

# Молекулярная биология



# Некоторые параметры молекул ДНК и белка:

- Один шаг это полный виток спирали ДНК-поворот на  $360^\circ$
- Один шаг составляют **10** пар нуклеотидов
- Длина одного шага - **3,4** нм
- Расстояние между двумя нуклеотидами - **0,34** нм
- Молекулярная масса одного нуклеотида - **345** г/моль
- Молекулярная масса одной аминокислоты – **100** г/мол
- В молекуле ДНК:  **$A+G=T+C$**  (Правило Чаргаффа)
- Комплементарность нуклеотидов:  **$A=T$ ;  $G=C$**
- Цепи ДНК удерживаются водородными связями, которые образуются между комплементарными азотистыми основаниями: аденин с тиминном соединяются **2** водородными связями, а гуанин с цитозином **тремя**.
- В среднем один белок содержит **400** аминокислот

## Задачи

- Последовательность нуклеотидов одной из цепей двух молекул ДНК следующая:
  - 1) ...А А А Т Ц Т Т Т Т Т Т А Т А А Т Г Т...
  - 2) ...А Г Г Ц Ц Г Ц Г Ц Т Ц Ц Ц Ц Г А Г Г Г...
- В какой молекуле двойная спираль устойчивее к повышению температуры?



# Задачи

- **Фрагмент молекулы ДНК состоит из 6000 нуклеотидов. Определите длину данного фрагмента ДНК.**



# Задачи

- Фрагмент молекулы ДНК состоит из 6000 нуклеотидов. Определите длину данного фрагмента ДНК.
- Решение:
- $6000 : 2 = 3000$  - длина одной цепи ДНК.
- 0,34 - длина одного нуклеотида
- $3000 * 0,34 \text{ нм} = 1020 \text{ нм}$
- **Ответ: 1020 нм.**

## Задачи

- **Фрагмент молекулы ДНК состоит из 4500 нуклеотидов. Определите длину данного фрагмента ДНК.**



# Задачи

- **Фрагмент молекулы ДНК состоит из 3000 нуклеотидов, из них цитидиловых нуклеотидов 650. Определите длину данного фрагмента и количество адениловых, тимидиловых и гуаниловых нуклеотидов.**

## Задачи

- Фрагмент молекулы ДНК состоит из 3000 нуклеотидов, из них цитидиловых нуклеотидов 650. Определите длину данного фрагмента и количество адениловых, тимидиловых и гуаниловых нуклеотидов.

- Решение:

- Длина одной пары нуклеотидов 0.34 нм.
- $3000/2 \times 0.34 = 510$  нм.
- За закону Чаргафа:
- $A = T$
- $G = Ц = 650$  нуклеотидов
- $G + Ц = 1300$ .
- $3000 - 1300 = 1700$ .
- $1700 = A + T$
- $A = T = 850$  нуклеотидов.

## Задачи

- **Фрагмент молекулы ДНК состоит из 5760 нуклеотидов, из них тимидиловых нуклеотидов 1125. Определите длину данного фрагмента и количество адениловых, гуаниловых и цитидиловых нуклеотидов.**

## Задачи

- Определите количество водородных связей во фрагменте ДНК – ГТЦАТГГАТАГТЦЦТАТ
- Решение:
- Цепи ДНК комплементарны , значит суммы А+Т и Г+Ц равны суммам второй цепи. Между А и Т образуется 2 водородные связи, между Г и Ц – 3 СВЯЗИ.
- $A+T=10 \cdot 2=20$ ,  $G+C=7 \cdot 3=21$ ,  $20+21= 41$  водородных связей

## Задачи

- Определите количество водородных связей во фрагменте ДНК – ТЦГАГТАЦЦТАТГАТЦЦЦ



## Задачи

- Молекула ДНК состоит из 4000 нуклеотидов. Определите число полных спиральных витков в данной молекуле
- Решение:
- Полный виток или один шаг в молекуле ДНК составляют 10 пар нуклеотидов. В данной молекуле 4000 нуклеотидов, что составляет 2000 пар, следовательно:  $2000:10=200$  полных витков.
- Ответ: **200 полных спиральных витков в молекуле ДНК**

## Задачи

- Молекула ДНК состоит из 3600 нуклеотидов. Определите число полных спиральных витков в данной молекуле.



