

Муниципальное бюджетное общеобразовательное учреждение «Лицей
современных технологий управления № 2» г. Пензы.

Индивидуальный проект по теме:
**«Создание прайм-системы для определения
детекции бактериофагов на промышленном
этапе производства кисломолочных
продуктов»**

Автор проекта:
обучающаяся 10 класса
Суслина Софья

Руководитель проекта:
К.б.н., доцент
Золина Н. Ф.

г. Пенза, 2023 г.

Оглавление

Введение.....	3
ПЦР-тестирование: что это такое, принцип метода и варианты его проведения.....	5
Нахождение генома образцов	6
Выравнивание.....	6
Филогения.....	7
Характеристика праймеров	8
Поиск праймеров и создание прайм-системы	10
Выводы.....	11
Список использованных источников	12
Приложения	13

Введение

Бактериофаги (-фаги) – это распространённый вид вирусов, заражающих бактерии. Они являются облигатными паразитами, т.е. существование бактериофага вне живого организма невозможно. Большая часть цикла репродукции вирусов заканчивается лизисом и высвобождением новых вирионов, способных инфицировать соседние клетки. Это многочисленная группа (к примеру, в одной капле морской воды может содержаться до 10⁸ фагов), которая играет одну из ключевых ролей – поддержание баланса бактериальной популяции, посредством сокращения численности видов. В медицине умение бактериофагов бороться с болезнетворными бактериями активно используется, однако, в условиях пищевой промышленности фаговая атака приводит к огромным потерям, а главная проблема заключается в трудности борьбы нею. Бактериофаги, как и многие другие вирусы способны к быстрым и многочисленным мутациям, но, кроме этого, у них имеется система защиты, которая может адаптироваться под тот тип бактерии, которую фаг заражает.

Кисломолочные продукты оказывают огромное влияние на здоровье человека. Из них люди получают кальций, незаменимые аминокислоты, витамины, а также лакто- и бифидобактерии. Для их получения в пищевом производстве используются молочнокислые бактерии. Однако, в процессе производства заражение бактериофагами может привести к ряду проблем. Например, таких как увеличение потерь в производстве, уменьшение качества продукции, усложнение внедрения новых штаммов с улучшенными свойствами, а также сложности в обработке заражённых культур.

Кроме того, актуальность данной проблемы возросла в последние годы из-за ухода с рынка многих компаний, предоставляющих услуги по культурам бактериофагов и соответствующего оборудования по сохранению чистых полезных человеку культур. Детекция различных патогенных бактериофагов снижает качество молочной продукции.

Для борьбы с ними учёные разработали множество способов обнаружения. Один из наиболее часто используемых непосредственно на производстве – метод ПЦР-тестирования. Для его проведения требуются специфичные праймеры, подбираемые генетиками.

Главной целью данного проекта является создание теоретически рабочей системы праймеров для выявления бактериофагов в заквасках с помощью ПЦР-тестирования.

Для достижения поставленной цели были сформулированы следующие задачи:

- Изучение ПЦР-тестирования, его видов и выбор наиболее эффективного;
- Нахождение генома бактериофагов, выделенных из заквасок на производстве;
- Подбор праймеров для проведения ПЦР-тестирования позволяющего обнаружить вирусы с использованием филогении;
- Создание специфичной прайм-системы.

В результате реализации поставленных задач должна получиться прайм-система, с помощью которой можно определить наиболее часто встречающихся видов бактериофагов. Будущее внедрение этой системы обеспечит уменьшение затрат на ПЦР, реактивы и время, необходимое для выявления вирусов без готовых пар праймеров, а также обеспечит более качественное изготовление кисломолочных продуктов питания и продуктов, использующих их в своём производстве. В перспективе это улучшит здоровье населения и сэкономит ресурсы, которые могут быть потрачены на другие важные сферы жизни.

На просторах интернета данная тема не поднималась, однако существует множество сайтов, описывающих данную проблему без реализации её решения. В реализации проекта участвуют определённые геномы, которые были реально получены из заражённых заквасок.

