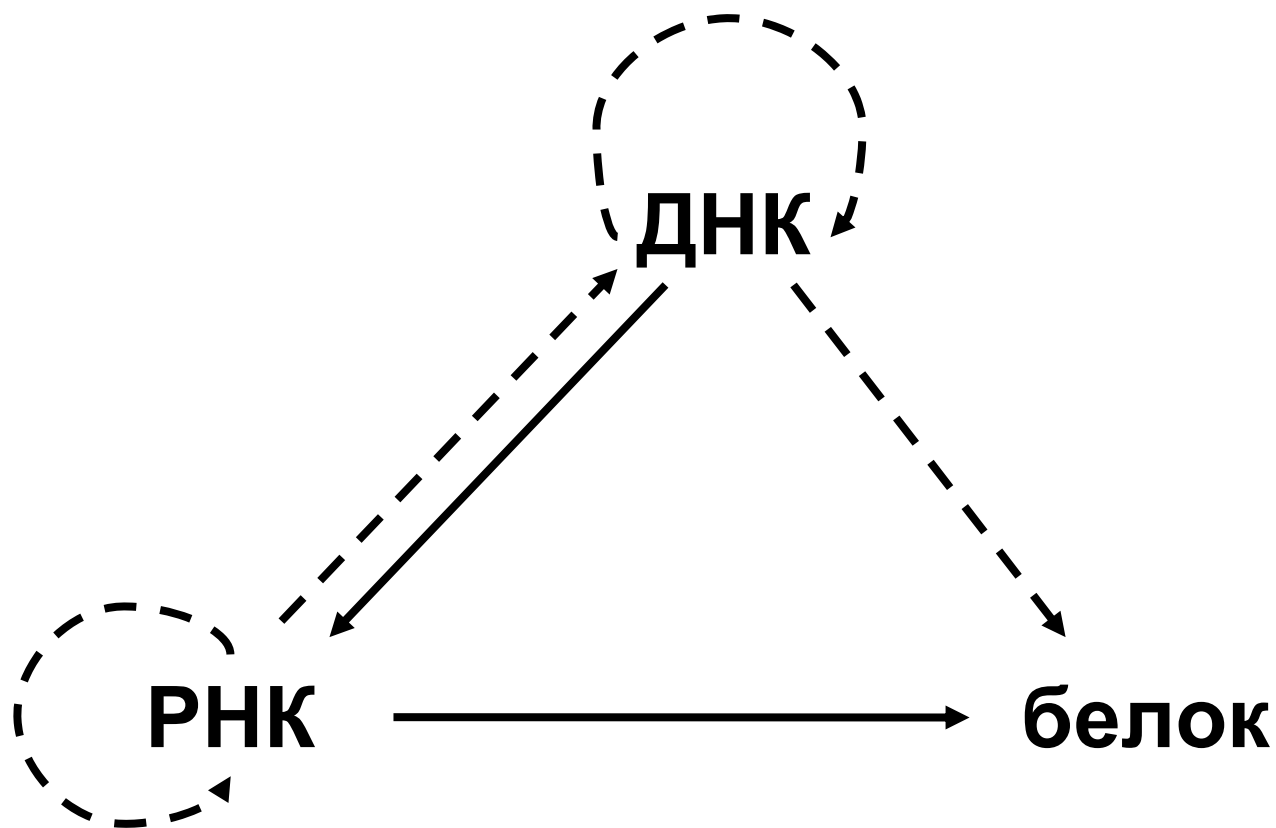


Молекулярные механизмы прогрессивной эволюции многоклеточных организмов:

роль некодирующих последовательностей ДНК

Алсу Фаритовна Сайфитдинова

Центральная догма молекулярной биологии, 1958 г, Френсис Крик

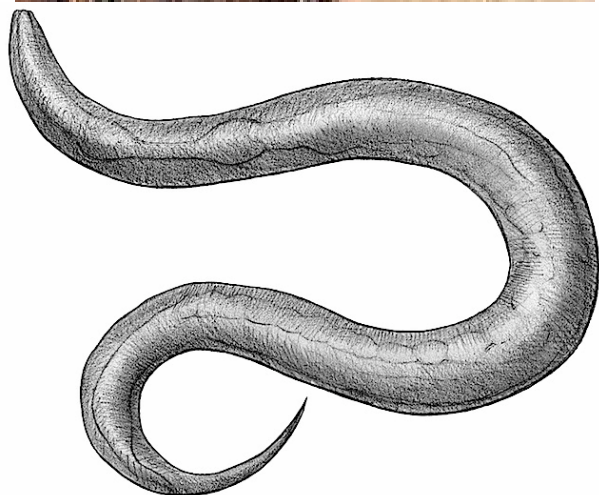


1957 г. Дж. Б. С. Холдейн установил несоответствие реальной скорости видообразования и ожидаемой в соответствии с популяционно-генетической моделью. Мутации не могут обеспечить приспособление и выживаемость популяции при быстром изменении условий внешней среды

1968 г. Мото Кимура выдвинул теорию «нейтральной эволюции», согласно которой ДНК может свободно мутировать на протяжении долгого времени без всякого воздействия со стороны внешней среды.

В 1972 г. Сусуми Оно предложил термин «мусорная ДНК» (*junk DNA*)

1980 г. Лесли Оргель и Френсис Крик закрепили за некодирующей ДНК роль эволюционного балласта в высказывании «It has little specificity and conveys little or no selective advantage to the organism»



Геном человека – организма, состоящего более чем из 100 триллионов клеток, различных по дифференцировке и координировано работающих для обеспечения не только жизнедеятельности, но и мыслительного процесса – включает около 20 тысяч белок-кодирующих последовательностей.

Приблизительно из того же количества генов, причем имеющих значительное сходство с генами человека, состоит геном свободноживущей нематоды *Caenorhabditis elegans* длиной около 1мм, тело которой состоит из одной тысячи соматических клеток, включая 302 нейрона.

Archaea:	Размер генома п.о.	Число цистронов
<i>Aeropyrum pernix</i> K1	1,669,696	1,700
<i>Candidatus Korarchaeum cryptofilum</i> OPF8	1,590,757	1,602
<i>Methanocaldococcus fervens</i> AG86	1,507,251	1,581
<i>Methanococcus maripaludis</i> C5	1,789,046	1,822
<i>Methanosarcina barkeri</i> str. Fusaro	4,873,766	3,625
<i>Nanoarchaeum equitans</i> Kin4-M	490,885	536
<i>Pyrococcus abyssi</i> GE5	1,768,562	1,784
<i>Sulfolobus acidocaldarius</i> DSM 639	2,225,959	2,223
<i>Thermococcus gammatolerans</i> EJ3	2,045,438	2,157
<i>Thermofilum pendens</i> Hrk 5	1,813,393	1,876

Bacteria:	Размер генома п.о	Число цистронов
<i>Acetobacter pasteurianus</i> IFO 3283-01	3,340,249	3,050
<i>Burkholderia cenocepacia</i> MC0-3	7,971,389	7,008
<i>Escherichia coli</i> ATCC 8739	4,746,218	4,200
<i>Candidatus Phytoplasma mali</i>	601,943	497
<i>Chlamydophila pneumoniae</i> CWL029	1,230,230	1,052
<i>Lactobacillus fermentum</i> IFO 3956	2,098,685	1,843
<i>Prochlorococcus marinus</i> str. MIT 9301	1,641,879	1,907
<i>Staphylococcus aureus</i> subsp. aureus ST398	2,885,367	2,710
<i>Streptomyces scabiei</i> 87.22	10,148,695	8,809
<i>Streptomyces bingchenggensis</i> BCW-1	11,936,683	10,022
<i>Treponema pallidum</i> subsp. pallidum str. Nichols	1,138,011	1,031

	Размер генома млн.п.о.	Число предсказанных генов тыс
<i>Saccharomyces cerevisiae</i>	12.1	5.6
<i>Schizosaccharomyces pombe</i>	14	4.8
<i>Arabidopsis thaliana</i>	145	25
<i>Caenorhabditis elegans</i>	97	20.1
<i>Drosophila melanogaster</i>	165	14
<i>Mus musculus</i>	2600	22
<i>Homo sapiens</i> XY / XX	3080/3022	23

Плотность генов в геномах разных организмов

Haemophilus influenzae 1 ген на 1000 п.о.

