

***L31*-ТРАНСПОЗОНЫ ШЕСТИЛУЧЕВЫХ КОРАЛЛОВ (*Hexacorallia*): РАСПРОСТРАНЕНИЕ, РАЗНООБРАЗИЕ И ЭВОЛЮЦИЯ**

© 2024 г. Л. В. Пузакова¹, М. В. Пузаков¹ *, П. М. Пузакова²

¹Федеральный исследовательский центр “Институт биологии южных морей
им. А.О. Ковалевского” Российской академии наук, Севастополь, 299011 Россия

²Филиал Московского государственного университета им. М.В. Ломоносова,
в г. Севастополе, Севастополь, 299001 Россия

*e-mail: puzakov.mikh@yandex.ru

Поступила в редакцию 09.11.2023 г.

После доработки 23.12.2023 г.

Принята к публикации 28.12.2023 г.

Мобильные генетические элементы (МГЭ) эукариот — ретротранспозоны и ДНК-транспозоны — это нуклеотидные последовательности, способные перемещаться из локуса в локус генома, а также между геномами различных организмов. ДНК-транспозоны *L31* являются древней и разнообразной группой, входящей в большую группу *IS630/Tc1/mariner*. *L31*-транспозоны не получили широкого распространения и присутствуют у ограниченного числа таксонов. Помимо последовательности, кодирующей DDE/D-транспозазу, *L31*-транспозоны несут еще одну ОРС (ОРС2). Детальный анализ элементов *L31* в геномах шестилучевых кораллов позволил получить подробную информацию о распространении, разнообразии и структуре элементов. Были выявлены две большие группы, *L31-duo* и *L31-uno*, отличающиеся и по паттерну каталитического домена и по структуре. В результате реконструкции эволюции *L31*-транспозонов было предположено, что шестилучевые кораллы получили *L31*-транспозоны от двустворчатых моллюсков. При этом вышеперечисленная группа *L31-uno*, возможно, была получена моллюсками в результате горизонтального переноса уже от кораллов. Исследования распространения и разнообразия МГЭ у морских беспозвоночных будут способствовать лучшему пониманию процессов эволюции МГЭ и их роли в эволюционной истории видов.

Ключевые слова: ДНК-транспозоны, *L31*-транспозоны, коралловые полипы, эволюция геномов, биоразнообразие.

DOI: 10.31857/S0016675824060027 **EDN:** BYEJMV

Мобильные генетические элементы (МГЭ) эукариот — ретротранспозоны и ДНК-транспозоны — это нуклеотидные последовательности, структура которых обеспечивает возможность их перемещений из локуса в локус генома, и даже между геномами различных организмов [1, 2]. Функциональные ретротранспозоны несут гены, кодирующие ферменты, которые обеспечивают создание РНК-копии МГЭ (посредника) и последующий синтез комплементарной ДНК и ее интеграцию в геном. Такой механизм перемещения принято называть “копирование—вставка”. Ретротранспозоны можно разделить на три основных подкласса в соответствии с их механизмом репликации и интеграции: 1) элементы с длинными концевыми повторами (*LTR*), 2) элементы без длинных концевых повторов (*non-LTR*) и 3) элементы, использующие тирозин-рекомбиназу (*YR*) [1–4].

ДНК-транспозоны не создают себе копии-посредника и вырезаются из прежнего локуса геномной последовательности и встраиваются в новый. Такой способ перемещения называется (но не для всех МГЭ этого класса) “вырезание—вставка”. В настоящее время известно о четырех основных группах ДНК-транспозонов: 1) элементы, использующие DDE/D-транспозазу (DDE/D-транспозоны), 2) элементы, использующие тирозин-рекомбиназу (*Cryptons*), 3) элементы “катыщегося круга” (*Helitrons*) и 4) “самосинтезирующие” транспозоны (*Mavericks* или *Polintons*) [1–4].

DDE/D-транспозоны являются широко распространенными и наиболее разнообразными из всех МГЭ [2]. Они включают в себя около двух десятков надсемейств [4–10]. Распространение этой группы МГЭ достигло таких масштабов, что ген DDE/D-транспозазы является претендентом на

титул самого старого и наиболее распространеного гена на земле [11]. Надсемейство DDE/D-транспозонов *L31* является достаточно древней и разнообразной группой, входящей в объединение надсемейств (инфракласс) *IS630/Tc1/mariner* [9]. Первые представители надсемейства *L31* DDE/D-транспозонов были выявлены при исследовании генома тихоокеанской устрицы *Crassostrea gigas* [12]. Было установлено, что *L31*-транспозоны не получили широкого распространения и присутствуют только у двустворчатых моллюсков (*Bivalvia*) и шестилучевых кораллов (*Hexacorallia*). Впоследствии появились данные, что *L31*-транспозоны присутствуют еще и у представителей нематод (*Nematoda*) [13]. Детальный анализ геномов двустворчатых моллюсков показал, что *L31*-транспозоны имеются только в подклассе *Autobranchia*, с преобладанием по количеству и разнообразию в инфраклассе *Pteriomorphia* [9]. В надсемействе транспозонов *L31* моллюсков было выделено пять филогенетических кластеров, включающих в себя элементы со значительными вариациями в структуре.

Элементы *L31*, как и все DDE/D-транспозоны, имеют концевые инвертированные повторы (КИП) и открытую рамку считывания (ОРС), кодирующую фермент транспозазу (рис 1). Транспозаза содержит ДНК-связывающий (PAIRED) и каталитический (DDE/D) домены. Домен PAIRED представлен двумя триадами α -спиралей, расположенными в первой половине аминокислотной последовательности и между которыми имеется GRPR-подобный мотив (так называемый ТА-крюк). Во второй половине последовательности транспозазы локализован DDE/D-домен, паттерн которого является одной из классификационных черт DDE/D-транспозонов и отвечает за разрезание и перемещение цепей ДНК в процессе внедрения. Этот домен имеет три характерных аминокислотных остатка — два аспартата (D) и один глутамат (E) или аспаргат (D) в качестве третьего, отсюда и название этой триады — мотив или домен DDE/D. На основе различного количества аминокислотных остатков между вторым аспаргатом и третьим глутаматом/аспаргатом в доменах DDE/D суперсемейство *IS630/Tc1/mariner* классифицируется на группы. В частности, *L31*-транспозоны имеют паттерн DD37E. Домен PAIRED связывается с КИП, ТА-крюк опосредует связывание домена PAIRED с ДНК-мишенью, а DDE/D-домен, обладающий эндонуклеазной и лигирующей активностью, обеспечивает как эксцизии, так и инсерции элемента [14, 15]. *L31*-транспозоны несут еще одну ОРС (ОРС2), что встречается достаточно редко у DDE/D-транспозонов [16–19]. Среди предполагаемых функций белка ОРС2 рассматривались такие, как возможность участия в регуляции транскрипции транспозазы или в обеспечении транспорта транспозазы через ядерный поровый комплекс или даже в

транспорте нуклеотидной последовательности самого *L31*-транспозона через клеточную мембрану [9].

Исследование представленности МГЭ у различных таксонов способствует большему пониманию как молекулярной эволюции геномов, так и эволюционной истории видов. Данная работа посвящена изучению элементов надсемейства *L31* у шестилучевых кораллов. *L31*-транспозоны были обнаружены только в двух исследованных отрядах — *Actiniaria* и *Scleractinia* подкласса *Hexacorallia*, однако внутри этих таксонов данная группа представлена множеством филогенетических групп.

