GBE

The Golden Lancehead Genome Reveals Distinct Selective Processes Acting on Venom Genes of an Island Endemic Snake

Pedro G. Nachtigall [1,2], Andrew J. Mason [3], Darin R. Rokyta [2], H. Lisle Gibbs [3], Felipe G. Grazziotin [4], Inácio L. M. Junqueira-de-Azevedo [1,*]

[1]Laboratório de Toxinologia Aplicada, CeTICS, Instituto Butantan, São Paulo, SP, BR

[2]Department of Biological Science, Florida State University, Tallahassee, FL, USA

[3]Department of Evolution, Ecology, and Organismal Biology, Ohio State University, Columbus, OH, USA

[4]Laboratório de Coleções Zoológicas, LCZ, Instituto Butantan, São Paulo, SP, BR

*Corresponding author: E-mail: inacio.azevedo@butantan.gov.br.

Accepted: December 04, 2025

Abstract

High-quality genomic resources are important for accurate assessments of adaptive evolution in rapidly evolving island endemic species. The golden lancehead (*Bothrops insularis*) is a critically endangered venomous species endemic to the Queimada Grande island located on the southeast coast of Brazil, with no reference genome available. Here, we present a high-quality near chromosome-level and well-annotated genome for the golden lancehead. A macrosyntenic analysis using genomes from other viper species revealed that microchromosomes present higher rearrangements consisting of fission and fusion events. Using our genome and genomic data from eight individuals, we conducted a survey of the genetic variation of toxin genes, which included the nucleotide diversity and copy-number variation (CNV). We also inferred a demographic history for the species in the last 100,000 years. The genetic variation analysis revealed that major components of *B. insularis* venom appear to be evolving largely under natural selection processes rather than genetic drift as expected for an insular species. PLA2s and CTLs are under balancing selection, whereas SVMPs and SVSPs are under positive selection. The CNV suggests recent duplication events in SVMPs and CTLs and deletion events in SVSPs and PLA2s. The demographic history indicates a stable population size over the last 10,000 years, suggesting that *B. insularis* is both genetically and demographically healthy. Altogether, we provide a genomic resource to better understand the differentiation of an iconic snake and evidence that selection has driven the evolution of diverse venom genes over short evolutionary timescales in an insular species.

Key words: snake, toxin evolution, island, venomics, conservation.

Significance

Understanding how natural selection and genetic drift shape the variation of adaptive and complex traits is a central topic in biology. Snake venoms comprise a complex phenotype that is often assumed to be under strong positive selection and to evolve rapidly. Studies focusing on the recent insularization of venomous snakes may help us better depict the evolutionary processes in shaping a complex phenotype. Here, we present the first chromosome-level genome assembly of the golden lancehead and use intraspecific genomic data to assess the roles played by evolutionary processes in toxin genes. Our study reveals that the golden lancehead toxin genes face distinct selection forces rather than genetic drift, which could result from the availability of prey items on the island. We provide a genomic resource for an iconic snake species and evidence that selection may play an important role in the evolution of diverse venom genes over short evolutionary timescales in an island species with a small population size.

Introduction

Snakes compose a monophyletic group of Squamata, which represents one of the most speciose vertebrate radiations (Bellairs and Underwood 1951; Stanley 1985; Title et al. 2024). This expansion resulted in more than 3,900 snake species inhabiting nearly all environments worldwide (Uetz et al. 2021). The evolutionary expansion of snakes is associated with a suite of phenotypic lineage-specific innovations, including an elongated and limbless body, highly kinetic skulls, and advanced chemosensory systems, which were accompanied by a sophisticated venom delivery system in some lineages (Rieppel 1988; Da Silva et al. 2018; Title et al. 2024). Such traits have promoted the evolution and diversification of specialized predatory behaviors and diets in snakes (Greene 1997; Grundler and Rabosky 2021). Despite their remarkable diversity and broad distribution, some snake species face severe threats and require urgent extinction risk assessments and targeted conservation actions, particularly those restricted to small or fragmented areas, such as island endemics (Roll et al. 2017; Cox et al. 2022).

The golden lancehead, *Bothrops insularis*, is a critically endangered species endemic to the Queimada Grande Island, 30 km off the southeastern coast of Brazil. The species belongs to the most diverse viperid genus in South America, which is responsible for most envenoming cases in the region (Magalhães et al. 2019). *B. insularis* is part of the *B. jararaca* species complex, a diverse group distributed across the Brazilian Atlantic Forest that includes four other insular species (*B. alcatraz*, *B. germanoi*, *B. otavioi*, and *B. sazimai*) and the mainland *B. jararaca* (Marques et al. 2002; Barbo et al. 2012, 2016, 2022). Although closely related, these species exhibit markedly distinct ecology and phenotypic traits (Martins et al. 2002; Barbo et al. 2022). For instance, *B. insularis* is commonly found foraging in trees, is highly active during the day, and displays a yellowish-golden coloration. All these traits are absent in other insular species and extremely rare in mainland populations (Marques et al. 2002). Unlike other species in the complex, *B. insularis* relies predominantly on migrant passerine birds, which account for over 80% of its diet, making it the only member of the complex classified as a bird specialist (Martins et al. 2002).

The phenotypic uniqueness of *B. insularis* becomes even more remarkable given its relatively recent speciation time (reviewed in Barbo et al. 2022). The most probable evolutionary scenario for the high diversity of insular species in the complex comprises the effects of recent sea-level rises, which isolated ancestral populations of the mainland distributed along the coast (Marques et al. 2002). These populations became trapped on the easternmost highlands of the continental shelf that now form the series of continental islands off the Brazilian coast (Vanzolini 1973; Leite et al. 2016).

Queimada Grande Island has a small area of approximately 0.43 km^2, although its size changed considerably over its geological history due to repeated insularization events associated with the last glacial-interglacial cycles during the Pleistocene and early Holocene (Furtado et al. 1998). Sea-level records and bathymetric data indicate three probable periods for recent insularization, approximately 9.5 kya (9–11 kya), 12 kya (10–15 kya) and 55 kya (50–60 kya) (reviewed in Barbo et al. 2022). Even if the oldest estimate marks the speciation time of *B. insularis*, the number of fixed phenotypic differences from *B. jararaca* implies an unusually rapid rate of divergence in the golden lancehead. On the other hand, studies based on a limited number of genetic markers have suggested a close genetic proximity between *B. insularis* and other species of the complex, with mitochondrial phylogenies allocating the species to a position nested within the genetic diversity of mainland *B. jararaca* (Grazziotin et al. 2006; Barbo et al. 2012, 2022).

In contrast to the well-documented morphological and ecological differences (Martins et al. 2002; Wüster et al. 2005), the uniqueness of venom activity and composition of *B. insularis* has never been properly evaluated within a framework accounting for the extreme population-level variation of mainland *B. jararaca* (e.g., Gonçalves-Machado et al. 2016). Consequently, all

findings reported to date may be biased by the inclusion of different geographic variants from the continent in the comparative analyses. Despite this, much attention has been given to the venom composition of *B. insularis*, and the species was the first snake to have its venom gland transcriptome analyzed through Sanger-based Expressed Sequence Tags (ESTs) in early 2000s (Junqueira-de Azevedo and Ho 2002), and later characterized using MS/MS-based proteomics (Valente et al. 2009). These studies revealed a venom rich in snake venom metalloproteinases (SVMPs), snake venom serine proteases (SVSPs), C-type lectins (CTLs), phospholipase A_2 (PLA2), bradykinin-potentiating peptide (BPP), and other components, which were later found to be prevalent in other viper venoms (Junqueira-de Azevedo et al. 2015; Freitas-de Sousa et al. 2020; Sousa et al. 2021; Nachtigall et al. 2022; Oliveira et al. 2022). Other molecules exclusive to snake venom glands and unknown at the time, such as snake venom vascular endothelial growth factor (Junqueira-de Azevedo et al. 2001) and calglandulin (Junqueira-de Azevedo et al. 2003), were first identified in *B. insularis*. Functional assays showed that the diverse set of toxins in *B. insularis* venom leads to hemorrhagic, anticoagulant, myotoxic, and hypotensive effects (Selistre et al. 1990; Cogo 1991; Cogo et al. 1993; Barbosa et al. 2003; Braga et al. 2006; Cogo et al. 2006; Machado Braga et al. 2008).

