

УДК 577.1

ПОИСК АЛЬТЕРНАТИВНЫХ ВТОРИЧНЫХ СТРУКТУР РНК, РЕГУЛИРУЮЩИХ ЭКСПРЕССИЮ БАКТЕРИАЛЬНЫХ ГЕНОВ

© 2003 г. Е. В. Любецкая*, Л. А. Леонтьев, М. С. Гельфанд¹, В. А. Любецкий*

Институт проблем передачи информации Российской академии наук, Москва, 101447

¹ Федеральное государственное унитарное предприятие “ГосНИИгенетика”, Москва, 113545

Поступила в редакцию 18.02.2003 г.

Экспрессия многих генов бактерий регулируется посредством образования альтернативных вторичных структур в лидерной области мРНК. Приведены результаты применения разработанного нами алгоритма поиска (по одной нуклеотидной последовательности) таких структур РНК для анализа оперонов биосинтеза аминокислот альфа- и гамма-протеобактерий. Предсказаны аттенюаторы этих оперонов в геномах малоизученных гамма-протеобактерий, в том числе *Shewanella putrefaciens*, и аттенюаторы триптофанового оперона ряда альфа-протеобактерий.

Ключевые слова: аттенюаторная структура, протеобактерии, регуляция экспрессии, алгоритм поиска.

ПОСТАНОВКА ЗАДАЧИ И МЕТОД ПОИСКА

Регуляция экспрессии многих бактериальных генов осуществляется на уровне трансляции или в результате взаимодействия процессов трансляции и транскрипции. В большинстве случаев в рассматриваемой области РНК образуются альтернативные вторичные структуры, которые служат основным регуляторным сигналом. В качестве примера можно привести аттенюаторы оперонов синтеза аминокислот [1] и регуляторные структуры некоторых оперонов рибосомных белков *Escherichia coli* [2].

Стандартный подход к поиску таких регуляторных сигналов состоит в построении вторичных структур для новых геномов с помощью шаблонов (распознающих правил), основанных на уже известных подобных структурах из хорошо изученных геномов. В частности, такой метод применен в работах [1, 2]. Однако он имеет естественное ограничение: позволяя переносить известные данные о регуляции экспрессии на новые геномы, он не дает возможности изучать новые системы регуляции. В то же время среди многочисленных групп бактерий существуют и такие, которые не содержат хорошо изученных видов. Более того, не исключена возможность, что даже известные геномы обладают новыми регуляторными системами.

Поэтому возникает задача поиска регуляторных структур в РНК без использования подобных шаблонов. В общем виде эта задача, по-видимому, не может быть решена, так как число вторичных структур, потенциально способных возник-

нуть даже в относительно коротких фрагментах РНК, велико и неясно, как выделить среди них регуляторные. Однако можно воспользоваться тем, что многие регуляторные взаимодействия основаны на образовании альтернативных структур. Поэтому при наличии соответствующей компьютерной программы поиск альтернативных структур в РНК может производиться на следующей основе:

(1) сканируют достаточно протяженные участки, расположенные перед генами, и в них ищут потенциальные альтернативные вторичные структуры;

(2) если найденная структура настолько “хороша” (включает протяженные сильно и “правильно” перекрывающиеся спирали), что ее случайное появление кажется невозможным, то она подлежит экспериментальной проверке;

(3) если же предсказание альтернативной структуры менее надежно (т.е. неясен точный вид альтернативных спиралей или структура содержит короткие или несовершенные спирали), то следует анализировать возможность образования *аналогичной* (в некотором неформальном смысле) структуры перед тем же геном в ряде родственных геномов. При этом предполагается, что истинные вторичные структуры консервативны даже при изменении нуклеотидной последовательности (это предположение выполняется во многих случаях [1, 2]). Анализ производят сначала с помощью алгоритма, а затем вручную, поскольку нет четко сформулированного критерия консервативности.

Механизм регуляции биосинтеза некоторых аминокислот и аминоацил-тРНК-синтетаз в про-

* Эл. почта: lin@iitp.ru; lyubetska@iitp.ru

теобактериях основан на соотношении скоростей транскрипции и трансляции. Он экспериментально показан для таких оперонов, как *trp*, *his*, *ilvGMEDA*, *ilvBN*, *phe*, *thr*, *leu* в *E. coli* и *Salmonella typhimurium* [2]. В этой регуляции роль посредника выполняет рибосома.

На рис. 1 схематически показана такая регуляция: изображены две альтернативные конформации вторичной структуры лидерной (некодирующей) части мРНК, расположенной на 5'-конце перед генами биосинтеза аминокислот. Цифрами обозначены различные участки мРНК, вовлеченные в формирование вторичной структуры. Первая конформация (обозначенная 1 : 2 и 3 : 4) называется *терминирующей*, вторая (обозначенная 2 : 3) – *анти-терминирующей*.

