

Выравнивание ...



340

*

... последовательностей белков
и его биологический смысл

1zjh:A : FG-VEQDVMVFASFIRKASDVHEVRKVLGE-KGKNIKII SKIENHEGVRRFDEIL
1a3x:B : FG-VKNGVHMVFASFIRTANDVLTIREVLGE-QGKDVKII VKIENQQGVNNFDEIL
3eoe:B : NFgIPMGCNFIAASFVQSADDVRYIRGLLGP-RGRHIRIIPKIENVEGLVNFDEIL
1aqf:B : FG-VEQDVMVFASFIRKAADVHEVRKILGE-KGKNIKII SKIENHEGVRRFDEIL
1pk1:G : FG-VEQGVDMI FASFIRSAEQVGDVRKALGP-KGRDIMIICKIENHQGVQNI DSI
1e0t:A : FG-CEQGVDFVAASFIRKRSDVIEIREHLKAhGGENIHII SKIENQEG LNNFDEIL
fg v ASfir a dV R Lg G i II KIEN G fDeIl

* 300 * 320 *

Мутации и отбор

1zjh:A : EASDGI MVARGDLGIEI pA-EKVFLA QKMMI GRCNRAGKPVICATQmlesmikkpR
1a3x:B : KVDVIVRQD GIEI pA-PEVFAVQK IAKSNFAKRVICATQmlesmtynpR
3eoe:B : AEIDGIMTIRQD GIEI pA-PEVFLA QKMMI GRCNRAGKPVICATQmlesmiknpR
1aqf:B : EASDGI MVARGDLGIEI pA-EKVFLA QKMMI GRCNRAGKPVICATQmlesmikkpR
1pk1:G : EESDGI MVARGDLGVEI -PaEKVVVAQKIL I SKCNVAGKPVICATQmlesmtynpR
1e0t:A : EASDGI MVARGDLGVEI pV-EEVIFAQKMMIEK CIRARKVVITATM-----R
DGilVARGDLGIEI pV-EEVIFAQKMMIEK CIRARKVVITATM-----R

Сегодня основное внимание

уделим мутациям

340 360 380 *

1zjh:A : PTRAE GSDVANAVLDGADCIMLSGE takgdy PLEAVRMQH LIAREAE Aaiyhqlf
1a3x:B : PTRAE VSDVGNAILD GADCVMLSGE takgny PINAVTTMAETAVIAEQ aiaylpny
3eoe:B : PTRAE AADVANAVLDGTDCVMSLGE tangef PVITVETMARICYEAET cvdypaly
1aqf:B : PTRAE GSDVANAVLDGADCIMLSGE takgdy PLEAVRMQH LIAREAE Aamfhrklf
1pk1:G : PTRAE VSDVANAVFNGADCVMLSGE takgky PNEVVQYMARICLEAQ S alneyvff
1e0t:A : PTDAEAGDVANAILDGTDAVMSLGE ----- PLEAVSIMATI CERTDR-----
PTrAE DVaNA ldG Dc MLSGETa q P V i a

Мутации белка – следствия мутаций кодирующей последовательности его ген

• Последовательность одного и того же белка у потомка может отличаться от последовательности предка

