

 pismavavilov.ru

DOI 10.18699/Letters2020-6-12

Original article

Classification of LTR retrotransposons in the flatworm *Macrostomum lignano*

M. Biryukov¹, E. Berezikov^{1,2}, K. Ustyantsev ¹ 

Abstract: Transposable elements (TEs) are important drivers of eukaryotic genome evolution. However, their movement can be deleterious for the genome of an individual cell or organism. Therefore, host developed numerous mechanisms to control activity of TEs, and characterizing the repertoire of TEs in a given species is an essential component of understanding the biology of its genome. *Macrostomum lignano* is a free-living flatworm that is increasingly used as a model organism to study various biological phenomena, such as regeneration, ageing, and resilience to genotoxic stress. The genome of *M. lignano* was recently published, providing an opportunity to study the repertoire of its TEs. Almost 21% of *M. lignano* genome is comprised by long terminal repeat retrotransposons (LTR-RTs). Therefore, LTR-RTs are the main players in the host/TEs arms race in *M. lignano* and should be under tight control of the host's defense system. Here, we provide the first phylogeny-based classification of LTR-RTs, their abundance and diversity in *M. lignano* genome. Interestingly, we found evidence that several *M. lignano* LTR-RTs acquired heat shock response elements in their promoters. The direct regulation of TEs by heat stress was previously described only in plants, and *M. lignano* is the first example that LTR-RTs in animals might also use heat stress to circumvent host's defense system.

Key words: transposable elements; heat stress; heat shock; phylogeny; flatworms; retrotransposons.

Acknowledgments: The work on LTR retrotransposon search and heat shock elements identification and annotation was supported by the Russian Foundation for Basic Research (RFBR, grant No. 18-34-00288) to KU. Classification of LTR retrotransposons was supported by the Russian State Budget project No. 0324–2019-0040-C-01 to EB and KU.

For citation: Biryukov M., Berezikov E., Ustyantsev K. Classification of LTR retrotransposons in the flatworm *Macrostomum lignano*. Pisma v Vavilovskii Zhurnal Genetiki i Seleksii = Letters to Vavilov Journal of Genetics and Breeding. 2020;6(2):54-59. DOI 10.18699/Letters2020-6-12.

Оригинальное исследование

Классификация LTR-ретротранспозонов плоского червя *Macrostomum lignano*

М.Ю. Бирюков¹, Е.В. Березиков^{1,2}, К.В. Устьянцев ¹ 

Аннотация: Мобильные элементы (МЭ) – важные драйверы эволюции геномов эукариот. Однако их перемещение может быть вредоносным для генома отдельной клетки или организма. В связи с этим геномы хозяев развили различные механизмы контроля активности МЭ, и, таким образом, характеристика разнообразия МЭ отдельных видов составляет важную часть понимания биологии их генома. *Macrostomum lignano* – свободноживущий плоский червь, который все чаще используют в качестве модельного организма для исследования различных биологических явлений, таких как регенерация, старение и устойчивость к генотоксическому стрессу. Геном *M. lignano* недавно опубликован, что создает возможность для изучения разнообразия его мобильных элементов. Почти 21% генома *M. lignano* занимают LTR-ретротранспозоны (LTR-РТ). Следовательно, именно LTR-РТ становятся основными игроками в гонке вооружений хозяин/МЭ у *M. lignano*, и поэтому они должны находиться под жестким контролем системы защиты хозяина. В данной работе мы представляем первую филогенетически опосредованную классификацию LTR-РТ, их разнообразие и распространение в геноме *M. lignano*. Мы обнаружили интересное свидетельство того, что некоторые LTR-РТ приобрели элементы чувствительности к тепловому стрессу в своих промоторах. Прямая регуляция МЭ тепловым стрессом ранее была описана лишь у растений, и *M. lignano* демонстрирует первый пример того, что LTR-РТ животных также могут использовать тепловой стресс для обхода защитных систем организма-хозяина.


Ключевые слова: мобильные элементы; температурный стресс; тепловой шок; филогения; плоские черви; ретротранспозоны.


