

Московский Государственный Университет им. М.В. Ломоносова

Факультет Биоинженерии и Биоинформатики

**Регуляция дыхания энтеробактерий:
сопоставление данных по анализу экспрессии на биочипах и сравнительного-
геномного анализа**

Отчет по курсовой работе студентки
третьего курса Цыгановой Марины

Научный руководитель: Равчеев Д.А.

Консультант: д.б.н. проф. Гельфанд М.С.

Москва, 2006

Введение

Известно, что многие прокариотические организмы, например *Escherichia coli*, способны адаптироваться к различным условиям существования. В частности, наличие множества разнообразных дыхательных систем позволяет прокариотам использовать много разнообразных энергетических субстратов - доноров и акцепторов электронов.

Функция дыхательных систем. Функция дыхательных систем состоит в преобразовании потенциальной энергии протонного градиента в энергию химической связи в молекуле АТФ. В случае прокариотической клетки протонный градиент создается на уровне дыхательных цепей, локализованных в цитоплазматической мембране клетки. В процессе переноса электронов по дыхательным цепям энергия окислительно-восстановительных процессов преобразуется в потенциальную энергию протонного градиента, который формируется за счет разницы концентраций протонов между периплазмой и цитоплазмой, разделенных цитоплазматической мембраной. За синтез АТФ отвечает сложный ферментативный комплекс - АТФ-синтетаза, субъединицы которого в геноме *E.coli* кодируются генами оперона *atpIBEFHAGDC*.

Особенности строения дыхательных цепей бактерий. Главной особенностью прокариотических дыхательных цепей является их модульное строение: эти цепи состоят из отдельных компонентов, каждый из которых может быть заменен в соответствии с изменениями условий среды. В составе любой дыхательной цепи можно выделить три модуля:

- дегидрогеназы доноров электронов;
- хиноны, представляющие собой мембрано-растворимые переносчики электронов;
- редуктазы акцепторов электронов.

Дегидрогеназы и редуктазы представлены в мембране в виде белковых комплексов, отдельные компоненты которых кодируются, как правило, генами одного оперона (Gennis and Stewart, 1996). Разнообразие дыхательных цепей, необходимое для приспособления к условиям окружающей среды, достигается комбинацией различных дегидрогеназ и редуктаз. Разнообразие хинонов не столь велико,

и использование тех или иных хинонов зависит от присутствующих в мембране дегидрогеназ и редуктаз (Unden and Bongaerts, 1997).

Концентрация различных компонентов цепи зависит от условий роста и физиологических потребностей клетки. Например, в анаэробных условиях в присутствии нитрата преимущественно синтезируются нитрат-редуктазы. Они принимают электроны от хинонов и являются, таким образом, хинон-оксидазами, в то время как дегидрогеназы (в случае нитратного дыхания это преимущественно формиат-дегидрогеназы) являются хинон-редуктазами. Разнообразие дыхательных цепей для *E.coli* представлено на Рис.1.

Одним из свойств дыхательных цепей является наличие нескольких ферментов с одинаковой субстратной специфичностью (Gennis and Stewart, 1996). Например, *E.coli* способна синтезировать две различные НАДН-дегидрогеназы (НДГ): НДГ-I и НДГ-II, и две терминальные аэробные редуктазы, содержащие цитохром *bd* или *bo₃*. Все четыре фермента участвуют в формировании протонного градиента. Количество протонов, которое фермент может перенести из цитоплазмы в периплазму в расчете на один электрон, называют биоэнергетической эффективностью. В случае НДГ-II, кодируемой геном *ndh*, биоэнергетическая эффективность имеет значение $0 \text{ H}^+/\text{e}^-$, то есть данная дегидрогеназа не является протон-транслоцирующей (Рис.2а). В то же время, НДГ-I, субъединицы которой кодируются генами оперона *nuoABCDEFGHIJKLMN*, переносит два протона в расчете на один электрон (Рис.2б). Вышеупомянутые аэробные редуктазы различаются также по значениям биоэнергетической эффективности. Так, для редуктазы *bo₃*, кодируемой опероном *cyoABCDE*, биоэнергетическая активность составляет $2\text{H}^+/\text{e}^-$ (Рис. 2в), тогда как для редуктазы *bd*, кодируемой опероном *cydAB*, – всего $1\text{H}^+/\text{e}^-$ (Рис. 2д). Таким образом, количество протонов, которое может быть перенесено в пределах этой системы, колеблется от четырех до одного. Путем комбинации этих четырех элементов возможна тонкая настройка системы в зависимости от условий среды и физиологических потребностей клетки (Calhoun et al, 1993) (Рис.3). Например, при низкой концентрации кислорода в среде необходимо снижение внутриклеточного содержания молекулярного кислорода. Это связано с тем, что в данных условиях предпочтительно

функционируют чувствительные к кислороду белки, отвечающие за ассимиляцию альтернативных акцепторов электронов. Защита этих белков обеспечивается работой цитохром-*bd* содержащей оксидазы, которая обладает очень высоким сродством к кислороду и поэтому значительно снижает содержание кислорода в клетке, тем самым предотвращая нарушение функционирования чувствительных белков (Hill et al, 1990). Биоэнергетическая эффективность цитохром-*bd* содержащей оксидазы составляет только $1\text{H}^+/\text{e}^-$, поэтому предпочтительнее в качестве дегидрогеназы использовать НДГ-I, обладающую максимальной биоэнергетической эффективностью.

Регуляция дыхания гамма-протеобактерий. В ряде экспериментальных исследований было показано существование определенной иерархии в использовании различных акцепторов электронов. Например, в присутствии кислорода анаэробные пути метаболизма подавляются, так как кислород является наилучшим акцептором электронов, поскольку имеет наибольшее значение окислительно-восстановительного потенциала. В анаэробных условиях наиболее выгодным акцептором является нитрат, и в его присутствии подавляются все другие анаэробные пути. Подобная иерархия осуществляется за счет регуляции на уровне транскрипции соответствующих редуктаз ферментами, реагирующими на присутствие различных акцепторов электронов (Unden and Bongaerts, 1997). Предпочтительность использования более электроположительных акцепторов объясняется высоким энергетическим выходом реакций с участием этих веществ (Табл.1).

Так, восстановление акцепторов с более высоким окислительно-восстановительным потенциалом позволяет протонным насосами использовать большую химическую энергию для транспорта протонов через мембрану.

В зависимости от акцептора перестраивается не только дыхательная цепь, но и значительная часть катаболизма, включая центральные метаболические пути. Использование акцептора и донора связано со значением окислительно-восстановительного потенциала ΔE° (ОВП) для реакции их восстановления. Если ОВП имеет положительное значение, то реакция идет в сторону восстановления, если отрицательное - в сторону окисления. Чем выше абсолютное значение ОВП, тем выгоднее использование этого донора или акцептора.

Факторы транскрипции, осуществляющие глобальную регуляцию дыхания в *E.coli*. Как уже говорилось, энтеробактерии являются факультативными анаэробами. Так, в случае *E.coli* и родственных микроорганизмов, переключение между аэробным и анаэробным типами обмена осуществляется несколькими регуляторами транскрипции (Iuchi and Weiner, 1996). Воздействию этих регуляторов подвергаются не только гены терминальных редуктаз, но также гены дегидрогеназ доноров электронов и ряд других генов, задействованных в процессах дыхания и клеточного метаболизма.

Реакция на молекулярный кислород в *E.coli* осуществляется в основном двумя регуляторными системами: белком Fnr и двухкомпонентной системой ArcB-ArcA, в которой ArcB – сенсорная киназа, а ArcA – регулятор транскрипции. FNR и ArcA могут служить как активаторами, так и репрессорами, в зависимости от расположения сайтов их связывания относительно промоторов регулируемого оперона.

Белок Fnr - это глобальный регулятор, осуществляющий переключение между аэробным и анаэробным метаболизмом. В анаэробных условиях Fnr осуществляет активацию генов анаэробного дыхания и репрессию генов аэробного метаболизма. Этот белок состоит из сенсорного и ДНК-связывающего доменов. Сигнал от сенсорного к ДНК-связывающему домену передается посредством внутримолекулярных конформационных изменений, в результате которых в отсутствие кислорода возможно образование гомодимеров, способных связываться с ДНК (Melville *et al.*, 1996). Инактивация Fnr кислородом включает в себя окисление и разрушение железо-серных кластеров. Такой кластер связан с четырьмя консервативными остатками цистеина. Под действием кислорода три цистеиновых остатка становятся открытыми, что вызывает частичное разрушение кластера (Рис.4). В таком состоянии Fnr не способен к димеризации и находится в неактивной форме (Trageser and Uden, 1991). Был показано, что сайт связывания Fnr представляет собой инвертированный повтор длиной четырнадцать пар нуклеотидов (п.н) со спейсером, состоящим из четырех п.н. (Герасимова и др, 2001).

ArcA является важнейшим регулятором аэробного метаболизма и репрессором генов аэробного метаболизма в анаэробных условиях. Он входит в состав двухкомпонентной регуляторной системы ArcA-ArcB. В анаэробных условиях ArcB, убихинон-зависимая трансмембранная киназа чувствительная к присутствию кислорода, автофосфорилируется, а затем трансфосфорилирует ArcA (Georgellis *et al.*, 1999), в результате чего последний становится способным связываться с ДНК (Рис.5). В аэробных же условиях ArcB действует как ArcA-Р-специфичная фосфатаза, катализируя дефосфорилирование ArcA-Р (Georgellis *et al.*, 1998). Чувствительность ArcB к содержанию кислорода осуществляется за счет взаимодействия с хинонами (Malpica *et al.*, 2004). Интенсивность потока электронов через электрон-транспортную сеть зависит от вида акцептора, преимущественно содержащегося в среде. Если условия аэробные, то этот поток будет наиболее интенсивным. Хинон, как связующее звено электрон-транспортной цепи, в свою очередь регистрирует изменение интенсивности потока электронов при недостатке кислорода в среде и становится способным принимать электроны от цистеинов ArcB. Образующиеся в результате окисления цистенинов дисульфидные мостики меняют конформацию ArcB так, что тот становится неспособным к проявлению киназной активности. Сайт связывания ArcA представляет собой нестрогий прямой повтор (Favorov *et al.*, 2005). В системе регуляции дыхания *E.coli* белок ArcA занимает подчиненное положение относительно Fnr: известно, что транскрипция гена *arcA* регулируется последним (Comran and Touati, 1994). Этот процесс называется Fnr-ArcA регуляторным каскадом. Некоторые гены, участвующие в процессе дыхания, находятся под совместной регуляцией Fnr и ArcA (Unden *et al.*, 1995). Лого для сайтов Fnr и ArcA представлены на Рис.6.

Методы исследования регуляции дыхания. Известно, что как в осуществлении процессов дыхания, так и в его регуляции, участвует большое количество белков. Экспериментальный анализ отдельно каждого белка и соответствующего гена представляет собой чрезвычайно трудоемкий процесс. Поэтому для исследования дыхания необходимо использование методов, позволяющих получать, хранить и обрабатывать значительные массивы данных.

Метод биочипов. Долгое время в молекулярной биологии для изучения процессов, происходящих в организме, исследовались одиночные гены. Однако, организм – сложная система взаимосвязанных процессов. Поэтому в настоящее время для исследования экспрессии генов часто применяется метод биочипов, позволяющий одновременно получить информацию о работе большого числа генов и обработать эту информацию. Таким образом, исследователь имеет возможность оценивать организм как целостную систему, а не как сумму его составляющих. Дизайн эксперимента в общем случае состоит в следующем: в заданные точки стеклянной подложки помещаются образцы олигонуклеотидов, соответствующие исследуемым генам, затем из культуры клеток, выращенных в определенных условиях, выделяется фракция РНК. Полученная смесь РНК, так называемый “транскриптом”, предварительно обрабатывается флуоресцирующими красителями и наносится на подложку, в результате чего РНК гибридизуется с соответствующими олигонуклеотидами, иммобилизованными на биочипе. Готовый чип сканируется лазером для измерения интенсивности сигнала флуоресценции каждой точки. В результате по изменению уровня флуоресценции можно судить об изменениях экспрессии исследуемых генов. Данный метод предоставляет данные о составе модулона. Модулон регулятора – это группа генов, каждый из которых имеет существенно разный уровень экспрессии в диком штамме и штамме, мутантном по гену этого регулятора. Данный метод характеризуется достаточно большим процентом шума, недостоверных данных, так как процесс обработки и анализа сигналов лазерного сканирования очень сложен. Кроме того, этот процесс очень сложен технологически, так всегда существует некоторое фоновое излучение от РНК, которая не связалась с ДНК, но тем не менее осталась на подложке после промывания (Hatfield *et al.*, 2003).

Биоинформатический метод. Помимо экспериментальных технологий, в настоящее время широко используются биоинформатические методы исследований. Они дают результат значительно быстрее по сравнению с экспериментальными методами и активно развиваются. Предсказания, сделанные данными методами, облегчают работу экспериментального биолога, поскольку указывают на потенциальный объект для исследования. Применительно к регуляции, биоинформатический метод состоит в поиске потенциальных регуляторных сайтов. Так, с конца девяностых годов в

экспериментальных работах после установления факта регуляции поиск сайтов проводился преимущественно биоинформатическими методами. Сравнительно-геномный подход, заключающийся в проверке консервативности потенциальных сайтов связывания регуляторов, значительно повышает качество предсказаний. В случае если перед ортологичными генами в родственных геномах сохраняются потенциальные сайты, ген считается принадлежащим к регулону. Непрерывный процесс увеличения числа геномов значительно расширяет возможности сравнительной геномики как в количественном, так и качественном отношении (Gelfand *et al.*, 2000).