МАТЕРИАЛЫ И МЕТОДЫ

Поиск последовательностей *L31*-транспозонов

Для поиска *L31*-транспозонов применяли *tblastn* со стандартными настройками (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) [20]. В качестве образца использовали аминокислотные последовательности транспозазы и белка ОРС2 (доп. материалы 1), входящих в структуру элемента *Mariner-31_CGi* устрицы *Crassostrea gigas*. Полногеномные нуклеотидные последовательности шестилучевых кораллов были взяты из базы данных NCBI (National Center for Biotechnology Information) Assembly (<https://www.ncbi.nlm.nih.gov/assembly/>). Поиск полноразмерных МГЭ осуществляли на основе гомологии к образцу. Последовательности с процентом соответствия образцу по длине (Query Coverage) ниже 45 при анализе исключались. На следующем этапе учитывали только последовательности, длина транспозазы которых превышала 266 а.о. Затем выбирали последовательности, наиболее протяженные и с наибольшим сходством. Отобранные гомологичные образцу последовательности были взяты вместе с фланкирующими зонами с общей протяженностью 3000 пн.

Если сходство (по результатам анализа в *blastn*) между нуклеотидными последовательностями выбранных гомологичных фрагментов было ниже 65% при значении статистического параметра *E* более $1e-20$, то такие фрагменты (элементы) называли уникальными. Копиями уникального элемента считали фрагменты, идентичность нуклеотидной последовательности с последовательностью уникального элемента которых была 65% и выше при значении статистического параметра *E* равного $1e-20$ и меньше. При оценке количества копий учитывались только элементы, протяженность которых выше 10% от полноразмерного *L31*-транспозона. Для поиска условно потенциально-функциональных транспозаз для каждого обнаруженного элемента мы брали в качестве образца аминокислотную последовательность

неповрежденной транспозазы этого элемента и осуществляли поиск в tblastn с настройками фильтров: Query Coverage 99–100 и Percent Identity 90–100. Среди результатов поиска при подсчете учитывали только те транспозазы, идентичность которых составляет не менее 90%, а соответствие образцу по длине не менее 99% (Query Coverage), с неповрежденной OPC со стартовым кодоном и стоп-кодоном.

Анализ структуры последовательностей

КИП обнаруживали с помощью blastn путем анализа нуклеотидных последовательностей [20]. Границы гипотетических OPC определяли с помощью ORF Finder (<https://www.ncbi.nlm.nih.gov/orffinder/>) и далее уточняли визуально. Для анализа доменной структуры последовательностей выборочно было взято по три элемента из каждого кластера (L31-1 – L31-9). Для анализа выбирали условно потенциально-функциональные элементы. Расположение GRPR-подобного мотива и маркерных аминокислотных оснований: аспартат (D), аспарат (D) и глутамат (E) каталитического домена идентифицировали визуально по гомологии. ДНК-связывающий домен PAIRED выявляли, анализируя вторичную структуру транспозазы, предсказанную с помощью программы PSIPRED v4.0 [21]. Последовательность сигнала ядерной локализации (NLS) определяли с помощью программы PSORT (<https://www.genscript.com/psort.html>). Графическое представление обобщенных последовательностей фрагмента области второго аспартата каталитического домена было сгенерировано с помощью WebLogo [22].

Филогенетический анализ

Филогенетическое дерево было рассчитано с использованием метода максимального

правдоподобия в программе IQ-TREE [Nguyen L.T., 2015] со сверхбыстрым бутстреп-анализом (UFBoot) (1000 повторов) [23], а модель LG+F+I+G4 была выбрана с помощью ModelFinder [24]. Множественное выравнивание было выполнено с помощью MAFFT с применением метода G-INS-I [25]. В филогенетический анализ были включены все выявленные элементы кораллов, L31-транспозазоны моллюсков, обнаруженные ранее [9], а также представители надсемейств *Gambol*, *pogo*, *Tc1/mariner*, *Sailor*, *HvSm*, *Tec*, *TBE* и в качестве внешней группы – элементы IS630. (доп. материалы 1).

РЕЗУЛЬТАТЫ

Представленность L31-транспозонов у шестилучевых кораллов

В ходе исследования были изучены полногеномные последовательности стрекающих подкласса Hexacorallia. Из пяти отрядов шестилучевых кораллов (Actiniaria, Antipatharia, Corallimorpharia, Scleractinia, Zoantharia) в геномных коллекциях NCBI были представлены (на июль 2022 г.) только два – Actiniaria и Scleractinia. Были изучены 12 видов (6 семейств) отряда Actiniaria и 30 видов (7 семейств) – Scleractinia (доп. материалы 2). Элементы надсемейства L31 были обнаружены во всех исследованных сборках, однако у трех видов *Nematostella vectensis* (Actiniaria), *Stylophora pistillata* и *Orbicella faveolata* (Scleractinia) были обнаружены лишь короткие фрагменты (доп. материалы 2), наличие которых оценивалось как отсутствие (утрата) элемента. Всего было описано 172 уникальных L31-транспозона, при этом количество уникальных элементов у видов варьировало от нуля (*Nematostella vectensis*, *Stylophora pistillata*, *Orbicella faveolata*) до семи (*Diadumene lineata*, *Acropora microphthalma*, *Acr. yongei*, *Montipora efflorescens*) (доп. материалы 2). Количество копий каждого уникального L31-транспозона было невысоким и превышало 20 лишь в нескольких случаях: L31-4_EPal (29), L31-2_PSpe (24), L31-1_MCap (24), L31-4_MCac (26) и L31-4_CGra (31). У двух элементов количество копий превысило 40 – L31-6_AYon (43) и L31-5_DLin (49), а у L31-5_CGra достигло 96 (доп. материалы 2).

Разнообразие L31-транспозонов стрекающих

Для оценки разнообразия L31-транспозонов был проведен филогенетический анализ, в который были включены все выявленные элементы кораллов, представители надсемейства L31 моллюсков, обнаруженные ранее [9], а также представители надсемейств *Gambol*, *pogo*, *Tc1/mariner*, *Sailor*, *HvSm*, *Tec*, *TBE* и в качестве внешней группы

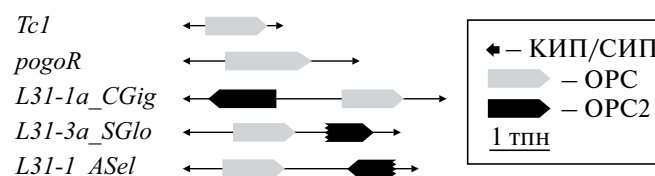


Рис. 1. Структура транспозонов суперсемейства L31. 1 тпн—1000 пн; КИП/СИП — концевые инвертированные повторы или субконцевые инвертированные повторы; OPC — открытая рамка считывания, кодирующая транспозазу; OPC2 — открытая рамка считывания, кодирующая белок с неизвестной функцией.

— элементы *IS630* (доп. материалы 1). Преобладающее большинство изучаемых элементов стрекающих сформировало единую кладу с *L31*-транспозонами моллюсков (рис. 2). Однако внутри сформировались две ярко выраженные подгруппы. Первая (названная *L31-duo*) включала элементы с паттерном каталитического домена транспозазы DD37E, вторая (названная *L31-uno*) объединяла транспозазы с паттерном DD38E. Обе подгруппы включали элементы и кораллов, и моллюсков. Практически все элементы группы *L31-duo* имели фрагменты OPC2, тогда как из 39 элементов *L31-uno* только

L31-5_DLin имел фрагмент OPC2. Три элемента стрекающих с паттерном DD34E сформировали отдельную кладу (*L31-like*), далекую от остальных *L31*-транспозонов (рис. 2). Эти элементы, также как и *L31-uno* (за одним исключением), не имели OPC2. С классификацией и принадлежностью данных элементов предстоит разобраться в последующих исследованиях.

