

проф. доктор физ-мат наук **В.А. Любецкий:**

vasiliy.lyubetsky@math.msu.ru (мех-мат МГУ)

и <http://lab6.iitp.ru/ru/pub/>

(Институт Проблем Передачи Информации РАН).

В ближайшем семестре и учебном году я читаю курсы

«Современные методы обработки данных» и

«Математическая биология и дискретная оптимизация»

(**по понедельникам, с 16.45** по выбору кафедры, их можно сдавать как два полугодовых или как один годовой).

Я занимаюсь задачами, возникши-ми в естествознании и прикладными:

**«в Море, вокруг нас, много интересного и это можно описать
математически».**

В 20-м веке было два великих открытия (на века):

Вселенная имеет абсолютное Начало (из «тяжёлого ничто»)

$\approx 13,8$ млрд лет назад;

И

все Живое имеет абсолютное Начало ≈ 4 млрд лет назад и устроено

единообразно, суть биологической Жизни –

ДНК; РНК-полимераза – Рибосома – белок; и Эволюция

(хранить записи; читать их – чинить; и реагировать/меняться).

Последний непонятый вызов такого масштаба – **сознание и психика**

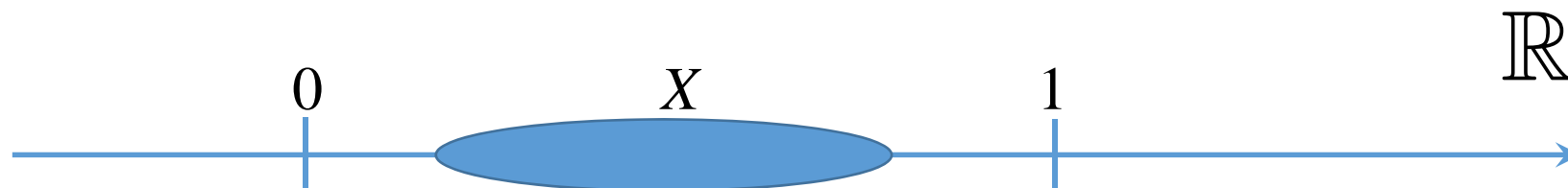
(абсолютно неразрешимые проблемы).

Со мной можно заниматься по трём связанным между собой темам (хотя они кажутся, на первый взгляд разными):

1) Теория множеств, теория моделей и нестандартный анализ:

здесь изучают **абсолютно неразрешимые проблемы**,
случайные числа,
связь вещественных и комплексных чисел.

Абсолютно неразрешимые проблемы:



1) Континуум-проблема состоит в том, верно ли:

$$\exists X (\forall n \in \mathbb{N} \neg (X \leftrightarrow \{0, 1, \dots, n\}), \neg (X \leftrightarrow \mathbb{N}), \neg (X \leftrightarrow [0, 1]))$$

2) Описано индивидуальное множество $X \subseteq [0, 1]$. Это X удивительно простое: оно – проекция дополнения аналитически(!) описанного множества в обычной плоскости чисел.

Измеримо ли оно по Лебегу?, т.е. существует ли $\int_0^1 \chi(x) dx$, где χ – характеристическая функция этого множества?

Проблемы оказались абсолютно неразрешимыми !

Более того, **разные абсолютно неразрешимые проблемы неожиданно оказались связанными между собой !**

Более того, очень похожие проблемы (мера и категория, например) оказались: одна разрешимой, а другая нет! //

Вещественное число называется случайным, если его нельзя локализовать «просто» описанным множеством меры 0.

**У кольца есть локализации, и свойства локализаций переносятся на само исходное кольцо
(=Локализации).**

**Связь вещественных и комплексных чисел, или
рациональных и алгебраических чисел (в \mathbb{R} и в \mathbb{C})**

дополняют друг друга –

оказывается, что эту связь можно перенести на кольца и

даже группы

(= Модельный компаньон).

задачи по Дескриптивной теории множеств: 1) Множество алгебраических чисел биективно с натуральным рядом N ?

2) Нигде не плотным называется такое множество X , что любое открытое множество в отрезке $[0,1]$ содержит его открытое подмножество, не содержащее точек из X . Является ли отрезок $[0,1]$ счётным объединением нигде не плотных множеств?

3) Можно ли представить отрезок $[0,1]$ как объединение двух или счётного числа множеств, каждое из которых не биективно с этим отрезком? (Или с числом слагаемых строго меньше, чем мощность $[0,1]$?)

4) Какие счётно-аддитивные неотрицательные функции можно определить на множестве $P([0,1])$ всех подмножеств отрезка $[0,1]$ как группе по сложению $+$ по модулю 1. Такие функции называются мерами. Указание. В этой группе определим фактор-множество по множеству рациональных чисел.

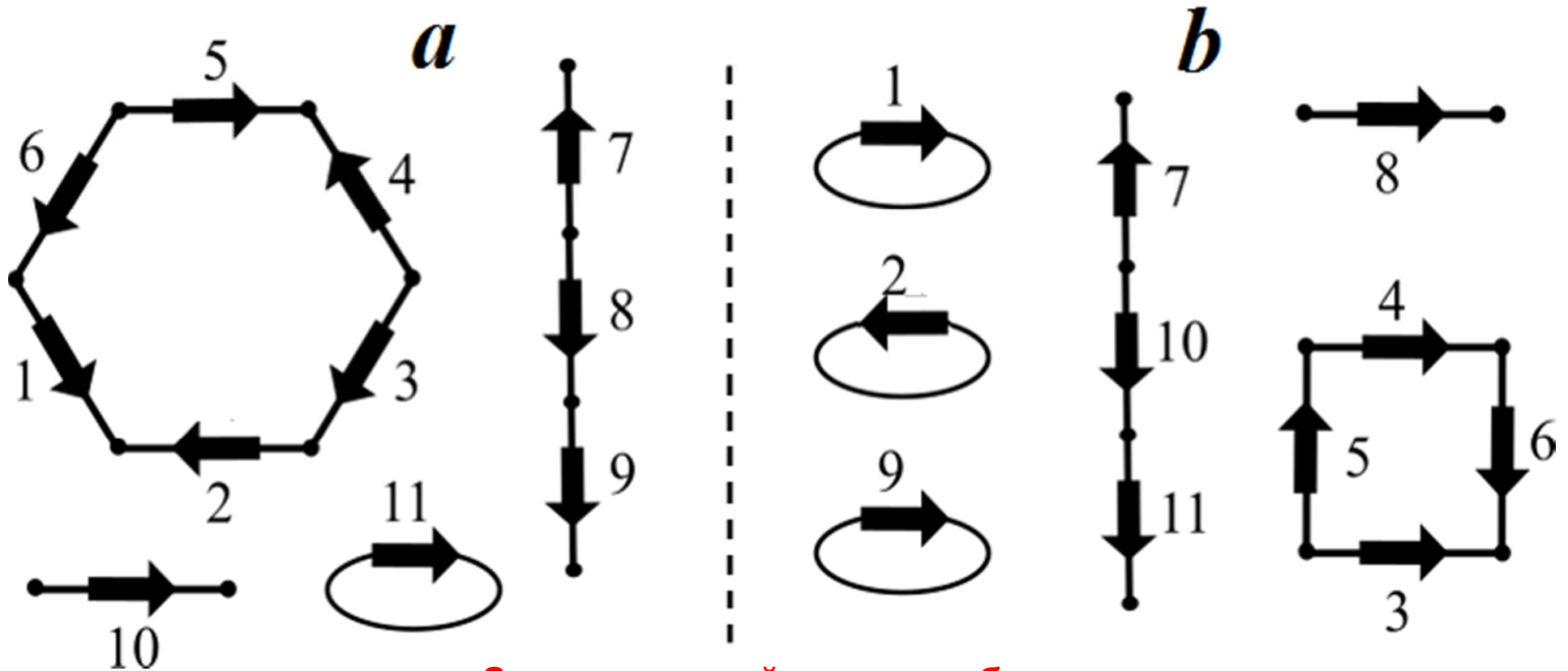
по Дескриптивная теория множеств:

5) В множестве $P(N)$ всех подмножеств натурального ряда N определим отношение эквивалентности: $X \approx Y$, если их симметрическая разность $X \Delta Y$ конечна, и каждое конечное множество эквивалентно каждому ко-конечному (т.е. его дополнение конечно). Определить действие группы $Z = (+, -)$ целых чисел на $P(N)$, орбиты которого – классы эквивалентности этого отношения и только они.

Указание. Классы эквивалентности отношения – счётные плотные множества, орбиты же действия Z видятся как дискретные множества, подобные Z .

