

А. И. Ковалевская¹, Т. В. Романовская²

¹Институт биоорганической химии НАН Беларуси, Минск, Республика Беларусь

²Белорусский государственный университет, Минск, Республика Беларусь

ЭВОЛЮЦИОННАЯ ИСТОРИЯ ГЕНОВ СЕМЕЙСТВА *MTG* У ПОЗВОНОЧНЫХ

Аннотация. Высококонсервативное генное семейство *MTG* у позвоночных включает три гомолога – *MTG8*, *MTGR1*, *MTG16*, которые кодируют транскрипционные корепрессоры, играющие важную роль в гемопоэзе, нейрогенезе и дифференцировке эпителиальных стволовых клеток. Эти гены являются чрезвычайно важными, потому что они способны вовлекаться в транслокации, ассоциированные с различными типами рака. Изучение эволюции этого семейства способно пролить свет на то, как происходила их структурно-функциональная дивергенция. Для изучения эволюции этого генного семейства был проведен филогенетический анализ нуклеотидных и аминокислотных последовательностей, выяснена доменная организация продуктов генов семейства *MTG*, определен механизм появления первого гена *MTG* и таксон – родоначальник этого гена. Кроме того, установлен механизм экспансии генного семейства, проведен анализ скоростей эволюции, действующей на отдельные домены белков семейства *MTG*, и определены консервативные позиции внутри каждого гена.

Ключевые слова: генное семейство *MTG*, филогенетический анализ, доменная организация, эволюция

Для цитирования: Ковалевская, А. И. Эволюционная история генов семейства *MTG* у позвоночных / А. И. Ковалевская, Т. В. Романовская // Вест. Нац. акад. наук Беларусі. Сер. біял. навук. – 2019. – Т. 64, № 4. – С. 391–402. <https://doi.org/10.29235/1029-8940-2019-64-4-391-402>

A. I. Kavaleuskaya¹, T. V. Ramanouskaya²

¹Institute of Bioorganic Chemistry of the National Academy of Sciences of Belarus, Minsk, Republic of Belarus

²Belarusian State University, Minsk, Republic of Belarus

EVOLUTIONARY HISTORY OF THE *MTG* GENE FAMILY IN VERTEBRATES

Abstract. The highly conserved *MTG* gene family includes three homologs in vertebrates (*MTG8*, *MTGR1*, *MTG16*) encoding transcriptional corepressors, which are important in haemopoiesis, neurogenesis and epithelial stem cell differentiation. These genes are of particular interest because they are involved in translocations, associated with different types of cancer. Looking at how this gene family evolved might provide insights into history of its structural and functional diversification. We have performed a phylogenetic analysis of *MTG* nucleotide and protein sequences to examine the evolutionary events. The domain organization of *MTG* gene products was clarified, the mechanism of appearance of the first *MTG* gene was revealed and the ancestor taxon was determined. Also the mechanism of *MTG* gene family emergence was established. In addition, analysis of the rates of evolution acting on individual domains was made, and conservative positions within each gene of *MTG* family were determined.

Keywords: *MTG* gene family, phylogenetic analysis, domain organization, evolution

For citation: Kavaleuskaya A. I., Ramanouskaya T. V. Evolutionary history of the *MTG* gene family in vertebrates. *Vesti Natsyynal'noi akademii navuk Belarusi. Seryya biyalagichnych navuk = Proceedings of the National Academy of Sciences of Belarus. Biological series*, 2019, vol. 64, no. 4, pp. 391–402 (in Russian). <https://doi.org/10.29235/1029-8940-2019-64-4-391-402>

Введение. Высококонсервативное семейство генов *MTG* (myeloid translocation gene) у мыши и у человека включает три гена, продукты которых являются транскрипционными корепрессорами: *CBFA2T1* (синонимы *MTG8*, *RUNX1T1*, *ETO*), *CBFA2T2* (синоним *MTGR1*) и *CBFA2T3* (синонимы *MTG16*, *ETO-2*). Представители этого семейства регулируют транскрипцию генов-мишеней, образуя комплексы с ДНК-связывающими белками, и рекрутируют другие корепрессоры и гистоновые деацетилазы [1].

Семейство *MTG* было открыто при изучении острого миелоидного лейкоза [2, 3]. Тогда же было показано, что одной из базовых мутаций, приводящих к этой неоплазии, является транслокация t(8;21), затрагивающая гомолог гена *runt* дрозофилы (*AML1*, 21-я хромосома) и ген на

8-й хромосоме, который назвали *MTG8* (myeloid translocation gene on chromosome 8). Позже было выяснено, что все три гена семейства могут вовлекаться в транслокации, ассоциированные со злокачественными заболеваниями: *MTG16* – в образование гибридного гена *AML1-MTG16* t(16;21) у пациентов с индуцированным терапией лейкозом, *MTGR1* – в образование гена *AML1-MTGR1* t(20;21) при колоректальном раке [4].

У дрозофилы в геноме обнаруживается единственный ген, имеющий высокий процент сходства и идентичную с белками семейства *MTG* позвоночных доменную организацию. Высокая консервативность указывает на функциональную значимость белка. Действительно, гены данного семейства задействованы в генетическом контроле морфогенеза различных тканей на этапе эмбрионального развития. Нокаут этих генов в эксперименте приводит к гибели эмбрионов мышей [5].

Как известно, семейства генов формируются вследствие дупликаций с последующей дивергенцией генов по структуре и функции. Наше исследование эволюционной истории генов семейства *MTG* основывалось на представленных в базах данных последовательностях геномов животных.

Цель работы – детальное изучение эволюции структуры и функции генов данного семейства для углубления понимания процессов, связанных как с дифференцировкой тканей, так и с их злокачественной трансформацией у позвоночных.

Материалы и методы исследования. Последовательности генов и кодируемых ими белков были получены из баз данных NCBI (<https://www.ncbi.nlm.nih.gov>), UniProt (<https://www.uniprot.org>), Ensemble (<http://www.ensembl.org/index.html>).

Домены были аннотированы с помощью скрытых марковских моделей (HMMs) в Pfam (<http://pfam.xfam.org>); домен NHR3 был аннотирован в COILS (http://embnet.vital-it.ch/software/COILS_form.html).

Полученный набор последовательностей выравнивали, используя алгоритм MUSCLE из пакета MEGA7, и редактировали с помощью Gblocks (http://molevol.cmima.csic.es/castresana/Gblocks_server.html) с минимальной жесткостью селекции позиций. Отредактированные последовательности использовали для построения деревьев программой BEAST2. Достоверность филогенетических реконструкций оценивали по апостериорной вероятности узлов.

Результаты и их обсуждение. Все белки семейства *MTG* характеризуются наличием 4 консервативных доменов, названных на основании гомологии с доменами белка *neuvy* у дрозофилы (рис. 1): NHR1 (nervy homology region 1), NHR2, NHR3, NHR4. Домен NHR1 (TAFH, от англ. TATA-box associated factor homology), гомологичный TAF4-субъединице TFIID, обеспечивает избирательное связывание белка с некоторыми транскрипционными факторами, определяющими в норме спектр генов-мишеней репрессорной активности белков *MTG* [6]. Домен NHR2 содержит гидрофобный гептадный повтор и обеспечивает гомо- и гетероолигомеризацию белков семейства *MTG*, что многократно повышает их сродство к корепрессорам и транскрипционным факторам и является критичным для реализации их регуляторной функции. Вместе с тем совместно NHR1 и NHR2 образуют платформу для других белок-белковых взаимодействий [7].