¹ Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

² European Research Institute for the Biology of Ageing, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

¹ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Европейский исследовательский институт биологии старения, Университет Гронингена, Университетский медицинский центр Гронингена, Гронинген, Нидерланды

 e-mail: ustyantsev@bionet.nsc.ru

 © Biryukov M., Berezikov E., Ustyantsev K., 2020

Introduction

Transposable elements (TEs) are important components of eukaryotic genomes. Through their mobility in a genome, TEs can give rise to a variety of genetic alterations, which can be either deleterious or advantageous for the host genome (Feschotte, 2008). For their ability to generate genetic variation in populations and to shape the genome structure, TEs are now recognized as key drivers of evolution of eukaryotes (Kazazian, 2004; Koonin, 2016). However, from the perspective of individual genome stability, TEs movement and propagation are mostly posing a threat if left uncontrolled (Kidwell, Lisch, 2001). Therefore, host genomes developed numerous mechanisms to suppress and regulate TEs activity (Obbard et al., 2009; Goodier, 2016; Tóth et al., 2016).

A substantial part of eukaryotic genomes is constituted by retrotransposons, a class of TEs which replicate through reverse transcription of their messenger RNA, with subsequent integration of new copies into the genome (Wicker et al., 2007). Such 'copy and paste' mechanism underlies fast propagation of retrotransposons. Among retrotransposons, two distinct groups, which differ both in their principal life cycle organization and structural composition, are defined: long terminal repeat (LTR) retrotransposons (LTR-RTs) and non-LTR retrotransposons (lacking LTRs) (Wicker et al., 2007). LTR-RTs are phylogenetic relatives of vertebrate retroviruses and structurally are also similar to them (Xiong, Eickbush, 1990). They usually encode for two genes positioned between the LTRs: *gag* (codes for the structural protein forming virus-like particle) and *pol* (polyprotein, codes for essential enzymatically active proteins: protease (PR), ribonuclease H (RNH), reverse transcriptase (RT), and integrase (INT)). PR processes the Pol polyprotein and the Gag protein into functional units. RT-RNH enzyme performs the reverse transcription process. INT integrates new LTR-RTs copy into the genome. LTRs contain regulatory sequences, including promoter region recognized by RNA polymerase II, transcription termination signals and other regulatory motifs (Kumar, Bennetzen, 1999; Havecker et al., 2004). LTR-RTs are generally classified based on the phylogenetic relationships of the amino-acid sequences of the RT domain into four major families: *Ty1/Copia*, *Ty3/Gypsy*, *Bel/Pao* and *Retroviridae* (vertebrate retroviruses) (Eickbush, Malik, 2002; Havecker et al., 2004). There are other types of LTR-RTs-like elements, *DIRS* and plant caulimoviruses, which lack LTRs and have distinct structural features, but still share a common origin with other LTR-RTs (Xiong, Eickbush, 1990; Poulter, Goodwin, 2005).

Recently, we assembled and analyzed the genome of a free-living flatworm *M. lignano* (Wudarski et al., 2017), a model organism used to study the mechanisms of stem cells regulation, regeneration and ageing (Wudarski et al., 2020). We found that about 25% of the *M. lignano* genome is comprised by various TEs, among which 83% (~21% of the genome) are LTR-RTs (Wudarski et al., 2017). Thus, LTR-RTs are the main players in the host/TEs arms race in *M. lignano* and should be under tight control of the host's defense system. However, absence of detailed classification of the identified LTR-RTs prevents the studies of how their activity is regulated in the *M. lignano* genome, which is required for comprehensive understanding of the biology of this model animal.

