

О СХОДСТВЕ СТРАТЕГИЙ КОДИРОВАНИЯ МИТОХОНДРИАЛЬНЫХ БЕЛКОВ ЧЕЛОВЕКА И ТРИХИНЕЛЛЫ В УСЛОВИЯХ КОЭВОЛЮЦИИ ИХ ГЕНОМОВ.

Е.А.Черноус, В.Э. Бутвиловский, А.В. Бутвиловский

Белорусский государственный медицинский университет

Изучены мРНК, кодирующие 12 ферментов дыхательной цепи человека, трихинеллы и цианорабдитис. Установлено, что насыщенность гуанином и цитозином изученных мРНК человека достоверно ближе к таковым трихинеллы. Выявлено сходство долей ГЦЗ-кодонов и соотношения наблюдаемой и теоретической частот претерминальных кодонов в изученных мРНК человека и трихинеллы. Установлено, что аминокислотный состав изученных белков и картина использования синонимичных кодонов в мРНК человека более сходны с таковыми трихинеллы.

Ключевые слова: стратегия кодирования, ГЦ-насыщенность, доля ГЦЗ-кодонов, претерминальные кодоны, картина использования синонимичных кодонов, аминокислотный состав, ферменты дыхательной цепи.

E.A. Chernous, V.E. Butvilovsky, A.V. Butvilovsky

ABOUT A SIMILARITY OF THE CODING STRATEGIES OF HUMAN AND TRICHINELLA'S MITOCHONDRIAL PROTEINS IN CONDITIONS OF COEVOLUTION OF THEIR GENOMES

We investigated mRNA, coding 12 respiratory chain's enzymes of man, trichinella and caenorhabditis. It was determined that guanine and cytosine content in investigated human mRNA was significantly closer to those of trichinella. The similarity of GC3-codons portion and relation of observed and theoretical frequencies of preterminal codons in investigated human and trichinella's mRNA was detected. It was determined that amino acid composition of the investigated proteins and synonymous codon's usage pattern in human mRNAs were more similar to those of trichinella.

Key words: coding strategy, GC-saturation, portion of GC3-codons, preterminal codons, synonymous codon's usage pattern, amino acid composition, respiratory chain's enzymes.

Трихинеллез – тяжелое паразитарное заболевание человека и животных, возбудителем которого является трихинелла. Трихинеллез является характерным для Республики Беларусь природно-очаговым заболеванием.

Коэволюция трихинеллы и ее хозяев хорошо изучена на видовом, организменном, тканевом, клеточном уровнях организации живого. Коэволюционным механизмам в системе “паразит–хозяин”, формирующейся при трихинеллезе на молекулярно-генетическом уровне, посвящены лишь единичные работы [17]. Изучение коэволюции паразитов и их хозяев имеет фундаментальное и прикладное значение, так как взаимосвязанные изменения компонентов систем “паразит-хозяин” (в том числе и на молекулярно-генетическом уровне) могут являться одной из причин изменения чувствительности возбудителей заболеваний к действию противопаразитарных препаратов.

На основании ранее проведенных исследований по изучению нуклеотидного состава мРНК, кодирующих 12 ферментов дыхательной цепи человека, трихинеллы и цианорабдитис, нами выдвинуто предположение о сходстве стратегий кодирования

соответствующих митохондриальных белков в системе “паразит-хозяин”, формирующейся при трихинеллезе [9].

Стратегия кодирования белка определяется картиной использования кодонов в соответствующей ему мРНК. Анализ картин использования кодонов в последовательностях нуклеиновых кислот стал возможным в 70-80-х годах прошлого столетия, когда в международных базах данных появилось достаточное количество секвенированных последовательностей РНК и ДНК [15]. С тех пор был установлен ряд важных закономерностей стратегий кодирования белков [1, 10, 16]. Многообразие возможных стратегий кодирования связано с вырожденностью генетического кода (в среднем на каждую из 20 аминокислот приходится три синонимичных кодона) [13].

В 1980-м году Р. Грэнтсем предположил, что каждый вид организмов имеет оригинальную стратегию кодирования белков [12]. Позднее установлена вариация использования кодонов и у организмов одного вида, что связано с уровнем экспрессии гена [14, 20, 27], его размером [18], структурой мРНК [19], аминокислотным составом кодируемого белка [23] и другими факторами [13].

При изучении стратегии кодирования наиболее часто анализируются следующие показатели [3, 4]:

1. Частота использования претерминальных кодонов (ПТК; кодонов, способных стать терминальными в результате одношаговой мутации).
2. Доля ГЦЗ-кодонов (кодонов, содержащих в третьем положении гуанин или цитозин, за исключением терминальных [8]).
3. Картина использования синонимичных кодонов (анализируется путем вычисления показателя относительного использования синонимичных кодонов, *Relative synonymous codon`s usage, RSCU* [4]).

Стратегия кодирования может оказывать определенное влияние на помехоустойчивость процесса трансляции (путем уменьшения или увеличения частоты претерминальных кодонов), а также его скорость и точность (путем неравномерного использования синонимичных кодонов и неодинакового содержания в клетке изоакцепторных тРНК) [17].

Важнейшим фактором, определяющим стратегию кодирования того или иного белка, является ГЦ-насыщенность (суммарное содержание гуанина и цитозина) его мРНК (и соответствующего кодирующего участка ДНК) [13].

Типичная мтДНК многоклеточных животных (*Metazoa*) высоко консервативна – это циркулярная молекула ДНК размером 14-18 тысяч пар нуклеотидов, содержащая участок для инициации репликации и транскрипции и 37 генов.

Митохондриальные ДНК секвенированы у разных таксономических групп животных (млекопитающие, птицы, земноводные, насекомые, круглые черви, кольчатые черви, кишечнополостные и т.д.). Продолжающееся в настоящее время секвенирование последовательностей геномов различных организмов позволяет изучать картины использования кодонов в большем числе нуклеотидных последовательностей РНК и ДНК, что обеспечивает биологическую значимость получаемых выводов.