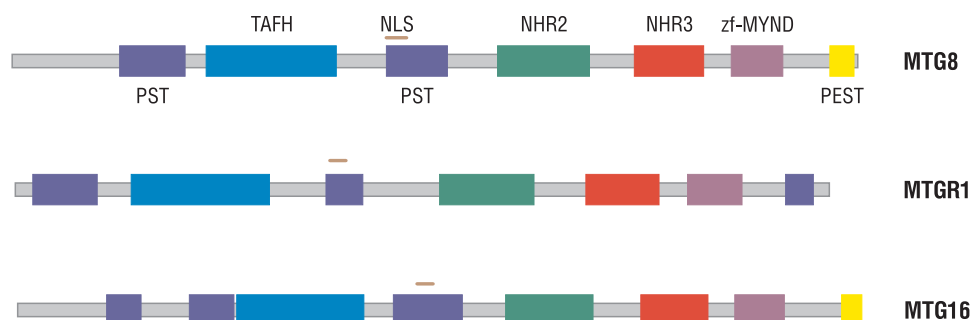


Рис. 1. Доменная структура трех белков семейства *MTG* человека. NHR1, NHR2, NHR3, zf-MYND – консервативные домены, PST – пролин/серин/треонин-богатый тракт, PEST – пролин/глутаминовая кислота/серин/треонин-богатый тракт, NLS – сигнал ядерной локализации
Fig. 1. Domain structure of three proteins of the human *MTG* family. NHR1, NHR2, NHR3, zf-MYND – conservative domains, PST – proline/serine/threonine-rich tract, PEST – proline/ glutamic acid/serine/threonine-rich tract, NLS – nuclear localization signal

Домен NHR3 образует специфическую структуру типа суперскрученных α -спиралей, а также участвует в олигомеризации белка и рекрутировании белков-корепрессоров [8]. Домен NHR4 (zf-MYND, zinc fingers of Myeloid translocation protein, Nervy, Deaf1) содержит два высококонсервативных мотива цинковых пальцев типа SxxSxxxxxxxSxxC и SxxSxxxxxxxHxxC. Несмотря на то что обычно цинковые пальцы связываются с ДНК и регулируют экспрессию генов, у белков семейства *MTG* домен zf-MYND осуществляет только белок-белковые взаимодействия и является центром рекрутирования белков-корепрессоров N-CoR и SMRT [9].

Специфической характеристикой генов семейства *MTG* является сочетание в определенном порядке четырех основных доменов (TAFH, NHR2, NHR3 и zf-MYND). Ни у одного из представителей одноклеточных гена с таким сочетанием доменов не встречается, однако уже у *Cnidaria* есть полный ген, продукт которого содержит все четыре домена. Следовательно, вероятный момент формирования гена-родоначальника приходится на эволюцию наиболее ранних представителей *Metazoa*. Таковыми являются губки (*Porifera*), гребневники (*Ctenophora*) и *Placozoa* (таксон, представленный в настоящее время несколькими видами трихоплаксов). К сожалению, до настоящего времени нет однозначного решения вопроса о том, в каком порядке происходило формирование и обособление данных групп. Нами проанализированы геномы некоторых представителей этих таксонов на наличие генов-гомологов семейства *MTG*.

Было установлено, что у представителей каждого таксона имеется ген, содержащий как минимум два домена (TAFH и zf-MYND), первый из которых также присутствует в составе транскрипционного фактора TFIIID, а второй – в белке DEAF-1, имеющихся у многих одноклеточных организмов, включая хоанофлагеллят – предковый для *Metazoa* таксон. Можно предположить, что объединение этих доменов в одной общей рамке считывания произошло вследствие хромосомной перестройки, породившей новый ген с регуляторной функцией.

Домены NHR2 и NHR3 по гомологии последовательностей у примитивных *Metazoa* не выявляются, равно как и в геномах одноклеточных, однако у *Amphimedon queenslandica* и *Trichoplax adhaerens* участок с соответствующей структурой расположен как раз там, где у более высокоорганизованных животных располагается домен NHR3. Один из вероятных сценариев появления гена семейства *MTG* показан на рис. 2. Альтернативный сценарий может предполагать утрату ранее сформированных доменов: NHR2 – у *Porifera* и *Placozoa*, NHR2 и NHR3 – у *Ctenophora*.

Для изучения истории семейства *MTG* и раскрытия основных эволюционных процессов, приведших к его образованию, нами были выбраны виды позвоночных и беспозвоночных с наиболее полно прочитанными и аннотированными геномами. Представителями базальных ветвей были выбраны губка *Amphimedon queenslandica*, гребневик *Mnemiopsis leydi*, трихоплакс *Trichoplax belcheri* и морской анемон *Nematostella victensis*. Клада первичноротых представлена *Lophotrochozoa*

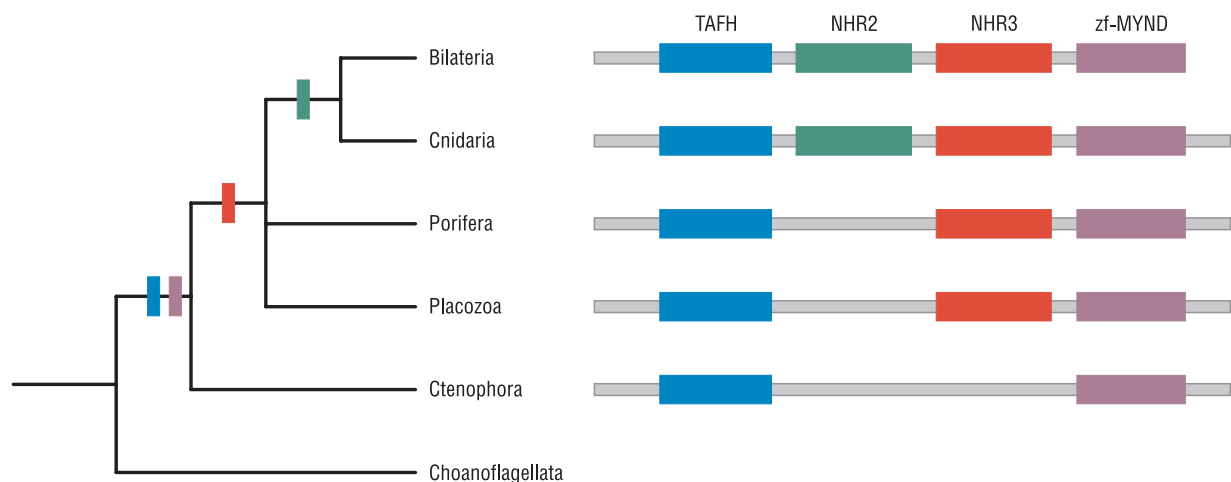


Рис. 2. История появления гена – родоначальника семейства *MTG* у ранних *Metazoa*. Вертикальные цветные метки обозначают моменты включения новых доменов в структуру белка

Fig. 2. Origin history of the ancestral *MTG* family gene in basal *Metazoa*. Colored bars indicate the domain incorporation in the protein structure

(моллюски *Aplysia californica*, *Lottia gigantean* и *Mizuhopecten yessoensis*) и *Ecdysozoa* (*Drosophila arizonae*, *Atta cephalotes*, *Dufourea novaengliae*, *Galendromus occidentalis*); клада вторичноротых – головохордовыми (ланцетник *Branchiostoma floridae*), иглокожими (морской еж *Strongylocentrotus purpurea*, морская звезда *Patiria miniata*) и позвоночными (полный список см. в таблице).