На основании филогенетического анализа, а также различий в консервативных локусах аминокислотной последовательности транспозаз среди

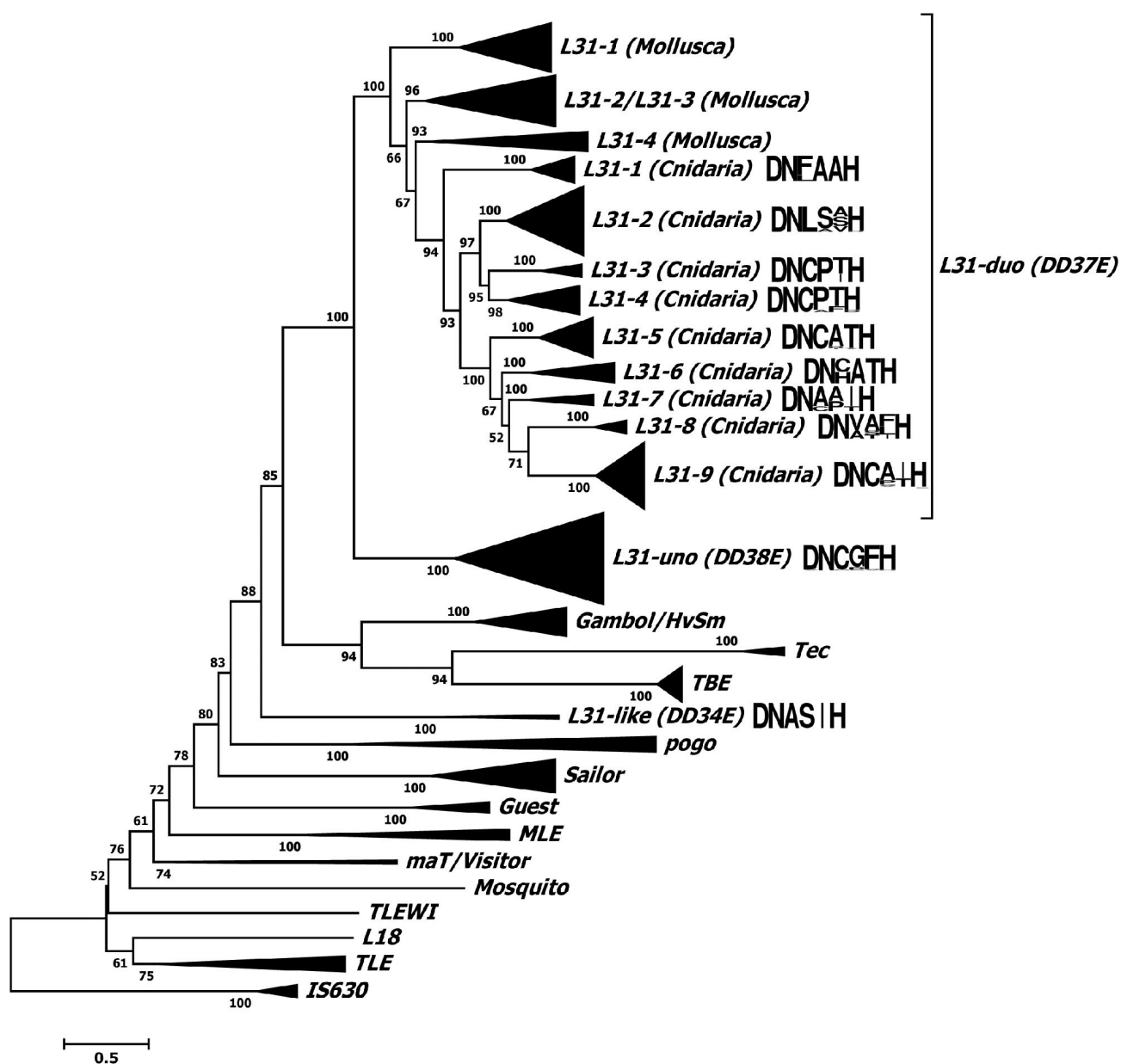


Рис. 2. Филогенетическое разнообразие *L31*-транспозонов. Графическое представление обобщенных последовательностей фрагмента области второго аспартата каталитического домена транспозазы шестиручковых кораллов указаны справа от дендрограммы. Бутстреп-значения менее 50% на дендрограмме не указаны.

