

УТВЕРЖДАЮ

Зам. директора Института общей  
генетики им. Н.И. Вавилова РАН,

д.б.н., проф. С.К. Абилев



5 ноября 2014

### ОТЗЫВ

ведущей организации

на диссертационную работу Цой Ольги Владиславовны  
«Эволюция систем регуляции транскрипции в геномах бактерий»,  
представленную на соискание учёной степени кандидата биологических наук  
по специальности 03.01.09 «Математическая биология, биоинформатика»

**Актуальность** темы диссертации О.В. Цой связана с исследованием эволюции систем, регулирующих транскрипцию в геномах бактерий. В связи с развитием технологий секвенирования нуклеотидных последовательностей число известных бактериальных геномных последовательностей в международных базах данных возрастает с экспоненциальной скоростью. Хотя большая часть бактериальных геномов представлена в виде контигов, но даже такой информации достаточно для проведения сравнительного эволюционного анализа не только отдельных семейств генов, но и целых биологических систем.

Основное внимание в работе уделяется анализу сети регуляции транскрипции. Как известно большинство биологических систем в живых организмах или сообществах образуют так называемые генные сети. Они представляют собой совокупность координированно взаимодействующих генов, взаимодействующих между собой. Основным предназначением таких взаимодействий является адаптация к изменяющимся условиям внешней

среды. Одним из наиболее важных путей адаптации является процесс регуляции транскрипции генов. Биологическая система, состоящая из транскрипционных факторов и регулируемых ими генов и образует транскрипционную сеть. Исследование такого рода сетей позволяет понять какими механизмами пользуются живые организмы для адекватного реагирования на изменяющиеся факторы окружающей среды.

Автором отмечается, что несмотря на обилие данных по строению геномных последовательностей большого числа бактериальных видов. Анализ систем регуляции транскрипции проводился только на нескольких модельных организмах или изучаются только отдельные функциональные системы. Конечно, экспериментальное исследование каждого отдельного штамма задача трудоемкая и требующая огромных финансовых затрат, но применение компьютерных методов сравнительной геномики позволяет на основании известных у ряда организмов транскрипционных сетей реконструировать их строение у близкородственных организмов. Таким образом возникает на данный момент возможность крупномасштабного эволюционного анализа строения транскрипционных сетей у бактерий.

Работа диссертанта посвящена реконструкции и анализу сетей транскрипционных регуляторных элементов в близкородственных геномах с помощью методов компьютерной биологии для немодельных организмов. Автором одним из первых в мире была проведена работа по предсказанию «de novo» сайтов связывания белков регуляторов пути утилизации этаноламина.

Немаловажной частью работы является изучение эволюции регуляторных блоков. Генную сеть можно разбить на подструктуры называемые блоками. В частности для транскрипционной сети наиболее интересным является блок типа «треугольник», представляющий собой взаимодействующую систему, состоящую из двух транскрипционных факторов и регулируемого ими гена. Автором проводится изучение эволюции блоков типа «треугольник» в транскрипционных сетях бактерий.

**Основные научные результаты**, полученные О.В. Цой заключаются в следующем:

1. Путем анализа таксономического распределения генов утилизации этаноламина показано, что существует два возможных пути катаболизма использования этаноламина, связанных с типом оперона.
2. Проведено исследование эволюции генов утилизации этаноламина. В результате было установлено, что короткий оперон является предковым, из которого путем добавления новых генов образовался длинный тип, так же обнаруживаются признаки, как минимум трех событий горизонтального переноса.
3. Был «de novo» предсказан мотив участка связывания транскрипционного фактора EutR в Enterobacteriales и Burkholderiales. В Enterobacteriales участки связывания EutR обнаружены в промоторной области семи оперонов, среди которых есть непосредственно гены утилизации этаноламина, а также гены синтеза кофактора основного фермента пути этаноламинлиазы – кобаламина. В Burkholderiales участки связывания EutR обнаружены только непосредственно в промоторной области генов утилизации этаноламина.
4. Показано, что в штаммах *Escherichia coli* взаимодействия типа Л→ген (локальный транскрипционный фактор→ген) чаще сохраняются в регуляторных блоках, чем вне их, но на уровнях порядка Enterobacteriales эти взаимодействия консервативнее в парных взаимодействиях, по сравнению с регуляторными блоками, что может быть результатом изменчивости транскрипционной сети.
5. Установлено, что разные типы треугольников на уровне порядка Enterobacteriales эволюционируют по-разному. Показано, что регуляция как глобальными, так и локальными транскрипционными факторами в несогласованном треугольнике оказывается консервативнее, чем в согласованном.

**Новизна** этих результатов состоит в том, что впервые был проведен анализ эволюции генов утилизации этаноламина. Установлены гены

входящие в предковый оперон. Представлена модель возникновения существующих на данный момент оперонов. Предсказан «de novo» мотив связывания участка связывания транскрипционного фактора EutR в Enterobacteriales и Burkholderiales. Проанализирована эволюция блока типа «треугольник» на уровне порядка Enterobacteriales и штаммов *Escherichia coli*. Установлено что что регуляция как глобальными, так и локальными транскрипционными факторами в несогласованном треугольнике оказывается консервативнее, чем в согласованном. Впервые при изучении эволюции блоков регуляции были учтены события, связанные с появлением мутаций в сайтах связывания транскрипционных факторов.

**Структура работы.** Диссертация состоит из введения, четырех глав, осуждения, выводов и списка литературы из 166 наименований Работа изложена на 100 страницах машинописного текста, включает 21 рисунок и 14 таблиц.