У большинства беспозвоночных выявлен только один гомолог генов семейства *MTG*, в то время как у всех позвоночных обнаружено по три гомологичных гена. Однако в геноме четырех представителей хелицерных присутствуют два гомолога (у клещей родов *Galendromus* и *Varroa*, паука *Stegodyphus* и мечехвоста *Limulus*). Мы построили филогенетическое дерево, которое включало гены семейства *MTG* нескольких позвоночных и нескольких беспозвоночных, имеющих один или два гена-гомолога. Результаты анализа отображены на рис. 3.

Исходя из топологии дерева, можно сделать заключение, что дубликации, сформировавшие семейства генов у позвоночных, произошли независимо в группе вторичноротых после обособления ветви хордовых от ветви иглокожих. Что касается представителей хелицерных, то ввиду достаточно сильной дивергенции последовательностей сложно сказать однозначно, имела ли место одна дубликация, произошедшая в основании хелицерных (недавнее исследование показало, что обособление данной группы сопровождалось полногеномной дубликацией [10]), или две независимые дубликации в каждой из ветвей. Следует отметить, что в базе данных по геномам

**Виды вторичноротых, чьи последовательности белков *MTG* использовались для построения филогении
Species of Deuterostomes used to reconstruct *MTG* proteins phylogeny**

Название (рус.)	Название (лат.)	Таксономия
Ланцетник	<i>Branchiostoma belcheri</i>	Chordata, Cephalochordata
Данио	<i>Danio rerio</i>	Chordata, Vertebrata, Gnathostomata, Teleostomi, Actinopterygii
Латимерия	<i>Latimeria chalumnae</i>	Chordata, Vertebrata, Gnathostomata, Teleostomi, Sarcopterygii
Ксенопус	<i>Xenopus laevis</i>	Chordata, Vertebrata, Gnathostomata, Tetrapoda, Amphibia
Анолис	<i>Anolis carolinensis</i>	Chordata, Vertebrata, Gnathostomata, Tetrapoda, Sauropsida, Lepidosauria, Squamata, Bifurcata, Unidentata, Episquamata, Iguania
Зеленая черепаха	<i>Chelonia midas</i>	Chordata, Vertebrata, Gnathostomata, Tetrapoda, Sauropsida, Archelosauria, Testudines
Острочешуйная куфия	<i>Protobothrops mucrosquamatus</i>	Chordata, Vertebrata, Gnathostomata, Tetrapoda, Sauropsida, Lepidosauria, Squamata, Bifurcata, Unidentata, Episquamata, Serpentes
Гребнистый крокодил	<i>Crocodylus porosus</i>	Chordata, Vertebrata, Gnathostomata, Tetrapoda, Sauropsida, Archelosauria, Archosauria, Crocodylia
Банкивская джунглевая курица	<i>Gallus gallus</i>	Chordata, Vertebrata, Gnathostomata, Tetrapoda, Sauropsida, Archelosauria, Archosauria, Aves
Африканский слон	<i>Loxodonta africana</i>	Chordata, Vertebrata, Gnathostomata, Tetrapoda, Mammalia, Theria, Eutheria, Afrotheria
Косатка	<i>Orcinus orca</i>	Chordata, Vertebrata, Gnathostomata, Tetrapoda, Mammalia, Theria, Eutheria, Boroetheria, Laurasiatheria, Cetartiodactyla, Cetacea
Дикая свинья	<i>Sus scrofa</i>	Chordata, Vertebrata, Gnathostomata, Tetrapoda, Mammalia, Theria, Eutheria, Boroetheria, Laurasiatheria, Cetartiodactyla, Suina
Гималайский листонос	<i>Hipposideros armiger</i>	Chordata, Vertebrata, Gnathostomata, Tetrapoda, Mammalia, Theria, Eutheria, Boroetheria, Laurasiatheria, Chiroptera
Собака	<i>Canis lupus familiaris</i>	Chordata, Vertebrata, Gnathostomata, Tetrapoda, Mammalia, Theria, Eutheria, Boroetheria, Laurasiatheria, Carnivora
Норвежская крыса	<i>Rattus norvegicus</i>	Chordata, Vertebrata, Gnathostomata, Tetrapoda, Mammalia, Theria, Eutheria, Boroetheria, Euarchontoglires, Glires, Rodentia
Домовый опоссум	<i>Monodelphis domestica</i>	Chordata, Vertebrata, Gnathostomata, Tetrapoda, Mammalia, Theria, Metatheria
Обыкновенный шимпанзе	<i>Pan troglodytes</i>	Chordata, Vertebrata, Gnathostomata, Tetrapoda, Mammalia, Theria, Eutheria, Boroetheria, Euarchontoglires, Primates, Catarrhini
Человек	<i>Homo sapiens</i>	Chordata, Vertebrata, Gnathostomata, Tetrapoda, Mammalia, Theria, Eutheria, Boroetheria, Euarchontoglires, Primates, Catarrhini
Боливийский саймири	<i>Saimiri boliviensis boliviensis</i>	Chordata, Vertebrata, Gnathostomata, Tetrapoda, Mammalia, Theria, Eutheria, Boroetheria, Euarchontoglires, Primates, Platyrrhini

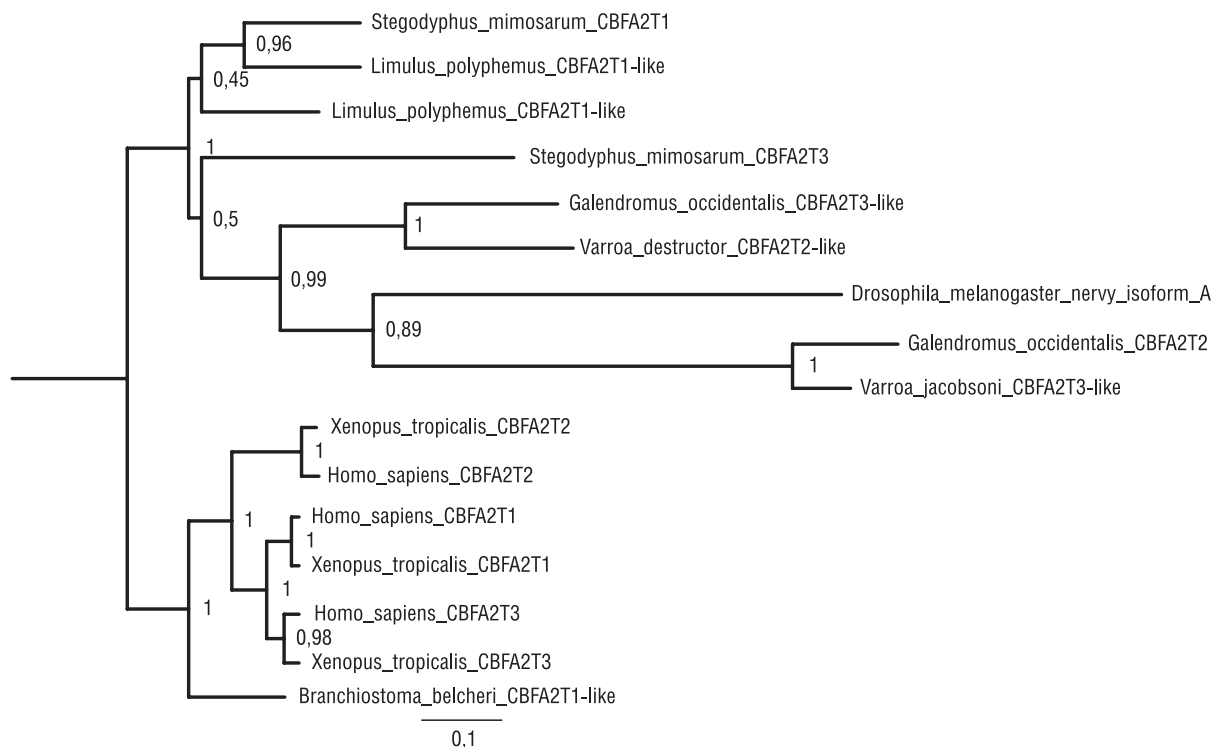


Рис. 3. Филогенетическое дерево белков *MTG* у членистоногих и позвоночных. Длина ветвей пропорциональна количеству аминокислотных замен, шкала соответствует 1 аминокислотной замене на 10 сайтов. Указаны апостериорные вероятности узлов

Fig. 3. Phylogenetic tree of *MTG* proteins of arthropods and chordates. Branch length are proportional to the number of amino acid substitutions, scale bar corresponds to 1 amino acid substitution per 10 sites. Numbers at nodes indicate posterior probability

различных беспозвоночных для гомологов генов семейства *MTG* необоснованно используют названия *CBFA2T1-like*, *CBFA2T2-like* и *CBFA2T3-like*, что вносит путаницу, поскольку это не соответствует истинному ходу эволюции.

