

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
имени М.В. ЛОМОНОСОВА

*На правах рукописи*

**ЕВСЮТИНА ДАРЬЯ ВИКТОРОВНА**

**Регуляция передачи генетической информации у бактерий  
с редуцированным геномом**

1.5.3 – Молекулярная биология

**ДИССЕРТАЦИЯ**  
на соискание ученой степени  
кандидата биологических наук

Научный руководитель:  
академик РАН,  
доктор биологических наук,  
профессор  
Говорун Вадим Маркович

МОСКВА – 2022

## ОГЛАВЛЕНИЕ

<b>СПИСОК СОКРАЩЕНИЙ</b> .....	<b>5</b>
<b>ВВЕДЕНИЕ</b> .....	<b>6</b>
<b>1. ОБЗОР ЛИТЕРАТУРЫ</b> .....	<b>11</b>
1.1 Передача генетической информации .....	11
1.2 Общие представления о транскрипции у бактерий .....	14
1.3 Основные механизмы регуляции экспрессии генов у бактерий .....	15
1.3.1 Регуляция на уровне транскрипции .....	15
1.3.2 Посттранскрипционная регуляция экспрессии генов .....	17
1.3.3 Посттрансляционная регуляция .....	17
1.4 Методы поиска регуляторов и их мишеней .....	18
1.4.1 Футпринтинг ДНК .....	18
1.4.2 Анализ сдвига электрофоретической подвижности (EMSA) .....	20
1.4.3 Систематическая эволюция лигандов экспоненциальным обогащением (SELEX) .....	20
1.4.4 ChiP-chip, ChiP-seq, ChiP-exo.....	22
1.4.5 Биоинформатические методы .....	23
1.5 Особенности регуляции передачи генетической информации у бактерий с редуцированным геномом класса Mollicutes, на примере микоплазм.....	25
<b>2. МАТЕРИАЛЫ И МЕТОДЫ</b> .....	<b>31</b>
2.1 Поиск потенциальных факторов регуляции транскрипции <i>M. gallisepticum</i> S6 .....	31
2.2 Оценка консервативности белков и геномного контекста.....	31
2.3 Поиск сайтов связывания потенциальных транскрипционных факторов <i>MraZ</i> , <i>Xre</i> , <i>Fur</i> , <i>WhiA</i> .....	31
2.4 Питательные среды.....	31
2.5 Выделение геномной ДНК.....	32
2.6 Компетентные клетки и трансформация <i>E. coli</i> .....	32
2.7 Выделение плазмидной ДНК щелочным гидролизом.....	33
2.8 Получение чистого препарата целевых белков из <i>M. gallisepticum</i> в клетках <i>E. coli</i> .....	34
2.9 Электрофорез белков в полиакриламидном геле по методу Лэммли.....	35
2.10 Вестерн-блот анализ .....	35
2.11 Метод торможения ДНК в геле (EMSA) .....	36
2.12 Расчет константы диссоциации .....	37
2.13 Эксклюзионная хроматография.....	37
2.14 Измерение уровня АТФ.....	38
2.15 Конструирование векторов для сверхэкспрессии генов в <i>M. gallisepticum</i> .....	38

2.16 Трансформация плазмидной ДНК клеток <i>M. gallisepticum</i> методом электропорации ..	39
2.17 Выделение РНК и синтез кДНК .....	40
2.18 Количественная ОТ-ПЦР в реальном времени .....	40
2.19 Разработка вектора для оценки влияния генетических элементов на экспрессию гена <i>egfp</i> .....	41
2.20 Таргетная репрессия транскрипции генов у <i>M. gallisepticum</i> с помощью системы CRISPR-интерференции (CRISPRi) .....	42
2.21 Выделение рибосом и фракции мРНК, связанной с рибосомами <i>M. gallisepticum</i> .....	46
2.22 Подготовка библиотек кДНК и секвенирование .....	47
2.23 Анализ данных секвенирования РНК .....	48
<b>3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ .....</b>	<b>49</b>
3.1 Поиск потенциальных факторов регуляции транскрипции у <i>M. gallisepticum</i> на основании сходства первичной структуры .....	49
3.2 Поиск сайтов связывания белка <i>MraZ</i> .....	51
3.2.1 Белок <i>MraZ</i> связывает серию прямых повторов, расположенных в промоторной области DCW кластера .....	53
3.2.2 Сайт связывания <i>MraZ</i> консервативен у бактерий класса Mollicutes .....	54
3.2.3 Другие мишени белка <i>MraZ</i> в геноме <i>M. gallisepticum</i> .....	54
3.2.4 Определение стехиометрии комплекса белка <i>MraZ</i> с дцДНК.....	56
3.2.5 Сверхэкспрессия белка <i>MraZ</i> влияет на форму клеток бактерии .....	57
3.3 Поиск сайтов связывания белка <i>Xre</i> ( <i>HsdC</i> ).....	57
3.3.1 Распределение гомологов белка C ( <i>Xre</i> ; <i>HdsC</i> ) у бактерий класса Mollicutes .....	59
3.3.2 Белок <i>Xre</i> связывает прямые повторы, расположенные в промоторной области оперона <i>hsd</i> .....	60
3.3.3 <i>Xre</i> , вероятно, выступает в качестве репрессора транскрипции <i>hsd</i> оперона.....	61
3.3.4 Другие мишени белка <i>Xre</i> ( <i>HsdC</i> ) в геноме <i>M. gallisepticum</i> S6 .....	62
3.4 Определение сайта связывания и генов-мишеней белка семейства FUR .....	63
3.4.1 Распределение белков семейства FUR у молликут .....	64
3.4.2 <i>Fur</i> – регулятор экспрессии генов, кодирующих транспорт катионов металлов .....	68
3.5 Поиск сайтов связывания белка <i>WhiA</i> .....	68
3.5.1 Промоторная область оперона <i>rpsJ</i> у разных бактерий консервативна и содержит core мотив.....	68
3.5.2 <i>WhiA</i> <i>M. gallisepticum</i> специфически связывается с промоторной областью оперона <i>rpsJ</i> .....	70
3.5.3 АТФ влияет на связывание <i>WhiA</i> с ДНК-субстратом.....	71
3.5.4 Репрессия транскрипции <i>whiA</i> приводит к увеличению экспрессии <i>rpsJ</i> .....	73

3.6 Применение системы для поиска и валидации новых генов – мишеней регуляторов ....	74
3.6.1 Тестирование системы для известного фактора регуляции транскрипции и его гена-мишени .....	76
3.6.2 Влияние разных элементов промотора на экспрессию репортерного гена <i>egfp</i> у <i>M. gallisepticum</i> .....	77
3.6.3 Примеры репрессии .....	78
3.6.4 Пример активации.....	84
3.7 Анализ фракции мРНК, связанной с рибосомами .....	85
3.7.1 Уровень транскрипции пропорционален уровню трансляции в нормальных условиях .....	85
3.7.2 Модель предсказания обогащения мРНК во фракции, связанной с рибосомами ....	87
<b>ЗАКЛЮЧЕНИЕ .....</b>	<b>90</b>
<b>ВЫВОДЫ .....</b>	<b>92</b>
<b>СПИСОК ЛИТЕРАТУРЫ .....</b>	<b>93</b>
<b>ПРИЛОЖЕНИЕ.....</b>	<b>107</b>
<b>БЛАГОДАРНОСТИ .....</b>	<b>145</b>

## СПИСОК СОКРАЩЕНИЙ

- EMSA – Electrophoretic mobility shift assay
- EGFP – Enhanced Green fluorescent protein
- HEPES – 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
- IPTG – Isopropyl  $\beta$ -D-1-thiogalactopyranoside
- LB – Lysogeny broth, Luria Bertani
- MOPS – (3-(N-morpholino)propanesulfonic acid
- RBS – Ribosome-binding site
- TSS – Transcription start site
- TCC – Сайт начала транскрипции
- АДФ – Аденозиндифосфат
- АМФ – Аденозинтрифосфат
- АТФ – Аденозинмонофосфат
- ДТТ – Дитиотреитол
- ДСН (SDS) – Додecilсульфат натрия
- ПААГ – Полиакриламидный гель
- п.н. – пар нуклеотидов
- РБ – фракция мРНК, связанная с рибосомами
- РБС – сайт связывания рибосомы
- Р-М – Рестрикция–модификация
- СЦ – суммарная цитоплазматическая фракция мРНК
- ТЕМЕД – Tetramethylethylenediamine
- т.п.н – тысяча пар нуклеотидов
- ТФ – Транскрипционный фактор
- ЭДТА – Этилендиаминтетрауксусная кислота

## ВВЕДЕНИЕ

**Актуальность работы.** Передача генетической информации – один из центральных процессов в живых организмах, связывающий между собой ДНК, РНК и белки. Её регуляция обеспечивает пластичность живой системы и может осуществляться на разных этапах. Бактерии с редуцированным геномом, такие как представители класса Mollicutes, способны к самостоятельной репликации своего генетического материала на искусственных средах. Уменьшение размера генома у этих бактерий привело и к редукции репертуара регуляторов экспрессии генов, но, несмотря на это, Молликуты сохранили широкие адаптивные способности. Изучение регуляции передачи генетической информации у бактерий с редуцированным геномом представляет интерес как с фундаментальной (обеспечение стабильности клетки и процессов в ней при наличии малого числа известных регуляторов), так и с прикладной точек зрения (изучение правил, которые помогут создать искусственную клетку с заданным набором свойств и функций).

**Степень разработанности темы.** Микоплазмы были одними из первых бактерий, чьи геномы были полностью секвенированы и со временем, как бактерии с редуцированным геномом, стали главными модельными объектами системной биологии. Ученые полагали расшифровать основные принципы организации и функционирования бактериальной клетки на этих «просто устроенных» организмах. Однако, чем больше накапливалось мультиомиксных данных о микоплазмах, тем больше появлялось вопросов, в частности о том, как осуществляется регуляция процессов в отсутствие классических регуляторов, известных для других модельных бактерий. В работах, опубликованных группой Л. Серрано, была предложена модель регуляции экспрессии генов, где основной вклад в регуляцию вносит пространственная организация ДНК, вторичные структуры РНК, малые РНК, а действием транскрипционных факторов может быть объяснена лишь малая доля изменений уровня мРНК. Описанные Л. Серрано и коллегами процессы действительно могут вносить свой вклад в регуляцию экспрессии генов. В свою очередь, в представленной работе мы попытались найти и охарактеризовать мишени как предсказанных транскрипционных факторов, так и новых, еще неизвестных, и выяснить их вклад в регуляцию транскрипции.

**Цели и задачи исследования.** Целью работы было выяснение механизмов регуляции передачи генетической информации у бактерий с редуцированным геномом на примере *Mycoplasma gallisepticum* S6.

Для достижения цели были поставлены следующие задачи:

1. Из всех генов *M. gallisepticum* S6 отобрать гены потенциальных факторов регуляции транскрипции (ТФ)
2. Провести поиск генов-мишеней потенциальных транскрипционных факторов
3. Экспериментально подтвердить сайты связывания ТФ-ов
4. Предложить способ валидации новых генов – мишеней регуляторов
5. Рассмотреть связывание рибосомы с мРНК как возможный механизм регуляции передачи информации от мРНК к белку. В качестве модели использовать ответ микоплазмы на тепловой стресс

**Научная новизна работы.** В результате проделанной работы описаны и охарактеризованы четыре новых транскрипционных фактора *M. gallisepticum* S6 – MraZ Xre, WhiA и Fur. Найдены и экспериментально подтверждены их сайты связывания. Разработана и валидирована система для поиска мишеней регуляторов на транскрипционном уровне, а также найден новый регулятор – ParX. Удалось установить избирательность рибосом в отношении к мРНК при тепловом стрессе. Впервые показано, что адаптивный ответ микоплазмы на тепловой стресс состоит в изменении транскрипции генов, кодирующих шапероны, иммуноглобулин-связывающие белки, белки, участвующие в делении клетки, и несколько белков с неизвестной функцией. Изменение экспрессии остальных генов, вероятно, представляет собой транскрипционный шум, поскольку не приводит к существенному изменению абсолютного содержания транскриптов.

**Теоретическая и практическая значимость работы.** Полученные результаты вносят вклад в раскрытие регуляторных событий у бактерий с редуцированным геномом и могут быть использованы как для создания искусственных живых систем с заданными функциями, так и для поиска новых препаратов для борьбы с микоплазмами.

**Методология и методы исследования.** Работа выполнена с использованием стандартных молекулярно-биологических и биохимических методов. Для получения чистых препаратов потенциальных транскрипционных факторов были сконструированы вектора, содержащие гены интересующих белков. Генетический код микоплазм отличается от кода других бактерий. Так, кодон TGA не является стоп-кодоном, а кодирует триптофан. Таким образом, для использования векторов в стандартной экспрессионной системе в *E. coli* необходимо было провести замену TGA кодонов в кодирующих последовательностях генов на TGG. Экспериментальное подтверждение ДНК-связывающих свойств белков-кандидатов оценивали с помощью электрофореза в нативных условиях. Для каждого белка были подобраны условия для его взаимодействия с субстратом – дцДНК. Анализ консервативности генов-регуляторов и их сайтов связывания проводили стандартными биоинформатическими

методами. В работе были созданы новые генетические конструкции для сверхэкспрессии генов в Молликутах. Такие суперпродуценты позволили выявить фенотипические эффекты действия новых транскрипционных факторов. Для поиска и подтверждения мишеней ТФ-ов была разработана система на основе CRISPR интерференции для подавления экспрессии целевых генов у *M. gallisepticum*. Для валидации новых генов – мишеней регуляторов – была разработана система оценки влияния вносимого генетического элемента на уровень транскрипции. В эту систему входит вектор на основе транспозона, который может быть использован для микоплазм. Вектор позволяет вносить желаемые генетические элементы (участки промотора, потенциальный сайт связывания ТФ-а) в промоторную область гена *egfp* и затем с помощью ПЦР в реальном времени оценить влияние элемента на экспрессию репортерного гена. Для анализа фракции мРНК, связанной с рибосомами, были выделены целые рибосомы, проведено секвенирование РНК, оценена дифференциальная экспрессия генов. Также в работе были использованы рутинные микробиологические и биохимические подходы, которые подробно описаны в разделе «Материалы и методы».

#### **Положения, выносимые на защиту:**

1. Используя базы данных функциональных доменов, было предсказано 10 потенциальных факторов регуляции транскрипции у *M. gallisepticum* S6.
2. Белок *MraZ* связывает серию прямых повторов – AAAGTG[T/G], разделенных спейсером в три нуклеотида. *MraZ* активирует транскрипцию генов кластера деления. Сверхэкспрессия *mraZ* отражается на фенотипе клеток микоплазмы – клетки удлиняются, образуя филаменты.
3. Белок *Xre* (*HsdC*) связывает повторы вида (GTGTTAN<sub>5</sub>)<sub>2</sub>. *Xre* выступает репрессором экспрессии генов одной из систем рестрикции–модификации.
4. Белки семейства *FUR* неравномерно распределены среди представителей семейства *Mycoplasmataceae*. Несмотря на консервативность оперона, содержащего *fur*, у трёх хорошо изученных видов микоплазм ортологичные мишени ТФ-а не обнаружены. У *M. gallisepticum* сайт связывания *Fur* представляет собой инвертированные повторы (TTATTTWDAAWWWTTTHWAAATAA). *Fur* является репрессором гена *zip*, кодирующего мембранный белок – транспортер ионов металлов.
5. Ген, кодирующий транскрипционный фактор *WhiA*, консервативен среди грамположительных бактерий. *WhiA* связывается с последовательностью GATACACCN<sub>7</sub>GTTGTC, расположенной в промоторной области генов рибосомных белков, подавляя их экспрессию.

6. Разработан вектор на основе транспозона Tn4001 для оценки влияния генетических элементов на экспрессию репортерного гена. С помощью данной системы найден набор генов *M. gallisepticum*, экспрессия которых зависит от влияния регуляторов. К таким генам относятся: *uvrB*, *parA*, *rplJ*. Был идентифицирован новый белок – ParX, способный связываться с 5'-областью *parA* и подавлять его экспрессию.
7. При действии теплового стресса рибосомы проявляют избирательность в отношении к мРНК. Основываясь на абсолютной представленности транскрипта и изменении частоты связывания рибосомы с транскриптом при стрессе, ответ на стресс делится на две части. Первая часть – это шум. Вторая – адаптивный ответ. Рибосома избирательно транслирует мРНК, кодирующие шапероны, иммуноглобулин-связывающие белки, белки, участвующие в делении клетки, и несколько белков с неизвестной функцией.

**Личный вклад автора.** Автор принимал активное участие в постановке научных задач, проведении экспериментов по генной инженерии *E. coli* и *M. gallisepticum*, оценке уровня экспрессии генов, биоинформатическому анализу данных, что вносило решающий вклад в исследование. Основные результаты, представленные в работе, получены лично автором. Текст диссертации написан самостоятельно. Порядок авторов, указанных в научных трудах, соответствует их вкладу. Вклад автора в научных трудах, где автор указан первым составляет 80%; вторым – 40%; третьим и последующим – до 20%. Планирование экспериментов было выполнено совместно с д.б.н. В.М. Говоруном и к.б.н. Г.Ю. Фисуновым. Выделение чистых препаратов белков и эксклюзионная хроматография выполнены совместно с к.б.н. В.А. Мануверой. Микроскопия проводилась при участии А.В. Летарова. Подготовка библиотек кДНК для высокопроизводительного секвенирования проводилась совместно с к.х.н. Т.А. Семашко и к.б.н. А.С. Никитиной. Биоинформатический анализ выполнен совместно с И.А. Гараниной. Подготовка к публикации полученных результатов проводилась при участии всех соавторов.

**Степень достоверности.** Результаты были получены с применением классических методов молекулярной биологии, с использованием качественных реактивов. Работа выполнялась на современном оборудовании, изготовленном ведущими мировыми производителями. Все эксперименты были повторены несколько раз и хорошо воспроизводимы.

**Апробация результатов.** По материалам диссертации опубликовано 8 статей в международных рецензируемых журналах, рекомендованных для защиты в диссертационном совете МГУ по специальности 1.5.3 (03.01.03) – Молекулярная биология. Также результаты были представлены на всероссийских и международных конференциях: XXVIII Зимняя

молодежная научная школа "Перспективные направления физико-химической биологии и биотехнологии" (Москва, 2016 г.); Научная конференция молодых ученых по медицинской биологии ФГБУ ФНКЦ ФХМ ФМБА России (Москва, 2016 г.); The 7th EMBO Meeting (Мангейм, 2016 г.); V Съезд биохимиков России (Сочи, 2016 г.); Systems Biology and Bioinformatics, The Ninth International Young Scientists School SBB-2017 (Ялта, 2017 г.); Международная научная конференция по биоорганической химии «XII чтения памяти академика Ю.А. Овчинникова», VIII Российский симпозиум «Белки и пептиды» (Москва, 2017 г.); Итоговая научно-практическая конференции ФГБУ ФНКЦ ФХМ ФМБА (Москва, 2021 г.).

**Структура и объем диссертации.** Материалы диссертации изложены на 145 страницах текста, включают 50 рисунков и 6 таблиц. Диссертация состоит из следующих разделов: список сокращений, введение, обзор литературы, материалы и методы, результаты и обсуждение, заключение, выводы, список литературы и приложение. Список литературы включает 163 источника.

## 1. ОБЗОР ЛИТЕРАТУРЫ

### 1.1 Передача генетической информации

Основным правилом передачи генетической информации долгое время считалась центральная догма молекулярной биологии, сформулированная Френсисом Криком в 1958 году. В своей статье «On protein synthesis», посвящённой механизмам белкового синтеза и взаимосвязям между РНК, ДНК и белками, автор излагает два основных принципа: «Гипотеза последовательности» и «Центральная догма» [1].

Согласно гипотезе последовательности, специфичность нуклеиновой кислоты проявляется исключительно в определённой последовательности оснований, из которых она состоит. И эта последовательность – код для аминокислотной последовательности конкретного белка [1]. Нужно отметить, что не было достоверно известно, как передаётся информация от гена к белку. Верный генетический код – правило, по которому нуклеотидным триплетам мРНК ставятся в соответствие аминокислоты – был открыт только в 1965 году. Хотя гипотезы кодирования аминокислот высказывались и раньше [2]. Теория последовательности была широко распространена среди учёных.

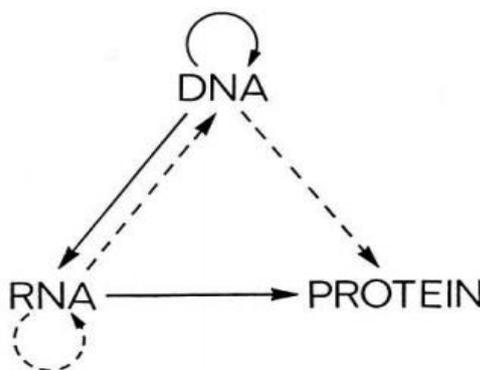
Центральная догма в оригинальной формулировке Крика: «This states that once 'information' has passed into protein it cannot get out again. In more detail, the transfer of information from nucleic acid to nucleic acid, or from nucleic acid to protein may be possible, but transfer from protein to protein, or from protein to nucleic acid is impossible. Information means here the precise determination of sequence, either of bases in the nucleic acid or of amino acid residues in the protein». В ней говорится, что когда «информация» перешла в белок, она не может выйти обратно. То есть возможна передача информации от нуклеиновой кислоты к нуклеиновой кислоте, или от нуклеиновой кислоты к белку, а передача от белка к белку, или от белка к нуклеиновой кислоте, невозможна. Под словом «информация» Крик понимал точную последовательность, либо оснований в нуклеиновой кислоте или аминокислотных остатков в белке [1]. Сформулированная догма привлекла к себе внимание. Так, критике подвергалось само использование слова «догма», узость понятия «информация» (в основном в конце XX-ого, начале XXI в.), ставились вопросы, в чём же заключается «передача информации».

Догма (греч. dogma) – система основных положений кого-нибудь учения или научного направления; положение, утверждение, не допускающее возражений и не требующее доказательств. Сам Крик писал в своей автобиографии ‘What Mad Pursuit’, что использование слова «догма» вызвало, чуть ли не больше вопросов, чем её содержание. В ответ на вопросы о причинах использования слова «догма» учёный писал: «Я назвал эту идею «центральной

догмой», по двум причинам. Во-первых, я уже использовал слово «гипотеза» при формулировке первого принципа – «гипотеза последовательности». Во-вторых, я хотел сделать акцент на то, что моё новое предположение более могущественное. Я применял это слово не в том самом смысле, в котором его применяют в религиозных учениях, однако это предположение действительно имело слабую экспериментальную поддержку» [3].

Определение понятия «информация» дано в статье Крика довольно чётко. Со временем в контексте догмы, к слову «информация», помимо последовательности нуклеотидов нуклеиновых кислот и последовательности аминокислот в белках, предлагали добавить трёхмерную структуру нуклеиновых кислот, белков, а также малые молекулы. Был опубликован комментарий, в котором говорится о том, что маленькие молекулы являются ключевым элементом биологии. Они взаимодействуют с макромолекулами – РНК, ДНК и белками. Маленькие молекулы способны связываться с ДНК и модулировать её, регулировать функции РНК, связываясь с рибопереключателем мРНК [4], [5].

В 1970 года Френсис Крик публикует статью «The Central Dogma of Molecular Biology», посвященную центральной догме и её трактовке. Здесь он рассматривает все возможные пути передачи информации (Рисунок 1.1) [6].



**Рисунок 1.1 «Треугольник Крика».** Сплошными стрелками указаны общие пути передачи информации, пунктирными стрелками – специальные. Отсутствие стрелок, ведущих от белка к ДНК и к РНК, и от белка к белку говорит о невозможности таких переходов [6]

С развитием молекулярной биологии догма устарела, и уже доказано ее опровержение (например, эпигенетические модификации ДНК и редактирование РНК, где изменяется последовательность ДНК и РНК, соответственно). То есть, иными словами, меняется информация в понимании Крика. Рассмотренные примеры – это примеры передачи информации от белка к нуклеиновым кислотам. Однако переноса информации от белка к белку (если вкладывать в «слово» информация тот смысл, который в него вкладывал Френсис Крик) пока не открыто. То есть пока неизвестны молекулярные машины, которые могли бы

перевести аминокислотную последовательность в нуклеотидную или такие силы, которые заставили бы рибосому работать в обратном направлении.

Одним из следствий центральной догмы молекулярной биологии является прямая зависимость уровня мРНК соответствующего белка. Но все больше накапливается данных, где это соответствие не подтверждается, что указывает на наличие регуляции передачи генетической информации. Изучение таких процессов в настоящее время ведется на бактериях с редуцированным геномом, таких как представители класса Mollicutes. Эти данные позволят составить представление о заведомо необходимом и достаточном наборе связей, позволяющим клетке выживать.

Центральная догма молекулярной биологии постулирует наличие связи между репликацией, транскрипцией и трансляцией. Развитие синтетической биологии, одной из целей которой является создание системы, с предсказуемой работой генетических программ: экспрессией генов, белков, и в конечном счете функцией [7], поставило вопрос о формулировании ортогональной центральной догмы. То есть системы, которая бы не зависела от хозяина. Сейчас, при создании генно-инженерных конструкций, зачастую используются ресурсы хозяина (организма, в который интегрируется конструкция) – например, аппарат трансляции, репликации и зачастую транскрипции. В тоже время, регуляторные системы хозяина могут непредсказуемо влиять на экспрессию генов в конструкциях, перенос конструкции из одного организма в другой зачастую невозможен. Ортогональная центральная догма подразумевает, что молекулярные аппараты репликации, транскрипции и трансляции целевых конструкций и хозяина будут отличаться и не перекрываться. Один из примеров ортогональной системы репликации ДНК описан у дрожжей *Kluyveromyces lactis* [8]. Ортогональная система транскрипции используется учеными довольно давно, например, в бактериальных системах применяют аппарат транскрипции бактериофагов. Для создания регуляторной сети можно использовать разные промоторы или искусственно сконструированные регуляторы на основе цинковых пальцев, TALE белков и технологии CRISPR [9]. С ортогональной трансляцией все намного сложнее. Сам комплекс трансляции огромен и состоит из многих компонентов. Основные современные успехи связаны с перекодированием кодонов [10], созданием нового генетического кода – четырех буквенного [11]. Уже созданы синтетические тРНК. Однако трансляция осуществляется рибосомами и всем сопутствующим белковым аппаратом хозяина, и еще только предстоит определить минимальный необходимый набор факторов (белковых или нет), обеспечивающий успешный синтез белка [12].

Существует альтернативный способ создания системы, с предсказуемой работой генетических программ. Это сборка искусственной клетки, состоящей только из внесенных извне элементов. Для того, чтобы запрограммировать работу такой клетки, необходимо знать, как будут работать все ее компоненты, причем не только по отдельности, но и вместе. В таком случае изучение процессов передачи генетической информации – это важный аспект синтетической биологии. Представители класса Mollicutes обладают самым маленьким по размеру геномом среди известных организмов, способных к самостоятельной репликации своего генетического материала. Подобное свойство делает их удобным объектом для исследования на системном уровне, а также для разработки на их основе искусственных систем.

Таким образом, изучение регуляции передачи генетической информации у бактерий с редуцированным геномом представляет интерес как с фундаментальной (обеспечение стабильности клетки и процессов в ней при наличии малого числа известных регуляторов), так и с прикладной точек зрения (изучение правил, которые помогут создать искусственную клетку, с заданным набором свойств и функций).

В силу обширности темы, в данной работе будет подробно рассмотрена регуляция передачи информации от ДНК к РНК – транскрипции и менее подробно от РНК к белку – трансляции.

## **1.2 Общие представления о транскрипции у бактерий**

Транскрипция – это процесс синтеза РНК по ДНК матрице. Транскрипция является одним из наиболее регулируемых процессов в клетке и включает в себя три основных стадии: инициация, элонгация и терминация. Регуляция может происходить на всех этапах. У бактерий транскрипция осуществляется единственной ДНК-зависимой РНК полимеразой, в то время как у эукариот существует три разных РНК полимеразы. РНК полимеразы бактерий представляет собой мультисубъединичный фермент. Обычно его коровая (core) часть, необходимая для синтеза РНК, включает в себя две  $\alpha$  субъединицы,  $\beta$ ,  $\beta'$ , и  $\omega$ . Однако для инициации транскрипции необходимо взаимодействие коровой части РНК полимеразы с  $\sigma$  фактором [13]

Типичный промотор обычно содержит два гексануклеотидных мотива: -10 и -35 (координаты указаны относительно старта транскрипции) и находящийся между ними спейсер длиной около 17 п.о. Также важными элементами промотора, которые могут входить в его состав, являются «upstream» (UP) элемент длиной около 20 п.о., «extended -10» (TGn) элемент (участок длиной 3-4 п.о., расположенный сразу после -10 мотива). Сила промотора (т.е.

эффективность инициации транскрипции) определяется длиной его элементов и их нуклеотидной последовательностью. Большинство бактерий имеет один главный  $\sigma$  фактор (например, у *Escherichia coli* это  $\sigma 70$ , а у *Bacillus subtilis* –  $\sigma A$ ) и один или несколько альтернативных факторов инициации для транскрипции специфичных классов генов [14].

Кратко цикл транскрипции у бактерий включает три основных стадии: инициация, элонгация и терминация. Для того чтобы РНК полимеразы начала транскрипцию с нужного промотора, необходимо ее связывание с  $\sigma$  фактором. Сигма фактор узнает промотор и образуется закрытый комплекс, в котором ДНК сохраняет двуспиральную структуру. В таком состоянии РНК полимеразы не способна к синтезу транскрипта. Комплекс нестабилен и легко диссоциирует при повышении ионной силы [13]. Закрытый комплекс может обратимо превращаться в открытый, в котором происходит расплетание, плавление двойной спирали ДНК. Наступает стадия инициации. На этой стадии происходит присоединение нуклеотидов. РНК продукт на этой стадии непрочен связан с РНК полимеразой и ДНК матрицей и может высвободиться (так называемая абортивная транскрипция), при этом РНК полимеразы, не покидая ДНК матрицы, снова иницирует транскрипцию. Когда РНК достигнет определенной длины, абортирования транскрипции не происходит. Транскрибирующий комплекс стабилизируется, диссоциирует  $\sigma$  фактор, что является маркером начала стадии элонгации – синтез полноценного транскрипта. Завершается синтез РНК на особых участках ДНК – терминаторах. У бактерий существует два основных класса терминации –  $\rho$ -зависимая и  $\rho$ -независимая [15].

Регуляция процесса транскрипции очень важна для клетки и может осуществляться с помощью различных молекулярных механизмов.

### **1.3 Основные механизмы регуляции экспрессии генов у бактерий**

#### **1.3.1 Регуляция на уровне транскрипции**

Регуляция транскрипции является наиболее распространенным механизмом реагирования на окружающую среду. Это позволяет регулировать экспрессию генов без ненужной траты энергии и других ресурсов. Транскрипция может регулироваться несколькими разными механизмами и часто гены являются мишенями для более чем одного регуляторного события. У большинства бактерий самые большие регулоны находятся под контролем одного и того же сигма фактора [16]. Основные функции  $\sigma$  фактора заключаются в обеспечении узнавания специфичных последовательностей промотора, расположении холофермента РНК полимеразы на целевом промоторе и в облегчении расплетания двойной спирали ДНК около старта транскрипции [17]. Часто сигма-факторы сами являются объектами

регуляторных событий, либо на уровне экспрессии, либо с помощью анти-сигма факторов [18]. Такими анти-сигма факторами могут быть как белки, например, белок UsfX *Mycobacterium tuberculosis* ингибирует  $\sigma^F$  [19], так и РНК, например 6S *E. coli*. 6S РНК *E. coli* стабильна и имеет сложную вторичную структуру, мимикрируя под структуру промотора 6S РНК конкурирует за связывание с основным сигма-фактором [20]. Помимо сигма- и анти-сигма факторов, другие белки и молекулы могут влиять на сродство РНК-полимеразы к специфическим промоторам. Регуляторы транскрипции ответственны за большинство специфических регуляторных событий. Такие белковые транскрипционные факторы имеют ДНК-связывающий домен. В настоящее время расшифровано более 30 структур прокариотических ДНК-связывающих белков, и большинство из них (84%) имеют структурный мотив спираль–поворот–спираль (НТН) [21]. Транскрипционные факторы можно разделить на три группы: активаторы, репрессоры и плейотропные регуляторы, которые в зависимости от условий, могут выступать как в качестве активатора, так и в качестве репрессора. Механизмы действия активаторов могут быть разные, однако большинство из них связывается с операторными последовательностями, расположенными выше (upstream) от промотора-мишени [22]. Такое связывание приводит к изменению топологии ДНК в районе промотора, что позволяет РНК-полимеразе связаться с промотором или с дополнительными cis-активирующими элементами. Другие активаторы влияют на время жизни открытого комплекса, тем самым способствуя инициации транскрипции. Также активаторы могут модулировать действие репрессоров. Существуют активаторы, которые связываются с ДНК (downstream) от старта начала транскрипции, но это скорее исключение, чем правило [23]. В отличие от активаторов, репрессоры связываются либо непосредственно с промотором, либо между лидерной последовательностью и старт-кодоном. Таким образом, репрессор создает стерические помехи для связывания промотора РНК-полимеразой, закрывая -35 и -10 боксы или выветливая участок промотора. Также репрессоры могут модулировать действие активаторов. Плейотропные регуляторы, такие как ArgR *Streptomyces coelicolor* [24] действуют как активаторы или репрессоры в зависимости от внешних условий. Действие транскрипционных факторов только кажется просто устроенным, в живой клетке, где число таких регуляторов может достигать нескольких десятков и больше, а их мишени могут пересекаться между собой, все это образует сложную регуляторную сеть, которая позволяет эффективно адаптироваться к разным внешним воздействиям. Помимо описанных выше участников регуляции экспрессии генов, существуют белки, которые помимо способности связываться с ДНК (регуляторной функции) обладают и ферментативной активностью (по сути, это двухкомпонентные системы, где сенсор и регулятор соединены в одном белке). Одним из представителей таких бифункциональных белков является белок *putA E. coli*,

который катализирует двухстадийную деградацию пролина до глутамата, а также связывается с промоторной областью генов *putA* и *putP* (переносчик пролина), подавляя их экспрессию в отсутствие пролина [25]. Белок RocG *Bacillus subtilis* пример фермента, который косвенно влияет на транскрипцию генов. Глутаматдегидрогеназа RocG катализирует реакцию дезаминирования глутамта с образованием альфа-кетоглутарата, но также взаимодействует с белком CltC – фактором транскрипции, который регулирует биосинтез глутамата из альфа-кетоглутарата, таким образом связывая два метаболических пути [26].

Еще один механизм регуляции экспрессии генов – регуляция с помощью РНК-переключателей (аттенюаторов) – цис- регуляторных элементов РНК, которые контролируют транскрипцию или трансляцию с помощью анти-терминации. Рибопереключатели располагаются в 5'-лидерной последовательности транскрипта и формируют стабильную шпильку (терминатор) в отсутствие эффектора. Если эффектор присутствует, он связывается с терминатором транскрипции, что приводит к структурным изменениям и продолжении транскрипции. В качестве эффектора могут выступать метаболиты [27], белки [28] и тРНК [29]. Помимо взаимодействия с молекулой эффектора, на работу РНК-переключателей может влиять температура, так как стабильность образованной вторичной структуры конечна [30].

### **1.3.2 Посттранскрипционная регуляция экспрессии генов**

Посттранскрипционная регуляция у бактерий, помимо деградации РНК РНКазами, включает в себя регуляцию с помощью малых РНК (sRNA). Действие этих sRNA можно разделить на несколько групп: маленькие, анти-смысловые РНК, которые связываются с мРНК-мишенью, влияя на инициацию трансляции, маскируя или раскрывая последовательность Шайна-Дальгарно своих мишеней, либо изменяя их стабильность [31]; sRNA, которые взаимодействуют с белками (например, ранее описанная 6S РНК). На стабильность РНК влияет полиаденилирование, которое приводит у бактерий к более быстрой деградации [32].

### **1.3.3 Посттрансляционная регуляция**

Посттрансляционная регуляция у бактерий осуществляется путем модификации белка или путем его деградации при взаимодействии с другим белком или РНК. Известно много разных модификаций белков – фосфорилирование, ацетилирование, биотинилирование, метилирование, гликозилирование. Эти модификации могут участвовать в передаче сигнала путем запуска ферментативной активности, например как ацетилирование у *Salmonella* [33]. Также к посттрансляционной регуляции можно отнести аллостерическую регуляцию белка

при взаимодействии с низкомолекулярными соединениями-эффе́кторами (как правило, такими способами регулируется активность ферментов, а также ДНК-связывающие свойства факторов транскрипции) [34], [35].

Таким образом, регуляция экспрессии генов может происходить на разных этапах и с разной скоростью, что в конечном итоге ведет к изменению содержания функционально активного продукта экспрессии гена. Аллостерическая регуляция имеет максимальную скорость ответа на внешние условия [35], ей лишь немного уступают механизмы регуляции, основанные на действии вторичных структур мРНК, а также малых РНК. Регуляция, основанная на действии специфических транскрипционных факторов, обеспечивает точный ответ на стресс, но является более медленным механизмом.

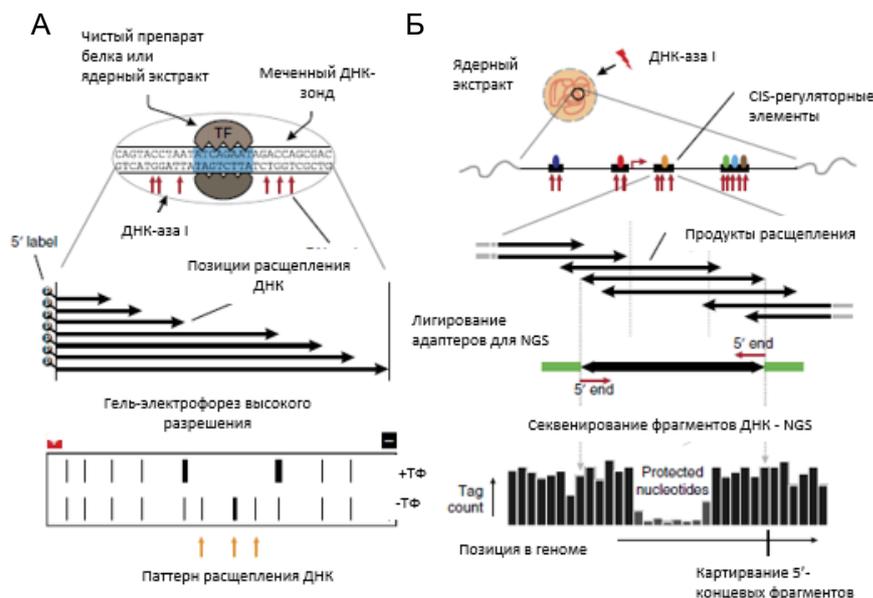
#### **1.4 Методы поиска регуляторов и их мишеней**

Развитие высокопроизводительных методов секвенирования привело к тому, что скорость накопления полногеномных данных секвенирования значительно опережает понимание того, как устроена регуляция генов в организмах. В попытке восстановить и описать регуляторную сеть бактерий, в которой участвуют белковые факторы транскрипции, глобально можно использовать два подхода: от ТФ к мишени и от мишени к ТФ. Первый подход предполагает наличие белка-кандидата в транскрипционный фактор. Среди экспериментальных методов поиска и подтверждения мишеней такого кандидата выделяют методы, для поиска одиночных мишеней – ДНК-футпринтинг, сдвиг электрофоретической подвижности (EMSA) и высокопроизводительные способы – SELEX, ChiP-chip, ChiP-seq, ChiP-exo

##### **1.4.1 Футпринтинг ДНК**

Этот метод появился в конце 70-ых годов XX века [36], и состоит из двух этапов: 1) воздействие на ДНК-белковый комплекс агентом, способным неспецифически разрушать или модифицировать ДНК, в то время как участок ДНК, связанный с белком не подвергается воздействию, 2) визуализация и детекция полученных паттернов расщепления или модификации с разрешением до единичных нуклеотидов. Таким образом можно узнать места в ДНК, с которыми связывается белок. Классическое исполнение футпринтинга ДНК – использование ДНКазы I в качестве расщепляющего агента и анализ результата с помощью секвенирования по методу Макса-Гилберта [36] (Рисунок 1.2 А). Со временем этот метод модифицировался (применялись другие агенты – гидроксил радикалы –  $Fe^{2+}$ -ЭДТА, алкилирующие ДНК агенты – диметилсульфат) и развился [37]. Использование футпринтинга ДНК для эукариот *in vivo* продемонстрировало роль *cis*-регуляторных элементов [38].

Кульминацией такого развития стало появления метода цифрового геномного футпринтинга ДНК [39]. Этот метод совмещает обработку ДНКазой I интактных ядер, выделенных из эукариотического организма, массивное параллельное секвенирование фрагментов ДНК, освобождённых в результате ферментативной обработки, и биоинформатическую обработку данных, давая количественную оценку относительных скоростей расщепления по каждому нуклеотиду по всему геному (Рисунок 1.2 Б).



**Рисунок 1.2 Принципы проведения ДНК футпринтинга.** А – Классический метод – очищенный белок или ядерный экстракт смешивали *in vitro* с радиоактивно меченным ДНК-зондом, обрабатывали образец ДНКазой I. Полученную смесь фрагментов разделяли с помощью гель-электрофореза высокого разрешения. Сравнивая паттерны расщепления ДНК-зонда в образце, где был добавлен белок-кандидат в ТФ и где белок не добавляли, можно определить место на ДНК, где связывается белок; Б – Цифровой геномный футпринтинг ДНК сочетает в себе обработку ДНК ДНКазой I, очистку маленьких фрагментов ДНК, полученных в результате расщепления, массивное параллельное секвенирование концов фрагментов (мест расщепления ДНКазой I), биоинформатический анализ данных для получения покрытия частот расщепления нуклеотидов по всему геному. Изображение адаптировано из [40].

Сравнивая распределение частоты расщепления нуклеотидов по всему геному можно получить позиции ДНК, которые защищены от расщепления вследствие связывания с белками или конформационных изменений [40]. Совмещая такие данные, например, с данными о начале старта транскрипции, анализируя мотивы, чьи последовательности обогащены в защищенных от ДНКазы местах генома, можно найти новые сайты связывания транскрипционных факторов [39].

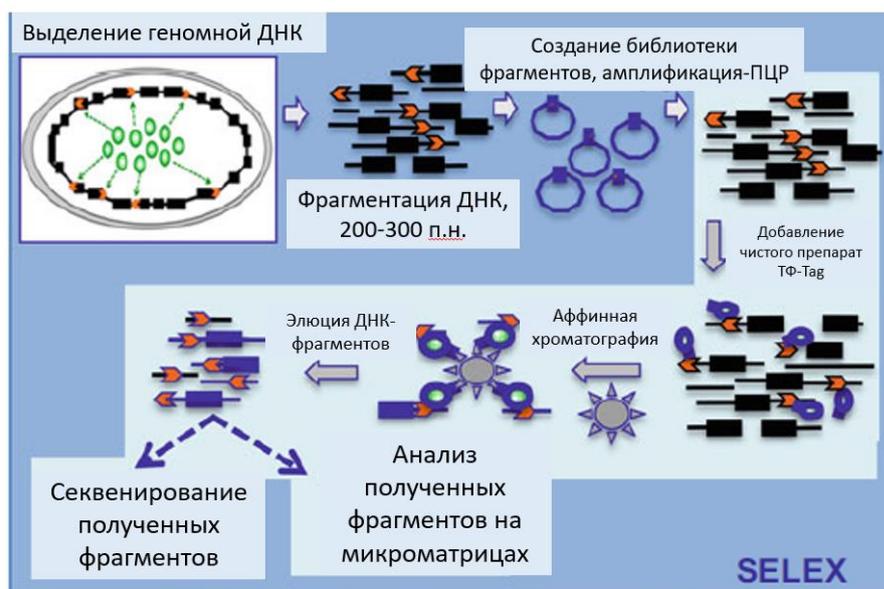
#### 1.4.2 Анализ сдвига электрофоретической подвижности (EMSA)

Анализ сдвига электрофоретической подвижности основан на том, что в нативных условиях подвижность комплекса ДНК-белок меньше, чем подвижность свободной ДНК. У

этого метода существует как ряд преимуществ, так и ряд ограничений. Этот метод прост в исполнении, позволяет проводить *in vitro* анализ влияния различных кофакторов и условий на силу взаимодействия белка с ДНК [41]. Использование меченных изотопами нуклеиновых кислот делает анализ высокочувствительным (концентрации используемых субстратов порядка 0.1 нМ и менее). Для анализа можно использовать как короткие последовательности, так и длинные, а также сложные структуры – триплексы, квадруплексы [42] и т.д.. Метод может быть применен как для очищенного белка, так и для смеси белков, например, клеточного экстракта [43], тогда для определения белка, который связывается с ДНК-субстратом применяют масс-спектрометрию [44] или вестерн-блот. Ограничения, связанные с применением этого метода, заключаются в возможной диссоциации комплекса при прохождении через гель, если константа связывания белка с ДНК-субстратом низкая, также некоторые ДНК-белковые комплексы стабильнее в геле, чем внутри клетки [45]. Помимо этого, EMSA не дает прямое указание о расположении последовательности, которая занята белком (в случае, если для связывания используют протяженный фрагмент ДНК). Эту информацию можно получить, применяя метод фузринтинга ДНК. Тем не менее EMSA остается довольно популярным методом подтверждения и поиска ДНК-связывающих белков [41].

#### **1.4.3 Систематическая эволюция лигандов экспоненциальным обогащением (SELEX)**

Метод SELEX основан на обогащении последовательностей нуклеиновой кислоты (РНК, дцДНК или оцДНК) с которыми связывается тестируемый белок из смеси всевозможных синтетических последовательностей нуклеиновых кислот, с последующим секвенированием [46], [47]. В случае поиска мотива связывания для ДНК-связывающих белков, библиотека состоит из синтетических фрагментов ДНК, количество уникальных последовательностей –  $4^n$ , где  $n$  – длина последовательности. Чем больше длина тестируемых последовательностей, тем сложнее получить эффективную концентрацию ДНК-субстрата для скрининга. Поэтому вместо синтетических последовательностей используют смесь фрагментов геномной ДНК, полученной путем неспецифического расщепления [48]. Akira Ishihama с коллегами развил классический метод SELEX и предложил новую модификацию [49], [50] (Рисунок 1.3).



**Рисунок 1.3** Схема проведения эксперимента SELEX. Поскольку сайты связывания ТФ исследуемого организма расположены в геноме этого организма, геномная ДНК выделяется и фрагментируется на участки длиной 200-300 п.о.. Затем создается библиотека фрагментов путем их клонирования в вектор, фрагменты нарабатываются методом ПЦР и используются как субстрат для связывания с выделенным белком-транскрипционным фактором, иммобилизованным на сорбенте. Элюируют фрагменты ДНК с последующим их секвенированием или анализе на ДНК-микрочипах. Зеленые круги – очищенный белок – ТФ, содержащий Тап из 6-остатков гистидина, необходимый для проведения аффинной хроматографии; оранжевые пятиугольники – места связывания белка. Серая восьмиконечная звезда – сорбент Ni-NTA агароза [50]

Безусловно этот метод не лишен недостатков. Необходимы априорные знания, о том, что белок является ТФ, об активирующих условиях для ТФ, кофакторах, влияющих на связывание белка с ДНК. Более того, метод дает большое количество ложноположительных результатов из-за высокого уровня неспецифического связывания белка с ДНК-субстратом.

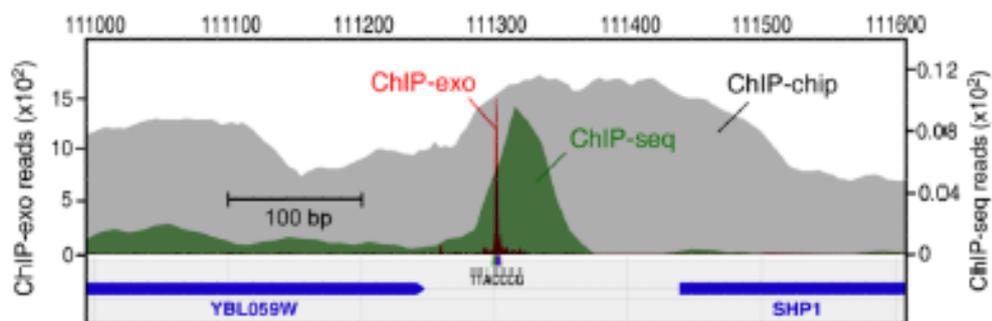
#### 1.4.4 ChiP-chip, ChiP-seq, ChiP-exo

Места взаимодействия ТФ с ДНК могут быть обнаружены с помощью иммунопреципитации хроматина (ChiP). Преимущество этого подхода заключается в его использовании в живых клетках, *in vivo* [51]. Отличие Chip методов – в способе детекции места связывания белка и основные усилия в модификации методики направлены на снижение фона, ложноположительных (сигнал от не связавшейся с белком фракции ДНК) и ложноотрицательных (если аффинность белка к определённым последовательностям мала, то сигнал будет слаб и можно принять его за шум) результатов, увеличение разрешения метода – если сайты связывания расположены близко друг к другу [52].

В экспериментах на основе иммунопреципитации хроматина транскрипционный фактор сшивается с ДНК в месте их взаимодействия, с помощью специфичных к ТФ антител обогащают целевые последовательности ДНК, с которыми был связан белок. Затем

обогатившие последовательности ДНК идентифицируют, соотносят с позициями на референсном геноме организма, чей ТФ исследовался, и определяют места связывания ТФ. В качестве сшивающего агента, как правило, используют формальдегид, который способен проникать внутрь клеток и образовывать сшивки между макромолекулами, в том числе между ДНК и белком. Небольшой размер формальдегида накладывает ограничения на расстояние между реакционными группами – 2 ангстрема, подходящее для сшивки двух взаимодействующих макромолекул [53].

Исторически, первым методом для широкомасштабных экспериментов по поиску сайтов связывания белков был метод на основе микроматриц ДНК (ДНК-микрочипов) – ChIP-chip [54]. Этот метод детекции сайтов связывания оказался не очень надежным, имел низкую чувствительность и разрешение [55], хоть иногда успешно применялся [56]. Следующим шагом к улучшению детекции мест связывания белков с ДНК стало применение технологии высокопроизводительного секвенирования – ChIP-seq [55]. Использование технологии секвенирования, в особенности секвенирования парных концов – ChIP-PET, позволило более точно картировать места связывания ТФ, повысить чувствительность метода [57]. Разработка метода ChIP-exo, где после этапа иммунопреципитации образец обрабатывают 5'-3' экзонуклеазой, которая расщепляет 5'-конец захваченной ДНК и останавливается на фиксированном расстоянии от места, где пришит белок [52] позволило еще точнее определять все места связывания белка с ДНК, значительно снизив количество ложноположительных и ложноотрицательных результатов. На Рисунке 1.4 изображено сравнение результатов идентификации сайтов связывания белка Reb1 пекарских дрожжей с помощью методов ChIP-chip, ChIP-seq и ChIP-exo на одном из локусов.



**Рисунок 1.4 Сравнение результатов ChIP-exo, ChIP-chip и ChIP-seq для Reb1 в определенных локусах.** Серый, зеленые и красный графики, соответственно, показывают распределение необработанных сигналов, измеренных с помощью ChIP-chip с использованием микрочипов Affymetrix, имеющих расстояние между зондами 5 п.н. [56], ChIP-seq и ChIP-exo. Метки секвенирования на каждой цепи были смещены в направлении 3'-конца на 14 п.н., чтобы максимизировать перекрытие противоположной нити

Дальнейшее развитие этой технологии направлено на упрощение подготовки образцов, в частности обогащенных библиотек ДНК-фрагментов для секвенирования (использование транспозазы Tn5 для присоединения адаптеров или лигирование адаптера к оцДНК после расщепления экзонуклеазой фага лямбда), снижение требуемого количества материала (количество клеток) для проведения анализа и стоимости [58].

Каждый из описанных методов имеет свои ограничения. Для применения методов EMSA и цифрового геномного футпринтинга ДНК необходимы априорные знания о ДНК-связывающих белках и условиях, при которых связывание белка с ДНК происходит, разрешение и чувствительность этих методов не всегда позволяет определить все мотивы, с которыми может связываться белок. Методы на основе хроматиниммуопреципитации сложны в исполнении и обработке данных, от чего напрямую зависит результат.

#### **1.4.5 Биоинформатические методы**

Для поиска потенциальных регуляторов и предсказания их сайтов связывания используют сравнительно-геномный анализ [59]. При аннотации генов или белков путем сравнения с последовательностями ранее изученных генов или белков других организмов на основе сходства можно сделать предположение о функции гена (белка). При поиске транскрипционных факторов у бактерий облегчает задачу наличие структурно-функциональных доменов. В настоящее время расшифровано более 30 структур прокариотических ДНК-связывающих белков, и большинство из них 84% имеют структурный мотив спираль–поворот–спираль (НТН) [21]. Однако, не все ТФ бактерий имеют аннотированные структурно-функциональные домены.

У бактерий многие гены организованы в опероны, и транскрибируются единой молекулой РНК. Коэкспрессия генов может быть обусловлена пространственной структурой ДНК в клетке, в результате чего даже удаленные на линейной молекуле гены оказываются сближенными в пространстве, согласно их транскрипционной активности [60]. Как правило, такие гены образуют единый функциональный контур, например, являются участниками одного метаболического пути [61].

Для поиска потенциальных сайтов связывания в биоинформатике используют метод филогенетического футпринтинга или ортологичных последовательностей [59]. Метод основан на предположении, что функционально важные участки ДНК консервативны, поэтому выравнивание промоторных областей большого набора ортологичных генов, для которых предполагается ко-регуляция, позволяет найти консервативные участки, которые, возможно, являются сайтами связывания ТФ [59]. Однако, у этого метода есть ограничения. В

частности, при наличии малой выборки последовательностей геномов, ее смещении в сторону анализа геномов близкородственных видов, или эволюционно далеких организмов [62] сложно получить информативную последовательность.

Данные о локализации сайтов связывания транскрипционных факторов необходимы для реконструкции регуляторных транскрипционных сетей. Для этого используют биоинформатические подходы, основанные на анализе полногеномных данных об экспрессии генов, организации оперонов, и метаболических путей, но для построения модели ответа клетки на воздействия необходимы не отрывочные знания об устройстве конкретного регуляторного пути, а полногеномные регуляторные сети. Несмотря на развитие экспериментальных и биоинформатических подходов для поиска регуляторов транскрипции, их сайтов связывания даже для модельных организмов, таких как *E. coli* и *B. subtilis* все еще нет ответов как регулируется и регулируется ли вообще половина генов в геноме. Например, в геноме одной из самой изученных бактерий *Escherichia coli* K-12 MG1655 около 30% генов (более 1000) не аннотировано [63], и из них 50-80 это предполагаемые не охарактеризованные транскрипционные факторы [64]. В своей работе [65] Y. Gao с коллегами предложили интегрированный подход, сочетающий в себе биоинформатические и экспериментальные методы для поиска и реконструкции регуляторных сетей не охарактеризованных факторов транскрипции *E. coli* K-12 MG1655, который также может быть применен для других бактерий. Этот подход включает несколько этапов:

1. Компьютерное предсказание кандидатов в транскрипционные факторы. Для этого используется программа TFpredict, основанная на машинном обучении. Используя гомологию первичных структур белков, алгоритм с некой вероятностью определяет является ли белок транскрипционным фактором.
2. Классификация кандидатов. Авторы выделили три класса белков: класс I: белки, чья способность связываться с ДНК была продемонстрирована в условиях *in vitro*; класс II: белки, которые охарактеризованы в близкородственных штаммах; класс III: кандидаты, для которых не показано ДНК-связывающих свойств.
3. Экспериментальная проверка кандидатов в ТФ на полногеномном уровне с помощью методов ChIP-ехо и RNA-seq, предсказание функции ТФ путем сравнения фенотипа клеток дикого типа и дефектных по гену кандидатного ТФ.

Стоит отметить, что в нормальных условиях культивирования, не охарактеризованные ТФ-ы, скорее всего, будут экспрессироваться на низком уровне и для того, чтобы увеличить их экспрессию, для каждого кандидата в ТФ, необходимо подбирать «активирующие» условия. Для *E. coli* накоплено большое количество данных транскриптомного [66] и

протеомного [67] профилирования в различных условиях, анализируя эти данные можно выяснить, в каких условиях наблюдается активация кандидата в ТФ и даже предположить мишени этого ТФ. Однако, даже в работе [65] Y. Gao для *E. coli* из 16 отобранных транскрипционных факторов для 6 не удалось выяснить определенного сайта связывания. Авторы предполагают несколько причин: в условиях при которых проводился эксперимент не произошла активация ТФ, либо, алгоритм, который предсказывает кандидаты в ТФ дает ложноположительные результаты.

Поиск новых факторов регуляции транскрипции и их мишеней, даже при наличии априорных знаний о ТФ, все еще остается трудоемкой задачей, для решения которой необходимо использование сочетания биоинформатических и молекулярно-биологических методов. Внедрение в исследование высокопроизводительных методов секвенирования, масс-спектрометрии и алгоритмов машинного обучения позволит решить эту задачу быстрее.

### **1.5 Особенности регуляции передачи генетической информации у бактерий с редуцированным геномом класса Mollicutes, на примере микоплазм**

Бактерии рода *Mycoplasma* входят в класс Mollicutes. Для них характерен малый размер генома с низким содержанием (около 30%) ГЦ пар нуклеотидов, отсутствие клеточной стенки, редукция метаболических и регуляторных путей [68], [69], [70].

ДНК-зависимая РНК полимераз микоплазм имеет сходное строение с РНК полимеразами большинства бактерий [71], но не чувствительна к рифампицину в следствие мутации в гене *rpoB* – бета-субъединицы РНК-полимеразы [72]. Известно о наличии главного сигма фактора. Первые данные об устройстве промотора и транскрипции у микоплазм были получены Ричардом Херманном и коллегами в середине 80-ых годов XX века. Изучая промотор оперона рибосомных генов Херманн показал, что в отличие от промотора оперона рибосомных генов *E. coli* промотор у *Mycoplasma hyopneumoniae* не имеет выраженного -35 участка, а терминация транскрипции осуществляется по Rho-независимому механизму [73]. В 1988 году Рон Гафни описал первую простую систему регуляции транскрипции у микоплазм. Промотор оперона рибосомных генов (*rrnA*) микоплазмы распознается РНК-полимеразой *E. coli*, однако повышение уровня аллармона ppGpp в клетках ведет к активации транскрипции с промотора *rrnA* микоплазмы, тогда как на экспрессию собственного рибосомного оперона ppGpp действует негативно [74].

Глобальные регуляторы скудно представлены у микоплазм. Так у некоторых видов известно о наличии одного альтернативного  $\sigma$  фактора [75], что возможно указывает на доминирование других механизмов ответа на стрессы. Описанный альтернативный сигма

фактор – MG428 *Mycoplasma genitalium* – активирует транскрипцию генов, участвующих в рекомбинации [75]. У некоторых микоплазм обнаружены белки теплового шока. У грамотрицательных бактерий ответ на тепловой шок регулируется альтернативными сигма-факторами. У грамположительных бактерий ответ на тепловой шок регулируется взаимодействием предполагаемого репрессора (гомолог HrcA) с CIRCE-подобным элементом, находящимся в промоторной области. Однако, CIRCE элемент найден не перед всеми генами теплового шока *Mycoplasma pneumonia*, что указывает на не универсальность этого механизма регуляции [76]. Работы, посвященные исследованию систем репарации у микоплазм, показали отсутствие белка-репрессора генов SOS- ответа LexA, характерного для большинства бактерий [77].

Представители класса молликут – *Mycoplasma genitalium* и *Mycoplasma pneumonia* были одними из первых бактерий, чьи геномы были полностью секвенированы. Это произошло в 1995 и 1996 годах, соответственно [78]. Тогда же началось активное развитие системной биологии, области, зародившейся в середине двадцатого века. Это связано с увеличением вычислительных мощностей, появлением высокопроизводительных методов анализа в биологии. Микоплазмы, как бактерии с редуцированным геномом, способные самореплицироваться на питательной среде, стали главными объектами системной биологии. Так, группы ученых, полагали расшифровать основные принципы организации и функционирования бактериальной клетки, на этих «просто устроенных» организмах.

В 2000 году основываясь на полных последовательностях геномов двух представителей класса Mollicutes было подтверждено отсутствие консенсусной последовательности -35 бокса в структуре промотора, при наличии консервативного -10 участка, а также наличие большого количества транскриптов, лишенных 5'-нетраслируемого участка мРНК, так называемых – безлидерных РНК. Тогда же была предпринята попытка составить матрицу для полногеномного поиска стартов начала транскрипции у микоплазм [79]. С наступлением XXI века в биологии все чаще стал применяться системный – омиксный подход, сначала для одного класса объектов, например – транскриптома, потом сразу для несколько измерений – транскриптом, протеом и т.д. Интеграция таких данных позволит создать модель, которая может описать и предсказать поведение живой системы [80].

В 2009 году группа во главе с Льюисом Серрано опубликовала серию работ, посвященную изучению транскриптома [70], протеома [81] и метаболома *Mycoplasma pneumoniae*. В работе использовались пертурбационные модели – стандартный подход для изучения бактерий. Ученые определили изменения транскриптома на микроматрицах в 62 условиях – при различной температуре культивации, осмотическом стрессе, в присутствии

этанола, при повреждениях ДНК, аресте клеточного цикла, на разных стадиях роста, в «голодной» среде, при добавлении разных источников углерода, воздействии антибиотиков. Амплитуда изменений уровня мРНК белок-кодирующих транскриптов не превышала пяти раз. Было найдено 117 ранее неизвестных транскриптов, большинство из них – 89 являются бессмысленными, комплементарными участкам функционально-смысловых мРНК. Число бессмысленных транскриптов у *M. pneumoniae* составляло 13% от числа белок-кодирующих транскриптов. Было обнаружено, что гены, локализованные на субопераонах, как правило, принадлежат к различным функциональным категориям. Показана неожиданная сложность организации регуляции транскрипции у бактерий с редуцированным геномом, причем наличие предсказанных девяти факторов регуляции транскрипции не могло объяснить этого феномена. Основные обнаруженные особенности транскрипции у микоплазм – гетерогенность экспрессии внутри одного оперона, изменение структуры оперона, ведущее к синтезу альтернативных транскриптов в ответ на изменение условий и наличие большого числа бессмысленных молекул РНК [70]. Фисунов Г.Ю. с соавторами подробно описал структуру промотора микоплазм, выделив, как основные элементы, так и расстояние между ними. Оптимальный промотор микоплазм состоит из -10 бокса вида TAWAAT (W=A или T). Инициаторный нуклеотид вносит большой вклад в частоту инициации транскрипции, наиболее предпочтительными являются G или A. Длина спейсера между -10 боксом и стартом начала транскрипции – 6 п.н., для небольшого числа промоторов это значение составляет 5 или 7. На «силу промотора», то есть частоту инициации транскрипции влияет окружение промотора. Сильные промоторы микоплазм имеют классические для бактерий EHT-элемент -10 бокса вида TRTG (R=A или G) со спейсером в один нуклеотид [82].