Регуляторные каскады и их предположительная роль в расширении модулонов. Синтез большинства генов, белки которых участвуют в различных биологических процессах, например в процессе дыхания, контролируется более чем одним регулятором, которые формируют сложные системы регуляторных каскадов, и, действуя в различных комбинациях, могут вызывать активацию или репрессию определенных генов (Unden and Bongaerts, 1997).

Системы регуляторных каскадов построены по иерархическому принципу. Они состоят из последовательно подчиненных регуляторов и разветвляется от регулятора уровня I, так называемого глобального регулятора, подобно дереву, каждый узел которого – регулятор уровня n. Если регулятор n-го уровня влияет на экспрессию каких-либо K генов, то регулятор уровня n-1 не только регулирует экспрессию своих генов, но и опосредованно влияет на экспрессию этих K генов. Таким образом, можно ожидать, что в состав модулона регулятора более высокого уровня будут входить не только гены, напрямую регулируемые этим регулятором, то есть члены регулона, но и гены, находящиеся под влиянием регуляторов более низких уровней. Если двигаться по этой схеме от самых низких уровней к первому, модулон каждого последующего уровня будет шире предыдущего, а модулон глобального регулятора будет помимо соответствующего регулона включать модулоны всех уровней. То есть ожидается, что регуляторные каскады вносят вклад в расширение модулона по сравнению с соответствующим регулоном.

Однако существуют и другие причины расширения модулонов. Так, если регулятор транскрипции, регулятор I, регулирует метаболический путь синтеза вещества, которое является

эффектором для другого регулятора транскрипции, регулятора II, то регулон регулятора II будет входить в модулон I.

В нашей работе мы рассматривали только роль регуляторных каскадов в расширении модулонов.

Цели и задачи

Целью данной работы является оценка соотношения генов в сопряженных модулоне и регулоне. Сопряженными называются модулон и регулон одного регулятора, например говорят о сопряженных модулонах и регулонах регулятора Fnr. Как было сказано выше, модулон представляет собой группу генов, уровень экспрессии которых зависит от некоторого регулятора. Регулон же - группа генов, непосредственно регулируемых некоторым регулятором. Интуитивно понятно, что в состав модулона должно входить большее число генов, чем в регулон. Этому явлению существует два объяснения: наличие экспериментального «шума», который, как правило, присутствует в данных, полученных с помощью метода биочипов, и, как уже было сказано, возможное влияние регуляторных каскадов. Определение роли регуляторных каскадов в расширении модулона по отношению к соответствующему регулону также является целью нашей работы.

В настоящей работе были поставлены следующие задачи:

1. Обработка экспериментальных данных о влиянии белков Fnr (Salmon *et al*, 2003) и ArcA (Salmon *et al*, 2005) на экспрессию генов *E.coli*, полученных с помощью биочипов. Выделение Fnr и ArcA модулонов.

2. Выделение Fnr и ArcA регулонов на основе:

- анализа литературных данных об исследовании регуляции экспрессии отдельных оперонов различными молекулярно-биологическими методами; как правило, такие исследования дают более точные результаты, чем биочипы.

- сравнительно-геномного анализа, проведенного в данной работе (см. Материалы и методы).

3. Сопоставление полученных результатов, сравнение модулонов и регулонов.

Материалы и методы

Данная работа построена на сравнительном анализе экспериментально полученных с помощью метода биочипов данных и результатов сравнительной геномики.

Исследованные геномы. Полные последовательности: *Escherichia coli* K12 (EC) (Blattner *et al.*, 1997), *Salmonella typhi* Ty2 (STY) (Deng *et al.*, 2003), *Salmonella typhimurium* LT2 (STM) (McClelland *et al.*, 2001), *Yersinia pestis* KIM (YP) (Deng *et al.*, 2002), *Pectobacterium carotovorum* subsp. *atroseptica* SCRI1043 (PCA) (Bell *et al.*, 2004), *Photobacterium luminescens* subsp. *laumondii* TTO1 (PHL) (Duchaud *et al.*, 2003) были взяты из базы данных GenBank (<http://www.ncbi.nlm.nih.gov/Genbank>).

Предварительные последовательности геномов: *Yersinia enterocolitica* (YE), *Serratia marcescens* Db11 (SM), *Photobacterium asymbiotica* ATCC43949 (PHA), были взяты с интернет-сайта Wellcome Trust Sanger Institute (<http://www.sanger.ac.uk/>). Предварительная последовательность *Pectobacterium chrysanthemi* str. 3937 (PCH), была взята с интернет-сайта The Institute For Genomic Research (<http://www.tigr.org>). Во всех этих последовательностях были найдены ортологи белков Fng и ArcA.

Выделение модулонов. В качестве экспериментальных данных о составе модулона были взяты результаты двух исследований. В обоих исследованиях был использован метод биочипов для выявления влияния белков-регуляторов Fng (Salmon *et al.*, 2003) и ArcA (Salmon *et al.*, 2005) на уровень экспрессии большого числа генов *E. coli*.

В указанных работах приводились значения изменения интенсивности экспрессии генов. По значениям изменения уровня экспрессии генов эффекты каждого фактора были разделены на 3 группы:

- активация – при увеличении уровня экспрессии от двух раз;
- отсутствие эффекта – при увеличении до двух раз и уменьшении до двух раз;
- репрессия – при уменьшении уровня экспрессии от двух раз.

Гены, уровень экспрессии которых менялся в два и более раз, были отнесены к модулону.

Гены, уровень экспрессии которых изменялся менее, чем в два раза, в состав модулона не

включались. Подобный критерий разделения является общепринятым в области изучения регуляции экспрессии генов (Neighart, 1996).

На следующем этапе работы на основе этой информации были выделены опероны, данные об изменении экспрессии которых можно считать противоречивыми. Данные считались противоречивыми, если для разных генов в пределах одного оперона указывались противоположные эффекты одного регулятора, или для одного из генов оперона указывалось отсутствие эффекта в отличие от других генов этого оперона.

Поиск потенциальных сайтов связывания. Для поиска сайтов связывания был применен метод позиционных весов нуклеотидов (Гельфанд и Миронов, 1999). В работе использовались матрицы для поиска сайтов связывания белков Fnr (Gerasimova *et al*, 2001) и для ArcA (Favorov *et al.*, 2005), ранее полученные в нашей лаборатории. Поиск сайтов производился в потенциальных регуляторных областях генов – в пределах -400 ...+100 п.н. относительно старта трансляции гена. Для гена *cydA* было экспериментально показано наличие двух сайтов связывания ArcA на расстояниях -637 п.н. и -719 п.н. от старта трансляции. Поэтому для данного гена поиск потенциальных ArcA-сайтов проводился в пределах -750...+100 п.н. относительно старта трансляции гена. Сайты считались значимыми, если их вес превышал пороговое значение: 3.50 для Fnr-сайтов и 4.25 для ArcA-сайтов. В случае Fnr потенциальный сайт регуляции был обнаружен в регуляторной области 15% генов *E.coli*, в случае ArcA – в регуляторной области 7% генов. В этой ситуации снижается возможность недопредсказания.

Определение состава регулона. Для определения принадлежности гена к регулону был использован метод проверки соответствия. В его основе лежит сравнительный анализ геномных последовательностей родственных организмов. Ген считался принадлежащим к регулону, если сайт перед ним сохранялся более чем в трех родственных геномах. Однако если сайт был обнаружен только перед генами в *Escherichia coli* K12, *Salmonella tiphimurium* LT2 и *Salmonella typhi* Ty1, ген не включался в регулон, поскольку степень родства между данными организмами очень высока, и

сохранение сайтов может являться артефактом, вызванным высокой степенью сходства некодирующих областей.

Выделение регулонов. В данной работе мы несколько модифицировали критерий, по которому классифицировали ген как принадлежащий или не принадлежащий к регулону. Ген считался членом регулона, если он входил в состав модулона и, одновременно с этим, имел в регуляторной области консервативный сайт. То есть описанный подход базируется на совмещении экспериментальных данных и данных сравнительной геномики.

Программное обеспечение. Для поиска ортологичных генов и потенциальных сайтов использовался пакет программ “Genome Explorer” (Миронов *и др.*, 2000).

Результаты и обсуждение.

Определение членов модулона. На основе результатов двух исследований, в которых был использован метод биочипов для выявления влияния белков-регуляторов Fnr и ArcA на уровень экспрессии генов *E.coli*, были выделены Fnr- и ArcA-модулоны. В общей сложности в двух исследованиях предоставляются данные об изменении уровня экспрессии в два и более раз для 174 оперонов. Из них, по сформулированному в разделе “Материалы и методы” критерию были выделены опероны, данные об изменении экспрессии которых противоречивы. В эту группу вошли 4 оперона (Таб.2).

В пределах этой группы можно выделить две подгруппы оперонов:

- данные о влиянии регулятора Fnr противоречивы – 3 оперона (*ylcBCD-ybdE*, *rpsJ-rplCDWB-rpsS-rolV-rpsC-rplP-rpmC-rpsQ*, *rplNXE-rpsNH-rplFR-rpsE-rpmD-rplO-prlA-rpmJ-rpsMKD-rpoA-rplQ*);
- данные о влиянии регулятора ArcA противоречивы – 3 оперона (*nuoABCDEFGHJKLMN*, *rpsJ-rplCDWB-rpsS-rolV-rpsC-rplP-rpmC-rpsQ*, *rplNXE-rpsNH-rplFR-rpsE-rpmD-rplO-prlA-rpmJ-rpsMKD-rpoA-rplQ*).

При этом перед опероном *ylcBCD-ybdE* был найден консервативный сайт для Fnr, а перед опероном *tnoABCDEFGHIJKLMN*, гены которого кодируют субъединицы НДГ-I, - как для Fnr, так и для ArcA. Кроме того, для оперона *tnoABCDEFGHIJKLMN* был экспериментально показан сайт связывания белка ArcA (Bongaerts et al, 1995), что согласуется с нашими результатами.

Несмотря на противоречия данных микрочипов, мы включили оперон *tnoABCDEFGHIJKLMN* в состав Fnr и ArcA модулонов, руководствуясь литературными данными, и оперон *ylcBCD-ybdE* – в Fnr модулон.

Кроме того, по данным Salmon *et al*, 2005 оперон *cydAB* не относится к ArcA модулону, хотя экспериментально было показано наличие сайта связывания ArcA для этого оперона (Lynch and Lin, 1996). Поэтому оперон *cydAB* мы также включили в ArcA модулон.

Из оставшихся после удаления противоречий 172 оперонов были выделены 151 оперонов, относящихся к Fnr-модулону, и 135 оперона, относящихся к ArcA-модулону (Табл.3).

Поиск потенциальных сайтов Fnr и ArcA перед генами *E.coli*. Для каждого гена, входящего в Fnr-модулон или ArcA-модулон, был произведен поиск сайтов связывания Fnr и ArcA (см. Материалы и методы). В результате потенциальный сайт связывания Fnr был обнаружен перед 38 оперонами, а сайт связывания ArcA – перед 23 опероном (Таб. 4, 5).

Процедура проверки найденных потенциальных сайтов связывания Fnr и ArcA на консервативность методом проверки соответствия показала, что в 27 случаях из 38 для Fnr и в 15 случаях из 23 для ArcA сайт сохраняется перед ортологичными генами как минимум в трех родственных геномах (Таб. 4, 5).

Данные опероны с высокой вероятностью входят в соответствующие регулоны, а сайты перед ними консервативны. Для этих оперонов в 15 из 41 случаев сайт сохраняется перед ортологичными генами в восьми, девяти или десяти родственных геномах, в 17 случаях сайт сохраняется в пяти и более родственных геномах, и только в 10 случаях сайт сохраняется в четырех или трех родственных геномах.

Так как эта процедура проверки потенциальных сайтов на консервативность проводилась для оперонов, которые являются членами Fnr-модулона и ArgA-модулона по данным биочипов, то все опероны, перед которыми был найден такой сайт, можно отнести к Fnr-регулону или ArgA-регулону.

Структура Fnr-регулона и ArgA-регулона представлена в Табл.4 и Табл.5.

Среди оперонов Fnr-регулона и ArgA-регулона может быть выделена группа из десяти оперонов, которые входят одновременно в Fnr-регулон и ArgA-регулон (Табл.2), так называемое “ядро”. Опероны, составляющие эту группу, характеризуются рядом особенностей. Сайты перед этими оперонами имеют наибольшие из возможных веса. Так, сайт связывания Fnr перед геном *ndh* имеет вес 5.03, в остальных случаях веса сайтов связывания Fnr значительно превышают значение 3.75, либо приближаются к этому значению. Что касается сайтов связывания ArgA, то только в случае оперона *atpIBEFHAGDC* вес ArgA-сайта ниже 4.75, все же остальные опероны имеют в своих регуляторных областях сайты связывания ArgA с весом значительно превышающим эту величину и достигающими, в случае *mdh*, значения 5.19.

Практически все сайты перед рассматриваемыми оперонами высококонсервативны. Так, сайт связывания Fnr перед оперонами *pdhR-aceEF-lpdA*, *ndh*, *atpIBEFHAGDC* и сайт связывания ArgA перед оперонами *pdhR-aceEF-lpdA*, *icdA*, *cydAB* сохраняются в десяти родственных геномах, а сайт связывания ArgA перед оперонами *ndh* сохраняется в девяти родственных геномах.