## Задачи

- Длина участка молекулы ДНК составляет 850 нм. Определите количество нуклеотидов в одной цепи ДНК



## Задачи

- Длина участка молекулы ДНК составляет 850 нм. Определите количество нуклеотидов в одной цепи ДНК
- Решение:
- Расстояние между двумя нуклеотидами - 0,34 нм
- $850/0,34 = 2500$  нуклеотидов

## Задачи

- **Фрагмент молекулы ДНК состоит из 1000 нуклеотидов, из них адениловых нуклеотидов 23%. Определите количество гуаниловых, тимидиловых и цитидиловых нуклеотидов.**

## Задачи

- Фрагмент молекулы ДНК состоит из 1000 нуклеотидов, из них адениловых нуклеотидов 23%. Определите количество гуаниловых, тимидиловых и цитидиловых нуклеотидов.
- Решение:
- 1000 нуклеотидов это 100%, определим, сколько нуклеотидов составляет 23% адениловых:  
 $23 \cdot 1000 : 100 = 230$  нуклеотидов,
- $A = T = 230$ ,  $1000 - 460 (A + T) = 540 (G + C)$ ;  $G = C = 270 (540 : 2)$
- Ответ: **G-270, T-230, Ц-270 нуклеотидов в фрагменте ДНК**

## Задачи

- **Фрагмент молекулы ДНК состоит из 2000 нуклеотидов, из них гуаниловых нуклеотидов 18%. Определите количество адениловых, тимидиловых и цитидиловых нуклеотидов.**

## Задачи

- Определите молекулярную массу фрагмента ДНК если он состоит из 900 нуклеотидов.



## Задачи

- Определите молекулярную массу фрагмента ДНК если он состоит из 900 нуклеотидов.
- Решение:
- Молекулярная масса одного нуклеотида составляет 345 г/моль, следовательно молекулярная масса фрагмента ДНК:  $900 \cdot 345 = 310500$  г/моль
- Ответ: молекулярная масса фрагмента ДНК- **310500 г/моль**

## Задачи

- Определите молекулярную массу фрагмента ДНК если он состоит из 1400 нуклеотидов.



## Задачи

- Молекулярная масса молекулы ДНК составляет 17250 г/моль. Определите количество нуклеотидов в молекуле и её длину.



## Задачи

- Молекулярная масса молекулы ДНК составляет 17250 г/моль. Определите количество нуклеотидов в молекуле и её длину.

- Решение:

- Молекулярная масса одного нуклеотида - 345 г/моль
- $17250 / 345 = 50$  нуклеотидов
- Всего в молекуле ДНК  $50 / 2 = 25$  пар нуклеотидов
- Расстояние между двумя нуклеотидами - 0,34 нм
- $25 \times 0,34 = 8,5$  нм

## Задачи

- Молекулярная масса молекулы ДНК составляет 27600 г/моль. Определите количество нуклеотидов в молекуле и её длину.



F (Forward)



5' ЦЦГТАЦГТГААТГЦ



3' ГГЦАТГЦАЦТТАЦГ



R (Reverse)

# Вторичная структура ДНК

## Правила расположения нуклеотидов в ДНК:

5' CATGTA 3'

3' GTACAT 5'



CATGTA

### Homo sapiens insulin (INS), transcript variant 3, mRNA

NCBI Reference Sequence: NM\_001185098.2

[GenBank](#) [Graphics](#)

```
>NM_001185098.2 Homo sapiens insulin (INS), transcript variant 3, mRNA
AGCCCTCCAGGACAGGCTGCATCAGAAGAGGCCATCAAGCAGGTCTGTTCCAAGGGCCTTTGCGTCAGGT
GGGCTCAGGATTCCAGGGTGGCTGGACCCAGGCCCCAGCTCTGCAGCAGGGAGGACGTGGCTGGGCTCG
TGAAGCATGTGGGGGTGAGCCCAGGGGCCCAAGGCAGGGCACCTGGCCTTACGCCTGCCTCAGCCCTGC
CTGTCTCCCAGATCACTGTCCTTCTGCCATGGCCCTGTGGATGCGCCTCCTGCCCTGCTGGCGCTGCTG
GCCCTCTGGGGACCTGACCCAGCCGACGCTTTGTGAACCAACACCTGTGCGGCTCACACCTGGTGGAAAG
CTCTCTACCTAGTGTGCGGGGAACGAGGCTTCTTCTACACACCCAAGACCCGCCGGGAGGCAGAGGACCT
GCAGGTGGGGCAGGTGGAGCTGGGCGGGGGCCCTGGTGCAGGCAGCCTGCAGCCCTTGGCCCTGGAGGGG
TCCCTGCAGAAGCGTGGCATTGTGGAACAATGCTGTACCAGCATCTGCTCCCTCTACCAGCTGGAGAACT
ACTGCAACTAGACGCAGCCCGCAGGCAGCCCCACACCCGCCGCTCCTGCACCGAGAGAGATGGAATAAA
GCCCTTGAACCAGC
```

## F (Forward)

- 5' CCGTACGTGAATGC 3'
- 5' GCATTCACGTACGG 3'

