

СКОРОСТЬ МОЛЕКУЛЯРНОЙ ЭВОЛЮЦИИ РАСТВОРИМОЙ
ТИАМИНТРИФОСФАТАЗЫ МЛЕКОПИТАЮЩИХ

А. Ф. МАКАРЧИКОВ

Гродненский государственный аграрный университет, Беларусь;
Институт фармакологии и биохимии НАН Беларуси, Гродно;
e-mail: a_makarchikov@yahoo.com

Тиаминтрифосфат (ТТФ) присутствует в клетках различных типов — от бактерий до млекопитающих. У млекопитающих гидролиз ТТФ осуществляется специфичной растворимой ТТФ-азой, которая экспериментально не обнаружена у представителей других групп организмов. В базах данных NCBI и Ensembl мы нашли информацию о полноразмерных и частичных аминокислотных последовательностях ТТФ-азы 38 видов млекопитающих. Скорость эволюционных замен аминокислот на сайт в год (k_{aa}) в молекуле ТТФ-азы, рассчитанная исходя из имеющихся данных, равна $1,41 \cdot 10^{-9}$, что соответствует элементарному периоду эволюции примерно в 4,4 млн лет. Для разных участков полипептидной цепи этого энзима характерна различная скорость эволюции. Более вариабельны С-концевой фрагмент ($k_{aa} = 3,76 \cdot 10^{-9}$ рассчитана только для видов с длиной полипептида ТТФ-азы 230 ао и слона) и центральная область (остатки 69–141; $k_{aa} = 1,95 \cdot 10^{-9}$), тогда как k_{aa} для N-концевой последовательности (остатки 1–68) составляет $1,03 \cdot 10^{-9}$, а для фрагмента 142–210 — $0,81 \cdot 10^{-9}$.

Ключевые слова: тиаминтрифосфатаза, скорость молекулярной эволюции, млекопитающие.

В объектах живой природы витамин В₁ (тиамин) присутствует в нефосфорилированной форме и в виде трех фосфорных эфиров: ТМФ, ТДФ и ТТФ. В настоящее время установлена лишь функция ТДФ, являющегося кофактором более чем 25 энзимов [1]; биологическая роль ТМФ и ТТФ все еще остается неизвестной.

На протяжении последних 30 лет ТТФ рассматривался в качестве нейроактивной формы тиамина, принимающей участие в функционировании возбудимых мембран [2]. Однако недавние исследования показали, что ТТФ является универсальным компонентом организмов, расположенных на разных ступенках эволюционной лестницы — от бактерий до млекопитающих, вне зависимости от типа их питания, клеточной организации или каких-либо других признаков, совокупность которых используется для систематики и классификации всего многообразия форм жизни на Земле. Этот факт может свидетельствовать о том, что ТТФ представляет собой филогенетически древнюю молекулу, которая была отобрана на начальных этапах эволюции и играет фундаментальную роль в жизнедеятельности клет-

ки как элементарной единицы живой материи [3]. Эксперименты, проведенные на *Escherichia coli* [4] и растительном организме *Arabidopsis thaliana* [5], позволяют предполагать участие ТТФ в молекулярных механизмах адаптации к действию стрессорных факторов [6].

В клетках млекопитающих гидролиз ТТФ катализируется растворимой ТТФ-азой (3.6.1.28) [7, 8]. Учитывая абсолютную специфичность ТТФ-азы к субстрату и ее высокую каталитическую эффективность, можно полагать, что функция этого энзима состоит в быстрой регуляции концентраций ТТФ в зависимости от физиологического состояния клетки. В настоящее время клонированы кДНК ТТФ-азы нескольких видов млекопитающих [9, 10]. Энзимы различных видов отличаются друг от друга количеством аминокислотных остатков в полипептидной цепи и каталитическими свойствами. Было показано, что ТТФ-аза вместе с СуаВ-подобной аденилатциклазой входит в суперсемейство протеинов, содержащих СУТН-домен, общий эволюционный предшественник которых, по-видимому, участвовал в метаболизме нуклеотидов и органических полифосфатов [11].

Сокращения: ао — аминокислотный остаток, ОРС — открытая рамка считывания, пн — пара нуклеотидов, ТМФ — тиаминмонофосфат, ТДФ — тиаминдифосфат, ТТФ — тиаминтрифосфат, ТТФ-аза — тиаминтрифосфатаза, UEP — единичный (элементарный) период эволюции

Специфичная растворимая ТТФ-аза экспрессирована в различных органах и тканях млекопитающих, но экспериментально не обнаружена у представителей других групп живых организмов [5, 9, 12], т. е. в отличие от ТТФ данный фермент является эволюционно молодым. В клетках бактерий, растений, птиц и рыб гидролиз ТТФ протекает под действием менее специфичных фосфатаз [5, 12, 13]. Таким образом, можно говорить о кардинальных различиях системы обмена ТТФ в разных филогенетических линиях. При этом достаточно очевидным кажется вывод о том, что механизмы регуляции метаболизма ТТФ проделали эволюционный путь от контроля скорости его синтеза к контролю скорости гидролиза.

Цель настоящей работы заключалась в исследовании темпов молекулярной эволюции растворимой ТТФ-азы млекопитающих.

Материалы и методы

В работе использовались информационные ресурсы NCBI (<http://www.ncbi.nlm.nih.gov>), геномного браузера Ensembl (release 54; <http://www.ensembl.org>) и базы данных TimeTree (<http://www.timetree.net>).

Поиск, обработку и анализ аминокислотных и нуклеотидных последовательностей проводили с помощью алгоритма BLAST [14] (<http://www.ncbi.nlm.nih.gov/blast>) и программ сервера ExPASy (<http://www.expasy.ch>). Множественное выравнивание последовательностей осуществлялось программой BioEdit, v. 7.0.9.0. Для построения филогенетического дерева ТТФ-азы млекопитающих применяли программы из пакета PHYLIP, v. 3.68 (<http://evolution.genetics.washington.edu/phyliip>).

Скорость эволюционных замен аминокислот на сайт в год (k_{aa}) рассчитывали по тангенсу угла наклона прямой в координатах *время дивергенции* — K_{aa} , учитывая, что $k_{aa} = K_{aa}/(2T)$, где K_{aa} — среднее число аминокислотных замен, приходящееся на пару гомологичных сайтов двух сравниваемых полипептидов; T — число лет, прошедших после эволюционной дивергенции от протеина-предшественника. Величина K_{aa} определяется выражением $K_{aa} = -\ln(1 - p_d - 0,2p_d^2)$, в котором $p_d = d_{aa}/n_{aa}$ — доля аминокислотных различий, n_{aa} — полное число аминокислотных сайтов, по которым сравниваются две гомологичные последовательности, d_{aa} — число сайтов, по которым они отличаются друг от друга [15].

UEP (Unit Evolutionary Period) рассчитывали по уравнению прямой в координатах *время дивергенции* — $p_d \cdot 100$.

Результаты и обсуждение

В базах данных NCBI и Ensembl (release 54) мы нашли информацию о первичной структуре ТТФ-азы 38 видов млекопитающих — человека (*Homo sapiens*), яванского макака (*Macaca fascicularis*), борнеанского орангутана (*Pongo pygmaeus*), суматранского орангутана (*Pongo abelii*), макака-резус (*Macaca mulatta*), свиньи (*Sus scrofa*), собаки (*Canis familiaris*), крысы (*Rattus norvegicus*), мыши (*Mus musculus*), быка (*Bos taurus*), кролика (*Oryctolagus cuniculus*), свинохвостого макака (*Macaca nemestrina*), морской свинки (*Cavia porcellus*), девятипоясного броненосца (*Dasypus novemcinctus*), малого иглистого тенрека (*Echinops telfairi*), европейского ежа (*Erinaceus europaeus*), овцы (*Ovis aries*), кошки (*Felis catus*), лошади (*Equus caballus*), короткохвостого опоссума (*Monodelphis domestica*), шимпанзе (*Pan troglodytes*), туайи белангера (*Tupaia belangeri*), серого мышинного лемура (*Microcebus murinus*), галаго (*Otolemur garnettii*), кенгурового прыгуна Орда (*Dipodomys ordii*), гориллы (*Gorilla gorilla*), скалистого домана (*Procapra capensis*), черноморской афалины (*Tursiops truncatus*), альпаки (*Vicugna pacos*), бурой ночницы (*Myotis lucifugus*), суслика тринадцатиполосого (*Spermophilus tridecemlineatus*), африканского слона (*Loxodonta africana*), американской пищухи (*Ochotona princeps*), кустарникового валлаби (*Macropus eugenii*), калонга (*Pteropus vampyrus*), филиппинского долгопята (*Tarsius syrichta*), ленивца Гоффмана (*Choloepus hoffmanni*) и уткуноса (*Ornithorhynchus anatinus*). На схемах 1, 2, 3 и 4 показаны выровненные с помощью программы ClustalW [16] (<http://www.expasy.ch>) аминокислотные последовательности фермента всех перечисленных выше видов. Идентификационные номера (коды) последовательностей приведены в табл. 1.

Структуры, представленные на схеме 1, являются полноразмерными полипептидами, кодируемыми ОРС. У большинства видов их длина колеблется в пределах 219–230 аа (молекулярная масса ≈ 25 кДа) и соответствует каталитически активному ферменту, экспериментально обнаруженному в различных органах и тканях быка, крысы, мыши, человека и свиньи [5]. В данном отношении удивительным выглядит тот факт, что в сердце собаки выявлена мРНК (EST DN374055), которая кодирует «урезанную» изоформу ТТФ-азы (схема 2). Эта мРНК, являющаяся результатом альтернативного сплайсинга, несет информацию о полипептиде длиной в 117 аа. Пока ничего неизвестно об экспрессии изоформы 2 на уровне

Таблица 1. Аминокислотные последовательности ТТФ-азы 38 видов млекопитающих, зарегистрированные в базах данных NCBI и Ensembl

Вид млекопитающего	Длина, ао	ID в базе данных Protein NCBI ^{a)}	ID в базе данных EST NCBI ^{b)}	ID в Ensembl
<i>Homo sapiens</i>	230 230	NP_077304		ENSP00000288014 ENSP00000384580
<i>Macaca fascicularis</i>	230	Q9BGW0		
<i>Pongo pygmaeus</i>	230			ENSPPYP00000007067
<i>Pongo abelii</i>	230	CAH92146		
<i>Macaca mulatta</i>	230	XP_001109333		ENSMMUP00000005536
<i>Sus scrofa</i>	230	XP_001928758	BP144175	
<i>Canis familiaris</i>	230 117	XP_547736 XP_859333	DN374055	ENSCAFP000000017125
<i>Rattus norvegicus</i>	224	Q8CGV7		ENSRNOP000000038507
<i>Mus musculus</i>	224	Q8JZL3		ENSMUSP000000056026
<i>Bos taurus</i>	219	Q8MKF1		ENSBTAP000000009334
<i>Equus caballus</i>	230	XP_001492080		ENSECAP000000012204
<i>Monodelphis domestica</i>	227	XP_001380088		ENSMODP000000004692
<i>Pan troglodytes</i>	230	XP_522803		ENSPTRP000000010488
<i>Cavia porcellus</i>	220 229		EB369051	ENSCPOP000000000640
<i>Felis catus</i>	230			ENSFCAP000000011188
<i>Tupaia belangeri</i>	229			ENSTBEP000000001326
<i>Microcebus murinus</i>	228			ENSMICP000000009229
<i>Otolemur garnettii</i>	230			ENSOGAP000000007835
<i>Echinops telfairi</i>	225			ENSETEP000000016215
<i>Erinaceus europaeus</i>	229			ENSEEUP000000009249
<i>Loxodonta africana</i>	234			ENSLAFP000000002658
<i>Myotis lucifugus</i>	226			ENSMLUP000000015254
<i>Dasyurus novemcinctus</i>	203			ENSDNOP000000011298
<i>Ochotona princeps</i>	230			ENSOPRP000000015179
<i>Spermophilus tridecemlineatus</i>	182			ENSSTOP000000001982
<i>Oryctolagus cuniculus</i>	84 84 ???		EC623422 EB379139	ENSOCUP000000006669
<i>Ovis aries</i>	132 46		EE814858 EE820268	
<i>Macropus eugenii</i>	179		EX202350	
<i>Macaca nemestrina</i>	156 22		EB526073 EB526074	
<i>Ornithorhynchus anatinus</i>	453	XM_001519599		
<i>Dipodomys ordii</i>	228			ENSDORP000000008334
<i>Gorilla gorilla</i>	230			ENSGGOP000000014419

Таблица 1. (Продолжение)

<i>Procavia capensis</i>	229		ENSPCAP00000004155
<i>Pteropus vampyrus</i>	???		ENSPVAP000000016540
<i>Tursiops truncatus</i>	229		ENSTTRP000000010107
<i>Lama pacos</i>	230		ENSVPAP000000002769
<i>Tarsius syrichta</i>	???		ENSTSYR000000006192
<i>Choloepus hoffmanni</i>	219		ENSCHOP000000002953

^{a)} один из возможных нескольких кодов; ^{b)} нуклеотидные последовательности

протеина, так же как и о том, обладает ли изоэнзим 2 каталитической активностью.

Очевидный интерес представляет ТТФ-аза кролика, состоящая всего из 84 ао (схема 2). Ее первичная структура была выведена нами с помощью трехрамочной трансляции фрагментов мРНК EC623422 (497 пн) и EB379139 (719 пн) из базы данных EST [17] (<http://www.ncbi.nlm.nih.gov>). В обоих случаях нуклеотидные последовательности заканчиваются стоп-кодонами. Поскольку названные EST получены в результате секвенирования двух различных библиотек кДНК — из тройничного нерва (клон nbc45d04) и сердца — ошибки секвенирования, по всей видимости, должны исключаться. Вполне возможно, что эта «урезанная» мРНК также представляет собой продукт альтернативного сплайсинга. С-концевая последовательность ТТФ-азы кролика (остатки 61–84) гомологична участку (остатки 78–101) молекулы изоэнзима 2 собаки. В то же время, данный фрагмент отсутствует у ТТФ-азы других видов млекопитающих. В молекуле ТТФ-азы кролика указанному участку предшествует делеция 17 ао. Если энзим кролика окажется каталитически активным, то можно надеяться, что его исследования способны дать ключ к скорейшему определению того минимального числа остатков (и идентифицировать их), которое необходимо для поддержания структурно-функциональной организации и, следовательно, реализации биологической функции ТТФ-азы, поскольку общее число потенциальных кандидатов при этом сокращается как минимум в 2,7 раза.

Следует обратить внимание на различия представленных на схемах 1 и 3 структур. Так, по информации, находящейся на браузере Ensembl, энзим *C. porcellus* состоит из 220 ао, при этом его С-концевой дипептид (-LG) идентичен одноименным участкам ТТФ-аз большинства других видов (схема 3). Однако при трехрамочной трансляции EST EB369051

(730 пн) из базы данных NCBI нами реконструирована несколько иная непрерывная последовательность протяженностью в 229 ао, не завершающаяся стоп-кодом (схема 1). Поскольку цепь разорвана после ТТ-дуплета, очевидно, что кодируемый ОРС полипептид должен быть длиннее, по крайней мере, на один остаток *Phe* (TTT/C) или *Leu* (TTA/G). Учитывая сказанное, представляется более вероятным, что энзим морской свинки построен из 231 ао, а его структура в базе данных Ensembl содержит неточности вследствие неправильного предсказания экзонно-интронной организации гена. Действительно, при трансляции нуклеотидной последовательности гена ТТФ-азы (ENSCPOG00000000719) легко обнаружить непрерывные последовательности, которые кодируют полипептиды длиной соответственно в 183 ао (1-я рамка) и 48 ао (3-я рамка).

Можно также полагать, что ТТФ-аза *M. lucifugus* построена не из 226, а из 230 ао. При анализе фрагмента геномной ДНК (AAPE01464562) нам не удалось восстановить лишь один остаток полипептидной цепи энзима (остаток 10 на схеме 3 обозначен X). Поскольку при виртуальной трансляции в указанном месте располагается стоп-кодон, этим остатком может быть *Tyr*, *Trp* либо *Cys*.

На схеме 4 приведены частичные аминокислотные последовательности ТТФ-азы, информацию о которых можно найти в базах данных NCBI и Ensembl. Судя по имеющимся данным, энзим *P. vampyrus*, *T. syrichta*, *M. nemestrina*, *C. hoffmanni* и *D. novemcinctus* (в представленной структуре, по-видимому, отсутствует N-концевой участок) должен иметь характерную для большинства видов млекопитающих длину в 228–230 ао. Также кажется вполне обоснованным предположение о том, что ОРС мРНК ТТФ-азы овцы, как и ТТФ-азы быка (схема 1) кодирует протеин, состоящий из 219 ао. Изображенная на схеме 4 структура полипептида *O. anatinus* выведена по фрагменту

M. musculus	1	MAQGLIEVERKFA	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTSELSIMLS	DHMLRQREGSGWEILKCPGVTG	SGPNEYYEVTSEAAI	VAQLFELLGSGBQK	-----	PAGVAAVLGS	LKL	113		
R. norvegicus	1	MAQGLIEVERKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTSELSIMLS	DHMLRQREGSGWEIFKCPGVTG	SGPNEYYEVTSESAI	VTQLFELLGSGRQE	-----	TAGVAAVLGR	LKL	113		
C. porcellus	1	MAQGMIEVERKFT	PGDPTDEERLQELGGSLEHQITFRDI	YDTPELTIMSDHMLRQREGSGWEILKCPGAVG	SGAHTKYVELTSDPAI	VARLQVLEAEBLG	-----	SGGVAAVLSP	LGL	113			
S. scrofa	1	MALGLIEVERKFL	PGPGIEERLQELGLLEHVRVTFRTDY	YDTPELSIMSDHMLRQREGSGWQLKYGAASV	SEPCKYKTEL	TAKSAI	VAQLCEVIRAEVPT	-----	AGGVAAVLDR	LGL	113		
C. familiaris	1	MAQALIEVERKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELSIMRADYWLQREGSGWEILKCPGR	PGVSGPHTYEIV	VELIVESDI	VAQLCEVILGTEVLG	-----	AGGVAAVLGP	LKL	113		
F. catus	1	MAHALIEVERRFT	IPGPGTEKRLQELGGLLEHVRVTFRTDY	YDTPELSIMRADYWLQREGSGWEILKCPGTAG	SGVSGPHTYEIV	VELIVESDI	VAQLCEVILGTEVLG	-----	AGGVAAVLGP	LGL	113		
B. taurus	1	MAQGLIEVERKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELSIMRADYWLQREGSGWEILKCPGAA	SGVSGPHTYEIV	VELIVESDI	VAQLCEVILGTEVLG	-----	AGGVAAVLGP	LGL	113		
M. fascicularis	1	MAQGLIEVERKFL	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELSIMQADHMLRRRDS	GWEILKCPGAA	SGVLGHHTTEYKEL	TAEPTI	VAQLCKVLGADGLG	-----	AGDVAAVLDP	LGL	113	
M. mulatta	1	MAQGLIEVERKFL	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELSIMQADHMLRRRDS	GWEILKCPGAA	SGVLGHHTTEYKEL	TAEPTI	VAQLCKVLGADGLG	-----	AGDVAAVLDP	LGL	113	
P. pygmaeus	1	MAQGLIEVERKFL	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELSIMQADHMLRRRDS	GWEILKCPGAA	SGVLGHHTTEYKEL	TAEPTI	VAQLCKVLGADGLG	-----	AGDVAAMLG	LGL	113	
H. sapiens	1	MAQGLIEVERKFL	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELSIMQADHMLRRRDS	GWEILKCPGAA	SGVLGHHTTEYKEL	TAEPTI	VAQLCKVLGADGLG	-----	AGDVAAVLGP	LGL	113	
E. caballus	1	MAQGLIEVERKFL	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELSIMQADHMLRRRDS	GWEILKCPGAA	SGVLGHHTTEYKEL	TAEPTI	VAQLCKVLGADGLG	-----	AGHVAAVLGP	LGL	113	
P. troglodytes	1	MAQGLIEVERKFL	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTSELSIMQADHMLRRRDS	GWEILKCPGAA	SGVLGHHTTEYKEL	TAEPTI	VAQLCKVLGADGLG	-----	AGDVAAVLGP	LGL	113	
T. belangeri	1	MAQGLIEVERKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELSIMQADHMLRRRDS	GWEILKCPGAA	SGVSGLHTEYKEL	TAEPTI	VAQLCKVLGADGLG	-----	ACGVAAVLGP	LGL	113	
M. murinus	1	MAQGLIEVELKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELIRIMRADYWLRRQ	-DSGWEILKCPGAA	SGVLGHHTQYVEL	TAEPTI	VAQLCEVILETEGFG	-----	DGGVAAVLGP	MGL	112	
O. garnettii	1	MAQGLIEVELKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELIMRADYWLRRQ	DRGWELIPC	GAAVSGPHTQYKEL	TAEPTI	VAQSVQILETEGLG	-----	AGGVAAVLS	LGL	113	
M. domestica	1	MTHPWIEVEKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELTIMRADHMLRRENG	WMLKCPGAGSGPHTKYVEL	TAEAAI	AARLCEQLGSGELG	-----	IETVSDMLG	QLKL	113		
E. telfairi	1	MAPGSIEVERKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELIMRADHMLRHRREG	WMLKFLGAA	SGSAHTYREL	TAEPTI	VAQLCEVILGTEGQKQ	-----	TGCVAAAMLG	LGL	113	
E. europaeus	1	MAQGLIEVERKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELSIMRADHMLRRRDS	GWEILKCPGAG	PGVSGPHTKYVEL	TAEPTI	VAQLCEVILGTEGQKQ	-----	AEAGVAAVLGP	LGL	113	
L. africana	1	MAQGLIEVERKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTSELSIMKANHMLRQREGSGWEILKCPG	VAGSGPHTYEIV	REL	TAEPTI	VAQLCEVILGAEGLGAGG	VAALLG	LGL	118		
P. abelii	1	MAQGLIEVERKFL	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELSIMQADHMLRRRDS	GWEILKCPGAA	SGVLGHHTTEYKEL	TAEPTI	VAQLCKVLGADGLG	-----	AGDVAAMLG	LGL	113	
O. princeps	1	MAQGLIEVERKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELSIMQADHMLRRRDS	GWEILKCPGAA	SGVLGHHTTEYKEL	TAEPTI	VAQLCKVLGADGLG	-----	AECSVAALG	LGL	112	
D. ordii	1	MAQGLIEVERKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELSIMQADHMLRRRDS	GWEILKCPGAG	PGVSGPHTYEIV	ELTSEPTI	VAQLRLLGABEGL	-----	AGSVAALG	LGL	113	
G. gorilla	1	MAQGLIEVERKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELSIMQADHMLRRRDS	GWEILKCPGAA	SGVLGHHTTEYKEL	TAEPTI	VAQLCKVLGADGLG	-----	AGDVAAVLGP	LGL	113	
T. truncatus	1	MAQGLIEVERKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELSIMRADYWLRRQREGSGWEILKCPGAA	SVSGPHTKYVEL	TAEPTI	VAQLCEVIRAEVPT	-----	AGGVAAVLGP	LGL	113		
M. eugenii	1	MTHPWIEVEKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELTIMRADHMLRHRREGSGWEILKCPGAA	SVSGPHTKYVEL	TAEPTI	VAQLCEVIRAEVPT	-----	AGGVAAVLGP	LGL	113		
L. pacos	1	MAQGLIEVERKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELTIMRADHMLRRRDS	GWEILKCPGAG	SVSGPHTKYVEL	TAEPTI	VAQLCEVILGTEGQKQ	-----	TESVSEMLGR	LKL	113	
P. capensis	1	MAQGLIEVERKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELSIMRADHMLRRRDS	GWEILKCPGAG	SVSGPHTYEIV	REL	TAEPTI	VAQLCEVILGTEGQKQ	-----	AGGVAALG	LGL	113
M. lucifugus	1	MAHGLIEVERKFT	PGDPTDEERLQELGATLEHVRVTFRTDY	YDTPELSIMRADHMLRRRDS	GWEILKCPGAG	SVSGPHTYEIV	REL	TAEPTI	VAQLCEVILGTEGQKQ	-----	TGDLAAQGL	LGL	113
M. musculus	114	QEVASFITR	SSWKALVLSG-AHGQEPQL	TIDLSADFG	YAVGEV	EAHVPAAL	EKI	ITVSSMLG	VPAQEAPAKIMVYLQ	FRPDYQRLLEAA	-SSGE	224	
R. norvegicus	114	QEVASFITR	SSWKALVLSG-AHEES	LLITVDLSDT	FGAVGEV	EAHVPAAL	EKI	ISVSSMLG	VPAQEAPAKIMVYLQ	FRPDYQRLLEAA	-SSGE	224	
C. porcellus	114	QEVASFVTKR	SSWKALVLSG-ADEEP	PLLRVDLDTAD	FGAVGEV	EAHVPAAL	EKI	QSLSSMLG	VPAQEAPAKIMVYLQ	FRPDYQRLLEAA	-SSGE	231	
S. scrofa	114	QEVASFVTKR	SSWKALVLSG-ADEEP	PLLRVDLDTAD	FGAVGEV	EAHVPAAL	EKI	HSVSSMLG	VPAQEAPAKIMVYLQ	FRPDYQRLLEAA	-SSGE	230	
C. familiaris	114	QEVASFVTKR	SSWKALVLSG-ADEEP	PLLRVDLDTAD	FGAVGEV	EAHVPAAL	EKI	HSVSSMLG	VPAQEAPAKIMVYLQ	FRPDYQRLLEAA	-SSGE	230	
F. catus	114	QEVASFVTKR	SSWKALVLSG-ADEEP	PLLRVDLDTAD	FGAVGEV	EAHVPAAL	EKI	HSVSSMLG	VPAQEAPAKIMVYLQ	FRPDYQRLLEAA	-SSGE	230	
C. familiaris	114	QEVASFVTKR	SSWKALVLSG-ADEEP	PLLRVDLDTAD	FGAVGEV	EAHVPAAL	EKI	HSVSSMLG	VPAQEAPAKIMVYLQ	FRPDYQRLLEAA	-SSGE	230	
B. taurus	114	QEVASFVTKR	SSWKALVLSG-ADEEP	PLLRVDLDTAD	FGAVGEV	EAHVPAAL	EKI	HSVSSMLG	VPAQEAPAKIMVYLQ	FRPDYQRLLEAA	-SSGE	230	
M. fascicularis	114	QEVASFVTKR	SSWKALVLSG-ADEEP	PLLRVDLDTAD	FGAVGEV	EAHVPAAL	EKI	HSVSSMLG	VPAQEAPAKIMVYLQ	FRPDYQRLLEAA	-SSGE	230	
M. mulatta	114	QEVASFVTKR	SSWKALVLSG-ADEEP	PLLRVDLDTAD	FGAVGEV	EAHVPAAL	EKI	HSVSSMLG	VPAQEAPAKIMVYLQ	FRPDYQRLLEAA	-SSGE	230	
P. pygmaeus	114	QEVASFVTKR	SSWKALVLSG-ADEEP	PLLRVDLDTAD	FGAVGEV	EAHVPAAL	EKI	HSVSSMLG	VPAQEAPAKIMVYLQ	FRPDYQRLLEAA	-SSGE	230	
H. sapiens	114	QEVASFVTKR	SSWKALVLSG-ADEEP	PLLRVDLDTAD	FGAVGEV	EAHVPAAL	EKI	HSVSSMLG	VPAQEAPAKIMVYLQ	FRPDYQRLLEAA	-SSGE	230	
E. caballus	114	QEVASFVTKR	SSWKALVLSG-ADEEP	PLLRVDLDTAD	FGAVGEV	EAHVPAAL	EKI	HSVSSMLG	VPAQEAPAKIMVYLQ	FRPDYQRLLEAA	-SSGE	230	
P. troglodytes	114	QEVASFVTKR	SSWKALVLSG-ADEEP	PLLRVDLDTAD	FGAVGEV	EAHVPAAL	EKI	HSVSSMLG	VPAQEAPAKIMVYLQ	FRPDYQRLLEAA	-SSGE	230	
T. belangeri	114	QEVASFVTKR	SSWKALVLSG-ADEEP	PLLRVDLDTAD	FGAVGEV	EAHVPAAL	EKI	HSVSSMLG	VPAQEAPAKIMVYLQ	FRPDYQRLLEAA	-SSGE	230	
M. murinus	113	QEVASFVTKR	SSWKALVLSG-ADEE	-RLRRVDLDTAD	FGAVGEV	EAHVPAAL	EKI	HSVSSMLG	VPAQEAPAKIMVYLQ	FRPDYQRLLEAA	-SSGE	228	

Схема 1. Аминокислотные последовательности ТТФ-азы 29 видов млекопитающих, выровненные с помощью программы CLUSTALW с ручной коррекцией. Последовательность валлаби выведена по фрагменту геномной ДНК ABQ001014192 (wgs database, <http://www.ncbi.nlm.nih.gov>), морской свинки – путем виртуальной трансляции EST EB369051 (<http://www.ncbi.nlm.nih.gov>) и гена ENSCPOG00000000719 (<http://www.ensembl.org>)

<i>O. garnettii</i>	114	QEVASFVTVNRSWKLELLG-GDEEKSPFLRVLDLDTADFGYAVGEVEVLVHEEAEPAALEKIHRLSNMLGVQVQETVPKLIIVYLQRFHPQHYQRLLLEVNSRER-PQGTKHPDSSLG-	230
<i>M. domestica</i>	114	QEVASYVTDRSVWKALISG-AGDEEEPLTVLDLDTADFGYAVGEVEALVREAAEVPAALEKIHSLSLLGVSEQEQTPGKLIIVYLKRFREDDYQRLLLEKTKGNDEQ-GSEATEN----	227
<i>E. telfairi</i>	114	EQVASFVTVTRRSWKLMFLG-A-EQESPLRVLDLDTADFGYAVGEVEALVHEEAEPAALEKIHSLSLLGVQALEKPAPAKLIIVYQRFQPHYQILLKVKRARER-PPGTQQPD-----	225
<i>E. europaeus</i>	114	QEVASFVTVTRRSWKVLPVPH-VEESDPLRVLDLDTADFGYAVGEVEALVHKEAQVPAALEKINSLSLLGVLVQEKAPAKLIIVYLQRFREDDYIRLLEVNDFREMP-PQGTQQP-GSLY-	229
<i>L. africana</i>	119	QEVAKFVTVKRSWKLVLLG-A-EEELPLKVDLDTADFGYAVGEVEALVHEEAEPVPTALGKIHSLSLLGVPAQEKPPAKLIIVYLQRFREDDYQILLLESGNSFREK-LQETKHPDSSLG-	234
<i>P. abelii</i>	114	QEVASFVTVKRSWKLVLLG-ADEEFPQLRVLDLDTADFGYAVGEVEALVHEEAEPVAVEKIHRLSMGLVPAQETAAPAKLIIVYLQRFREDDYQRLLLEVNSRER-PQETEDPDNCLG-	230
<i>O. princeps</i>	113	QETASFVTVKRSWKLVLPGRAGEEPLRVLDLDTADFGYAVGEVEALVREAAEVPAALEKIHSLSLLGVPAQEKAPAKLIIVYLQRFREDDYQILLLEAANLGRPGCTERPD-LV-	230
<i>D. ordii</i>	114	QEVASFVTVNRSWKLMFPFG-AREK-EPLLRVLDLDTADFGYAVGEVEALVHEEAEPVPALEKIHSLSMGLVPTQEKAPAKLIIVYLQRFREDDYQRLLLE-NTFRE--PRGAKHSDRSLDY	228
<i>G. gorilla</i>	114	QEVASFVTVKRSWKLVLLG-ADEEFPQLRVLDLDTADFGYAVGEVEALVHKEAEVPTALEKIHRLSMGLVPAQETAAPAKLIIVYLQRFREDDYQRLLLEVNSRER-PQETEDPDHCLG-	230
<i>T. truncatus</i>	114	QEVASFVTVKRSWKLMFPG-ADEEFPPLRVLDLDTADFGYAVGEVEALVQEEAEVPALEKINSLSMGLVLAQETAAPAKLIIVYLQRFREDDYQRLLLEVYSTEK-P-GTDFDVTILG-	229
<i>M. eugeni</i>	114	QEVASFVTVTRSVWKALISG-AGDEEPLTVLDLDTADFGYAVGEVEALVREAAEVPAALEKIHSLSLLGVSSQEQTPGKLIIVYLKRFREDDYQRLLLEVATGKNDQOQSAAMEN-----	228
<i>L. pacos</i>	114	QEVASFVTVKRSWKLVLLG-ADEEFPPLRVLDLDTADFGYAVGEVEALVHKEAEVPTALKKIHSLSMGLVPAQETAAPAKLIIVYLQRFREDDYQRLLLEVYGSREK-PQGTQKDFDMRLG-	230
<i>P. capensis</i>	114	QEVASFVTVKRSWKLVLLG-A-DEEFPPLRVLDLDTADFGYAVGEVEALVHEEAEPVPALEKINSLSLLGVPAQEKPPAKLIIVYLQRFREDDYQILLLEANSRER-PQGTKYPDGSLG-	229
<i>M. lucifugus</i>	110	QEVASFVTVNRSWKLVLSR-ADREEPPLRVLDLDTADFGYAVGEVEALVREAAEVPAALEKINKLGSMLGVLAQETAAPAKLIIVYLQRFREDDYQRLLLEVHSSREK-PGPTNDADSSLG-	226

Схема 1. (Продолжение)

NW_001740506 геномной ДНК и представляет собой N-концевую часть молекулы протяженностью в 69 ао, а ТТФ-аза *S. tridecemlineatus* — это участок молекулы, гомологичный кодируемому первым экзонном ТТФ-азы человека.

В геноме человека единственный локус ТТФ-азы — ТНТРА (ID 79178) — расположен в большом плече 14-й хромосомы (позиция 14q11.2) по соседству с локусами APIG2 и ZNF409 (EntrezGene [18], <http://www.ncbi.nlm.nih.gov>). Возможны по крайней мере два пути сплайсинга первичного транскрипта. Один из них связан с удалением центрального интрона (транскрипт ENST00000288014, схема 5), причем размер участка полипептидной цепи, кодируемого первым экзонном, равен 182 ао, а участка, зашифрованного во втором экзоне — 48 ао. При альтернативном сплайсинге распознается еще один интрон в 5'UTR мРНК (транскрипт ENST00000404535; <http://www.ensembl.org>). Таким образом, эти альтернативные способы экспрессии гена не затрагивают аминокислотной последовательности протеина: в любом случае зрелая мРНК несет информацию о полипептиде длиной в 230 ао. Сведения об экзонно-интронной организации генов ТТФ-азы 32-х видов млекопитающих представлены в табл. 2.

Из табл. 2 видно, что хотя в гене ТТФ-азы быка содержится 5 экзонов, лишь 2 из них (экзоны 3 и 4) несут информацию о нуклеотидной последовательности ОРС транскрипта, при этом размер участка полипептидной цепи, кодируемого экзонном 3, как и в случае ТТФ-азы человека, равен 182 ао. Аналогичным образом только в 2-х из 3-х экзонов зашифрована структура энзима свиньи. Наличие 3-х и более экзонов предсказано также для генов ТТФ-азы *O. anatinus*, *T. belangeri*, *O. princeps*, *C. porcellus*, *D. novemcinctus*, *E. telfairi*, *E. europaeus*, *M. murinus*, *M. lucifugus*, *O. garnettii*, *O. anatinus*, *D. ordii*, *P. capensis*, *P. vampyrus*, *T. truncatus* и *O. cuniculus*, причем у этих видов все экзоны содержат информацию о структуре протеина.

Следует отметить, что предсказания экзонно-интронной организации генов ТТФ-азы, по-видимому, не всегда надежны. Так, например, в гене (ENSMICG00000010142) *M. murinus* предсказано наличие 4-х экзонов: первый из них кодирует N-концевой участок полипептидной цепи длиной 12 ао (<http://www.ensembl.org>). В то же время при трансляции фрагмента геномной ДНК ABDC01231528 из базы данных wgs NCBI в направлении 5'→3' во второй рамке нами обнаружена непрерывная аминокислотная последовательность из 180 ао, гомологичная N-концевому участку ТТФ-азы

<i>C. familiaris</i>	1	MAQALIEVER	KFVPGPDTE	RLQELGATLE	HRVTFRDSYY	DPELSLMRA	50
<i>O. cuniculus</i>	1	MAQGLIEVER	KFIPGPGTEQ	RLQELGATLE	RRVIFRDITY	DPELSLMRA	50
<i>C. familiaris</i>	51	DYWLRQREGS	GWELKCPGRP	GVSGPHTECA	STGESTCQAD	CVPTALPASG	100
<i>O. cuniculus</i>	51	DHWLRQREDR	-----	-----GCA	SPGDSRQAD	RVPPALPAPG	83
<i>C. familiaris</i>	101	LSAPARSTQP	QQEATGD				117
<i>O. cuniculus</i>	84	L-----	-----				84

Схема 2. Множественное выравнивание аминокислотных последовательностей ТТФ-азы кролика и собаки (изоформа 2)

человека, кодируемому первым экзоном. Введенная по фрагментам ДНК ABDC01231528 и ABDC01152870 последовательность ТТФ-азы *M. murinus* отличается от зарегистрированной в Ensembl четырьмя позициями (остатки 13 (Ile) и 214-216 (Ser, Gln, Arg) на схеме 3). Сходным образом фрагмент геномной ДНК (AAKN01731967, 3711 пн) *C. porcellus* несет информацию о непрерывной из 183 ао N-концевой последовательности ТТФ-азы, хотя исходя из данных браузера Ensembl эта часть молекулы энзима должна кодироваться двумя экзона-

ми. Можно также показать, что информация о структуре ТТФ-азы *D. ordii* зашифрована в двух непрерывных участках геномной ДНК (фрагменты ABRO01236453 и ABRO01236454; wgs database, <http://www.ncbi.nlm.nih.gov>). К аналогичному выводу приводит и виртуальная трансляция гена ENSDORG00000008883 (<http://www.ensembl.org>), т. е. синтезируемый на рибосомах полипептид состоит не из 228 (схема 1), а из 229 а. о. (схема 3). Следовательно, нельзя исключать, что полипептидная цепь энзима у всех перечисленных в табл. 2 млеко-

<i>C. porcellus</i>	1	MAQGMIEVER	KFVPGPGTEE	RLQELGGSLE	HQITFRDIYY	DPELTLMQS	50
<i>D. ordii</i>	1	MAQGLIEVER	KFVPGPDTEK	RLQELGGTLV	HRFTFRDITY	DTTELSLMRS	50
<i>M. murinus</i>	1	MAQGLIEVEL	KFIPGPGTEE	RLQELGSTLE	HRVTFRDITY	DPELRLMRA	50
<i>M. lucifugus</i>	1	MAQGLIEMEX	ELISGPETEE	RLQELGATLG	YWGTFRDSYY	VIPELSLMRA	50
<i>C. porcellus</i>	51	DHWLRQREGS	GWELKCPGAV	GVSGAHTKYV	ELTSDPAIVA	RLCQVLEAEE	100
<i>D. ordii</i>	51	DHWLRQREGS	GWELKCPGKA	GIPGPHTEYI	ELTSEPEIVA	QLRELLGAEE	100
<i>M. murinus</i>	51	DYWLRQRD-S	GWELKCPGAA	GVLGPHTQYV	ELTAEPAlVA	QLCEVLETEG	99
<i>M. lucifugus</i>	51	DHCLRQRENS	EWELKCPGTA	GVSTPHTEYL	ELTAEPAlVA	QLYKVLGAEV	100
<i>C. porcellus</i>	101	LGSGGVA AVL	SPLGLQE VAS	FVTKR-----	-----ED-EEP	LLKVDLDTAD	140
<i>D. ordii</i>	101	LGAGSVA AVL	GSLGLQE VAS	FVTNRS AWKL	MFPGA E-KEP	LLRVDLDTAD	149
<i>M. murinus</i>	100	PGDGGVA AVL	GPMGLQE VAS	FVTNRS AWKL	VLLGAD-EER	PLRVDLDTAD	148
<i>M. lucifugus</i>	101	PGAEGVA AVL	GPLGLQE VAS	FVTNRS AWKL	VLSRADREEP	PLRVDLDTAD	150
<i>C. porcellus</i>	141	FGYTVGEVEA	LVREEAEVPS	ALEKIQSLSK	MLGVPAHEKA	PAKLIVYLQR	190
<i>D. ordii</i>	150	FGYAVGEVEA	LVHEVAEVPA	ALEKILSLSS	MLGVPTQEKA	PAKLIVYLQR	199
<i>M. murinus</i>	149	FGYTVGEVEV	LVHEEA EVPT	ALEKIHRLSN	MLGVQVQETV	PAKLTVYLQR	198
<i>M. lucifugus</i>	151	FGYAVGEIEA	LVREEAEVPA	ALEKINKLGS	MLGVLAQERA	PAKLIVYLQR	200
<i>C. porcellus</i>	191	FRPQDYQRL	LQAHTTG EK PQ	GTEHPDNHLG	-	220	
<i>D. ordii</i>	200	FRPEDYQRL	LES-TLSEKPR	GAKHS DRSLD	Y	229	
<i>M. murinus</i>	199	FRPQVYQRL	I EVNNS SQRPQ	GTKHPDSSLG	-	228	
<i>M. lucifugus</i>	201	FRPQDYQRL	L EVHSSREKPE	GTNDADSSLG	-	230	

Схема 3. Множественное выравнивание аминокислотных последовательностей ТТФ-азы морской свинки (ENSCPOP00000000640), кенгурового прыгуна орда (виртуальная трансляция фрагментов геномной ДНК ABRO01236453, ABRO01236454 и гена ENSDORG00000008883), серого мышинного лемура (фрагменты геномной ДНК ABDC01231528 и ABDC01152870) и бурой ночницы (фрагмент геномной ДНК AARE01464562)

<i>O. aeris</i>	MAQGLIEVER	KFVPGPGTEE	RLQELGGTLE	HRVTFRDSYY	DTPELSLMRA
<i>S. tridecemliatus</i>	MAQGLIEVER	KFIPGPGIEE	RLQELGGILE	HRITFRDITY	DTPELSLMRS
<i>D. novemcinctus</i>	-----	-----	-----GGILE	HQVTFRDSYY	DT-ELSLMRA
<i>O. anatinus</i>	MASGRIEVEC	KFVPGPGTEE	RLRALGGTLE	HRTSFSDSY	DTAELTLMRA
<i>P. vampyrus</i>	MAQGLIEVER	KFIPGPGTEE	RLLELGGTLE	HQVTFRDSYY	DTPELSLMRT
<i>T. syrichta</i>	MAPGLIELER	KFIPGPGIEE	RLQELGGTLE	HQVTFRDITY	DNSELGLMRA
<i>M. nemestrina</i>	MAQGLIEVER	KFLPGPGTEE	RLQKLGGTLE	HRVTFRDITY	DTPELSLMQA
<i>C. hoffmanni</i>	MAHGLIEVER	KFIPGPGTEE	RLRELGGTLE	HRVTFRDSYY	DTPKLSLMRA
<i>O. aeris</i>	DYWLRQREGS	GWELKCPGAA	SVSGPHTEYT	ELTAEPSIVV	QLCEVLGAVV
<i>S. tridecemliatus</i>	NHWLRHREGS	GWELKCPGAS	GISGPHTEYK	ELTSEPAVVA	QLCEVLGAKG
<i>D. novemcinctus</i>	DYWLRQREGS	GWELKCPRAT	LVSGPHTEYM	ELTVERAVVA	QVCEVLGTED
<i>O. anatinus</i>	DHWLRRREGS	GWELKCPGT-	-----	-----	-----
<i>P. vampyrus</i>	DHWLRQRESS	GWELKCP-AA	-VSLHTEYL	ELTAEPAlVA	KLCEVLGADV
<i>T. syrichta</i>	DYWLRQRENS	GWELKCPGAA	GVLGPHTKYI	ELTAEPAlVA	QLCEVLGAEA
<i>M. nemestrina</i>	DHWLRRREDS	GWELKCPGAA	GVLGHHTYK	ELTAEPAlVA	QLCKVLGADG
<i>C. hoffmanni</i>	DYWLRQREGS	GWELKCPGTA	GVSGSHTEYK	ELTVECAIVA	QLCEVLDTFV
<i>O. aeris</i>	PGAGVVA AVL	GPLGLQLVAS	FVTKRSAWKL	VLK-----	-----
<i>S. tridecemliatus</i>	LGDGSA AVL	GPLGLQEVAS	FVTKRSSWKL	VFPGVNEEEP	MLKVDLDTAD
<i>D. novemcinctus</i>	LWAGGVA AVL	RPLGLEEVAS	FVTTRSSWKL	VLFGP-EEEL	LLRIDLDTAD
<i>O. anatinus</i>	-----	-----	-----	-----	-----
<i>P. vampyrus</i>	LGAGVVA AVL	SPLGLQE---	-----	-----	-----
<i>T. syrichta</i>	LGAGGVA AVL	GSLGLQEVAS	FVTKRSAWKL	VLLGADQEEQ	PLRV-----
<i>M. nemestrina</i>	LGAGDVA AVL	CPLGLQEVAS	FVTKRSAWKL	VLLGTDEEEP	QLRVDLDTAD
<i>C. hoffmanni</i>	LGTGGVA AVL	GPLGLQEVAS	FVTKRSAWKL	VLLGA-EEEL	PLRVDLDTAD
<i>O. aeris</i>	-----	-----	----IHHLSS	LLGVLAQEKA	PAKLIVYLQR
<i>S. tridecemliatus</i>	FGYAVGEVEA	LVQKEAEVPT	ALKKINSLS	ML-----	-----
<i>D. novemcinctus</i>	FGYAVGEVEA	LVHDEAEVPT	VLEKIHSLSS	LLGVSEQEKA	PSKLIVYLQR
<i>O. anatinus</i>	-----	-----	-----	-----	-----
<i>P. vampyrus</i>	-----	-----	-----	--GVLAQEVA	PAKLIVYLQR
<i>T. syrichta</i>	-----	-----	-----	--GVPAQEKA	PAKLIMYIQR
<i>M. nemestrina</i>	FGYAVG----	-----	-----	-----	-----
<i>C. hoffmanni</i>	FGYAVGEVEA	LVHEEAEVPT	ALE-----	--GVSEQEKA	PAKLIVYLQH
<i>O. aeris</i>	FRPDYQRL	EVYGSKEQP-	-----	----	
<i>S. tridecemliatus</i>	-----	-----	-----	----	
<i>D. novemcinctus</i>	FRPDYQHLL	EANHSEERPR	GTKHSDSSLS	----	
<i>O. anatinus</i>	-----	-----	-----	----	
<i>P. vampyrus</i>	FRPDYQRL	DVHSSREKPQ	GTKDADSSLD	----	
<i>T. syrichta</i>	FRPDYQRL	EVNGSREKPQ	GTKHPDGILD	----	
<i>M. nemestrina</i>	-----LL	EVNSSKQRPQ	GTEDRDNCLG	----	
<i>C. hoffmanni</i>	FRPDYQHLL	EVNSSKERPQ	GTQHSDCSLG	----	

Схема 4. Множественное выравнивание частичных аминокислотных последовательностей ТТФ-азы 8 видов млекопитающих



Таблица 2. Экзонно-интронная структура генов ТТФ-азы млекопитающих

Вид млекопитающего	Число экзонов	Число участков полипептида, кодируемых разными экзонами	ID в базе данных EntrezGene NCBI	ID в Ensembl
<i>Homo sapiens</i>	2, 3	2	79178	ENSG00000157306
<i>Pongo pygmaeus</i>	2	2		ENSPPYG00000006225
<i>Macaca mulatta</i>	2, 3	2	714403	ENSMMUG00000004142
<i>Sus scrofa</i>	3	2	396635	
<i>Canis familiaris</i>	2, 3	2	490614	ENSCAFG00000011651
<i>Rattus norvegicus</i>	2	2	305889	ENSRNOG00000018105
<i>Mus musculus</i>	2	2	105663	ENSMUSG00000045691
<i>Bos Taurus</i>	5	2	282090	ENSBTAG00000007097
<i>Equus caballus</i>	2	2	100052214	ENSECAG00000014559
<i>Monodelphis domestica</i>	2	2	100030627	ENSMODG00000003828
<i>Pan troglodytes</i>	2	2		ENSPTRG00000006180
<i>Cavia porcellus</i>	3	3		ENSCPOG00000000719
<i>Felis catus</i>	2	2		ENSFCAG00000012059
<i>Tupaia belangeri</i>	3	3		ENSTBEG00000001547
<i>Microcebus murinus</i>	4	4		ENSMICG00000010142
<i>Otolemur garnettii</i>	5	5		ENSOGAG00000008737
<i>Echinops telfairi</i>	3	3		ENSETEG00000019964
<i>Erinaceus europaeus</i>	3	3		ENSEEUG00000010169
<i>Loxodonta africana</i>	2	2		ENSLAFG00000003200
<i>Myotis lucifugus</i>	6	6		ENSMUG00000016741
<i>Dasyurus novemcinctus</i>	3	3		ENSDNOG00000014589
<i>Ochotona princeps</i>	6	6		ENSOPRG00000016630
<i>Ornithorhynchus anatinus</i>	6	6	100090555	
<i>Dipodomys ordii</i>	4	4		ENSDORG00000008883
<i>Gorilla gorilla</i>	2	2		ENSGGOG00000014783
<i>Procavia capensis</i>	8	8		ENSPCAG00000004443
<i>Pteropus vampyrus</i>	4	4		ENSPVAG00000017536
<i>Tursiops truncatus</i>	4	4		ENSTTRT00000010660
<i>Lama pacos</i>	2	2		ENSVPAG00000002998
<i>Tarsius syrichta</i>	2	2		ENSTSYG00000006757
<i>Oryctolagus cuniculus</i>	6	6		ENSOCUG00000007718
<i>Choloepus hoffmanni</i>	2	2		ENSCHOG00000003338

питающих кодируется только двумя экзонами, а гены ТТФ-азы в большем числе случаев, чем это вытекает из приведенных данных, имеют одинаковое строение и состоят из 2-х экзонов и интрона.

Как уже упоминалось, растворимая ТТФ-аза относится к обширному суперсемейству протеинов, содержащих СУТН-домен и обнаруженных у представителей всех трех суперцарств живых организмов [11]. В настоя-

щее время число известных аминокислотных последовательностей, архитектура которых основана на единственном СУТН-домене, приближается к 600 (Pfam 23, release 1.7[19], <http://pfam.sanger.ac.uk>). В случае ТТФ-азы в формировании пространственной структуры домена задействовано 84–87% от общего количества аминокислотных остатков (позиции 5–198 в молекуле энзима человека).

Хотя обладающий ТТФ-азной активностью энзим с молекулярной массой ≈ 25 кДа экспериментально найден только у млекопитающих [5, 9, 12, 13], существование генов-ортологов ТТФ-азы предсказано также для нескольких видов костных рыб (*Salmo salar* (LOC100194700), *Danio rerio* (GeneID 100048920; ENSDARG00000063049), *Gasterosteus aculeatus* (ENSACG00000002180), *Oryzias latipes* (ENSORLG00000006150), *Takifugu rubripes* (ENSTRUG00000006456), *Tetraodon nigroviridis* (GSTENG000009631001), земноводного *Xenopus tropicalis* (ENSXETG00000014559) и рептилии *Anolis carolinensis*. Кроме того, в базе данных nr NCBI мы обнаружили ряд гипотетических протеинов (AAI33924.1 (*D. rerio*), CAF93810.1 (*T. nigroviridis*), NP_001133227.1 (*S. salar*), AAI35218.1 (*X. tropicalis*), AAI06327.1 (*Xenopus laevis*), XP001627536 (*Nematostella vectensis*), XP001192841.1 (*Strongylocentrotus purpuratus*), соответствующие структуре ТТФ-азы человека ($E = 10^{-18}$ – 10^{-34}). Тем не менее, ввиду отсутствия сведений об энзиматической активности указанных протеинов и их не слишком высокой степени идентичности энзиму человека, на наш взгляд, еще рано причислять эти протеины к функциональным гомологам специфичной растворимой ТТФ-азы. Действительно, например последовательности протеина AAI33924.1 (228 ао; GeneID 100048920) *D. rerio* и ТТФ-азы человека совпадают по аминокислотным позициям на 34,6%, а аденилатциклазы *Pyrobaculum aerophilum* (176 ао; NP558881.1) и ТТФ-азы — на 31,8%. Аденилатциклазам (4.6.1.1) однако не присуща фосфатазная активность (BRENDA [20], <http://www.brenda-enzymes.info>). Если все же допустить, что перечисленные выше гипотетические протеины наделены ТТФ-азной активностью, то появление растворимой ТТФ-азы по-видимому должно относиться не ранее, чем к периоду дивергенции Animalia, поскольку гомологичные ТТФ-азе последовательности не выявлены у прокариот, простейших, грибов и растений (по состоянию на 15.05.2009 секвенированы геномы соответственно 571, 32, 77 и 12 видов). С другой стороны, гомологи ТТФ-азы пока не обнаружены у таких классов

животных, как насекомые (21 вид) и птицы (2 вида), а также у червей (11 видов), что может указывать на более позднее происхождение этого энзима.

Располагая информацией об аминокислотных последовательностях ТТФ-азы нескольких видов млекопитающих, мы можем оценить скорость молекулярной эволюции энзима в пределах этой таксономической группы. Исходные данные для таких расчетов представлены на рис. 1 и в табл. 3. Матрицу попарных расстояний рассчитывали с помощью программы Protdist из пакета PHYLIP; времена дивергенции взяты из электронной базы данных TimeTree [21] и работ [22–26]. Здесь следует отметить, что на сегодня наименее надежной является оценка момента расхождения Glires (отряды Rodentia и Lagomorpha)/Euarchonta (отряды Primates и Scandentia). Например, в статье Douzery & Huchon [27] приведена цифра 68,4 млн лет, основанная на анализе аминокислотных последовательностей 20 ядерных протеинов у 6 видов млекопитающих. По другим сведениям, которые были получены на более массивных выборках (52–360 протеинов), указанное событие может относиться к периоду 90,8–120 млн лет [23, 24, 28]. Хотя средневзвешенное время расхождения групп Glires/Euarchonta, исходя из базы данных TimeTree, составляет 106,69 млн лет, очевидно, что это несколько завышенное значение, поскольку обе группы входят в надотряд Euarchontoglires, дивергировавший с сестринской группой Laurasiatheria примерно 76,7–144,1 млн лет назад [24–29] (средневзвешенное время — 97,96 млн лет [21]). В связи с этим при построении филогенетического дерева видов и дальнейших расчетах мы принимали момент расхождения Glires/Euarchonta в 95,5 млн лет [23]. Существуют также неоднозначности в оценках времени дивергенции отрядов внутри Laurasiatheria. В данном случае были выбраны значения, позволяющие сохранить топологию соответствующего участка филогенетического дерева, приведенного в работе Springer et al. [25].

Из табл. 3 видно, что скорость эволюции ТТФ-азы несколько отличается в разных филогенетических линиях млекопитающих. Действительно, судя по представленным данным, эволюция энзима должна быть более быстрой в семействе Muridae и более медленной у тушканчиков и лошадей. Однако повышенные на фоне других видов темпы эволюции (если это на самом деле имеет место) ТТФ-азы мыши и крысы, надо полагать, никак не связаны с

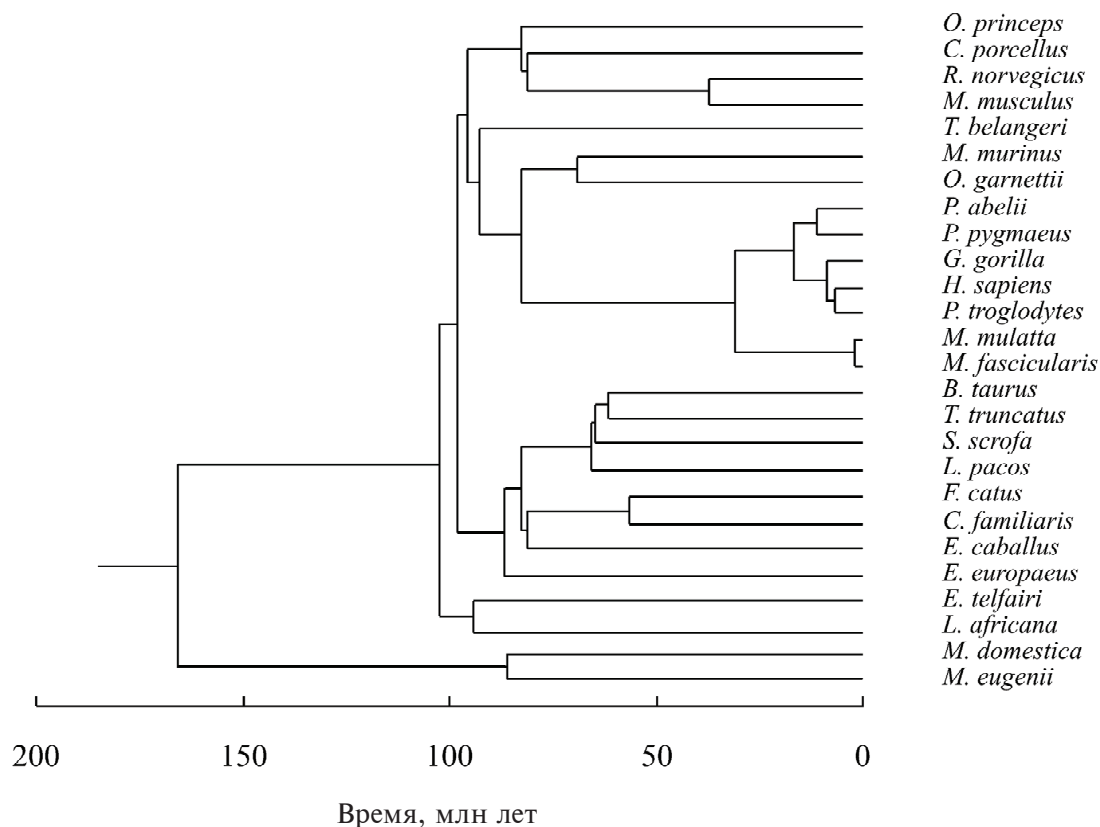


Рис. 1. Филогенетическое дерево 26 видов млекопитающих, построенное с помощью Тахопому-браузера NCBI, базы данных TimeTree и работ [22–26]

кардинальными структурными перестройками энзима, а отражают определенную общую тенденцию. Так, например, было показано, что скорость эволюции ДНК у антропоидов ниже, чем у крыс и мышей, причем соотношение нуклеотидных замен в линии грызунов и человека со времени их дивергенции составляет 1,3 для несинонимических (замещения аминокислот) и 2,0 — для синонимических [30, 31]. Хотя бы отчасти это обстоятельство может объясняться различиями в продолжительности жизни поколения, интенсивности метаболизма или эффективности репарации ДНК [30, 32, 33]. Кроме того, характер межвидовых вариаций в скоростях молекулярной эволюции различается для ядерных и митохондриальных генов [34].

На рис. 2 приведен график, на котором каждая точка представляет собой расчетное значение числа аминокислотных замен (табл. 3), приуроченное ко времени дивергенции попарно сравниваемых видов (рис. 1). Исходя из тангенса угла наклона прямой (a), построенной по уравнению линейной регрессии, средняя k_{aa} равна $1,41 \cdot 10^{-9}$ замен аминокислот на сайт в год. Учитывая размеры полипептид-

ной цепи ТТФ-азы, это должно соответствовать одному молекулярному событию приблизительно за 3 млн лет. Можно также оценить скорость дивергенции ТТФ-азы в терминах UEP, определяемого как «время в миллионах лет, необходимое для возникновения различия в 1% между аминокислотными последовательностями двух дивергировавших линий» [35]. UEP ТТФ-азы, рассчитанный по уравнению прямой в координатах *время дивергенции* — $p_d \cdot 100$ (данные не приведены), составляет 4,4 млн лет.

Хорошо известно, что макромолекулы могут эволюционировать с разными скоростями. Например, UEP цитохрома *c* равен примерно 20 млн лет, гемоглобина — 5,8 млн лет, фибринопептида — 1,1 млн лет [36]. Скорость накопления мутаций является индивидуальной особенностью каждого протеина, по-видимому зависящей от его чувствительности к изменениям. С позиций теории нейтральности здесь применимо понятие о селективных ограничениях: для молекулярной эволюции характерна повышенная частота мутационных замен в менее важных в функциональном отношении

Таблица 3. Матрица долей аминокислотных различий D^p , (выше диагонали) между последовательностями ТТФ-азы 26 видов млекопитающих и расчетных значений K_{aa} (ниже диагонали)

	<i>M. musculus</i>	<i>R. norvegicus</i>	<i>C. porcellus</i>	<i>S. scrofa</i>	<i>C. familiaris</i>	<i>F. catus</i>	<i>B. taurus</i>	<i>M. fascicularis</i>	<i>M. mulatta</i>	<i>P. pygmaeus</i>	<i>H. sapiens</i>	<i>E. caballus</i>	<i>P. troglodytes</i>	<i>T. belangeri</i>	<i>M. murinus</i>	<i>O. garnettii</i>	<i>M. domestica</i>	<i>E. telfairi</i>	<i>E. europaeus</i>	<i>L. africana</i>	<i>P. abelii</i>	<i>O. princeps</i>	<i>G. gorilla</i>	<i>T. trincatus</i>	<i>M. eugeni</i>	<i>L. pascos</i>
<i>M. musculus</i>																										
<i>R. norvegicus</i>	,107																									
<i>C. porcellus</i>	,386	,371																								
<i>S. scrofa</i>	,416	,408	,331																							
<i>C. familiaris</i>	,342	,328	,304	,311																						
<i>F. catus</i>	,378	,378	,266	,338	,195																					
<i>B. taurus</i>	,320	,327	,276	,224	,188	,200																				
<i>M. fascicularis</i>	,349	,364	,279	,298	,266	,248	,194																			
<i>M. mulatta</i>	,328	,342	,272	,285	,242	,236	,170	,018																		
<i>P. pygmaeus</i>	,321	,335	,260	,285	,242	,224	,170	,055	,036																	
<i>H. sapiens</i>	,335	,349	,272	,285	,248	,236	,176	,059	,041	,022																
<i>E. caballus</i>	,308	,308	,224	,260	,195	,206	,153	,178	,156	,150	,161															
<i>P. troglodytes</i>	,335	,349	,285	,311	,254	,242	,188	,069	,050	,031	,018	,172														
<i>T. belangeri</i>	,342	,335	,249	,306	,237	,231	,177	,173	,151	,130	,140	,140	,151													
<i>M. murinus</i>	,421	,421	,328	,349	,244	,256	,221	,220	,202	,191	,197	,244	,208	,215												
<i>O. garnettii</i>	,447	,431	,352	,366	,279	,279	,256	,230	,218	,206	,212	,254	,230	,231	,125											
<i>M. domestica</i>	,520	,538	,426	,533	,458	,482	,423	,482	,458	,434	,466	,411	,482	,437	,523	,560										
<i>E. telfairi</i>	,458	,442	,313	,391	,354	,340	,298	,340	,319	,286	,313	,306	,327	,288	,344	,340	,497									
<i>E. europaeus</i>	,403	,380	,375	,389	,306	,319	,269	,347	,319	,306	,319	,280	,319	,314	,365	,382	,507	,391								
<i>L. africana</i>	,359	,351	,293	,375	,237	,274	,213	,249	,237	,196	,213	,207	,213	,208	,270	,280	,461	,293	,335							
<i>P. abelii</i>	,328	,342	,266	,292	,248	,230	,176	,059	,041	,004	,027	,156	,036	,135	,197	,212	,442	,293	,313	,201						
<i>O. princeps</i>	,405	,398	,342	,384	,335	,335	,277	,282	,256	,244	,256	,269	,269	,264	,367	,377	,472	,364	,379	,330	,250					
<i>G. gorilla</i>	,330	,344	,274	,293	,243	,231	,166	,055	,036	,018	,004	,157	,013	,136	,198	,213	,461	,386	,308	,275	,275	,270				
<i>T. truncatus</i>	,364	,356	,255	,184	,213	,120	,207	,184	,173	,168	,146	,146	,190	,180	,221	,243	,445	,308	,282	,232	,179	,277	,174			
<i>M. eugeni</i>	,547	,547	,447	,513	,447	,479	,413	,463	,439	,424	,447	,409	,463	,419	,511	,539	,116	,469	,513	,458	,432	,502	,442	,419		
<i>L. pascos</i>	,328	,315	,266	,242	,189	,189	,136	,206	,183	,150	,167	,156	,178	,168	,226	,236	,458	,286	,268	,207	,156	,282	,157	,140		

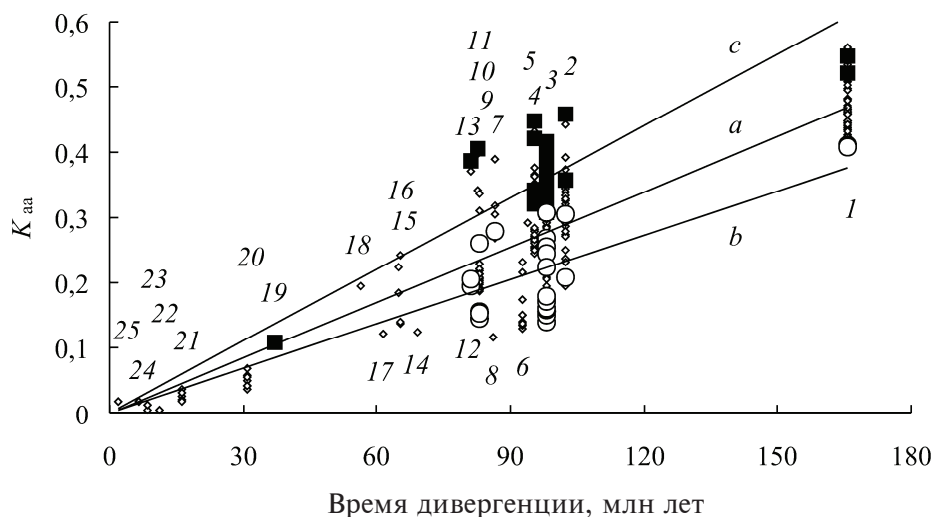


Рис. 2. Зависимость между K_{aa} и временем дивергенции: 1 – *Marsupialia/Eutheria*, 2 – *Afrotheria/Boreoeutheria*, 3 – *Euarchontoglires/Laurasiatheria*, 4 – *Glires/Euarchonta*, 5 – мадагаскарский еж/африканский слон, 6 – *Scandentia/Primates*, 7 – *Eulipotyphila/Cetartiodactyla + Carnivora + Perissodactyla*, 8 – опоссум/валлаби, 9 – *Cetartiodactyla/Carnivora + Perissodactyla*, 10 – *Strepsirrhini/Catarrhini*, 11 – *Rodentia/Lagomorpha*, 12 – *Carnivora/Perissodactyla*, 13 – *Sciurognathi/Hystricognathi*, 14 – лемур/галаго, 15 – *Ruminantia + Cetacea + Suina/Tylopoda*, 16 – *Ruminantia + Cetacea/Suina*, 17 – черноморская афалина/бык, 18 – кошка/собака, 19 – крыса/мышь, 20 – *Hominidae/Cercopitheciidae*, 21 – (группа *Homo/Pan/Gorilla*)/*Pongo*, 22 – борнеанский орангутан/суматранский орангутан, 23 – *Gorilla/Homo + Pan*, 24 – человек/шимпанзе, 25 – макак-резус/яванский макак

молекулах или их частях по сравнению с более важными [15, 35, 37]. Иными словами, вероятность нейтральных мутаций жестко связана с лимитами, налагаемыми на пространственную структуру молекулы. Действительно, с этой точки зрения цитохром *c* – более древний «испытанный» фермент, структура которого сформировалась с возникновением эукариот и прошла эволюционный путь более чем в 10^9 лет, оставив возможность лишь для незначительных изменений. Цитохром *c* играет фундаментальную роль в жизнедеятельности эукариотной клетки, участвуя в переносе электронов по дыхательной цепи от комплекса III (убихинол-цитохром *c*-редуктаза (1.10.2.2)) к комплексу IV (цитохром *c*-оксидаза (1.9.3.1)) [38]. С другой стороны, такие быстро эволюционирующие молекулы как фибринопептиды, которые отщепляются от фибриногена в процессе образования кровяного сгустка, не несут на себе столь существенных функций. По этой причине любая мутация, не препятствующая отщеплению фибринопептидов, оказывается приемлемой для вида, а скорость эволюционных замен может быть очень близкой к истинной скорости мутирования. Фактически же в продуктивном взаимодействии с тромбином, по видимому, участвуют 5 остатков фибринопеп-

тида А (*Phe*-8, *Gly*-12, *Gly*-13, *Arg*-16 и *Arg*-19) и 2 остатка фибринопептида В (*Arg*-14 и *Gly*-15) человека [39]. Основная функция гемоглобина состоит в транспорте кислорода. Этот протеин, построенный из четырех субъединиц, может находиться в двух конформационных состояниях, обозначаемых как Т- и R-формы. В процессе «работы» гемоглобина имеют место кооперативные взаимодействия между субъединицами и, кроме того, на равновесие между Т- и R-формами влияют различные аллостерические эффекторы, регулирующие связывание O_2 [38]. Поэтому, можно ожидать более строгих, чем у фибринопептидов, требований к его структуре и соответственно более низких темпов эволюции. Рассматривая в данном аспекте скорость мутационных замен в молекуле ТТФ-азы, следует, вероятно, говорить о том, что по сравнению с гемоглобином функциональные ограничения, налагаемые на ее пространственную укладку, не столь жестки.

Частота мутационных замен может быть неодинаковой не только для разных протеинов, но и для разных участков одной молекулы. Например, по оценке Kimura & Ohta [40], скорость эволюции гемового «кармана» α -цепи гемоглобина – величина порядка $0,165 \cdot 10^{-9}$, а функционально менее важных наружных

участков молекулы — $1,35 \cdot 10^{-9}$. Еще один характерный случай — эволюция среднего сегмента проинсулина. Эта часть молекулы, отщепляющаяся при образовании активного гормона, накапливает мутации в несколько раз быстрее, чем сам инсулин (пептиды А и В) [15]. Приведенными примерами наглядно иллюстрируется тот факт, что менее значимые в функциональном отношении участки молекул подвержены более быстрой эволюции. На схеме 1 видно, что если исключить из рассмотрения последовательность *M. lucifugus*, то в N-концевом сегменте молекулы энзима протяженностью 68 ао находятся 26 из 65 инвариантных позиций. Другой высоко консервативный участок включает остатки 142-210, тогда как центральная зона энзима и С-концевой фрагмент более вариабельны. Расчеты показывают, что скорость эволюции С-концевой последовательности ТТФ-азы (остатки 211-230) составляет $3,76 \cdot 10^{-9}$, центрального участка (остатки 69-141) — $1,95 \cdot 10^{-9}$, а N-концевой трети (остатки 1-68) — $1,03 \cdot 10^{-9}$ замен на аминокислотный сайт в год. Самая же низкая частота мутаций у фрагмента 142-210. Здесь в среднем имела место лишь одна замена за 17,9 млн лет ($k_{aa} = 0,81 \cdot 10^{-9}$). Опираясь на сказанное выше, можно ожидать, что именно на эту область мо-

лекулы накладываются самые строгие функциональные ограничения, так что большинство мутаций в силу их неблагоприятного влияния на операционные свойства энзима подвергаются отсеиванию в результате действия селективного отбора. Пока, однако, неясно, как все это соотносится с существованием «укороченных» вариантов ТТФ-азы, мРНК которых обнаружена у собаки и кролика (схема 2). Если эти формы окажутся каталитически состоятельными, то наличие высококонсервативных медленно эволюционирующих «лишних» частей молекулы полноразмерного энзима по-видимому нельзя будет объяснить иначе, чем их важной ролью в детерминации специфичности или регуляторных свойств.

Ни один из алгоритмов, широко применяемых в молекулярной филогении (Maximum Likelihood, Maximal Parsimony, Neighbor-joining, UPGMA), не дает дерева ТТФ-азы, эквивалентного видовому дереву, изображенному на рис. 1. В качестве примера на рис. 3 показано дерево ТТФ-азы, построенное методом Neighbor-joining. Не трудно заметить, что в обоих деревьях достаточно хорошо соответствуют лишь топологии частей, объединяющих группы Euarchonta и Carnivora + Cetartiodactyla (за исключением лошади). Устойчивость боль-

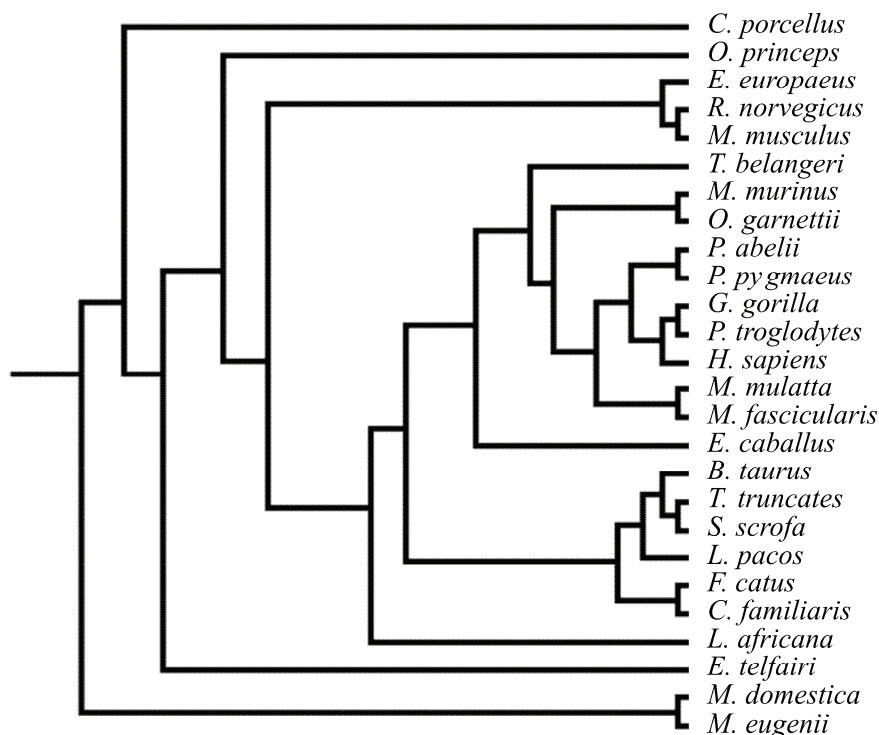


Рис. 3. Филогенетическое дерево ТТФ-азы, построенное методом Neighbor-joining

шинства узлов в этих участках реконструированного дерева ТТФ-азы подтверждается результатами бутстреп-анализа по 1000 репликам, однако в целом топология данного дерева (как и построенных с помощью других методов) не достоверна.

Стохастический характер процесса молекулярной эволюции [41] едва ли позволяет рассчитывать на то, что форма дерева, реконструированного исходя из различий в аминокислотных последовательностях одного протеина (гена), могла бы в точности отвечать дереву видов. Более того, молекулярные и видовые деревья, как правило, должны отличаться, поскольку узлы первых соответствуют молекулярным событиям — дивергенции гена, осуществляемой в результате мутационного процесса, тогда как узлы вторых отражают события видообразования, т.е. расщепления популяции исходного вида на две группы, не способные к скрещиванию например в результате географической изоляции. Маловероятно, чтобы эти события осуществились одновременно. К несоответствию молекулярного и видового деревьев могут также приводить различия в скорости накопления мутаций в исследуемом гене (протеине) у разных видов организмов. Использование likelihood ratio теста (<http://evolution.genetics.washington.edu/phylip>) указывает на то, что гипотеза “молекулярных часов” [41], по-видимому, неприменима в отношении эволюции ТТФ-азы ($p = 0,001$). Как уже отмечалось выше, скорость эволюции энзима в разных филогенетических линиях не совсем постоянна. По данным рис. 2, например, можно рассчитать, что в родословной лошади (кружки, прямая b) $k_{aa} = 1,13 \cdot 10^{-9}$. Это в 1,62 раза ниже скорости эволюции ТТФ-азы мыши ($k_{aa} = 1,83 \cdot 10^{-9}$, квадраты, прямая c).

В заключение следует кратко коснуться еще одной проблемы, имеющей непосредственное отношение к эволюции макромолекул. Это — полиморфизм протеинов. Хотя споры о причинах внутривидовой генетической изменчивости на молекулярном уровне не утихли до сих пор [42], уже к началу 1980-х были получены свидетельства того, что характер и степень протеинового полиморфизма в пределах вида хорошо согласуются с предсказаниями теории нейтральности, утверждающей, что данный феномен является переходной фазой молекулярной эволюции [43–45]. С точки зрения теории нейтральности для локусов, которые кодируют протеины с жесткими функциональными ограничениями, характерен более низкий уровень генетической из-

менчивости, т.е. молекулы (или их части), важные в функциональном отношении и поэтому медленно эволюционирующие, должны иметь низкий уровень гетерозиготности. Основываясь на приведенных выше расчетах, мы вправе говорить о том, что скорость эволюции ТТФ-азы несколько выше или приблизительно равна скорости эволюции гемоглобина. Следовательно, можно было бы ожидать и сходного уровня полиморфизма этих протеинов. В настоящее время в популяции человека обнаружено более 300 генетических вариантов гемоглобина, большинство из которых представляют собой результат единичных аминокислотных замен, не оказывающих существенного влияния на структурно-функциональные свойства молекулы и не имеющих фенотипического выражения [46]. Каждый из этих вариантов гемоглобина — продукт измененного гена. Однако в базе данных SNP (Build 130 <http://www.ncbi.nlm.nih.gov>) зарегистрировано всего 2 аллельных варианта ТНТРА, полиморфных по 1 нуклеотиду в транслируемом регионе мРНК (rs34015250). Это противоречие может найти простое объяснение с тех же позиций времени: если история структурных исследований гемоглобина продолжается почти полвека, то истоки исследований ТТФ-азы на тонком молекулярном уровне берут свое начало лишь в преддверии нынешнего столетия [7].

ШВИДКІСТЬ МОЛЕКУЛЯРНОЇ ЕВОЛЮЦІЇ РОЗЧИННОЇ ТІАМІНТРИФОСФАТАЗИ ССАВЦІВ

А. Ф. Макаричиков

Гродненський державний аграрний
університет, Беларусь;
Інститут фармакології та біохімії
НАН Беларусі, Гродно;
e-mail: a_makarchikov@yahoo.com

Тіамінтрифосфат (ТТФ) присутній у клітинах різних типів — від бактерій до ссавців. У ссавців гідроліз ТТФ здійснюється специфічною розчинною ТТФ-азою, яка експериментально не виявлена у представників інших груп організмів. У базах даних NCBI і Ensembl ми знайшли інформацію щодо повнорозмірних і часткових амінокислотних послідовностей ТТФ-ази 38 видів ссавців. Швидкість еволюційних замінів амінокислот на сайт за 1 рік (k_{aa}) у молекулі ТТФ-ази, розрахована виходячи з наявних даних, дорівнює $1,41 \cdot 10^{-9}$, що відповідає елементарному періоду еволюції приблизно в 4,4 млн років. Для різних ділянок

поліпептидного ланцюга ензиму характерною є різна швидкість еволюції. Більш варіабельні 3-кінцевий фрагмент ($k_{aa} = 3,76 \cdot 10^{-9}$ розрахована тільки для видів з довжиною поліпептиду ТТФ-ази 230 ао і слона), а також центральна область (залишки 69-141; $k_{aa} = 1,95 \cdot 10^{-9}$), тоді як k_{aa} для N-кінцевої послідовності (залишки 1-68) складає $1,03 \cdot 10^{-9}$, а для фрагмента 142-210 — $0,81 \cdot 10^{-9}$.

Ключові слова: тіамінтрифосфатаза, швидкість молекулярної еволюції, ссавці.

THE RATE OF MOLECULAR EVOLUTION OF MAMMALIAN SOLUBLE THIAMINE TRIPHOSPHATASE

A. F. Makarchikov

Grodno State Agricultural University, Belarus;
Institute of Pharmacology and Biochemistry, National
Academy of Sciences of Belarus, Grodno, Belarus;
e-mail: a_makarchikov@yahoo.com

Summary

Thiamine triphosphate (ThTP) exists in various living cells — from bacteria to mammals. ThTP concentration in mammals is regulated by a specific soluble ThTPase, which has not been revealed experimentally, however, in other organisms. In NCBI and Ensembl databases we have found information about full-size or partial amino acid sequences of the enzyme from 38 mammal species. An average rate of amino acid substitutions (k_{aa}) in ThTPase molecule was estimated from the data available to be $1.41 \cdot 10^{-9}$ per site per year. This corresponds to unit evolutionary period of about 4.4 million years. The evolutionary rate varies for different portions of the enzyme, C-terminal fragment being the most variable ($k_{aa} = 3.76 \cdot 10^{-9}$; calculated only for 230 aa species + elephant). An average replacement rate of $1.95 \cdot 10^{-9}$ per amino acid site per year was calculated for the central portion of the enzyme (residues 69-141), while N-terminal sequence (residues 1-68) and 142-210 fragment evolved with k_{aa} of $1.03 \cdot 10^{-9}$ and $0.81 \cdot 10^{-9}$, respectively.

Key words: thiamine triphosphatase, molecular evolution rate, mammals.

1. *Enzyme nomenclature* / www.expasy.ch
2. Bettendorff L., Wins P. // *Recent Res. Devel. Neurochem.* — 1999. — 2. — P. 37–62.
3. Макарович А. Ф., Беттендорфф Л. // *Новости мед.-биол. наук.* — 2004. — № 2. — С. 70–76.
4. Lakaye B., Wirtzfeld B., Wins P. et al. // *J. Biol. Chem.* — 2004. — 279. — P. 17142–17147.
5. Makarchikov A. F., Lakaye B., Gulyai I. E. et al. // *Cell. Mol. Life Sci.* — 2003. — 60. — P. 1477–1488.
6. Макарович А. Ф. // *Тиамінтрифосфат: индикатор клеточного стресса? XI съезд Белорус. об-ва физиол.* — Минск. — 2006. — С. 81–82.
7. Makarchikov A. F., Chernikevich I. P. // *Biochim. Biophys. Acta.* — 1992. — 1117. — P. 326–332.
8. Makarchikov A. F. // *J. Biochem. Mol. Biol. Biophys.* — 2001. — 5. — P. 75–82.
9. Lakaye B., Makarchikov A. F., Fernandes Antunes A. et al. // *J. Biol. Chem.* — 2002. — 277. — P. 13771–13777.
10. Szyniarowski P., Lakaye B., Czerniecki J. et al. // *Biochim. Biophys. Acta.* — 2005. — 1725. — P. 93–102.
11. Iyer L. M., Aravind L. // *BMC Genomics.* — 2002. — 3:33.
12. Макарович А. Ф., Беттендорфф Л. // *Новости мед.-биол. наук.* — 2005. — № 2. — С. 55–60.
13. Гуляй И. Э., Макарович А. Ф. // *Весті НАН Беларусі. Сер. біял. навук.* — 2002. — № 4. — С. 78–81.
14. Altschul S. F., Gish W., Miller W. et al. // *J. Mol. Biol.* — 1990. — 215. — P. 403–410.
15. Кумура М. *Молекулярная эволюция: теория нейтральности.* — М.: Мир, 1985. — 394 с.
16. Thompson J. D., Higgins D. G., Gibson T. J. // *Nucleic Acids Res.* — 1994. — 22. — P. 4673–4680.
17. Boguski M. S., Lowe T. M., Tolstoshev C. M. // *Nat. Genet.* — 1993. — 4. — P. 332–333.
18. Maglott D., Ostell J., Pruitt K. D., Tatusova T. // *Nucleic Acids Res.* — 2005. — 33. — P. D54–D58.
19. Finn R. D., Mistry J., Tate J. et al. // *Ibid.* — 2008. — 36. — P. D281–D288.
20. Schomburg I., Hofmann O., Baensch C. // *Gene Funct. Dis.* — 2000. — 3–4. — P. 109–118
21. Hedges S. B., Dudley J., Kumar S. // *Bioinformatics.* — 2006. — 22. — P. 2971–2972.
22. Hayasaka K., Fujii K., Horai S. // *Mol. Biol. Evol.* — 1996. — 13. — P. 1044–1053.
23. Nei M., Glazko G. A. // *J. Hered.* — 2002. — 93. — P. 157–164.
24. Kumar S., Hedges B. // *Nature.* — 1998. — 392. — P. 917–919.
25. Springer M. S., Murphy W. J., Eizirik E., O'Brien S. J. // *Proc. Natl. Acad. Sci. USA.* — 2003. — 100. — P. 1056–1061.

26. Hasegawa M., Thorne J. L., Kishino H. // *Genes Genet. Syst.* — 2003. — **78**. — P. 267–283.
27. Douzery E. J., Huchon D. // *Mol. Phylogenet. Evol.* — 2004. — **33**. — P. 922–935.
28. Blair J. E., Shah P., Hedges S. B. // *BMC Bioinformaticings.* — 2005. — 6:53.
29. Pereira S. L., Baker A. J. // *Mol. Biol. Evol.* — 2006. — **23**. — P. 1731–1740.
30. Britten J. R. // *Science.* — 1986. — **231**. — P. 1393–1398.
31. Wu C. I., Li W. H. // *Proc. Natl. Acad. Sci. USA.* — 1985. — **82**. — P. 1741–1745.
32. Martin A. P., Palumbi S. R. // *Ibid.* — 1993. — **90**. — P. 4087–4091.
33. Bromham L. // *Mol. Biol. Evol.* — 2002. — **19**. — P. 309–309.
34. Vawter L., Brown M. W. // *Science.* — 1986. — **234**. — P. 194–196.
35. Dickerson R. E. // *J. Mol. Evol.* — 1971. — **1**. — P. 26–45.
36. Хадорн Э., Венер Р. *Общая зоология.* М.: Мир, 1989. — 528 с.
37. King J. L., Jukes T. H. // *Science.* — 1969. — **164**. — P. 788–798.
38. Кольман Я., Рем К.-Х. *Наглядная биохимия.* — М.: Мир, 2000. — 469 с.
39. Binnie C. G., Lord S. T. // *Blood.* — 1993. — **81**. — P. 3186–3192.
40. Kimura M., Ohta T. // *Genetics.* — 1973. — **73**. — P. 19–35.
41. Zuckerkandl E., Pauling L. *Evolutionary divergence and convergence in proteins / Evolving genes and proteins*, V. Bryson, H. J. Vogel, Eds. NY: Academic Press. — 1965. — P. 97–166.
42. Nei M. // *Mol. Biol. Evol.* — 2005. — **22**. — P. 2318–2342.
43. Chakraborty R., Fuerst P. A., Nei M. // *Genetics.* — 1978. — **88**. — P. 367–390.
44. Skibinski D. O. F., Ward R. D. // *Genet. Res.* — 1981. — **38**. — P. 71–92.
45. Yamazaki T., Maruyama T. // *Science.* — 1972. — **178**. — P. 56–58.
46. Nelson D. L., Cox M. M. *Lehninger principles of biochemistry.* NY: Worth Publishers. — 2000. — 1232 p.

Получено 28.08.2007