Some studies have suggested that the biological effects of *B. insularis* venom may be more toxic to birds than to mammals (Cogo 1991; Cogo et al. 1993; Zelanis et al. 2008), likely the result of positive selection driven by the adaptation to its specialized diet. However, compositional analyses have not identified any significant differences in its toxin arsenal compared to the venom of the mainland *B. jararaca* (Zelanis et al. 2008). Minor structural differences may underlie this supposed specificity, but this remains unknown. One of the main barriers to clarifying these differences is the absence of a comprehensive genome detailing venom variation and genomic structure within the species complex. Such genomic information could also provide an opportunity to investigate the mechanisms that generated trait differences over a very short evolutionary timescales (Grazziotin et al. 2006).

In addition, the conservation status of *B. insularis* creates the need for genomic information even more pressing, as the golden lancehead is currently listed as "Critically Endangered" on the International Union for Conservation of Nature (IUCN) Red List (Silveira Aea 2025) and by the Brazilian environmental agency "Instituto Chico Mendes de Conservação da Biodiversidade" (ICMBio). Some studies have also raised concerns about a possible population decline in recent years (Martins et al. 2008), but no solid conclusion has been reached (Guimarães et al. 2014). Deterministic threats caused by anthropogenic factors, such as illegal trade (Martins et al. 2008) and climate changes, may be directly affecting the species, while also triggering stochastic threats through changes in the genetic diversity and demography. Characterizing patterns of genomic variation—particularly fitness-related variation, such as in toxin genes—can provide a foundation for improved management strategies for this endangered species.

Genome sequencing has proven essential for a comprehensive understanding of the evolutionary processes underlying multiple phenotypic variations, and it is particularly crucial in cases of insular rapid evolution (e.g., Card et al. 2019). The current capacity to generate high-quality genomic data provides powerful tools for more accurate estimations of genetic diversity, population structure, and the demographic history of target species. Such high-quality reference genomes are vital for guiding modern conservation strategies, as they allow precise tracking of genetic variation within and between populations, which is critical for understanding adaptive processes and assessing the evolutionary genetic repertoire of threatened taxa (Supple and Shapiro 2018; Hohenlohe et al. 2021; Hogg 2024). Advances in sequencing and bioinformatics technologies have greatly facilitated the assembly of these genomes (Giani et al. 2020). In particular, integrating accurate long-reads with chromatin conformation capture and chromosome interaction data allows the production of highly contiguous, near telomere-to-telomere, and error-free genome assemblies (Rhie et al. 2021).

Here, we leverage the first chromosome-level genome assembly of *B. insularis*, providing a valuable resource for genetic analyses aiming to elucidate its evolutionary history and enhance the conservation management of this critically endangered species. Additionally, we compare this genome with that of *B. jararaca* and use genomic data of other *B. insularis* individuals to assess the genomic context and genetic variation of toxin genes in the species. We analyzed the nucleotide diversity and copy number variation (CNV) of toxin genes to trace the selective pressure and gene-family evolution modes acting upon the venom evolution and diversification in the species. We also inferred the recent demographic history to better understand the population size dynamics since the last glacial period and check for the need for better conservation acts. Finally, we performed a macrosyntenic analysis among available chromosome-level genomes of vipers to assess large-scale genomic evolution in the clade.

Results

Chromosome-level Genome Assembly

The *B. insularis* genome assembly, derived from a homogametic male, produced a total of 110 scaffolds and presented higher quality than other available *Bothrops* genomes (Almeida et al. 2021; Nachtigall et al. 2024;

Table 1. Genomics statistics comparing the *B. insularis* genome to the closely related *Bothrops* species with available genomes

	B. insularis	*B. jararaca*	*B. alternatus*
Assembly level	Chromosome	Draft	Draft
Assembly size (Gb)	1.6	1.7	1.7
# of scaffolds	110	27,698	–
Scaffold N50	206 Mb	164 Kb	–
# of contigs	698	28,752	1,557
Contig N50	14 Mb	163 Kb	14 Mb
GC content (%)	40.5	40.5	40.5
BUSCO*	95.8%	79.3%	95.8%
QV score	38.26	NA	36.78
Genes**	21,887	NA	29,245
Venom genes	72	51	59
Reference	Present study	Almeida et al. (2021)	Nachtigall et al. (2024)

NA: Not analyzed. * Percentage of complete BUSCO genes using the Tetrapoda gene set (odb10; total of 5,310 genes). ** Distinct gene annotation approaches were applied among studies.

Table 1). We were able to reach the chromosome-level genome assembly for the *B. insularis* and properly assign their chromosomes based on chromosomal markers (Matsubara et al. 2006) and genomic alignments to other two chromosome-level genome assemblies of *Crotalus* species (Schield et al. 2019; Hogan et al. 2024; Nachtigall et al. 2025). We detected that the longest 18 scaffolds represent the 18 chromosomes expected for *Bothrops* male individuals (Beçak et al. 2008). The 18 chromosomes recovered consisted of seven autosomal macrochromosomes, one sex macrochromosome (Z), and 10 autosomal microchromosomes (Figs. 1a and 1b). The haploid genome assembly size of chromosomes achieved a size of 1.54 Gb and a scaffold N50 of 206.9 Mb. The assembly was highly accurate with a QV score of 38.27 and it was shown to have high completeness by acquiring a BUSCO of 95.7% complete genes (92.2% single-copy and 3.5% duplicated), 1.0% fragmented, and 3.3% missing when using the Tetrapoda gene set, which contains a total of 5,310 genes. Moreover, we identified telomeric repeats at both ends of four chromosomes and on one end of additional 11 chromosomes, further attesting to the completeness of the assembly (Supplementary Fig. S1a). The chromosome-level genome assembly of *B. insularis* present a similar quality to that obtained for previously published chromosome-level genome assemblies of snakes (Schield et al. 2019; Peng et al. 2020; Hogan et al. 2024). The assembled mitochondrial genome obtained a size of 17,519 bp (Supplementary Fig. S2) and returned a total of 38 genes, comprising 13 protein-coding genes, two ribosomal RNAs, and 23 transporter RNAs as expected for snake species (Yan et al. 2008). It presented a percent identity of 98.56% to the *B. jararaca* mitochondrial genome (Almeida et al. 2016), which is in accordance with previous studies using mitochondrial genes for phylogenetic inferences of the *B. jararaca* complex (Grazziotin et al. 2006).

The repeat annotation revealed 48.54% of repetitive sequences, which comprises 42.90% of transposable elements (TEs) and 5.64% of tandem repeats (Figs. S1b and S1c). Among TEs, LINEs were the most abundant elements (20.22%), which is in accordance with previous studies analyzing the repeat content of squamates and snakes (Castoe et al. 2013; Pasquesi et al. 2018). We also tested for recent and ancient bursts of TE activity by calculating the Kimura substitution levels of repeat sequences. This analysis revealed recent bursts of TEs mainly accounted for LINE elements, which were previously discussed to be involved in genome rearrangements of snake genomes (Pasquesi et al. 2018).

