

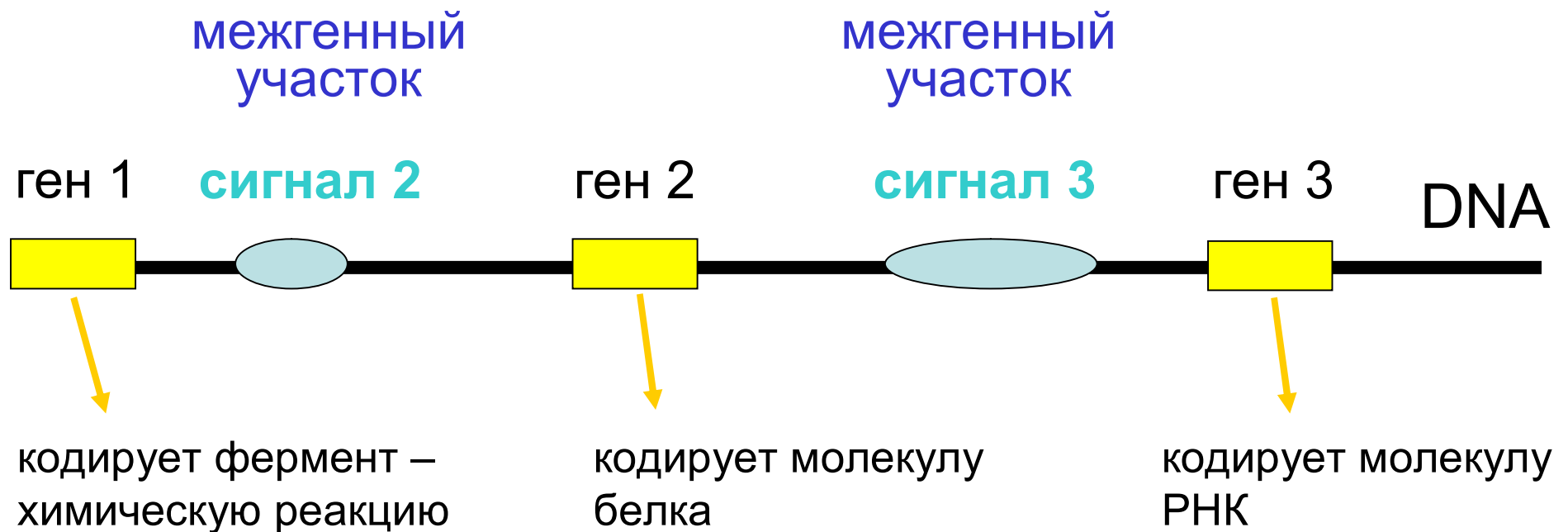
**В. Любецкий, А. Селиверстов**

лаборатория 6

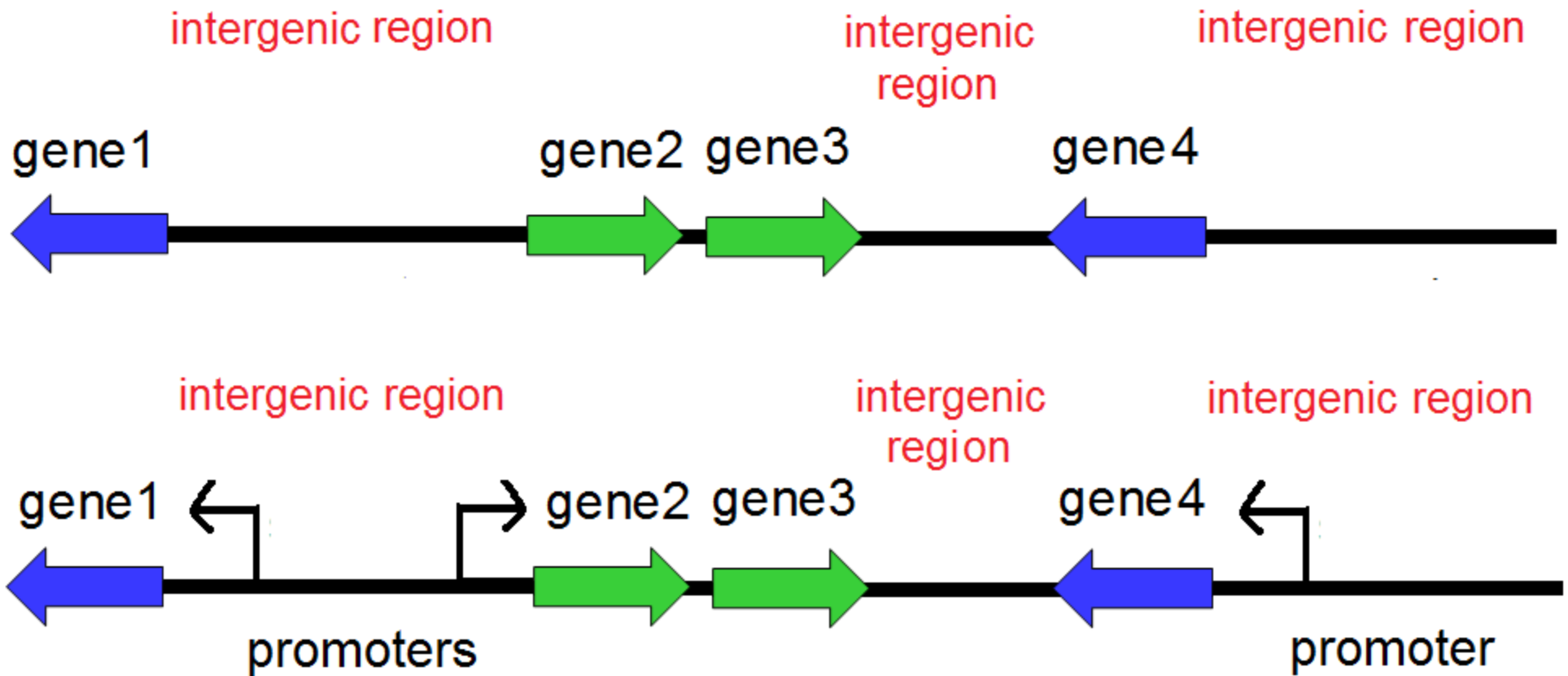
<http://lab6.iitp.ru/>

**Тканеспецифичная транскрипция в  
митохондриях хордовых  
в процессе онтогенеза**

**ДНК** – последовательность в 4-буквенном алфавите {**A, C, T, G**} с характерной длиной 3 миллиона – 6 миллиардов (бывает и всего 17тыс) позиций. Каждая **буква** называется **нуклеотид**:  
**TTGACATGGCTATATAAGTTCATGTTATACT = 30 позиций**



**«Суть жизни»:** ген **считывается или не считывается**  
по **сигналу** из лидерной области



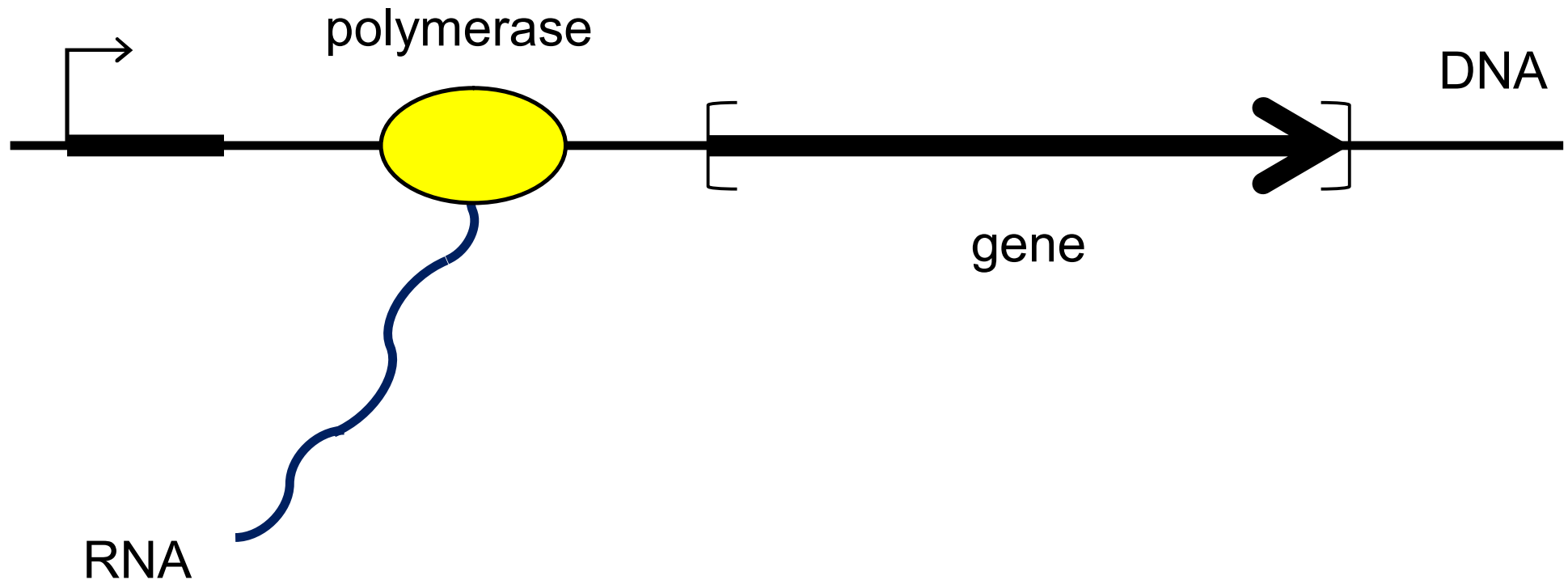
Пример сигнала – **промотор**, участок с определенными буквенным составом и с направлением. Например:

human **CAAACCCCAAAGACA** →

bacteria **TTGACA -17..18- TATAAT -4..7- R** →

frog **ACRTTATA** (R=A или G) →

## Что такое «чтение гена»?

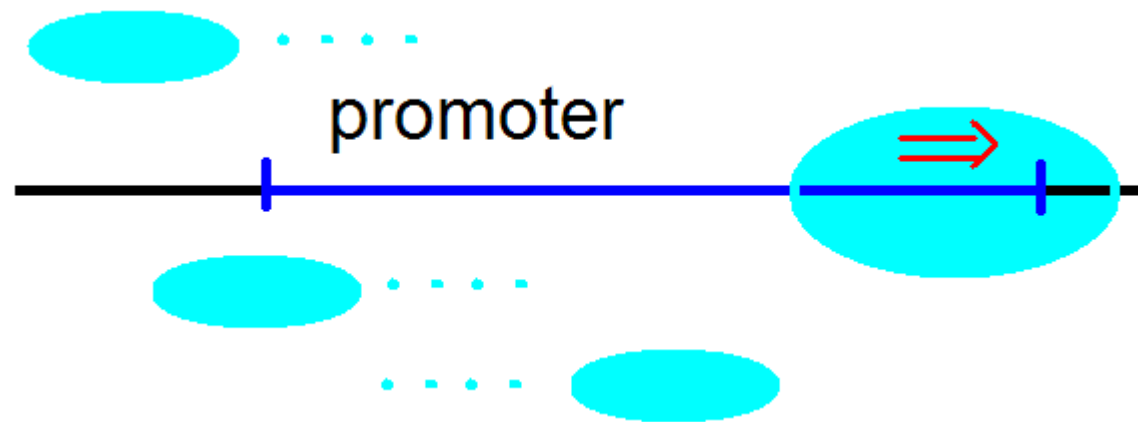


Основной процесс в «текущей жизни клетки» – **транскрипция**, он состоит: **1)** в **связывании** с промотором молекулярной машины, называемой **полимеразой**, **2)** ее **движении** в направлении промотора, **3)** если полимераза **проходит** любой **со-направленный** с ее движением ген, она его «читает» («копирует в РНК»)

## Что такое: «связывание с промотором»?

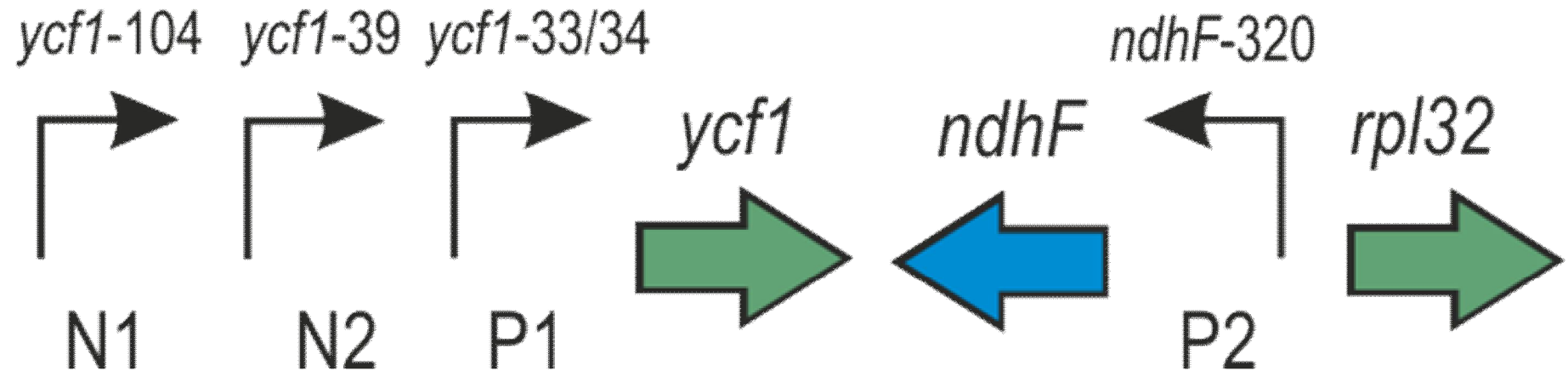
Промотор **свободен**, если на нем не находится ни одной полимеразы (и ни одной другой молекулы).

**Связывание** полимеразы с промотором происходит, если в момент попытки связывания он **свободен**:



Каждый промотор характеризуется своей **ИНТЕНСИВНОСТЬЮ ПОПЫТОК СВЯЗЫВАНИЯ**

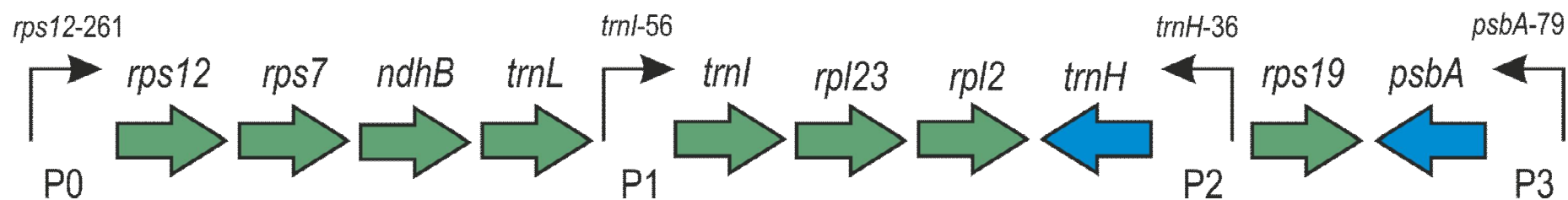
## Пример: 3 гена и 4 промотора



**Взаиморасположение** генов и промоторов (и других сигналов) может быть очень разнообразной.

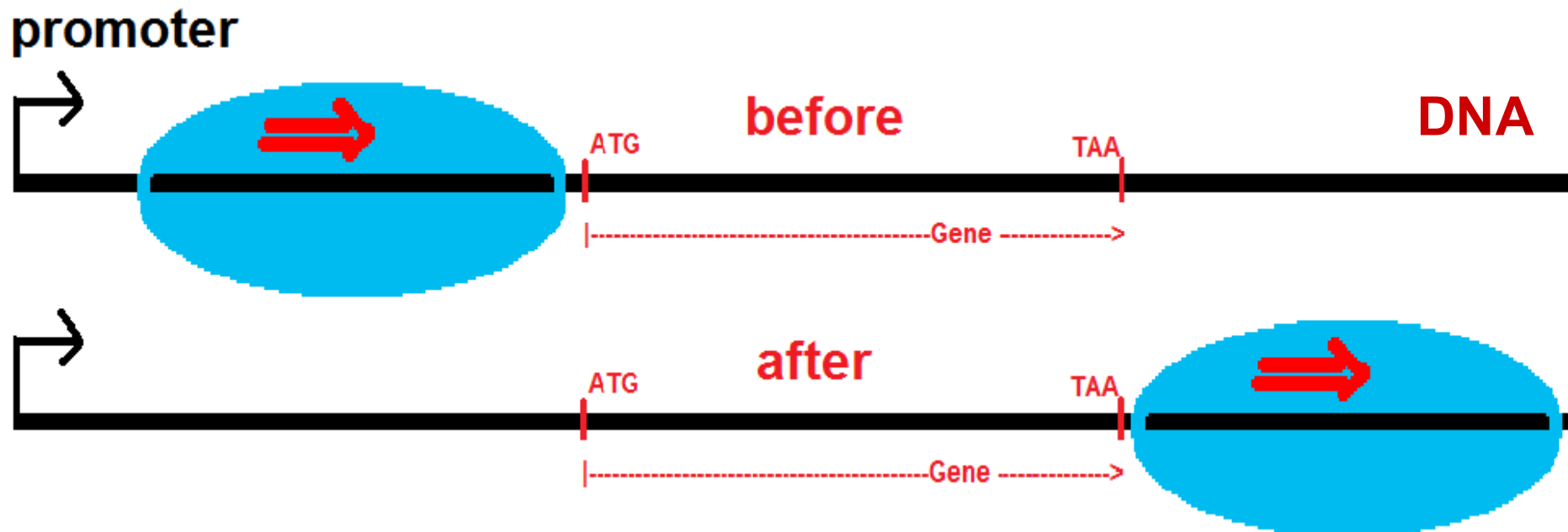
В этом одна из трудностей математического изучения ДНК (геометрия бывает очень разная)

## Еще пример: 10 генов и 4 промотора



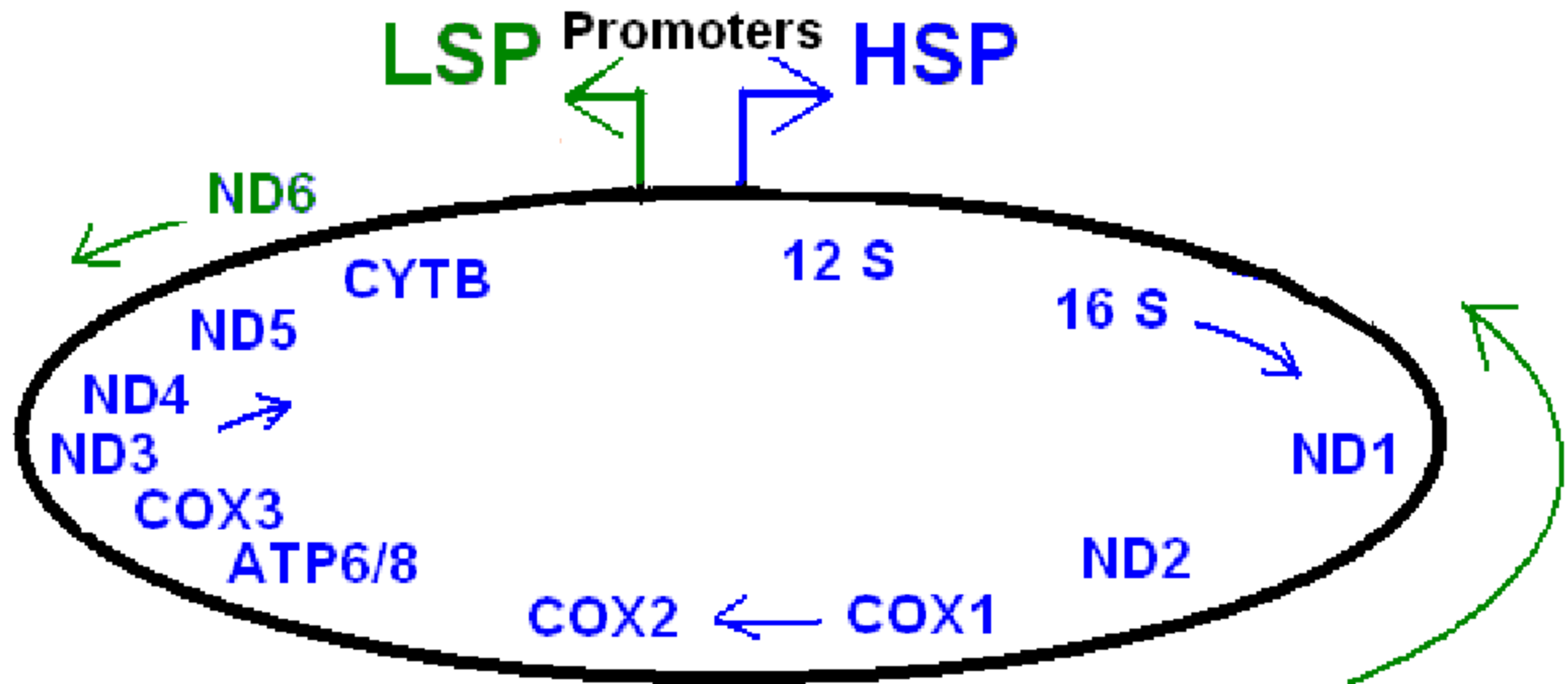
С каждым геном связан «счётчик числа прочтений»:

ген «**прочитан однократно**», если полимераза **прошла** его от начала до конца один раз. При каждом гене находится свой счётчик **числа его чтений (=прохождений)** в ед. времени. **Частота** чтения гена называется **уровнем его транскрипции**





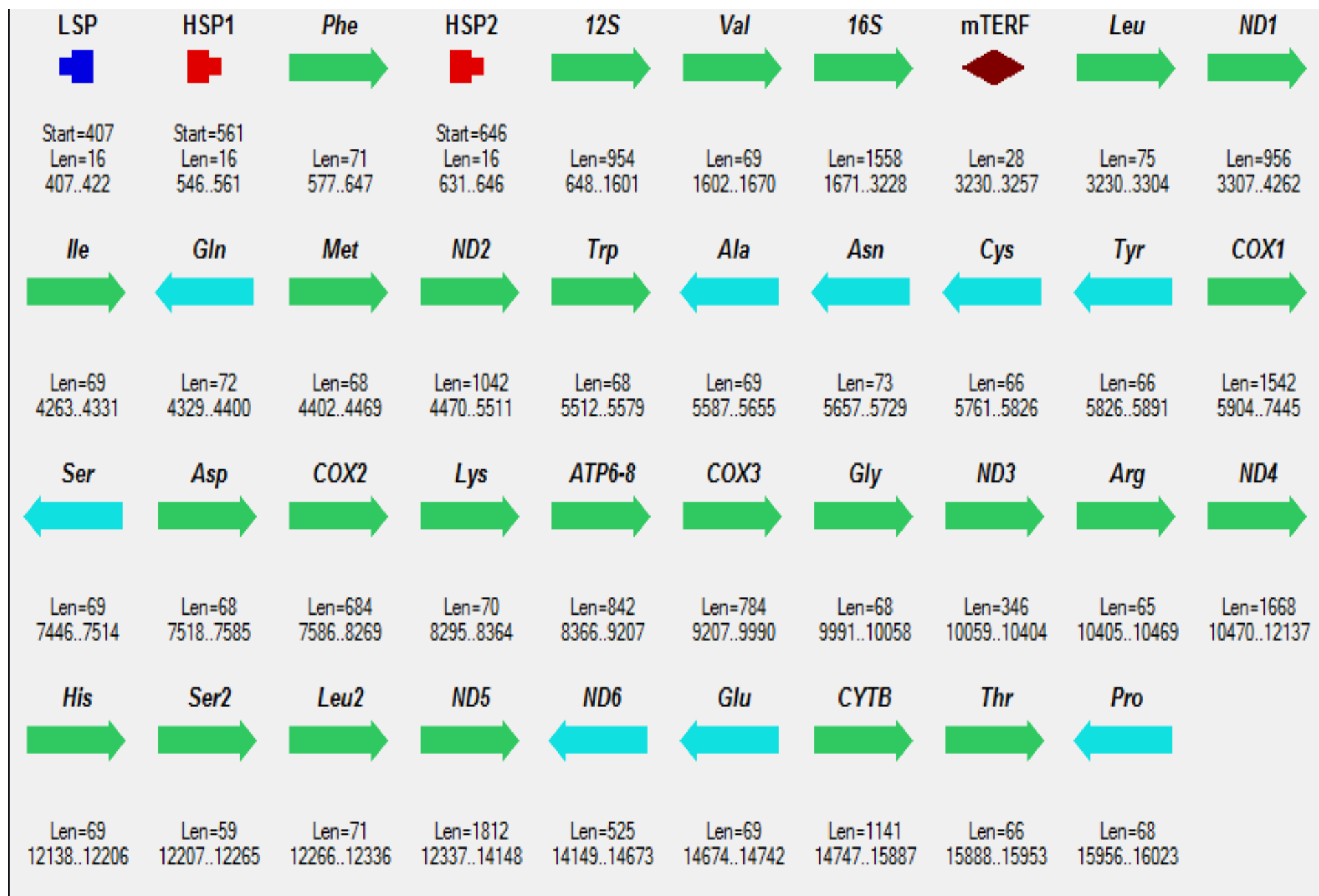
**Митохондриальная ДНК человека.** В ней порядок генов одинаков у человека, крысы и шпорцевой лягушки, но несколько отличается у курицы. У разных видов промоторы LSP и HSP заметно отличаются



У человека, крысы, лягушки показаны число промоторов и координаты их «концов» (стартов транскрипции):

<i>Homo sapiens</i>	<b>HSP1</b>	<b>561</b>
	<b>HSP2</b>	<b>646</b>
	<b>LSP</b>	<b>407</b>
<i>Rattus norvegicus</i>	<b>HSP1</b>	<b>16298</b>
	<b>HSP2</b>	<b>66</b>
	<b>LSP</b>	<b>16193</b>
<i>Xenopus laevis</i>	<b>HSP1</b>	<b>2102</b>
	<b>HSP2</b>	<b>2049</b>
	<b>LSP1</b>	<b>2103</b>
	<b>LSP2A</b>	<b>2042</b>
	<b>LSP2B</b>	<b>2033</b>

# Митохондриальный геном человека:

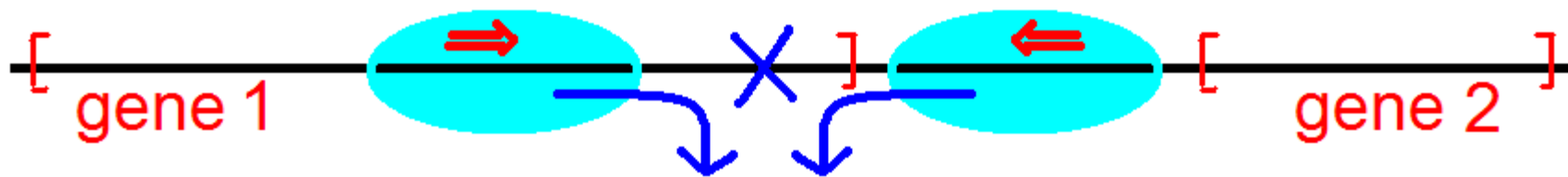


# Митохондриальный геном лягушки:



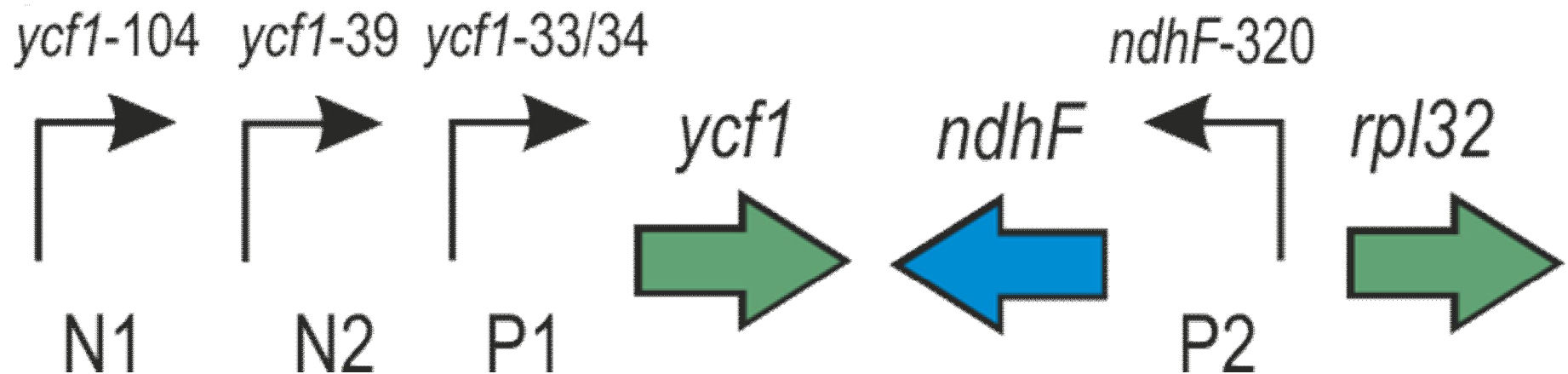
Если одна полимераза догоняет другую, то вторая движется вслед за первой, не влияя на нее и только снижая скорость своего движения.

При лобовом столкновении, обе полимеразы сваливаются:



Как автомобили на дороге

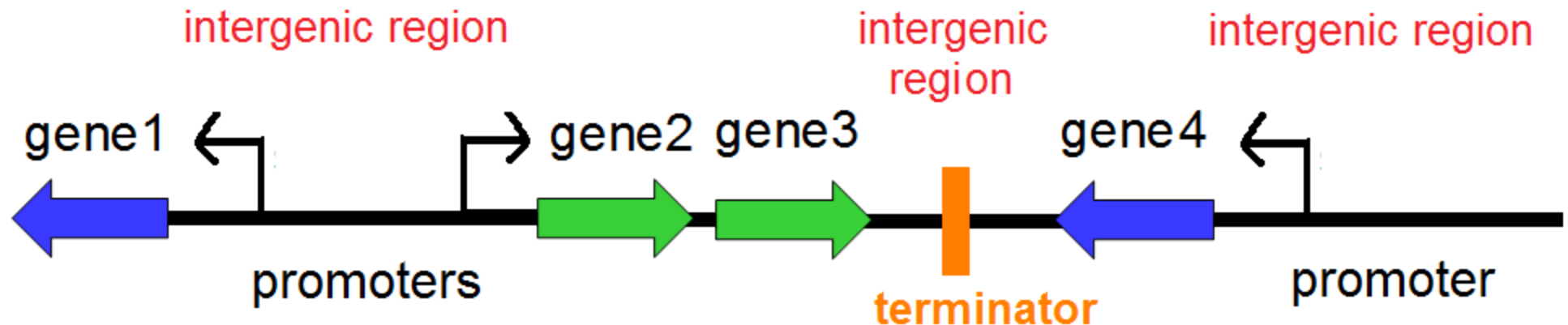
Итак, имеем **ситуацию для моделирования**: много полимераз одновременно связываются с ДНК и движутся по ней по **определенным правилам**. При этом каждый **ген считывается свое число раз**, полимеразами идущими по нему с сонаправленного промотора:



Как автомобили на дороге

Еще одно действующее лицо: **терминатор**,

он **скидывает** определенную долю полимераз  $p$ , идущих с одной стороны, и определенную долю полимераз  $q$ , идущих с другой стороны (**два автоинспектора** на разных сторонах дороги)



Различие в числах  $p$  и  $q$  называется степенью поляризации терминатора. Если одно из них равно 1, то он 100% поляризованный в соответствующем направлении

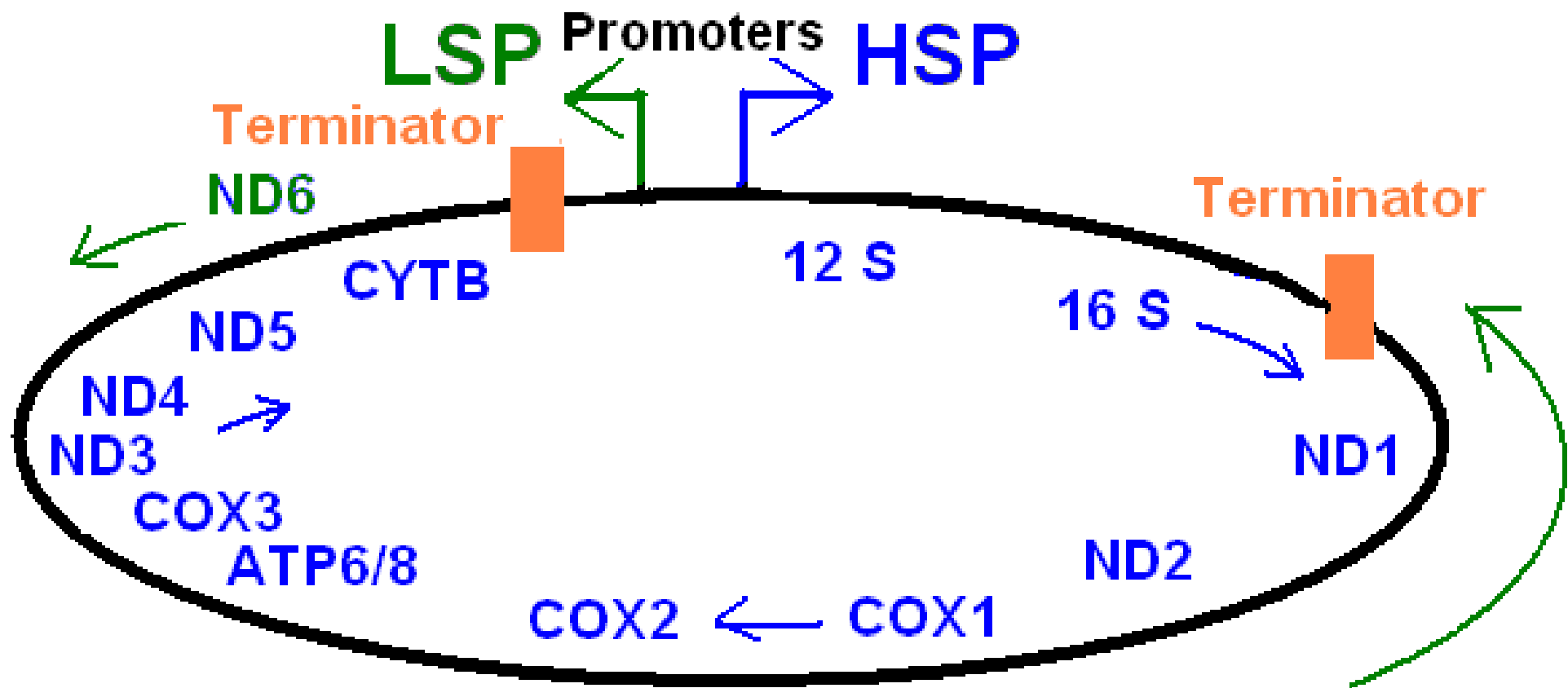
В ряде случаев терминаторы и их положения были  
впервые обнаружены нами  
именно с помощью моделирования!

Как это возможно?

Если терминатор отсутствует в данном месте и с  
данными значениями  $p$  и  $q$ , то модель **не**  
**предсказывает опытных значений** про другие  
объекты расположенные в других местах на ДНК  
– те и там, которые и где мерить легче



У человека, крысы и лягушки известно два терминатора:



**Один** из этих двух **терминаторов** («**терминатор1**») обусловлен связыванием белка mTERF с сайтом длиной 28 на ДНК непосредственно после гена 16S.

Например, транзиция **A→G** в позиции **3243** снижает связь mTERF с ДНК, что вызывает тяжелые генетических заболевания человека. Вообще различные мутации в ДНК митохондрий приводят к **MELAS-болезням (энцефаломиопатии)**.

**Проблема:** какие **изменения ферментного состава** произойдут **в случае замен букв в разных позициях?**

Таким образом моделирование позволяет изучать митохондриальные генетические заболевания, в частности, MELAS-болезни

Терминатор2 имеет другой механизм работы, который был нами выяснен (формирование гуанилового тетрамера на РНК).

Точные позиции и буквенный состав терминаторов были предсказаны нами.

Например, для терминатора2 имеем:

<i>Homo sapiens</i>	GGGGGAGGGGGGG
<i>Rattus norvegicus</i>	GGGGGTGGGGGGGG
<i>Xenopus laevis</i>	GGGGGGTAGGGGGG
<i>Artemia franciscana</i>	GGGGGGGGGGGGGG
	GGGGGGGGGGGGGG

## Повторение

Каждая полимераза сначала связывается со своим промотором, а затем движется в соответствующем направлении.

Полимеразы, движутся друг за другом, не влияют на впереди идущую, а только снижают скорость, если впереди идущая движется медленнее. При лобовом столкновении полимеразы сваливаются.

Терминаторы также сваливают полимеразы.

Эти **процессы** нами **математически описаны** и в **совокупности моделируются**

Модель реализована программой для многопроцессорной вычислительной системы; программа доступна по адресу <http://lab6.iitp.ru/ru/rivals/> и снабжена полным и подробным описанием, включая примеры моделирования.

Там же имеется вариант программы для персонального компьютера, позволяющий следить за ходом моделирования в наглядном графическом режиме.

Вычисления выполнялись на суперкомпьютере МВС-100К Межведомственного Суперкомпьютерного Центра РАН с привлечением 2048 процессоров

Пример численного результата:

вероятности протекания **терминатора**<sup>1</sup> «от» и «к» промотору HSP соответственно равны **0.217** (20%) и **0.048** (5%) – доли при условии, что белок mTERF находится на сайте.

Эта – **единственная пара чисел**, которая согласует **все опытные данные** по лягушкам. Поскольку участки этих промоторов и сам связывающий белок mTERF почти совпадают у всех хордовых, то **эти числа служат ответом и для других хордовых.**

**Хотя опыты с другими хордовыми не проводились и их трудно провести! В этот смысл моделир-ия**

Таким образом, **терминатор1** достаточно поляризован, но не на 100%, как предполагалось в литературе.

Имеется еще второй ответ, который указывает на неполяризованность **терминатора1**.

Таким образом, модель указывает на критически важный опыт!

Мы в соответствии с имеющимися данными  
верим в первый ответ

Наше моделирование транскрипции **без терминаторов** показало **поведение, невозможное с биологической точки зрения**. А именно, после некоторого начального периода ДНК начинает транскрибироваться только в одном направлении, т.е. гены в другом направлении перестают читаться. А модель «правильная».

Таким образом, была **«доказана» существенная биологическая роль терминаторов**, что не было очевидно только из опытов



## Найденное решение для эмбриона лягушки (*Xenopus laevis* mitochondrion)

time	QH	QL	Frog3		Frog1		Frog2	
			mTERF	LSP1	mTERF	LSP1	mTERF	LSP1
0	0.2165	0.0480	0.0173	0.0021	0.0067	0.0023	0.0031	0.0021
5	0.2165	0.0480	0.0190	0.0055	0.0118	0.0061		
6	0.2165	0.0480					0.0041	0.0023
9	0.2165	0.0480					0.0055	0.0041
10	0.2165	0.0480			0.0157	0.0061		
14	0.2165	0.0480	0.0278	0.0097	0.0209	0.0107		
16	0.2165	0.0480			0.0253	0.0157		
18	0.2165	0.0480			0.0306	0.0190		
20	0.2165	0.0480	0.0370	0.0118	0.0336	0.0209	0.0067	0.0118
23	0.2165	0.0480			0.0370	0.0542		
28	0.2165	0.0480	0.0407	0.1056				
30	0.2165	0.0480					0.0130	0.0253
48	0.2165	0.0480	0.0278	0.0655	0.0407	0.0960	0.0336	0.0872
96	0.2165	0.0480			0.0306	0.0872		
7 days	0.2165	0.0480					0.0089	0.0190

**Методика исследования показана на  
следующем слайде**

**(изучается зависимость от стадии развития  
организма и от ткани);**

**заметим** на одном организме выясняются  
одни параметры, а на другом – другие,  
и в результате они переносятся с  
одного организма на другой:

На первом шаге работы у **лягушки** определили **вероятности терминаторов 1 и 2**. Для терминатора<sup>1</sup> известны относительные интенсивности связывания с промоторами [Bogenhagen, Romanelli, 1988] и также многочисленные данные об уровнях транскрипции генов [Ammini, Hauswirth, 1999]. Эти **данные известны в онтогенезе**: на 0, 5, 10, 14, 16, 18, 20, 23, 48, 96 часов развития.

На втором шаге в предположении тех же вероятностях терминаторов, **у человека** определили **интенсивности попыток связывания** с промоторами по данным из [R.Gelfand, Attardi, 1981] и [Piechota *et al.*, 2006].

**У человека** определили **влияние мутаций** и сравнили его с известными данными [Chomyn *et al.*, 1992].

На третьем шаге **у крысы** определили **интенсивности попыток связывания в разных тканях: мозгу, сердце, печени, мышцах и почках** по данным из [Enríquez *et al.* 1999] и [Fernández-Vizarra *et al.*, 2011]

## Выводы о MELAS-мутациях

Транзиция А→G внутри сайта связывания mTERF в клетках печени вызывает многие генетические заболевания человека. Но при этом уровни транскрипции мРНК почти не меняются [Chomyn *et al.*, 1992]. **Откуда же происходит болезнь?**

1) Наше (на основании моделирования) объяснение этого эффекта состоит в **существенном уменьшении уровня транскрипции фенилаланиновой тРНК**, которую трудно померить и это не было сделано.

2) Нами определено влияние и **других мутаций в других позициях и тканях**

Терминатор1 расположен в области, кодирующей лейциновую тРНК (т.е. внутри гена лейциновой тРНК) и потому думали, что причина болезни в мутации лейцинового гена!

Моделирование показало, что мутация влияет именно на терминатор1 и сказывается на уровне транскрипции фенилаланинового гена

В работах [Enríquez *et al.* 1999] и [Shock *et al.* 2011] обсуждается роль **метилирования митохондриальной ДНК** в регуляции транскрипции у млекопитающих. Моделирование позволяет уточнить влияние эпигенетических нарушений на экспрессию генов митохондрий.

С тех же позиций изучены и другие виды, для которых измерена транскрипция генов митохондрий: радужной форели *Oncorhynchus mykiss* [Leary *et al.*, 1998], других рыб [O'Brien, 2011], креветки *Artemia franciscana*, морских ежей *Paracentrotus lividus* и *Strongylocentrotus purpuratus*

Спасибо за Ваше внимание

Методика вычислений:

у лягушки 4 переменных: промотор LSP1 (через него выражаются остальные четыре),  $p$  и  $q$  терминатора1 (терминатор 2 не рассматривается, он ослабляет полимеразы с LSP1, пропуская 35%, в обратную сторону все), связывание mTERF;

у человека 3 переменных:  $p$  и  $q$  уже заданы от лягушки, теперь это – промоторы LSP1 и HSP1 (промотор HSP2 - фиксированная доля от HSP1, сейчас в диком типе суммарная  $1/10$ , а при мутации нужно смотреть отдельно HSP1 и HSP2(?)), связывание mTERF.

У лягушек и человека имеются общие условия. Кроме них дано следующее.

У трех лягушек:  $p$  и  $q$  терминатора1 не меняются во времени,

у 3й: относительные (к яйцу) уровни транскрипции генов 16S и ND6;

у 1й: относительные количества мРНК, пересчитываемые через времена полураспада в уровни транскрипции, генов ND1, COX2, ATP6, ND4, ND6, CYTB в одни моменты времени

и у 2й: тоже те же гены в другие моменты времени,

связывания LSP1 в моменты времени 0, 20 и 48 час одинаковы («одинаково развиваются»).

Итак, хорошо, что из лягушки находится фактически одна пара  $p$  и  $q$ , которую принимаем для человека.

У человека здорового (в диком типе):

у здорового и больного считаем сумму HSP1+HSP2 (?);

1) относительные (ND1) количества РНК генов ND1, ND2, COX1, COX2, ATP6/8, ND3, ND5, CYTB – пишем функционал, который отражает отклонение;

2) превышение суммарного числа транскриптов рРНК, т.е. числа транскриптов 16S, над мРНК цитохрома  $b$  как будто в 10-15 раз на клетку (но можно перейти к уровням транскрипции, так как коэффициент сокращается);

= как обеспечить двухкритериальную оптимизацию здесь?, методика: сначала откидываем кандидатов для которых не выполняется  $16S > x=?$ , затем на отобранных кандидатах находим минимум значений функционала, для которого кандидатов, уклоняющихся более квантили, удаляем; повышаем  $x$  и т.д.

Отсюда определяем параметры здорового человека, из которых используем: связывания промотора1 и терминатора1 (сама mTERF у больного не меняется?). Затем их уменьшаем на разные доли  $y$  и  $z$  (перебор), чтобы удовлетворить условиям у больного человека. Эти последние состоят в ровно одном!: «RNA-Leu отклонение до 20% от нормы, для tRNA-Lys отклонение до 40-50% от нормы [Chomyn *et al.*, 1992]».

3) У нас при некоторых  $y$  и  $z$  уровни Phe, 12S, Val, 16S падают в два раза, что хорошо, тк объясняет суть болезни. Но хорошо бы еще получить шевеление на RNA-Leu и tRNA-Lys, хотя Саша говорит, что и ноль на них годится.

Итак, методика поиска решения состоит в следующем: применяем желтое, до тех пор пока находятся искомые  $y$  и  $z$ .



## Выводы об mTERF

В литературе отмечалась отсутствие разумного механизма работы терминатора<sup>1</sup> без предположения о его строгой поляризации.

Но моделирование показывает, что возможно объяснение всех известных данных без этого предположения. Это согласуется с опытами [Ammini, Hauswirth, 1999], что у лягушки незначительная часть РНК-полимераз транскрибирует ген 16S рРНК в сторону якобы строгой поляризации

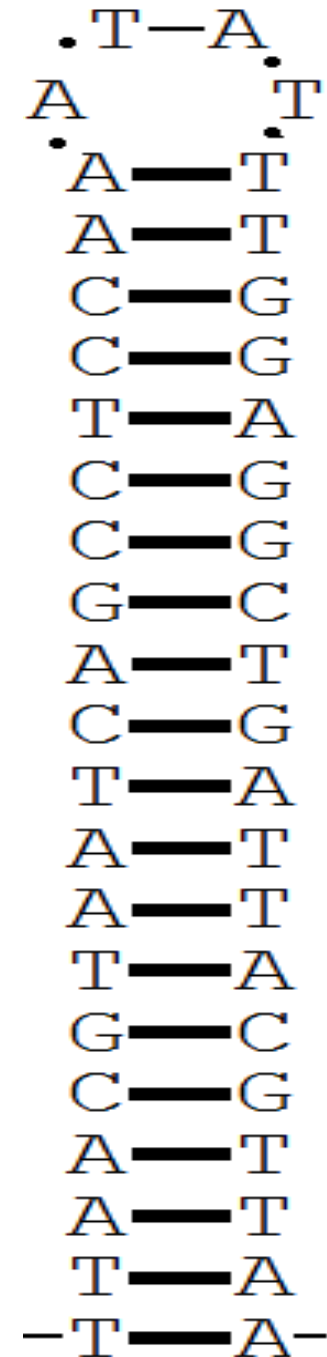
## How a terminator works?

**Terminator** forms a «**helix**»

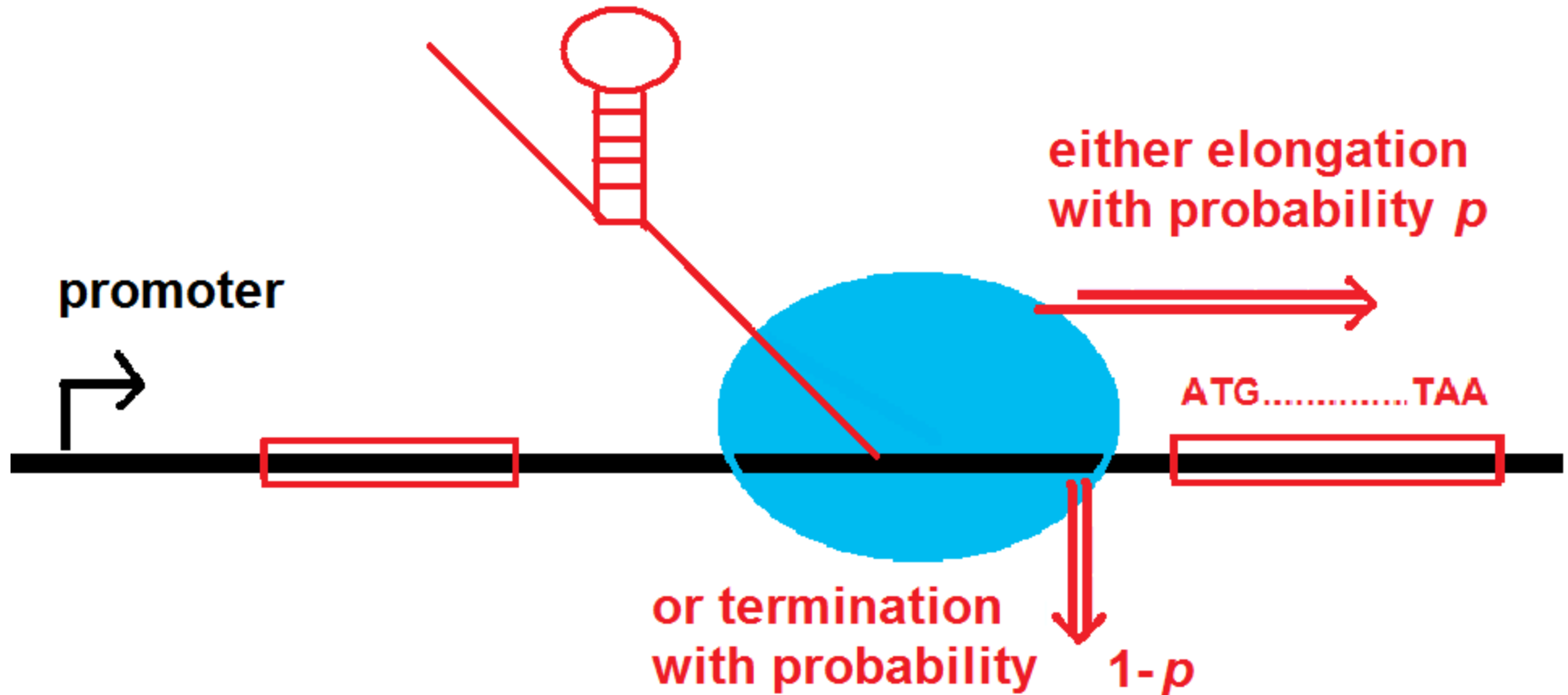
(in yellow is its left shoulder, in blue – the right shoulder).

Paired are G to C, and A to T

TTAACGTAATCAGCCTCCAAATATTTGGAGGCTGATTACGTTAA



Each terminator has a certain intensity  $1-p$  of the polymerase **detachments**

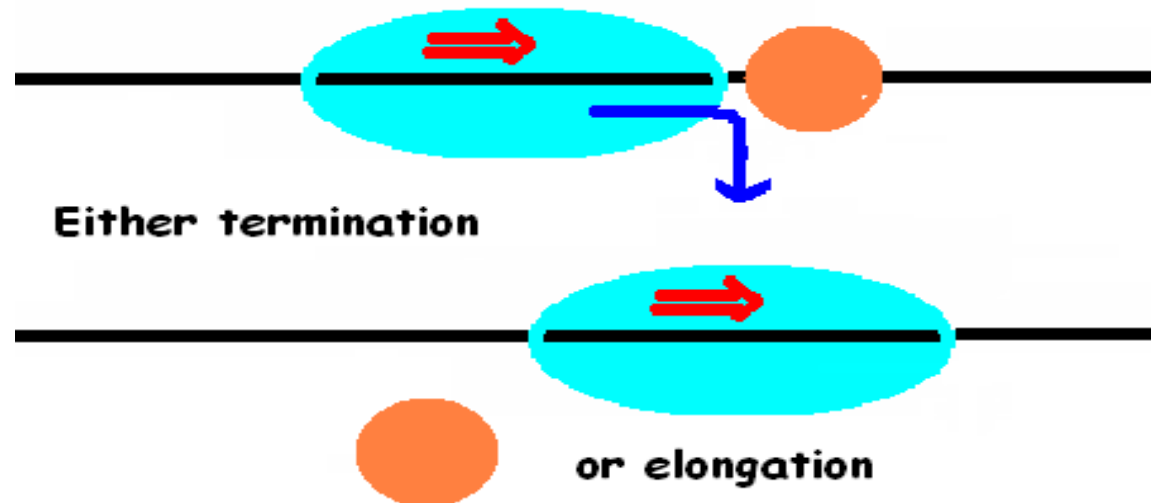


For example, **mitochondrial promoters** of:

**human**    **5'-CANACC(G)CC(A)AAAGAPyA-3'**

**frog**                                    **ACRTTATA**

У человека, крысы и лягушки известно два терминатора, которых происходит:



Один терминатор обусловлен связыванием белка mTERRF с сайтом длиной 28 на ДНК непосредственно после гена 16S rRNA и внутри гена tRNA-Leu. Транзиция A→G в позиции 3243 снижает связь mTERRF с последовательностью ДНК и вызывает тяжелые заболевания человека

Второй терминатор – строго поляризованный, поскольку терминация обусловлена формированием гуанилового квадруплекса (тетрамера) на РНК. Этот терминатор вызывает терминацию 65% транскриптов, начинающихся с LSP [Wanrooij *et al.*, 2010]. У креветки предсказано два таких терминатора

<b><i>Homo sapiens</i></b>	GGGGGAGGGGGGG
<b><i>Rattus norvegicus</i></b>	GGGGGTGGGGGGGG
<b><i>Xenopus laevis</i></b>	GGGGGGTAGGGGG
<b><i>Artemia franciscana</i></b>	GGGGGGGGGGGGGG
	GGGGGGGGGGGGGG

Полимеразы бывают двух типов: бактериальные и фаговые. Каждая сначала связывается со своим промотором (бактериального или фагового типа), а затем движется в соответствующем направлении.

Полимеразы, движущиеся друг за другом, не влияют друг на друга; вторая снижает скорость, если первая движется медленнее. При лобовом столкновении полимеразы бактериального типа всегда обе сваливаются. Фагового не обязательно.

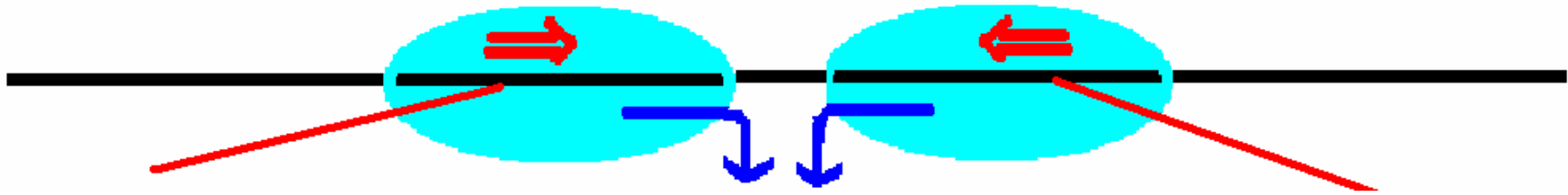
Все эти процессы совместно моделируются

1) Столкновение РНК-полимераз фагов Т3 и Т7 может не приводить к немедленной терминации [Ma, McAllister, 2009]. В этом случае, происходит одновременная транскрипция двух комплементарных цепей ДНК, что приводит к образованию дуплексов РНК, которые не транслируются и быстро распадаются.

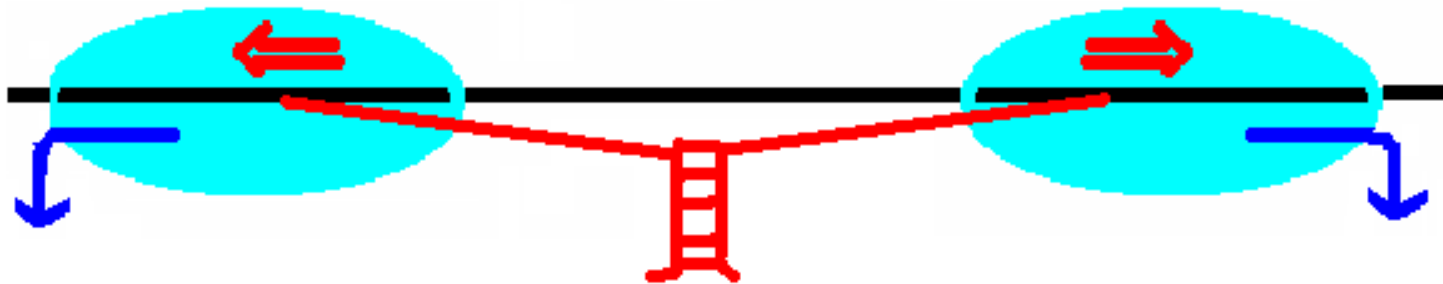
2) Образование таких дуплексов может приводить к терминации транскрипции через конкуренцию между энергетически более выгодным РНК·РНК дуплексом и РНК·ДНК дуплексом **внутри РНК-полимеразы**



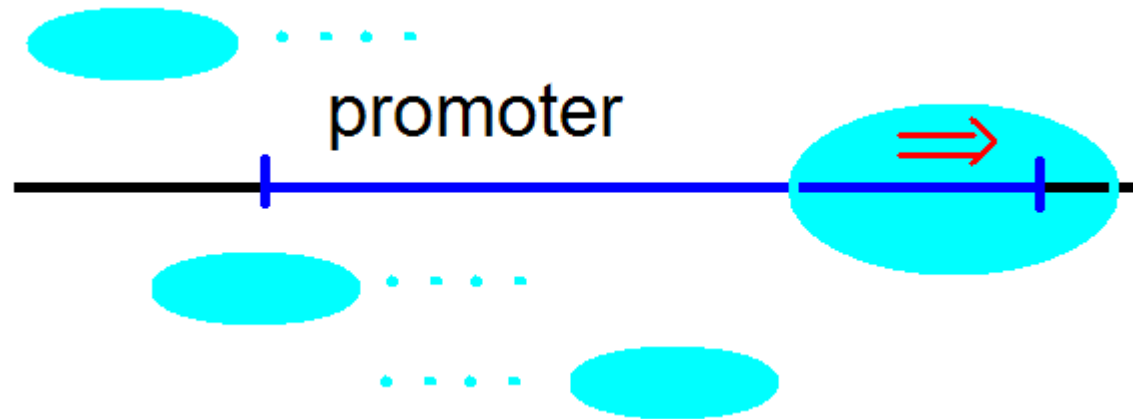
In case of a front collision both polymerases either  
**detach immediately:**



or produce two complementary RNA strands:



The promoter is ***available*** if **none of polymerases overlap with it**. Binding occurs only if the promoter is available



Each promoter, for each polymerase type, is characterized by the ***intensity of binding attempts***.

