Genomics 113 (2021) 4163–4172

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Comparative genomics of tadpole shrimps (Crustacea, Branchiopoda, Notostraca): Dynamic genome evolution against the backdrop of morphological stasis

Andrea Luchetti [a,*], Giobbe Forni [a,1], Jacopo Martelossi [a], Castrense Savojardo [b], Pier Luigi Martelli [b], Rita Casadio [b], Alyza M. Skaist [c], Sarah J. Wheelan [c], Barbara Mantovani [a]

[a] Department of Biological, Geological and Environmental Sciences, University of Bologna, via Selmi 3, 40126 Bologna, Italy
[b] Biocomputing Group, Department of Pharmacy and Biotechnology, University of Bologna, Italy
[c] Department of Oncology, Sidney Kimmel Comprehensive Cancer Center, The Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

ARTICLE INFO

Keywords:
Arthropoda
Genome sequencing
Living fossils
Parthenogenesis
Tadpole shrimps
Substitution rate

ABSTRACT

This analysis presents five genome assemblies of four Notostraca taxa. Notostraca origin dates to the Permian/Upper Devonian and the extant forms show a striking morphological similarity to fossil taxa. The comparison of sequenced genomes with other Branchiopoda genomes shows that, despite the morphological stasis, Notostraca share a dynamic genome evolution with high turnover for gene families' expansion/contraction and a transposable elements content comparable to other branchiopods. While Notostraca substitutions rate appears similar or lower in comparison to other branchiopods, a subset of genes shows a faster evolutionary pace, highlighting the difficulty of generalizing about genomic stasis *versus* dynamism. Moreover, we found that the variation of *Triops cancriformis* transposable elements content appeared linked to reproductive strategies, in line with theoretical expectations. Overall, besides providing new genomic resources for the study of these organisms, which appear relevant for their ecology and evolution, we also confirmed the decoupling of morphological and molecular evolution.

1. Introduction

The class Branchiopoda consists of small crustaceans distributed world-wide and living in fresh waters, including extreme habitats like temporary ponds and hypersaline lakes, with few species inhabiting marine environment [8,20,67]. One of the most interesting aspects of these crustaceans is found in their adaptations to ephemeral environments, such as resting-eggs: these are drought-resistant and can survive years before hatching. As they usually do not hatch at the same time, they constitute an egg bank composed by different generations and genotypes: the diversity retained by these egg banks could, therefore, facilitate their resilience to environmental changes [7].

The deep phylogeny and the systematics of the class are the subject of a long-standing debate: although recent phylogenomic analyses allowed to draw a clearer picture of the relationships among major clades, with Anostraca sister to all other branchiopod taxa and Notostraca sister to Diplostraca (which includes Laevicaudata, Spinicaudata and Cladoceromorpha; [71]), the higher order systematics of Branchiopoda is far from being clearly established. Four major monophyletic taxa are currently ascribed to this class: the Anostraca (fairy shrimps), Laevicaudata (clam shrimps), Onychocaudata (Spinicaudata, clam shrimps, + Cladoceromorpha, water fleas), and Notostraca (tadpole shrimps) [60,71]. Their origin dates back to the Middle Cambrian, ~500 million years ago (Mya), and the four main lineages were already established by the Early Silurian (~450 Mya; [76]). The order Notostraca includes the two extant genera *Triops* and *Lepidurus* and, together with the fossil group Kazacharthra, they form the Calmanostraca: despite the apparent simplicity of their relationship, the consideration of fossil taxa complicates the phylogeny of the group [86]. Overall, Notostraca is known to be an ancient lineage as suggested by the analysis of Permian (~300–250 Mya) fossils *Triops cancriformis permiensis* (recently elevated to the species status as *Triops permiensis* [24] [stat. nov.]; [39]) and *Lepidurus occitaniacus* [24], and of the Upper Devonian (~420–360 Mya) taxon *Strudops goldenbergi* [43].

* Corresponding author.
E-mail address: andrea.luchetti@unibo.it (A. Luchetti).
[1] Present address: Dipartimento di Scienze Agrarie e Ambientali - Università degli Studi di Milano, Via Celoria 2, 20,133 Milano, Italy.

https://doi.org/10.1016/j.ygeno.2021.11.001
Received 14 May 2021; Received in revised form 26 October 2021; Accepted 2 November 2021
Available online 5 November 2021

The remarkable morphological conservation of extant Notostraca species, closely resembling fossil taxa, is one of the most interesting aspects of tadpole shrimps and contributed to the attribution of the controversial designation of "living fossils" [21]. This concept was first introduced by Darwin in its *Origin of Species* and, since then, it became a common epithet for organisms mainly showing striking morphological conservatism and/or poor lineage diversity. For example, beside tadpole shrimps, the definition of "living fossil" is used for other metazoans like the coelacanth, the tuatara, the horseshoe crab, and the platypus [45]. In line with this notion, the morphological stasis has been correlated with a slower molecular evolution in the coelacanth [1] and in the tuatara [25], although for the latter there is not a consensus [29]. However, the "part-whole ambiguity" problem (*i.e.*, considering some "parts" of an organism to draw conclusions about the "whole" organism) has been often overlooked: for example, many "living fossils" are actually characterized by both ancient and more recently originated traits [45]. In the coelacanth, for instance, although a genome-wide lower nucleotide substitutions rate has been observed with respect to other fishes [1], it is also true that some genes showed higher substitution rates. Moreover, transposable element (TE) content is far from being static but they, instead, proved to be active in a recent past [11,58]. When considering Notostraca, the notion of living fossil has been challenged on a phylogenetic ground, showing that the whole order Notostraca underwent a recent evolutionary radiation [53]. On the other hand, recent genomic analyses of two *Lepidurus* genomes indicated a lower rate of molecular evolution with respect to *Daphnia* genomes [68]. Also at the mitochondrial genome level, Notostraca showed significantly lower evolutionary rate than other branchiopods, but this has been correlated with an ancestral change in nucleotide composition in Anostraca and Onychocaudata [46].

In this work, we report about the sequencing and analysis of five draft genomes of four tadpole shrimp taxa: *Lepidurus apus apus, Lepidurus couesii*, *Triops longicaudatus* and *Triops cancriformis*. For the latter species, two samples were sequenced: one belonging to the Italian population of Novara, characterized by a parthenogenetic reproduction, and the other collected from Espolla (Girona), in Northern Spain, known to be a bisexually reproducing population [38]. These data add, therefore, to the so far only sequenced *T. cancriformis* from a Japanese, putative hermaphroditic population [31,79]. Notostraca species have very small genomes, probably the smallest ones among crustaceans, with an average C-value (*i.e.*, the amount of DNA contained in the nucleus of a haploid cell) of ~0.11 pg [34]: such limited genome size makes these organisms particularly suitable to be completely sequenced with small efforts [68]. In addition to notostracan samples, we also provide here the draft genome of the spinicaudatan *Leptestheria dahalacensis*, collected in Italy (Isola della Scala, Verona), to further improve the branchiopod genomes dataset for the comparative analyses. Our aim, besides contributing to the genomic resources available for these organisms and laying the ground plan for further studies on their ecology and evolution, is also to provide the first comparative analysis of Branchiopoda genomes comprising taxa distributed across the main lineages of this class. These analyses showed that the well-known morphological conservatism of Notostraca contrasts with a dynamic genome turnover.

2. Material and methods

2.1. Sample collection, DNA extraction and sequencing

Total DNA was extracted from individual specimens using the DNA extraction kit (STRATEC), after dissection for gut removal. Whole genome sequencing has been carried out on Illumina HiSeq2000 platform on parthenogenetic *T. cancriformis* (Johns Hopkins University Experimental and Computational Genomics Core, SKCCC, USA; paired end, 2 × 100 bp; insert size = 150 bp) and on HiSeqX platform on bisexual *T. cancriformis*, *T. longicaudatus, L. apus apus, L. couesii* and *Leptestheria dahalacensis* (Macrogen Inc., South Korea; paired end, 2 × 150 bp; insert size = 350 bp).

2.2. Genome assembly and annotation

Quality trimming and Illumina adapter removal have been carried out using Trimmomatic v. 0.39 [5] requiring a minimum quality score of 20 and a minimum read length after trimming of 30 bp. Assembly of Notostraca genomes has been performed using ABySS assembler version 1.5.8 [32] with default parameters. The genome of the spinicaudatan *L. dahalacensis* was, instead, assembled using Platanus Genome Assembler [36]. The optimal value for the *k-mer* size has been determined by testing different values in the range 32–128 bp. The optimal *k-mer* for each genome has been then selected according to the N50 statistics computed on the resulting assembly. Only scaffolds longer than 1000 bp were retained in the final assemblies. After assembly, each genome was processed using the Redundans tools v.0.14a [62], with default parameters, to perform: i) redundancy reduction at 51% identity and 80% overlap and 200 bp minimum contig length; ii) short-read scaffolding using the respective short-read library; iii) two iterations of gap closing using the reads libraries.