Установлено, что митохондриальные геномы нематод объединяет ряд уникальных свойств [28]:

1. Меньший размер большинства генов по сравнению с таковыми других многоклеточных животных.
2. Отсутствие гена, кодирующего субъединицу 8 F₀ АТФ-синтазы.
3. Ни один из генов не использует АУГ в качестве старт-кодона, а ряд генов и вовсе использует для инициации необычные кодоны ГТТ и ТТГ.
4. В кодируемых тРНК отсутствуют дигидроуридиловые или псевдоуридиловые петли и т.д.

Известно [22], что мтДНК трихинеллы сходна с мтДНК других нематод (размер и структура кодируемых рРНК), с типичной мтДНК (37 генов, сходство их расположения, инициация АУХ-кодонами, 5'-АЦГ-3' антикодон в аргининовой тРНК), а по ряду параметров занимает промежуточное между ними положение (размер белок-кодирующих генов, структура тРНК).

Цель исследования: сравнить ГЦ-насыщенность, долю ГЦЗ-кодонов и частоту использования претерминальных кодонов, картины использования синонимичных кодонов в мРНК, кодирующих митохондриальные белки человека и трихинеллы, а также аминокислотный состав данных белков.

Материалы и методы исследования. Проанализированы взятые с сервера NCBI (National Center for Biotechnology information, www.ncbi.nlm.nih.gov) нуклеотидные последовательности мРНК, кодирующих ряд митохондриальных белков (субъединицы 1–4, 4L, 5, 6 НАДН-дегидрогеназы, цитохром *b*, субъединицы 1–3 цитохром-*c*-оксидазы, субъединица 6 АТФ-синтазы) человека (*Homo sapiens* (H.s), [26]) и трихинеллы (*Trichinella spiralis* (Tr.sp.), [22]). Нуклеотидные последовательности мРНК, кодирующих субъединицы 3,4L,6 НАДН-дегидрогеназы, субъединицу 2 цитохром-*c*-оксидазы трихинеллы получены самостоятельно в рамках гранта БРФФИ [7]. В качестве контроля использованы мРНК, кодирующие аналогичные белки свободноживущего круглого червя цианорабдитис (*Caenorhabditis elegans* (C.el), [28]). Нуклеотидный состав, использование кодонов, показатели относительного использования синонимичных кодонов (RSCU, relative synonymous codon's usage) и аминокислотный состав изучены при помощи программы MEGA 3 [21, 24]. Сходство стратегий кодирования изучаемых белков в мРНК определено по дистанции МакАйнерни:

$$D_{jk} = \sum_{i=1}^n \frac{abs(RSCU_{ji} - RSCU_{ki})}{n},$$

где *n* – число синонимичных вырожденных кодонов для определенной таблицы генетического кода, *j_i* – кодон *i* последовательности *j*, *k_i* – кодон *i* последовательности *k* [25].

Полученные результаты были обработаны статистически [2, 6] с помощью пакета программ Microsoft Excel 2000, достоверность различий определена по критерию Стьюдента.

Результаты и обсуждение. На основании полученных данных по нуклеотидному составу изученных мРНК вычислена их ГЦ-насыщенность (таблица 1).

Таблица 1

ГЦ-насыщенность в мРНК, кодирующих ряд митохондриальных белков человека, трихинеллы, цианорабдитис

Фермент/организм	<i>Homo sapiens</i>	<i>Trichinella spiralis</i>	<i>Caenorhabditis elegans</i>
НАДН-дегидрогеназа-1	47,6±1,61 ^{2,3}	35,6±1,60 ^{1,3}	25,2±1,47 ^{1,2}

НАДН-дегидрогеназа-2	43,1±1,53 ^{2,3}	35,3±1,60 ^{1,3}	19,8±1,37 ^{1,2}
НАДН-дегидрогеназа-3	40,5±2,64 ^{2,3}	29,6±2,44 ^{1,3}	21,4±2,24 ^{1,2}
НАДН-дегидрогеназа-4	44,4±1,34 ^{2,3}	38,7±1,39 ^{1,3}	22,2±1,18 ^{1,2}
НАДН-дегидрогеназа-4L	43,0±2,87 ³	36,6±3,07 ³	18,3±2,53 ^{1,2}
НАДН-дегидрогеназа-5	45,0±1,17 ^{2,3}	38,2±1,23 ^{1,3}	22,7±1,05 ^{1,2}
НАДН-дегидрогеназа-6	43,0±2,16 ^{2,3}	28,3±3,60 ^{1,3}	20,0±1,92 ^{1,2}
Цитохром <i>b</i>	45,9±1,48 ^{2,3}	35,2±1,43 ^{1,3}	26,2±1,32 ^{1,2}
Цитохром- <i>c</i> -оксидаза-1	46,2±1,27 ^{2,3}	37,6±1,23 ^{1,3}	29,7±1,15 ^{1,2}
Цитохром- <i>c</i> -оксидаза-2	46,2±1,91 ^{2,3}	36,5±1,85 ^{1,3}	25,9±1,66 ^{1,2}
Цитохром- <i>c</i> -оксидаза-3	46,9±1,79 ^{2,3}	34,9±1,71 ^{1,3}	27,9±1,62 ^{1,2}
АТФ-синтаза 6	44,5±1,90 ^{2,3}	32,3±1,76 ^{1,3}	24,0±1,74 ^{1,2}

Примечание – Здесь и далее знаком ¹ обозначено достоверное ($p < 0,05$) различие по сравнению с *Homo sapiens*, ² – *Trichinella spiralis*, ³ – *Caenorhabditis elegans*.

Наибольшая ГЦ-насыщенность характерна для изученных мРНК человека, меньшая – трихинеллы, а наименьшая – для мРНК цианорабдитис. Установлено, что ГЦ-насыщенность изученных мРНК человека, более сходна с таковой мРНК трихинеллы. Различия содержания гуанина и цитозина в данных мРНК человека и трихинеллы достоверны для 11 из 12 сравниваемых мРНК. Обращает на себя внимание тот факт, что между ГЦ-насыщенностью всех изученных мРНК трихинеллы и цианорабдитис существуют достоверные различия ($p < 0,05$). Полученные данные позволяют сделать предположение о том, что сходство ГЦ-насыщенности мРНК человека и трихинеллы, по-видимому, является одним из проявлений коэволюции в системе “паразит-хозяин” на молекулярно-генетическом уровне.

