

# 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 MVARGDLGIEIpA-EKVFLA QKMMI GRCNRAGKPVICATQmlesmikkpR  
1a3x:B : KVIDVIVRQD GIEIpA-PEVFAVQK IAKSNFAKRVICATQmlesmtynpR  
3eoe:B : AEIDGIMTIRQD GIEIpA-EKVFLA QKMMI GRCNRAGKPVICATQmlesmiknpR  
1aqf:B : EASDGI MVARGDLGIEIpA-EKVFLA QKMMI GRCNRAGKPVICATQmlesmikkpR  
1pk1:G : EESDGI MVARGDLGVEI-PaEKVVVAQKILISKCNVAGKPVICATQmlesmtynpR  
1e0t:A : EASDGI MVARGDLGVEIpV-EEVIFAQKMMIEK CIRARKVVITATM-----R  
DGilVARGDLGVEIIP eV aqR I eR aqR VI ATqmlesm pR

Сегодня основное внимание

уделим мутациям

340 360 380 \*

1zjh:A : PTRAE GSDVANAVLDGADCIMLSGEtakgdyPLEAVRMQH LIAREAEAAiyhlqlf  
1a3x:B : PTRAEVSDVGNAILD GADCVMLSGEtakgnyPINAVTTMAETA VIAEQaiaylpny  
3eoe:B : PTRAEAADVANAVLDGTDCVMLSGETangefPVITVETMARICYEAETcvdypaly  
1aqf:B : PTRAE GSDVANAVLDGADCIMLSGEtakgdyPLEAVRMQH LIAREAEAAamfhrklf  
1pk1:G : PTRAEVSDVANAVFNGADCVMLSGEtakgkyPNEVVQYMARICLEAQSalneyvff  
1e0t:A : PTDAEAGDVANA I LDGTDAVMLSGE-----PLEAVSIMATICERTDR-----  
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**

