



Институт проблем передачи информации
им. А.А.Харкевича РАН

Биоинформатика, или молекулярная биология *in silico*

М.С.Гельфанд

15 января 2008

Расшифрован геном!

расшифрован геном человека - Поиск в Google - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search

Address http://www.google.ru/search?complete=... Go Links

Google.ru/расшифрован геном человека

Веб Картинки Карты Новости Группы Почта Gmail

Google.ru/расшифрован геном человека

Поиск в Интернете Поиск страниц

Веб Результаты 1 - 10 из примерно 73 200 для **расшифрован геном человека**

[Расшифрована структура генома человека](#)
Статьи. Наука и техника

[Расшифрована структура генома человека](#). Татьяна БАТЕНЕВА. С одной стороны, можно считать, что начался отсчет новой эпохи, сулящей победу над болезнями, ...
n-t.ru/tp/in/rs.htm - 13k -
Сохранено в кэше - Похожие страницы

[Расшифрован геном человека](#)
Фактически это означает завершение первого этапа грандиозного исследовательского проекта, известного как "Геном Человека". ...
www.ma-ma.ru/news/167 - 25k -
Сохранено в кэше - Похожие страницы

[РИА Новости - Наука - Расшифрован геном человека](#)
textLooks span date { font-size:11px; color:#384787; padding-right: 10px; } </style> <div class="textLooks" >
<h5>Расшифрован геном человека</h5> <p><span ...
www.rian.ru/science/20030415/366929.html - 64k -
Сохранено в кэше - Похожие страницы

[Диабет-Новости | ДОСТИЖЕНИЯ / ОТКРЫТИЯ](#)
Специалисты утверждали, что на реализацию проекта "Геном человека" понадобится более 20 лет, но его **расшифровали** за 6. Уже в ближайшие 10 лет ученые ...
www.diabet-news.ru/info/otkr/otkr10.shtm - 37k -
Сохранено в кэше - Похожие страницы

[Впервые расшифрован геном конкретного человека](#)
В июне 2000 года совместно с руководителем проекта "Геном человека" Френсисом Коллинсом доктор Вентер объявил о первой реконструкции полного генома человека ...
www.bilogs.net/2007-09-04/vpervie_rasshifrovan_genom_konkretnogo_cheloveka.html - 11k -

genome deciphered - Поиск в Google - Microsoft Internet Explorer

View Favorites Tools Help

Back Forward Stop Home Search Favorites

Address http://www.google.ru/search?complete=1&... Go Links

Google.ru/human genome deciphered

Картинки Карты Новости Группы Почта Gmail

Google.ru/human genome deciphered

Поиск в Интернете Поиск страниц

результаты 1 - 10 из примерно 373 000 для **human genome deciphered**

[e.gov | 2000 Release: Two-Thirds Human Genome Sequenced](#) - [Перевести эту страницу]
Is of **Human DNA Script Deciphered** by Human Genome Project. Public Consortium To Complete "Working Draft" of Human Genome. March 2000 ...
genome.gov/10002080 - 15k -
Сохранено в кэше - Похожие страницы

[Ma Mania — Deciphering the human genome: what does it mean](#)
[Перевести эту страницу]
... gave extraordinary attention to the recent announcement of the draft sequence of the **human genome** by the Human Genome Project. ...
www.wersingenesi.org/docs2001/0309_genome.asp - 31k -
Сохранено в кэше - Похожие страницы

[NewsHour: Breaking the Code -- December 2000](#)
[Перевести эту страницу]
... genetic sequences would be comparatively easy to decipher with the aid of computers. If you printed out the whole **human genome** on pages-- and I don't know if ...
www.newshour.com/bb/health/july-dec99/dna_12-2.html - 27k -
Сохранено в кэше - Похожие страницы

[The Genome Deciphered, May Herald Medical Breakthroughs](#) - [Перевести эту страницу]
The **Genome Deciphered**, May Herald Medical Breakthroughs ... Compared to the **human genome**, the **genome** is only about 1.5 percent different, ...
nationalgeographic.com/news/2007/04/070412-macaque-genome.html - 27k -
Сохранено в кэше - Похожие страницы

[Scientists Start Deciphering Dog Genome](#) - [Перевести эту страницу]
A dog's best friend may soon have its **genome** deciphered. The National **Human Genome** Research Institute (NHGRI) in Bethesda, Maryland, recently added a dog to ...

расшифрован геном - Поиск в Google - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites

Address http://www.google.ru/search?complete=1&hl=ru&rlz=1T4GFRC_ruRU2... Go Links

Google.ru/расшифрован геном

Веб Картинки Карты Новости Группы Почта Gmail

Google.ru/расшифрован геном

Поиск в Интернете Поиск страниц на русском

Веб Результаты 1 - 10 из примерно 87 900 для **расшифрован геном**. (0,27 секунд)

[Расшифрован геном кошки](#) - Медицинская библиотека сервера MedLinks.ru

[Расшифрован геном кошки](#) - Международная группа ученых, координируемая специалистами американского Национального института рака (National Cancer Institute), ...
www.medlinks.ru/article.php?sid=30736 - 46k -
Сохранено в кэше - Похожие страницы

[Расшифрован геном винограда Pinot Noir](#) | Газета.Ru: Новости

[Расшифрован геном винограда Pinot Noir](#). Группа генетиков из Италии, США, Бельгии и Великобритании завершила предварительную расшифровку генома одного из ...
www.gazeta.ru/news/science/2007/12/19/n_1154765.shtm - 79k -
Сохранено в кэше - Похожие страницы

[Элементы - новости науки: Полностью расшифрован геном собаки](#)
Наряду с полной расшифровкой генома боксера были частично прочтены генетические коды еще нескольких пород собак, а также родственных видов — волка и койота. ...
elementy.ru/news/165035 - 30k -
Сохранено в кэше - Похожие страницы

[Коммерческая биотехнология | Расшифрован геном опоссума](#)
Коммерческая биотехнология - сайт для всех, кто участвует в развитии российской биотехнологии: ученых, бизнесменов, производителей и продавцов оборудования ...
www.cbio.ru/modules/news/article.php?storyid=2784 - 36k -
Сохранено в кэше - Похожие страницы

[Расшифрован геном устойчивого к антибиотикам штамма туберкулеза](#)
Сегодня поступило сообщение, что исследователям из ЮАР удалось полностью секвенировать геном устойчивой к антибиотику формы туберкулеза. ...
md.cnews.ru/natur_science/news/line/index_science.shtml?2007/10/12/270287 - 27k -
Сохранено в кэше - Похожие страницы

[ПОЛИТ.РУ \ НАУКА \ Расшифрован геном червя, вызывающего слоновью ...](#)

[Расшифрован геном червя, вызывающего слоновью болезнь](#). Как сообщает Reuters, международная группа ученых под руководством Элоди Гедин (Elodie Gedin) ...