Binding attempts are allowed to form a Poisson process, with a RNA polymerase moving at a predetermined rate until colliding with another polymerase

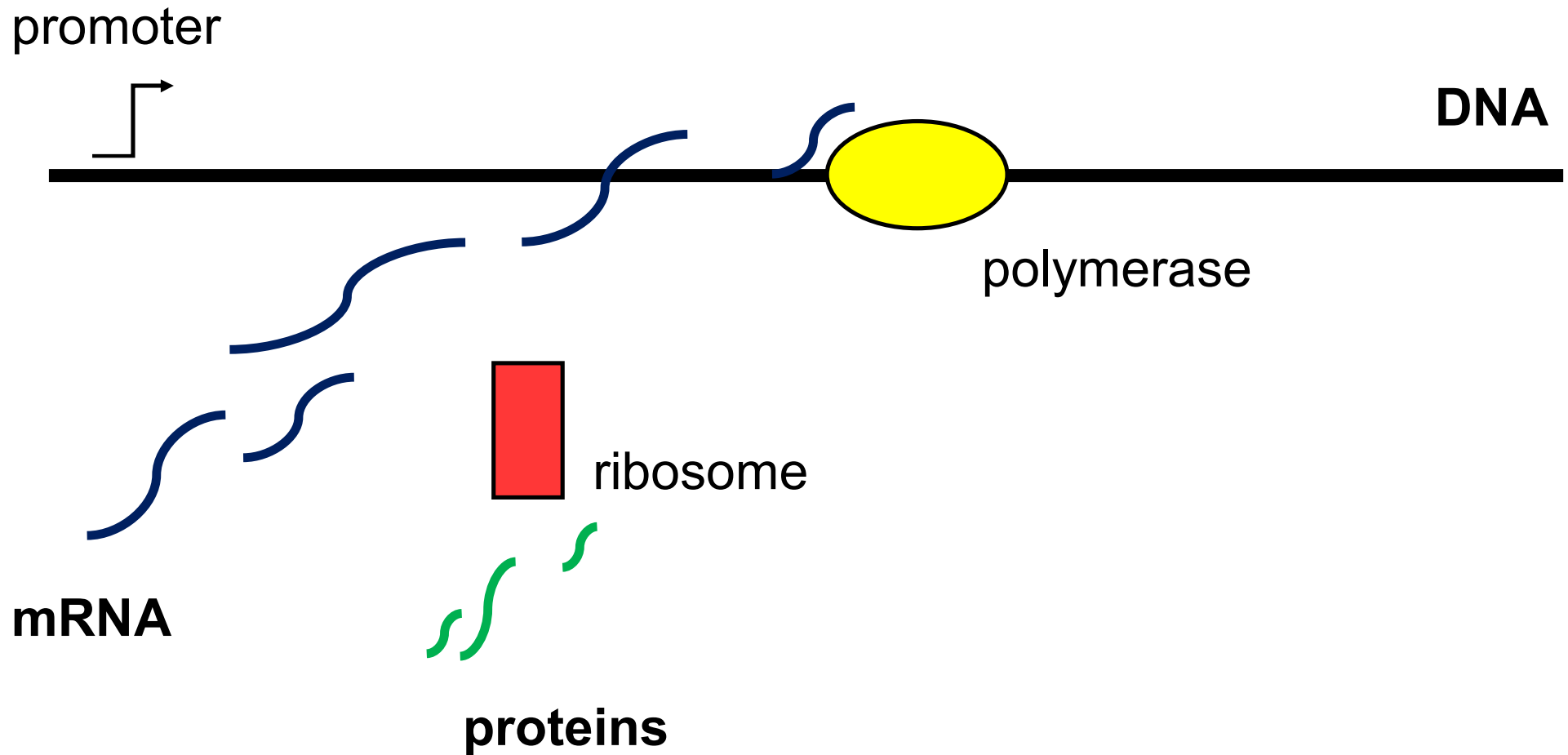
## Найденное решение для *Xenopus laevis* mitochondrion

time	QH	QL	Frog3		Frog1		Frog2	
			mTERF	LSP1	mTERF	LSP1	mTERF	LSP1
0	0.2165	0.0480	0.0173	0.0021	0.0067	0.0023	0.0031	0.0021
5	0.2165	0.0480	0.0190	0.0055	0.0118	0.0061		
6	0.2165	0.0480					0.0041	0.0023
9	0.2165	0.0480					0.0055	0.0041
10	0.2165	0.0480			0.0157	0.0061		
14	0.2165	0.0480	0.0278	0.0097	0.0209	0.0107		
16	0.2165	0.0480			0.0253	0.0157		
18	0.2165	0.0480			0.0306	0.0190		
20	0.2165	0.0480	0.0370	0.0118	0.0336	0.0209	0.0067	0.0118
23	0.2165	0.0480			0.0370	0.0542		
28	0.2165	0.0480	0.0407	0.1056				
30	0.2165	0.0480					0.0130	0.0253
48	0.2165	0.0480	0.0278	0.0655	0.0407	0.0960	0.0336	0.0872
96	0.2165	0.0480			0.0306	0.0872		
7 days	0.2165	0.0480					0.0089	0.0190

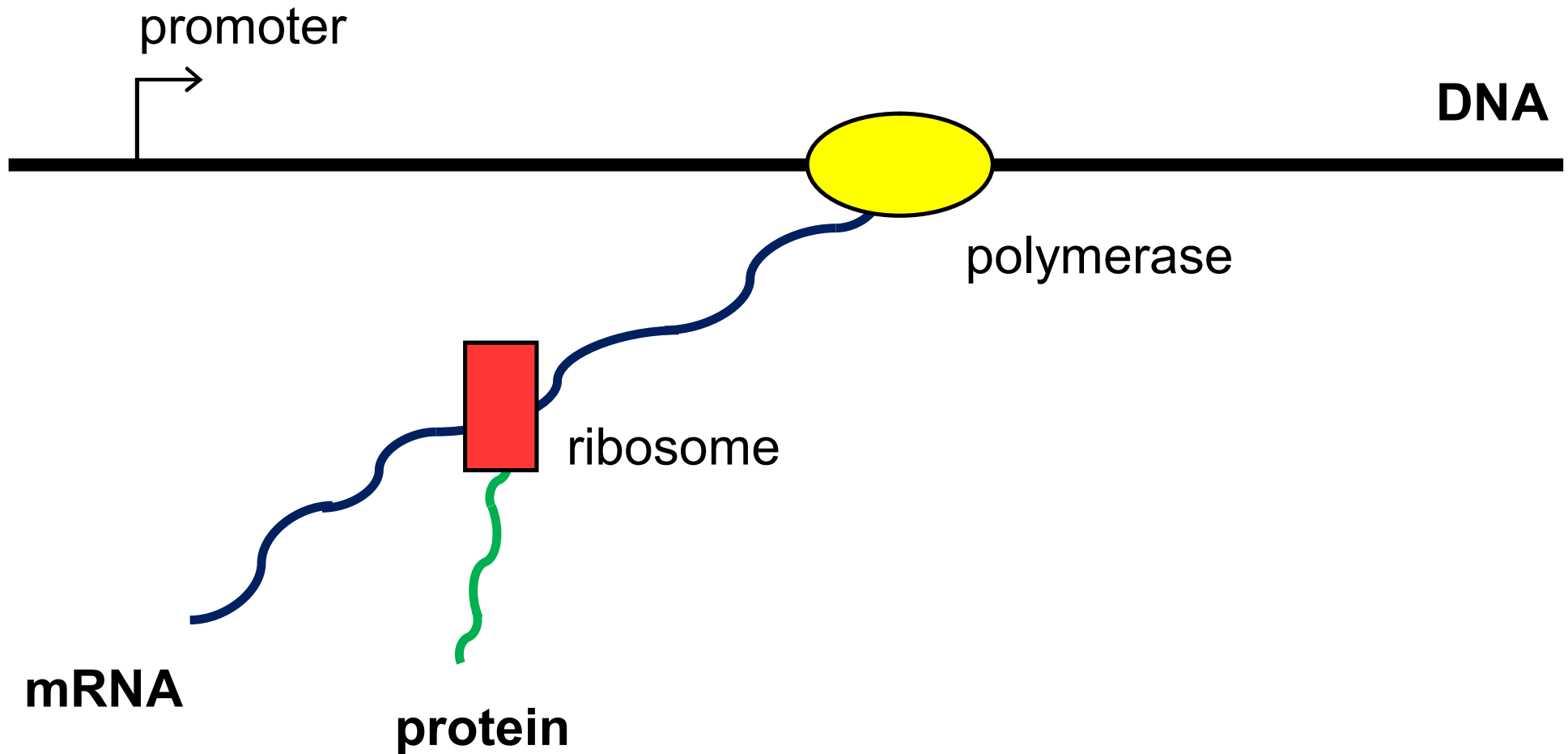
**Геном – это несколько таких ДНК (= несколько последовательностей).**

**ДНК расположены в ядре, в пластидах, в митохондриях (органеллы клетки) и все вместе эти ДНК составляют геном клетки.**

**У бактерий нет органелл и ДНК плавают прямо в цитоплазме**



Имеется **два** основных процесса в «текущей жизни клетки»: **транскрипция** и затем **трансляция**. Эти процессы могут протекать согласованно или независимо. Некоторые гены подвергаются только транскрипции



Имеется **два** основных процесса в «текущей жизни клетки»: **транскрипция** и затем **трансляция**. Эти процессы могут протекать согласованно или независимо. Некоторые гены подвергаются только транскрипции