Естественно-научные проблемы: изучение Природы – математическое описание явлений Природы (физика, биология). В таком описании центральный вопрос – оптимизация: найти $\min_{x \in X} f(x)$

2) **Дискретная оптимизация**: трудные задачи, решаемые простыми алгоритмами (по сложности вычислений линейными, квадратичными и т.п.). Например, даны две картинки **a** и **b**;



Задача кратчайшего преобразования

Найти простой алгоритм, который наименьшим числом операций (которые **заданы**) преобразует картинку **a** в картинку **b**.

Дискретная оптимизация есть
оптимизация функции, заданной на графе.

Граф – по сути произвольное конечное множество с
произвольными отношениями на нём.

Дан граф **G** *полный* (любые разные вершины соединены ребром).

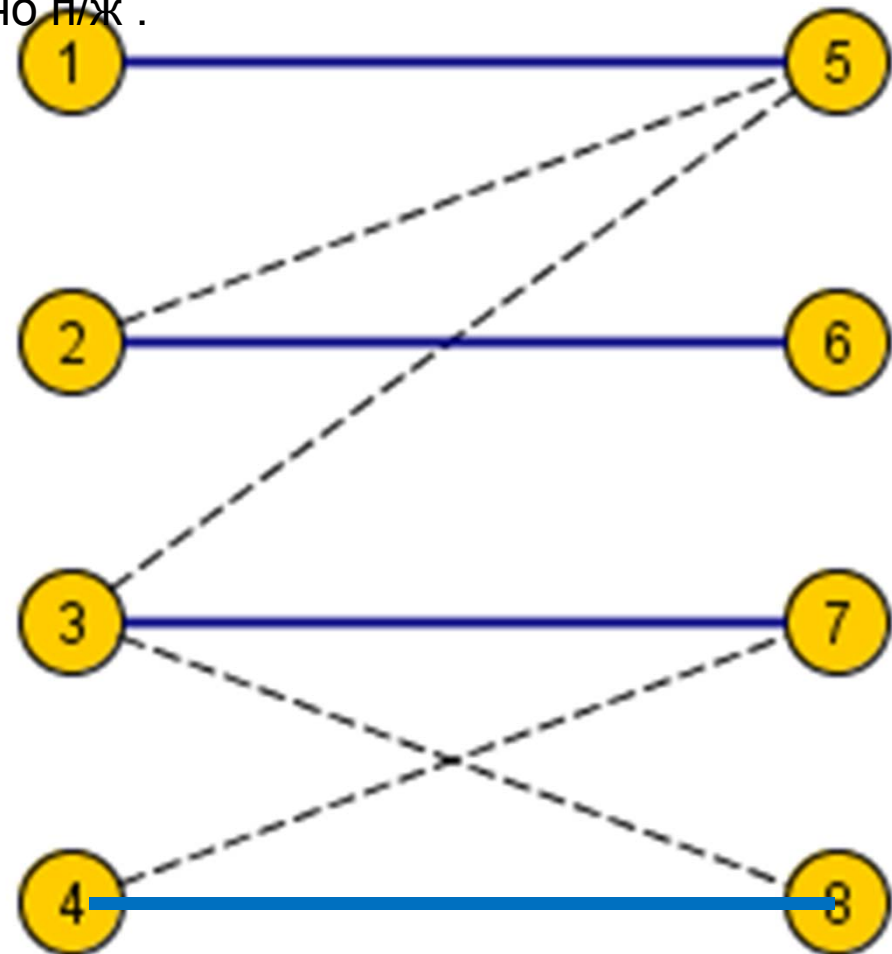
Каждому ребру приписано число – его вес. Найти **множество**

непересекающихся рёбер с максимальным *суммарным* весом.

исходные рёбра в **G** с весами 1 показаны не все, максимальное паросочетание показано п/ж .

Множество непересекающихся рёбер в **G** называется **паросочетанием** **P**. Сумма весов рёбер в **P** называется **весом P**.

Задача: найти простым алгоритмом паросочетание **P** с **максимальным весом**.



задачи по Дискретная оптимизация:

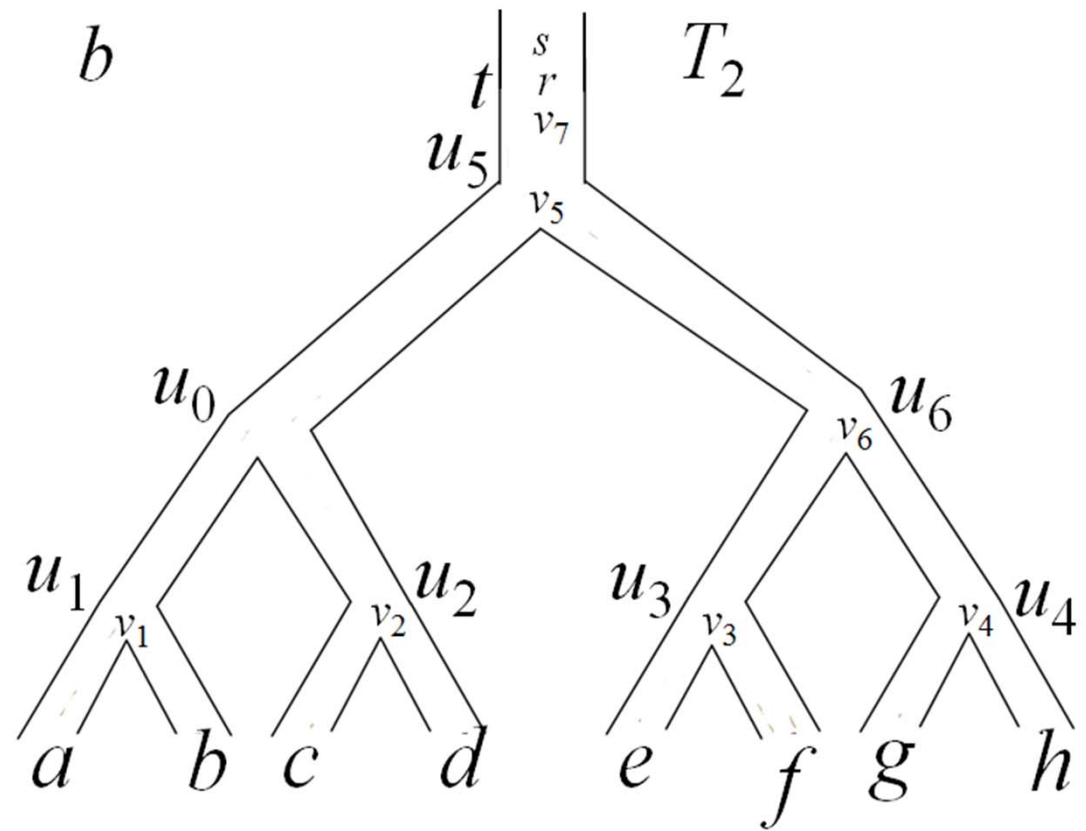
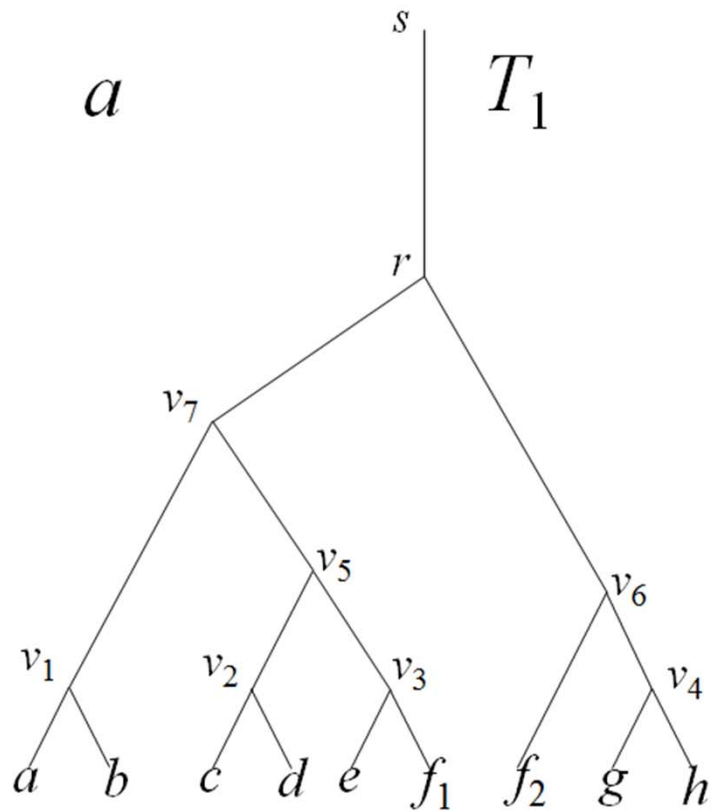
дан граф *полный* (= проведены все рёбра между разными вершинами); *паросочетанием* называется множество его попарно не пересекающихся рёбер; *весом* паросочетания называется сумма весов его рёбер. В полном графе с 4 вершинами найти **все** максимального веса паросочетания. Затем **все** минимального веса *полные* паросочетания (=любая вершина принадлежит какому-то ребру паросочетания).