L31-2 *AMur* --MSKVNKCGRTYNPGEAIGESLRSASIDRMLLDGGDPATGFFGGFRKFNIGDCEGVSAPFVSKLWKTFCIQGDHMPQ-KRSS**CNPS**HLKPEDVQVLEFLKKEKESATLAS
 L31-2 *AEch* --MSKVNKCGRTYNPGEAIGESLRSASIDRMLLDGGDPATGFFGGFRKFNIGDCEGVSAPFVSKLWKTFCIQGDHMPQ-KRSS**CNPS**HLKPEDVQVLEFLKKEKESATLAS
 L31-2 *PSin* --MAYSVNRKGRSLIRGADISLGRGSIIDSILLEGDDPASGFFPGFSGYSDVADRFVSNQFVSKLWQNFCTTGEHLS-KKK**CNPS**HLKPELMFLKKEKESATLAS
 L31-3 *ATene* -----MNVYGRGTENQCKTFFGSDLSLIIIDEILRQGGDVSTGYFAGSFRTIGSRKXISGVAVSKIWQTCFCKXGELFPL-HPVR**GHLK**KLEPDLDLIPLMIKSRFSITYKE
 L31-1 *AEqu* -----MNVYGRAYNQKTLGSDLSLIIIDEILRQGGDVSTGYFAGSFRTIGSRKXISGVAVSKIWQTCFCKXGELFPL-HPVR**GHLK**KLEPDLDLIPLMIKSRFSITYKE
 L31-1 *PCru* -----MINTYGRFTTKGKALGKDVSLSLIIIDEILRQGGDVSTGYFAGSFRTIGSRKXISGVAVSKIWQTCFCKXGELFPL-HPVR**GHLK**KLEPDLDLIPLMIKSRFSITYKE
 L31-2 *Avir* ---MEVNSFGRTYARGKVLPEDEFSLIIIDEILRQGGDVSTGYFAGSFRTIGSRKXISGVAVSKIWQTCFCKXGELFPL-HPVR**GHLK**KLEPDLDLIPLMIKSRFSITYKE
 L31-2 *CJar* ---MAHMSMAGRVYERKALNEDLRNIIQDVVEKGGDLVTGYPFGSEFLIANKRTTYNTYVVKIKWQFEBTGSTRY--ESHAAG**SKHL**QDDDLDFIRPLKTRASMATGE
 L31-3 *CGra* --MAEATSKRGRCYINGRQLQDLRRKIVEDIVDGGDVTGFFPGFSATAEARNRVKFDVTRKIKWKEVVEKGDTEIT-KTRAA**GKHI**QDDDLDFIRPLKTRASMATGE
 L31-3 *EPal* ---MIVTKKGRTFHRGKSLGFDVVRTAIVDHMTSKGGDIVTXHPPGSLREVAHFHKVSKTIVVKTRWQCSSETADIDVQ-WKG**CNPP**PHLQSHHLDFTIRGLTIKIPSPMPYK
 L31-1 *PSpe* --MNYCVSKGRILVPGRAIDDFRLVVDHMSNGDVTGYSFGSSVSDADHEKLSRSCVAKLWNGACERATIDPQ-WKG**CNPP**THLQSDHLDLLEALKSAKPSITYNE
 L31-4 *HMag* ---MLSAKGRFTTNGKALDEDRSNIIISIVKGGDHSIGGFFAGNFSGIKNYVSGQTIKKVWKRTQDGEIGPRDKK**SONP**SHLTPTLEFIEVLKHSFSPSPMYK
 L31-1 *SMer* ---MAAFNELGRYIERGKCLNTHIKGDIVDFIITNGGDSITGYPFAKWTGDKFVSGKTAKNVWQKFFVMDGTVSPK-KRIS**CNPP**KLSFGDLQIETMTKIKPSTSSKN
 L31-1 *AMyr* ---MAFNEKGRFYERGGKLSDEFKGOIIDKILETGDDRISGYFPAKWTGDKFVSGKTAKNVWQKFFVMDGTVSPK-KRIS**CNPP**KLSFGDLQIETMTKIKPSTSSKN
 L31-1 *Ahya* ---MAFNKGRFYERGGKLSDEFKGOIIDKILETGDDRISGYFPGKWTGDKFVSGKTAKNVWQKFFVMDGTVSPK-KRIS**CNPP**KLSFGDLQIETMTKIKPSTSSKN
 L31-5 *EPal* ---MINTYGRFTTKGKALGKDVSLSLIIIDEILRQGGDVSTGYFAGSFRTIGSRKXISGVAVSKIWQTCFCKXGELFPL-HPVR**GHLK**KLEPDLDLIPLMIKSRFSITYKE
 L31-3 *AYon* ---MKRLNASGRTYVKGRLSADFRSLIIIDEILRQGGDVSTGYFAGSFRTIGSRKXISGVAVSKIWQTCFCKXGELFPL-HPVR**GHLK**KLEPDLDLIPLMIKSRFSITYKE
 L31-4 *ATene* ---MAVNGQKRTYRKGSPLDPLRLSLIFDEILLNGGHAQTGYPFGSEFLIANKRTTYNTYVVKIKWQFEBTGSTRY--ESHAAG**SKHL**QDDDLDFIRPLKTRASMATGE
 L31-1 *MCap* MEARKVNQFGRSYRPGVALAQDLKFLIIDSILRQGGDRTGYIPRSVTOFARELIRSVSNVTSKVWFRYCEEMITTPK-PK**GLTF**EELKEDDERELIEVLKHSFSPMSLSE
 L31-1 *AMic* MEALKVNFGRYRTYTGIALQDMRSLIIDRLIQEGGDRVTGYIPRSYFSEELKLSVNIITKIKWKFCEELSINPR-AG**GTNW**SKLTGDDLELIEILKIEKPSVSLAE
 L31-4 *MCac* MEALKVNFGRYRTYTGIALQDMRSLIIDRLIQEGGDRVTGYIPRSYFSEELKLSVNIITKIKWKFCEELSINPR-AG**GTNW**SKLTGDDLELIEILKIEKPSVSLAE
 L31-3 *DLin* ---MFKVSKKGRTYCCGRALSBDTRIGIIDIINIAARGNNHVNGVQGFTEVAQLNVSASVSKIWKRCDFHRTISPK-KHSG**GHPS**SLNSGDLQLEVLKHSFSPMSLSE
 L31-5 *ACyt* ---MLTVSKSGRSYIPGKPLSBDRLHLIIIDKIVGGGDLPTMIFPGFVDIANEINVSATVSKIWKNYCEFGSLSP-LKHG**CNPS**RLSNGDLLELIEVLKHSFSPMSLSE
 L31-1 *AAwi* ---MLTVSKSGRSYIPGKPLSBDRLHLIIIDKIVGGGDLPTMIFPGFVDIANEINVSATVSKIWKNYCEFGSLSP-LKHG**CNPS**RLSNGDLLELIEVLKHSFSPMSLSE
 L31-3 *AIint* ---MFSINKYGRNFKPGIPLSMDLRNEVIQMAATR-----PISIEICERYNRSTVHKYTKQYRETNIEPLSRKHIRTSLKSGFSDMLLETIQGGKSTSLKE
 L31-6 *AYon* ---MFSINKYGRNFKPGIPLSKDFRNEVIQMAATR-----PISIEICERYNRSTVHKYTKQYRETNIEPLSRKHIRTSLKSGFSDMLLETIQGGKSTSLKE
 L31-1 *PAus* ---MFSINKYGRNFKPGIPLSMDLRNOV-----KELAQDY---CFSEVGRRIRISKSGSVSKIWKQYNTLGSTAPKLNHVTRVPKCTFODSILLETIMVQARGSSSLKE

