scientific reports

Check for updates

OPEN

Structure and genetic diversity of *Macrobrachium amazonicum* complex

Gabriel Monteiro de Lima[1], Beatriz Rochitti Boza[2], Luciano Domingues Queiroz[3], Gabriel Iketani[4], Carlos Murilo Tenório Maciel[1], Claudio Oliveira[2], Fausto Foresti[2], Wagner C. Valenti[5], Iracilda Sampaio[1], Vanessa Paes da Cruz[2]✉ & Cristiana Ramalho Maciel[1]✉

The present study evaluated the phylogenetic relationships and the population structure in the *'Macrobrachium amazonicum'* species complex, including *M. amazonicum* and *M. pantanalense* based on Single Nucleotide Polymorphism (SNP) markers. We analyzed specimens from three main genetic lineages (Clade I: *M. amazonicum* – Santarém-PA; Clade III: *M. amazonicum* – Mosqueiro-PA; Clade II: *M. pantanalense* – Bela Vista de Goiás-GO) and one artificially isolated population (*M. amazonicum* – Tucuruí-PA). After ddRAD sequencing, a final set of 969 SNPs were genotyped for all populations/species. High levels of genetic structure were observed among the populations of *M. amazonicum*, particularly when individuals from inland freshwater habitats in Amazon were compared to those from coastal basins, reinforcing the morphophysiological differences among both groups. Moreover, differently from the phylogenetic inferences using mitochondrial DNA markers, all analyses by SNPs confirmed the taxonomic status of *M. pantanalense* as a distinct species from *M. amazonicum*. Therefore, we provided the first molecular evidence of speciation in this group, being useful to resolve the taxonomic uncertainties in *M. pantanalense*. This study also indicates a putative process of ongoing speciation involving the clades I and III of *M. amazonicum*, thus emphasizing the need for further studies to better assess the relationships between both lineages.

Keywords Amazon river prawn, Genetic structure, Single nucleotide polymorphism, Phylogenetic relationships

Freshwater prawns of the genus *Macrobrachium* Spence Bate, 1868 are significant both ecologically and economically. Extensive research has been conducted on their genomics to understand their genetic diversity, evolutionary adaptations, and responses to environmental changes[1]. *Macrobrachium* are widespread throughout tropical and subtropical regions, being represented by more than 260 described species[2,3].

Nonetheless, the true species richness within this genus remains a contentious issue due to significant taxonomic ambiguities, which are often linked to their extensive geographical distribution, substantial morphological variability, and notably high genetic divergence rates[4]. Studies have demonstrated considerable intraspecific genetic differentiation, ranging from 0.0 to 12.6%, and interspecific genetic divergence, spanning from 15.1 to 25.5%, as revealed by DNA barcoding techniques[4–7]. These findings collectively illustrate the high rates of genetic divergence within the genus *Macrobrachium*, which are pivotal for understanding their evolutionary history, taxonomy, and conservation strategies. Nonetheless, these issues might determine biased taxonomic inferences related to the description of new species and synonymization[8–10].

The ability of *Macrobrachium* species to adapt to varying salinity levels is further demonstrated by transcriptomic studies, which identify specific genes involved in osmoregulation, cell volume regulation, and osmotic stress response, showing differential expression patterns under different salinity conditions[11]. The plasticity of morphophysiological features in *Macrobrachium* is particularly evident when species from estuarine

[1]Institute of Coastal Studies, Federal University of Pará, Bragança 68600-000, Brazil. [2]Paulista State University, Botucatu Institute of Biosciences, Botucatu 18618-689, Brazil. [3]Federal Institute of Education Science and Technology of Pará, Tucuruí 68464-000, Brazil. [4]Institute of Educacional Sciences, Federal University of Western Pará, Santarém 68040-255, Brazil. [5]Aquaculture Center, São Paulo State University, Jaboticabal 14884-900, Brazil. ✉email: vanessa.paes@unesp.br; cmaciel@ufpa.br

regions are compared to those from inland freshwater habitats. For instance, the coastal species present greater fecundity, smaller eggs, extended larval development (ELP) and more larvae stages. On the other hand, the inland freshwater species are characterized by lower fecundity, larger eggs and short larval development (SLD)[12]. In addition, morphotypes have been described within some valid species, reinforcing the role of morphological variation in taxonomic misidentification[13,14].

The Amazon river prawn *Macrobrachium amazonicum* (Heller, 1862) represents an uncommon case in the genus since this taxon is widespread over most coastal and inland hydrographic basins from South America, showing ELD in populations from both environments[15]. In addition, this species is divided into three main lineages, as follows: the hololimnetic populations from the continental Amazon basin (I), Paraná-Paraguay basin (II), and the diadromous populations from coastal systems in north/northeastern Brazil (III)[6]. The lineages I e II are physiologically adapted to freshwater life cycles, being found in rivers, lakes and other inland aquatic environments[16]. The growth and fertility rates of lineages I and II are lower than those reported in coastal and estuarine populations (lineage III)[17].

Furthermore, the populations from the Paraná-Paraguay basin were later described as a new and endemic species to the Pantanal region, named *Macrobrachium pantanalense*[18]based on their unique morphological, physiological, reproductive isolation and life history traits[16,19,20]. Nevertheless, molecular analyses based on mitochondrial DNA markers revealed low levels of genetic divergence between the main lineages of *M. amazonicum* (I and III) and *M. pantanalense* (II). For instance, the variation in the cytochrome c oxidase subunit I (COI) and 16 S rDNA sequences ranged from 0.0 to 3.3% and 0.0 to 1.1%, respectively[6,21]. As a result of the incongruencies between the genetic data because are rates considered low divergence for the *Macrobrachium* group and other biological aspects, both species have been referred to as the '*M. amazonicum* complex'.

Accordingly, the utilization of traditional molecular markers based on mitochondrial (mtDNA) and nuclear sequences for phylogenetic inferences has proved to have some limitations to resolve the taxonomic uncertainties of some groups, including *Macrobrachium*, since these markers allow analyzing the polymorphic sites from restricted DNA regions[22]. For instance, the COI marker in *M. amazonicum* encompasses only ~ 4% (650 to 15000 bp) of the total mtDNA, which in turn is remarkable smaller that the nuclear genome[6]. Other issues that interfere in the accuracy of mtDNA-based inferences include the occurrence of heteroplasmy and the presence of nuclear insertions of mitochondrial DNA or Numts[23]. Both events have been reported in populations of *M. amazonicum*[24]. Therefore, the true status of species differentiation and reliable phylogenetic reconstructions among the main lineages *M. amazonicum* should rely on alternative and refined approaches to elucidate their evolutionary history.

In this sense, the analyses of Single Nucleotide Polymorphisms (SNPs) stand out as a powerful tool to genetic studies since these markers are highly abundant and provide a broad coverage of genomes, besides being usually associated with neutral loci[25]. Because of these features, the SNPs enable a more accurate assessment of genomic variability, being particularly efficient to estimate the levels of intraspecific genetic diversity and population structure[22,26]. Recently, these markers have been easily obtained via Next Generation Sequencing (NGS), using techniques such as Restriction Site Associated DNA Sequencing (RADseq) and its variations such as ddRAD[27,28]. As expected, the large dataset generated by these methods have been helpful in solving phylogenetic problems in cryptic species complexes. In *Macrobrachium*, this approach has been successfully applied to population studies of *M. rosenbergii* in Asia[29] and to the delimitation of closely related species along African coast[30].

Therefore, the aim of the present study was to obtain an SNP dataset based on the ddRAD technique to reassess the population structure of the main lineages of the *M. amazonicum* complex in Brazil, including the hololimnetic and diadromous populations of *M. amazonicum* and *M. pantanalense*. The results provided new insights into the taxonomic status of *M. pantanalense*, as well as the structure patterns and genetic diversity levels in coastal and inland populations of the Amazon river prawn. Additionally, these findings contributed to a deeper understanding of the evolutionary relationships among the distinct lineages within this species complex, while also providing data to support the proper management and conservation of the group's natural populations. This is particularly relevant as some of these populations are extensively exploited by artisanal fisheries in northern Brazil, highlighting the potential need for differential management strategies based on genetic parameters and population structures.