The gene annotation identified 21,887 protein-coding genes, of which 19,882 (90.83%) matched with high similarity to entries in the ENSEMBL database. Through homology and orthology inferences, we assigned meaningful functional annotations to 19,955 genes (91.17% of the total). ToxCodAn-Genome (Nachtigall et al. 2024) with manual curation allowed us to identify 72 venom protein-coding genes (Supplementary Table S1) comprising a set of known major and minor components of viper venom (reviewed in Oliveira et al. 2022). The venom-gland transcriptome from the genome individual (Fig. 1c; Supplementary Table S2) showed that the seven major components of the venom were snake venom metalloproteinases (SVMPs), snake venom serine proteases (SVSPs), C-type lectins (CTLs), bradykinin-potentiating peptide (BPP), L-amino acid oxidase (LAAO), snake venom vascular endothelial growth factor (VEGF), and phospholipase A_2 (PLA2) as previously observed (Junqueira-de Azevedo and Ho 2002; Valente et al. 2009). The major components are highly expressed in venom glands when compared to other tissues (Fig. 1d). SVMPs, SVSPs, and PLA2s are organized in tandem arrays and located in microchromosomes, CTLs are mainly organized in tandem array in a macrochromosome, LAAO was detected as a duplicated locus in a macrochromosome, and all other toxin genes were mainly detected in macrochromosomes and organized as single loci as previously observed in other Crotalinae species (Margres et al. 2021; Hogan et al. 2024; Supplementary Table S1).

Macrosynteny

To check for syntenic blocks and chromosome rearrangements in the *B. insularis* genome when compared to other viperids, we performed a whole-genome synteny analysis of viper species with near chromosome-level genome assemblies available (Fig. 2a). We noticed large syntenic

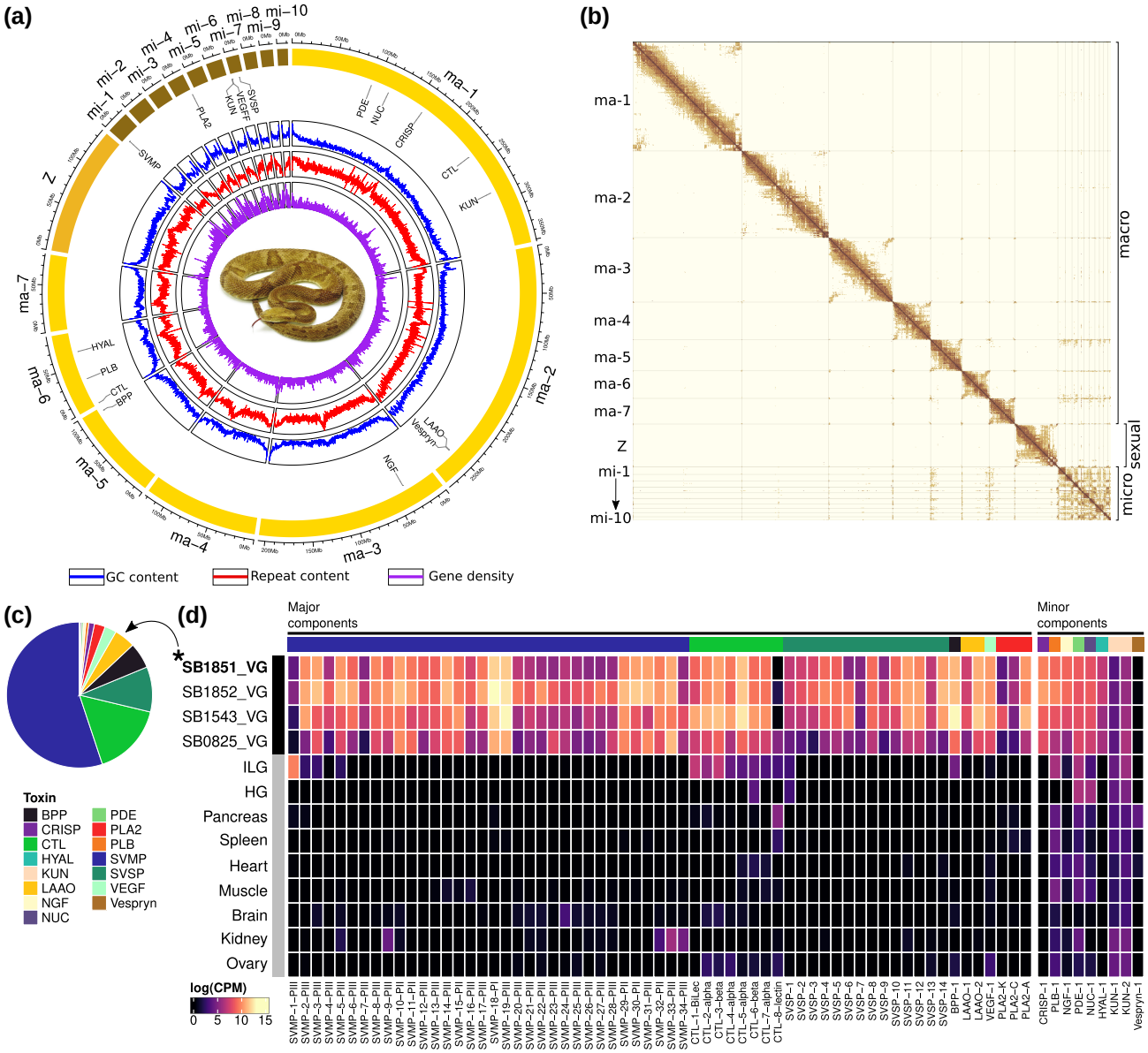

Fig. 1. An overview of the chromosome-level genome assembly of *B. insularis* and the toxin gene annotation. a) A circos plot shows the distribution of venom genes across the genome. Circular rings from inner to outer display the gene density (purple), repeat content (red), and GC content (blue) within 100-Kb windows. The macro-, micro-, and sex chromosomes are colored in light yellow, brown, and dark yellow, respectively. b) A Hi-C contact map for all 18 assembled chromosomes with darker colors indicating stronger interactions. c) Venom gland expression profile of the genome individual (highlighted in bold and marked with an asterisk). d) Expression profile of toxin genes among several tissues and individuals. BPP, Bradykinin-potentiating peptide. CRISP, Cysteine-rich secretory protein. CTL, C-type lectin. HYAL, Hyaluronidase. KUN, Kunitz-type protease inhibitor. LAAO, L-amino acid oxidase. NGF, Nerve growth factor, NUC, 5′-Nucleotidase. PDE, Phosphodiesterase. PLA2, Phospholipase A_2. PLB, Phospholipase B. SVMP, snake venom metalloproteinase. SVSP, snake venom serine protease. VEGF, snake venom vascular endothelial growth factor. VG, venom gland. ILG, infralabial gland. HG, harderian gland.

blocks within macro- and microchromosomes. There are few rearrangements in macrochromosomes, like translocations and inversions in macrochromosomes 1 and 2, mainly affecting the end of chromosomes. There is a specific fission event in macrochromosome 3 detected in *Gloydius shedaoensis*, as previously observed (Peng et al. 2023), which represents a lineage-specific fission event. A few rearrangements in the Z chromosome were observed, including a large inversion in *Sistrurus catenatus* and a small inversion and translocation in *G. shedaoensis*. In microchromosomes, which present the most rearrangement events, we observed fission and fusion events among distinct chromosomes and translocation events within homologous chromosomes.

