

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
БИОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ
Кафедра генетики

ОВЧАРОВА
Анна Игоревна

БИОИНФОРМАТИЧЕСКИЙ АНАЛИЗ ЭВОЛЮЦИИ
СЕМЕЙСТВА ГЕНОВ MTG

Аннотация
к дипломной работе

Научный руководитель:
кандидат биологических
наук,
доцент Т. В. Романовская

Минск, 2018

РЕФЕРАТ

Дипломная работа включает: страниц – 40, рисунка – 17, таблиц – 2, источников – 38.

Ключевые слова: MTG, ОСТРЫЙ МИЕЛОИДНЫЙ ЛЕЙКОЗ, ГИБРИДНЫЙ ГЕН AML1/ETO, ЭВОЛЮЦИЯ, ФИЛОГЕНЕТИЧЕСКОЕ ДРЕВО, БИОИНФОРМАТИЧЕСКИЙ АНАЛИЗ, ПОЗВОНОЧНЫЕ.

Объект исследования: последовательности генов MTG и их продуктов у человека и других позвоночных.

Цель: изучение доменной организации продуктов генов семейства MTG у позвоночных животных и истории их возникновения в ходе эволюции, анализ скоростей эволюции и действия отбора на отдельные домены и аминокислотные позиции.

Методы: специализированные инструменты биоинформационного анализа.

Было определено, что родоначальный ген сформировался у общего предка всех Metazoa из последовательностей, кодирующих TAFH-домен TAF4-субъединицы транскрипционного фактора TFIID и высококонсервативный домен цинковых пальцев zf-MYND из гена DEAF1. Окончательное формирование полного гена и фиксация границ доменов относительно экзонов произошли у общего предка Cnidaria и Bilateria. После разделения Bilateria на Proto- и Deuterostomia в линии Vertebrata произошли две последовательные полногеномные дупликации, приведшие к образованию исследуемого семейства. Эволюционный анализ скоростей эволюции по каждому домену выявил действие движущего отбора на последовательности домена TAFH гена MTGR1 у всех Vertebrata, домена NHR4 у предкового для MTG8 и MTG16 гена, и домена NHR3 гена MTGR1 у плацентарных млекопитающих. Анализ действия отбора на отдельные кодоны методом Ka/Ks выявил одну позицию в составе домена NHR3 гена MTG8, которая может или могла подвергаться движущему отбору, однако дальнейший биоинформационный анализ показал отсутствие статистически значимых отличий по скорости накопления несинонимичных замен в этой позиции. Были определены позиции внутри каждого домена, консервативные внутри одного гена, но отличающиеся у двух других. Эти позиции вместе с рассчитанными коэффициентами Ka/Ks могут помочь в обнаружении сайтов связывания специфических белков-партнеров.

РЕФЕРАТ

Дыпломная праца ўключае: старонак - 43, малюнка - 17, табліц - 2, крыніц - 38.

Ключавыя словы: MTG, ВОСТРЫ МІЕЛОІДНЫ ЛЕЙКОЗ, ГІБРЫДНЫ ГЕН AML1/ETO, ЭВАЛЮЦЫЯ, ФЛАГЕНЕТЫЧНАЕ ДРЭВА, БІЯІНФАРМАЦЫЙНЫ АНАЛІЗ, ПАЗВАНОЧНЫЯ.

Аб'ект даследавання: паслядоўнасці генаў MTG і іх прадуктаў у чалавека і іншых пазваночных.

Мэта: вывучэнне даменнай арганізацыі прадуктаў генаў сямейства MTG у пазваночных жывёл і гісторыі іх узнікнення ў ходзе эвалюцыі, аналіз хуткасцяў эвалюцыі і дзеянні адбору на асобныя дамены і амінакіслотныя пазіцыі.

Метады: спецыялізаваныя інструменты біяінфармацыйнага аналізу.

