



Сложность генных сетей прокариот и эукариот.

Валентин Владимирович Суслов, м.н.с. ЛТГ ИЦиГ

Кафедра информационной биологии ФЕН НГУ



Сравнительная характеристика геномов про- и эукариот

Таксон ^α	Вид ^α	Гаплоидный геном·млн п.н. ^α	Число генов·в геноме ^α
ПРОКАРИОТЫ ^α			
МИКОПЛАЗМЫ ^α	<i>Mycoplasma genitalium</i> ^α	0,58 ^α	470 ^α
	<i>Mycoplasma pneumoniae</i> ^α	0,82 ^α	~670 ^α
РИККЕТСИИ ^α	<i>Rickettsia prowazekii</i> ^α	1,1 ^α	834 ^α
АРХЕОБАКТЕРИИ ^α	<i>Archaeoglobus fulgidus</i> ^α	2,18 ^α	2436 ^α
	<i>Methanopyrus kandleri</i> ^α	1,69 ^α	1738 ^α
ЦИАНОБАКТЕРИИ ^α	<i>Synechocystis</i> sp. ^α	3,57 ^α	3168 ^α
ЭУБАКТЕРИИ ^α	<i>Escherichia coli</i> ^α	4,6—5,5 ^α	4288 ^α
	<i>Campylobacter jejuni</i> ^α	1,64 ^α	1654 ^α
	<i>Aquifex aeolicus</i> ^α	1,55 ^α	1512 ^α
	<i>Neisseria meningitidis</i> ^α	2,27 ^α	2121 ^α
	<i>Bacillus subtilis</i> ^α	4,2 ^α	4100 ^α
НИЗШИЕ ЭУКАРИОТЫ ^α			
ГРИБЫ ^α	<i>Saccharomyces cerevisiae</i> ^α	11,4 ^α	6241 ^α
	<i>Schizosaccharomyces pombe</i> ^α	13,8 ^α	4824 ^α
	<i>Aspergillus nidulans</i> ^α	31 ^α	α
ПРОТИСТЫ ^α	<i>Amoeba dubia</i> ^α	670000 ^α	α
	<i>Entamoeba histolytica</i> ^α	20 ^α	α
	<i>Dictyostelium discoideum</i> ^α	32 ^α	11000 ^α
ВЫСШИЕ ЭУКАРИОТЫ ^α			
ВЫСШИЕ РАСТЕНИЯ ^α	<i>Lilium longiflorum</i> ^α	90000 ^α	α
	<i>Arabidopsis thaliana</i> ^α	115,7 ^α	27540 ^α
	<i>Oryza sativa</i> ^α	466 ^α	46022·-·55615 ^α
ПЕРВИЧНОРОТЫЕ ^α	<i>Caenorhabditis elegans</i> ^α	97 ^α	19049 ^α
	<i>Drosophila melanogaster</i> ^α	120 ^α	13600 ^α
ВТОРИЧНОРОТЫЕ ^α	<i>Protopterus aethiopicus</i> ^α	139000 ^α	α
	<i>Fugu rubripes</i> ^α	365-400 ^α	30-40·тыс. ^α
	<i>Homo sapiens</i> ^α	3000 ^α	30000 ^α
	<i>Mus musculus</i> ^α	2500 ^α	37000 ^α

1 протист, имеющий сложный жизненный цикл и многоклеточную стадию



Выводы из геномных проектов



Расшифровка геномов выявила

- 1) сложность прокариот в целом коррелирует с размерами геномов и числом генов;
- 2) наблюдаются рост размера геномов и числа генов при переходе от прокариот к эукариотам и от одноклеточных к многоклеточным;
- 3) у эукариот отсутствует связь между биологической сложностью, размерами геномов и числом генов .



Характеристики сложности генных сетей и других биологических объектов (клеток, организмов, экосистем)

Можно выделить ряд инвариантных характеристик, связанных с ростом сложности биологических объектов :

- 1) увеличение количества элементов;
- 2) увеличение числа связей между элементами;
- 3) увеличение числа уровней иерархии
- 4) увеличение количества элементов и связей, работающих в единицу времени и/или в единице пространства;
- 5) рост разнообразия режимов динамического поведения.



Структура генной сети



В любой генной сети можно выделить

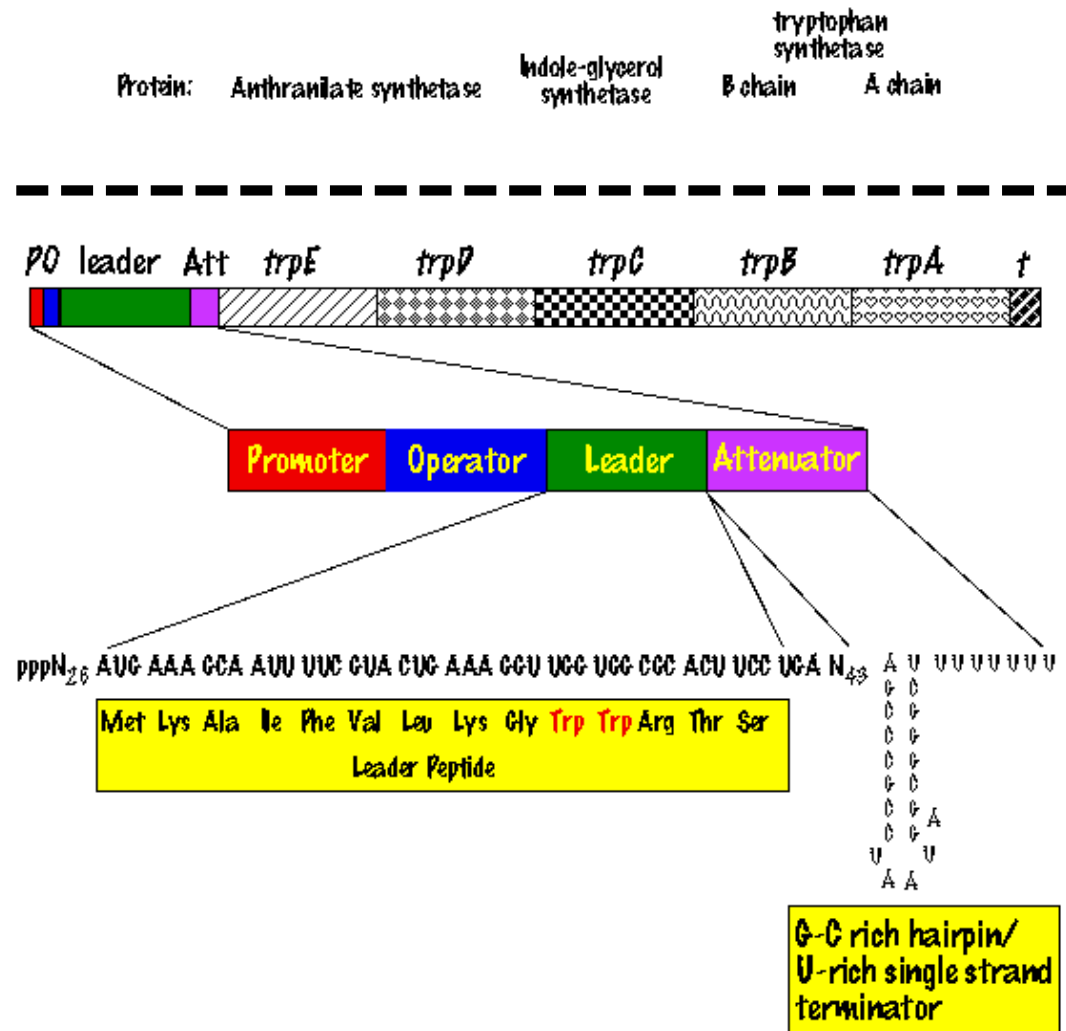
- 1) ядро сети – собственно группу последовательностей ДНК (РНК), ответственных за формирование признака;
- 2) пути передачи сигналов от рецепторов к регуляторным районам элементов ядра сети;
- 3) центральные регуляторы, регулирующие сразу несколько элементов ядра сети;
- 4) набор положительных или отрицательных обратных связей, обеспечивающих функционирование сети в определенном режиме, или наоборот, запрограммированное отклонение от этого режима.



Элемент ядра сети прокариот



Триптофановый оперон *E.coli*



Оперон - группа генов с общими регуляторными районами, соответственно транскрибирующаяся как один цистрон.

Преимущества:

- 1) быстрота и четкость регуляции – белковые продукты сразу синтезируются в стехиометрических количествах;
- 2) экономия на размере ДНК.

Недостатки:

- 1) затруднения в тонкой регуляции каждого гена;
- 2) чувствительность к повреждениям (цис-мутации).



Элемент ядра сети эукариот



Кассеты генов клеточного цикла:

- а) репрессия транскрипционным фактором E2F1/DP1/pRB,**
- б) активация транскрипционным фактором E2F1/DP1.**

Кассета генов: группа функционально связанных генов. Каждый имеет свои регуляторные районы, в которых есть одинаковые сайты связывания, что и обеспечивает активацию всех генов кассеты при появлении определенного транскрипционного фактора. Каждый ген кассеты транскрибируется как отдельный цистрон.

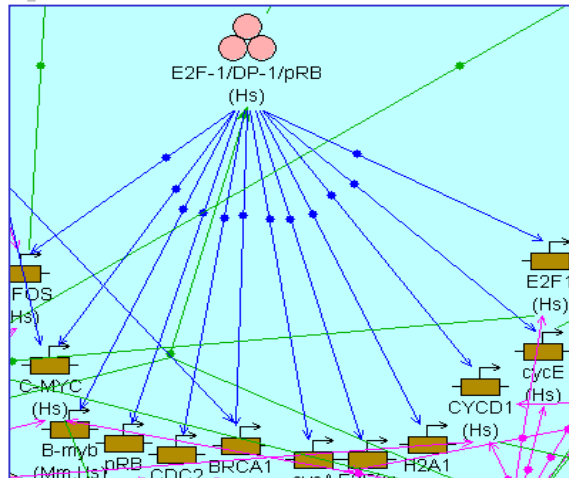
Преимущества:

- 1) тонкая регуляция каждого гена;**
- 2) помехоустойчивость;**
- 3) любой ген может участвовать в разные моменты времени в разных кассетах.**

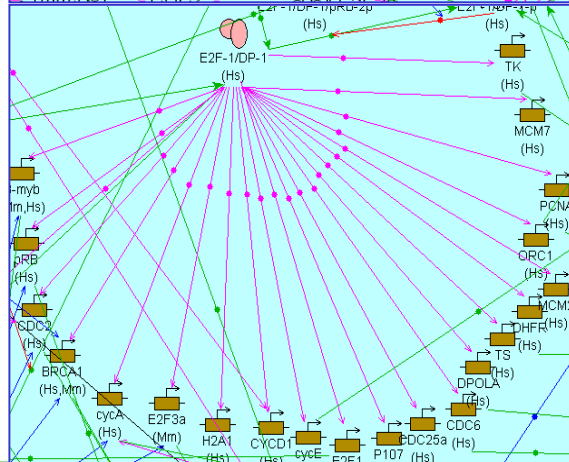
Недостатки:

- 1) отсутствие стехиометрии, требуется дополнительная подгонка;**
- 2) сложность регуляции;**
- 3) занимают много места, требуют большого генома.**

а)



б)





Центральные регуляторы сети прокариот



Сигма-факторы

субъединицы белкового комплекса РНК-полимеразы, обеспечивающие специфичное связывание комплекса с сайтом связывания.

