0
SQ6	CTGCGTTGC	0	0	4.4	0
SQ7	CGGTATCAC	3.6	0	0	0
Diversity, <i>h</i>		0.25 ± 0.16	0.33 ± 0.19	0.26 ± 0.16	0.2 ± 0.13
Number of genotypes		11	12	16	11

Note. The order of loci in haplotypes follows Table 2.



Median network of haplotypes for the *TREH* gene in Siberian populations.

Haplotype designations follow Table 3. Polymorphic loci are characterized in Table 2.

в интроне гена *TREH* (см. табл. 3). В Приложениях 1–3¹ приводятся сведения о распространенности вариантов полиморфизма в локусах rs7389, rs2276064 и rs140687140. Так, вариант rs7389-G довольно широко распространен в различных регионах мира, но максимум частоты (66 %) наблюдается в Северо-Восточной Сибири (см. Приложение 1). Вариант rs2276064-A относится также к числу довольно частых в различных регионах мира, но самые высокие его частоты отмечаются в популяциях Восточной Азии, с максимумом на Северо-Востоке Сибири (см. табл. 1; Приложение 2). Вариант rs140687140-C в региональных группах мира намного более редок: чаще всего он обнаруживается в Восточной Азии (1.6 %), а в остальных регионах его частота ниже 0.1 % (данные базы ExAC). В Сибири наиболее высокие частоты rs140687140-C характерны для коренного населения северо-восточной части региона (16.1 %) (см. Приложение 3).

Следует отметить, что только для локуса rs140687140 в южносибирской выборке наблюдалось достоверное от-

клонение от равновесия Харди–Вайнберга ($p = 0.045$) (см. Приложение 3). В этом случае уровень наблюдаемой гетерозиготности оказался достоверно ниже уровня ожидаемой. Такой результат может быть связан с генетической гетерогенностью южносибирской выборки, представленной коренным населением как Байкальского, так и Алтае-Саянского регионов. Тем не менее для подтверждения отклонения от равновесного распределения частот генотипов необходимо расширение выборок из этих популяций.

Кроме формирующей целую ветвь гаплотипов SQ2, SQ3, SQ4 и SQ7, у коренного населения Сибири присутствует гаплотип SQ5, выявленный в различных сибирских популяциях с частотой от 10 до 32.8 % (см. табл. 3). Этот гаплотип интересен тем, что он отличается от центрального гаплотипа SQ1 заменами сразу в четырех нуклеотидных позициях. По данным «Проекта 1000 геномов» (<http://www.ensembl.org>), гаплотип SQ5 присутствует в различных региональных группах мира. Максимальные частоты наблюдаются в Юго-Восточной Азии (37 %), минимальные – в Африке (3 %), а в Америке, Европе и Восточной Азии его частота составляет примерно 20 %.

¹ Приложения 1–3 см. по адресу:
<http://www.bionet.nsc.ru/vogis/download/pict-2017-21/appx16.pdf>

Результаты проведенного исследования показывают, что среди коренного населения Северо-Восточной Азии наблюдаются наиболее высокие частоты (около 60 %) варианта rs2276064-А гена *TREH*, ассоциирующегося с самой низкой активностью трегалазы. Это позволяет понять, почему коренные северяне избегают пищи, содержащей трегалозу, однако не объясняет причин высокой распространенности в популяциях малоактивной трегалазы. Повышенная частота варианта rs2276064-А характерна в целом для населения Восточной Азии (см. табл. 1), и поэтому возможно, что в популяциях коренного населения Северо-Восточной Сибири произошло еще большее увеличение частоты этого варианта в силу случайных причин, связанных с действием дрейфа генов в популяциях малой эффективной численности. Традиция отказа от грибов могла дополнительно снизить давление эволюционных факторов на этот локус и привести к увеличению частоты малоактивной формы трегалазы, особенно если такая традиция существует на протяжении многих поколений.

Acknowledgments

This study was supported by the RAS Presidium program "Pilot basic research for the development of Arctic regions of Russia". The authors acknowledge the assistance of Drs. T. Kivisild and F. Clemente (Cambridge University, UK).

Conflict of interest

The authors declare no conflict of interest.

References

Adzhubei I., Jordan D.M., Sunyaev S.R. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr. Protoc. Hum. Genet.* 2013;7:Unit7.20. DOI 10.1002/0471142905.hg0720s76.

Arola H., Koivula T., Karvonen A.L., Jokela H., Ahola T., Isokoski M. Low trehalase activity is associated with abdominal symptoms caused by edible mushrooms. *Scand. J. Gastroenterol.* 1999;34:898-903. PMID: 10522609.

Borinskaya S.A., Kozlov A.I., Yankovsky N.K. Genes and food traditions. *Etnograficheskoe obozrenie = Ethnographic Review.* 2009;3: 117-137. (in Russian)

Clemente F.J., Cardona A., Inchley C.E., Peter B.M., Jacobs G., Pagani L., Lawson D.J., Antão T., Vicente M., Mitt M., DeGiorgio M., Faltyskova Z., Xue Y., Ayub Q., Szpak M., Mägi R., Eriksson A., Manica A., Raghavan M., Rasmussen M., Rasmussen S., Willerslev E., Vidal-Puig A., Tyler-Smith C., Vilems R., Nielsen R., Metspalu M., Malyarchuk B., Derenko M., Kivisild T. A selective sweep on a deleterious mutation in the *CPT1A* gene in Arctic populations. *Am. J. Hum. Genet.* 2014;95:584-589. DOI 10.1016/j.ajhg.2014.09.016.

Excoffier L., Laval G., Balding D. Gametic phase estimation over large genomic regions using an adaptive window approach. *Hum. Genomics.* 2003;1:7-19. DOI 10.1186/1479-7364-1-1-7.

Excoffier L., Laval G., Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinf. Online.* 2007;1:47-50. PMID: PMC2658868.

Inchley C.E., Larbey C.D.A., Shwan N.A.A., Pagani L., Saag L., Antão T., Jacobs G., Hudjashov G., Eichstaedt T., Malyarchuk B., Derenko M., Wee J., Abdullah S., Ricaut F.-X., Mormina M.E., Vilems R., Metspalu M., Jones M.K., Armour J.A.L., Kivisild T. Selective sweep on human amylase genes postdates the split with Neanderthals. *Sci. Rep.* 2016;6:37198. DOI 10.1038/srep37198.

Karasov W.H., Martinez del Rio C., Caviades-Vidal E. Ecological physiology of diet and digestive systems. *Annu. Rev. Physiol.* 2011;73: 69-93. DOI 10.1146/annurev-physiol-012110-142152.

Kozlov A., Vershubsky G., Borinskaya S., Sokolova M., Nuvano V. Activity of disaccharidases in arctic populations: evolutionary aspects disaccharidases in arctic populations. *J. Physiol. Anthropol. Appl. Hum. Sci.* 2005;24:473-476. DOI 10.2114/jpa.24.473.

Malyarchuk B.A., Derenko M.V., Denisova G.A. The frequency of the inactive sucrase-isomaltase variant in indigenous populations of Northeastern Asia. *Russian Journal of Genetics.* 2017;53:1052-1054. DOI 10.1134/S1022795417090095.

Marcadier J.L., Boland M., Scott C.R., Issa K., Wu Z., McIntyre A.D., Hegele R.A., Geraghty M.T., Lines M.A. Congenital sucrase-isomaltase deficiency: identification of a common Inuit founder mutation. *Canad. Med. Assoc. J.* 2015;187:102-107. DOI 10.1503/cmaj.140657.

Muller Y.L., Hanson R.L., Knowler W.C., Fleming J., Goswami J., Huang K., Traurig M., Sutherland J., Wiedrich C., Wiedrich K., Mahke D., Ossowski V., Kobes S., Bogardus C., Baier L.J. Identification of genetic variation that determines human trehalase activity and its association with type 2 diabetes. *Hum. Genet.* 2013;132:697-707. DOI 10.1007/s00439-013-1278-3.

Pagani L., Lawson D.J., Jagoda E., Mörseburg A., Mitt M., Clemente F., Hudjashov G., DeGiorgio M., Eriksson A., Saag L., Wall J.D., Cardona A., Mägi R., Sayres M.A., Kaewert S., Inchley C., Scheib C.L., Järve M., Karmin M., Jacobs G.S., Antao T., Iliescu F.M., Kushniarevich A., Ayub Q., Tyler-Smith C., Xue Y., Yunusbayev B., Tambets K., Mallick C.B., Saag L., Pocheshkhova E., Andriadze G., Muller C., Westaway M., Lambert D., Zoraqi G., Turdikulova S., Dalimova D., Sabitov Z., Sultana G.N.N., Lachance J., Tishkoff S., Momynaliev K., Isakova J., Damba L.D., Gubina M., Nymadawa P., Evseeva I., Atramentova L., Utevska O., Ricaut F.-X., Brucato N., Sudoyo H., Letellier T., Cox M.P., Barashkov N.A., Skaro V., Mulahasanovic L., Primorac D., Sahakyan H., Mormina M., Eichstaedt C.A., Lichman D.V., Abdullah S., Chaubey G., Wee J.T.S., Mihailov E., Karunas A., Litvinov S., Khusainova R., Ekomasova N., Akhmetova V., Khidiyatova I., Marjanovic D., Yepiskoposyan L., Behar D.M., Balanovska E., Metspalu A., Derenko M., Malyarchuk B., Voevoda M., Fedorova S.A., Osipova L.P., Lahr M.M., Gerbault P., Leavesley M., Migliano A.B., Petraglia M., Balanovsky O., Khusnutdinova E.K., Metspalu E., Thomas M.G., Manica A., Nielsen R., Vilems R., Willerslev E., Kivisild T., Metspalu M. Genomic analyses inform on migration events during the peopling of Eurasia. *Nature.* 2016;538:238-242. DOI 10.1038/nature19792.

Pedersen C.T., Lohmueller K.E., Grarup N., Bjerregaard P., Hansen T., Siegmund H.R., Moltke I., Albrechtsen A. The effect of an extreme and prolonged population bottleneck on patterns of deleterious variation: Insights from the Greenlandic Inuit. *Genetics.* 2017;205:787-801. DOI 10.1534/genetics.116.193821.

Perry G.H., Dominy N.J., Claw K.G., Lee A.S., Fiegler H., Redon R., Werner J., Villanea F.A., Mountain J.L., Misra R., Carter N.P., Lee C., Stone A.C. Diet and the evolution of human amylase gene copy number variation. *Nat. Genet.* 2007;39:1256-1260. DOI 10.1038/ng2123.

Tishkoff S.A., Reed F.A., Ranciaro A., Voight B.F., Babbitt C.C., Silverman J.S., Powell K., Mortensen H.M., Hirbo J.B., Osman M., Ibrahim M., Omar S.A., Lema G., Nyambo T.B., Ghorji J., Bumpstead S., Pritchard J.K., Wray G.A., Deloukas P. Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Genet.* 2007;39:31-40. DOI 10.1038/ng1946.

White M.J., Kodaman N.M., Harder R.H., Asselbergs F.W., Vaughan D.R., Brown N.J., Moore J.H., Williams S.M. Genetics of plasminogen activator inhibitor-1 (PAI-1) in a Ghanaian population. *PLoS ONE.* 2015;10:e0136379. DOI 10.1371/journal.pone.0136379.



Методы геномного редактирования дрожжей: история и современное состояние

А.С. Розанов¹✉, В.Н. Шляхтун¹, Л.А. Текутьева^{2,3}, О.М. Сон^{2,3}, С.В. Сизова^{4,5}, С.Е. Пельтек¹

¹ Федеральное исследовательское учреждение Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Дальневосточный федеральный университет, Владивосток, Россия

³ ООО «АРНИКА», Владивосток, Россия

⁴ ЗАО «Центр новых технологий и бизнеса», Москва, Россия

⁵ Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова Российской академии наук, Москва, Россия

Дрожжи являются модельным эукариотическим организмом, на котором отрабатываются многие предположения о работе генома, а также методы его редактирования. Наиболее часто в исследовательских работах используют *Saccharomyces cerevisiae*, которые очень хорошо приспособлены физиологически к культивированию в условиях биореактора и признаны абсолютно безопасными. В последнее десятилетие методы генетической инженерии дрожжей претерпели значительные изменения. Появились новые инструменты, которые пришли из смежных направлений и позволили значительно ускорить процесс получения новых штаммов. Прежде всего это белки для направленного внесения изменений в последовательность ДНК. Длительное время методы редактирования генома дрожжей базировались на использовании их собственной системы гомологичной рекомбинации. Она удобна и удовлетворяла потребности исследователей на протяжении нескольких десятилетий, до того времени, когда на первый план стали выходить высокопроизводительные методы. Во втором десятилетии XXI века произошло бурное развитие высокопроизводительных подходов, в первую очередь методов анализа в биологии: геномики, транскриптомики, протеомики, метаболомики, интерактомики и др. Сформировалась биоинформационная база, которая позволила быстро обрабатывать растущий поток информации и моделировать клеточные процессы. В результате скорость анализа и предсказания мишеней для редактирования генома стала превышать скорость их получения, что, естественно, обусловило поиск новых методов генетической инженерии. Особенно сильно этот процесс затронул работы по модификации свойств микроорганизмов. Современные задачи стали требовать не единичных модификаций, а десятков и сотен, а иногда тысяч модификаций. В результате исследователи, занимающиеся дрожжами, стали вовлекать в работу новые инструменты геномного редактирования, которые ранее развивались для изучения более сложных объектов, таких как животные, растения, клеточные линии и др. Современные методы геномной инженерии дрожжей позволяют вносить несколько модификаций в геном за один шаг. В данном обзоре рассматривается вопрос применения и перспектив дальнейшего развития методов направленного редактирования генома в инженерии дрожжей.

Ключевые слова: дрожжи; *Saccharomyces cerevisiae*; генетическая инженерия; ZFN; цинковые пальцы; TALENS; CRISPR/Cas; Аргонавт.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Розанов А.С., Шляхтун В.Н., Текутьева Л.А., Сон О.М., Сизова С.В., Пельтек С.Е. Методы геномного редактирования дрожжей: история и современное состояние. Вавиловский журнал генетики и селекции. 2017;21(8):969-978. DOI 10.18699/VJ17.321

HOW TO CITE THIS ARTICLE:

Rozanov A.S., Shlyahntun V.N., Tekutieva L.A., Son O.M., Sizova S.V., Peltek S.E. Methods of yeast genome editing. Vavilovskii Zhurnal Genetiki i Selektcii = Vavilov Journal of Genetics and Breeding. 2017;21(8):969-978. DOI 10.18699/VJ17.321 (in Russian)

Received 26.07.2017

Accepted for publication 11.10.2017

© AUTHORS, 2017

✉ e-mail: sibiryak.n@gmail.com

Methods of yeast genome editing

A.S. Rozanov¹✉, V.N. Shlyahntun¹, L.A. Tekutieva^{2,3}, O.M. Son^{2,3}, S.V. Sizova^{4,5}, S.E. Peltek¹

¹ Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

² Far Eastern Federal University, Vladivostok, Russia

³ Arnika, Ltd., Vladivostok, Russia

⁴ Closed Joint Stock Company "Center for New Technologies and Business", Moscow, Russia

⁵ Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry RAS, Moscow, Russia

Yeasts are a convenient model eukaryote used for genome studies and genome editing. *Saccharomyces cerevisiae* is the species most widely employed in biotechnology, since it is easily cultivated in bioreactors and is absolutely safe. The last decade saw a significant development of methods of yeast genetic engineering and the creation of novel instruments adapted from other fields, which allowed one to significantly accelerate the construction of new strains. The most prominent examples are the proteins used for directed DNA editing. For a long time, yeast genome engineering was based on the yeasts' system of homologous recombination. It was sufficient for several decades before the development of high-throughput methods. Many high-throughput methods were developed in the second decade of the XXI century, including those used in genomics, transcriptomics, proteomics, metabolomics, interactomics, etc. Modern bioinformatic databases now allow one to rapidly process the increasing flow of information and model cellular processes. As a result, the rate of analysis and prediction of targets for genome editing is currently higher than the rate of genome editing, which led to the development of new methods of genetic engineering. This process was particularly pronounced for microorganisms. Modern tasks require tens, hundreds, sometimes even thousands of genome modifications, which made researchers to look for new techniques. As a result, the instruments used for more complex objects, such as animals, plants, and cell lines, were adapted for yeasts. Modern methods for yeast genome editing allow introducing several modifications into the genome in a single step. In this study, we review the methods of directed genome editing and their applications and perspectives for yeasts.

Key words: yeast; *Saccharomyces cerevisiae*; genetic engineering; ZFN; zinc fingers; TALENS; CRISPR/Cas; Argonaut; NgAgo.

Дрожжи – модельный эукариотический организм, на котором отрабатываются многие предположения о работе генома, а также методы его редактирования. Чаще всего в исследовательских работах используют *Saccharomyces cerevisiae*, которые физиологически очень хорошо приспособлены к культивированию в условиях биореактора и признаны абсолютно безопасными. История их применения насчитывает не одно тысячелетие. С развитием промышленной биотехнологии дрожжи также нашли свою нишу (Botstein et al., 1997; Jensen, Keasling, 2015). Использование дрожжей в современной биотехнологии весьма распространено (рис. 1), и если раньше это были природные штаммы, то теперь уровень их модификации бывает очень высоким.

Устойчивое изменение свойств организма возможно только за счет модификации его генома. На первых этапах исследователи занимались изучением свойств генома посредством нокаута отдельных генов или изменения экспрессией какого-то одного гена. По мере накопления опыта и знаний требования к методам направленного редактирования генома все более усложняются, биомодификаций становится больше и растут требования к надежности и точности. В настоящее время наиболее процессивные методы редактирования генома используют для бактерий. Более сложно организованные объекты, такие как животные и растения, со временем тоже требуют большей процессивности. Особенно остро вопрос эффективности методов преобразования генома для микроорганизмов встал, когда методы биоинформатики и аналитические подходы позволили предсказывать и анализировать сложные модельные эксперименты, требующие модификации и тонкой регуляции большого числа генов и метаболических путей (Lechner et al., 2016). Уже сейчас для отдельных задач в биологии необходимы сотни и тысячи модификаций. Из числа решенных одной из самых объемных задач можно считать получение штамма *Escherichia coli* с уменьшенным числом кодонов. В ходе достижения этой цели был получен геном, содержащий более 62 000 модификаций (Ostrov et al., 2016). Эта работа представляет вариант синтеза генома *de novo*, но хорошо демонстрирует, какие объемы модификаций в настоящее время задействованы в передовых исследованиях.

Наиболее известной работой, выполненной на дрожжах, является воссоздание метаболического пути биосинтеза артемизинина в *S. cerevisiae*. Изначально этот метаболит был обнаружен в полыни *Artemisia annua*. Артемизинин оказался очень эффективен при малярии, но из биомассы полыни невозможно извлекать его в количествах, достаточных для обеспечения широкого круга заболевших. Тогда группа ученых под руководством J. Keasling получила рекомбинантный штамм дрожжей, способный производить артемизининовую кислоту, из которой химически получали артемизинин в количестве 25 г/л. Благодаря этому способу новое лекарство стало доступным даже для жителей беднейших стран, где распространена малярия (Paddon et al., 2013). Для этого в геном пекарских дрожжей были внесены 12 генов и отрегулирована их совместная работа.

Кроме того, в настоящее время выполнены или выполняются несколько проектов, связанных с перестройкой

метаболизма дрожжей, получением продуцентов биотоплива и компонентов химического синтеза (Nielsen et al., 2013), синтезом терпеноидов (Wriessnegger, Pichler, 2013) и др. Подобные задачи требуют сотен модификаций, которые должны выполняться быстро и надежно, и соответствующих инструментов, развитием которых занимается множество лабораторий.

Современные инструменты для направленного редактирования генома дрожжей

Методы направленного редактирования генома дрожжей развивались параллельно с общими методами молекулярной биологии. С точки зрения редактирования генома, дрожжи оказались очень хорошим объектом. Благодаря высокому уровню гомологичной рекомбинации, на начальных этапах необходимость развития методов направленного редактирования генома дрожжей не волновала исследователей, вплоть до второго десятилетия XXI века. Первые инструменты, ZFN и TALENS, оказались слишком трудозатратными, чтобы вытеснить классические методики. Дрожжи при этом использовались как вспомогательный материал для редактирования генома более сложных объектов (Kuijpers et al., 2013). Однако требования к количеству вводимых в геном изменений, их точности и скорости введения постоянно растут, и для современного процесса получения модифицированных штаммов необходимы большие возможности, чем в случае применения естественных механизмов клетки дрожжей. В 2012 г. для редактирования генома эукариот начали использовать систему CRISPR/Cas, которая по своей эффективности конкурировала с классическими методами редактирования генома дрожжей. В 2016 г. был открыт еще один белок, перспективный для геномного редактирования, – Аргонавт. К сожалению, в настоящий момент достоверно показана работоспособность только его термофильных вариантов. Несмотря на это, подобный белок мог бы значительно обогатить возможности генетической инженерии свободноживущих одноклеточных организмов.

Цинковые пальцы

Одним из первых современных инструментов для направленного редактирования генома стали так называемые цинковые пальцы (ZFNs) – нуклеазы, способные узнавать определенный сайт ДНК за счет группы ДНК-связывающих белковых доменов, имеющих аналогичное название. Они содержат характерный участок, который включает два цистеиновых и два гистидиновых остатка. Эти аминокислоты образуют хелатный комплекс с ионом цинка, а расположенная между ними полипептидная цепочка выделяется в виде пальца. Каждый цинковый палец может узнавать определенную последовательность ДНК из трех нуклеотидов и специфично связываться с ней. Цинк здесь играет структурную роль, способствуя стабильности домена (Kim et al., 1996). Цинковые пальцы входят в состав белков, в основном транскрипционных факторов, которые играют важную роль в различных клеточных процессах, таких как репликация, репарация, транскрипция, трансляция, метаболизм, передача сигналов, пролиферация клеток, апоптоз и др.

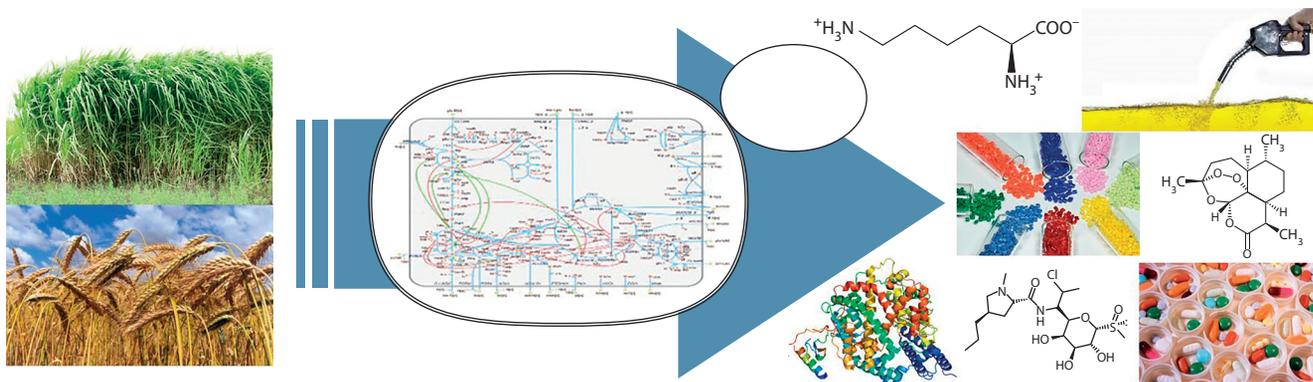


Fig. 1. Use of yeast in the conversion of renewable sources of energy and matter.

Идея комбинирования нескольких цинковых пальцев для получения домена, способного к узнаванию целевой последовательности, появилась в конце XX века. Согласно этой идее, домен цинкового пальца может быть спроектирован так, чтобы узнавать желаемую последовательность ДНК и связываться с ней. Если к этому домену добавить нуклеазный домен, то можно получить инструмент для направленного разрезания цепи ДНК. Такие белки были предложены в 1996 г. (Kim et al., 1996). Для генерации двунитовых разрывов ДНК использовали бинарную конструкцию (Vitinaite et al., 1998). Поскольку направленное редактирование генома очень перспективно, к работе над развитием нового инструмента параллельно приступили несколько групп, которые применяли множество различных молекулярно-биологических техник для повышения точности и эффективности ZFNs (Kim et al., 1996; Durai et al., 2005; Guo et al., 2010).

На первых порах внимание было сфокусировано на получении специфических белков, содержащих цинковые пальцы (ZFPs), через бактериальный двугибридный метод с последующим соединением ZFPs с доменом нуклеазы *FokI* (Kim et al., 1996; Durai et al., 2005). Но в большинстве случаев в условиях *in vivo* ZFNs оказались не способны разрезать сайты-мишени. Тогда ученые из Китая провели исследование, направленные на разработку совместного скрининга и проверки для получения эффективных ZFP нуклеаз в дрожжах. Для этого были использованы штаммы *S. cerevisiae*, которые несут репортерные системы, позволяющие детектировать необходимую активность. Промежуточная репортерная плаزمид имеет два гомологичных плеча в 30 п. н. *Gal4* и сайт-мишень ZFN. Двойная цепь ДНК разрезалась по целевой последовательности с помощью ZFNs и восстанавливалась механизмом одноцепочечного отжига (single strand annealing – SSA). В случае удачного восстановления начиналась наработка транскрипционного фактора *Gal4*, необходимого для запуска двух ауксотрофных маркеров (Wang et al., 2013).

Однако у метода «цинковых пальцев» обнаружились серьезные недостатки: во-первых, это не вполне строгое распознавание тринуклеотидных повторов, что приводит к заметному числу нецелевых расщеплений ДНК. Во-вторых, метод оказался весьма трудозатратным и дорогостоящим, поскольку для каждой последовательности ДНК необходимо создать свою оптимизированную бел-

ковую структуру zinc-finger нуклеазы. Поэтому система «цинковые пальцы» широкого распространения не получила, и в 2010 г. научный интерес к этому объекту начал резко снижаться. Из-за сложности использования ZFNs не нашли широкого применения в генетической инженерии дрожжей, так как они не имели очевидных преимуществ в сравнении с простой рекомбинацией. Но, несмотря на это, система была опробована на дрожжах, которые выступали в роли модельных объектов.

TALENs

Более перспективным, чем ZFNs, средством избирательного воздействия на ДНК оказались конструкции на основе химерных нуклеаз, названные TALENs (transcription activator-like effector nucleases) (Sanjana et al., 2012). TALE нуклеазы – это полноценный белок TALE или его участок, объединенный с доменом, расщепляющим ДНК *FokI*-нуклеазой. Они были разработаны как высокоэффективный инструмент направленной модификации генома. Повторяющиеся домены TALE обычно состоят из 34 аминокислот, две из которых, расположенные в позициях 12 и 13, высоко вариабельные (repeat variable di-residue – RVD), и именно они отвечают за узнавание определенного нуклеотида. Идентичность этих двух остатков определяет специфичность связывания каждого TALE-повтора. Простой шифр указывает целевую базу каждого RVD (NI = A, HD = C, NG = T, NN = G или A). Также в природе были найдены TALENs, имеющие переменное число мономеров, от 1.5 до 33.5. Таким образом, каждый мономер предназначен для одного нуклеотида, и линейная последовательность мономеров в TALE определяет последовательность ДНК-мишени в направлении 5'–3' (Sanjana et al., 2012).

Природным прототипом TALENs стали белки (TAL-effectors) некоторых бактерий рода *Xanthomonas*, паразитирующих в клетках сельскохозяйственных растений. Попадая в ядро растительной клетки, эти бактериальные белки имитируют транскрипционные факторы и связываются с определенными участками ДНК, активируя таким образом гены, необходимые для выживания паразита (Sanjana et al., 2012).

С помощью искусственных нуклеаз TALENs стало возможным внести двунитовые разрывы в любой участок генома. В 2011 г. на страницах авторитетного журнала

Nature Methods методы геномной инженерии, в первую очередь использование TALENs, были названы методом года, благодаря колоссальному спектру возможных применений в самых разных областях фундаментальной и прикладной науки: от функциональной геномики и биологии развития до сельскохозяйственной биотехнологии (Joung, Sander, 2013).

С 2010 по 2013 г. велись работы по модификации эндогенных генов дрожжей, плодовой мушки, аскариды, лягушек, крыс, свиней и многих других организмов при помощи TALENs с использованием технологий и методик, ранее разработанных для ZFNs (Joung, Sander, 2013). Как и ZFNs, TALENs не нашли широкого применения в генетической инженерии дрожжей из-за трудоемкости их использования. Однако из-за простоты работы с дрожжами их применяли для развития и тестирования метода (Li et al., 2011).

