

**РЕГУЛЯЦИЯ ТРАНСКРИПЦИИ В СИСТЕМЕ ГЕНОВ,
ОТВЕТСТВЕННЫХ ЗА ПРОДУКЦИЮ БАКТЕРИОЦИНОВ
*Streptococcus equi***

© 2002 г. Е. А. Котельникова¹, М. С. Гельфанд²

¹ Государственный научно-исследовательский институт генетики и селекции промышленных микроорганизмов, Москва 113545; e-mail: kkot2001@yandex.ru

² Integrated Genomics-Moscow, Москва 117333; e-mail: integratedgenomics@pochtamt.ru

Поступила в редакцию 30.07.2001 г.

Окончательный вариант получен 28.01.2002 г.

Производство бактериоцинов многими грам-положительными бактериями контролируется при помощи двухкомпонентной регуляторной системы, в которую входят сенсорный белок и регулятор ответа. В данной работе методами компьютерного анализа описан локус генов, ответственных за синтез бактериоцинов класса II, в геноме *Streptococcus equi*. Предсказаны потенциальные регуляторные сайты, узнаваемые ДНК-связывающим белком соответствующей двухкомпонентной системы, представляющие собой прямые повторы.

Большинство бактерий обладают способностью влиять на рост и развитие в среде других, обычно близкородственных, видов. Часто этот процесс связан с продукцией этими бактериями белковых молекул, называемых бактериоцинами грам-положительных бактерий, что в основном связано с их использованием в пищевой промышленности, например для консервирования продуктов, а также в медицине в качестве антибиотиков. Бактериоциты грам-положительных бактерий обычно подразделяются на три класса: класс I – лантибиотики [1, 2]; класс II – небольшие (<10 кДа), относительно термоустойчивые немодифицированные бактериоцины [3]; класс III – большие, неустойчивые к изменениям температуры белки.

Многие из грам-положительных бактерий, вырабатывающих бактериоцины, обладают системой “quorum sensing”, т.е. способностью контролировать синтез различных белков в зависимости от плотности клеток в среде [1, 4]. Для обеспечения этого механизма используется двухкомпонентная регуляторная система.

Клетка синтезирует и преобразовывает небольшие сигнальные молекулы, после чего секретирует их при помощи АТФ-зависимых транспортеров (ABC-транспортеров). Эти молекулы-феромоны являются входным сигналом для специфических мембранных сенсорных белков, относящихся к бактериальной двухкомпонентной системе, в состав которой входят сенсор и регулятор ответа. Сенсор, гистидинкиназа, при взаимодействии с феромоном фосфорилируется, после чего фосфорилирует регуляторный белок, который становится активным и запускает транскрипцию регулируемого оперона. Для некоторых локусов

бактериоцинов второго класса известно, что при этом регуляторный белок связывается в промоторной области с прямыми повторами определенного вида, а именно 9–11 нуклеотидов, разделенных 10–24 нуклеотидами [1, 2]. Как правило, ген, кодирующий сигнальный пептид (часто это сам бактериоцин), при помощи которого организм и узнает о плотности клеток в среде, сам находится в регулируемом опероне, и таким образом осуществляется авторегуляция.

В настоящей работе сделана попытка применить методы сравнительной геномики [5–7] для описания локусов выработки бактериоцинов и их регуляции в геномах стрептококков, для которых отсутствуют экспериментальные данные.

МАТЕРИАЛЫ И МЕТОДЫ

Для поиска регуляторных последовательностей и сравнения геномов использовалась программа GenomeExplorer [8]. Поиск по базам данных проводился при помощи программы Blast [9]. Исследовались геномы *Streptococcus pneumoniae* [10] и *S. equi* [11], а также участки геномов *S. pneumoniae* (номер доступа GenBank AJ276410) и *S. thermophilus* [12], ответственные за синтез бактериоцинов, имеющиеся в базе данных GenBank [13].

РЕЗУЛЬТАТЫ

Сначала с использованием программ Blast и GenomeExplorer были идентифицированы гены, предположительно значимые для производства бактериоцинов. Для этого проводилось тотальное сканирование геномов, при котором идентифици-