 L31-2 *AMur* IKETVENYCNLNGG-TLSAIGNTVRNRLPDG-PFTRKKLTRPKAEKFTFQNLAYCQGLNFIISALPPEKIKFFDEAGVN-SGTGNPVYSSSLKGTKSIEIAA-IPRGVN
 L31-2 *AEch* IKETVENYCNLNGG-TLSAIGNTVRNRLPDG-PFTRKKLTRPKAEKFTFQNLAYCQGLNFIISALPPEKIKFFDEAGVN-SGTGNPVYSSSLKGTKSIEIAA-IPRGVN
 L31-2 *PSin* IKETVENYCNLNGG-TLSAIGNTVRNRLPDG-PFTRKKLTRPKAEKFTFQNLAYCQGLNFIISALPPEKIKFFDEAGVN-SGTGNPVYSSSLKGTKSIEIAA-IPRGVN
 L31-3 *ATene* IKENVDHAPT--AE-VSISTIGRAAREHLEPEG-KMSWKVIRPAAEKFSFNDIAHCQSYINFMSTEDPYRVKFFDEAGFELPDCAKPKYGHSVICTPCVEVMR-NSQTPN
 L31-1 *AEqu* IKENVDHAPT--AG-VSISTIGRAAREHLEPEG-KMSWKVIRPAAEKFSFNDIAHCQSYINFMSTEDPYRVKFFDEAGFELPDCAKPKYGHSVICTPCVEVMR-NSQTPN
 L31-1 *PCru* IKENVDHAGSN--AR-VSVTITIGRAAREHLEPEG-QMTWKKMVRPAAEKFTFQNLAYCQSYINFMSTEDPYRVKFFDEAGFELPDCAKPKYGHSVICTPCVEIQR-NVQTPN
 L31-2 *Avir* LKKNVNFECQLPTG--TSQAQANRNVNRCMGGG-KWSLKMVRPAAEKFTFQNLAYCQSYINFMSTEDPYRVKFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-2 *CJar* VRYVYVNRFCNVACG--TSNAATQIRATRNHITDG-SWTWKRTQPVAFKFSFNDIAHCQSYINFMSTEDPYRVKFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-3 *CGra* LYKHNNHFCNIAAG--TSNTAIRRVLRNNMTDC-KWTWKKLTRPVTEKFTFQNLAYCQSYINFMSTEDPYRVKFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-3 *EPal* IHEAVNHMCPIPAG--TSRSISGKAVQNRHOGG-RWTWKKMRNPAAEKFSFNDIAHCQSYINFMSTEDPYRVKFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-1 *PSpe* IREVINANCAIPSG-TSISAIGRAVQRRLSGG-PWTWKKMRSTFKYKFTFQNLAYCQSYINFMSTEDPYRVKFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-4 *HMag* VHDVAVSQCNVIGN--TSQAQANRNVNRCMGGG-KWSLKMVRPAAEKFTFQNLAYCQSYINFMSTEDPYRVKFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-1 *SMer* LLSHVLYQGNLQGTG-TSISAIGRTVRNSILDG-SWSWKMSRNVHKKFTFQNLAYCQSYINFMSTEDPYRVKFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-1 *AMyr* IREQQLQHGNFPSSG--TSISTINRAIRTSLSG-KWSWKMSRNVHKKFTFQNLAYCQSYINFMSTEDPYRVKFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-1 *Ahya* LTELQQLHGNFPSSG--TSISTINRAIRTSLSG-KWSWKMSRNVHKKFTFQNLAYCQSYINFMSTEDPYRVKFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-5 *EPal* IHEKLEDFGNIPNG--TSYSATSRALKRMLSGKQYSRKKVSPIAQERFTVENMAYTQLFIDYLYSKDFPKLFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-3 *AYon* LHDVINEEFGDIPNG--TSISAISRRLNMLSGLYSKRKISSFAQERFTVENMAYTQIFIDDLHAKNPYKLFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-4 *ATene* LNLNELQYQGNIGG--TSISAISKALCERMFTGKAYSRKKVYVAERFTANVMYQDFLDYMHQVDPFKLFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-1 *MCap* IMEELQQLG--GOEISMAVSATRAKSLPSSGQYSRKKLTVMAMERTPNLFTYQTLFINVSSKDPRLKFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-1 *AMic* IISOLEEMD--GEEVSMATVSRAIKHRLPSG-PYTRKKITKIALERFTQNTNFTYQTFNNTYATKNPRRIKFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-4 *MCac* LDDNLNFPGLPYGTSRSSVGRVANSNLPSC-KFSYKKITKIAQERFTQNLAYCQSYINFMSTEDPYRVKFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-3 *DLin* VMDCLYQGLDLPFGTSTVAVCNVAVRNLPSGKGFTEFKKIAHAQERFTANMAYTQLFIDYLYSKDFPKLFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-5 *ACyt* VMDCLYQGLDLPFGTSTVAVCNVAVRNLPSGKGFTEFKKIAHAQERFTANMAYTQLFIDYLYSKDFPKLFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-1 *AAwi* VMDCLYQGLDLPFGTSTVAVCNVAVRNLPSGKGFTEFKKIAHAQERFTANMAYTQLFIDYLYSKDFPKLFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-3 *AIint* IKESLSNFGD--CGELSLSTLSRHVRK**NLP**SGKDYSRKRLGKAGERFTTHENLVYQTLFDYLYSKDFPKLFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-6 *AYon* MKSELSNFGD--CGELSLSTLSRHVRK**NLP**SGKDYSRKRLGKAGERFTTHENLVYQTLFDYLYSKDFPKLFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN
 L31-1 *PAus* LRDLLIATHGD--CGELSLSTLSRHVRK**NLP**SGKDYSRKRLGKAGERFTTHENLVYQTLFDYLYSKDFPKLFFDEAGFELPDCAKPKYGHSVICTPCVEIMR-NAQSPN

 L31-2 *AMur* ITLNLVLGLEGILYANTLDGASNTFLNFFGEAGQATSAL-GNPAIEGQDGIIMDNCA**TH**HEAGTALQTLWMDMGANVITYTPSLSPFNPAVEVLVFNKLKILKREEYV
 L31-2 *AEch* ITLNLVLGLEGILYANTLDGASNTFLNFFGEAGQATSAL-GNPAIEGQDGIIMDNCA**TH**HEAGTALQTLWMDMGANVITYTPSLSPFNPAVEVLVFNKLKILKREEYV
 L31-2 *PSin* VTLNLLCGLEGVLYANTVEGASDSINFLNFFAEAGRVTTAL-GNPAIEFGDYIILDNCP**TH**RYDTGNILQRLWLMQAGAIITYTPSLSPFNPAVEVLVFNKLKILKREEYV
 L31-3 *ATene* VTLNLLCGRKGIMYANTLKEASNTVEFLNFFYEAASQFTQPD-GNSILYEGYIIVVDNV**AH**RFDEGGEILTEWLSFGMTILYLPVSPSELNPVELVFNKLKILKREEYV
 L31-1 *AEqu* VTLNLLCGRKGIMYANTLKEASNTVEFLNFFYEAASQFTQPD-GNSILYEGYIIVVDNV**AH**RFDEGGEILTEWLSFGMTILYLPVSPSELNPVELVFNKLKILKREEYV
 L31-1 *PCru* VTLNLLCGRKGIMYANTLKEASNTVEFLNFFYEAASQFTQPD-GNSILYEGYIIVVDNV**AH**RFDEGGEILTEWLSFGMTILYLPVSPSELNPVELVFNKLKILKREEYV
 L31-2 *Avir* ITLNLMLCGLNGIMYANTIRGASDTLNFLEFEEASTCQLN-GRPVLYEGDHIILDNCA**TH**HEAGTALQTLWMDMGANVITYTPSLSPFNPAVEVLVFNKLKILKREEYV
 L31-2 *CJar* ITLNLMLCGLNGIMYANTIRGASDTLNFLEFEEASTCQLN-GRPVLYEGDHIILDNCA**TH**HEAGTALQTLWMDMGANVITYTPSLSPFNPAVEVLVFNKLKILKREEYV
 L31-3 *CGra* ITLNLMLCGLNGIMYANTIRGASDTLNFLEFEEASTCQLN-GRPVLYEGDHIILDNCA**TH**HEAGTALQTLWMDMGANVITYTPSLSPFNPAVEVLVFNKLKILKREEYV
 L31-3 *EPal* ITLNLMLCGLNGIMYANTIRGASDTLNFLEFEEASTCQLN-GRPVLYEGDHIILDNCA**TH**HEAGTALQTLWMDMGANVITYTPSLSPFNPAVEVLVFNKLKILKREEYV
 L31-1 *PSpe* ITLNLMLCGLNGIMYANTIRGASDTLNFLEFEEASTCQLN-GRPVLYEGDHIILDNCA**TH**HEAGTALQTLWMDMGANVITYTPSLSPFNPAVEVLVFNKLKILKREEYV
 L31-4 *HMag* VTLNMLAGMEGVMAFANTIDGATTLDFLNFFGEASNSFLRN-GEPLVYEGDHIILDNCA**TH**HEAGTALQTLWMDMGANVITYTPSLSPFNPAVEVLVFNKLKILKREEYV
 L31-1 *SMer* ATLNLVLGLEGVLYANTLDGASNTLEFLNFFDEATKATQIN-GNPVLMAGDILVLDNCA**TH**HNAGGFALGOWLDTMGIDVYLYTPSPSELNPVELVFNKLKILKREEYV
 L31-1 *AMyr* ITLNLFLAGLEGVLYANTLDGASNTLEFLNFFDEATKATQIN-GNPVLMAGDILVLDNCA**TH**HNAGGFALGOWLDTMGIDVYLYTPSPSELNPVELVFNKLKILKREEYV
 L31-5 *EPal* ITLNLFLAGLEGVLYANTLDGASNTLEFLNFFDEATKATQIN-GNPVLMAGDILVLDNCA**TH**HNAGGFALGOWLDTMGIDVYLYTPSPSELNPVELVFNKLKILKREEYV
 L31-3 *AYon* ITVNVLAGLNGIEMNTIHGSDTLIFLDFFSQAGKAANIEGRPALEVGDIIVMDNCA**TH**HNAGGFALGOWLDTMGIDVYLYTPSPSELNPVELVFNKLKILKREEYV
 L31-4 *ATene* ITVNVLAGLNGIEMNTIHGSDTLIFLDFFSQAGKAANIEGRPALEVGDIIVMDNCA**TH**HNAGGFALGOWLDTMGIDVYLYTPSPSELNPVELVFNKLKILKREEYV
 L31-1 *MCap* TTLNMLVSLNGPEYYSIVSGATNTARLSFFQEAESGVNITGRPCLEVGDIIVMDNCA**TH**HNAGGFALGOWLDTMGIDVYLYTPSPSELNPVELVFNKLKILKREEYV
 L31-1 *AMic* TTLNMLVSLNGPEYYSIVSGATNTARLSFFQEAESGVNITGRPCLEVGDIIVMDNCA**TH**HNAGGFALGOWLDTMGIDVYLYTPSPSELNPVELVFNKLKILKREEYV
 L31-4 *MCac* TTLNMLVSLNGPEYYSIVSGATNTARLSFFQEAESGVNITGRPCLEVGDIIVMDNCA**TH**HNAGGFALGOWLDTMGIDVYLYTPSPSELNPVELVFNKLKILKREEYV
 L31-3 *DLin* TTLNMLVSLNGPEYYSIVSGATNTARLSFFQEAESGVNITGRPCLEVGDIIVMDNCA**TH**HNAGGFALGOWLDTMGIDVYLYTPSPSELNPVELVFNKLKILKREEYV
 L31-5 *ACyt* TTLNMLVSLNGPEYYSIVSGATNTARLSFFQEAESGVNITGRPCLEVGDIIVMDNCA**TH**HNAGGFALGOWLDTMGIDVYLYTPSPSELNPVELVFNKLKILKREEYV
 L31-1 *AAwi* TTLNMLVSLNGPEYYSIVSGATNTARLSFFQEAESGVNITGRPCLEVGDIIVMDNCA**TH**HNAGGFALGOWLDTMGIDVYLYTPSPSELNPVELVFNKLKILKREEYV
 L31-3 *AIint* TTLNMLVSLNGPEYYSIVSGATNTARLSFFQEAESGVNITGRPCLEVGDIIVMDNCA**TH**HNAGGFALGOWLDTMGIDVYLYTPSPSELNPVELVFNKLKILKREEYV
 L31-6 *AYon* TTLNMLVSLNGPEYYSIVSGATNTARLSFFQEAESGVNITGRPCLEVGDIIVMDNCA**TH**HNAGGFALGOWLDTMGIDVYLYTPSPSELNPVELVFNKLKILKREEYV
 L31-1 *PAus* TTLNMLVSLNGPEYYSIVSGATNTARLSFFQEAESGVNITGRPCLEVGDIIVMDNCA**TH**HNAGGFALGOWLDTMGIDVYLYTPSPSELNPVELVFNKLKILKREEYV