Results

Molecular identification of specimens

A total of 39 samples of *Macrobrachium* were sequenced and analyzed using the gene mitochondrial COI marker, with mean fragment size = 572 base pairs (bp). The identification of specimens of *Macrobrachium amazonicum* ($N = 35$) from Mosqueiro, Tucuruí, and Santarém, as well as *M. pantanalense* ($N = 4$) were confirmed, revealing 94 to 100% of similarity with the available sequences in GenBank (Supplementary Table S1). Lower levels of similarity were recorded mainly in Mosqueiro, associated with higher percentages of heteroplasmies in the generated sequences. The pairwise genetic distance between populations of *M. amazonicum* ranged from 0.00 (Santarém x Tucuruí) to 0.32% (Santarém x Mosqueiro). In comparison to *M. pantanalense*, the genetic divergence varied from 0.03 to 0.23% in relation to samples of *M. amazonicum* from Mosqueiro and Santarém, respectively (Supplementary Table S2).

Identification of SNP loci

The sequencing of ddRAD libraries generated a total of 260,096,350 raw paired-end reads from 39 sequenced individuals corresponding to *M. amazonicum* and *M. pantanalense*. The variation in the number of reads per sample ranged from 2,029,956 to 6,863,355; except for a single sample that presented a lower amount of data, comprising 842,738 sequences. After applying quality filters, approximately 99,6% of the data were retained, and

all sequences were standardized to 143 bp, resulting in a total of 259,206,820 reads (Fig. 1; Supplementary Table S3).

A total of 3,676 SNPs was identified using the software STACKs (*refmap*). From this number, we selected the SNPs that have been successfully genotyped in, at least, 70% of individuals from each population ($r = 0.7$) and

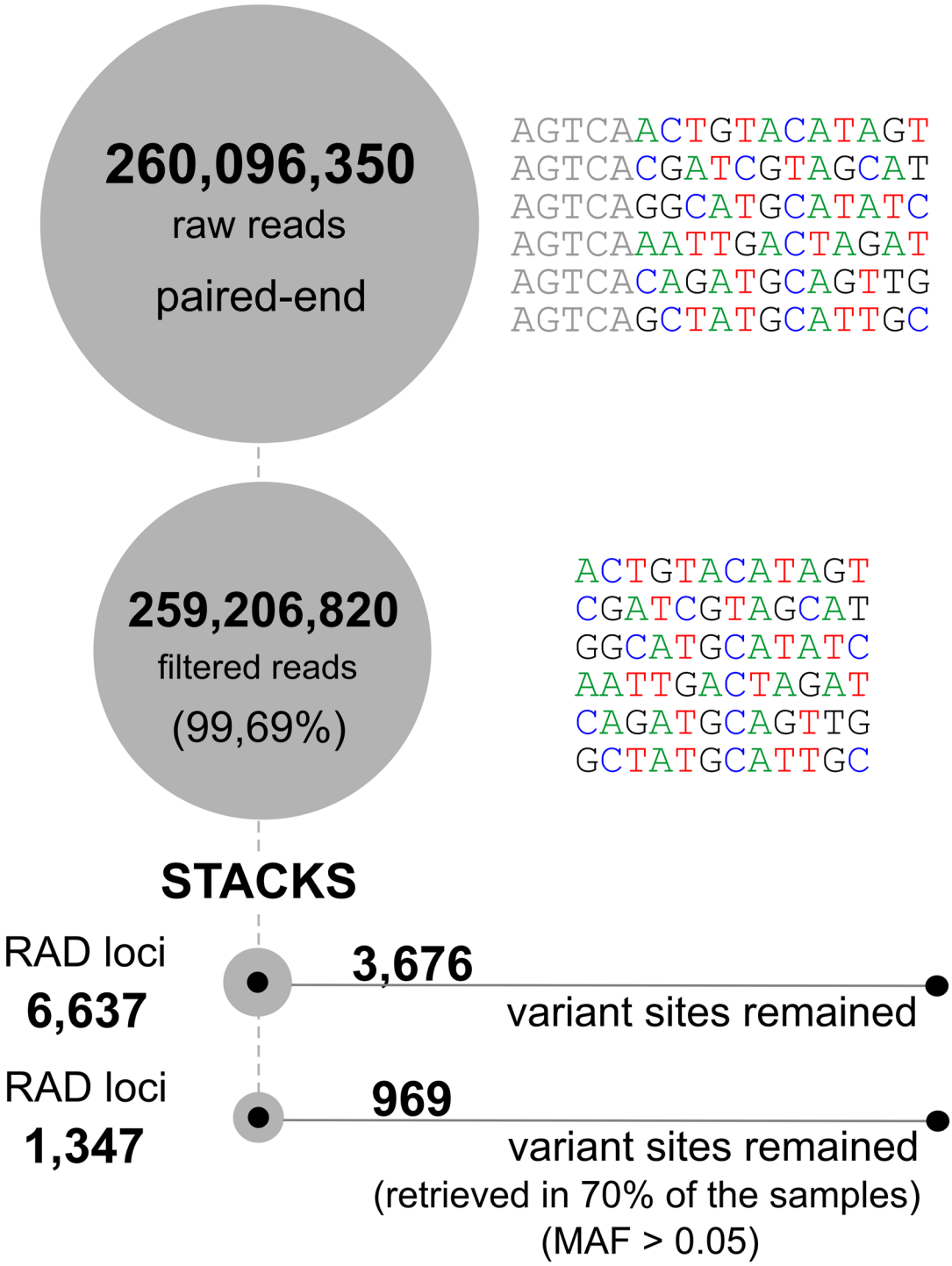

Fig. 1. Overview workflow of the data obtained from ddRAD sequencing, highlighting the quantity of SNP markers obtained at each filtering stage, and the final set of markers (969 SNPs) for the *M. amazonicum* and *M. pantanalense* species.

presenting a minor allele frequency (MAF) above 0.05. In the end, 969 SNPs were obtained, encompassing 562 transitions, 407 transversions (Supplementary Figure S1) and 36% of missing data.

Population genetic differentiation

The highest number of private alleles (Ap) was recorded in individuals of *M. pantanalense* ($n = 4$), totaling Ap = 235. As for *M. amazonicum*, three populations were analyzed, corresponding to the collection sites in Mosqueiro ($n = 11$), Tucuruí ($n = 11$) and Santarém ($n = 12$). The individuals from Mosqueiro and Santarém showed similar values of private alleles (Ap = 107 and Ap = 105, respectively). The lowest value was observed in the individuals from Tucuruí (Ap = 89) (Table 1).

The values of observed heterozygosity (H_O) were lower than those estimated for the expected heterozygosity (H_E) in all populations of both species (Table 1). *M. pantanalense* exhibited the highest H_O value (0.423), while the H_E was equal to 0.445. Even lower values of H_O and H_E were recorded in *M. amazonicum* populations, with H_O ranging from 0.211 (Mosqueiro) to 0.121 (Tucuruí), while the H_E varied between 0.247 (Santarém) and 0.142 (Tucuruí). In general, significant differences were reported between H_O and H_E values (Bartlett's K-squared = 267.97, df = 1, p-value < 2.2e-16), indicating an excess of homozygotes within populations.

The fixation indexes (F_{IS}) were positive in the three populations of *M. amazonicum* (0.075–0.180), corroborating the excess of homozygotes estimated by heterozygosity values assuming random mating within populations. In the case of *M. pantanalense*, a negative F_{IS} was recorded (−0.1375). The nucleotide diversity ranged from 0.026 to 0.034 in *M. amazonicum* and equal to 0.018 in *M. pantanalense* (Table 1).