Additionally, we calculated the GC content and the mean gene density in 1Mb windows per chromosome for

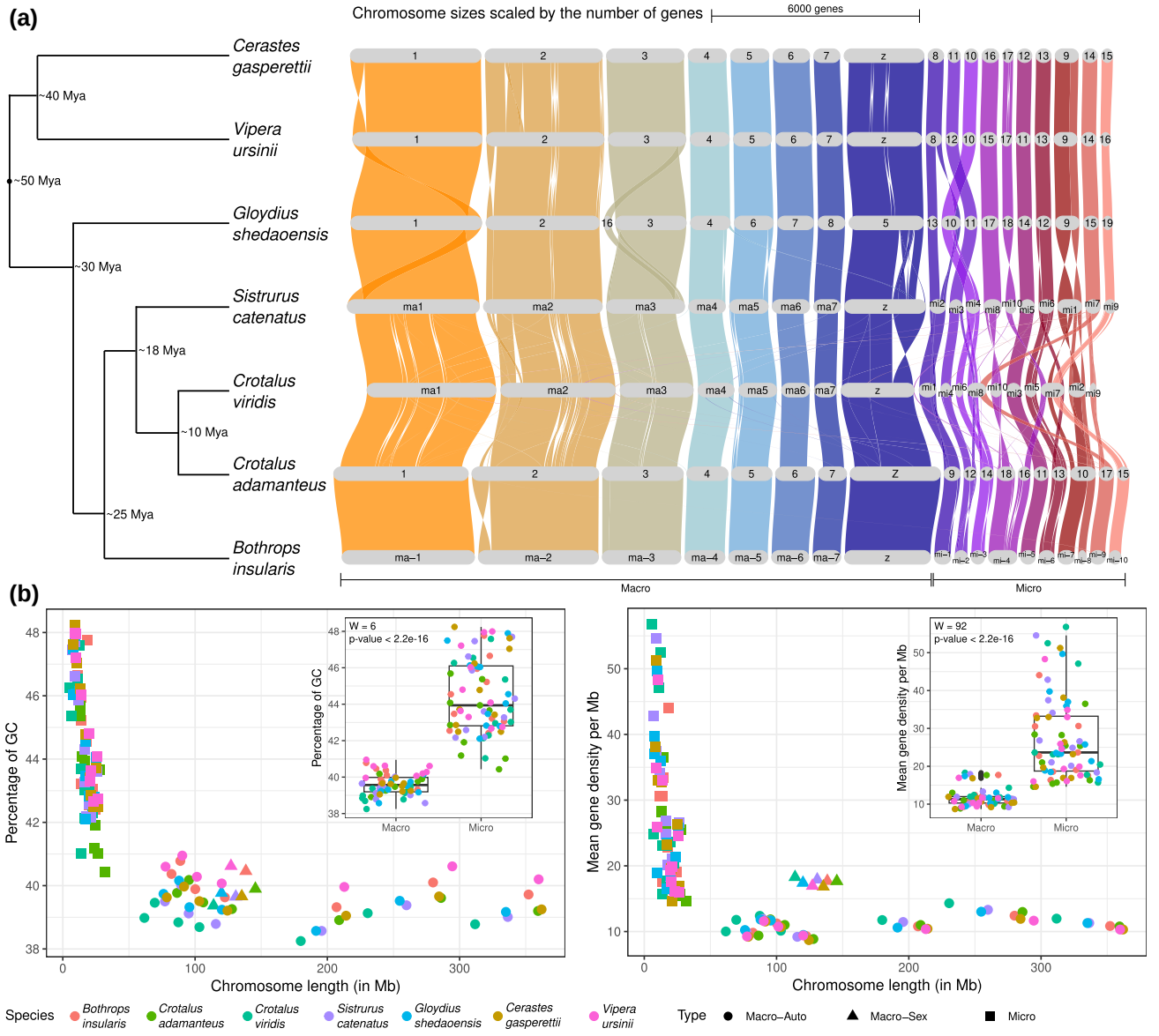

Fig. 2. Macrosynteny of *B. insularis* genome and other viper chromosome-level genomes. a) Riparian plot showing syntenic blocks and rearrangements among the species analyzed. Chromosomes are color-coded by chromosomes using the *B. insularis* genome as a reference. Chromosome sizes are scaled by the number of genes. A schematic representation of the phylogenetic tree with an estimated time of divergence between species based on previous work (Alencar et al. 2016) is shown at the left. b) Percentage of GC content versus chromosome size (at the left) showing statistically significant differences between macro- and microchromosomes (Wilcoxon test: W = 6, *p*-value<2.2e−16). Mean gene density versus chromosome size (at the right) shows statistically significant differences between macro- and microchromosomes (Wilcoxon test: W = 92, *p*-value<2.2e−16).

each species. We noticed that microchromosomes are GC-rich and gene-dense when compared to macrochromosomes (Fig. 2b). Interestingly, a previous report observed a similar pattern in other species containing microchromosomes, like in squamates, turtles, and birds (Waters et al. 2021). The microchromosomes present high levels of interchromosomal contact frequency when compared to macrochromosomes, which is a pattern broadly conserved among species containing microchromosomes (Schield et al. 2019; Perry et al. 2020). A recent report assembling the haplotype-resolved genome of a viper species revealed that microchromosomes are enriched for rearrangements (Nachtigall et al. 2025), whereas another report revealed a similar pattern in sea snakes (Ludington et al. 2023). It suggests a dynamic genome organization in microchromosomes of snakes. Our analysis, together with previous findings, indicate that microchromosomes in vipers are GC-rich and gene-dense, have low repeat content, high

