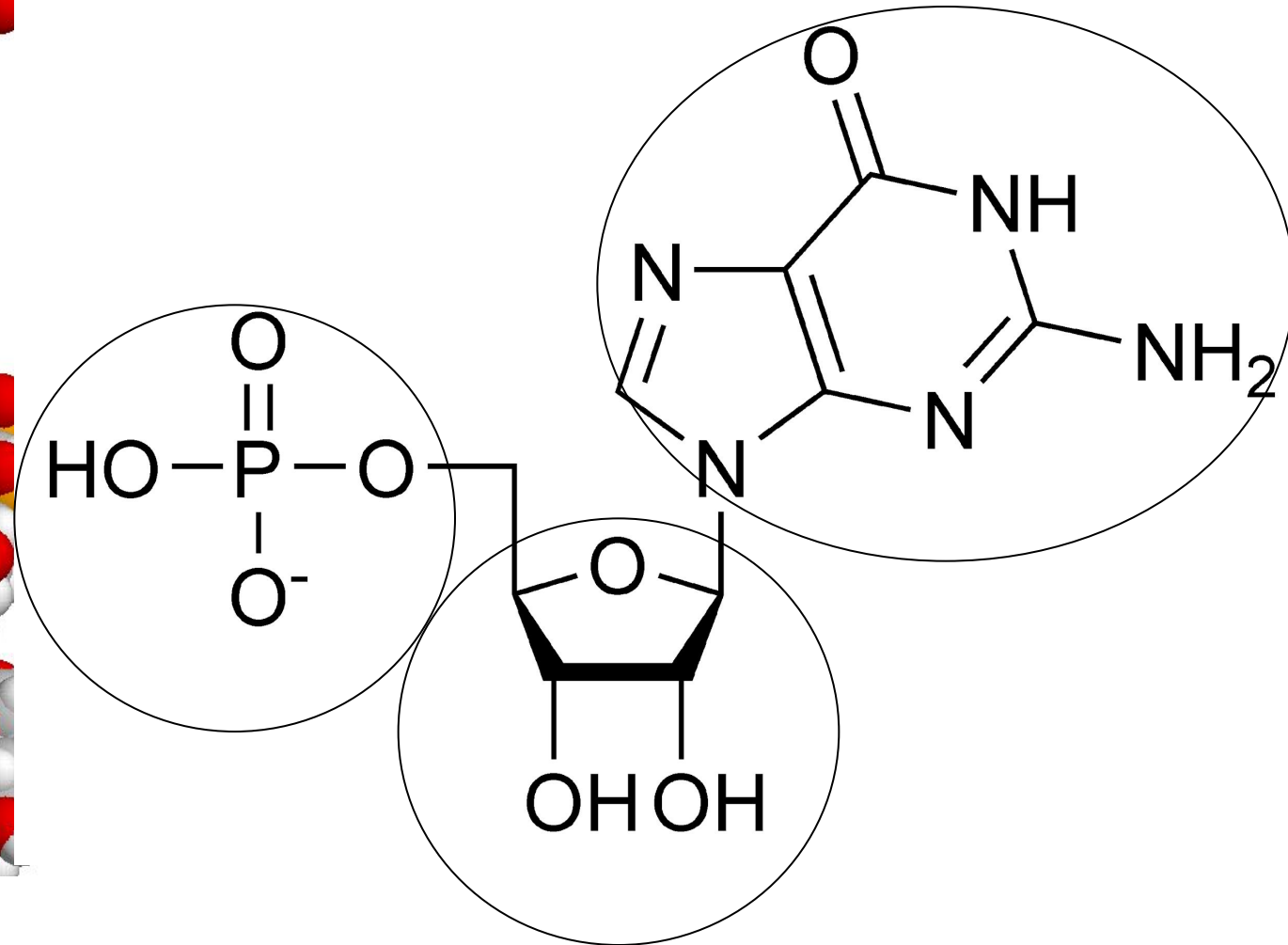
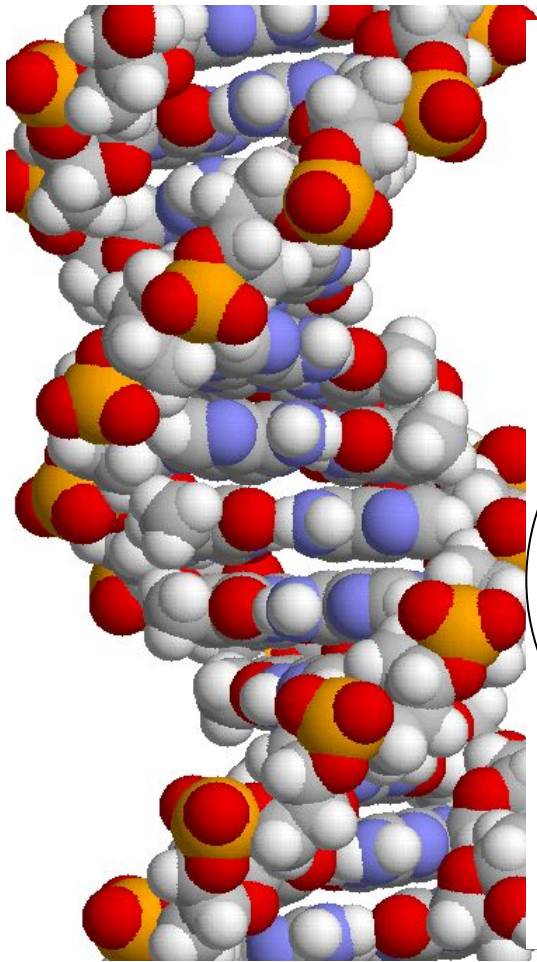
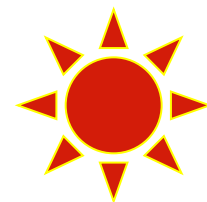
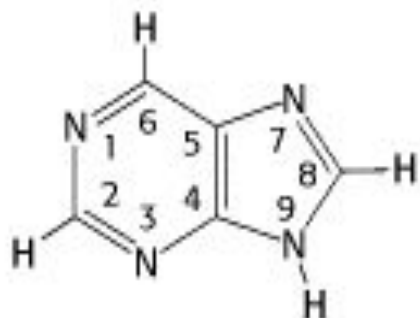


