

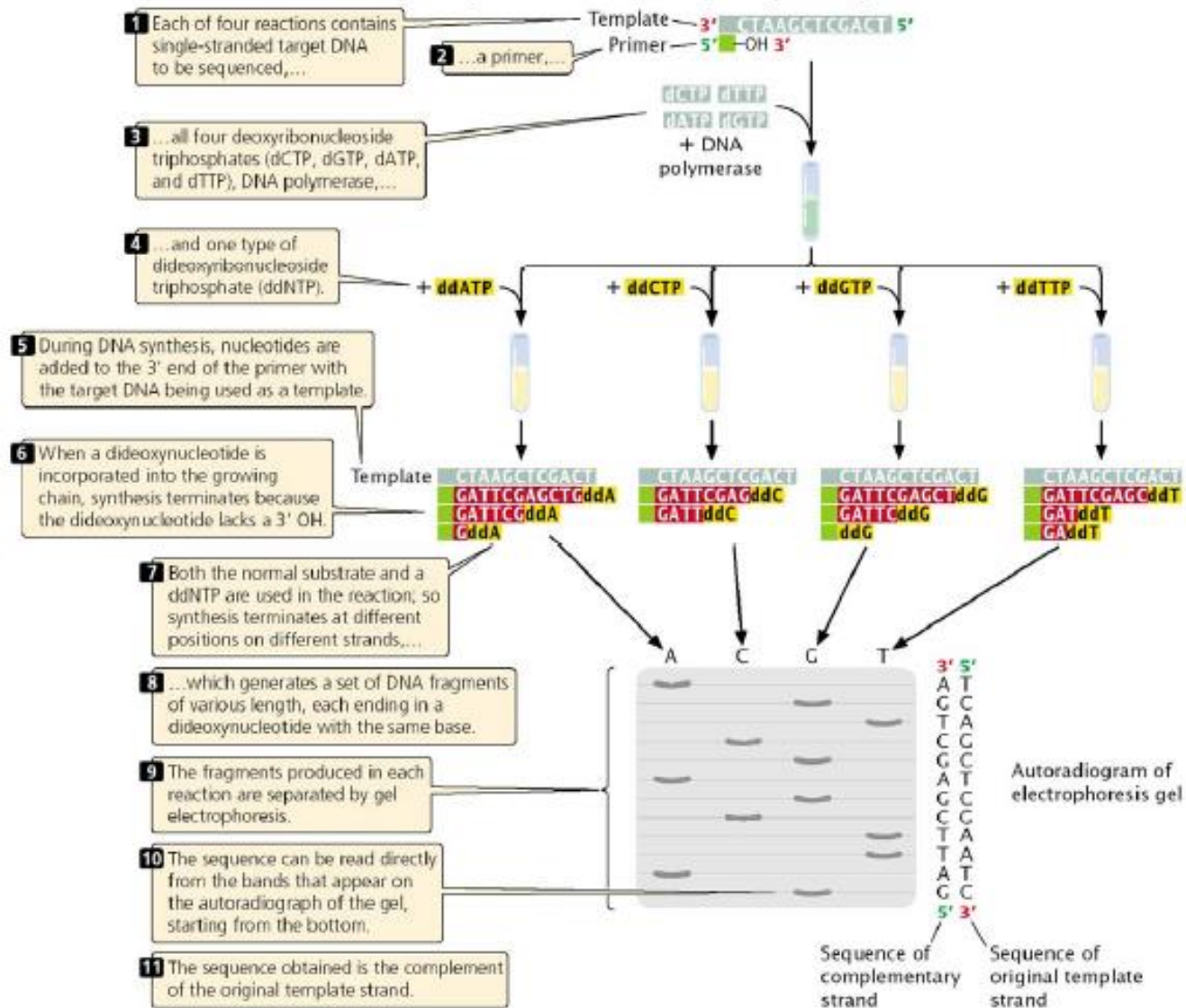
# Изучение эволюции геномов

# 1. Сравнение нуклеотидных последовательностей

Классификация и сравнительные характеристики  
моноклусных и мультилокусных ДНК-маркеров

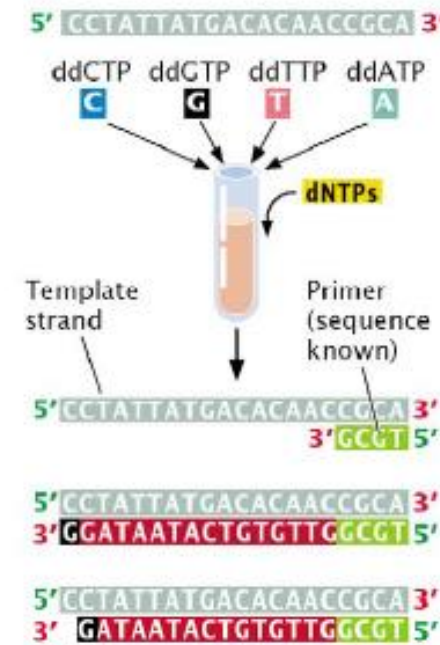
Классификация	Моноклусные маркеры	Мультилокусные маркеры
Классификация		
Методы, основанные на блот-гибридизации	<b>RFLP (restriction fragment length polymorphism)</b> – полиморфизм длины рестриционных фрагментов (Botstein <i>et al.</i> , 1980)	<b>Минисателлиты</b> (Jeffreys <i>et al.</i> , 1985)
Методы, основанные на ПЦР	<p><b>SSR (simple sequences repeats)</b> – простые повторяющиеся последовательности (микросателлиты) (Tautz, Renz, 1984)</p> <p><b>STS (sequences tagged site)</b> – последовательности, характеризующие locus (Olson <i>et al.</i>, 1989)</p> <p><b>SCAR (sequence characterized amplified region)</b> – последовательность, характеризующая амплифицированную область (Paran, Michelmore, 1993)</p> <p><b>SSCP (single strand conformation polymorphism)</b> – полиморфизм конформации одноцепочечной ДНК (Orita <i>et al.</i>, 1989)</p> <p><b>CAPS (cleaved amplified polymorphic sequences)</b> – расщепленные амплифицированные полиморфные последовательности (Konieczny, Ausubel, 1993)</p>	<p><b>RAPD (random amplified polymorphic DNA)</b> – случайно амплифицированная полиморфная ДНК (Welsh <i>et al.</i>, 1990; Williams <i>et al.</i>, 1990)</p> <p><b>ISSR (inter simple sequence repeats)</b> – межмикросателлитные последовательности (Zietkiewicz <i>et al.</i>, 1994)</p> <p><b>IRAP (inter-retrotransposon amplified polymorphism)</b> – полиморфизм амплифицированных последовательностей между ретротранспозонами (Kalendar, Schulman, 2006)</p> <p><b>AFLP (amplified fragment length polymorphism)</b> – полиморфизм длины амплифицированных фрагментов (Vos <i>et al.</i>, 1995)</p> <p><b>SSAP (sequence-specific amplification polymorphism)</b> – полиморфизм специфично амплифицированных последовательностей (Waugh <i>et al.</i>, 1997)</p>
Методы, основанные на применении ДНК-чипов	<b>SNP (single-nucleotide polymorphism)</b> – однонуклеотидный полиморфизм (Wang <i>et al.</i> , 1998)	<b>DArT (diversity array technology)</b> – ДНК-чип технология для изучения разнообразия (Jaccoud <i>et al.</i> , 2001)

# Метод секвенирования ДНК, основанный на терминеции синтеза дидезоксинуклеотидтрифосфатами



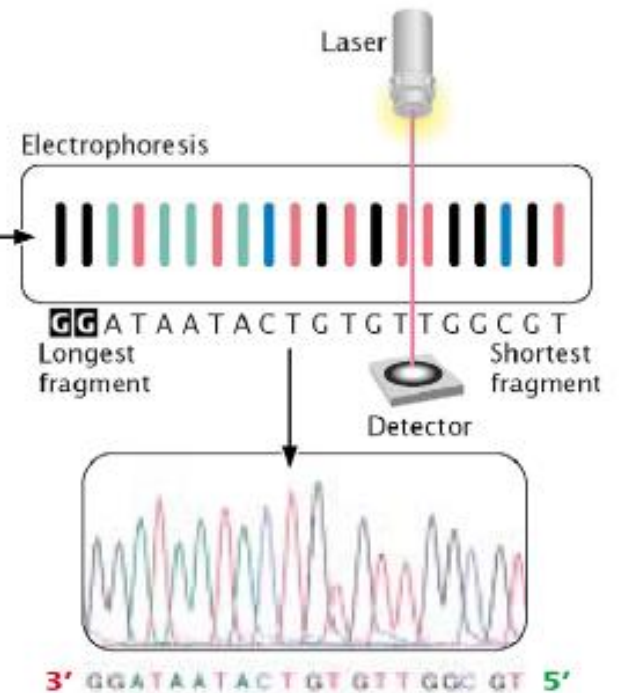
19.6 The dideoxy method of DNA sequencing is based on the termination of DNA synthesis.

- 1 A single-stranded DNA fragment whose base sequence is to be determined (the template) is isolated.
- 2 Each of the four ddNTPs is tagged with a fluorescent dye, and the Sanger sequencing reaction is carried out.
- 3 The fragments that end in the same base have the same colored dye attached.
- 4 The products are denatured, and the DNA fragments produced by the four reactions are mixed and loaded into a single well on an electrophoresis gel. The fragments migrate through the gel according to size,...
- 5 ...and the fluorescent dye on the DNA is detected by using a laser beam and detector.
- 6 Each fragment appears as a peak on the computer printout; the color of the peak indicates which base the peak represents.
- 7 The sequence information is read directly into the computer, which converts it into the complementary—target—sequence.



**Терминированные фрагменты ДНК фракционируются электрофоретически**

**Автоматизация секвенирования: каждый из четырех ddNTP имеет флюоресцентную метку своего цвета**



**Фрагмент каждого размера имеет цвет, соответствующий одному из четырех ddNTP**

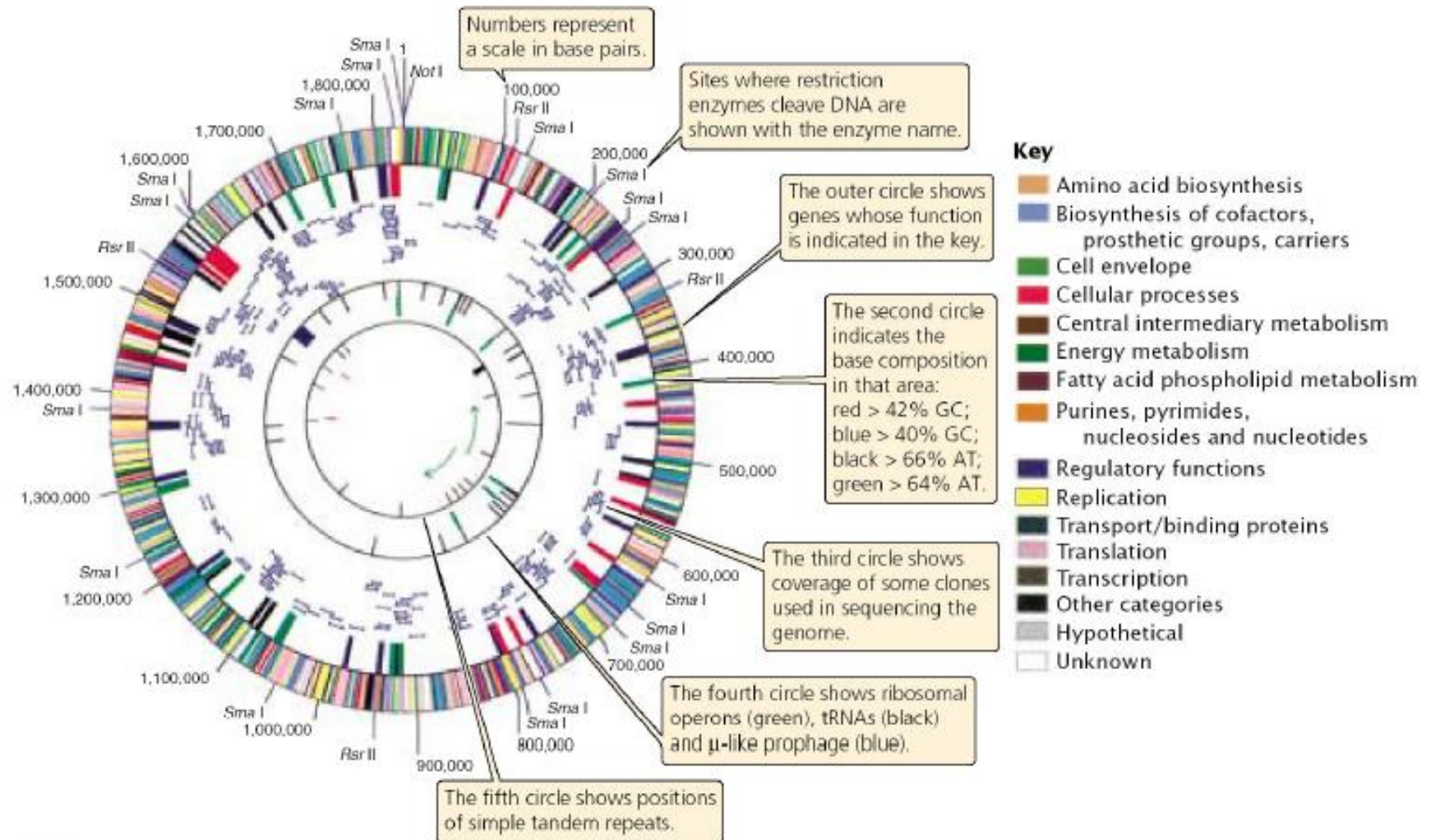
**19.8 The dideoxy sequencing method can be automated. СООТВЕТСТВУЮЩИЙ ОДНОМУ ИЗ ЧЕТЫРЕХ**

**ddNTP**



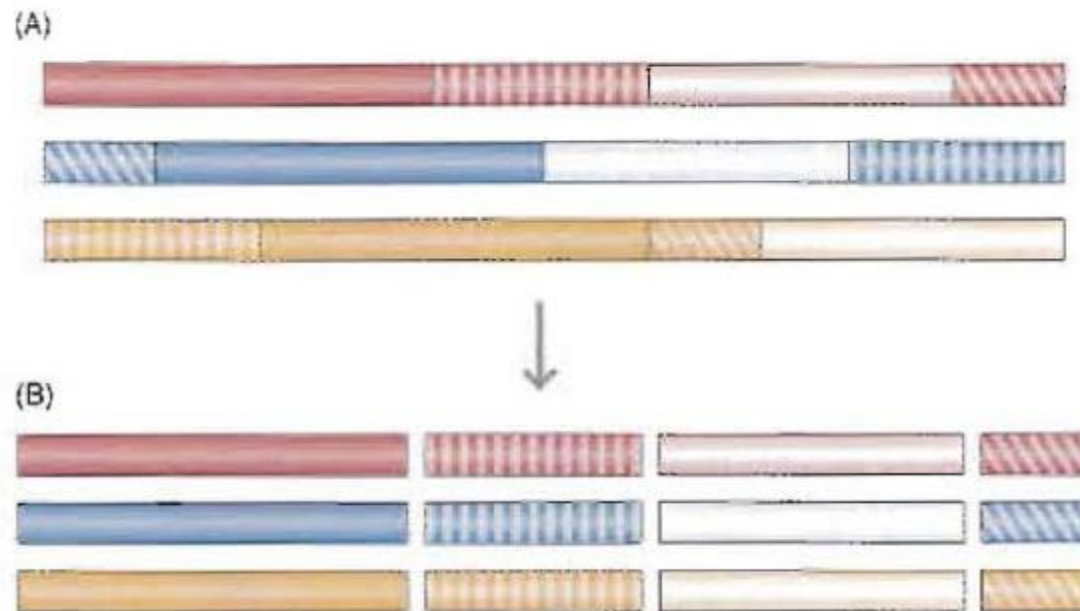


# Бактерия *H.influenzae* была первым свободно живущим организмом, геном которого был секвенирован (TIGR, США)



**19.9** The bacterium *Haemophilus influenzae* was the first free-living organism to be sequenced. (From R.D. Fleischman et al., 1993, *Science* 269:469; Scan courtesy of TIGR.)

