

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ НАУКИ
ИНСТИТУТ ОБЩЕЙ ГЕНЕТИКИ им. Н.И. ВАВИЛОВА
РОССИЙСКОЙ АКАДЕМИИ НАУК

На правах рукописи

ДЬЯЧКОВА МАРИНА СЕРГЕЕВНА

**СРАВНИТЕЛЬНЫЙ ЭВОЛЮЦИОННЫЙ И ФУНКЦИОНАЛЬНЫЙ АНАЛИЗ ГЕНОВ
КЛАСТЕРА PFNA У ПРЕДСТАВИТЕЛЕЙ РОДА *BIFIDOBACTERIUM***

специальность 03.02.07 – генетика

ДИССЕРТАЦИЯ

на соискание ученой степени кандидата биологических наук

Научный руководитель:
д.б.н., проф. Даниленко Валерий Николаевич

Москва – 2019

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	4
ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ.....	10
1.1 Бифидобактерии как важный компонент комменсальной микробиоты человека.....	10
1.2 Характеристика рода <i>Bifidobacterium</i>	12
1.3 Молекулярные механизмы, обеспечивающие коммуникацию представителей микробиоты с факторами занимаемой экологической ниши в организме хозяина.....	17
1.3.1 Системы сигнальной трансдукции бифидобактерий.....	19
1.3.2 Механизмы адгезии бифидобактерий.....	25
1.4 Быстрая молекулярная эволюция генов как инструмент адаптации к занимаемой экологической нише.....	27
1.4.1 Быстрая эволюция генов, участвующих в реализации механизмов антагонистических взаимодействий.....	28
1.4.2 Быстрая эволюция генов, участвующих в реализации механизмов взаимодействия с другими факторами занимаемой экологической ниши.....	30
1.4.3 Методы и подходы, используемые для детекции событий положительного отбора в молекулярной эволюции генов.....	31
ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ.	33
2.1 Анализ <i>in silico</i>	33
2.1.1 Биоинформатический анализ генетического окружения гена <i>pkb2</i> в геномах представителей различных видов рода <i>Bifidobacterium</i>	33
2.1.2 Анализ молекулярной эволюции генов кластера PFNA в геномах представителей различных видов рода <i>Bifidobacterium</i>	35
2.1.2.1 Подготовка данных для последующего анализа молекулярной эволюции генов кластера PFNA представителей различных видов рода <i>Bifidobacterium</i>	35
2.1.2.2 Филогенетический анализ, основанный на сравнении последовательностей генов кластера PFNA представителей различных видов рода <i>Bifidobacterium</i>	37
2.1.2.3 Подбор параметров анализа молекулярной эволюции и поиск событий продолжительного/ эпизодического положительного отбора.....	38
2.2 Экспериментальный анализ.....	42
2.2.1 Изучение транскрипционной организации кластера PFNA в геноме <i>B. longum</i> subsp. <i>longum</i> GT15.	42
2.2.2 Сравнительный анализ уровня экспрессии гена <i>fn3</i> у представителей различных видов рода <i>Bifidobacterium</i>	46
2.2.4 Оценка влияния IL-6 как фактора иммунного ответа на скорость роста культуры и	

экспрессию генов штамма <i>B. longum</i> subsp. <i>longum</i> GT15.....	50
2.2.5 Определение полногеномной нуклеотидной последовательности штамма <i>B. angulatum</i> GT102.....	51
ГЛАВА 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ.....	53
3.1 Биоинформатический анализ генетического окружения гена <i>pkb2</i> в геномах представителей различных видов рода <i>Bifidobacterium</i>	53
3.2 Изучение транскрипционной организации кластера PFNA в геноме <i>B. longum</i> subsp. <i>longum</i> GT15.....	61
3.3 Анализ молекулярной эволюции последовательностей генов кластера PFNA в геномах представителей различных видов рода <i>Bifidobacterium</i>	64
3.3.1 Филогенетический анализ, основанный на сравнении последовательностей генов кластера PFNA представителей различных видов рода <i>Bifidobacterium</i>	65
3.3.2 Тестирование гипотез о наличии событий продолжительного положительного отбора в молекулярной эволюции последовательностей генов кластера PFNA.....	68
3.3.3 Тестирование гипотез о наличии событий эпизодического положительного отбора в молекулярной эволюции последовательностей генов кластера PFNA.....	70
3.4 Сравнительный анализ уровня экспрессии гена <i>fn3</i> у представителей различных видов рода <i>Bifidobacterium</i>	85
3.5 Изучение видоспецифичности киназно-субстратного взаимодействия в реакции перекрестного фосфорилирования <i>in vitro</i>	87
3.6 Оценка влияния IL-6 как фактора иммунного ответа на рост культуры и экспрессию генов штамма <i>B. longum</i> subsp. <i>longum</i> GT15.....	90
ЗАКЛЮЧЕНИЕ.....	94
ВЫВОДЫ.....	98
СПИСОК ИСПОЛЬЗУЕМЫХ СОКРАЩЕНИЙ.....	99
СПИСОК ЛИТЕРАТУРЫ.....	101
ПРИЛОЖЕНИЯ.....	116
Приложение 1.	116
Приложение 2.	125

ВВЕДЕНИЕ

Бифидобактерии являются анаэробами и считаются одними из самых древних представителей актинобактерий [Gao, Parmanathan and Gupta, 2006], существовавшими во времена низкого содержания кислорода в атмосфере Земли. В процессе колонизации новой анаэробной ниши бифидобактерии смогли эффективно приспособиться к обитанию в желудочно-кишечном тракте (ЖКТ) хозяина за счет утилизации гликанов мукоидного слоя кишечного эпителия [Koropatkin, Cameron and Martens, 2012; Grimm *et al.*, 2014]. В настоящее время бифидобактерии входят в состав облигатной комменсальной микробиоты человека и других животных, включая насекомых, птиц и млекопитающих. Современные представления об экологии бифидобактерий предполагают строгую ассоциацию между видами бифидобактерий и занимаемыми экологическими нишами в организмах хозяев [D'Aimmo *et al.*, 2014; Michelini *et al.*, 2016 a, b; Modesto *et al.*, 2014; Bottacini *et al.*, 2012], что обеспечивается, вероятно, длительной коэволюцией и особыми механизмами видоспецифической адаптации, которые на сегодняшний день изучены весьма слабо.

Среди анаэробных комменсальных микроорганизмов микробиоты человека бифидобактерии являются наиболее ранними колонизаторами [Alderberth *et al.*, 2007]. Являясь доминирующей компонентой комменсальной микробиоты младенцев, бифидобактерии способствуют становлению иммунной системы человека на ранних этапах постнатального развития [Kalliomäki *et al.*, 2001; Fujimura *et al.*, 2016]. Бифидобактерии как компонент комменсальной микробиоты оказывают положительный эффект на здоровье человека и поддерживают нормальное функционирование организма хозяина. Ключевыми факторами в этом процессе, по-видимому, являются адгезионные антагонистические свойства представителей рода *Bifidobacterium*, а также способность осуществлять коммуникацию с иммунной системой, стимулируя выработку как противовоспалительных, так и провоспалительных цитокинов и других факторов иммунного ответа [Ruiz *et al.*, 2017; Аверина и др., 2015]. Однако механизмы, обеспечивающие выживание бифидобактерий в условиях воспалительного процесса и постоянство репертуара нормальной микробиоты кишечника, остаются малоизученными.

Ранее в геномах бифидобактерий обнаружены и охарактеризованы шесть СТПК, среди которых наибольший интерес представляет видоспецифическая СТПК Pkb2, демонстрирующая аномально высокий уровень дивергенции последовательностей различных видов по сравнению с другими охарактеризованными СТПК [Nezametdinova *et al.*, 2014; Zakharevich *et al.*, 2015; Алексеева и др., 2015]. Анализ субстратов фосфорилирования позволил выдвинуть предположение, что функции СТПК Pkb2 могут быть связаны с адгезией и коммуникацией бифидобактерий с эпителиальными клетками кишечника [Nezametdinova *et*

al., 2018]. Высокий уровень межвидовой дивергенции последовательностей генов может быть следствием воздействия положительного естественного отбора, направленного на сохранение мутаций, приводящих к изменению структуры кодируемых белков, и обеспечивающего быструю молекулярную эволюцию генов. Быстрая эволюция и положительный отбор нередко затрагивают гены, вовлеченные в коммуникацию с факторами внешней среды, и особенно, биотическими экологическими факторами [Brockhurst *et al.*, 2014].

Объектом изучения данной диссертационной работы является ген *pkb2*, кодирующий видоспецифическую СТПК Pkb2 в геномах различных видов бифидобактерий, а также его генетическое окружение. Ранее было показано, что генетическое окружение гена *pkb2* достаточно консервативно [Nezametdinova *et al.*, 2014]. Генетическое окружение гена *pkb2* включает в себя группу генов, совместно обнаруживаемых в геномах представителей различных видов рода *Bifidobacterium*: гены *pkb2*, *fn3*, *aaa-atp*, *duf58* и *tgm*. Кластер генов получил название PFNA (*pkb2*, *fn3*, *aaa-atp*). Гены кластера PFNA предположительно являются одной из систем, обеспечивающих коммуникацию бифидобактерий с факторами занимаемой экологической ниши, поэтому исследование их функциональной активности и особенностей эволюции является важной ступенью в понимании не только процессов адаптации и взаимодействия с организмом хозяина, но и их значения в поддержании здоровья и гомеостаза организма человека и других носителей в целом.

Цель исследования:

Целью настоящей работы является функциональная характеристика и анализ молекулярной эволюции генов кластера PFNA у представителей различных видов рода *Bifidobacterium*.

Задачи исследования:

1. Структурно-функциональная характеристика генов кластера PFNA (*pkb2*, *fn3*, *aaa-atp*, *duf58* и *tgm*) в геномах представителей различных видов рода *Bifidobacterium*.
2. Изучение транскрипционной организации кластера генов PFNA в геноме *B. longum* subsp. *longum* GT15.
3. Анализ молекулярной эволюции последовательностей генов кластера PFNA в геномах представителей различных видов рода *Bifidobacterium*.
4. Анализ уровня экспрессии гена *fn3* в геномах представителей различных видов рода *Bifidobacterium*.

5. Изучение функциональной активности и субстратной специфичности полноразмерных СТПК Pkb2 штаммов *B. longum* subsp. *longum* GT15 и *B. choerinum* DSM 20434.
6. Изучение влияния провоспалительного цитокина IL-6 на рост культуры и экспрессию генов штамма *B. longum* subsp. *longum* GT15.

Научная новизна и практическая значимость исследования:

В рамках диссертационной работы впервые была определена полногеномная последовательность штамма *B. angulatum* GT 102 в статусе «complete». Сборка генома в статусе «complete» имеет практическую значимость, так как обеспечивает формирование референсной последовательности для полногеномного секвенирования и сборки геномов новых штаммов вида *B. angulatum* и близкородственных видов бифидобактерий.

Впервые было дано объяснение феномену быстрой эволюции и высокой степени межвидовой дивергенции последовательностей генов кластера PFNA бифидобактерий, что расширяет представления о роли положительного отбора в молекулярной эволюции генов, вовлеченных в коммуникацию с факторами занимаемой экологической ниши в организме хозяина. Сайты, на которые осуществляется давление отбора, впервые были локализованы в первичной структуре белков, кодируемых генами кластера. Полученные данные могут быть использованы для расширения структурно-функциональной аннотации данных белков.

Впервые была изучена экспериментально и *in silico* транскрипционная организация кластера PFNA в геноме *B. longum* subsp. *longum* GT15. В результате анализа было установлено, что гены *pkb2*, *fn3*, *aaa-atp*, *duf58*, *tgm*, *prpC*, *fha* и ген BLGT_RS02790 в геноме *B. longum* subsp. *longum* GT15 транскрибируются в составе единого оперона, а также были определены точки старта и терминации транскрипции.

Впервые была проведена оценка влияния IL-6 как фактора иммунного ответа на рост культуры и экспрессию генов штамма *B. longum* subsp. *longum* GT15. Полученные данные могут быть использованы для изучения генов и генных сетей, участвующих во взаимодействии с иммунной системой организма хозяина.

Впервые была изучена видоспецифичность киназно-субстратного взаимодействия на примере штаммов *B. longum* subsp. *longum* GT15 и *B. choerinum* DSM 20434 в ходе киназной реакции. В результате анализа впервые экспериментально была показана функциональная активность полноразмерных СТПК Pkb2 представителей видов *B. longum* subsp. *longum* и *B. choerinum*, субстратная видоспецифичность для СТПК Pkb2 *B. longum* subsp. *longum* GT15 и ее отсутствие для СТПК Pkb2 *B. choerinum* DSM 20434.

Положения, выносимые на защиту:

1. Гены кластера PFNA (*pkb2*, *fn3*, *aaa-atp*, *duf58* и *tgm*) представляют собой эволюционно стабильную группу функционально связанных генов, обнаруживаемых совместно в геномах представителей различных видов рода *Bifidobacterium*, и транскрибируются в составе единого оперона.
2. Высокий уровень межвидовой дивергенции нуклеотидных последовательностей генов кластера PFNA обусловлен событиями продолжительного и эпизодического положительного отбора, провоцирующих быструю молекулярную эволюцию генов.
3. Быстрая эволюция последовательностей генов *pkb2* и *aaa-atp* обуславливает видоспецифичность киназно-субстратного взаимодействия у *B. longum* subsp. *longum* GT15.

Личный вклад автора:

Все основные результаты были получены лично автором, либо при его участии в планировании и проведении экспериментов. Часть экспериментов была проведена совместно с сотрудниками лаборатории Генетики микроорганизмов ИОГен им. Н. И. Вавилова РАН. В частности, эксперимент по оценке влияния провоспалительного цитокина IL-6 как фактора иммунного ответа на скорость роста культуры и экспрессию генов штамма *B. longum* subsp. *longum* GT15 был проведен совместно с к.б.н. Климиной К. М. Эксперимент по изучению видоспецифичности киназно-субстратного взаимодействия на примере штаммов *B. longum* subsp. *longum* GT15 и *B. choerinum* DSM 20434 в ходе киназной реакции был проведен совместно с к. б. н. Мавлетовой Д. А. и к. б. н. Алексеевой М. Г. на базе Лаборатории изотопных методов ИБХ им. М. М. Шемякина и Ю. А. Овчинникова РАН.

Апробация диссертационной работы:

Доклады по теме диссертации проводились на ежегодных отчетах аспирантов ФГБУН ИОГен РАН в 2014-2018 г. Автор опубликовано 4 статьи по теме диссертации в научных рецензируемых изданиях, индексируемых в базе данных Scopus и отвечающих требованиям Высшей Аттестационной Комиссии Министерства образования и науки Российской Федерации, а также 2 статьи в сборниках материалов конференций. Промежуточные и итоговые результаты диссертационной работы были представлены на российских и международных конференциях. Устный доклад был представлен на II Объединенном научном форуме (Сочи-Дагомыс, Россия, 2019). Стендовые доклады были представлены на 7-м международном съезде ПНМС (Килларни, Ирландия, 2018) и 43-м международном конгрессе FEBS Congress (Прага, Чехия, 2018). Апробация диссертационной работы

проведена на межлабораторном семинаре ИОГен РАН (протокол №15 от 29 октября 2018 г.).

Публикации по теме диссертации в научных журналах:

1. **Dyachkova, M. S.**, Klimina, K. M., Kovtun, A. S., Zakharevich, N. V., Nezametdinova, V. Z., Averina, O. V., and Danilenko, V. N. (2015). Draft genome sequences of *Bifidobacterium angulatum* GT102 and *Bifidobacterium adolescentis* 150: focusing on the genes potentially involved in the gut-brain axis. *Genome Announcement*, 3(4), e00709-15.
2. Nezametdinova, V. Z., Mavletova, D. A., Alekseeva, M. G., **Chekalina, M. S.**, Zakharevich, N. V., and Danilenko, V. N. (2018). Species-specific serine-threonine protein kinase Pkb2 of *Bifidobacterium longum* subsp. *longum*: Genetic environment and substrate specificity. *Anaerobe*, 51, 26-35.
3. Zakharevich, N. V., Nezametdinova, V. Z., Averina, O. V., **Chekalina, M. S.**, Alekseeva, M. G., Danilenko, V. N. (2019). Complete genome sequence of *Bifidobacterium angulatum* GT102: potential genes and systems of communication with host. *Russian Journal of Genetics*, 55(7), 847–864.
4. **Dyachkova, M. S.**, Chekalin, E. V., Danilenko, V. N. (2019). Positive selection in *Bifidobacterium* genes drives species-specific host-bacteria communication. *Frontiers in microbiology*, 10, 2374.

Публикации по теме диссертации в сборниках материалов конференций:

1. **M. Chekalina**, V. Danilenko. Study of the PFNA cluster in bifidobacteria: structure, evolution and possible functions. FEBS Open Bio, Volume 8, Issue S1. Supplement: 43rd FEBS Congress, Biochemistry Forever, Prague, Czech Republic, July 7-12, 2018.
2. **М. С. Чекалина**, К. М. Климина, В. Н. Даниленко. Изменение профиля экспрессии генов штамма *Bifidobacterium longum* GT15 в условиях воздействия провоспалительных цитокинов IL-6 и TNF α . Acta Naturae, Спецвыпуск, Том 1. Научные труды II Объединенного научного форума, Сочи-Дагомыс, Россия, 1-6 октября, 2019.

Участие в конференциях с докладами по теме исследования:

1. **M. S. Chekalina**, V. N. Danilenko. The molecular evolution of the PFNA gene cluster in bifidobacteria. 7th IHMC International Human Microbiome Congress, Killarney, Ireland, June 26th-28th, 2018, стендовый доклад.
2. **M. Chekalina**, V. Danilenko. Study of the PFNA cluster in bifidobacteria: structure, evolution and possible functions. 43rd FEBS Congress, Biochemistry Forever, Prague,

Czech Republic, July 7th-12th, 2018, стендовый доклад.

3. **М. С. Чекалина**, К. М. Климина, В. Н. Даниленко. Изменение профиля экспрессии генов штамма *Bifidobacterium longum* GT15 в условиях воздействия провоспалительных цитокинов IL-6 и TNF α . II Объединенный научный форум, Сочи-Дагомыс, Россия, 1-6 октября, 2019, устный доклад.

Структура и объем диссертационной работы:

Диссертация включает в себя следующие главы: «Введение», «Обзор литературы», «Материалы и методы», «Результаты и обсуждение», «Заключение», «Выводы» и «Список литературы». Работа изложена на 129 страницах машинописного текста, включая 14 страниц Приложений, содержит 30 рисунков и 19 таблиц. Список цитируемых литературных источников включает 179 наименований, из которых 177 – на английском языке.

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

1.1 Бифидобактерии как важный компонент комменсальной микробиоты человека.

Кишечной микробиотой человека называется совокупность микроорганизмов, населяющих пищеварительный тракт. Большая часть микроорганизмов кишечной микробиоты вовлечены в комменсальные или мутуалистические отношения с организмом человека. Микробиота кишечника выполняет ряд важных физиологических функций в организме хозяина, включая иммунную и моторную функции, обеспечение целостности эпителиального барьера, антагонистическую функцию, расщепление сложных полисахаридов и синтез витаминов. Микробиота кишечника способна оказывать как непосредственное воздействие на слизистую кишечника и энтеральную вегетативную нервную систему, так и метаболическое косвенное воздействие на функционирование дистальных органов и систем [Clarke *et al.*, 2014].

Среди анаэробных комменсальных микроорганизмов бифидобактерии являются наиболее ранними колонизаторами. Бифидобактерии обнаруживаются в значительном количестве у 2/3 младенцев в возрасте одной недели [Alderberth *et al.*, 2007]. В процессе дальнейшего онтогенеза численность бифидобактерий сокращается с 11% у детей в возрасте от 1 до 4 лет до 3% у взрослых [Ringel-Kulka *et al.*, 2013]. Видами, характеризующимися наибольшей численностью среди бифидобактерий кишечной микробиоты детей, являются *B. bifidum*, *B. infantis*, *B. breve* и *B. parvulorum*, взрослых – *B. adolescentis*. *B. bifidum* и *B. longum* также встречаются в обеих группах [Reuter, 2001]. Согласно другому исследованию, *B. longum* наиболее представлен в обеих группах; во взрослом возрасте существенно возрастает представленность видов *B. adolescentis* и *B. animalis* subsp. *lactis* (рис. 1) [Turroni *et al.*, 2012].

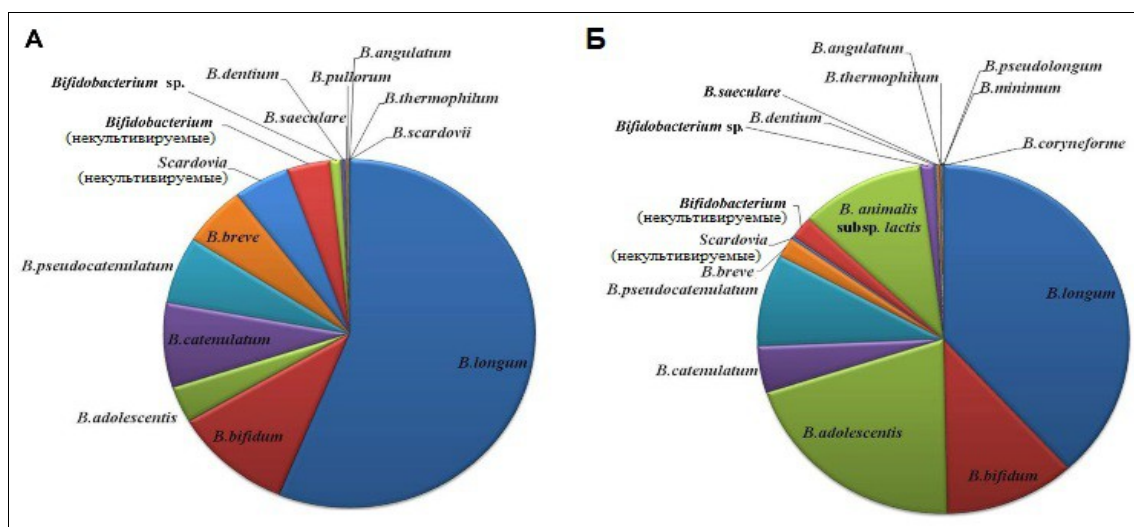


Рисунок 1. Представленность различных видов рода *Bifidobacterium* в микробиоте младенцев (А) и матерей (Б) [Turroni *et al.*, 2012, с изменениями].

Бифидобактерии обладают ярко выраженными иммуномодулирующими свойствами. Являясь доминирующей компонентой комменсальной микробиоты младенцев, бифидобактерии способствуют становлению иммунной системы человека на ранних этапах постнатального развития. Понижение уровня представленности бифидобактерий в кишечнике младенцев сопряжено с высоким риском atopических заболеваний для групп младенцев в возрасте трех недель и трех месяцев; повышение риска atopических заболеваний сохраняется до возраста 1 года [Kalliomäki *et al.*, 2001]. Похожим образом, понижение уровня представленности бифидобактерий в составе кишечной микробиоты было выявлено для младенцев в возрасте трех месяцев, у которых впоследствии развилась атопия в возрасте двух лет или астма в возрасте четырех лет [Fujimura *et al.*, 2016]. Эти данные указывают на критическую роль бифидобактерий в формировании иммунной системы в раннем возрасте. В настоящее время штаммоспецифичное иммуномодулирующее действие бифидобактерий изучено в моделях *in vitro* и *in vivo* во множестве опубликованных исследований [Ruiz *et al.*, 2017].

Одной из важнейшей функций бифидобактерий в составе комменсальной микробиоты организма хозяина является способность проявлять антагонистические свойства по отношению к патогенным микроорганизмам. Антагонистические свойства могут быть обусловлены реализацией конкурентных взаимоотношений между бифидобактериями и патогенами, а также адгезионными свойствами бифидобактериальных штаммов [Bibiloni *et al.*, 2001; Westermann, 2015]. Адгезия комменсальных микроорганизмов к эпителиальным клеткам и мукоидному слою кишечника предотвращает колонизацию экологической ниши патогенными микроорганизмами. Ингибиторный эффект штаммов *B. breve* и *B. infantis* был продемонстрирован на примере связывания энтеротоксигенных и энтеропатогенных штаммов *Escherichia coli* и *Salmonella enterica* с культурой клеток энтероцитов Caco-2 [Bernet *et al.*, 1993]. Штамм *B. lactis* DR10 ингибирует связывание *E. coli* O157:H7 с монослоем кишечного эпителия, уменьшая также инвазивность данного патогенного штамма [Gopal *et al.*, 2001]. Более того, было показано, что штаммы *Bifidobacterium* spp. CA1 и F9 ингибируют проникновение *S. enterica* в клетки Caco-2 [Lievin *et al.*, 2000].

Антагонистический эффект бифидобактерий проявляется также в способности стимулировать иммунный ответ в присутствии патогенов. Было показано влияние живых бактериальных клеток и супернатантов их культур на экспрессию генов цитокинов IL-8, IL-10, TNF α в результате коинкубации с клеточной культурой THP-1 (клеточная культура человека) в присутствии липополисахарида *Escherichia coli* 0127: B8 в качестве индуктора. У штамма *B. longum* GT15 выявлена иммуномодулирующая активность, проявляемая в индукции экспрессии как противовоспалительного цитокина IL-10, так и провоспалительных

цитокинов IL-8 и TNF α с постепенным увеличением в течение всего времени коинкубации клеток. Наибольшая экспрессия цитокинов наблюдалась при коинкубации THP-1 клеток с живыми клетками бифидобактерий, причем максимально она проявилась в случае TNF α через 6 часов коинкубации [Аверина и др., 2015].

Таким образом, бифидобактерии как компонент комменсальной микробиоты оказывают положительный эффект на здоровье человека и поддерживают нормальное функционирование организма хозяина. Ключевыми факторами в этом процессе, по-видимому, являются адгезионные антагонистические свойства представителей рода *Bifidobacterium*, а также способность осуществлять коммуникацию с иммунной системой, стимулируя выработку как противовоспалительных, так и провоспалительных цитокинов и других факторов иммунного ответа [Ruiz *et al.*, 2017; Аверина и др., 2015]. Однако механизмы, обеспечивающие выживание комменсалов в условиях воспалительного процесса и постоянство репертуара нормальной микробиоты кишечника, остаются малоизученными. В данной диссертационной работе мы предлагаем новые гены в качестве кандидатов для реализации механизмов устойчивости к факторам иммунного ответа организма хозяина.

1.2 Характеристика рода *Bifidobacterium*.

Согласно современной таксономической классификации, род *Bifidobacterium* (бифидобактерии) относится к семейству *Bifidobacteriaceae* типа *Actinobacteria* (актинобактерии). Актинобактерии представляют собой одну из наиболее крупных таксономических единиц домена Bacteria [Ludwig, Euzéby and Whitman, 2015]. Представители рода *Bifidobacterium* являются грамположительными неспорообразующими неподвижными микроорганизмами [Sgorbati *et al.*, 1995]. Впервые бифидобактерии были описаны как палочкообразные бактерии бифидной Y-образной морфологии и классифицированы изначально как *Bacillus bifidus communis* [Tissier, 1900]. В результате реклассификации бифидобактерии были выделены в отдельный таксон [Poupard *et al.*, 1973]. К настоящему времени описано множество видов и подвидов бифидобактерий, выделенных из разных источников (рис. 2). Согласно современной таксономической классификации наиболее близкородственными бифидобактериям организмами являются представители рода *Gardnerella*, также принадлежащие семейству *Bifidobacteriaceae* [Mattarelli *et al.*, 2017]. Род *Bifidobacterium* является типовым для семейства *Bifidobacteriaceae* и порядка *Bifidobacteriales*.

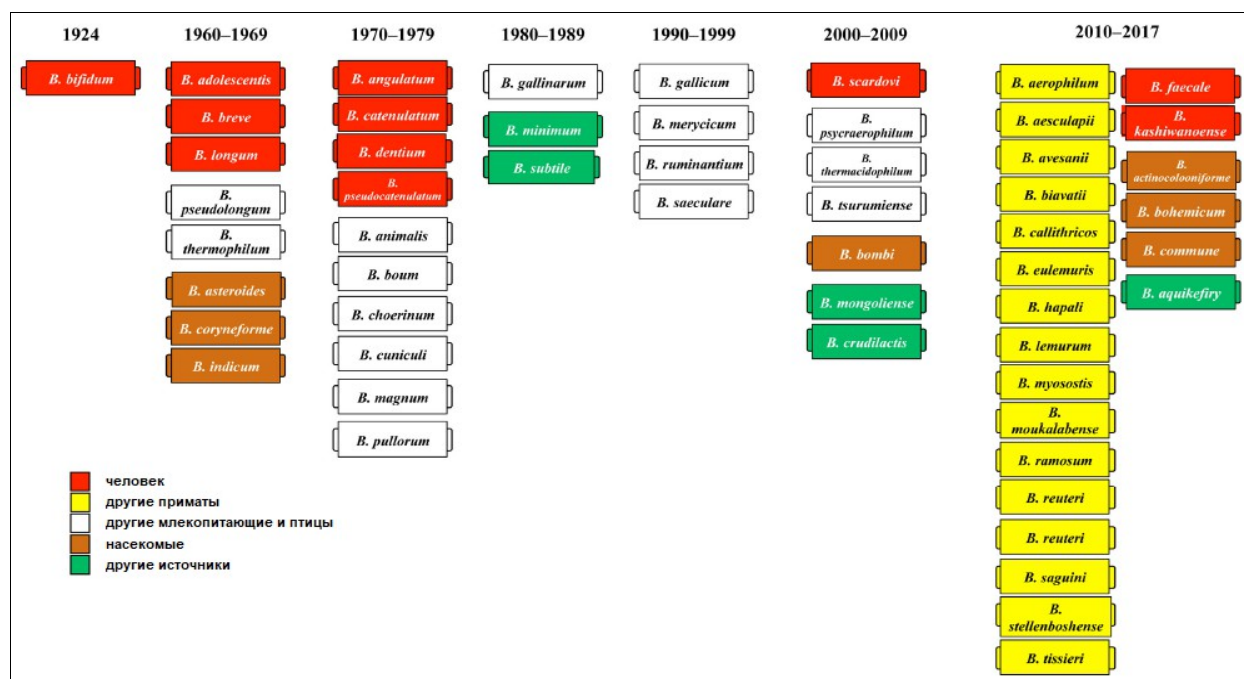


Рисунок 2. Описанные виды бифидобактерий и источники выделения [Mattarelli *et al.*, 2017, с изменениями].

Геном бифидобактерий характеризуется высоким значением GC-состава, варьирующим от 55 до 67% [Ventura *et al.*, 2004; Klijn *et al.*, 2005]. Исследования последних лет показали высокую степень консервативности нуклеотидных последовательностей бифидобактерий. Род обладает низким уровнем геномной и филогенетической изменчивости. Между геномами различных видов прослеживается высокая степень синтении. Адаптация к обитанию в толстом кишечнике, где концентрации некоторых субстратов довольно низка, включает в себя прототрофность по широкому спектру необходимых веществ. Геномы бифидобактерий содержат гены, необходимые для биосинтеза, по меньшей мере, 19 аминокислот, пиримидиновых и пуриновых нуклеотидов, фолиевой кислоты, тиамина. [Ventura *et al.*, 2007; Schell, 2002]. Более 8% генома *Bifidobacterium* кодирует ферменты необходимые для метаболизма углеводов, что на 30% больше, чем у *E. coli* или *Enterococcus faecium*. Данная особенность позволяет усваивать более широкий спектр сложных сахаров, что является преимуществом, позволяющим конкурировать с другими представителями кишечной микробиоты [Ventura *et al.*, 2007].

Род *Bifidobacterium* характеризуется значительным видовым разнообразием. На сегодняшний день (октябрь, 2019 г.) описано 70 видов в составе рода *Bifidobacterium* [<http://www.bacterio.net/bifidobacterium.html>], для большей части которых доступна информация о геномной последовательности, секвенированной, по крайней мере, в статусе «draft».

Принимая во внимание информацию об экологических нишах, а также данные сравнительного анализа последовательностей маркерных генов, бифидобактерии были разделены на 6 филогенетических групп, объединяющие близкородственные виды и подвиды [Ventura *et al.*, 2006]. Опубликованные современные филогенетические схемы (деревья), построенные на основе множественных выравниваний объединенных последовательностей коровых генов отдельных представителей различных видов бифидобактерий, воспроизводят традиционно выделяемые филогруппы, объединяющие близкородственные виды и подвиды: *B. asteroides*, *B. adolescentis*, *B. boum*, *B. longum*, *B. pseudolongum*, *B. pullorum* [Sun *et al.*, 2015], а также выделяют новую филогруппу *B. bifidum* [Lugli *et al.*, 2014; Lugli *et al.*, 2017; Lugli *et al.*, 2018] (рис. 3).

Бифидобактерии являются анаэробами и считаются одними из самых древних представителей актинобактерий [Gao, Parmanathan and Gupta, 2006], существовавшими во времена низкого содержания кислорода в атмосфере Земли. В результате изменений условий внешней среды, бифидобактерии были вынуждены занимать новые анаэробные экологические ниши. Бифидобактерии смогли колонизировать и эффективно приспособиться к обитанию в ЖКТ хозяина за счет утилизации гликанов мукоидного слоя кишечного эпителия [Koropatkin, Cameron and Martens, 2012; Grimm *et al.*, 2014]. В настоящее время бифидобактерии входят в состав облигатной комменсальной микробиоты человека и других животных, включая насекомых, птиц и млекопитающих. Некоторые виды бифидобактерий были выделены из внешних источников, в частности *B. minimum*, *B. subtile*, *B. mongoliense*, *B. aquikefiry*, *B. crudilactis*, *B. animalis* subsp. *lactis* (рис. 2). Это может быть объяснено, однако, контаминацией внешних источников бифидобактериями, которые в нормальных условиях являются частью комменсальной микробиоты в организме хозяина.

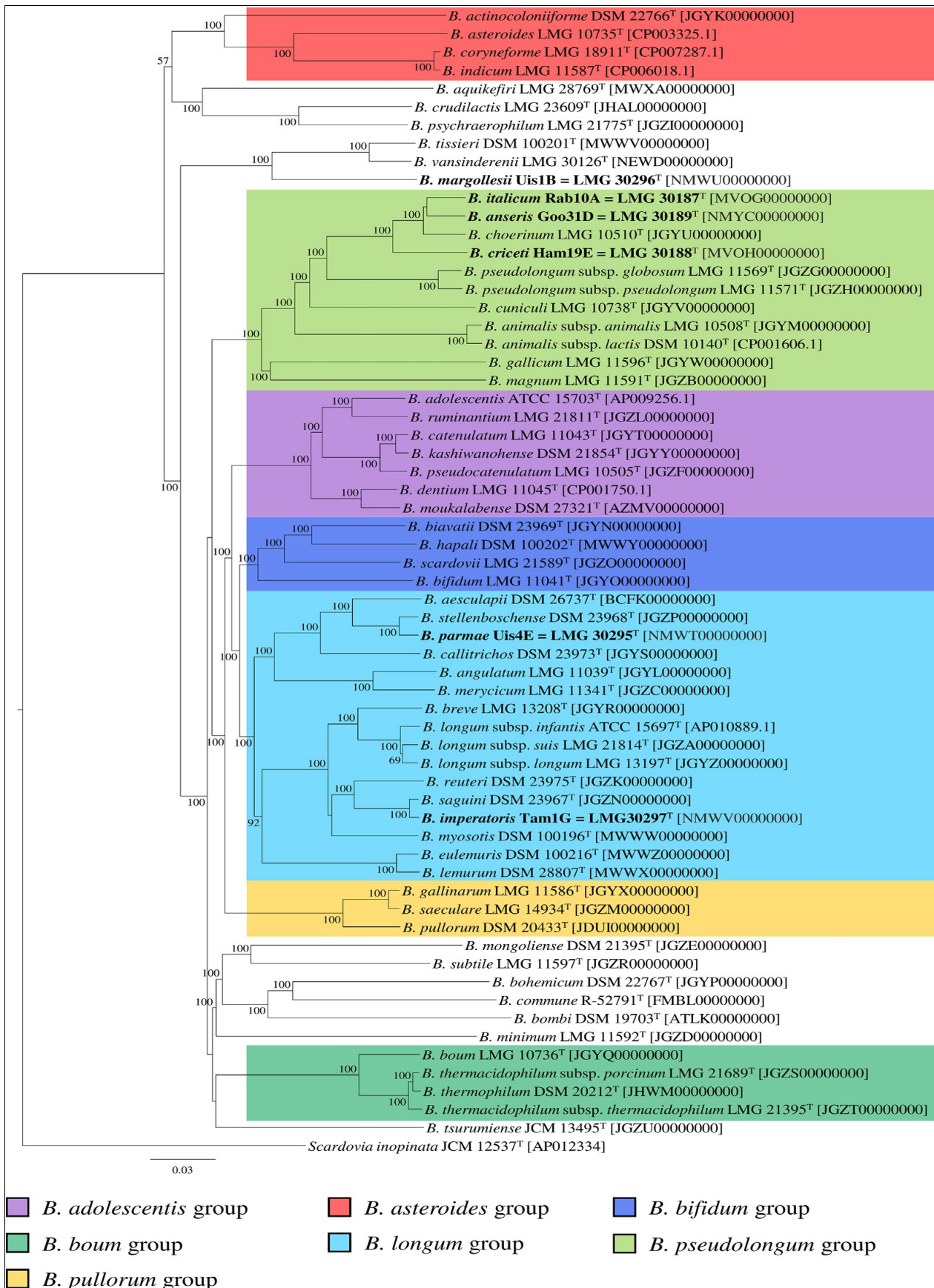


Рисунок 3. Филогенетическое дерево рода *Bifidobacterium*, построенное на основе объединенных последовательностей 233 коровых генов бифидобактерий. Указана принадлежность отдельных видов бифидобактерий к филогруппам [Lugli *et al.*, 2018].

Современные представления об экологии бифидобактерий предполагают строгую ассоциацию между видами бифидобактерий и занимаемыми экологическими нишами. Для многих видов бифидобактерий, по-видимому, характерна видоспецифичность в отношении хозяев. В частности, виды *B. magnum* и *B. cuniculi* были обнаружены только в ЖКТ кролика, *B. pullorum* и *B. gallinarum* только в ЖКТ цыплят и *B. longum* subsp. *suis* только в ЖКТ свиней. За исключением человеческих видов *B. adolescentis* и *B. dentium*, обнаруженных также в организмах орангутана и шимпанзе, соответственно [D'Aimmo *et al.*, 2014], все другие виды бифидобактерий, описанные у нечеловеческих приматов, таких как гориллы [Tsuchida *et al.*, 2014], обыкновенные игрунки и тамарины [Michelini *et al.*, 2016 a, b; Modesto *et al.*, 2014], никогда не были обнаружены в организме человека или других животных. Гипотеза о коэволюции между комменсальными микроорганизмами и хозяевами поддерживается схемой видообразования внутри рода бифидобактерий, что было подтверждено в результате анализа геномной последовательности вида *B. asteroides*, обитающего в организме пчелы. Проведенный анализ подтвердил ранее предсказанную способность *B. asteroides* к респираторному аэробному метаболизму. Толерантность *B. asteroides* к относительно высокому уровню содержания кислорода в среде связана, по-видимому, с адаптацией к обитанию в ЖКТ пчелы, характеризующегося высокой степенью аэрации. Согласно современным представлениям, филогруппа *B. asteroides*, включающая в себя виды *B. asteroides*, *B. actinocoloniiforme*, *B. indicum* и *B. coryneforme*, является эволюционно наиболее древней. Консервативность генных кластеров, обеспечивающих реализацию данной функции у разных штаммов *B. asteroides*, является косвенным подтверждением гипотезы о том, что респираторный аэробный тип обмена веществ является общей метаболической особенностью наиболее древних представителей бифидобактерий. Данная способность была утрачена в процессе видообразования и колонизации новых экологических ниш, представленных организмами современных млекопитающих [Bottacini *et al.*, 2012]. Большинство бактериальных видов, обитающих в организмах шмелей и пчел, не встречаются в других экологических нишах и принадлежат к филогенетически древним линиям, что указывает в данном случае на длительные взаимоотношения между комменсальными микроорганизмами и организмом хозяина, а также на предполагаемую длительную коэволюцию [Mattarelli *et al.*, 2017]. Данный процесс обеспечивается, вероятно, особыми механизмами адаптации комменсалов к факторам занимаемой экологической ниши в организме хозяина.

1.3 Молекулярные механизмы, обеспечивающие коммуникацию представителей микробиоты с факторами занимаемой экологической ниши в организме хозяина.

Выживание организмов зависит от умения быстро отвечать и адаптироваться к постоянно изменяющимся условиям окружающей среды. Адаптивный потенциал клеток заключается в способности воспринимать и передавать внешние и внутренние сигналы. Исследования в области бактериальной эндокринологии показали, что микроорганизмы в ходе длительной коэволюции сформировали сенсорные структуры для детекции молекул, продуцируемых клетками организма хозяина [Hughes and Sperandio, 2008; Lesouhaitier *et al.*, 2009; Freestone, 2013]. Микроорганизмы способны воспринимать сигнальные молекулы иммунной системы, пептидные гормоны и нейромедиаторы и отвечать на воздействие изменением скорости роста и ряда других характеристик (табл. 1) [Lesouhaitier *et al.*, 2009]. Механизмы, обеспечивающие данную коммуникацию, остаются малоизученными.

Таблица 1. Чувствительность грамотрицательных микроорганизмов к воздействию различных сигнальных молекул эукариотического происхождения [Lesouhaitier *et al.*, 2009, с изменениями].

группа молекул	сигнальная молекула	вид бактерий	наблюдаемый эффект
сигнальные молекулы иммунной системы	интерлейкин-1 β	<i>Escherichia coli</i>	увеличение скорости роста культуры
		<i>Acinetobacter</i> spp	увеличение скорости роста культуры
	фактор некроза опухоли- α	<i>Shigella flexneri</i>	увеличение инвазивных способностей
	интерлейкин-2	<i>Escherichia coli</i>	увеличение скорости роста культуры
	гранулоцитарно-макрофагальный колониестимулирующий фактор	<i>Escherichia coli</i>	увеличение скорости роста культуры
	интерлейкин-6	<i>Pseudomonas aeruginosa</i>	увеличение скорости роста культуры
	интерферон- γ	<i>Klebsiella pneumoniae</i>	увеличение скорости роста культуры
		<i>Escherichia coli</i>	увеличение скорости роста культуры
		<i>Pseudomonas aeruginosa</i>	увеличение чувства кворума/вирулентности
	динорфин	<i>Pseudomonas aeruginosa</i>	увеличение чувства кворума/вирулентности
нейромедиаторы	норадреналин	<i>Acinetobacter lwoffii</i>	увеличение скорости роста культуры
		<i>Bordetella bronchiseptica</i>	увеличение скорости роста культуры
		<i>Campylobacter jejuni</i>	увеличение скорости роста культуры, вирулентности
		<i>Citrobacter freundii</i>	увеличение скорости роста культуры

группа молекул	сигнальная молекула	вид бактерий	наблюдаемый эффект
		<i>Hafnia alvei</i>	увеличение скорости роста культуры
		<i>Morganella morganii</i> <i>Proteus mirabilis</i>	увеличение скорости роста культуры
		<i>Shigella sonnei</i>	увеличение скорости роста культуры
		<i>Vibrio splendidus</i>	увеличение скорости роста культуры
		<i>Vibrio parahaemolyticus</i>	увеличение скорости роста культуры, вирулентности
		<i>Xanthomonas maltophilia</i>	увеличение скорости роста культуры
	катехоламины	<i>Escherichia coli</i>	увеличение скорости роста культуры
		<i>Aeromonas hydrophila</i>	увеличение скорости роста культуры
		<i>Pseudomonas aeruginosa</i>	увеличение скорости роста культуры, вирулентности
		<i>Salmonella enterica</i>	увеличение скорости роста культуры, адгезии
		<i>Yersinia enterocolitica</i>	увеличение скорости роста культуры
	гамма-аминомасляная кислота	<i>Agrobacterium tumefaciens</i>	увеличение вирулентности
		<i>Pseudomonas aeruginosa</i>	увеличение вирулентности
пептидные гормоны	субстанция Р	<i>Escherichia coli</i>	антибактериальный эффект
		<i>Pseudomonas aeruginosa</i>	антибактериальный эффект
	инсулин	<i>Burkholderia pseudomallei</i>	уменьшение скорости роста культуры
	кальцитонин-ген-связанный пептид	<i>Escherichia coli</i>	антибактериальный эффект
	хромогранины	<i>Escherichia coli</i>	антибактериальный эффект
	соматостатин	<i>Helicobacter pylori</i>	уменьшение скорости роста культуры
	мозговой натрийуретический пептид	<i>Pseudomonas aeruginosa</i>	увеличение вирулентности
		<i>Pseudomonas fluorescens</i>	увеличение вирулентности
		<i>E. coli</i> ; <i>P. aeruginosa</i>	антибактериальный эффект
	натрийуретический пептид типа С	<i>Pseudomonas fluorescens</i>	увеличение вирулентности
		<i>Pseudomonas aeruginosa</i>	увеличение вирулентности
	гастрин	<i>Helicobacter pylori</i>	увеличение скорости роста культуры
	адренокортикотропный гормон	<i>Vibrio splendidus</i>	увеличение скорости роста культуры
	нейропептид Y	<i>Escherichia coli</i>	антибактериальный эффект

Наряду с сигнальными молекулами, продуцируемыми непосредственно клетками организма хозяина, микроорганизмы, являясь частью микробного сообщества, способны воспринимать другие сигналы занимаемой экологической ниши [Bivar Xavier, 2018]. Способность комменсальных микроорганизмов адгезировать к мукоидному слою кишечного

эпителия также является важным фактором, обеспечивающим колонизацию и способность взаимодействовать с организмом хозяина.

1.3.1 Системы сигнальной трансдукции бифидобактерий.