Escherichia coli 1 ген на 1400 п.о.

Saccharomyces cerevisiae 1 ген на 2000 п.о.

Caenorhabditis elegans 1 ген на 5000 п.о.

Arabidopsis thaliana 1 ген на 4000 п.о.

Drosophila melanogaster 1 ген на 9000 п.о.

Mus musculus 1 ген на 100,000 п.о.

Homo sapiens 1 ген на 100,000 п.о.

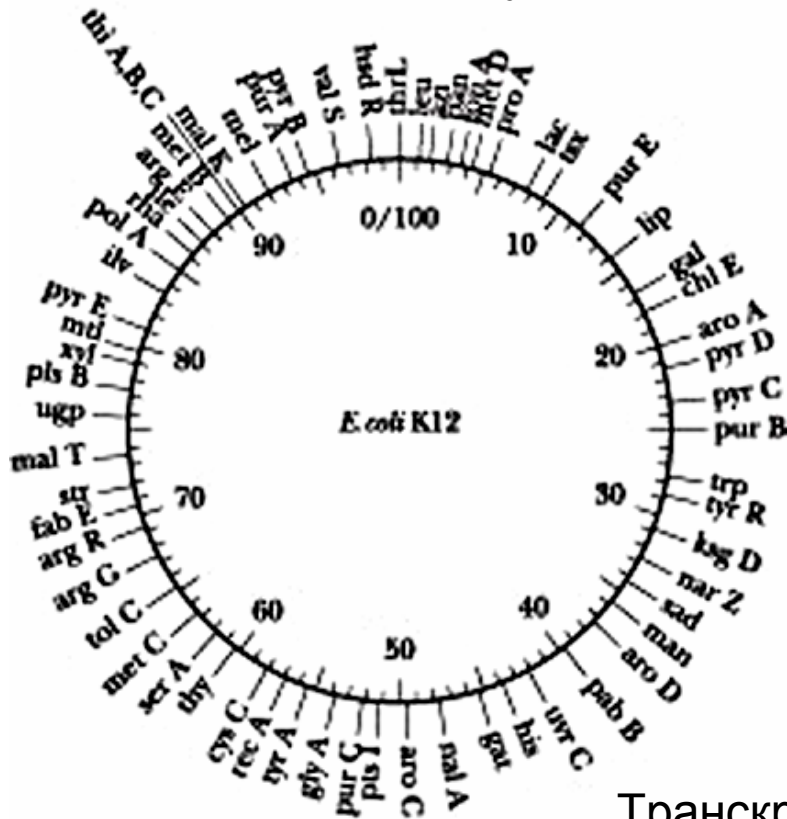
Escherichia coli

87,8% генома – реальные и вероятные белок-кодирующие последовательности

0,8% генома – гены стабильных фракций РНК (21 ген тРНК, 86 генов рРНК)

11,0% генома – регуляторные элементы (репликация, транскрипция)