Нуклеиновые кислоты - линейные гетерополимеры нуклеотидов

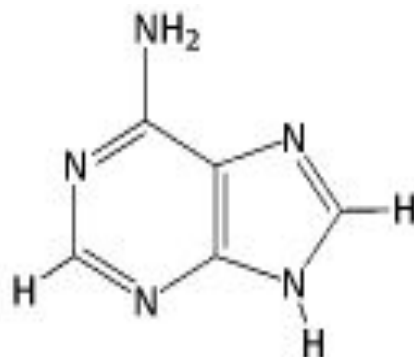




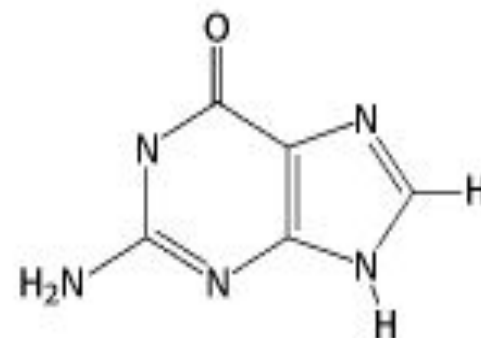
PURINES



Purine

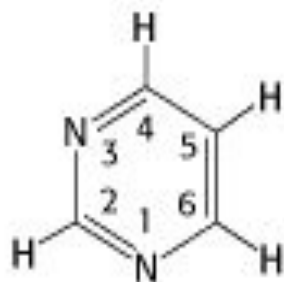


Adenine

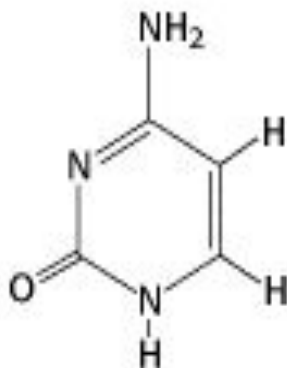


Guanine

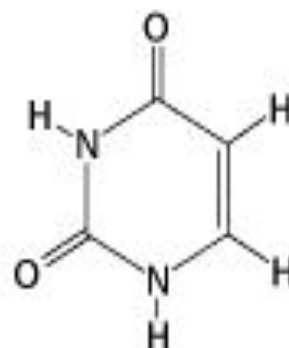
PYRIMIDINES



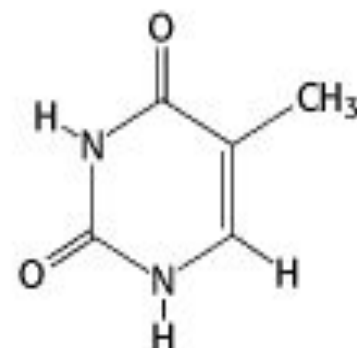
Pyrimidine



Cytosine