• Отличия происходят из-за мутаций в кодирующей последовательности гена

CDS, coding sequence –

кодирующая

последовательность гена

atgсссаааgсс.....tga

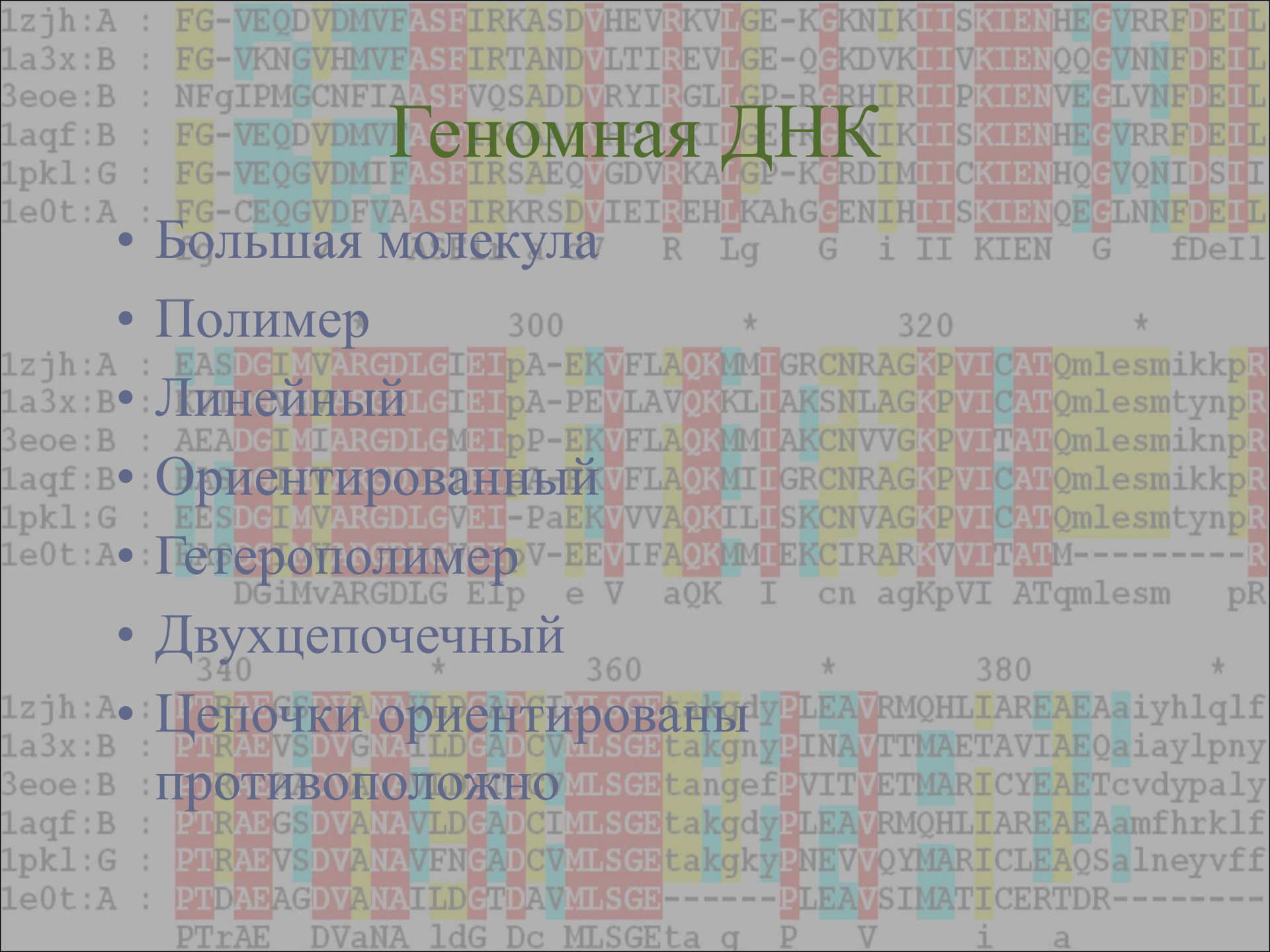
Начало CDS: **atg**

Стоп-кодоны: **taa, tag, tga**

Геномная ДНК

- Большая молекула
- Полимер
- Линейный
- Ориентированный
- Гетерополимер
- Двухцепочечный

- Цепочки ориентированы противоположно



Каждая комплементарная пара нуклеотидов ДНК потомка происходит из комплементарной пары ДНК предка