Среди известных бактериальных белков, обеспечивающих коммуникацию с внешней средой, которая заключается в восприятии внешних стимулов и ответе на них, одними из наиболее изученных являются системы сигнальной трансдукции. Бактериальные системы сигнальной трансдукции, как правило, представляют собой трансмембранные (ТМ) белки, проводящие сигнал от внеклеточного сенсорного модуля к внутриклеточному модулю ответа. Они подразделяются на двухкомпонентные системы, включающие гистидиновые киназы (*англ.* histidine protein kinases, НК) совместно с белками-регуляторами ответа (*англ.* response regulators, RR), и однокомпонентные системы, представленные серин-треониновыми протеинкиназами (СТПК) [Mascher *et al.*, 2006; Huang *et al.*, 2015]. Протеинкиназы совместно с сопряженными протеинфосфатазами играют ключевую роль в передаче сигнала, катализируя процессы обратимого фосфорилирования субстратов. Киназы воспринимают внешние стимулы, подвергаются автофосфорилированию и затем приобретают способность к фосфорилированию субстратов, модулируя их активность. Фосфорилирование по специфическим аминокислотным остаткам, в частности, серину (Ser), треонину (Thr), тирозину (Tyr), гистидину (His) и аспартату (Asp), может контролировать активность белка-мишени напрямую, например, индуцируя конформационные изменения в активном сайте, или косвенно, например, регулируя белок-белковое взаимодействие [Pereira, Goss and Dworkin, 2011].

Долгое время считалось, что обратимое фосфорилирование у прокариот осуществляется при участии двухкомпонентных систем сигнальной трансдукции, в частности, гистидиновых и аспартокиназ. Фосфорилирование по остаткам Ser, Thr и Tyr впервые было описано у эукариот. Эукариотические серин-треониновые и тирозиновые киназы были сгруппированы в суперсемейство эукариотических протеинкиназ на основании гомологии последовательностей между каталитическими доменами [Hanks and Hunter, 1995]. Позднее представители суперсемейства эукариотических протеинкиназ были обнаружены также у прокариот. В частности, Pkn1 *Mycoboccus xanthus* стала первой охарактеризованной СТПК эукариотического типа у бактерий [Munoz-Dorado, Inouye and Inouye, 1991]. СТПК данного суперсемейства широко распространены у грамположительных бактерий [Pereira, Goss and Dworkin, 2011]. Они обнаружены у почвенных микроорганизмов, патогенных, комменсальных и пробиотических бактерий [Zakharevich *et al.*, 2012]. Современное состояние исследований в области сигнальной трансдукции предполагает, что однокомпонентные системы, представленные СТПК, являются доминирующими системами

передачи сигнала у прокариот [Ulrich *et al.*, 2009].

Каталитические домены СТПК эукариотического типа, как правило, содержат в своей структуре 12 субдоменов, которые формируют двулопастную каталитическую кор-структуру с активным сайтом внутри [Hanks and Hunter, 1995]. Более компактная N-терминальная лопасть вовлечена преимущественно в связывание и ориентирование донорной молекулы АТФ, в то время как более крупная С-терминальная лопасть участвует в связывании субстрата и инициирует перенос фосфатной группы. Киназные домены СТПК эукариотического типа демонстрируют высокую степень консервативности пространственной структуры. Тем не менее, степень гомологии последовательностей каталитических доменов внутри суперсемейства может быть очень невысока. Консервативность третичной структуры киназных доменов обеспечивается наличием специфических мотивов и двенадцати практически инвариантных аминокислотных остатков, которые напрямую или косвенно участвуют в позиционировании субстрата фосфорилирования и донорных молекул АТФ [Hanks and Hunter, 1995] (рис. 4).

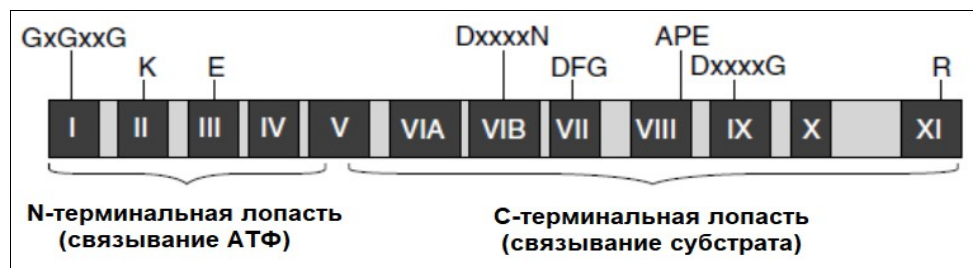


Рисунок 4. Строение каталитического домена СТПК эукариотического типа. Указана локализация двенадцати консервативных субдоменов (римские цифры) и консервативных аминокислотных мотивов с использованием однобуквенного аминокислотного кода, где «х» обозначает произвольную аминокислоту. Организация субдоменов обеспечивает формирование двулопастной пространственной структуры [Hanks, 2003].

Киназы представляют собой молекулярные переключатели, способные находиться в двух состояниях: активном и неактивном [Huse and Kuriyan, 2002]. Переход в активное состояние контролируется различными механизмами, в результате чего происходят обширные конформационные изменения в структуре обеих лопастей. Активация многих киназ происходит с помощью фосфорилирования по крайней мере одного Ser/Thr (Tyr для тирозиновых протеинкиназ) остатка активационной петли в результате автофосфорилирования или трансфосфорилирования другой киназой. Данная модификация способствует стабилизации петли в конформационном состоянии, обеспечивающим связывание субстрата и катализ [Huse and Kuriyan, 2002].

Впервые структура бактериальных СТПК эукариотического типа была описана на примере PknB *Mycobacterium tuberculosis*. Она оказалась сходной со структурой СТПК эукариот, за исключением некоторых отличий. Было показано, что PknB кристаллизуется в виде димера, благодаря взаимодействию между N-терминальными лопастями каталитических доменов двух субъединиц [Ortiz-Lombardia *et al.*, 2003; Young *et al.*, 2003]. Димеризационный интерфейс оказался консервативным среди ортологов PknB [Young *et al.*, 2003], а также был обнаружен в структуре СТПК PknE *M. tuberculosis*, несмотря на низкий уровень гомологии последовательностей данного региона между двумя киназами [Gay, Ng and Alber, 2006]. Экспериментальные данные подтверждают участие механизма лиганд-индуцированной димеризации в активации СТПК (рис. 5). В частности, замещение лиганд-связывающего домена PknD *M. tuberculosis* на домены, способные к рапамицин-индуцируемой димеризации, приводило к активации нефосфорилированной киназы. В свою очередь, замещение консервативных остатков в N-терминальной лопасти PknD приводило к падению уровня автофосфорилирования и изменению субстратной специфичности [Greenstein *et al.*, 2007]. Похожие наблюдения были также получены для PrkA *Pseudomonas aeruginosa* [Hsu, Schwarz and Mougous, 2009].

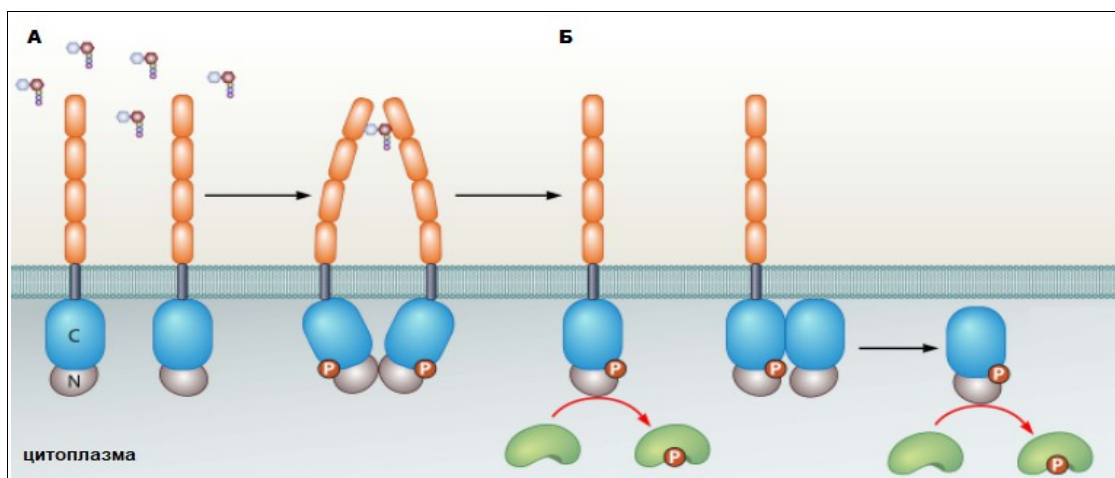


Рисунок 5. Гипотетическая модель активации СТПК эукариотического типа на примере PknB *M. tuberculosis*. В присутствии лиганда два или более мономера PknB связывают лиганд с помощью внеклеточного сигнального домена, что приводит к сближению внутриклеточных каталитических доменов и формированию димеризационного интерфейса с последующей активации СТПК в результате автофосфорилирования (А). Активированная PknB может напрямую фосфорилировать субстрат или активировать сигнальный каскад через фосфорилирование внутриклеточной протеинкиназы (Б) [Pereira, Goss and Dworkin, 2011, с изменениями].

Помимо каталитического домена СТПК эукариотического типа часто содержат дополнительные домены, обеспечивающие, в основном, связывание лигандов или реализацию белок-белкового взаимодействия. В частности, практически все грамположительные бактерии содержат СТПК, в структуре которой обнаруживается внеклеточный регион, содержащий PASTA (*англ.* penicillin-binding protein and Ser/Thr kinase-associated) повторы, связывающие, по-видимому, пептидогликаны [Yeats, Finn and Bateman, 2002]. Данная гипотеза была позднее подтверждена в исследовании, которое показало связывание пептидогликана PASTA-повторами PrkC *B. subtilis* [Shah *et al.*, 2008]. Таким образом, впервые были определены лиганды связывания внеклеточного домена СТПК эукариотического типа и подтверждена его сигнальная роль.

Анализ серин-треонинового фосфопротеома выявил мотив фосфорилирования у *M. tuberculosis*. Идентифицированный мотив имеет следующую структуру: ХааааТХ(Х/У)ф(Р/Р)I, где Т соответствует фосфотреонину, а соответствует кислым остаткам, ф соответствует крупному гидрофобному остатку. Данный мотив обнаруживается в структуре субстратов шести из одиннадцати СТПК *M. tuberculosis* (PknA, PknB, PknD, PknE, PknF и PknH) [Prisic *et al.*, 2010]. Также были проанализированы сайты фосфорилирования нескольких субстратов PknB *S. aureus*. Несмотря на то, что наиболее часто фосфорилированию подвергаются остатки серина и треонина, также было обнаружено неспецифическое фосфорилирование тирозина [Miller *et al.*, 2010].

Фосфорилированные остатки Ser, Thr и Tyr не подвергаются быстрому гидролизу, в отличие от лабильных фосfogистидина и аспартил-фосфата, поэтому для корректного функционирования механизма сигнальной трансдукции на основе обратимого фосфорилирования серина и треонина, требуется наличие фосфопротеиновой фосфатазы в качестве функционального партнера СТПК [Alber, 2009].

Бактериальные СТПК эукариотического типа участвуют в формировании сложных сигнальных сетей, которые включают фосфозависимое белок-белковое взаимодействие при участии консервативных белковых модулей или доменов. Одним из примеров является белковый домен FHA (*англ.* fork-head-associated), взаимодействующий, преимущественно, с фосфорилированными остатками треонина [Hammet *et al.*, 2003].

Представляя собой систему сигнальной трансдукции бактерий, СТПК эукариотического типа совместно со своими функциональными партнерами вовлечены в реализацию широкого спектра клеточных функций, в том числе, в процессы развития, метаболизма, клеточного деления и синтеза клеточной стенки, вирулентности и многие другие [Pereira, Goss and Dworkin, 2011].

В геномах бифидобактерий ранее были идентифицированы и охарактеризованы гены, кодирующие СТПК эукариотического типа. Аннотацию проводили на основе наличия в структуре двенадцати консервативных субдоменов. Кластерный анализ показал, что обнаруженные СТПК подразделяются на шесть индивидуальных групп ортологичных белков (Pkb1-Pkb6) (рис. 6), обладающих различной степенью внутригрупповой гомологии. На основе полученной кладограммы были выделены «консервативные» СТПК (Pkb1, Pkb3, Pkb5 и Pkb6), «уникальная» СТПК Pkb4, обнаруженная в геномах видов *B. longum*, *B. bifidum* и *B. breve*, а также «видоспецифическая» СТПК Pkb2, демонстрирующая аномально высокий уровень межвидовой дивергенции последовательностей [Nezametdinova *et al.*, 2014].

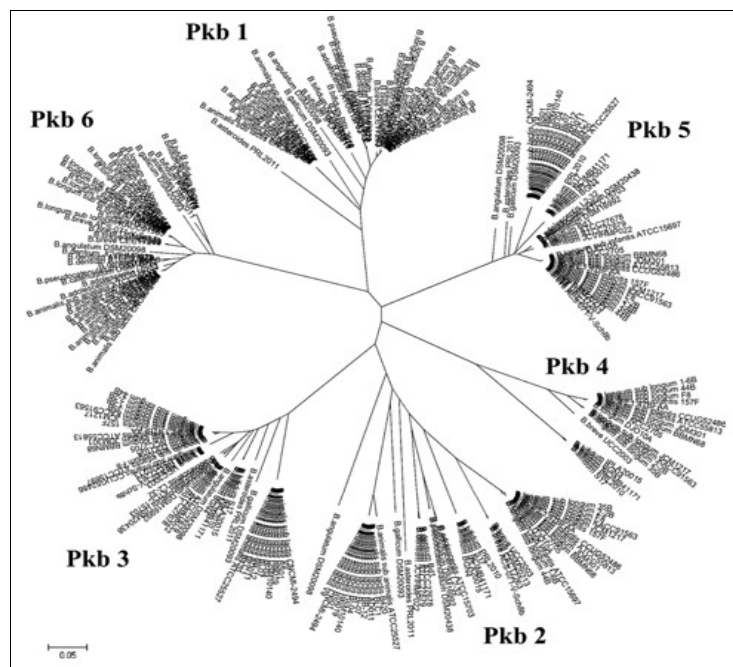


Рисунок 6. Кластерная дендрограмма, полученная на основе выравниваний последовательностей каталитического домена СТПК Pkb1-Pkb6 48 бифидобактериальных штаммов [Nezametdinova *et al.*, 2014, с изменениями].

В структуре каждой СТПК был аннотирован каталитический домен длиной 218-267 а. о. в N-концевом регионе. В С-концевом регионе СТПК Pkb1 и Pkb5 были аннотированы три и четыре PASTA-домена, соответственно. Другие регуляторные домены в структуре исследуемых протеинкиназ аннотированы не были [Nezametdinova *et al.*, 2014] (рис. 7).

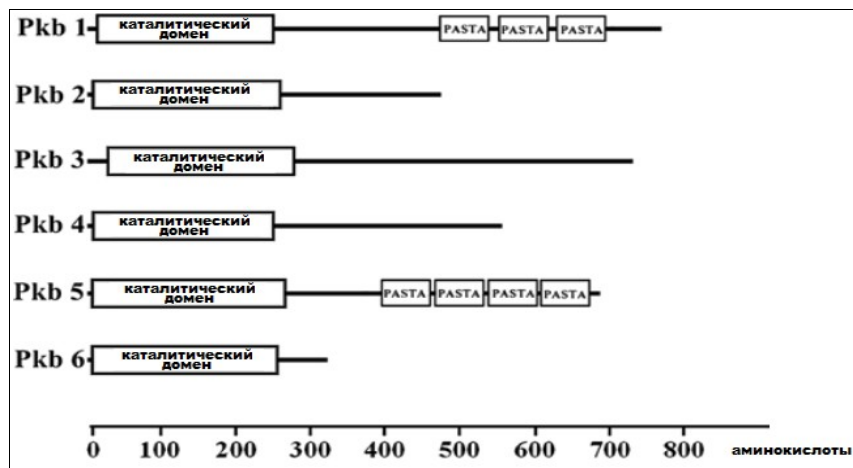


Рисунок 7. Доменная структура охарактеризованных СТПК эукариотического типа бифидобактерий [Nezametdinova *et al.*, 2014, с изменениями].

Функциональность каталитических доменов СТПК Pkb1-Pkb6 *B. longum* B379M была исследована *in vitro*. Способность к автофосфорилированию была продемонстрирована для СТПК Pkb2 и Pkb5 [Nezametdinova *et al.*, 2014], а также Pkb3 [Алексеева и др., 2015]. Также были определены субстраты фосфорилирования СТПК Pkb2 *B. longum* GT15. Для идентификации субстратов были применены два принципиально разных подхода: анализ генетического окружения гена *pkb2* и анализ спектра фосфорилированных мембранных везикул. В результате анализа генетического окружения было показано, что субстратом фосфорилирования Pkb2 является АТФаза, кодируемая геном *aaa-atp*. Масс-спектрометрический анализ позволил выявить конкретные аминокислотные сайты фосфорилирования ААА-АТР. Анализ спектра фосфорилированных мембранных везикул выявил одиннадцать предполагаемых субстратов, среди которых оказались белки, относящиеся к различным функциональным группам. Глутаминсинтаза GlnA1, а также α и β субъединицы FoF1-АТФ-синтазы были отобраны для дальнейшего анализа фосфорилирования *in vitro*, который подтвердил, что данные белки действительно являются субстратами фосфорилирования СТПК Pkb2. Анализ субстратов фосфорилирования позволил выдвинуть предположение, что функции СТПК Pkb2 могут быть связаны с адгезией и коммуникацией бифидобактерий с эпителиальными клетками кишечника [Nezametdinova *et al.*, 2018].

На сегодняшний день системы сигнальной трансдукции бифидобактерий изучены слабо, несмотря на, по-видимому, ключевую роль в реализации механизмов коммуникации с факторами внешней среды, представленной организмом хозяина. Характер дивергенции СТПК Pkb2 представляет особый интерес, так как может указывать на видоспецифичность вышеупомянутых механизмов коммуникации.

1.3.2 Механизмы адгезии бифидобактерий.

Адгезия микроорганизмов к мукоидному слою кишечного эпителия является важным фактором, обеспечивающим колонизацию и способность взаимодействовать с организмом хозяина. Мукоидный слой богат гликопротеинами и гликолипидами и обеспечивает значительное количество мишеней адгезии для комменсальной и патогенной микробиоты. Не исключено также, что бактерии способны преодолевать мукоидный слой, достигая поверхности эпителиальных клеток в норме [McGuckin *et al.*, 2011] и при патологиях, обусловленных деградацией мукоидного слоя [Swidsinski *et al.*, 2007; McGuckin *et al.*, 2011; Hoskins, 2012]. Микроорганизмы демонстрируют широкий спектр механизмов адгезии, среди которых пили и фимбрии, адгезины, распознающие разнообразные поверхностные соединения, в том числе компоненты внеклеточного матрикса (коллагены, ламинины, эластин, протеогликаны и гиалуроновую кислоту) и другие.

Адгезионные гликопротеины, такие как витронектин, фибриноген и особенно фибронектин также распознаются многими бактериальными патогенами [Pizarro-Cerdá and Cossart, 2006]. Фибронектин локализован преимущественно в областях межклеточных контактов кишечного эпителия и является одним из главных компонентов внеклеточного матрикса. Фибронектин содержит фибронектиновые домены трех типов (FNI, FNII и FNIII). Фибронектиновые домены типа III широко распространены в структуре эукариотических молекул клеточной адгезии (*англ.* cell adhesion molecules, CAMs) суперсемейства IgSF [Preedy, 2016]. FNIII домены содержат семь антипараллельных β -листов, чья топология напоминает строение домена IgC [Harpaz and Chothia, 1994]. Филогенетический анализ показал, что FNIII домены эукариот и архей имеют бактериальное происхождение [Valk, van der Kaaij and Dijkhuizen, 2017]. Бактериальные FNIII домены были впервые описаны в структуре белков, участвующих в утилизации углеводных субстратов [Watanabe *et al.*, 1990], однако на сегодняшний день FNIII домены обнаружены также во множестве других бактериальных белков [Henderson *et al.*, 2011; Letunic, Doerks and Bork, 2014]. В частности, белок FlpA *Campylobacter jejuni*, содержащий по крайней мере три FNIII домена, отвечает за адгезию к эпителиальным клеткам организма хозяина за счет связывания фибронектина. Было показано, что мутанты по гену *flpA* демонстрируют значительное снижение способности к адгезии в сравнении с диким типом. Кроме того, поликлональные антитела к FlpA блокировали адгезионные свойства в зависимости от концентрации [Konkel, Larson and Flanagan, 2010]. Наличие FNIII домена в белках, связывающих фибронектин, может быть объяснено предполагаемой способностью FNIII доменов к гомофильному взаимодействию.

Бифидобактерии обладают адгезионными свойствами в отношении мукоидного слоя [He *et al.*, 2001; Izquierdo *et al.*, 2008], а также способностью утилизировать муцины в

качестве субстрата [Tailford *et al.*, 2015]. Бифидобактерии демонстрируют широкое разнообразие механизмов, обеспечивающих адгезию у различных видов (табл. 2).

Таблица 2. Механизмы, обеспечивающие адгезионные свойства бифидобактерий [Westermann *et al.*, 2016, с изменениями].

наименование механизма	вид бифидобактерий
пили типа IVb (Tad)	<i>B. bifidum</i> , <i>B. breve</i> , <i>B. longum</i> subsp. <i>longum</i> , <i>B. adolescentis</i>
пили типа IVa	<i>B. adolescentis</i>
сортаза-зависимые пили	<i>B. adolescentis</i> , <i>B. animalis</i> subsp. <i>lactis</i> , <i>B. bifidum</i> , <i>B. breve</i> , <i>B. dentium</i> , <i>B. longum</i> subsp. <i>longum</i> , <i>B. longum</i> subsp. <i>infantis</i>
ВорА	<i>B. bifidum</i>
трансальдолаза	<i>B. bifidum</i> , <i>B. longum</i> subsp. <i>longum</i>
DnaK	<i>B. animalis</i> subsp. <i>lactis</i>
енолаза	<i>B. bifidum</i> , <i>B. animalis</i> subsp. <i>lactis</i> , <i>B. longum</i> subsp. <i>longum</i>
гидрофобность	<i>Bifidobacterium</i> sp.

Адгезионные свойства белка BBIF_0450, содержащего FNIII домен, к клеточной культуре Сасо-2 были изучены в рекомбинантных штаммах *B. bifidum* S17 и *B. longum* E18. В качестве контроля были использованы штаммы дикого типа, а также рекомбинантные штаммы, содержащие плазмиду pCW без вставки. Адгезионные свойства штаммов измеряли путем коинкубирования культуры бифидобактерий совместно с культурой эпителиальных клеток человека Сасо-2 и подсчета количества колониеобразующих единиц после отмывки бифидобактериальных клеток. Рекомбинантный штамм *B. bifidum* S17, содержащий плазмиду pCW_450 со вставкой последовательности гена *bbif_0450* *B. bifidum* S17, показал значительное увеличение адгезионных свойств по сравнению со штаммом дикого типа и рекомбинантным штаммом *B. bifidum* S17, содержащим плазмиду без вставки ($P < 0,01$) (рис. 8, А), что является экспериментальным подтверждением участия белка BBIF_0450, содержащего FNIII домен, в реализации механизмов адгезии. В свою очередь, экспрессия гена *bbif_0450* *B. bifidum* S17 в рекомбинантном штамме *B. longum* E18 не привела к статистически значимым изменениям адгезионных свойств по сравнению со штаммом дикого типа и рекомбинантным штаммом *B. longum* E18, содержащим плазмиду без вставки (рис. 8, Б). Таким образом, адгезионные свойства *bbif_0450* *B. bifidum* S17 демонстрируют видоспецифичность и проявляются только в штамме *B. bifidum* S17 [Westermann, 2015].

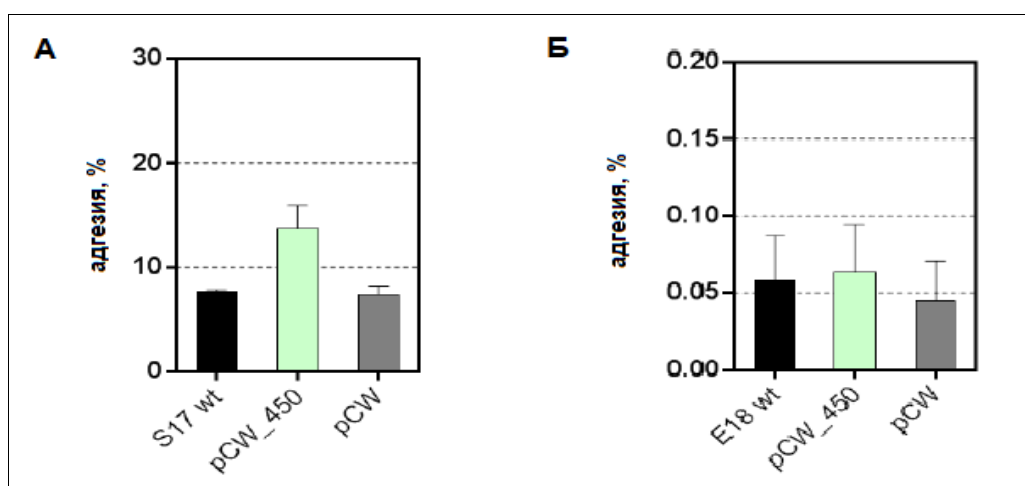


Рисунок 8. Адгезия штаммов *B. bifidum* S17 (А) и *B. longum* E18 (Б) к культуре эпителиальных клеток человека Сасо-2. Установлено процентное соотношение между изначальным количеством культивируемых клеток бифидобактерий и количеством колониеобразующих единиц после отмывки. Указаны средние значения со стандартным отклонением для трех независимых экспериментов, каждый из которых выполнен в трех повторностях; S17 wt и E18 wt — штаммы дикого типа (черный цвет), pCW_450 – рекомбинантные штаммы, содержащие последовательности гена *bbif_0450* *B. bifidum* S17 (зеленый цвет), pCW – рекомбинантные штаммы, содержащие плазмиду без вставки (серый цвет) [Westermann, 2015, с изменениями].

Вышеупомянутый ген *bbif_0450*, кодирующий FNIII-домен-содержащий белок, и его ортологи в геномах бифидобактерий являются объектом изучения данной диссертационной работы. Экспериментально показанные видоспецифические свойства представляют особый интерес в контексте изучения молекулярной эволюции данного белка.

1.4 Быстрая молекулярная эволюция генов как инструмент адаптации к занимаемой экологической нише.

В предыдущем разделе были описаны различные молекулярные механизмы, обеспечивающие коммуникацию представителей микробиоты с факторами занимаемой экологической ниши в организме хозяина. Разнообразие занимаемых бифидобактериями экологических ниш, а также тот факт, что даже наиболее близкородственные виды и подвиды были выделены из источников с различными физиологическими и биохимическими свойствами (например, подвиды *B. longum* subsp. *longum* и *B. longum* subsp. *suis* являются представителями микробиоты ЖКТ человека и свиньи, соответственно; похожим образом, подвиды *B. pseudolongum* subsp. *pseudolongum* и *B. pseudolongum* subsp. *globosum* являются представителями микробиоты ЖКТ свиньи и крупного рогатого скота, соответственно [Sun *et al.*, 2014]) указывают на то, что бифидобактерии сталкивались с необходимостью быстро

адаптироваться к новым условиям в процессе специализации. На молекулярном уровне адаптивные изменения в структуре белков, участвующих в реализации механизмов взаимодействия с факторами внешней среды, могли происходить в результате воздействия положительного отбора. Положительный естественный отбор, в отличие от отрицательного отбора, направлен на формирование новых адаптивных аллельных вариантов генов [Page and Holmes, 2009] и обеспечивает быструю эволюцию кодирующих последовательностей.

1.4.1 Быстрая эволюция генов, участвующих в реализации механизмов антагонистических взаимодействий.

Одними из хорошо изученных случаев быстрой эволюции генов, вовлеченных в процесс коммуникации с факторами внешней среды, являются эффекты, описанные гипотезой Черной Королевы. Эволюционная гипотеза Черной Королевы может быть сформулирована следующим образом: относительно эволюционной системы виду необходимы постоянное изменение и адаптация, чтобы поддерживать его существование в окружающем биологическом мире, постоянно эволюционирующем вместе с ним. Данная гипотеза предполагает, что взаимодействующие виды в ходе совместной эволюции могут становиться объектом постоянного воздействия естественного отбора с целью адаптации и контрадаптации [Van Valen, 1973; Van Valen, 1974; Stenseth and Smith, 1984]. Антагонистическая коэволюция, по-видимому, является одной из главных движущих сил эволюционных изменений внутри вида. Экспериментально было продемонстрировано, что скорость молекулярной эволюции генов паразита гораздо выше в случае коэволюции с хозяином [Paterson *et al.*, 2010]. Развитие функциональной генетики взаимодействий и сравнительного анализа позволило установить, что среди быстроэволюционирующих генов наиболее часто встречаются участники биотических взаимодействий [Brockhurst *et al.*, 2014]. Высокая степень дивергенции последовательностей, наблюдаемая в структуре некоторых генов устойчивости хозяев [Hedrick, 1994; Obbard *et al.*, 2006; Clark *et al.*, 2007] и инфективности патогенов [Blanc *et al.*, 2005; Mu *et al.*, 2006; Barrett *et al.*, 2009], подтверждает данную гипотезу.

Иммунная система является центральным медиатором антагонистических взаимодействий между патогенной микробиотой и организмом хозяина. Функционирование иммунной системы хозяина направлено на распознавание и элиминирование потенциальной патогенной инфекции. Патогены, таким образом, находятся под постоянным селективным давлением, направленным на обеспечение выживаемости в условиях иммунного ответа. Результатом естественного отбора является развитие механизмов, затрудняющих распознавание патогенов, супрессирующих сигнальную трансдукцию или увеличивающих резистентность к факторам иммунного ответа. Успешная реализация данных механизмов

патогенов, однако, приводит к развитию селективного давления, обеспечивающего эволюцию иммунитета хозяина. Таким образом, взаимоотношения между патогенными микроорганизмами и организмом хозяина могут быть охарактеризованы как непрерывный процесс реципрокной адаптации под воздействием естественного отбора [Singh, Xu and Kulathinal, 2012]. Механизмы адаптации могут включать события потери и приобретения генов, а также быстрой эволюции последовательностей. Гены, вовлеченные в функционирование иммунной системы, часто эволюционируют гораздо быстрее, чем гены других систем организма, предположительно в результате вышеупомянутой «гонки вооружений». Молекулярная эволюция генов иммунной системы подробно изучена у насекомых. Сигнальные гены Toll и Imd путей, которые могут быть активированы путем распознавания хозяином компонентов клеточной стенки грамположительных (преимущественно Toll) и грамотрицательных (преимущественно Imd) бактериальных клеток, эволюционируют на уровне последовательностей очень быстро и обнаруживаются среди наиболее дивергентных генов иммунной системы у *D. melanogaster*, *Anopheles gambiae* и *Aedes aegypti* [Waterhouse *et al.*, 2007]. Несколько индивидуальных сигнальных генов Toll и Imd путей *A. mellifera* [Viljakainen *et al.*, 2009] и *D. melanogaster* [Begun and Whitley 2000; Schlenke and Begun 2003; Jiggins and Kim 2007; Sackton *et al.* 2007] демонстрируют значимые доказательства положительного отбора. Гены *Тер*, которые кодируют опсонины, способные помечать патогенные агенты для последующего фагоцитоза и иммунологической элиминации, эволюционируют под воздействием положительного отбора в геномах представителей родов *Drosophila* [Jiggins and Kim 2007; Sackton *et al.* 2007], *Anopheles* [Little and Cobbe 2005] и *Daphnia* [Little *et al.* 2004]. Гены семейства рецепторов *scavenger* в группе видов *melanogaster* демонстрируют быструю эволюцию [Lazzaro, 2005; Sackton *et al.* 2007], также как некоторые другие гены, кодирующие рецепторы фагоцитоза [Sackton *et al.* 2007]. Опсонины и рецепторы фагоцитоза связывают множество патогенных молекул, среди которых могут быть эволюционно лабильные, что обеспечивает, по-видимому, быструю эволюцию рецепторных генов. Положительный отбор также был обнаружен в последовательности гена антимикробного пептида дефенсина муравьев [Viljakainen and Pamilo, 2008].

Как было отмечено ранее, гены микроорганизмов, участвующие в реализации механизмов антагонистических взаимодействий, также могут находиться под воздействием быстрой эволюции. Скрининг генов, находящихся под давлением положительного отбора в геномах шести штаммов *E. coli*, пять из которых являются патогенами человека, выявил гены, кодирующие поверхностные белки, такие как порины (*lamB*, *ompC*, *ompF*, *ompW*, и *fhuA*), предположительно участвующие в эволюционной «гонке

вооружений» с другими микроорганизмами, фагами или иммунной системой организма хозяина. Идентификация сайтов под воздействием положительного отбора позволила установить, что данные сайты локализованы почти исключительно во внеклеточных участках трансмембранных β -листов. Данные участки обогащены сайтами, которые являются мишенями для фагов, колицинов или иммунной системы организма человека [Petersen *et al.*, 2007].

1.4.2 Быстрая эволюция генов, участвующих в реализации механизмов взаимодействия с другими факторами занимаемой экологической ниши.

Быстрая эволюция генов патогенных микроорганизмов, отвечающих за адаптацию к занимаемой экологической нише в организме хозяина, хорошо изучена на примере возбудителя чумы *Yersinia pestis*. Молекулярный филогенетический анализ показал, что дивергенция видов *Y. pestis* и *Y. pseudotuberculosis* произошла крайне недавно, около 6500-20000 лет назад. Данный процесс видообразования сопровождался горизонтальным переносом генов и плазмид, а также нейтральной эволюцией или положительным отбором. *Y. pestis* характеризуется широким спектром организмов, используемых в качестве постоянного хозяина или вектора. В результате процесса адаптации к большому количеству экологических ниш, *Y. pestis* также подвергалась, по-видимому, быстрым эволюционным изменениям [Singh, Xu and Kulathinal, 2012]. На примере *Y. pestis* экспериментально показан блестящий пример адаптивной псевдогенизации. Авторы обнаружили, что ген *rcaA*, являющийся негативным регулятором образования биопленок, функционален в геноме *Y. pseudotuberculosis*, в то время как в геноме *Y. pestis* *rcaA* представляет собой псевдоген. В результате псевдогенизации *rcaA* *Y. pestis* приобрели способность к образованию биопленок в организме вектора по сравнению с *Y. pseudotuberculosis*. Экспериментально было показано, что замена псевдогенизированной копии *rcaA* на функциональную репрессирует формирование биопленок у *Y. pestis*, что могло бы привести к снижению трансмиссивности патогена [Sun, Hinnebusch and Darby, 2008]. Таким образом, псевдогенизация произошла в результате воздействия положительного отбора, обеспечившего повышение адаптивности и распространение *Y. pestis* [Zhang, 2008].

Геномный скрининг двух видов кораллов *Acropora millepora* и *Acropora palmata* позволил идентифицировать гены, ответственные за коммуникацию с факторами занимаемой экологической ниши и находящиеся под давлением положительного отбора. Среди генов, находящихся под давлением положительного отбора, оказались гены и генные пути, ответственные за сенсорное восприятие (в том числе, трансмембранные рецепторы и ассоциированные сигнальные пути), а также же гены, ответственные за предполагаемую коммуникацию с эндосимбионтом *Symbiodinium*. В частности, наиболее обогащенной фракцией быстро эволюционирующих белков оказались молекулы клеточной адгезии.

Данная категория белков может участвовать в реализации симбиотических взаимоотношений, так как данные белки связываются с широким спектром поверхностных молекул и могут быть медиаторами распознавания между эндосимбионтом *Symbiodinium* и организмом хозяина [Voolstra *et al.*, 2011]. Описанный случай быстрой эволюции генов клеточной адгезии представляет особый интерес в контексте изучения молекулярной эволюции гена *fn3* и его окружения.

Таким образом, быстрая эволюция генов может играть важную роль в развитии механизмов коммуникации микроорганизмов с факторами занимаемой экологической ниши в организме хозяина. Мы предполагаем реализацию данного сценария в молекулярной эволюции отдельных генов бифидобактерий, ответственных, по-видимому, за выполнение данной функции.

1.4.3 Методы и подходы, используемые для детекции событий положительного отбора в молекулярной эволюции генов.

Естественный отбор является одним из движущих факторов, обеспечивающих молекулярную эволюцию генов. Судьба возникающей *de novo* мутации в последовательности гена, то есть потеря мутации или ее фиксация, зависит от соотношения между естественным отбором и случайным дрейфом генов. Популяционная генетика традиционно классифицирует мутации как вредные, полезные и нейтральные, оказывающие минимальный эффект на организм. Скорость накопления нейтральных мутаций соответствует общему геномному уровню накопления мутаций μ . В общем случае, полезные мутации накапливаются быстрее ($\mu_+ > \mu$), вредные — медленнее ($\mu_- < \mu$). Синонимичные замены в нуклеотидных сайтах, в основном, являются нейтральными, в то время как несинонимичные часто могут выступать в роли вредных или полезных мутаций [Peacock, 2015]. Как было отмечено ранее, естественный отбор, воздействующий на нуклеотидную последовательность белок-кодирующих участков генов, может быть отрицательным (очищающий отбор, стабилизирующий отбор), направленным на элиминирование возникающих несинонимичных замен с целью сохранить консервативную структуру белка, а также положительным (адаптивный отбор, дарвиновский отбор), направленным на закрепление несинонимичных замен и формирование новых адаптивных аллельных вариантов генов [Page and Holmes, 2009]. Положительный отбор подразделяют на продолжительный и эпизодический. Продолжительный отбор означает, что в процессе эволюции на сайт оказывалось постоянное давление положительного отбора (например, в результате «гонки вооружений» патогенов с иммунной системой хозяина). Эпизодический отбор предполагает однократное возникновение несинонимичной мутации с ее последующим закреплением.

Одним из способов тестирования положительного отбора на уровне нуклеотидных последовательностей является оценка значения ω , рассчитываемого как соотношение количества несинонимичных замен, нормированных на количество несинонимичных сайтов (d_N , K_a) к количеству синонимичных замен, нормированных на количество синонимичных сайтов (d_S , K_s) в белок-кодирующих последовательностях гомологичных генов за одинаковый период времени [Nei and Kumar, 2000]. Значение $\omega > 1$ интерпретируется как доказательство воздействия положительного отбора, в то время как значения $\omega < 1$ и $\omega = 1$ указывают на отрицательный отбор и нейтральную эволюцию, соответственно. Как было отмечено ранее, синонимичные замены в основном являются эволюционно нейтральными, так как не приводят к изменению первичной структуры кодируемого белка, в отличие от несинонимичных замен. Большинство несинонимичных замен в кодирующих регионах генов оказывают негативный эффект на структуру и функции белков. В результате большинство генов в геномах организмов находятся под давлением отрицательного отбора и демонстрируют значение $\omega < 1$. Положительный отбор является относительно редким событием, поэтому детекция положительного отбора и выяснение причин оказываемого давления представляет особый интерес для исследователей [Peacock, 2015]. В общем случае детектирование положительного отбора является сложной задачей, так как положительный отбор часто может затрагивать отдельные сайты последовательности или короткий эволюционный период, в то время как значение ω для полноразмерной последовательности остается низким. Положительный отбор может также происходить в отдельных ветвях филогенетического дерева. В последние годы были разработаны подходы, которые позволяют тестировать положительный отбор в индивидуальных ветвях или наборе ветвей филогенетического дерева (т. н. «branch» модели) [Yu and Irwin, 1996; Messier and Stewart, 1997; Zhang, Kumar, and Nei, 1997; Yang 1998; Zhang, Rosenberg, and Nei 1998] с помощью метода максимального правдоподобия. Также разработаны модели, позволяющие тестировать положительный отбор в отдельных кодонах белок-кодирующей последовательности гена (т. н. «site» модели) [Nielsen and Yang, 1998; Suzuki and Gojobori, 1999; Yang *et al.*, 2000]. Тесты на основе «site» моделей позволяют детектировать события продолжительного положительного отбора. Наиболее реалистичными являются модели, разработанные для тестирования эпизодического положительного отбора в индивидуальных кодонах последовательностей, соответствующих отдельным ветвям филогенетического дерева (т. н. «branch-site» модели) [Yang and Nielsen, 2002]. Модификация данной модели имеет низкий уровень ложноположительных результатов и обладает большей статистической мощностью по сравнению с тестом на основе «branch» моделей [Zhang, Nielsen and Yang, 2005], поэтому была использована в текущей работе для детекции положительного отбора в интересующих нас генах.

ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ.

2.1 Анализ *in silico*.

2.1.1 Биоинформатический анализ генетического окружения гена *pkb2* в геномах представителей различных видов рода *Bifidobacterium*.

Принципиальная схема анализа филогенетических профилей групп ортологичных белков бифидобактерий представлена на рис. 9.

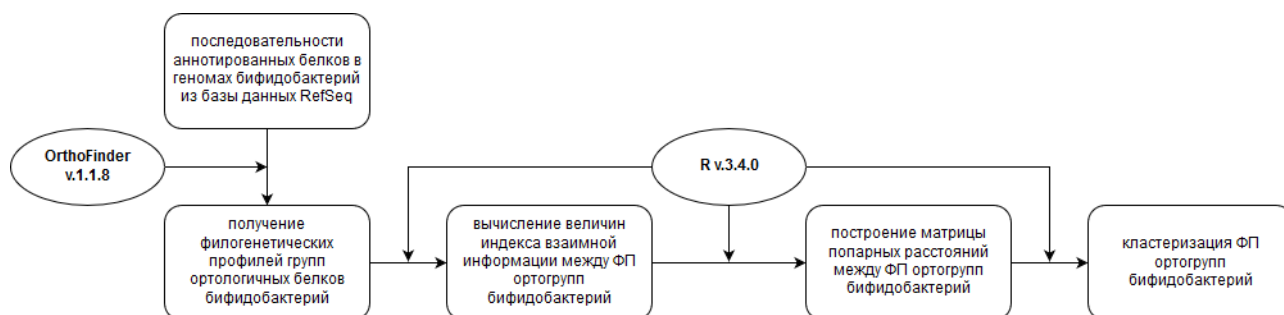


Рисунок 9. Принципиальная схема анализа филогенетических профилей групп ортологичных белков бифидобактерий.

Определение групп ортологичных белков (ортогрупп) бифидобактерий проводили с помощью программы OrthoFinder (v.1.1.8) [Emms and Kelly, 2015]. Программа OrthoFinder использует метод Reciprocal Best Hit (RBH) на основе алгоритма локального выравнивания аминокислотных последовательностей BLAST [Altschul *et al.*, 1990]. Эффективность анализа филогенетических профилей быстро возрастает с увеличением количества геномов в анализе, однако по достижении порогового значения ~100 геномов скорость возрастания эффективности при добавлении в анализ новых геномов начинает снижаться и практически исчезает на отметке в ~400 геномов [Skunca and Dessimoz, 2015]. Для определения ортогрупп белок-кодирующих последовательностей генов бифидобактерий было использовано 109 геномных аннотаций из базы данных референсных последовательностей RefSeq NCBI (<https://www.ncbi.nlm.nih.gov/refseq/>), включающих в себя белковые аннотации последовательностей геномов всех секвенированных на декабрь 2017 г. видов (75 геномов в статусе "complete" и 34 генома в статусе "draft", относящихся к 51 виду рода *Bifidobacterium*) (табл. 1 Приложений). Точность анализа филогенетических профилей зависит также от качества и полноты сборки геномов. В анализ были включены преимущественно полногеномные последовательности в статусе «complete», кроме случаев отсутствия в базе данных RefSeq полной сборки геномных последовательности отдельных видов бифидобактерий.

Визуализацию ФП генов проводили с использованием программы Count (v9.1106 RC1) [Csuros, 2010]. Матрицу попарных расстояний между ФП получали с использованием значений индекса взаимной информации (англ. mutual information, MI): $D_{MI}=1-MI$. Кластерный анализ проводили с помощью метода невзвешенного попарного среднего (UPGMA) в программной среде R (v.3.4.0).

Из полученного списка групп ортологичных белков бифидобактерий были отобраны ортогруппы, соответствующие белковым продуктам, кодируемым генами кластера PFNA: OG0000354 (*aaa-atp*), OG0001048 (*pkb2*), OG0001049 (*duf58*), OG0001050 (*igm*), OG0001061 (*fn3*).

Филогенетическое дерево бифидобактерий на основе наличия/ отсутствия определенных ортогрупп в отдельных геномах также было получено в программе OrthoFinder (v.1.1.8) с использованием алгоритма DendroBLAST [Kelly and Maini, 2013].

Кластеризацию последовательностей белка, кодируемого геном *aaa-atp* кластера PFNA бифидобактерий с другими представителями суперсемейства AAA+ АТФаз в целях определения принадлежности исследуемого белка к определенному семейству и подсемейству проводили в программе CLANS. Программа использует алгоритм на основе графов Фрухтермана-Рейнгольда для отображения попарного сходства последовательностей в двумерном/ трёхмерном пространстве [Frickey and Lupas, 2004]. Для оценки принадлежности анализируемой АТФазы, кодируемой геном *aaa-atp* кластера PFNA к определенному семейству AAA+ АТФаз, была проведена кластеризация 14 аминокислотных последовательностей AAA-АТФ представителей различных видов бифидобактерий (*B. longum*, *B. bifidum*, *B. ruminantium*, *B. stercoris*, *B. dentium*, *B. saguini*, *B. adolescentis*, *B. pseudocatenulatum*, *B. boum*, *B. aesculapii*, *B. thermacidophilum*, *B. animalis*, *B. merycicum*, *B. sp.* MSTE12) с 4428 аминокислотными последовательностями опубликованного списка белков суперсемейства AAA+ АТФаз [Iyer *et al.*, 2004]. Определение принадлежности исследуемой АТФазы к определенному подсемейству проводили в результате кластеризации с 596 аминокислотными последовательностями опубликованного списка белков семейства MoxR AAA+ АТФаз [Snider and Houry, 2006]. Аминокислотные последовательности были получены с использованием авторских скриптов, реализованных в программной среде R (v.3.4.0).

Поиск ортологов гена, кодирующего BT1854 *B. thetaiotaomicron*, во всех доступных для анализа (октябрь, 2019 г.) секвенированных полногеномных последовательностях бифидобактерий проводили с помощью программы BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) [Altschul *et al.*, 1990].

2.1.2 Анализ молекулярной эволюции генов кластера PFNA в геномах представителей различных видов рода *Bifidobacterium*.

Анализ молекулярной эволюции генов кластера PFNA проводили с помощью метода максимального правдоподобия. Принципиальная схема алгоритма анализа молекулярной эволюции генов кластера PFNA бифидобактерий представлена на рис. 10.