Группа исследователей из Франции изучала TALENs с использованием *S. cerevisiae* в качестве модельного объекта (Richard et al., 2014). Экспансия тринуклеотидных повторов является причиной более 20 тяжелых неврологических расстройств у людей. Большое количество исследований было посвящено изучению механизмов, влияющих на экспансию многочисленных CAG/CTG повторов. В упомянутой работе TALEN применялся для распознавания и вырезания CAG/CTG тринуклеотидных повторов. В результате показано, что применение TALENs эндонуклеаз имеет 100 % эффективность в сокращении числа повторов, не вызывая какой-либо другой мутации в геноме дрожжей (Richard et al., 2014).

В работе (Li et al., 2011) описаны модульные технологии сборки заказных dTALENs и характеристики десяти таких dTALENs для направленной модификации генов. В методе использованы четыре основных повтора совместно с RVDs: NI, NG, NN и HD. Метод позволяет генерировать 48 готовых к использованию модулей, которые могут быть использованы для сборки до 23 повторяющихся единиц в любом заданном порядке. К преимуществам метода модульной сборки относятся простота, скорость и экономичность, благодаря чему этот метод может быть освоен в большинстве академических или промышленных лабораторий молекулярной биологии. Все десять dTALENs продемонстрировали эффективность нокаута и/или замены гена в тестах с тремя различными хромосомными генами *S. cerevisiae*. Данный анализ является надежным и легким индикатором активности TALEN и подтверждает, что функциональные dTALENs могут использоваться для широкого спектра геномных локусов (Li et al., 2011).

Технология обладает значительным потенциалом в области экспериментальной биологии и медицины и в широком диапазоне отраслей наук о жизни. В течение двух лет TALENs были самым востребованным методом генетической инженерии. Но с 2012 г. популярность стала набирать другая система редактирования генов – CRISPR/Cas, которая в 2015 г. окончательно заменила TALENs как метод генетической модификации.

CRISPR/Cas

Система CRISPR (clustered regularly interspaced short palindromic repeats – короткие палиндромные повторы, рас-

положенные группами, равномерно удаленные друг от друга) – это естественная система адаптивного иммунитета бактерий и архей, защищающая их от проникновения чужеродной ДНК.

Специфические нуклеотидные повторы были обнаружены в 1987 г. случайно в последовательностях, фланкирующих 3'-концы *iap*-генов *E. coli* (Ishino et al., 1987). В 2002 г. R. Jansen с коллегами предложили акроним CRISPR. Ими было охарактеризовано новое семейство повторяющихся последовательностей ДНК, присутствующее у доменов бактерий и архей, но отсутствующее у эукариот и вирусов. Выявлены ассоциированные с повторами гены *cas* (CRISPR-associated), которые располагались неизменно рядом с локусом CRISPR, а также показана их взаимосвязь (Jansen et al., 2002). Структура CRISPR-кассеты представляет собой набор участков уникальной ДНК (захваченных последовательностей ДНК вирусов и плазмид) примерно одинаковой длины (обычно 32–38 п. н.), названных спейсерами. Спейсеры отделены друг от друга CRISPR-повторами, палиндромное строение которых способствует формированию вторичных структур в виде «шпилек». Размер CRISPR-повтора может варьировать от 21 до 37 п. н. (Jansen et al., 2002). Как правило, CRISPR находятся в бактериальной хромосоме, но могут быть идентифицированы и в плазмидной ДНК (Godde, Bickerton, 2006), их число может достигать 375, но обычно меньше 50 (Hogvath, Barrangou, 2010). В 2005 г. было показано, что белковые продукты расположенных рядом *cas*-генов обладают хеликазной и нуклеазной активностью (Haft et al., 2005). Дальнейшие исследования этой системы позволили установить ее роль в обеспечении защиты микроорганизмов от вторжения чужеродной ДНК (Boltin et al., 2005; Barrangou et al., 2007; Mojica et al., 2009). В 2012 г. две группы исследователей смогли направить нуклеазную активность Cas9 специфичным по последовательности образом *in vitro* (Gasiunas et al., 2012; Jinek et al., 2012). Сразу после этого в 2013 г. были опубликованы работы по применению CRISPR/Cas для инженерии генома клеток животных (Cong et al., 2013; Mali et al., 2013). С тех пор метод был оптимизирован и адаптирован для редактирования геномов разных организмов: животных, растений, насекомых, грибов, дрожжей, бактерий.

Эффективность CRISPR/Cas превзошла возможности естественной рекомбинации дрожжей, что способствовало ее активному вовлечению в разработку методов их геномного редактирования. Применение Cas9 для внесения разрывов в геном *S. cerevisiae* увеличило эффективность гомологичной рекомбинации пятикратно в присутствии одноцепочечных олигонуклеотидов в качестве матрицы, а в присутствии двуцепочечных олигонуклеотидов – 130-кратно (DiCarlo et al., 2013). Впоследствии метод CRISPR/Cas был использован для модификации других видов дрожжей (см. таблицу).

В природной системе CRISPR/Cas белок Cas9 направляется к целевому участку ДНК гибридом из двух молекул РНК – CRISPR RNA (crRNA) и tracrRNA, транскрибирующей crRNA. Вместе эти две молекулы РНК формируют вторичную структуру, которая связывается с белком Cas9. crRNA направляет систему к целевому участку генома из примерно 20 пар оснований, комплементарных фрагменту

Yeast species for which CRISPR/Cas systems for targeted genome editing have been developed

Species	Reference(s)
<i>Saccharomyces cerevisiae</i>	DiCarlo et al., 2013; Gao, Zhao, 2014; Ryan et al., 2014; Zhang et al., 2014; Bao et al., 2015; Horwitz et al., 2015; Jakočiūnas et al., 2015; Mans et al., 2015; Stovicek et al., 2015; Generoso et al., 2016; Jensen et al., 2017; Liu et al., 2017; Ng, Dean, 2017; Reider Apel et al., 2017; et al.
<i>Yarrowia lipolytica</i>	Schwartz et al., 2016
<i>Kluyveromyces lactis</i> , <i>Kluyveromyces marxianus</i>	Horwitz et al., 2015; Nambu-Nishida et al., 2017
<i>Komagataella phaffii</i> (before <i>Pichia pastoris</i>)	Weninger et al., 2016
<i>Schizosaccharomyces pombe</i>	Jacobs et al., 2014
<i>Ogataea polymorpha</i>	Numamoto et al., 2017
<i>Candida albicans</i> , <i>C. glabrata</i> , <i>C. lusitaniae</i>	Vyas et al., 2015; Cen et al., 2017; Norton et al., 2017
<i>Cryptococcus neoformans</i>	Wang et al., 2016

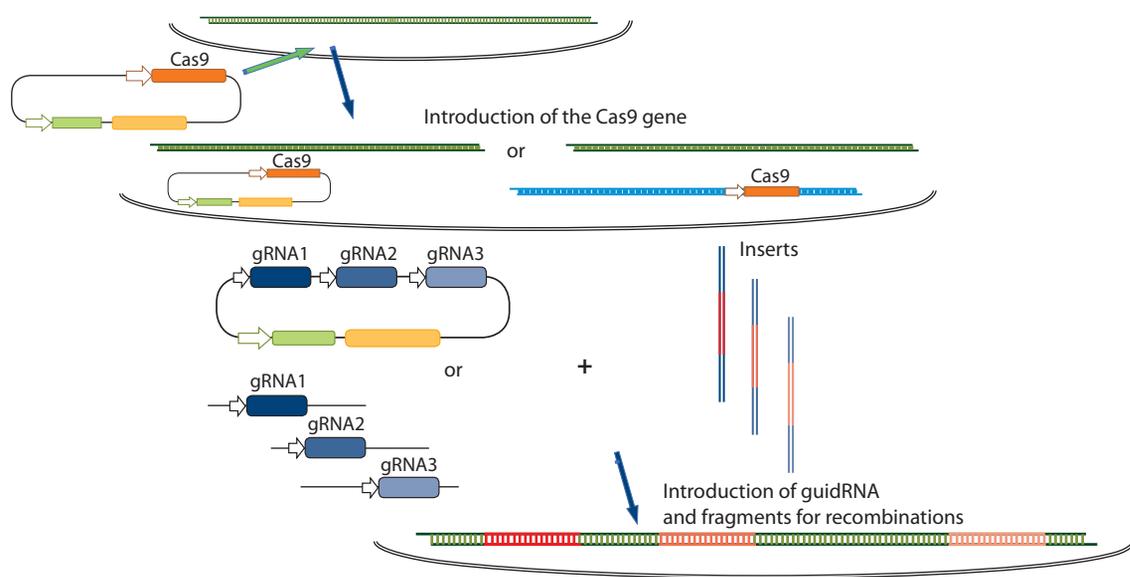


Fig. 2. Genetic engineering of yeast with the CRISPR/Cas system.

на crRNA. За целевой последовательностью должен следовать мотив 5'-NGG-3', названный PAM (protospacer adjacent motif). ДНК разрезается за три нуклеотида до PAM последовательности (Gasiunas et al., 2012). Для большего удобства в генетической инженерии crRNA и tracrRNA были объединены в одну guiding RNA (гидРНК) через линкер (Jinek et al., 2012).

Чтобы осуществить изменения в геноме с помощью системы CRISPR, в клетке одновременно должны присутствовать белок Cas9, направляющая РНК (гидРНК), а также, если необходимо, ДНК-матрица для гомологичной рекомбинации. Наиболее часто используемый белок в генетической инженерии *S. cerevisiae* – Cas9 из *Streptococcus pyogenes*. Последовательность гена белка может быть природной (Ryan et al., 2014; Bao et al., 2015), оптимизированной для экспрессии в клетках человека (DiCarlo et al., 2013; Gao, Zhao, 2014; Zhang et al., 2014; Jakočiūnas et al., 2015; Mans et al., 2015; Stovicek et al., 2015) и дрожжей (Horwitz et al., 2015; Generoso et al., 2016). Ген белка Cas9 экспрессировали под контролем конститутивных промоторов разной силы. Наиболее часто используемый про-

мотор – TEF1 (DiCarlo et al., 2013; Zhang et al., 2014; Bao et al., 2015; Jakočiūnas et al., 2015; Mans et al., 2015; Shi et al., 2016; Jensen et al., 2017; Liu et al., 2017; Ng, Dean, 2017), второй по частоте использования промотор – ADH1 (Gao, Zhao, 2014; Jacobs et al., 2014; Fernandez, Berro, 2016; Reider Apel et al., 2017), далее следуют RNR2 (Ryan et al., 2014), FBA1 (Horwitz et al., 2015), TDH3 (Laughery et al., 2015), PGK1 (Lee et al., 2015) и др.

Чаще всего ген Cas9 вводили в клетку с использованием плазмидных векторов, низкокопийных самореплицирующихся центромерных векторов (DiCarlo et al., 2013; Zhang et al., 2014; Jakočiūnas et al., 2015; Stovicek et al., 2015) или многокопийных векторов на основе 2μ вектора (Gao, Zhao, 2014; Ryan et al., 2014; Bao et al., 2015; Horwitz et al., 2015; Laughery et al., 2015; Generoso et al., 2016). В работе (Mans et al., 2015) были описаны сконструированные штаммы *S. cerevisiae*, в которых ген Cas9 был клонирован в геном, что позволяет использовать меньше селективных маркеров, а также сократить число необходимых стадий. Схема эксперимента по генетической инженерии дрожжей приведена на рис. 2.

ГидРНК – это второй важный элемент метода CRISPR/Cas редактирования генома. Введение и правильная транскрипция гидРНК имеют наибольшее значение во всем процессе клонирования с использованием системы CRISPR/Cas. Введение в *S. cerevisiae* чаще всего проводят в составе 2 μ вектора. Обычно Cas9 и гидРНК располагают в разных конструкциях, но в ряде работ их располагали в одной плазмиде (Ryan, Cate, 2014), хотя ввиду очень большого размера гена Cas9 такой вариант достаточно труден в исполнении. К тому же экспрессия Cas9 должна предшествовать трансфекции генов гидРНК (Walter et al., 2016). Использование ПЦР-фрагментов для конструирования гидРНК кассет также не получило распространения из-за большей эффективности использования плазмид (DiCarlo et al., 2013). Функциональные гидРНК получают с помощью разных конструкций, включающих промотор, терминатор и функциональные элементы для созревания РНК: 1) промотор для полимеразы III, транскрибирующий РНК с лидерной последовательностью, которая отрезается в процессе созревания (DiCarlo et al., 2013; Farzadfard et al., 2013); 2) промоторы для полимеразы III, содержащие *cis*-регуляторные элементы, транскрибирующие последовательность гидРНК с рибозимом, отрезающим 5'-конец (Ryan et al., 2014); 3) промотор для полимеразы II, транскрибирующий гидРНК с двумя рибозимами, отщепляющими дополнительные последовательности РНК с обоих концов (Gao, Zhao, 2014).

Эффективная работа системы гомологичной рекомбинации дрожжей позволяет параллельно вносить несколько изменений в геном с использованием CRISPR/Cas. Решающим этапом при этом является одновременная экспрессия нескольких гидРНК, для достижения чего используют несколько путей: несколько векторов с различными маркерами устойчивости (Mans et al., 2015); один вектор (Ryan et al., 2014; Jakočiūnas et al., 2015; Lee et al., 2015); трансфекция нескольких линейаризованных кассет (Horwitz et al., 2015).

Примеры использования CRISPR/Cas в работе с дрожжами

Получение сложных веществ при помощи микробной синтеза – задача, которая требует большого объема изменений в геноме. За последние годы сделано несколько подобных работ с использованием метода CRISPR/Cas. Очень интересное исследование выполнено под руководством J.D. Keasling и A. Mukhopadhyay (Reider Apel et al., 2017). Ими разработан набор инструментов для изучения метаболических путей в клетках *S. cerevisiae*, включая набор промоторов, выявлена зависимость интенсивности экспрессии от позиции на хромосомах, создан программный продукт для разработки плазмид и последовательностей для выполнения конкретных задач, направленных на оптимизацию метаболических путей. Так как для экспрессии белков в рабочем состоянии важна локализация, был получен набор лидерных последовательностей для локализации белка в разных частях клетки. Этот набор инструментов был применен для разработки промышленного штамма-продуцента таксадиена. В итоге получен штамм, способный производить таксадиен, а в результате оптимизации метаболических путей удалось

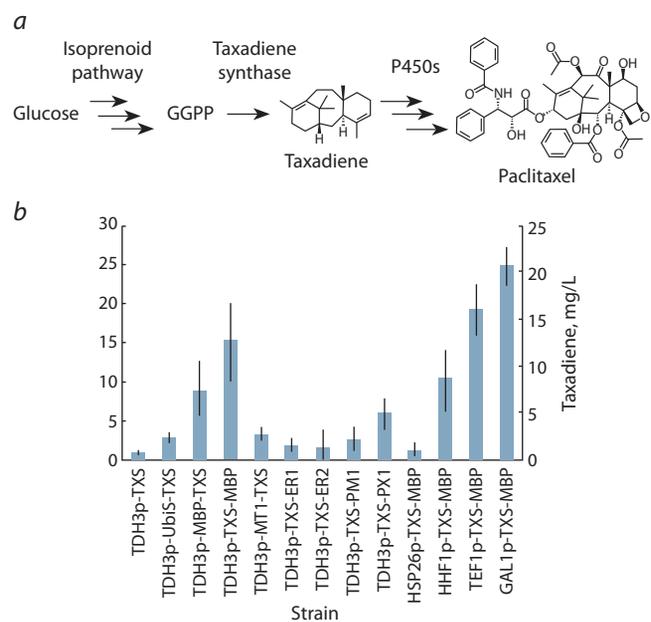


Fig. 3. (a) Pathways of taxadiene biosynthesis and further metabolism in the recombinant *S. cerevisiae* strain. (b) Taxadiene production by strain obtained in the study according to (Reider Apel et al., 2017).

увеличить его концентрацию в культуре в 25 раз (рис. 3) (Reider Apel et al., 2017). В другой работе авторы получили штамм-продуцент нарингенина, который впоследствии был оптимизирован для снижения концентрации побочного продукта (phloretic acid) (Vanegas et al., 2017).

Получение функциональных белков – еще одна задача, которая может быть эффективно решена с использованием дрожжей и требует значительного объема работ по генетической инженерии. В публикации (Sheng et al., 2017) проведена оптимизация выхода Hepatitis B Small Antigen при помощи направленного геномного редактирования *S. cerevisiae*. С одной стороны, работа не содержит оригинальных решений, основанных на моделировании и глубоком понимании процессов, с другой – она демонстрирует возможности использования технологии CRISPR/Cas для высокоэффективного редактирования генома. В ходе исследования авторы параллельно проводили нокаут нескольких генов в разных вариантах для проверки эффективности наработки целевого белка. В результате удалось повысить конечную концентрацию Hepatitis B Small Antigen (HBsAg) в 2.7 раза (Sheng et al., 2017).

Другое важное направление современной биотехнологии – это получение компонентов для крупнотоннажного химического синтеза. Так, при помощи разработанного набора инструментов для выполнения метаболической инженерии *Schizosaccharomyces pombe* с использованием CRISPR/Cas был выделен штамм-продуцент молочной кислоты (Ozaki et al., 2017). Для этого в геноме были нокаутированы шесть генов, встроены три гена и проведена оптимизация их экспрессии. Итоговый штамм позволил получить молочную кислоту в концентрации 25 г/л с выходом 0.71 г/г. Дальнейшая оптимизация может увеличить выход конечного продукта до значения, достаточного для

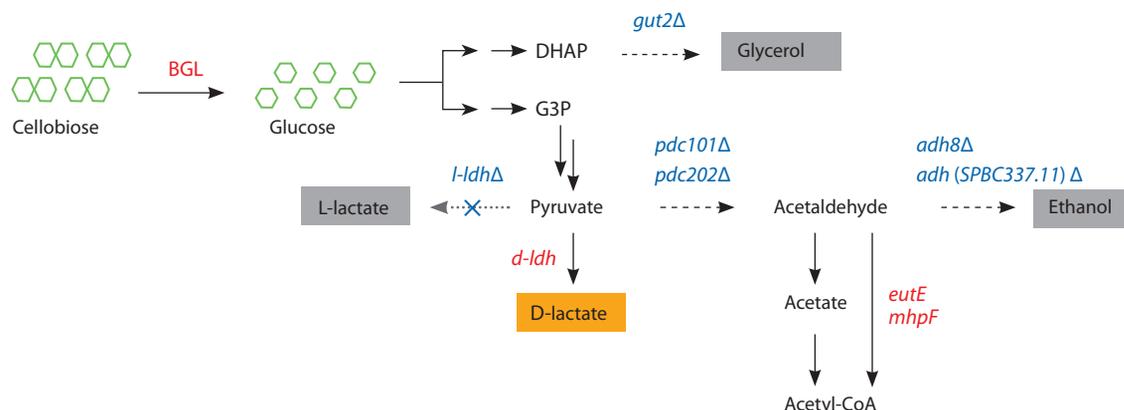


Fig. 4. Changes in the *Schizosaccharomyces pombe* strain introduced for the development of a lactic acid producer according to (Ozaki et al., 2017).

Genes cloned into the genome are shown in red, and knocked-out genes, in blue.

организации технологического процесса (Ozaki et al., 2017). Изменения, которые были внесены в ходе выполнения работы, представлены на рис. 4. В исследовании (Shi et al., 2016) был получен штамм-производитель (R,R)-2,3-бутандиола, способный к коферментации глюкозы и ксилитозы, а также проведено многократное клонирование 24 кб фрагмента, несущего необходимые последовательности.

Аргонавт

Уже на протяжении шести лет система CRISPR/Cas доминирует в качестве инструмента для геномного редактирования эукариот, ее использование весьма эффективно и достаточно просто. Несмотря на это, в мире продолжается поиск других инструментов, которые были бы более удобны или просты в использовании. В 2016 г. опубликовано сообщение об обнаружении белка Аргонавт (Gao et al., 2016), который может существенно дополнить возможности методов геномного редактирования. Подобный белок мог бы оказаться особенно эффективным при работе с одноклеточными организмами или в клеточных культурах. К сожалению, активность белка не подтверждена независимыми исследованиями других лабораторий (Burgess et al., 2016; Javidi-Parsijani et al., 2017; Khin et al., 2017).

Белок Аргонавт впервые упоминается в работе, описывающей мутацию *Arabidopsis thaliana*. Своё название ген и соответствующий белок получили по латинскому названию осьминога *Argonauta argo* – листья мутантного растения свернулись, как щупальца у кальмара (Bohmert et al., 1998). Позднее выяснили, что Аргонавт – это белок, играющий ключевую роль в эукариотической РНК-интерференции. При интерференции Аргонавт использует короткие 5'-фосфорилированные РНК для нацеливания на мРНК. Белки Аргонавт относятся к суперсемейству PIWI (Swarts et al., 2014).

Большое количество геномов прокариот также содержат гены белка Аргонавт (Cerutti et al., 2000; Shabalina, Koonin, 2008; Makarova et al., 2009). В работе (Vogel, 2014) показано, что роль этих белков в клетках прокариот заключается в защите от чужеродной ДНК, попадающей в клетку.

Белки Аргонавт обладают различной специфичностью: часть белков образует комплексы с 5'-фосфорилированными олиго-РНК, другая – с 5'-фосфорилированными олиго-ДНК. Первыми описанными белками Аргонавт прокариот были белки термофильных микроорганизмов, для которых показана способность к внесению разрывов в ДНК *in vitro*. Это были белки из *Thermus thermophilus* (Swarts et al., 2015b) и *Pyrococcus furiosus* (Swarts et al., 2015a). Белки осуществляли защиту клетки от проникновения чужеродной ДНК и проявляли способность к направленному разрушению плазмидных векторов строго в местах, комплементарных направляющим 5'-фосфорилированным ДНК-олигонуклеотидам. В частности, широко используемый в молекулярной биологии для изучения генома термофильных бактерий штамм *Thermus thermophilus* HB27, легко поддающийся трансфекции, несёт поврежденный ген белка Аргонавт (Swarts et al., 2015b). Изучение свойств найденных белков выявило их способность к связыванию с 5'-фосфорилированными олигонуклеотидами и направленному внесению разрывов в нуклеотидную последовательность.

Способность белков Аргонавт к внесению разрывов в полинуклеотидные последовательности по строго определенным сайтам, комплементарным направляющим 5'-фосфорилированным ДНК-олигонуклеотидам (гидДНК), заинтересовала исследователей как новая возможность для создания систем направленного редактирования генома. Однако белки термофильных организмов оказались непригодны для внесения изменений в клетки мезофильных организмов, так как для их работы требовалась температура более 65 °С, что неприемлемо – клетки животных, растений и большинства микроорганизмов не выживают в таких условиях.

В сравнении с CRISPR/Cas, система на основе белка Аргонавт может обладать рядом преимуществ, особенно для одноклеточных организмов, таких как дрожжи. Ключевыми для системы модификации являются всего два элемента: сам белок и направляющий 5'-фосфорилированный олигонуклеотид. Направляющая гидДНК для белков Аргонавт значительно проще в сравнении с гидРНК для Cas9. Размер известных белков Аргонавт со-

ставляет 2/3 от размеров белка Cas9, что важно в связи с тем, что у Cas9 очень большой размер и его ген неудобен в работе.

Однако в настоящее время вопрос о возможности использования белка Аргонавт для редактирования генома эукариот вновь оказался открытым. Описание фермента Аргонавт, способного работать в физиологических для животных температурах, естественно, заинтересовало научное сообщество. Сразу несколько групп начали разрабатывать методы геномного редактирования с использованием нового фермента. Неожиданностью стало то, что ни одной группе исследователей не удалось показать его активность (Burgess et al., 2016; Javid-Parsijani et al., 2017; Khin et al., 2017).

Заключение

Инженерия генома дрожжей имеет длинную и успешную историю благодаря наличию собственных механизмов гомологичной рекомбинации. С одной стороны, дрожжи *S. cerevisiae* долгое время были основным модельным объектом при изучении эукариот, что объяснялось относительной простотой устройства и методов культивирования. С другой стороны, это замечательная платформа для получения штаммов для биотехнологического применения. Дрожжи служили моделью и вспомогательным инструментом при разработке технически сложных методов для модификации генома животных и растений, а также были целью работ. В настоящее время наиболее процессивные инструменты для направленного геномного редактирования разработаны именно для дрожжей. В первую очередь это параллельное введение мутаций при помощи CRISPR/Cas, которое реализовано уже сегодня. В обзоре приведены примеры успешного использования этого метода и описаны существующие инструменты.

Поскольку методы генетической инженерии одноклеточных организмов несколько отличаются от принципов работы с многоклеточными, возможно появление новых инструментов, специализирующихся именно на работе с ними. Одним из таких вариантов мог бы стать белок Аргонавт. К сожалению, в настоящее время достоверно показана активность только высокотемпературных вариантов этого белка.

Acknowledgments

This study was supported by the Russian Federation Ministry of Education and Science, agreement 02.G25.31.0172 of December 1, 2015.

Conflict of interest

The authors declare no conflict of interest.

References

Bao Z., Xiao H., Liang J., Zhang L., Xiong X., Sun N., Si T., Zhao H. Homology-integrated CRISPR-Cas (HI-CRISPR) system for one-step multigene disruption in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* 2015;4:585-594. DOI 10.1021/sb500255k.

Barrangou R., Fremaux C., Deveau H., Richards M., Boyaval P., Moineau S., Romero D.A., Horvath P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*. 2007;315(5819):1709-1712. DOI 10.1126/science.1138140.

Bitinaite J., Wah D.A., Aggarwal A.K., Schildkraut I. *FokI* dimerization is required for DNA cleavage. *Proc. Natl. Acad. Sci. USA*. 1998;95:10570-10575. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC27935/>.

Bohmer K., Camus I., Bellini C., Bouchez D., Caboche M., Benning C. *AGO1* defines a novel locus of *Arabidopsis* controlling leaf development. *EMBO J.* 1998;17(1):170-180. DOI 10.1093/emboj/17.1.170.

Bolotin A., Quinquis B., Sorokin A., Ehrlich S.D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*. 2005;151(8):2551-2561. DOI 10.1099/mic.0.28048-0.

Botstein D., Chervitz S.A., Cherry M.J. Yeast as a model organism. *Science*. 1997;277(5330):1259-1260. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3039837/>.

Burgess S., Cheng L., Gu F., Huang J., Huang Z., Lin S., Li J., Li W., Qin W., Sun Y., Songyang Z., Wei W., Wu Q., Wang H., Wang X., Xiong J.W., Xi J., Yang H., Zhou B., Zhang B. Questions about NgAgo. *Protein Cell*. 2016;7(12):913-915. DOI 10.1007/s13238-016-0343-9.

Cen Y., Timmermans B., Souffriau B., Thevelein J.M., Van Dijk P. Comparison of genome engineering using the CRISPR-Cas9 system in *C. glabrata* wild type and *lig4* strains. *Fungal Genet. Biol.* 2017;107:44-50. DOI 10.1016/j.fgb.2017.08.004.

Cerutti L., Mian N., Bateman A. Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the Piwi domain. *Trends Biochem. Sci.* 2000;25:481-482. DOI 10.1016/S0968-0004(00)01641-8.

Cong L., Ran F.A., Cox D., Lin S., Barretto R., Habib N., Hsu P.D., Wu X., Jiang W., Marraffini L.A., Zhang F. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339(6121):819-823. DOI 10.1126/science.1231143.

DiCarlo J.E., Norville J.E., Mali P., Rios X., Aach J., Church G.M. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* 2013;41(7):4336-4343. DOI 10.1093/nar/gkt135.

Durai S., Mani M., Kandavelou K., Wu J., Porteus M.H., Chandrasegaran S. Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. *Nucleic Acids Res.* 2005;33:5978-5990. DOI 10.1093/nar/gki912.

Farzadfard F., Perli S.D., Lu T.K. Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas. *ACS Synth. Biol.* 2013;2:604-613. DOI 10.1021/sb400081r.

Fernandez R., Berro J. Use of a fluoride channel as a new selection marker for fission yeast plasmids and application to fast genome editing with CRISPR/Cas9. *Yeast*. 2016;33:549-557. DOI 10.1002/yea.3178.

Gao F., Shen X.Z., Jiang F., Wu Y., Han Ch. DNA-guided genome editing using the *Natronobacterium gregoryi* Argonaute. *Nat. Biotechnol.* 2016;34(7):768-773. DOI 10.1038/nbt.3547.