# Выравнивание последовательностей геномов



Основная задача – нахождение минимального набора фрагментов последовательностей, которые гомологичны и коллинеарны друг другу

- Поиск сегментов, произошедших от общего предка
- Отслеживание эволюционных событий в геноме



# Сравнение гомологичных последовательностей

- Сходные участки – поддерживающий отбор – функционально важные
- Различающиеся участки – направленный отбор – определяют различия видов  
...или отсутствие отбора

# Purifying (negative) selection

(A)

species A	ATG	CTG	GAG	ACT	GGA	TGG	ATC	→	M	L	E	T	G	W	I
species B	ATG	CTG	GAA	ACC	GGG	TGG	ATT		M	L	E	T	G	W	I
species C	ATG	CTC	GAC	ACT	AGA	TGG	ATA		M	L	D	T	R	W	I

(B)

species A	GTTGGC-CCA	ACTGAC	→	GUUGG	C	C
species B	GTCGGG-CCT	TCCGGT	→	GUCGG	G	C
species C	GTTGCCACCA	TGTAAC	→	UGGCC	U	U
				GUUGC	C	A
				CAAUG	U	A

(C)

species A	ATGAATATT---	TTGGCCAT
species B	ATTA--ATCATCT	TGACCAT
species C	ATTCATAAAATATT--	CCAT

→ Стабилизация последовательности

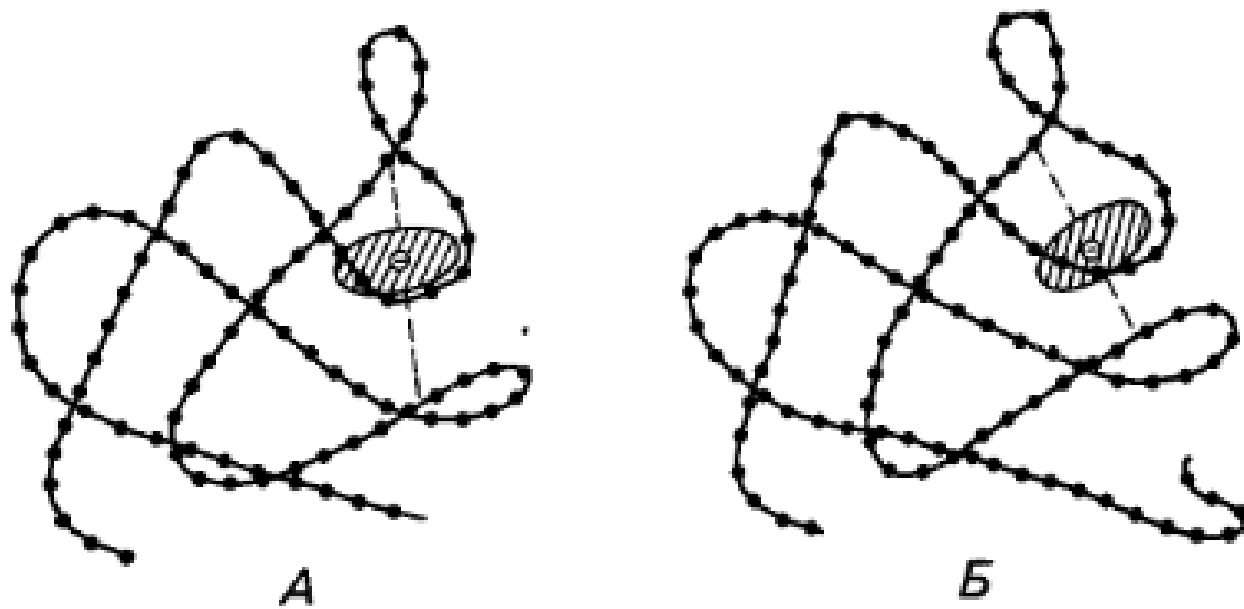


Рис. 19.5. Складывание полипептидных цепей (по Е. Пукеркандлю, 1966): *A* —  $\alpha$ -гемоглобина человека; *B* — миоглобина кита

# Гены эмбрионального развития крайне консервативны

(A)



Мутация *apterous*

(B)



Замена мутантной копии гена нормальной

(C)

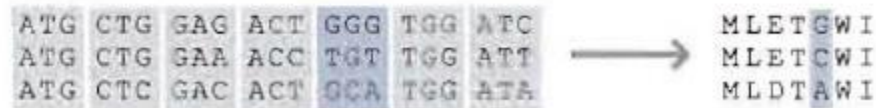


Замена мутантной копии гена ортологом из генома человека

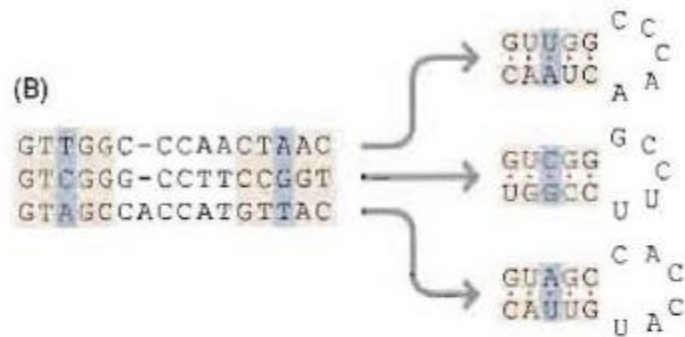


# Positive selection

(A)



(B)



→ Адаптивная эволюция



# Реаннотация геномов

- Валидация предсказанных генов

```

SC ATGACTCTCCTACTTCAAA -AATGTATCAGAGTAGAACTGC -GGAAAGATCCAATATTTA
SP ACGACTCTCCTACCACAAAATGATGCGTCGGAAAAGAACTGC TGGCAAATCCAATGTTTG
SM TTGATTTCACTTACCACAAACTGTATCAAAAAGGAAAAGT TAAACAAACA TAA TGTTC
SB -----GAACTTTTCAACAAACGTAGTT TAACTTGGTGCCGTCTTAA
      * * * * *
    
```

```

SC ACGCTTCT -TCTCACTCGAAATATTTCTGTCAAGTTT TAGAATCTGTTATATAT -CCCTT
SP ACACTGCC -TCTGGACCTGAAATGTCCCGTCTATTCCAGAAGCTCTTATAGCTCCCTT
SM ATATTTGTCTCTGGTATTCACCTGGATC ACCTGTCTTGGAACTCTTATATCT -TCTC
SB GATGTT - -AAGCCAATCTAAATCTTCCGGTGTGTTTGGTCTCGAGAATGTTGTTATCG
      * * * * *
    
```

```

SC TAAAGGAAACTTGTATCTCTAGATTTTCATCTT -GTC -GATCTTCAAACAACTGTTA-
SP CAAAGGAAATTTCTGATATTCATA TCTCATCTTATC -AATGTTCAAACAGCGTGTTA
SM CAAAGTAGATCTGATGTCACTAGA TCTCATTATGTCBAATGATTGAACAATGATTTCG
SB CATTGCTCAAACAAAGTCTATATTTGTACCTATCTC -AATGGCAGGGCAACAAGTTGT
      * * * * *
    
```

```

SC - - - -TCAGAGAAACAGCTCTATGAGAAATCAGCTGATG -TCAGTATCTGTCACTCACT
SP CAGACACAGAGAAACAGCTTTCATGAGAAATCAGCCCGTG -TCAGTCTCTGCATCCGCT
SM CAGACACCAGTTAAACAACAGCAAGATAAATTGAGGGACG - - -GTGCTCTACTGCCCTGCT
SB AAG CATGAGCAGCACTAGCGAAAGGAACCCCTGAAATATAGTACTCGGTAATCTTCA
      * * * * *
    
```

```

SC TCGTGAATTTACAT -TGACTGCTCTTTCTGTGCAATTCATCTGTATTATTATAGTTGCAA
SP CAGTGAACCCATGTATGACTGCTCTTTCTGTGCAATTCATCTGTA - - -TTC TAGCCGCAA
SM CAGATAACTCATCG -TGACTGCTCTTCTGTACCGTCCAACCTGTATTCTCTTCTTAA
SB CATACTTCTTAC -CAGCTGTTCTCTGTGTGGGCCAATCTGTATCTGACTATCATCTA
      * * * * *
    
```

```

SC CTAGATGGCAAAAATT - -GATAATGCCCTAAAAATGACAGGCCG TAA
SP TTAACGACAGAAAATT TATAATGCCCTAATGATGACAGGTTGATAA
SM TTAGAAGACGGGAAGTT - -GATAATGCCCTAATGACGACAGGGCATAAA
SB CTAGCC TCAAAGAATT - -GATAATGCCCTAACGGCGACAGGCCACAA
      * * * * *
    
```

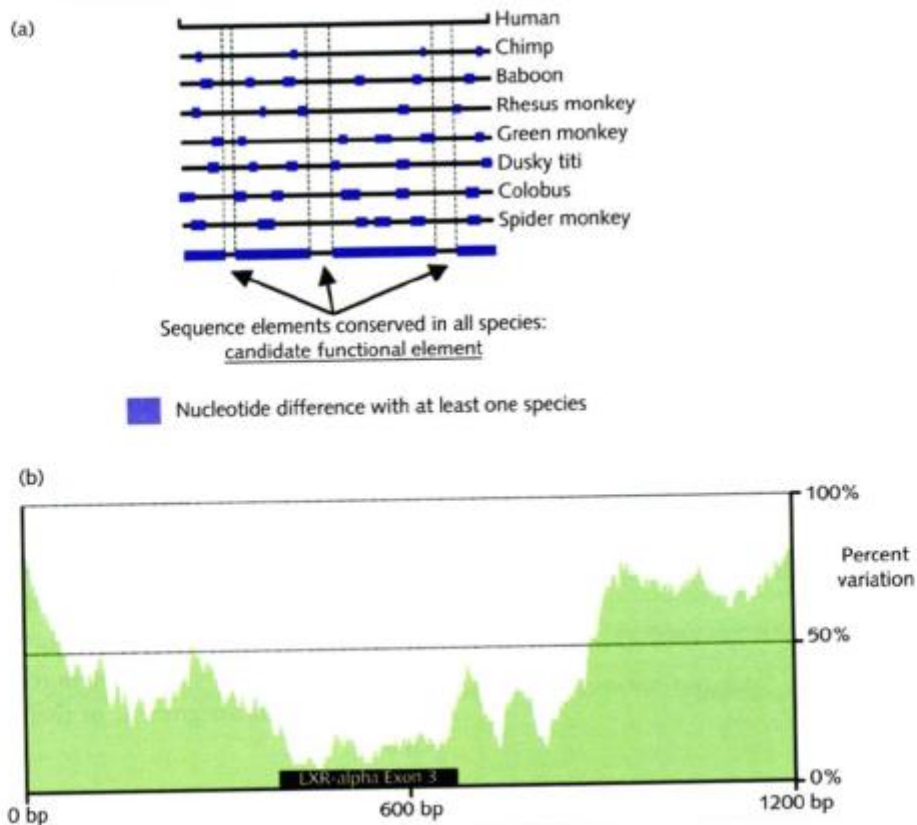
TAA - start  
TATAATGCCCTAATGATGACAGGTTGATAA - stop  
▲ - ins/del

SC – *Saccharomyces cerevisiae*  
 SP – *S. paradoxus*  
 SM – *S. mikatae*  
 SB – *S. bayanus*

# Реаннотация геномов

- Идентификация новых генов

Выявление негативной селекции на некодирующих участках позволило обнаружить ряд генов малых РНК

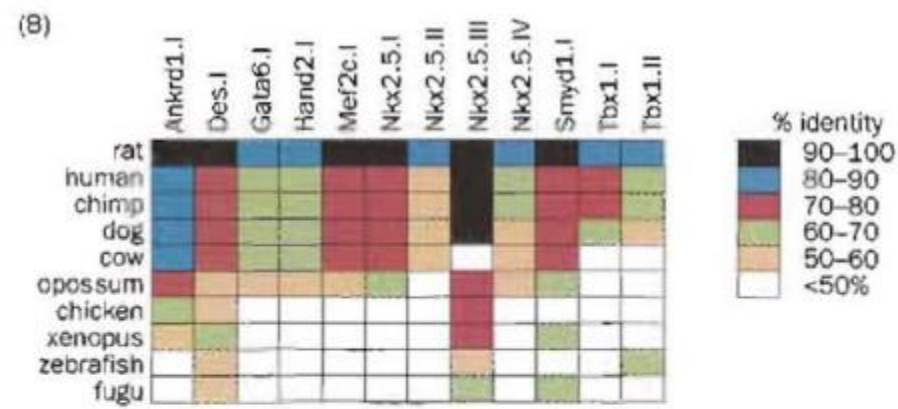
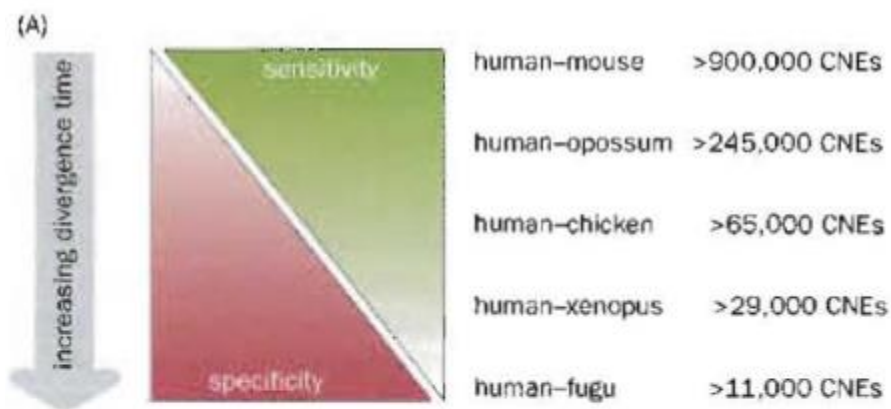




# Реаннотация геномов

- Идентификация регуляторных последовательностей

CNE – conserved noncoding element



Могут быть как консервативные, так и переменные регуляторные элементы

# Реаннотация геномов

## • Валидация предсказанных генов

SC ATGACTCTCCTACTTCAAA-AATGTATCAGAGTAGAACTGC-GGAAAGATCCAATATTTA  
 SP ACGACTCTCCTACCACAAATAATGCGTCGGAAAAGAAGCTGGGCAAATCCAATGPTTG  
 SM TTGATTTCACTAACCCAAACACTGTATCAAAAAGGAAAAGTAAACAAACATAATGTTC  
 SB -----GAACTTTTCAACAAACGTAGTTAAACTTGGTGCCGCTTAA  
 \* \* \* \* \* \* \* \* \* \* \* \* \*

SC ACGCTTCT-TCTTCACTCGAAATATTTCTGTCTAGTTTITAGAATCTGTTATATAT-CCCTT  
 SP ACACTGCC-TCTGGACTGAAATGTCCCGCTCTATTCCAGAAGCTCTTATAGCTCCCTT  
 SM ATATTTGTCTCTGGTATTCACCTGGATCACCTGTCTTGGAAATCTCTTATATCT-TCTC  
 SB GATGTT--AAGCCAATCTAATTCTCCGGTGTGTTTGGTCTCTCGAGAATGTTGATCG  
 \* \* \* \* \* \* \* \* \* \* \* \* \*

SC TAAAGGAAACTTGATACTTCTAGATTTTCATCTT-GTC-GATCTTCAAACAACGTGTTA-  
 SP CAAAGGAAATTCGTATATTCATA-TCTCATCTTATC-AATGTTCAAACAGCGTGTTA-  
 SM CAAAGTAGATCTGATGTCACTAGA-TCTCATTTATGTCAAATGATTGAACAATGATTTCG  
 SB CATTGCTCAAACAAGTCTCTATATTTGTACCTATCTC-AATGGCAGGGCAACAAGTTGT  
 \* \* \* \* \* \* \* \* \* \* \* \* \*

SC ----TCAGAGAAACAGCTCTATGAGAAATCAGCTGATG-TCAGTATTCTGTCACTCACT  
 SP CAGACACAGAGAAACAGCTTCATGAGAAGTCAGCCGGTG-TCAGTGCTCTGACATCCGCT  
 SM CAGACACCAGTAAACAACAGCAAGATAATTGAGGGACG----GTGCTCTACTGCCTGCT  
 SB AAGCATGAGCAGCAGTAGCGAAAGGAACCGCCTGAAAATAGTACTCGGTAATCTTCA  
 \* \* \* \* \* \* \* \* \* \* \* \* \*