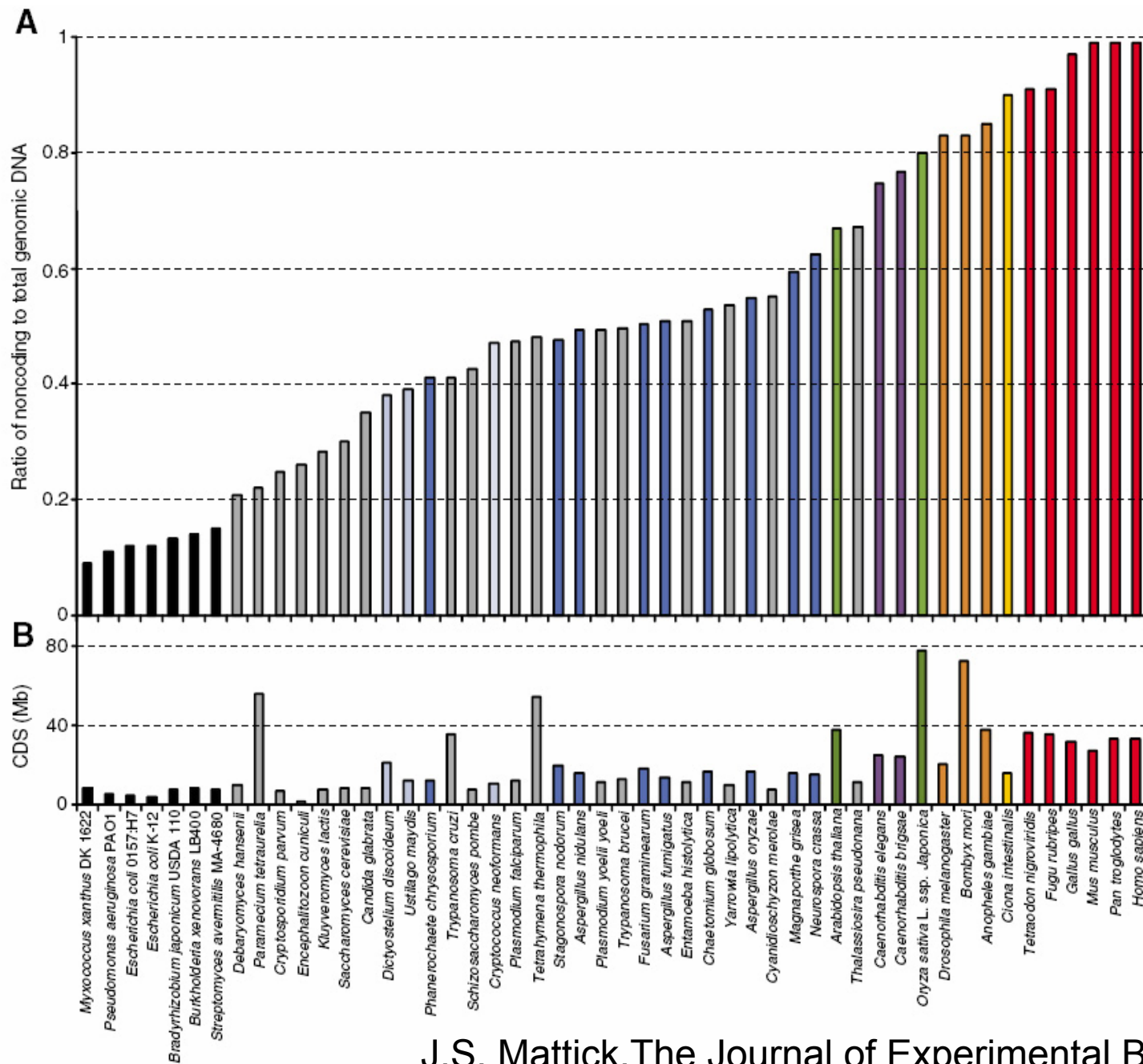
0,7% генома – некодирующие повторы:



Октамер GCTGGTGG – встречается в 186 позициях в прямой ориентации и 148 в обратной ориентации, играет ключевую роль в конъюгационной рекомбинации, соответствует "горячим точкам рекомбинаций"

Повтор REP – палиндром, 40 п.о., 581 копия на геном (составляет 0,54% всей ДНК), функция неизвестна

Транскрибирующиеся регуляторные некодирующие последовательности в составе полицистронных мРНК



Homo sapiens

размер генома	3.02 млрд п.о.,
число предсказанных генов	23 000
средняя длина гена	27 000 нт
среднее число экзонов на ген	8.8
средний размер экзона	145 нт
средний размер интрона	3 365 нт
средняя длина 5'-НТО	300 нт
средняя длина 3'-НТО	770 нт
средний размер кодирующей части	1300 нт

http://www.ornl.gov/sci/techresources/Human_Genome/project/info.shtml

кодирующие белки последовательности составляют 1,2% всего генома

Показано наличие транскриптов для

93% генома *Homo sapiens*

70% генома *Mus musculus*

85% генома *Drosophila melanogaster*

70% генома *Caenorhabditis elegans*

E. Birney, et al. Nature, 2007

P. Carninci et al. Science, 2005

J.R. Manak et al. Nat Genet, 2007

H. He et al Genome Research, 2007

По данным базы данных FANTOM из всех транскриптов из генома мыши длиной более 200 нт около 34% не кодируют белок.

60% из них подвергнуты сплайсингу, 44% полиаденилированы.