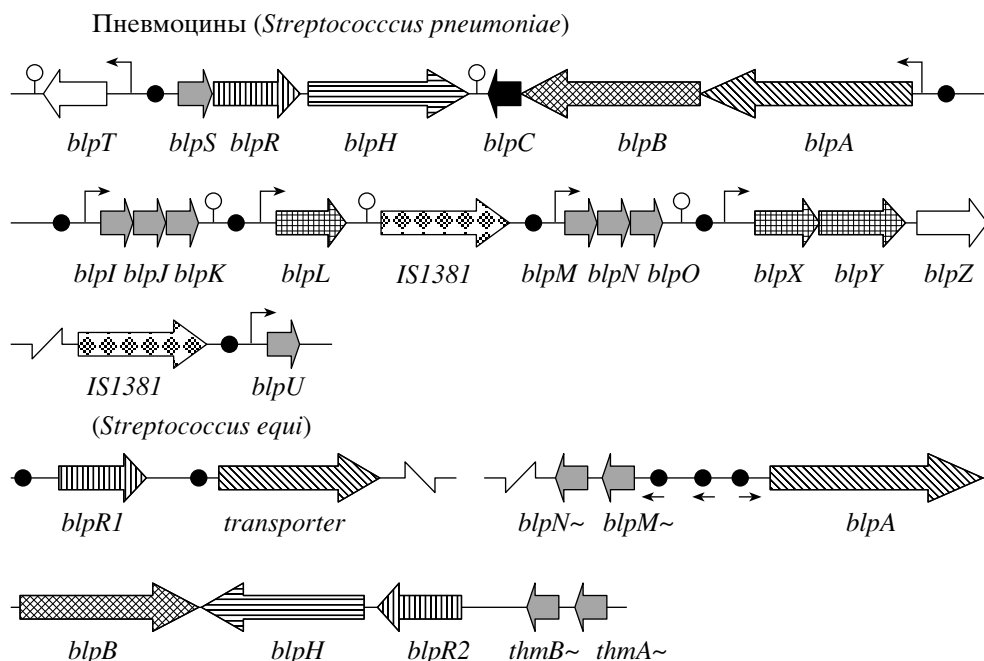


Рис. 1. Схематическое представление генетической организации локусов, ответственных за продукцию бактериоцинов в *S. pneumoniae* и *S. equi*. Последовательности в *S. equi*, гомологичные генам *rpc* из *S. pneumoniae* или *thm* из *S. thermophilus*, обозначены при помощи тильды. Черным цветом обозначен ген, кодирующий предшественника пептид-феромона, серым – бактериоцин-подобные пептиды, белым – неизвестные гены; вертикальная штриховка – регуляторы ответа, горизонтальная штриховка – сенсорные гены, наклонная штриховка и наклонная сетка – транспортные гены и вспомогательные к ним белки соответственно, вертикальная сетка – гены иммунитета. *IS1381* – мобильный элемент. Промоторы обозначены прямоугольными стрелками, прямые повторы – черными кружками, белые кружки – известные терминаторы транскрипции.

ровались все гены, кодирующие белки, сходные с известными пептидами-предшественниками бактериоцинов и сигнальных пептидов, а также белками транспорта, регуляторной системы, процессинга и иммунитета. Таким образом были описаны локусы выработки бактериоцинов класса II в геноме *S. equi*. Один из этих локусов содержит транскрибируемые в одном направлении гены регулятора ответа двухкомпонентной системы и транспортного белка, другой – транскрибируемые навстречу друг другу гены регулятора ответа двухкомпонентной системы и транспортного белка, другой – транскрибируемые навстречу друг другу гены транспорта, сходные с системой *comAB* [4] в *S. pneumoniae*, и гены двухкомпонентной системы (рис. 1).

Для выделения предположительных мест связывания регулятора ответа с ДНК в *S. pneumoniae* производился поиск повторов заданного вида в некодирующих областях перед уже известными генами, предположительно значимыми для продукции бактериоцинов. Для этого использовалась программа SignalX (входящая в комплект поставки GenomeExplorer).

В тех случаях, когда расстояние между генами, транскрибируемыми в одном направлении, меньше 100 нуклеотидных пар, считалось, что они

входят в один оперон, и повторы искались уже перед этим опероном.

В результате были найдены повторы длиной 11 нуклеотидов, разделенные 10 нуклеотидами, в областях перед различными генами (таблица, рис. 1). В таблице приведены указанные повторы, которые в дальнейшем использовались для поиска регуляторных сигналов в *S. equi*. Уже после завершения этого исследования нам стала известна работа [14], в которой делается предположение о значимости данных участков *S. pneumoniae*.