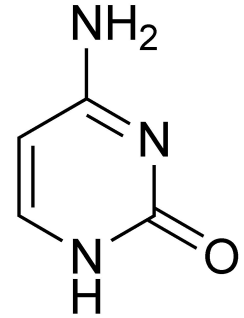


Uracil

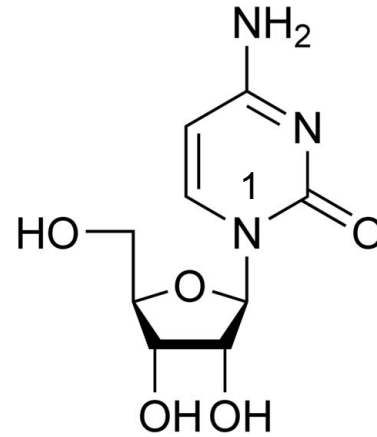


Thymine

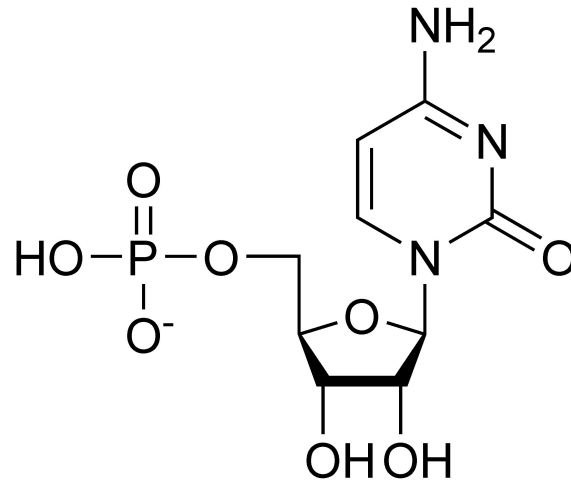
Азотистое основание цитозин

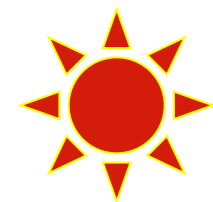
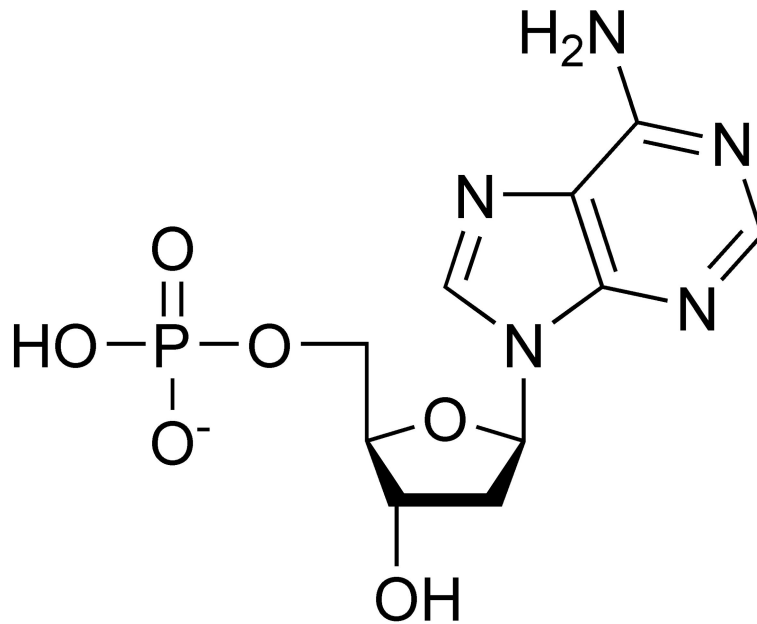
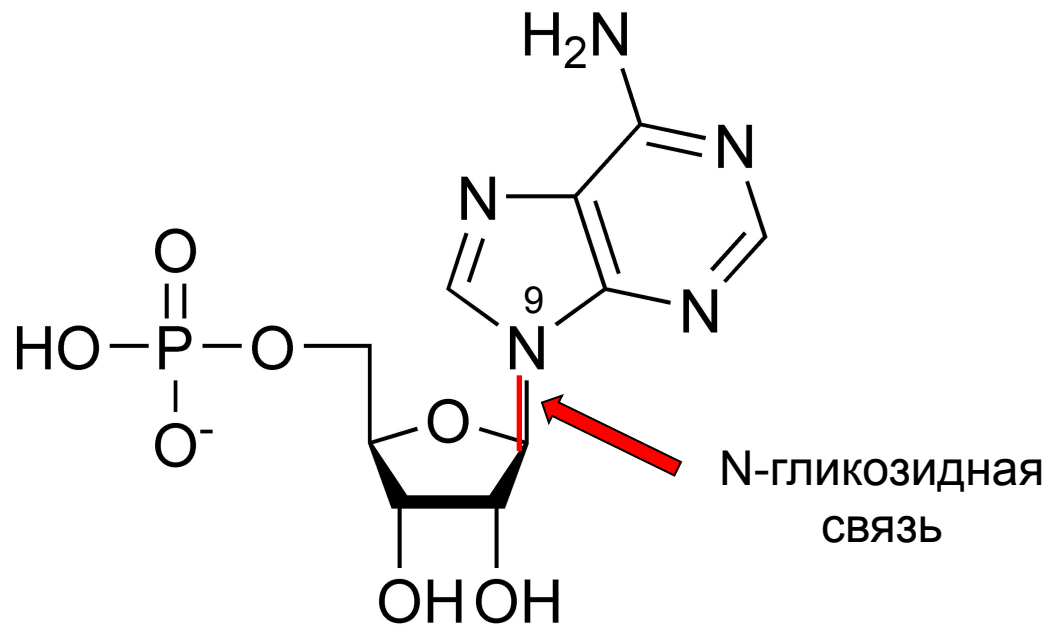


Нуклеозид цитидин



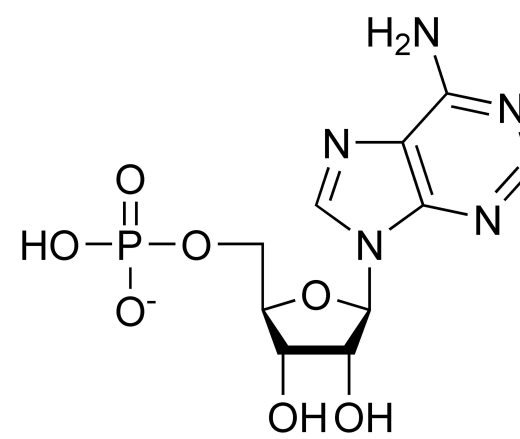
Нуклеотид цитидинмонофосфат (ЦМФ)