Перехватить зашифрованное сообщение -

еще не значит его понять

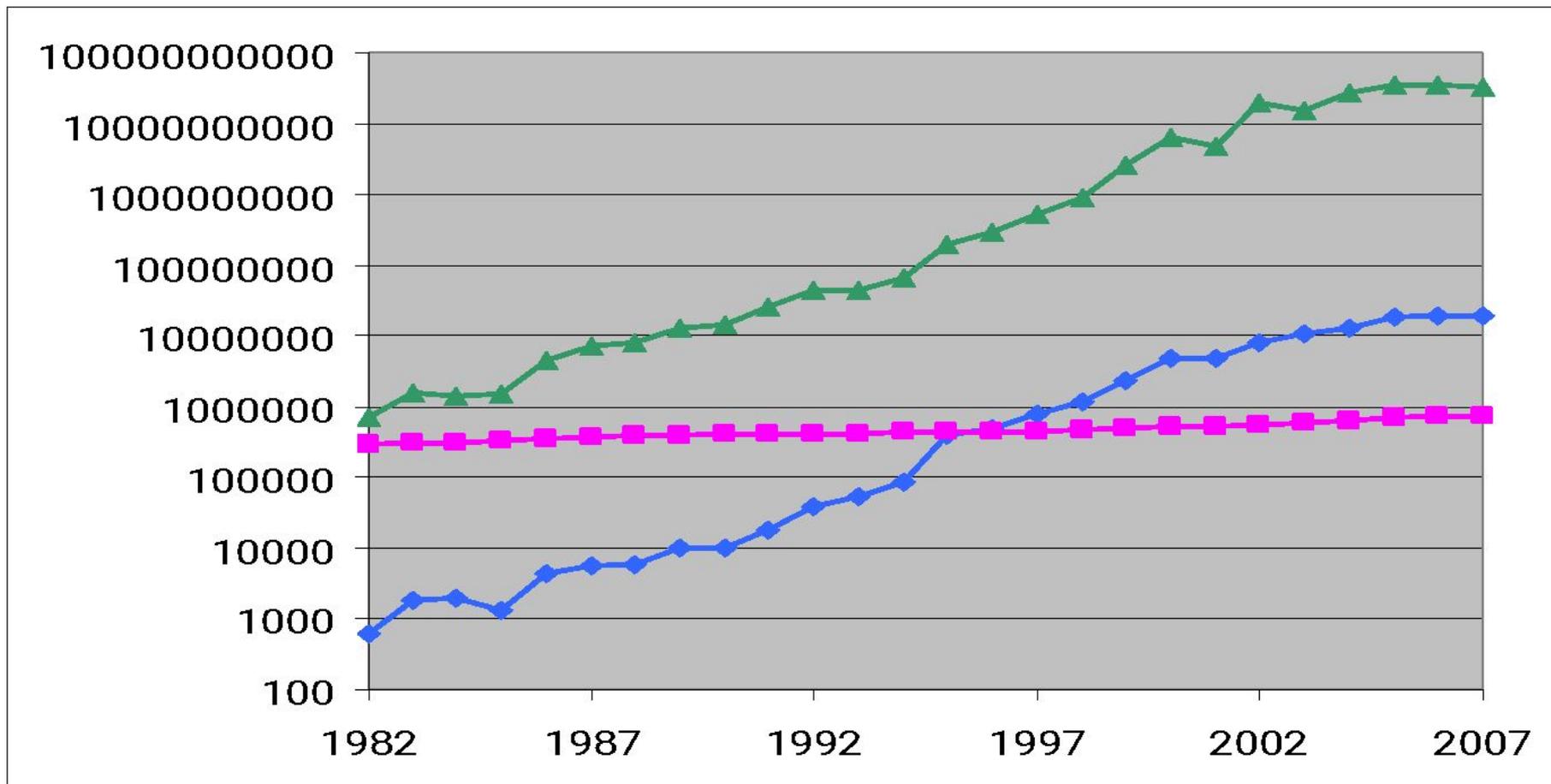
aacgggcaatatgtctctgtgtggattaaaaaagagtgctgatagcagctcttgaactggttacc tggcgtgagtaaataaaa ttttat ttagac ttaggtcactaaatactttaaccaat atagga tagcgcaagacagataaaaat tacagagtac acaac atcca tgaac cgc at tagca cccacc attac caccaccatc accat tacc acaggtaacggtgctggggctgacgggtacaggaaac acagaaaa aagcccgccacctgacagtgccgggttttttttcgacc aaaggtaacagaggta acaac catgc gagggttgaa gttcggcgggtacatcagtgggcaaaa tgcagaacgt tttctgccgggttgcc gat at tct ggaagc aatgc caggc agggg caggt ggcgaccgtc ctctc tgc cc ccgcc aaaa tacca accat cttgt agcga t gatt gaaaaaac attagcggcc aagat gcttt accca atatc agcga tgc cgaacgtat ttt tggcgaactctgac gggac tgc cc gccgc ccagc cggga tttcc gctgg cacaat tga aact t tgc tgc gacc aggaattgccc aaaa taaaa catgt cctgc atggc attag ttt gttggggcagtgccggatagc atcaacgctgctgatttggcctggtggagaaaaatgtcga tgc cc attatggcggcgtgttgaagcgcgtggtcacaacgtt accgt tatc gatccgggtcga aaaa actgcttgacgtggggc attac ctogaatctaccgttgat at tgc t g agtcc acccgccgtat tggcggcaagccgcattccggctgaccac atgggtgctga tggctggttt cactgccggtaatgaaaa aggcgagctggtggttc tgggacgcaacggttccgac tactc cgtcgcgggtgctggc ggccgtttta cgcgc c gattgttgc gaga tctggacggga tgttgacgggtgttta tactgcgatccggctca ggtgccc gatgc gagggtgttgaaagtc gatgt cctatcagga agoga tggagc tttc t tact tggc gctaaagttc ttcac cccc gcacc atcac cccca tgc cc cagttcaga tccc tgcctgat taaaaat accggaatc ctcaagctcc aggtacgctc attgggtgccagcgtgatga agacgaatta cgggtcaagggcatt tccaatctgaataa catggcaatg ttcagcgttccggcc cgggg gat gaaaaggatggt tggca tggcggcgcgtct tggcagc gatgcaac ggc cc gbtat tccg tgggtgctgat tacgc aatca tctccgaa tacag taca gtttc tgcgt tccgc aaagc gactggt gcgagctgaacggcaatgcaggaaga gttctacctggaactgaaagaaggc ttactggagc cgttggcgggtgacggaacggctggccattatctc ggtgg taggt gatggtatgc gcacc ttactggt gatctcggcga aattct tggcc ggcgtggccc ggc cc aatatcaaca ttgct gccat tgc t c aggga tcttc tgaac gctcaatctctgtc tgggtcaataacgat gatgc gaccactggc gtcgcgt tactcatcagatgctgt tcaat accga t cagg ttatc gaagt gtttg gattggcgt cgggtg cgttggcgg tgcgc tgc t g agc aactgaagcgtcagc aaagc tgggtgaagaataaacat atcga ctta cgtgtcgcgggtttgctaac t c gaaggcactgctcac caatgtacatggcct taatc tggaa aactg gcaggaagaa ctggcgc aagccaa agagc cgtttaatc tgggc gctta at tgcct cgtgaaagaatac atc tgc tgaaccc ggtca ttgttgactgt tacttccagc caggc agtggcggatcaata tggc gacttc ctgc gcgaa ggttccacgtt gttacgcc gaaca aaaaaggccaaac ctgc t c gatgga t t act accat cagttgcgttatgcggcggaaaaatcgcggcgtaaattcc tctatgacac caacgttggggcgtgattaccgggtatc gagaac ctgca aaatc tgc t c aatgc tgg tgatgaa ttgat gaagtctcc ggcatctcttcaggttcgctttcttatatc ttcggcaagttagac gaaggcatgagtttc tccgagggcaccac actggcggg gaaatgggttatacc gaacc gga cccgcgagatgactcttctggatgga tgtggcgcgt aagctattgattctc gctcgtgaaa cgggacgtgaaactggagctg cgggatattgaaat tgaac ctgtgctgccc gcagagtttaacgc cga ggggtgatgctgc cgtt ttagggca tctgtcacagctcgcagatc tcttggccgc gctgtggcgaaggc cgtgatgaa gaaaagtt tggcgtatg ttggc aatat tgatgaa gatggcgtctg cccgctgaagat tggcgaagtgatgg taatgatccgtgttcaaag tga aaaaatggcgaaa acgcc ctggc ctctc atagc cactattatcagcc gctgc cgttggtactgcccgat atgg tgc ggg caatgac gttac agctgcgggtgtctt tgc t gatctgctacgtacc tctca tggaa gttag gagtc tgaca tgggttaagttatgccc ggcttccagtgccaa tatgagcgtc ggggttgatgtgc tgggggc ggcgg tgaca cctgt tgatgtgca ttgctcggagatgta gtcac ggttgaggcg gcagagacat t cagtc tcaacaac ctggcagcct tggcc gataa gctgc cgtcagagcc acgggaaa atatcgtttatc agtgc tgggagcgtt tttgc cagga gcttggcaagcaaat tccag tggcg atgac tctgg aaaaagaaat gccgatcgg ttcgggcttaaggctc agcgcctgt t cagtggtcggcgg cgtgatggcgaatgaaca ctgcccgaagcgcct taatgacactcgtttgctggcctttg atgggcgagttggaa gggcgtatc tccggcagca t tcat tacga caacgtggcaccgtgttttc ttg gttggtatgcagttgatgattgaagaaa acgac atcatcagtc agcaagtgcc aggggtttgat gagggtgctgt ggggtgctggc gtatccggggatta aagtc tgcac ggcagaagcc agggc tttttac cggcgcagtatc gccgc cagga ttgca ttggcgcagggcgcac atctggcaggcttca ttcac gccctgctattcccgtcagcc tgagcttgc cgcga agctgatgaa agatgttatc gctga accctacc gtgaacggttac tggca ggttccggc aggcgcggca ggcgg ttgog gaaatcggcgcggta gcgagcggta tctcc ggctc cggc ccgac tttgt tgcctctgtg tgaca agccggatc cggcc agc gcgttgc c gactggttgggtbaa aaaaactacctgcaaaa t caggaaggt tttgt tcaata tttgc cggctggata cggcggggcgc acgag tactggaaa actaa atgaa actctacaatctgaa agatcaca atgagcaggtcagctttgcgca agccgt aacc caggggttgggcaaaa aatca ggggc tgttt tccc gcacg accctggcggaa atc agcctgactgaaattgatgagatgc tgaagctggat tttgtca cccgcag tgcga agatcctctc ggcgt ttattgggtgatgaaa tcccgcagga aatcc tggaa gagcgcgtac ggcgcgctt tggc t tccc ggctc cggtc gccaa tgttgaaagc gatgtcggttgct tggaaattgtcc acgggccaac gctggcat ttaaaa gtttgggggtcgtt t atggcaca aat actgacc atattggggcgttaagcc agtgaacattctgacccgcagatcc ggttatctggc cggcagtggtc atgcttt tgcacggcgatttcgatgc actacttgaagctggttgc gttttatgctgcgaccaa gttgtggaagagtgttccg ctgcccgtagcttatcgtgc caaaagagctggcagaacg tgc ttttatgca gcccggc

Геном бактерии: несколько миллионов нуклеотидов

От 600 до 9 тысяч генов (примерно 90% генома кодирует белки)

На этом слайде – 0,1% генома *Escherichia coli*

Экспоненциальный рост объема данных



красный - статьи (PubMed)

синий - последовательности (GenBank)

зеленый - объем в нуклеотидах (GenBank)

из 18 миллионов ссылок,
~675 тыс. отвечают на
"bioinforma* OR computa*"

16 тыс. "bioinforma*"