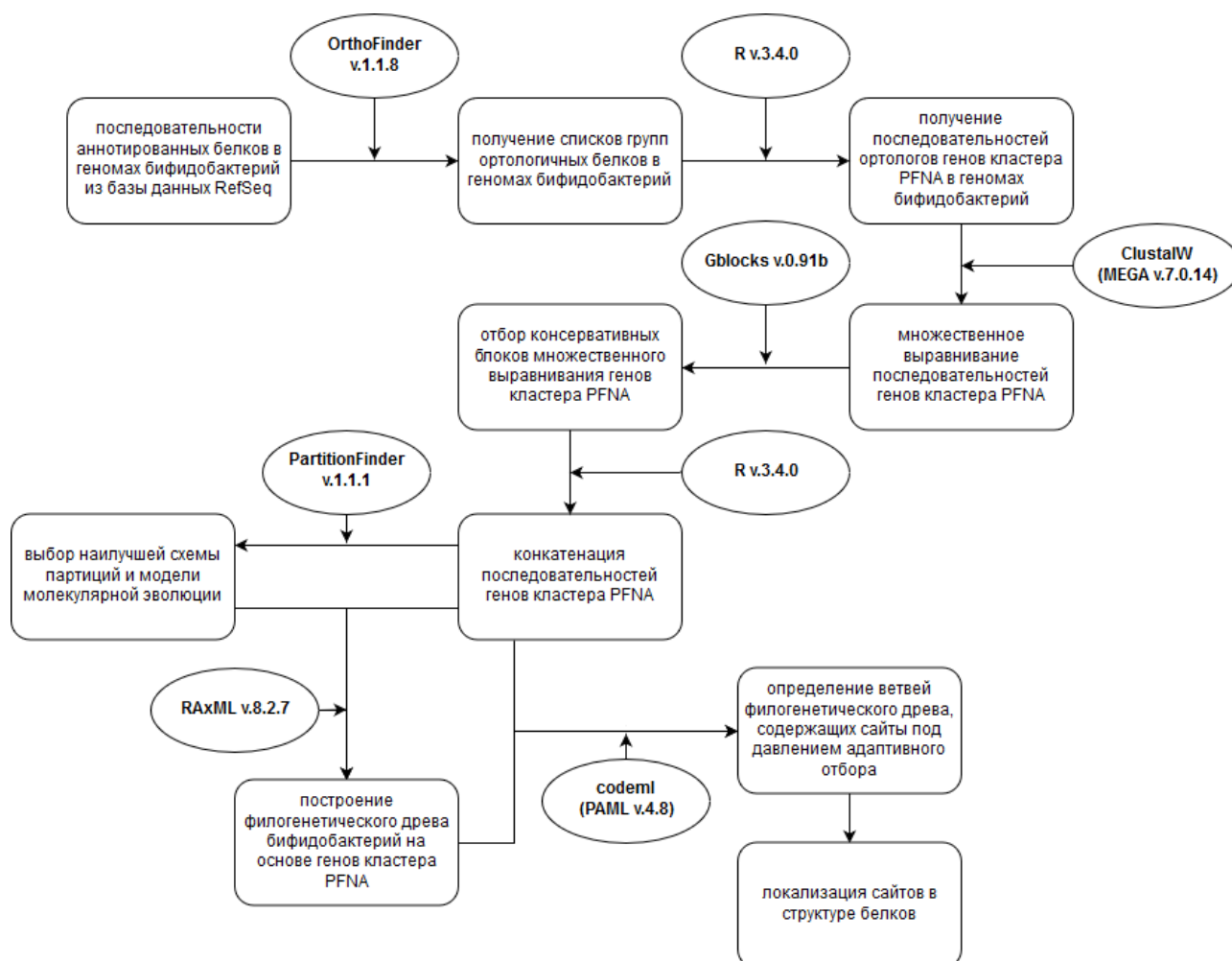


Рисунок 10. Принципиальная схема анализа молекулярной эволюции последовательностей генов кластера PFNA бифидобактерий.

2.1.2.1 Подготовка данных для последующего анализа молекулярной эволюции генов кластера PFNA представителей различных видов рода *Bifidobacterium*.

Подготовка данных для последующего построения филогенетического дерева и анализа молекулярной эволюции включала в себя получение нуклеотидных последовательностей групп ортологичных генов кластера PFNA представителей различных видов рода *Bifidobacterium*, множественное выравнивание последовательностей и отбор консервативных блоков выравнивания, определение наилучшей схемы партиций и модели молекулярной эволюции.

Полученные ранее с помощью программы OrthoFinder списки ортологичных белков представителей различных видов бифидобактерий, относящихся к ортогруппам OG0000354, OG0001048, OG0001049, OG0001050 и OG0001061, использовали для формирования базы данных нуклеотидных последовательностей генов кластера PFNA бифидобактерий. Были отобраны последовательности кодирующих регионов генов *pkb2*, *fn3*, *aaa-atp*, *duf58* и *tgm* в геномах представителей 43 различных видов и подвидов бифидобактерий. Список и характеристика используемых в анализе геномов приведена в табл. 2 Приложений. Нуклеотидные последовательности кодирующих регионов вышеупомянутых генов были получены с использованием авторских скриптов, реализованных в программной среде R (v.3.4.0).

Множественное выравнивание 43 нуклеотидных последовательностей генов *pkb2*, *fn3*, *aaa-atp*, *duf58* и *tgm* осуществляли независимо без смещения кодонов в программе ClustalW [Thompson, Higgins and Gibson, 1994] в составе программного обеспечения MEGA (v.7.0.14) [Kumar, Stecher and Tamura, 2016] с использованием настроек по умолчанию.

Для последующего анализа молекулярной эволюции производили отбор консервативных блоков множественного выравнивания в программе Gblocks (v.0.91b) (рис. 11) с использованием настроек по умолчанию (табл. 3). Отбор блоков позволяет удалить регионы с неудовлетворительным качеством выравнивания, что ведет к улучшению последующего филогенетического анализа и получению более надежных результатов [Talavera and Castresana, 2007; Castresana, 2000].

Таблица 3. Значения параметров отбора консервативных блоков множественного выравнивания генов кластера PFNA бифидобактерий в программе Gblocks (v.0.91b).

наименование параметра	значение параметра
минимальное количество последовательностей для консервативной позиции блока	22
минимальное количество последовательностей для фланкирующей позиции блока	36
максимальное количество смежных неконсервативных позиций в блоке	8
минимальная длина блока	10
допустимое количество пропусков в блоке	0

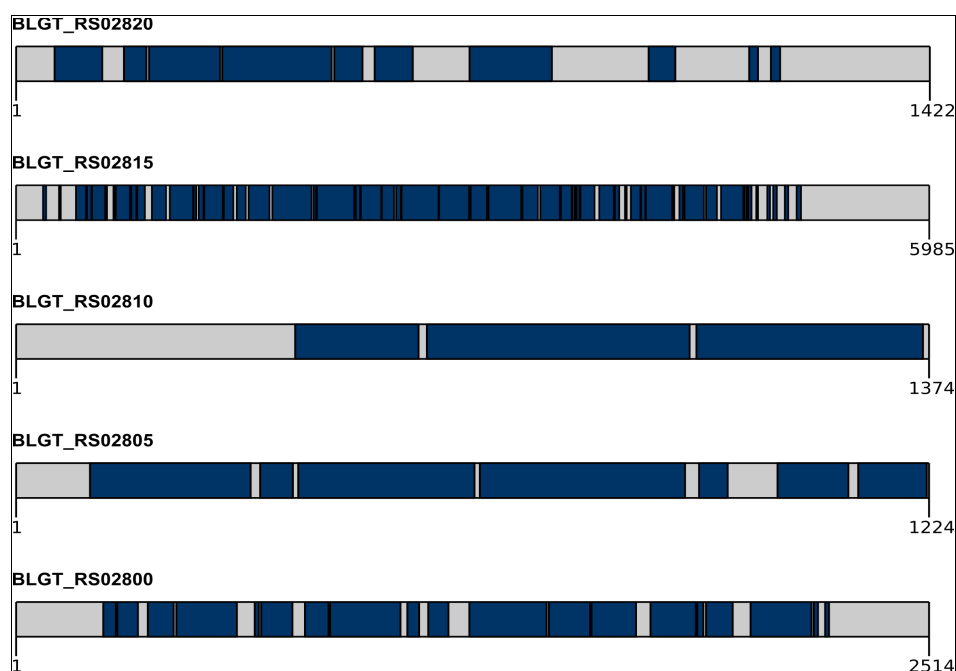


Рисунок 11. Локализация отобранных консервативных блоков выравнивания (синий цвет) в последовательностях генов *pkb2* (BLGT_RS02820), *fn3* (BLGT_RS02815), *aaa-atp* (BLGT_RS02810), *duf58* (BLGT_RS02805) и *tgm* (BLGT_RS02800) кластера PFNA *B. longum* subsp. *longum* GT15.

Определение наилучшей схемы партиций и модели молекулярной эволюции проводили в программе PartitionFinder (v.1.1.1) [Lanfear *et al.*, 2016] с использованием следующих информационных критериев для выбора из нескольких статистических моделей: информационного критерия Акаике (AIC), Байесовского информационного критерия (BIC) и скорректированного критерия Акаике (AICc). В качестве наилучшей модели молекулярной эволюции была определена модель GTR+Г+I.

2.1.2.2 Филогенетический анализ, основанный на сравнении последовательностей генов кластера PFNA представителей различных видов рода *Bifidobacterium*.

Построение филогенетических деревьев проводили с использованием метода максимального правдоподобия в программе RaxML (v.8.2.7) [Stamatakis, 2014] на основе множественных выравниваний белок-кодирующих последовательностей генов *pkb2*, *fn3*, *aaa-atp*, *duf58* и *tgm* 43 представителей различных видов рода *Bifidobacterium* (табл. 2 Приложений). Для построения общего филогенетического дерева бифидобактерий на основе последовательностей пяти генов кластера PFNA, множественные выравнивания последовательностей индивидуальных генов предварительно объединяли. Устойчивость кластеризации оценивали по значениям поддержек ветвей после 10000 итераций бутстреп-анализа. Так как ортологи генов кластера PFNA бифидобактерий не были обнаружены в

геномах других организмов, назначение аутгруппы при построении филогенетических деревьев не представлялось возможным.

Анализ коэволюции последовательностей генов кластера PFNA проводили в программе MirrorTree (<http://csbg.cnb.csic.es/mtserver/>) [Ochoa and Pazos, 2010]. Метод MirrorTree основан на оценке коэффициента корреляции Пирсона между матрицами попарных эволюционных расстояний двух сравниваемых филогенетических деревьев. Матрицы эволюционных расстояний получали путем суммирования длин ветвей, разделяющих листья филогенетического дерева.

2.1.2.3 Подбор параметров анализа молекулярной эволюции и поиск событий продолжительного/ эпизодического положительного отбора.

Подбор параметров анализа молекулярной эволюции и поиск событий продолжительного и эпизодического положительного отбора проводили с помощью программы codeml в составе пакета PAML (v.4.8) [Yang, 2007]. Для поиска событий продолжительного положительного отбора использовали «site» модели, позволяющие тестировать положительный отбор в индивидуальных кодонах белок-кодирующих последовательностей генов (табл. 4). Для поиска событий эпизодического положительного отбора использовали «branch-site» модели, позволяющие тестировать положительный отбор в индивидуальных кодонах последовательностей, соответствующих отдельным ветвям филогенетического дерева (табл. 4).

Таблица 4. Описание и значение параметров моделей, используемых для поиска продолжительного/ эпизодического положительного отбора в белок-кодирующих последовательностях генов кластера PFNA бифидобактерий.

наименование модели	описание модели	значение параметров модели
«site» модели		
модель M0	модель предполагает одинаковое значение ω для всех сайтов последовательности во всех ветвях филограммы	model = 0 NSsites = 0 fix_omega = 0
модель M1a	модель предполагает два класса сайтов последовательности со значениями $\omega < 1$ и $\omega = 1$ во всех ветвях филограммы (модель не предполагает наличие положительного отбора)	model = 0 NSsites = 1 fix_omega = 0
модель M3	модель предполагает несколько классов сайтов последовательности с различными значениями ω во всех ветвях филограммы	model = 0 NSsites = 3 fix_omega = 0
модель M8	модель предполагает 11 классов сайтов последовательности со значениями $\omega < 1$ и $\omega \geq 1$ во всех ветвях филограммы (модель предполагает наличие положительного отбора)	model = 0 NSsites = 8 fix_omega = 0
модель M8a	модель предполагает 11 классов сайтов последовательности со значениями $\omega < 1$ и $\omega = 1$ во всех ветвях филограммы (модель не предполагает наличие положительного отбора)	model = 0 NSsites = 8 fix_omega = 1 omega = 1

«branch-site» модели		
модель А	модель предполагает две категории ветвей: тестируемые ветви интереса (т. н. «foreground») и остальные ветви филограммы (т. н. «background»); модель предполагает три класса сайтов для тестируемых ветвей со значением $\omega < 1$, $\omega = 1$ и $\omega > 1$ и два класса сайтов для остальных ветвей филограммы со значением $\omega < 1$ и $\omega = 1$ (модель предполагает наличие положительного отбора только в тестируемых ветвях)	model = 2 NSsites = 2 fix_omega = 0
модель A1	модель предполагает две категории ветвей: тестируемые ветви интереса (т. н. «foreground») и остальные ветви филограммы (т. н. «background»); модель предполагает два класса сайтов для тестируемых и остальных ветвей филограммы со значением $\omega < 1$ и $\omega = 1$ (модель не предполагает наличие положительного отбора)	model = 2 NSsites = 2 fix_omega = 1 omega = 1

Программа *codeml* в составе пакета PAML (v.4.8) использует метод максимального правдоподобия, что позволяет описать вышеупомянутые модели как математическую сумму стохастических процессов молекулярной эволюции. Данный метод подбирает значения различных параметров наблюдаемых данных в соответствии с выбранной моделью, а также вычисляет значение логарифма функции правдоподобия $\ln L$, которая представляет собой вероятность наблюдения данных при сгенерированных параметрах эволюционной модели.

В целях повышения качества анализа подбирали оптимальное значение параметра CodonFreq с использованием следующих информационных критериев для выбора из нескольких статистических моделей: информационного критерия Акаике (AIC), Байесовского информационного критерия (BIC) и скорректированного критерия Акаике (AICc). Были протестированы все возможные значения параметра CodonFreq в модели A1. В качестве тестового набора ветвей была назначена клада, соответствующая виду *B. longum*, включающему в себя подвиды *B. longum* subsp. *infantis* и *B. longum* subsp. *longum*. Значение информационных критериев рассчитывали по следующим формулам: $AIC = 2 \times (k - \ln L)$, $AICc = AIC + (2 \times k \times (k + 1) / (n - k - 1))$, $BIC = k \times \ln n - 2 \times \ln L$, где k – количество параметров модели (np), n – размер выборки (количество нуклеотидов в выравнивании), L – значение функции правдоподобия. Наилучшая модель традиционно имеет наименьшее значение информационного критерия. По результатам всех информационных критериев было подобрано значение параметра модели CodonFreq=7 (табл. 5).

Таблица 5. Оптимизация значения параметра CodonFreq для последующего анализа молекулярной эволюции.

значение параметра CodonFreq модели	количество параметров модели (np)	значение $\ln L$ для модели	значение информационного критерия AIC	значение информационного критерия AICc	значение информационного критерия BIC
CodonFreq=0:	4	-223014,19	446036,38	446036,38	446064,48
CodonFreq=1:	4	-221082,69	442173,39	442173,39	442201,49

CodonFreq=2:	4	-219215,20	438438,40	438438,41	438466,50
CodonFreq=3:	4	-216426,68	432861,35	432861,36	432889,45
CodonFreq=4:	4	-219235,41	438478,81	438478,82	438506,92
CodonFreq=5:	4	-217637,24	435282,48	435282,48	435310,58
CodonFreq=6:	26	-215027,37	430106,74	430106,91	430289,41
CodonFreq=7:	67	-212064,17	424262,34	424263,45	424733,06

Длины ветвей филогенетического дерева вычисляются в программе *RaxML* как предполагаемое количество замен на нуклеотидный сайт, в то время как длины ветвей при анализе молекулярной эволюции в программе *codeml* интерпретируются как предполагаемое количество замен на кодон. Таким образом, длины ветвей, полученные в программе *RaxML* не могут быть напрямую использованы для поиска событий положительного отбора. Для вычисления корректных длин ветвей филогенетического дерева использовали модель M0 ($\text{fix_blength} = 0$), и затем фиксировали длины ветвей во всех последующих тестах (значение параметра $\text{fix_blength} = 2$).

На основе значений логарифма функции правдоподобия $\ln L$, полученных с помощью реализации моделей, описанных в табл. 4, были осуществлены тесты отношения правдоподобия (*англ.* likelihood ratio test, LRT) (табл. 6) и оценены P-значения (*англ.* P-value, P), представляющие собой вероятности ошибки первого рода при соответствующем количестве степеней свободы распределения. LRT является статистическим критерием, используемым для сравнения соответствия двух моделей (нулевой модели и альтернативной модели) наблюдаемым данным и вычисляется по формуле $\text{LRT} = 2 \times (\ln L - \ln L_0)$, где L_0 и L это значения функции правдоподобия нулевой и альтернативной модели, соответственно. LRT-статистика имеет распределение χ^2 в случае «site» тестов и смешанное распределение $\frac{1}{2}\chi^2 + \frac{1}{2}\chi_0^2$ в случае «branch-site» тестов [Zhang, Nielsen and Yang, 2005]. Авторы рекомендуют использовать χ^2 в обоих случаях, однако в случае «branch-site» это увеличивает консервативность тестов, поэтому данное распределение применяли только в строгих условиях тестирования.

Таблица 6. Описание проводимых тестов отношения правдоподобия, используемых для поиска продолжительного/ эпизодического положительного отбора в белок-кодирующих последовательностях генов кластера PFNA бифидобактерий.

нулевая гипотеза	альтернативная гипотеза	количество степеней свободы распределения	тестируемая гипотеза
«site» тесты			
M0	M3	4	тестируется предположение о наличии в последовательности нескольких категорий сайтов с различными значениями ω

M8a	M8	1	тестируется предположение о наличии в последовательности сайтов, находящихся под давлением положительного отбора ($\omega > 1$) во всех ветвях филограммы
«branch-site» тесты			
M1a	A	2	тестируется предположение о наличии в последовательности сайтов, находящихся под давлением положительного отбора ($\omega > 1$)/ под ослабленным давлением отрицательного отбора в тестируемой ветви/ наборе ветвей по сравнению с остальными ветвями филограммы
A1	A	1	тестируется предположение о наличии в последовательности сайтов, находящихся под давлением положительного отбора ($\omega > 1$) в тестируемой ветви/ наборе ветвей по сравнению с остальными ветвями филограммы

Тесты «branch-site» на наличие положительного отбора/ ослабление отрицательного отбора проводили независимо для 35 индивидуальных ветвей и клад филогенетического дерева бифидобактерий, полученного на основе объединенных последовательностей белок-кодирующих участков генов кластера PFNA. Низкие значения длин тестируемых ветвей могут приводить к значительному ослаблению статистической мощности теста и неспособности детектировать положительный отбор. Короткие ветви были протестированы в составе следующих клад филогенетического дерева, включающих в себя группу близкородственных видов или подвидов бифидобактерий: *B. angulatum*, *B. merycicum*; *B. animalis* subsp. *animalis*, *B. animalis* subsp. *lactis*; *B. catenulatum*, *B. kashiwanohense*, *B. pseudocatenulatum*; *B. coryneforme*, *B. indicum*; *B. longum* subsp. *infantis*, *B. longum* subsp. *longum*; *B. thermacidophilum* subsp. *porcinum*, *B. thermacidophilum* subsp. *thermacidophilum*, *B. thermophilum*. Для уменьшения вероятности ошибок первого рода, т. е. вероятности ложноотрицательного результата в результате множественного тестирования, были использованы поправки на множественное сравнение [Anisimova and Yang, 2007], в частности, поправка Бонферрони и процедура Бенджамини-Хохберга (критерий FDR, англ. False Discovery Rate). Вычисление поправки Бонферрони проводили следующим образом: $p_i < \alpha/m$, где α — уровень значимости, m — количество испытаний. Вычисление поправки FDR проводили в программной среде R (v.3.4.0).

Тесты на наличие эпизодического положительного отбора проводили в строгих и нестрогих условиях (табл. 7). Нестрогие условия тестирования применяли с целью уменьшить значение ошибки второго рода, т. е. вероятность ложноотрицательного результата при поиске редких событий, так как в строгих условиях тесты достаточно консервативны [Zhang, Nielsen and Yang, 2005].

Таблица 7. Значения параметров «branch-site» тестов в строгих и нестрогих условиях.

параметры тестирования	строгие условия	нестрогие условия
распределение	χ^2	$\frac{1}{2}\chi^2 + \frac{1}{2}\chi_0^2$
уровень значимости α	0,01	0,05
поправка на множественное сравнение	поправка Банферрони	метод Бенджамини-Хохберга

Для верификации полученных результатов проводили дополнительные тесты на наличие эпизодического положительного отбора с использованием программы MEME [Murrell *et al.*, 2012] в составе пакета HyPhy (v.2.2.4).

Для ветвей филогенетического дерева, в которых были обнаружены классы сайтов под давлением положительного отбора, проводили процедуру поиска конкретных аминокислотных сайтов с помощью эмпирического Байесовского подхода (*англ.* Bayes empirical Bayes, BEB), в программе codeml пакета PAML (v.4.8). Данный подход оценивает апостериорные вероятности (*англ.* posterior probability, PP) принадлежности аминокислотного сайта к классу сайтов со значением $\omega > 1$ [Yang, Wong and Nielsen, 2005]. В качестве кандидатов для воздействия положительного отбора были назначены сайты со значениями PP $> 0,7$ [Pavlova *et al.*, 2017]. Оригинальные координаты аминокислотных сайтов под давлением положительного отбора были получены с использованием авторских скриптов, реализованных в программной среде R (v.3.4.0). Визуализацию первичной структуры белков с указанием локализации сайтов под давлением положительного отбора (PP = 0,7-0,95 и PP $> 0,95$) проводили в программе IBS (v.1.0). Для локализации доменов в первичной структуре белков использовали белковую аннотацию базы данных RefSeq NCBI, а также базу данных SMART (<http://smart.embl-heidelberg.de/>) [Letunic, Doerks and Bork, 2014; Letunic and Bork, 2017].

2.2 Экспериментальный анализ.

2.2.1 Изучение транскрипционной организации кластера PFNA в геноме *B. longum* subsp. *longum* GT15.

Экспериментальный анализ включал в себя культивирование штамма *B. longum* subsp. *longum* GT15, выделение тотальной РНК *B. longum* subsp. *longum* GT15, проведение реакции обратной транскрипции, постановка ПЦР на матрице полученной кДНК с использованием набора специфических праймеров с целью подтверждения предполагаемой оперонной организации, а также определение точки старта транскрипции методом быстрой амплификации 5'-концов кДНК (*англ.* rapid amplification of cDNA 5'-ends, 5'-RACE).

Бифидобактерии выращивали на твердой агаризованной среде Lactobacillus MRS Agar и в жидкой среде Bifidobacterium Broth (HiMedia, Индия) с добавлением цистеина (0,5 г/л) (HiMedia, Индия). Культивирование штаммов бифидобактерий проводили в анаэробных условиях (HiAnaerobic System – Mark III, AnaeroHiGas Pack 3,5L; HiMedia, Индия) при температуре 37°C до середины логарифмической (экспоненциальной) фазы роста культуры клеток. Данный протокол культивирования бифидобактериальных штаммов является общим для описанных ниже экспериментов.

Выделение геномной ДНК из осажденных на центрифуге Beckman J2-21(США) клеток проводили с использованием набора GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich, США) в соответствии с протоколом производителя. Концентрацию выделенной ДНК измеряли с использованием набора Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, США) на приборе Qubit 2.0 Fluorometer (Thermo Fisher Scientific, США). Данный протокол выделения и очистки геномной ДНК является общим для описанных ниже экспериментов.

Для выделения тотальной РНК штаммы бифидобактерий выращивали до середины логарифмической (экспоненциальной) фазы роста культуры клеток. Далее клетки обрабатывали протектором РНК RNA Protect Bacteria Reagent (QIAGEN, США). Выделение и очистку тотальной РНК из осажденных на центрифуге Beckman J2-21(США) клеток проводили с использованием набора RNeasy Mini Kit (QIAGEN, США) в соответствии с протоколом производителя. Клетки разрушали с помощью гомогенизатора Speed Mill plus (Analytik Jena AG, Германия) с использованием пробирок Lysis Tubes B (Analytik Jena AG, Германия). Удаление остаточной геномной ДНК проводили с использованием наборов RNase-Free DNase Set (QIAGEN, США) и TURBO DNA-free Kit (Invitrogen, США) в соответствии с протоколом производителей. Концентрацию выделенной РНК измеряли с использованием набора Qubit RNA HS Assay Kit (Thermo Fisher Scientific, США) на приборе Qubit 2.0 Fluorometer (Thermo Fisher Scientific, США). Данный протокол выделения и очистки тотальной РНК является общим для описанных ниже экспериментов.

Реакцию обратной транскрипции проводили с использованием обратной транскриптазы SuperScript III Reverse Transcriptase и случайных гексануклеотидных праймеров [d(N)6] (random hexamers) (Invitrogen, США) в соответствии с протоколом производителя. Перед постановкой реакции обратной транскрипции производили нормирование количества тотальной РНК в различных пробах (начальное количество матрицы в реакционной смеси составило 380,6 нг во всех пробах). Концентрацию полученной кДНК измеряли с использованием набора Qubit ssDNA HS Assay Kit (Thermo Fisher Scientific, США) на приборе Qubit 2.0 Fluorometer (Thermo Fisher Scientific, США). Данный протокол получения кДНК в ходе реакции обратной транскрипции является общим

для описанных ниже экспериментов.

Амплификацию геномной ДНК и кДНК, полученной в ходе реакции обратной транскрипции, проводили с использованием набора реактивов «Амплификация» (Dialat Ltd., Москва, Россия) согласно протоколу производителя на приборе Терцик ТП4-ПЦР01 (ДНК-Технология, Россия). Используемое количество $MgCl_2$ (50мМ) в реакционной смеси составило 1,0 мкл. Также для уменьшения температуры отжига праймеров к реакционной смеси добавляли 1,25 мкл аддитива диметилсульфоксида (ДМСО). Условия амплификации включали следующие этапы: начальную денатурацию 3 мин. при 95°C, затем 30 циклов – денатурация 30 сек. при 95°C, отжиг 30 сек. при оптимальной температуре отжига, элонгация 1 мин./1 кн. п. при 72°C, далее финальная элонгация 5 мин. при 72°C. Данный протокол постановки реакции ПЦР является общим для описанных ниже экспериментов.

Праймеры для изучения транскрипционной организации генов кластера PFNA *B. longum* subsp. *longum* GT15 подбирали комплементарно внутренним участкам соседних генов так, чтобы амплификат соответствовал спейсерным последовательностям (табл. 8). Олигонуклеотиды для постановки данного эксперимента, а также экспериментов, описанных ниже, подбирали в программе Primer-BLAST (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>). Олигонуклеотиды синтезировали в ЗАО «Синтол» (Москва, Россия), а также в ЗАО Евроген (Москва, Россия).

Таблица 8. Праймеры для изучения транскрипционной организации генов кластера PFNA *B. longum* subsp. *longum* GT15.

ампликон	праймер	размер продукта	последовательность (5' → 3')
sp1	sp1-fn3-F	920 п. о.	GTCGTTTTCGGTCAGGATGGT
	sp1-pkb2-R		TATTCGCCTTCGTGGTCGTT
sp2	sp2-mrp-F	903 п. о.	TGCCGATAACCACCTTGGAG
	sp2-fn3-R		AATCAACAAACCGCAGGAGC
sp3	sp3-duf58-F	925 п. о.	TCCACATCGGTTTGCTTGGA
	sp3-mrp-R		CGCAGATGGACCGATTCTT
sp4	sp4-transglutaminase-F	776 п. о.	GCAAACGCCAGAAACACGAT
	sp4-duf58-R		CTCCTCCATCGGCGTCAAAT
sp5	sp5-prpc-F	847 п. о.	GGA CTGAGATGGTAGGTGCG
	sp5-transglutaminase-R		TTCAAGGCCATACTGCTGGC
sp6	hypo7-F	932 п. о.	ATTTGCATCTGGTCGGTGGT
	prpc-R		CCGCACCTACCATCTCAGTC
sp7	fha-F	763 п. о.	TCTACGATACGCACCGCCTT
	hypo7-R		AATACGGGTTCAC TGCCGAG

В качестве отрицательного контроля была использована тотальная РНК *B. longum* subsp. *longum* GT15. В качестве контроля длин ампликонов была использована геномная ДНК *B. longum* subsp. *longum* GT15.

Электрофорез проводили в 0,8-2% агарозном геле в трис-ацетатном буфере (ТАЕ). В качестве маркеров длины фрагментов ДНК использовали маркеры молекулярной веса Gene Ruler 100bp Plus DNA Ladder (Fermentas, Литва) и Lambda DNA/EcoRI plus HindIII (Thermo Fisher Scientific, США). Данный протокол визуализации ПЦР-продуктов является общим для описанных ниже экспериментов.

Для выявления 5'-концевых последовательностей использовали метод Step-Out RACE [Matz *et al.*, 1999]. Данный метод заключается в применении пошагового «внешнего праймирования» универсальными смесями адаптерных праймеров (step-out ПЦР) одновременно с пошаговым «заглублением» геноспецифических праймеров (nested ПЦР). На матрице тотальной РНК *B. longum* subsp. *longum* GT15 была проведена пошаговая амплификация с использованием геноспецифических праймеров (табл. 9) и набора Mint RACE cDNA amplification set (Евроген, Россия).

Таблица 9. Праймеры для постановки 5'-RACE с целью поиска точек старта транскрипции оперона PFNA *B. longum* subsp. *longum* GT15.

ген	раунд ПЦР	праймер	последовательность (5' → 3')
<i>pkb2</i>	I	pkb2-real-C	CAGCACCGCTTCCACATCG
	II	sp0-pkb2-F	TGGGCTTGATGTCGTGATGG
<i>prpC</i>	I	prpC-F	GGGTGATGATGTTGCGTGGA
	II	sp5-prpC-F	GGACTGAGATGGTAGGTGCG
	III	sp5-prpC-F1	TATCCTGACGGGTGCGTGA

Далее проводили очистку ПЦР-продуктов с помощью набора QIAquick PCR Purification Kit (QIAGEN, США) или QIAquick Gel Extraction Kit (QIAGEN, США) согласно протоколу производителя. Данный протокол очистки ПЦР-продуктов является общим для последующих экспериментов. Определение нуклеотидной последовательности проводили в ЗАО Евроген (Москва, Россия). Визуализацию нуклеотидной последовательности проводили в программе FinchTV (v.1.4.0).

Поиск точки терминации транскрипции проводили *in silico* с помощью выравнивания последовательностей в программе ClustalW [Thompson, Higgins and Gibson, 1994] в составе программного обеспечения MEGA (v.7.0.14) [Kumar, Stecher and Tamura, 2016] с использованием настроек по умолчанию.

2.2.2 Сравнительный анализ уровня экспрессии гена *fn3* у представителей различных видов рода *Bifidobacterium*.

Количественный анализ уровня экспрессии гена *fn3* у представителей различных видов бифидобактерий включал в себя построение кривых роста культивируемых штаммов бифидобактерий, выделение и очистку тотальной РНК, получение кДНК в ходе реакции обратной транскрипции, амплификацию фрагмента гена *fn3* представителей различных видов бифидобактерий в ходе реакции ПЦР в реальном времени с использованием видоспецифических праймеров и количественный анализ результатов.

Анализ транскрипционной активности проводили для следующих штаммов бифидобактерий из лабораторной коллекции: *B. animalis* subsp. *animalis* ATCC 25527, *B. choerinum* DSM 20434, *B. pseudolongum* subsp. *globosum* DSM 20092, *B. angulatum* GT102, *B. breve* DSM 20213, *B. longum* subsp. *longum* GT15, *B. thermophilum* DSM 20210, *B. ruminantium* AC1786, *B. dentium* Bifido 9. Культивирование штаммов, выделение и очистку тотальной РНК, получение кДНК в ходе реакции обратной транскрипции проводили согласно описанным ранее протоколам.

ПЦР в реальном времени (ПЦР-РВ) проводили на матрице кДНК с использованием набора 2.5x Реакционная смесь для проведения ПЦР-РВ в присутствии Eva Green и ROX (СИНТОЛ, Россия) согласно протоколу производителя на приборе CFX96 Touch Real-Time PCR Detection System (Bio-Rad, США). Перед постановкой ПЦР в реальном времени проводили нормирование количества кДНК в различных пробах (начальное количество матрицы в реакционной смеси составило 41,65 нг во всех пробах). Условия амплификации включали следующие этапы: начальную денатурацию 6 мин. при 95°C, затем 40 циклов – денатурация 15 сек. при 95°C, отжиг и элонгация 40 сек. при 60-62°C (температура отжига подбиралась отдельно для каждой пары праймеров) и регистрация сигнала флуоресценции, затем построение кривой плавления продуктов амплификации. По форме графика первой производной кривой плавления оценивали специфичность отжига праймеров и амплификации целевого продукта. Праймеры для постановки ПЦР подбирали на внутренние участки гена *fn3* таким образом, чтобы размер амплификата не превышал 200 п. о. (табл. 10).

Таблица 10. Праймеры для постановки ПЦР в реальном времени с целью изучения транскрипционной активности гена *fn3* у представителей различных видов рода *Bifidobacterium*.

штамм	праймер	размер продукта	последовательность (5' → 3')
<i>B. animalis</i> ATCC 25527	fn3-ani-F	179 п. о.	GGGCATAGTGTCTGGTGGATG

	fn3-ani-R		CCGTGTAGGTGAGGGGGTAA
<i>B. choerinum</i> DSM 20434	fn3-cho-F	126 п. о.	CATCCCCAATGACGGCGAAA
	fn3-cho-R		GCATCAACGACGGGTTTGTCTC
<i>B. breve</i> DSM 20213	fn3-bre-F	184 п. о.	TACGCATTCCCATCACGCTG
	fn3-bre-R		GTCCAATCCTCCACGGCATAG
<i>B. thermophilum</i> DSM 20210	fn3-thermo-F	172 п. о.	CTCCGCCACCATCACCATT
	fn3-thermo-R		GCTACCCAATCCGAGCAGTT
<i>B. angulatum</i> GT102	fn3-ang-F	179 п. о.	TACGCATTCCCATCGCACTT
	fn3-ang-R		CCAATCCTCCACCGCATAGG
<i>B. pseudolongum</i> DSM 20092	fn3-pseudo-F	145 п. о.	ACGAGCAACAAGCACACCTT
	fn3-pseudo-R		GCACCGAGCAATCCGTCTG
<i>B. ruminantium</i> AC1786	fn3-rumi-F	155 п. о.	ACACCGTCAAAGACAAGGCA
	fn3-rumi-R		CTGCGACACATCCACATCCA
<i>B. dentium</i> Bifido 9	fn3-dent-F	148 п. о.	CTGAAGGTCAAGGCTGGGTC
	fn3-dent-R		TTATCATTCGGTTGGCGGGG
<i>B. longum</i> GT15	Fibronectin-F	176 п. о.	TACGACTATCGTGTGCCGTC
	Fibronectin-R		ACGGTGATGGTGGTCAATG

Расчет значений начальной концентрации матрицы N_0 в пробе проводили по следующей формуле: $N_0 = N_q / E^{C_q}$, где N_q — значение сигнала флуоресценции в пробе на цикле C_q , E — эффективность ПЦР. Эффективность ПЦР вычисляли с помощью построения стандартной кривой эффективности. Для этого проводили серию (N) десятикратных разбавлений матрицы и затем строили прямую в координатах $\log[N]/C_q$. Значение эффективности амплификации вычисляли по следующей формуле: $E = -1/\log(k)$, где k — тангенс угла наклона прямой ($y=kx+B$).

Нормирование уровня экспрессии генов по концентрации суммарной РНК [Bustin *et al.*, 2002] было выбрано нами в качестве наиболее оптимального варианта. Так как уровень экспрессии генов домашнего хозяйства (*англ.* housekeeping genes, HG), традиционно используемых в качестве эндогенного контроля, в клетках представителей различных видов рода *Bifidobacterium* различается, нормирование данных с использованием референсных генов не представлялось возможным.

2.2.3 Изучение видоспецифичности киназно-субстратного взаимодействия в реакции перекрестного фосфорилирования *in vitro*.

Принципиальная схема экспериментального изучения видоспецифичности киназно-субстратного взаимодействия в реакции перекрестного фосфорилирования *in vitro* представлена на рис. 12.

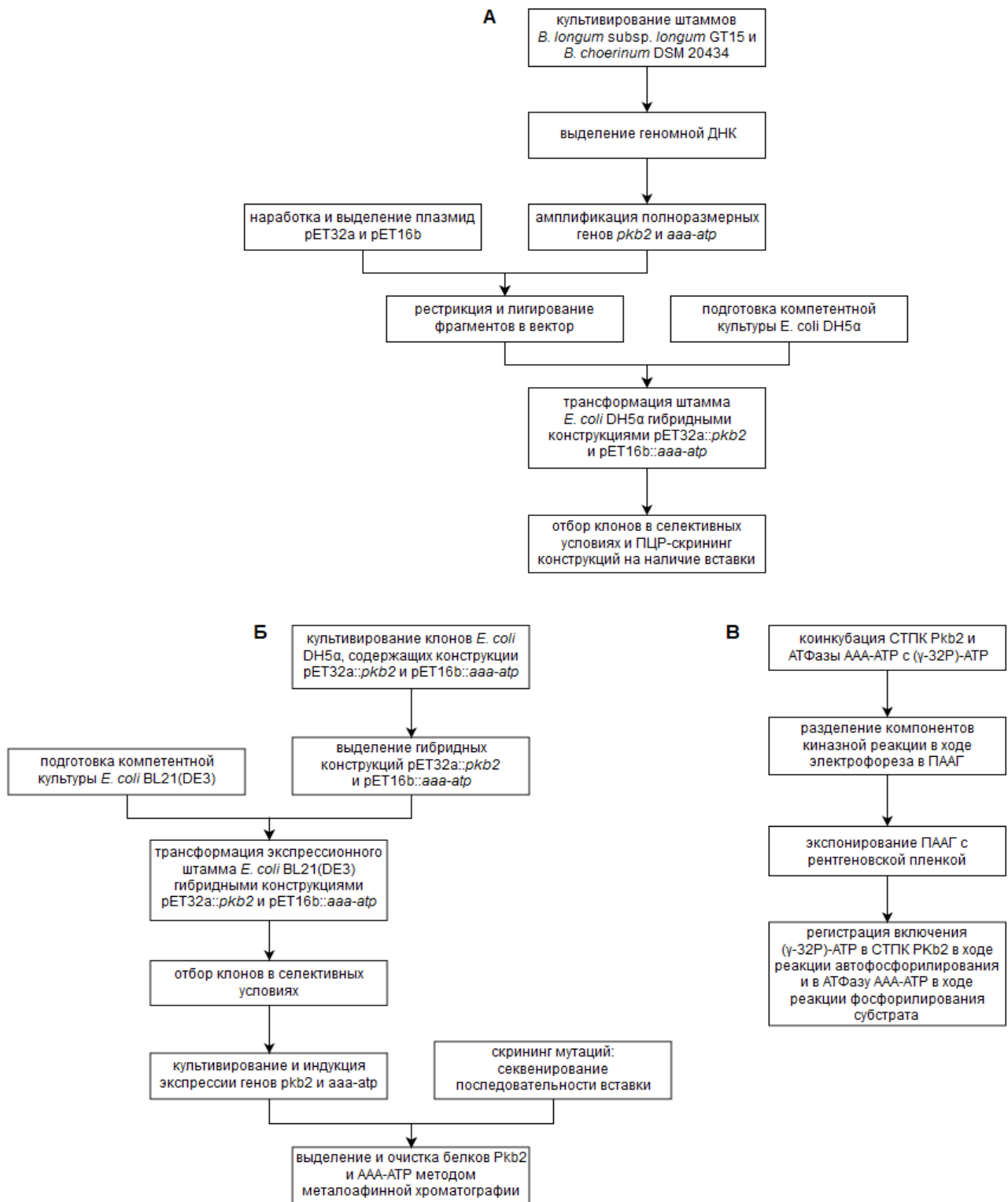


Рисунок 12. Принципиальная схема экспериментального изучения видоспецифичности киназно-субстратного взаимодействия *in vitro*. А - клонирование полноразмерных генов *pkb2* и *aaa-atp* *B. longum* subsp. *longum* GT15 и *B. choerinum* DSM 20434 в штамм *E. coli* DH5α, Б - выделение и очистка полноразмерных белков Pkb2 и AAA-ATP из экспрессионного штамма *E. coli* BL21(DE3) методом металоафинной хроматографии в нативных условиях, В - коинкубация СТПК Pkb2 и АТФазы AAA-ATP с (γ-32P)-АТФ и выявление фосфорилированных продуктов киназной реакции методом автордиографии.

Культивирование штаммов *B. longum* subsp. *longum* GT15 и *B. choerinum* DSM 20434, выделение геномной ДНК, амплификацию генов *pkb2* и *aaa-atp* и очистку ПЦР-продуктов проводили согласно описанным ранее протоколам. Праймеры подбирали комплементарно фланкирующим участкам полноразмерной последовательности генов. Для возможности последующего клонирования к 5'-концам праймеров присоединяли нуклеотидные последовательности, содержащие сайты рестрикции (табл. 11).

Таблица 11. Праймеры для молекулярного клонирования генов *pkb2* и *aaa-atp* *B. choerinum* DSM 20434.

ген	праймер	сайт рестрикции	последовательность (5' → 3')
<i>pkb2</i>	cho-pkb2-F	BamHI	TATCGGATCCCCACAGAGGACCGCATCCACG
	cho-pkb2-R	HindIII	CCGCAAGCTTCTTGGCCGCGCAGGCG
<i>prpC</i>	cho-ATPase-F	NdeI	GCTGCATATGATGACGAAGCACAACGAGACACC
	cho-ATPase-R	BamHI	AGCCGGATCCGTGCACGTCGAGCTGCGG

Для клонирования и экспрессии последовательностей использовали плазмидные векторы pET32a и pET16b (Novagen, Германия), содержащие ген тиоредоксина и ген устойчивости к ампициллину в качестве селективного маркера. Для рестрикции амплифицированных фрагментов и плазмид использовали ферменты BamHI, HindIII, NdeI (Fermentas, Литва); для лигирования ДНК использовали T4 ДНК-лигазу (Promega, США).

Клонирование генов *pkb2* и *aaa-atp* проводили в штамме *E. coli* DH5α: *F*-, Φ 80 *ΔlacZΔM15*, *Δ(lacZYA-argF)*, *U169* (Promega). Клетки *E. coli* культивировали на твердой агаризованной среде LB Agar (Sigma-Aldrich, США) и в жидкой среде LB Broth (Sigma-Aldrich, США) в течение ночи при температуре 37°C при постоянном встряхивании со скоростью 200 оборотов в минуту с добавлением ампициллина (100 мкг/мл). Плазмидную ДНК из клеток *E. coli* DH5α выделяли с использованием набора GeneJet Plasmid Miniprep Kit (Fermentas, Литва) в соответствии с инструкцией производителя. Экспрессию генов проводили в штамме *E. coli* BL21 (DE3): *F*-, *dcm*, *ompT*, *hsdS* (*rB-mB*-), *galλ* (DE3) (Novagen). Индукцию экспрессии генов осуществляли в присутствии изопропил-β-D-1-тиогактопиранозида (ИПТГ) (250 нг/мл) в течение ночи при температуре 28°C при постоянном встряхивании со скоростью 250 оборотов в минуту с добавлением ампициллина (150 мкг/мл). Приготовление компетентной культуры *E. coli* DH5α и BL21(DE3) и трансформацию проводили стандартными методами [Maniatis, Fritsch and Sambrook, 1982].

Для выделения и очистки белков Pkb2 и AAA-ATP *B. choerinum* DSM 20434 методом металоафинной хроматографии использовали набор Ni-NTA Fast StartKit (Qiagen, США). Электрофорез белковых фракций проводили в ПААГ с использованием маркера

молекулярного веса Prestained Protein Molecular Weight Marker (Fermentas, Литва).

Киназную реакцию проводили в течение 30 мин при 37°C в 20 мкл буфера, содержащего 25 мМ Tris-HCl (pH 7,4), 5 мМ MgCl₂, 5 мМ MnCl₂, 1 мМ ДТТ, 1 мМ ЭДТА, 1 мМ ФМСФ. Реакцию автофосфорилирования проводили с использованием эквимольных количеств СТПК Pkb2 *B. choerinum* DSM 20434 и *B. longum* subsp. *longum* GT15. Реакцию фосфорилирования и перекрестного фосфорилирования проводили путем коинкубации СТПК Pkb2 и субстрата ААА-АТР, принадлежащих одному/ разным видам бифидобактерий. Реакцию начинали добавлением АТФ до конечной концентрации 10 мкМ (содержащего 10 мкКи/мкл [γ -³²P]-АТФ). Реакцию останавливали добавлением Ds-Na-ПААГ-буфера для образцов, реакционную смесь прогревали в течение 5 мин при 95°C. Образцы анализировали электрофоретически в 12,5%-ном Ds-Na-ПААГ. Для авторадииографии гели фиксировали в трихлоруксусной кислоте (ТХУ) 50%, окрашивали Кумасси R250 0,1% и экспонировали с рентгеновской пленкой.

2.2.4 Оценка влияния IL-6 как фактора иммунного ответа на скорость роста культуры и экспрессию генов штамма *B. longum* subsp. *longum* GT15.

Принципиальная схема транскриптомного анализа в эксперименте по оценке влияния IL-6 на экспрессию генов *B. longum* subsp. *longum* GT15 представлена на рис. 13.