С началом транскрипции диссоциируют из комплекса.

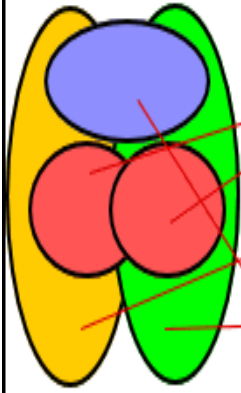
Преимущества:

- 1) быстрота регуляции;
- 2) быстрое рециклирование, вследствие чего можно обойтись небольшим числом молекул.

Недостатки:

- 1) ограничения на размер белка, а следовательно и на размер сайта связывания ДНК;
- 2) невозможность взаимодействия нескольких сигма-факторов в одном комплексе;
- 3) временные ограничения на связь с ДНК.

**Prokaryotic RNA Polymerase:
Holoenzyme Enzyme**



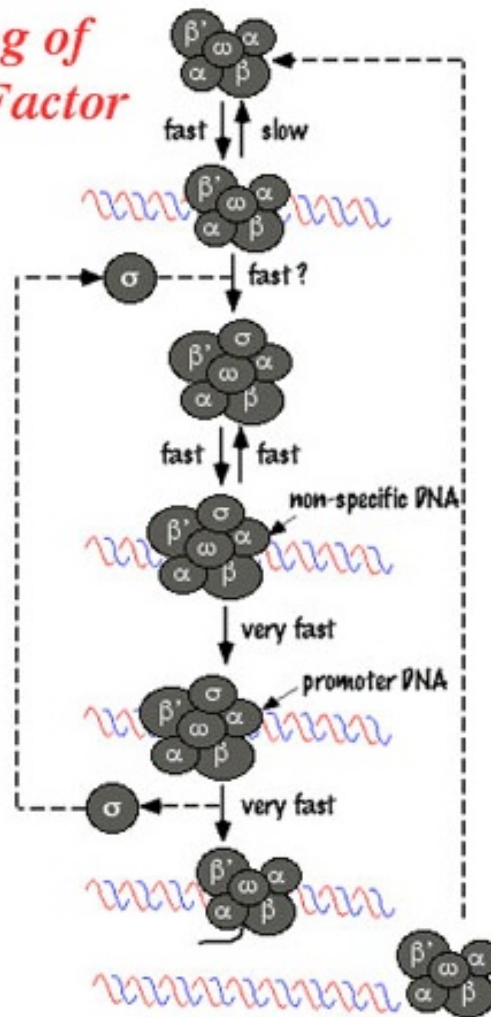
<u>Subunit</u>	<u>Size</u>	<u>#/Molecule</u>	<u>Function</u>
α	36.5 kD	2	chain initiation and interaction with regulatory proteins
β	151 kD	1	chain initiation and elongation
β'	155 kD	1	DNA binding
σ	70 kD	1	promoter recognition



Центральные регуляторы сети прокариот



Cycling of Sigma Factor



Core enzyme stored on DNA

Sigma factor associates with core enzyme

Holoenzyme has reduced affinity for loose sites

Holoenzyme has increased affinity for promoters

Sigma is released

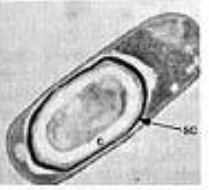
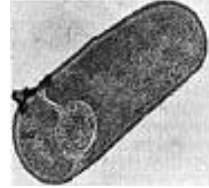
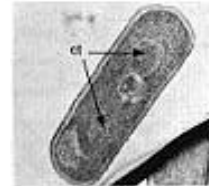
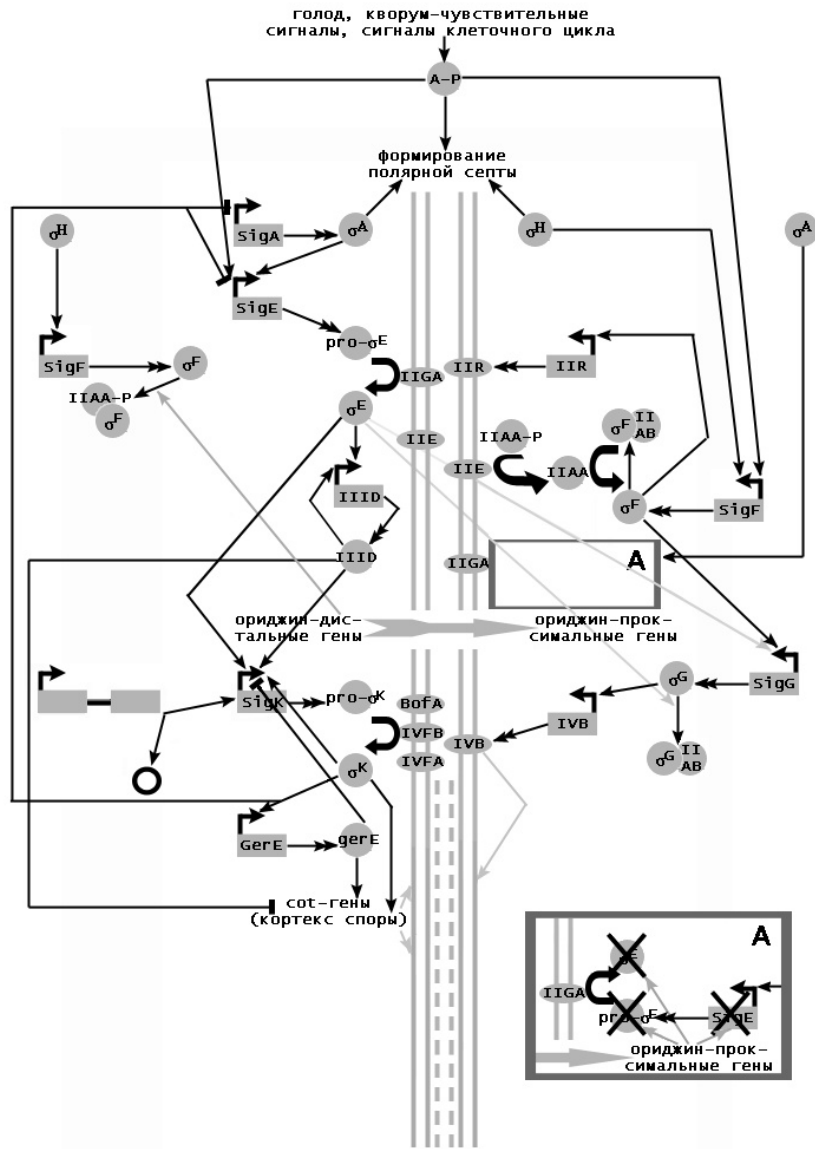
Core enzyme synthesizes RNA

Core enzyme terminates and is released

DIJ



Генная сеть формирования споры *Bacillus subtilis*



10)

Слева материнская клетка, справа – проспора. До споруляции в клетке присутствуют σ_A и σ_H – факторы. Голод, кворум-чувствительные сигналы, сигналы клеточного цикла активируют белок SpoA-P, запускающий сборку полярной септы и активирующий экспрессию σ_A и σ_H – факторов, которые активируют экспрессию σ_E и σ_F – факторов соответственно. σ_F связывается в цитоплазме антифактором SpoIIAB, от которого его освобождает белок SpoIIAA. Септальный SpoIIIE переводит фосфорилированный SpoIIAA-P в SpoIIAA. SpoIIIE локализован на обеих частях полярной септы, но по-видимому из-за того, что материнская клетка много больше проспоры высвобождение σ_F преимущественно идет в проспоре. σ_E синтезируется в виде профактора, расщепляемого септальной протеазой SpoIIIGA. SpoIIIGA материнской клетки активируется септальным SpoIIIR проспоры, синтез которого активируется σ_F проспоры. Тем временем ген SpoIIIE участвует в праспределении копий поделившейся хромосомы между проспорой и материнской клеткой. Проникшие первыми в проспору ориджин-проксимальные гены ингибируют там σ_E , тогда как двойная порция ориджин-дистальных генов (хромосомы материнской клетки и ориджин-дистальной части хромосомы, перемещающейся в проспору) блокируют σ_F в материнской клетке. Далее σ_E в материнской клетке активирует синтез SpoIIID и σ_K , а в проспоре неизученным пока образом – активирует синтез σ_G или его высвобождение из комплекса с анти-фактором SpoIIAB.



Центральные регуляторы сети про- и эукариот

Транскрипционные факторы

белковые молекулы, связывающиеся с сайтами ДНК и взаимодействующие с комплексом РНК-полимеразы. Взаимодействие с сайтами ДНК и с комплексом РНК-полимеразы могут происходить независимо друг от друга.

Преимущества:

- 1) тонкая регуляция транскрипции;
- 2) возможность образования сложных мультимерных комплексов на нескольких сайтах (со всеми их преимуществами - кооперативное связывание, синэргизм и т.д.);
- 3) время существования может не зависеть от времени существования комплекса РНК-полимеразы;
- 4) нет ограничений на размер белка.

Недостатки:

- 1) медленное рециклирование.



Прокариоты (*Bacillus subtilis*)



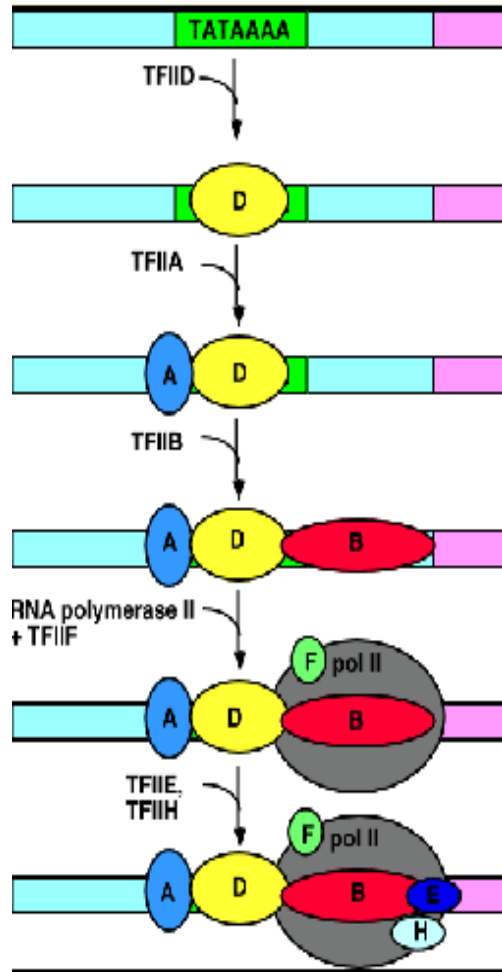
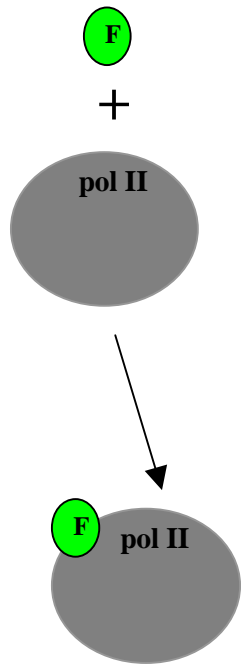