70% транскриптов перекрываются с транскриптами в другой ориентации, образуя дуплексы с 58.7 % некодирующих РНК и 87% всех кодирующих белки РНК (18,021 генов имеют участки , перекрывающиеся другими кодирующими и некодирующими последовательностями).

некодирующие участки мРНК: 5'-НТО, 3'-НТО, интроны
регуляция трансляции, предшественники других типов РНК

длинные некодирующие РНК (от 200нт и больше)
изменение состояния хроматина, ретротранспозиция, регуляция транскрипции, трансляции, компартментализация ядра, хранение сигналов быстрого ответа на стресс, позиционная информация, межклеточные взаимодействия, предшественники других типов РНК

рибосомальные РНК
каталитическая составляющая рибосом

малые ядерные РНК (до 300нт)
сплайсинг, регуляция транскрипции, поддержание целостности теломер

малые ядрышковые РНК (60-300нт)
редактирование РНК, процессинг рРНК, тРНК, мяРНК

транспортные РНК (около 80нт)
перенос специфических аминокислот к местам синтеза белка на рибосоме

микроРНК (21-22нт)

образуются из коротких несовершенных шпилек

модуляция трансляции, направленная деградация мишени - мРНК

малые интерферирующие РНК (20-25нт)

образуются из дуплексов и длинных шпилек

деградация вирусных РНК, изменение состояния хроматина, регуляция транскрипции

взаимодействующие с PIWI РНК (28-33нт) и ассоциированные с повторами РНК (23-28нт)

образуются из однонитевых предшественников

инактивация транспозонов в половых и соматических клетках

сканирующие РНК (около 25 нт)

транскрибируются с генома микронуклеуса во время мейоза,

транскрипция контролируется РНК интерференцией

определяют точные границы разрезания хроматина во время реконструкции макронуклеуса инфузорий

Chlamydomonas reinhardtii

размер генома 121 млн п.о., 17 хромосом
15,143 белок-кодирующих последовательностей,
большинство из них содержит интроны (в
среднем 8 интронов на ген)

3 кластера рРНК, 259 генов тРНК.

70% составляют некодирующие
последовательности:

61 класс простых повторов; около 100 семейств
ретротранспозонов; 64 семейства коротких
промежуточных элементов (SINEs)
происходящих от генов тРНК



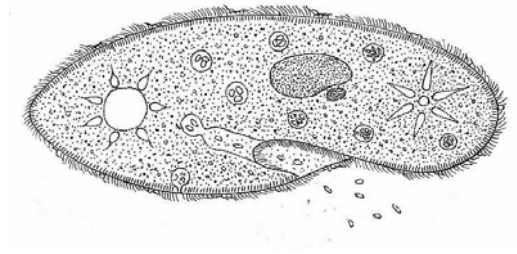
Merchant, S.S. et al. (117 authors) Science, 2007

Длинные шпилечные структуры 150 – 729 нт. служат основой для получения микроРНК, которая участвует в регуляции гаметогенеза.

Molnar A, et al. Nature, 2007

Zhao T, et al. Genes and Development, 2007

Инфузории



Paramecium tetraurelia – короткие сканирующие РНК 25 нт. точно определяют границы последовательностей, которые должны остаться в новом макронуклеусе