SC TCGTGAATTTACAT-TGACTGTCTTTTCGTGCAATTCATACTGTATTATTATAGTTGCAA  
 SP CAGTGAACCCATGTATGACTGTCTTTTCGTGCAATTCATCTGTA--TTCTAGCCGCAA  
 SM CAGATAACTCATCG-TGACTGTCTTCTGTACGGTCCAACCTGTATTCTTTCTTAA  
 SB CATACTTCTTAC-CAGCTGTTCTCTCGTGTGGCCAATCTGTATCTGACTATCATCTA  
 \* \* \* \* \* \* \* \* \* \* \* \* \*

SC CTAGATGGCAAAAATT--GATAATGCCCTAATAATGACAGGCCGTAA  
 SP TTAAACGACAGAAAATTATATAATGCCCTAATGATGACAGGTTGATAA  
 SM TTAGAAGACGGGAAGTT--GATAATGCCCTATTGACGACAGGCCATAAA  
 SB CTAGCCTCAAAGAATT--GATAATGCCCTAACGGCGACAGGCCACAA  
 \*\* \* \* \* \* \* \* \* \* \* \* \* \* \*

□ - start  
 □ - stop  
 ▲ - in/del

SC – *Saccharomyces cerevisiae*  
 SP – *S. paradoxus*  
 SM – *S. mikatae*  
 SB – *S. bayanus*

## **2. Выявление тенденций в эволюции геномов и структурных генов**

# Сравнительные размеры геномов

organism	estimated size	estimated	average gene density	chromosome
		gene number		number
<i>Homo sapiens</i> (human)	2900 million bases	~30,000	1 gene per 100,000 bases	46
<i>Rattus norvegicus</i> (rat)				
<i>Mus musculus</i> (mouse)	2500 million bases	~30,000	1 gene per 100,000 bases	40
<i>Drosophila melanogaster</i> (fruit fly)				
<i>Arabidopsis thaliana</i> (plant)	180 million bases	13 600	1 gene per 9,000 bases	8
<i>Caenorhabditis elegans</i> (roundworm)	125 million bases	25 500	1 gene per 4000 bases	5
<i>Saccharomyces cerevisiae</i> (yeast)	97 million bases	19 100	1 gene per 5000 bases	6
<i>Escherichia coli</i> (bacteria)	12 million bases	6300	1 gene per 2000 bases	16
<i>H. influenzae</i> (bacteria)	4.7 million bases	3200	1 gene per 1400 bases	1
	1.8 million bases	1700	1 gene per 1000 bases	1



# Тенденции в эволюции генов

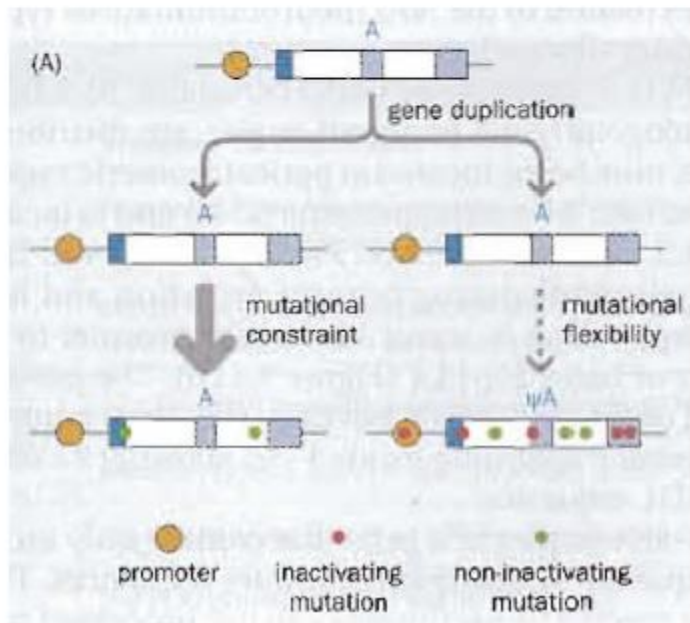
1. Автономизация генов
2. Олигомеризация генов
3. Появление мозаичной структуры гена
4. Появление перекрывающихся генов
5. Дупликации генов

# Усложнение генов

- Дупликация экзонов (неравный кроссинговер)
  - возможности для дивергенции последовательности и появления новых экзонов
  - возможности для альтернативного сплайсинга
  - увеличение структурного домена белка
  - вариабильность генов иммуноглобулинов
- Перетасовка экзонов (неаллельная рекомбинация, транспозоны)

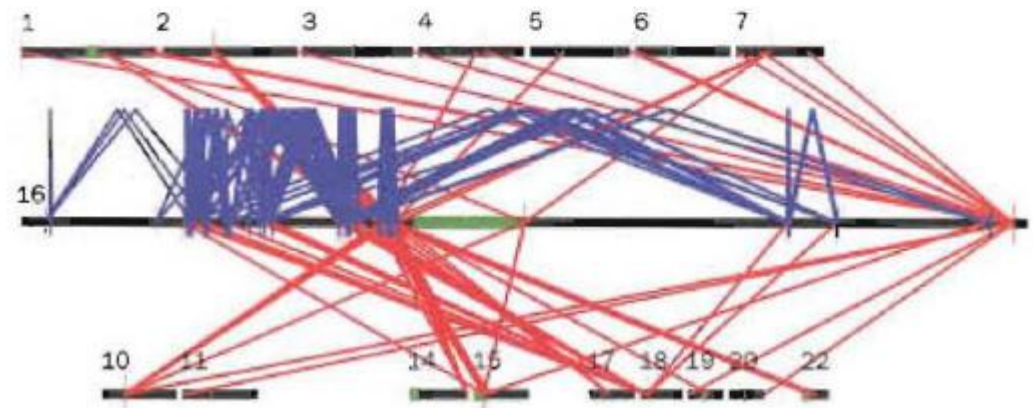
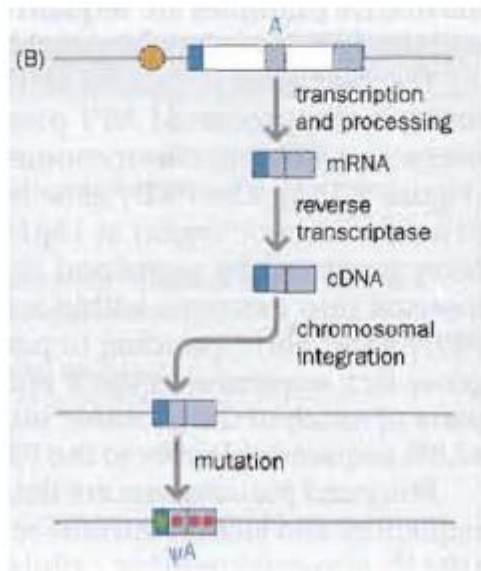
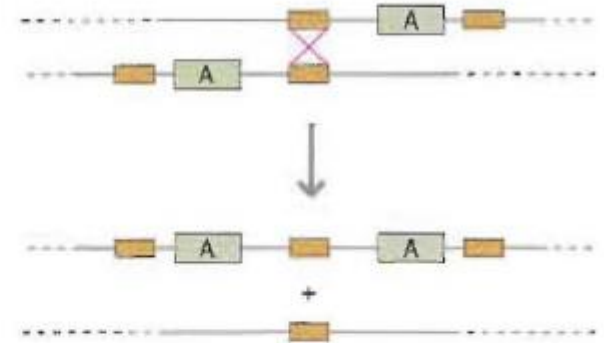
# Следствие дупликации гена

- Увеличение дозы гена
- Дивергенция гена
  - псевдогены



# Дупликация генов

- Тандемная дупликация
- Транспозиция  
(ретрогены – нет промотора, интронов)
- Сегментные дупликации
- Полногеномные дупликации

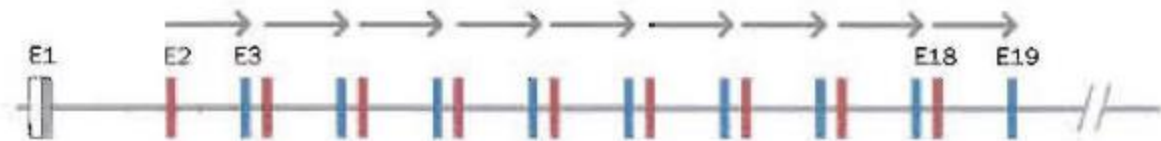


# Дупликация экзонов

(около 10% генов человека)

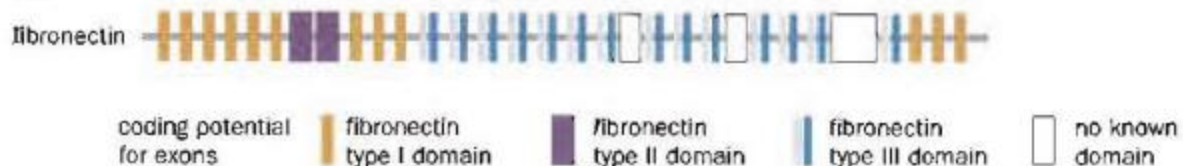
Ген *LPA*

(A)



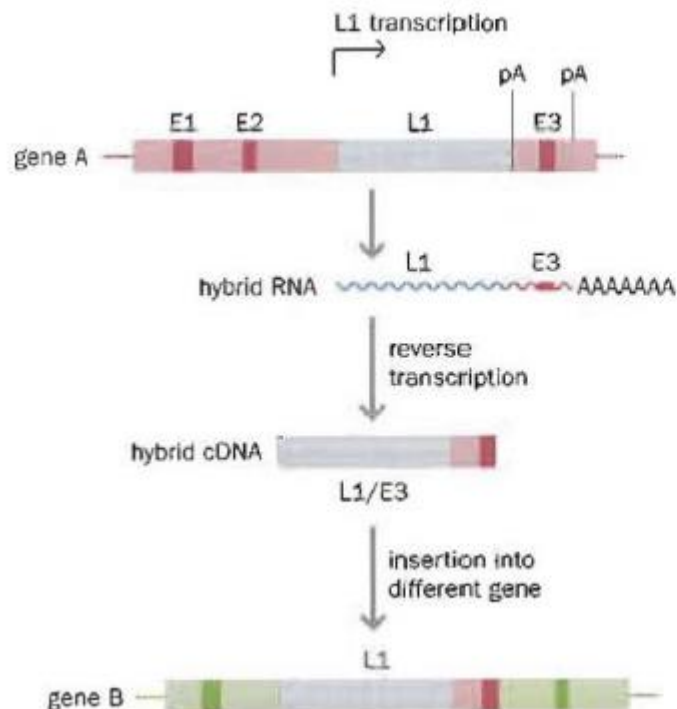
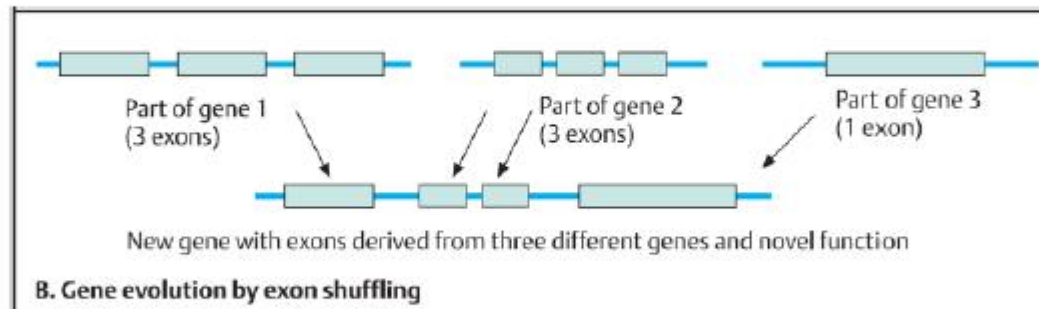
```
E2 ----- 17 AAPEQSHVVQDCYHGDGQSYRGTYSTTVTGRTCQAWSMTPHQHNRTTENYPN
E3 ----- 131 AGLIMNYCRNPDAVAAPYCYTRDPGVRWEYCNLTQCSDAEGTAVAPPTVTPVPSLEAPSEQ 130
          245 AAPEQSHVVQDCYHGDGQSYRGTYSTTVTGRTCQAWSMTPHQHNRTTENYPN
          244 AGLIMNYCRNPDAVAAPYCYTRDPGVRWEYCNLTQCSDAEGTAVAPPTVTPVPSLEAPSEQ 244
          359 AAPEQSHVVQDCYHGDGQSYRGTYSTTVTGRTCQAWSMTPHQHNRTTENYPN
          358 AGLIMNYCRNPDAVAAPYCYTRDPGVRWEYCNLTQCSDAEGTAVAPPTVTPVPSLEAPSEQ 358
          473 AAPEQSHVVQDCYHGDGQSYRGTYSTTVTGRTCQAWSMTPHQHNRTTENYPN
          472 AGLIMNYCRNPDAVAAPYCYTRDPGVRWEYCNLTQCSDAEGTAVAPPTVTPVPSLEAPSEQ 472
          587 AAPEQSHVVQDCYHGDGQSYRGTYSTTVTGRTCQAWSMTPHQHNRTTENYPN
          586 AGLIMNYCRNPDAVAAPYCYTRDPGVRWEYCNLTQCSDAEGTAVAPPTVTPVPSLEAPSEQ 586
          701 AAPEQSHVVQDCYHGDGQSYRGTYSTTVTGRTCQAWSMTPHQHNRTTENYPN
          700 AGLIMNYCRNPDAVAAPYCYTRDPGVRWEYCNLTQCSDAEGTAVAPPTVTPVPSLEAPSEQ 700
          815 AAPEQSHVVQDCYHGDGQSYRGTYSTTVTGRTCQAWSMTPHQHNRTTENYPN
          814 AGLIMNYCRNPDAVAAPYCYTRDPGVRWEYCNLTQCSDAEGTAVAPPTVTPVPSLEAPSEQ 814
          929 AAPEQSHVVQDCYHGDGQSYRGTYSTTVTGRTCQAWSMTPHQHNRTTENYPN
E18 ----- 928 AGLIMNYCRNPDAVAAPYCYTRDPGVRWEYCNLTQCSDAEGTAVAPPTVTPVPSLEAPSEQ 928
E19 ----- 1042 AGLIMNYCRNPDAVAAPYCYTRDPGVRWEYCNLTQCSDAEGTAVAPPTVTPVPSLEAPSEQ 1042
```

(B)





# Перетасовка экзонов



Ретротранспозиция

# Гомологи

- Гены, обладающие эволюционным родством и сходными последовательностями
- Паралоги – гомологичные гены в одном геноме, возникшие в результате дупликации
- Ортологи – гомологичные гены разных видов, произошедшие от общего предшественника и передавались вертикально при эволюции

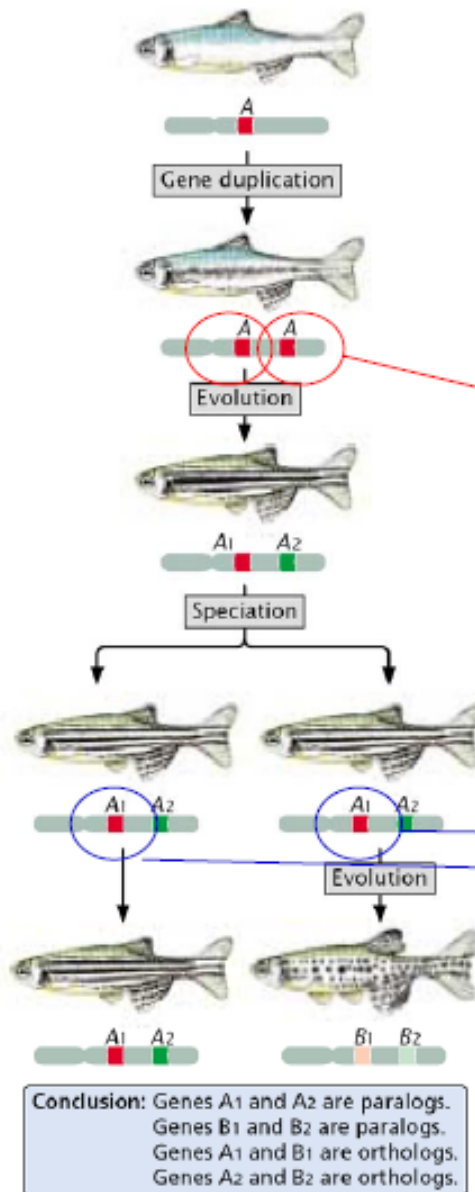
Гомологичные последовательности родственны по происхождению.

Для генома человека две последовательности могут быть общими не по происхождению, а по причинам случайного характера, если длина последовательности менее 15 нуклеотидов.