Gao Y., Zhao Y. Self-processing of ribozyme-flanked RNAs into guide RNAs *in vitro* and *in vivo* for CRISPR-mediated genome editing. *J. Integr. Plant Biol.* 2014;56:343-349. DOI 10.1111/jipb.12152.

Gasiunas G., Barrangou R., Horvath P., Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. USA*. 2012;109:E2579-E2586. DOI 10.1073/pnas.1208507109.

Generoso W.C., Gottardi M., Oreb M., Boles E. Simplified CRISPR-Cas genome editing for *Saccharomyces cerevisiae*. *J. Microbiol. Methods*. 2016;127:203-205. doi.org/10.1016/j.mimet.2016.06.020.

Godde J.S., Bickerton A. The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. *J. Mol. Evol.* 2006;62(6):718-729. DOI 10.1007/s00239-005-0223-z.

Guo J., Gaj T., Barbas C.F. Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases. *J. Mol. Biol.* 2010;400(1):96-107. DOI 10.1016/j.jmb.2010.04.060.

Haft D.H., Selengut J., Mongodin E.F., Nelson K.E. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/

- Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* 2005;1(6):E60. DOI 10.1371/journal.pcbi.0010060.
- Horvath P., Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. *Science*. 2010;327(5962):167-170. DOI 10.1126/science.1179555.
- Horwitz A.A., Walter J.M., Schubert M.G., Kung S.H., Hawkins K., Platt D.M., Hernday A.D., Mahatdejkul-Meadows T., Szeto W., Chandran S.S., Newman J.D. Efficient multiplexed integration of synergistic alleles and metabolic pathways in yeasts via CRISPR-Cas. *Cell Systems*. 2015;1(1):88-96. doi.org/10.1016/j.cels.2015.02.001.
- Ishino Y., Shinagawa H., Makino K., Amemura M., Nakata A. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* 1987;169(12):5429-5433. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC213968/>.
- Jacobs J.Z., Ciccaglione K.M., Tournier V., Zaratiegui M. Implementation of the CRISPR-Cas9 system in fission yeast. *Nat. Commun.* 2014;5:5344. DOI 10.1038/ncomms6344.
- Jakočiūnas T., Bonde I., Herrgård M., Harrison S.J., Kristensen M., Pedersen L.E., Jensen M.K., Keasling J.D. Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae*. *Metab. Eng.* 2015;28:213-222. DOI 10.1016/j.ymben.2015.01.008.
- Jansen R., Embden J.D., Gaastra W., Schouls L.M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* 2002;43(6):1565-1575. <https://www.ncbi.nlm.nih.gov/pubmed/11952905>.
- Javidi-Parsijani P., Niu G., Davis M., Lu P., Atala A., Lu B. No evidence of genome editing activity from *Natronobacterium gregoryi* Argonaute (NgAgo) in human cell. *PLoS ONE*. 2017;12(5):e0177444. DOI 10.1371/journal.pone.0177444.
- Jensen E.D., Ferreira R., Jakočiūnas T., Arsovska D., Zhang J., Ding L., Smith J.D., David F., Nielsen J., Jensen M.K., Keasling J.D. Transcriptional reprogramming in yeast using dCas9 and combinatorial gRNA strategies. *Microb. Cell Fact.* 2017;16(1):46. DOI 10.1186/s12934-017-0664-2.
- Jensen M.K., Keasling J.D. Recent applications of synthetic biology tools for yeast metabolic engineering. *FEMS Yeast Res.* 2015;15(1):1-10. DOI 10.1111/1567-1364.12185.
- Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337(6096):816-821. DOI 10.1126/science.1225829.
- Joung J.K., Sander J.D. TALENs: a widely applicable technology for targeted genome editing. *Nat. Rev. Mol. Cell Biol.* 2013;14(1):49-55. DOI 10.1038/nrm3486.
- Khin N.C., Lowe J.L., Jensen L.M., Burgio G. No evidence for genome editing in mouse zygotes and HEK293T human cell line using the DNA-guided *Natronobacterium gregoryi* Argonaute (NgAgo). *PLoS ONE*. 2017;12(6):e0178768. DOI 10.1371/journal.pone.0178768.
- Kim Y.G., Cha J., Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. USA*. 1996;93:1156-1160. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC40048/>.
- Kuijpers N.G., Solis-Escalante D., Bosman L., van den Broek M., Pronk J.T., Daran J.-M., Daran-Lapujade P. A versatile, efficient strategy for assembly of multi-fragment expression vectors in *Saccharomyces cerevisiae* using 60 bp synthetic recombination sequences. *Microb. Cell Fact.* 2013;12(1):47. DOI 10.1186/1475-2859-12-47.
- Laughery M.F., Hunter T., Brown A., Hoopes J., Ostbye T., Shumaker T., Wyrick J.J. New vectors for simple and streamlined CRISPR-Cas9 genome editing in *Saccharomyces cerevisiae*. *Yeast*. 2015;32:711-720. DOI 10.1002/yea.3098.
- Lechner A., Brunk E., Keasling J.D. The need for integrated approaches in metabolic engineering. *Cold Spring Harb. Perspect. Biol.* 2016;8(11):a023903. DOI 10.1101/cshperspect.a023903.
- Lee M.E., DeLoache W.C., Cervantes B., Dueber J.E. A highly characterized yeast toolkit for modular, multipart assembly. *ACS Synth. Biol.* 2015;4:975-986. DOI 10.1021/sb500366v.
- Li T., Huang Sh., Zhao X., Wright D.A., Carpenter S., Spalding M.H., Weeks D.P., Yang B. Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. *Nucleic Acids Res.* 2011;39(14):6315-6325. DOI 10.1093/nar/gkr188.
- Liu Z., Liang Y., Ang E.L., Zhao H. A new era of genome integration – simply cut and paste! *ACS Synth. Biol.* 2017;6(4):601-609. DOI 10.1021/acssynbio.6b00331.
- Makarova K.S., Wolf Y.I., Van der Oost J., Koonin E.V. Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biol. Direct*. 2009;4:29. DOI 10.1186/1745-6150-4-29.
- Mali P., Yang L., Esvelt K.M., Aach J., Guell M., DiCarlo J.E., Norville J.E., Church G.M. RNA-guided human genome engineering via Cas9. *Science*. 2013;339:823-826. DOI 10.1126/science.1232033.
- Mans R., van Rossum H.M., Wijsman M., Backx A., Kuijpers N.G., van den Broek M., Daran-Lapujade P., Pronk J.T., van Maris A.J., Daran J.M. CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 2015;15(2):Fov004. DOI 10.1093/femsyr/fov004.
- Mojica F.J., Díez-Villaseñor C., García-Martínez J., Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*. 2009;155(3):733-740. DOI 10.1099/mic.0.023960-0.
- Nambu-Nishida Y., Nishida K., Hasunuma T., Kondo A. Development of a comprehensive set of tools for genome engineering in a cold- and thermo-tolerant *Kluyveromyces marxianus* yeast strain. *Sci. Rep.* 2017;7(1):8993. DOI 10.1038/s41598-017-08356-5.
- Ng H., Dean N. Dramatic improvement of CRISPR/Cas9 editing in *Candida albicans* by increased single guide RNA expression. *mSphere*. 2017;2(2):e00385-16. DOI 10.1128/mSphere.00385-16.
- Nielsen J., Larsson C., van Maris A., Pronk J. Metabolic engineering of yeast for production of fuels and chemicals. *Curr. Opin. Biotechnol.* 2013;24(3):398-404. DOI 10.1016/j.copbio.2013.03.023.
- Norton E.L., Sherwood R.K., Bennett R.J. Development of a CRISPR-Cas9 system for efficient genome editing of *Candida lusitanae*. *mSphere*. 2017;2(3):e00217-17. DOI 10.1128/mSphere.00217-17.
- Numamoto M., Maekawa H., Kaneko Y. Efficient genome editing by CRISPR/Cas9 with a tRNA-sgRNA fusion in the methylotrophic yeast *Opuntia polymorpha*. *J. Biosci. Bioeng.* 2017;124(5):487-492. DOI 10.1016/j.jbiosc.2017.06.001.
- Ostrov N., Landon M., Guell M., Kuznetsov G., Teramoto J., Cervantes N., Zhou M., Singh K., Napolitano M.G., Moosburner M., Shrock E., Pruitt B.W., Conway N., Goodman D.B., Gardner C.L., Tyree G., Gonzales A., Wanner B.L., Norville J.E., Lajoie M.J., Church G.M. Design, synthesis, and testing toward a 57-codon genome. *Science*. 2016;353(6301):819-822. DOI 10.1126/science.aaf3639.
- Ozaki A., Konishi R., Otomo C., Kishida M., Takayama S., Matsu-moto T., Tanaka T., Kondo A. Metabolic engineering of *Schizosaccharomyces pombe* via CRISPR-Cas9 genome editing for lactic acid production from glucose and cellobiose. *Metab. Eng. Commun.* 2017;5:60-67. DOI 10.1016/j.meteno.2017.08.002.
- Paddon C.J., Westfall P.J., Pitera D.J., Benjamin K., Fisher K., McPhee D., Leavell M.D., Tai A., Main A., Eng D., Polichuk D.R., Teoh K.H., Reed D.W., Treynor T., Lenihan J., Jiang H., Fleck M., Bajad S., Dang G., Dengrove D., Diola D., Dorin G., Ellens K.W., Fickes S., Galazzo J., Gaucher S.P., Geistlinger T., Henry R., Hepp M., Horning T., Iqbal T., Kizer L., Lieu B., Melis D., Moss N., Regentin R., Secrest S., Tsuruta H., Vazquez R., Westblade L.F., Xu L., Yu M., Zhang Y., Zhao L., Lievens J., Covello P.S., Keasling J.D., Reiling K.K., Renninger N.S., Newman J.D. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature*. 2013;496(7446):528-532. DOI 10.1038/nature12051.
- Reider Apel A., d'Espaux L., Wehrs M., Sachs D., Li R.A., Tong G.J., Garber M., Nnadi O., Zhuang W., Hillson N.J., Keasling J.D., Mukhopadhyay A. A Cas9-based toolkit to program gene expression in

- Saccharomyces cerevisiae*. Nucleic Acids Res. 2017;45(1):496-508. DOI 10.1093/nar/gkw1023.
- Richard G.-F., Viterbo D., Khanna V., Mosbach V., Castelain L., Dujon B. Highly specific contractions of a single CAG/CTG trinucleotide repeat by TALEN in yeast. PLoS ONE. 2014;9(4):e95611. DOI 10.1371/journal.pone.0095611.
- Ryan O.W., Cate J.H. Multiplex engineering of industrial yeast genomes using CRISPRm. Methods Enzymol. 2014;546:473-489. DOI 10.1016/B978-0-12-801185-0.00023-4.
- Ryan O.W., Skerker J.M., Maurer M.J., Li X., Tsai J.C., Poddar S., Lee M.E., DeLoache W., Dueber J.E., Arkin A.P., Cate J.H.D. Selection of chromosomal DNA libraries using a multiplex CRISPR system. eLife. 2014;3:e03703. DOI 10.7554/eLife.03703.
- Sanjana N.E., Cong L., Zhou Y., Cunniff M.M., Feng G., Zhang F. A transcription activator-like effector toolbox for genome engineering. Nat. Protocols. 2012;7(1):171-192. DOI 10.1038/nprot.2011.431.
- Schwartz C.M., Hussain M.S., Blenner M., Wheeldon I. Synthetic RNA polymerase III promoters facilitate high-efficiency CRISPR-Cas9-mediated genome editing in *Yarrowia lipolytica*. ACS Synth. Biol. 2016;5(4):356-359. DOI 10.1021/acssynbio.5b00162.
- Shabalina S.A., Koonin E.V. Origins and evolution of eukaryotic RNA interference. Trends Ecol. Evol. 2008;23:578-587. DOI 10.1016/j.tree.2008.06.005.
- Sheng J., Flick H., Feng X. Systematic optimization of protein secretory pathways in *Saccharomyces cerevisiae* to increase expression of Hepatitis B Small Antigen. Front. Microbiol. 2017;8:875. DOI 10.3389/fmicb.2017.00875.
- Shi S., Liang Y., Zhang M.M., Ang E.L., Zhao H. A highly efficient single-step, markerless strategy for multi-copy chromosomal integration of large biochemical pathways in *Saccharomyces cerevisiae*. Metab. Eng. 2016;33:19-27. DOI 10.1016/j.ymben.2015.10.011.
- Stovicek V., Borodina I., Forster J. CRISPR-Cas system enables fast and simple genome editing of industrial *Saccharomyces cerevisiae* strains. Metab. Eng. Commun. 2015;2:13-22. doi.org/10.1016/j.meteno.2015.03.001.
- Swarts D.C., Hegge J.W., Hinojo I., Shiimori M., Ellis M.A., Dumrongkulraksa J., Terns R.M., Terns M.P., van der Oost J. Argonaute of the archaeon *Pyrococcus furiosus* is a DNA-guided nuclease that targets cognate DNA. Nucleic Acids Res. 2015a;43(10):5120-5129. DOI 10.1093/nar/gkv415.
- Swarts D.C., Koehorst J.J., Westra E.R., Schaap P.J., van der Oost J. Effects of Argonaute on gene expression in *Thermus thermophilus*. PLoS ONE. 2015b;10(4):e0124880. DOI 10.1371/journal.pone.0124880.
- Swarts D.C., Makarova K., Wang Y., Nakanishi K., Ketting R.F., Koonin E.V., Patel D.J., van der Oost J. The evolutionary journey of Argonaute proteins. Nat. Struct. Mol. Biol. 2014;21(9):743-753. DOI 10.1038/nsmb.2879.
- Vanegas K.G., Lehka B.J., Mortensen U.H. SWITCH: a dynamic CRISPR tool for genome engineering and metabolic pathway control for cell factory construction in *Saccharomyces cerevisiae*. Microb. Cell Fact. 2017;16(1):25. DOI 10.1186/s12934-017-0632-x.
- Vogel J. A bacterial seek-and-destroy system for foreign DNA. Science. 2014;344:972-973. DOI 10.1126/science.1252962.
- Vyas V.K., Barrasa M.I., Fink G.R. A *Candida albicans* CRISPR system permits genetic engineering of essential genes and gene families. Sci. Adv. 2015;1(3):e1500248. DOI 10.1126/sciadv.1500248.
- Walter J.M., Chandran S.S., Horwitz A.A. CRISPR-cas-assisted multiplexing (CAM): simple same-day multi-locus engineering in yeast. J. Cell Physiol. 2016;231:2563-2569. DOI 10.1002/jcp.25375.
- Wang L., Lin J., Zhang T., Xu K., Ren Ch., Zhang Zh. Simultaneous screening and validation of effective zinc finger nucleases in yeast. PLoS ONE. 2013;8(5):e64687. DOI 10.1371/journal.pone.0064687.
- Wang Y., Wei D., Zhu X., Pan J., Zhang P., Huo L., Zhu X. A 'suicide' CRISPR-Cas9 system to promote gene deletion and restoration by electroporation in *Cryptococcus neoformans*. Sci. Rep. 2016;6:31145. DOI 10.1038/srep31145.
- Weninger A., Hatzl A.-M., Schmid C., Vogl T., Glieder A. Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast *Pichia pastoris*. J. Biotechnol. 2016;235:139-149. DOI 10.1016/j.jbiotec.2016.03.027.
- Wriessnegger T., Pichler H. Yeast metabolic engineering—targeting sterol metabolism and terpenoid formation. Progr. Lipid Res. 2013;52(3):277-293. DOI 10.1016/j.plipres.2013.03.001.
- Zhang G.C., Kong I.I., Kim H., Liu J.J., Cate J.H., Jin Y.S. Construction of a quadruple auxotrophic mutant of an industrial polyploidy *Saccharomyces cerevisiae* using RNA-guided Cas9 nuclease. Appl. Environ. Microb. 2014;80:7694-7701. DOI 10.1128/AEM.02310-14.



Растительные системы экспрессии в качестве продуцентов рекомбинантных фармацевтически ценных белков

Е.В. Дейнеко , А.А. Загорская

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

Рынок фармацевтически ценных белков – наиболее быстро развивающийся сегмент экономики. Большая часть биофармацевтиков получена в клетках млекопитающих и микроорганизмов, однако обе системы обладают рядом недостатков. Растительные клетки сочетают в себе достоинства эукариотической системы наработки белка и простоту и дешевизну бактериальной. Использование растений для получения рекомбинантных белков – экономически значимое и перспективное направление. Преимуществом растительных систем является более низкая стоимость культивирования клеток. Они свободны от нежелательных компонентов, таких как эндотоксины бактерий, гипергликозилированные белки, продуцируемые дрожжами, патогены животных и человека в клеточных культурах трансгенных животных. Растения относятся к высшим эукариотам, поэтому в их клетках происходит полноценный фолдинг и образование сложных мультимерных белковых комплексов, а также значительная часть посттрансляционных модификаций аналогично таковым в клетках млекопитающих. Развиваемые ныне растительные системы экспрессии рекомбинантных белков чрезвычайно разнообразны и насчитывают более 100 различных технологий, основанных на разных видах растений, способах переноса генов, экспрессионных стратегиях, методах последующего извлечения целевого белка и пр. К ним относятся ядерная и пластидная трансформация, транзистная и стабильная экспрессия при трансформации с помощью агробактериального переноса, бомбардировки или электропорации, культивирование целых наземных или водных растений, растительных тканей или суспензионных клеточных культур в качестве экспрессионных систем. В обзоре анализируется современное состояние исследований в области использования растительных систем экспрессии для наработки рекомбинантных фармацевтических белков. Сделан акцент на преимуществах культур растительных клеток по сравнению с другими системами экспрессии. Описаны растительные системы для наработки рекомбинантных белков, такие как транспластомные растения, культуры мхов и водных растений, а также суспензионные культуры клеток высших растений. Рассмотрено современное состояние рынка рекомбинантных белков, полученных с применением растительных систем экспрессии. Обсуждаются перспективы растительных («съедобных») вакцин, созданных на основе генетически модифицированных растений.

Ключевые слова: системы экспрессии; трансгенные растения; биопродуценты; рекомбинантные белки; растительные вакцины.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Дейнеко Е.В., Загорская А.А. Растительные системы экспрессии в качестве продуцентов рекомбинантных фармацевтически ценных белков. Вавиловский журнал генетики и селекции. 2017;21(8):979-985. DOI 10.18699/VJ17.322

HOW TO CITE THIS ARTICLE:

Deineko E.V., Zagorskaya A.A. Plant expression systems for production of recombinant pharmaceutically important proteins. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2017;21(8):979-985. DOI 10.18699/VJ17.322 (in Russian)

Received 08.11.2017

Accepted for publication 30.11.2017

© AUTHORS, 2017

Plant expression systems for production of recombinant pharmaceutically important proteins

E.V. Deineko , A.A. Zagorskaya

Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

The market of pharmaceutically valuable proteins is the fastest growing segment of the economy. Most biopharmaceuticals have been obtained in mammalian and microorganism cells, but both systems have a number of disadvantages. Plant cells combine the advantages of the eukaryotic system of protein production and the simplicity and cheapness of the bacterial, and the use of plants for the production of recombinant proteins is an economically important and promising direction. The advantage of plant systems is the lower cost of cell cultivation. They are free from unwanted components, such as bacterial endotoxins, hyperglycosylated proteins produced by yeast, animal and human pathogens in cell cultures of transgenic animals. In addition, plants are higher eukaryotes, and therefore full-value folding and the formation of multimeric protein complexes occur in their cells, as well as a significant portion of post-translational modifications similar to those in mammalian cells. The currently developed plant expression systems for recombinant proteins are extremely diverse and number more than 100 different technologies based on different plant species, gene transfer methods, expression strategies, methods for the subsequent extraction of the target protein, etc. This is nuclear and plastid transformation, transient and stable expression during transformation using agrobacterial transport, bombardment or electroporation, cultivation of whole terrestrial or aquatic plants, plant tissues or suspension cell cultures as expression systems. The review examines the current state of research in the use of plant expression systems for the production of recombinant proteins for pharmaceuticals. The emphasis was placed on the advantages of plant cell cultures in comparison with other expression systems. Specific examples discuss promising plant systems for the production of recombinant proteins, such as transplastomic plants, moss and aquatic plant cultures, as well as suspension cultures of cells of higher plants. The current state of the market for recombinant proteins obtained using plant expression systems is considered. The prospects of creating plant ("edible") vaccines based on genetically modified plants are discussed.

Key words: expression systems; transgenic plants; bioproducers; recombinant proteins; plant vaccines.

В настоящее время большое число белков медицинского назначения получают не из природных источников, а при синтезе их рекомбинантных аналогов. Для этих целей используются различные системы экспрессии: бактериальные (*Escherichia coli*), дрожжевые (*Saccharomyces cerevisiae*, *Pichia pastoris* и *Hansenula polymorpha*), клетки животных (клетки яичников китайского хомячка) и др. Это позволяет нарабатывать в промышленных масштабах белки, которые было бы невозможно в достаточном количестве получить традиционными методами экстракции (например, инсулин или гормон роста человека). Получение рекомбинантных белков основано на технологии рекомбинантных ДНК, включающей клонирование целевого гена, кодирующего фармацевтически ценный (целевой) белок, и его перенос в геном клеток, в которых данный целевой белок будет синтезирован.

Для бактериальных систем экспрессии характерны относительная простота генно-инженерных манипуляций, высокая скорость деления клеток и, соответственно, значительный выход биомассы, высокий уровень экспрессии рекомбинантных белков и возможность масштабного производства белка в биореакторах. Однако прокариотическая клетка не в состоянии осуществлять многие посттрансляционные модификации синтезируемого рекомбинантного белка, включая прежде всего его гликозилирование и корректное образование дисульфидных связей. Следует отметить, что некорректные посттрансляционные модификации или их отсутствие могут сильно повлиять на свойства рекомбинантного белка, в том числе на его биологическую активность и фармакокинетику. Таким образом, прокариотические системы экспрессии используются для синтеза относительно простых терапевтических белков, таких как инсулин, интерферон или гормон роста человека. Немаловажным является и этап выделения рекомбинантного белка, синтезируемого в прокариотической системе экспрессии, поскольку именно этот этап связан с очисткой целевого продукта от пирогенных эндотоксинов и других продуктов метаболизма бактериальных клеток в культуре. В среднем выход рекомбинантного белка в бактериальной системе экспрессии составляет от 20 до 400 мг/л культуральной среды.

Наиболее приближенно к человеческому типу посттрансляционные модификации рекомбинантного белка осуществляются в системах экспрессии, основанных на культурах клеток млекопитающих (Casteleijn, Richardson, 2014), но и эти системы не лишены недостатков. К ним относятся высокая стоимость культивирования, трудности с масштабированием процесса и потенциальная заражаемость клеточной культуры патогенами человека и животных. Использование дрожжей для наработки рекомбинантных белков часто сопровождается некорректными посттрансляционными модификациями, связанными с образованием гипергликозилированных рекомбинантных белков.

В последние 10–15 лет интенсивно возрастает интерес к альтернативным системам экспрессии, в частности к растительным клеткам, культивируемым в биореакторах. Растительные суспензионные клеточные культуры лишены вышеперечисленных недостатков и чрезвычайно

перспективны для этих целей, сочетая достоинства эукариотических систем экспрессии и простоту и дешевизну бактериальных.

Рынок фармацевтически ценных белков растет быстрее, чем фармацевтический рынок в целом, и по прогнозам специалистов будет достигать к 2020-му году 278.2 млрд долл. США (Casteleijn, Richardson, 2014). Крупнейшие биотехнологические и фармацевтические компании проявляют большую заинтересованность и инвестируют значительные средства в развитие научных исследований по разработке новых платформ для производства рекомбинантных белков и внедрения их в производство. Весьма привлекательными для исследователей и фармацевтических компаний становятся растительные системы экспрессии. Преимущество этих систем – относительно невысокая стоимость их культивирования. Они свободны от нежелательных компонентов, таких как эндотоксины бактерий или гипергликозилированные целевые белки, продуцируемые дрожжами, и в отличие от клеточных культур животного происхождения – от патогенов животных и человека. В клетках растений происходит корректный фолдинг и образование сложных мультимерных белковых комплексов, а также большая часть посттрансляционных модификаций целевых белков, необходимых для их биологической активности (Twyman et al., 2003; Nagels et al., 2012). В растительных клетках могут синтезироваться такие сложные белки млекопитающих, как коллагены, гемоглобин, иммуноглобулины. В ближайшем будущем получение рекомбинантных белков в культурах клеток высших растений, вероятнее всего, станет наиболее часто используемой из всех ныне применяемых растительных систем экспрессии.

Рекомбинантные белки в растительных системах могут синтезироваться при использовании различных платформ, основанных на разных видах растений, способах переноса генов, экспрессионных стратегиях, методах последующего извлечения целевого белка и пр. Это ядерная и пластидная трансформация, транзientная (Gleba et al., 2007; Huang, McDonald, 2009) и стабильная экспрессия при трансформации с помощью агробактериального переноса (Gelvin, 2003), бомбардировки или электропорации (Rosales-Mendoza, Tello-Olea, 2015), культивирование целых наземных или водных растений, растительных тканей или суспензионных клеточных культур в качестве экспрессионных систем.

Синтезированные рекомбинантные белки могут быть направлены в различные компартменты растительной клетки (вакуоли или люмены эндоплазматического ретикулума), а также в апопласт и различные органы растения (семена, клубни, плоды и т. д.). Благодаря этому рекомбинантные белки в растительных тканях могут быть длительное время сохранены без каких-либо изменений и снижения биологической активности (Daniell et al., 2005; Gleba et al., 2005).

Немаловажен и тот факт, что разработанные к настоящему времени методы агроботанического возделывания хозяйственно важных видов растений, а также системы семеноводства для той или иной культуры делают растения привлекательными для использования их в качестве биофабрик белков медицинского назначения.

Перспективные растительные системы экспрессии для наработки рекомбинантных белков

К настоящему времени разработаны технологии получения генетически модифицированных (трансгенных) растений, в геном которых перенесены гены, кодирующие различные белки для медицинских целей, в том числе и белки человека. Более 200 биофармацевтиков уже представлены на рынке, и еще большее их количество проходит доклинические испытания. Крупнейшие биотехнологические и фармацевтические компании, такие как Elicyte, Ventria, Medicago, Greenovation, LSBC и Pfizer, проявляют огромную заинтересованность и инвестируют в развитие научных исследований по разработке новых платформ для получения фармацевтически ценных белков и внедрения их в производство. Данные технологии основаны на прямом или векторном переносе целевых генов в ядерный геном растения. Получены десятки видов трансгенных растений, в геном которых перенесены последовательности антигенов различных возбудителей инфекционных заболеваний, разнообразных терапевтических белков, моноклональных антител (Howard, 2005; Tekoah et al., 2015). Однако использование растительных систем для производства рекомбинантных белков различного назначения сдерживается частично из-за недостаточно высокого уровня их накопления, который составляет, как правило, не более 1 % общего растворимого белка (ОРБ) (Пермякова и др., 2015).

В качестве альтернативных систем экспрессии рассматриваются внеядерные геномы (пластомы) хлоропластов растений. Технология трансформирования хлоропластов разрабатывалась как многообещающий подход в получении рекомбинантных белков. При трансформации пластида не наблюдается проблем, характерных для ядерной трансформации, таких как замолкание генов, эпигенетические эффекты или вариабельность экспрессии трансгенов. Высокий уровень экспрессии генов, перенесенных в хлоропластный геном, достигается высокой копийностью пластидной ДНК, а экологическая безопасность обеспечивается отсутствием чужеродных генов в пыльце. Однако необходимо подчеркнуть, что эти преимущества нивелируются методическими сложностями, связанными с переносом генов в пластомы растений, а также тем фактом, что транспластомные системы перспективны лишь в случае, когда целевой белок не подвергается сложным посттрансляционным модификациям.

Мощным толчком к развитию технологии получения рекомбинантных белков на основе хлоропластного генома послужило сообщение о создании транспластомных растений табака с выходом целевого белка (Cry2Aa2-белок из *Bacillus thuringiensis*) на уровне 46.1 % ОРБ (DeCosa et al., 2001). В последнее время интенсивно развиваются технологии создания транспластомных растений многих видов: сои (Dufourmantel et al., 2004), хлопка (Daniell, 2007), салата (Lelivelt et al., 2005) и др. Получено более 20 видов транспластомных растений, уровень накопления гетерологичных белков у которых составил (в расчете от ОРБ) от 6 % для интерферона-гамма и 19 % для интерферона-альфа человека до 33 % для инсулин-подобного фактора роста (IGF-1) человека.