Рис. 3. Множественное выравнивание последовательностей транспозаз *L31*-транспозонов шестилучевых кораллов. Названия потенциально-функциональных элементов обозначены полужирным курсивом. α -спирали ДНК-связывающего домена выделены серым. Предполагаемый NLS обозначен полужирным курсивом. Триада DDE каталитического домена показана черным цветом. GPRK-мотив обозначен полужирным и подчеркнут.

L31-duo было выделено девять кластеров (рис. 2, доп. материалы 3).

Структурные особенности *L31-транспозонов стрекающих*

В работе [9] описано, что практически все изученные авторами элементы *L31* моллюсков имели белок *OPC2*, при этом довольно часто этот белок был потенциально-функциональным. У изученных кораллов, из 172 выявленных элементов, 96 имели белок *OPC2*, но следует отметить, что ни один из этих белков не был потенциально-функциональным. Все белки *OPC2* имели существенные повреждения (множественные стоп-кодоны, сдвиги рамки считывания). У 76 элементов белок *OPC2* не обнаружен совсем. Таким образом, в отличие от моллюсков, элементы стрекающих подкласса *Hexacorallia* гораздо хуже сохранились.

Существенное различие в структуре между *L31*-транспозонами моллюсков и стрекающих заключается в направлении *OPC* транспозазы и *OPC2* (рис. 1). Во всех изученных элементах шестилучевых кораллов *OPC* транспозазы и *OPC2* направлены к центру (друг к другу стоп-кодонами, в направлении 5'–3'... 3'–5'), тогда как в элементах *L31* моллюсков они направлены вовне (друг к другу старт-кодонами, в направлении 3'–5'... 5'–3').

По первичному анализу (доп. материалы 2) 104 элемента имеют потенциально-функциональную транспозазу, а 68 не имеют таковой. У некоторых элементов количество копий, имеющих потенциально-функциональную транспозазу, колебалось от одной до пяти, но у большинства элементов было представлено одна–две.

Общая длина элементов колебалась от 774 пн (у *L31-1.3_SMer*) до 7598 пн (у *L31-1.2_HCri*). Такой большой разброс длины элементов можно связать с тем, что элементы (будучи неподверженными отбору) сохраняются даже при значительных нарушениях (делециях и инсерциях). КИП обнаружены у 119 элементов. Длина их варьировала от 19 пн до 212 пн. У 53 транспозонов *L31* отсутствуют КИП. У элемента *L31-1_AMur* семь нуклеотидов КИП “наползает” на начало *OPC* транспозазы. Длина потенциально функциональной транспозазы была типична и колебалась в районе 318–367 а.о. Длина нетранспозазного белка от –73 а.о. (у *L31-2_AMyr*) до 406 а.о. (у *L31-1_ATene*), но, как было сказано выше, все эти белки имели делеции. Расстояние между нетранспозазным белком и транспозазой колебалось от 385 пн (у *L31-3_ANas*) до 4755 пн (у *L31-2_AEqu*). У 14 элементов расстояние между транспозазой и *OPC2* равнялось 447 пн. По-видимому, такое расстояние могло быть у предкового элемента, позже оно подверглось вставкам или делециям, что мы наблюдаем у других элементов, а у этих 14 сохранилось без изменений.

Проведенный доменный анализ был выборочным из-за большого количества обнаруженных транспозонов (рис. 3). Мы взяли 27 условно потенциально-функциональных элементов. У 13 элементов присутствуют все шесть α -спиралей ДНК-связывающегося домена, каталитический домен и неповрежденная структура *OPC*. Однако GRPR-подобный мотив обнаружен не у всех, у элемента *L31-3_AInt* его нет. Таким образом, по результатам доменного анализа только 12 элементов имеют рабочую структуру. Если сделать проекцию полученных результатов на все 104 элемента, обозначенных нами потенциально-функциональными (доп. материалы 2), то после доменного анализа больше половины кандидатов, вероятно, не имели бы рабочей структуры.