В случае укладки мРНК в анти-терминирующую конформацию РНК-полимераза синтезирует весь транскрипт; если же мРНК приобретает терминирующую конформацию, то транскрипция преждевременно обрывается, не доходя до структурных генов оперона. Как видно из рис. 1, спираль, называемая *анти-терминатором*, складывается из частей “паузной” спирали и *терминатора*. Показанная вторичная структура мРНК (т.е. анти-терминатор или терминатор вместе с паузной спиралью) альтернативны друг другу. Атенуаторы транскрипции оперонов биосинтеза ароматических аминокислот (*trp*, *pheA*, *pheST*) нескольких гамма-протеобактерий предсказаны ранее с помощью сравнительного анализа [1].

Нами разработан алгоритм и реализующая его компьютерная программа [3], которая позволяет отобразить для последующего ручного анализа набор попарно наиболее сходных вторичных структур (и составляющих их спиралей) по данному набору регуляторных областей гомологичных оперонов в родственных геномах. Конечно, критерии “хорошести” (пункт 2) и “аналогичности” (пункт 3), использованные при поиске альтернативных структур, по необходимости приблизительны, и относительно них “наилучшая” и “наиболее консервативная” (т.е. даже оптимальные по соответствующим критериям) альтернативные структуры являются лишь предположительными. Для их подтверждения требуется экспериментальная проверка либо использование дополнительных соображений, таких как наличие последовательности, кодирующей сигнальный пептид.

Принципиальная особенность подхода, примененного в настоящей работе, состоит в том, что мы проверяли впервые предложенный нами алгоритм решения задачи из пункта 1, который работает на *одной* исходной последовательности. Набор родственных регуляторных областей из пункта 3 привлекается только для последующего анализа и оценки результатов работы этого алго-

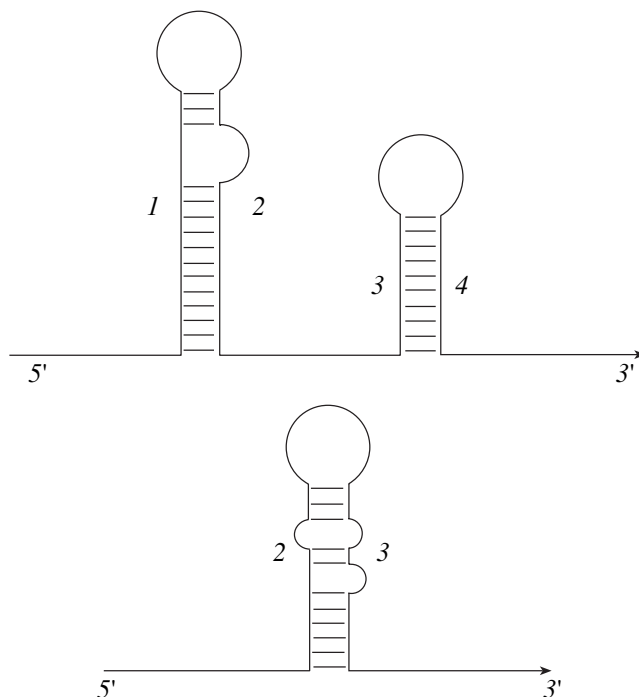


Рис. 1. Альтернативные вторичные структуры мРНК, составляющие типичный аттенуатор транскрипции.

ритма, при этом используются самостоятельные алгоритмы из [3].

В этом состоит основное отличие предлагаемого здесь подхода от тех, которые опубликованы, в частности, в [4–7]. В этих работах либо используется сравнение с уже известным шаблоном, либо с помощью анализа компенсаторных замен в выравненных нуклеотидных последовательностях строится консервативная вторичная структура.

Напомним стандартные определения. Вторичная структура состоит из *спиралей*, а спираль состоит из двух упорядоченных – левого и правого – наборов (каждый из одинакового числа) *отрезков* нуклеотидов. В каждом наборе соседние отрезки разделены какими-то *промежутками*, которые либо не имеют пары в другом наборе (тогда этот промежуток называется *вытягиванием*), либо имеют такую пару (и тогда эта пара промежутков называется *внутренней петлей*). Промежуток между самими наборами называется *внешней петлей*. Каждый *i*-ый отрезок от начала левого набора должен быть комплементарным *i*-ому отрезку от конца правого набора. Эти отрезки удобно нумеровать комплементарными парами, начиная от внешней петли. Такая нумерация связана с представлением о том, что первые от внешней петли отрезки (точнее, пары таких комплементарных отрезков) обычно определяются алгоритмами более точно, а за ними идет “лапша” – гораздо менее точно определяемые пары отрезков.