По обучающей выборке, приведенной для *S. pneumoniae* в таблице, строилась матрица значимости нуклеотидов для каждой из позиций (весовая матрица) [6]. Далее, используя процедуру поиска последовательностей регуляторных сайтов по весовой матрице, реализованную в программе GenomeExplorer, производился поиск похожих сайтов в доступных полных геномах грамположительных бактерий. Варьируя параметры поиска, мы нашли в геноме *S. equi* повторы, сходные с повторами пневмоцинового локуса. Эти повторы располагались в областях 3' от генов, связанных с синтезом бактериоцинов. При более внимательном исследовании с использованием процедуры ДНК-белкового выравнивания, на близлежащих участках генома, не содержащих

B1pM~ MNTTLMKQFNIIDADKLAHIEGGKNNWQANVWEGGTAAVSGWGLGTAICAASGAGAPFMGACGYIGAKFGVSLWAGVGTGATGGF
 B1pM MDTKIMEQFHEMDITMLSSIEGGKNNWQTNVLEGGAAFGGWGLGTAICAASGVGAPFMGACGYIGAKFGVDLWAGVGTGATGGF

 B1pN~ MTHYYQIKEEELNDICGGNPAGAAVVGGLTCAVGGIKLGSRIQPWGAVIGGVGSAAVCGYLAYKASS
 B1pN MNTYCNINETMLSEVYCGNSGGAAVVAALGCAAGGVKYGRLLGFPWGAAGGIGGAVVCGYLAYTATS

 ThmA~ MNTETINQFDVMNIEALATVEGGYSGKDKLKDMDGMYMLIGAGSG ALWGPAPAGGIGVLPGAFVGAHVGAIGGGFACMGGMIGNRFN
 ThmA MNTITICKFDVLDAEILLSTVEGGYSGKDKLKDMDGMYALAGAGSG ALWGPAPAGVGAIPGAFVGAHVGAIAAGGFACMGGMIGNKFN
 B1pJ MNTKMLSQLEVMDEMLAKVEGGYSSDTCQNALITGVTTGITGGTGAGLATLVAGLAGAFVGAHVGAIAIGGGITCLIGGMVGDKLGLSW

 ThmB~ MKQYEDFQMLYELDLASVNGGKINWGSVAGHCAGGAI VTGAFSGPLIYVLGAGIGCIVGAGQAI INGL
 ThmB MKQYNGFEVLHELDLANVTGGQINWGSVWGHCI GGAI IGGAFSG GAAAGVGCIVGSGKAI INGL
 B1pI MNTKMMEQFSVMDNEELEIVS GGRGNLGS AIGGCIG AVLLAAATG PITGGAATLIVGSGGIMSSL

Рис. 2. Белковые последовательности пептидов-предшественников B1pM, B1pN и B1pI, B1pJ из *S. pneumoniae*, ThmA и ThmB из *S. thermophilus*, а также гомологичных участков *S. equi*. Серым цветом отмечены потенциальные Gly-Gly-сайты расщепления, характерные для бактериоцин-подобных пептидов класса II.

Регуляторные повторы в *S. pneumoniae* и *S. equi*

Организм	Гены	Последовательности
<i>S. pneumoniae</i>	<i>blpX/pncM</i>	АТТСААГАТГТttcgatgacaАТТСААГАТТТ
	<i>BlpU</i>	АТТСААГАСГТttcgatgccaАТТСААГАТТТ
	<i>blpI/pncA</i>	АТТСААГАСГТttcgatgacaАТТСААГАТСТ
	<i>blpM/pncI</i>	АТТСААГАСГТttcgatgactАТТСАААТСТ
	<i>blpA/iA</i>	АТТСААГААГТttaaagactАТТСААГАТТТ
	<i>blpT</i>	АТТСААГАСАТtccaatgacaАТТСААГАТТТ
	<i>blpL/pncG</i>	АТТСААГААГТttagatgaccАТТТААГАТТТ
<i>S. equi</i>	<i>blpA</i>	АТТТААГАСГТtccaacgactАТТСААГАСТА
	<i>blpA</i>	АТТСААГАСГТttcgacgacaТТТТААГАСТТ
	<i>blpM~pncI~</i>	АТТСААГАСГТttcgacgacaТТТТААГАСТТ
	<i>blpRI</i>	АТТТААГАСАТaatcagtaccАТТТААГАТТТ
	<i>transporter</i>	АТТСААГАСААaataagaaccАТТСААГАТТТ

Примечание. Заглавными буквами обозначены нуклеотиды, входящие в повтор. Обозначения генов даны в соответствии с двумя вариантами аннотации соответствующего участка генома: в одном случае *pnc* (от английского названия бактериоцина pneumocin), а в другом *blp* (бактериоцин-подобный пептид, bacteriocin-like peptide).

аннотированных генов, были обнаружены небольшие последовательности, гомологичные последовательностям, кодирующим бактериоцин-подобные пептиды *BlpM*, *BlpN*, вырабатываемые *S. pneumoniae*. Гены *blpI* и *blpJ*, транскрибируемые совместно в *S. pneumoniae*, схожи с компонентами *thmA* и *thmB* двухпептидного бактериоцина термофилина 13 в *S. thermophilus*, которые в свою очередь являются наиболее близкими гомологами для пары предположительных бактериоцин-подобных пептидов в *S. equi* (рис. 2).