$4^{15} = 10^9 =$  длина генома человека в нуклеотидах.

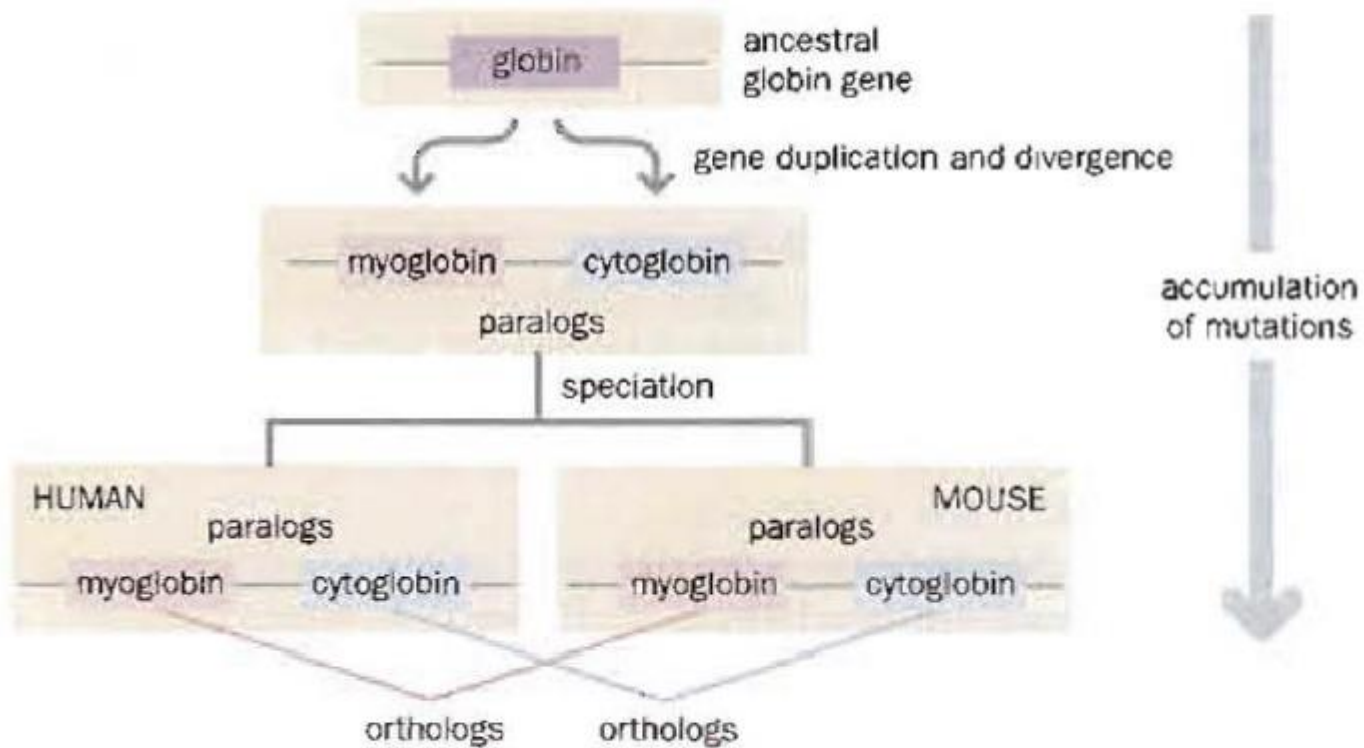
Гомологичные последовательности нуклеотидов в геноме одного вида называются паралоги (A1 и A2; B1 и B2). Появление двух генов позволяет одному из них меняться, приобретая измененную функцию

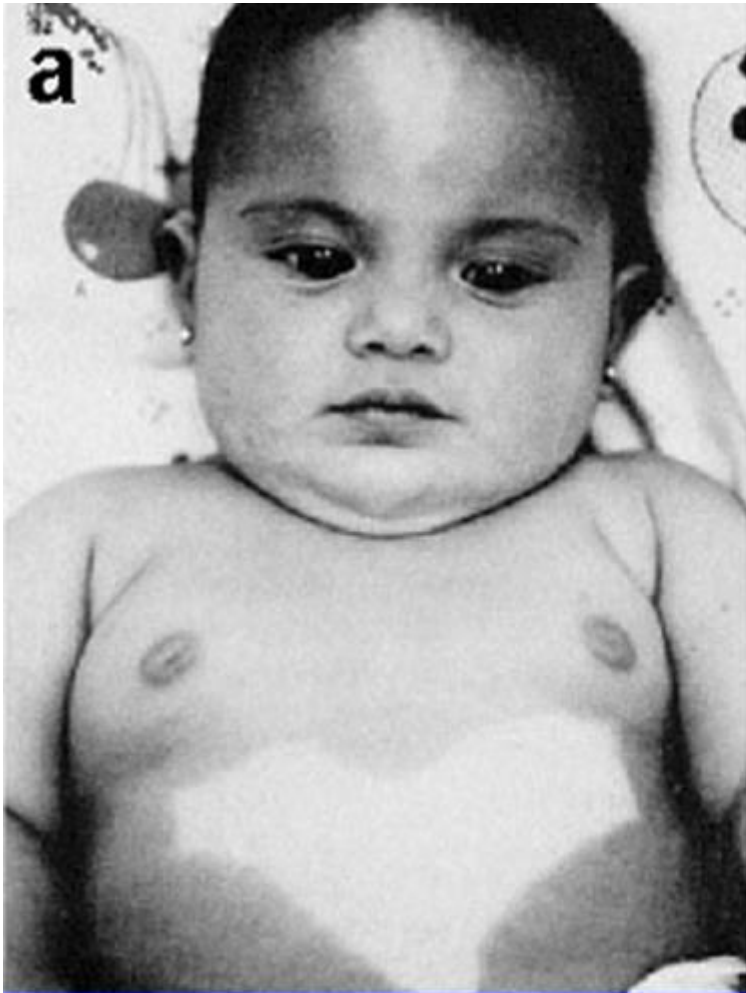
Гомологичные последовательности нуклеотидов в геномах разных видов называются ортологи (A1 и B1; A2 и B2). Знание гомологичного (ортологичного) гена у одного вида, позволяет изучать его функцию на другом, виде – модельном, более удобном для работы



**19.15 Homologous sequences are evolutionarily related.** Orthologs are homologous sequences found in different species; paralogs are homologous genes in the same species that arise from gene duplication.

# Ортологи и паралоги

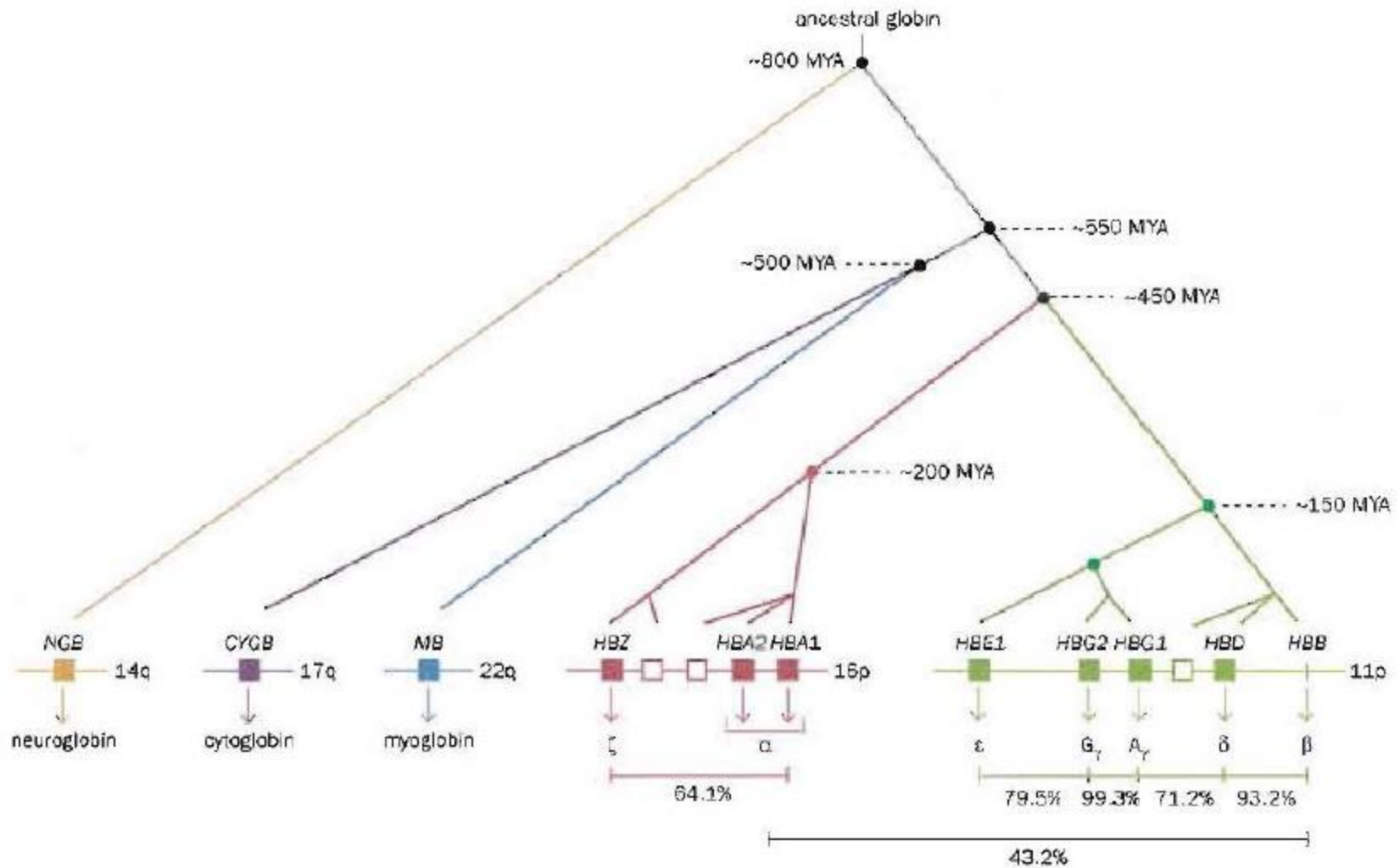




**Mutations in the same gene can result in strikingly similar consequences for humans and mice**



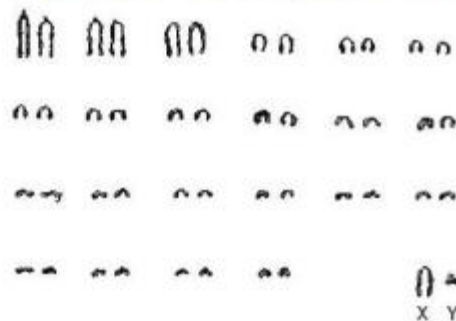
# Эволюция суперсемейства глобинов



# Хромосомные перестройки

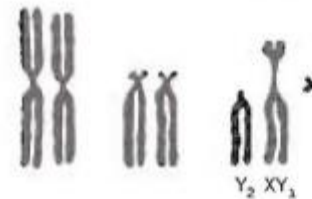
*Muntiacus reevesi*

(A)



*Mutiacus muntijak*

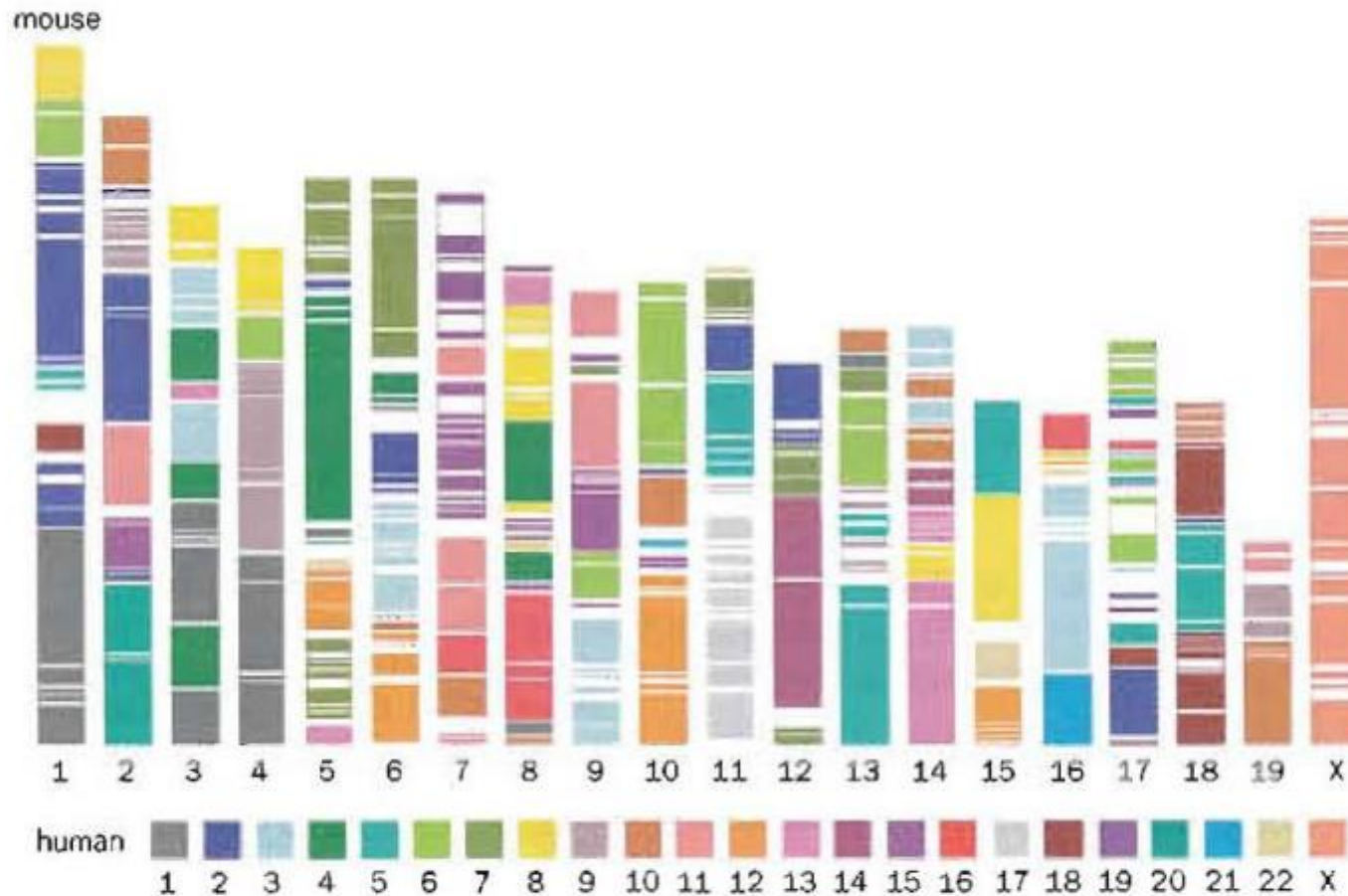
(B)



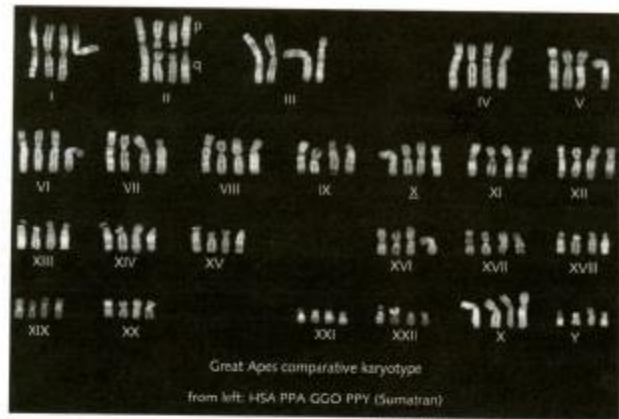
# Хромосомные перестройки

- Инверсии, транслокации характерны для эволюции млекопитающих
- Формирование неоцентромеры – изменение положение центромеры в хромосоме