В первой главе приводится обзор литературы. В данном разделе достаточно подробно освещается вводная информация о генных сетях. Приводится обзор работ, посвященных системам утилизации этаноламина. Описываются основные способы идентификации мотивов связывания транскрипционных факторов. Описываются типы регуляторных блоков типа «треугольник». Данная глава дает необходимый минимум сведений о состоянии исследований в областях рассматриваемых в работе.

Во второй главе приводится информация об основных материалах и методах, использованных во время работы. Представлен список бактериальных геномов, которые использовались при выполнении исследований. Приведен список программ, которые использовались при проведении анализа.

В третьей главе приводятся результаты исследований, посвященных предсказанию регуляторных элементов транскрипционной сети на примере пути деградации этаноламина. В первом подразделе приводятся результаты поиска ортологов генов утилизации этаноламина. Приводится структура *eut-*

оперона. Во втором подразделе рассматриваются результаты изучения эволюции системы утилизации этаноламина. Приводится филогенетическое дерево для белка eutB, а так же результаты реконструкция состава eut-оперонов у предков порядка Enterobacteriales и типов Firmicutes и Proteobacteria. В третьем подразделе рассматриваются результаты предсказания «de novo» мотивов участка связывания транскрипционного фактора EutR в Enterobacteriales и Burkholderiales.

В четвертой главе изложены результаты изучения эволюции регуляторных блоков на примере блока типа «треугольник». В первом подразделе определяются какие факторы транскрипции считаются в дальнейшем глобальными, а какие локальными. Вводится понятие глобального и локального транскрипционных факторов. Во втором разделе описывается функциональная аннотация генов, входящих в состав регуляторных блоков типа «треугольник». В третьем и четвертом подразделах приводятся результаты изучения эволюции регуляторного блока типа «треугольник» в близкородственных организмах на примере штаммов Escherichia coli и в Enterobacteriales.

**Практическая ценность работы.** Обнаружено, что большинство бактерий, вызывающих пищевые отравления, имеют ферменты утилизации этаноламина. Таким образом изучение пути утилизации этаноламина может иметь ценность при проведении биомедицинских исследований направленных на разработку новых лекарственных средств.

Несмотря на общее положительное впечатление, работа содержит ряд недостатков:

1. Избранный автором способ представления подписей к рисункам не сильно удачен. Ряд подписей находятся на двух страницах, что несколько затрудняет понимание содержания рисунка.
2. Рисунки 19-22 не имеют подписей на осях. Нужно сказать, что в подписи к рисунку 19 приводится информация о том что отложено по каждой из осей и в подписях к остальным рисункам имеются ссылки на

рисунок 19. Но так как рисунки однотипные они сливаются друг с другом и понять чем они отличаются друг от друга очень сложно.

3. В таблице 2 не приводится в каком состоянии были геномы на момент их использования в работе. То есть были ли они опубликованы в виде полногеномной последовательности или же в виде контигов.
4. В главе «Материалы и методы» указываются три программы для построения выравниваний ClustalX, Muscle и MEGA. Из дальнейшего текста не понятно какая из них использовалась на какой стадии выполнения работы, а так же не указаны параметры с которыми они запускались.
5. На стр. 40 приводится ссылка на описание метода «neighbour-joining» в главе «Материалы и методы», но к сожалению там упоминание об этом методе отсутствует.

Отмеченные недостатки носят скорее редакционный характер и не снижают практическую ценность результатов, полученных в работе.

**Рекомендации по практическому использованию.** Результаты диссертации могут использоваться при проведении научных исследований в области молекулярной биологии и биоинформатики. Полученные в диссертации статистические закономерности и модели могут быть использованы в Федеральном государственном бюджетном учреждении науки Институт общей генетики им. Н.И. Вавилова РАН, Федеральном государственном бюджетном учреждении науки Институт молекулярной генетики РАН, Федеральном государственном бюджетном учреждении науки Институт математических проблем биологии РАН, Федеральном государственном бюджетном учреждении науки Институт биологии гена РАН, Федеральном государственном бюджетном учреждении науки Институт проблем передачи информации им. А.А. Харкевича РАН, Московском государственном университете имени М.В. Ломоносова, Новосибирском государственном университете и других учебных и научно-исследовательских организациях.

**Общая оценка работы.** Диссертация представляет собой завершённое

научное исследование, выполненное автором самостоятельно и на высоком уровне. Основные результаты работы опубликованы в рецензируемых научных изданиях и представлены на ведущих международных конференциях в области биоинформатики и системной биологии. Результаты достоверны и достаточно полно опубликованы, автореферат адекватно отражает содержание диссертации. Решение поставленной в диссертации задачи имеет большое значение для понимания эволюции систем регуляции транскрипции, а также задач функциональной и эволюционной геномики и генетической биоинформатики, что соответствует требованиям «Положения о порядке присуждения ученых степеней», утвержденного Постановлением Правительства РФ от 24.09.2013 г. № 842, предъявляемым к диссертациям на соискание ученой степени кандидата наук, а ее автор заслуживает присуждения заслуживает присуждения ученой степени кандидата биологических наук по специальности 03.01.09 - математическая биология, биоинформатика. Материалы диссертации были рассмотрены и обсуждены на Межлабораторном семинаре ИОГен РАН 17 октября 2014 года (протокол №2).

05 ноября 2014

Н.с. лаборатории системной биологии и вычислительной генетики ИОГен  
РАН



к.ф.-м.н. Касьянов А.С.

Подпись научного сотрудника лаборатории системной биологии и вычислительной генетики ИОГен РАН, кандидата физико-математических наук Касьянова Артема Сергеевича удостоверяю.

Ученый секретарь ИОГен РАН  
д.б.н. Огаркова О.А.