The presence of possible contaminants in the assemblies was assessed through BlobTools v. 2 [42]. A first assembly of each genome was BLASTed against the NCBI *nt* database with the *blastn* parameters and a stringent e-value of 1×10^{25}, and then annotated using the *bestsum taxrule* at the Phylum level. Reads coverage and mapping statistics were calculated aligning the short reads with Minimap2 [44] to the corresponding assembly. Reads that mapped to regions annotated as Bacteria or Proteobacteria were removed, and the remaining were re-assembled as previously described. A second BlobTools run was performed on the new assemblies to check for permanence of contaminant reads.

The completeness of the assembled genomes was assessed by checking for the presence of conserved representative genes using BUSCO v. 5 [94], with the Arthropoda orthologue set (N = 1013), implemented by the gVOLANTE online platform v. 1.0 (https://gvolante.riken.jp/; [59]; last accessed in August 2021).

Annotation of transposable elements (TEs) was carried out using a *de novo* approach. RepeatModeler v. 1.0 [72] was run with default parameters. Moreover, TE-specific search has been carried out using LTR_Finder v. 1.06 [91] and SINE_Scan v. 1.1.1 [51], with default parameters. Final TE libraries were obtained merging the TE consensus sequences from the three software used and redundancy was reduced with the CD-HIT program [23], setting 80% of sequence identity and 80% of sequence overlap. For *T. cancriformis* genomes, all searches were merged in a single library. Reduced TE libraries were then used as a database for RepeatMasker v. 4.0 [73] to calculate the relative abundance of TE families and the repeat landscape (*i.e.*, the distribution of the Kimura divergence of each TE copy from the respective consensus sequence). To perform a comparative analysis, the same procedure was carried out for all other available branchiopod genomes. A hierarchical clustering of TE content was then performed using the ComplexHeatmap R package v. 3.12 [28] with a distance matrix based on Kendall's τ.

After repeat masking, protein coding genes were predicted using Augustus v. 2.5 [74], with gene models previously built on *Lepidurus apus lubbocki* and *Lepidurus arcticus* in Savojardo et al. [68].

2.3. Orthologous gene prediction, phylogenomic inference and comparative analyses

In order to perform genome comparative analyses with other branchiopods, assemblies and gene annotations of seven additional branchiopod representatives were obtained from the NCBI database (*Daphnia pulex*, *Daphnia magna*, *Eulimnadia texana*, a Japanese sample of *T. cancriformis*, *L. arcticus* and *L. apus lubbocki*; last accessed on June 2020) and from the Korea Polar Research Institute (*Artemia franciscana* at https://antagen.kopri.re.kr/project/genome_info_iframe.php?Code=AF01#, last accessed August 2021; [35]) (Suppl. Table S1). Moreover, two hexapod species were used as

outgroups: the red flour beetle *Tribolium castaneum* (Insecta, Coleoptera; Genbank acc. no. GCA_000002335.3; *Tribolium* Sequencing Consortium, [81]) and the springtail *Orchesella cincta* (Entognatha, Collembola; Genbank acc. no. GCA_001718145.1; [18]).

Clustering of orthologous gene (OG) families has been carried out using Orthofinder v. 1.0.6 [17], with default parameters. OGs consisting of single-copy genes found in at least 70% of taxa (*i.e.*, at least 10 species) and at least 300 nucleotides long were retained for downstream phylogenomic analyses. Protein alignments were obtained using MAFFT v. 7 with the *–auto* strategy [37] and then retro translated to nucleotide alignments using pal2nal v. 14 [77]. The Maximum Likelihood phylogenetic inference was carried out using IQ-TREE v. 2.0.3 [56] with ModelFinder model selection. Nodal support was assessed with 1000 Ultrafast Bootstrap approximation (UFBoot) and 1000 Shimodaira-Hasegawa approximate Likelihood Ratio Test (sh-aLRT) replicates. Moreover, genes and sites concordance factors (gCF and sCF, respectively; [57]) were calculated with IQ-TREE to analyze the support of the obtained species tree at genes and sites level. Subsequently we carried out phylogenetic dating of the inferred tree using the least square dating (LSD2) method [80], implementing a lognormal relaxed clock, with the nodes constrained on fossil records as indicated in Suppl. Table S2.

For the estimation of differential nucleotide substitutions rate across clades, we used the software package RRTree v. 1.1 [65] with default parameters. Basically, the algorithm implements the relative-rate test [89] at lineage level [64], without the need to specify a tree topology. RRTree compares nucleotide or amino acid substitution rates between sequences grouped in defined lineages with respect to given outgroup sequences and calculate the exact probability for significance of observed differences [65]. Given that RRTree requires at least one sample for each tested lineage (including the outgroups), the analysis was restricted to 698 OGs.

Analyses of selective pressure were carried out considering the ω parameter, calculated as the ratio (dN/dS) between the proportion of non-synonymous changes over non-synonymous sites (dN) and the proportion of synonymous changes over synonymous sites (dS). Using the species tree topology, with gene-optimized branch lengths, we inferred a model with a single ω value shared by all branches in the phylogeny and an alternative model with an ω value inferred for each clade (or branch when comparing species pairs). The two models were then compared using a Likelihood Ratio Test and the resulting *p*-values were adjusted using FDR correction. The whole analysis was performed using *codeml* of the PAML v. 4.9j package [92], as implemented in BASE v. 1.0 [93]. To avoid possible biases in ω calculation, we excluded values having $dS < 0.01$, which may lead to inaccurate ω estimation, and $dS > 2$, that may imply potential substitutions saturation [85]. Discordance between gene trees and the species tree can underlie a wide range of technical and biological phenomena - such as sequence misalignment, non-orthology and incomplete lineage sorting - which may bias evolutionary rate inference (*e.g.,* [54]). As such, we only considered genes that have a tree topology identical to that of the species tree, as determined by having a normalized Robinson-Foulds distance of 0 (nRF, [30]). For the orders-level comparison and due to the requirement of BASE, we only considered OGs with at least two species per clade, limiting the comparison to 676 OGs.

To perform a gene family expansion and contraction analysis, we used CAFE v. 5 [55] with a birth and death rate model estimated over the time-calibrated species tree. Basically, CAFE infers ortholog groups (hence indicated as gene families for consistency with CAFE terminology) already present at the root of the time-tree, based on a parsimony principle, and analyzes the extent of contraction/expansion of each of them along the phylogeny. In particular, CAFE estimates the number of rapidly evolving gene families, defined as the gene families with a significant shift in their evolutionary rate (*i.e.,* expansion or contraction) with respect to a given rate, indicated as λ; rapidly evolving gene families are identified on branches using a Viterbi assignment to a node and computing a *p*-value for the transition from parent to child node (Viterbi $p < 0.05$; [14]). To consider assembly errors in the dataset, models were corrected with an error distribution inferred by the software.

Nucleotide divergence has been calculated using distmat from the EMBOSS package [63] on the single-copy genes shared by Notostraca and separately by the two genera *Lepidurus* and *Triops*, respectively.

3. Results

3.1. Genome assembly and genes annotation

Sequenced reads were cleaned from adapters, trimmed from low-quality bases, and checked for contamination before the final assembly. This led to a reduction of raw data ranging from 0.5% (*L. apus apus*) to 13% (*T. cancriformis* ITA and *L. dahalacensis*). Considering the Notostraca average genome size (~107.5 Mb; [34]), a theoretical genome coverage ranging from 661× (*L. dahalacensis*) to 1129× (*L. couesii*) was obtained (Suppl. Table S3).

Optimal genome assemblies were found after checking different *k-mer* lengths and selecting the best one based on N50 statistics: this led to the use of different *k-mer* lengths among the reconstructed genomes (Table 1). Genomes assembly lengths of *L. couesii* and *L. apus apus* species resulted in 74.1 Mb and 78.8 Mb (Table 1), respectively. On average, *Triops* assemblies were slightly larger: the smallest being the *T. longicaudatus* one (83.7 Mb) and the larger being the *T. cancriformis* ESP one (115.2 Mb) (Table 1). The *L. dahalacensis* genome resulted 103.5 Mb long (Table 1).

The quality of assembled genomes was then checked for completeness using the arthropods' gene set of BUSCO ($N = 1013$ orthologs): reported statistics indicated the presence of 92.89%–97.63% complete genes and 96.25%–98.41% complete+partial genes (Table 2). Gene predictions led to the annotation of 10,109 (*T. longicaudatus*) to 16,336 (*L. dahalacensis*) genes (Table 1).