Регуляция транскрипции эукариот – сборка комплекса РНК-полимеразы II



TFIIF –
аналог сигма-
факторов
прокариот

Транскрипци-
онные факторы



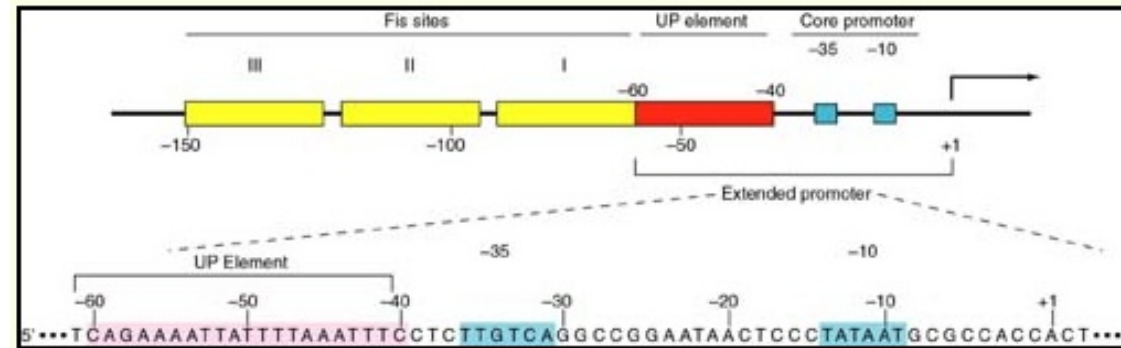
Factor (Abbreviation)	Composition	Function
TFIIA (IIA)	2 or 3 subunits	stabilizes binding between TFIID and promoters
TFIIB (IIB)	single subunit	interaction between TFIID and polII-TFIIF
TFIID (IID)	TBP (TATA box-binding protein) 8 - 10 TAF _{II} 's (TBP-associated proteins)	binding to TATA box interaction with promoter elements and with gene-specific transcription factors
TFIIF (IIF)	2 subunits	somewhat like sigma in prokaryotes, this protein causes RNA pol II to bind to the complex assembly at the promoter
TFIIE (IIE)	2 subunits	required for binding and stimulation of transcription
TFIIH (IIH)	complex	kinase activity (associated kinase activates polII by phosphorylation), helicase activity



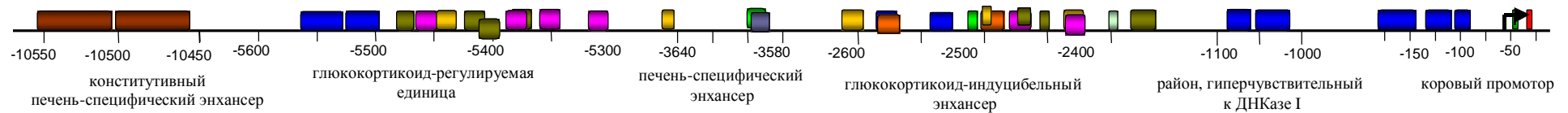
Регуляторные районы про- и эукариот



Ген рибосомальной РНК *rrnB* *E.coli*



5'-регуляторный район гена тирозинаминотрансферазы крысы



В то время как у прокариот характерный размер регуляторного района гена составляет ~60–100 п.о. У эукариот он намного больше и может достигать десятков тысяч п.о. и содержать десятки сайтов связывания транскрипционных факторов.



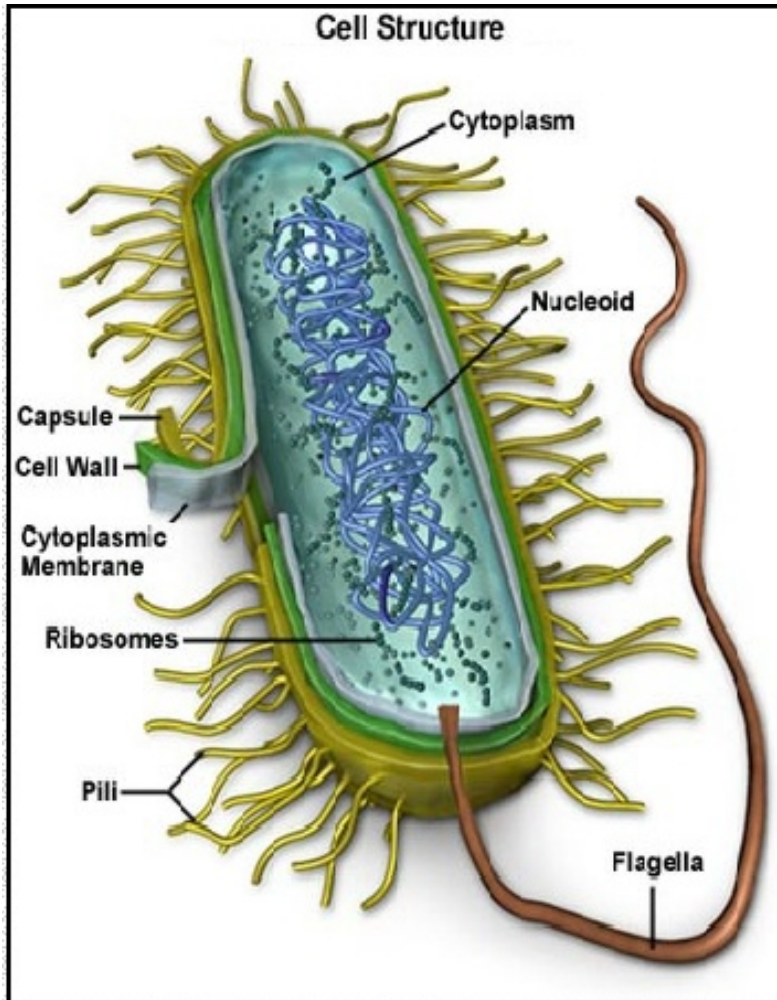
Длина регуляторного района и сложность регуляции транскрипции



Рассмотрим регуляторный район, содержащий J сайтов связывания транскрипционных факторов. Предполагая, что каждый сайт может находиться в двух состояниях: (1) свободном и (2) связанном с соответствующим транскрипционным фактором, можно оценить W - полное количество состояний регуляторного района - как 2^J . При $J=20$, $W \sim 1000$. Иными словами, даже в простейшем варианте с двумя состояниями каждого сайта емкость такого кода регуляции транскрипции весьма велика. Значит, даже при небольшом числе транскрипционных факторов, но с множеством сайтов связывания в регуляторном районе гена, его кодирующая емкость будет огромна.



Особенности клетки прокариот



- 1) гаплоидность,
- 2) пассивный механизм сегрегации хромосом, ассоциированный с мембраной,
- 3) совмещение транскрипции и трансляции во времени и пространстве,
- 4) малые размеры клеток,
- 5) узкоспециализированный тип питания – пиноцитоз, связанный с наличием жесткого экзоскелета,
- 6) отсутствие компартментов и, как следствие, невозможность разделения метаболических процессов в пространстве,
- 7) отсутствие активного внутриклеточного транспорта, роль которого играет диффузия.



Катастрофа мутационных ошибок Эйгена



$$V_m < V_{\max} = \frac{\ln S_m}{1 - q}$$

Любая популяция генетических самовоспроизводящихся систем имеет верхнюю границу темпов мутирования. Гаплоидные популяции достигают ее, когда за один цикл репликации возникает одна летальная мутация на геном [Eigen, 1971]. Отсюда, чем выше средняя вероятность мутирования ($1 - q$), тем ниже предел размера генома. Значит, усложнение генетических самовоспроизводящихся систем, требующее роста генома, невозможно без роста надежности хранения и копирования генетической информации. В ходе эволюции бактерии вплотную подошли к этой границе мутационной катастрофы ошибок, что запретило рост их гаплоидных геномов. Преодоление этого запрета стало главной эволюционной мотивацией усложнения организмов.



Минимизация – тренд эволюции некодирующей ДНК в геномах прокариот



Table 1. Various statistics of prokaryotic genomes

Species	Genome size (kb)	No. of genes	Gene density (per 1 kb)	Fraction of non-coding DNA
<i>Aeropyrum pernix</i>	1670	1688	1.01	0.14
<i>Sulfolobus solfataricus</i>	2592	3012	1.16	0.15
<i>Sulfolobus tokodaii</i>	2695	2956	1.10	0.15
<i>Methanococcus jannaschii</i>	1665	1828	1.10	0.12
<i>Methanobacterium thermoautotrophicum</i>	1751	1917	1.09	0.09
<i>Pyrococcus horikoshii</i>	1739	1796	1.03	0.09
<i>Pyrococcus abyssi</i>	1765	1802	1.02	0.08
<i>Archaeoglobus fulgidus</i>	2178	2467	1.13	0.07
<i>Thermoplasma acidophilum</i>	1565	1528	0.98	0.12
<i>Thermoplasma volcanium</i>	1585	1548	0.98	0.14
<i>Halobacterium</i> sp. ^a	2570	2640	1.03	0.14
<i>Escherichia coli</i> K12	4639	4375	0.94	0.12
<i>Buchnera</i> sp.	641	610	0.95	0.12
<i>Salmonella typhi</i>	4809	4696	0.98	0.13
<i>Vibrio cholerae</i> ^b	4033	3949	0.98	0.13
<i>Yersinia pestis</i>	4654	4096	0.88	0.19
<i>Haemophilus influenzae</i>	1830	1746	0.96	0.11
<i>Pasteurella multocida</i>	2258	2064	0.91	0.10
<i>Pseudomonas aeruginosa</i>	6264	5642	0.90	0.10
<i>Xylella fastidiosa</i>	2679	2886	1.08	0.16
<i>Neisseria meningitidis</i> MC58	2272	2150	0.95	0.20
<i>Caulobacter crescentus</i>	4017	3782	0.94	0.09
<i>Mesorhizobium loti</i> ^a	7596	7296	0.96	0.14
<i>Sinorhizobium meliloti</i> ^a	6691	6258	0.94	0.13
<i>Agrobacterium tumefaciens</i> ^a	5674	5357	0.94	0.10
<i>Rickettsia prowazekii</i>	1112	920	0.83	0.24
<i>Ureaplasma urealyticum</i>	752	650	0.86	0.08
<i>Mycobacterium tuberculosis</i>	4404	3961	0.90	0.09
<i>Mycobacterium leprae</i>	3288	1653	0.50	0.50
<i>Synechocystis</i> PCC6803	3574	3203	0.90	0.12
<i>Deinococcus radiodurans</i> ^a	3242	3229	1.00	0.10

Паразитические прокариоты с аномально высоким содержанием некодирующей ДНК



Количество регуляторов связано с числом генов аллометрической зависимостью (Крофт)