Материалы: геномы бактериофагов, полученных из международной библиотеки молекулярной биологии NCBI.

Методы: работа с программами Mega, Ugene и сайтами NCBI, OligoCalc (ОлигоКалькулятор).

ПЦР-тестирование: что это такое, принцип метода и варианты его проведения

ПЦР (Полимеразная цепная реакция) – это один из методов молекулярной биологии, благодаря которому можно создать копии определённого фрагмента ДНК с использованием исходного образца, повысив его содержание в пробе.

ПЦР состоит из трёх основных этапов:

1. Денатурация.

Пробу нагревают и ДНК-молекулы распадаются на две нити. Если рабочим материалом является РНК (вместо ДНК), то перед денатурацией проводится ПЦР с обратной транскрипцией, который позволяет превратить молекулу РНК в ДНК.

2. Отжиг праймеров.

Температуру снижают, и если в пробе присутствует целевая ДНК, то праймеры присоединяются к концам двух нитей (праймеры – это небольшие фрагменты одноцепочной ДНК).

3. Элонгация.

Далее при повторном нагреве активируется фермент ДНК-полимераза и начинается достраивание второй нити ДНК (началом служит праймер). Для восстановления цепи ДНК-полимераза использует свободные нуклеотиды.

Вариантов проведения ПЦР существует огромное множество. Для наглядности рассмотрим несколько:

- ПЦР с горячим стартом (hot-start PCR).
В образец добавляется полимеразы с блокирующими её активность антителами. Благодаря этому свойству уменьшается присутствие нежелательных продуктов, образовавшихся за счёт неспецифичной амплификации ДНК при обычных температурах. ПЦР с горячим стартом используется в тех случаях, когда образцов много и некоторые будут стоять уже с добавленной полимеразой.
- «Холодная» ПЦР (COLD-PCR).
Используется двухступенчатый цикл температур, позволяющих точнее определить специфичность или неспецифичность амплификации. Благодаря «холодной» ПЦР определяется однонуклеотидная мутация. Принцип заключается в обнаружении минимальных изменений температуры при модификации одного нуклеотида на другой.
- Ступенчатая ПЦР (touchdown PCR).
При увеличении числа циклов – меняется температура, пока она не будет соответствовать температурам последующих циклов. При этом вероятность неспецифичной амплификации уменьшается, так как праймеры будут легче связываться с ДНК.
- Мультиплексная ПЦР (multiplex PCR).
В одну пробирку с ДНК помещают сразу несколько праймеров. Используется при определении нескольких патогенов, при этом экономятся время и реактивы. При этом t-плавления, ГЦ-состав, количество нуклеотидов и длина ампликона не должны сильно различаться.

Нахождение генома образцов

Для реализации будущих праймеров необходимы реальные образцы фагов, полученных из заражённых заквасок. Для этого были проанализированы виды молочнокислых бактерий и выяснилось, что класс *Lactobacillus* не только один из самых распространённых, но и активно используемых в промышленности.

Далее необходимы были геномы бактериофагов. Для их нахождения использовался сайт NCBI (NCBI – это сайт, в котором хранится общедоступная база данных молекулярной биологии).

В результате были отобраны 10 последовательностей заражённых фагами видов. Из них 7 одного вида: *Lactobacillus phage c5*, *Lactobacillus phage Ld3*, *Lactobacillus phage LL-Ku*, *Lactobacillus phage phiLdb*, *Lactobacillus phage Silenus*, *Lactobacillus phage Bassarid*, *Lactobacillus phage CL1*. А также 3 отличных: *Pediococcus phage clP1*, *Bacteriophage phage SPP1*, *Oenococcus phage phiOE33PA*.

Их геномы были скачены в варианте FAST. FAST – это способ записи текстовых данных, распознаваемый самыми разными текстовыми редакторами, а также любым программами для работы с геномами.

Теперь, когда геномы были получены необходимо выравнить их для дальнейшего анализа и работы.