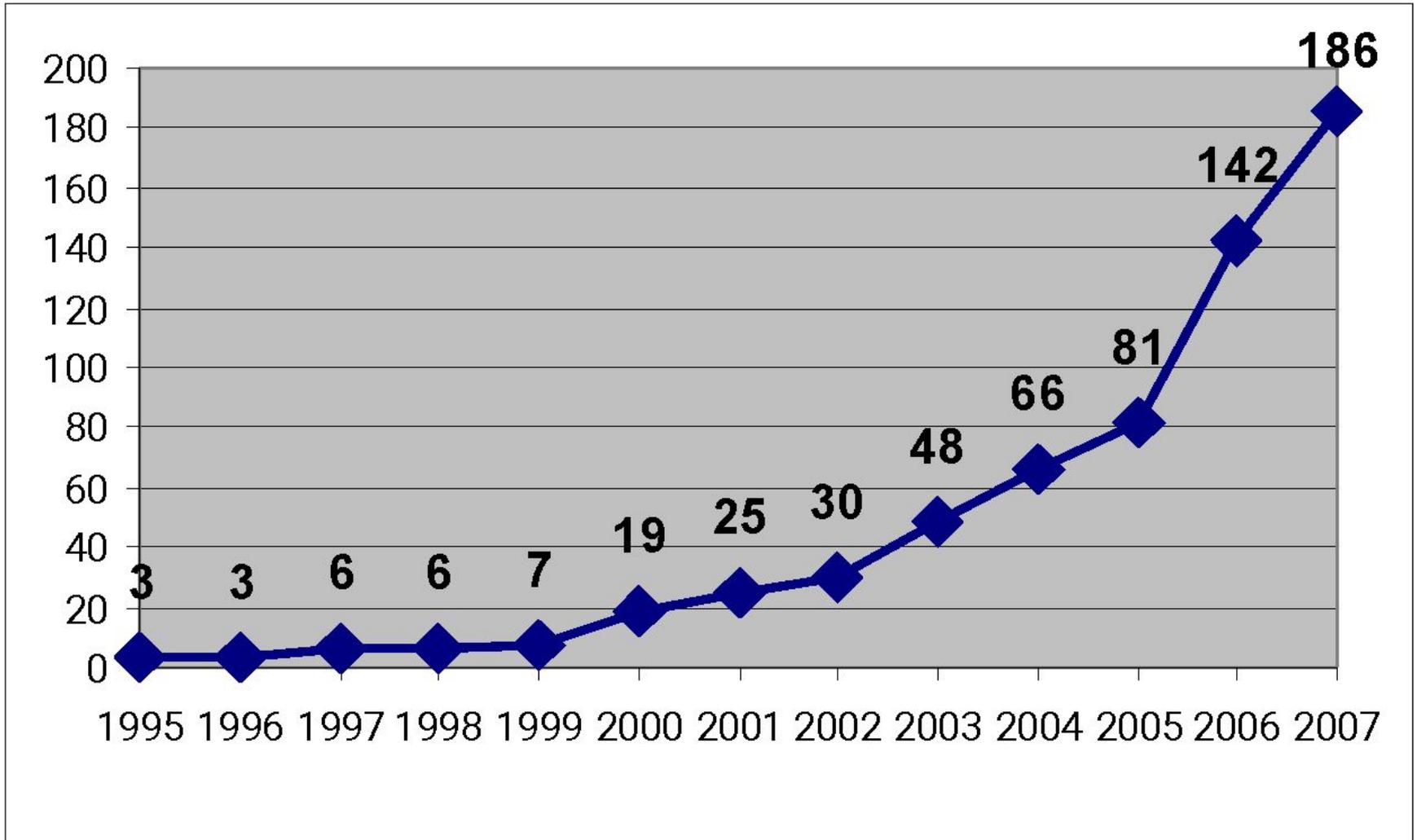
65 тыс. "bioinforma* OR computa*"

Цель (локальная): аннотировать гены / белки *in silico*

- Что?
 - (биохимическая) функция
 - клеточная роль
- Когда?
 - Регуляция
 - Экспрессия
 - Время жизни (мРНК, белка)
- Где?
 - Локализация
 - Внутри/снаружи
 - Органеллы и компартменты
- Как?
 - Механизм
 - Специфичность, регуляция

Наиболее важные предсказания затем
проверяются экспериментально

622 полных генома (прокариот)



Цель (глобальная)

Предсказать свойства организма путем
(компьютерного) анализа его генома

(возможно, с использованием дополнительной информации: эпигенетика, белок-белковые взаимодействия и т.п.)

сейчас: метаболическая реконструкция,
транспортные системы, ответ на стресс и т.д.

“Понять” эволюцию геномов/организмов

Сравнительная геномика

Базовые постулаты:

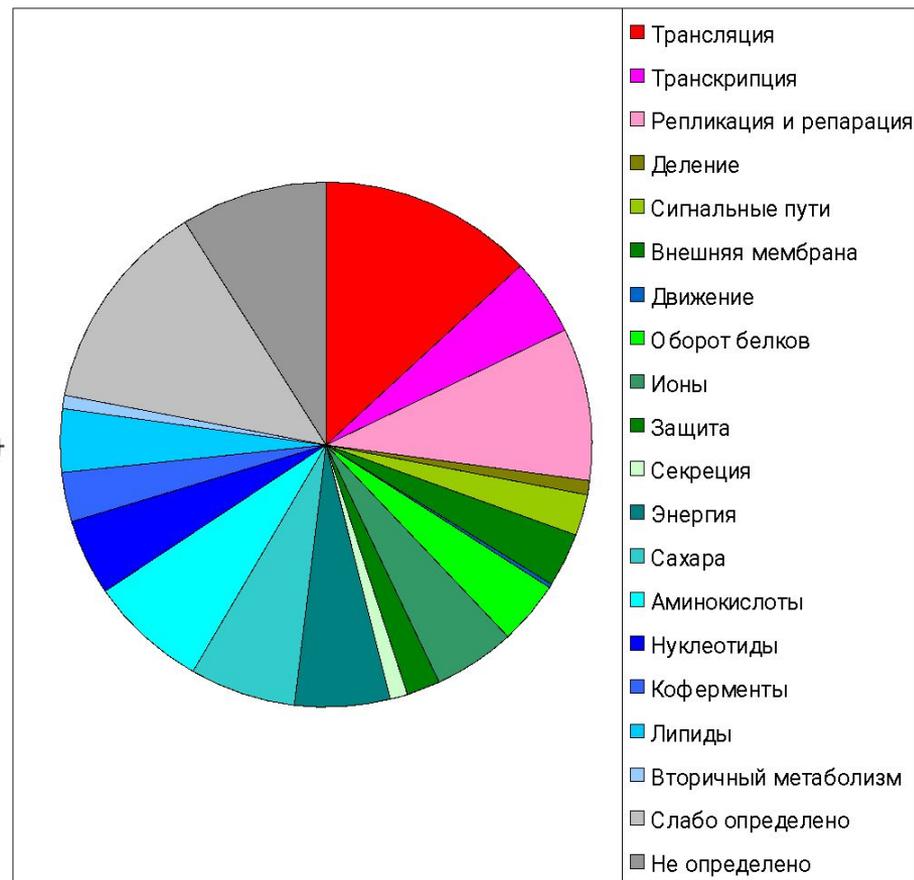
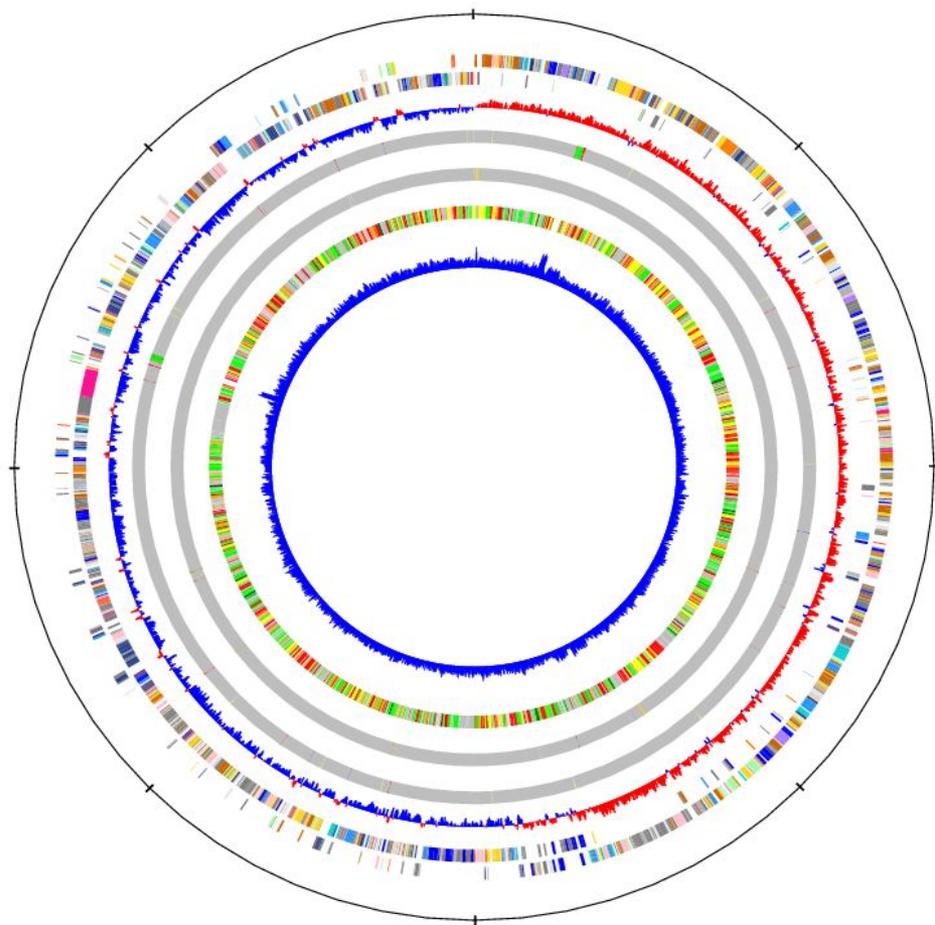
- Сходство => гомология
(общность происхождения)
- Гомология => сходная функция
- Консервативно то, что важно
 - структурные и функциональные мотивы в белках
 - регуляторные сайты в ДНК

Математические и алгоритмические проблемы

- **Формализация понятия сходства**
 - теория вероятностей: значимость наблюдаемого уровня сходства
 - вычислительная геометрия: сходство структур
- **Алгоритмические проблемы:**
 - быстрый поиск сходных последовательностей
 - большой объем базы данных (растет быстрее, чем быстродействие процессоров)
 - множественное выравнивание
 - оптимальный алгоритм имеет полиномиальное время работы, но степень равна числу последовательностей
 - построение эволюционных деревьев
 - баланс между биологическими соображениями и вычислительными возможностями
- **Идентификация функциональных и регуляторных мотивов в последовательностях**
 - теория предсказания образов: нейронные сети, поддерживающие вектора и т.п.