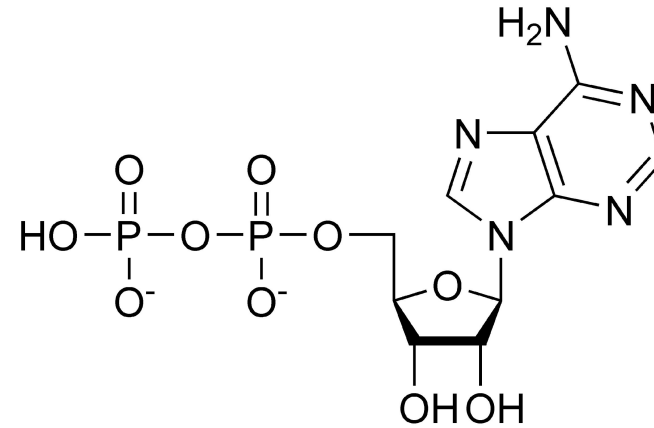




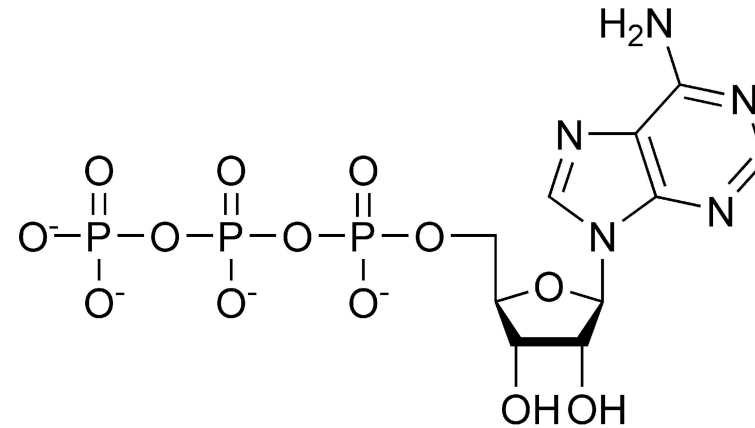
Аденозин-5'-монофосфат
(АМФ),



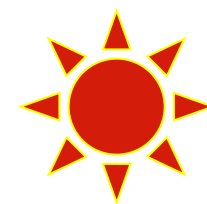
Аденозин-5'-дифосфат (АДФ),



Аденозин-5'-трифосфат (АТФ)