Выравнивание

Выравнивание необходимо для соотношения одинаковых нуклеотидов разных последовательностей. Так как наши геномы имеют очень большие размеры (от 20 000 до 40 000), то онлайн-сайты не смогут обработать такое огромное количество информации, поэтому было решено воспользоваться программой Mega для выравнивания и филогении.

В программу Mega встроены 2 алгоритма выравнивания – MUSCLE и ClustalW. Это одни из самых распространённых и эффективных алгоритма. Они оба способны к выравниванию больших геномов, но для дальнейшей работы нужно использовать один. Сравним их:

1. MUSCLE

Работает с большим количеством последовательностей достаточно быстро, но качество может быть не очень хорошим, особенно при выравнивании дальнородственных последовательностей. Имеет ограниченное количество настроек.

2. ClustalW

Работает медленнее, но зачастую качественнее, особенно при наличии дальнородственных связей. Также он обладает большей гибкостью настроек (например, матрица замен, каркасные последовательности и т.д.).

Для выравнивания были опробованы оба варианта, но лучшим оказался ClustalW, так как выравнивание почти сразу началось с двух строк последовательности (в Mega 2 строки последовательностей начали идти одновременно только с 200 нуклеотида), а также образовалось гораздо меньше гэпов (*).

*Хорошим можно назвать то выравнивание, которое имеет меньший вес. Вес увеличивается за счёт разных штрафов. Одним из штрафов являются гэпы. Гэпы – это пропуски, обозначаемые горизонтальными чертами в белом квадрате. Алгоритм создаёт гэпы, если считает, что сдвиг последовательности на одну или нескольких пар нуклеотидов позволит выравнить большее количество строчек, нежели при варианте, где сдвига не было. При выравнивании желательно

избегать гэпов, но зачастую – это невозможно. В таких случаях, необходимо минимизировать их количество.

Затем при просмотре готового выравнивания были обнаружены множественные участки совпадающих нуклеотидов. Это означает, что некоторые из выбранных видов являются близкородственными и к ним возможен подбор общей пары праймеров.

Для определения родства воспользуемся филогенетическим деревом.

Филогения

Деревья, построенные методами филогении, разделены на составляющие части, каждая из которых соответствует определённым терминам, обозначающим их:

- Листья – название группы или единицы организмов, чьё родство анализируется.
- Клада – группа организмов, объединённая общим узлом.
- Узел – общий предок рассматриваемой отдельно группы.
- Ветви – отрезки, соединяющие корень с узлами или листьями.
- Корень – предположительный предок, от которого отходят все ветви.

В филогении существуют методы построения деревьев и деление их на две группы: дистанционные и методы, основанные на характере и морфологических признаках.

1. Дистанционные методы.

Деревья строятся на основе генетических или морфологических расстояний. Эти методы используют попарные сравнения последовательностей или признаков для вычисления расстояния. Выделяют метод Neighbor-Joining (NJ) и метод невзвешенных парных групп со средним арифметическим (UPGMA).

- Присоединение к соседям (Нью-Джерси).
Алгоритм начинается со звёздообразного древа. Далее многократно объединяются пары таксонов с наименьшим попарным расстоянием, пока не будет сформировано полное древо. Метод последовательно объединяет данные до тех пор, пока все виды не объединятся в некорневое древо. При этом происходит минимизация общей длины ветвей древа.
- Метод невзвешенных парных групп со средним арифметическим (UPGMA).
Предполагает постоянную одинаковую скорость эволюции или молекулярные часы. Принцип поиска пар тот же, что и у Нью-Джерси. Алгоритм создаёт бинарное древо с длинами ветвей, представляющими средние генетические или морфологические расстояния.

Дистанционные методы имеют преимущество в работе с большим количеством информации. Однако, из-за необъективности данных такие методы могут давать относительные результаты без использования дополнительных методов. Поэтому при изучении деревьев исследователи сравнивают результаты с помощью методов, основанных на расстоянии с методами на основе признаков или иными подходящими методами, чтобы оценить точность и надёжность филогенетических отношений.

2. Методы, основанные на характере.

Деревья, строят на основе анализа определённых признаков и черт. Методы используют наличие или отсутствие общих производных признаков (синапоморфий) для вывода об эволюционных отношениях.