Кроме того, большое количество экспериментальных данных указывает на существование или возможность существования предсказанных нами сайтов. Регуляция белком Fnr была ранее показана для оперонов *pdhR-aceEF-lpdA* (Quail *et al*, 1994), *dcuB-fumB* (Golby *et al*, 1998), *ndh* (Green and Guest, 1994) и *cydAB* (Cotter and Gunsalus, 1992), а модуляция белком Fnr – для оперонов *cyoABCDE* (Cotter and Gunsalus, 1992), *icdA* (Chao *et al*, 1997), *gltA* (Park *et al*, 1997). Регуляция белком ArgA была показана для оперонов *nuoABCDEFGHIJKLMN* (Bongaerts *et al*, 1995), *cyoABCDE* (Tseng *et al*, 1996), *ndh* (Liu and Wulf, 2004), *icdA* (Chao *et al*, 1997), *gltA* (Lynch and Lin, 1996), в то время как модуляция этим регулятором – для оперонов *pdhR-aceEF-lpdA* (Quail *et al*, 1994), *dcuB-fumB* (Golby *et al*, 1998), *mdh* (van der Rest *et al*, 2000).

Рассматриваемые опероны участвуют в ключевых этапах процесса дыхания (Рис.7). Они были разделены нами на следующие группы.

1. Гены НАДН-дегидрогеназы II и I: *ndh* (Calhoun and Gennis, 1993) и *nnoABCDEFGHIJKLMN* (Weidner *et al*, 1993). Подробно функции НАДН-дегидрогеназы II и I рассматривается в разделе “Введение”.
2. Гены аэробных редуктаз, содержащих цитохромы *bo* и *bd*: соответственно, *cyoABCDE* и *cydAB* (Cherugi *et al*, 1990). Функция этих редуктаз также подробно описана в разделе “Введение”.
3. Гены метаболизма четырехуглеродных дикарбоновых кислот: *dcuB-fumB* (Bell *et al*, 1989; Six *et al*, 1993). Продукты этих генов – ферменты и белки-транспортеры, которые поставляют субстраты для фумаратного дыхания. Кроме того, FumB участвует в цикле Кребса.
4. Ген цикла трикарбоновых кислот: *mdh* (McAlister-Henn *et al*, 1987), *icdA* (Oshima *et al*, 1996), *gltA* (Ner *et al*, 1983).
5. Гены субъединиц АТФ-синтетазы: *atpIBEFHAGDC* (Kanazawa *et al*, 1982).
6. Гены, обеспечивающие регуляцию экспрессии генов, отвечающих за реакцию образования ацетил-кофермента А из пирувата: *pdhR-aceEF-lpdA* (Haydon *et al*, 1993; Stephens *et al*, 1983).

Кроме рассмотренных выше десяти оперонов “ядра” в состав Fnr-регулона входят еще 16 оперонов (Таб.4). Регуляция белком Fnr была ранее показана для двух таких оперонов: *narXL* (Takahashi *et al*, 1994), *nirBDC-cysG* (Jayaraman *et al*, 1988).

Эти 17 оперонов были разделены на три группы (Табл.6):

1. Гены, участвующие в процессе дыхания: *narXL*, *glgCAP*, *nirBDC-cysG*, *pstSCAB-phoU*.
2. Гены, не участвующие в процессе дыхания: *tgt-yajC-secD-secF*, *purMN*, *thrS-infC-rpmI-rplT*, *rpsT*, *iadA*, *nanA-nanT*, *upp-uraA*, *ibpB-yidE*, *rhaBA*, *metBL*, *aroP*.
3. Гены, функция которых неизвестна или экспериментально не определена: *ybhK*.

Аналогично существует 4 оперона ArcA-регулона, которые не входят в состав “ядра” (Табл.5).

Для оперона *lpdA* экспериментально была показана регуляция белком ArcA (Cunningham and Guest, 1998). Рассматриваемые опероны были также разделены на три группы (Табл.6):

1. Ген, участвующий в процессе дыхания: *gapA*.
2. Гены, не участвующие в процессе дыхания: *acrEF*.
3. Гены, функция которых неизвестна или экспериментально не определена: *yigI-rarD-yigG-yigF, yadF*.

Опероны, которые не имеют в своей регуляторной области консервативного сайта по формальному критерию не могут быть отнесены к регулону. Таких оперонов в случае Fnr – семь (Табл.4), а в случае ArcA – пять (Табл.5). Замечено, что, данные опероны не относятся к оперонам центрального метаболизма, а для некоторых из них функция вовсе не известна (Табл.6). Кроме того, практически все сайты перед этими генами имеют низкий вес, менее 3.75 для Fnr и менее 4.75 для ArcA. Экспериментальных данных, подтверждающих модуляцию или прямую регуляцию этих оперонов Fnr или ArcA, так же не существует. Таким образом, нет противоречий, препятствующих исключению этой группы оперонов из соответствующих потенциальных регулонов.

В случае, когда гены оперона имеют ортологи только в *E.coli*, *S. typhimurium* и *S. typhi*, мы не можем говорить о принадлежности этих генов к регулонам. В случае Fnr было обнаружено четыре таких оперона (Табл.4), в случае ArcA – два (Табл.5).

Соотношение сопряженных модулонов и регулонов. Вклад регуляторных каскадов в расширение модулонов. Как уже было отмечено, системы регуляторных каскадов построены по иерархическому принципу. В случае дыхания Fnr является глобальным регулятором, который напрямую или опосредованно влияет на экспрессию не только генов ферментов дыхания, но и генов, кодирующих регуляторы более низкого уровня, например ArcA. ArcA таким образом является регулятором второго уровня. Ожидалось, что Fnr-модулон помимо генов Fnr-регулона будет включать гены ArcA-модулона. Для определения роли регуляторного каскада Fnr-ArcA в расширении Fnr-модулона по сравнению с Fnr-регулоном мы сопоставили полученные нами данные о составе

регулона с данными метода биочипов и составили схему, визуализирующую соотношение Fnr-регулона, Fnr-модулона, ArcA-регулона и ArcA-модулона (Рис. 8). Как было сказано, к Fnr-регулону нами было отнесено 26 оперонов. Следовательно, остальные 124 оперона не находятся под прямой Fnr-регуляцией. В тоже время, 90 из этих оперонов входят в ArcA-модулон. Эти соотношения иллюстрируют идею расширения модулона регулятора более высокого уровня за счет модулона регулятора более низкого уровня. 20 оперонов из ArcA-модулона не входят в Fnr-модулон. Скорее всего, это является следствием высокого процента «шума» в данных биочипов.

Известно, что регулятор нитрат-нитритного дыхания *narXL* регулируется белком Fnr.

Ожидается, что он будет вносить вклад в расширение Fnr-модулона как регулятор второго уровня.

Для того чтобы проследить вклад регуляторов третьего и более уровней в расширение Fnr- и ArcA-модулонов, мы обратились к литературным данным. Экспериментально было установлено, что белок ArcA контролирует экспрессию гена *lctR*, кодирующего регулятор третьего уровня, который регулирует экспрессию генов оперона *lctPRD*, отвечающих за процесс молочно-кислого брожения (Lynch and Lin, 1996). Ожидается, что эти гены будут входить модулоны Fnr и ArcA. Однако, в работах Salmon *et al*, 2003 и Salmon *et al*, 2005 экспрессия данных генов не исследовалась.

Кроме того, нельзя исключить возможность расширения модулонов за счет неизвестных регуляторов. Так, в геноме *E.coli* было предсказано наличие 150 белков-регуляторов транскрипции, точная функция которых неизвестна (Browning and Busby, 2004).

Выводы

1. На примере Fnr- и ArcA-регуляторных систем *E.coli* приведено сравнение данных микрочипов и результатов сравнительной геномики.
2. Выделены Fnr- и ArcA-регулоны энтеробактерий, обнаружено 20 новых членов для Fnr-регулона и 9 новых членов для ArcA-регулона
3. Показана роль Fnr-ArcA-регуляторного каскада в расширении Fnr-регулона.
4. Выделен ряд генов, так называемое “ядро”, для которых характерно сохранение Fnr- и ArcA регуляции в эволюции.

Список литературы:

- Allen S.P., Polazzi J.O., Gierse J., Easton A.M., 1992. **Two novel heat shock genes encoding proteins produced in response to heterologous protein expression in *Escherichia coli*.** *J. Bacteriol.* **174**; 6938-6947.
- Amemura M., Makino K., Shinagawa H., Kobayashi A., Nakata A., 1985. **Nucleotide sequence of the genes involved in phosphate transport and regulation of the phosphate regulon in *Escherichia coli*.** *J. Mol. Biol.* **184**; 241-250.
- Anderson P.S., Smith J.M., Mygind B., 1992. **Characterization of the *upp* gene encoding uracil phosphoribosyltransferase of *Escherichia coli* K12.** *Eur. J. Biochem.* **204**; 51-56.
- Aronson B.D., Ravnkar P.D., Somerville R.L., 1988. **Nucleotide sequence of the 2-amino-3-ketobutyrate coenzyme A ligase (*kbl*) gene of *E. coli*.** *Nucleic Acids Res.* **16**; 3586-3586.
- Baecker P.A., Furlong C.E., Preiss J., 1983. **Biosynthesis of bacterial glycogen. Primary structure of *Escherichia coli* ADP-glucose synthetase as deduced from the nucleotide sequence of the *glgC* gene.** *J. Biol. Chem.* **258**; 5084-5088.
- Bell K.S., Sebahia M., Pritchard L., Holden M.T., Hyman L.J., Holeva M.C., Thomson N.R., Bentley S.D., Churcher L.J., Mungall K., Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J., Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H., Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S., Salmond G.P., Birch P.R., Parkhill J., Toth I.K., 2004. **Genome**

sequence of the enterobacterial phytopathogen *Erwinia carotovora* subsp. *atroseptica* and characterization of virulence factors. *Proc. Natl. Acad. Sci. U. S. A.* **101**; 11105-11110.

Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B. and Shao Y., 1997. **The complete genome sequence of *Escherichia coli* K-12.** *Science.* **277**; 1453-1474.

Bell P.J., Andrews S.C., Sivak M.N., Guest J.R., 1989. **Nucleotide sequence of the Fnr-regulated fumarase gene (*fumB*) of *Escherichia coli* K-12.** *J. Bacteriol.* **171**; 3494-3503.

Boehm R., Sauter M., Boeck A., 1997. **Nucleotide sequence and expression of an operon in *Escherichia coli* coding for formate hydrogenlyase components.** *Science* **277**; 1453-1474.

Bongaerts J., Zoske S., Weidner U., Uden G., 1995. **Transcriptional regulation of the proton translocating NADH dehydrogenase genes (*nuoA-N*) of *Escherichia coli* by electron acceptors, electron donors and gene regulators.** *Mol Microbiol.* **16**; 521-34.

Branlant G., Branlant C., 1985. **Nucleotide sequence of the *Escherichia coli* *gap* gene. Different evolutionary behavior of the NAD⁺-binding domain and of the catalytic domain of D-glyceraldehyde-3-phosphate dehydrogenase.** *Eur. J. Biochem.* **150**; 61-66.

Browning D.F., Busby S.J., 2004. **The regulation of bacterial transcription initiation.** *Nat Rev Microbiol.* **2**; 57-65. Review.

Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R., 1993. **DNA sequence and analysis of 136 kilobases of the *Escherichia coli* genome: organizational symmetry around the origin of replication.** *Genomics* **16**; 551-561.

Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L., Blattner F.R., 1995. **Analysis of the *Escherichia coli* genome VI: DNA sequence of the region from 92.8 through 100 minutes.** *Nucleic Acids Res.* **23**; 2105-2119.

- Cain B.D., Simoni R.D., 1989. **Proton translocation by the F1F0ATPase of *Escherichia coli*. Mutagenic analysis of the a subunit.** *J Biol Chem.* **264**; 3292-300.
- Calhoun M. W., Oden K. L., Gennis R. B., Teixeira de Mattos M. J., and Neijssel O. M., 1993. **Energetic efficiency of *Escherichia coli*: effects of mutations in components of the aerobic respiratory chain.** *J. Bacteriol.* **175**; 3020–3025.
- Calhoun M.W., Gennis R.B., 1993. **Demonstration of separate genetic loci encoding distinct membrane-bound respiratory NADH dehydrogenases in *Escherichia coli*.** *J Bacteriol.* **175**; 3013-9.
- Chao G.C., Shen J., Tseng C. P., Park S. J., Gunsalus R. P., 1997. **Aerobic regulation of isocitrate dehydrogenase gene (*icd*) expression in *Escherichia coli* by the *arcA* and *fnr* gene products.** *J Bacteriol.* **179**; 4299–4304.
- Chepuri V., Lemieux L., Au D.C.T., Gennis R.B., 1990. **The sequence of the *cyo* operon indicates substantial structural similarities between the cytochrome o ubiquinol oxidase of *Escherichia coli* and the aa3-type family of cytochrome c oxidases.** *J. Biol. Chem.* **265**; 11185-11192.
- Choi Y.-L., Kawamukai M., Utsumi R., Sakai H., Komano T., 1989. **Molecular cloning and sequencing of the glycogen phosphorylase gene from *Escherichia coli*.** *FEBS Lett.* **243**; 193-198.
- Compan I., Touati D., 1994. **Anaerobic activation of *arcA* transcription in *Escherichia coli*: roles of Fnr and ArcA.** *Mol Microbiol.* **11**; 955-64.
- Cotter P.A., Gunsalus R.P., 1992. **Contribution of the *fnr* and *arcA* gene products in coordinate regulation of cytochrome o and d oxidase (*cyoABCDE* and *cydAB*) genes in *Escherichia coli*.** *FEMS Microbiol Lett.* **70**; 31-36.
- Cunningham L., Guest J.R., 1998. **Transcription and transcript processing in the *sdhCDAB-sucABCD* operon of *Escherichia coli*.** *Microbiology.* **144**; 2113-2123.
- Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R., 1992. **Analysis of the *Escherichia coli* genome: DNA sequence of the region from 84.5 to 86.5 minutes.** *Science* **257**; 771-778.