Далее мы провели более детальное исследование филогении членов семейства в пределах позвоночных, в том числе представителей группы амфибий, рептилий и разных отрядов млекопитающих. В качестве представителя внешней группы при построении дерева был включен ланцетник, который имеет только один гомолог для генов семейства *MTG* (рис. 4). По результатам анализа было получено дополнительное подтверждение того, что генное семейство *MTG*, члены которого имеются у всех позвоночных, появилось в ходе двух раундов дупликаций, предшествовавших дивергенции данной клады. Топология оставалась устойчивой при выборе в качестве представителя внешней группы других беспозвоночных, таких как полип *Nematostella victensis* или морской еж *Strongylocentrotus purpuratus*.

Проведенный анализ согласуется с аннотацией семейства генов *MTG* как онологов, образованных в результате двух раундов полногеномной дупликации, предшествовавших появлению позвоночных, в базе данных OHNOLOGS (<http://ohnologs.curie.fr/cgi-bin/SearchPage.cgi>).

Известно, что разные участки генов и, соответственно, их белковых продуктов эволюционируют с разной скоростью, что в большинстве случаев определяется различиями в отборе, действующем на данные участки последовательностей. Действие движущего отбора ускоряет накопление несинонимичных замен, что проявляется в удлинении ветвей дерева. Движущий отбор подразумевает изменение функциональной нагрузки и белков-партнеров у изучаемого белка. Для анализа скорости эволюции каждого индивидуального домена белков семейства *MTG* нами было использовано дерево с фиксированной топологией. Информация о филогенетических взаимоотношениях организмов, включенных в анализ, была получена из проекта Taxonomy базы NCBI. Ланцетник служил в качестве представителя внешней группы. Филогенетические построения по каждому домену представлены на рис. 5–8 соответственно. Дерево по домену TAFH не

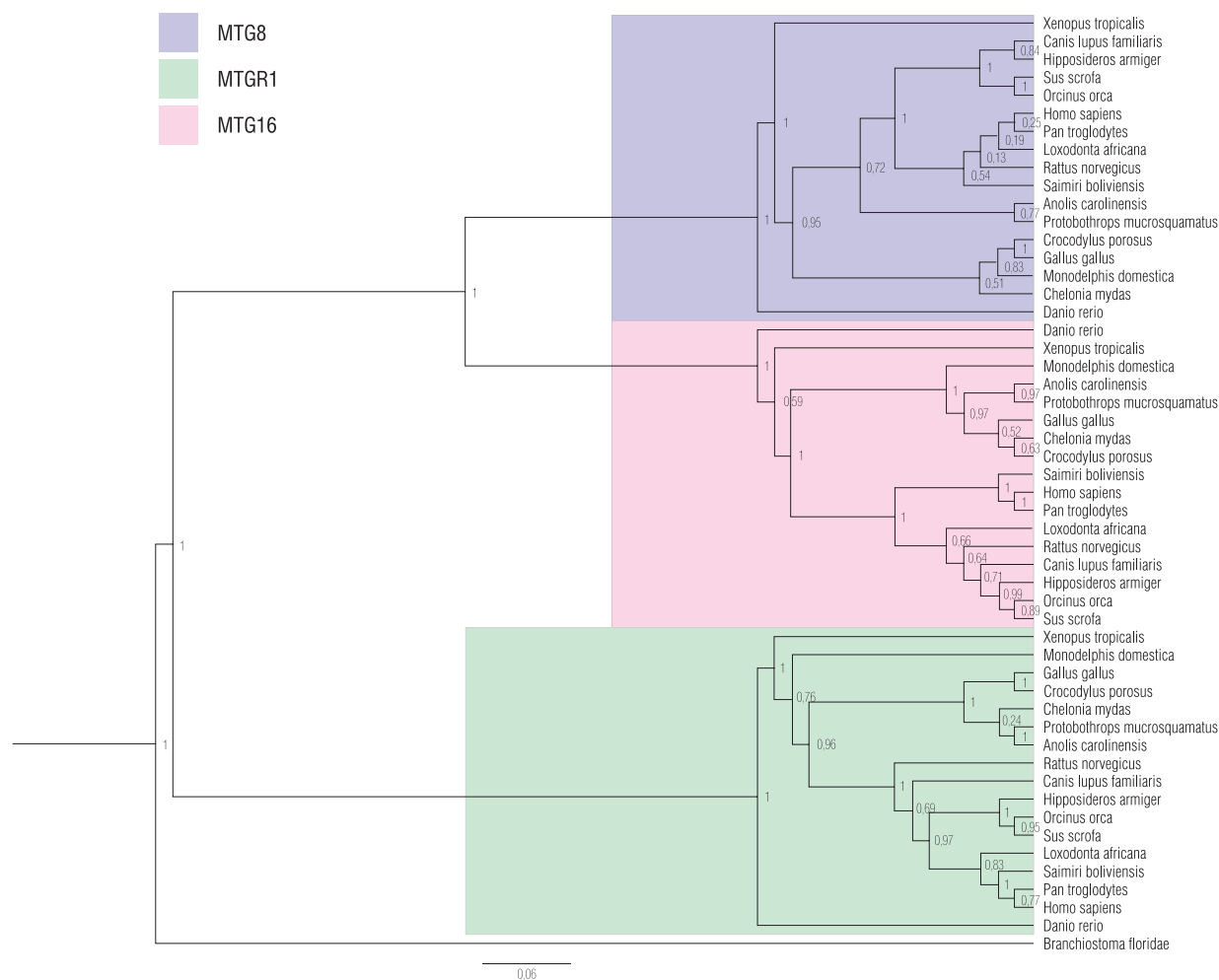


Рис. 4. История дупликаций гена семейства *MTG* у Vertebrata с ланцетником *Branchiostoma belcheri* в качестве внешней группы. Указаны апостериорные вероятности узлов. Шкала соответствует 6 аминокислотным заменам на 100 сайтов

Fig. 4. Duplication history of the *MTG* gene family in the Vertebrata with *Branchiostoma floridae* as an outgroup. Numbers at nodes indicate posterior probability. Scale bar corresponds to 6 amino acid substitutions per 100 sites

включает этот домен в гене *CBFA2T3* латимерии из-за содержащейся в нем крупной делеции, которая сильно нарушает топологию дерева.

Полученные результаты позволяют предположить действие движущего отбора по домену NHR1 (TAFH) в гене *CBFA2T2* до начала дивергенции позвоночных и в несколько меньшей степени по этому же домену в гене *CBFA2T1* (см. рис. 5).

