

**Федеральное государственное бюджетное учреждение науки
«Научно-исследовательский институт физико-химической медицины
Федерального медико-биологического агентства»
(НИИ ФХМ ФМБА России)**

На правах рукописи

ГОРБАЧЕВ АЛЕКСЕЙ ЮРЬЕВИЧ

**СИСТЕМА РЕПАРАЦИИ ДНК У БАКТЕРИИ MYCOPLASMA
GALLISEPTICUM**

03.01.04 – биохимия

03.01.03 – молекулярная биология

диссертация на соискание ученой степени
кандидата биологических наук

Научные руководители:

доктор биологических наук,
профессор, член-корреспондент РАН

Говорун В. М.

кандидат биологических наук

Камашев Д. Э.

Москва – 2014

Оглавление

Список сокращений	5
Актуальность работы	6
Цель работы:	8
Научная новизна и практическая значимость работы.....	8
Апробация работы.....	9
Публикации по теме работы	10
1. Введение	11
1.1. Общая характеристика и классификация Молликут	11
1.2. Репарация ДНК и поддержание геномной стабильности у бактерий .	12
1.2.1. Система SOS ответа	13
1.2.2. Гомологическая рекомбинация.....	19
1.2.3. Система эксцизионной репарации нуклеотидов	20
1.2.4. Система эксцизионной репарации азотистых оснований	22
1.2.5. Система репарации мисматчей	25
1.2.6. Устойчивость ДНК к повреждениям и синтез с низкой точностью	27
1.3. Гистоноподобный белок HU	31
2. Материалы и методы	34
2.1. Олигонуклеотиды	34
2.2. Культивирование <i>Mycoplasma gallisepticum</i> и получение клеточного экстракта	35
2.3. Метод торможения ДНК в геле.....	36
2.4. Идентификация белков, связывающих мисматчи, из клеточного экстракта <i>M. gallisepticum</i>	37
2.5. Гидролиз белков в полиакриламидном геле.....	38
2.6. Идентификация белков методом МАЛДИ-МС	38
2.7. Расчет констант диссоциации	39
2.8. Получение чистого препарата целевого белка Hup ₂ из <i>M. gallisepticum</i> в клетках <i>E. coli</i>	41
2.9. Выделение ДНК.....	42
2.10. Комплементационный тест генов hup ₁ и hup ₂ из <i>M. gallisepticum</i> в <i>E. coli</i>	42

2.11.	Выделение РНК и синтез кДНК.....	44
2.12.	Количественная ПЦР в реальном времени	44
2.13.	Капельно-цифровая ПЦР	45
2.14.	Определение числа копий РНК и ДНК	45
2.15.	Определение предельно переносимых воздействий.....	46
2.16.	Определение кинетики роста культуры	46
2.17.	Праймеры и молекулярные зонды для количественной ПЦР	47
2.18.	Получение клеточного экстракта и разделение белков в одномерном полиакриламидном геле	47
2.19.	Хромато-масс-спектрометрия.....	48
2.20.	Анализ масс-спектрометрических данных	49
2.21.	Двумерный гель-электрофорез с дифференциальной окраской цианинами.....	50
2.22.	Сравнительный анализ и реконструкция системы репарации	51
2.23.	Статистический анализ изменений уровней мРНК	52
2.24.	Кластеризация генов по паттернам изменения экспрессии при тепловом шоке.....	53
2.25.	Измерение АТФ	53
2.26.	Флуорометрическое определение скорости образования перекиси водорода	53
3.	Результаты и обсуждение.....	54
3.1.	<i>M. gallisepticum</i> обладает белками, способными распознавать ошибочно спаренные основания в ДНК.....	54
3.2.	Очистка и идентификация мисматч-связывающих белков <i>M. gallisepticum</i>	56
3.4.	Связывание белка mgHU с поврежденной ДНК сильно при физиологическом значении ионной силы	61
3.5.	Связывание различных ДНК-структур, типичных для репарационного пути MMR.....	63
3.6.	Комплементационный тест.....	67
3.8.	Реконструкция системы репарации ДНК у <i>M. gallisepticum</i>	73
3.9.	Представленность транскриптов генов систем репарации в клетке ...	76

3.10. Представленность участников систем репарации в клетке на белковом уровне.....	78
3.12. Протеомное профилирование в условиях теплового шока.....	86
3.13. Тепловой стресс «разгоняет» клеточный метаболизм приводя к повышенному уровню внутриклеточной АТФ, окислительному стрессу, сопровождающемуся повреждениями ДНК.....	89
Заключение	93
Список литературы	96
Приложение 1.	111
Приложение 2.	111
Приложение 3.	113
Приложение 4.	113

Список сокращений

ДНК – дезоксирибонуклеиновая кислота

оцДНК – одноцепочечная ДНК

дцДНК – двуцепочечная ДНК

РНК – рибонуклеиновая кислота

рРНК – рибосомальная РНК

АТФ – аденозинтрифосфат

дАТФ – дезокси АТФ

ПЦР – полимеразная цепная реакция

ЭДТА – этилендиаминтетраацетат

ДТТ – дитиотреитол

ТФУ – трифторуксусная кислота

ПААГ – полиакриламидный гель

АФК – активные формы кислорода

НАД – никотинамид-аденин-динуклеотид

п.о. – пара оснований

а.о. – аминокислотные остатки

SDS – sodiumdodecylsulfate, додецилсульфат натрия

IPTG – isopropyl β -D-1-thiogalactopyranoside, изопропил- β -D-1-тиогалактопиранозид

СТАВ – cetyltrimethylammonium bromide, цетил-триметил-аммоний бромид

CHAPS – 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate

NP40 – nonyl phenoхypolyethoxylethanol

Актуальность работы

Бактерии, принадлежащие к классу Mollicutes, являются наименьшими известными организмами, способными к автономному делению на искусственных питательных средах. Для них характерны: значительная редукция генома, размеры которого варьируют от 580 тыс. до 1.4 млн. пар оснований; низкое содержание Г-Ц оснований (31 % для *Mycoplasma gallisepticum*), отсутствие клеточной стенки и многих метаболических систем [1].

Инфицирование *M. gallisepticum* является причиной хронических респираторных заболеваний у кур и инфекционного синусита у индеек. Возбудитель наносит серьезный экономический ущерб сельскому хозяйству. Согласно официальным данным, ежегодный ущерб от микоплазмоза для сельского хозяйства США составляет 140 млн. долларов [2] и около 780 млн. долларов для мирового производства птицы [3].

Следует также отметить, что в последнее время (в период с 1994) зарегистрированы случаи микоплазмозов у диких птиц, например, зябликов вида *Carpodacus mexicanus* [4]. По этой причине *M. gallisepticum* является важным объектом исследования и с точки зрения экологии.

M. gallisepticum характеризуется малым размером генома (986 тыс. п.о.), а также небольшим числом генов, кодирующих 836 открытых рамок трансляции. Кроме того, согласно филогенетической классификации Молликут, основанной на результатах анализа последовательностей 16S рРНК, *M. gallisepticum* является ближайшим родственником двух человеческих патогенов *Mycoplasma genitalium* и *Mycoplasma pneumoniae* (ветвь pneumoniae) [5]. При этом *M. gallisepticum* обладает сравнительно коротким временем деления – 4 часа (Приложение 1), не патогенна для человека и легко культивируется в жидкой, полужидкой и на твердой средах, в отличие от близкородственных видов. Все вышеперечисленные свойства делают *M. gallisepticum* чрезвычайно интересным и удобным объектом для

системной биологии микоплазм и других исследований современной молекулярной биологии прокариот. Обладая средней частотой однонуклеотидных мутаций [6,7], соизмеримой с таковой для *Escherichia coli*, она имеет геномные зоны, подверженные чрезвычайно высокой изменчивости (10^{-5} мутаций на одну п.о. в геноме за поколение) [8]. Данные зоны несут в своем составе гены, кодирующие поверхностные антигены, и эта изменчивость необходима для быстрой адаптации к иммунной атаке хозяина или даже для возможности заражения новых хозяев, как в случае с *Carpodacus mexicanus* [9].

Поскольку молликуты (в частности *M. gallisepticum*) – это бактерии с минимальным набором генов, при этом способные к размножению в бесклеточной среде, мы полагаем, что имеющийся у них состав репаративных белков является необходимым и достаточным для поддержания целостности генома.

Цель работы:

Целью настоящей работы было определение состава и функциональной активности генов системы репарации ДНК у бактерии *Mycoplasma gallisepticum*.

В ходе работы необходимо было решить следующие задачи:

1. Идентифицировать и охарактеризовать белок *M. gallisepticum*, способный специфически связывать ДНК, содержащую неправильно спаренные нуклеотиды.
2. Провести биоинформатический анализ генома *M. gallisepticum* для выявления полного набора участников системы репарации ДНК.
3. Определить, экспрессируются ли найденные гены на уровне мРНК и белков. Определить количественную представленность исследуемых транскриптов в расчете на клетку.
4. Провести протеомное профилирование и измерение уровня транскрипции исследуемых генов у *M. gallisepticum* в условиях стрессовых воздействий.

Научная новизна и практическая значимость работы

Впервые идентифицирован и охарактеризован белок, способный связывать ДНК, содержащую ошибочно-спаренные нуклеотиды. Показано наличие SOS-ответа (ранее считавшегося отсутствующим) у микоплазм в условиях стресса. Показана активация системы коррекции повреждений ДНК в ситуациях, способствующих мутационному процессу.

Изучение системы репарации ДНК, как участника системы генерирования устойчивости к антибиотикам имеет высокую практическую значимость для здравоохранения и сельского хозяйства. В связи с тем, что микоплазмы являются паразитами человека и сельскохозяйственных животных, полученные данные могут быть использованы в будущем для

создания новых антибиотических препаратов, с учетом специфики организации микоплазм.

Апробация работы

Результаты работы были доложены на конференции Европейского молекулярно-биологического общества «От функциональной геномики к системной биологии (Гейдельберг, 17-20 ноября 2012), на III Международной научно-практической конференции «Постгеномные методы анализа в биологии, лабораторной и клинической медицине», (Казань, 22-24 ноября 2012), на VI-ом Российском симпозиуме «Белки и пептиды» (Уфа, 11-15 июня 2013), 38-ом Конгрессе Федерации европейских биохимических обществ, FEBS (Санкт-Петербург, 6-11 июля 2013).

Объем работы

Диссертационная работа изложена на 141 странице, состоит из введения, обзора литературы, экспериментальной части, результатов и их обсуждения, основных выводов, списка литературы и приложений. Диссертация содержит 7 таблиц и 17 рисунков.

По теме диссертационной работы опубликовано 3 статьи в зарубежных научных журналах. Результаты работы представлены в виде тезисов на 5 российских и международных конференциях.

Публикации по теме работы

1. Kamashev D., Oberto J., Serebryakova M., Gorbachev A., Zhukova Y., Levitskii S., Mazur A., Govorun V. Mycoplasma gallisepticum produces a histone-like protein that recognizes base mismatches in DNA // *Biochemistry*. 2011. V. 50, P. 8692–8702.
2. 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*. 2013. V. 14. № 1. P. 726-737.
3. Vanyushkina A. A., Fisunov G. Y., Gorbachev A. Y., Kamashev D. E., Govorun V. M. Metabolomic Analysis of Three Mollicute Species. // *PLoS One*. 2014. V. 9. № 3. P. 1-16 (e89312).
4. Фисунов Г. Ю., Горбачёв А. Ю., Дёмина И. А., Говорун В. М. Регуляция экспрессии генов у *Mycoplasma gallisepticum* – бактерии с редуцированным геномом. // *Acta Naturae*. 2013. спецвыпуск №1 С. 45-46
5. Горбачев А.Ю., Камашев Д.Э., Говорун В.М. Метод идентификации ДНК-связывающих белков и его применение для изучения системы репарации ДНК Молликут. // III Международная научно-практическая конференция «Постгеномные методы анализа в биологии, лабораторной и клинической медицине» (Казань, 22-24 ноября 2012). С. 153.
6. Фисунов Г.Ю., Горбачёв А.Ю., Алексеев Д.Г., Мазин П.В., Побегуц О.В., Горшкова Т.Н., Израельсон М.А., Ковальчук С.И., Ванюшкина А.А., Карпова И.Ю., Семашко Т.А., Камашев Д.Э., Кострюкова Е.С., Говорун В.М. Системный подход к изучению экспрессии генов в минимальной клетке. // VI Российский симпозиум «Белки и пептиды» (Уфа, 11-15 июня 2013). С. 57.
7. Gorbachev A., Fisunov G., Govorun V. Regulation of gene expression in *Mycoplasma gallisepticum*: heat-stress model // *EMBO Conference From Functional Genomics to Systems Biology* (Гейдельберг, 17-20 ноября 2012). P. 156
8. Gorbachev A., Fisunov G., Govorun V. Regulation of gene expression in *Mycoplasma gallisepticum*: heat-stress model // *38th FEBS Congress* (Санкт-Петербург, 6-11 июля 2013). P. 560

1. Введение

1.1. Общая характеристика и классификация Молликут

Грам-положительные микроорганизмы, принадлежащие к классу Mollicutes, для представителей которого характерно отсутствие клеточной стенки, являются наименьшими по размеру известными бактериями, способными к самостоятельному существованию. Для них характерны: значительная редукция генома, размеры которого варьируют от 580 тыс. до 1.4 млн. пар оснований; низкое содержание Г-Ц оснований (31 % для *M. gallisepticum*) [1].

Согласно современной таксономической классификации, класс Mollicutes относится к отдельному типу Tenericutes (домен Bacteria), в состав которого входит только этот класс. В настоящее время описано около 200 видов, относящихся к классу Mollicutes, которые объединяют в пять следующих семейств:

- Acholeplasmatales
- Anaeroplasmatales
- Entomoplasmatales
- Haloplasmatales
- Mycoplasmatales

Семейство Mycoplasmatales объединяет представителей родов *Mycoplasma* и *Ureaplasma*. На сегодняшний день описано более 100 видов и подвидов рода *Mycoplasma*, к ним относится множество патогенов человека и животных [5].

Согласно филогенетическому анализу, основанному на сравнении последовательностей генов, кодирующих 16S рРНК, *M. gallisepticum* принадлежит к кластеру pneumoniae вместе с человеческими патогенами

Mycoplasma pneumoniae и *Mycoplasma genitalium*. Уровень гомологии генов 16S рРНК (99 %) указывает на высокую степень родства между этими видами [5].

1.2. Репарация ДНК и поддержание геномной стабильности у бактерий

Все клетки должны аккуратно копировать и сохранять (поддерживать) последовательность ДНК, для того, чтобы обеспечивать правильную передачу генетического материала следующему поколению. Организмы от бактерий до человека содержат множество систем репарации ДНК, ответственных за специфическое распознавание и устранение многочисленных повреждений ДНК или неправильных спариваний, образующихся в течение жизненного цикла клетки. Большинство этих повреждений предположительно возникают из эндогенных источников, таких как химически активные побочные продукты нормального клеточного метаболизма [10]. Повреждение ДНК и накопление мутаций может снижать приспособленность клеток и потенциально влиять на их выживаемость. С другой стороны, мутагенез также служит материалом для эволюции, так, например, однонуклеотидные замены могут давать селективное преимущество бактериальным клеткам при изменении условий окружающей среды [11].

Системы репарации и мутагенеза наиболее изучены для *E.coli*, при этом белковые семейства, включающие в себя участников репарации, являются высоко консервативными и распространенными среди всех живых организмов.

Развитие полногеномного секвенирования дало возможность для изучения эволюции и оценки времени дивергенции Грам-положительных и Грам-отрицательных бактерий: более двух миллиардов лет назад [12]. Такая

длительная сепарация привела к расхождению во многих процессах репарации ДНК у Грам-положительных и Грам-отрицательных бактерий, включая молекулярные механизмы и способы их регуляции. Исследования последних лет делают все более ясным понимание того, что репарация ДНК у Грам-положительных бактерий значительно отличается от того, что было ранее описано для *E. coli*. В ряде случаев у Грам-положительных бактерий присутствуют механизмы репарации, полностью отсутствующие у *E. coli* [13].

В настоящем обзоре мы рассматриваем различные механизмы репарации ДНК, представленные как у Грам-положительных (на примере *Bacillus subtilis*), так и у Грам-отрицательных (на примере *E. coli*) бактерий.

1.2.1. Система SOS ответа

Активация SOS ответа представляет собой цепь транскрипционных событий, которые происходят в результате повреждения ДНК, остановки репликативной вилки и многих других воздействий, нарушающих целостность генома [14]. Система SOS ответа хорошо охарактеризована для *E. coli* [14]. При появлении в клетке одноцепочечной ДНК, происходит ее связывание белком RecA (образование нуклео-белкового филамента), что в итоге приводит к стимуляции автокаталитического расщепления белка LexA – транскрипционного репрессора, негативно регулирующего SOS индукцию. В результате такого расщепления происходит дерепрессия генов, находящихся под контролем белка LexA и запускается SOS ответ. В *E. coli* обнаружено 56 генов, репрессированных белком LexA, которые составляют SOS регулон [14].

Одним из первых прямых доказательств наличия у Грам-положительных бактерий системы, индуцируемой повреждениями ДНК, может служить эксперимент со случайными вставками гена *LacZ* без

промотора в хромосому *B. subtilis*, чтобы определить повышает ли повреждение ДНК экспрессию индуцируемых повреждениями генов (англ. *din*, damage-inducible) [15]. В результате такого анализа было идентифицировано 15 генов *din*, экспрессия которых возрастает при повреждениях ДНК, вызванных УФ-излучением, митомицином С, а также этилметансульфонатом (EMS). Данные эксперименты продемонстрировали наличие SOS-подобной системы у *B. subtilis*.

Основное отличие между бактериальными SOS системами состоит в перечне генов, находящихся под контролем LexA-репрессора и варьирующих от организма к организму.

У *B. subtilis* высоко консервативные белки RecA и LexA играют центральную роль в регуляции SOS транскрипционного ответа [16] (рис. 1). Белок RecA представляет собой мультифункциональный полипептид, необходимый для осуществления гомологической рекомбинации. Этот белок положительно регулирует SOS ответ у *B. subtilis*, как и у *E. coli* [14,16].

Таблица 1. SOS-боксы Грам-положительных бактерий в сравнении с боксом *E. coli* [13].

Организм	Консенсус SOS-бокса*
<i>Escherichia coli</i>	CTGT-(AT) ₄ -ACAG
<i>Bacillus subtilis</i>	CGAAC-RNRY-GTTYC
<i>Staphylococcus aureus</i>	CGAAC-AAAT-GTTCG
<i>Listeria monocytogenes</i>	AATAAGAACATATGTTCGTTT
<i>Corynebacterium glutamicum</i>	TCGAA(A/C)ANNTGTTCGA

*Обозначения: R – пурины, Y – пиримидины.

Белок RecA образует комплекс с оцДНК, формируя нуклео-протеиновый филамент, который стимулирует автокаталитическое расщепление белка LexA, транскрипционного репрессора SOS регулона [17]. Белок LexA репрессирует экспрессию 63 генов, входящих в состав 23 оперонов у *B. subtilis*, связываясь с их промоторами и предотвращая транскрипцию.

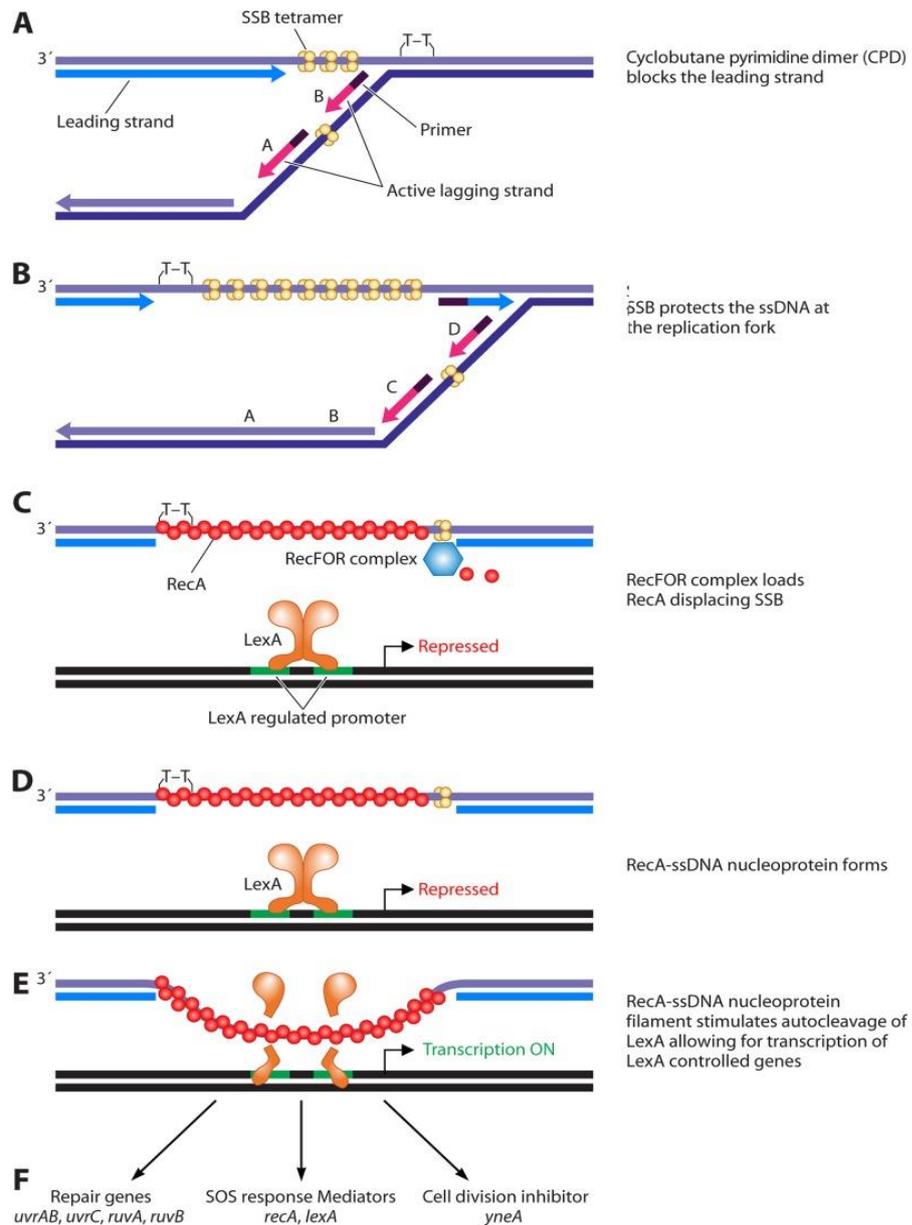


Рис. 1. Модель активации SOS ответа у *B. subtilis*. (A) В данной модели УФ-излучение вызывает образование циклобутанового пиримидинового димера в лидирующей матричной цепи ДНК при репликации, в результате чего образуется одноцепочечный «зазор» на дочерней цепи после репраймирования и продолжения репликации за повреждением. (B) Белок SSB связывается с одноцепочечным участком на дочерней цепи. (C) Белки-посредники рекомбинации RecF, RecO, RecR и, возможно, дополнительные факторы стимулируют связывание белка RecA с одноцепочечным участком при одновременном замещении белка SSB. (D) Белок RecA формирует нуклео-протеиновый филамент с оцДНК. (E) Комплекс RecA с оцДНК взаимодействует с белком LexA, в результате происходит активация скрытой протеазной активности и авторасщепление белка LexA. В результате инактивации белка LexA происходит дерепрессия транскрипции SOS генов и индуцируется глобальный транскрипционный ответ. (F) SOS-зависимые изменения экспрессии генов помогают клетке выжить при повреждении ДНК с помощью повышения уровня белков репарации, остановки клеточного деления, и, затем, повышения уровня LexA и RecA для прекращения SOS ответа после восстановления повреждений (адаптировано из [13]).

Консенсус сайта связывания белка LexA был определен для *B. subtilis* и некоторых других бактерий [16] (табл. 1). Многие LexA-регулируемые гены являются функциональными участниками репарации и репликации ДНК, а также участвуют в остановке клеточного деления [18]. Следует отметить, что примерно 25% генов, участвующих в SOS ответе у *B. subtilis* имеют неизвестную функцию [16]. При этом только 10 генов из состава SOS системы *B. subtilis* имеют гомологов или аналогов среди SOS-генов *E. coli* [13].

Известно, что SOS индукция очень важна для выживания клеток *E. coli* при экзогенном повреждении ДНК [19]. Сравнение относительного количества клеток, в которых индуцируется SOS система под действием ионизирующего излучения, между *E. coli* и *B. subtilis* демонстрирует, что клетки *E. coli* имеют гораздо более низкий порог повреждений, вызывающий SOS ответ [16]. Кроме того, данное исследование показывает, что сайт-специфические двуцепочечные разрывы, генерируемые с помощью эндонуклеазы I – SceI, вызывают SOS индукцию менее, чем в 5% клеток. Более того, в отсутствие SOS индукции, *B. subtilis* способна выживать при более высоких дозах радиации, чем переносимые *E. coli*, что может свидетельствовать о более эффективной системе репарации ДНК у *B. subtilis* (даже без SOS индукции) по сравнению с *E. coli* [16]. В другом эксперименте был использован подход с использованием репрессора TetR и его оператора (в англоязычной литературе – *tetO* array) для блокировки движения ДНК-полимеразы и остановки репликативной вилки [20]. Данное исследование демонстрирует, что не происходит значительной индукции SOS ответа при блокировании репликации клеточной ДНК. Принимая во внимание оба представленных выше исследования, можно заключить, что *B. subtilis* может эффективно репарировать повреждения ДНК, а также переживать остановку репликации без быстрой активации SOS ответа. В дополнение к этому, стоит

сказать, что клетки *B. subtilis*, содержащие нерасщепляемый вариант белка LexA, демонстрируют выживаемость в условиях значительных повреждений ДНК, что указывает на эффективную систему репарации в отсутствие включенного SOS ответа [16].

Система SOS ответа была исследована у некоторых патогенных и непатогенных Грам-положительных микроорганизмов. Условный человеческий патоген *Staphylococcus aureus* содержит гены *lexA* и *recA* [21]. При воздействии субингибирующих концентраций ДНК-повреждающих антибиотиков, таких как фторхинолоны (ингибиторы ДНК-гиразы), у *S. aureus* наблюдается SOS ответ [22]. С использованием технологии микрочипов был проведен транскриптомный анализ SOS ответа у данного микроорганизма под действием цiproфлоксацина (фторхинолон, индуцирующий двуцепочечные разрывы ДНК и остановку репликации) [23]. В данном исследовании был проведен анализ штамма *S. aureus*, содержащего нерасщепляемый вариант LexA, в сравнении с диким типом. В результате было идентифицировано 16 генов, находящихся под контролем LexA-репрессора. Данное число генов является небольшим относительно числа генов, находящихся под SOS контролем у *B. subtilis*. Среди генов, входящих в SOS систему *S. aureus*, были идентифицированы как повышающие экспрессию: гены *recA* и *lexA*, так и гены, продукты которых участвуют в системе репарации по пути эксцизии нуклеотидов (NER) – *uvrA* и *uvrB*, гены топоизомеразы IV – *parC* и *parE*, а также гены, кодирующие нуклеазу SbcCD – *sbcC* и *sbcD*. Связывание белка LexA из *S. aureus* с промотором гена *recA* было продемонстрировано в [21], что согласуется со способом регуляции *recA* у других бактерий. Интересно, что среди генов, индуцируемых фторхинолонами у *S. aureus*, присутствуют гены, кодирующие фибронектин-связывающие белки, необходимые для прикрепления к внеклеточному матриксу и плазматической мембране клеток хозяина [21].

Кроме того, промотор гена, кодирующего фибронектин-связывающий белок В (*fnbB*), связывается белком LexA. Подобные результаты позволяют предполагать, что повреждение ДНК может влиять на способность *S. aureus* формировать сгустки или прикрепляться к клеточной поверхности – важные свойства в инфекционном процессе [21].

Другой человеческий патоген – бактерия *Listeria monocytogenes* также содержит гены *lexA* и *recA* [24]. Исследование действия ДНК-повреждающего агента митомицина С на *Listeria* позволило идентифицировать 29 индуцируемых генов, входящих в состав 16 оперонов [24]. Большинство этих генов вовлечены в репарацию ДНК и регуляцию клеточного деления [24].

Система SOS ответа у других Грам-положительных бактерий также регулируется с помощью белков RecA и LexA [25]. В целом, число и функциональная активность участников данной системы варьирует от вида к виду. Однако практически у всех исследованных микроорганизмов в систему SOS входят гены, ответственные за репарацию ДНК и контроль клеточного деления.

В геноме *M. gallisepticum* присутствуют следующие участники SOS-системы: рекомбиназа А (RecA), хеликазный комплекс (RuvAB), нуклеазно-хеликазный комплекс UvrABC, а также ДНК-зависимая ДНК-полимераза IV типа (DinB), с предсказанной по гомологии мутаторной активностью. При этом гомологи всех известных регуляторов SOS-системы не были найдены ни в одном из проанализированных геномов микоплазм [26]. Отсутствие у всех представителей класса Mollicutes LexA репрессора долгое время служило доказательством того, что регуляция SOS-ответа у этих бактерий отсутствует [26,27].

1.2.2. Гомологическая рекомбинация

Система гомологической рекомбинации занимает центральное место в репарации двуцепочечных разрывов ДНК и интеграции ДНК после генетической трансформации [28].

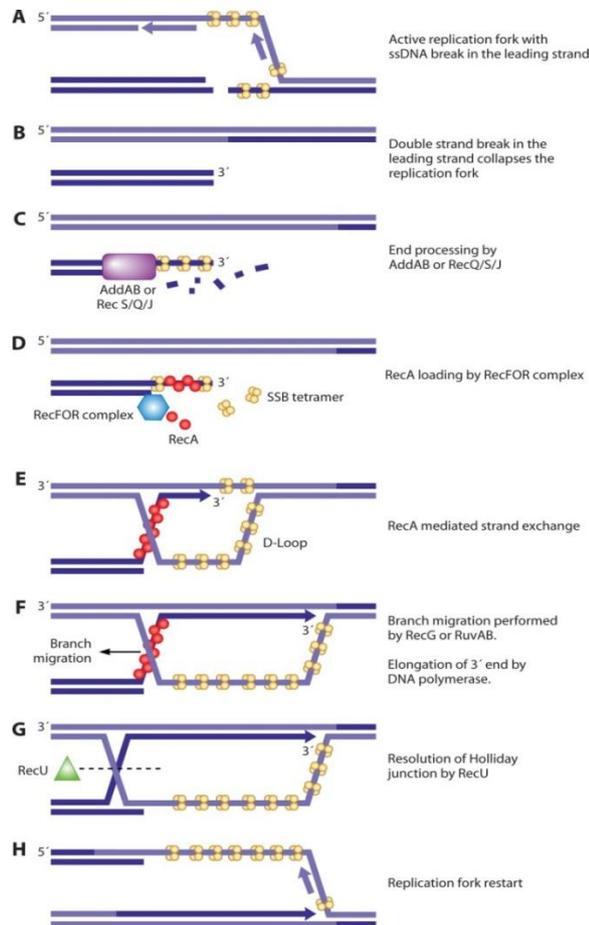


Рис. 2. Модель репарации одного двуцепочечного разрыва путем гомологической рекомбинации у *B. subtilis*. (A) Активная репликативная вилка с одноцепочечным разрывом (nick) в лидирующей цепи матрицы. (B) При движении репликативной вилки через место повреждения происходит ее разрушение и образование двуцепочечного разрыва ДНК. (C) Свободный двуцепочечный конец процессируется с помощью нуклеазно-хеликазного комплекса AddAB или сочетанием хеликазы RecS или RecQ с нуклеазой RecJ. В случае AddAB комплекса, он деградирует обе цепи (по 3' и 5'-концам) пока не достигнет сайта Chi, после чего происходит формирование 3'-выступающего одноцепочечного конца ДНК. (D) Белковый комплекс-посредник рекомбинации – RecFOR производит загрузку рекомбиназы RecA на одноцепочечный участок ДНК. В результате образуется оцДНК-RecA нуклео-белковый филамент. (E) оцДНК-RecA филамент формирует D-петлю, при этом одна из цепей двуцепочечной матрицы ДНК замещается оцДНК-RecA филаментом. (F) Свободный 3'-ДНК конец из филамента достраивается с помощью ДНК-полимеразы, гомологичная цепь используется в качестве матрицы для синтеза. Белок RecG или RuvAB комплекс обеспечивает миграцию структуры Холлидея. (G) После того, как поврежденная цепь значительно удлинилась, структура Холлидея разрешается с помощью эндонуклеазы RecU или, возможно, RecV (прерывистой линией показано место разрезания). (H) Происходит восстановление репликативной вилки и достройка отстающей цепи ДНК (адаптировано из[13]).

Система гомологической рекомбинации занимает центральное место в репарации двуцепочечных разрывов ДНК и интеграции ДНК после генетической трансформации [28]. Можно выделить следующую последовательность действий при репарации двуцепочечных разрывов: (i) распознавание и расщепление свободных двуцепочечных концов с образованием 3'- выступающего (из дцДНК) одноцепочечного конца ДНК, (ii) образование комплекса рекомбиназы (RecA) с образовавшейся в результате процессинга оцДНК, (iii) спаривание оцДНК с неповрежденным гомологичным участком ДНК («внедрение» оцДНК в гомологичную дцДНК), образование «D-петли», (iv) синтез ДНК с использованием 3'-ОН конца «внедряемой» цепи в качестве затравки, (v) эндонуклеолитическое разрешение «D-петлевой» структуры, с образованием двух неповрежденных хромосом [13] (рис. 2). Перечисленные шаги, необходимые для гомологической рекомбинации являются общими среди всех живых организмов, с отличиями в белковых участниках на каждом из этапов.

В геноме *M. gallisepticum* присутствуют гены, кодирующие ключевые белки рекомбинации – рекомбиназы RecA и RecR, а также гены RuvA и RuvB, кодирующие ДНК-хеликазу RuvAB (участвует в миграции ДНК-цепей), а также ген, кодирующий фермент, способный разрешать структуру Холлидея – ДНК-резольваза RecU [29].

1.2.3. Система эксцизионной репарации нуклеотидов

Система эксцизионной репарации нуклеотидов (NER) необходима для высокоточной репарации различных повреждений, вызванных химическими веществами и УФ-излучением [30]. Высококонсервативный нуклеазный комплекс UvrABC осуществляет распознавание и вырезание фрагмента ДНК длиной в 10-15 нуклеотидов, содержащий поврежденные нуклеотиды [13]. У

B. subtilis гены *uvrBA* входят в один оперон и их экспрессия регулируется SOS системой [31]. Инактивация гена *uvrA* делает клетки высокочувствительными к различным повреждающим ДНК агентам, включая УФ-излучение, 4NQO (4-нитрохинолин-1-оксид), митомицин С [32]. Ген *uvrC* локализован отдельно от оперона *uvrBA* и демонстрирует очень низкий уровень SOS индукции [32]. У *E. coli* репарация по пути эксцизии нуклеотидов осуществляется следующим образом: в первую очередь белковый комплекс UvrA₂B (димерUvrA с мономером UvrB) осуществляет распознавание поврежденных нуклеотидов [33]. После распознавания и связывания повреждения, происходит диссоциация белков UvrA и привлечение UvrC на место повреждения [34]. Белок UvrC в комплексе с UvrB вносит два одноцепочечных разрыва по бокам от повреждения в репарируемую цепь ДНК на расстоянии 10-15 нуклеотидов друг от друга [35]. Затем поврежденный участок удаляется с помощью ДНК-хеликазы II (UvrD), после чего образовавшаяся брешь застраивается ДНК-полимеразой I [13]. Оставшийся одноцепочечный разрыв (nick) восстанавливается с помощью ДНК-лигазы [33]. Большинство биохимических экспериментов по изучению системы NER были выполнены с белками *E. coli*. Для гена *uvrC* из *B. subtilis* показано с помощью комплементационного теста, что он может заменять функции удаленного *uvrC* у *E. coli*. Кроме того, очищенный белок UvrC *B. subtilis* может выполнять функции UvrC *E. coli* в реконструированной *in vitro* репарационной реакции с очищенными компонентами. Эксперименты по локализации слитого с флуоресцентным белком UvrA (UvrA-GFP) демонстрируют, что белок UvrA-GFP колокализуется с нуклеотидом и, кроме того, флуоресценция белка UvrA-GFP усиливается после УФ-облучения нуклеотида [36]. Данный эксперимент позволяет предполагать, что UvrA может сканировать хромосомную ДНК с целью идентификации поврежденных нуклеотидов.

В геноме *Mycoplasma gallisepticum* присутствуют практически все гены, кодирующие белки основного пути NER, за исключением ДНК-полимераз классов I и II [29]. При этом присутствует ген, частично гомологичный ДНК-полимеразе I *E.coli* и *B. subtilis* своим экзонуклеазным доменом, но не содержащий полимеразного домена (рис.3).

1.2.4. Система эксцизионной репарации азотистых оснований

Система эксцизионной репарации азотистых оснований (BER) специализируется на небольших повреждениях ДНК. В отличие от системы NER, которая функционирует на нуклеотид/олигонуклеотидном уровне восстановления более серьезных повреждений [37]. Мелкие повреждения азотистых оснований могут возникать в результате многочисленных химических реакций, включая алкилирование, окисление, депуринизацию/депиримидинизацию, дезаминирование и встраивание dUTP нуклеотидов в процессе репликации [37]. С учетом многочисленных источников потенциальных повреждений, система BER считается одной из наиболее часто используемых путей репарации ДНК *in vivo* [38]. Основной процесс репарации по пути BER начинается с детекции повреждения с помощью одной из гликозилаз, ферментов, которые гидролизуют β -N-гликозидную связь между дезоксирибозой и поврежденным азотистым основанием с высвобождением последнего. В результате образуется апуриновый/апиримидиновый сайт – не содержащий азотистого основания (далее называется AP-сайт). Подобные сайты высоко мутагенны и могут быть потенциальной причиной одноцепочечных разрывов ДНК [37,39]. На следующем этапе AP-сайт узнается белками: AP-эндонуклеазой или AP-лиазой, которые могут вносить одноцепочечный разрыв (nick) с 5'- или 3'-конца от AP-сайта, соответственно. Далее происходит процессинг AP-сайта с помощью экзонуклеазы или дезоксирибофосфодиэстеразы (dRpase), с

образованием небольшой однонуклеотидной брешы. Затем брешь застраивается ДНК-полимеразой I и лигируется, в результате сайт восстанавливается в первоначальном неповрежденном виде [37,39].

Следует отдельно остановиться на рассмотрении репарации, связанной с окисленными формами гуанина, такими как 7, 8 – дигидро-8-оксогуанин (далее 8-оксогуанин), а также гуанин с открытыми имидазольными кольцами – формамидопиримидин [40,41]. Образование 8-оксогуанина является следствием окисления гуанина по 8 атому углерода, в то время как формамидопиримидин образуется под действием ионизирующей радиации, метилирования азота N⁷ или окислительного повреждения [13]. Окислительное повреждение ДНК может вызываться экзогенными источниками, однако, основной причиной окисления являются эндогенные активные формы кислорода (АФК), образующиеся в результате клеточного метаболизма [40]. Основной формой окислительного повреждения ДНК является 8-оксогуанин и, в случае отсутствия его репарации, он может являться мутагеном, поскольку репликативные ДНК-полимеразы могут встраивать в новую цепь дАТФ в качестве комплементарной пары в процессе синтеза ДНК [42,43]. В результате образуется 8-оксо-G-A спаривание и, если такое спаривание не будет исправлено, на следующем цикле произойдет трансверсия пары GC в AT[44–46]. У *E. coli* гликозилазы MutM (*fpg*) и MutY снижают мутагенный потенциал 8-оксогуанина в ДНК за счет его прямого удаления (MutM) или удаления аденина (MutY) из 8-оксо-G-A пары [47,48].

Делеционный анализ по генам *mutM* и *mutY* у *B. subtilis* показывает, что *mutM*-дефицитный штамм имеет небольшое увеличение уровня мутагенеза (примерно в 5 раз по сравнению с диким типом), в то время как генный нокаут по *mutY* вызывает возрастание уровня мутагенеза почти в 100 раз. В случае двойного мутанта (*mutM* и *mutY*), который неспособен удалять 8-оксогуанин и 8-оксо-G-A спаривание одновременно, частота мутаций возрастает примерно в 1000 раз [49].

Еще одной мишенью для окисления является нуклеотидный пул в клетке. Например, dGTP может быть окислен с образованием 8-оксо-dGTP [10]. Окисленные нуклеотиды имеют высокий мутагенный потенциал, поскольку могут встраиваться в ДНК в процессе репликации. При встраивании 8-оксо-dGTP он может спариваться с аденином, что может привести к АТ=>ГС трансверсии [50]. Для того, чтобы убирать окисленные нуклеотиды из нуклеотидного пула, в клетках *E. coli* существует специальный белок – MutT [38]. Клетки, дефицитные по гену *mutT*, показывают повышенную частоту спонтанного мутагенеза – до 10000 раз выше, чем в диком типе, при этом спектр мутаций представлен в основном АТ=>ГС трансверсиями [51]. По своей ферментативной активности MutT является нуклеозид триозофосфатазой, которая селективно гидролизует 8-оксо-dGTP с образованием пирофосфата и 8-оксо-dGMP, который не может встраиваться в ДНК при репликации [52]. Для *B. subtilis* было показано, что функцию гидролиза 8-оксо-dGTP может выполнять белок YtkD, кодируемый геном *ytkD*, ортологичным гену *mutT* у *E. coli*. Кроме того, было показано, что ген *ytkD* может заменять функцию гена *mutT* у *E. coli* в комплементационном тесте [53–55]. Однако в другом исследовании было показано отсутствие предпочтений у белка YtkD к 8-оксо-dGMP относительно dGTP в качестве субстрата, также была показана низкая частота мутагенеза у штамма с нокаутом по гену *ytkD*. Подобные результаты позволяют утверждать, что белок YtkD у *B. subtilis* не является функциональным аналогом белка MutM для *E. coli*, а является неспецифической нуклеотид-гидролазой, которая может снижать содержание окисленных нуклеотидтрифосфатов наравне со всеми остальными в общем пуле.

В геноме *M. gallisepticum* содержится три основных фермента эксцизионной репарации азотистых оснований: две гликозилазы (урацил-ДНК-гликозилаза (Ung) и формаидопиримидин-ДНК-гликозилаза (MutM))

и одна 5' AP эндонуклеаза (Nfo). Следует отметить, что две ДНК-гликозилазы, присутствующие в геноме *M. gallisepticum*, являются наиболее часто используемыми ферментами, участвующими в устранении ДНК-повреждений, вызванных эндогенным окислительным стрессом [56].

1.2.5. Система репарации мисматчей

Система репарации мисматчей (под «мисматчем» понимается пара ошибочно спаренных нуклеотидов) – процесс коррекции ошибок репликации, возникающих в процессе синтеза ДНК [57,58]. Для *E. coli* описана следующая последовательность действий: белок MutS связывает ошибочное спаривание, привлекая при этом белок MutL [57,58], затем белок MutL привлекает и активирует эндонуклеазу MutH, которая вносит одноцепочечный разрыв (nick) в неметилованную (дочернюю) цепь ДНК. После этого хеликаза UvrD приходит на место разрыва и раскручивает цепь ДНК, содержащую неправильное основание [57,58]. Затем цепь, содержащая ошибку, деградирует при участии одной из многих нуклеаз, в зависимости от направления деградации цепи (3'=>5' или 5'=>3'). Образовавшаяся в результате деградации брешь застраивается с помощью ДНК-полимеразы III и лигируется ДНК-лигазой [57,58]. Для прохождения репарации ДНК по пути MMR у *E. coli* требуются следующие белки: MutS, MutL, MutH, UvrD, а также ДНК-метиلاза Dam, способная метилировать аденин в GATC сайтах ДНК. Наличие метилированного аденина в данном сайте определяет правильную дискриминацию цепи ДНК, по которой будет происходить мисматч-репарация (табл. 2).

Таблица 2. Сравнение белков, участвующих в мисматч-репарации, у *E. coli* и *B. subtilis* (адаптировано из[10]).

Белки <i>E. coli</i>	Функциональная роль	Белки <i>B. subtilis</i>
MutS	Распознавание мисматча	MutS
MutL	Белок-посредник	MutL (обладает)

		эндонуклеазной активностью)
MutH	Метил-чувствительная эндонуклеаза	отсутствует
Dam	Метилаза (GATC)	отсутствует
Эксонуклеазы I, X, VII, RecJ	Эксонуклеазы, деградирующие репарируемую цепь ДНК	Нет эксонуклеаз, для которых показано участие в мисматч-репарации ¹
UvrD	ДНК-хеликаза	YrrC ²

¹Эксонуклеазы I и X отсутствуют у *B. subtilis*, Exo VII и RecJ найдены у *B. subtilis*, однако, их функция в мисматч-репарации не известна.

²YrrC может функционировать как хеликаза в мисматч репарации у Грам-положительных бактерий, является ортологом белка RecD2 с 5'-3' ДНК-хеликазной полярностью, основано на исследовании белка RecD2 у *D. radiodurans*[59].

Для *B. subtilis* показано присутствие высоко консервативных белков MutS и MutL [60], однако, отсутствуют белки MutH, Dam, кроме того, функцию UvrD выполняет его ортолог – белок YrrC [10]. Белок MutS представляет собой «сенсор», который распознает мисматчи, в то время как белок MutL является «линкером», осуществляющим взаимодействие между белками мисматч-репарации для правильной работы по исправлению ошибок репликации у *B. subtilis* [57,58]. Поскольку в геноме *B. subtilis* отсутствуют гены *dam* и *mutH*, была высказана гипотеза, что *B. subtilis* реализует независимый от метилирования ДНК путь мисматч-репарации, что является существенным отличием от MMR пути репарации, описанным для *E. coli*. Кроме того, анализ метилирования сайтов GATC у *B. subtilis* и *S. aureus* указывает на отсутствие функционального аналога *dam* метилазы у этих микроорганизмов [61].

Недавно было показано, что опосредованная метилированием система мисматч-репарации ДНК, охарактеризованная для *E. coli*, отсутствует у большинства изученных бактерий. В качестве гипотезы высказывается предположение, что большая часть бактерий использует независимую от метилирования ДНК систему репарации мисматчей [10]. Предполагается, что дискриминация правильной цепи при репарации осуществляется по наличию

одноцепочечных разрывов вблизи репликативной вилки, по механизму сходному с эукариотическим [62,63].

Система репарации неспаренных оснований («мисматч»-репарация) до последнего времени считалась наиболее редуцированной у всех микоплазм. У них не было обнаружено гомологов ни одного из ключевых генов – *MutH*, *MutL* и *MutS*. Ряд авторов [64] спекулировал о повышенном уровне спонтанного мутагенеза у микоплазм, однако позднее было показано, что уровень мутагенеза существенно не отличается от такового у *E. coli* [6,7].

Характерной особенностью *M. gallisepticum* является отсутствие «канонических» экзонуклеаз (ExoI, ExoVII, RecJ, ExoX [65]), участвующих в пути MMR, подробно изученного для *E. coli*. Однако, следует отметить, что все микоплазмы имеют «DNaPolII-подобную» экзонуклеазу Exo, фермент, гомологичный ДНК-полимеразе I *E. coli*, который потенциально может работать в системе «мисматч»-репарации. При этом в белке Exo отсутствует ДНК-полимеразный домен, присутствующий в DNaPolII *E. coli* (рис.3).

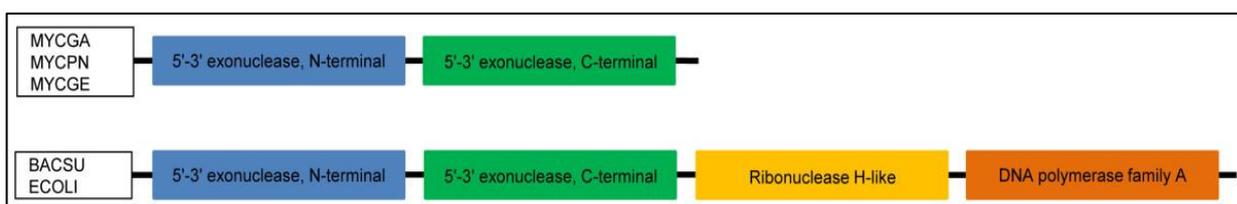


Рис.3. Доменная организация белка Exo *M. gallisepticum*, *M. genitalium* и *M. pneumoniae* (сверху) в сравнении с ДНК-полимеразой I (DNaPolII) *E. coli* и *B. subtilis* (снизу) [66].

1.2.6. Устойчивость ДНК к повреждениям и синтез с низкой точностью

Низкоточные ДНК-полимеразы получили свое название благодаря способности осуществлять репликацию через некодирующие основания, которые в норме блокируют прохождение репликативной ДНК-полимеразы [67,68]. У *B. subtilis* и многих других Грам-положительных бактерий присутствуют две полимеразы Y семейства: PolY1 (YqjW) и PolY2 (YqjH)

[69]. ДНК полимеразы PolY1 демонстрирует высокое сходство аминокислотной последовательности с белком UmuC из *E. coli*, в то время как полимеразы PolY2 ближе по последовательности к ДНК-полимеразе IV (DinB) из *E. coli* [69]. В клетках *E. coli* белок UmuC взаимодействует с пострансляционно-модифицированной версией белка UmuD, называемой UmuD' [70]. У данного белка удалены 24 аминокислотных остатка с N-конца белка UmuD. Два мономера процессированного белка - UmuD' образуют димер и в таком виде связывают белок UmuC, формируя ДНК-полимеразу V (UmuD'₂C) [71,72]. Стоит отметить, что белок UmuD отсутствует у неродственных для *E. coli* бактерий [73]. В частности, белковый аналог UmuD не был идентифицирован в *B. subtilis* [74]. Предполагается, что PolY1 работает без дополнительного белкового фактора, аналогичного белку UmuD в *E. coli*. Кроме того, для обеих альтернативных полимераз *B. subtilis* – PolY1 и PolY2 было показано (с использованием двугибридной дрожжевой системы) взаимодействие с фактором процессивности репликации – β-кольцом (β-clamp), кодируемым геном *dnaN* [75].

Следует отметить, что для альтернативной полимеразы DinB *E. coli* была показана функциональная роль в адаптивном мутагенезе при различных стресс-вызывающих воздействиях [76]. Для полимеразы PolY2 также показано участие в адаптивном мутагенезе при стационарной фазе роста [74].

Синтез «через повреждения» (англ. translesion synthesis) не ограничивается участием только ДНК-полимераз Y-семейства, поскольку ДНК-полимеразы C-семейства также могут играть роль при репликации поврежденных участков ДНК и мутагенезе у некоторых Грам-положительных бактерий. Белок DnaE является репликативной полимеразой у *E. coli* и других Грам-отрицательных бактерий [77]. Для *B. subtilis* и многих Грам-положительных бактерий с низким Г-Ц составом показано наличие двух ДНК-полимераз C-семейства: PolC и DnaE [78]. Состав репликативного комплекса Грам-положительных бактерий с низким содержанием Г-Ц

оснований отличается от охарактеризованного для *E. coli* (для обзора см. [77,79,80]). Стоит отметить, что делеция по гену *dnaE* является летальной для таких Грам-положительных бактерий, как *B. subtilis*, *Streptomyces coelicolor*, *Streptococcus pyogenes*, *S. aureus*, *Enterococcus faecalis* [81–84]. Жизненно необходимая роль DnaE заключается в способности синтезировать ДНК с РНК-затравки на отстающей цепи ДНК в процессе репликации, в отличие от ДНК-полимеразы PolC, нуждающейся в ДНК-дуплексе для начала синтеза [85]. Напротив, у Грам-отрицательных бактерий белок DnaE ассоциирован с «корректирующей» ε-субъединицей (DnaQ) ДНК-полимеразного комплекса, которая обладает 3'-5' экзонуклеазной активностью, необходимой для коррекции ошибок и высокой точности репликации ДНК [86]. В то же время отсутствуют сообщения о связи DnaE Грам-положительных бактерий с корректирующей ε-субъединицей ДНК-полимеразы.