Среди других альтернативных систем экспрессии рекомбинантных белков для фармакологии следует назвать культуры водных растений: водоросли, микроводоросли и ряска, обладающие высокими темпами роста на простых по составу и дешевых питательных средах или просто на воде, однако требующие освещения.

Ряска (*Lemna gibba*, ряска горбатая и *L. minor*, ряска малая) – многолетнее растение, относящееся к однодольным покрытосеменным растениям, которые обитают на поверхности стоячих пресных вод. Ряска привлекает внимание исследователей как потенциальная высокоэффективная система экспрессии гетерологичных белков благодаря своей способности к быстрому накоплению биомассы. Завершены работы по секвенированию хлоропластного генома ряски; для переноса в ее геном целевых генов разработаны методы агробактериальной трансформации и биобаллистики. Показано, что моноклональные антитела человека, синтезированные в тканях ряски, проявляют более высокую активность при связывании с соответствующими рецепторами по сравнению с их гомологами, полученными в клетках СНО-линии клеток яичников китайских хомяков (Cox et al., 2006). При наличии в генетической конструкции соответствующих сигнальных последовательностей рекомбинантный белок может быть секретирован в культуральную среду (Rosales-Mendoza, 2016).

Одноклеточные зеленые водоросли *Chlamydomonas reinhardtii*, *Phaeodactylum tricorutum*, *Tetraselmis suecica* и *Odontella aurita* рассматриваются как альтернативные системы экспрессии рекомбинантных белков (Taunt et al., 2017). Наиболее перспективна среди этих водорослей *C. reinhardtii*, получение рекомбинантных белков в которой основано на трансформации хлоропластов (Almaraz-Delgado et al., 2014). Около 40 % объема клетки этой водоросли занимает один хлоропласт большого размера. Первые генетически модифицированные водоросли на основе *C. reinhardtii* получены в 1988 г. (Boynnton et al., 1988). Интеграция фрагментов экзогенных ДНК в геном этой водоросли протекает по механизму гомологичной рекомбинации. Клеточные популяции водорослей гомогенны по размеру клеток, клеточная масса удваивается через каждые 4–8 ч культивирования (Franklin, Mayfield, 2004). Для культивирования *C. reinhardtii* возможно использование биореакторов больших объемов (до 500 тыс. л). Более того, способность секретировать белки в культуральную среду может существенно снизить стоимость рекомбинантного белка (Almaraz-Delgado et al., 2014).

Зеленый мох *Physcomitrella patens* – единственный представитель мохообразных, геном которого полностью секвенирован (Reinsing et al., 2008). Разработана методика трансформации *P. patens*. Первым успешным примером использования этого вида мха в качестве системы экспрессии рекомбинантных белков стало получение рекомбинантного эритропоэтина человека (Weise et al., 2007). Фирма Greenovation (Германия) разрабатывает технологию наработки биофармацевтических белков на основе *P. patens* в клеточной суспензионной культуре с применением биореакторов. Привлекательность этого вида мхов в качестве системы экспрессии рекомбинантных белков состоит в том, что фрагменты экзогенных ДНК могут быть

интегрированы в его геном по механизму гомологичной рекомбинации, что исключает возможное инактивирование в дальнейшем экспрессии перенесенных генов (трансгенов). *P. patens* рассматривается как перспективный кандидат на «молекулярное фермерство», поскольку в этой системе экспрессии белки эукариотического происхождения претерпевают посттрансляционную модификацию (гликозилирование, образование дисульфидных связей и т. д.). Если проводить сравнение белкового продукта, синтезируемого в клетках животных и в клетках *P. patens*, то последние имеют очевидное преимущество, поскольку отсутствует риск заражения культуры патогенами животного происхождения. К тому же клетки *P. patens* можно поддерживать в виде суспензионной культуры в контролируемых условиях (Reski et al., 2015). Однако большая часть уже существующих терапевтических белков получена с помощью стабильной ядерной трансформации с последующим выделением и их очисткой из трансгенных растений-регенерантов (Desai et al., 2010). В качестве вектора, как правило, используется *Agrobacterium tumefaciens*, способная трансформировать широкий спектр двудольных и однодольных растений. В растениях кукурузы, риса и ячменя синтезируются некоторые технические реагенты для диагностики (Howard, 2005; Fischer et al., 2012), а также ряд фармацевтических продуктов, находящихся на этапе клинических испытаний. Несмотря на очевидные преимущества использования растительных систем, внедрение новых методических разработок в производство происходит довольно медленно, и только немногие рекомбинантные белки, полученные из целых трансгенных растений, достигли рынка. Одна из причин этого – негативное общественное мнение, сложившееся вокруг генно-модифицированных растений, а также несовершенство законодательной базы, на основании которой такие продукты могут попасть к потребителю.

Суспензионные клеточные культуры

Перспективным направлением, позволяющим преодолеть существенные недостатки использования трансгенных растений в качестве биофабрик, представляется культивирование растительных клеток в ферментерах. Возможность получения рекомбинантных белков в суспензиях растительных клеток показана более 25 лет назад, однако долгое время исследования были сосредоточены на использовании для синтеза целых растений. Отношение к использованию суспензий растительных клеток начало меняться после всплеска коммерческого интереса к получению рекомбинантных белков, столкнувшегося с отсутствием регулирующих законов и настороженностью по отношению к ГМО, особенно в Европе. Технология культивирования растительных клеток обеспечивает точное соблюдение условий выращивания клеток в отличие от полевых условий, где существует погодная, климатическая, почвенная составляющие, а также влияние вредителей, травоядных животных и различных микроорганизмов (Rybicki, 2010; Fischer et al., 2012). Выращивание суспензионных культур клеток в стерильных реакторах не только элиминирует риск заражения культуры клеток микотоксинами и пестицидами (Hellwig et al., 2004; Fischer et al., 2012), но также снижает до минимума возможность

переноса генетически модифицированных клеток в окружающую среду.

Неоспоримым преимуществом системы культивирования суспензионных культур клеток по сравнению с выращиванием целых растений является их быстрый рост. В экспоненциальной фазе роста для удвоения клеток необходимо 2–3 суток, один цикл культивирования клеток ВУ-2 табака, например, составляет 1–2 недели, в то время как на получение растений уходит несколько месяцев (Kaldis et al., 2013).

Существенное достоинство культуры растительных клеток как экспрессионной системы – их способность продуцировать и секретировать биологически активные белки через мембрану и клеточную стенку в межклеточное пространство. Этот процесс метаболически зависим и может обеспечиваться специфическими лидерными последовательностями пептидов как растительного, так и животного происхождения (Magnuson et al., 1998). Благодаря накоплению рекомбинантных белков в среде при культивировании суспензий растительных клеток упрощается процесс их извлечения и очистки. Переработка растительных тканей и целых растений связана с трудоемкой и затратной процедурой экстракции, а выделение из среды облегчается в связи с возможностью быстрого отделения клеток, в которых содержится большее количество сопутствующих белков. Целевые белки, секретируемые в культуральную среду, характеризуются также высокой степенью целостности и однородностью, так как их транспорт из клетки происходит после полного процессирования, т. е. после полного удаления сигнальных пептидов и присоединения гликановых структур, если конечный продукт является гликопротеином. В том случае, если рекомбинантные белки накапливаются в больших количествах внутри клеток, их получение связано с очисткой от непротессированных, незрелых и сигнальных белков и гетерологичных гликанов. По неопубликованным данным (цит. по: Schillberg et al., 2013), антитела, извлеченные из культуральной среды при выращивании клеток табака ВУ-2, представляют собой гомогенный комплекс из трех гликоформ с доминирующей формой, представленной 87 %, тогда как эти же антитела синтезируются в целых растениях в виде шести различных гликановых форм.

Перспективность суспензионных культур растительных клеток в качестве платформы для получения рекомбинантных белков стала очевидной после сообщения крупнейшей фармацевтической компанией США (Pfizer) о выпуске на рынок α -галиглюцефразы, синтезированной клетками моркови в биореакторах. Именно после внедрения производства α -галиглюцефразы в клетках моркови был открыт путь к полному принятию этой технологии и пересмотру главенствующей поначалу идеи использования целых растений. К настоящему времени созданы и успешно используются для производства биофармацевтиков клеточные линии табака ВУ-2 (Bright Yellow-2) (Nagels et al., 2012), риса (Schiermeyer, Schillberg, 2012), люцерны (Huang, McDonald, 2009), моркови (Rosales-Mendoza, Tello-Olea, 2015) и др. Эти клеточные культуры отличаются высокими темпами роста и восприимчивостью к трансформации с помощью *A. tumefaciens*.

Несмотря на успешность использования суспензионных клеточных культур растений для коммерческого получения фармацевтических белков, существует еще много нерешенных проблем, наиболее важной из которых является недостаточно высокий выход рекомбинантного белка. Именно на этих проблемах сосредоточено внимание исследовательских групп и биотехнологических компаний.

Современное состояние рынка рекомбинантных белков медицинского назначения, полученных на основе генетически модифицированных растений

В последние годы около 15 зарубежных фирм связывают свою деятельность с наработкой фармацевтических белков медицинского назначения в генетически модифицированных растениях (Tiwari et al., 2009; Martinez et al., 2012). Среди них следует отметить Planet Biotechnology, Dow AgroSciences, Protalix Biotherapeutics, Medicago, Biolex, Novoplant и др. Деятельность этих фирм основана на использовании таких растений, как табак, люцерна, рис, подсолнечник, ячмень, горох, кукуруза и *Arabidopsis thaliana*. Необходимо отметить, что в большей части работ по получению рекомбинантных белков используются генетически модифицированные растения с ядерной трансформацией, т. е. доставкой чужеродного гена в ядерный геном растения. Фирмы Chlorogen и Bayer для получения рекомбинантных белков используют транспластомные растения, а также метод агроинфильтрации, базирующийся на транзитной (временной) экспрессии чужеродных генов в растительных клетках. Часто для одних и тех же ценных фармацевтических белков фирмами разрабатываются различные оригинальные методики получения. Ярким примером может служить производство апротинина – поливалентного ингибитора протеиназ, относящегося к антиферментным препаратам. Этот белок применяется в качестве препарата, оказывающего антипротеолитическое, антифибринолитическое и гемостатическое действие и используется в медицине уже более 40 лет. Источником апротинина являются органы (легкие и др.) крупного рогатого скота. Разработан способ получения рекомбинантного апротинина с использованием *S. cerevisiae* (Apeler et al., 2004). Путем сравнительного анализа систем экспрессии на основе растительных клеток и тканей для наработки апротинина установлено, что в тканях табака при ядерной трансформации накапливалось 0.03 % этого белка от ОБР, в тканях люцерны – 0.1, в семенах кукурузы – 8.9, в тканях ряски – 3.7, а в тканях табака при агроинфильтрации – 4.2 % (Sourrouille et al., 2009). Рекомбинантный апротинин не отличался по качеству от соответствующего белка, выделенного из тканей крупного рогатого скота (Sourrouille et al., 2009). Такие фирмы, как ProdiGene, Medicago и Large Scale Biology Corporation используют кукурузу, люцерну и табак для получения рекомбинантного апротинина для коммерческих целей.

Перспективы создания растительных («съедобных») вакцин на основе генетически модифицированных растений

Идея использования клеток растений для наработки рекомбинантных антигенов впервые успешно реализована в

1992 г. группой исследователей под руководством Ч. Арнтзена (Mason et al., 1992). В их работе установлено, что поверхностный HBsAg-антиген вируса гепатита В не только накапливается в тканях трансгенных растений табака, но и способен к самосборке в вирусоподобные частицы размером около 22 нм. Такие частицы были идентичны рекомбинантным вирусоподобным частицам HBsAg-антигена, выделенным из промышленной рекомбинантной вакцины на основе дрожжей, а также вирусоподобным частицам из плазмы крови больных вирусом гепатита В. На основании полученных данных стало очевидным, что чужеродные белки способны синтезироваться в клетках трансгенных растений в их природной иммунологически активной форме. Это открывало новые возможности использования растений как более дешевых систем экспрессии для создания рекомбинантных вакцин.

Следующим принципиальным шагом в разработке концепции «съедобных» вакцин на основе генетически модифицированных растений были работы по созданию трансгенных растений картофеля, продуцирующих термолabileный энтеротоксин из *E. coli* (Haq et al., 1995; Mason et al., 1998) и В-субъединицу холерного токсина (Arakawa et al., 1998).

Термолabileный энтеротоксин *E. coli* состоит из двух частей: LT-A (фермент) и LT-B (пентамер из рецептор-связывающих полипептидов). LT-B взаимодействует с рецепторами ганглиозидов на поверхности мембраны эпителиоцитов тонкого кишечника млекопитающих и транспортирует LT-A в клетки кишечника. В эпителиоцитах LT-A вызывает изменение клеточного метаболизма и обезвоживание клеток. Если обе части термолabileного энтеротоксина отделить друг от друга, то презентация LT-B белкового комплекса на поверхности эпителиоцитов будет стимулировать сильный иммунный ответ слизистой оболочки кишечника без проявления каких-либо признаков заболевания. Именно эта особенность положена в основу исследований группы Ч. Арнтзена (Haq et al., 1995) по созданию «съедобных» вакцин. В их работе установлено, что LT-B, синтезируемый в трансгенных растениях табака и картофеля, а также LT-B, выделенный из *E. coli*, вызывают однотипные иммунные реакции у мышей.

В дальнейшем LT-B-последовательность была оптимизирована для экспрессии в растительных клетках и перенесена в геном растений картофеля (Mason et al., 1998). В клубнях картофеля белок корректно собирался в олигомеры и накапливался в достаточно больших количествах. На основании клинических испытаний рекомбинантной LT-B-вакцины показано, что поедание добровольцами сырых клубней картофеля, содержащих 0.3–10 мг LT-B, приводило к образованию мукозных и системных антител с высокими титрами (Tacket et al., 2000).

Опираясь на проведенные исследования, можно заключить, что «съедобные» вакцины на основе трансгенных растений способны вызывать защитный иммунитет и открывают новые возможности на пути создания недорогих и простых в обращении вакцин против инфекционных болезней животных и человека. По данным (Yusibov, Rabin-dran, 2008; Yusibov et al., 2011), около шести антигенов и антител (для пассивной иммунизации), синтезированных

в генетически модифицированных растениях, достигли первой и второй фаз клинических испытаний.

Механизм иммунизации «съедобными» вакцинами основан на антигенпредставляющей способности перитонеальных макрофагов тонкого кишечника млекопитающих. В кишечнике чужеродный белок, обладающий антигенными свойствами, распознается специальными М-клетками, которые широко представлены в толще слизистого эпителия. М-клетки транспортируют захваченный антиген к перитонеальным макрофагам и В-лимфоцитам, находящимся в лимфоидных образованиях тонкого кишечника (пейеровых бляшках). В результате презентации антигена на поверхности антиген-представляющих клеток происходит активация Т-лимфоцитов-хелперов, которые в сочетании с антигеном активируют В-лимфоциты. Дифференцированные В-клетки выходят из лимфоидных фолликулов слизистой оболочки и поступают через общую циркуляцию в мезентральные лимфатические узлы, где происходит их созревание и превращение в плазматические клетки, синтезирующие специфические антитела. Плазматические клетки способны снова мигрировать к слизистым оболочкам дыхательных путей, желудочно-кишечного и мочеполового трактов. Секреторные иммуноглобулины IgA транспортируются на поверхность слизистых оболочек, где они связываются с чужеродными антигенами и препятствуют их проникновению в организм. Следует отметить, что мукозная вакцинация стимулирует как иммунный ответ слизистых оболочек – первого защитного барьера на пути патогенных агентов, так и общий иммунный ответ организма. Описана работа по изучению иммуногенности «съедобной» вакцины на основе трансгенных растений томатов, экспрессирующих HBsAg-антиген вируса гепатита В (Schelkunov et al., 2006). Показано, что при пероральном введении мышам гомогената плодов трансгенных растений томатов уровень антител в сыворотке крови возрастал после второго кормления и оставался высоким до конца эксперимента. Мукозный иммунный ответ формировался после первого кормления и оставался высоким в течение всего эксперимента.

К настоящему времени созданы трансгенные растения на основе многих видов, таких как табак, томаты, салат-латук, *A. thaliana* и турнепс, в которые перенесены гены, контролирующие синтез различных антигенов и антител. В некоторых случаях антигены «сшиваются» с другими белками для упрощения процедуры детекции продуктов целевого гена, например с геном β -глюкуронидазы (*uidA*). По активности фермента β -глюкуронидазы можно легко определить уровень накопления целевого антигена в тканях генетически модифицированного растения.

Заклучение

Растения используются человечеством для медицинских целей уже многие тысячи лет. Однако только в начале 21-го века с помощью методов генетической инженерии стало возможным создавать новые типы растений, в тканях которых могут синтезироваться и накапливаться белки из различных гетерологичных систем. Созданы трансгенные растения, в ядерный и хлоропластный геномы которых перенесены гены, контролирующие синтез белков, важных в терапии различных заболеваний. Благодаря

развитию методов генетической инженерии за последние 20 лет открылись перспективы в области создания вакцин нового поколения. Возможность изолировать гены патогенных вирусов и бактерий, а также манипулировать ими и переносить в геномы других организмов (например, в геномы растений) привела к созданию генетически модифицированных организмов, у которых накапливаемый в тканях белок-антиген служит лишь «опознавательным знаком» патогена и не способен вызывать болезнь. В настоящее время более десятка зарубежных фирм используют растительные системы экспрессии для наработки рекомбинантных белков, в том числе и белков-антигенов различных возбудителей инфекционных заболеваний.

Acknowledgments

This study was supported by State Budgeted Project 0324-2016-008.

Conflict of interest

The authors declare no conflict of interest.

References

- Almaraz-Delgado A.L., Flores-Urbe J., Perez-Espana V.H., Salgado-Manjarrez E., Badillo-Corona J.A. Production of therapeutic proteins in the chloroplast of *Chlamydomonas reinhardtii*. *AMB Express*. 2014;4:57. DOI 10.1186/s13568-014-0057-4. PMID:25136510.
- Apeler H., Peters J., Schroder W. Expression, purification, and pharmacological characterization of a recombinant aprotinin variant. *Drug Res*. 2004;54(8):483-497. DOI 10.1055/s-0031-1297003.
- Arakawa T., Chong D.K.X., Langridge H.R. Efficacy of food plant-based oral cholera toxin B subunit vaccine. *Nat. Biotechnol*. 1998; 16:292-297. DOI 10.1038/nbt0398-292.
- Boynton J.E., Gilham N.W., Harris S.E. Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles. *Science*. 1988; 240:1534-1538. DOI 10.1126/science.2897716.
- Casteleijn M., Richardson D. Engineering cells and proteins – creating pharmaceuticals. *Eur. Pharm. Rev*. 2014;19(4):12-19.
- Cox K.M., Sterling J.D., Regan J.T. Glycan optimization of a human monoclonal antibody in the aquatic plant *Lemna minor*. *Nat. Biotechnol*. 2006;24:1591-1597. DOI 10.1038/nbt1260.
- Daniell H. Transgenic containment by maternal inheritance: effective or elusive. *Proc. Natl. Acad. Sci. USA*. 2007;1104:6879-6880. DOI 10.1073/pnas.0702219104.
- Daniell H., Chebolu S., Kumar S., Singleton M., Falconer R. Chloroplast-derived vaccine antigens and other therapeutic proteins. *Vaccine*. 2005;23:1779-1783. DOI 10.1016/j.vaccine.2004.11.004.
- DeCosa B., Moar W., Lee S.B., Miller M., Daniell H. Overexpression of the Bt cry2Aa2 operon in chloroplasts leads to formation of insecticidal crystals. *Nat. Biotechnol*. 2001;19:71-74. DOI 10.1038/83559.
- Desai P.N., Shrivastava N., Padh H. Production of heterologous proteins in plants: Strategies for optimal expression. *Biotechnol. Adv*. 2010;28:427-435. DOI 10.1016/j.biotechadv.2010.01.005.
- Dufourmantel N., Pelissier B., Garçon F., Peltier G., Ferullo J.M., Tissot G. Generation of fertile transplastomic soybean. *Plant Mol. Biol*. 2004;55:479-489. DOI 10.1007/s11103-004-0192-4.
- Fischer R., Schillberg S., Hellwig S., Twyman R.M., Drossard J. GMP issues for recombinant plant-derived pharmaceutical proteins. *Biotechnol. Adv*. 2012;30:434-439. DOI 10.1016/j.biotechadv.2011.08.007.
- Franklin S.E., Mayfield S.P. Prospects for molecular farming in the green algae *Chlamydomonas*. *Curr. Opin. Plant Biol*. 2004;7:159-165. DOI 10.1016/j.pbi.2004.01.012.
- Gelvin S.B. Agrobacterium-mediated plant transformation: the biology behind the “gene-jockeying” tool. *Microbiol. Mol. Biol. Rev*. 2003;67(1):16-37. DOI 10.1128/MMBR.67.1.16-37.2003.

- Gleba Y., Klimyuk V., Marillonnet S. Magniffection – a new platform for expressing recombinant vaccines in plants. *Vaccine*. 2005;23:2042-2048. DOI 10.1016/j.vaccine.2005.01.006.
- Gleba Y., Klimyuk V., Marillonnet S. Viral vectors for the expression of proteins in plants. *Curr. Opin. Biotechnol.* 2007;18:134-141. DOI 10.1016/j.copbio.2007.03.002.
- Haq T.A., Mason H.S., Clements J.D., Arntzen C.J. Oral immunization with a recombinant bacterial antigen produced in transgenic plants. *Science*. 1995;268:714-719. DOI 10.1126/science.7732379.
- Hellwig S., Drossard J., Twyman R.M., Fischer R. Plant cell cultures for the production of recombinant proteins. *Nat. Biotechnol.* 2004; 22:1415-1422. DOI 10.1016/j.biotechadv.2011.08.007.
- Howard J.A. Commercialization of biopharmaceutical and bioindustrial proteins from plants. *Crop Sci.* 2005;45:468-472. DOI 10.2135/cropsci2005.0468.
- Huang T.-K., McDonald K.A. Bioreactor engineering for recombinant protein production in plant cell suspension cultures. *Biochem. Eng. J.* 2009;45:168-184. DOI 10.1016/j.bej.2009.02.008.
- Kaldis A., Ahmad A., Reid A., McGarvey B., Brandle J., Ma Sh., Jevnikar A., Kohalmi S.E., Menassa R. Highlevel production of human interleukin-10 fusions in tobacco cell suspension cultures. *Plant Biotechnol. J.* 2013;11:535-545. DOI 10.1111/pbi.12041.
- Lelivelt C., McCabe M., Newell C. Stable plastid transformation in lettuce (*Lactuca sativa* L.). *Plant Mol. Biol.* 2005;58:763-774. DOI 10.1007/s11103-005-7704-8.
- Magnuson N.S., Linzmaier P.M., Reeves R., An G., Hay-Glass K., Lee J.M. Secretion of biologically active human interleukin-2 and interleukin-4 from genetically modified tobacco cells in suspension culture. *Protein Expr. Purif.* 1998;13:45-52. DOI 10.1006/prep.1998.0872.
- Martinez C.A., Guilietti A.M., Talou R. Research advances in plant-made flavivirus antigens. *Biotechnol. Adv.* 2012;30:1493-1505. DOI 10.1016/j.biotechadv.2012.03.004.
- Mason H.S., Haq T.A., Clements J.D., Arntzen C.J. Edible vaccine protects mice against *Escherichia coli* heat-labile enterotoxin (LT): potatoes expressing a synthetic LT-B-gene. *Vaccine*. 1998;16:1336-1343. DOI 10.1016/S0264-410X(98)80020-0.
- Mason H.S., Lam D.M., Arntzen C.J. Expression of hepatitis B surface antigen in transgenic plants. *Proc. Natl. Acad. Sci. USA*. 1992;89: 11745-11749.
- Nagels B., Weterings K., Callewaert N., van Damme E.J.M. Production of plant made pharmaceuticals: from plant host to functional protein. *Crit. Rev. Plant Sci.* 2012;31:148-180. DOI 10.1080/07352689.2011.616075.
- Permyakova N.V., Uvarova E.A., Deineko E.V. State of research in the field of the creation of plant vaccines for veterinary use. *Russian Journal of Plant Physiology*. 2015;62(1):23-38. DOI 10.1134/S1021443715010100.
- Reinsing S., Lang D., Knight C. The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science*. 2008;319:64-69. DOI 10.1126/science.1150646.
- Reski R., Parsons J., Decker E.L. Moss-made pharmaceuticals: from bench to bedside. *Plant Biotechnol. J.* 2015;13(8):1191-1198. DOI 10.1111/pbi.12401.
- Rosales-Mendoza S. *Algae-Based Biopharmaceuticals*. Springer, 2016. DOI 10.1007/978-3-319-32232-2.
- Rosales-Mendoza S., Tello-Olea M.A. Carrot cells: a pioneering platform for biopharmaceuticals production. *Mol. Biotechnol.* 2015;57: 219-232. DOI 10.1007/s12033-014-9837-y.
- Rybicki E.P. Plant-made vaccines for humans and animals. *Plant Biotechnol. J.* 2010;8:620-637. DOI 10.1111/j.1467-7652.2010.00507.x.
- Schelkunov S.N., Salyaev R.K., Pozdnyakov S.G., Rekoslavskaya N.I., Nesterov A.E. Immunogenicity of a novel, bivalent, plant-based oral vaccine against hepatitis B and human immunodeficiency viruses. *Biotechnol. Lett.* 2006;28(13):959-967. DOI 10.1007/s10529-006-9028-4.
- Schiermeyer A., Schillberg S. Plant molecular pharming – pharmaceuticals for human health. *Encyclopedia of Sustainability Science and Technology*. Ed. R.A. Meyers. N. Y.: Springer, 2012;8126-8141.
- Schillberg S., Raven N., Fischer R., Twyman R., Schiermeyer A. Molecular farming of pharmaceutical proteins using plant suspension cell and tissue cultures. *Curr. Pharm. Des.* 2013;19:5531-5542.
- Sourrouille C., Marshall B., Lienard D., Faye L. From Neanderthal to nanobiotech: From plant potions to pharming with plant factories. Ed. L. Faye, V. Gomord. *Methods in Molecular Biology: Recombinant Proteins From Plants*. Humana Press, a part of Springer Science+Business Media, 2009;1-23. DOI 10.1007/978-1-59745-407-0_1.
- Tacket C.O., Mason H.S., Losonsky G., Estes M.K., Arntzen C.J. Human immune responses to a novel Norwalk virus vaccine delivered in transgenic potatoes. *J. Infect. Dis.* 2000;182:302-305. DOI 10.1086/315653.
- Taunt H., Stoffels L., Purton S. Green biologics: The algal chloroplast as a platform for making biopharmaceuticals. *Bioengineered*. 2017. DOI 10.1080/21655979.2017.1377867.
- Tekoah Y., Shulman A., Kizhner T., Ruderfer I., Fux L., Nataf Y., Bartfeld D., Ariel T., Gingis-Velitski S., Hanania U., Shaaltiel Y. Large-scale production of pharmaceutical proteins in plant cell culture – the protalix experience. *Plant Biotechnol. J.* 2015;13:1199-1208. DOI 10.1111/pbi.12428.
- Tiwari S., Verma P.C., Singh P.K., Tuli R. Plants as bioreactors for the production of vaccines and antigens. *Biotechnol. Adv.* 2009;27:449-467. DOI 10.1016/j.biotechadv.2009.03.006.
- Twyman R.M., Stoger E., Schillberg S., Christou P., Fischer R. Molecular farming in plants: host systems and expression technology. *Trends Biotechnol.* 2003;21:570-578. DOI 10.1016/j.tibtech.2003.10.002.
- Weise A., Altmann F.M., Rodriguez-Franco M. High level expression of secreted complex glycosylated recombinant human erythropoietin in the *Physcomitrella* delta-fuc-t and delta-xyl-t mutant. *Plant Biotechnol. J.* 2007;5:389-401. DOI 10.1111/j.1467-7652.2007.00248.x.
- Yusibov V., Rabindran S. Recent progress in the development of plant-derived vaccines. *Expert Rev. Vaccin.* 2008;7:1173-1183. DOI 10.1586/14760584.7.8.1173.
- Yusibov V., Streatfield S., Kushnir N. Clinical development of plant-produced recombinant pharmaceuticals. *Hum. Vaccin.* 2011;7(3):313-321. DOI 10.4161/hv.7.3.14207.