Опероны биосинтеза аминокислот гамма-протеобактерий и предсказанные аттенуаторы

Организм	Оперон, ген							
<i>Escherichia coli</i>	**trpEDCBA	**pheST	**pheA	**hisGDCBHAFI	**thrABC	**leuABCD	**ilvGMEDA	**ilvBN
<i>Salmonella typhi</i>	**trpED #trpCBA	**pheST	**pheA	*hisGDCBHAFI	*thrABC	*leuABCDxx	*ilvGMEDA	*ilvBN
<i>Yersinia pestis</i>	**trpEGDCBA	*pheST	**pheA1 **pheA2	*hisGDCBAHAFI	*thrABC	*leuABCD	*ilvGMEDA	*ilvBN
<i>Vibrio cholerae</i>	**trpEGDCBA	#pheST	**pheA	*hisGDCBHAFI	*thrABC	*leuABCDx	*ilvGMEDA	Нет гена
<i>Haemophilus influenzae</i>	#trpEGDC #trpBA	#pheST	#pheA	*hisGDCBHAFI	*thrABC	*leuABCD	Нет гена	#ilvBN
<i>Shewanella putrefaciens</i>	*trpEGDCBA	#pheST	*pheA	*hisGDCBHAFI	*thrABC	*leuABCD	*ilvGDA	Нет гена
<i>Actinobacillus actinomycetemcomitans</i>	#trpEG #trpD #trpFC #trpBA	#pheST	#pheA	#hisGDC #hisBH #hisF	*thrAB #thrCx	Нет гена	#ilvGE	#ilvBN
<i>Pasteurella multocida</i>	#trpEG #trpDC #trpBA	#pheSTxx	#pheA	*hisG #hisDCBxHAFxA	*thrABCxxx	*leuABCD	*ilvGMxDA	Нет гена
<i>Klebsiella pneumoniae</i>	#trpEDC #trpB	*pheS pheT	*pheA	#hisD #hisBHA #hisFI	#thrABC	#leuA #leuC	*ilvGxEDA	*ilvBN
<i>Pseudomonas aeruginosa</i>	#trpE	#pheST	#pheA	#hisGDC #BH #hisAF	#thrA #thrC	#leuA #leuB #leuC #leuD	Нет гена	#ilvBN
<i>Xanthomonas axonopodis</i>	#trpExGxDCx #trpB #trpA	#pheST	#pheA	#xhisGDCBHAFI	*thrAB #thrC	Нет гена	*ilvCGMxleuA	Нет гена
<i>Erwinia carotovora</i>	*trpEGDCBA	*pheST	*pheA	*hisGDCBHA	*thrABC	*leuABCD	*ilvGED	#ilvBN
<i>Xylella fastidiosa</i>	#trpEGDC #trpBAx	#pheST	#pheA	#xhisGDCBHAFI	#thrABC	Нет гена	*xilvGxleuA	Нет гена

* Потенциальные кандидаты с лидерными пептидами, предсказанные нашим алгоритмом на роль регуляторных структур.

** Известные структуры с лидерными пептидами, найденные нашим алгоритмом и приведенные в [6, 7].

Наш алгоритм не дал предсказание, так как не выполняется одно из следующих условий: нет лидерного пептида при хорошей альтернативной структуре, нет хорошей альтернативной структуры, нет потенциального терминатора (polyT), слишком короткая область перед первым геном оперона.

Здесь “первые” отрезки – это обычно один–три отрезка (считая от внешней петли).

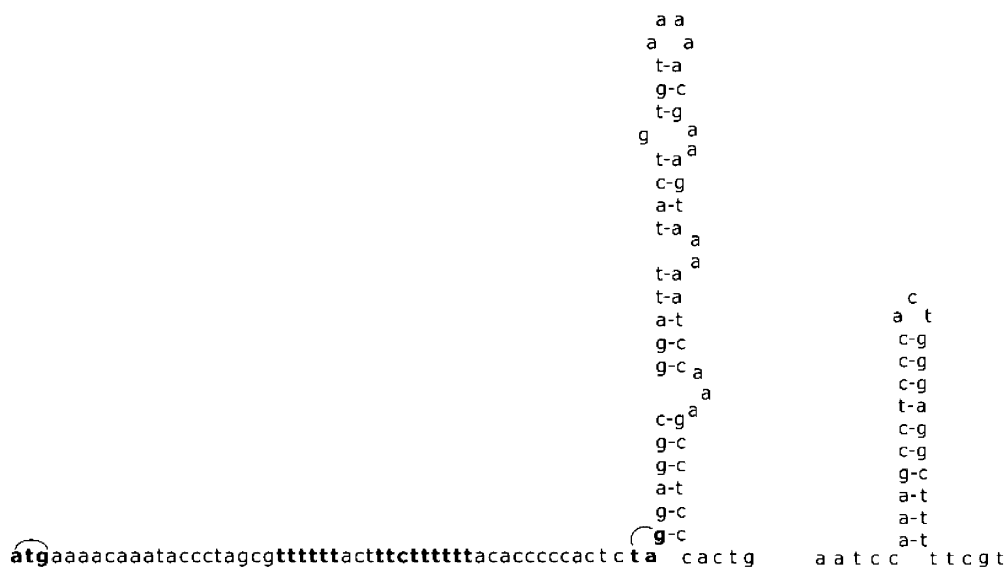
Обозначим концы левого набора A и B , а концы правого набора – C и D (нумерация нуклеотидов везде идет от начала исходной последовательности к ее концу). Отрезки от A до B и от C до D назовем *плечами* спирали (соответственно левым и правым). Таким образом, плечо – это набор отрезков вместе со всеми выпячиваниями и внутренними петлями (точнее, “половинками” этих петель) в нем.