Номенклатура стандартных азотистых оснований, нуклеозидов и нуклеотидов

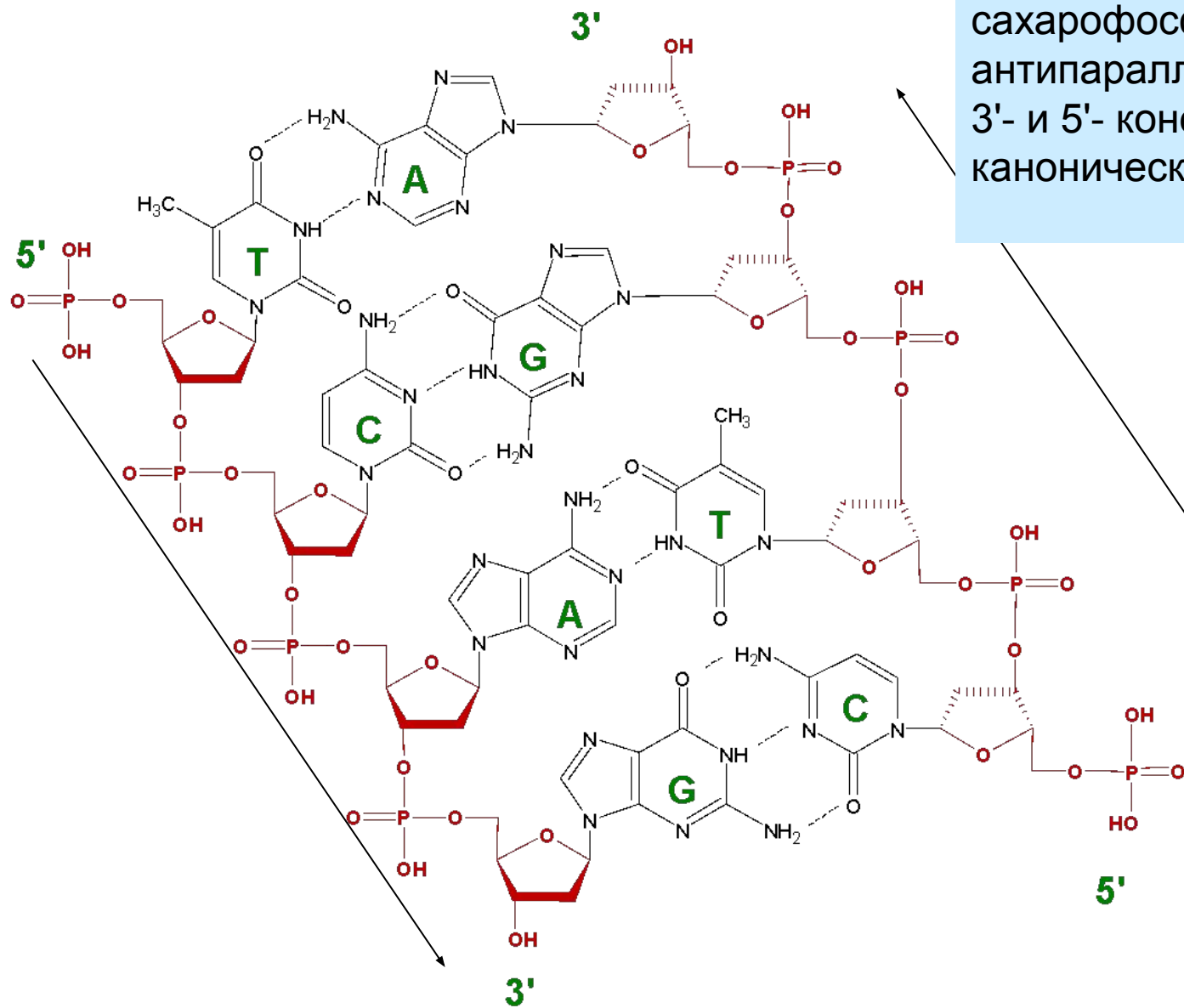


РНК		
Азотистое основание	Рибонуклеозид	Рибонуклеотид = рибонуклеозид-5'-фосфат
Аденин(A)	Аденозин (Adenosine)	аденозин-5'-фосфат, адениловая кислота, АМФ (Adenylate, AMP)
Гуанин (G)	Гуанозин (Guanosine)	гуанозин-5'-фосфат, гуаниловая кислота, ГМФ (Guanylate, GMP)
Урацил (U)	Уридин (Uridine)	уридин-5'-фосфат, уридиловая кислота, УМФ (Uridylate, UMP)
Цитозин (C)	Цитидин (Cytidine)	цитидин-5'-фосфат, цитидиловая кислота, ЦМФ (Cytidylate, CMP)
ДНК		