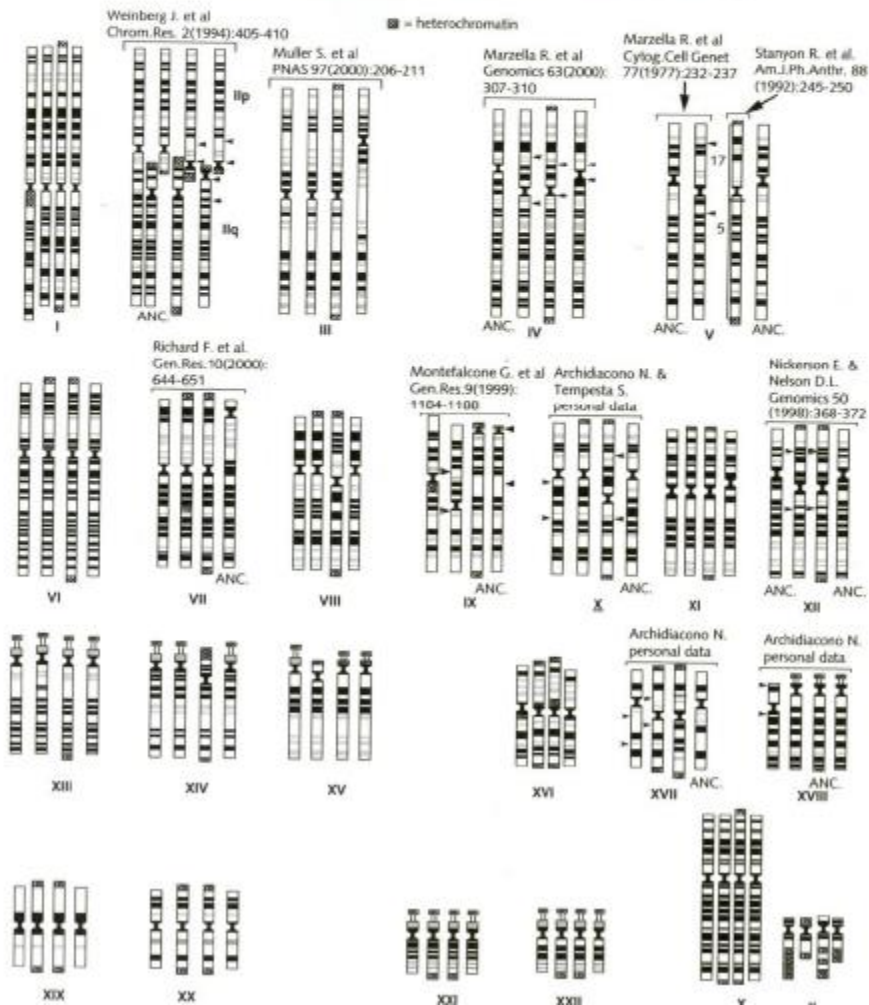
# Хромосомные перестройки



Средний размер синтенных участков хромосом человека и мыши – 10 Мб  
Межхромосомные перестройки не затрагивают половые хромосомы



- человек
- шимпанзе
- горилла
- орангутанг





# Определение пола

- Генное
- Хромосомное
  - самки гомогаметны, самцы гетерогаметны (XY или XO)
  - самцы гетерогаметны, самки гомогаметны (ZW или ZO)
- Гапло-диплоидное
- Средовое

System	Sex chromosome constitutions	Distribution/examples	Comments
X-Y	XY (male); XX (female)	prevalent in mammals, but also in some amphibians and in some fish <sup>b</sup>	Y chromosome carries male-determining factor; sperm are X or Y and so determine the sex
Z-W	ZW (female); ZZ (male)	prevalent in birds and reptiles, but also in some amphibians and in some fish <sup>b</sup>	eggs are Z or W and so determine the sex
X-autosome ratio	XX (female/hermaphrodite); X0 <sup>a</sup> /XY (male)	in many types of insect and some nematodes	<i>C. elegans</i> is an example in which XX individuals are hermaphrodites; <i>D. melanogaster</i> has a Y chromosome but it does not confer maleness
Haplo-diploid	relies on ploidy only	in many types of insect	unfertilized eggs produce haploid fertile males; fertilized eggs are diploid and give rise to females or sterile males

<sup>a</sup>The 0 denotes zero, so that X0 means a single X. <sup>b</sup>There can be different systems in the same kind of organism. For example, some medaka species have the X-Y system and others have the Z-W system.

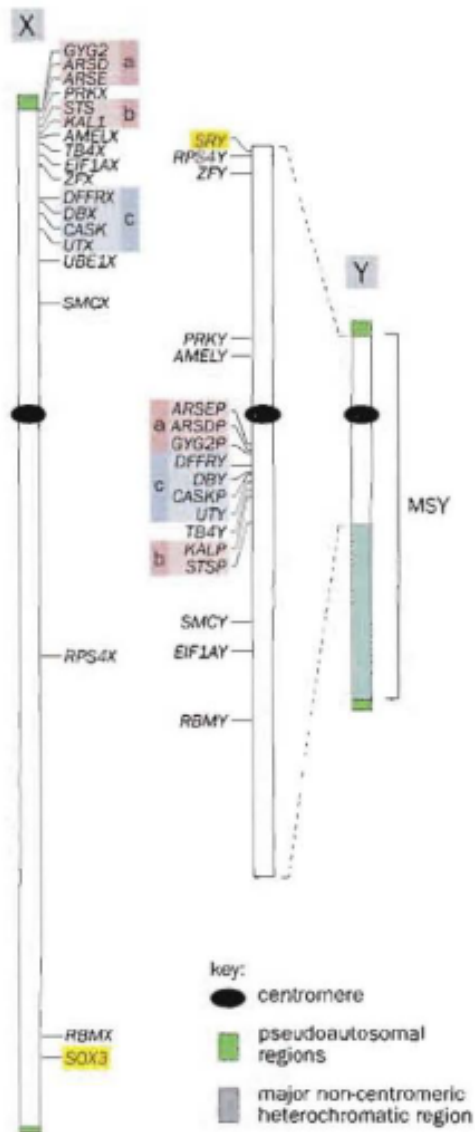
# Гетерогаметный пол

Хромосома определяющая пол (Y, W), как правило, отличается маленьким размером, малым количеством генов, большим количеством повторов

Человек:

- X-хромосома – 900 белок-кодирующих генов, 155 Mb
- Y-хромосома – 80 белок-кодирующих генов (30 уникальных), 59 Mb
- Рекомбинация между X и Y в мейозе происходит только в идентичных прителомерных областях (псевдоаутосомные регионы)

# X и Y хромосомы человека



MSY – male-specific region on the Y chromosome.

Gametologs – гомологичные гены X и Y хромосом вне псевдоаутосомных регионов. В Y хромосоме несколько таких генов стало псевдогенами.

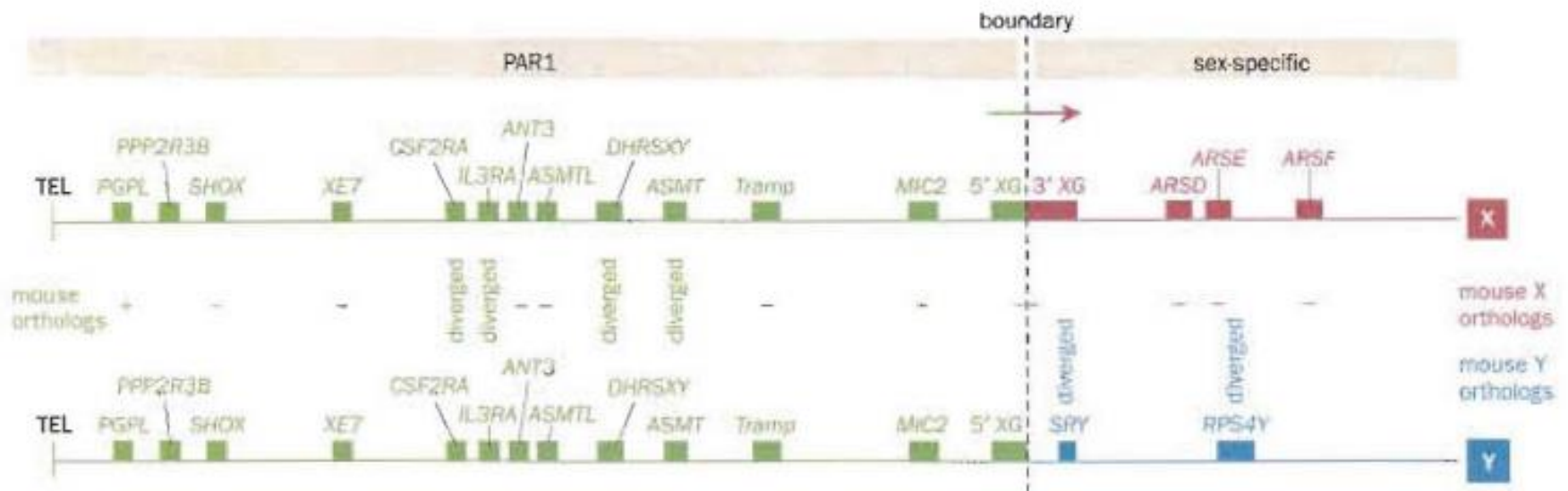
*SRY* и *SOX3* изначально произошли от одного гена, но отличаются довольно сильно.

# Pseudoautosomal regions (PAR1, PAR2)

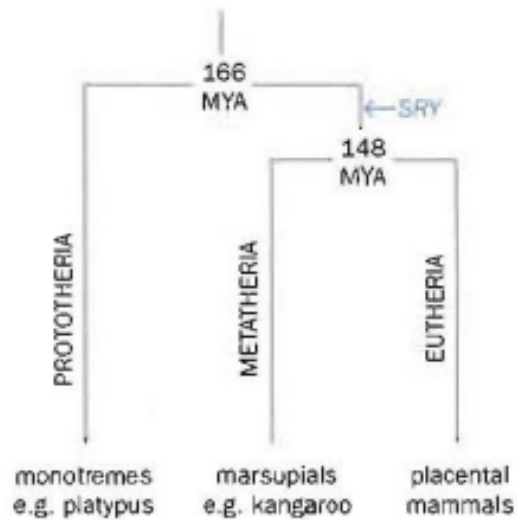
PAR1 – 2,6 Mb

PAR2 – 330 kb

PAR1 – высокая скорость эволюции

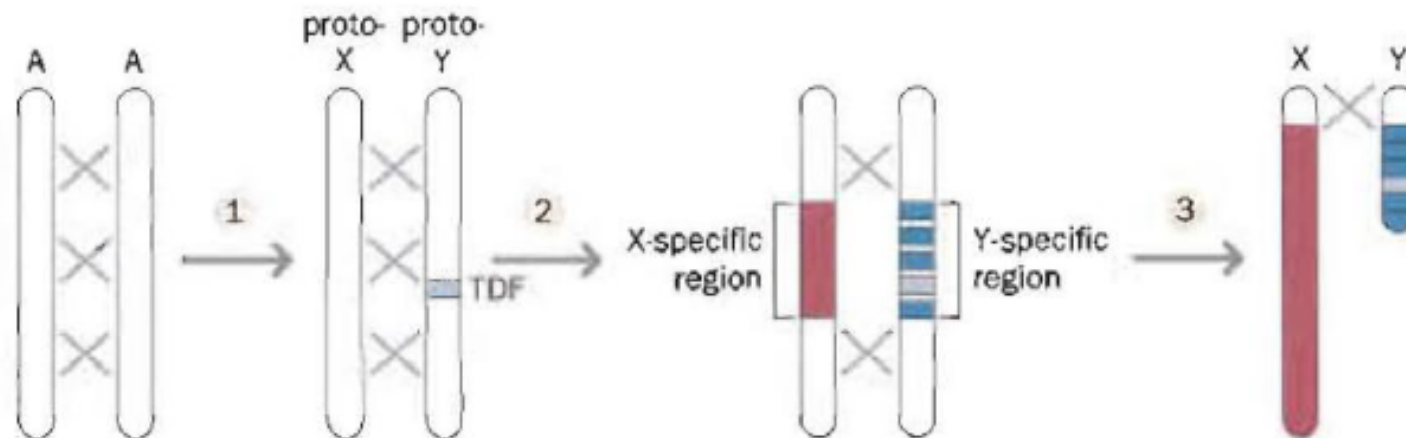


# X и Y хромосомы произошли от пары аутосом



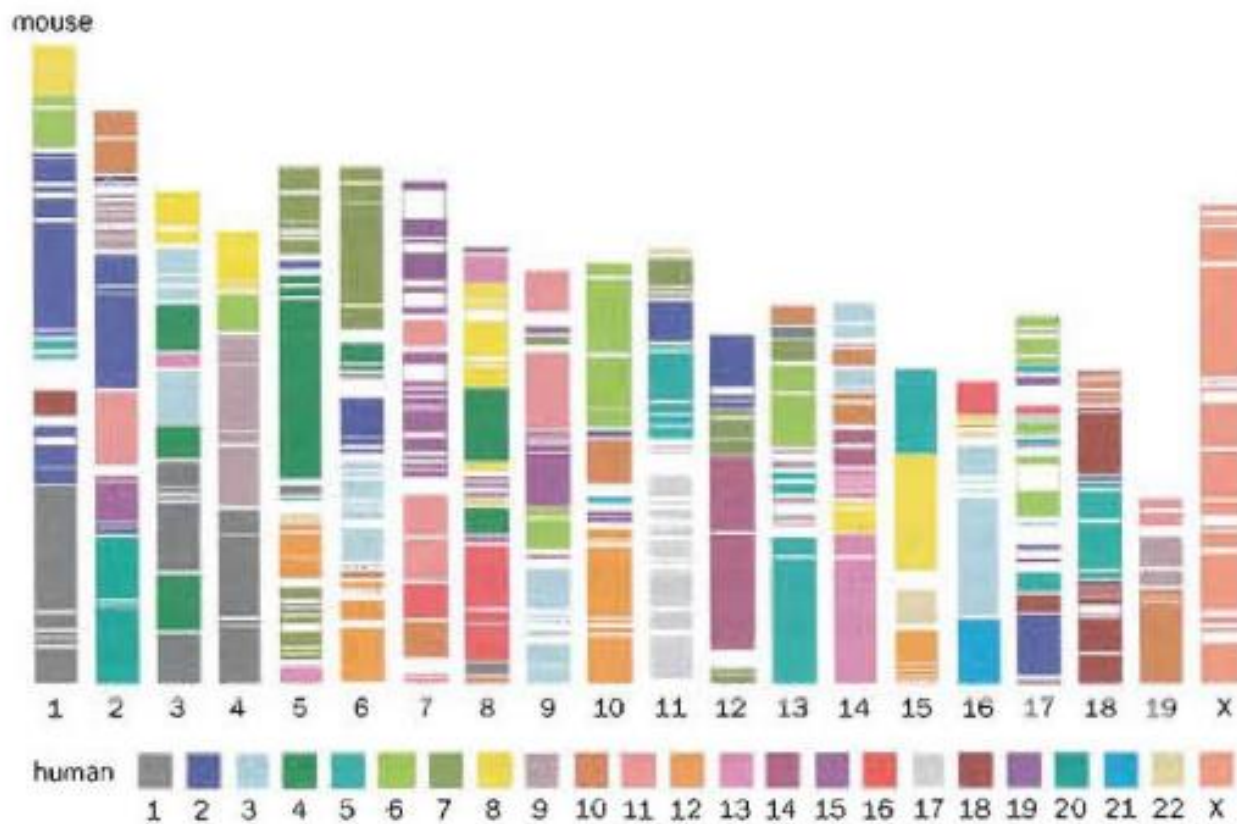
1. Появление testis-determining factor (TDF)
2. Рекомбинация подавляется инверсией
3. Дальнейшая дивергенция из-за отсутствия рекомбинации и накопления мутаций