Организм

Гены

<i>Mycoplasma genitalium</i>	480	2
<i>Buchnera aphidicola</i> Sg	545	7
<i>Buchnera</i> sp.	574	9
<i>Ureaplasma urealyticum</i>	611	3
<i>Wigglesworthia brevipalpis</i>	611	12
<i>Mycoplasma pneumoniae</i>	688	2
<i>Mycoplasma pulmonis</i>	782	4
<i>Tropheryma whippelii</i> test	808	6
<i>Rickettsia prowazekii</i>	834	12
<i>Chlamydia trachomatis</i>	894	10
<i>Chlamydia muridarum</i>	916	10
<i>Treponema pallidum</i>	1031	14
<i>Mycoplasma penetrans</i>	1037	16
<i>Chlamydia pneumoniae</i> AR39	1110	10
<i>Rickettsia conorii</i>	1374	15
<i>Thermoplasma acidophilum</i>	1478	28
<i>Helicobacter pylori</i> J99	1491	9
<i>Thermoplasma volcanium</i>	1526	26
<i>Aquifex aeolicus</i>	1553	39
<i>Mycobacterium leprae</i>	1605	43
<i>Campylobacter jejuni</i>	1634	27
<i>Borrelia burgdorferi</i>	1637	12
<i>Methanopyrus kandleri</i>	1687	15
<i>Streptococcus pyogenes</i>	1696	78
<i>Haemophilus influenzae</i>	1709	63
<i>Bifidobacterium longum</i>	1729	78
<i>Pyrococcus abyssi</i>	1765	40
<i>Methanococcus jannaschii</i>	1770	28
<i>Thermotoga maritima</i>	1846	59
<i>Methanobacterium thermoautotrophicum</i>	1869	45
<i>Streptococcus mutans</i>	1960	111
<i>Pasteurella multocida</i>	2014	68
<i>Neisseria meningitidis</i> Mc58	2025	40
<i>Streptococcus pneumoniae</i> R6	2043	80
<i>Pyrococcus horikoshii</i>	2064	33
<i>Pyrococcus furiosus</i>	2065	39
<i>Pisobacterium nucleatum</i>	2068	53
<i>Streptococcus agalactiae</i> 2603	2124	102
<i>Chlorobium tepidum</i> TLS	2252	34
<i>Lactococcus lactis</i>	2266	106
<i>Clostridium tetani</i> E88	2373	95

Организм

Гены Регуляторы

<i>Archaeoglobus fulgidus</i>	2407	85
<i>Staphylococcus epidermidis</i> ATCC 12228	2419	72
<i>Thermosynechococcus elongatus</i>	2475	60
<i>Thermotamaerobacter tengcongensis</i>	2588	115
<i>Halobacterium</i> sp.	2605	69
<i>Pyrobaculum aerophilum</i>	2605	33
<i>Staphylococcus aureus</i> MW2	2632	98
<i>Aeropyrum pernix</i>	2694	19
<i>Clostridium perfringens</i>	2723	121
<i>Sulfolobus tokodaii</i>	2826	68
<i>Xylella fastidiosa</i>	2831	71
<i>Listeria monocytogenes</i>	2846	177
<i>Corynebacterium efficiens</i> TS-314	2950	121
<i>Sulfolobus solfataricus</i>	2977	67
<i>Lactobacillus plantarum</i>	3009	183
<i>Listeria innocua</i>	3043	175
<i>Deinococcus radiodurans</i>	3102	99
<i>Synechocystis</i> PCC6803	3169	111
<i>Brucella melitensis</i>	3198	169
<i>Brucella suis</i> 1330	3264	164
<i>Methanocaldococcus jensenii</i>	3371	76
<i>Oceanobacillus ihemyensis</i>	3496	183
<i>Caulobacter crescentus</i>	3737	237
<i>Vibrio cholerae</i>	3828	223
<i>Clostridium acetobutylicum</i>	3848	221
<i>Bacillus halodurans</i>	4066	252
<i>Yersinia pestis</i> KIM	4090	196
<i>Bacillus subtilis</i>	4100	247
<i>Shigella flexneri</i> 2a	4180	216
<i>Xanthomonas campestris</i>	4181	225
<i>Mycobacterium tuberculosis</i> CDC1551	4187	160
<i>Escherichia coli</i> K12	4289	275
<i>Xanthomonas citri</i>	4312	233
<i>Vibrio vulnificus</i> CMCP6	4537	308
<i>Methanocaldococcus acetivorans</i>	4540	109
<i>Salmonella typhimurium</i> LT2	4553	304
<i>Leptospira interrogans</i>	4727	117
<i>Salmonella typhi</i>	4767	279
<i>Shewanella oneidensis</i>	4778	233
<i>Ralstonia solanacearum</i>	5116	345
<i>Agrobacterium tumefaciens</i> C58	5301	392
<i>Pseudomonas putida</i> KT2440	5350	353

$$c = vp$$

$$R = \sum_{n=0}^N cn = \frac{cN(N+1)}{2} \approx \frac{c}{2} N^2$$

R – гены-регуляторы

N – общее число регулируемых единиц (генов или оперонов)

v – фракция генов с уникальными взаимодействиями

p – селективное преимущество



Оперонная структура генома и ограничения на его размер препятствуют аллометрическому росту числа регуляторов (Крофт).

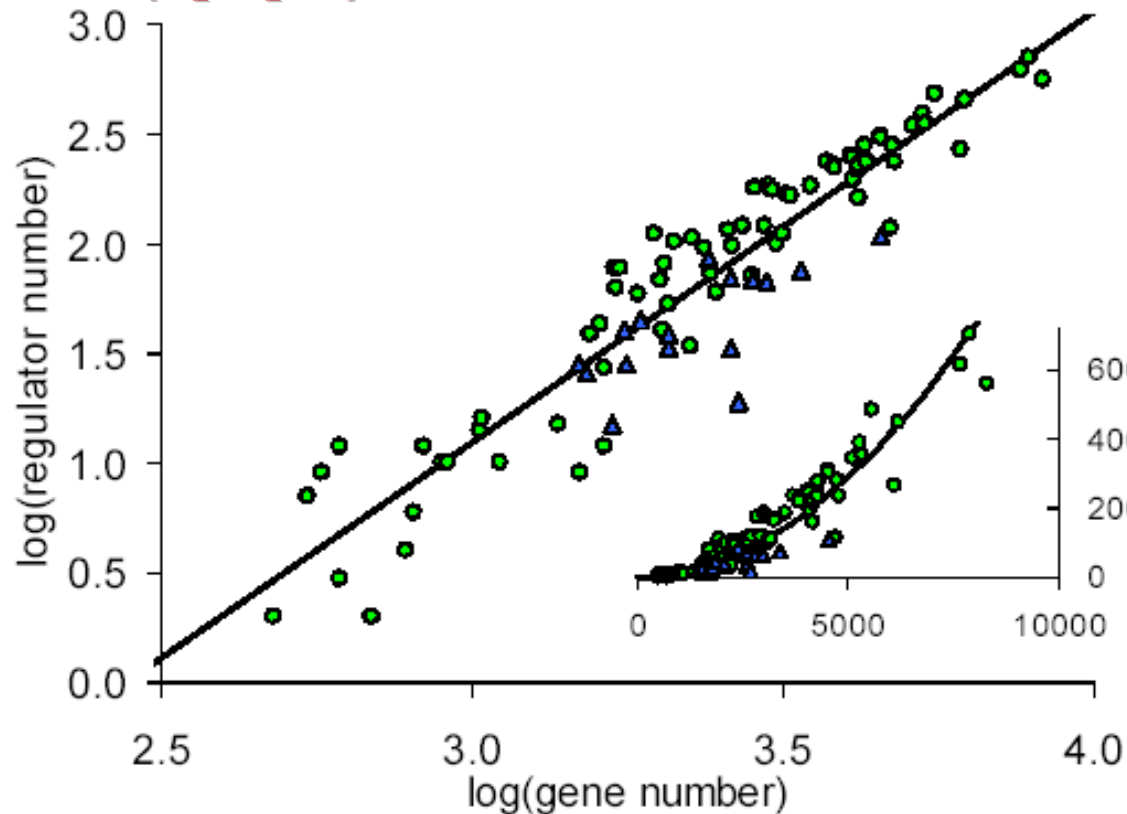


Figure 1: Double-logarithmic plot of transcriptional regulator number against total gene number for bacteria (green circles) and archaea (blue triangles). The overall distribution is well described by a straight line with slope 1.96 ($r^2 = 0.88$, 95% confidence interval: 1.81 – 2.11), corresponding to a quadratic relationship between regulator number and genome size. The inset shows the same data before log-transformation.

$$R = \sum_{n=0}^N cn = \frac{cN(N+1)}{2} \approx \frac{c}{2} N^2$$

Чтобы увеличить число генов-регуляторов, надо увеличить число регулируемых единиц. Иными словами – разбить опероны на более мелкие регулируемые единицы. Но это должно вызвать рост размера генома, так как каждый ген потребует своего регуляторного района. Рост размера генома запрещен катастрофой мутационных ошибок Эйгена. Таким образом, число генов-регуляторов у бактерий должно быть ограничено сверху.

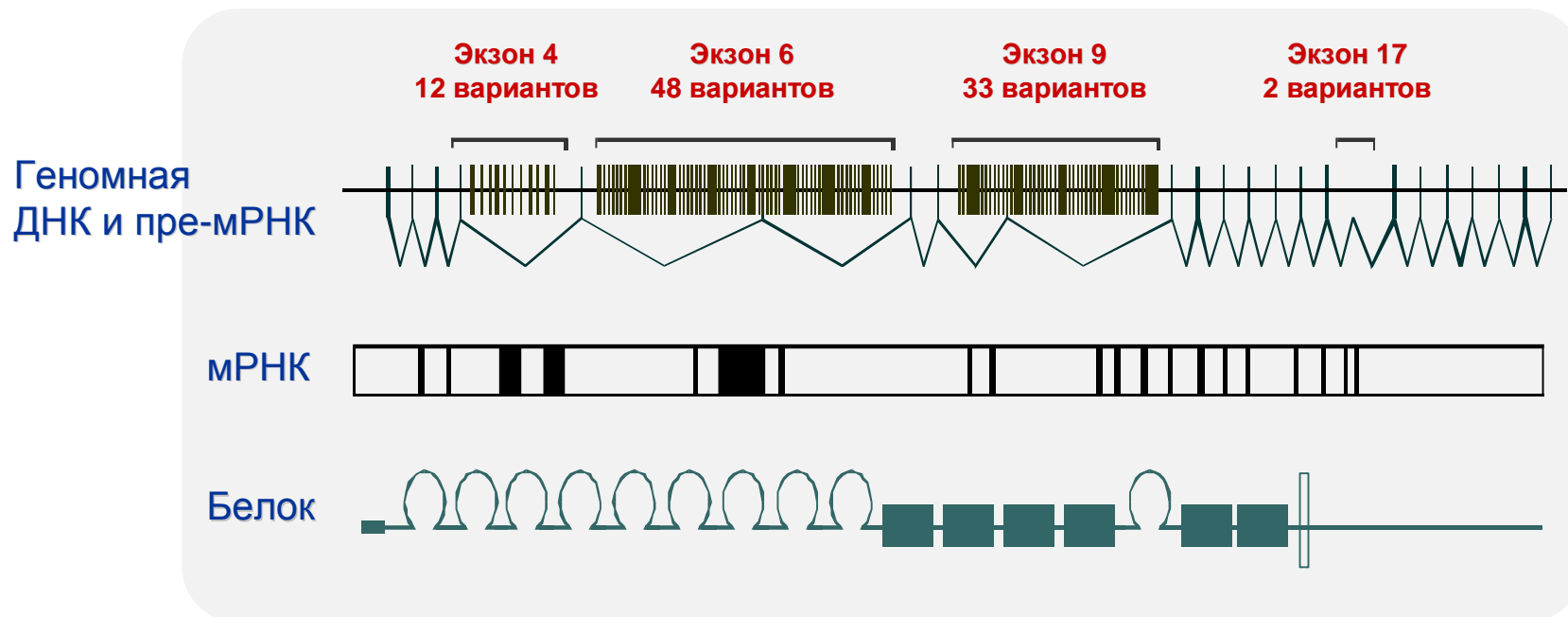


Экзон-интронная структура и альтернативный сплайсинг эукариот обеспечивают огромное число регулируемых единиц без значительного роста числа генов и размеров генома, что позволяет эукариотам отодвинуть верхнюю границу для числа генов-регуляторов.