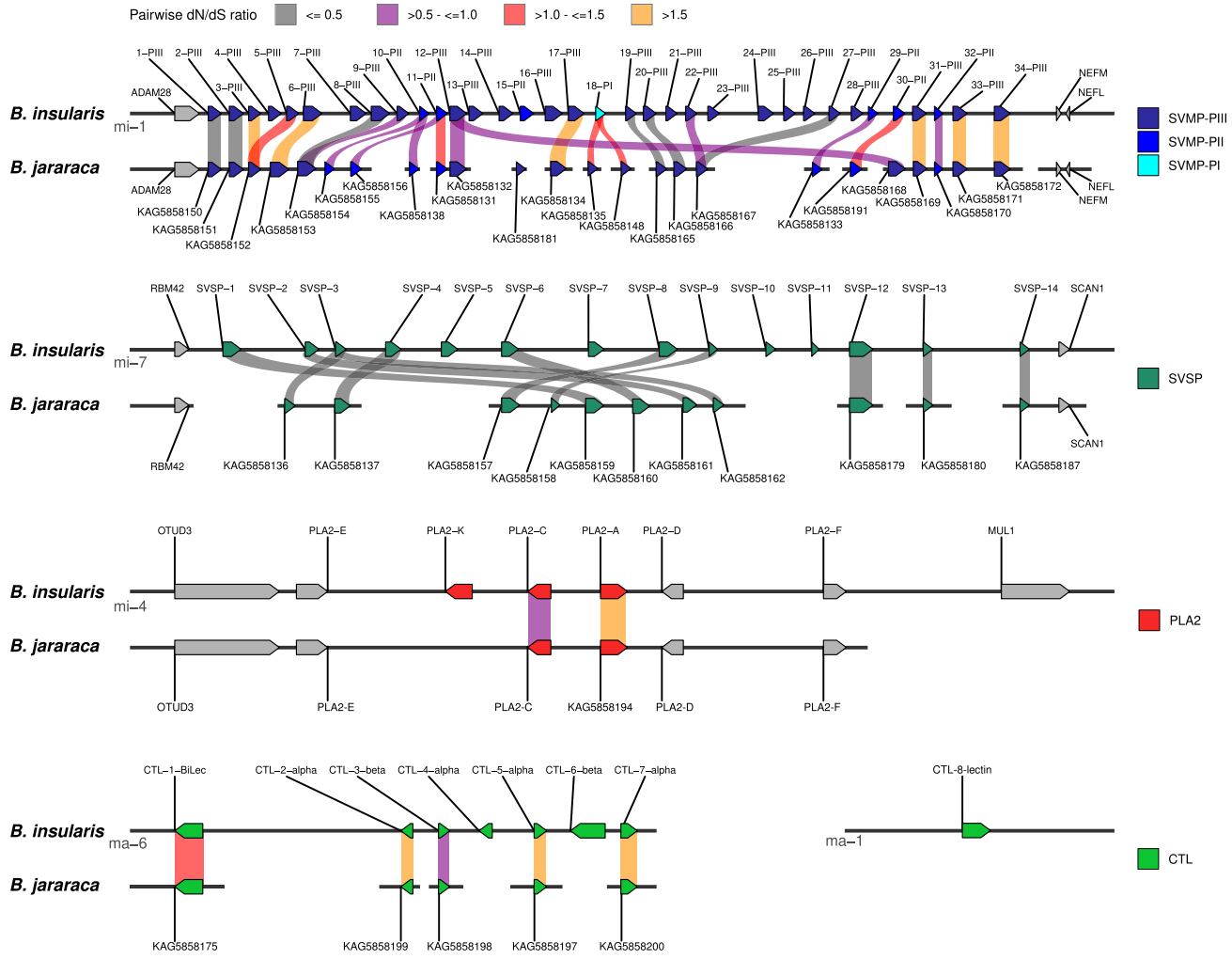

Fig. 3. Genomic comparison of multi-loci toxin families (i.e., SVMP, SVSP, PLA2, and CTL) between *B. insularis* and the mainland counterpart *B. jararaca*. The riparian plots show the relationships of toxin genes between species based on orthology inference, phylogenetic tree, and genomic alignments. The shaded areas highlight the orthologous toxin genes and are color-coded based on their pairwise dN/dS ratios. Arrows represent genes with toxins color-coded by their toxin family and nontoxin genes colored in light gray.

interchromosomal contact frequency, and high rates of recombination, which are widely detected in other microchromosome-containing species (Consortium ICGS 2004; Schield et al. 2019; Perry et al. 2020; Waters et al. 2021).

Genomic Landscape of Toxin Genes

The toxin annotation revealed 72 genes from 15 toxin families (Fig. 1d), which comprise major and minor components of viper venoms (Oliveira et al. 2022). Similar to what was previously observed in *B. jararaca* (Almeida et al. 2021) and *B. alternatus* (Nachtigall et al. 2024), most toxin families were represented by a single locus (i.e., VEGF-F, BPP, PLB, NUC, PDE, CRISP, Vespryn, HYAL, and NGF), whereas two toxin families were detected with two paralog copies (i.e., KUN and LAAO) and the other toxin families were organized as tandem arrays with variable number of paralog copies (i.e., SVMP, PLA2, SVSP, and CTL). We detected homologous to the 72 toxin genes in the ToxProt database and identified 60 toxin genes matching the venom gland expression sequence tag (EST) data of *B. insularis* available (Supplementary Table S1). The toxin genes in the *B. insularis* genome not identified in the EST dataset consisted of toxins lowly expressed in the venom gland transcriptome, which may reflect the low throughput of the EST approach. Next, we performed a microsynteny analysis of the multi-loci toxin arrays (i.e., SVMP, SVSP, PLA2, and CTL) comparing *B. insularis* with the mainland *B. jararaca* (Fig. 3).

The SVMP loci in *B. insularis* are composed of 34 SVMPs flanked by the nontoxin genes ADAM28 and NEFM, as previously observed in vipers and elapids (Giorgianni et al.

2020; Almeida et al. 2021). Of the total 34 SVMPs, 27 represented a PIII class, 6 represented PII class, and one represented a PI class and it is in accordance with a previous transcriptomic study using EST approach (Junqueira-de Azevedo and Ho 2002). Despite this high number of SVMP genes, only two of them (SVMP-18-PI and SVMP-19-PIII) are considerably higher expressed than the others (i.e., SVMP-18-PI and SVMP-19-PIII represent 20–30% of expression of all toxin genes; Supplementary Table S2), a common pattern observed for the three individuals with venom gland transcriptomes sequenced. These two genes are located next to each other in the central region of the SVMP array. The most expressed toxin gene in *B. insularis* (SVMP-18-PI), is a PI type of SVMP, a kind of gene not reported in the sequenced genome of the mainland *B. jararaca*. However, there are previous reports of PI genes in other individuals of *B. jararaca* (Junqueira-de Azevedo et al. 2015), indicating that this PI may represent a population variant in the *B. jararaca* complex. The other highly expressed SVMP is the SVMP-19-PIII. This protein corresponds to the well characterized Bothropasin (Mandelbaum et al. 1982), the major component of *B. jararaca* venom and a potent hemorrhagic enzyme. In *B. cotiara*, this protein was shown to contain a hypotensive peptide (position 191–198) (Miyamoto et al. 2024) that is also present in the *B. insularis* sequence (Supplementary Fig. S3). Moreover, although this gene encodes a major classic PIII SVMP found in many *Bothrops* species, we identified an alternative allele resulting in a PI type transcript in the *B. insularis* genome, as discussed later.

Other SVMP genes were just medial to low expression in the venom glands of *B. insularis* (i.e., those SVMP genes represent 0.01–5% of expression of all toxin genes; Supplementary Table S2). This includes the gene (SVMP-13-PII) for the only *B. insularis* SVMP characterized so far, insularinase, a PII shown to be a prothrombin activator (de Albuquerque Modesto et al. 2005), whose disintegrin domain is a potent inhibitor of human platelet aggregation and fibrinogen-adhesion of endothelial cells (Della-Casa et al. 2011). The gene SVMP-4-PIII coding for HF3, which is a potent hemorrhagic SVMP (Menezes et al. 2008), was also mapped and shown to have several paralog copies ranging from 89% to 94% of similarity. The SVMP-17-PIII shares 92% identity to the Berythractivase (Q8UVG0), which is a non-hemorrhagic SVMP capable of activating prothrombin in the absence of calcium, first discovered in *B. erythromelas* (Silva et al. 2003). Most of the other lowly expressed SVMP genes were similar to various proteins of different Viperidae, especially uncharacterized SVMPs from *B. jararaca*. Indeed, the microsyntenic analysis revealed 22 SVMPs of *B. insularis* orthologous to SVMPs in the *B. jararaca* genome (Fig. 3; Supplementary Fig. S4) with few putative duplication and deletion events. However, those apparent gains and losses can't be confirmed using the available *B. jararaca* genome assembly and may instead reflect its fragmentation level. This scenario reveals a diverse functional set of SVMPs in the *B. insularis* genome, although only two of them seem to account for the majority of proteins in the venom.

The SVSP loci are composed of 14 SVSP genes arranged in tandem arrays. This pattern was also reported in *B. jararaca*, *B. alternatus*, and other vipers (Schield et al. 2019; Almeida et al. 2021; Margres et al. 2021; Hogan et al. 2024). The number of SVSP genes detected is compatible with that observed in the venom gland transcriptomic data of *B. insularis* (Junqueira-de Azevedo and Ho 2002). Most *B. insularis* SVSPs matched to known SVSPs exerting procoagulant activity (Serrano and Maroun 2005). The SVSP-6 gene is 95% similar to PABJ isoforms, well characterized venom serine proteases that induces platelet aggregation through activation of protease-activated platelet receptors (Serrano et al. 1995). The SVSP-12 gene codes for BjSP, a serine protease with the special property of inducing fibrinolysis without promoting plasma coagulation or platelet aggregation (Paes Leme et al. 2008; Carone et al. 2018). SVSP-8 shares 98% similarity with an uncharacterized SVSP cDNA obtained from the *B. jararacussu* venom gland transcriptome, a species distantly related within the genus. The other SVSPs showed generally low levels of conservation (80–90%) with other viperid species and exhibited low expression levels. The microsynteny comparison with *B. jararaca* revealed 10 SVSPs with clear orthology and some rearrangements in gene order (Fig. 3; Supplementary Fig. S5), but the absence of some SVSP genes in *B. jararaca* may just reflect the fragmentation of the available *B. jararaca* genome.