Ген *SRY* произошел от гена *SOX3*

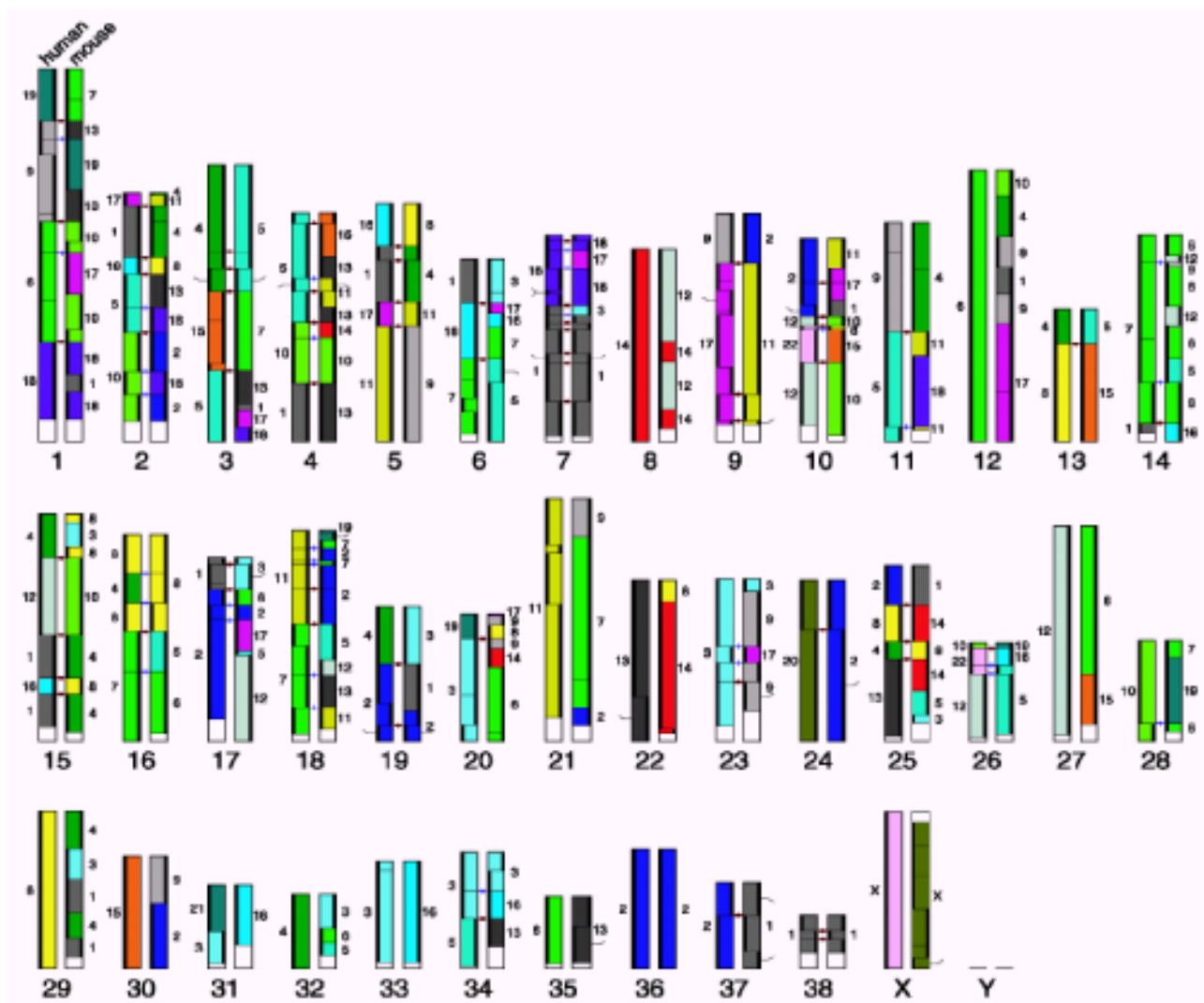




# Межхромосомные перестройки не характерны для половых хромосом



## Сравнение хромосом собаки, человека и мыши



# **3. Изучение регуляторных механизмов работы генома**

# Усложнение на уровне регуляции

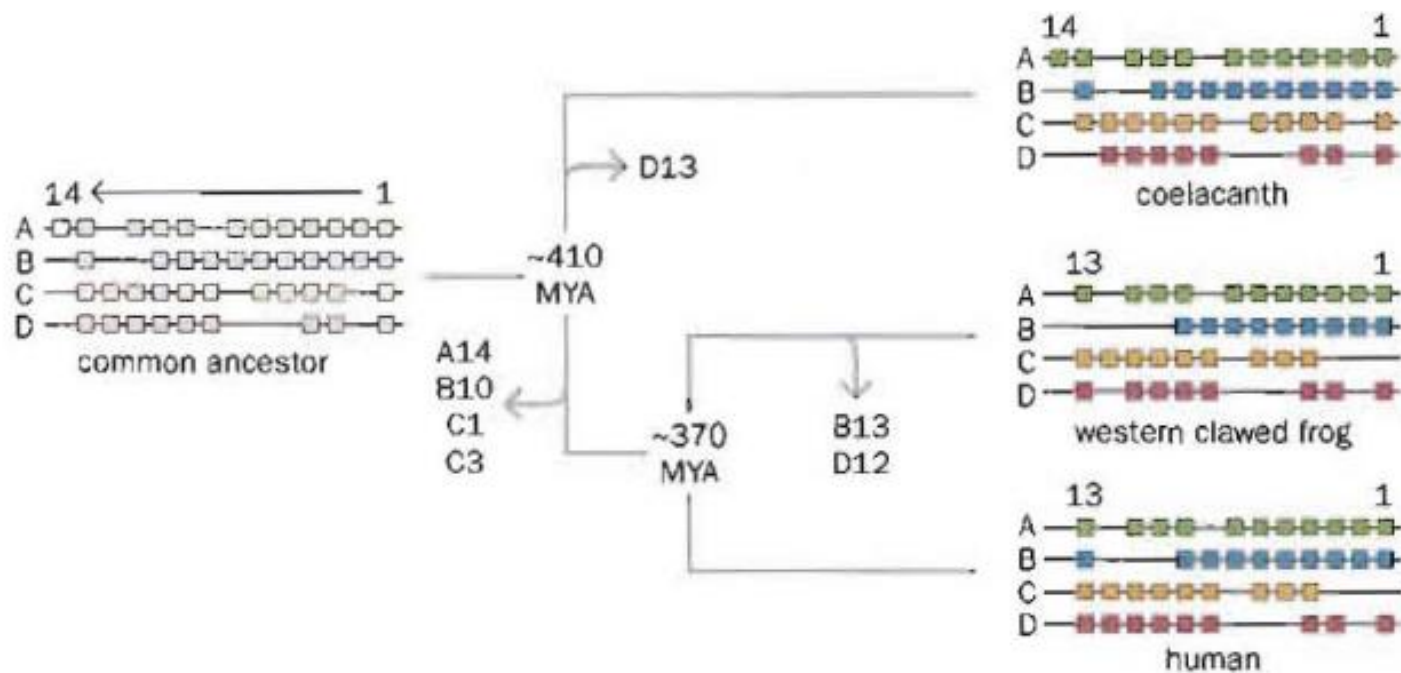
- Альтернативный сплайсинг
- Некодирующая (регуляторная) ДНК





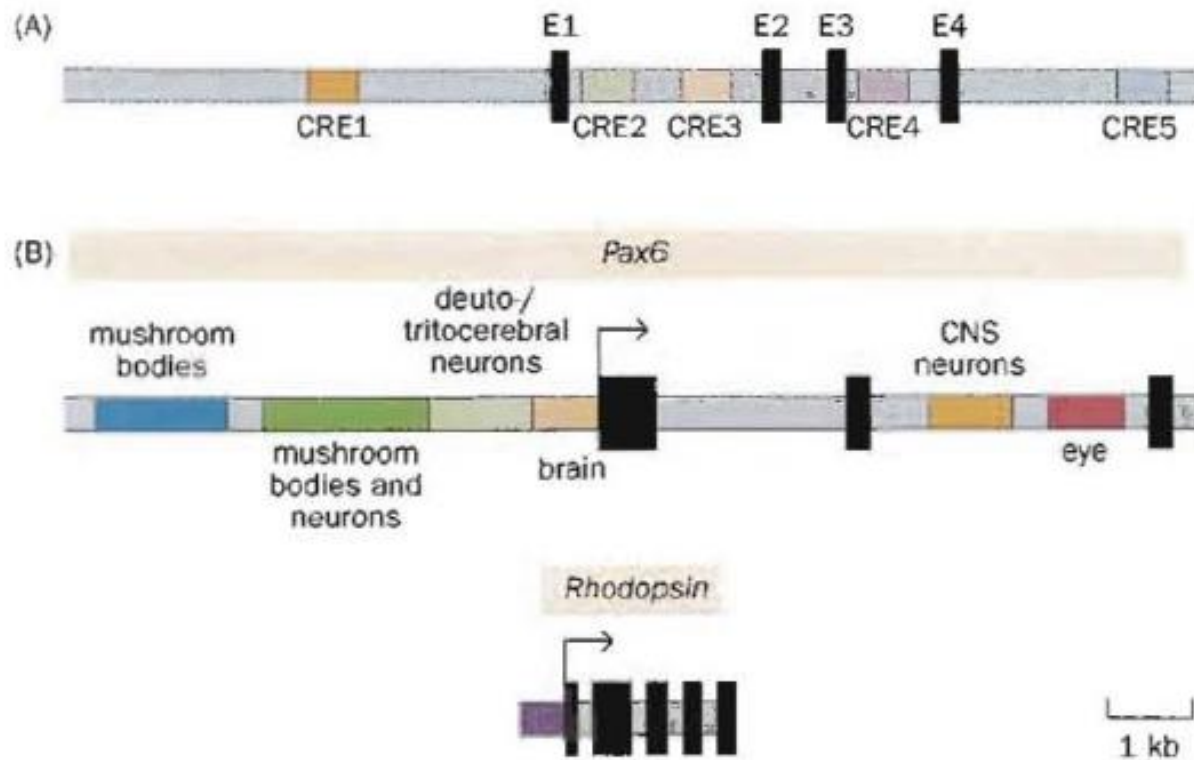
# Гены эмбрионального развития крайне консервативны

Дупликация *Нох* генов не происходила на протяжении сотен миллионов лет



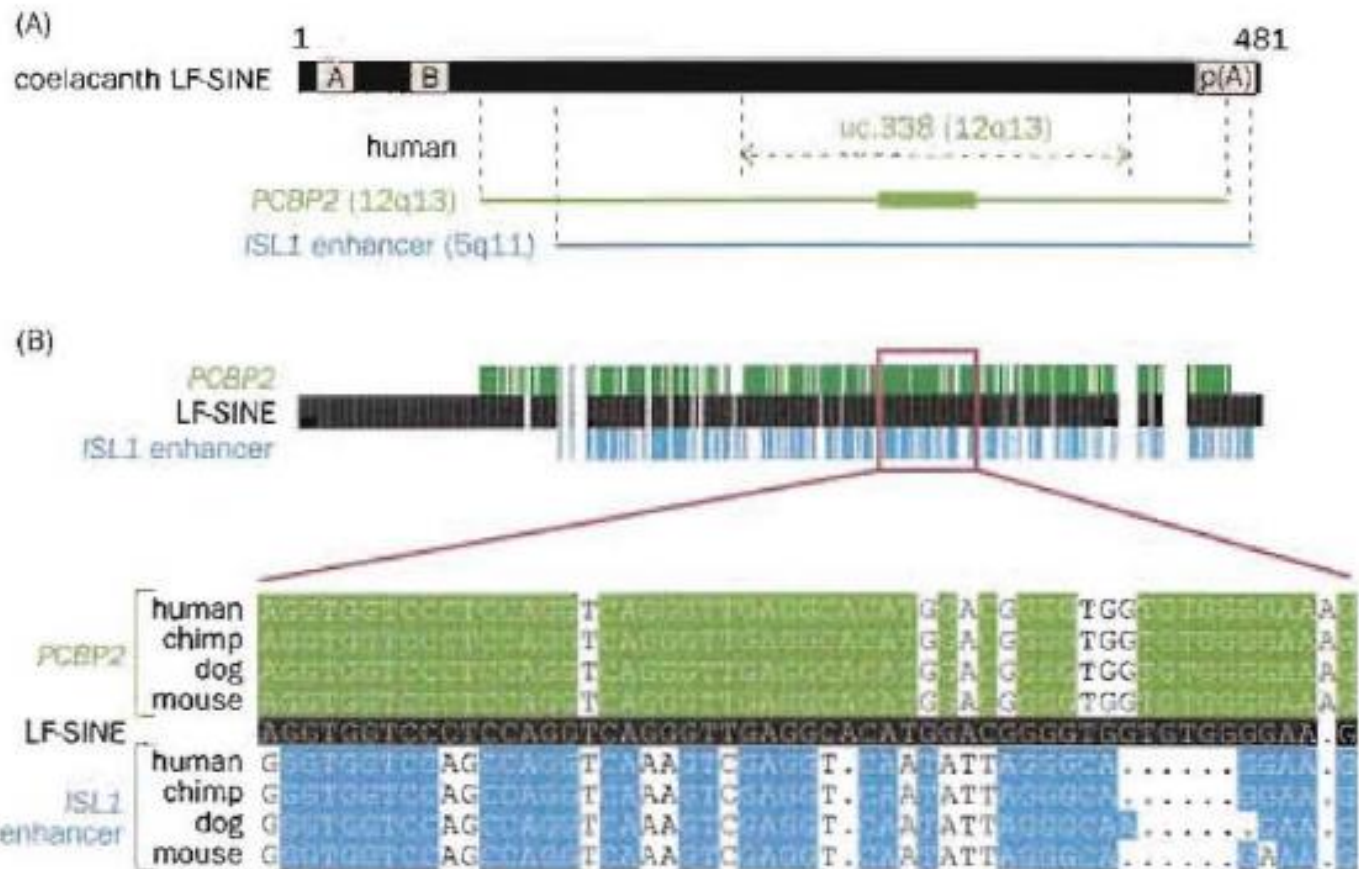
# Cis-regulatory elements (CRE)

- CRE отвечают за экспрессию генов в определенных тканях и/или в определенное время
- Новые CRE → новые функции гена



*Pax6* – транскрипционный фактор, отвечающий за развитие отделов мозга, ЦНС, глаз  
*Rhodopsin* – мишень *Pax6*, выполняет одну функцию (фоторецептор в глазу)

# Гомология экзона гена *PCBP2* и энхансера гена *ISL1* с ретротранспозоном LF-SINE



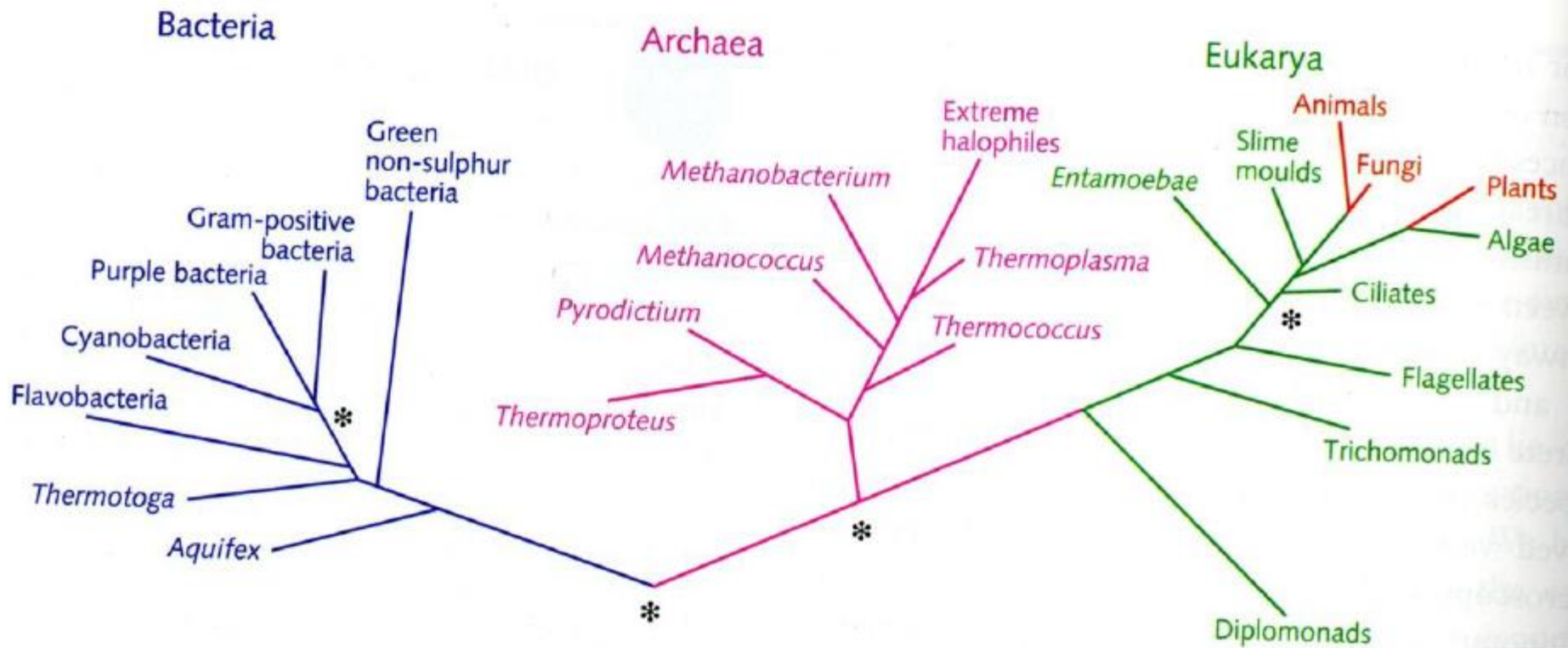
# 4. Филогенетические построения

# Наше место на эволюционном древе

- Филогенез – эволюционные отношения между организмами
- Классический подход основан на анатомических и морфологических признаках
- Молекулярная филогенетика использует сравнение нуклеотидных и аминокислотных последовательностей
- Изначально использовалось небольшое количество локусов
- Филогеномика использует знания о полных последовательностях геномов