$$R = \sum_{n=0}^N cn = \frac{cN(N+1)}{2} \approx \frac{c}{2} N^2$$

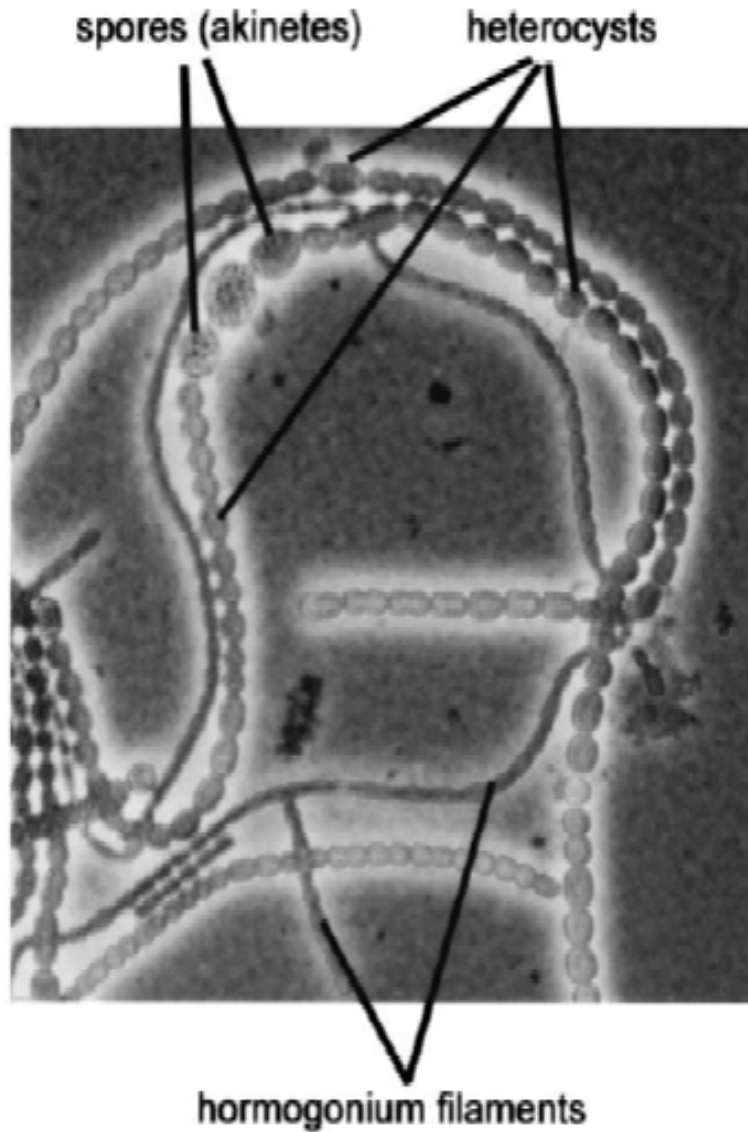
Альтернативный сплайсинг обеспечивает продукцию огромного количества вариантов белка DSCAM, участвующего в формировании тонкой нервной системы дрозофилы:

$N = 12 \times 48 \times 33 \times 2 \dots$

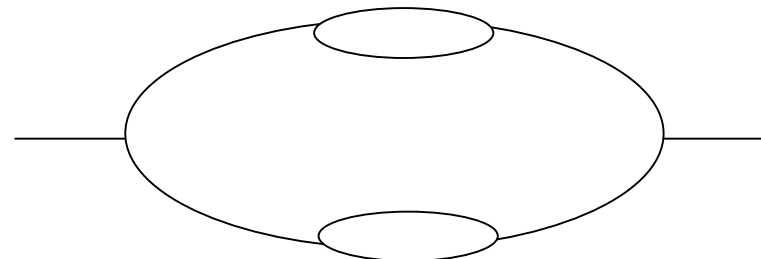




Рекомбинация по повторам ограничивает их распространение в геномах прокариот

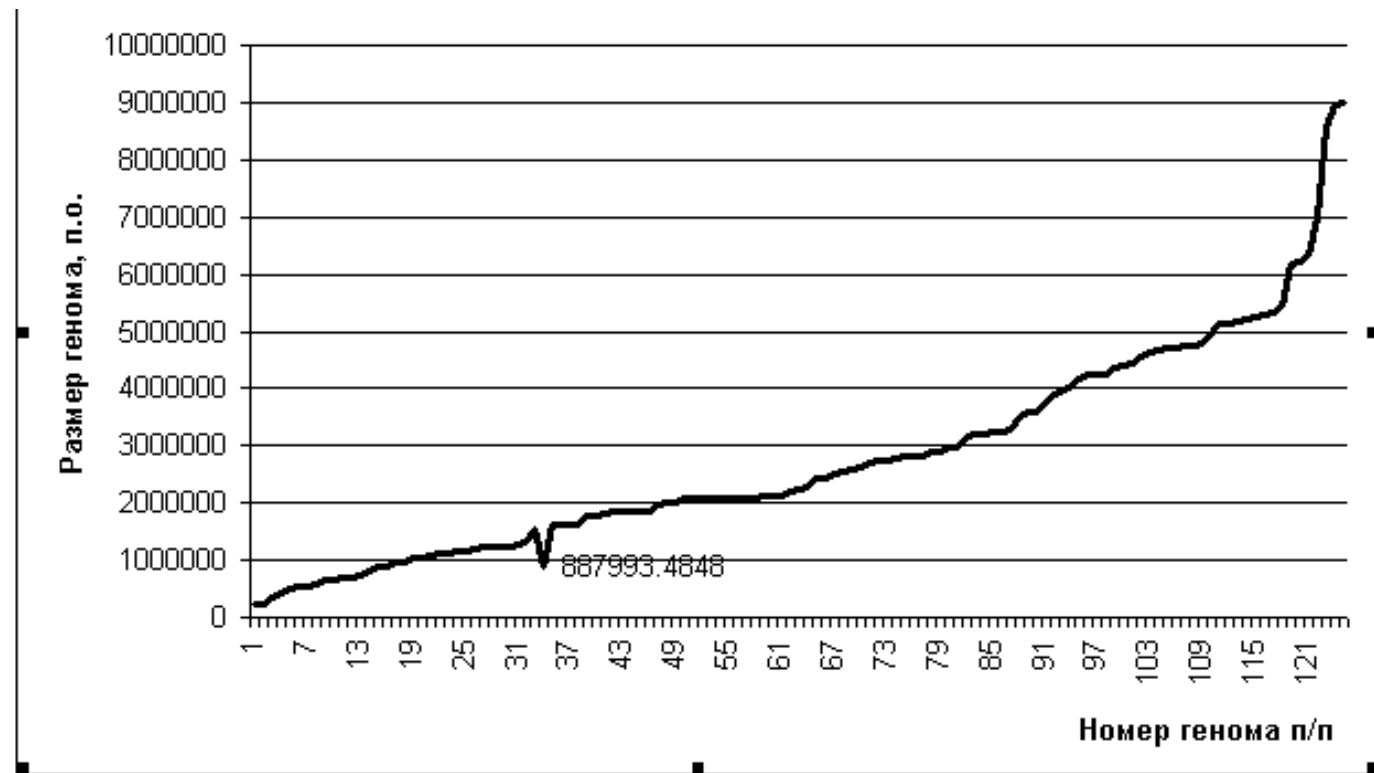


- 1) Гаплоидность и оперонная структура делает прокариот чувствительными к делециям.
- 2) Неразвитая компактизация ДНК в нуклеоиде делает длину однонитевой ДНК в репликативной вилке прокариот (1000 – 2000 п.н.) в 10 раз больше, чем у эукариот (100 – 200 п.н.)
- 3) При высокой скорости роста инициация нового раунда репликации может начаться еще до завершения предыдущего.





Распределение размеров геномов у эубактерий



Распределение размеров генома у эубактерий. ¶

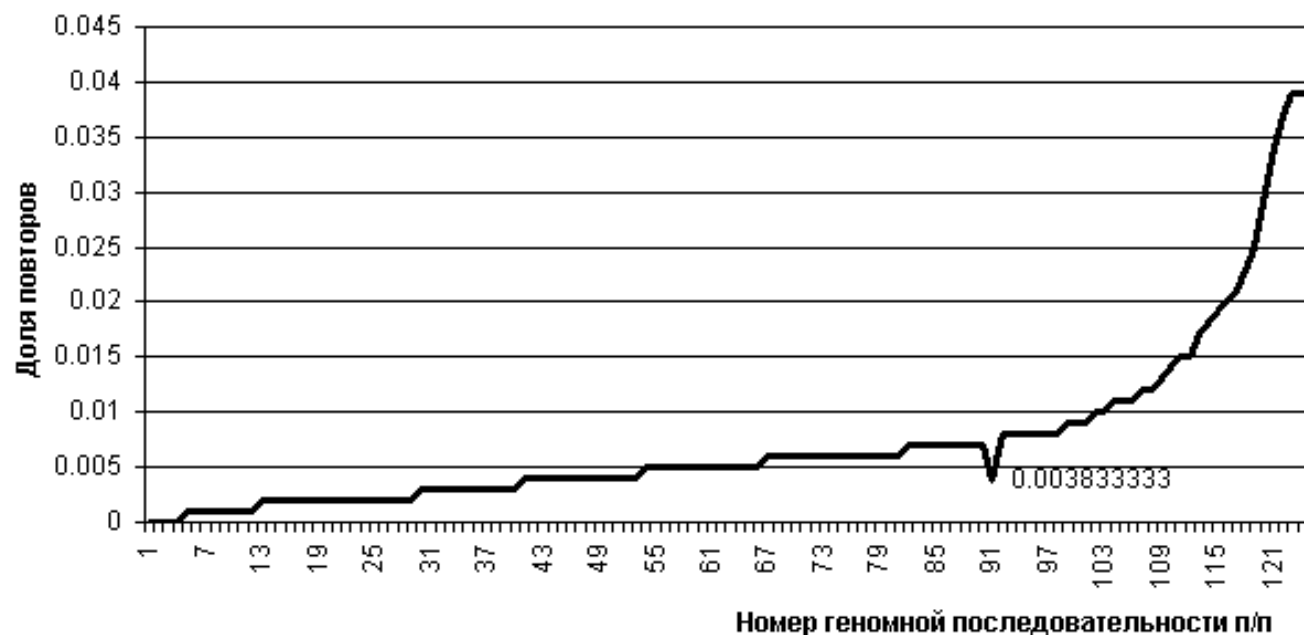
Среднее значение показано цифрой - менее 1 Мб. ¶