Алгоритм [3], примерно квадратичный от размера исходных данных, имеет на входе любую нуклеотидную последовательность и выдает список потенциальных вторичных структур (в ней) типа альтернативных двухшпильчатой (терминатор–антитерминатор) или трехшпильчатой (терминатор–антитерминатор–пауза). В двухшпильчатом случае (с терминатором A' , B' , C' , D' и антитерминатором A , B , C , D) алгоритм проверяет

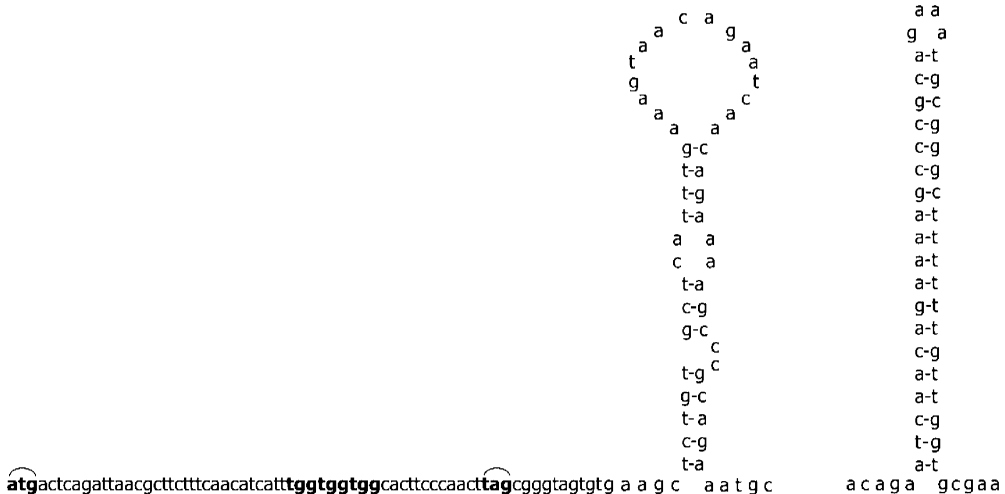
выполнение, например, таких естественных условий: $B < A'$, $D < D'$ и отрезки этих двух спиралей имеют не менее какого-то заданного числа общих нуклеотидов. Аналогично проверяются подобные условия в случае трех спиралей трехшпильчатой структуры.

В настоящей работе на примере регуляции экспрессии генов биосинтеза аминокислот оперонов *his*, *ilv*, *leu* и *thr* в альфа- и гамма-протеобактериях представлены результаты применения упомянутого алгоритма для поиска потенциальных альтернативных вторичных структур РНК.

В то же время алгоритм применим и к некоторым другим типам вторичных структур, но в этой статье мы рассматриваем только упомянутый тип. Алгоритм использует некоторые подходы, предложенные в [8]. Его систематическое описание приведено в [9], а исполнимый файл доступен на сайте <http://www.iitp.ru/lyubetsky>, поэтому здесь не приводится его описание. Заметим, что он хо-



pheA



trpE

Рис. 2. Предсказанные аттенуаторы транскрипции *Shewanella putrefaciens*.

рошо распараллеливается, т.е. вычисления могут выполняться одновременно на большом числе процессоров с общей памятью, что существенно для полногеномного анализа. Это основано, в частности, на следующем обстоятельстве: алгоритм просматривает входную последовательность с рамкой фиксированной ширины и последующие вычисления проводятся им параллельно для каждого окна.

Заметим еще одну особенность алгоритма. Поиск в каком-то смысле локально оптимальных вторичных структур обычно предполагает оптимизацию некоторого (числового) функционала,

зависящего от аргумента – переменной, которая пробегает множество вторичных структур или, по крайней мере, составляющих их спиралей. Известно, что такая оптимизация для сложных функционалов или обширных областей изменения аргумента алгоритмически сложна. Идея нашего алгоритма, в частности, состоит в том, чтобы вести оптимизацию не функционала, а на некотором (нелинейном) отношении порядка в множестве специально подобранных характеристик вторичных структур и составляющих их спиралей. Представляется, что легче указать правильное соотношение между характеристиками