ОБСУЖДЕНИЕ

Используя сравнительный анализ геномов, а также поиск регуляторных последовательностей, реализованный в программе GenomeExplorer, мы нашли предположительные места связывания регулятора ответа двухкомпонентной системы *S. equi* с ДНК. Эти прямые повторы, схожие с повторами в *S. pneumoniae*, имеют структуру, характерную для регуляции продукции бактериоцинов грам-положительных бактерий второго класса. Кроме того, они располагаются в непосредственной близости от генов, гомологичных генам транспорта и двухкомпонентной системы *S. pneumoniae*. Это дает основание предполагать, что в *S. equi* указанные повторы могут отвечать за регуляцию продукции неизвестных ранее бактериоцин-подобных пептидов в зависимости от клеточной плотности.

Для локусов, связанных с синтезом бактериоцинов грам-положительными бактериями, характерна такая организация, когда гены двухкомпонентной системы располагаются последовательно в одном опероне, а ген феромона находится

либо в том же, либо в близлежащем опероне, регулируемом совместно с предыдущим. В случае *S. equi* нам не удалось обнаружить ген, который мог бы играть роль гена феромона. Кроме того, вблизи гена регулятора ответа *blpRI* двухкомпонентной системы не найдено соответствующей гистидинкиназы. Все это может быть связано как с обрывами контиг в местах, где могли бы быть данные гены, и расположением соответствующих генов в несеквенированных областях, так и с тем, что в данном организме регуляторная система кодируется иначе, нежели в известных ранее организмах.

Работа была частично финансово поддержана грантами INTAS (№ 99-1476) и ННМИ (№ 55000309).

СПИСОК ЛИТЕРАТУРЫ

1. Jack R.W., Tagg J.R., Ray B. Bacteriocins of gram-positive bacteria // Microbiol. Reviews. 1995. V. 59. № 2. P. 171–200.
2. Guder A., Wiedemann I., Sahl H.-G. Posttranslationally modified bacteriocins – the lantibiotics // Biopolymers (Peptide Science). 2000. V. 55. P. 62–73.
3. Nes I.F., Holo H. Class II antimicrobial peptides from lactic acid bacteria // Biopolymers (Peptide Science). 2000. V. 55. P. 50–61.
4. Kleerebezem M., Quadri L.E.N., Kuipers O.P., Vos W.M. Quorum sensing by peptide pheromones and two-component signal-transduction systems in gram-positive bacteria // Mol. Microbiology. 1997. V. 24. № 5. P. 895–904.
5. Gelfand M.S. Recognition of regulatory sites by genomic comparsion // Res. Microbiol. 1999. V. 150. P. 755–771.
6. Gelfand M.S., Koonin E.V., Mironov A.A. Prediction of transription regulatory sites in Archaea by a comparative

- genomic approach // Nucl. Acids Res. 2000. V. 28. P. 695–705.
7. *Overbeek R., Fonstein M., D'Souza M. et al.* The use of gene clusters to infer functional coupling // Proc. Natl Acad. Sci. USA. 1999. V. 96. P. 2896–2901.
 8. *Миронов А.А., Винокурова Н.П., Гельфанд М.М.* Программное обеспечение анализа бактериальных геномов // Молекуляр. биология. 2000. Т. 34. № 2. С. 253–262.
 9. *Altschul S.F., Gish W., Miller W. et al.* Basic local alignment tool // J. Mol. Biol. 1990. V. 215. № 3. P. 403–410.
 10. *Tettelin H., Nelson K.E., Paulsen I.T. et al.* Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae* // Science. 2001. V. 293. № 5529. P. 498–506.
 11. [ftp: // ftp.sanger.ac.uk/pub/pathogens.se](ftp://ftp.sanger.ac.uk/pub/pathogens.se)
 12. *Marciset O., Jeronimus-Stratingh M.C., Mollet B., Poolman B.* Thermophilin 13, a nontypical antilisterial poration complex bacteriocin, that functions without a receptor // J. Biol. Chem. 1997. V. 272. № 22. P. 14 277–14 284.
 13. *Benson D., Karsch-Mizrachi I., Lipman D.J. et al.* GenBank // Nucl. Acids Res. 2000. V. 28. № 1. P. 15–18.
 14. *Zaizieu A., Gardes C., Flint N. et al.* Microarray-based identification of a novel *Streptococcus pneumoniae* regulon controlled by an autoinduced peptide // J. Bacteriology. 2000. V. 182. № 17. P. 4696–4703.