Размеры варьируют от 500 Кб до 9 Мб. ¶

Примечание - каждая хромосома одного организма учитывалась по отдельности. ¶



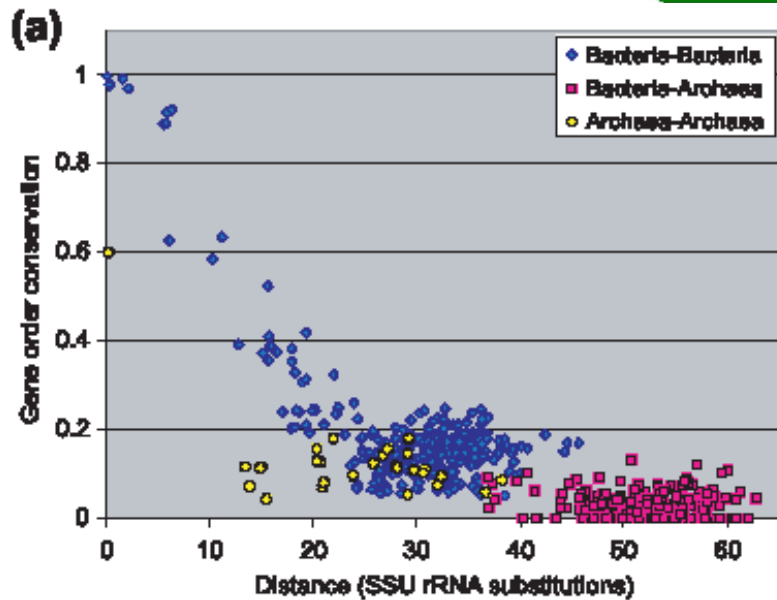
Доля неслучайных совершенных повторов относительно полного размера геномов у эубактерий



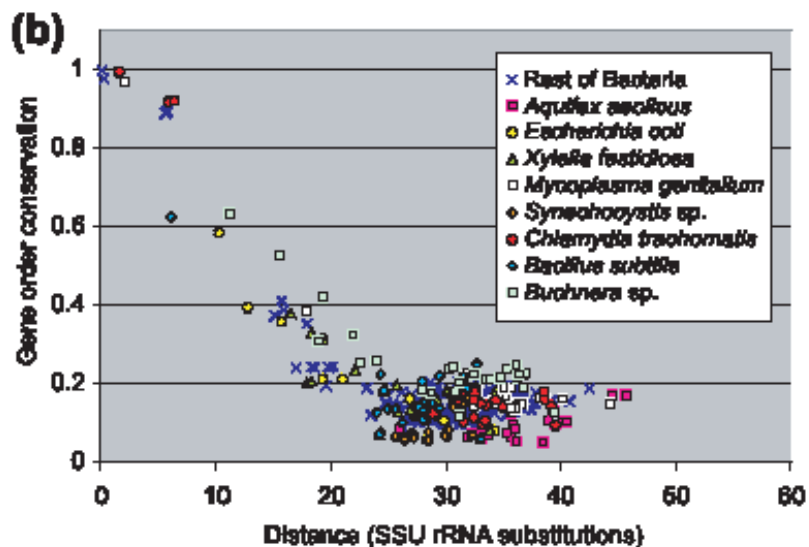
Оценка доли избыточности как общей длины неслучайных совершенных повторов относительно размера генома дана ниже. В среднем доля таких повторов (прямых и инвертированных повторов более 20 п.о.) составляет менее 1%, точнее 0.38%. ¶



Консервативность порядка генов в эволюции: прокариоты.



Сигмовидный тип кривой показывает, что оперонная структура стабилизирует порядок генов в геноме.





Консервативность порядка генов в эволюции: прокариоты

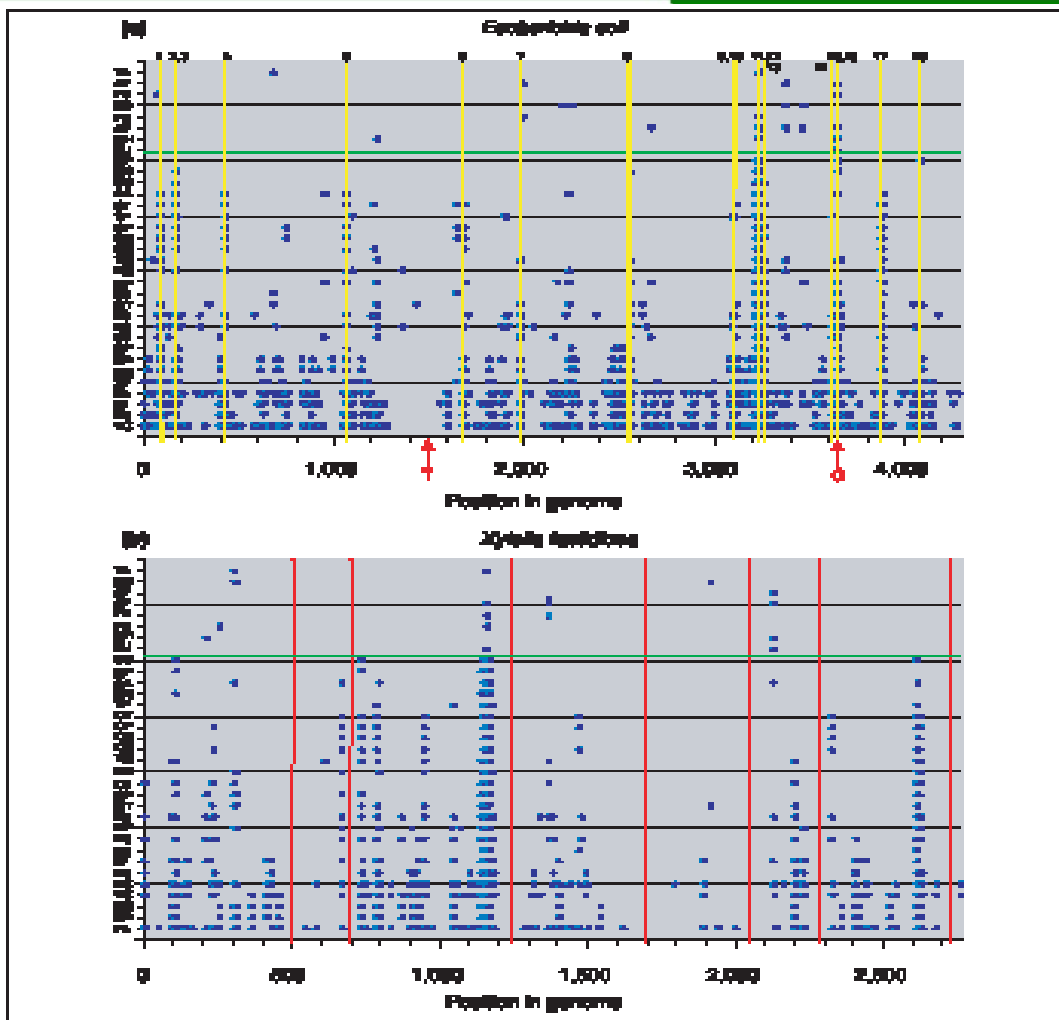


Figure 4
Gene order conservation in the species studied, using (a) *Escherichia coli* and (b) *Xylella fastidiosus* as a reference. Position in the reference genome means number of genes from minute zero. Individual species are plotted in the y axis and are ordered according to their phylogenetic distance (estimated by SSU rRNA substitutions) to the reference species. The more closely related species are shown lower down and more distantly related species higher up the axis. Species names are listed in Table 2. Blue dots indicate genes belonging to conserved runs for each species. A horizontal green line separates Bacteria from Archaea. (a) For *E. coli*, yellow lines show the regions with especially high conservation of gene order. A detailed study of these regions can be found in Table 1. The origin and terminus of replication are marked O and T, respectively, at the bottom of this graph. (b) For *X. fastidiosus*, red lines indicate regions of high frequency of unique genes [25]. A low degree of gene order conservation was found in these regions.

Наиболее эволюционно стабилен порядок генов с оптимизированными функциями, особенно, если они находятся вдали от горячих точек рекомбинации (верхний рисунок). Наоборот, порядок эволюционно недавно приобретенных генов, или генов с уникальными функциями нестабилен (нижний рисунок).



Прокариоты: связь оперонов и консервативности порядка генов.



Position*	Gene	% Conservation†	Function	Functional class‡
Group 12				
3217	<i>rplQ</i>	73	50S ribosomal subunit protein L17	Translation
3218	<i>rpoA</i>	76	RNA polymerase, alpha subunit	Transcription
3219	<i>rpsD</i>	39	30S ribosomal subunit protein S4	Translation
3220	<i>rpsK</i>	76	30S ribosomal subunit protein S11	Translation
3221	<i>rpsM</i>	76	30S ribosomal subunit protein S13	Translation
3222	<i>rpmJ</i>	42	50S ribosomal subunit protein L36	Translation
3223	<i>prlA</i>	70	Putative ATPase subunit of translocase	Cellular processes (translocation)
3224	<i>rplO</i>	15	50S ribosomal subunit protein L15	Translation
3225	<i>rpmD</i>	33	50S ribosomal subunit protein L30	Translation
3226	<i>rpsE</i>	73	30S ribosomal subunit protein S5	Translation
3227	<i>rplR</i>	64	50S ribosomal subunit protein L18	Translation
3228	<i>rplF</i>	70	50S ribosomal subunit protein L6	Translation
3229	<i>rpsH</i>	82	30S ribosomal subunit protein S8, and regulator	Translation
3230	<i>rpsN</i>	27	30S ribosomal subunit protein S14	Translation
3231	<i>rplE</i>	88	50S ribosomal subunit protein L5	Translation
3232	<i>rplX</i>	58	50S ribosomal subunit protein L24	Translation
3233	<i>rplN</i>	88	50S ribosomal subunit protein L14	Translation
3234	<i>rpsQ</i>	67	30S ribosomal subunit protein S17	Translation
3235	<i>rpmC</i>	24	50S ribosomal subunit protein L29	Translation
3236	<i>rplP</i>	76	50S ribosomal subunit protein L16	Translation
3237	<i>rpsC</i>	82	30S ribosomal subunit protein S3	Translation
3238	<i>rplV</i>	58	50S ribosomal subunit protein L22	Translation
3239	<i>rpsS</i>	70	30S ribosomal subunit protein S19	Translation
3240	<i>rplB</i>	82	50S ribosomal subunit protein L2	Translation
3241	<i>rplW</i>	33	50S ribosomal subunit protein L23	Translation
3242	<i>rplD</i>	73	50S ribosomal subunit protein L4	Translation
3243	<i>rplC</i>	76	50S ribosomal subunit protein L3	Translation
3244	<i>rpsJ</i>	61	30S ribosomal subunit protein S10	Translation