ОБСУЖДЕНИЕ

На данный момент установлено, что *L31*-транспозоны присутствуют в геномах двустворчатых моллюсков (*Bivalvia*) [9, 12] и шестилучевых кораллов (*Hexacorallia*) [12], а также есть свидетельства их наличия у нематод (*Nematoda*) [13]. Тогда как *L31*-транспозоны двустворчатых моллюсков и шестилучевых кораллов изучены достаточно детально, с присутствием и разнообразием данных элементов у нематод еще только предстоит разобраться. По усредненным оценкам, время дивергенции стрекающих (*Cnidaria*), к которым относятся коралловые полипы, и двусторонне-симметричных (*Bilateria*), к которым относятся моллюски, составляет 685 миллионов лет назад (млн. лет назад) (604–1250 млн лет назад) [26]. Присутствие *L31*-транспозонов в таких эволюционно отдаленных группах может свидетельствовать о том, что один или оба таксона получили данную группу МГЭ через горизонтальный перенос. Горизонтальный перенос играет значимую роль в распространении и эволюции МГЭ [27, 28].

Среди *L31*-транспозонов моллюсков был выявлен элемент *L31-3a_SGlo* [9], у которого *OPC2* направлена как у всех моллюсков (от центра к КИП), тогда как *OPC*-транспозазы – как у шестилучевых кораллов (к центру) (рис. 1). Наличие такого промежуточного варианта дает основания предполагать, что возможно, среди моллюсков появился предковый вариант *L31-duo* (с *OPC*, направленными к центру), который в результате горизонтального переноса проник в геном предка шестилучевых кораллов. На это указывает и структура филогенетического древа (рис. 2). Моллюски в группе *L31-duo* формируют три клады (*L31-1*, *L31-2/L31-3* и *L31-4*), тогда как кораллы одну, близкую к кладе моллюсков *L31-4*. Однако общая ветвь элементов *L31-4* моллюсков и всех *L31-duo* кораллов имеют довольно низкую бутстреп-поддержку, что не дает оснований утверждать об эволюционной близости

данных групп. Вполне вероятно, что кораллы могли получить *L31*-транспозон из клады *L31-3* моллюсков, к которой относится элемент *L31-3a_SGlo*.

Пока нет возможности однозначно установить какая из групп, *L31-duo* или *L31-uno*, является первичной. Однако наличие в группе *L31-uno* элемента кораллов *L31-5_DLin* с OPC2 (доп. материалы 3) указывает на то, что эта *L31-uno* возникла в результате появления среди копий предкового *L31*-транспозона кораллов варианта с каталитическим доменом DD38E и с последующей потерей OPC2 большинством потомков. При этом моллюски получили *L31-uno* обратно через горизонтальный перенос.

Однако не исключен и следующий сценарий. *L31*-транспозоны включают две группы *L31-duo* и *L31-uno*, отличающиеся и по паттерну каталитического домена, и по структуре. При этом в обеих группах есть элементы и моллюсков, и кораллов. Это дает основания предполагать, что разделение данных групп произошло до дивергенции таксонов животных и, соответственно, все разнообразие *L31*-транспозонов является следствием вертикального наследования. При этом у преобладающего большинства таксонов животных *L31*-транспозоны были потеряны. Однако, мы все же склоняемся к первому варианту. Кроме того, есть еще *L31*-транспозоны нематод, и по мере поступления новых данных картина будет проясняться.

После переноса в новый геном МГЭ проходят череду этапов, названных “жизненным циклом” [29]. При успешной колонизации генома нового хозяина, для МГЭ может наступить этап диверсификации, когда под воздействием мутационных процессов в отсутствие отбора появляется множество различных вариантов элемента. Чем выше транспозиционная активность на ранних порах, тем, соответственно, будет выше разнообразие копий-потомков. Элементы актиний и склерактиний представлены во всех девяти кластерах *L31-duo*, а также не создают очевидных коррелирующих с таксонами клад в группе *L31-uno* (рис. 2, доп. материалы 3), соответственно основная диверсификация *L31*-транспозонов кораллов произошла до дивергенции (436.6–538.7 млн. лет назад [26]) этих двух групп шестилучевых кораллов. Высокое разнообразие *L31*-транспозонов кораллов свидетельствуют об очень высокой активности ранних копий. В то же время, в отличие от моллюсков, элементы шестилучевых кораллов гораздо хуже сохранились. Возможно, это связано с более ранним подавлением активности *L31*-транспозонов.

В результате настоящего исследования *L31*-транспозонов шестилучевых кораллов была получена детальная информация о распространении, разнообразии и структуре элементов. Были выявлены две группы: *L31-duo* и *L31-uno*,

отличающиеся и по паттерну каталитического домена, и по структуре. В результате реконструкции эволюции *L31*-транспозонов, основанной на данных о филогении, распространении, разнообразии и структуре мы предполагаем, что шестилучевые кораллы получили *L31*-транспозоны от двустворчатых моллюсков. При этом вышеупомянутая группа *L31-uno* возможно была получена моллюсками в результате горизонтального переноса уже от кораллов. Исследования распространения и разнообразия МГЭ у морских беспозвоночных будут способствовать лучшему пониманию процессов эволюции МГЭ и их роли в эволюционной истории видов.

Работа выполнялась в рамках государственного задания ФГБУН ИМБИ “Функциональные, метаболические и токсикологические аспекты существования гидробионтов и их популяций в биотопах с различным физико-химическим режимом”, номер гос. регистрации 121041400077-1.

Исследование одобрено Этическим комитетом Института биологии южных морей имени А.О. Ковалевского РАН, от 25.09.2023 г., номер протокола 5.

Все применимые международные, национальные и/или институциональные принципы ухода и использования животных были соблюдены.

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Kojima K.K. Structural and sequence diversity of eukaryotic transposable elements // *Genes Genet. Syst.* 2020. V. 94. № 6. P. 233–252. <https://doi.org/10.1266/ggs.18-00024>
2. Wells J.N., Feschotte C.A. Field guide to eukaryotic transposable elements // *Annu. Rev. Genet.* 2020. V. 54. P. 539–561. <https://doi.org/10.1146 annurev-genet-040620-022145>
3. Wicker T., Sabot F., Hua-Van A. et al. A unified classification system for eukaryotic transposable elements // *Nat. Rev. Genet.* 2007. V. 8. № 12. P. 973–982. <https://doi.org/10.1038/nrg2165>
4. Kapitonov V.V., Jurka J. A universal classification of eukaryotic transposable elements implemented in Repbase // *Nat. Rev. Genet.* 2008. V. 9. P. 411–412. <https://doi.org/10.1038/nrg2165-cl>
5. Yuan Y.W., Wessler S.R. The catalytic domain of all eukaryotic cut-and-paste transposase superfamilies // *Proc. Natl Acad. Sci. USA.* 2011. V. 108. № 19. P. 7884–7889. <https://doi.org/10.1073/pnas.1104208108>
6. Arkhipova I.R. Using bioinformatic and phylogenetic approaches to classify transposable elements and