3.2. Phylogenomics and genome evolution

Our ortholog search among the 15 genomes (branchiopods+outgroups) resulted in 21,121 orthologs groups (OGs). We selected OGs containing single-copy genes for at least 10 taxa ($N = 1001$) to build a Maximum Likelihood phylogenetic tree. Nodal supports of the resulting tree were maximum for all nodes; Notostraca taxa formed a clearly monophyletic clade in sister relationship with Onychocaudata (Fig. 1a). Concordance factors resulted lower at deepest nodes, with gCF = 55.5% and sCF = 35.4% supporting the sister relationship of Anostraca with the other branchiopods, and gCF = 47.7% and sCF = 38.6% supporting the split between Notostraca and Onychocaudata (Suppl. Fig. S1). In the inferred time-tree, the *Triops* clade originated about 48.1 Mya (C.I. = 34.5–61.4 Mya) while the *Lepidurus* one dates back to 20.7 Mya (C.I. = 8.99–29.1 Mya). The overall nucleotide substitution rate was 2.06×10^{-3} substitutions/site/million year (C.I. = 1.80×10^{-3} - 2.35×10^{-3} substitutions/site/million year), but it widely varied among branches (Fig. 1a).

The analysis of transposable elements (TEs) of newly sequenced notostracan genomes, excluding the unclassified repeats, showed the occupancy varying between 2.2% in *T. longicaudatus* to 8.7% in *T. cancriformis* ESP. The genomic landscape was dominated by both DNA and LTR elements, with a small percentage of LINEs and the absence of SINEs (Fig. 1b; Suppl. Table S4). These features are shared with previously sequenced notostracan genomes of *L. apus lubbocki*, *L. arcticus* and the Japanese *T. cancriformis* sample. In *Daphnia* spp., *E. texana* and *A. franciscana*, LTR elements represented the great majority of retrieved TEs, while the *L. dahalacensis* genome showed a higher contribution of DNA elements (Fig. 1b; Suppl. Table S4). Notostraca showed the lowest percentages of TEs occupancy with respect to other branchiopods, except for *D. magna* and *L. dahalacensis* (Fig. 1b; Suppl. Table S4). When TEs occupancy was analyzed at the family level, a generally more complex pattern emerged (Fig. 2). The clustering of genomes based on

Table 1
General statistics of obtained optimal assemblies.

Species	*k-mer*	No. Contigs (N50) †	No. Scaffold (N50) †	Length (bp)	GC (%)	No. of genes
Lepidurus apus apus	96 bp	4988 (81,312)	2573 (136,761)	78,857,107	41.5	11,441
Lepidurus couesii	88 bp	10,883 (16,518)	3516 (42,063)	74,083,348	41.7	11,870
Triops longicaudatus	96 bp	1954 (113,776)	636 (380,856)	83,658,226	40.2	10,109
Triops cancriformis (ESP)	64 bp	21,317 (13,074)	10,570 (22,383)	115,234,264	41.3	14,801
Triops cancriformis (ITA)	64 bp	11,354 (18,497)	7548 (29,572)	99,177,902	41.2	12,747
Leptestheria dahalacensis	32 bp	35,651 (3137)	7559 (29,515)	103,541,535	40.5	16,336

† =length ≥ 1.0 kb.

Table 2
Number of arthropods' orthologs set found in genome assemblies with BUSCO v. 5.

Species	Complete (%)	Complete + Partial (%)	Missing (%)
Lepidurus apus apus	989 (97.63%)	1004 (99.11%)	9 (0.89%)
Lepidurus couesii	972 (95.95%)	998 (98.52%)	15 (1.48%)
Triops longicaudatus	986 (97.33%)	1007 (99.41%)	6 (0.59%)
Triops cancriformis (ESP)	971 (95.85%)	997 (98.42%)	16 (1.58%)
Triops cancriformis (ITA)	979 (96.64%)	1003 (99.01%)	10 (0.99%)
Leptestheria dahalacensis	941 (92.89%)	975 (96.25%)	38 (3.75%)

TEs' family abundance revealed a pattern consistent with phylogenetic relationships among the clades Anostraca, Onychocaudata and Notostraca. Moreover, within Onychocaudata, the two *Daphnia* genomes, on one side, and the two spinicaudatan genomes (*E. texana* and *L. dahalacensis*), on the other side, clustered together and resulted in sister relationship, according to species phylogeny. On the other hand, the relationships between and within genera in Notostraca are not maintained, except for the monophyletic group of *T. cancriformis* genomes (Fig. 2).

The relative abundance of TE families resulted significantly different both among *Triops* and among *Lepidurus* genomes (Friedman test, $p < 0.001$ for both comparisons). The *T. cancriformis* ESP sample showed the highest TE content (8.7%) and the *T. longicaudatus* genome has the lowest TE occupancy (2.2%); in *Lepidurus* genomes, the TE content ranged from 1.9% in *L. arcticus* to 6.4% in *L. apus apus* (Suppl. Table S4). All pairwise comparisons among *Triops* samples were significant (post-hoc pairwise Mann-Whitney test, with Bonferroni correction, $p < 0.001$; Suppl. Table S5). On the other hand, pairwise comparisons were found significant (post-hoc pairwise Mann-Whitney test, with Bonferroni correction, $p < 0.01$) between both *L. apus apus* and *L. apus lubbocki vs* both *L. arcticus* and *L. couesii*, but no differences were found between *L. arcticus* and *L. couesii* or between *L. apus apus* and *L. apus lubbocki* (Suppl. Table S6).

3.3. Analysis of substitutions rate and gene families' evolutionary dynamics

The relative-rate test implemented in RRTree showed that the nucleotide substitutions rate of the Anostraca lineage is significantly higher than those calculated for both Onychocaudata and Notostraca (exact probability, $p < 0.001$ for both comparisons); the same results were obtained analyzing amino acid sequences ($p < 0.001$ for both comparisons). No differences were observed between Onychocaudata and Notostraca. Yet, a gene-by-gene approach revealed a slightly different picture. Out of the 698 OGs analyzed, 230 *A. franciscana* genes and 168 onychocaudatan genes showed a significantly higher nucleotide substitutions rate than their notostracan ortholog. However, 9 and 70 notostracan genes showed a substitutions rate higher than their orthologs in *A. franciscana* and Onychocaudata, respectively (Fig. 3).

The GC content of analyzed genes varies among lineages, with Onychocaudata showing the highest proportion (GC = 46.8% ± 0.1%), and Anostracan the lowest one (GC = 40.3% ± 0.1%) (Suppl. Fig. S2). Overall, the GC content among lineages appeared significantly different (Kruskal-Wallis test, $p < 0.001$; post-hoc pairwise Mann-Whitney test, with Bonferroni correction, $p < 0.05$ for all comparisons). Though, when the GC content of genes with higher substitutions rate was compared to the GC proportion of genes not showing differential rates no significant differences were found (Suppl. Table S7).

The natural selection analysis indicated the alternative model of a clade-specific ω as the most likely for 452 OGs. When excluding dS values that may bias ω estimates and considering only those genes with gene tree identical to the species tree (nRF = 0), the count was 78 OGs. Both Onychocaudata and Notostraca showed a single (different) gene with $\omega > 1.0$ ($\omega = 1.389$ and $\omega = 1.097$, respectively). The average ω values of Anostraca, Onychocaudata and Notostraca resulted significantly different from each other, with ω values equal to 0.246, 0.335 and 0.274, respectively (Friedman test, $p < 0.001$; post-hoc pairwise Mann-Whitney test, $p < 0.001$ for all comparisons; Suppl. Fig. S3). However, genes with higher substitutions rate did not show significant differences in the average ω values in any considered comparison (Suppl. Table S8).

Overall, CAFE analyzed 7591 gene families which were reconstructed as present at the root of the species tree. We tested two different models, corrected with an error rate of 5% estimated by the software, with different rates of evolution (λ). First, we considered a single $\lambda = 8.57 \times 10^{-4}$ ($-lnL = 159{,}007.0$) for the entire tree. Then, we compared it to a two-rates model, with a λ value for Notostraca ($\lambda = 1.09 \times 10^{-3}$) and one for the remaining of the tree ($\lambda = 8.46 \times 10^{-4}$) ($-lnL = 158{,}885.0$). The Likelihood Ratio Test showed the two-rates model as the most likely ($p < 0.0001$), therefore indicating Notostraca as having a significantly higher rate of gene family evolution with respect to other branchiopods. In fact, *Triops* and *Lepidurus* showed, on average, a higher number of rapidly evolving gene families (Fig. 1a).

3.4. Notostraca divergence analysis

The overall divergence between *Triops* and *Lepidurus* species, calculated over 2677 orthologs, was 25.05% ± 0.13%, with values ranging from 6.20% (*T. longicaudatus vs L. apus lubbocki*) to 64.60% (*T. cancriformis* ESP *vs L. couesii*) (Suppl. Fig. S4).

To analyze the divergence among the bisexual *T. cancriformis* ESP, the parthenogenetic *T. cancriformis* ITA and the putative hermaphrodite *T. cancriformis* JAP, we considered all OGs which included single-copy genes for all the three samples and for the outgroup species *T. longicaudatus* ($N = 4403$). The ITA and JAP samples diverged by 0.24% ± 0.01%, while both diverged from the ESP sample by 1.98% ± 0.04% and 1.97% ± 0.03%, respectively. The divergence between *T. longicaudatus* and *T. cancriformis* samples ranged from 14.27% ± 0.08% (*vs* the JAP sample) to 14.32% ± 0.08% (*vs* the ESP sample) (Fig. 4a).