Консервативность порядка генов в эволюции: эукариоты



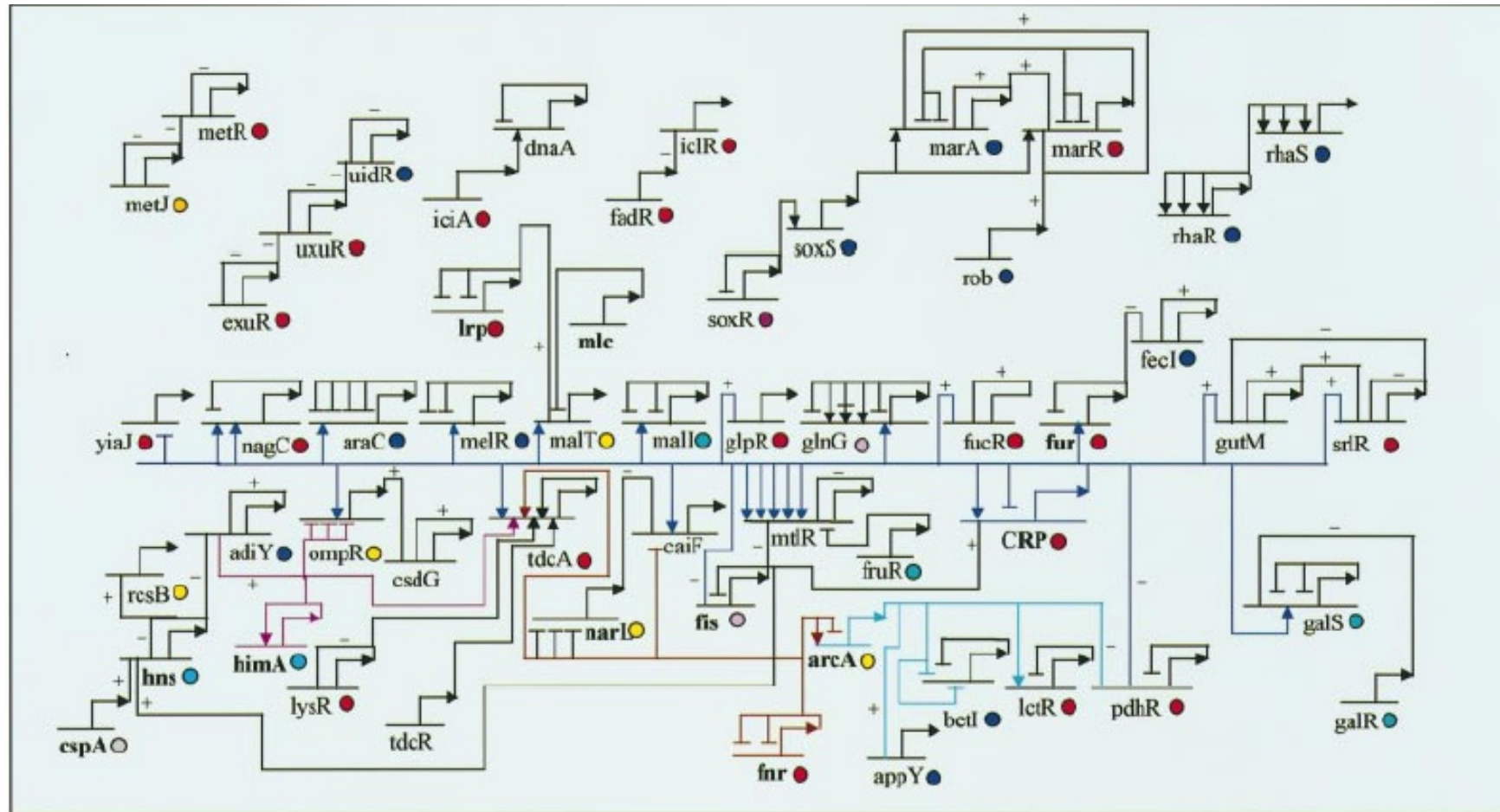
Table 1. Stable complexes with conserved co-regulation in yeast and worm.

	Complex	No. of co-regulated proteins
Translation	Cytoplasmic ribosomal large subunit	23
	Cytoplasmic ribosomal small subunit	21
	Translation complexes (eEF1, eEF2, eIF3)	5
Transcription	RNA polymerase I	2
RNA processing	Nop56p–Nop1p complex	2
Replication	Replication fork complexes	4
	Pre-replication complex	2
Electron transfer chain	Cyt. bc1 complex (complex III), Cyt. c oxidase (complex IV)	2
	F0–F1 ATP synthase (complex V)	2
	Proteasome	19–22S regulator
	20S proteasome	4
	Chaperonin containing T-complex TRiC	3
	Total	76

Порядок наиболее жестко корегулируемых генов эволюционно стабилен. Как правило белковые продукты этих генов тесно (порой физически) взаимодействуют.



Регуляция транскрипционными факторами у *E.coli*: генная сеть



● Homeodomain-like	● FIS-like	● Nucleic acid-binding proteins
● winged helix DNA binding domain	● Putative DNA-binding domain	● C-terminal effector domain of the Bipartite response regulator
● Met repressor-like	● IHF like DNA-binding domain	● Lambda repressor-like DNA binding domain



Регуляция транскрипционными факторами у *E.coli*: прямая и непрямая регуляция

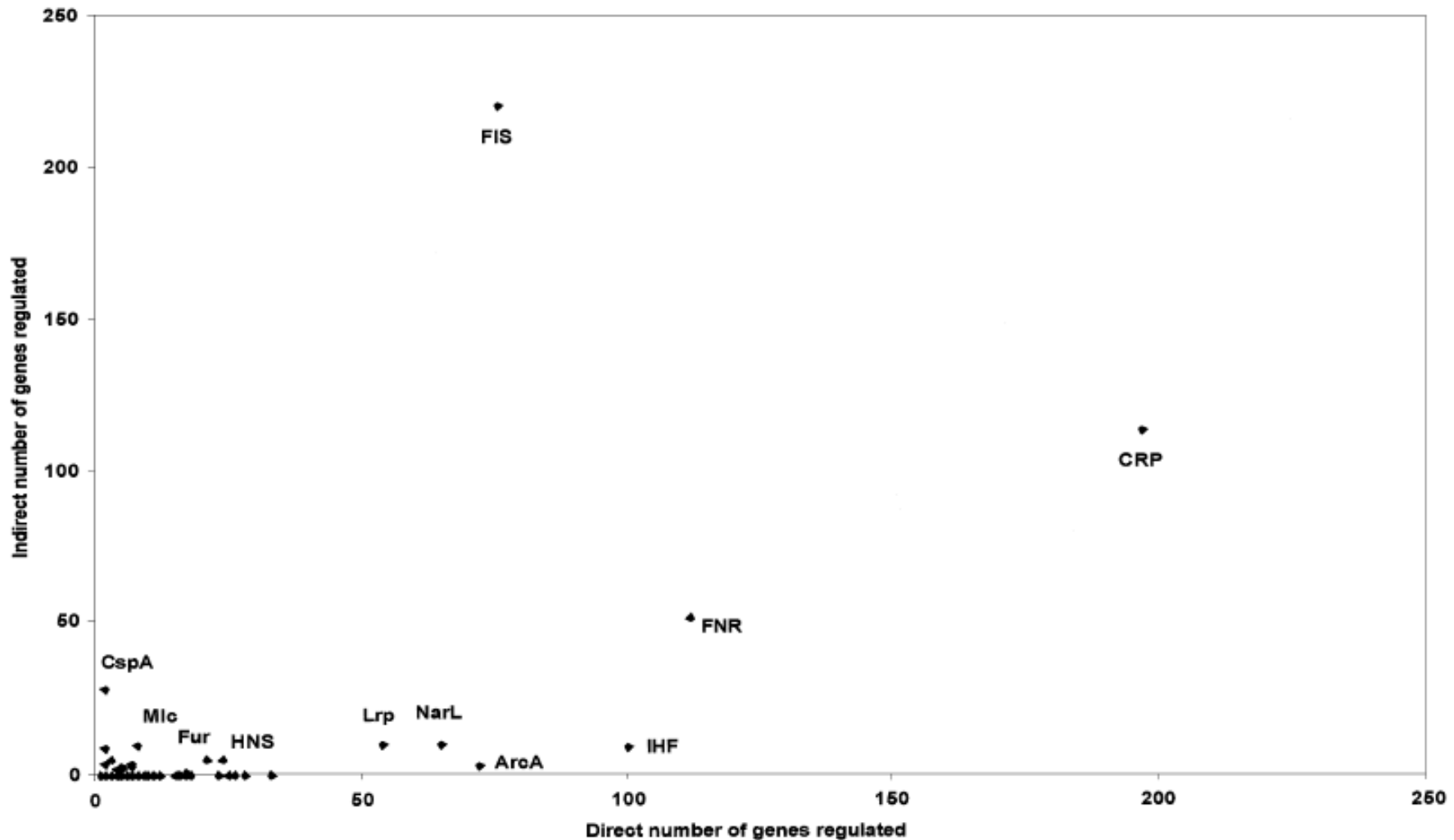


Figure 4. Direct and indirect gene regulation by *E.coli* transcription factors. The direct number of genes regulated is represented on the *x* axis and the indirect number of genes on the *y* axis. The global regulators, which are marked on the graph regulate a large number of genes, and participate in regulatory cascades, resulting in indirect regulation of genes.