Deng W., Burland V., Plunkett G., Boutin A., Mayhew G.F., Liss P., Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C., Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V., Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R., Perry R.D., 2002. **Genome sequence of *Yersinia pestis* KIM.** *J. Bacteriol.* **184**; 4601-4611.

Deng W., Liou S.R., Plunkett G., Mayhew G.F., Rose D.J., Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R., 2003. **Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2 and CT18.** *J. Bacteriol.* **185**; 2330-2337.

Duchange N., Zakin M.M., Ferrara P., Saint-Girons I., Park I., Tran S.V., Py M.-C., Cohen G.N., 1983. **Structure of the *metJBLF* cluster in *Escherichia coli* K12. Sequence of the *metB* structural gene and of the 5'- and 3'-flanking regions of the *metBL* operon.** *J. Biol. Chem.* **258**; 14868-14871.

Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A., Taourit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.F., Dassa E., Deroose R., Derzelle S., Freyssinet G., Gaudriault S., Medigue C., Lanois A., Powell K., Siguier P., Vincent R., Wingate V., Zouine M., Glaser P., Boemare N., Danchin A., Kunst F., 2003. **The genome sequence of the entomopathogenic bacterium *Photorhabdus luminescens*.** *Nat. Biotechnol.* **21**; 1307-1313.

Favorov A.V., Gelfand M.S., Gerasimova A.V., Ravcheev D.A., Mironov A.A., Makeev V.J., 2005. **A Gibbs sampler for identification of symmetrically structured, spaced DNA motifs with improved estimation of the signal length.** *Bioinformatics.* **21**(10); 2240-5.

Fayat G., Mayaux J.-F., Sacerdot C., Fromant M., Springer M., Grunberg-Manago M., Blanquet S., 1983. ***Escherichia coli* phenylalanyl-tRNA synthetase operon region. Evidence for an attenuation mechanism. Identification of the gene for the ribosomal protein L20.** *J. Mol. Biol.* **171**; 239-261.

Gardel C., Johnson K., Jacq A., Beckwith J., 1990. **The *secD* locus of *E.coli* codes for two membrane proteins required for protein export.** *EMBO J.* **9**; 3209-3216.

- Gay N.J., 1984. **Construction and characterization of an *Escherichia coli* strain with a *uncl* mutation.** *J Bacteriol*; **158**; 820-5.
- Gelfand M.S., Koonin E.V., Mironov A.A., 2000. **Prediction of transcription regulatory sites in Archaea by a comparative genomic approach.** *Nucleic Acids Res.* **28**; 695-705.
- Gennis R. B. and Stewart V., 1996. **Respiration.** In Neidhart F.C. (eds.) *Escherichia coli* and *Salmonella*. Cellular and Molecular Biology, ASM Press, Washington , pp. 217-286.
- Georgellis D., Kwon O., De Wulf P., Lin E.C., 1998. **Signal decay through a reverse phosphorelay in the Arc two-component signal transduction system.** *J Biol Chem.* **273**; 32864-9.
- Gerasimova A.V., Ravcheyev D.A., Gelfand M.S., Rakhmaninova A.B, 2004. **Comparative Genomic Analysis of Respiration switch in Gamma-Proteobacteria.** Proceedings for *The Fourth International Conference on Bioinformatics of Genome Regulation and Structure.* **2**; 195-198.
- Gill D.R., Hatfull G.F., Salmond G.P.C., 1986. **A new cell division operon in *Escherichia coli*.** *Mol. Gen. Genet.* **205**; 134-145.
- Golby P., Kelly D.J., Guest J.R., Andrews S.C., 1998. **Transcriptional regulation and organization of the *dcuA* and *dcuB* genes, encoding homologous anaerobic C4-dicarboxylate transporters in *Escherichia coli*.** *J Bacteriol.* **180**; 6586-96.
- Green J. and Guest J.R., 1994. **Regulation of transcription at the *ndh* promoter of *Escherichia coli* by Fnr and novel factors.** *Mol Microbiol*; **12**; 433-44.
- Green G.N., Fang H., Lin R.-J., Newton G., Mather M., Georgiou C.D., Gennis R.B., 1988. **The nucleotide sequence of the *cyd* locus encoding the two subunits of the cytochrome d terminal oxidase complex of *Escherichia coli*.** *J. Biol. Chem.* **263**; 13138-13143.

Haydon D.J., Quail M.A., Guest J.R., 1993. **A mutation causing constitutive synthesis of the pyruvate dehydrogenase complex in *Escherichia coli* is located within the *pdhR* gene.** *FEBS Lett.* **336**; 43-47.

Hatfield G. Wesley, Hung She-pin and Baldi Pierre, 2003. **Differential of DNA microarray gene expression data.** *Molecular Microbiology.* **47**; 871-877.

Hill, S., Viollet S., Smith A. T., and Anthony C., 1990. **Roles for enteric d-type cytochrome oxidase in N₂ fixation and microaerobiosis.** *J. Bacteriol.* **172**; 2071–2078.

Honore N., Cole S.T. , 1990. **Nucleotide sequence of the *aroP* gene encoding the general aromatic amino acid transport protein of *Escherichia coli* K-12: homology with yeast transport proteins.** *Nucleic Acids Res.* **18**; 653-653.

Iuchi S. and Weiner L., 1996. **Cellular and molecular physiology of *Escherichia coli* in the adaptation to aerobic environments.** *J Biochem (Tokyo).* **120**;1055-63. Review.

Jayaraman P.S., Gaston K.L., Cole J.A., Busby S.J., 1988. **The *nirB* promoter of *Escherichia coli*: location of nucleotide sequences essential for regulation by oxygen, the Fnr protein and nitrite.** *Mol Microbiol.* **2**; 527-30.

Liu Xueqiao and Wulf Peter, 2004. **Probing the ArcA-P modulon of *Escherichia coli* by whole genome transcriptional analysis and sequence recognition profiling.** *J Biochem.* **279**; 12588-12597.

Kanazawa H., Mabuchi K., Futai M., 1982. **Nucleotide sequence of the promoter region of the gene cluster for proton-translocating ATPase from *Escherichia coli* and identification of the active promoter.** *Biochem. Biophys. Res. Commun.* **107**; 568-575.

Kitagawa Y., Akaboshi E., Shinagawa H., Horii T., Ogawa H., Kato T., 1985. **Structural analysis of the *umu* operon required for inducible mutagenesis in *Escherichia coli*.** *Proc. Natl. Acad. Sci. U.S.A.* **82**; 4336-4340.

Kumar A., Larsen C.E., Preiss J., 1986. **Biosynthesis of bacterial glycogen. Primary structure of *Escherichia coli* ADP-glucose:alpha-1,4-glucan, 4-glucoyltransferase as deduced from the nucleotide sequence of the *glgA* gene.** *J. Biol. Chem.* **261**; 16256-16259.

Leveque F., Plateau P., Dessen P., Blanquet S., 1990. **Homology of *lysS* and *lysU*, the two *Escherichia coli* genes encoding distinct lysyl-tRNA synthetase species.** *Nucleic Acids Res.* **18**; 305-312.

Lynch A.S. and Lin E.C., 1996. **Transcriptional control mediated by the ArcA two-component response regulator protein of *Escherichia coli*: characterization of DNA binding at target promoters.** *J. Bacteriol.* **178**; 6238-49.

Mackie G.A., 1981. **Nucleotide sequence of the gene for ribosomal protein S20 and its flanking regions.** *J. Biol. Chem.* **256**; 8177-8182.

Magota K., Otsuji N., Miki T., Horiuchi T., Tsunasawa S., Kondo J., Sakiyama F., Amemura M., Morita T., Shinagawa H., Nakata A., 1984. **Nucleotide sequence of the *phoS* gene, the structural gene for the phosphate-binding protein of *Escherichia coli*.** *J. Bacteriol.* **157**; 909-917.

Malpica R., Franco B., Rodriguez C., Kwon O., Georgellis D., 2004. **Identification of a quinone-sensitive redox switch in the ArcB sensor kinase.** *Proc Natl Acad Sci U S A.* **101**; 13318-23.

Martinez J., Steenbergen S., Vimr E., 1995. **Derived structure of the putative sialic acid transporter from *Escherichia coli* predicts a novel sugar permease domain.** *J. Bacteriol.* **177**; 6005-6010.

Mayaux J.-F., Fayat G., Fromant M., Springer M., Grunberg-Manago M., Blanquet S., 1983. **Structural and transcriptional evidence for related *thrS* and *infC* expression.** *Proc. Natl. Acad. Sci. U.S.A.* **80**; 6152-6156.

McAlister-Henn L., Blaber M., Bradshaw R.A., Nisco S.J., 1987. **Complete nucleotide sequence of the *Escherichia coli* gene encoding malate dehydrogenase.** *Nucleic Acids Res.* **15**; 4993-4993.

McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P., Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D., Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K., 2001.

Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature.* **413**; 852-856

Melville S.B. and Gunsalus R.P., 1996. **Isolation of an oxygen-sensitive Fnr protein of *Escherichia coli*: interaction at activator and repressor sites of Fnr-controlled genes.** *Proc Natl Acad Sci U S A.* **93**; 1226-1231.

Merlin C., Masters M., McAteer S., Coulson A., 2003. **Why is carbonic anhydrase essential to *Escherichia coli*?** *J Bacteriol.***185**; 6415-6424

Murakami S., Nakashima R., Yamashita E., Yamaguchi A. 2002. **Crystal structure of bacterial multidrug efflux transporter AcrB.** *Nature.* **419** (6907); 587-593.

Neighart F.C. (eds.), 1996. ***Escherichia coli* and *Salmonella*. Cellular and Molecular Biology.** ASM Press, Washington.

Moralejo P., Egan S.M., Hidalgo E.F., Aguilar J., 1993. **Sequencing and characterization of a gene cluster encoding the enzymes for L-rhamnose metabolism in *Escherichia coli*.** *J. Bacteriol.* **175**; 5585-5594.

Navarro C., Wu L.-F., Mandrand-Berthelot M.-A., 1993. **The *nik* operon of *Escherichia coli* encodes a periplasmic binding-protein-dependent transport system for nickel.** *Mol. Microbiol.* **9**; 1181-1191.

Ner S.S., Bhayana V., Bell A.W., Giles I.G., Duckworth H.W., Bloxham D.P., 1983. **Complete sequence of the *glcA* gene encoding citrate synthase in *Escherichia coli*.** *Biochemistry.* **22**; 5243-5249.

Nohno T., Noji S., Taniguchi S., Saito T., 1989. **The *narX* and *narL* genes encoding the nitrate-sensing regulators of *Escherichia coli* are homologous to a family of prokaryotic two-component regulatory genes.** *Nucleic Acids Res.* **17**; 2947-2957.

Nobelmann B., Lengeler J.W., 1995. **Sequence of the *gat* operon for galactitol utilization from a wild-type strain EC3132 of *Escherichia coli*.** *Biochim. Biophys. Acta* **1262**; 69-72.

Ohta Y., Watanabe K., Kimura A., 1985. **Complete nucleotide sequence of the *E. coli* N-acetylneuraminate lyase.** *Nucleic Acids Res.* **13**; 8843-8852.

Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A., Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K., Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N., Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M., Horiuchi T., 1982. **A 718-kb DNA sequence of the *Escherichia coli* K-12 genome corresponding to the 12.7-28.0 min region on the linkage map.** *DNA Res.* **3**(3):137-55.

Peakman T., Crouzet J., Mayaux J.F., Busby S.J.W., Mohan S., Harborne N., Wootton J., Nicolson R., Cole J.A., 1990. **Nucleotide sequence, organisation and structural analysis of the products of genes in the *nirB-cysG* region of the *Escherichia coli* K-12 chromosome.** *Eur. J. Biochem.* **191**; 315-323.

Park S.J., Chao G., Gunsalus R.P., 1997. **Aerobic regulation of the *sucABCD* genes of *Escherichia coli*, which encode alpha-ketoglutarate dehydrogenase and succinyl coenzyme A synthetase: roles of *ArcA*, *Fnr*, and the upstream *sdhCDAB* promoter.** *J Bacteriol.* **179**; 4138-42.

Park S.J., Cotter P.A., Gunsalus R.P., 1995. **Regulation of malate dehydrogenase (*mdh*) gene expression in *Escherichia coli* in response to oxygen, carbon, and heme availability.** *J Bacteriol.* **177**: 6652-6.

Pistocchi R., Kashiwagi K., Miyamoto S., Nukui E., Sadakata Y., Kobayashi H., Igarashi K., 1993. **Characteristics of the operon for a putrescine transport system that maps at 19 minutes on the *Escherichia coli* chromosome.** *J. Biol. Chem.* **268**; 146-152.

Quail M.A., Haydon D.J., Guest J.R., 1994. **The *pdhR-aceEF-lpd* operon of *Escherichia coli* expresses the pyruvate dehydrogenase complex.** *Mol Microbiol.* **12**; 95-104.

van der Rest M.E., Frank C., Molenaar D., 2000. **Functions of the membrane-associated and cytoplasmic malate dehydrogenases in the citric acid cycle of *Escherichia coli*.** *J Bacteriol.* **182**; 6892-6899.

Reuter K., Slany R., Ullrich F., Kersten H., 1991. **Structure and organization of *Escherichia coli* genes involved in biosynthesis of the deazaguanine derivative queuine, a nutrient factor for eukaryotes.** *J. Bacteriol.* **173**; 2256-2264.

Sacerdot C., Fayat G., Dessen P., Springer M., Plumbridge J.A., Grunberg-Manago M., Blanquet S., 1982. **Sequence of a 1.26-kb DNA fragment containing the structural gene for *E.coli* initiation factor IF3: presence of an AUU initiator codon.** *EMBO J.* **1**; 311-315.