To gain more details on the divergence between *T. cancriformis* ESP *vs* ITA + JAP, we searched for possible selective events along the respective branches. After filtering for dS and for gene trees' nRF, 21 OGs

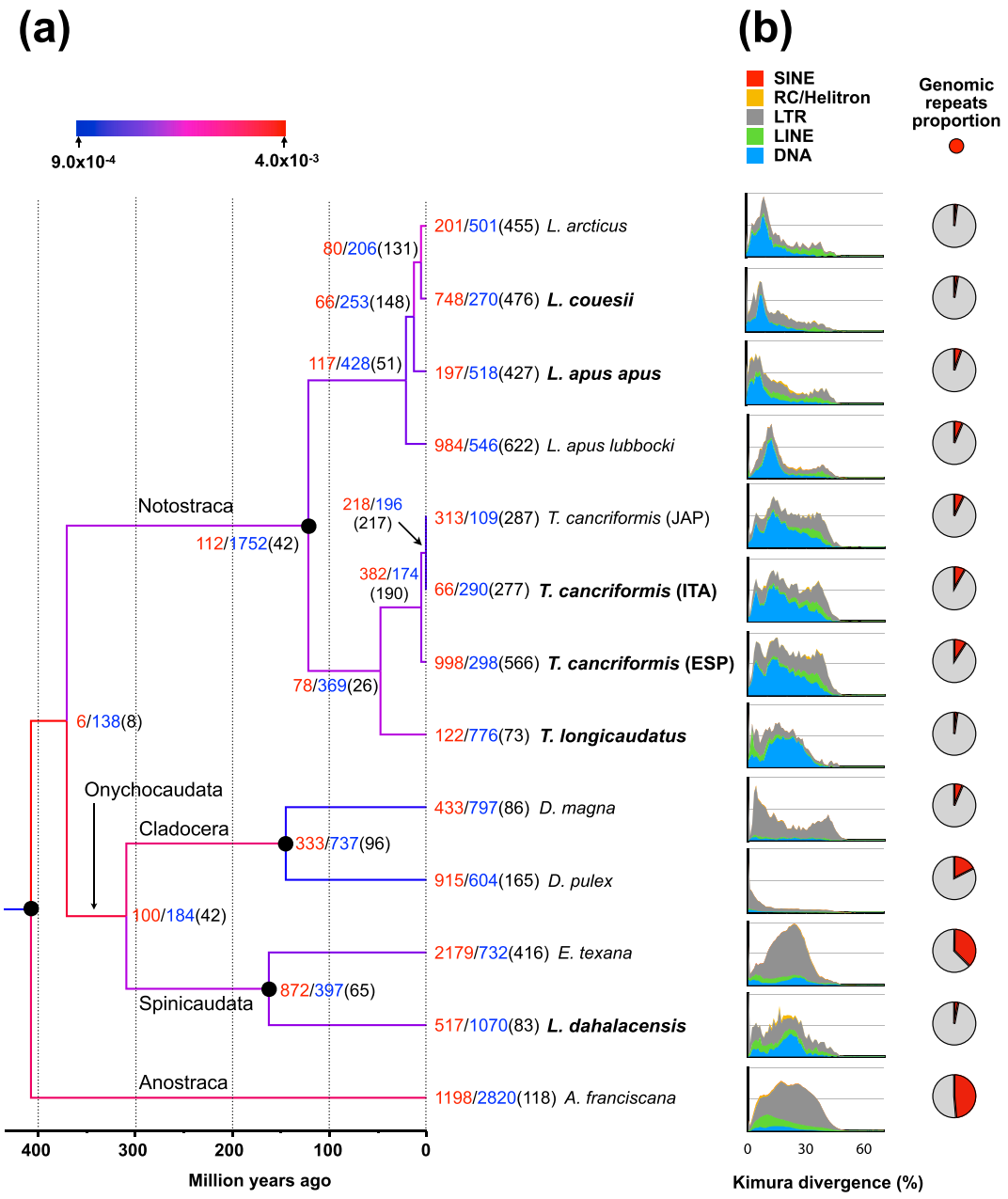

Fig. 1. Phylogenomics and transposable elements content. (a) Time-calibrated phylogeny based on 1001 ortholog groups. Outgroups are omitted for graphical purposes. Newly sequenced genomes are indicated in bold. All nodes received maximum support (sh-aLRT = 100; UFBoot = 100). Branches are coloured according to the relative nucleotide substitutions rate (number of substitutions/site/Million year), as indicated in the upper left legend. Black dots indicate age calibration points. Numbers at nodes and tips represent estimated gene families' expansions (red), contractions (blue), and the number of estimated rapidly evolving families (between parentheses). (b) Transposable elements landscapes and their occupancy in the respective genomes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

were analyzed. Both lineages showed a single gene with $\omega > 1.0$, but they also showed approximately the same average ω (ESP, $\omega = 0.594 \pm 0.226$; ITA + JAP, $\omega = 0.425 \pm 0.06$; paired Wilcoxon test, $p = 0.562$) (Suppl. Fig. S5).

The divergence among *Lepidurus* genomes was estimated on 4195 orthologs, selected as for *Triops* samples. *Lepidurus arcticusand L. couesii* resulted the most similar, with an observed divergence of 2.28% ± 0.03%, while *L. apus lubbocki* was the most differentiated, with divergence ranging from 6.05% ± 0.05% (*vs L. apus apus*) to 6.55% ± 0.06% (*vs L. arcticus*; Fig. 4b).

Since *L. arcticus* and *L. couesii* are sister species but were collected in completely different environments (Iceland *vs* Southern Italy), we tested for possible selective events along the respective branches. Of the analyzed OGs, 117 fitted the model with branch-specific ω and, after filtering, a data set of 22 OGs was retained. Although, *L. arcticus* showed a single gene with $\omega > 1.0$, the average ω values on the two branches resulted nearly identical (*L. arcticus* $\omega = 0.369 \pm 0.060$; *L. couesii*, $\omega = 0.394 \pm 0.050$; paired Wilcoxon test, $p = 0.074$) (Suppl. Fig. S5).

4. Discussion

In this work, we report the sequencing of five draft genomes from the four tadpole shrimp taxa *L. apus apus*, *L. couesii*, *T. longicaudatus* and *T. cancriformis*. Moreover, to enrich the Branchiopoda genomic

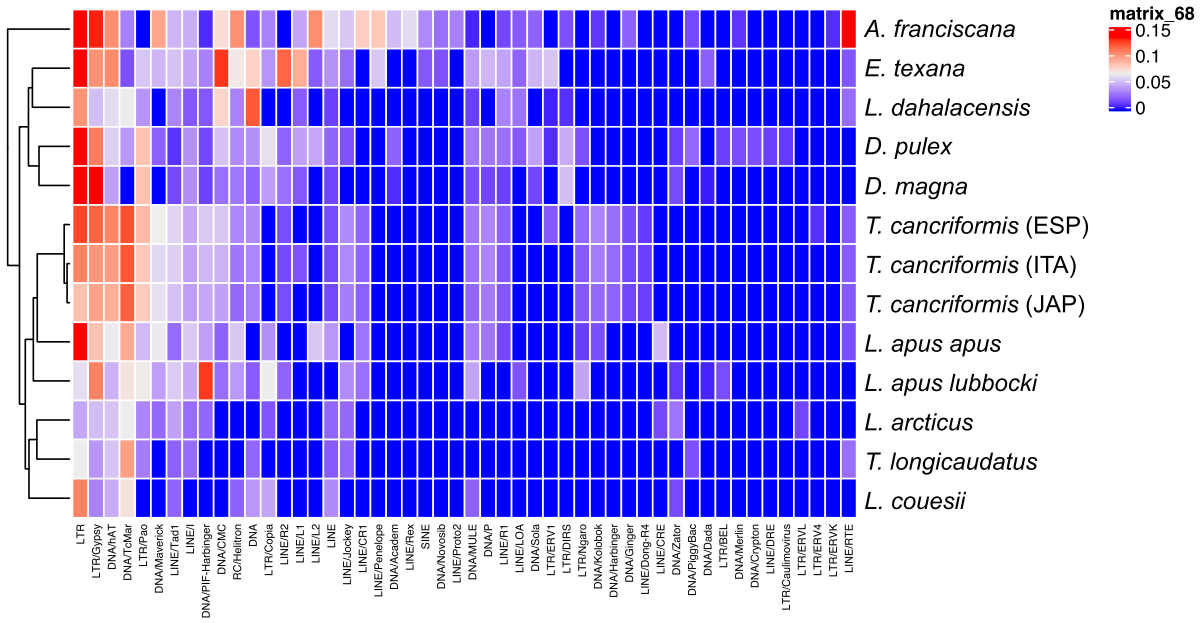

Fig. 2. Analysis of transposable element (TE) families. The heatmap shows the distribution of TE families' occurrence in analyzed genomes; the cladogram shows genomes clustering based on Kendall's τ.