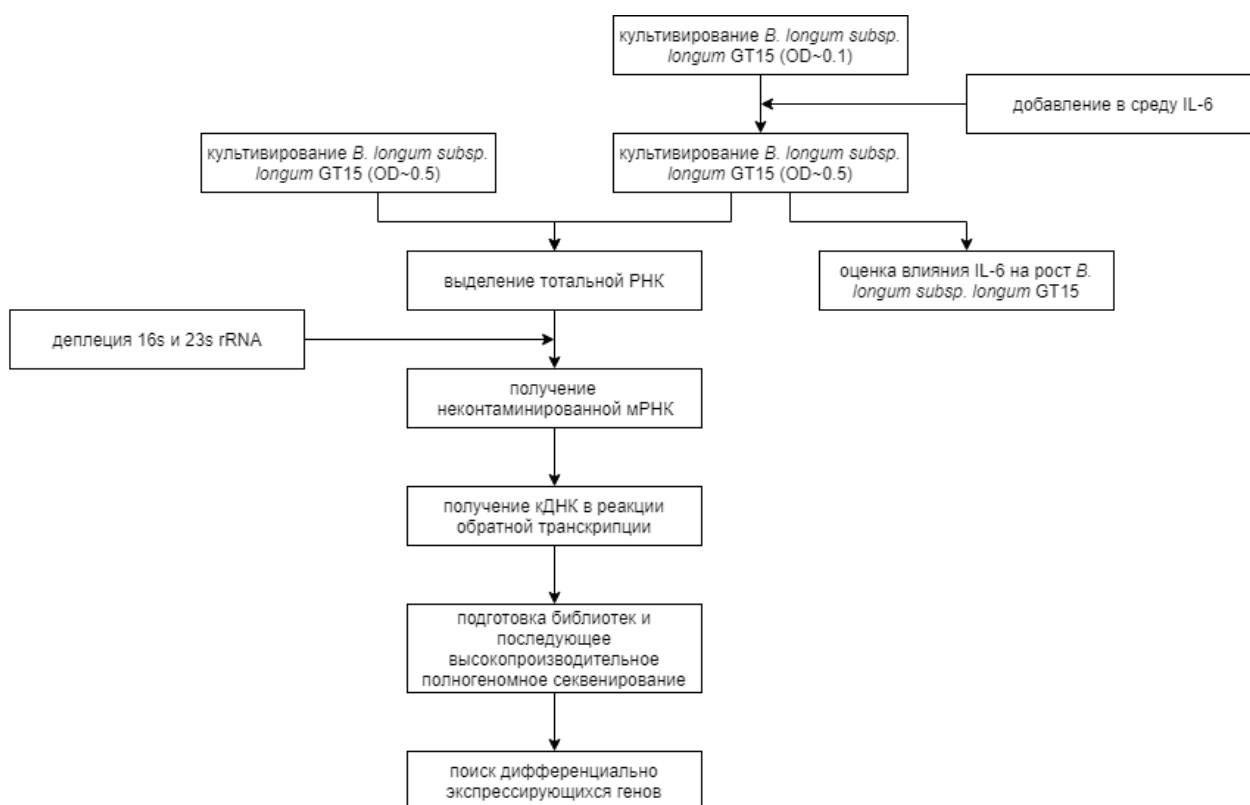


Рисунок 13. Принципиальная схема экспериментального анализа влияния IL-6 на рост и экспрессию генов штамма *B. longum* subsp. *longum* GT15.

Культивирование штамма *B. longum* subsp. *longum* GT15 проводили до значения оптической плотности OD~0,1, затем к экспериментальным культурам добавляли IL-6 (Invitrogen, США) в растворе CH₃COOH (0,1 М) (раствор, в котором производитель IL-6 рекомендует растворять лиофилизат) в различных концентрациях (1 нг/мл, 10 нг/мл), а к контрольным культурам добавляли соответствующий объем CH₃COOH (0,1 М). Выделение тотальной РНК проводили согласно описанному ранее протоколу.

Для обогащения выделенной тотальной РНК *B. longum* subsp. *longum* GT15 последовательностями мРНК проводили деплецию рРНК. 16S и 23S РНК удаляли с использованием Ribo-Zero Bacteria rRNA Removal Kit набора ScriptSeq Complete Kit (Bacteria) (Epicentre, США) в соответствии с инструкцией производителя. Последующую очистку РНК проводили с использованием набора Agencourt RNA Clean XP (Beckman Coulter, США) в соответствии с инструкцией производителя. Качество полученной РНК проверяли на биоанализере Agilent 2100 (Agilent Technologies, США) с использованием набора RNA 6000 Pico Kit (Agilent Technologies, США). Подготовку библиотеки кДНК проводили с использованием набора ScriptSeq Complete Kit (Bacteria) (Illumina, США) в соответствии с инструкциями производителя. Библиотеку подвергали финальной очистке на магнитных частицах Agencourt AMPure XP System (Beckman Coulter, США). Размер и качество библиотеки оценивали на биоанализере Agilent 2100 (Agilent Technologies, США) с использованием чипа High Sensitivity DNA Chip (Agilent Technologies, США) согласно протоколу производителя. Количественно библиотеку оценивали с использованием набора Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, США) на приборе Qubit 2.0 Fluorometer (Thermo Fisher Scientific, США). Секвенирование полученной библиотеки (12 pM) проводили на приборе Illumina MiSeq, используя набор парных прочтений 2×250 bp с добавлением 10% Phix в качестве контроля.

Поиск дифференциально экспрессирующихся генов (ДЭГ) проводили согласно опубликованному стандартному протоколу [Yalamanchili, Wan and Liu, 2017]. Эксперимент проводили в трех повторностях.

2.2.5 Определение полногеномной нуклеотидной последовательности штамма *B. angulatum* GT102.

Культивирование штамма *B. angulatum* GT102 проводили согласно описанному ранее протоколу. Была произведена паспортизация и депонирование штамма *B. angulatum* GT102 во Всероссийской коллекции промышленных микроорганизмов (ВКПМ) (Москва, Россия). Идентификационный номер штамма в коллекции ВКПМ: Ас-1973.

Выделение геномной ДНК *B. angulatum* GT102 проводили согласно описанному ранее протоколу.

Для проведения секвенирования генома *B. angulatum* GT102 была использована платформа 454 GS Junior (Roche, Швейцария). Пробоподготовка для секвенирования была проведена в соответствии с протоколом производителя с использованием рекомендованных реагентов. Библиотеки фрагментов ДНК (ShotGun) готовили по стандартному протоколу (Rapid Library). Секвенирование библиотек фрагментов на приборе Roche 454 GS Junior проводили с использованием версии реактивов Titanium (Roche). В результате полногеномного секвенирования было получено 105590 коротких прочтений. Сборка отдельных прочтений фрагментов ДНК в контиги проводилась с помощью программного обеспечения 454 GS De Novo Assembler (v. 3.0) с использованием параметров по умолчанию. В результате сборки были получены 17 контигов общей длиной 2046935 п. н. со средним 34-кратным покрытием.

Сборку генома в статусе «complete» проводили методом секвенирования по Сэнгеру недостающих фрагментов. Амплификацию фрагментов проводили с использованием набора реактивов FastStart High Fidelity PCR System (Sigma-Aldrich, США). Визуализацию и очистку ПЦР-продуктов проводили согласно описанным ранее протоколам. Праймеры для соединения контигов геномной последовательности *B. angulatum* GT102 подбирали на фланкирующие участки соседних контигов. Порядок контигов определяли с помощью программы Mauve (v. 2.3.1) [Darling, Mau and Perna, 2010].

Геномная последовательность была депонирована в базу данных GenBank (NCBI) под номером CP014241. Автоматическую аннотацию генома проводили с помощью NCBI Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP) (<http://www.ncbi.nlm.nih.gov/genomes/static/Pipeline.html>).

ГЛАВА 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

3.1 Биоинформатический анализ генетического окружения гена *pkb2* в геномах представителей различных видов рода *Bifidobacterium*.

Ранее в геномах бифидобактерий были обнаружены и охарактеризованы шесть СТПК [Nezametdinova *et al.*, 2014; Zakharevich *et al.*, 2015; Алексеева и др., 2015]. Видоспецифическая СТПК *Pkb2*, демонстрирующая аномально высокий уровень дивергенции последовательностей различных видов по сравнению с другими охарактеризованными СТПК, и ее генетическое окружение являются объектом исследования данной диссертационной работы.

Эволюционно стабильная группа генов в генетическом окружении *pkb2* была определена на основе анализа филогенетических профилей *in silico*, включавшего в себя получение списков групп ортологичных белков для 109 представителей различных видов бифидобактерий (табл. 1 Приложений), построение филогенетических профилей ортогрупп и вычисление величин индекса взаимной информации на основе попарного сравнения филогенетических профилей ортогрупп с последующей кластеризацией (рис. 9). Ортогруппой называется группа генов в геномах существующих видов, которые происходят от одного гена у последнего общего предка. Филогенетический профиль ортогруппы представляет собой бинарный вектор, отображающий схему возникновения и потери гена в геномах группы исследуемых организмов. В анализ была включена геномная последовательность штамма *B. angulatum* GT 102, для которого было проведено полногеномное секвенирование и сборка генома в статусе «complete», что обеспечило более высокую точность анализа.

Группы ортологичных белков OG0001048, OG0001061, OG0001049 и OG0001050, кодируемых генами *pkb2*, *fn3*, *duf58* и *tgm*, соответственно, обладают схожими филогенетическими профилями (рис. 14), что обеспечивает высокие значения индекса взаимной информации (табл. 12). Предполагается, что эволюция генов, принадлежащих к одному функциональному пути, происходит совместно, поэтому гены, характеризующиеся сходными или инвертированными ФП, могут выступать в качестве кандидатов на роль функциональных партнеров или компенсаторных механизмов, соответственно [Pellegrini *et al.*, 1999; Pellegrini, 2012]. Филогенетический профиль группы ортологичных белков OG0000354, кодируемых геном *aaa-atp* в составе кластера PFNA, отличается от филогенетических профилей ортогрупп OG0001048, OG0001061, OG0001049 и OG0001050 (рис. 14). Согласно проведенному анализу, помимо ортологов гена *aaa-atp* ортогруппа OG0000354 содержит также паралоги данного гена, обнаруживаемые в геномах

бифидобактерий, у которых отсутствует кластер PFNA. Кроме того, в геномах *B. breve* гены *fn3* и *aaa-atp* объединены, вследствие чего гибридный белок был отнесен к ортогруппе OG0001061. Это приводит к формированию отличного филогенетического профиля и, следовательно, низкому значению индекса взаимной информации ($MI = 0,05-0,06$) (табл. 12).



Рисунок 14. Филогенетические профили групп ортологичных белков, кодируемых генами кластера PFNA. Указано количество геномов, содержащих данный ген, а также вектор представленности гена в геномах представителей различных видов бифидобактерий, где пропуски указывают на отсутствие гена в отдельном геноме.

Таблица 12. Значение величин индекса взаимной информации на основе попарного сравнения филогенетических профилей групп ортологичных белков, кодируемых генами кластера PFNA.

	OG0000354	OG0001048	OG0001049	OG0001050	OG0001061	OG0001260	OG0001261	OG0001487
	4	8	9	0	1	0	1	7
OG0000354	1							
OG0001048	0,06	1						
OG0001049	0,06	0,77	1					
OG0001050	0,06	0,77	0,77	1				
OG0001061	0,05	0,70	0,70	0,70	1			
OG0001260	0,18	0,09	0,09	0,09	0,07	1		
OG0001261	0,19	0,10	0,10	0,10	0,08	0,93	1	
OG0001487	0,16	0,14	0,14	0,08	0,15	0,28	0,27	1

Таким образом, с помощью анализа филогенетических профилей мы установили, что генетическое окружение гена *pkb2* включает в себя следующую эволюционно стабильную группу генов, совместно обнаруживаемых в геномах представителей различных видов рода *Bifidobacterium*: гены *pkb2*, *fn3*, *aaa-atp*, *duf58* и *tgm*. Кластер генов получил название PFNA (*pkb2*, *fn3*, *aaa-atp*). Данные гены кодируют СТПК Pkb2 (*pkb2*); протяженный белок FN3, в структуре которого аннотированы FNIII домены, содержащие мотивы, характерные для цитокиновых рецепторов (*fn3*); предполагаемую MoxR АТФазу AAA-АТР (*aaa-atp*); белок Duf58 с неизвестной функцией, в структуре которого аннотирован домен DUF58 (*duf58*), а также предполагаемую трансглутаминазу Tgm (*tgm*). Гены кластера имеют общее

направление транскрипции и являются, по-видимому, функциональными партнерами.

Системы сигнальной трансдукции обеспечивают связь бифидобактерий с внешней средой, представленной организмом хозяина, и могут быть потенциальными участниками формирования механизмов адаптации. Однокомпонентные сигнальные системы, представленные серин-треониновыми протеинкиназами и обеспечивающие передачу сигнала путем обратимого фосфорилирования белков, по всей видимости, являются доминирующими у прокариот [Alber, 2009; Ruggiero *et al.*, 2012; Ulrich *et al.*, 2009]. СТПК Pkb2 представляет собой трансмембранный белок и имеет внеклеточный регион, отвечающий, по-видимому, за связывание лигандов и выполняющий сигнальную функцию.

Как было отмечено ранее, одним из субстратов фосфорилирования исследуемой СТПК Pkb2 является глутаминсинтаза GlnA1. Ортологи глутаминсинтазы GlnA1 были обнаружены во внеклеточных протеомах нескольких штаммов бифидобактерий [Vazquez-Gutierrez *et al.*, 2017]. Белок GlnA1 *B. animalis* subsp. *lactis* способен связываться с плазминогеном человека и обеспечивать адгезию бифидобактерий к эпителию кишечника [Candela *et al.*, 2007]. Белок GlnA1 патогенной бактерии *Mycobacterium tuberculosis* способен связываться с плазминогеном и фибронектином [Kainulainen and Korhonen, 2014]. Анализ субстратов фосфорилирования позволил выдвинуть предположение, что функции СТПК Pkb2 могут быть связаны с адгезией и коммуникацией бифидобактерий с эпителиальными клетками кишечника [Nezametdinova *et al.*, 2018].

Одним из субстратов фосфорилирования СТПК Pkb2 является также АТФаза, кодируемая геном *aaa-atp* в составе кластера PFNA [Nezametdinova *et al.*, 2018]. Продукт гена *aaa-atp* аннотирован как АТФаза, ассоциированная с широким спектром клеточных активностей (*англ.* ATPases associated with diverse cellular activities, AAA+). Суперсемейство АТФаз AAA+ обладает значительным функциональным разнообразием [Snider J, Thibault and Houry, 2008]. Нами была проведена кластеризация с другими АТФазами данного суперсемейства, в результате чего мы отнесли АТФазу AAA-АТР к подсемейству MRP (*англ.* MoxR Proper) семейства MoxR. Согласно опубликованным данным, подсемейство MRP является наиболее распространенным внутри семейства MoxR. К данному подсемейству принадлежат 48,2% всех исследуемых белков MoxR [Snider and Houry, 2006]. Экспериментально показанные функции для АТФаз подсемейства MRP заключаются в моделировании активности субстратов. В частности, у *Paracoccus denitrificans* ген *moxR* локализован рядом с кластером MDH, отвечающим за экспрессию метанолдегидрогеназы. Данный ген, по-видимому, участвует в регуляции MDH путем модификации и активации субъединиц MDH или регулируя гены, ответственные за экспрессию активаторов MDH [Van Spanning *et al.*, 1991]. У *Methylobacterium extorquens* АТФаза подсемейства MRP участвует в

инсерции кофактора Ca^{2+} в MDH [Toyama, Anthony, Lidstrom, 1998]. В общем случае белки MRP могут быть вовлечены в процессы инсерции кофакторов, обеспечивая, вероятно, функцию молекулярного шаперона [Snider and Houry, 2006].

Анализ генетического окружения генов, кодирующих MRP в геномах представителей трехсот видов микроорганизмов, показал, что в большинстве случаев данные гены сопровождается генами, кодирующими белок, в структуре которого аннотирован домен DUF58, а также примерно в трети случаев генами, кодирующими предполагаемую транскляминазу [Wong and Houry, 2012]. Аналогичное строение демонстрирует кластер генов PFNA. Известные функции транскляминаз заключаются в посттрансляционной модификации субстратов с образованием изопептидных связей, устойчивых к протеиназному расщеплению [Griffin *et al.*, 2002]. Предполагаемая транскляминаза, кодируемая геном *igm* в составе кластера PFNA, представляет собой политопический трансмембранный белок, подобно многим рецепторам, ионным каналам и транспортерам, что указывает на возможное участие во взаимодействиях с факторами внешней среды, представленной организмом хозяина.

Как было отмечено ранее, функция белка, кодируемого геном *fn3* в составе кластера генов PFNA бифидобактерий, связана с реализацией механизмов адгезии. Данная функция была экспериментально показана для штамма *B. bifidum* S17 [Westermann, 2015]. Фибронектиновые домены типа III широко распространены в структуре эукариотических молекул клеточной адгезии (*англ.* cell adhesion molecules, CAMs) суперсемейства IgSF [Preedy, 2016], а также обнаруживаются в структуре бактериальных белков, которые отвечают за адгезию к эпителиальным клеткам организма хозяина за счет связывания фибронектина [Konkel, Larson and Flanagan, 2010].

Аннотированные в структуре FNIII доменов мотивы, характерные для цитокин-связывающего региона рецептора gp-130, указывают также на возможное взаимодействие белка FN3 с цитокинами, продуцируемыми иммунной системой хозяина. Семейство лигандов рецептора gp-130 включает IL-6 и некоторые другие цитокины, принадлежащие к группе IL-6 [Bravo *et al.*, 1998]. Как было отмечено ранее, микроорганизмы способны воспринимать сигнальные молекулы иммунной системы организма хозяина и отвечать на воздействие изменением скорости роста и ряда других характеристик [Lesouhaitier *et al.*, 2009]. Механизмы, обеспечивающие выживание комменсалов в условиях воспалительного процесса и постоянство репертуара нормальной микробиоты кишечника, остаются малоизученными. На сегодняшний день исследовательскими группами опубликованы немногочисленные работы, посвященные изучению данных механизмов у отдельных представителей комменсальной микробиоты. В частности, в качестве механизма

устойчивости комменсалов к факторам иммунного ответа предполагается модификация липополисахаридов, что повышает резистентность к антимикробным пептидам у представителей типа *Bacteroidetes* [Cullen *et al.*, 2015]. Предполагаемая ассоциированная с мембраной фосфолипидфосфатаза BT1854 у *B. thetaiotaomicron* участвует в реализации вышеупомянутого механизма устойчивости. Среди актинобактерий повышенная резистентность к антимикробным пептидам была показана только для *C. aerofaciens* [Cullen *et al.*, 2015], механизмы устойчивости представителей рода *Bifidobacterium* к факторам иммунного ответа не изучались. Мы провели поиск ортологов гена, кодирующего BT1854 *B. thetaiotaomicron*, во всех доступных для анализа (октябрь, 2019 г.) секвенированных полногеномных последовательностях бифидобактерий. Данный ген в геномах бифидобактерий обнаружен не был, что позволяет предполагать существование иных механизмов и генных путей, выполняющих данную функцию. Вероятным кандидатом является ген *fn3* в составе кластера генов PFNA.

В результате анализа филогенетических профилей было установлено, что кластер PFNA *B. breve* содержит гибридный ген, полученный в результате объединения генов *fn3* и *aaa-atp* без сдвига рамки считывания. Следует отметить, что эволюционные события объединения генов часто трактуются как доказательство физического взаимодействия кодируемых ими белков, так как в результате объединения транслируется гибридный белок, содержащий функциональные домены обоих типов [Enright *et al.*, 1999; Marcotte *et al.*, 1999 a, b; Sali, 1999]. Таким образом, благодаря обнаружению события объединения генов *fn3* и *aaa-atp* в геномах *B. breve*, мы предсказали *in silico* взаимодействие между белками AAA-АТР и FN3.

Мы показали, что гены в составе кластера PFNA, подобно *pkb2*, обладают высоким уровнем межвидовой дивергенции последовательностей по сравнению с коровыми генами (табл. 13).

Таблица 13. Минимальный процент идентичности аминокислотных последовательностей, полученный в результате множественного выравнивания белков, кодируемых генами кластера PFNA у представителей различных видов рода *Bifidobacterium*. Для сравнения представлены данные анализа уровня гомологии консервативной СТПК Pkb5.

	Pkb5	Pkb2	FN3	AAA-АТР	Duf58	Tgm
минимальный процент идентичности	0,68	0,37	0,38	0,5	0,32	0,38

В геномах представителей ряда видов бифидобактерий, в том числе в геноме *B. longum* subsp. *longum* GT15, в генетическом окружении представлены дополнительно гены, кодирующие предполагаемую серин-треониновую фосфатазу ортогруппы OG0001260 (*prpC*) и гипотетический белок ортогруппы OG0001261, в структуре которого известные домены не аннотированы, а также в отдельных случаях ген *fha*, кодирующий белок ортогруппы OG0001487, в структуре которого аннотирован домен FHA. Данные гены расположены в направлении транскрипции генов кластера PFNA непосредственно после гена *tgm*. Серин-треониновая фосфатаза, кодируемая геном *prpC*, может обеспечивать функцию сигнальной трансдукции совместно с СТПК Pkb2, осуществляя дефосфорилирование фосфорилированных субстратов. Домен FHA в составе гипотетического белка, кодируемого геном *fha*, взаимодействует с белковым партнером в процессе, регулируемом обратимым фосфорилированием, связывая остатки фосфотреонина, а также другие фосфорилированные остатки [Pallen, Chaudhuri and Khan, 2002]. Таким образом, данный белок предположительно участвует в процессе СТПК-опосредованного сигналинга. Филогенетические профили ортогрупп OG0001260 и OG0001261 демонстрируют высокое значение индекса взаимной информации (MI = 0,93) и также представляют собой эволюционно стабильную группу генов, являясь, по-видимому, функциональными партнерами. Несмотря на относительно невысокое значение индекса взаимной информации между ортогруппой OG0001487 и ортогруппами OG0001260 и OG0001261 (MI = 0,28 и MI = 0,27, соответственно), в результате кластеризации данные ортогруппы оказываются в составе одной клады (рис. 15), что косвенно может свидетельствовать об участии вышеупомянутых белков в реализации общей функции. Анализ филогенетических профилей позволил обнаружить также предполагаемого функционального партнера (OG0001137) для белков ортогрупп OG0001260 и OG0001261. Потеря гена, кодирующего белки ортогруппы OG0001137 в ряде геномов бифидобактерий происходит одновременно с потерей генов, кодирующих белки ортогрупп OG0001260 и OG0001261, в результате чего данные белки оказываются в составе общей клады (рис. 15). Белки ортогруппы OG0001137 аннотированы как белки семейства гистидиновых фосфатаз. Гистидиновые киназы и фосфатазы наряду с серин-треониновыми киназами и фосфатазами обеспечивают сигнальную трансдукцию у прокариот [Wolanin, Thomason and Stock, 2002].

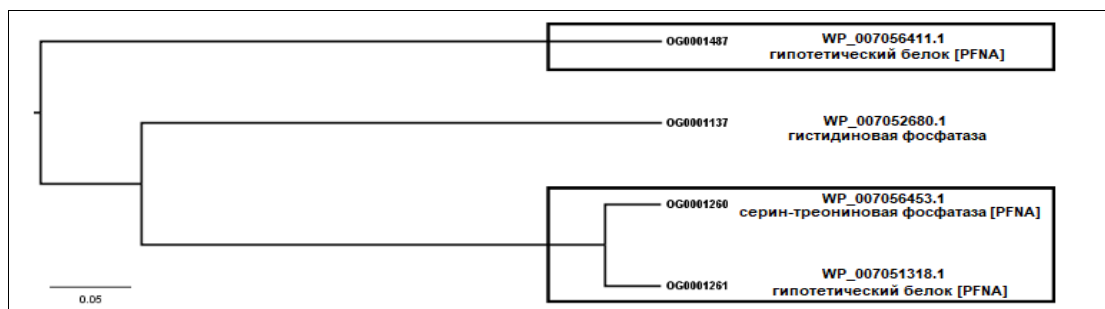


Рисунок 15. Фрагмент кладограммы, полученной в результате кластеризации на основе матрицы расстояний между филогенетическими профилями групп ортологичных белков бифидобактерий. В результате кластеризации ортогруппы OG0001487, OG0001260 и OG0001261 оказываются в составе одной клады.

Кластер генов PFNA является уникальным для рода *Bifidobacterium* и не обнаруживается в секвенированных геномах других организмов, в том числе, наиболее близкородственных [Nezametdinova *et al.*, 2018]. Мы показали, что кластер содержит эволюционно стабильную группу генов, а также в некоторых случаях дополнительные гены, обнаруживаемые в геномах отдельных представителей бифидобактерий (рис. 16). В общем случае, функцией кластера генов PFNA является, по-видимому, взаимодействие с факторами внешней среды, представленной организмом хозяина. Высокая степень межвидовой дивергенции последовательностей генов кластера PFNA у разных видов указывает на видоспецифичность выполняемых функций.

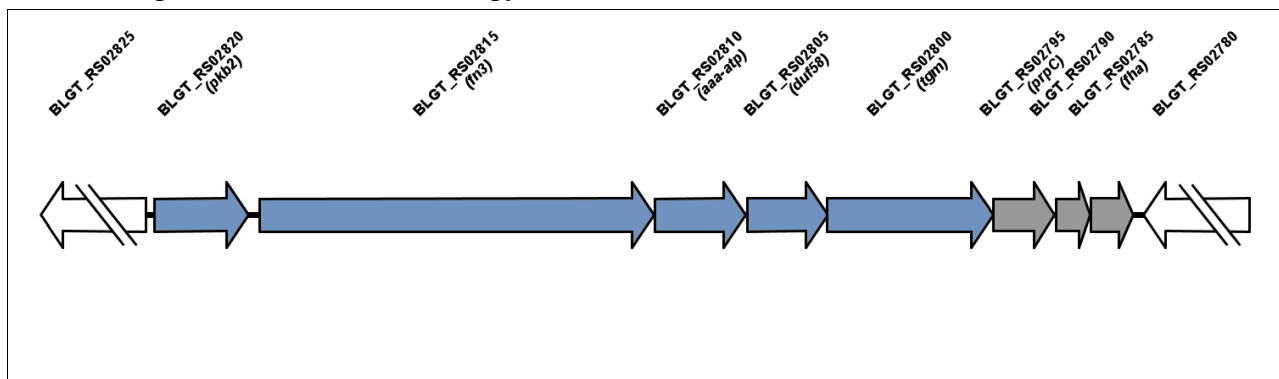


Рисунок 16. Структура кластера генов PFNA в геноме *B. longum* subsp. *longum* GT15. Указаны гены, формирующие коровую структуру кластера PFNA у представителей различных видов бифидобактерий (голубой цвет), а также дополнительные гены, представленные в геномах отдельных видов (серый цвет).

В результате анализа филогенетических профилей было получено также филогенетическое дерево анализируемых геномов. Филогенетическое дерево, полученное на основе данных о наличии/отсутствии определенных ортогрупп в отдельных геномах считается менее точным, чем филограммы, построенные на основе множественных выравниваний последовательностей коровых генов [Kelly and Maini, 2013], однако данное

дерево воспроизводит внешние ветви и клады опубликованных филограмм бифидобактерий с высокой точностью [Lugli *et al.*, 2014; Lugli *et al.*, 2017; Lugli *et al.*, 2018; Sun *et al.*, 2015].

На основе филогенетических профилей групп ортологических белков, кодируемых генами кластера PFNA, была проведена реконструкция эволюционных событий потери генов (англ. gene loss) в различных ветвях и кладах вышеупомянутого филогенетического дерева (рис. 17).

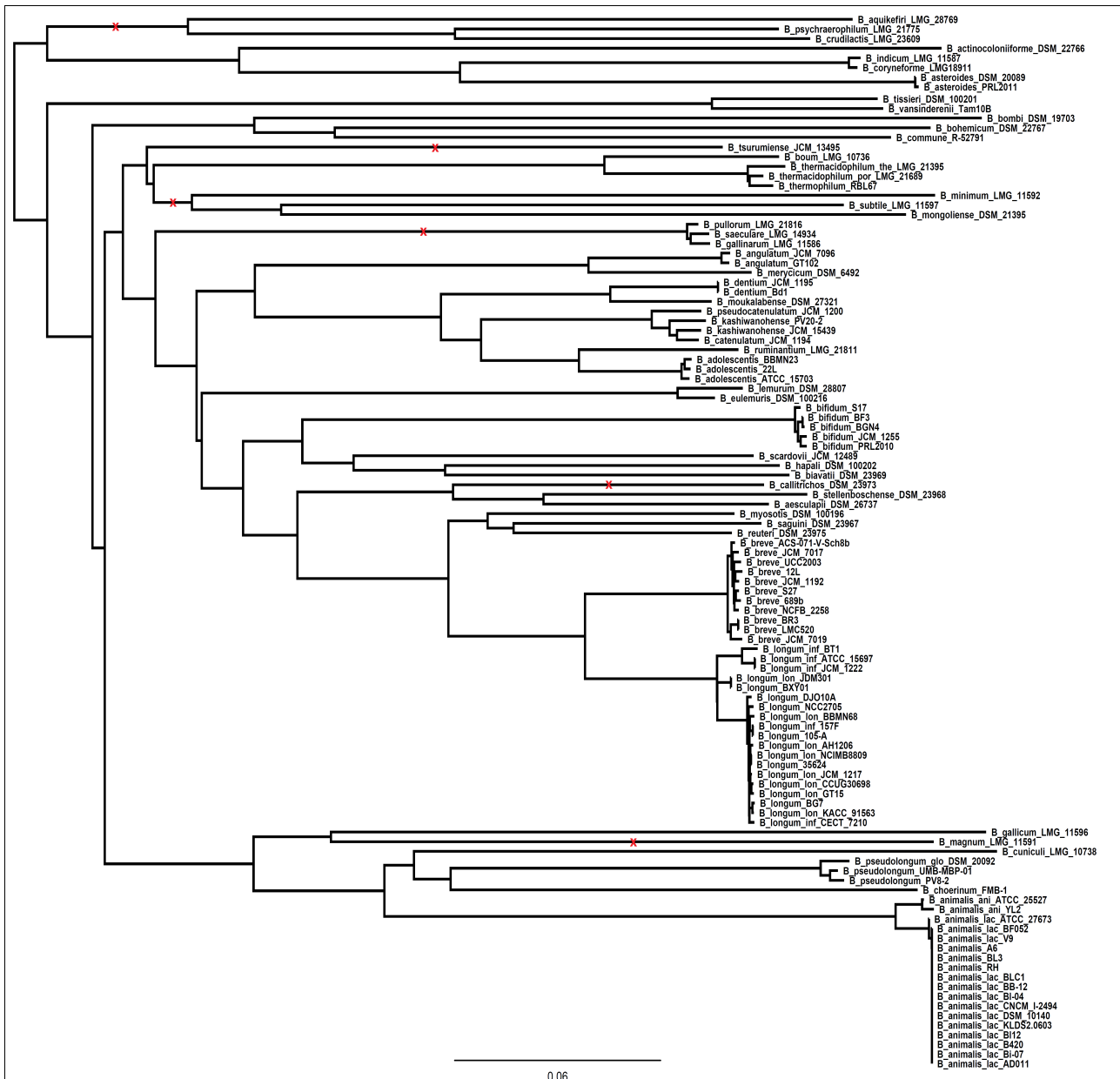


Рисунок 17. Филогенетическое дерево, полученное в результате применения алгоритма DendroBLAST для 109 геномов представителей различных видов рода *Bifidobacterium*. Указаны ветви, в которых локализованы события потери кластера генов PFNA.

В результате эволюционных событий потери генов в филогении бифидобактерий, кластер PFNA отсутствует у ряда видов: *B. aquikefiri*, *B. psychraerophilum*, *B. crudilactis*, *B. tsurumiense*, *B. minimum*, *B. subtile*, *B. mongoliense*, *B. pullorum*, *B. saeculare*, *B. gallinarum*, *B. callitrichos* и *B. magnum*. Вероятно, данные события потери генов сопряжены с реализацией механизмов компенсации функции. В частности, у представителей отдельных видов бифидобактерий (*B. stellenboschense*, *B. gallinarum*, *B. pullorum*, *B. saeculare*, *B. bohemicum*, *B. minimum*) нами была обнаружена ранее неохарактеризованная СТПК (OG0003463), образующая эволюционно стабильную группу с белком, в составе которого аннотирован бактериальный Ig-подобный домен (OG0003460). Ig-подобные домены входят в состав молекул клеточной адгезии суперсемейства IgSF, к которому предположительно относится также белок FN3 кластера PFNA. У видов, относящихся к филогруппе *B. pullorum* (*B. gallinarum*, *B. pullorum*, *B. saeculare*), а также вида *B. minimum* белки ортогрупп OG0003463 и OG0003460 предположительно могут участвовать в реализации компенсаторных механизмов, так как геномы данных видов кластер PFNA не содержат.

3.2 Изучение транскрипционной организации кластера PFNA в геноме *B. longum* subsp. *longum* GT15.

Кластер PFNA представляет собой эволюционно стабильную группу генов, вовлеченных, по-видимому, в выполнение общей функции. Наиболее эволюционно устойчивые функциональные ассоциации демонстрируют, как правило, гены в оперонах [Moreno-Hagelsieb, 2015]. Для подтверждения данной гипотезы мы изучили транскрипционную организацию кластера PFNA в геноме *B. longum* subsp. *longum* GT15.

Экспериментальный анализ включал в себя выделение тотальной РНК *B. longum* subsp. *longum* GT15, проведение реакции обратной транскрипции, постановка ПЦР на матрице полученной кДНК с использованием набора специфических праймеров с целью подтверждения предполагаемой оперонной организации, а также определение точки старта транскрипции методом быстрой амплификации 5'-концов кДНК (англ. rapid amplification of cDNA 5'-ends, 5'-RACE). Поиск точки терминации транскрипции проводили *in silico*.

В результате транскрипционного анализа были получены ампликоны, соответствующие спейсерным участкам, что подтверждает полицистронность матрицы мРНК (рис. 18). Таким образом, гены *pkb2*, *fn3*, *aaa-atp*, *duf58*, *tgm*, а также гены *prpC*, *fha* и ген BLGT_RS02790 в геноме *B. longum* subsp. *longum* GT15 транскрибируются в составе единого оперона. Оперонная организация подтверждает участие генов кластера PFNA в реализации общей функции.

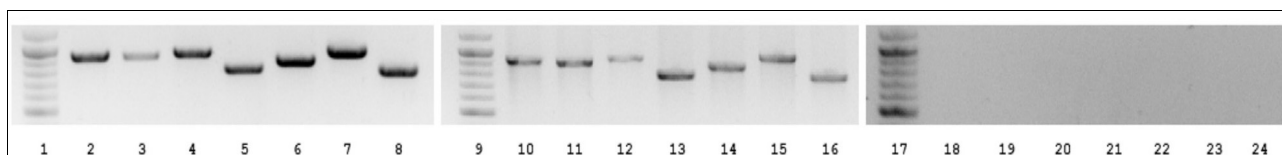


Рисунок 18. Электрофореграмма результатов амплификации фрагментов sp1-sp7. Дорожки: 1, 9, 17 - маркер длин ДНК (Fermentas 100bp Plus); 2-8, 10-16, 18-24 — ампликоны на матрице геномной ДНК, кДНК и тотальной РНК, соответственно; 2, 10, 18 — ампликон sp1; 3, 11, 19 — ампликон sp2; 4, 12, 20 — ампликон sp3; 5, 13, 21 — ампликон sp4; 6, 14, 22 — ампликон sp5; 7, 15, 23 — ампликон sp6; 8, 16, 24 — ампликон sp7.

Метод 5'-RACE был использован нами для поиска точки старта транскрипции оперона PFNA. В результате пошагового «заглубления» геноспецифических праймеров в нескольких раундах ПЦР-амплификации (рис. 19), были получены ампликоны, фланкированные с 3'-конца последовательностью геноспецифического праймера, а с 5'-конца последовательностью нетранслируемого участка, включая точку старта транскрипции (англ. 5'-untranslated region, 5'-UTR). Для полученных индивидуальных ПЦР-ампликонов была определена нуклеотидная последовательность для локализации точки старта транскрипции. В результате была обнаружена точка старта транскрипции оперона, находящаяся на расстоянии -40-41 п.о. от стартового кодона белок-кодирующей последовательности гена *pkb2* (рис. 21). Также дополнительно была обнаружена точка старта транскрипции, локализованная внутри последовательности гена *tgm*, что может свидетельствовать о комплексной регуляции транскрипции генов оперона. Дополнительная точка старта транскрипции находится на расстоянии -58-61 п.о. от стартового кодона белок-кодирующей последовательности гена *prpC* (рис. 21).

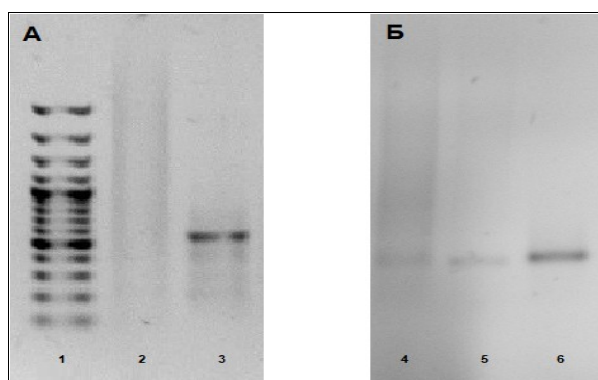


Рисунок 19. Электрофореграмма результатов пошагового заглубления геноспецифических праймеров в нескольких раундах амплификации 5'-RACE. А — определение точки старта транскрипции гена *pkb2*, Б — определение точки старта транскрипции гена *prpC*. Дорожки: 1 — маркер длин ДНК (Fermentas 100bp Plus), 2 — первый раунд амплификации 5'-RACE с геноспецифическим праймером *pkb2*-realC, 3 — второй раунд амплификации 5'-RACE с геноспецифическим праймером *sp0*-*pkb2*-F; 4 — первый раунд

амплификации 5'-RACE с геноспецифическим праймером prpc-F, 5 – второй раунд
 амплификации 5'-RACE с геноспецифическим праймером sp5-prpC-F, 6 – третий раунд
 амплификации 5'-RACE с геноспецифическим праймером sp5-prpC-F1.

Точка терминации транскрипции была предсказана на основании анализа структуры последовательностей спейсеров и генов оперона PFNA и его окружения в геномах *B. longum* subsp. *longum* GT15 и других представителей *B. longum* subsp. *longum*. Нами было проведено сравнение последовательностей, содержащих р-независимый терминатор транскрипции кластера генов PFNA в геномах штаммов LMG 13197, BBMN60 и KACC 91563 *B. longum* subsp. *longum* (рис. 20, А) с последовательностью участка, содержащего предполагаемый терминатор в геноме *B. longum* subsp. *longum* GT15. Анализ выравнивания выявил абсолютную консервативность последовательности анализируемого участка в исследуемых штаммах (рис. 20, Б). Таким образом, в результате был обнаружен р-независимый терминатор, находящийся на расстоянии +65 п.о. от стоп-кодона белок-кодирующей последовательности гена *fha* (рис. 21). Другие р-независимые терминаторы в структуре оперона обнаружены не были.

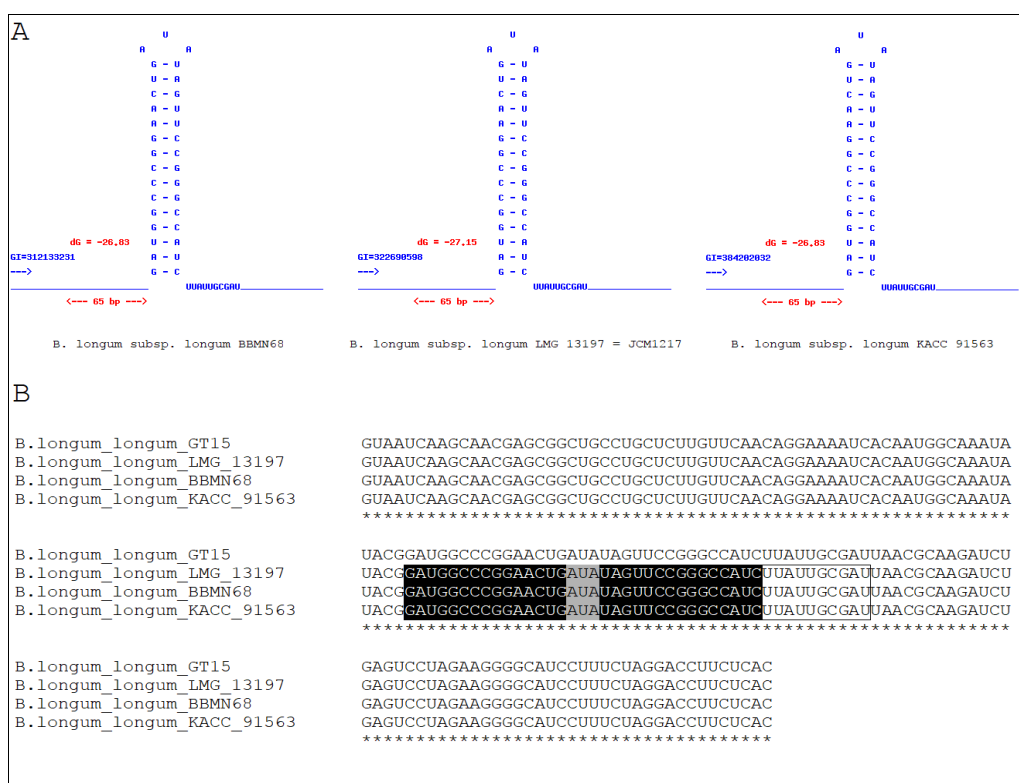


Рисунок 20. Поиск точек терминации транскрипции *in silico* внутри видоспецифического кластера *B. longum subsp. longum* GT15; А – предсказанные р-независимые терминаторы внутри региона интереса в геномах *B. longum subsp. longum*; В – выравнивание спейсерной области, содержащей предсказанные терминаторы, в геномах *B. longum subsp. longum*; указаны последовательности, формирующие стебель (черный бокс), петлю (серый бокс) и U-богатый участок (белый бокс) терминатора.

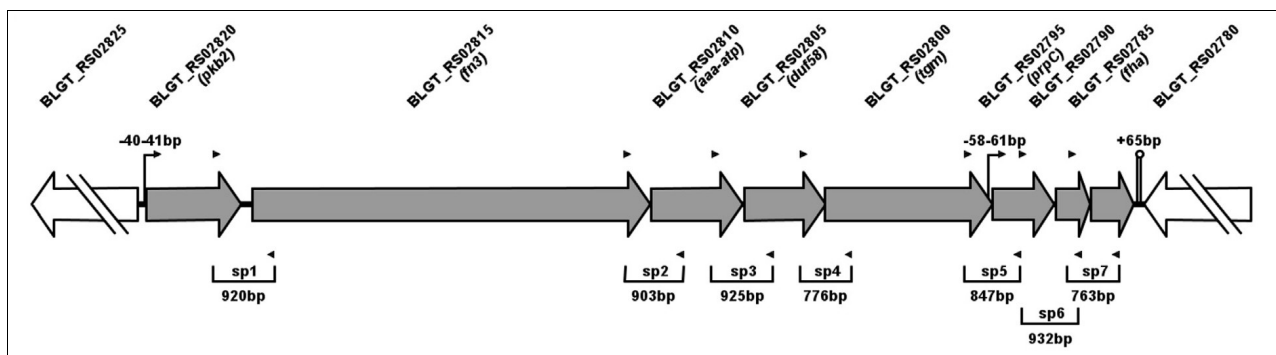


Рисунок 21. Выявленная транскрипционная организация кластера генов PFNA *B. longum* subsp. *longum* GT15. Для подтверждения транскрипционной организации были получены ампликоны sp1-sp7, соответствующие последовательностям межгенных спейсеров, а также участкам прилегающих генов. Указаны сайты посадки специфических праймеров (черные треугольники), длины амплифицируемых спейсерных регионов, локализация точек старта транскрипции (черные стрелки), а также локализация точки ρ -независимой терминации транскрипции (черная шпилька).

3.3 Анализ молекулярной эволюции последовательностей генов кластера PFNA в геномах представителей различных видов рода *Bifidobacterium*.

Как было отмечено ранее, быстрая эволюция и положительный отбор нередко затрагивают гены, вовлеченные в коммуникацию с факторами внешней среды, и особенно, биотическими экологическими факторами. Гипотеза Черной Королевы описывает эффекты, оказываемые на кодирующие последовательности генов в результате совместной эволюции взаимодействующих видов. Аномально высокий уровень дивергенции последовательностей указывает на вероятные события положительного отбора/ ослабления отрицательного отбора в молекулярной эволюции кластера PFNA, обеспечившие быструю эволюцию генов. В контексте роли генов кластера PFNA в коммуникации бифидобактерий с факторами занимаемой экологической ниши в организме хозяина, обнаружение положительного отбора могло бы послужить ярким примером эффекта Черной Королевы. Возможной причиной вступления комменсалов в «гонку вооружений» в рамках гипотезы Черной Королевы является адаптация к условиям обитания в микробном сообществе ЖКТ, например, конкуренция с другими комменсалами и патогенами за экологическую нишу, а также давление, оказываемое со стороны иммунной системы организма хозяина.

Для подтверждения данной гипотезы мы провели *in silico* поиск событий положительного отбора в молекулярной эволюции генов кластера PFNA. Анализ молекулярной эволюции проводили с помощью метода максимального правдоподобия. Используемый метод позволяет детектировать эволюционные события продолжительного [Yang *et al.*, 2000] или эпизодического [Zhang, Nielsen and Yang, 2005] положительного отбора

в нуклеотидных последовательностях белок-кодирующих генов. Подготовка данных для последующего анализа молекулярной эволюции включала в себя получение нуклеотидных последовательностей групп ортологичных генов кластера PFNA представителей различных видов рода *Bifidobacterium*, множественное выравнивание последовательностей и отбор консервативных блоков выравнивания, определение наилучшей схемы партиций и модели молекулярной эволюции, построение филогенетического дерева бифидобактерий на основе последовательностей генов кластера PFNA, а также подбор оптимальных значений параметров анализа (рис. 10).

3.3.1 Филогенетический анализ, основанный на сравнении последовательностей генов кластера PFNA представителей различных видов рода *Bifidobacterium*.