В 2016 году было проверено предположение об участии бессмысленных транскриптов в регуляции транскрипции у микоплазм. В результате, не удалось выявить общие детерминанты, определяющие уровень экспрессии бессмысленных транскриптов и их влияние на представленность и время жизни белок-кодирующих РНК [83]. В 2017 году той же группой ученых после 8 лет изучения бактерии *M. pneumoniae*, в том числе регуляции транскрипции и трансляции, была предложена модель регуляции, где нет классических транскрипционных факторов [84]. Авторы описали три уровня регуляции. Первый уровень – геномный. Здесь основным фактором является структура хромосомы, организованная в контактирующие домены и белки, ассоциированные с нуклеоидом. Немалую роль играют супервитки ДНК, а также гиразы и топоизомеразы, которые регулируют их образование и разрешение. Уровень АТФ в свою очередь может влиять на активность АТФ-зависимых топоизомераз. Второй уровень – оперонный. Здесь регуляция может осуществляться за счет инициации и

терминации транскрипции – влияет последовательность промотора, первый нуклеотид транскрипта (A/G в +1 позиции транскрипта влияет на скорость инициации), как показал Фисунов Г.Ю.. Внутри оперона, состоящего из нескольких генов, уровень мРНК падает при отдалении от точки инициации транскрипции. В последовательности ДНК существуют внутренние терминаторы, которые могут срабатывать или нет в зависимости от внешних условий [82]. И третий уровень – уровень РНК. На этом уровне вклад вносят структуры РНК – такие как рибопереключателы, а также малые РНК. Малые РНК могут контролировать деградацию мРНК, защищая их от рибонуклеаз, а также обеспечивать связывание рибосомы с конкретной мРНК [84]. мРНК, синтезируемые РНК-полимеразой не обязательно транслируются рибосомами в белки. Регуляция инициации трансляции может происходить с помощью белков рибосомы, а также модификации рибосомальных белков. Одна из последних работ группы Л. Серрано посвящена организации регуляторной сети у бактерий *M. pneumoniae* [85]. В этой работе, в лучших традициях системной биологии, авторы объединили накопленные данные своих многочисленных экспериментов. На первом этапе с помощью ДНК-аффинной хроматографии с последующим масс-спектрометрическим анализом были определены ДНК и/или РНК-связывающие белки *M. pneumoniae*, таких белков оказалось 174. Для 128 был выполнен ChiP-seq эксперимент, для 23 белков были найдены специфические пики – предположительные сайты связывания этих белков, среди этих белков 6 – предполагаемые ТФ-ы. Затем были получены штаммы *M. pneumoniae* со сверхэкспрессией или инактивацией предполагаемых ДНК-связывающих белков, в большинстве случаев не наблюдалось корреляции между экспрессией регулятора и уровнем генов-мишеней, найденных с помощью технологии ChiP-seq. Был проведен транскрипционный анализ при различных воздействиях на клетки *M. pneumoniae* (37 состояний). Используя данные о мишенях транскрипционных факторов, наличии рибопереключателы, топологии ДНК, позиции старта транскрипции, скорости деградации РНК, структуре промоторов и еще 11 различных детерминант, с помощью алгоритма random forest была построена модель, предсказывающая уровень экспрессии генов в различных состояниях *M. pneumoniae*. В результате было обнаружено, что действия ТФ-ов могут объяснить в среднем лишь 9% изменений на уровне мРНК в любом эксперименте, в то время как 36% изменений могут быть результатом действия других регуляторных механизмов. Таким образом, хотя некоторые транскрипционные факторы могли быть упущены в исследовании, вывод о том, что регуляция транскрипции у *M. pneumoniae* в основном осуществляется не ТФ-ами является верным по мнению авторов работы [85]. Стоит отметить, что эта работа была выполнена после того, как были опубликованы научные статьи нашей группы, посвящённые ТФ-ам и регуляторным сетям *M. gallisepticum*. Часть этих результатов описаны в данной диссертации.

Биосинтез белков сложный процесс, в котором задействовано большое количество участников. Для микоплазм, более продукты более 25% генов участвует в трансляции [86]. Рибосома микоплазм в целом имеет классическую структуру, характерную для рибосом прокариот. Сравнительный анализ аппарата трансляции у 39 представителей молликут, а также модельных бактерий, показал консервативность 104 генов, что составляет 40% и 49% от числа генов системы трансляции *E. coli* и *B. subtilis*, соответственно [87]. Наиболее консервативными оказались рибосомные белки, аминоацил-тРНК-синтетазы, факторы трансляции, но и здесь есть исключения. У молликут отсутствует пять генов рибосомных белков – S14b/RpsNb, S22/RpsV, L31b/RpmEb, L7b/RplGb и L25/RplY, впрочем, как и большинства бактерий. Набор из этих белков обеспечивает гетерогенность рибосом в ответ на стресс [88]. Почти все молликуты, кроме бактерий кластера *Nominis*, лишены белка S1 (RpsA), который участвует в сборке 30S субъединицы рибосомы на последовательности Шайна — Дальгарно. При это трансляция безлидерных мРНК эффективно осуществляется и без S1 белка, посредством сборки целой рибосомы в районе старт-кодона. Другая интересная особенность большей части (11 из 52) рибосомных белков микоплазм – наличие бесструктурных концевых участков, которые отсутствуют в гомологичных последовательностях *E. coli*. Функции этих участков неизвестны, но их повреждение влияет на приспособленность и выживаемость клеток. Факторы трансляции обеспечивают скорость и точность синтеза белка, большинство из них встречается у молликут, исключение составляет ГТФ-аза ТурА (VirA), которая ассоциирована с рибосомами во время стресса. Микоплазмы утратили ген фактора терминации RF2, необходимый для считывания стоп-кодона UGA, а также систему *selA*, *selB*, *selC* и *selD*. Кодон UGA у молликут кодирует триптофан. У микоплазм высокая вероятность ошибочного включения аминокислот при синтезе белка, вследствие упрощения систем редактирования. Так у этих бактерий отсутствуют белки семейства YbaK, ProX или AlaX, а редактирующий домен некоторых аминоацил-тРНК-синтетаз имеет мутации, что ведет к некорректному синтезу [89]. Менее всего консервативны гены, кодирующие ферменты, участвующие в посттранскрипционной модификации т и рРНК, в биогенезе и созревании рибосом – включая посттрансляционные модификации рибосомных белков и рибонуклеазы, которые процессируют т, р и мРНК. У бактерий транскрипция и трансляция сопряжены. Для *M. pneumoniae* показано, что фактор терминации NusA, который стимулирует регуляторные паузы в процессе транскрипции, а также усиливает Rho-независимую терминацию, может действовать как сенсор РНК-полимеразы, который обнаруживает приближающуюся рибосому и модулирует элонгацию транскрипции [90]. Регуляция трансляции может происходить на любом из этапов. Уровень трансляции определяется доступностью сайта связывания рибосомы на мРНК, природой стартового

кодона, положением последовательности Шайна-Дальгарно (SD) относительно стартового кодона, его комплементарностью анти-SD последовательности 16S рРНК и наличием А/Т богатых последовательностей. На стадии элонгации регуляции трансляции осуществляется за счет пауз, длительность которых в свою очередь зависит от доступности субстрата, действия факторов элонгации. На стадии терминации – за счет скорости рециркуляции рибосом [91].

Таким образом, регуляции передачи генетической информации у микоплазм имеет ряд особенностей – мало представлены глобальные регуляторы и классические ТФ-ы, что может быть следствием редукции метаболических путей и известных механизмов борьбы со стрессом, экспрессия участников которые, как правило, находится под регуляцией. Высокая гетерогенность экспрессии внутри оперона, синтез большого количества анти-смысловых транскриптов. На уровне трансляции, регуляция, опосредованная действием факторов трансляции, хорошо изученных на *E. coli*, может быть реализована иначе, за счет отсутствия некоторых участников. Нам представляется особенно важным провести исследование по поиску новых транскрипционных факторов в бактерия с редуцированным геномом, выявлению их основных мишеней, поскольку это существенно продвинет исследователей в понимании организации и регуляции наиболее «просто устроенных» живых клеток, что в дальнейшем может послужит основой для создания искусственных систем.

## **2 МАТЕРИАЛЫ И МЕТОДЫ**

### **2.1 Поиск потенциальных факторов регуляции транскрипции у бактерий класса Mollicutes**

Для поиска потенциальных факторов регуляции транскрипции было выбрано 50 видов бактерий класса Mollicutes. Последовательности были извлечены из базы Genbank. Для каждого вида были получены наборы последовательностей белков в формате .fasta. Для поиска ортологов использовали программу Proteinortho, BLAST. Поиск функциональных, консервативных доменов проводили с помощью CD-Search (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) и InterPro (<https://www.ebi.ac.uk/interpro/>). Затем белки были отсортированы по наличию ДНК-связывающих доменов, аннотации, приписанной найденным доменам. Помимо транскрипционных факторов, связываться с ДНК могут ДНКазы, ДНК-зависимая РНК-полимераза и т.п. Такие белки были удалены из списка потенциальных факторов регуляции транскрипции.

### **2.2 Оценка консервативности белков и геномного контекста**

Множественное выравнивание последовательностей проводили с помощью программы MUSCLE [92]. Для построения и визуализации филогенетических деревьев максимального правдоподобия, ближайших соседей, а также проведения бутстреп анализа использовали программы MEGA [93], PHYLIP. Последовательности сайтов связывания для изученных регуляторов были получены из базы данных RegPrecise [94]. Лого последовательностей были созданы с использованием пакета WebLogo. Геномный контекст оценивали, сравнивая расположение интересующих генов в геноме разных организмов с помощью геномного браузера VAC-browser. [95].

### **2.3 Поиск сайтов связывания потенциальных транскрипционных факторов MraZ, HsdC, Fur, WhiA, ParX**

Используя наши ранее полученные данные о сайте начала транскрипции для *M.gallisepticum* S6 [96], промоторные области были извлечены и проанализированы на наличие повторов – типичных сайтов связывания транскрипционных факторов и консервативность среди других представителей молликут. Для каждого найденного мотива, была составлена нуклеотидная позиционная матрица весов, которую использовали для поиска дополнительных мишеней ТФ-а, членов регулона.

### **2.4 Питательные среды и культивирование бактерий**

В качестве жидкой питательной среды для культивирования *E. coli* использовали среду LB (на 1 литр): 10 г триптона; 5 г дрожжевого экстракта; 10 г NaCl. Доводили pH среды до 7.4 добавлением 2 М раствора NaOH. Для получения твердых сред в LB добавляли 2% агара.

*Mycoplasma gallisepticum* S6 культивировали в жидкой питательной среде согласно [97] (на 1 литр): 20 г триптозы; 5 г NaCl; 3 г Трис; 1,3 г KCl, pH 7.4 с добавлением 1% глюкозы, 50 мл 10% жидкого дрожжевого экстракта, 100 мл лошадиной сыворотки (термически инактивированной при 56°C в течение 30 мин), 200 ед/мл пенициллина, 0.002% индикатора фенолового красного. Для приготовления полужидкой среды, добавляли 0,3% агара. Определение сублетальных доз 2,2'-дипиридила для *M. gallisepticum* проводили следующим образом. Аликвоты суспензионной культуры клеток *M. gallisepticum* в логарифмической фазе роста подвергались различным дозам воздействия 2,2'-дипиридила в течение 30 мин. Затем с помощью метода кратных разведений и изменению цвета индикатора среды, определяли титр жизнеспособных клеток. Используемые концентрации хелатирующего агента – 1,5 мМ; 3 мМ; 4.5 мМ, 6 мМ; 7,7 мМ. Определенная сублетальная доза 2,2'-дипиридила для *M. gallisepticum* составила – 3 мМ.

## 2.5 Выделение геномной ДНК

Клеточный осадок *M.gallisepticum*, полученный центрифугированием 1 мл 12 часовой культуры (середина логарифмической фазы роста) при 10000 g 4°C в течение 10 мин., лизировали в 500 мкл СТАВ буфера (состав приведен ниже), после чего образцы инкубировали при 60°C в течение 30 мин. Затем к образцам добавляли 500 мкл хлороформа, активно встряхивали и центрифугировали при 16000 g, 4°C, в течение 20 минут. Водную фазу отбирали в отдельную пробирку и осаждали нуклеиновые кислоты добавлением равного количества изопропилового спирта и 5% по объему ацетата натрия (3М ацетата натрия, pH 5.2). Смесь инкубировали при -20°C в течение как минимум часа и центрифугировали при 16000 g, 4°C, в течение 20 минут. Полученный осадок промывали раствором 80% этилового спирта. Осадок растворяли в деионизованной воде, измеряли концентрацию флуориметрически на флуориметре Qubit (Invitrogen, США) согласно инструкции производителя и использовали как матрицу для ПЦР.

СТАВ буфер: 2% СТАВ; 100 мМ Трис-HCl pH=8.0; 20 мМ ЭДТА; 1,4 М NaCl)

## 2.6 Компетентные клетки и трансформация *E.coli*

Химически компетентные клетки *E. coli* получали по модифицированной методике, описанной в [98]. Кратко, 2 мл ночной культуры клеток *E. coli*, штамм TOP10 (Invitrogen, США) пересеивали в 100 мл свежей LB-среды и инкубировали при 37°C на качалке (140 об./мин) до достижения оптической плотности 0,5 при длине волны – 600 нм. Стерильно

делили культуру на две пробирки по 50 мл и осаждали клетки центрифугированием 3500 g в течение 10 мин при 4°C. Тщательно отбирали надосадочную жидкость и стерильно ресуспендировали клетки в 17 мл раствора RF1 (состав приведен ниже), инкубировали 15 мин во льду. Затем клетки осаждали, отбирали супернатант, ресуспендировали в 4 мл раствора RF2 (состав приведен ниже), инкубировали 15 мин и разделили на аликвоты по 200 мкл. Хранили при -75°C до востребования.

Раствор RF1: 100 mM RbCl<sub>2</sub>; 50 mM MnCl<sub>2</sub>\*4H<sub>2</sub>O; 10 mM CaCl<sub>2</sub>\*2H<sub>2</sub>O; 30 mM ацетата калия; глицерин 15% (об/об). Доводили pH полученного раствора ледяной уксусной кислотой до 5.8. Полученный раствор стерилизовали путём фильтрования через 0,22 μM фильтр (Milipore).

Раствор RF2: 10 mM MOPS; 75 mM CaCl<sub>2</sub>\*2H<sub>2</sub>O; 10 mM RbCl<sub>2</sub>; глицерин 15% (об/об). Доводили pH полученного раствора 10 M NaOH до 6.5.

Полученные раствор стерилизовали путём фильтрования через 0,22 μM фильтр (Milipore, США). Хранили при -20°C, перед использованием размораживали при +4°C.

Химическую трансформацию компетентных клеток проводили следующим образом. Около 10-20 нг плазмидной ДНК (или 5 мкл лигазной смеси) добавляли к 50 мкл компетентных клеток и инкубировали при 4°C, 10 минут. После этого клетки подвергали тепловому шоку при 42°C в течение 2 мин., затем помещали в лёд на 2 минуты, далее добавляли 500 мл предварительно прогретой до 37°C среды LB и инкубировали 40 мин при 37°C, после чего высевали на чашки Петри с твердой питательной средой, содержащей ампициллин.

## **2.7 Выделение плазмидной ДНК щелочным гидролизом**

Ночную культуру *E. coli* объёмом 7-10 мл центрифугировали 5 мин. при 5000 g и 4°C. Супернатант удаляли, осадок клеток суспендировали в 150 мкл буфера 20T, затем добавляли 250 мкл р-ра «А», осторожно перемешивали, наклоняя пробирку, и выдерживали 4 мин при 4°C (составы растворов приведены ниже). После этого добавляли 350 мкл 3 M ацетата натрия pH 5.2, смесь инкубировали 10 мин при 4°C. Появлялся хлопьевидный осадок, который удаляли центрифугированием в течение 20 мин при 12000 g 4°C. Супернатант переносили в чистую пробирку и добавляли равный объём изопропилового спирта. Смесь инкубировали при -20°C в течение как минимум часа и центрифугировали при 16000 g, 4°C, в течение 20 минут. Полученный осадок промывали раствором 80% этилового спирта. Осадок растворяли в 50 мкл деионизованной воды, добавляли 0,5 мкл рибонуклеазы А (Sigma-Aldrich, США), инкубировали 30 мин при комнатной температуре. Очищали плазмиду коммерческим набором GeneJet Kit (Thermo Fisher Scientific, США), согласно инструкциям производителя.

Буфер 20T: 20 mM Трис-HCl, pH 7.5

Раствор «А»: 0,2 M NaOH; 1% SDS

## **2.8 Получение чистого препарата целевых белков из *M. gallisepticum* в клетках *E. coli***

В работе получали очищенные препараты белков MraZ, HsdC (Xre), WhiA, ParX, Fur *M. gallisepticum* S6. Для ПЦР-амплификации фрагментов генов с целью последующего клонирования использовали смесь полимераз Taq (НПФ «Литех», Россия) и Pfu (Thermo Fisher Scientific, США). В качестве матрицы использовали геномную ДНК *M. gallisepticum* S6 (п. 2.5). Праймеры подбирали исходя из последовательности генов, на концах праймеры содержали сайты для эндонуклеаз рестрикции SalI и BamHI. Список праймеров представлен в приложении, Таблица П1. Амплификацию проводили на приборе Tetrad 2 Thermal Cycler (Bio-Rad, США), условия были подобраны, согласно рекомендациям производителя полимераз. ПЦР-продукт очищали экстракцией хлороформом и преципитацией изопропанолом с добавлением ацетата натрия. Размер продукта, его чистоту и количество оценивали путём электрофореза в агарозном геле с окрашиванием бромистым этидием. Полученный ПЦР-продукт был подвергнут обработке эндонуклеазами рестрикции SalI и BamHI (Thermo Fisher Scientific, США), после чего проводили лигазную реакцию с вектором pETmin (Novagen, Германия) или pET15, предварительно подвергнутым обработке теми же эндонуклеазами.

После реакции лигирования вектора и целевого гена, полученным образцом трансформировали компетентные клетки *E. coli* (см. п. 2.6), штамм TOP10 (Invitrogen, США), высевали на чашку с LB агаром и ампициллином, 100 мкг/мл. После 15 часов инкубации при 37°C, колонии *E. coli* скалывали и проводили скрининг методом ПЦР. Длины продуктов оценивали электрофорезом в агарозном геле с окрашиванием бромистым этидием. Колонии бактерий, содержащие плазмиду с интересующим нас геном высевали в жидкой среде LB с добавлением ампициллина 100 мкг/мл, растили при 37°C и постоянном покачивании (140 об./мин) в течение ночи. Затем клетки осаждали (5000 g в течение 5 мин при 4°C) и выделяли плазмиду щелочным гидролизом (п. 2.7). Измеряли концентрацию плазмиды флуориметрически на флуориметре Qubit (Invitrogen, США) согласно инструкции производителя. Нуклеотидная последовательность клонированного гена была определена по каждой из цепей методом секвенирования по Сэнгеру в лаборатории постгеномных исследований в биологии ФГБУ ФНКЦ ФХМ ФМБА России. Для выделения целевого белка, штамм B834 (DE3) *E. coli* (Life Technologies, США) трансформировали полученной плазмидой. Клетки инкубировали в 300 мл LB-среды, содержащей 100 мкг/мл ампициллина при 37°C на качалке (140 об./мин) до достижения оптической плотности 0,8 при длине волны

– 600 нм. Затем для индукции экспрессии добавляли индуктор ИПТГ до финальной концентрации 0,1мМ и инкубировали клетки еще 3 часа. Клетки осаждали и разрушали ультразвуком. Целевой белок очищали с использованием набора для аффинной хроматографии HisTrap histidine-tagged protein purification kit (Amersham Biosciences, Германия) в соответствии с протоколом производителя. Концентрация полученного белка была измерена по методу Бредфорда с помощью Quickstart Bradford dye reagent (BioRad, США). Целостность полученного белка оценивали с помощью гель-электрофореза по методу Лэммли [99].

## 2.9 Электрофорез белков в полиакриламидном геле по методу Лэммли

Для анализа белков применяли систему денатурирующего ступенчатого электрофореза [99]. В качестве концентрирующего геля использовали 4% ПААГ (соотношение акриламида и метиленбисакриламида 29:1; 0,125 М Трис-НСl, рН 6.8; 0,1 % персульфата аммония; 0,1 % ТЕМЕДа), в качестве разделяющего – 12% ПААГ (0,375 М Трис-НСl, рН 8.8; 0,05 % персульфата аммония; 0,05 % ТЕМЕДа). Оба геля содержали ДСН (SDS) в концентрации 0,1%. К исследуемым образцам добавляли равный объем буфера для образцов (состав приведен ниже). Перед нанесением на гель пробы прогревали при 95 °С, 6 мин. и центрифугировали 5 мин. при 16000 g. Электрофорез проводили при напряжённости электрического поля 10 В/см в буфере для электрофореза (состав приведен ниже). По окончании электрофореза блок геля окрашивали с использованием 0,4% Кумасси-R250 в растворе, содержащего изопропанол, этанол, уксусную кислоту и дистиллированную воду в соотношении 2:1:1:6. Отмывку осуществляли 10% раствором уксусной кислоты при нагревании.

Буфер для образцов: 0,25 М Трис-НСl рН 6.8; 4% ДСН; 10% глицерина; 0,004% бромфенолового синего; 5% ДТТ

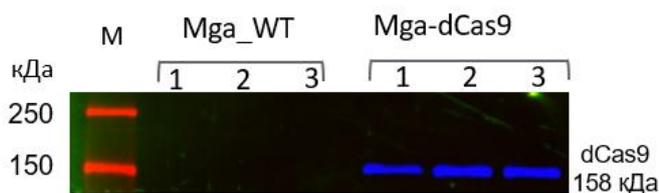
Буфер для электрофореза: 25 мМ Трис; 0,25 М глицина; 0.1% ДСН

## 2.10 Вестерн-блот анализ

Для получения клеточного лизата, 20 мл суспензионной культуры *M. gallisepticum* дикого типа, а также трёх трансформантов, содержащих *dcas9* *S. pyogenes*, осаждали центрифугированием 10000 g 10 мин 4°C. Надосадочную жидкость отбирали, клеточный осадок ресуспендировали в 10 мл охлажденного буфера (20 мМ Tris-НСl рН 7.4, 150 мМ NaCl, 3 мМ MgCl<sub>2</sub>), снова осадил и ресуспендировали 200 мкл того же буфера. После добавляли 5 мкл 10% детергента Nonidet P-40 (Sigma-Aldrich, США), тщательно перемешивали. Клеточный лизат замораживали при –75°C не менее 1 ч. Полученный клеточный лизат осветляли центрифугированием при 20000 g в течение 20 мин при 4°C. Супернатант отбирали,

измеряли концентрацию белка по методу Брэдфорд. Анализируемые образцы, содержащие 20 мкг белка, разделяли в 6-12% градиентном денатурирующем ПААГ по Лэммли. В качестве маркера масс наносили 5 мкл Precision Plus Protein™WesternC™ Blotting Standards (Bio-Rad, США). Переносили белки на PVDF мембрану (Bio-Rad, США) методом полусухого переноса в приборе Trans-Blot Turbo Transfer System (Bio-Rad, США). Мембрану дважды промывали в буфере TBS, инкубировали в 3 % растворе бычьего сывороточного альбумина в течение 30 мин в буфере TBST (TBS, содержащий 0,1% Tween-20) и инкубировали с первичными антителами Cas9 Monoclonal Antibody 10C11-A12 (Invitrogen, США) в разведении 1:1000 в течение 1 часа. Мембрану отмывали в буфере TBST, 3 раза по 10 мин и инкубировали со вторичными антителами Anti-Mouse IgG (Fc specific)-Peroxidase, Cat. # A2554 (Sigma-Aldrich, США), в разведении 1:100000 в течение 1 часа, отмывали мембрану в буфере TBST 3 раза. Визуализировали целевые белковые полос с помощью субстрата пероксидазы хрена Amersham™ ECL Western Blotting Detection Reagents (Little Chalfont, Великобритания) в системе детекции люминесценции ChemiDoc MP system (Bio-Rad, США) (Рисунок 2.1).

Буфер TBS: 100 mM Tris-HCl, 150 mM NaCl



**Рисунок 2.1** Вестерн-блот анализ экспрессии dCas9 *S. pyogenes* в трансформантах *M. gallisepticum*. Mga\_WT – лизат клеток *M. gallisepticum* дикого типа, приведен в качестве контроля; Mga-dCas9 – лизат клеток трансформантов, экспрессирующих dCas9 *S. pyogenes*, M - маркер молекулярных масс. Опубликовано в [100]

## 2.11 Метод торможения ДНК в геле (EMSA)

Связывание очищенных белков *M. gallisepticum* с фрагментами дцДНК было протестировано с помощью торможения в полиакриламидном геле в нативных условиях [41]. Белок, связываясь с ДНК-субстратом, тормозит его миграцию сквозь полиакриламидный гель в процессе электрофореза. При этом в качестве контроля служит образец фрагмента дцДНК без добавления белка. Различные количества чистого препарат белка (от 0,5 до 40 пмоль) инкубировали с субстратом – фрагментом дцДНК (длиной от 20 до 40 п.о). В качестве субстрата использовали либо 0,1 – 5 пмоль (50 – 250 нМ) фрагмента дцДНК, меченного с одного из 5'-концов флуоресцеином (FAM), либо 5-20 пмоль (250-1000 нМ) немеченного фрагмента дцДНК с последующей окраской геля 1% раствором бромистого этидия. К пробам добавляли глицерин (8 % от общего объема образца), и буфер 20 mM Трис-HCl pH 8,0, инкубировали в течение 15 мин при 37°C. Общий объем каждого образца составлял 20 мкл.

Перед нанесением образцов проводили префорез – 20 В/см при 10°C в течение 30 мин. Образцы вносили в лунки 6% или 8% полиакриламидного геля (200x180x1 мм), в качестве буфера использовали 100 мМ Трис-борат pH 8.3, электрофорез проводили в течение 1 часа при 10°C и напряженности электрического поля 40 В/см.

Для белка WhiA реакцию связывания с дцДНК проводили в 20 мМ карбонатном буфере pH 9.8, электрофорез в 100 мМ Трис-борат (pH 10.0). При стандартных условиях проведения эксперимента, белок WhiA агрегирует, в силу своих физико-химических свойств.

В качестве кофакторов для белка WhiA были протестированы нуклеотиды: АМФ, ГМФ, ЦТФ, УМФ, АДФ, ГДФ, ЦДФ, АТФ, ГТФ, УТФ и ЦТФ. Количество АТФ и АДФ варьировали от 0,1 до 5 мМ.

После электрофореза проводили детекцию ДНК и ДНК-белковых комплексов в геле путем измерения флуоресценции красителя FAM (длины волн возбуждения и эмиссии – 494 нм и 528 нм, соответственно), либо бромистого этидия в комплексе с ДНК (длины волн возбуждения и эмиссии – 300 нм и 605 нм, соответственно) при помощи сканера Typhoon 9410 (Amersham, Германия). Интенсивность флуоресценции оценивали с помощью программного пакета ImageQuant (GE Healthcare, США).

### **2.12 Расчет константы диссоциации**

Для расчета константы диссоциации комплекса, образцы, содержащие равное количество дцДНК-субстрата (250 нМ) и разное количество белка (от 0 до 1435 нМ) подвергали разделению в нативных условиях (EMSA), а затем оценивали интенсивности флуоресценции несвязанной с белком ДНК и ДНК в комплексе с ТФ с помощью программного пакета ImageQuant (GE Healthcare, США). Экспериментальные данные интенсивности флуоресценции были аппроксимированы функцией уравнения  $F=1/(1+(Kd/P)^h)$ , где F – доля субстрата, в комплексе с белком; Kd – константа диссоциации; h – коэффициент Хилла [101]. Значение Kd и h представлены как среднее значение трех экспериментов со стандартным отклонением

### **2.13 Эксклюзионная хроматография**

Измерение стехиометрии ДНК-белкового комплекса и определение олигомерного состояния белка MraZ было выполнено с помощью эксклюзионной хроматография. Для проведения хроматографии колонка Tricorn 10/ 300 была заполнена носителем – Superdex 200 (GE Healthcare, США), уравновешена буфером, содержащим 20 мМ Трис-HCl, 150 мМ KCl, 5 мМ MgCl<sub>2</sub>, pH 7.5. Скорость потока составляла 2 мл/мин. Колонку калибровали с помощью набора LMW Gel Filtration Calibration Kit (GE Healthcare, США), согласно рекомендациям производителя. Образец объемом 330 мкл, содержащий 230 мкг белка MraZ (14 нмоль) и 1

нмоль дцДНК подвергали разделению на колонке. Пробы, содержащие только белок или только ДНК-субстрат, были использованы в качестве контроля. Разделение проводили с помощью хроматографической системы АКТА FPLC (GE Healthcare, США). Выполнено совместно с Манувера В.А. (лаб. генной инженерии бактерий ФНКЦ ФХМ при ФМБА России).

**Таблица 2.1** Расчетные значения для построения калибровочной кривой для хроматографической колонки Tricorn 10/ 300, заполненной носителем – Superdex 200

Mr, кДа	V <sub>элюции</sub> , мл	K <sub>av</sub>	LgMr
75	15,26	0,465	1,875
44	16,63	0,542	1,643
29	17,67	0,643	1,462
13,7	19,10	0,748	1,137
6,5	20,81	0,874	0,813

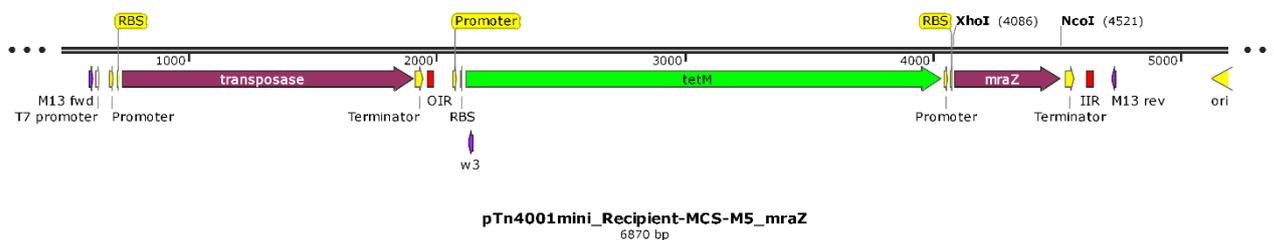
#### 2.14 Измерение уровня АТФ

Концентрацию АТФ в образцах определяли методом биолюминесцентной АТФ-метрии. Для этого готовили пробы объемом 200 мкл, которые содержали 50 мМ калий-фосфатного буфера (рН 7.8), 5 мМ MgCl<sub>2</sub>, 5 мкл смеси люциферин-люцифераза (Lumtek, Россия) и 1 мкл тестируемого образца. В качестве образцов использовали 20 мкл смеси, содержащей 1 мМ АТФ, 1 пмоль дцДНК в 20 мМ карбонатного буфера рН 9.8 с добавлением или без 20 пмоль белка WhiA. Затем образцы инкубировали в течение 60 мин при 37°C и 1 мкл добавляли в пробы для измерения уровня АТФ. Измерения проводились в трех повторностях. Измерение сигнала проводили на хемилюминометре Lum 1200 (DiSoft). Калибровочная кривая для оценки количества АТФ была построена с использованием образцов, с известной концентрацией АТФ.

#### 2.15 Конструирование векторов для сверхэкспрессии генов в *M. gallisepticum*

ДНК фрагмент, кодирующий интересующий ген (*mraZ*, *fur*) амплифицировали с использованием пары олигонуклеотидов (Таблица П1 Приложения) и геномной ДНК *M. gallisepticum* в качестве матрицы. Амплификацию проводили на приборе Tetrad 2 Thermal Cycler (Bio-Rad, США). ПЦР-продукт очищали экстракцией хлороформом и преципитацией изопропанолом с добавлением ацетата натрия. Длину продуктов, их чистоту и количество оценивали путём электрофореза в агарозном геле с окрашиванием бромистым этидием. Полученный ПЦР-продукт подвергали обработке эндонуклеазами рестрикции XhoI и NcoI (Thermo Fisher Scientific, США), после чего проводили лигазную реакцию с ранее разработанным в нашей лаборатории вектором на основе транспозона Tn4001[82], содержащим ген устойчивости к тетрациклину (*tetM*) с оптимизированным кодонным

составом для микоплазмы, в качестве селективного маркера (Рисунок 2.2). Последовательность промотора, сайта связывания рибосомы и терминатора были подобраны так, чтобы экспрессия вносимого в конструкцию гена была максимальна.



**Рисунок 2.2** Схема участка плазмиды для сверхэкспрессии гена *mraZ*. Вектор pTn4001mini\_Recipient-MCS-M5 разработан на основе транспозона. Встраивание в геном микоплазмы происходит участка, расположенного между повторами OIR и IIR (обозначены красным), которые узнает транспозаза. Промоторы, сайты связывания рибосомы и терминаторы обозначены желтым. Ген устойчивости к тетрациклину – зеленым. Полную схему плазмиды представлена в Приложении, Рисунок П1.

После реакции лигирования, полученным вектором трансформировали компетентные клетки *E. coli* (см. п. 2.6), штамм TOP10 (Invitrogen) и высевали на чашку с LB агаром и ампициллином (100 мкг/мл). После 15 часов инкубации при 37°C, колонии *E. coli* скольвали и проводили скрининг методом ПЦР. Длины продуктов оценивали электрофорезом в агарозном геле с окрашиванием бромистым этидием. Колонии бактерий, содержащие плазмиду с интересующим нас геном высевали в жидкой среде LB с добавлением ампициллина 100мкг/мл, растили 37°C на качалке (140 об./мин) в течении ночи. Затем клетки осаждали (5000 g в течение 5 мин при 4°C) и выделили плазмиду щелочным гидролизом (см п. 2.7). Измеряли концентрацию плазмиды флуориметрически на флуориметре Qubit (Invitrogen) согласно инструкции производителя. Последовательность вставки проверяли с помощью секвенирования по методу Сэнгера (лаб. постгеномных исследований в биологии ФГБУ ФНКЦ ФХМ ФМБА России). Плазмиды, не содержащие мутаций использовали для электропорации клеток *M. gallisepticum* S6.

## 2.16 Трансформация плазмидной ДНК клеток *M.gallisepticum* методом электропорации

Трансформацию клеток микоплазмы плазмидой проводили по методике, описанной в [102] с некоторыми модификациями. Один мл суспензионной культуры в логарифмической фазе роста осаждали центрифугированием 10000 g 10 мин 4°C. Надосадочную жидкость стерильно отбирали и ресуспендировали клеточный осадок в 250 мкл стерильного охлаждённого буфера для электропорации. Повторяли процедуру дважды. Затем ресуспендировали клетки в 250 мкл стерильного охлаждённого буфера для электропорации, добавляли 200-400 нг плазмиды, переносили в предварительно охлажденные кюветы для

электропорации с шириной зазора 2 мм. Проводили электропорацию на приборе MicroPulser Electroporator (Bio-Rad) со следующими параметрами:  $V=2500$  В,  $R=100$  Ом,  $t=5$  ms. Затем переносили клетки в 1 мл среды для культивирования микоплазмы, инкубировали клетки в течение 4 часов при  $37^{\circ}\text{C}$ . После высевали культуру в 10 мл полужидкой среды, содержащей тетрациклин с конечной концентрацией 2 мкг/мл. Спустя 4-5 суток, выкалывали выросшие колонии, переносили в жидкую среду, подращивали и использовали для дальнейших экспериментов по оценки экспрессии генов.

Буфер для электропорации: 8 мМ NEPEES; 272 мМ сахараза, pH = 7.4

### **2.17 Выделение РНК и синтез кДНК**

Тотальную фракцию РНК выделяли из суспензии клеток *M. gallisepticum* с помощью реагента Trizol LS (Invitrogen) по протоколу фирмы-производителя. Для этого к 100 мкл жидкой культуры микоплазмы в логарифмической фазе роста добавляли 300 мкл реагента Trizol LS, образец интенсивно перемешивали, инкубировали при комнатной температуре 5 мин, добавляли 70 мкл хлороформа, и центрифугировали при 16000 g,  $4^{\circ}\text{C}$ , в течение 15 минут. Водную фазу, содержащую нуклеиновые кислоты, отбирали в отдельную пробирку и осаждали добавлением равного количества изопропилового спирта. Инкубировали при  $-20^{\circ}\text{C}$  не менее часа, центрифугировали образец при 16000 g,  $4^{\circ}\text{C}$ , в течение 20 минут. Осадок промывали 80% этиловым спиртом. Полученный осадок растворяли в 10 мкл деионизированной воды (получали около 1 мкг РНК), обрабатывали ДНКазой I (Thermo Fisher Scientific, США) и использовали для синтеза кДНК с помощью обратной транскриптазы Maxima H Minus Reverse Transcriptas (Thermo Fisher Scientific, США), согласно протоколу фирмы производителя. Полученную кДНК разводили в 10 раз в деионизированной воде и использовали как матрицу для ПЦР в реальном времени из расчета 1 мкл на одно измерение.

### **2.18 Количественная ОТ-ПЦР в реальном времени**

ПЦР проводили в реакционной смеси, содержащей 1x ПЦР буфер (НПФ «Литех», Россия); 0,2 мМ каждого дНТФ; 5 пмоль каждого из праймеров; 0,75 ед. Taq-полимеразы (НПФ «Литех», Россия); 0,5x SYBR Green (Promega, США); 2% формамида и 1 мкл матрицы кДНК. Суммарный объем смеси составлял 20 мкл. Амплификацию и детекцию флуоресценции проводили на приборе C1000 Touch с оптическим модулем CFX96 (Bio-Rad, США) по следующей программе: первичное плавление  $96^{\circ}\text{C}$  1 мин, плавление  $96^{\circ}\text{C}$  15 с, отжиг праймеров  $58^{\circ}\text{C}$  20 с, синтез  $65^{\circ}\text{C}$  60 с, снятие спектра флуоресценции, возврат к этапу плавления, количество циклов ПЦР – 40. Для определения специфичности реакции после проведения ПЦР определялась кинетика плавления ПЦР-продукта. Последовательности

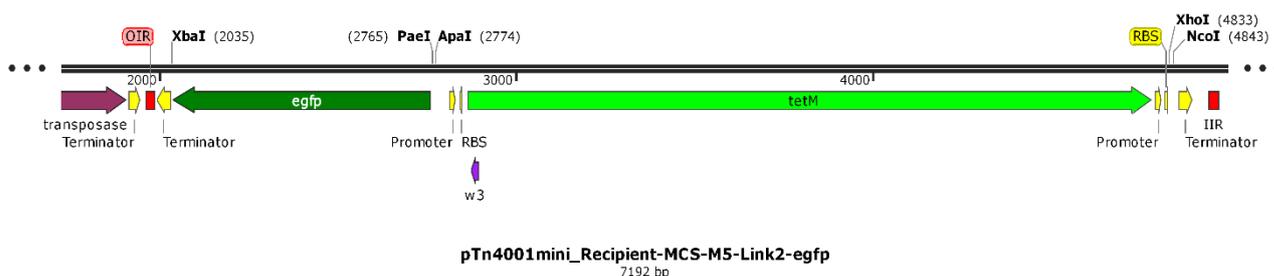
праймеров представлены в Таблице П1 Приложения. Уровни экспрессии целевых генов внутри каждого образца нормализовали на уровень мРНК гена «домашнего хозяйства» *tuf*.

Экспериментальные данные ПЦР в реальном времени, представленные в данной работе, были получены в не менее трёх независимых биологических повторах и представлены в виде среднего всех повторных экспериментов с определенными стандартными ошибками. Достоверность различия двух средних значений оценивалась с помощью рангового непарного критерия Вилкоксона, реализованного в программных пакетах R. На всех графиках:  $p < 0,05$  (\*), n.s. – не значимое различие.

### **2.19 Разработка вектора для оценки влияния генетических элементов на экспрессию гена *egfp***

Плазмиду pTn4001mini\_Recipient-MCS-M5 модифицировали, добавив три сайта для эндонуклеаз рестрикции *ApaI*, *PaeI* и *XbaI* перед геном *tetM*, получив плазмиду pTn4001mini\_Recipient-MCS-M5-Link2. Ген *egfp* амплифицировали с плазмиды pEGFP-C1. Полученный ПЦР-продукт и вектор pTn4001mini\_Recipient-MCS-M5-Link2 подвергали обработке эндонуклеазами рестрикции *PaeI* и *XbaI* (Thermo Fisher Scientific, США), после чего проводили лигазную реакцию. Нарработку плазмиды проводили в *E. coli* (см. п. 2.6), штамм TOP10 (Invitrogen, США) как описано выше (см. п. 2.6). Генетические элементы (участки промотора, потенциальные сайты связывания ТФ) получали путем химического синтеза (НПФ «Литех», Россия или ЗАО «Евроген», Россия). Олигонуклеотиды разводили в деионизированной воде до концентрации 10-50 пмоль/мкл, подвергали обработке полинуклеотид киназой T4 (Thermo Fisher Scientific, США), на 20 мкл реакционной смеси добавляли 10-50 пмоль олигонуклеотида, 2 мкл 10x буфера А, 2 мкл 10 мМ АТФ, 1 мкл (10 Units) T4 полинуклеотид киназы (Thermo Fisher Scientific, США). Образцы инкубировали при 37°C 1 час. Затем комплементарные олигонуклеотиды смешивали, отжигали на приборе Tetrad 2 Thermal Cycler (Bio-Rad, США). Для это образец нагревали до 95°C в течение 2 мин и затем охлаждали до 20°C со скоростью 0,1°C/с. Олигонуклеотиды подбирали так, чтобы после отжига, дцДНК фрагмент имел липкие концы, пригодные для проведения лигазной реакции с вектором pTn4001mini\_Recipient-MCS-M5-Link2-eGFP (Рисунок 2.3), предварительно обработанным рестриктазами *ApaI* и *PaeI*. Длинные фрагменты 5'-областей генов получали с помощью ПЦР с использованием ДНК полимеразы Phusion (Thermo Fisher Scientific, США), согласно протоколу производителя. Список праймеров представлен в Таблице П1 Приложения. Реакцию проводили в 100 мкл смеси, затем ПЦР продукты очищали от белков добавляя равный объем хлороформа. Смесь активно встряхивали и центрифугировали при 16000 g, 4°C, в течение 20 минут. Водную фазу отбирали в отдельную пробирку и осаждали

нуклеиновые кислоты добавлением равного количества изопропилового спирта и 5% по объему ацетата натрия (3М ацетата натрия, pH 5.2). Смесь инкубировали при -20°C в течение как минимум часа и центрифугировали при 16000 g, 4°C, в течение 20 минут. Полученный осадок промывали раствором 80% этилового спирта. Осадок растворяли в 20 мкл деионизированной воды, измеряли концентрацию флуориметрически на флуориметре Qubit (Invitrogen, США) согласно инструкции производителя. Подвергали ампликоны обработкой эндонуклеазами рестрикции *ApaI* и *PaeI*, после чего проводили лигазную реакцию с вектором pTn4001mini\_Recieipient-MCS-M5-Link2-eGFP.



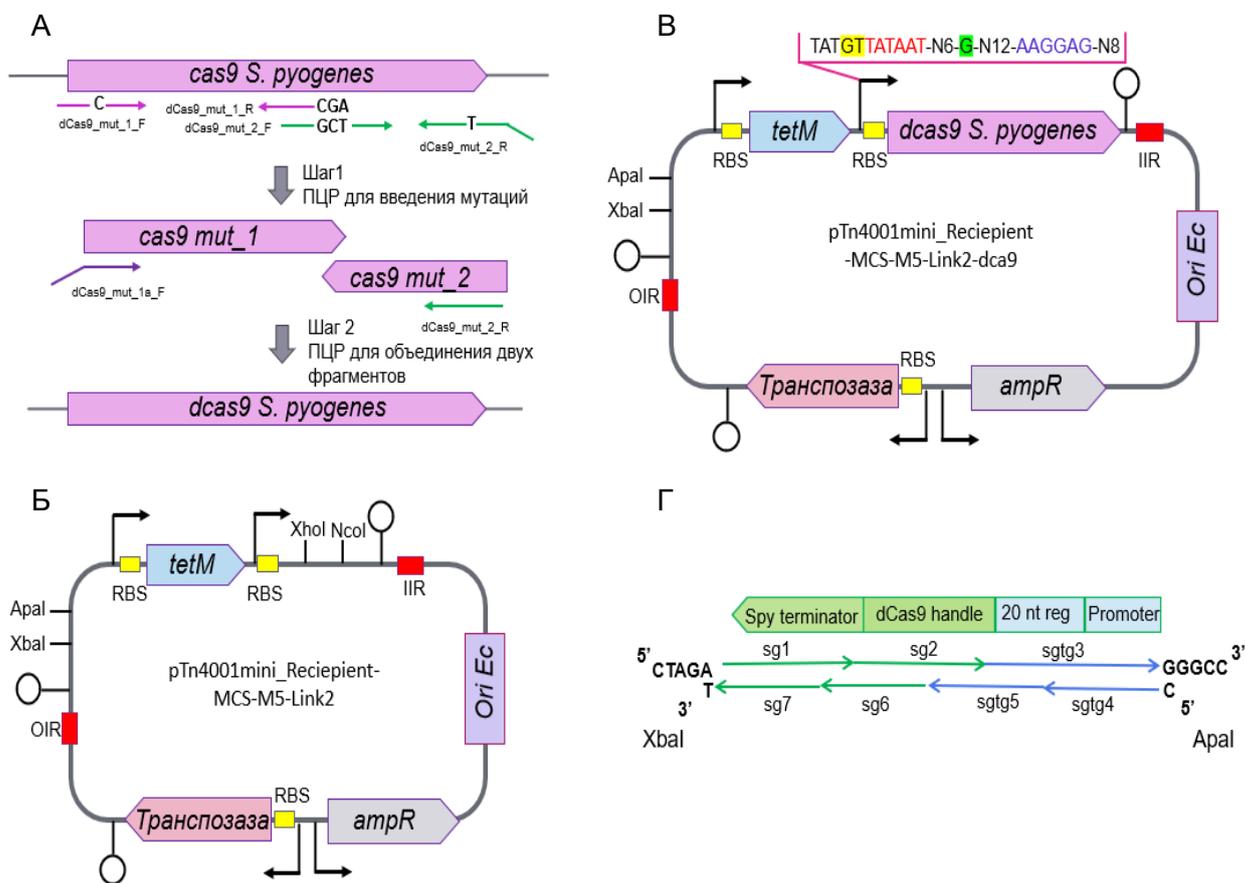
**Рисунок 2.3** Схема участка плазмиды pTn4001mini\_Recipient-MCS-M5-Link2-egfp для проверки влияния генетических элементов на экспрессию гена *egfp*. Интересующий фрагмент дцДНК, содержащий последовательности промоторов, предсказанных сайтов связывания регуляторов клонировали по сайтам рестрикции *ApaI* и *PaeI*. Встраивание в геном микоплазмы происходит участка, расположенного между повторами OIR и IIR (обозначены красным), которые узнает траспозаза. Промоторы, сайты связывания рибосомы и терминаторы обозначены желтым. Ген устойчивости к тетрациклину – зеленым. Полная схема плазмиды представлена в Приложении, Рисунок П2

После проводили трансформацию *M. gallisepticum* S6 полученным вектором, как описано в 2.15, выделяли РНК, получали кДНК (п. 2.17) и оценивали экспрессию генов «домашнего хозяйства», а также гена *egfp* с помощью ПЦР в реальном времени (п. 2.18). Последовательности праймеров представлены в Приложении, Таблица П1.

## 2.20 Таргетная репрессия транскрипции генов у *M. gallisepticum* с помощью системы CRISPR-интерференции (CRISPRi)

Для репрессии транскрипции генов в *M. gallisepticum* был создан вектор на основе pTn4001mini\_Recipient-MCS-M5-Link2, содержащий ген *dcas9* *Streptococcus pyogenes*, а также кодирующий sgРНК (small guide RNA) участок ДНК для нацеливания РНК-белкового комплекса на ДНК-мишень. Исходно в нашем распоряжении была плаزمида, содержащая ген нуклеазы *cas9* *S. pyogenes* дикого типа. Для того, чтобы получить мутантный вариант белка, способный связываться с sgРНК и ДНК-мишенью, но не гидролизовать её, необходимо было провести замену 10-ого а.о. аспарагиновой кислоты и 840-ого а.о. гистидина на остатки аланина [103], а также заменить стоп-кодон TGA на TAA. Мутагенез проводили с помощью ПЦР в два этапа, сборку полноразмерного гена *dcas9* с линейным вектором

pTn4001mini\_Recieipient-MCS-M5-Link2 проводили по методу Гибсона с помощью реагентов NEBuilder HiFi DNA assembly master mix (New England BioLabs, США) (Рисунок 2.4 А). Реакцию амплификации проводили с использованием высокоточной ДНК-полимеразы Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific, США), согласно протоколу производителя. На первом этапе в качестве праймеров использовали олигонуклеотиды, содержащие точечные замены (Приложение, Таблица П1). Полученные ПЦР-продукты очищали экстракцией хлороформом и преципитацией изопропанолом с добавлением ацетата натрия. Длину продуктов, их чистоту и количество оценивали путём электрофореза в агарозном геле с окрашиванием бромистым этидием. На втором раунде амплификации использовали олигонуклеотиды с 5'-концами, гомологичными концам вектора pTn4001mini\_Recieipient-MCS-M5-Link2 после обработки эндонуклеазами рестрикции NcoI и XhoI.



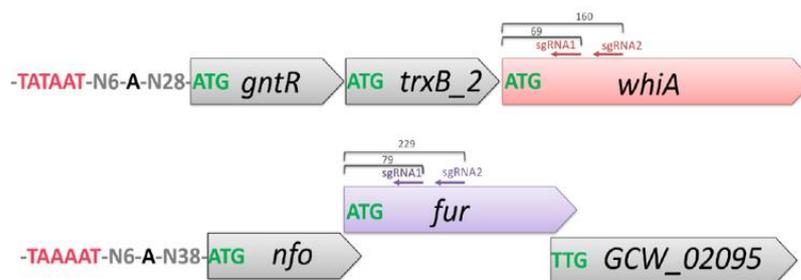
**Рисунок 2.4** Схема сборки плазмиды *pTn4001mini\_Recieipient-MCS-M5-Link2-dcas9 S. pyogenes* и последовательностей ДНК, кодирующих sgРНК для репрессии транскрипции генов *M. gallisepticum*. **А** – Последовательность проведения сайт-направленного мутагенеза для получения гена *dcas9*, в котором заменены 10-ый а.о. аспарагиновой кислоты и 840-ой а.о. гистидина на остатки аланина (GAT → GCT и CAC → GCT соответственно); праймеры указаны стрелками; буквы обозначают замещающие нуклеотиды; **Б** – Схема вектора *pTn4001mini\_Recieipient-MCS-M5-Link2*; положение промоторов показаны стрелками, терминаторы – кружками, RBS (сайт связывания рибосомы) – желтым, OIR и IIR-инвертированные повторы для интеграции транспозона в геном; *ampR*, *tetM* – гены устойчивости к ампициллину и тетрациклину, соответственно. **В** – Карта полученной

вектора, для проведения CRISPR-интерференции; выделены нуклеотидные последовательности промотора и сайта связывания рибосомы; элемент промотора extended элемент выделен желтым, -10-бокс красным, инициаторный нуклеотид зеленым, RBS синим; Г. – Стратегия сборки фрагментов двуцепочечной ДНК, кодирующих sgРНК. Каждый фрагмент состоял из семи перекрывающихся олигонуклеотидов, четыре из которых были общими для всех фрагментов, кодирующих sgРНК, и три уникальными для каждой конкретной sgРНК. Все олигонуклеотиды фосфорилировали, лигировали вместе, и полученный продукт клонировали между сайтами *ApaI* и *XbaI* в вектор pTn4001mini\_Recipient-MCS-M5-Link2-dcas9. Общие олигонуклеотиды отмечены зелеными стрелками, уникальные — синими. Опубликовано в [100]

Полученным вектором трансформировали компетентные клетки *E. coli* (см. п. 2.6), штамм TOP10 (Invitrogen, США), проводили скрининг колоний, наработку плазмид и секвенирование по методу Сэнгера как описано выше (п. 2.15). Плазмиду, не несущую мутаций использовали для трансформации клеток *M. gallisepticum* (п. 2.16). Полученные колонии трансформантов выращивали в 10 мл жидкой среды, выделяли ДНК и с помощью секвенирования по методу Сэнгера оценивали позицию интеграции транспозона в геномную ДНК микоплазмы. Три трансформанта, у которых интеграция в геном транспозона не затронула гены, использовали для подтверждения экспрессии dCas9 на уровне белка с помощью вестрен-блота (п. 2.10)

Согласно работам Martin Jinek [104], химерная молекула sgРНК состоит из трех частей – регион нацеливания на ДНК-мишень (20 н.о.), шпилька для связывания dCas9 – dCas9-binding handle (42 н.о.), терминатор транскрипции (40 н.о.). Существует несколько способов ввести sgРНК в клетки. Было показано, что эффективность репрессии транскрипции генов у *E. coli* зависит от позиции мишени на ДНК [103]. Для того, чтобы легко менять регион нацеливания на ДНК мишень в sgРНК, мы решили проводить сборку фрагмента ДНК, кодирующего sgРНК вместе с сильным промотором из олигонуклеотидов, а затем клонировать такие фрагменты в плазмиду pTn4001mini\_Recipient-MCS-M5-Link2-dcas9 по сайтам рестрикции *XbaI* и *ApaI*. Фрагмент такой ДНК имел длину 129 п.н.. Мы разбили фрагмент на 7 пересекающихся олигонуклеотидов таким образом, что каждый фрагмент, нацеленный на новую ДНК-мишень будет состоять из 4-ёх общих олигонуклеотидов и трёх уникальных (Рисунок 2.4 Г).

Для генов *whiA* и *fur* *M. gallisepticum* было выбрано две позиции нацеливания sgРНК-dCas9 в теле гена с учётом положения РАМ (protospacer adjacent motif) на ДНК. Вторичная структура каждой sgРНК была проверена с помощью сервиса The ViennaRNA Web Services (<http://rna.tbi.univie.ac.at/>), специфичность (количество мишеней) – с помощью blastn (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). На Рисунке 2.5 представлена схема положения sgРНК (sgРНК1 и sgРНК2) относительно детерминант промотора и стар-кодона.



**Рисунок 2.5 Структура оперонов, содержащих гены-мишени.** Стрелки указывают положение региона нацеливания на ДНК-мишень сконструированных sgРНК; сверху отмечено расстояние от старт-кодона; -10-бокс промотора выделен красным; инициаторный нуклеотид – черным, старт-кодон – зеленым. Опубликовано в [100]

Олигонуклеотиды разводили в деионизированной воде до концентрации 100 пмоль/мкл, подвергали обработке полинуклеотид киназой T4 (Thermo Fisher Scientific, США), на 20 мкл реакционной смеси добавляли 100 пмоль олигонуклеотида, 2 мкл 10x буфера А, 2 мкл 10 мМ АТФ, 1 мкл (10 Units) T4 полинуклеотид киназы (Thermo Fisher Scientific, США). Образцы инкубировали при 37°C 1 час. Затем отбирали по 5 мкл каждого олигонуклеотида, смешивали, отжигали на приборе Tetrad 2 Thermal Cycler (Bio-Rad, США). Для это образец нагревали до 95°C в течение 2 мин и затем охлаждали до 25°C со скоростью 0,1°C/с.. Затем добавляли 2 мкл 10 мМ АТФ, 1 мкл T4 ДНК-лигазы (5 Units) (Thermo Fisher Scientific, США), доводили объем реакционной смеси до 40 мкл добавлением деионизированной воды, инкубировали 1 час при комнатной температуре. Полученные ДНК-фрагменты очищали экстракцией хлороформом и преципитацией изопропанолом с добавлением ацетата натрия.

Длину продуктов оценивали с помощью электрофореза в агарозном геле, измеряли концентрацию флуориметрически на флуориметре Qubit (Invitrogen, США) согласно инструкции производителя. Полученные фрагменты ДНК имели липкие концы, пригодные для проведения лигазной реакции с вектором pTn4001mini\_Recieipient-MCS-M5-Link2-dcas9, предварительно обработанным рестриктазами XbaI и ApaI. Полученными векторами трансформировали компетентные клетки *E. coli* (см. п. 2.6), штамм TOP10 (Invitrogen, США), проводили скрининг колоний, наработку плазмид и секвенирование по методу Сэнгера как описано выше (п. 2.15). Плазмиды, не содержащие мутаций использовали для электропорации клеток *M. gallisepticum*. В Таблице 2.2. представлены характеристики полученных последовательностей sgРНК в составе системы для CRISPR-интерференции.

**Таблица 2.2. Характеристика сконструированных sgРНК и их активности**

Название sgРНК	Регион нацеливания на ДНК-мишень (20 н.о.)	T <sub>m</sub> , °C	ГЦ состав, %	Средний уровень репрессии целевого гена, Fold ± SE
fur-sg1	ACCTTCTAGGTTTGGTGATT	58,5	40	23 ± 0,6
fur-sg2	ACGAATCAACAAGATTGAGT	57,4	35	43 ± 8,8
whiA-sg1	CATTTTGCACCAACAAGTCG	60,9	45	5±0.2
whiA-sg2	AGATGATCGTTTTTGGGTGT	59,9	40	48± 5.3

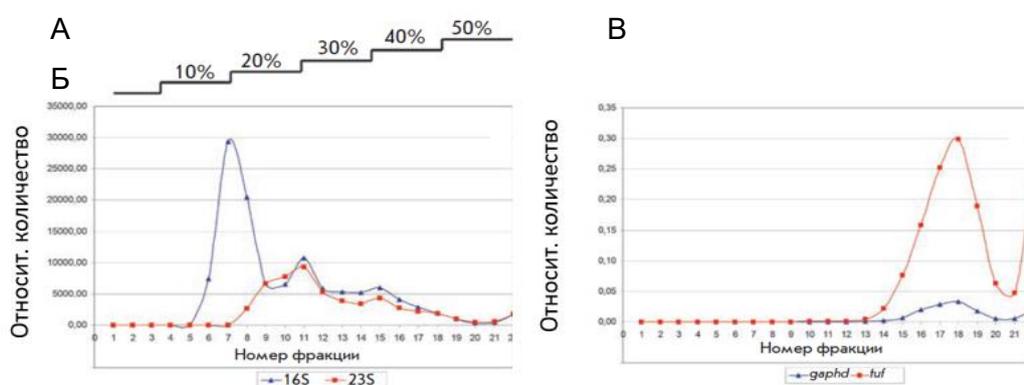
Высокая активность протестированных последовательностей sgРНК не связана с высоким содержанием ГЦ пар, однако наличие динуклеотида ГУ в 3' области нацеливания sgРНК ассоциирован с большим уровнем подавления экспрессии целевых генов. Эта связь подтверждается у большей выборки сконструированных sgРНК [100]. Стоит отметить, что при подавлении экспрессии *fur*, уровень экспрессии нижестоящего гена *GCW\_02095* оперона снижается в 3,5 раза независимо от уровня репрессии *fur*. Уровент мРНК *nfo* не меняется Рисунок ПЗ Приложения.

### **2.21 Выделение рибосом и фракции мРНК, связанной с рибосомами *M.gallisepticum***

Совместно с Фисуновым Г.Ю. (лаб. протеомного анализа ФНКЦ ФХМ при ФМБА России). проводили выделение рибосом и фракции мРНК, связанной с рибосомами. Образцы были получены для культуры *M.gallisepticum* S6 выращенной в нормальных условиях (37°C), так и культуры, перенесшей тепловой стресс (42°C в течение 30 мин). Для выделения рибосом к 50 мл суспензионной культуры *M. gallisepticum* S6 в логарифмической фазе роста добавили хлорамфеникол до конечной концентрации 100 мкг/мл, тщательно перемешали и инкубировали в течение 5 мин на льду. Затем клетки осаждали центрифугированием 4500 g 20 мин при 4°C. Супернатант отбирали, клеточный осадок ресуспендировали в 500 мкл буфера для лизиса. После добавляли 15 мкл 10% детергента Nonidet P-40 (Sigma-Aldrich, США), тщательно перемешивали. Клеточный лизат замораживали при -75°C не менее 1 ч. Полученный клеточный лизат осветляли центрифугированием при 20000 g в течение 20 мин при 4°C. Супернатант отбирали и фракционировали с помощью центрифугирования в ступенчатом градиенте сахарозы от 10% до 50% с шагом 10%. Ступенчатый градиент сахарозы создавали в 5-мл поликарбонатной пробирке путем послойного нанесения раствора сахарозы разной плотности. Объем каждого слоя – 750 мкл. Сахарозу растворяли в буфере для лизиса без добавления Nonidet P-40, хлорамфеникола, смеси ингибиторов протеаз и РНКаз. Разделение проводили с помощью ультрацентрифуги Optima (Beckman Coulter, США) при 200 000 g в течение 1 ч при 4°C. Фракции объемом 200 мкл отбирали с помощью пипетки.

Из каждой полученной фракции выделяли РНК, как описано в п. 2.17. Часть выделенной РНК использовали для синтеза кДНК и оценки представленности рибосомной РНК и мРНК генов «домашнего хозяйства» (глицеральдегид-3-фосфат дегидрогеназы – *gapd*; фактора трансляции EF-Tu – *tuf*) в каждой фракции методом количественной ОТ-ПЦР как описано в п.2.14 и п.2.15 (Рисунок 2.6).

РНК фракций, где наблюдалось эквимолярное соотношение 16S и 23S рРНК, а также высокое содержание мРНК генов домашнего хозяйства (фракции 17 и 18) объединили (фракция мРНК, связанная с рибосомами) и использовали для получения библиотек кДНК и последующего секвенирования на платформе SOLiD 5500 (Life Technologies, США).