Salmon K., Hung S. P., Mekjian K., Baldi P., and Hatfield G. W., 2003. **Global gene expression profiling in *Escherichia coli* K12. The effects of oxygen availability and Fnr.** *J. Biol. Chem.* **278**; 29837-29855.

Salmon K. A., Hung S. P., Steffan N. R., Krupp R., Baldi P., Hatfield G. W., and Gunsalus R. P., 2005. **Global gene expression profiling in *Escherichia coli* K12: The effects of oxygen availability and ArcA.** *J. Biol. Chem.* **280**; 15084-15096.

Six S., Andrews S.C., Roberts R.E., Uden G., Guest J.R., 1993. **Construction and properties of *Escherichia coli* mutants defective in two genes encoding homologous membrane proteins with putative roles in anaerobic C4-dicarboxylic acid transport.** *Biochem. Soc. Trans.* **21**; 342S-342S.

Smith J.M., Daum H.A. III, 1986. **Nucleotide sequence of the *purM* gene encoding 5'-phosphoribosyl-5-aminoimidazole synthetase of *Escherichia coli* K12.** *J. Biol. Chem.* **261**; 10632-10636.

Stephens P.E., Darlison M.G., Lewis H.M., Guest J.R., 1983. **The pyruvate dehydrogenase complex of *Escherichia coli* K12. Nucleotide sequence encoding the pyruvate dehydrogenase component.** *Eur. J. Biochem.* **133**; 155-162.

Stephens P.E., Lewis H.M., Darlison M.G., Guest J.R., 1983. **Nucleotide sequence of the lipoamide dehydrogenase gene of *Escherichia coli* K12.** *Eur. J. Biochem.* **135**; 519-527.

Takahashi K., Hattori T., Nakanishi T., Nohno T., Fujita N., Ishihama A., Taniguchi S., 1994. **Repression of in vitro transcription of the *Escherichia coli* *fnr* and *nar X* genes by FNR protein.** *FEBS Lett.* **340**; 59-64.

Trageser M. and Unden G., 1991. **Oxygen regulated gene expression in *Escherichia coli*: control of anaerobic respiration by the Fnr protein.** *Antonie Van Leeuwenhoek.* **59**; 65-76.

Unden G. and Bongaerts J., 1997. **Alternative respiratory of *Escherichia coli*: energetics and transcriptional regulation in response to electron acceptors.** *Biochimica et Biophysica Acta.* **1320**; 217-234.

Tseng C.P., Albrecht J., Gunsalus R.P., 1996. **Effect of microaerophilic cell growth conditions on expression of the aerobic (*cyoABCDE* and *cydAB*) and anaerobic (*narGHJI*, *frdABCD*, and *dmsABC*) respiratory pathway genes in *Escherichia coli*.** *J Bacteriol.* **178**; 1094-8.

Unden G., Becker S., Bongaerts J., et al, 1995. **O₂-sensing and O₂-dependent gene regulation in facultatively anaerobic bacteria.** *Arch Microbiol.* **164**; 81-90.

Wada A., Sako T., 1987. **Primary structures of and genes for new ribosomal proteins A and B in *Escherichia coli*.** *J. Biochem.* **101**; 817-820.

Weidner U., Geier S., Ptock A., Friedrich T., Leif H., Weiss H., 1993. **The gene locus of the proton-translocating NADH: ubiquinone oxidoreductase in *Escherichia coli*. Organization of the 14 genes and relationship between the derived proteins and subunits of mitochondrial complex I.** *J. Mol. Biol.* **233**(1); 109-122.

Woods S.A., Schwartzbach S.D., Guest J.R., 1988. **Two biochemically distinct classes of fumarase in *Escherichia coli*.** *Biochim Biophys Acta.* **954**; 14-26.

Xu J., Bertrand K.P. **Nucleotide sequence of the *acrEF* operon from *Escherichia coli*.** *Submitted (MAY-1993) to the EMBL/GenBank/DDBJ databases.*

Jiang Y., Yang F., Zhang X., Yang J., Chen L., Yan Y., Nie H., Xiong Z., Wang J., Dong J., Xue Y., Xu X., Zhu Y., Chen S., Jin Q., 2005. **The complete sequence and analysis of the large virulence plasmid pSS of *Shigella sonnei*.** *Plasmid* **54**; 149-159.

Young J.G., Rogers B.L., Campbell H.D., Jaworowski A., Shaw D.C, 1981. **Nucleotide sequence coding for the respiratory NADH dehydrogenase of *Escherichia coli*. UUG initiation codon.** *Eur. J. Biochem.* **116**; 165-170.

Zakin M.M., Duchange N., Ferrara P., Cohen G.N., 1983. **Nucleotide sequence of the *metL* gene of *Escherichia coli*. Its product, the bifunctional aspartokinase ii-homoserine dehydrogenase II, and the bifunctional product of the *thrA* gene, aspartokinase I-homoserine dehydrogenase I, derive from a common ancestor.** *Biol Chem.* **258**; 3028-31

Гельфанд М.С., Миронов А.А. ,1999. **Компьютерный анализ регуляторных сигналов в полных бактериальных геномах.** *Молекулярная биология.* **33**; 772-778.

Герасимова А.В., Родионов Д.А., Миронов А.А., Гелфанд М.С., 2001. **Компьютерный анализ регуляторных сигналов в бактериальных геномах. Участки связывания Fnr.** *Мол Биол (Моск).* **35**; 1001-1009.

Гусев М.В., Минеева Л.Ф., 1992. **Микробиология. Издательство Московского Государственного Университета. Москва.**

Табл.1. Таблица окислительно-восстановительных потенциалов для акцепторов и доноров электронов (Гусев и Минеева, 1992). * - ОВП FeS-кластера зависит от аминокислотного окружения

Реакция	ΔE°
Пируват/ацетат + CO ₂	-700
H ⁺ / ¹ / ₂ H ₂	-420
НАД(Ф) ⁺ /НАД(Ф)·H ₂	-320
S ⁰ /HS ⁻	-270
SO ₄ ²⁻ /HS ⁻	-220
ФАД/ФАД·H ₂	-220
ФМН/ФМН·H ₂	-190
Менахинон окисл/восст	-74
Рубедоксин окисл/восст	-57
Ферредоксины Fe ₂ S ₂ -типа окисл/восст	-225 ... -455*
Ферредоксины Fe ₄ S ₄ -типа окисл/восст	+350 ... -480*
Фумарат/сукцинат	+30
Цитохром <i>b</i> окисл/восст	+70
Убихинон окисл/восст	+100
Цитохром <i>c</i> окисл/восст	+220
Цитохром <i>a</i> окисл/восст	+290
NO ₃ ⁻ /NO ₂ ⁻	+433
Fe ³⁺ /Fe ²⁺	+772
¹ / ₂ O ₂ /H ₂ O	+820

Табл.2. Противоречивые данные о влиянии регулятора на уровень экспрессии генов. Обозначения: “-”

– репрессия, “0” – нет эффекта, “+” – активация, *- нет сайта.

Опероны	Эффект Fnr (Salmon et al,2003)	Эффект ArcA (Salmon et al, 2005)
Оперон <i>ylcBCD-ybdE</i>		
<i>ybdE</i>	+	+
<i>ylcD</i>	+	+
<i>ylcB</i>	0	Не исследовался
Оперон <i>nuoABCDEFGHIJKLMN</i>		
<i>nuoE</i>	--	--
<i>nuoB</i>	Не исследовался	-
<i>nuoF</i>	Не исследовался	0
<i>nuoG</i>	Не исследовался	0
Оперон <i>rplCDWB-rpsS-rolV-rpsC-rplP-rpmC-rpsQ</i>		
<i>rpsJ</i>	-	-
<i>rplB</i>	0	0
<i>rplC</i>	0	0
<i>rplD</i>	0	0
<i>rpmC</i>	+	+
Оперон <i>rplNXE-rpsNH-rplFR-rpsE-rpmD-rplO-prlA-rpmJ-rpsMKD-rpoA-rplQ</i>		
<i>prlA</i>	-	-
<i>rplX</i>	Не исследовался	-
<i>rplO</i>	0	0
<i>rpsE</i>	0	0
<i>rplQ</i>	0	0
<i>rpsH</i>	0	0
<i>rplE</i>	0	0
<i>rplF</i>	Не исследовался	0

Табл.3. Сопоставление Fnr- и ArcA- модулонов и регулонов. Условные обозначения: “+” – регулятор активирует экспрессию гена, “-” – регулятор репрессирует экспрессию оперона, “0” – регулятор не влияет на экспрессию оперона, “#” – данные по модуляции противоречивы. Серым цветом отмечены гены, входящие в регулон, зеленым – не входящие. ¹⁾ – в данном опероне регуляция происходит в трех промоторах.

Оперон	Эффект Fnr	Fnr-регулон	Эффект ArcA	ArcA-регулон
<i>xylE</i>	+		Не исследовался	
<i>mhpABCDFE</i>	+		Не исследовался	
<i>b2878</i>	+		Не исследовался	
<i>fdhF</i>	+		Не исследовался	
<i>yhiUV</i>	+		Не исследовался	
<i>caiTABCDE</i>	+		Не исследовался	
<i>yjiZ</i>	+		Не исследовался	
<i>yjeH</i>	+		Не исследовался	
<i>araBAD</i>	+		Не исследовался	
<i>yhcL</i>	+		Не исследовался	
<i>fliLMNOPQR</i>	+		Не исследовался	
<i>nadC</i>	+		Не исследовался	
<i>yjeT</i>	+		Не исследовался	
<i>b2438</i>	+		Не исследовался	
<i>aat</i>	+		Не исследовался	
<i>yhjRQONML</i>	+		0	
<i>ybeRS</i>	+		0	
<i>cmtBA-yggP-b2931-yggFDC</i>	+		0	
<i>b3050-b3051</i>	+		0	
<i>sbm-ygfDGH</i>	+		0	
<i>yjeMNO</i>	+		0	
<i>proVWX</i>	+		0	
<i>thiMD-yegX</i>	-		0	
<i>rpoN-yhbH-ptsN-yhbJ</i>	-		0	
<i>ycbPONME</i>	+		0	
<i>emrKY</i>	+		0	
<i>parC</i>	+		0	
<i>yfbTS</i>	+		0	
<i>gcvTH</i>	+		0	
<i>yceL-grxB</i>	+		0	
<i>dfp-dut-ttk</i>	+		0	
<i>aroKB-damX-dam-rpe-gph-trpS</i>	+		0	
<i>dadAX</i>	+		0	
<i>ylcBCD-ybdE</i>	+		#	
<i>queA-tgt-yajC-secD-secF</i>	-		0	
<i>narXL</i>	+		0	
<i>purMN</i>	-		0	
<i>metBL</i>	+		+	
<i>thrS-infC-rpmI-rplT</i>	-		-	
<i>rpsT</i>	-		-	
<i>yjiHG-iadA-yjiE</i>	+		+	
<i>nanA-nanT</i>	+		+	
<i>upp-uraA</i>	+		+	

<i>ibpB-yidE</i>	+		+	
<i>glgCAP</i>	+		+	
<i>ybhK</i>	+		+	
<i>nirBDC-cysC</i>	+		+	
<i>rhaBA</i>	+		+	
<i>pstSCAB-phoU</i>	+		+	
<i>cydAB</i>	-		0	
<i>nuoABCDEFGHIJKLMN</i>	-		#	
<i>aroP</i>	+		+	
<i>pdhR-aceEF-lpdA¹⁾</i>	-		-	
<i>lpdA</i>	-		-	
<i>cyoABCDE</i>	-		-	
<i>potFGHI</i>	-		-	
<i>dcuB-fumB</i>	+		+	
<i>ndh</i>	-		-	
<i>atpIBEFHAGDC</i>	-		-	
<i>icdA</i>	-		-	
<i>mdh</i>	-		-	
<i>gltA</i>	-		-	
<i>yigI-rarD-yigG-yigF</i>	+		+	
<i>gapA-yeaD</i>	-		-	
<i>acrEF</i>	+		+	
<i>yadF</i>	-		-	
<i>hycBCDEFGHI</i>	+		+	
<i>lysU</i>	+		+	
<i>gatYZABCDR_2</i>	-		-	
<i>yhiW</i>	+		+	
<i>b1012</i>	+		+	
<i>ydbA_2</i>	+		+	
<i>ftsYEX</i>	+		+	
<i>yjcRQP</i>	+		+	
<i>yjcP</i>	+		+	
<i>ybdL</i>	+		+	
<i>umuDC</i>	+		+	
<i>ybjX</i>	-		-	
<i>kbl-tdh</i>	+		+	
<i>nikABCDE</i>	+		Не исследовался	
<i>yidKJ</i>	+		Не исследовался	
<i>yfiA</i>	-		-	
<i>yceD-rpmF</i>	-		-	
<i>trpLEDCBA</i>	-		-	
<i>gpmA</i>	-		-	
<i>ptsI-crr</i>	-		-	
<i>rpsL-rpsG-fusA-tufA</i>	-		-	
<i>rplM-rpsI</i>	-		-	
<i>rpsP-yfiA-trmD-rplS</i>	-		-	
<i>yajG-ampG</i>	-		-	
<i>oppABCDEF</i>	-		-	
<i>rpsA</i>	-		-	
<i>b2866-ygeT</i>	+		+	
<i>pyrD</i>	+		+	
<i>mobABC</i>	+		+	
<i>yhjW</i>	+		+	
<i>recC</i>	+		+	
<i>narU</i>	+		+	