Создание кандидатной вакцины против клещевого энцефалита на основе гибридного рекомбинантного flagG-protE-белка

П.А. Белавин¹✉, Д.А. Кунык¹, Е.В. Протопопова², В.Б. Локтев^{1, 2}, Е.В. Дейнеко¹

¹ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Государственный научный центр вирусологии и биотехнологии «Вектор» Роспотребнадзора, р.п. Кольцово, Новосибирская область, Россия

В настоящей работе описаны конструирование гена, кодирующего гибридный белок flagG-protE, синтез и очистка полученного рекомбинантного белка, а также исследование его антигенных характеристик на панели моноклональных антител (МКА). Рекомбинантный белок flagG-protE является перспективной молекулой для создания кандидатной рекомбинантной вакцины против клещевого энцефалита благодаря способности к связыванию с МКА против природного белка E вируса клещевого энцефалита. Проведено исследование антигенных детерминант двух рекомбинантных белков protE и flagG-protE с помощью панели из восьми МКА. Рекомбинантный белок protE представлен белком оболочки вируса клещевого энцефалита, в рекомбинантном белке flagG-protE к нему добавлен домен flagG, кодирующий флагеллин G *Salmonella typhi*. Установлено, что изучаемые МКА связывались с эпитопами рекомбинантного белка protE. Это свидетельствует о том, что исследуемый рекомбинантный белок имеет антигенную структуру, схожую с антигенной структурой нативного белка E вируса клещевого энцефалита. При исследовании рекомбинантного белка flagG-protE по способности к связыванию с панелью из восьми МКА только пять из них были способны связываться с эпитопами рассматриваемого белка. МКА 4F6, 7F10 и 6B9 не узнавали соответствующий эпитоп рекомбинантного белка flagG-protE, тогда как в рекомбинантном белке protE эти эпитопы выявлялись успешно. Полученные данные свидетельствуют о том, что антигенная структура рекомбинантного protE-белка может быть изменена под влиянием флагеллинового домена, что, в свою очередь, может привести к недоступности некоторых антигенных детерминант. Это обстоятельство необходимо учитывать при конструировании рекомбинантных антигенов. Тем не менее принципиально важные районы в области пептида слияния и домена III оказались доступны для антител. Это должно обеспечить формирование нейтрализующих антител, а наличие полной аминокислотной последовательности белка E в составе рекомбинантного белка будет индуцировать формирование Т-клеточного иммунного ответа. Появление нового поколения вакцин против клещевого энцефалита с более высоким уровнем безопасности и иммуногенности позволит усовершенствовать вакцинопрофилактику населения от клещевого энцефалита.

Ключевые слова: флагеллин G; белок оболочки; адъювант; ВКЭ; иммуногенность; вакцина.

Candidate vaccine construction against tick-borne encephalitis based on hybrid recombinant flagG-protE-protein

P.A. Belavin¹✉, D.A. Kunyk¹, E.V. Protopopova², V.B. Loktev^{1, 2}, E.V. Deineko¹

¹ Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

² State Research Center of Virology and Biotechnology "Vector", Novosibirsk region, Russia

The present work describes the construction of the gene encoding the recombinant protein flagG-protE, its synthesis, purification and study. The recombinant flagG-protE protein is a promising molecule for developing a candidate recombinant vaccine against tick-borne encephalitis by the ability to bind to monoclonal antibodies (MCA) against native protein E of tick-borne encephalitis virus. The antigenic determinants of two recombinant proteins were studied: protE and flagG-protE using a panel of 8 MCA. The recombinant protein protE comprises the tick-borne encephalitis virus envelope protein and the flagG-protE recombinant protein has an additional flagG domain encoding flagellin G of *Salmonella typhi*. It was found that the MCA tested revealed epitopes on the recombinant protein protE. This indicates that the investigated recombinant protein has an antigenic structure similar to the antigenic structure of the native tick-borne encephalitis virus protein E. In the study of the recombinant protein flagG-protE by the ability to bind a panel of 8 MCA, only five of them react with epitopes of the tested protein. MCA 4F6, 7F10, and 6B9 did not recognize the corresponding epitope in the recombinant flagG-protE protein, while in the recombinant protein protE, these epitopes were detected successfully. Our data indicate that the antigenic structure of recombinant protE-protein can be changed under the influence of the flagellin domain, which in turn can lead to the unavailability of some antigenic determinants. This fact must be taken into account when constructing recombinant molecules with antigenic properties. Nevertheless, the fundamentally important regions in the region of the fusion peptide and III domain are antigenically present on the surface of the recombinant protein. This should ensure the formation of neutralizing antibodies, and the presence of a complete amino acid sequence of protein E in the recombinant protein induces the formation of a T-cell immune response. The emergence of a new generation of vaccines against tick-borne encephalitis

with a higher level of safety and immunogenicity will improve the vaccine prevention of the population from tick-borne encephalitis.

Key words: flagellin G; envelope protein; adjuvant; TBE; immunogenicity; vaccine.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Белавин П.А., Кунык Д.А., Протопопова Е.В., Локтев В.Б., Дейнеко Е.В. Создание кандидатной вакцины против клещевого энцефалита на основе гибридного рекомбинантного flagG-protE-белка. Вавиловский журнал генетики и селекции. 2017; 21(8):986-992. DOI 10.18699/VJ17.323

HOW TO CITE THIS ARTICLE:

Belavin P.A., Kunyk D.A., Protopopova E.V., Loktev V.B., Deineko E.V. Candidate vaccine construction against tick-borne encephalitis based on hybrid recombinant flagG-protE-protein. Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding. 2017;21(8):986-992. DOI 10.18699/VJ17.323 (in Russian)

Флавивirusы способны инфицировать широкий круг организмов, которые включают в себя различные виды млекопитающих (в том числе человека), а также насекомых, птиц, рептилий. Сотни миллионов случаев заболевания человека, вызываемые различными флавивirusами, регистрируются в разных географических регионах мира (Grard et al., 2007; Локтев, 2011). Для России наибольшее медицинское значение имеет вирус клещевого энцефалита (Korenberg, Kovalevskii, 1999), который был открыт в 1937 г. на Дальнем Востоке. Фактически одновременно были созданы первые инактивированные вакцины против этой инфекции (Зильбер, 1939). Вирус клещевого энцефалита (ВКЭ) вызывает у людей заболевание центральной нервной системы с уровнем смертности 1–2 % для европейского и сибирского генотипов и до 30 % для дальневосточного (Gritsun et al., 2003). Вирус широко распространен на севере Евро-Азиатского континента, где ежегодно регистрируется до 14 тыс. случаев заболевания. Заболеваемость клещевым энцефалитом регистрируется более чем в 25 европейских и семи азиатских странах (Charrel et al., 2004; Amicizia et al., 2013). Недавно опубликованы данные о широком распространении и заболеваемости КЭ в Китае (Yoshii et al., 2017). На территории России ежегодно регистрируется от 1.5 до 10 тыс. случаев заболевания (Злобин, 2010). В последние годы наблюдается очень существенный рост – более чем в 50 раз, заболеваемости клещевым энцефалитом в северных регионах европейской части страны (Глушкова и др., 2011; Tokarevich et al., 2011; Микрюкова и др., 2014).

Для вакцинации против КЭ широко используют четыре типа инактивированных вакцин, производимых в России (штаммы 205 и Софьин), Австрии (Neudorfl) и Германии (K23) (Leonova et al., 2007; Morozova et al., 2014). На территории Китая с 1953 г. применяют оригинальную инактивированную вакцину против ВКЭ на основе штамма Zen-Zhang (Yoshii et al., 2017). Фактически только в Австрии удалось установить контроль над заболеваемостью КЭ при помощи вакцинации (Heinz et al., 2013). В России, в Свердловской области иммунизировано 86.1 % населения, что позволило также резко снизить заболеваемость КЭ в этом регионе (Государственный доклад..., 2017).

Основным иммуногеном, индуцирующим появление вируснейтрализующих антител, является гликопротеин E вируса клещевого энцефалита (Heinz, Stiasny, 2012). Бактериальный флагеллин – перспективный и эффективный природный адъювант, усиливающий иммунный ответ против флавивirusов (McDonald et al., 2007). Гибридная

молекула, несущая основные домены вирусного белка E и флагеллина G, может стать основой для конструирования кандидатной рекомбинантной вакцины против ВКЭ с усиленной иммуногенностью. Можно предположить, что применение рекомбинантной вакцины приведет к упрощению схемы иммунизации, обеспечит формирование длительного иммунитета, повысит безопасность вакцины и позволит создать новые биотехнологические схемы для наработки кандидатной вакцины, в том числе с использованием клеток растений.

В настоящей работе описано создание гена белка flagG-protE, его экспрессия, очистка белка flagG-protE как основного специфического компонента рекомбинантной вакцины против КЭ, исследование с помощью панели моноклональных антител к гликопротеину E ВКЭ сохранности основных антигенных детерминант рекомбинантной молекулы. Обсуждается перспективность гибридной молекулы flagG-protE для возможного использования в качестве кандидатной вакцины против ВКЭ с усиленной иммуногенностью.

Материалы и методы

Клонирование целевых генов. ДНК выделяли с помощью коммерческого «набора для выделения плазмидной ДНК» (ООО «Цитокин», г. Санкт-Петербург) согласно протоколу фирмы-изготовителя.

Для получения фрагментов ДНК генов *TBEVgp1* и гена *fliG* применяли ПЦР с праймерами, содержащими линкерные последовательности для эндонуклеаз рестрикции: *Bse3DI*, *BglII* и *HindIII* (табл. 1). В качестве матрицы использовали pHis6-protE и pHis6-flagG. Компетентные клетки получали по стандартной методике (Гловер, 1988). Клетки хранили при –70 °С.

Наработка и очистка белка. Клетки *Escherichia coli* BL21(DE3), несущие соответствующую плазмиду, засеивали в ночь в 5 мл LB, содержащей ампициллин (50 мг/мл). На следующий день инокулировали 1 мл ночной культуры в 100 мл LB с ампициллином (50 мкг/мл). Инкубировали на качалке в течение 2 ч при 37 °С и 180 об/мин. Индукцию биосинтеза рекомбинантного белка проводили добавлением 1 мл раствора 100 мМ ИПТГ. Продолжали культивировать на протяжении 5 ч. Биомассу собирали центрифугированием: 4000 об/мин, 10 мин (центрифуга Avanti, ротор JLA-16-250, США).

Полученную биомассу ресуспендировали в 4 мл фосфатно-солевого буфера, содержащего 10 мМ имидазола и 2 М мочевины (pH 8.0), и обрабатывали ультразвуком

Table 1. PCR-primers for cloning *fliG-TBEVgp1* gene

Primer	Nucleotide sequence
U- <i>Bgl</i> II-flaG	5'- CCCAGATCT ATGGCGCAGGTGATTAACACCA-3'
L- <i>Bse</i> 3DI-link fla	5'- CCC GCAATG CCC CGCGCCGCGCAGCAGGCTCAGCACG-3'
U- <i>Bse</i> 3DI_protE	5'- CCC GCAATGGG TAT GCCTCACGATGCACACATCTG-3'
L- <i>Hind</i> III_protE	5'- CCCAAGCTT ATATGCCTTTCCTGGTTTT-3'

Amplicon lengths: the flagellin gene, 1518 bp; the glycoprotein E gene, 1257 bp; the hybrid gene (at cloning restriction sites), 2770 bp. The recognition sites for restriction endonucleases are shown in boldface; hydrolysis sites for *Bse*3DI, in gray; and complementarity regions are underlined.

на ультразвуковой установке для лизиса клеток Cole-Parmer. Режим работы: мощность 35 Вт, 5 импульсов по 30 с с промежуточным охлаждением во льду 1 мин. Лизат центрифугировали при 10000 об/мин в течение 10 мин, супернатант удаляли (центрифуга Eppendorf miniSpin plus, ротор f-45-12-11). Осадок (тельца включения) суспендировали в 4 мл фосфатно-солевого буфера, содержащего 10 мМ имидазола и 8 М мочевины (pH 8.0), обрабатывали ультразвуком и центрифугировали в течение 5 мин при 10000 об/мин. Супернатант наносили на уравновешенную колонку с Ni-NTA-смолой (1 мл). После промывки связавшийся белок элюировали фосфатно-солевым буфером, содержащим 8 М мочевины и 250 мМ имидазола. Концентрацию белка в растворе измеряли методом Бредфорда (Bradford, 1976). Белковый электрофорез проводили по методике, описанной в работе (Laemmli, 1970).

МКА и иммуноферментный анализ. Моноклональные антитела (МКА) против ВКЭ получали с использованием гибридом, описанных ранее (Гайдамович и др., 1990; Протопопова и др., 1996; Романова и др., 2006). Иммуноферментный анализ (ИФА) проводили по общепринятым методикам с применением инактивированных вирусных и рекомбинантных антигенов (Разумов и др., 1991). С этой целью в лунки полистироловых планшетов вносили по 100 мкл очищенного антигена в концентрации 5–10 мкг/мл и сорбировали при 4 °С в течение ночи. Места неспецифического связывания насыщали 0.5 % раствором казеина в буфере ТСБ-Твин (0.145 М хлористого натрия, 20 мМ трис-НСl, 5 мМ PMSF (Sigma, США), содержащем 0.1 % Твин 20 (Serva, Германия), pH 7.4, 1 ч при 37 °С. Затем инкубировали с МКА 1 ч при 37 °С. Специфическое связывание выявляли антивидовыми мечеными пероксидазой хрена антителами против IgG. В качестве хромогена использовали раствор О-фенилендиамина (1 мг/мл орто-фенилендиамина, 0.03 % перекиси водорода) в цитратно-фосфатном буфере (0.2 М лимонной кислоты, 0.5 М Na₂HPO₄, pH 5.0). Выдерживали 20 мин в темноте, останавливали реакцию добавлением 100 мкл 1 N HCl на лунку и измеряли оптическую плотность на спектрофотометре Uniscan (Финляндия) со светофильтром с максимумом пропускания 450 нм.

Результаты

Конструирование гена, кодирующего гибридный белок flagG-protE, было выполнено при помощи метода Golden Gate Cloning (Engler et al., 2008). Метод основан на использовании эндонуклеаз рестрикции IIS-типа для получения фрагментов с уникальными липкими концами и

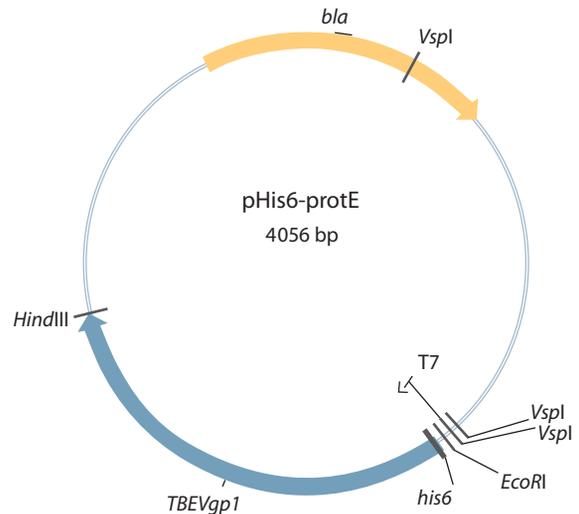


Fig. 1. Map of plasmid pHis6-protE.

Designations: *Hind*III, *Eco*RI, and *Vsp*I are restriction sites with corresponding endonucleases; *bla*, gene for ampicillin resistance; T7, phage T7 promoter; *his*6, a tag of 6 histidines; *TBEVgp1*, gene encoding protE. Arrows indicate the direction of transcription.

при лигировании позволяет оперировать более чем двумя фрагментами одновременно. Фрагмент гена *flaG* получен с помощью ПЦР с использованием праймеров с введенными сайтами узнавания для рестриктаз *Bgl*II и *Bse*3DI. Второй фрагмент гена *TBEVgp1* размером 1257 п. о., кодирующий белок protE, синтезирован с использованием праймеров с введенными сайтами узнавания для рестриктаз *Bse*3DI и *Hind*III и амплифицирован с плазмиды pHis6-protE (рис. 1).

Фрагменты ДНК после гидролиза соответствующими рестриктазами и очистки в 1 % агарозном геле при помощи «набора для очистки ДНК» («Цитокин», Санкт-Петербург) лигировали с вектором, полученным из pHis6-protE после обработки *Hind*III и *Bgl*II. Лигазную смесь использовали для трансформации клеток *E. coli*. По результатам рестрикционного анализа с участием эндонуклеаз рестрикции *Vsp*I, *Eco*RI и *Hind*III среди 24 независимо полученных клонов бактерий, устойчивых к ампициллину, выделено 8 клонов с ожидаемой для гена pHis6-flagG-protE картиной рестрикции (рис. 2).

Сравнивая расчетные данные и данные электрофореза, можно сделать вывод, что все сайты рестрикции ДНК плазмид располагаются правильно и отобранные клоны

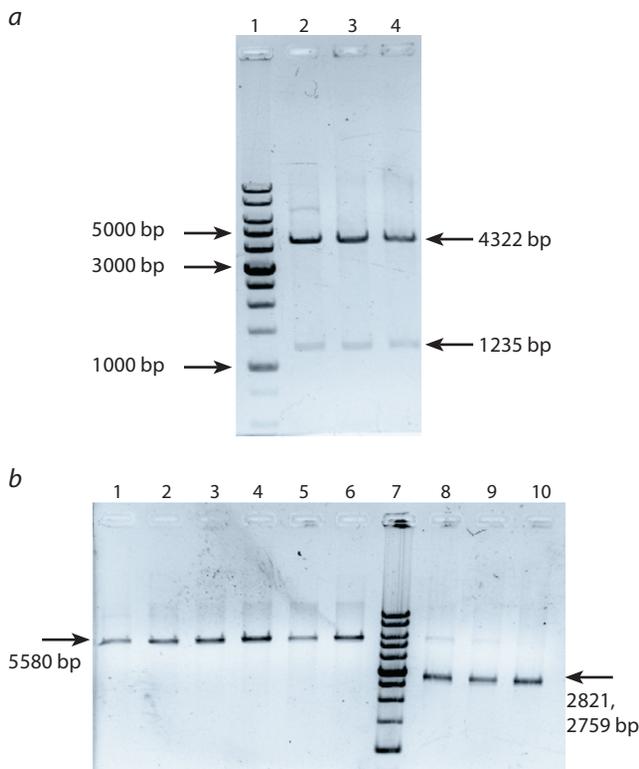


Fig. 2. Electrophoretic analysis of pHis6-flagG-protE clones: (a) Analysis of clones using the restriction enzyme *VspI* (lanes 2–4). Lane 1: 1-kb molecular weight ladder, SibEnzyme, Russia (16 fragments, bp: 10,000, 8000, 6000, 5000, 4000, 3000×2, 2500, 2000, 1500, 1000×2, 750, 500×2, 250). (b) Analysis of clones using restriction enzymes *EcoRI* (lanes 1–3) and *HindIII* (lanes 4–6), double hydrolysis with these endonucleases (lanes 8–10), SibEnzyme molecular weight ladder (lane 7). Expected fragment sizes, bp: with *VspI*, 4322 and 1235 bp; with *EcoRI*, 5580 bp; with *HindIII*, 5580 bp; with *EcoRI* + *HindIII*, 2821 and 2759 bp.

содержат целевую плазмиду pHis6-flagG-protE, которая включает ген, кодирующий рекомбинантный гибридный белок flagG-protE (рис. 3). Секвенирование плазмиды pHis6-flagG-protE подтвердило отсутствие мутаций и правильность сборки конструкции.

Обе конструкции (а именно *his6_TBEVgp1* и *fliG_TBEVgp1*) депонированы в GenBank: MG458224 – ген *his6_TBEVgp1*, кодирующий protE; MG458225 – ген *fliG_TBEVgp1*, кодирующий flagG-protE.

Анализ препаратов рекомбинантных белков проводили в 12 % полиакриламидном денатурирующем белковом электрофорезе. Расчетные молекулярные массы flagG-protE (101275.08 Да) и protE (47832.62 Да) хорошо коррелируют с полученными в эксперименте. На электрофореграмме (рис. 4) видно, что на дорожке FE белковая полоса имеет подвижность около 100 кДа, а на дорожке E – около 50 кДа. Кроме того, белковые полосы не имеют видимых примесей, что свидетельствует о чистоте полученных препаратов flagG-protE и protE.

Очищенные рекомбинантные белки protE и flagG-protE были использованы в иммуноферментном анализе с панелью из восьми моноклональных антител, полученных к вирусному гликопротеину E. Антигены сорбировали

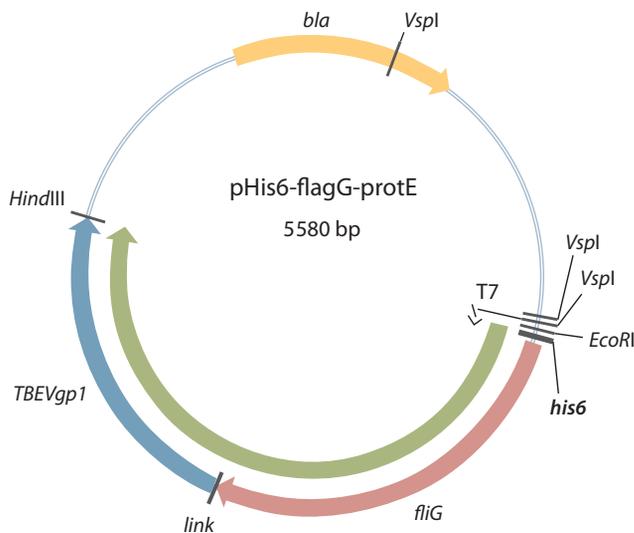


Fig. 3. Map of plasmid pHis6-flagG-protE.

Designations: *HindIII*, *EcoRI*, and *VspI* are sites of the corresponding restriction endonucleases; *bla*, gene for ampicillin resistance; T7, phage T7 promoter; *his6*, a tag of 6 histidines; *link*, sequence encoding 3 amino acids Gly-Ala-Gly; flagG-E, *fliG* and *TBEVgp1* are genes encoding the flagG and protE domains, respectively. Arrows indicate the direction of transcription.

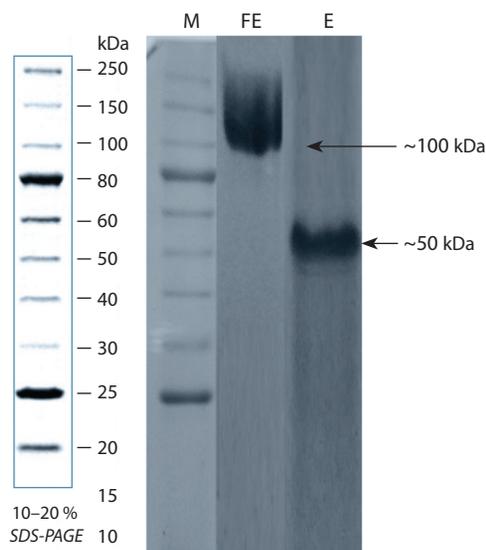


Fig. 4. Electrophoretic analysis of recombinant proteins flagG-protE and protE.

Designations: M, molecular weight ladder; FE, protein flagG-protE; E, protein protE.

на полистирольные планшеты, подвергали стандартным процедурам промывки, сорбции с антителами, конъюгатам и с последующим измерением оптической плотности на спектрофотометре. Результаты ИФА приведены в табл. 2.

По результатам анализа установлено, что панель из восьми моноклональных антител выявляла соответствующие эпитопы рекомбинантного белка protE. Это свидетельствует об антигенной структуре исследуемого рекомбинантного гибридного белка, схожей с антигенной структурой нативного вирусного белка E ВКЭ. Гибридный

Table 2. Interaction of monoclonal antibodies with recombinant proteins protE and flagG-protE and with viral antigens from prototype strains of three major TBEV genotypes

MCA	MCA-binding epitope on protein E (a. o.)	Recombinant proteins		TBEV genotype			
		flagG-protE	protE	europaean	siberian	far-East	
		Strains					
				Absettarov, KJ000002	C11-13, MF043953	205, JX498939	4072, KF951037
10H10	98–113	+++	+++	++	++	++	+++
4F6	19–273	–	+	+++	+++	+++	–
7F10	19–273	–	+	++	++	++	–
6B9	19–273	–	++	Н.и.	Н.и.	++	Н.и.
13F6	273–429	+++	++	++	++	+	+++
EB1	273–429	++	+	+++	+++	++	+++
E6B	273–429	++	+	++	++	++	++
7D3	273–429	+++	+++	+	++	+	+

The antigen was immobilized on a plate, 200 ng/well. MCA against TBEV protein E were diluted at 1 : 300. Optical densities: +, 0.3–0.8; ++, 0.8–1.5; +++, > 1.5; n.d., not determined. GenBank accession numbers are indicated for TBEV strains. The antigen was obtained by lysing a purified TBEV strain (Leonova et al., 2007) from the collection of virus strains at the Vector Center after genome-wide sequencing.

белок flagG-protE проявляет свои антигенные свойства аналогично рекомбинантному protE и нативному белку E вируса клещевого энцефалита, однако только пять видов МКА взаимодействовали с гибридным белком flagG-protE. МКА 4F6, 7F10 и 6B9 не узнавали эпитоп в гибридном белке при его наличии в составе белка protE. Идентичность аминокислотной последовательности фрагмента белка E в двух рекомбинантных полипептидах позволяет предположить, что эпитопы для МКА 4F6, 7F10 и 6B9 сохраняются в гибридном белке и, по всей вероятности, прикрыты полипептидной цепью флагеллина.

Принципиально важно отметить, что МКА 13F6, EB1, E6B и 10H10 взаимодействовали с гибридной рекомбинантной молекулой flagG-protE белка. МКА 13F6, EB1, E6B опознавали эпитопы домена III белка E, расположенного между 273–429 а. о. Домен III флавивирусов обеспечивает индукцию вируснейтрализующих антител и рецепторное взаимодействие (Ershova et al., 2016). МКА 10H10 распознают так называемый пептид слияния, расположенный в районе 98–113 а. о., который является высококонсервативным для большинства флавивирусов (Морозова и др., 2009). Этот район обеспечивает рецепторное взаимодействие вирусной частицы с ламининсвязывающим белком, участвует в процессе взаимодействия клеточных и вирусных мембран и формирует протективный иммунный ответ в организме (Ershova et al., 2016). Сохранность конформации эпитопов домена III и пептида слияния флавивирусов в составе гибридного белка позволяет предположить, что гибридный белок flagG-protE будет способен индуцировать образование вируснейтрализующих и антирецепторных антител и тем самым формировать полноценный противовирусный иммунитет, защищающий организм от развития инфекционного процесса. При этом флагеллин в составе гибридной молекулы будет дополнительно обладать сильными адьювантными свойствами, что сформирует выраженный и продолжи-

тельный противовирусный иммунитет. Эти свойства гибридного белка flagG-protE дают основание рассматривать его как весьма перспективный для создания кандидатной вакцины против клещевого энцефалита.

Обсуждение

Перспективным направлением для создания новых рекомбинантных вакцин является конструирование новых гибридных иммуногенов, несущих в своем составе эффективные природные адьюванты. В составе рекомбинантных вакцин в качестве иммуногенов чаще всего используют белки оболочек вирусов, трансмембранные белки различных инфекционных агентов, несущие эпитопы, узнаваемые антителами. В качестве адьювантов в последнее время используют различные полипептиды, такие как эндотоксины (В субъединица холерного токсина (СТВ), термолабильный эндотоксин сальмонеллы (LTB)), интерфероны и флагеллины. Например, в работе (Taylor et al., 2011) описано создание новой эффективной рекомбинантной вакцины против гриппа. При использовании флагеллина в качестве адьюванта уровень антител в крови пациентов вырос более чем в 10 раз. Кроме того, даже в дозе 5 мкг вакцина показала свою эффективность для пожилых людей (опытная группа со средним возрастом 71 год). До применения флагеллина пожилые люди слабо реагировали на вакцинацию и заболели гриппом. Таким образом, продемонстрировано, что флагеллин обладает потенциалом для стимулирования врожденного и приобретенного иммунитета, причем формирует гуморальный и клеточный иммунный ответы при вакцинации (Mizel, Bates, 2010; Skountzou et al., 2010; Girard et al., 2011).