Вторая группа не подходит по определению, т.к. рассматриваются только геномы бактериофагов, а информация о внешних признаках не имеется. Поэтому строим дерево с использованием дистанционных методов.

Для построения древа использовался NJ-метод. В данном случае использование только этого метода не показало бы реальной вероятности такой филогении, поэтому вдобавок к NJ, был использован и бутстреп (*). Из-за его появления при описании родства опираемся именно на него. Только при совмещении этих методов и анализе полученных цифр, можно говорить о достоверности результатов.

* Bootstrap метод (или бутстреп метод) – статистический метод, используемый для оценки уверенности в филогенетических исследованиях.

Суть бутстреп метода заключается в создании большого количества псевдовыборок из исходного набора данных путем выбора элементов с замещением. Затем для каждой псевдовыборки строится филогенетическое дерево с использованием выбранного метода построения таксономических деревьев.

Далее, используя большое количество построенных филогенетических деревьев, проводится высчитывание процентов вероятности появления той или иной ветви. Чем выше такой процент, тем больше уверенность в том, что данная ветвь является надежной и соответствует истинной филогении. Процент выше 70 говорит о близкородственности.

Этот метод позволяет оценить стабильность филогенетической реконструкции, учитывая различные факторы, такие как недостаточное количество данных или неправильная модель эволюции.

Таким образом, в программе Mega было создано древо с использованием двух методов NJ-метода и бутстреп. (Приложение 1).

Для облегчения анализа оно было перерисовано. Каждому виду был присвоен свой цвет. (Приложение 2). Затем были выделены три группы родства на основе процентного соотношения:

- 1 группа (100%) – близкородственная, к ней можно подобрать общую пару праймеров.
- 2 группа (78%) – близкородственная, но из-за геномных мутаций подбор общей пары может быть невозможен.
- 3 группа (процент не указан) – пара праймеров будет подбираться отдельно.

Теперь группируем листы по ближайшим узлам, нумерует образованные клады и с учётом промежуточного анализа предполагаем прайм-систему, состоящую из 4 пар. Для 1, 2 и 3 клады – по одной общей, а для 4 – отдельную.

Характеристика праймеров

Чтобы в результате ПЦР-тестирования получить правдивые результаты, нужно не только подбирать качественные реагенты и устанавливать правильную температуру, но и нельзя забывать про праймеры. Праймеры – основа ПЦР. А грамотно подобранные праймеры – половина сделанного качественного анализа.

Для создания подходящей пары праймеров необходимо знать пару основ.

Во-первых, праймеры – это сразу две последовательности прямая (Forward primer) и обратная (Reverse primer). Прямой праймер ищется от 5`- конца матричной (транскрибируемой) цепи и полностью соответствует её последовательности нуклеотидов. Обратный праймер также ищется от 5`-конца, но противоположной цепи ДНК – смысловой (кодирующей). При этом обратный праймер является комплементарной обратной последовательностью выбранного участка матричной цепи.

Во-вторых, праймеры должны соответствовать основным характеристикам:

- Длина праймера – 18-25 нуклеотидов
- t плавления – 45-65 градусов
- ГЦ-состав – 40%-60%
- Разница в t между Forward primer и Reverse primer - <5 градусов
- Длина ампликона (*) – 100-1000 нуклеотидов
- Наличие специфичных 3`- концов
- Г и Ц на 3`- конце (при этом 3`- конец не должен содержать ГЦ-богатого участка)

* Ампликон – это участок ДНК, который удваивается в результате проведения ПЦР. Расположен между Forward primer и Reverse primer и ими же определяется.

Стоит заметить, что для создания «идеального» праймера характеристики более узкие (например, разница в t между праймерами <2 градусов) и добиться их намного сложнее. Иногда подобрать праймер, соответствующий характеристикам невозможно. В таких случаях в ПЦР используются специфичные условия или в праймер добавляются нуклеотиды.

В-третьих, необходимо избегать прогоны, повторы и уметь предсказать появление нежелательных продуктов шпилек и димеров.