Второй домен не претерпевал значимых замен и эволюционировал довольно стабильно, хотя большая длина ветви к предшественнику генов *MTG8* и *MTG16* свидетельствует о слабом движущем отборе. В целом, домен NHR2 эволюционировал с более высокой скоростью, чем все остальные домены, при этом во всех ветвях дерева эта скорость оставалась более или менее равномерной. Это может указывать на то, что домен NHR2 в меньшей степени подвергся действию стабилизирующего отбора, вероятно, из-за того, что способность к олигомеризации зависит скорее от общих физико-химических свойств включенных аминокислот, чем от наличия конкретных последовательностей. У ланцетника домен NHR2 претерпел масштабные перестройки, поэтому он был исключен из набора последовательностей для данного домена.

В эволюции домена NHR3 интерес представляет значительное изменение аминокислотной последовательности в ветви, предшествовавшей дивергенции плацентарных млекопитающих. Наблюдаемое далее резкое замедление эволюции, очевидно, указывает скорее на включение стабилизирующего отбора после некоторого эпизода движущего отбора, чем на ослабление стабилизирующего отбора в связи с потерей функции.

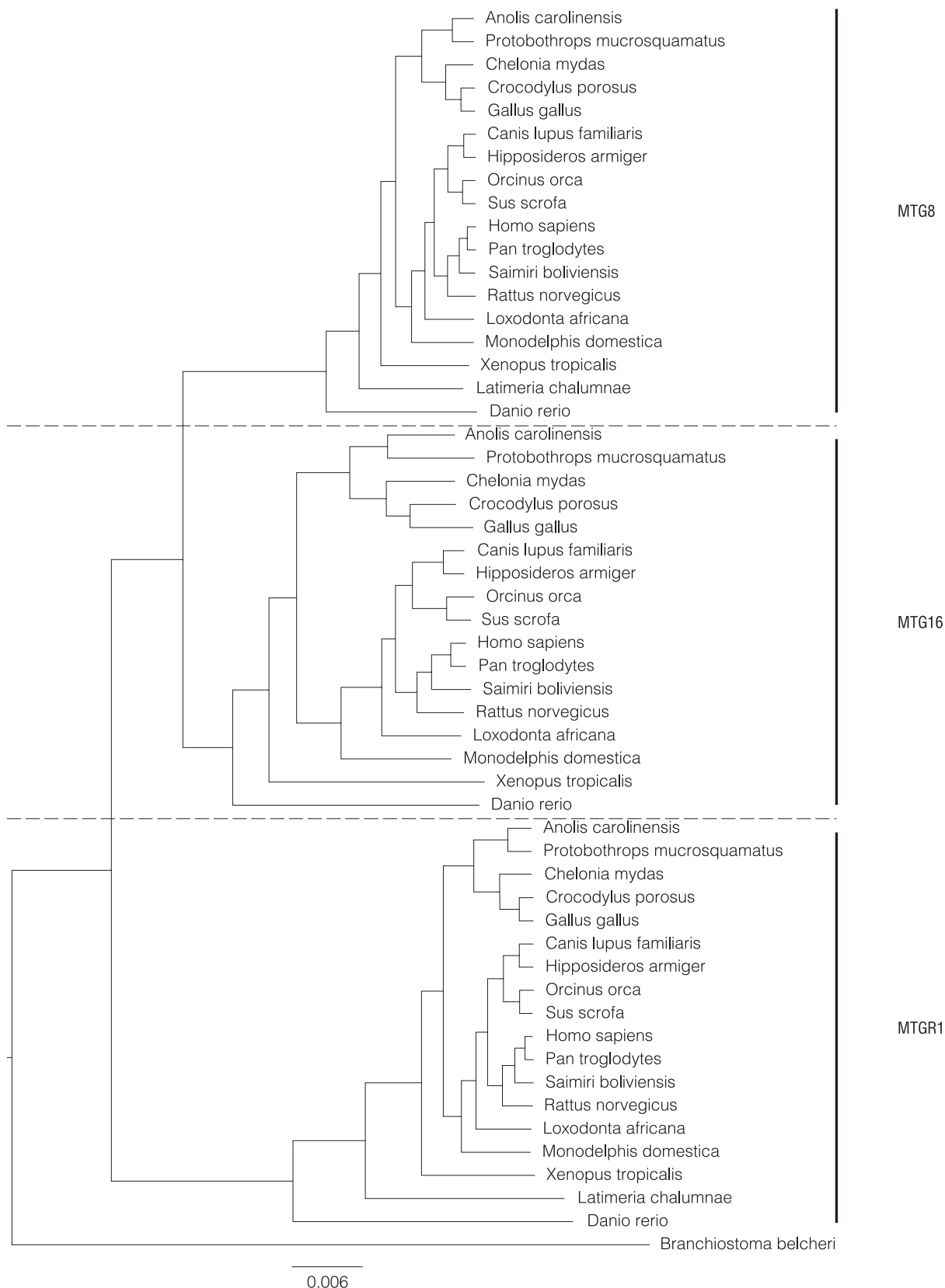


Рис. 5. Эволюция домена TAFH генов семейства *MTG* позвоночных с ланцетником *Branchiostoma belcheri* в качестве внешней группы. Длина ветвей пропорциональна количеству аминокислотных замен, шкала соответствует 6 аминокислотным заменам на 1000 сайтов

Fig. 5. Evolution of the TAFH domain of vertebrate *MTG* family genes with *Branchiostoma belcheri* as an outgroup. Branch length are proportional to the number of amino acid substitutions, scale bar corresponds to 6 amino acid substitutions per 1000 sites