Population genetic structure

The fixation coefficient values (F_{ST}) indicated significant differences among all groups (p-value < 0.05), revealing a high genetic structure within *M. amazonicum* (0.184 to 0.237). The lowest value was reported in pairwise comparisons between the populations from Tucuruí and Santarém populations (F_{ST} = 0.184), whereas the highest divergence was identified between samples from Mosqueiro and Santarém (F_{ST} = 0.237). The three populations of *M. amazonicum* showed remarkable high F_{ST} values when compared to *M. pantanalense* (F_{ST} = 0.870–0.928), being the highest divergence recorded between the latter and the population from Mosqueiro in northern/northeastern Brazil (F_{ST} = 0.928) (Fig. 2). These data reveal a deep genetic differentiation between *M. amazonicum* and *M. pantanalense*. The Mantel's test indicated a strong but no signification correlation (0.9454; p-value = 0.204) between the genetic and geographic distances among populations (Supplementary Figure S2).

The Bayesian analysis in the software STRUCTURE was carried out to test the partitioning of the 39 individuals from both species into distinct genetic clusters (K = 1–6). Accordingly, four genetic clusters were identified based on the method suggested by Puechmaille[31] with an optimal value of K = 4 (Supplementary Figure S3). It should be pointed out that the presence of the four genetic groups remained unchanged, regardless of the tested K values (K = 4 to 6) (Supplementary Figure S4).

Three out of the four genetic clusters identified in the present study corresponded to different populations of *M. amazonicum*, as follows: (1) Mosqueiro; (2) Tucuruí; and (3) Santarém. In addition, a greater admixture was observed within the population from Tucuruí (cluster 2), particularly with samples from Mosqueiro (cluster 1). The cluster 4 was strictly composed of samples of *M. pantanalense*, revealing no shared alleles with any population of *M. amazonicum* (Fig. 3a).

The inferences based on Discriminant Analysis of Principal Components (DAPC) also identified four genetic clusters, with 100% of probability of assigning the samples to their clusters (K = 4; BIC = 164). Again, these results reinforce the genetic division among the three populations of *M. amazonicum* (clusters 1 to 3) and *M. pantanalense* (cluster 4) as well (Fig. 3b).

The genetic clusters identified by the abovementioned approaches are reinforced by the divMigrate analysis (Fig. 3c), which revealed high rates of gene flow between the samples from Mosqueiro (cluster 1) and Tucuruí (cluster 2), resulting in a greater admixture between both clusters. In turn, the estimated gene flow between *M. pantanalense* (cluster 4) and the populations of *M. amazonicum* was extremely low (0.04–0.07), suggesting that *M. pantanalense* represents a distinct group in relation to *M. amazonicum*.

Phylogenetic inference by SNPs

The phylogenetic reconstruction based on the 969 SNPs by Maximum Likelihood (ML) (Supplementary Figure S5) and Bayesian Inference (BI) (Supplementary Figure S6) also corroborated the results obtained from other analyses. Both methods supported the presence of two distinct groups: *M. amazonicum* and *M. pantanalense*, with maximum support levels (100 for ML and 1 for BI). The former comprised the 35 representatives from

Species	Samples location	N	Ap	Np	H_O + sd	H_E + sd	F_{IS}	π + sd
M. amazonicum	Mosqueiro – Pará	11	107	378	0.211 ± 0.177	0.221 ± 0.138	0.1804	0.026 ± 0.014
	Tucuruí – Pará	12	89	348	0.121 ± 0.071	0.142 ± 0.086	0.0754	0.027 ± 0.015
	Santarém – Pará	12	105	316	0.203 ± 0.188	0.247 ± 0.136	0.1448	0.034 ± 0.018
M. pantanalense	Bela Vista de Goiás - Goiás	4	235	75	0.423 ± 0.413	0.445 ± 0.144	-0.1375	0.018 ± 0.010

Table 1. Information regarding individuals from the *M. amazonicum* and *M. pantanalense* species used in ddRAD sequencing and statistics of genetic diversity estimated based on the final set of 969 SNP markers. N: Number of individuals per population, Ap: number of private alleles, Np: number of polymorphic loci, H_O: observed heterozygosity, H_E: expected heterozygosity, F_{IS}: inbreeding coefficient, π: nucleotide diversity.

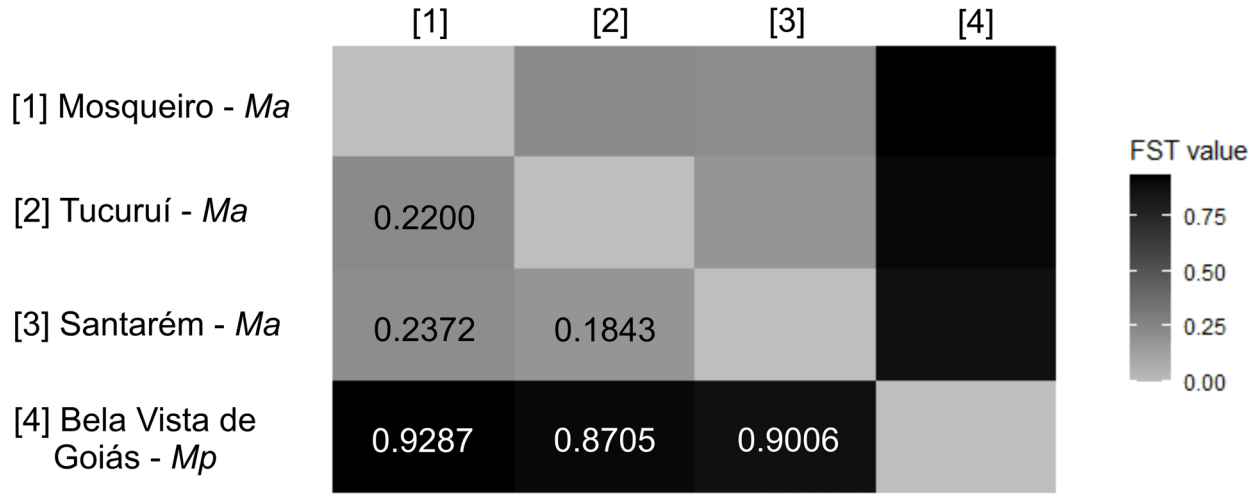

Fig. 2. Pairwise F_{ST} values between the different populations of the two species *M. amazonicum* and *M. pantanalense*, indicating high genetic structuring in *M. amazonicum*, as well as a clear separation between *M. amazonicum* and *M. pantanalense* (highlighted in gray). The comparison between all groups showed significant differences (*p-value* < 0.05).

distinct populations of *M. amazonicum* (Mosqueiro, Tucuruí, and Santarém), while the latter encompassed the four specimens of *M. pantanalense* (Fig. 3d). In addition, the population from Santarém showed high genetic structure (posterior probability = 1 in BI) in relation to the other samples of *M. amazonicum*.

Species delimitation test

The delimitation test measuring the limit of species boundaries, conducted with Bayes Factor Delimitation (BFD*), showed strong support from the Marginal Likelihood (ML) for the model with the hypothesis of the existence of three distinct species (ML: −3881.7764): *M. pantanalense* (Bela Vista de Goiás), *M. amazonicum* coastal cluster (Mosqueiro and Tucuruí) and *M. amazonicum* continental cluster (Santarém). The models with hypotheses of two species (I: Mosqueiro, Tucuruí, Santarém + II: Bela Vista de Goiás) and a single species were the poorly supported (ML: 2940.8119 and 3165.5122, respectively) (Table 2).

Discussion

In the present study, we performed a new evaluation of the phylogenetic relationships within the '*Macrobrachium amazonicum*' complex, using SNP markers obtained via ddRAD sequencing for representatives of the three main lineages of the group, including *M. pantanalense*. Our findings revealed new perspectives on patterns of differentiation, genetic diversity and high population structuring within the complex, allowing us for the first time, based on genetic data, to classify *M. pantanalense* as a distinct species from *M. amazonicum*, as well as providing evidence of a potential new case of ongoing speciation in the continental Amazon population.

Species complex in *M. amazonicum*

Previous molecular reports based on COI sequences have identified three main lineages within *M. amazonicum*, distributed as follows: clade I - continental Amazon basins, clade II - Paraná-Paraguay basin, and clade III – coastal systems from northern and northeastern Brazil. The clades I and II are represented by hololimnetic populations from inland regions, and clade III is composed of diadromous animals dependent on brackish water for their larval development. Such environmental variation across the range of these lineages implies distinct patterns of osmotic regulation accompanied by differential selective pressures that can lead to speciation[16]. Furthermore, the Amazon river prawn also presents a remarkable reproductive plasticity. In fact, some populations from clade III (coastal lineage) have been found in areas with no access to estuarine waters (Rio Grande-São Paulo and Tucuruí-Pará) and are characterized by reduced growth rates, thus resembling inland specimens and hindering the definition of their taxonomic status[21,24].