Шпильки – это вторичные структуры, которые образуются во время ПЦР, когда праймер комплементарен ДНК-матрице. Таким образом образуется вторичный продукт, который увеличит время ПЦР, уменьшит эффективность и приведёт к ложноположительным или ложноотрицательным результатам.

Димеры – это такие же вторичные структуры, но в отличие от шпилек они образуются, когда праймеры образуют комплементарные участки. Димеры делятся на 2 вида – self-димеры и hetero-димеры. В self-димерах комплементарными являются одинаковые праймеры, а в hetero-димерах – разные.

Для дальнейшей работы с подбором пар праймеров и поиском нежелательных продуктов будет использоваться программа Ugene. Но перед тем, как проверять пару праймеров в Ugene, она сначала будет проанализирована на сайте OligoCalc. В нём можно также определить наличие димеров и шпилек, но его точность не высока. Однако при обнаружении данных структур праймер далее оцениваться не будет. Кроме продуктов в работе с OligoCalc анализироваться будут ГЦ-состав и t плавления. Также сайт помогает с поиском Reverse primer. Данная функция будет использована тоже.

Поиск праймеров и создание прайм-системы

На практике выяснилось, что 3 клада не имеет длинных общих участков нуклеотидов и пары праймеров, будут подбираться отдельно. Всё выравнивание было разделено на участки для подбора пар. (Приложение 3).

Для подбора праймеров можно воспользоваться двумя способами: ручным (подбор человеком) и механическим (подбор с использованием сторонних программ). Разница между ними заключается в сложности и затраченном времени, при этом решено было опробовать оба. Для общих пар – ручной, для отдельных – механический.

Таким образом, было найдено 6 пар праймеров, которые полностью соответствуют основным характеристикам.

Далее праймеры были занесены в таблицу с указанием соответствующих ему геномов. Это и есть прайм-система. (Приложение 4).

На основе созданной системы был выбран вариант проведения – мультиплексная ПЦР. С его помощью можно амплифицировать сразу все 6 пар праймеров для получения множественных результатов. Этот способ более эффективен и затрачивает меньше времени, нежели обычная ПЦР. Также критерии, необходимые для её проведения, в системе соблюдены. Следовательно, при помощи системы можно сэкономить ещё больше ресурсов.