Three PLA2 genes were arranged in tandem array between other nontoxic PLA2 genes (i.e., PLA2-E, PLA2-D, and PLA2-F) and flanked by OTUD3 and MUL1 genes, which is a pattern broadly conserved across vipers (Dowell et al. 2016; Margres et al. 2021). The toxic PLA2s comprised one acidic D49 type (PLA2-A), one basic D49 (PLA2-C), and one basic K49 (PLA2-K). The PLA2-A is homologous to a calcium-dependent PLA2 described to induce myonecrosis in mice and chickens (Cogo et al. 2006). The PLA2-C exhibited low expression level and it is orthologous to the PLA2 that may have originated the other toxic PLA2 in vipers (Dowell et al. 2016), which generally acts on the inhibition of blood coagulation (Wang et al. 2005). The less expressed PLA2-K is homologous to the Bothropstoxin-1 (Q90249), which is a basic PLA2 K49 type with myotoxic activity (Homsi-Brandeburgo et al. 1988). The microsynteny comparison with *B. jararaca* revealed that PLA2-A and PLA2-C are conserved, whereas *B. jararaca* lost the PLA2-K. The absence of PLA2-K in *B. jararaca* may reflect intraspecific variation, once it was detected in other *B. jararaca* individuals (Junqueira-de Azevedo et al. 2015). Moreover, several lineage-specific losses were previously observed in

rattlesnakes (Dowell et al. 2016) and the PLA2 loci in lanceheads may be facing a similar evolutionary history as previously suggested (Nachtigall et al. 2022).

We identified eight CTL genes, from which seven were highly expressed. The highly expressed CTLs were located in tandem array in ma-6 consisting of one monomeric CTL (CTL-1-BiLec), which is homologous to the galactose-binding monomeric CTL with a hemagglutinating activity (Guimarães-Gomes et al. 2004), and six heterodimeric CTLs (CTL-2 to −7), which are homologous to four alpha and two beta subunits. Interestingly, two pairs of alpha and beta CTLs (i.e., CTL-2-alpha and CTL-3-beta, and CTL-6-beta, and CTL-7-alpha) were organized in pairs with an inverted pattern as previously described in *B. alternatus* (Nachtigall et al. 2024) and observed in other viperids (Margres et al. 2021; Hogan et al. 2024; Nachtigall et al. 2025). The lowly expressed CTL (CTL-8) was located as a single-locus in the ma-1, and it is homologous to monomeric galactose-binding lectins. Its isolated positioning in another chromosome and its lower expression level indicate that this CTL gene is likely a non-venom CTL that has been propagated as a possible toxin in venom gland transcriptome annotations. Comparison of CTLs with those of *B. jararaca*, revealed five homologous genes, whereas three had no detectable homologs (Fig. 3; Supplementary Fig. S6). However, the absence of those three CTLs in the *B. jararaca* genome may reflect genome fragmentation rather than a true deletion event.

Other relevant components of *Bothrops* venom include BPPs, CRISP, VEGF-F, LAAO, NGF and HYAL. The position of those genes follows what have been described in *B. jararaca* genome, with the additional confirmation that BPP gene is placed between DIS3L2 and COPS7B in ma-5, LAAO genes are placed between CKR1 and UQCRC1 at ma-2, NGF-1 is between TEMEM9 and TSPAN2 at ma-3 and HYAL gene is flanked by WASL and HYAL4 at ma-6.

For a better understanding of the evolution of major venom components of *B. insularis*, we measured the pairwise synonymous substitution rate (dS), nonsynonymous substitution rate (dN), and dN/dS ratio of the inferred orthologous toxin genes of *B. insularis* and *B. jararaca* (Fig. 3; Supplementary Table S3). Among SVMPs, 15 pairs showed a dN/dS ratio lower than one, and 11 pairs exhibited a dN/dS ratio higher than one, which reveals that specific SVMPs present a higher number of nonsynonymous substitutions. All SVSP pairs displayed a dN/dS ratio lower than 1 suggesting an absence of protein changes in this toxin family. PLA2-A and PLA2-C showed a dN/dS ratio higher and lower than 1, respectively. It reveals that the PLA2-A accumulated nonsynonymous substitutions affecting its final protein sequence. Most CTL pairs presented dN/dS ratio higher than 1, which suggests higher amino acid changes in these loci in response to the adaptation process. This analysis suggests that SVMP, PLA2, and CTL genes show higher nonsynonymous substitution rates when compared to the mainland species, possibly reflecting *B. insularis* adaptation to the prey available on the island.

Genetic Variation of Toxin Genes

To investigate population genomics in *B. insularis*, we sequenced and analyzed low-coverage whole genomes from seven additional individuals (total of eight individuals), assessing toxin gene variation and demographic history (discussed in the next section). We mapped genome resequencing data to the assembled *B. insularis*, with variant calling and filtering steps yielding 9,542,686 high-quality biallelic variants representing single nucleotide polymorphisms (SNPs). To assess the quality of SNPs and ensure it is not affecting downstream analyses, we measured the distribution of depth (DP), missing data, minimum allele frequency (MAF), and singletons of variants detected in the genome and within toxin genes (see Material and methods for further details). Then, we measured the number of SNPs, nucleotide diversity (π), and Tajima's D within the coding sequence (CDS), introns, and flaking regions (i.e., 500 bp upstream and downstream of the CDS to account for noncoding regions surrounding the gene). Additionally, we calculated the copy-number variation (CNV) of toxin genes by using the genomic read coverage obtained with the mapped data.

We evaluated nucleotide-level diversity across toxin families by comparing their diversity distribution. PLA2, CTL, and LAAO are the major components with higher nucleotide diversity, whereas PDE, PLB, KUN, and Vespryn are the minor components with higher nucleotide diversity (Fig. 4a). These genes also presented a positive Tajima's D values when compared with other toxin families, suggesting potential balancing selection. In contrast, SVMP and SVSP exhibited very low diversity and negative Tajima's D values, consistent with possible positive selection on these toxin families. We also evaluated whether the observed diversity affected CDSs, introns, and/or flanking regions (Fig. 4b). When analyzing the features separately, we found that only PLA2s and CTLs showed higher nucleotide diversity in CDSs compared with introns and flanking regions (Supplementary Fig. S7). As expected, we detected a positive correlation among the diversity indexes analyzed, with PLA2 and CTLs showing higher diversity than other toxin genes (Fig. 4c). Indeed, these genes showed positive Tajima's D values, confirming the signals of balancing selection, whereas SVMP and SVSP showed low values of nucleotide diversity and a negative Tajima's D, supporting the signals of positive selection in both toxin families.

To assess functional variation in toxin genes, we analyzed amino acid changes resulting from the nucleotide diversity observed in the previous analyses. This revealed that CTLs, PLA2, LAAO, PLB, PDE, KUN, and Vespryn show