Было вызначана, што продкавы ген сфармаваўся ў агульнага продка ўсіх Metazoa з паслядоўнасцяў, якія кадзіруюць TAFH-дамен TAF4-субадзінкі транскрыпцыйнага фактара TFIID і высока кансерватыўны дамен цынкавых пальцаў zf-MYND з гена DEAF1. Канчатковае фарміраванне поўнага гена і фіксацыя межаў даменаў адносна экзонаў адбыліся ў агульнага продка Cnidaria і Bilateria. Пасля падзелу Bilateria на Proto- і Deuterostomia ў лініі Vertebrata адбыліся дзве паслядоўныя поўнагенамыя дуплікацыі, якія прывялі да ўтварэння доследнага сямейства. Пры параўнальным аналізе трох паралогаў *Homo sapiens* і іх адзінага гамолага ў *Branchiostoma belcheri* было выяўлена, што гены MTG маюць кансерватыўную экзон-інтронную структуру і падобныя мяжы бялковых даменаў. Акрамя таго, было высветлена, што і паралагі, і артолагі сямейства MTG варыябельныя на N-канцах, якія не ўваходзяць у склад даменаў. Эвалюцыйны аналіз хуткасцяў эвалюцыі па кожным дамену выявіў дзеянне рухаючага адбору на паслядоўнасці дамена TAFH гена MTGR1 ва ўсіх Vertebrata, дамена NHR4 у продкаў для MTG8 і MTG16 генаў, і дамена NHR3 гена MTGR1 ў плацэнтарных млекакормячых. Аналіз дзеяння адбору на асобныя кадоны метадам Ka/Ks выявіў адну пазіцыю ў дамене NHR3 гена MTG8, якая можа ці магла падвяргацца рухаючаму адбору, аднак далейшы біяінфармацыйны аналіз паказаў адсутнасць статыстычна значных адрозненняў па хуткасці назапашвання несінанімічных замен ў гэтай пазіцыі. Былі вызначаны пазіцыі у кожным дамене, кансерватыўныя ўнутры аднаго гена, якія але адрозніваюцца у двух іншых генаў. Гэтыя пазіцыі разам з разлічанымі каэфіцыентамі Ka/Ks могуць дапамагчы ў выяўленні сайтаў звязвання спецыфічных бялкоў-партнёраў.

ABSTRACT

The graduation project includes: pages – 43, figure –17, tables – 2, sources – 38.

Key words: MTG, ACUTE MYELOID LEUKEMIA, HYBRID GENE AML1/ETO, EVOLUTION, PHYLOGENETIC TREE, BIOINFORMATIC ANALYSIS, VERTEBRATES.

Object of investigation: sequences of MTG genes and their products in humans and other vertebrates.

Aim of work: to investigate the domain organization of products of MTG family genes in vertebrates and the history of their origin during evolution, analysis of evolutionary rates and selection pressure on individual domains and amino acid positions.

Methods: specialized tools of bioinformatics analysis.

It was shown that the ancestral gene first appeared in the last common ancestor of basal Metazoans from the sequences encoding the TAFH domain of the TAF4 subunit of the TFIID transcription factor and the highly conserved zf-MYND zinc finger domain of the DEAF1 gene. The resulting formation of a complete gene and fixation of domain boundaries towards exons occurred in the last common ancestor of Cnidaria and Bilateria. The separation of Bilateria into Proto- and Deuterostomia were followed by two whole genome duplications occurred in Vertebrata clade, leading to the formation of MTG gene family. In a comparative analysis of three paralogs of *Homo sapiens* and their single homolog in *Branchiostoma belcheri* it was found that MTG genes have a conservative exon-intron structure and similar boundaries of protein domains. In addition, it was found that both the paralogs and the orthologs of the MTG family are variable at the N-terminus, which is not part of the domain structure. Evolutionary analysis of evolution rates for each domain revealed the effect of driving selection on the TAFH domain of MTG8 gene in all Vertebrata, the NHR4 domain in the ancestral MTG8 and MTG16 genes, and the NHR3 domain of the MTGR1 gene in placental mammals. The analysis selection pressure on individual codons using the Ka/Ks method revealed one position in the composition of the NHR3 domain of the MTG8 gene, which may or might have been subjected to the driving selection, however further bioinformatic analysis showed no statistically significant differences in the rate of accumulation of non-synonymous substitutions in this position. Amino acid positions within individual domains that have shown conservatism within one of three paralogs were defined. These positions together with the calculated coefficients Ka/Ks can help in the detection of binding sites for specific protein partners.