**Рисунок 2.6 Состав фракций, после разделения цитоплазматической фракции *M. gallisepticum* в градиенте сахарозы.** С помощью метода количественной ОТ-ПЦР в реальном времени измеряли представленность в каждой фракции – 16S рРНК, 23S рРНК, мРНК глицеральдегид-3-фосфат-дегидрогеназы (*gapdh*) и фактора трансляции EF-Tu (*tuf*). Все результаты (для рРНК и мРНК) нормировали на количество 16S рРНК во фракции номер 1. **А** – профиль ступенчатого градиента в масштабе относительно полученных фракций; **Б** – распределение рРНК по фракциям; **В** – распределение мРНК по фракциям. Опубликовано в [105]

Буфер для лизиса: 20 мМ HEPES; 100 мМ NaCl; 6 мМ MgCl<sub>2</sub>; 2 мМ спермидина; 100 мкг/мл хлорамфеникола; 5 мкл смесь ингибиторов протеаз (GE Healthcare, США); 200 ед. ингибитора РНКаз RiboLock (Thermo Fisher Scientific, США), pH 7.5

## 2.22 Подготовка библиотек кДНК и секвенирование

Подготовку библиотек кДНК и секвенирование проводили совместно с Никитиной А.С. (лаб. постгеномных исследований в биологии ФНКЦ ФХМ при ФМБА России).

Рибосомально-связанную фракцию РНК очищали от тРНК с помощью набора PureLink RNA Mini Kit (Ambion, США). Полученную фракцию РНК химически фрагментировали, добавляя буфер, содержащий Zn<sup>2+</sup> и инкубировали образцы при 70°C в течение 15 мин. Реакцию фрагментации останавливали, добавляя 20 мМ раствор ЭДТА. Длина полученных – фрагментов оставляла 200 п.н. Затем образцы обрабатывали полинуклеотид киназой T4

(Thermo Fisher Scientific, США). Подготовку библиотеки кДНК проводили в соответствии со стандартным протоколом для подготовки библиотеки секвенирования РНК (RNA-seq) на платформе SOLiD с использованием наборов Total RNA-Seq Kit и SOLiD RNA Barcoding Kit (Ambion, США). Качество РНК, фрагментированной РНК, библиотек кДНК оценивали с помощью Agilent 2100 Bioanalyzer system (Agilent, США). Дц-кДНК амплифицировали и нормализацию библиотек дц-кДНК проводили с помощью дуплекс-специфической нуклеазы DSN (Евроген, Россия) согласно рекомендациям производителя. Затем образцы были проанализированы с помощью электрофореза в агарозном геле. Блок геля, содержащий фрагменты библиотеки длиной 200-300 п.н. был извлечен и дц-кДНК экстрагировали из геля с использованием набора SOLiD Library Quick Gel Extraction Kit (Life Technologies, США). Все образцы были подготовлены в трех биологических повторностях.

Буфер для фрагментации: 100 мМ ZnSO<sub>4</sub>; 100 мМ Трис, pH = 7.0

### **2.23 Анализ данных секвенирования РНК**

Качество прочтений, полученных в результате секвенирования, проверяли с помощью программы FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>). С помощью программы Trimmomatic [106] прочтения с качеством ниже 20 были удалены, с 5'- концов прочтений были удалены последовательности адаптеров. Используя программу Bowtie2 [107] оставшиеся прочтения были картированы на последовательность генома *Mycoplasma gallisepticum* S6 (GenBank: CP006916.3). Значения покрытия прочтениями каждого гена *M. gallisepticum* S6, а также нормализованные значения RPKM (reads per kilobase per million of mapped reads — число прочтений на тысячу нуклеотидов на миллион картированных прочтений) были получены с помощью программы SAJR [82]. Для расчёта дифференциальной экспрессии генов использовали пакет SAJR, а также пакеты статистической среды R – DESeq. В качестве порогов статистически достоверной дифференциальной экспрессии были взяты значения  $p\text{-value} < 0.05$  (с учетом поправки Бенджамини – Хохберга) и двухкратный уровень изменения экспрессии. Результаты подсчета дифференциальной экспрессии генов представлены в Таблице П2 Приложения. Данные секвенирования суммарной цитоплазматической фракции РНК *M.gallisepticum* S6 культивированной при нормальных условиях (37°C), и после теплового воздействия (42°C в течение 30 мин), были взяты из ранее опубликованного исследования [82].

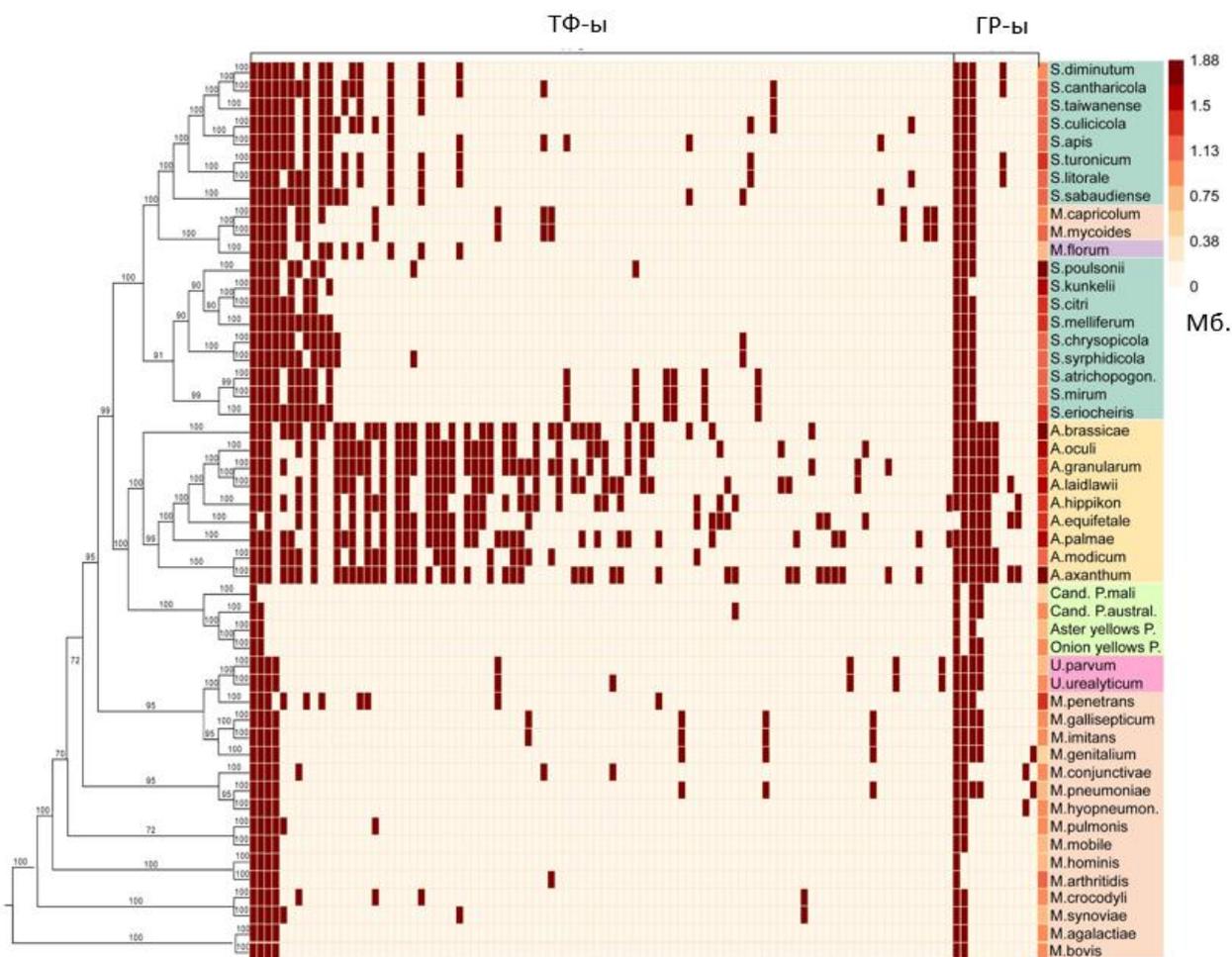
### 3 РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

#### 3.1 Поиск потенциальных факторов регуляции транскрипции у *Mycoplasma gallisepticum* на основании сходства первичной структуры

На первом этапе был произведен поиск предполагаемых факторов регуляции транскрипции, а также оценено их распределение среди 50 представителей класса Mollicutes. На основе сравнения белковых последовательностей у выбранных бактерий были составлены кластеры ортологов. Затем для участника кластера проведен поиск функциональных доменов. Предполагаемые факторы регуляции транскрипции были отобраны по наличию ДНК-связывающих мотивов вида спираль–поворот–спираль (НТН), самому распространённому структурному элементу белков, связывающих нуклеиновые кислоты [21], а также с учетом аннотации белка, для фильтрации ложноположительных находок. То есть белков, связывающих ДНК, но не являющиеся регуляторами экспрессии генов. Такой отбор был проведен вручную. Например, у *M. gallisepticum* S6 согласно найденным доменам и аннотации 91 белок способен связываться с ДНК, из них 18 – транспозазы; и только 6 содержали классический структурный мотив НТН спираль–поворот–спираль, также в список регуляторов были включены белки, аннотированные как транскрипционные регуляторы (Таблица 3.1). На Рисунке 3.1 представлено распределение ДНК-связывающих белков, отнесенных к потенциальным факторам регуляции транскрипции. Помимо классических транскрипционных факторов, в исследование были включены глобальные регуляторы – SpxA, SpoT, HprK у которых не предсказано наличие НТН-доменов, но от действия которых зависит экспрессия множества генов, а также предполагаемые альтернативные сигма факторы.

**Таблица 3.1. Список потенциальных белковых регуляторов транскрипции *M.gallisepticum***

Название белка	Наличие НТН-домена
HrcA	Содержит НТН домен; регулятор генов теплового стресса; консервативен
WhiA	Содержит НТН домен; консервативен
YebC	Не содержит НТН домен; консервативен
MraZ	Не содержит НТН домен; консервативен
GntR	Содержит НТН домен
Fur	Содержит НТН домен
Xre	Содержит НТН домен
YlxM	Содержит НТН домен
Глобальные регуляторы	
SpoT	Не содержит НТН домен; ppGpp-синтаза II/ ppGpp-гидролаза
HprK	Не содержит НТН домен; киназа Hpr
GCW_00440	Не содержит НТН домен; альтернативный сигма-фактор
SpxA	Не содержит НТН домен; глобальный регулятор Spx



**Рисунок 3.1. Распределение факторов регуляции транскрипции у молликут.** ТФ – транскрипционные факторы, ГР – глобальные регуляторы. Ортологичные группы регуляторов образуют вертикаль; наличие регулятора отмечено коричневым цветом, отсутствие – желтым. Филогенетическое дерево построено на основе выравнивания ортологов, консервативных между всеми видами. Панель справа – виды бактерий класса Mollicutes, для которых проведен поиск регуляторов; цвет выделяет принадлежность к разным родам: желтый – *Acholeplasma* spp., бирюзовый – *Spiroplasma* spp., оранжевый – *Mycoplasma* spp., лимонный – *Phytoplasma* spp., фиолетовый – *Mesoplasma florum*, розовый – *Ureaplasma*. Размеры геномов бактерий отмечены градиентом красного цвета согласно легенде. Работа выполнена совместно с И.А. Гараниной [96].

Среди предсказанных регуляторов только четыре белка – HrcA, MraZ, WhiA и YebC – были представлены у более 70% видов. HrcA – единственный фактор регуляции транскрипции *M. gallisepticum* с известным механизмом действия. Этот белок связывается с CIRCE участком генов шаперонов, выступая в качестве репрессора [76]. Сайты связывания MraZ, WhiA и YebC у молликут оставались неизвестными на момент начала выполнения диссертационной работы. Больше всего регуляторов было предсказано у представителей рода *Acholeplasma*. Среди молликут ахелеплазмы обладают наибольшим размером генома, большинство видов – сапрофиты. В целом размер генома молликут коррелирует с количеством предсказанных факторов регуляции транскрипции: чем меньше размер генома, тем меньше регуляторов, – стабилизация числа регуляторов на геном происходит при размере генома бактерии от 1 Мб.

К глобальным регуляторам мы отнесли белок SpxA – хорошо изученный у *B. subtilis* регулятор ответа на окислительный стресс, который не содержит НТН-мотив и модулирует активность РНК-полимеразы за счёт взаимодействия с её С-концевым доменом  $\alpha$ -субъединицы и участком промотора [108]. SpoT – фермент синтеза гуанозин-тетрафосфата (-пентафосфата) или (p)ppGpp. У грамположительных бактерий основная роль этого метаболита в регуляции транскрипции заключается в модуляции концентрации пуринов, за счёт чего меняется доступность нуклеотидов при инициации транскрипции [109]. Фосфорилированная по остатку серина форма другого глобального регулятора – белка HprK – связывается с транскрипционными факторами катаболизма углеводов, влияя на сродство ТФ-ов к ДНК. У *B. subtilis* экспрессия около 10% генов регулируется с помощью подобного механизма [110]. HprK найдена почти у всех молликут, кроме фитоплазм.

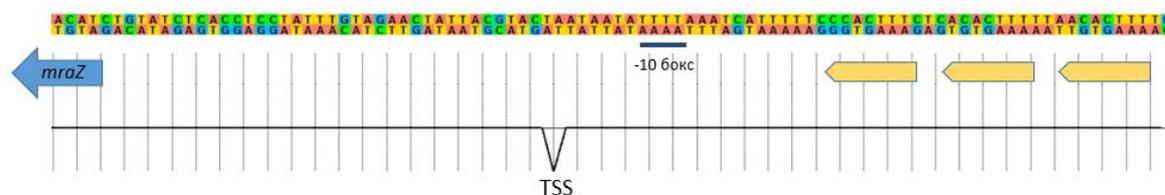
Далее экспериментальная работа по поиску и валидации мишеней ТФ-ов была проведена для консервативных белков MraZ и WhiA, а для также специфических факторов регуляции транскрипции – Xre и Fur. Изучение регуляторов SpxA и YebC также проводилось в нашей лаборатории Д.С. Матюшкиной. Показана их роль при взаимодействии микоплазмы с клеткой хозяина и адаптации к внутриклеточным условиям.

### **3.2 Поиск сайтов связывания белка MraZ**

Ген *mraZ* консервативен у бактерий и, как правило, является первым геном в кластере деления DCW (Division and Cell Wall). У *E. coli* в состав этого кластера входит 16 генов, чьи белковые продукты участвуют в биосинтезе пептидогликана и сборке дивисомы – аппарата клеточного деления [111]. У *Mycoplasma gallisepticum*, как и у всех представителей класса Mollicutes, отсутствует клеточная стенка, в состав DCW кластера входит только четыре гена – *mraZ*, *mraW*, *ftsA* и *ftsZ*. Долгое время функция MraZ оставалась неясна, и лишь в 2014 году для *Escherichia coli* была показана способность MraZ связываться с ДНК и выступать в качестве репрессора генов DCW кластера [111]. Сохранение этого гена у молликут, вероятно, указывает на важную роль этого белка и на фоне общей редукции регуляторного аппарата может вносить больший вклад в регуляцию, чем у бактерий с расширенным спектром модуляторов.

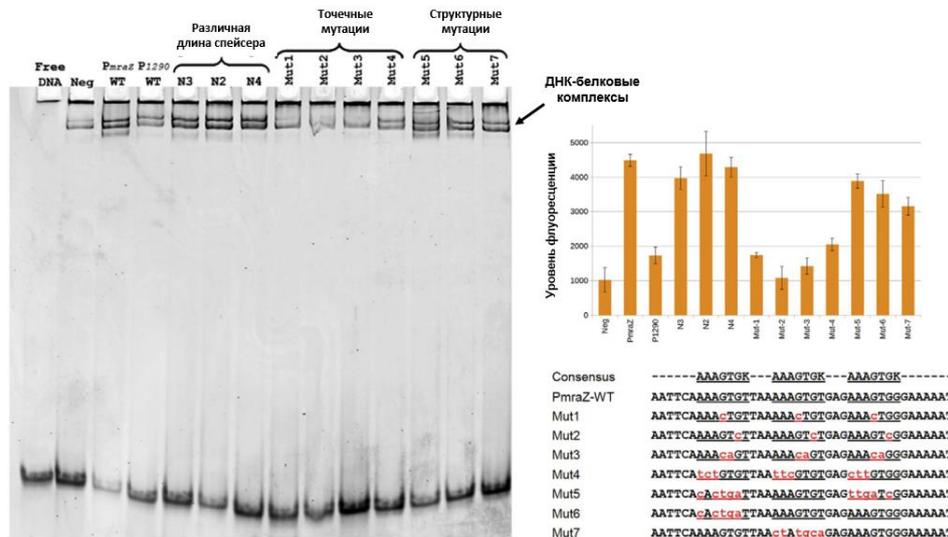
#### **3.2.1 Белок MraZ связывает серию прямых повторов, расположенных в промоторной области генов DCW кластера**

Используя ранее полученные данные о сайте начала транскрипции, мы определили промоторную область генов DCW кластера *M. gallisepticum*. Рядом с промотором расположены прямые повторы – AAAGTG[T/G], разделенные спейсером в три нуклеотида (Рисунок 3.2).



**Рисунок 3.2. Последовательность промоторной области генов DCW кластера.** Черный пик – разметка позиции старта транскрипции – TSS; повторы AAAGTG[T/G] обозначены жёлтым; начало кодирующей последовательности гена *mraZ* показано синей стрелкой.

Консервативность расположения генов DCW кластера, наличие повторов – типичные сайты связывания ТФ-ов, а также тот факт, что зачастую ТФ-ы бактерий регулируют оперон, в котором они располагаются, указывает на то, что эти повторы – сайты связывания MraZ. Это предположение мы проверили экспериментально. Белок MraZ *M. gallisepticum* был получен с рекомбинантной молекулы ДНК в кишечной палочке. С помощью метода торможения ДНК в полиакриламидном геле, где в качестве субстрата использовали участок промоторной области DCW кластера длиной 40 п.о. и содержащих три повтора, было подтверждено, что белок связывается с дцДНК из промоторной области DCW кластера, состоящей из трёх прямых повторов. Длина спейсера между повторами – 3 п.н. Также мы показали, что длина спейсера между повторами от 2 до 4 п.н. (N2 и N4 соответственно) не влияет на эффективность связывания белка с ДНК (Рисунок 3.3 А-В).



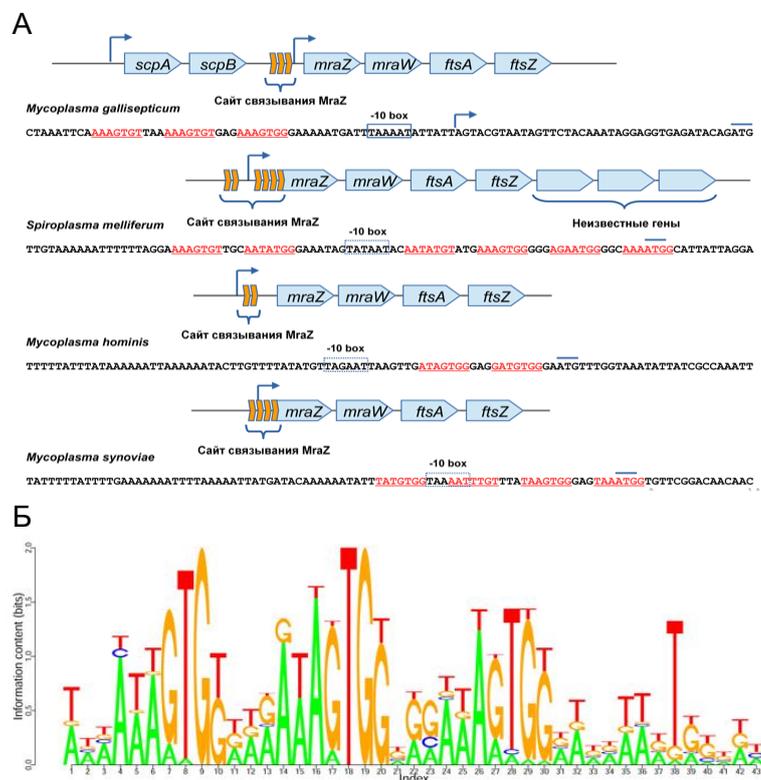
**Рисунок 3.3. ДНК-связывающие свойства белка MraZ.** А – Анализ эффективности связывания 500 нМ различных фрагментов дцДНК с 20 нМ белка MraZ методом торможения ДНК в 8% ПААГ в нативных условиях; P<sub>mraZ</sub>-WT, P<sub>01290</sub>WT – фрагменты промоторных областей генов *mraZ* и *GCW\_01290* соответственно; Б – Нормализованные интенсивности флуоресценции ДНК и ДНК-белковых комплексов; В – Последовательности ДНК, использованные в эксперименте; нуклеотиды сайта связывания MraZ, подвергшиеся замене, выделены красным цветом. Опубликовано в [112]

Белок MraZ также способен неспецифически связываться с ДНК. Для определения вклада нуклеотидов сайта в специфичность взаимодействия белка с субстратом (дцДНК) мы вносили мутации в разные позиции сайта и оценивали эффективность связывания белка MraZ

с субстратом. Наибольший вклад в специфичность взаимодействия вносит вторая тройка нуклеотидов повтора сайта связывания. При введении мутаций в такое положение сайта всех трёх повторов (Mut1, Mut2, Mut3) эффективность связывания становится сопоставима со связыванием с неспецифической дцДНК (Рисунок 3.3 А-Б).

### 3.2.2 Сайт связывания MraZ консервативен у бактерий класса Mollicutes

Мы провели поиск сайтов связывания ТФ MraZ у 30 представителей класса молликут, включая рода *Mycoplasma*, *Ureaplasma*, *Acholeplasma* и *Spiroplasma*. Повторы очень консервативны, они были обнаружены у всех представителей Mollicutes. Количество повторов, с которыми может связаться белок, а также расположение повторов относительно сайта начала транскрипции отличается у разных видов (Рисунок 3.4 А). Сайты связывания всегда локализованы рядом с опероном генов DCW кластера. Используя предсказанные сайты связывания белка MraZ у 30 представителей класса молликут, нами была построена лигатура, которая указывает на вклад каждой из букв в консервативность сайта – пятая и шестая позиции (Т, G соответственно), наиболее консервативны. Эти данные подтверждают описанные в пункте 3.2.1 экспериментальные данные о важности второй тройки нуклеотидов повтора сайта (Рисунок 3.4 Б).



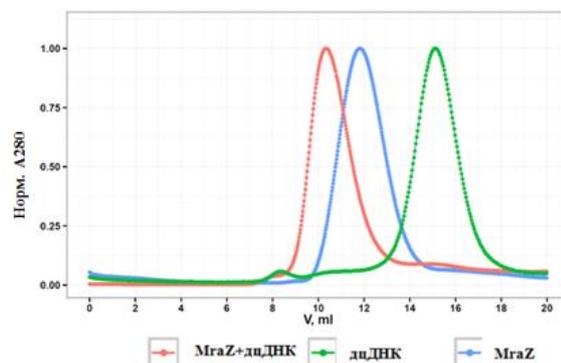
**Рисунок 3.4.** Структура расположения и консервативность сайтов связывания MraZ. А – Структура DCW кластера у разных представителей класса молликут. Сайт связывания MraZ выделен красным цветом. Сплошной линией указан старт-кодон; Б – Лигатура (LOGO) для набора последовательностей промотора гена *mraZ* 30 видов молликут. Опубликовано в [112]

### 3.2.3 Другие мишени белка MraZ в геноме *M.gallisepticum*

Мы также провели поиск сайтов связывания белка MraZ в промоторных областях всех генов *M.gallisepticum*. Сайт, состоящий из двух повторов и спейсера в 2 нуклеотида и заменой последнего нуклеотида [T/G] на А во втором повторе, обнаружен вблизи оперона, содержащего три гена с неизвестной функцией – *GCW\_01290- GCW\_01295- GCW\_01300*. MraZ в 2,5 раза менее эффективно взаимодействует с этим сайтом (Рисунок 3.3). Сайты, состоящие из одного повтора, расположены в промоторных областях генов *GCW\_00555* (функция неизвестна) и *GCW\_00595* (кодирует сериновую протеазу). Оба белка, кодируемые этими генами, содержат предсказанный трансмембранный участок и локализованы на поверхности цитоплазматической мембраны. Белок MraZ связывает фрагмент ДНК, содержащий единственный повтор вида AAAGTG[T/G], эффективнее, чем последовательности, не содержащие мотив (Рисунок 3.4 А). Наличие элементарных мотивов сайта связывания MraZ указывает на возможную ко-регуляцию экспрессии описанных генов и генов кластера деления.

### 3.2.4 Определение стехиометрии комплекса белка MraZ с дцДНК

Совместно с В.А. Манувера (лаб. генной инженерии ФГБУ ФНКЦ ФХМ ФМБА России) была определена стехиометрия ДНК-белкового комплекса с использованием эксклюзионной хроматографии на колонке Tricorn 10/300, заполненной смолой Superdex 200. С использованием готового набора белков с известной молекулярной массой была построена калибровочная кривая для расчёта масс анализируемого образца и рассчитаны молекулярные массы макромолекул. Хроматографическому разделению подвергали смесь MraZ и ДНК-субстрата – фрагмент ДНК длиной 40 пар нуклеотидов и содержащий три сайта связывания MraZ, а также отдельно образец белка и ДНК-субстрата. Теоретическая расчётная масса мономера белка MraZ – 18,4 кДа. Масса белка в чистом препарате, определенная в нашем эксперименте, составила 228,8 кДа. Из этого можно сделать вывод, что в растворе MraZ существует в виде 12-членного олигомера. Масса ДНК-белкового комплекса, вычисленная с использованием данных эксклюзионной хроматографии – 506,9 кДа; это указывает на то, что с одним участком дцДНК связывается 24 мономера MraZ (Рисунок 3.5), Таблица 3.2.

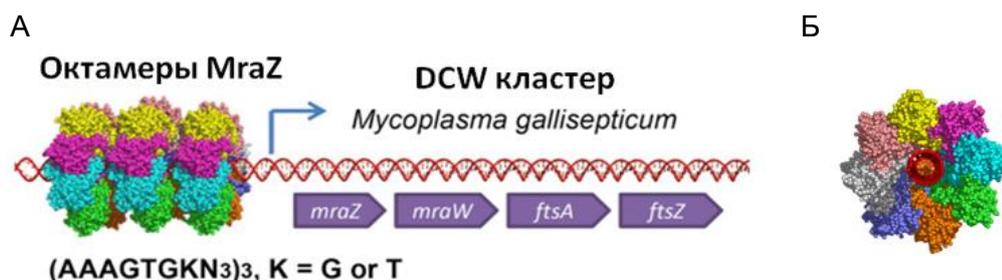


**Рисунок 3.5. Определение массы ДНК-белкового комплекса методом эксклюзионной хроматографии.** Разделение образца с соотношением MraZ:дцДНК равное 14:1, чистого препарата MraZ и дцДНК проводили с помощью эксклюзионной хроматографии на колонке Tricorn 10/300, заполненной смолой Superdex 200, в растворе 20 mM Tris-HCl, 150 mM KCl, и 5 mM MgCl<sub>2</sub>, pH 7,5. Нормализованные профили элюции дцДНК, белка MraZ и дцДНК-белкового комплекса обозначены в соответствии с легендой.

**Таблица 3.2. Рассчитанные молекулярные массы белка MraZ, дцДНК и дцДНК-белкового комплекса**

Образец	Объем элюции, мл	Kav	Mr расчетная, кДа
дцДНК	15,23	0,463	72,7
MraZ	12,62	0,270	228,8
дцДНК+MraZ	10,82	0,137	506,9

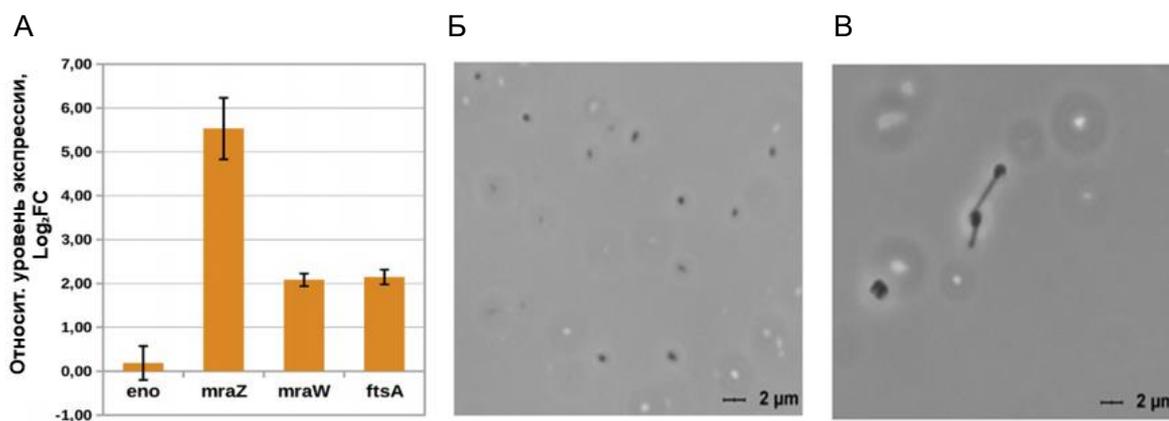
В 2004 году была расшифрована структура белка MraZ *Mycoplasma pneumoniae* [113], а в 2005 году – кишечной палочки [114]. Тогда еще функция этого белка была неизвестна, было лишь показано, что ген *mraZ* очень консервативен среди бактерий. Структура белка оказалась хеликазо-подобной, мономеры белка MraZ смыкались в кольцо. Используя наши данные о том, что с участком дцДНК, содержащим три прямых повтора AAAGTG[T/G], связывается 24 мономера белка, а олигомеры MraZ образуют кольцо, с помощью программы PyMOL (<https://pymol.org/2/>), мы представили гипотетическую схему взаимодействия белка MraZ с дцДНК (Рисунок 3.6). Мономеры ТФ образуют кольцо из восьми субъединиц на каждом повторе сайта связывания, диаметр полости в кольце (3,4 нм, сопоставим с диаметром дцДНК (2 нм)).



**Рисунок 3.6. Предполагаемая схема взаимодействия MraZ с дцДНК.** **А** – Расположение комплекса MraZ-дцДНК; вид вдоль цепи ДНК. Отмечена консенсусная последовательность сайта связывания и состав оперона DCW кластера. Синяя стрелка указывает направление синтеза мРНК. **Б** – Тот же комплекс, вид на полость, образуемую восемью мономерами MraZ. Для построения схемы использовали расшифрованную структуру белка MraZ *M. pneumoniae* (PDB ID: 1N0G).

### 3.2.5 Сверхэкспрессия белка MraZ влияет на форму клеток бактерии

Для подтверждения влияния белка MraZ на экспрессию генов DCW кластера *in vivo* и выявления фенотипических эффектов были получены трансформанты *M. gallisepticum* со сверхэкспрессией гена *mraZ*. Уровень экспрессии оценивали методом количественной ПЦР в реальном времени. В трансформантах со сверхэкспрессией гена *mraZ* уровень мРНК *mraZ* был в 40 раз выше, чем у клеток дикого типа. Уровень экспрессии остальных генов DCW кластера у трансформантов был выше в 4 раза, а представленность транскриптов генов «домашнего хозяйства» во всех типах клеток была одинакова (Рисунок 3.7 А). Форма клеток трансформантов отличалась от формы клеток бактерий дикого типа – клетки трансформантов вытягиваются, происходит филаментация (Рисунок 3.7 В,С). В статье [111] показано, что у *E. coli* MraZ выступает в качестве репрессора транскрипции генов DCW кластера, однако, как мы показали, у *M. gallisepticum* MraZ – активатор транскрипции генов кластера деления. Известно, что баланс между концентрациями MraZ и MraW является решающим фактором для выживания клеток. У микоплазм сверхэкспрессия MraZ приводит к увеличению экспрессии DCW кластера, что частично восстанавливает баланс и не ведет к гибели клеток. Фенотипически супер-продукция белка MraZ у кишечной палочки и микоплазмы проявляется сходным образом. Нарушение деления клеток, вероятно, происходит на заключительных этапах цитокинеза, что приводит к удлинению клеток.



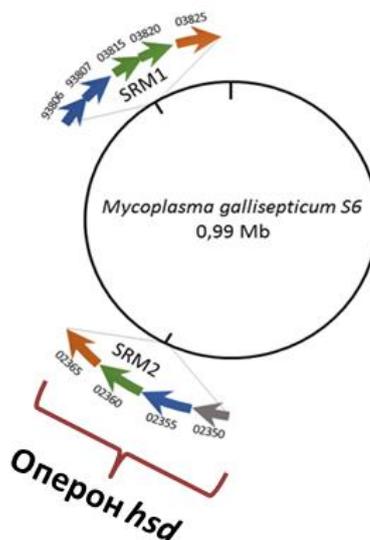
**Рисунок 3.7 Эффект сверхэкспрессии *mraZ* в клетках *M. gallisepticum*.** **А** – Влияние сверхпродукции MraZ на транскрипцию генов DCW кластера, измеренное с помощью ОТ-ПЦР в реальном времени. Ген енолазы (*eno*) приведен в качестве контроля; **Б, В** – Фазово-контрастные микрофотографии клеток *M. gallisepticum* дикого типа и со сверхэкспрессией *mraZ*, соответственно. Опубликовано в [112]

Таким образом, для *M. gallisepticum* экспериментально подтверждено наличие белка-активатора клеточного деления, что вносит существенный вклад в раскрытие регуляторных

событий у бактерий с редуцированным геномом и указывает на важность синхронизации процесса деления.

### 3.3 Поиск сайтов связывания белка Xre (HsdC)

Как и у большинства бактерий, у микоплазм есть системы рестрикции–модификации (P-M), которые защищают бактерию от экзогенной ДНК, в частности ДНК бактериофагов. Кроме этого, системы P-M могут контролировать репликацию ДНК, экспрессию генов и зачастую перемещаются между геномами бактерий за счёт горизонтального переноса. Системы P-M разделяют на несколько типов согласно их структурно-функциональной организации, однако все системы имеют в своем составе сайт-специфическую эндонуклеазу (R-субъединица) и сайт-специфическую ДНК-метилтрансферазу (M-субъединица). У *M. gallisepticum* S6 аннотировано две системы P-M (Рисунок 3.8). Гены обеих систем собраны в опероны. У некоторых бактерий системы P-M имеют специальный фактор регуляции транскрипции – С-белок. Сайт связывания С-белка – С-бокс консервативен и, как правило, состоит из двух инвертированных повторов – GACT—AGTC. У одной из систем P-M *Mycoplasma gallisepticum* S6 первым геном оперона *hsd* является ген *xre* (*GCW\_02350/hsdC*), который гомологичен ранее описанным С-белкам других бактерий.

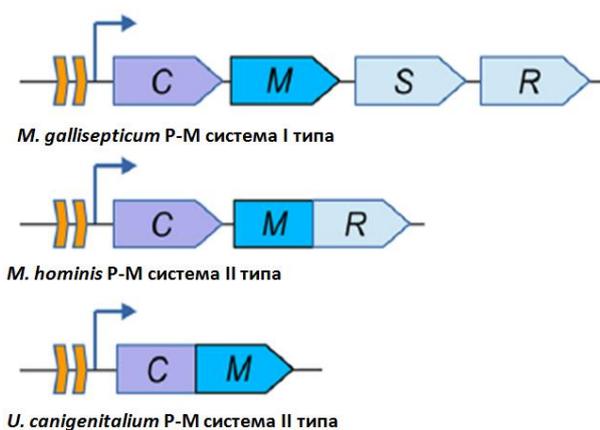


**Рисунок 3.8.** Схема расположения генов систем рестрикции–модификации у *M. gallisepticum* S6. Стрелками отмечены гены P-M систем. Синим цветом выделены гены, кодирующие M-субъединицу; зеленым – сайт-специфический ДНК-связывающий белок (S-субъединицу); оранжевым – R-субъединицу. Ген С-белка отмечен серым цветом.

#### 3.3.1 Распределение гомологов белка С (Xre; HdsC) у бактерий класса Mollicutes

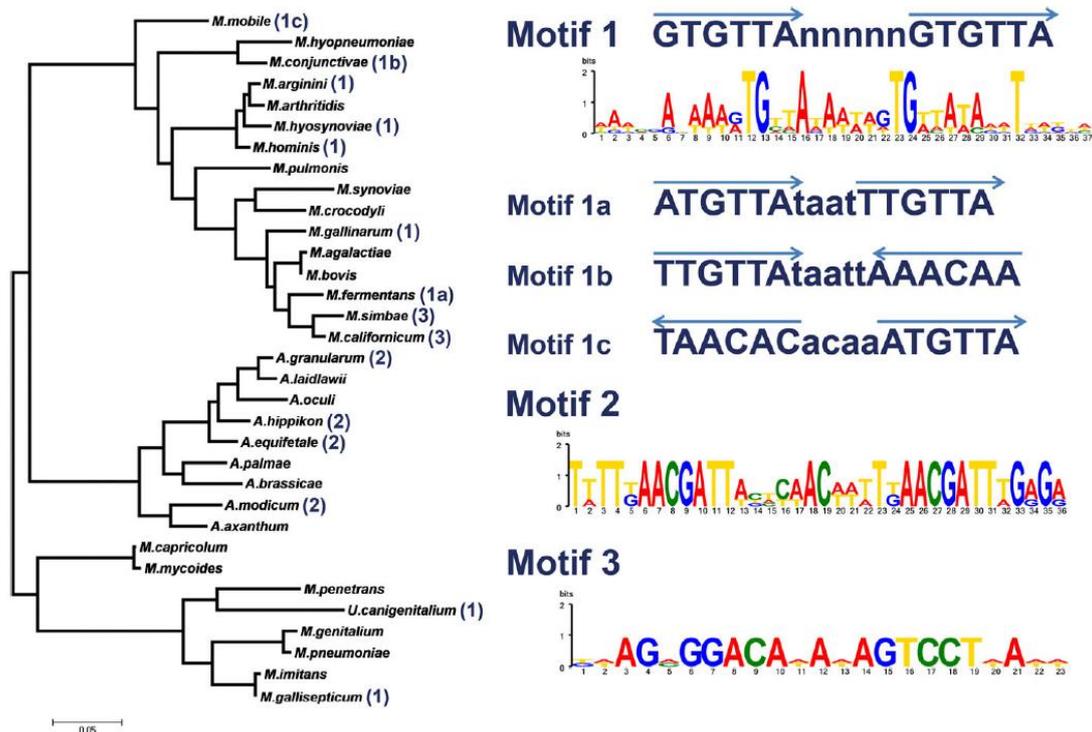
Мы провели поиск белков, гомологичных Xre. Среди молликут этот белок встречается у разных семейств, однако сходство последовательностей не высоко – от 40 до 60%. Все идентифицированные гены С-белков молликут ассоциированы с P-M системами первого, либо

второго типов. Гены регуляторов либо входят с генами Р-М систем в один оперон, либо стоят отдельно, либо слиты с другими генами системы Р-М (Рисунок 3.9).



**Рисунок 3.9. Варианты расположения гена С-белка (*xre*) у некоторых представителей молликут.** М – ген, кодирующий сайт-специфическую ДНК-метилтрансферазу; R – ген, кодирующий сайт-специфическую эндонуклеазу; S – сайт-специфический ДНК – связывающий белок. Оранжевые сегменты – консервативные повторы, возможный сайт связывания белка С. Опубликовано в [115]

В промоторных областях генов С-белков мы обнаружили консервативные области – возможные сайты связывания. Мотивы этих сайтов отличаются, они могут быть представлены как прямыми, так и инвертированными повторами (Рисунок 3.10).



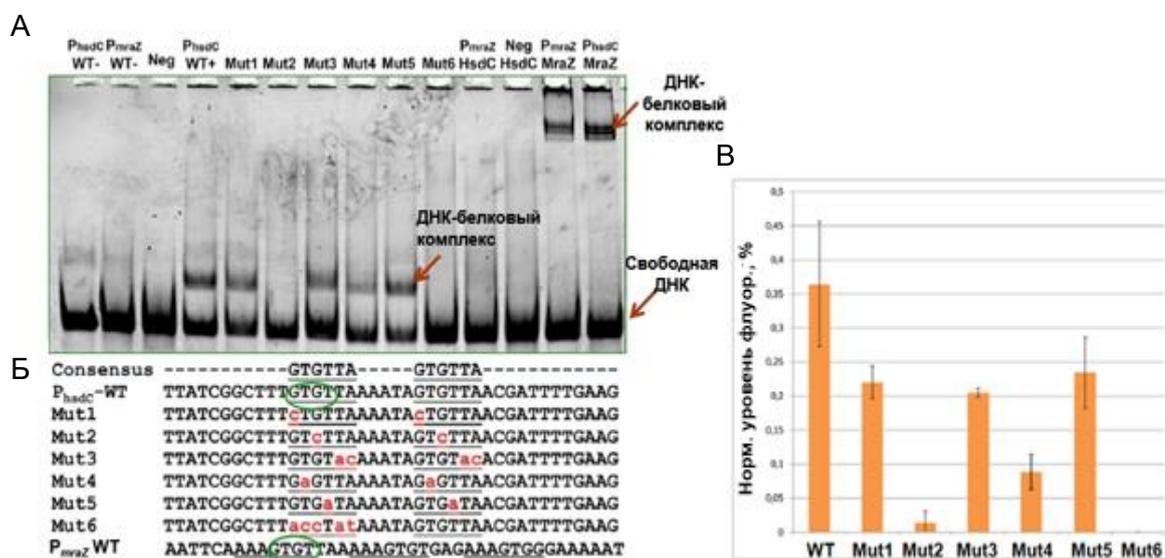
**Рисунок 3.10. Филогенетическое дерево представителей класса Mollicutes, в геноме которых есть ген С-белка.** Дерево построено по последовательностям 16S РНК. В скобках указан номер мотива, предсказанного как сайт связывания С-белка, у микоплазм и ахолеплазм [115].

Выделяется три группы мотивов. У представителей рода *Achleplasma* сайт представлен длинными прямыми повторами, разделенные спейсером в три нуклеотида

(AACGAATN<sub>12</sub>)<sub>3</sub>, хотя встречаются и спейсеры длиной в 1 п.о. (Motif 2). Первую группу мотивов можно разбить на четыре подгруппы: Motif 1, Motif 1a, Motif 1b, Motif 1c. Их отличает первый нуклеотид повтора, ориентация и длина спейсера. Третий тип мотива (Motif 3) найден у некоторых микоплазм и кардинально отличается от других мотивов – GGACN<sub>5</sub>GTCC. Этот мотив похож на известный С-бокс белка AhdI – AGTCCN<sub>2</sub>GGACT, но с противоположной ориентацией повторов [116].

### 3.3.2 Белок Xre связывает прямые повторы, расположенные в промоторной области оперона *hsd*

Мы рассмотрели промоторную область оперона генов системы Р-М *Mycoplasma gallisepticum* S6 для того, чтобы экспериментально доказать способность белка Xre (С-белок) связываться с предсказанной нами последовательностью. Предсказанный сайт связывания белка Xre – (GTGTTAN<sub>5</sub>)<sub>2</sub>, схож с сайтом связывания белка MraZ – (AAAGTG[T/G]N<sub>3</sub>)<sub>3</sub>, где наибольший вклад в ДНК-белковое взаимодействие вносила последовательность GTG[T/G]. Поэтому мы также решили проверить, способен ли С-белок связываться с промоторной областью генов DCW кластера. Оказалось, С-белок связывается с промотором гена *xre* (*hsdC*), но не связывает P<sub>mraZ</sub>\_WT, в то время как белок MraZ связывал оба типа последовательностей, но P<sub>hsdC</sub>\_WT – менее эффективно (Рисунок 3.11).

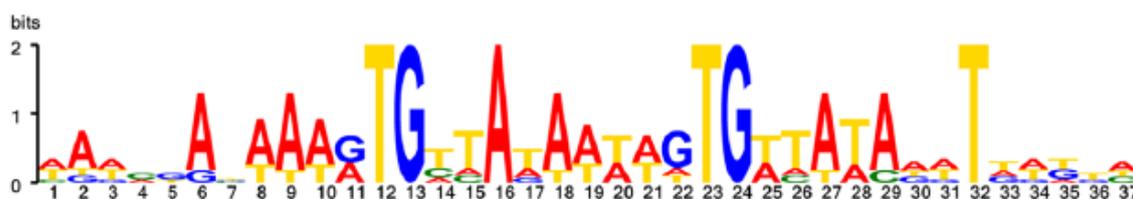


**Рисунок 3.11.** ДНК-связывающие свойства С-белка (Xre). **А** – Анализ эффективности связывания 500 нМ различных дцДНК фрагментов 1000 нМ белка Xre методом торможения ДНК в 8% ПААГ в нативных условиях; P<sub>hsdC</sub>\_WT, P<sub>mraZ</sub>\_WT– фрагменты промоторных областей генов *hsdC* и *mraZ* соответственно; Neg – фрагмент дцДНК, не содержащий предсказанный сайт связывания. **Б** – Последовательности ДНК, использованные в эксперименте; сайт связывания подчеркнут. **В** – Нормализованные значения интенсивности флуоресценции ДНК и ДНК-белковых комплексов [115]

Эти данные указывают, что Xre (HsdC) избирателен к длине спейсера, и при уменьшении его длины даже на один нуклеотид Xre становится неспособен связать ДНК.

Внося мутации в разные позиции сайта и оценивая эффективность связывания белка Xre (HsdC) с субстратом, мы показали, что наибольший вклад во взаимодействие вносят TG нуклеотиды сайта. При введении мутаций в такие положения сайта (Mut2, Mut4), эффективность связывания падала в 15 и 4,5 раза соответственно. В отличие от MraZ, HsdC связывает дцДНК специфично. Потеря одного из повторов сайта драматически сказывается на способности белка связывать субстрат (Mut6) (Рисунок 3.11).

Для предсказанных сайтов связывания белка С молликут, состоящих из прямых повторов, была построена консенсусная последовательность, отражающая консервативность TG оснований в каждом повторе, что подтверждается экспериментальными данными о вкладе TG нуклеотидов во взаимодействие с белком-регулятором (рисунок 3.12).



**Рисунок 3.12.** Лигатура (LOGO) для набора последовательностей промотора генов *xre* молликут. Наиболее консервативные нуклеотиды сайта – TG основания в каждом повторе [115].

### 3.3.3 Xre, вероятно, выступает в качестве репрессора транскрипции *hsd* оперона

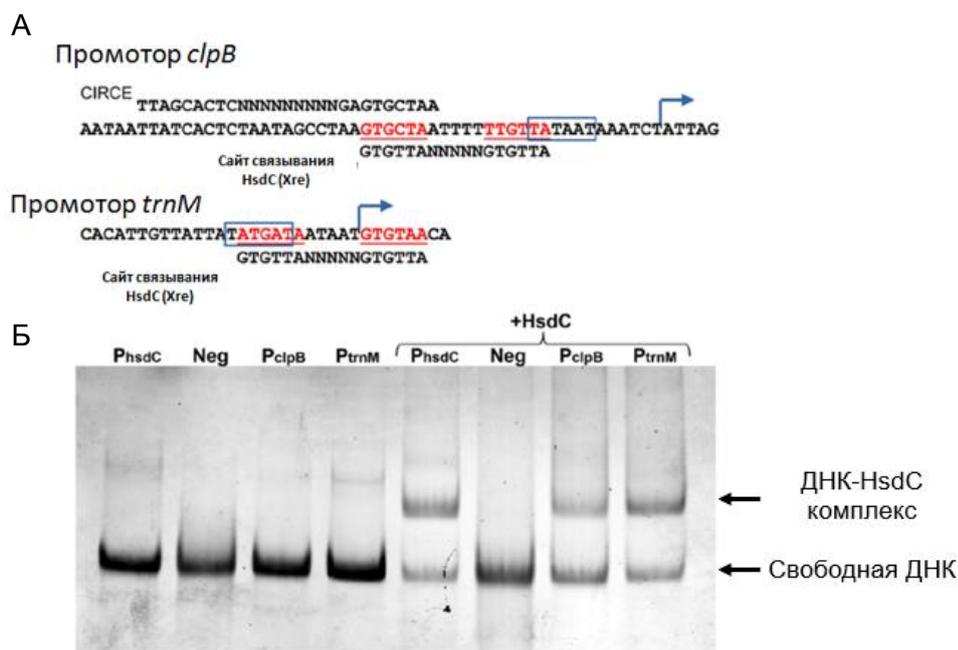
Используя данные о позиции старта начала транскрипции (TSS – Transcription Start Site) [82], структуре и силе промоторов [96], мы определили положение сайта связывания Xre относительно TSS у *M.gallisepticum* S6 в опероне *hsd*. Оперон *hsd* имеет два промотора. У первого детерминанты сильного промотора – классический -10 бокс с extended участком. Если сопоставить теоретическую силу промотора *hsd*, рассчитанную из силы составляющих его детерминант (неопубликованные данные), и экспериментальную, рассчитанную как покрытие прочтениями, полученными при секвенировании библиотеки кДНК с обогащением по 5'-концевым фрагментам РНК [82], то эта экспериментальная сила промотора будет на несколько порядков ниже, чем теоретическая. Это указывает на наличие репрессора. Оперон *hsd* имеет несколько промоторов. Первый включает в себя сайт связывания Xre, который частично перекрывает -10 бокс. Такое расположение сайта типично для репрессоров [22] (Рисунок 3.13). Второй промотор слабый и находится в пределах кодирующей последовательности (не отмечен).



**Рисунок 3.13.** Взаимное расположение сайта связывания Xre и детерминант промотора оперона *hsd*. Сайт Xre выделен красным и подчеркнут, сайт начала транскрипции (TSS) обозначен синей стрелкой.

### 3.3.4 Другие мишени белка Xre (HsdC) в геноме *M. gallisepticum* S6

Мы провели поиск сайтов связывания белка HsdC в промоторных областях всех генов *M. gallisepticum* S6. Мы рассматривали только те промоторы, где сайт связывания находился в окрестностях -10 бокса или перекрывался с TSS. Таких промоторов оказалось два – оперон, состоящий из генов тРНК (первый ген – *trnM*, *GCW\_00940*), и ген шаперона *clpB* (*GCW\_02785*). В обоих случаях сайт связывания перекрывался с -10 боксом. С помощью метода торможения ДНК мы показали, что белок HsdC связывается с найденными сайтами с такой же силой, что и с промотором оперона *hsd* (Рисунок 3.14).



**Рисунок 3.14.** Анализ способности HsdC связываться с дцДНК дополнительных мишеней в геноме *M. gallisepticum*. **А** – Структура промоторных областей генов *clpB* и *trnM*, содержащих сайт связывания HsdC. TSS обозначен синей стрелкой, потенциальные сайты связывания выделены красным. **Б** – Анализ связывания HsdC с найденными последовательностями, проведенный методом торможения ДНК в полиакриламидном геле. Опубликовано в [115]

Участок связывания Xre в промоторе *clpB* перекрывается с последовательностью CIRCE – местом связывания ТФ HrcA. Такое пересечение усложняет регуляторную сеть шаперонов – единственную из описанных ранее. Промоторные области *clpB* и *trnM* у *M. gallisepticum* S6 и *M. gallisepticum* R(low) не отличаются, однако в геноме штамма S6

содержится ген *xre*, а у *R(low)* – нет. Это означает, что способность связывать новую мишень, не относящуюся напрямую к системе Р-М, была приобретена случайно и закреплена у штамма S6.

Таким образом, нами впервые был проведен сравнительный анализ систем рестрикции–модификации бактерий класса Mollicutes, предсказаны уникальные сайты связывания факторов регуляции транскрипции (С-белков) генов систем Р-М. Экспериментально подтвержден и охарактеризован сайт связывания С-белка *M. gallisepticum* S6. Большинство С-белков молликут распознают мотив, отличающийся от известного С-бокса. Наиболее часто встречающийся у молликут мотив состоит из прямых, а не инвертированных повторов, что указывает на альтернативный механизм димеризации белка. Точная роль С-белка в контроле экспрессии генов оперона *hsd* не ясна. Изученные С-белки Р-М систем AhdI и PvuII являются активаторами при связывании с последовательностью, расположенной выше -35 бокса. AhdI выступает репрессором, связываясь между -10 и -35 боксами, мешая связаться РНК-полимеразе. Обе хорошо изученные системы имеют две транскрипционные единицы, только R-субъединица находится под контролем С-белка. Это приводит к тому, что S и M субъединицы экспрессируются первыми и модифицируют геномную ДНК до того, как R субъединица сможет ее расщепить. Белок HsdC *M. gallisepticum*, вероятно, действует как авторепрессор всего оперона, что в свою очередь отличает его от других С-белков Р-М систем.

#### 3.4. Определение сайта связывания и генов-мишеней белка семейства FUR

Белки FUR (ferric uptake regulator) представляют собой распространённое семейство прокариотических регуляторов транскрипции. Участники этого семейства белков имеют ряд общих особенностей. Их последовательности, как правило, содержат два домена – расположенный на N-конце ДНК-связывающий НТН домен (спираль–поворот–спираль) и на С-конце домен димеризации, в шарнирной области между доменами находится гистидин-богатый мотив ННХНХ2СХ2С [117]. Это разнообразная группа регуляторов, специализирующихся на контроле метаболизма железа (Fur), цинка (Zur), марганца (Mur), никеля (Nur), меди (BosR), гема (Irr), а также реакции на окислительный стресс (PerR) [118]. Fur, фактор транскрипции, давший название семейству, первоначально характеризовался как регулятор гомеостаза железа. Поступление железа в клетку, а также его хранение должны находиться под жёстким контролем из-за высокой токсичности гидроксильных радикалов, образующихся в результате реакции Фентона при взаимодействии  $Fe^{2+}$  с супероксидом и перекисью водорода [119]. Долгое время считалось, что это специализированный регулятор системы обмена  $Fe^{2+}$ , однако в последнее время, появляется всё больше работ, указывающих

на то, что Fur – глобальный регулятор [120]. Особый интерес представляет изучение Fur белков у патогенных бактерий [121].

### 3.4.1 Распределение белков семейства FUR у молликут

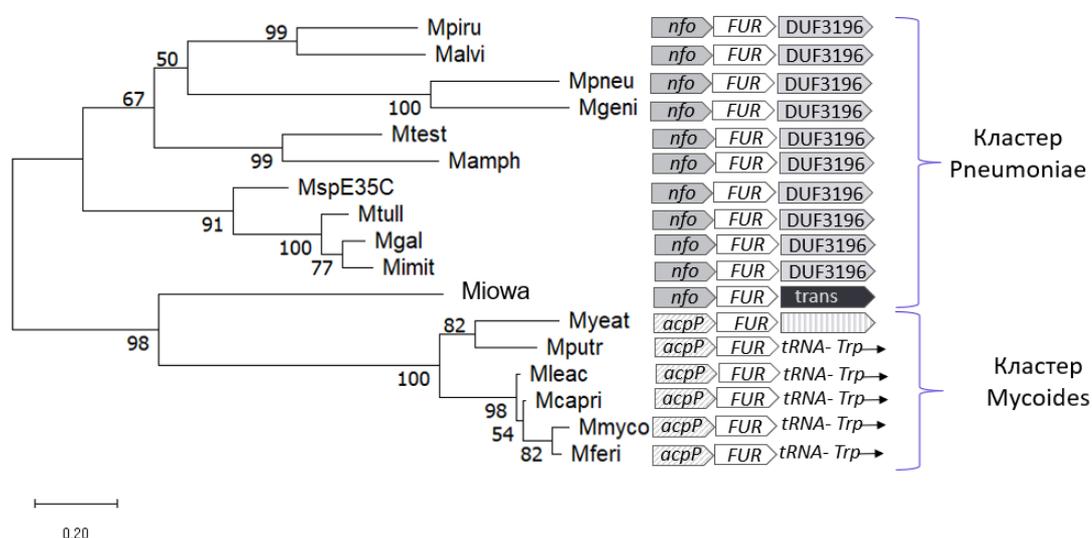
Для того чтобы оценить, насколько белки семейства FUR представлены и консервативны в бактериях класса Mollicutes, мы провели поиск белков семейства FUR у молликут. Несмотря на то, что белки семейства FUR считаются широко распространёнными у прокариот, среди микоплазм многие представители не имеют FUR белков. Представители рода *Acholeplasma*, как правило, имеют набор из трёх белков FUR-семейства, кроме *A. axanthum* и *A. hippikon*, у этих бактерий 2 и 1 FUR белок, соответственно. FUR белки отличаются по кофактору, входящему в их состав, поэтому в одном организме может находиться более одного белка семейства FUR разной или одинаковой специализации. В геномах микоплазм найдена одна копия гена, кодирующего белки семейства FUR. Далее мы подробнее рассмотрели распределение белков семейства FUR у бактерий семейства *Mycoplasmataceae*, патогенов многих животных. Из 160 видов бактерий семейства *Mycoplasmataceae* только 17 видов (Таблица 3.3) имели гомологи белка семейства FUR *Mycoplasma gallisepticum*.

**Таблица 3.3. Представители семейства *Mycoplasmataceae*, у которых обнаружены гомологи белка Fur**

Вид	Условное обозначение	Вид	Условное обозначение
<i>Mycoplasma pirum</i>	Mpiru	<i>Mycoplasma gallisepticum</i>	Mgal
<i>Mycoplasma alvi</i>	Malvi	<i>Mycoplasma imitans</i>	Mimit
<i>Mycoplasma pneumoniae</i>	Mpne	<i>Mycoplasma iowa</i>	Miowa
<i>Mycoplasma genitalium</i>	Mgeni	<i>Mycoplasma yeatsii</i>	Myeat
<i>Mycoplasma testudinis</i>	Mtest	<i>Mycoplasma putrefaciens</i>	Mputr
<i>Mycoplasma amphoriforme</i>	Mamph	<i>Mycoplasma leachii</i>	Mleac
<i>Mycoplasma</i> sp. E35C	MspE35C	<i>Mycoplasma capricolum</i>	Mcapri
<i>Mycoplasma tullyi</i>	Mtull	<i>Mycoplasma mycoides</i>	Mmyco
<i>Mycoplasma gallisepticum</i>	Mgal	<i>Mycoplasma feriruminatoris</i>	Mferi
<i>Mycoplasma imitans</i>	Mimit		

Эти 17 видов являются представителями кластеров *Pneumoniae* или *Mycoides*. Белки семейства FUR не обнаружены у представителей кластеров *Hominis* и *Mesoplasma*. У *M. gallisepticum fur* входит в состав оперона вместе с *nfo*, кодирующим АП-эндонуклеазу IV и геном *GCW\_02095*, кодирующим белок с неизвестной функцией, но содержащим консервативный домен DUF3196. Анализ геномного окружения позволил выделить ортологичные группы и показал консервативность оперона внутри кластера *Pneumoniae*.

Исключение составляет *Miowa*. Среди представителей кластера *Mycoides* у близкородственных видов наблюдается неоднородность FUR-кодирующего локуса, вероятно, вследствие горизонтального переноса генов, действия мобильных элементов (Рисунок 3.15).



**Рисунок 3.15.** Филогенетическое дерево максимального правдоподобия для всех проанализированных гомологов белков семейства FUR *Mycoplasmataceae*. Уровни бутстрэп поддержки указаны числами у каждого узла. На правой панели указан геномный контекст генов, кодирующих белки семейства FUR. *nfo* – ген, кодирующий эндонуклеазу IV; DUF3196 – ген, кодирующий белок с неизвестной функцией; *trans* – ген, кодирующий транспозазу; *acpP* – ген, кодирующий ацилпереносающий белок; *tRNA-Trp* – ген тРНК<sup>Трп</sup>.

### 3.4.2 Fur – регулятор экспрессии генов, кодирующих транспорт катионов металлов

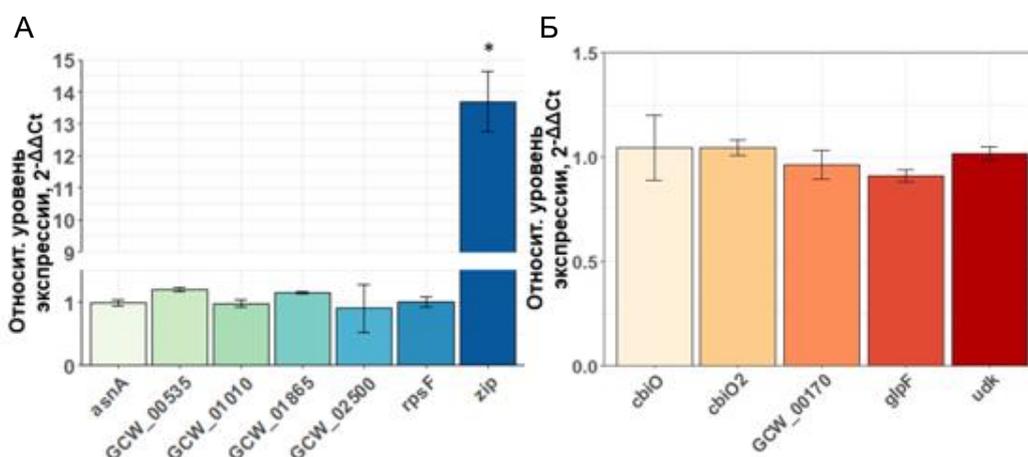
Консервативность ортологичных групп у представителей кластера *Pneumoniae* позволила использовать методы сравнительной геномики для анализа транскрипционных регулонов белка семейства FUR. Ранее проведенные работы по реконструкции регуляторных сетей у *M. pneumoniae* [85] и *M. genitalium* [122] выявили потенциальные мишени Fur. Все четыре мишени *M. genitalium* (в том числе оперон *cbio/MG302-MG304*, содержащий гены субъединиц ABC транспортера металлов) имели ортологов среди предсказанных мишеней *M. pneumoniae*. Мы не обнаружили у *M. gallisepticum* гомологов белков двух мишеней *M. genitalium* (MG149 и MG302) (Таблица 3.4). То есть в группе родственных бактерий, принадлежащих одному кластеру внутри семейства *Mycoplasmataceae*, при сохранении гена-регулятора внутри оперона, не происходит сохранения всех ортологичных генов-мишеней. Общая консенсусная последовательность сайта связывания транскрипционного фактора Fur у *M. pneumoniae* и *M. genitalium* представляет собой А/Т-богатый палиндром вида ТТАТТТWКААААWТТМWАААТАА. Мы обнаружили точное вхождение сайта в промоторной области гена *zip* *M. gallisepticum*. Промоторные области генов *GCW\_01010*,

*GCW\_02500*, *GCW\_01865*, *rpsF*, *asnA*, *GCW\_00535* содержали этот сайт, но с несколькими нуклеотдными заменами.