Для полимеразы DnaE из *Streptococcus pyogenes* показано наличие «склонной к ошибкам» активности (англ. “error prone”) за способность реплицировать через AP-сайты, а также достраивать праймеры, содержащие ошибочные спаривания. Предполагается, что DnaE у *S. pyogenes* может играть роль в репликации поврежденных участков ДНК и мутагенезе [87].

Для *B. subtilis* ген *dnaE* является жизненно необходимым, благодаря роли продукта этого гена в репликации отстающей цепи ДНК. Кроме того, показано увеличение уровня белка DnaE при повреждении ДНК, а также в клетках с нокаутом по гену *lexA* [88]. Также показано наличие сайта связывания LexA в промоторной зоне гена *dnaE*, а также 11-кратное увеличение представленности транскрипта при повреждении ДНК [89]. Эти данные подтверждают гипотезу о том, что DnaE может участвовать в репарации ДНК, репликации поврежденных участков. Недавно, в экспериментах по экспрессии белка DnaE, слитого с зеленым флуоресцентным белком (DnaE-GFP), было продемонстрировано, что при

обработке клеток ДНК-повреждающим агентом (митомицином С) наблюдаются увеличение клеток, содержащих флуоресцирующие точки, что также указывает на роль DnaE в репликации поврежденной ДНК. Таким образом, можно обобщить, что DnaE является SOS индуцируемой ДНК-полимеразой, способной осуществлять синтез через поврежденные участки ДНК [90].

1.3. Гистоноподобный белок HU

У бактерий, правильная упаковка и структурирование генома, необходимые для поддержания его целостности, требуют специальных белков. Нуклеоид-ассоциированные белки (NAPs), включая LRP, FIS, H-NS, IHF и HU, вовлечены в упаковку и образование суперспирализации ДНК. Они модулируют важнейшие клеточные функции, такие как репликация, рекомбинация, репарация и транскрипция клеточной ДНК [91–93]. Каждый вид характеризуется своим набором NAPs, при этом только HU-подобные белки имеют повсеместное распространение среди бактерий. Гистоноподобные свойства HU заключаются в его способности вносить, как это делает гистоновый октамер, отрицательную суперспирализацию в релаксированную замкнутую кольцевую молекулу ДНК в присутствии топоизомеразы I и конденсировать ДНК [94,95], а также связывать ДНК неспецифически.

HU-белок *E. coli* имеет предпочтение к связыванию суперскрученной ДНК [96], что приводит к зависимости активности топоизомеразы I от количества HU-белка в клетке [97]. Белок HU также вовлечен в репарацию [98,99], рекомбинацию и репликацию ДНК [100]. Транскриптомное профилирование HU-дефицитного штамма показало, что HU модулирует транскрипцию 8% генов *E. coli*, при этом только 10% из них могли подвергнуться косвенному воздействию через суперспирализацию [101]. Функции и ДНК-связывающие свойства HU-подобных белков могут варьировать в зависимости от специфического NAP контента и условий жизни микроорганизма [102]. Делеция по гену, кодирующему белок HU в *E. coli* не является летальной при отсутствии делеций по генам IHF и H-NS [103]. Напротив, отсутствие белка HU является летальным в случае, когда это единственный NAP [104,105].

У большинства бактерий, HU белок представляет собой гомодимер, при этом у энтеробактерий HU является гетеродимером [106]. Кристаллическая структура комплекса HU-белка с ДНК была разрешена для *Anabaena* и *Borrelia burgdoferi* [107,108]. Две HU субъединицы переплетены между собой с образованием компактного альфа-спирального тела, от которого отходят две бета-слоиные «руки», взаимодействующие с малой бороздкой ДНК (рис.4).

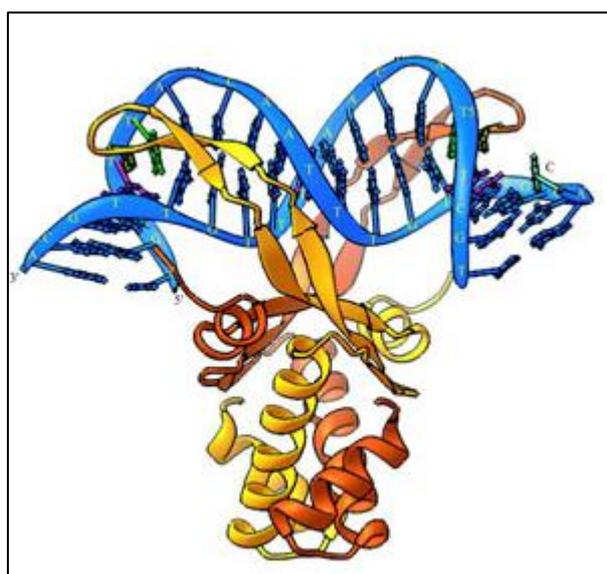


Рис.4. Схематическая структура кокристалла белка HU с ДНК (адаптировано из[107]).

Большинство изученных HU-подобных белков характеризуется слабым и неспецифическим связыванием с ДНК-дуплексом и более сильным и структуро-специфичным связыванием с неканоническими ДНК-структурами, такими как вилки, дцДНК с одноцепочечным разрывом, структуры Холлидея и т.п. [109–112]. При этом репертуар предпочитаемых субстратов для связывания значительно варьирует от вида к виду [102]. HU-белок *E. coli* связывает ДНК-дуплекс (В-форму), также как РНК-дуплекс (А-форма) и ДНК-РНК-дуплекс [96,113] и, кроме того, связывает одноцепочечную ДНК [114,115]. Однако, неизвестно, какие из свойств, перечисленных выше,

являются необходимыми для функционирования гистоноподобного HU-белка.

В геноме *M. gallisepticum* найдены два гена, гомологичных гену, кодирующему белок HU *E. coli* – *HinA/hup_1* и *HinA/hup_2* [116]. Функциональная роль ни одного из данных белков у микоплазм не была известна до настоящего исследования.

2. Материалы и методы

2.1. Олигонуклеотиды

Олигонуклеотиды были синтезированы в компании «Литех» Карповым В. А. на приборе ASM 800 DNAsynthesizer (BiossetLtd).

Последовательности олигонуклеотидов, меченных FAM (6-карбоксихлорофлуоресцеин) приведены в табл. 3, жирным курсивом с нижним подчеркиванием выделены места мисматчей. Для получения дуплексов по 3-12 μM каждого олигонуклеотида в паре инкубировали в 50 мкл буфера, содержащего 20 mM Трис-НСl (рН 8.0) и 200 mM NaCl, при 90⁰С в течение 3 мин, затем жидкость медленно охлаждали в лабораторном помещении в течение 4 часов.

Таблица 3. Олионуклеотиды, используемые для экспериментов по торможению ДНК в геле.

Название	Прямой (5'-3')	Обратный (5'-3')
ds	СТGACCTAAGGGATCCGTCGT AGCAAG <u>G</u> ACTCAACTGCACTC TAGACT	AGTCTAGAGTGCAGTTGAG T <u>C</u> CTTGCTACGACGGATCCC TTAGGTCAG
CC	СТGACCTAAGGGATCCGTCGT AGCAAG <u>C</u> ACTCAACTGCACTC TAGACT	AGTCTAGAGTGCAGTTGAG T <u>C</u> CTTGCTACGACGGATCCC TTAGGTCAG
CA	СТGACCTAAGGGATCCGTCGT AGCAAG <u>A</u> ACTCAACTGCACTC TAGACT	AGTCTAGAGTGCAGTTGAG T <u>C</u> CTTGCTACGACGGATCCC TTAGGTCAG
CT	СТGACCTAAGGGATCCGTCGT AGCAAG <u>T</u> ACTCAACTGCACTC TAGACT	AGTCTAGAGTGCAGTTGAG T <u>C</u> CTTGCTACGACGGATCCC TTAGGTCAG
AA	СТGACCTAAGGGATCCGTCGT AGCAAG <u>A</u> ACTCAACTGCACTC TAGACT	AGTCTAGAGTGCAGTTGAG T <u>A</u> CTTGCTACGACGGATCCC TTAGGTCAG
GG	СТGACCTAAGGGATCCGTCGT AGCAAG <u>G</u> ACTCAACTGCACTC TAGACT	AGTCTAGAGTGCAGTTGAG T <u>G</u> CTTGCTACGACGGATCC CTTAGGTCAG
GT	СТGACCTAAGGGATCCGTCGT AGCAAG <u>G</u> ACTCAACTGCACTC TAGACT	AGTCTAGAGTGCAGTTGAG T <u>T</u> CTTGCTACGACGGATCCC TTAGGTCAG
TT	СТGACCTAAGGGATCCGTCGT	AGTCTAGAGTGCAGTTGAG

	AGCAAG <u>T</u> ACTCAACTGCACTC TAGACT	<u>T</u> TCTTGCTACGACGGATCCC TTAGGTCAG
A1	CTGACCTAAGGGATCCGTCGT AGCAA <u>A</u> GGACTCAACTGCACT CTAGACT	AGTCTAGAGTGCAGTTGAG <u>T</u> CCTTGCTACGACGGATCCC TTAGGTCAG
C1	CTGACCTAAGGGATCCGTCGT AGCAA <u>C</u> GGACTCAACTGCACT CTAGACT	AGTCTAGAGTGCAGTTGAG <u>T</u> CCTTGCTACGACGGATCCC TTAGGTCAG
G1	CTGACCTAAGGGATCCGTCGT AGCAA <u>G</u> GGACTCAACTGCACT CTAGACT	AGTCTAGAGTGCAGTTGAG <u>T</u> CCTTGCTACGACGGATCCC TTAGGTCAG
T1	CTGACCTAAGGGATCCGTCGT AGCAAG <u>G</u> TACTCAACTGCACT CTAGACT	AGTCTAGAGTGCAGTTGAG <u>T</u> CCTTGCTACGACGGATCCC TTAGGTCAG
m2	CTGACCTAAGGGATCCGTCGT AGCAA <u>A</u> CACTCAACTGCACTC TAGACT	AGTCTAGAGTGCAGTTGAG <u>T</u> CCTTGCTACGACGGATCCC TTAGGTCAG
m3	CTGACCTAAGGGATCCGTCGT AGCAA <u>A</u> CGTCAACTGCACTC TAGACT	AGTCTAGAGTGCAGTTGAG <u>T</u> CCTTGCTACGACGGATCCC TTAGGTCAG
m4	CTGACCTAAGGGATCCGTCGT AGCAA <u>A</u> TGTTCAACTGCACTCT AGACT	AGTCTAGAGTGCAGTTGAG <u>T</u> CCTTGCTACGACGGATCCC TTAGGTCAG

2.2. *Культивирование Mycoplasma gallisepticum и получение клеточного экстракта*

Культура клеток *M. gallisepticum* штамм S6 выращивалась в жидкой среде, содержащей (на 1 литр): 20 г триптозы, 5 г NaCl, 3г Трис, 1,3 г KCl, pH 7.4 с добавлением 1% глюкозы, 50 мл жидкого дрожжевого экстракта, 100 мл лошадиной сыворотки (термически инактивированной при 56⁰С в течение 30 мин), 200 ед/мл пенициллина. Для экстракции белка 400 мл клеточной культуры (клетки были посеяны 1:100, росли в течение 48 часов) центрифугировали при 4⁰С, при 16000g, в течение 25 минут. Клеточный осадок трижды промывали буфером, содержащим 200 мМ Трис-HCl pH 8.0, 150мМ NaCl и 5 мкг/мл PMSF (фенилметансульфонилфторид). Затем клеточный осадок ресуспендировали в 8 мл буфера, содержащем 20 мМ Трис-HCl pH 8.0, 0,1мМ ЭДТА и 5 мкг/мл PMSF.

Клеточную суспензию, содержащую порядка 10^{12} клеток, подвергали трем последовательным размораживаниям-оттаиваниям с добавлением детергента Nonidet P-40 (Sigma Aldrich) до конечной концентрации 0,24% (V/V) после первого цикла. После второго цикла замораживания-оттаивания добавляли к суспензии NaCl до 800мМ и ЭДТА до 1 мМ после третьего цикла. После этого суспензию центрифугировали при 4°C , при 16000g, в течение 25 минут. Полученный супернатант диализовали против буфера, содержащего 10 мМ Трис-HCl pH 8.0, 0,1мМ ЭДТА и 0,1 мкг/мл PMSF, после чего концентрировали в пять раз на вакуумном концентраторе (Eppendorf) при комнатной температуре, диализовали повторно против того же буфера и хранили при температуре -75°C .

2.3. Метод торможения ДНК в геле

Связывание белков *M.gallisepticum* с ДНК было протестировано с помощью торможения в геле [117]. Белок, связываясь с олигонуклеотидом, тормозит его миграцию сквозь полиакриламидный гель в процессе электрофореза в нативных условиях. При этом в качестве контроля служит образец олигонуклеотида без добавления белкового экстракта. Различные количества клеточного экстракта *M. gallisepticum* (0,01-5мкл от 1мл экстракта, полученного с 10^{12} клеток) были инкубированы с 5'-FAM-меченным олигонуклеотидным зондом (50 нМ) в течение 15 мин в 8мкл связывающего буфера, содержащего 25 мМ Трис-HCl pH 8; 0, 9% глицерин и различные концентрации NaCl. Образцы были внесены в лунки 8% полиакриламидного геля (200x180x1 мм), в качестве буфера использовали 100мМ Трис-борат (pH 8,3), после чего проводили электрофорез при 555В и 11°C в течение 1 часа.

2.4. Идентификация белков, связывающих мисматчи, из клеточного экстракта *M. gallisepticum*

Белки, связывающие ДНК-мисматчи, были идентифицированы с помощью техники двумерного гель-электрофореза.

В первом направлении белки разделялись в нативных условиях: 40 мкл от 1 мл клеточного экстракта *M. gallisepticum*, полученного из 10^{12} клеток, были смешаны со 100 пмоль FAM-меченого двуцепочечного олигонуклеотида (приготовление описано выше), содержащего ошибочные спаривания. В качестве контроля использовался клеточный экстракт без добавления ДНК. Электрофоретическое разделение проводилось в полиакриламидном геле (акриламид к бисакриламиду = 30:1) с буфером, содержащим 100 мМ Трис-борат pH 8,3, как описано в предыдущем разделе.

Для дальнейшего разделения белков, связывающих специфические ДНК-структуры, проводили электрофорез во втором направлении в денатурирующих условиях. Для этого вырезанные после первого направления полоски геля инкубировали в буфере, содержащем 25 мМ Трис-HCl pH 8,0 и 0,2% SDS, в течение 10 минут, после чего помещали каждую полоску в залитый концентрирующий гель (4% акриламид, 0,13% бисакриламид, 125 мМ Трис, pH 6,8). Электрофоретическое разделение проводили в денатурирующих условиях в градиентном геле (9–16% акриламида, 0,3 – 0,53% бисакриламида, 25 мМ Трис, 192 мМ глицин, pH 8,3) длиной 20 см, прохождение электрофореза контролировали визуально по миграции красителя бромфенолового синего. По окончании электрофореза проводили окрашивание серебром (см. раздел Двумерный дифференциальный гель-электрофорез) для определения локализации белков и ДНК. Соответствующие белковые пятна были вырезаны и использовались для получения триптических пептидов.

2.5. Гидролиз белков в полиакриламидном геле

По окончании электрофореза, гель фиксировали в растворе, содержащем 20% этанола и 10% уксусной кислоты, в течение 30 мин, а затем тщательно отмывали от фиксирующего раствора водой. Участок геля, содержащий белки, разрезали на кусочки размером 1 x 1 мм и переносили в пробирки объемом 200 мкл. Кусочки геля с белками обрабатывали раствором 10 mM ДТТ в 100 mM NH_4HCO_3 в течение 30 мин при 56°C для восстановления дисульфидных связей, а затем алкилировали, с добавлением йодацетамида до конечной концентрации 55 mM в 100 mM NH_4HCO_3 в темноте, в течение 20 мин. После этого из геля удаляли воду 100% ацетонитрилом.

Для гидролиза к высушенным кусочкам геля добавляли по 150 мкл раствора, содержащего 40 mM NH_4HCO_3 , 10% ацетонитрил и 20 нг/мкл трипсина (TrypsinGold, mass spectrometry grade, Promega). Образцы инкубировали сначала 60 мин при 4°C , а затем 16-18 ч при 37°C . Для последующего МАЛДИ-МС анализа пептиды экстрагировали из геля 0,5% ТФУ. Для хромато-МС-анализа на квадруполь-времяпролётном масс-спектрометре пептиды экстрагировали из гелей трехкратно следующими растворами: 5% HCOOH , и дважды 50% ацетонитрилом с 5% муравьиной кислотой. Экстракты объединяли и высушивали досуха на вакуумном концентраторе, при 45°C . Полученные осадки растворяли в 50 мкл раствора, содержащего 95% H_2O , 5% ацетонитрила и 0,1% муравьиной кислоты.

2.6. Идентификация белков методом МАЛДИ-МС

Аликвоты (1мкл) образцов были нанесены на стальную мишень GroundSteel с добавлением 0,3мкл полунасыщенного раствора 2,5-дигидроксibenзойной кислоты (Aldrich) в 30% ацетонитриле с 0,5% ТФУ к каждому, после чего капли высушивали при комнатной температуре.

Ионизацию проводили методом МАЛДИ (матрица-активированной лазерной десорбционной ионизации). Масс-спектры регистрировали с помощью времяпролетного масс-спектрометра MALDI-TOF UltraflexII (Bruker Daltonik) снабженного Nd лазером. $[MH]^+$ ионы измеряли в рефлекторной моде, точность моноизотопной массы пика составляла 50 ppm. После этого проводили пептидный фингерпринт – поиск с помощью сервиса Mascot (<http://www.matrixscience.com>) по белковой базе данных NCBI. В качестве допустимых модификаций пептидов выбирали окисление метионина и пропионамидирование цистеина. Кандидатный белок считался значимым, если его счет был больше 80 ($p < 0.01$). Комбинированный MS+MS/MS поиск с помощью программного обеспечения Biotools 3.0 (Bruker Daltonik) использовался для повышения вероятности идентификации, если это было необходимо.

2.7. Расчет констант диссоциации

Для расчета констант диссоциации мы использовали образец, содержащий ДНК-связывающий белок и ДНК-субстрат, который может формировать только один тип ДНК-белкового комплекса. Общая концентрация белка – P_0 была известна, также, как и концентрация добавленной ДНК – S_0 .

Разделяя свободную и связавшуюся с белком флуоресцентно меченую ДНК (5'-FAM) в полиакриламидном геле, мы измеряли величины соответствующих флуоресцентных сигналов, которые затем использовали для расчета абсолютного количества связанной и свободной ДНК. При этом отношение флуоресцентных сигналов связанной и свободной ДНК принимали за r ($r = [\text{связанная ДНК}]/[\text{свободная ДНК}]$).

Тогда концентрацию свободной ДНК (S_f) можно было вычислить по формуле $S_f = S_0/(1+r)$, а связанной – $S_b = S_0 \cdot r/(1+r)$.

По определению, константа диссоциации это:

$$K_d = P_f \cdot S_f / S_b$$

Где P_f – это концентрация свободного белка. Если белок формирует только специфический 1:1 комплекс с ДНК, то

$$P_0 = P_f + S_b$$

Комбинируя приведенные выше уравнения получаем:

$$S_b = P_0 - K_d \cdot r$$

И, далее

$$K_d = P_0/r - S_0/(1+r) = P_0 \cdot [свободная ДНК] / [связанная ДНК] - S_0 \cdot [свободная ДНК] / ([связанная ДНК] + [свободная ДНК]),$$

где коэффициент r – получен для каждого ДНК-субстрата в отдельном эксперименте по торможению ДНК. Данные соотношения верны даже в случае если $S_0 > K_d$, наиболее важным является точность измерения соотношения: $[свободная ДНК] / [связанная ДНК]$.

Концентрацию белка mgHU в клеточном экстракте *M. gallisepticum* оценивали путем сравнения способностей рекомбинантного белка mgHU и клеточного экстракта к связыванию одинаковых ДНК-субстратов, при допущении, что нативный (содержащийся в экстракте) и рекомбинантный белки имеют сходные константы связывания. Подробное обсуждение расчета констант диссоциации с помощью метода торможения в геле можно найти в статье [115].

2.8. Получение чистого препарата целевого белка *Hup_2* из *M. gallisepticum* в клетках *E. coli*

Для ПЦР-амплификации *hup_2* гена *M. gallisepticum* использовали следующие праймеры: прямой, NdeI сайт рестрикции подчеркнут, tgscacatatgtttattatggcaaaaatcaaatc; обратный, BamHI сайт подчеркнут, atgcggatccctatttgtgcgaatctac. Праймеры были подобраны на основании последовательности гена *hup_2* (Swiss-ProtentryQ49504), полученной при геномной аннотации. В качестве матрицы для ПЦР-амплификации использовали геномную ДНК *M. gallisepticum* штамм S6. Полученный ПЦР-продукт был подвергнут обработке эндонуклеазами рестрикции NdeI и BamHI, после чего проводили лигазную реакцию с вектором pET15b (Novagen), предварительно подвергнутым обработке теми же эндонуклеазами. После реакции лигирования, полученным вектором трансформировали компетентные клетки *E. coli*, штамм DH5 α (Life Technologies). Нуклеотидная последовательность клонированного гена была определена по каждой из цепей и совпадала для пяти независимых клонов *E. coli*. Полученную таким образом рекомбинатную плазмиду назвали pET15b-hin. Для выделения целевого белка, плазмидой pET15b-hin был трансформирован B834 (DE3) штамм *E. coli* (Life Technologies). Бактерии B834(DE3)/pET15b-hin инкубировали в 300мл LB-среды, содержащей 100мкг/мл ампициллина при 37⁰С на качалке (140 об./мин) до достижения оптической плотности 0,8 при длине волны – 600нм. Затем для индукции экспрессии добавляли IPTG до финальной концентрации 0,1мМ и инкубировали клетки еще 3 часа. Далее клетки были разрушены ультразвуком и целевой белок *hup_2* был очищен с использованием набора для аффинной хроматографии HisTrap histidine-tagged protein purification kit (Amersham Biosciences) в соответствии с протоколом производителя. Концентрация полученного белка была измерена по методу Бредфорда с помощью Quickstart Bradford dye reagent (BioRad).

Целостность полученного белка была оценена с помощью SDS-гель-электрофореза по методу Леммли [118]. Для оценки чистоты выделенного белка и идентификации последовательности аминокислот мы анализировали выделенный белок методом МАЛДИ-МС без предварительного протеолитического расщепления. Молекулярная масса исследуемого полипептида составила $13386 \pm 0,5$ Да. Следов HU-белка *E. coli* или других белков обнаружено не было.

2.9. Выделение ДНК

Клеточный осадок *M. gallisepticum*, полученный центрифугированием (10000g, 10 мин, 4⁰C) из 1 мл 24 часовой культуры, растворяли в 500 мкл СТАВ буфера (2% СТАВ, 100 мМ Трис-НСl рН=8, 20 мМ ЭДТА, 1,4 М NaCl), после чего образцы инкубировали при 60⁰C в течение 30 мин. Затем к образцам добавляли 500 мкл СНCl₃, активно встряхивали и центрифугировали при 14000g, 4⁰C, в течение 20 минут, после чего отбирали водную фазу и осаждали ДНК добавлением равного количества изопропилового спирта. Полученную ДНК растворяли в деионизованной воде, измеряли концентрацию и использовали как матрицу для ПЦР в реальном времени, а также для капельно-цифровой ПЦР [119].

2.10. Комплементационный тест генов *hur_1* и *hur_2* из *M. gallisepticum* в *E. coli*

Клетки *E. coli*, несущие делеции одновременно в двух генах (*hurA* и *hurB*), кодирующих белок HU, фенотипически характеризуются медленным ростом [120]. При этом делеция в одном из генов *hurA* или *hurB* не приводит к значительному снижению скорости роста [120], что дает возможность использовать комплементационный тест с единственным геном *hur* из другой бактерии [106,121].

Для того чтобы провести комплементационный тест с генами *hup_1* и *hup_2* из *M. gallisepticum*, мы провели следующую процедуру. Геномная ДНК *M. gallisepticum* была использована для амплификации целевых генов. Полимеразную цепную реакцию для амплификации генов *hup_1* и *hup_2* проводили с использованием Pfu-полимеразы (Promega), в течение 30 следующих циклов: 94⁰С – 1 мин, 52⁰С – 1 мин, 72⁰С – 5 мин. Для амплификации использовали следующие праймеры: для *hup_2*: GGAATTCCATATGGCAAAAATCAAATCATTAAGTGCTGCTG и GCTCTAGAGTATTATATAAGATCTTGATTAACSTATTTGTGC, для *hup_1*: GGAATTCCATATGATGCTAACTAAATCTGAAATTTGC и GCTCTAGACTAGCGGTTCTTAAAGAATCCAGCGGC.

Полученные ПЦР-фрагменты, соответствующие открытым рамкам трансляции *hup_1* и *hup_2* были очищены в агарозном геле, обработаны эндонуклеазами рестрикции NdeI и XbaI, а затем клонированы в экспрессионный вектор pBADN [101].

Мы экспрессировали по отдельности два гена, потенциально кодирующих HU-белок в *M. gallisepticum*– *HimA/Hup_2* и *HimA/Hup_1*, клонированных в плазмиды под контролем арабинозо-индуцибельного P_{ARA} промотора, названные pJO193 и pJO201, соответственно. В качестве отрицательного контроля мы использовали исходный вектор pBADN [101]. В качестве положительного контроля был использован вектор pBAD-6H-*hupA*, несущий ген *hupA* из *E. coli*, под контролем того же (P_{ARA}) промотора [121]. Четырьмя вышеописанными плазмидами были трансформированы клетки *E. coli* штамма JO3020, полученные в результате трансформации штаммы были названы JO193, JO201, JO215 и JO217. Данные штаммы были высеяны на полную среду, содержащую 0,9% агара в отсутствие и присутствии арабинозы, после 14 часов роста детектировали фенотип выросших колоний.

2.11. Выделение РНК и синтез кДНК

Тотальную РНК выделяли из суспензии клеток *M. gallisepticum* с помощью реагента Trizol LS (Invitrogen) по протоколу фирмы-производителя. Для этого к 300 мкл жидкой культуры *M. gallisepticum* добавляли 900 мкл Trizol LS, после чего образец интенсивно перемешивали до гомогенного состояния, инкубировали при комнатной температуре 15 мин, добавляли 240 мкл CHCl_3 , и центрифугировали при 14000 g, 4⁰С, в течение 15 минут. Затем отбирали водную фазу и осаждали РНК добавлением равного количества изопропилового спирта. Полученную РНК растворяли в деионизованной воде и обрабатывали ДНКазой I (Fermentas) и использовали для синтеза кДНК с помощью обратной транскриптазы Mu-MLV Reverse Transcriptase (Fermentas), согласно протоколу фирмы производителя. Полученную кДНК использовали как матрицу для ПЦР в реальном времени, а также для капельно-цифровой ПЦР [119].

2.12. Количественная ПЦР в реальном времени

ПЦР в реальном времени проводили с использованием готовой ПЦР-системы iQ SYBR Green Supermix (Bio-Rad) с добавлением праймеров (Приложение 1) по 5 пмоль каждого, а также образцов кДНК или геномной ДНК в количестве 10 нг на одну реакцию. ПЦР-реакцию проводили на амплификаторе CFX96™ Real-Time PCR Detection System (Bio-Rad) при следующих условиях: 95⁰С – 10мин, затем 94⁰С – 15с, 58⁰С – 20с, 65⁰С – 1мин, в течение 40 циклов. Детекция флуоресценции проводилась при 65⁰С в конце каждого цикла.

Экспериментальные данные ПЦР в реальном времени, представленные в данной работе были получены в трех независимых биологических повторах, для каждого из которых были выполнены также два технических повтора. Результаты представляют собой среднее всех повторных

экспериментов с определенными стандартными отклонениями (Приложение 4).

2.13. Капельно-цифровая ПЦР

Технология капельно-цифровой ПЦР (droplet-digital PCR) была использована для подтверждения результатов, полученных с помощью ПЦР в реальном времени, поскольку эта технология позволяет напрямую определять число копий молекул ДНК в образце, как описано в [119].

Для приготовления реакционной эмульсии использовали двукратную ПЦР-смесь ddPCR™ Supermix for Probes (Bio-Rad), минеральное масло Droplet Generator Oil (Bio-Rad) и пробы кДНК или геномной ДНК в количестве 10 нг на одну реакцию, а также праймеры по 5 пмоль каждого и молекулярный зонд – 2,5 пмоль на реакцию (Приложение 1). Приготовление эмульсии осуществляли на приборе QX100 Droplet Generator компании (Bio-Rad). ПЦР проводили на амплификаторе DNA Engine Tetrad 2 (Bio-Rad). Считывание результатов проводили на приборе QX100 Droplet Reader (Bio-Rad). Обработку данных проводили с помощью программы QuantaSoft (Bio-Rad). Условия получения эмульсии, а также проведения эмульсионной ПЦР соответствовали протоколу фирмы-производителя.

2.14. Определение числа копий РНК и ДНК

Измерение количества копий ДНК и РНК проводили методом ПЦР в реальном времени и методом капельно-цифровой ПЦР. Полученные данные были нормализованы на единицу объёма культуральной среды и было рассчитано соотношение числа копий РНК к числу копий геномной ДНК для каждого исследованного гена. При этом мы предполагали, что на каждую клетку приходится в среднем единственная копия геномной ДНК.

2.15. Определение предельно переносимых воздействий

Определение уровня предельно переносимых (сублетальных) воздействий было необходимо для стандартизации условий опыта и избегания гибели клеток в процессе эксперимента. Клетки *M. gallisepticum* подвергались воздействию стрессорных факторов различной интенсивности, после чего проводился тест на количество колониеобразующих единиц. Сублетальная доза воздействия определялась как наибольшая, при которой число колониеобразующих единиц оставался таким же, как и в контроле.

В результате уровни сублетальных воздействий составили: 1 час при 46⁰С – для теплового стресса, 1,2М NaCl – для осмотического стресса, 0,02% H₂O₂ – для перекисного стресса, 2 мкг в течение часа – для ципрофлоксацина, 8 мкг в течение 4 часов – для тетрациклина.

2.16. Определение кинетики роста культуры

В процессе роста культуры проводили отбор проб через каждые 3 часа, в течение 24 часов, после пересева культуры на свежую среду. В качестве параметров, отражающих процесс роста культуры, нами были выбраны скорость накопления геномной ДНК и рибосомальной РНК, которые определяли с помощью ПЦР в реальном времени, использовали праймеры к генам 16S и 23S рРНК (Приложение 1). Кинетику накопления геномной ДНК использовали в качестве индикатора скорости репликации генома. Кинетика накопления рибосомальной РНК отражает скорость роста биомассы, если исходить из предположения о том, что на определённый объём цитоплазмы приходится постоянное количество рибосом. При этом, поскольку РНК легко деградирует, в её прирост должны вносить вклад только живые клетки. В количество геномной ДНК вносят вклад, как живые, так и мёртвые клетки. Оптическую плотность культуры не измеряли, поскольку клетки *M. gallisepticum* имеют тенденцию к агрегированию по достижении

определённой плотности культуры, что может выглядеть как резкое увеличение оптической плотности, в то время как реальный прирост числа клеток гораздо меньше.

Было установлено, что время достижения культурой стационарной фазы составляет 21 час после разведения клеток средой для культивирования в соотношении 1:10. Таким образом, для проведения теплового шока на кривой роста культуры были выбраны две точки: 12 часов, когда культура находится в стадии активного роста и 24 часа, когда культура входит в стационарную фазу.

Для остальных воздействий (антибиотики, соль, перекись) измерение уровня мРНК проводили только в логарифмической фазе роста (12 часовая культура).

2.17. Праймеры и молекулярные зонды для количественной ПЦР

Нуклеотидные последовательности праймеров и зондов приведены в Приложении 1. Дизайн праймеров производился с использованием открытого программного обеспечения PerlPrimer [<http://perlprimer.sourceforge.net/>].

2.18. Получение клеточного экстракта и разделение белков в одномерном полиакриламидном геле

Клеточный осадок дважды отмывали в среде, содержащей 150мМ NaCl, 50 мМ Трис, 2 мМ MgCl₂, рН 7,4. К клеточному осадку добавляли 20 мкл 1% раствора SDS (Helicon) в 100 мМ NH₄HCO₃, обрабатывали 15 мин в ультразвуковой бане и центрифугировали 5 мин при 10000 g и 4°C. Супернатант отбирали и определяли количество белка в нем с помощью набора Bicinchoninic acid protein assay kit (Sigma). После этого к образцам добавляли по 20 мкл 2-х кратного буфера Лэммли, содержавшего 2% SDS и

2,5% ДТТ и выдерживали при 95⁰С, в течение 5 мин. Полученный образец вносили в гель в объеме, соответствующем содержанию 50 мкг белка на лунку. Разделение белков проводили с помощью одномерного электрофореза в денатурирующих условиях в 7,5% ПААГ [122], при силе тока 10 мА на гель длиной 10 см и толщиной 0,1 см. Электрофорез останавливали, когда фронт краски достигал 1,5 см ниже концентрирующего геля.

Всего было получено 2 образца – два технических повтора электрофоретического разделения в геле.

2.19. Хромато-масс-спектрометрия

Хромато-масс-спектрометрические эксперименты были проведены на квадруполь-времяпролётном масс-спектрометре TripleTOF 5600+ (ABSciex) с источником ионов NanoSprayIII, совмещенном с хроматографической системой NanoLC Ultra 2D+ (Eksigent). Каждый образец анализировали методом хромато-масс-спектрометрии в одном техническом повторе. Хроматографическое разделение проводили в градиенте ацетонитрила в воде с добавлением 0,1% муравьиной кислоты (от 5 до 40% ацетонитрила за 120 мин) на колонках 75 мкм×150 мм, содержащих сорбент Phenomenex Luna C18 3мкм, при скорости потока 300 нл/мин.

Для анализа образцов триптических гидролизатов использовали IDA-режим масс-спектрометра. На основании первого обзорного MS1 спектра (диапазон масс измерения и отбора ионов для последующего MS2 анализа 300-1250 m/z, накопление сигнала 250 мс) отбирались 50 родительских ионов с максимальной для текущего MS спектра интенсивностью для последующего фрагментного MS/MS анализа (разрешение квадрупольного UNIT – 0,7 а.е.м., диапазон измерения масс: 200-1800 m/z, оптимизация фокусировки ионного пучка на получение высокой чувствительности, накопление сигнала 50 мс для каждого отобранного родительского иона). Для столкновительной

диссоциации использовался азот и фиксированная средняя энергия столкновений 40 В вне зависимости от параметров отобранных ионов. При этом за время накопления сигнала (50 мс) энергия столкновений линейно изменялась от 25 до 55 В. Проанализированные в MS2 режиме родительские ионы (их m/z) исключались для повторного отбора на 15 сек.

2.20. Анализ масс-спектрометрических данных

Для идентификации белков микоплазмы (идентификации пептидов – триптических фрагментов белков) каждый образец был проанализирован в одном техническом повторе. Полученные сырые данные .wiff были проанализированы в программе ProteinPilot 4.5 revision 1656 (ABSciex) с помощью поискового алгоритма Paragon 4.5.0.0 revision 1654 (ABSciex) с использованием стандартных параметров для поиска по базе данных белков на основе полногеномной сборки *Mycoplasma gallisepticum* S6 (genbankid: AFFR01000000). Были использованы следующие параметры поиска: алкилирование цистеинов – йодацетамид, гидролиз с помощью трипсина; оборудование: TripleTOF5600, вид организма не конкретизировался, так как использованная база данных содержала последовательности белков только вида *M. gallisepticum*, глубокий поиск с последующим дополнительным статистическим анализом достоверности идентификации. Группировка спектров осуществлялась с параметрами по умолчанию алгоритмом ProGroup, встроенным в ProteinPilot. Статистический анализ достоверности идентификаций (и определение порогового значения unusedscore) проводился с помощью алгоритма ProteomicS Performance Evaluation Pipeline Software (PSPEP), также встроенного в программу ProteinPilot.

Полные данные масс-спектрометрического анализа сохранены на сервере веб-сервиса Proteome Xchange Consortium (<http://proteomecentral.proteomexchange.org>) с помощью хранилища

PRIDE[123]. Идентификаторы набора данных: PXD000249 и DOI 10.6019/PXD000249.

2.21. Двумерный гель-электрофорез с дифференциальной окраской цианинами

Перед постановкой двумерного электрофореза к клеточному осадку, полученному с 4 мл жидкой культуры *M. gallisepticum*, добавляли смесь нуклеаз. Затем клетки солюбилизировали в буфере для приготовления электрофоретических образцов, содержащим 8 М мочевины, 2 М тиомочевину, 4% раствора 30% CHAPS + 10% NP40 в 10 мМТрис (pH 8,0) и смесь ингибиторов протеаз ProteaseInhibitorMix (GE Healthcare), в течение 5 мин при +4⁰С. Образцы центрифугировали при 15000 g, в течение 15 мин. Полученные супернатанты отбирали и измеряли концентрацию белка по методу Бредфорда с помощью Quickstart Bradford dye reagent (BioRad). Мечение белков цианинами CyDye3-DIGE ($\lambda_{ex}=(512);550$ nm $\lambda_{em}=570;(615)$ nm) и CyDye5-DIGE ($\lambda_{ex}=(625);650$ nm, $\lambda_{em}=670$ nm) проводили согласно рекомендациям фирмы-изготовителя (GE Healthcare). После остановки реакции связывания цианинов с белком с помощью 10мМ раствора лизина проводили одномерный электрофорез окрашенных образцов для проверки эффективности окрашивания. Перед постановкой первого направления электрофореза смешивали образцы один к одному, добавляли ДТТ (дителиотреитол) до 100 мМ и амфолины 3-10 до 1%.

Изоэлектрофокусирование проводили в стеклянных восемнадцатисантиметровых трубочках ($d_{внутр.}= 1,5$ мм). Приготовленный раствор полиакриламидного геля для ИЭФ вносили в трубочки с помощью иглы диаметром 1 мм. После окончания полимеризации геля шприцем наносили заранее подготовленные пробы и сразу же аккуратно наслаивали верхний буфер до края трубочки. Фокусировку осуществляли в следующем режиме: 100 В – 200 В– 300 В – 400 В – 500 В – 600 В – по 45 мин; 700 В – 10

ч; 900 В – 1 – 2 ч. Состав «верхнего буфера»: деионизированная вода, 50 мМ NaOH. Состав «нижнего буфера»: деионизированная вода, 20 мМ H_3PO_4 .

После изоэлектрофокусирования трубочки уравнивали в буфере, содержащем 6 М мочевины, 30% глицерин, 6,25 М Трис – HCl, pH 6.8, 2% SDS, бромфеноловый синий, в течение 20 минут. Затем трубочки переносили на поверхность градиентного полиакриламидного геля и закрепляли 0.9% агарозой, приготовленной на том же буфере, что и гель. Электродный буфер содержал 25 мМ Трис, 192 мМ глицин, 0,1% SDS, pH=8,3. Электрофорез проводили при охлаждении (10°C) в следующем режиме из расчета на одно стекло (200 x 200 x 1 мм): 20 мА– 20 мин; 40 мА– 2 ч; 35 мА– 2.5 ч.

По окончании электрофореза гель вынимали из стекол и сканировали на сканере TyphoonTrio (Amersham) при длинах волны лазера 532 нм (зеленая флуоресценция) и 633 нм (красная флуоресценция). Затем гели маркировали и окрашивали серебром [124] для последующего вырезания белковых пятен и их идентификации. Полученные изображения анализировали с помощью программы PDQuest 8.0 (Bio-Rad). Точками отличия считались точки, интенсивность флуоресценции которых в группах контроля и опыта отличалась в среднем более, чем в 2 раза (по итогам трех независимых повторов).

После окрашивания вырезали соответствующие белковые точки из геля, подвергали гидролизу трипсином и последующей МАЛДИ-МС детекции, согласно описанным выше процедурам.

2.22. Сравнительный анализ и реконструкция системы репарации

На основании анализа литературы и базы данных Gene Ontology (<http://www.geneontology.org/>) мы составили список всех генов, вовлеченных в репарацию ДНК у *E. coli* и (или) *Bacillus subtilis*. Для каждого гена из этого списка мы провели поиск гомологов в геноме *M. gallisepticum*, используя

алгоритмы blastn и blastp (e-value <1e-25). Для всех найденных гомологов были произведены выравнивания аминокислотных последовательностей (ClustalW2 алгоритм) с соответствующими им гомологами из *E. coli* и (или) *B. subtilis* с целью анализа аминокислотных замен в активных центрах белков. Аминокислоты активного центра белков определяли с помощью базы данных PDB (<http://www.rcsb.org/pdb/home/home.do>). Перечень и возможная функциональная роль найденных участников системы репарации ДНК у *M. gallisepticum* представлены в табл.5 раздела «Результаты и обсуждения». Результаты множественных выравниваний представлены в Приложении 4.

2.23. Статистический анализ изменений уровней мРНК

Для определения генов, значимо меняющих уровень экспрессии при стрессе, использовались значения, полученные методом ПЦР в реальном времени. Данные представляют среднее значение, полученное в результате трех индивидуальных биологических повторов с двумя техническими повторами каждого (результаты с рассчитанными стандартными отклонениями представлены в Приложении 2). Уровень мРНК (log₂), определенный в каждом типе стресса, сравнивался с уровнем мРНК для каждого из исследованных генов в контроле – культуре клеток в экспоненциальной фазе роста. Для теплового шока в стационарной фазе в качестве контроля использовался уровень мРНК клеток в стационарной фазе без дополнительных воздействий. Для определения значимости изменения использовался t-тест с последующей коррекцией на множественное тестирование методом Бенджамини-Хохберга [125], гены с q-value не более 0,05 считались значимо меняющими экспрессию при данном стрессе.

2.24. Кластеризация генов по паттернам изменения экспрессии при тепловом шоке

Чтобы объединить гены, похожим образом меняющие экспрессию на уровне мРНК при тепловом шоке, использовалась следующая процедура. Численные значения (\log_2), соответствующие уровням мРНК генов при 15-ти и 30-ти минутном тепловом шоке, были усреднены. Далее для каждого гена в каждом из трех состояний (контроль, 5 минут стресса и 15-30 минут стресса) был вычислен порядок уровня мРНК (состояние, соответствующее максимальному уровню мРНК, получало порядок равный 1, состояние, соответствующее минимальному уровню мРНК, получало порядок равный 3). Гены, которые во всех трех состояниях имели идентичные порядки уровней мРНК, объединялись в один паттерн (рис. 13В).

2.25. Измерение АТФ

Для измерения внутриклеточного уровня АТФ клеточную суспензию *M. gallisepticum*, содержащую порядка 10^8 клеток растворяли в стерильном DMSO (Диметилсульфоксид). Далее концентрацию АТФ измеряли с помощью фирменного набора «Люмтек» и люминометра модели ЛЮМ-1 согласно инструкциям фирмы производителя (Люмтек). В качестве отрицательного контроля при измерении уровня АТФ использовали среду для культивирования *M. gallisepticum*.

2.26. Флуориметрическое определение скорости образования перекиси водорода

Продукция H_2O_2 была измерена с использованием набора Amplex Red hydrogen peroxide assay kit в соответствии с инструкциями фирмы-производителя (Invitrogen). Измерения осуществляли с помощью спектрофлуориметра F-2700 (Hitachi).

3. Результаты и обсуждение

Одна из важнейших биохимических систем репарации ДНК – система mismatch-репарации (MMR) не обнаружена у микоплазм. Данная система необходима для коррекции замен оснований, инсерций, делеций как у бактерий, так и у высших организмов. Поэтому в первой части нашего исследования мы решили провести экспериментальный поиск белков, способных специфически узнавать и связывать типичные для пути MMR повреждения ДНК.

Для того, чтобы реализовать задачу поиска белков, связывающих ДНК с неправильным спариванием, мы разработали экспериментальный подход, состоящий из двух этапов. На первом этапе, используя специфический, содержащий ошибочно спаренные нуклеотиды, ДНК-субстрат, мы проводили его торможение в геле клеточным экстрактом *M. gallisepticum*. На втором этапе мы проводили разделение белков, связывающих ДНК, в денатурирующих условиях с их последующим вырезанием из геля, трипсинолизом и масс-спектрометрической идентификацией методом пептидного фингерпринта.

3.1. *M. gallisepticum* обладает белками, способными распознавать ошибочно спаренные основания в ДНК

Эксперименты с клеточным экстрактом *M. gallisepticum* по торможению в геле 5'-меченной дцДНК, содержащей однонуклеотидное неправильное спаривание продемонстрировали наличие двух основных белковых факторов, способных связывать и тормозить подвижность меченной ДНК. Результат этих экспериментов представлен на рис. 5. Верхняя полоса имеет одинаковую интенсивность для всех ДНК-субстратов вне зависимости от наличия ошибочного спаривания. Подобный паттерн соответствует неспецифическому связыванию двуцепочечной ДНК.

Напротив, второй белковый компонент, ответственный за нижнюю «заторможенную» полосу связывает олигонуклеотиды, содержащие ошибочное спаривание, более эффективно по сравнению с каноническим дуплексом. При этом связывающая активность является сиквенс-специфичной: содержащий СС пару олигонуклеотид демонстрирует лучшее связывание с белком.

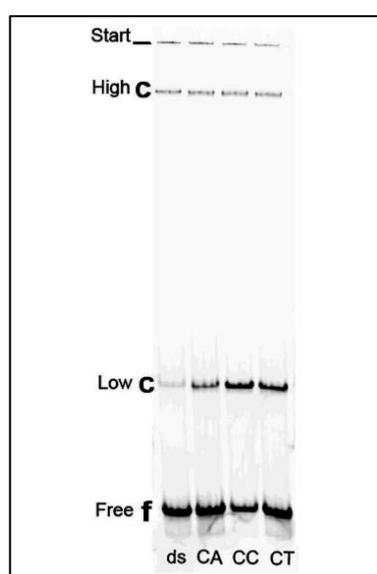


Рис.5. Анализ связывающей активности ДНК с неправильным спариванием в клеточном экстракте *M. gallisepticum*. ДНК-белковые комплексы мигрируют в геле медленнее, чем свободная ДНК, создавая дополнительные «заторможенные» полосы. Символы f и c обозначают свободную и связанную с белком ДНК, соответственно. Использовали меченый олигонуклеотид длиной 48 п.о., содержащий цитозиновый остаток в середине, который гибридизовали с комплементарным олигонуклеотидом, содержащим в качестве основания в соответствующей позиции: гуанин (линия ds), аденин (CA), цитозин (CC) или тимин (CT). Данные дцДНК-фрагменты были инкубированы с экстрактом *M. gallisepticum* в буфере, содержащем 40мМ NaCl (детальные пояснения в разделе «Материалы и методы»).

3.2. Очистка и идентификация мисматч-связывающих белков *M. gallisepticum*

Идентификация белковых компонентов, которые распознают ошибочные спаривания в ДНК, была проведена путем очистки целевых белков двумерным гель-электрофорезом с последующей масс-спектрометрической детекцией белковых пятен методом белкового фингерпринта с использованием времяпролетной масс-спектрометрии с МАЛДИ ионизацией. Метод МАЛДИ-масс-спектрометрической детекции обладает высокой чувствительностью и позволяет идентифицировать белки в окрашенных серебром гелях по характерному паттерну триптических пептидов. Результаты эксперимента представлены на рис.6.

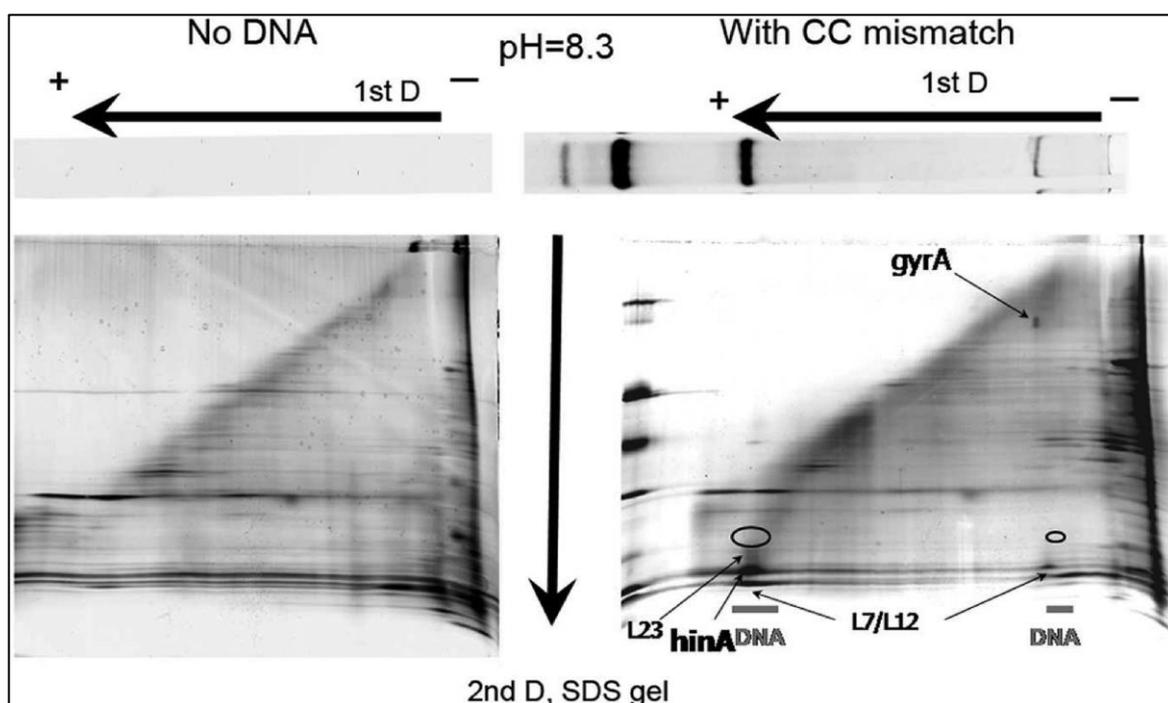


Рис. 6. Анализ «мисматч»-связывающих белков *M. gallisepticum* и их очистка для МС-идентификации. Клеточный экстракт *M. gallisepticum* без ДНК или с добавлением ДНК, содержащей СС-пару (указано) был нанесен на полиакриламидный гель и подвергнут электрофоретическому разделению в первом направлении в нативных условиях. ДНК была визуализирована (верхняя часть рисунка), после чего обе полоски геля были вырезаны и прикреплены к полиакриламидному гелю, содержащему SDS, после чего подвергнуты гель-электрофорезу в денатурирующих условиях. Полученные гели были отсканированы для определения положения ДНК (помечены овалами) и, затем, окрашены серебром для детектирования белков. Белковые пятна, помеченные стрелками, были вырезаны из геля, подвергнуты трипсинолизу. Полученные после триптического

гидролиза пептиды анализировали с помощью масс-спектрометрической методики пептидного фингерпринта или МС/МС-анализа.

Электрофорез в первом направлении проводили в нативных условиях при рН 8.3 (две верхние полосы на рис.6). При данном значении рН только отрицательно заряженные белки с $pI < 8$ могут входить в гель. ДНК-связывающие белки обычно характеризуются высоким значением $pI (> 8)$ и должны оставаться на входе в гель, если к ним не добавить ДНК-субстрат. В данном случае ДНК-субстрат (дуплекс, содержащий СС-пару) был добавлен только к образцу, соответствующему правой части рис.6, в то время как в другой образец (левая часть), ДНК не добавляли. После проведения первого направления электрофореза, гели были отсканированы для позиционирования меченой ДНК. Четырем флюоресцентно-окрашенным полоскам на правой части соответствуют слева-направо: свободная оцДНК, свободная дцДНК, специфический ДНК-белковый комплекс и неспецифический ДНК-белковый комплекс.