Most likely, the group *M. amazonicum* complex has undergone a recent separation or speciation that was insufficient to determine deep morphological changes or high levels of genetic divergence in the COI marker. In addition, all populations in this group share an ELD (extended larval development), typical of coastal prawns. On the other hand, our results present population analysis based in SNPs markers revealed, for the first time, significant levels of genetic divergence between *M. pantanalense* and *M. amazonicum*, corroborating the speciation in this species complex. The three structured populations of *M. amazonicum* showed high levels of divergence when compared to *M. pantanalense*, being compatible with interspecific differentiation and reinforcing that both taxa should represent distinct species, corroborating the previous morphological and reproductive analyses[18,20].

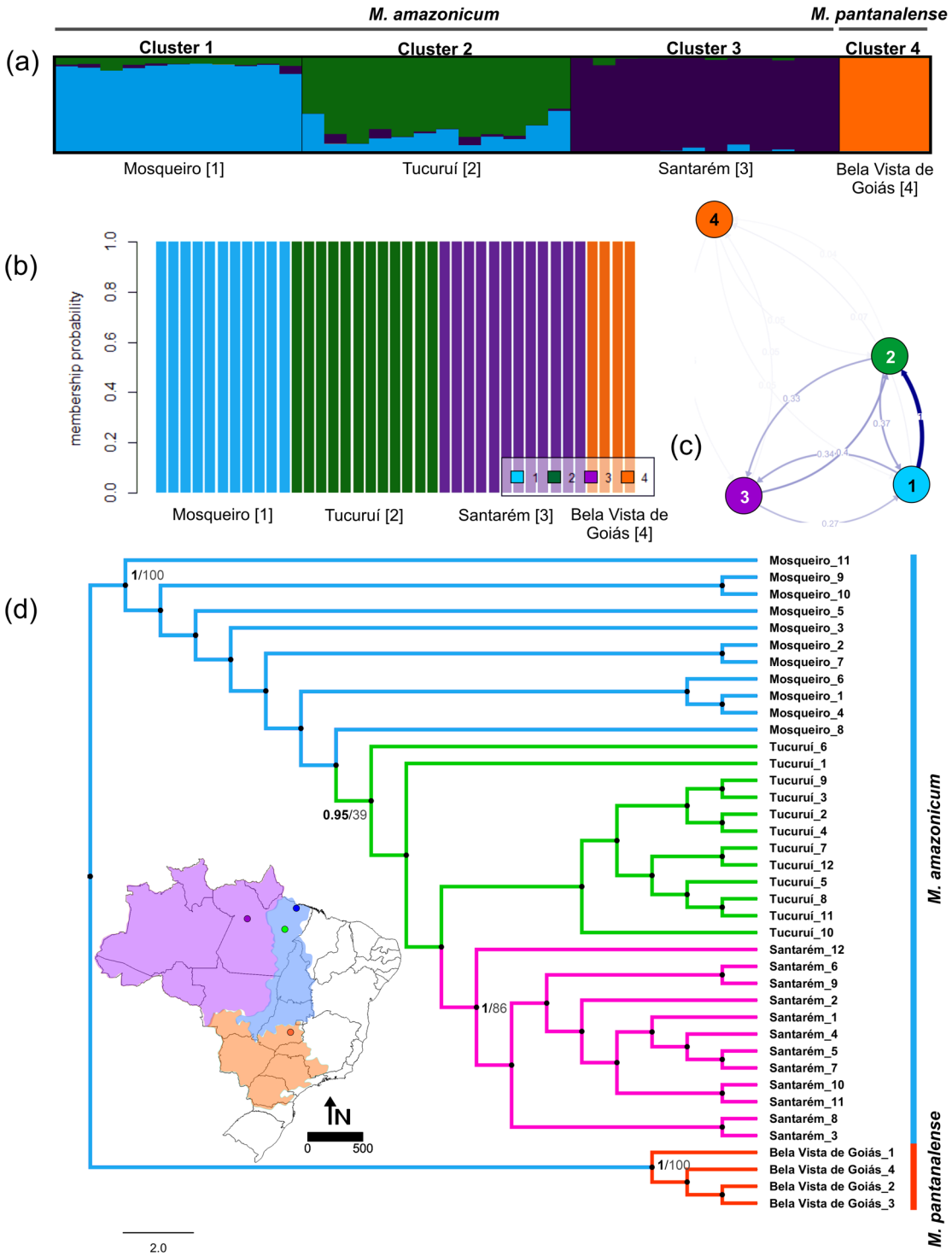

Fig. 3. Structuring of genetic clusters among the 39 individuals of the *M. amazonicum* and *M. pantanalense* species, defined with optimal K = 4, using the Puechmaille method (**a**) and DAPC (**b**), with an extremely low gene flow recorded between the two species (**c**), wherein *M. pantanalense* does not share ancestral alleles with *M. amazonicum* (a). The two species exhibited clear separation in the phylogeny constructed through Maximum Likelihood (support = 100) and Bayesian Inference (posterior probability = 1) based on the 969 SNPs (**d**).

Likewise, this evidence was further supported by the phylogenetic inferences, resulting in maximum support values independent of the tested method. The four sequenced samples from Bela Vista de Goiás (GO) were previously defined as belonging to clade II, now validated as *M. pantanalense*. Nevertheless, since the recognition of *M. pantanalense* as a distinct species is a recent debate associated with the low divergence in COI marker, most of the sequences available in NCBI are referred to as *M. amazonicum* (last access on April, 2nd 2024). In turn, the identification of samples used in the present study as *M. pantanalense* were confirmed based on their high levels of similarity with a single haplotype from Avaré-São Paulo (GU929532.1), showing that the genetic divergence herein inferred could not be attributed to artifact or misidentification. Currently, the *M. pantanalense* lineages comprises regions recorded in Mato Grosso do Sul (Aquidauana and Miranda), Goiás (Bela Vista de Goiás), Minas Gerais (Araguari) and São Paulo (Avaré)[6,18,21].

Accordingly, biological, physiological and morphological reports also support the allocation of *M. pantanalense* as a distinct species from the *M. amazonicum* complex. In relation to their morphology, the populations of the *M. pantanalense* are characterized by shorter total length in relation to other lineages along with the presence of unique structures in the carapace and telson of adults[18,32], and a distinct larval development pattern[19]. Another peculiarity of *M. pantanalense* refers to their reproductive biology since females exhibit lower fecundity associated with the production of large eggs and a similar sex ratio between males and females[21,33,34]. Furthermore, experimental hybridization between hololimnetic individuals from both species within the *M. amazonicum* complex proved to be unfeasible, representing a reproductive barrier putatively associated with recent speciation[20].

In general, the relationships inferred by SNPs among the lineages within the *M. amazonicum* complex were significantly different from those based on COI and 16 S sequences. Previous reports recovered *M. amazonicum* as a monophyletic group whose origin was associated with the continental population of the Amazon region (clade I), followed by subsequent diversification of populations in the Pantanal (Clade II) and northern and northeastern coastal systems (clade III)[6]. The phylogenetic inferences based on SNPs also revealed three main clusters in this group, but characterized by higher levels of genetic differentiation. These data were also reflected in the BFD* species delimitation test, which recorded the highest support for the three-species scenario, grouping *M. pantanalense* as a distinct species, also supporting the separation of the coastal (Mosqueiro and Tucuruí) and continental Amazon groups (Santarém). Furthermore, the model with the hypothesis of a single species in the *M. amazonicum* complex has the lowest support value.