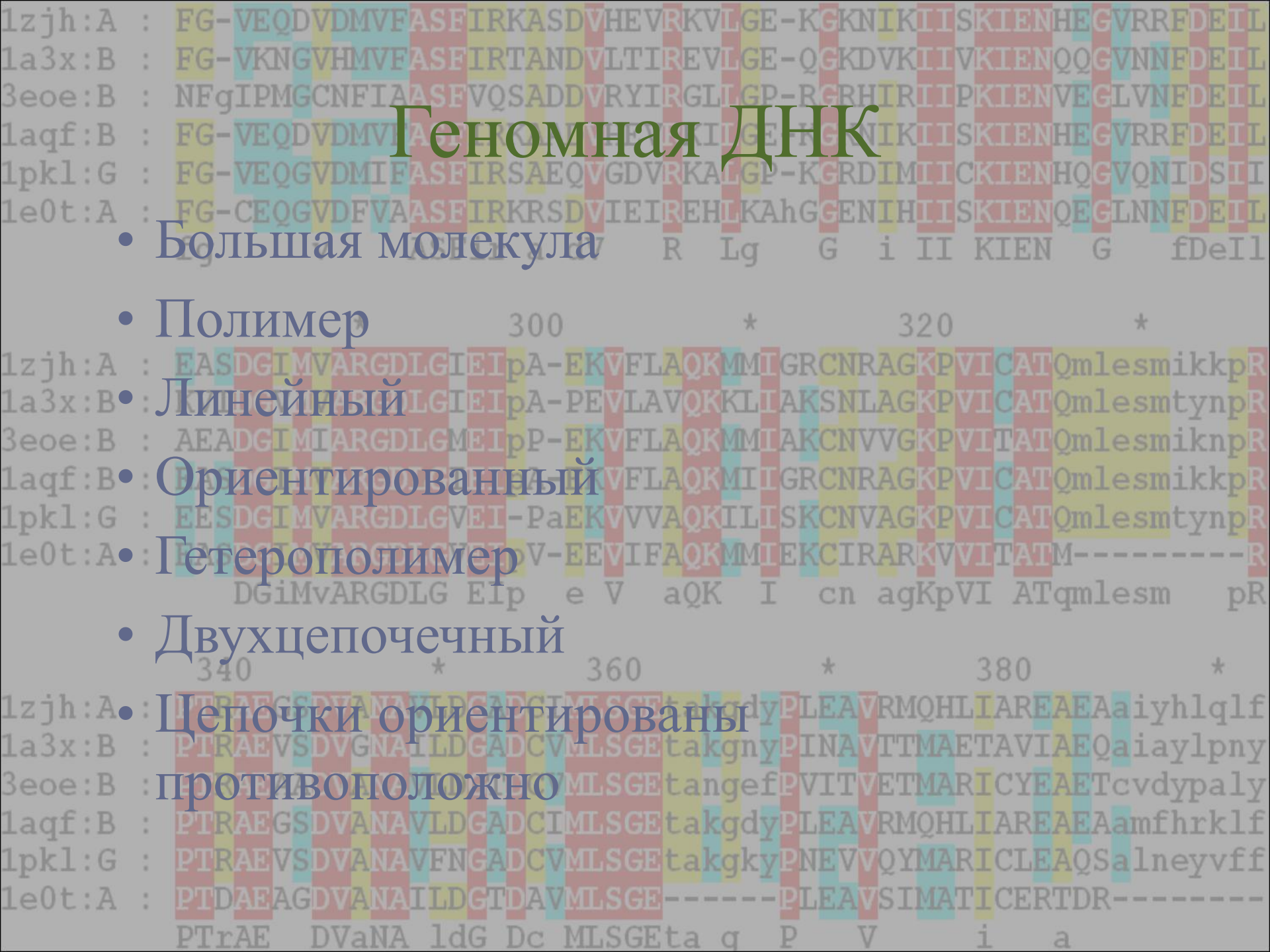
Стоп-кодоны: **taa, tag, tga**



# 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