**Таблица 3.4. Сравнение консервативности аминокислотных последовательностей мишеней ТФ Fur трёх представителей микоплазм. Каждая строка соответствует одной мишени. В скобках указан процент схожих по свойствам аминокислот первичной структуры белка в выравнивании при поиске программой Blast**

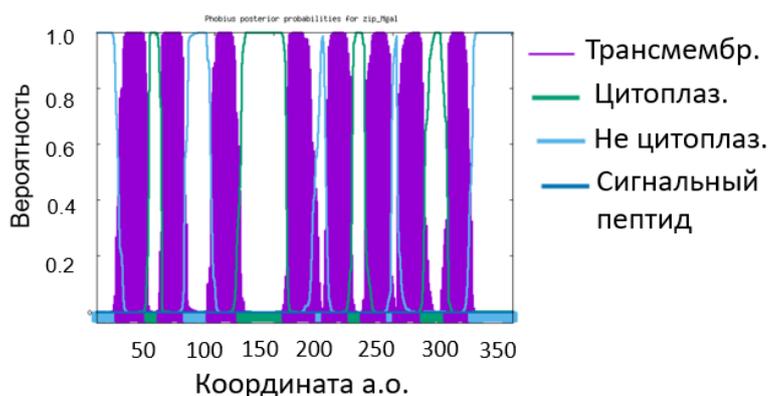
<i>M. genitalium</i> (% сходства)	Аннотация	<i>M. pneumoniae</i> (% сходства)	<i>M. gallisepticum</i> (% сходства)
MG149	Липопротеин MG149 (Hrl)	MPN162 (52%)	-
MG302/ <i>cbiO</i>	Субъединица ABC транспортера CbiQ	MPN431 (77%)	-
MG303	Субъединица ABC транспортера MG303	MPN432 (75%)	<i>GCW_03275/cbiO</i> (54%)
MG304	Субъединица ABC транспортера MG304	MPN433 (82%)	<i>GCW_03280/cbiO2</i> (58%)
MG260 (74%)	Липопротеин	MPN363	<i>GCW_00170</i> (68%)
MG382 (88%)	Уридинкиназа Udk	MPN561	<i>GCW_02565</i> (62%)
MG260 (61%)	Липопротеин	MPN408	<i>GCW_00170</i> (48%)
MG033 (70%)	Транспортер глицерола GlpF	MPN043	<i>GCW_00065</i> (54%)

Для того чтобы экспериментально проверить, регулируются ли отобранные нами мишени белком Fur у *M. gallisepticum*, с помощью CRISPR интерференции были получены трансформанты микоплазмы, где экспрессия *fur* была в среднем ниже в 43 раза, по сравнению с экспрессией этого гена в штамме микоплазмы, экспрессирующем *dcas9*, но не экспрессирующем гидовую РНК, нацеленную на *fur*. Обозначим такие трансформанты *fur\_sg2*. Далее мы оценили влияние нокдауна *fur* на уровень мРНК предполагаемых генов-мишеней. Экспрессию генов-мишеней внутри каждого образца нормализовали на уровень транскрипта *tuf*. Изменение экспрессии генов рассчитано относительно образцов мРНК трансформантов *M. gallisepticum*, не экспрессирующих гидовую РНК (Рисунок 3.16). Обозначим такие трансформанты *Mga\_dcas9*.



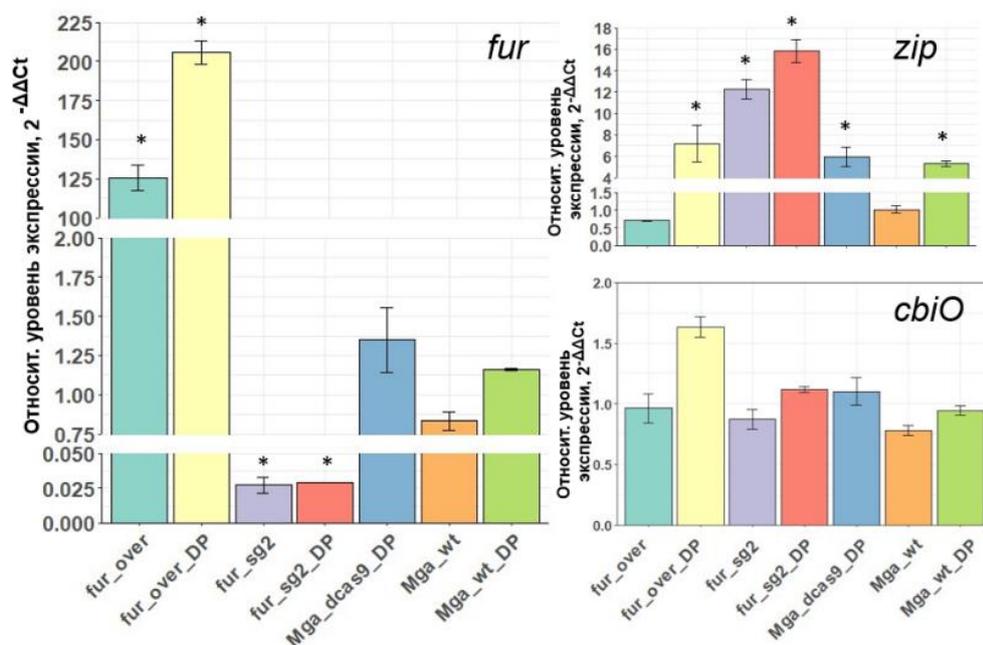
**Рисунок 3.16 Влияние нокдауна *fur* на экспрессию предполагаемых генов-мишеней.** **А** – Относительный уровень мРНК генов *M. gallisepticum*, в промоторной области которых обнаружен предполагаемый сайт связывания Fur. **Б** – Относительный уровень мРНК генов *M. gallisepticum fur\_sg2*, ортологичных генам-мишеням Fur *M. pneumoniae* и *M. genitalium*. Приведено среднее не менее 3 независимых экспериментов со стандартным отклонением. \* –  $p < 0,05$  согласно ранговому непарному критерию Вилкоксона.

Среди проверенных предполагаемых мишеней только уровень мРНК *zip* значительно увеличился в среднем в 13 раз при репрессии *fur* у *M. gallisepticum*. Белки семейства ZIP участвуют в транспорте катионов металлов, что функционально объединяет их с транспортером CbiO- мишенями Fur *M. pneumoniae* и *M. genitalium*. Согласно сервису Phobius, инструменту по предсказанию локализации белков, белок Zip *M. gallisepticum* состоит из 8 трансмембранных альфа-спиралей (Рисунок 3.17), что характерно для полных последовательностей белков семейства ZIP прокариот и эукариот.



**Рисунок 3.17. Предсказание наличия трансмембранных участков белка Zip *M. gallisepticum*.** Трансмембранные участки выделены фиолетовым цветом.

Металлы являются кофакторами белков семейства FUR, которые в апо-форме или в комплексе с катионом металла ингибируют или активируют транскрипцию генов-мишеней. Мы сравнили уровни экспрессии *fur*, *zip* и *cbiO* у трансформантов микоплазмы с нокдауном по *fur*, сверхэкспрессией, а также клеток дикого типа в условиях истощения среды культивирования по железу. Для этого к клеткам микоплазмы добавляли сублетальную дозу раствора 2,2'-дипиридила и инкубировали культуру в течение 30 мин. Экспрессию генов внутри каждого образца нормализовали на уровень *tuf*. Изменение уровня мРНК рассчитывали относительно образцов мРНК трансформантов *M. gallisepticum*, экспрессирующих *dcas9*, но не экспрессирующих гидовую РНК, нацеленную на *fur* в нормальных условиях (Рисунок 3.18).



**Рисунок 3.18.** Анализ влияния дефицита железа на экспрессию *fur*, *zip*, *cbiO*. Относительный уровень мРНК генов *fur*, *zip* и *cbiO* у разных штаммов *M. gallisepticum* при нормальных условиях и в условиях хелатирования железа (DP) измерен методом количественной ОТ-ПЦР в реальном времени. Приведено среднее не менее 3 независимых экспериментов со стандартным отклонением. \* –  $p < 0,05$  согласно ранговому непарному критерию Вилкоксона. Условные обозначения: Mga\_wt – *M. gallisepticum* дикого типа; fur\_over – трансформанты *M. gallisepticum* с сверхэкспрессией *fur*; Mga\_dcas9 -трансформанты *M. gallisepticum*, экспрессирующие *dcas9*; fur\_sg2 – трансформанты *M. gallisepticum* с нокадауном по *fur*; DP – 2,2'-дипиридил.

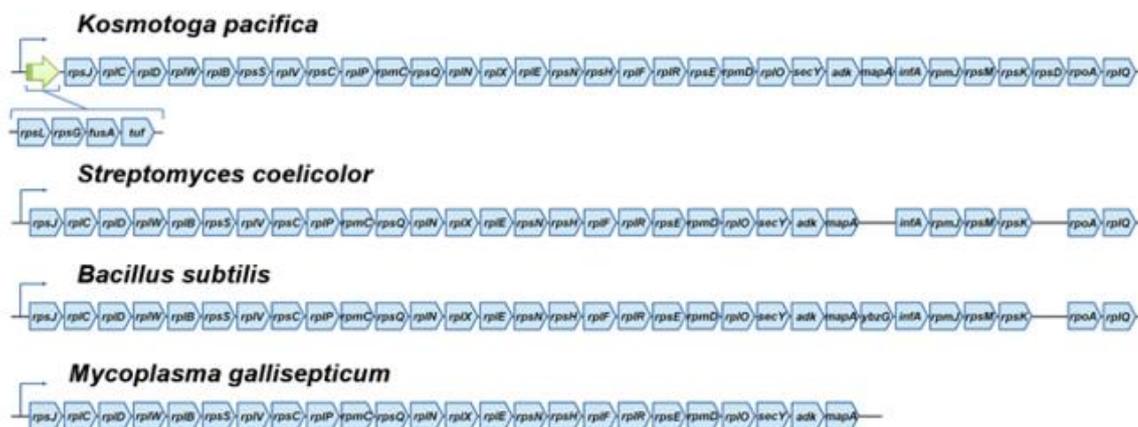
Уровень экспрессии *zip* в клетках Mga\_wt, Mga\_dcas9 увеличился в 5–6 раз в ответ на добавление 2,2'-дипиридила. При этом уровень *fur* не менялся. В трансформантах со сверхэкспрессией *fur* при базальном высоком уровне регулятора (в 125 раз выше, чем в трансформантах Mga\_dcas9), дипиридил несущественно влиял на уровень ТФ-а, представленность мРНК *zip* возросла в 7 раз. При подавлении экспрессии *fur* хелатирование железа ведет к незначительной активации транскрипции гена транспортера. Уровень транскрипции другого предполагаемого транспортера металлов – CbiO – не менялся как у разных трансформантов, так и при добавлении 2,2'-дипиридила. В совокупности эти результаты подтверждают классическую модель работы белков семейства FUR, где кофактором является железо. В нормальных условиях Fur репрессирует транскрипцию *zip*. Дефицит железа – кофактора ТФ-а – ведет к диссоциации ДНК-белкового комплекса и активации транскрипции *zip*, чей белковый продукт обеспечивает экспорт железа. Организация регулонов Fur у микоплазм служит хорошим примером того, как быстро эволюционируют регуляторные сети в бактериях.

### 3.5 Поиск сайтов связывания белка WhiA

Ген *whiA* консервативен у грамположительных бактерий [123]. Кристаллическая структура белка WhiA *Thermatoga maritima* была расшифрована в 2009 году [124]; было показано, что белок имеет два домена – N-концевой домен хоминг-эндонуклеазы семейства LAGLIDADG (HEN домен) и C-концевой домен спираль–поворот–спираль (HTH). Эндонуклеазный домен утратил свою активность вследствие замены а.о аспарагиновой кислоты активного сайта, однако сохранил способность связывать ДНК [125]. Для некоторых бактерий была изучена функция WhiA. У бактерий рода *Streptomyces* WhiA участвует в сложной регуляторной сети. Вместе с другим транскрипционным фактором – WhiB, WhiA осуществляет корегуляцию экспрессии генов, участвующих в споруляции и сегрегации хромосом [126], [127], [128]. С помощью метода ChiP-seq был определен сайт связывания WhiA – мотив GACAC. Таких сайтов в геноме *Streptomyces venezuelae* около 15 тыс., однако связывание WhiA происходит только с малой долей этих сайтов, что указывает на наличие дополнительных неизвестных детерминант, обеспечивающих специфичность связывания WhiA с ДНК [129], [128]. У *Bacillus subtilis* роль WhiA не связана с регуляцией споруляции, но белок важен для сегрегации хромосом, влияя на экспрессию гена *ftsZ*, а также *parA* и *parB* [128], [130]. У *Corynebacterium glutamicum* WhiA связывается с промоторной областью генов деления (например, *ftsZ*) и с белком WhcD, играя важную роль в делении клеток [131], [132]. Функция WhiA у представителя класса Mollicutes – *Mycoplasma pneumoniae* – была предложена Эйлерсом в кандидатской диссертации [133]. Он показал, что инактивация *whiA* путем транспозонного мутагенеза (штамм  $\Delta whiA$  *M. pneumoniae*) приводит к повышению экспрессии одного из рибосомных оперонов (первый ген – *rpsJ*), таким образом, возможно, WhiA является регулятором оперона *rpsJ*. Эйлерс также обнаружил, что в стационарной фазе роста, не происходит подавления экспрессии оперона *rpsJ* у  $\Delta whiA$  штамма *M. pneumoniae*, в отличие от штамма *M. pneumoniae* дикого типа. Возможно, WhiA – сенсор фазы роста бактерии.

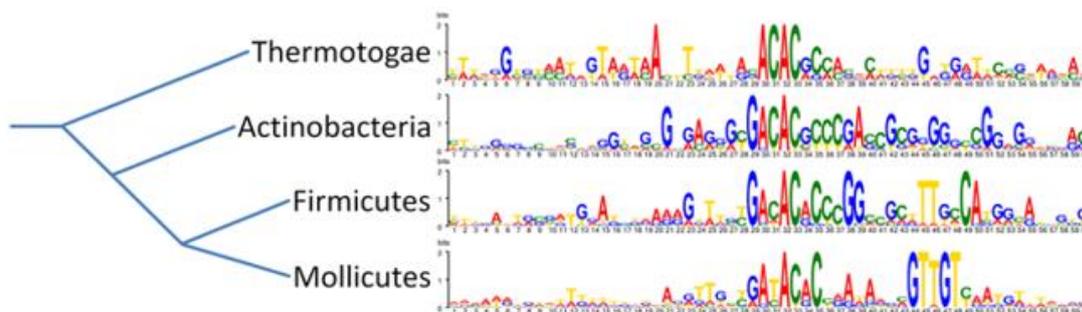
### **3.5.1 Промоторная область оперона *rpsJ* у разных бактерий консервативна и содержит core мотив**

Согласно предположению Эйлерса, WhiA *M. pneumoniae* регулирует экспрессию оперона рибосомных генов (первый ген – *rpsJ*, будем называть так весь оперон). Ген *whiA* консервативен, так же как и гены рибосомных белков, таким образом, сайт связывания белка WhiA должен быть сохранен у бактерий, у которых есть *whiA*. Мы провели сравнительный анализ промоторных областей рибосомного оперона *rpsJ* и состава самого оперона у разных представителей грамположительных бактерий – тип Firmicutes, класс Mollicutes, тип Actinobacteria, семейство Thermotogae, тип Bacteroidetes и тип Chlamydiae (Рисунок 3.19).



**Рисунок 3.19.** Структуры оперонов – предположительных мишеней WhiA грамположительных бактерий. Строение оперонов, содержащих ген *rpsJ*, у разных бактерий, сохранивших ген *whiA*.

У всех представителей рассмотренных бактерий оперон *rpsJ* содержит большое количество генов – до 30. Точный генный состав варьирует у разных видов, но в начале оперона, как правило, располагаются гены *rpsJ* (кодирует 30S рибосомный белок S10), *rplC* (кодирует 50S рибосомный белок L3), *rplD* (кодирует 50S рибосомный белок L4). Затем мы проанализировали промоторные области оперона рибосомных генов на наличие консервативных участков. На Рисунке 3.20 представлены обнаруженные консервативные элементы.



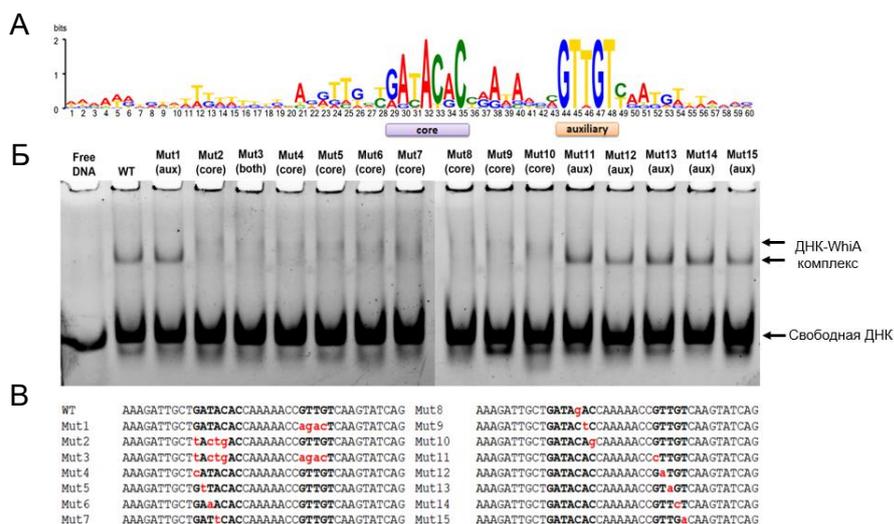
**Рисунок 3.20.** Консервативности сайта связывания WhiA у грамположительных бактерий. Лигатуры консервативных участков промоторных областей рибосомного оперона *rpsJ* построены из промоторных последовательностей оперона рибосомных генов

У большинства видов промоторная область имеет консервативную последовательность с ядром (core) – GAYACRCY, где Y = C или T, R = A или G. Различия в core мотиве могут быть обусловлены разным содержанием ГЦ пар в геноме рассматриваемых бактерий (Рисунок 3.19 Б). У представителей типа Actinobacteria содержание ГЦ пар в геноме может достигать 70% [134], а у представителей класса Mollicutes – 30% [135]. Обнаруженный core мотив гомологичен сайту, найденному с помощью ChiP-seq для WhiA *Streptomyces* [129], [128]. Помимо core мотива у некоторых групп бактерий выделяется другой консервативный мотив.

Например, у Mollicutes – GTTGT, последовательность схожа с каноническим -35 боксом бактериальных промоторов. В дальнейшем будем называть этот мотив дополнительным (aux). Основываясь на данных картирования позиций сайтов начала транскрипции для *M. gallisepticum* S6 [82], core и aux мотивы после точки начала транскрипции оперона *rpsJ*, поэтому aux мотив не может быть -35 боксом оперона *rpsJ*.

### 3.5.2 WhiA *M. gallisepticum* специфически связывается с промоторной областью оперона *rpsJ*

С помощью метода торможения ДНК в геле в нативных условиях мы подтвердили способность WhiA связываться с консервативной некодирующей областью оперона *rpsJ*. Для изучения специфичности связывания WhiA мы использовали набор фрагментов дцДНК длиной 40 п.н. с точечными мутациями либо в core мотиве, либо в aux мотиве, либо в обоих мотивах сразу (Рисунок 3.21 А-В).

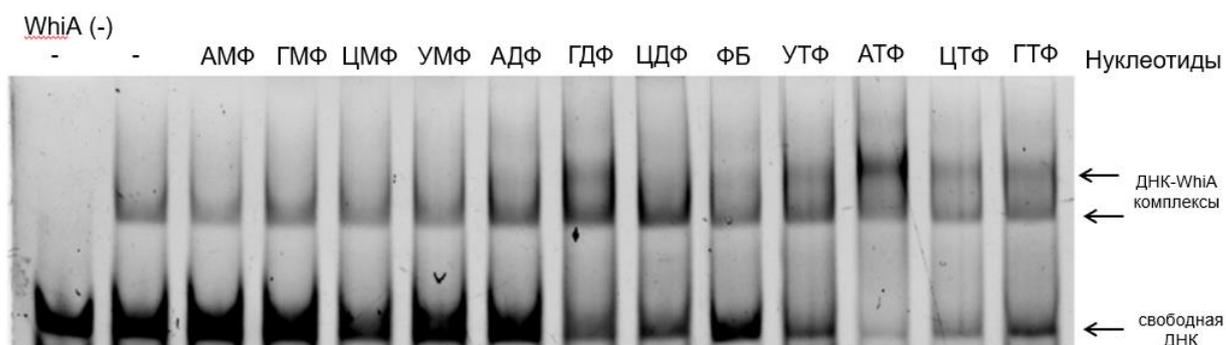


**Рисунок 3.21** А – лигатура (LOGO) для набора промоторных последовательностей оперона *rpsJ* бактерий класса Mollicutes. Б – оценка влияния мутаций в core и aux мотивах на связывание WhiA с дцДНК; каждый образец содержал 250 нМ дцДНК-субстрата и 600 нМ белка. В – последовательности ДНК, использованные в эксперименте; нуклеотиды, подвергшиеся замене, выделены красным.

В результате белок WhiA связывал все фрагменты дцДНК, но только с теми, где мутация не затронула core мотив, WhiA образовывал комплекс с большей электрофоретической подвижностью. Любая мутация в core мотиве мешала образованию этого ДНК-белкового комплекса. Было показано для WhiA *Streptomyces*, что с мотивом GACAC (core) связывается НТН-домен белка, в то же время с помощью биоинформатических методов предсказано сохранение ДНК-связывающей активности для НЕН домена, однако последовательность, с которой связывается НЕН домен, WhiA неизвестен [125].

### 3.5.3 АТФ влияет на связывание WhiA с ДНК-субстратом

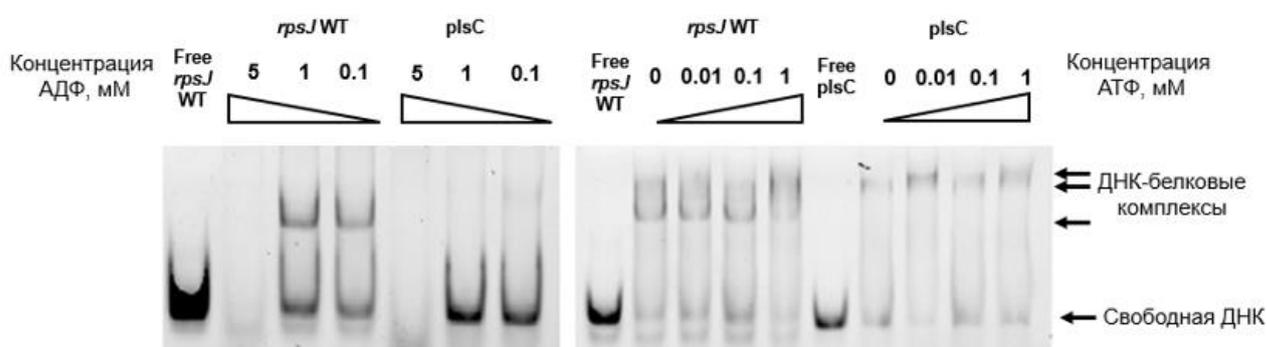
Для некоторых хоминг эндонуклеаз семейства LAGLIDAG показана их способность гидролизовать АТФ, а также показано влияние кофакторов (катионов металлов) на эту способность и на специфичность связывания и расщепления ДНК [136]. Помимо этого, ген, кодирующий аденилаткиназу (фермент, который превращает АТФ и АМФ в АДФ и наоборот), является консервативным компонентом оперона *rpsJ*. Принимая во внимание эти два факта, мы предположили, что HEN домен WhiA может иметь АТФ-азную, АТФ-связывающую или АДФ-связывающую функцию. Чтобы проверить, влияет ли наличие нуклеотидов на связывание WhiA с ДНК, был проведен EMSA эксперимент при добавлении различных нуклеозид -моно, -ди и -трифосфатов к образцам (Рисунок 3.22).



**Рисунок 3.22 Влияние нуклеотидов на ДНК-связывающие свойства WhiA.** Каждый образец содержал 250 нМ ДНК-субстрата *rpsJ*\_WT и 600 нМ белка WhiA (кроме дорожки, обозначенной WhiA (-), где белок отсутствовал). Концентрация нуклеотидов составляла 1 мМ. В качестве контроля использовали образец без добавления нуклеотидов, а также образец, содержащий фосфатный буфер ФБ.

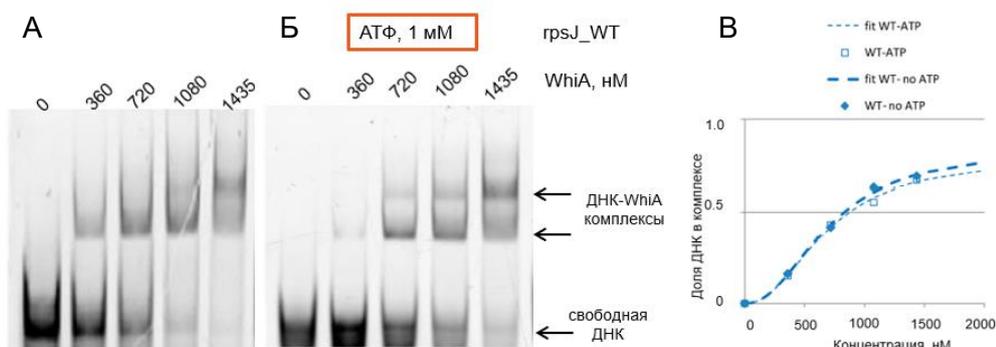
Добавление нуклеозид-монофосфатов, так же как и присутствие фосфатного буфера не влияло на способность WhiA взаимодействовать с ДНК. Присутствие нуклеозид-дифосфатов и трифосфатов, в особенности АТФ, влияло на подвижность ДНК-белкового комплекса – наряду с низкомолекулярным комплексом присутствовал комплекс с большей молекулярной массой. АТФ был выбран нами как наиболее вероятный лиганд WhiA. Далее мы варьировали концентрации АТФ и АДФ для анализа влияния на ДНК-связывающие свойства WhiA как со специфическим ДНК-субстратом *rpsJ*\_WT, так и с неспецифическим – *plsC*. Последовательность неспецифического фрагмента ДНК не содержала сайт связывания WhiA. АДФ был выбран как нуклеотид, который не влиял на связывание WhiA с ДНК, при этом его концентрация в клетке связана с концентрацией АТФ (Рисунок 3.22). Добавление АДФ не влияло на связывание WhiA с ДНК как со специфическим, так и с неспецифическим ДНК-субстратами, однако при высоких концентрациях АДФ белок агрегирует и не подвергается разделению в геле. При добавлении возрастающего количества АТФ, напротив, происходит перестройка ДНК-белкового комплекса, если в качестве субстрата выступал фрагмент дцДНК – *rpsJ*\_WT; при концентрации АТФ 1 мМ доля комплекса с меньшей электрофоретической

подвижностью была максимальна (Рисунок 3.23). Добавление различных концентраций АТФ не оказывало существенного влияния на связывание WhiA с неспецифическим субстратом – *plsC*.



**Рисунок 3.23** Влияние АТФ и АДФ на способность WhiA связывать ДНК. Анализ связывания 600 нм WhiA с 250 нм *rpsJ\_WT* – ДНК-субстрат, соответствующий промоторной области *rpsJ*. *plsC* – ДНК-субстрат, не содержащий сайты связывания WhiA при добавлении возрастающей концентрации АТФ и АДФ.

Далее мы рассчитали константы диссоциации комплекса WhiA с *rpsJ\_WT* в двух условиях – в отсутствие АТФ и при наличии 1 мМ АТФ в образцах. Такая концентрация АТФ близка к физиологической для бактерий [137]. Для этого были использованы данные о доли ДНК-субстрата в комплексе с WhiA, при повышающейся концентрации белка (Рисунок 3.24 А-В).

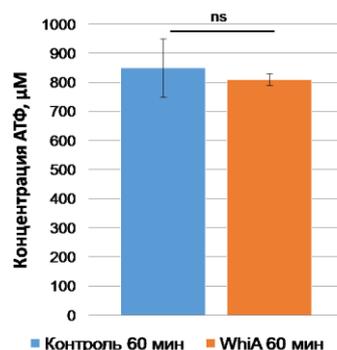


**Рисунок 3.24** А – Анализ связывания 250 нм ДНК-субстрата *rpsJ\_WT* при возрастающей концентрации WhiA (концентрация белка указана в верхней части рисунка) без добавления АТФ. Б – в присутствии 1 мМ АТФ. В – Кривые связывания возрастающей концентрации WhiA с 250 нм ДНК-субстрата *rpsJ\_WT*, рассчитанные из данных эксперимента по торможению ДНК в 8% ПААГ в нативных условиях.

Полученные в обоих случаях константы диссоциации комплекса WhiA с *rpsJ\_WT* оказались одинаковы; для комплекса WhiA-*rpsJ\_WT* в условиях отсутствия АТФ  $K_d$  составила  $700 \pm 80$  нМ, коэффициент Хилла  $1,8 \pm 0,2$ , что свидетельствует о положительном кооперативном эффекте. При определении констант не учитывалось наличие двух видов ДНК-белковых комплексов. То есть АТФ не влияет на общую константу диссоциации

специфического дцДНК-WhiA комплекса, но влияет на соотношение между высокомолекулярным и низкомолекулярным комплексами.

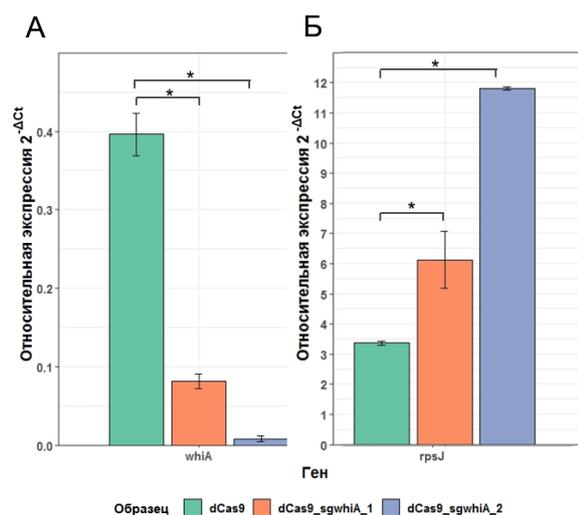
Для того чтобы проверить, способен ли WhiA гидролизовать АТФ, мы провели измерение уровня АТФ методом биолюминесцентной АТФ-метрии в образцах, содержащих белок WhiA и ДНК-субстрат *rpsJ*\_WT. В качестве контрольного образца использовали пробу, содержащую ДНК-субстрат, но без добавления WhiA. Белок WhiA не обладает детектируемой способностью гидролизовать АТФ (Рисунок 3.25).



**Рисунок 3.25. Оценка АТФ-азной активности WhiA.** Анализ уровня АТФ в контрольном образце и образце, содержащем белок WhiA, измеренный методом биолюминесцентной АТФ-метрии. Измерение АТФ проводили после Инкубации образцов в течение 60 мин при 37°C.

### 3.5.4 Репрессия транскрипции *whiA* приводит к увеличению экспрессии *rpsJ*

Для демонстрации влияния WhiA на экспрессию *rpsJ in vivo* мы подавляли экспрессию гена *whiA* у *M. gallisepticum* S6 с помощью системы CRISPR интерференции. Среди двух сконструированных молекул sgРНК наибольший эффект оказала вторая – sgwhiA2, нацеленная на 140–160 позиции гена *whiA* относительно первого нуклеотида открытой рамки считывания. В трансформантах *M. gallisepticum*, экспрессирующих dCas9 и sgwhiA2, удалось снизить уровень экспрессии *whiA* в среднем в 48 раз (Рисунок 3.26 А) по сравнению с экспрессией этого гена в штамме дикого типа. При этом экспрессия гена *rpsJ* возросла в 3,6 раза (Рисунок 3.23 Б). Падение уровень экспрессии *whiA* в трансформантах dCas9\_sgwhiA1 в пять раз привело к росту уровня *rpsJ* в два раза.



**Рисунок 3.26 Анализ влияния репрессии *whiA* на экспрессию *rpsJ*** А, Б – Уровень экспрессии транскрипционного фактора *whiA* и его мишени – *rpsJ* в трансформантах dCas9, dCas9\_sgwhiA\_1 и dCas9\_sgwhiA\_2 относительно представленности этих транскриптов в штамме микоплазмы дикого типа. Измерение проводили с помощью количественной ОТ-ПЦР в реальном времени. Уровни экспрессии генов внутри каждого образца нормализовали на уровень мРНК гена «домашнего хозяйства» *tuf*. Приведено среднее не менее 3 независимых экспериментов со стандартным отклонением. \* –  $p < 0,05$  согласно ранговому непарному критерию Вилкоксона.

Таким образом, среди 10 найденных потенциальных факторов регуляции транскрипции *M. gallisepticum* S6, нам удалось подробно охарактеризовать 4 из них, представив экспериментальные доказательства их мишеней. Консервативные ТФ-ы, такие как MraZ и WhiA, регулируют гены, которые поддерживают основные клеточные функции – деление и синтез белка. Приобретенная штаммом S6 система рестрикции–модификации, содержащая С-белок (Xre), вероятно, усложняет регуляторную сеть этого штамма вследствие наличия в геноме микоплазмы дополнительных мишеней, которые могут распознаваться этим ТФ-ом. Белок Fug в нормальных условиях репрессирует ген *zip*, кодирующий транспортёр металлов. При сохранении и консервативности гена ТФ-а эта мишень не утрачена у близкородственных видов – *M. genitalium* и *M. pneumoniae*.

### 3.6 Поиск и валидации новых генов – мишеней регуляторов

Несколько лет работы над изучением транскрипции у представителей класса Mollicutes в нашей лаборатории показали, что у микоплазм структура промотора упрощена, по сравнению с промоторами других представителей класса, такими как ахлеплазмы и спироплазмы. В структуре промотора микоплазм можно выделить единственный консервативный элемент -10 бокс. Такое упрощение внутри класса происходит постепенно, коррелируя с уменьшением размера генома и количеством известных факторов регуляции транскрипции [96], [82]. Несмотря на это, при воздействии стрессов на *M. gallisepticum* изменяется экспрессия до половины генов, вместе с тем микоплазмы проявляют удивительные

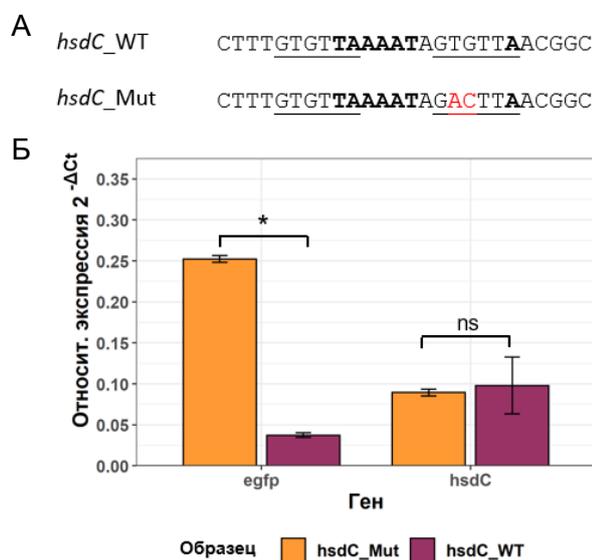
адаптационные способности [82], [70]. Таким образом, мы считаем, что существуют новые неизвестные механизмы контроля экспрессии генов у микоплазм, которые могут быть реализованы с помощью белков, РНК, топологии ДНК и т.п. В нашей лаборатории для *M. gallisepticum* была построена модель, которая по последовательности промотора предсказывает его «силу» (неопубликованные данные). Эта модель построена с помощью методов машинного обучения (алгоритм random forest), было использовано несколько характеристик промотора, оценивающих соответствие оптимальной последовательности оптимальным расстояниям между различными детерминантами промотора и нуклеотидный состав спейсеров между детерминантами. Соответствие оптимальным последовательностям детерминант промотора выразилось через вес выравнивания последовательности с позиционной весовой матрицей оптимального мотива. Модель для каждого промотора *M. gallisepticum* сообщает предсказанную «силу». Сравнивая теоретическую (предсказанную моделью) и экспериментально измеренную «силу» промотора (данные 5'-концевого секвенирования РНК), можно определить набор промоторов (генов), которые находятся под регуляцией (т.е. активированы или репрессированы). Для промоторов, которые находятся под репрессией, предсказанная «сила» будет больше, чем экспериментально измеренная, для активированных – наоборот. Существенный плюс этого подхода – отсутствие априорных знаний о том, как именно реализован механизм регуляции. Применяя этот подход, было предсказано, что 40% генов *M. gallisepticum* S6 находятся под действием регуляторов, из них экспрессия двух третей подавлена репрессорными механизмами (неопубликованные данные).

Для того чтобы валидировать модель и подтвердить ее результаты, нами был создан вектор на основе транспозона (п. 2.19), который позволяет вносить желаемые генетические элементы (участки промотора, потенциальный сайт связывания ТФ) в промоторную область гена *egfp* и затем с помощью ПЦР в реальном времени оценить влияние элемента на экспрессию гена *egfp*. Промоторные области интересующих нас генов были разбиты на перекрывающиеся участки. Серию фрагментов, содержащих необходимые и достаточные элементы для инициации транскрипции, мы назвали mini. Дополнительные участки, как правило, места посадки транскрипционных факторов, либо последовательности рибопереключателей мы обозначили цифрами. Помимо экспрессии *egfp*, мы также отслеживали экспрессию генов «домашнего хозяйства» и генов интереса.

### **3.6.1 Тестирование системы для известного фактора регуляции транскрипции и его гена-мишени**

Для проверки нашей системы мы использовали промоторную область оперона *hsd* *M. gallisepticum* S6, описанного в п. 3.3. Мы создали две конструкции с геном *egfp*. В первую

мы внесли фрагмент длиной 23 п.н., соответствующий природной последовательности промоторной области оперона *hsd* и содержащий -10 бокс промотора, а также два повтора GTGTTA – сайт связывания транскрипционного фактора HsdC (белок C). Такие трансформанты *M. gallisepticum* мы обозначили (*hsdC\_WT*) (Рисунок 3.24). Вторая конструкция содержала тот же фрагмент, длиной 23 п.н., но с мутацией во втором повторе двух наиболее консервативных нуклеотидов сайта связывания HsdC (TG заменили на AC). Такие трансформанты *M. gallisepticum* мы назвали (*hsdC\_Mut*) (Рисунок 3.27). Экспрессия гена *egfp*, чья промоторная область содержала мутированный сайт связывания регулятора HsdC, была достоверно выше в 6,8 раз, по сравнению с экспрессией этого гена в трансформантах *hsdC\_WT* (Рисунок 3.27). Экспрессия самого гена *hsdC* у трансформантов не отличалась.



**Рисунок 3.27 Проверка работы системы для оценки влияния вносимых детерминант промотора на экспрессию репортерного гена *egfp*.** **А** – *hsdC\_WT* – природная последовательность промоторной области оперона *hsd* *M. gallisepticum* S6; *hsdC\_Mut* – последовательность, содержащая мутации в сайте связывания ТФ HsdC. -10 бокс промотора и инициаторный нуклеотид выделены жирным шрифтом; сайт связывания HsdC – подчеркнут; мутированные нуклеотиды выделены красным. **Б** – Уровни относительной экспрессии генов *egfp* и *hsdC* в трансформантах *hsdC\_WT* и *hsdC\_Mut* *M. gallisepticum* S6. Измерение проводили с помощью количественной ОТ-ПЦР в реальном времени. Уровни экспрессии генов внутри каждого образца нормализовали на уровень мРНК гена «домашнего хозяйства» *tuf*. Приведено среднее не менее 3 независимых экспериментов со стандартным отклонением. \* –  $p < 0,05$  согласно ранговому непарному критерию Вилкоксона.

То есть в трансформантах *hsdC\_Mut* транскрипционный фактор HsdC не может связаться с промоторной областью гена *egfp* и репрессировать его транскрипцию. Данный факт указывает на то, что предложенная нами система работает.

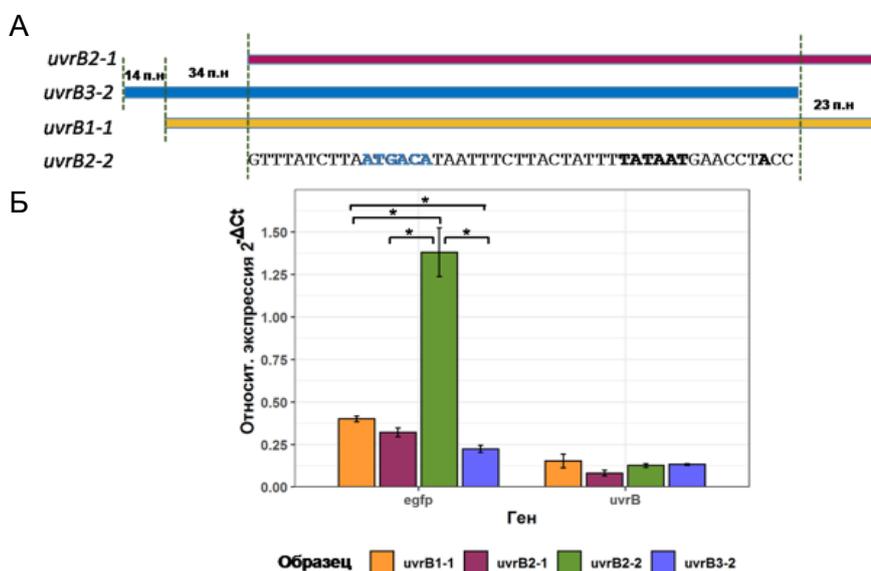
### 3.6.2 Влияние разных элементов промотора на экспрессию репортерного гена *egfp* у *Mycoplasma gallisepticum*



генов бактерий с редуцированным геномом в силу снижения сложности структуры промотора это влияние отсутствует или уменьшено. Редукция генома и смещение состава генома в сторону преобладания АТ-пар привели к упрощению структуры промотора не только у микоплазм, но также у симбионта тли – *Buchnera aphidicola* [138] и *Helicobacter pylori* [139]. Снижение сложности структуры промотора связывают с отсутствием альтернативных сигма-факторов.

### 3.6.3 Примеры репрессии

В качестве примера гена, находящегося под репрессией у *Mycoplasma gallisepticum* S6, мы приводим ген *uvrB*, который кодирует белок UvrB, один из ключевых компонентов эксцизионной репарации нуклеотидов [140]. Нами были созданы четыре конструкции, содержащие фрагменты промоторной области гена *uvrB* *M. gallisepticum*. Все четыре фрагмента перекрывались и имели общую область длиной 47 п.н., где находится -10, -35 боксы промотора (Рисунок 3.29). Анализ экспрессии гена *egfp* показал, что у трансформантов *M. gallisepticum uvrB2-2*, уровень *egfp* был максимален. Добавление фрагментов с до и/или с после относительно внесенной последовательности *uvrB2-2*, снижало экспрессию репортерного гена (Рисунок 3.29).



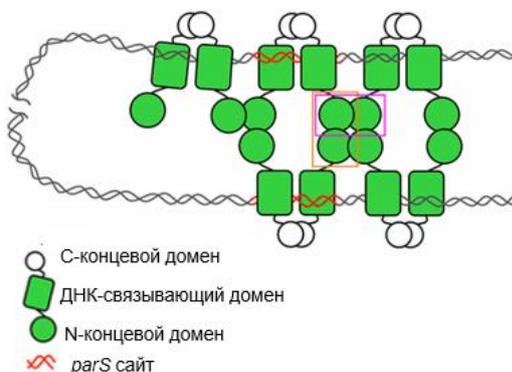
**Рисунок 3.29** Анализ влияния последовательностей промоторной области гена *uvrB* на экспрессию репортерного гена *egfp*. **А** – Последовательности промоторной области гена *uvrB*, внесенные в конструкцию. -10 бокс промотора и инициаторный нуклеотид выделены жирным шрифтом черного цвета; последовательность -35 бокса выделена жирным шрифтом синего цвета. **Б** – Уровни экспрессии генов *egfp* и *uvrB* в трансформантах *M. gallisepticum* S6. Измерение проводили с помощью количественной ОТ-ПЦР в реальном времени. Уровни экспрессии целевых генов внутри каждого образца нормализовали на уровень мРНК гена «домашнего хозяйства» *tuf*. Приведено среднее не менее 3 независимых экспериментов со стандартным отклонением. \* –  $p < 0,05$  согласно ранговому непарному критерию Вилкоксона.

Полученные данные указывают на то, что экспрессия гена *uvrB* *M. gallisepticum* находится под влиянием репрессора, причем механизм репрессии связан с влиянием последовательностей, находящихся как до промотора, так и после него.

Репарация ДНК имеет важное значения для поддержания стабильности генома у всех живых организмов, поэтому не удивительно, что этот процесс регулируется. Например, у *E. coli* экспрессия *uvrB* находится под регуляцией двух репрессоров – *lexA* и *dnaA* [141]. У *M. gallisepticum* отсутствует глобальный регулятор *lexA*, промоторная область гена *uvrB* не содержит *dnaA*-боксов.

Другой пример репрессии у *M. gallisepticum* – ген *parA*, расположенный в одном опероне с геном с неизвестной функцией и геном *dnaA* – инициатором репликации ДНК. Белок ParA у многих бактерий входит в систему ParABS, которая участвует в сегрегации геномной, а также плазмидной ДНК во время клеточного деления [142]. *parS* – это нуклеотидная последовательность вида TGTTCACGTGAAACA (некоторые ее вариации допустимы и описаны у разных видов бактерий), как правило, такие сайты (от 1 до 20) локализованы вблизи точки начала репликации (*ori*), а также в зависимости от вида на расстоянии от 10 т.п.н у *Caulobacter crescentus* до 200 т.п.н у *Streptomyces coelicolor*, или даже до 650 т.п.н у *Pseudomonas aeruginosa* [143], [144]. Изученные белки ParB разных организмов не консервативны по аминокислотной последовательности [145], [144], но имеют общие структурные особенности: N-концевой домен обеспечивает белок-белковые взаимодействия с ParA; центральный домен содержит мотив спираль-поворот-спираль или лента-спираль-спираль, ответственный за связывание с *parS*; и C-концевой домен обеспечивает димеризацию ParB и неспецифическое связывание с ДНК [146]. Белок ParB способен специфически связываться с сайтами *parS* и распространяться вдоль ДНК за счет неспецифического взаимодействия с ДНК, впоследствии образуя комплексы высшего порядка. Было показано, что для сборки комплексов ParB высшего порядка необходимо взаимодействие белка с отдаленными *parS* сайтами, а за счет димеризации богатого остатками аргинина ParB бокса II образуются белок-белковые «мостики» (Рисунок 3.30) [147]. Недавно было показано, что N-концевой домен ParB способен связывать и гидролизовать цитидинтрифосфат, распространение ParB происходит в присутствии ЦТФ [148]. Белок ParA – АТФ-аза, движущая сила сегрегации хромосом. Современные представления о работе системы ParABS таковы: димер белка ParA, несущий АТФ, неспецифически связывается с ДНК. ParB имеет сродство к ParA, ParB-*parS* комплекс взаимодействует с димерами ParA, стимулируя его АТФ-зную активность. Гидролиз АТФ вызывает диссоциацию димера и высвобождение мономеров ParA и ДНК. В непосредственной близости от комплекса ParB создается область, истощенная по

димерам ParA. Из-за высокого сродства ParB-parS к димерам ParA, ParB-parS движется в сторону более высокой концентрации ParA. Таким образом, ParA формирует градиент концентрации с максимумом на полюсе будущей новой клетки и минимумом в районе начала репликации (*ori*), направляя движение ParB-parS комплекса [149], [150].

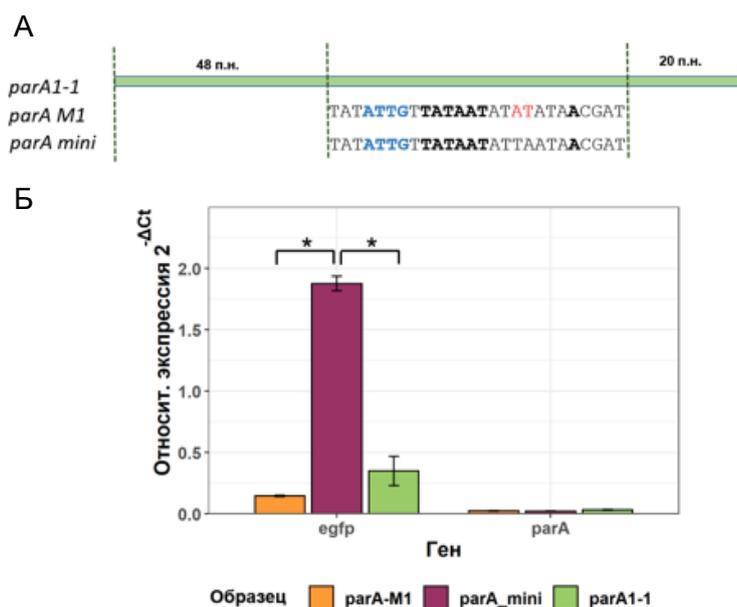


**Рисунок 3.30. Модель действия ParB *Bacillus subtilis*.** Белый круг – С-концевой домен димеризации ParB, зеленый прямоугольник – ДНК-связывающий домен, и зеленый круг – гибкие N-концевые домены ParB. Димер ParB образуется через С-концевой домен димеризации. ParB связывает хромосомную ДНК в определенном сайте *parS* (красный) или в неспецифических сайтах (серый). Множественные молекулы ParB распространяются вдоль хромосомной ДНК через N-концевой домен с помощью смежных взаимодействий (обведены красным цветом), образуется комплекс ParB-ДНК высшего порядка. Адаптировано из [147].

У многих бактерий процесс сегрегации поддерживается полярными белками-ориентирами, таким как PopZ у *Caulobacter crescentus*, которые изолируют комплекс ParB-parS на полюсах клетки [151]. Некоторые бактерии (например,  $\gamma$ -протеобактерии, включая *E. coli*) не имеют систему ParABS, у этих бактерий сегрегацию хромосом обеспечивают другие белки – конденсины – SMC [152].

У микоплазм не описано системы ParABS [144], [153], однако у родов *Acholeplasma* и *Spiroplasma* класса Mollicutes и некоторых микоплазм – *Mycoplasma gallisepticum*, *M. imitans*, *M. testudinis*, *M. pirum*, *M. amphoriforme*, *M. alvi*, *M. pneumoniae*, *M. genitalium*, *M. iowae*, *M. penetrans*, *M. phocicerebrale*, *M. conjunctivae* в геноме закодирован гомолог хорошо изученного белка ParA (Spo0JA) *Bacillus subtilis*, но нет ParB (Spo0JB).

Используя наш вектор, мы показали, что, находясь под регуляцией только промотора гена *parA*, состоящего из консенсусной последовательности -10 бокса и extended элемента (*parA\_mini*), экспрессия *egfp* в 72 раза больше, чем экспрессия природного гена *parA* *M. gallisepticum*. Добавляя 20 п.н. в область после промотора и в 48 п.н. в область до промотора т.е. относительно тестируемого участка последовательности *parA\_mini*, удалось снизить экспрессию репортерного гена в 5,4 раза. Уровень гена *parA* не изменился (Рисунок 3.31).



**Рисунок 3.31 Анализ влияния последовательностей промоторной области гена *parA* на экспрессию репортерного гена *egfp*.** А – Последовательности промоторной области гена *parA*, внесенные в конструкцию; -10 бокс промотора и инициаторный нуклеотид выделены жирным шрифтом черного цвета; последовательность EХТ выделена жирным шрифтом синего цвета; мутации, внесенные в природную последовательность, выделены красным. Б – Уровень экспрессии генов *egfp* и *parA* в трансформантах *M. gallisepticum* S6. Измерение проводили с помощью количественной ОТ-ПЦР в реальном времени. Уровни экспрессии целевых генов внутри каждого образца нормализовали на уровень мРНК гена «домашнего хозяйства» *tuf*. Приведено среднее не менее 3 независимых экспериментов со стандартным отклонением. \* –  $p < 0,05$  согласно ранговому непарному критерию Вилкоксона.

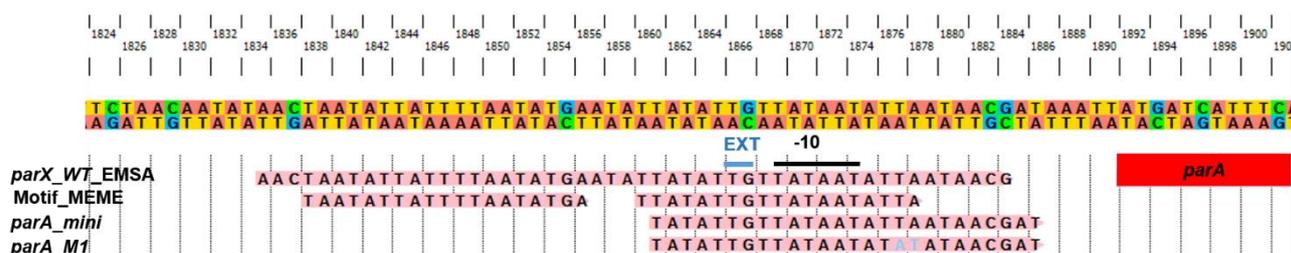
Изучая промоторную область гена *parA* у разных микоплазм, мы обнаружили консервативные области. С помощью программы MEME у 10 из 13 видов микоплазм, у которых присутствует ген *parA*, был найден мотив – два несовершенных бипалиндрома (см. Рисунок 3.32).



**Рисунок 3.32 Поиск консервативного мотива в промоторных областях генов *parA* микоплазм.** Лигатура (LOGO) найденного мотива среди набора промоторных последовательностей гена *parA* 13 видов микоплазм, у которых обнаружен гена *parA*, гомологичный хорошо изученному белку ParA (Spo0JA) *Bacillus subtilis*. Для поиска мотивов программе MEME подавались области длиной 100 нуклеотидов (координаты от -100 до 0 н. считая от старт-кодона *parA*).

Поиск такого мотива в геноме *M. gallisepticum* с помощью программы FIMO выявил 210 мотивов с  $p$ -value меньше 0,0001. Из них 29 расположены в районе *ori* –в пределах 1900 п.н.. Возможно, это связано с тем, что район начала репликации АТ-богатый, так же как и найденный мотив. В промоторной области гена *parA* *M. gallisepticum* было найдено два

мотива. Они не идентичны, отличаются несколькими нуклеотидными заменами, один из мотивов включает в себя -10 бокс промотора *parA* (Рисунок 3.33).

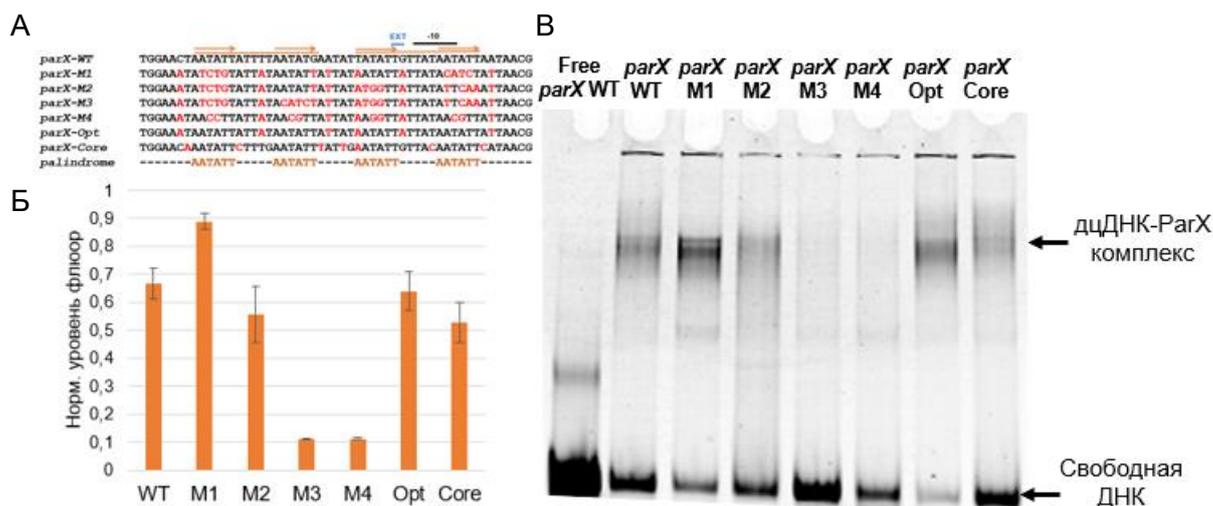


**Рисунок 3.33** Последовательность промоторной области гена *parA* *M. gallisepticum* S6. Отмечены элементы промотора, мотивы, найденные программой MEME, а также последовательности, используемые в экспериментах по оценке экспрессии *egfp* и EMSA (описаны ниже). Белок-кодирующая часть гена *parA* отмечена красным прямоугольником.

Г.Ю. Фисуновым совместно с О.В. Побегуц (лаб. протеомного анализа ФГБУ ФНКЦ ФХМ ФМБА России) был выполнен эксперимент по обогащению фракции растворимых белков *M. gallisepticum* белками, способными связываться с промоторной областью гена *parA*. Эта область содержала два найденных мотива (Рисунок 3.33). С помощью аффинной хроматографии белков лизата *M. gallisepticum*, где в качестве субстрата был использован фрагмент дцДНК промоторной области гена *parA*, меченный биотином и иммобилизованный на магнитных шариках со стрептавидином, были получены белки, связывающиеся с данной последовательностью. Среди самых обогащенных в связанной с ДНК фракцией белков, относительно представленности белков тотального лизата, оказался белок с неизвестной функцией. В геноме *M. gallisepticum* ген этого белка располагается после гена *parA* в одном с ним опероне, что характерно для гена *parB*. Мы назвали этот ген с неизвестной функцией *parX*. У *M. gallisepticum* оперон, содержащий ген *parA*, находится рядом с точкой начала репликации (*ori*), однако вблизи гена *parA* не было найдено *parS* сайтов, схожих с *parS* сайтами, обнаруженными у других бактериях. Сайт TGTTCACGTGAAACA, схожий с *parS* сайтом *B. subtilis*, найден в геноме *M. gallisepticum* на расстоянии 0,53 от размера генома микоплазмы, выраженного в п.н., относительно *ori* (то есть, если заключить геном микоплазмы в ровное кольцо, сайт TGTTCACGTGAAACA будет находиться фактически напротив *ori*). Интересно, что этот сайт располагается в теле гена *scpA/scpB*, чей белковый продукт имеет два домена: один участвует в сегрегации и конденсации ДНК, вместе с *ScpB* и *Smc* формирует конденсин-подобный комплекс, а другой имеет дигидрофолатредуктазную активность.

Используя метод сдвига электрофоретической подвижности (EMSA), мы проверили способность *ParX* связываться с промоторной областью собственного оперона, а также оценили влияние мутаций в этой последовательности на аффинность белка к ДНК (Рисунок

3.34). Длина тестируемого участка составила 53 п.н., найденные мотивы входили в этот фрагмент (Рисунок 3.33). Фрагмент *parX\_M1* содержал такие мутации, чтобы в каждом из повторов осталось только по одному совершенному палиндрому – ААТАТТ, расположенные рядом; *parX\_M2* – один ААТАТТ в первом повторе; *parX\_M3* и *parX\_M4* – ни одного ААТАТТ, но разное количество замен; *parX\_Opt* и *parX\_Core* – четыре ААТАТТ, но разный нуклеотидный контекст.



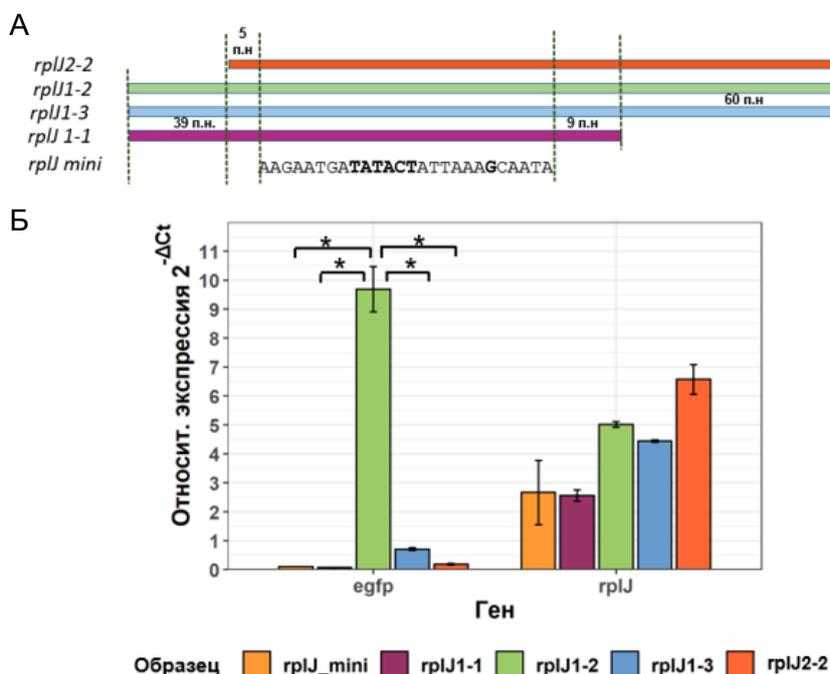
**Рисунок 3.34** Оценка способности ParX связывать промоторную область *parA*. **А** – Последовательности фрагментов дцДНК, используемые в эксперименте; *parX\_WT*– фрагмент промоторной области гена *parA* дикого типа; M1-M4 – варианты последовательности, содержащие нуклеотидные замены в различных позициях повторов. **Б, В** – Анализ связывания 100нМ различных фрагментов дцДНК 600 нМ белка ParX методом торможения ДНК в 8% ПААГ.

ParX эффективно связывал фрагменты дцДНК *parX\_M1* и *parX\_Opt*, то есть участки с близко расположенными совершенными палиндромными последовательностями в определенном контексте. С помощью нашего вектора мы проверили влияние мутации в одном из палиндромных участков обнаруженного нами мотива на экспрессию репортерного гена *egfp* (Рисунок 3.30) – последовательность *parA\_M1*. Нарушение палиндрама с сохранением содержания количества АТ оснований во вносимом в вектор участке снижало экспрессию гена *egfp* в 12,8 раз, по сравнению с экспрессией этого гена, где вносимый участок содержал «природную» последовательность – *parA\_mini* (Рисунок 3.33). Вероятно, ParX является репрессором оперона *parA-parX-dnaA*, что не исключает его роли в сегрегации ДНК, аналогично белку ParV других бактерий.

### 3.6.4 Пример активации

Ген рибосомного белка L10 большой субъединицы рибосомы (*rplJ*) – один из генов, чья экспрессия находится под действием активатора. Сравнивая экспрессию гена *egfp* у

трансформантов *M. gallisepticum*, содержащих разные варианты промоторных областей гена *rplJ* перед репортерным геном, мы показали, что разница в экспрессии *egfp* может достигать 95,5 раз. Такая разница была между трансформантами *rplJ\_mini* и *rplJ1-2* (Рисунок 3.35).



**Рисунок 3.35** Анализ влияния последовательностей промоторной области гена *rplJ* на экспрессию репортерного гена *egfp*. А – Последовательности промоторной области гена *rplJ*, внесенные в конструкцию; -10 бокс промотора и инициаторный нуклеотид выделены жирным шрифтом черного цвета; последовательность ЕХТ выделена жирным шрифтом синего цвета; мутации, внесенные в природную последовательность, выделены красным. Б – Уровень экспрессии генов *egfp* и *rplJ* в трансформантах *M. gallisepticum* S6. Измерение проводили с помощью количественной ОТ-ПЦР в реальном времени. Уровни экспрессии целевых генов внутри каждого образца нормализовали на уровень мРНК гена «домашнего хозяйства» *tuf*. Приведено среднее не менее 3 независимых экспериментов со стандартным отклонением. \* –  $p < 0,05$  согласно ранговому непарному критерию Вилкоксона.

Регуляция транскрипции *rplJ* организована сложно. При добавлении длинных последовательностей только до или только после промотора не ведет к значимому изменению экспрессии репортерного гена (трансформанты *rplJ1-1* и *rplJ2-2*). Но, если расширять эти области в обе стороны, наблюдается увеличение экспрессии гена *egfp* (для трансформанта *rplJ1-3* в 7 раз, для трансформанта *rplJ1-2* в 95,5 раз относительно *rplJ\_mini*). Возможно, участки до и после промотора гена *rplJ* содержат сайты связывания активатора, который действует как энхансеры у эукариот.

У бактерий развит механизм регуляции экспрессии генов за счет образования вторичных структур РНК - РНК-переключателей (аттенуаторов) – цис- регуляторных элементов РНК. Например, у *E. coli* для многих оперонов рибосомных генов, описана авторегуляция, как на уровне транскрипции, так и на уровне трансляции [154]. Дальнейшие исследования будут

направлены на выяснение точного механизма регуляции экспрессии у описанных выше генов – *uvrB* и оперона рибосомных генов, содержащего *rplJ* у микоплазм.

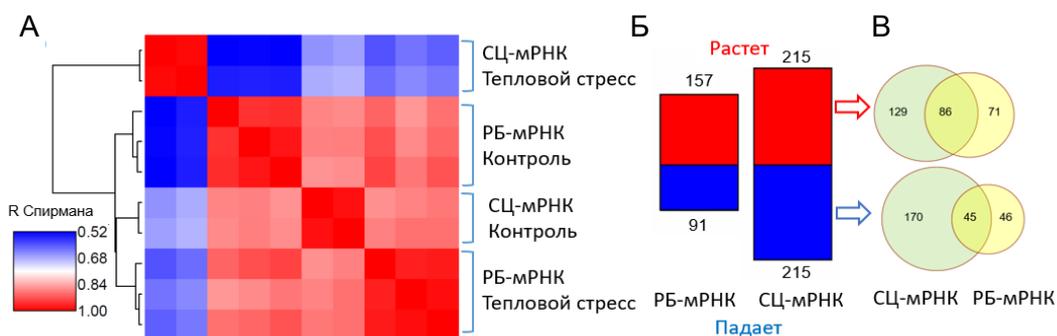
### **3.7 Анализ фракции мРНК, связанной с рибосомами**

Регуляция передачи генетической информации от ДНК к РНК – транскрипция, у *Mycoplasma gallisepticum* не может быть объяснена только работой известных транскрипционных факторов, вероятно, есть другие механизмы. Мы не знаем пока, какие именно, но знаем их мишени – около 40% генов *M. gallisepticum*. В живых организмах основной функциональный ответ осуществляют белки. Многие работы, направленные на изучение взаимосвязи уровня мРНК и белка в клетке как эукариот так и прокариот, указывают на расхождение между этими уровнями [155], [156], [157], [158]. Некоторые исследователи объясняют такое расхождение специфичной для каждой мРНК и белка скоростью синтеза и деградации [159], [160], а некоторые – ошибками применяемых методов, то есть измеряемая представленность мРНК и белка искажена за счет смещений при пробоподготовке образцов и за счет методов анализа данных [161].

Ранее в нашей лаборатории было показано, что при тепловом стрессе ответ на уровне транскрипции колоссальный, что не отражается на представленности соответствующих белков [82]. Используя модель теплового стресса и профилирование рибосом, мы решили выяснить механизмы регуляции передачи генетической информации от мРНК к белку.

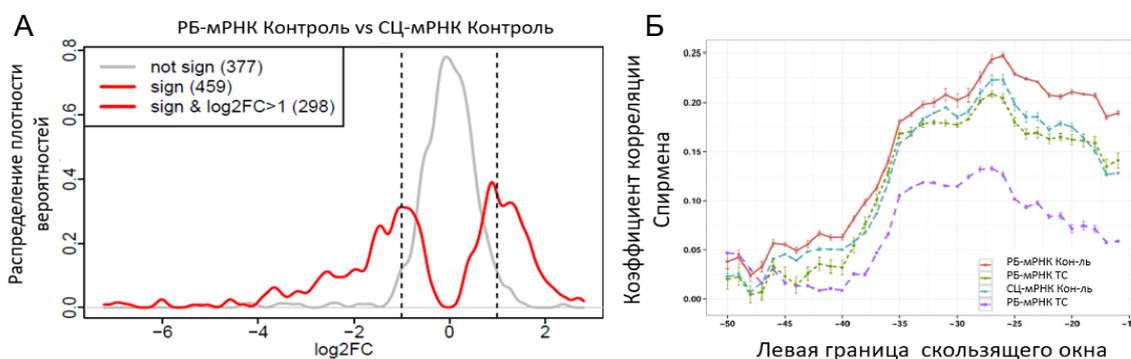
#### **3.7.1 Уровень транскрипции пропорционален уровню трансляции в нормальных условиях**

Представленность всех транскриптов, выраженная в нормализованных значениях RPKM (reads per kilobase per million of mapped reads — число прочтений на тысячу нуклеотидов на миллион картированных прочтений) была получена для образцов суммарной цитоплазматической фракции мРНК (СЦ) и связанной с рибосомами фракции мРНК (РБ) *M. gallisepticum* в двух состояниях – логарифмический рост и после теплового стресса (п. 2.21). При экспоненциальном росте содержание мРНК, связанной с рибосомами, имело высокую корреляцию с общим пулом мРНК (коэффициент корреляции по Спирмену 0,87) (Рисунок 3.36).



**Рисунок 3.36** Связь представленности транскриптов в анализируемых образцах. **А** – Тепловая карта значений коэффициентов ранговой корреляции по Спирмену рассчитанных для значений RPKM всех транскриптов во всех образцах транскриптома (СЦ– суммарная цитоплазматическая фракция мРНК) и транслятома (РБ – рибосомально-связанная фракция мРНК). **Б** – Количество дифференциально экспрессирующийся генов ( $|\text{Log}_2\text{FC}| > 2$ ,  $p\text{-value}_{\text{adj}}$ ) при тепловом стресс в двух фракциях – РБ и СЦ. **В** – Количество общих и уникальных дифференциально экспрессирующийся генов при тепловом стрессе в РБ и СЦ фракциях, представлены в виде диаграмм Венна [162].

Затем мы сравнили представленность мРНК в РБ и СЦ фракциях в нормальных условиях. При нормальных условиях, транскрипты, чья представленность в РБ была меньше, чем в СЦ фракции, имели широкий диапазон изменений ( $-1 > \text{Log}_2\text{FC} > -7.2$ ). Для тех мРНК, чей уровень в РБ-фракции возрос, изменения не превышали 7 раз ( $1 > \text{Log}_2\text{FC} > 2.8$ ) (Рисунок 3.37 А)



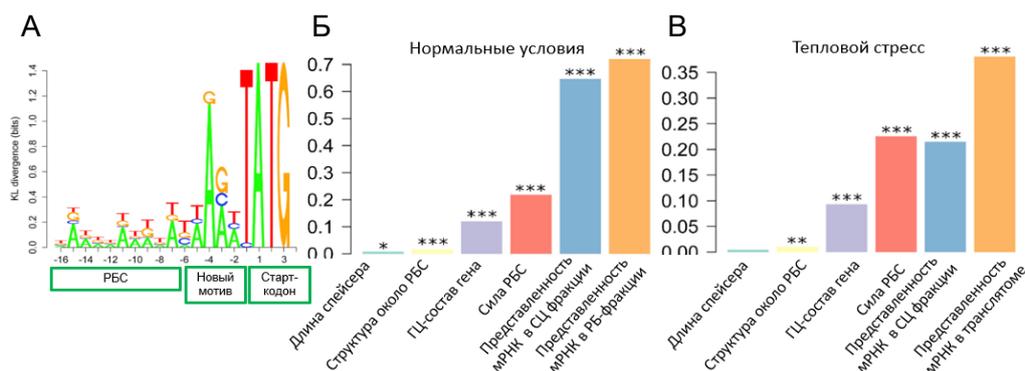
**Рисунок 3.37** Анализ распределения и свойств мРНК в РБ и СЦ фракциях в нормальных условиях. **А** – Распределение диапазона изменений уровня мРНК между РБ и СЦ фракций. Красным показано распределение для статистически значимых измерений (sign),  $p\text{-value}_{\text{adj}} < 0,05$ ; серым – для не значимых (not sign). На панели сверху указано количество генов, чьи транскрипты были отнесены к описанным группам. **Б** – Коэффициент корреляции Спирмена между значением представленности мРНК и значением dG структур в 5'-нетранслируемой области мРНК. Корреляция была рассчитана для значение dG в скользящем окне 30 п.н. в области от -50 до +14 относительно старт-кодона; ТС – тепловой стресс, Кон-ль – контроль, нормальные условия. Опубликовано в [162]

То есть в нормальных условиях всё же есть некая избирательность рибосом по отношению к мРНК. Мы предположили, что наличие вторичных структур в 5'-нетранслируемой области может оказывать влияние на эффективность связывания рибосомы с мРНК. Чтобы это выяснить, с помощью программы quickfold мы рассчитали свободную энергию  $-dG$ - образования вторичной структуры в скользящем окне 30 п.н. вблизи старт-кодона (область от -50 до +14 п.н.). Мы продемонстрировали, что корреляция содержания

мРНК во фракции, связанной с рибосомой, со стабильностью вторичной структуры имеет пик в области от -26 до +4 вокруг стартового кодона (Рисунок 3.37 Б). Чем менее стабильна вторичная структура вокруг стартового кодона, тем такой транскрипт более представлен. Для СЦ фракции мРНК мы наблюдали сходную тенденцию с несколько меньшей корреляцией (0,25 и 0,22 соответственно). Такой подход – попытка выявить общие детерминанты. Также мы использовали предыдущие данные [82], чтобы оценить силу сайта связывания рибосомы (РБС), выраженную в значении свободной энергии образования дуплекса между 16S рРНК и 5'-областью мРНК. Корреляция между силой РБС и содержанием мРНК в суммарном пуле была 0,20, для мРНК в РБ фракции немного выше – 0,23. Среди генов, чьи мРНК обогащены в РБ фракции в нормальных условиях, мы выявили дополнительную детерминанту связывания рибосомы, расположенную внутри спейсера между РБС и старт-кодоном (Рисунок 3.38 А).

### 3.7.2 Модель предсказания обогащения мРНК во фракции, связанной с рибосомами

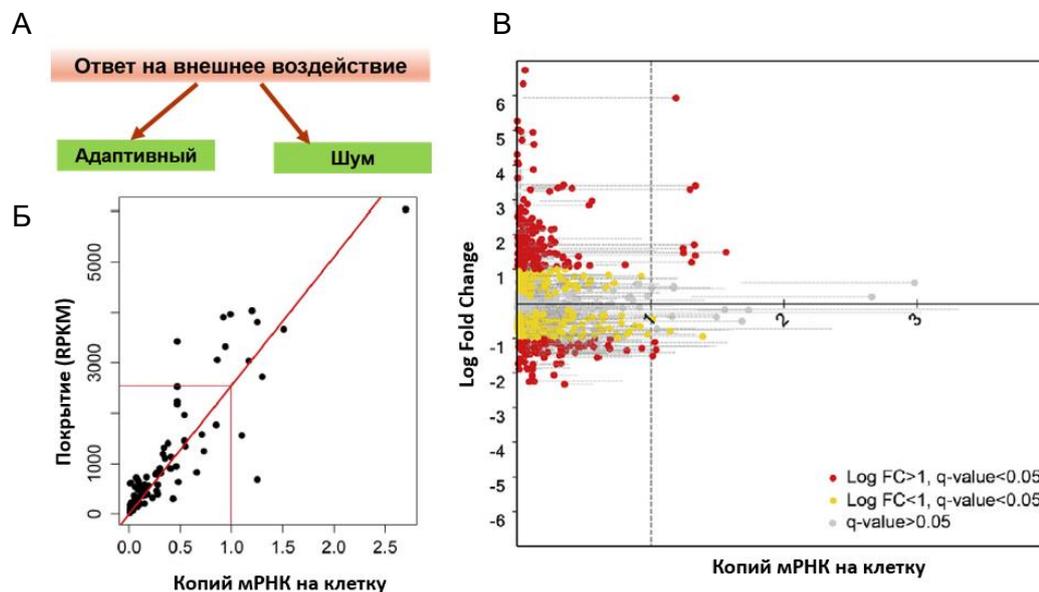
С использованием метода множественной регрессии мы построили модель, предсказывающую связывание рибосомы с мРНК (т.е. представленность мРНК во фракции, связанной с рибосомами). В качестве предикторов мы использовали представленность мРНК в СЦ фракции, силу РБС, стабильность вторичной структуры 5'-нетранслируемой области мРНК, ГЦ состав гена. Наша модель объясняет 72% представленности мРНК при нормальных условиях и 38% – при тепловом стрессе. Лучшим предиктором для экспоненциальной фазы является количество тотальной мРНК, а для теплового стресса – сила РБС (Рисунок 3.38).



**Рисунок 3.38 А – Мотив, идентифицированный у генов, чьи мРНК обогащены в РБ фракции при нормальных условиях. Б – Предсказание представленности мРНК в рибосомально-связанной фракции с помощью множественной регрессии при нормальных условиях и при тепловом стрессе. Звездочки показывают статистическую значимость соответствующего фактора (предиктора), \*\*\* – p-value < 0.01, \*\* – p-value < 0.05. Опубликовано в [162]**

Тепловой стресс значительно влияет на представленность мРНК как в суммарной цитоплазматической фракции, так и во фракции мРНК, связанной с рибосомами (Рисунок 3.36

Б). Вектора изменения половины (53%) от общего пула мРНК в РС и СЦ сонаправлены (Рисунок 3.36 В). Однако, как известно из данных протеомного анализа *M. gallisepticum*, представленность только нескольких белков меняется в два и более раз [82] при тепловом стрессе. При проведении эксперимента мы использовали сублетальные значения стрессового воздействия, то есть такого, при котором не меняется количество жизнеспособных клеток. Для микоплазм показано, что количество копий большинства мРНК в клетке не превышает нескольких штук [163], [140]. Мы предположили, что для сохранения популяции, при изменении внешних условий все клетки должны адаптироваться, то есть белки, необходимые для адаптации, должны экспрессироваться во всех клетках во время стресса. Используя ранее полученные данные цифровой ПЦР в капле об абсолютной представленности мРНК в клетках *M. gallisepticum* [140], мы произвели калибровку покрытия, выраженного в RPKM в количестве копий мРНК на клетку (Рисунок 3.39). И далее рассматривали транслятом в контексте копий на клетку. Основываясь на абсолютной представленности мРНК и изменении частоты связывания рибосомы с транскриптом при тепловом стрессе (диапазон изменения мРНК, мы смогли разделить ответ на стресс на две части. Первая часть – это шум. Вторая – адаптивный ответ – транскрипты, чья представленность увеличивается до одной копии на клетку и более (Рисунок 3.39).



**Рисунок 3.39.** Анализ данных эксперимента по профилированию рибосом. А – Предположительное деление ответа клетки на стресс; Б – Связь покрытия (RPKM) с представленностью транскрипта в клетке – калибровочная кривая. В – Изменение представленности пула рибосомально-связанной мРНК на клетку при тепловом стрессе; Красные круги – транскрипты, чья представленность изменилась в два и более раз, желтые круги – менее двух раз; серые – измерены не достоверно; серыми пунктирными линиями показано исходное значение представленности мРНК в экспоненциальной фазе роста. Опубликовано в [162]

Представленность только 9 мРНК возросла во фракции мРНК, связанной с рибосомами до одной и более копий на клетку. При тепловом стрессе рибосома избирательно транслирует мРНК, кодирующие шаперон (*clpB*), иммуноглобулин-связывающие белки (*GCW\_93958*, *GCW\_93946*, *GCW\_93948*), белки, участвующие в делении клетки (*mraZ*, *mraW*, *ftsZ*), несколько белков с неизвестной функцией (*GCW\_03070*, *GCW\_02935*). Таким образом, мы обнаружили, что при действии теплового стресса на *M. gallisepticum* рибосомы проявляют избирательность в отношении к мРНК. В нормальных условиях представленность транскрипта в суммарной цитоплазматической фракции определяет связывание с рибосомой. При тепловом стрессе действуют иные механизмы, возможно, за счет белков, ассоциированных с рибосомой.

## ЗАКЛЮЧЕНИЕ

В настоящей работе были исследованы механизмы регуляции экспрессии генов – передачи генетической информации у бактерий с редуцированным геномом, на примере *Mycoplasma gallisepticum*. Мы обнаружили лишь 10 транскрипционных факторов, которые могут быть найдены с помощью сравнительно-геномного анализа, и лишь 4 из них были высококонсервативны среди бактерий в целом. Мы показали, что консервативные ТФ-ы, такие как MraZ и WhiA, регулируют гены, которые поддерживают основные клеточные функции – деление клетки и синтез белка. Другой рассмотренный в работе транскрипционный фактор – Fug – является примером того, что ортологичные транскрипционные факторы бактерий не всегда регулируют ортологичные гены-мишени даже внутри одного семейства, при этом функция мишеней остается схожей. Мы продемонстрировали, как сильно транскрипция *M. gallisepticum* может зависеть от наличия и последовательности классических детерминант промотора (-10 и -35 боксов). Однако для значительного большинства генов бактерий с редуцированным геномом в силу снижения сложности структуры промотора это влияние отсутствует или уменьшено. Всё это указывает на то, что существуют иные механизмы регуляции. Л. Серрано была предложена модель регуляции экспрессии генов, где основной вклад в регуляцию вносит пространственная организация ДНК, вторичные структуры РНК, малые РНК, а действием транскрипционных факторов может быть объяснена лишь малая доля изменений уровня мРНК. Мы показали, что последовательности, окружающие промоторы генов микоплазм, содержат довольно короткие мотивы, наличие или отсутствие которых способно изменить уровень мРНК на два порядка. Эти последовательности могут выступать местами посадки регуляторов, образовывать вторичные структуры или участвовать в регуляции посредством других механизмов. Почему эволюция микоплазмы не избрала путь регуляции, основанный на изменении детерминант промотора и использования классических ТФ-ов, как у многих бактерий, а привела к использованию других механизмов, и насколько эти механизмы представлены в остальных живых организмах (сосуществуя вместе с уже известными), остается фундаментальным вопросом. Вместе с тем на примере С-белка приобретенной штаммом *M. gallisepticum* S6 системы рестрикции–модификации мы продемонстрировали, как существующая регуляторная сеть может усложняться вследствие наличия в геноме микоплазмы дополнительных мишеней, которые могут распознаваться этим ТФ-ом.

Отвечая на вопрос, как изменение уровня мРНК при тепловом стрессе связано с адаптацией бактерии, мы показали, что в нормальных условиях связывание рибосомы с мРНК напрямую зависит от уровня мРНК, в то время как при тепловом стрессе – в большей степени

от силы РБС (сайт связывания рибосомы) и ГЦ-состава мРНК. Проведя калибровку покрытия мРНК, выраженного в РРКМ в количестве копий мРНК на клетку, мы смогли разделить ответ на стресс на две части. Первая часть – это шум. Вторая – адаптивный ответ – транскрипты, чья представленность во фракции мРНК, связанной рибосомами, увеличивается до одной копии на клетку и более. Мы показали, что при тепловом стрессе рибосома избирательно транслирует мРНК, кодирующие шаперон (*clpB*), иммуноглобулин-связывающие белки (*GCW\_93958*, *GCW\_93946*, *GCW\_93948*), белки, участвующие в делении клетки (*mraZ*, *mraW*, *ftsZ*), несколько белков с неизвестной функцией (*GCW\_03070*, *GCW\_02935*). Вероятно, в естественной среде обитания *M. gallisepticum* (организме птиц), повышение температуры – признак воспаления, что является сигналом для микоплазмы для борьбы с денатурацией собственных белков (с помощью шаперонов), активации системы избегания иммунного ответа хозяина (с помощью IgG-связывающих белков) и запуска активного размножения (при успешной работе механизмов борьбы с иммунной системой птиц, нахождение внутри эукариот – это идеальное место для жизни с постоянным и стабильным доступом питательных веществ).

Дальнейшие исследования будут направлены на установление точных механизмов регуляции, чьи мишени были обнаружены в нашей работе, а также на тщательное описание детерминант эффективности трансляции у микоплазм с целью использования для нужд синтетической биологии.

## ВЫВОДЫ

1. На основании аннотированных функциональных доменов белков *M. gallisepticum* обнаружено 10 потенциальных факторов регуляции транскрипции.
2. Установлен сайт связывания белка MraZ, состоящий из серии прямых повторов – AAAGTG[T/G]. MraZ активирует транскрипцию генов кластера деления. Сверхэкспрессия *mraZ* приводит к филаментации клеток.
3. Найден и подтвержден сайт связывания белка Xre (HsdC) – повторы вида (GTGTTAN<sub>5</sub>)<sub>2</sub>. Xre выступает репрессором генов одной из систем рестрикции–модификации.
4. Найдена мишень белка Fur *M. gallisepticum*. На примере регулона Fur показано, что ортологичные регуляторы даже внутри одного семейства бактерий могут иметь разные мишени, но сохранять мотив сайта связывания. При этом функции мишеней в клетке остаются схожими.
5. Охарактеризован ассиметричный сайт связывания белка WhiA, состоящий из последовательности GATACACCN<sub>7</sub>GTTGTC, расположенной в промоторной области генов рибосомных белков. В нормальных условиях WhiA является репрессором.
6. Экспериментально подтверждено, что экспрессия генов *uvrB*, *parA*, *rplJ* зависит от действия регуляторов. Найден новый белок ParX, который регулирует экспрессию *parA*.
7. Тепловой стресс вызывает значительное увеличение транскрипционного шума в популяции *M. gallisepticum*, при этом транслируются преимущественно мРНК, кодирующие белки, которые обеспечивают адаптивный ответ на стресс.

## СПИСОК ЛИТЕРАТУРЫ

1. Crick F. On Protein Synthesis // Symp. Soc. Exp. Biol. 1958. P. 138–166.
2. Gamow G., Rich A., Yčas M. The Problem of Information Transfer from the Nucleic Acids to Proteins // Proc. Natl. Acad. Sci. 1956. Vol. 41, № 12. P. 23–68.
3. Crick F. What Mad Pursuit: A Personal View of Scientific Discovery // New York: Basic Books. 1989. Vol. 1, № 1. 182 p.
4. Schreiber S.L. Small molecules: the missing link in the central dogma. // Nat. Chem. Biol. 2005. Vol. 1, № 2. P. 64–66.
5. Winkler W.C., Breaker R.R. Genetic control by metabolite-binding riboswitches // ChemBioChem. 2003. Vol. 4, № 10. P. 1024–1032.
6. Crick F. Central Dogma of Molecular Biology // Nature. 1970. Vol. 227, № 5258. P. 561–563.
7. Chen Y.Y., Galloway K.E., Smolke C.D. Conceptual frameworks for biological design Synthetic biology: advancing biological frontiers by building synthetic systems // Genome Biol. 2012. Vol. 13. P. 240.
8. Irene C., Maciariello C., Cioci F., Camilloni G., Newlon C.S., Fabiani L. Identification of the sequences required for chromosomal replicator function in *Kluyveromyces lactis* // Mol. Microbiol. 2004. Vol. 51, № 5. P. 1413–1423.
9. Zhang S., Guo F., Yan W., Dai Z., Dong W., Zhou J., Zhang W., Xin F., Jiang M. Recent Advances of CRISPR/Cas9-Based Genetic Engineering and Transcriptional Regulation in Industrial Biology // Front. Bioeng. Biotechnol. 2020. Vol. 7, № January. P. 1–11.
10. Lajoie M.J., Rovner A.J., Goodman D.B., Aerni H., Haimovich A.D., Kuznetsov G., Mercer J.A., Wang H.H., Carr P.A., Mosberg J.A., Rohland N., Schultz P.G., Jacobson J.M., Rinehart J., Church G.M., Isaacs F.J. Genomically Recoded Organisms Expand Biological Functions // Science (80-. ). 2013. Vol. 342, № 6156. P. 357–360.
11. Chatterjee A., Lajoie M.J., Xiao H., Church G.M., Schultz P.G. A bacterial strain with a unique quadruplet codon specifying non-native amino acids // ChemBioChem. 2014. Vol. 15, № 12. P. 1782–1786.
12. Liu C.C., Jewett M.C., Chin J.W., Voigt C.A. Toward an orthogonal central dogma // Nat. Chem. Biol. Nature Publishing Group, 2018. Vol. 14, № 2. P. 103–106.
13. Saecker R.M., Record M.T., Dehaseth P.L. Mechanism of bacterial transcription initiation: RNA polymerase - Promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis // J. Mol. Biol. Elsevier Ltd, 2011. Vol. 412, № 5. P. 754–771.

14. Yang X., Lewis P.J. The interaction between bacterial transcription factors and RNA polymerase during the transition from initiation to elongation // *Transcription*. 2010. Vol. 1, № 2. P. 66–69.
15. Peters J.M., Vangeloff A.D., Landick R. Bacterial transcription terminators: The RNA 3'-end chronicles // *J. Mol. Biol. Elsevier B.V.*, 2011. Vol. 412, № 5. P. 793–813.
16. Bervoets I., Charlier D. Diversity, versatility and complexity of bacterial gene regulation mechanisms: Opportunities and drawbacks for applications in synthetic biology // *FEMS Microbiol. Rev.* 2019. Vol. 43, № 3. P. 304–339.
17. Browning D.F., Busby S.J. The regulation of bacterial transcription initiation. // *Nat. Rev. Microbiol.* Nature Publishing Group, 2004. Vol. 2, № 1. P. 57–65.
18. Hughes K.T., Mathee K. The anti-sigma factors. // *Annu. Rev. Microbiol.* 1998. Vol. 52. P. 231–286.
19. Beaucher J., Rodrigue S., Pierre-Étienne J., Smith I., Brzezinski R., Gaudreau L. Novel *Mycobacterium tuberculosis* anti- $\sigma$  factor antagonists control  $\sigma$ F activity by distinct mechanisms // *Mol. Microbiol.* 2002. Vol. 45, № 6. P. 1527–1540.
20. Wassarman K.M. 6S RNA, a Global Regulator of Transcription // *Regul. with RNA Bact. Archaea*. 2018. Vol. 6, № 3. P. 355–367.
21. Ulrich L. One-component systems dominate signal transduction in prokaryotes // *Trends Microbiol.* NIH Public Access, 2005. Vol. 22, № 13(2). P. 52–56.
22. Browning D.F., Busby S.J.W. Local and global regulation of transcription initiation in bacteria // *Nat. Rev. Microbiol.* Nature Publishing Group, 2016. Vol. 14, № 10. P. 638–650.
23. Sperandio V., Mellies J.L., Delahay R.M., Frankel G., Adam Crawford J., Nguyen W., Kaper J.B. Activation of enteropathogenic *Escherichia coli* (EPEC) LEE2 and LEE3 operons by Ler // *Mol. Microbiol.* 2000. Vol. 38, № 4. P. 781–793.
24. Botas A., Pérez-Redondo R., Rodríguez-García A., Álvarez-álvarez R., Yagüe P., Manteca A., Liras P. ArgR of *Streptomyces coelicolor* is a pleiotropic transcriptional regulator: Effect on the transcriptome, antibiotic production, and differentiation in liquid cultures // *Front. Microbiol.* 2018. Vol. 9, № MAR. P. 1–18.
25. De Spicer P.O., Maloy S. PutA protein, a membrane-associated flavin dehydrogenase, acts as a redox- dependent transcriptional regulator // *Proceedings of the National Academy of Sciences of the United States of America*. 1993. Vol. 90, № 9. P. 4295–4298.
26. Gunka K., Newman J.A., Commichau F.M., Herzberg C., Rodrigues C., Hewitt L., Lewis R.J., Stülke J. Functional dissection of a trigger enzyme: Mutations of the *Bacillus subtilis* glutamate dehydrogenase RocG that affect differentially its catalytic activity and regulatory

- properties // *J. Mol. Biol.* Elsevier Ltd, 2010. Vol. 400, № 4. P. 815–827.
27. Sudarsan N., Lee E.R., Weinberg Z., Moy R.H., Kim J.N., Link K.H., Breaker R.R. Riboswitches in Eubacteria Sense the Second Messenger Cyclic Di-GMP // *Science* (80-. ). 2008. Vol. 321, № 5887. P. 411–413.
  28. Nomura M., Yates J.L., Dean D., Post L.E. Feedback regulation of ribosomal protein gene expression in *Escherichia coli*: Structural homology of ribosomal RNA and ribosomal protein mRNA // *Proc. Natl. Acad. Sci. U. S. A.* 1980. Vol. 77, № 12 II. P. 7084–7088.
  29. Vitreschak A.G., Mironov A.A., Lyubetsky V.A., Gelfand M.S. Comparative genomic analysis of T-box regulatory systems in bacteria // *Rna*. 2008. Vol. 14, № 4. P. 717–735.
  30. Naville M., Gautheret D. Transcription attenuation in bacteria: Theme and variations // *Briefings Funct. Genomics Proteomics*. 2009. Vol. 8, № 6. P. 482–492.
  31. Oliva G., Sahr T., Buchrieser C. Small RNAs, 5' UTR elements and RNA-binding proteins in intracellular bacteria: Impact on metabolism and virulence // *FEMS Microbiol. Rev.* 2015. Vol. 39, № 3. P. 331–349.
  32. Andrade J.M., Hajnsdorf E., Régnier P., Arraiano C.M. The poly(A)-dependent degradation pathway of rpsO mRNA is primarily mediated by RNase R // *Rna*. 2009. Vol. 15, № 2. P. 316–326.
  33. Wang Q., Zhang Y., Yang C., Xiong H., Lin Y., Yao J., Li H., Xie L., Zhao W., Yao Y., Ning Z.-B., Zeng R., Xiong Y., Guan K.-L., Zhao S., Zhao G.-P. Acetylation of Metabolic Enzymes Coordinates Carbon Source Utilization and Metabolic Flux // *Science* (80-. ). 2010. Vol. 327, № 5968. P. 1004–1007.
  34. Grant G.A. Contrasting catalytic and allosteric mechanisms for phosphoglycerate dehydrogenases // *Arch. Biochem. Biophys.* 2012. Vol. 519, № 2. P. 175–185.
  35. Wilson C.J., Zhan H., Swint-Kruse L., Matthews K.S. The lactose repressor system: Paradigms for regulation, allosteric behavior and protein folding // *Cell. Mol. Life Sci.* 2007. Vol. 64, № 1. P. 3–16.
  36. Galas D.J., Schmitz A. DNAase footprinting a simple method for the detection of protein-DNA binding specificity // *Nucleic Acids Res.* 1978. Vol. 5, № 9. P. 3157–3170.
  37. Hampshire A.J., Rusling D.A., Broughton-Head V.J., Fox K.R. Footprinting: A method for determining the sequence selectivity, affinity and kinetics of DNA-binding ligands // *Methods*. 2007. Vol. 42, № 2. P. 128–140.
  38. Strauss E.C., Orkin S.H. In vivo protein-DNA interactions at hypersensitive site 3 of the human beta-globin locus control region. // *Proc. Natl. Acad. Sci.* 1992. Vol. 89, № 13. P. 5809–5813.

39. Hesselberth J.R., Chen X., Zhang Z., Sabo P.J., Sandstrom R., Reynolds A.P., Thurman R.E., Neph S., Kuehn M.S., Noble W.S., Fields S., Stamatoyannopoulos J.A. Global mapping of protein-DNA interactions in vivo by digital genomic footprinting // *Nat. Methods*. 2009. Vol. 6, № 4. P. 283–289.
40. Vierstra J., Stamatoyannopoulos J.A. Genomic footprinting // *Nat. Methods*. 2016. Vol. 13, № 3. P. 213–221.
41. Hellman L.M., Fried M.G. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. // *Nat. Protoc. Nature Publishing Group*, 2007. Vol. 2, № 8. P. 1849–1861.
42. Tolstonog G. V., Li G., Shoeman R.L., Traub P. Interaction in vitro of type III intermediate filament proteins with higher order structures of single-stranded DNA, particularly with G-quadruplex DNA // *DNA Cell Biol.* 2005. Vol. 24, № 2. P. 85–110.
43. Varshavsky A. Electrophoretic Assay for DNA-Binding Proteins // *Methods Enzymol.* 1987. Vol. 151, № C. P. 551–565.
44. Woo A.J., Dods J.S., Susanto E., Ulgiati D., Abraham L.J. A proteomics approach for the identification of DNA binding activities observed in the electrophoretic mobility shift assay. // *Mol. Cell. Proteomics*. 2002. Vol. 1, № 6. P. 472–478.
45. Fried M.G., Liu G. Molecular sequestration stabilizes CAP-DNA complexes during polyacrylamide gel electrophoresis // *Nucleic Acids Res.* 1994. Vol. 22, № 23. P. 5054–5059.
46. Ellington A.D., Szostak J.W. In vitro selection of RNA molecules that bind specific ligands // *Nature*. 1990. Vol. 346, № 6287. P. 818–822.
47. Lorenz C., von Pelchrzim F., Schroeder R. Genomic systematic evolution of ligands by exponential enrichment (Genomic SELEX) for the identification of protein-binding RNAs independent of their expression levels // *Nat. Protoc.* 2006. Vol. 1, № 5. P. 2204–2212.
48. Singer B.S., Shtatland T., Brown D., Gold L. Libraries for genomic SELEX // *Nucleic Acids Res.* 1997. Vol. 25, № 4. P. 781–786.
49. Shimada T., Fujita N., Maeda M., Ishihama A. Systematic search for the Cra-binding promoters using genomic SELEX system // *Genes to Cells*. 2005. Vol. 10, № 9. P. 907–918.
50. Shimada T., Ogasawara H., Ishihama A. Genomic SELEX screening of regulatory targets of *Escherichia coli* transcription factors // *Methods Mol. Biol.* 2018. Vol. 1837. P. 49–69.
51. Solomon M.J., Varshavsky A. Formaldehyde-mediated DNA-protein crosslinking: A probe for in vivo chromatin structures // *Proc. Natl. Acad. Sci. U. S. A.* 1985. Vol. 82, № 19. P. 6470–6474.

52. Rhee H.S., Pugh B.F. Comprehensive Genome-wide Protein-DNA Interactions Detected at Single-Nucleotide Resolution // *Cell*. 2011. Vol. 147, № 6. P. 1408–1419.
53. Hoffman E.A., Frey B.L., Smith L.M., Auble D.T. Formaldehyde crosslinking: A tool for the study of chromatin complexes // *J. Biol. Chem.* 2015. Vol. 290, № 44. P. 26404–26411.
54. Lieb J.D., Liu X., Botstein D., Brown P.O. Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association // *Nat. Genet.* 2001. Vol. 28, № 4. P. 327–334.
55. Johnson D.S., Mortazavi A., Myers R.M., Wold B. Genome-Wide Mapping of in Vivo Protein-DNA Interactions // *Science (80-. )*. 2007. Vol. 316, № 5830. P. 1497–1502.
56. Venters B.J., Pugh B.F. A canonical promoter organization of the transcription machinery and its regulators in the *Saccharomyces* genome // *Genome Res.* 2009. Vol. 19, № 3. P. 360–371.
57. Kuznetsov V.A. Relative Avidity, Specificity, and Sensitivity of Transcription Factor–DNA Binding in Genome-Scale Experiments // *Methods in molecular biology (Clifton, N.J.)*. 2009. Vol. 563. P. 15–50.
58. Rossi M.J., Lai W.K.M., Pugh B.F. Simplified ChIP-exo assays // *Nat. Commun.* 2018. Vol. 9, № 1. P. 1–13.
59. Rodionov D.A. Comparative Genomic Reconstruction of Transcriptional Regulatory Networks in Bacteria // *Chem. Rev.* 2007. Vol. 107, № 8. P. 3467–3497.
60. Tung B.K. Le, Maxim V. Imakaev, Leonid A. Mirny M.T.L. High-resolution mapping of the spatial organization of a bacterial chromosome // *Science (80-. )*. 2014. Vol. 342, № 6159. P. 731–734.
61. Rogozin I.B. Connected gene neighborhoods in prokaryotic genomes // *Nucleic Acids Res.* 2002. Vol. 30, № 10. P. 2212–2223.
62. Das M.K., Dai H.K. A survey of DNA motif finding algorithms // *BMC Bioinformatics*. 2007. Vol. 8, № SUPPL. 7. P. 1–13.
63. Chang Y.C., Hu Z., Rachlin J., Anton B.P., Kasif S., Roberts R.J., Steffen M. COMBREX-DB: An experiment centered database of protein function: Knowledge, predictions and knowledge gaps // *Nucleic Acids Res.* 2016. Vol. 44, № D1. P. D330–D335.
64. Ishihama A., Shimada T., Yamazaki Y. Transcription profile of *Escherichia coli*: Genomic SELEX search for regulatory targets of transcription factors // *Nucleic Acids Res.* 2016. Vol. 44, № 5. P. 2058–2074.
65. Gao Y., Yurkovich J.T., Seo S.W., Kabimoldayev I., Dräger A., Chen K., Sastry A. V., Fang X., Mih N., Yang L., Eichner J., Cho B.K., Kim D., Palsson B.O. Systematic discovery of

- uncharacterized transcription factors in *Escherichia coli* K-12 MG1655 // *Nucleic Acids Res.* 2018. Vol. 46, № 20. P. 10682–10696.
66. Janga S.C., Contreras-Moreira B. Dissecting the expression patterns of transcription factors across conditions using an integrated network-based approach // *Nucleic Acids Res.* 2010. Vol. 38, № 20. P. 6841–6856.
67. Schmidt A., Kochanowski K., Vedelaar S., Ahrné E., Volkmer B., Callipo L., Knoop K., Bauer M., Aebersold R., Heinemann M. The quantitative and condition-dependent *Escherichia coli* proteome // *Nat. Biotechnol.* 2016. Vol. 34, № 1. P. 104–110.
68. Sirand-Pugnet P., Lartigue C., Marena M., Jacob D., Barré A., Barbe V., Schenowitz C., Mangenot S., Couloux A., Segurens B., De Daruvar A., Blanchard A., Citti C. Being pathogenic, plastic, and sexual while living with a nearly minimal bacterial genome // *PLoS Genet.* 2007. Vol. 3, № 5. P. 744–758.
69. Razin S. The mycoplasmas. // *Microbiol. Rev.* 1978. Vol. 42, № 2. P. 414–470.
70. Güell M., Van Noort V., Yus E., Chen W.-H., Leigh-Bell J., Michalodimitrakis K., Yamada T., Arumugam M., Doerks T., Kühner S., Rode M., Suyama M., Schmidt S., Gavin A.-C., Bork P., Serrano L. Transcriptome complexity in a genome-reduced bacterium. // *Science* (80-. ). American Association for the Advancement of Science, 2009. Vol. 326, № 5957. P. 1268–1271.
71. Weiner J., Herrmann R., Browning G.F. Transcription in *Mycoplasma pneumoniae*. // *Nucleic Acids Res.* Oxford University Press, 2000. Vol. 28, № 22. P. 4488–4496.
72. Bébéar C., Pereyre S., Peuchant O. *Mycoplasma pneumoniae* : susceptibility and resistance to antibiotics // *Future Microbiol.* 2011. Vol. 6, № 4. P. 423–431.
73. Taschke C., Klinkert M.Q., Wolters J., Herrmann R. Organization of the ribosomal RNA genes in *Mycoplasma hyopneumoniae*: The 5S rRNA gene is separated from the 16S and 23S rRNA genes // *MGG Mol. Gen. Genet.* 1986. Vol. 205, № 3. P. 428–433.
74. Gafny R., Hyman H.C., Razin S., Glaser G. Promoters of *Mycoplasma capricolum* ribosomal RNA operons: identical activities but different regulation in homologous and heterologous cells // *Nucleic Acids Res.* 1988. Vol. 16, № 1. P. 61–76.
75. Torres-Puig S., Broto A., Querol E., Piñol J., Pich O.Q. A novel sigma factor reveals a unique regulon controlling cell-specific recombination in *Mycoplasma genitalium* // *Nucleic Acids Res.* 2015. Vol. 43, № 10. P. 4923–4936.
76. Chang L.-J., Chen W.-H., Minion F.C., Shiuan D. Mycoplasmas regulate the expression of heat-shock protein genes through CIRCE-HrcA interactions. // *Biochem. Biophys. Res. Commun.* 2008. Vol. 367, № 1. P. 213–218.

77. Carvalho F.M., Fonseca M.M., Batistuzzo De Medeiros S., Scortecci K.C., Blaha C.A.G., Agnez-Lima L.F. DNA repair in reduced genome: the Mycoplasma model. // *Gene*. 2005. Vol. 360, № 2. P. 111–119.
78. Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A., Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M., Fritchman J.L., Weidman J.F., Small K. V., Sandusky M., Fuhrmann J., Nguyen D., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M., Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S. The Minimal Gene Complement of Mycoplasma genitalium // *Science* (80-. ). 1995. Vol. 270, № 5235. P. 397–404.
79. III J.W., Herrmann R., Browning G.F. Transcription in Mycoplasma pneumoniae. // *Nucleic Acids Res.* 2000. Vol. 28, № 22. P. 4488–4496.
80. Ge H., Walhout A.J.M., Vidal M. Integrating “omic” information: A bridge between genomics and systems biology // *Trends Genet.* 2003. Vol. 19, № 10. P. 551–560.
81. Yus E., Maier T., Michalodimitrakis K., Van Noort V., Yamada T., Chen W.-H., Wodke J. a H., Güell M., Martínez S., Bourgeois R., Kühner S., Raineri E., Letunic I., Kalinina O. V., Rode M., Herrmann R., Gutiérrez-Gallego R., Russell R.B., Gavin A.-C., Bork P., Serrano L. Impact of genome reduction on bacterial metabolism and its regulation. // *Science* (80-. ). Aaas, 2009. Vol. 326, № 5957. P. 1263–1268.
82. Mazin P. V., Fisunov G.Y., Gorbachev A.Y., Kapitskaya K.Y., Altukhov I.A., Semashko T.A., Alexeev D.G., Govorun V.M. Transcriptome analysis reveals novel regulatory mechanisms in a genome-reduced bacterium // *Nucleic Acids Res.* 2014. Vol. 42, № 21. P. 13254–13268.
83. Lloréns-Rico V., Cano J., Kamminga T., Gil R., Latorre A., Chen W.H., Bork P., Glass J.I., Serrano L., Lluch-Senar M. Bacterial antisense RNAs are mainly the product of transcriptional noise // *Sci. Adv.* 2016. Vol. 2, № 3. P. 1–9.
84. Miravet-Verde S., Lloréns-Rico V., Serrano L. Alternative transcriptional regulation in genome-reduced bacteria // *Curr. Opin. Microbiol.* 2017. Vol. 39. P. 89–95.
85. Yus E., Lloréns-Rico V., Martínez S., Gallo C., Eilers H., Blötz C., Stülke J., Lluch-Senar M., Serrano L. Determination of the Gene Regulatory Network of a Genome-Reduced Bacterium Highlights Alternative Regulation Independent of Transcription Factors // *Cell Syst.* 2019. Vol. 9, № 2. P. 143-158.e13.
86. Razin S., Yogev D., Naot Y. Molecular Biology and Pathogenicity of Mycoplasmas // *Molecular Biology and Pathogenicity of Mycoplasmas*. 1998. Vol. 62, № 4.
87. Grosjean H., Breton M., Sirand-Pugnet P., Tardy F., Thiaucourt F., Citti C., Barré A., Yoshizawa S., Fourmy D., de Crécy-Lagard V., Blanchard A. Predicting the Minimal

- Translation Apparatus: Lessons from the Reductive Evolution of Mollicutes // *PLoS Genet.* 2014. Vol. 10, № 5.
88. Byrgazov K., Vesper O., Moll I. Ribosome heterogeneity : another level of complexity in bacterial translation regulation // *Curr. Opin. Microbiol.* Elsevier Ltd, 2013. Vol. 16, № 2. P. 133–139.
  89. Li L., Boniecki M.T., Jaffe J.D., Imai B.S., Yau P.M., Luthey-Schulten Z.A., Martinis S.A. Naturally occurring aminoacyl-tRNA synthetases editing-domain mutations that cause mistranslation in *Mycoplasma* parasites // *Proc. Natl. Acad. Sci. U. S. A.* 2011. Vol. 108, № 23. P. 9378–9383.
  90. Reilly F.J.O., Xue L., Graziadei A., Sinn L., Lenz S., Tegunov D., Blötz C., Singh N., Hagen W.J.H., Cramer P., Stülke J. In-cell architecture of an actively transcribing- translating expressome // *Science* (80-. ). 2020. Vol. 557, № July. P. 554–557.
  91. Samatova E., Daberger J., Liutkute M., Rodnina M. V. Translational Control by Ribosome Pausing in Bacteria : How a Non-uniform Pace of Translation Affects Protein Production and Folding // *Front. Microbiol.* 2021. Vol. 11, № January.
  92. Edgar R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput // *Nucleic Acids Res.* 2004. Vol. 32, № 5. P. 1792–1797.
  93. Tamura K., Peterson D., Peterson N., Stecher G., Nei M., Kumar S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods // *Mol. Biol. Evol.* 2011. Vol. 28, № 10. P. 2731–2739.
  94. Novichkov P.S., Kazakov A.E., Ravcheev D.A., Leyn S.A., Kovaleva G.Y., Sutormin R.A., Kazanov M.D., Riehl W., Arkin A.P., Dubchak I., Rodionov D.A. RegPrecise 3.0--a resource for genome-scale exploration of transcriptional regulation in bacteria. // *BMC Genomics.* BMC Genomics, 2013. Vol. 14, № 1. P. 745.
  95. Garanina I.A., Fisunov G.Y., Govorun V.M. BAC-BROWSER: The tool for visualization and analysis of prokaryotic genomes // *Front. Microbiol.* 2018. Vol. 9, № NOV. P. 1–6.
  96. Fisunov G.Y., Garanina I.A., Evsyutina D. V., Semashko T.A., Nikitina A.S., Govorun V.M. Reconstruction of transcription control networks in mollicutes by high-throughput identification of promoters // *Front. Microbiol.* 2016. Vol. 7, № DEC. P. 1–15.
  97. Morowitz H.J., Maniloff J. Analysis of the life cycle of *Mycoplasma gallisepticum*. // *J. Bacteriol.* 1966. Vol. 91, № 4. P. 1638–1644.
  98. Green R., Rogers E.J. Transformation of chemically competent *E. coli* // *Methods in Enzymology.* 1st ed. Elsevier Inc., 2013. Vol. 529. 329–336 p.
  99. Laemmli U.K. Cleavage of structural proteins during the assembly of the head of

- bacteriophage T4 // *Nature*. 1970. Vol. 227, № 5259. P. 680–685.
100. Evsyutina D. V., Fisunov G.Y., Pobeguts O. V., Kovalchuk S.I., Govorun V.M. Gene Silencing through CRISPR Interference in *Mycoplasmas* // *Microorganisms*. 2022. Vol. 10, № 6.
  101. Ryder S.P., Recht M.I., Williamson J.R. Quantitative analysis of protein-RNA interactions by gel mobility shift // *Methods Mol. Biol.* 2010. Vol. 488. P. 99–115.
  102. Lee S.W., Browning G.F., Markham P.F. Development of a replicable oriC plasmid for *Mycoplasma gallisepticum* and *Mycoplasma imitans*, and gene disruption through homologous recombination in *M. gallisepticum* // *Microbiology*. 2008. Vol. 154, № 9. P. 2571–2580.
  103. Qi L.S., Larson M.H., Gilbert L.A., Doudna J.A., Weissman J.S., Arkin A.P., Lim W.A. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression // *Cell*. Elsevier, 2013. Vol. 152, № 5. P. 1173–1183.
  104. Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity // *Science* (80- ). 2012. Vol. 337, № 6096. P. 816–821.
  105. Fisunov G.Y., Evsyutina D. V, Arzamasov A.A., Butenko I.O., Govorun V.M. Profiling of *Mycoplasma gallisepticum* Ribosomes // *Acta Naturae*. 2015. Vol. 7, № 27. P. 107–112.
  106. Bolger A.M., Lohse M., Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data // *Bioinformatics*. 2014. Vol. 30, № 15. P. 2114–2120.
  107. Langmead B., Salzberg S.L. Fast gapped-read alignment with Bowtie 2 // *Nat. Methods*. 2012. Vol. 9, № 4. P. 357–359.
  108. Shi J., Li F., Wen A., Yu L., Wang L., Wang F., Jin Y., Jin S., Feng Y., Lin W. Structural basis of transcription activation by the global regulator Spx // *Nucleic Acids Res. Oxford University Press*, 2021. Vol. 49, № 18. P. 10756–10769.
  109. Kriel A., Bittner A.N., Kim S.H., Liu K., Tehranchi A.K., Zou W.Y., Rendon S., Chen R., Tu B.P., Wang J.D. Direct Regulation of GTP Homeostasis by (p)ppGpp: A Critical Component of Viability and Stress Resistance // *Mol. Cell*. 2013. Vol. 48, № 2. P. 231–241.
  110. Nessler S., Fieulaine S., Poncet S., Galinier A., Deutscher J. HPr Kinase / Phosphorylase , the Sensor Enzyme of Catabolite Repression in Gram-Positive Bacteria : Structural Aspects of the Enzyme and the Complex with Its Protein Substrate // *J. Bacteriol.* 2003. Vol. 185, № 14. P. 4003–4010.
  111. Eraso J.M., Markillie L.M., Mitchell H.D., Taylor R.C., Orr G., Margolin W. The highly conserved MraZ protein is a transcriptional regulator in *Escherichia coli* // *J. Bacteriol.* 2014.

Vol. 196, № 11. P. 2053–2066.

112. Fisunov G.Y., Evsyutina D. V., Semashko T.A., Arzamasov A.A., Manuvera V.A., Letarov A. V., Govorun V.M. Binding site of MraZ transcription factor in Mollicutes // *Biochimie. Elsevier B.V*, 2016. Vol. 125. P. 59–65.
113. Chen S., Jancrick J., Yokota H., Kim R., Kim S.-H. Crystal structure of a protein associated with cell division from *Mycoplasma pneumoniae* (GI: 13508053): A novel fold with a conserved sequence motif // *Proteins Struct. Funct. Bioinforma.* 2004. Vol. 55, № 4. P. 785–791.
114. Adams M.A., Udell C.M., Pal G.P., Jia Z. MraZ from *Escherichia coli*: Cloning, purification, crystallization and preliminary X-ray analysis // *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 2005. Vol. 61, № 4. P. 378–380.
115. Fisunov G.Y., Evsyutina D. V., Manuvera V.A., Govorun V.M. Binding site of restriction-modification system controller protein in Mollicutes // *BMC Microbiol. BMC Microbiology*, 2017. Vol. 17, № 1. P. 1–7.
116. Streeter S.D., Papapanagiotou I., McGeehan J.E., Kneale G.G. DNA footprinting and biophysical characterization of the controller protein C.AhdI suggests the basis of a genetic switch // *Nucleic Acids Res.* 2004. Vol. 32, № 21. P. 6445–6453.
117. Fillat M.F. The fur (ferric uptake regulator) superfamily: Diversity and versatility of key transcriptional regulators // *Arch. Biochem. Biophys. Elsevier Inc.*, 2014. Vol. 546. P. 41–52.
118. Sevilla E., Bes M.T., Peleato M.L., Fillat M.F. Fur-like proteins: Beyond the ferric uptake regulator (Fur) paralog // *Arch. Biochem. Biophys. Elsevier Inc.*, 2021. Vol. 701, № September 2020. P. 108770.
119. Andrews S.C., Robinson A.K., Rodríguez-Quñones F. Bacterial iron homeostasis // *FEMS Microbiol. Rev.* 2003. Vol. 27, № 2–3. P. 215–237.
120. Troxell B., Hassan H.M. Transcriptional regulation by Ferric Uptake Regulator (Fur) in pathogenic bacteria. // *Front. Cell. Infect. Microbiol.* 2013. Vol. 3, № October. P. 59.
121. Yu C., Genco C.A. Fur-mediated global regulatory circuits in pathogenic *Neisseria* species. // *J. Bacteriol.* 2012. Vol. 194, № 23. P. 6372–6381.
122. Martínez-Torró C., Torres-Puig S., Monge M., Sánchez-Alba L., González-Martín M., Marcos-Silva M., Perálvarez-Marín A., Canals F., Querol E., Piñol J., Pich O.Q. Transcriptional response to metal starvation in the emerging pathogen *Mycoplasma genitalium* is mediated by Fur-dependent and –independent regulatory pathways // *Emerg. Microbes Infect.* 2020. Vol. 9, № 1. P. 5–19.
123. Ainsa J.A., Ryding N.J., Hartley N., Findlay K.C., Bruton C.J., Chater K.F. WhiA, a protein

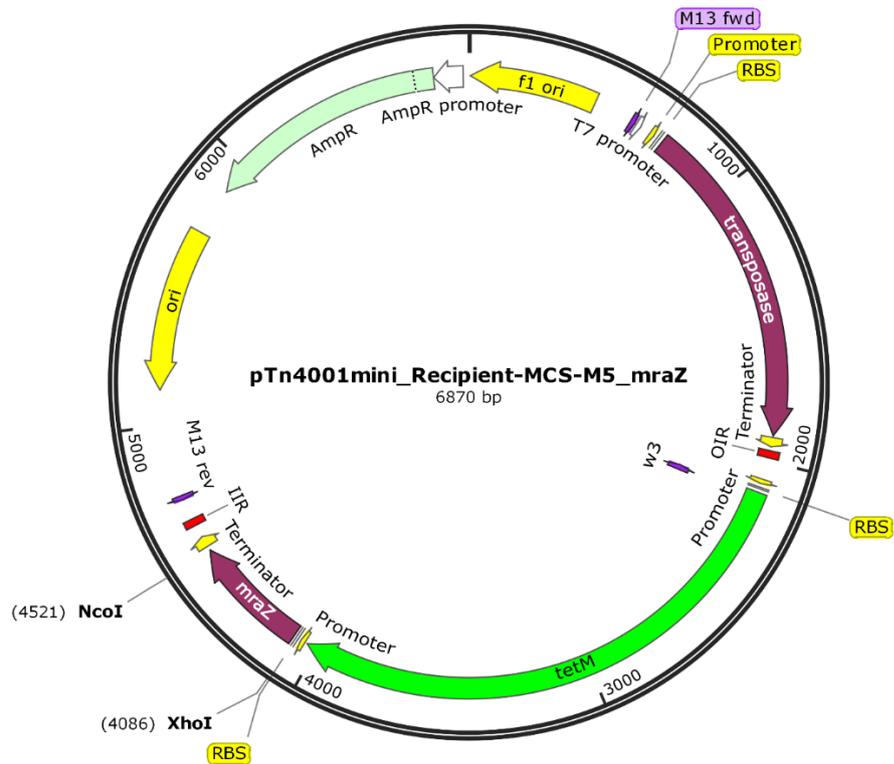
- of unknown function conserved among gram-positive bacteria, is essential for sporulation in *Streptomyces coelicolor* A3(2) // *J. Bacteriol.* 2000. Vol. 182, № 19. P. 5470–5478.
124. Kaiser B.K., Clifton M.C., Shen B.W., Stoddard B.L. The Structure of a Bacterial DUF199/WhiA Protein: Domestication of an Invasive Endonuclease // *Structure.* 2009. Vol. 17, № 10. P. 1368–1376.
  125. Knizewski L., Ginalski K. Bacterial DUF199/COG1481 proteins including sporulation regulator WhiA are distant homologs of LAGLIDADG homing endonucleases that retained only DNA binding // *Cell Cycle.* 2007. Vol. 6, № 13. P. 1666–1670.
  126. Jakimowicz D., Mouz S., Zakrzewska-Czerwińska J., Chater K.F. Developmental Control of a parAB Promoter Leads to Formation of Sporulation-Associated ParB Complexes in *Streptomyces coelicolor* // *J. Bacteriol.* 2006. Vol. 188, № 5. P. 1710–1720.
  127. Buttner M.J. Genes Required for Aerial Growth , Cell Division , and Chromosome *venezuelae* // *MBio.* 2013. Vol. 4, № 5. P. 1–18.
  128. Kaiser B.K., Stoddard B.L. DNA recognition and transcriptional regulation by the WhiA sporulation factor // *Sci. Rep.* 2011. Vol. 1. P. 1–9.
  129. Bush M.J., Chandra G., Bibb M.J., Findlay K.C., Buttner M.J. Genome-wide chromatin immunoprecipitation sequencing analysis shows that WhiB is a transcription factor that cocontrols its regulon with WhiA to initiate developmental cell division in *Streptomyces* // *MBio.* 2016. Vol. 7, № 2. P. 1–13.
  130. Bohorquez L.C., Surdova K., Jonker M.J., Hamoen L.W. The Conserved DNA Binding Protein WhiA Influences Chromosome Segregation in *Bacillus subtilis* // *J. Bacteriol.* / ed. Henkin T.M. 2018. Vol. 200, № 8. P. 1–17.
  131. Lee D.S., Kim P., Kim E.S., Kim Y., Lee H.S. *Corynebacterium glutamicum* WhcD interacts with WhiA to exert a regulatory effect on cell division genes // *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* Springer International Publishing, 2018. Vol. 111, № 5. P. 641–648.
  132. Lee J.H., Jeong H., Kim Y., Lee H.S. *Corynebacterium glutamicum* whiA plays roles in cell division, cell envelope formation, and general cell physiology // *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* Springer International Publishing, 2020. Vol. 113, № 5. P. 629–641.
  133. Eilers H. Transcription in *Mycoplasma pneumoniae*. the Georg-August-Universität Göttingen, 2010. 1–193 p.
  134. Ser H.L., Tan W.S., Ab Mutalib N.S., Yin W.F., Chan K.G., Goh B.H., Lee L.H. Genome sequence of *Streptomyces mangrovisoli* MUSC 149T isolated from intertidal sediments // *Brazilian J. Microbiol. Sociedade Brasileira de Microbiologia*, 2018. Vol. 49, № 1. P. 13–15.

135. Dybvig K., Voelker L.L. Molecular biology of Mycoplasmas // *Annu. Rev. Microbiol.* 1996. Vol. 50, № 1. P. 25–57.
136. Guhan N., Muniyappa K. Mycobacterium tuberculosis RecA intein, a LAGLIDADG homing endonuclease, displays Mn<sup>2+</sup> and DNA-dependent ATPase activity // *Nucleic Acids Res.* 2003. Vol. 31, № 14. P. 4184–4191.
137. Yaginuma H., Kawai S., Tabata K. V., Tomiyama K., Kakizuka A., Komatsuzaki T., Noji H., Imamura H. Diversity in ATP concentrations in a single bacterial cell population revealed by quantitative single-cell imaging // *Sci. Rep.* 2014. Vol. 4.
138. Clark M.A., Baumann L., Baumann P. Sequence analysis of a 34.7-kb DNA segment from the genome of Buchnera aphidicola (endosymbiont of aphids) containing groEL, dnaA, the atp operon, gidA, and rho // *Curr. Microbiol.* 1998. Vol. 36, № 3. P. 158–163.
139. Sharma C.M., Hoffmann S., Darfeuille F., Reignier J., Findeiß S., Sittka A., Chabas S., Reiche K., Hackermüller J., Reinhardt R., Stadler P.F., Vogel J. The primary transcriptome of the major human pathogen Helicobacter pylori // *Nature.* 2010. Vol. 464, № 7286. P. 250–255.
140. Gorbachev A.Y., Fisunov G.Y., Izraelson M., Evsyutina D. V, Mazin P. V, Alexeev D.G., Pobeguts O. V, Gorshkova T.N., Kovalchuk S.I., Kamashev D.E., Govorun V.M. DNA repair in Mycoplasma gallisepticum. // *BMC Genomics.* BMC Genomics, 2013. Vol. 14, № 1. P. 726.
141. Wurihan, Gezi, Brambilla E., Wang S., Sun H., Fan L., Shi Y., Sclavi B., Morigen. DnaA and LexA proteins regulate transcription of the uvrB gene in Escherichia coli: The role of DnaA in the control of the SOS regulon // *Front. Microbiol.* 2018. Vol. 9, № JUN. P. 1–16.
142. Xindan W., Paula Montero L., David Z.R. Organization and segregation of bacterial chromosomes // *Nat Rev Genet.* 2008. Vol. 86, № 12. P. 3279–3288.
143. Jakimowicz D., Chater K., Zakrzewska-Czerwińska J. The ParB protein of Streptomyces coelicolor A3(2) recognizes a cluster of parS sequences within the origin-proximal region of the linear chromosome // *Mol. Microbiol.* 2002. Vol. 45, № 5. P. 1365–1377.
144. Kawalek A., Wawrzyniak P., Bartosik A.A., Jagura-Burdzy G. Rules and exceptions: The role of chromosomal ParB in DNA segregation and other cellular processes // *Microorganisms.* 2020. Vol. 8, № 1. P. 19–22.
145. Yamaichi Y., Niki H. Active segregation by the Bacillus subtilis partitioning system in Escherichia coli // *Proc. Natl. Acad. Sci. U. S. A.* 2000. Vol. 97, № 26. P. 14656–14661.
146. Schumacher M.A. Structural biology of plasmid partition: Uncovering the molecular mechanisms of DNA segregation // *Biochem. J.* 2008. Vol. 412, № 1. P. 1–18.

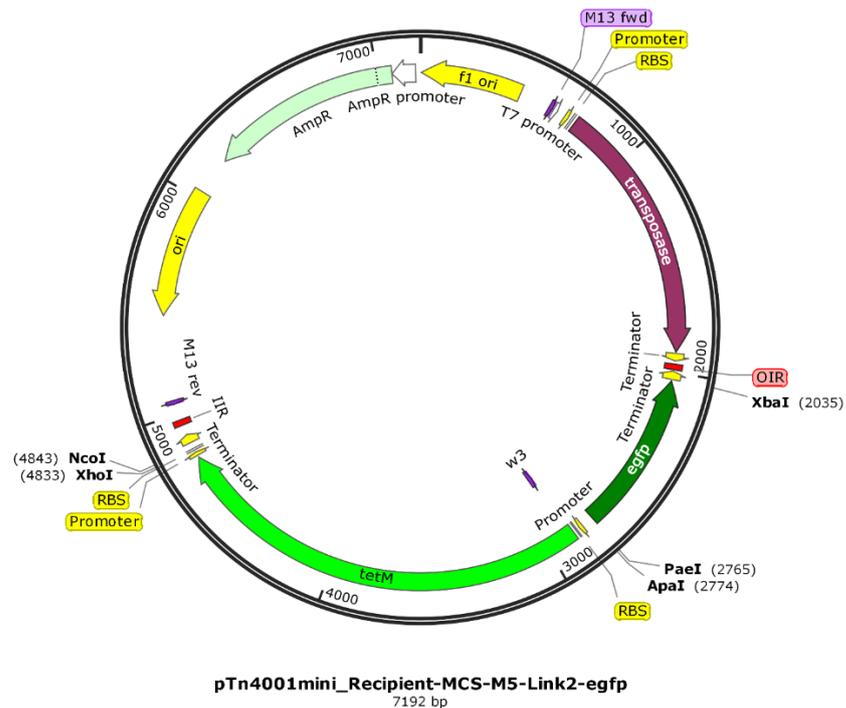
147. Chen B.W., Lin M.H., Chu C.H., Hsu C.E., Sun Y.J. Insights into ParB spreading from the complex structure of Spo0J and parS // *Proc. Natl. Acad. Sci. U. S. A.* 2015. Vol. 112, № 21. P. 6613–6618.
148. Soh Y.M., Davidson I.F., Zamuner S., Basquin J., Bock F.P., Taschner M., Veening J.W., de Los Rios P., Peters J.M., Gruber S. Self-organization of parS centromeres by the ParB CTP hydrolase // *Science* (80-. ). 2019. Vol. 366, № 6469. P. 1129–1133.
149. Pióro M., Jakimowicz D. Chromosome Segregation Proteins as Coordinators of Cell Cycle in Response to Environmental Conditions // *Front. Microbiol.* 2020. Vol. 11, № April.
150. Jung A., Raßbach A., Pulpetta R.L., van Teeseling M.C.F., Heinrich K., Sobetzko P., Serrania J., Becker A., Thanbichler M. Two-step chromosome segregation in the stalked budding bacterium *Hyphomonas neptunium* // *Nat. Commun. Springer US*, 2019. Vol. 10, № 1. P. 1–16.
151. Ehrle H.M., Guidry J.T., Iacovetto R., Salisbury A.K., Sandidge D.J., Bowman G.R. Polar organizing protein PopZ is required for chromosome segregation in *Agrobacterium tumefaciens* // *J. Bacteriol.* 2017. Vol. 199, № 17. P. 1–14.
152. Nolivos S., Sherratt D. The bacterial chromosome: Architecture and action of bacterial SMC and SMC-like complexes // *FEMS Microbiol. Rev.* 2014. Vol. 38, № 3. P. 380–392.
153. Livny J., Yamaichi Y., Waldor M.K. Distribution of centromere-like parS sites in bacteria: Insights from comparative genomics // *J. Bacteriol.* 2007. Vol. 189, № 23. P. 8693–8703.
154. Freedman L.P., Zengel J.M., Archer R.H., Lindahl L. Autogenous control of the S10 ribosomal protein operon of *Escherichia coli* : Genetic dissection of transcriptional and posttranscriptional regulation // *Proc. Natl. Acad. Sci.* 1987. Vol. 84, № September. P. 6516–6520.
155. Maier T., Güell M., Serrano L. Correlation of mRNA and protein in complex biological samples. // *FEBS Lett. Federation of European Biochemical Societies*, 2009. Vol. 583, № 24. P. 3966–3973.
156. Vogel C., Marcotte E.M. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses // *Nat. Rev. Genet.* 2012. Vol. 13, № 4. P. 227–232.
157. Pospíšil J., Strunin D., Ziková A., Hubálek M., Vohradský J. A comparison of protein and mRNA expression during development of the soil dwelling prokaryote (*S. coelicolor*) // *Proteomics*. 2020. P. 2000032.
158. Erdmann J., Preusse M., Khaledi A., Pich A., Häussler S. Environment-driven changes of mRNA and protein levels in *Pseudomonas aeruginosa* // *Environ. Microbiol.* 2018. Vol. 20, № 11. P. 3952–3963.

159. Cheng Z., Teo G., Krueger S., Rock T.M., Koh H.W., Choi H., Vogel C. Differential dynamics of the mammalian mRNA and protein expression response to misfolding stress // *Mol. Syst. Biol.* 2016. Vol. 12, № 1. P. 855.
160. Gingold H., Pilpel Y. Determinants of translation efficiency and accuracy // *Mol. Syst. Biol.* Nature Publishing Group, 2011. Vol. 7, № 481. P. 1–13.
161. Li J.J., Biggin M.D. Statistics requantitates the central dogma // *Science* (80-. ). 2015. Vol. 347, № 6226. P. 1066–1067.
162. Fisunov G.Y., Evsyutina D. V., Garanina I.A., Arzamasov A.A., Butenko I.O., Altukhov I.A., Nikitina A.S., Govorun V.M. Ribosome profiling reveals an adaptation strategy of reduced bacterium to acute stress // *Biochimie. Elsevier B.V.*, 2017. Vol. 132. P. 66–74.
163. Ku S., Gavin A., Aebersold R., Serrano L. Quantification of mRNA and protein and integration with protein turnover in a bacterium // *Mol. Syst. Biol.* 2011. № 511. P. 1–12.

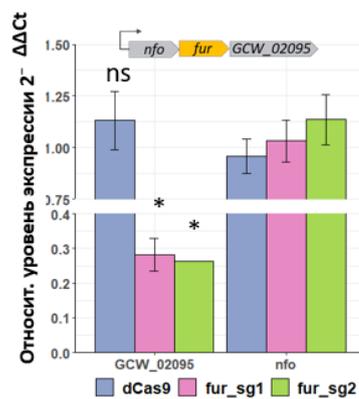
## ПРИЛОЖЕНИЕ



**Рисунок П1. Полная схема плазмиды pTn4001mini\_Recipient-MCS-M5 для сверхэкспрессии генов в микоплазмах.** На приведенной схеме плазмиды для сверхэкспрессии гена *mraZ*: *tetM* – ген устойчивости к тетрациклину, используется как селективный маркер для отбора колоний микоплазмы, содержащих плазмиду, кодонный состав оптимизирован для экспрессии в микоплазмах; *AmpR* – ген устойчивости к ампициллину; *f1 ori*, *ori* – ориджины репликации; *OIR* и *IIR* – инвертированные повторы по которым выщепляется транспозон; *transposase* – ген транспозазы, кодонный состав оптимизирован для экспрессии в микоплазмах. Промоторы, сайты связывания рибосомы и терминаторы обозначены желтым цветом; *mraZ* – целевой ген для сверхэкспрессии



**Рисунок П2.** Полная схема плазмиды pTn4001mini\_Recipient-MCS-M5-Link2-egfp для проверки влияния элементов промотора на экспрессию гена *egfp* в *Mycoplasma gallisepticum*. *tetM* – ген устойчивости к тетрациклину, используется как селективный маркер для отбора колоний микоплазмы, содержащих плазмиду, кодонный состав оптимизирован для экспрессии в микоплазмах; AmpR – ген устойчивости к ампициллину; f1 ori, ori – ориджины репликации; OIR и IIR – инвертированные повторы по которым выщепляется транспозон; transposase – ген транспозазы, кодонный состав оптимизирован для экспрессии в микоплазмах. Промоторы, сайты связывания рибосомы и терминаторы обозначены желтым цветом; *egfp* – ген зеленого флуоресцентного белка



**Рисунок П3.** Влияние нокадауна *fur* на экспрессию других участников оперона. Относительный уровень мРНК генов *M. gallisepticum*, входящих в один оперон с геном *fur*. Экспрессию генов внутри каждого образца нормализовали на уровень транскрипта *tuf*. Изменение экспрессии генов рассчитано относительно образцов мРНК *M. gallisepticum* дикого типа. Приведено среднее не менее 3 независимых экспериментов со стандартным отклонением. \* –  $p < 0,05$  согласно ранговому непарному критерию Вилкоксона. Панель сверху – структура оперона, содержащего *fur*

**Таблица П1** Последовательности синтетических олигонуклеотидов, используемых в работе

Название	Последовательность
<b>Для клонирования генов и получения чистых белков</b>	
MraZF	ATTAGGATCCATGTTTATTGGCAACSTATCAACA
MraZR	TATAGTCGACATGATCATCATCCATTGATTCAG
MraZ_correct_F	CATTGAATTGTGGGATGTCAATG
MraZ_correct_R	CATTGACATCCCAACAATTCAATG
HsdC-F	ATTAGGATCCATGTTTGATTATGCAAAGAAAATTA
HsdC-R	TATAGTCGACATCATCTAATTTTCATGCCAATCT
WhiA-F	ATTAGGATCCATGGTCACSTTTAGTCAAGAG
WhiA-R	TATAGTCGACTTAATTTTTTAATTGCTTATATAGCGTC
ParX-F	TTAAGGATCCATGGAAAATAAAAAAGCTAAAAAAAAGAACGAAG
ParX-R	TATAGTCGACTTCCTCTTTAATACTTTTATCAAGAATCATTG
<b>Для EMSA</b>	
<b>ParX</b>	
ParX-WT-F	TGGAACSTAATATTTTAAATATGAATATTATATTGTTATAATATTAATAACG
ParX-WT-RC	CGTTATTAATATTTATAACAATATAATATTCATATTTAAAATAATATTAGTTCCA
ParX-Core-F	TGGAACAAATATTTCTTTGAATATTTATTGAATATTGTTACAATATTCATAACG
ParX-Core-RC	CGTTATGAATATTGTAACAATATTCATAAAATATTCAAAGAATATTTGTTCCA
ParX-Opt-F	TGGAATAATATTTATTTATAAATATTTATAAATATTTATAAATATTTAATTAACG
ParX-Opt-RC	CGTTAATAATATTTATAAATAATTTATAAATAATTTATAAATAATATTTATTTCCA
ParX-M1-F	TGGAATATCTGTATTATAAATATTTATAAATATTTATAACATCTATTAACG
ParX-M1-RC	CGTTAATAGATGTATAAATAATTTATAAATAATTTATAATACAGATATTTCCA
ParX-M2-F	TGGAATATCTGTATTATAAATATTTATAATGGTTATTATATTCAAATTAACG
ParX-M2-RC	CGTTAATTTGAATATAAATAACCATATAAATAATTTATAATACAGATATTTCCA
ParX-M3-F	TGGAATATCTGTATTATAACATCTATTATAATGGTTATTATATTCAAATTAACG
ParX-M3-RC	CGTTAATTTGAATATAAATAACCATATAATAGATGTATAATACAGATATTTCCA
ParX-M4-F	TGGAATAACSTTATTATAACGTTATTATAAGGTTATTATAACGTTATTAACG
ParX-M4-RC	CGTTAATAACGTTATAAATAACSTTATAAATAACGTTATAAATAAGGTTATTTCCA
<b>MraZ</b>	
PmraZWT-F	AATTCAAAGTGTTAAAAAGTGTGAGAAAGTGGGAAAAAT
PmraZWT-RC	ATTTTTCCCACCTTCTCACACTTTTTAACACTTTTGAATT
Mut1-F	AATTCAAACAGTGTAAAAAGTGTGAGAAAGTGGGAAAAAT
Mut1-RC	ATTTTTCCCAGTTTCTCACAGTTTTTAACAGTTTTGAATT
Mut2-F	AATTCAAAGTCTTAAAAAGTGTGAGAAAGTCGGAAAAAT
Mut2-RC	ATTTTTCCGACTTCTCAGACTTTTTAAGACTTTTGAATT
Mut3-F	AATTCAAACAGTGTAAAAACAGTGTGAGAAACAGGAAAAAT
Mut3-RC	ATTTTTCCCTGTTTCTCACTGTTTTTAACGTTTTGAATT
Mut4-F	AATTCATCTGTGTTAATTCGTGTGAGCTTGTGGGAAAAAT
Mut4-RC	ATTTTTCCCACAAGCTCACACGAATTAACACAGATGAATT
Mut5-F	AATTCACACTGATTTAAAAAGTGTGAGTTGATCGGAAAAAT
Mut5-RC	ATTTTTCCGATCAACTCACACTTTTTAATCAGTGTGAATT
Mut6-F	AATTCACACTGATTTAAAAAGTGTGAGAAAGTGGGAAAAAT
Mut6-RC	ATTTTTCCCACCTTCTCACACTTTTTAATCAGTGTGAATT
Mut7-F	AATTCAAAGTGTTAACTATGCAGAGAAAGTGGGAAAAAT

Mut7-RC	ATTTTTCCCACCTTCTCTGCATAGTTAACACTTTTGAATT
P1290WT-F	AGTTTATTTTCATAAAAGTGTTTAAAGTGAAAATAGCTAAA
P1290WT-RC	TTTAGCTATTTTCACTTTAAACACTTTTATGAAATAAACT
N3-F	GATATACGTCTAAAAAGTGTGAGAAAGTGGGAATAAGTAC
N3-RC	GTACTTATTCCCACCTTCTCACACTTTTTAGACGTATATC
N2-F	GATATACGTCTAAAAAGTGTAGAAAGTGGGAATAAGTAGC
N2-RC	GCTACTTATTCCCACCTTCTACACTTTTTAGACGTATATC
N4-F	GATATACGTCTAAAAAGTGTGCAGAAAGTGGGAATAAGTA
N4-RC	TACTTATTCCCACCTTCTGCACACTTTTTAGACGTATATC
Neg-F	AAAACACCCTATTTTTGATATGATATAGTCATACAAAGGA
Neg-RC	TCCTTTGTATGACTATATCATATCAAAAATAGGGTGTTTT

### HsdC

hsdC-WT-F	TTATCGGCTTTGTGTTAAAATAGTGTTAACGATTTTGAAG
hsdC-WT-RC	CTTCAAATCGTTAACACTATTTTAAACACAAAGCCGATAA
hsdC-mut1-F	TTATCGGCTTTCTGTTAAAATACTGTTAACGATTTTGAAG
hsdC-mut1-RC	CTTCAAATCGTTAACAGTATTTTAAACAGAAAGCCGATAA
hsdC-mut2-F	TTATCGGCTTTGTCTTAAAATAGTCTTAAACGATTTTGAAG
hsdC-mut2-RC	CTTCAAATCGTTAAGACTATTTTAAAGACAAAGCCGATAA
hsdC-mut3-F	TTATCGGCTTTGTGTACAAATAGTGTACACGATTTTGAAG
hsdC-mut3-RC	CTTCAAATCGTGTACACTATTTGTACACAAAGCCGATAA
hsdC-mut4-F	TTATCGGCTTTGAGTTAAAATAGAGTTAACGATTTTGAAG
hsdC-mut4-RC	CTTCAAATCGTTAACTCTATTTTAACTCAAAGCCGATAA
hsdC-mut5-F	TTATCGGCTTTGTGATAAAAATAGTGATAACGATTTTGAAG
hsdC-mut5-RC	CTTCAAATCGTTATCACTATTTTATCACAAAGCCGATAA
hsdC-mut6-F	TTATCGGCTTTACCTATAAATAGTGTTAACGATTTTGAAG
hsdC-mut6-RC	CTTCAAATCGTTAACACTATTTATAGGTAAAGCCGATAA
Neg-F	AAAACACCCTATTTTTGATATGATATAGTCATACAAAGGA
Neg-RC	TCCTTTGTATGACTATATCATATCAAAAATAGGGTGTTTT
PclpB-F	TAATAGCCTAAGTGCTAATTTTTTGTATAATAAATCTAT
PclpB-RC	ATAGATTTATTATAACAAAAATAGCACTTAGGCTATTA
PtrnM-F	ATTGTTATTATATGATAATAATGTGTAACACATCGCGGGA
PtrnM-RC	TCCCGCGATGTGTTACACATTATTATCATATAATAACAAT

### WhiA

WhiA-WT-F	AAAGATTGCTGATACACCAAAAACCGTTGTCAAGTATCAG
WhiA-WT-RC	CTGATACTTGACAACGGTTTTTGGTGTATCAGCAATCTTT
WhiA-M1-F	AAAGATTGCTGATACACCAAAAACCGACTCAAGTATCAG
WhiA-M1-RC	CTGATACTTGAGTCTGGTTTTTGGTGTATCAGCAATCTTT
WhiA-M2-F	AAAGATTGCTTACTGACCAAAAACCGTTGTCAAGTATCAG
WhiA-M2-RC	CTGATACTTGACAACGGTTTTTGGTCAGTAAGCAATCTTT
WhiA-M3-F	AAAGATTGCTTACTGACCAAAAACCGACTCAAGTATCAG
WhiA-M3-RC	CTGATACTTGAGTCTGGTTTTTGGTCAGTAAGCAATCTTT
WhiA-M4-F	AAAGATTGCTCATAACCAAAAACCGTTGTCAAGTATCAG
WhiA-M4-RC	CTGATACTTGACAACGGTTTTTGGTGTATGAGCAATCTTT
WhiA-M5-F	AAAGATTGCTGTTACCAAAAACCGTTGTCAAGTATCAG
WhiA-M5-RC	CTGATACTTGACAACGGTTTTTGGTGTAAACAGCAATCTTT
WhiA-M6-F	AAAGATTGCTGAAACCAAAAACCGTTGTCAAGTATCAG

WhiA-M6-RC	CTGATACTTGACAACGGTTTTTTGGTGTTCAGCAATCTTT
WhiA-M7-F	AAAGATTGCTGATTCACCAAAAACCGTTGTCAAGTATCAG
WhiA-M7-RC	CTGATACTTGACAACGGTTTTTTGGTGAATCAGCAATCTTT
WhiA-M8-F	AAAGATTGCTGATAGACCAAAAACCGTTGTCAAGTATCAG
WhiA-M8-RC	CTGATACTTGACAACGGTTTTTTGGTCTATCAGCAATCTTT
WhiA-M9-F	AAAGATTGCTGATACTCCA AAAACCGTTGTCAAGTATCAG
WhiA-M9-RC	CTGATACTTGACAACGGTTTTTTGGAGTATCAGCAATCTTT
WhiA-M10-F	AAAGATTGCTGATACAGCAAAAACCGTTGTCAAGTATCAG
WhiA-M10-RC	CTGATACTTGACAACGGTTTTTTGCTGTATCAGCAATCTTT
WhiA-M11-F	AAAGATTGCTGATACACCAAAAACCCTTGTCAAGTATCAG
WhiA-M11-RC	CTGATACTTGACAAGGGTTTTTTGGTGTATCAGCAATCTTT
WhiA-M12-F	AAAGATTGCTGATACACCAAAAACCGATGTCAAGTATCAG
WhiA-M12-RC	CTGATACTTGACATCGGTTTTTTGGTGTATCAGCAATCTTT
WhiA-M13-F	AAAGATTGCTGATACACCAAAAACCGTAGTCAAGTATCAG
WhiA-M13-RC	CTGATACTTGACTACGGTTTTTTGGTGTATCAGCAATCTTT
WhiA-M14-F	AAAGATTGCTGATACACCAAAAACCGTTCTCAAGTATCAG
WhiA-M14-RC	CTGATACTTGAGAACGGTTTTTTGGTGTATCAGCAATCTTT
WhiA-M15-F	AAAGATTGCTGATACACCAAAAACCGTTGACAAGTATCAG
WhiA-M15-RC	CTGATACTTGTCAACGGTTTTTTGGTGTATCAGCAATCTTT
<b>Для сверхэкспрессии</b>	
mraZ_over-F	TATACTCGAGATGTTTTATTGGCAACTATCAACA
mraZ_over-R	TAATCCATGGTTAATGATCATCATCCATTGATTC
fur_over-F	TATACTCGAGATGAAGAATCATAATCAACTAGATT
fur_over-R	TAATCCATGGTTAATGATCTCATTGATCGTATGAAATG
<b>Для CRISPRi</b>	
dCas9_mut_1_F	TAGTATGGATAAGAAATACTCAATAGGACTGGCTATTGGC
dCas9_mut_1_R	AGCATCGACATCATAATCACTCAGCCTATTAATATC
dCas9_mut2_F	GATTATGATGTCGATGCTATTGTTCCACAAAGTTTCCTT
dCas9_mut2_R	GATTTTCTGCAGATGTCTTAGTCACCTCCCAGCTG
dCas9_mut_1a_F	AACTGAAAGGAGGATACTAGTATGGATAAGAAATACTCAATAGGAC
sg_1	CTAGAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTT
sg_7	CGAGTCGGTGCTTTTTTTT
sg_6	AAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCAC
sg_2	GATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCT
whiAtg1_3	AAAACCGACTTGTTGGTGCAAAATGTGAAATATTATAACATAGGGCC
whiAtg1_4	CTATGTTATAATATTTTCACATTTTGAC
whiAtg1_5	CAACAAGTCGGTTTTAGAGCTAGAAATAGCAAGTT
whiAtg2_3	AAAACACACCCAAAAACGATCATCTTGAAATATTATAACATAGGGCC
whiAtg2_4	CTATGTTATAATATTTCAAGATGATCGT
whiAtg2_5	TTTTGGGTGTGTTTTAGAGCTAGAAATAGCAAGTT
furtg1a_5	TTTGGTGATTGTTTTAGAGCTAGAAATAGCAAGTT
furtg1a_3	AAAACAATCACCAACCTAGAAGGTTGAAATATTATAACATAGGGCC
furtg1a_4	CTATGTTATAATATTTCAACCTTCTAGG
furtg2a_5	AAGATTGAGTGTTTTAGAGCTAGAAATAGCAAGTT
furtg2a_3	AAAACACTCAATCTTGTTGATTCGTTGAAATATTATAACATAGGGCC

furtg2a_4	СТАТГТТАТААТАТТТСААСГААТСААС
Для вектора с egfr	
hsdC-WT-F	СТТТГТГТТААААТАГТГТТААСГГСАТГ
hsdC-WT-RC	СГТТААСАСТАТТТТААСАСАААГГГГСС
hsdC-Mut-F	СТТТГТГТТААААТАГАСТТААСГГСАТГ
hsdC-Mut-RC	СГТТААГТСТАТТТТААСАСАААГГГГСС
Mini-1-F	АТТГАСАТАСАТГАТТСАТТАТГТТТААТАТГСАТГСАСАТГ
Mini-1-RC	ТГСАТГСАТАТТАТААСАТААТГААТСАТГТАТГТСААТГГГСС
Mini-2-F	АСГАТТСТАСАТГАТТСАТТАТГТТТААТАТГСАТГСАСАТГ
Mini-2-RC	ТГСАТГСАТАТТАТААСАТААТГААТСАТГТАГААТСГТГГСС
Mini-6-F	АТТГАСАТАСАТГАТТСАТАТТАТААТАТГСАТГСАСАТГ
Mini-6-RC	ТГСАТГСАТАТТАТААТГАТАТГААТСАТГТАТГТСААТГГГСС
Mini-7-F	АТТГСАТТАСАТГАТТСАТАТТАТААТАТГСАТГСАСАТГ
Mini-7-RC	ТГСАТГСАТАТТАТААТГАТАТГААТСАТГТААТГСААТГГГСС
uvrB-1-F	ТТААГГГСССГАААТТТСАГАТААГТАГТСАААТТС
uvrB-1-R	ТАТАГСАТГСАТТТТТАТСТАТТСАТТАСАТТАГ
uvrB-2-R	ТАТАГСАТГСАТТАГТТСАТТАТААААТАГТААГА
uvrB-2-F	ТТААГГГСССГТТТАТСТАТТААТГАСАТААТТТСТАТА
uvrB-3-F	ТТААГГГССССАГАТАСТТТТТТГАААТТТСАГАТААГ
uvrB-3-R	ТАТАГСАТГСАТТТАААТТТТТАТСТАТТАСАТТАТТ
parA-1-F	САТТГТТТГТТТТСТААСАТАТААСТАГАТГ
parA-1-R	САААТГАААТГАТСАТААТТТАТСАТТАГГСС
parA-mini-F	СТАТАТГТТТААТААТААТААТААСТАГАТГСАТГ
parA-mini-RC	САТСАТТАТТААТАТТАТААСАТАТАГГСС
parA-M1-mini-F	СТАТАТГТТТААТААТАТАТААСТАГАТГСАТГ
parA-M1-mini-RC	САТСАТТАТАТАТАТТАТААСАТАТАГГСС
rplJ-mini-F	СААГААТГАТАТАСТАТТАААГААТАГАТГ
rplJ-mini-R	СТАТТГСТТТААТАГАТАТАТТАТТТТГГСС
rplJ-1-F	ТТААГГГСССАГАТГААГАТАГСТААТТССАТААГ
rplJ-1-R	ТАТАГСАТГСАТТТТГГТТАТТГСТТТААТАГТ
rplJ-2-R	ТАТАГСАТГСАТТАТТТААААТТААТАГАТАААГТГ
rplJ-2-F	ТТААГГГСССААГАТАААГААТГАТАТАСТАТТАААГА
rplJ-3-R	ТАТАГСАТГСТАГАТАТГГТТАТСТААССТАСА
Для количественной ОТ-ПЦР	
MGA_23S-F	АГАААТАСАТТАГАТГСАТГГААГ
MGA_23S-R	ГАГТТССТТАГАСТАТАГАТТТАСТАТ
MGA_eno-F	ГСАТТСТАГАСАТТАСААТГГ
MGA_eno-R	АССТТАСАТТАССТТТГТТАГАТА
MGA_gapd-F	АСГАСТТААСТАГАТГСТААГАС
MGA_gapd-R	ТГГТТАСАСАТТГТАААСТАСАГ
MGA_tuf-F	ТСААААТСАТТАСАСТАТТАСАСАС
MGA_tuf-R	ТСТАГАТТАСТАГГАТГТТАТТАСС
MGA_mraZ-F	АТГТТТАТТГГСААСТАТТАСАСА
MGA_mraZ-R	АТГАТТАСАТТАСАТТАГАТТТАСА
MGA_mraW-F	ТСАГАТТАТТААСТАГАТТАСАГ

MGA_mraW-R	GGAAATATCTTCTTGCAGGATG
MGA_ftsA-F	GGCTGATCAAAACTATTATGCAG
MGA_ftsA-R	TGATCAGATCTTCGTTACGGT
MGA_whiA-F	CAAGAGGTTAAAAGTGAGATCTGTG
MGA_whiA-R	GCTGTTGTTCTTCTAGTTCGTG
MGA_rpsJ-F	TTAAACAAGCCGCAATTAGGTC
MGA_rpsJ-R	TACCCTTCATAACTCTTTGAGC
MGA_hsdC-F	TATTCGATCCAGCTTGTGGTAG
MGA_hsdC-R	CATAATCAGTCCATCTAGGGTCA
MGA_uvrB-F	TGTGTTGTCGGAATTAACCT
MGA_uvrB-R	ATCGCTTTAGTCATCTCATCAG
MGA_parA-F	ATCATTTTCAATTTATTAACAATAAAGGT
MGA_parA-R	TTTCCATTGTTCTTGTTATCC
MGA_rplJ-F	ACAGCAAAGCTATTACAGCAC
MGA_rplJ-R	CATCTCTTGATGGAAGAACAGC
MGA_zip_F	CACCTTAAGTAGTTACACCGAG
MGA_zip_R	GGGTTAATAAAGAAACCGTTCC
MGA_rpsF_F	ATTATGCTAATGGTAAGCGGTC
MGA_rpsF_R	TCTTTGCATTTTGGCTAGCTC
MGA_GCW_02500_F	CTGCTTTAGAAGTTGGTTATCGTC
MGA_GCW_02500_R	CCTTCAAGATAAAACTCTTCCATCG
MGA_GCW_01865_F	ATAAGAACAAAGTTCACGGTGG
MGA_GCW_01865_R	AATCCTGCATAGATTGCTGA
MGA_GCW_00535_F	AATGATAACTCCTGCTGTTCC
MGA_GCW_00535_R	CATAAAGCTCATCGTAAGAACC
MGA_cbiO_F	ACCTTGATTACGTAGTTGAAGC
MGA_cbiO_R	CAGATAACCAGTTCATTGGGAT
MGA_GCW_00170_F	GGTGTATTGGAATTGACTCTC
MGA_GCW_00170_R	ACAAAGCTTTGGTAGTAACCTG
MGA_glpF_F	CAGCTAAAAGAATCATAGTGCC
MGA_glpF_R	GTGGTGTGGTTATTAGTATCCAAG
MGA_udk_F	CACTAGATCAGATCAGTTTACTGAG
MGA_udk_R	CAAGGTAAGATCAGATTAGCGT
egfp-F	GACCCTGAAGTTCATCTGCACC
egfp-R	TAGTTGTACTCCAGCTTGTGCC

**Таблица  
П2**

		Значения RPKM										Суммарная цитоплазматическая фракция мРНК		Рибосомально- связанная фракция мРНК	
		Суммарная цитоплазматическая фракция мРНК				Рибосомально-связанная фракция мРНК									
		Контроль		Тепловой стресс		Контроль			Тепловой стресс						
TAG	Название гена	с.1	с.2	hs30.1	hs30.2	Ribo с.1	Ribo с.2	Ribo с.3	Ribo hs30.1	Ribo hs30.2	Ribo hs30.3	q-value	log2FC	q-value	log2F C
GCW_00005	parA	83,4	105,4	86,4	92,9	183,9	146,5	142,6	123,8	102,7	164,7	6,26E-01	-0,06	3,84E-01	-0,25
GCW_00010	dnaA	54,9	74,3	91,3	92,3	40,1	60,6	42,9	107,3	109,8	123,6	4,04E-02	0,52	4,48E-15	1,31
GCW_00015	GCW_00015	0,0	2,5	17,1	18,1	0,0	0,0	0,0	0,0	0,0	0,0	2,07E-07	5,68	1,00E+00	0,00
GCW_00030	rpmH	1310,7	1018,4	269,9	270,3	1504,8	1402,7	334,8	544,8	561,1	673,3	7,51E-22	-2,10	1,44E-01	-0,55
GCW_00035	rnpA	495,5	300,2	138,3	102,8	274,7	230,2	131,2	432,9	470,4	451,9	3,90E-04	-1,69	3,43E-06	1,20
GCW_00040	yidC	801,1	933,2	326,6	359,3	754,9	775,3	851,1	506,4	512,3	480,4	5,66E-13	-1,34	9,51E-22	-0,63
GCW_00045	ksgA	133,9	106,5	93,6	80,2	66,3	113,4	56,2	202,2	230,8	145,9	9,15E-03	-0,46	7,87E-04	1,38
GCW_00050	ispE	48,1	50,9	76,5	83,6	21,9	39,8	10,7	54,0	73,3	31,9	1,07E-07	0,69	7,41E-02	1,29
GCW_00055	argS	158,7	203,8	125,7	124,5	309,7	297,5	377,6	281,3	246,1	283,8	1,38E-02	-0,52	6,93E-02	-0,24
GCW_00060	mnuA	173,6	229,9	157,6	153,4	310,0	372,7	183,5	287,1	244,6	293,0	9,95E-02	-0,36	8,27E-01	0,02
GCW_00065	glpF	116,0	168,8	127,4	165,5	51,7	29,1	40,4	83,7	98,8	64,7	9,47E-01	0,05	1,65E-03	1,09
GCW_00075	glpO	346,7	442,9	89,9	103,4	442,3	221,1	405,7	177,7	179,2	194,7	4,24E-12	-2,02	1,02E-02	-0,85
GCW_00080	GCW_00080	408,5	497,8	205,4	227,7	362,6	442,7	433,6	516,2	461,7	787,8	7,28E-07	-1,06	1,34E-01	0,52
GCW_00085	GCW_00085	747,9	717,1	2037,1	1980,5	901,4	480,2	700,2	474,6	542,6	829,7	0,00E+00	1,46	7,98E-01	-0,13
GCW_00095	GCW_00095	417,2	561,7	91,2	106,0	586,8	568,1	740,6	605,4	576,6	560,5	8,68E-11	-2,30	5,96E-01	-0,07
GCW_00100	GCW_00100	5,7	9,3	12,5	13,9	3,2	0,0	0,0	0,0	1,0	0,0	3,85E-02	0,86	6,03E-01	-0,53
GCW_00110	GCW_00110	104,9	116,6	220,7	225,4	36,5	49,5	34,4	81,9	103,0	63,8	9,93E-22	1,01	1,36E-03	1,08
GCW_00115	GCW_00115	193,6	209,2	265,8	283,6	105,4	104,7	80,9	137,0	160,7	89,9	2,76E-04	0,45	1,89E-01	0,42
GCW_00120	GCW_00120	431,7	675,9	574,4	590,5	135,5	188,9	222,7	355,1	277,0	200,2	9,51E-01	0,11	8,48E-02	0,64
GCW_00125	GCW_00125	110,6	104,6	377,6	356,4	19,7	34,7	42,9	117,4	130,1	57,0	4,01E-143	1,77	3,89E-03	1,67
GCW_00135	hpt	758,3	605,4	474,7	494,2	1203,8	1350,3	1507,6	1242,5	1233,3	1250,5	2,14E-03	-0,48	3,67E-01	-0,08
GCW_00140	ftsH	1285,8	1168,9	735,4	811,7	2020,1	1908,0	3071,8	1536,3	1654,3	1605,5	6,47E-11	-0,67	5,79E-02	-0,47
GCW_00145	GCW_00145	200,8	169,5	140,3	100,4	42,6	36,0	75,7	183,0	209,2	134,8	8,37E-03	-0,64	6,57E-06	1,86
GCW_00150	GCW_00150	103,6	73,8	132,6	96,7	18,6	28,8	41,8	85,1	105,8	60,7	1,57E-01	0,37	5,63E-04	1,58
GCW_00155	fruB	1452,5	615,6	199,4	174,5	3586,9	2407,6	1339,6	504,7	401,9	1096,9	7,10E-03	-2,34	4,85E-03	-1,86

GCW_00160	tktA_1	624,0	653,6	363,8	372,6	924,8	715,8	885,8	688,2	658,4	633,9	7,79E-23	-0,79	3,05E-02	-0,30
GCW_00165	secA	607,3	697,7	424,4	486,5	646,9	689,0	980,1	865,5	864,7	838,5	3,29E-03	-0,52	3,04E-01	0,21
GCW_00170	GCW_00170	2253,9	1792,0	325,9	336,6	4600,9	4915,8	8136,6	1158,7	872,8	1577,9	2,21E-34	-2,60	1,38E-07	-2,24
GCW_00175	GCW_00175	923,4	792,6	82,1	78,2	243,6	198,8	205,7	285,5	281,0	195,4	5,36E-129	-3,42	2,78E-01	0,26
GCW_00180	ugpC	776,1	734,0	109,2	110,4	220,2	230,0	439,7	151,1	142,2	125,0	0,00E+00	-2,78	1,31E-02	-0,98
GCW_00185	ugpA	589,9	588,5	112,8	121,6	335,9	266,7	387,6	167,1	190,2	142,5	5,17E-247	-2,33	7,25E-05	-0,94
GCW_00190	ugpE	548,0	414,8	115,2	109,8	347,4	395,5	478,9	170,5	198,2	160,9	1,39E-17	-2,08	2,16E-10	-1,16
GCW_00195	GCW_00195	353,8	306,7	69,0	64,7	240,9	175,3	277,3	96,8	96,8	89,2	4,88E-101	-2,30	2,39E-07	-1,23
GCW_00200	GCW_00200	383,3	332,7	102,9	89,8	322,7	266,1	270,8	104,1	103,2	118,0	1,40E-64	-1,89	1,41E-23	-1,36
GCW_00205	uvrB	113,5	152,5	25,5	30,6	90,6	134,2	196,2	167,2	119,3	180,0	2,60E-10	-2,23	7,04E-01	0,24
GCW_00210	pstS	134,9	156,6	36,1	35,2	102,8	83,1	89,9	60,9	75,9	64,3	2,12E-27	-2,03	2,64E-02	-0,42
GCW_00215	pstC/pstA	144,2	152,8	26,6	30,7	129,0	238,5	218,9	141,3	164,3	132,9	7,07E-68	-2,37	2,12E-01	-0,34
GCW_00220	pstB	333,4	417,8	75,0	87,8	656,7	887,0	994,9	563,4	498,8	504,1	9,30E-15	-2,20	2,17E-04	-0,64
GCW_00225	phoU	237,2	252,4	50,2	52,1	540,4	662,1	950,5	516,0	490,5	493,4	4,07E-90	-2,26	6,63E-02	-0,44
GCW_00230	nrdA	532,1	642,2	390,7	441,8	848,8	1091,7	902,3	699,4	816,0	1165,9	1,24E-02	-0,49	8,83E-01	-0,07
GCW_00235	nrdI	689,9	981,1	242,6	330,2	830,0	973,2	597,9	541,4	451,4	674,2	9,48E-05	-1,54	6,55E-02	-0,48
GCW_00240	nrdF	705,4	738,7	155,5	166,5	2009,4	1595,9	1655,1	581,0	453,6	859,6	1,34E-97	-2,16	1,05E-06	-1,48
GCW_00245	thyA	119,8	201,1	62,7	66,7	425,1	196,5	169,6	193,7	168,7	229,9	5,18E-03	-1,26	4,81E-01	-0,27
GCW_00260	GCW_00260	30,4	24,4	127,3	115,8	6,5	18,4	15,4	33,3	57,6	18,9	3,36E-87	2,15	5,27E-02	1,48
GCW_00265	GCW_00265	54,9	29,1	93,5	71,7	10,6	15,1	10,3	45,0	61,0	43,1	1,11E-02	1,03	3,78E-10	2,09
GCW_00270	GCW_00270	149,0	192,9	206,2	178,3	102,6	78,7	109,5	69,0	77,1	75,0	5,03E-01	0,18	5,66E-02	-0,34
GCW_00275	rpsJ	2917,9	2088,5	262,6	208,3	1953,6	2254,5	1168,3	1354,3	1832,8	1330,3	7,35E-19	-3,40	6,11E-01	-0,17
GCW_00280	rplC	2326,4	1555,1	347,2	352,9	2472,7	1978,1	1335,0	1316,6	1504,9	1181,4	7,60E-10	-2,44	1,25E-01	-0,45
GCW_00285	rplD	1581,5	1157,8	291,5	295,0	1961,6	1308,8	1103,4	873,3	953,0	851,6	1,08E-14	-2,21	3,43E-02	-0,63
GCW_00290	rplW	1807,7	1543,6	344,2	312,3	1871,9	1539,5	1555,0	674,8	652,1	689,5	8,15E-76	-2,35	1,01E-22	-1,26
GCW_00295	rplB	1686,5	1542,4	416,7	458,3	2684,9	2082,0	1863,4	1015,0	917,3	947,3	1,54E-105	-1,88	1,11E-08	-1,15

GCW_00300	rpsS	1373,4	1225,8	311,6	304,8	1733,1	1535,7	1188,4	964,5	1105,4	884,7	2,64E-171	-2,07	1,25E-02	-0,54
GCW_00305	rplV	1262,0	1072,9	292,6	257,2	1923,9	1322,8	1190,4	820,0	836,7	688,1	2,33E-56	-2,08	2,28E-03	-0,85
GCW_00310	rpsC	1860,0	1554,1	487,7	481,1	2652,8	2128,6	3960,7	1173,1	976,5	901,8	1,96E-40	-1,81	3,04E-05	-1,44
GCW_00315	rplP	1697,4	1509,2	490,9	565,2	2927,5	2543,1	2205,6	1323,5	1029,3	973,0	5,19E-38	-1,60	1,04E-09	-1,17
GCW_00320	rpmC	1132,9	1354,6	293,2	301,8	1954,5	1338,2	1854,7	978,8	822,9	750,6	5,05E-21	-2,06	2,60E-05	-0,96
GCW_00325	rpsQ	1775,7	1328,7	445,6	467,8	3249,2	2351,2	1960,4	1415,8	1070,6	1089,3	1,37E-12	-1,75	3,17E-04	-1,02
GCW_00330	rplN	1776,0	1623,4	479,1	710,4	3595,9	2525,1	2782,3	1363,0	1013,1	910,1	1,09E-09	-1,54	8,39E-08	-1,40
GCW_00335	rplX	1630,6	1736,3	439,3	486,9	2470,7	1663,9	1907,7	1141,8	805,1	802,4	2,97E-50	-1,86	4,80E-05	-1,10
GCW_00340	rplE	1698,7	1487,6	532,8	515,0	2762,2	1964,9	2501,0	1346,3	1046,4	970,0	4,00E-77	-1,60	2,84E-06	-1,06
GCW_00345	rpsN	1976,6	1898,6	633,2	645,8	4008,4	3959,5	2693,5	2005,4	1401,1	1425,6	0,00E+00	-1,60	1,16E-05	-1,10
GCW_00350	rpsH	1451,8	873,5	425,4	420,6	2164,6	1487,2	1207,2	1414,1	1340,5	1002,0	1,12E-03	-1,41	3,40E-01	-0,31
GCW_00355	rplF	2003,1	1796,1	669,6	806,9	1818,4	1629,9	2072,8	1673,2	1391,1	1118,9	8,65E-21	-1,37	9,52E-02	-0,37
GCW_00360	rplR	1648,6	1183,3	578,5	660,1	1929,8	1947,3	1651,4	1303,2	965,9	922,2	2,45E-05	-1,18	1,45E-05	-0,77
GCW_00365	rpsE	1951,8	2146,9	812,4	942,4	2593,3	2173,1	2376,9	1544,6	1096,4	1190,0	2,81E-14	-1,23	9,30E-07	-0,87
GCW_00370	rplO	1857,2	1407,4	887,2	952,1	1698,6	1550,9	827,6	1501,1	1288,6	1122,4	1,14E-04	-0,81	8,87E-01	0,04
GCW_00375	secY	1162,2	1048,0	846,5	931,1	679,2	721,8	1384,1	930,7	868,8	607,4	3,42E-03	-0,31	7,41E-01	-0,12
GCW_00380	adk	616,7	466,4	598,3	529,7	322,1	482,5	503,9	526,9	579,1	403,7	6,38E-01	0,07	3,51E-01	0,26
GCW_00385	map	711,1	616,6	750,7	743,7	418,1	491,0	925,7	495,9	471,0	303,8	2,09E-02	0,17	3,40E-01	-0,43
GCW_00390	ldh	4840,7	2007,2	1174,5	1305,8	5785,9	7291,2	17482,2	4593,8	7223,4	4615,1	7,11E-02	-1,33	2,57E-01	-0,72
GCW_00405	GCW_00405	6,3	7,0	23,5	19,1	1,6	1,3	0,6	3,3	3,0	1,7	4,55E-21	1,67	1,61E-02	1,26
GCW_00410	lysS	357,8	400,3	29,7	37,8	589,1	665,0	936,8	524,1	430,6	555,5	5,57E-56	-3,49	4,21E-02	-0,48
GCW_00415	dnaJ	299,5	426,9	65,0	86,2	154,2	141,4	182,5	203,6	156,8	151,4	1,06E-07	-2,25	5,96E-01	0,13
GCW_00420	GCW_00420	232,9	239,9	230,5	244,5	628,6	692,1	352,4	649,5	660,9	772,1	9,30E-01	0,01	2,25E-01	0,41
GCW_00430	GCW_00430	27,7	18,3	14,9	10,3	41,9	61,9	48,2	29,7	28,3	37,1	3,39E-02	-0,86	2,06E-03	-0,63
GCW_00435	ptsA	1429,1	1377,2	904,4	1142,1	2395,8	1990,7	3399,1	1231,4	1202,0	1412,4	1,29E-02	-0,46	5,81E-04	-0,95
GCW_00440	GCW_00440	214,0	203,9	309,8	296,1	87,9	64,4	56,8	305,8	309,7	316,8	4,03E-46	0,54	2,16E-59	2,22
GCW_00445	rpsO	1972,8	769,5	5492,0	3103,1	1447,3	852,5	1560,2	667,0	805,4	1153,3	1,10E-02	1,75	2,19E-01	-0,51

GCW_00450	GCW_00450	236,7	301,7	214,8	215,6	237,8	222,7	232,5	294,2	249,5	269,6	1,06E-01	-0,31	6,04E-03	0,27
GCW_00455	ftsK	277,1	269,5	402,3	357,4	114,7	174,8	123,7	226,2	253,9	209,7	4,70E-15	0,47	1,56E-04	0,80
GCW_00460	trxA_1	564,0	615,2	795,3	893,0	1007,5	798,2	1296,2	468,1	396,3	386,9	1,62E-03	0,52	1,08E-06	-1,25
GCW_00465	GCW_00465	83,8	75,8	342,8	271,9	31,1	52,2	47,4	88,4	106,8	77,7	1,01E-23	1,94	4,49E-05	1,12
GCW_00470	uvrA	195,6	260,3	543,9	568,6	211,5	209,9	225,5	264,1	248,4	216,6	4,48E-10	1,30	7,04E-02	0,21
GCW_00475	tsf	2254,3	1675,3	1299,2	1220,5	3045,4	3397,5	5371,5	1697,3	1740,9	2491,3	3,31E-03	-0,63	7,45E-03	-0,93
GCW_00480	pyrH	1522,4	1255,3	1190,5	1163,7	810,3	573,3	827,9	1437,0	1671,0	1510,0	5,48E-02	-0,23	2,16E-10	1,12
GCW_00485	frr	814,8	738,7	794,0	774,7	1085,0	441,3	1054,5	571,0	816,7	622,4	6,83E-01	0,02	6,09E-01	-0,23
GCW_00490	cdsA	266,6	285,3	281,2	257,0	98,8	198,9	172,1	246,1	340,1	233,5	6,45E-01	-0,04	1,46E-02	0,88
GCW_00495	dxr	141,1	143,3	275,2	232,8	34,3	52,1	56,8	138,0	205,3	119,1	4,55E-14	0,83	1,01E-04	1,72
GCW_00500	GCW_00500	249,7	326,8	500,3	553,5	331,0	237,8	322,0	191,1	206,4	215,2	2,70E-04	0,88	8,05E-03	-0,49
GCW_00510	GCW_00510	210,0	236,1	549,6	542,3	166,5	241,6	249,6	282,8	297,1	274,9	4,31E-43	1,29	7,80E-03	0,44
GCW_00515	GCW_00515	17,5	21,2	134,5	128,2	16,5	21,9	39,5	60,2	64,2	67,3	1,82E-139	2,77	6,56E-06	1,43
GCW_00520	GCW_00520	236,7	191,1	560,7	578,3	409,2	524,7	271,0	356,6	436,7	540,6	8,38E-33	1,42	5,96E-01	0,22
GCW_00525	GCW_00525	351,8	307,9	457,8	429,6	543,1	477,2	211,8	466,3	494,0	500,5	6,70E-12	0,43	4,51E-01	0,40
GCW_00535	GCW_00535	28,1	47,6	298,7	331,2	36,4	26,5	32,0	20,4	16,0	12,5	6,32E-25	3,10	4,16E-04	-0,93
GCW_00545	GCW_00545	71,8	76,8	146,4	138,8	132,1	138,4	71,6	134,7	111,7	132,4	2,76E-46	0,94	6,26E-01	0,24
GCW_00550	GCW_00550	65,1	80,8	154,4	159,8	275,7	261,5	85,6	267,8	186,3	227,2	5,97E-11	1,11	8,37E-01	0,33
GCW_00555	GCW_00555	98,8	131,8	209,3	208,5	246,3	287,3	217,3	334,9	245,5	296,3	4,98E-05	0,87	2,89E-01	0,26
GCW_00565	GCW_00565	105,0	106,9	186,4	162,9	276,5	333,0	184,5	370,2	242,4	346,6	1,67E-17	0,72	4,93E-01	0,33
GCW_00570	GCW_00570	57,5	86,0	214,5	216,0	50,1	78,2	25,3	146,8	128,8	113,9	1,17E-10	1,61	2,45E-04	1,52
GCW_00575	GCW_00575	9,1	12,1	147,8	163,3	1,6	0,0	0,0	7,6	15,2	7,1	3,43E-61	3,88	4,37E-04	6,74
GCW_00580	GCW_00580	782,7	679,2	1237,4	1207,6	937,6	805,4	1236,1	528,8	453,9	533,9	1,03E-30	0,75	3,30E-05	-0,92
GCW_00585	GCW_00585	460,2	425,8	2546,3	2183,4	427,4	419,7	191,2	508,0	536,8	606,1	5,01E-84	2,41	2,04E-02	0,80
GCW_00590	GCW_00590	674,2	443,3	2884,7	2274,8	2056,2	2366,7	1653,3	1327,2	1579,8	1427,8	6,08E-16	2,23	1,47E-02	-0,44
GCW_00595	GCW_00595	1143,1	790,3	2653,5	2100,7	2663,7	2376,6	4039,8	1285,2	1149,7	1092,7	1,98E-07	1,31	1,39E-05	-1,29
GCW_00600	GCW_00600	1043,4	690,8	2942,0	2348,6	3307,6	3497,6	5602,4	1569,4	1750,2	1365,5	6,99E-10	1,63	4,24E-05	-1,33
GCW_00605	nusA	592,0	685,2	380,8	423,4	1006,0	881,5	704,4	704,3	599,5	644,2	1,04E-04	-0,66	2,80E-02	-0,36

GCW_00610	GCW_00610	258,8	282,3	163,1	177,8	488,3	354,1	72,8	656,8	713,8	525,3	2,09E-07	-0,67	3,95E-02	1,47
GCW_00615	infB	490,0	659,1	327,7	357,9	1220,9	1079,1	923,9	1017,0	881,6	721,9	5,97E-03	-0,73	2,06E-01	-0,26
GCW_00620	rbfA	234,6	235,3	178,9	125,8	362,3	374,7	251,0	743,8	778,3	515,7	3,85E-03	-0,65	4,36E-04	1,08
GCW_00625	GCW_00625	96,8	165,5	84,8	100,1	40,1	109,9	177,6	43,7	40,6	98,4	2,17E-01	-0,46	3,03E-01	-0,68
GCW_00630	GCW_00630	23,2	52,6	79,5	145,1	18,4	12,5	15,5	19,8	22,8	10,9	6,60E-02	1,62	5,44E-01	0,19
GCW_00640	GCW_00640	934,5	552,9	427,3	419,7	1016,9	1253,0	1274,9	443,7	431,9	575,4	5,90E-02	-0,76	1,18E-15	-1,25
GCW_00650	GCW_00650	508,9	412,7	191,2	171,1	606,8	470,0	486,5	199,3	174,8	299,1	3,46E-17	-1,34	3,20E-05	-1,21
GCW_00655	truB	202,5	244,3	225,0	226,3	50,4	34,4	51,4	123,2	120,3	104,6	9,51E-01	0,02	6,00E-14	1,41
GCW_00660	ribF/trmU	59,2	65,2	200,7	187,5	22,5	44,6	31,1	53,3	57,6	40,0	8,35E-100	1,64	4,88E-02	0,70
GCW_00665	ach1	72,5	110,6	104,6	105,5	144,3	128,3	136,5	89,6	67,6	78,6	6,43E-01	0,23	1,08E-08	-0,76
GCW_00670	alaS	190,9	252,1	66,5	76,5	464,1	577,1	502,3	478,8	383,0	420,8	9,24E-08	-1,62	5,50E-02	-0,23
GCW_00675	GCW_00675	145,9	234,6	49,0	54,5	112,0	132,9	119,6	395,0	365,1	341,3	1,07E-04	-1,84	3,70E-56	1,64
GCW_00685	GCW_00685	108,0	141,8	56,8	47,1	160,2	144,4	99,0	202,3	183,2	188,8	7,96E-06	-1,26	5,20E-03	0,58
GCW_00690	GCW_00690	157,7	212,7	71,3	71,4	646,0	683,3	770,7	474,9	335,9	462,6	3,31E-06	-1,36	6,92E-05	-0,70
GCW_00695	rpsT	3265,6	1963,7	4637,6	3293,1	2108,3	988,6	839,3	808,7	1058,4	1308,4	8,86E-02	0,63	6,68E-01	-0,18
GCW_00700	GCW_00700	302,7	382,1	126,7	164,2	416,0	336,4	448,3	211,1	224,8	289,4	1,97E-05	-1,24	9,06E-04	-0,69
GCW_00705	fmt	128,7	149,3	29,4	32,9	76,1	104,1	129,9	90,7	72,3	76,9	5,05E-26	-2,16	1,62E-01	-0,30
GCW_00710	GCW_00710	69,7	63,3	24,5	23,4	61,1	90,2	64,3	79,5	72,9	75,6	4,83E-192	-1,47	6,03E-01	0,14
GCW_00715	GCW_00715	181,7	217,0	32,5	27,2	537,2	224,4	427,3	102,8	68,9	183,8	6,04E-28	-2,74	4,70E-03	-1,73
GCW_00720	GCW_00720	115,4	130,8	48,6	47,9	86,9	100,8	128,4	105,5	83,8	85,4	2,72E-20	-1,35	4,06E-01	-0,15
GCW_00725	GCW_00725	463,2	525,3	131,9	130,7	664,2	330,8	532,4	454,9	395,6	374,5	2,51E-32	-1,91	4,32E-01	-0,23
GCW_00760	GCW_00760	3738,8	1609,1	382,8	385,1	372,5	376,9	320,1	884,2	1367,8	809,0	2,95E-03	-2,67	1,92E-04	1,52
GCW_00765	ptsG_1	4361,6	2462,2	842,0	958,3	3281,6	3589,5	6301,0	2034,3	2368,2	2207,9	4,99E-04	-1,87	1,27E-02	-0,90
GCW_00770	putA	734,8	815,1	403,6	548,1	1191,3	1300,0	3523,1	936,9	830,6	1349,4	4,15E-03	-0,72	2,00E-01	-0,75
GCW_00775	rpmE	5439,0	3145,4	8865,0	6356,2	2858,9	1615,1	2022,2	2400,1	3558,1	3533,4	2,62E-02	0,86	5,30E-02	0,60
GCW_00780	prfA	184,9	279,7	395,9	611,5	335,9	368,2	419,8	293,4	235,8	325,8	4,65E-02	1,11	3,71E-02	-0,36
GCW_00785	hemK	39,8	34,4	44,4	49,4	17,0	22,6	19,6	66,2	69,5	59,8	1,78E-02	0,34	3,76E-36	1,77

GCW_00790	GCW_00790	26,2	24,9	37,3	37,0	19,9	28,3	20,8	44,7	41,8	41,0	7,98E-07	0,54	5,05E-11	0,94
GCW_00795	rpiB	1041,4	771,4	742,0	914,5	1004,4	536,5	862,3	526,9	536,9	821,0	6,43E-01	-0,12	4,64E-01	-0,29
GCW_00800	GCW_00800	525,2	497,2	635,2	664,0	750,2	471,5	1379,4	456,5	430,8	667,8	1,02E-06	0,35	2,39E-01	-0,59
GCW_00805	recU	31,5	17,9	61,9	43,1	33,9	67,9	26,6	50,6	55,3	68,8	6,24E-03	1,12	2,98E-01	0,59
GCW_00810	hup_1	785,5	1026,2	292,8	321,7	1219,6	1061,3	1508,0	364,2	244,3	522,2	4,99E-08	-1,55	4,72E-07	-1,76
GCW_00815	GCW_00815	352,5	219,7	351,1	265,4	220,0	111,0	78,7	80,1	107,1	86,4	6,43E-01	0,13	3,69E-01	-0,42
GCW_00820	GCW_00820	348,3	423,7	522,0	641,7	124,6	112,1	124,6	81,1	106,5	81,2	3,66E-02	0,59	7,19E-02	-0,40
GCW_00825	rplJ	2188,5	1222,0	745,1	792,4	1925,5	888,8	910,1	599,3	746,0	750,0	2,22E-02	-1,09	1,25E-01	-0,70
GCW_00830	rplL	1848,0	1107,6	835,6	986,6	7475,1	5353,9	5576,7	1080,1	1023,4	1478,4	1,04E-01	-0,66	1,45E-17	-2,33
GCW_00835	rpmF	2674,5	1604,0	885,8	1028,6	1869,9	1236,9	1125,6	1262,1	1249,6	1529,9	9,25E-03	-1,12	9,16E-01	0,00
GCW_00840	GCW_00840	1704,2	903,0	619,6	614,4	1238,8	1017,9	473,8	878,5	937,7	1213,6	4,84E-02	-1,01	6,74E-01	0,29
GCW_00845	dnaJ_2	473,0	602,3	917,7	981,7	728,9	859,5	940,8	1096,6	1095,4	1712,4	7,81E-05	0,83	4,06E-02	0,64
GCW_00850	GCW_00850	2204,5	1858,6	2299,4	2190,8	1521,9	1493,0	1910,4	1106,3	940,2	1012,6	9,16E-02	0,15	4,55E-06	-0,64
GCW_00855	GCW_00855	831,1	695,5	1435,9	1391,9	405,7	597,0	635,3	783,3	991,5	775,8	1,69E-26	0,90	3,26E-03	0,70
GCW_00860	metG_1	282,5	296,6	980,2	892,6	344,1	403,4	416,8	384,5	460,0	562,4	3,26E-132	1,69	1,04E-01	0,30
GCW_00865	dnaJ_3	303,6	343,3	1184,1	1371,7	301,7	306,3	271,3	442,3	498,7	568,1	1,14E-19	1,98	1,43E-09	0,81
GCW_00875	GCW_00875	1226,7	1017,0	472,8	496,2	1405,3	1835,7	1376,0	1198,7	1038,5	1279,0	7,94E-18	-1,20	2,34E-02	-0,35
GCW_00885	metG_2	447,1	501,1	538,2	537,0	992,2	783,5	911,2	626,7	670,0	838,5	1,52E-01	0,18	1,39E-01	-0,30
GCW_00890	der	214,3	304,4	171,2	193,3	209,2	224,3	199,4	262,3	205,0	177,7	8,54E-02	-0,49	8,65E-01	0,05
GCW_00895	GCW_00895	201,4	254,8	161,4	159,3	218,6	188,1	126,4	167,5	134,9	148,3	1,33E-02	-0,50	3,84E-01	-0,17
GCW_00900	GCW_00900	87,0	111,2	97,7	83,2	50,1	36,2	45,9	43,5	43,3	38,0	5,28E-01	-0,13	8,69E-01	-0,03
GCW_00905	GCW_00905	722,5	682,5	581,8	725,2	2065,0	1176,5	2428,0	703,3	665,6	1760,2	5,24E-01	-0,11	2,11E-01	-0,91
GCW_00960	GCW_00960	950,4	660,7	1515,2	1409,4	1585,4	1455,7	1352,6	1056,0	943,7	1286,3	1,08E-05	0,88	2,08E-02	-0,39
GCW_00965	GCW_00965	843,8	578,6	1563,1	1326,9	2035,7	2546,1	2197,4	1552,9	2130,6	2238,6	3,83E-06	1,04	5,46E-01	-0,17
GCW_00975	GCW_00975	180,6	147,3	578,1	509,9	252,7	205,9	143,9	201,0	242,2	413,2	7,24E-49	1,74	2,75E-01	0,51
GCW_00980	GCW_00980	995,6	994,1	419,8	486,5	1578,6	1259,6	2412,4	695,8	526,5	1004,7	6,78E-22	-1,14	6,31E-03	-1,20
GCW_00985	GCW_00985	111,0	114,8	126,1	116,8	26,4	27,4	40,5	45,3	46,4	35,5	8,27E-02	0,11	5,66E-02	0,49
GCW_01000	ftsY	406,4	380,2	298,4	306,9	473,7	501,8	260,9	635,3	602,6	473,9	3,33E-13	-0,38	1,04E-01	0,56

GCW_01005	ylxM	217,6	223,0	216,2	180,3	134,7	94,8	92,9	407,4	415,5	329,1	2,06E-01	-0,16	5,69E-16	1,89
GCW_01010	GCW_01010	1451,0	1235,0	502,0	512,6	5853,0	6885,4	8377,6	2838,8	3476,5	3868,4	4,65E-34	-1,40	4,98E-07	-1,01
GCW_01015	GCW_01015	728,2	473,1	269,7	288,0	561,4	355,6	296,6	1179,9	1689,8	1147,0	2,07E-03	-1,07	8,04E-06	1,80
GCW_01020	thrS	760,5	807,7	519,0	498,3	1159,2	996,9	1820,1	1042,9	1169,6	896,9	1,91E-15	-0,62	4,31E-01	-0,28
GCW_01025	GCW_01025	419,3	388,3	409,7	363,4	183,4	180,5	240,5	426,4	502,5	283,6	4,32E-01	-0,06	4,69E-03	1,02
GCW_01030	hlp3	954,4	1131,1	798,3	807,8	1061,8	1402,4	1651,7	852,1	648,1	732,4	1,60E-02	-0,37	7,21E-05	-0,83
GCW_01035	licA	578,1	476,4	507,5	427,4	266,8	346,4	280,8	895,8	1029,6	674,9	3,36E-01	-0,17	2,93E-07	1,57
GCW_01040	mgc2	4479,0	3707,7	1356,5	1595,9	2332,8	2106,3	4909,6	1845,9	1647,3	1510,0	2,87E-16	-1,47	8,77E-02	-0,76
GCW_01045	gapA	2094,8	1695,9	1101,6	1227,5	1212,2	979,5	1527,5	1370,2	1396,4	1163,0	4,65E-05	-0,70	5,62E-01	0,14
GCW_01050	crmA	2202,7	1798,2	1326,6	1402,8	1456,7	1340,0	1683,9	1304,0	1001,7	946,6	1,87E-04	-0,54	2,20E-02	-0,43
GCW_01055	crmB	169,1	219,0	322,1	290,7	96,2	88,5	66,2	184,4	171,7	90,0	1,16E-03	0,67	5,15E-02	0,82
GCW_01060	crmC	96,4	133,1	309,8	270,1	33,3	35,0	39,1	82,6	88,3	42,3	2,31E-09	1,35	2,14E-02	0,96
GCW_01065	hisS	259,1	343,6	145,7	135,8	140,6	144,2	240,3	208,7	220,6	144,3	4,57E-05	-1,08	6,03E-01	0,19
GCW_01070	aspS	243,3	288,5	122,4	123,5	350,1	370,5	589,5	299,6	278,8	290,3	5,57E-10	-1,11	4,73E-02	-0,51
GCW_01075	spoT	127,5	172,3	96,1	86,5	61,9	72,2	121,8	128,1	134,4	93,5	7,79E-03	-0,70	1,39E-01	0,56
GCW_01080	pyrG	218,6	230,2	109,2	93,1	422,2	541,2	483,1	534,2	421,8	451,0	3,15E-27	-1,15	8,43E-01	0,00
GCW_01085	GCW_01085	633,8	571,5	189,2	191,4	469,9	588,7	645,0	641,6	656,0	526,8	9,30E-140	-1,66	4,37E-01	0,14
GCW_01090	apt	764,5	867,7	305,8	331,9	988,5	1067,1	2341,7	751,2	661,2	799,3	7,46E-17	-1,35	6,56E-02	-0,84
GCW_01095	GCW_01095	789,3	650,5	186,2	197,1	1813,9	1137,1	1464,9	480,0	487,7	731,3	8,09E-33	-1,90	3,09E-05	-1,34
GCW_01100	GCW_01100	987,4	984,4	433,0	490,4	529,6	288,6	382,7	391,4	408,4	395,5	1,99E-25	-1,10	8,63E-01	0,08
GCW_01105	greA	1173,9	1127,6	626,6	681,4	2200,6	1454,2	1515,1	717,7	761,0	880,2	1,40E-24	-0,82	4,55E-05	-1,07
GCW_01135	potE	293,8	268,2	404,9	393,2	580,7	548,3	477,6	379,2	566,5	1100,0	5,92E-50	0,51	5,28E-01	0,25
GCW_01140	dps	539,2	407,2	2392,9	2271,0	620,7	314,5	557,8	324,0	398,8	620,9	8,02E-91	2,32	8,62E-01	-0,11
GCW_01145	GCW_01145	1316,5	918,2	106,3	130,6	1424,8	2355,8	1419,4	1668,1	1955,5	2007,6	3,18E-15	-3,22	5,86E-01	0,19
GCW_01150	GCW_01150	84,0	72,0	43,8	51,2	167,0	197,4	212,6	103,8	94,6	101,6	7,60E-06	-0,71	2,32E-19	-0,90
GCW_01155	GCW_01155	195,9	181,5	37,8	45,2	530,6	686,6	615,5	383,2	306,1	443,8	1,00E-79	-2,19	4,14E-04	-0,66
GCW_01160	GCW_01160	150,3	106,4	69,3	77,8	283,4	360,3	380,5	210,5	190,6	266,8	4,22E-03	-0,78	3,04E-03	-0,58
GCW_01165	GCW_01165	107,9	97,0	362,2	614,8	51,8	56,8	126,2	361,9	346,8	251,7	5,98E-04	2,21	4,19E-07	2,18

GCW_01170	GCW_01170	93,9	92,3	42,2	54,2	416,9	553,8	570,5	538,1	440,1	485,7	1,42E-07	-0,96	7,31E-01	-0,03
GCW_01175	GCW_01175	105,1	118,8	31,8	35,8	327,2	420,9	434,6	255,6	211,2	282,5	4,70E-23	-1,73	2,60E-04	-0,62
GCW_01180	GCW_01180	67,4	74,7	256,2	426,4	41,6	40,7	100,7	253,2	265,7	181,0	4,14E-04	2,22	6,84E-06	2,09
GCW_01190	GCW_01190	68,9	76,6	274,8	453,7	46,7	41,3	104,6	272,9	282,8	198,1	2,70E-04	2,28	1,77E-06	2,12
GCW_01195	GCW_01195	200,6	184,5	65,9	74,9	572,0	659,7	615,6	437,2	362,4	499,7	5,40E-44	-1,45	1,49E-03	-0,48
GCW_01200	GCW_01200	36,6	23,6	40,2	28,9	15,0	18,2	23,0	20,3	22,3	16,6	4,72E-01	0,21	5,93E-01	0,13
GCW_01210	dut	63,4	101,6	329,1	370,7	147,4	107,7	44,4	65,4	57,7	112,4	1,60E-11	2,12	5,92E-01	-0,20
GCW_01215	GCW_01215	192,7	162,6	75,7	66,6	272,9	303,9	344,2	327,3	322,3	369,1	4,49E-25	-1,32	1,00E-01	0,19
GCW_01220	rluA_1	140,2	161,8	49,9	53,8	145,1	192,3	185,1	254,2	232,8	256,5	4,92E-18	-1,54	5,66E-07	0,56
GCW_01225	lspA	165,4	159,7	33,8	27,5	460,8	448,7	330,7	217,3	190,7	292,7	9,63E-209	-2,41	1,65E-03	-0,79
GCW_01230	GCW_01230	845,5	752,1	1235,3	1252,2	2368,5	1753,5	1769,7	960,1	731,3	1314,3	1,35E-21	0,64	2,76E-03	-0,96
GCW_01235	GCW_01235	289,6	319,0	206,5	200,3	880,7	538,3	651,5	561,2	452,4	757,8	4,05E-08	-0,58	5,96E-01	-0,19
GCW_01240	rpoB	988,2	993,8	1133,5	1078,1	1569,6	1743,8	1810,0	1783,9	1777,8	1723,8	2,47E-06	0,16	1,28E-01	0,09
GCW_01245	rpoC	1097,6	1012,9	1018,7	971,7	1284,4	1281,9	1586,1	1323,1	1284,8	1070,3	7,20E-03	-0,08	4,63E-01	-0,13
GCW_01250	GCW_01250	641,6	559,5	627,1	521,3	206,1	203,8	180,4	323,9	311,0	192,6	7,56E-01	-0,07	9,34E-02	0,49
GCW_01255	GCW_01255	421,3	429,1	1207,4	1106,0	228,8	278,4	303,0	412,2	346,9	313,8	2,10E-158	1,44	1,59E-02	0,44
GCW_01260	GCW_01260	514,8	517,7	1192,4	1125,1	282,2	309,8	439,3	341,5	350,8	336,3	0,00E+00	1,17	8,15E-01	0,06
GCW_01270	hatA	258,2	258,3	255,8	258,7	276,5	245,8	2262,7	259,5	201,3	307,4	7,87E-01	0,00	2,30E-01	-1,05
GCW_01275	hatB	122,7	90,6	184,9	151,7	86,2	130,2	381,9	282,3	259,0	282,6	1,11E-03	0,67	4,28E-01	0,80
GCW_01285	proS	208,2	199,8	208,8	185,4	287,5	407,0	364,4	301,8	270,0	312,3	4,81E-01	-0,05	1,16E-01	-0,21
GCW_01290	GCW_01290	80,5	84,8	422,0	397,8	74,8	78,6	28,5	100,1	65,7	66,6	0,00E+00	2,31	5,14E-01	0,50
GCW_01295	GCW_01295	60,5	60,4	556,6	472,8	11,8	9,1	10,3	38,2	33,0	25,0	1,36E-93	3,08	9,57E-09	1,65
GCW_01300	GCW_01300	15,0	21,4	412,4	329,1	2,2	2,8	3,2	10,1	13,4	10,9	2,59E-46	4,36	1,26E-17	2,10
GCW_01305	acpP	528,8	324,4	1579,8	1299,8	168,2	128,9	48,7	570,7	978,3	836,8	1,31E-10	1,79	5,50E-09	2,97
GCW_01315	tuf	7894,6	4166,0	2482,0	2845,5	10540,8	10933,2	21730,0	6840,2	8864,9	8315,0	3,65E-02	-1,11	8,41E-02	-0,73
GCW_01325	GCW_01325	721,0	658,1	913,3	933,3	445,1	475,8	492,1	573,8	674,1	615,1	6,03E-12	0,42	7,48E-06	0,44
GCW_01330	GCW_01330	213,8	250,7	725,1	679,0	61,2	95,7	99,7	170,1	201,0	125,6	8,40E-48	1,60	2,99E-03	1,00

GCW_01350	GCW_01350	491,2	374,6	188,2	159,6	1373,2	1508,4	937,9	787,8	876,4	974,3	5,03E-09	-1,31	2,82E-02	-0,47
GCW_01355	GCW_01355	1345,0	986,8	608,0	611,9	1149,2	637,2	1392,9	704,6	800,7	842,6	7,79E-05	-0,92	3,28E-01	-0,33
GCW_01360	dnaE	242,6	234,0	332,0	299,3	146,3	190,1	176,1	318,7	292,0	238,9	4,23E-17	0,40	4,45E-05	0,77
GCW_01365	exo	180,5	209,9	318,9	288,0	160,0	153,0	217,4	167,1	133,8	107,1	2,17E-06	0,64	2,13E-01	-0,35
GCW_01370	mutM	232,1	289,2	295,0	331,6	217,2	254,9	260,2	196,7	164,2	140,5	3,10E-01	0,27	1,57E-03	-0,52
GCW_01375	coaE	124,2	114,0	206,8	182,2	109,2	142,2	124,5	119,3	109,9	81,8	2,72E-21	0,71	2,77E-01	-0,24
GCW_01380	dnaB	115,3	155,7	233,3	208,1	81,9	69,4	90,0	52,5	55,7	51,8	2,26E-03	0,72	7,04E-05	-0,55
GCW_01385	GCW_01385	73,1	75,9	215,8	187,8	72,7	120,7	76,7	123,0	99,4	139,4	1,35E-46	1,43	1,63E-01	0,49
GCW_01390	mgtA	163,6	240,9	338,4	374,3	244,2	261,4	235,1	175,1	171,8	177,4	8,72E-03	0,84	1,31E-31	-0,46
GCW_01395	asnC	650,7	746,3	689,3	747,7	1212,7	1138,8	1972,6	745,5	712,0	1000,3	9,43E-01	0,04	2,63E-02	-0,75
GCW_01400	pgsA	196,9	279,4	419,4	397,2	146,0	228,6	254,7	384,2	371,7	316,0	2,13E-03	0,80	2,64E-04	0,84
GCW_01405	GCW_01405	150,8	139,0	359,9	313,6	128,5	182,6	147,8	302,4	336,5	258,0	6,44E-43	1,21	1,65E-06	1,01
GCW_01410	ispD	149,0	160,1	419,3	394,6	58,4	90,0	70,0	187,0	195,7	146,8	4,57E-103	1,40	6,32E-08	1,33
GCW_01415	GCW_01415	349,6	344,7	669,8	571,9	399,9	341,7	614,9	323,1	317,0	270,2	1,74E-17	0,83	9,93E-02	-0,50
GCW_01420	GCW_01420	1228,4	1008,7	1005,6	1016,1	1091,0	1735,2	1868,2	1673,5	1756,5	2000,2	2,78E-01	-0,14	2,66E-01	0,28
GCW_01430	GCW_01430	127,5	123,2	466,5	366,1	93,5	131,0	54,5	87,9	74,8	88,1	6,67E-19	1,72	6,76E-01	-0,03
GCW_01435	GCW_01435	59,0	54,3	290,2	213,9	16,4	33,5	26,9	26,3	30,8	23,6	3,68E-15	2,14	7,07E-01	0,16
GCW_01440	GCW_01440	68,3	78,7	391,5	369,1	64,5	94,2	64,5	47,4	49,2	61,4	4,40E-163	2,37	4,33E-02	-0,44
GCW_01445	GCW_01445	49,2	50,2	320,5	271,8	9,6	14,3	7,5	32,9	41,7	46,9	5,49E-74	2,57	5,03E-12	2,03
GCW_01450	GCW_01450	165,5	129,8	348,8	326,4	278,3	457,1	190,1	464,9	492,2	492,1	3,92E-24	1,20	3,87E-02	0,78
GCW_01455	GCW_01455	66,5	62,9	44,0	49,1	51,4	91,0	31,4	90,1	89,5	85,2	1,41E-05	-0,48	1,16E-01	0,78
GCW_01460	GCW_01460	52,5	20,5	49,9	42,2	7,4	11,6	5,4	43,7	64,7	19,5	4,92E-01	0,48	5,73E-03	2,34
GCW_01470	GCW_01470	23,7	14,5	31,2	19,9	214,0	186,8	55,1	93,9	103,9	179,3	2,64E-01	0,43	6,86E-01	-0,07
GCW_01475	GCW_01475	19,0	5,5	24,6	16,3	87,4	122,8	26,6	101,2	149,1	154,1	2,69E-01	0,96	1,04E-01	1,05
GCW_01480	GCW_01480	326,7	389,0	254,4	320,3	1064,1	969,3	1314,7	545,5	478,9	794,4	1,60E-01	-0,32	2,76E-03	-0,86
GCW_01490	GCW_01490	153,5	66,6	216,4	149,5	74,3	130,2	41,6	161,6	192,0	229,2	1,45E-01	0,83	1,87E-03	1,42
GCW_01495	GCW_01495	126,3	112,1	287,0	210,6	38,4	70,2	21,1	105,4	124,1	109,9	5,15E-06	1,05	2,29E-04	1,59

GCW_01500	GCW_01500	296,2	329,2	1346,8	1412,3	24,7	26,7	19,2	167,8	173,4	146,4	7,01E-71	2,14	6,21E-57	2,84
GCW_01505	GCW_01505	154,5	109,2	853,6	876,5	3,3	14,0	23,9	87,9	125,7	102,9	7,63E-67	2,74	2,32E-10	3,38
GCW_01515	asnS_2	124,4	109,7	100,7	88,6	20,7	37,6	17,7	45,8	48,4	42,9	1,13E-03	-0,31	4,13E-03	0,97
GCW_01520	GCW_01520	80,4	47,1	137,3	107,8	111,9	128,0	50,3	102,4	100,5	126,8	4,14E-03	0,98	6,15E-01	0,33
GCW_01525	GCW_01525	80,5	53,4	244,4	227,6	152,6	250,4	126,8	180,1	234,8	303,6	2,05E-21	1,85	2,26E-01	0,51
GCW_01530	GCW_01530	68,1	85,4	459,1	483,4	48,7	25,2	14,3	89,2	74,8	116,2	2,40E-58	2,63	6,99E-04	1,86
GCW_01535	GCW_01535	502,5	619,3	905,9	1063,0	1312,6	1052,2	1597,0	941,5	1017,1	1928,6	1,07E-03	0,82	9,70E-01	-0,05
GCW_01540	GCW_01540	195,0	167,6	284,9	244,9	736,8	518,9	372,6	566,9	604,9	957,1	7,67E-07	0,55	3,48E-01	0,44
GCW_01545	GCW_01545	312,0	355,5	378,3	357,0	391,3	351,6	236,5	306,9	316,3	304,5	3,08E-01	0,14	8,60E-01	-0,01
GCW_01550	GCW_01550	149,7	207,6	130,7	125,1	77,2	86,4	118,0	174,9	162,7	146,5	7,08E-02	-0,46	1,21E-05	0,84
GCW_01555	GCW_01555	356,6	405,7	167,8	169,8	1309,6	1076,2	1189,2	521,9	373,1	691,7	1,60E-15	-1,17	3,34E-05	-1,17
GCW_01560	GCW_01560	57,1	87,2	94,3	102,9	28,8	58,0	28,0	122,8	111,3	74,0	2,23E-01	0,48	9,03E-04	1,52
GCW_01570	GCW_01570	179,6	116,6	73,3	82,8	145,0	168,3	153,3	124,1	213,1	195,3	1,04E-02	-0,89	3,18E-01	0,19
GCW_01575	ppa	1016,3	989,2	667,3	841,7	3431,0	1642,8	4188,0	683,4	678,2	1251,6	2,48E-02	-0,42	1,55E-03	-1,74
GCW_01580	GCW_01580	23,8	28,9	7,3	6,7	135,7	113,7	99,4	73,0	61,6	74,5	1,33E-17	-1,90	5,41E-05	-0,69
GCW_01585	sufC	487,6	436,3	164,0	194,8	1087,4	1057,7	2957,8	640,6	505,8	1006,2	1,16E-22	-1,37	9,93E-02	-1,09
GCW_01590	sufB	324,4	276,4	154,6	173,9	655,4	877,8	765,9	652,8	601,1	879,4	1,02E-09	-0,87	7,29E-01	-0,08
GCW_01595	csdB	237,8	223,6	158,3	171,7	110,8	133,2	255,1	292,9	332,3	346,6	1,09E-08	-0,48	2,00E-03	1,09
GCW_01600	GCW_01600	243,1	239,0	188,1	177,9	445,9	375,6	398,9	377,1	384,2	437,9	4,58E-18	-0,40	8,83E-01	0,01
GCW_01605	sufB	419,8	478,3	332,0	408,7	1557,4	1713,8	2955,0	884,4	701,8	930,7	1,58E-01	-0,28	8,94E-04	-1,22
GCW_01610	dnaJ_4	268,5	337,8	315,6	357,2	118,3	116,7	167,0	221,8	193,6	143,3	6,43E-01	0,16	7,86E-02	0,51
GCW_01615	gpsA	282,7	359,0	403,1	460,7	302,1	193,5	247,9	232,4	245,6	170,4	1,03E-01	0,44	6,74E-01	-0,15
GCW_01620	dnaJ_5	187,3	240,2	373,6	396,3	134,2	224,4	170,2	225,2	176,7	131,8	4,18E-05	0,86	9,61E-01	0,05
GCW_01625	GCW_01625	3104,6	1721,5	1425,6	1557,7	6189,4	6062,1	9632,3	3188,9	3875,2	3244,9	1,57E-01	-0,63	4,48E-04	-1,02
GCW_01630	GCW_01630	2310,7	1731,5	1458,5	1490,1	5032,4	5007,1	8162,8	3652,1	4578,3	3578,7	3,02E-02	-0,44	8,71E-02	-0,55
GCW_01635	osmC	4174,7	1940,7	1074,1	1224,7	4200,6	3870,3	12886,3	2504,4	2390,9	5594,3	4,36E-02	-1,31	2,89E-01	-0,84
GCW_01640	ffh	203,5	226,0	155,9	159,1	803,2	892,4	820,1	436,5	350,5	356,5	1,13E-04	-0,44	2,62E-29	-1,10
GCW_01645	tyrS	147,3	190,0	137,7	111,9	591,2	593,9	499,2	234,1	238,6	261,2	8,11E-02	-0,43	5,44E-31	-1,16

GCW_01650	glyA	222,9	289,1	165,7	154,2	843,3	472,4	433,4	281,7	270,8	287,2	3,67E-03	-0,67	1,09E-02	-0,95
GCW_01660	GCW_01660	221,9	182,3	264,5	312,1	404,1	446,5	256,5	349,1	376,9	489,8	1,24E-02	0,52	5,95E-01	0,20
GCW_01665	GCW_01665	797,8	691,7	1252,8	1228,4	954,4	820,3	1256,5	540,0	468,2	547,1	7,28E-29	0,74	3,39E-05	-0,90
GCW_01670	GCW_01670	18,7	21,6	57,7	34,8	13,5	6,0	1,7	21,0	25,6	28,2	3,85E-03	1,16	6,48E-04	2,30
GCW_01675	GCW_01675	6,0	5,2	27,0	27,3	14,6	11,4	0,0	25,5	14,0	24,7	7,10E-12	2,27	9,34E-02	3,81
GCW_01680	GCW_01680	1043,1	796,2	916,2	877,0	2087,2	2038,4	2267,3	1281,4	1249,5	1595,5	9,40E-01	-0,02	4,22E-06	-0,60
GCW_01685	rplM	2188,7	1492,9	422,3	403,1	1723,9	1318,5	1563,8	821,9	1023,6	923,7	2,10E-09	-2,13	1,66E-04	-0,69
GCW_01690	rpsI	1951,9	1140,6	406,3	398,0	1921,9	1582,1	1102,5	675,4	774,6	892,3	1,45E-04	-1,89	1,02E-03	-0,91
GCW_01695	ispG	496,5	554,3	45,3	56,9	112,3	74,5	92,1	214,1	194,4	130,6	1,19E-55	-3,37	3,31E-03	0,98
GCW_01700	upp	487,8	579,0	77,8	89,8	615,6	285,5	575,0	198,1	161,6	200,5	2,19E-27	-2,67	1,01E-03	-1,29
GCW_01705	GCW_01705	794,6	830,9	456,0	433,7	581,6	600,6	747,4	755,8	662,4	838,3	7,76E-36	-0,87	1,31E-01	0,27
GCW_01710	atpI	1969,7	1313,7	474,4	482,4	179,6	165,0	228,6	344,0	388,4	264,2	9,99E-07	-1,75	1,68E-03	0,83
GCW_01715	atpB	1554,8	1249,2	558,5	559,4	643,3	403,1	1284,8	294,0	314,2	288,1	5,12E-16	-1,32	3,72E-02	-1,18
GCW_01720	atpE	1247,8	1112,1	624,3	798,4	1964,0	1412,7	1688,0	378,1	281,0	561,2	1,95E-04	-0,74	2,19E-10	-2,06
GCW_01725	atpF	820,7	600,4	369,9	359,6	247,5	169,2	342,4	228,3	283,3	276,5	5,64E-05	-0,94	6,74E-01	0,15
GCW_01730	atpH	653,6	604,3	425,5	363,7	195,7	169,2	427,9	262,6	305,0	271,7	4,17E-16	-0,68	7,29E-01	0,25
GCW_01735	atpA	879,2	916,9	636,2	702,5	875,4	724,7	1621,0	541,4	537,2	557,6	1,55E-04	-0,42	3,72E-02	-0,85
GCW_01740	atpG	788,7	758,2	555,5	504,9	315,3	312,2	724,9	496,8	496,6	411,7	8,45E-52	-0,55	7,98E-01	0,21
GCW_01745	atpD_1	993,4	987,3	672,7	736,9	1112,5	805,3	2248,1	620,3	492,8	575,2	1,07E-07	-0,49	2,91E-02	-1,13
GCW_01750	atpC	576,6	500,1	457,1	438,0	153,0	148,2	351,8	337,4	328,3	270,8	9,10E-04	-0,26	1,87E-01	0,68
GCW_01755	GCW_01755	267,5	249,0	401,0	341,6	62,8	74,6	161,7	158,5	167,9	91,8	3,56E-08	0,52	3,18E-01	0,60
GCW_01775	GCW_01775	619,2	493,1	324,1	297,3	1648,4	1283,4	1558,7	777,5	767,4	945,9	2,33E-07	-0,83	1,40E-06	-0,81
GCW_01780	gapd	5010,0	2607,5	1267,6	1319,0	7028,2	6643,1	17478,5	3293,6	3824,9	4647,0	9,38E-03	-1,48	3,63E-02	-1,23
GCW_01785	pgk	5184,5	2752,7	1332,3	1461,0	2711,5	1649,0	2852,6	2533,1	3991,6	2905,9	9,19E-03	-1,44	1,80E-01	0,44
GCW_01790	GCW_01790	318,0	423,5	236,1	242,0	978,5	961,7	562,1	658,7	750,3	767,2	1,13E-02	-0,62	5,67E-01	-0,12
GCW_01795	GCW_01795	195,1	186,5	230,2	187,4	109,7	103,7	102,1	247,7	318,6	256,9	2,47E-01	0,12	6,73E-13	1,41
GCW_01800	rpsU	2341,5	1705,2	1737,5	1771,1	1280,6	1501,3	1186,0	1101,2	1232,3	1319,8	3,97E-01	-0,19	5,89E-01	-0,08
GCW_01810	GCW_01810	1046,8	976,6	353,8	333,7	1362,5	1625,1	1539,1	1039,1	1034,7	1266,3	0,00E+00	-1,56	7,28E-04	-0,40

GCW_01815	GCW_01815	110,6	113,3	46,0	36,9	263,8	243,6	219,3	109,2	86,2	108,3	3,12E-34	-1,44	5,10E-19	-1,22
GCW_01820	cspR	145,5	162,5	39,6	40,7	118,0	108,8	71,9	83,6	77,5	88,3	1,08E-37	-1,94	3,04E-01	-0,19
GCW_01825	ileS	402,8	405,9	147,2	146,8	884,1	948,7	1300,9	830,4	801,6	1022,0	2,91E-198	-1,46	3,99E-01	-0,19
GCW_01830	fldA	335,8	432,3	152,9	149,7	209,5	325,2	257,9	598,7	603,4	501,4	1,38E-07	-1,33	3,47E-09	1,16
GCW_01835	plpA	1921,1	1988,3	1186,9	1212,1	2568,9	3249,2	6499,0	1706,8	1396,1	1730,9	6,45E-23	-0,70	1,31E-02	-1,20
GCW_01840	hlp2	500,9	523,8	590,3	583,9	154,7	221,4	242,3	346,0	376,5	311,1	3,88E-03	0,20	4,79E-05	0,80
GCW_01845	GCW_01845	325,9	354,9	551,6	527,1	110,0	184,7	184,0	274,0	404,6	312,2	2,29E-16	0,67	3,27E-04	1,11
GCW_01850	GCW_01850	209,8	162,2	400,2	389,9	46,5	74,2	123,9	178,8	283,7	199,8	1,33E-17	1,10	8,31E-04	1,56
GCW_01860	GCW_01860	76,2	94,1	34,1	37,1	8,6	8,2	3,7	25,9	23,2	12,8	3,01E-08	-1,25	1,93E-03	1,67
GCW_01865	GCW_01865	96,2	139,8	34,1	55,9	10,1	10,7	7,2	43,3	32,0	22,2	2,54E-03	-1,41	8,02E-05	1,80
GCW_01870	GCW_01870	378,0	383,4	234,8	265,3	1289,9	1093,4	993,2	1173,2	1232,6	1350,4	1,34E-07	-0,61	1,63E-01	0,20
GCW_01875	arcA_1	848,3	917,0	1028,7	1111,5	963,8	890,7	1497,7	792,6	762,5	691,2	4,29E-02	0,28	6,27E-02	-0,50
GCW_01880	trxB_1	905,2	841,1	940,9	940,5	783,3	1021,6	1487,2	962,8	860,4	771,3	7,13E-03	0,11	3,32E-01	-0,26
GCW_01885	GCW_01885	275,6	333,2	988,2	988,2	293,5	441,9	380,8	234,7	236,5	311,9	2,16E-39	1,71	3,13E-02	-0,47
GCW_01890	GCW_01890	499,3	555,3	1608,0	1540,6	706,8	930,5	1248,3	644,2	508,7	563,1	3,80E-88	1,58	7,33E-03	-0,68
GCW_01895	GCW_01895	139,3	185,0	1161,3	1205,3	35,5	26,5	75,6	130,3	113,4	76,0	1,69E-64	2,88	1,31E-02	1,37
GCW_01900	dnaJ_6	1127,7	1080,1	2520,1	2448,3	511,1	589,8	813,6	690,1	804,1	644,8	0,00E+00	1,17	3,54E-01	0,22
GCW_01905	grpE	1424,6	1727,0	2484,3	2685,6	1635,5	1804,6	4585,8	764,1	761,3	932,9	1,46E-04	0,72	1,54E-02	-1,51
GCW_01910	GCW_01910	105,4	138,1	299,0	281,5	28,3	44,0	38,0	211,0	239,0	145,7	3,54E-12	1,27	2,02E-09	2,46
GCW_01920	GCW_01920	299,0	335,6	536,9	575,3	463,5	402,3	405,3	391,7	384,4	627,8	7,72E-09	0,81	6,16E-01	0,15
GCW_01925	GCW_01925	146,5	131,3	487,1	817,8	71,2	76,1	170,7	475,8	445,9	339,0	5,05E-04	2,19	2,42E-07	2,13
GCW_01930	GCW_01930	275,7	277,1	130,8	128,0	514,5	677,8	614,7	391,5	404,3	597,4	8,57E-160	-1,09	1,89E-01	-0,35
GCW_01935	GCW_01935	90,0	72,4	230,1	371,1	37,9	44,4	81,7	261,1	266,3	189,5	9,12E-04	1,86	2,92E-09	2,23
GCW_01940	GCW_01940	11799,5	5095,8	481,6	650,7	7394,7	10193,5	15964,9	7256,1	11103,0	9899,9	4,89E-04	-3,79	7,31E-01	-0,16
GCW_01945	GCW_01945	148,0	96,9	153,8	248,2	34,7	39,9	56,5	306,6	381,3	246,7	1,81E-01	0,71	1,98E-12	2,88
GCW_01960	GCW_01960	153,2	155,6	17,4	26,2	298,9	398,2	371,0	219,3	165,7	269,3	1,56E-41	-2,85	4,80E-03	-0,69
GCW_01965	GCW_01965	2,2	2,3	18,0	15,9	0,4	0,0	0,9	0,7	1,6	4,5	3,61E-25	2,88	9,71E-02	3,13

GCW_01970	GCW_01970	76,3	82,4	154,4	162,3	107,8	75,2	104,8	76,2	80,4	133,0	4,89E-20	1,00	8,87E-01	0,02
GCW_01975	beta-pgm	30,6	44,2	298,6	353,8	6,7	0,0	5,4	15,9	26,9	23,6	2,25E-22	3,14	4,46E-05	4,60
GCW_01980	GCW_01980	810,3	706,8	1289,4	1264,5	964,8	825,8	1265,4	552,7	484,5	563,3	1,76E-33	0,76	5,11E-05	-0,87
GCW_01985	GCW_01985	37,2	20,5	102,6	63,0	38,8	59,2	8,7	54,8	58,9	40,8	2,05E-03	1,54	3,65E-01	0,94
GCW_01990	GCW_01990	40,9	92,7	131,7	126,5	85,6	48,3	108,2	68,6	44,5	40,5	7,25E-02	1,07	1,63E-01	-0,58
GCW_01995	GCW_01995	27,4	37,4	74,5	64,6	16,9	17,8	6,4	30,2	34,7	25,7	1,15E-06	1,11	1,58E-03	1,31
GCW_02000	GCW_02000	161,7	178,2	66,0	50,7	323,7	323,4	368,8	422,7	490,7	396,1	2,45E-18	-1,55	5,35E-03	0,40
GCW_02005	hrcA	183,6	247,2	60,8	50,2	787,7	600,0	874,3	399,5	390,6	400,2	5,80E-09	-1,95	3,10E-06	-0,87
GCW_02010	uvrC	55,4	56,8	68,8	56,8	54,8	73,6	95,5	101,3	88,8	74,8	1,73E-01	0,16	3,12E-01	0,31
GCW_02015	GCW_02015	140,3	188,9	70,9	84,1	67,3	71,5	43,7	76,6	58,8	43,4	3,39E-04	-1,07	9,31E-01	0,00
GCW_02020	infC	1715,1	1435,5	585,8	857,0	2876,6	3372,9	2861,2	1435,4	1353,5	1940,8	6,67E-05	-1,15	4,22E-07	-0,92
GCW_02025	rpmI	796,9	806,3	331,8	416,2	1679,4	1528,8	665,9	564,5	465,0	850,5	4,44E-11	-1,11	3,57E-02	-0,94
GCW_02030	rplT	1392,8	1531,4	632,5	846,8	2330,5	1453,9	1362,4	825,5	801,4	1022,1	2,03E-05	-1,00	6,18E-03	-0,88
GCW_02035	GCW_02035	1016,7	635,1	491,1	551,4	559,1	306,9	536,3	665,5	726,4	706,5	8,44E-02	-0,63	8,92E-03	0,67
GCW_02055	GCW_02055	3,6	3,6	13,2	15,4	0,4	0,0	0,8	0,9	0,9	0,5	1,95E-15	1,97	2,92E-01	2,03
GCW_02060	GCW_02060	65,1	57,9	44,3	38,8	24,0	45,6	29,8	33,8	41,6	41,5	1,04E-10	-0,57	3,49E-01	0,32
GCW_02065	GCW_02065	40,4	42,2	53,4	46,3	18,8	37,7	28,1	48,6	51,3	38,5	5,36E-03	0,27	1,29E-02	0,80
GCW_02070	rplU	2033,4	1465,0	1104,3	1205,5	1683,0	1363,6	613,4	878,3	1025,1	1052,1	1,94E-02	-0,58	5,47E-01	-0,15
GCW_02075	GCW_02075	2400,8	1355,5	1343,3	1451,8	1027,7	463,9	313,2	980,1	1141,9	1218,5	3,68E-01	-0,37	4,87E-02	1,10
GCW_02080	rpmA	1195,7	902,1	961,6	1003,3	2037,3	1062,3	635,4	646,5	705,7	916,2	6,83E-01	-0,08	2,55E-01	-0,53
GCW_02085	nfo	165,1	190,6	251,3	250,6	186,7	157,6	414,8	182,2	142,8	235,3	1,45E-04	0,50	5,11E-01	-0,29
GCW_02090	fur	87,8	105,3	196,3	158,0	63,5	108,9	199,1	153,7	120,4	198,7	1,58E-05	0,87	5,12E-01	0,51
GCW_02095	GCW_02095	55,6	67,6	191,8	159,5	83,3	111,2	162,5	110,1	83,6	136,5	5,19E-16	1,51	8,15E-01	-0,05
GCW_02100	tig	864,8	1025,8	753,2	784,5	1880,1	1394,7	2809,4	801,8	755,9	1017,1	5,86E-02	-0,29	1,91E-03	-1,15
GCW_02105	lon	885,4	912,7	3081,3	2894,5	1586,5	1787,1	2251,8	1942,5	2232,4	3481,5	0,00E+00	1,73	1,93E-01	0,45
GCW_02110	GCW_02110	143,1	161,5	314,4	329,5	85,5	125,7	143,1	185,7	227,4	236,1	1,33E-17	1,08	9,99E-06	0,94
GCW_02115	GCW_02115	133,1	158,6	350,1	377,6	134,0	124,4	114,9	190,0	192,2	236,5	3,93E-15	1,32	2,69E-07	0,76
GCW_02120	GCW_02120	23,5	29,9	50,7	48,9	2,6	1,4	4,8	5,8	11,5	5,0	3,72E-07	0,91	5,25E-02	1,44

GCW_02125	GCW_02125	17,8	30,7	101,4	115,3	3,5	0,8	5,0	9,8	6,0	4,2	3,43E-10	2,21	1,28E-01	1,41
GCW_02130	GCW_02130	16,0	11,1	93,6	98,8	0,6	0,0	0,6	4,0	10,3	3,6	4,73E-54	2,85	2,94E-03	4,71
GCW_02135	GCW_02135	242,0	264,0	290,5	301,6	1108,2	566,0	683,1	285,5	203,6	462,2	5,78E-02	0,23	1,44E-02	-1,29
GCW_02140	GCW_02140	72,0	74,4	161,6	133,7	45,1	77,2	26,2	173,9	157,2	134,5	6,44E-14	1,01	3,79E-06	1,82
GCW_02150	GCW_02150	81,8	88,2	120,4	122,3	179,1	188,1	127,6	173,4	137,2	125,4	3,02E-08	0,51	4,85E-01	-0,14
GCW_02155	GCW_02155	210,8	179,0	119,1	119,0	617,0	607,3	461,8	365,7	230,1	371,7	9,83E-12	-0,71	2,67E-03	-0,78
GCW_02160	GCW_02160	105,4	86,6	103,9	84,0	75,7	107,8	86,0	86,4	78,8	91,7	9,96E-01	-0,03	7,71E-01	-0,02
GCW_02165	GCW_02165	106,9	143,3	147,9	145,4	189,9	271,1	242,4	290,1	264,1	298,7	4,01E-01	0,24	3,25E-02	0,33
GCW_02170	GCW_02170	273,6	294,0	1014,7	1109,8	452,5	404,9	434,8	385,9	403,9	370,9	2,03E-37	1,90	2,57E-01	-0,12
GCW_02175	GCW_02175	259,3	281,4	954,5	1013,6	200,4	233,0	292,5	292,6	283,4	243,6	5,27E-52	1,86	2,45E-01	0,23
GCW_02180	GCW_02180	278,9	202,3	788,2	741,6	105,4	144,0	134,6	265,8	278,1	198,1	3,70E-29	1,69	4,73E-05	0,99
GCW_02185	GCW_02185	554,7	496,6	899,1	945,6	141,0	262,2	217,7	443,7	427,2	310,9	3,37E-19	0,81	1,51E-03	1,00
GCW_02190	rpoD	593,0	540,2	522,5	504,8	421,0	385,7	274,4	320,2	274,0	290,3	5,34E-04	-0,14	1,90E-01	-0,23
GCW_02195	dnaG	106,5	121,0	95,3	76,3	78,1	119,1	108,1	103,0	96,3	70,8	2,50E-02	-0,41	6,03E-01	-0,13
GCW_02200	glyS	200,9	222,1	112,8	117,0	130,0	132,1	217,3	123,9	130,6	84,2	1,09E-12	-0,88	2,46E-01	-0,45
GCW_02205	recO	51,3	50,5	78,3	64,7	14,9	39,8	25,1	67,0	57,2	41,7	3,38E-05	0,48	1,29E-02	1,18
GCW_02210	era	82,4	72,7	89,0	75,4	79,3	78,9	193,5	51,0	37,9	37,5	3,59E-01	0,08	1,76E-02	-1,32
GCW_02215	GCW_02215	128,2	148,9	90,0	83,5	89,1	63,7	79,1	68,1	51,0	48,7	7,55E-06	-0,67	6,28E-02	-0,43
GCW_02220	GCW_02220	4,0	5,3	30,5	26,1	5,7	4,5	6,0	3,1	4,0	5,6	1,16E-39	2,60	4,19E-01	-0,35
GCW_02225	GCW_02225	14,6	12,2	19,4	14,5	1,6	7,1	2,7	6,8	7,8	4,2	8,11E-02	0,33	2,79E-01	1,00
GCW_02230	GCW_02230	10,1	10,9	29,3	30,6	3,3	3,6	0,6	5,6	5,4	7,7	5,75E-26	1,51	7,77E-03	1,68
GCW_02235	ftsZ	1099,8	912,9	3565,1	3238,3	168,3	142,7	182,7	428,4	475,9	356,7	1,19E-80	1,76	2,09E-10	1,39
GCW_02240	ftsA	171,0	195,2	2756,3	2228,1	376,1	336,9	191,5	185,7	110,6	172,0	4,12E-54	3,76	1,18E-02	-0,89
GCW_02245	mraW	462,3	518,7	3195,6	3154,1	66,8	53,6	45,7	181,5	174,8	98,4	7,97E-271	2,70	5,54E-04	1,46
GCW_02250	mraZ	905,1	697,4	3513,8	2819,7	239,1	179,7	99,9	497,5	528,1	417,4	5,37E-21	1,99	1,24E-06	1,60
GCW_02265	plsC	1294,7	1042,0	1861,9	1786,8	754,1	982,0	1335,1	1038,9	1368,0	1341,6	6,60E-09	0,65	1,79E-01	0,35
GCW_02270	acpS	1448,0	1176,2	1610,6	1626,8	2455,1	1771,9	1757,8	947,8	1116,1	1322,3	1,11E-02	0,31	1,65E-03	-0,78
GCW_02275	GCW_02275	293,3	267,7	277,5	320,6	719,3	324,4	823,2	117,5	122,3	244,8	6,08E-01	0,09	1,37E-03	-1,88

GCW_02285	GCW_02285	320,4	327,2	383,8	348,7	444,1	582,0	612,1	365,6	311,5	384,5	3,06E-03	0,18	1,96E-04	-0,58
GCW_02290	GCW_02290	334,3	232,5	217,4	178,8	174,0	102,5	305,2	74,6	89,7	77,8	9,18E-02	-0,50	3,68E-02	-1,09
GCW_02295	GCW_02295	1026,8	949,5	681,6	768,5	1903,4	1194,7	1622,7	629,6	779,9	920,0	8,72E-05	-0,45	7,05E-04	-0,97
GCW_02300	GCW_02300	50,6	79,0	312,8	322,1	86,4	84,0	30,5	105,0	75,2	62,1	4,54E-21	2,33	6,20E-01	0,42
GCW_02310	mtn	77,1	75,3	110,5	91,1	48,8	40,1	46,8	76,0	79,0	69,4	5,66E-04	0,40	3,78E-10	0,77
GCW_02315	nadD	139,7	122,4	135,0	123,5	124,2	109,3	184,9	122,8	137,9	146,4	9,45E-01	-0,02	9,10E-01	0,03
GCW_02320	yqeH	148,4	182,0	102,0	88,7	60,2	51,0	119,0	137,6	140,9	111,3	1,15E-04	-0,79	3,28E-02	0,89
GCW_02325	parC	138,0	182,5	84,8	87,2	146,4	168,6	131,6	314,9	305,0	258,0	5,16E-04	-0,88	5,05E-11	1,02
GCW_02330	parE	306,3	427,1	110,4	132,4	406,9	315,3	306,7	440,2	379,0	333,0	7,18E-06	-1,58	3,85E-01	0,20
GCW_02335	hup_2	2520,2	1843,5	557,5	477,2	3643,4	2836,3	7355,4	1212,8	1073,7	2111,6	5,96E-13	-2,06	1,23E-02	-1,56
GCW_02340	GCW_02340	390,2	478,1	107,6	94,3	470,7	483,5	215,1	465,4	354,1	463,2	1,70E-17	-2,10	7,53E-01	0,25
GCW_02350	GCW_02350	99,7	192,9	70,7	72,9	126,3	167,9	34,8	142,3	119,1	97,2	5,72E-02	-0,95	8,47E-01	0,42
GCW_02355	GCW_02355	111,6	166,7	96,9	97,3	145,4	192,7	140,1	271,3	238,1	210,6	1,02E-01	-0,49	1,33E-03	0,64
GCW_02360	GCW_02360	132,6	177,6	176,9	182,8	92,4	117,0	67,8	276,2	311,4	230,5	4,66E-01	0,23	3,89E-09	1,63
GCW_02365	GCW_02365	370,4	458,4	764,0	853,2	245,2	292,9	280,6	458,5	500,3	435,3	5,80E-06	0,97	2,01E-13	0,81
GCW_02370	GCW_02370	266,3	279,0	1167,6	1140,9	126,3	200,8	181,4	336,3	407,3	296,5	0,00E+00	2,08	3,61E-05	1,09
GCW_02375	GCW_02375	370,7	392,6	40,6	38,9	119,7	69,3	50,1	107,1	97,0	128,0	1,03E-165	-3,26	2,23E-01	0,60
GCW_02380	GCW_02380	306,3	194,7	44,7	55,1	298,4	142,5	37,2	96,6	104,4	112,1	5,16E-07	-2,30	4,88E-01	-0,12
GCW_02390	GCW_02390	252,0	199,9	27,0	35,1	588,0	557,1	677,7	357,8	279,7	362,1	3,26E-28	-2,86	3,93E-08	-0,83
GCW_02395	GCW_02395	99,1	85,3	316,5	531,2	45,1	49,0	102,9	308,4	306,9	229,8	5,75E-04	2,16	5,03E-09	2,23
GCW_02405	GCW_02405	178,8	188,8	81,3	85,9	350,8	450,0	491,4	241,8	230,1	279,5	1,39E-30	-1,13	3,85E-06	-0,73
GCW_02410	GCW_02410	1053,6	1012,3	294,1	468,9	953,5	1216,4	2576,4	574,6	457,9	534,8	5,55E-07	-1,47	7,92E-03	-1,43
GCW_02415	GCW_02415	157,3	134,1	532,4	894,7	74,4	80,9	173,4	504,3	474,2	353,0	4,19E-04	2,25	1,38E-07	2,15
GCW_02425	GCW_02425	17,9	16,0	45,3	55,7	2,9	2,3	1,7	19,0	18,5	9,4	4,77E-10	1,57	2,86E-06	2,77
GCW_02430	GCW_02430	792,5	769,3	311,0	347,2	951,7	1130,5	1055,9	1098,8	1043,0	1110,0	8,46E-44	-1,25	1,90E-01	0,09
GCW_02440	GCW_02440	450,7	485,0	155,3	255,7	380,0	459,3	1159,4	191,0	121,1	157,9	3,06E-04	-1,23	6,36E-03	-1,89
GCW_02455	GCW_02455	314,7	324,3	37,5	48,1	439,7	504,7	580,8	372,2	377,9	450,9	4,18E-82	-2,91	4,13E-02	-0,30
GCW_02465	GCW_02465	536,2	233,0	402,9	651,2	653,1	607,1	198,7	924,4	1101,3	1414,1	5,00E-01	0,54	2,73E-03	1,44

GCW_02470	GCW_02470	134,9	106,7	98,6	97,9	101,9	106,4	102,1	190,0	280,8	212,9	6,25E-02	-0,29	7,89E-06	1,16
GCW_02475	GCW_02475	307,8	354,0	316,6	345,2	219,3	271,3	109,1	440,1	468,8	384,2	8,66E-01	0,00	2,29E-04	1,24
GCW_02480	trkA	574,8	550,9	134,7	134,6	800,2	580,2	797,0	484,5	534,4	461,6	0,00E+00	-2,06	1,23E-02	-0,50
GCW_02485	leuS	459,0	498,3	151,5	154,6	481,0	565,6	738,7	558,0	631,3	603,5	5,04E-44	-1,64	7,31E-01	0,07
GCW_02490	GCW_02490	1692,7	1228,5	765,5	777,6	3086,2	3351,4	8711,5	2365,2	2555,1	3955,2	1,88E-04	-0,90	2,95E-01	-0,60
GCW_02495	GCW_02495	1461,5	1222,7	611,1	698,4	4100,1	4067,4	11212,6	2121,4	2246,9	3800,6	1,14E-10	-1,03	9,88E-02	-1,08
GCW_02500	GCW_02500	390,3	393,6	183,0	178,3	1427,5	1243,1	2086,1	570,8	396,5	799,7	6,73E-151	-1,12	4,55E-04	-1,41
GCW_02505	GCW_02505	103,4	118,3	207,9	207,2	49,1	19,2	27,3	53,8	64,0	54,3	1,86E-15	0,91	1,96E-02	0,99
GCW_02510	GCW_02510	919,4	975,4	3520,9	3571,6	2438,1	1962,9	2322,5	1044,5	951,6	1452,0	1,42E-134	1,91	4,80E-05	-0,94
GCW_02515	GCW_02515	681,4	599,4	1144,8	1055,2	658,0	556,3	729,4	1076,2	1213,5	962,6	1,66E-35	0,78	1,51E-05	0,78
GCW_02520	pepC_1	844,2	702,1	962,8	880,9	1997,1	2439,1	3072,2	1836,1	2012,4	1784,1	1,11E-02	0,26	7,49E-02	-0,36
GCW_02525	pepC_2	693,1	571,0	808,4	738,4	1961,0	2185,1	3535,6	1149,5	1249,6	1386,4	6,02E-03	0,30	2,92E-03	-0,94
GCW_02530	smpB	666,1	530,7	776,9	684,5	266,1	180,3	173,4	587,0	848,7	611,4	3,17E-02	0,30	3,45E-07	1,77
GCW_02535	GCW_02535	743,6	605,3	640,3	607,8	414,0	471,8	433,8	1085,2	1487,3	1029,5	4,51E-01	-0,10	4,33E-07	1,47
GCW_02540	GCW_02540	1730,2	1190,1	1034,8	970,5	1381,7	1708,6	2005,0	1797,1	2001,3	1895,0	5,90E-02	-0,52	1,63E-01	0,21
GCW_02545	rpmG	1783,5	1466,5	873,6	979,5	2154,0	2300,0	1068,8	1174,2	1292,1	1669,8	8,08E-07	-0,81	2,99E-01	-0,32
GCW_02560	arcA_2	456,6	388,4	1783,8	1577,6	545,8	563,8	818,4	543,6	528,8	633,4	2,70E-84	1,99	5,85E-01	-0,12
GCW_02565	udk	302,5	374,9	175,0	174,2	198,5	173,7	258,9	122,6	99,0	110,3	4,74E-06	-0,95	3,45E-05	-0,87
GCW_02570	GCW_02570	92,9	90,2	68,1	70,7	77,7	93,9	125,4	78,2	74,3	70,6	7,02E-06	-0,40	9,12E-02	-0,35
GCW_02580	GCW_02580	417,3	489,8	220,3	266,2	99,5	64,9	86,5	356,0	489,8	234,2	2,74E-05	-0,90	1,48E-04	2,10
GCW_02585	GCW_02585	2676,0	3435,2	823,0	879,9	1489,1	1446,1	1572,5	2070,9	2768,4	2475,3	4,61E-11	-1,83	7,11E-06	0,73
GCW_02590	pepA	2458,3	1787,8	800,4	911,7	6561,7	6552,4	12231,7	2412,4	2617,5	3563,0	1,77E-06	-1,29	5,07E-04	-1,48
GCW_02595	GCW_02595	137,6	151,1	144,9	136,6	167,9	141,5	144,2	84,7	104,6	124,9	6,48E-01	-0,03	1,50E-02	-0,50
GCW_02600	GCW_02600	2,1	1,1	30,0	18,5	0,0	0,0	1,1	2,5	3,0	2,0	5,74E-10	3,96	1,95E-02	5,03
GCW_02605	VlhA.2.01	726,2	756,4	1736,0	1638,1	863,6	672,8	671,6	595,1	576,1	848,8	2,92E-129	1,19	7,35E-01	-0,10

GCW_02610	GCW_02610	847,6	868,3	1617,2	1848,0	604,6	451,9	609,5	814,6	982,8	1076,6	8,09E-10	1,01	9,99E-06	0,83
GCW_02615	VlhA.2.02	1247,9	1084,3	2011,0	2219,6	1136,9	1199,3	1661,7	1196,2	1369,0	1793,1	4,64E-10	0,86	5,35E-01	0,16
GCW_02620	disA	251,6	275,9	170,1	172,4	193,2	323,9	265,7	492,7	465,6	358,1	9,42E-09	-0,62	2,67E-03	0,81
GCW_02625	rnr	609,5	632,2	255,0	253,2	741,6	825,0	880,5	977,5	936,0	869,3	5,01E-75	-1,29	1,18E-02	0,23
GCW_02630	secG	582,5	530,2	183,2	195,6	601,4	631,2	856,9	504,2	437,5	485,3	2,98E-89	-1,55	5,35E-03	-0,50
GCW_02635	GCW_02635	162,0	186,6	121,1	119,5	72,2	85,3	109,6	240,9	224,4	160,6	1,19E-04	-0,53	1,45E-05	1,27
GCW_02640	trxB_2	441,6	467,8	160,5	176,2	1008,9	730,7	976,6	476,6	424,3	517,4	4,53E-35	-1,43	3,72E-06	-0,89
GCW_02645	GCW_02645	507,1	461,0	165,7	171,9	353,6	322,2	740,0	268,2	266,0	327,1	7,48E-115	-1,52	1,79E-01	-0,58
GCW_02650	GCW_02650	111,6	151,1	201,4	176,2	91,4	60,8	80,2	57,7	70,5	76,7	3,41E-02	0,54	6,56E-01	-0,13
GCW_02660	ygjD	412,8	480,1	444,4	436,1	96,7	62,0	86,2	269,8	345,3	214,3	7,86E-01	-0,02	2,52E-06	1,80
GCW_02665	potD	478,2	553,0	392,2	417,5	541,2	666,6	270,0	566,8	592,9	519,7	1,96E-02	-0,35	5,67E-01	0,32
GCW_02670	potC	439,9	560,7	313,2	353,0	669,4	838,2	326,0	650,6	861,1	746,5	1,22E-02	-0,58	3,61E-01	0,44
GCW_02675	potB	441,7	479,9	297,3	338,8	683,2	860,0	280,9	767,1	1005,4	850,0	2,51E-04	-0,54	1,63E-01	0,70
GCW_02680	potA	724,2	784,1	366,0	430,8	1144,2	1167,1	512,7	821,3	824,0	870,8	7,41E-09	-0,92	6,86E-01	-0,03
GCW_02685	GCW_02685	193,2	213,1	199,9	205,9	95,2	102,5	151,6	168,4	158,3	152,9	8,58E-01	0,00	1,41E-02	0,53
GCW_02690	GCW_02690	130,7	140,6	161,2	193,1	237,7	375,6	394,2	310,6	280,1	345,9	7,52E-02	0,38	7,41E-01	-0,04
GCW_02695	recA	59,5	94,5	102,5	129,3	76,7	74,7	60,5	48,5	37,0	51,5	1,93E-01	0,62	1,35E-03	-0,60
GCW_02700	GCW_02700	631,4	547,3	1072,7	1060,8	727,3	632,0	921,6	426,5	378,8	439,9	1,18E-36	0,86	1,86E-05	-0,82
GCW_02705	msrB	212,2	251,8	154,5	197,0	1424,6	1259,7	1650,9	426,9	296,1	985,8	8,69E-02	-0,41	1,59E-02	-1,48
GCW_02710	groEL	1561,5	1344,1	436,3	449,7	3881,2	3743,1	5378,9	2386,4	2871,7	3313,3	2,69E-53	-1,71	2,21E-02	-0,56
GCW_02715	groES	1026,6	895,7	271,1	244,2	2362,6	1308,0	1482,3	1713,0	1990,6	1982,2	1,98E-80	-1,90	5,28E-01	0,23
GCW_02720	pykF	1318,9	1171,2	474,4	506,7	2782,8	2499,7	4609,8	1788,5	1861,1	2119,3	1,50E-46	-1,34	3,29E-02	-0,69
GCW_02725	pfkA	1352,0	1276,9	473,4	534,4	2329,8	1482,8	3287,2	1263,2	1249,0	1371,5	3,79E-48	-1,38	3,31E-02	-0,76
GCW_02730	lplA	703,4	555,1	253,9	217,2	759,8	628,2	735,8	570,4	698,7	625,6	3,12E-13	-1,41	5,82E-01	-0,13
GCW_02735	lpd	3288,7	2159,0	1036,7	1133,1	1700,1	1501,2	2840,9	2112,5	2566,5	2368,7	2,43E-04	-1,30	3,43E-01	0,31
GCW_02740	aceF	3772,3	2304,4	1041,4	1188,7	5634,6	5146,8	11692,4	2349,0	2353,3	2999,7	1,06E-03	-1,41	4,70E-03	-1,41
GCW_02745	acoB	5143,8	2675,1	1476,5	1576,8	5073,4	4815,9	9958,3	3294,2	3880,3	3612,4	2,04E-02	-1,28	6,25E-02	-0,76

GCW_02750	acoA	5280,6	2790,6	1349,1	1443,1	6213,2	5484,4	11891,3	3001,2	3582,3	3328,2	8,73E-03	-1,46	1,09E-02	-1,13
GCW_02755	GCW_02755	4878,5	2455,3	1164,9	1303,6	6400,9	6017,7	13797,3	3123,7	3438,5	3457,7	1,36E-02	-1,49	1,18E-02	-1,24
GCW_02760	ackA	806,7	827,2	844,0	783,6	2690,2	1962,1	3322,7	889,5	673,5	1210,6	8,91E-01	-0,01	3,20E-05	-1,49
GCW_02765	trkG	165,1	188,2	615,4	655,5	239,5	304,2	279,0	304,4	257,3	261,4	1,31E-39	1,85	8,87E-01	0,04
GCW_02770	GCW_02770	142,8	190,7	460,5	475,8	70,0	155,9	107,7	218,2	223,7	198,1	4,12E-14	1,50	4,37E-04	1,05
GCW_02775	GCW_02775	199,7	210,4	62,1	57,6	76,8	72,1	83,1	57,4	59,3	55,0	2,04E-87	-1,78	4,78E-06	-0,39
GCW_02780	GCW_02780	87,2	82,2	38,8	29,4	41,9	72,1	24,3	33,2	39,2	28,7	1,87E-21	-1,32	4,21E-01	-0,29
GCW_02785	clpB	787,6	827,8	4267,4	3721,0	1342,7	1961,3	2703,1	3637,4	4630,3	8530,9	1,78E-94	2,30	1,54E-02	1,48
GCW_02795	plsX	197,5	206,7	129,2	113,1	95,5	123,8	114,6	295,3	326,1	275,2	3,78E-16	-0,74	1,16E-23	1,47
GCW_02800	GCW_02800	514,1	554,7	236,5	214,3	506,1	533,7	660,3	578,6	618,7	523,1	1,57E-31	-1,25	6,51E-01	0,06
GCW_02805	GCW_02805	879,2	927,4	232,2	290,0	403,1	401,8	463,0	478,0	428,4	409,2	4,46E-27	-1,80	4,46E-01	0,09
GCW_02810	rluA_2	64,8	92,1	57,7	46,6	17,8	24,2	23,7	70,4	88,9	41,1	6,46E-02	-0,57	1,52E-03	1,59
GCW_02815	spoU	104,4	123,9	70,6	51,1	45,2	60,6	45,9	150,8	164,6	108,4	2,88E-04	-0,92	9,26E-07	1,51
GCW_02820	GCW_02820	156,4	239,3	109,8	114,5	181,2	132,1	122,8	203,6	195,0	171,3	2,32E-02	-0,79	4,33E-02	0,44
GCW_02825	thil	100,2	139,9	84,3	89,4	99,3	108,1	141,8	129,7	99,6	138,5	8,70E-02	-0,45	7,31E-01	0,12
GCW_02830	GCW_02830	133,6	204,7	96,1	101,4	170,5	100,5	182,0	70,1	56,1	72,7	3,03E-02	-0,74	4,42E-04	-1,11
GCW_02835	GCW_02835	62,0	52,2	70,7	62,6	71,7	58,4	127,3	31,3	25,8	55,0	2,89E-02	0,23	3,74E-02	-1,15
GCW_02845	GCW_02845	1703,7	1582,4	875,3	1123,4	716,3	928,0	830,3	801,0	1340,9	741,4	1,58E-04	-0,73	4,45E-01	0,22
GCW_02855	GCW_02855	78,4	113,4	40,5	37,5	133,4	136,1	93,4	134,0	95,8	77,5	2,13E-04	-1,27	4,70E-01	-0,22
GCW_02860	eno	2832,1	2227,6	1728,7	1807,6	5069,4	4073,6	8086,9	2268,9	2464,3	3204,7	2,96E-03	-0,51	6,99E-03	-1,03
GCW_02865	GCW_02865	63,8	89,5	243,1	210,7	70,8	62,6	45,4	50,1	64,7	64,6	5,69E-12	1,58	7,88E-01	0,06
GCW_02870	GCW_02870	169,7	210,2	42,5	42,5	137,2	70,7	375,1	28,0	36,6	63,6	2,98E-17	-2,15	4,29E-02	-1,90
GCW_02875	GCW_02875	283,1	349,2	512,4	489,7	316,4	429,0	310,4	375,3	397,8	284,4	5,59E-05	0,67	8,59E-01	0,04
GCW_02880	nusB	338,6	309,4	380,7	376,0	142,4	203,4	124,3	351,9	380,8	274,2	4,16E-08	0,23	4,80E-05	1,16
GCW_02885	GCW_02885	414,6	298,2	291,2	272,3	242,1	314,8	214,9	296,0	330,0	345,0	1,73E-01	-0,32	2,75E-02	0,39
GCW_02890	efp	819,6	742,5	518,4	640,4	606,4	386,1	641,8	388,5	417,0	450,7	1,51E-02	-0,44	2,23E-01	-0,31
GCW_02895	GCW_02895	848,3	247,2	662,3	575,8	2828,9	4966,3	698,8	2019,3	2035,7	2772,2	7,24E-01	0,43	6,51E-01	0,11

GCW_02910	dppF/oppF_1	242,5	249,3	128,5	114,9	255,0	369,6	423,1	416,2	396,0	269,2	9,26E-46	-1,02	7,86E-01	0,09
GCW_02915	dppD/oppD_1	346,2	484,4	128,3	141,1	272,3	311,3	492,8	513,6	428,2	344,7	2,92E-06	-1,60	4,05E-01	0,33
GCW_02920	dppC/oppC_1	463,6	586,2	127,9	142,1	525,7	519,1	831,9	498,2	412,3	379,8	1,21E-12	-1,95	8,13E-02	-0,47
GCW_02925	dppB/oppB_1	441,7	478,2	91,3	93,3	517,6	351,3	553,2	318,1	336,8	244,0	2,83E-73	-2,31	3,53E-02	-0,61
GCW_02930	GCW_02930	662,3	704,1	101,0	123,4	2169,4	2090,2	2909,4	1167,3	840,5	812,6	6,49E-62	-2,61	4,44E-08	-1,31
GCW_02935	GCW_02935	1404,5	903,5	3868,0	2791,6	1057,6	952,5	580,2	1792,8	2013,8	1795,9	5,64E-06	1,55	1,45E-07	1,20
GCW_02940	dppF/oppF_2	3155,8	2240,6	2258,5	2291,3	1173,6	1788,1	869,7	1110,8	2770,6	1783,0	3,49E-01	-0,22	2,41E-01	0,57
GCW_02945	dppD/oppD_2	4215,6	3167,1	2452,9	2682,5	1482,7	1687,3	807,9	1233,6	2570,6	1517,8	1,82E-02	-0,51	3,28E-01	0,46
GCW_02950	dppC/oppC_2	4459,7	2712,0	2153,4	2245,8	1732,4	2603,8	1282,0	1392,7	2959,2	1594,3	7,99E-02	-0,66	7,31E-01	0,10
GCW_02955	dppB/oppB_2	4297,7	2595,4	1776,9	1806,0	2000,1	3373,6	1874,0	1618,2	3697,1	1994,4	2,35E-02	-0,90	8,15E-01	0,01
GCW_02960	oppA	4656,5	2996,0	1596,4	1664,4	1899,4	3067,6	1794,3	1341,5	2514,8	1465,5	9,61E-04	-1,20	6,29E-01	-0,32
GCW_02965	GCW_02965	3764,3	2237,0	860,4	863,0	1643,2	3256,0	1517,1	1317,7	2406,9	2052,2	2,08E-04	-1,75	8,87E-01	-0,07
GCW_02970	GCW_02970	0,8	0,7	7,1	1,6	0,0	0,0	0,0	0,0	1,2	0,0	5,22E-02	2,12	2,23E-01	1,98
GCW_02975	rpsB	1557,1	1296,0	2978,3	2857,4	1618,3	1357,7	2060,8	811,0	949,3	1045,0	1,14E-34	1,04	5,81E-04	-0,79
GCW_02980	azoR	613,6	623,3	145,5	167,5	2306,1	2138,1	3400,8	895,2	671,8	2021,1	6,34E-74	-1,98	5,85E-02	-1,22
GCW_02985	GCW_02985	283,5	350,1	852,8	864,2	263,8	201,6	175,9	217,8	246,0	509,6	1,16E-21	1,45	3,00E-01	0,55
GCW_02990	rarA	104,2	124,0	264,5	270,7	80,2	86,5	144,6	130,8	141,2	153,3	9,66E-19	1,24	5,98E-02	0,54
GCW_02995	rsml	166,7	156,1	226,5	236,1	86,5	109,7	124,2	126,1	138,7	144,7	2,20E-13	0,52	4,50E-03	0,41
GCW_03005	GCW_03005	2788,8	1672,5	2928,8	2978,9	5582,2	5804,0	5580,2	3296,6	4072,2	5705,6	1,75E-01	0,45	2,18E-01	-0,37
GCW_03010	rpmB	4118,8	2749,7	4569,0	4276,6	2259,5	1889,0	1146,3	1158,9	1527,2	1726,3	1,10E-01	0,39	5,71E-01	-0,19
GCW_03015	GCW_03015	183,4	180,9	33,8	28,1	577,8	581,2	717,6	328,7	232,6	642,3	2,03E-276	-2,56	2,39E-01	-0,73
GCW_03020	pvpA	2893,2	2787,6	3986,9	3628,8	1418,9	1247,5	2041,0	1155,8	1100,1	1176,4	3,42E-25	0,42	9,71E-02	-0,38
GCW_03025	fusA	1795,3	1516,2	2527,2	2163,6	1877,8	1882,4	2310,0	1310,4	1286,3	1140,1	2,71E-05	0,50	5,38E-07	-0,66