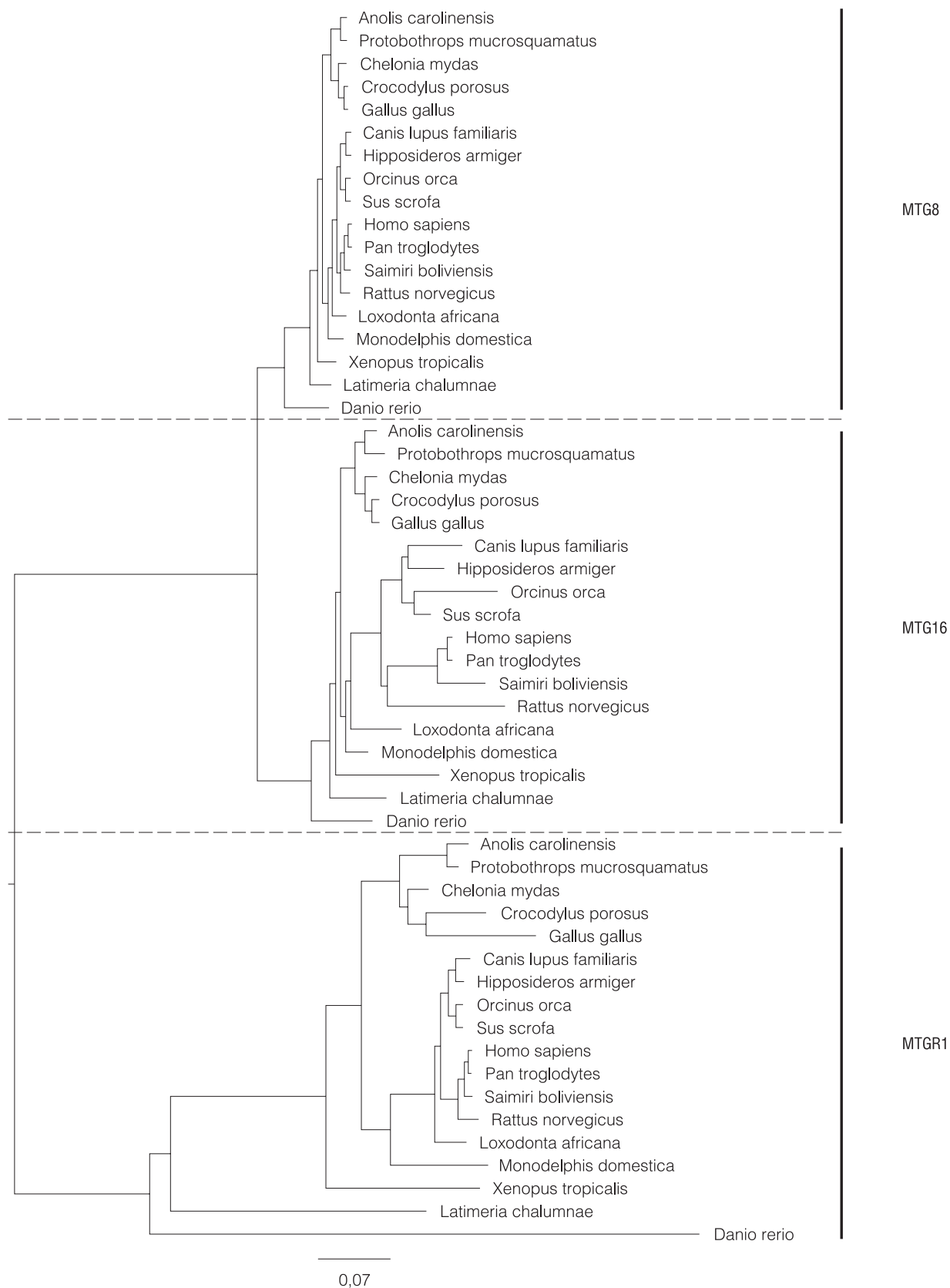


Рис. 6. Эволюция домена NHR2 генов семейства *MTG* позвоночных с ланцетником *Branchyostoma belcheri* в качестве внешней группы. Длина ветвей пропорциональна количеству аминокислотных замен, шкала соответствует 7 аминокислотным заменам на 100 сайтов

Fig. 6. Evolution of the NHR2 domain of vertebrate *MTG* family genes with *Branchyostoma belcheri* as an outgroup. Branch length are proportional to the number of amino acid substitutions, scale bar corresponds to 7 substitutions per 100 sites

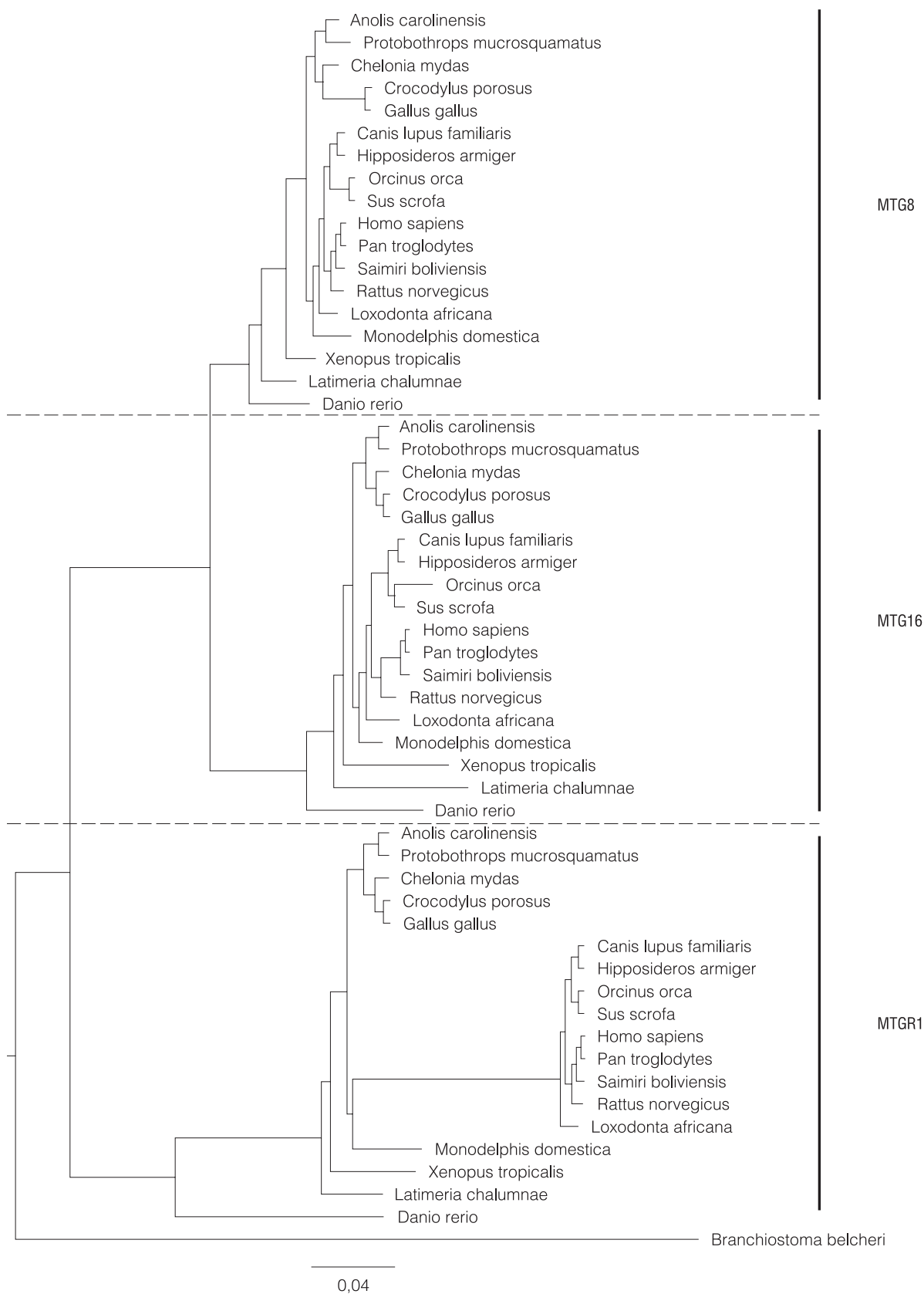


Рис. 7. Эволюция домена NHR3 генов семейства *MTG* позвоночных с ланцетником *Branchiostoma belcheri* в качестве внешней группы. Длина ветвей пропорциональна количеству аминокислотных замен, шкала соответствует 4 аминокислотным заменам на 100 сайтов

Fig. 7. Evolution of the NHR3 domain of vertebrate *MTG* family genes with *Branchiostoma belcheri* as an outgroup. Branch length are proportional to the number of amino acid substitutions, scale bar corresponds to 4 substitutions per 100 sites