• ДНК потомка всегда получается из предковой ДНК репликацией

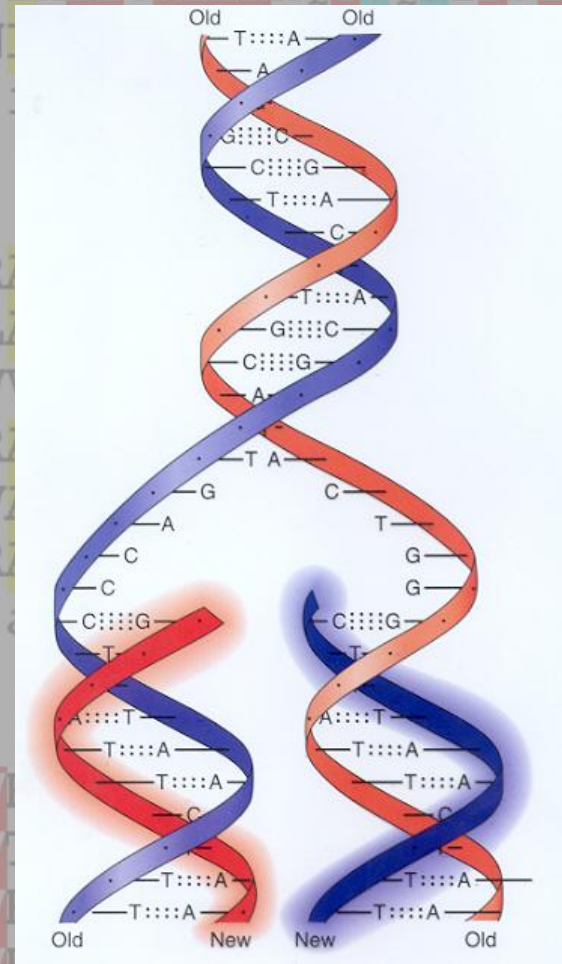
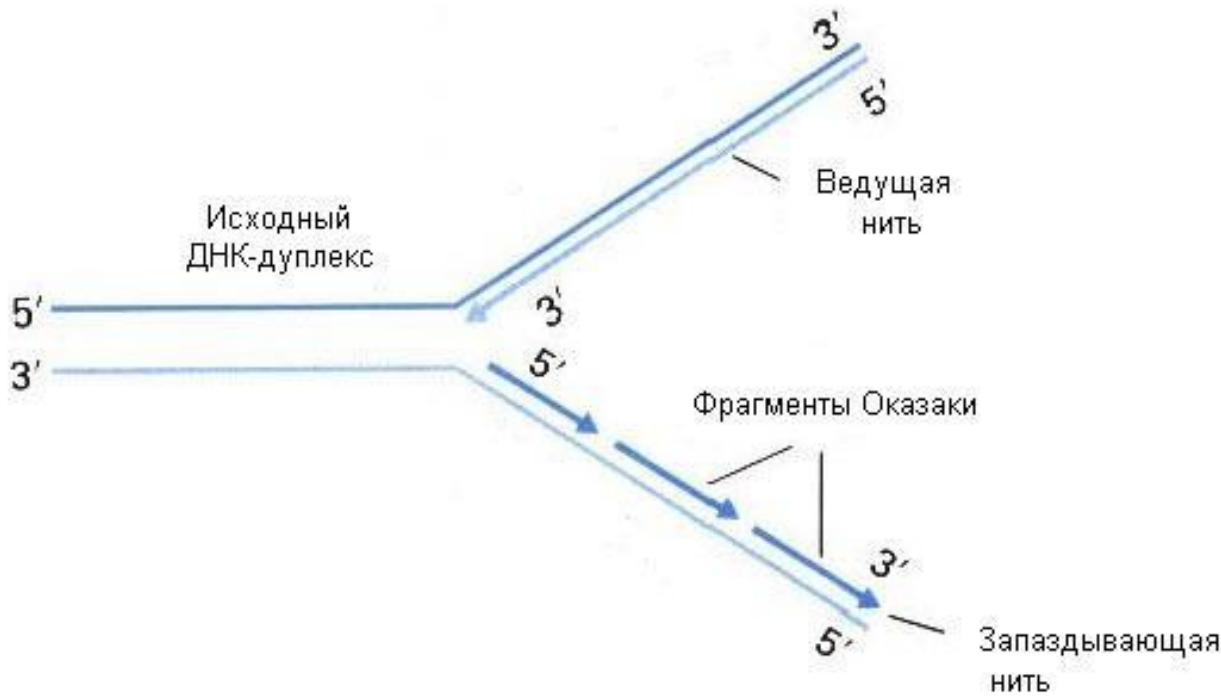
• Поэтому ДНК-потомок имеет ту же последовательность, что и ДНК-предок