Lepe`re G. et al. Genes and Development, 2008

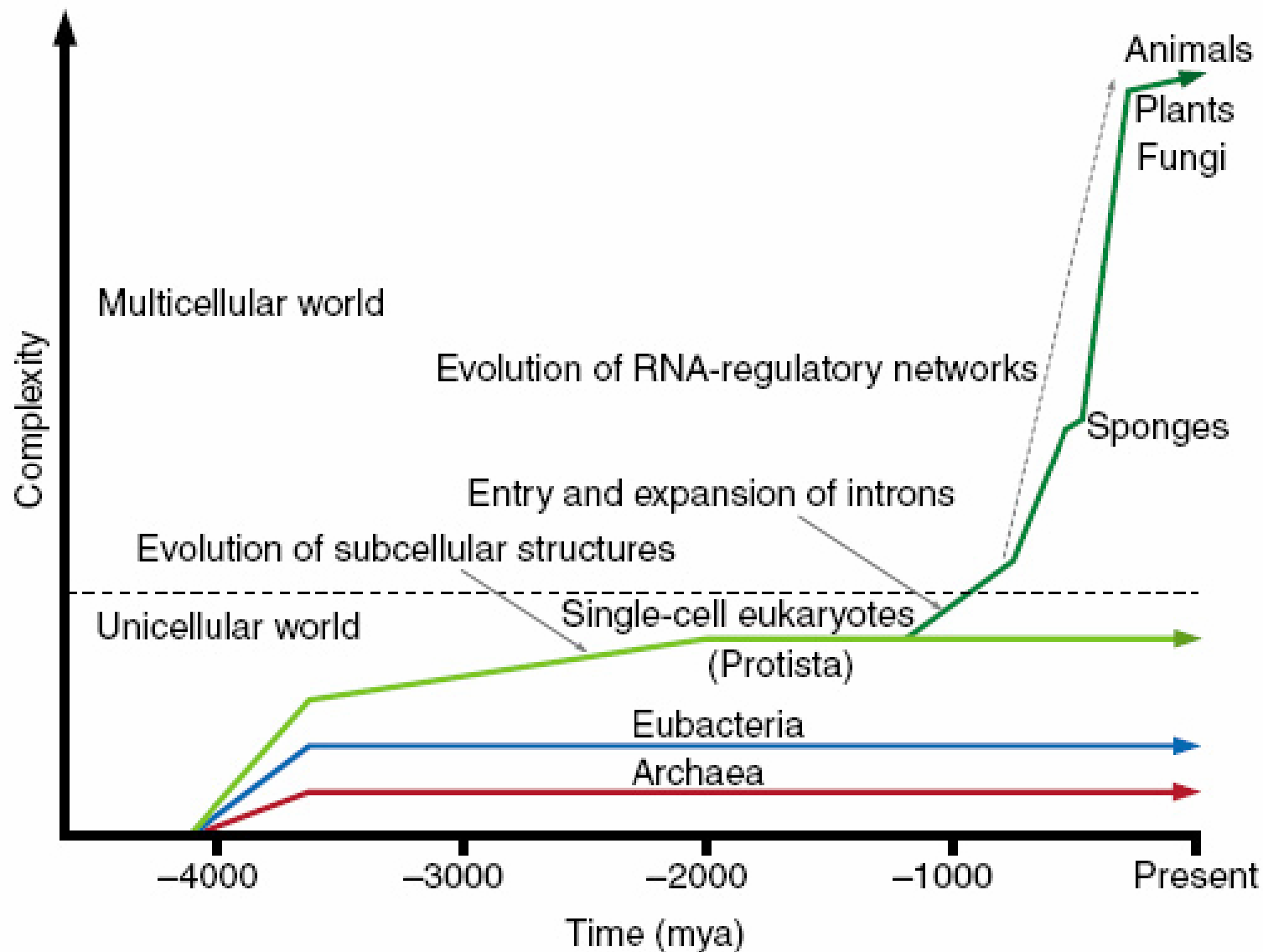
Tetrahymena – малые интерферирующие РНК участвуют в разметке ДНК перед элиминацией хроматина созревающего макронуклеуса

Liu Y, et al. Genes and Development, 2007

Oxytricha trifallax – при созревании макронуклеуса длинные материнские некодирующие РНК покрывают участки генома, которые не должны быть элиминированы.

Nowacki M. et al. Nature, 2007

- ✓ реакция на изменяющиеся условия окружающей среды
- ✓ защита от вирусов
- ✓ хранение сигналов быстрого ответа на стресс
- ✓ компарментализация ядра и клетки
- ✓ структурная организация хроматина
- ✓ компенсация дозы генов, как за счет активации транскрипции, так и за счет гетерохроматинизации
- ✓ импринтинг
- ✓ организация хроматина гамет
- ✓ дифференцировка половых клеток
- ✓ ответ на межклеточные взаимодействия
- ✓ формирование позиционной информации
- ✓ регулирование индивидуального развития
- ✓ формирование осей тела
- ✓ сегментация
- ✓ дифференцировка органов и тканей
- ✓ развитие нервной системы



У человека и мыши доля высококонсервативных последовательностей генома составляет 5%.

R.H. Waterston et al. Nature, 2002

Длинная некодирующая РНК Xist млекопитающих имеет непротяженные высококонсервативные участки.