Весы рёбер взять в следующих вариантах:

- 1) $(a,b)=(a,c)=(c,d)=(b,d)=1, (a,d)=(c,b)=2.$
- 2) $(a,b)=(a,c)=-1, (c,d)=(b,d)=2, (a,d)=(c,b)=1.$
- 3) $(a,b)=(c,d)=-1, (a,c)=-2, (b,d)=2, (a,d)=(c,b)=1.$
- 4) $(a,b)=-2, (c,d)=2, (a,c)=(b,d)=1, (a,d)=2, (c,b)=-1.$

Дискретная оптимизация:

задачи реконструкции вдоль дерева или внутри дерева.

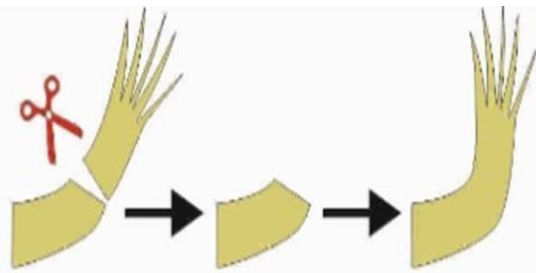




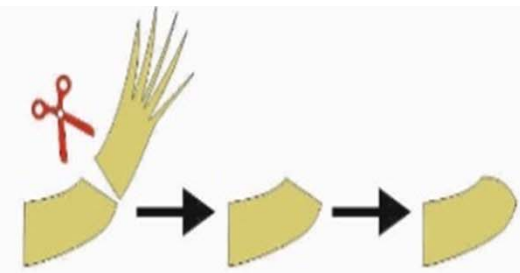
3) **Математическая биология** —

приложение математики, программирования,
биоинформатики):

**математические модели биологических процессов в
живой клетке и в эволюции организмов/видов.**



good regeneration



bad regeneration



черепахи,

утконос,

крокодилы

ехидны

сумчатые

Fishes

Amphibians

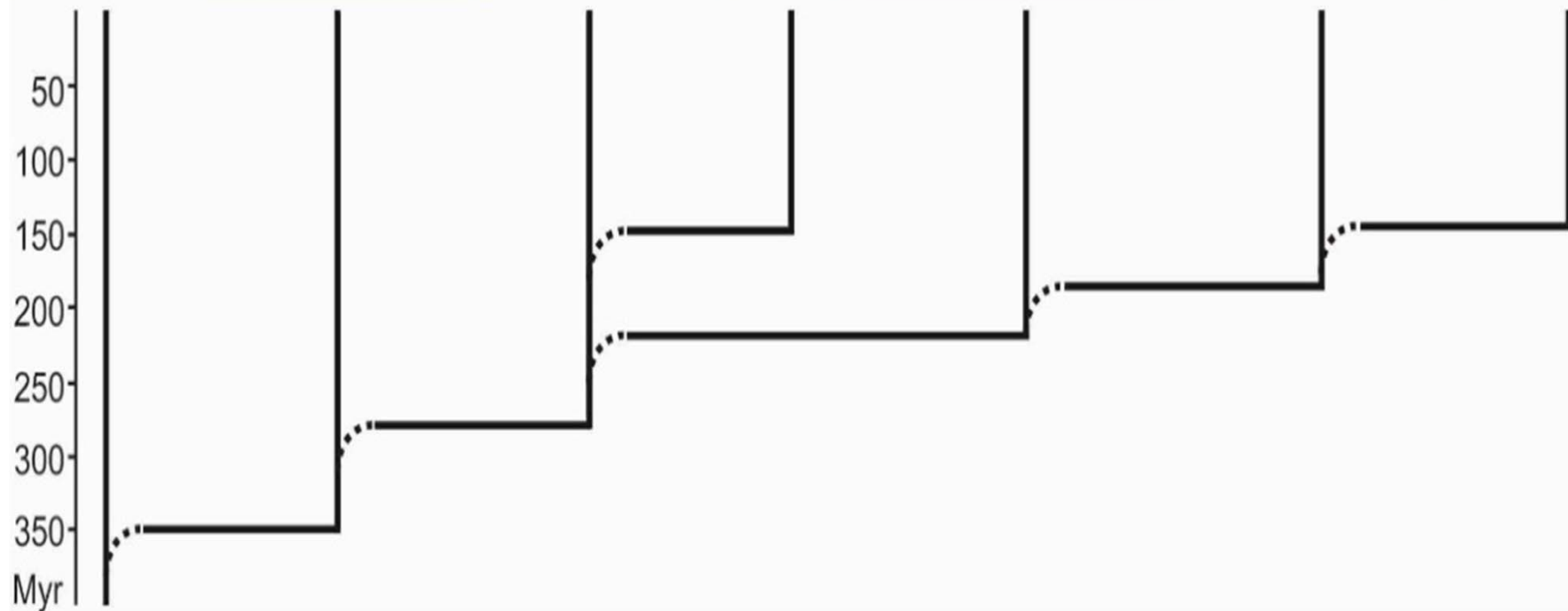
Reptilians

Birds

Monotremes

Marsupials

Placental
mammals



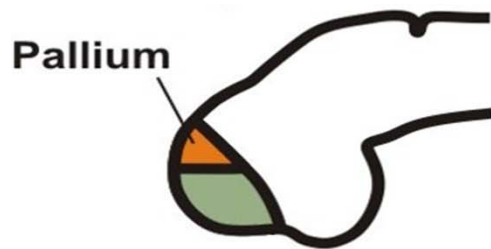
Математическая биология –

алгоритмическая работа с генами

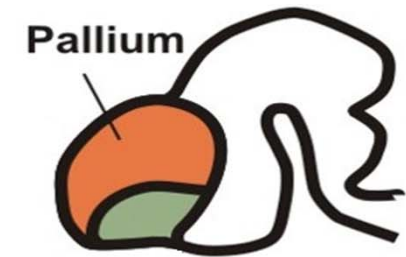
– длинными словами из четырёх букв А, С, G, Т.

Пример:

примитивный мозг
холоднокровных
(лягушки)



развитый мозг
теплокровных
(человека)



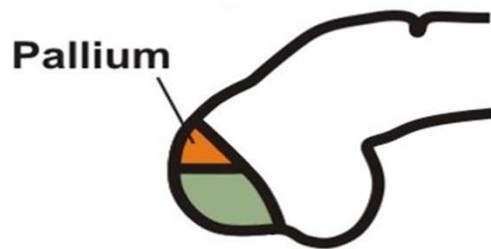
Математическая биология –

алгоритмическая работа с генами

– длинными словами из четырёх букв А, С, G, Т.

Пример:

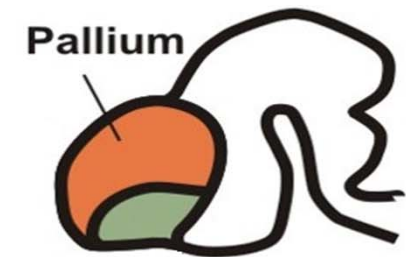
примитивный мозг
холоднокровных
(лягушки)



потеря алгоритмически
найденного нами **гена c-answer**
приводит в эволюции к
потере/приобретению
фенотипического признака:

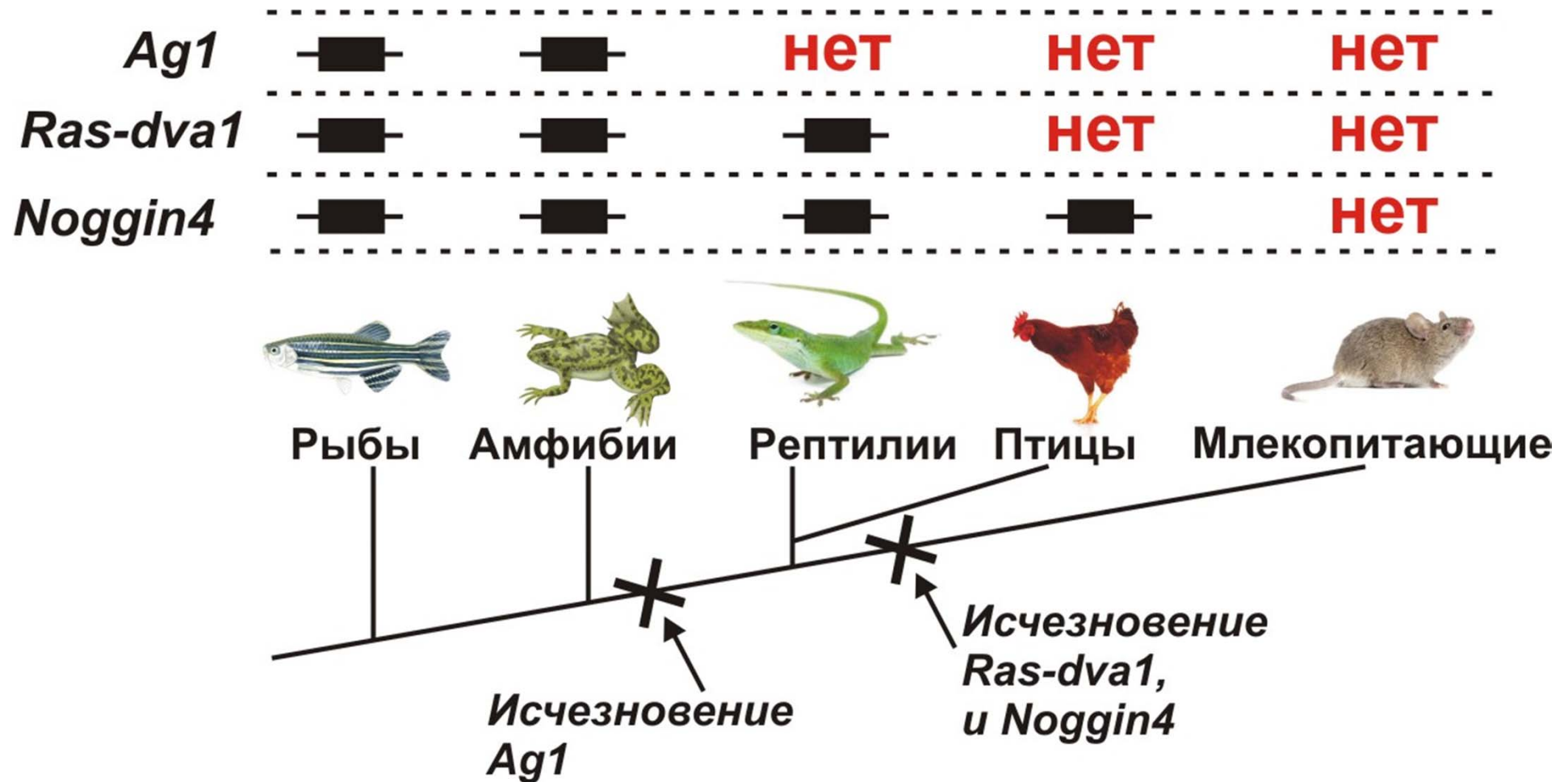


развитый мозг
теплокровных
(человека)



Точнее: общий предок современных лягушек и человека имел этот ген, но у одного потомка он остался и от него пошли лягушки, а у другого потомка он потерялся и от него пошли люди.

Время идёт, гены теряются/приобретаются и соответственно признаки организма/вида теряются:



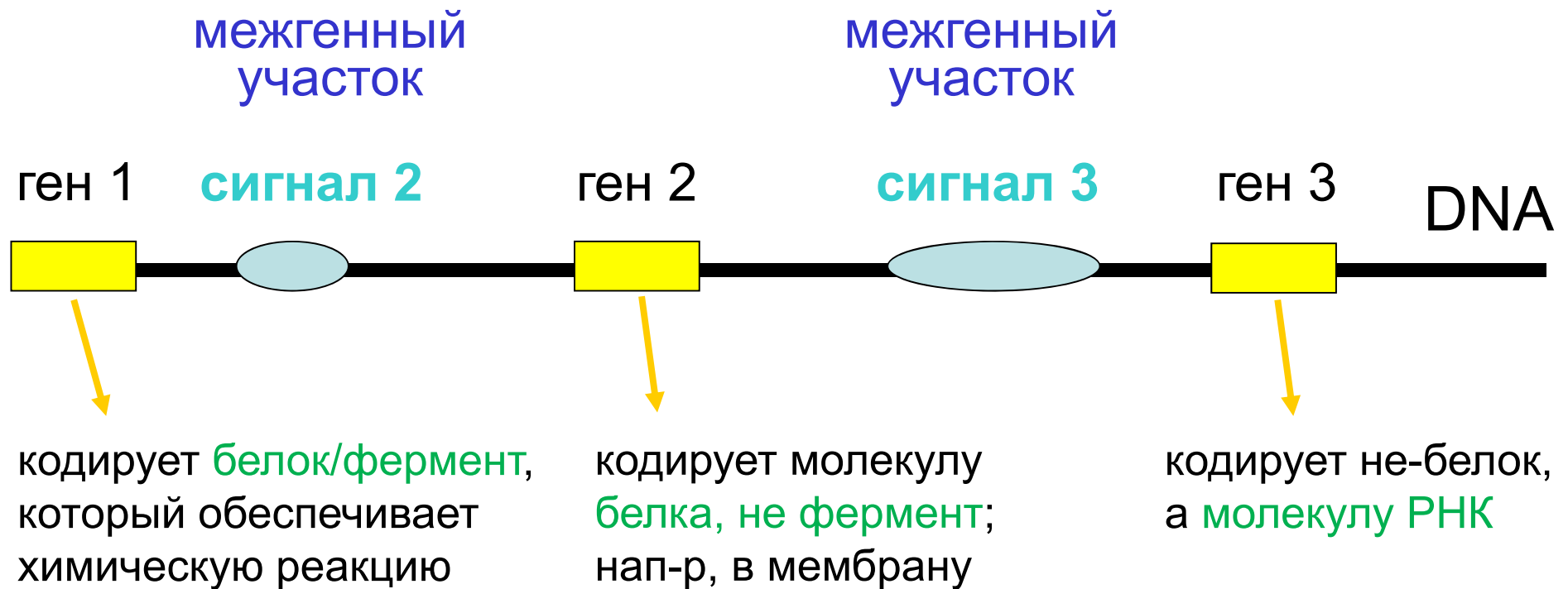
Задачи по Математической биологии:

подключитесь к базе данных Ensembl

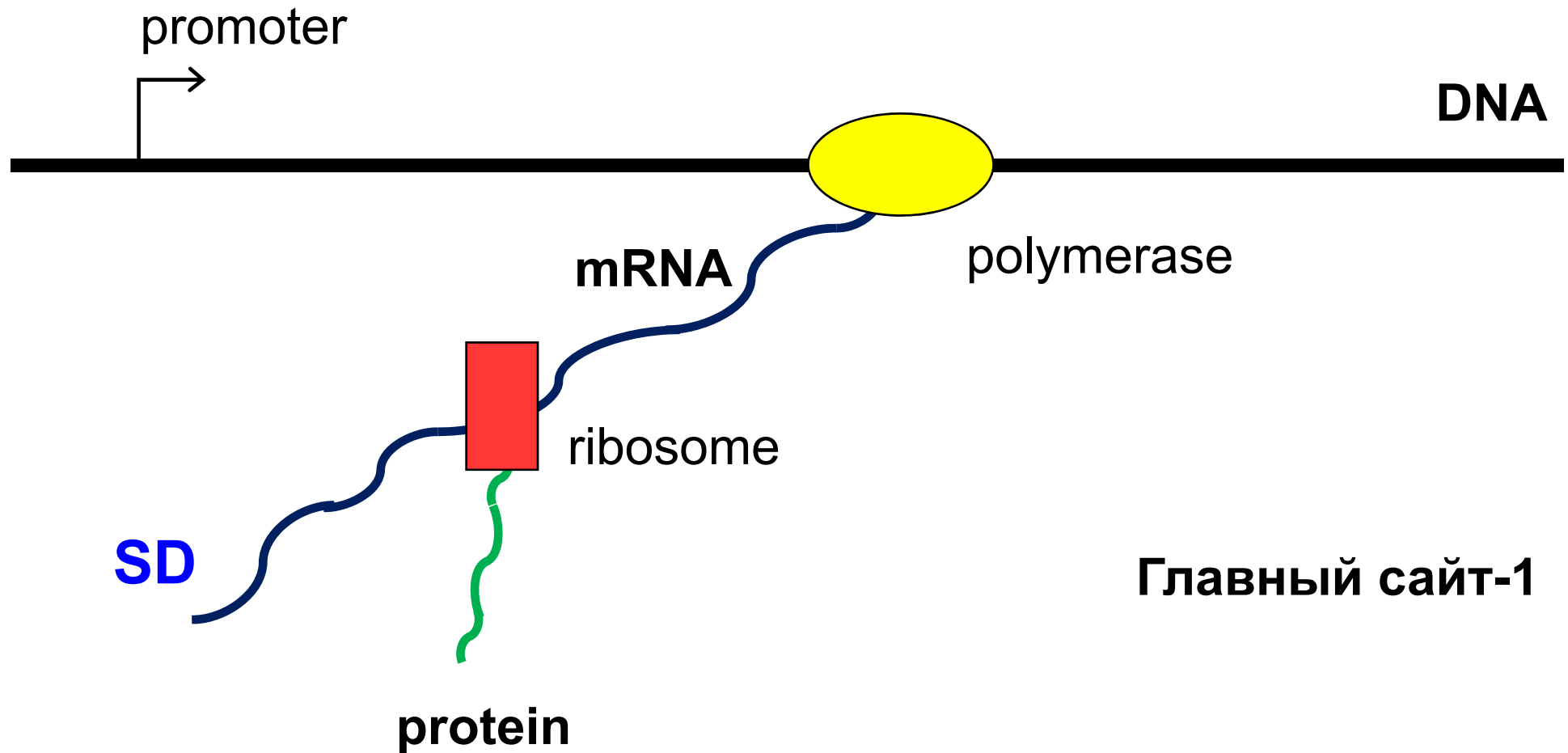
<https://www.google.com/search?q=Ensembl&oq=Ensembl&aqs=chrome..69i57j0i271.14623j0j7&sourceid=chrome&ie=UTF-8>

Найдите там геном человека (*Homo sapiens*) и в нём геном митохондрии и в нём самый длинный и самый короткий гены.