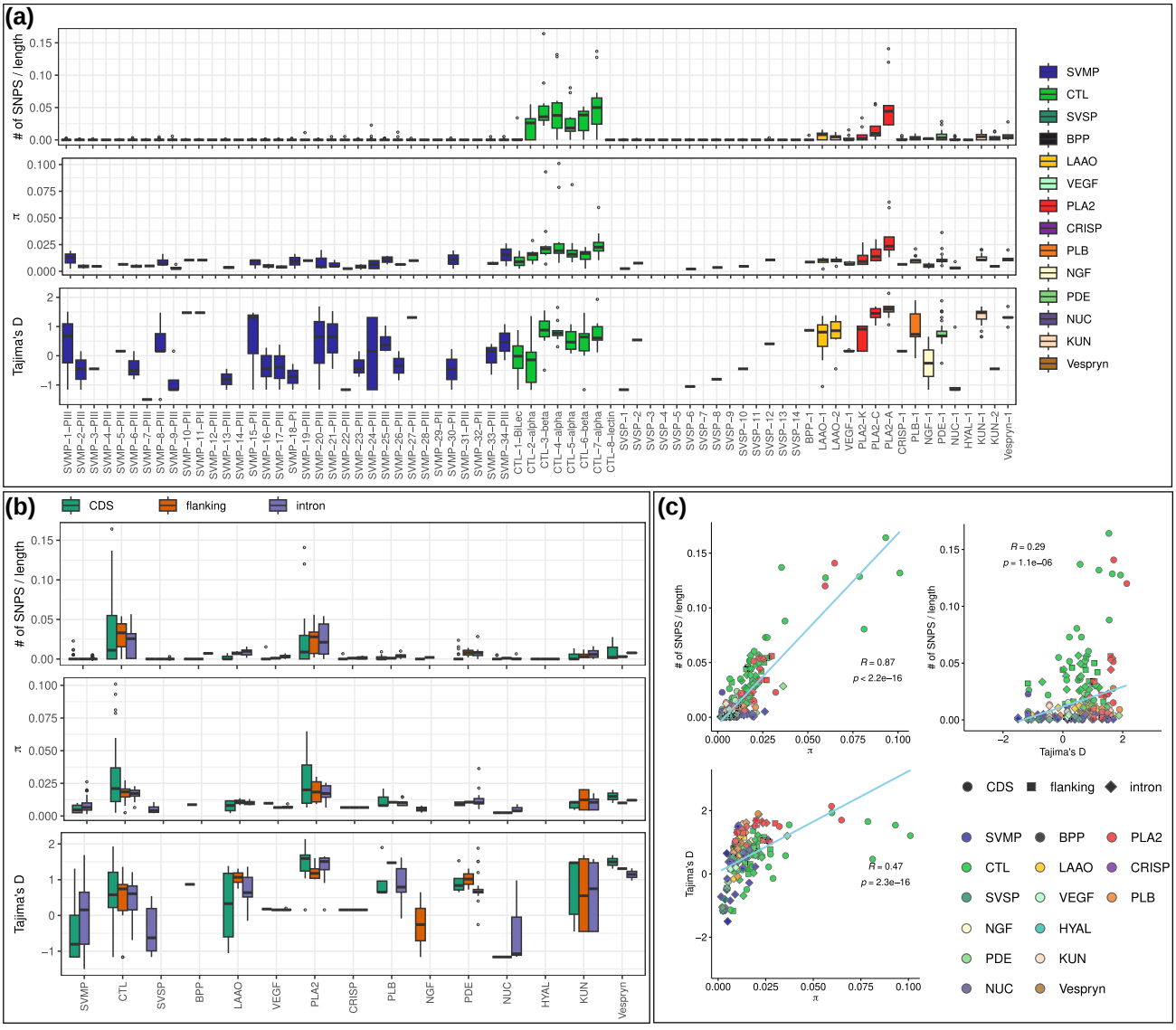

Fig. 4. Number of SNPs, nucleotide diversity (π), and Tajima's D in toxin genes. a) Distribution of measures in each toxin gene accounting for the CDS, introns, and flanking regions. b) Distribution of measures within CDS, intron, and flaking regions of each toxin family. c) Pairwise comparison of measures showing a positive correlation (R > 0 and *p*-value < 0.001).

mutations causing amino acid changes at the population level (Fig. 5). Most amino acid changes in PLA2 were nonsynonymous and affected the PLA2-A gene, which is the PLA2 with higher Tajima's D value among PLA2s, suggesting a stronger signal of positive selection on this toxin gene. Among CTLs, there are no amino acid changes detected in the monomeric forms (i.e., CTL-1 and CTL-8), whereas the heterodimeric CTLs showed higher levels of changes, predominantly nonsynonymous. Only one SVMP gene showed synonymous mutation (SVMP-23) and no amino acid-altering mutations were detected in SVSPs.

Notably, two SVMP genes (SVMP-19 and SVMP-21) contain nonsense mutations. These genes are located near the PI gene (i.e., SVMP-18), which is itself derived from a premature stop codon (Giorgianni et al. 2020). In the SVMP-19 of the reference genome individual (SB1851), we identified a SNP that truncates the CDS (Fig. 6a). This truncated variant occurs in around 50% of the reads mapping to this gene (Fig. 6b). It is an A>T substitution in exon 13 that introduces a stop codon replacing a Lys at position 435 of the precursor protein, terminating translation at the beginning of the disintegrin-like domain and producing a PI-like protein instead of the classic PIII bothropasin (Fig. 6c). The sequence downstream of the stop codon, which originally encoded the ancillary domains, now corresponds to the 3'UTR of a PI without any additional

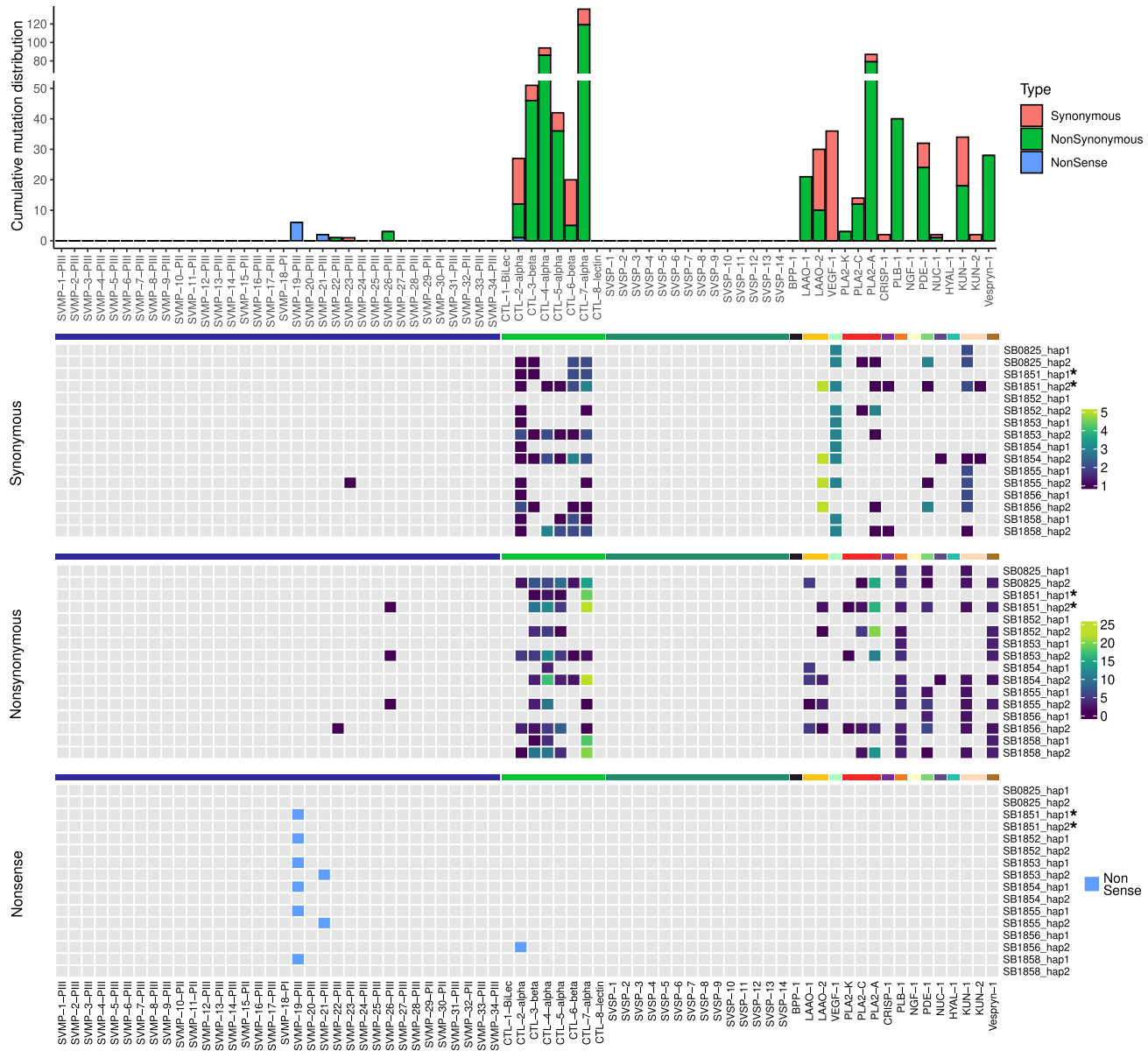

Fig. 5. Mutation type distribution in toxin genes. At the top, the barplot shows the cumulative mutation distribution among all samples analyzed. Different types of mutations are shown in different colors. The heatmap charts show the number of each mutation type in both haplotypes of each sample analyzed. The substitutions are shown from top to bottom as synonymous, nonsynonymous, and nonsense. The individual used to assemble the reference genome is marked with an asterisk. The toxin families are color-coded as in Fig. 1.