<i>ygiF-glnE-b3052</i>	+		+	
<i>map-glnD-dapD-b0165-yaeI</i>	+		+	
<i>yjcS</i>	+		+	
<i>b0540-tra5_2</i>	+		+	
<i>tynA</i>	+		+	
<i>mrcA</i>	+		+	
<i>yheEFG-hofFGH-yheHIJK-pshM-hofD-bfr</i>	+		+	
<i>yadQR</i>	+		+	
<i>speC</i>	+		+	
<i>yhjJ</i>	+		+	
<i>yabM</i>	+		+	
<i>yhhTU</i>	+		+	
<i>dinG</i>	+		+	
<i>yhiIH-yhhJ</i>	+		+	
<i>sbcDC</i>	+		+	
<i>ydcF</i>	-		-	
<i>ycfT</i>	+		+	
<i>yaaU</i>	+		+	
<i>yaaJA</i>	+		+	
<i>yi4I</i>	+		+	
<i>yi21_5-yi22_5-yqiGHI</i>	+		+	
<i>ompA</i>	-		-	
<i>argT-hisJQMP</i>	-		-	
<i>frdABCD</i>	+		+	
<i>yafZ</i>	+		+	
<i>aroM</i>	+		+	
<i>yghQ</i>	+		+	
<i>nadA-pnuC</i>	+		+	
<i>yhjCD</i>	+		+	
<i>degQS</i>	+		+	
<i>yhdRP</i>	+		+	
<i>yihR</i>	+		+	
<i>cdh</i>	+		+	
<i>menFD-yfbB-menBCE-pmrD</i>	+		+	
<i>yhcRQPo</i>	+		+	
<i>yjhIH</i>	+		+	
<i>manXYZ</i>	-		-	
<i>ptsG</i>	-		-	
<i>yggTUV</i>	-		-	
<i>rpmBG</i>	-		-	
<i>b2146-yeiA</i>	+		+	
<i>yjdHG</i>	+		+	
<i>yicONM</i>	+		+	
<i>yhfZYXWVUTS</i>	+		+	
<i>sbp</i>	+		+	
<i>nrdAB</i>	-		-	
<i>gapC_1-gapC_2</i>	-		-	
<i>artPIQM</i>	-		-	
<i>ilvY</i>	+		+	
<i>viaY</i>	+		+	
<i>fhuACDB</i>	+		+	
<i>yigN</i>	+		+	
<i>fliFGHIJK</i>	+		+	
<i>wecDE-wzxE-yifM_1</i>	0		+	

<i>xylFGHR</i>	0		+	
<i>ptr-recBD</i>	0		+	
<i>tyrB</i>	0		+	
<i>yhjX</i>	0		+	
<i>cvpA-purF</i>	0		+	
<i>yraPONM</i>	0		+	
<i>yjfHIJKLMC</i>	0		+	
<i>yphGFEDCB</i>	0		+	
<i>yhgF</i>	0		+	
<i>dnaX-ybaB-recB</i>	0		+	
<i>hisS-b2513-b2512-10-xseA</i>	0		+	
<i>rbsDACBKR</i>	0		-	
<i>yjiON</i>	0		+	
<i>idnDOTR-yjgR</i>	0		+	
<i>yjhATS</i>	0		+	
<i>ycdC</i>	0		-	
<i>b1172-b1171</i>	Не исследовался		+	
<i>yhjE</i>	Не исследовался		+	
<i>aceBAK</i>	Не исследовался		+	
<i>yjcR</i>	Не исследовался		-	

Табл.4. Предполагаемый FNR-регулон. Условные обозначения для геномов: см. Материалы и методы. Указаны веса сайтов, в скобках приводится положение сайта относительно рамки считывания. “-” – Сайт перед опероном отсутствует, “0” – ортологов данных генов не обнаружено. Жирным шрифтом отмечены члены “ядра”.

Оперон	EC	STM	STY	YE	YP	SM	PHL	PHA	PCA	PCH
Гены, вошедшие в регулон										
<i>tgt-yajC-secD-secF</i>	4.12 (-31)	3.83 (-322)	3.73 (-244)	-	4.27 (41)	3.56 (-319)	4.14 (-49)	3.56 (-392)	-	-
	3.83 (-187)									
<i>purMN</i>	4.46 (-244)	4.28 (-325)	4.28 (-322)	4.09 (-263)	4.09 (-311)	-	5.03 (-138)	5.08 (-125)	4.01 (-310)	4.09 (-303)
<i>narXL</i>	5.15 (-278)	5.05 (-282)	5.05 (-281)	-	3.53 (-279)	5.00 (-291)	4.03 (46)	4.03 (41)	4.70 (-203)	5.00 (-352)
	4.20 (-240)									4.41 (-89)
<i>thrS-infC-rpmI-rplT</i>	3.67 (-69)	-	-	3.59 (-258)	3.57 (-319)	-	3.73 (-278)	3.79 (-64)	3.49 (21)	-
				3.98 (-73)	3.81 (-73)		3.79 (-65)			
					3.57 (-77)					
<i>rpsT</i>	3.85 (1)	3.85 (1)	3.85 (1)	0	3.85 (-1)	0	4.62 (-265)	0	0	0
<i>potFGHI</i>	3.51 (-398)	3.79 (-99)	3.79 (-97)	-	-	-	0	0	-	-
<i>yjiHG-iadA-yjiE</i>	3.60 (8)	4.41 (-69)	4.41 (-281)	-	0	0	0	0	3.49 (-4)	0
<i>nanA-nanT</i>	4.36 (47)	3.70 (47)	3.70 (47)	-	3.87 (-98)	0	0	0	0	0
<i>upp-uraA</i>	4.46 (-67)	4.28 (-102)	4.28 (-104)	4.09 (-99)	4.09 (-66)	-	5.03 (-102)	5.08 (-30)	4.01 (-167)	4.09 (-96)
<i>ibpB-yidE</i>	3.98 (-173)	-	3.40 (-179)	3.89 (-286)	4.25 (-306)	0	0	0	3.62 (-144)	3.96 (-103)
				4.25 (-191)						
<i>glgCAP</i>	3.93 (-169)	3.58 (-169)	3.58 (-159)	-	3.76 (-169)	3.53 (-234)	0	0	-	-
<i>ybhK</i>	4.42 (-146)	4.32 (-145)	4.32 (-147)	3.67 (-183)	3.99 (-191)	-	3.59 (-266)	3.67 (-100)	3.54 (-118)	-
				3.99 (-115)						
<i>nirBDC-cysG</i>	5.15 (-73)	5.15 (-73)	5.15 (-71)	5.15 (-72)	5.15 (-70)	4.70 (-70)	0	0	3.73 (-94)	3.48 (-175)
				3.70 (-321)						
				3.67 (-1)						
				3.66 (-343)						
<i>rhaBA</i>	3.64 (-397)	3.64 (-397)	3.64 (-399)	0	3.58 (-128)	0	0	0	-	-
		3.52 (-323)	3.52 (-325)							
<i>pstSCAB-phoU</i>	4.22 (-270)	3.56 (-111)	3.56 (-113)	3.45 (-128)	3.72 (-278)	-	4.04 (-34)	3.84 (-105)	3.53 (-50)	

<i>metBL</i>	3.79 (-73)	-	-	4.03 (-213)	3.61 (-16)	-	4.50 (-65)	0	3.69 (50)	3.92 (-171)
<i>aroP</i>	3.80 (-91)	-	-	-	-	-	3.62 (-78)	-	3.67 (0)	3.61 (86)
<i>pdhR-aceEF-lpdA</i>	4.43 (-116)	4.43 (-117)	4.43 (-117)	3.84 (-397)	4.04 (-64)	4.35 (-118)	3.50 (-380)	3.82 (-354)	3.87 (-114)	3.40 (-118)
				4.11 (-148)						
<i>gltA</i>	3.67 (-172)	3.69 (-212)	3.69 (-214)	3.45 (-366)	0	-	3.75 (-109)	-	-	-
<i>nuoABCDEFGHIJKLMN</i>	3.59 (-432)	4.34 (-198)	4.34 (-198)	3.50 (-382)	-	0	3.69 (-392)	-	0	0
<i>cyoABCDE</i>	3.37 (-155)	-	-	-	3.50 (-222)	3.67 (86)	3.80 (-337)	-	-	3.77 (83)
<i>dcuB-fumB</i>	3.73 (-282)	-	-	0	0	4.11 (-79)	0	0	3.64 (-277)	4.31 (-107)
									4.31 (-241)	
									3.54 (-98)	
<i>ndh</i>	5.03 (-151)	3.58 (-200)	3.58 (-198)	-	4.90 (-149)	4.64 (-109)	4.18 (-149)	4.52 (-151)	4.90 (-336)	5.05 (-340)
		4.65 (-151)	4.65 (-149)							
		3.94 (-71)	3.94 (-69)						3.52 (62)	
<i>atpIBEFHAGDC</i>	3.92 (-110)	3.92 (-121)	3.92 (-123)	3.93 (-224)	3.51 (-307)	3.65 (80)	3.64 (-224)	3.53 (-225)	3.70 (83)	3.78 (74)
<i>icdA</i>	3.60 (-147)	-	-	-	-	3.51 (-239)	4.50 (-248)	3.92 (-250)	3.60 (-1)	3.56 (-317)
						4.41 (-4)	3.60 (-2)	3.60 (-4)		
<i>mdh</i>	3.93 (-258)	3.93 (-236)	3.93 (-236)	3.79 (-255)	-	-	3.87 (-85)	3.45 (-46)	-	-
<i>cydAB</i>	4.78 (-346)	5.08 (-348)	5.08 (-350)	3.90 (-376)	4.84 (-313)	4.08 (-356)	-	-	4.84 (-399)	3.87 (-397)
	4.57 (-294)	4.14 (-296)	4.14 (-298)	4.84 (-312)	3.55 (-347)				4.84 (-374)	
	3.98 (-249)	3.81 (-251)	3.81 (-253)	4.31 (-261)	4.48 (-262)					
Гены, не вошедшие в регулон										
<i>nikABCDE</i>	3.73 (-87)	0	0	0	0	0	0	0	0	-
<i>kbl-tdh</i>	4.37 (-122)	-	-	3.95 (-315)	0	0	0	0	0	0
<i>ybjX</i>	3.79 (-257)	-	-	3.53 (-333)	0	0	0	0	0	0
<i>ybdL</i>	3.54 (-343)	-	-	-	0	-	0	0	-	3.54 (-333)
<i>ftsYEX</i>	3.65 (-399)	-	-	-	-	-	-	0	-	-
<i>yjcRQP</i>	4.02 (-387)	0	0	0	0	3.52 (56)	0	0	0	0

<i>aceEF-lpdA</i>	4.29 (-132)	-	-	-	-	0	-	-	-	-
Гены, определение принадлежности которых к регулону невозможно										
<i>yidKJ</i>	3.86 (-239)	0	0	0	0	0	0	0	0	0
<i>ylcBCD-ybdE</i>	3.55 (-184)	0	0	0	0	0	0	0	0	0
<i>yjcP</i>	4.42 (-146)	0	0	0	0	0	0	0	0	0
<i>umuDC</i>	4.20 (-247)									
	4.02 (-1)	-	-	0	0	0	0	0	0	0

Таб.5. Предполагаемый ArgA-регулон. Условные обозначения для геномов: см. Материалы и методы. Указаны веса сайтов, в скобках приводится положение сайта относительно рамки считывания. “-” – Сайт перед опероном отсутствует, “0” – ортологов данных генов не обнаружено. Жирным отмечены члены “ядра”

Оперон	EC	STM	STY	YE	YP	SM	PHL	PHA	PCA	PCH
Гены, вошедшие в регулон										
<i>pdhR-aceEF-lpdA</i>	4.42 (-230)	4.85 (-85)	4.85 (-85)	4.85 (-116)	4.85 (-32)	4.85 (-86)	4.85 (-348)	4.85 (-322)	4.85 (-82)	4.85 (-86)
	5.06 (-232)									
<i>lpdA</i>	4.57 (-223)	4.39 (-108)	4.39 (-105)	4.68 (-248)	4.68 (-138)	-	-	-	4.39 (-139)	4.39 (-120)
	4.85 (-84)									
<i>cyoABCDE</i>	4.92 (-260)	4.97 (-261)	4.97 (-263)	4.46 (-302)	4.71 (-301)	4.53 (-319)	4.77 (-185)	4.83 (-186)	-	-
	4.95 (-249)	4.69 (-250)	4.69 (-252)			4.82 (-286)				
<i>dcuB-fumB</i>	4.28 (-395)	-	-	4.34 (-285)	-	4.11 (-79)	0	0	3.64 (-277)	4.31 (-107)
	4.53 (-274)								4.31 (-241)	
	4.32 (-256)								3.54 (-98)	
	4.95 (-390)									
<i>ndh</i>	5.15 (-45)	5.21 (-45)	5.21 (-43)	-	4.94 (-31)	4.71 (-388)	4.75 (-31)	5.13 (-33)	4.27 (-218)	4.27 (-221)
	4.65 (-34)	4.65 (-34)	4.65 (-32)							
		4.58 (-66)	4.58 (-64)							
<i>atpIBEFHAGDC</i>	4.59 (-339)	-	-	-	-	4.28 (-88)	4.60 (-290)	-	-	-
						4.28 (-163)				
<i>icdA</i>	4.91 (-111)	4.61 (-111)	4.61 (-109)	4.35 (-40)	4.73 (11)	4.73 (-41)	4.91(-46)	4.91 (-48)	4.53 (-38)	4.53 (-38)
<i>mdh</i>	5.19 (-263)	5.19 (-241)	5.19 (-241)	-	-	-	4.71 (-285)	4.41 (-375)	-	5.05 (-318)
<i>cydAB</i>	5.94 (-586)	4.91 (-408)	4.91 (-406)	5.32 (-382)	4.60 (-343)	5.30 (-477)	4.92 (-537)	4.85 (-259)	4.47(-617)	4.70 (-442)
<i>nuoABCDEFGHIJKLMN</i>	4.46 (-163)	4.46 (-163)	4.46 (-163)	4.46 (-171)	4.46 (-172)	0	5.70 (-540)	5.50 (-541)	0	0
<i>gltA</i>	4.34 (-393)	4.24 (-129)	4.24 (-131)	4.40 (-365)	4.83 (-250)	-	-	-	-	-
	5.15 (-654)			4.34 (-191)						
<i>yigI-rarD-yigG-yigF</i>	4.91 (-16)	4.91 (-16)	4.91 (-36)	-	5.15 (-85)	4.30 (-66)	4.91 (-72)	5.52 (-140)	-	5.52 (-73)
<i>gapA-yeaD</i>	4.99 (-296)	4.99 (-294)	4.32 (-294)	5.29 (-297)	5.29 (-288)	-	-	-	5.29 (-314)	5.35 (-299)
<i>acrEF</i>	4.63 (-288)	-	-	-	-	4.35 (-79)	-	-	4.54 (-90)	-