Первый российский бактериальный геном - *Acholeplasma laidlawii*

Секвенирование: ИФХМ МЗ РФ, аннотация: ИППТИ РАН

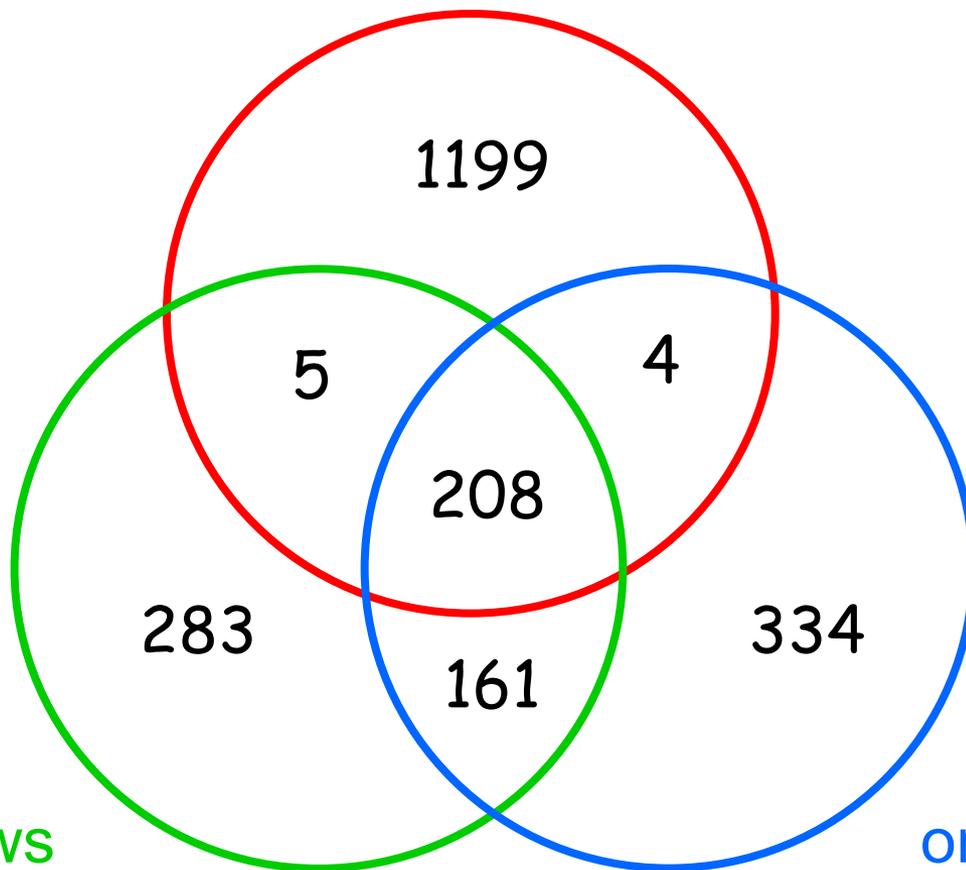


~1,5 Мб; ~1400 генов.

Установлены функции ~80% генов; проведена метаболическая реконструкция

Сравнение с родственными геномами

Acholeplasma laidlawii



aster yellows
Phytoplasma

onion yellows
Phytoplasma

Сравнительная геномика - 2

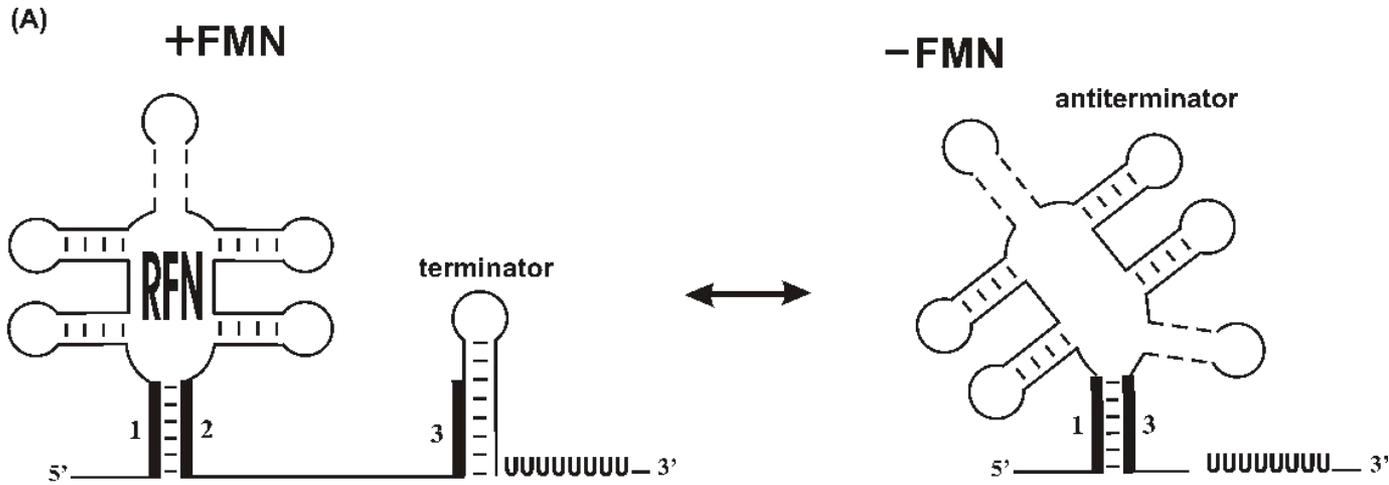
- Не обязательно последовательности:
 - структура белка и РНК
 - расположение генов на хромосоме (ко-локализация)
 - ко-регуляция и ко-экспрессия генов
 - филогенетические образцы (совместное появление в геномах)
- Предсказав структурные особенности белка, можно определить его функциональный класс
- Изучение геномного контекста позволяет отнести ген (белок) к функциональной подсистеме
- Задача: формализация этих подходов
 - Полногеномные сравнения
 - Статистическая значимость
 - Распознавание образов и экспертные системы

РНК-переключатели: от биоинформатического анализа к экспериментальной проверке

- Новый универсальный механизм регуляции экспрессии генов за счет формирования альтернативных структур РНК и прямого связывания малых молекул
- Структуры и механизм предсказаны биоинформатически и затем подтверждены экспериментально

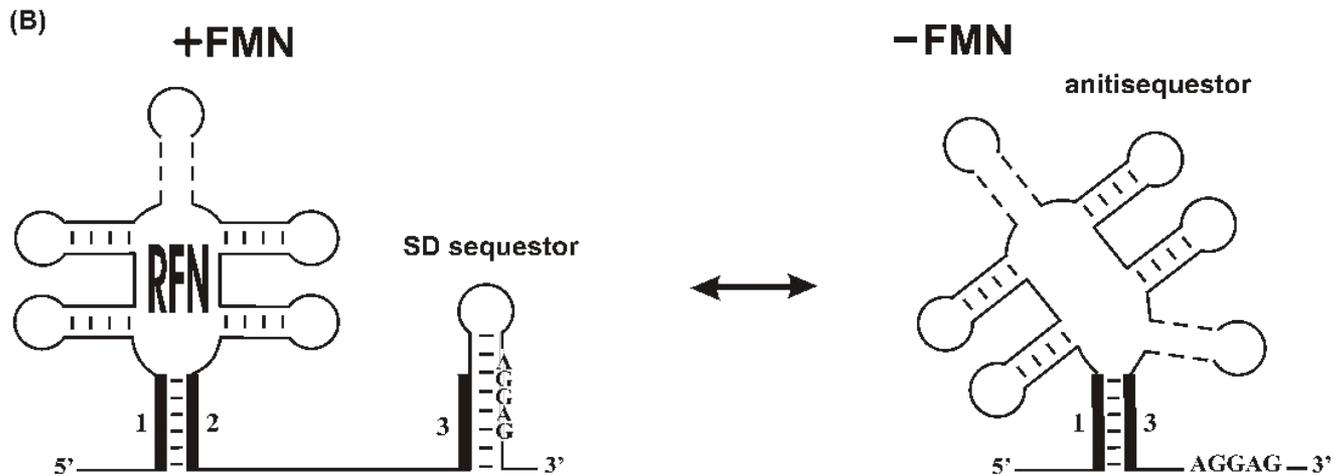
Регуляция экспрессии генов за счет формирования альтернативных структур РНК

• Transcription attenuation



Структура предсказана на основе сравнительного анализа выравненных последовательностей

• Translation attenuation



Механизм предсказан на основе литературных данных и анализа структурных особенностей