После прохождения второго направления (в присутствии SDS) гель был зафиксирован и окрашен серебром (две нижние части на рис.6). Положения пятен, соответствующих ДНК, были определены с помощью сканирования и помечены на рис.6 овалами. Сравнение двумерных гелей, изображенных на рис.6, показывает, что некоторые белковые пятна, присутствуют на правом и отсутствуют на левом геле. Данные белки считали кандидатами в ДНК-связывающие белки, соответствующие им пятна вырезали, подвергали трипсинолизу в геле и идентификации с использованием технологии пептидного фингерпринта. В качестве анализатора служил времяпролетный масс-спектрометр UltraFlexII (BrukerDaltonics). Белок, соответствующий верхнему комплексу при первом направлении электрофореза был идентифицирован как А субъединица ДНК гиразы. Белок, соответствующий специфическому ДНК-белковому комплексу в первом направлении, был идентифицирован как продукт гена *hinA/hup_2* (Swiss-ProtentryQ49504). Ген

hup_2 кодирует полипептид с молекулярной массой 9 кДа, аннотированный ранее как гомолог HU белка *E. coli* (cd00591 SequenceCluster, HU_INF) [126]. Следует отметить, что в геноме *M. gallisepticum* представлен еще один гомолог – *hinA/hup_1* (Swiss-ProtentryQ7NBW7), кодирующий полипептид массой 10кДа. Интересно, что мы не нашли ни одного пептида от белка HimA/Hup_1 в пятне на геле, соответствующем белку *hup_2*. Мы считаем, что HimA/Hup_1 не связывает ДНК, содержащую ошибочные спаривания и, при этом, не образует гетеродимера с белком HimA/Hup_2, обладающим таковой активностью.

		1	10	20	30	40
consensus	-	*n*n***nn*nnnnn.....*n*nnnnnn*nn**nn*nnn*nn**n				
mgHU	MAKIKS	LSAAEYLKEMADETN.....LKVQDIRLVVTSLQKVLAKELATTGE				
hup_1	M	LTKSEICKIIAECTG.....VSPKLVKACFQVYSDLVKKKIKSQGQ				
aclHU	-	MNKTELVALVADKAE.....VTQAMAQKVVNSFVDVVTETLSKDEK				
bsuHU	-	MNKTELINAVAEASE.....LSKKDATKAVDSVFDITILDALKNQDK				
TF1	-	MNKTELKAIQAQDTE.....LTQVSVSKMLASFEEKITTEVAKGDK				
ecoHUα	-	MNKTLQIDVIAEKAE.....LSKTQAKAALESTLAAITESLKEGDA				
ecoHUβ	-	MNKSQIDKIAAGAD.....ISKAAAGRALDAIIASVTESLKEGDD				
T.maritima	-	MTKKELIDRVAKKAG.....AKKKDVKLILDITILETITEALAKGEK				
Hbb	MSFSRRPK	VTKSDIVDQIALNIKNNNLKLEKKYIRLVIDAFFEELKSNLCSNNV				
Anabaena	-	MNKGELVDAVAEKAS.....VTKKQADAVLTAALETIIEAVSSGDK				
		50	60	70	80	90
consensus		*n*n***n*n*nn*nn*n.n*****nn*n*nnnnnnn**n***n**nn*				
mgHU		VRLFDIGKFKLVATKPRT.GINPKTKQKIQIPAGKKIKLTVSKILTDVDSHK				
hup_1		VRLPELGTFRVTIGRERI.SVNPITGAQTRIPPKPKVKFRAAKPLKEVTATIKWKYVSEDELLQPKRKATVFFKNR				
aclHU		VVVTGFGTFFVNRVARR.GKNPRTGEEIIVPAQKTPAFKAGKLLKDAVK				
bsuHU		IQLIGFGNFVRRERSARK.GRNPQTGEEIEIPASKVPAFKPGKALKDAVAGK				
TF1		VQLTGFLNIKPVARQARK.GFNPQTQEALEIAPSVGVSVKPGESLKKAAEGLKYEDFAK				
ecoHUα		VALVGFGTFFKVNHRAERT.GRNPQTGKEIKIAAANVPAPFVSGKALKDAVK				
ecoHUβ		VALVGFGTFFAVKERAART.GRNPQTGKEITIAAAKVPSPFRAGKALKDAVN				
T.maritima		VQIVGFSGFVVRKAAARK.GVNPQTRKPIITIPERKVPKFKPGKALKEKVK				
Hbb		IEFRSFGTFFVRRKGRNLARNPQTGEYVKVLDHNVAYFRPGKDLKERVWGIKG				
Anabaena		VTLVGFSGSFESRERKARE.GRNPKTNEKMEIPATRVPAFSAAGKLFREKVAPP				

Рис. 7. Выравнивание аминокислотных последовательностей гомологов HU. Нумерация аминокислотных остатков соответствует таковой в белке HU *B.subtilis* (bsuHU). В консенсусной последовательности консервативные и неконсервативные остатки обозначены звездочкой и буквой n, соответственно. Точками обозначены разрывы в выравниваниях. Для выравнивания использовали гомологи HU-белка из следующих микроорганизмов: *M. gallisepticum* (mgHU), *Acholeplasma laidlawii* (aclHU), *B. subtilis* (bsuHU), *B. subtilis bacteriophage SPO1* (TF1), *E. coli* (ecoHUα и ecoHUβ), *Thermotogamaritima*, *Borrelia burgdorferi* (Hbb), *Anabaena*.

Далее белок *HinA/Hup_2* мы будем называть *mgHU*. В пятнах геля, соответствующих ДНК-белковым комплексам в первом направлении, мы также идентифицировали другие белки – это рибосомальные белки L23 и L7/12 (рис.6). Мы предполагаем, что они неспецифически связывают ДНК за счет своего высокого положительного заряда.

На рис.7 приведено сравнение аминокислотной последовательности белков *mgHU* и *Hup_1 M. gallisepticum* с гомологами из других бактериальных видов. Некоторые свойства последовательности *mgHU* отличают ее от гомологичных: белок сильно заряжен (+10) по сравнению с каноническими HU (+2 у *bsuHU* и +4 у *ecoHU*). Часть этого избыточного заряда происходит от дополнительного N-концевого гексапептида МАКІKS. Похожий N-терминальный «хвост» наблюдается также у HU-гомологов у других молликут, относящихся к кластеру *Pneumoniae group* [5]. N- и C-концы взаимодействуют с ДНК и могут значительно влиять на свойства HU-подобных белков [101]. В отличие от белка *mgHU*, продукт гена *hup_1* обладает удлиненным C-концом из 24 аминокислотных остатков и не имеет дополнительного МАКІKS-мотива, что позволяет предполагать различную функциональную активность двух белков.

3.3. Связывание ДНК, содержащей неправильные спаривания, клеточным экстрактом *M. gallisepticum* и очищенным белком *mgHU*

Способность *mgHU* связывать «поврежденную» ДНК была подтверждена с помощью очищенного белка. Для этого ген *himA/hup_2* из *M. gallisepticum* был клонирован и экспрессирован в *E. coli*, соответствующий белок был очищен, как описано в разделе «Материалы и методы». Связывание *mgHU* с различными дцДНК-субстратами, содержащими типичные повреждения, продемонстрировано на рис.8. Панели А и В показывают сравнение способности клеточного экстракта *M. gallisepticum* и

чистого препарата белка mgHU вызывать торможение ДНК в геле. Представленные на рис. 8 паттерны связывания демонстрируют сходные свойства очищенного белка и клеточного экстракта. Исходя из полученных нами данных, был сделан вывод, что белок HimA/Hup₂, специфически связывающий СС-спаривание, также обладает способностью связывать и ДНК, содержащую другие ошибочные спаривания, представленные на рис.8. Подобный эксперимент с продуктом гена *himA/hup₁* продемонстрировал неспособность данного белка связывать ДНК (данные не показаны).

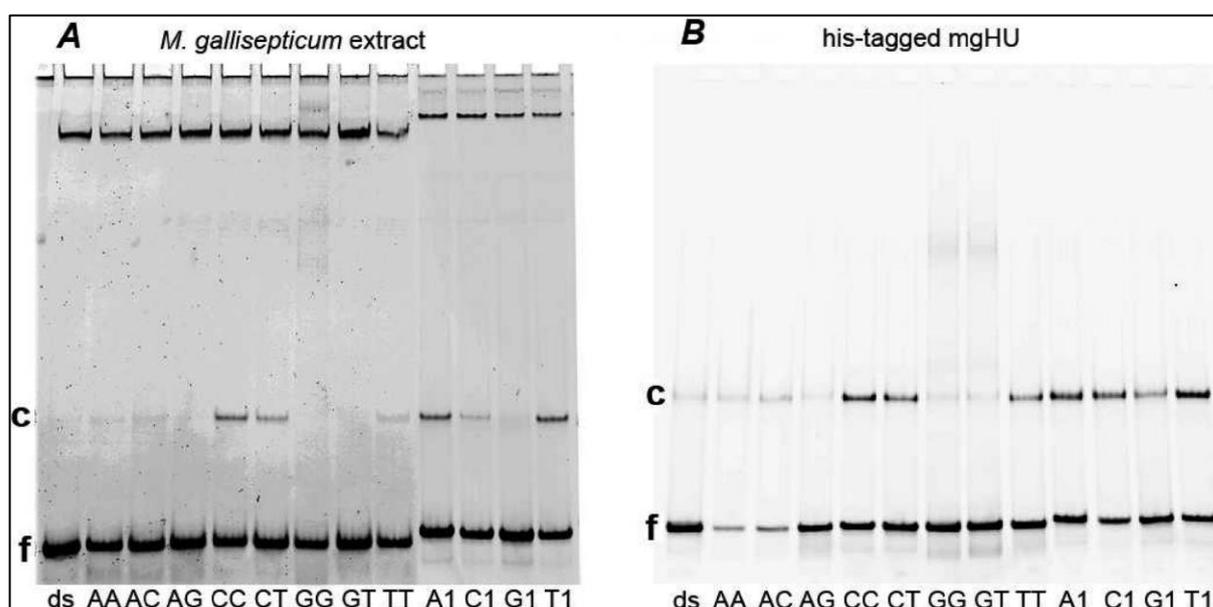


Рис.8. Сравнение профилей связывания ДНК клеточным экстрактом *M. gallisepticum* и очищенным белком mgHU в геле, содержащем 100мМ Трис-бората. На панели А представлен паттерн связывания с клеточным экстрактом *M. gallisepticum* (0,25мкл на дорожку) с различными дцДНК-фрагментами, содержащими ошибочное спаривание (замену или инсерцию одного нуклеотида) в центре. Метки А1, С1, G1 и Т1 означают наличие инсерции нуклеотида с соответствующим азотистым основанием. Панель В представляет паттерн связывания для очищенного his-HU (2мМ на дорожку) с теми же ДНК-субстратами, что и на панели А. Все пробы содержали 150мМ NaCl. Остальные обозначения такие же, как на рис.5.

Белок mgHU демонстрирует способность эффективно и структуро-специфически связывать дцДНК, содержащие инсерции А, С или Т (рис.8). Репарация данных повреждений чрезвычайно важна, поскольку эти повреждения приводят к летальным мутациям со сдвигом рамки трансляции

белка. Продемонстрированное связывание однонуклеотидных неправильных спариваний является сиквенс-специфичным: пиримидин-пиримидиновые спаривания связываются значительно лучше. На основании этих результатов, мы смогли идентифицировать mgHU как белковый фактор, способный специфически связывать ДНК, содержащую ошибочные спаривания (экспериментальные результаты приведены на рис.5 и 6). Используя белок HU из *E. coli* в качестве образца, мы показали, что mgHU, как и все HU-подобные белки, взаимодействует с ДНК в виде димера (данные не приведены).

При физиологическом значении ионной силы (150мМ NaCl) белок mgHU связывает ДНК, содержащую ошибочное спаривание, по крайней мере, в 100 раз сильнее, чем каноническую дцДНК. Мы предполагаем, что mgHU является основным белком в экстракте *M. gallisepticum*, связывающим ДНК с ошибочными спариваниями. Другие белки, способные связывать неправильное спаривание (если такие белки есть) должны иметь очень низкую концентрацию и/или малое значение константы связывания. Хотя очищенный из *E. coli* белок mgHU демонстрирует сходный ДНК-связывающий паттерн с клеточным экстрактом, для более детальной характеристики нативных свойств mgHU мы использовали клеточный экстракт *M. gallisepticum*.

3.4. Связывание белка mgHU с поврежденной ДНК сильно при физиологическом значении ионной силы

Из предыдущих исследований свойств белка HU *E. coli* известно, что неспецифическое связывание белка с ДНК-дуплексом значительно ослабевает с увеличением концентрации NaCl (ионной силы) и практически отсутствует при «физиологических» концентрациях одновалентных ионов [96,112,114]. Напротив, специфическое связывание HU-белка с неканоническими дцДНК-структурами остается значительным даже при

высоких концентрациях соли [109,110]. Зависимость связывания белка mgHU с некоторыми «повреждениями» дцДНК от концентрации соли представлена на рис. 9. По относительной интенсивности полос, отмеченных на рис. 9 можно увидеть, что связывание белка mgHU с «поврежденной» дцДНК является сильным и специфичным при физиологических концентрациях соли (100-200мМ NaCl). При этом профиль связывания при разной концентрации соли может меняться в зависимости от конкретной структуры ДНК-субстрата.

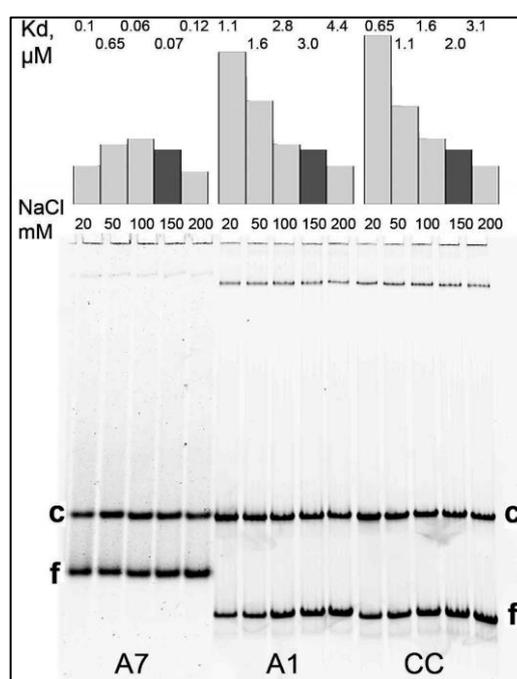


Рис.9. Зависимость связывания белка mgHU с ДНК от концентрации соли. Три дцДНК-субстрата, указанных в нижней части рисунка, были смешаны с клеточным экстрактом и нанесены на гель (0,1; 2,5 и 2,5мкл экстракта на дорожку для A7, A1 и CC соответственно). Буфер для связывания содержал различные концентрации NaCl (20, 50, 100, 150 и 200мМ). В качестве буфера для приготовления геля и проведения электрофореза использовали 100мМ Трис-борат. На верхней части рисунка представлены диаграммы относительной интенсивности полос, соответствующих mgHU-дцДНК комплексам. Столбцы в диаграммах для разных ДНК-субстратов нормализованы по отдельности и не могут быть сравнены между собой. Над каждым столбцом в диаграммах указаны значения соответствующих констант диссоциации в микромолях.

Диаграммы на рис. 9 показывают, что с увеличением концентрации NaCl от 20 до 200мМ, наблюдается 4-х и 4.5-кратное снижение аффинности

для структур, содержащих инсерцию A1 и неканоническую пару СС, соответственно. Связывание с ДНК-субстратом, содержащим вставку из 7 адениловых нуклеотидов (A7), качественно отличается от двух других. В данном случае аффинность белка к ДНК возрастает с увеличением концентрации соли, достигая максимума при 100мМ NaCl. Причины данного эффекта остаются для нас неясными.

3.5. Связывание различных ДНК-структур, типичных для репарационного пути MMR

Ошибки репликации приводят к повреждениям ДНК нескольких определенных типов. Среди них наиболее частыми являются однонуклеотидные инсерции и ошибочные спаривания. Более протяженные повреждения происходят с более низкой вероятностью.

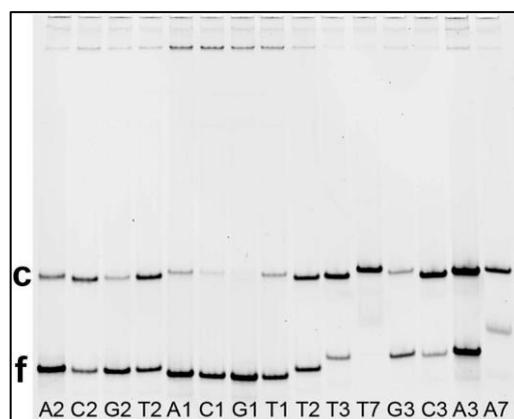


Рис.10. Паттерн связывания белка mgHU с ДНК-субстратами, содержащими нуклеотидные инсерции различной длины и последовательности. Экстракт *M. gallisepticum* (0.25мкл на одну дорожку) был инкубирован в буфере, содержащем 150мМ NaCl, с серией дцДНК-структур, в состав которых инсерция одного, двух, трех или семи нуклеотидов в центре. Затем ДНК-содержащие компоненты были разделены электрофорезом в геле. Обозначения такие же как на рис.5.

У *E. coli* и большинства известных бактерий подобные повреждения распознаются с помощью белка MutS, после чего запускается MMR-путь репарации ДНК (см. Введение). В геноме *M. gallisepticum* отсутствуют гены,

кодирующие гомологи белка MutS [116], при этом в предыдущих экспериментах мы показали, что большинство соответствующих повреждений ДНК может быть опознано белком mgHU.

На рис.10 показано специфическое связывание белка mgHU с фрагментами дцДНК, содержащими нуклеотидные инсерции различной длины. Данный эффект определяется последовательностью ДНК, взаимодействующей с белком mgHU. Для гомополимерных инсерций от двух до семи нуклеотидов, аффинность белка mgHU к ДНК возрастает в следующем порядке: T = C > A > G. Для наиболее частых однонуклеотидных инсерций, связывающие предпочтения белка mgHU несколько отличаются: A = T > C (Значения K_d – 1, 0.7 и 3 мкМ, соответственно). При этом инсерция G не распознается (нет отличий от неспецифического связывания с канонической дцДНК).

Отсутствие связывания с ДНК, содержащей G-инсерцию может иметь важные биологические последствия. Поэтому мы проверили, влияет ли нуклеотидный контекст на связывание G-мисматчей. На рис.11А представлено сравнение связывания белка mgHU с G-мисматчами различной длины внутри двух разных полинуклеотидных контекстов в одном и том же дцДНК-фрагменте. При этом результаты сходны и показывают, что эффект является локальным и что белок mgHU не распознает однонуклеотидные G-мисматчи. Однонуклеотидный G-мисматч был также вставлен в 10 различных позиций дцДНК-фрагмента, и в каждом случае только слабое связывание с mgHU было детектировано (не показано).

Связывание белка mgHU с ДНК, содержащей неправильно спаренные нуклеотиды, с различной длиной и составом неправильно спаренного участка было проанализировано методом торможения в геле и показано на рис. 8 и 11В. Если использовать эти данные и расположить мисматчи по степени ослабления связывания, получим следующий ряд: мисматч 4 (4 неправильно спаренных нуклеотида) = мисматч 3 = мисматч 2 = CC > CT = TT > AA = AC (в

табл. 1 представлены значения констант дисоциации K_d). Связывание с парами AG, GG и TG не отличается от неспецифического взаимодействия с двухцепочечной ДНК. Как и в случае с инсерциями нуклеотидов, связывание белка mgHU с ДНК, содержащей неправильно спаренные нуклеотиды, не зависит от нуклеотидного контекста (рис. 11С).

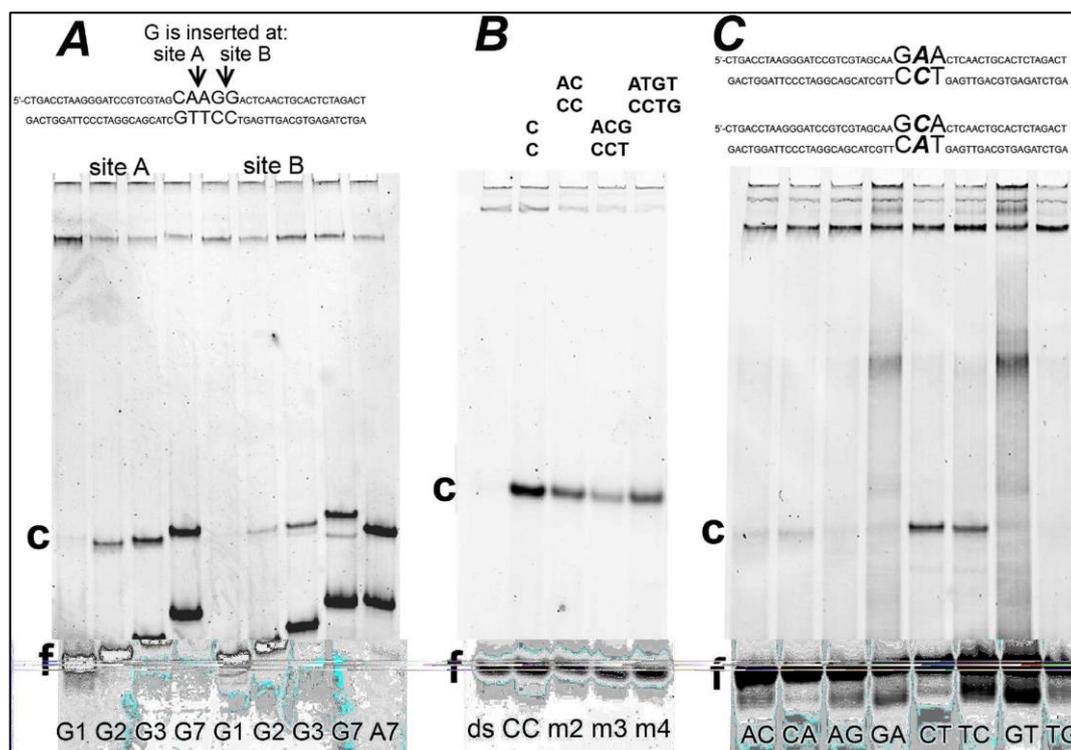


Рис.11. Паттерн связывания белка mgHU с ДНК-структурами, содержащими неправильно спаренные основания и инсерции различной длины и последовательности в разном нуклеотидном контексте. Клеточный экстракт *M. gallisepticum* был инкубирован в буфере, содержащем 150 mM NaCl с серией меченых ДНК-структур, после чего ДНК-содержащие компоненты были разделены электрофорезом в геле. Последовательности использованных дцДНК представлены в разделе Материалы и методы. Обозначения сходны с таковыми для рис.5. (А) G-инсерции различной длины, которые были введены в два различных нуклеотидных контекста, представлены в верхней части рисунка. Длина G-инсерции указана в нижней части рисунка по каждой из дорожек. Для сравнения приведены результаты аналогичного эксперимента с инсерцией семи адениловых нуклеотидов – A7. (В) Слева направо: дорожки ds, CC, m2, m3 и m4 показывают торможение в геле следующих ДНК структур: каноническая дцДНК, ДНК, содержащая неправильное спаривание одного, двух, трех и четырех нуклеотидов, соответственно. Неспаренные основания представлены в верхней части рисунка. (С) Сравнение связывания белка mgHU с ДНК-структурами, содержащими неправильное спаривание одной пары оснований, помещенной в различный нуклеотидный контекст. Последовательности приведены в верхней части рисунка для AC и CA-мисматчей.

Абсолютные значения констант диссоциации для различных комплексов белка mgHU с ДНК были рассчитаны путем сравнения относительной интенсивности полосы, соответствующий комплексу, с таковой интенсивностью комплекса белка mgHU с ДНК, содержащей A7-инсерцию, нанесенным на тот же гель. Для ДНК с A7-инсерцией значение константы диссоциации было точно рассчитано, как описано выше (см. табл. 4 и Материалы и методы).

Таблица 4. Константы диссоциации комплекса белка mgHU с ДНК, содержащей неправильно спаренные нуклеотиды и инсерции в 150 mM NaCl (значения указаны в наномолях)

Неправильные спаривания		Инсерции (выпетливания)		
Однонуклеотидное		1 нуклеотид	A	1000
CC	2000		T	700
CT	5000		C	3000
TT	10000		G	дцДНК ^a
AC	30000	2 нуклеотида	A	1300
AG	дцДНК ^a		T	160
GT	дцДНК ^a		C	150
AA	дцДНК ^a		G	2500
2-4 нуклеотида		3 нуклеотида	A	200
5'-AC-3'/5'-CC-3'	2500		T	100
5'-ACG-3'/5'-TCC-3'	3500		C	35
5'-ATGT-3'/5'-GTCC-3'	2000		G	1000
		7 нуклеотидов	A	70

			T	20
			C	20
			G	50

^aСлабое взаимодействие сходное с неспецифическим связыванием с двухцепочечной ДНК.

3.6. Комплементационный тест

Результаты комплементационного теста были получены и любезно предоставлены нашим французским коллегой – Jacques Oberto (Institut de Génétique et Microbiologie, CNRS UMR 8621, Université Paris XI, Paris, France).

Клетки *E. coli*, несущие делеции одновременно в двух генах (*hupA* и *hupB*), кодирующих белок HU, фенотипически характеризуются медленным ростом [120]. Подобный фенотип можно объяснить недавними результатами по исследованию HU-регулона у *E. coli*, в состав которого, как было показано в работе [101], входят четыре класса генов, отвечающих за приспособление к анаэробнозису, кислотному стрессу, повышенной осмолярности среды и SOS-ответ [101]. Поскольку делеция в одном из генов *hupA* или *hupB* не приводит к значительному снижению скорости роста [120], существует возможность использовать комплементационный тест с единственным геном *hup* из другой бактерии [106,121].

В нашем исследовании ген *HimA/Hup_2* был выбран благодаря своей способности распознавать повреждения ДНК, специфичные для MMR-пути репарации ДНК. Эти данные позволяют предположить, что белок mgHU является участником системы репарации ДНК. Учитывая известную плеiotропность функций белка HU *E. coli*, необходимо было проверить, обладает ли теми же способностями белок mgHU. Руководствуясь этой целью, мы протестировали, способны ли белки *HimA/Hup_2* и *HimA/Hup_1* функционально заменить HU-белок в клетках *E. coli*.

В присутствии арабинозы в качестве индуктора мы наблюдали для штамма JO193, экспрессирующего ген *HimA/Hup_2*, такой же фенотип как в положительном контроле (штамм JO215), который экспрессировал ген *hupA* (рис.12). При этом штамм JO201, несущий ген *HimA/Hup_1*, не показал комплементацию в присутствии арабинозы. В случае отсутствия арабинозы все четыре штамма проявляли типичный *hupAB* фенотип медленного роста.

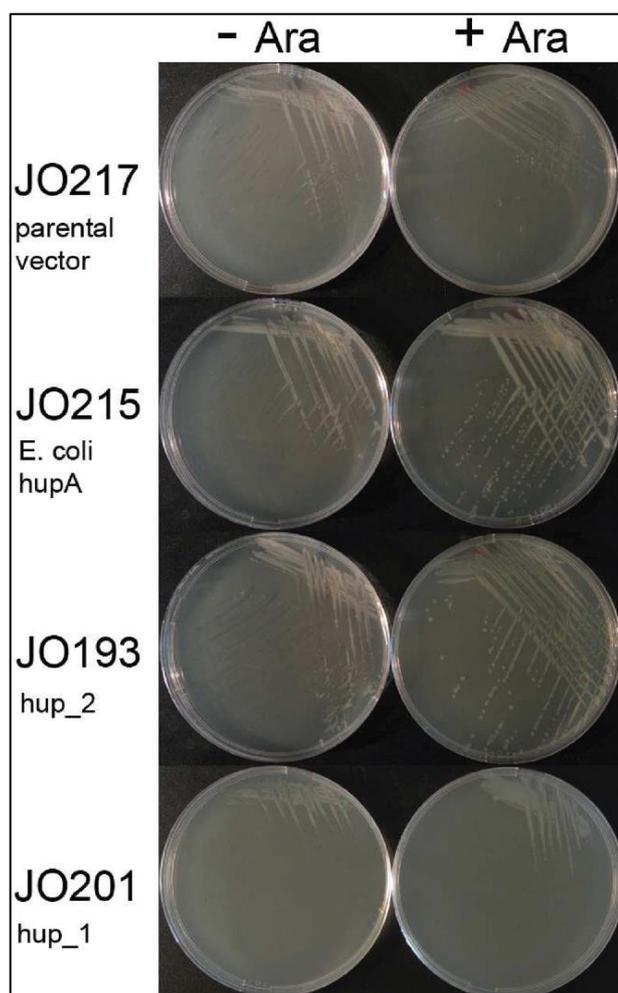


Рис.12. Комплементационный тест. Штамм *E. coli hupAB* JO3020 был трансформирован следующими плазмидами: содержащими гены *M. gallisepticum hup_2* (JO193) и *hup_1*(JO201), ген *hupAE. coli* (JO215) под контролем арабинозо-индуцируемого промотора P_{ARA} , а также исходным вектором (JO217). В отсутствии арабинозы (- Ara) все четыре полученных после трансформации штамма показывают типичный для *hupAB* фенотип медленного роста. В присутствии 0,2% арабинозы (+Ara) штамм JO193, несущий ген *hup_2* из *M. gallisepticum* показывает фенотип сходный с фенотипом штамма JO215 (положительный контроль), несущем функциональный ген *hupA* из *E. coli*. При этом штамм JO201, содержащий ген *hup_1* из *M. gallisepticum* демонстрирует фенотип медленного роста.

3.7. *Возможная роль белка mgHU в мисматч-репарации M. gallisepticum*

Микоплазмы – наименьшие известные микроорганизмы, способные к росту на искусственных питательных средах. Они характеризуются помимо прочего отсутствием ключевых элементов MMR системы репарации, необходимой для распознавания и коррекции ошибок репликации, например, таких как некорректное спаривание [127]. Тем не менее исследование клеточного экстракта *M. gallisepticum* позволило выявить белок, который способен связывать однонуклеотидные мисматчи в дцДНК-фрагментах. Данный белковый фактор был идентифицирован как гомолог HU-белка *E. coli*. Подобные белки имеют широкое распространение среди бактерий. Белок HU из *E. coli* активно исследовался на протяжении последних четырех десятилетий [101,107,108,128–133]. В последнее время стало понятным, что HU-подобные белки различных видов бактерий не идентичны. При этом характер связывания ДНК для этих белков довольно сильно варьируют среди разных видов [102].

Как было показано в серии работ, специфичность связывания HU-подобных белков может сильно и непредсказуемо меняться за счет всего лишь одной аминокислотной замены в первичной последовательности белка [134–138]. Сравнение последовательностей показывает, что белок mgHU имеет множество модификаций в консервативных участках, которые считаются критическими для структуры и специфичности белка (см. рис.7). Три особенности белка mgHU заслуживают отдельного внимания.

(I) Консервативный мотив GF_{Gn}F (позиции 46-50), а также инварианта F79 формируют гидрофобное ядро димера HU белка [107]. Выравнивание последовательностей приведенное на рис.7 позволяет предполагать, что данное гидрофобное ядро mgHU значительно отличается от белков, для которых разрешена трехмерная структура.

(II) Кроме того, N-концевая последовательность белка mgHU значительно отличается от ранее изученных HU-подобных белков. Высоко консервативная аминокислота КЗ обычно играет ключевую роль в стабилизации изогнутой ДНК-конформации [134]. Результаты компьютерного выравнивания позволяют предполагать, что в белке mgHU данная аминокислота отсутствует и добавляется гексапептид МАКИКС с N-конца. Несмотря на чрезвычайно низкий уровень гомологии последовательности, нельзя исключить, что КЗ в МАКИКС-мотиве по сути соответствует консервативному лизину из других бактерий. Однако подобная дополнительная гексапептидная петля встречается у других белков этого класса рядом с N-концом.

(III) В последовательности белка mgHU присутствует важная R61I замена в GRNPnT-мотиве (позиции 60-65). В соответствии с данными рентгеноструктурного анализа интеркалирующая аминокислота Р63 играет ключевую роль в обеспечении излома связанной молекулы ДНК [107]. Аминокислотная замена R61V была ранее изучена для белка HU из термофильной археи *Thermotoga maritima* [137]. Оказалось, что замена аргинина в 61 положении на неполярную аминокислоту снижает сродство HU-белка к интактной дцДНК. Приведенные выше три особенности повторяются для гомологов HU у других молликут.

Несмотря на значительные изменения в последовательности, белок mgHU обладает основными свойствами HU-подобных белков и соответствует ранее изученным гомологам HU из других видов. Данный белок активно связывает дцДНК с одноцепочечным разрывом (nick), оцДНК и различные неканонические ДНК-структуры, такие как структура Холлидея, а также одно- и двуцепочечные ДНК вилки (неопубликованные данные Д. Камашева). Связывание всех этих ДНК-субстратов сохраняет свою силу в широком диапазоне концентраций соли, включая физиологическую концентрацию. Хотя белок mgHU слабо связывает дцДНК, это, очевидно,

гистоно-подобный белок, поскольку обладает способностью «заменять» HU-белок *E. coli* в *in vivo* экспериментах.

Напротив, второй гомолог HU-белка, аннотированный в геноме *M. gallisepticum* (HimA/Hup_1), по всей видимости, имеет другую функцию. Вне зависимости от концентрации соли данный белок не связывает ДНК-субстраты, используемые в нашем исследовании и не способен «заменить» белок HU из *E. coli* в экспериментах по коплементации. Само наличие двух гомологов HU-белка в микоплазме вызывает удивление. Одного подобного гена достаточно для большинства исследованных бактерий, в то время как род *Mycoplasma* является продуктом редуktивной эволюции, ввиду утраты большого количества генов, не нужных для специфических условий существования этих организмов [139]. Однако, другие бактерии обычно имеют группу нуклеоид-ассоциированных белков, включая HU, IHF, Fis, H-NS и Lrp. Данные белки сильно разнятся между собой и часто выполняют взаимодополняющие функции. Например, белок H-NS способен супрессировать транскрипцию генов, вызывая образования петли ДНК рядом с промоторными регионами, в то же время белок HU способен выпрямлять ДНК и, возможно, работает как антагонист белка H-NS [140]. Поскольку в геноме *M. gallisepticum* отсутствуют дополнительные нуклеоид-ассоциированные белки помимо HU, возможно, что продукт гена HimA/Hup_1 участвует в дополнительных, так называемых «архитектурных» активностях.

Следует отметить, что биологически важные однонуклеотидные мисматчи повышают «гибкость» ДНК. У большинства организмов эти повреждения специфически распознаются участником MMR системы репарации ДНК – белком MutS. Интересно, что способность белка MutS связывать ДНК субстраты падает с увеличением «гибкости» ДНК и наиболее гибкий СС-мисматч является худшим MMR-субстратом [141]. Это наблюдение послужило толчком для скринирования *E. coli* на

специфическую СС-мисматч связывающую активность [142]. Оказалось, что СС-мисматчи могут распознаваться белками FabA и MutM, но не белком HU. В экспериментах *in vitro* была проверена способность белка HU связывать одиночные ТТ-мисматчи [143] и двойные TG-СТ-мисматчи [115]. В обоих случаях связывание оказалось сходным с таковым для неспецифической дцДНК.

Среди всех известных белков, способных связывать ДНК-мисматчи, для *M. gallisepticum*, согласно геномной аннотации, был предсказан только белок MutM. Поскольку белок MutM не был детектирован в наших экспериментах по скринингу мисматч-связывающей активности, его активность незначительна ввиду низкой концентрации или измененной специфичности. Последнее не так уж удивительно, если посмотреть на свойства белка NimA/Hup_1. В то же самое время белок mgHU способен связывать однонуклеотидные неправильные спаривания и инсерции в присутствии избытка дцДНК. При этом белок mgHU высоко представлен в клетке, поскольку это единственный нуклеоид-ассоциированный белок, и, кроме того, повреждения ДНК, возникающие в результате ошибок репликации должны быть распознаны очень эффективно. Мы считаем, что приобретенная специфичность является следствием отсутствия системы MMR у микоплазм. К примеру, бактерия *Acholeplasma laidlawii* относится к тому же классу – Mollicutes, что и *M. gallisepticum*, однако она имеет систему репарации MMR-пути. Исследование HU-белка *A. laidlawii* показывают отсутствие специфических предпочтений к MMR-субстратам [144]. Возможно, что такие свойства mgHU необходимы для снижения частоты мутагенеза, поскольку частота мутаций *M. gallisepticum* сходна с таковой для *E. coli* [6,7].

3.8. Реконструкция системы репарации ДНК у *M. gallisepticum*

На основании анализа литературы и базы данных Gene Ontology мы составили список всех генов, вовлеченных в репарацию ДНК у *E. coli* и (или) *Bacillus subtilis*. Для каждого гена из этого списка мы провели поиск гомологов в геноме *M. gallisepticum*, для того чтобы *in silico* реконструировать систему репарации ДНК в исследуемой бактерии (см. Материалы и методы).

Система эксцизионной репарации нуклеотидов

Система эксцизионной репарации нуклеотидов (NER) представлена у *M. gallisepticum* полностью как на уровне генома, так и на уровне протеома (табл. 5 и 6). Она включает в себя белки: UvrA (узнавание повреждения), UvrB (локальное плавление ДНК), UvrC (вырезание повреждённого участка), хеликазу UvrD (удаление повреждённого участка) а также ДНК-полимеразу и ДНК-лигазу (синтез ДНК и восстановление целостной двойной спирали) [30]. При этом у *M. gallisepticum*, по-видимому, используется ДНК-полимераза III (DnaE) или ДНК-полимераза IV (DinB), поскольку другие типы полимераз (ДНК-полимеразы I, II и V) отсутствуют в геноме.

Система эксцизионной репарации оснований

Система эксцизионной репарации оснований (BER) состоит из белков, узнающих специфическую модификацию ДНК и расщепляющих N-гликозидную связь, при этом поврежденное основание удаляется. Соответственно, эти белки названы гликозилазами. У *M. gallisepticum* мы обнаружили только две гликозилазы: урацил-гликозилазу Ung и формамидопиримидин-гликозилазу MutM. После удаления поврежденного основания гликозилазой образуется так называемый

апуриновый/апиримидиновый сайт. Для его репарации AP-эндонуклеаза вносит одноцепочечный разрыв 5' от AP-сайта. У *M. gallisepticum* мы обнаружили единственную AP-эндонуклеазу – Nfo. В бактериях существуют два семейства AP-эндонуклеаз. Nfo – представитель эндонуклеаз IV типа, которые имеют также 3'-фосфатазную активность, но не имеют 5'-фосфатазной активности и 3'-5' экзонуклеазной активности [37]. 3'-5' экзонуклеазной активностью обладают AP-эндонуклеазы III, найденные у всех организмов от *B. subtilis* (exoA) и *E. coli* (xthA и Nth) до человека; у *M. gallisepticum* мы не обнаружили гомолога AP-эндонуклеазы III. AP-лиаз, которые вносят одноцепочечный разрыв 3' от AP-сайта, мы также не обнаружили. Используя освободившийся 3'-ОН конец ДНК, ДНК-полимераза III и ДНК-лигаза завершают репарацию. На уровне белка у *M. gallisepticum* ранее были обнаружены все компоненты BER, кроме урацил-гликозилазы [145,146].

Система репарации ошибочно спаренных нуклеотидов

Система репарации неспаренных оснований («мисматч»-репарация) до последнего времени считалась наиболее редуцированной у всех микоплазм. У них не обнаружено гомологов ни одного из ключевых генов MutH, MutL и MutS. Ряд авторов [64] спекулировал о повышенном уровне спонтанного мутагенеза у микоплазм, однако позднее было показано, что уровень мутагенеза существенно не отличается от такового у *E. coli* [6],[7].

В нашем исследовании была показана способность гистоноподобного HU-белка (hup2) *M. gallisepticum* узнавать и связывать неспаренные основания *in vitro*, а также восстанавливать нормальный рост HU-дефицитной (hupAB(-)) *E. coli*, что может указывать на его функциональную роль в системе «мисматч»-репарации.

Мы обнаружили еще два гена, продукты экспрессии которых по своей доменной организации потенциально могут быть вовлечены в систему

репарации ДНК-«мисматчей». Это ген MGA_0195, содержащий эндонуклеазный домен, относящийся к тому же суперсемейству, что и соответствующий домен у белка MutH. А также ген MGA_0793, который содержит домен vsr, необходимый для распознавания мисматчей, содержащих неспаренный гуаниловый нуклеотид.

Характерной особенностью *M. gallisepticum* является отсутствие «канонических» экзонуклеаз (ExoI, ExoVII, RecJ, ExoX[65]), участвующих в пути MMR у *E. coli*. Однако, следует отметить, что все микоплазмы имеют «DNAPoII-подобную» экзонуклеазу Exo, фермент, гомологичный ДНК-полимеразе I *E. coli*, который потенциально может работать в системе «мисматч»-репарации. При этом в белке Exo отсутствует ДНК-полимеразный домен, присутствующий в DNAPoII *E. coli* (рис.3).

Следует отметить, что все найденные потенциальные участники «мисматч»-репарации были детектированы нами на белковом уровне (табл. б).

Система рекомбинационной репарации

В геноме *M. gallisepticum* присутствуют гены, кодирующие ключевые белки рекомбинации – рекомбиназы RecA и RecR, RecO (MGA_0016), а также гены *ruvA* и *ruvB*, кодирующие ДНК-хеликазу RuvAB (участвует в миграции ДНК-цепей), а также два гена, кодирующие ферменты, способные разрешать структуру Холлидея – ДНК-резольвазы RecU и MGA_0836. Кроме того, присутствует ген *smc*, кодирующий Smc-когезин, способный осуществлять когезию хромосом после репликации, участвуя таким образом в рекомбинационной репарации [147]. Из них на уровне белка ранее был обнаружен только RecR [145,146]. В настоящей работе нам удалось идентифицировать также белки RecA и Smc (табл. б).

Система SOS ответа

В геноме *M. gallisepticum* были найдены следующие участники SOS-системы: рекомбиназа А (*recA*), рекомбиназа R (*recR*), хеликазный комплекс (*ruvA*, *ruvB*), нуклеазно-хеликазный комплекс UvrABC, а также ДНК-зависимая ДНК-полимераза IV типа (*dinB*) с предсказанной по гомологии мутаторной активностью. В бактериях роль этой полимеразы (вместе с ДНК-полимеразой V, отсутствующей у всех представителей класса Mollicutes) – синтез ДНК с поврежденной ДНК матрицы [148]. При этом гомологи известных регуляторов бактериальной SOS-системы (LexA [149], HdiR [150]) не были найдены ни в одном из проанализированных геномов молликут. Все потенциальные участники системы SOS-ответа были обнаружены нами на уровне мРНК, при этом на белковом уровне не удалось идентифицировать только хеликазу RuvAB (табл. 6).

3.9. Представленность транскриптов генов систем репарации в клетке

Сравнительный анализ, основанный только на геномных данных, не позволяет однозначно утверждать об экспрессии гена и функциональной активности кодируемого им белка. Для того, чтобы оценить активность аннотированных генов, мы провели транскрипционный анализ генов репарации методами количественной ПЦР (капельно-цифровая ПЦР и ПЦР в реальном времени – см. Материалы и методы). Мы проанализировали транскрипцию всех генов репарации (список в табл.5) и показали, что все гены репарации транскрибируются в *M. gallisepticum* при 37⁰С в логарифмической фазе роста культуры. Мы также пересчитали количество детектируемых нами транскриптов в число копий на один бактериальный геном (табл. 5).

Таблица 5. Список генов репарации, их присутствие в транскриптоме и протеоме

Система	Название гена	Функция	Присутствие в геноме <i>M. gallisepticum</i>	Число копий мРНК на один геном	Присутствие в протеоме
Все системы	<i>ligA</i>	ДНК лигаза	+	0.03	+
«Мисматч»-репарация	<i>exo</i>	5' - 3'-экзонуклеаза	+	нд	+
«Мисматч»-репарация	<i>hup2</i>	Связывание ДНК-мисматчей	+	0.4	+
Экцизия нуклеотидов + SOS ответ	<i>uvrA</i>	эксцинуклеаза ABC субъединица А	+	0.02	+
Экцизия нуклеотидов + SOS ответ	<i>uvrB</i>	эксцинуклеаза ABC субъединица В	+	0.02	+
Экцизия нуклеотидов + SOS ответ	<i>uvrC</i>	эксцинуклеаза ABC субъединица С	+	0.02	+
Экцизия нуклеотидов + «мисматч»-репарация	<i>uvrD</i>	ДНК-хеликаза II	+	0.04	+
Экцизия оснований	<i>fpg (mutM)</i>	формапидопиримидин-ДНК гликозилаза	+	0.09	+
Экцизия оснований	<i>ung</i>	Урацил-ДНК гликозилаза	+	0.01	+
Экцизия оснований	<i>nfo</i>	эндонуклеаза IV	+	0.08	+
Рекомбинация	<i>recA</i>	рекомбиназа RecA	+	0.01	+
Рекомбинация + SOS ответ	<i>ruvA</i>	АТФ-зависимая хеликаза структуры Холлидея субъединица А	+	0.01	-
Рекомбинация + SOS ответ	<i>ruvB</i>	АТФ-зависимая хеликаза структуры Холлидея субъединица В	+	0.01	-
Рекомбинация	<i>smc</i>	Когезия хромосом	+	нд	+
Рекомбинация + SOS ответ	<i>recR</i>	Рекомбиназа RecR	+	0.02	+
Рекомбинация	<i>MGA_0016</i>	рекомбиназа RecO	+	нд	+
Рекомбинация	<i>recU</i>	Резольваза структуры Холлидея	+	нд	+
Рекомбинация	<i>MGA_0836</i>	Резольваза структуры Холлидея	+	нд	+
SOS ответ	<i>dinB</i>	ДНК-полимераза IV	+	0.01	+
Ген домашнего хозяйства	<i>16S rRNA</i>	16S рРНК	+	600	
Ген домашнего хозяйства	<i>23S rRNA</i>	23S рРНК	+	155	

*нд – анализ не был произведен

Следует отметить, что число копий мРНК варьирует от одной на сто (*recA*, *ruvA*, *ung*, *dinB*) до одной на 2,5 копии геномной ДНК (*hup2*). Если считать, что в каждой клетке в среднем содержится по одной копии геномной ДНК, то представленность транскриптов достаточно низка и в большей части клеток транскрипты генов, кодирующих белки репарации, отсутствуют.

3.10. Представленность участников систем репарации в клетке на белковом уровне

Для того чтобы оценить представленность участников систем репарации на белковом уровне мы использовали методы хромато-масс-спектрометрии для возможно более полной идентификации всех белков в клетках *M. gallisepticum*.

Всего был идентифицирован 561 белок (данные не приведены) при использовании следующих критериев отбора: число уникальных пептидов на белок – не менее двух, global FDR составлял 1% (исходя из анализа в PSPEP, пороговое значение “unusedscore” для белка 0.4). Всего был идентифицирован 17221 уникальный пептид (globalFDR 1% по PSPEP). Необходимо отметить, что алгоритм Paragon, использованный нами для поиска, помимо немодифицированных триптических пептидов позволяет также выявить полутриптические и нетриптические пептиды, а также пептиды, содержащие все возможные модификации аминокислот, учитываемые этим алгоритмом по умолчанию.

Анализ представленности белков из табл.6 показал наличие в протеоме микоплазмы большинства из них – 17 белков для 19 генов, представленных в табл.5. Все белки в табл. 6 были уникальными при группировании спектров с помощью алгоритма ProGroup.

Таблица 6. Идентификация белков – продуктов генов репарации методом LC-MS/MS

Идентификатор в Базе Данных	Название белка	Неиспользованный Score при идентификации*	Score в программе ProteinPilot	Количество уникальных пептидов (с достоверностью $\geq 95\%$)	Процент покрытия последовательности белка уникальными пептидами (с достоверностью $\geq 95\%$)
gi 31541218	SMC (когезин)	84.01	84.01	55	54.43
gi 284811881	UvrA (эксцинуклеаза ABC субъединица А)	75.8	75.8	50	55.67
gi 284811888	MGA_0793 (ДНК-хеликаза, содержит vsg-домен)	62.07	62.07	41	36.41
gi 284812070	UvrD (ДНК-хеликаза II)	51.65	51.65	27	47.54
gi 284811857	UvrB (эксцинуклеаза ABC субъединица В)	41.75	41.75	24	39.01
gi 31541419	Nfo (эндонуклеаза IV)	31.18	31.18	17	63.04
gi 284812280	LigA (ДНК-лигаза)	24.41	24.83	13	23.08
gi 284812220	RecR (рекомбиназа RecR)	24.06	24.06	15	72.82
gi 284811982	MutM (формаидопиримидин-ДНК гликозилаза)	20.02	20.02	10	41.97
gi 31541551	MGA_0195 (белок, содержащий эндонуклеазный домен II типа)	18.9	18.9	18	49.18
gi 284812101	Hup2 (гистоноподобный белок XY)	16.7	18.1	23	71.72
gi 284811981	Exo (5'-3'-эксонуклеаза)	12.01	12.01	8	37.85
gi 284812049	UvrC (эксцинуклеаза ABC субъединица С)	6.19	6.19	4	5.672
gi 31541441	MGA_0016 (рекомбиназа RecO)	5.29	5.29	3	19.5
gi 31541171	putative Holliday junction resolvase	4.9	4.9	5	31.69
gi 284812207	DinB (ДНК-полимераза IV)	4.05	4.05	2	5.985
gi 31541522	RecA (рекомбиназа RecA)	4.01	4.01	2	8.547
gi 31541659	Ung (урацил-ДНК гликозилаза)	3.13	3.86	3	13.85

*Алгоритм идентификации белков предполагает получение значения score для белка как суммы значений score для всех относящихся к нему пептидов (значение score в программе ProteinPilot является прямым производным достоверности идентификации). В том случае, когда какой-либо пептид является общим для двух белков, его вклад в score белка, имеющего меньшую достоверность идентификации (меньший суммарный score) будет меньше максимально возможного значения, рассчитанного на основе достоверности его идентификации. Таким образом, значение неиспользованного score отражает использование тех же самых пептидов (а точнее, спектров, на основе которых были идентифицированы пептиды) при идентификации других белков. Чем ближе значение неиспользованного score к полному score белка (который, в свою очередь, является просто суммой максимально возможных вкладов всех идентифицированных в нем пептидов), тем более специфичной и достоверной является его идентификация, т.к. тем более уникальными и специфичными являются пептиды, его составляющие.

3.11. Ответ генов систем репарации на стрессорные воздействия

Поскольку по литературным данным считается, что система SOS у молликут подверглась редукции [26], мы решили проверить, меняется ли уровень мРНК генов, кодирующих белки репарации в ответ на стрессорные воздействия, чтобы понять, имеет ли место SOS-ответ у *M. gallisepticum*. Для этого мы подвергли клетки критическим воздействиям (см. «Материалы и методы») – температуры, осмотическому и перекисному стрессу, а также действию антибиотиков. Такие воздействия были выбраны потому, что они являются стандартными в исследовании паразитических бактерий, так как при взаимодействии с организмом-хозяином клетки сталкиваются с реакциями воспаления: высокой температурой, перекисной атакой иммунной системы, а также подвергаются действию антибиотиков в процессе терапии. На рис. 13А представлена цветовая карта, отражающая изменения транскрипционного профиля для генов, кодирующих белки репарации ДНК. Исходные данные представлены в Приложении 2.

В случае обработки ципрофлоксацином мы наблюдали повышение уровней мРНК генов *ruvA*, *recA*, *recR*, а также значительный рост (15 раз) уровня мРНК *dinB*. В условиях стресса, вызванного тетрациклином, мы обнаружили индукцию генов рекомбинации – *recA* и *recR*, а также повышение уровня мРНК *gyrA* и *gyrB*, генов, кодирующих ДНК-гиразу. В условиях солевого стресса наблюдали индукцию генов *dinB* и *ung*, а в случае перекисного стресса, практически не наблюдали значимых изменений (рис. 13А).

Среди всех типов стресса отдельно следует выделить стационарную фазу роста культуры, а также тепловой шок в условиях стационарной фазы. В первом случае наблюдается репрессия транскрипции большинства исследованных генов. Во втором случае ответ на тепловой стресс не происходит (рис. 13А).