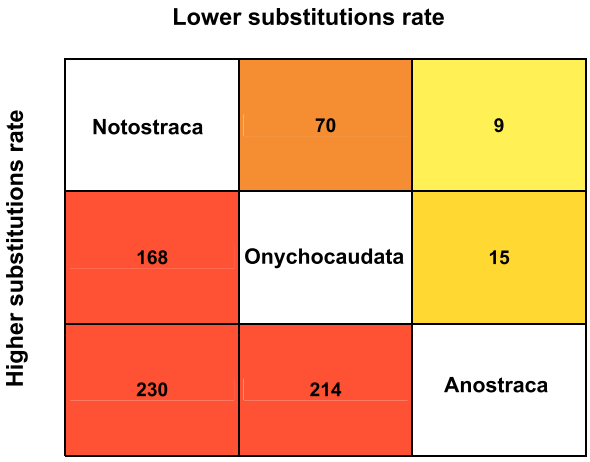

Fig. 3. Gene-by-gene differential substitutions rate analyses using the relative-rate test implemented in RRTree. The heatmap shows the number of genes with significant differential nucleotide substitutions rate (exact probability, $p < 0.001$) among lineages when comparing the three main branchiopod clades: Notostraca, Onychocaudata and Anostraca.

resources for further comparisons, we also included the draft genome of the spinicaudatan *L. dahalacensis*.

In line with data of previous sequencing [31,68], the assembly size of notostracan genomes ranged between ~74 Mb and ~ 115 Mb. Genome size estimates vary between 0.09 pg and 0.10 pg (except for *Triops australiensis* with a genome size of 0.16 pg, likely due to a supernumerary chromosome; [34]), approximately corresponding to ~88 Mb and ~ 98 Mb, respectively. These data suggest that the new assemblies covered from 84.1% up to 100% of the expected genome size. The completeness of sequencing is also confirmed by the BUSCO analysis results, indicating that almost all arthropod core genes were recovered. As a comparison, previously sequenced *Lepidurus* genomes showed completely overlapping BUSCO statistics [68]; moreover, the same analysis on the Japanese *T. cancriformis* genome [31] yielded 95.97% of complete genes and 98.03% complete+partial genes. The assembly size of *L. dahalacensis* (103.5 Mb) is larger than the average size of notostracan assemblies but it is smaller than the other Spinicaudata genome sequenced so far, *E. texana* (120.5 Mb; [2]). This suggests we may have missed a portion of the genome, which is a common issue in genome sequencing [2]. Though, the BUSCO analysis for completeness indicated that the *L. dahalacensis* assembly harbours up to the 96% of arthropod orthologs set, suggesting it can be considered a reliable source of genomic information. Overall, the length of Notostraca genomes appeared generally shorter than those of other branchiopods, such as *Daphnia* (120.0 Mb–197.2 Mb; [13]), the spinicaudatan *E. texana* (120.5 Mb) and *L. dahalacensis* (103.5 Mb), and the anostracan *A. franciscana*, (938.0 Mb; [35]) reflecting the trend of genome size reduction already observed based on genome size estimates [2,34].

The TEs content of newly sequenced genomes resulted approximately of the same level of magnitude previously scored in *L. arcticus* and *L. apus lubbocki* [68], with *T. cancriformis* harbouring slightly more TE insertions. When compared with other branchiopod genomes, notostracans appeared to host decidedly less LTR elements which, on the contrary, represent the majority of TEs in the other genomes, except *L. dahalacensis*, as also reported in previous analysis on *Daphnia* species [3,13]. Substantial changes of TEs' landscape have been commonly observed among both vertebrate and invertebrate lineages, with abrupt variation also among closely related taxa [12,90]. Present results on TEs' landscape and family abundance evidenced a variation of TEs content among genera and species in Notostraca, with TEs family occupancy showing a distribution pattern not consistent with the species tree. This suggests an extensive TEs genomic turnover, which contributed to the differential elimination/accumulation of insertions in different genomes during the evolution of the lineage.

The phylogenomic analysis fully reflected the known relationships among analyzed branchiopod lineages [71]. Although all nodes in the tree received the maximum support, concordance factors at deepest nodes indicated that a substantial part of genes and sites does not support the species tree topology. These conflicts across genes and sites can be due to a scarce phylogenetic signal associated to a high level of substitutions saturation, or due to events of incomplete lineage sorting or gene introgression. Further analyses on discordant genes and sites may shed light on this latter phenomenon and can contribute to widen the knowledge on branchiopods evolutionary history.

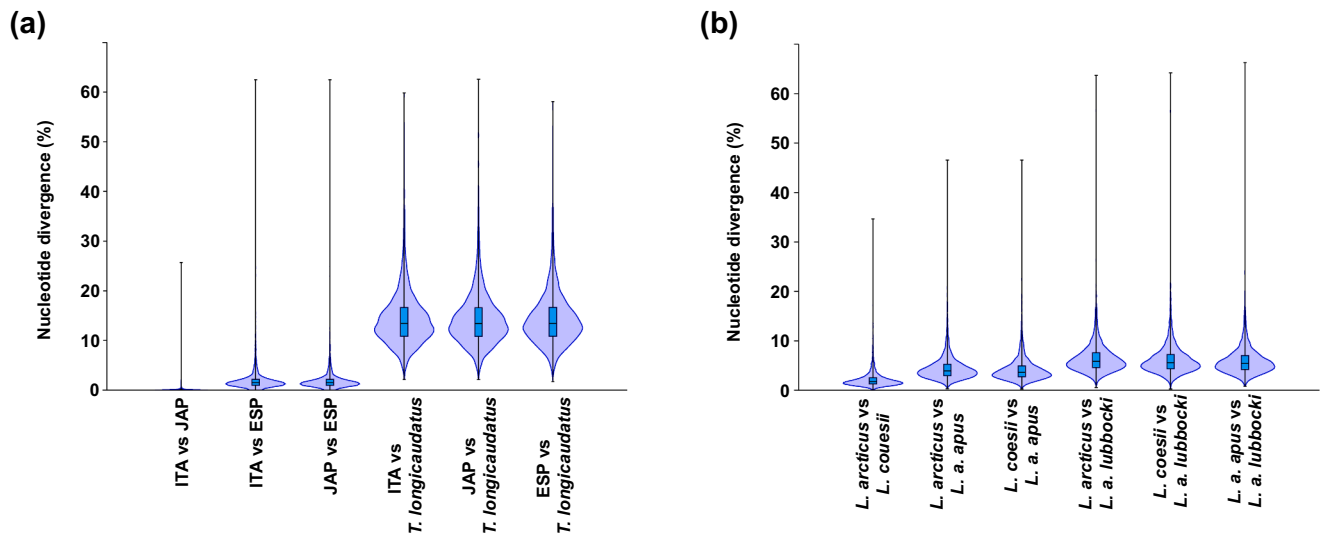

Fig. 4. *Triops* and *Lepidurus* nucleotide divergence analysis. Extent of nucleotide divergence among (a) *Triops cancriformis* samples and with *T. longicaudatus*, (b) *Lepidurus* samples.

The analysis of the substitutions rate showed a clear difference among the clade Notostraca+Onychocaudata and Anostraca (here represented by *A. franciscana*), with the latter showing a higher substitutions rate. These data are in line with a recent branchiopod mitochondrial genome comparison [46], where Notostraca also showed a lower substitutions rate than Onychocaudata. In a previous genome survey, *L. arcticus* and *L. apus lubbocki* were also found evolving slower than *Daphnia* species [68]. Differential substitutions rate can be linked to many life-history traits, such as generation time, differences in body size, metabolic rates, fecundity or, at the clade level, speciation rate [15,26,52,87,88]. Although these differences in life-history traits would likely affect the whole genome, in the present analysis the gene-by-gene analysis clearly evidenced that the differential substitutions rate only affect a relative proportion of genes rather than the whole dataset. This suggests that factors acting more locally could be responsible for the observed pattern. The faster substitutions rate observed in Anostraca and Onychocaudata mitochondrial genomes with respect to the Notostraca ones may have been driven by an ancestral change in nucleotide composition [46], as also observed in other organisms [22] and predicted by mathematical models [75]. Yet, no significant change in GC content can be observed in our dataset concerning genes with differential substitutions rate. Moreover, we found no correlation between substitutions rate and the natural selection regime, as Anostraca appeared to be subject to stronger purifying selection. When dealing specifically with genes showing higher substitutions rate, we also did not find any significant difference in the extent of selective pressures. This, therefore, suggests natural selection does not affect the genes substitutions rate.