5'-нетранслируемые области бактериальных генов биосинтеза рибософина

	1	2	2'	3	Add.	3'	Variable	4	4'	5	5'	1'														
BS	TTGTATCTTC	GGGG-CAGGGT	TGAAAT	CCC	GACC	GGCCGT	21	AGCC	CCGT	GAC--	8	4	8	-----TGGATTCA	TTTAA-G	CTGAAG	CCGC	CACT	GAA-AGT	CTGGAT	GGGA	GAAG	GGAT	GAT		
BQ	AGCATCTTTC	GGGG-TCCGGT	TGAAAT	CCC	AACC	GGCCGT	19	AGTC	CCGT	GAC--	8	5	8	-----TGGATCTAGT	TGAAAT	CTAGGG	CCGC	CACT	AGT-AT	AGT	CTGGAT	GGGA	GAAG	GAT	ATG	
BE	TGCATGCTTC	GGGG-CAGGGT	TGAAAT	CCC	GACC	GGCCGT	20	AGCC	CCGC	GAC--	3	4	3	-----AGGATCCGT	TGAAAT	CTAGGG	CCGC	CACT	AGT-AT	AGT	CTGGAT	GGGA	GAAG	GAT	GGC	
HD	TTTATGCTTC	GGGG-CTGGGT	TGAAAT	CCC	GACC	GGCCGT	19	AGTC	CCGT	GAC--	10	4	10	-----TGGACTCGT	TGAAAT	CTAGGG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGGA	GAAG	GAT	GAAG	
Bam	TGTATCTTTC	GGGG-CTGGGT	TGAAAT	CCC	GACC	GGCCGT	23	AGCC	CCGT	GAC--	8	4	8	-----TGGATTCTAGT	TGAAAG	CTAGGG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGGA	GAAG	GAT	GAAG	
CA	GATGCTTTC	AGGG-ATGGGT	TGAAAT	CCC	AAT	CGCCGT	2	AGCC	CCGC	AA--	3	4	3	-----AGATCCGT	TAAACT	CTAGGG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG	
DF	CTTAACTTTC	GGGG-TAGGGT	TGAAAT	CCC	AAT	CGCCGT	2	AGCC	CCGC	G--	7	6	7	-----ATTTCGT	TAAAT	CTAAAG	CCGC	CACT	AGT-AA	AGT	CTGGAT	GGAA	GAAG	AAT	ATTT	
SA	TAAATCTTTC	GGGG-CAGGGT	TGAAAT	CCC	AACC	GGCCAGT	6	AGCT	CCGC	GAC--	11	3	11	-----CTGACTAGT	TGAGAT	CTAGAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGGA	GAAG	AAT	GT	
LLX	ATAAATCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	2	AGTC	CCGC	GA--	4	4	4	-----ATGATTCGT	TGAAAT	CTAGGG	CCGC	CACT	AGT-AT	AGT	CTGGAT	GGAA	GAAG	AAT	AAT	
PN	AACTATCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	2	AGCC	CCGC	GA--	3	4	3	-----ATGATTTCT	TGAAAT	CTAAAG	CCGC	CACT	AGT-AT	AGT	CTGGAT	GGAA	GAAG	AAT	AAAA	
TM	AAACCTCTTC	GGGG-CAGGGT	TGAAAT	CCC	GACC	GGCCGT	3	AGCC	CCGC	GAG--	5	4	5	-----TTGACTCCGT	TGAAAT	CTAGGG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGGA	GAAG	AGC	CTGA	
DR	GACCTCTTTC	GGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	15	AGCC	CCGC	GA--	8	12	9	-----CCGATCCG	CTGAACT	CTGGAG	CCGC	CACT	AGT-AT	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG	
TQ	CACCTCTTTC	GGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	3	AGCC	CCGC	GA--	5	4	5	-----CCGACCCG	CTGAAAT	CTGGAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGGA	GAAG	AGC	AGCC	
AO	AATAATCTTC	AGGG-CAGGGT	TGAAAT	CCC	GAT	CGCCGT	2	AGTC	CCGC	GA--	7	7	7	-----AGGAACCCGT	TGAAAT	CTGGAG	CCGC	CACT	AGT-AT	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG	
DU	TTTAACTTTC	AGGG-CAGGGT	TGAAAT	CCC	GAT	CGCCGT	2	AGTC	CCGC	GA--	13	4	12	-----AGGAACCT	AGT	CTGAAAT	CTAGT	ACCG	CACT	AGT-AT	AGT	CTGGAT	GGAA	GAAG	AGC	GAAG
CAU	GAAGACCTTC	GGGG-CAGGGT	TGAAAT	CCC	GAT	CGCCGT	20	AGCC	CCGC	GA--	3	4	3	-----AGGAACCCG	CTGAT	CTGGAG	CCGC	CACT	AGT-AT	AGT	CTGGAT	GGGA	GAAG	AGC	CTGGC	
EN	TAAAGCTTTC	AGGG-CAGGGT	TGAAAT	CCC	GACC	GGCTGAG	2	AGTC	CCGC	G--	5	4	5	-----GATTTCT	TGAAAT	CTAAAG	CCGC	CACT	AGT-AG	AGT	CTGGAT	GGGA	GAAG	AAT	GAAG	
TFU	ACGGCTCTTC	CGGG-GTCGGT	TGAAAT	CCC	GAA	CGCCGT	3	AGTC	CCGC	GAC--	8	5	8	-----TGGAACTCGT	TGAAAT	CTGGAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGGA	GAAG	AAT	CTGGC	
SX	-AGCCACTTC	CGGG-GTCGGT	TGAAAT	CCC	AAT	CGCCGT	3	AGTC	CCGC	GAC--	8	5	8	-----TTGACCACT	TGAAAT	CTGGAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGGA	GAAG	AGC	AGT	
BU	GTGCTCTTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	30	AGCC	CCGC	GAGCC	137	4	8	GTCAAGCAGAT	CTGGT	TGAAAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GTGGC	
BPS	GTGCTCTTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	21	AGCC	CCGC	GAGCC	8	4	8	GTCAAGCAGAT	CTGGT	TGAAAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GTGGC	
REU	TTACGTCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	31	AGCC	CCGC	GAGCC	7	5	7	GTCAAGCAGAT	CTGGT	TGAAAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GTGGC	
RSO	GTACATCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	21	AGCC	CCGC	GAGCC	11	3	11	GTCAAGCAGAT	CTGGT	TGAAAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GTGGC	
EC	GCTTATCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	17	AGCC	CCGC	GAGCC	8	4	8	GACACAGAT	CTGGT	TGAAAT	CTGGAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG
TY	GCTTATCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	67	AGCC	CCGC	GAGCC	8	3	8	GTCAAGCAGAT	CTGGT	TGAAAT	CTGGAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG
KP	GCTTATCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	20	AGCC	CCGC	GAGCC	8	4	8	GTCAAGCAGAT	CTGGT	TGAAAT	CTGGAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG
HI	TCCGATCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	2	AGCC	CCGC	GAGCC	26	9	30	GTCAAGCAGAT	CTGGT	TGAAAT	CTAAAG	CCGC	CACT	AGT-AA	AGT	CTGGAT	GGAA	GAAG	AAT	AAAA
VK	GCGCATCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	14	AGCC	CCGC	GAGCC	11	9	11	GTCAAGCAGAT	CTGGT	TGAAAT	CTAAAG	CCGC	CACT	AGT-AT	AGT	CTGGAT	GGAA	GAAG	AAT	AAGC
VC	CAATATCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	13	AGCC	CCGC	GAGCC	5	4	5	GTCAAGCAGAT	CTGGT	TGAAAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG	
YP	GCCTTATCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	40	AGCC	CCGC	GAGCC	16	6	16	GTCAAGCAGAT	CTGGT	TGAAAT	CTGGAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG
AB	GCGCATCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	25	AGCC	CCGC	GAGCC	16	4	27	GTCAAGCAGAT	CTGGT	TGAAAT	CTAAAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	AAAA
BP	GTACATCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	18	AGCC	CCGC	GAGCC	10	4	10	GTCAAGCAGAT	CTGGT	TGAAAT	CTGGAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GTGGC
AC	ACATGCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	16	AGCC	CCGC	GAGCA	10	3	11	---CGCAGAT	CTGGT	TGAAAT	CTAGAG	CCGC	CACT	AGT-AT	AGT	CTGGAT	GGAA	GAAG	AAT	GAGC
Spu	AACTATCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	34	AGCC	CCGC	GAGCC	6	6	6	GTCAAGCAGAT	CTGGT	TGAAAT	CTAAAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GTAA
PP	GTGCTCTTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	13	AGCC	CCGC	GAGCC	7	3	7	GTCAAGCAGAT	CTGGT	TGAAAT	CTAGAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GTAA
AU	GGTTGCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	17	AGCC	CCGC	GAGCC	7	9	7	GTCAAGCAGAT	CTGGT	TGAAAT	CTAGAG	CCGC	CACT	AGT-AT	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG
PU	AAACGTCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	19	AGCC	CCGC	GAGCC	19	4	18	GTCAAGCAGAT	CTGGT	TGAAAT	CTGGAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG
PY	TAACTCTTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	19	AGCC	CCGC	GAGCC	15	4	16	GTCAAGCAGAT	CTGGT	TGAAAT	CTGGAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG
PA	TAACTCTTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	19	AGCC	CCGC	GAGCC	14	4	13	GTCAAGCAGAT	CTGGT	TGAAAT	CTGGAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG
ML0	TAAAGTCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	16	AGCC	CCGC	GAGCC	8	5	8	GTCAAGCAGAT	CTGGT	TGAAAT	CTGGAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG
SM	AAAGCTCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	34	AGCC	CCGC	GAGCC	8	3	8	GTCAAGCAGAT	CTGGT	TGAAAT	CTGGAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG
BME	GCTTGTCTTC	GGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	17	AGCC	CCGC	GAGCC	10	15	10	GTCAAGCAGAT	CTGGT	TGAAAT	CTGGAG	CCGC	CACT	AGT-AGT	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG
BS	ATCAATCTTC	GGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	18	AGCC	CCGC	GA--	5	4	5	-----AGGATTCGT	TGAAAT	CTGGAG	CCGC	CACT	AGT-AC	AGT	CTGGAT	GGGA	GAAG	AAT	GAAG	
BQ	GTCTATCTTC	GGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	27	AGCC	CCGC	GA--	3	5	3	-----AGGATTCGT	TGAAAT	CTAAAG	CCGC	CACT	AGT-AT	AGT	CTGGAT	GGGA	GAAG	AAT	GAAG	
BE	ATTCATCTTC	GGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	20	AGCC	CCGC	GA--	3	4	3	-----AGGATTCGT	TGAAAT	CTGGAG	CCGC	CACT	AGT-AT	AGT	CTGGAT	GGGA	GAAG	AAT	GAAG	
CA	AATGATCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	2	AGCC	CCGC	GAG--	3	4	3	-----TATGATCCG	TGAAAT	CTGGAG	CCGC	CACT	AGT-AA	AGT	CTGGAT	GGAA	GAAG	AAT	ATAT	
DF	GAAGATCTTC	GGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	2	AGCC	CCGC	G--	6	4	6	-----GATTTCT	TGAAAT	CTAAAG	CCGC	CACT	AGT-AA	AGT	CTGGAT	GGGA	GAAG	AAT	ATTT	
EF	GTCTCTTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	3	AGTC	CCGC	GAC--	5	3	5	-----ATTGATTCGT	TGAAAT	CTAAAG	CCGC	CACT	AGT-AT	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG	
LLX	AAATATCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	21	AGTC	CCGC	GAT--	4	4	4	-----TTGAACTG	TGAAAT	CTGGAG	CCGC	CACT	AGT-AA	AGT	CTGGAT	GGAA	GAAG	AAT	GAAG	
LO	GTCTATCTTC	GGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	3	AGTC	CCGC	GAT--	3	10	3	-----TTGACTCTG	TGAAAT	CTGGAG	CCGC	CACT	AGT-AT	AGT	CTGGAT	GGGA	GAAG	AAT	GTGGC	
PN	AAAGATCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	125	AGTC	CCGC	GAT--	3	4	3	-----GATGATTCGT	TGAAAT	CTAGAG	CCGC	CACT	AGT-AT	AGT	CTGGAT	GGGA	GAAG	AAT	GAAG	
ST	AAATGCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	14	AGTC	CCGC	G--	3	4	3	-----GATGATTCGT	TGAAAT	CTAGAG	CCGC	CACT	AGT-AT	AGT	CTGGAT	GGGA	GAAG	AAT	GAAG	
MN	AAATGCTTC	AGGG-CAGGGT	TGAAAT	CCC	AAT	CGCCGT	104	AGTC	CCGC	G--	3	4	3	-----GATGATTCGT	TGAAAT	CTAGAG	CCGC	CACT	AGT-AA	AGT	CTGGAT	GGGA	GAAG	AAT		

