

*В.Э. Бутвиловский,
Ю.И. Линник,
А.В. Бутвиловский,
Е.В. Барковский*

О сходстве стратегий кодирования митохондриальных белков человека и трихинеллы. Часть 1: гц-насыщенность, доля гцз-кодонов и частота использования претерминальных кодонов

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. Ключевые слова: стратегия кодирования, ГЦ-насыщенность, доля ГЦЗ-кодонов, претерминальные кодоны, ферменты дыхательной цепи.

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 [12]. 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 [7].

Стратегия кодирования белка определяется картиной использования кодонов в соответствующей ему мРНК. Многообразие возможных стратегий кодирования связано с вырожденностью генетического кода (в среднем на каждую из 20 аминокислот приходится три синонимичных кодона) [10].

В 1980-м году Р. Грэнтсем предположил, что каждый вид организмов имеет оригинальную стратегию кодирования белков [9]. Позднее установлена вариация использования кодонов и у организмов одного вида, что связано с

уровнем экспрессии гена [11, 15, 21], его размером [13], структурой мРНК [14], аминокислотным составом кодируемого белка [18] и другими факторами [10].

При изучении стратегии кодирования наиболее часто анализируются следующие показатели [2, 3]:

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 [6]).

3. Картина использования синонимичных кодонов (анализируется путем вычисления показателя относительного использования синонимичных кодонов, *Relative synonymous codon's usage, RSCU* [3]).

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) [12].

Важнейшим фактором, определяющим стратегию кодирования того или иного белка, является ГЦ-насыщенность (суммарное содержание гуанина и цитозина) его мРНК (и соответствующего кодирующего участка ДНК) [10].

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. Данные гены можно разделить на 2 группы:

1) гены, кодирующие структурные РНК митохондриального белок-синтезирующего аппарат (большую и малую субъединицы рРНК и 22 тРНК, включая по две лейциновых и сериновых);

2) гены, кодирующие белки (субъединицы 1-6 и 4L НАДН-дегидрогеназы, цитохром b, субъединицы 1-3 цитохром – с – оксидазы и субъединицы 6 и 8 FO АТФ-синтазы).

Митохондриальные ДНК секвенированы у разных таксономических групп животных (млекопитающие, птицы, земноводные, насекомые, круглые черви, кольчатые черви, кишечнополостные и т.д.). Установлено, что митохондриальные геномы нематод объединяет ряд уникальных свойств [22]:

1. Меньший размер большинства генов по сравнению с таковыми других многоклеточных животных.

2. Отсутствие гена, кодирующего субъединицу 8 FO 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 [17], 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 стратегии кодирования митохондриальных белков человека и трихинеллы.

Цель данной части исследования: сравнить ГЦ-насыщенность, долю ГЦЗ-кодонов и частоту использования претерминальных кодонов в мРНК, кодирующих митохондриальные белки человека и трихинеллы.

Материал и методы

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 – с – оксидазы, субъединица 6 АТФ-синтазы) человека (*Homo sapiens*, [20]) и трихинеллы (*Trichinella spiralis*, [17]). В качестве контроля использованы мРНК, кодирующие аналогичные белки свободноживущего круглого червя цианорабдитис (*Caenorhabditis elegans*, [22]). Нуклеотидный состав и использование кодонов изучены при помощи программы MEGA 3 [16, 19]. Полученные результаты были обработаны статистически [1, 5] 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 (табл. 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 ⁴	36,6±3,07 ⁴	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 b	45,9±1,48 ^{2,3}	35,2±1,43 ^{1,3}	26,2±1,32 ^{1,2}
Цитохром-с-оксидаза-1	46,2±1,27 ^{2,3}	37,6±1,23 ^{1,3}	29,7±1,15 ^{1,2}
Цитохром-с-оксидаза-2	46,2±1,91 ^{2,3}	36,5±1,85 ^{1,3}	25,9±1,66 ^{1,2}
Цитохром-с-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, а не цианорабдитис. Вместе с тем 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.

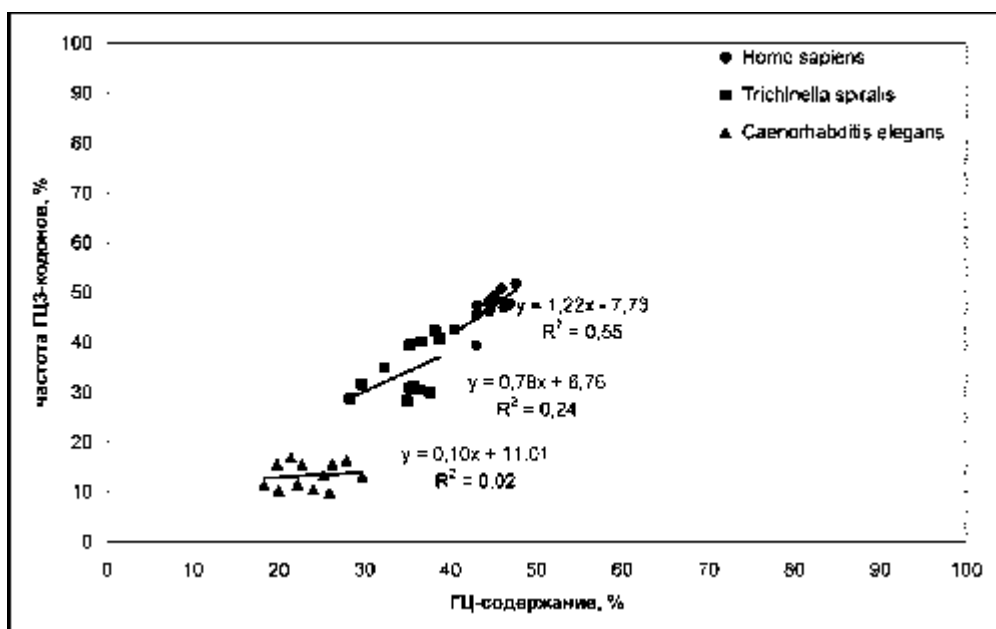


Рис. 1. Зависимость между ГЦ-насыщенностью и долей ГЦ3-кодонов в изученных мРНК человека, трихинеллы и цианорабдитис

Таблица 2

Доля ГЦ3-кодонов в мРНК, кодирующих ряд митохондриальных белков человека, трихинеллы и цианорабдитис

Фермент / организм	Homo sapiens	Trichinella spiralis	Caenorhabditis eli
НАДН-дегидрогеназа-1	51,7±2,80 ²³	31,3±2,68 ¹³	13,4±1,99 ¹⁷
НАДН-дегидрогеназа-2	47,4±2,68 ²³	39,5±2,84 ¹³	15,6±2,16 ¹⁷
НАДН-дегидрогеназа-3	42,6±4,61 ³	31,6±4,30 ³	17,0±3,55 ¹²
НАДН-дегидрогеназа-4	48,1±2,33 ²³	40,8±2,42 ¹³	11,5±1,58 ¹²
НАДН-дегидрогеназа-4L	45,5±5,00 ³	40,2±5,41 ³	11,5±3,61 ¹³
НАДН-дегидрогеназа-5	49,1±2,03 ²³	42,2±2,17 ¹³	15,5±1,57 ¹³
НАДН-дегидрогеназа-6	39,4±3,69 ²³	28,7±3,61 ¹³	10,3±2,52 ¹³
Цитохром b	50,8±2,57 ²³	30,9±2,40 ¹³	15,6±1,88 ¹²
Цитохром-с-оксидаза-1	47,1±2,20 ²³	30,1±2,02 ¹³	13,1±1,47 ¹²
Цитохром-с-оксидаза-2	48,3±3,31 ²³	30,5±3,06 ¹³	9,9±1,96 ¹⁷
Цитохром-с-оксидаза-3	47,7±3,10 ²³	28,3±2,80 ¹³	16,4±2,31 ¹⁷
АТФ-синтаза 6	46,3±3,31 ²³	35,0±2,87 ¹³	10,5±2,17 ¹²

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.

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	<i>Homo sapiens</i>	<i>Trichinella spiralis</i>	<i>Caenorhabditis elegans</i>
НАДН-дегидрогеназа-1	31,4±2,60 ²	21,3±2,36 ¹³	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 ²³	24,2±3,42 ¹	27,6±3,71 ¹
Цитохром <i>b</i>	31,8±2,39 ²	22,9±2,18 ¹³	29,7±2,37 ²
Цитохром-с-оксидаза-1	30,4±2,03 ²	17,5±1,67 ¹³	25,7±1,91 ²
Цитохром-с-оксидаза-2	29,8±3,03	25,2±2,89	31,9±3,06
Цитохром-с-оксидаза-3	33,9±2,93 ²³	20,5±2,51 ¹	25,0±2,71 ¹
АТФ синтаза <i>b</i>	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 [4, 8]. 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 – $24/60 \cdot 100\% = 40,0\%$. 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 (табл. 4).

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- <i>b</i>	0,80	1,01	1,31
Цитохром- <i>c</i> -оксидаза-1	0,76	0,77	1,14
Цитохром- <i>c</i> -оксидаза-2	0,75	1,12	1,41
Цитохром- <i>c</i> -оксидаза-3	0,85	0,91	1,11
АТФаза 6	0,66	1,00	1,33

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.

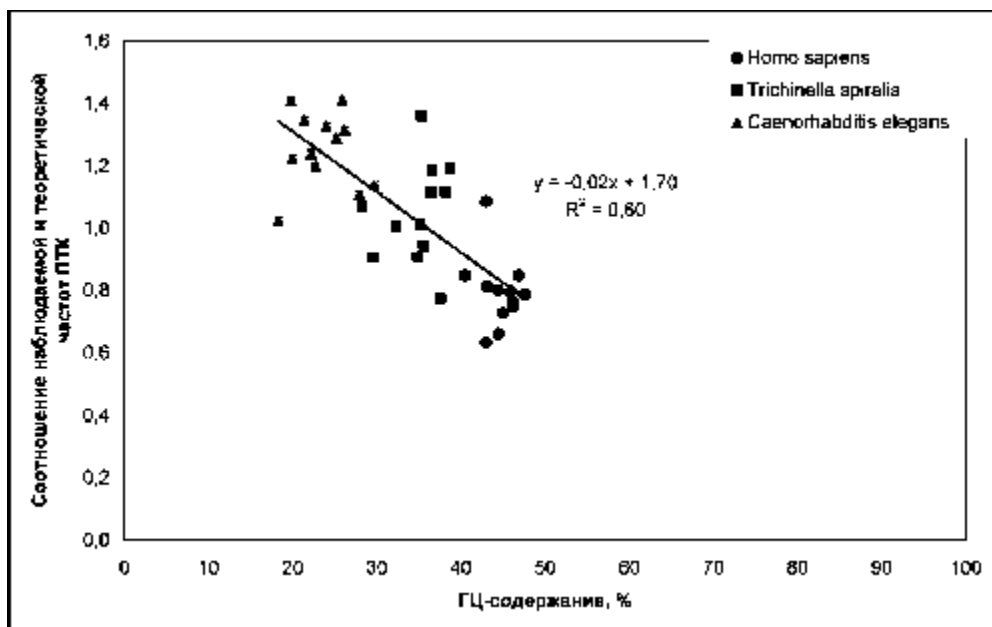


Рис. 2. Зависимость соотношения наблюдаемых и теоретических частот ПТК от ГЦ-содержания изученных мРНК человека, трихинеллы и цианорабдитис

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 ($r = -0,78 \pm 0,107$, $p < 0,001$). Получено уравнение регрессионного анализа для данных показателей: $y = -0,02x + 1,70$, где y – соотношение наблюдаемых и теоретических частот ПТК, а x – ГЦ-содержание. По данному уравнению вычислено, что при ГЦ-содержании, равном 35%, наблюдаемая частота ПТК равна теоретической (их соотношение равно 1). Следовательно, можно предположить, что при содержании гуанина и цитозина в мРНК более 35% частота претерминальных кодонов будет меньше таковой согласно таблицам генетического кода. Данное предположение по-новому раскрывает сущность взаимосвязи между двумя данными показателями и нуждается в тщательной проверке на большем количестве мРНК, транслируемым по разным таблицам генетического кода.

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.

Литература

1. Бредфорд Хилл, А. Основы медицинской статистики. – М., 1958.
2. 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.
3. Бутвиловский, А.В. Стратегия кодирования алкогольдегидрогеназ человека / А.В. Бутвиловский, Е.В. Барковский, В.Э. Бутвиловский // Вестник ВГМУ.-Витебск, 2006, №3. – С. 24 – 29.
4. 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.
5. 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.
6. Справочник терминов молекулярной эволюции и филогенетики. Учебно-методическое пособие / В.Э. Бутвиловский [и др.] // Минск.: БГМУ, 2006. – 40 с.
7. 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.
8. Хрусталева, В.В. Частота использования претерминальных кодонов в матричных РНК бактериальных аденилатциклаз / В.В. Хрусталева, Е.В. Барковский // Здоровоохранение. – 2006. – №2. – С. 17 – 20.
9. Codon catalog usage and the genome hypothesis / R. Grantham [et al]. // Nucl. Acids Res. – 1980. – Vol. 8. – P. 49 – 62.
10. 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.
11. 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.
12. 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.
13. 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.

14. 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.
15. Ikemura, T. Codon usage and tRNA content in unicellular and multicellular organisms / T. Ikemura // Mol. Biol. Evol. – 1985. – V. 2. – P. 13 – 34.
16. 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.
17. 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.
18. 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.
19. Nei, M. Molecular Evolution and Phylogenetics / M. Nei, S. Kumar // Oxford University Press, New York, 2000.
20. Mitochondrial genome variation and the origin of modern humans / Ingman M. [et. al] // Nature. – 2000. – Vol. 408. – P. 708-713.
21. 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.
22. The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum* / R. Okimoto [et. al] // Genetics. – 1992. – Vol. 130 (3). – P. 471-498.