The correlation between rates of molecular evolution and the rate of morphological changes has been debated in the past [9,61] and it was never completely ruled out. Some recent studies, in fact, suggested a causal relationship between a comparatively low nucleotide substitutions rate and the morphological stasis in the so-called living fossils [1,25]. However, this correlation has been questioned in the coelacanth on genomic and paleontological grounds [11,58]. In consideration of the problem "part-whole ambiguity" in defining living fossils [45], the gene-by-gene approach clearly evidenced that the substitutions rate varies between genes among branchiopod lineages, suggesting that local molecular mechanisms may be more effective in shaping the rate of molecular evolution rather than a consistent genome-wide process [11]. In this view, further evidence can be drawn from the genome turnover observed in Notostraca when considering the differential occurrence of TE families and the higher rate of gene families' evolution. Beside showing a different TE landscape with respect to other Branchiopoda, the analysis of TEs families showed their significantly different distribution both between and within *Triops* and *Lepidurus* genera, thus suggesting that TEs in Notostraca genomes are subject to active turnover. Moreover, the estimate of the rate of gene family evolution indicated that Notostraca has a significantly higher evolutionary rate, which is evident by analyzing the number of rapidly evolving gene families (*i.e.*, those gene families showing a significant shift in the extent of expansion/contraction with respect to the expectation). Altogether, these observations clearly point to a non-static genome in Notostraca. Therefore, along with analyses on the lineage radiation [53], data presented here contribute to cast doubt about the notion of "living fossil" applied to tadpole shrimps and, more generally, to the coupling of morphological and genomic conservatism.

Notwithstanding the clear geographical proximity, the Italian (ITA) and the Spanish (ESP) *T. cancriformis* samples showed a higher genetic divergence with respect to the Japanese sample. This differs from data on mitochondrial genomes, where they resulted only 0.18% divergent in contrast to the Japanese sample whose divergence from the Italian and Spanish samples was 0.24% and 0.32%, respectively [46]. These data are in line with previous analyses performed on both mitochondrial DNA fragments and microsatellite loci: despite a close similarity of the mitochondrial markers, microsatellite loci indicated a more clear-cut divergence of the Spanish population from the Italian one [49,84]. Although generally more variable than nuclear genes, the possibility of strong purifying selection on the mitochondrial genome could explain the lower genetic divergence. On the other hand, it is also possible that a relaxation of selection regime, or even positive selection, acting on the nuclear genome can hamper genetic divergence. Although limitations of ω in estimating population-level selective pressures change [41], in the present analysis we found no significant change of average ω between the ESP and ITA + JAP lineages, once again suggesting that natural selection does not affect the rate of molecular evolution in tadpole shrimps. It is to be noted, though, that the taxonomic status of Spanish populations could be questioned: most of them, in fact, were formerly ascribed to the *Triops cancriformis simplex* subspecies, but mitochondrial DNA markers and a morphological revision indicated that the Northern Spain population of Girona province, where the present ESP sample was collected, should actually belong to the *T. cancriformis cancriformis* lineage [38]. Notably, the only character that appears to clearly distinguish the Northern Spain *T. cancriformis* population is related to its

gonochoric reproduction rather than to phylogenetic position [38]. At present it is not possible to draw further conclusions about the observed high divergence of the ESP sample from the Italian parthenogenetic *T. cancriformis*, although on microsatellites and genome-wide data we may speculate about a possible introgression of mitochondrial *T. cancriformis cancriformis* haplotypes in a divergent *T. cancriformis* lineage.

An interesting aspect emerging from the analysis of *T. cancriformis* genomes concerns the link between the observed differential TE content and the reproductive strategies. The Spanish ESP sample belongs to a bisexual population [38,84], while the Italian and the Japanese populations are parthenogenetic and hermaphroditic, respectively. Theoretically speaking, TE content should be lower in selfers and parthenogenetic taxa than in bisexual ones; yet, this was only confirmed in a few instances, like in *D. pulex* [70,82], *Caenorhabditis* spp. [16] and stick insects [6], while in other instances no differences emerged [3,78] or even more TEs were retrieved in parthenogenetic lineages [4,40]. In line with theoretical expectations, data presented here indicate that the bisexual Spanish *T. cancriformis* sample has significantly higher TEs contents than the parthenogenetic Italian sample and the hermaphroditic Japanese one. Moreover, the parthenogenetic sample shows higher TEs content than the hermaphroditic one. Generally speaking, beside the organism breeding system, other factors may intervene on the variation of the TE genomic content, such as the extent of population size: the larger the population size, the higher the possibility of TEs purging even in selfers and parthenogenetic taxa. Another explanation could be the possibility of selective forces acting on specific target as in the "deleterious mutations" model, where TE insertions are counter-selected because of gene disruption, or the "ectopic recombination" model, where TE insertions may trigger unequal recombination events resulting in deleterious chromosomal rearrangement (reviewed in [27]). Recent analyses on 26 parthenogenetic animals suggested that observed genomic patterns maybe lineage specific, depending on the accumulation of active/non-active TEs, the evolution of efficient TEs suppression mechanisms and the time since the transition to parthenogenesis, therefore not necessarily reflecting general outcomes of parthenogenesis [33]. Further studies on *T. cancriformis* genomes from populations with different breeding systems could potentially clarify the TE dynamics in this context.

The phylogenomic relationships and the observed divergence between *Lepidurus* genomes are in line with previous analyses based on mitochondrial and nuclear gene fragments [39,48,50,83]. The species *Lepidurus apus* is here confirmed paraphyletic, and the divergence between the subspecies *L. apus apus* and *L. apus lubbocki* is similar to that between the other two *Lepidurus* species analyzed. This confirms the previously suggested differentiation at species level, which also seems supported by morphological differences of the supra-anal plate [50]. *L. arcticus* and *L. couesii* resulted the most closely related samples and exhibited the lowest genome-wide divergence. This, again, further confirms previous analyses and their possible common origin [39,48,50,83], also reflected by their close geographic distribution. In fact, *L. arcticus* is a circumpolar species and *L. couesii* is a North American and Asian species reaching Eastern Europe [10,66]. Yet, the presently analyzed *L. couesii* sample has been collected in Southern Italy, where it was found for the first time in the 2004–2005 [69]. This constitute an interesting finding because the climatic and environmental conditions in Southern Italy are strikingly different from those in Northern Europe. Although we do not find evidence of positive selective pressures which may underlie *L. couesii* adaptation to new environmental conditions, this sister species pair may be a reliable framework where to address deeper studies on environmental adaptation genomics and transcriptomics.

Data availability

Raw reads data are available in NCBI SRA database under the following accession numbers: SRR14118908, SRR14119275, SRR14127293, SRR14127286, SRR14127427, SRR15808114. Genome assemblies have been submitted to NCBI Genbank under the BioProject PRJNA417576. Genome assemblies, gene annotations and repeat libraries are available on Figshare under the DOI: https://doi.org/10.6084/m9.figshare.14420669.

Authors' contribution

AL and BM: Conceptualization, Investigation; AL: Data curation, Formal analysis, Supervision, Writing - original draft; GF, JM, CS and AMS: Data curation, Validation, Formal analysis, Methodology, Software, Writing - review & editing; PLM, RC, SJW and BM: Funding acquisition, Resources, Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no competing interests.

Acknowledgments

This work has been supported by Canziani funding to BM and by the National Institute of Health/National Cancer Institute grant NIH P30 CA006973 to SJW. Authors want to thank Franca Scanabissi, Giuseppe Alfonso, Erich Eder, Dani Boix and Jordi Sala for providing samples. *D. magna* data were produced by The Center for Genomics and Bioinformatics at Indiana University and distributed *via* wFleaBase in collaboration with the *Daphnia* Genomics Consortium. We are grateful to the Editor and to the two anonymous Reviewers whose comments and suggestions helped to improve the present analysis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygeno.2021.11.001.