Доля ГЦЗ-кодонов в изученных мРНК и ее зависимость от ГЦ-содержания представлены в таблице 2 и на рисунке 1, соответственно.

Таблица 2

Доля ГЦЗ-кодонов в мРНК, кодирующих ряд митохондриальных белков человека, трихинеллы и цианорабдитис

Фермент / организм	<i>Homo sapiens</i>	<i>Trichinella spiralis</i>	<i>Caenorhabditis elegans</i>
НАДН-дегидрогеназа-1	51,7±2,80 ^{2,3}	31,3±2,68 ^{1,3}	13,4±1,99 ^{1,2}
НАДН-дегидрогеназа-2	47,4±2,68 ^{2,3}	39,5±2,84 ^{1,3}	15,6±2,16 ^{1,2}
НАДН-дегидрогеназа-3	42,6±4,61 ³	31,6±4,30 ³	17,0±3,55 ^{1,2}
НАДН-дегидрогеназа-4	48,1±2,33 ^{2,3}	40,8±2,42 ^{1,3}	11,5±1,58 ^{1,2}
НАДН-дегидрогеназа-4L	45,5±5,00 ³	40,2±5,41 ³	11,5±3,61 ^{1,2}
НАДН-дегидрогеназа-5	49,1±2,03 ^{2,3}	42,2±2,17 ^{1,3}	15,5±1,57 ^{1,2}
НАДН-дегидрогеназа-6	39,4±3,69 ^{2,3}	28,7±3,61 ^{1,3}	10,3±2,52 ^{1,2}
Цитохром <i>b</i>	50,8±2,57 ^{2,3}	30,9±2,40 ^{1,3}	15,6±1,88 ^{1,2}
Цитохром- <i>c</i> -оксидаза-1	47,1±2,20 ^{2,3}	30,1±2,02 ^{1,3}	13,1±1,47 ^{1,2}
Цитохром- <i>c</i> -оксидаза-2	48,3±3,31 ^{2,3}	30,5±3,06 ^{1,3}	9,9±1,96 ^{1,2}
Цитохром- <i>c</i> -оксидаза-3	47,7±3,10 ^{2,3}	28,3±2,80 ^{1,3}	16,4±2,31 ^{1,2}
АТФ-синтаза 6	46,3±3,31 ^{2,3}	35,0±2,87 ^{1,3}	10,5±2,17 ^{1,2}

Установлено, что, сходно с ГЦ-насыщенностью, максимальная доля ГЦЗ-кодонов наблюдается в изученных мРНК человека, меньшая – трихинеллы, а минимальная – в мРНК цианорабдитис. Доля ГЦЗ-кодонов в изученных мРНК человека более сходна с таковой в мРНК трихинеллы по сравнению с контролем.

На рисунке 1 видно сходство по ГЦ-насыщенности и доли ГЦЗ-кодонов мРНК человека и трихинеллы. Вместе с тем обнаружено, что характер связи между данными показателями у человека и трихинеллы сходен ($r=0,89\pm 0,097$) и достоверно ($p<0,05$) отличается от такового у цианорабдитис ($r=0,13\pm 0,314$), что подтверждается и различиями наклонов соответствующих линейных трендов.

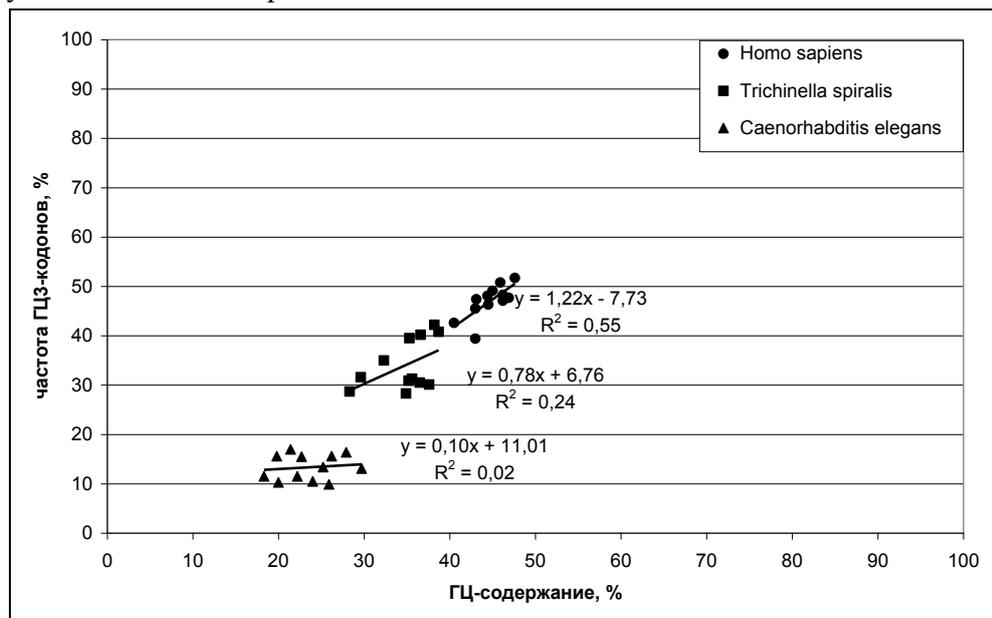


Рис. 1. Зависимость между ГЦ-насыщенностью и долей ГЦЗ-кодонов в изученных мРНК человека, трихинеллы и цианорабдитис

Частота использования ПТК в изученных мРНК представлена в таблице 3.