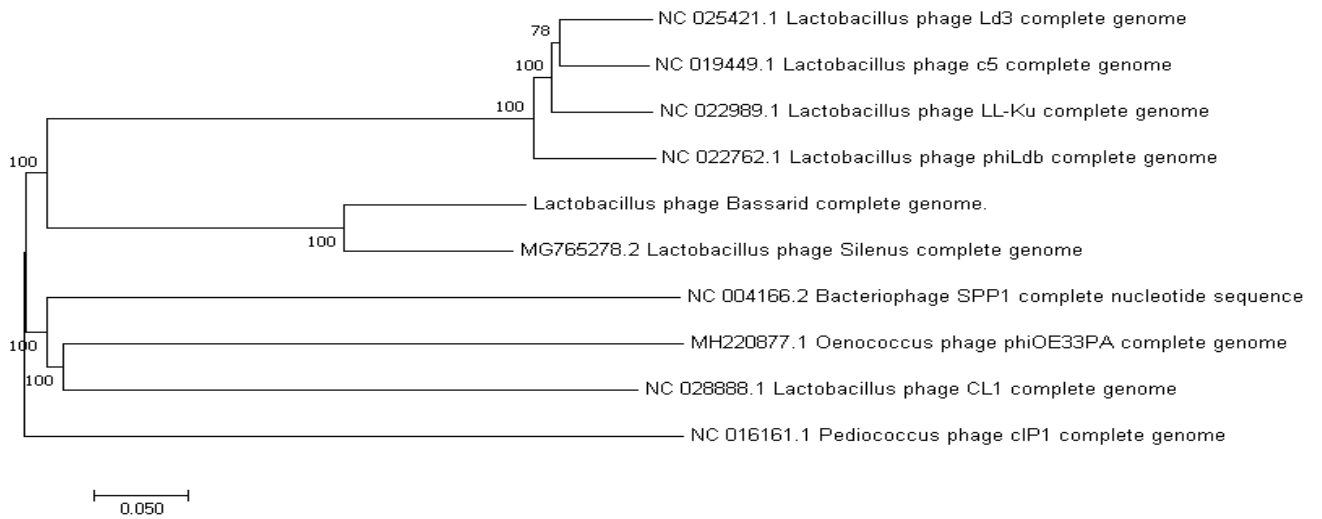
Выводы

1. Идея проекта была направлена на помощь производителям по улучшению качества кисломолочных продуктов и уменьшение затрат с ресурсами посредством создания прайм-системы по детекции бактериофагов.
2. В результате проделанной работы было сделано следующее:
 - С помощью сайта NSBI были найдены геномы бактериофагов, полученные из реально заражённых заквасок;
 - Пользуясь программами Mega и Ugene, сделано выравнивание ClustalW и проведён анализ геномов;
 - Благодаря методам Neighbor-Joining и Bootstrap, было сокращено количество возможных пар праймеров и найдены участки геномов для каждой подбираемой пары;
 - Были использованы сайты NSBI, OligoCalc и программа Mega для создания специфичной прайм-системы, подразумевающей проведение мультиплексной ПЦР для большей эффективности метода.