Флагеллин G (flagG) из *Salmonella typhi*, используемый нами как адьювант при создании кандидатной вакцины против клещевого энцефалита, входит в состав жгутиков бактерий, где выполняет структурную функцию. При попадании в организм теплокровных флагеллин взаимодей-

ствует с toll-like-рецептором (TLR5) клеток врожденной иммунной системы, что приводит к развитию воспаления как защитной реакции организма со стороны врожденного иммунитета и дальнейшему формированию специфического (адаптивного) иммунитета. Благодаря этому свойству флагеллин рассматривается как адъювант, т.е. усилитель иммунных ответов организма на введение чужеродного белка. Предложенный нами подход к созданию новых субъединичных вакцин против клещевого энцефалита предполагает объединение нескольких компонентов в одном полипептидном рекомбинантном белке: flagG, соединенный «мостиком» из трех аминокислот с protE, а также включение полигистидинового тракта на N-конце для улучшения эффективности выделения и очистки рекомбинантного белка. Такого рода вакцины потенциально должны минимизировать негативные эффекты, связанные с повышением температуры после вакцинации и развитием неспецифических воспалительных процессов в месте введения, а также усиливать иммунный ответ организма человека при меньшей дозе препарата.

Гликопротеин E входит в состав вириона вируса клещевого энцефалита и является единственным экспонируемым на его поверхности белком (Mandl et al., 1988). В своем составе гликопротеин E имеет три основных домена, причем домен III обеспечивает индукцию вируснейтрализующих антител (Heinz, Stiasny, 2012). Иммунизация даже одним этим доменом обеспечивает защиту животных от летальной инфекции (Ershova et al., 2016). Полученные нами данные свидетельствуют о том, что антигенная структура рекомбинантного protE-белка может быть изменена под влиянием флагеллинового домена, что в свою очередь может привести к недоступности некоторых антигенных детерминант. Тем не менее принципиально важные районы в области пептида слияния и домена III представлены на поверхности гибридного белка. Это должно обеспечить формирование нейтрализующих антител, а наличие полной аминокислотной последовательности белка E в составе гибридного белка также индуцирует формирование Т-клеточного иммунного ответа. Полученные результаты показывают, что гибридный белок flagG-protE имеет хорошие перспективы для использования в качестве кандидатной вакцины против вируса клещевого энцефалита.

Таким образом, сконструирован рекомбинантный ген, кодирующий флагеллин *G. S. typhi*, слитый с геном гликопротеина E (protE) ВКЭ. Между генами заложен «шарнир» из трех аминокислотных остатков и полигистидиновый тракт для аффинной очистки на колонке Ni-NTA. С помощью панели МКА показана сохранность основных антигенных детерминант в составе гибридного белка. Можно ожидать усиления иммуногенных характеристик рекомбинантного белка flagG-protE за счет введения адъюванта. Кроме того, мы предполагаем, что белки с большой молекулярной массой обладают повышенной иммунностью, что должно повысить эффективность потенциальной вакцины против ВКЭ. Появление нового поколения вакцин против КЭ с более высоким уровнем безопасности и иммуногенности позволит усовершенствовать вакцинопрофилактику населения от клещевого энцефалита.

Acknowledgments

This study was supported by State Budgeted Project 0324-2016-008 “Genetic Foundations of Biotechnologies and Bioinformatics”.

Conflict of interest

The authors declare no conflict of interest.

References

- Amicizia D., Domnich A., Panatto D., Lai P.L., Cristina M.L., Avio U., Gasparini R. Epidemiology of tick-borne encephalitis (TBE) in Europe and its prevention by available vaccines. *Hum. Vaccin. Immunother.* 2013;9(5):1163-1171. DOI 10.4161/hv.23802.
- Bradford M.M. A rapid and sensitive method for the quantitation microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 1976;72:248-254. DOI 10.1016/0003-2697(76)90527-3.
- Charrel R.N., Attoui H., Butenko A.M., Clegg J.C., Deubel V., Frolova T.V., Gould E.A., Gritsun T.S., Heinz F.X., Labuda M., Lashkevich V.A., Loktev V., Lundkvist A., Lvov D.V., Mandl C.W., Niedrig M., Papa A., Petrov V.S., Plyusnin A., Randolph S., Süß J., Zlobin V.I., de Lamballerie X. Tick-borne virus diseases of human interest in Europe. *Clin. Microbiol. Infect.* 2004;10:1040-1055. DOI 10.1111/j.1469-0691.2004.01022.x.
- Engler C., Kandzia R., Marillonnet S.A. One pot, one step, precision cloning method with high throughput capability. *PLoS ONE.* 2008; 3(11):e3647. DOI 10.1371/journal.pone.0003647.
- Ershova A.S., Gra O.A., Lyaschuk A.M., Grunina T.M., Tkachuk A.P., Bartov M.S., Savina D.M., Sergienko O.V., Galushkina Z.M., Gudov V.P., Kozlovskaya L.I., Kholodilov I.S., Gmyl L.V., Karganova G.G., Lunin V.G., Karyagina A.S., Gintsburg A.L. Recombinant domains III of Tick-Borne Encephalitis Virus envelope protein in combination with dextran and CpGs induce immune response and partial protectiveness against TBE virus infection in mice. *BMC Infect. Dis.* 2016;16(1):544. DOI 10.1186/s12879-016-1884-5.
- Gaidamovich S.Ya., Loktev V.B., Lavrova N.A., Maksyutov A.Z., Mel'nikova E.E., Pereboev A.V., Protopenova E.V., Razumov I.A., Svshnikova N.A., Khusainova A.D. Monoclonal antibodies cross-reacting with the tick-borne encephalitis virus and with the Venezuelan equine encephalomyelitis virus. *Voprosy virusologii = Problems of Virology.* 1990;35(3):221-225. (in Russian)
- Girard A., Saron W., Bergeron-Sandoval L.P., Sarhan F., Archambault D. Flagellin produced in plants is a potent adjuvant for oral immunization. *Vaccine.* 2011;29(38):6695-6703. DOI 10.1016/j.vaccine.2011.06.092.
- Glover D.M. DNA Cloning. A Practical Approach, vols. 1, 2. Oxford: IRL Press Publ., 1985. [Russ. ed.: Glover D.M., Klonirovanie DNK. Metody. Moscow: Mir Publ., 1988. (in Russian)].
- Glushkova L.I., Korabel'nikov I.V., Egorova Iu.I. Occurrence of *Ixodes persulcatus* P. Sch. in the southern and central Republic of Komi. *Meditinskaya parazitologiya i parazitarnye bolezni = Medical Parasitology and Parasitological Diseases.* 2011;3:48-50. (in Russian)
- Gosudarstvennyy doklad “O sostoyanii sanitarno-epidemiologicheskogo blagopoluchiya naseleniya v Sverdlovskoy oblasti v 2016 godu” [State report “On the state of the sanitary and epidemiological welfare of the population in the Sverdlovsk oblast in 2016”]. Yekaterinburg, 2017. (www.66.rospotrebнадзор.ru). (in Russian)
- Grard G., Moureau G., Charrel R.N., Lemasson J.J., Gonzalez J.P., Gallian P., Gritsun T.S., Holmes E.C., Gould E.A., de Lamballerie X. Genetic characterization of tick-borne flaviviruses: new insights into evolution, pathogenetic determinants and taxonomy. *Virology.* 2007; 361(1):80-92. DOI 10.1016/j.virol.2006.09.015.
- Gritsun T.S., Lashkevich V.A., Gould E.A. Tick-borne encephalitis. *Antiviral Res.* 2003;57:129-146. DOI 10.1016/S0166-3542(02)00206-1.
- Heinz F.X., Stiasny K. Flaviviruses and their antigenic structure. *J. Clin. Virol.* 2012;55(4):289-295. Epub 2012; Sep 21. DOI 10.1016/j.jcv.2012.08.024.

- Heinz F.X., Stiasny K., Holzmann H., Grgic-Vitek M., Kriz B., Essl A., Kundi M. Vaccination and tick-borne encephalitis, central Europe. *Emerg. Infect. Dis.* 2013;19(1):69-76. DOI 10.3201/eid1901.120458.
- Korenberg E.I., Kovalevskii Y.V. Main features of tick-borne encephalitis eco-epidemiology in Russia. *Zentralblatt Bakteriologie*. 1999; 289(5-7):525-539. DOI 10.1016/S0934-8840(99)80006-1.
- Laemmli U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 1970;227:680-685.
- Leonova G.N., Ternovoi V.A., Pavlenko E.V., Maistrovskaya O.S., Protopopova E.V., Loktev V.B. Evaluation of vaccine Encepur® Adult for induction human neutralizing antibodies against recent Far Eastern subtype strains of tick-borne encephalitis virus. *Vaccine*. 2007;25(5):895-901. DOI 10.1016/j.vaccine.2006.09.014.
- Loktev V.B. Taxonomy of flaviviruses and their genetic diversity. *In: Infektsii, peredavaemye kleshchami v Sibirskom regione. (Red. Vlasova V.V., Repina V.E.) [Vlasova V.V., Repina V.E. (Eds). Infections Transmitted by Ticks in the Siberian Region]. Novosibirsk: SB RAS Publ., 2011;257-279. (in Russian)*
- Mandl C.W., Guirakhoo F., Holzmann H., Heinz F.X., Kunz Ch. Identification of dengue sequences by genomic amplification: rapid diagnosis of dengue virus serotypes in peripheral blood. *J. Virol.* 1988; 63:564-571. DOI 10.1016/0166-0934(90)90042-E.
- McDonald W.F., Huleatt J.W., Foellmer H.G., Hewitt D., Tang J., Desai P., Price A., Jacobs A., Takahashi V.N., Huang Y., Nakaar V., Alexopoulou L., Fikrig E., Powell T.J. A West Nile virus recombinant protein vaccine that coactivates innate and adaptive immunity. *J. Infect. Dis.* 2007;195(11):1607-1617.
- Mikryukova T.P., Chausov E.V., Konovalova S.N., Kononova Yu.V., Protopopova E.V., Kartashov M.Yu., Ternovoy V.A., Glushkova L.I., Korabel'nikov I.V., Egorova Yu.I., Loktev V.B. Genetic diversity of the tick-borne encephalitis virus in *Ixodes persulcatus* ticks in northeastern European Russia. *Parazitologiya = Parasitology*. 2014;48(2):131-149. (in Russian)
- Mizel S.B., Bates J.T. Flagellin as an adjuvant: cellular mechanisms and potential. *J. Immunol.* 2010;185(10):5677-5682. DOI 10.4049/jimmunol.1002156.
- Morozova O.V., Bakhvalova V.N., Matveyev L.E., Shevtsova A.S., Isaeva E.I., Zlobin V.I., Protopopova E.V., Seligman S. Antigenic and immunogenic properties of multiple antigenic peptides including the flavivirus fusion peptide. *Epidemiologiya i vaktsinoprofilaktika = Epidemiology and Vaccinal Prevention*. 2009;6(49):44-50. (in Russian)
- Morozova O.V., Bakhvalova V.N., Potapova O.F., Grishechkin A.E., Isaeva E.I., Aldarov K.V., Klinov D.V., Vorovich M.F. Evaluation of immune response and protective effect of four vaccines against the tick-borne encephalitis virus. *Vaccine*. 2014;32(25):3101-3106. DOI 10.1016/j.vaccine.2014.02.046.
- Protopopova E.V., Khusainova A.D., Konovalova S.N., Loktev V.B. Preparation and study of properties of anti-idiotypic antibodies, carrying hemagglutinating paratopes of tick-borne encephalitis virus on their surface. *Voprosy virusologii = Problems of Virology*. 1996; 41(2):50-53. (in Russian)
- Razumov I.A., Agapov E.V., Pereboevoy A.V., Protopopova E.V., Lebedeva S.D., Loktev V.B. Antigenic structure of the Venezuelan equine encephalomyelitis E2 glycoprotein as analyzed with rat monoclonal antibodies. *Voprosy virusologii = Problems of Virology*. 1991;36(1): 34-37. (in Russian)
- Romanova L.Yu., Gmyl' L.V., Loktev V.B., Protopopova E.V., Dzhivanyan T.I., Lashkevich V.A., Karganova G.G. Change in the antigenic structure of surface glycoprotein E of tick-borne encephalitis virus during its adaptation to ticks and mammals. *Voprosy virusologii = Problems of Virology*. 2006;51(6):31-34. (in Russian)
- Skountzou I., Martin M.P., Wang B., Ye L., Koutsonanos D., Weldon W., Jacob J., Compans R.W. *Salmonella* flagellins are potent adjuvants for intranasally administered whole inactivated influenza vaccine. *Vaccine*. 2010;28(24):4103-4112. DOI 10.1016/j.vaccine.2009.07.058.
- Taylor D., Treanor J., Strout C., Johnson C., Fitzgerald T., Kavita U., Ozer K., Tussey L., Shaw A. Induction of a potent immune response in the elderly using the TLR-5 agonist, flagellin, with a recombinant hemagglutinin influenza-flagellin fusion vaccine (VAX125, STF2. HA1 SI). *Vaccine*. 2011;29:4897-4902. DOI 10.1016/j.vaccine.2011.05.001.
- Tokarevich N.K., Tronin A.A., Blinova O.V., Buzinov R.V., Boltenev V.P., Yurasova E.D., Nurse J. The impact of climate change on the expansion of *Ixodes persulcatus* habitat and the incidence of tick-borne encephalitis in the north of European Russia. *Glob. Health Action*. 2011;4:8448.
- Yoshii K., Song J.Y., Park S.B., Yang J., Schmitt H.J. Tick-borne encephalitis in Japan, Republic of Korea and China. *Emerg. Microbes Infect.* 2017;6(9):e82. DOI 10.1038/emi.2017.69.
- Zil'ber L.A. Epidemic spring-summer tick-borne encephalitis. *Sovremennaya meditsina = Modern Medicine*. 1939;23:11-15. (in Russian)
- Zlobin V.I. Tick-borne encephalitis in the Russian Federation: etiology, epidemiology and prevention strategy. *Terra Medica*. 2010;2:13-21. (in Russian)



Создание стабильного штамма-продуцента полноразмерного антитела человека на примере антитела против вируса экстремелии

А.Л. Матвеев¹, Я.А. Хлусевич¹, И.К. Байков¹, И.В. Бабкин^{1, 2}, Е.П. Гончарова¹, В.В. Морозова^{1, 2}, Н.В. Тикунова^{1, 2}

¹ Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук, Новосибирск, Россия

² Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия

Полноразмерные антитела человека обладают большим терапевтическим потенциалом, однако разработка стабильных штаммов, обеспечивающих высокий уровень продукции полноразмерных антител, является непростой задачей, поскольку молекулы антител содержат два типа полипептидных цепей. При получении штамма-продуцента чаще всего используют подход, основанный на случайной интеграции в геном плазмиды, содержащей ген, который кодирует целевой белок. Цель данного исследования – разработка оригинальной экспрессионной системы на основе направленной рекомбинации (генный таргетинг) для интеграции гена, кодирующего полноразмерное антитело человека, в транскрипционно активную область генома эукариотических суспензионных клеток CHO-S. Для создания высокопродуктивного стабильного штамма на первом этапе была сконструирована кассетная векторная плаزمиды pCDNA5/FRT-DHFR-CH-CL, содержащая сайт гомологичной рекомбинации и гены, кодирующие тяжелую и легкую цепи полноразмерного антитела человека класса IgG1/каппа. ДНК плазмиды pCDNA5/FRT-DHFR-CH-CL организовали таким образом, что перед последовательностями, кодирующими константные домены тяжелых и легких цепей антитела человека, находились сайты узнавания эндонуклеаз рестрикции для удобного встраивания фрагментов ДНК, кодирующих соответствующие варибельные домены тяжелых и легких цепей. На втором этапе в кассетную плазмиду pCDNA5/FRT-DHFR-CH-CL были встроены ДНК-фрагменты, кодирующие варибельные домены тяжелых и легких цепей антитела человека против ортопоксвирусного белка р35. Затем полученной плазмидой трансфицировали эукариотические клетки CHO-S/FRT, содержащие FRT-сайт гомологичной рекомбинации и экспрессирующий белок GFP. При встройке целевых генов, кодирующих тяжелые и легкие цепи антитела в FRT-сайт, продукция GFP должна была прекратиться. При использовании такой системы отбора был получен стабильный клон, продуцирующий целевое антитело fh8E с уровнем продукции около 100 мкг/мл. Аффинность связывания очищенного антитела fh8E с таргетным белком, измеренная методом поверхностного плазмонного резонанса, составила 12 нМ. Антитело fh8E продемонстрировало вируснейтрализующие свойства в реакции ингибирования бляшкообразования вируса осповакцины в экспериментах *in vitro*.

Ключевые слова: полноразмерное антитело человека; штамм-продуцент; клетки CHO; плазмиды; геномная амплификация; вирус экстремелии.

Development of a stable eukaryotic strain producing fully human monoclonal antibody on the basis of the human antibody against ectromelia virus

A.L. Matveev¹, Ya.A. Khlusevich¹, I.K. Baykov¹, I.V. Babkin^{1, 2}, E.P. Goncharova¹, V.V. Morozova^{1, 2}, N.V. Tikunova^{1, 2}

¹ Institute of Chemical Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia

² Novosibirsk State University, Novosibirsk, Russia

Fully-human antibodies have a great therapeutic importance; however, the development of stable strains providing a high level of production of full-size antibodies is a challenging task, as antibody molecules contain two types of polypeptide chains. To develop the producing strain, random integration of the plasmid containing the gene encoding the target protein into the genome of the host cells is commonly used. The aim of this study was the development of an original expression system, using gene targeting to integrate the gene encoding the fully-human antibody into the transcriptionally active region of the genome of eukaryotic suspension cells CHO-S. To develop a stable strain, the cassette vector plasmid pCDNA5/FRT-DHFR-CH-CL containing the site of homologous recombination and the genes encoding heavy and light chains of the fully human antibody of the IgG1/kappa class was constructed at the first step. Notably, DNA of the plasmid pCDNA5/FRT-DHFR-CH-CL was organized in such a way that the restriction sites for rapid cloning of DNA fragments encoding the variable domains of heavy and light chains were inserted upstream of the sequences encoding constant domains of the heavy and light chains of the antibody. Secondly, DNA fragments encoding the variable domains of the heavy and light chains of antibody against orthopoxvirus protein p35 were inserted into the pCDNA5/FRT-DHFR-CH-CL cassette plasmid. Then, CHO-S/FRT cells, which contain the FRT-site for homologous recombination and are able to produce green fluorescence protein GFP, were transfected with the constructed plasmid. After the insertion of the target genes into the FRT-site, GFP production was supposed to stop. Using this selection system, a stable clone producing target antibody fh8E was selected with the level of production of about 100 µg/ml. The binding affinity of purified antibody fh8E with the targeted protein, measured

by surface plasmon resonance, was 12 nM. In addition, antibody fh8E demonstrated anti-vaccinia virus activity in the plaque reduction neutralization test *in vitro*.

Key words: fully human antibody; producing strain; CHO cells; plasmid; genomic amplification; ectromelia virus.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Матвеев А.Л., Хлусевич Я.А., Байков И.К., Бабкин И.В., Гончарова Е.П., Морозова В.В., Тикунова Н.В. Создание стабильного штамма-продуцента полноразмерного антитела человека на примере антитела против вируса эктромелии. Вавиловский журнал генетики и селекции. 2017;21(8):993-1000. DOI 10.18699/VJ17.324

HOW TO CITE THIS ARTICLE:

Matveev A.L., Khlusevich Ya.A., Baykov I.K., Babkin I.V., Goncharova E.P., Morozova V.V., Tikunova N.V. Development of a stable eukaryotic strain producing fully human monoclonal antibody on the basis of the human antibody against ectromelia virus. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2017;21(8):993-1000. DOI 10.18699/VJ17.324 (in Russian)

В последние годы все большее место на фармацевтическом рынке занимают препараты на основе терапевтических антител. Это связано с исключительными свойствами этих молекул – высокой специфичностью взаимодействия с антигеном, способностью вовлекать различные звенья иммунной системы, а также проникать в ткани и выводиться из организма с помощью естественных механизмов. Первыми были разработаны терапевтические и профилактические препараты на основе сывороток иммунизированных людей или животных. Так, в Российской Федерации успешно применяют иммуноглобулин человека против гепатита В, противоклещевой и противостолбнячный иммуноглобулины, полученные на основе донорской крови (<http://www.microgen.ru/products>); в США для лечения осложнений после вакцинации вирусом осповакцины используют сывороточный вакцинный иммуноглобулин (VIG) (Hopkins, Lane, 2004). Вместе с тем в настоящее время наблюдается тенденция отказа от сывороточных препаратов из-за возможного биологического риска, связанного с их применением, и в клинической практике в основном используют рекомбинантные полноразмерные антитела, которые нарабатывают в эукариотических клеточных линиях (Reichert, 2016, 2017). Однако из более чем сорока терапевтических антител, представленных на мировом фармацевтическом рынке, большинство – это противцитокинные и противораковые антитела, и имеется лишь одно противовирусное терапевтическое антитело (против респираторного синциального вируса).

Ключевым этапом в процессе разработки терапевтических антител является создание стабильного высокопродуктивного штамма-продуцента, обеспечивающего коммерческую привлекательность производства необходимого для медицины антитела. При этом получить клеточные линии, стабильно экспрессирующие рекомбинантные антитела с высоким уровнем продукции, бывает трудно из-за того, что такие клоны встречаются с частотой 1/10000, и во время культивирования клеток чужеродные гены легко элиминируются (Lucas et al., 1996). Продуктивность рекомбинантных клеточных линий напрямую зависит от места интеграции целевых генов в геном, но вероятность получения стабильных клеток, у которых целевой ген интегрирован в активный участок генома, довольно низка. Известно, что только 0.1 % геномной ДНК содержит транскрипционно активные последовательности, поэтому на уровень продукции влияют геномная среда, так называемый эффект положения и число копий

целевого гена в геноме (Little et al., 1993). Однако гомологичная рекомбинация между геномом клетки хозяина и плазмидной ДНК, используемой для трансфекции, происходит с низкой частотой (Hirata et al., 2002).

Одним из способов повышения вероятности направленной интеграции является применение специфических ферментов, таких как Flp-рекомбиназы дрожжей. Эти ферменты катализируют гомологичную направленную рекомбинацию с высокой частотой, если в донорной плазмидной ДНК и реципиентной геномной ДНК есть конкретные сайты для рекомбинации (Wilson et al., 2001). Flp-рекомбиназы были широко использованы для направленной рекомбинации целевых генов в геном клеток животных (Wirth et al., 2007). Сайт-специфичные системы рекомбинации Cre-LoxP и FLP-FRT были хорошо изучены и часто применяются для сайт-специфической рекомбинации клеток животных (Kito et al., 2002; Coroadinha et al., 2006; Huang et al., 2007; Kameyama et al., 2010). Тем не менее интеграция в эухроматин не является гарантией обеспечения долгосрочной экспрессии рекомбинантного гена. Во многих случаях реципиентные гены быстро инактивируются (замолкают), что может быть ассоциировано с гипоацетилизацией гистонов, метилированием лизина-9 гистона H3 и увеличением CpG-метилирования в промоторной области трансгена (Richards, 2002; Mutskov, Felsenfeld, 2004).

Разработано несколько стратегий для уменьшения негативного влияния эффекта положения при случайной интеграции (Zahn-Zabal et al., 2001; Fernandez et al., 2002; Kwaks et al., 2003). Одна из возможных стратегий преодоления эффекта положения – это использование направленной рекомбинации (генный таргетинг) для интеграции гена, кодирующего целевой белок, по специфическим сайтам в транскрипционно активную область генома.

Как правило, разработкой препаратов на основе полноразмерных антител занимаются частные фармацевтические компании, поэтому многие детали, связанные с получением штаммов-продуцентов, остаются неизвестными. При получении штамма-продуцента чаще всего применяют традиционный подход, основанный на случайной интеграции в геном последовательности ДНК, кодирующей целевой белок. Клеточные линии для трансфекции с помощью направленной рекомбинации, созданные на основе суспензионных клеток, не описаны. Однако именно такой подход является, на наш взгляд, наиболее перспективным для получения промышленных штаммов-продуцентов. Цель данного исследования – разработка оригинальной

экспрессионной системы для создания на ее основе высокопродуктивных стабильных штаммов-продуцентов полноразмерных антител человека на примере антитела против белка р35 ортопоксвирусов.

Материалы и методы

Клетки *Escherichia coli* XL1-blue, использовавшиеся при конструировании плазмидной ДНК, суспензионные эукариотические клетки CHO-S/FRT, содержащие FRT-сайт гомологичной рекомбинации и продуцирующие зеленый флуоресцентный белок GFP (green fluorescence protein), и эукариотические клетки *Vero* Е6 были получены из Коллекции экстремофильных микроорганизмов и типовых культур Института химической биологии и фундаментальной медицины (ИХБФМ) СО РАН. Штамм вируса осповакцины, несущего в составе генома ген *gfp* (BOB-LIVP-GFP), был создан ранее (Петров и др., 2013).

Конструирование кассетной векторной плазмиды рCDNA5/FRT-DHFR-CH-CL, содержащей сайт гомологичной рекомбинации и гены, кодирующие тяжелую и легкую цепи полноразмерного антитела человека, проводили в несколько этапов (рис. 1). На первом этапе в плазмиду рCDNA5/FRT (Life Technologies), содержащую FRT-сайт для гомологичной рекомбинации, была встроена последовательность промотора гена, кодирующего фактор элонгации рEf1 α . Для этого ДНК-фрагмент, содержащий в своем составе промотор *Pr/Ef1 α* , полученный в ПЦР с использованием праймеров Ef1 α _U17 и Ef1 α _L22 (см. таблицу) и плазмиды рbudCE4.1 (Life Technologies), обрабатывали эндонуклеазами рестрикции *KpnI* и *BamHI* и объединяли в реакции лигирования с плазмидой рCDNA5/FRT, обработанной теми же рестриктазами. На втором этапе в полученную плазмиду рCDNA5/FRT-рEf1 α последовательно встраивали гены, кодирующие константные домены тяжелой и легкой цепей антитела человека, с сайтами полиаденилирования бычьего гормона роста (BST) и вируса SV40 соответственно. Соответствующие фрагменты ДНК (*CH*+polyA_{BST} и *CL*+polyA_{SV40}) были получены в ПЦР с использованием плазмид рCH и рCL, сконструированных ранее (Байков и др., 2013), и пар праймеров CH_U22 и BGH_polyA_L18, а также CL_U21 и SV40_polyA_L30 (см. таблицу). В результате была создана промежуточная плаزمида рCDNA5/FRT-CH-CL, организованная таким образом, что перед последовательностями, кодирующими константные домены тяжелых и легких цепей антитела человека, находились сайты рестрикции *EcoRV* и *NotI*, а также *NheI* и *AflIII* соответственно для быстрого встраивания фрагментов ДНК, кодирующих эти переменные фрагменты.

На следующем этапе в плазмиде рOptiVec-ТОРО (Life Technologies), содержащей ген, кодирующий дегидрофолат редуктазу (*dhfr*), был заменен IRES-элемент вируса энцефаломиокардита на слабый эукариотический промотор фосфоглицераткиназы – *Pr/pgk*. Для этого с помощью праймеров рGK_U20 и рGK_L17 (см. таблицу) был получен ПЦР-фрагмент, содержащий *Pr/pgk*, причем на 5'- и 3'-концах ПЦР-фрагмента находились сайты *HpaI* и *BglII* соответственно. В качестве матрицы для получения этого ПЦР-продукта использовали плазмиду рExT (Tasic et al., 2012). Затем плазмидная ДНК рOptiVec-ТОРО и полу-

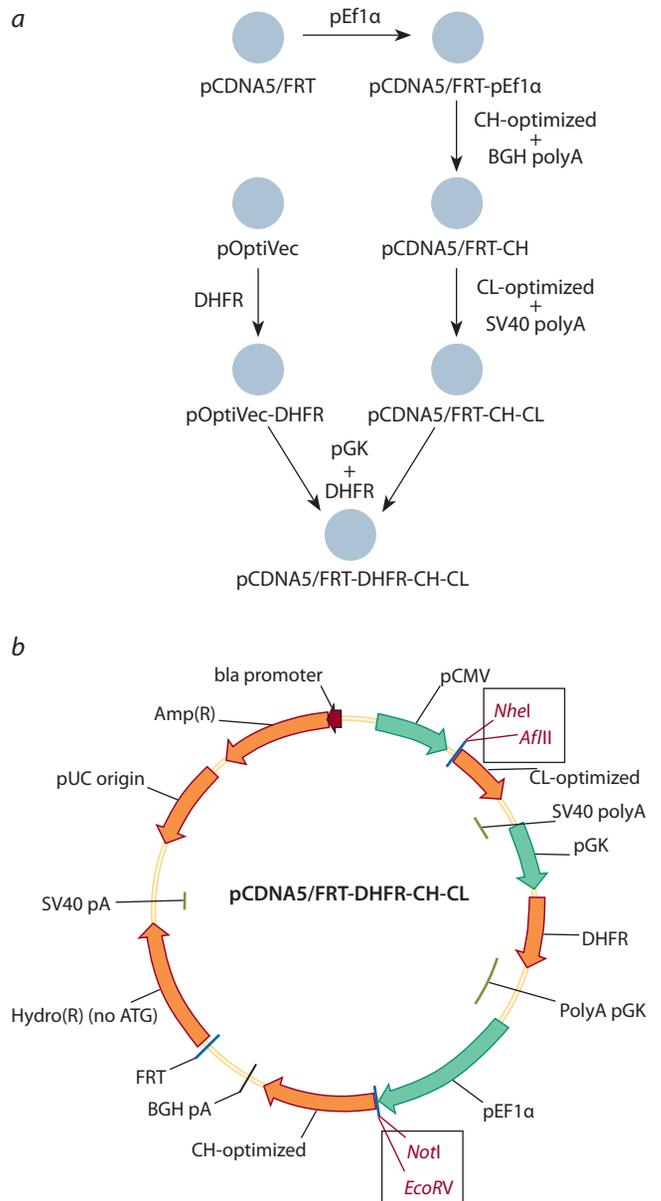


Fig. 1. (a) Schematic presentation of the construction of the cassette vector plasmid pCDNA5/FRT-DHFR-CH-CL. (b) Map of pCDNA5/FRT-DHFR-CH-CL plasmid.