Таблица 3

Частота использования претерминальных кодонов в мРНК, кодирующих ряд митохондриальных белков человека, трихинеллы и цианорабдитис

Фермент / организм	Homo sapiens	Trichinella spiralis	Caenorhabditis elegans
НАДН-дегидрогеназа-1	31,4±2,60 ²	21,3±2,36 ^{1,3}	29,1±2,66 ²
НАДН-дегидрогеназа-2	32,5±2,51	30,7±2,68	31,8±2,77
НАДН-дегидрогеназа-3	33,9±4,41 ²	20,5±3,73 ¹	30,4±4,35
НАДН-дегидрогеназа-4	32,0±2,18	26,9±2,18	28,0±2,22
НАДН-дегидрогеназа-4L	25,3±4,37	26,8±4,89	23,1±4,77
НАДН-дегидрогеназа-5	29,1±1,85	25,2±1,91	27,1±1,93
НАДН-дегидрогеназа-6	43,4±3,75 ^{2,3}	24,2±3,42 ¹	27,6±3,71 ¹
Цитохром <i>b</i>	31,8±2,39 ²	22,9±2,18 ^{1,3}	29,7±2,37 ²
Цитохром- <i>c</i> -оксидаза-1	30,4±2,03 ²	17,5±1,67 ^{1,3}	25,7±1,91 ²
Цитохром- <i>c</i> -оксидаза-2	29,8±3,03	25,2±2,89	31,9±3,06
Цитохром- <i>c</i> -оксидаза-3	33,9±2,93 ^{2,3}	20,5±2,51 ¹	25,0±2,71 ¹
АТФ-синтаза 6	26,4±2,93	22,7±2,52	30,0±3,24

Из приведенных данных видно, что различия между частотами использования ПТК в изучаемых мРНК, как правило, недостоверны. Между ГЦ-содержанием мРНК и частотой претерминальных кодонов обнаружена слабая недостоверная корреляционная связь. Необходимо отметить, что корреляционная связь между частотой претерминальных кодонов мРНК и их ГЦ-насыщенностью была ранее охарактеризована как сильная, достоверная и обратная [5, 11]. Несовпадение полученных нами данных с данными литературы вероятно связано с отличиями таблиц генетического кода для мРНК, кодирующих митохондриальные белки позвоночных (человека) и беспозвоночных животных (трихинеллы и цианорабдитис). Так, в первой из них кодоны АГА и АГГ являются терминальными, а во второй – соответствуют серину. Разное количество терминальных кодонов в этих двух таблицах генетического кода обуславливает различия числа претерминальных кодонов в них. В таблице генетического кода для митохондриальных белков позвоночных терминальными являются 4 кодона, смысловыми – 60 кодонов, претерминальными – 24 кодона, поэтому теоретическая частота ПТК – $24/60 \times 100\% = 40,0\%$. В таблице генетического кода для митохондриальных белков беспозвоночных животных терминальными являются 2 кодона, смысловыми – 62 кодона, претерминальными – 14 кодонов, что дает теоретическую частоту ПТК, равную $14/62 \times 100\% = 22,6\%$. Для корректного сравнения частоты использования претерминальных кодонов в изучаемых мРНК человека и круглых червей, с нашей точки зрения, необходимо вычислить соотношения наблюдаемой и теоретической частот ПТК (таблица 4).

Установлено, что среднее соотношение наблюдаемой и теоретически ожидаемой частот претерминальных кодонов в изученных мРНК человека равно $0,79 \pm 0,034$, в мРНК трихинеллы – $1,05 \pm 0,048$, в мРНК цианорабдитис – $1,25 \pm 0,036$. При этом различия между этими показателями человека и *C.elegans* достоверно ($p < 0,05$) отличаются, а человека и *T.spiralis* неразличимы статистически. Это свидетельствует о наличии сходства стратегий кодирования митохондриальных белков человека и трихинеллы в мРНК.

Таблица 4

Соотношение наблюдаемой и теоретической частот использования претерминальных кодонов в мРНК, кодирующих ряд митохондриальных белков человека, трихинеллы и цианорабдитис

Фермент/организм	<i>Homo sapiens</i>	<i>Trichinella spiralis</i>	<i>Caenorhabditis elegans</i>
НАДН-дегидрогеназа-1	0,79	0,94	1,29
НАДН-дегидрогеназа-2	0,81	1,36	1,41
НАДН-дегидрогеназа-3	0,85	0,91	1,35
НАДН-дегидрогеназа-4	0,80	1,19	1,24
НАДН-дегидрогеназа-4L	0,63	1,19	1,02
НАДН-дегидрогеназа-5	0,73	1,12	1,20
НАДН-дегидрогеназа-6	1,09	1,07	1,22
Цитохром-b	0,80	1,01	1,31
Цитохром-c-оксидаза-1	0,76	0,77	1,14
Цитохром-c-оксидаза-2	0,75	1,12	1,41
Цитохром-c-оксидаза-3	0,85	0,91	1,11
АТФаза 6	0,66	1,00	1,33

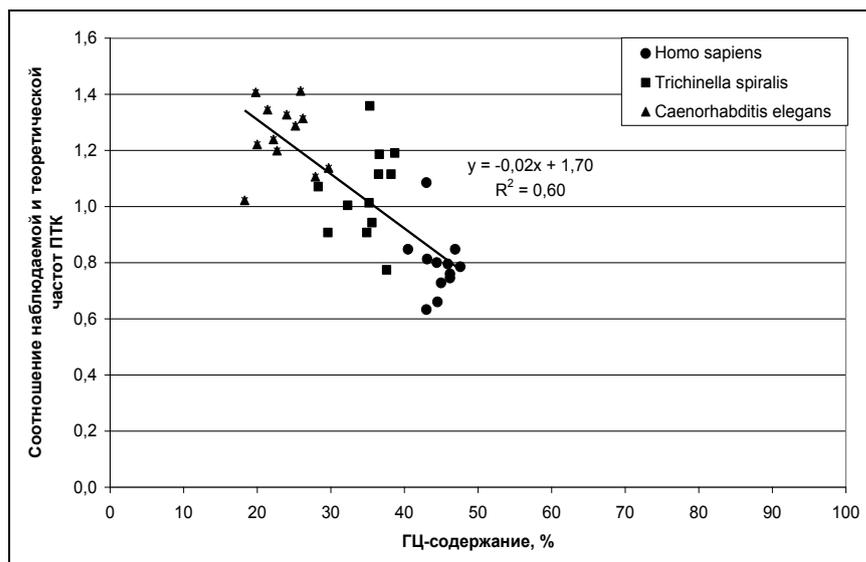


Рис. 2. Зависимость соотношения наблюдаемых и теоретических частот ПТК от ГЦ-содержания изученных мРНК человека, трихинеллы и цианорабдитис