The separation between *M. pantanalense* and *M. amazonicum* in the phylogenetic inferences was supported by maximum support values. The separation between *M. pantanalense* and *M. amazonicum* in the phylogenetic inference was supported by maximum support values. The Santarém population (Clade I) was defined as an internal group within *M. amazonicum*, with a branch having the maximum posterior probability value from the BI, being another insight into the potential speciation event[35]. Despite the absence of the outgroup, the F_{ST} differentiation values, as well as the BFD* analysis, corroborate the phylogenetic arrangement obtained.

Population structure and genetic diversity

In turn, the populations belonging to *M. amazonicum* are genetically structured, particularly when inland and coastal Amazon lineages were compared to each other, evidenced by the high F_{ST} values and high number of private alleles. Contrary to the pattern observed in *M. pantanalense*, the specific status of the lineages from Santarém, Tucuruí and Mosqueiro in Pará could not be resolved because they shared several ancestor alleles with evidence of gene flow. However, the lower levels of gene flow recorded between Santarém and the other *M. amazonicum* populations (Mosqueiro and Tucuruí) were reflected in the reduction of ancestral alleles shared with these populations. Additionally, the high F_{ST} values and BFD* support recorded between Santarém and the other locations reinforce the question of whether the continental Amazon populations (clade I) might be involved in an ongoing speciation process.

Recently, Pereira et al.[35] investigated the phylogenetic relationships of isopod parasites of the genus *Probopyrus* that infect the gills of *M. amazonicum* and revealed the existence of one species strictly related to coastal (*P. bithynis*) and other to continental (*Probopyrus* sp.) populations in the Amazon. This report reinforces the hypothesis of isolation among the populations of the Amazon river prawn related to the species-specific coevolution between both genera of hosts and parasites (coastal *M. amazonicum* x *P. bithynis*; *M. acanthurus* x *P. pandalicola*)[36–38]. These authors also recorded a divergence of ~3% in COI sequences between the prawn samples from Santarém and the coastal groups of *M. amazonicum*, considered one of the highest levels of genetic differentiation ever reported for this marker in the *M. amazonicum* complex. Moreover, the lineage from continental Amazonian regions is composed of unique haplotypes[6,35]suggesting a reduced rate of gene flow with the remaining lineages. Thus, the genetic structuring inferred from the SNPs in the present study corroborates previous studies suggesting a new speciation event among the eastern Amazon populations.

On the other hand, even though the lineages I and III identified in rivers along the Amazon showed high levels of connectivity, they are differentiated by their life histories. The individuals from coastal regions depend on estuarine environments to complete their larval development, where adults reach a larger final length (94 to 160 mm). On the other hand, the prawns from inland waters have osmoregulatory adaptations to complete their whole life cycle in freshwater habitats[15,39–41]. The adults from continental regions are also differentiated, presenting smaller sizes (44 to 136 mm) and early sex maturation. However, the samples from Tucuruí can be regarded as an artificial population since they have become isolated in a lake after the construction of the Tucuruí hydroelectric power Station (UHE) in the 1980s. This event has drastically changed some traits in this local population of the Amazon river prawn, including early sex maturation, reduction in body size and deviation in the sex ratio for females[15,42–44].

The changes in the biological characteristics of the animals from Tucuruí, accompanied by the development of trade-off strategies between energy allocation and reproductive success[45]may be associated with some level of

genetic diversity loss. However, the F_{IS} value indicated a low level of inbreeding. On the other hand, short-term isolation events generally have minimal effects on species diversity and genetic structure. The impacts of habitat fragmentation may take hundreds of years to become identifiable through population genetic parameters, depending on factors such as generation intervals, effective population size, and migration rates[46].

Another factor to be discussed is the F_{ST} values recorded between Mosqueiro and the artificially isolated Tucuruí population, which were similar to those observed between Mosqueiro and Santarém, representing distinct clades, indicating that the Tucuruí population could represent a genetic stock alternative to clades I and III, as the F_{ST} values do not align with the isolation period of approximately 40 years. Another factor supporting this hypothesis is that the number of polymorphic loci identified in Tucuruí is similar to those observed in other *M. amazonicum* populations. Additional studies on the Upper Tocantins populations are necessary to more consistently assess the genetic diversity of this group.

The high population structure within the *M. amazonicum* complex recovered in the present study is likely to reflect the increased resolution of SNPs[47,48]thus providing a deeper understanding about the evolutionary history and the partition of genetic diversity among populations and species. Another aspect to be considered is the absence of Numts and heteroplasmy events in present analyses. On the contrary, both events are frequently recorded in the mtDNA sequences of *M. amazonicum*, which jeopardize the phylogenetic inferences of this group using traditional markers such as COI and 16 S[24]. Furthermore, incomplete lineage sorting of the mtDNA might also account for the relatively low levels of genetic divergence previously recorded in this complex[6,21]as usually observed by the utilization of mitochondrial markers in related lineages under recent speciation processes[49]. Based on our results, it is necessary to expand the sampling regions for the application of this approach, in order to verify if the genetic differentiation recorded with the SNPs is consolidated in different areas of species occurrence, as well as to increase the sampling of coastal and continental populations of *M. amazonicum*.

The genetic diversity indices obtained from the SNPs corroborate that the population clades of *M. amazonicum* (Mosqueiro, Tucuruí, Santarém), as well as *M. pantanalense* (Bela Vista de Goiás), may represent distinct genetic stocks and should be treated differently based on population characteristics and group structure, as previously proposed[6]. These findings may also aid in the proper management of natural populations of both species, being useful for the selection and assembly of breeder panels, as well as for capture management. Furthermore, this study provides data that may serve as a reference for future comparisons with other populations.

Therefore, the analyses of SNP profiles represent a potential tool to resolve the controversial phylogenetic relationships in the *M. amazonicum* complex and other cryptic congeneric groups in the Americas, characterized by high morphological similarity and low molecular divergence in mtDNA markers. Among them, we highlight the debate about *M. americanus* and *M. carcinus*, as well as the putative synonyms between *M. petronioi* vs. *M. potiuna*[4,50] and *M. tenellum* vs. *M. acanthurus*[4,51] or the *M. olfersi* species complex[4,7,52].

Conclusion

We conducted the first analysis of the population structure among the main lineages of *M. amazonicum* in Brazil based on 969 SNPs obtained via ddRAD sequencing. The preset dataset showed a clear separation between *M. amazonicum* and *M. pantanalense* at species level from a genetic perspective, corroborating the bioecological aspects of each, thus addressing one of the ultimate inconsistencies regarding the taxonomic status of this recently described species, evidencing a recent speciation process. Furthermore, we revealed a high genetic structure among populations of *M. amazonicum*, indicating ongoing speciation processes between inland and coastal groups. These results reinforce the need of further studies comprising an increased number of populations to assess the connectivity and levels of effective genetic diversity in populations of *M. amazonicum*.

Materials and methods

Ethics statement

All sampling activities were conducted under the permanent license for the collection of zoological material (SISBIO License No. 12773-1), granted by the Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio) and the Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA), authorizes the collection of specimens in accordance with Brazilian environmental regulations.

Sampling and DNA extraction

A total of 39 samples belonging to the two species of the '*Macrobrachium amazonicum*' complex (*M. amazonicum* and *M. pantanalense*) was analyzed. The specimens of *M. amazonicum* were collected along three regions in Eastern Amazon Basin, Northern Brazil: Mosqueiro-PA (N = 11) (01°04'17.3" S, 48°18'36.3" W), Tucuruí-PA (N = 12) (03°48'22.9" S, 49°44'1.3" W), and Santarém-PA (N = 12) (02°24'58.5" S, 54°33'54.0" W). The individuals of *M. pantanalense* (N = 4) were obtained in Bela Vista de Goiás-GO (16°34'49.7" S, 49°18'38.1" W), Paraná River basin. Afterwards, ~ 100 mg of muscle tissue from each specimen was stored in 70% ethanol for further molecular analyses.

The total genomic DNA was isolated using the Wizard® Genomic DNA Purification kit (Promega, USA), following the manufacturer's instructions. The DNA quality was assessed via agarose gel electrophoresis and quantified using the Qubit 4.0 Fluorometer (Thermo Fisher Scientific, USA), standardized to 6.00 ng/µl. The extracted DNA was used as template for the amplification of COI fragment and preparation of the genomic library.