ченный ПЦР-фрагмент были обработаны эндонуклеазами *HpaI* и *BglII* и объединены в реакции лигирования. В результате была сконструирована промежуточная плазмида рOptiVec-DHFR.

На последнем этапе в плазмиду рCDNA5/FRT-CH-CL был встроен ген *dhfr*. Для этого получали ПЦР-фрагмент, содержащий промотор *Pr/pgk*, ген *dhfr* и сайт полиаденилирования гена тимидинкиназы. В ПЦР-фрагмент с помощью праймеров ввели на 5'- и 3'-концы сайты *ApaI* и *Ksp22I* соответственно. В качестве матрицы для синтеза этого ПЦР-продукта использовали промежуточную плазмиду рOptiVec-DHFR и пары праймеров рGK_APA_U20 и DHFR_L21. В результате реакции лигирования была сконструирована кассетная плазмиды рCDNA5/FRT-DHFR-CH-CL, содержащая уникальные сайты *EcoRV* и

Primers used in the study

Primer	Nucleotide sequence
Ef1 α _U17	5'-AATGGTACCCGTGAGGCTCCGGTGCC-3'
Ef1 α _L22	5'-TGGGGATCCTCACGACACCTGAAATGGAAGA-3'
CH_U22	5'-AATCTCGAGTCATCTGCCAGTACCAAAGGCC-3'
BGH_polyA_L18	5'-AATCTCGAGCCATAGAGCCCACCGCAT-3'
CL_U21	5'-GCAAAGCTTCAAGAGGACTGTAGCAGCTCC-3'
SV40_polyA_L30	5'-GCAAAGCTTAACCTGTTTATTGCAGCTTATAATGGTTAC-3'
pGK_U20	5'-GGGTTGTTAACGAATTCTACCGGGTAGGGGA-3'
pGK_L17	5'-CCCAAAGATCTGCAGGTCGAAAGGCC-3'
pGK_APA_U20	5'-ATTGGGCCAATTCTACCGGGTAGGGGA-3'
DHFR_L21	5'-GCCTGATCATAGCTAGAGGTGCAGCGTAAT-3'
VH8E_U24	5'-GGGATATCGCCGCCACCATGGCTTGGGTGTGGACCTTGCTA-3'
VH8E_L45	5'-AGCGGCCGCGTACTGGCTTGGCCCCAGTAGTCAAAGTAGCCTC-3'
VL8E_L46	5'-GAAGCTAGCTGCTACAGTCTGGCCAAAAGTGTACGGGGTACTGTA-3'
VL8E_U45	5'-CCTTAAGAGCCACCATGAAGTCACAGACCCAGGTCTTCGTATTTC-3'

NotI, а также *NheI* и *AflIII* соответственно, для клонирования генов, кодирующих вариабельные домены тяжелых и легких цепей антитела человека.

Для наработки полноразмерного антитела человека fh8E (fully human 8E) в сконструированную плазмиду pCDNA5/FRT-DHFR-CH-CL встроили фрагменты ДНК, кодирующие V_H и V_L домены фагового антитела 8E. Эти фрагменты ДНК были синтезированы в ПЦР с помощью фагмидной ДНК рНЕН-8E, полученной ранее (Khlusevich et al., 2014б), и праймеров VH8E_U24 и VH8E_L45, содержащих на 5'-концах сайты эндонуклеаз рестрикции *EcoRV* и *NotI* соответственно для встраивания V_H-гена, а также праймеров VL8E_L46 и VL8E_U45, содержащих на 5'-концах сайты эндонуклеаз рестрикции *NheI* и *AflIII* соответственно (см. таблицу) для встраивания V_L-гена. В результате была получена плазида pCDNA5/FRT-DHFR-full_8E, кодирующая тяжелые и легкие цепи полноразмерного антитела fh8E. Плазмидную ДНК нарабатывали в клетках *E. coli*, очищали с помощью набора PureYield Plasmid Midiprep (Promega) и использовали для последующей трансфекции эукариотических клеток.

Для получения штамма, продуцирующего антитело fh8E, суспензионные клетки линии CHO-S/FRT одновременно трансфицировали двумя плазмидами: вспомогательной плазмидой pOG44 (Life Technologies) с геном, кодирующим флиппазу, и сконструированной плазмидой pCDNA5/FRT-DHFR-full_8E, кодирующей целевое антитело fh8E. Трансфекцию проводили с помощью трансфектанта Lipofectamine 2000 (Life Technologies) согласно рекомендациям производителя. Эффективность гомологичной рекомбинации оценивали через 72 ч после трансфекции методом проточной цитофлуориметрии с использованием цитофлуориметра NovoCyte (ACEA).

Полученные пулы клеток в концентрации 0.5×10^6 кл./мл засеяли в селективную среду CD OptiCHO, содержащую 80 мкг/мл селективного антибиотика Гигромицин В и все необходимые ростовые добавки. В выживших клетках проводили амплификацию целевого гена. Для этого

к аликватам клеток добавляли метотрексат (TEVA, Израиль) в концентрациях 200, 400, 800 и 1600 нМ. После этого отдельные клоны получали методом предельных разведений.

Для очистки антитела fh8E культуральную среду, содержащую это антитело, центрифугировали 10 мин при 12000 g. Хроматографическую полипропиленовую колонку (QIAgen), содержащую белок А-сефарозу CL-4B (GE Healthcare), предварительно уравновешивали фосфатно-солевым буферным раствором (ФСБР) в составе: 100 mM NaCl, 50 mM Na₂HPO₄ при pH 7.4. Супернатант наносили на колонку со скоростью 0.5 мл/мин при температуре 4 °C. Колонку промывали тремя объемами ФСБР и элюировали антитело fh8E четырьмя объемами 0.1 М цитратного буфера, pH 3.0. К элюату добавляли 1 М буфер трис-HCl, pH 8.8, в соотношении 1:10 для изменения pH до физиологических значений. Раствор антитела fh8E концентрировали с помощью фильтров Amicon Ultra-4 30K (Millipore) и заменяли буфер на ФСБР с добавлением 0.05 % азида натрия. Концентрацию белка в препарате определяли спектрофотометрически, проводя измерение при длине волны 280 нм.

Для выявления иммуноглобулинов человека в культуральной жидкости эукариотических клеток, продуцирующих антитело fh8E, в 96-луночные иммунологические планшеты («Медполимер», Россия) сорбировали по 100 нг поликлональных антител козы («БиоСан», Россия) против Fab-фрагмента IgG человека в 0.1 М карбонатном буфере, pH 9.4. После блокировки мест неспецифического связывания 3 % раствором бычьего сывороточного альбумина (БСА) в ФСБР, pH 7.5, в лунки добавляли 5-кратные разведения супернатантов клеток, продуцирующих антитело fh8E, и инкубировали их в течение 1 ч. Связавшиеся антитела выявляли поликлональными антителами козы («БиоСан», Россия) против Fc-фрагмента IgG человека, конъюгированными с щелочной фосфатазой (Sigma Aldrich). В качестве хромогена использовали пара-нитрофенилфосфат (Sigma, США).

Для анализа специфичности антитела fh8E в 96-луночные иммунологические планшеты («Медполимер», Россия) сорбировали по 100 нг ВОВ-LIVP-GFP в ФСБР, рН 7.5. После блокировки мест неспецифического связывания 3 % раствором БСА в ФСБР в иммунологические планшеты добавляли двукратные разведения антитела fh8E (исходная концентрация 20 мкг/мл) в ФСБР, содержащем 0.1 % Твин 20. Антитела, связавшиеся с вирусными антигенами, выявляли поликлональными антителами козы («БиоСан», Россия) против Fc-фрагмента IgG человека, конъюгированными со щелочной фосфатазой (Sigma Aldrich). В качестве хромогена использовали пара-нитрофенилфосфат (Sigma, США). Контролем неспецифического связывания служило химерное антитело ch14D5 против гликопротеина Е вируса клещевого энцефалита, содержащее константные домены, аналогичные fh8E (Baykov et al., 2014).

Для проведения вестерн-блот анализа лизаты ВОВ-LIVP-GFP или *E. coli* разделяли электрофоретически в 12 % полиакриламидном геле с 0.1 % додецилсульфатом натрия и переносили на нитроцеллюлозную мембрану (Sigma, США). Места неспецифического связывания на мембране насыщали 5 % раствором сухого молока в ФСБР, затем мембрану инкубировали 1 ч при 37 °С с 1 мкг антитела в 0.1 % Твин 20 в ФСБР. Связавшиеся антитела проявляли поликлональными антителами козы против антител человека (Sigma Aldrich, США), конъюгированными со щелочной фосфатазой, в разведении 1:12000 в 0.1 % растворе Твин 20 в ФСБР. Визуализацию иммунных комплексов проводили, добавляя 0.5 мг 5-бromo-3-индолил фосфата (Sigma, США) и 1.0 мг нитро-тетразолиевого синего (Roche).

Константу связывания антитела fh8E с белком p35delta12 (Хлусевич и др., 2014а) определяли на оптическом биосенсоре ProteOn XPR 36 (Bio-Rad). Измерения вели в ФСБР с добавлением 0.005 % Твин 20. Белок *A. Staphylococcus aureus* ковалентно иммобилизовали на поверхность чипа GLC до уровня около 2000 ед. отклика. Антитело fh8E иммобилизовали на поверхность вертикальных каналов L1 и L2 до уровня около 1500 ед. отклика (RU). Последовательные двукратные разведения белка p35delta12 (Хлусевич и др., 2014а), начиная с концентрации 500 мМ, пропускали через биосенсор в горизонтальной ориентации со скоростью потока 25 мкл/мин. Связывание длилось 10 мин, диссоциация комплексов – 30 мин. Для вычисления константы аффинности проводили глобальный анализ данных в программе ProteOnManager v.3.1.0 (Bio-Rad) с использованием модели односайтового связывания.

Для исследования вируснейтрализующей активности антитела fh8E *in vitro* суспензию ВОВ разводили до уровня 320 БОЕ/мл и смешивали с равным объемом соответствующего разведения fh8E. Смеси антитела с вирусом инкубировали в течение 1 ч при 37 °С, наносили на монослой клеток линии *Vero E6* в 24-луночные планшеты (Orange Scientific, Бельгия) и инкубировали 1 ч при 37 °С. После этого смеси удаляли и в лунки вносили среду DMEM, содержащую 2 % эмбриональную сыворотку теленка и 0.24 % агар. Через 3–4 дня монослой клеток окрашивали 0.1 % раствором кристаллического фиолетового с 10 % формалином, визуализируя бляшки. Титр нейтрализации

считали по формуле $N = (V_0 - V_n) / V_0 \times 100 \%$, где V_0 и V_n – количество бляшек в контрольных и тестируемых лунках соответственно.

Результаты и обсуждение

Полноразмерные антитела человека обладают большим терапевтическим потенциалом, поэтому в настоящее время сывороточные препараты повсеместно заменяют на препараты на основе рекомбинантных антител. В частности, для замены VIG разрабатывают противоортопоксвирусные полноразмерные человеческие или химерные антитела (McCausland et al., 2010; Crickard et al., 2012; Tikunova et al., 2012; Matho et al., 2015). Как правило, для исследования противовирусных свойств полученных антител используют экспериментальную наработку антител в системе транзientной экспрессии при одновременной трансфекции эукариотических клеток двумя плазмидами, каждая из которых несет ген, кодирующий тяжелую или легкую цепи целевого антитела. При этом эффективность получения трансфектантов, несущих обе плазмиды, довольно низка, и продуктивность получаемых целевых клеток варьирует от 0.1 до 5 мг/л.

Для создания стабильного штамма, продуцирующего полноразмерное человеческое антитело против ортопоксвирусов, на основе коммерчески доступных плазмид pCDNA5/FRT и pbudCE4.1, а также сконструированных нами ранее плазмид pCH и pCL (Байков и др., 2013) была создана кассетная плаزمида pCDNA5/FRT-DHFR-CH-CL, содержащая сайт гомологичной рекомбинации FRT и гены, кодирующие константные домены тяжелой и легкой цепей полноразмерного антитела человека класса IgG1/kappa (см. рис. 1). ДНК плазмиды pCDNA5/FRT-DHFR-CH-CL сконструировали таким образом, чтобы перед последовательностями, кодирующими константные домены тяжелых и легких цепей антитела человека, находились сайты рестрикции для удобного встраивания фрагментов ДНК, кодирующих соответствующие варибельные домены тяжелых и легких цепей. Затем в плазмиду pCDNA5/FRT-DHFR-CH-CL были встроены последовательности ДНК, кодирующие варибельные домены тяжелой и легкой цепей антитела человека против ортопоксвирусного белка p35. Эти последовательности ДНК были отобраны ранее в результате фагового дисплея комбинаторной иммунной библиотеки одноцепочечных антител человека, сконструированной нами ранее на основе РНК периферических лимфоцитов доноров, вакцинированных вирусом осповакцины (Tikunova et al., 2012). В результате биопэннинга этой библиотеки и последующей селекции фаговых антител, специфически связывающих вирус экстремелии, были отобраны несколько одноцепочечных антител (scFv) (Хлусевич и др., 2014б), представляющих собой варибельные домены тяжелой и легкой цепей, объединенных гибким пептидным линкером в единую полипептидную последовательность. Одно из отобранных антител, scFv-8E, послужило источником варибельных доменов легкой и тяжелой цепи сконструированного антитела. В результате встраивания генов, кодирующих варибельные домены scFv-8E, в кассетную плазмиду pCDNA5/FRT-DHFR-CH-CL была сконструирована плазмиды pCDNA5/FRT-DHFR-full_8E,

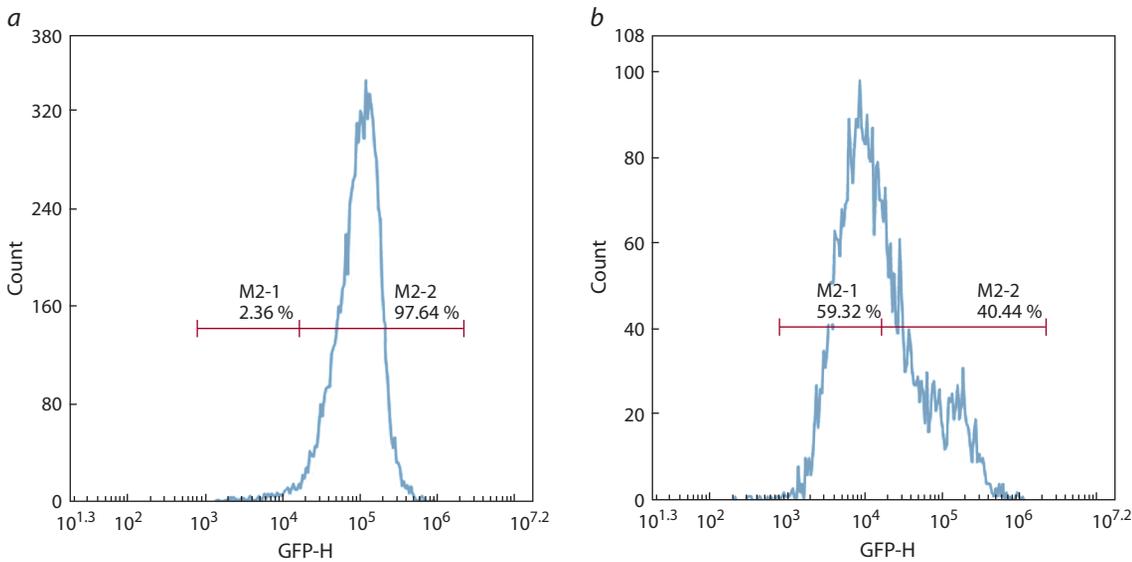


Fig. 2. Changes in GFP-fluorescence in CHO-S/FRT cells recorded by flow-cytometry:

(a) the population of ancestral nontransfected CHO-S/FRT cells; (b) CHO-S/FRT cells 48 h after transfection with pOG44 and pCDNA5\FRT-DHFR-full_8E plasmids used in the ratio 14 : 1.

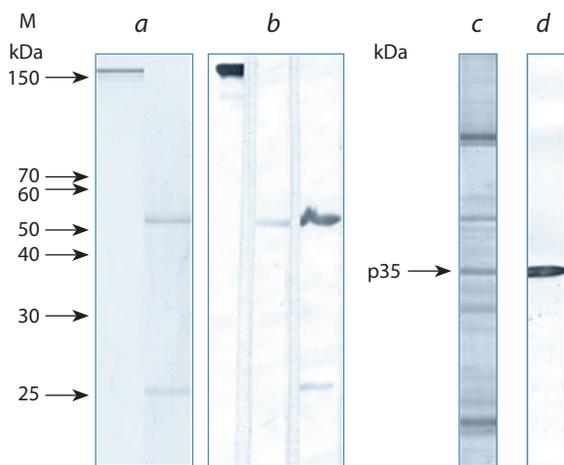


Fig. 3. Analysis of the purity and specificity of recombinant antibody fh8E:

(a) Electrophoretic image of the purified antibody fh8E under nonreducing and reducing conditions. (b) Western blot analysis of purified antibody fh8E revealed by polyclonal goat antibodies against the Fc fragment of human IgG conjugated with alkaline phosphatase and against the whole human IgG molecule (H+L) conjugated with alkaline phosphatase. (c) VACV-LIVP-GFP proteins fractionated by PAGE. (d) Western blot analysis of VACV-LIVP-GFP proteins fractionated by PAGE and visualized with full-size human antibody fh8E. M, molecular weight ladder #26614 (Fermentas, Lithuania).

несущая гены тяжелой и легкой цепей полноразмерного антитела fh8E.

Для создания стабильного штамма-производителя полноразмерного антитела человека fh8E суспензионную клеточную линию CHO-S/FRT одновременно трансфицировали коммерческой плазмидой pOG44 и сконструированной плазмидой pCDNA5\FRT-DHFR-full_8E. При этом транс-

фекцию клеток проводили в четырех вариантах с использованием различных массовых соотношений плазмид pOG44 и pCDNA5\FRT-DHFR-full_8E: 19:1, 14:1, 9:1 и 4:1.

Поскольку FRT-сайт в геноме клеточной линии CHO-S/FRT расположен между стартовым кодоном ATG и слитым белком GFP-Zeo, встраивание в геном плазмиды, содержащей FRT-сайт, инактивирует слитый белок GFP-Zeo. Таким образом, стабильные клеточные линии, в которых произошла гомологичная рекомбинация по FRT-сайту, могут быть отобраны по чувствительности к зеоцину и по отсутствию продукции белка GFP. Следовательно, у трансформированных клеток со встройкой в FRT-сайт должно наблюдаться уменьшение флуоресценции в зеленой области спектра. В связи с этим флуоресценцию пулов трансфицированных клеток анализировали на проточном цитофлуориметре по сравнению с флуоресценцией исходных нетрансфицированных клеток (рис. 2). Уровень уменьшения флуоресценции варьировал от 6 % (соотношение плазмид pOG44 и pCDNA5\FRT-DHFR-full_8E было 4:1) до 70 % (соотношение 14:1).

Пулы, содержащие наименьшее число клеток, экспрессирующих GFP, использовали для дальнейшей селекции. Для этого клетки высевали на селективную среду с гиромидином В и повторно анализировали по снижению уровня флуоресценции в зеленой части спектра с помощью проточного цитофлуориметра. Аликвоты, содержащие клетки отобранных клонов, замораживали, а оставшиеся части культивировали в течение 21 дня в среде CD OptiCHO со 100 мкг/мл селективного антибиотика зеоцина для подтверждения точности гомологичной рекомбинации. При обнаружении клонов, чувствительных к зеоцину, клетки из соответствующих пробирок размораживали и подвергали процедуре амплификации генов.

Для амплификации генов клетки, продуцирующие антитело fh8E, высевали в концентрации 5×10^5 кл./мл в

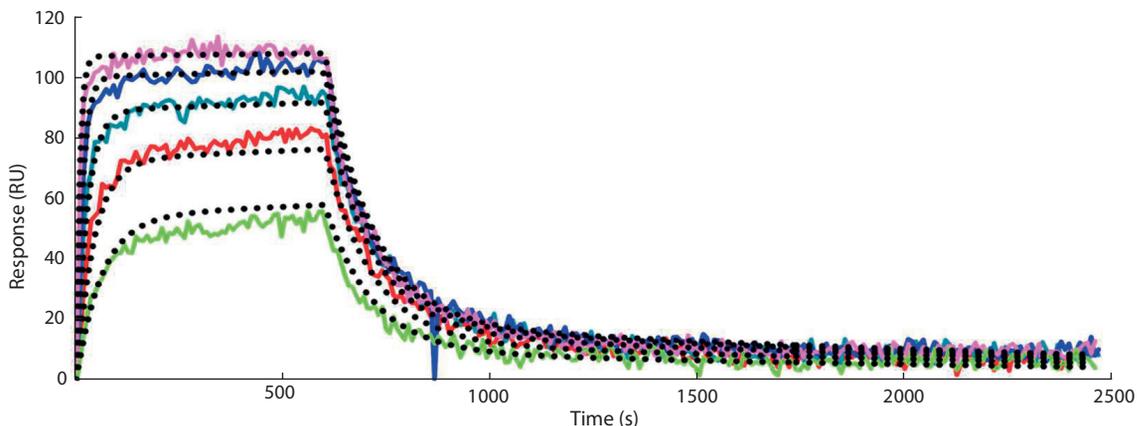


Fig. 4. Measurement of the affinity constant of the binding of fh8E with recombinant protein p35delta12.

Antibody fh8E was immobilized on the surface of a GLC chip; p35delta12 was used as the analyte at concentrations 31.2 nM, 62.5 nM, 125 nM, 250 nM, and 500 nM. The predicted curves are dotted.

селективную среду с гиромидином В в четырех повторах. В каждый повтор добавляли 200, 400, 800 и 1600 нМ метотрексата (МТХ) (TEVA, Израиль). После культивирования уровень амплификации оценивали по продуктивности полученных пулов клеток. Продуктивность составила (мг/л):

Исходный пул клеток	200 нМ МТХ	400 нМ МТХ	800 нМ МТХ	1600 нМ МТХ
6.2	5.3	9.6	15.2	11.3

Затем методом предельных разведений были получены и проанализированы по 96 отдельным клоном для пулов «800 нМ МТХ» и «1600 нМ МТХ». Наибольшая продуктивность отобранных клонов составила 110.2 мг/л.

Полученный клон охарактеризовали по стабильности экспрессии, для чего его растили в течение 35 пассажей, оценивая продуктивность. После 35 пассажей продуктивность одного из клонов составила 102.1 мг/л.

Следует отметить, что при создании штамма-продуцента важно, чтобы целевой белок обладал правильной конформацией, обеспечивающей сохранение его иммунохимических свойств. В случае противовирусных антител существенными характеристиками являются аффинность и наличие противовирусных свойств. Для оценки иммунохимических свойств антитела fh8E проводили аффинную очистку антитела (рис. 3, а), подтверждали принадлежность константных доменов к константным доменам человека (рис. 3, б), а также тестировали способность антитела fh8E выявлять ортопоксвирусный белок р35 в вестерн-блот анализе (рис. 3, в, г).

Константу аффинности полноразмерного антитела fh8E измеряли методом поверхностного плазмонного резонанса по связыванию с рекомбинантным ортопоксвирусным белком р35delta12, сконструированным ранее (Хлусевич и др., 2014а). Ранее было продемонстрировано, что белок р35delta12 специфически выявляется сыворотками доноров, вакцинированных вирусом осповакцины, с помощью ELISA и вестерн-блот анализа, а также конкурирует с вирусом осповакцины за связывание с моноклональным антителом против ортопоксвирусов (Хлусевич и др., 2014а). Измерение динамических констант показало, что константы ассоциации и диссоциации составили:

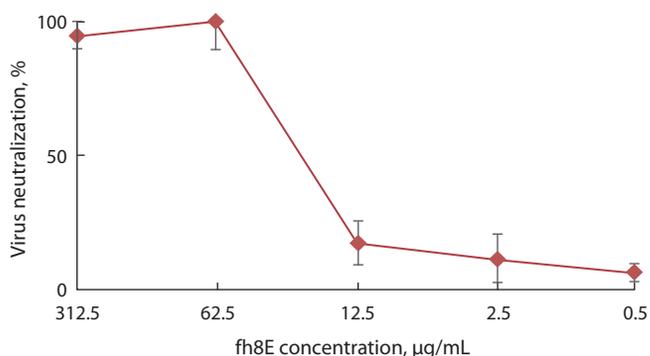


Fig. 5. Analysis of the neutralizing activity of serial dilutions of fh8E against VACV-LIVP-GFP taken at the dilution 320 PFU/mL.

Results of two independent experiments performed in three replications are shown.

$k_{on} = 4.9 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ и $k_{off} = 5.6 \times 10^{-3} \text{ s}^{-1}$ соответственно; равновесная константа аффинности антитела fh8E равна 12 нМ (рис. 4).

Противовирусные свойства антитела fh8E протестировали в реакции подавления бляшкообразования ВОВ *in vitro*. В эксперименте антитело fh8E последовательно разводили в ФСБР (рН 7.2) с шагом 1:5, начиная с концентрации 312 мкг/мл. Антитело fh8E продемонстрировало способность нейтрализовать инфекционность ВОВ дозозависимым образом (рис. 5). Титр нейтрализации IC_{50} определяли как разведение антитела, при котором наблюдалось 50 % уменьшение количества бляшек. В результате IC_{50} составил около 25 мкг/мл.

Заключение

В результате проведенных экспериментов была разработана оригинальная экспрессионная система на основе направленной рекомбинации (генный таргетинг) для интеграции гена, кодирующего полноразмерное антитело человека, в транскрипционно активную область генома эукариотических суспензионных клеток CHO-S/FRT. С помощью разработанной экспрессионной системы получен

стабильный клон, обеспечивающий продукцию полно-размерного антиортопоксвирусного антитела человека с продуктивностью около 100 мг/л. Очищенное целевое антитело обладало наномолярной аффинностью и продемонстрировало вируснейтрализующую активность *in vitro*.

Следует отметить, что достигнутый уровень продуктивности в настоящее время не является достаточным для биотехнологического производства рекомбинантного полноразмерного антитела – для коммерческой привлекательности продуктивность должна превышать 1 г/л. Вместе с тем достигнутый уровень экспрессии обеспечивает получение целевого полноразмерного антитела в количествах, достаточных для выполнения экспериментов *in vitro* и *in vivo* и проведения доклинических испытаний. При разработке противовирусных антител это позволяет в короткие сроки оценить перспективность дальнейшей разработки терапевтического препарата на основе исследуемого антитела.

Acknowledgments

This study was supported by the Russian Science Foundation, project 16-14-00083. The maintenance of the strain producing multisubunit proteins was supported by the Fundamental Research Program of the Siberian Branch of the RAS, project 0309-2016-0003.

Conflict of interest

The authors declare no conflict of interest.

References

Baykov I.K., Khlusevich Ya.A., Matveev A.L., Tikunova N.V. Construction of cassette vector plasmids for production of full-size recombinant antibodies. *Vestnik NGU: Seriya biologicheskaya, klinicheskaya meditsina* = Herald of the Novosibirsk State University. Series: Biology and clinical medicine. 2013;11(3):56-64. (in Russian)

Baykov I.K., Matveev A.L., Stronin O.V., Ryzhikov A.B., Matveyev L.E., Kasakin M.F., Richter V.A., Tikunova N.V. A protective chimeric antibody to tick-borne encephalitis virus. *Vaccine*. 2014; 32(29):3589-3594.