- understand their complex evolutionary histories // *Mob. DNA*. 2017. V. 8. № 19.
<https://doi.org/10.1186/s13100-017-0103-2>
7. Gao B., Wang Y.L., Diaby M. et al. Evolution of *pogo*, a separate superfamily of *IS630-Tc1-mariner* transposons, revealing recurrent domestication events in vertebrates // *Mob. DNA*. 2020. V. 11. № 25.
 8. Shi S., Puzakov M., Guan Z. et al. Prokaryotic and eukaryotic horizontal transfer of *Sailor* (dd82e), a new superfamily of *IS630-Tc1-Mariner* DNA-transposons // *Biology (Basel)*. 2021. V. 10. № 10.
<https://doi.org/10.3390/biology10101005>
 9. Puzakov M.V., Puzakova L.V. Structure and evolution of DNA transposons of the L31 superfamily in *Bivalves* // *Mol. Biol.* 2024. V. 58. № 1. P. 57–75.
<https://doi.org/10.1134/S0026893324010114>
 10. Shi S., Puzakov M.V., Puzakova L.V. et al. *Hiker*, a new family of DNA transposons encoding transposases with DD35E motifs, displays a distinct phylogenetic relationship with most known DNA transposon families of *IS630-Tc1-mariner* (*ITm*) // *Mol. Phylog. Evol.* 2023. V. 188.
<https://doi.org/10.1016/j.ympev.2023.107906>
 11. Aziz R.K., Breitbart M., Edwards R.A. Transposases are the most abundant, most ubiquitous genes in nature // *Nucl. Ac. Res.* 2010. V. 38. № 13. P. 4207–4217.
<https://doi.org/10.1093/nar/gkq140>
 12. Puzakov M.V., Puzakova L.V., Cheresiz S.V. An Analysis of *IS630/Tc1-mariner* transposons in the genome of a pacific oyster *Crassostrea gigas* // *J. Mol. Evol.* 2018. V. 86. № 8. P. 566–580.
<https://doi.org/10.1007/s00239-018-9868-2>
 13. Dupeyron M., Baril T., Bass C., Hayward A. Phylogenetic analysis of the *Tc1-mariner* superfamily reveals the unexplored diversity of *pogo*-like elements // *Mob. DNA*. 2020. V. 11. № 21.
<https://doi.org/10.1186/s13100-020-00212-0>
 14. Tellier M., Bouuaert C.C., Chalmers R. *Mariner* and the *ITm* superfamily of transposons // *Microbiol. Spectr.* 2015. V. 3. № 2.
<https://doi.org/10.1128/microbiolspec.MDNA3-0033-2014>
 15. Ivics Z., Izsvák Z. *Sleeping Beauty* transposition // *Microbiol. Spectr.* 2015. V. 3. № 2.
<https://doi.org/10.1128/microbiolspec.MDNA3-0042-2014>
 16. Jahn C.L., Doktor S.Z., Frels J.S. et al. Structures of the *Euplotes crassus* *Tec1* and *Tec2* elements: Identification of putative transposase coding regions // *Gene*. 1993. V. 133. № 1. P. 71–78.
[https://doi.org/10.1016/0378-1119\(93\)90226-s](https://doi.org/10.1016/0378-1119(93)90226-s)
 17. Chen X., Landweber L.F. Phylogenomic analysis reveals genome-wide purifying selection on *TBE* transposons in the ciliate *Oxytricha* // *Mob. DNA*. 2016. V. 7. № 2.
<https://doi.org/10.1186/s13100-016-0057-9>
 18. Dupeyron M., Singh K.S., Bass C., Hayward A. Evolution of *Mutator* transposable elements across eukaryotic diversity // *Mob. DNA*. 2019. V. 10. № 12.
<https://doi.org/10.1186/s13100-019-0153-8>
 19. Doak T.G., Witherspoon D.J., Jahn C.L., Herrick G. Selection on the genes of *Euplotes crassus* *Tec1* and *Tec2* transposons: Evolutionary appearance of a programmed frameshift in a *Tec2* gene encoding a tyrosine family site-specific recombinase // *Eukaryot. Cell*. 2003. V. 2. № 1. P. 95–102.
<https://doi.org/10.1128/EC.2.1.95-102.2003>
 20. Altschul S.F., Madden T.L., Schäffer A.A. et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs // *Nucl. Ac. Res.* 1997. V. 25. № 17. P. 3389–3402.
<https://doi.org/10.1093/nar/25.17.3389>
 21. Buchan D.W.A., Jones D.T. The PSIPRED protein analysis workbench: 20 years on // *Nucl. Ac. Res.* 2019. V. 47. P. 402–407.
<https://doi.org/10.1093/nar/gkz297>
 22. Crooks G.E., Hon G., Chandonia J.M., Brenner S.E. WebLogo: A sequence logo generator // *Genome Res.* 2004. V. 14. № 6. P. 1188–1190.
<https://doi.org/10.1101/gr.849004>
 23. Hoang D.T., Chernomor O., von Haeseler A. et al. UFBoot2: Improving the ultrafast bootstrap approximation // *Mol. Biol. Evol.* 2018. V. 35. № 2. P. 518–522.
<https://doi.org/10.1093/molbev/msx281>
 24. Kalyaanamoorthy S., Minh B.Q., Wong T.K.F. et al. ModelFinder: Fast model selection for accurate phylogenetic estimates // *Nat. Methods*. 2017. V. 14. № 6. P. 587–589.
<https://doi.org/10.1038/nmeth.4285>
 25. Yamada K.D., Tomii K., Katoh K. Application of the MAFFT sequence alignment program to large data – Reexamination of the usefulness of chained guide trees // *Bioinformatics*. V. 32. № 21. P. 3246–3251.
<https://doi.org/10.1093/bioinformatics/btw4122016>
 26. Kumar M., Suleski J.E., Craig A.E. et al. TimeTree 5: An expanded resource for species divergence times // *Mol. Biol. Evol.* 2022. V. 39(8).
<https://doi.org/10.1093/molbev/msac174>
 27. Wallau G. L., Ortiz M. F., Loreto E. L. Horizontal transposon transfer in eukarya: detection, bias, and perspectives // *Genome Biol. Evol.* 2012. V. 4. № 8. P. 689–699.
<https://doi.org/10.1093/gbe/evs055>
 28. Melo E.S., Wallau G.L. Mosquito genomes are frequently invaded by transposable elements through horizontal transfer // *PLoS Genet.* 2020. V. 16(11).
<https://doi.org/10.1371/journal.pgen.1008946>
 29. Blumenstiel J.P. Birth, school, work, death, and resurrection: The life stages and dynamics of transposable element proliferation // *Genes (Basel)*. 2019. V. 10. № 5.
<https://doi.org/10.3390/genes10050336>

***L31* Transposons of Hexacorallia: Distribution, Diversity and Evolution**

L. V. Puzakova¹, M. V. Puzakov^{1, *}, P. M. Puzakova²

¹*Kovalevsky Institute of Biology of the Southern Seas of Russian Academy of Sciences, Sevastopol, 299011 Russia*

²*Lomonosov Moscow State University, Branch in Sevastopol, Sevastopol, 299001 Russia*

**e-mail: puzakov.mikh@yandex.ru*

Transposable elements (TE) of eukaryotes – retrotransposons and DNA transposons – are nucleotide sequences that can move from locus to locus of the genome, as well as between the genomes of different organisms. *L31* DNA transposons are an ancient and diverse group belonging to the large IS630/Tc1/mariner group. *L31* transposons are not widespread and are present in a limited number of taxa. In addition to the sequence encoding the DDE/D transposase, *L31* transposons carry another ORF (ORF2). Detailed analysis of *L31* elements in the genomes of six-rayed corals has provided detailed information on the distribution, diversity and structure of the elements. Two large groups, *L31*-duo and *L31*-uno, were identified, differing in both catalytic domain pattern and structure. As a result of reconstruction of the evolution of *L31* transposons, it was suggested that six-rayed corals received *L31* transposons from bivalves. At the same time, the split-off group *L31*-uno may have been obtained by mollusks as a result of horizontal transfer from corals. Studies of the distribution and diversity of TE in marine invertebrates will contribute to a better understanding of the evolutionary processes of TE and their role in the evolutionary history of species.

Keywords: DNA transposons, *L31* transposons, coral polyps, genome evolution, biodiversity.