Сходный уровень дивергенции последовательностей является косвенным подтверждением коэволюции генов в составе кластера PFNA бифидобактерий. Экспериментально установленная оперонная организация подтверждает участие в реализации общей функции, что также указывает на совместную эволюцию вышеупомянутых генов. Для подтверждения коэволюции генов кластера PFNA была проведена оценка корреляционной зависимости между значениями попарных эволюционных расстояний внутри индивидуальных филогенетических деревьев, построенных на основе выравниваний 43 последовательностей генов *pkb2* (рис. 1 Приложений), *fn3* (рис. 2 Приложений), *aaa-atp* (рис. 3 Приложений), *duf58* (рис. 4 Приложений) и *tgm* (рис. 5 Приложений). Значения величин коэффициента корреляции Пирсона, полученные в результате попарного сравнения филогенетических деревьев, составили 0,800-0,978 (табл. 14) при значении $P < 1,00E-006$. *B. actinocoloniiforme* демонстрирует наиболее высокое значение длины ветви во всех анализируемых филограммах, кроме филогенетического дерева, построенного на основе множественного выравнивания последовательностей гена *pkb2*, где длина ветви *B. actinocoloniiforme* имеет невысокое значение по сравнению со значениями длин других ветвей филограммы. Таким образом, скорость накопления замен в гене *pkb2* *B. actinocoloniiforme* оказывается ниже, чем в других генах кластера PFNA. Удаление ветви *B. actinocoloniiforme* улучшило значения коэффициентов корреляции между вышеупомянутым деревом и филогенетическими деревьями, построенными на основе множественных выравниваний последовательностей генов *fn3*, *aaa-atp*, *duf58* и *tgm* с 0.800-0.814 до 0.883-0.912.

Таблица 14. Значения величин коэффициента корреляции Пирсона, полученные в результате попарного сравнения филогенетических деревьев, построенных на основе последовательностей генов кластера PFNA. Для филогенетического дерева, построенного на основе множественного выравнивания последовательностей гена *pkb2* представлены значения до и после удаления ветви *B. actinocoloniiforme*, соответственно.

	<i>pkb2</i>	<i>fn3</i>	<i>aaa-atp</i>	<i>duf58</i>	<i>tgm</i>
<i>pkb2</i>	1				
<i>fn3</i>	0.800/ 0.900	1			
<i>aaa-atp</i>	0.814/ 0.902	0.966	1		
<i>duf58</i>	0.814/ 0.883	0.955	0.967	1	
<i>tgm</i>	0.813/ 0.912	0.976	0.978	0.976	1

Высокие значения величин коэффициента корреляции Пирсона, полученные в результате попарного сравнения филогенетических деревьев, подтверждают гипотезу о коэволюции генов кластера PFNA, что дает возможность объединить множественные выравнивания последовательностей генов *pkb2*, *fn3*, *aaa-atp*, *duf58* и *tgm* для последующего построения общего филогенетического дерева и анализа молекулярной эволюции. Использование последовательностей объединенных генов увеличивает статистическую мощность тестов по сравнению с использованием последовательностей индивидуальных генов, а также увеличивает аккуратность построения филогенетического дерева за счет увеличения количества анализируемых замен.

На основе объединенных белок-кодирующих последовательностей генов кластера PFNA 43 представителей различных видов рода *Bifidobacterium* было построено неукорененное филогенетическое дерево (рис. 22).

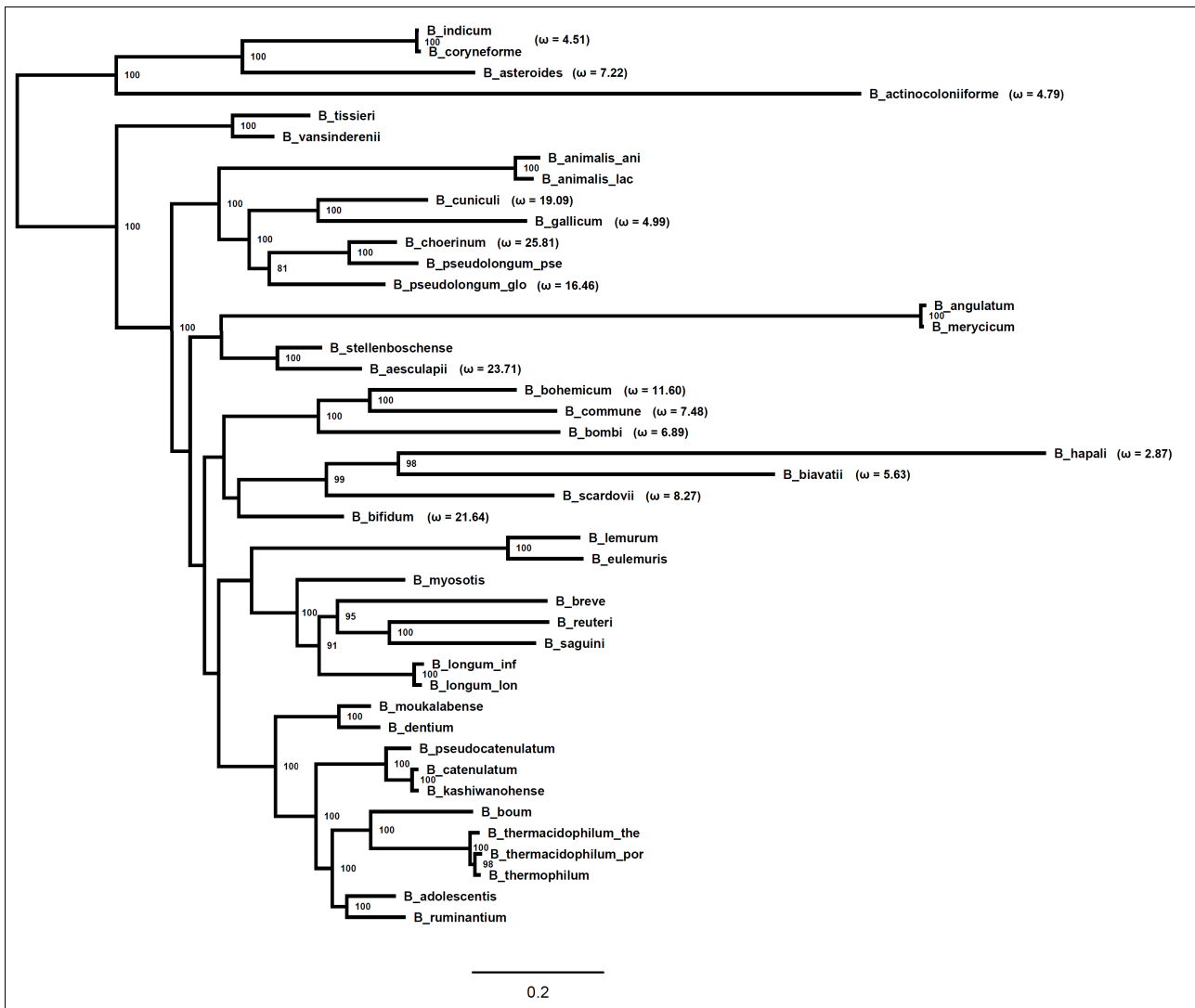


Рисунок 22. Филогенетического дерева, построенное на основе объединенных белок-кодирующих последовательностей генов кластера PFNA 43 представителей различных видов рода *Bifidobacterium*. Значения поддержек внутренних узлов филограммы получены после 10000 итераций бутстреп-анализа в программе RaxML. Поддержки в узлах филограммы со значением менее 70 на рисунке не представлены. Указаны конкретные значения ω для класса сайтов со значением $\omega > 1$ в тех ветвях филограммы, где показано наличие эпизидического положительного отбора в строгих условиях. Представлены оригинальные длины ветвей филогенетического дерева, масштабная линейка соответствует значению две предполагаемые замены на десять нуклеотидных сайтов.

Топология внешних клад полученного филогенетического дерева демонстрирует высокие значения поддержек узлов и с высокой точностью воспроизводит опубликованные филограммы бифидобактерий, несмотря на имеющиеся незначительные отличия. Топология полученного дерева воспроизводит следующие филогруппы рода *Bifidobacterium*: группы *B. asteroides*, *B. pseudolongum*, *B. longum*, *B. boum* [Sun *et al.*, 2015], а также группу *B. bifidum* [Lugli *et al.*, 2014; Lugli *et al.*, 2017; Lugli *et al.*, 2018]. Рассчитанные нами эволюционные

дистанции (E_D) между парами таксонов *B. indicum* и *B. coryneforme* ($E_D = 0,0089$) и *B. catenulatum* и *B. kashiwanohense* ($E_D = 0,02$) оказались меньше, чем для пар таксонов *B. animalis* subsp. *animalis* и *B. animalis* subsp. *lactis* ($E_D = 0.0654$) и *B. longum* subsp. *infantis* и *B. longum* subsp. *longum* ($E_D = 0.0262$), относящихся к одному виду, что свидетельствует о высоком уровне генетического родства. Мы предполагаем, что исследуемые штаммы видов *B. indicum* и *B. coryneforme*, а также *B. catenulatum* и *B. kashiwanohense* должны быть реклассифицированы и объединены в один вид. Напротив, топология ветвей и эволюционные дистанции для пары таксонов *B. pseudolongum* subsp. *pseudolongum* и *B. pseudolongum* subsp. *globosum* ($E_D = 0.4054$) свидетельствуют о несоответствии между полученными нами результатами и текущей классификацией бифидобактерий. Мы предполагаем, что исследуемые штаммы подвидов *B. pseudolongum* subsp. *pseudolongum* и *B. pseudolongum* subsp. *globosum* должны быть отнесены к разным видам. Наши результаты согласуются с ранее предложенной реклассификацией [Lugli *et al.*, 2014].

Топология глубинных ветвей полученного филогенетического дерева является нестабильной, о чем свидетельствуют невысокие значения поддержек узлов. Стоит отметить, однако, что существующие опубликованные филогенетические деревья бифидобактерий, построенные на основе объединенных последовательностей коровых генов также имеют различающиеся конфигурации глубинных узлов и ветвей, несмотря на высокие значения поддержек в каждом отдельном случае [Lugli *et al.*, 2014; Lugli *et al.*, 2017; Lugli *et al.*, 2018; Sun *et al.*, 2015].

Полученное филогенетическое дерево имеет высокие значения длин ветвей по сравнению с длинами ветвей опубликованных филогенетических деревьев, построенных на основе объединенных последовательностей коровых генов бифидобактерий [Lugli *et al.*, 2014; Lugli *et al.*, 2017; Lugli *et al.*, 2018; Sun *et al.*, 2015]. Данный факт указывает на высокую скорость молекулярной эволюции нуклеотидных последовательностей белок-кодирующих участков генов кластера PFNA по сравнению с консервативными последовательностями белок-кодирующих участков коровых генов, на которые, как правило, действует давление отрицательного отбора. Быстрая эволюция генов кластера PFNA обуславливает высокий уровень межвидовой дивергенции последовательностей и может быть следствием воздействия положительного отбора или ослабления воздействия отрицательного отбора.

3.3.2 Тестирование гипотез о наличии событий продолжительного положительного отбора в молекулярной эволюции последовательностей генов кластера PFNA.

В качестве объяснения феномену быстрой эволюции белок-кодирующих участков генов кластера PFNA бифидобактерий была выдвинута гипотеза о наличии событий продолжительного положительного отбора в молекулярной эволюции последовательностей. Были получены значения функций максимального правдоподобия четырех «site» моделей (табл. 15) и затем проведены тесты LRT на наличие в последовательности нескольких классов сайтов с разными значениями ω (M0 в качестве нулевой модели против M3 в качестве альтернативной), а также на наличие в последовательности сайтов, находящихся под давлением положительного отбора ($\omega > 1$) во всех ветвях филогенетического дерева (M8a в качестве нулевой модели против M8 в качестве альтернативной). Значения LRT для обоих тестов оказались статистически значимыми (табл. 16).

Таблица 15. Значения логарифмов функций правдоподобия «site» моделей, тестируемых на наличие продолжительного положительного отбора в последовательностях генов кластера PFNA во всех ветвях филогенетического дерева бифидобактерий.

наименование теста	количество параметров (np)	значение lnL для теста
M0	148	-215057,27
M3	69	-206592,56
M8	68	-206235,92
M8a	67	-206323,52

Таблица 16. Значения тестов LRT на наличие продолжительного положительного отбора в последовательностях генов кластера PFNA во всех ветвях филогенетического дерева бифидобактерий.

наименование теста	значение теста LRT	количество степеней свободы (df)	P
M3 vs. M0	16929,43	79	0
M8 vs. M8a	175,19	1	0

Таким образом, экспериментально *in silico* показано наличие в последовательности сайтов, находящихся под давлением положительного отбора во всех ветвях филогенетического дерева *Bifidobacterium*. Данные сайты были затем идентифицированы с помощью эмпирического Байесовского подхода. Сайты со значением PP > 0,7 были назначены в качестве кандидатов для воздействия положительного отбора. В результате были обнаружены два аминокислотных сайта, находящихся под давлением положительного отбора во всех ветвях филогенетического дерева бифидобактерий: 97T (PP=0.744, $\omega=1.322\pm0.309$) и 100I (PP=0.858, $\omega=1.401\pm0.248$). Координаты сайтов указаны для последовательности первичной структуры белка, кодируемого геном *tgm* кластера PFNA *B. longum* subsp. *longum* GT15 (рис. 23). Оба сайта-кандидата локализованы в трансмембранном домене (англ.

transmembrane domain, TM) белка, кодируемого геном *tgm* кластера PFNA бифидобактерий. Аминокислотные остатки, формирующие вторичную структуру ТМ, как правило, находятся под давлением отрицательного отбора вследствие биофизических и функциональных ограничений аминокислотного состава трансмембранных α -спиралей. Однако в особых случаях трансмембранные регионы белков могут находиться под воздействием положительного отбора. В частности, трансмембранные регионы, участвующие в связывании с лигандами, могут демонстрировать быструю эволюцию, расширяя репертуар связываемых лигандов семейства белков [Spielman and Wilke, 2013]. Предполагаемая транскляминаза, кодируемая геном *tgm* кластера PFNA бифидобактерий, представляет собой политопический альфаспиральный интегральный мембранный белок, подобно многим рецепторам, ионным каналам и транспортерам. В первичной структуре данного белка различных видов бифидобактерий аннотирован N-концевой регион, содержащий до восьми сопряженных трансмембранных α -спиралей (рис. 24, Д), которые могут способствовать формированию третичной структуры, участвующей в связывании лигандов.



Рисунок 23. Локализация аминокислотных сайтов, находящихся под давлением продолжительного положительного отбора во всех ветвях филогенетического дерева бифидобактерий, в первичной структуре белка WP_038426324.1, кодируемого геном *tgm* кластера PFNA в геноме *B. longum* subsp. *longum* GT15. На рисунке представлена доменная организация белка, а также локализация сайтов-кандидатов со значением $PP > 0,7$.

3.3.3 Тестирование гипотез о наличии событий эпизодического положительного отбора в молекулярной эволюции последовательностей генов кластера PFNA.

В дополнении к обнаруженным событиям продолжительного положительного отбора, были протестированы гипотезы о наличии независимых событий эпизодического положительного отбора в молекулярной эволюции последовательностей генов кластера PFNA различных видов бифидобактерий. Эпизодический отбор, воздействующий на отдельные сайты в индивидуальных ветвях и кладах филогенетического дерева, представляет

собой наиболее распространенный случай положительного отбора.

Низкие значения длин тестируемых ветвей приводят к значительному ослаблению статистической мощности теста и неспособности детектировать положительный отбор. Короткие ветви были протестированы в составе следующих клад филогенетического дерева, включающих в себя группу близкородственных видов или подвидов бифидобактерий: *B. angulatum*, *B. merycicum*; *B. animalis* subsp. *animalis*, *B. animalis* subsp. *lactis*; *B. catenulatum*, *B. kashiwanohense*, *B. pseudocatenulatum*; *B. coryneforme*, *B. indicum*; *B. longum* subsp. *infantis*, *B. longum* subsp. *longum*; *B. thermacidophilum* subsp. *porcinum*, *B. thermacidophilum* subsp. *thermacidophilum*, *B. thermophilum*. Таким образом, для последующего анализа были отобраны 35 индивидуальных ветвей и клад филогенетического дерева бифидобактерий (табл. 17).

Были получены значения функций максимального правдоподобия двух «branch-site» моделей для тестируемых ветвей и клад, и затем проведены тесты LRT в строгих и нестрогих условиях. Нестрогие условия тестирования (табл. 7) применялись с целью уменьшить вероятность ошибки второго рода, т. е. вероятность ложноотрицательного результата при поиске редких событий, так как в строгих условиях тест достаточно консервативен [Zhang, Nielsen and Yang, 2005]. Вначале был проведен т. н. тест 1, в процессе которого тестируется предположение о наличии в последовательности сайтов, находящихся под давлением положительного отбора ($\omega > 1$)/ под ослабленным давлением отрицательного отбора в тестируемой ветви/ наборе ветвей по сравнению с остальными ветвями филогенетического дерева (M1a в качестве нулевой модели против A в качестве альтернативной) [Zhang, Nielsen and Yang, 2005]. Значения LRT для ряда тестируемых ветвей и клад филогенетического дерева оказались статистически значимыми, в частности, были обнаружены события отбора в 20 тестируемых ветвях в строгих условиях и в 27 тестируемых ветвях в нестрогих условиях (табл. 17).

Так как разрешающая способность теста 1 недостаточна для того, чтобы отличить детектируемые события положительного отбора от событий ослабления отрицательного отбора, был проведен т. н. тест 2, рекомендованный авторами в качестве прямого теста на положительный отбор [Zhang, Nielsen and Yang, 2005]. Для ветвей и клад филогенетического дерева, прошедших тест 1, было независимо протестировано предположение о наличии в последовательностях сайтов, находящихся под давлением положительного отбора ($\omega > 1$) в тестируемой ветви/ наборе ветвей по сравнению с остальными ветвями филогенетического дерева (A1 в качестве нулевой модели против A в качестве альтернативной) [Zhang, Nielsen and Yang, 2005]. Были обнаружены события положительного отбора в 15 из 20 тестируемых ветвей в строгих условиях и в 26 из 27 тестируемых ветвей в нестрогих условиях (табл. 18).

Таблица 17. Значения тестов LRT на наличие эпизодического положительного отбора/ ослабления отрицательного отбора в последовательностях генов кластера PFNA в различных ветвях филогенетического дерева бифидобактерий в строгих и нестрогих условиях. Указаны соответствующие Р-значения; тесты LRT со значениями $P < \alpha$ являются статистически значимыми.

тестируемая ветвь/ клада филогенетического дерева	значение lnL для теста M1a	значение lnL для теста H1	значение теста LRT	Р (строгие условия, $\alpha = 2,86E-004$)	Р (нестрогие условия, $\alpha = 0,05$)
<i>B. actinocoloniiforme</i>	-212064,17	-211851,62	425,10	0	0
<i>B. adolescentis</i>	-212064,17	-212064,17	0	1	1
<i>B. aesculapii</i>	-212064,17	-212052,63	23,08	9,74E-006	8,5225E-006
<i>B. angulatum</i> , <i>B. merycicum</i>	-212064,17	-211963,15	202,03	0	0
<i>B. animalis</i> subsp. <i>animalis</i> , <i>B. animalis</i> subsp. <i>lactis</i>	-212064,17	-212000,55	127,23	0	0
<i>B. asteroides</i>	-212064,17	-211985,51	157,31	0	0
<i>B. biavatii</i>	-212064,17	-211976,88	174,58	0	0
<i>B. bifidum</i>	-212064,17	-212047,66	33,01	7,00E-008	6,805556E-008
<i>B. bohemicum</i>	-212064,17	-212044,93	38,48	0	0
<i>B. bombi</i>	-212064,17	-212027,85	72,63	0	0
<i>B. boum</i>	-212064,17	-212060,78	6,78	3,37E-002	2,27E-002
<i>B. breve</i>	-212064,17	-212050,48	27,38	1,13E-006	1,040789E-006
<i>B. catenulatum</i> , <i>B. kashiwanohense</i> , <i>B. pseudocatenulatum</i>	-212064,17	-212060,94	6,46	3,95E-002	2,56E-002
<i>B. choerinum</i>	-212064,17	-212042,51	43,32	0	0
<i>B. commune</i>	-212064,17	-212047,15	34,03	4,00E-008	4,117647E-008

<i>B. coryneforme</i> , <i>B. indicum</i>	-212064,17	-212022,76	82,82	0	0
<i>B. cuniculi</i>	-212064,17	-212007,55	113,23	0	0
<i>B. dentium</i>	-212064,17	-212064,17	0	1	1
<i>B. eulemuris</i>	-212064,17	-212015,17	98,00	0	0
<i>B. gallicum</i>	-212064,17	-212041,48	45,37	0	0
<i>B. hapali</i>	-212064,17	-211948,99	230,36	0	0
<i>B. lemurum</i>	-212064,17	-212016,75	94,84	0	0
<i>B. longum</i> subsp. <i>infantis</i> , <i>B. longum</i> subsp. <i>longum</i>	-212064,17	-212064,17	0	1	1
<i>B. moukalabense</i>	-212064,17	-212064,17	0	1	1
<i>B. myosotis</i>	-212064,17	-212058,82	10,69	4,76E-003	3,63E-003
<i>B. pseudolongum</i> subsp. <i>globosum</i>	-212064,17	-211994,42	139,50	0	0
<i>B. pseudolongum</i> subsp. <i>pseudolongum</i>	-212064,17	-212059,50	9,33	9,43E-003	6,88E-003
<i>B. reuteri</i>	-212064,17	-212062,75	2,84	2,42E-001	1,40E-001
<i>B. ruminantium</i>	-212064,17	-212059,56	9,22	9,94E-003	6,96E-003
<i>B. saguini</i>	-212064,17	-212064,17	0	1	1
<i>B. scardovii</i>	-212064,17	-211991,13	146,07	0	0
<i>B. stellenboschense</i>	-212064,17	-212056,48	15,37	4,59E-004	3,83E-004
<i>B. thermacidophilum</i> subsp. <i>porcinum</i> , <i>B. thermacidophilum</i> subsp. <i>thermacidophilum</i> , <i>B. thermophilum</i>	-212064,17	-212062,36	3,62	1,64E-001	9,90E-002
<i>B. tissieri</i>	-212064,17	-212058,77	10,79	4,53E-003	3,60E-003
<i>B. vansinderenii</i>	-212064,17	-212061,75	4,84	8,91E-002	5,57E-002

Таблица 18. Значения тестов LRT на наличие эпизодического положительного отбора в последовательностях генов кластера PFNA в различных ветвях филогенетического дерева бифидобактерий в строгих и нестрогих условиях. Указаны соответствующие Р-значения; тесты LRT со значениями $P < \alpha$ являются статистически значимыми. Указаны значения ω тестируемой ветви для группы сайтов со значением $\omega > 1$.

тестируемая ветвь филогенетического дерева	значение lnL для теста	значение lnL для теста H1	значение теста LRT	P (строгие условия, $\alpha = 5,00E-004$)	P (нестрогие условия, $\alpha = 5,00E-002$)	значение ω тестируемой ветви для группы сайтов
<i>B. actinocoloniiforme</i>	-211910,16	-211851,62	117,09	0	0	4,79
<i>B. aesculapii</i>	-212063,36	-212052,63	21,46	3,62E-006	3,76E-006	23,71
<i>B. angulatum</i> , <i>B. merycicum</i>	-211965,13	-211963,15	3,97	4,64E-002	2,51E-002	1,46
<i>B. animalis</i> subsp. <i>animalis</i> , <i>B. animalis</i> subsp. <i>lactis</i>	-212004,17	-212000,55	7,24	7,14E-003	4,19E-003	1,60
<i>B. asteroides</i>	-212022,65	-211985,51	74,27	0	0	7,22
<i>B. biavatii</i>	-212011,54	-211976,88	69,31	0	0	5,63
<i>B. bifidum</i>	-212062,27	-212047,66	29,22	6,00E-008	7,36E-008	21,64
<i>B. bohemicum</i>	-212059,73	-212044,93	29,61	5,00E-007	6,75E-008	11,60
<i>B. bombi</i>	-212047,48	-212027,85	39,26	0	0	6,89
<i>B. boum</i>	-212064,16	-212060,78	6,77	не тестировали в строгих условиях	5,22E-003	12,52
<i>B. breve</i>	-212050,48	-212050,48	0	1	1	
<i>B. catenulatum</i> , <i>B. kashiwanohense</i> , <i>B. pseudocatenulatum</i>	-212062,61	-212060,94	3,36	не тестировали в строгих условиях	3,48E-002	
<i>B. choerinum</i>	-212061,10	-212042,51	37,19	0	0	25,81
<i>B. commune</i>	-212057,18	-212047,15	20,04	7,57E-006	6,81E-006	7,48
<i>B. coryneforme</i> , <i>B. indicum</i>	-212037,27	-212022,76	29,01	7,00E-008	7,875E-008	4,51

<i>B. cuniculi</i>	-212053,51	-212007,55	91,91	0	0	19,09
<i>B. eulemuris</i>	-212020,79	-212015,17	11,24	7,99E-004	6,34E-004	3,57
<i>B. gallicum</i>	-212052,12	-212041,48	21,28	3,97E-006	3,83E-006	4,99
<i>B. hapali</i>	-211966,66	-211948,99	35,35	0	0	2,87
<i>B. lemurum</i>	-212021,01	-212016,75	8,53	3,50E-003	2,25E-003	3,64
<i>B. myosotis</i>	-212063,87	-212058,82	10,09	не тестировали в строгих условиях	1,06E-003	23,68
<i>B. pseudolongum subsp. globosum</i>	-212050,38	-211994,42	111,93	0	0	16,46
<i>B. pseudolongum subsp. pseudolongum</i>	-212064,06	-212059,50	9,12	не тестировали в строгих условиях	1,71E-003	22,29
<i>B. ruminantium</i>	-212063,64	-212059,56	8,17	не тестировали в строгих условиях	2,62E-003	8,64
<i>B. scardovii</i>	-212029,70	-211991,13	77,14	0	0	8,27
<i>B. stellenboschense</i>	-212064,13	-212056,48	15,30	не тестировали в строгих условиях	7,74E-005	44,86
<i>B. tissieri</i>	-212064,06	-212058,77	10,57	не тестировали в строгих условиях	8,62E-004	14,09

Для подтверждения полученных результатов мы применили дополнительный подход [Murrel *et al.*, 2012]. Мы обнаружили события эпизодического положительного отбора в молекулярной эволюции генов кластера PFNA во всех тестируемых ветвях и кладах филогенетического дерева *Bifidobacterium*. Таким образом, полученные нами результаты являются надежными и воспроизводимыми.

Среди ветвей, для которых значение теста 2 оказалось статистически значимым в строгих условиях, крайне высокое значение ω тестируемой ветви для группы сайтов со значением $\omega > 1$ наблюдается для ветви *B. choerinum* ($\omega = 25,81$), а также *B. aesculapii* ($\omega = 23,71$), *B. bifidum* ($\omega = 21,64$), *B. cuniculi* ($\omega = 19,09$), *B. pseudolongum* subsp. *globosum* ($\omega = 16,46$), *B. bohemicum* ($\omega = 11,60$) и некоторых других ветвей. Вышеперечисленные ветви имеют сравнительно невысокую длину, поэтому статистическая значимость полученных результатов не может быть следствием ошибки первого рода в результате насыщения последовательностей синонимичными заменами (d_s) и переоценки отношения d_N/d_S . Данная проблема хорошо известна для ветвей с высоким значениям длин и часто приводит к получению ложноположительных результатов [Cannarozzi and Schneider, 2012].

Значения LRT для следующих ветвей и клад оказались статистически незначимыми даже при тестировании в нестрогих условиях (табл. 18): *B. adolescentis*; *B. catenulatum*, *B. kashiwanohense*, *B. pseudocatenulatum*; *B. dentium*; *B. longum* subsp. *infantis*, *B. longum* subsp. *longum*; *B. moukalabense*; *B. reuteri*; *B. saguini*; *B. thermacidophilum* subsp. *porcinum*, *B. thermacidophilum* subsp. *thermacidophilum*, *B. thermophilum*; *B. vansinderenii*, однако это может быть объяснено снижением статистической мощности тестов за счет сравнительно невысоких значений длин ветвей в отдельных случаях или неизбежным увеличением значения ошибки второго рода за счет использования поправки на множественное сравнение.

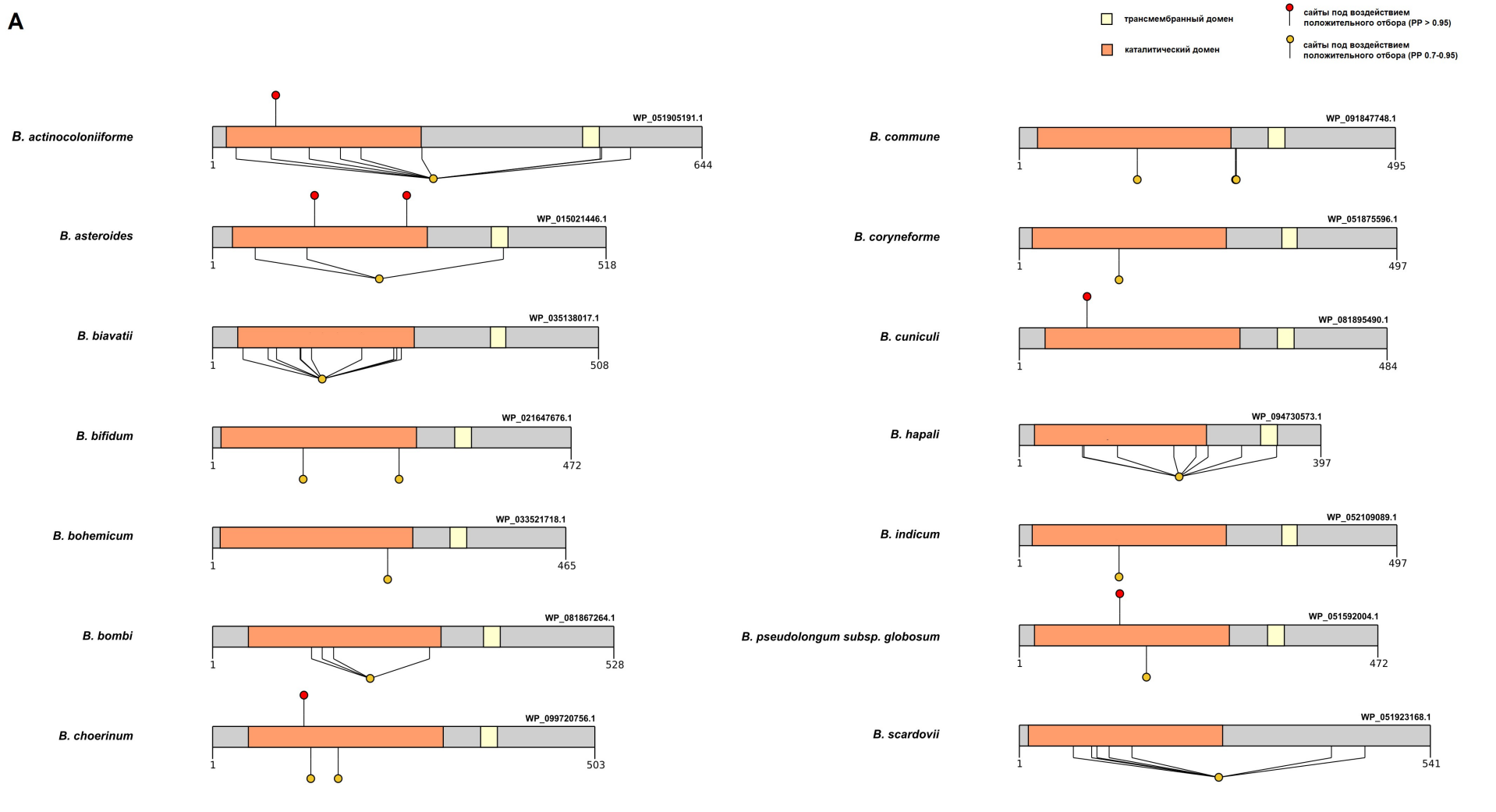
Отдельный интерес представляют ветви филогенетического дерева, для которых значение теста 1 оказалось статистически значимым, в то время как значение теста 2 — статистически незначимым. Это указывает на вероятные события ослабления воздействия отрицательного отбора в молекулярной эволюции последовательностей. В частности, значение LRT теста 1 для тестируемой ветви *B. breve* оказалось статистически значимым в строгих условиях (табл. 17), в то время как значение теста 2 оказалось статистически незначимым даже в нестрогих условиях (табл. 18). Рассчитанное значение $\omega = 1$ для класса сайтов $\omega \geq 1$ также указывает на ослабление давления отрицательного отбора в пользу нейтральной эволюции. Кластер PFNA *B. breve* содержит гибридный ген, полученный в результате объединения генов *fn3* и *aaa-atp*. Данное событие является уникальным для рода *Bifidobacterium* и не обнаружено в геномах представителей других видов. Эволюционное событие объединения генов *fn3* и *aaa-atp* *B. breve* является сравнительно недавним, так как

даже наиболее близкородственные виды бифидобактерий в составе филогруппы *B. longum* содержат в геномах последовательности индивидуальных генов *fn3* и *aaa-atp*. Хорошо известно, что молодые белки имеют тенденцию к ослаблению отрицательного отбора, что обеспечивает им быструю эволюцию по сравнению с эволюционно более древними белками [Domazet-Loso and Tautz, 2003; Daubin and Ochman, 2004; Alba and Castresana 2005; Alba and Castresana, 2007; Garcia-Vallve *et al.*, 2005; Wolf *et al.*, 2009].

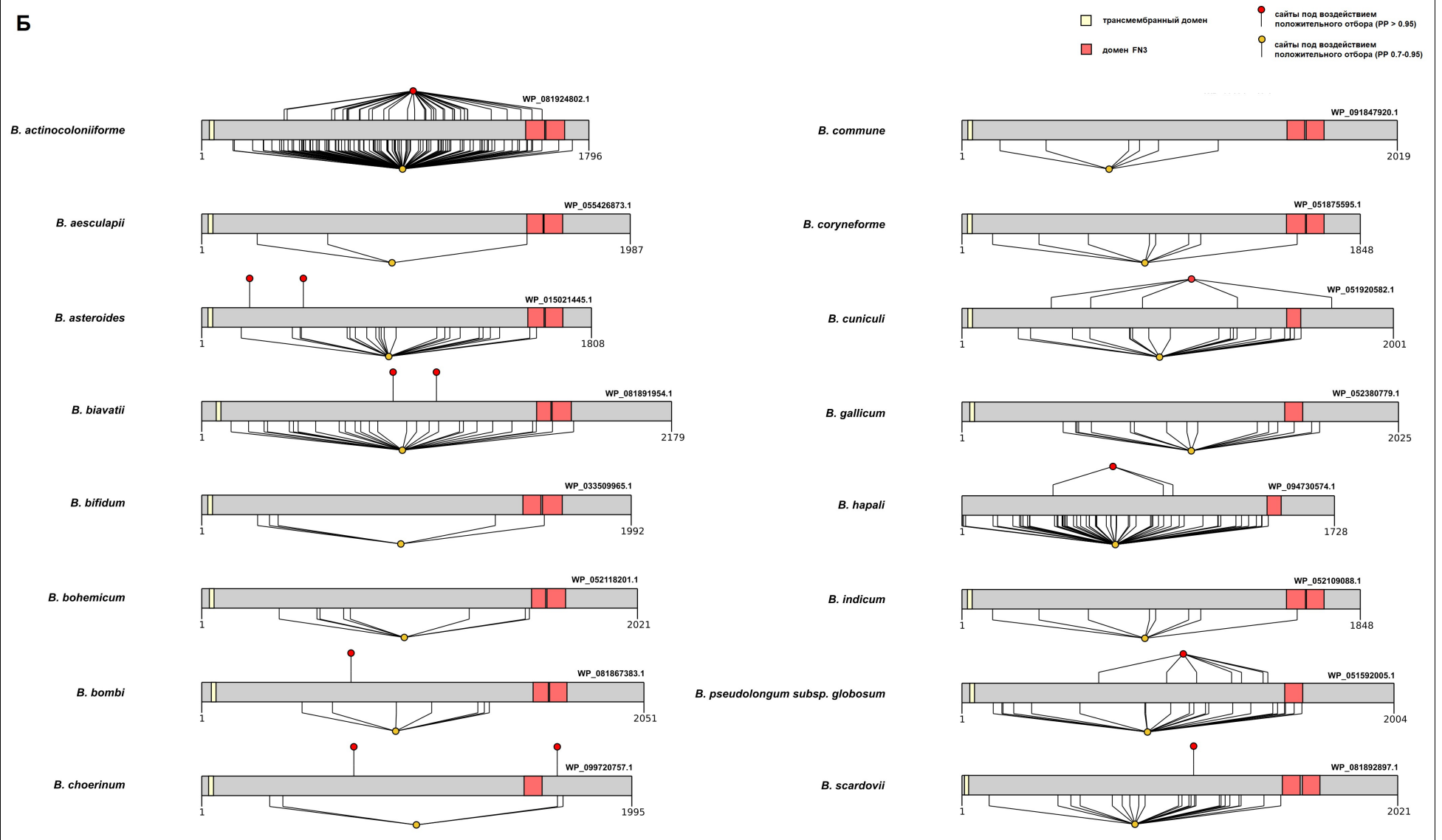
Таким образом, экспериментально *in silico* показано наличие независимых событий положительного отбора в молекулярной эволюции генов кластера PFNA для большинства ветвей филогенетического дерева бифидобактерий (26 из 35 тестируемых ветвей в нестрогих условиях), что объясняет высокую степень межвидовой дивергенции последовательностей. Проведенные тесты предполагают наличие групп сайтов со значением $\omega > 1$ в тестируемых ветвях по сравнению с остальными ветвями филогенетического дерева *Bifidobacterium*, следовательно, положительный отбор в молекулярной эволюции последовательностей генов кластера PFNA проявлял видоспецифический характер, воздействуя на группы сайтов в последовательностях, соответствующих отдельным ветвям филогении независимо от других ветвей.

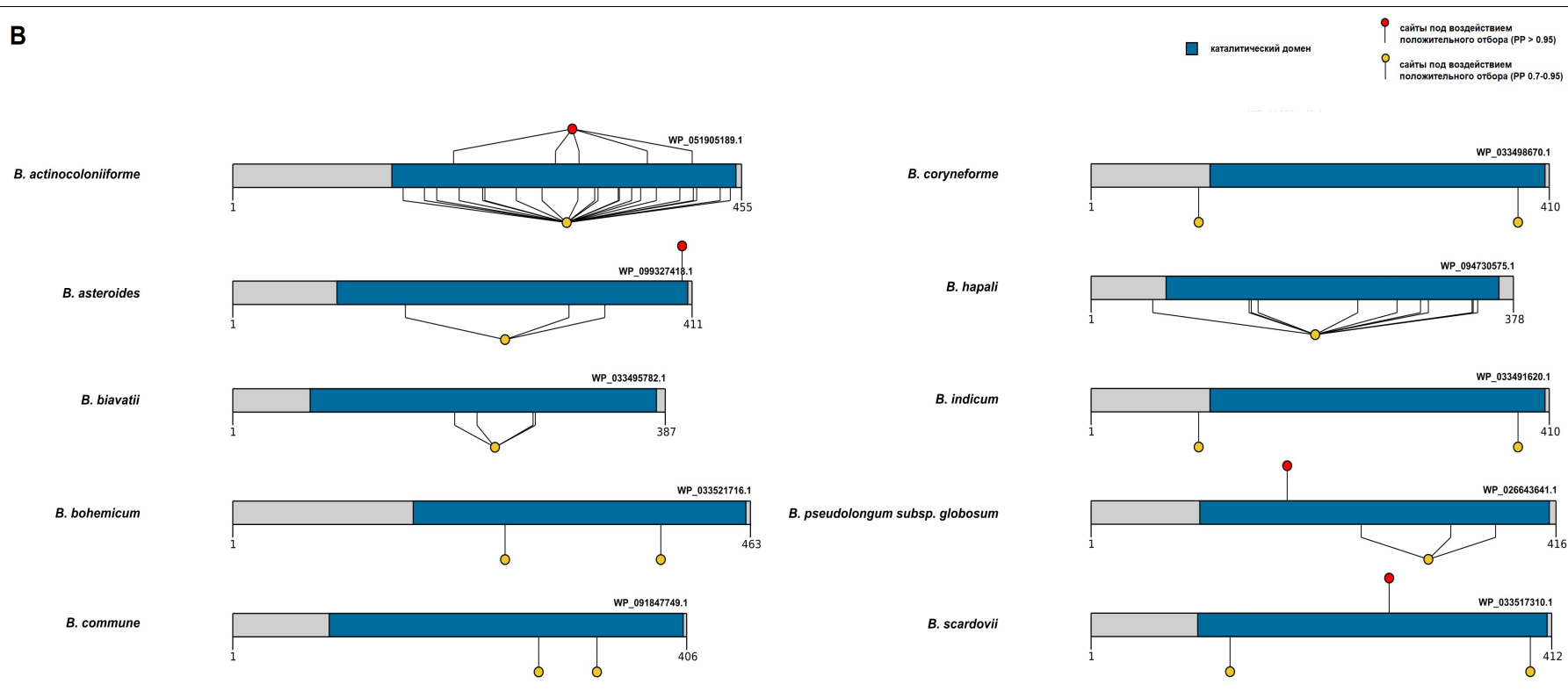
Сайты-кандидаты для воздействия эпизодического положительного отбора в молекулярной эволюции последовательностей белок-кодирующих участков генов кластера PFNA были затем идентифицированы с помощью эмпирического Байесовского подхода ($PP > 0,7$) (табл. 3 Приложений) и локализованы в первичной структуре белков, кодируемых генами *pkb2* (рис. 24, А), *fn3* (рис. 24, Б), *aaa-atp* (рис. 24, В), *duf58* (рис. 24, Г) и *tgm* (рис. 24, Д) представителей различных видов рода *Bifidobacterium*.

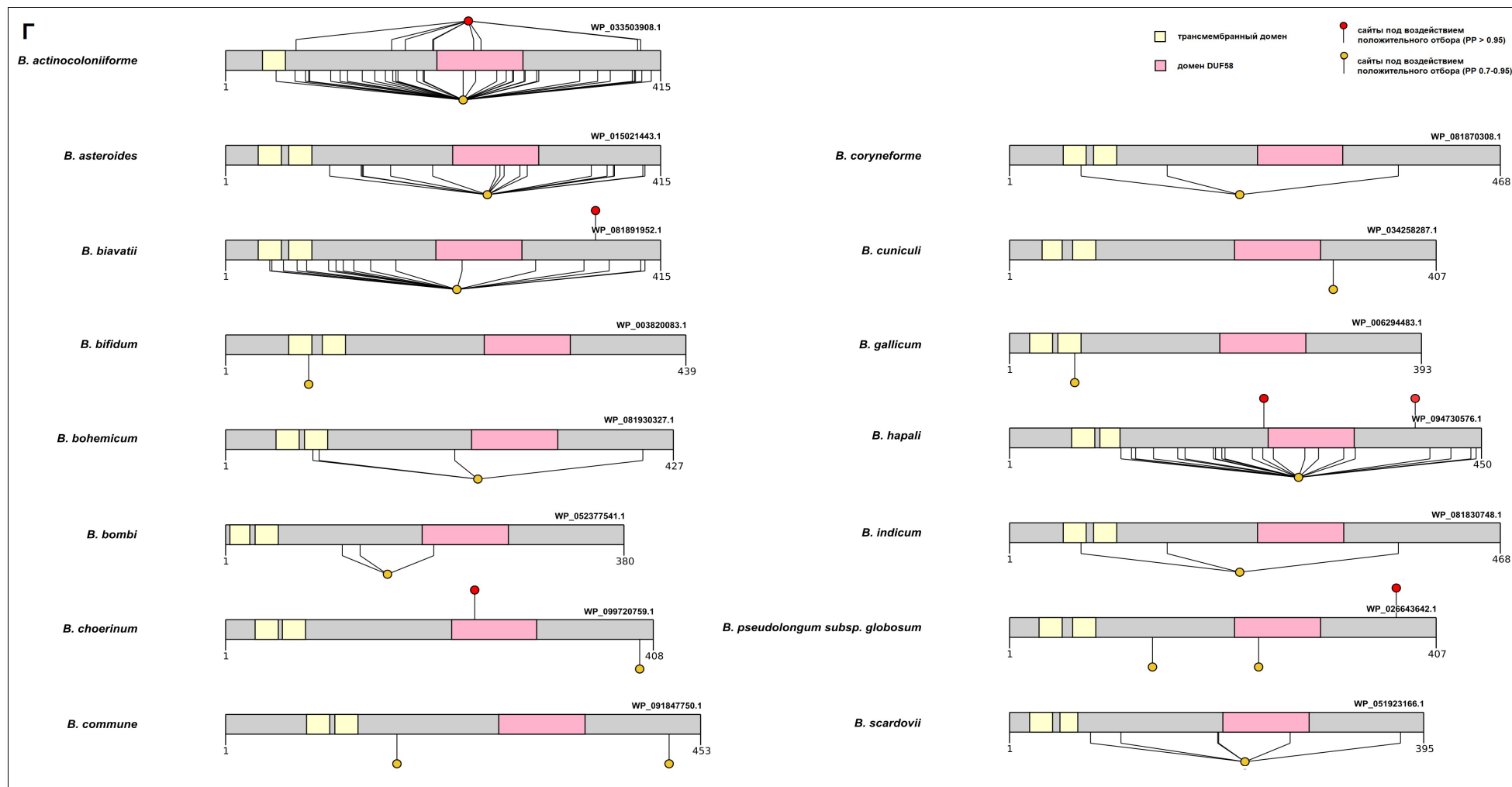
A



Б



B



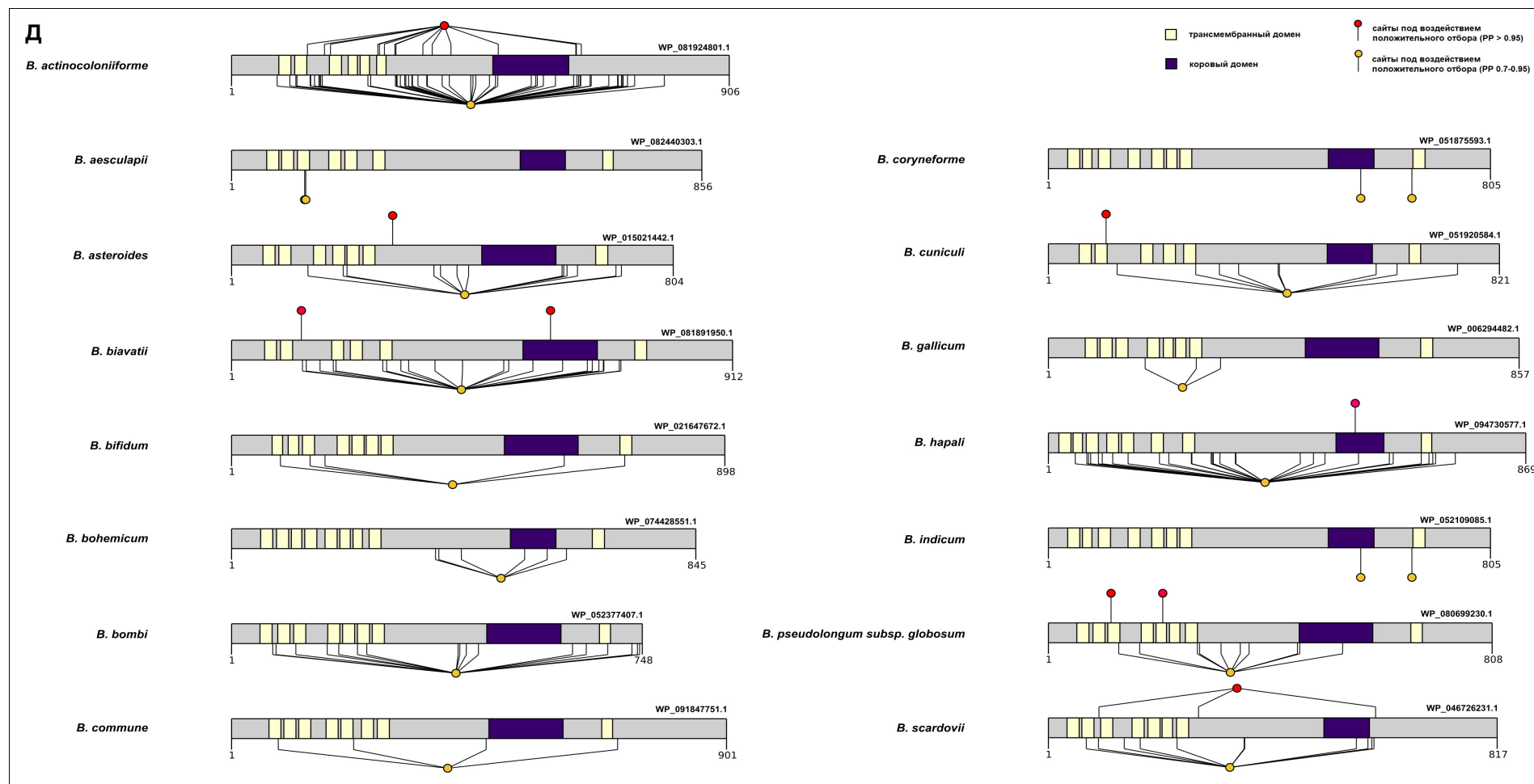


Рисунок 24. Локализация сайтов под давлением эпизодического положительного отбора в первичной структуре белков, кодируемых генами *pkb2* (А), *fn3* (Б), *aaa-atp* (В), *duf58* (Г) и *tgm* (Д) кластера PFNA представителей различных видов бифидобактерий. На рисунке представлена доменная организация белков, а также локализация сайтов-кандидатов со значением PP = 0,7-0,95 и PP > 0,95.