References

[1] C.T. Amemiya, J. Alföldi, A.P. Lee, S. Fan, H. Philippe, I. MacCallum, K. Lindblad-Toh, The African coelacanth genome provides insights into tetrapod evolution, Nature 496 (2013) 311–316.
[2] J.G. Baldwin-Brown, S.C. Weeks, A.D. Long, A new standard for crustacean genomes: the highly contiguous, annotated genome assembly of the clam shrimp *Eulimnadia texana* reveals HOX gene order and identifies the sex chromosome, Genome Biol. Evol. 10 (2018) 143–156.
[3] J. Bast, I. Schaefer, T. Schwander, M. Maraun, S. Scheu, K. Kraaijeveld, No accumulation of transposable elements in asexual arthropods, Mol. Biol. Evol. 33 (2016) 697–706.
[4] R. Blanc-Mathieu, L. Perfus-Barbeoch, J.-M. Aury, M. Da Rocha, J. Gouzy, E. Sallet, E.G.J. Danchin, Hybridization and polyploidy enable genomic plasticity without sex in the most devastating plant-parasitic nematodes, PLoS Genet. 13 (6) (2017), e1006777.
[5] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics 30 (2014) 2114–2120.
[6] L. Bonandin, C. Scavariello, A. Luchetti, B. Mantovani, Evolutionary dynamics of R2 retroelement and insertion inheritance in the genome of bisexual and parthenogenetic *Bacillus rossius* populations (Insecta Phasmida), Insect Mol. Biol. 23 (2014) 808–820.
[7] L. Brendonck, L. De Meester, Egg banks in freshwater zoo- plankton: evolutionary and ecological archives in the sediment, Hydrobiologia 491 (2003) 65–84.
[8] L. Brendonck, D.C. Rogers, J. Olesen, S. Weeks, W.R. Hoeh, Global diversity of large branchiopods (Crustacea: Branchiopoda) in freshwater, Hydrobiologia 595 (2008) 167–176.
[9] L. Bromhann, M. Woolfit, M.S.Y. Lee, A. Rambaut, Testing the relationship between morphological and molecular rates of change along phylogenies, Evolution 56 (2002) 1921–1930.
[10] J. Brtek, A. Thiéry, The geographic distribution of the European branchiopods (Anostraca, Notostraca, Spinicaudata, Laevicaudata), Hydrobiologia 298 (1995) 263–280.
[11] D. Casane, P. Laurenti, Why coelacanths are not 'living fossils', Bioessays 35 (2013) 332–338.
[12] D. Chalopin, M. Naville, F. Plard, D. Galiana, J.-N. Volff, Comparative analysis of transposable elements highlights mobilome diversity and evolution in vertebrates, Genome Biol. Evol. 7 (2015) 567–580.

[13] J.K. Colbourne, M.E. Pfrender, D. Gilbert, W.K. Thomas, A. Tucker, T.H. Oakley, J. L. Boore, The ecoresponsive genome of *Daphnia pulex*, Science 331 (2011) 555–561.

[14] T. De Bie, N. Cristianini, J.P. Demuth, M.W. Hahn, CAFE: a computational tool for the study of gene family evolution, Bioinformatics 22 (2006) 1269–1271.

[15] A.R. De La Torre, Z. Li, Y. Van de Peer, P.K. Ingvarsson, Contrasting rates of molecular evolution and patterns of selection among gymnosperms and flowering plants, Mol. Biol. Evol. 34 (2017) 1363–1377.

[16] E.S. Dolgin, B. Charlesworth, A.D. Cutter, Population frequencies of transposable elements in selfing and outcrossing *Caenorhabditis* nematodes, Genet. Res. 90 (2008) 317–329.

[17] D.M. Emms, S. Kelly, OrthoFinder: phylogenetic orthology inference for comparative genomics, Genome Biol. 20 (2019) 238.

[18] A. Faddeeva-Vakhrusheva, M.F.L. Derks, S.Y. Anvar, V. Agamennone, W. Suring, S. Smit, D. Roelofs, Gene family evolution reflects adaptation to soil environmental stressors in the genome of the collembolan *Orchesella cincta*, Genome Biol. Evol. 8 (2016) 2106–2117.

[20] L. Forró, N.M. Korovchinsky, A.A. Kotov, A. Petrusek, Global diversity of cladocerans (Cladocera; Crustacea) in freshwater, Hydrobiologia 595 (2008) 177–184.

[21] G. Fryer, Studies on the functional-morphology and biology of the Notostraca (Crustacea, Branchiopoda), Phil. Trans. R Soc. London B 321 (1988) 27–124.

[22] M. Friedrich, D. Tautz, An episodic change of rDNA nucleotide substitution rate has occurred during the emergence of the insect order Diptera, Mol. Biol. Evol. 14 (1997) 644–653.

[23] L. Fu, B. Niu, Z. Zhu, S. Wu, W. Li, CD-HIT: accelerated for clustering the next generation sequencing data, Bioinformatics 28 (2012) 3150–3152.

[24] G. Gand, J. Garric, J. Lapeyrie, Biocénoses à triopsidés (Crus- tacea, Branchiopoda) du Permien du bassin de Lodève (France), Geobios 30 (1997) 673–700.

[25] N.J. Gemmell, K. Rutherford, S. Prost, M. Tollis, D. Winter, J.R. Macey, Ngatiwai Trust Board, The tuatara genome reveals ancient features of amniote evolution, Nature 584 (2020) 403–409.

[26] J.F. Gillooly, A.P. Allen, G.B. West, J.H. Brown, The rate of DNA evolution: effects of body size and temperature on the molecular clock, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 140–145.

[27] S. Glémin, C.M. François, N. Galtier, Genome evolution in outcrossing vs. Selfing vs. asexual species, in: M. Anisimova (Ed.), Evolutionary Genomics: Statistical and Computational Methods, Methods in Molecular Biology, Humana Press, 2019, pp. 331–369.

[28] Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data, Bioinformatics 32 (2016) 2847–2849.

[29] J.M. Hay, S. Subramanian, C.D. Millar, E. Mohandesan, D.M. Lambert, Rapid molecular evolution in a living fossil, Trends Genet. 24 (2008) 106–109.

[30] J. Huerta-Cepas, F. Serra, P. Bork, ETE 3: reconstruction, analysis, and visualization of phylogenomic data, Mol. Biol. Evol. 33 (2016) 1635–1638.

[31] K.T. Ikeda, Y. Hirose, K. Hiraoka, E. Noro, K. Fujishima, M. Tomita, A. Kanai, Identification, expression, and molecular evolution of microRNAs in the "living fossil" T*riops cancriformis* (tadpole shrimp), RNA 21 (2015) 230–242.

[32] S.D. Jackman, B.P. Vandervalk, H. Mohamadi, J. Chu, S. Yeo, S.A. Hammond, I. Birol, ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter, Genome Res. 27 (2017) 768–777.

[33] K.S. Jaron, J. Bast, R.W. Nowell, T. Rhyker Ranallo-Benavidez, M. Robinson-Rechavi, T. Schwander, Genomic features of parthenogenetic animals, J. Hered. 112 (2021) 19–33.

[34] N.W. Jeffery, *Genome size diversity and evolution in the Crustacea* (PhD Thesis), University of Guelph, Ontario, Canada, 2015.

[35] E. Jo, S.J. Lee, E. Choi, J. Kim, S.G. Lee, J.H. Lee, J.H. Kim, H. Park, Whole genome survey and microsatellite motif identification of *Artemia franciscana*, Biosci. Rep. 41 (2021), BSR20203868.

[36] R. Kajitani, K. Toshimoto, H. Noguchi, A. Toyoda, Y. Ogura, M. Okuno, T. Itoh, Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads, Genome Res. 24 (2014) 1384–1395.

[37] K. Katoh, D.M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability, Mol. Biol. Evol. 30 (2013) 772–780.

[38] M. Korn, F. Marrone, J.L. Pérez-Bote, M. Machado, M. Cristo, L.C. Da Fonseca, A. K. Hundsdoerfer, Sister species within the *Triops cancriformis* lineage (Crustacea, Notostraca), Zool. Scr. 35 (2006) 301–322.

[39] M. Korn, N. Rabet, H.V. Ghate, F. Marrone, A.K. Hundsdoerfer, Molecular phylogeny of the Notostraca, Mol. Phylogenet. Evol. 69 (2013) 1159–1171.

[40] K. Kraaijeveld, B. Zwanenburg, B. Hubert, C. Vieira, S. De Pater, J.J.M. Van Alphen, P. De Knijff, Transposon proliferation in an asexual parasitoid, Mol. Ecol. 21 (2012) 3898–3906.

[41] S. Kryazhimskiy, J.B. Plotkin, The population genetics of dN/dS, PLoS Genet. 4 (2008), e1000304.

[42] D.R. Laetsch, M.L. Blaxter, BlobTools: interrogation of genome assemblies, F1000 Res 6 (2017) 1287.

[43] L. Lagebro, P. Gueriau, T.A. Hegna, N. Rabet, A.D. Butler, G.E. Budd, The oldest notostracan (Upper Devonian Strud locality, Belgium), Palaeontology 58 (2015) 497–509.

[44] H. Li, Minimap2: pairwise alignment for nucleotide sequences, Bioinformatics 34 (2018) 3094–3100.

[45] S. Lidgard, A.C. Love, Rethinking living fossils, BioScience 68 (2018) 760–770.

[46] A. Luchetti, G. Forni, A.M. Skaist, S.J. Wheelan, B. Mantovani, Mitochondrial genome diversity and evolution in Branchiopoda (Crustacea), Zool. Lett. 5 (2019) 15.

[48] B. Mantovani, M. Cesari, F. Scanabissi, Molecular taxonomy and phylogeny of the 'living fossil' lineages *Triops* and *Lepidurus* (Branchiopoda: Notostraca), Zool. Scr. 33 (2004) 367–374.

[49] B. Mantovani, M. Cesari, A. Luchetti, F. Scanabissi, Mitochondrial and nuclear DNA variability in the living fossil *Triops cancriformis* (Bosc, 1801) (Crustacea, Branchiopoda, Notostraca), Heredity 100 (2008) 496–505.