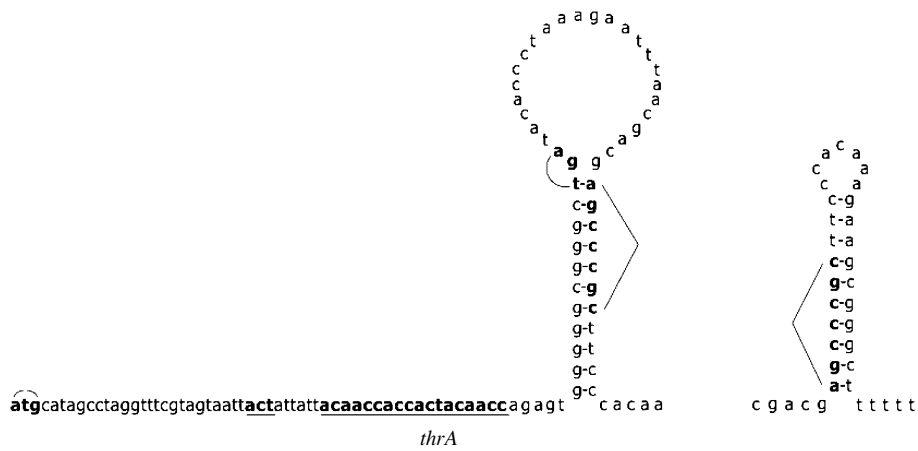
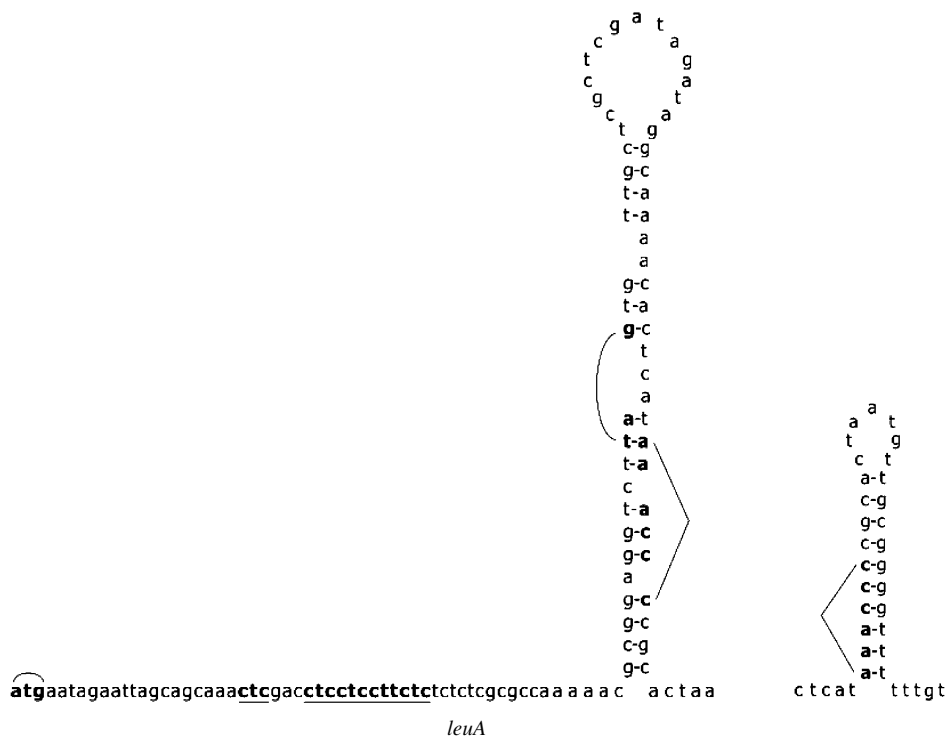
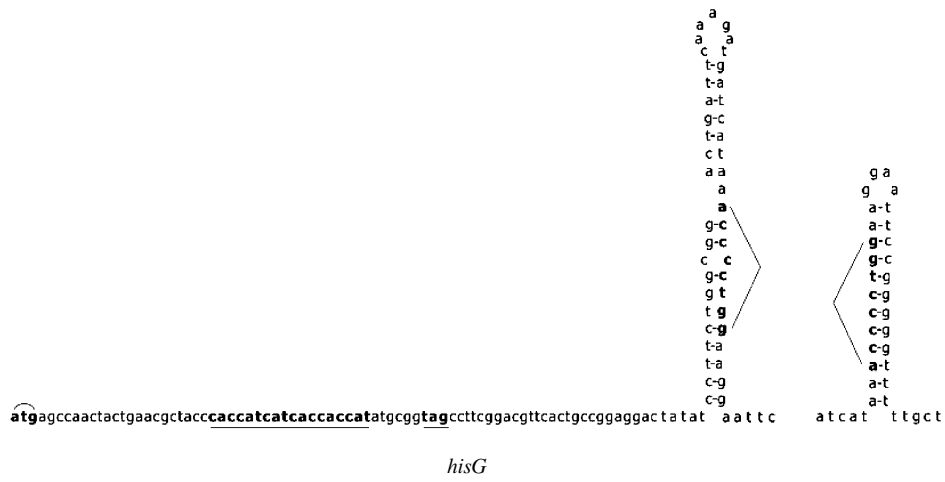


Рис. 2. Окончание.