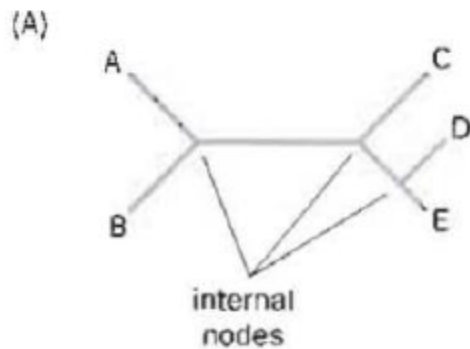


# Филогенетические деревья на основе 16s РНК



# Построение эволюционных деревьев

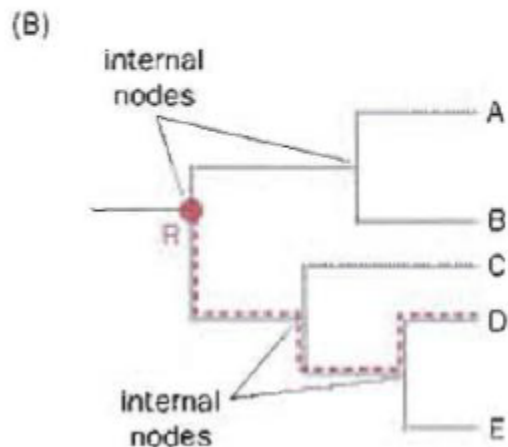
- Строятся исходя из гомологии анализируемых последовательностей нуклеиновых кислот или белков

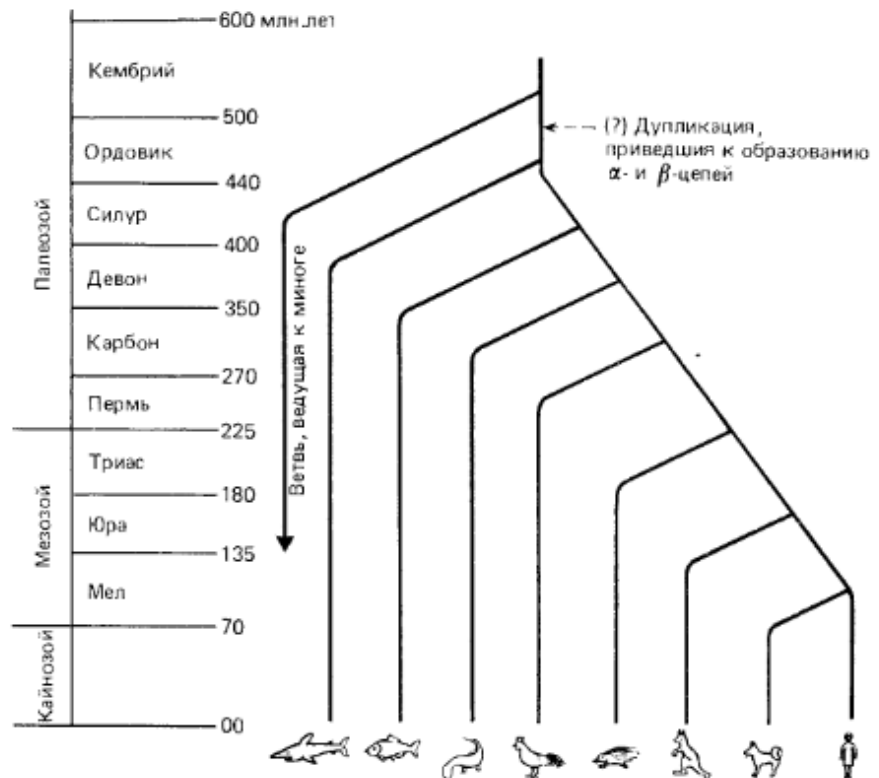


Branches & Nodes

A – unrooted tree

B – rooted tree (cladogram)





	Акула	Карп	Тритон	Курица	Ехидна	Кенгуру	Собака	Человек
Акула		59,4	61,4	59,7	60,4	55,4	56,8	53,2
Карп			53,2	51,4	53,6	50,7	47,9	48,6
Тритон				44,7	50,4	47,5	46,1	44,0
Курица					34,0	29,1	31,2	24,8
Ехидна						34,8	29,8	26,2
Кенгуру							23,4	19,1
Собака								16,3
Человек								

$\alpha$ -цепь

Различия между аминокислотными последовательностями, %

# Метод матрицы расстояний

## Distance matrix method

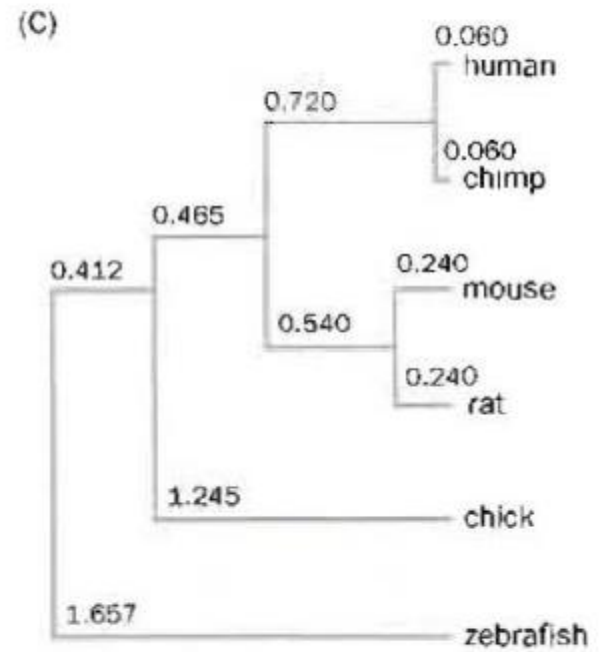
(A)

```

human  MSLTKTERTIIIVSMWAKISTQADTIGTETIERLFLSHPQTKTYFPHPDLH
chimp  .....G.....
mouse  ...M.N..A..M...E.MAA..EP.....C.Y.....
rat    ...M.N..A..M...D.MAPH.EP.....S.Y.....
chick  .A..QA.KAAVTTI...VA..IES..L.S....A.Y.....VS
zebrafish  ...AKDKAV.KGF.G..AS...S..Q.AMG.MLTVY....I..A.WPD.
    
```

(B)

	human	chimp	mouse	rat	chick	zebrafish
human		0.02	0.24	0.26	0.42	0.54
chimp			0.26	0.28	0.42	0.54
mouse				0.08	0.40	0.54
rat					0.42	0.54
chick						0.56



# Другие методы

- Варианты distance matrix
  - Neighbor relation
  - Neighbor joining (учитывает отличия в скорости эволюции)
- Maximum-parsimony – поиск дерева с минимальным количеством эволюционных шагов
- Maximum-likelihood – поиск оптимального дерева из всех возможных вариантов
- Эвристические методы (для большого объема данных)





- 5. Изучение механизмов геномных перестроек

# Негомологичная рекомбинация

- группа рекомбинационных процессов, основанных на очень ограниченной гомологии между рекомбинирующими ДНК или вообще происходящих без нее.

1. Сайт-специфическая рекомбинация
2. Транспозиции
3. Незаконная рекомбинация

# Сайт-специфическая рекомбинация

Происходит между специфическими последовательностями ДНК в пределах очень коротких участков гомологии, обычно 15-30 п.н.

- обеспечивает интеграцию (включение) ДНК умеренных фагов в хромосомы бактерий
- обеспечивает инверсию отдельных участков ДНК в хромосомах бактерий и бактериофагов и в плазмидах дрожжей
- обеспечивает перестройки в последовательностях ДНК, кодирующих иммуноглобулины

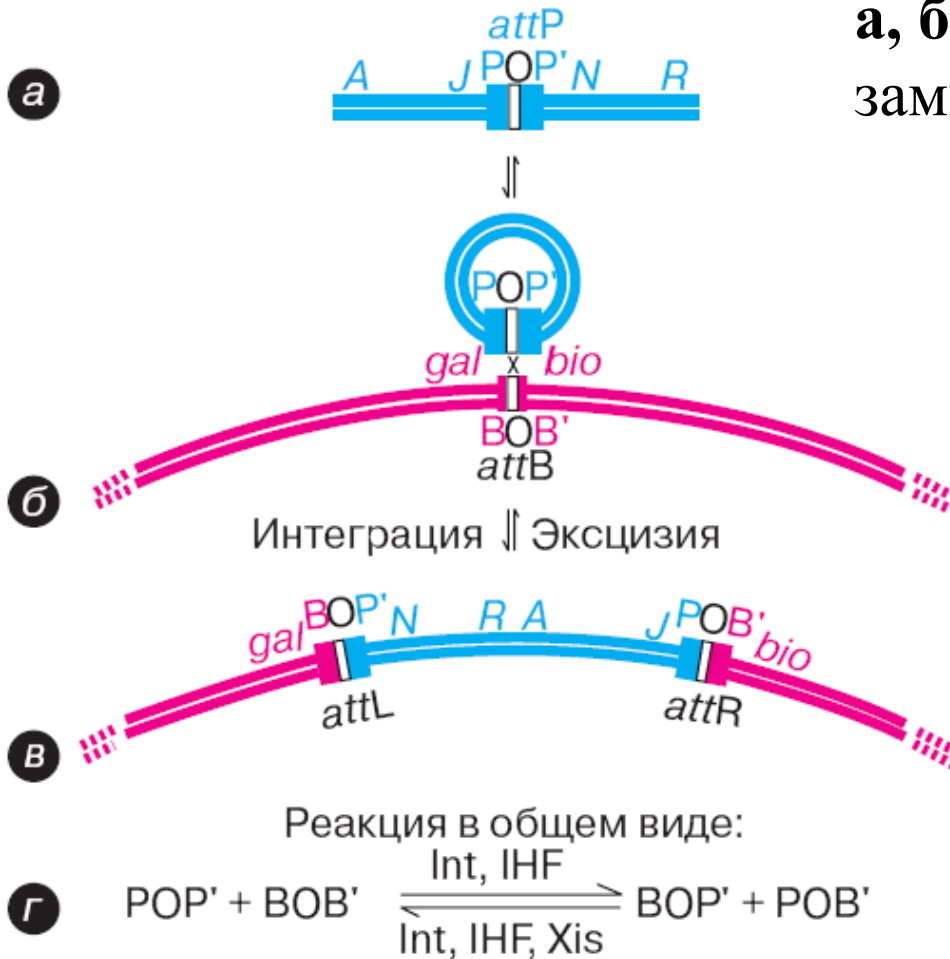
# Схема сайт-специфической рекомбинации у фага λ

**а, б** - двуцепочечная ДНК фага замыкается в кольцо

**б, в** - интеграция ДНК фага в хромосому бактерии между генами *gal* и *bio*, путем рекомбинации между особыми *att* (attachment)-сайтами

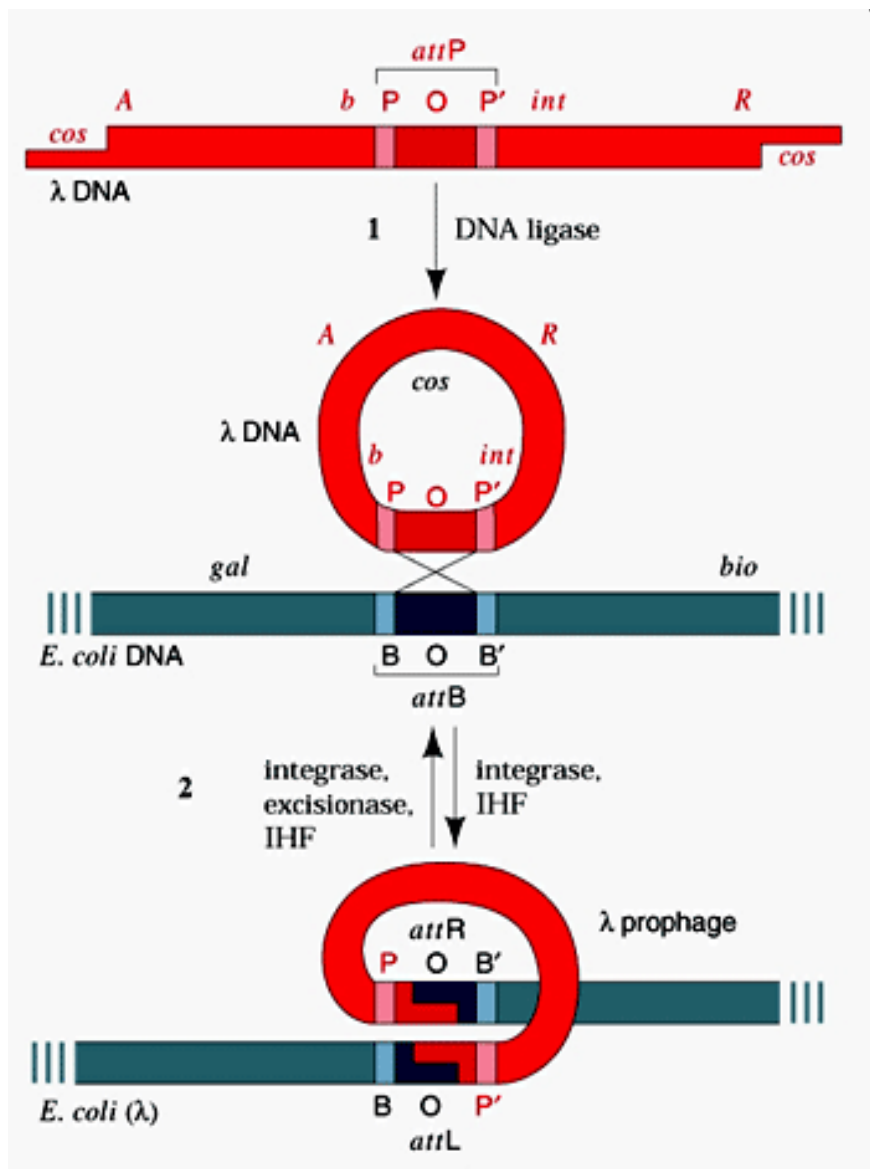
**в** – профаг

**г** – запись процесса рекомбинации в общем виде



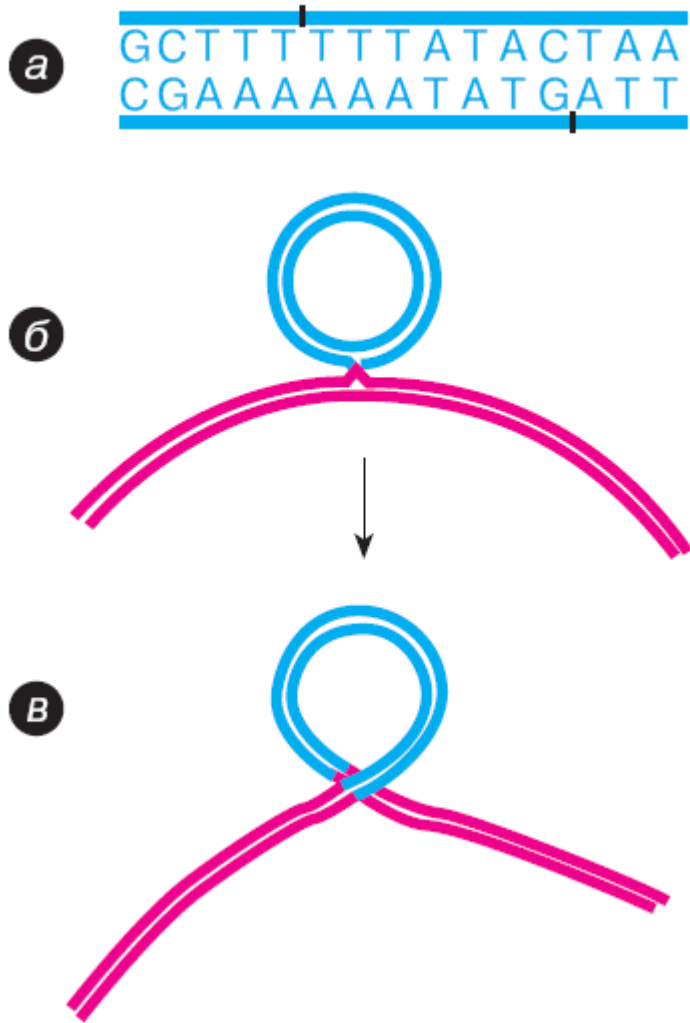


# Сайт-специфическая рекомбинация



Встраивание кольцевой ДНК фага  $\lambda$  в хромосому *E. coli* и ее обратное выщепление