<i>potFGHI</i>	4.36 (-216)	4.34 (-284)	4.34 (-282)	-	-	-	0	0	-	-
<i>yadF</i>	4.39 (-18)	4.39 (-18)	4.39 (-18)	-	-	-	-	4.44 (-74)	-	-
	4.27 (-17)									
Гены, не вошедшие в регулон										
<i>aroP</i>	4.57 (-332)	-	-	-	-	-	-	-	-	-
<i>gatYZABCDR</i>	4.63 (-534)	-	-	0	-	-	0	0	0	0
	4.38 (-529)									
<i>lysU</i>	4.47 (-215)	-	-	-	-	0	0	-	-	-
<i>ycdM</i>	4.56 (-86)	0	0	4.40 (-125)	0	0	0	0	0	0
	4.67 (-88)									
<i>hycBCDEFGHI</i>	4.85 (-655)	-	-	0	-	-	0	0	0	0
Гены, определение принадлежности которых к регулону невозможно										
<i>ydbA</i>	4.38 (22)	-	0	0	0	0	0	0	0	0
<i>yhiW</i>	4.32 (-145)	0	0	0	0	0	0	0	0	0

Таб.6. Функции генов, входящих в FNR- или ArcA-регулоны. “+” – оперон относится к регулону, “-” – оперон не относится к регулону, “0” – определение принадлежности оперона к регулону не возможно. Жирным шрифтом выделены члены “ядра”.

Оперон	Функция	Ссылка	Fnr-регулон	ArcA-регулон
<i>pdhR-aceEF-lpdA</i>	Пируват-дегидрогеназный комплекс	Haydon <i>et al</i> , 1993; Stephens <i>et al</i> , 1983	+	+
<i>cyoABCDE</i>	Аэробная цитохром <i>bo3</i> -содержащая оксидаза	Chepuri <i>et al</i> , 1990	+	+
<i>dcuB-fumB</i>	Транспортер четырехуглеродных дикарбоновых кислот	Bell <i>et al</i> , 1989; Six <i>et al</i> , 1993	+	+
	Фумарат-гидратаза		+	+
<i>ndh</i>	НАДН-дегидрогеназа II	Calhoun and Gennis, 1993	+	+
<i>atp1BEFHAGDC</i>	АТФ-синтетаза	Kanazawa <i>et al</i> , 1982	+	+
<i>icdA</i>	НАДФН-зависимая изоцитратат-дегидрогеназа	Oshima <i>et al</i> , 1996	+	+
<i>mdh</i>	Малат-дегидрогеназа	McAlister-Henn <i>et al</i> , 1987	+	+
<i>gltA</i>	Цитрат синтаза	Ner <i>et al</i> , 1983	+	+
<i>cydAB</i>	Аэробная цитохром <i>bd</i> -содержащая оксидаза	Chepuri <i>et al</i> , 1990	+	+
<i>nuoABCDEFGHIJKLMN</i>	НАДН-дегидрогеназы I	Weidner <i>et al</i> , 1995	+	+
<i>yigI-rarD-yigG-yigF</i>	Фунция неизвестна	Blattner <i>et al</i> , 1997	-	+
<i>gapA</i>	Глицеральдегид-3-фосфат-дегидрогеназа	Branlant and Branlant, 1985	-	+
<i>acrEF</i>	Гены устойчивости к акрифлавину	Haydon <i>et al</i> , 1993; Stephens <i>et al</i> , 1983	-	+
<i>yadF</i>	Ангидраза	Merlin <i>et al</i> , 2003	-	+
<i>aroP</i>	Транспортер аминокислот	Honore and Cole, 1990	+	-
<i>gatYZABCDR</i>	Утилизация галактозы	Nobelmann and Lengeler, 1995	-	-
<i>lysU</i>	Лизил-тРНК синтетаза	Leveque <i>et al</i> , 1990	-	-
<i>ycdM</i>	Предполагаемая монооксигеназа	Blattner <i>et al</i> , 1997	-	-
<i>potFGHI</i>	Система транспорта путресцина	Pistocchi <i>et al</i> , 1993	+	-
<i>hycBCDEFGHI</i>	Формиат-гидрогенилаза	Boehm <i>et al</i> , 1997	-	-
<i>ydbA</i>	Функция неизвестна	Blattner <i>et al</i> , 1997	-	0
<i>yhiW</i>	Предполагаемый регулятор транскрипции	Jiang, 2005	-	0
<i>tgt-yajC-secD-secF</i>	тРНК транс-гликозилаза	Reuter <i>et al</i> , 1991; Gardel <i>et al</i> , 1990	+	-
	Функция неизвестна		+	-
	Система секреции белков		+	-
<i>purMN</i>	PurM - фосфорибозил аминоксидаз-синтаза	Smith and Daum, 1986	+	-
	PurN - фосфорибозилглицинамид формилтрансфераза			
<i>narXL</i>	Двухкомпонентная система регуляции нитрат-нитритного дыхания	Nohn <i>et al</i> , 1989	+	-
<i>thrS-infC-rpmI-rplT</i>	ThrS - треонил-тРНК-синтетаза	Mayaux <i>et al</i> , 1983; Sacerdot <i>et al</i> , 1982; Wada and Sako, 1987; Fayat <i>et al</i> , 1983	+	-
	InfC - фактор инициации трансляции IF-3			
	RpmI - рибосомный белок L35			
	RplT - рибосомный белок L20			
<i>rpsT</i>	Рибосомный белок S20	Mackie, 1981	+	-

<i>iadA</i>	Изоаспартил дипептидаза	Burland <i>et al</i> , 1995	+	-
<i>nanA-nanT</i>	NanA - N-ацетилнейрамин лиаза	Ohta <i>et al</i> , 1985; Martinez <i>et al</i> , 1995	+	-
	NanT - предполагаемый транспортер сиаловой кислоты		+	-
<i>upp-uraA</i>	Upp - урацил фосфорибозил трансфераза	Anderson <i>et al</i> , 1992	+	-
	UraA - транспортер урацила		+	-
<i>ibpB-yidE</i>	IbpB - белок теплового шока	Allen <i>et al</i> , 1992; Burland <i>et al</i> , 1993	+	-
	YidE - предполагаемый транспортный белок		+	-
<i>glgCAP</i>	Гены биосинтеза гликогена	Baecker <i>et al</i> , 1983; Kumar <i>et al</i> , 1986; Choi <i>et al</i> , 1989	+	-
<i>ybhK</i>	Функция неизвестна	Oshima <i>et al</i> , 1996	+	-
<i>nirBDC-cysG</i>	NirBC - нитрит редуктаза	Peakman <i>et al</i> , 1990	+	-
	NirD - транспортер нитрита		+	-
	CysG - сирогем синтаза		+	-
<i>rhaBA</i>	Гены метаболизма рамнозы	Moralejo <i>et al</i> , 1993	+	-
<i>pstSCAB-phoU</i>	Система транспорта фосфатов	Magota <i>et al</i> , 1984; Amemura <i>et al</i> , 1985	+	-
<i>metBL</i>	Гены биосинтеза метионина	Duchange <i>et al</i> , 1983	+	-
<i>nikABCDE</i>	Гены транспорта никеля	Navarro <i>et al</i> , 1993	-	-
<i>kbl-tdh</i>	Kbl - 2-амино-3-кетобутират КоА лигаза	Aronson <i>et al</i> , 1988	-	-
	Tdh - L-треонин 3-дегидрогеназа		-	-
<i>ybjX</i>	Функция неизвестна	Oshima <i>et al</i> , 1996	-	-
<i>ybdL</i>	Функция неизвестна	Oshima <i>et al</i> , 1997	-	-
<i>ftsYEX</i>	Гены регуляции деления клетки	Gill <i>et al</i> , 1986	-	-
<i>yjcRQP</i>	Функция неизвестна	Oshima <i>et al</i> , 1996	-	-
<i>yidKJ</i>	Функция неизвестна	Burland <i>et al</i> , 1993	0	-
<i>yjcP</i>	Функция неизвестна	Oshima <i>et al</i> , 1996	0	-
<i>umuDC</i>	Мутагенная ДНК-полимераза	Kitagawa <i>et al</i> , 1985	0	-
<i>cusCFBA</i>	Система экспорта меди	Murakami <i>et al</i> , 2002	-	-

Рис.1. Многообразие дыхательных систем *E.coli*. Для доноров и акцепторов электронов указаны окислительно-восстановительные потенциалы: слева – для реакций окисления доноров электронов, справа – для реакций восстановления акцепторов электронов. Стрелками показано направление увеличения окислительно-восстановительного потенциала (Gennis and Stewart, 1996).

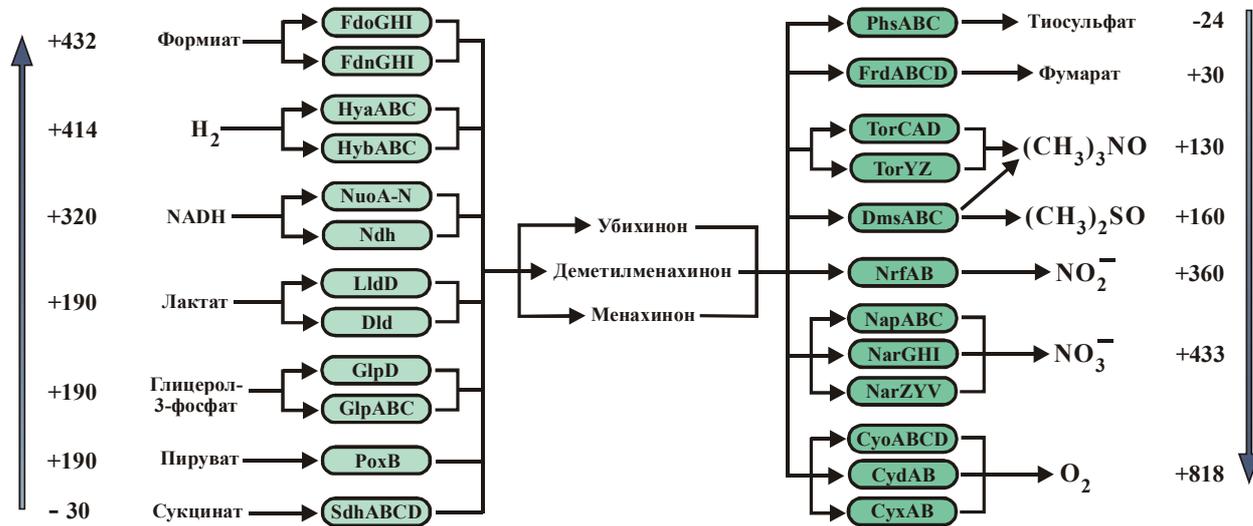


Рис.2. Различные механизмы создания протонного градиента. а) НАДН-дегидрогеназа I; б) НАДН-дегидрогеназа II; в) цитохром bo_3 -содержащая редуктаза; д) цитохром bd -содержащая редуктаза.