Обнаруженные сайты-кандидаты локализованы в различных участках исследуемых белков, включая аннотированные функциональные домены и трансмембранные регионы, а также регионы с отсутствием аннотированных функциональных участков. Сайты со значениями $PP > 0,95$ являются наиболее надежными кандидатами для воздействия положительного отбора.

Такие сайты были обнаружены в каталитическом домене СТПК, кодируемой геном *pkb2*, и АТФазном домене АТФазы, кодируемой геном *aaa-atp* у представителей различных видов бифидобактерий. В белках, кодируемых генами *fn3* и *duf58* кластера PFNA, такие сайты были обнаружены как в аннотированных доменах за исключением ТМ, так и вне аннотированных доменов. В предполагаемой транслютаминазе, кодируемой геном *tgm*, сайты со значением $PP > 0,95$ были обнаружены на всем протяжении первичной структуры белка, включая трансмембранные домены (рис. 24, А-Д).

Эпизодический положительный отбор в функциональных доменах вышеупомянутых белков может обеспечивать быструю эволюцию в ответ на воздействие внешних факторов. Несмотря на то, что гипотеза Черной Королевы обычно описывает эффекты бинарных антагонистических взаимодействий, необходимо учитывать также контекст микробного сообщества [Brockhurst *et al.*, 2014]. Иммунная система организма хозяина нацелена на элиминацию патогенов, однако в ходе воспалительного процесса полезные микроорганизмы также могут подвергаться воздействию факторов иммунного ответа. Таким образом, комменсальные представители микробиоты, подобно патогенным, находятся под давлением иммунной системы хозяина. Характер молекулярной эволюции генов, участвующих в реализации механизмов устойчивости комменсалов, может быть сходен с таковым у патогенных микроорганизмов. Мы предполагаем, что адаптивные изменения в генах иммунной системы хозяина в ходе «гонки вооружений» могут оказывать влияние не только на патогенов, но и на комменсальных представителей микробиоты, которые в результате данного процесса также могут становиться участниками эволюционной гонки. В частности, предполагаемое взаимодействие с цитокинами белка FN3 может служить объяснением феномену быстрой эволюции в ответ на быструю эволюцию генов цитокинов, обеспечиваемую воздействием положительного отбора [Zelus *et al.*, 2000; Neves *et al.*, 2014; Daub *et al.*, 2017]. Адгезионные свойства данного белка также могут обеспечивать быструю эволюцию в ответ на видоспецифическое изменение репертуара субстратов адгезии [Voolstra *et al.*, 2011]. В общем случае функцией белка FN3 является, по-видимому, взаимодействие с факторами внешней среды, представленной организмом хозяина. Показано, что белки,

обеспечивающие взаимодействие с факторами внешней среды, могут подвергаться воздействию положительного отбора [Voolstra *et al.*, 2011].

С-концевой сигнальный участок СТПК Pkb2, который отвечает за распознавание внешних стимулов, демонстрирует высокую вариабельность у различных видов (рис. 25). Различия в структуре данного участка даже у близкородственных видов, принадлежащих одной филогруппе, могут обуславливать видоспецифичность связывания лигандов, и, следовательно, сигналинга в целом.

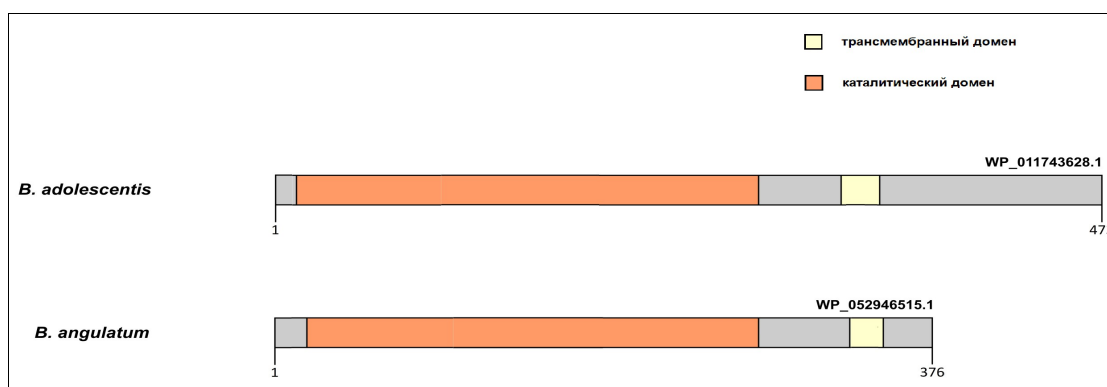


Рисунок 25. Различия в первичной структуре СТПК Pkb2 представителей видов *B. angulatum* и *B. adolescentis* филогруппы *B. adolescentis*. В структуре Pkb2 представлены каталитический киназный домен, трансмембранный домен (ТМ) и С-концевой сигнальный регион.

Как упоминалось ранее, сложное строение трансмембранного региона предполагаемой транслютаминазы, включающего в себя до восьми аннотированных трансмембранных α -спиралей, может обеспечивать способность связывания с лигандами. Эпизодический положительный отбор в аминокислотных сайтах трансмембранных доменов, наряду с обнаруженным продолжительным отбором, может обеспечивать быструю эволюцию лиганд-связывающих участков белка в ответ на видоспецифическое изменение репертуара лигандов. Таким образом, положительный отбор в молекулярной эволюции генов *pkb2* и *tgm*, мог участвовать в формировании механизмов видоспецифического сигналинга.

Быстрая эволюция по крайней мере одного из генов кластера PFNA под воздействием внешних факторов может провоцировать быструю эволюцию коэволюционирующих генов, что приводит к высокой степени межвидовой дивергенции последовательностей всех взаимодействующих генов. В частности, в процессе фосфорилирования СТПК Pkb2 и АТФаза AAA-АТР физически взаимодействуют друг с другом. Изменение структуры одного из белков в результате воздействия положительного отбора может привести к нарушению

белок-белкового интерфейса взаимодействующих белков, что, в свою очередь, провоцирует быстрые изменения в структуре второго белка для восстановления эффективности киназно-субстратного взаимодействия. Данный сценарий также может быть реализован в результате взаимодействия между белками FN3 и AAA-АТР, предсказанного нами *in silico*.

Таким образом, нами было установлено, что гены кластера PFNA, участвуя в реализации механизмов коммуникации с факторами внешней среды, демонстрируют эффекты, описываемые гипотезой Черной Королевы. Необходимость адаптироваться к новым условиям в результате специализации и смены экологической ниши также могла увеличивать скорость эволюции в анализируемых генах.

Сайты, на которые осуществляется давление отбора, могут быть структурно или функционально значимыми, и, таким образом, представляют собой вероятных кандидатов для расширения структурно-функциональной аннотации белков. Аннотация белков, кодируемых генами кластера PFNA, представляет собой важную задачу, так как структура и функции данных белков малоизучены.

3.4 Сравнительный анализ уровня экспрессии гена *fn3* у представителей различных видов рода *Bifidobacterium*.

Согласно литературным данным, гены, в молекулярной эволюции которых обнаружен положительный отбор, экспрессируются на более низком уровне, чем гены под воздействием отрицательного отбора или нейтральной эволюции [Kosiol *et al.*, 2008; Koester *et al.*, 2012]. Например, в геномах *Saccharomyces cerevisiae* и *Thalassiosira pseudonana* уровень экспрессии генов, по-видимому, отрицательно коррелирует со скоростью эволюции: высокоэкспрессирующиеся гены эволюционируют более медленно и наоборот [Drummond *et al.*, 2005]. Это может быть объяснено тем фактом, что высокоэкспрессирующиеся в нормальных условиях гены, как правило, являются наиболее функционально важными и, следовательно, находятся под сильным давлением отрицательного отбора. Накопление несинонимичных замен может приводить к ошибкам фолдинга и формирования нативной структуры белков, что является критическим фактором, обуславливающим воздействие отрицательного отбора на коровые гены [Drummond *et al.*, 2005].

В данной диссертационной работе мы проверили гипотезу о наличии зависимости между уровнем экспрессии гена *fn3* и силой отбора в геномах различных видов бифидобактерий. Нами был проведен количественный анализ экспрессии генов с помощью метода ПЦР в реальном времени. Анализ транскрипционной активности включал в себя

построение кривых роста культивируемых бифидобактерий, выделение и очистку тотальной РНК, получение кДНК в ходе реакции обратной транскрипции, амплификацию фрагмента гена *fn3* представителей различных видов бифидобактерий в ходе реакции ПЦР в реальном времени с использованием видоспецифических праймеров и количественный анализ результатов. Анализ транскрипционной активности был проведен для следующих штаммов бифидобактерий из лабораторной коллекции: *B. animalis* subsp. *animalis* ATCC 25527, *B. choerinum* DSM 20434, *B. pseudolongum* subsp. *globosum* DSM 20092, *B. angulatum* GT102, *B. breve* DSM 20213, *B. longum* subsp. *longum* GT15, *B. thermophilum* DSM 20210, *B. ruminantium* AC1786, *B. dentium* Bifido 9. Для корректного количественного анализа результатов было определено значение эффективности амплификации в каждой из исследуемых проб.

В результате анализа транскрипционной активности мы показали, что ген *fn3* транскрибируется в геномах всех исследуемых штаммов. Сравнительный анализ начальных значений сигнала флуоресценции показал, что уровень экспрессии гена *fn3* широко варьирует в различных пробах и не зависит от силы и характера естественного отбора (рис. 26).

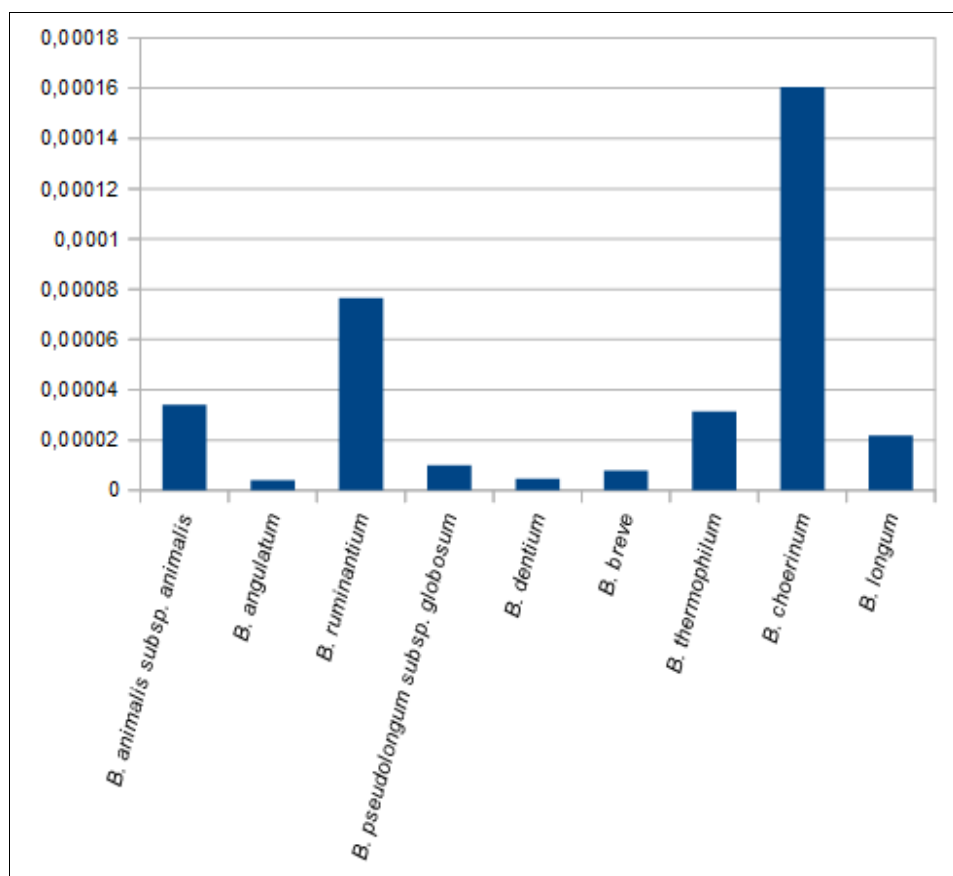


Рисунок 26. Сравнение начальных значений сигнала флуоресценции в пробах кДНК.

По оси Y отложены значения начальных значений сигнала флуоресценции в единицах флуоресценции (англ. relative fluorescence units, RFU).

Мы показали, что наиболее высокий уровень экспрессии гена *fn3* демонстрирует штамм *B. choerinum* DSM 20434, для которого было получено наибольшее значение ω тестируемой ветви для группы сайтов со значением $\omega > 1$ в строгих условиях тестирования ($\omega = 25,81$), что является, по-видимому, исключением из общего правила обнаруживаемой зависимости между уровнем экспрессии генов и силой отбора. Эпизодический положительный отбор в молекулярной эволюции генов кластера PFNA *B. choerinum* мог быть направлен на увеличение функциональной активности кодируемых белков, что, в свою очередь, могло оказывать влияние на уровень экспрессии и приводить к увеличению уровня транскрипционной активности анализируемых генов.

3.5 Изучение видоспецифичности киназно-субстратного взаимодействия в реакции перекрестного фосфорилирования *in vitro*.

Положительный отбор и быстрая эволюция вносят изменения в структуру ортологичных белков различных видов бифидобактерий, что может приводить к формированию видоспецифичности белок-белкового взаимодействия коэволюционирующих функциональных партнеров.

Как было отмечено, для СТПК Pkb2 ранее были определены субстраты фосфорилирования [Nezametdinova *et al.*, 2018], среди которых оказалась АТФаза AAA-АТР. Таким образом, экспериментально было показано взаимодействие между продуктами генов *pkb2* и *aaa-atp* кластера PFNA. В данной диссертационной работе была изучена видоспецифичность киназно-субстратного взаимодействия на примере штаммов *B. longum* subsp. *longum* GT15 и *B. choerinum* DSM 20434 в ходе киназной реакции с (γ -32P)-АТР в качестве маркера.

Экспериментальный анализ включал в себя следующие этапы: клонирование полноразмерных генов *pkb2* и *aaa-atp* *B. longum* subsp. *longum* GT15 и *B. choerinum* DSM 20434 в штамм *E. coli* DH5 α (рис. 12, А), выделение и очистку полноразмерных белков Pkb2 и AAA-АТР из экспрессионного штамма *E. coli* BL21(DE3) методом металоаффинной хроматографии в нативных условиях (рис. 12, Б), коинкубацию СТПК Pkb2 и АТФазы AAA-АТР с (γ -32P)-АТР и выявление фосфорилированных продуктов киназной реакции методом автордиографии (рис. 12, В).

Нами была отработана методика выделения СТПК Pkb2 и АТФазы AAA-АТР *B. choerinum* DSM 20434 в нативных условиях (рис. 27).

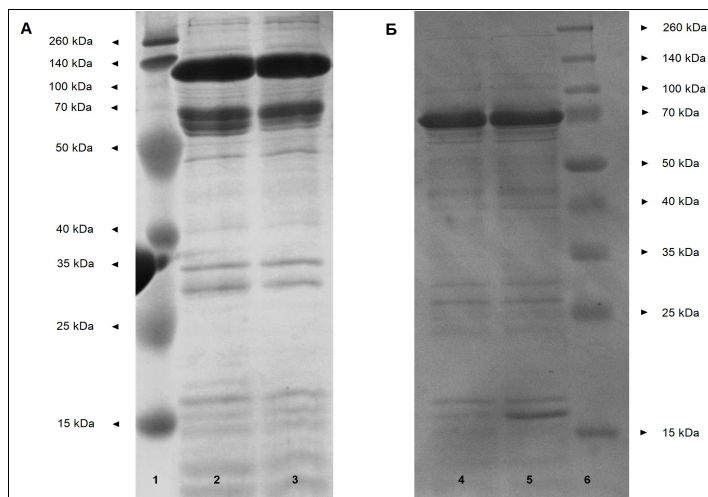


Рисунок 27. Электрофореграмма для визуализации результатов выделения белков Pkb2 (A) и AAA-ATP (Б) *B. choerinum* DSM 20434 из экспрессионного штамма BL21(DE3) *E. coli* в нативных условиях. Дорожки: 1, 6 – маркер длин белкового продукта; 2, 3 – Pkb2 *B. choerinum* DSM 20434; 4, 5 – AAA-ATP *B. choerinum* DSM 20434.

Способность к автофосфорилированию каталитического домена СТПК Pkb2 *B. longum* subsp. *longum* была показана ранее в штамме B379M [Nezametdinova *et al.*, 2014]. Нами была продемонстрирована способность к автофосфорилированию полноразмерных СТПК Pkb2 *B. longum* subsp. *longum* GT15 и *B. choerinum* DSM 20434 (рис. 29, дорожки 7, 8).

Для оценки способности к фосфорилированию субстратов была проведена киназная реакция для реакционных смесей, содержащих СТПК Pkb2 и АТФазу AAA-АТР из одного вида бифидобактерий. Была показана способность к фосфорилированию субстрата в обоих исследуемых видах (рис. 29, дорожки 1, 3).

Для оценки способности к перекрестному фосфорилированию субстратов была проведена киназная реакция для реакционных смесей, содержащих СТПК Pkb2 и АТФазу AAA-АТР из разных видов бифидобактерий. В реакции перекрестного фосфорилирования СТПК Pkb2 *B. longum* subsp. *longum* GT15 и АТФазы AAA-АТР *B. choerinum* DSM 20434 способность киназы к фосфорилированию субстрата из другого вида показана не была (рис. 29, дорожка 5), что указывает на видоспецифичность киназно-субстратного взаимодействия. Однако, в реакции перекрестного фосфорилирования СТПК Pkb2 *B. choerinum* DSM 20434 и АТФазы AAA-АТР *B. longum* subsp. *longum* GT15 нами была продемонстрирована способность киназы к фосфорилированию АТФазы из другого вида (рис. 29, дорожка 6), указывающая на отсутствие субстратной видоспецифичности. Это может быть связано с более высокой функциональной активностью СТПК Pkb2 *B. choerinum* DSM 20434 по

сравнению с СТПК Pkb2 *B. longum* subsp. *longum* GT15, о чем косвенно может свидетельствовать качественно более высокий уровень киназной активности в реакциях автофосфорилирования (рис. 29, дорожки 7, 8). Это согласуется с ранее полученными данными об уровне экспрессии гена *fn3* *B. choerinum* DSM 20434 и выводами о роли положительного отбора в молекулярной эволюции генов кластера PFNA *B. choerinum*.

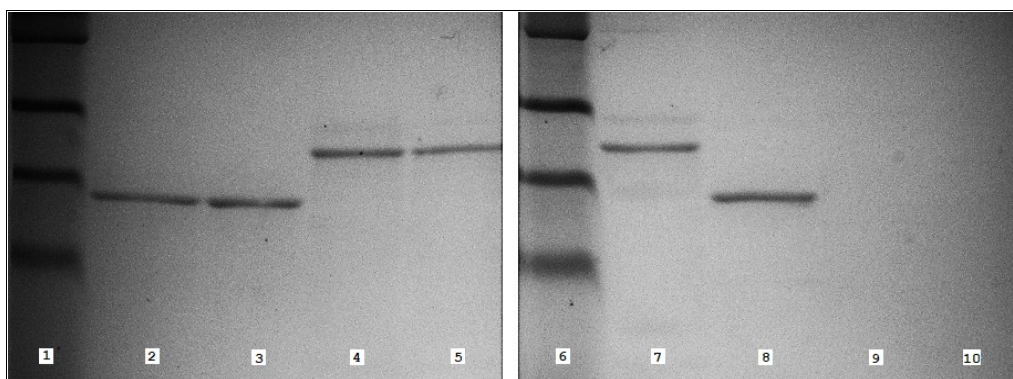


Рисунок 28. Электрофореграмма разделения компонентов киназной реакции. 1,6 – маркер молекулярного веса; 2 – Pkb2 GT15 + AAA-ATP GT15; 3 – AAA-ATP GT15; 4 – Pkb2 DSM 20434 + AAA-ATP DSM 20434; 5 – AAA-ATP DSM 20434; 7 – Pkb2 GT15 + AAA-ATP DSM 20434; 8 – Pkb2 DSM 20434 + AAA-ATP GT15; 9 – Pkb2 GT15; 10 – Pkb2 DSM 20434.

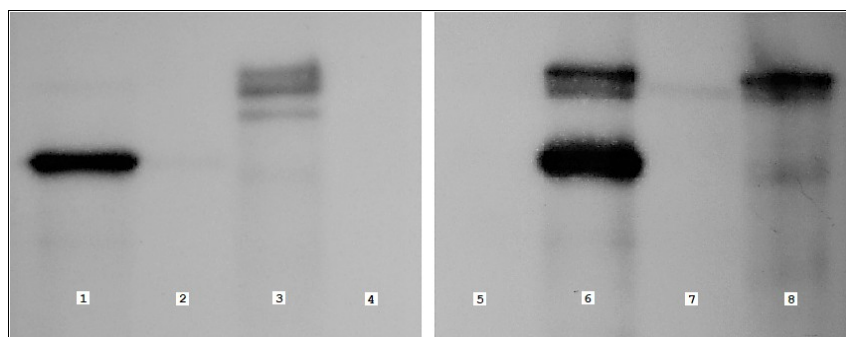


Рисунок 29. Выявление фосфорилированных продуктов киназной реакции методом автордиографии. 1 – Pkb2 GT15 + AAA-ATP GT15; 2 – AAA-ATP GT15; 3 – Pkb2 DSM 20434 + AAA-ATP DSM 20434; 4 – AAA-ATP DSM 20434; 5 – Pkb2 GT15 + AAA-ATP DSM 20434; 6 – Pkb2 DSM 20434 + AAA-ATP GT15; 7 – Pkb2 GT15; 8 – Pkb2 DSM 20434.

Таким образом, в результате анализа нами впервые экспериментально была показана функциональная активность полноразмерных СТПК Pkb2 представителей видов *B. longum* subsp. *longum* и *B. choerinum*, субстратная видоспецифичность для СТПК Pkb2 *B. longum* subsp. *longum* GT15 и ее отсутствие для СТПК Pkb2 *B. choerinum* DSM 20434.

3.6 Оценка влияния IL-6 как фактора иммунного ответа на рост культуры и экспрессию генов штамма *B. longum* subsp. *longum* GT15.

Как было отмечено ранее, быстрая эволюция и обнаруженный нами положительный отбор в последовательностях гена *fn3* бифидобактерий могут быть обусловлены участием в реализации механизмов коммуникации с иммунной системой организма хозяина в рамках гипотезы Черной Королевы. В структуре FNIII доменов белка, кодируемого геном *fn3* кластера PFNA, аннотированы мотивы, характерные для цитокин-связывающего региона рецептора gr-130. Рецептор gr-130 связывает IL-6 и некоторые другие цитокины данного семейства. Мы предполагаем, что ген *fn3* может быть медиатором взаимодействия между иммунной системой организма хозяина и бифидобактериальными клетками, опосредуя механизмы цитокин-зависимого сигналинга. Для проверки данной гипотезы мы провели оценку влияния IL-6 как фактора иммунного ответа на рост и экспрессию генов *B. longum* subsp. *longum* GT15. Как было отмечено ранее, добавление цитокинов к культурам клеток микроорганизмов может оказывать влияние на скорость роста культуры и другие характеристики (табл. 1). В частности экспериментально было показано, что патогенные микроорганизмы *Staphylococcus aureus*, *Acinetobacter* spp. и *Pseudomonas aeruginosa* демонстрируют увеличение роста в ответ на добавление в среду провоспалительных цитокинов TNF α , IL-1 β , and IL-6 в различных концентрациях в модели воспаления *in vitro* [Meduri *et al.*, 1999]. Таким образом, была показана чувствительность отдельных представителей патогенной микробиоты к факторам иммунного ответа, что указывает на возможность реализации сходных механизмов цитокиновой индукции у представителей комменсальной микробиоты.

Эксперимент по изучению влияния IL-6 на скорость роста и экспрессию генов *B. longum* subsp. *longum* GT15 включал следующие этапы: культивирование исследуемого штамма в нормальных условиях и в присутствии провоспалительного цитокина IL-6, оценка влияния на IL-6 скорость роста бактериальной культуры, выделение тотальной РНК, получение неконтаминированной мРНК, получение кДНК в ходе реакции обратной транскрипции, подготовка геномных библиотек и полногеномное секвенирование и затем поиск дифференциально экспрессирующихся генов (рис. 13).

В результате оценки влияния различных концентраций провоспалительного цитокина IL-6 на рост штамма *B. longum* subsp. *longum* GT15 мы показали, что добавление IL-6 в различных концентрациях не оказывает дозозависимый эффект на скорость роста бифидобактериальной культуры. При этом кривые роста штамма, культивируемого в

экспериментальных условиях, не демонстрировали значимых отличий от кривых роста штамма, культивируемого в контрольных условиях (рис. 30). На основании полученных результатов нами сделан вывод о том, что присутствие в среде IL-6 в различных концентрациях не влияет на скорость роста культуры *B. longum* subsp. *longum* GT15.

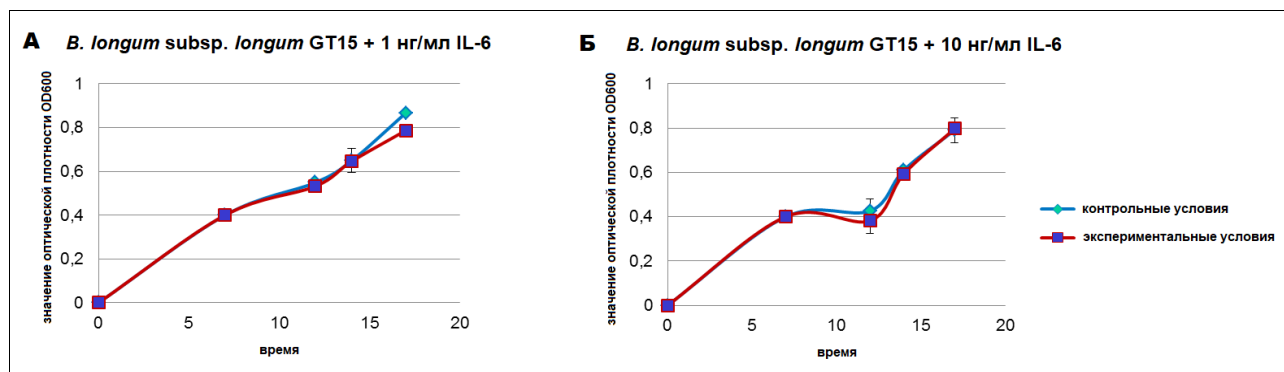


Рисунок 30. Скорость роста культуры штамма *B. longum* subsp. *longum* GT15 в присутствии провоспалительного цитокина IL-6 в концентрации 1нг/мл (А) и 10нг/мл (Б) по сравнению с культивированием в нормальных условиях. По оси Y отложены значения оптической плотности (OD).

Нами была отработана методика RNA-Seq для грамположительных бактерий с высоким GC-составом геномных последовательностей. Проведенный нами транскриптомный анализ позволил оценить уровень экспрессии генов в экспериментальных условиях при добавлении провоспалительного цитокина IL-6 в концентрации 1нг/мл по сравнению с культивированием в контрольных условиях и охарактеризовать обнаруженные дифференциально экспрессирующиеся гены (табл. 19).

Таблица 19. Дифференциально экспрессирующиеся гены в геноме *B. longum* subsp. *longum* GT15 в модели воздействия факторов иммунного ответа, представленных провоспалительным цитокином IL-6, по сравнению с нормальными условиями.

наименование локуса	продукт гена и аннотация	log2FoldChange	P ($\alpha=0,05$)
BLGT_00140	субъединица UgpC импортера UgpABCE, отвечающего за транспорт глицерол-3-фосфата (англ. glycerol 3-phosphate, G3P); G3P используется клеткой в качестве дополнительного источника фосфатов	0,86	5,28E-04
BLGT_00625	молекулярный шаперон Hsp20, супрессирующий агрегацию белка в условиях стрессового ответа	0,75	3,63E-02
BLGT_01150	продукт содержит домен, характерный для белков, которые осуществляют деградацию гликопротеинов	1,02	1,14E-04

	мукоидного слоя		
BLGT_04920	α -субъединица глутаматсинтазы	-0,66	3,24E-02
BLGT_06345	ассоциированный с рибосомой ингибитор RaiA, осуществляющий арест трансляции в условиях стрессового ответа	0,57	3,75E-02
BLGT_08010	белок семейства нуклеотидилтрансфераз с широким спектром возможных функций	0,83	3,63E-02
BLGT_08670	белок семейства гликозилтрансфераз, участвующая предположительно в утилизации гликопротеинов мукоидного слоя	0,98	1,62E-04
BLGT_08680	пермеазный компонент системы транспорта сахаров ABC-типа; обеспечивает транспорт и метаболизм углеводов	1,27	9,30E-08
BLGT_08975	50S рибосомальный белок L3	-0,93	3,63E-02
BLGT_09920	белок, содержащий Ig-подобные домены типов II и III; предположительно участвует в адгезии	0,73	1,04E-03
BLGT_01155	мембранный белок с неизвестной функцией	1,06	2,47E-03
BLGT_04965	белок семейства антиоксидантных ферментов пероксиредоксинов, участвующий в реализации механизмов стрессового ответа	0,86	4,28E-02

Ген *fn3* и другие гены кластера PFNA бифидобактерий не были обнаружены среди дифференциально экспрессирующихся генов. Регуляция экспрессии генов оперона PFNA может не включать цитокиновую индукцию. Выполняя рецепторную роль, ген *fn3* может запускать сигнальный механизм, результатом которого, в свою очередь, может являться изменение транскрипционной активности ряда других генов.

Функциональная характеристика дифференциально экспрессирующихся генов показала, что добавление в среду IL-6 приводит к увеличению уровня экспрессии генов, участвующих в реализации механизмов стрессового ответа (BLGT_00625, BLGT_06345, BLGT_04965). Увеличение уровня экспрессии белка RaiA, участвующего в аресте трансляции в условиях стрессового ответа, сопровождается уменьшением уровня экспрессии 50S рибосомального белка L3, что приводит к снижению уровня трансляции и, следовательно, базового метаболизма клеток. Пероксиредоксин, кодируемый геном BLGT_04965, относится к семейству белков, которые контролируют уровень цитокин-индуцированных пероксидов в клетках млекопитающих [Paula *et al.*, 2013]. Сходные механизмы могут быть реализованы в бактериальных клетках. Также происходит увеличение уровня экспрессии группы генов, ответственных за утилизацию и потребление гликопротеинов и других субстратов (BLGT_00140, BLGT_01150, BLGT_08670,

BLGT_08680), что может быть связано с необходимостью накопления веществ для увеличения жизнеспособности в стрессовых условиях под воздействием факторов иммунного ответа.

Таким образом, нами экспериментально было показано, что добавление в среду провоспалительного цитокина IL-6 не оказывает влияния на рост штамма *B. longum* subsp. *longum* GT15. Транскриптомный анализ не выявил изменения в уровне экспрессии генов кластера PFNA по сравнению с нормальными условиями, однако выявил изменения в уровне экспрессии ряда других генов, предположительно участвующих в реализации механизмов устойчивости к факторам иммунного ответа бифидобактерий как комменсальных представителей микробиоты.

ЗАКЛЮЧЕНИЕ

С помощью анализа филогенетических профилей групп ортологичных белков у представителей различных видов рода *Bifidobacterium* нами была определена эволюционно стабильная группа генов в генетическом окружении гена *pkb2* (*pkb2*, *aaa-atp*, *fn3*, *duf58*, *tgm*). В анализ была включена геномная последовательность штамма *B. angulatum* GT 102, для которого мы провели полногеномное секвенирование и сборку последовательности в статусе «complete». Кластер генов получил название PFNA (*pkb2*, *fn3*, *aaa-atp*). На основе функциональной аннотации генов мы предположили, что функции кластера PFNA связаны с реализацией механизмов коммуникации с факторами занимаемой экологической ниши в организме хозяина, в частности, с механизмами сигнальной трансдукции, адгезии и взаимодействия с иммунной системой. Нами была проведена реконструкция эволюционных событий потери генов в различных ветвях и кладах филогенетического дерева бифидобактерий и установлены предполагаемые механизмы компенсации функции генов кластера PFNA у отдельных видов бифидобактерий.

Транскрипционная организация кластера PFNA была изучена в геноме *B. longum* subsp. *longum* GT15. В результате анализа нами было установлено, что гены *pkb2*, *fn3*, *aaa-atp*, *duf58*, *tgm*, а также гены *prpC*, *fha* и ген BLGT_RS02790 в геноме *B. longum* subsp. *longum* GT15 транскрибируются в составе единого оперона. Оперонная организация подтверждает участие генов кластера PFNA в реализации общей функции. Нами была обнаружена точка старта транскрипции оперона, находящаяся на расстоянии -40-41 п. о. от стартового кодона белок-кодирующей последовательности гена *pkb2*. Также дополнительно была обнаружена точка старта транскрипции, локализованная внутри последовательности гена *tgm* на расстоянии -58-61 п. о. от стартового кодона белок-кодирующей последовательности гена *prpC*, что может свидетельствовать о комплексной регуляции транскрипции генов оперона в геноме *B. longum* subsp. *longum* GT15. В результате анализа *in silico* был обнаружен р-независимый терминатор транскрипции оперона, находящийся на расстоянии +65 п. о. от стоп-кодона белок-кодирующей последовательности гена *fha*.

Нами было показано, что гены в составе кластера PFNA, подобно *pkb2*, обладают высоким уровнем межвидовой дивергенции последовательностей по сравнению с коровыми генами. Высокая степень межвидовой дивергенции последовательностей генов кластера PFNA у разных видов указывает на видоспецифичность выполняемых функций. Анализ молекулярной эволюции генов кластера PFNA бифидобактерий был проведен с целью выяснения причин аномально высокого уровня межвидовой дивергенции

последовательностей. Нами была подтверждена гипотеза о коэволюции последовательностей генов *pkb2*, *fn3*, *aaa-atp*, *duf58*, *tgm* кластера PFNA бифидобактерий и проведено построение филогенетического дерева бифидобактерий на основе объединенных последовательностей вышеупомянутых генов. Топология внешних клад полученного филогенетического дерева демонстрирует высокие значения поддержек узлов и с высокой точностью воспроизводит опубликованные филограммы бифидобактерий, несмотря на имеющиеся незначительные отличия. На основе анализа топологии ветвей и эволюционных дистанций было выдвинуто предположение о необходимости реклассификации исследуемых пар таксонов *B. indicum* и *B. coryneforme*, *B. catenulatum* и *B. kashiwanohense*, а также *B. pseudolongum* subsp. *pseudolongum* и *B. pseudolongum* subsp. *globosum*, что согласуется с литературными данными. Полученное филогенетическое дерево имеет более высокие значения длин ветвей по сравнению с длинами ветвей опубликованных филогенетических деревьев, построенных на основе объединенных последовательностей коровых генов. Данный факт указывает на высокую скорость молекулярной эволюции нуклеотидных последовательностей белок-кодирующих участков генов кластера PFNA по сравнению с консервативными последовательностями белок-кодирующих участков коровых генов, на которые, как правило, действует давление отрицательного отбора.

В качестве объяснения феномену быстрой эволюции белок-кодирующих участков генов кластера PFNA бифидобактерий была выдвинута гипотеза о наличии событий продолжительного положительного отбора в молекулярной эволюции последовательностей. Экспериментально *in silico* были обнаружены сайты, находящиеся под давлением положительного отбора во всех ветвях филогенетического дерева. Данные сайты были затем идентифицированы и локализованы в последовательности трансмембранных доменов первичной структуры белка, кодируемого геном *tgm* кластера PFNA бифидобактерий.

В дополнении к обнаруженным событиям продолжительного положительного отбора, нами также были протестированы гипотезы о наличии независимых событий эпизодического положительного отбора в молекулярной эволюции последовательностей генов кластера PFNA различных видов бифидобактерий. Эпизодический отбор, воздействующий на отдельные сайты в индивидуальных ветвях и кладах филогенетического дерева, представляет собой наиболее распространенный случай положительного отбора. Экспериментально *in silico* было показано наличие независимых событий эпизодического положительного отбора в молекулярной эволюции генов кластера PFNA для большинства ветвей филогенетического дерева бифидобактерий, что объясняет высокую степень межвидовой дивергенции

последовательностей. Проведенные тесты предполагают наличие групп сайтов под давлением положительного отбора в тестируемых ветвях по сравнению с остальными ветвями филогенетического дерева, следовательно, положительный отбор в молекулярной эволюции последовательностей генов кластера PFNA проявлял видоспецифический характер, воздействуя на группы сайтов в последовательностях, соответствующих отдельным ветвям филогении, независимо от других ветвей. Сайты-кандидаты для воздействия эпизодического положительного отбора в молекулярной эволюции последовательностей белок-кодирующих участков генов кластера PFNA были затем идентифицированы и локализованы в первичной структуре белков, кодируемых генами *pkb2*, *fn3*, *aaa-atp*, *duf58* и *tgm* представителей различных видов рода *Bifidobacterium*. В частности, такие сайты были обнаружены в каталитическом домене СТПК, кодируемой геном *pkb2*, и АТФазном домене АТФазы, кодируемой геном *aaa-atp* у представителей различных видов бифидобактерий. В белках, кодируемых генами *fn3* и *duf58* кластера PFNA, такие сайты были обнаружены как в аннотированных доменах за исключением ТМ, так и вне аннотированных доменов. В предполагаемой трансглутаминазе, кодируемой геном *tgm*, сайты под давлением положительного отбора были обнаружены на всем протяжении первичной структуры белка, включая трансмембранные домены.

Продолжительный и эпизодический положительный отбор в последовательностях вышеупомянутых белков может обеспечивать быструю эволюцию в ответ на воздействие внешних факторов. Быстрая эволюция и положительный отбор нередко затрагивают гены, вовлеченные в коммуникацию с факторами внешней среды, и особенно, биотическими экологическими факторами. Гипотеза Черной Королевы описывает эффекты, оказываемые на кодирующие последовательности генов в результате совместной эволюции взаимодействующих видов. Нами было установлено, что гены кластера PFNA, участвуя в реализации механизмов коммуникации с факторами внешней среды, представленной организмом хозяина, демонстрируют эффекты, описываемые гипотезой Черной Королевы. Необходимость адаптироваться к новым условиям в результате специализации и смены экологической ниши также могла увеличивать скорость эволюции в анализируемых генах. Сайты, на которые осуществляется давление отбора, могут быть структурно или функционально значимыми, и, таким образом, представляют собой вероятных кандидатов для расширения структурно-функциональной аннотации белков, кодируемых генами кластера PFNA, что представляет собой важную задачу, так как структура и функции данных белков малоизучены.

В данной диссертационной работе мы проверили гипотезу о наличии зависимости между уровнем экспрессии гена *fn3* и силой отбора в геномах представителей различных видов бифидобактерий. В результате количественного анализа экспрессии генов мы показали, что ген *fn3* транскрибируется в геномах всех исследуемых штаммов. Сравнительный анализ начальных значений сигнала флуоресценции показал, что уровень экспрессии гена *fn3* широко варьирует в различных пробах и не зависит от силы и характера естественного отбора.

Положительный отбор и быстрая эволюция вносят изменения в структуру ортологичных белков различных видов бифидобактерий, что может приводить к формированию видоспецифичности белок-белкового взаимодействия коэволюционирующих функциональных партнеров. Нами впервые была изучена видоспецифичность киназно-субстратного взаимодействия на примере штаммов *B. longum* subsp. *longum* GT15 и *B. choerinum* DSM 20434 в ходе киназной реакции. В результате анализа экспериментально была показана функциональная активность полноразмерных СТПК Pkb2 представителей видов *B. longum* subsp. *longum* и *B. choerinum*, субстратная видоспецифичность для СТПК Pkb2 *B. longum* subsp. *longum* GT15 и ее отсутствие для СТПК Pkb2 *B. choerinum* DSM 20434.

Быстрая эволюция и обнаруженный нами положительный отбор в последовательностях гена *fn3* бифидобактерий могут быть обусловлены участием данного гена в реализации механизмов коммуникации с иммунной системой организма хозяина в ходе цитокин-опосредованного сигналинга. В последней части работы нами была проведена оценка влияния IL-6 как фактора иммунного ответа на рост и экспрессию генов *B. longum* subsp. *longum* GT15. Экспериментально было показано, что добавление в среду провоспалительного цитокина IL-6 не оказывает влияния на рост штамма *B. longum* subsp. *longum* GT15. Транскриптомный анализ не выявил изменения в уровне экспрессии генов кластера PFNA по сравнению с нормальными условиями, однако выявил изменения в уровне экспрессии ряда других генов, предположительно участвующих в реализации механизмов устойчивости к факторам иммунного ответа бифидобактерий как комменсальных представителей микробиоты, в частности, группы генов стрессового ответа, а также группы генов, ответственных за утилизацию и потребление гликопротеинов и других субстратов.

ВЫВОДЫ

1. Установлено, что кластер генов PFNA (*pkb2*, *fn3*, *aaa-atp*, *duf58* и *tgm*) представляет собой эволюционно стабильную группу функционально связанных генов, обнаруживаемых совместно в геномах представителей различных видов рода *Bifidobacterium*, и, подобно *pkb2*, демонстрирующих высокую степень межвидовой дивергенции последовательностей;
2. Впервые установлена оперонная организация кластера генов PFNA в геноме *B. longum* subsp. *longum* GT15. Обнаружена точка старта транскрипции и ρ -независимый терминатор транскрипции оперона PFNA. Также была обнаружена дополнительная точка старта транскрипции внутри последовательности гена *tgm* на расстоянии -58-61 п. о. от стартового кодона белок-кодирующей последовательности гена *prpC*, что может свидетельствовать о комплексной регуляции транскрипции генов оперона в геноме *B. longum* subsp. *longum* GT15.
3. Установлены причины высокого уровня межвидовой дивергенции последовательностей генов кластера PFNA. В молекулярной эволюции гена *tgm* обнаружены события продолжительного положительного отбора во всех ветвях филогенетического дерева. В молекулярной эволюции генов *pkb2*, *fn3*, *aaa-atp*, *duf58* и *tgm* обнаружены независимые события эпизодического положительного отбора для большинства ветвей филогенетического дерева *Bifidobacterium*.
4. Ген *fn3* транскрибируется в геномах девяти исследуемых штаммов бифидобактерий различных видов. Сравнительный анализ начальных значений сигнала флуоресценции показал, что уровень экспрессии гена *fn3* широко варьирует в различных пробах и не зависит от силы и характера естественного отбора.
5. Впервые экспериментально показана функциональная активность полноразмерных СТПК Pkb2 представителей видов *B. longum* subsp. *longum* и *B. choerinum*, субстратная видоспецифичность для СТПК Pkb2 *B. longum* subsp. *longum* GT15 и ее отсутствие для СТПК Pkb2 *B. choerinum* DSM 20434.
6. Добавление в среду провоспалительного цитокина IL-6 не оказывает влияния на рост *B. longum* subsp. *longum* GT15 и экспрессию генов кластера PFNA, однако оказывает влияние на экспрессию ряда других генов, предположительно участвующих в реализации механизмов устойчивости к факторам иммунного ответа бифидобактерий.