[50] B. Mantovani, M. Cesari, F. Scanabissi, Molecular taxonomy and phylogeny of Italian *Lepidurus* taxa (Branchiopoda: Notostraca), It. J. Zool. 76 (2009) 358–365.

[51] H. Mao, H. Wang, SINE_scan: an efficient tool to discover short interspersed nuclear elements (SINEs) in large-scale genomic datasets, Bioinformatics 33 (2017) 743–745.

[52] A.P. Martin, S.R. Palumbi, Body size, metabolic rate, generation time, and the molecular clock, Proc. Natl. Acad. Sci. U. S. A. 90 (1993) 4087–4091.

[53] T.C. Mathers, R.L. Hammond, R.A. Jenner, B. Hänfling, A. Gómez, Multiple global radiations in tadpole shrimps challenge the concept of 'living fossils', PeerJ 1 (2013), e62.

[54] F.K. Mendes, M.W. Hahn, Gene tree discordance causes apparent substitution rate variation, Syst. Biol. 65 (2016) 711–721.

[55] F.K. Mendes, D. Vanderpool, B. Fulton, M.W. Hahn, CAFE 5 models variation in evolutionary rates among gene families, Bioinformatics 36 (2020) 5516–5518.

[56] B.Q. Minh, H.A. Schmidt, O. Chernomor, D. Schrempf, M.D. Woodhams, A. von Haeseler, R. Lanfear, IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era, Mol. Biol. Evol. 37 (2020) 1530–1534.

[57] B.Q. Minh, M.W. Hahn, R. Lanfear, New methods to calculate concordance factors for phylogenomic datasets, Mol. Biol. Evol. 37 (2020) 2727–2733.

[58] M. Naville, D. Chalopin, D. Casane, P. Laurenti, J.N. Volff, The coelacanth: can a "living fossil" have active transposable elements in its genome? Mob. Genet. Elem. 5 (2015) 55–59.

[59] O. Nishimura, Y. Hara, S. Kuraku, gVolante for standardizing completeness assessment of genome and transcriptome assemblies, Bioinformatics 33 (2017) 3635–3637.

[60] J. Olesen, S. Richter, Onychocaudata (Branchiopoda: Diplostraca), a new high-level taxon in branchiopod systematics, J. Crustac. Biol. 33 (2013) 62–65.

[61] K.E. Omland, Correlated rates of molecular and morphological evolution, Evolution 51 (1997) 1381–1393.

[62] L.P. Pryszcz, T. Gabaldón, Redundans: an assembly pipeline for highly heterozygous genomes, Nucleic Acids Res. 44 (2016), e113.

[63] P. Rice, I. Longden, A. Bleasby, EMBOSS: the European molecular biology open software suite, Trends Genet. 16 (2000) 276–277.

[64] M. Robinson, M. Gouy, C. Gautier, D. Mouchiroud, Sensitivity of the relative-rate test to taxonomic sampling, Mol. Biol. Evol. 15 (1998) 1091–1098.

[65] M. Robinson-Rechavi, D. Huchon, RRTree: relative-rate tests between groups of sequences on a phylogenetic tree, Bioinformatics 16 (2000) 296–297.

[66] D.C. Rogers, Revision of the Nearctic *Lepidurus* (Notostraca), J. Crustac. Biol. 21 (2001) 991–1006.

[67] D.C. Rogers, Branchiopoda (Anostraca, Notostraca, Laevicaudata, Spinicaudata, Cyclestherida), in: G.E. Likens (Ed.), Encyclopedia of Inland Waters vol. 2, Elsevier Science, Oxford, 2009, pp. 242–249.

[68] C. Savojardo, A. Luchetti, P.L. Martelli, R. Casadio, B. Mantovani, Draft genomes and genomic divergence of two *Lepidurus* tadpole shrimp species (Crustacea, Branchiopoda, Notostraca), Mol. Ecol. Resour. 19 (2019) 235–244.

[69] F. Scanabissi, G. Alfonso, S. Bergamaschi, B. Mantovani, Primo ritrovamento di *Lepidurus couesii* Packard, 1875 in Italia, Thalassia Salentina 29 (2006) 113–124.

[70] S. Schaack, E. Choi, M. Lynch, E.J. Pritham, DNA transposons and the role of recombination in mutation accumulation in *Daphnia pulex*, Genome Biol. 11 (2010) R46.

[71] M. Schwentner, S. Richter, D.C. Rogers, G. Giribet, Tetraconatan phylogeny with special focus on Malacostraca and Branchiopoda: highlighting the strength of taxon-specific matrices in phylogenomics, Proc. R. Soc. B 285 (2018) 20181524.

[72] A.F.A. Smit, R. Hubley, RepeatModeler Open-1.0, Retrieved from, https://www.repeatmasker.org, 2008–2015.

[73] A.F.A. Smit, R. Hubley, P. Green, RepeatMasker Open-4.0, Retrieved from, http://www.repeatmasker.org, 2013–2015.

[74] M. Stanke, O. Keller, I. Gunduz, A. Hayes, S. Waack, B. Morgenstern, AUGUSTUS: *ab initio* prediction of alternative transcripts, Nucleic Acids Res. 4 (2006) W435–W439.

[75] N. Sueoka, Directional mutation pressure, mutator mutations, and the dynamics of molecular evolution, J. Mol. Evol. 37 (1993) 137–153.

[76] X.-Y. Sun, X. Xia, Q. Yang, Dating the origin of the major lineages of Branchiopoda, Palaeoworld 25 (2016) 303–317.

[77] M. Suyama, D. Torrents, P. Bork, PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments, Nucleic Acids Res. 34 (2006) W609–W612.

[78] A. Szitenberg, S. Cha, C.H. Opperman, D.M. Bird, M.L. Blaxter, D.H. Lunt, Genetic drift, not life history or RNAi, determine long-term evolution of transposable elements, Genome Biol. Evol. 8 (2016) 2964–2978.

[79] F. Takahashi, Use of the tadpole shrimp (*Triops* spp.) as a biological agent to control paddy weeds in Japan, in: J. Bay-Petersen (Ed.), Integrated Management of Paddy and Aquatic Weeds in Asia, FFTC, Taipei, 1994, pp. 128–138.

[80] T.-H. To, M. Jung, S. Lycett, O. Gascuel, Fast dating using least-squares criteria and algorithms, Syst. Biol. 65 (2016) 82–97.

[81] Tribolium Genome Sequencing Consortium, The genome of the model beetle and pest *Tribolium castaneum*, Nature 452 (2008) 949–955.

[82] P. Valizadeh, T.J. Crease, The association between breeding system and transposable element dynamics in *Daphnia pulex*, J. Mol. Evol. 66 (2008) 643–654.

[83] B. Vanschoenwinkel, T. Pinceel, M.P.M. Vanhove, C. Denis, M. Jocque, B.V. Timms, L. Brendock, Toward a global phylogeny of the "living fossil" crustacean order of the Notostraca, PLoS One 7 (2012), e34998.
[84] A. Velonà, A. Luchetti, F. Scanabissi, B. Mantovani, Genetic variability and reproductive modalities in European populations of Triops cancriformis (Crustacea, Branchiopoda, Notostraca), It. J. Zool. 76 (2009) 366–375.
[85] J.L. Villanueva-Cañas, S. Laurie, M.M. Albà, Improving genome-wide scans of positive selection by using protein isoforms of similar length, Genome Biol. Evol. 5 (2013) 457–467.
[86] P. Wagner, J.T. Haug, C. Haug, A new calmanostracan crustacean species from the Cretaceous Yixian Formation and a simple approach for differentiating fossil tadpole shrimps and their relatives, Zool. Lett. 5 (2019) 20.
[87] A.J. Webster, R.J.H. Payne, M. Pagel, Molecular phylogenies link rates of evolution and speciation, Science 301 (5632) (2003) 478.
[88] J.J. Welch, O.R.P. Bininda-Emonds, L. Bromham, Correlates of substitution rate variation in mammalian protein-coding sequences, BMC Evol. Biol. 8 (2008) 53.
[89] C.I. Wu, W.H. Li, Evidence for higher rates of nucleotide substitution in rodents than in man, Proc. Natl. Acad. Sci. U. S. A. 82 (1985) 1741–1745.
[90] C. Wu, J. Lu, Diversification of transposable elements in arthropods and its impact on genome evolution, Genes 10 (2019) 338.
[91] Z. Xu, H. Wang, LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons, Nucleic Acids Res. 35 (2007) W265–W268.
[92] Z. Yang, PAML 4: phylogenetic analysis by maximum likelihood, Mol. Biol. Evol. 24 (2007) 1586–1591.
[93] G. Forni, A.A. Ruggieri, G. Piccinini, A. Luchetti, BASE: a novel workflow to integrate non-ubiquitous genes in genomics analyses for selection, Ecol. Evol. 11 (2021) 13029–13035.
[94] G. Manni, M.R. Berkeley, M. Seppey, F.A. Simão, E.M. Zdobnov, BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes, Mol. Biol. Evol. 38 (2021) 4647–4654.