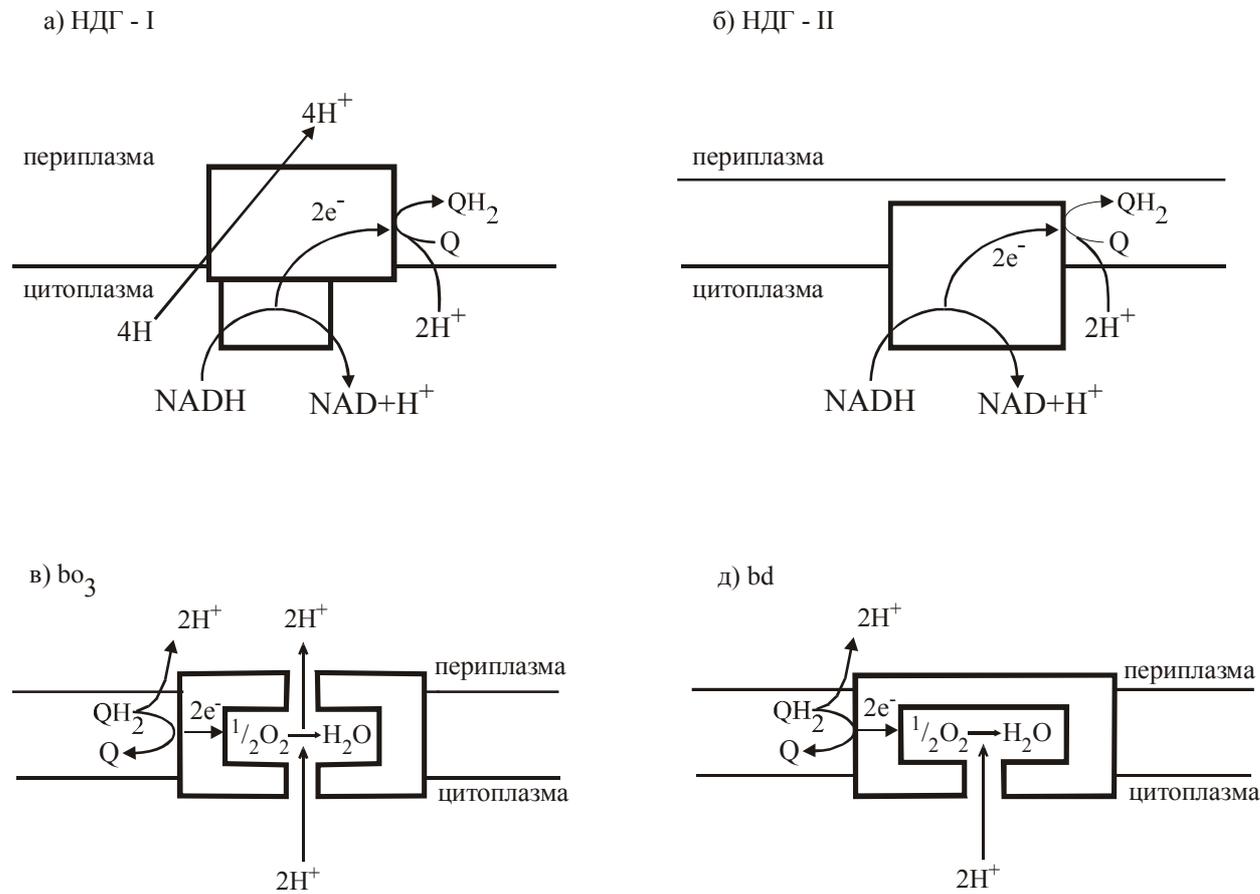


Рис.3. Четыре различных варианта дыхательных цепей, ферменты которых используют НАДН в качестве донора электронов и молекулярный кислород в качестве акцептора электронов (по Calhoun et al, 1993).

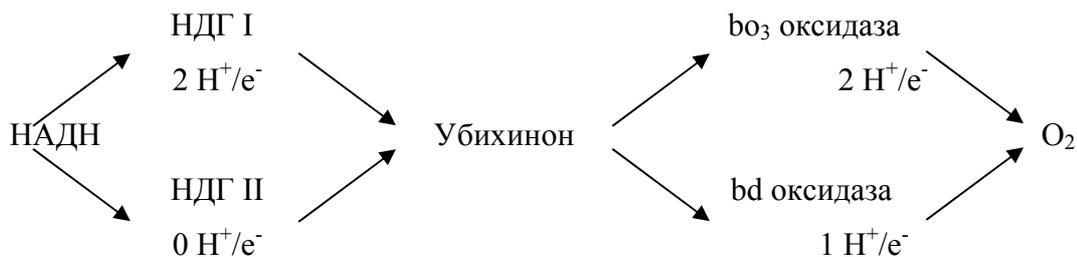


Рис.4. Механизм активации и инактивации FNR в зависимости наличия или отсутствия молекулярного кислорода в среде (см. пояснения в тексте).

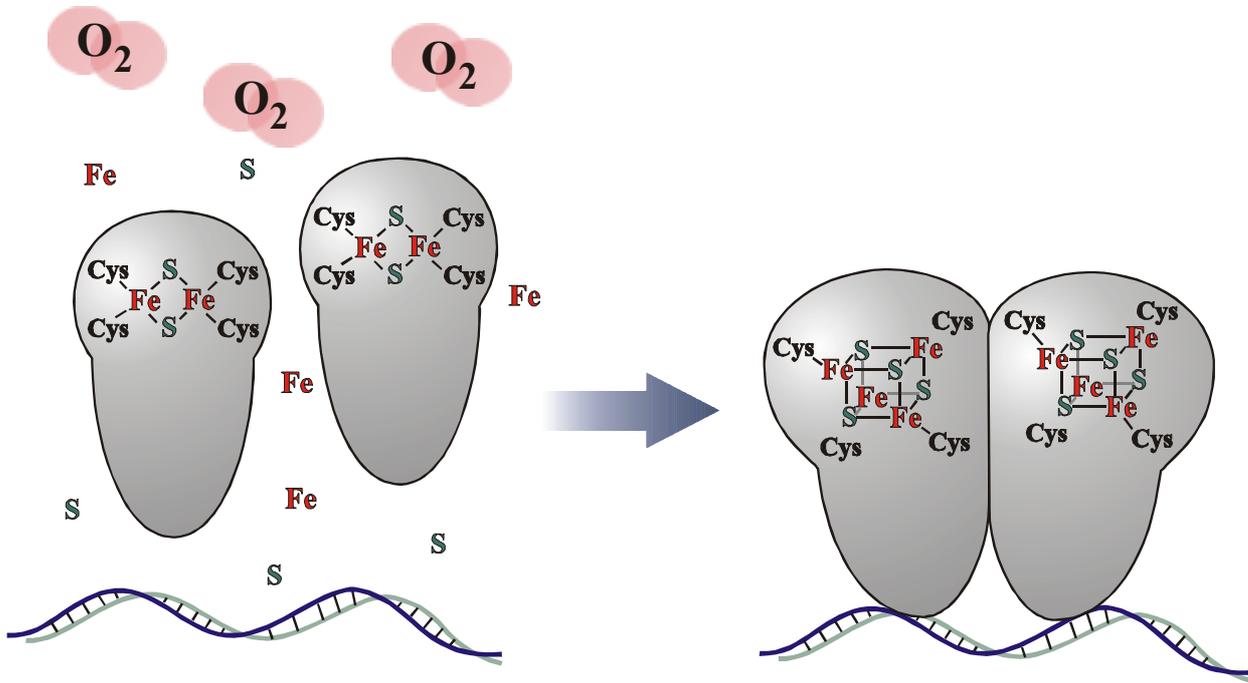


Рис.5. Механизм активации и инактивации ArcA (см. пояснения в тексте).

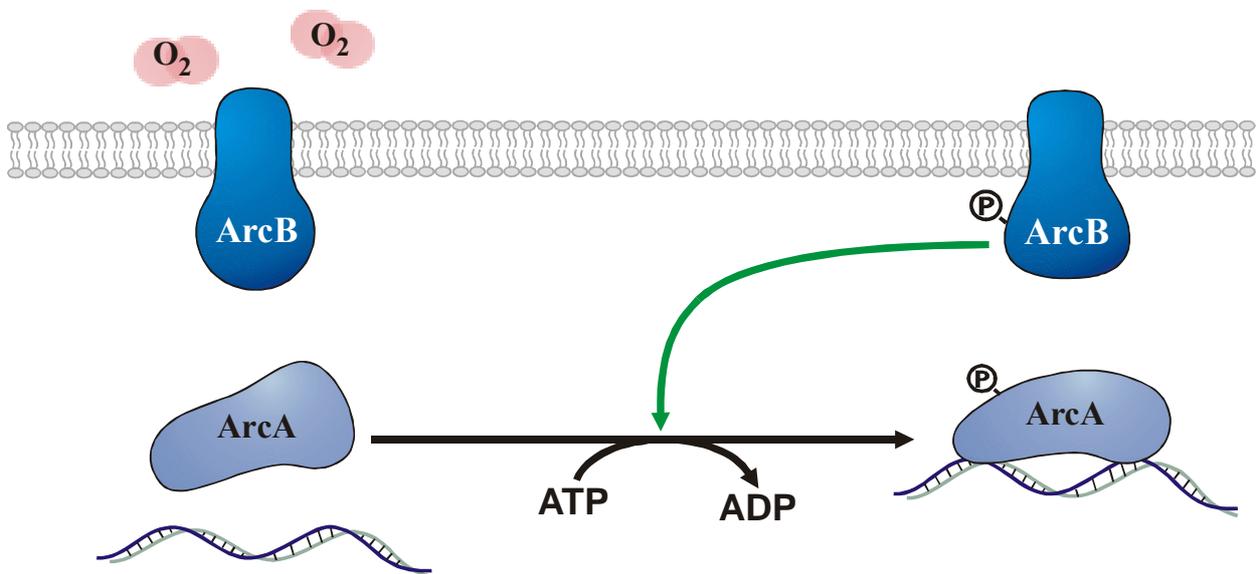


Рис.5. Диаграмма Logo для сайтов связывания Fnr и ArcA. По горизонтальной оси указан номер позиции нуклеотида, по вертикали – информационное содержание позиции в битах. Высота столбца пропорциональна информационному содержанию данной позиции, относительная высота каждой буквы соответствует частоте нуклеотида в данной позиции.

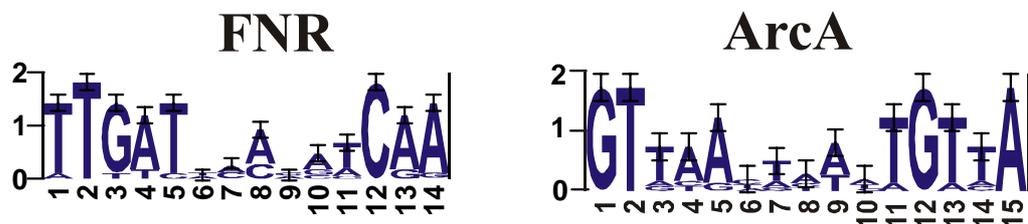


Рис.7. Функции генов ядра.

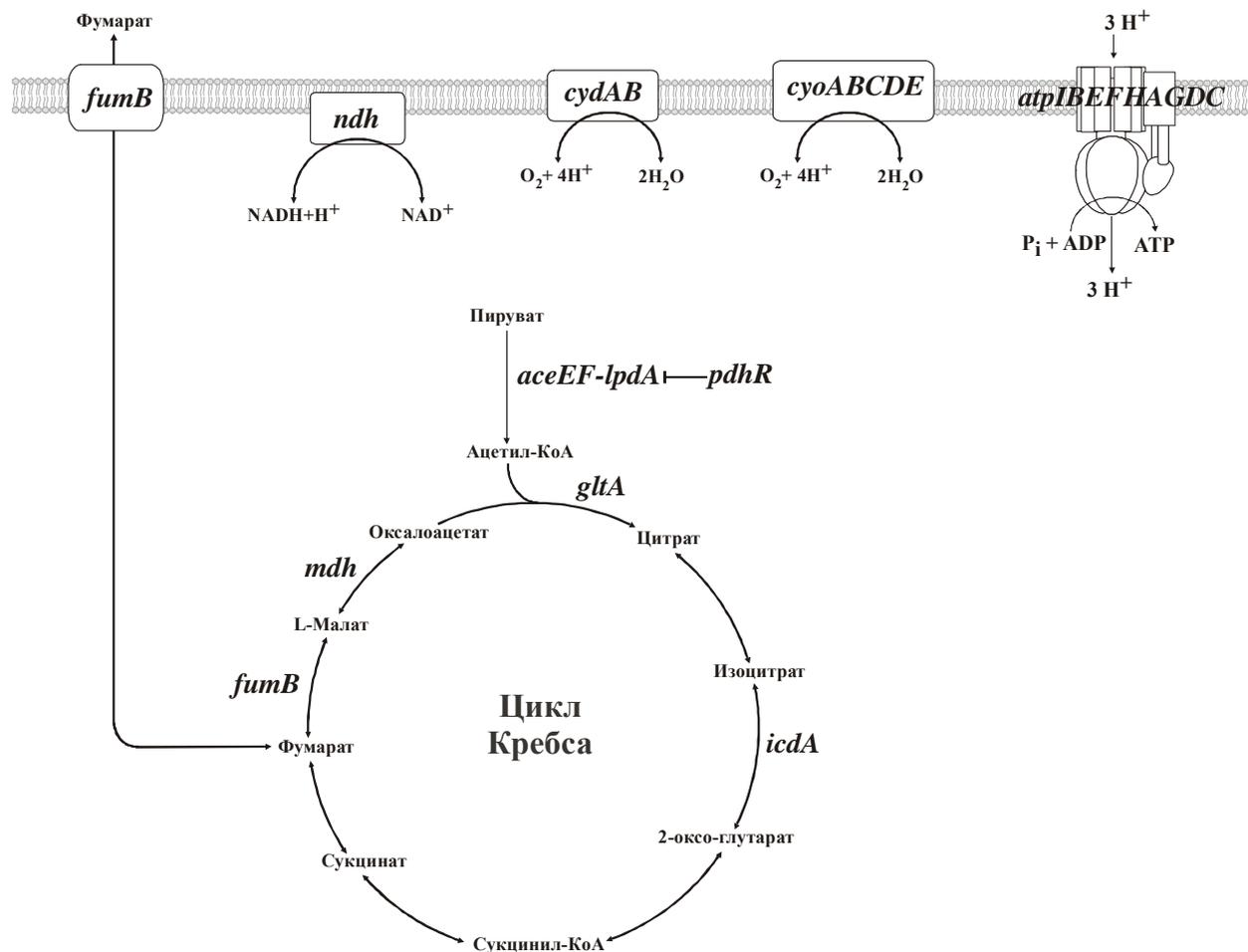


Рис.8. Схема, визуализирующая соотношение Fnr-регулона, Fnr-модулона, ArcA-регулона и ArcA-модулона. Ниже указано количество генов, в скобках оперонов, входящих в состав Fnr-модулона, Fnr-регулона, ArcA-регулона и ArcA-модулона соответственно.

Модулон Fnr 151 (413)		
	Регулон Fnr 26 (83)	
	Регулон ArcA 16 (54)	
	Модулон ArcA 135 (371)	