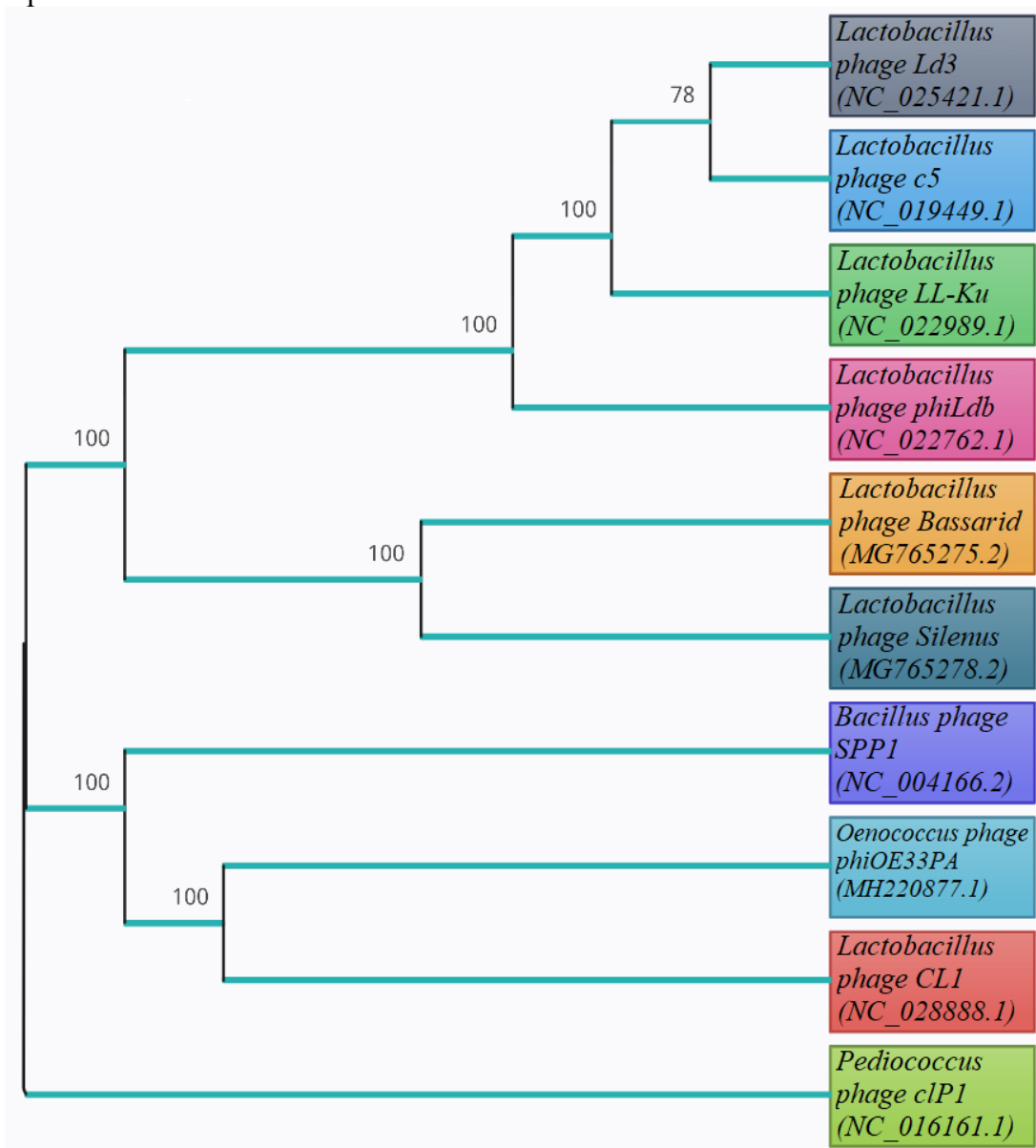
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- Cécile Philippe/Jeffrey K Cornuault/Alessandra G de Melo/Rachel Morin-Pelchat/Alice P Jolicoeur/Sylvain Moineau//The never-ending battle between lactic acid bacteria and their phages. *FEMS Microbiology Reviews* – 2023. – URL: <https://academic.oup.com/femsre/article-abstract/47/4/fuad035/7206401?redirectedFrom=fulltext>
- Smith, J. //Methods in Pictures: Polymerase Chain Reaction. *Biomolecula*. – 2022. – URL: <https://biomolecula.ru/articles/metody-v-kartinkakh-polimeraznaia-tsepnaia-reaktsiia>
- Premier Biosoft. PCR Primer Design. URL: https://premierbiosoft.com/tech_notes/PCR_Primer_Design.html
- UGENE//PCR Primer Design for DNA Assembly. Официальный сайт UGENE. – 2021. – URL: <https://doc.ugene.net/wiki/display/UM/PCR+Primer+Design+for+DNA+Assembly>
- Чемерис Д.А./Кирьянова О.Ю./Губайдуллин И.М./Чемерис А.В.//Дизайн праймеров для полимеразной цепной реакции (краткий обзор компьютерных программ и баз данных). *Биомика*. – 2016. – URL: <https://biomicsj.ru/upload/iblock/b05/3.pdf>
- Josiane E Garneau/Sylvain Moineau.//Bacteriophages of lactic acid bacteria and their impact on milk fermentations. *PubMed*. – 2011. – URL: <https://pubmed.ncbi.nlm.nih.gov/21995802/>
- Бу Сурав Био//Определение праймера, типы дизайна праймера, онлайн инструменты, используемые. *Microbiology Note*. – 2023. – URL: <https://microbiologynote.com/ru/определение-праймера%2C-типы-дизайна-праймера%2C-онлайн-инструменты%2C-используемые/#TypesofPrimerinPCR>
- Бу Сурав Био//Типы определения филогенетического дерева, шаги, методы, использование. *Microbiology Note*. – 2023. – URL: <https://microbiologynote.com/ru/типы-определения-филогенетического-дерева%2C-шаги%2C-методы%2C-использование/>