СПИСОК ИСПОЛЬЗУЕМЫХ СОКРАЩЕНИЙ

АТФ – аденозинтрифосфат

ВКПМ – всероссийская коллекция промышленных микроорганизмов

ДМСО – диметилсульфоксид

ДНК – дезоксирибонуклеиновая кислота

ДТТ – дитиотреитол

ДЭГ – дифференциально экспрессирующиеся гены

ЖКТ – желудочно-кишечный тракт

ИПТГ – изопропил- β -D-1-тиогалактопиранозид

кДНК – комплементарная ДНК

ПААГ – полиакриламидный гель

ПЦР – полимеразная цепная реакция

ПЦР-РВ – ПЦР в реальном времени

РНК – рибонуклеиновая кислота (рРНК – рибосомная РНК)

СТПК – серин-треониновые протеинкиназы

ТМ – трансмембранный домен

ТХУ – трихлоруксусная кислота

ФМСФ – фенилметилсульфонилфторид

ФП – филогенетический профиль

ЭДТА – этилендиаминтетраацетат

AIC – Akaike information criterion (информационный критерий Акаике)

AICc – Akaike information criterion corrected (скорректированный информационный критерий Акаике)

БЕБ – Bayes empirical Bayes (эмпирический Байесовский подход)

BIC – Bayesian information criterion (Байесовский информационный критерий)

DUF – domain of unknown function (домен с неизвестной функцией)

FDR – false discovery rate

FHA – forkhead associated

FN – fibronectin (фибронектин)

HG – housekeeping genes (гены домашнего хозяйства)

HK – histidine protein kinase (гистидиновая протеинкиназа)

Ig – immunoglobulin (иммуноглобулин)

IL – interleukin (интерлейкин)

LRT – likelihood ratio test (тест отношения правдоподобия)

MI - mutual information (взаимная информация)

OD – optical density (оптическая плотность)

PASTA – penicillin-binding protein and Ser/Thr kinase-associated

PFNA – *pkb2, fn3, aaa-atp*

PP – posterior probability (апостериорная вероятность)

RACE – rapid amplification of cDNA ends (быстрая амплификация концов кДНК)

RBH – reciprocal best hit (наилучшее реципрокное попадание)

RFU – relative fluorescence unit (относительная единица флуоресценции)

RR – response regulator (регулятор ответа)

TAE – tris-acetate-EDTA (трис-ацетатный буфер)

TNF – tumor necrosis factor (фактор некроза опухоли)

UPGMA - unweighted pair group method with arithmetic mean (метод невзвешенного попарного среднего)

wt – wild type (дикий тип)

СПИСОК ЛИТЕРАТУРЫ

1. Аверина, О. В., Ермоленко, Е. И., Ратушный, А. Ю., Тарасова, Е. А., Борщев, Ю. Ю., Леонтьева, Г. Ф., Крамская, Т. А., Котылева, М. П., Даниленко, В. Н., Суворов, А. Н. (2015). Влияние пробиотиков на продукцию цитокинов в системах *in vitro* и *in vivo*. Медицинская иммунология, 17 (5), 443-454.
2. Алексеева, М. Г., Мавлетова, Д. А., Колчина, Н. В., Незаметдинова, В. З., Даниленко, В. Н. (2015). Выделение и очистка рекомбинантных белков серин-треониновых протеинкиназ штамма *Bifidobacterium longum* b379m и изучение их активности. Биохимия, 80 (10), 1578 – 1587.
3. Adlerberth, I., Strachan, D. P., Matricardi, P. M., Ahrne, S., Orfei, L., Aberg, N., Perkin, M. R., Tripodi, S., Hesselmar, B., Saalman, R., Coates, A. R., Bonanno, C. L., Panetta, V. and Wold, A. E. (2007). Gut microbiota and development of atopic eczema in 3 European birth cohorts. Journal of allergy and clinical immunology, 120(2), 343-350.
4. Alba, M. M., and Castresana, J. (2004). Inverse relationship between evolutionary rate and age of mammalian genes. Molecular biology and evolution, 22(3), 598-606.
5. Alba, M. M., and Castresana, J. (2007). On homology searches by protein Blast and the characterization of the age of genes. BMC evolutionary biology, 7(1), 53.
6. Alber, T. (2009). Signaling mechanisms of the *Mycobacterium tuberculosis* receptor Ser/Thr protein kinases. Current opinion in structural biology, 19(6), 650-657.
7. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. Journal of molecular biology, 215(3), 403-410.
8. Anisimova, M., and Yang, Z. (2007). Multiple hypothesis testing to detect lineages under positive selection that affects only a few sites. Molecular biology and evolution, 24(5), 1219-1228.
9. Barrett, L. G., Thrall, P. H., Dodds, P. N., Van der Merwe, M., Linde, C. C., Lawrence, G. J., and Burdon, J. J. (2009). Diversity and evolution of effector loci in natural populations of the plant pathogen *Melampsora lini*. Molecular biology and evolution, 26(11), 2499-2513.
10. Begun, D. J., and Whitley, P. (2000). Adaptive evolution of Relish, a *Drosophila* NF- κ B/I κ B protein. Genetics, 154(3), 1231-1238.
11. Bernet, M. F., Brassart, D., Neeser, J. R., and Servin, A. L. (1993). Adhesion of human bifidobacterial strains to cultured human intestinal epithelial cells and inhibition of enteropathogen-cell interactions. Applied and environmental microbiology, 59(12), 4121-4128.

12. Bibiloni, R., Perez, P. F., Garrote, G. L., Disalvo, E. A., and De Antoni, G. L. (2001). Surface characterization and adhesive properties of bifidobacteria. In *Methods in enzymology*, 336, 411-427. Academic Press.
13. Bivar Xavier, K. (2018). Bacterial interspecies quorum sensing in the mammalian gut microbiota. *Comptes rendus biologiques*, 341(5), 297-299.
14. Blanc, G., Ngwamidiba, M., Ogata, H., Fournier, P. E., Claverie, J. M., and Raoult, D. (2005). Molecular evolution of rickettsia surface antigens: evidence of positive selection. *Molecular biology and evolution*, 22(10), 2073-2083.
15. Bottacini, F., Milani, C., Turrone, F., Sanchez, B., Foroni, E., Duranti, S., Serafini, F., Viappiani, A., Strati, F., Ferrarini, A., Delledonne, M., Henrissat, B., Coutinho, P., Fitzgerald, G. F., Margolles, A., van Sinderen, D. and Ventura M. (2012). *Bifidobacterium asteroides* PRL2011 genome analysis reveals clues for colonization of the insect gut. *PLoS one*, 7(9), e44229.
16. Bravo, J., Staunton, D., Heath, J. K., and Jones, E. Y. (1998). Crystal structure of a cytokine-binding region of gp130. *The EMBO Journal*, 17(6), 1665-1674.
17. Bustin, S. A. (2002). Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *Journal of molecular endocrinology*, 29(1), 23-39.
18. Candela, M., Bergmann, S., Vici, M., Vitali, B., Turrone, S., Eikmanns, B. J., Hammerschmidt, S. and Brigidi, P. (2007). Binding of human plasminogen to *Bifidobacterium*. *Journal of bacteriology*, 189(16), 5929-5936.
19. Cannarozzi G. M., and Schneider A. (2012). *Codon evolution: mechanisms and models*. Oxford: Oxford University Press.
20. Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular biology and evolution*, 17(4), 540-552.
21. Clark, A. G., Eisen, M. B., Smith, D. R., Bergman, C. M., Oliver, B., Markow, T. A., *et al.* (2007). Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature*, 450(7167), 203-18.
22. Clarke, G., Stilling, R. M., Kennedy, P. J., Stanton, C., Cryan, J. F., and Dinan, T. G. (2014). Minireview: gut microbiota: the neglected endocrine organ. *Molecular endocrinology*, 28(8), 1221-1238.
23. Csuros, M. (2010). Count: evolutionary analysis of phylogenetic profiles with parsimony and likelihood. *Bioinformatics*, 26(15), 1910-1912.
24. Cullen, T. W., Schofield, W. B., Barry, N. A., Putnam, E. E., Rundell, E. A., Trent, M. S., Degnan, P. H., Booth, C. J., Yu, H. and Goodman, A. L. (2015). Antimicrobial peptide

- resistance mediates resilience of prominent gut commensals during inflammation. *Science*, 347(6218), 170-175.
25. Darling, A. E., Mau, B., and Perna, N. T. (2010). ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PloS one*, 5(6), e11147.
 26. Daub, J. T., Moretti, S., Davydov, I. I., Excoffier, L., and Robinson-Rechavi, M. (2017). Detection of pathways affected by positive selection in primate lineages ancestral to humans. *Molecular biology and evolution*, 34(6), 1391-1402.
 27. Daubin, V., and Ochman, H. (2004). Bacterial genomes as new gene homes: the genealogy of ORFans in *E. coli*. *Genome research*, 14(6), 1036-1042.
 28. Domazet-Loso, T., and Tautz, D. (2003). An evolutionary analysis of orphan genes in *Drosophila*. *Genome research*, 13(10), 2213-2219.
 29. Drummond D. A., Bloom J. D., Adami C., Wilke C. O., and Arnold F. H. (2005). Why highly expressed proteins evolve slowly. *Proceedings of the National Academy of Sciences*, 102, 14338–14343.
 30. Emms, D. M., and Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome biology*, 16(1), 157.
 31. Enright A. J., Iliopoulos I., Kyripides N. C., and Ouzounis C. A. (1999). Protein interaction maps for complete genomes based on gene fusion events. *Nature*, 402, 86–90.
 32. Freestone, P. (2013). Communication between bacteria and their hosts. *Scientifica (Cairo)*, 2013, 361073.
 33. Frickey, T., and Lupas, A. (2004). CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics*, 20(18), 3702-3704.
 34. Fujimura, K. E., Sitarik, A. R., Havstad, S., Lin, D. L., Levan, S., Fadrosh, D., Panzer, A. R., LaMere, B., Rackaityte, E., Lukacs, N. W., Wegienka, G., Boushey, H. A., Ownby, D. R., Zoratti, E. M., Levin, A. M., Johnson, C. C. and Lynch, S. V. (2016). Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nature medicine*, 22(10), 1187.
 35. Gao, B., Paramanathan, R., and Gupta, R. S. (2006). Signature proteins that are distinctive characteristics of *Actinobacteria* and their subgroups. *Antonie Van Leeuwenhoek*, 90(1), 69-91.
 36. Garcia-Vallve, S., Alonso, A., and Bravo, I. G. (2005). Papillomaviruses: different genes have different histories. *Trends in microbiology*, 13(11), 514-521.
 37. Gay, L. M., Ng, H. L., and Alber, T. (2006). A conserved dimer and global conformational

- changes in the structure of apo-PknE Ser/Thr protein kinase from *Mycobacterium tuberculosis*. *Journal of molecular biology*, 360(2), 409-420.
38. Gopal, P. K., Prasad, J., Smart, J., and Gill, H. S. (2001). In vitro adherence properties of *Lactobacillus rhamnosus* DR20 and *Bifidobacterium lactis* DR10 strains and their antagonistic activity against an enterotoxigenic *Escherichia coli*. *International journal of food microbiology*, 67(3), 207-216.
 39. Greenstein, A. E., Echols, N., Lombana, T. N., King, D. S., and Alber, T. (2007). Allosteric activation by dimerization of the PknD receptor Ser/Thr protein kinase from *Mycobacterium tuberculosis*. *Journal of biological chemistry*, 282(15), 11427-11435.
 40. Griffin, M., Casadio, R., and Bergamini, C. M. (2002). Transglutaminases: nature's biological glues. *Biochemical journal*, 368(Pt 2), 377-96.
 41. Grimm, V., Gleinser, M., Neu, C., Zhurina, D., and Riedel, C. U. (2014). Expression of fluorescent proteins in bifidobacteria for analysis of host-microbe interaction. *Applied and environmental microbiology*, AEM-04261.
 42. Hammet, A., Pike, B., McNees, C., Conlan, L., Tennis, N., and Heierhorst, J. (2003). FHA domains as phospho-threonine binding modules in cell signaling. *IUBMB life*, 55(1), 23-27.
 43. Hanks, S. K. (2003). Genomic analysis of the eukaryotic protein kinase superfamily: a perspective. *Genome biology*, 4(5), 111.
 44. Hanks, S. K., and Hunter, T. (1995). Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *The FASEB journal*, 9(8), 576-596.
 45. Harpaz, Y., and Chothia, C. (1994). Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains. *Journal of molecular biology*, 238(4), 528-539.
 46. He, F., Ouwehand, A. C., Hashimoto, H., Isolauri, E., Benno, Y., and Salminen, S. (2001). Adhesion of *Bifidobacterium* spp. to human intestinal mucus. *Microbiology and immunology*, 45(3), 259-262.
 47. Hedrick, P. W. (1994). Evolutionary genetics of the major histocompatibility complex. *The american naturalist*, 143(6), 945-964.
 48. Henderson, B., Nair, S., Pallas, J., and Williams, M. A. (2011). Fibronectin: a multidomain host adhesin targeted by bacterial fibronectin-binding proteins. *FEMS microbiology reviews*, 35(1), 147-200.
 49. Hoskins, L. C. (2012). Degradation of mucus glycoproteins in the gastrointestinal tract. *The glycoconjugates*, 2, 235-253.

50. Hsu, F., Schwarz, S., and Mougous, J. D. (2009). TagR promotes PpkA-catalysed type VI secretion activation in *Pseudomonas aeruginosa*. *Molecular microbiology*, 72(5), 1111-1125.
51. Huang, X., Pinto, D., Fritz, G., and Mascher, T. (2015). Environmental sensing in *Actinobacteria*: A comprehensive survey on the signaling capacity of this phylum. *Journal of bacteriology*, JB-00176.
52. Hughes, D. T., and Sperandio, V. (2008). Inter-kingdom signalling: communication between bacteria and their hosts. *Nature Reviews Microbiology*, 6(2), 111-20.
53. Huse, M., and Kuriyan, J. (2002). The conformational plasticity of protein kinases. *Cell*, 109(3), 275-282.
54. Iyer, L. M., Leipe, D. D., Koonin, E. V., and Aravind, L. (2004). Evolutionary history and higher order classification of AAA+ ATPases. *Journal of structural biology*, 146(1-2), 11-31.
55. Izquierdo, E., Medina, M., Ennahar, S., Marchioni, E., and Sanz, Y. (2008). Resistance to simulated gastrointestinal conditions and adhesion to mucus as probiotic criteria for *Bifidobacterium longum* strains. *Current microbiology*, 56(6), 613-618.
56. Jiggins, F. M., and Kim, K. W. (2007). A screen for immunity genes evolving under positive selection in *Drosophila*. *Journal of evolutionary biology*, 20(3), 965-970.
57. Kainulainen, V., and Korhonen, T. K. (2014). Dancing to another tune - adhesive moonlighting proteins in bacteria. *Biology*, 3 (2014), 178e204.
58. Kalliomaki, M., Kirjavainen, P., Eerola, E., Kero, P., Salminen, S., and Isolauri, E. (2001). Distinct patterns of neonatal gut microflora in infants in whom atopy was and was not developing. *Journal of allergy and clinical immunology*, 107(1), 129-134.
59. Kelly, S., and Maini, P. K. (2013). DendroBLAST: approximate phylogenetic trees in the absence of multiple sequence alignments. *PloS one*, 8(3), e58537.
60. Klijn, A., Mercenier, A., and Arigoni, F. (2005). Lessons from the genomes of bifidobacteria. *FEMS Microbiology Reviews*, 29(3), 491-509.
61. Koester J. A., Swanson W. J., and Armbrust E. V. (2012). Positive selection within a diatom species acts on putative protein interactions and transcriptional regulation. *Molecular biology and evolution*, 30, 422-434.
62. Konkel, M. E., Larson, C. L., and Flanagan, R. C. (2010). *Campylobacter jejuni* FlpA binds fibronectin and is required for maximal host cell adherence. *Journal of bacteriology*, 192(1), 68-76.
63. Koropatkin, N. M., Cameron, E. A., and Martens, E. C. (2012). How glycan metabolism shapes the human gut microbiota. *Nature reviews microbiology*, 10(5), 323.

64. Kosiol C., Vinar T., Da F. R., Hubisz M. J., Bustamante C. D., Nielsen R., and Siepel, A. (2008). Patterns of positive selection in six mammalian genomes. *PLoS genetics*, 4(8), e1000144.
65. Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular biology and evolution*, 33(7), 1870-1874.
66. Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., and Calcott, B. (2016). PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, 34(3), 772-773.
67. Lazzaro, B. P. (2005). Elevated polymorphism and divergence in the class C scavenger receptors of *Drosophila melanogaster* and *D. simulans*. *Genetics*, 169(4), 2023–2034.
68. Lesouhaitier, O., Veron, W., Chapalain, A., Madi, A., Blier, A. S., Dagorn, A., Connil, N., Chevalier, S., Orange, N., and Feuilloley, M. (2009). Gram-negative bacterial sensors for eukaryotic signal molecules. *Sensors*, 9(9), 6967-6990.
69. Letunic, I., and Bork, P. (2017). 20 years of the SMART protein domain annotation resource. *Nucleic acids research*, 46(D1), D493-D496.
70. Letunic, I., Doerks, T., and Bork, P. (2014). SMART: recent updates, new developments and status in 2015. *Nucleic acids research*, 43(D1), D257-D260.
71. Lievin, V., Peiffer, I., Hudault, S., Rochat, F., Brassart, D., Neeser, J. R., and Servin, A. L. (2000). *Bifidobacterium* strains from resident infant human gastrointestinal microflora exert antimicrobial activity. *Gut*, 47(5), 646-652.
72. Little, T. J., and Cobbe, N. (2005). The evolution of immune-related genes from disease carrying mosquitoes: diversity in a peptidoglycan-and a thioester-recognizing protein. *Insect molecular biology*, 14(6), 599-605.
73. Little, T. J., Colbourne, J. K., and Crease, T. J. (2004). Molecular evolution of *Daphnia* immunity genes: Polymorphism in a Gram-negative binding protein gene and an α -2-macroglobulin gene. *Journal of Molecular Evolution*, 59(4), 498-506.
74. Ludwig, W., Euzéby, J. P., and Whitman, W. B. (2015). Taxonomic outline of the phylum *Actinobacteria*. *Bergey's Manual of Systematics of Archaea and Bacteria*, 1-4.
75. Lugli, G. A., Mangifesta, M., Duranti, S., Anzalone, R., Milani, C., Mancabelli, L., Alessandri, G., Turrone, F., Ossiprandi, M. C., van Sinderen, D. and Ventura, M. (2018). Phylogenetic classification of six novel species belonging to the genus *Bifidobacterium* comprising *Bifidobacterium anseris* sp. nov., *Bifidobacterium criceti* sp. nov., *Bifidobacterium imperatoris* sp. nov., *Bifidobacterium italicum* sp. nov., *Bifidobacterium*

- margollesii* sp. nov. and *Bifidobacterium parmae* sp. nov. Systematic and applied microbiology, 41(3), 173-183.
76. Lugli, G. A., Milani, C., Turrone, F., Duranti, S., Ferrario, C., Viappiani, A., Mancabelli, L., Mangifesta, M., Taminiau, B., Delcenserie, V., van Sinderen, D. and Ventura, M. (2014). Investigation of the evolutionary development of the genus *Bifidobacterium* by comparative genomics. Applied and environmental microbiology, 80(20), 6383-6394.
 77. Lugli, G. A., Milani, C., Turrone, F., Duranti, S., Mancabelli, L., Mangifesta, M., Ferrario, C., Modesto, M., Mattarelli, P., Jiri, K., van Sinderen, D. and Ventura, M. (2017). Comparative genomic and phylogenomic analyses of the *Bifidobacteriaceae* family. BMC genomics, 18(1), 568.
 78. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular cloning: a laboratory manual, 545. Cold Spring Harbor, NY: Cold spring harbor laboratory.
 79. Marcotte E. M., Pellegrini M., Ng H-L., Rice D. W., Yeates T. O., and Eisenberg D. (1999). Detecting protein function and protein-protein interactions from genome sequences. Science, 285, 751–753.
 80. Marcotte E. M., Pellegrini M., Thompson M. J., Yeates T. O., and Eisenberg D. (1999). A combined algorithm for genome-wide prediction of protein function. Nature, 402, 83–86.
 81. Mascher, T., Helmann, J. D., and Uden, G. (2006). Stimulus perception in bacterial signal-transducing histidine kinases. Microbiology and molecular biology reviews, 70(4), 910-938.
 82. Mattarelli, P., Biavati, B., Holzapfel, W. H., and Wood, B. J. (2017). The bifidobacteria and related organisms: biology, taxonomy, applications. Academic Press.
 83. Matz, M., Shagin, D., Bogdanova, E., Britanova, O., Lukyanov, S., Diatchenko, L., and Chenchik, A. (1999). Amplification of cDNA ends based on template-switching effect and step-out PCR. Nucleic acids research, 27(6), 1558-1560.
 84. McGuckin, M. A., Linden, S. K., Sutton, P., and Florin, T. H. (2011). Mucin dynamics and enteric pathogens. Nature Reviews Microbiology, 9(4), 265.
 85. Meduri, G. U., Kanangat, S., Stefan, J., Tolley, E., and Schaberg, D. (1999). Cytokines IL-1 β , IL-6, and TNF- α enhance *in vitro* growth of bacteria. American journal of respiratory and critical care medicine, 160(3), 961-967.
 86. Messier, W., and Stewart, C. B. (1997). Episodic adaptive evolution of primate lysozymes. Nature, 385(6612), 151.
 87. Michelini, S., Modesto, M., Filippini, G., Spiezio, C., Sandri, C., Biavati, B., Pisi, A. and Mattarelli, P. (2016). *Bifidobacterium aerophilum* sp. nov., *Bifidobacterium avesanii* sp. nov. and *Bifidobacterium ramosum* sp. nov.: three novel taxa from the faeces of cotton-top

- tamarin (*Saguinus oedipus* L.). Systematic and applied microbiology, 39(4), 229-236.
88. Michelini, S., Oki, K., Yanokura, E., Shimakawa, Y., Modesto, M., Mattarelli, P., Biavati, B. and Watanabe, K. (2016). *Bifidobacterium myosotis* sp. nov., *Bifidobacterium tissieri* sp. nov. and *Bifidobacterium hapali* sp. nov., isolated from faeces of baby common marmosets (*Callithrix jacchus* L.). International journal of systematic and evolutionary microbiology, 66(1), 255-265.
 89. Miller, M., Donat, S., Rakette, S., Stehle, T., Kouwen, T. R., Diks, S. H., Dreisbach, A., Reilman, E., Gronau, K., Becher, D., Peppelenbosch, M. P., van Dijl, J. M. and Ohlsen, K. (2010). Staphylococcal PknB as the first prokaryotic representative of the proline-directed kinases. PloS one, 5(2), e9057.
 90. Modesto, M., Michelini, S., Stefanini, I., Ferrara, A., Tacconi, S., Biavati, B., and Mattarelli, P. (2014). *Bifidobacterium aesculapii* sp. nov., from the faeces of the baby common marmoset (*Callithrix jacchus*). International journal of systematic and evolutionary microbiology, 64(8), 2819-2827.
 91. Moreno-Hagelsieb, G. (2015). The power of operon rearrangements for predicting functional associations. Computational and structural biotechnology journal, 13, 402-406.
 92. Mu, J., Awadalla, P., Duan, J., McGee, K. M., Keebler, J., Seydel, K., McVean G. A., and Su X. Z. (2007). Genome-wide variation and identification of vaccine targets in the *Plasmodium falciparum* genome. Nature genetics, 39(1), 126.
 93. Munoz-Dorado, J., Inouye, S., and Inouye, M. (1991). A gene encoding a protein serine/threonine kinase is required for normal development of *M. xanthus*, a gram-negative bacterium. Cell, 67(5), 995-1006.
 94. Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., Kosakovsky Pond, S. L. (2012). Detecting individual sites subject to episodic diversifying
 95. Nei, M., and Kumar, S. (2000). Molecular evolution and phylogenetics. Oxford university press.
 96. Neves, F., Abrantes, J., Steinke, J. W., and Esteves, P. J. (2014). Maximum-likelihood approaches reveal signatures of positive selection in IL genes in mammals. Innate immunity, 20(2), 184-191.
 97. Nezametdinova, V. Z., Mavletova, D. A., Alekseeva, M. G., Chekalina, M. S., Zakharevich, N. V., and Danilenko, V. N. (2018). Species-specific serine-threonine protein kinase Pkb2 of *Bifidobacterium longum* subsp. longum: Genetic environment and substrate specificity. Anaerobe, 51, 26-35.
 98. Nezametdinova, V. Z., Zakharevich, N. V., Alekseeva, M. G., Averina, O. V., Mavletova, D.

- A., and Danilenko, V. N. (2014). Identification and characterization of the serine/threonine protein kinases in *Bifidobacterium*. *Archives of microbiology*, 196(2), 125-136.
99. Nielsen, R., and Yang, Z. (1998). Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics*, 148(3), 929-936.
 100. Obbard, D. J., Jiggins, F. M., Halligan, D. L., and Little, T. J. (2006). Natural selection drives extremely rapid evolution in antiviral RNAi genes. *Current Biology*, 16(6), 580-585.
 101. Ochoa, D., and Pazos, F. (2010). Studying the co-evolution of protein families with the Mirrortree web server. *Bioinformatics*, 26(10), 1370-1371.
 102. Ortiz-Lombardia, M., Pompeo, F., Boitel, B., and Alzari, P. M. (2003). Crystal structure of the catalytic domain of the PknB serine/threonine kinase from *Mycobacterium tuberculosis*. *Journal of Biological Chemistry*, 278(15), 13094-13100.
 103. Page, R. D., and Holmes, E. C. (2009). *Molecular evolution: a phylogenetic approach*. John Wiley and Sons.
 104. Pallen, M., Chaudhuri, R., and Khan, A. (2002). Bacterial FHA domains: neglected players in the phospho-threonine signalling game? *Trends in microbiology*, 10(12), 556-563.
 105. Paterson, S., Vogwill, T., Buckling, A., Benmayor, R., Spiers, A. J., Thomson, N. R., Quail, M., Smith, F., Walker, D., Libberton, B., Fenton, A., Hall, N., and Brockhurst, M. A. (2010). Antagonistic coevolution accelerates molecular evolution. *Nature*, 464(7286), 275.
 106. Paula, F. M., Ferreira, S. M., Boschero, A. C., and Souza, K. L. (2013). Modulation of the peroxiredoxin system by cytokines in insulin-producing RINm5F cells: down-regulation of PRDX6 increases susceptibility of beta cells to oxidative stress. *Molecular and cellular endocrinology*, 374(1-2), 56-64.
 107. Pavlova, A., Gan, H. M., Lee, Y. P., Austin, C. M., Gilligan, D. M., Lintermans, M., and Sunnucks, P. (2017). Purifying selection and genetic drift shaped Pleistocene evolution of the mitochondrial genome in an endangered Australian freshwater fish. *Heredity*, 118(5), 466.
 108. Peacock, C. (2015). *Parasite Genomics Protocols*. Humana Press.
 109. Pellegrini, M. (2012). Using phylogenetic profiles to predict functional relationships. In *Bacterial molecular networks*. Springer, New York, NY.
 110. Pellegrini, M., Marcotte, E. M., Thompson, M. J., Eisenberg, D., and Yeates, T. O. (1999). Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. *Proceedings of the National Academy of Sciences*, 96(8), 4285-4288.

111. Pereira, S. F., Goss, L., and Dworkin, J. (2011). Eukaryote-like serine/threonine kinases and phosphatases in bacteria. *Microbiology and Molecular Biology Reviews*, 75(1), 192-212.
112. Petersen, L., Bollback, J. P., Dimmic, M., Hubisz, M., and Nielsen, R. (2007). Genes under positive selection in *Escherichia coli*. *Genome Research*, 17(9), 1336–1343.
113. Pizarro-Cerda, J., and Cossart, P. (2006). Bacterial adhesion and entry into host cells. *Cell*, 124(4), 715-727.
114. Poupard, J. A., Husain, I. N. T. I. S. A. R., and Norris, R. F. (1973). Biology of the bifidobacteria. *Bacteriological reviews*, 37(2), 136.
115. Preedy, V. R. (2016). Adhesion molecules. CRC Press.
116. Prsic, S., Dankwa, S., Schwartz, D., Chou, M. F., Locasale, J. W., Kang, C. M., Bemis, G., Church, G. M., Steen, H. and Husson, R. N. (2010). Extensive phosphorylation with overlapping specificity by *Mycobacterium tuberculosis* serine/threonine protein kinases. *Proceedings of the National Academy of Sciences*, 107(16), 7521-7526.
117. Reuter, G. (2001). The *Lactobacillus* and *Bifidobacterium* microflora of the human intestine: composition and succession. *Current issues in intestinal microbiology*, 2(2), 43-53.
118. Ringel-Kulka, T., Cheng, J., Ringel, Y., Salojärvi, J., Carroll, I., Palva, A., de Vos, W. M. and Satokari, R. (2013). Intestinal microbiota in healthy US young children and adults—a high throughput microarray analysis. *PLoS one*, 8(5), e64315.
119. Ruggiero, A., De Simone, P., Smaldone, G., Squeglia, F., and Berisio, R. (2012). Bacterial cell division regulation by Ser/Thr kinases: a structural perspective. *Current protein and peptide science*, 13(8), 756-766.
120. Ruiz, L., Delgado, S., Ruas-Madiedo, P., Sanchez, B., and Margolles, A. (2017). Bifidobacteria and their molecular communication with the immune system. *Frontiers in microbiology*, 8, 2345.
121. Sackton, T. B., Lazzaro, B. P., Schlenke, T. A., Evans, J. D., Hultmark, D., and Clark, A. G. (2007). Dynamic evolution of the innate immune system in *Drosophila*. *Nature genetics*, 39(12), 1461.
122. Sali A. (1999). Functional links between proteins. *Nature*, 402, 23–26.
123. Schell, M. A., Karmirantzou, M., Snel, B., Vilanova, D., Berger, B., Pessi, G., Zwahlen, M. C., Desiere, F., Bork, P., Delley, M., Pridmore, R. D. and Arigoni, F. (2002). The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. *Proceedings of the National Academy of Sciences*, 99(22), 14422-14427.

124. Schlenke, T. A., and Begun, D. J. (2003). Natural selection drives *Drosophila* immune system evolution. *Genetics*, 164(4), 1471-1480.
- selection. *PLoS genetics*, 8(7), e1002764.
125. Sgorbati, B., Biavati, B., and Palenzona, D. (1995). The genus *Bifidobacterium*. In *The genera of lactic acid bacteria*, 279-306. Springer, Boston, MA.
126. Shah, I. M., Laaberki, M. H., Popham, D. L., and Dworkin, J. (2008). A eukaryotic-like Ser/Thr kinase signals bacteria to exit dormancy in response to peptidoglycan fragments. *Cell*, 135(3), 486-496.
127. Singh, R. S., Xu, J., and Kulathinal, R. J. (2012). *Rapidly evolving genes and genetic systems*. Oxford University Press.
128. Skunca, N., and Dessimoz, C. (2015). Phylogenetic profiling: how much input data is enough? *PloS one*, 10(2), e0114701.
129. Snider, J., and Houry, W. A. (2006). MoxR AAA+ ATPases: a novel family of molecular chaperones? *Journal of structural biology*, 156(1), 200-209.
130. Snider, J., Thibault, G., and Houry, W. A. (2008). The AAA+ superfamily of functionally diverse proteins. *Genome biology*, 9(4), 216.
131. Spielman, S. J., and Wilke, C. O. (2013). Membrane environment imposes unique selection pressures on transmembrane domains of G protein-coupled receptors. *Journal of molecular evolution*, 76(3), 172-182.
132. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312-1313.
133. Stenseth, N. C., and Smith, J. M. (1984). Coevolution in ecosystems: Red Queen evolution or stasis? *Evolution*, 38(4), 870-880.
134. Sun, Y. C., Hinnebusch, B. J., and Darby, C. (2008). Experimental evidence for negative selection in the evolution of a *Yersinia pestis* pseudogene. *Proceedings of the National Academy of Sciences*, 105(23), 8097-101.
135. Sun, Z., Zhang, W., Guo, C., Yang, X., Liu, W., Wu, Y., Song, Y., Kwok, L. Y , Cui, Y., Menghe, B., Yang, R., Hu, L. and Zhang, H. (2015). Comparative genomic analysis of 45 type strains of the genus *Bifidobacterium*: a snapshot of its genetic diversity and evolution. *PLoS one*, 10(2), e0117912.
136. Suzuki, Y., and Gojobori, T. (1999). A method for detecting positive selection at single amino acid sites. *Molecular biology and evolution*, 16(10), 1315-1328.
137. Swidsinski, A., Loening-Baucke, V., Theissig, F., Engelhardt, H., Bengmark, S., Koch, S., Lochs, H., and Dorffel, Y. (2007). Comparative study of the intestinal mucus

- barrier in normal and inflamed colon. *Gut*, 56(3), 343-350.
138. Tailford, L. E., Crost, E. H., Kavanaugh, D., and Juge, N. (2015). Mucin glycan foraging in the human gut microbiome. *Frontiers in genetics*, 6, 81.
 139. Talavera, G., and Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic biology*, 56(4), 564-577.
 140. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research*, 22(22), 4673-4680.
 141. Tissier, H. (1900). *Recherches sur la flore intestinale des nourrissons (e'tat normal et pathologique)* (Doctoral dissertation).
 142. Toyama, H., Anthony, C., and Lidstrom, M. E. (1998). Construction of insertion and deletion *mx*A mutants of *Methylobacterium extorquens* AM1 by electroporation. *FEMS Microbiology Letters*, 166(1), 1-7.
 143. Tsuchida, S., Takahashi, S., Nguema, P. P. M., Fujita, S., Kitahara, M., Yamagiwa, J., Ngomanda, A., Ohkuma, M. and Ushida, K. (2014). *Bifidobacterium moukalabense* sp. nov., isolated from the faeces of wild west lowland gorilla (*Gorilla gorilla gorilla*). *International journal of systematic and evolutionary microbiology*, 64(Pt 2), 449-455.
 144. Turroni, F., Peano, C., Pass, D. A., Foroni, E., Severgnini, M., Claesson, M. J., Kerr, C., Hourihane, J., Murray, D., Fuligni, F., Gueimonde, M., Margolles, A., De Bellis, G., O'Toole, P. W., van Sinderen, D., Marchesi, J. R. and Ventura, M. (2012). Diversity of bifidobacteria within the infant gut microbiota. *PloS one*, 7(5), e36957.
 145. Ulrich, L. E., Koonin, E. V., and Zhulin, I. B. (2005). One-component systems dominate signal transduction in prokaryotes. *Trends in Microbiology*, 13(2), 52-56.
 146. Valk, V., van der Kaaij, R. M., and Dijkhuizen, L. (2017). The evolutionary origin and possible functional roles of FNIII domains in two *Microbacterium aurum* B8. A granular starch degrading enzymes, and in other carbohydrate acting enzymes. *Amylase*, 1(1), 1-11.
 147. Van Spanning, R. J., Wansell, C. W., De Boer, T., Hazelaar, M. J., Anazawa, H., Harms, N., Oltmann, L. F. and Stouthamer, A. H. (1991). Isolation and characterization of the *moxJ*, *moxG*, *moxI*, and *moxR* genes of *Paracoccus denitrificans*: inactivation of *moxJ*, *moxG*, and *moxR* and the resultant effect on methylotrophic growth. *Journal of bacteriology*, 173(21), 6948-6961.

148. Van Valen, L. (1973). A new evolutionary law. *Evolutionary Theory*, 1, 1-30.
149. Van Valen, L. (1974). Molecular evolution as predicted by natural selection. *Journal of molecular evolution*, 3(2), 89-101.
150. Vazquez-Gutierrez, P., Stevens, M. J. A., Gehrig, P., Barkow-Oesterreicher, S., Lacroix, C., and Chassard, C. (2017). The extracellular proteome of two *Bifidobacterium* species reveals different adaptation strategies to low iron conditions. *BMC Genomics*, 18, 41.
151. Ventura, M., Canchaya, C., Del Casale, A., Dellaglio, F., Neviani, E., Fitzgerald, G. F., and Van Sinderen, D. (2006). Analysis of bifidobacterial evolution using a multilocus approach. *International journal of systematic and evolutionary microbiology*, 56(12), 2783-2792.
152. Ventura, M., Canchaya, C., Fitzgerald, G. F., Gupta, R. S., and van Sinderen, D. (2007). Genomics as a means to understand bacterial phylogeny and ecological adaptation: the case of bifidobacteria. *Antonie Van Leeuwenhoek*, 91(4), 351-372.
153. Ventura, M., van Sinderen, D., Fitzgerald, G. F. and Zink, R. (2004). Insights into the taxonomy, genetics and physiology of bifidobacteria. *Antonie Van Leeuwenhoek*, 86(3), 205-223.
154. Viljakainen, L. and Pamilo, P. (2008). Selection on an antimicrobial peptide defensin in ants. *Journal of molecular evolution*, 67(6), 643.
155. Viljakainen, L., Evans, J. D., Hasselmann, M., Rueppell, O., Tingek, S. and Pamilo, P. (2009). Rapid evolution of immune proteins in social insects. *Molecular biology and evolution*, 26(8), 1791-1801.
156. Voolstra, C. R., Sunagawa, S., Matz, M. V., Bayer, T., Aranda, M., Buschiazzi, E., Desalvo, M. K., Lindquist, E., Szmant, A. M., Coffroth, M. A. and Medina, M. (2011). Rapid evolution of coral proteins responsible for interaction with the environment. *PloS one*, 6(5), e20392.
157. Watanabe, T., Suzuki, K., Oyanagi, W., Ohnishi, K., and Tanaka, H. (1990). Gene cloning of chitinase A1 from *Bacillus circulans* WL-12 revealed its evolutionary relationship to *Serratia* chitinase and to the type III homology units of fibronectin. *Journal of Biological Chemistry*, 265(26), 15659-15665.
158. Waterhouse, R. M., Kriventseva, E. V., Meister, S., Xi, Z., Alvarez, K. S., Bartholomay, L. C., Barillas-Mury, C., Bian, G., Blandin, S., Christensen, B. M., Dong, Y., Jiang, H., Kanost, M. R., Koutsos, A. C., Levashina, E. A., Li, J., Ligoxygakis, P., Maccallum, R. M., Mayhew, G. F., Mendes, A., Michel, K., Osta, M. A., Paskewitz, S., Shin,

- S. W., Vlachou, D., Wang, L., Wei, W., Zheng, L., Zou, Z., Severson, D. W., Raikhel, A. S., Kafatos, F. C., Dimopoulos, G., Zdobnov, E. M. and Christophides, G. K. (2007). Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science*, 316(5832), 1738-1743.
159. Westermann, C. (2015). Analysis of potential host-colonization factors in *Bifidobacterium bifidum* S17 (Doctoral dissertation, Universitat Ulm).
 160. Westermann, C., Gleinser, M., Corr, S. C., and Riedel, C. U. (2016). A critical evaluation of bifidobacterial adhesion to the host tissue. *Frontiers in microbiology*, 7, 1220.
 161. Wolanin, P. M., Thomason, P. A., and Stock, J. B. (2002). Histidine protein kinases: key signal transducers outside the animal kingdom. *Genome biology*, 3(10), reviews3013-1.
 162. Wolf, Y. I., Novichkov, P. S., Karev, G. P., Koonin, E. V., and Lipman, D. J. (2009). The universal distribution of evolutionary rates of genes and distinct characteristics of eukaryotic genes of different apparent ages. *Proceedings of the national academy of sciences*, 106(18), 7273-7280.
 163. Wong, K. S., and Houry, W. A. (2012). Novel structural and functional insights into the MoxR family of AAA+ ATPases. *Journal of structural biology*, 179(2), 211-221.
 164. Yalamanchili, H. K., Wan, Y. W., and Liu, Z. (2017). Data analysis pipeline for RNA-seq experiments: from differential expression to cryptic splicing. *Current protocols in bioinformatics*, 59(1), 11-15.
 165. Yang, Z. (1998). Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Molecular biology and evolution*, 15(5), 568-573.
 166. Yang, Z. (2007). PAML: phylogenetic analysis by maximum likelihood. *Molecular biology and evolution*, 24(8), 1586-1591.
 167. Yang, Z., and Nielsen, R. (2002). Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Molecular biology and evolution*, 19(6), 908-917.
 168. Yang, Z., Nielsen, R., Goldman, N., and Pedersen, A. M. K. (2000). Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics*, 155(1), 431-449.
 169. Yang, Z., Wong, W. S., and Nielsen, R. (2005). Bayes empirical Bayes inference of amino acid sites under positive selection. *Molecular biology and evolution*, 22(4), 1107-1118.
 170. Yeats, C., Finn, R. D., and Bateman, A. (2002). The PASTA domain: a β -lactam-

- binding domain. Trends in biochemical sciences, 27(9), 438-440.
171. Young, T. A., Delagoutte, B., Endrizzi, J. A., Falick, A. M., and Alber, T. (2003). Structure of *Mycobacterium tuberculosis* PknB supports a universal activation mechanism for Ser/Thr protein kinases. Nature structural and molecular biology, 10(3), 168.
 172. Yu, M., and Irwin, D. M. (1996). Evolution of stomach lysozyme: the pig lysozyme gene. Molecular phylogenetics and evolution, 5(2), 298-308.
 173. Zakharevich, N. V., Averina, O. V., Klimina, K. M., Kudryavtseva, A. V., Kasianov, A. S., Makeev, V. J., and Danilenko, V. N. (2015). Complete genome sequence of *Bifidobacterium longum* GT15: identification and characterization of unique and global regulatory genes. Microbial ecology, 70(3), 819-834.
 174. Zakharevich, N. V., Osolodkin, D. I., Artamonova, I. I., Palyulin, V. A., Zefirov, N. S., and Danilenko, V. N. (2012). Signatures of the ATP-binding pocket as a basis for structural classification of the serine/threonine protein kinases of gram-positive bacteria. Proteins: structure, function, and bioinformatics, 80(5), 1363-1376.
 175. Zelus, D., Robinson-Rechavi, M., Delacre, M., Auriault, C., and Laudet, V. (2000). Fast evolution of interleukin-2 in mammals and positive selection in ruminants. Journal of molecular evolution, 51(3), 234-244.
 176. Zhang, J. (2008). Positive selection, not negative selection, in the pseudogenization of *rcaA* in *Yersinia pestis*. Proceedings of the national academy of sciences, 105(42), E69-E69.
 177. Zhang, J., Kumar, S., and Nei, M. (1997). Small-sample tests of episodic adaptive evolution: a case study of primate lysozymes. Molecular biology and evolution, 14(12), 1335-1338.
 178. Zhang, J., Nielsen, R., and Yang, Z. (2005). Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Molecular biology and evolution, 22(12), 2472-2479.
 179. Zhang, J., Rosenberg, H. F., and Nei, M. (1998). Positive Darwinian selection after gene duplication in primate ribonuclease genes. Proceedings of the national academy of sciences, 95(7), 3708-3713.

ПРИЛОЖЕНИЯ

Приложение 1.

Таблица 1. Список используемых геномных последовательностей для анализа филогенетических профилей генов представителей различных видов рода *Bifidobacterium*.