Аттенюация транскрипции

Antiterminator

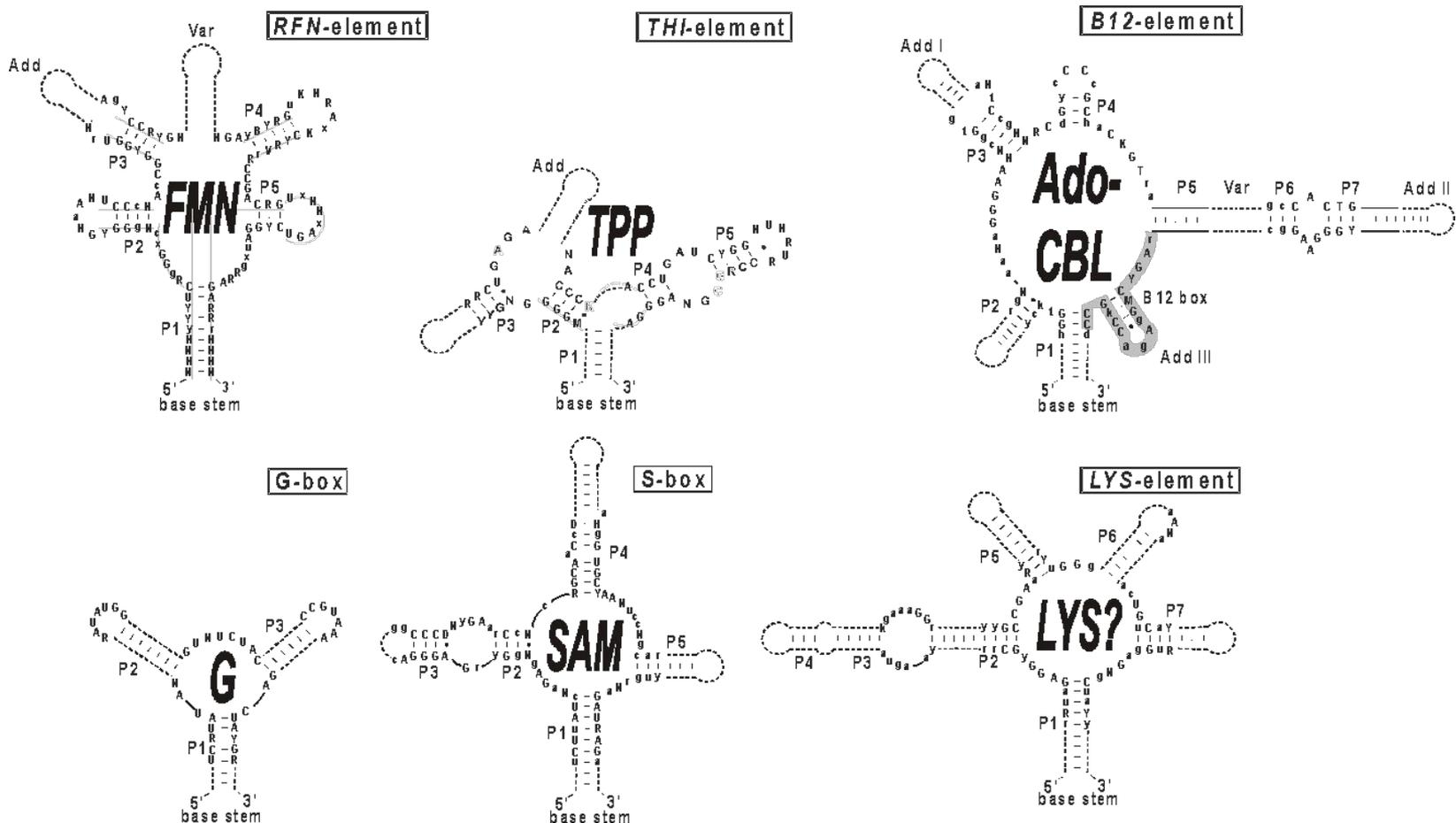
The *RFN* element

Terminator

Bam	GACAAAAAATATTTGATTGTA TCTTCGGGGCT TGGGTG	---	TCTGGATGGGGAAGGATGA	59	-----GTAA AGCCCGAA TGTGTAA--	ACATTCGGGGCTTTTTGACGCCAAAT
BS	GGACAAATGAATAAAGATTGTAT TTCGGGGC AGGGTG	---	TCTGGATGGGAGAAGGATGA	59	-----CTAA AGCCCGAA TTTTTTA--	TAAATTCGGGGCTTTTTGACGGTAAA
BQ	CTATAAATTTGAGCAAACAGC ATCCTTCGGGGT CGGGTG	---	TCTGGATGGGAGAAGGATAT	250	-----CCAA AGCCCAAGGAT ATTAAA--	ATCCTTGGGGTTTTTTGTTTTTTTT
BE	ACATAACGATATAGTGAT GCACTCCTTCGGGGC AGGGTG	---	TCTGGATGGGGAAGGATGC	155	-----TG AGCCCGGGG CAT-----	CCCGGGGTTTCATTTTTATTG
HD	AAATTTGAATAATTAATTTTTA TCCCTTCGGGGC TGGGTG	---	TCTGGATGGGAGAAGGAAAAC	148	-----AT CCCCCGT GAGAACAAAA--	TCTCTGGGGCTTTTTTGCGCGC
CA	TAATGGTAATTTAATAGGAT TCTTCAGGGAT TGGGTG	---	TCTGGATGAAAGAAGAAATA	34	----- AATCTC GAAGGAATTACC--	TTTCTTTGAGATTTTTTTATTG
DF	TAAATATAAATTTAA TACTTAA TCTTCGGGGT AGGGTG	---	TCTGGATGGAAGAAGATAAT	63	-----TAA ACCTGAG TTAATT-----	CTCAGGGTTTTTTGTTTTAAAAA
LLX	ACTTTAGCTACAA TTGAATAA TCTTCAGGGC CAGGGTG	---	TCTGGATGAAAGAAGATAAT	127	-----AAAAG ACCTGAA ATTTT-----	ATTTTAGGGTCTTATTTTTTATTAG
PN*	ATCATCTGTAA TTGAATAA TCTTCAGGGC AGGGTG	---	TCTGGATGAAAGAAGATAAA	81	-----TGTA GCCTTGAG TAGTCCCC--	TATTCAGGTATATTTTTTTGGAGG
PN*	ATCATCTGTAA TTGAATAA TCTTCAGGGC AGGGTG	---	TCTGGATGAAAGAAGATAAA	19	-----CG TGCTTGAATGA TACTTTG	TCATTTTCAGAGCATTTTTGTTAATC
TM	AAAAC TGAATACA AAAGAA CGCTCCTTCGGGGC AGGGTG	---	TCCGGATGGGAGAGAGCGTG	13	-----ATGGGA CCCGAGA -----	G GGTCCCTTTTCTTTTACA
AO	ATTTGCAACAATTTTTTAATA ATCTTCAGGGC AGGGTG	---	TCTGGATGGAAGAAGATGAA	33	-----TTTACA AGCCTTGAGAT CGAAAG--	ATTTCAAGGCTTTTTTCATCATT
DU	AATTTTTTTAACTAATTTTTA ATCTTCAGGGC AGGGTG	---	TCTGGATGGAAGAAGAAGAG	47	-----TGCATA AGCCTTGAGAT CTTAG--	GATTTCAAGGCTTTTTTCATAGTTA
FN	TAATCGAATA TGTA AAATAAAG TCTTCAGGGC AGGGTG	---	TCTGGATGGGAGAAGAATTA	18	-----ATAT TGCTC AGACTTT-----	GTTTGAGCATTTTTTTATTAA
SA	TAATAACAATTTCA TA TATAAT TCTTTCGGGGC AGGGTG	---	TCTGGAT GGGAGAAA GAAATG	74	----- TTTTCTCC TTGCACTTAATT--	GATGTGAGGATTTTTGTTTATA
DHA	ACTCTTTTTAGATGAATACGA ACTTCGAGGTA AGGGTG	---	TCCGGATGGGAGAAGGTACA	43	-----GTTTT ATGCCCTCGAGGA CACCAT	TTCCCTCGAGGCATTTTTGTTCTTTC
FN	GAAAAATAAATATTA AAAATA TCTTTCGGGGC AGGGTG	---	TCTGGATGAGAGAAGAAAAG	40	-----CTT ACCCGAAT TTCTAT-----	AATTCGGTTTTTTTTATTTT
CA	AAATAAAAAAATAAAGAATGAT TCTTCAGGGC AGGGTG	---	TCTGGATGAAAGAAGATA TA	19	-----TA TGCCCTGAC CTTTTT--	CGTTGGGGCTTTTTTAATGCT
DF	AAAATTA AAAAATCAAGAAG ATCTTCGGGGC AGGGTG	---	TCTGGATGAGAGAAGATAAT	45	-----ATAAAAA CTCGAAGAT AGGG--	TCTTCGAGTTTTTTGTTTTTCCTAA
BS	TAATTAATTTCA TA TGATCAAT TCTTTCGGGGC AGGGTG	---	TCTGGATGGGGAAGATGGA	103	--AAAGAAC CTTTCGGTTTTTC GA	GTAAGATGTGA TCGAAAAGG AGAGAATGA AGTG AAA
BQ	GGGAAAATAGAAATATCGG TCTATCTTCGGGGC AGGGTG	---	TCTGGAT GGGAGAAG ATGGA	54	----- ATTCCTCC TTTTGTTGTA--	-----ACACA AGGG TTTTTTTCGTTCT ATG
BE	ATAAAAAATGTAATAAGCGA TTTCACTCTCGGGC AGGGTG	---	TCTGGAT GGGAGAAG ATGAA	114	-----GGCAGC CTTCTCTT TGTA	GGATGAAT TCACGAGAAGGGGAGGAG AACAAGC ATG
PN	GTTTTTTGTTATGATAAAAGAG TCTTTCAGGGC AGGGTG	---	TCTGGATGGGAGAAACGAA	137	--AA CTTCTCTC TGATTTTAG-----	-----AAAAT TGGAG GAACCTGTT ATG ACA
ST	TAAATCTGCTATGCTAGAAG TGCTTTCAGGGC AGGGTG	---	TCTGGAT TGAGAGAAG ACCGG	130	--GGAA CTTCTTTT CAATTTGAAA--	-----AAAT TGGAG GAATTTTTTA ATG TC
MN	ATTTTTTGATA TGCTATAAG TGCTTTCAGGGC AGGGTG	---	TCTGGAT GGGAGAAG ACTGA	138	--GG CTTCTTTT CGATTTGTAA--	-----AAAT TGGAG GAATTTTTTT ATG AA
SA	AAATTTAATAATGTA AAA TTCACTTTCGGGGC AGGGTG	---	TCTGGAT GGGAG AAGATGGA	17	----- TCCTCTATTC TTACG-----	AGATGAATGGA AGGAG AAAAATGAAT ATG
EF	AAAAAATAATA CAAGGTT TCTTTCAGGGC AGGGTG	---	GTCTGGATAAAGAAGATAGG	33	---CTACTCTATTTTT TCCCTGCAG A-----	-----AAAA TGGGG TTTTTTTTG ATG A
LLX	TTTTTGTGCTATAATAAAAATA TCTTTCAGGGC ACCGTG	---	TCTGGATGGAAGAAGATGAA	66	--TCAACT TCTC GAAA TTGAAGAA T-TATT	TTCTCATATTT GGAG GTTTTTTTT ATG T
LO	ATTGTAAGAAAATATTCGTT CATCTTTCGGGGC AGGGTG	---	TCTGGAT GGGAGAAG ATGTTG	79	---ATGCACAAAC TCTCC TCAACTTTTTTTA-----	----- GTTGAGG TTTTTTATT TCG