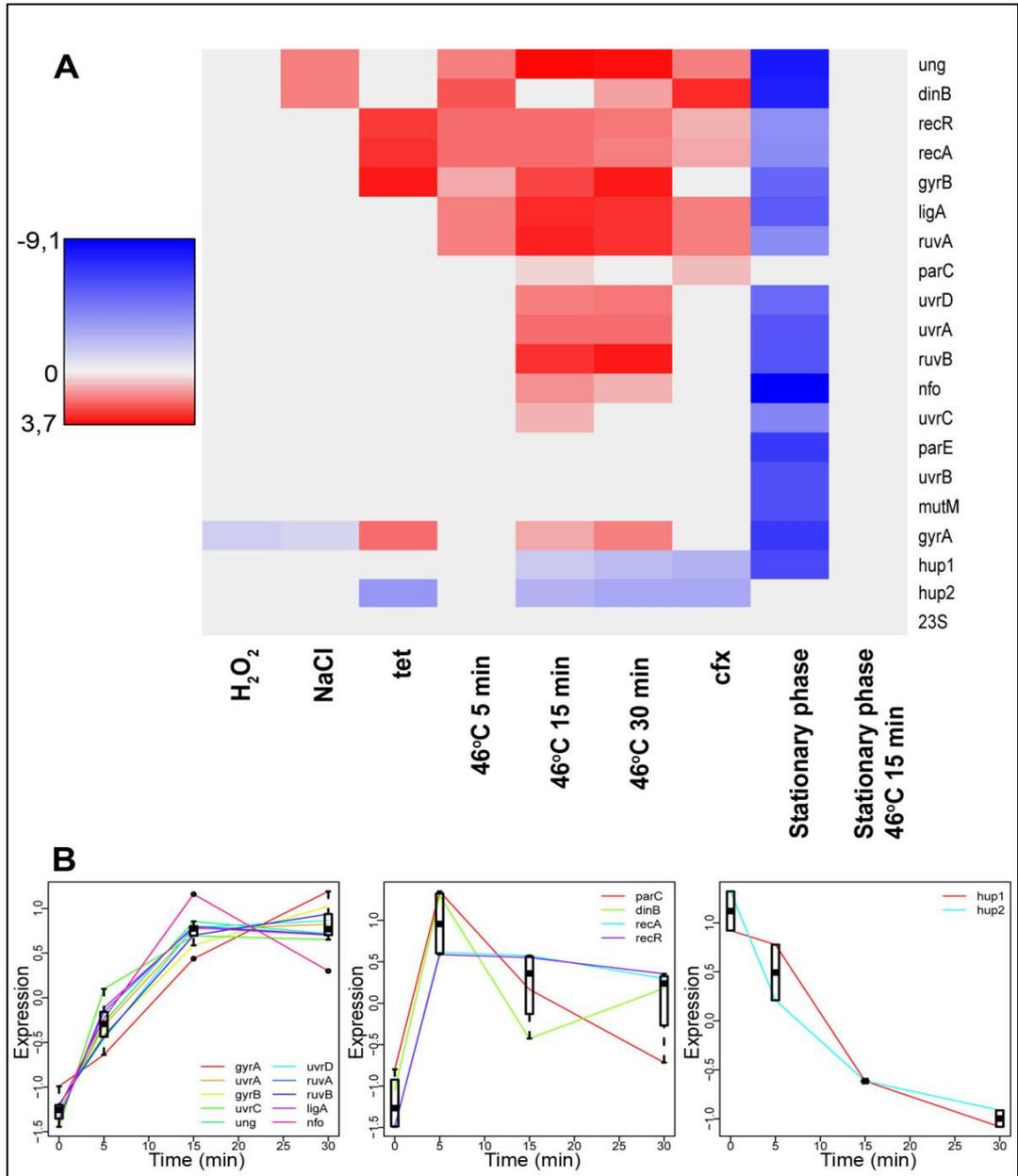


Рис.13. (А) Транскрипционный профиль генов, кодирующих белки репарации, при различных стрессорных воздействиях. Гены расположены по строкам, условия воздействия – по столбцам. Цвета обозначают значимость (см. материалы и методы) и направление изменения уровня мРНК по сравнению с контролем. Серым обозначены гены, не меняющие значимо экспрессию на уровне мРНК при указанном воздействии, красным обозначены гены, для которых наблюдается рост уровня мРНК, а синим – падение. Яркость цвета пропорциональна логарифму (\log_2) отношения уровней мРНК в контроле и стрессе (исходные данные приведены в Приложении 2). В качестве контроля показан ген 23S рРНК.

(В) Динамика изменения транскрипционного профиля при тепловом стрессе. Гены из системы репарации разделены на три паттерна (см. материалы и методы), которые показаны отдельными графиками. Один ген показан одной линией. По оси абсцисс отложено время, по оси ординат – нормированный уровень мРНК.

Исследование динамики изменения транскрипционного профиля генов репарации *M. gallisepticum* при тепловом шоке (рис.13В) позволяет выделить три группы генов. К первой группе относится большинство из исследованных нами генов. Для неё характерно последовательное увеличение уровня транскрипции. Вторая группа включает гены, отвечающие на тепловой стресс немедленно. При этом в дальнейшем уровень мРНК может как снижаться (*parC*, *dinB*), так и выходить на плато (*recA*, *recR*). В третью группу попадают два гена гистоноподобных белков *hup1* и *hup2*. Для них характерно последовательное снижение транскрипции. Быстрая индукция на уровне транскрипции у генов второй группы может свидетельствовать о наличии репрессора, действующего аналогично описанному LexA-репрессору. Последующее падение уровня мРНК генов *parC* и *dinB* может быть обусловлено наличием системы отрицательной обратной связи. Это является вероятным особенно в случае гена *dinB*, кодирующего альтернативную ДНК-полимеразу, поскольку её активация может привести к опасно-высокому уровню мутагенеза. При этом увеличение уровня мРНК большинства генов репарации может быть скорее следствием действия глобального механизма регуляции, чем одного транскрипционного фактора.

Исследование системы репарации микоплазм, бактерий с редуцированными геномами, позволяет судить о минимальном числе генов, необходимых для поддержания геномной стабильности. Сравнительный анализ показывает, что система репарации *M. gallisepticum* представлена меньшим числом участников, чем мы наблюдаем в случае *E. coli*. Однако если от численных характеристик перейти к функциональным, то можно заметить присутствие ключевых звеньев всех основных репаративных систем в отсутствие дублирующих друг друга звеньев.

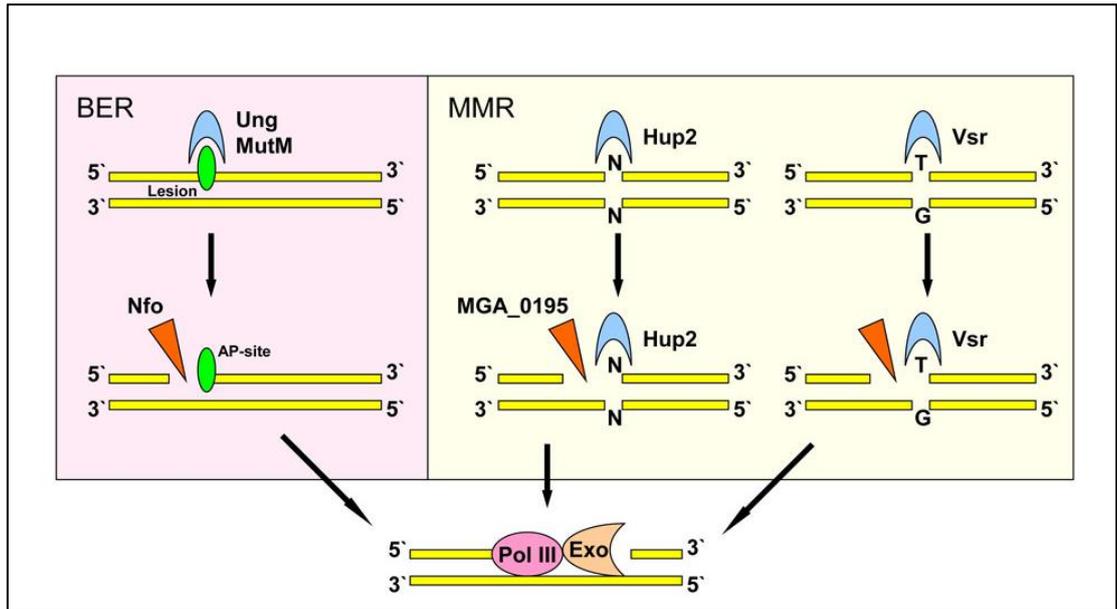


Рис.14. Возможная модель работы MMR и BER систем репарации ДНК у *M. gallisepticum*. BER: ДНК-гликозилаза (MutM или Ung) распознает и удаляет поврежденное основание, после чего AP-эндонуклеаза вносит одноцепочечный разрыв с 5'-конца от поврежденного сайта. MMR: белок Hup2 или Vsr (в случае T-G мисматча) узнает неправильное спаривание, затем эндонуклеаза MGA_0125 вносит одноцепочечный разрыв в дочернюю цепь ДНК. На заключительном этапе двух типов репарации экзонуклеаза (Exo) осуществляет гидролиз цепи ДНК (в направлении 5'-3'), содержащей разрыв, который застраивается ДНК-полимеразой III (DnaE).

Результаты нашей работы показывают, что возможный репертуар репаративных систем *M. gallisepticum* может быть представлен большим числом белков, чем считалось до последнего времени. В частности, мы обнаружили несколько ранее неизвестных для микоплазм белков. Один из них – белок MGA_0195, аннотированный как гипотетический белок с неизвестной функцией, по результатам выравнивания имеет гомологию с белком Vsr, вовлеченным в репарацию «мисматчей» ДНК, содержащих неспаренный гуанин у *E. coli* [151]. Интересно, что идентифицированный нами в *M. gallisepticum* белок mgHU (hup2), способный связывать ДНК-«мисматчи», не связывает неспаренные Т-Г, А-Г и Г-Г (рис. 14).

Поскольку метилирование GATC сайтов отсутствует, дискриминация цепей может осуществляться за счет взаимодействия репаративного комплекса с В-субъединицей ДНК-полимеразы III при репликации – такой

механизм показан для ряда микроорганизмов (Грам-положительных бактерий) [10]. Таким образом, можно по крайней мере утверждать, что *M. gallisepticum* имеет фермент, способный узнавать и связывать ДНК-«мисматчи», а также ферменты, гипотетически способные осуществлять удаление поврежденного участка ДНК. Однако остается открытым вопрос, какой из ферментов микоплазм является функциональным аналогом белка mutH, необходимого для внесения одноцепочечного разрыва в репарируемую цепь ДНК.

Мы обнаружили, что *M. gallisepticum* обладает полными путями эксцизионной репарации нуклеотидов и рекомбинации ДНК. Особый интерес для анализа генов представляет система эксцизионной репарации поврежденных оснований ДНК, поскольку она всегда представлена большим числом белков, каждый из которых узнает свой тип повреждения. Известно три типа повреждений оснований: окисление, алкилирование и дезаминирование. ДНК-гликозилаза MutM, присутствующая в геноме *M. gallisepticum*, исправляет одно из самых часто возникающих ДНК-повреждений, вызванных эндогенным окислительным стрессом [152]. Урацил-гликозилаза устраняет из ДНК урацил, который возникает при спонтанном дезамидировании цитозина или при его ошибочном включении в ДНК при репликации. Из двух AP-эндонуклеаз, III и IV, *M. gallisepticum* обладает лишь одной, эндонуклеазой IV. Интересно, что именно эндонуклеаза IV (Nfo) обладает дополнительной активностью – она распознает окисленные основания (гидроксицитозин, дигидроксиурацил и дигидрокситимидин) и вносит одноцепочечный разрыв с 5'-стороны от повреждения, который служит затравкой для репарации посредством полимеразы и лигазы [153].

Полагая, что молликуты (микоплазмы) – это бактерии с минимальным геномом, способные к самостоятельному размножению, мы приходим к

выводу, что имеющийся у них состав репаративных белков является необходимым и достаточным для жизнедеятельности на бесклеточной среде.

Результаты транскрипционного анализа исследуемых в работе генов указывают на то, что молекулы мРНК представлены далеко не в каждой клетке популяции (табл. 5). Эти данные согласуются с литературными и могут быть связаны с большим временем жизни функционального белка в отличие от короткоживущих мРНК [154].

На протеомном уровне (табл. 5, 6) нам удалось идентифицировать большую часть (80%) белков-участников систем репарации, в том числе участников системы SOS-ответа – DinB и RecA, одних из самых низкопредставленных белков.

Интересным результатом транскрипционного профилирования оказалась индукция генов SOS-ответа при различных шоковых воздействиях – в тепловом шоке, а также под действием ципрофлоксацина и тетрациклина (рис.13А). Этот факт заинтересовал нас по нескольким причинам. В первую очередь, подобной реакции не было показано ранее для Mollicutes, в том числе в системных работах по анализу транскрипционных ответов у *M. pneumoniae* [155], одного из ближайших родственников *M. gallisepticum* по данным филогенетических исследований [156]. Во-вторых, в геноме всех представителей класса Mollicutes отсутствует какой бы то ни было известный регулятор SOS-ответа, в связи с чем ряд авторов считал SOS-систему нефункциональной у микоплазм [26,27]. Однако, полученные нами результаты согласуются с литературными данными, полученными по транскрипционному анализу бактерий, не являющихся родственниками микоплазм и имеющим описанные регуляторные механизмы для генов SOS [10,89,157,158]. Данные наблюдения могут указывать на функциональность системы SOS-ответа с одной стороны, а также на присутствие какого-то неизвестного регулятора с другой. В пользу гипотезы о наличии такого регулятора может также свидетельствовать наличие в геноме ряда генов,

белковые продукты которых, согласно аннотации, обладают сиквенс-специфическими ДНК-связывающими доменами и потенциально могут играть роль транскрипционных факторов (данные не представлены). Продемонстрированная здесь индукция ДНК-полимеразы IV типа в различных типах стресса (рис.13А) согласуется с литературными данными и может быть механизмом приспособления к стрессам через повышение эндогенного уровня мутагенеза [159–162].

3.12. Протеомное профилирование в условиях теплового шока

Поскольку транскрипционный ответ наибольшего числа генов, участвующих в системе репарации *M. gallisepticum*, наблюдался в условиях теплового шока (в логарифмической фазе роста культуры), мы провели протеомное профилирование в данных условиях методом двумерного гелевого электрофореза с дифференциальной окраской цианинами с целью проверить, каким образом изменяется белковый состав клеток. Кроме того, мы также хотели понять причины наличия SOS-подобного ответа при тепловом стрессе.

На рис. 15 и в табл. 7 представлены результаты протеомного профилирования. Среди белков, уровень которых возрастает, в условиях теплового стресса присутствуют известные белки теплового шока – шаперон ClpB и ко-шаперонин GroES. Возрастание представленности данных белков согласуется с ранее проведенными исследованиями ответа на тепловой шок многих бактерий, в том числе описано для представителей рода *Mycoplasma* [163].

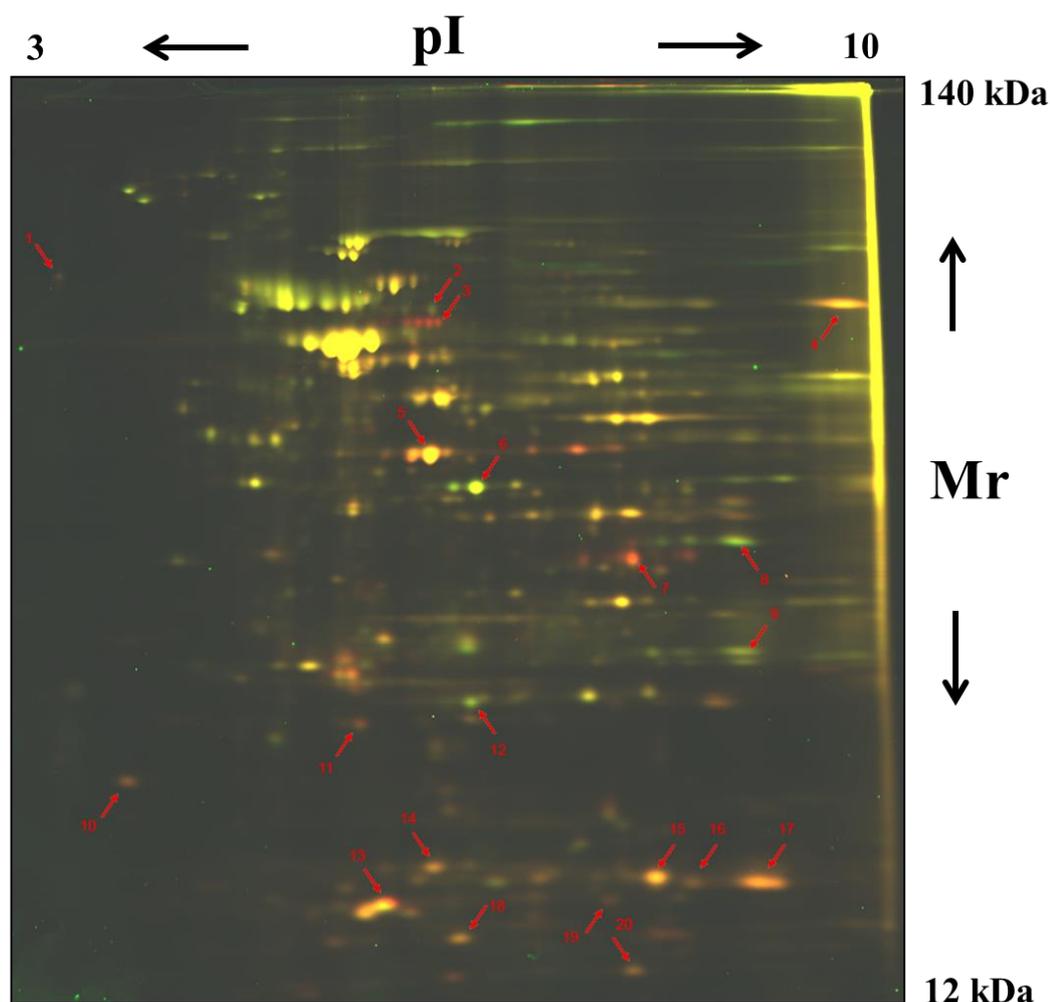


Рис. 15. Протеомное профилирование *M. gallisepticum* (логарифмическая фаза роста культуры) в условиях теплового стресса методом двумерного гель-электрофореза с дифференциальной окраской цианинами. Разделение в первом направлении проводилось по изоэлектрической точке (pI), во втором по молекулярной массе белка (Mr). Стрелками указаны белки, представленность которых меняется. Красный цвет – представленность при тепловом стрессе растет, зеленый – падает. Цифры показывают отдельные белки, список белков представлен в табл.7.

К другой группе белков, представленность которых растет, относятся белки, участвующие в трансляции: факторы EF-Tu и EF-Ts, а также транскрипции: δ - субъединица РНК полимеразы RpoE. Следует отметить, что для белка RpoE показано участие в глобальной регуляции ответа на стресс у *Streptococcus mutants* [164].

Анализируя данные протеомного профилирования, мы заинтересовались возрастанием представленности белка НАДН-оксидазы. Функция данного белка заключается в окислении НАДН до НАД⁺ при

участии кислорода. Акцептором электрона в данной реакции является молекула воды, в результате чего образуется перекись водорода, которая потенциально может служить источником активных форм кислорода (далее АФК), основных эндогенных мутагенов. Из литературных данных известно, что в условиях теплового шока, в клетках происходит повышение потребности в АТФ, который активно расходуется шаперонами для восстановления правильного фолдинга клеточных белков.

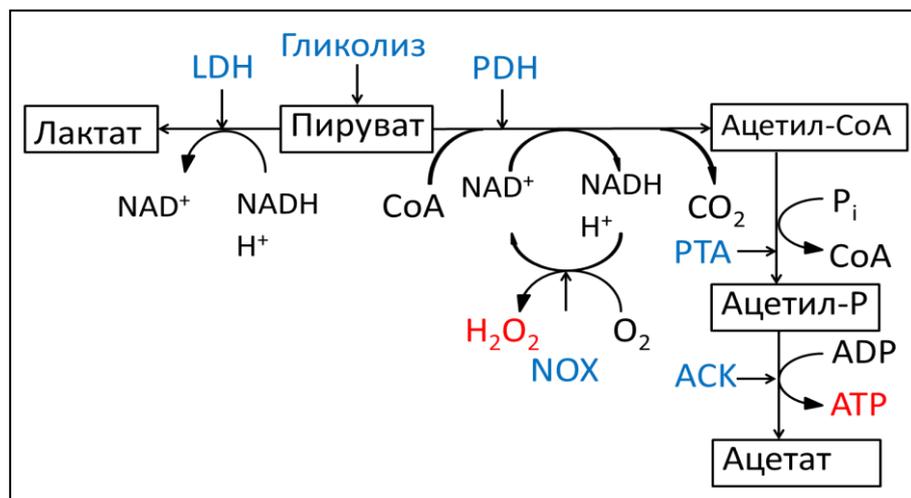


Рис. 16. Схема метаболизма пирувата у *M. gallisepticum*. Реконструирована на основании геномных данных. Обозначения: NOX – НАДН-оксидаза, LDH – лактат-дегидрогеназа, PDH – пируват-дегидрогеназа, PTA – фосфотрансацетилаза, АСК – ацетат-киназа.

Поскольку единственным способом генерирования АТФ для *M. gallisepticum* является гликолитическое расщепление глюкозы, а также из ранних исследований гликолиза известно, что лимитирующей его стадией является регенерация НАД^+ , мы предположили, что возрастание уровня внутриклеточной НАДН-оксидазы необходимо клеткам для ускорения гликолиза и синтеза АТФ. Кроме того, утилизация пирувата по пути его превращения в ацетат вместо лактата дает дополнительную молекулу АТФ в расчете на одну молекулу пирувата (рис. 16).

Таблица 7. Изменение представленности белков *M. gallisepticum* в условиях теплового стресса (номера в левом столбце соответствуют номерам белков на рис. 15)

№	Локус в геноме	Белок (функция)	Изменение уровня
1	MGA_0131	RotD спермидин/путресцинABC транспортер	рост
2	MGA_0495	Гипотетический белок (функция неизвестна)	падение
3	MGA_0167	НАДН-оксидаза	рост
4	MGA_0178	ClpB шаперон	рост
5	MGA_1033	Фактор трансляции EF-Tu	рост
6	MGA_0165	Пируват-дегидрогеназа E1 α -субъединица	падение
7	MGA_0782	Фактор трансляции EF-Ts	рост
8	MGA_0164	Пируват-дегидрогеназа E1 β - субъединица	падение
9	MGA_0357	Триозо-фосфат изомеразы	падение
10	MGA_0497	RpoE (δ - субъединица РНК полимеразы)	рост
11	MGA_1017	HatA ABC транспортер	рост
12	MGA_0091	Фосфолипид-связывающий белок	падение
13	MGA_0865	Фактор созревания рибосом	рост
14	MGA_0870	Гипотетический белок (функция неизвестна)	рост
15	MGA_0579	Гипотетический белок (функция неизвестна)	рост
16	MGA_1125	SufC	рост
17	MGA_1142	OsmC (пероксиредоксин)	рост
18	MGA_0452	Тиоредоксин (антиоксидант)	рост
19	MGA_0774	Тиоредоксин (антиоксидант)	рост
20	MGA_0153	Ко-шаперонин GroES	рост

3.13. Тепловой стресс «разгоняет» клеточный метаболизм приводя к повышенному уровню внутриклеточной АТФ, окислительному стрессу, сопровождающемуся повреждениями ДНК

Для того, чтобы проверить, возрастает ли уровень внутриклеточного АТФ, мы провели его измерение с помощью биолуминесцентной системы на основе люцеферин/люцеферазной реакции.

Полученные результаты приведены на рис. 17. В условиях инкубации клеток *M. gallisepticum* мы наблюдали семикратное увеличение уровня АТФ в процессе теплового стресса.

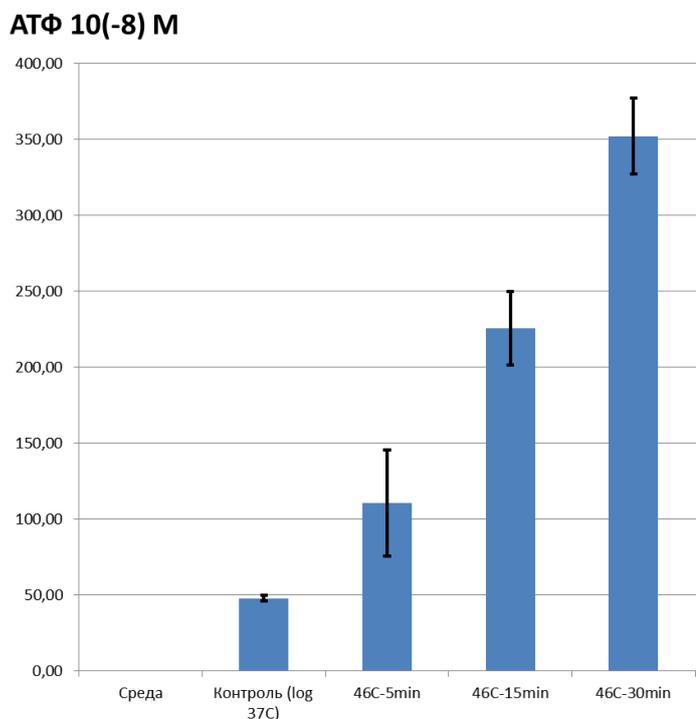


Рис. 17. Изменение уровня АТФ при тепловом стрессе. В качестве «нулевого» уровня была использована стерильная среда для культивирования *M. gallisepticum*. Образцы подвергнутые тепловому стрессу (5, 15 и 30 мин при 46⁰С) демонстрируют повышение уровня внутриклеточного АТФ по сравнению с контролем (клетки в логарифмической фазе роста). Эксперимент был проведен в виде трех независимых биологических повторах с тремя техническими повторами в каждом.

Мы также провели измерение скорости генерации перекисных форм кислорода внутри клеток *M. gallisepticum* в условиях теплового стресса результаты приведены на рис. 18.

Как видно из рис. 18, в клетках *M. gallisepticum* происходит значительное возрастание скорости генерации АФК. Поскольку АФК являются основными эндогенными источниками повреждений ДНК, было интересно детектировать наличие таковых.

В условиях теплового стресса для *M. gallisepticum* мы провели метаболомное профилирование (неопубликованные данные, полученные

совместно с Ванюшкиной А.). В результате анализа метаболома мы детектировали высокий уровень 8-оксогуанина в клетках *M. gallisepticum*, подвергнутых тепловому стрессу (данные не приведены). Присутствие 8-оксогуанина указывает на то, что клетки находятся в условиях окислительного стресса, что также подтверждается данными по измерению содержания перекиси в клетках, подвергнутых тепловому шоку (рис. 18).

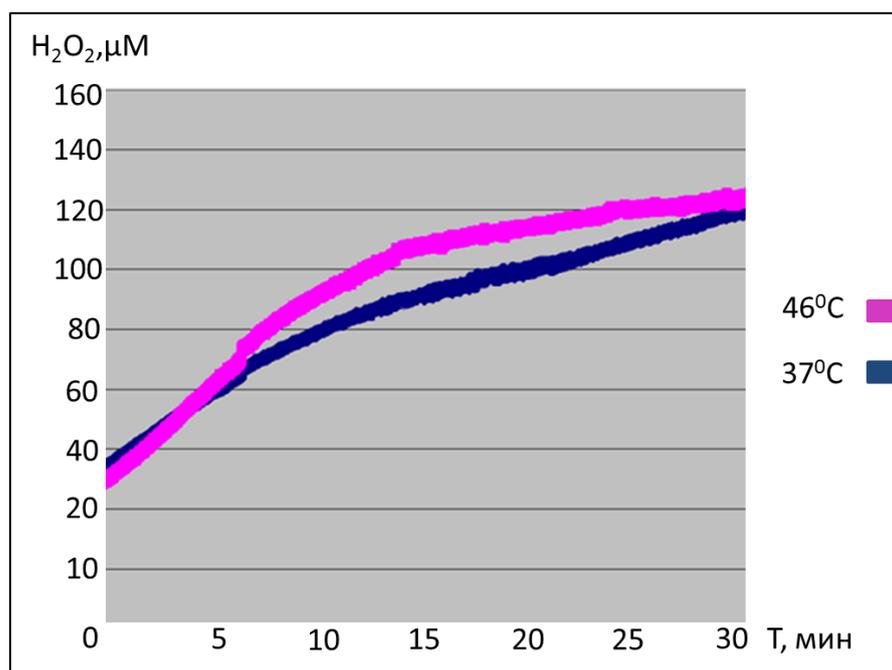


Рис. 18. Изменение скорости образования перекиси в клетках *M. gallisepticum* в логарифмической фазе роста (37°C) и в условиях теплового стресса (46°C). По оси абсцисс указано время в мин, по оси ординат – условные единицы флуоресценции, отражающие накопление H₂O₂.

Повышение уровня клеточной АТФ вместе с возрастанием скорости образования перекиси, а также появление 8-оксогуанина дает возможность сделать два предположения: (i) клетки *M. gallisepticum* находятся в условиях окислительного стресса, (ii) В условиях окислительного стресса может возрасть количество повреждений клеточной ДНК.

Имея косвенные указания на то что, в результате теплового стресса создаются предпосылки для повреждения клеточной ДНК (8-оксогуанин, АФК, индукция мутаторной ДНК-полимеразы *dinB*) мы решили определить,

имеет ли место возрастание уровня мутагенеза в данных условиях. В качестве критерия, отражающего уровень мутагенеза, мы использовали количество колоний *M. gallisepticum*, устойчивых к антибиотику фторхинолонового ряда – ципрофлоксацину. Данный антибиотик был выбран по причине его высокой активности в отношении микоплазм [165]. Кроме того, устойчивость к нему определяется наличием мутаций в генах, кодирующих ДНК-гиразу – *gyrA* и *gyrB*, а также ДНК-топоизомеразу – *parE* и *parC*. По указанным причинам мы предполагали, что изменение числа выросших колоний в среде, содержащей летальную дозу антибиотика, будет отражать изменение в уровне мутагенеза в клетках *M. gallisepticum*. Однако, в многочисленных экспериментах нами не было детектировано сколько-нибудь значимого возрастания уровня мутагенеза (данные не приведены).

Таким образом, можно заключить, что в условиях теплового стресса клетки *M. gallisepticum* для повышения уровня АТФ значительно увеличивают скорость окисления НАДН, что приводит в качестве побочной реакции к усилению образования эндогенных АФК. Такое значительное образование АФК вызывает повреждения ДНК, в основном в виде окисленного гуанина (8-оксогуанин). В подобных условиях становится достаточно логичным активация SOS-ответа в клетках *M. gallisepticum*. При этом активация систем репарации защищает клетки *M. gallisepticum* от гибели, не приводя к значимому возрастанию уровня мутагенеза, что является доказательством функциональной активности системы репарации ДНК у исследуемого микроорганизма.

Заключение

Организмы, принадлежащие к классу Mollicutes, для представителей которого характерно отсутствие клеточной стенки, являются наименьшими по размеру известными организмами, способными к самостоятельному делению на бесклеточных питательных средах. Один из представителей данного класса – *M. gallisepticum* характеризуется, помимо прочего, отсутствием ключевых элементов системы мисматч-репарации ДНК, необходимой для распознавания и коррекции ошибок репликации, например, таких как некорректное спаривание [31].

Проведенный в нашем исследовании скрининг клеточного экстракта *M. gallisepticum* выявил белок, который способен связывать одонуклеотидные мисматчи в дцДНК-фрагментах. Данный белковый фактор был идентифицирован как гомолог HU-белка *E. coli*.

Во второй части нашего исследования мы стремились наиболее полно охарактеризовать систему репарации *M. gallisepticum* в целом. Основываясь на геномных и литературных данных, мы составили список участников различных путей репарации ДНК. Экспрессия каждого гена из полученного *минимального* набора была проанализирована на уровне транскрипции по представленности их мРНК (количество копий на клетку) в нормальных условиях и под влиянием различных стрессоров. При этом было показано, что даже в отсутствие известных для других организмов регуляторов у *M. gallisepticum* происходит индукция SOS-ответа на транскрипционном уровне в ответ на стрессорные воздействия.

Для подтверждения того, что транскрипция исследуемых генов действительно приводит к синтезу белкового продукта, способного устранять возможные повреждения ДНК, мы провели исчерпывающий протеомный анализ и идентифицировали большую часть участников систем репарации ДНК на белковом уровне.

В третьей части нашего исследования мы продемонстрировали повышение уровня АФК внутри клеток *M. gallisepticum*, приводящее к повреждениям ДНК с образованием 8-оксогуанина. При этом система репарации успевает устранить все повреждения, поскольку значимого возрастания уровня мутагенеза не происходит.

Основные выводы исследования

1. Идентифицирован и охарактеризован белокHU (Hup2) из *M. gallisepticum*, способный специфически связывать ДНК, содержащую неправильно спаренные нуклеотиды.
2. На основе биоинформатического анализа была произведена *in silico* реконструкция системы репарации ДНК, предложена модель репарации ДНК, содержащей ошибочно-спаренные нуклеотиды. Найдены новые, ранее не аннотированные, потенциальные участники системы репарации: белки MGA_0793, MGA_0195, MGA_0836, MGA_0016.
3. Показана экспрессия генов, кодирующих участников репарации ДНК, на транскрипционном и протеомном уровне. Для всех исследуемых генов определена количественная представленность транскриптов на один клеточный геном.
4. Показано, что при тепловом стрессе у *M. gallisepticum* происходит возрастание уровня внутриклеточной АТФ, повышение скорости образования внутриклеточной перекиси, а также наблюдается SOS-ответ.

Список литературы

1. Razin S., Yogev D. Molecular Biology and Pathogenicity of Mycoplasmas // Society. 1998. Vol. 62. № 4. P. 1094–1156.
2. Peebles, E.D., Branton S.L. Mycoplasma gallisepticum in the commercial egg-laying hen: an historical perspective considering effects of pathogen strain, age of bird at inoculation, and diet on performance and physiology // J. Appl. Poult. Res. 2012. Vol. 212. P. 897–914.
3. Hennigan S.L. et al. Detection and differentiation of avian mycoplasmas by surface-enhanced Raman spectroscopy based on a silver nanorod array. // Appl. Environ. Microbiol. 2012. Vol. 78. № 6. P. 1930–1935.
4. Thomas NJ, Hunter DB A.C. Infectious Diseases of Wild Birds. Wiley-Blackwell. 2007. P. 496.
5. Thompson C.C. et al. Towards a genome based taxonomy of Mycoplasmas // Infect. Genet. Evol. 2011. Vol. 11. № 7. P. 1798–1804.
6. Gruson D. et al. In Vitro Development of Resistance to Six and Four Fluoroquinolones in Mycoplasma pneumoniae and Mycoplasma hominis , Respectively. 2005. Vol. 49. № 3. P. 1190–1193.
7. Curti E. et al. DNA polymerase switching: effects on spontaneous mutagenesis in Escherichia coli. // Mol. Microbiol. 2009. Vol. 71. № 2. P. 315–331.
8. Delaney N.F. et al. Ultrafast evolution and loss of CRISPRs following a host shift in a novel wildlife pathogen, Mycoplasma gallisepticum. // PLoS Genet. 2012. Vol. 8. № 2. P. e1002511.
9. Hawley D.M. et al. Parallel patterns of increased virulence in a recently emerged wildlife pathogen. // PLoS Biol. 2013. Vol. 11. № 5. P. e1001570.
10. Lenhart J.S. et al. DNA repair and genome maintenance in Bacillus subtilis. // Microbiol. Mol. Biol. Rev. 2012. Vol. 76. № 3. P. 530–564.
11. Zinser E.R., Kolter R. Mutations Enhancing Amino Acid Catabolism Confer a Growth Advantage in Stationary Phase // J. Bacteriol. 1999. Vol. 181. № 18. P. 5800–5807.
12. Ciccarelli F.D. et al. Toward automatic reconstruction of a highly resolved tree of life. // Science. 2006. Vol. 311. № 5765. P. 1283–1287.

13. Lenhart J.S. et al. DNA repair and genome maintenance in *Bacillus subtilis*. // *Microbiol. Mol. Biol. Rev.* 2012. Vol. 76. № 3. P. 530–564.
14. Michel B. After 30 years of study, the bacterial SOS response still surprises us. // *PLoS Biol.* 2005. Vol. 3. № 7. P. e255.
15. Lovett C.M. et al. SOS-like induction in *Bacillus subtilis*: induction of the RecA protein analog and a damage-inducible operon by DNA damage in Rec⁺ and DNA repair-deficient strains. // *J. Bacteriol.* 1988. Vol. 170. № 4. P. 1467–1474.
16. Simmons L.A. et al. Comparison of Responses to Double-Strand Breaks between *Escherichia coli* and *Bacillus subtilis* Reveals Different Requirements for SOS Induction □ // *Society*. 2009. Vol. 191. № 4. P. 1152–1161.
17. Lovett C.M.J., Cho K.C., O’Gara T.M. Purification of an SOS repressor from *Bacillus subtilis*. // *J. Bacteriol.* 1993. Vol. 175. № 21. P. 6842–6849.
18. Goranov A.I. et al. Characterization of the global transcriptional responses to different types of DNA damage and disruption of replication in *Bacillus subtilis*. // *J. Bacteriol.* 2006. Vol. 188. № 15. P. 5595–5605.
19. Witkin E.M. et al. Recovery from ultraviolet light-induced inhibition of DNA synthesis requires umuDC gene products in recA718 mutant strains but not in recA⁺ strains of *Escherichia coli*. // *Proc. Natl. Acad. Sci.* 1987. Vol. 84. № 19. P. 6805–6809.
20. Bernard R., Marquis K.A., Rudner D.Z. Nucleoid occlusion prevents cell division during replication fork arrest in *Bacillus subtilis*. // *Mol. Microbiol.* 2010. Vol. 78. № 4. P. 866–882.
21. Bisognano C. et al. A recA-LexA-dependent pathway mediates ciprofloxacin-induced fibronectin binding in *Staphylococcus aureus*. // *J. Biol. Chem.* 2004. Vol. 279. № 10. P. 9064–9071.
22. Mesak L.R., Miao V., Davies J. Effects of subinhibitory concentrations of antibiotics on SOS and DNA repair gene expression in *Staphylococcus aureus*. // *Antimicrob. Agents Chemother.* 2008. Vol. 52. № 9. P. 3394–3397.
23. Cirz R.T. et al. Complete and SOS-mediated response of *Staphylococcus aureus* to the antibiotic ciprofloxacin. // *J. Bacteriol.* 2007. Vol. 189. № 2. P. 531–539.

24. Van der Veen S. et al. The SOS response of *Listeria monocytogenes* is involved in stress resistance and mutagenesis. // *Microbiology*. 2010. Vol. 156. № Pt 2. P. 374–384.
25. Jochmann N. et al. Genetic makeup of the *Corynebacterium glutamicum* LexA regulon deduced from comparative transcriptomics and in vitro DNA band shift assays. // *Microbiology*. 2009. Vol. 155. № Pt 5. P. 1459–1477.
26. Carvalho M. et al. DNA repair in reduced genome : The *Mycoplasma* model // *DNA Repair (Amst)*. 2005. Vol. 360. P. 111 – 119.
27. Razin S., Yogevev D., Naot Y. Molecular biology and pathogenicity of mycoplasmas // *Microbiol. Mol. Biol. Rev.* 1998. Vol. 62. № 4. P. 1094–1156.
28. Dubnau D. DNA uptake in bacteria. // *Annu. Rev. Microbiol. Annual Reviews* 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA, 1999. Vol. 53. P. 217–244.
29. Carvalho F.M. et al. DNA repair in reduced genome: the *Mycoplasma* model. // *Gene*. 2005. Vol. 360. № 2. P. 111–119.
30. Lage C. et al. New insights on how nucleotide excision repair could remove DNA adducts induced by chemotherapeutic agents and psoralens plus UV-A (PUVA) in *Escherichia coli* cells // *Mutat. Res. Mutat. Res.* 2003. Vol. 544. № 2-3. P. 143–157.
31. Au N. et al. Genetic Composition of the *Bacillus subtilis* SOS System // *J. Bacteriol.* 2005. Vol. 187. № 22. P. 7655–7666.
32. Friedman B.M., Yasbin R.E. The genetics and specificity of the constitutive excision repair system of *Bacillus subtilis*. // *Mol. Gen. Genet.* 1983. Vol. 190. № 3. P. 481–486.
33. Sancar A. DNA excision repair. // *Annu. Rev. Biochem. Annual Reviews* 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA, 1996. Vol. 65. P. 43–81.
34. Orren D. et al. Post-incision steps of nucleotide excision repair in *Escherichia coli*. Disassembly of the UvrBC-DNA complex by helicase II and DNA polymerase I // *J. Biol. Chem.* 1992. Vol. 267, № 2. P. 780–788.

35. Orren D.K., Sancar A. The (A)BC excinuclease of *Escherichia coli* has only the UvrB and UvrC subunits in the incision complex. // *Proc. Natl. Acad. Sci.* 1989. Vol. 86, № 14. P. 5237–5241.
36. Smith B.T., Grossman A.D., Walker G.C. Localization of UvrA and Effect of DNA Damage on the Chromosome of *Bacillus subtilis* // *J. Bacteriol.* 2002. Vol. 184, № 2. P. 488–493.
37. Dalhus B. et al. DNA base repair--recognition and initiation of catalysis. // *FEMS Microbiol. Rev.* 2009. Vol. 33, № 6. P. 1044–1078.
38. Bjelland S., Seeberg E. Mutagenicity, toxicity and repair of DNA base damage induced by oxidation // *Mutat. Res. Mol. Mech. Mutagen.* 2003. Vol. 531, № 1. P. 37–80.
39. Baute J., Depicker A. Base Excision Repair and its Role in Maintaining Genome Stability. Informa UK Ltd London, UK, 2008.
40. H. Kasai, M. H. Chung, D. S. Jones, H. Inoue, H. Ishikawa, H. Kamiya, E. Ohtsuka S.N. 8-Hydroxyguanine, a DNA adduct formed by oxygen radicals: its implication on oxygen radical-involved mutagenesis/carcinogenesis. [Online] // *J. Toxicol. Sci.* 1991. P. 95–105. URL: https://www.jstage.jst.go.jp/article/jts1976/16/SupplementI/16_SupplementI_95/_article.
41. Tchou J. et al. 8-oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity. // *Proc. Natl. Acad. Sci.* 1991. Vol. 88, № 11. P. 4690–4694.
42. Hsu G.W. et al. Error-prone replication of oxidatively damaged DNA by a high-fidelity DNA polymerase. // *Nature.* 2004. Vol. 431, № 7005. P. 217–221.
43. Shibutani S., Takeshita M., Grollman A.P. Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. // *Nature.* 1991. Vol. 349, № 6308. P. 431–434.
44. Michaels M.L. et al. Evidence that MutY and MutM combine to prevent mutations by an oxidatively damaged form of guanine in DNA. // *Proc. Natl. Acad. Sci. U. S. A.* 1992. Vol. 89, № 15. P. 7022–7025.
45. Michaels M.L. et al. MutM, a protein that prevents G C→T A transversions, is formamidopyrimidine-DNA glycosylase // *Nucleic Acids Res.* 1991. Vol. 19, № 13. P. 3629–3632.

46. Michaels M.L. et al. A repair system for 8-Oxo-7,8-dihydrodeoxyguanine // *Biochemistry*. American Chemical Society, 1992. Vol. 31, № 45. P. 10964–10968.
47. Qi Y. et al. Encounter and extrusion of an intrahelical lesion by a DNA repair enzyme. // *Nature*. Macmillan Publishers Limited. All rights reserved, 2009. Vol. 462, № 7274. P. 762–766.
48. Qi Y. et al. Entrapment and structure of an extrahelical guanine attempting to enter the active site of a bacterial DNA glycosylase, MutM. // *J. Biol. Chem.* 2010. Vol. 285, № 2. P. 1468–1478.
49. Sasaki M., Kurusu Y. Analysis of spontaneous base substitutions generated in mutator strains of *Bacillus subtilis* // *FEMS Microbiol. Lett.* 2004. Vol. 234, № 1. P. 37–42.
50. Maki H., Sekiguchi M. MutT protein specifically hydrolyses a potent mutagenic substrate for DNA synthesis. // *Nature*. 1992. Vol. 355, № 6357. P. 273–275.
51. R M Schaaper B.I.B. A.T----C.G transversions and their prevention by the *Escherichia coli* mutT and mutHLS pathways. // *Mol. Gen. Genet.* 1989. Vol. 219, № 1-2. P. 256 – 62.
52. Cox E.C., Yanofsky C. Mutator Gene Studies in *Escherichia coli* // *J. Bacteriol.* 1969. Vol. 100, № 1. P. 390–397.
53. Castellanos-Juárez F.X. et al. YtkD and MutT protect vegetative cells but not spores of *Bacillus subtilis* from oxidative stress. // *J. Bacteriol.* 2006. Vol. 188, № 6. P. 2285–2289.
54. Ramirez M.I. et al. The ytkD (mutTA) Gene of *Bacillus subtilis* Encodes a Functional Antimutator 8-Oxo-(dGTP/GTP)ase and Is under Dual Control of Sigma A and Sigma F RNA Polymerases // *J. Bacteriol.* 2004. Vol. 186, № 4. P. 1050–1059.
55. Vidales L.E. et al. Defects in the error prevention oxidized guanine system potentiate stationary-phase mutagenesis in *Bacillus subtilis*. // *J. Bacteriol.* 2009. Vol. 191, № 2. P. 506–513.
56. Wang D., Kreutzer D.A., Essigmann J.M. Mutagenicity and repair of oxidative DNA damage: insights from studies using defined lesions // *Mutat. Res. Mol. Mech. Mutagen.* 1998. Vol. 400, № 1. P. 99–115.

57. Kunkel T.A., Erie D.A. DNA mismatch repair. // *Annu. Rev. Biochem. Annual Reviews*, 2005. Vol. 74. P. 681–710.
58. Schofield M.J., Hsieh P. DNA mismatch repair: molecular mechanisms and biological function. // *Annu. Rev. Microbiol. Annual Reviews* 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA, 2003. Vol. 57. P. 579–608.
59. Yang H. et al. The role of *Bacillus anthracis* RecD2 helicase in DNA mismatch repair // *DNA Repair (Amst)*. 2011. Vol. 10, № 11. P. 1121–1130.
60. Ginetti F. et al. *Bacillus subtilis* mutS mutL operon: identification, nucleotide sequence and mutagenesis // *Microbiology*. 1996. Vol. 142, № 8. P. 2021–2029.
61. Dreiseikelmann B., Wackernagel W. Absence in *Bacillus subtilis* and *Staphylococcus aureus* of the sequence-specific deoxyribonucleic acid methylation that is conferred in *Escherichia coli* K-12 by the dam and dcm enzymes. // *J. Bacteriol.* 1981. Vol. 147, № 1. P. 259–261.
62. Eisen J. A phylogenomic study of the MutS family of proteins // *Nucleic Acids Res.* 1998. Vol. 26, № 18. P. 4291–4300.
63. Fukui K. DNA mismatch repair in eukaryotes and bacteria. // *J. Nucleic Acids*. 2010. Vol. 2010, № c.
64. Woese C. Bacterial evolution. // *Microbiol. Rev.* 1987. Vol. 51, № 2.
65. Iyer R.R. et al. DNA mismatch repair: functions and mechanisms. // *Chem. Rev.* 2006. Vol. 106, № 2. P. 302–323.
66. Gorbachev A.Y. et al. DNA repair in *Mycoplasma gallisepticum*. // *BMC Genomics*. 2013. Vol. 14, № 1. P. 726.
67. Berdis A., Sutton M.D. Coordinating DNA polymerase traffic during high and low fidelity synthesis // *Biochim. Biophys. Acta - Proteins Proteomics*. 2010. Vol. 1804, № 5. P. 1167–1179.
68. Sutton M.D., Walker G.C. Managing DNA polymerases: coordinating DNA replication, DNA repair, and DNA recombination. // *Proc. Natl. Acad. Sci. U. S. A.* 2001. Vol. 98, № 15. P. 8342–8349.
69. Sung H.-M. et al. Roles of YqjH and YqjW, Homologs of the *Escherichiacoli* UmuC/DinB or Y Superfamily of DNA Polymerases, in Stationary-Phase

- Mutagenesis and UV-Induced Mutagenesis of *Bacillus subtilis* // *J. Bacteriol.* 2003. Vol. 185, № 7. P. 2153–2160.
70. Ollivierre J.N., Fang J., Beuning P.J. The Roles of UmuD in Regulating Mutagenesis. // *J. Nucleic Acids.* 2010. Vol. 2010.
 71. Reuven N.B. The Mutagenesis Protein UmuC Is a DNA Polymerase Activated by UmuD', RecA, and SSB and Is Specialized for Translesion Replication // *J. Biol. Chem.* 1999. Vol. 274, № 45. P. 31763–31766.
 72. Tang M. et al. UmuD'(2)C is an error-prone DNA polymerase, *Escherichia coli* pol V. // *Proc. Natl. Acad. Sci. U. S. A.* 1999. Vol. 96, № 16. P. 8919–8924.
 73. Eisen J.A., Hanawalt P.C. A phylogenomic study of DNA repair genes, proteins, and processes // *Mutat. Res. Repair.* 1999. Vol. 435, № 3. P. 171–213.
 74. Duigou S. et al. Distinctive genetic features exhibited by the Y-family DNA polymerases in *Bacillus subtilis*. // *Mol. Microbiol.* 2004. Vol. 54, № 2. P. 439–451.
 75. Noirot-Gros M.-F. et al. An expanded view of bacterial DNA replication. // *Proc. Natl. Acad. Sci. U. S. A.* 2002. Vol. 99, № 12. P. 8342–8347.
 76. McKenzie G.J. et al. SOS Mutator DNA Polymerase IV Functions in Adaptive Mutation and Not Adaptive Amplification // *Mol. Cell.* 2001. Vol. 7, № 3. P. 571–579.
 77. McHenry C.S. Breaking the rules: bacteria that use several DNA polymerase IIIs. // *EMBO Rep. European Molecular Biology Organization*, 2011. Vol. 12, № 5. P. 408–414.
 78. Koonin E. V., Bork P. Ancient duplication of DNA polymerase inferred from analysis of complete bacterial genomes // *Trends Biochem. Sci.* 1996. Vol. 21, № 4. P. 128–129.
 79. Bruck I., O'Donnell M. The DNA replication machine of a gram-positive organism. // *J. Biol. Chem.* 2000. Vol. 275, № 37. P. 28971–28983.
 80. McHenry C.S. DNA replicases from a bacterial perspective. // *Annu. Rev. Biochem. Annual Reviews*, 2011. Vol. 80. P. 403–436.

81. Dervyn E. et al. Two essential DNA polymerases at the bacterial replication fork. // *Science*. 2001. Vol. 294, № 5547. P. 1716–1719.
82. Flett F. et al. A “Gram-negative-type” DNA polymerase III is essential for replication of the linear chromosome of *Streptomyces coelicolor* A3(2) // *Mol. Microbiol.* 1999. Vol. 31, № 3. P. 949–958.
83. Foster K.A. et al. DNA polymerase III of *Enterococcus faecalis*: expression and characterization of recombinant enzymes encoded by the *polC* and *dnaE* genes // *Protein Expr. Purif.* 2003. Vol. 27, № 1. P. 90–97.
84. Inoue R. et al. Genetic identification of two distinct DNA polymerases, DnaE and PolC, that are essential for chromosomal DNA replication in *Staphylococcus aureus*. // *Mol. Genet. Genomics*. 2001. Vol. 266, № 4. P. 564–571.
85. Sanders G.M., Dallmann H.G., McHenry C.S. Reconstitution of the *B. subtilis* Replisome with 13 Proteins Including Two Distinct Replicases // *Mol. Cell*. 2010. Vol. 37, № 2. P. 273–281.
86. Scheuermann R.H., Echols H. A separate editing exonuclease for DNA replication: the epsilon subunit of *Escherichia coli* DNA polymerase III holoenzyme. // *Proc. Natl. Acad. Sci.* 1984. Vol. 81, № 24. P. 7747–7751.
87. Bruck I., Goodman M.F., O’Donnell M. The essential C family DnaE polymerase is error-prone and efficient at lesion bypass. // *J. Biol. Chem.* 2003. Vol. 278, № 45. P. 44361–44368.
88. Le Chatelier E. et al. Involvement of DnaE, the second replicative DNA polymerase from *Bacillus subtilis*, in DNA mutagenesis. // *J. Biol. Chem.* 2004. Vol. 279, № 3. P. 1757–1767.
89. Au N. et al. Genetic Composition of the *Bacillus subtilis* SOS System // *J. Bacteriol.* 2005. Vol. 187, № 22. P. 7655–7666.
90. Klocko A.D. et al. Mismatch repair causes the dynamic release of an essential DNA polymerase from the replication fork. // *Mol. Microbiol.* 2011. Vol. 82, № 3. P. 648–663.
91. Anuchin A.M. et al. Histone-like proteins of bacteria (review) // *Appl. Biochem. Microbiol.* 2011. Vol. 47, № 6. P. 580–585.
92. Pruss G.J., Drlica K. DNA supercoiling and prokaryotic transcription // *Cell*. 1989. Vol. 56, № 4. P. 521–523.