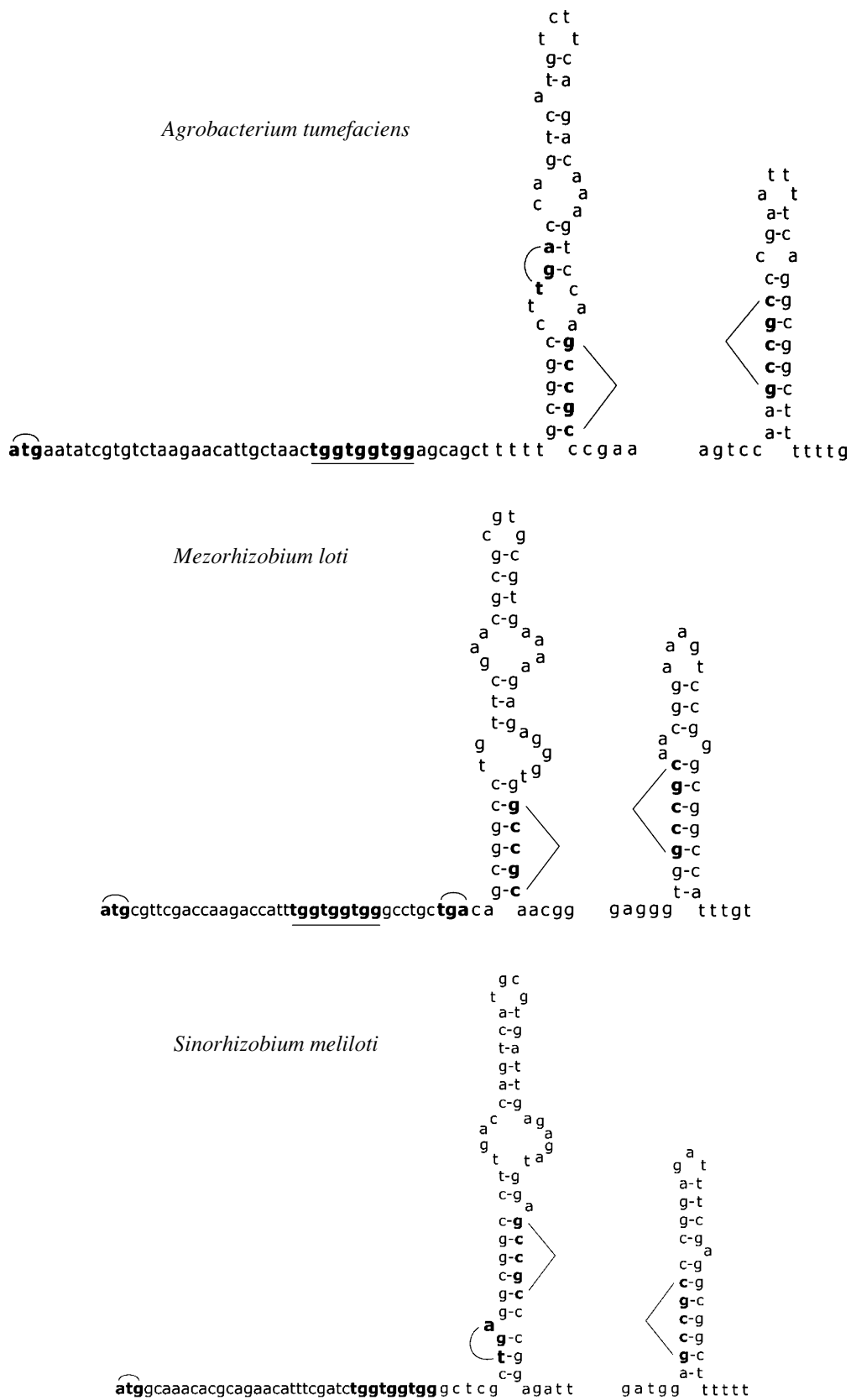


Рис. 3. Аттенуаторы оперонов *trpEG* альфа-протеобактерий.

вторичной структуры, чем числовой функционал, оптимизация которого приводила бы к разумному ответу.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Мы изучали геномы гамма-протеобактерий *E. coli*, *S. typhi*, *Klebsiella pneumoniae*, *Yersinia pestis*, *Vibrio cholerae*, *Haemophilus influenzae*, *Actinobacillus actinomycetemcomitans*, *Pasteurella multocida*, *S. putrefaciens*, *Pseudomonas aeruginosa* и альфа-протеобактерий (*Agrobacterium tumefaciens*, *Sinorhizobium meliloti*, *Mezorhizobium loti*, *Rhodospseudomonas palustris*, *Caulobacter crescentus*, *Brucella melitensis*), а также бета-протеобактерии *Bordetella pertussis*, в которых анализировали опероны, ортологичные оперонам синтеза аминокислот в *E. coli* (а именно *trpEDCBA*, *pheA*, *pheST*, *hisGDCB*, *HAFI*, *ilvBN*, *ilvGMEDA*, *leuABCD*, *thrABC*).

Поиск ортологов проводили с помощью программы GenomeExplorer [7]. После нахождения нашим алгоритмом вторичных структур мы проводили выравнивание нуклеотидных регуляторных областей с целью контроля результатов и использовали для этого программу Clustal [10].