Amplification and sequencing of the COI gene

A partial fragment of the COI gene was amplified using the COI-A and COI-F primers[53] to confirm the identity of samples. The amplicons were verified by electrophoresis in 1% agarose gel. Subsequently, they were purified

using PEG 8000 following Paithankar and Prasad (1991). The Sanger's sequencing was performed in ABI Genetic Analyzer 3500 XL automatic sequencer (Applied Biosystems).

The COI sequences were aligned in the software Bioedit 7.0.9.0 (Hall, 1999) and compared to the available datasets from the National Center for Biotechnology Information (NCBI) to confirm their identities (https://www.ncbi.nlm.nih.gov/, accessed on March 12th, 2023). Next, the genetic distances were estimated using the Kimura 2-parameter (K2P) model in the software MEGA X[54]. All sequences were submitted to GenBank (PQ426859.1 - PQ426897.1).

Preparation of ddRAD libraries and NGS sequencing

The ddRAD genomic libraries were constructed using EcoRI and MpsI enzymes for genome fragmentation according to the methodology reported by[28] and adapted by[55]. After digestion, P1 and P2 adapters were ligated to the fragments. Subsequently, the amplification was carried out using the Illumina I5 and I7 index set. The purification process was performed after each step using the AMPure® XP beads (Beckman Coulter, USA) to eliminate residual reagents such as enzymes and primers.

The final products were quantified and standardized for pool construction. Next, size selection of fragments (300–500 bp) was carried out in 1% agarose gel. After being selected, the fragments were purified from the agarose gel using the Wizard® SV Gel and PCR Clean-Up System kit (Promega, USA).

Finally, the library was quantified using the Step One Real-Time PCR System (Applied Biosystems) for normalization of the data to be sequenced. The generated product was sequenced along with 2 × 150 bp paired-end (PE) reads following a high-output flow cell, using the Illumina NextSeq500 NGS platform.

Bioinformatic Analyses - SNPs

The quality of the sequenced data was evaluated using FastQC[56] and MultiQC[57] on the distribution of read size, GC%, and presence of adapters. Initially, low-quality reads ($Q < 20$), residual adapters, and reads were standardized to 143 bp using the software TRIMMOMATIC v.0.32[58].

The construction of SNP profile was performed using the pipeline STACKS v.2.66[59], where SNP calling and genotype identification were established under a reference-based approach using the *ref_map.pl* tool. The reference genome of the *Macrobrachium nipponense*, available in NCBI (accession number: GCA_015110555.1), was used to accomplish this step. Subsequently, the database was mapped and organized for SNP calling using the software packages Bowtie2 v.2.2.4[60] and SAMtools[61]respectively. After being identified, the steps of SNP filtering were carried out using the Populations module implemented in Stacks by (i) selecting the genotyped SNPs in more than 70% of samples ($r = 0.7$) and (ii) excluding those SNPs with minimum allele frequency (MAF) < 0.05.

Population genetic differentiation

For population analyses, we initially identified the private alleles and the polymorphic sites using the Populations tool. The software PGDSpider v.2.1.1.5[62] was employed to convert the data into input matrices for analysis in the software packages ARLEQUIN v.3.5.2.2[63] and GENEPOP[64]. In ARLEQUIN, we estimated the observed (H_O) and expected (H_E) heterozygosity values, as well as pairwise F_{ST} genetic distances. In GENEPOP, we calculated the population inbreeding coefficients (F_{IS}). The Mantel's test was carried out to verify the genetic isolation by distance (IBD) by comparing the F_{ST} values and geographical distances between groups based on 1,000 randomizations.

Population genetic structure

The number of genetic clusters within populations was assessed using the software ADEGENET v.2.1.1[65] in R package based on Discriminant Analysis of Principal Components (DAPC). Subsequently, the number of clusters within groups was evaluated by Bayesian Analysis using the software STRUCTURE v.2.3.4[66] following the methodology by Puechmaille[31]. Distinct K values (K = 1 to 6) were tested based on 1,000,000 Markov chain Monte Carlo (MCMC) steps, 10% of burn-in (100,000) and 10 independent iterations for each k-value, assuming the Admixture model and correlated frequencies. The summary of the STRUCTURE results and plot generation were performed using the software Structure Selector[67].

Phylogenetic inferences based on SNPs

Phylogenetic trees were built to represent the relationships among lineages of the '*M. amazonicum*' complex based on Maximum Likelihood (ML) and Bayesian Inference (BI) methods. The phylogenetic reconstruction by ML was performed in the software RAxML v.8.2.1[68] following the GTR-GAMMA evolutionary model from an initial maximum parsimony tree and bootstrap analysis using the autoMRE function to obtain the best tree. The BI tree was generated in the software MrBayes v.3.2.7[69]by performing four simultaneous runs with four Markovian chains each. Initially, a random tree search was performed over 50 million generations, sampling at every 5000 generations and assuming a burn-in of 25%.

Bayes factor delimitation (BFD*)

The Bayes Factor Delimitation (BFD[70]; analysis modified for genome-wide SNP data (BFD*[71]; was carried out to validate the different species identified, using BEAST v2.6.6[72], by SNAPP method[73] to estimate the number of species directly from the SNPs. Due to SNAPP's extensive computational time, a reduced dataset was built containing only four samples from each population group (Mosqueiro, Tucuruí, Santarém and Bela Vista de Goiás - Supplementary Table S1). The different species number scenarios were chosen based on the group's recent definitions and problems, with four distinct scenarios being tested (Table 2). Marginal likelihoods were estimated for each model using 24 path sampling steps, an α-value of 0.3 and a MCMC chain length of 1 000 000 generations (20% burn-in). Tracer v1.7.2[74] was used to check whether the number of generations was

Species scenario	Number of species	Rank	Marginal Likelihood	Bayes Factor
Ma (M, T, S, G) *	1	4	3165. 5122	14,094.578
Ma (M, T, S) + Mp (G)	2	3	2940.8119	13,645.176
Ma (M, T) + Ma (S) + Mp (G)	3	1	-3881.7764	-
Ma (M) + Ma (S) + Ma (T) + Mp (G)	4	2	-3525.6437	712.265

Table 2. Summary of the BFD* species scenarios, including the number of species, rank, marginal likelihood and the Bayes factor. Species in the delimitation scenarios are depicted as initials, as follows: *Ma* = *M. amazonicum*; Mp = *M. pantanalense*; m = mosqueiro, t = tucuruí, s = santarém, b = bela Vista de goiás. *Current scenario by genetics data (mtDNA).

sufficient for the trees to converge. The support for each alternative species delimitation model was assessed by calculating and comparing Bayes Factors (BFs) according to the equation: BF = 2 × (marginal likelihood of model 1 – marginal likelihood of model 2). In this comparison, model 1 represents the alternative that is richer in species. A negative BF value indicates support in favor of model 1, whereas a positive BF value indicates support in favor of model 2.

Data availability

The SRA datasets analyzed during the current study are available for academic purposes upon signing a Material Transfer Agreement with the corresponding author (cmaciel@ufpa.br; vanessa.paes@unesp.br). The COI sequences generated in this study has been deposited at the National Center for Biotechnology Information (NCBI) under the access numbers PQ426859 - PQ426897. The sequences are currently private and will be made public with the publication of this study.