ДНК – последовательность=слово в 4-буквенном алфавите {**А,С,Т,Г**} с длиной до сотен миллионов букв. Каждая **буква** называется **нуклеотид**. Например, 30 позиций:
TTGACATGGCTATATAAGTTCATGTTATACT

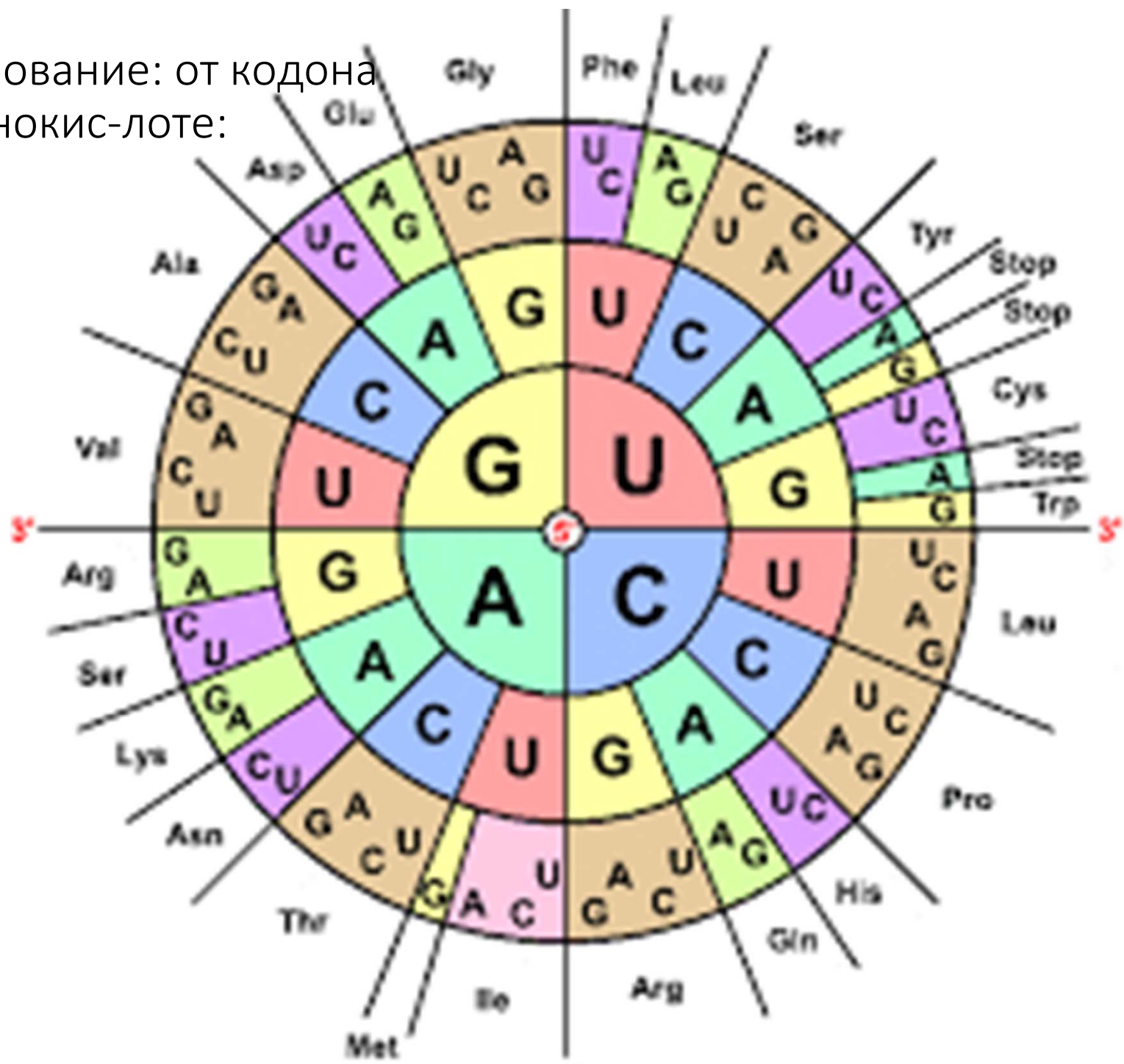


«Суть жизни»: ген **считывается или не считывается** по **сигналу** из лидерной области в зависимости от обстоя-тв!

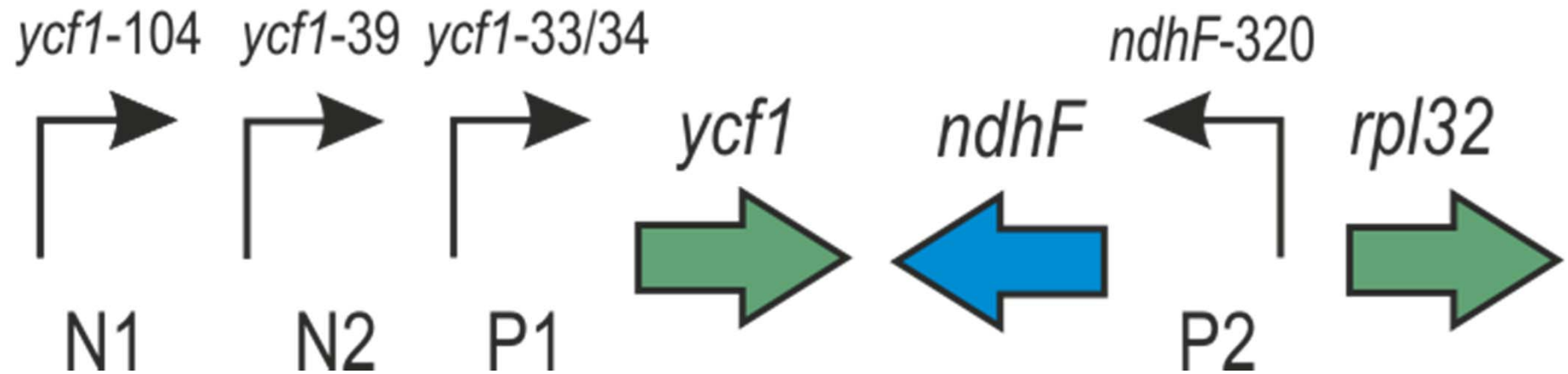


Итак, имеется **два основных процесса** в «текущей жизни клетки»: **ТРАНСКРИПЦИЯ** и затем (иногда) **ТРАНСЛЯЦИЯ**. Эти процессы протекают согласованно и независимо: рибосома не может начать работу до полимеразы, но начав работу, «не интересуется» судьбой полимеразы. Но у них есть связь!

Кодирование: от кодона
к аминокис-лоте:

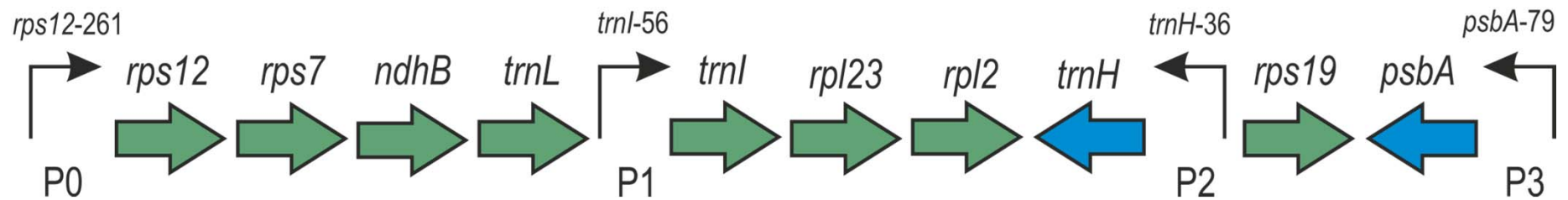


Пример: 3 гена и 4 промотора

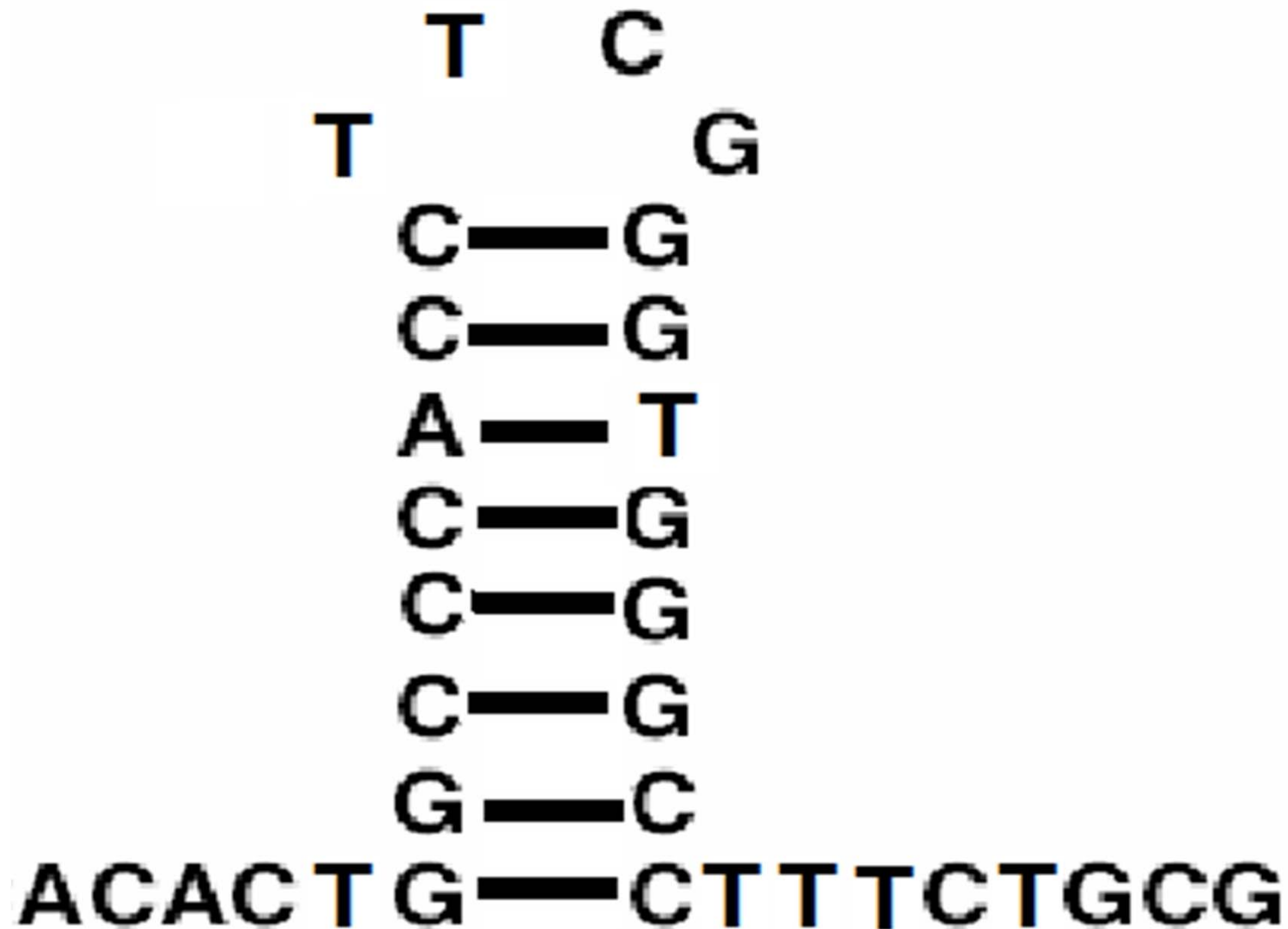


Взаиморасположение генов и промоторов может быть очень разнообразным.

Еще пример: 10 генов и 4 промотора



Показана **шпилька с черенком, плечами и петлёй**,
в которой, как положено, склеиваются **G** и **C**, **A** и **T**:

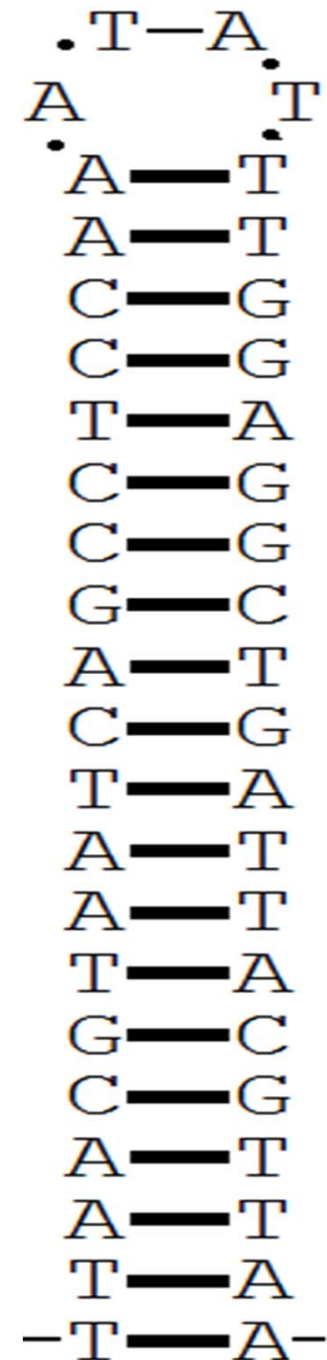


Снизу показан участок **РНК (ДНК)**: жёлтый участок называется **левым плечом**, голубой участок – **правым плечом**; то, что между плечами, – **петлей**. Справа показано, как такой участок сворачивается в **шпильку (=спираль)**. У шпильки различают **черенок** и **петлю**.

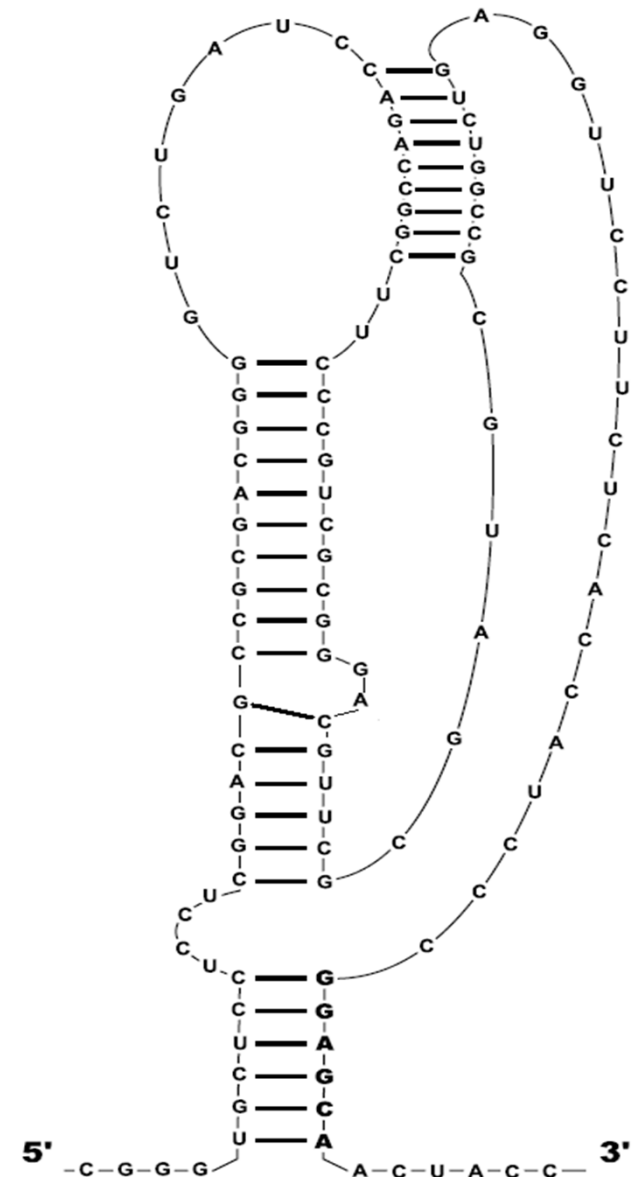
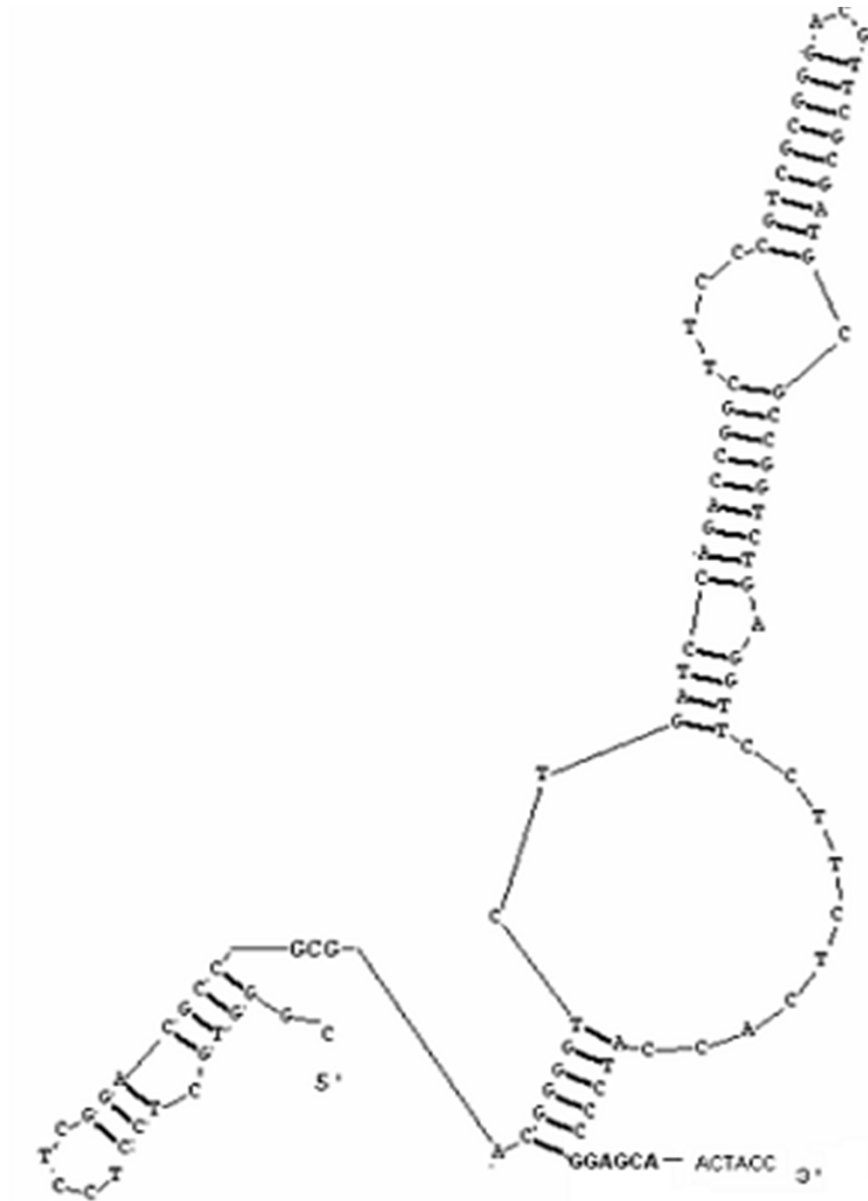
Сотни, тысячи шпилек образуют **вторичную структуру данного участка РНК**.