Antiterminator

Другие РНК-переключатели, найденные методами сравнительной геномики



- Есть во всех трех основных царствах (бактерии, археи, эукариоты)
- Древнейшие регуляторные элементы: реликт «РНКового мира»?

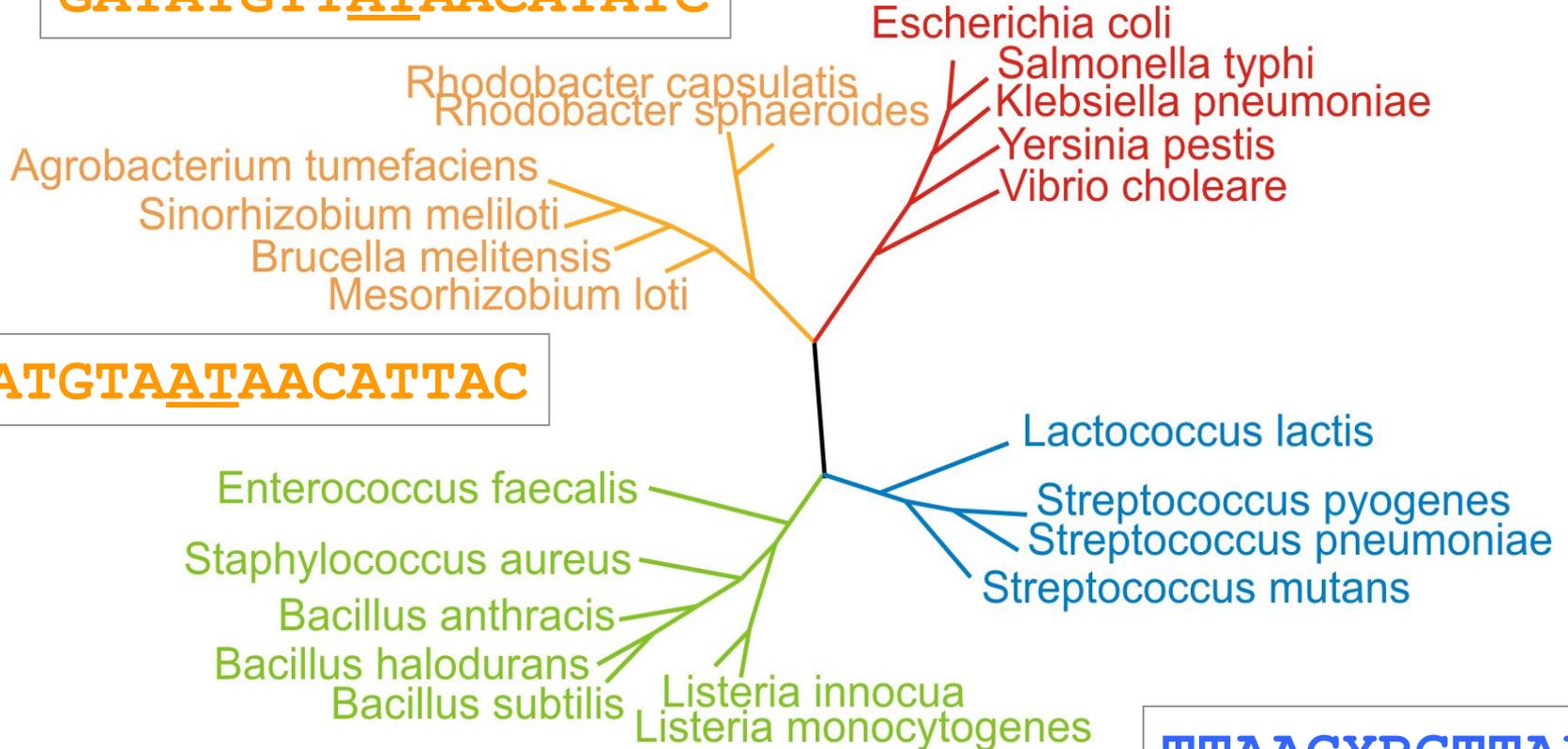
Регуляторы гомеостаза цинка

nZUR- α

nZUR- γ

GATATGTTATAACATATC

GAAATGTTATATAACATTC



GTAATGTAATAACATTAC

TTAACYRGTTAA

pZUR

TAAATCGTAATNATTACGATTTA

AdcR

Регуляция гомологов рибосомальных белков

	L36	L33	L31	S14
<i>E. coli, S.typhi</i>	(-)	-	(-) +	-
<i>K. pneumoniae</i>	(-)	-	(-) -	-
<i>Y. pestis, V. cholerae</i>	(-) x	-	(-) +	-
<i>B subtilis</i>	(-)	(-) + -	(-) +	(-) +
<i>S. aureus</i>	(-)	(-) - -	-	(-) +
<i>Listeria spp.</i>	(-)	(-) -	-	(-) +
<i>E. faecalis</i>	(-)	(-) x - -	-	(-) + -
<i>S. pne., S. mutans</i>	(-)	(-) - -	-	(-)
<i>S. pyo., L. lactis</i>	(-)	(-) - -	-	(-) +

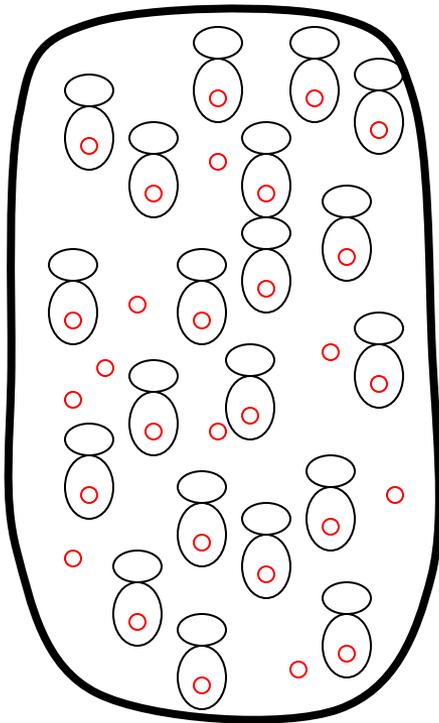
nZUR

pZUR

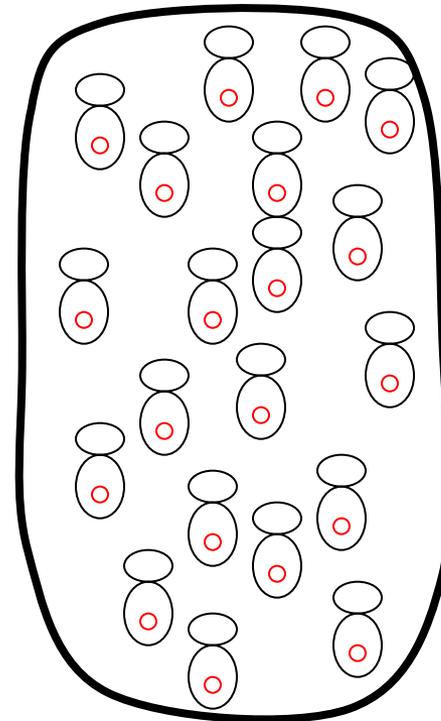
AdCR

Плохой сценарий

достаточно
цинка



недостаточно цинка: весь цинк
использован рибосомами, не
хватает цинка для ферментов