Регуляция транскрипционными факторами у *E.coli*: корегуляция несколькими факторами

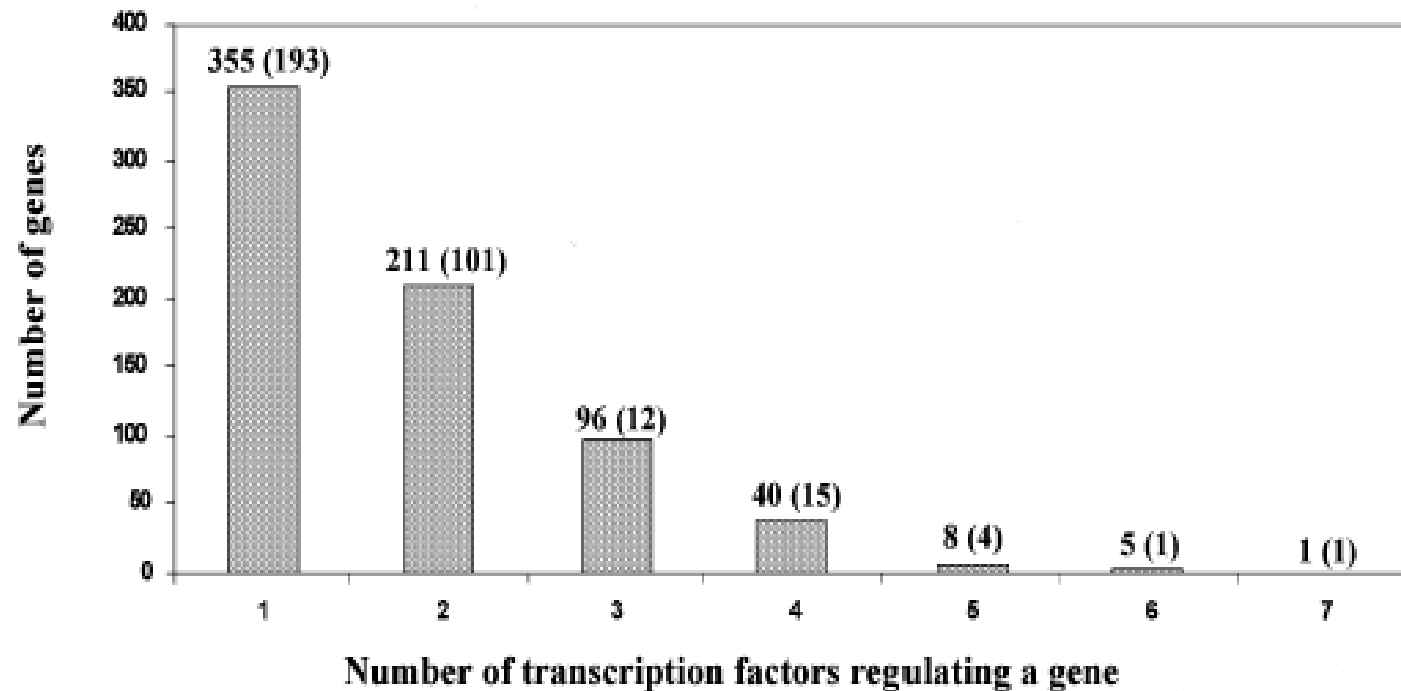


Figure 5. Distribution of the number of transcription factors regulating a gene. Numbers in parentheses represent the number of operons.



Регуляция транскрипционными факторами у *E.coli*: корегуляция несколькими факторами

Table 3. Numbers of transcription factors active at the same promoter

No. of transcription factors	No. of co-regulating transcription factors
26	0
35	1
18	2
12	3
4	4
6	5
2	6
1	7
2	8
2	9
2	10
1	11
1	12
1	13
1	19
1	20
1	20
1	52

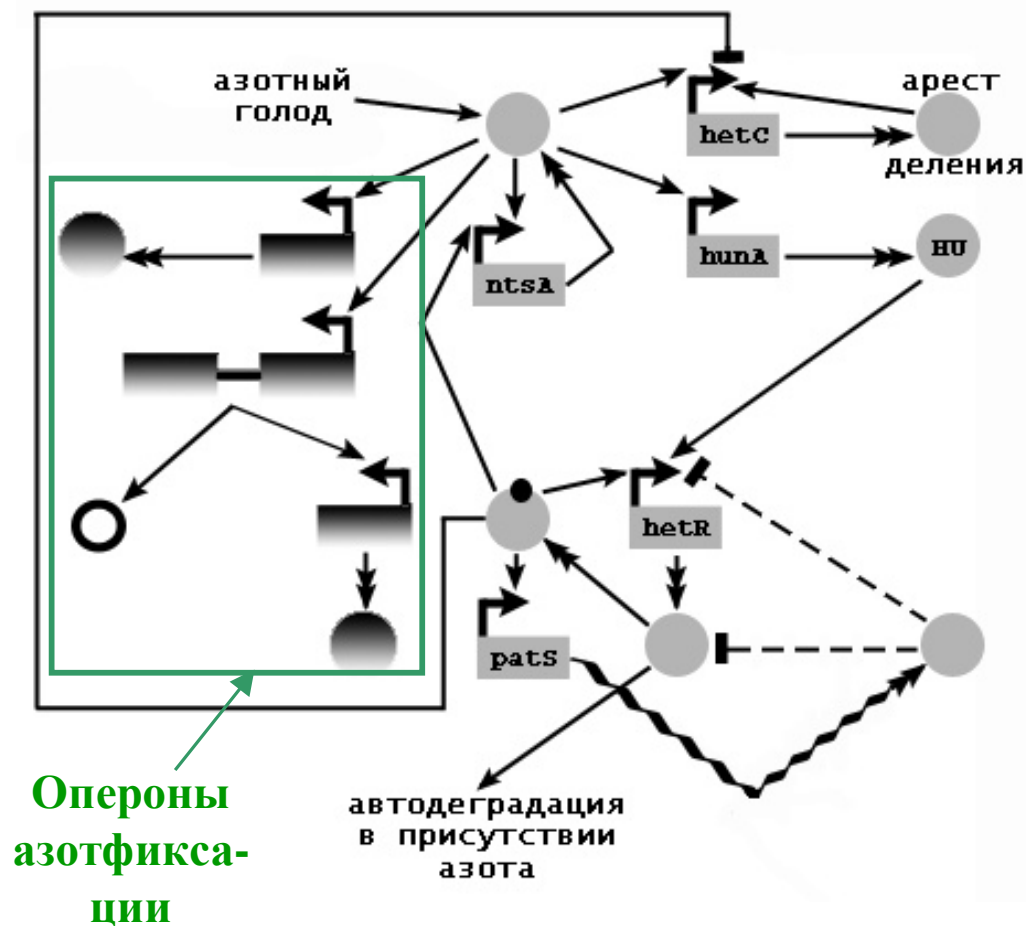
There is one transcription factor that has 52 co-regulating transcription factors (CRP) and 26 transcription factors that have no co-regulating transcription factors at all.

Table 4. Homologous genes regulated by the same transcription factor

Combination of transcription factors regulating operons	Homologous gene(s) in first operon	Homologous gene(s) in second operon	
BirA	BioF	BioA	
Fnr	ArcA	NarL	
HimA	HimA	HimD	
CysB	CysA	CysC	Figure 6A
CysB	CysM	CysK	Figure 6A
CysB	CysH	CysD	Figure 6A
TyrR	AroF	AroG	
ArgR	ArgF	ArgI	
Fur	FepC	FhuC	Figure 6A
Fur	FepB	FhuD	Figure 6A
Fur	FepA	FhuA	Figure 6A
Hns	MukB	ProV	
PhoB	PhnCDKLN	PstBS	
PhoB	PhnCKLN	PhoH	
CpxR	LpxD	LpxA	
LexA	RecN	UvrD	
PurR	PurD	PurK	
Crp, Fis	HupA	HupB	Figure 6B
Fnr, NarL	DmsAB	FdnGH	Figure 6B
Fnr, NarL	DmsB	NrfC	Figure 6B



Генная сеть дифференцировки азотфиксирующих клеток - гетероцист у цианобактерий.



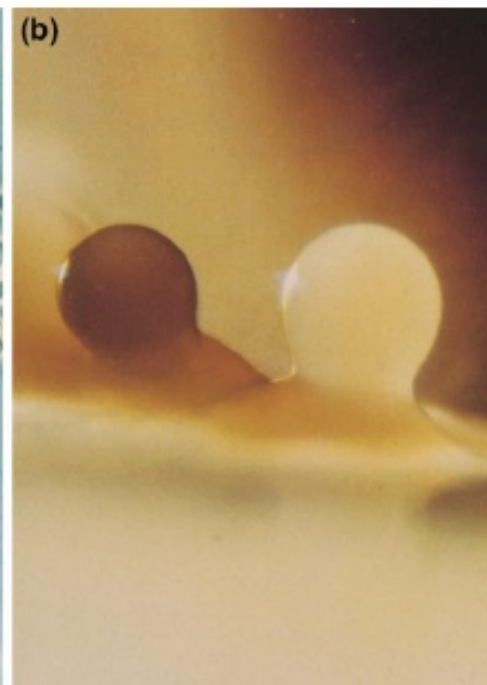
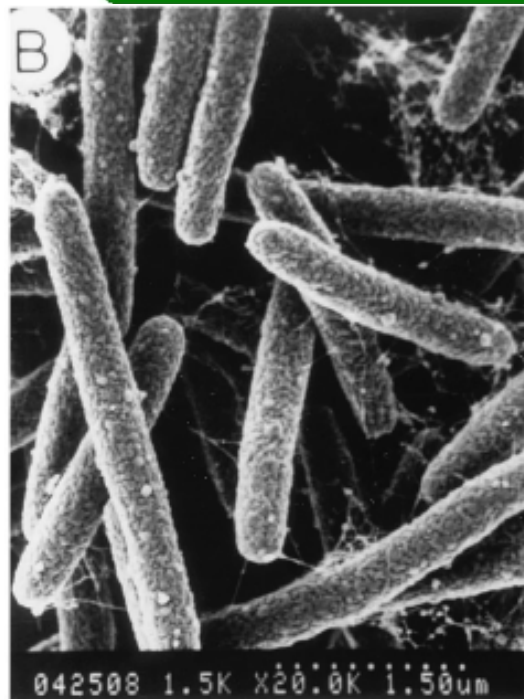
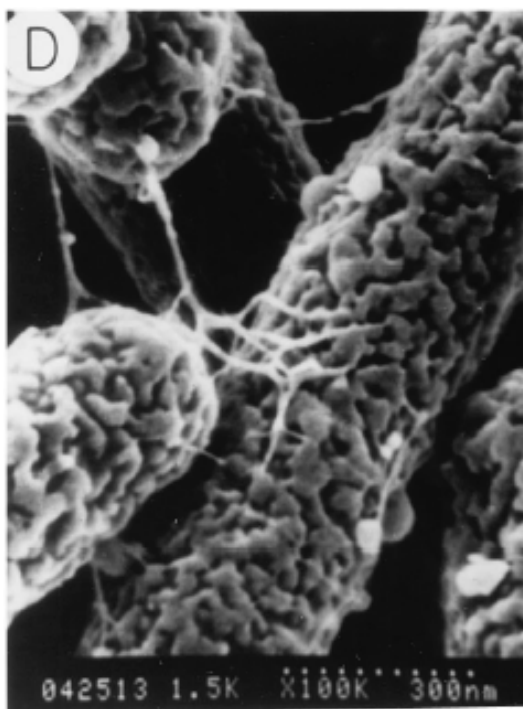
Азотный голод вызывает дифференцировку гетероцист, активируя белок NtcA, активирующий синтез белков HetC, HetR и самого себя. HetC связан с арестом деления. HetR подавляет HetC, разрешая синтез ДНК и запуская транскрипцию оперонов азотфиксации, белков PatS и NtcA. Положительная обратная связь HetR -> NtcA -> HetC и отрицательные связи HetR -> HetC стабилизируют процесс. PatS, благодаря малым размерам, диффундирует в клетки-соседи гетероцисты, подавляя HetR (аналог градиента морфогенов эукариот), чем по-видимому и исчерпывается межклеточная коммуникация.

PatS –
"морфоген" цианобактерий.

AGATTATGAAGGCAATTATGTTAGTGAATTTCTGTGATGAGCCGGTAGTGGTAGATAGAACGA
 M K A I M L V N F C D E R G S G R *
 | | | |
 C D S L



Формирование плодовых тел у миксобактерий





Благодарю за внимание.