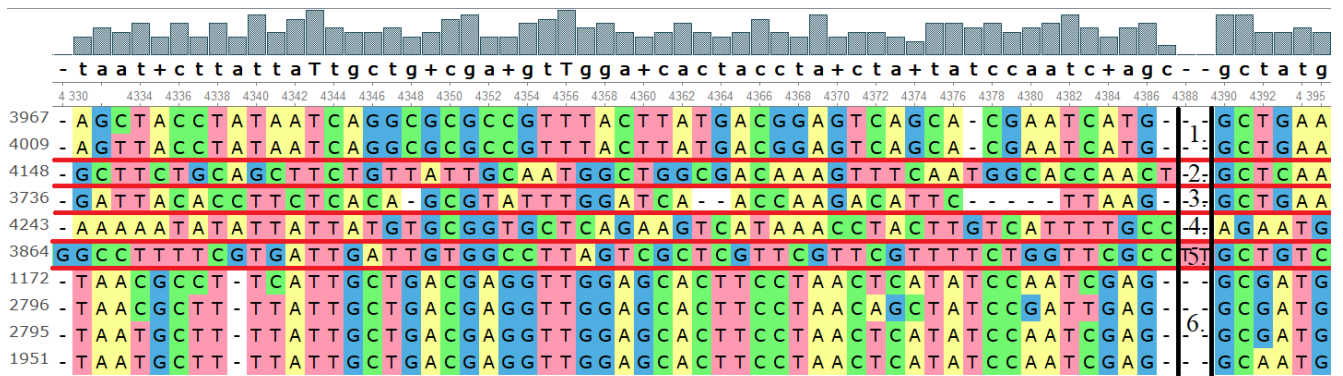
Приложение 1.



Приложение 2.



Приложение 3.



1 и 6 возможно подобрать общую систему пар праймеров;
2 - 5 нужно подбирать разные системы праймеров.

Приложение 4.

№	Название(-я)	Праймеры
1	Lactobacillus phage Bassarid (MG765275.2) Lactobacillus phage Silenus (MG765278.2)	F: 5`GATGGATAAGCAGCGGGAAC3` R: 5`ACATATAGCCTGCTAGACGC3`
2	Oenococcus phage phiOE33PA (MH220877.1)	F: 5`GTCTTGCCGCCAAAAGATGG3` R: 5`CGCCCCTTATCAAACCACGA3`
3	Lactobacillus phage CL1 (NC_028888.1)	F: 5`AATTGCCGATGCTCCATTGC3` R: 5`GATTCGCCGTTTTGCTGGTT3`
4	Bacteriophage phage SPP1 (NC_004166.2)	F: 5`TCCATCTACGGTGGGGTAGG3` R: 5`GCTTTTGATGTGTTTCGGCGT3`
5	Pediococcus phage cIP1 (NC_016161.1)	F: 5`GGTTCGACCCACGAACATA3` R: 5`ATCAGTCCCGCCGTTTGTTA3`
6	Lactobacillus phage Ld3 (NC_025421.1) Lactobacillus phage phiLdb (NC_022762.1) Lactobacillus phage c5 (NC_019449.1) Lactobacillus phage LL-Ku (NC_022989.1)	F: 5`AGGTTGGAGCACTTCCTAAC3` R: 5`ACCGCAAGCCATGACTTTTC3`

Рецензия

по теме «Создание прайм-системы для определения детекции бактериофагов на промышленном этапе производства кисломолочных продуктов.»

Автор проекта: обучающаяся 10 класса

МБОУ ЛСТУ №2 г.Пензы

Суслина Софья

В работе «Создание прайм-системы для определения детекции бактериофагов на промышленном этапе производства кисломолочных продуктов» поднята серьезная проблема сохранения кисломолочной продукции.

Исследование представляет собой значимый вклад в область ПЦР-методики, который позволяет уменьшить затраты ресурсов при детекции бактериофагов посредством создания и улучшения системы праймеров, что позволяет повысить качество и безопасность готовой продукции.

Для получения нужного результата были использованы различные программы в сфере биоинформатики: программы Mega, Ugene и сайты NCBI, OligoCalc (ОлигоКалькулятор).

Проведена большая практическая работа, которая включает в себя поиск геномов бактериофагов и их анализ, а также создание системы праймеров и выявление лучших методов для её эксплуатации.

Благодаря этой работе, предлагаются новые возможности для контроля за процессом производства, который может значительно повысить качество конечного продукта.

В исследовании приведено множество подробной теоретической информации, сравнение различных методик и полученных результатов, что гарантирует её научную надёжность.

Работа имеет большой круг реализации и обладает потенциалом для широкого применения в пищевой промышленности. Результаты исследования могут быть полезны как для производителей кисломолочной продукции, так и для научного сообщества, заинтересованного в совершенствовании методов контроля за качеством пищевых продуктов.