В геномах гамма-протеобактерий найден 121 оперон. Эти результаты приведены в таблице. Видно, что алгоритм нашел потенциальные аттенуаторные структуры в 59 случаях из 121. В 62 случаях, когда алгоритм не нашел трехшпильчатой регуляторной структуры в соответствующей регуляторной области и которые мы классифицировали как отрицательный результат работы нашего алгоритма, может оказаться, что в части случаев такая структура действительно отсутствует.

Все известные опубликованные аттенуаторные структуры в оперонах видов, указанных в таблице, были найдены с помощью нашего алгоритма. Помимо относительно простых предсказаний в геномах, таксономически близких к уже изученным, сделано несколько нетривиальных наблюдений. В частности, впервые предсказано существование аттенуаторов в геноме *S. putrefaciens*, которые регулируют опероны *ilvGMEDA*, *ilvB*, *trp*, *pheA*, *his* (рис. 2). Следует отметить, что соответствующие регуляторные области не могут быть выравнены.

Построены также потенциальные аттенуаторные области оперонов *trpEG* альфа-протеобактерий *A. tumefaciens*, *S. meliloti*, *M. loti*, *R. palustris*, *B. melitensis*, *B. pertussis* (часть из них приведена на рис. 3). Как показал анализ опубликованных данных, это первые примеры аттенуаторов в геномах альфа-протеобактерий. Эти регуляторные области не могут быть выравнены как друг с другом, так и с соответствующими областями гамма-протеобактерий (рис. 4). В то же время GC-бога-

тые комплементарные области антитерминатора и терминатора консервативны – похожи даже соответствующие тройки слов (будем обозначать их слева направо цифрами 2, 3, 4) в гамма-протеобактериях CGGGc-GCCCG-CGGGc и альфа-протеобактериях GCGGC-GCCGC-GCGGC (эти “слова 2-3-4” входят в конец правого плеча паузной шпильки, в начало левого плеча терминатора и в конец его правого плеча и определяют спаривание соответствующих плеч 2 и 3 или 3 и 4 аттенуаторной структуры). Это указывает на гомологичность аттенуаторов, т.е. на их происхождение от общего предка. Повышенная консервативность в этих областях может объясняться тем, что для сохранения обеих регуляторных спиралей, терминаторной и антитерминаторной, должны произойти сразу три (а не две, как в одной спирали) комплементарные замены.

Наконец, показан аттенуатор оперона *hisSxG* *C. crescentus*. Во всех других случаях наш алгоритм указал на мотивированное отсутствие аттенуаторной регуляции в классе протеобактерий.

Итак, тестирование показало, что предложенный алгоритм пригоден для предварительного распознавания альтернативных структур РНК, играющих роль в регуляции экспрессии генов.

В дальнейшем планируется провести систематический анализ полных бактериальных геномов, в том числе представителей малоизученных таксономических групп; усовершенствовать алгоритм, добавив в него автоматический анализ лидерных пептидов и других биологически значимых структур; рассмотреть структуры более сложные, чем две альтернативные спирали; изучить эволюцию аттенуаторных структур.

Авторы благодарят А.А. Миронова и А.Г. Витрецака за помощь и обсуждение результатов. Визуализация результатов была выполнена М.А. Ширшиным, авторы также благодарят его за помощь в программировании и счете. Авторы благодарны рецензенту, который указал на статью [11], не известную нам.

Работа получила частичную поддержку Медицинского Института Ховарда Хьюза (55000309).

СПИСОК ЛИТЕРАТУРЫ

1. Витрецак А.Г., Гельфанд М.С. 2000. Компьютерный анализ регуляторных сигналов в бактериальных геномах. Аттенуаторы оперонов метаболизма ароматических аминокислот. *Молекуляр. биология.* **34**, 545–552.
2. Landick R., Yanofsky C. 1996. Transcription Attenuation. In *Escherichia coli and Salmonella. Cellular and Molecular Biology*. Ed. Neidhardt F.C. Washington DC: ASM Press, 1, Ch. **81**, 1263–1286.
3. Горбунов К.Ю., Любецкая Е.В., Любецкий В.А. 2001. О двух алгоритмах поиска альтернативной