## R (Reverse)

- 5' GCATTCACGTACGG 3'

## C (Complement)

CGTAAGTGCATTGCC

## RC (Reverse-Complement)

CCGTACGTGAATGC

# Выравнивание генетических последовательностей

- В эволюции генетических последовательностей происходят как **замены**, так и **вставки** и **делеции**. Первым этапом филогенетического анализа является идентификация вставок и делеций, имевших место в эволюционной истории анализируемой группы последовательностей. Эту процедуру называют выравниванием (to align, alignment) последовательностей.
- **Выравнивание последовательностей** направлено на выявление гомологичных (имеющих общее эволюционное происхождение) позиций анализируемых последовательностей, установление наиболее вероятного, т.е. требующего наименьшего числа эволюционных событий, сценария эволюции анализируемой группы.

# Выравнивание генетических последовательностей

```

111111111122
123456789012345678901
1  ATACCTGCGATAGCTTCTGAT
   |||||*****
2  ATACCTGCGAAGCTTCTGAT.
```



# Выравнивание генетических последовательностей

- **Clustal** - это одна из самых широко используемых компьютерных программ для множественного выравнивания нуклеотидных и аминокислотных последовательностей (multiple sequence alignment).
- <https://www.ebi.ac.uk/Tools/msa/clustalo/>

# BLAST

- **BLAST** (англ. **B**asic **L**ocal **A**lignment **S**earch **T**ool — средство поиска основного локального выравнивания) — семейство компьютерных программ, служащих для поиска гомологов белков или нуклеиновых кислот, для которых известна первичная структура (последовательность) или её фрагмент.
- Используя BLAST, исследователь может сравнить имеющуюся у него последовательность с последовательностями из базы данных и найти последовательности предполагаемых гомологов.
- Является важнейшим инструментом для молекулярных биологов, биоинформатиков и систематиков.

# Классификация программ серии BLAST

## Нуклеотидные

- предназначены для сравнения изучаемой нуклеотидной последовательности с базой данных секвенированных нуклеиновых кислот и их участков:
- **blastn** — медленное сравнение с целью поиска всех сходных последовательностей и др.
- **megablast** — быстрое сравнение с целью поиска высоко сходных последовательностей,
- **dmegablast** — быстрое сравнение с целью поиска дивергировавших последовательностей, обладающих незначительным сходством,

# Классификация программ серии BLAST

## Белковые

- предназначены для сравнения изучаемой аминокислотной последовательности белка с имеющейся базой данных белков и их участков.
- **blastp** — медленное сравнение с целью поиска всех сходных последовательностей,
- **cdart** — сравнение с целью поиска гомологичных белков по доменной архитектуре,
- **rpsblast** — сравнение с базой данных консервативных доменов,
- **psi-blast** — сравнение с целью поиска последовательностей, обладающих незначительным сходством,
- **phi-blast** — поиск белков, содержащих определённый пользователем паттерн и др.

# Классификация программ серии BLAST

## Транслирующие

- способны транслировать нуклеотидные последовательности в аминокислотные:
- blastx — переводит изучаемую нуклеотидную последовательность в кодируемые аминокислоты, а затем сравнивает её с имеющейся базой данных аминокислотных последовательностей белков,
- tblastn — изучаемая аминокислотная последовательность сравнивается с транслированными последовательностями базы данных секвенированных нуклеиновых кислот,
- tblastx — переводит изучаемую нуклеотидную последовательность в аминокислотную, а затем сравнивает её с транслированными последовательностями базы данных секвенированных нуклеиновых кислот.

# Классификация программ серии BLAST

## Геномные

- предназначены для сравнения изучаемой нуклеотидной последовательности с базой данных секвенированного генома какого-либо организма (человека, мыши и др.)
- **magicblast** — картирует прочтения (риды) на полный геном или транскриптом.



## Задачи

1. Найти в **ГенБанке** нуклеотидную последовательность гена (части гена) инсулина человека <https://www.ncbi.nlm.nih.gov/gene/>
2. С помощью программы **blastn** найти в нуклеотидном виде гомологи этой последовательности  
<https://blast.ncbi.nlm.nih.gov/Blast.cgi>
3. Создать новый файл с расширением **.fasta** в программе **UltraEdit**
4. Выбрать среди найденных гомологов несколько от **пяти разных видов** организмов
5. Открыть их нуклеотидную последовательность в формате **FASTA**

## Задачи

6. Копировать последовательности в созданный файл
7. Открыть файл с помощью программы MEGA5
8. Выровнять последовательности в программе MEGA5
9. Найти наилучшую модель построения филогенетического дерева
10. Построить филогенетическое дерево для выбранных нуклеотидных последовательностей