Received: 4 November 2024; Accepted: 26 August 2025
Published online: 29 August 2025

References

1. Molina, W. F. et al. Molecular cytogenetic analysis in freshwater prawns of the genus *Macrobrachium* (Crustacea: decapoda: Palaemonidae). *Int J. Mol. Sci* **21** (7), 2599 (2020).
2. Wowor, D. et al. Evolution of life history traits in Asian freshwater prawns of the genus *Macrobrachium* (Crustacea: decapoda: Palaemonidae) based on multilocus molecular phylogenetic analysis. *Mol. Phylogenet Evol.* **52**, 340–350 (2009).
3. De Grave, S. et al. Benchmarking global biodiversity of decapod crustaceans (Crustacea: Decapoda). *Journal Crustacean Biology* **43**, 1–9 (2023).
4. Pileggi, L. G. & Mantelatto, F. L. Molecular phylogeny of the freshwater Prawn genus *Macrobrachium* (Decapoda, Palaemonidae), with emphasis on the relationships among selected American species. *Invertebr Syst.* **24**, 194–208 (2010).
5. Siriwut, W. et al. Evidence of cryptic diversity in freshwater *Macrobrachium* prawns from Indochinese riverine systems revealed by DNA barcode, species delimitation and phylogenetic approaches. *PLoS One* **16**, (2021).
6. Vergamini, F. G., Pileggi, L. G. & Mantelatto, F. L. Genetic variability of the Amazon river Prawn *Macrobrachium Amazonicum* (Decapoda, caridea, Palaemonidae). *Contrib. Zool.* **80**, 67–83 (2011).
7. Rossi, N. & Mantelatto, F. L. Molecular analysis of the freshwater Prawn *Macrobrachium olfersii* (Decapoda, Palaemonidae) supports the existence of a single species throughout its distribution. *PLoS One* **8**, (2013).
8. Castelin, M., de Mazancourt, V., Marquet, G., Zimmerman, G. & Keith, P. Genetic and morphological evidence for cryptic species in *Macrobrachium australe* and resurrection of *M. ustulatum* (Crustacea, Palaemonidae). *Eur. J. Taxon.* **289**, 1–27 (2017).
9. Siriwut, W., Jeratthitikul, E., Panha, S., Chanabun, R. & Sutcharit, C. Molecular phylogeny and species delimitation of the freshwater prawn *Macrobrachium pilimanus* species group, with descriptions of three new species from Thailand. *PeerJ* 8, (2020).
10. Chaowvieng, A. et al. Molecular phylogeny and taxonomic position of macrobrachium lanchesteri (De man, 1911), with descriptions of two new species from Thailand (Decapoda, caridea, Palaemonidae). *Zookeys* **2024**, 163–193 (2024).
11. Rahi, M. L., Amin, S., Mather, P. B. & Hurwood, D. A. Candidate genes that have facilitated freshwater adaptation by palaemonid prawns in the genus *Macrobrachium*: identification and expression validation in a model species (*M. koombooloomba*). *PeerJ* 5, e2977 (2017).
12. Anger, K. Neotropical Macrobrachium (Caridea: Palaemonidae):on the biology, origin, and radiation of freshwater-invading shrimp. *Journal of Crustacean Biology* vol. 33 151–183 Preprint at (2013). https://doi.org/10.1163/1937240X-00002124
13. Moraes-Riodades, P. M. C. & Valenti, W. C. Morphotypes in male Amazon river prawns, *Macrobrachium Amazonicum*. *Aquaculture* **236**, 297–307 (2004).
14. Karplus, I. & Barki, A. Male morphotypes and alternative mating tactics in freshwater prawns of the genus *Macrobrachium*: a review. *Reviews in Aquaculture* vol. 11 925–940 Preprint at (2019). https://doi.org/10.1111/raq.12273
15. Ramalho Maciel, C. & Valenti, W. C. Biology, fisheries, and aquaculture of the Amazon river Prawn *Macrobrachium amazonicum*: a review. *Nauplius* **17**, 61–79 (2009).
16. Fabri, L. M. et al. A kinetic characterization of the gill V(H+)-ATPase from two hololimnetic populations of the Amazon river shrimp *Macrobrachium Amazonicum*. *Comp Biochem. Physiol. B Biochem. Mol. Biol* **268**, (2023).
17. Pantaleão, J. A. F., Hirose, G. L. & Costa, R. C. Relative growth, morphological sexual maturity, and size of *Macrobrachium Amazonicum* (heller 1862) (crustacea, decapoda, palaemonidae) in a population with an entirely freshwater life cycle. *Invertebr Reprod. Dev.* **56**, 180–190 (2012).
18. Dos Santos, A., Hayd, L. & Anger, K. A new species of *Macrobrachium* spence bate, 1868 (Decapoda, Palaemonidae), *M. pantanalense*, from the Pantanal, Brazil. *Zootaxa* **3700**, 534–546 (2013).
19. Nogueira, C. S., Costa, R. C. & Pantaleão, J. A. F. Variation in larval traits between closely related species of freshwater prawns (*Macrobrachium Amazonicum* and *M. pantanalense*). *Acta Zool.* https://doi.org/10.1111/azo.12489 (2023).
20. Nogueira, C. S., Pantaleão, J. A. F. & Costa, R. C. Hybridisation experiments between freshwater prawns *Macrobrachium Amazonicum* and *M. pantanalense* (Decapoda: Palaemonidae), and the effects of geographical isolation. *Mar. Freshw. Res.* **72**, 520–525 (2021).