- вторичной структуры РНК. *Информационные процессы*. **1**, 178–187. (<http://www.jip.ru/>).
4. Gorodkin J., Stricklin S.L., Stormo G.D. 2001. Discovering common stem-loop motifs in unaligned RNA sequence. *Nucleic Acids Res.* **29**, 2035–2044.
 5. Eddy S.R. 2002. A memory-efficient dynamic programming algorithm for optimal alignment of a sequence to an RNA secondary structure. *BMC Bioinformatics*. **3**, 1–16.
 6. Waterman M.S. 1989. In *Mathematical Methods for DNA Sequences*. Ed. Waterman M.S. Boca Raton, FL: CRC Press, 185–224.
 7. Миронов А.А., Винокурова Н.П., Гельфанд М.С. 2000. Программное обеспечение анализа бактериальных геномов. *Молекуляр. биология*. **34**, 253–262.
 8. Верещагин Н.К., Любецкий В.А. 2000. Алгоритм определения вторичной структуры РНК. *Труды научно-исследовательского семинара логического центра ИФ РАН*. Вып. 14. М.: Изд-во РАН, 99–109.
 9. Леонтьев Л.А., Любецкая Е.В., Любецкий В.А. 2002. Модифицированный алгоритм поиска альтернативных вторичных структур РНК и результаты счета. *Информационные процессы*. **2**, 100–105 (<http://www.jip.ru/>).
 10. Jeanmougin F., Thompson J.D., Gouy M., Higgins D.G., Gibson T.J. 1998. Multiple sequence alignment with Clustal X. *Trends Biochem. Sci.* **23**, 403–405.
 11. Lathe W., Suyama M., Bork P. 2002. Identification of attenuation and antitermination regulation in prokaryotes. *Genome Biology*. **3**, preprint 0003. (<http://genomebiology.com/2002/3/6/preprint/0003>).
 12. Верещагин Н.К., Любецкий В.А. 2000. Алгоритм определения вторичной структуры РНК. *Труды научно-исследовательского семинара логического центра ИФ РАН*. Вып. 14. М.: Изд-во РАН, 99–109.
 13. Pesole G., Liuni S., D'Souza M. 2000. PatSearch: a pattern matcher software that finds functional elements in nucleotide and protein sequences and assesses their statistical significance. *Bioinformatics*. **16**, 439–450.
 14. Vitreschak A.G., Rodionov D.A., Mironov A.A., Gelfand M.S. 2002. Regulation of riboflavin biosynthesis and transport genes in bacteria by transcriptional and translational attenuation. *Nucleic Acids Res.* **30**, 3141–3151.
 15. Rodionov D.A., Vitreschak A.G., Mironov A.A., Gelfand M.S. 2002. Comparative genomics of thiamin biosynthesis in prokaryotes: new genes and regulatory mechanisms. *J. Biol. Chem.* **277**, 48949–48959.
 16. Витрещак А.Г., Бансал А.К., Титов И.И., Гельфанд М.С. 1999. Компьютерный анализ регуляторных сигналов в полных бактериальных геномах. Инициация трансляции оперонов рибосомальных белков. *Биофизика*. **44**, 601–610.
 17. Switzer R.L., Turner R.J., Lu Y. 1999. Regulation of the *Bacillus subtilis* pyrimidine biosynthetic operon by transcriptional attenuation: control of gene expression by an mRNA-binding protein. *Progress Nucleic Acid Res. Mol. Biol.* **62**, 329–367.
 18. Liu J., Turnbough C.L. Jr. 1994. Effects of transcriptional start site sequence and position on nucleotide-sensitive selection of alternative start sites at the *pyrC* promoter in *Escherichia coli*. *J. Bacteriol.* **176**, 2938–2945.
 19. Rutberg B. 1997. Antitermination of transcription of catabolic operons. *Mol. Microbiol.* **23**, 413–421.
 20. Perkins J.B., Pero J.G. 2001. In *Bacillus subtilis and its relatives: from genes to cells*. Eds Sonenshein A.L., Hoch J.A., Losick R. Washington, DC: American Soc. Microbiol., 279–293.
 21. Miranda-Rios J., Navarro M., Soberon M. 2001. A conserved RNA structure (thi box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. *Proc. Natl. Acad. Sci. USA*. **98**, 9736–9741.
 22. Stasinopoulos S.J., Farr G.A., Bechhofer D.H. 1998. *Bacillus subtilis* tetA(L) gene expression: evidence for regulation by translational reinitiation. *Mol. Microbiol.* **30**, 923–932.
 23. Keener J., Nomura M. 1996. Regulation of ribosome synthesis. In *Escherichia coli and Salmonella. Cellular and Molecular Biology*. Ed. Neidhardt F.C. Washington DC: ASM Press, 1, Ch. **90**, 1417–1431.
 24. Decatur A., McMurry M.T., Kunkel B.N., Losick R. 1997. Translation of the mRNA for the sporulation gene *spoIIID* of *Bacillus subtilis* is dependent upon translation of a small upstream open reading frame. *J. Bacteriol.* **179**, 1324–1328.

Search for Alternative RNA Secondary Structures Regulating Expression of Bacterial Genes

E. V. Lyubetskaya¹, L. A. Leont'ev¹, M. S. Gelfand², and V. A. Lyubetsky¹

¹*Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, 101447 Russia;*
E-mail: lin@iitp.ru, lyubetsk@iitp.ru

²*State Research Center GosNIIGenetika, Moscow, 113545 Russia*

Abstract—Expression of many bacterial genes is regulated by formation of alternative secondary RNA structure within the leader mRNA sequence. Our algorithm designed to search for these structures (basing on analysis of one nucleotide sequence) was applied to analyze operons of amino acid biosynthesis in alpha- and gamma- proteobacteria. The attenuators of these operons are predicted for genomes of some poorly known gamma-proteobacteria including *Shewanella putrefaciens*, attenuators of the tryptophan operon in some alpha-proteobacteria are also predicted.