

# Занятие 4. Геномные браузеры.

Московский Государственный Университет имени М.В.Ломоносова  
Факультет Биоинженерии и Биоинформатики

27 сентября 2013 г.

# Кодировки

Кодирование — правило сопоставления символу некоторого числа.

Кодировка DOS-866										Кодировка Windows-1251											
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9		
3	!	"	#	\$	%	&	'			3	!	"	#	\$	%	&	'				
4	(	)	*	+	,	-	/	0	1	4	(	)	*	+	,	-	/	0	1		
5	2	3	4	5	6	7	8	9	:	;	5	2	3	4	5	6	7	8	9		
6	<	=	>	?	@	A	B	C	D	E	6	<	=	>	?	@	A	B	C		
7	F	G	H	I	J	K	L	M	N	O	7	F	G	H	I	J	K	L	M		
8	P	Q	R	S	T	U	V	W	X	Y	8	P	Q	R	S	T	U	V	W		
9	Z	[	\	]	^	_	`	a	b	c	9	Z	[	\	]	^	_	`	a		
10	d	e	f	g	h	i	j	k	l	m	10	d	e	f	g	h	i	j	k		
11	n	o	p	q	r	s	t	u	v	w	11	n	o	p	q	r	s	t	u		
12	x	y	z	{		}	~	▯	А	В	12	x	y	z	{		}	~	▯		
13	В	Г	Д	Е	Ж	З	И	Й	К	Л	13	,	’	”	…	†	‡	£	¤		
14	М	Н	О	П	Р	С	Т	У	Ф	Х	14	Ь	К	Ъ	У	Ь	’	”	•		
15	Ч	Ц	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	15	-	-	▯	▯	▯	▯	▯	▯		
16	а	б	в	г	д	е	ж	з	и	й	16	у	у	Ј	а	г	;	\$	Е	Ф	
17	к	л	м	н	о	п	р	с	т	у	17	ё	ё	ч	-	а	у	’	”	і	
18	ј	ј	ј	ј	ј	ј	ј	ј	ј	ј	18	г	и	ї	•	е	н	е	»	ј	ѕ
19	ѕ	ї	А	В	В	Г	Д	Е	Ж	З	19	ѕ	ї	А	В	В	Г	Д	Е	Ж	З
20	И	Й	К	Л	М	Н	О	П	Р	С	20	И	Й	К	Л	М	Н	О	П	Р	С
21	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	21	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы
22	Ь	Э	Ю	Я	а	б	в	г	д	е	22	Ь	Э	Ю	Я	а	б	в	г	д	е
23	ж	з	и	й	к	л	м	н	о	п	23	ж	з	и	й	к	л	м	н	о	п
24	р	с	т	у	ф	х	ц	ч	ш	щ	24	р	с	т	у	ф	х	ц	ч	ш	щ
25	ъ	ы	ь	э	ю	я					25	ъ	ы	ь	э	ю	я				

Один и тот же набор бит в разных кодировках означает разные символы!

Наиболее популярные кодировки: ASCII, UTF-8, Win-1251.

## Немного биологии

В геноме гены располагаются:

- ▶ на прямой и обратных цепях;
- ▶ в трех рамках считывания на одной цепи.

Так же бывают плазмиды, опероны, экзон-интронные структуры, псевдогены и т.д.

# Плазмида

Плазмида - небольшая молекула ДНК, физически отделенная от хромосом. Автономно реплицируется, транслируется. Клетки умеют ее передавать друг другу.