На рисунке 2 видно, что по соотношению наблюдаемых и теоретических частот ПТК и ГЦ-содержанию изученные мРНК человека ближе к таковым трихинеллы, а не контроля. Между соотношением наблюдаемых и теоретических частот ПТК и ГЦ-содержанием во всех изученных мРНК обнаружена сильная достоверная отрицательная корреляционная связь ($r = -0,78 \pm 0,107$, $p < 0,001$).

Для каждой из 20 аминокислот вычислены суммарные RSCU соответствующих ей ГЦЗ-кодонов. Средние значения показателя относительного использования синонимичных кодонов в мРНК, кодирующих ряд митохондриальных белков человека, трихинеллы и цианорабдитис, представлены в таблице 5.

Таблица 5

Средние значения RSCU для мРНК, кодирующих ряд митохондриальных белков человека, трихинеллы и цианорабдитис

Аминокислота (соответствующие ГЦЗ-кодоны)	Среднее RSCU для мРНК H.s.	Среднее RSCU для мРНК Tr.sp.	Среднее RSCU для мРНК C.el.	p для RSCU по мРНК H.s. и мРНК Tr.sp.	p для RSCU по мРНК H.s. и мРНК C.el.
Фен (УУЦ)	1,08±0,146	0,84±0,198	0,10±0,033	>0,05	<0,001
Лей (ЦУЦ, ЦУГ, УУГ)	2,13±0,161	2,00±0,441	0,93±0,099	>0,05	<0,001
Иле (АУЦ)	1,07±0,128	0,79±0,179	0,18±0,037	>0,05	<0,001
Мет (АУГ)	0,45±0,119	0,65±0,235	0,53±0,082	>0,05	>0,05
Вал (ГУЦ, ГУГ)	1,53±0,180	1,26±0,282	0,52±0,075	>0,05	<0,001
Сер (УЦЦ, УЦГ, АГЦ+АГГ для круглых червей)	3,29±0,265	2,40±0,129	1,04±0,180	<0,01	<0,001
Про (ЦЦЦ, ЦЦГ)	2,08±0,263	0,75±0,173	0,64±0,239	<0,001	<0,001
Тре (АЦЦ, АЦГ)	1,77±0,196	1,09±0,143	0,58±0,343	<0,05	<0,01

Ала (ГЦЦ, ГЦГ)	2,20±0,201	1,07±0,355	0,62±0,221	<0,05	<0,001
Тир (УАЦ)	1,29±0,140	0,88±0,212	0,29±0,053	>0,05	<0,001
Гис (ЦАЦ)	1,38±0,207	0,92±0,253	0,31±0,177	>0,05	<0,001
Глн (ЦАГ)	0,13±0,049	0,37±0,233	0,33±0,117	>0,05	>0,05
Асн (ААЦ)	1,55±0,120	0,90±0,208	0,11±0,036	<0,05	<0,001
Лиз (ААГ)	0,24±0,084	0,67±0,253	0,18±0,062	>0,05	>0,05
Асп (ГАЦ)	1,37±0,226	0,65±0,224	0,45±0,212	<0,05	<0,05
Глу (ГАГ)	0,45±0,154	0,70±0,267	0,52±0,111	>0,05	>0,05
Цис (УГЦ)	1,19±0,276	0,50±0,229	0,08±0,087	>0,05	<0,001
Три (УГГ)	0,21±0,087	0,57±0,221	0,32±0,177	>0,05	>0,05
Арг (ЦГЦ, ЦГГ)	2,17±0,327	0,90±0,369	0,05±0,050	<0,05	<0,001
Гли (ГГЦ, ГГГ)	2,41±0,162	1,09±0,154	0,56±0,148	<0,001	<0,001

Установлено, что по 15-ти аминокислотам (75,0±9,93%) средние значения RSCU для мРНК человека достоверно ($p<0,001$) ближе к таковым трихинеллы по сравнению с контролем. Это свидетельствует о наличии коэволюции компонентов изучаемой системы “паразит-хозяин” на молекулярно-генетическом уровне. В пользу такой точки зрения говорит и то, что по 12-ти аминокислотам (60,0±11,24%) средние значения RSCU для мРНК человека и трихинеллы не являются статистически различимыми. Для сравнения, статистически незначимыми являются отличия средних значений RSCU для мРНК человека и цианорабдитис лишь по 5-ти аминокислотам (25,0±9,93%).

Для учета особенностей кодирования отдельных белков в мРНК нами вычислены дистанции МакАйнерни для мРНК, кодирующих каждый из 12 изученных ферментов дыхательной цепи (табл. 6).

