

# Стоп-кодон – не повод останавливаться

Лавренова Виктория, Максакова Мария, Нуждина Екатерина, Беседина Елизавета, Меер Маргарита

Лаборатория эволюционной биологии

проект был выполнен на средства фонда «Династия»



## Аннотация

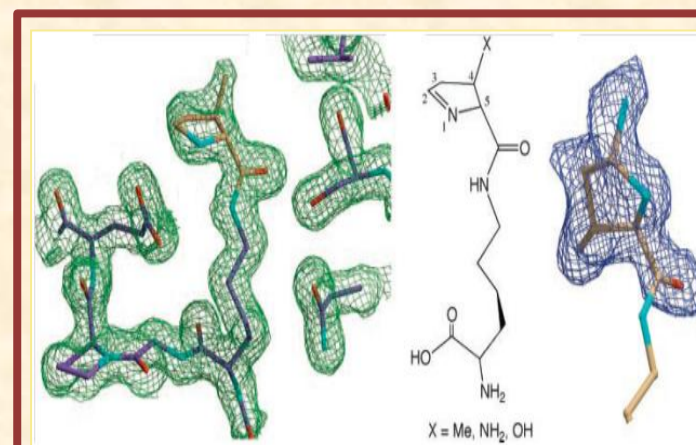
Пирролизин (Pyl, O) входит в состав некоторых белков, являясь 22-й аминокислотой генетического кода. Она зашифрована триплетом TAG, который обычно является стоп-кодonom.

В ходе работы были проанализированы стоп-кодоны в геномах 10 организмов, содержащих и не содержащих Pyl. Аннотации белков для этих организмов из GenBank'a не содержат ни одного Pyl. Ряд подобных белков был заново аннотирован.

## Введение

Аминокислота **пирролизин** была впервые обнаружена в 2002 в *Methanosarcina barkeri*. [1] Она состоит из лизинового остатка, связанного с метилированным пирролиновым кольцом. Пирролизин встречается, в основном, у археобактерий-метаногенов (в семействе *Methanosarcinaceae*), а также у двух видов бактерий: *Desulfitobacterium hafniense* и дельта-протеобактерии – симбионта червя *Olavius algarvensis*. Пирролизин нужен, в основном, для катаболизма метиламинов и входит в активный центр метиламин-метилтрансфераз. [2].

Пирролизин является «полноценной» аминокислотой, включающейся в белок на уровне трансляции, и имеет собственную аминоацил-тРНК-синтетазу (PylS). Он кодируется кодоном TAG, который в то же время является стоп-кодonom. Таким образом, у этого кодона оказывается две функции. Остаётся неизвестным механизм распознавания TAG как кодона пирролизина клеткой, что делает задачу предсказания Pyl-содержащих белков очень сложной.



Пирролизин в активном центре монометиламин-метилтрансферазы

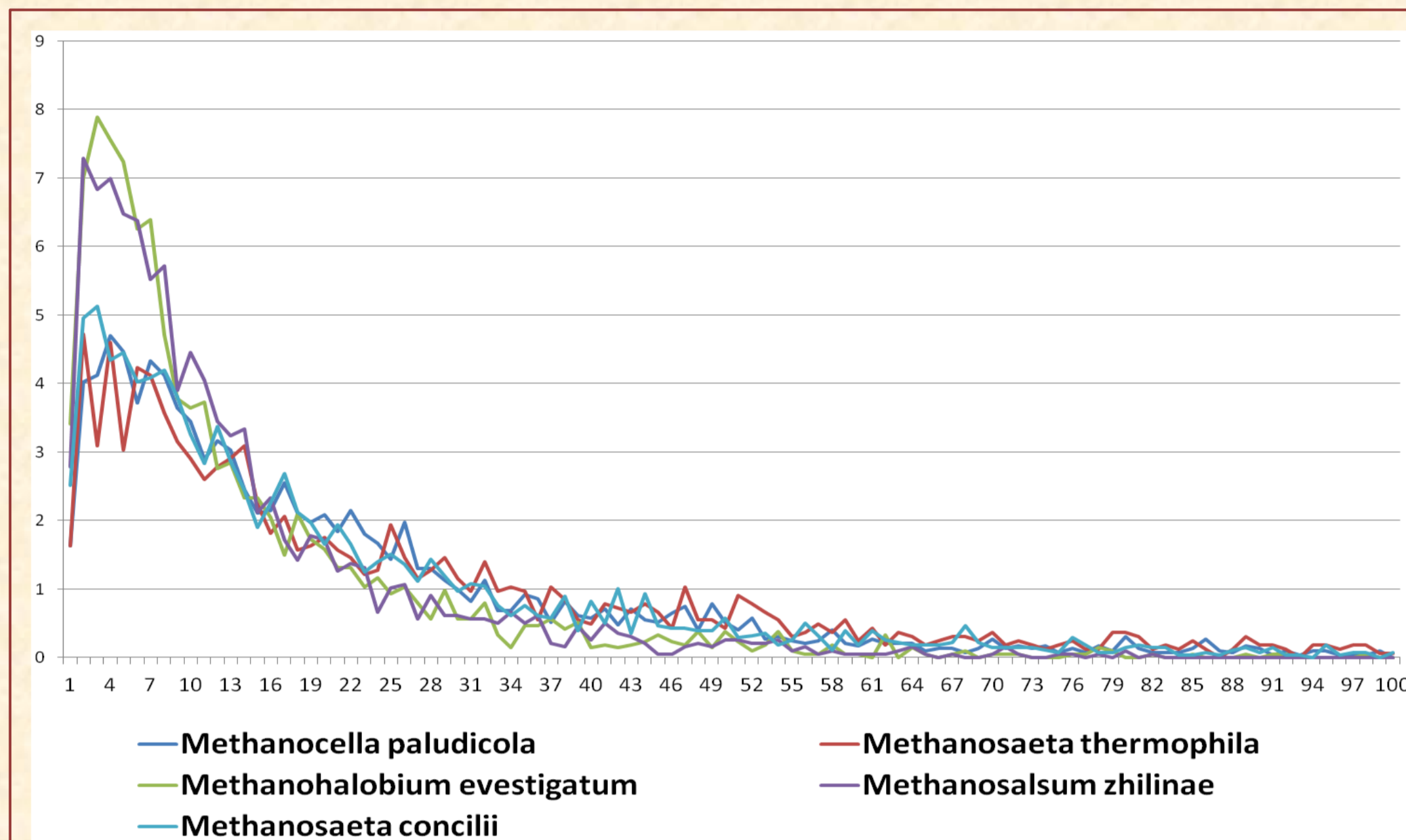
## Результаты

В аминокислотных последовательностях 10 метаногенных архей и бактерий из Genbank ни разу не встретился пирролизин, что показалось нам довольно удивительным. Мы решили:

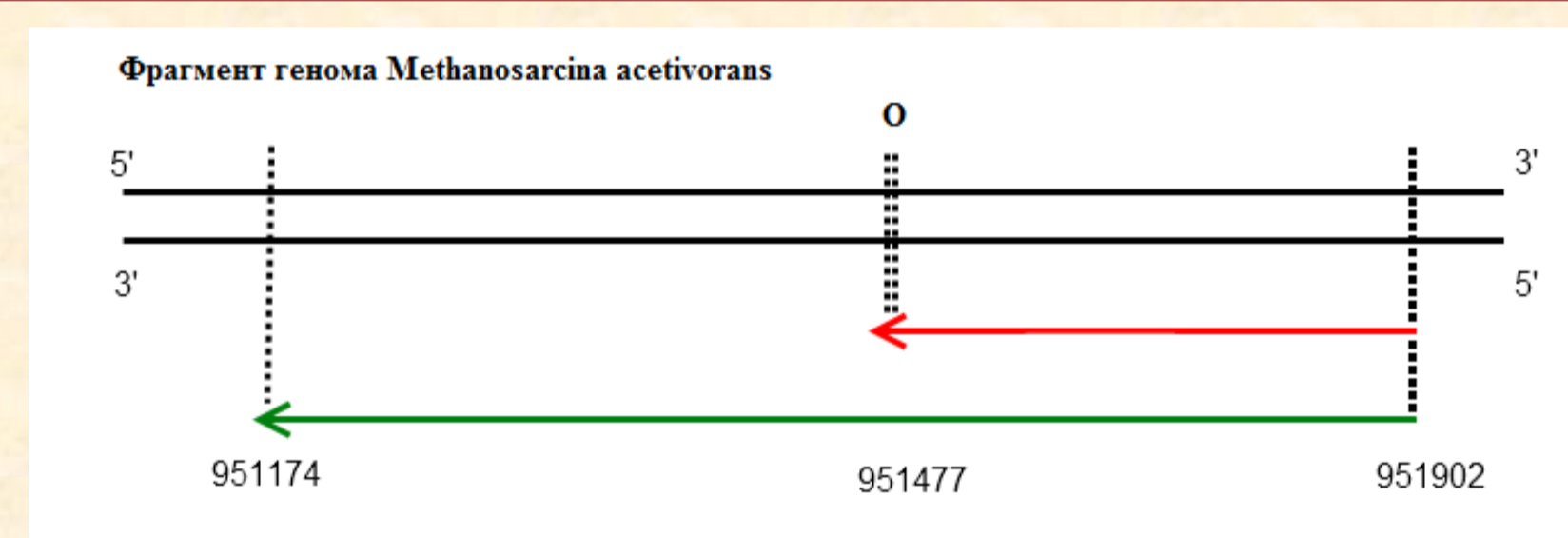
- ➔ Найти TAG в кодирующих последовательностях генов наших организмов и попытаться «транслировать» их в белки.
- ➔ Найти в литературе примеры экспериментально определённых последовательностей белков, точно содержащих пирролизин.

### Белки, содержащие пирролизин:

- ММА-метилтрансферазы (mtmB);
- ДМА-метилтрансферазы (mtbB);
- ТМА-метилтрансферазы (mttB);
- гистидил-тРНК-гуанилилтрансфераза (Thg 1); [3]
- транскрипционный фактор из семейства TetR (MA2902); [4]
- некоторые транспозазы.



Распределение расстояний между основным и дополнительным стоп-кодонами



В гене Thg 1 у *Methanosarcina acetivorans* произошла точечная мутация – замена триптофанового кодона UGG на пирролизиновый UAG. Ген Thg 1 в геноме у *Methanosarcina acetivorans* транскрибируется с комплементарной цепи, он был аннотирован с 951902 по 951477 нуклеотид. Мы нашли гомологичный ему ген у *Methanosarcina barkeri fusaro*, выровняли их и выяснили, что фермент у *M. acetivorans* транскрибируется, по-видимому, с 951902 по 951174 и приводит к синтезу более длинного белка.

Организм	N генов	%TGA	% TAA	% TAG	TAG внутри генов	Pyl внутри генов (Genbank)	%TGA(2)	%TAA(2)	%TAG(2)
<i>Methanocella paludicola</i>	3004	34	45	21	0	0	34	46	20
<i>Methanococcoides burtonii</i>	2272	53	45	2	1	0	41	42	17
<i>Methanohalobium evestigatum</i>	2148	31	68	1	8	0	28	57	15
<i>Methanohalophilus mahii</i>	1987	51	48	1	0	0	39	46	15
<i>Methanosaeta concilii</i>	2829	49	25	26	0	0	47	31	22
<i>Methanosaeta thermophila</i>	1696	67	14	19	0	0	62	21	18
<i>Methanosalsum zhilinae</i>	1976	49	46	5	14	0	40	44	16
<i>Methanosarcina acetivorans</i>	4540	46	49	5	13	0	40	44	16
<i>Methanosarcina barkeri fusaro</i>	3607	38	57	4	2	0	35	48	16
<i>Methanosarcina mazei</i>	3371	43	53	4	0	0	37	49	14

## Материалы и методы

- Программирование на языке Perl
- Нахождение гомологичных последовательностей при помощи Blast
- Выравнивание последовательностей с помощью MUSCLE
- Полные геномные последовательности из GenBank следующих видов:
 

<i>Methanocella paludicola</i>	<i>Methanococcoides burtonii</i>
<i>Methanohalobium evestigatum</i>	<i>Methanohalophilus mahii</i>
<i>Methanosaeta concilii</i>	<i>Methanosaeta thermophila</i>
<i>Methanosalsum zhilinae</i>	<i>Methanosarcina acetivorans</i>
<i>Methanosarcina barkeri fusaro</i>	<i>Methanosarcina mazei</i>
- Использование баз данных (PubMed, Protein)

## Литература

1. A New UAG-Encoded Residue in the Structure of a Methanogen Methyltransferase. Bing Hao *et al.* *Science* 296, 1462 (2002).
2. Review Article. Selenocysteine, Pyrrolysine, and the Unique Energy Metabolism of Methanogenic Archaea. Michael Rother, and Joseph A. Krzycki. *Archaea* (2010).
3. The appearance of pyrrolysine in tRNA<sup>His</sup> guanylyltransferase by neutral evolution. Ilka U. Heinemann, Patrick O'Donoghue, Catherine Madinger, Jack Benner, Lennart Randau, Christopher J. Noren, and Dieter Soll. *PNAS* (2009).
4. Mining prokaryotic genomes for unknown amino acids: a stop-codon-based approach. Masashi Fujita, Hisaaki Mihara, Susumu Goto, Nobuyoshi Esaki and Minoru Kanehisa. *BMC Bioinformatics* (2007).

## Выводы

1. Белки, содержащие пирролизин, плохо аннотированы в базах данных. В частности в GenBank.
2. TAG реже других стоп-кодonom бывает дополнительным (2).
3. Расстояние между основным стоп-кодonom и дополнительным меньше у пирролизин-содержащих организмов, чем у других.

## Благодарности

Авторский коллектив выражает искреннюю признательность фонду «Династия» за финансовую поддержку, а также Фёдору Кондрашову и Алексею Шипилёву за помощь в решении технических проблем.