Также важно последующее **спаривание** между шпильками, или между плечом на РНК и плечом на ДНК.

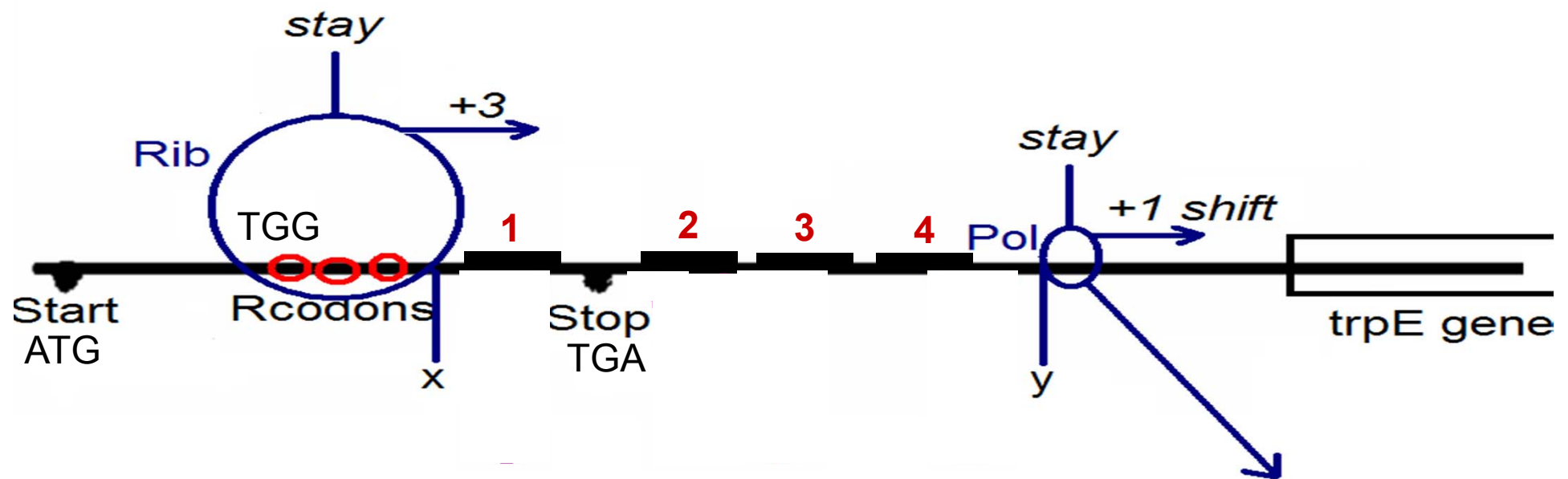
ТТААСГТААТСАГССТССАААТАТТТГГАГГСТГАТТАСГТТАА



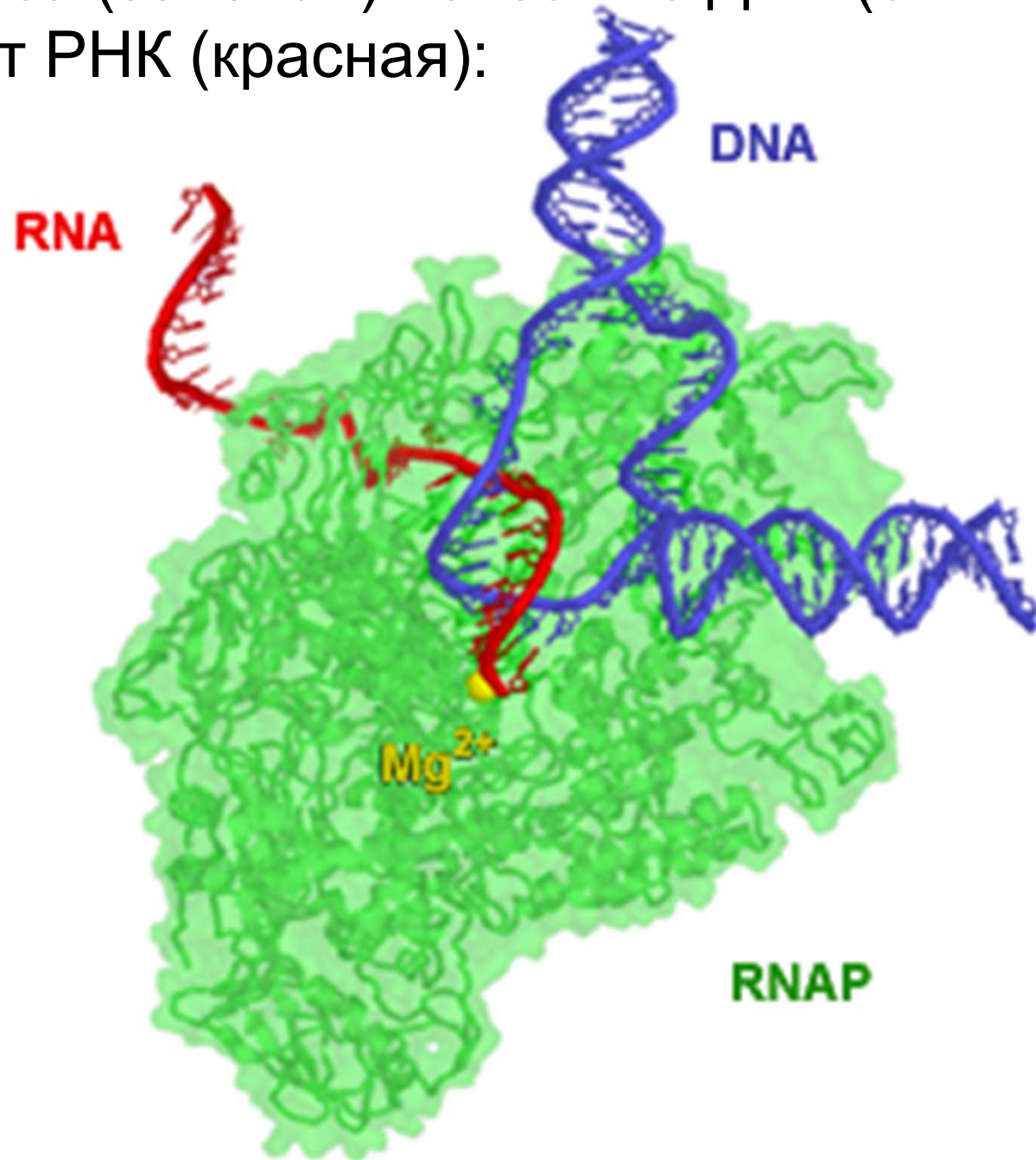
Примеры очень простых вторичных структур (шпильки «с выпячиваниями» и «псевдоузел»):