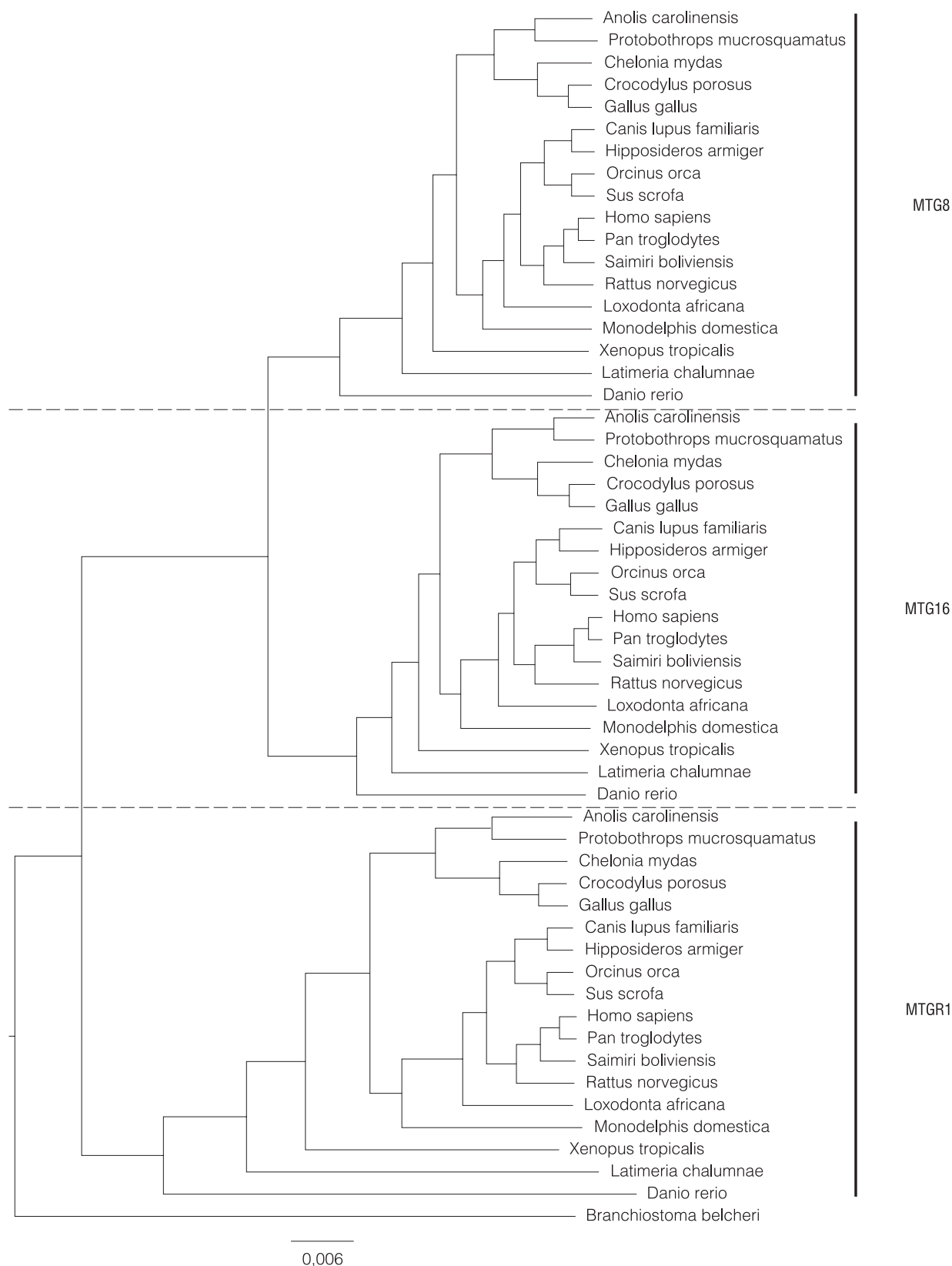


Рис. 8. Эволюция домена NHR4 (zf-MYND) генов семейства *MTG* позвоночных с ланцетником *Branchiostoma belcheri* в качестве внешней группы. Длина ветвей пропорциональна количеству аминокислотных замен, шкала соответствует 6 аминокислотным заменам на 1000 сайтов

Fig. 8. Evolution of the NHR4 (zf-MYND) domain of vertebrate *MTG* family genes with *Branchiostoma belcheri* as an outgroup. Branch length are proportional to the number of amino acid substitutions, scale bar corresponds to 6 substitutions per 1000 sites

Наконец, в эволюции домена NHR4 наблюдается существенное ускорение эволюции в ветви, которая дала ген-родоначальник для *CBFA2T1* и *CBFA2T3*. Таким образом, можно предположить, что движущий отбор по данному домену действовал в промежутке между первой и второй дубликациями, сформировавшими данное генное семейство. Одновременно этот домен сохранял крайне высокую консервативность для каждого из членов семейства, что указывает на важность выполняемых им функций.

Следует отметить, что окончательное заключение о влиянии отбора на белковые последовательности не может быть сделано исключительно на основании анализа скоростей эволюции. В дальнейшем планируется провести дополнительную оценку пропорции синонимичных и несинонимичных замен, что не только даст более ясное представление о влиянии отбора на разные участки белка, но и позволит предсказать наиболее критичные сайты, определяющие функциональность каждого домена. В сочетании с анализом паттернов коэволюции это даст возможность уточнить список белков-партнеров для генов семейства *MTG* и механизмы реализации их регуляторной функции.

Заключение. Первый ген, кодирующий характерный для семейства *MTG* кор доменов TAFH+zf-MYND, появился у предков первых многоклеточных – Placozoa, Stenophora и Porifera. Его формирование происходило из последовательностей, кодирующих TAFH-домен транскрипционного фактора TFIIID и высококонсервативный древний домен zf-MYND, который у эукариот входит в состав гена *DEAF-1*. Объединение этих двух доменов инициировало дальнейшее насыщение последовательности мутациями и формирование доменов NHR2 и NHR3, участвующих в гомо- и гетероолигомеризации и рекрутировании корепрессоров и гистоновых деацетилаз. Несмотря на то что обычно домены цинковых пальцев участвуют в связывании с ДНК, домен zf-MYND белков семейства *MTG* выступает «платформой» для белок-белковых взаимодействий между многочисленными белками репрессирующего комплекса. У некоторых организмов домен NHR3 не имеет гомологии с ортологами других организмов, однако программа-анализатор находит его по сохраняющемуся паттерну гидрофобных и полярных остатков, необходимых для формирования структуры суперскрученных α -спиралей. Из-за не очень хорошего прочтения и сборки генов базальных Metazoa, а также из-за их древнего происхождения невозможно установить точную филогению и, соответственно, определить, в какой последовательности появлялись домены в белках *MTG*, однако вероятнее всего, что ген, кодирующий кор белков семейства, появился у общего предка всех Metazoa. Он сформировал домен NHR3, вслед за чем произошли линиеспецифические перестройки и потеря этого домена домена у гребневика, а также мутационное насыщение в этом же домене у трихоплакса, из-за чего стало невозможным определить гомологию. У более совершенных Bilateria стал формироваться домен NHR2, после чего произошла окончательная фиксация относительных границ экзонов и доменов. В линии Protostomia ген остался в единственном числе. У нескольких хелицеровых удалось найти по два гена семейства, которые являются паралогами: у клещей *Galendromus occidentalis* и *Varroa destructor*, мечехвоста *Limulus polyphemus* и у паука *Stegodyphus mimosarum*, в базе проходящих под названиями *CBFA2T1-like*, *CBFA2T2-like* и *CBFA2T3-like*. Однако филогенетические построения показали, что эти паралоги не имеют общего происхождения с паралогами позвоночных. В кладе Vertebrata генное семейство остается очень консервативным, несмотря на редкие различия между более древними таксонами.

Кроме того, нами проанализированы скорости эволюции отдельных доменов, а также действие отбора на эти домены. Показано, что в ходе эволюции гены *MTG8* и *MTG16* быстрее эволюционировали по доменам NHR2 и zf-MYND, а ген *MTGR1* – по домену TAFH у всех позвоночных и по домену NHR3 у плацентарных млекопитающих. Анализ скоростей эволюции отдельных доменов дает представление о том, в каком домене происходило накопление специфических сайтов.