Азотистое основание	2'-дезоксирибонуклеозид	Дезоксирибонуклеотид = 2'-дезоксирибонуклеозид-5'-фосфат
Аденин(A)	2'-дезоксиаденозин (Deoxyadenosine)	2'-дезоксиаденозин-5'-фосфат, 2'-дезоксиадениловая кислота, (Deoxyadenylate, dAMP)
Гуанин (G)	2'-дезоксигуанозин (Deoxyguanosine)	2'-дезоксигуанозин-5'-фосфат, 2'-дезоксигуаниловая кислота (Deoxyguanylate, dGMP)
Тимин (T)	Тимидин (Thymidine)	Тимидиловая кислота (Thymidylate, TMP)
Цитозин (C)	2'-деоксицитидин (Deoxycytidine)	2'-деоксицитидин-5'-фосфат, 2'-деоксицитидиловая кислота, (Deoxycytidylate, dCMP)

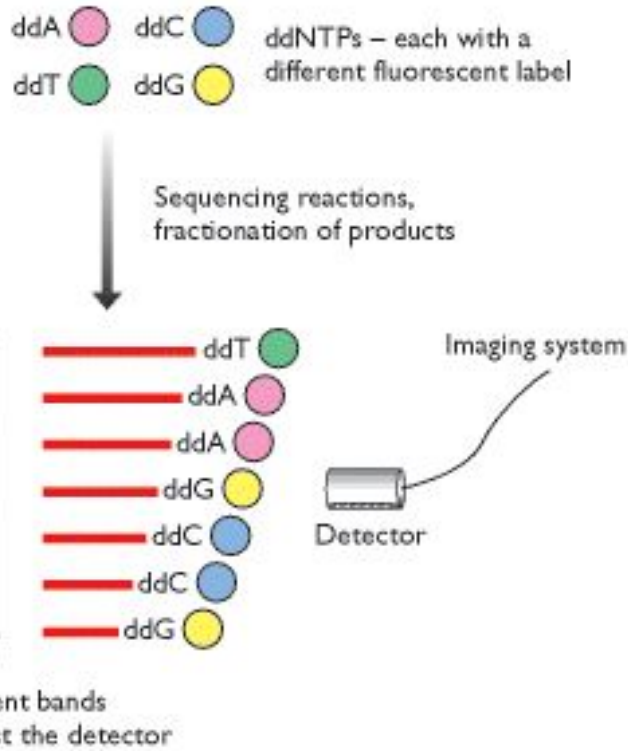
ДНК

Повторяем:

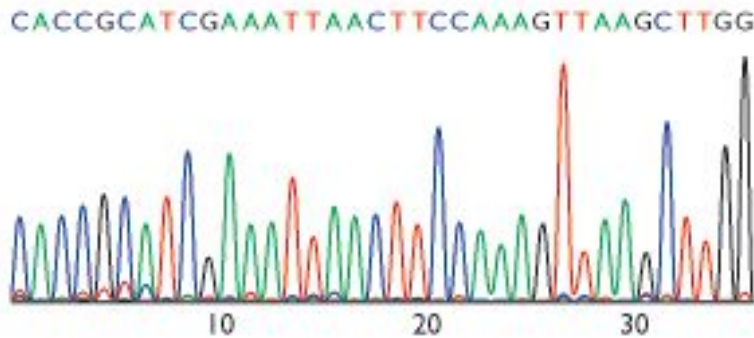
фосфодиэфирные связи,
сахарофосфатный остов,
антипараллельные цепи,
3'- и 5'- конец,
канонические пары.



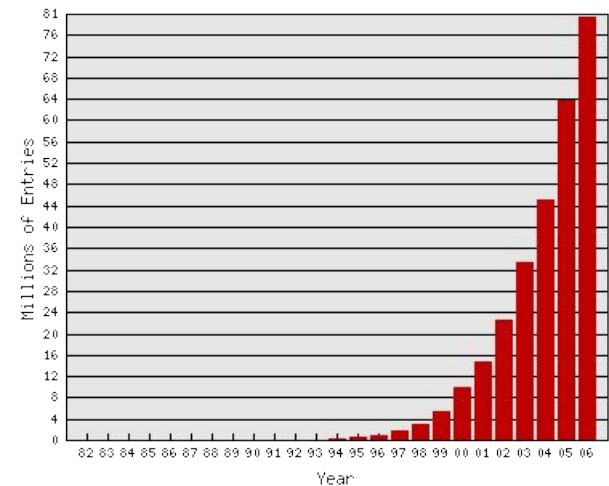
(A)



(B)



Разработка эффективных методов секвенирования привела к быстрому росту известных последовательностей



Как записывают последовательности нуклеиновых кислот ?

1. Последовательность = последовательность однобуквенных символов.
Никаких дефисов и обозначений фосфодиэфирных связей.
2. Одни и те же однобуквенные символы для последовательностей РНК и ДНК (при записи РНК обычно 'U' \Rightarrow 'T').
Любая последовательность по умолчанию считается ДНК (т.е. полимером 2'-дезоксирибонуклеотидов).
3. Одни и те же символы используются для обозначения азотистых оснований, нуклеозидов и нуклеотидов
Допустимы заглавные и строчные буквы, хотя рекомендованы заглавные.
4. **Последовательность записывается в направлении 5'→3'**

Пример:

5'-CTCGAC-3'

Nomenclature Committee of the International Union of Biochemistry (NC-IUB)
Nomenclature for incompletely specified bases in nucleic acid sequences
Recommendations 1984
Biochem. J. (1985) 229, 281-286

Описание сайтов
связывания
с регуляторными белками

Восстановление предковой
последовательности

Описание сайтов
рестрикции

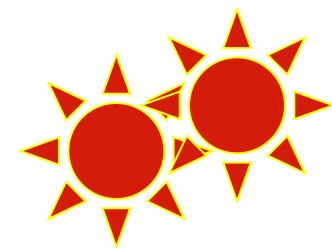
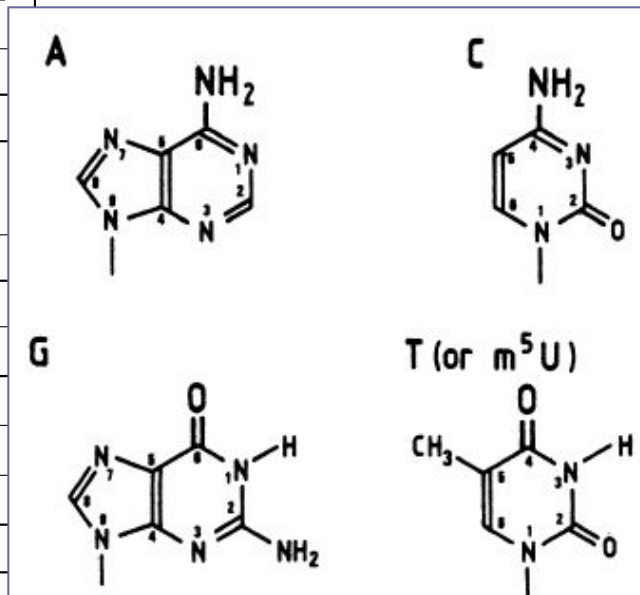
Описание вырожденности
генетического кода

```
1 ----TGGtACAGCATTTGCA
2 ----TGGCACAGCcTTcGCA
3 ----TGGCAttaGcTTTGCA
4 ----TGGCACgatAgTcGCA
5 ----TGGCACAGGcTgTGct
6 ----TGGCACAGatTTcGct
7 ----TGGtACAaGAccTGCA
8 ----TGGCACgattTTTtCA
9 ----TGGCAagcaAaTTGCA
10 ----gGGCgCAGCcTTcGCA
11 ----TGGtAtcGCAaTTGct
12 ----TGGagCgcGAaTTGCA
13 ----TGGtAtgttcccTGCA
```

CONSENSUS TGGCACrrsmtTTGCA

Общепринятые однобуквенные обозначения для стандартных азотистых оснований (остатков нуклеозидов и нуклеотидов) и вырожденных позиций в выравниваниях нуклеиновых кислот

Символ	Расшифровка	Происхождение обозначения
G	G, guanine (гуанин)	
A	A, adenine, (аденин)	
T	T/U, thymine/uracil (тимин в ДНК и урацил в РНК)	
C	C, cytosine (цитозин)	
R	A или G, purine(пурины)	pu R ine
Y	C или T или U, pyrimidine (пиримидины)	p Y rimidine
M	A или C	a M ino
K	G или T	K eto
S	G или C	Strong interaction (3 H bonds)
W	A или T	Weak interaction (2 H bonds)
H	A или C или T, но не G	в алфавите 'H' следует за 'G'
B	G или T или C, но не A	'B' следует за 'A'
V	G или C или A, но не T или U	'V' следует за 'U'
D	G или A или T, но не C	'D' следует за 'C'
N	G или A или T или C	a N y



ttttacctcttttagtgatattgtgatagagcaaaaatcccgcacattgtgctgggattgttttaaaccttgttgatttaattttcaatcgcttctttataaagaagtagtggtgccc
acaacactcacattgcatatcaatacggcctttatgttcggctaataatttcgcaatttctcatcagagatgagcagtagatgcagaactagaacgctcagcagagcagccaca
gaaaaattgtacatcttgtgctggataaagattaacggtttctctgtagataaacgataggagtaacttctgacgggagaccaataattcttcatctttactgttgctgagc
gtagttaaatgctcaaaatcttctggtgtaccagaaccatcaggcataattgtaataacatacctgctgccaactggctgacctcatattctccagtacgaataattaattgagttg



GenBank



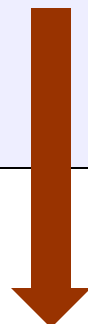
EMBL



DDBJ



компьютерный поиск гена, трансляция и компьютерная аннотация



Базы данных научной литературы



~2 500 000 последовательностей

Экспертиза



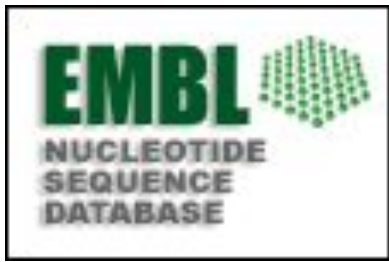
~200 000 последовательностей



UniParc
(UniProt Archive)

UniRef
(UniProt non-redundant Reference databases)





The EMBL Nucleotide Sequence Database (также просто БД EMBL)