• *Закон биологии: из каждого закона биологии существуют исключения*

• *Вопрос: существуют ли исключения из этого закона?*

Схема репликации

РЕПЛИКАТИВНАЯ ВИЛКА



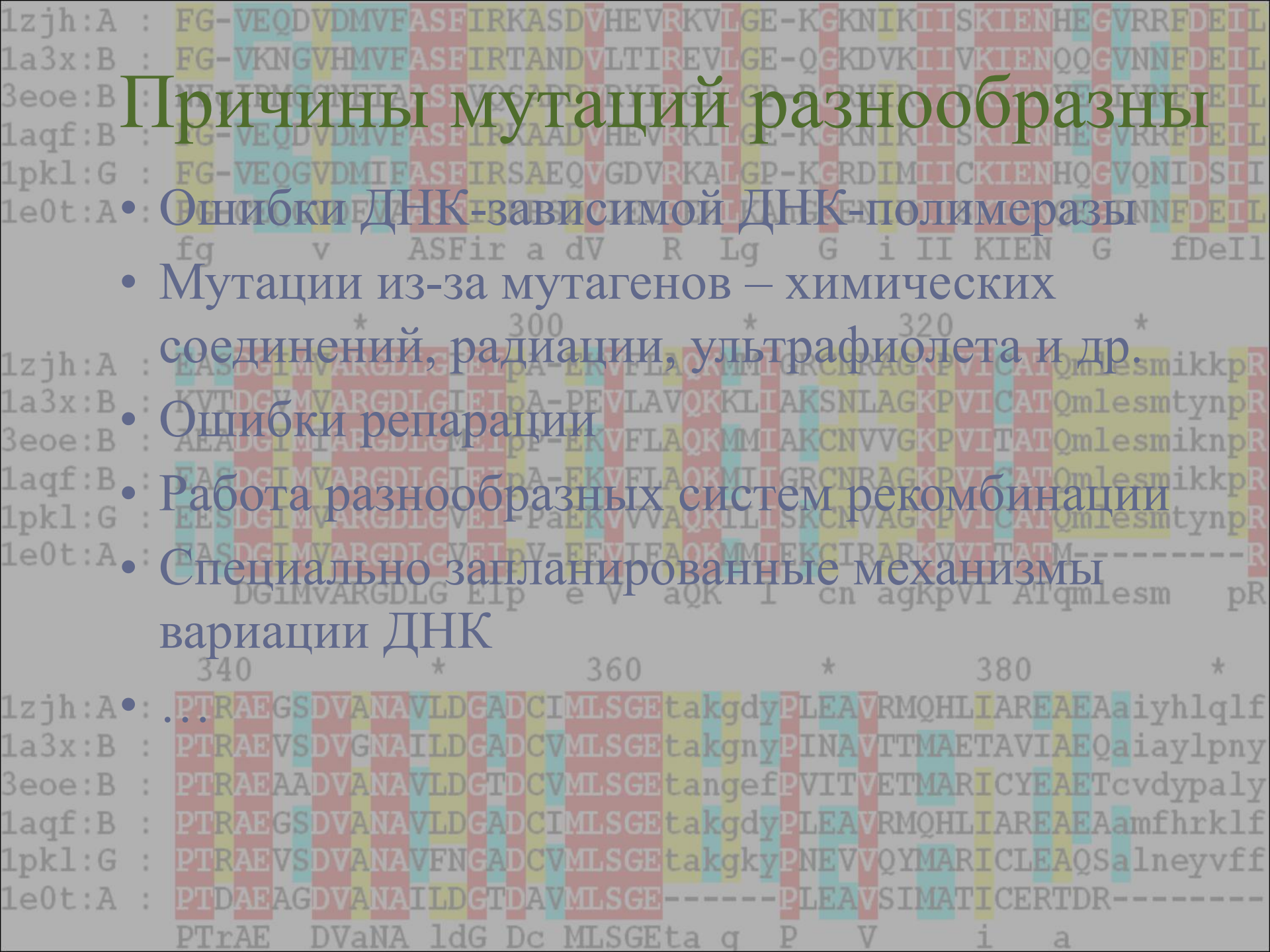
Виды мутаций

- Замена нуклеотида на другой
- Делеция одного или нескольких подряд нуклеотидов
- Вставка одного или нескольких подряд нуклеотидов

- Дупликация участка
- Перестановка участка - рекомбинация

Причины мутаций разнообразны

- Ошибки ДНК-зависимой ДНК-полимеразы
- Мутации из-за мутагенов – химических соединений, радиации, ультрафиолета и др.
- Ошибки репарации
- Работа разнообразных систем рекомбинации
- Специально запланированные механизмы вариации ДНК



Мутации CDS и отбор

- Мутации белка происходят из-за мутаций кодирующей последовательности ДНК

- Последовательность белка находится под прессом отбора:

- каждый белок имеет свои биологические функции
- функция белка зависит от структуры белка и от остатков, напрямую участвующих в выполнении функции
- свойства и структура белка зависят от последовательности

- Малое число мутаций CDS имеют шанс закрепиться в эволюции

- Мутации последовательности белка происходят редко; *но если подождать 10 млн лет, то можно обнаружить довольно много!*

Как отобразить отношение предок — потомок для нуклеотидов?

ДНК предка: **AAAC TGAT GCAACGTGA**



ДНК потомка: **...AAtCttTGATAcCTGA...**

Выравнивание – общепринятый способ отражения “родства” нуклеотидов или аминокислотных остатков

seq1:.....AAAC--TGATGCAACGTGA.....

seq2:.....AAtCttTGAT---ACcTGA.....

“Наследование”

аминокислотного остатка белка

- Если все три нуклеотида **кодона** потомка унаследованы от трех нуклеотидов **кодона**

предка, то можно говорить о том, что остаток потомка произошел от остатка предка

- Такие остатки следует помещать в одну колонку выравнивания аминокислотных последовательностей

- Это еще не вся правда про выравнивание последовательностей белков!

Проблема выравнивания

- Время наблюдения за последовательностью

много меньше 10 млн лет 😞

- Мы наблюдаем только потомков общего предка, а самого предка не знаем

- Про родство аминокислотных остатков приходится догадываться

- Как догадываться — об этом на следующих занятиях

Аминокислотные остатки помещают в одну колонку выравнивания если они

- происходят от одного предкового остатка последовательности белка – общего предка (эволюция)
- их C_α атомы находятся в участках полипептидной цепи сходной конформации (структура)
- играют сходную роль в белке (функция)

Пример выравнивания

```

          *           240           *           260           *           280
1zjh:A : FG-VEQDVDMVFASFIRKASDVHEVRKVLGE-KGKNIKIISKIENHEGVRRFDEIL : 260
1a3x:B : FG-VKNGVHMVFASFIRTANDVLTIREVLGE-QGKDVKIIIVKIENQQGVNMFDEIL : 254
3eoe:B : NFgIPMGCNFIAASFVQSADDVRYIRGLLGP-RGRHIRIIPKIENVEGLVNFDEIL : 249
1aqf:B : FG-VEQDVDMVFASFIRKAADVHEVRKILGE-KGKNIKIISKIENHEGVRRFDEIL : 272
1pkl:G : FG-VEQGVDMIFASFIRSAEQVGDVRKALGP-KGRDIMIICKIENHQGVQNIDSI : 252
1e0t:A : FG-CEQGVDFVAASFIRKRSDVIEIREHLKAhGGENIHIIISKIENQEGLNMFDEIL : 230
fg      v      ASfir a dV      R      Lg      G      i      II      KIEN      G      fDeil

```

```

          *           300           *           320           *
1zjh:A : EASDGI MVAR GDLGIEI pA -EKVFLAQKMMI GRCNRAGKPVICATQmlesmikkpR : 315
1a3x:B : KVT DGV MVAR GDLGIEI pA -PEVLAVQK KLI AKSNLAGKPVICATQmlesmtynpR : 309
3eoe:B : AEADGI MIAR GDLGMEI pP -EKVFLAQKMMI AKCNVVGKPVITATQmlesmiknpR : 304
1aqf:B : EASDGI MVAR GDLGIEI pA -EKVFLAQKMI I GRCNRAGKPVICATQmlesmikkpR : 327
1pkl:G : EESDGI MVAR GDLGVEI -PaEKVVVAQKIL I SKCNVAGKPVICATQmlesmtynpR : 307
1e0t:A : EASDGI MVAR GDLGVEI pV -EEVIFAQKMMI EK CIRARKVVI TATM-----R : 276
DGiMvARGDLG EIp e V aQK I cn agKpVI ATqmlesm pR

```

```

          340           *           360           *           380           *
1zjh:A : PTRAE GSDVANAVLDGADCI MLSGETakgdy PLEAVRMQH LIAREAEAaiyhlqlf : 371
1a3x:B : PTRAE VSDVGNAILD GADCV MLSGETakgny PINAVTTMAETAVIAEQaiaylpny : 365
3eoe:B : PTRAE AADVANAVLDGTD CV MLSGETangef PVITVETMARICYEAE Tcvdypaly : 360
1aqf:B : PTRAE GSDVANAVLDGADCI MLSGETakgdy PLEAVRMQH LIAREAEAamfhrklf : 383
1pkl:G : PTRAE VSDVANAVFNGADCV MLSGETakgky PNEVVQYMARICLEAQSalneyvff : 363
1e0t:A : PTDAE AGDVANAILDGTDAV MLSGE----- PLEAVSIMATICERTDR----- : 318
PTrAE DVaNA ldG Dc MLSGETa g P V i a

```


Выравнивание и эволюция

```
          *          20          *          4
POLG_CXB4J : GAQVSTQKTGAHETSLASGNSIIHYTNINYYKDAASNS : 39
POLG_CXB4E : GAQVSTQKTGAHETSLSATGNSIIHYTNINYYKDAASNS : 39

          0          *          60
POLG_CXB4J : ANRQDFTQDPSKFTPEPVKDVMIKSLPALN : 68
POLG_CXB4E : ANRQDFTQDPSKFTPEPVKDVMIKSLPALN : 68
```

Последовательности белка оболочки из двух штаммов
вируса Коксаки

```
          340          *          360          *          380          *
1zjh:A : PTRAE GSDVANAVLDGADCIMLSGEtakgdyPLEAVRMQHLLIAREAEAAaiyhqlf
1a3x:B : PTRAE GSDVANAVLDGADCIMLSGEtakgnyPINAVTTMAETA VIAEQaiaylpny
3eoe:B : PTRAE AADVANAVLDGTDCVMLSGETangefPVITVETMARICYEAETcvdypaly
1aqf:B : PTRAE GSDVANAVLDGADCIMLSGEtakgdyPLEAVRMQHLLIAREAEAAamfhrklf
1pk1:G : PTRAE VSDVANAVFNGADCVMLSGETakgkyPNEVVQYMARICLEAQSalneyvff
1e0t:A : PTDAEAGDVANAAILDGTDAVMLSGE-----PLEAVSIMATICERTDR-----
          PTrAE DVaNA ldG Dc MLSGETa q P V i a
```

		*		20		*		4	
POLG_CXB4J :	GAQVSTQKTGAHETSLSASGNSIIHYTNINYYKDAASNS	:							39
POLG_CXB4E :	GAQVSTQKTGAHETSLSATGNSIIHYTNINYYKDAASNS	:							39
POLG_HE71B :	GSQVSTQRS GSHENSNSATEGSTINYYTTINYYKDSYAAT	:							39

	0		*		60			
POLG_CXB4J :	ANRQDFTQDPSKFTPEVKDVMIKSLPALN	:						68
POLG_CXB4E :	ANRQDFTQDPSKFTPEVKDVMIKSLPALN	:						68
POLG_HE71B :	AGKQSLKQDPDKFANPVKDI FTEMAAPLK	:						68

Последовательности белка оболочки из двух штаммов
 вируса Коксаки и энтеровируса человека

1zjh:A : PTRAE GSDVANAVLDGADCIMLSGE takgdy PLEAVRMQH LIAREAE Aaiyh lqlf
 1a3x:B : PTRAE VSDVANAVLDGADCIMLSGE takgny PNEVVQYMARIC LEAQ S alneyvff
 3eoe:B : PTRAE GSDVANAVLDGADCIMLSGE takgdy PLEAVRMQH LIAREAE Aamfhrklf
 1aqf:B : PTRAE GSDVANAVLDGADCIMLSGE takgdy PLEAVRMQH LIAREAE Aamfhrklf
 1pk1:G : PTRAE VSDVANAVFNGADCIMLSGE takgky PNEVVQYMARIC LEAQ S alneyvff
 1e0t:A : PTDAEAGDVANA ILDGTDAVMLS GE----- PLEAVSIMATI CERTDR-----
 PTrAE DVaNA ldG Dc MLSGEta q P V i a

Что есть что (редактор GeneDoc)

Номер столбца выравнивания

```

                *                20                *
MTA1_YEAST : ---KSSISPOARAFLEQVRRK---QSLNS : 24
MAT2_YEAST : KPYRGHRETKENVRILESWEAKNIENPYLDT : 31
                3 2      LE  F  4      L13
    
```

```

                40                *                60
MTA1_YEAST : KEKEEVAKCGITPLQVRVWFINKRMRSK- : 53
MAT2_YEAST : KGLENIIMKNTSLSRIQIKNWVSNRRRKEKT : 61
                K  E  6  K      63  6Q64  W  N4R  4  K
    
```

Название последовательности

Функционально консервативная позиция

Консервативный остаток

Номер последнего в строке остатка ИЗ ЭТОЙ ПОСЛЕДОВАТЕЛЬНОСТИ

1zjh:A : FG-VEQDVMVFASFIRKASDVHEVRKVLGE-KGKNIKII SKIENHEGVRRFDEIL
 1a3x:B : FG-VKNGVHMVFASFIRTANDVLTIREVLGE-QGKDVKII VKIENQQGVNNFDEIL
 3eoe:B : NFgIPMGCNFIAASVQSADDVRYIRGLLGP-RGRHIRIIPKIENVEGLVNFDEIL
 1aqf:B : FG-VEQDVMVFASFIRKAADVHEVRKILGE-KGKNIKII SKIENHEGVRRFDEIL
 1pk1:G : FG-VEQGVDMI FASFIKIQVGIYFAIIPGGRDIMIICKIENHQGVQNI DSI
 1e0t:A : FG-CEQGVDFVAASFIRKRSDVIEIREHLKAhGGENIHII SKIENQEGLNNFDEIL
 fg v ASfir a dV R Lg G i II KIEN G fDeIl

Сегодня все

* 300 * 320 *

1zjh:A : EASDGI MVARGDLGIEIpA-EKVFLAQKMMI GRCNRAGKPVICATQmlesmikkpR
 1a3x:B : KVTDGVMVARGDLGIEIpA-PEVLAVQKKLI AKSNLAGKPVICATQmlesmtynpR
 3eoe:B : AEADGIMIARGDLGMEIpP-EKVFLAQKMMIAKCNVVGKPVITATQmlesmiknpR
 1aqf:B : EASDGI MVARGDLGIEIpA-EKVFLAQKMI I GRCNRAGKPVICATQmlesmikkpR
 1pk1:G : EESDGI MVARGDLGVEI-PaEKVVVAQKILISKCNVAGKPVICATQmlesmtynpR
 1e0t:A : EASDGI MVARGDLGVEIpV-EEVIFAQKMMIEK CIRARKVVITATM-----R
 DGiMvARGDLG EIp e V aQK I cn agKpVI ATqmlesm pR

340 * 360 * 380 *

1zjh:A : PTRAE GSDVANAVLDGADCIMLSGEtakgdyPLEAVRMQH LIAREAEAAaiyhqlf
 1a3x:B : PTRAEVSDVGNAILDGADCVMLSGETakgnyPINAVTTMAETA VIAEQaiaylpny
 3eoe:B : PTRAEAADVANAVLDGTDCVMLSGEtangefPVITVETMARICYEAETcvdypaly
 1aqf:B : PTRAE GSDVANAVLDGADCIMLSGEtakgdyPLEAVRMQH LIAREAEAAamfhrklf
 1pk1:G : PTRAEVSDVANAVFNGADCVMLSGETakgkyPNEVVQYMARICLEAQSalneyvff
 1e0t:A : PTDAEAGDVANAILDGTDAVMLSGE-----PLEAVSIMATICERTDR-----
 PTrAE DVaNA ldG Dc MLSGETa q P V i a