93. Dillon S.C., Dorman C.J. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. // *Nat. Rev. Microbiol.* Nature Publishing Group, 2010. Vol. 8, № 3. P. 185–195.
94. Rouvière-Yaniv J., Yaniv M., Germond J.-E. E. coli DNA binding protein HU forms nucleosome-like structure with circular double-stranded DNA // *Cell.* Elsevier, 1979. Vol. 17, № 2. P. 265–274.
95. Rouvière-Yaniv J. Localization of the HU Protein on the Escherichia coli Nucleoid // *Cold Spring Harb. Symp. Quant. Biol.* 1978. Vol. 42, № 0. P. 439–447.
96. Balandina A., Kamashev D., Rouvière-Yaniv J. The bacterial histone-like protein HU specifically recognizes similar structures in all nucleic acids. DNA, RNA, and their hybrids. // *J. Biol. Chem.* 2002. Vol. 277, № 31. P. 27622–27628.
97. Bensaid A. et al. Cross-talk Between Topoisomerase I and HU in Escherichia coli // *J. Mol. Biol.* 1996. Vol. 256, № 2. P. 292–300.
98. Boubrik F., Rouvière-Yaniv J. Increased sensitivity to gamma irradiation in bacteria lacking protein HU. // *Proc. Natl. Acad. Sci. U. S. A.* 1995. Vol. 92, № 9. P. 3958–3962.
99. Li S., Waters R. Escherichia coli strains lacking protein HU are UV sensitive due to a role for HU in homologous recombination. // *J. Bacteriol.* 1998. Vol. 180, № 15. P. 3750–3756.
100. Bramhill D., Kornberg A. A model for initiation at origins of DNA replication // *Cell.* 1988. Vol. 54, № 7. P. 915–918.
101. Oberto J. et al. The HU regulon is composed of genes responding to anaerobiosis, acid stress, high osmolarity and SOS induction. // *PLoS One* / ed. Imhof A. Public Library of Science, 2009. Vol. 4, № 2. P. e4367.
102. Grove A. Functional evolution of bacterial histone-like HU proteins. // *Curr. Issues Mol. Biol.* 2011. Vol. 13, № 1. P. 1–12.
103. Yasuzawa K. et al. Histone-like proteins are required for cell growth and constraint of supercoils in DNA. // *Gene.* 1992. Vol. 122, № 1. P. 9–15.
104. Micka B., Marahiel M.A. The DNA-binding protein HBSu is essential for normal growth and development in Bacillus subtilis // *Biochimie.* 1992. Vol. 74, № 7. P. 641–650.

105. Glass J.I. et al. Essential genes of a minimal bacterium. // Proc. Natl. Acad. Sci. U. S. A. 2006. Vol. 103, № 2. P. 425–430.
106. Oberto J., Rouviere-Yaniv J. *Serratia marcescens* contains a heterodimeric HU protein like *Escherichia coli* and *Salmonella typhimurium*. // J. Bacteriol. 1996. Vol. 178, № 1. P. 293–297.
107. Swinger K.K. et al. Flexible DNA bending in HU-DNA cocrystal structures. // EMBO J. European Molecular Biology Organization, 2003. Vol. 22, № 14. P. 3749–3760.
108. Mouw K.W., Rice P.A. Shaping the *Borrelia burgdorferi* genome: crystal structure and binding properties of the DNA-bending protein Hbb. // Mol. Microbiol. 2007. Vol. 63, № 5. P. 1319–1330.
109. Pontiggia A. et al. Protein HU binds specifically to kinked DNA // Mol. Microbiol. 1993. Vol. 7, № 3. P. 343–350.
110. Zelwer C. HU Protein of *Escherichia coli* Binds Specifically to DNA That Contains Single-strand Breaks or Gaps // J. Biol. Chem. 1995. Vol. 270, № 17. P. 10291–10296.
111. Kamashev D., Balandina A., Rouviere-Yaniv J. The binding motif recognized by HU on both nicked and cruciform DNA. // EMBO J. European Molecular Biology Organization, 1999. Vol. 18, № 19. P. 5434–5444.
112. Kamashev D., Rouviere-Yaniv J. The histone-like protein HU binds specifically to DNA recombination and repair intermediates. // EMBO J. European Molecular Biology Organization, 2000. Vol. 19, № 23. P. 6527–6535.
113. Balandina A. et al. The *Escherichia coli* histone-like protein HU regulates rpoS translation // Mol. Microbiol. 2001. Vol. 39, № 4. P. 1069–1079.
114. Bonnefoy E., Rouvière-Yaniv J. HU and IHF, two homologous histone-like proteins of *Escherichia coli*, form different protein-DNA complexes with short DNA fragments. // EMBO J. 1991. Vol. 10, № 3. P. 687–696.
115. Kamashev D. et al. HU binds and folds single-stranded DNA. // Nucleic Acids Res. 2008. Vol. 36, № 3. P. 1026–1036.

116. Papazisi L. The complete genome sequence of the avian pathogen *Mycoplasma gallisepticum* strain Rlow // *Microbiology*. 2003. Vol. 149, № 9. P. 2307–2316.
117. Garner M.M., Revzin A. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. // *Nucleic Acids Res.* 1981. Vol. 9, № 13. P. 3047–3060.
118. Laemmli U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. // *Nature*. 1970. Vol. 227, № 5259. P. 680–685.
119. Hindson B.J. et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. // *Anal. Chem.* 2011. Vol. 83, № 22. P. 8604–8610.
120. Huisman O. et al. Multiple defects in *Escherichia coli* mutants lacking HU protein. // *J. Bacteriol.* 1989. Vol. 171, № 7. P. 3704–3712.
121. Pellegrini O. et al. Overproduction and improved strategies to purify the threenative forms of nuclease-free HU protein // *Biochimie*. 2000. Vol. 82, № 8. P. 693–704.
122. Cleavage of structural proteins during the assembly of the head of bacteriophage // *Nature*. 1970. Vol. 15, № 227. P. 680–685.
123. Vizcaíno J.A. et al. The PRoteomics IDentifications (PRIDE) database and associated tools: status in 2013. // *Nucleic Acids Res.* 2013. Vol. 41, № Database issue. P. D1063–9.
124. Shevchenko A. et al. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. // *Anal. Chem.* 1996. Vol. 68, № 5. P. 850–858.
125. Benjamini Y and Hochberg Y. Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing // *J. R. Stat. Soc.* 1995. Vol. 57, № 1. P. 289–300.
126. Marchler-Bauer A. et al. CDD: specific functional annotation with the Conserved Domain Database. // *Nucleic Acids Res.* 2009. Vol. 37, № Database issue. P. D205–10.
127. Kunkel M.T. et al. Spatio-temporal dynamics of protein kinase B/Akt signaling revealed by a genetically encoded fluorescent reporter. // *J. Biol. Chem.* 2005. Vol. 280, № 7. P. 5581–5587.

128. Balandina A. et al. The Escherichia coli histone-like protein HU regulates rpoS translation // Mol. Microbiol. 2001. Vol. 39.
129. Rouvière-Yaniv J., Gros F. Characterization of a novel, low-molecular-weight DNA-binding protein from Escherichia coli. // Proc. Natl. Acad. Sci. U. S. A. 1975. Vol. 72, № 9. P. 3428–3432.
130. Chodavarapu S. et al. Escherichia coli DnaA interacts with HU in initiation at the E. coli replication origin. // Mol. Microbiol. 2008. Vol. 67, № 4. P. 781–792.
131. Ryan V.T. et al. IHF and HU stimulate assembly of pre-replication complexes at Escherichia coli oriC by two different mechanisms // Mol. Microbiol. 2002. Vol. 46, № 1. P. 113–124.
132. Giangrossi M. et al. Selective expression of the β -subunit of nucleoid-associated protein HU during cold shock in Escherichia coli // Mol. Microbiol. 2002. Vol. 44, № 1. P. 205–216.
133. Liu S.-T. A HU-like Protein Binds to Specific Sites within nod Promoters of Rhizobium leguminosarum // J. Biol. Chem. 1998. Vol. 273, № 32. P. 20568–20574.
134. Grove A., Saavedra T.C. The Role of Surface-Exposed Lysines in Wrapping DNA about the Bacterial Histone-Like Protein HU † // Biochemistry. American Chemical Society, 2002. Vol. 41, № 24. P. 7597–7603.
135. C C., S G., A G. Substrate specificity of Helicobacter pylori histone-like HU protein is determined by insufficient stabilization of DNA flexure points. Portland Press Ltd., 2004.
136. E K. et al. Surface salt bridges modulate the DNA site size of bacterial histone-like HU proteins. Portland Press Ltd., 2005.
137. Grove A., Lim L. High-affinity DNA binding of HU protein from the hyperthermophile Thermotoga maritima11Edited by T. Richmond // J. Mol. Biol. 2001. Vol. 311, № 3. P. 491–502.
138. Kar S., Edgar R., Adhya S. Nucleoid remodeling by an altered HU protein: reorganization of the transcription program. // Proc. Natl. Acad. Sci. U. S. A. 2005. Vol. 102, № 45. P. 16397–16402.

139. Razin S., Yogev D., Naot Y. Molecular Biology and Pathogenicity of Mycoplasmas // *Microbiol. Mol. Biol. Rev.* 1998. Vol. 62, № 4. P. 1094–1156.
140. Dorman C.J., Deighan P. Regulation of gene expression by histone-like proteins in bacteria // *Curr. Opin. Genet. Dev.* 2003. Vol. 13, № 2. P. 179–184.
141. Kunkel T.A., Erie D.A. DNA mismatch repair. // *Annu. Rev. Biochem. Annual Reviews*, 2005. Vol. 74. P. 681–710.
142. Nakahara T. et al. Identification of proteins of *Escherichia coli* and *Saccharomyces cerevisiae* that specifically bind to C/C mismatches in DNA // *Nucleic Acids Res.* 2000. Vol. 28, № 13. P. 2551–2556.
143. Arthanari H. et al. Effects of HU binding on the equilibrium cyclization of mismatched, curved, and normal DNA. // *Biophys. J.* 2004. Vol. 86, № 3. P. 1625–1631.
144. Levitskiy S.A. et al. Purification and functional analysis of recombinant *Acholeplasma laidlawii* histone-like HU protein // *Biochimie.* 2011. Vol. 93, № 7. P. 1102–1109.
145. Fisunov G.Y. et al. Core proteome of the minimal cell: comparative proteomics of three mollicute species. // *PLoS One* / ed. Brusica V. Public Library of Science, 2011. Vol. 6, № 7. P. e21964.
146. Рогова М.А. et al. ПРОТЕОМ БАКТЕРИИ *Mycoplasma gallisepticum* *. 2009.
147. Graumann P.L., Knust T. Dynamics of the bacterial SMC complex and SMC-like proteins involved in DNA repair. // *Chromosome Res.* 2009. Vol. 17, № 2. P. 265–275.
148. Wolff E. et al. Polymerases Leave Fingerprints : Analysis of the Mutational Spectrum in *Escherichia coli* rpoB To Assess the Role of Polymerase IV in Spontaneous Mutation // *J. Bacteriol.* 2004. Vol. 186. P. 2900–2004.
149. Sciences M.L. Review The bacterial LexA transcriptional repressor // *Cell. Mol. Life Sci.* 2009. Vol. 66. P. 82 – 93.
150. Savijoki K. et al. Heat and DNA damage induction of the LexA-like regulator HdiR from *Lactococcus lactis* is mediated by RecA and ClpP // *Mol. Microbiol.* 2003. Vol. 50, № 2. P. 609–621.

151. Robertson A.B., Matson S.W. Reconstitution of the very short patch repair pathway from *Escherichia coli*. // *J. Biol. Chem.* 2012. Vol. 287, № 39. P. 32953–32966.
152. Wang D., Kreutzer D. a, Essigmann J.M. Mutagenicity and repair of oxidative DNA damage: insights from studies using defined lesions. // *Mutat. Res.* 1998. Vol. 400, № 1-2. P. 99–115.
153. Ischenko A. a, Sapparbaev M.K. Alternative nucleotide incision repair pathway for oxidative DNA damage. // *Nature.* 2002. Vol. 415, № 6868. P. 183–187.
154. Maier T., Schmidt A., Güell M. Quantification of mRNA and protein and integration with protein turnover in a bacterium // *Mol. Syst.* 2011. Vol. 7, № 511. P. 1–12.
155. Güell M. Transcriptome Complexity in a // *Science* (80-.). 2009. Vol. 1268.
156. Mandelco L. et al. A Phylogenetic Analysis of the Mycoplasmas : Basis for Their Classification // *J. Bacteriol.* 1989. Vol. 171, № 12. P. 6445–6467.
157. Sioud M., Boudabous A., Cekaite L. Transcriptional responses of *Bacillus subtilis* and *thuringiensis* to antibiotics and anti-tumour drugs // *Int. J. Mol. Med.* 2009. Vol. 23. P. 33–39.
158. Dörr T., Lewis K., Vulić M. SOS response induces persistence to fluoroquinolones in *Escherichia coli*. // *PLoS Genet.* 2009. Vol. 5, № 12. P. e1000760.
159. Janion C. Inducible SOS response system of DNA repair and mutagenesis in *Escherichia coli*. // *Int. J. Biol. Sci.* 2008. Vol. 4, № 6. P. 338–344.
160. Dörr T., Lewis K., Vulić M. SOS response induces persistence to fluoroquinolones in *Escherichia coli* // *PLoS Genet.* 2009. Vol. 5, № 12. P. e1000760.
161. Giuliodori A.M. et al. Review on bacterial stress topics. // *Ann. N. Y. Acad. Sci.* 2007. Vol. 1113. P. 95–104.
162. Layton J.C., Foster P.L. Error-Prone DNA Polymerase IV Is Regulated by the Heat Shock Chaperone GroE in *Escherichia coli* // *J. Bacteriol.* 2005. Vol. 187, № 2. P. 449–457.

163. Dascher C.C., Poddar S.K., Maniloff J. Heat shock response in mycoplasmas, genome-limited organisms. // *J. Bacteriol.* 1990. Vol. 172, № 4. P. 1823–1827.
164. Xue X. et al. The delta subunit of RNA polymerase, RpoE, is a global modulator of *Streptococcus mutans* environmental adaptation. // *J. Bacteriol.* 2010. Vol. 192, № 19. P. 5081–5092.
165. Mowles J.M. The use of ciprofloxacin for the elimination of mycoplasma from naturally infected cell lines. // *Cytotechnology.* 1988. Vol. 1, № 4. P. 355–358.

Приложение 1.

Синтетические олигонуклеотиды

Ген	Прямой праймер 5'-3'	Обратный праймер 5'-3'	Молекулярный зонд 5'-3'
<i>clpB</i>	GAAAGATTACAGGCAAAAGGTG	GCGCTTCTCCTCTTAGTACTGC	FAM-ATCAAATTGTCGGTAAATTCGCTCAATTCT-BHQ1
<i>16SrRNA</i>	CCGTGTCTCAGTCCCATTGT	GCAAGTCGATCGGATGTAGC	FAM-CGGGTAACTGAAATTCCTTCACGACCAATAGTAAC-BHQ1
<i>23SrRNA</i>	CATCCCGTAAGTTCGCAAGA	ACCCGTACCTGGATTTCACC	FAM-ATGGAGTGACGGAGAAGGTTAATGCATC-BHQ1
<i>hup2</i>	ATTTGTGCGAATCTACTGCA	AATGGCAGACGAAACTAACC	FAM-TTGAACAGTTAGTTTGATTTTCTTCCTGCTG-BHQ1
<i>hup1</i>	AGGCGGAATTCTAGTTTGTG	ATTATCGCTGAATGACTGGAG	FAM-CGGGTAACTGAAATTCCTTCACGACCAATAGTAAC-BHQ1
<i>parE</i>	CATGACTTCACCTTCAGAG	CATTTATTGCCAAAGACGGGA	FAM-CAGCGTTTTGAATTGAGACTTGGATTAGTTGG-BHQ1
<i>uvrB</i>	TGTGTGTGCGAATTAACCT	ATCGCTTATGTCATCTCATCAG	FAM-TGACGCTGACAACCTGGTTATTTTCAGAAG-BHQ1
<i>parC</i>	AAACAACCTAATCCCAAACC	CTTAAACTAATGCCAGGACCA	FAM-AGTTTGTGTTGCTTCTAAGATCTCTTAATCGAACG-BHQ1
<i>uvrD</i>	GTGTTGTAAGTTATGGTCCGA	CAGTGGTTTGACGATCATAGG	FAM-TGGTGTTAATCGTTTAAATGTAGCAATCACTCGAG-BHQ1
<i>uvrC</i>	TATGAAGTGATCTATCGCAGGT	AGTGACGATCTTATCTGTTGG	FAM-TTGTGATCTACCTGATCTGATTATCTAGATGGTGG-BHQ1
<i>gyrA</i>	TACCACTGAACCATTTGCT	CCCTCATAGTGATAGTGC GA	FAM-CAATCGGTTCTGTTCTGAAGCATCATAGTTATCA-BHQ1
<i>dinB</i>	CCAAGTCTAATGGTCTATCGT	ATCGCTTGCAGATATCTCAC	FAM-TAGCTAGAAAATTTGGCGTTCCGCTGTC-BHQ1
<i>uvrA</i>	AAACCCTGCTTCACTTACTG	TTTCCACTACCAGAAACTCC	FAM-TCCACCCAGGGAACGGACAAAAGATC-BHQ1
<i>recA</i>	TGGTATGAGAATCAACTACCTG	GTAATCCCGAGGTTACAACCTG	FAM-TCGCTTTTGATCTAATCCCTATGCATTATTGC-BHQ1
<i>nei</i>	GAACGTCATGTGTTAGTACGA	GCATAAATATTACCGATCCAG	FAM-TTGAAGTGCATATCATGACACAAGAAGATTGG-BHQ1
<i>nfo</i>	ATTATGTTGTTACAGCTCCT	ATCAGCCATTGTTTCAAGAC	FAM-AGCTAATGGTGATTCAACTAAACGAGAACGC-BHQ1
<i>recR</i>	TAATTTGCTGTTGTTTGGCC	GTCAGGGTTGATCAAATCGT	FAM-ACCTACCATGATAAGCTTGGGATTCTTCAATCAC-BHQ1
<i>ligA</i>	CTTAGGGATGATCTCAGCAG	GTTACACCTTAAGTTTCCCA	FAM-AGTGATAATCGAACCTTCTAATCAATTGGGGC-BHQ1
<i>gyrB</i>	GTGTTGTGTGAGAACTATTGGA	GACCAGATTTAGCTGAACCA	FAM-CTAAAGCAGTTTCAGCAATGTTTTACGTAAACG-BHQ1
<i>ruvB</i>	TTGTGGGTGATAACGAATGG	AGCCAGAATAAAGCATTAGACC	FAM-TTCATTAGCAATAATCTGTGCCAACGATGTTTTG-BHQ1
<i>ruvA</i>	TGCTCAATCAGCATCTCTG	GCACTGTTAACGATCAATGG	FAM-AGGTCTTTGCTATCATCAATTTCTCATCGTCATC-BHQ1
<i>ung</i>	TGATTAATTCTTCCCAACCCAG	GTAATCATTGGTCAAGATCCGT	FAM-TTGATTAATAACACCCCTTGGTTAGCCAG-BHQ1

Приложение 2.

Результаты транскрипционного профилирования системы репарации ДНК *M. gallisepticum* при различных воздействиях

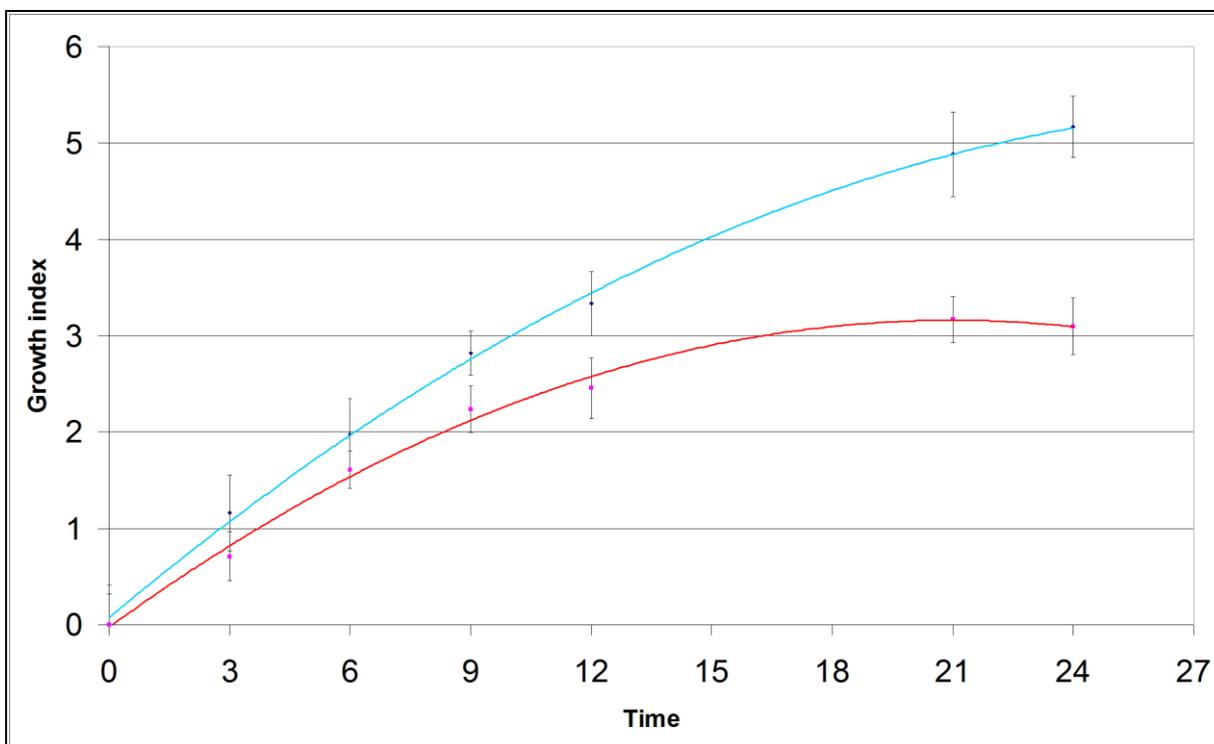
Ген	Контроль (лог-фаза), log2	Ст. откл.	Ципрофлоксацин, log2	Ст. откл.	Тетрациклин, log2	Ст. откл.	46С 5 мин, log2	Ст. откл.	46С 15 мин, log2	Ст. откл.
<i>clpB</i>	23,71	0,22	25,84	0,20	23,05	0,07	19,18	0,35	19,32	0,68
<i>16S</i>	10,07	0,37	10,20	0,04	9,89	0,14	9,69	0,05	11,09	0,68
<i>23S</i>	11,28	0,04	11,29	0,01	11,33	0,04	11,29	0,01	11,17	0,07
<i>hup2</i>	24,88	0,45	27,64	0,44	28,26	0,17	26,21	0,14	27,21	0,72
<i>hup1</i>	24,45	0,20	26,87	0,45	24,46	0,16	24,59	0,47	25,95	0,20
<i>parE</i>	25,67	1,31	27,28	0,32	25,41	0,15	26,61	0,58	26,60	1,28
<i>uvrB</i>	27,39	0,33	26,80	0,56	26,74	0,09	27,01	0,50	28,02	0,31
<i>parC</i>	27,27	0,07	26,37	0,09	26,22	0,77	25,99	0,54	26,69	0,12

uvrD	25,84	0,41	25,74	0,14	24,42	0,54	25,14	0,44	23,95	0,09
uvrC	27,40	0,30	26,81	0,22	26,06	0,69	26,67	0,38	26,39	0,03
gyrA	25,55	0,11	25,98	0,31	23,49	0,32	25,24	0,27	24,32	0,38
dinB	29,18	0,11	25,99	0,21	27,20	0,56	26,72	0,44	28,52	0,34
uvrA	26,06	0,31	27,87	1,48	24,80	1,31	25,08	0,23	23,99	0,06
recA	27,52	0,24	26,38	0,23	24,54	0,09	25,40	0,28	25,44	0,20
nei	27,08	0,43	27,75	0,12	25,03	1,19	27,45	0,33	26,82	0,35
nfo	26,25	0,27	27,10	0,63	23,98	1,06	25,59	0,11	24,66	0,33
recR	27,48	0,28	26,43	0,06	24,60	0,64	25,33	0,34	25,38	0,31
ligA	27,54	0,17	25,72	0,27	24,87	0,87	25,68	0,53	24,34	0,15
gyrB	26,71	0,19	25,81	0,32	23,34	0,41	25,50	0,18	24,00	0,44
ruvB	26,90	0,52	25,94	0,24	24,30	1,54	25,70	0,36	23,90	0,16
ruvA	27,75	0,21	25,90	0,27	25,04	1,58	25,98	0,33	24,54	0,52
ung	27,55	0,17	25,70	0,15	23,04	1,27	25,78	0,17	23,82	0,37
Ген	46С 30 мин, log2	Ст. откл.	H2O2, log2	Ст. откл.	NaCl, log2	Ст. откл.	Стац. Фаза, log2	Ст. откл.	46С 15 мин. стац. фаза, log2	Ст. откл.
clpB	21,03	0,11	23,22	0,03	21,62	0,31	19,83	0,59	19,66	0,55
16S	13,05	0,80	9,85	0,04	10,26	0,01	10,57	0,04	10,37	0,24
23S	11,98	0,69	11,27	0,01	11,29	0,01	11,26	0,03	10,89	0,42
hup2	27,57	0,13	25,94	0,72	25,63	0,05	24,50	1,33	24,51	1,40
hup1	26,40	0,02	26,15	0,27	25,47	0,21	30,77	1,16	30,92	1,00
parE	26,39	0,02	25,35	0,31	24,48	0,15	32,57	2,24	32,96	2,37
uvrB	28,16	0,01	27,04	0,33	27,24	0,47	33,47	1,60	33,48	1,22
parC	27,21	0,06	27,58	0,15	27,34	0,14	35,79	4,21	36,68	3,75
uvrD	23,88	0,23	26,69	0,24	26,11	0,07	30,99	1,55	31,58	1,56
uvrC	26,41	0,17	27,26	0,38	26,87	0,08	31,45	1,20	32,28	1,21
gyrA	23,68	0,19	26,83	0,03	26,62	0,17	32,56	1,34	33,56	1,21
dinB	27,90	0,09	29,56	0,14	27,38	0,28	37,03	1,22	37,49	1,13
uvrA	23,93	0,12	26,08	0,22	25,90	0,00	31,96	1,14	32,48	1,46
recA	25,72	0,31	26,34	0,12	26,41	0,16	31,27	1,20	32,41	1,30
nei	25,97	0,29	27,45	0,41	26,17	0,23	33,14	0,18	33,54	0,84
nfo	25,23	0,10	25,19	0,15	25,06	0,27	35,37	1,26	34,87	1,03
recR	25,58	0,25	26,34	0,04	26,23	0,14	31,17	1,21	32,21	1,30
ligA	24,46	0,22	27,57	0,07	27,19	0,26	33,11	1,15	33,62	1,14
gyrB	23,33	0,13	27,68	0,20	26,23	0,33	31,98	0,89	32,60	0,91
ruvB	23,52	0,12	27,31	0,08	25,92	0,11	32,76	0,59	32,47	0,65
ruvA	24,68	0,07	28,00	0,13	26,74	0,16	31,59	0,36	31,94	0,26
ung	24,06	0,09	27,44	0,10	25,75	0,07	35,69	0,96	35,71	1,92

*Ст. откл. – стандартное отклонение

Данные в таблице представляют среднее значение для трех биологических повторов. В каждом из биологических повторов были выполнены два технических повтора.

Приложение 3.



Определение кинетики роста жидкой культуры *M. gallisepticum*. Красная кривая отражает накопление рибосомальной РНК, синяя – геномной ДНК. По горизонтальной оси указано время в часах, по вертикальной – индекс роста в логарифмической шкале, прирост на 1 единицу соответствует двукратному увеличению.

Приложение 4.

Множественные выравнивания аминокислотных последовательностей белков репарации ДНК у *M. gallisepticum* соответствующими гомологами из *E. coli* (или) *B. subtilis*.

1. DNA ligase (gene name *ligA*)

```
>sp|P15042|DNLJ_ECOLI DNA ligase OS=Escherichia coli (strain K12) GN=ligA
PE=1 SV=2
MESIEQQLTELRTTLRHHEYLYHVMDAPEIPDAEYDRIMRELRELETKHPELITPDSPTQ
RVGAAPLAAFSQIRHEVPMLSLDNVFDDEESFLAFNKRVDRLKNNKVTWCCELKLDGLA
VSILYENGLVLSAATRGGDTTGEDITSNVRTIRAIPLKLGHENIPARLEVRGEVFLPQAG
FEKINEDARRTGKVFANPRNAAAGSLRQLDPRITAKRPLTFFCYGVGVLEGGELPDTHL
GRLLQFKKWGLPVSDRVTLCESAEVLAFLYHKVEEDRPTLGFDIDGVVIVKVNLSLAQQEQL
GFVARAPRWAVAFKFFPAQEQTFFVRDVEFQVGRGTGAIPTVARLEPVHVAGVLSNATLHN
ADEIERLGLRIGDKVVIRRAGDVIPQVVNVVLSERPEDTREVVFPTHCPVCGSDVERVEG
EAVARCTGGLICGAQRKESLKHVFSRRAMDVDGMDKIIDQLVEKEYVHTPADLFLKLTAG
KLTGLERMGPKSAQNVNALEKAKETTFFARFLYALGIREVGEATAAGLAAYFGTLEALEA
```

ASIEELQKVPDVGIVVASHVHNFFAEESNRNVEISELLAEGVHWPAPIVINAEEDIDSPFAG
 KTVVLTGSLSQMSRDDAKARLVELGAKVAGSVSKKTDLVIAGEAAGSKLAKAQELGIEVI
 DEAEMLRLLGS

>sp|O31498|DNLJ_BACSU DNA ligase OS=Bacillus subtilis (strain 168) GN=ligA
 PE=3 SV=1

MDKETAKQRAEELRRTINKYSYEYTLDEPSVDPDAEYDRMLQELIAIEEEHPDLRTPDSP
 TQRVGGAVLEAFQKVTHGTPMLSLGNAFNADDLDRDFRRVRQSVGDDVAYNVELKIDGLA
 VSLRYEDGYFVRGATRGGDTTGEDITENLKTIRNIPKMNRELSIEVRGEAYMPKRSFEA
 LNEERIKNEEEPPFANPRNAAAGSLRQLDPKIAAKRNLDIFVYSIAELDEMGETVQSQGLD
 FLDELGFKTNQERKKCGSIEEVITLIDELQAKRADLPYEIDGIVIKVDSLDQQEELGFTA
 KSPRWAIAYKFPAAEVVTKLLDIELNVGRTGVITPTAILEPVKVAGTTVSRASLHNECLI
 KEKDIRILDKVVKKAGDIIPEVVNVLDVQRTGEEKEFSMPTECPECGSELVRIEVEVAL
 RCINPECPAQIREGLIHFVSRNAMNIDGLGERVITQLFEENLVRNVADLYKLTKERVIQL
 ERMGEKSTENLISSIQKSKENSLERLLFGLGIRFIGSKAAKTLAMHFESLENLKKASKEE
 LLAVDEIGEKMADEVITYFHKEEMLELLNELQELGVNTLYKGPKKVKAEDSDSYFAGKTI
 VLTGKLEELSRNEAKAQIEALGGKLTGVSVKNTDLVIAGEAAGSKLTKAQELNIEVWNEE
 QLMGELKK

>sp|Q7NAF8|DNLJ_MYCGA DNA ligase OS=Mycoplasma gallisepticum (strain R(low /
 passage 15 / clone 2)) GN=ligA PE=3 SV=2

MIEKKNESIKEIIEDLVKCLKWEHEYYVLSNPSVSDEVYDNTYRLLGYERKYPQYVLS
 YSPTQRVGSSISNKFIVKHLYMLSLGNCNFEEELNFNENIAKISKQEDNPYVLEPKI
 DGLSISLIYLDGVLSEALTRGDGVFGESVIANIKTIKTIPLKINTTIKKIVIRGEVYVSN
 QDFEAINASRDEDKKFANSRNYASGLRNIDVSEVAKRKLNAFFYYIPNAYELGFETQYQ
 VIQQLKEWGFNVAKEIKLFSNIKELYTSLKELENNKKNLDYRIDGAVIKYNNFKDYEIIG
 YTSKFPKWAIAAYKFAPTQVQTQLKDIILNVGRTGKLTFFVAQLAPIELEGSIITYATLHNL
 EYINDLDIRINDYVYLKAAEII PKVIGVNLDKRPNNAKKLEFDYNCPSCHQPLVKKPEE
 VDWCYIYDQCKQKQLQYLIYYCSKPIMNIEGLSESTLALFFNTKVNVDVIRECNELITNQ
 VSEDLSLFKLLDNEEQTFVNSVLDIYQLERYKEIIPKWLKKGFSKSSLYNFRFQEKSF
 DKLINSINESKNRELYRLLAALNIKYIGIATAKSANTYHDIQDLKNTVEDYMRADIS
 SITANSLFSFFSDEKNWELIEQLKTLINTKDEINNDLVDSSSIYYDKKFVITGFSFISR
 NDIKKLKLKIKFVSGVSKNVDFVLAGNSPTAKKINQAKVLNIPIIQEEIWNK

sp P15042 DNLJ_ECOLI	----MESIEQQLTELRTTLRHHEYLHVMDAPEIPDAEYDRMLRELRELETKHPELITP	55
sp O31498 DNLJ_BACSU	---MDKETAKQRAEELRRTINKYSYEYTLDEPSVDPDAEYDRMLQELIAIEEEHPDLRTP	57
sp Q7NAF8 DNLJ_MYCGA	MIEKKNESIKEIIEDLVKCLKWEHEYYVLSNPSVSDEVYDNTYRLLGYERKYPQYVLS	60
	*: :: * * : : : * : : : * : : : * : : : * : : : * : : : * : : : *	

sp P15042 DNLJ_ECOLI	DSPTQRVGAAPLAAFSQIRHEVPMLSLDNVDFDEESFLAFNKRVDRLKNNKVTWCCEL	115
sp O31498 DNLJ_BACSU	DSPTQRVGGAVLEAFQKVTHGTPMLSLGNAFNADDLDRDFRRVRQSVG--DDVAYNVELK	115
sp Q7NAF8 DNLJ_MYCGA	YSPTQRVGSSISNKFIVKHLYMLSLGNCNFEEELNFNENIAKISKQ-EDNPYVLEPKI	119
	*****: * : : * * : : : * : : : * : : : * : : : * : : : * * *	

sp P15042 DNLJ_ECOLI	LDGLAVSILYENGVLVSAATRGGDTTGEDITSNVRTIRAIPLKLHGENIPARLEVRGEVF	175
sp O31498 DNLJ_BACSU	IDGLAVSLRYEDGYFVRGATRGGDTTGEDITENLKTIRNIPKMNRELS---IEVRGEAY	172
sp Q7NAF8 DNLJ_MYCGA	IDGLSISLIYLDGVLSEALTRGDGVFGESVIANIKTIKTIPLKINTTIK--KIVIRGEVY	177
	:***: : * : * : . ***** * : : * : : * : : * : : * : : * : : *	

sp P15042 DNLJ_ECOLI	LPQAGFEKINEDARRTGKVFANPRNAAAGSLRQLDPRIATAKRPLTFFCYGVGVLEGGEL	235
sp O31498 DNLJ_BACSU	MPKRSFEALNEERIKNEEEPPFANPRNAAAGSLRQLDPKIAAKRNLDIFVYSIAELDEMGV	232
sp Q7NAF8 DNLJ_MYCGA	VSNQDFEAIN--ASRDEDKKFANSRNYASGLRNIDVSEVAKRKLNAFFYYIPNAYELGF	235
	: : . * : * : : : * : * * * : * : * : * : * : * : * : * : *	

sp P15042 DNLJ_ECOLI	PDTHLGRLLQFKKWLVPVSDRVTLCESAEVLAIFYHKVEEDRPTLGFIDIDGVVIVNSLA	295
sp O31498 DNLJ_BACSU	ETQSQG-LDFLDELGFKTNQERKKCGSIEEVITLIDELQAKRADLPYEIDGIVIKVDSLD	291
sp Q7NAF8 DNLJ_MYCGA	ETQYQV-IQQLKEWGFNVAKEIKLFSNIKELYTSLKELENNKKNLDYRIDGAVIKYNNFK	294
	: : : * : : * : : : : : : : * : * * * * * : :	

sp P15042 DNLJ_ECOLI	QQEQLGFVARAPRWAVAFKFPQEQMTFVRDVEFQVGRGTGAIPTVARLEPVHVAGVLVSN	355
sp O31498 DNLJ_BACSU	QQEELGFTAKSPRWAIAYKFPAAEVVTKLLDIELNVGRTGVITPTAILEPVKVAGTTVSR	351
sp Q7NAF8 DNLJ_MYCGA	DYEIIGYTSKFPKWAIAAYKFAPTQVQTQLKDIILNVGRTGKLTFFVAQLAPIELEGSIITY	354
	: * : : : : * : * : * : * : : * : * : * : * : * : * : * : * : *	

sp P15042 DNLJ_ECOLI	ATLHNADEIERLGLRIGDKVIVIRAGDVIPQVVNVVLSERPEDTREVVFPTHCPVCGSDV	415
sp O31498 DNLJ_BACSU	ASLHNECLIKEKDIRILDKVVKKAGDIIPEVVNVLDVQRTGEEKEFSMPTECPECGSEL	411
sp Q7NAF8 DNLJ_MYCGA	ATLHNLLEYINDLDIRINDYVYLKAAEII PKVIGVNLDKRPNNAKKLEFDYNCPSCHQPL	414
	* : * * : * : . * * * * : * : : * : * : * : * : * : * : * : * : *	

sp P15042 DNLJ_ECOLI	ERVEGEAVARCTGGGLICGAQRKESLKHVFSRRAMDVDGMGDKIIDQLV-----	463
sp O31498 DNLJ_BACSU	VRIEVEVALRCIN-PECPAQIREGLIHFVSRNAMNIDGLGERVITQLF-----	458

```

sp|Q7NAF8|DNLJ_MYCGA      VKKPEEVDWYCIY-DQCKQKQLQYLIYYCSKPIMNIEGLSESTLALFFNTKVNVDVIRECN 473
      :  * .  *      * : : * :: * : *::*:*: : : .

sp|P15042|DNLJ_ECOLI      -----EKEYVHTPADLFKLTAGKLTGLE----- 486
sp|O31498|DNLJ_BACSU      -----EENLVRNVADLYKLTKERVIQLE----- 481
sp|Q7NAF8|DNLJ_MYCGA      ELITNQNVSEDLSLFKLLDNEEQTFVNSVLDIYQLERYKEII IKPWLKKGFSKSSLYNF 533
      *: *.. *::*: : :

sp|P15042|DNLJ_ECOLI      RMGPKSAQNVVNALEKAKETT FARFLYALGIREVGEATAAGLAAYFGTLEALEAASIEEL 546
sp|O31498|DNLJ_BACSU      RMGEKSTENLISSIQKSKENSLERLLFGLGIRFIGSKAAKTLAMHFESLENLKKASKEEL 541
sp|Q7NAF8|DNLJ_MYCGA      RFQEKSFDKLINSINESKNRELYRLLAALNIKYIGIATAKSIANTYHDIDQLKNTVEDY 593
      *: ** ::::::::::*: : :* .**: : * : * : : : * : * :

sp|P15042|DNLJ_ECOLI      QKVPDVGIVVASHVHNFFAEESNRNVISELLAEGVH--WPAPIVINAEEIDSPFAGKTVV 604
sp|O31498|DNLJ_BACSU      LAVDEIGEKMADAVITYFHKEEMLELNLQELGVNTLYKGPVKVAEDSDSYFAGKTIV 601
sp|Q7NAF8|DNLJ_MYCGA      MRLADISSITANSLSFFSDEKNWELIEQLKTLN--TKDEINNDLVDSSSIYDKKFV 651
      : :.. * . : .* .* : : : * : : : : . : . * : . * . *

sp|P15042|DNLJ_ECOLI      LTGSLSQMSRDDAKARLVELGAKVAGSVSKKTDLVIAG-EAAGSKLAKAQELGIEVIDEA 663
sp|O31498|DNLJ_BACSU      LTGKLEELSRNEAKAIEALGGKLTGSVSKNTDLVIAG-EAAGSKLTKAQELNIEVWNEE 660
sp|Q7NAF8|DNLJ_MYCGA      ITGSFISISRNDIIKKLSLYKIKFVSGVSKNVDFVLGNSPTAKKINQAKVLNIPIIQEE 711
      : ** : . : : *      * . . . . * : * : * * . . . . * : : * : * : : *

sp|P15042|DNLJ_ECOLI      EMLRLLGS 671
sp|O31498|DNLJ_BACSU      QLMGELKK 668
sp|Q7NAF8|DNLJ_MYCGA      IWNK---- 715
    
```

Organism	Active site (N6-AMP-lysine intermediate)	Metal binding (Zn)	Binding site (NAD)	Site (Interaction with target DNA)
E. coli (strain K12)	K115	408, 411, 426, 432	113,136, 173, 290, 314	487, 492
B. subtilis (strain 168)	K115	404, 407, 422, 427	113, 136, 170, 286, 310	No information
M. gallisepticum (strain R(low / passage 15 / clone 2))	K119	407, 410, 425, 430	117, 140, 175, 289, 313	No information

2. DNA-methyltransferase (gene name *hsdM*)

Gene is absent in *B. subtilis* (strain 168) genome.

```

>sp|P08957|T1MK_ECOLI      Type I restriction enzyme EcoKI M protein
OS=Escherichia coli (strain K12) GN=hsdM PE=1 SV=1
MNNNDLVAKLWKLCDNLRDGGVSYQNYVNELASLLFLMKCKETGQAEYLPEGYRWDDLK
SRIGQEQLQFYRKMVLVHLGEDDKKLVQAVFHNVSTTITEPKQITALVSNMDSLWDYNGAH
GKSRDDFDGMYEGLLQKNANETKSGAGQYFTPRPLIKTIIHLLKPQPREVVQDPAAGTAG
FLIEADRYVKSQTNDLDDLDGDTQDFQIHRAFIGLELVPGTRRLALMNCLLHDIEGNLDH
GGAIRLGNLTGSDGENLPKAHIVATNPPFGSAAGTNI TRTFVHPTS NKQLCFMQHIETL
HPGGRAAVVVPDNLVLEGGKGTDIRRDLMDKCHLHTILRLPTGIFYAQGVKTNVLFFTKG
TVANPNQDKNCTDDVWVYDLR TNMPSFGKRT PFTDEHLQPFERYGEDPHGLSPRTEGEW
SFNAEETEVAADSEENKNTDQHLATSRWRKFSREWIRTAKSDSLDISWLKDKDSIDADSLP
EPDVLAAEAMGELVQALSELDALMREL GASDEADLQRQLLEEAFFGGVKE
>tr|Q7NAH5|Q7NAH5_MYCGA      Type I restriction-modification system
methyltransferase (M) subunit OS=Mycoplasma gallisepticum (strain R(low /
passage 15 / clone 2)) GN=hsdM PE=4 SV=2
MTKQELAREIWAMANEMRGNIEANDYKDYILGFLFYKYLSDKQDEYFASKNVVKDEDKKQ
YLVALAEADKKGIASIIHKCKDLGYIAYENLFSWTWIKNYNPGDDLSDKVSTALNSFER
SILEKYEESFKDIFKDLQAGIQKLGNTAYERSEAIWNICNLINKIPITSKQDYDILGFVY
EYLISMFAANAGKKAGEFYTPHEVSQ LMSVIAANHLKGLKNVSIYDPTSGSLLITLGREL
KKIDKNVKIQYYAQEVIDT TYNITRMNLLMNDVH SVNMFACGDTL KEDWPFVYEEQKYK
SKRTDAVVS NPPYSLAWNTENKENDPRFRYGLAPKSKSELAFLLSLYHLEDHGILTIVL
PHGVLFRRGGSELQIRQNLISHDHIDAIIGLPSNIFFGTGIPTIIMVLKRSKTKKEKNNVL
    
```

FIDASKYFTKEGNKKNLQSSDIVRIYDAFSAREDIPGFARVVSHEEIKANEYNLNI PKYI
 DLVDNGDNHNL YSSIFSGIPHN DIDKLSDFWSTFPTLKKALLNDNGKNYQLKDHDVEKVI
 SNNEEVKKYLSDFNKSLESRLTYFKKSLIDTDLNVIDLYNVYQQFLDKIQSVLKS YKLLD
 YYQAFQLFDNEW TIIENDLKVIKSADNKNSFDTIRELKEHDSSTISAKADKKLVSKNTTY
 GIYQTPVIPF EFVTKLKFNNQLEALEINSNKVEEINARLEELLNEVAGYESDVVNNFYKK
 EENKLNFD EIKKQLKNLSVVAKSQPESVEALLVEALSIDKEKRALNSAIRKAKLQLEKDT
 IQAYSKLTDEEAKTLLCLKWIDPLITAI SKLTKYKIENLASELNRLDKKYIDKLS DLEVQ
 IIQTQNSLIELINQLEGPDL DMEGLNELKKILGKK

Organism	Binding site (S-adenosyl-L-methionine)	Active site
E. coli (strain K12)	E216	No information
M. gallisepticum (strain R(low / passage 15 / clone 2))	No information	No information

```

sp|P08957|T1MK_ECOLI      MNNNDLVAKLWKLCDNLRDGGVSYQNYVNELASLLFLKMCKETG----- 44
tr|Q7NAH5|Q7NAH5_MYCGA  MTKQELAREIWAMANEMR-GNIEANDYKDYILGFLFYKYLSDKQDEYFAS 49
*.:*:* .:* :*: * *.. :* : : :*: * ..

sp|P08957|T1MK_ECOLI      -----QEAEYLP 51
tr|Q7NAH5|Q7NAH5_MYCGA  KNVVKDEDKKQYLVALAEADKKGIASIIHKCKDLGYIAYENLFWSTWIK 99
: :

sp|P08957|T1MK_ECOLI      EGYRWDDLKSRIGQEQLQFYRKMLVHLGEDDKKLVQAVFHNVSTT-ITEP 100
tr|Q7NAH5|Q7NAH5_MYCGA  NYNPGDDLSDKVVSTALNSFERSILEKEYEESFKDIFKDLQAGIQKLGNTAY 149
:   ***.:. . . * *:* : * . *.: : : . . . *

sp|P08957|T1MK_ECOLI      KQITALVSNMDSLWDYNGAHGKSRDDFGDMYEGLLQKNANETKSGAGQYF 150
tr|Q7NAH5|Q7NAH5_MYCGA  ERSEAIWNICNLINKIPITSKQYDILGFVYEYLISMFAANAGKKAGEFY 199
: : * : . : : : : : . * : * : * * : . * : : . * : : :

sp|P08957|T1MK_ECOLI      TPRPLIKTI IHLKQPPEVVQDPAAGTAGFLIEADRYVKSQTNLDDLD 200
tr|Q7NAH5|Q7NAH5_MYCGA  TPHEVSQLMSVIAANHLKGLKNVSIYDPT-----SGSLITLQRELKKID 244
** : : : : : : : : : . . : : : : : : * : : :

sp|P08957|T1MK_ECOLI      GDTQDFQIHRAFIGL LVPGTRRLALMNCLLHDIEGN---LDHGGAIRLG 247
tr|Q7NAH5|Q7NAH5_MYCGA  KN-----VKIQYYAQVIDTTYNITRMNLLMNDVHSVNMFAKCGDTLKED 289
:   : : : . * : : * : : * * : * : * : . . * : : : .

sp|P08957|T1MK_ECOLI      NTLGSDGENLP--KAHIVATNPPFGSAAGTN-----ITRTFVHPTSNIKQ 289
tr|Q7NAH5|Q7NAH5_MYCGA  WPFVYEEQKYKSKRTDAVSNPPYSLAWNTENKENDPRFRYGLAPKSKSE 339
. : : : : : . * : * : * : . * : * : * : * : * : * : :

sp|P08957|T1MK_ECOLI      LCFMQHIIETLHPGGRAAVVVDNVLFEGGKGTDIRRDLMDKCHLHTILR 339
tr|Q7NAH5|Q7NAH5_MYCGA  LAFLHSLYHLEDHGILTI VLPHGVLFRGGSELQIRQNLISHDHIDAIIG 389
* . * : * : * : : * : * : * : * : * : * : * : * : * :

sp|P08957|T1MK_ECOLI      LPTGIFYAQGVKTNVLFFTKG-----TVANPNQDKNC 371
tr|Q7NAH5|Q7NAH5_MYCGA  LPSNIFFGTGIP TIIMVLKRSKTKKEKNVLFIDASKYFTKEGNKKNLQS 439
** : * : . * : * : : : : . . * . : : : .

sp|P08957|T1MK_ECOLI      TDDVWVYDLRTN--MPSFGKRTFPTDEHLQPFER-----VYGEDPH 410
tr|Q7NAH5|Q7NAH5_MYCGA  SDIVRIYDAFSAREDIPGFARVVSHEEIKANEYNLNI PKYIDLVDNGDNH 489
: * * : * * : : * : * : . . : : : : : * . * *

sp|P08957|T1MK_ECOLI      GLSPRTEG----- 418
tr|Q7NAH5|Q7NAH5_MYCGA  NLYSSIFSGIPHN DIDKLSDFWSTFPTLKKALLNDNGKNYQLKDHDVEKV 539
.* . .

sp|P08957|T1MK_ECOLI      ----- 589
tr|Q7NAH5|Q7NAH5_MYCGA  ISNNEEVKKYLSDFNKSLESRLTYFKKSLIDTDLNVIDLYNVYQQFLDKI

sp|P08957|T1MK_ECOLI      -----EWSFNAEETEVAADSEENKN----- 437
tr|Q7NAH5|Q7NAH5_MYCGA  QSVLKS YKLLDYYQAFQLFDNEW TIIENDLKVIKSADNKNSFDTIRELKE 639
** : : : * . * : * *
    
```

```

sp|P08957|T1MK_ECOLI      -----TDQHLATS--RWRKFSREWIRTAKSDSLDISWLKDKDSIDA 476
tr|Q7NAH5|Q7NAH5_MYCGA  HDSSSTISAKADKKLVSKNTTYGIYQTPVIPFEFVTKLKFENNQLEALEINS 689
                          *:::*:::  :  :.  *  .*:::  :  .*:::

sp|P08957|T1MK_ECOLI      DSLPEPDVLAEEAMGELVQALS-----ELDALMRELG---- 508
tr|Q7NAH5|Q7NAH5_MYCGA  NKVEEINARLELLNEVAGYESDVVNNFYKKEENKLNFDI IKKQLKNLSV 739
                          :.: * :.  * :.*::  *  ::* : :.*

sp|P08957|T1MK_ECOLI      -ASDEADLQRQLLEEAFFGGVKE----- 529
tr|Q7NAH5|Q7NAH5_MYCGA  VAKSQPESVEALLVEALSIDKEKRALNSAIRKAKLQLEKDTIQAYSKLT 789
                          *..:..:  . ** **:.  **

sp|P08957|T1MK_ECOLI      -----
tr|Q7NAH5|Q7NAH5_MYCGA  EEAKTLLCLKWIDPLITAI SKLTKYKIENLASELNRLDKKYIDKLSLEV 839

sp|P08957|T1MK_ECOLI      -----
tr|Q7NAH5|Q7NAH5_MYCGA  QIIQTQNSLIELINQLEGPDL DMEGLNELKILGKK 875