In eukaryotes, piRNA-mediated (Piwi-interacting RNA, a class of small RNAs) post-transcriptional silencing is one of the major mechanisms to control TEs activity (Iwasaki et al., 2015; Tóth et al., 2016; Dechaud et al., 2019). The piRNA pathway was shown to be positively regulated by Hsp90 (heat shock protein 90) protein in *Drosophila*, and mutations in *Hsp90* result in dysregulation of piRNA biogenesis and subsequent TEs activation (Specchia et al., 2010). In nematodes, environmental stress such as heat or oxidative stress can destabilize the function of Hsp90, leading to increase in TEs movement (Ryan et al., 2016). In line with this, previous studies reported activation of LTR retrotransposons under heat shock in *Drosophila*, plants and fungi (Vasilyeva et al., 1999; Ikeda et al., 2001; Cavrak et al., 2014). Interestingly, in *Arabidopsis thaliana* and other Brassicaceae species specific heat shock sensitive motifs (heat shock elements, HSEs) were found in some LTR-RTs promoter sequences, which rendered them directly transcriptionally activated by heat stress conditions (Cavrak et al., 2014; Pietzenek et al., 2016). Recently, we showed that long-term culturing of *M. lignano* under elevated temperatures is detrimental for the development of the worm (Wudarski et al., 2019), and it is possible that activation of LTR-RTs is the reason for this observation.

Here, we provide the first classification of LTR-RTs in the *M. lignano* genome, which can serve as a basis to study complex relationships between the host/TEs inter-relationships in this model animal. Additionally, we investigate the presence of HSEs in the promoter sequences of the identified LTR-RTs as a first step towards understanding of how heat stress can regulate LTR-RTs in *M. lignano*.

Materials and methods

Computational mining, structural annotation, and clustering of LTR-RTs. For the analysis, the latest published (Wudarski et al., 2017) genome assembly of *M. lignano*, Mlig_3_7, was used (GenBank accession number GCA_002269645.1). For identification of full-length LTR-RTs in the genome assembly, the following computational pipeline was implemented. First, de novo search for long terminal repeat (LTRs)-like sequences was performed using the LTRharvest tool (Ellinghaus et al., 2008) with the following set of parameters: -vic 10, -seed 20, -minlenltr 100, -maxlenltr 7000, -motif TGCA, -similar 85.0, -mintsd 4, -maxtsd 6, -motifmis 1. Second, to remove false positive hits, including nested LTR-RTs insertions and tandem repeat sequences, the results of LTRharvest were further processed using LTR_retriever software using the default parameters (Ou, Jiang, 2018). Third, the resulted set of LTR-RTs-like sequences was screened for the presence of the LTR-RTs-specific protein domains (RT, INT, RNH, PR, and GAG) using standalone RPS-BLAST rpstblastn (Reverse-Position-Specific BLAST with Translated nucleotide vs protein domain search) software (Marchler-Bauer et al., 2009), and the sequences without the RT domain were filtered-out from further analysis. Each element was given a score based on the following criteria: integrity of the predicted RT domain, presence and integrity of other non-RT domains, similarity of the LTRs, and integrity of the polyprotein open reading frame. Next, passed LTR-RTs sequences were clustered into lineages using MMseqs2 (Steinegger, Söding, 2017) with easy-cluster parameter by their RT domain amino acid sequences, requiring

up to 50% identity and 80% coverage. Lineages comprised of only 1 or 2 elements were discarded from further analysis. Representative LTR-RTs sequences from each lineage were selected based on the maximum score assigned.

Phylogenetic analysis and classification. Amino acid sequences of RT, RNH, and INT domains of the selected representative *M. lignano* LTR-RTs were aligned using MAFFT (Katoh, Standley, 2013) to the corresponding "RT/RNaseH/INT"_retroelement profile alignments retrieved from GyDB (http://gydb.org/index.php/Collection_alignments). Next, the resulted alignments were concatenated together and the final RT-RNH-INT alignment was used for phylogenetic analysis. Phylogenetic analysis by maximum likelihood was performed using standalone version of IQ-TREE (Nguyen et al., 2015) with automatic substitution model selection (-m TEST parameter) applying SH-aLRT (Guindon et al., 2010) and UFBoot (Minh et al., 2013) (-alrt 1000, -bb parameters) to estimate branch support values.

Classification of *M. lignano* LTR-RTs sequences was performed based on their phylogenetic position on the resulted tree and vicinity to known sequences from GyDB (<http://gydb.org>).

The identified representative LTR-RTs nucleotide sequences, their structural annotation and phylogenetic based classification (see below) are available at http://gb.macgenome.org/downloads/Mlig_3_7/Repeats/LTR.

Identification of heat shock elements in the LTRs. Using the in-house written Python script, LTRs sequences of the identified *M. lignano* LTR-RTs were checked for the presence of the canonical heat shock element (HSE) palindromic motif nTTCnnGAAnnTTCn (Mittal et al., 2011) using the following regular expression pattern: TTC[GATC][GATC]GAA[GATC][GATC]TTC, allowing for overlapping matches and searching in both DNA strands. The search was considered significant when more than 1 perfect match was found per LTR and when it was repeated in both LTRs.

Results

Initial search using LTRharvest and LTR_retriever in *M. lignano* genome assembly resulted in 1225 LTR-RTs-like sequences. After screening for LTR-RTs-specific protein domains and filtering by the presence of the RT domain, 997 full-length LTR-RTs sequences were retained. Subsequent clustering by the RT domain sequences followed by filtering of singletons and clusters with only 2 sequences resulted in 983 LTR-RTs distributed among 33 lineages (Table 1).

Phylogenetic analysis using amino acid sequences of LTR-RTs conserved protein domains (RT, RNH, and INT) was performed to classify the identified set of representative *M. lignano* LTR-RTs (see Figure). The overwhelming majority (90.8%) of *M. lignano* LTR-RTs belong to the *Ty3/Gypsy* group (893 elements) and spread among 5 clusters, 4 of which are known (*CsRN1*, *Mag*, *Gmr1*, *Cer1*), and 1 paraphyletic cluster appears to be specific to *M. lignano* (see Table 1, Figure). The highest diversity, with 18 clustered lineages, was found for the *Mag* cluster, 7 lineages for the paraphyletic new cluster, and 1 lineage for the *CsRN1*, *Gmr1*, and *Cer1* clusters (see Table 1, Figure). Only one *M. lignano*-specific cluster of LTR-RTs, comprised of a single lineage with 17 elements, was found in the *Ty1/Copia* group near the known

Hydra and *Copia* clusters. In the *Bel/Pao* group, *M. lignano* LTR-RTs clustered together with the *Bel* and *Tas* clusters comprised of 1 and 3 lineages, respectively (see Figure).

Multiple canonical HSEs ranging from 3 to 4 perfect motif matches were identified in 5' region of the LTR promoters of two clustered lineages of *M. lignano* *Ty3/Gypsy* LTR-RTs: *Mlig_LTR-RT_73* (*Mag*, 9 out of 13 elements with HSEs) and *Mlig_LTR-RT_986* (*CsRN1*, 9 out of 39 elements with HSEs) (Table 2, see Figure).

Discussion

TEs and LTR-RTs in particular are important components of eukaryotic genomes. In *M. lignano*, LTR-RTs comprise the majority of its TEs and constitute almost 21% of its genome (Wudarski et al., 2017). Providing detailed information on abundance, diversity and classification of LTR-RTs is a necessary step towards more comprehensive understanding of various biological aspects of this model animal.

We identified 983 *M. lignano* full-length LTR retroelements belonging to the 3 major phylogenetic branches of LTR-RTs: *Ty3/Gypsy*, *Bel/Pao*, and *Ty1/Copia*, with the apparent prevalence of two *Ty3/Gypsy* clusters: *Mag* (489 elements) and a new one (new *Mlig*-specific, 319 elements) (see Table 1, Figure). This new cluster is paraphyletic, as it does not have credible enough aLRT (Guindon et al., 2010) and UFBoot (Minh et al., 2013) values (aLRT \geq 80 and UFBoot \geq 95) to support the monophyly and rather represents a set of closely-related *M. lignano* LTR-RTs sequences which can be members of distinct clusters (see Figure). Therefore, assignment of distinct *M. lignano* LTR-RTs to this new cluster should be taken with care, and we advise to refer to one of its 7 lineages, since their relationships to each other are still not fully confirmed. The new *Mlig*-specific cluster in the *Ty1/Copia* group is represented by a single lineage and cannot be unambiguously assigned to any of the existent clusters due to insufficient phylogenetic support (see Figure).

Previous publications report that heat stress can influence either directly or indirectly the activation of TEs and LTR-RTs in particular in various animal, plant and fungi species (Vasilyeva et al., 1999; Ikeda et al., 2001; Specchia et al., 2010; Pietzenuk et al., 2016; Ryan et al., 2016). Recent studies on a model plant *Arabidopsis thaliana* and other Brassicaceae species indicate that some *Ty1/Copia* LTR-RTs, namely *ONSEN*, have acquired canonical HSEs motifs (Mittal et al., 2011) in their promoter sequences that rendered them directly activated upon heat stress (Cavrak et al., 2014; Pietzenuk et al., 2016). Inspired by this and taking into account our own study, which showed deleterious effect of elevated, heat shock range, temperatures on the development in *M. lignano* (Wudarski et al., 2019), we attempted a search for HSEs motifs in *M. lignano* LTR-RTs. Intriguingly, we found that *M. lignano* LTR-RTs belonging to two phylogenetically distinct *Ty3/Gypsy* clusters (*Mag* and *CsRN1*) acquired multiple (from 3 to 4) HSEs in their promoter sequences (see Table 2). Therefore, it can be hypothesized that the observed developmental and growth anomalies in *M. lignano* (Wudarski et al., 2019) can also be a result of the increased LTR-RTs transcription and movement. This interesting fact definitely requires further studies.

Table 1. Summary of the LTR-RTs repertoire and classification in *M. lignano*

Group	Cluster	Num of Lineages	Num of Elements
<i>Ty3/Gypsy</i>	<i>Mag</i>	18	489
	new <i>Mlig</i> -specific	7	319
	<i>CsRN1</i>	1	39
	<i>Gmr1</i>	1	6
	<i>Cer1</i>	1	40
<i>Ty1/Copia</i>	new <i>Mlig</i> -specific	1	17
<i>Bel/Pao</i>	<i>Bel</i>	1	52
	<i>Tas</i>	3	21
	Total	33	983

Group and cluster names are according to GyDB (<http://gydb.org>)

Table 2. *M. lignano* *Ty3/Gypsy* LTR-RTs with canonical HSEs in their LTRs

Cluster Lineage representative*	Element ID**	Num of HSEs
<i>Mag</i> <i>Mlig_LTR-RT_73</i> [13]	73	3
	118	3
	168	3
	224	3
	254	3
	544	3
	821	3
	943	3
	952	3
	1029	3
<i>CsRN1</i> <i>Mlig_LTR-RT_986</i> [39]	8	4
	151	4
	282	3
	290	3
	360	3
	648	4
	980	3
	986	4
	1105	3

* Lineage representative names are as follows: *Mlig_LTR-RT_NUM1* [NUM2], where NUM1 stands for the element ID, and NUM2 indicates the number of elements clustered within the lineage

** Corresponds to NUM1 in an element name

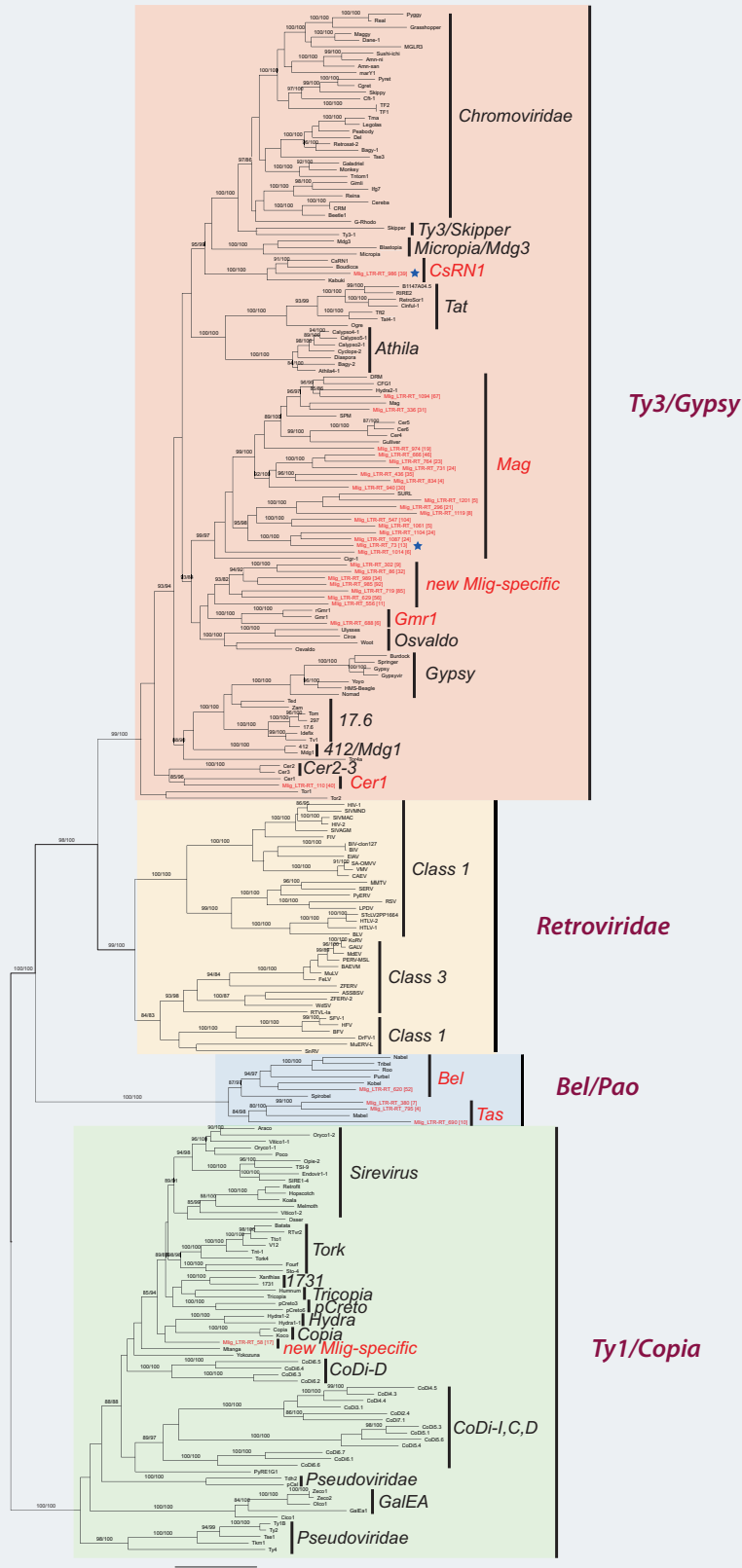
Conclusion

Our analysis lays the ground for research on LTR-RTs and TEs in *M. lignano*. It opens opportunity to study host defense mechanisms that *M. lignano* employs to restrict TEs activity, as well as adaptations that TEs evolved to circumvent these restrictions. The identification of heat stress responsive elements in promoters of the *M. lignano* LTR-RTs indicates that these retroelements may be directly activated by heat stress, as it was previously shown for LTR-RTs in plants. It supports the notion that there is a tight connection between TEs activation

and host's stress response machinery, the question which is now can be addressed in *M. lignano*.

References

- Cavrak V.V., Lettner N., Jamge S., Kosarewicz A., Bayer L.M., Mittelsten Scheid O. How a retrotransposon exploits the plant's heat stress response for its activation. *PLoS Genet.* 2014;10(1):e1004115. DOI: 10.1371/journal.pgen.1004115.
- Dechaud C., Volf J.-N., Schartl M., Naville M. Sex and the TEs: transposable elements in sexual development and function in animals. *Mobile DNA.* 2019;10(1):42. DOI: 10.1186/s13100-019-0185-0.
- Eickbush T.H., Malik H.S. Origins and evolution of retrotransposons.



Unrooted maximum likelihood tree of phylogenetic relationships among representatives of four major LTR-RTs groups reconstructed based on concatenated amino-acid sequences of their RT, RNH and INT protein domains. Significant (higher than 80) branch support values (aLRT/UFBoot) are indicated above the branches leading to corresponding cluster. Major phylogenetic groups are highlighted with colored blocks. Borders of phylogenetic clusters are drawn as vertical black bars to the right from the tree. Names of the clusters and LTR-RTs are according to GyDB (<http://gydb.org>). Names of *M. lignano* LTR-RTs lineage representatives are highlighted in red and include element ID (first number) and the number of elements clustered within the lineage (second number). The clusters that include *M. lignano* LTR-RTs are highlighted in red. *M. lignano* lineages that have elements with HSE motifs in their promoters are indicated with blue star-shaped marks.

- In: Mobile DNA II2002; American Society of Microbiology. 2002;1111-1144. DOI: 10.1128/9781555817954.ch49.
- Ellinghaus D., Kurtz S., Willhoeft U. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinform.* 2008;9(1):18. DOI: 10.1186/1471-2105-9-18.
- Feschotte C. Transposable elements and the evolution of regulatory networks. *Nat Rev Genet.* 2008;9(5):397-405. DOI: 10.1038/nrg2337.
- Goodier J.L. Restricting retrotransposons: a review. *Mobile DNA.* 2016;7(1):16. DOI: 10.1186/s13100-016-0070-z.
- Guindon S., Dufayard J.-F., Lefort V., Anisimova M., Hordijk W., Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* 2010;59(3):307-321. DOI: 10.1093/sysbio/syq010.
- Havecker E.R., Gao X., Voytas D.F. The diversity of LTR retrotransposons. *Genome Biol.* 2004;5(6):225. DOI: 10.1186/gb-2004-5-6-225.
- Ikeda K., Nakayashiki H., Takagi M., Tosa Y., Mayama S. Heat shock, copper sulfate and oxidative stress activate the retrotransposon MAGGY resident in the plant pathogenic fungus *Magnaporthe grisea*. *Mol Genet Genom.* 2001;266(2):318-325.
- Iwasaki Y.W., Siomi M.C., Siomi H. PIWI-Interacting RNA: Its biogenesis and functions. *Annu Rev Biochem.* 2015;84:405-433. DOI: 10.1146/annurev-biochem-060614-034258.
- Katoh K., Standley D.M. MAFFT Multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30(4):772-780. DOI: 10.1093/molbev/mst010.
- Kazian H.H. Mobile elements: drivers of genome evolution. *Science.* (New York, N.Y.). 2004;303(5664):1626-1632. DOI: 10.1126/science.1089670.
- Kidwell M.G., Lisch D.R. Perspective: transposable elements, parasitic DNA, and genome evolution. *Evolution.* 2001;55(1):1-24.
- Koonin E.V. Viruses and mobile elements as drivers of evolutionary transitions. *Philos Trans R Soc B.* 2016;371(1701). DOI: 10.1098/rstb.2015.0442.
- Kumar A., Bennetzen J.L. Plant retrotransposons. *Annu Rev Genet.* 1999;33(1):479-532. DOI: 10.1146/annurev.genet.33.1.479.
- Marchler-Bauer A., Anderson J.B., Chitsaz F., Derbyshire M.K., DeWeese-Scott C., Fong J.H., Geer L.Y., Geer R.C., Gonzales N.R., Gwadz M., He S., Hurwitz D.I., Jackson J.D., Ke Z., Lanczycki C.J., Liebert C.A., Liu C., Lu F., Lu S., Marchler G.H., Mullokandov M., Song J.S., Tasneem A., Than-ki N., Yamashita R.A., Zhang D., Zhang N., Bryant S.H. CDD: specific functional annotation with the Conserved Domain Database. *Nucleic Acids Res.* 2009;37(suppl_1):D205-D210. DOI: 10.1093/nar/gkn845.
- Minh B.Q., Nguyen M.A.T., von Haeseler A. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol.* 2013;30(5):1188-1195. DOI: 10.1093/molbev/mst024.
- Mittal D., Enoki Y., Lavania D., Singh A., Sakurai H., Grover A. Binding affinities and interactions among different heat shock element types and heat shock factors in rice (*Oryza sativa* L.). *FEBS J.* 2011;278(17):3076-3085. DOI: 10.1111/j.1742-4658.2011.08229.x.
- Nguyen L.-T., Schmidt H.A., von Haeseler A., Minh B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 2015;32(1):268-274. DOI: 10.1093/molbev/msu300.
- Obbard D.J., Gordon K.H.J., Buck A.H., Jiggins F.M. The evolution of RNAi as a defence against viruses and transposable elements. *Philos Trans R Soc B.* 2009;364(1513):99-115. DOI: 10.1098/rstb.2008.0168.
- Ou S., Jiang N. LTR_retriever: A highly accurate and sensitive program for identification of long terminal repeat retrotransposons. *Plant Physiol.* 2018;176(2):1410-1422. DOI: 10.1104/pp.17.01310.
- Pietzenek B., Markus C., Gaubert H., Bagwan N., Merotto A., Bucher E., Pecinka A. Recurrent evolution of heat-responsiveness in Brassicaceae COPIA elements. *Genome Biol.* 2016;17(1):209. DOI: 10.1186/s13059-016-1072-3.
- Poulter R.T.M., Goodwin T.J.D. DIRS-1 and the other tyrosine recombinase retrotransposons. *Cytogenet Genome Res.* 2005;110(1-4):575-588. DOI: 10.1159/000084991.
- Ryan C.P., Brownlie J.C., Whyard S. Hsp90 and physiological stress are linked to autonomous transposon mobility and heritable genetic change in Nematodes. *Genome Biol Evol.* 2016;8(12):3794-3805. DOI: 10.1093/gbe/evw284.
- Specchia V., Piacentini L., Tritto P., Fanti L., D'Alessandro R., Palumbo G., Pimpinelli S., Bozzetti M.P. Hsp90 prevents phenotypic variation by suppressing the mutagenic activity of transposons. *Nature.* 2010;463(7281):662-665. DOI: 10.1038/nature08739.
- Steinberger M., Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol.* 2017;35(11):1026-1028. DOI: 10.1038/nbt.3988.
- Tóth K.F., Pezic D., Stuwe E., Webster A. The piRNA pathway guards the germline genome against transposable elements. *Adv Exp Med Biol.* 2016;886:51-77. DOI: 10.1007/978-94-017-7417-8_4.
- Vasilyeva L.A., Bubenshchikova E.V., Ratner V.A. Heavy heat shock induced retrotransposon transposition in *Drosophila*. *Genet Res.* 1999;74(2):111-119.
- Wicker T., Sabot F., Hua-Van A., Bennetzen J.L., Capy P., Chalhoub B., Flavell A., Leroy P., Morgante M., Panaud O., Paux E., SanMiguel P., Schulman A.H. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet.* 2007;8(12):973-982. DOI: 10.1038/nrg2165.
- Wudarski J., Egger B., Ramm S.A., Schärer L., Ladurner P., Zadesenets K.S., Rubtsov N.B., Mouton S., Berezikov E. The free-living flatworm *Macrostomum lignano*. *EvoDevo.* 2020;11(1):5. DOI: 10.1186/s13227-020-00150-1.
- Wudarski J., Simanov D., Ustyantsev K., de Mulder K., Grelling M., Grudniewska M., Beltman F., Glazenburg L., Demircan T., Wunderer J., Qi W., Vizoso D.B., Weissert P.M., Olivieri D., Mouton S., Guryev V., Aboobaker A., Schärer L., Ladurner P., Berezikov E. Efficient transgenesis and annotated genome sequence of the regenerative flatworm model *Macrostomum lignano*. *Nat Commun.* 2017;8(1):2120. DOI: 10.1038/s41467-017-02214-8.
- Wudarski J., Ustyantsev K., Glazenburg L., Berezikov E. Influence of temperature on development, reproduction and regeneration in the flatworm model organism, *Macrostomum lignano*. *Zoological Letters.* 2019;5(1):7. DOI: 10.1186/s40851-019-0122-6.
- Xiong Y., Eickbush T.H. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J.* 1990;9(10):3353-3362.

Conflict of interest. The authors declare no conflict of interest.

Received April 03, 2020. Revised April 30, 2020. Accepted May 06, 2020.