K.C.Pang et al. Nrendt in Genetics, 2006

В среднем степень сходства в длинных некодирующих РНК составляет 4,1-5,5 %, что практически соответствует таковому в мРНК – 5,2%.

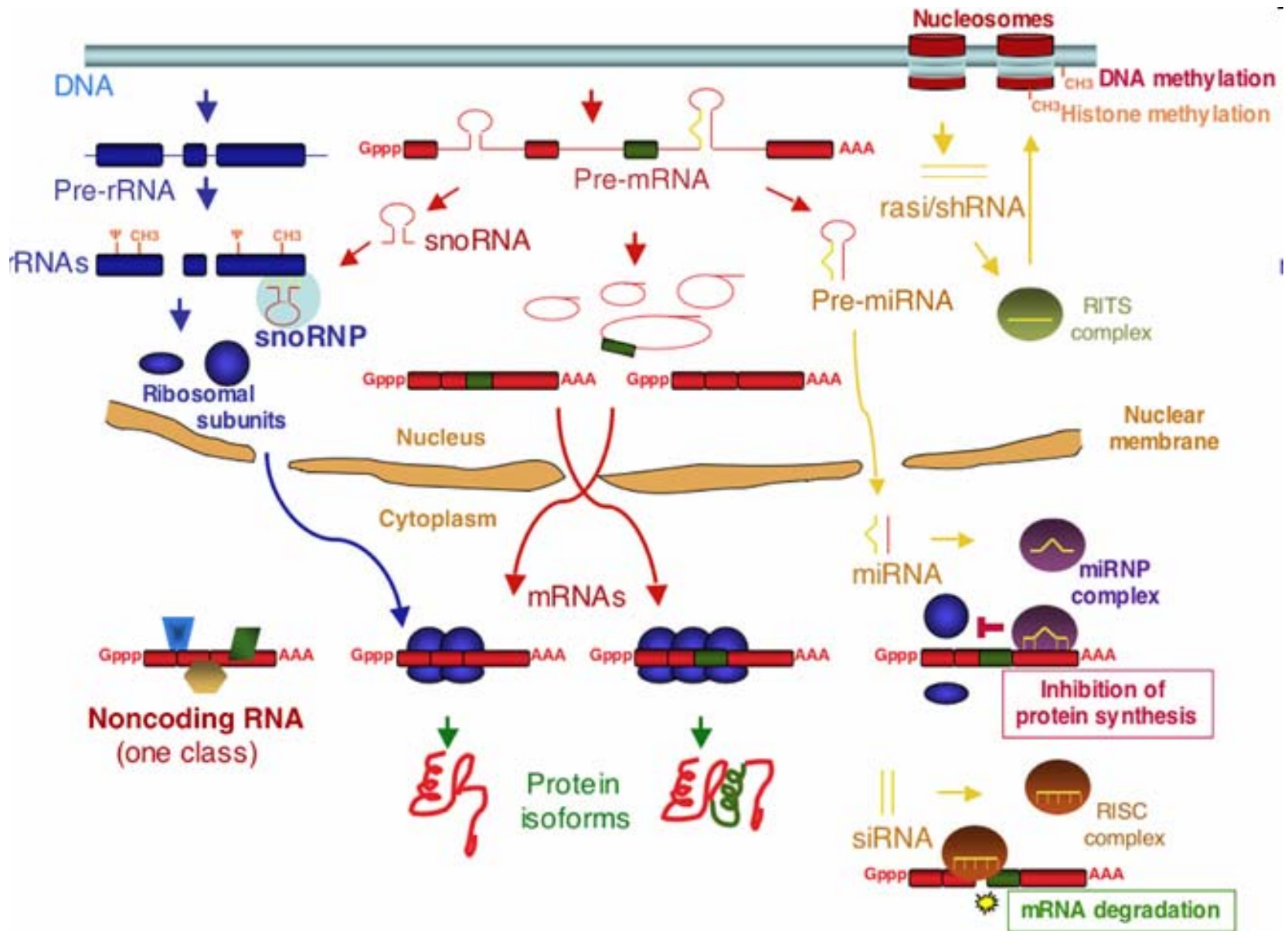
J. Ponjavic et al. Genome Research, 2007