21. Calixto-Cunha, M., Rodrigues, T. S., Ueira-Vieira, C., Alves, R., de Almeida, A. C. & D. F. & Genetic and phenotypic variability in populations of the *Macrobrachium Amazonicum* complex: new findings to the upper Paraná hydrographic basin. *Zool. Anz.* **293**, 26–36 (2021).
22. Morin, P. A., Luikart, G. & Wayne, R. K. SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution* 19, 208–216 Preprint at (2004). https://doi.org/10.1016/j.tree.2004.01.009
23. Chow, S., Yanagimoto, T. & Takeyama, H. Detection of heteroplasmy and nuclear mitochondrial pseudogenes in the Japanese spiny Lobster *Panulirus japonicus*. *Sci Rep* **11**, 21780 (2021).
24. Iketani, G., Pimentel, L., Torres, E. S., do Rêgo, P. S. & Sampaio, I. Mitochondrial heteroplasmy and pseudogenes in the freshwater prawn, macrobrachium Amazonicum (Heller, 1862): DNA barcoding and phylogeographic implications. *Mitochondrial DNA DNA Mapp. Seq. Anal.* **32**, 1–11 (2021).
25. Garvin, M. R., Saitoh, K. & Gharrett, A. J. Application of single nucleotide polymorphisms to non-model species: a technical review. *Mol. Ecol. Resour.* **10**, 915–934 (2010).
26. Wellenreuther, M., Mérot, C., Berdan, E. & Bernatchez, L. Going beyond snps: the role of structural genomic variants in adaptive evolution and species diversification. *Mol. Ecol.* **28**, 1203–1209 (2019).
27. Baird, N. A. et al. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* **3** (10), e3376 (2008).
28. Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S. & Hoekstra, H. E. Double digest radseq: an inexpensive method for de Novo SNP discovery and genotyping in model and non-model species. *PLoS One* **7** (5), e37135 (2012).
29. Alam, M. M. M., Westfall, K. M. & Pálsson, S. Historical demography and genetic differentiation of the giant freshwater Prawn *Macrobrachium Rosenbergii* in Bangladesh based on mitochondrial and DdRAD sequence variation. *Ecol. Evol.* **7**, 4326–4335 (2017).
30. Makombu, J. G. et al. Species-informative SNP markers for characterising freshwater prawns of genus *Macrobrachium* in Cameroon. *PLoS One* **17**, 10 (2022).
31. Puechmaille, S. J. The program STRUCTURE does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Mol. Ecol. Resour.* **16**, 608–627 (2016).
32. Nogueira, C. S., Camargo, N. F., Pantaleão, J. A. F. & Costa, R. C. Elucidating taxonomic problems of two closely related freshwater Prawn lineages of the genus *Macrobrachium* (Caridea: Palaemonidae): A geometric morphometrics approach. *Zool. Anz.* **304**, 73–83 (2023).
33. Hayd, L. & Anger, K. Reproductive and morphometric traits of *Macrobrachium Amazonicum* (Decapoda: Palaemonidae) from the pantanal, brazil, suggests initial speciation. *Rev. Biol. Trop.* **61**, 39–57 (2013).
34. Meireles, A. L., Valenti, W. C. & Mantelatto, F. L. Reproductive variability of the Amazon river prawn, *Macrobrachium Amazonicum* (Caridea, Palaemonidae): influence of life cycle on egg production. *Lat Am. J. Aquat. Res.* **41**, 718–731 (2013).
35. Pereira, R. I. S., Maciel, C. R. & Iketani, G. Molecular features of probopyrus sp. (Isopoda: Bopyridae) from Brazilian Amazonia and the parasitism of inland populations of macrobrachium Amazonicum (Decapoda: Palaemonidae). *Parasitology* **149**, 203–208 (2022).
36. Odinetz Collart, O. Interactions entre Le parasite *Probopyrus bithynis* (Isopoda, Bopyridae) et l'un de Ses hôtes, La Crevette *Macrobrachium Amazonicum* (Decapoda, Palaemonidae). *Crustaceana* **58**, 258–269 (1990).
37. Corrêa, L. L. et al. Histopathological alterations in gills of Amazonian shrimp *Macrobrachium Amazonicum* parasitized by isopod *Probopyrus bithynis* (Bopyridae). *Dis. Aquat. Organ.* **129**, 117–122 (2018).
38. de Barros, M. S. F., da Silva Neto, L. S. & Calado, T. C. S. First record of parasitism by probopyrus Pandalicola (Isopoda, Bopyridae) on the freshwater Prawn macrobrachium acanthurus (Decapoda, Palaemonidae) and ecological interactions. *J. Parasitic Dis.* **45**, 273–278 (2021).
39. Magalhães, C. Desenvolvimento larval obtido Em laboratório de palaemonídeos Da Região Amazônica. *Amazoniana* **9**, 247–274 (1985).
40. Guest, W. C. Laboratory life history of the palaemonid shirmp *Macrobrachium Amazonicum* (Heller) (Decapoda, Palaemonidae). *Crustaceana* **37**, 141–152 (1979).
41. Soeiro, R. K. S., Rocha, C. P., Maciel, M., Abrunhosa, F. A. & Maciel, C. R. Relationship between the coastal origin of the freshwater Prawn *Macrobrachium Amazonicum* and salinity levels in the hatchery. *Bol. Inst. Pesca*. **42**, 691–703 (2016).
42. Odinetz Collart, O. Tucurui dam and the populations of the Prawn *Macrobrachium Amazonicum* in the lower Tocantins (Pa-Brazil): a four year study. *Arch. Hydrobiol.* **121**, 213–227 (1990).
43. Silva, G. M. F. et al. Has a river dam affected the life-history traits of a freshwater prawn? *Ecol. Evol.* **10**, 6536–6548 (2020).
44. Freire, J. L., Bentes, B. & Fontes, V. B. Da silva, E. M. Morphometric discrimination among three stocks of *Macrobrachium Amazonicum* in the Brazilian Amazon. *Limnologica* **64**, 1–10 (2017).
45. Silva, G. M. F. et al. Energy allocation trade-off in *Macrobrachium amazonicum*, with no resting stage. *Zool. Anz.* **293**, 1–8 (2021).
46. Allendorf, F. W., Luikart, G. & Aitken, S. N. Conservation and the genetics of populations. *John Wiley & Sons* (2013).
47. Davey, J. L., Blaxter, M. W. & RADseq Next-generation population genetics. *Brief. Funct. Genomics*. **9**, 416–423 (2010).
48. Wenne, R. Single Nucleotide Polymorphism Markers with Applications in Conservation and Exploitation of Aquatic Natural Populations. *Animals* vol. 13 Preprint at (2023). https://doi.org/10.3390/ani13061089
49. Steenwyk, J. L., Li, Y., Zhou, X., Shen, X. X. & Rokas, A. Incongruence in the phylogenomics era. *Nature Reviews Genetics* vol. 24 834–850 Preprint at (2023). https://doi.org/10.1038/s41576-023-00620-x
50. de Carvalho, F. L., Pileggi, L. G. & Mantelatto, F. L. Molecular data Raise the possibility of cryptic species in the Brazilian endemic Prawn *Macrobrachium Potiuna* (Decapoda, Palaemonidae). *Lat Am. J. Aquat. Res.* **41**, 707–717 (2013).
51. Meza-Sánchez, I. G. et al. The systematics of the amphidromous shrimp *Macrobrachium tenellum* (Smith, 1871) (Decapoda: caridea: Palaemonidae) from the Mexican Pacific slope. *Journal Crustacean Biology* **44** (1), (2024).
52. Pileggi, L. G. & Mantelatto, F. L. Taxonomic revision of doubtful Brazilian freshwater shrimp species of genus *Macrobrachium* (Decapoda, Palaemonidae). *Iheringia Ser. Zool.* **102**, 426–437 (2012).
53. Palumbi, S. R. & Benzie, J. Large mitochondrial DNA differences between morphologically similar Penaeid shrimp. *Mol. Mar. Biol. Biotechnol.* **1**, 27–34 (1991).
54. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
55. Campos, M. et al. Microgeographical structure in the major Neotropical malaria vector *Anopheles Darlingi* using microsatellites and SNP markers. *Parasit. Vectors*. **10**, 1–8 (2017).
56. Andrews, S. FastQC: a quality control tool for high throughput sequence data. *Babraham Bioinformatics* Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (2010).
57. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
58. Bolger, A. M., Lohse, M., Usadel, B. & Trimmomatic A flexible trimmer for illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
59. Rochette, N. C., Rivera-Colón, A. G. & Catchen, J. M. Stacks 2: analytical methods for paired-end sequencing improve RADseq-based population genomics. *Mol. Ecol.* **28**, 4737–4754 (2019).
60. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with bowtie 2. *Nat. Methods*. **9**, 357–359 (2012).
61. Li, H. et al. The sequence alignment/map format and samtools. *Bioinformatics* **25**, 2078–2079 (2009).

62. Lischer, H. E. L. & Excoffier, L. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* **28**, 298–299 (2012).
63. Excoffier, L. & Lischer, H. E. L. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and windows. *Mol. Ecol. Resour.* **10**, 564–567 (2010).
64. Rousset, F. Genepop'007: a complete re-implementation of the Genepop software for windows and Linux. *Mol. Ecol. Resour.* **8**, 103–106 (2008).
65. Jombart, T. & Ahmed, I. Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**, 3070–3071 (2011).
66. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
67. Li, Y. & Liu, J. STRUCTURE SELECTOR: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Mol. Ecol. Resour.* **18**, 176–177 (2018).
68. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
69. Ronquist, F. et al. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
70. Grummer, J. A., Bryson, R. W. & Reeder, T. W. Species delimitation using Bayes factors: simulations and application to the *Sceloporus scalaris* species group (Squamata: Phrynosomatidae). *Syst. Biol.* **63**, 119–133 (2014).
71. Leaché, A. D., Fujita, M. K., Minin, V. N. & Bouckaert, R. R. Species delimitation using Genome-Wide SNP data. *Syst. Biol.* **63**, 534–542 (2014).
72. Bouckaert, R. et al. BEAST 2.5: an advanced software platform for bayesian evolutionary analysis. *PLoS Comput. Biol.* **15**, e1006650 (2019).
73. Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A. & RoyChoudhury, A. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Mol. Biol. Evol.* **29**, 1917–1932 (2012).
74. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior summarization in bayesian phylogenetics using tracer 1.7. *Syst. Biol.* **67**, 901–904 (2018).
75. Kass, R. E. & Raftery, A. E. Bayes factors. *J. Am. Stat. Assoc.* **90**, 773–795 (1995).

Acknowledgements

The authors would like to thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, grant number: 88887.511314/2020-00). This work has been granted by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq – grant number: 407952/2023-3), FAPESPA: Fundação Amazônia de Amparo a Estudos e Pesquisas do estado do Pará (grant number 001/2017 FAPESPA/SECTET), Amazonia +10 (FAPESPA – grant number 173/2023). The authors would also like to thank the Institute of Biotechnology (IBTEC) - Botucatu and the technical team.

Author contributions

G. M. L., L.D.Q., G.I., conceived the study and provided samples. G.M.L., V.P.C. B.R.B., C.R.M. generated the data. G.M.L., B.R.B., C.M.T.M., analysed the data. G.M.L., V.P.C., C.R.M., wrote the manuscript. G.M.L., C.O., F.F., W.C.V., I.S., V.P.C., C.R.M., revised the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-17666-y.

Correspondence and requests for materials should be addressed to V.P.C. or C.R.M.