# Схема двух основных этапов интегративной рекомбинации у фага лямбда

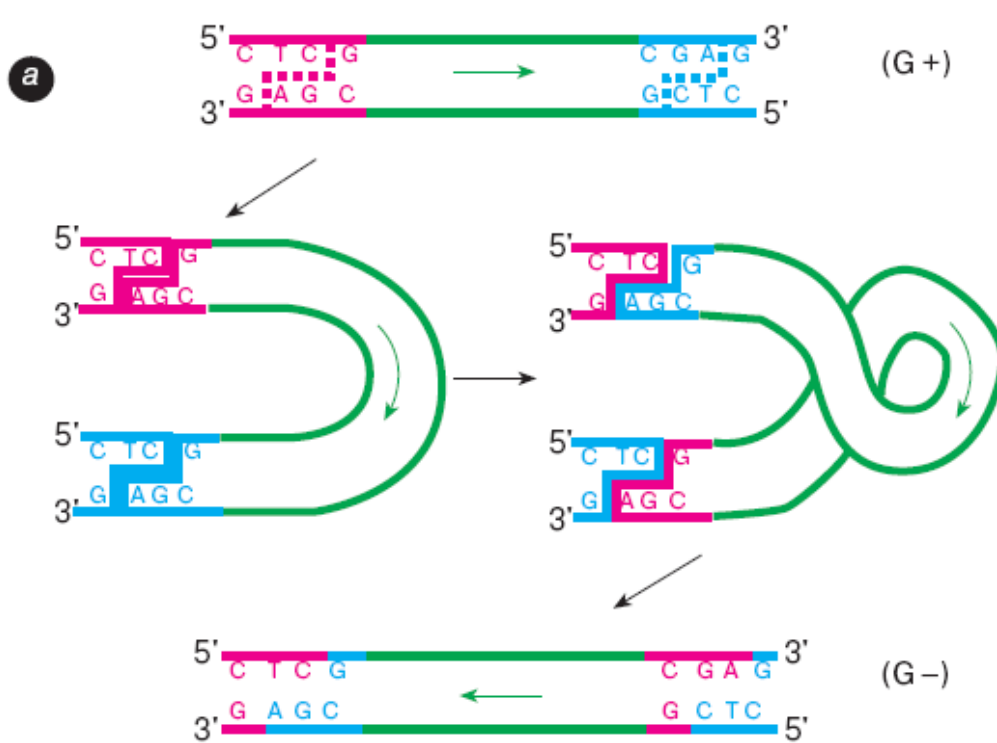


**а** - нуклеотидная последовательность att-сайтов в центральной части O

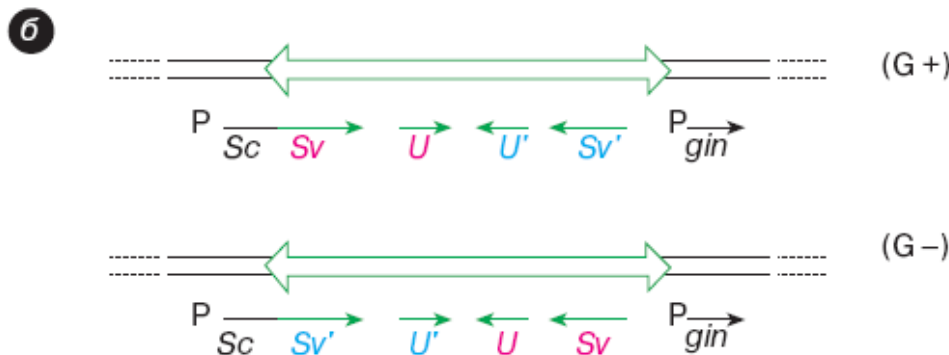
**б** – промежуточная структура, образовавшаяся после обмена двумя цепями ДНК одинаковой полярности в att-сайтах

**в** – продукт завершенной рекомбинации

# Схема сайт-специфической инверсии сегмента G в ДНК фага Му



**а** - сегмент G содержит инвертированные повторы со специфическими рекомбинационными сайтами. Инвертаза проводит рекомбинацию между этими сайтами.



**б** - G-сегмент содержит четыре гена: U, U', Sv и Sv'. При одной ориентации сегмента (G+) транскрибируются гены Sv и U, при противоположной ориентации (G-) функционируют гены Sv' и U.

# Транспозиции

Лежат в основе передвижения подвижных (мобильных) генетических элементов.

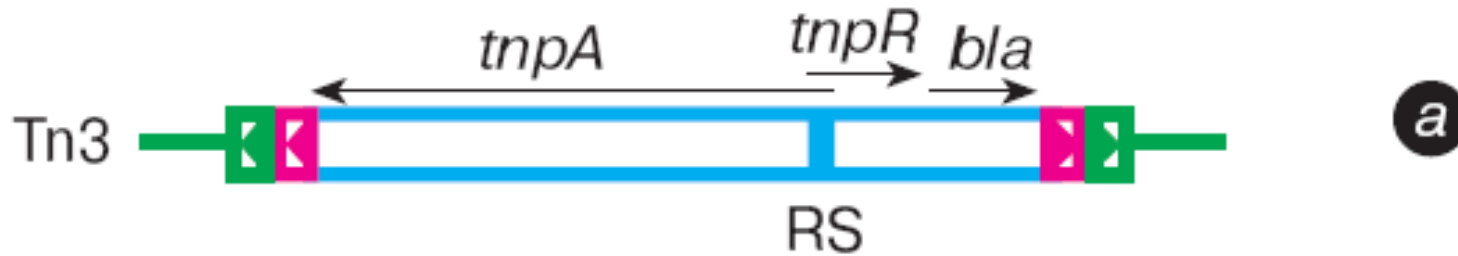
**Подвижные элементы** - это особые последовательности ДНК, способные к перемещениям из одного участка молекулы ДНК (хромосомы или плазмиды) в другой, или в другую молекулу в той же клетке, или даже в клетки другого организма.

Подвижные элементы, как правило, не существуют автономно, а находятся в составе хромосом или плазмид.

Главный белок транспозиции – транспозаза.

# Структурная организация некоторых подвижных элементов

а – бактериальный транспозон Tn3.



б – бактериальный транспозон Tn5.

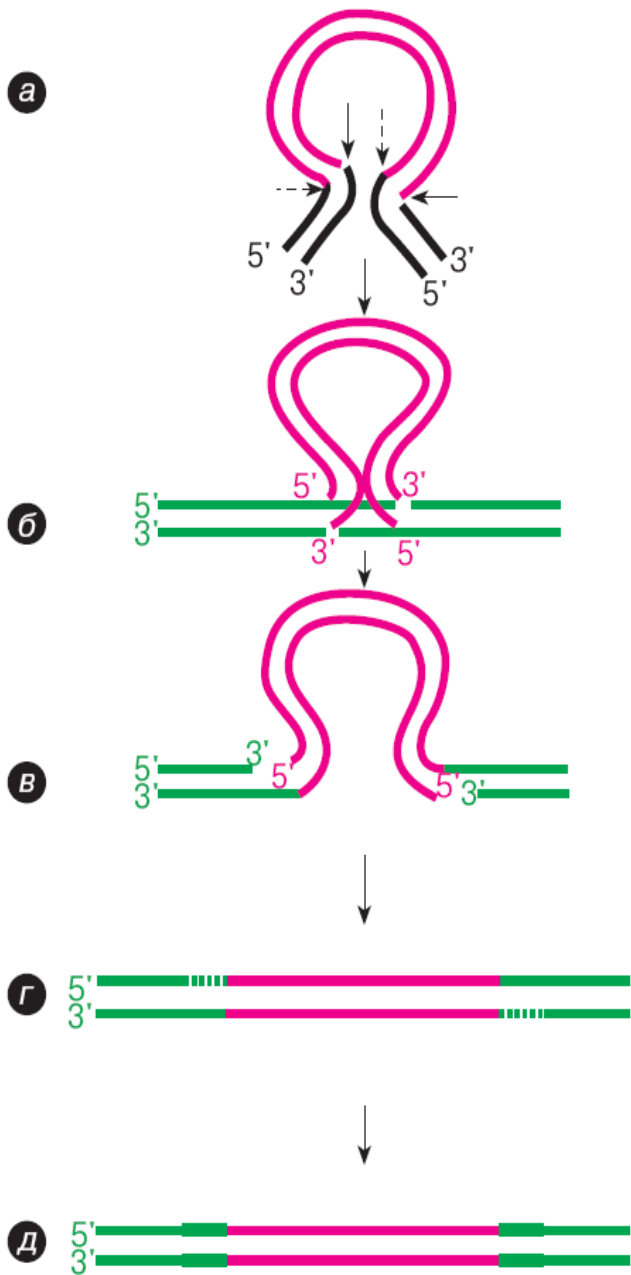


в – дрожжевой ретротранспозон Ty1.



Синим обозначена центральная часть элементов, красным - обращенные концевые повторы бактериальных подвижных элементов, зеленым - прямые повторы ДНК-мишени.

# Общая схема рекомбинационных реакций при транспозициях



**а** - транспозаза (у ретротранспозонов - интегразы) сводит вместе концы подвижного элемента и делает разрывы точно по этим концам

**б** - транспозаза сводит в контакт концы элемента и дуплекс ДНК-мишени. При этом она делает в обеих цепях ДНК-мишени ступенчатые разрывы

**в** - обмен цепями, приводящий к рекомбинации между ДНК элемента и мишени, остаются бреши

**г** - бреши заполняются путем репаративной репликации ДНК по матрице ДНК-мишени

**д** - возникновение прямых повторов ДНК-мишени на концах элемента

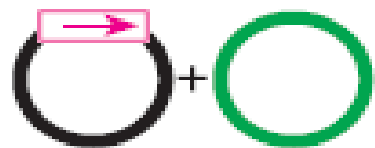


# **Выделяют три основных механизма рекомбинации при транспозициях:**

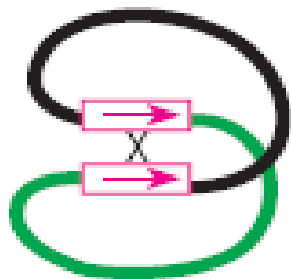
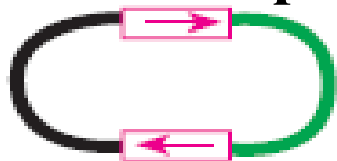
- репликативная транспозиция
- нерепликативная транспозиция
- перемещение ретротранспозонов

## Репликативная транспозиция

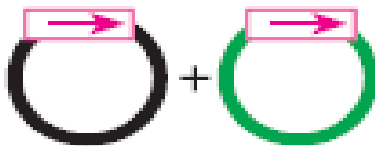
Подвижный элемент, перемещаясь в другую молекулу, оставляет свою копию в исходной ДНК.



коинтеграт



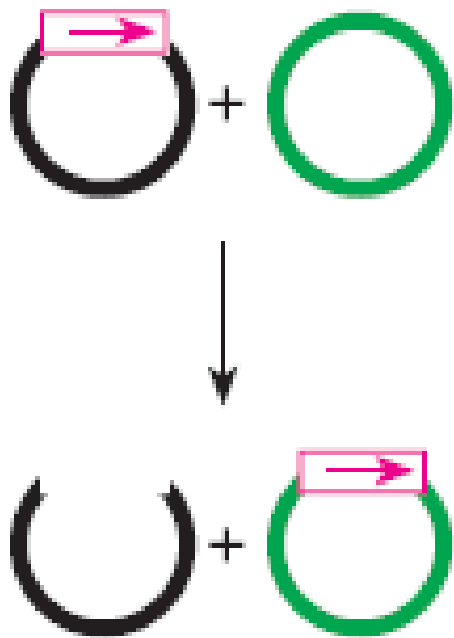
резолваза



Обнаружена у фага М<sub>1</sub> и бактериальных транспозонов семейства Tn3 с короткими обращенными повторами.

# Нерепликативная транспозиция

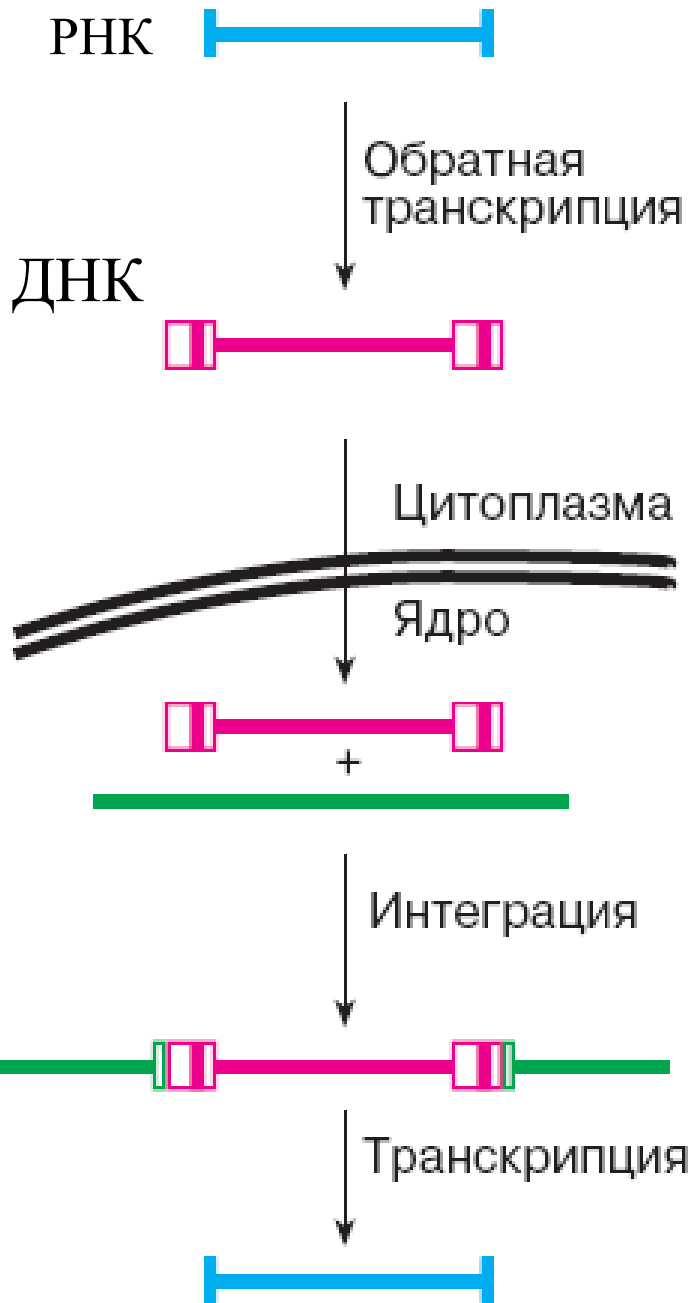
Заключается в вырезании элемента и его перемещении в новое место.



Характерна для большинства подвижных элементов бактерий и эукариотических элементов с короткими обращенными повторами

# Перемещение ретротранспозонов

Ген *pol* ретротранспозона кодирует несколько ферментов: интегразу, обратную транскриптазу, РНКазуН и протеазу.



# Незаконная рекомбинация

**Сборная группа процессов, где рекомбинация происходит без гомологии между молекулами ДНК, и при этом без участия механизмов сайт-специфической рекомбинации или транспозиций. Общим для них является соединение концов негомологичных молекул ДНК.**

- захват ретровирусом некоторых клеточных генов при его эксцизии из хромосомы хозяйской клетки
- интеграция фрагментов ДНК, вводимых в клетки позвоночных с помощью микроинъекций

Впервые описана японским исследователем Х. Икедой с сотрудниками в 1982 году.