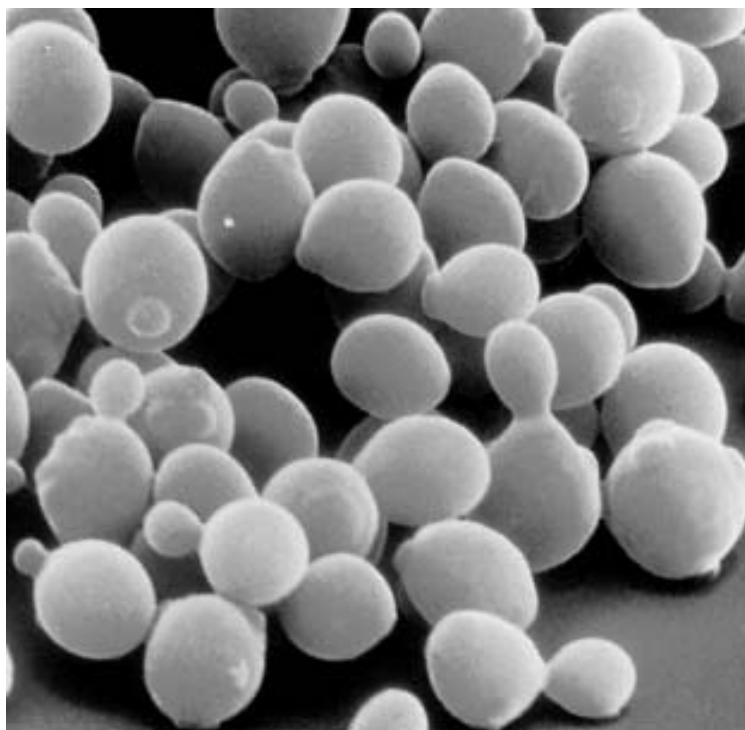
В геномах млекопитающих наиболее консервативные участки кодируют микро РНК, регулирующие тканевую дифференцировку и процессы развития.

K.C.Pang et al. Nrendt in Genetics, 2006

Наименее консервативные участки генома кодируют быстро эволюционирующие видоспецифичные микро РНК, к которым относятся, например, специфичные для мозга человека транскрипты.

P.P. Amaral, J.S.Mattick. Mammalian Genome, 2008

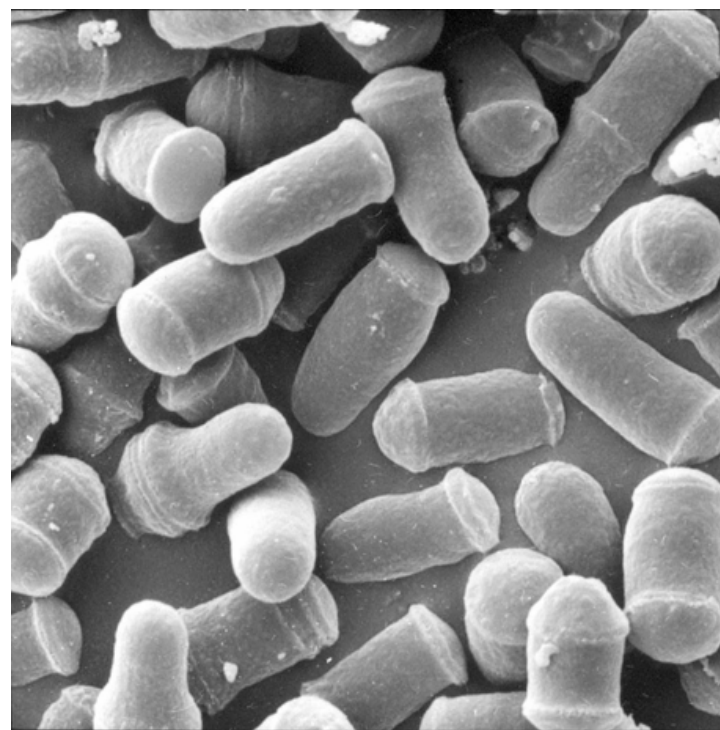




Saccharomyces cerevisiae

Размер генома 12.1 млн.п.о.

Число генов 5600



Schizosaccharomyces pombe

Размер генома 14 млн.п.о.

Число генов 4800