Список использованных источников

1. Rossetti, S. The *MTG* proteins: chromatin repression players with a passion for networking / S. Rossetti, A. T. Hoogeveen, N. Sacchi // Genomics. – 2004. – Vol. 84, N 1. – P. 1–9. <https://doi.org/10.1016/j.ygeno.2004.02.011>
2. Associations between morphology, karyotype and clinical features in myeloid leukemias / M. A. Bitter [et al.] // Human Pathol. – 1987. – Vol. 18, N 3. – P. 211–225. [https://doi.org/10.1016/s0046-8177\(87\)80002-3](https://doi.org/10.1016/s0046-8177(87)80002-3)
3. Therapy-related acute myeloid leukemia/myelodysplasia with balanced 21q22 translocations / D. A. Arber [et al.] // Am. J. Clin. Pathol. – 2002. – Vol. 117, N 2. – P. 306–313. <https://doi.org/10.1309/c3g2-cxa0-he9j-tkdr>

4. Myeloid translocation genes differentially regulate colorectal cancer programs / B. Parang [et al.] // *Oncogene*. – 2016. – Vol. 35, N 49. – P. 6341–6349. <https://doi.org/10.1038/onc.2016.167>
5. Identification of homeotic target genes in *Drosophila melanogaster* including *nervy*, a proto-oncogene homologue / P. G. Feinstein [et al.] // *Genetics*. – 1995. – Vol. 140, N 2. – P. 573–586.
6. Alternative splicing targeting the hTAF4-TAFH domain of TAF4 represses proliferation and accelerates chondrogenic differentiation of human mesenchymal stem cells / J. Kazantseva [et al.] // *PLoS ONE*. – 2013. – Vol. 8, N 10. – P. e74799. <https://doi.org/10.1371/journal.pone.0074799>
7. Oligomerization of ETO is obligatory for corepressor interaction / J. Zhang [et al.] // *Mol. Cell. Biol.* – 2001. – Vol. 21, N 1. – P. 156–163. <https://doi.org/10.1128/mcb.21.1.156-163.2001>
8. Cloning, expression, purification and crystallization of NHR3 domain from acute myelogenous leukemia-related protein AML1-ETO / H.-T. Yang [et al.] // *Acta Biochim. Biophys. Sinica*. – 2004. – Vol. 36, N 8. – P. 566–570. <https://doi.org/10.1093/abbs/36.8.566>
9. The MYND motif is required for repression of basal transcription from the multidrug resistance 1 promoter by the t(8;21) fusion protein / B. Lutterbach [et al.] // *Mol. Cell. Biol.* – 1998. – Vol. 18, N 6. – P. 3604–3611. <https://doi.org/10.1128/mcb.18.6.3604>
10. The house spider genome reveals an ancient whole-genome duplication during arachnid evolution / E. E. Schwager [et al.] // *BMC Biology*. – 2017. – Vol. 15, N 1. – Art. 62. <https://doi.org/10.1186/s12915-017-0399-x>

References

1. Rossetti S., Hoogveen A. T., Sacchi N. The *MTG* proteins: chromatin repression players with a passion for networking. *Genomics*, 2004, vol. 84, no. 1, pp. 1–9. <https://doi.org/10.1016/j.ygeno.2004.02.011>
2. Bitter M. A., Le Beau M. M., Rowley J. D., Larson R. A., Golomb H. M., Vardiman J. W. Associations between morphology, karyotype and clinical features in myeloid leukemias. *Human Pathology*, 1987, vol. 18, no. 3, pp. 211–225. [https://doi.org/10.1016/s0046-8177\(87\)80002-3](https://doi.org/10.1016/s0046-8177(87)80002-3)
3. Arber D. A., Slovak M. L., Popplewell L., Bedell V., Ikle D., Rowley J. D. Therapy-related acute myeloid leukemia/myelodysplasia with balanced 21q22 translocations. *American Journal of Clinical Pathology*, 2002, vol. 117, no. 2, pp. 306–313. <https://doi.org/10.1309/c3g2-cxa0-he9j-tkdr>
4. Parang B., Bradley A. M., Mittal M. K., Short S. P., Thompson J. J., Barrett C. W. [et al.]. Myeloid translocation genes differentially regulate colorectal cancer programs. *Oncogene*, 2016, vol. 35, no. 49, pp. 6341–6349. <https://doi.org/10.1038/onc.2016.167>
5. Feinstein P. G., Kornfeld K., Hogness D. S., Mann R. S. Identification of homeotic target genes in *Drosophila melanogaster* including *nervy*, a proto-oncogene homologue. *Genetics*, 1995, vol. 140, no. 2, pp. 573–586.
6. Kazantseva J., Kivil A., Tints K., Kazantseva A., Neuman T., Palm K. Alternative splicing targeting the hTAF4-TAFH domain of TAF4 represses proliferation and accelerates chondrogenic differentiation of human mesenchymal stem cells. *PLoS ONE*, 2013, vol. 8, no. 10, p. e74799. <https://doi.org/10.1371/journal.pone.0074799>
7. Zhang J., Hug B. A., Huang E. Y., Chen C. W., Gelmetti V., Maccarana M., Minucci S., Pelicci P. G., Lazar M. A. Oligomerization of ETO is obligatory for corepressor interaction. *Molecular and Cellular Biology*, 2001, vol. 21, no. 1, pp. 156–163. <https://doi.org/10.1128/mcb.21.1.156-163.2001>
8. Yang H.-T., Wu D.-H., Xue X.-Y., Liang W.-X., Miao X.-Y., Pang H., Chen S.-J. Cloning, expression, purification and crystallization of NHR3 domain from acute myelogenous leukemia-related protein AML1-ETO. *Acta Biochimica et Biophysica Sinica*, 2004, vol. 36, no. 8, pp. 566–570. <https://doi.org/10.1093/abbs/36.8.566>
9. Lutterbach B., Sun D., Schuetz J., Hiebert S. W. The MYND motif is required for repression of basal transcription from the multidrug resistance 1 promoter by the t(8;21) fusion protein. *Molecular and Cellular Biology*, 1998, vol. 18, no. 6, pp. 3604–3611. <https://doi.org/10.1128/mcb.18.6.3604>
10. Schwager E. E., Sharma P. P., Clarke T., Leite D. J., Wierschin T., Pechmann M. [et al.]. The house spider genome reveals an ancient whole-genome duplication during arachnid evolution. *BMC Biology*, 2017, vol. 15, no. 1, art. 62. <https://doi.org/10.1186/s12915-017-0399-x>

Информация об авторах

Ковалевская Анна Игоревна – мл. науч. сотрудник. Институт биоорганической химии НАН Беларуси (ул. Академика В.Ф. Купревича, 5, корп. 2, 220141, г. Минск, Республика Беларусь). E-mail: bio.ovcharovaa@gmail.com

Романовская Татьяна Владимировна – канд. биол. наук, доцент, преподаватель. Белорусский государственный университет (пр. Независимости, 4, 220030, г. Минск, Республика Беларусь). E-mail: romanovskaya.t@mail.ru

Information about the authors

Anna I. Kavaleuskaya – Junior Researcher. Institute of Bioorganic Chemistry of the National Academy of Sciences of Belarus (5/2, Academician V. F. Kuprevich Str., 220141, Minsk, Republic of Belarus). E-mail: bio.ovcharovaa@gmail.com

Tatsiana V. Ramanouskaya – Ph. D. (Biol.), Assistant Professor, Lecturer. Belarusian State University (4, Nezavisimosti Ave., 220030, Minsk, Republic of Belarus). E-mail: romanovskaya.t@mail.ru