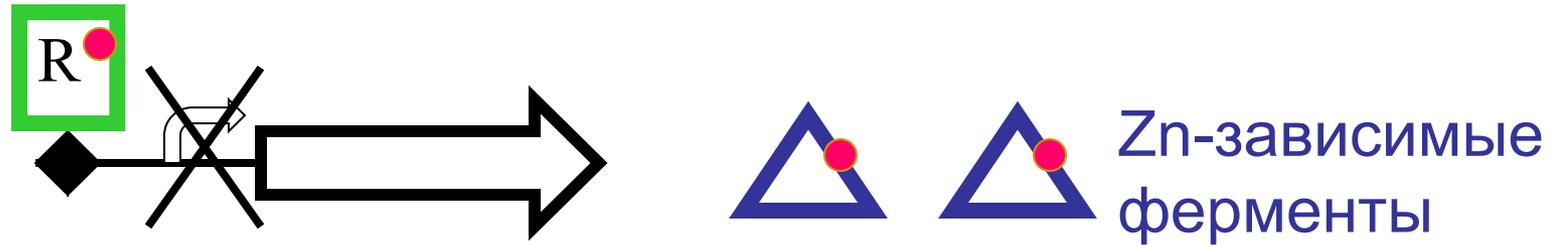


Регуляторный механизм

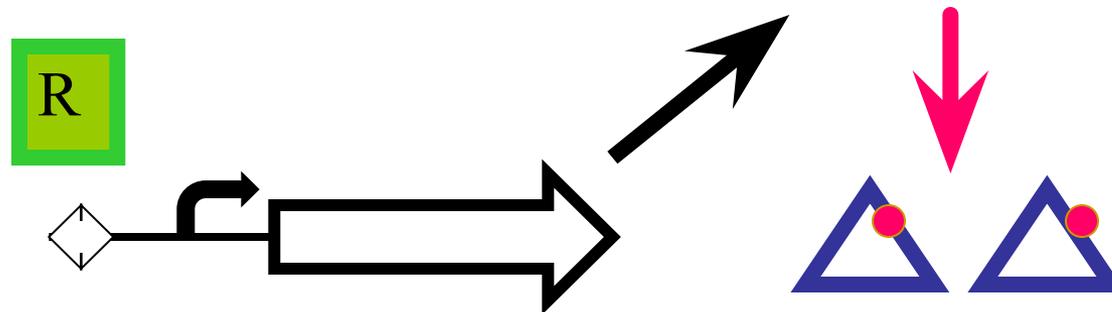
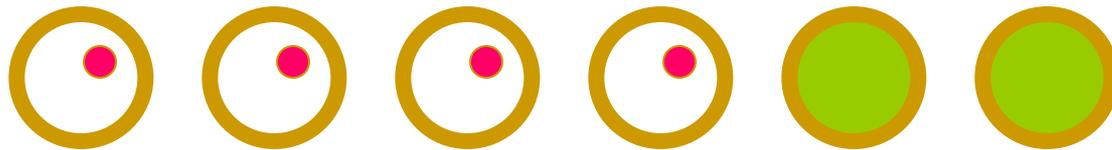
Достаточно **Zn**



репрессор



Голодание по **Zn**



Предсказание ...

(Proc Natl Acad Sci U S A. 2003 Aug 19;100(17):9912-7.)

Comparative genomics of bacterial zinc regulons: Enhanced ion transport, pathogenesis, and rearrangement of ribosomal proteins

Ekaterina M. Panina*[†], Andrey A. Mironov*[‡], and Mikhail S. Gelfand*^{‡§}

... и подтверждение

(Mol Microbiol. 2004 Apr;52(1):273-83.)

Zinc is a key factor in controlling alternation of two types of L31 protein in the *Bacillus subtilis* ribosome

Hideaki Nanamiya,^{1†} Genki Akanuma,^{1†}
Yousuke Natori,¹ Rikinori Murayama,¹ Saori Kosono,²
Toshiaki Kudo,² Kazuo Kobayashi,³
Naotake Ogasawara,³ Seung-Moon Park,⁴ Kozo Ochi⁴
and Fujio Kawamura^{1*}

(+ еще пять статей в последующие годы):
другие гены, другие бактерии

Сводка подтвердившихся предсказаний

- Регуляторы
 - РНК-переключатели
 - витамины: рибофлавин, тиамин, кобаламин
 - аминокислоты: лизин, метионин
 - Факторы транскрипции
 - NrdR: рибонуклеотид-редуктазы
 - MtaR, CmbR: метионин и цистеин
 - NiaR, NrtR: метаболизм NAD
 - NsrR, NnrA: нитрозативный стресс
- Регуляторные взаимодействия
 - регуляторные мотивы в ДНК (>10)
 - отдельные сайты связывания (>20)
- Ферменты
 - ThiN и TenA (биосинтез тиамина)
 - CobX, CobZ (биосинтез кобаламина)
 - FadE (синтез жирных кислот)
 - AbnA, Xca (катаболизм арабинозы)
 - NagK, NagBII (катаболизм N-ацетилглюкозамина)
- Транспортеры
 - витамины и кофакторы
 - YpaA и RibM: рибофлавин
 - BioMNY: биотин
 - ThiXYZ: тиамин
 - NiaP: ниацин
 - Vng1369-71: корриноиды
 - сахара и полисахариды
 - OgtABCD: продукты деградации пектина
 - NagP: N-ацетилглюкозамин
 - аминокислоты
 - MetD: метионин
 - SteT: треонин
 - ионы металлов
 - CbiMNQO, HoxN: кобальт
 - NikMNQO: никель
 - нуклеотиды:
 - YicE: ксантин

Не только тексты

Другие типы массовых экспериментов:

- Транскриптомика
 - «выстилающие массивы»: полная карта транскриптов
 - уровень экспрессии и время жизни мРНК
 - ДНК-белковые взаимодействия
- Протеомика
 - концентрации белков
 - белок-белковые взаимодействия, белковые комплексы
 - структуры белков
- Эпигенетика
 - метилирование ДНК
 - положение и модификации нуклеосом
- Генетика
 - летальность мутаций
 - фенотипы
 - синтетические летали

«Неприкладная» биоинформатика

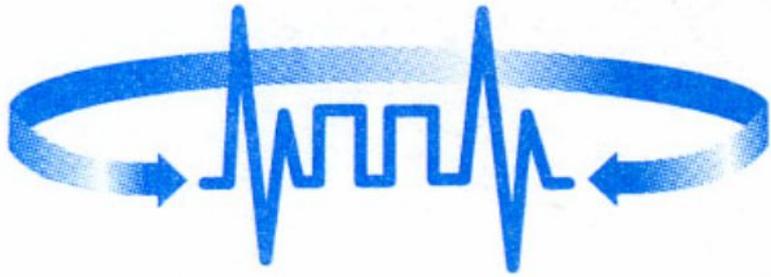
- Молекулярная эволюция
 - филогения генов
 - таксономия организмов
 - горизонтальные переносы и т.п.
 - положительный и отрицательный отбор
 - что сделало нас людьми?
 - лекарственная устойчивость
 - эволюция геномов
- Системная биология
 - строение геномов
 - сети взаимодействий
 - белок-белковые
 - регуляция транскрипции
 - сигнальные пути

Перспективы

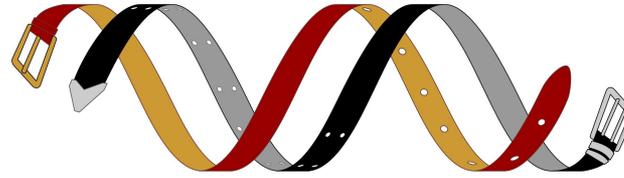
- **Индивидуальные геномы**
 - персональные человеческие геномы
 - геномы штаммов
- **Метагеномы**
 - некультивируемые бактерии
 - метагеномы экологических ниш
 - бактериальный метагеном человека
- **Другие виды данных**
 - интеграция
- **Моделирование (пока рано)**
 - построение полных карт

Биоинформатика в России

- Сильные устойчивые школы
 - Москва+Пущино: содружество лабораторий в разных учреждениях
 - Новосибирск: вертикальная структура
- Высокий уровень
 - конкурентоспособный в мире
- Интеграция с образованием
 - Факультет биоинженерии и биоинформатики МГУ
 - Кафедра информационной биологии ФЕН НГУ
- Перспективная область:
 - относительно дешево
 - общедоступная информация - можно использовать чужие данные
 - связь с экспериментальными исследованиями, ведущимися на современном уровне
- Как развивать:
 - технические разработки (базы данных, пакеты программ) - контракты, наличие потребителя
 - интегрированные компоненты в медицинских и молекулярно-биологических проектах (крупные проекты должны иметь биоинформатическую поддержку)
 - самостоятельные исследовательские работы - гранты



Учебно-Научный Центр



Биоинформатика



- Российский фонд фундаментальных исследований
- РАН, программа «Молекулярная и клеточная биология»
- INTAS
- Howard Hughes Medical Institute