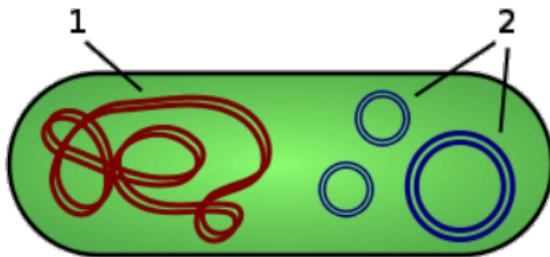
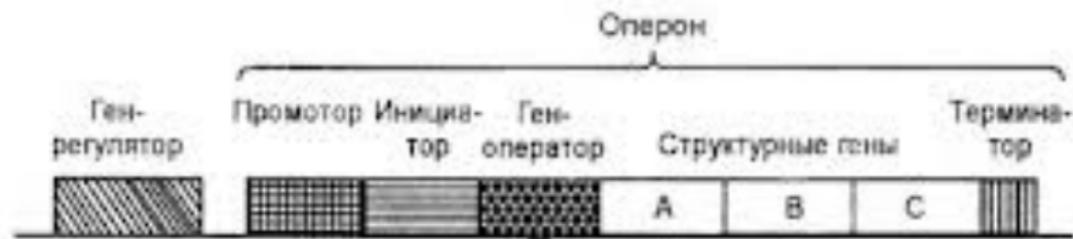


Рис. 1: 1. Геномная ДНК. 2. Плазмиды

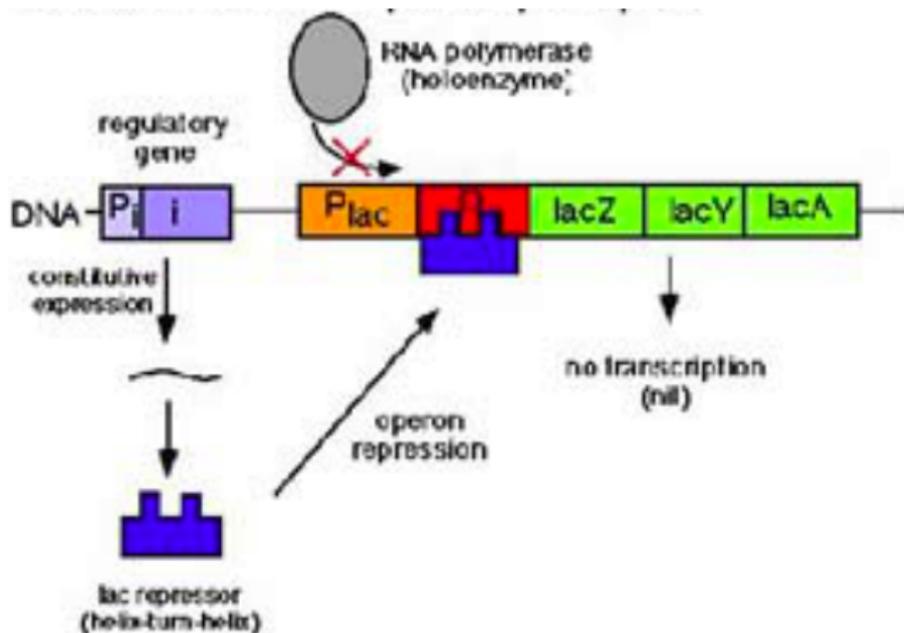
# Оперон

Оперон - группа генов, транскрибируемых и регулируемых вместе.



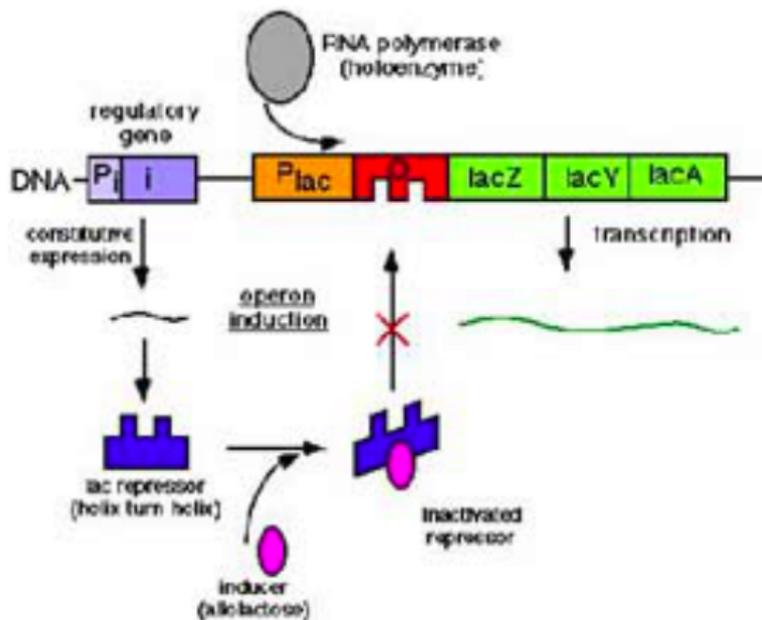
## Пример оперона. Лактозный оперон

В случае отсутствия в бактериальной клетке лактозы на промоторном участке находится белок, препятствующий транскрипции. При появлении лактозы происходит ее связывание с супрессором, его диссоциация и начинается транскрипция.



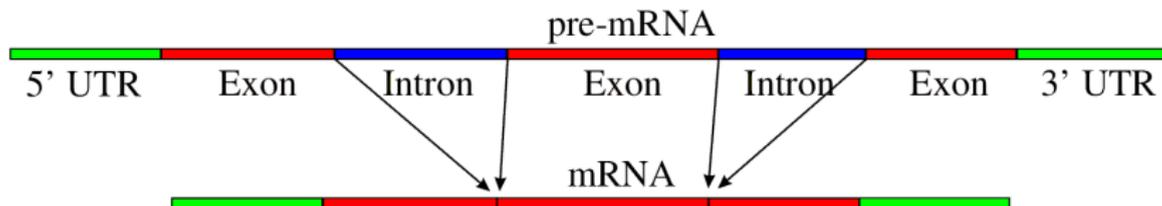
## Лактозный оперон. Стимуляция

Транскрипция слабая, ее можно усилить: в голодной клетке накапливается цАМФ, который связывается со специальным белком. Такой комплекс усиливает транскрипцию.



## Экзоны и интроны, псевдогены

1. Сплайсинг: после транскрипции вырезаются части последовательности (интроны), находящиеся между старт- и стоп-кодоном. Оставшиеся (экзоны) лигируются.



2. Псевдогены. Некоторые участки ДНК, ограниченные старт- и стоп-кодонами не транскрибируются. С некоторых РНК не транслируется белок. С некоторых РНК не транслируется белок - псевдогены.

# Визуализация геномных данных

Геномные браузеры (Genome Browser):

- ▶ NCBI (National Center for Biotechnology Information, U.S.A)
- ▶ Ensembl (The European Bioinformatics Institute)
- ▶ UCSC (University of California)
- ▶ и многие другие

# Примеры использования геномных браузеров

## Seal anellovirus TFFN/USA/2006, complete genome

NCBI Reference Sequence: NC\_015212.1

[GenBank](#) [FASTA](#)

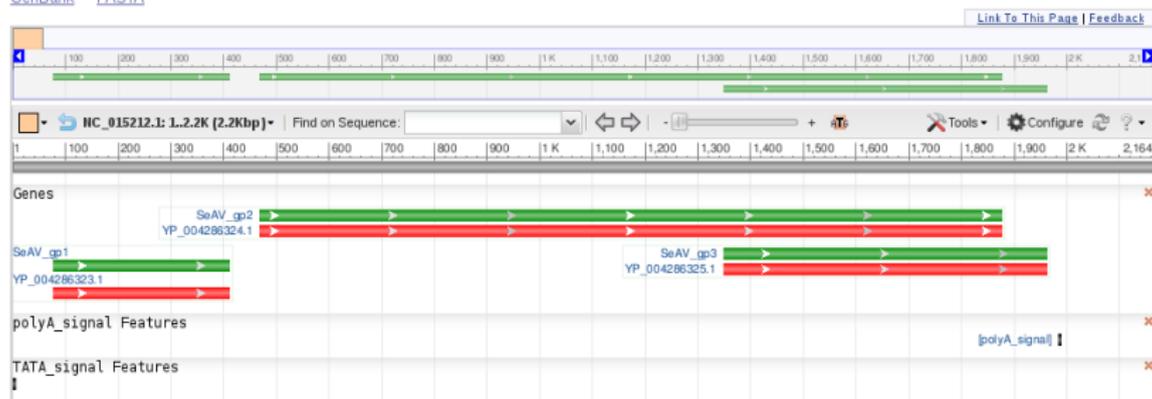


Рис. 2: Визуализация генома Seal anellovirus.