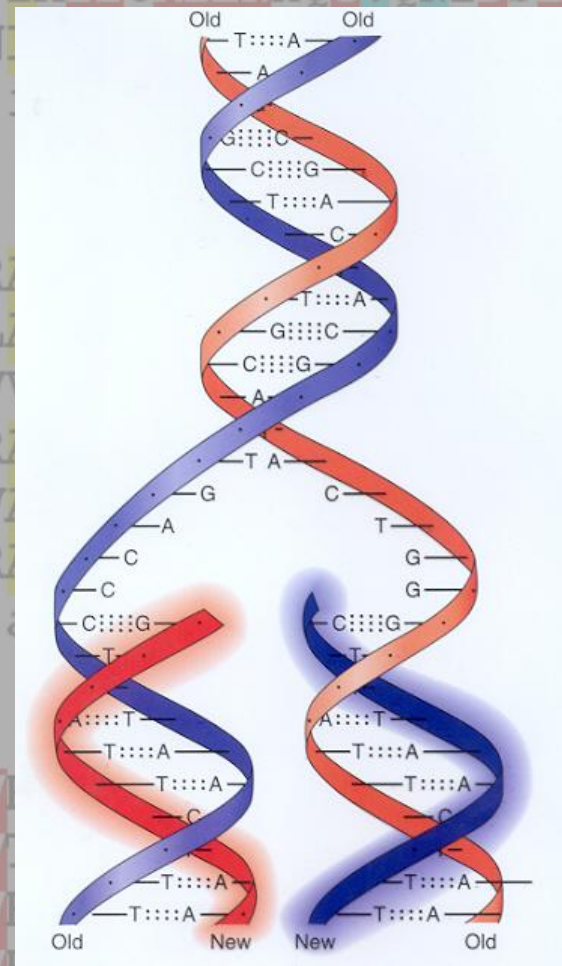
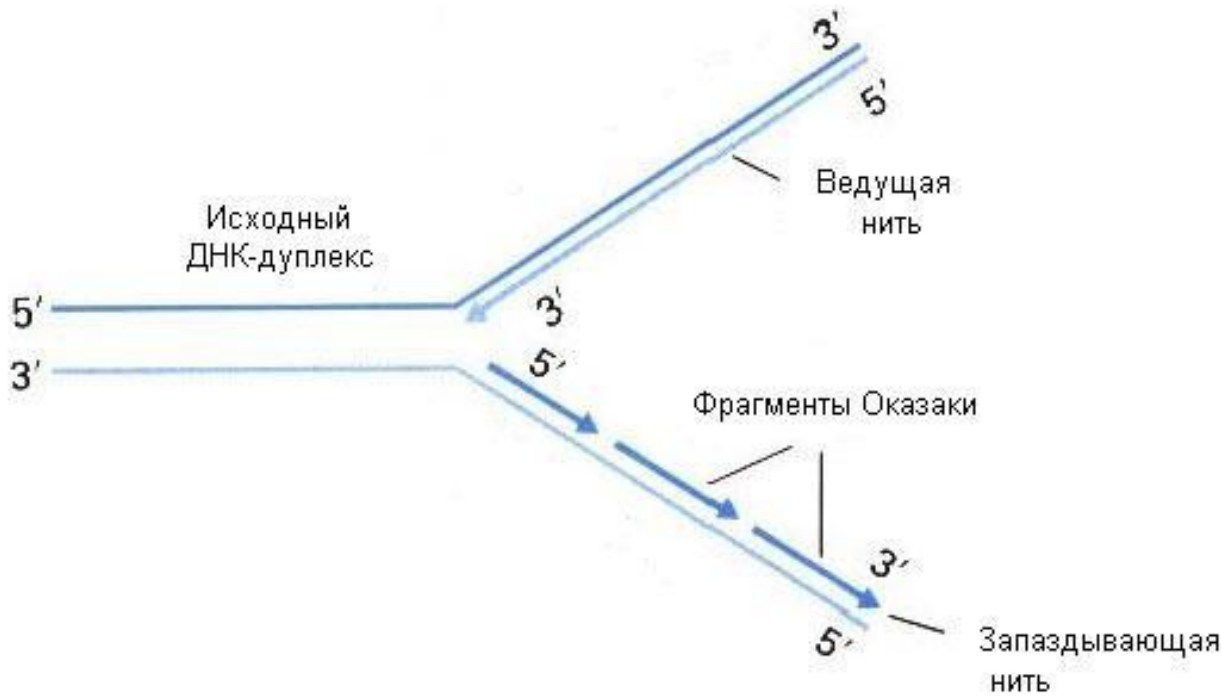
• *Закон биологии: из каждого закона биологии существуют исключения*

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



# Схема репликации

## РЕПЛИКАТИВНАЯ ВИЛКА



# Виды мутаций

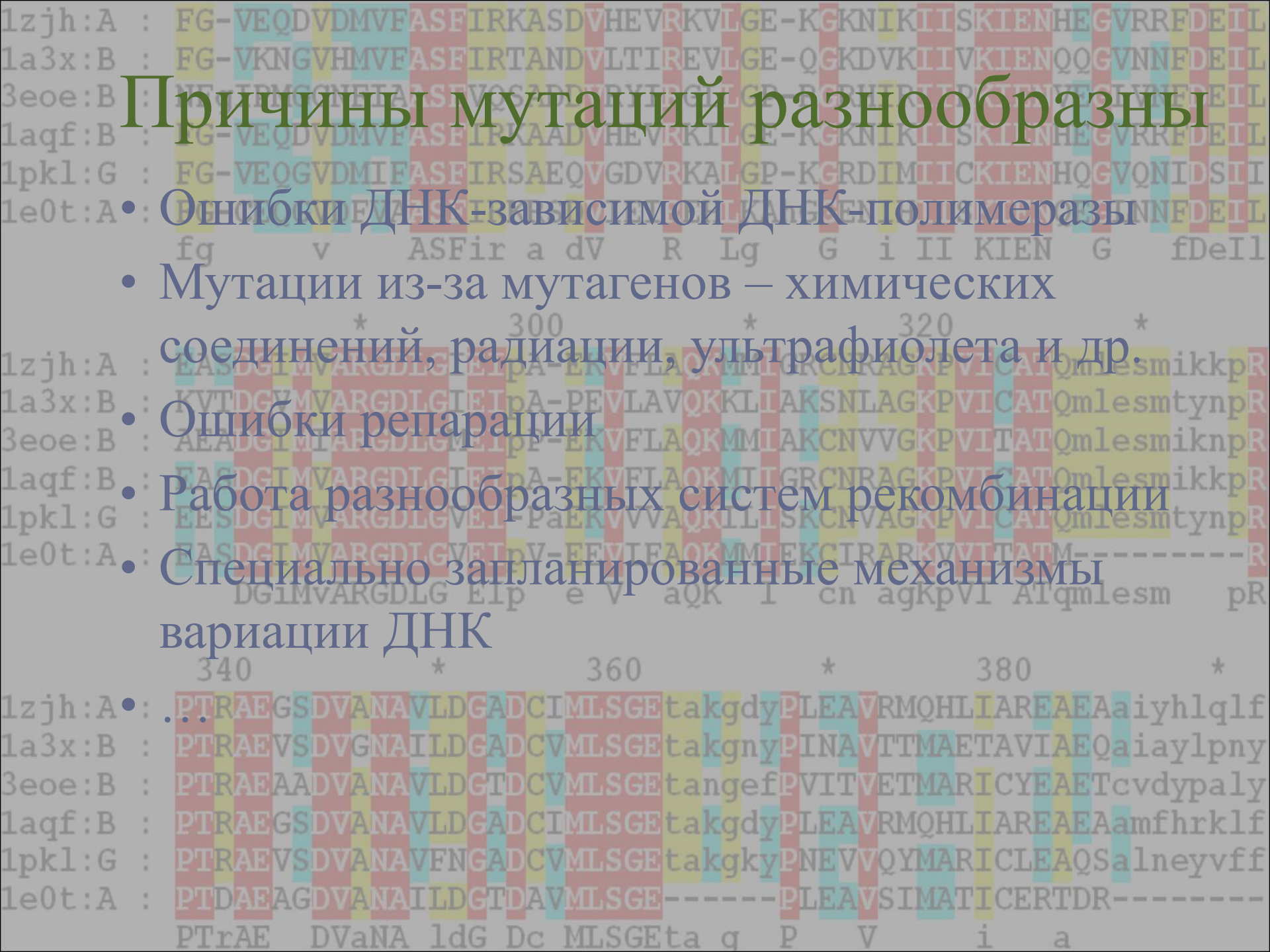
- Замена нуклеотида на другой
- Делеция одного или нескольких подряд нуклеотидов
- Вставка одного или нескольких подряд нуклеотидов

- 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<sub>α</sub> атомы находятся в участках полипептидной цепи сходной конформации (структура)
- играют сходную роль в белке (функция)



# Пример выравнивания

	*	240	*	260	*	280	
1zjh:A	:	FG-VEQDVMVFASFI	RKASDVHEVRKVL	LGE-KGKNIKIIS	SKIENHEG	VRRFDEIL	: 260
1a3x:B	:	FG-VKNGVHMVFASFI	RTANDVLTIREV	LGE-QGKDVKIIV	KIENQQGV	NNFDEIL	: 254
3eoe:B	:	NFGIPMGCNFI	ASFVQSADDVRYI	RGLLGP-RGRHIRI	IPKIENVEGL	VNFDEIL	: 249
1aqf:B	:	FG-VEQDVMVFASFI	RKAADVHEVRKI	LGE-KGKNIKIIS	SKIENHEG	VRRFDEIL	: 272
1pkl:G	:	FG-VEQGVDMIFASFI	RSAEQVGDVRKA	LGP-KGRDIMI	ICKIENHQGV	QNIDSI	: 252
1e0t:A	:	FG-CEQGVDFVAASFI	RKRSDVIEIREH	LKAhGGENIH	IISKIENQEG	LNNFDEIL	: 230
		fg	v	ASfir a dV	R	Lg	G i II KIEN G fDeil

	*	300	*	320	*		
1zjh:A	:	EASDGI	MVARGDLGIEI	pA-EKVFLAQKMMI	IGRCNRAGKPVICAT	QmlesmikkpR	: 315
1a3x:B	:	KVTDGVMVARGDLGIEI	pA-PEVLAVQK	KLIAKSNLAGKPVICAT	QmlesmtynpR		: 309
3eoe:B	:	AEADGIMIARGDLGMEI	pP-EKVFLAQKMMI	IAKCNVVGKPVITAT	QmlesmiknpR		: 304
1aqf:B	:	EASDGI	MVARGDLGIEI	pA-EKVFLAQKMMI	IGRCNRAGKPVICAT	QmlesmikkpR	: 327
1pkl:G	:	EESDGI	MVARGDLGVEI	-PaEKVVVAQKIL	ISKCNVAGKPVICAT	QmlesmtynpR	: 307
1e0t:A	:	EASDGI	MVARGDLGVEI	pV-EEVIFAQKMMI	EKCIRARKVVITATM	-----R	: 276
		DGI	MvARGDLG EIp	e V aQK	I	cn agKpVI ATqmlesm pR	

	340	*	360	*	380	*		
1zjh:A	:	PTRAE	GS	DVANAVLDGAD	CIMLSGETakgdy	PLEAVRMQH	LIAREAEAaiyhlqlf	: 371
1a3x:B	:	PTRAE	VSDVGNAILD	GADCVMLSGETakgny	PINAVTTMA	ETAVIAEQaiaylpny		: 365
3eoe:B	:	PTRAE	AADVANAVLDG	TDCVMLSGETangef	PVITVETMAR	ICYEAETcvdypaly		: 360
1aqf:B	:	PTRAE	GS	DVANAVLDGAD	CIMLSGETakgdy	PLEAVRMQH	LIAREAEAamfhrklf	: 383
1pkl:G	:	PTRAE	VSDVANAVFNG	ADCVMLSGETakgky	PNEVVQYMAR	ICLEAQSalneyvff		: 363
1e0t:A	:	PTDAE	AGDVANAILD	GTDAVMLSGET	-----PLEAVS	SIMATICERTDR	-----	: 318
		PtrAE	DVaNAlDg	Dc MLsGETa	g	P	V i a	



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-VEOGVDMT FASFIRSAEOVGDVRKALGP-KGRDIMTICKTENHOGVONTDSTI  
1e0t:A : FG-CEGSDVDFASFIRKASDVHEVRKVLGE-KGKNIKII SKIENHEGVRRFDEIL  
fg v ASFir a dV R Lg G i II KIEN G fDeIl  
\* 300 \* 320 \*

Что делать с остатками, которые  
не должны выравниваться?

1zjh:A : EASDGI MVARGDLGIEI pA-EKVFLAQKMMI GRCNRAGKPVICATQmlesmikkpR  
1a3x:B : KVTDGVMVARGDLGIEI pA-PEVLAVQKKLI AKSNLAGKPVICATQmlesmtynpR  
3eoe:B : AEADGIMIARGDLGMEI pP-EKVFLAQKMMI AKCNVVGKPVITATQmlesmiknpR  
1aqf:B : EASDGI MVARGDLGIEI pA-EKVFLAQKMMI GRCNRAGKPVICATQmlesmikkpR  
1pk1:G : EESDGI MVARGDLGVEI -PaEKVVVAQKILISKCNVAGKPVICATQmlesmtynpR  
1e0t:A : EASDGI MVARGDLGVEI pV-EEVIFAQKMMIEKCIRARKVVITATM-----R  
DGiMvARGDLG EI p e V aQK I cn agKpVI ATqmlesm pR

340 \* 360 \* 380 \*  
1zjh:A : PTRAE GSDVANAVLDGADCIMLSGE takgdy PLEAVRMQH LIAREAE Aaiyhqlf  
1a3x:B : PTRAE VSDVGNAILD GADCVMLSGE takgny PINAVTTMAETAVIAEQ aiaylpny  
3eoe:B : PTRAE AADVANAVLDGTDCVMSLGE tange fPVITVETMARICYEAET cvdypaly  
1aqf:B : PTRAE GSDVANAVLDGADCIMLSGE takgdy PLEAVRMQH LIAREAE Aamfhrklf  
1pk1:G : PTRAE VSDVANAVFNGADCVMLSGE takgky PNEVVQYMARICLEAQSalneyvff  
1e0t:A : PTDAEAGDVANA ILDGTDAVMSLGE ----- PLEAVSIMATICERTDR-----  
PTrAE DVaNA ldG Dc MLSGETa q 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 GSDVANAILD GADCVMLSGEtakgnyPINAVTTMAETA VIAEQaiaylpny
3eoe:B : PTRAE AADVANA VLDGTDCVMLSGETangefPVITVETMARICYEAETcvdypaly
1aqf:B : PTRAE GSDVANAVLDGADCIMLSGEtakgdyPLEAVRMQHLLIAREAEAAamfhrklf
1pk1:G : PTRAE VSDVANAVFNGADCVMLSGETakgkyPNEVVQYMARICLEAQSalneyvff
1e0t:A : PTDAEAGDVANA I LDGTDAVMLSGE-----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	:	AGKQSLKQDPDKEANPVKDI 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 LEAQSalneyvff  
 3eoe:B : PTRAE GSDVANAVLDGADCIMLSGE takgdy PLEAVRMQH LIAREAE Aamfhrklf  
 1aqf:B : PTRAE GSDVANAVLDGADCIMLSGE takgdy PLEAVRMQH LIAREAE Aamfhrklf  
 1pk1:G : PTRAE VSDVANAVFNGADCIMLSGE takgky PNEVVQYMARIC LEAQSalneyvff  
 1e0t:A : PTDAEAGDVANA ILDTDAVMSLGE ----- PLEAVSIMATI CERTDR-----  
 PTrAE DVaNA ldG Dc MLSGETa q P V i a



# Что есть что (редактор GeneDoc)

Номер столбца выравнивания

```

                *                20                *
MTA1_YEAST : ---KSSIS P Q A R A F L E Q V E R R K --- Q S L N S : 24
MAT2_YEAST : K P Y R G H R F T K E N V R I L E S W E A K N I E N P Y L D T : 31
                3 2                L E    F    4                L13
    
```

```

                40                *                60
MTA1_YEAST : K E K E E V A K K C G I T P L Q V R V W F I N K R M R S K - : 53
MAT2_YEAST : K G L E N I M K N T S L S R I Q I K N W V S N R R R K E K T : 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 : NFgIPMGCNFIAASFVQSADDVRYIRGLLGP-RGRHIRIIPKIENVEGLVNFDEIL  
 1aqf:B : FG-VEQDVMVFASFIRKAADVHEVRKILGE-KGKNIKII SKIENHEGVRRFDEIL  
 1pk1:G : FG-VEQGVDMI FASFIKIQVGIYFAIYRGRDIMIICKIENHQGVQNI DSI  
 1e0t:A : FG-CEQGVDFVAASFIRKRSDVIEIREHLKAhGGENIHII SKIENQEG LNNFDEIL  
 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 : PTDAEAGDVANA ILDGTDVAVMLSGE-----PLEAVSIMATI CERTDR-----  
 PTrAE DVaNA ldG Dc MLSGETa q P V i a