Таблица 6

Дистанции МакАйнерни для мРНК, кодирующих митохондриальные белки человека и трихинеллы (опыт), а также человека и цианорабдитис (контроль)

Фермент / организм	Опыт	Контроль	p
НАДН-дегидрогеназа-1	0,010±0,0015	0,018±0,0020	<0,01
НАДН-дегидрогеназа-2	0,022±0,0016	0,016±0,0016	<0,05
НАДН-дегидрогеназа-3	0,013±0,0022	0,017±0,0021	>0,05
НАДН-дегидрогеназа-4	0,024±0,0017	0,017±0,0018	<0,05
НАДН-дегидрогеназа-4L	0,023±0,0023	0,020±0,0024	>0,05
НАДН-дегидрогеназа-5	0,023±0,0013	0,020±0,0021	>0,05
НАДН-дегидрогеназа-6	0,021±0,0021	0,015±0,0019	<0,05
Цитохром <i>b</i>	0,009±0,0013	0,018±0,0017	<0,001
Цитохром- <i>c</i> -оксидаза-1	0,008±0,0009	0,016±0,0016	<0,001
Цитохром- <i>c</i> -оксидаза-2	0,010±0,0014	0,017±0,0016	<0,01
Цитохром- <i>c</i> -оксидаза-3	0,010±0,0014	0,017±0,0016	<0,01
АТФ-синтаза 6	0,008±0,0013	0,018±0,0019	<0,001

Из приведенных в таблице 6 данных видно, что достоверные отличия дистанций МакАйнерни характерны для 9 изученных мРНК, кодирующих (75,0±13,06%) белки человека и трихинеллы (по сравнению с таковыми для мРНК человека и цианорабдитис). При этом полученные дистанции для шести из них (для мРНК, соответствующих субъединице 1

НАДН-дегидрогеназы, цитохрому *b*, субъединицам 1-3 цитохром-*c*-оксидазы, субъединице 6 АТФ-синтазы) меньше таковых контроля, а для трех из них (соответствующих субъединицам 2, 4, 6 НАДН-дегидрогеназы) выше таковых контроля.

Однако при коэволюции человека и трихинеллы на молекулярно-генетическом уровне следовало бы ожидать достоверно меньших дистанций МакАйнерни опыта по сравнению с контролем для всех изученных мРНК. Это наводит на мысль о существовании некоторых особенностей такой методики изучения картин использования кодонов.

Для их определения нами проведен тщательный сравнительный анализ показателей RSCU для кодонов изучаемых мРНК человека, трихинеллы и цианорабдитис. Установлено, что особенностью анализа МакАйнерни является то, что биологическая значимость результатов напрямую связана с числом кодонов нуклеиновой кислоты, по которым вычисляется дистанция.

Рассмотрим эту особенность на примере кодонов УГУ и УГЦ (соответствуют цистеину) мРНК, кодирующих субъединицу 6 НАДН-дегидрогеназы человека, трихинеллы и цианорабдитис (табл. 7).

Таблица 7

Показатели RSCU и количество кодонов УГУ и УГЦ в мРНК, кодирующих ряд митохондриальных белков человека, трихинеллы и цианорабдитис

Организм/ показатель	мРНК, кодирующие субъединицу 6 НАДН-дегидрогеназы		12 изученных мРНК	
	Кодон УГУ	Кодон УГЦ	Кодон УГУ	Кодон УГЦ
Человек	n=1 RSCU=2,0	n= 0 RSCU=0	n=4 RSCU=0,4	n=17 RSCU=1,6
Трихинелла	n= 0 RSCU=0	n=2 RSCU=2,0	n=44 RSCU=1,7	n=7 RSCU=0,3
Цианорабдитис	n=3 RSCU=2,0	n= 0 RSCU=0	n=39 RSCU=1,9	n=2 RSCU=0,1

Анализируя значения показателя относительного использования синонимичных кодонов в мРНК, кодирующих субъединицу 6 НАДН-дегидрогеназы, можно прийти к выводу, что картины использования кодонов УГУ и УГЦ в данной мРНК человека сходны с таковыми в мРНК цианорабдитис и диаметрально противоположны таковым мРНК трихинеллы. На этом примере четко видно, что данные значения RSCU получены по малому числу синонимичных кодонов (1–3) и, следовательно, обладают низкой репрезентативностью. Такие значения RSCU будут вносить существенный вклад в вычисляемую на их основании дистанцию МакАйнерни, анализ которой может привести к ошибочным выводам.

Очевидным путем устранения данного недостатка методики МакАйнерни является совместный анализ и вычисление дистанций максимально возможного (в пределах поставленной цели исследования) количества последовательностей нуклеиновых кислот. При учете установленной особенности дистанция МакАйнерни может быть рекомендована в качестве эффективного способа определения сходства картин использования синонимичных кодонов в последовательностях РНК и ДНК.

В таблице 7 видно, что значения показателей RSCU по 12 изученным мРНК человека ближе к таковым трихинеллы по сравнению с контролем. Вычисленная нами общая дистанция МакАйнерни для 12 изучаемых мРНК человека и трихинеллы ($0,010 \pm 0,0009$) в 1,5 раза меньше таковой для мРНК человека и цианорабдитис ($0,015 \pm 0,0017$; различия достоверны, $p < 0,05$). Этот факт свидетельствует о статистически значимом сходстве картин использования синонимичных кодонов в мРНК, кодирующих исследуемые митохондриальные белки человека и трихинеллы. Наличие такого сходства, с нашей точки зрения, является признаком сходства аминокислотного состава кодируемых белков.

В таблице 8 показана доля аминокислот группы GARP, кодируемых ГЦ-богатыми кодонами – глицина (G; ГГУ, ГГЦ, ГГА, ГГГ), аланина – (A; ГЦУ, ГЦЦ, ГЦА, ГЦГ), аргинина (R; ЦГУ, ЦГЦ, ЦГА, ЦГГ) и пролина (P; ЦЦУ, ЦЦЦ, ЦЦА, ЦЦГ).

Таблица 8

Доля аминокислот GARP в митохондриальных белках человека, трихинеллы и цианорабдитис

Фермент / показатель	Доля GARP у H.s., %	Доля GARP у Tr.sp., %	Доля GARP у C.el., %	p по доле GARP у H.s. и Tr.sp	p по доле GARP у H.s. и C.el.
НАДН-дегидрогеназа-1	21,4±2,30	14,7±2,05	13,8±2,02	<0,05	<0,05
НАДН-дегидрогеназа-2	17,3±2,03	7,8±1,56	5,0±1,30	<0,001	<0,001
НАДН-дегидрогеназа-3	18,3±3,61	9,5±2,72	8,1±2,59	>0,05	<0,05
НАДН-дегидрогеназа-4	16,3±1,72	14,6±1,74	10,8±1,53	>0,05	<0,05
НАДН-дегидрогеназа-4L	16,3±1,72	11,1±3,49	6,5±2,81	>0,05	<0,05
НАДН-дегидрогеназа-5	18,3±1,57	14,1±1,53	9,1±1,25	>0,05	<0,001
НАДН-дегидрогеназа-6	25,3±3,30	8,3±2,21	7,6±2,21	<0,001	<0,001
Цитохром <i>b</i>	20,9±2,09	15,9±1,90	14,9±1,85	>0,05	<0,05
Цитохром- <i>c</i> -оксидаза-1	24,0±1,89	20,4±1,78	18,7±1,70	>0,05	<0,05
Цитохром- <i>c</i> -оксидаза-2	20,3±2,67	13,8±2,30	13,4±2,24	>0,05	>0,05
Цитохром- <i>c</i> -оксидаза-3	20,0±2,48	12,8±2,08	14,9±2,23	<0,05	>0,05
АТФ-синтаза 6	20,8±2,70	8,7±1,70	10,1±2,14	<0,001	<0,01

Установлено, что содержание аминокислот группы GARP неразличимо статистически в 7-ми изученных белках ($58,3 \pm 14,87\%$) человека и трихинеллы и лишь в 2-х белках человека и цианорабдитис ($16,7 \pm 11,25\%$; $p < 0,05$). Исходя из предположения о коэволюции человека и трихинеллы (как компонентов системы «паразит-хозяин», формирующейся при трихинеллезе) на молекулярно-генетическом уровне следовало бы ожидать, что данный параметр у человека и трихинеллы будет неразличимым для всех изученных белков. Отличия полученных данных от теоретически ожидаемых связаны, вероятно, с функциональными особенностями изучаемых ферментов.

Перейдем к анализу закономерностей содержания кодируемых ГЦ-бедными кодонами аминокислот FYMINK (табл. 9) – фенилаланина (F; УУУ, УУЦ), тирозина – (Y; УАУ, УАЦ), метионина (M; АУА, АУГ), изолейцина (I; АУУ, АУЦ), аспарагина (N; ААУ, ААЦ) и лизина (K; ААА, ААГ)

Таблица 9

Доля аминокислот FYMINK в митохондриальных белках человека, трихинеллы и цианорабдитис

Фермент / показатель	Доля FYMINK у H.s., %	Доля FYMINK у Tr.sp., %	Доля FYMINK у C.el., %	р по доле FYMINK у H.s.и Tr.sp.	р по доле FYMINK у H.s.и C.el.
НАДН-дегидрогеназа-1	28,0±2,52	34,1±2,74	39,5±2,87	>0,05	<0,01
НАДН-дегидрогеназа-2	32,6±2,52	30,2±2,67	46,1±2,97	>0,05	<0,001
НАДН-дегидрогеназа-3	30,5±4,29	44,0±4,61	45,9±4,73	<0,05	<0,05
НАДН-дегидрогеназа-4	29,0±2,12	25,1±2,14	42,3±2,44	>0,05	<0,001
НАДН-дегидрогеназа-4L	30,6±4,66	29,6±5,07	53,3±5,69	>0,05	<0,01
НАДН-дегидрогеназа-5	31,0±1,88	29,5±2,00	43,3±2,16	>0,05	<0,001
НАДН-дегидрогеназа-6	27,0±3,37	43,6±3,97	44,5±4,14	<0,01	<0,01
Цитохром <i>b</i>	31,5±2,39	36,7±2,50	37,3±2,51	>0,05	>0,05
Цитохром- <i>c</i> -оксидаза-1	31,2±2,05	33,1±2,08	30,9±2,02	>0,05	>0,05
Цитохром- <i>c</i> -оксидаза-2	27,3±2,96	30,2±3,06	32,9±3,09	>0,05	>0,05
Цитохром- <i>c</i> -оксидаза-3	26,2±2,73	32,3±2,92	36,1±3,01	>0,05	<0,05
АТФ-синтаза <i>b</i>	31,0±3,08	42,4±2,97	36,7±3,42	<0,05	>0,05

Содержание аминокислот FYMINK неразличимо статистически в 9-ти изученных белках (75,0±13,06%) человека и трихинеллы и лишь в 4-х белках человека и цианорабдитис (33,3±14,21%, $p < 0,05$).

Полученные данные по содержанию аминокислот FYMINK и GARP свидетельствуют о том, что аминокислотный состав изучаемых белков также является объектом коэволюции человека и трихинеллы как компонентов системы «паразит-хозяин», формирующейся при трихинеллезе на молекулярно-генетическом уровне.

Выводы:

1. Насыщенность гуанином и цитозином, доля ГЦЗ-кодонов и соотношение наблюдаемой и теоретической частот претерминальных кодонов в мРНК, кодирующих митохондриальные белки человека, достоверно ближе к таковым трихинеллы по сравнению с контролем.

2. Характер зависимости между ГЦ-насыщенностью и долей ГЦЗ-кодонов в изученных мРНК человека и трихинеллы сходен и отличается от такового цианорабдитис.

3. Картина использования синонимичных кодонов в мРНК, кодирующих исследуемые митохондриальные белки человека достоверно более сходна с таковой трихинеллы по сравнению с контролем.

4. Содержание аминокислот групп GARP и FYMINK в изученных митохондриальных белках человека достоверно ближе к таковым трихинеллы по сравнению с цианорабдитис.

Литература

1. Бутвиловский А.В. Динамика изменений нуклеотидных последовательностей матричных РНК и аминокислотных последовательностей алкогольдегидрогеназ в процессе эволюции: автореф. ... дис. канд. мед. наук: 03.00.04. – Минск, 2007. – 22 с.

2. Бредфорд Хилл А. Основы медицинской статистики. – М., 1958.

3. Бутвиловский, А.В. Об использовании претерминальных кодонов и кодонов, содержащих гуанин и цитозин в нуклеотидных последовательностях мРНК алкогольдегидрогеназ человека 1-7 типов / А.В.

Бутвиловский, Е.В. Барковский // Материалы международного симпозиума “Молекулярные механизмы регуляции функции клетки”. – Тюмень: Издательство “ВекторБук”, 2005. – С. 275-277.

4. Бутвиловский, А.В. Стратегия кодирования алкогольдегидрогеназ человека / А.В. Бутвиловский, Е.В. Барковский, В.Э. Бутвиловский // Вестник ВГМУ. - Витебск, 2006, №3. – С. 24–29.

5. Изменения в процессе эволюции мутационного давления в последовательностях мРНК, кодирующих алкогольдегидрогеназы класса 3 хордовых животных / Бутвиловский А.В. [и др.] // Медицинский журнал. – Минск, 2007. – №1. – С. 22–25.

6. Медик, В.А. Статистика в медицине и биологии: руководство в 2-х томах / В.А. Медик, М.С. Токмачев, Б.Б. Фишман // Под ред. Ю.М. Комарова. Т.1. Теоретическая статистика. – М.: Медицина, 2000. – С. 283–289.

7. Секвенирование и анализ нуклеотидных последовательностей генов, кодирующих ряд ферментов дыхательной цепи круглых червей: Отчет о НИР (заключ.) / Грант БРФФИ №Б08М-084 Руководитель работы Е.А. Черноус. - Мн.: 2010. 59 с.

8. Справочник терминов молекулярной эволюции и филогенетики. Учебно-методическое пособие / В.Э. Бутвиловский [и др.] // Мн.: БГМУ, 2006. – 40 с.

9. Стратегия кодирования ряда митохондриальных белков в системе «паразит-хозяин», формирующейся при трихинеллезе / В.Э. Бутвиловский [и др.] // Труды V Республиканской научно-практической конференции «Достижения и перспективы развития современной паразитологии». – Витебск: ВГМУ, 2006. – С. 140–143.

10. Черноус Е.А., Бутвиловский А.В., Линник Ю.И., Давыдов А.В. Стратегия кодирования NADH-дегидрогеназ круглых червей // Материалы XI Міжнародного медичного конгресу студентів і молодих вчених – Тернопіль: Укрмедкнига, 2007. – С. 290.

11. Хрусталева, В.В. Частота использования претерминальных кодонов в матричных РНК бактериальных аденилатциклаз / В.В. Хрусталева, Е.В. Барковский // Здоровоохранение. – 2006. – №2. – С. 17–20.

12. Codon catalog usage and the genome hypothesis / R. Grantham [et al]. // Nucl. Acids Res. – 1980. – Vol. 8. – P. 49–62.

13. Codon usage between genomes is constrained by genome-wide mutational process / S.L. Chen [et. al] // Proc. Natl. Acad. Sci. – 2004. – V. 101. – P. 3480–3485.

14. Codon usage patterns in *Escherichia coli*, *Bacillus subtilis*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Drosophila melanogaster* and *Homo sapiens*; a review of the considerable within-species diversity / P.M. Sharp [et. al]. // Nucl. Acids Res. – 1988. – V.16. – P. 8207–8211.

15. Codon catalog usage is a genome strategy modulated for gene expressivity / R.Grantham [et. al] // Nucleic Acids Res. - 1981. - Vol. 9. – P. 43-74.

16. Codon usage: mutational bias, translational selection or both? / P.M. Sharp [et. al] // Biochem. Soc. Trans. – 1993. – Vol. 21. – P. 835-841.

17. Cutter, A.D. The evolution of biased codon and amino acid usage in nematode genomes / A.D. Cutter, J.D. Wasmuth, M.L. Blaxter // Mol. Biol. Evol. – 2006. – V. 23. – P. 2302–2315.

18. Eyre-Walker, A. Synonymous codon bias is related to gene length in *Escherichia coli*: selection for translational accuracy? / A. Eyre-Walker // Mol. Biol. Evol. – 1996. – Vol. 13. – P. 864–872.

19. Hasegawa, M. Secondary structure of MS2 phage RNA and bias in code word usage / M. Hasegawa, T. Yasunaga, T. Miyata // Nucl. Acids Res. – 1979. – V. 7. – P. 2073–2079.

20. Ikemura, T. Codon usage and tRNA content in unicellular and multicellular organisms / T. Ikemura // Mol. Biol. Evol. – 1985. – V. 2. – P. 13–34.

21. Kumar, S. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment / S. Kumar, K.Tamura, M. Nei // Brief. Bioinform.. – 2004. – Vol. 5. – P. 150-163.

22. Lavrov, D.V. *Trichinella spiralis* mtDNA. A nematode mitochondrial genome that encodes a putative ATP8 and normally structured tRNAs and has a gene arrangement relatable to those of coelomate metazoans / D.V. Lavrov, W.M. Brown // Genetics. – 2001. – Vol. 157 (2). – P. 621-637.

23. Lobry, J.R. Hydrophobicity, expressivity and aromaticity are the major trends of amino-acid usage in 999 *Escherichia coli* chromosome-encoded genes / J.R. Lobry, C. Gautier // Nucl. Acids Res. – 1994. – V. 22. – P. 3174–3180.

24. Nei, M. Molecular Evolution and Phylogenetics / M. Nei, S. Kumar // Oxford University Press, New York, 2000.

25. McInerney, J.O. GCUA: General codon usage analysis / J.O. McInerney // Bioinformatics. – 1998. – Vol. 14(4). – P.372-373.

26. Mitochondrial genome variation and the origin of modern humans / Ingman M. [et. al] // Nature. – 2000. – Vol. 408. – P. 708-713.

27. Rocha, E.P.C. Codon usage bias from tRNA`s point of view: redundancy, specialization and efficient decoding for translation optimization / E.P.C. Rocha // Genome Res. – 2004. – V. 14. – P. 2279–2286.

28. The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum* / R. Okimoto [et. al] // Genetics. – 1992. – Vol. 130 (3). – P. 471-498.