```

3. 5'-3'-exonuclease (gene name *exo*)

```

>sp|P00582|DPO1_ECOLI DNA polymerase I OS=Escherichia coli (strain K12)
GN=polA PE=1 SV=1
MVQIPQNPLIILVDGSSYLYRAYHAFPLTNSAGEPTGAMYGVNLMLRSLIMQYKPTHAAV
VFDAKGGKTFRDELFEHYKSHRPPMPDDLRAQIEPLHAMVKAMGLPLLAVSGVEADDVIGT
LAREA EKAGRPVLIISTGDKMAQLVTPNITLINTMTNTILGPEEVNKNYGVPELIIIDFL
ALMGDSSDNIPGVPVGEKTAQALLQGLGGLDTLYAEPEKIAGLSFRGAKTMAAKLEQNK
EVAYLSYQLATIKTDVELELTCEQLEVQQPAAEELLGLFKKYEFKRWTADVEAGKWLOAK
GAKPAAKPEQETSVADEAPEVTATVISYDNYVTILDEETLKAWIAKLEKAPVFAFDTETDS
LDNISANLVGLSFAIEPGVAAYIPVAHDYLDAPDQISRERALELLKPLEDEKALKVGGQN
LKYDRGILANYGIELRGIAFDTMLESYILNSVAGRHDMDSLAERWLKHKITITFEEIAGKG
KNQLTFNQIALEEAGRYAAEDADVTLQLHLKMWPDLQKHGKPLNVFENIEMPLVPVLSRI
ERNGVKIDPKVLHNSHEELTLRLAELEKKAHEIAGEEFNLSSTKQLQTI LFEKQGIKPLK
KTPGGAPSTSEEVLEELALDYPLPKVILEYRGLAKLKSTYTDKLP LMINPKTGRVHTSYH
QAVTATGRLSSTDPNLQNI PVRNEEGRRIRQAFIAPEDYVIVSADYSQIELRIMAHLSRD
KGLLTAF AEGKDIHRATAAEVFGLEPLETVTSEQRSAKAINFGLIYGMSAFGLARQLNIP
RKEAQKYMDLYFERYPGVLEYMERTRAQAKEQGYVETLDGRRLYLPDIKSSNGARRAAAE
RAAINAPMQGTAADI IKRAMIAVDAWLQAEQPRVRMIMQVHDEL VFEVHKDDVDVAVAKQI
HQLMENCNTRLDVPLLVEVGS GENWDQAH
>sp|O34996|DPO1_BACSU DNA polymerase I OS=Bacillus subtilis (strain 168)
GN=polA PE=3 SV=1
MTERKKLVLDGNSLAYRAFFALP LLSNDKGVHTNAVYGFAMILMKMLEDEKPTHMLVAF
DAGKTTFRHGTFKEYKGG RQKTPPELSEQMPFIRELLDAYQISRYELEQYEADDIIGTLA
KSAEKDGFVVKVFSGDKDLTQLATDKTTVAITRKGITDVEFYTPPEHVKEKYGLTPEQIID
MKGLMGDSSDNIPGVPVGEKTAIKLLKQFDSVEKLLS IDEVSGKLLKEKLEEFKDQAL
MSKELATIMTDAPIEVSVSGLEYQGFNREQVIAIFKDLGFNTLLERLGEDSAEAEQDQSL
EDINVKTVDVTS DILVSPSAFVVEQIGDNYHEEPILGFSIVNETGAYFIPKDIAVESEV
FKEWVENDEQKKWVFD SKRAVVALRWQGIELKGA EFDTLAAYI INPGNSYDDVASVAKD
YGLHIVSSDES VYGKAKRAVPSEDLSEHLGRKALAIQSLREKLVQELENNDQLELFEE
LEMLLALILGEMESTGVKVDVDR LKRMGEELGAKLKEYE EKIHEIAGEPFNINSPKQLGV
ILFEKIGLPVVKTKTKTG YSTSADVLEKLDKHDIVDYILQYRQIGK LQSTYIEGLLVTR
PDSHKVHTRFNQAL TQTGRLSSTDPNLQNIPIRLEEGRKIRQAFVPSEKDWLIFAADYSQ
IELRVLAHISKDENLIEAFTNDMDIHTKTAMDVFHVAKDEVTSAMRRQAKAVNFGIVYGI
SDYGLSQNLGITRKEAGAFIDRYLESFQGVKAYMEDSVQEAQKQGYVTTLMHRRRYIPEL
TSRNFNIRSFAERTAMNTP IQGSAADI IKKAMIDMAAKLKEKQLKARLLLQVHDELIFEA
PKEEIEILEKLVPEVMEHALALDVPLKVDVFASGPSWYDAK
>tr|Q7NBN5|Q7NBN5_MYCGA 5'-3' exonuclease OS=Mycoplasma gallisepticum (strain
R(low / passage 15 / clone 2)) GN=exo PE=4 SV=2
MKSTKKALI IDGNSLVFRAFYATLSMYEYAIKKGIRPSNGIKTSLKMINKILNSDQYDYA
LVAFDSKEKTDRAKIYEGYKATRKKPVEGLIEQLVALQDGF SYLGLNVLSSPGIEADDLI
GSFSALANKDQITCHIYTSDDIFQLVNQYNVYQFVKGVS VFNQVHERNFQEHFHDLPK
EDVIQYKALVGDSSDNIPGVKIGEKTA VQLIKDYLNIDNIYANLDQIKPSIKDKLVANK
ANCYLSKELATIRTDCLVDQYINNFKLKLPLDQQNYFAFCEYYKISHLD

```

```

sp|P00582|DPO1_ECOLI      MVQIPQNPILILVDGSSYLYRAYHAFPPLTNSAGEP----TGAMYGVNLML 46
sp|O34996|DPO1_BACSU     --MTERKKLVLDVGNLAYSRAFFALPLLSNDKGVH----TNAVYGFAMIL 44
tr|Q7NBN5|Q7NBN5_MYCGA  --MKSTKKALIIDGNSLVFRAFATLSMEYAIKKGIRPSNGIKTSLKMI 48
      :  ::**.*  :**.*  :  :  :  :  :

sp|P00582|DPO1_ECOLI      RSLIMQYKPTHAAVVFDAKGKTRFDELFEHYKSHRPPMPDDLRAQIEPLH 96
sp|O34996|DPO1_BACSU     MKMLEDEKPTHMLVAFDAGKTTFRHGTFKEYKGGRRQKTPPELSEQMPFIR 94
tr|Q7NBN5|Q7NBN5_MYCGA  NKILNSDQYDYALVAFDSKEKTDRAKIYEGYKATRKKPVEGLIEQLVALQ 98
      .: . :  :  *.*:  .* *  : : *. *  * * : :

sp|P00582|DPO1_ECOLI      AMVKAMGLPLLAVSGVEADDVIGTLAREAEKAGRPVLISTGDKDMAQLVT 146
sp|O34996|DPO1_BACSU     ELLDAYQISRYELQYEADDIIGTLAKSAEKDGFVKVFSGDKDLTQLAT 144
tr|Q7NBN5|Q7NBN5_MYCGA  DGFSYLGVLNVLSSPGIEADDLIGSFSALANKDQITCHIYTSQDIFQLVN 148
      ..  :  :  ****:**:  : * *  :  : . : * : * ..

sp|P00582|DPO1_ECOLI      PNITLINTMT----NTILGPEEVVNKYG-VPELIIIDFLALMGDSSDNIP 191
sp|O34996|DPO1_BACSU     DKTTVAITRKGITDVEFYTPHVKEKYG-LTPEQIIDMKGLMGDSSDNIP 193
tr|Q7NBN5|Q7NBN5_MYCGA  QYNVVYQFVKGVSFVNQVHERNFQEHFHLKPEDVIQYKALVGDSSDNIP 198
      . :  .  . . . : :  : * * * : * . : *****

sp|P00582|DPO1_ECOLI      GVPGVGEKTAQALLQGLGLDTLYAEPEKIAGLSFRGAKTMAAKLEQNKE 241
sp|O34996|DPO1_BACSU     GVPGVGEKTAIKLLKQFDSVEKLLSIDEVS-----GKKLKEKLEEFKD 237
tr|Q7NBN5|Q7NBN5_MYCGA  GVKGIGEKTAVQLIKDYLNIDNIYANLDQIK-----PSIKDKLVANKA 241
      * * * : ***** * :  . : . :  . : :  . :  * * *

sp|P00582|DPO1_ECOLI      VAYLSYQLATIKTDVELELTCEQLEVQQPAAEELLGLFKKYEFKRWTADV 291
sp|O34996|DPO1_BACSU     QALMSKELATIMTDAPIEVSVSGLLEYQGFNREQVIAIFKDLGFN----- 281
tr|Q7NBN5|Q7NBN5_MYCGA  NCYLSKELATIRTDCLVDQYINNFKLPLDQQNYFAFCEYYKIS----- 285
      . : * : ***** * *  : :  . : :  :  : :  :  : .

sp|P00582|DPO1_ECOLI      EAGKWLQAKGAKPAAKPQETSVADEAPEVTATVISYDNYVTILDEETLKA 341
sp|O34996|DPO1_BACSU     ---TLLERLGEDSAEAEQDQSLLEDINVKTVTDVTS----- 313
tr|Q7NBN5|Q7NBN5_MYCGA  ---HLD----- 288
      * :

sp|P00582|DPO1_ECOLI      WIAKLEKAPVFAFDTETDSDLNISANLVGLSFAIEPGVAAYIIPVAHDYLD 391
sp|O34996|DPO1_BACSU     --DILVSPSAFVVEQIGDNYH--EPIILGFSIVNETG--AYF----- 349
tr|Q7NBN5|Q7NBN5_MYCGA  -----

sp|P00582|DPO1_ECOLI      APDQISRERALELLKPLEDEKALKVGNLKYDRGILANYGIELRGI AFD 441
sp|O34996|DPO1_BACSU     IPKDIAVES--EVFKWVENDEQKKWVFDKRAVVALRWQGIELKGA EFD 397
tr|Q7NBN5|Q7NBN5_MYCGA  -----

sp|P00582|DPO1_ECOLI      TMLESYILNSVAGRHMDSLAERWLKHKTTITFEEIAGKGNQLTFNQIAL 491
sp|O34996|DPO1_BACSU     TLLAAYIINPGNSYDDVASVAKDYGLHIVSSDES SVYGKAKRAVPSEVDL 447
tr|Q7NBN5|Q7NBN5_MYCGA  -----

sp|P00582|DPO1_ECOLI      EEAGRYAAEDADVTLQLHLKMWPDQLKHKGPLNVFENIEMPLVPVLSRIE 541
sp|O34996|DPO1_BACSU     SEH---LGRKALAIQSLREKLVQELNND-QLELFEELEMP LALILGEME 493
tr|Q7NBN5|Q7NBN5_MYCGA  -----

sp|P00582|DPO1_ECOLI      RNVGKIDPKVLHNSHEELTLRLAELEKKAHEIAGEEFNLSSTKQLQTILF 591
sp|O34996|DPO1_BACSU     STGVKVDVDR LKRMGEELGAKLKEYEEKIHEIAGEPFNINSPQLGVILF 543
tr|Q7NBN5|Q7NBN5_MYCGA  -----

sp|P00582|DPO1_ECOLI      EKQGIKPLKKTGGAPSTSEEVLEELALDYPLPKVILEYRGLAKLKSTYT 641
sp|O34996|DPO1_BACSU     EKIGLPVVKKTGTGY-STSA DVLEKLADKHDIVDYILQYRQIGLQSTYI 592
tr|Q7NBN5|Q7NBN5_MYCGA  -----

sp|P00582|DPO1_ECOLI      DKLPLMINPKTRGVHTSYHQAVTATGRLSSTDPNLQNI PVRNEEGRRI RQ 691
sp|O34996|DPO1_BACSU     EGLLKVTRPDSHKVHTRFNQALTQTGRLSSTDPNLQNIPIRLEEGRKIRQ 642
tr|Q7NBN5|Q7NBN5_MYCGA  -----

sp|P00582|DPO1_ECOLI      AFIAP-EDYVIVSADYSQIELRIMAHLSRDKGLLTAFAEGKDIHRATAAE 740
sp|O34996|DPO1_BACSU     AFVPSEKDWLIFAADYSQIELRVLAHISKDENLIEAFTNDMDIHTKTAMD 692
tr|Q7NBN5|Q7NBN5_MYCGA  -----

```

```

sp|P00582|DPO1_ECOLI      VFGLPLETVTSEQRRSAKAINFGLIYGMSAFGLARQLNIPRKEAQKYMDL 790
sp|O34996|DPO1_BACSU     VFHVAKDEVTSAMRRQAKAVNFGIVYGISDYGLSQNLGITRKEAGAFIDR 742
tr|Q7NBN5|Q7NBN5_MYCGA  -----

sp|P00582|DPO1_ECOLI      YFERYPGVLEYMERTRAQAKEQGYVETLDGRRLYLDPDIKSSNGARRAAE 840
sp|O34996|DPO1_BACSU     YLESFQGVKAYMEDSVQEAQKQGYVTTLMHRRRYIPELTSRNFNIRSF AE 792
tr|Q7NBN5|Q7NBN5_MYCGA  -----

sp|P00582|DPO1_ECOLI      RAAINAPMQGTAADI IKRAMIAVD AWLQAEQPRVRMIMQVHDELVFV EHK 890
sp|O34996|DPO1_BACSU     RTAMNTPIQGSAADI IKKAMIDMAAKLKEKQLKARLLLQVHDELIF EAPK 842
tr|Q7NBN5|Q7NBN5_MYCGA  -----

sp|P00582|DPO1_ECOLI      DDVDAVAKQIHQLMENCTRLDVP LLVEVGSGENWDQAH 928
sp|O34996|DPO1_BACSU     EETEILEKLVPEVMEHALALD VPLKVD FASGPSWYDAK 880
tr|Q7NBN5|Q7NBN5_MYCGA  -----

```

4. Putative vsr protein (gene name MGA_0793)

Gene is absent in *B. subtilis* (strain 168) genome.

```

>sp|P09184|VSR_ECOLI Very short patch repair protein OS=Escherichia coli (strain K12)
GN=vsr PE=1 SV=3
MADVHDKATRSKNMRAIATRDTAIEKRLASLLTGQGLAFRVQDASLPGRPDFVVD EYRCV
IFTHGCFWHHHHCYLFKVPATRTEFWLEKIGKNVERDRRDISRLQELGWRVLIWECALR
GREKLTDEALTERLEEWICGEGASAQIDTQGIHLLA
>tr|Q7NC14|Q7NC14_MYCGA Putative helicase superfamily protein OS=Mycoplasma
gallisepticum (strain R(low / passage 15 / clone 2)) GN=MYCGA0950 PE=4 SV=2
MANDWTWLRTRLINNKTKSNTFWIRTRGSSVMDIGVLLKLTSNIDLSIKGILSYLNSND
TLELDRKRLRELSDL EAFNLFGIETAKELYKEYTDQFSKIFKKLKVEERVTDGTSLFIGL
PIIEGCNQWGSYRAPLLYVEVVLYPVNQYQKIVFKINRSEFWINTTILAVEASKRGILF
ENKYDSSKLDQALEIFRTFDIGFKKPSSTNELINFKEMTKKAFLENWEQNGGINNIVNN
VVLGNFDIKGDKLLKDFTEILDKDPD TVDEVFNKKD LLLFNYEKFA NEYSLSDIYLTSHL
DFFQQLAVKHALEGDVVIEGPPGTGKSETILNINIALKGTALFVSEKTTATEVVYNR
LGKFKHLALYIPSLNKEPGKFYRQFSDYENYFSENYDQVHKTPNAKFDPPDY LKRYLEQS
YIIHKIYNYEINSGENVYSFLNLIILNYKPMVDVHINIDDYTRFDEWLKIYTNQDWMTKHQ
EYIALFNEIDTKWKASTFATFLKIHQKDPNDIKTLLYAIHLYAKKGIKDEYRVSFFFRP
HEKVIESAKLVTEQINKFIELEQYKSETKKRTILKNLEIDIKRYHKQYFNSWFVQNHSGA
FLSKLNSAQSTLDNLDNYSSD VVYIQSCKRNKAEI IKNFYELYRHDKRGLLDVCRQG
RNKSFKNIAWFWKLNREI IKKMFKIHIMSFETASILLENKKDLYDYVIVDEASQVFLERA
IPALYRAKKYI IAGDTKQIQPSSFFSSRSDYDDVALDKLADEEILEVEESVNAVSLIHYL
KERSRINVVLRHYRHSNFGDLIAFTNDHVYDNELI FMNKAIKQKESFIVHDIIDGKWKDR
KNIPEAQAIVSRIQRLTKTADYQKSLGIVAFNRSQADLIELMLDKLNDPLVNEWRERNND
NGEYIGL FVKSVENVQGD ERDIIIFSVAYDKSVVSYGPISSTNGVNRNLNVAITRAKDRI
ELFKTNKASEYNGWGSSAGTRLFVEYLSYCENVANHSDYTTYDRQTTEIEEKLKDKSLI
FDDVKSTLEKAFGQYFTIKRNV DNGSYNFD FVIYHEEVPFLVIDLDIKPFKGMADFNESF
IYRNIFLKNRGWKHFIIWSTEWKLNKRKVVLLAIKDILDKRIQMNQK

```

```

sp|P09184|VSR_ECOLI      1 -----MADV 4
tr|Q7NC14|Q7NC14_MYCGA 1 MANDWTWLRTRLINNKTKSNTFWIRTRGSSVMDIGVLLKLTSNIDLSIKGILSYLNSNDTLELDRKRLRELSDL EAFNLF 80

sp|P09184|VSR_ECOLI      5 HDKATRSKNMRAIATRDTAIEKRLASLLTGQGLAFRVQDASLPGRPDFVVD EYRCVIFTHGCFWHHHHCY---LFKVPAT 81
tr|Q7NC14|Q7NC14_MYCGA 81 FGIETAKELYKEYTDQFSKIFKKLKVEERVTDGTSLFIGLPIIEGCNQWGSYRAPLLYVEVVLYPVNQYQKIVFKI--N 158

sp|P09184|VSR_ECOLI      82 RTEFWLEKIGKNVERDRR-----DISRLQ-ELGWRVLIWECALRG---REKLTDEALTER--LEEWICGEGAS--- 144
tr|Q7NC14|Q7NC14_MYCGA 159 RSEFWINTTILAVEASKRGILFENKYDSSKLDQALEIFRTFDIGFKKPSSTNELINFKEMTKKAFLENWEQNGGINNIV 238

sp|P09184|VSR_ECOLI      145 -----AQIDTQGIHLLA----- 156
tr|Q7NC14|Q7NC14_MYCGA 239 NNVLGNFDIKGDKLLKDFTEILDKDPD TVDEVFNKKD LLLFNYEKFA NEYSLSDIYLTSHLDFFQQLAVKHALEGDVVI 318

sp|P09184|VSR_ECOLI      ----- 319
tr|Q7NC14|Q7NC14_MYCGA 319 EGGPPGTGKSETILNINIALKGTALFVSEKTTATEVVYNRNLGKFKHLALYIPSLNKEPGKFYRQFSDYENYFSENYD 398

sp|P09184|VSR_ECOLI      ----- 399
tr|Q7NC14|Q7NC14_MYCGA 399 QVHKTPNAKFDPPDY LKRYLEQSYIIHKIYNYEINSGENVYSFLNLIILNYKPMVDVHINIDDYTRFDEWLKIYTNQDWMTK 478

sp|P09184|VSR_ECOLI      ----- 479
tr|Q7NC14|Q7NC14_MYCGA 479 HQEYIALFNEIDTKWKASTFATFLKIHQKDPNDIKTLLYAIHLYAKKGIKDEYRVSFFFRPHEKVIESAKLVTEQINKF 558

sp|P09184|VSR_ECOLI      ----- 638
tr|Q7NC14|Q7NC14_MYCGA 559 IELEQYKSETKKRTILKNLEIDIKRYHKQYFNSWFVQNHSGAFLSKLNSAQSTLDNLDNYSSD VVYIQSCKRNKAEI 638

```

```

sp|P09184|VSR_ECOLI -----
tr|Q7NC14|Q7NC14_MYCGA 639 IKNFYELYRHDKRGLLDVCVRQGRNKSFKNIAWWFKLNREI IKKMFKIHMSPETASILLENKKDLYDYVIVDEASQVFLE 718

sp|P09184|VSR_ECOLI -----
tr|Q7NC14|Q7NC14_MYCGA 719 RAIPALYRAKKYIIAGDTKQLQPSFFSSRSYDDVALDKLADEEILEVEESVNAVSLIHYLKERSRINVVLRVYHRSNF 798

sp|P09184|VSR_ECOLI -----
tr|Q7NC14|Q7NC14_MYCGA 799 GDLIAFTNDHVVDNELIFMKAIKQKESFIVHDIIDGKWKDRKNIPEAQIVSRIQRLTKTADYQKSLGIVAFNRSQADL 878

sp|P09184|VSR_ECOLI -----
tr|Q7NC14|Q7NC14_MYCGA 879 IELMLDKLNDPLVNEWRERNDNGEYIIGLFVKSVENVQGDERRDIIIFSVAYDKSVVSYGPISSSTNGVNRNLNVAITRAKD 958

sp|P09184|VSR_ECOLI -----
tr|Q7NC14|Q7NC14_MYCGA 959 RIELFKTNKASEYNGWSSSAGTRLFVEYLSYCENVANHSDYTTYDRQTTEIEEKLKDKSLIFDDVKSTLEKAFQYFTI 1038

sp|P09184|VSR_ECOLI -----
tr|Q7NC14|Q7NC14_MYCGA 1039 KRNVDNGSYNPFVVIYHEEVPFLVIDLDIKPFKGMADFNESFIYRNIFLKNRGWKHFIIWSTEWKLNKRKRVLLAIKIDILD 1118

sp|P09184|VSR_ECOLI -----
tr|Q7NC14|Q7NC14_MYCGA 1119 KRIQMNQK 1126

```

5. Putative MutH analogue (gene name *MGA_0195*)

Gene is absent in *B. subtilis* (strain 168) genome.

```

>sp|P06722|MUTH_ECOLI DNA mismatch repair protein MutH OS=Escherichia coli
(strain K12) GN=mutH PE=1 SV=3
MSQPRPLLSPPETEEQLLAQAQQLSGYTLGELAAALVGLVTPENLKRDKGWIGVLEIWLG
ASAGSKPEQDFAALGVELKTIPVDSLGRPLETTFVCVAPLTGNSGVTWETSHVRHKLKRV
LWIPVEGERSIPLAQRVVGSPLLWSPNEEEDRQLREDWEEELMDMIVLGQVERITARHGEY
LQIRPKAANAKALTEAIGARGERILTLPGRFYLKKNFTSALLARHFLIQ
>tr|Q7NAY4|Q7NAY4_MYCGA Uncharacterized protein OS=Mycoplasma gallisepticum
(strain R(low / passage 15 / clone 2)) GN=MYCGA5010 PE=4 SV=1
MLINYISFIKHGFYIIILINKNMPFKKDLNLSDCFKIVDNQLVLSEKYKQEYGSKFKKIT
GSRFPNVLGINEFNTPFIEWLKMVNLYYETMDPILSKAGVVIEPKVREYIMNKFKINYKT
YDPIKVGFDLFDKDNQIFGGIPDGEVPDQDGNLLYDENHPMLEIKTTSIDKLSYKKVDGLL
RMMVDQSGLPPIVAKAKREKYFEWYDENKQIKVKKEYVLQLSLYLYLRNAKYGRFGVIFLRP
EDYQNPFAIDLSQRLIDVVDKVEKSSIEPYIEQATQWYQDHI IKGISPKMTRSDFLEFLK
LHKII

```

```

p|P06722|MUTH_ECOLI -----MSQPRPLLSPPETEE 15
tr|Q7NAY4|Q7NAY4_MYCGA MLINYISFIKHGFYIIILINKNMPFKKDLNLSDCFKIVDNQLVLSEKYKQ 50
. *: . : : :

sp|P06722|MUTH_ECOLI QLLAQAQQLSGYTLGELAAALVGLVTP--ENLKRDKGWIGVLEIWLGASA 63
tr|Q7NAY4|Q7NAY4_MYCGA EYGSKFKKITGSRFPNVLGINEFNTPFIEWLKMVNLYYETMDPILSKAGV 100
: : : : : * : : : : : * * * * : : : * * ..

sp|P06722|MUTH_ECOLI GSKP---EQDFAALGVELKTIPVDSLGRPLETTFVCVA-----PLTGNS 104
tr|Q7NAY4|Q7NAY4_MYCGA VIEPKVREYIMNKFKINYKYDPIKVGFDLFDKDNQIFGGIPDGEVPDQDQ 150
:* * : : : : * * .:* * . . . * : :

sp|P06722|MUTH_ECOLI GVTWETSHVRHKLKRVLWIPVEGERS---IPLAQRVVGSPLLWSPNE--- 148
tr|Q7NAY4|Q7NAY4_MYCGA NLLYDENHPMLEIKTTSIDKLSYKKVDGLLRMMVDQSGLPPIVAKAKREKYF 200
.: : : .* : : * . : : : : : : : : : * * : : : *

sp|P06722|MUTH_ECOLI ---EEDRQLREDWEEELMDMIVLGQVERITARHGEYLQIRPKAAN----- 189
tr|Q7NAY4|Q7NAY4_MYCGA EWYDENKQIKVKKEYVLQLSLYLYLRNAKYGRFGVIFLRPEQNPFAID 250
: * : * : : * : : : : : : : : : : : * : : : :

sp|P06722|MUTH_ECOLI -AKALTEAIGARGERILTLP-----RGFYLKKNFTSALLAR 224
tr|Q7NAY4|Q7NAY4_MYCGA LSQRLIDVVDKVEKSSIEPYIEQATQWYQDHI IKGISPKMTRSDFLEFLK 300
: : * . : . : * * : : * * : : : :

sp|P06722|MUTH_ECOLI HFLIQ 229
tr|Q7NAY4|Q7NAY4_MYCGA LHKII 305
. *

```

6. Excinuclease ABC subunit A (gene name *uvrA*)

```

>sp|P0A698|UVRA_ECOLI UvrABC system protein A OS=Escherichia coli (strain
K12) GN=uvrA PE=1 SV=1

```

MDKIEVRGARTHNLKININLVIIPRDKLIIVVTGLSGSGKSSLAFDFTLYAEGQRRYVESLSAY
 ARQFLSLMEKPDVDHIEGLSPAISIEQKSTSHNPRSTVGTITEIHDYLRLLLFARVGEPRC
 PDHDVPLAAQTVSQMVDNVLSPQEGKRLMLLAPIIKERKGEHTKTLENLASQGYIRARID
 GEVCDLSDPPKLELQKKHTIEVVDRFKVRDDLTQRLAESFETALELSGGTAVVADMDDP
 KAEELLSANFACPICGYSMRELEPRLFSFNNPAGACPTCDGLGVQQYFDPDRVIQNP
 SLAGGAIIRGWDRRNFFYFQMLKSLADHYKFDVEAPWGSLSANVHKVVLYGSGKENIEFKY
 MDRGDTSIIRHPFEGVLHNMERRYKETEESAVREELAKFISNRPCASCEGTRLRREARH
 VYVENTPLPAISDMSIGHAMEFFNNLKLQAGQRAKIAEKILKEIGDRLKFLVNVGLNYLTL
 SRSAETLSGGEAQRIRLASQIGAGLVGMVYLDEPSIGLHQDNERLLGTLIHLRDLGNT
 VIVVEHDEDAIRAADHVIDIGPGAGVHGGEVVAEGPLEAIMAVPESLTGQYMSGKRKIEV
 PKKRV PANPEKVLKLTGARGNNLKDVTTLTPVGLFTCITGVSGSGKSTLINDTLFPIAQR
 QLNATIAEPAPYRDIQGLEHFDKVIDIDQSPIGRTPRSNPATYTGVTFTVRELFAGVPE
 SRARGYTPGRFSFNVRGGRCEACQGDGVIKVMHFPLDIYVPCDQCKGKRYNRETLEIKY
 KGKTIHEVLDMTIEEAREFFDAVPALARKLQTLMDVGLTYIRLQGSATTLSSGGEAQRV
 ARELSKRGTGQTLYLDEPTTGLHFADIQQLLDVHLKLRDQNTIVVIEHNLDVIKTADW
 IVDLGPPEGSGGGEILVSGTPEVVAECEASHTARFLK PML

>sp|O34863|UVRA_BACSU UvrABC system protein A OS=Bacillus subtilis (strain 168) GN=uvrA PE=3 SV=1

MAMDRIEVKGARAHNLKINIDVTIPRDQLVVTGLSGSGKSSLAFDTIYAEGQRRYVESLS
 AYARQFLGQMDKPDVDAIEGLSPAISIDQKTSRNPRSTVGTVTEIYDYLRLLYARVGP
 HCPEHGIEITSQTIEQMVDRIEYPERTKLQVLAPIVSGRKGAVKQVLEQIRKQGYVRVR
 IDGEMAELSDDIELEKNKKHSIEVVIDRIVVKEGVAARLSDSLETALRLGEGRVMIDVIG
 EEELMFSEHHACPHCGFSIGELEPRLFSFNSPFGACPTCDGLGMKLEVDADLVI PNQDLS
 LKENAVAPWTPISSQYYPQLLEAVCTHYGIDMDVPVKDLPKHQLDKVLVYSGDDLIYFRY
 ENDFGQIREGEIQFEGVLRNIEERRYKETGSDFIREQMEQYMSQKSCPTCKGYRLKKEALA
 VLIDGRHIGKITELSVADALAFFKDLTLSEKDMQIANLILREIVERLSFLDKVGLDYLTTL
 SRAAGTLSGGEAQRIRLATQIGSRLSGVLYILDEPSIGLHQDNDRLISALKNMRLGNT
 LIVVEHDEDTMMAADYLIDIGPGAGIHGGQVISAGTPEEVMEPNLSLTGSYLSGKFIPL
 PPERRKPDGRYIEIKGASENNLKKVNAKFPPLGTF TAVTGVSGSGKSTLVNEILHKALAQK
 LHKAKAKPGSHKEIKGLDHLKVIDIDQAPIGRTPRSNPATYTGVTDDIRVFAQTNEAK
 VRGYKGRFSFNVKGGRCEACRGDGIKIKEMHFPLDVYVPCVCHGKRYNRETLEVTYK
 KSIDVLDMTVEDALSFFENIPKIKRKLQTLVDVGLGYITLQGPATTLSSGGEAQRVKLAS
 ELHKRSTGRTLYILDEPTTGLHVDDIARLLVVLQRLVDNGDVTLVIEHNLDI IKTADYIV
 DLGPEGGAGGGTIVASGTPEEITEVEESYTG RYLKPVIERDKTRMKSLLKAKETATS

>tr|Q7NC22|Q7NC22_MYCGA UvrABC system protein A OS=Mycoplasma gallisepticum (strain R(low / passage 15 / clone 2)) GN=uvrA PE=3 SV=2

MKKIKKDSANYISIVGARQNNLKNISLDIPKNQLVVTGLSGSGKSSLAFKTIYAEGQRR
 YLESLSPYARQFLGNNDKPDVDSIEGLSPSISIDQKSTSHNPRSTVGTVTEVYDFLRLLLW
 SRVGDAYCINGHMIKT'TTIKQIIDHVLELEDDSKLQILAPVIKQKGTFKNEFEKFKYKQ
 GFMRVLVDGVVYSLDDKIELDKNQKHDISIVIDRLIILNKDNQTKLRITDAIETALTVSNG
 LIQIISNDQAKYEFSLNHSCDQCFFIPELEPRLFSFNSPIGACDYCKGLGFTYEPDVK
 IIPNKDLTINEGAIDYFKNRINTSSQDWQRFYSIRHYQIDLNTPIKNLSKKEINYLLEG
 SDEPIEIVIESANRNGISSRLDYVEGIAKLIQRRHLETKSSAARDNYSKYTSEQKCKTCD
 GKKLSPAALSVKIGGLDIEFTNLNVNKALDFILGLEFNEEKTKIAKFLKEILDRLYFL
 VNVGLEYLTLNRNASTLSGGESQRIRLATQIGSRLSGVLYVLDEPSIGLHQDNDKLIK
 LLSMRDLGNSLIVVEHDEETMMSADYLIDIGPGAGTYGGKVVAAGTVEEVMKNPASLTGQ
 YLSKKLEIEQPKKLHPGNGQKIVLKGASANNLKNINVEFPLNKLVVVTGVSGSGKSTLIN
 QTLVNGIEKALFNKHVEVGKYKSLIGINNIDKVIKVSQDPIGRTPRSNPATYVSVFDDIR
 EVFANVFEAKARGYTKSRFSFNVSGGRCDCCQGDGKCIEMHFPLDVYVVKCSCSSCKKYN
 EATLEIKYKNKSIYDVLEMSCEEALEFFKVIIPAINRKLQLMCDVGLGYMKLSTNATELSG
 GEARIKLAKYLQRKATGNTIYVLDEPTTGLHAHDIKKLLSVLNRLVDNGDSVIVIEHNLD
 ELIKVADHIDLGPNGGDDGGYLICAGTPQELVKNYTDSSYTARYLAKIMKS

sp P0A698 UVRA_ECOLI	-----MDKIEVRGARTHNLKININLVIIPRDKLIIVVTGLSGSG	SSLAF	42
sp O34863 UVRA_BACSU	-----MAMDRIEVKGARAHNLKINIDVTIPRDQLVVTGLSGSG	SSLAF	44
tr Q7NC22 Q7NC22_MYCGA	MKKIKKDSANYISIVGARQNNLKNISLDIPKNQLVVTGLSGSG	SSLAF	50
	: *.: ** :*****.: **::*.:*****		
sp P0A698 UVRA_ECOLI	DTLYAEGQRRYVESLSAYARQFLSLMEKPDVDHIEGLSPAISIEQKSTSH		92
sp O34863 UVRA_BACSU	DTIYAEGQRRYVESLSAYARQFLGQMDKPDVDAIEGLSPAISIDQKTSR		94
tr Q7NC22 Q7NC22_MYCGA	KTIYAEGQRRYVESLSAYARQFLGNNDKPDVDSIEGLSPSISIDQKSTSH		100
	.*:*****:****.*****. :***** *****:***:*.**:		

```

sp|P0A698|UVRA_ECOLI      NPRSTVGTITEIHDYLRLLLFARVGEPRCPDHDVPLAAQTVSQMVDNVLSQ 142
sp|O34863|UVRA_BACSU      NPRSTVGTVTEYDYDLRLLLYARVKGPHCEHGIEITSQTIEQMVDRILEY 144
tr|Q7NC22|Q7NC22_MYCGA    NPRSTVGTVTEVYDFLRLLSRVGDAYCINGHGMIKTTTIKQIIDHVLLEL 150
*****:***:***:***:***. * : : *.:***:*.

sp|P0A698|UVRA_ECOLI      PEGKRLMLLAPIIKERKGEHTKTLENLASQGYIRARIDGVECDLSDPPKL 192
sp|O34863|UVRA_BACSU      PERTKLQVLAPIVSGRKGAVKVLQIRKQGYVVRVIDGEMAELSDDIEL 194
tr|Q7NC22|Q7NC22_MYCGA    EDDSKLQILAPVIKLGKTFKNEFEKFKYKQGFMRVLVDGVVYSLDDKIEL 200
: .:* :***:. :** . : :*: .**:* .** : .* * :*

sp|P0A698|UVRA_ECOLI      ELQKKHTIEVVVDRFKVRDD--LTQRLAESFETALELSGGTAVVADMDDP 240
sp|O34863|UVRA_BACSU      EKNKKHSIEVVVIDRIVVKEG--VAARLSDSLETALRLGEGRVVIDVIGE- 241
tr|Q7NC22|Q7NC22_MYCGA    DKNQKHDISIVIDRLILNKDNQTKLRITDAIETALTVSNGLIQIISNDQ- 249
: :.* *.:**:* :... *:::*** . * : :

sp|P0A698|UVRA_ECOLI      KAEELLFSANFACPICGYSMRELEPRLFSFNPPAGACPTCDGLGVQYFD 290
sp|O34863|UVRA_BACSU      --EELMFSEHHACPHCGFSIGELEPRLFSFNPPGACPTCDGLGMKLEVD 289
tr|Q7NC22|Q7NC22_MYCGA    --AKYEFSLNHSDCQCGFFIPELEPRLFSFNPIGACDYCKGLGFTYEPD 297
: * * :.* ** : *****.* *** *.*. *

sp|P0A698|UVRA_ECOLI      PDRVIQNPESLAGGAIARGWD--RRNFYY--FQMLKSLADHYKFDVEAPWG 337
sp|O34863|UVRA_BACSU      ADLVI PNQDLSLKENAVAPWTPISSQYY--PQLEAVCTHYGIDMDVPVK 337
tr|Q7NC22|Q7NC22_MYCGA    VDKIIPNKDLTINEGADYFKNRINTSSQDWQRFYSIIRHYQIDLNTPIK 347
* * * *:*: .*: : . * : : * * :*.:*

sp|P0A698|UVRA_ECOLI      SLSANVHKVVLYGSGKENIEFKYMNDRGDTSIRRHFPFEGVLHNMERRYKE 387
sp|O34863|UVRA_BACSU      DLPKHQLDKVLYGSGDLLIYFRYENDFGQIREGEIQFEGVLRNIERRYKE 387
tr|Q7NC22|Q7NC22_MYCGA    NLSKKEINYLLEGSDEPIEIVIESANRNGISSRLDYVEGIAKLIQRHLE 397
.* : . : * * * . . . . . * * : : : * * :

sp|P0A698|UVRA_ECOLI      TESSAVREELAKFISNRPCASCEGTRLRREARHVYVENTPLPAISDMSIG 437
sp|O34863|UVRA_BACSU      TGSDFIREQMEQYMSQKSCPTCKGYRLKKEALAVLIDGRHIGKITELVA 437
tr|Q7NC22|Q7NC22_MYCGA    TKSSAARDNYSKYTSEQKCTCDGKKSIPAALSVKIGGLDIEFTNLNVN 447
* * . * : : : * : * * * * * * : . : : : :

sp|P0A698|UVRA_ECOLI      HAMEFFNNLKLQGRAKIAEKILKEIGDRLKFLVNVGLNYLTLRSRAETL 487
sp|O34863|UVRA_BACSU      DALAFFKDLTLEKDMQIANLILREIVERLSFLDKVGLDYLTLRAAGTL 487
tr|Q7NC22|Q7NC22_MYCGA    KALDFILGLEFNEKTKIAKFVLKEILDRLYFLVNVGLEYLTLRSNASTL 497
.* : * : * : : : * : * : * * * * * * * * * * * *

sp|P0A698|UVRA_ECOLI      SGGEAQRIRLASQIGAGLVGMVYVLDPEPSIGLHQRDNERLLGLTLHLRDL 537
sp|O34863|UVRA_BACSU      SGGEAQRIRLATQIGSRLSGVLYILDEPSIGLHQRDNDRLISALKNMARDL 537
tr|Q7NC22|Q7NC22_MYCGA    SGGESQRIRLATQIGSRLSGVLYVLDPEPSIGLHQKDNDKLIKTLMSRDL 547
*****:***: * * * :*****:***:*. : * :***

sp|P0A698|UVRA_ECOLI      GNTVIVVEHDEDAIRAADHVIDIGPGAGVHGGEVVAEGPLEAIMAVPESL 587
sp|O34863|UVRA_BACSU      GNTLIVVEHDEDTMAADYLDIDIGPGAGIHGGQVISAGTPEEVMEVDPNSL 587
tr|Q7NC22|Q7NC22_MYCGA    GNSLIVVEHDEETMSADYLDIDIGPGAGYGGKVVAAGTVEEVMMKNPASL 597
*.:***:***: : : * : * : * : * : * : * . * * * * *

sp|P0A698|UVRA_ECOLI      TGQYMSGKRKIEVPPKRVANPEKVLKLTGARGNNLKDVTTLTPVGLFTC 637
sp|O34863|UVRA_BACSU      TGSYLSGKFKFIPPPERKPD--GRYIEIKGASENNLKKVNAKFPGLGFTA 636
tr|Q7NC22|Q7NC22_MYCGA    TGQYLSKLEIEQPKKLHPGN--GQKIVLKGASANNLKNINVEFPLNKLIV 646
*.*:* * * * * : : : : * * * * . . . * : . .

sp|P0A698|UVRA_ECOLI      ITGVSGSGLSTLINDTLFPFIAQRQLNGATIAEPAPYRDIQGLEHFDKVID 687
sp|O34863|UVRA_BACSU      VTGVSGSGLSTLVNEILHKALAKLH--KAKAKPGSHKEIKGLDHLDKVID 685
tr|Q7NC22|Q7NC22_MYCGA    VTGVSGSGLSTLINQTLVNGIEKALF--NKHVEVGKYSLIGINNIDKVIK 695
:*****:***: * : * : . : : : : * : : : *

sp|P0A698|UVRA_ECOLI      IDQSPIGRTPRSNPATYTGVFVTPVRELFAGVPESRARGYTPGRFSFNVRG 737
sp|O34863|UVRA_BACSU      IDQAPIGRTPRSNPATYTGVFDDIRDVFAQTNEAKVRGYKKGRFSFNVKG 735
tr|Q7NC22|Q7NC22_MYCGA    VSQDPIGRTPRSNPATYVSFDDIREVFANVFEAKARGYTKSRFSFNVSG 745
.* *****.*** :*: * . *.:***. ***** *

sp|P0A698|UVRA_ECOLI      GRCEACQGDGVIVKEMHFLPDIYVPCDQCKGKRYNRETLEIKYKGTIHE 787
sp|O34863|UVRA_BACSU      GRCEACRGDGIKIEMHFLPDVYVPCVCHGKRYNRETLEVYKGSISD 785
tr|Q7NC22|Q7NC22_MYCGA    GRCDDCQGDGVKCIEMHFLPDVYVKCSSCNGKYNNEATLEIKYKNKSIYD 795
* * : * : * : * * * * * * * * * * * * * * *

sp|P0A698|UVRA_ECOLI      VLDMTIEEAREFFDAVPALARKLQTLMDVGLTYIRLQGSATTLSGGEAQR 837
sp|O34863|UVRA_BACSU      VLDMTVEDALSFFENIPKIKRKLQTLYDVGLGYITLQGPATTLSGGEAQR 835
tr|Q7NC22|Q7NC22_MYCGA    VLEMSCEEALEFFKVIIPAINRKLQLMCDVGLGMYKMLSTNATELSGGEAQR 845
*.:* : * * . * : * * * : * * * * * * * * * * * *

sp|P0A698|UVRA_ECOLI      VKLARELSKRGTGQTLYLDEPTTGLHFADIQQLLDVLHKLDRDQNTIVV 887
sp|O34863|UVRA_BACSU      VKLASELHKRSTGRTLYLDEPTTGLHVDDIARLLVVLQRLVDNGDVLV 885

```

```
tr|Q7NC22|Q7NC22_MYCGA      IKLAKYLQRKATGNTIYVLDEPTTGLHAHDIKLLSVLNRLVDNGDSVIV 895
:*** * :.:.**.*:*:***** ** :** **::: *:*:::

sp|P0A698|UVRA_ECOLI        IEHNLDVIKTADWIVDLGPEGGSGGGEILVSGTPETVAE--CEASHTARF 935
sp|O34863|UVRA_BACSU        IEHNLDIIKTADYIVDLGPEGGAGGGTIVASGTPEEITE--VEESYTGRY 933
tr|Q7NC22|Q7NC22_MYCGA      IEHNLELIKVADHIIDLGPNGGDDGGYLICAGTPQELVKNYTDSSTARY 945
*****:*.**.*:*:****:**.* **:: :***: :. : *:*.*:

sp|P0A698|UVRA_ECOLI        LKPML----- 940
sp|O34863|UVRA_BACSU        LKPVIERDKTRMKSLKAKETATS 957
tr|Q7NC22|Q7NC22_MYCGA      LAKIMKS----- 952
*  ::
```

Organism	ATP hydrolysis*
E. coli (strain K12)	K37, K646
B. subtilis (strain 168)	No information
M. gallisepticum (strain R(low / passage 15 / clone 2))	No information

*Site-specific mutagenesis of conserved residues within Walker A and B sequences of Escherichia coli UvrA protein. (<http://pubs.acs.org/doi/pdf/10.1021/bi00230a004>)

7. Excinuclease ABC subunit B (gene name *uvrB*)

```
>sp|P0A8F8|UVRB_ECOLI UvrABC system protein B OS=Escherichia coli (strain
K12) GN=uvrB PE=1 SV=2
MSKPFKLNSAFKPSGDQPEAIRRLEEGLDGLAHQTLGVTGSGKTFITIANVIADLQRPT
MVLAPNKTLAAQLYGEMKEFFPENAVEYFVSYDYQPEAYVPSSDTFIEKDASVNEHIE
QMRLSATKAMLERRDVVVVASVSAIYGLGDPDLYLKMLHLTVGMIIDQRAILRRLAELQ
YARNDQAFQRGTFRVRGEVIDIFPAESDDIALRVELFDEEVERLSLFDPLTGQIVSTIPR
FTIYPKTHYVTPRERIVQAMEEIKEELAARRKVLENNKLLLEEQRLTQRTQFDLEMMNEL
GYCSGIENYSRFLSGRGPGEPPPTLFDYLPADGLLVVDESHVTIPQIGMYRGDRARKET
LVEYGFRLPSALDNRPLKFEEFEALAPQTIYVSATPGNYELEKSGGDVVDQVVRPTGLLD
PIIEVRPVATQVDDLLSEIRQAAINERVLVTTLTKRMAEDLTEYLEEHGERVRYLHSDI
DTVERMEIIRDRLRGEFDVLVGINLLREGLDMPEVSLVAILDADKEGFLRSERSLIQTIG
RAARNVNGKAILYGDKITPSMAKAIGETERREKQQKYNEEHGITPQGLNKKVVDILALG
QNIAKTKAKGRGKSRPIVEPDNVPMDMSPKALQQKIHELEGLMMQHAQNLEFEEAAQIRD
QLHQLRELFIAAS
>sp|P37954|UVRB_BACSU UvrABC system protein B OS=Bacillus subtilis (strain
168) GN=uvrB PE=1 SV=2
MKDRFELVSKYQPQGDQPKAIEKLVKGIQEGKKHQTLLGATGTGKTFITVSNLIKEVNKPT
LVIAHNKTLAAGLYSEFKEFFPNNAVEYFVSYDYQPEAYVPQDTDFIEKDASINDEID
KLRHSATSALFERRDVI IASVSCIYGLGSPEEYREMVVSLRTEMEIERNELLRKLVDIQ
YARNDIDFQRGTFRVRGDVVEIFPASRDEHCVRVEFFGDEIERIREVDALTGEILGDRDH
VAIFPASHFVTRAEKMEKAIQNIKEKELEEQLKVMHENGKLEAQRLEQRTRYDLEMMREM
GFCSGIENYSRHLTLRPPGSTPYTLDDYFPDDFMIVVDESHVTIPQVRGMFNGDQARKQV
LVDHGFRLPSALDNRPLRFEEFEKHMHNIVYVSATPGPYEIEHTDEMVEQIIRPTGLLDP
LIDVRPIEGQIDDLIGEIQARIERNERVLVTTLTKKMSEDLTDYLKEIGIKVNYLHSEIK
TLERIEIIRDRLRGLKYDVLVGINLLREGLDIPEVSLVAILDADKEGFLRSERSLIQTIGR
AARNAEGRVIMYADKITSMEIAINETKRREQQERFNEEHGITPKTINKEIRDVIRATV
AAEDKAEYKTKAAPKLSKMTKKERQKVVEQMEHEMKEAAKALDFERAAELRDLLLLKAE
G
>tr|Q7NC43|Q7NC43_MYCGA UvrABC system protein B OS=Mycoplasma gallisepticum
(strain R(low / passage 15 / clone 2)) GN=uvrB PE=3 SV=2
MAEKKFKLVSKNKPAGHQPEAIKKLVDGINKNKKYQTLGATGTGKTFITIANVIEKTQKK
TLILAHNKTLAAQLYLEFKELFPNNAVEYFVSYFDYQPEAYIPRTDMYIEKSSVTNDEI
EMLRLASLNSLSTRNDVIVVASVACIYPAANPEDFDIYRIILKVGNTLKLSDLKENLIRL
NYARSPENCNEPGTFRIKGDVVDIFPGYVSDHIIRLSFFGDELEEIRKIHPDSSVIEKYT
SYVLGPANEYILNFERKDTAIIKRIQEELMFRVQEFKNQQKLVEAQRQQRTEYDIDAIE
FGFCNGIENYAFHLELREKGSTPWTLFDFFGDDWLMVIDESHISVPQVKGMFNTDKSRKT
TLVEYGFRLPSALENRPLNYDEFSNKSQVIFVSATPNDEEIKLSNNEIEQIVRPTGLL
DPTVEIRPRLDQINDLMNELKKQKDKNERFTITVTTIKMAEDLTEYLKERNFKCAYIHNE
```

LKTLERSLILNDLRRGKYDCVVGINLLREGLDIPVSLVCFDADKPGYFRSDKALIQTI
 GRAARNQNGRVMYADEMTKAMKIAVDETNRRRKIQEKFNKDHKITPKTIKPIYDDLKN
 KASHKQIEEVMRKTAKGDKFIKMIEDLRNEMLEAAKNQNYEHAASLRDLIELETQQLS
 KTNK

```

sp|P0A8F8|UVRB_ECOLI      -MSKPFKLNSAFKPSGDQPEAIRRLEEGLEDGLAHQTLTGVTGSGKTFTI 49
sp|P37954|UVRB_BACSU      -MKDRFELVSKYQPQGDQPKAIEKLVKGIQEGKKHQTLLGATGTGKTFTV 49
tr|Q7NC43|Q7NC43_MYCGA    MAEKKFKLVSKNKPAGHQPEAIKKLVGDINKNKYQTLGATGTGKTFTI 50
      ..*:* * : * * **:*:* . * : * : . . : * * * * * . * * * * * :

sp|P0A8F8|UVRB_ECOLI      ANVIADLQRPTMVLAPNKTLLAAQLYGEMKEFFPENAVEYFVSYDYDQPE 99
sp|P37954|UVRB_BACSU      SNLIKEVKNKPTLVIAHNKTLAGQLYSEFKEFFPNNAVEYFVSYDYDQPE 99
tr|Q7NC43|Q7NC43_MYCGA    ANVIEKTQKKTLLIAHNKTLAAQLYLEFKELFPNNAVEYFVSYDFDYQPE 100
      : * : * . : : * : : * * * * * . * * * * * : * * * * * : * * * * *

sp|P0A8F8|UVRB_ECOLI      AYVPSSDTFIEKDASVNEHIEQMRLSATKAMLERRDVVVVASVSAIYGLG 149
sp|P37954|UVRB_BACSU      AYVPQTDTFIEKDasINDEIDKLRHSATSALFERRDVIIIASVSCIYGLG 149
tr|Q7NC43|Q7NC43_MYCGA    AYIPRTDMYIEKSSVTNDEIEMLRASLSLSSTRNDVIVVASVACIYPAA 150
      ** : * : * : * * * . : * . : * : * : : : * . * * : : * * * . *

sp|P0A8F8|UVRB_ECOLI      DDPDLYLKMMHLTLVGMIIQRAILRRLAELQYARNDQAFQRTGTRVVRGEV 199
sp|P37954|UVRB_BACSU      SPEEYREMVVSLRTEMEIERNELLRKLVDIQYARNDIDFQRTGTRVVRGDV 199
tr|Q7NC43|Q7NC43_MYCGA    NPEDFDIYRIILKVGNTLKLSDLKENLIRLNYARSPECNEPGTFRIKGDV 200
      . * : : : * . . : . : * : * * * . : * * * : : * *

sp|P0A8F8|UVRB_ECOLI      IDIFPAESDDIALRVELFDEEVERLSLFDPLTGQIVSTIPRFTIYPKTHY 249
sp|P37954|UVRB_BACSU      VEIFPASRDEHCVRVEFFGDEIERIREVDALTEILGDRDHVAIFPASHF 249
tr|Q7NC43|Q7NC43_MYCGA    VDIFPGYVSDHIIRLSFFGDELEBEIRKIHPDSSVIEKYTSYVLGPANEY 250
      : * * * . : : * : : * : * * : * . . . . : : : * * * : :

sp|P0A8F8|UVRB_ECOLI      VTPRERIVQAMEEIKEEELAARRKVLLENKLLLEEQLRTQRTQFDLEMMNE 299
sp|P37954|UVRB_BACSU      VTRAEKMEKAIQNIKEELEEQKVMHENGKLLAQRLQRTQRTQFDLEMMRE 299
tr|Q7NC43|Q7NC43_MYCGA    ILNFERKDTAIRKIQEELMFRVQEFKNQKQKLVQAEQRLQRTQRTQFDLE 300
      : * : * * : : * * * * * : : : : * * * * * * * * . : : *

sp|P0A8F8|UVRB_ECOLI      LGYCSGIENYSRFLSGRGPGEPPPTLFDYLPADGLLVVDESHVITPQIGG 349
sp|P37954|UVRB_BACSU      MGFCSGIENYSRHLLTLPPEPSTPYTLDDYFPDDFMIVVDESHVITPQVRG 349
tr|Q7NC43|Q7NC43_MYCGA    FGFCSGIENYAFHLELREKGSTPWTLLDFDFGDDWLMVIDESHISVPQVKG 350
      : * : * . * * * * : * * * * * : * : * * * * * : : * * *

sp|P0A8F8|UVRB_ECOLI      MYRGDRARKETLVEYGFRLPSALDNRPLKFEEFEALAPQTIYVSATPGNY 399
sp|P37954|UVRB_BACSU      MFNGDQARKQVLDHGFRLPSALDNRPLRFEEFEKHMHNIVYVSATPGPY 399
tr|Q7NC43|Q7NC43_MYCGA    MFNTDKSRKTTLVEYGFRLPSALENRPLNYDEFNSKSDQVIFVSATPNDE 400
      * . : * : * * . * : * * * * * * * * * * * * . : : * * * * *

sp|P0A8F8|UVRB_ECOLI      ELEKSGGDVVDQVVRPTGLLDPIIEVRPVATQVDDLLSEIRQRAAINERV 449
sp|P37954|UVRB_BACSU      ETEHT-DEMVEQIIRPTGLLDPLIDVRIEIQIDDLIGEIQARIERNERV 448
tr|Q7NC43|Q7NC43_MYCGA    EIKLSNNEIEIQIVRPTGLLDPTVEIRPRLDQINDLMNELKKQKDKNERT 450
      * : : : . : : * : * * * * * * * * * * * * * * * * * * .

sp|P0A8F8|UVRB_ECOLI      LVTTLTKRMAEDLTEYLEEHGERVRYLHSDIDTVERMEIIRDRLRGEFDV 499
sp|P37954|UVRB_BACSU      LVTTLTKKMSEDLTDYLKEIGIKVNYLHSEIKTLERIEIIRDRLRKYD 498
tr|Q7NC43|Q7NC43_MYCGA    FTTVTTIKMAEDLTEYELKERNFKCAYIHNELKTLERSLILNDLRRGKYDC 500
      : * . * : * : * * * * * * * . : * : * . : * : * * * * : * : *

sp|P0A8F8|UVRB_ECOLI      LVGINLLREGLDPEVSLVAILDADKEGFLRSERSLIQTI GRAARNVNGK 549
sp|P37954|UVRB_BACSU      LVGINLLREGLDPEVSLVAILDADKEGFLRSERSLIQTI GRAARNAEGR 548
tr|Q7NC43|Q7NC43_MYCGA    VVGINLLREGLDPEVSLVCFDADKPGYFRSDKALIQTI GRAARNQNGR 550
      : * * * * * * * * * * * * * * * * * * * * * * * * * * * * * :

sp|P0A8F8|UVRB_ECOLI      AILYGDKITPSMAKAIGETERRREKQKQYNEEHGITPQGLNKKVVDILAL 599
sp|P37954|UVRB_BACSU      VIMYADKITKSMIAINETKRRREQERFNEEHGITPKTINKEIRDVIRA 598
tr|Q7NC43|Q7NC43_MYCGA    VIMYADEMTKAMKIAVDETNRRRKIQEKFNKDHKITPKTIKPIYDDLKN 600
      . * : * * : * * * * * * * * * * * * * * * * * * * * * :

sp|P0A8F8|UVRB_ECOLI      GQNIAKTKAKGRGKSRPIVEPDNVPMDMSPKALQKQIHELEGLMMQHAQN 649
sp|P37954|UVRB_BACSU      -TVAAEKAEYKTKAAPKLS-----KMTKKERQKVVEQMEHEMKEAAGA 641
tr|Q7NC43|Q7NC43_MYCGA    -KASHKQIEEVMRKTAKGD-----KFIKMIEDLRNEMLEAAKN 638
      : : * : . . : : : : . . * : * :

sp|P0A8F8|UVRB_ECOLI      LEFEEAAQIRDQLHLRELFIAS-- 673
sp|P37954|UVRB_BACSU      LDFERAAELRDLLLELKAEG----- 661
tr|Q7NC43|Q7NC43_MYCGA    QNYEHAASLRDLIELETQQLS KTNK 664
      : * . * * . * * : * : .
    
```

Organism	Nucleotide binding (ATP) - potential	Motif (Beta-hairpin)	Domain (Helicase ATP-binding, Helicase C-terminal, UVR)
E. coli (strain K12)	39-46	92-115	26-415, 431-597, 633-668
B. subtilis (strain 168)	39-46	92-115	26-413, 430-596, 625-660
M. gallisepticum (strain R(low / passage 15 / clone 2))	40-47	93-116	27-184, 432-594, 622-657

8. Excinuclease ABC subunit C (gene name *uvrC*)

>sp|P0A8G0|UVRC_ECOLI UvrABC system protein C OS=Escherichia coli (strain K12) GN=uvrC PE=1 SV=1

```
MSDQFDAKAFKTVTSQPGVYRMYDAGGTVIYVGKAKDLKKRLSSYFRSNLASRKTEALV
AQIQQIDVTVTHTTETEALLLEHNYIKLYQPRYNVLLRDDKSYPFIFLSGDTHPRLAMHRG
AKHAKGEYFGPFNGYAVRETLALLQKIFPIRQCENSVYRNRSRCLQYQIGRCLGPCVE
GLVSEEEYAQQVEYVRLFLSGKDDQVLTQLISRMETASQNFEEAARIRDQIQAVRRVT
EKQFVSNTGDDLDVIGVAFDAGMACVHVLFIRQGKVLGSRSYFPKVPGGTELSEVVETV
GQFYLQGSQMRTLPGEILLDFNLSDKTLLADSLSELAGRKINVQTKPRGDRARYLKLART
NAATALTSKLSQQSTVHQRLTALASVLKLEPKRMECFDISHTMGEQTVASCVVDANGP
LRAEYRRYNTGITPGDDYAAMNQVLRRLRYGKAIDDSKIPDVILIDGGKGLAQAKNVFA
ELDVSWDKNHPLLLGVAKGADRKAGLETLFFFEPEGEGFSLPPDSPALHVIQHIRDESHDH
AIGGHRKKRAKVNTSSLETIEGVGPKRRQMLLKYMGGLQGLRNASVEEIAKVPGISQGL
AEKIFWSLKH
```

>sp|P14951|UVRC_BACSU UvrABC system protein C OS=Bacillus subtilis (strain 168) GN=uvrC PE=3 SV=2

```
MNKQLKEKLALLPDQPGCYLMKDRQQTVIYVGKAKVLKNRVRSYFTGSHDAKTQRLVTEI
EDFEYIVTSSNLEALILEMNLIKKHDPKYNVMLKDDKTYPFIKLTHERHPRILIVTRNVKK
DKGRYFGPYPNVQAARETKKLLDRLYPLRKC SKLPDRVCLYHGLQCLAPCVKDISETN
RELVESITRFLRGGYNEVKKELEEKMHAAENLEFERAKELRDQIAHIESTMEKQKMTMN
DLVDRDVFAYAYDKGWMCVQVFFIRQKLIERDVSMFPLYQEADDEEFLTFIQGFYSKNNH
FLPKEILVPSIDQSMIEQLLETNVHQPKKGPKELEMLLAHKNAKIALKEKFSlierdee
RSIGAVQKLGALNIYTPHRIEAFDNSNIQGTNPVSAMIVFIDGKPYKKEYRKYKIKTVT
GPDDYGSMREVVRRRYTRVLRENLPPLDLIIIDGGKGINAARDVIENELGLDPIAGLA
KDEKHRTSNLLIGDPLEVAYLERNSSQEFYLLQRIQDEVHRFAISFHRQIRGKSAFQSVLD
DIPGIGEKRKMLLKHFGSVKMKKEASLEDIKKAGVPAAAAQLLYDKLQK
```

>sp|Q7NBC4|UVRC_MYCGA UvrABC system protein C OS=Mycoplasma gallisepticum (strain R(low / passage 15 / clone 2)) GN=uvrC PE=3 SV=2

```
MNFNLKQKLDLAPKKPGCYLWKNHLNEIYIGKAKNIYKRVHQYFNGPKDLKTSKLVNEI
FYVEFIEVNNENEALLLEANLIKHKPRYNILLKDNNGYPYILMTKEKYPRLIYTRNFD
KKGKHYGPFASSEMKAIDLYNLLKLFPLKNCFNKGRKCEFYDLNLCMKACTHEVSEAD
YEVMMKKIDYFFHNGADQVLKDLKEKESIASEKFDQAKKYLDLQKAINLIFDKQIINL
YSAKERIDVLAYQIKENVICIVLFSYVSSQLVSKNTICDFYYGEEQEVITSYLSQYYKDN
IKPKILYASLDQANATLLKDSLGIIEINPTSGKMNEIMSLALQNVTNELSQKYDSLKKE
QRINLALDQLKLIKVDKLNHLEVYDNSNLFNTDKVSAMIVFENNQFNKKKYRKYKIKDQ
QALGDYHYMYEVIYRRLYQALKNNFVDLPDLIILDGGKHQVLAAKKAIVDLQIDKKINLI
GLAKNNKHQTDKIVTFDLDEISLDKSSALYFFLANLQEEVHKFAISFFRKTAKSLYDSI
LDQIKGLGKRRKQQLIEHFKTIDEIKKASIASLSQVLPPIEIAKKLKQKLDQS
```

```
sp|P0A8G0|UVRC_ECOLI MSDQFDAKAFKTVTSQPGVYRMYDAGGTVIYVGKAKDLKKRLSSYFRSNLASRKTEALV 60
sp|P14951|UVRC_BACSU MNKQLKEK--LALLPDQPGCYLMKDRQQTVIYVGKAKVLKNRVRSYFTG-SHDAKTQRLV 57
sp|Q7NBC4|UVRC_MYCGA --MNFNLKQKLDLAPKKPGCYLWKNHLNEIYIGKAKNIYKRVHQYFNG-PKDLKTSKLV 57
::: * * ..:** * : :***:***: :*: **. . **. **
```

```
sp|P0A8G0|UVRC_ECOLI AQIQQIDVTVTHTTETEALLLEHNYIKLYQPRYNVLLRDDKSYPFIFLSGDTHPRLAMHRG 120
sp|P14951|UVRC_BACSU TEIEDFEYIVTSSNLEALILEMNLIKKHDPKYNVMLKDDKTYPFIKLTHERHPRILIVTRN 117
sp|Q7NBC4|UVRC_MYCGA NEIFYVEFIEVNNENEALLLEANLIKHKPRYNILLKDNNGYPYILMTKEKYPRLIYTRN 117
:* .: .: .: ***:** * ** :*:***:***: ***: * : : **** *
```

```

sp|P0A8G0|UVR_C_ECOLI AKHAKGEYFGPPFN-GYAVRETLALLQKIFPIRQCENSVYRNRSRCLQYQIGRCLGPCV 179
sp|P14951|UVR_C_BACSU VKKDKGRYFGPPFN-VQAAARETKKLLDRLYPLRKCS---KLPDRVCLYHLGQCLAPCV 172
sp|Q7NBC4|UVR_C_MYCGA FDPKKGKHYGPFASSEMKAADLYNLLKLFPLKNCFN---KGRKCEFYDLNLCMKACT 173
. **.:**.. . : ** :.:**:* . * * *.. * : .*.

sp|P0A8G0|UVR_C_ECOLI EGLVSEEEYAQQVEYVRLFLSGKDDQVLTQLISRMETASQNLEFEEAARIRDQIQAVRRV 239
sp|P14951|UVR_C_BACSU K-DISEETNRELVESITRFLRGGYNEVKKLEEKMEHAAENLEFERAKELRDQIAHIEST 231
sp|Q7NBC4|UVR_C_MYCGA H-EVSEADYEVMKKKIDYFFHNGADQVLKDLKEKESIASEKFDQAKKYLDLQKAINLI 232
. :** : : * : . :.* . : * : : : : : . * . * . : .

sp|P0A8G0|UVR_C_ECOLI TEKQ-FVSN TGDD-LDVI GVAFDAGMACVHVL FIRQ GKVLG-SRSYFPKVPGGTELSEVV 296
sp|P14951|UVR_C_BACSU MEKQKMTMNDLVD-RDVFAYAYDKGWMCVQVFFIRQ GKLI ERDVS MFPLYQ---EAD EEF 287
sp|Q7NBC4|UVR_C_MYCGA FDKQIINLYSAKERIDVLAYQIKENVICIVLFSYVSSQLVSKNTICDFYYG---EEQ EVI 289
:* : : : * : . . * : : : . : : : . * . * .

sp|P0A8G0|UVR_C_ECOLI ETFVQGQFYLQGSQMRTLPGEILLDFNLSDKTLLADSLSELAGRKINVTQKPRGDRARYLK 356
sp|P14951|UVR_C_BACSU LTFIQGFYSKNN--HFLPKELVP--DSIDQSMIEQLL-----ETNVHQPKKGPCKELLM 338
sp|Q7NBC4|UVR_C_MYCGA TSYLSQYYKDNK----PKILYASLDQANATLLKDSLGL-----IEIINPTSGKMMNEIMS 339
: : . : * . : : : : : * : : : * . : .

sp|P0A8G0|UVR_C_ECOLI LARTNAATALTSKLSQQSTVHQR----LTALASVLKLPVEVRMECFDISHMTGMEQTVASC 412
sp|P14951|UVR_C_BACSU LAHKNAKIALKEKFS LIERDEERSIGAVQKLGEALNIYTPHRIEAFDNSNIQGTNPVSAM 398
sp|Q7NBC4|UVR_C_MYCGA LALQNVTNELSQYKSLVKKEQRINLALDQLKLLIKVDKLNHLEVYDNSNLFNTDKVSAM 399
** * . * . * . . : * : : * : : : : : * : * : : : * : : :

sp|P0A8G0|UVR_C_ECOLI VVFDANGPLRAEYRRYNTGITPGDDY AAMNQVLR RRYGKAIDDSK--IPDVLIDGGKG 470
sp|P14951|UVR_C_BACSU IVFIDGKPYKKEYRKYKIKVTGPDY GSMREVRRRYTRVLENLP-LPDLIIIDGGKG 457
sp|Q7NBC4|UVR_C_MYCGA IVFENNQFNKKYRKYKIKDQQA LGDYHYMYEVIYRRLYQALKNNFVDLPDLIILDGGKH 459
:* . : : * : * : . ** * : * : * : : : . : * : * : * : * : *

sp|P0A8G0|UVR_C_ECOLI QLAQAKNVFAELDVSWDKNHPLLGVAKGADRKAGLETLFFEPEGEGFSLPPDSPALHVI 530
sp|P14951|UVR_C_BACSU QINAARDVIEN-ELGLDIP---IAGLAK--DEKHRTSNLLIGDPLEVAYLERN SQEFYLL 511
sp|Q7NBC4|UVR_C_MYCGA QVLAAKKAIVDLQIDKKIN---LIGLAK---NNKHQTDKIVTFDLDEISL DKSSALYFFL 513
* : * : . : : : . : * : * : . : : . : * . * . : .

sp|P0A8G0|UVR_C_ECOLI QHIRDESHDHAIGHRKKRAKVKNTSSLETIEGVGPKRRQMLLKYMGGQLGRNASVEEI 590
sp|P14951|UVR_C_BACSU QRIQDEVHRFAISFHRQIRGKSAFQSVLDDIPGIGEKRRKMLLKHFGSVKMKKEASLEDI 571
sp|Q7NBC4|UVR_C_MYCGA ANLQEVHFKFAISFFRKTKAKSLYSDILDQIKGLGKRRKQLIEHFKTID EIKKASIASL 573
. : : * * . * . : * : * * : * * : * : : : . : : : * : : .

sp|P0A8G0|UVR_C_ECOLI AKVPGISQGLAEKIFWSLKH- 610
sp|P14951|UVR_C_BACSU KKA-GVPAAAAQLLYDKLQK- 590
sp|Q7NBC4|UVR_C_MYCGA SQV--LPIETAKKLLKQLDQS 592
. . . . * : : . * : .

```

9. DNA helicase II (gene name *uvrD*)

```

>sp|P03018|UVRD_ECOLI DNA helicase II OS=Escherichia coli (strain K12)
GN=uvrD PE=1 SV=1
MDVSYLLDSLNDKQREAVAAPRSNLLVLAGAGSGKTRVLVHRIAWLMSVENCSPYSIMAV
TFTNKAAAEMRHRIGQLMGT SQGGMWVGT FHGLAHRLLRAHHMDANLPQDFQILDSEDL
RLKRLIKAMNLDKQWPPRQAMWYINSQKDEGLRPHHIQS YGNPVEQTWQKVYQAYQEA
CDRAGLVDFAE LLLRAHELWLNKPHILQHYRERFTN I LVDEFQDTNNIQYAWIRLLAGDT
GKVMIVGDDDDQSIYGWRGAQVENIQRF LNDFPGAETIRLEQNYRSTSNILSAANALIENN
NGRLGKKLWTDGADGEPISLYCAFNE LDEARFV VNR IKTWQDNGGALAECAILYRSNAQS
RVLEEALLQASMPYRIYGGMRFFERQEIKDALS YLRLIANRNDAAFERVVNTPTRGIGD
RTL DVVRQTSRDRQLTLWQACRELLQE KALAGRAASALQRFMELIDALAQETADMPLHVQ
TDRVIKDSGLRTMYEQEKGEKGQTR IENLEELVTATRQFSYNEEDEDLMPLQAFLSHAAL
EAGEGQADTWQDAVQ LMTLHSAKGLEFPQVFIVGMEEGMFP SQMSLDEGGRL EEEERLAY
VGVTRAMQKLTLYAETRRLYGKEVYHRPSRFIGELPEECVEEVRLRATVSRPVSHQRMG
TPMVENDSGYKLGQRVRHAKFGE GTIVNMEGSGEHSRLQVAFQGGQIKWLVAAYARLESV
>sp|O34580|PCRA_BACSU ATP-dependent DNA helicase PcrA OS=Bacillus subtilis
(strain 168) GN=pcrA PE=1 SV=1
MNYISNQLLSGLNPVQQA VKT TDG P LLLMAGAGSGKTRVLTHRIAYLMAEKHVAPWNIL
AITFTNKAAAREMKERVESILGPGADDI WISTFHSMCVRILRRDIDRIGINRNF SILD TAD
QLSVIKGILKERNLDPKKFDPRSILGTISSAKNELTEPEEFSKVAGGYDQVVS DVYADY
QKKLLKNQSLDFDDLIMTTIKL FDRVPEVLEFYQRKFQYI HVDEYQDTNRAQYMLVKQLA
ERFQNL CVVGDSDQSIYRWRGADITN ILSFEKDPNASVILLEQNYRSTKRILRAANEVI
KNNSNRKPKNLWTE NDEGIKISY YRGDNEFGEGQFVAGKIHQLHSTGKRKLS DIAILYRT

```

NAQSRVIEETLLKAGLNYNIVGGTKFYDRKEIKDILAYLRLVSNPDDDISFTRIVNVPKR
 GVGATSLEKIASYAAINGLSFFQAIQQVDFIGVSAKAANALDSFRQMIENLTNMQDYL
 TELTEEILDKTEYREMLKAEKSIEAQSRLENIDEFLSVTKNFEQKSEDKTLVAF
 IADIDQLDQKEEESGGKDAITLMTLHAAKGLEFPVFLMGLEEGVFPHSRSLMEEAE
 MEEERRLAYVGITRAEQELYLTNAKMRTLFRGTNMNPESTRFIAEIPDDLLENLNEK
 KETRATSARKMQPRRGPVSRPVSASYASKTGGDTLNWAVGDKAGHKKWGTGTVVSVK
 GEGEGTELDIAFPSPVGVKRLLAFAPIEKQ

>tr|Q7NB99|Q7NB99_MYCGA DNA helicase II OS=Mycoplasma gallisepticum (strain
 R(low / passage 15 / clone 2)) GN=uvrD PE=4 SV=2

MQDYLKSLNKQQYDVVTSDLIPFVAVAGAGTGKTKVLTSTRYALIEHFKIPEYKILAITF
 TNKAAKEMQHRLEKLLNKEKTQVSFRTHFGFCAQVLRREEVNNVDRLNDRFNILDEVDQAK
 LIEDLLKSQKYEYYSQYTFDFKKNKVMISIINDAKTYNLDVAEFLASDLNKLGEDHILTPN
 LVQPLSNFYHDYEKALKELNAIDFNLDLNIAYNLFNDPIILKKWQNRYEAILVDEFQDA
 NEIQYKIVKLLREKNNNFLFVGDQSIYGWRGANSEIGDSIRDFDNDLVVKYLTQNYRS
 KOSILNLANDAIIKMNSRYFKSLLSHDLTDLGPKPIWINFNSNIEYQNRVMDKIKELVAS
 KQYTYGDFAILYRTNFSVSLERLIKENRIPYEIFGGYKFFLRKEIKDLIGYKLVDTNN
 DIAFDRIINTPRRMIGDTSIEI IKELANKKSITEYEALDYLDESNIKANVKKSAQNFKKM
 IEDLRANQGNWSVYQTINEI IKRINYDYLDNEPTKHDSVNEFIDFLNKYEKEYENDFGTK
 LTINDFIQNLALEGLDLDNNQPNHNKNAKLMTHSAKGLEFKNVFVINMNENILPSSRSI
 AATNNKAKLAEERRIVVAYTRAKHNLWLCSNQDYDARTKEPYQPSRFLYELSDLVLDKQ
 DQAKTYFDKHNFLTDDGWFNSKKSPSKLDWNSTNEVEHNYFVGEIIYHQLYGEGIVRE
 IDDLTIKVSFKDKKAGTKDLIKNHKLITYAK

sp|P03018|UVRD_ECOLI --MDVSYLLDLSLNDKQREAVAAPRSNLLVLAGAGSGKTRVLVHRIAWLMS 48
 sp|O34580|PCRA_BACSU MNYISNQLLSGLNPVQQEAVKTTDGPLLLMAGAGSGKTRVLTHRIAYLMA 50
 tr|Q7NB99|Q7NB99_MYCGA ---MQDYLKSLNKQQYDVVTSDLIPFVAVAGAGTGKTKVLTSTRYALIE 46
 . *..** * :.* : :.:****:***:*. ***:.*:

sp|P03018|UVRD_ECOLI VENCSPYSIMAVTFTNKAAAEMRHRIGQLMGTSQGGMWVGTFFHGLAHRLL 98
 sp|O34580|PCRA_BACSU EKHVAPWNILAITFTNKAAAREMKERVESILGPGADDIWIWSTFHSMVCVRI 100
 tr|Q7NB99|Q7NB99_MYCGA HFKIPEYKILAITFTNKAAKEMQHRLEKLLNKEKTQVSFRTHFGFCAQV 96
 : . :.*:*** **:*: :.: . ***:.. :.*

sp|P03018|UVRD_ECOLI RAHMDAN-LPQDFQILDSEDLRLKRLIKAMNLD-----KQWPPRQA 142
 sp|O34580|PCRA_BACSU RRDIDRIG-INRNFSLDTADQLSVIKGILKERNLDP-----KKFDPRSI 144
 tr|Q7NB99|Q7NB99_MYCGA REEVNNVDRLNDRFNILDEVDQAKLIEDLLKSQKYEYYSQYTFDFKKNKV 146
 * . . : *..** ** :.: :.* : : . . : .

sp|P03018|UVRD_ECOLI MWYINSQKDEGLRP-HHIQSYG-----NPVEQTWQKVYQAYQEAC 181
 sp|O34580|PCRA_BACSU LGTISSAKNELTEPEEFSKVAG-----GYDQVVSVDVYADYQKKL 184
 tr|Q7NB99|Q7NB99_MYCGA MSIINDAKTYNLDVAEFLASDLNKLGEDHILTPNLVQPLSNFYHDYEKAL 196
 : *.. * .. * ..* **:

sp|P03018|UVRD_ECOLI DRAGLVDFAEALLRAHELWLNKPHILQHYRERFTNILVDEFQDTNNIYA 231
 sp|O34580|PCRA_BACSU LKNQSLDFDDLIMTTIKLFDVRVPEVLEFYQRKFQYIHVDEYQDTNRAQYM 234
 tr|Q7NB99|Q7NB99_MYCGA KELNAIDFNLDLNIAYNLFNDPIILKKWQNRYEAILVDEFQDANEIQYK 246
 . :** *: : :.* . * * : :.: : * ***:**:* **

sp|P03018|UVRD_ECOLI WIRLLAGDTGKVMIVGDDQSIYGWRGAQVENIQRFLNDFPGAETIRLEQ 281
 sp|O34580|PCRA_BACSU LVKQLAERFQNLVCGSDQSIYRWGADITNLLSFEKDYPNASVILLLEQ 284
 tr|Q7NB99|Q7NB99_MYCGA IVKLLREKNNNFLFVGDQSIYGWRGANSEIGDSIRDFDNDLVVKYLTQ 296
 : : * . : .*** ***: ***: : : * : . . * *

sp|P03018|UVRD_ECOLI NYRSTSNILSAANALIENNGRLGKKLWTD--GADGEPISLYCAFNELE 329
 sp|O34580|PCRA_BACSU NYRSTKRILRAANEVIKNNSNRKPKNLWTE--NDEGIKISYYRGDNEFGE 332
 tr|Q7NB99|Q7NB99_MYCGA NYRSKQSILNLANDAIIKMNSRYFKSLLSHDLTDLGPKPIWINFNSNIEYQ 346
 ****.. ** ** * :.* * * : . * * :

sp|P03018|UVRD_ECOLI ARFVVNRIKTWQDNG-GALAECAILYRSNAQSRVLEEALLQASMPYRIYG 378
 sp|O34580|PCRA_BACSU GQFVAGKIHQLHSTGKRKLSDIAILYRTNAQSRVIEETLLKAGLNYNIVG 382
 tr|Q7NB99|Q7NB99_MYCGA NRVVMDKIKELVASKQYTYGDFAILYRTNFSVSLERLIKENRIPYEIFG 396
 :** :.*: . :.: ****.* * :.* : : : * * *

sp|P03018|UVRD_ECOLI GMRFFERQEIKDALSYLRLIANRNDAAFERVVNTPTRGIGDRITLDVVVQ 428
 sp|O34580|PCRA_BACSU GTKFYDRKEIKDILAYLRLVSNPDDDISFTRIVNVPKRGVATSLEKIAS 432
 tr|Q7NB99|Q7NB99_MYCGA GYKFFLRKEIKDLIGYKLVDTNN-DIAFDRIINTPRRMIGDTSIEIIE 445
 * :*: *:**: ***: : . : * * * :.* * * * : : .

sp|P03018|UVRD_ECOLI TSRDRQLTLWQACRELLQEKALAGRAASALQRFMELIDALAQETADMPLH 478
 sp|O34580|PCRA_BACSU YAAINGLSFFQAIQQ-VDFIGVSAKAANALDSFRQMIENLTNMQDYL 481
 tr|Q7NB99|Q7NB99_MYCGA LANKKSITEYEALDY-LDESNIKANVKKSAQNFKKMIEDLRANQGNWSVY 494

```

: . :: :*   :: : ... .: : * ::*: *   .:
sp|P03018|UVRD_ECOLI      VQTRDRIKDSGLRMTMYEQEKGEKQOTRIENLEELVTATRQFSYNEEDED- 527
sp|O34580|PCRA_BACSU     ELTEEIILDKTEYREMLKAEKSIEAQSRLNIDEFLSVTKNFEQKSEDK-- 529
tr|Q7NB99|Q7NB99_MYCGA  QTINEI IKRINYDYDLN-----EPTKHDSVNEFIDFLNKYEKEYENDFG 538
      :...:   :           : : :*: :   .: : *:.

sp|P03018|UVRD_ECOLI      -LMPLQAFLSHAALEAGEGQADT-----WQDAVQLMTLHSAKGLEFPQV 570
sp|O34580|PCRA_BACSU     ---TLVAFITDLALADIQDLDQKEEESGGKDAITLMTLHAAKGLEFPVV 576
tr|Q7NB99|Q7NB99_MYCGA  TKLTINDFIQNLALEGDLNQP----NHKNKALKLMTIHSKAGLEFKNV 584
      .: * : . ** . . : :           :*: **:*:***** *

sp|P03018|UVRD_ECOLI      FIVGMEEGMFSPQMSLDEG---GRLEEERRLAYVGVTRAMQKLTLYAET 617
sp|O34580|PCRA_BACSU     FLMGLEEGVFPHSRSLMEE---AEMEEERRLAYVGITRAEQELYLTNAKM 623
tr|Q7NB99|Q7NB99_MYCGA  FVINMNENILPSSRSIAATNNKAKLAEERRIVYVAYTRAKHNLWLCNSQD 634
      *:::.*:.*: * . *           .: *::.*: * . * * : * * :

sp|P03018|UVRD_ECOLI      RRLYGKEVYHRPSRFIGELPEECVVEVR----LRATVSR-----PVSH 656
sp|O34580|PCRA_BACSU     RTLFRGTNMNPESRFIAEIPDDLLENLNEKKESTRATSARKMQPRRGPVSR 673
tr|Q7NB99|Q7NB99_MYCGA  YDARTKEPYQ-PSRFLYELSDLVLDKQDQAKTYFDKHNFLTDDDGWFNSK 683
      : : *:: *::: : :           . *

sp|P03018|UVRD_ECOLI      QRMGTPMVENDSG-YKLGQVRVHAKFGEGTIVNMEGSGEHSRLQVAFQG- 704
sp|O34580|PCRA_BACSU     PVSYASKTGGDTLNWAVGDKAGHKKWGTGTVVSVKGEGBTELDIAFPSP 723
tr|Q7NB99|Q7NB99_MYCGA  KSPSKLDAWNSTNEVEHNYFVGEI IYHQLYGEIVREIDDLTIKVSFKDK 733
      . .: . . . :           .: . : :*: .

sp|P03018|UVRD_ECOLI      -QGIKWLVAAYARLESV- 720
sp|O34580|PCRA_BACSU     -VGVKRLAFAPIEKQ- 739
tr|Q7NB99|Q7NB99_MYCGA  KAGTKDLIKNHKLITYAK 751
      * * * : . :

```

10. Formamidopyrimidine-DNA glycosylase (gene name *fpg* (*mutM*))

```

>sp|P05523|FPG_ECOLI Formamidopyrimidine-DNA glycosylase OS=Escherichia coli
(strain K12) GN=mutM PE=1 SV=3
MPELPEVETSRRGIEPHLVGATILHAVVRNGLRWRPVSEEIYRLSDQPVLSVQRRAKYLL
LELPEGWII IHLGMSGSLRILPEELPPEKHDHVDLMSNGKVLRYTDPFRFGAWLWTKEL
EGHNVLTHLGPPEPLSDDFNGEYLHQKCAKKKTAIKPWLMNDKLVVGVGNIYASESLFAAG
IHPDRLASSLSLAECCELLARVIKAVLLRSIEQGGTTLKDFLQSDGKPGYFAQELQVYGRK
GEPCRVCGTPIVATKHAQRATFYCRQCQK
>sp|O34403|FPG_BACSU Formamidopyrimidine-DNA glycosylase OS=Bacillus subtilis
(strain 168) GN=mutM PE=3 SV=4
MPELPEVETVRRTLTGLVKGKTIKSVEIRWPNI IKRPAEPEEFARKLAGETIQSIGRRGK
FLLFHLHDHYVMVSHLRMEGKYGLHQAEPEDDKHVHVIF TMTDGTQLRYRDVRKFGTMHLF
KPGE EAGELPLSQLGPEPDAEEFTSAYLKDR LAKTNRAVKTALLDQKT VVGLGNIYVDEA
LFRAGVHPETKANQLSDKTIKTLHAEIKNTLQEAIDAGGSTVRSYINSQGEIGMFQLQHF
VYGKKDEPCKNCGTMI SKIVVGGRGTHFCAKCQTKK
>tr|Q7NBN4|Q7NBN4_MYCGA Formamidopyrimidine-DNA glycosylase OS=Mycoplasma
gallisepticum (strain R(low / passage 15 / clone 2)) GN=mutM PE=3 SV=2
MPELPEVQTVINYLKTKI INQKINNVIVSALKVLKNATAKEFKKFLVNEHFVDIKRIGKY
IIFILSNNKVLVSHLRMEGKYKISQFKAKYDERHVLVRFILDDFELHYHDTRRFGTFHIH
SVLDYQDQDYLLKLAIDPTQQEWDWKYLKNNAQKSSRVIKSVLLDQSVVAGIGNIYADEI
LFLSKINPAKKANELTDQQFKEISK NATKVLLKAIELNGTTIFSYQFKENHAGSYQDYLN
VHLQKDKPCKVCGNLVKKTKLNNRGTYCYCAKCQK

sp|P05523|FPG_ECOLI      MPELPEVETSRRGIEPHLVGATILHAVVRN-GRLRWPVSEEIYR--LSDQ 47
sp|O34403|FPG_BACSU     MPELPEVETVRRTLTGLVKGKTIKSVEIRWPNI IKRPAEPEEFARKLAGE 50
tr|Q7NBN4|Q7NBN4_MYCGA  MPELPEVQTVINYLKTKI INQKINNVIVSALKVLKN-ATAKEFKKFLVNE 49
      *****:* . : : . .* . :           : : . : : * . :

sp|P05523|FPG_ECOLI      PVLSVQRRAKYLLLELPEG-WII IHLGMSGSLRILPEELPPEKHDHVDLV 96
sp|O34403|FPG_BACSU     TIQSIGRRGKFLFLHDHY-VMVSHLRMEGKYGLHQAEPEDDKHVHVIF T 99
tr|Q7NBN4|Q7NBN4_MYCGA  HFVDIKRIGKYIIFILSNNKVLVSHLRMEGKYKISQFKAKYDERHVLVRF 99
      . .: * .*: : : * .           : : * * *. * . : : : : :

sp|P05523|FPG_ECOLI      MSNGKVLRYTDPFRFGAWLWTK--ELEGHNVLTHLGPPEPLSDDFNGEYLH 144
sp|O34403|FPG_BACSU     MTDGTQLRYRDVRKFGTMHLFKPGE EAGELPLSQLGPEPDAEEFTSAYLK 149
tr|Q7NBN4|Q7NBN4_MYCGA  ILDDFELHYHDTRRFGTFHIH SVLDYQDQDYLLKLAIDPTQQEWDWKYLK 149
      : .: *:* * *:*: . :           .: *:*.*: * : : : *:*

```

```

sp|P05523|FPG_ECOLI      QKCAKKKTAIKPWLMDNKLVVGVGNIYASESLFAAGIHPDRLASSLSLAE 194
sp|O34403|FPG_BACSU     DRLAKTNRVAVKTTALDQKTVVGLGNIYVDEALFRAGVHPETKANQLSDKT 199
tr|Q7NBN4|Q7NBN4_MYCGA NNAQKSSRVIKSVLLDQSVVAGIGNIYADEILFLSKINPAKKANELTDQQ 199
: . * . . . : * . : * . : * . : * . : * . : * . : * . : * . :
: . * . . . : * . : * . : * . : * . : * . : * . : * . : * . :

sp|P05523|FPG_ECOLI      CELLARVIKAVLLRSIEQGGTTLKDFLQSDGKPGYFAQELQVYGRKGEPC 244
sp|O34403|FPG_BACSU     IKTLHAEIKNTLQEAIDAGGSTVRSYINSQGEIGMFQLQHFVYGKKDEPC 249
tr|Q7NBN4|Q7NBN4_MYCGA FKEISKATKVLKAIELNGTTIFSYQFKENHAGSYQDYLVNHLQKDKPC 249
: : . . * . : * . : * . : * . : * . : * . : * . : * . : * . :

sp|P05523|FPG_ECOLI      RVCGTPIVATKHAQRATFYCRQCQK-- 269
sp|O34403|FPG_BACSU     KNCGTMISKIVVGGRGTHFCAKQCQTK 276
tr|Q7NBN4|Q7NBN4_MYCGA KVCGNLVKKTTLNRRGTYYCAKQCQK-- 274
: * . : . * . : * . : * . : * . : * . : * . : * . : * . :

```

11. Uracil-DNA glycosylase (gene name *ung*)

```

>sp|P12295|UNG_ECOLI Uracil-DNA glycosylase OS=Escherichia coli (strain K12)
GN=ung PE=1 SV=2
MANELTWHADVLAEEKQQPYFLNTLQTVASERQSGVTIYPPQKDFVNAFRFTELGDVKVVI
LGQDPYHGPQQAHLAFSVRPGIAIPPSLLNMYKELENTIPGFTRPNHGYLESWARQGV
LLNTVLTVRAGQAHSHASLWETFTDKVISLINQHREGVVFLWGSQAQKKGAIIDKQRH
HVLKAPHPSPLSAHRGFFGCNHFVLANQWLEQRGETPIDWMPVLPASEE
>sp|P39615|UNG_BACSU Uracil-DNA glycosylase OS=Bacillus subtilis (strain 168)
GN=ung PE=1 SV=1
MKQLLQDSWVNQLKEEFEKPYQELREMLKREYAEQTIYPDSRDIFNALHYTSYDDVKV
ILGQDPYHGPQQAQGLSFSVKPGVKQPPSLKNIFLELQDDIGCSIPNHGSLVSWAKQGV
LLNTVLTVRRGQANSHKKGWERLTDRIIDVLSERERPVIIFILWGRHAQMKKERIDTSKH
FIIESTHPSPF SARNGFFGSRPF SRANAYLEKMGAPIDWCIKDL
>tr|Q7NAM7|Q7NAM7_MYCGA Uracil-DNA glycosylase OS=Mycoplasma gallisepticum
(strain R(low / passage 15 / clone 2)) GN=ung PE=3 SV=1
MLEQLIGEIQTNWKLINQFFATHKTIYHQLDQLIKNRSEKNELIPKKELI FNAFNFFDY
QETKVVIIIGQDPYADLKKANGLAFGVDNPNPVS LRNI IKELINNLKLDEQQLDDFDYSL
KSWANQGVLLINTILT VVKQNPLSDQNLGWEELIKFLILKLENTQTPVFVWLGKKAQGF
LEPYQLKHV LKSAHPSFFSAKQFFNNHFNLINE LLKTKNEQLIQVVKQNK

```

```

sp|P12295|UNG_ECOLI      ---MANELT---WHDVLAEEKQ-QPYFLNTLQTVASERQSGVTIYPPQK 43
sp|P39615|UNG_BACSU     ---MKQLLQDSWVNQLKEEFE-KPYQ-ELREMLKREYAEQTIYPDSRD 44
tr|Q7NAM7|Q7NAM7_MYCGA MLEQLIGEIQTNWKLINQFFATHKTIYHQLDQLIKNRSEKNELIPKKE 50
: : * : : : : * : . . . : * . .

sp|P12295|UNG_ECOLI      VFNAFRFTELGDVKVVI LQ Q PYHGPQQAHLAFSVRPGIAIPPSLLNMY 93
sp|P39615|UNG_BACSU     IFNALHYTSYDDVKVVI LQ Q PYHGPQQAQGLSFSVKPGVKQPPSLKNIF 94
tr|Q7NAM7|Q7NAM7_MYCGA IFNAFNFFDYQETKVVII G Q PYADLKKANGLAFGVDNPNPVS LRNI I 99
: * * . : . : . * * . : * * . : * * . : * * . : * * . :

sp|P12295|UNG_ECOLI      KELENTIPGFTRPNHG---YLESWARQGVLLNTVLTVRAGQAHSHASLG 140
sp|P39615|UNG_BACSU     LE LQDDI -GCSIPNHG---SLVSWAKQGVLLNTVLTVRRGQANSHKKG 140
tr|Q7NAM7|Q7NAM7_MYCGA KELINNLKLDEQQLDDFDYSLKSWANQGVLLINTILT VVKQNPLSDQNLG 149
* * : : . . * * * . * * * . : * * . : . * . . *

sp|P12295|UNG_ECOLI      WETFTDKVISLINQHREGVVFLWGSQAQKKGAIIDKQRHHVLKAPHPSP 190
sp|P39615|UNG_BACSU     WERLTDRIIDVLSERERPVIIFILWGRHAQMKKERIDTSKHFIESTHPS 190
tr|Q7NAM7|Q7NAM7_MYCGA WEELIKFLILKLENTQTPVFVWLGKKAQG--FLEPYQLKHV LKSAHPSF 197
* * : . : * : . . : * * * : * * . : . : . : . : * *

sp|P12295|UNG_ECOLI      LSAHRGFFGCNHFVLANQWLEQRGETPIDWMPVLPASEE 229
sp|P39615|UNG_BACSU     FSARNGFFGSRPF SRANAYLEKMGAPIDWCIKDL---- 225
tr|Q7NAM7|Q7NAM7_MYCGA FSAKQ-FNNHFNLINE LLKTKNEQLIQVVKQNK---- 231
: * . : * . . * * * : . * * :

```

Organism	Active site (Proton acceptor)
E. coli (strain K12)	D64
B. subtilis (strain 168)	D65
M. gallisepticum (strain R(low / passage 15 / clone 2))	No information

12. Endonuclease IV (gene name *nfo*)

```

>sp|P0A6C1|END4_ECOLI Endonuclease 4 OS=Escherichia coli (strain K12) GN=nfo
PE=1 SV=1
MKYIGAHVSAAGGLANAAIRAAEIDATAFALFTKNQRQWRAAPLTTQTIDEFKAAACEKYH
Y TSAQILPHDSYLINLGHVPTEALEKSRDAFIDEMQRCEQLGLSLLNFHPGSHLMQISEE
DCLARIAESINIALDKTQGVTA VIENTAGQGSNLGFKFEHLAAIIDGVEDKSRVGCIDT
CHAF AAGYDLRTPAECEKTFADFARTVGFKYLRGMHLNDAKSTFGSRVDRHHS LGE GNIG
HDAFRWIMQDDRFDGIPLILETINPDIWAE EIAWLKAQQTEKAVA
>sp|P54476|END4_BACSU Probable endonuclease 4 OS=Bacillus subtilis (strain
168) GN=nfo PE=3 SV=1
MLRIGSHVSMGKHM LLAASQEA VSYGANTFMIYTGAPQNT RRKKIEDLNIEAGRAHMQE
NGIDEIIVHAPYIINIGNTTNPSTFELGVDFLRSEIERTAAIGAKQIVLHPGAHV GAGAE
AGIKKIIIEGLNEVIDPNQNVQIALETMAGKGSECGRSFEELAQIIIEGVTHNEQLSVCFDT
CHTHDAGYNIVEDFDGVLNEFDKIIGIDRIKVLHINDSKNVKGARKDRHENIGFGEIGFD
ALQYVVHHEQLKDI PKILET PYVGEDKKNKKPPYRFEIEMLKEKQFDDTLLEKILQQ
>sp|Q7NBA9|END4_MYCGA Probable endonuclease 4 OS=Mycoplasma gallisepticum
(strain R(low / passage 15 / clone 2)) GN=nfo PE=3 SV=1
MKS NKIKYLGCFV GATKPDFMLGMVKT VVDYGAT SFMFYSGPPQSFRRTPTAQFKLDLAK
AYLAKHNLGDLGDN YVVHAPYLINLANGDSTKRERSFNFFLDELKRTNELGAKYFVLHPG
SALNVKDKTQALDHLATELNRAISM TKDTIICLETMADKGQQICSKFEELRYVIDQISDK
SRIGVCFDTCHVHDAGYDLAKTQELIDHFDQVIGLKYLYVIHLNDSKNPMGARKDRHANI
GYGKIGFENLLNFIYHKEICNKII ILET PWIDDP IRGEVPLYKEE IEMIRNKKFVEGLVN
EES

sp|P54476|END4_BACSU      -----MLRIGSHVSMGKHM LLAASQEA VSYGANTFMIYTGAPQNT RRKKIEDLNIEAGR 55
sp|Q7NBA9|END4_MYCGA    MKSNKIKYLGCFV GATKPDFMLGMVKT VVDYGAT SFMFYSGPPQSFRRTPTAQFKLDLAK 60
sp|P0A6C1|END4_ECOLI    -----MKYIGAHVSAAGG--LANAAIRAAEIDATAFALFTKNQRQWRAAPLTTQTIDEFK 53
      :  :*..*  :      :      ...  .*..*  :::  :  *      :  :  :

sp|P54476|END4_BACSU    AHMQENGI----DEIIVHAPYIINIGNTTNPSTFELGVDFLRSEIERTAAIGAKQIVLHP 111
sp|Q7NBA9|END4_MYCGA    AYLAKHNLGDLGDN YVVHAPYLINLANGDSTKRERSFNFFLDELKRTNELGAKYFVLHP 119
sp|P0A6C1|END4_ECOLI    AACEKYHYTS--AQILPHDSYLINLGH P--VTEALEKSRDAFIDEMQRCEQLGLSLLNFHP 110
      *      :      :  :  *  .*:**:  :  .  *  .  :  :  .*:**  :  *  .  :  :  **

sp|P54476|END4_BACSU    GAHV GAG-AEAGIKKIIIEGLNEVIDPNQNVQIALETMAGKGSECGRSFEELAQIIIEGVTH 170
sp|Q7NBA9|END4_MYCGA    GSALNVKDKTQALDHLATELNRAISM TKDTIICLETMADKGQQICSKFEELRYVIDQISD 179
sp|P0A6C1|END4_ECOLI    GSHLMQISEEDCLARIAESINIALDKTQGVTA VIENTAGQGSNLGFKFEHLAAIIDGVED 170
      *:  :      :  :  :  *  ..  ....  .*  .*:**:  .*..*  :*:  :  .

sp|P54476|END4_BACSU    NEQLSVCFDTCHTHDAGYNI--VEDFDGVLNEFDKIIGIDRIKVLHINDSKNVKGARKDR 228
sp|Q7NBA9|END4_MYCGA    KSRIGVCFDTCHVHDAGYDL--AKTQE-LIDHFDQVIGLKYLYVIHLNDSKNPMGARKDR 236
sp|P0A6C1|END4_ECOLI    KSRVGCIDTCHAF AAGYDLRTPAECEKTFADFARTVGFKYLRGMHLNDAKSTFGSRVDR 230
      :  :  :  .*:**:  *:*:  :      :  :  *  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

sp|P54476|END4_BACSU    HENIGFGEIGFDALQYVVHHEQLKDI PKILET PYVGEDKKNKKPPYRFEIEMLKEKQFDD 288
sp|Q7NBA9|END4_MYCGA    HANIGYGKIGFENLLNFIYHKEICNKII ILET PWIDDP IRGEVPLYKEE IEMIRNKKFVE 296
sp|P0A6C1|END4_ECOLI    HHS LGE GNIGHDAFRWIMQDDRFDGIPLILETINPD-----IWAEEIAWLKAQQTEK 282
      *  .*  *:*:**:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

sp|P54476|END4_BACSU    TLEKILQQ 297
sp|Q7NBA9|END4_MYCGA    GLVNEES-- 303
sp|P0A6C1|END4_ECOLI    AVA----- 285
    
```

13. Recombinase RecA (gene name *recA*)

```

>sp|P0A7G6|RECA_ECOLI Protein RecA OS=Escherichia coli (strain K12) GN=recA
PE=1 SV=2
MAIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMG
RIVEIYGPESSGKTTTLTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQP
DTGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRCLA
GNLKQSN TLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG
SETRVKVVKNKIAAPFKQAEFQILYGEGINFYGELVDLGVKEK LIEKAGAWYSYKGEKIG
QGGANATAWLKDNPETAKEIEKKVRELLLSNPSTPDFSVDDSEGVAETNEDEF
>sp|P16971|RECA_BACSU Protein RecA OS=Bacillus subtilis (strain 168) GN=recA
PE=1 SV=2
    
```

MSDRQAALDMALKQIEKQFGKGSIMKLGKTDTRISTVPSGSLALDALTALGIGGYPRGRII
 EVYGPRESSGKTTVALHAIAEVQQGGQAAFIDAEHALDPVYAQKLGVNIEELLSQPDTG
 EQALEIAEALVRSQAVDIVVSDVAALVPKAEIEGDMGDSHVGLQARLMSQALRKLSGAI
 NKSKTIAIFINQIREKVGVMFGNPETTPGGRALKFYSSVRLEVRRAEQLKQGNVDMGNKT
 KIKVVKNKVAPPFRTAEVDIMYEGEGISKEGEIIDLGTELDIDVQKSGSWYSYEEERLGQGR
 ENAKQFLKENKDIMLMIQEQIREHYGLDNNGVVQQQAEETQEELEFEE
 >tr|F8WJY9|F8WJY9_MYCGA Protein RecA OS=Mycoplasma gallisepticum (strain
 R(low / passage 15 / clone 2)) GN=recA PE=3 SV=1
 MFNKKNYDKIMNNIKNSKRNNDRITNITNYDVLKLLQKFGKTNIYLNEKDELKDLEAIS
 TGSIKLDHALGTDGFIKGRIVEIYGNESCGKTTLALSTIKQAIDRNMRVAFIDAEHALDL
 RYVKRLGIDLTKLI IARP DYGEQGFEI IKS LIKTELIDLIVVSDVAALVPKVEIEGKMED
 QTMGTHARMSRGLSRIQPLLAKHNVSVIFINQLREKVGIMFGNPEVTTGGKALKFYSS
 RLELRRAEIIKDAANNAIGIRSKATITKNKLSTPMTTTYIDFYFKSGISEVNEIIDLAID
 YQIIIEQSGSWFSYQKEKIAQGKANLITKLGEESEELYKLIKTEVLNKLKDCQ

```

sp|P0A7G6|RECA_ECOLI      -----MAIDENKQKALA--AALGQIEKQFGKGSIMRLGE 32
sp|P16971|RECA_BACSU      -----MSDRQAALD--MALKQIEKQFGKGSIMKLGE 29
tr|F8WJY9|F8WJY9_MYCGA  MFNKKNYDKIMNNIKNSKRNNDRITNITNYDVLKLLQKFGKTNIYLNEK 50
                           .:. : . * :.:** . * :

```

```

sp|P0A7G6|RECA_ECOLI      DRSMDETISTGSLSLDIALGAGGLPMGRIVEIYGPRESSGKTTTLQVIA 82
sp|P16971|RECA_BACSU      KTDTRISTVPSGSLALDALTALGIGGYPRGRIIEVYGPRESSGKTTVALHAI 79
tr|F8WJY9|F8WJY9_MYCGA  DELKDLEAISTGSIKLDHALGTDGFIKGRIVEIYGNESCGKTTLALSTIK 100
                           . :.:** : * * * * . * * * * : * * * * : * * * * : * *

```

```

sp|P0A7G6|RECA_ECOLI      AAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPDTEQALEICDA 132
sp|P16971|RECA_BACSU      EVQQGGQAAFIDAEHALDPVYAQKLGVNIEELLSQPDTGEQALEIAEA 129
tr|F8WJY9|F8WJY9_MYCGA  QAIDRNMRVAFIDAEHALDLRYVKRLGIDLTKLI IARP DYGEQGFEI IKS 150
                           . . . * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

```

sp|P0A7G6|RECA_ECOLI      LARSGAVDIVVSDVAALTPKAEIEGEIGDSHMLAARMMSQAMRKLGN 182
sp|P16971|RECA_BACSU      LVRSGAVDIVVSDVAALVPKAEIEGDMGDSHVGLQARLMSQALRKLSGA 179
tr|F8WJY9|F8WJY9_MYCGA  LIKTELIDLIVVSDVAALVPKVEIEGKMEDQTMGTHARMSRGLSRIQPL 200
                           * : : . * : * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

```

sp|P0A7G6|RECA_ECOLI      LKQSNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAV 232
sp|P16971|RECA_BACSU      INKSKTIAIFINQIREKVGVMFGNPETTPGGRALKFYSSVRLEVRRAEQL 229
tr|F8WJY9|F8WJY9_MYCGA  LAKHNVSVIFINQLREKVGIMFGNPEVTTGGKALKFYSSRLELRRAEII 250
                           : : : . * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

```

sp|P0A7G6|RECA_ECOLI      KEGEN-VVGSETRVVKVKNKIAAPFKQAEFQIILYGEINFYGELVDLGVK 281
sp|P16971|RECA_BACSU      KQGN-VMGNKTKIKVVKNKVAPPFRTAEVDIMYEGEGISKEGEIIDLGT 278
tr|F8WJY9|F8WJY9_MYCGA  KDAANNAIGIRSKATITKNKLSTPMTTTYIDFYFKSGISEVNEIIDLAID 300
                           * : . : * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

```

sp|P0A7G6|RECA_ECOLI      EKLIKAGAWYSYKGEKIGQGANATAWLKDNPETAKEIEKKVRELLLSN 331
sp|P16971|RECA_BACSU      LDIVQKSGSWYSYEEERLGQRENKQFLKENKDIMLMIQEQIREHYGLD 328
tr|F8WJY9|F8WJY9_MYCGA  YQIIIEQSGSWFSYQKEKIAQGKANLITKLGEESEELYKLIKTEVLN---- 345
                           . : : : * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

```

sp|P0A7G6|RECA_ECOLI      PNSTPDFSVDDSEGVAETNEDF 353
sp|P16971|RECA_BACSU      NNGVVQQQAEETQEELEFEE-- 348
tr|F8WJY9|F8WJY9_MYCGA  -----KLKDCQ----- 351
                           . : : :

```

14. Holliday junction ATP-dependent DNA helicase subunit A (gene name *ruvA*)

```

>sp|P0A809|RUVA_ECOLI Holliday junction ATP-dependent DNA helicase RuvA
OS=Escherichia coli (strain K12) GN=ruvA PE=1 SV=1
MIGRLRGIIEKQPPVLVIEVGGVGYEVHMPMTCFYELPEAGQEAI VTFHFVVREDAQLL
YGFNNKQERTLKFELIKTNGVGPKLALAILSGMSAQQFVNAVEREEV GALVKLPGIGKKT
AERLIVEMKDRFKGLHGLDFTPAADLVLTSPASPATDDAEQEAVAALVALGYKPQEASRM
VSKIARPDASSETLIREALRAAL
>sp|O05392|RUVA_BACSU Holliday junction ATP-dependent DNA helicase RuvA
OS=Bacillus subtilis (strain 168) GN=ruvA PE=3 SV=2
MIEFVKGTIDYVSPQYIVIENGGIGYQIFTPNPFYKERSQETIFTYHHIREDAFSLYGF
STREEKALFTKLLNVTGIGPKGALAILGSGDPAVIAIENEDEAF LVKFPGVGKKTARQ

```



```

GIVQDFIKVEPEYELAISTALNKAANKNIIVDSNQDAINAVNFLKANKAGRATFLPLANLK
DRDVKPEHLEVLEQVEGYLGI AANLVNYHDQYDPAIRALLGQIIIASDLEAATKISKFTY
QLYRVISLGGDIVNAGGAI TGGAESKQTHSLFNLDEKIDTLKNE LLVAEKNINELNKKLE
FLTADYTKKDKEFNEQKIAIQRYQDLIVIEQKKLDDYKIQYEQLTDKTFDGKDVKWDDKK
IKDKLFSLETKKATLVQDLKINQEAKDMYQKQVNQLEKDVTLFYKEIDEDKNDKLRREQ
LTKHENTIYLA KSKINESYNMAIEFAIENYNKPLPISLSQARSEVVKLQSTLNNLGA INM
EAIQELDIKKERYEKLYSQQQELINARERINQAIIRLDEKAI FEFDQLINNLNKELPKTF
YYLFGGNC EIRYSNPEEKLTSGIEVFASPPGKNIGNLNL LSGGEKALVALSVLFSILKV
SSFPLVVLDEAESALDLANVERFANI IKNSSDQQTQFLIITHREGTMVKCDKLI GATMQTK
GVTKMLSVSLHQAKDMAEEIESQ

sp|P51834|SMC_BACSU      -MFLKRLDVIGFKSFAERISVDFVKGVTAVVGPNGSGKSNITDAIRWVLG 49
tr|Q7NBU0|Q7NBU0_MYCGA  MLFLKKFHAQGFKSYADNISFTFDEHVTGIVGPNGSGKSNVVDALKWVLG 50
      :***:.. ***:*. ** * : ** :*****:..**:*

sp|P51834|SMC_BACSU      EQSARSLRGGKMEIDIIFAGSDSRKRLNLAEVTLTLDNDHFLPIDFHEVS 99
tr|Q7NBU0|Q7NBU0_MYCGA  ERSMKNLRGKTSDDVIFFGSQEKPAKFAEVSLTFDNSQGYLHDKRKEIT 100
      *. * :.*** . :*:** **:. : ***:**:* : * . :*:

sp|P51834|SMC_BACSU      VTRRVYR-SGESEFLINNQPCLKDIIDL FMD SGLGKEAFSII SQKVEE 148
tr|Q7NBU0|Q7NBU0_MYCGA  VTRRVYRSGVSEYLI NNPESSLKEINDIFLDSGLTKGSLCIISQNTVSS 150
      ***** ** **:*:***:*. **:* *:***:*** * :.*****. *..

sp|P51834|SMC_BACSU      ILSSKAEDRRSIFEEAAGVLKYKTRKKAENKLFETQDNLNRVEDILHEL 198
tr|Q7NBU0|Q7NBU0_MYCGA  FIEAKPEDRRQIFEDAAGIGRYAKKKQDAIRQIARTNDNLKEITTVINEL 200
      :.:*.***.***:***: :* .:*. * :.: .*:***:.. *:*

sp|P51834|SMC_BACSU      EGQVEPLKIQASIAKDYLEKKKELEHVEIALTAYDIEELHGKWLSTLKEKV 248
tr|Q7NBU0|Q7NBU0_MYCGA  NRDLKKNLQQA EKAILYAETKEKLDLEITLSVNEYLISQKEIEALSEQI 250
      : :. : * : ** . * * *.*:***:***:.. : : : :.***:..

sp|P51834|SMC_BACSU      QMAKEEELAESSAISAKEAKIEDTRDKIQALDESVDLQVLLVTSEELE 298
tr|Q7NBU0|Q7NBU0_MYCGA  AEIDERLLKNDPQLQINQEKLEAFKKRYNSADQNVQKIQDELQKIYDEIV 300
      .*. * .. . :. : *:* :. : : *:.***: * * :*:

sp|P51834|SMC_BACSU      KLEGRKEVLKERKKNVQVQEQLEEAIVQFQKQETVLKEELSKQEA V FET 348
tr|Q7NBU0|Q7NBU0_MYCGA  LLE-KRNVFNDLQLKSDLSNDKKNKINALEQLLKSSEEQLKKYFELIST 349
      ** :.***: : : : : : : : * : ** . :*:*. * :. *

sp|P51834|SMC_BACSU      LQAEVQLRAQVKEKQALSLHNENVEEKIEQLKSDYFELLNSQASIRNE 398
tr|Q7NBU0|Q7NBU0_MYCGA  WEEELKEK-----DVDKTDLANELENLKKSLATFQVKRYEANLQ 388
      : *:* : : : : : : : : : : : : : : : : : : :

sp|P51834|SMC_BACSU      LQLLDDQMSQSAVTLQRLADNNEKHLQERHDISARKAACETEFARIEQEI 448
tr|Q7NBU0|Q7NBU0_MYCGA  VQFYQNQKIN----- 398
      :* : : * :

sp|P51834|SMC_BACSU      HSQVGAYRDMQTKYEQKKRQY EKNESALYQAYQYVQQARSKKDMLETMQG 498
tr|Q7NBU0|Q7NBU0_MYCGA  -----QFAQDAG----- 405
      *:.**

sp|P51834|SMC_BACSU      DFGFYQGVKEVLKAKERLGGIRGAVLELISTEQKYETAIEIALGASAQH 548
tr|Q7NBU0|Q7NBU0_MYCGA  -----VRTVLNNKDAIGGVHGI VQDFIKVEPEYELAISTALNKAANK 447
      * : ** : * : *** : * * : :. * * : ** ** . * : :

sp|P51834|SMC_BACSU      VVTDDEQSARKAIQYLKQNSFGRATFLPLSVIRDRQLQSRDAETAARHSS 598
tr|Q7NBU0|Q7NBU0_MYCGA  IIVDSNQDAINAVNFLKANKAGRATFLPLANLKDRDVKPEHLEVLEQVEG 497
      :. :*. * * : : : ** * . ***** : : * : : : . . * . : .

sp|P51834|SMC_BACSU      FLGVASELVTFDPA YRSVIQNL LGTVLITEDLKGANELAKLLGHRYRIVT 648
tr|Q7NBU0|Q7NBU0_MYCGA  YLGI AANLVNYHDQYDPAIRALLGQIIIASDLEAATKISKFTYQLYRVIS 547
      :* : * : * : . * . * : ** : * : * : * : * : * : * : * : * :

sp|P51834|SMC_BACSU      LEGDVVNP GSGMTGAVKKNNSLLGRSRELEDVTKRLAEMEETALLEQ 698
tr|Q7NBU0|Q7NBU0_MYCGA  LGGDIVNAGGAI TGGAESKQTHSLFN-----LDE 576
      * ** : * . * : * : * : * : * : * : * : * : * : * : * :

sp|P51834|SMC_BACSU      EVKTLKHSIQDMEKKLADLRETGEGLRLKQDVKGQLYELQVAEKNIN 748
tr|Q7NBU0|Q7NBU0_MYCGA  KIDTLKN-----ELLVAEKNIN-- 593
      :. :***: ** *****

sp|P51834|SMC_BACSU      LELYDQEK SALS EDEERKVRKRKLEEELS AVSEKMKQLEEDIDRLTKQK 798
tr|Q7NBU0|Q7NBU0_MYCGA  -----ELNKKL 599
      .*. * :

```

```

sp|P51834|SMC_BACSU      QTSSTKESLSNELTELKIAAAKKEQACKGEEDNLARLKKELTETELALK 848
tr|Q7NBU0|Q7NBU0_MYCGA  EFLTADYTKKDFNEQKIAIQRYQDLVIEQKLDYKIQY----- 641
      :  ::  .  .:*:*  ***  :  ::  *:*:*  *  :

sp|P51834|SMC_BACSU      EAKEDLSFLTSEMSSTSGEEKLEEAHKLNKDKTKTIELIALRRDQRIK 898
tr|Q7NBU0|Q7NBU0_MYCGA  ---EQLTDKTFDGKDVWDDKKIKDKLFSLETKKATLVQDLKINQEAADM 688
      *:*:*  *  :  .  .  .  .:*:*:*  :  :  .  .:*:*  :  :  :  :  :

sp|P51834|SMC_BACSU      LQHGLDLYERELKEMKRLYKQKTLLKDEEVKLRMEVELDNLQYLREE 948
tr|Q7NBU0|Q7NBU0_MYCGA  YQKQVNQLEKDVTLFYKEIDEDKNDKLRREQLTKHENTYIYLAISKINES 738
      *  :  :  *:*:*  :  :  .  .  .  .  .  .  .  .  *  :  *  :  :  .  .  .  .

sp|P51834|SMC_BACSU      YSLSFEGAKEKY--QLETDPEEARKRVKLIKLAIEELGTVNLGSLDEFER 996
tr|Q7NBU0|Q7NBU0_MYCGA  YNMAIEFAIENYKPLPISLSQARSEVVKLQSTLNNLGNLGAIMEAIEQLDI 788
      *  .  .  .  .  *  *  *  *  *  .  .  .  .  *  .  .  .  .  .  .  .  .  .  .  .  .

sp|P51834|SMC_BACSU      VNERKYFLSEQEDLTEAKNTLFQVIEEMDEEMTKRFNDTFVQIRSHFDQ 1046
tr|Q7NBU0|Q7NBU0_MYCGA  KKERYEKLYSQQELINARERINQAIIRLDEKAIFFEDQLINLNKELPK 838
      :  ***  :  *  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

sp|P51834|SMC_BACSU      VFRLSFGGGRAELRLTDPNDLLHSGVEIIAQPPGKKLQNLNLLSGGERAL 1096
tr|Q7NBU0|Q7NBU0_MYCGA  TFYYLFGGGNCEIRYSNPEEKLTSGIEVFASPPGKNIGNLNLSSGGEKAL 888
      .  *  *****  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

sp|P51834|SMC_BACSU      TAILLFSILKVRPVPFVCLDEVEAALDEANVFRFAQYLKKYSSDTQFIV 1146
tr|Q7NBU0|Q7NBU0_MYCGA  VALSVLFSILKVSSFPLVVLDEAESALDLANVERFANIKNSSDQTQFLI 938
      .  *  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

sp|P51834|SMC_BACSU      ITHRKGTMEEADVLYGVTMQESGVSKVISVKLEETKEFVQ----- 1186
tr|Q7NBU0|Q7NBU0_MYCGA  ITHREGTMVKCKDLIGATMQTKGVTKMLSVSLHQAKDMAEIEIESQ 983
      *****  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

```

```

>sp|P22523|MUKB_ECOLI Chromosome partition protein MukB OS=Escherichia coli
(strain K12) GN=mukB PE=1 SV=2
MIERGKFRSLTLINWNGFFARTFDLDELVTTLSSGGNGAGKSTTMAAFVTALIPDLTLLHF
RNTTEAGATSGSRDKGLHGKLGKAGVCYSMLDTINSRHQRVVVGVRLQVAVGRDRKVDIKP
FAIQGLPMSVQPTQLVTEITLNERQARVLPINELKDKLEAMEGVQFKQFNSITDYHSLMFD
LGI IARRLRASDRSKFYRLIEASLYGGISSAITRSLRDYLLPENSGVRKAFQDMEAAALR
ENRMTLEAIRVTQSDRDLFKHLISEATNYVAADYMRHANERRVHLDKALEFRRELHTSRQ
QLAAEQYKHVDMARELAEHNGAEGDLEADYQAASDHLNLVQTALRQQEKIEREADLDEL
QIRLEEQNEVVAAEAIERQQENEAREAAAELEVDELKSQ LADYQQALDVQQTRAIQYNQAI
AALNRAKELCHLPDLTADCAAEWLETFQAKELEATEKMLSLEQKMSMAQTAHSQFEQAYQ
LVVAINGPLARNEAWDVARELLREGVDQRHLAEQVQPLRMRLSELEQRLREQQEAERLLA
DFCKRQGNKFDIDELEALHQELEARIASLSDSVSNAREERMALRQEQEQLQSRIQSLMQR
APVWLAAQNSLNQLSEQCGEFTSSQDVTEYLQQLLEREREAIVERDEVGARKNAVDEEI
ERLSQPGGSEDQRLNALAERFGVLLSEIYDDVLEDPYFSALYGPSRHAI VVPDLSQV
TEHLEGLTDCPEDLYLIEGDPQSFDDSVFVSVDELEKAVVVKIADRQWRYSRFPEVPLFGR
AARESRIESLHAEREVLSEFATLSFDVQKTQRLHQAFSRFIGSHLAVAFESDPEAEIRQ
LNSRRVELERALS NHENDNQQR IQFEQAKEGV T ALNRI L PRLNLLADDSLADRVEIRE
RLDEAQEAARFVQQFGNQLAKLEPIVSVLQSDPEQFEQLKEDYAYSQQMQRDARQQAFAL
TEVVQRRAHFSYSDSAEMLSGNSDLNEKLRERLEQAEAEERTRAREALRGHAAQLSQYNQV
LASLKSSYDTKKE LLNDLQRELQDIGVRADSGAEERARIRRDELHAQLSNNSRRNQLK
ALTFCEAEMDNLTRKLRKLERDYFEMREQVVTA KAGWCAVMRMVKDNGVERRLHRRELAY
LSADDLRMSDKALGALRLAVADNEHLRDLRMS EDPKRPERKI QFFVAVYQHRLRERIRQ
DIIRTDDPVEAIEQMEI ELSRLTEELTSREQKLAISSRSVANI IRKTIQREQNRIRMLNQ
GLQNV SFGQVNSVRLNVNVRETHAMLLDV LSEQHEQHQDLFNSNRLTFSEALAKLYQRLN
PQIDMGQRTPTQ TIGEELLDYRNYLEMEVEVNRGSDGWLRAESGALSTGEAIGTGMSILVM
VVQSWEDESRRLRGKD ISPCRL LFLDEAARLDARS IATL FELCERLQMQLIIAAPENISP
EKGTTYKLVKRVFQNTTEHVHVGLRGFAPQLPETLPGTDEAPSQAS

```

```

sp|P51834|SMC_BACSU      -MFLKRLDVIGFKSFAERIS--VDFVKGVTAVVGPNGSGKSNITDAIRWV 47
tr|Q7NBU0|Q7NBU0_MYCGA  MLFLKKFHAQGFKSYADNIS--FTFDEHVTGIVGPNNGSGKSNVVDALKWV 48
sp|P22523|MUKB_ECOLI      MIERGKFRSLTLINWNGFFARTFDLDELVTTLSSGGNGAGKSTTMAAFVTA 50
      :  ::  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

sp|P51834|SMC_BACSU      LGEQSARSLRGGKMEIIFAGSDSR----KRLNLAEVLTLTDNDHFLP 92
tr|Q7NBU0|Q7NBU0_MYCGA  LGERSMKNLRGKTSDDVIFFGSQEK----PASKFAEVSLTFDNSQGYLH 93
sp|P22523|MUKB_ECOLI      LIPDLTLLHFRNTTEAGATSGSRDKGLHGKLGKAGVCYSMLDTINSRHQRV 100

```

```

* . : ** .. * *.
sp|P51834|SMC_BACSU      IDFHEVSVTRRVYR-SGESEFLINNQPC----- 119
tr|Q7NBU0|Q7NBU0_MYCGA  DKRKEITVTRRVYRGSVSEYLINNEPS----- 121
sp|P22523|MUKB_ECOLI     VVGVRLLQVAGRDRKVDIKPFAIQGLPMSVQPTQLVTEITLNERQARVLPL 150
      .: . * . . : *:. *

sp|P51834|SMC_BACSU      -----RLKDIIDLFMDSGLGKEAFSIIISQGK----- 145
tr|Q7NBU0|Q7NBU0_MYCGA  -----SLKEINDIFLDSGLTKGSLCIIISQNT----- 147
sp|P22523|MUKB_ECOLI     NELKDKLEAMEGVQFKQFNSITDYHSLMFDLGI IARRLRSASDRSKFYRL 200
      :: :.::* *: : *:.

sp|P51834|SMC_BACSU      -----VEEILSSKAEDRRSIFEEAAGVLKYKTRKKKAEN 179
tr|Q7NBU0|Q7NBU0_MYCGA  -----VSSFIEAKPEDRRQIFEDAAGIGRYAKKKQDAIR 181
sp|P22523|MUKB_ECOLI     IEASLYGGISSAITRSLRDYLLPENSGVRKAFQDMEAAALRENRTLEAIR 250
      : . : . : . * . *:. . : . * .

sp|P51834|SMC_BACSU      KLFETQDNLNRVEDILH----- 196
tr|Q7NBU0|Q7NBU0_MYCGA  QIARTNDNLKEITTVN----- 198
sp|P22523|MUKB_ECOLI     VTQSDRDLFKHLISEATNYVAADYMRHANERRVHLDKALEFRRELHTSRQ 300
      . * :.::

sp|P51834|SMC_BACSU      -----ELEGQVEPLK-IQASIAKDYLEKKKELEHVEIALTAY--- 232
tr|Q7NBU0|Q7NBU0_MYCGA  -----ELNRDLKKLN-QQAEKALYAETKEKLDLEITLSVN--- 234
sp|P22523|MUKB_ECOLI     QLAAEQYKHVDMARELAEHNGAEGDLEADYQAASDHLNLVQTALRQOEKI 350
      :: : : : :. . * . . *:. : : *

sp|P51834|SMC_BACSU      -----DIEELHGKWTSLKEKVQMAKEEELAESSAISAKEAKIEDTRDKIQ 277
tr|Q7NBU0|Q7NBU0_MYCGA  -----EYLISQKEIEALSEQIAEIDERLLKNDPQLQINQEKLEAFKKRYN 279
sp|P22523|MUKB_ECOLI     ERYEADLDELQIRLEEQNEVVAEAIERQQENEARAEAAELEVDELKSQLA 400
      : : . . * : * . . . . : : : :

sp|P51834|SMC_BACSU      ALDESVDELQVLLVTSEELEKLEGRKEVLKERKKNVQNQEQLE----- 322
tr|Q7NBU0|Q7NBU0_MYCGA  SADQNVQKIQDELQKIYDEIVLLE-KRNVFNDLQLKSDLDSNDKN----- 323
sp|P22523|MUKB_ECOLI     DYQQALDVQQTRAIQYQQAIAALNRAKELCHLPDLTADCAAEWLETFQAK 450
      : : : * : : * : : : . . : :

sp|P51834|SMC_BACSU      -----EAIVQFQKQETVLKEELSKQEAVFETLQAEVKQLRA---QVKEK 363
tr|Q7NBU0|Q7NBU0_MYCGA  -----KKINALEQLLKSSEEQLKQYFELISTWEEELKEK----- 357
sp|P22523|MUKB_ECOLI     ELEATEKMLSLEQKMSMAQTAHSQFEQAYQLVVAINGPLARNEAWDVARE 500
      : : : * . : . : .

sp|P51834|SMC_BACSU      QQALSLHNENVEEKIEQLKSDYFELLNSQASIRNELQLLDDQMSQSAVTL 413
tr|Q7NBU0|Q7NBU0_MYCGA  ----DVDKTDLANELENLKKSLATFQVKRYEANLQVQFYQNQKIN---- 398
sp|P22523|MUKB_ECOLI     LLREGVDQRHLAEQVQPLRMRLSELEQLRLEQQEAERLLADFCR----- 545
      . . : : : : * : : . . : :

sp|P51834|SMC_BACSU      QRLADNNEKHLQERHDISARKAACETEFARIEQEIHSQVGAYRDMQTKYE 463
tr|Q7NBU0|Q7NBU0_MYCGA  -----
sp|P22523|MUKB_ECOLI     QGKNFDIDELEALHQELEARIASLSDSVSNAREERMALRQEQEQQLSRIQ 595

sp|P51834|SMC_BACSU      QKKRQYEKNESALYQAYQYVQQARSKKDMLETMQGDFFSGFYQGVKEVLKA 513
tr|Q7NBU0|Q7NBU0_MYCGA  -----QFAQDAG-----VRTVLNN 412
sp|P22523|MUKB_ECOLI     SLMQRAPVWLAQNSLNLQLEQCGEEFTSSQDVTEYLLQQLLEREATIVE 645
      * :. : : : :

sp|P51834|SMC_BACSU      KERLGGIRGAVLELIS-----TEQ 532
tr|Q7NBU0|Q7NBU0_MYCGA  KDAIGGVHGIQDFIK-----VEP 431
sp|P22523|MUKB_ECOLI     RDEVGARKNAVDEEIERLSQPGGSEDQRLNALAERFGGVLLSEIYDDVSL 695
      : : * . . * : * .

sp|P51834|SMC_BACSU      KYETAIEIALGASAQHVVTDDQSARKAIQ----- 562
tr|Q7NBU0|Q7NBU0_MYCGA  EYELAISTALNKAANKNIIVDSNQDAINAVN----- 461
sp|P22523|MUKB_ECOLI     EDAPYFSAHYGPRHAIIVPDLQVTEHLEGLTDCPEDLYLIEGDPQSF 745
      : : . : : : . . . . : :

sp|P51834|SMC_BACSU      --YLKQNSFGRATFLPLSV-----IRDRQLQSRDAETA 593
tr|Q7NBU0|Q7NBU0_MYCGA  --FLKANKAGRATFLPLAN-----LKDRDVKPEHLEVL 492
sp|P22523|MUKB_ECOLI     DSVFSVDELEKAVVVKIADRQWRYSRFPEVPLFGRAARESRIESLHAERE 795
      : . : . * . : : : : : *

sp|P51834|SMC_BACSU      ARHSSFLGVASELVTDFPAYRSVIQNLGTVLITEDLKGANELAKLLGHR 643
tr|Q7NBU0|Q7NBU0_MYCGA  EQVEGYLGIAANLVNYHDQYDPAIRALLGQIIIASDLEAATKISKFTYQL 542
sp|P22523|MUKB_ECOLI     VLSERFATLSFDVQKTQRLHQAFSRFIGSHLAVAFESDPEAEIRQLNSRR 845
      : : : : . . : : : : : : : :

```

```

sp|P51834|SMC_BACSU      YRIVTLEGDVVNPGGG-----MTGGAVKKKK 669
tr|Q7NBU0|Q7NBU0_MYCGA YRVISLGGDIVNAGGA-----ITGGAESKQT 568
sp|P22523|MUKB_ECOLI    VELERALS NHENDNQOQRIQFEQAKEGVTALNRILPRLNLLADDSLADR 895
      .: .: * . : : : .:

sp|P51834|SMC_BACSU      NSLLGRSRELEDVTKRLAEMEEKTALLEQEVKTLKHSIQDMEKKLADLRE 719
tr|Q7NBU0|Q7NBU0_MYCGA HSLFN-----LDEKIDTLKN----- 583
sp|P22523|MUKB_ECOLI    DEIRERLDEAQAARFVQQFGNQLAKLEPIVSVLQSDPEQFEQLKEDYAY 945
      .: .: * : : : * :

sp|P51834|SMC_BACSU      TGEGLRLKQDVKGQLYELQVAEKNINHTHELYDQEKSALESSEDEERKVR 769
tr|Q7NBU0|Q7NBU0_MYCGA -----ELLVAEKNIN----- 593
sp|P22523|MUKB_ECOLI    S---QQMQRDARQQAFALTEVVQRRAHFYSYSDSAEMLSGNSDLNEKLRER 992
      * . : : . :

sp|P51834|SMC_BACSU      KRKLEEEELSAVSEKMKQLEEDIDRLTKQKQTSSTKESLSNELTELK--- 816
tr|Q7NBU0|Q7NBU0_MYCGA -----ELNKKLEFLTADYTKKDKFEQK--- 617
sp|P22523|MUKB_ECOLI    LEQAEAEERTRAREALRGHAAQLSQYNQVLASLKSYYDTKKELNLDLQREL 1042
      . : . : . : : :

sp|P51834|SMC_BACSU      --IAAAKKEQACKGEEDNLARLKKELTETELALKEAKEDLSFLTSEMSSS 864
tr|Q7NBU0|Q7NBU0_MYCGA --IAIQRYQDLIVIEQKKLDDYKIQY-----EQLTDTKTFDGKDV 654
sp|P22523|MUKB_ECOLI    QDIGVRADSGAEERARIRDELHAQLSNRRSRNQLKALTFCEAEMDNL 1092
      * . . . : : : * : : .

sp|P51834|SMC_BACSU      TSGEKLEEAHKHLDNDKTKTIELIALRRDQRIKQHGGLDYERELKEMK 914
tr|Q7NBU0|Q7NBU0_MYCGA KWDDKKIKDKLFSLETKKATLVQDLKINQEAQDMYQKQVNLQLEKDVTLFY 704
sp|P22523|MUKB_ECOLI    TRKLRKLERDYFEMREQVVTAKAGWCAVMRMVKDNGVERRLHRELAYLS 1142
      . * : : . . . : : :

sp|P51834|SMC_BACSU      RLYKQKTTLLKDEEVKLRMEVELDNLLQYLREEYSLSFEGAKEKY--QL 962
tr|Q7NBU0|Q7NBU0_MYCGA KEIDEDKNDKLRREQLTKHENTYLAQSKINESYNMAIEFAIENYNKPL 754
sp|P22523|MUKB_ECOLI    ADDLRMSDKALGALRLAVADNEHLRDVLRMSDDP-KRPERKIQFFVAVY 1191
      . . . : * : : * . * : :

sp|P51834|SMC_BACSU      ETDPEEARKRKLIKLAIEELGTVNLGSIDEFERNERYKFLSEQKE--- 1009
tr|Q7NBU0|Q7NBU0_MYCGA PISLSQARSEVVKLQSTLNNLGAINMEAIQELDIKKERYEKLYSQQQ--- 801
sp|P22523|MUKB_ECOLI    QHLRERIRQDIIRTDVPEAIEQMEIELSRLTEELTSREQKLAISSRSVA 1241
      . . * . : . : : : : : : : : * : * . .

sp|P51834|SMC_BACSU      -----DLTEAKNTLFQVIE 1023
tr|Q7NBU0|Q7NBU0_MYCGA -----ELINARERINQAI 815
sp|P22523|MUKB_ECOLI    NIIRKTIQREQNRIRMLNQGLQNVSVLNVNRETHAMLLDVLS 1291
      : : : : : : :

sp|P51834|SMC_BACSU      EMDEEMTKRFNDTFVQIRSHFDQVFRSLFGGGRaelR-----LTDPN 1065
tr|Q7NBU0|Q7NBU0_MYCGA RLDEKAI FEFDLINLNKELPKTFYLYLFGGNGCEIR-----YSNPE 857
sp|P22523|MUKB_ECOLI    EQHEQHQLDFNSNRLTFSEALAKLYQRINPQIDMGQRTPTIGEEELDYR 1341
      . * : * : . : : : : * * : .

sp|P51834|SMC_BACSU      DLLHSGVEIIAQPFGKQLNLLSGGERALTAIALLFSILK----- 1107
tr|Q7NBU0|Q7NBU0_MYCGA EKLTSGIEVFASPPGKNIGNLNLSSGGEKALVALSVLFSILK----- 899
sp|P22523|MUKB_ECOLI    NYLEMEVEVNRSGDWLRAESGALSTGEAIGTGMSILVMVQSWEDSRR 1391
      : * : * : . * : . * * . : : : * : :

sp|P51834|SMC_BACSU      -----VRVPFVCLDEVE--AALDEANVFRFAQYLKKYSSDTQFIVITHR 1150
tr|Q7NBU0|Q7NBU0_MYCGA -----VSSFPLVLDLAE--SALDLANVERFANIKNSSDQTFLIITHR 942
sp|P22523|MUKB_ECOLI    LRGKDISPCRLFLDEAARLDARSIALTFELCERLQMLIIAAPENISPE 1441
      : . : . * * . * . * : : : : : : : * : .

sp|P51834|SMC_BACSU      KGTMEADVLYGVTMQESGVSKVISVLEETKEFVQ----- 1186
tr|Q7NBU0|Q7NBU0_MYCGA EGTMVCKDKLIGATMQTKGVTKMLSVSLHQAQDMAEEIESQ---- 983
sp|P22523|MUKB_ECOLI    KGTTYKLVKRVFNTEHVHVGLRGFAPQLPETLPGTDEAPSQAS 1486
      : * : . : * : . . : :

```

17. Recombination protein RecR (gene name *recR*)

```

>sp|P0A7H6|RECR_ECOLI Recombination protein RecR OS=Escherichia coli (strain
K12) GN=recR PE=3 SV=1
MQTSPLLTQLMEALRCLPGVGPKSAQRMAFTLLQRDRSGGMRLAQALTRAMSEIGHCADC
RTFTTEQEVCNICSNPRRQENGQICVVESPADIIYAIEQTGQFSGRYFVLMGHLSPLDGIGP
DDIGLDRLEQRLAEEKITEVILATNPTEVEGEATANYIAELCAQYDVEASRIAHGVPVGGGE
LEMVDGTTLSHSLAGRHKIRF

```

```

>sp|P24277|RECR_BACSU Recombination protein RecR OS=Bacillus subtilis (strain
168) GN=recR PE=3 SV=2
MQYPEPISKLIDSFMKLPGIGPKTAVRLAFFVLGMKEDVVLDFAKALVNAKRNLTYCSVC
GHITDQDPCYICEDTRRDKSVICVVDQPKDVIAMEKMKYNGQYHVLHGAI SPMDGIGPE
DIKIPELLKRLQDDQVTEVILATNPNIEGEATAMYISRLKPSGIKLSRIAHGLPVGDDL
EYADEVTLSKALEGRREL
>tr|Q7NAN7|Q7NAN7_MYCGA Recombination protein RecR OS=Mycoplasma
gallisepticum (strain R(low / passage 15 / clone 2)) GN=recR PE=3 SV=2
MTSDLDLNEFNLLVEQISDLPSVSKKQAKKITQYLMTKSDRYVYDLIDVLKRAKLSIKIC
EMCQGWSNRSICSICSDESRRNNELCIVSFFDDLNVIEESQAYHGKYFILNHEISKKNKR
IIEEINFDLLLDLIKKQKIEKVIATNFTQDGQTTANYLRFLDDFDLKIYRLGMGLPYN
SSIDYADSFSLKGAFENKQLIKDKKA

sp|P0A7H6|RECR_ECOLI ---MQTSPLLTQLMEALRCLPGVGPKSAQRMAFTLLQRDRSGGMRLAQAL 47
sp|P24277|RECR_BACSU ---MQYPEPISKLIDSFMKLPGIGPKTAVRLAFFVLGMKEDVVLDFAKAL 47
tr|Q7NAN7|Q7NAN7_MYCGA MTSDLDLNEFNLLVEQISDLPSVSKKQAKKITQYLMTKSDRYVYDLIDVL 50
::*::: : **.. * * ::: : . : ..*

sp|P0A7H6|RECR_ECOLI TRAMSEIGHCADCRTFTEQEVCNICSNPRRQENGQICVVESPADIIAIEQ 97
sp|P24277|RECR_BACSU VNAKRNLTYCSVCGHITDQDPCYICEDTRR-DKSVICVVDQPKDVIAMEK 96
tr|Q7NAN7|Q7NAN7_MYCGA KRAKLSIKICEMCQGWSNRSICSICSDESRRNNELCIVSFFDDLNVIEE 99
.* .: * * ::: * **.: * :.. :*:.* * : :*:

sp|P0A7H6|RECR_ECOLI TGQFSGRYFVLMGHLSPLDGIGPDDIGLDRLEQRLAEEKITEVILATNPT 147
sp|P24277|RECR_BACSU MKEYNGQYHVLHGAI SPMDGIGPEDIKIPELLKRLQDDQVTEVILATNPN 146
tr|Q7NAN7|Q7NAN7_MYCGA SQAYHGKYFILNHEISKKNKRIIEEINFDLLLDLIKKQKIEKVIATNFT 149
: * :*:.* * : : :*: : * . : ::: :*:*** .

sp|P0A7H6|RECR_ECOLI VEGEATANYIAELCAQYDVEASRIAHGVPVGGGELEMVDGTTLSHSLAGRH 197
sp|P24277|RECR_BACSU IEGEATAMYISRLKPSGIKLSRIAHGLPVGGDLEYADEVTLSKALEGRR 196
tr|Q7NAN7|Q7NAN7_MYCGA QDGQTTANYLRFLDDFDLKIYRLGMGLPYNSSIDYADSFSLKGAFENKQ 199
:*:*** * : * ::: *:. *:* ..:: :.* :*. :. :.

sp|P0A7H6|RECR_ECOLI KIRF--- 201
sp|P24277|RECR_BACSU EL----- 198
tr|Q7NAN7|Q7NAN7_MYCGA LIKDKKA 206
:

```

18. Recombination protein RecO (gene name MGA_0016)

```

>sp|P0A7H3|RECO_ECOLI DNA repair protein RecO OS=Escherichia coli (strain
K12) GN=recO PE=1 SV=1
MEGWQRAFVLHSRPWSETSLMLDVFTEESGRVRLVAKGARSKRSTLKGALQPFTPLLLRF
GGRGEVKTLSRAEAVSALPLSGITLYSGLYINELLSRVLEYETRFSELFDFDYLHCIQSL
AGVTGTPEPALRRFELALLGHLGYGVNFTHCAGSGEPVDDTMTYRYREEKGFIASVVIDN
KTFTGRQLKALNAREFPDADTLRAAKRFRMALPKPYLGGKPLKSRELFRQFMPKRTVKTH
YE
>sp|P42095|RECO_BACSU DNA repair protein RecO OS=Bacillus subtilis (strain
168) GN=recO PE=1 SV=3
MLTKCEGIVLRTNDYGETNKIVTLLTREHGKIGVMARGAKKPNRSLSAVSQPFLYGSFLM
QKTSGLGLTQQGEMILSMRGIREDLFLTAYAAVAELVDRGTEEEKPNPYLFEFILESLK
QLNEGTDPDVITFIVQMKMLGVMGLYPELNHCVHCKSQDGTFFHFSVRDNGFICHRCFEKD
PYRIPIKPQTARLLRLFYFDLSRLGNVSLKEETKAELKQVIDLYEEYSGLYLKSKRFL
DQMESMKHLMGENKS
>tr|Q7NB87|Q7NB87_MYCGA Putative recombinational DNA repair protein RecO
OS=Mycoplasma gallisepticum (strain R(low / passage 15 / clone 2))
GN=MYCGA3920 PE=4 SV=1
MSSITKKGYYIIDYFDHNENDQIIKILFDDNTLTSLISVSGSKKILSKNGRYITLGS LHDFE
FFQARSIERLSKLLKKEIHEIDTKDATISESLPMVIMHYLLNKKSGELENNFFNFYDDVINY
VIKQRYSDETIIYILLNIINLEGIAFQLLNCGICNSKQVITLSFKKMYGLCEKCAYEQH
EFLYDKNFMRNIFWLIYKNDYEVSTLESRKYISLIKGLASAIYHNAGIYILEPVSFYLIK
K

sp|P42095|RECO_BACSU -MLTKCEGIVLRTNDYGETNKIVTLLTREHGKIGVMARGAKKPNRSLSAV 49
sp|P0A7H3|RECO_ECOLI -MEGWQRAFVLHSRPWSETSLMLDVFTEESGRVRLVAKGARSKRSTLKG 49
tr|Q7NB87|Q7NB87_MYCGA MSSITKKGYYIIDYFDHNENDQIIKILFDDNTLTSLISVSGSKKILSKNGRY 50
.. :. :*. . : : : : : : : : * : : . *

```

```

sp|P42095|RECO_BACSU      SQPFLYGSFLMQKTSGLGTLQQGEMILSMRGIREDLFLTAYAAVVAELVD 99
sp|P0A7H3|RECO_ECOLI     LQPFTPLLLRFQGRGEVKTLSRAEAVSLALPLSGITLYSG--LYINELLS 97
tr|Q7NB87|Q7NB87_MYCGA  ITLGSLHDFEFFQARSIERLSKLLKIHEDTKDATISESLPMVIMHYLYN 100
                          : :      : * . : :      :      : ..

sp|P42095|RECO_BACSU      RGTEEEKKPNPYLFEFILESLKQLNEGT-DPDVITFIVQMMLGVMGLYPE 148
sp|P0A7H3|RECO_ECOLI     RVLEYETRFSELFDDYLHCISLAGVTGTPEPALRRFELALLGHLGYGVN 147
tr|Q7NB87|Q7NB87_MYCGA  KKSGELENN--FFNFYDDVINYVIKQRYSDETIIYILLNIIINLEGIAFQ 148
                          :      : * . : :      :      : : . : : * :

sp|P42095|RECO_BACSU      LNHCVHCK-SQDGFHFSVRD-NGFICHRCFEKDPYRIPIKQPQTARLLRL 196
sp|P0A7H3|RECO_ECOLI     FTHCAGSGEPVDDTMTYRYREEKGFIVASVVIDN-----KTFTGRQLKA 190
tr|Q7NB87|Q7NB87_MYCGA  LLNCGICN--SKQVITLSFKKMYGLCEKCAEQHEF-----LYDKNFMRN 191
                          : : * . . . : . : * :      : :      : : :

sp|P42095|RECO_BACSU      FYYFDLSRLGNVSLKEETKAELKQVIDLYEEYSG-LYLKSKRFLDQMES 245
sp|P0A7H3|RECO_ECOLI     LNAREFP-----DADTLRAAKRFTRMALKPYLGGKPLKSRELFQFMP 233
tr|Q7NB87|Q7NB87_MYCGA  IFWLIYKN-----DYEVTLESRKYISLIKGLASA-IYHNAGIYLEPVFS 235
                          :      :      : : :      :      : : : . .

sp|P42095|RECO_BACSU      MKHLMGENKS 255
sp|P0A7H3|RECO_ECOLI     KRTVKTHYE- 242
tr|Q7NB87|Q7NB87_MYCGA  YLIKLK---- 241

```

19. Holliday junction resolvase (gene name *recU*)

Gene is absent in *E. coli* (strain K12) genome.

```

>sp|P39792|RECU_BACSU Holliday junction resolvase RecU OS=Bacillus subtilis
(strain 168) GN=recU PE=1 SV=1
MIRYPNGKTFQPKHSVSSQNSQKRAPSYSNRGMTLEDDLNETNKYYLTNQIAVIHKKPTP
VQIVNVHYPKRSAAVIKEAYFKQSSTTDYNGIYKGRYIDFEAKETKNKTSFPLQNFHDHQ
IEHMKQVKAQDQICFVVIISAFDQVYFLEADKLFYFWRKEKNGRKSIRKDELEETAYPIS
LGYAPRIDYISIEQLYFSPSSGAKG
>tr|Q7NBW8|Q7NBW8_MYCGA Holliday junction resolvase RecU OS=Mycoplasma
gallisepticum (strain R(low / passage 15 / clone 2)) GN=recU PE=3 SV=2
MHQNRGMFLETLINNTIKHNELAHKGLIFKRHLPIINVYNFANRRVTGWLKEKTQTDYYGL
YKGYFFDFDAKQSSKINYSLNKIQHQLDHLRKHIEQGGIAFILLLIVPKEEFYMIPIKK
IDSWLKNQESNTLKYEWIQKSSFKLELFYPGVIGIFEALQEWIDLITLRRSSGS

```

```

sp|P39792|RECU_BACSU      MIRYPNGKTFQPKHSVSSQNSQKRAPSYSNRGMTLEDDLNETNKYYLTNQ 50
tr|Q7NBW8|Q7NBW8_MYCGA  -----MHQNRGMFLETLINNTIKHNELAH 24
                          :.**** ** :*: * :

sp|P39792|RECU_BACSU      IAVIHKKPTPVQIVNVHYPKRSAAVIKEAYFKQSSTTDYNGIYKGRYIDF 100
tr|Q7NBW8|Q7NBW8_MYCGA  KGLIFKRHLPIINVYNFANRR-----VTGWLKEKTQTDYYGLYKGYFFDF 68
                          :*: *: * :      :      : : : : : * * : * : * :

sp|P39792|RECU_BACSU      EAKETKNKTSFPLQNFHDHQIEHMKQVKAQDQICFVVIISAF--DQVYFLE 148
tr|Q7NBW8|Q7NBW8_MYCGA  DAKQSS-KINYSLNKIQHQLDHLRKHIEQGGIAFILLLIVPKEEFYMIPI 117
                          :*: : * . : : * : : * : : : : * : * : : . : : * : :

sp|P39792|RECU_BACSU      ADKLFYFWRKEKNGRKSIRKDELEETAYPISLGYAPRIDYIS----IIE 194
tr|Q7NBW8|Q7NBW8_MYCGA  IKKIDSWLKNQESN---TLKYEWIQKSSFKLELFYPGVIGIFEALQEWID 164
                          . * : . : * * : : : : : : : : : * * . * . : * :

sp|P39792|RECU_BACSU      QLYFSPSSGAKG 206
tr|Q7NBW8|Q7NBW8_MYCGA  LITLRRSSGS-- 174
                          : : * * :

```

```

>sp|P0A814|RUV_C_ECOLI Crossover junction endodeoxyribonuclease RuvC
OS=Escherichia coli (strain K12) GN=ruvC PE=1 SV=2
MAIILGIDPGSRVTGYGVIRQVGRQLSYLGSQCIRTKVDDLPSRLKLIYAGVTEIITQFQ
PDYFAIEQVFMKNADSALKLQARGVAIVAAVNQELPVFEYAAARQVQKQTVVVGIGSAEKS
QVQHMVRTLLKLPANPQADAADALAIATHCHVSNAMQMSERLNLARGRLR

```

```

sp|P39792|RECU_BACSU      MIRYPNGKTFQPKHSVSSQNSQKRAPSYSNRGMTLEDDLNETNKYYLTNQ 49
tr|Q7NBW8|Q7NBW8_MYCGA  -----MHQNRGMFLETLINNTIKHNELA 23

```

```

sp|P0A814|RUVX_ECOLI -----MAILLGIDPGSRVTGYGVIRQVGR 24
                                     *:   :.

sp|P39792|RECU_BACSU      QIAVIHKKPTPVQIVNVHYPKRSAAVIKEAYFKQSSTTDYNGIYKGRYID 99
tr|Q7NBW8|Q7NBW8_MYCGA  HKGLIFKRHLPIVYNFANRR-----VTGWLKEKTQTDYYGLYKGYFFD 67
sp|P0A814|RUVX_ECOLI      QLSYLGSGCIRTKVDDLP SRLK-----LIYAGVTEIITQFQPDYFA 65
                             : . : .      : : : .      :      *:   : : : :

sp|P39792|RECU_BACSU      FEAKETKNKTSFPLQNFHDHQIEHMKQVKAQDGCIFVVISAF--DQVYFL 147
tr|Q7NBW8|Q7NBW8_MYCGA  FDAKQSS-KINYSLNKIKQHQLDHLRKHIEQGGIAFILLLLIVPKEEFYMI 116
sp|P0A814|RUVX_ECOLI      IEQVFMKNADSALKLQARGVAIVA AVNQELPVFEYAAARQVKQTVVGIG 115
                             : :      : . .* : : : : : : : : : . . :

sp|P39792|RECU_BACSU      EADKLFYFWRKEKNGRKSIRKDELEETAYPISLGYAPRIDYIS----II 193
tr|Q7NBW8|Q7NBW8_MYCGA  PIKKIDSWLKNQESN---TLKYEWIQKSSFKLELFYPGVIGIFEALQEWI 163
sp|P0A814|RUVX_ECOLI      SAEKSQ--VQHMVRT---LLKLPANPQADAADALAIAITHCHVSQNAMQM 160
                             .*      ..      .      : :      : :      * .      ..      :

sp|P39792|RECU_BACSU      EQLYFSPSSGAKG 206
tr|Q7NBW8|Q7NBW8_MYCGA  DLITLRRSSGS-- 174
sp|P0A814|RUVX_ECOLI      SESRLNLARGRLR 173
                             .      :      :      *

```

20. Holliday junction resolvase (gene name *MGA_0836*)

```

>sp|P0A811|RUVX_ECOLI Putative Holliday junction resolvase OS=Escherichia
coli (strain K12) GN=yqqF PE=1 SV=1
MSGTLLAFDFGTKSIGVAVGQRITGTARPLPAIKAQDGTDPWNI IERLLKEWQPDEIIVG
LPLNMDGTEQPLTARARKFANRIHGRFGVEVKLHDERLSTVEARSGLFEQGGYRALNKGK
VDSASAVIILESYFEQGY
>sp|O34634|RUVX_BACSU Putative Holliday junction resolvase OS=Bacillus
subtilis (strain 168) GN=yrrK PE=1 SV=1
MRILGLDLGKTKTLGVALSDEMGWTAQGIETIKINEAEGDYGLSRLSELIKDYTIDKIVLG
FPKNMNGTVGPRGEASQTFQVLETTYNVVPLWDERLTTMAAEKMLIAADVSRQKRKKV
IDKMAAVMILQGYLDSLN
>sp|Q7NBY7|RUVX_MYCGA Putative Holliday junction resolvase OS=Mycoplasma
gallisepticum (strain R(low / passage 15 / clone 2)) GN=MYCGA1230 PE=3 SV=1
MYVALDVGSRTLGIATGDGEFKIASPYCVISFNQYDFRQCLAE LKEKTASFFYDFKFVI
GMPKNIDQTKSSTTEMVENFIELLKANYKNEV IYDESYTSIIADQLLIDNQIKAKKRKE
KIDKLA AFVILQSFDDDDRYPK

sp|O34634|RUVX_BACSU      --MRILGLDLGKTKTLGVALSDEMGWTAQGIETIKINEAEGDYGLSRLSELIKDYTID-KI 57
sp|Q7NBY7|RUVX_MYCGA      --MYVALDVGSRTLGIATGDGEFKIASPYCVISFNQYDFRQCLAE LKEKTASFFYDFKF 58
sp|P0A811|RUVX_ECOLI      MSGTLLAFDFGTKSIGVAVGQRITGTARPLPAIKAQDGTDPWNI IER--LLKEWQPDEI 57
                             : : *. * : : : * . * : :      .      . : * : :

sp|O34634|RUVX_BACSU      VLGFPKNMNGTVGPRGEASQTFQVLETTYNVVPLWDERLTTMAAEKMLIAADVSRQKR 117
sp|Q7NBY7|RUVX_MYCGA      VIGMPKNIDQTKSSTTEMVENFIELLKANYKNEV IYDESYTSIIADQLLIDNQIKAKKR 118
sp|P0A811|RUVX_ECOLI      IVGLPLNMDGTEQPLTARARKFANRIHGRFGVEVKLHDERLSTVEARSGLFEQGGYRALN 117
                             : : * * : : * .      . * : : . : * : * * : : * . * :

sp|O34634|RUVX_BACSU      KKVIDKMAAVMILQGYLDSLN--- 138
sp|Q7NBY7|RUVX_MYCGA      KEKIDKLA AFVILQSFDDDDRYPK 142
sp|P0A811|RUVX_ECOLI      KGKVDASAVIILESYFEQGY--- 138
                             * . * . * : * * : : : .

```

21. DNA-polymerase IV (gene name *dinB*)

```

>sp|Q47155|DPO4_ECOLI DNA polymerase IV OS=Escherichia coli (strain K12)
GN=dinB PE=1 SV=1
MRKIIHVDMDCCFFAAVEMRDNPALRDIPIAIGGSRERRGVISTANY PARKFGVRSAMPTG
MALKLCPHLLTLLPGRF DAYKEASNHIREIFSRYSRIEPLSLDEAYLDVTD SVHCHGSAT
LIAQEIRQTI FNELQLTASAGVAPVKFLAKIASDMNKPNGQFVITPAEVP AFLQTLPLAK
IPGVGKVSAAKLEAMGLR TC GDVQKCDLVMLLKRFGKFGRI LWERSQGIDERDVNSERLR
KSVGVERTMAEDIHHWSECEAI IERLYPELEERRLAKVKPDL LIARQGVKLFDDDFQQTQ
EHVWPRLNKADLIATARKT WDERRGGRGVRLVGLHV TLLDPQMERQLVLGL
>sp|Q02886|DINB_BACSU Protein DinB OS=Bacillus subtilis (strain 168) GN=dinB
PE=3 SV=2
MSDFAFKLYEYNVWANQQIFNRLKELPK E IYHQEIQSVFPSISHVLSHVYLSDLGWIEVF
SGKTLSDALALAEQLKEQTEAKEIEEMEDLFLRLSERYILFLQQKEQLNKPLQIQNPSSG
IMKTTVSELLPHVNVNHGTYHRGNITAMLRQAGYASAPT DYGLYLFMTKTEKA

```