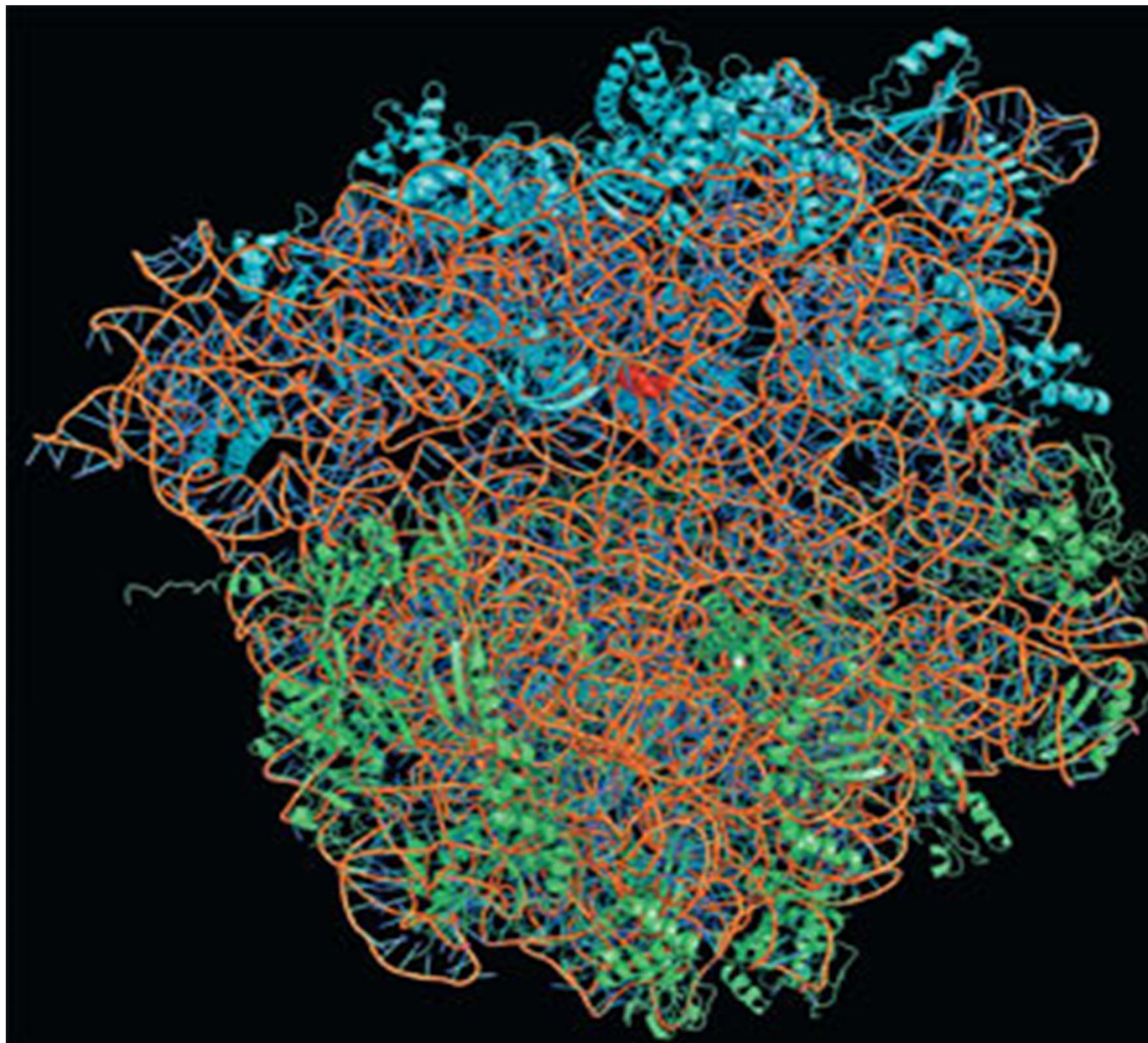
Сначала по ДНК ползёт **Полимераза** по той части, которая находится перед геном. По образующейся РНК ползёт **рибосома**. На участке **РНК между Rib и Pol** (т.е. от **x** до **y**) находится **море вторичных структур**, которое волнуется! Показаны возможные плечи 1, 2, 3, 4 шпилек, составляющих это море .

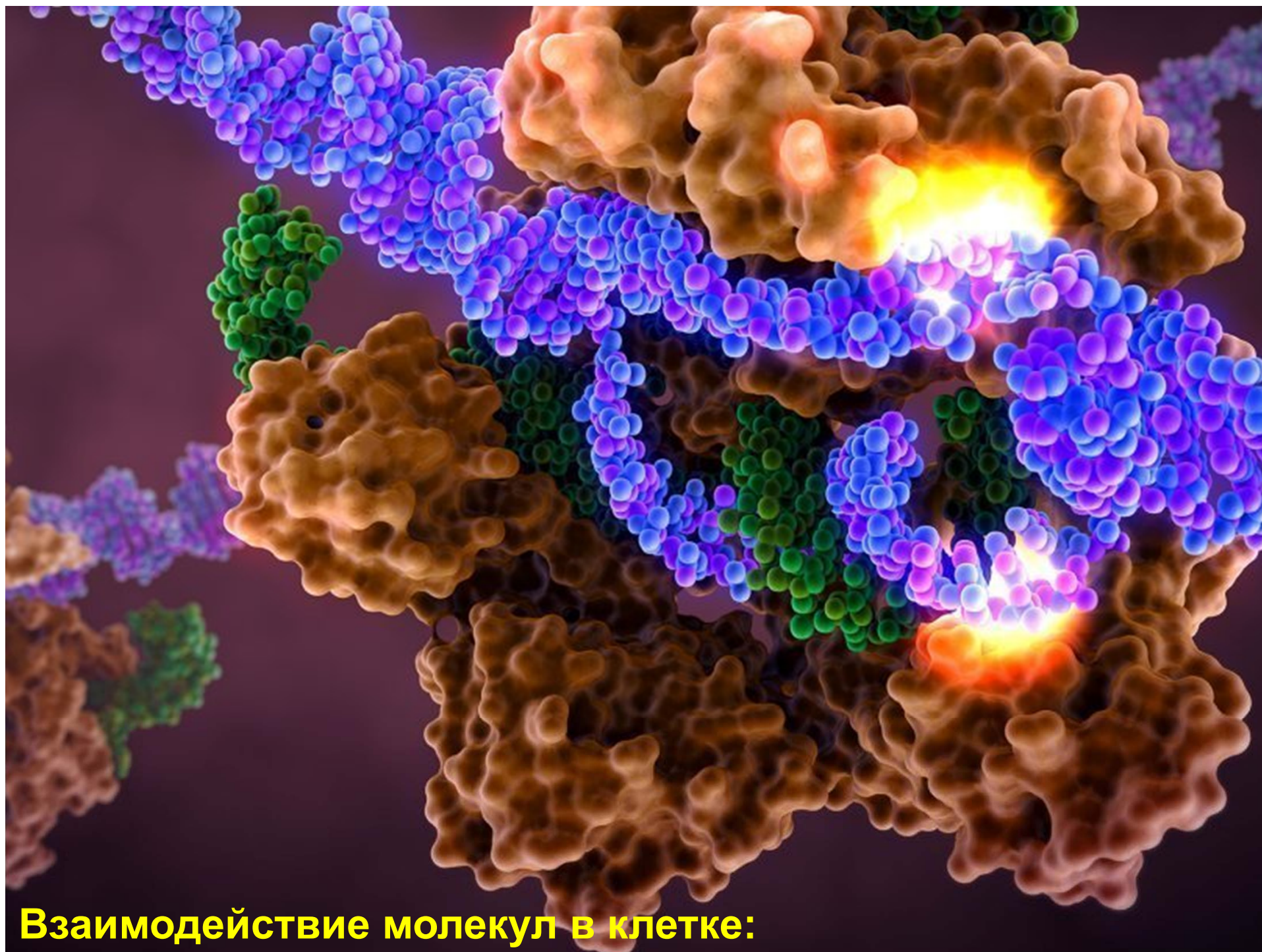


РНК-полимераза (зеленая) ползёт по ДНК (синяя) и из неё выходит РНК (красная):



Рибосома
ползёт по
мРНК и из
неё
выходит
белок:





Взаимодействие молекул в клетке: