Genetica (2021) 149:335–342
https://doi.org/10.1007/s10709-021-00130-w

SHORT COMMUNICATION

Transposable elements expression in *Rhinella marina* (cane toad) specimens submitted to immune and stress challenge

Adriana Ludwig[1] · Michelle Orane Schemberger[2] · Camilla Borges Gazolla[3] · Joana de Moura Gama[3] · Iraine Duarte[3] · Ana Luisa Kalb Lopes[1,4] · Carolina Mathias[3] · Desirrê Alexia Lourenço Petters-Vandresen[3] · Michelle Louise Zattera[3] · Daniel Pacheco Bruschi[3,5]

Received: 11 March 2021 / Accepted: 2 August 2021 / Published online: 12 August 2021

Abstract

Transposable elements (TEs) are important components of eukaryotic genomes and compose around 30% of the genome of *Rhinella marina*, an invasive toad species. Considering the possible role of TEs in the adaptation of populations, we have analyzed the expression of TEs in publicly available spleen tissue transcriptomic data generated for this species after immune and stress challenge. By analyzing the transcriptome assembly, we detected a high number of TE segments. Moreover, some distinct TE families were differentially expressed in some conditions. Our result shows that several TEs are capable of being transcribed in *R. marina* and they could help to generate a rapid response of specimens to the environment. Also, we can suggest that these TEs could be activated in the germinative cells as well producing variability to be selected and shaped by the evolutionary processes behind the success of this invasive species. Thus, the TEs are important targets for investigation in the context of *R. marina* adaptation.

Keywords RNA-seq · Transcriptome · Retrotransposon · Adaptation

Introduction

Rhinella marina is an invasive Anura species well adapted to diverse environments in different parts of the world and invasive populations have been well characterized at the phenotypic level, although the genetic basis of invasiveness was understudied until recently (Rollins et al. 2015). There is increasing evidence of rapid phenotypic adaptation of *R. marina* to heterogeneous climatic situations, e.g., extreme temperature and dehydrating conditions (Kosmala et al. 2017, 2020a; Mittan and Zamudio 2019). Although some markers have been associated with the *R. marina* adaptability (Rollins et al. 2015; Selechnik et al. 2019; Kosmala et al. 2020b), a reduced overall genetic diversity was observed for invasive populations suggesting that genetic changes may not be the only mechanisms by which adaptation occurs (Selechnik et al. 2019).

The draft genome assembly of *R. marina* indicates a large fraction of transposable elements (TEs), around 30% (Edwards et al. 2018). The TEs are mobile components of eukaryotic genomes and are considered an essential source of genetic variability being able to act as drivers in adaptation and evolution (Schrader et al. 2014; Schrader and Schmitz 2019). Thus, concerning the large fraction of TEs in this invasive species, it is easy to wonder about the impact of these sequences in the adaptation of *R. marina* populations.

The TEs are divided into two major classes (Finnegan 1989) based on their transposition intermediate: RNA (class I—retrotransposons) or DNA (class II—DNA transposons),

✉ Daniel Pacheco Bruschi
danielpachecobruschi@gmail.com

1 Laboratório de Ciências e Tecnologias Aplicadas em Saúde (LaCTAS), Instituto Carlos Chagas – Fiocruz-PR, Curitiba, Paraná, Brazil

2 Laboratório de Biotecnologia Aplicada a Fruticultura, Departamento de Fitotecnia e Fitossanidade, Universidade Estadual de Ponta Grossa (UEPG), Ponta Grossa, Paraná, Brazil

3 Programa de Pós-Graduação em Genética (PPG-GEN), Universidade Federal do Paraná (UFPR), Curitiba, Brazil

4 Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Paraná, Curitiba, Paraná, Brazil

5 Laboratório de Citogenética Evolutiva e Conservação Animal (LabCECA), Departamento de Genética, Universidade Federal do Paraná (UFPR), Curitiba, Brazil

this last being the most abundant category in *R. marina* genome (Edwards et al. 2018). They are further separated into additional categories (such as orders and superfamilies) (Wicker et al. 2007; Kapitonov and Jurka 2008).

The activity of TEs in a genome can provide an opportunity for more rapid adaptation than would be possible from other mutations and consequently TEs can facilitate the adaptive potential of invasive populations (Stapley et al. 2015). Several pieces of evidence had demonstrated an increase of transcriptional activity of TEs under stress conditions that could result in new TE insertions (Goubert et al. 2017; Li et al. 2018; Oliveira et al. 2021; Wos et al. 2021). Amplification waves can generate new opportunities to recruit regulatory elements from TEs as a substrate to create or improve regulatory networks on host genomes (exaptation). Indeed several retrotransposons and DNA transposons contain transcription factor-binding motifs that could serve as cis-elements (such as enhancers) on the regulatory network (Jacques et al. 2013; Sundaram et al. 2014; Sundaram and Wysocka 2020). Studies reveal TEs acting as cis-regulatory elements on the stress-response via in individuals submitted to biotic or abiotic stressors (Salazar et al. 2007; Naito et al. 2009; Pietzenuk et al. 2016; Salces-Ortiz et al. 2020). These outcomes reveal the impacts of these elements on genomic expression and their capacity to contribute to local adaptation.

Recently, the transcriptome of *R. marina* was evaluated under immune and stress challenge (Gardner et al. 2018) and based on the premise exposed above, we have evaluated the transcriptional activity of TEs found in this data. We observed that a great amount of TEs are capable to be expressed in somatic cells and some were differentially expressed in the treatments. Our work highlights the importance of studying TEs as possible facilitators of *R. marina* adaptation.

Material and methods

Summary of study design

We used the RNA sequencing data of *R. marina* published by Gardner et al. (2018) to search for transcripts containing TEs (Fig. 1). Staring from the fasta files of transcriptome

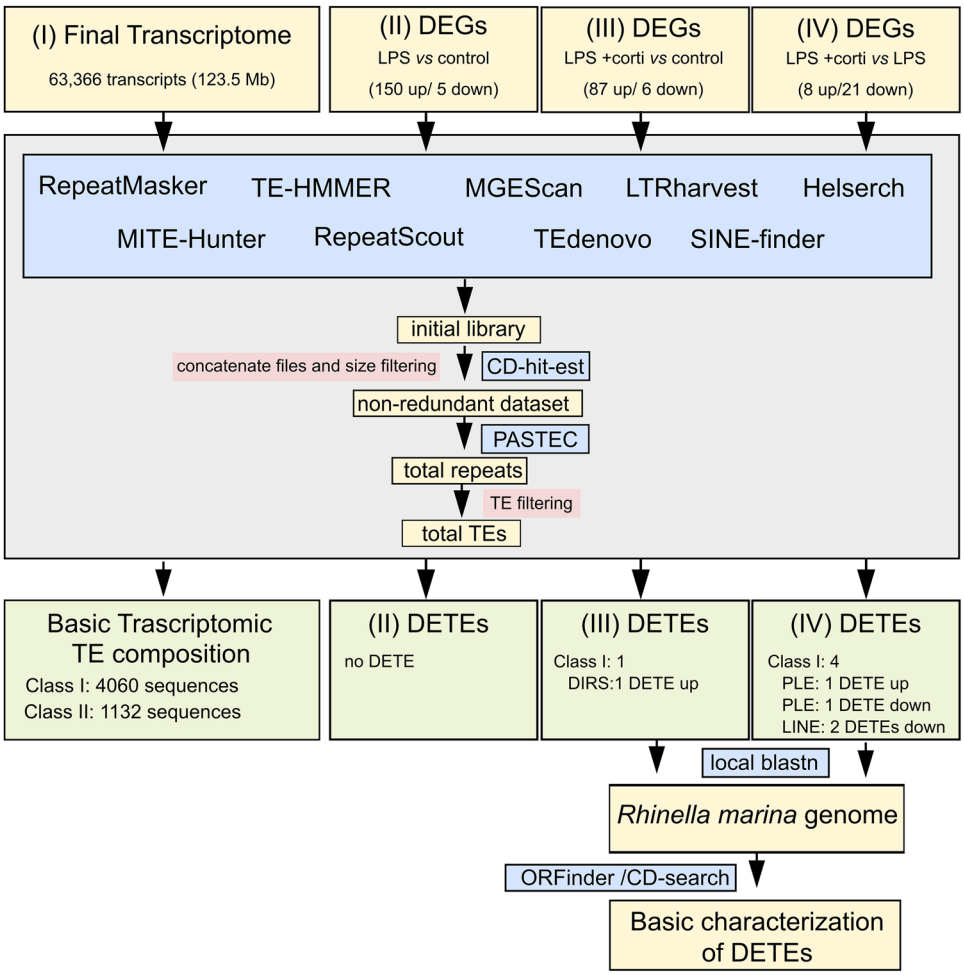

Fig. 1 The schematic workflow of the study. Four initial datasets based on Gardner et al. (2018) were submitted to searches of repeated sequences using nine tools using the PiRATE pipeline. Detected repeat sequences were filtered according to their length (minimum 300 bp) and clustered with CD-HIT-est to reduce redundancy. Putative TE sequences were automatically classified with PASTEC. By manual checking, sequences classified as noCat (uncategorized), SSR, Potential Chimeric and Potential Host Gene were eliminated and the remaining sequences correspond only to TEs (total TEs). The TEs identified from datasets III, and IV were considered Differentially Expressed TEs (DETEs) and were used as queries in local blastn searches against the *R. marina* genome. TEs copies were retrieved for basic characterization looking for the presence of ORFs, expected protein domains, and terminal repeats. Blue boxes represent bioinformatics tools, and gray boxes are tools implemented in PiRATE

assembly provided by the authors, we used different tools to identity TEs in the final transcriptome and among the differentially expressed genes (DEGs) determined by the authors in different conditions (see below). Differentially Expressed TEs (DETEs) were subsequently characterized based on genomic copies.

Datasets

Gardner et al. (2018) obtained the RNA-seq data from male and female spleen tissues of *R. marina* under two treatments: (1) specimens submitted to an acute immune challenge of lipopolysaccharide (LPS) and (2) specimens submitted to LPS plus corticosterone, the main glucocorticoid produced in amphibians under stress. The transcriptome was generated using resulting reads from two randomly selected trimmed libraries from each treatment generating, after some filtering, a transcriptome assembly of 123.5 Mb containing 63,366 transcripts. These authors found 155 DEGs comparing treatment 1 with the control, 93 DEGs comparing treatment 2 with control, and 29 comparing treatment 2 with treatment 1. As the identity of some DEGs was not confirmed by the authors, we tested if they could have been derived from TEs. Thus, based on the fasta sequence data provided by these authors, we constituted four initial datasets for the current work: (I) final transcriptome assembly (63,366 sequences) and three datasets for DEGs (II) treatment 1 (LPS) versus control, (III) treatment 2 (LPS with corticosterone) versus control and (IV) treatment 2 versus treatment 1.

The TE composition of *R. marina* genome was compiled from Edwards et al. (2018) that have evaluated the genome using RepeatMasker. Here, we reanalyzed the TE table content to complement the classification into orders.

Detection and analysis of TEs

The schematic workflow of our study is shown in Fig. 1. We performed the detection of TEs using the pipeline PiRATE (Berthelier et al. 2018), applying the following programs to detect TEs: RepeatMasker (Smit et al. 1996), TE-HMMER, LTRharvest (Ellinghaus et al. 2008), Helsearch (Yang and Bennetzen 2009), MGEScan-nonLTR (Rho and Tang 2009), MITE-Hunter(Han and Wessler 2010), SINE-Finder (Wenke et al. 2011), TEdenovo (Flutre et al. 2011) and Repeat Scout (Price et al. 2005). Detected sequences (only > 300 bp) from these tools were clustered with CD-HIT-est (Huang et al. 2010) to reduce redundancy and TE sequences were automatically classified with PASTEC (Hoede et al. 2014) following the Wicker et al. (2007) classification system. PASTEC results were manually filtered eliminating sequences classified as noCat (uncategorized), SSR (Simple Sequence Repeat), Potential Chimeric and Potential Host Gene. The remaining sequences correspond theoretically only to TEs (total TEs).

The transcripts of TEs identified as DETEs were used as queries in local blastn searches against the *R. marina* genome (GCA_900303285.1) for characterization of copies using a cutoff e-value threshold of e-10. All putative copies were recovered together with 5 kb of flanking regions. The presence of ORFs was analyzed using the NCBI ORFfinder tool, and the CD-search (Marchler-Bauer and Bryant 2004) was used to check for domains in the predicted protein sequences. *In-house* Python codes were used for filtering results and recovering sequences in different steps (available in https://github.com/adriludwig/simplebioscripts). For each DETE, the CENSOR tool was used to find the most similar TE available in the Repbase (Kohany et al. 2006). MAFFT was used to align DNA sequences from genomic copies (Rozewicki et al. 2019) to establish the border of elements. The presence of target site duplications (TSDs) was verified by manual inspection of aligned copies.

Results and discussion

Several TEs are expressed in spleen tissue of *R. marina*

Gardner et al. (2018) have studied the differential gene expression in spleen tissues of cane toad comparing treatments with LPS (immune challenge) and LPS plus exogenous corticosterone (immune and stress challenge). They observed that for the LPS challenge, genes encoding for cytokines that are involved in typical innate responses were significantly upregulated, while in the second treatment, genes involved with cell-mediated immunity were downregulated, suggesting that acute stress may alter the immune response. Here, we investigate if TEs were found to be expressed and differentially expressed in the data produced by these authors.

First, we investigated the content of TEs in the final transcriptome assembly generated with reads from all treatments, and thus, genes could have been expressed in one or more samples. Our searches against this transcriptome identified 16,867 fragments of repeats after redundancy and size filtering. From those, 4060 are classified as retrotransposons, 1132 as DNA transposons while the remaining are other non-TE categories (Table 1; Supplementary Table 1; Supplementary file 1). The most abundant TEs concerning the number of different sequences identified are from the main retrotransposons orders, LINE (Long Interspersed Nuclear Element), DIRS (*Dictyostelium* Intermediate Repeat Sequence) and LTR (Long Terminal Repeat) and the classical DNA transposons from order TIR (Terminal Inverted Repeat). Although PASTEC usually misclassifies fewer TEs

Table 1 Summary of TE sequences found in this work in the *R. marina* transcriptome and genomic TE content compiled from Edwards et al (2018)

Order/type	Number of distinct sequences	% in the genome
Class I		
LINE	1334	4.624
DIRS	1331	0.805
LTR	1004	3.164
PLE	176	0.760
LARD	110	na
TRIM	75	na
SINE	30	0.286
Class II		
TIR	838	18.64
MITE	149	na
Maverick	75	0.021
Helitron	58	0.656
Crypton	12	0.001

na not available

than some other tools (Hoede et al. 2014), it is still possible that some non-TE sequences were misclassified as TE. This condition is particularly higher for non-autonomous categories of retrotransposons LARD (Large Retrotransposon Derivatives) and TRIM (Terminal Repeat Retrotransposons in Miniature) and DNA transposons MITE (Miniature Inverted-repeat Transposable Element), but those represent less than 10% of TE sequences identified.

When we compare the result of TE content (number of distinct TE fragments) from the transcriptome with the proportion of the genome that each TE makes up (Table 1), it seems incongruous, since TIR comprises around 18%, while the others comprise less than 5%. However, it is important to mention that the TE content found in the assembly transcriptome does not take into account the intensity of expression of each TE neither their number of copies since this information is lost in the transcriptome assembly process. Thus, the discrepancies could be related to the proportion of transcriptionally active copies in each group (i.e., there are more transcriptionally active copies of retrotransposons than DNA transposons), or it could be only a matter of sequence diversity in each group (the diversity of transcriptionally active families is higher for retrotransposons than DNA transposons). Moreover, artifacts of transcriptome assembly could generate split fragments that could be found more often for larger TEs as the retrotransposons.

Recently, Pasquesi et al. (2020) have shown that several vertebrate lineages, including amphibious, display high levels of transcripts derived from TEs in somatic and gametic tissues. Even that most transcripts are derived from ancient TE families, the authors suggest the high level of TE-derived RNA in the cytoplasm may have secondary biological relevance.

Here, we also can predict that many or most of these transcripts are derived from truncated copies that are unable to transpose. Still, the identification of such a high number of TE-containing transcripts indicates several copies are proficient in being expressed at some level with potential consequences to the host.

Some TEs were differentially expressed in treatments

As mentioned above, Gardner et al. (2018) has found some DEGs in the comparisons of treatments with control and between the two treatments. Some of these DEGs were not mapped to any known protein in the Uniprot database and remained unidentified by the authors. Interestingly, we found that some of them are derived from TE sequences. In the first comparison (Fig. 1, dataset II), from 155 DEGs, none were classified as a TE. On the other hand, when exposed to immune plus stress challenge (Fig. 1, dataset III), one of the 93 DEGs corresponds to a retrotransposon from order DIRS and was upregulated. In the third comparison (Fig. 1, dataset IV), a fewer number of DEGs were identified and four of them correspond to TE sequences: two retrotransposon families from order PLE (*Penelope-like*), one up and the other downregulated, and two retrotransposon families from order LINE that were downregulated. The expression details are available in Table 2.

Analysis of the genomic copies revealed that most of the DETEs have genomic potentially active counterparts, i.e., structurally complete copies with preserved coding capacity (Fig. 2).

Table 2 Summary of expression result of DETEs identified in this work as compiled from Gardner et al. (2018)

Comparisons	DETE	ID transcript	logFC	P value	FDR	Direction
Dataset III	DIRS family 1	TRINITY_DN110424_c0_g1_i9|m.92261	3.292486	2.87E−05	0.027456212	Upregulated
Dataset IV	PLE family 1	TRINITY_DN112407_c1_g1_i5|m.52987	− 2.64145	9.56E−06	0.031864169	Downregulated
	LINE family 1	TRINITY_DN104205_c0_g2_i1|m.67970	− 2.58467	1.95E−05	0.045786477	Downregulated
	LINE family 2	TRINITY_DN112771_c1_g1_i1|m.78282	−1.58404	6.99E−06	0.026871663	Downregulated
	PLE family 2	TRINITY_DN109415_c2_g2_i1|m.87728	2.964099	1.23E−06	0.008733918	Upregulated

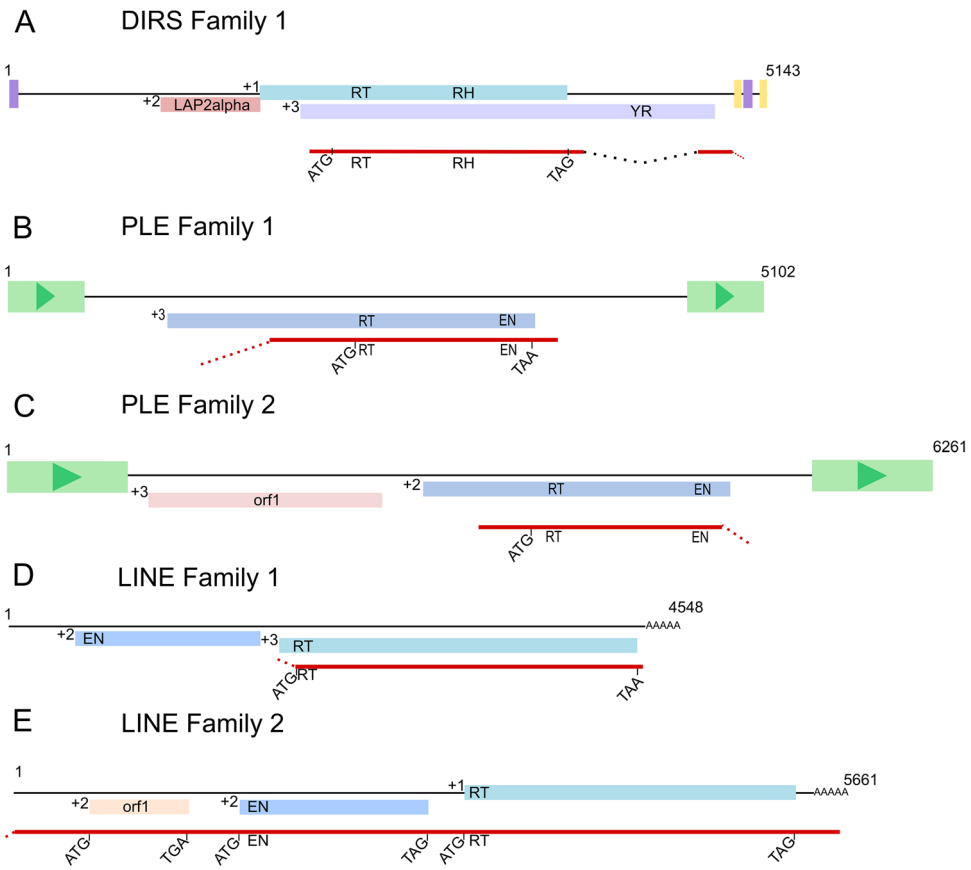

Fig. 2 Representation of genomic counterparts of the *R. marina* DETEs found in this work. Colored boxes represent the ORFs and the reading frame is indicated. Red lines indicate the alignment of assembled transcripts. Red dotted lines in the transcript are non-aligned regions and the dotted black line is a gap in the transcript. The elements and structures were designed nearly in scale. **a** DIRS family 1 containing three ORFs and the ICR. **b** PLE family 1 containing direct LTRs-like (green) and one ORF. **c** PLE family 2 containing direct LTRs-like (green) and two ORFs. ORF1 has no predicted domain. **d** LINE family 1 containing two ORFs and a poly-A tail. **e** LINE family 1 containing three ORFs and a poly-A tail. ORF1 has no predicted domain. Proteins and domains: LAP2alpha corresponds to the domain from Gag-like; RT and RH correspond to the Reverse Transcriptase and RNaseH domains; YR-Tyrosine recombinase holding the DNA_CRE_C domain; EN is the endonuclease domain that is distinct for PLE (GIY-YIG-PLEs) and LINE (L1-EN) elements

The DETE identified in Dataset III belongs to the order DIRS of retrotransposons since they encode a Tyrosine Recombinase (YR) protein instead of a DDE-integrase or endonuclease (EN), a diagnostic feature of this order (Poulter and Goodwin 2005; Poulter and Butler 2015) that is subsequently divided into four superfamilies (Ribeiro et al. 2019). We called it DIRS family 1 (Fig. 2a) and CENSOR result indicates the DIRS-24_XT (XT—from *Xenopus tropicalis*) as the most similar TE in the database (68% aa similarity. This element has the typical structure of members from the superfamily *DIRS-like* (Ribeiro et al. 2019), containing three preserved ORFs with a level of overlapping. The first ORF encodes a Gag-like protein that has the LAP2alpha domain, which is conserved in all *DIRS-like* families from *X. tropical* and *X. laevis* (Camilla Gazolla, personal communication). The second ORF encodes the Reverse Transcriptase (RT) and RNAseH (RH) domains, while the third ORF encodes the YR protein that contains the DNA_CRE_C domain in the C-terminal portion. Elements from superfamily *DIRS-like* have inverted terminal repeat (ITR) and an internal complementary region (ICR) that is complementary to the beginning and end of the element (Poulter and Goodwin 2005). The complete structure is not present in the copies that we have analyzed since we only found the ICR. Thus, these copies are probably no longer mobilized, but we cannot rule out the possibility that potentially active copies were not correctly assembled. The upregulated transcript covers the region corresponding to the RT/RH domains.

The PLE families 1 and 2 correspond to putative actives families (Fig. 2b and c). Elements from order PLE encodes an RT protein that is more closely related to telomerase than to the RT from other retrotransposons and frequently, these elements have direct or inverted long terminal repeats (known as LTR-like) (Wicker et al. 2007). Both PLE families 1 and 2 possess an ORF encoding the RT and EN (GIY-YIG_PLEs) domains and they harbor direct LTR-like structures of around

500 bp and 820 bp, respectively. PLE family 2 has an additional ORF with no predicted domain. Target site duplications (TSDs) were not found for these families suggesting either degeneration or that they do not produce TSDs such as some other PLE elements (Arkhipova et al. 2013). The predicted protein of family 1 has the highest similarity (64%) with Penelope-15_XL (XL – from *X. laevis*) while the proteins from family 2 have the highest similarity with Penelope-8_XT (46.5%) and Penelope -5_DR (59%) (DR—from *Danio rerio*). The assembled transcripts for both families cover the RT and EN domains.

The two last DETEs identified are from order LINE, the classical so-called non-LTR retrotransposons (Fig. 2d and e). Both families encode the RT and EN (L1-EN) domains in different ORFs while family 2 has an additional 5' ORF with no predicted domain. They have a poly-A tail in the 3' end and conserved 11 bp-TSDs were detected. Family 1 has the highest similarity (55%) with L1-11_XT and family 2 has the highest similarity with L1-36_XT (ORF1–64% aa similarity) and L1-51_XT (ORFs 2 and 3–59% aa similarity). The assembled transcript from family 1 covers only the RT domain and the one from family 2 covers the entire element.

Several TEs are shown to be differentially expressed in distinct tissues and/or conditions in different organisms (Horváth et al. 2017; Lerat et al. 2017; Oliveira et al. 2021; Wos et al. 2021). Recently, a single-cell RNA-seq approach indicates that many TEs are specifically expressed in distinct mammal cell types (including neural and immune cell lineages) and some are co-expressed with specific transcription factor genes in the corresponding cell types, suggesting these factors may be responsible for activating TEs (He et al. 2021).

Most DETEs transcripts that we identified contain start and stop codons suggesting they are functional protein-encoding transcripts, although we have no evidence of transposition activity. Further investigation should be carried out to understand the biological significance of the gene expression observed. Lanciano and Cristofari (2020) pointed that distinguishing transcripts derived from autonomous TE from those chimeric or pervasive intergenic transcripts is not a trivial task, but the TE expression itself can significantly influence cell physiology in different ways such as regulating the gene expression and chromatin accessibility and activating cellular signaling pathways.

Thus, finding TEs among the differentially expressed genes *in R. marina* spleen is an exciting result and intensifies the idea that TEs might have a role in rapid response to stress and new environments.

Conclusion

In the last years, a growing number of works are showing that increased TE activity might provide genetic diversification in populations having a role in habitat expansion and adaptation of species (Vieira et al. 1999; Vieira and Biémont 2004; Casacuberta and González 2013; Schrader et al. 2014; Stapley et al. 2015; Kofler et al. 2015; Schrader and Schmitz 2019). The results presented here, show that several distinct families of TEs can be transcribed in *R. marina* under stress conditions. Although we do not have the information concerning the coding conservation of all transcripts, we can expect that some are from active copies and could lead to TE amplification. We can expand the idea that beyond the somatic cells, these TEs could be also activated in the germinative cells generating new herdable insertions with the potential phenotypic consequences. Consistent with this idea, Pasquesi et al. (2020) have shown that recently active TEs are expressed in both germline and somatic tissues across vertebrates.

As we mentioned, Selechnik et al. (2019) observed a reduction in genetic diversity from native to invasive populations of *R. marina* and they suggest the importance of studies on plasticity and epigenetic changes to uncover the mysteries of rapid evolution. In this context, the expression of TEs could lead to regulatory changes helping to generate a rapid response of specimens to the environment. Thus, especially for species like *R. marina* that have a high TE content, these repetitive sequences are key targets for investigation in the context of species adaptation.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10709-021-00130-w.

Acknowledgements This work is a result of practical activities developed together with the students during the discipline "Tópicos especiais em genética III (Elementos Genéticos Móveis)" of Programa de Pós-Graduação em Genética (PPG-GEN)—UFPR. We thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/PROAP – Finance Code 001) for the scholarships provided to ALKL, MZ, JMG, ID, CM, DALPV.

Funding The authors received no specific funding for this work.

Declarations

Conflict of interest All authors declare that they have no conflict of interest.

References

Arkhipova IR, Yushenova IA, Rodriguez F (2013) Endonuclease-containing Penelope retrotransposons in the bdelloid rotifer *Adineta*

vaga exhibit unusual structural features and play a role in expansion of host gene families. Mob DNA 4:19. https://doi.org/10.1186/1759-8753-4-19
Berthelier J, Casse N, Daccord N et al (2018) A transposable element annotation pipeline and expression analysis reveal potentially active elements in the microalga *Tisochrysis lutea*. BMC Genomics 19:378. https://doi.org/10.1186/s12864-018-4763-1
Casacuberta E, González J (2013) The impact of transposable elements in environmental adaptation. Mol Ecol 22:1503–1517. https://doi.org/10.1111/mec.12170
Edwards RJ, Tuipulotu DE, Amos TG et al (2018) Draft genome assembly of the invasive cane toad, *Rhinella marina*. Gigascience. https://doi.org/10.1093/gigascience/giy095
Ellinghaus D, Kurtz S, Willhoeft U (2008) LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. BMC Bioinf 9:18. https://doi.org/10.1186/1471-2105-9-18
Finnegan DJ (1989) Eukaryotic transposable elements and genome evolution. Trends Genet 5:103–107. https://doi.org/10.1016/0168-9525(89)90039-5
Flutre T, Duprat E, Feuillet C, Quesneville H (2011) Considering transposable element diversification in de novo annotation approaches. PLoS ONE. https://doi.org/10.1371/journal.pone.0016526
Gardner S, Assis VR, Zhao H et al (2018) Differential gene expression to an LPS challenge in relation to exogenous corticosterone in the invasive cane toad (*Rhinella marina*). Dev Comp Immunol 88:114–123. https://doi.org/10.1016/j.dci.2018.07.016
Goubert C, Henri H, Minard G et al (2017) High-throughput sequencing of transposable element insertions suggests adaptive evolution of the invasive Asian tiger mosquito towards temperate environments. Mol Ecol 26:3968–3981. https://doi.org/10.1111/mec.14184
Han Y, Wessler SR (2010) MITE-Hunter: a program for discovering miniature inverted-repeat transposable elements from genomic sequences. Nucleic Acids Res 38:e199. https://doi.org/10.1093/nar/gkq862
He J, Babarinde IA, Sun L et al (2021) Identifying transposable element expression dynamics and heterogeneity during development at the single-cell level with a processing pipeline scTE. Nat Commun 12:1456. https://doi.org/10.1038/s41467-021-21808-x
Hoede C, Arnoux S, Moisset M et al (2014) PASTEC: an automatic transposable element classification tool. PLoS ONE 9:e91929. https://doi.org/10.1371/journal.pone.0091929
Horváth V, Merenciano M, González J (2017) Revisiting the relationship between transposable elements and the eukaryotic stress response. Trends Genet 33:832–841. https://doi.org/10.1016/j.tig.2017.08.007
Huang Y, Niu B, Gao Y et al (2010) CD-HIT Suite: a web server for clustering and comparing biological sequences. Bioinformatics 26:680–682. https://doi.org/10.1093/bioinformatics/btq003
Jacques P-É, Jeyakani J, Bourque G (2013) The majority of primate-specific regulatory sequences are derived from transposable elements. PLoS Genet 9:e1003504. https://doi.org/10.1371/journal.pgen.1003504
Kapitonov VV, Jurka J (2008) A universal classification of eukaryotic transposable elements implemented in Repbase. Nat Rev Genet 9:411–2; author reply 414. https://doi.org/10.1038/nrg2165-c1
Kofler R, Nolte V, Schlötterer C (2015) Tempo and mode of transposable element activity in Drosophila. PLoS Genet 11:e1005406. https://doi.org/10.1371/journal.pgen.1005406
Kohany O, Gentles AJ, Hankus L, Jurka J (2006) Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. BMC Bioinf 7:474. https://doi.org/10.1186/1471-2105-7-474
Kosmala G, Christian K, Brown G, Shine R (2017) Locomotor performance of cane toads differs between native-range and invasive populations. R Soc Open Sci 4:170517. https://doi.org/10.1098/rsos.170517
Kosmala GK, Brown GP, Shine R (2020a) Colonization history affects heating rates of invasive cane toads. Sci Rep 10:12553. https://doi.org/10.1038/s41598-020-69529-3
Kosmala GK, Brown GP, Shine R (2020b) Laid-back invaders: Cane toads (*Rhinella marina*) down-regulate their stress responses as they colonize a harsh climate. Glob Ecol Conserv 24:e01248. https://doi.org/10.1016/j.gecco.2020.e01248
Lanciano S, Cristofari G (2020) Measuring and interpreting transposable element expression. Nat Rev Genet 21:721–736. https://doi.org/10.1038/s41576-020-0251-y
Lerat E, Fablet M, Modolo L et al (2017) TEtools facilitates big data expression analysis of transposable elements and reveals an antagonism between their activity and that of piRNA genes. Nucleic Acids Res 45:e17. https://doi.org/10.1093/nar/gkw953
Li Z-W, Hou X-H, Chen J-F et al (2018) Transposable elements contribute to the adaptation of *Arabidopsis thaliana*. Genome Biol Evol 10:2140–2150. https://doi.org/10.1093/gbe/evy171
Marchler-Bauer A, Bryant SH (2004) CD-Search: protein domain annotations on the fly. Nucleic Acids Res 32:W327–W331. https://doi.org/10.1093/nar/gkh454
Mittan CS, Zamudio KR (2019) Rapid adaptation to cold in the invasive cane toad *Rhinella marina*. Conserv Physiol. https://doi.org/10.1093/conphys/coy075
Naito K, Zhang F, Tsukiyama T et al (2009) Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. Nature 461:1130–1134. https://doi.org/10.1038/nature08479
Oliveira DS, Rosa MT, Vieira C, Loreto ELS (2021) Oxidative and radiation stress induces transposable element transcription in *Drosophila melanogaster*. J Evol Biol 34:628–638. https://doi.org/10.1111/jeb.13762
Pasquesi GIM, Perry BW, Vandewege MW et al (2020) Vertebrate lineages exhibit diverse patterns of transposable element regulation and expression across tissues. Genome Biol Evol 12:506–521. https://doi.org/10.1093/gbe/evaa068
Pietzenuk B, Markus C, Gaubert H et al (2016) Recurrent evolution of heat-responsiveness in Brassicaceae COPIA elements. Genome Biol 17:209. https://doi.org/10.1186/s13059-016-1072-3
Poulter RTM, Butler MI (2015) Tyrosine recombinase retrotransposons and transposons. Microbiol Spectr. https://doi.org/10.1128/microbiolspec.MDNA3-0036-2014
Poulter RTM, Goodwin TJD (2005) DIRS-1 and the other tyrosine recombinase retrotransposons. Cytogenet Genome Res 110:575–588. https://doi.org/10.1159/000084991
Price AL, Jones NC, Pevzner PA (2005) De novo identification of repeat families in large genomes. Bioinformatics 21(Suppl 1):i351–i358. https://doi.org/10.1093/bioinformatics/bti1018
Rho M, Tang H (2009) MGEScan-non-LTR: computational identification and classification of autonomous non-LTR retrotransposons in eukaryotic genomes. Nucleic Acids Res 37:e143. https://doi.org/10.1093/nar/gkp752
Ribeiro YC, Robe LJ, Veluza DS et al (2019) Study of VIPER and TATE in kinetoplastids and the evolution of tyrosine recombinase retrotransposons. Mob DNA 10:34. https://doi.org/10.1186/s13100-019-0175-2
Rollins LA, Richardson MF, Shine R (2015) A genetic perspective on rapid evolution in cane toads (*Rhinella marina*). Mol Ecol 24:2264–2276. https://doi.org/10.1111/mec.13184
Rozewicki J, Li S, Amada KM et al (2019) MAFFT-DASH: integrated protein sequence and structural alignment. Nucleic Acids Res. https://doi.org/10.1093/nar/gkz342
Salazar M, González E, Casaretto JA et al (2007) The promoter of the TLC1.1 retrotransposon from Solanum chilense is activated

by multiple stress-related signaling molecules. Plant Cell Rep 26:1861–1868. https://doi.org/10.1007/s00299-007-0375-y

Salces-Ortiz J, Vargas-Chavez C, Guio L et al (2020) Transposable elements contribute to the genomic response to insecticides in *Drosophila melanogaster*. Philos Trans R Soc B Biol Sci 375:20190341. https://doi.org/10.1098/rstb.2019.0341

Schrader L, Schmitz J (2019) The impact of transposable elements in adaptive evolution. Mol Ecol 28:1537–1549. https://doi.org/10.1111/mec.14794

Schrader L, Kim JW, Ence D et al (2014) Transposable element islands facilitate adaptation to novel environments in an invasive species. Nat Commun 5:5495. https://doi.org/10.1038/ncomms6495

Selechnik D, Richardson MF, Shine R et al (2019) Increased adaptive variation despite reduced overall genetic diversity in a rapidly adapting invader. Front Genet. https://doi.org/10.3389/fgene.2019.01221

Smit A, Hubley R, Green P (1996) RepeatMasker. http://www.repeatmasker.org

Stapley J, Santure AW, Dennis SR (2015) Transposable elements as agents of rapid adaptation may explain the genetic paradox of invasive species. Mol Ecol 24:2241–2252. https://doi.org/10.1111/mec.13089

Sundaram V, Wysocka J (2020) Transposable elements as a potent source of diverse cis -regulatory sequences in mammalian genomes. Philos Trans R Soc B Biol Sci 375:20190347. https://doi.org/10.1098/rstb.2019.0347

Sundaram V, Cheng Y, Ma Z et al (2014) Widespread contribution of transposable elements to the innovation of gene regulatory networks. Genome Res 24:1963–1976. https://doi.org/10.1101/gr.168872.113

Vieira C, Biémont C (2004) Transposable element dynamics in two sibling species: *Drosophila melanogaster* and *Drosophila simulans*. Genetica 120:115–123. https://doi.org/10.1023/b:gene.0000017635.34955.b5

Vieira C, Lepetit D, Dumont S, Biémont C (1999) Wake up of transposable elements following *Drosophila simulans* worldwide colonization. Mol Biol Evol 16:1251–1255. https://doi.org/10.1093/oxfordjournals.molbev.a026215

Wenke T, Döbel T, Sörensen TR et al (2011) Targeted identification of short interspersed nuclear element families shows their widespread existence and extreme heterogeneity in plant genomes. Plant Cell 23:3117–3128. https://doi.org/10.1105/tpc.111.088682

Wicker T, Sabot F, Hua-Van A et al (2007) A unified classification system for eukaryotic transposable elements. Nat Rev Genet 8:973–982. https://doi.org/10.1038/nrg2165

Wos G, Choudhury RR, Kolář F, Parisod C (2021) Transcriptional activity of transposable elements along an elevational gradient in *Arabidopsis arenosa*. Mob DNA 12:7. https://doi.org/10.1186/s13100-021-00236-0

Yang L, Bennetzen JL (2009) Structure-based discovery and description of plant and animal Helitrons. Proc Natl Acad Sci USA 106:12832–12837. https://doi.org/10.1073/pnas.0905563106