GCW_03030	rpsG	1203,2	1065,6	1796,3	1655,6	1746,7	1605,0	1525,1	727,3	671,4	652,0	5,86E-24	0,61	1,91E-41	-1,21
GCW_03035	rpsL	1603,7	2077,5	2573,4	2676,9	1850,7	1986,5	1559,9	933,2	697,0	811,9	2,18E-02	0,52	5,90E-13	-1,11
GCW_03040	rpsD	1323,8	930,6	1877,1	1396,3	1192,6	890,4	885,9	726,3	735,7	647,8	4,41E-02	0,55	2,12E-02	-0,44
GCW_03045	GCW_03045	1028,2	1146,0	1508,6	1521,0	3265,4	2439,6	1375,0	1430,5	1285,1	1710,7	1,54E-05	0,48	1,03E-01	-0,56
GCW_03070	GCW_03070	361,0	426,0	3697,7	3066,4	173,2	197,6	321,8	601,1	836,0	688,1	6,38E-55	3,10	6,32E-08	1,70
GCW_03075	dnaK	3986,4	2661,8	7556,4	6021,0	4283,7	5001,0	12980,2	5554,5	7616,8	9163,3	8,15E-05	1,05	8,68E-01	0,20
GCW_03080	GCW_03080	39,3	44,9	15,3	12,0	37,2	31,6	22,7	18,6	28,3	19,9	7,07E-17	-1,63	2,90E-01	-0,41
GCW_03085	GCW_03085	49,6	58,8	260,2	244,0	86,3	62,7	40,2	33,2	33,9	61,7	4,26E-100	2,22	2,94E-01	-0,51
GCW_03090	mscL	212,2	311,0	658,7	758,7	257,5	176,3	116,6	121,7	156,9	256,2	2,04E-06	1,46	9,94E-01	0,00
GCW_03095	GCW_03095	196,9	263,0	1174,6	1329,3	502,4	334,7	341,9	253,7	250,9	246,5	2,30E-23	2,46	1,23E-02	-0,59
GCW_03100	ptsG_2	324,5	285,2	937,9	928,7	878,8	332,5	558,0	198,6	199,7	243,6	1,04E-178	1,62	7,45E-03	-1,32
GCW_03105	ppnK	426,0	510,5	1187,0	1281,2	303,0	314,5	336,1	323,1	297,1	290,5	2,09E-16	1,40	6,63E-01	-0,03
GCW_03110	spxA	726,6	587,2	1028,4	1020,7	663,9	411,7	752,0	221,5	211,7	320,7	1,19E-08	0,65	6,82E-04	-1,22
GCW_03115	trpS	371,0	427,4	337,9	404,8	766,5	544,8	1029,4	384,0	360,9	573,9	5,14E-01	-0,11	4,24E-02	-0,77
GCW_03120	GCW_03120	404,3	466,5	367,9	394,9	1168,5	1310,4	2112,6	1043,8	945,5	1286,7	1,76E-01	-0,19	1,88E-01	-0,41
GCW_03125	GCW_03125	257,7	238,1	790,9	663,0	94,2	118,8	125,3	278,6	352,8	269,1	1,89E-33	1,55	3,50E-10	1,45
GCW_03130	GCW_03130	495,2	500,9	945,8	1006,7	198,4	204,8	157,8	400,4	522,1	342,9	5,12E-23	0,97	7,12E-05	1,20
GCW_03135	GCW_03135	1077,8	1070,8	1128,0	1116,7	795,8	1121,6	1619,9	934,0	1091,0	865,4	9,93E-02	0,06	5,47E-01	-0,20
GCW_03140	GCW_03140	96,4	97,0	327,0	260,3	53,7	76,5	77,2	86,7	83,4	55,1	4,07E-19	1,59	6,04E-01	0,15
GCW_03145	lepA	270,3	299,1	493,6	458,0	85,7	128,5	149,8	254,8	256,0	181,4	2,63E-15	0,74	9,33E-04	0,99
GCW_03150	GCW_03150	232,8	214,1	336,8	258,0	257,6	225,0	312,2	210,3	221,5	171,4	1,55E-02	0,40	1,14E-01	-0,35
GCW_03155	GCW_03155	279,1	319,8	389,2	354,4	105,4	193,7	301,5	341,1	370,5	210,9	1,83E-02	0,31	1,72E-01	0,74
GCW_03160	GCW_03160	420,4	359,1	390,1	302,9	240,5	440,0	477,3	696,1	722,7	573,5	4,45E-01	-0,18	2,47E-03	0,88
GCW_03165	GCW_03165	767,7	759,3	536,8	475,7	1817,0	2154,9	2401,0	1635,1	1620,8	1759,9	6,07E-25	-0,59	7,34E-03	-0,30
GCW_03170	GCW_03170	504,4	525,7	271,7	256,2	813,2	990,2	975,3	1070,8	958,0	1090,9	5,47E-45	-0,96	5,19E-02	0,21
GCW_03180	GCW_03180	339,2	474,4	430,0	409,4	749,7	934,0	1159,3	600,0	498,7	747,7	9,66E-01	0,07	2,42E-02	-0,58
GCW_03185	GCW_03185	28,2	32,9	209,1	177,3	27,2	36,2	17,0	46,1	35,0	42,4	6,01E-58	2,66	4,99E-02	0,72

GCW_03195	GCW_03195	379,2	366,6	303,5	286,6	406,5	426,2	567,7	498,0	485,5	681,3	2,30E-35	-0,34	2,75E-01	0,29
GCW_03200	GCW_03200	239,4	242,5	277,0	265,4	241,8	430,0	424,8	447,4	472,1	436,6	8,80E-06	0,17	1,26E-01	0,39
GCW_03205	GCW_03205	504,3	635,4	390,6	367,1	1310,2	1395,0	1509,0	781,7	684,6	1063,4	4,54E-03	-0,58	8,85E-04	-0,72
GCW_03210	GCW_03210	82,1	72,5	104,8	76,5	140,6	86,1	61,3	84,0	63,8	97,9	2,32E-01	0,21	6,28E-01	-0,13
GCW_03215	GCW_03215	10,1	9,3	84,6	64,5	2,6	2,7	1,6	7,3	7,5	6,8	5,61E-27	2,92	1,26E-14	1,72
GCW_03220	valS	183,9	170,5	167,7	128,5	177,7	237,1	175,5	197,0	179,9	137,6	1,54E-01	-0,27	4,87E-01	-0,16
GCW_03225	ysxC	308,2	292,4	186,1	151,8	251,5	200,1	258,0	307,1	282,7	230,1	2,38E-15	-0,84	2,23E-01	0,25
GCW_03230	GCW_03230	714,9	1047,3	320,0	375,6	1231,4	979,9	1386,5	494,2	368,4	463,9	3,67E-04	-1,32	5,04E-11	-1,40
GCW_03235	tktA_2	183,6	194,6	164,1	155,5	179,6	221,8	124,5	177,3	185,6	177,4	7,82E-04	-0,24	7,78E-01	0,12
GCW_03240	GCW_03240	227,6	219,7	57,2	44,9	92,0	96,5	65,9	108,5	89,3	75,3	4,23E-102	-2,14	6,74E-01	0,14
GCW_03245	umuC	23,0	25,9	24,3	20,7	22,6	23,8	15,0	12,8	17,7	14,4	4,01E-01	-0,12	1,62E-01	-0,40
GCW_03250	GCW_03250	125,9	71,8	119,1	83,1	40,5	60,6	35,0	126,1	136,5	137,5	8,08E-01	0,07	4,80E-18	1,63
GCW_03255	prsA	377,3	367,7	306,4	246,9	713,2	839,8	688,5	456,4	361,3	598,9	5,77E-04	-0,44	8,96E-03	-0,65
GCW_03260	GCW_03260	1,5	1,2	4,3	1,4	0,0	0,0	1,2	0,9	0,0	0,0	1,64E-01	0,87	8,23E-01	-0,11
GCW_03265	truA	267,5	310,3	340,7	286,2	119,0	149,1	151,9	397,1	515,3	324,7	5,06E-01	0,12	3,30E-06	1,58
GCW_03270	cbiQ	404,0	378,8	369,5	315,4	383,9	383,7	549,5	459,6	551,8	392,7	3,95E-02	-0,20	5,54E-01	0,14
GCW_03275	cbiO	745,8	753,8	540,0	474,0	540,1	465,9	522,1	639,2	636,8	523,0	3,46E-15	-0,57	6,25E-02	0,27
GCW_03280	cbiO	1137,1	1060,0	603,7	564,1	811,0	639,6	1221,9	1043,5	1211,2	1011,4	3,26E-212	-0,91	2,29E-01	0,38
GCW_03285	gpml	1207,6	1168,5	638,6	610,0	1725,0	1861,1	4053,9	1393,5	1340,4	1525,5	0,00E+00	-0,93	1,14E-01	-0,69
GCW_03290	tpiA	1202,3	1062,2	623,1	574,2	3057,5	2182,7	3685,5	1207,5	1259,3	1557,2	5,55E-35	-0,92	9,64E-05	-1,09
GCW_03295	cpsG/manB	559,3	616,9	263,1	240,4	776,8	612,9	958,2	557,9	526,2	566,8	2,73E-23	-1,22	2,82E-02	-0,45
GCW_03300	cdd	675,3	752,7	295,1	316,5	581,7	413,4	493,2	468,1	444,2	421,1	1,47E-17	-1,22	5,47E-01	-0,11
GCW_03305	deoA	566,6	480,9	153,8	139,1	217,9	172,8	223,5	454,0	542,8	339,2	4,30E-49	-1,83	2,73E-04	1,14
GCW_03310	deoC_2	898,8	764,2	129,5	118,4	1847,8	1881,5	2184,1	999,1	1000,4	971,5	4,69E-86	-2,74	6,61E-29	-0,95
GCW_03315	deoD	789,8	706,3	120,6	106,1	2145,6	2074,1	2816,0	1229,6	1171,8	1305,9	5,40E-206	-2,72	1,12E-07	-0,87
GCW_03320	GCW_03320	619,7	694,6	90,2	98,5	680,6	638,9	660,7	618,1	669,6	584,3	5,10E-53	-2,80	8,15E-01	-0,04
GCW_03325	GCW_03325	496,2	554,6	49,4	52,2	533,5	662,7	542,6	617,3	737,7	628,8	9,63E-70	-3,37	1,10E-01	0,23

GCW_03330	GCW_03330	610,8	643,8	49,6	51,8	595,7	515,0	401,9	483,2	466,1	531,1	6,59E-188	-3,63	9,95E-01	0,02
GCW_03335	GCW_03335	99,1	114,8	33,3	40,6	597,6	802,6	701,7	413,6	397,5	563,7	9,46E-13	-1,54	5,62E-03	-0,58
GCW_03340	GCW_03340	30,1	47,2	47,9	75,6	29,2	30,9	11,5	89,8	91,5	61,2	2,73E-01	0,68	1,16E-05	1,91
GCW_03345	GCW_03345	305,0	363,7	119,2	131,3	110,5	147,7	117,9	240,7	277,1	221,1	6,12E-12	-1,41	4,44E-08	1,02
GCW_03350	GCW_03350	10679,9	4809,9	489,0	666,4	8657,2	11402,3	17604,0	7861,3	11620,6	10397,7	3,08E-04	-3,65	5,54E-01	-0,25
GCW_03360	GCW_03360	272,4	269,6	73,7	81,2	522,4	615,9	639,3	368,3	369,3	502,7	1,51E-113	-1,81	8,45E-03	-0,49
GCW_03370	GCW_03370	256,6	260,0	67,2	72,9	537,5	668,6	498,0	398,3	381,5	476,6	5,96E-115	-1,88	1,37E-02	-0,40
GCW_03380	GCW_03380	1559,1	1376,6	1425,6	1519,2	3200,0	2572,6	2220,1	1163,6	1000,3	1083,0	9,82E-01	0,01	3,78E-10	-1,25
GCW_03385	recR	218,9	196,4	611,6	575,7	80,2	74,5	101,4	130,0	138,2	118,5	2,13E-303	1,52	4,88E-05	0,64
GCW_03390	ybaB	426,8	477,5	790,5	900,1	563,9	581,5	1281,4	223,4	187,4	254,5	9,29E-07	0,90	1,32E-03	-1,73
GCW_03395	dnaX	184,4	182,1	294,8	250,6	97,6	130,8	124,7	153,8	160,4	168,1	5,28E-09	0,57	4,55E-06	0,50
GCW_03400	GCW_03400	681,1	836,8	898,9	912,2	382,2	374,5	482,1	291,9	299,1	436,3	2,01E-01	0,26	4,07E-01	-0,25
GCW_03405	ruvB	57,3	56,6	458,1	350,2	22,3	23,8	17,7	39,6	49,9	39,4	2,18E-26	2,81	5,41E-07	1,06
GCW_03410	ruvA	94,9	123,2	491,5	501,0	43,7	48,4	26,6	82,3	101,7	63,8	5,41E-42	2,20	2,31E-03	1,12
GCW_03415	GCW_03415	115,9	130,6	396,1	347,7	45,6	60,2	62,8	111,1	131,6	85,8	1,35E-39	1,59	5,54E-04	0,99
GCW_03420	gatB	278,4	323,8	539,0	509,1	246,1	310,1	304,7	423,6	420,3	366,6	4,27E-11	0,80	2,30E-05	0,54
GCW_03425	gatA	325,2	399,4	625,2	660,2	459,4	455,1	504,8	554,8	464,2	430,8	4,54E-06	0,83	7,20E-01	0,06
GCW_03430	gatC	332,3	409,5	724,5	772,9	296,6	268,3	374,3	333,9	333,7	280,4	2,68E-08	1,02	7,31E-01	0,06
GCW_03435	ung	100,5	117,3	531,1	463,9	26,1	30,5	22,0	93,8	94,7	59,1	1,04E-57	2,19	2,83E-06	1,67
GCW_03440	GCW_03440	411,6	527,1	1333,8	1271,0	377,8	500,4	448,9	446,7	356,3	344,6	2,21E-20	1,48	2,69E-01	-0,17
GCW_03445	dnaB	449,2	387,0	946,9	899,0	146,7	206,2	123,4	341,4	384,5	263,6	3,62E-73	1,15	3,43E-04	1,11
GCW_03450	rplI	1463,7	1090,1	1525,0	1441,5	2264,3	1616,2	2582,8	1115,8	1055,4	1129,9	1,82E-01	0,23	6,90E-05	-0,90
GCW_03455	rpsR	881,7	681,6	1243,1	1133,6	1858,0	2129,7	1223,7	766,8	584,3	750,4	1,94E-05	0,62	3,84E-06	-1,24
GCW_03460	ssb	783,8	728,6	1172,8	1084,2	471,8	469,2	425,0	696,2	571,0	635,4	1,10E-58	0,58	1,44E-05	0,51
GCW_03465	rpsF	1001,8	728,4	984,3	1015,0	2993,3	3083,1	1231,1	1059,2	885,4	1451,3	2,73E-01	0,23	1,83E-02	-0,98

GCW_03470	asnA	162,3	224,1	186,2	179,6	1014,9	523,1	446,3	334,2	233,3	619,1	6,78E-01	-0,06	2,78E-01	-0,73
GCW_03475	GCW_03475	164,0	176,2	78,4	76,2	394,9	391,7	427,3	287,6	224,1	287,8	2,15E-32	-1,14	8,66E-06	-0,57
GCW_03480	GCW_03480	65,5	78,9	70,6	59,0	48,0	29,9	21,9	24,7	20,0	36,6	4,11E-01	-0,15	5,84E-01	-0,23
GCW_03485	GCW_03485	57,4	75,0	61,7	51,2	245,8	190,9	138,9	62,3	41,1	86,3	3,40E-01	-0,22	1,00E-04	-1,59
GCW_03490	GCW_03490	879,7	624,3	1608,5	1463,3	1248,1	1256,5	1715,3	1080,1	1379,3	1744,4	1,55E-08	1,05	8,30E-01	0,02
GCW_03495	eutD	1743,6	1370,2	1919,6	1852,2	3423,8	2446,7	5450,4	1644,7	1659,6	2445,1	3,55E-02	0,29	4,13E-02	-0,89
GCW_03500	rbgA	163,7	164,1	1135,8	862,2	19,8	23,9	29,6	56,3	63,8	33,4	5,95E-23	2,59	8,10E-03	1,07
GCW_03505	GCW_03505	492,2	432,2	1720,5	1565,6	174,6	122,4	163,4	135,7	146,4	103,8	9,92E-153	1,83	5,06E-01	-0,21
GCW_03510	rplS	1102,7	798,4	1663,0	1421,3	2671,1	1497,9	780,6	447,1	478,1	532,6	3,49E-04	0,71	7,45E-03	-1,55
GCW_03515	trmD	1607,4	1222,6	2103,4	2005,7	423,9	227,7	166,9	546,9	556,2	505,8	2,81E-04	0,55	5,20E-03	1,12
GCW_03520	rpsP	1894,9	1485,8	2023,5	2014,2	2280,9	1129,5	735,4	575,2	585,2	705,1	6,22E-02	0,27	6,60E-02	-0,96
GCW_03525	rplQ	1559,8	1234,0	1869,2	1654,5	2129,6	1289,5	1096,6	580,8	564,5	568,6	1,48E-02	0,34	4,95E-04	-1,30
GCW_03530	rpoA	1580,4	1103,9	1843,8	1667,8	1999,6	1595,8	1755,1	1030,9	1095,1	1039,0	5,85E-02	0,41	5,46E-08	-0,71
GCW_03535	rpsK	1828,6	1749,9	2078,2	2091,8	2201,2	1531,1	1874,0	1099,0	872,4	970,2	8,16E-09	0,22	1,93E-05	-0,88
GCW_03540	rpsM	1942,7	1479,2	2025,3	2258,7	1560,5	1308,9	1705,2	875,8	841,4	826,8	1,03E-01	0,34	2,62E-09	-0,80
GCW_03545	rpmJ	2425,6	1783,2	2277,9	2626,6	1700,0	2418,7	910,0	1260,3	1063,2	1134,9	3,49E-01	0,23	1,91E-01	-0,39
GCW_03550	infA	1669,6	1057,9	1495,6	1406,1	829,2	2246,1	391,1	1294,2	1359,8	1182,4	6,72E-01	0,13	8,42E-01	0,54
GCW_03560	trxA_2	2682,1	1342,4	953,2	908,1	3919,1	3983,5	7583,3	1769,1	1422,4	2534,3	6,57E-02	-1,03	3,50E-03	-1,37
GCW_03565	topA	516,8	551,4	347,0	300,1	582,1	713,9	992,5	755,7	574,1	580,2	2,05E-11	-0,73	4,07E-01	-0,20
GCW_03570	rpe	254,1	234,2	78,7	63,1	103,7	95,2	57,9	286,0	299,0	179,7	1,02E-66	-1,79	6,13E-05	1,62
GCW_03575	pgi	510,8	631,2	127,0	126,6	1879,4	1688,9	1174,2	816,0	693,9	830,9	1,26E-17	-2,16	2,30E-05	-0,96
GCW_03580	rsgA	86,9	98,3	63,0	42,1	35,3	45,0	52,5	96,2	96,7	59,2	2,60E-03	-0,84	4,92E-03	0,95
GCW_03585	GCW_03585	207,9	241,3	87,4	83,6	120,1	190,8	182,0	250,6	237,2	182,0	9,36E-17	-1,39	5,99E-02	0,50
GCW_03590	ptc1	297,4	360,7	101,3	95,7	484,9	352,2	503,8	246,0	214,3	281,6	1,76E-15	-1,73	2,27E-04	-0,81
GCW_03595	gmk	343,4	428,7	93,2	104,1	427,4	286,4	478,2	198,1	139,0	245,4	1,15E-13	-1,96	4,35E-03	-1,00
GCW_03600	def	364,3	427,2	289,5	314,6	861,7	997,0	932,4	448,9	399,6	730,6	1,83E-02	-0,39	8,05E-03	-0,83
GCW_03605	GCW_03605	1461,9	845,9	415,6	465,2	639,7	531,8	533,9	302,6	440,9	358,5	4,39E-03	-1,34	1,27E-02	-0,60
GCW_03610	mmnG/gidA	264,5	288,4	299,4	264,7	306,2	288,8	214,5	269,8	242,4	317,0	8,21E-01	0,03	8,22E-01	0,08

GCW_03615	rsmG/gidB	181,3	165,7	347,7	271,8	70,4	92,1	73,5	162,0	158,9	183,4	1,02E-06	0,83	2,01E-20	1,14
GCW_03620	GCW_03620	271,7	317,2	792,7	675,6	140,2	112,8	151,6	159,8	152,8	176,1	1,70E-17	1,32	3,31E-02	0,32
GCW_03625	GCW_03625	87,4	79,4	435,3	377,0	102,3	113,4	39,6	160,0	164,0	133,5	4,27E-91	2,28	1,62E-02	1,02
GCW_03630	nusG	453,3	549,1	903,9	947,4	931,2	968,9	711,3	662,3	556,8	657,2	1,18E-07	0,89	4,84E-03	-0,43
GCW_03635	secE	382,2	376,2	532,8	550,7	275,4	279,6	158,4	283,1	273,1	233,7	1,02E-15	0,52	5,44E-01	0,23
GCW_03640	rpmG/rpl33	464,6	395,7	668,6	499,9	188,6	221,7	124,2	240,4	236,7	233,0	2,67E-02	0,43	4,19E-02	0,49
GCW_03645	GCW_03645	113,3	138,8	218,9	221,2	100,2	70,8	143,4	78,1	67,6	102,2	7,81E-07	0,81	4,57E-01	-0,27
GCW_03655	GCW_03655	315,2	338,9	365,5	377,8	357,6	341,4	458,5	304,9	240,2	313,3	9,30E-02	0,19	3,98E-02	-0,39
GCW_03660	GCW_03660	294,1	328,4	347,8	307,8	95,6	150,8	255,2	251,0	217,3	187,3	5,88E-01	0,07	3,10E-01	0,53
GCW_03665	GCW_03665	282,1	246,4	257,7	243,3	128,6	184,3	157,0	262,1	229,7	187,5	3,59E-01	-0,07	1,27E-02	0,57
GCW_03670	GCW_03670	327,7	441,1	328,9	332,4	164,5	222,1	195,8	276,4	179,5	107,3	3,26E-01	-0,20	8,90E-01	-0,10
GCW_03675	atpA	309,8	381,6	287,6	285,9	557,6	716,4	785,8	283,0	140,8	97,6	1,21E-01	-0,26	1,00E-06	-2,07
GCW_03680	atpD_2	341,8	410,1	235,6	232,5	621,7	863,0	899,8	272,7	132,8	113,5	8,45E-05	-0,68	2,98E-09	-2,25
GCW_03685	GCW_03685	203,1	251,5	187,8	175,4	87,1	90,9	86,8	69,7	47,9	22,7	7,75E-02	-0,32	3,21E-02	-1,02
GCW_03690	GCW_03690	734,8	701,3	419,8	377,7	527,7	800,1	801,1	1371,6	1467,4	2222,4	1,27E-96	-0,85	4,16E-04	1,28
GCW_03695	rpoE	958,0	689,1	128,0	110,1	2329,8	1735,0	2653,8	922,7	843,0	1150,2	1,88E-16	-2,77	4,22E-06	-1,16
GCW_03700	fba	1891,0	1647,2	226,0	277,9	4191,9	3890,3	6913,4	1746,4	1407,0	2027,2	1,63E-71	-2,82	5,19E-05	-1,46
GCW_03705	ychF	619,3	540,9	146,6	130,3	553,4	494,0	603,0	694,5	731,0	650,2	1,50E-88	-2,07	8,95E-04	0,37
GCW_03710	tdk	644,7	678,7	224,1	212,4	826,0	381,4	819,3	495,0	490,6	464,3	3,29E-79	-1,60	2,74E-01	-0,36
GCW_03715	rplK	1568,6	1393,1	2047,5	2128,1	4072,4	2506,5	1620,5	913,3	816,9	1226,5	2,87E-09	0,50	4,70E-03	-1,35
GCW_03720	rplA	1408,8	1131,5	1989,9	1842,3	2845,8	2194,1	2256,9	1202,0	1073,5	1311,8	3,05E-07	0,60	2,05E-08	-0,98
GCW_03725	pth	408,9	453,8	810,6	962,1	168,9	196,8	84,5	378,2	360,4	312,6	1,92E-06	1,04	1,16E-05	1,35
GCW_03730	tilS	92,1	79,1	481,0	377,1	10,1	15,4	12,6	61,7	76,8	49,9	1,76E-25	2,32	2,32E-10	2,34
GCW_03735	fruA	174,6	203,6	737,1	867,0	314,0	343,1	324,2	306,7	200,8	163,6	9,88E-18	2,08	6,23E-02	-0,56
GCW_03740	GCW_03740	187,5	158,3	772,7	919,6	129,0	163,6	146,7	165,4	136,7	84,8	3,89E-20	2,29	6,43E-01	-0,19
GCW_03745	manA	90,8	72,2	521,1	485,2	24,4	35,0	19,8	66,2	56,5	36,6	7,53E-187	2,63	1,12E-02	1,04
GCW_03750	GCW_03750	89,3	120,6	792,5	755,6	240,7	284,6	91,0	467,0	451,7	522,4	3,92E-88	2,90	2,14E-04	1,42
GCW_03760	csn1	326,7	386,1	355,2	373,7	972,7	1021,2	939,6	897,0	1069,8	1196,6	9,81E-01	0,04	2,56E-01	0,14

GCW_03765	cas1	104,0	90,6	36,1	17,0	101,8	72,1	206,1	111,8	119,6	71,1	8,02E-07	-1,97	6,97E-01	-0,18
GCW_03770	cas2	70,5	70,7	27,2	14,8	66,6	82,3	35,9	93,5	77,6	48,7	9,46E-09	-1,81	6,00E-01	0,32
GCW_03775	GCW_03775	56,4	35,7	23,3	19,1	161,7	226,2	118,9	136,0	94,3	71,7	5,06E-03	-1,09	5,85E-02	-0,71
GCW_03780	GCW_03780	151,2	151,8	51,1	49,5	90,4	147,7	144,5	204,8	197,9	169,1	6,38E-195	-1,59	3,84E-03	0,65
GCW_03785	cysS	331,2	360,3	115,9	130,1	863,2	933,7	1412,7	632,9	605,4	741,0	1,45E-25	-1,49	1,53E-02	-0,63
GCW_03790	GCW_03790	47,2	77,7	284,8	326,3	5,1	4,7	4,9	151,9	181,3	116,2	2,25E-12	2,33	6,05E-20	4,95
GCW_03795	GCW_03795	65,9	127,2	299,4	331,8	7,3	12,9	11,1	165,3	165,4	110,2	1,93E-05	1,78	6,74E-18	3,87
GCW_03800	GCW_03800	65,1	101,2	285,3	306,4	82,3	132,0	97,0	280,7	227,8	269,3	3,67E-11	1,87	2,54E-11	1,38
GCW_03805	GCW_03805	48,2	61,0	326,9	327,4	164,7	108,3	110,7	109,4	78,8	103,7	8,62E-83	2,59	1,93E-01	-0,34
GCW_03815	hsdS_1	21,8	33,4	318,3	300,7	42,9	30,5	28,1	26,6	22,2	27,3	3,41E-87	3,52	1,14E-01	-0,36
GCW_03820	hsdS_2	21,4	30,2	316,0	309,3	5,7	4,8	3,1	14,6	13,0	9,7	4,38E-135	3,62	5,59E-06	1,51
GCW_03825	hsdR	71,8	88,0	370,9	355,9	63,8	67,3	56,8	79,3	64,0	58,3	1,30E-76	2,19	5,59E-01	0,13
GCW_03835	GCW_03835	217,4	125,9	656,0	504,0	460,7	309,9	156,6	144,0	196,4	355,4	1,48E-07	1,80	5,82E-01	-0,35
GCW_03840	GCW_03840	461,6	468,6	1112,1	1199,3	487,5	598,2	1133,2	518,4	492,2	694,3	4,07E-30	1,31	4,93E-01	-0,26
GCW_03845	GCW_03845	176,0	208,3	680,2	618,5	196,2	212,7	126,8	235,6	204,8	219,8	1,04E-47	1,76	1,50E-01	0,37
GCW_03850	GCW_03850	168,1	181,3	416,7	387,0	86,4	109,1	73,2	166,4	179,7	118,0	1,82E-61	1,20	5,70E-03	0,82
GCW_03860	GCW_03860	1955,6	1264,2	1689,4	1625,9	2490,9	2353,6	1896,7	1620,7	1762,3	1810,6	7,90E-01	0,08	1,54E-02	-0,33
GCW_03865	GCW_03865	118,5	146,2	256,5	261,8	151,8	178,9	110,2	118,6	95,4	166,9	4,19E-09	0,98	5,75E-01	-0,18
GCW_03875	GCW_03875	9,3	5,3	31,1	29,1	0,5	2,4	0,0	4,2	7,0	3,7	1,42E-15	2,08	8,45E-03	4,03
GCW_03880	GCW_03880	10,7	9,7	49,8	51,5	2,3	3,5	0,8	5,6	5,1	3,5	7,35E-60	2,31	3,34E-02	1,34
GCW_03885	GCW_03885	717,9	611,8	1332,1	1155,9	610,1	1038,3	1084,5	1182,9	1234,5	1909,2	2,27E-16	0,91	6,56E-02	0,71
GCW_03890	GCW_03890	250,8	178,8	952,6	611,9	26,2	29,1	25,6	237,1	336,2	235,1	7,18E-06	1,85	3,77E-16	3,34
GCW_03895	GCW_03895	71,7	47,7	336,8	187,4	14,9	13,7	8,4	94,8	134,5	117,7	1,76E-04	2,10	1,16E-23	3,29
GCW_03900	GCW_03900	224,7	241,5	1016,5	935,8	304,9	256,1	141,8	313,3	258,3	406,3	6,12E-191	2,07	2,32E-01	0,56
GCW_03905	GCW_03905	20,8	31,4	22,0	16,2	4,6	3,6	1,4	10,0	12,8	8,8	2,37E-01	-0,44	3,23E-05	1,92
GCW_03910	GCW_03910	30,3	35,3	23,9	20,2	27,3	25,1	5,6	26,4	29,2	16,8	8,69E-04	-0,57	5,67E-01	0,62
GCW_03915	GCW_03915	109,9	122,0	49,2	42,9	126,4	66,1	41,2	41,4	36,5	37,7	8,86E-22	-1,33	8,38E-02	-0,82

GCW_03920	lig	88,2	101,5	37,6	37,6	114,6	109,6	119,7	111,8	110,7	95,3	7,01E-17	-1,33	5,73E-01	-0,08
GCW_03930	msrA	473,6	438,1	111,6	95,0	1341,3	1377,1	2529,7	826,9	859,9	1034,7	2,42E-201	-2,14	2,01E-02	-0,85
GCW_03935	GCW_03935	112,9	152,4	208,2	205,6	86,8	97,8	56,3	115,5	117,6	96,6	5,54E-03	0,66	4,53E-02	0,52
GCW_03940	GCW_03940	406,0	302,1	262,4	205,8	244,6	191,9	382,5	215,8	222,5	341,8	2,50E-02	-0,59	9,73E-01	0,00
GCW_03945	GCW_03945	606,9	674,4	3177,6	2799,5	347,8	592,3	1189,9	691,5	773,2	1186,0	1,39E-83	2,22	5,46E-01	0,50
GCW_03960	GCW_03960	99,7	65,1	1064,0	727,0	40,2	21,9	10,3	219,5	236,7	202,8	5,20E-14	3,45	1,44E-22	3,43
GCW_03965	GCW_03965	320,8	341,7	2494,9	2147,2	399,4	431,2	198,6	496,9	378,7	546,9	1,33E-95	2,81	2,14E-01	0,57
GCW_03970	GCW_03970	22,3	4,8	166,1	70,1	9,3	5,2	0,0	54,4	110,8	82,5	3,17E-03	3,38	1,06E-07	6,34
GCW_03980	GCW_03980	834,1	650,9	1994,0	1697,3	568,4	751,5	697,3	824,0	973,8	1180,5	8,42E-17	1,32	2,94E-03	0,60
GCW_03985	GCW_03985	808,3	817,0	1397,6	1250,3	469,3	809,6	1347,0	1011,1	860,0	1506,8	7,54E-27	0,70	4,85E-01	0,49
GCW_03990	GCW_03990	939,1	875,2	1499,5	1353,7	944,3	1474,0	1653,6	1618,0	1594,7	2492,2	8,50E-36	0,65	1,59E-01	0,53
GCW_03995	GCW_03995	0,0	0,0	18,7	19,5	0,0	0,0	0,0	0,0	0,0	1,4	1,50E-20	8,70	2,35E-01	2,02
GCW_04000	GCW_04000	298,1	371,2	580,1	586,6	837,4	652,2	752,0	391,5	293,5	509,8	3,79E-06	0,81	1,37E-03	-0,90
GCW_04005	GCW_04005	259,8	297,1	436,0	381,0	129,9	223,7	399,5	316,9	310,6	235,2	6,21E-05	0,55	6,08E-01	0,37
GCW_04010	gltX	358,7	336,4	503,9	441,3	398,9	434,5	840,2	370,0	379,7	333,9	7,77E-10	0,44	1,83E-01	-0,51
GCW_04015	fold	235,2	255,9	461,9	407,6	139,5	193,1	196,9	340,1	392,6	294,9	2,86E-15	0,82	6,62E-06	1,00
GCW_04020	lgt	445,5	408,4	616,2	591,1	671,4	703,4	841,5	793,5	877,1	842,8	7,17E-63	0,50	3,21E-02	0,23
GCW_04025	hprK	602,0	564,4	700,9	663,5	920,5	800,4	1292,8	934,0	937,5	1135,0	6,39E-34	0,23	8,53E-01	0,06
GCW_04030	GCW_04030	12,4	27,0	51,6	51,6	0,0	1,1	7,4	7,6	10,1	3,8	2,40E-03	1,49	2,01E-01	3,63
GCW_04035	GCW_04035	452,4	452,1	210,6	206,7	552,6	334,2	362,3	335,9	390,7	363,9	1,03E-202	-1,12	6,67E-01	-0,12
GCW_04045	trmE	92,3	105,8	652,7	619,8	52,6	64,1	61,8	67,3	76,9	47,8	1,19E-252	2,69	5,67E-01	0,12
GCW_04050	holB	64,1	58,8	459,7	421,8	23,6	54,0	38,1	58,7	66,8	66,2	0,00E+00	2,84	7,77E-03	0,84
GCW_04055	tmk	217,7	165,2	797,1	727,5	69,5	88,8	123,1	183,8	219,2	162,4	2,82E-55	2,01	1,41E-04	1,07
GCW_04060	serS	347,1	363,1	1062,6	1198,3	230,2	270,4	424,9	411,6	431,7	420,5	8,46E-23	1,67	4,46E-02	0,54
GCW_04065	gyrA	587,9	657,6	1529,3	1787,4	463,9	608,4	836,1	845,2	904,2	739,4	3,00E-11	1,41	8,59E-02	0,46
GCW_04070	gyrB	558,3	661,5	1835,3	2128,7	995,2	893,5	1307,1	741,9	644,5	698,1	4,76E-14	1,70	2,87E-03	-0,56
GCW_04075	dnaJ_7	411,5	376,3	1615,5	1786,9	315,1	445,9	464,6	534,6	595,1	511,0	5,19E-42	2,11	6,80E-03	0,48

GCW_04080	dnaN	635,4	509,0	1950,4	2032,8	780,3	628,2	892,0	577,5	550,2	610,6	4,09E-46	1,81	3,39E-02	-0,35
GCW_90014	GCW_90014	61,0	35,2	134,0	112,3	20,7	8,8	5,8	54,9	53,8	51,9	7,29E-06	1,41	8,71E-09	2,42
GCW_90015	GCW_90015	4,8	4,6	68,9	78,1	0,9	3,3	0,0	18,3	15,6	6,4	1,04E-58	3,94	1,37E-03	4,97
GCW_90016	salX	9,6	11,0	98,5	94,1	2,9	1,3	1,2	6,2	5,4	2,8	7,86E-314	3,22	2,01E-02	1,48
GCW_90017	cbiO	10,2	15,4	103,7	98,9	1,2	2,7	1,4	9,2	10,9	4,9	9,31E-57	3,01	3,73E-04	2,26
GCW_90066	glpK	252,1	329,0	152,9	183,8	66,8	30,5	53,5	161,4	194,9	128,4	4,20E-03	-0,78	7,46E-06	1,78
GCW_90068	GCW_90068	404,8	390,2	137,2	157,9	941,2	534,8	879,3	230,5	199,0	268,9	3,47E-43	-1,43	5,19E-07	-1,68
GCW_90093	GCW_90093	298,0	293,1	65,9	70,9	135,5	50,0	79,7	105,6	143,8	92,7	6,53E-218	-2,11	3,68E-01	0,50
GCW_90094	GCW_90094	145,0	148,0	62,5	63,7	25,3	20,8	21,9	82,6	117,7	54,6	9,35E-83	-1,21	5,23E-04	1,88
GCW_90102	GCW_90102	867,7	771,4	1171,2	1175,4	1155,8	902,7	1588,8	544,1	452,9	568,3	6,95E-17	0,52	8,97E-05	-1,15
GCW_90103	GCW_90103	502,1	398,1	997,8	940,3	521,5	519,3	427,4	420,1	415,1	404,8	5,16E-23	1,12	4,11E-02	-0,20
GCW_90134	ispF	921,7	743,7	489,4	527,4	1910,7	2101,4	2382,8	1186,1	1040,8	1293,5	1,34E-05	-0,70	6,55E-11	-0,82
GCW_90246	folA	121,5	168,6	99,1	95,9	106,6	133,3	84,7	202,1	197,9	250,8	3,62E-02	-0,55	1,65E-06	1,06
GCW_90292	GCW_90292	1160,1	817,4	174,4	188,2	2128,2	1798,2	471,8	790,0	793,8	748,2	1,14E-12	-2,43	1,22E-01	-0,61
GCW_90422	zip	2745,9	2017,2	460,8	455,2	236,7	170,8	160,7	79,8	96,0	83,5	2,63E-16	-2,36	1,44E-05	-1,08
GCW_90442	GCW_90442	2812,6	1913,4	680,5	628,4	434,3	258,4	338,9	1147,1	1413,8	1079,2	9,90E-08	-1,83	2,16E-11	1,88
GCW_90501	polC	74,4	92,1	478,9	464,0	63,3	88,2	64,7	109,2	88,5	86,6	4,80E-102	2,51	4,38E-02	0,44
GCW_90502	GCW_90502	77,5	84,4	294,5	259,5	56,9	31,3	36,8	55,5	82,1	60,4	3,63E-62	1,77	4,13E-02	0,73
GCW_90524	GCW_90524	24,2	35,2	305,3	332,0	30,2	30,9	10,0	14,4	11,7	14,8	2,24E-49	3,45	9,54E-02	-0,60
GCW_90526	GCW_90526	20,6	21,1	242,3	218,1	17,2	19,1	7,5	13,8	14,1	11,9	0,00E+00	3,46	7,77E-01	0,01
GCW_90537	GCW_90537	484,9	417,9	1120,5	1029,2	526,3	531,9	407,9	388,5	380,7	388,1	5,32E-67	1,25	1,45E-02	-0,29
GCW_90538	GCW_90538	313,5	285,0	839,4	786,2	337,1	301,7	300,7	176,6	161,4	198,1	0,00E+00	1,44	6,04E-12	-0,77
GCW_90556	GCW_90556	105,9	116,2	188,1	162,5	688,4	626,6	483,6	493,5	384,8	588,3	6,20E-08	0,66	2,89E-01	-0,26
GCW_90601	GCW_90601	823,5	934,4	1897,0	1528,1	1428,2	949,0	1237,3	1212,1	1253,7	940,9	1,07E-07	0,96	9,25E-01	-0,04
GCW_90632	GCW_90632	453,1	390,5	874,2	827,7	492,0	448,2	542,6	325,1	305,1	327,2	1,76E-55	1,02	5,68E-10	-0,59
GCW_90633	GCW_90633	266,0	200,1	438,6	356,1	518,6	420,1	390,9	218,9	192,3	247,1	1,12E-04	0,78	1,38E-07	-0,97
GCW_90642	GCW_90642	1,9	3,6	30,4	29,4	0,0	0,0	0,0	0,0	0,3	1,3	3,92E-30	3,51	7,66E-02	3,40

GCW_90643	GCW_90643	266,0	200,1	438,6	356,1	518,6	420,1	390,9	218,9	192,3	247,1	1,12E-04	0,78	1,38E-07	-0,97
GCW_90644	GCW_90644	374,7	348,1	708,2	667,0	481,5	384,2	614,9	244,0	220,1	258,8	2,35E-272	0,93	3,60E-05	-0,97
GCW_90646	GCW_90646	358,6	237,3	768,8	710,3	238,3	256,0	223,4	253,9	275,0	251,2	3,95E-10	1,34	5,03E-02	0,16
GCW_90676	smtA	81,5	114,5	39,0	45,1	58,8	43,1	66,8	240,2	191,4	134,8	2,77E-04	-1,20	1,01E-05	1,77
GCW_90677	GCW_90677	90,5	148,6	57,4	48,5	40,5	71,1	68,4	220,3	184,8	107,9	9,01E-03	-1,13	1,35E-03	1,53
GCW_90851	GCW_90851	2060,6	1904,7	2159,3	2149,1	974,1	1342,5	1761,1	1632,4	1487,6	1354,0	2,34E-03	0,12	5,21E-01	0,21
GCW_90867	GCW_90867	591,7	929,7	1245,7	1484,7	378,8	386,7	507,9	691,0	489,7	675,0	2,92E-02	0,88	2,80E-02	0,58
GCW_90874	GCW_90874	185,2	124,9	194,9	124,2	58,0	29,3	24,4	240,6	256,2	154,2	7,82E-01	0,03	2,28E-07	2,65
GCW_90876	GCW_90876	84,1	80,8	318,2	322,6	37,6	37,4	15,6	96,6	141,0	47,4	4,34E-253	1,96	2,30E-02	1,67
GCW_90877	GCW_90877	360,4	277,8	894,8	858,4	113,0	162,0	190,8	309,3	461,0	353,8	6,68E-34	1,47	2,11E-05	1,32
GCW_90959	GCW_90959	91,5	48,4	76,1	73,8	13,0	19,9	5,3	59,5	95,5	73,0	7,26E-01	0,17	6,32E-08	2,79
GCW_90986	GCW_90986	56,2	69,9	95,7	83,4	85,4	117,7	136,7	65,3	51,6	67,6	8,39E-03	0,51	2,24E-04	-0,83
GCW_90995	GCW_90995	76,2	110,1	116,1	128,2	116,1	242,8	153,5	154,6	95,4	106,6	2,53E-01	0,41	1,88E-01	-0,45
GCW_90997	GCW_90997	352,5	417,7	260,4	291,3	1097,1	342,4	320,1	256,2	201,3	238,8	8,85E-03	-0,48	1,10E-01	-1,06
GCW_90999	smc	260,2	318,2	234,3	245,2	567,5	886,7	274,9	584,6	506,6	464,7	1,21E-01	-0,26	7,46E-01	0,04
GCW_91142	GCW_91142	10,1	5,7	10,5	9,4	4,1	0,0	2,1	5,7	13,6	10,4	3,19E-01	0,38	3,29E-03	4,08
GCW_91176	VlhA	201,8	84,3	28,7	32,7	416,1	515,7	321,4	455,6	546,1	471,2	1,30E-02	-2,09	2,05E-01	0,29
GCW_91181	VlhA.3.07	325,4	245,2	27,4	45,6	1204,8	1358,2	1471,9	533,9	495,6	640,0	2,11E-15	-3,00	5,00E-22	-1,24
GCW_91197	GCW_91197	0,7	0,0	1,5	0,6	2,1	5,7	0,0	5,0	9,9	2,0	2,17E-01	2,48	3,20E-01	2,92
GCW_91202	GCW_91202	1586,2	857,5	835,9	856,5	1576,5	1590,4	2129,0	628,6	592,3	808,5	3,01E-01	-0,46	3,94E-10	-1,34
GCW_91203	GCW_91203	1377,9	929,6	1034,7	975,8	406,9	444,8	378,1	505,4	568,5	682,6	5,50E-01	-0,17	6,99E-04	0,55
GCW_91276	hatC	57,7	40,9	90,5	73,0	28,0	60,5	183,5	91,4	122,2	108,1	1,46E-03	0,74	6,56E-01	0,69
GCW_91278	GCW_91278	328,7	264,0	164,2	118,4	307,5	516,2	222,0	227,8	170,6	319,4	9,91E-06	-1,08	2,65E-01	-0,46
GCW_91316	degV	653,0	439,6	639,9	627,6	688,9	540,9	518,1	848,2	1313,7	1079,4	3,36E-01	0,24	7,08E-04	0,92
GCW_91421	GCW_91421	361,6	361,9	757,2	756,0	286,5	385,5	320,8	386,4	384,8	348,3	3,05E-132	1,07	1,18E-01	0,22
GCW_91423	GCW_91423	249,9	239,1	724,9	714,4	175,5	397,0	141,8	298,8	206,4	292,9	0,00E+00	1,56	8,32E-01	0,33
GCW_91455	GCW_91455	143,3	223,1	73,1	117,3	168,5	259,6	216,4	141,9	80,6	123,1	4,70E-02	-0,95	2,42E-03	-0,88

GCW_91461	GCW_91461	35,9	36,5	68,6	60,6	8,4	4,3	2,2	30,3	34,3	17,3	8,70E-28	0,83	8,12E-05	2,64
GCW_91481	salX	218,9	113,7	128,8	94,6	308,7	364,6	110,1	262,7	290,7	384,3	3,54E-01	-0,51	5,73E-01	0,45
GCW_91507	GCW_91507	945,8	865,9	1383,7	1427,2	1122,3	881,9	1782,0	555,0	456,1	576,9	9,34E-21	0,64	1,70E-03	-1,16
GCW_91508	GCW_91508	585,6	557,5	1142,6	1074,9	746,0	745,7	637,6	443,3	403,7	446,4	1,64E-297	0,96	2,73E-14	-0,68
GCW_91651	GCW_91651	626,5	511,8	807,7	740,8	795,0	648,8	1194,5	403,2	322,4	425,2	4,86E-05	0,45	6,75E-04	-1,12
GCW_91652	GCW_91652	550,6	452,1	1118,4	1032,9	633,3	623,5	568,0	462,2	438,0	454,4	3,68E-30	1,11	1,45E-11	-0,39
GCW_91756	GCW_91756	143,8	156,9	302,2	259,4	51,5	68,2	152,8	85,4	86,1	74,9	1,84E-13	0,90	8,87E-01	0,05
GCW_91851	GCW_91851	345,1	380,3	731,5	642,2	300,4	401,6	694,6	316,6	355,4	366,2	3,51E-16	0,92	3,85E-01	-0,30
GCW_91852	GCW_91852	365,7	381,1	768,1	743,3	536,5	846,4	825,8	463,9	542,5	509,3	7,31E-94	1,02	2,20E-02	-0,48
GCW_91909	GCW_91909	6,2	8,6	95,6	96,8	3,7	0,0	0,0	3,8	4,9	1,8	1,36E-74	3,71	2,41E-01	4,82
GCW_91912	GCW_91912	44,3	31,6	189,3	178,5	23,7	13,0	12,4	76,8	100,3	55,7	1,22E-57	2,30	8,85E-06	2,31
GCW_91946	VlhA	436,5	429,1	94,2	107,2	484,6	617,8	592,8	491,6	522,7	571,6	5,03E-123	-2,11	6,69E-01	-0,05
GCW_91947	VlhA	233,6	204,2	34,2	47,1	644,7	835,1	919,8	337,6	300,0	446,1	7,86E-34	-2,44	1,02E-06	-1,11
GCW_91948	VlhA	79,5	74,1	51,2	46,0	135,7	198,4	138,3	207,4	161,6	243,6	3,01E-31	-0,66	2,11E-01	0,42
GCW_91988	GCW_91988	52,4	83,7	143,4	169,8	59,3	25,9	78,7	68,7	43,0	44,5	1,16E-03	1,24	9,16E-01	0,08
GCW_92036	pheS	465,3	459,7	107,1	140,8	300,3	358,1	552,0	370,9	371,3	384,6	7,79E-30	-1,91	8,64E-01	-0,02
GCW_92037	pheT	314,9	306,9	107,1	121,3	134,1	184,1	243,3	298,7	328,6	273,1	8,45E-52	-1,45	1,95E-03	0,76
GCW_92038	GCW_92038	7,8	8,3	127,7	128,8	0,7	1,8	2,0	7,4	10,0	15,9	5,43E-151	3,98	1,16E-05	3,00
GCW_92039	GCW_92039	15,1	22,1	221,1	300,2	1,2	5,6	1,0	6,6	10,3	13,7	5,77E-13	3,82	6,77E-03	2,41
GCW_92117	GCW_92117	68,7	80,0	87,0	116,2	3,0	1,9	5,3	25,4	24,7	7,2	1,91E-01	0,44	5,62E-03	2,44
GCW_92148	GCW_92148	73,7	100,2	213,3	202,5	88,8	139,5	65,6	286,3	218,6	187,3	4,48E-10	1,27	8,45E-04	1,32
GCW_92149	uvrD	62,0	69,2	160,7	149,2	114,9	174,4	67,1	213,9	178,2	153,6	1,61E-42	1,24	1,16E-01	0,75
GCW_92263	scpB	265,7	351,4	117,8	110,9	836,2	479,9	459,7	392,1	280,4	372,5	3,94E-07	-1,42	5,20E-02	-0,68
GCW_92264	scpA/scpB	164,4	201,8	94,2	84,5	155,6	153,0	87,6	209,7	186,6	155,8	6,74E-07	-1,03	7,76E-02	0,56
GCW_92276	rnhC	44,0	56,9	36,4	41,3	201,8	132,6	82,8	66,0	49,2	109,8	1,03E-01	-0,37	1,21E-01	-0,84
GCW_92301	GCW_92301	11,5	11,1	88,2	76,4	2,0	5,2	2,4	3,9	4,9	2,3	1,50E-122	2,85	6,74E-01	0,31

GCW_92302	GCW_92302	16,4	21,6	78,4	73,7	13,1	5,8	5,5	8,9	8,8	4,4	6,83E-37	2,01	8,87E-01	-0,05
GCW_92396	VlhA	492,2	559,9	439,9	471,4	348,1	601,4	465,6	608,5	656,3	1184,9	1,24E-01	-0,20	1,03E-01	0,80
GCW_92397	VlhA.3.05	561,1	500,6	296,9	330,8	1245,3	1567,1	1318,1	580,2	667,7	1312,8	2,93E-11	-0,76	1,28E-01	-0,74
GCW_92416	VlhA	146,5	172,4	206,3	214,4	184,5	289,1	254,9	210,0	162,5	215,8	1,39E-02	0,40	2,00E-01	-0,26
GCW_92417	VlhA.4.04	92,3	112,7	193,8	204,3	391,8	415,5	145,4	187,6	171,1	212,8	2,93E-08	0,96	1,03E-01	-0,56
GCW_92433	GCW_92433	530,2	386,9	1050,5	979,0	542,6	543,7	434,9	425,1	416,5	413,5	6,87E-13	1,16	2,82E-02	-0,23
GCW_92452	VlhA	40,7	56,2	15,7	18,0	19,5	38,8	24,6	49,9	57,1	60,1	5,68E-06	-1,51	5,32E-05	1,11
GCW_92453	VlhA	407,1	304,0	45,9	68,1	534,0	589,7	805,3	321,8	345,4	439,4	3,13E-14	-2,65	1,99E-03	-0,75
GCW_92454	VlhA	262,6	261,1	114,8	129,5	576,3	746,0	735,1	387,5	330,4	456,5	3,62E-27	-1,10	9,00E-06	-0,77
GCW_92456	VlhA	249,2	271,2	155,7	148,0	288,8	425,0	316,4	336,9	295,3	307,4	9,30E-15	-0,78	5,46E-01	-0,08
GCW_92457	VlhA	352,2	360,6	106,9	120,1	766,4	999,9	899,2	413,0	299,4	443,1	4,47E-56	-1,65	7,09E-10	-1,18
GCW_92547	GCW_92547	807,1	734,3	1149,3	1166,8	999,0	790,6	1575,9	483,1	390,8	510,8	6,64E-24	0,59	1,26E-03	-1,19
GCW_92548	GCW_92548	522,7	376,8	1030,4	955,0	532,6	530,1	429,0	411,4	405,0	401,3	8,64E-12	1,16	1,61E-02	-0,25
GCW_92579	obgE	488,8	620,2	361,8	484,3	241,2	154,6	145,8	338,4	385,2	246,2	1,64E-01	-0,39	1,26E-02	0,89
GCW_92658	GCW_92658	317,7	330,9	280,2	326,0	324,1	321,0	374,9	309,4	259,7	399,5	4,40E-01	-0,10	8,32E-01	-0,05
GCW_92659	GCW_92659	349,3	316,2	226,4	276,9	425,3	369,0	268,3	308,2	309,9	426,0	1,89E-02	-0,41	9,94E-01	0,02
GCW_92786	GCW_92786	157,3	200,6	406,3	331,0	130,9	238,3	114,8	190,7	156,5	122,9	2,86E-06	1,05	9,10E-01	0,05
GCW_92794	GCW_92794	82,3	152,6	204,6	217,3	93,6	96,7	76,6	79,6	61,1	58,7	5,48E-02	0,91	2,82E-02	-0,39
GCW_92846	GCW_92846	1517,7	1353,5	865,5	940,2	765,0	1155,0	948,7	715,3	954,9	609,9	3,68E-11	-0,67	4,31E-01	-0,30
GCW_92854	GCW_92854	1015,4	1174,0	621,7	714,9	1466,6	2012,4	2697,0	965,5	1027,7	986,1	1,10E-04	-0,71	3,47E-04	-0,97
GCW_92861	GCW_92861	0,0	0,5	6,2	1,8	0,0	0,0	0,0	0,8	0,0	0,0	1,47E-02	4,42	4,01E-01	1,76
GCW_92898	GCW_92898	91,7	94,0	92,0	77,0	33,8	45,6	90,9	124,1	114,2	61,7	2,52E-01	-0,14	1,23E-01	0,92
GCW_92996	GCW_92996	26,7	23,5	100,8	79,3	2,2	2,9	9,1	10,7	14,4	6,1	3,59E-20	1,83	1,06E-01	1,37
GCW_93046	GCW_93046	3,5	6,5	15,1	14,3	0,2	0,6	0,2	2,5	3,8	0,5	9,66E-06	1,61	2,75E-02	2,51
GCW_93171	GCW_93171	127,9	126,1	207,3	195,1	125,7	226,7	233,9	239,9	271,5	242,0	1,88E-68	0,66	9,59E-02	0,45
GCW_93187	GCW_93187	25,4	46,1	206,6	196,3	10,1	4,1	9,1	37,0	26,7	28,3	2,08E-18	2,56	2,28E-08	2,10
GCW_93351	VlhA	42,0	33,0	35,3	35,0	11,8	21,3	19,9	29,2	34,1	30,8	6,25E-01	-0,08	5,90E-05	0,91
GCW_93352	GCW_93352	270,7	159,1	23,0	38,1	359,4	452,0	370,9	209,5	291,5	309,5	6,66E-06	-2,81	1,50E-02	-0,52

GCW_93361	VlhA	75,3	101,7	46,6	51,9	132,8	160,8	122,0	441,1	383,3	304,8	2,71E-03	-0,83	3,44E-09	1,47
GCW_93362	VlhA	1194,1	1018,9	219,2	451,0	2029,9	2433,8	5335,0	1475,4	1176,9	1313,0	4,55E-05	-1,81	2,64E-02	-1,14
GCW_93371	VlhA	211,5	208,1	125,7	151,9	381,0	527,4	413,3	343,8	359,0	534,0	9,38E-05	-0,60	8,27E-01	-0,07
GCW_93372	GCW_93372	126,5	33,2	17,6	16,3	156,0	166,8	61,5	220,5	380,2	354,5	9,70E-02	-1,93	1,52E-03	1,44
GCW_93396	GCW_93396	0,6	0,0	3,9	4,1	0,0	0,0	0,0	2,3	1,3	0,0	1,95E-04	4,54	2,85E-02	4,30
GCW_93654	GCW_93654	892,8	694,0	315,9	266,7	57,8	78,2	69,5	293,2	330,9	247,9	8,28E-12	-1,44	1,12E-18	2,12
GCW_93751	GCW_93751	65,9	77,6	670,5	620,6	74,7	65,0	20,7	162,6	136,3	145,8	6,60E-253	3,17	3,86E-05	1,71
GCW_93752	GCW_93752	46,0	35,8	408,6	340,9	51,8	15,5	2,6	25,2	29,8	33,8	2,33E-61	3,20	6,74E-01	1,24
GCW_93776	GCW_93776	129,4	234,8	87,0	87,3	20,8	73,6	51,2	87,5	92,2	89,3	3,35E-02	-1,00	1,68E-02	1,11
GCW_93806	hsdM	377,6	418,4	350,1	349,5	498,1	609,5	603,2	455,4	455,8	487,8	6,90E-02	-0,18	2,76E-03	-0,25
GCW_93807	GCW_93807	80,8	77,4	324,3	327,5	27,5	31,1	32,0	89,9	116,8	65,8	7,67E-311	2,04	5,38E-05	1,59
GCW_93849	GCW_93849	211,0	255,5	248,7	253,4	126,8	129,2	176,1	300,2	357,2	212,4	6,50E-01	0,11	2,21E-03	1,05
GCW_93851	GCW_93851	666,6	633,4	431,0	453,1	278,6	226,0	195,7	723,1	833,7	728,3	1,36E-19	-0,56	1,63E-24	1,76
GCW_93852	GCW_93852	1591,7	1508,2	484,4	583,0	1915,3	1714,3	2160,8	1235,6	1218,1	1771,4	1,61E-33	-1,54	7,88E-02	-0,43
GCW_93874	rlmB	17,0	17,5	161,8	130,9	3,6	2,6	5,2	10,4	18,7	13,8	4,00E-48	3,07	1,74E-05	1,96
GCW_93876	GCW_93876	8,8	11,1	28,5	32,0	5,0	13,8	0,0	8,1	12,3	4,8	7,96E-11	1,60	6,54E-01	2,84
GCW_93902	GCW_93902	121,4	112,6	689,1	564,3	53,3	35,4	10,3	258,0	240,6	245,1	5,16E-44	2,41	4,47E-20	3,24
GCW_93946	GCW_93946	139,8	126,4	3420,6	3182,2	12,9	6,4	8,7	99,8	98,4	64,3	0,00E+00	4,63	5,52E-13	3,30
GCW_93948	GCW_93948	97,0	89,9	3350,3	2709,1	6,9	0,0	5,1	61,9	48,3	52,4	6,50E-70	5,01	1,23E-16	5,93
GCW_93958	GCW_93958	213,7	314,3	3573,7	3237,4	25,0	19,3	11,0	224,1	207,5	124,2	4,88E-101	3,71	1,44E-09	3,40
GCW_93971	GCW_93971	225,7	145,1	1255,3	856,5	33,5	21,5	21,4	203,1	319,4	223,7	1,17E-09	2,52	5,88E-12	3,33
GCW_94036	tatD	83,9	82,0	691,7	576,5	101,0	93,1	72,1	69,4	50,6	59,4	6,17E-68	2,93	6,98E-03	-0,53
GCW_99016	GCW_99016	0,0	1,0	15,8	19,7	2,0	0,0	0,0	3,2	2,8	0,0	5,35E-09	6,31	3,35E-01	2,62
GCW_99017	GCW_99017	15,3	18,2	68,5	69,8	5,3	19,4	9,6	17,6	11,3	9,3	8,42E-41	2,05	8,63E-01	0,34
GCW_99018	GCW_99018	14,5	17,1	90,0	105,5	4,8	13,1	7,8	9,4	15,3	10,5	1,03E-24	2,63	2,92E-01	0,58
GCW_99045	GCW_99045	11,6	22,5	9,9	15,8	5,1	2,7	0,0	11,2	5,6	8,0	4,21E-01	-0,36	4,46E-02	3,63
GCW_99046	GCW_99046	9,0	13,0	9,9	5,2	0,0	0,0	0,9	3,4	3,6	1,6	3,04E-01	-0,58	2,48E-03	5,27

## БЛАГОДАРНОСТИ

Выражаю глубокую благодарность академику РАН, доктору биологических наук, профессору Говоруну Вадиму Марковичу и кандидату биологических наук Фисунову Глебу Юрьевичу за помощь, оказанную при работе над диссертацией, а также коллективу лаборатории протеомного анализа ФНКЦ ФХМ ФМБФ за рабочую атмосферу и научный дискурс. Кандидату биологических наук Матюшкиной Дарье Сергеевне и Третьяк Марине Викторовне за помощь в оформлении диссертации. Отдельная благодарность кандидату химических наук Семашко Татьяне Александровне за ценные советы и моему мужу за понимание и поддержку.