номер сборки	название вида	название подвида	название штамма	статус сборки	обозначение
GCF_001263395.1	<i>Bifidobacterium actinocoloniiforme</i>		DSM 22766	complete	GCF_001263395.1_B_actinocoloniiforme_DSM_22766
GCF_000737885.1	<i>Bifidobacterium adolescentis</i>		22L	complete	GCF_000737885.1_B_adolescentis_22L
GCF_000010425.1	<i>Bifidobacterium adolescentis</i>		ATCC 15703	complete	GCF_000010425.1_B_adolescentis_ATCC_15703
GCF_000817995.1	<i>Bifidobacterium adolescentis</i>		BBMN23	complete	GCF_000817995.1_B_adolescentis_BBMN23
GCF_001417815.1	<i>Bifidobacterium aesculapii</i>		DSM 26737	draft	GCF_001417815.1_B_aesculapii_DSM_26737
GCF_000966445.2	<i>Bifidobacterium angulatum</i>		GT102	complete	GCF_000966445.2_B_angulatum_GT102
GCF_001025155.1	<i>Bifidobacterium angulatum</i>		JCM 7096	complete	GCF_001025155.1_B_angulatum_JCM_7096
GCF_000260715.1	<i>Bifidobacterium animalis</i>	<i>animalis</i>	ATCC 25527	complete	GCF_000260715.1_B_animalis_animalis_ATCC_25527
GCF_001688645.2	<i>Bifidobacterium animalis</i>	<i>animalis</i>	YL2	complete	GCF_001688645.2_B_animalis_animalis_YL2
GCF_000021425.1	<i>Bifidobacterium animalis</i>	<i>lactis</i>	AD011	complete	GCF_000021425.1_B_animalis_lactis_AD011
GCF_000471945.1	<i>Bifidobacterium animalis</i>	<i>lactis</i>	ATCC 27673	complete	GCF_000471945.1_B_animalis_lactis_ATCC_27673
GCF_000277325.1	<i>Bifidobacterium animalis</i>	<i>lactis</i>	B420	complete	GCF_000277325.1_B_animalis_lactis_B420
GCF_000025245.1	<i>Bifidobacterium animalis</i>	<i>lactis</i>	BB-12	complete	GCF_000025245.1_B_animalis_lactis_BB-12
GCF_000818055.1	<i>Bifidobacterium animalis</i>	<i>lactis</i>	BF052	complete	GCF_000818055.1_B_animalis_lactis_BF052
GCF_000277345.1	<i>Bifidobacterium animalis</i>	<i>lactis</i>	Bi-07	complete	GCF_000277345.1_B_animalis_lactis_Bi-07
GCF_000022705.1	<i>Bifidobacterium animalis</i>	<i>lactis</i>	BI-04	complete	GCF_000022705.1_B_animalis_lactis_BI-04

GCF_000414215.1	<i>Bifidobacterium animalis</i>	<i>lactis</i>	BI12	complete	GCF_000414215.1_B_animalis_lac_BI12
GCF_000224965.2	<i>Bifidobacterium animalis</i>	<i>lactis</i>	BLC1	complete	GCF_000224965.2_B_animalis_lac_BLC1
GCF_000220885.1	<i>Bifidobacterium animalis</i>	<i>lactis</i>	CNCM I-2494	complete	GCF_000220885.1_B_animalis_lac_CNCM_I-2494
GCF_000022965.1	<i>Bifidobacterium animalis</i>	<i>lactis</i>	DSM 10140	complete	GCF_000022965.1_B_animalis_lac_DSM_10140
GCF_000816205.1	<i>Bifidobacterium animalis</i>	<i>lactis</i>	KLDS2.0603	complete	GCF_000816205.1_B_animalis_lac_KLDS2.0603
GCF_000092765.1	<i>Bifidobacterium animalis</i>	<i>lactis</i>	V9	complete	GCF_000092765.1_B_animalis_lac_V9
GCF_000817045.1	<i>Bifidobacterium animalis</i>		A6	complete	GCF_000817045.1_B_animalis_A6
GCF_002220485.1	<i>Bifidobacterium animalis</i>		BL3	complete	GCF_002220485.1_B_animalis_BL3
GCF_000695895.1	<i>Bifidobacterium animalis</i>		RH	complete	GCF_000695895.1_B_animalis_RH
GCF_002259795.1	<i>Bifidobacterium aquikefiri</i>		LMG 28769	draft	GCF_002259795.1_B_aquikefiri_LMG_28769
GCF_002715865.1	<i>Bifidobacterium asteroides</i>		DSM 20089	complete	GCF_002715865.1_B_asteroides_DSM_20089
GCF_000304215.1	<i>Bifidobacterium asteroides</i>		PRL2011	complete	GCF_000304215.1_B_asteroides_PRL2011
GCF_000741165.1	<i>Bifidobacterium biavatii</i>		DSM 23969	draft	GCF_000741165.1_B_biavatii_DSM_23969
GCF_001281345.1	<i>Bifidobacterium bifidum</i>		BF3	complete	GCF_001281345.1_B_bifidum_BF3
GCF_000265095.1	<i>Bifidobacterium bifidum</i>		BGN4	complete	GCF_000265095.1_B_bifidum_BGN4
GCF_001025135.1	<i>Bifidobacterium bifidum</i>		JCM 1255	complete	GCF_001025135.1_B_bifidum_JCM_1255
GCF_000165905.1	<i>Bifidobacterium bifidum</i>		PRL2010	complete	GCF_000165905.1_B_bifidum_PRL2010
GCF_000164965.1	<i>Bifidobacterium bifidum</i>		S17	complete	GCF_000164965.1_B_bifidum_S17
GCF_000741525.1	<i>Bifidobacterium bohemicum</i>		DSM 22767	draft	GCF_000741525.1_B_bohemicum_DSM_22767
GCF_000737845.1	<i>Bifidobacterium bombi</i>		DSM 19703	draft	GCF_000737845.1_B_bombi_DSM_19703
GCF_000741535.1	<i>Bifidobacterium boum</i>		LMG 10736	draft	GCF_000741535.1_B_boum_LMG_10736
GCF_000568955.1	<i>Bifidobacterium breve</i>		12L	complete	GCF_000568955.1_B_breve_12L

GCF_000569055.1	<i>Bifidobacterium breve</i>		689b	complete	GCF_000569055.1_B_breve_689b
GCF_000213865.1	<i>Bifidobacterium breve</i>		ACS-071-V-Sch8b	complete	GCF_000213865.1_B_breve_ACS-071-V-Sch8b
GCF_001281425.1	<i>Bifidobacterium breve</i>		BR3	complete	GCF_001281425.1_B_breve_BR3
GCF_001025175.1	<i>Bifidobacterium breve</i>		JCM 1192	complete	GCF_001025175.1_B_breve_JCM_1192
GCF_000568975.1	<i>Bifidobacterium breve</i>		JCM 7017	complete	GCF_000568975.1_B_breve_JCM_7017
GCF_000569015.1	<i>Bifidobacterium breve</i>		JCM 7019	complete	GCF_000569015.1_B_breve_JCM_7019
GCF_001990225.1	<i>Bifidobacterium breve</i>		LMC520	complete	GCF_001990225.1_B_breve_LMC520
GCF_000569035.1	<i>Bifidobacterium breve</i>		NCFB 2258	complete	GCF_000569035.1_B_breve_NCFB_2258
GCF_000569075.1	<i>Bifidobacterium breve</i>		S27	complete	GCF_000569075.1_B_breve_S27
GCF_000220135.1	<i>Bifidobacterium breve</i>		UCC2003	complete	GCF_000220135.1_B_breve_UCC2003
GCF_000741175.1	<i>Bifidobacterium callitrichos</i>		DSM 23973	draft	GCF_000741175.1_B_callitrichos_DSM_23973
GCF_001025195.1	<i>Bifidobacterium catenulatum</i>		JCM 1194	complete	GCF_001025195.1_B_catenulatum_JCM_1194
GCF_002761235.1	<i>Bifidobacterium choerinum</i>		FMB-1	complete	GCF_002761235.1_B_choerinum_FMB-1
GCF_900094885.1	<i>Bifidobacterium commune</i>		R-52791	draft	GCF_900094885.1_B_communne_R-52791
GCF_000737865.1	<i>Bifidobacterium coryneforme</i>		LMG18911	complete	GCF_000737865.1_B_coryneforme_LMG18911
GCF_000738005.1	<i>Bifidobacterium crudilactis</i>		LMG 23609	draft	GCF_000738005.1_B_crudilactis_LMG_23609
GCF_000741575.1	<i>Bifidobacterium cuniculi</i>		LMG 10738	draft	GCF_000741575.1_B_cuniculi_LMG_10738
GCF_000024445.1	<i>Bifidobacterium dentium</i>		Bd1	complete	GCF_000024445.1_B_dentium_Bd1
GCF_001042595.1	<i>Bifidobacterium dentium</i>		JCM 1195	complete	GCF_001042595.1_B_dentium_JCM_1195
GCF_002259685.1	<i>Bifidobacterium eulemuris</i>		DSM 100216	draft	GCF_002259685.1_B_eulemuris_DSM_100216
GCF_000741205.1	<i>Bifidobacterium gallicum</i>		LMG 11596	draft	GCF_000741205.1_B_gallicum_LMG_11596
GCF_000741215.1	<i>Bifidobacterium gallinarum</i>		LMG 11586	draft	GCF_000741215.1_B_gallinarum_LMG_11586

GCF_002259755.1	<i>Bifidobacterium hapali</i>		DSM 100202	draft	GCF_002259755.1_B_hapali_DSM_100202
GCF_000706765.1	<i>Bifidobacterium indicum</i>		LMG 11587	complete	GCF_000706765.1_B_indicum_LMG_11587
GCF_001042615.1	<i>Bifidobacterium kashiwanohense</i>		JCM 15439	complete	GCF_001042615.1_B_kashiwanohense_JCM_15439
GCF_000800455.1	<i>Bifidobacterium kashiwanohense</i>		PV20-2	complete	GCF_000800455.1_B_kashiwanohense_PV20-2
GCF_001895165.1	<i>Bifidobacterium lemurum</i>		DSM 28807	draft	GCF_001895165.1_B_lemurum_DSM_28807
GCF_000196575.1	<i>Bifidobacterium longum</i>	<i>infantis</i>	157F	complete	GCF_000196575.1_B_longum_inf_157F
GCF_000020425.1	<i>Bifidobacterium longum</i>	<i>infantis</i>	ATCC 15697	complete	GCF_000020425.1_B_longum_inf_ATCC_15697
GCF_001281305.1	<i>Bifidobacterium longum</i>	<i>infantis</i>	BT1	complete	GCF_001281305.1_B_longum_inf_BT1
GCF_001051015.2	<i>Bifidobacterium longum</i>	<i>infantis</i>	CECT 7210	complete	GCF_001051015.2_B_longum_inf_CECT_7210
GCF_000269965.1	<i>Bifidobacterium longum</i>	<i>infantis</i>	JCM 1222	complete	GCF_000269965.1_B_longum_inf_JCM_1222
GCF_001725985.1	<i>Bifidobacterium longum</i>	<i>longum</i>	AH1206	complete	GCF_001725985.1_B_longum_lon_AH1206
GCF_000166315.1	<i>Bifidobacterium longum</i>	<i>longum</i>	BBMN68	complete	GCF_000166315.1_B_longum_lon_BBMN68
GCF_001446275.1	<i>Bifidobacterium longum</i>	<i>longum</i>	CCUG30698	complete	GCF_001446275.1_B_longum_lon_CCUG30698
GCF_000772485.1	<i>Bifidobacterium longum</i>	<i>longum</i>	GT15	complete	GCF_000772485.1_B_longum_lon_GT15
GCF_000196555.1	<i>Bifidobacterium longum</i>	<i>longum</i>	JCM 1217	complete	GCF_000196555.1_B_longum_lon_JCM_1217
GCF_000092325.1	<i>Bifidobacterium longum</i>	<i>longum</i>	JDM301	complete	GCF_000092325.1_B_longum_lon_JDM301
GCF_000219455.1	<i>Bifidobacterium longum</i>	<i>longum</i>	KACC 91563	complete	GCF_000219455.1_B_longum_lon_KACC_91563
GCF_001446255.1	<i>Bifidobacterium longum</i>	<i>longum</i>	NCIMB8809	complete	GCF_001446255.1_B_longum_lon_NCIMB8809
GCF_001719085.1	<i>Bifidobacterium longum</i>		35624	complete	GCF_001719085.1_B_longum_35624
GCF_000829295.1	<i>Bifidobacterium longum</i>		105-A	complete	GCF_000829295.1_B_longum_105-A
GCF_001293145.1	<i>Bifidobacterium longum</i>		BG7	complete	GCF_001293145.1_B_longum_BG7
GCF_000730205.1	<i>Bifidobacterium longum</i>		BXY01	complete	GCF_000730205.1_B_longum_BXY01

GCF_000008945.1	<i>Bifidobacterium longum</i>		DJO10A	complete	GCF_000008945.1_B_longum_DJO10A
GCF_000007525.1	<i>Bifidobacterium longum</i>		NCC2705	complete	GCF_000007525.1_B_longum_NCC2705
GCF_000741255.1	<i>Bifidobacterium magnum</i>		LMG 11591	draft	GCF_000741255.1_B_magnum_LMG_11591
GCF_900129045.1	<i>Bifidobacterium merycicum</i>		DSM 6492	draft	GCF_900129045.1_B_merycicum_DSM_6492
GCF_000741645.1	<i>Bifidobacterium minimum</i>		LMG 11592	draft	GCF_000741645.1_B_minimum_LMG_11592
GCF_000741285.1	<i>Bifidobacterium mongoliense</i>		DSM 21395	draft	GCF_000741285.1_B_mongoliense_DSM_21395
GCF_000522505.1	<i>Bifidobacterium moukalabense</i>		DSM 27321	draft	GCF_000522505.1_B_moukalabense_DSM_27321
GCF_002259745.1	<i>Bifidobacterium myosotis</i>		DSM 100196	draft	GCF_002259745.1_B_myosotis_DSM_100196
GCF_001025215.1	<i>Bifidobacterium pseudocatenulatum</i>		JCM 1200	complete	GCF_001025215.1_B_pseudocatenulatum_JCM_1200
GCF_002706665.1	<i>Bifidobacterium pseudolongum</i>	<i>globosum</i>	DSM 20092	complete	GCF_002706665.1_B_pseudolongum_glo_DSM_20092
GCF_000800475.2	<i>Bifidobacterium pseudolongum</i>		PV8-2	complete	GCF_000800475.2_B_pseudolongum_PV8-2
GCF_002282915.1	<i>Bifidobacterium pseudolongum</i>		UMB-MBP-01	complete	GCF_002282915.1_B_pseudolongum_UMB-MBP-01
GCF_000741705.1	<i>Bifidobacterium psychraerophilum</i>		LMG 21775	draft	GCF_000741705.1_B_psychraerophilum_LMG_21775
GCF_000741335.1	<i>Bifidobacterium pullorum</i>		LMG 21816	draft	GCF_000741335.1_B_pullorum_LMG_21816
GCF_000741695.1	<i>Bifidobacterium reuteri</i>		DSM 23975	draft	GCF_000741695.1_B_reuteri_DSM_23975
GCF_000741365.1	<i>Bifidobacterium ruminantium</i>		LMG 21811	draft	GCF_000741365.1_B_ruminantium_LMG_21811
GCF_000741375.1	<i>Bifidobacterium saeculare</i>		LMG 14934	draft	GCF_000741375.1_B_saeculare_LMG_14934
GCF_000741715.1	<i>Bifidobacterium saguini</i>		DSM 23967	draft	GCF_000741715.1_B_saguini_DSM_23967
GCF_001042635.1	<i>Bifidobacterium scardovii</i>		JCM 12489	complete	GCF_001042635.1_B_scardovii_JCM_12489
GCF_000741785.1	<i>Bifidobacterium stellenboschense</i>		DSM 23968	draft	GCF_000741785.1_B_stellenboschense_DSM_23968
GCF_000741775.1	<i>Bifidobacterium subtile</i>		LMG 11597	draft	GCF_000741775.1_B_subtile_LMG_11597
GCF_000741445.1	<i>Bifidobacterium thermacidophilum</i>	<i>porcinum</i>	LMG 21689	draft	GCF_000741445.1_B_thermacidophilum_por_LMG_2168

GCF_000741455.1	<i>Bifidobacterium thermacidophilum</i>	<i>thermacidophilum</i>	LMG 21395	draft	GCF_000741455.1_B_thermacidophilum_the_LMG_21395
GCF_000347695.1	<i>Bifidobacterium thermophilum</i>		RBL67	complete	GCF_000347695.1_B_thermophilum_RBL67
GCF_002259645.1	<i>Bifidobacterium tissieri</i>		DSM 100201	draft	GCF_002259645.1_B_tissieri_DSM_100201
GCF_000741765.1	<i>Bifidobacterium tsurumiense</i>		JCM 13495	draft	GCF_000741765.1_B_tsurumiense_JCM_13495
GCF_002234915.1	<i>Bifidobacterium vansinderenii</i>		Tam10B	draft	GCF_002234915.1_B_vansinderenii_Tam10B

Таблица 2. Список используемых геномных последовательностей для анализа молекулярной эволюции генов кластера PFNA представителей различных видов рода *Bifidobacterium*.

номер сборки	название вида	название подвида	название штамма	статус сборки	обозначение
GCF_001263395.1	<i>Bifidobacterium actinocoloniiiforme</i>		DSM 22766	complete	B_actinocoloniiiforme
GCF_000010425.1	<i>Bifidobacterium adolescentis</i>		ATCC 15703	complete	B_adolescentis
GCF_001417815.1	<i>Bifidobacterium aesculapii</i>		DSM 26737	draft	B_aesculapii
GCF_000966445.2	<i>Bifidobacterium angulatum</i>		GT102	complete	B_angulatum
GCF_000260715.1	<i>Bifidobacterium animalis</i>	<i>animalis</i>	ATCC 25527	complete	B_animalis_ani
GCF_000022965.1	<i>Bifidobacterium animalis</i>	<i>lactis</i>	DSM 10140	complete	B_animalis_lac
GCF_002715865.1	<i>Bifidobacterium asteroides</i>		DSM 20089	complete	B_asteroides
GCF_000741165.1	<i>Bifidobacterium biavatii</i>		DSM 23969	draft	B_biavatii
GCF_001025135.1	<i>Bifidobacterium bifidum</i>		JCM 1255	complete	B_bifidum
GCF_000741525.1	<i>Bifidobacterium bohemicum</i>		DSM 22767	draft	B_bohemicum
GCF_000737845.1	<i>Bifidobacterium bombi</i>		DSM 19703	draft	B_bombi
GCF_000741535.1	<i>Bifidobacterium boum</i>		LMG 10736	draft	B_boum
GCF_001025175.1	<i>Bifidobacterium breve</i>		JCM 1192	complete	B_breve

GCF_001025195.1	<i>Bifidobacterium catenulatum</i>		JCM 1194	complete	B_catenulatum
GCF_002761235.1	<i>Bifidobacterium choerinum</i>		FMB-1	complete	B_choerinum
GCF_900094885.1	<i>Bifidobacterium commune</i>		R-52791	draft	B_commune
GCF_000737865.1	<i>Bifidobacterium coryneforme</i>		LMG 18911	complete	B_coryneforme
GCF_000741575.1	<i>Bifidobacterium cuniculi</i>		LMG 10738	draft	B_cuniculi
GCF_001042595.1	<i>Bifidobacterium dentium</i>		JCM 1195	complete	B_dentium
GCF_002259685.1	<i>Bifidobacterium eulemuris</i>		DSM 100216	draft	B_eulemuris
GCF_000741205.1	<i>Bifidobacterium gallicum</i>		LMG 11596	draft	B_gallicum
GCF_002259755.1	<i>Bifidobacterium hapali</i>		DSM 100202	draft	B_hapali
GCF_000706765.1	<i>Bifidobacterium indicum</i>		LMG 11587	complete	B_indicum
GCF_001042615.1	<i>Bifidobacterium kashiwanohense</i>		JCM 15439	complete	B_kashiwanohense
GCF_001895165.1	<i>Bifidobacterium lemorum</i>		DSM 28807	draft	B_lemorum
GCF_001281305.1	<i>Bifidobacterium longum</i>	<i>infantis</i>	BT1	complete	B_longum_inf
GCF_000772485.1	<i>Bifidobacterium longum</i>	<i>longum</i>	GT15	complete	B_longum_lon
GCF_900129045.1	<i>Bifidobacterium merycicum</i>		DSM 6492	draft	B_merycicum
GCF_000522505.1	<i>Bifidobacterium moukalabense</i>		DSM 27321	draft	B_moukalabense
GCF_002259745.1	<i>Bifidobacterium myosotis</i>		DSM 100196	draft	B_myosotis
GCF_001025215.1	<i>Bifidobacterium pseudocatenulatum</i>		JCM 1200	complete	B_pseudocatenulatum
GCF_002706665.1	<i>Bifidobacterium pseudolongum</i>	<i>globosum</i>	DSM 20092	complete	B_pseudolongum_glo
GCF_000741325.1	<i>Bifidobacterium pseudolongum</i>	<i>pseudolongum</i>	LMG 11571	draft	B_pseudolongum_pse
GCF_000741695.1	<i>Bifidobacterium reuteri</i>		DSM 23975	draft	B_reuteri
GCF_000741365.1	<i>Bifidobacterium ruminantium</i>		LMG 21811	draft	B_ruminantium

GCF_000771625.1	<i>Bifidobacterium saguini</i>		DSM 23967	draft	B_saguini
GCF_001042635.1	<i>Bifidobacterium scardovii</i>		JCM 12489	complete	B_scardovii
GCF_000741785.1	<i>Bifidobacterium stellenboschense</i>		DSM 23968	draft	B_stellenboschense
GCF_000741445.1	<i>Bifidobacterium thermacidophilum</i>	<i>porcinum</i>	LMG 21689	draft	B_thermacidophilum_por
GCF_000741455.1	<i>Bifidobacterium thermacidophilum</i>	<i>thermacidophilum</i>	LMG 21395	draft	B_thermacidophilum_the
GCF_000347695.1	<i>Bifidobacterium thermophilum</i>		RBL67	complete	B_thermophilum
GCF_002259645.1	<i>Bifidobacterium tissieri</i>		DSM 100201	draft	B_tissieri
GCF_002234915.1	<i>Bifidobacterium vansinderenii</i>		Tam10B	draft	B_vansinderenii

Таблица 3. Локализация предсказанных сайтов под давлением эпизодического положительного отбора ($PP > 0,7$) в первичной структуре белков, кодируемых генами кластера PFNA представителей различных видов рода *Bifidobacterium*. Сайты, представляющие собой наиболее надежных кандидатов ($PP > 0,95$), отмечены звездочкой.

тестируемая ветвь/ клада филогенетического дерева	<i>pkb2</i>	<i>fn3</i>	<i>aaa-atp</i>	<i>duf58</i>	<i>tgm</i>
<i>B. actinocoloniiforme</i>	32 78 84* 128 169 196 276 510 512 550	146 152 237 286 288 326 362 371 381 383* 397* 401 420 427 454 464 465 470 474 482 502 505 516 532 540 546 572 580 587 592 603* 618* 620 623 634 636 649 656* 664 674* 677* 680 681 682* 689 707 727 732* 733 734 735* 751 753 776* 778 792* 796 800 805 815 824 832 845* 859 862 865 869* 884 893 896 899 910 917 926 948 951 955* 957 966 969 988* 1006 1015 1027 1034* 1039* 1040* 1043 1045 1054 1063 1065* 1076* 1086* 1096 1100 1109 1114 1120* 1121 1127* 1132 1141 1150 1156 1164* 1172* 1173* 1199 1201* 1209 1216 1219* 1233* 1234* 1245 1260 1291 1299 1301 1311 1325 1330 1343* 1346 1359 1403 1417* 1440 1466 1474 1481* 1507 1511 1530 1532* 1535 1578 1579* 1656 1719	153 172 183 198* 203 224 226 254 277 289* 309 310* 324 327 345 346 357 365 371* 379 400 411* 412 415 436 445	49 67 68* 77 80 81 101 110 117 131 143 153 159* 160 163 165 172* 183 189 190 198* 199* 227 237* 239 244* 259 268 274 275 285 286 297 299 338 355 367 388 390 391 393* 396* 397 406	84 107 118 120 139* 144 145 152 160 161 163 166 170* 181* 249 254 267* 270* 279* 298* 299 300* 301 302 308 316 318 323 329 330 340 349* 356 357 366* 384 397 400* 436 440 445 470 481 485 486* 546 554 557 568 590 623 627* 628* 629 632 635 636 637* 653 655 687 703 704 710 712 713 721 734 788
<i>B. aesculapii</i>		258 585 1509			134 136

<i>B. asteroides</i>	57 125 135* 256* 383	184 223* 421 457 463 472* 675 688 713 766 808 823 833 847 902 1091 1128 1151 1199 1287 1305 1339 1382 1522 1553	155 301 333 402*	100 130 131 132 158 198 258 262 266 281 288 349 364 370 371 398 400	140 204 211 294* 369 381 411 432 602 605 611 630 700 710
<i>B. biavatii</i>	41 74 85 116 117 131 197 239 243 249	137 218 290 306 407 426 448 529 638 666 712 738 778 817 888* 951 997 1055 1089* 1099 1196 1214 1277 1350 1439 1535 1552 1627 1725	199 219 269 271	43 45 56 69 78 99 106 113 123 139 163 226 296 342 353* 366 396 400	128* 130 137 162 186 191 276 283 309 323 371 421 496 505 549 581* 603 648 657 669 677 678 706 70
<i>B. bifidum</i>	120 246	260 315 355 1362 1589		80	90 144 171 606 716
<i>B. bohemicum</i>	231	360 535 549 659 690 1241 1501 1520	244 383	84 90 219 398	372 378 419 534 575 610
<i>B. bombi</i>	131 145 160 286	466 607 693* 903 1067 1281 1311 1334		112 129 199	76 82 118 175 196 230 247 413 416 428 450 621 646 685 729 739 743
<i>B. choerinum</i>	121* 130 166	316 377 706* 1649* 1650 1676		238* 395	
<i>B. commune</i>	156 285 286	177 391 774 824 912 1189	274 326	164 423	86 223 464 703
<i>B. coryneforme</i> , <i>B. indicum</i>	132	143 358 587 870 899 1052 1108 1555	97 382	69 151 371	569 662
<i>B. cuniculi</i>	90*	262 319 415* 513 594 598* 778 781 792 838* 867 967 1192 1245 1259 1276 1278* 1363 1433 1497 1522 1523 1542 1573 1715*		309	106* 126 269 312 347 419 421 597 635 745
<i>B. gallicum</i>		470 528 534 542 601 782 800 950 1044 1094 1370 1423 1465 1541 1622 1659		63	177 270 314
<i>B. hapali</i>	84 86 130 204 233 249 293 339	3 8 18 145 166 237 268 274 284 305 314 316 423* 430 442 446 465 468 469 475 481 488 518 533 551 577 591 592 620 623 626 670 691 692 716 718 772 778 808 875 881 886 912 934* 936 971 979* 1008 1031 1118 1187 1189 1201 1271 1297 1330 1366 1390 1392 1421	56 142 144 150 239 274 295 302 341 342 346	107 117 121 138 161 168 195 197 203 206 232 242 243* 252 282 295 319 330 387* 401 421 440 445	50 71 74 78 92 99 118 146 203 220 261 262 297 299 301 313 341 342 459 483 504 559* 565 621 628 679 693 700 705 741
<i>B. pseudolongum</i> subsp. <i>globosum</i>	133* 168	146 186 191 241 324 548 635* 849 853 950* 969 1055 1076 1089* 1130 1145 1251* 1256 1297* 1340 1397* 1419* 1442 1464 1496 1501 1541 1577	176* 242 322 362	137 238 369*	115* 129 167 209* 272 276 315 350 366 454 458 536
<i>B. scardovii</i>	72 96 103 119 149 411 455	127 446 506 549 698 744 779 821 850 1005 1021 1053 1063 1075* 1079 1086 1194 1202 1221 1243 1285 1295 1401 1479	125 267* 393	78 107 199 200 268 373	70 92 93* 139 168 235 274* 357 358 561 582 589 593 596*

Приложение 2.

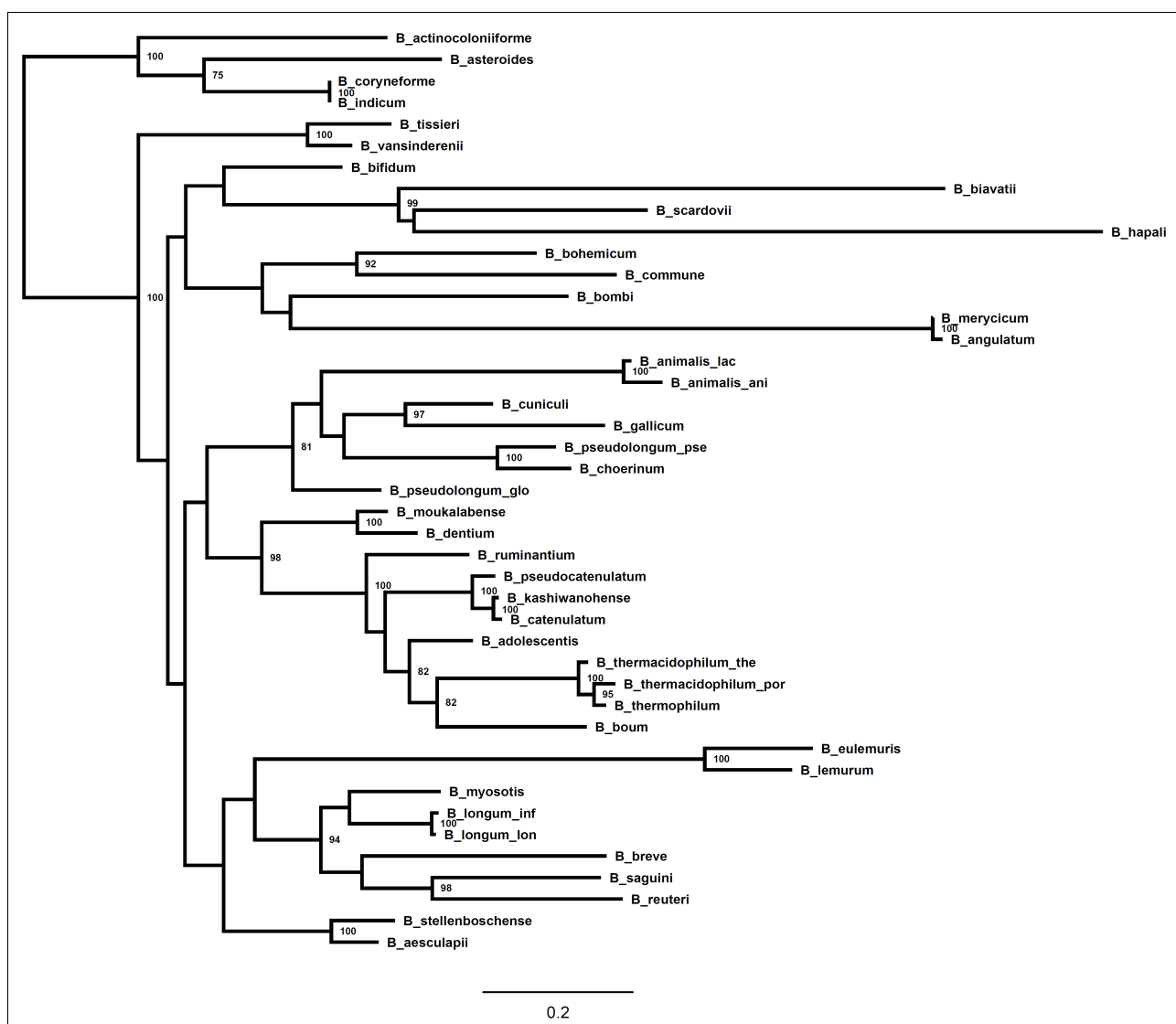


Рисунок 1. Филогенетического дерева, построенное на основе множественного выравнивания последовательностей гена *pkb2* кластера PFNA 43 представителей различных видов рода *Bifidobacterium*. Значения поддержек внутренних узлов филограммы получены после 10000 итераций бутстреп-анализа в программе RaxML. Поддержки в узлах филограммы со значением менее 70 на рисунке не представлены. Представлены оригинальные длины ветвей филогенетического дерева, масштабная линейка соответствует значению две предполагаемые замены на десять нуклеотидных сайтов.

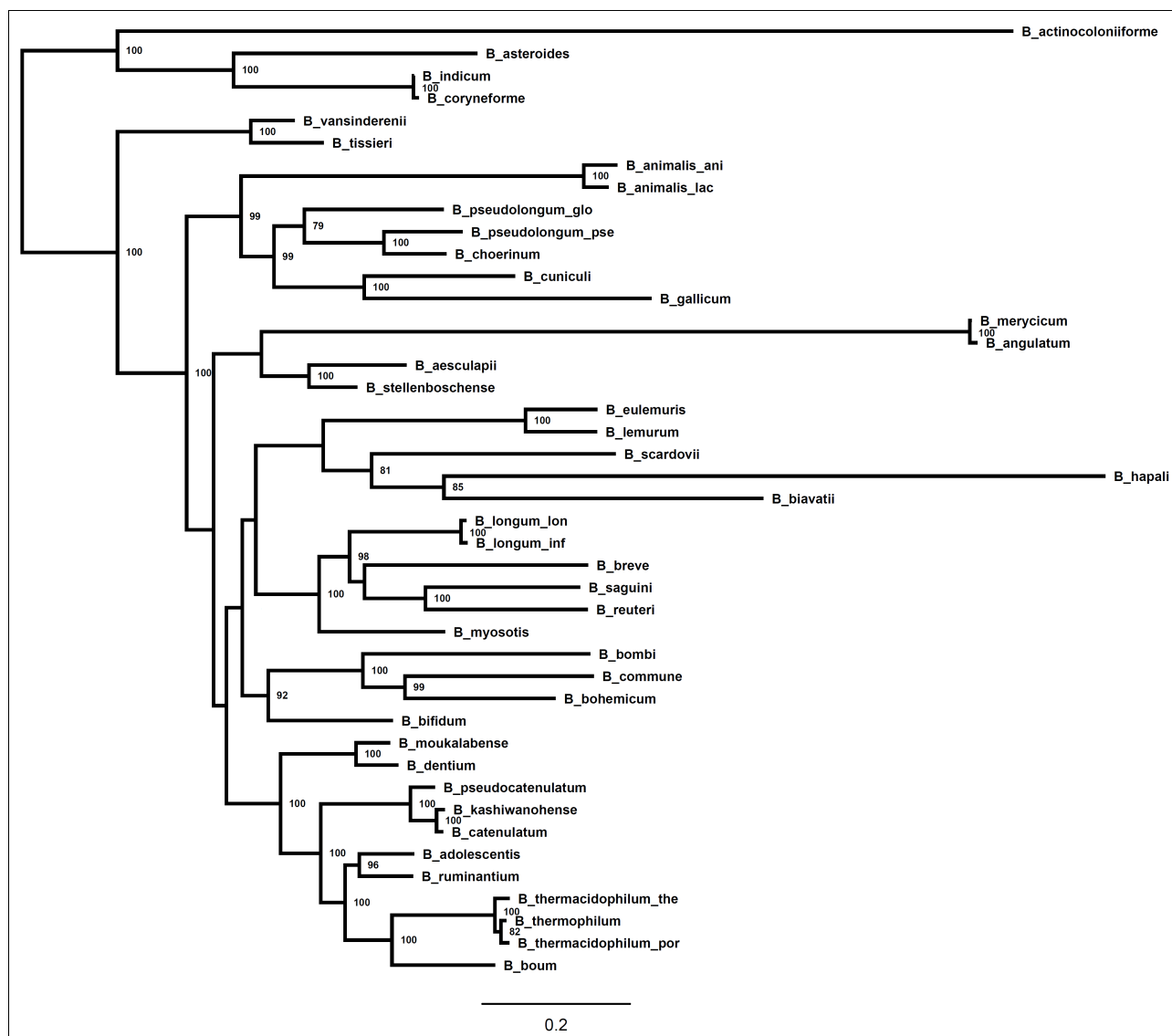


Рисунок 2. Филогенетического древо, построенное на основе множественного выравнивания последовательностей гена *fn3* кластера PFNA 43 представителей различных видов рода *Bifidobacterium*. Значения поддержек внутренних узлов филограммы получены после 10000 итераций бутстреп-анализа в программе RaxML. Поддержки в узлах филограммы со значением менее 70 на рисунке не представлены. Представлены оригинальные длины ветвей филогенетического древа, масштабная линейка соответствует значению две предполагаемые замены на десять нуклеотидных сайтов.

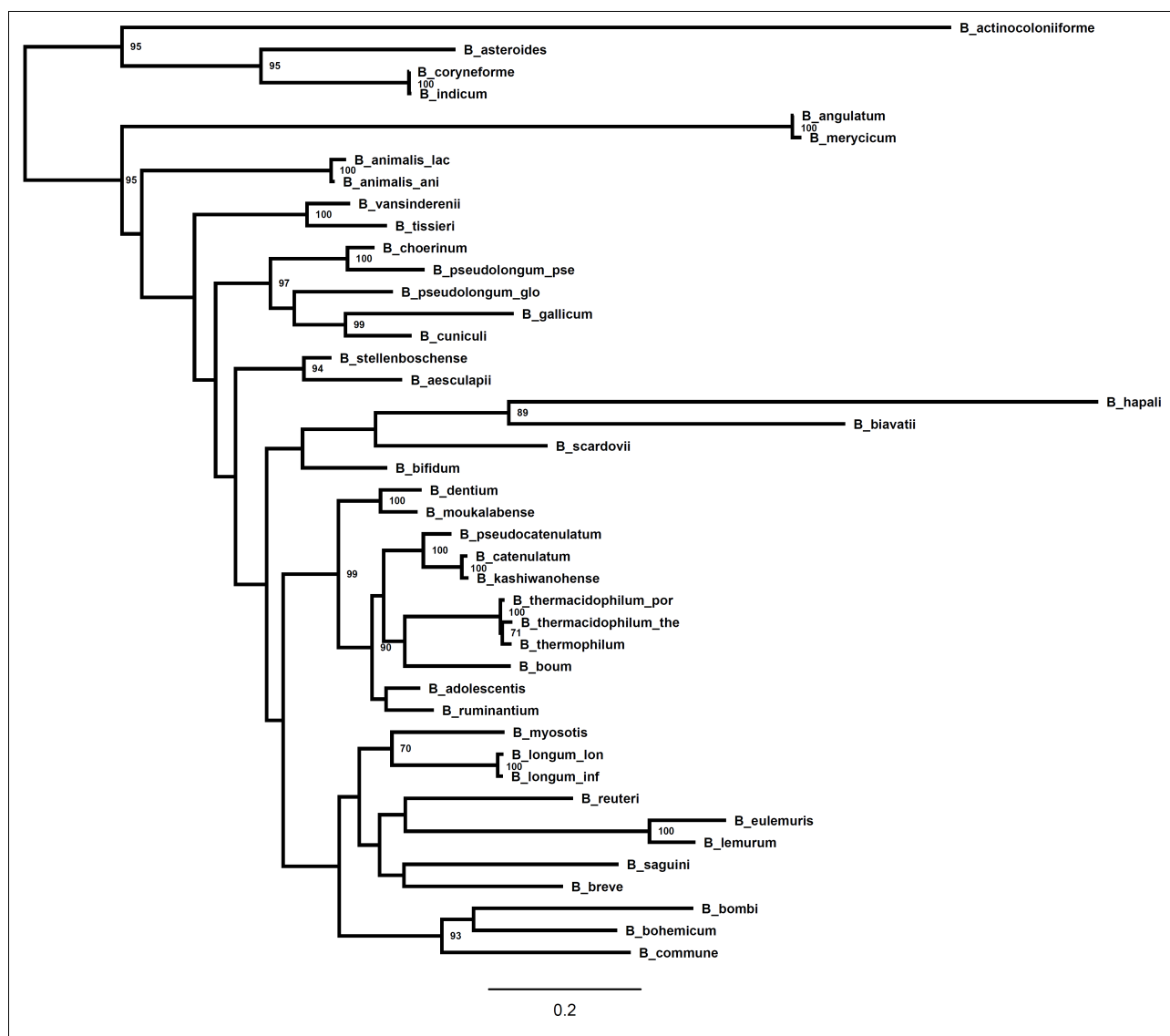


Рисунок 3. Филогенетического дерева, построенное на основе множественного выравнивания последовательностей гена *aaa-atp* кластера PFNA 43 представителей различных видов рода *Bifidobacterium*. Значения поддержек внутренних узлов филограммы получены после 10000 итераций бутстреп-анализа в программе RaхML. Поддержки в узлах филограммы со значением менее 70 на рисунке не представлены. Представлены оригинальные длины ветвей филогенетического дерева, масштабная линейка соответствует значению две предполагаемые замены на десять нуклеотидных сайтов.

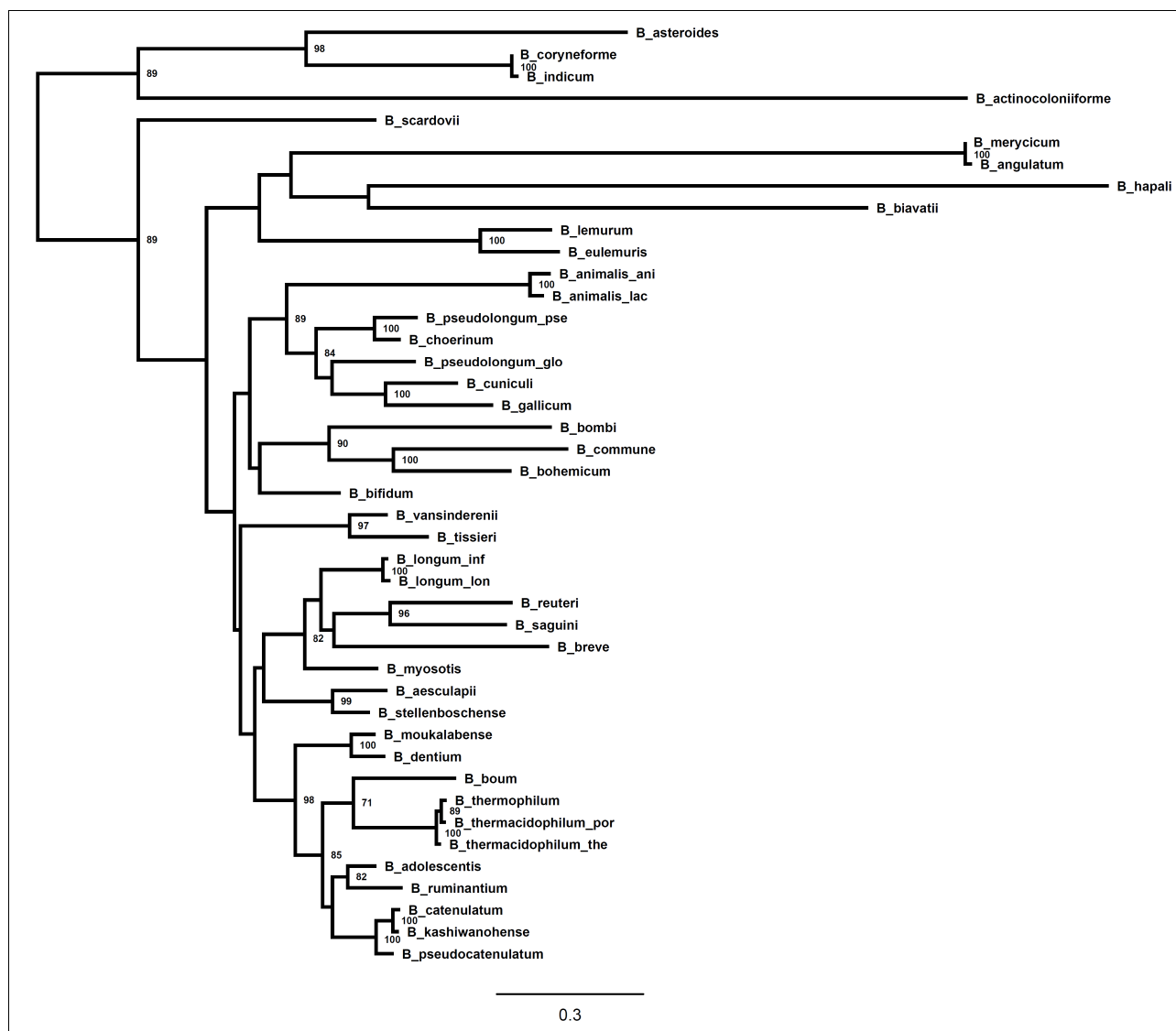


Рисунок 4. Филогенетического древо, построенное на основе множественного выравнивания последовательностей гена *duf58* кластера PFNA 43 представителей различных видов рода *Bifidobacterium*. Значения поддержек внутренних узлов филограммы получены после 10000 итераций бутстреп-анализа в программе RaxML. Поддержки в узлах филограммы со значением менее 70 на рисунке не представлены. Представлены оригинальные длины ветвей филогенетического древа, масштабная линейка соответствует значению три предполагаемые замены на десять нуклеотидных сайтов.

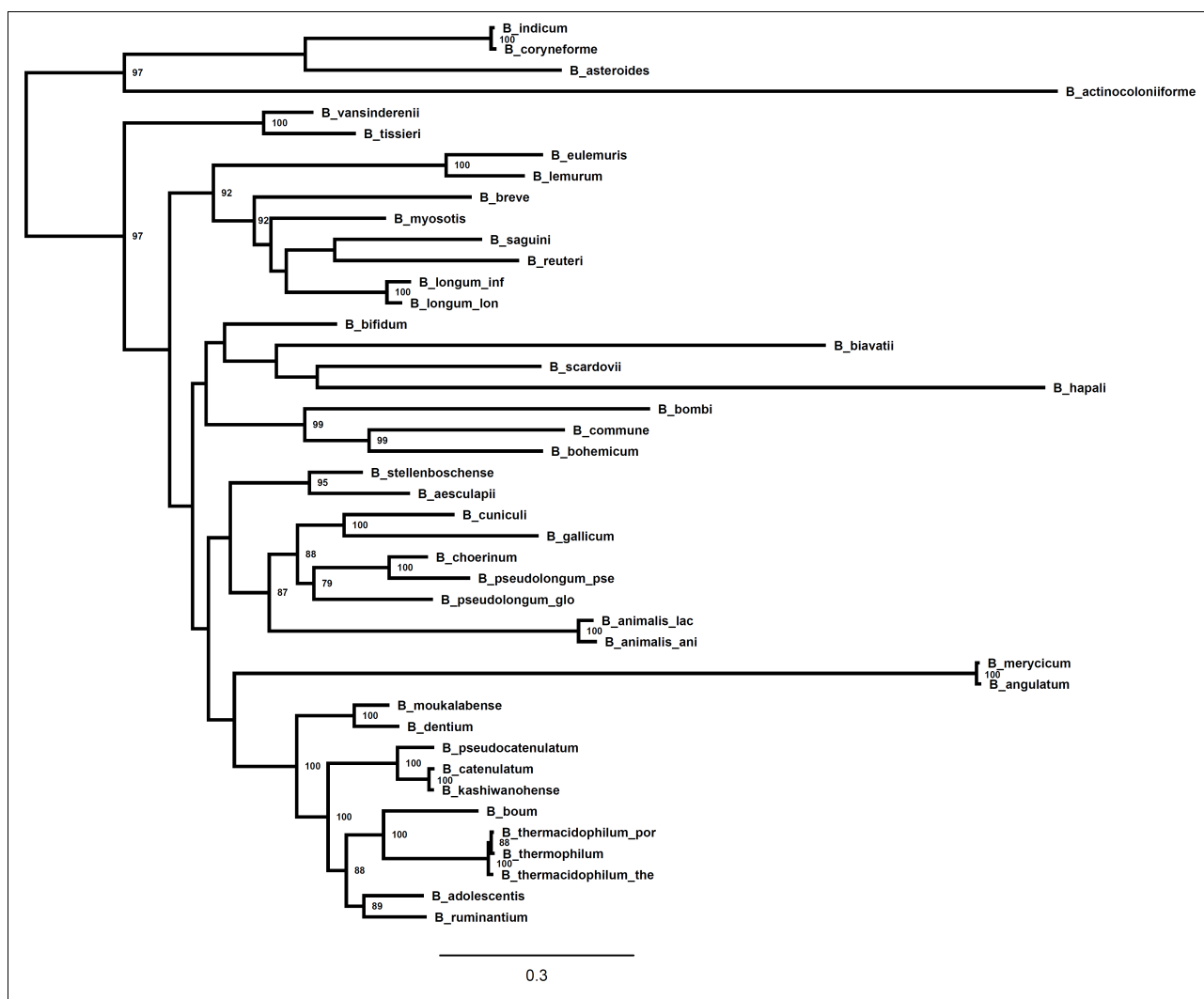


Рисунок 5. Филогенетического древо, построенное на основе множественного выравнивания последовательностей гена *tgm* кластера PFNA 43 представителей различных видов рода *Bifidobacterium*. Значения поддержек внутренних узлов филограммы получены после 10000 итераций бутстреп-анализа в программе RaxML. Поддержки в узлах филограммы со значением менее 70 на рисунке не представлены. Представлены оригинальные длины ветвей филогенетического древа, масштабная линейка соответствует значению три предполагаемые замены на десять нуклеотидных сайтов.