Coroadinha A.S., Schucht R., Gama-Norton L., Wirth D., Hauser H., Carroon M.J. The use of recombinase mediated cassette exchange in retroviral vector producer cell lines: predictability and efficiency by transgene exchange. *J. Biotechnol.* 2006;124(2):457-468.

Crickard L., Babas T., Seth S., Silvera P., Koriazova L., Crotty S. Protection of rabbits and immunodeficient mice against lethal poxvirus infections by human monoclonal antibodies. *PLoS ONE*. 2012; 7(11):e48706.

Fernandez J., Yaman I., Merrick W.C., Koromilas A., Wek R.C., Sood R., Hensold J., Hatzoglou M. Regulation of internal ribosome entry site-mediated translation by eukaryotic initiation factor-2 α phosphorylation and translation of a small upstream open reading frame. *J. Biol. Chem.* 2002;277(3):2050-2058.

Hirata R., Chamberlain J., Dong R., Russell D.W. Targeted transgene insertion into human chromosomes by adeno-associated virus vectors. *Nat. Biotechnol.* 2002;20:735-738.

Hopkins R.J., Lane J.M. Clinical efficacy of intramuscular vaccinia immune globulin: a literature review. *Clin. Infect. Dis.* 2004;39(6): 819-826.

Huang Y., Li Y., Wang Y.G., Gu X., Wang Y., Shen B.F. An efficient and targeted gene integration system for high-level antibody expression. *J. Immunol. Methods*. 2007;322(1-2):28-39.

Kameyama Y., Kawabe Y., Ito A., Kamihira M. An accumulative site-specific gene integration system using Cre recombinase-mediated cassette exchange. *Biotechnol. Bioeng.* 2010;105(6):1106-1114.

Khlusevich Ya.A., Morozova V., Pyshnyi D.V., Tikunova N.V. Antibodies against ectromelia virus capable of neutralizing variola virus: generation and application for epitope mapping. *FEBS J.* 2013; 280(S1):371-372.

Khlusevich Ya.A., Tikunova N., Morozova V., Bulychiev L., Bormotov N.,

Vlasov V., Sergeev A. Sredstvo dlya neytralizatsii virusa natural'noy ospy [The agent for neutralizing variola virus]. Patent RF No. 2515905, 20146. (in Russian)

Khlusevich Ya., Tikunova N., Morozova V., Grigor'eva A., Baykov I., P'yankov O. Rekombinantnaya plazmidnaya DNK pQE-p35d, obespечivayushchaya sintez rekombinantnogo belka p35d virusa ospy korov, shtamm bakteriy *Escherichia coli* – produktent rekombinantnogo belka r35d virusa ospy korov i rekombinantnyy belok r35d virusa ospy korov, ispol'zuemyy dlya sozdaniya test-sistem i konstruirovaniya sub'edinichnykh vaksin protiv ortopoksvirusnykh infektsiy [Recombinant plasmid DNA pQE-p35d providing synthesis of p35d recombinant protein of cowpox virus, *Escherichia coli* bacterial strain that is producer of p35d recombinant protein of cowpox virus and p35d recombinant protein of cowpox virus used to engineer test systems and to design orthopoxvirus split vaccines]. Patent RF No. 2511037, 2014a. (in Russian)

Kito M., Itami S., Fukano Y., Yamana K., Shibui T. Construction of engineered CHO strains for high-level production of recombinant proteins. *Appl. Microbiol. Biotechnol.* 2002;60(4):442-448.

Kwaks T.H., Barnett P., Hemrika W., Siersma T., Sewalt R.G., Satijn D.P., Brons J.F., van Blokland R., Kwakman P., Kruckeberg A.L., Kelder A., Otte A.P. Identification of anti-repressor elements that confer high and stable protein production in mammalian cells. *Nat. Biotechnol.* 2003;21(5):553-558.

Little M., Breitling F., Dübel S., Fuchs P., Braunagel M., Seehaus T., Klewinghaus I. Universal antibody libraries on phage and bacteria. *Year Immunol.* 1993;7:50-55.

Lucas B.K., Giere L.M., DeMarco R.A., Shen A., Chisholm V., Crowley C.W. High-level production of recombinant proteins in CHO cells using a dicistronic DHFR intron expression vector. *Nucleic Acids Res.* 1996;24(9):1774-1779.

Matho M.H., Schlossman A., Meng X., Benhnia M.R., Kaefer T., Buller M., Doronin K., Parker S., Peters B., Crotty S., Xiang Y., Zajonc D.M. Structural and functional characterization of anti-A33 antibodies reveal a potent cross-species orthopoxviruses neutralizer. *PLoS Pathog.* 2015;11.

McCausland M.M., Benhnia M.R., Crickard L., Laudenslager J., Granger S.W., Tahara T., Kubo R., Koriazova L., Kato S., Crotty S. Combination therapy of vaccinia virus infection with human anti-H3 and anti-B5 monoclonal antibodies in a small animal model. *Antivir. Ther.* 2010;15(4):661-675.

Mutskov V., Felsenfeld G. Silencing of transgene transcription precedes methylation of promoter DNA and histone H3 lysine 9. *EMBO J.* 2004;23(1):138-149.

Petrov I.S., Goncharova E.P., Kolosova I.V., Pozdnyakov S.G., Shchelkunov S.N., Zenkova M.A., Vlasov V.V. Antitumor effect of the L1VP-GFP recombinant vaccinia virus. *Doklady RAN = Proceedings of the Russian Academy of Sciences.* 2013;451(5):592-597. DOI 10.7868/S0869565213240274. (in Russian)

Reichert J.M. Antibodies to watch in 2016. *MAbs.* 2016;8(2):197-204.

Reichert J.M. Antibodies to watch in 2017. *MAbs.* 2017;9(2):167-181.

Richards E.J. Chromatin methylation: who's on first? *Curr. Biol.* 2002; 12(20):694-695.

Tasic B., Miyamichi K., Hippenmeyer S., Dani V.S., Zeng H., Joo W., Zong H., Chen-Tsai Y., Luo L. Extensions of MADM (Mosaic Analysis with Double Markers) in mice. *PLoS ONE.* 2012;7(3):e33332. DOI 10.1371/journal.pone.0033332.

Tikunova N., Dubrovskaya V., Morozova V., Yun T., Khlusevich Y., Bormotov N., Laman A., Brovko F., Shvalov A., Belanov E. The neutralizing human recombinant antibodies to pathogenic Orthopoxviruses derived from a phage display immune library. *Virus Res.* 2012;163(1):141-150.

Wilson C.J., Guglielmo C., Moua N.D., Tudor M., Grosveld G., Young R.A., Murray P.J. Yeast artificial chromosome targeting technology: an approach for the deletion of genes in the C57BL/6 mouse. *Anal. Biochem.* 2001;296(2):270-278.

Wirth D., Gama-Norton L., Riemer P., Sandhu U., Schucht R., Hauser H. Road to precision: recombinase-based targeting technologies for genome engineering. *Curr. Opin. Biotechnol.* 2007;18(5): 411-409.

Zahn-Zabal M., Kobr M., Girod P.A., Imhof M., Chatellard P., de Jesus M., Wurm F., Mermoud N. Development of stable cell lines for production or regulated expression using matrix attachment regions. *J. Biotechnol.* 2001;87(1):29-42.

Алфавитный указатель авторов статей, опубликованных в журнале в 2017 г.

- Агафонов А.В. 1, 135
Адолина И.Г. 6, 663
Айжаркын К.С. 1, 128
Айтназаров Р.Б. 7, 778
Алексеев Б.Я. 8, 932
Алексеев Я.И. 1, 124
Алехина Т.А. 7, 798
Алехна А.И. 5, 506
Аменов А.А. 1, 128
Амирханова Н.Т. 6, 677
Амстиславский С.Я. 5, 561; 6, 646; 7, 810; 8, 937
Анашкина А.А. 8, 887
Анисимова И.Н. 1, 83
Антонова О.Ю. 1, 51
Асбаганов С.В. 1, 135
Астахова Т.А. 5, 576
Аутеншлюс А.И. 3, 312
Афанасенко О.С. 1, 51
Афонников Д.А. 1, 96; 8, 895
- Б**
Бабкин И.В. 8, 993
Бадаева Е.Д. 2, 241; 6, 663
Бажан Н.М. 7, 833
Базаров С.М. 8, 959
Базарова Н.Э. 3, 374
Бай Х. 5, 581
Байков И.К. 8, 993
Банзаракцаева Т.Г. 8, 959
Баричева Э.М. 6, 710
Бархутова Д.Д. 8, 959
Батяновский А.В. 8, 887
Бебякина И.В. 6, 663
Белавин П.А. 8, 986
Белецкий А.В. 2, 234; 3, 334
Белова Ю.Н. 3, 374
Беляева Е.В. 5, 576
Бениаминов А.Д. 8, 920
Берман Д.И. 3, 374; 5, 588
Бессчетнов В.П. 2, 198
Бессчетнова Н.Н. 2, 198
Бёрнер А. 3, 341
Бидау К.Х. 4, 443
Блинов А.Г. 8, 943
Богачев С.С. 6, 717
Болоболова Е.У. 6, 710
Бондарь И.А. 8, 952
Борис К.В. 5, 596
Бородин П.М. 2, 259; 4, 385
Брагин А.О. 5, 581; 8, 911
Брусенцев Е.Ю. 5, 561; 6, 646; 7, 810; 8, 937
Брянская А.В. 6, 630
Бугун О.В. 5, 576
Букатич Е.Ю. 7, 862
- Булыгина Е.С. 3, 368
Булынецов С.В. 2, 170
Бурляева М.О. 2, 170
Быкова И.В. 1, 96
- В**
Вавилова В.Ю. 8, 943
Вараксин Н.А. 6, 717
Васько Н.И. 5, 537
Ваганабе Н. 2, 222
Велишаева Н.С. 1, 124
Вишнякова М.А. 2, 170
Владимирова А.В. 4, 392
Воевода М.И. 7, 778
Войцеховский М. 8, 943
Волкова О.Ю. 7, 764
Володин И.А. 4, 402
Володина Е.В. 4, 402
Волынкин В.А. 6, 608
Воронкова Е.В. 1, 42
- Г**
Гавриленко Т.А. 1, 16, 51, 83, 112
Галашева А.М. 2, 214
Ганич В.А. 6, 617
Ганичев Д.А. 8, 952
Гегентана Г. 5, 581
Герасимова С.В. 2, 250
Гербек Ю.Э. 4, 392
Гоголева С.С. 4, 402
Голованова Е.В. 3, 374; 5, 588
Гончар А.Л. 3, 296
Гончаренко Г.М. 6, 638
Гончаров Н.П. 6, 686
Гончарова Е.П. 8, 993
Гопций Т.И. 2, 189
Гордеева Е.И. 5, 545
Горчаков А.А. 7, 764
Гребенкина Н.А. 1, 104
Грунтенко Н.Е. 7, 825
Гуватова З.Г. 8, 932
Гукасян О.Н. 1, 42
Гуляева Е.И. 5, 523; 6, 671
Гунин А.В. 6, 623
- Давоян Р.О. 6, 663
Давоян Э.Р. 6, 663
Дейнеко Е.В. 8, 979, 986
Деменков П.С. 8, 911
Дементьева Н.В. 7, 770
Демидов Е.А. 6, 630
Деренко М.В. 3, 290; 8, 964
Держинский Е.А. 3, 374
Дмитриев А.А. 8, 920, 925
Добровольская О.Б. 2, 222
Долгова Е.В. 6, 717

Дорогина О.В. 1, 135; 3, 354; 5, 554
Дорогова Н.В. 6, 710
Доценко А.С. 7, 783
Драгович А.Ю. 2, 241
Дранов И.Г. 7, 816
Дременков И.В. 7, 783
Дружин А.Е. 2, 241
Друзьяка А.В. 2, 259
Дугаткин Л.А. 4, 487

Евдокимов М.Г. 5, 515
Ельцов И.В. 6, 686
Емельянов В.В. 1, 112
Емцева М.В. 5, 530
Еремина М.А. 7, 825
Ермишин А.П. 1, 16, 42
Ермолаев В.И. 5, 569; 6, 638
Ершова О.А. 5, 576
Есипова Н.Г. 8, 887

Жао Ж. 5, 581
Жао Ф. 5, 581
Жур К.В. 3, 296

Завьялов Е.Л. 7, 783, 841
Завьялова Я.Л. 7, 841
Загорская А.А. 8, 979
Запорожец В.И. 6, 638
Заседателев А.С. 8, 925
Захаренко Л.П. 7, 816
Захаров В.Г. 5, 534
Захаров И.К. 2, 269
Захаров-Гезехус И.А. 4, 477
Звягина Н.С. 5, 554
Зеленский Ю.И. 5, 515
Земцова А.Я. 6, 623
Зенкова Е.В. 8, 952
Зинченко А.С. 6, 663
Злобин Н.Е. 1, 104
Зорина З.А. 4, 421, 427
Зубанова Ю.С. 6, 663
Зубарев Ю.А. 6, 623

Иванисенко В.А. 8, 911
Иванова М.С. 6, 623
Игнатъева Е.В. 8, 895
Иголина Т.Н. 7, 810; 8, 937
Ильницкая Е.Т. 6, 617
Исмаил М.М. 3, 360

Кабиллов М.Р. 8, 959
Казанцева А.Ю. 7, 833
Казарцев И.А. 6, 671
Календарь Р.Н. 1, 128
Калинина Н.О. 1, 62
Каприн А.Д. 8, 932
Карпова Е.В. 6, 686
Карпова И.Ю. 8, 932
Кибальник О.П. 6, 651

Кизилова Е.А. 5, 561; 6, 646; 7, 810
Кикавский И.В. 5, 506
Кильчевский А.В. 3, 296
Киру С.Д. 1, 7
Киселева Е.В. 7, 848
Киселева М.В. 8, 925, 932
Климов Л.О. 6, 735
Климонтов В.В. 8, 952
Князев С.П. 5, 569; 6, 638
Кобозева Е.В. 1, 135
Коваль О.А. 7, 764
Ковнер А.В. 3, 312
Кожевникова В.В. 5, 561; 6, 646; 7, 810
Козаченко М.Р. 5, 537
Козлова Ю.Н. 8, 952
Колобова О.С. 1, 124
Колчанов Н.А. 8, 895
Комелькова М.В. 7, 783
Компанец Е.В. 5, 537
Кониченко Е.С. 3, 354
Коновалов А.А. 6, 686
Конопацкая И.Д. 8, 943
Концевая Г.В. 7, 778, 783, 841
Кораблев А.Н. 7, 758
Корнеева О.С. 2, 189
Коробан Н.В. 8, 925
Короткова А.М. 2, 250
Костерин О.Э. 2, 158; 6, 694
Костына З.А. 4, 427
Котенкова Е.В. 7, 788
Кочетков Д.В. 8, 932
Кочетов А.В. 1, 96
Кочиева Е.З. 2, 234; 3, 334
Красников А.А. 2, 222
Краснов Г.С. 8, 920
Красова Н.Г. 2, 214
Кудрявцева А.В. 8, 920, 925, 932
Кузнецов Е.Н. 8, 887
Кулемзин С.В. 7, 764
Кулигина Е.В. 7, 764
Куликов А.В. 7, 804
Куликова Е.А. 7, 804
Кулян Р.В. 5, 506
Кундас Л.А. 3, 296
Кунык Д.А. 8, 986
Купер К.Э. 7, 848

Лаврентьева Е.В. 8, 959
Лаврик И.Н. 8, 911
Лавриненко В.А. 7, 795
Лазарева Е.В. 6, 630
Лапшин М.С. 7, 783
Леонова И.Н. 7, 862
Леплина О.Ю. 6, 717
Лиманская С.В. 2, 189
Липатова А.В. 8, 925, 932
Лисачев А.П. 2, 259
Лиховской В.В. 6, 608
Локтев В.Б. 8, 986

Лузянин С.Л. 8, 943
Лукьянова Е.Н. 8, 925
Ляпунова О.А. 2, 152

Мак В.В. 7, 841
Макаркина М.А. 2, 207
Макаров В.В. 1, 62
Макарова Е.Н. 7, 833
Макарова С.С. 1, 62
Максютов Р.А. 3, 317
Маланханова Т.Б. 7, 856
Малахова А.А. 7, 856
Малиновская Л.П. 2, 259
Мальцев А.Н. 7, 788
Малюченко О.П. 1, 124
Малярчук Б.А. 3, 290; 8, 964
Маркель А.Л. 4, 492; 6, 735; 8, 911
Мартинек П. 2, 222
Мартинес П.А. 4, 443
Матвеев А.Л. 8, 993
Мензоров А.Г. 6, 701
Менибаев А.И. 5, 534
Миков Д.С. 6, 663
Минин С.Л. 3, 296
Мироненко Н.В. 1, 51
Мирошниченко Л.А. 2, 189
Михайлова Е.С. 3, 312
Можар Н.В. 2, 180
Мокроусова В.И. 5, 561; 6, 646; 7, 810
Моргунов А.И. 3, 347; 5, 515
Мордвинов В.А. 3, 312
Морозова В.В. 8, 952, 993
Морозова Е.В. 3, 341
Морозова К.Н. 7, 848
Моссэ И.Б. 3, 296
Мохамед С.Э. 3, 360
Мошкин М.П. 3, 304; 7, 757, 783, 848
Мошкин Ю.М. 7, 816, 848
Мурасева Д.С. 5, 554
Мустафин Р.Н. 6, 742
Мутерко А.Ф. 3, 323; 8, 878

Найденко С.В. 5, 561
Напримеров В.А. 5, 561; 6, 646; 7, 810; 8, 937
Наседкина Т.В. 8, 925
Наумов А.Г. 5, 537
Наумова Л.Г. 6, 617
Недолужко А.В. 3, 368
Непомнящих Т.С. 3, 317
Никитин С.В. 5, 569; 6, 638; 7, 778
Нимаев В.В. 8, 952
Новаковский Р.О. 8, 920
Новикова Т.И. 5, 554
Нуждин С.В. 2, 170
Нуштаева А.А. 7, 764
Нюшко К.М. 8, 925

Ожерельева З.Е. 2, 214
Окотруб К.А. 5, 561

Орищенко К.Е. 6, 717
Орлов Ю.Л. 5, 581; 8, 911
Орлова Е.А. 6, 686
Останин А.А. 6, 717
Отмахова Ю.С. 5, 545

Павлов В.В. 8, 952
Пахарукова М.Ю. 3, 312
Пельтек С.Е. 3, 374; 5, 588; 6, 630; 8, 969
Перепелкина О.В. 4, 421
Петрова О.М. 8, 937
Петровский Д.В. 3, 304; 7, 816, 841, 848
Пиндюрин А.В. 7, 816
Плеканчук В.С. 7, 798
Плеханова Е.С. 2, 170
Подколodная О.А. 8, 903
Подколodный Н.Л. 8, 903
Пожерукова В.Е. 3, 347
Полетаева И.И. 4, 421, 427
Полилов А.А. 3, 368
Полли П.Д. 4, 452
Полулях А.А. 6, 608
Полухович Ю.В. 1, 42
Попов А.Ю. 8, 925
Попова К.И. 2, 222
Попова Н.К. 4, 414
Потоцкая И.В. 3, 347
Поттер Е.А. 6, 717
Присяжнюк И.Е. 6, 701
Причко Т.Г. 2, 180
Прокудина О.И. 7, 798
Проскурина А.С. 6, 717
Протопопова Е.В. 8, 986
Прохорчук Е.Б. 3, 368
Пудова Е.А. 8, 925, 932
Пузаков М.В. 2, 269
Пузакова Л.В. 2, 269
Пузанов Г.А. 8, 920
Пузанский Р.К. 1, 112
Пшеничникова Т.А. 3, 341

Равин Н.В. 2, 234; 3, 334
Рагаева Д.С. 8, 937
Раднагуруева А.А. 8, 959
Радченко Е.Е. 1, 74
Расторгуев С.М. 3, 368
Рахмангулов Р.С. 5, 506
Риттер Г.С. 6, 717
Рихтер В.А. 7, 764
Рогаев Е.И. 8, 911
Рогозина Е.В. 1, 7, 30
Рожков О.А. 7, 841
Рожкова И.Н. 6, 646; 7, 810; 8, 937
Рожнова О.М. 8, 952
Розанов А.С. 6, 630; 8, 969
Ромашенко А.В. 3, 304; 7, 788, 848
Рсалиев А.С. 6, 677
Рубан А.С. 2, 241
Рувинский А.О. 4, 462

Рябичева Т.Г. 6, 717
Рязанова М.А. 6, 735; 7, 798

Сайк О.В. 8, 911
Салина Е.А. 3, 323; 6, 663; 7, 862
Самарина Л.С. 5, 506
Самохин А.Г. 8, 952
Саранина И.В. 8, 952
Святченко С.И. 5, 537
Седов Е.Н. 2, 207
Седышева Г.А. 2, 207
Селютина И.Ю. 3, 354
Серова З.М. 2, 207
Серова И.А. 7, 758
Сеферова И.В. 2, 170
Сибикеев С.Н. 2, 241
Сизова С.В. 8, 969
Симонов А.В. 3, 341
Сколотнева Е.С. 7, 862
Скрябин Б.В. 7, 758
Скрябин К.Г. 2, 234; 3, 334, 368
Слынько Н.М. 6, 630
Снежкина А.В. 8, 932
Солонечный П.Н. 5, 537; 6, 657
Сон О.М. 8, 969
Сорокин И.Е. 7, 804
Спасельникова А.В. 6, 717
Стабровская О.И. 5, 545
Старостин К.В. 6, 630
Степанова Е.В. 8, 925
Стёпочкин П.И. 5, 530
Столповский Ю.А. 4, 477
Супрун И.И. 6, 617
Сурина Н.М. 4, 427
Суялату С. 5, 581
Сюков В.В. 5, 534

Табиханова Л.Э. 5, 581
Тагиманова Д.С. 1, 128
Тальянский М.Э. 1, 62
Таран О.П. 6, 630
Таранин А.В. 7, 764
Таранов В.В. 1, 104
Таранов О.С. 3, 317
Твердохлеб Н.Н. 8, 903
Текутьева Л.А. 8, 969
Терновой В.В. 1, 104
Тикунов А.Ю. 8, 952
Тикунова Н.В. 8, 952, 993
Ткаченко А.В. 7, 764
Токмаков С.В. 6, 617
Торгашева А.А. 2, 259
Трапезов О.В. 7, 795
Трегубчак Т.В. 3, 317
Триголубов А.Н. 3, 312
Трут Л.Н. 4, 392, 402
Трущенко А.Ю. 3, 347
Туманян В.Г. 8, 887
Тычко Р.А. 8, 920
Тюнин В.А. 5, 523

У Н. 5, 581
У Х. 5, 581
У Ч. 5, 581
Уваров И.П. 7, 841
Уварова Ю.Е. 6, 630
Уилкинс А.С. 4, 435
Ульзетуева И.Д. 8, 959
Урошлев Л.А. 8, 887
Усенко Н.И. 5, 545

Фатьянова А.В. 7, 795
Федин Е.Н. 3, 312
Федорова М.С. 8, 925, 932
Федорова С.А. 7, 816
Федосеева Л.А. 6, 735
Федотова И.Б. 4, 427
Фекличева И.В. 7, 783
Филошин М.А. 2, 234; 3, 334; 5, 596
Фоменко Н.В. 8, 952

Хаблак С.Г. 2, 227
Хавкин Э.Е. 1, 30
Хапилина О.Н. 1, 128
Харламова А.В. 4, 392, 402
Хахинов В.В. 8, 959
Хлесткин В.К. 6, 728
Хлесткина Е.К. 1, 96; 2, 250; 5, 545
Хлусевич Я.А. 8, 993
Хоцкин Н.В. 7, 783, 804
Хуо В. 5, 581
Хуснутдинова Э.К. 6, 742
Хютти А.В. 1, 51

Цейликман В.Э. 7, 783
Цейликман О.Б. 7, 783
Цыганкова С.В. 3, 368

Чадаева И.В. 8, 911
Чайка А.В. 8, 932
Чалая Л.Д. 2, 180
Чен М. 5, 581
Черных Е.Р. 6, 717
Чернявский М.К. 8, 959
Чистякова А.К. 3, 341
Чурсин А.С. 3, 347

Шайдаюк Е.Л. 5, 523; 6, 671
Шалаева Т.В. 1, 124
Шаманин В.П. 3, 347
Шанина Е.П. 1, 124
Шарапова М.Б. 3, 304; 7, 848
Шарипова Д.В. 7, 856
Шарко Ф.С. 3, 368
Шарфунова И.Б. 5, 545
Шатохин К.С. 6, 638
Шашков М.П. 3, 374
Шевелев О.Б. 7, 783
Шепелев С.С. 3, 347

Шеховцов С.В. 3, 374; 5, 588
Шилов И.А. 1, 124
Шишкина А.А. 2, 241
Шишова М.Ф. 1, 112
Шляхтун В.Н. 6, 630; 8, 969
Шмаков Н.А. 1, 96
Шрейдер Е.Р. 5, 523
Штофин С.Г. 3, 312
Шульга О.А. 2, 234; 3, 334
Шумный В.К. 2, 250; 4, 387
Шундрин И.К. 6, 686

Шелкунов С.Н. 3, 317
Щенникова А.В. 2, 234; 3, 334
Щукина Л.В. 3, 341

Эрст Т.В. 6, 728

Юдин Н.С. 7, 778, 795, 816
Юсов В.С. 5, 515

Яковлев А.Ф. 7, 770
Яковлева Т.В. 7, 833
Якубицкий С.Н. 3, 317

Прием статей через электронную редакцию на сайте <http://vavilov.elpub.ru/index.php/jour>
Предварительно нужно зарегистрироваться как автору, затем в правом верхнем углу страницы выбрать «Отправить рукопись». После завершения загрузки материалов обязательно выбрать опцию «Отправить письмо», в этом случае редакция автоматически будет уведомлена о получении новой рукописи.

«Вавиловский журнал генетики и селекции»/“Vavilov Journal of Genetics and Breeding”
до 2011 г. выходил под названием «Информационный вестник ВОГиС»/
“The Herald of Vavilov Society for Geneticists and Breeding Scientists”.

Регистрационное свидетельство ПИ № ФС77-45870 выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций 20 июля 2011 г.

«Вавиловский журнал генетики и селекции» включен ВАК Минобрнауки России в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук, Российский индекс научного цитирования, ВИНТИ, базы данных Scopus, Zoological Record (Web of Science), Ebsco, DOAJ, Ulrich’s Periodicals Directory, Google Scholar, Russian Science Citation Index на платформе Web of Science.

Открытый доступ к полным текстам:
на сайте ИЦиГ СО РАН – bionet.nsc.ru/vogis/
платформе Elpub – vavilov.elpub.ru/index.php/jour
платформе Научной электронной библиотеки – elibrary.ru/title_about.asp?id=32440

Подписку на «Вавиловский журнал генетики и селекции» можно оформить в любом почтовом отделении России. Индекс издания 42153 по каталогу «Пресса России».

При перепечатке материалов ссылка на журнал обязательна.

✉ e-mail: vavilov_journal@bionet.nsc.ru

Издатель: Федеральное государственное бюджетное научное учреждение
«Федеральный исследовательский центр Институт цитологии и генетики
Сибирского отделения Российской академии наук»,
проспект Академика Лаврентьева, 10, Новосибирск, 630090.

Адрес редакции: проспект Академика Лаврентьева, 10, Новосибирск, 630090.

Секретарь по организационным вопросам С.В. Зубова. Тел.: (383)3634977.

Издание подготовлено информационно-издательским отделом ИЦиГ СО РАН. Тел.: (383)3634963*5218.

Начальник отдела: Т.Ф. Чалкова. Редакторы: В.Д. Ахметова, И.Ю. Ануфриева. Дизайн: А.В. Харкевич.

Компьютерная графика и верстка: Т.Б. Коняхина, О.Н. Савватеева.

Подписано в печать 26.12.2017. Выход в свет 30.01.2018. Формат 60 × 84 1/8. Усл. печ. л. 15,81.

Уч.-изд. л. 19,1. Тираж 200 экз. Заказ № 293. Цена свободная.

Отпечатано в типографии ФГУП «Издательство СО РАН», Морской проспект, 2, Новосибирск, 630090.