mismatches. We further confirmed that six other individuals carry this same SNP, all as heterozygotes with alleles encoding both PIII and PI-like forms, whereas two individuals are homozygotes containing only the PIII form (Fig. 6b). Transcriptome analysis of three of these individuals showed that the heterozygotes (SB1851 and SB1852) express a very low proportion of the PI-like form (1.2% in SB1851 and 0.8% in SB1852), whereas, as expected, the homozygous individual (SB0825) shows no evidence of the PI-like form. By contrast, the SVMP-21 presented a SNP resulting into a premature stop codon in only two individuals. It is a G>T substitution in exon 6 that replaces a Glu at position 191 of the precursor protein with a stop codon. This variant produces a truncated SVMP lacking active enzymatic domains, containing only the signal peptide and propeptide domains. Its low prevalence in the population may indicate a deleterious mutation maintained by genetic drift.

We used the mapped read coverage in the reference genome to estimate CNV of toxin genes (Fig. 7). This analysis revealed that most toxin genes present a similar copy

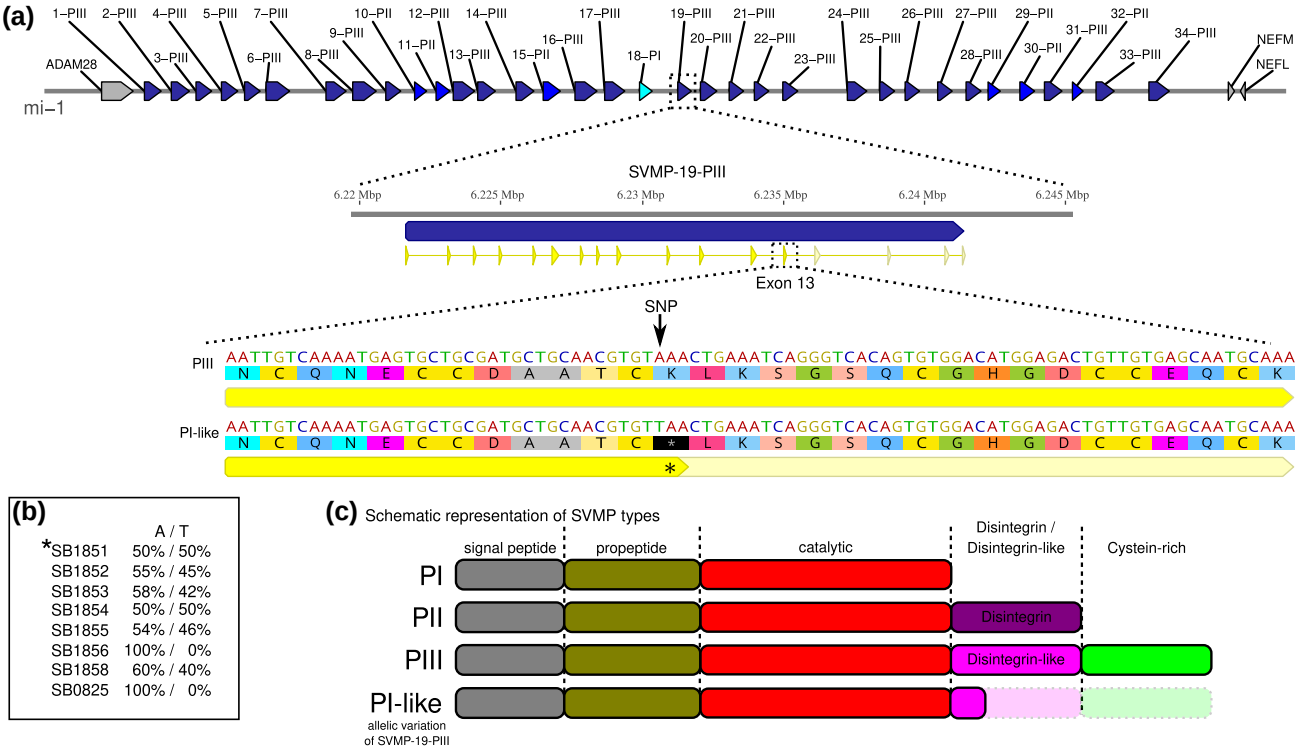

Fig. 6. The SVMP-19 gene structure and position of the allelic variation resulting in a PI-like protein. a) SVMP loci in mi-1 with zoom in the SVMP-19. At the bottom, a zoom into the exon 13, which contains the allelic variation detected. b) Percentage of each nucleotide in reads mapping into the allelic variation genomic position. The individual used for genome assembly is marked with an asterisk. c) Schematic representation of SVMP structure with the PI-like identified in the present study represented at the bottom.

number (i.e., values around one) to the reference genome, which suggests that those toxin genes are homozygotes similar to the reference genome (HomoRef). All toxin genes showing values different than one belong to the multicopy toxin families (i.e., SVMP, SVSP, CTL, and PLA2). Specifically, we detected six SVMP genes (i.e., SVMP-21, –22, –23, –24, –25, and –26) and six CTL genes (i.e., CTL-2, –3, –4, –5, –6, and –7) presenting double coverage (HomoDup), suggesting these genes may be duplicated in one of the haplotypes, whereas one SVSP gene (i.e., SVSP-6) and one PLA2 gene (i.e., PLA2-K) presenting half coverage, indicating a putative deletion event in one haplotype. The SVMP duplication of six genes in tandem was detected in three individuals (Supplementary Fig. S8), which indicate that heterozygous or homozygous state is spread in the population. The double coverage of six CTL genes was detected in four individuals (Supplementary Fig. S9), which may represent heterozygous states. The deletion of one SVSP gene was detected in seven individuals, which may represent a heterozygous state, whereas one individual may represent a homozygous state (Supplementary Fig. S10). The PLA2 deletion was detected in four individuals and likely represents a heterozygous state (Supplementary Fig. S11), in accordance with varying presence of this gene in *B. jararaca* population (Junqueira-de Azevedo et al. 2015). The CNVs detected in the present study likely represent confident duplication and deletion events between haplotypes, but further experimental validation is needed to confirm the haplotypes of each individual and also if those CNVs are affecting the final venom phenotype.

Demographic Inference of the Golden Lancehead

To investigate the demographic history of the golden lancehead over the last 100 kyr and assess potential effects of climatic changes and the insularization process, we inferred the effective population size (Ne) using the SMC++ method (Terhorst et al. 2017), which is indicated to estimate Ne between 100 and 100,000 years ago (Fig. 8). We used a generation time of 2 years and a mutation rate of 1.25 × 10e−8 per generation following previous studies using *Bothrops* and other snake species (Grazziotin et al. 2006; Pasquesi et al. 2018; Ludington and Sanders 2021). Our results indicate an ancestral Ne of approximately 140,000 individuals around 100 kya (Fig. 8) and a demographic history showing two periods of abrupt population decline, one around 50 kya (Ne ~30,000) and another around 11 kya

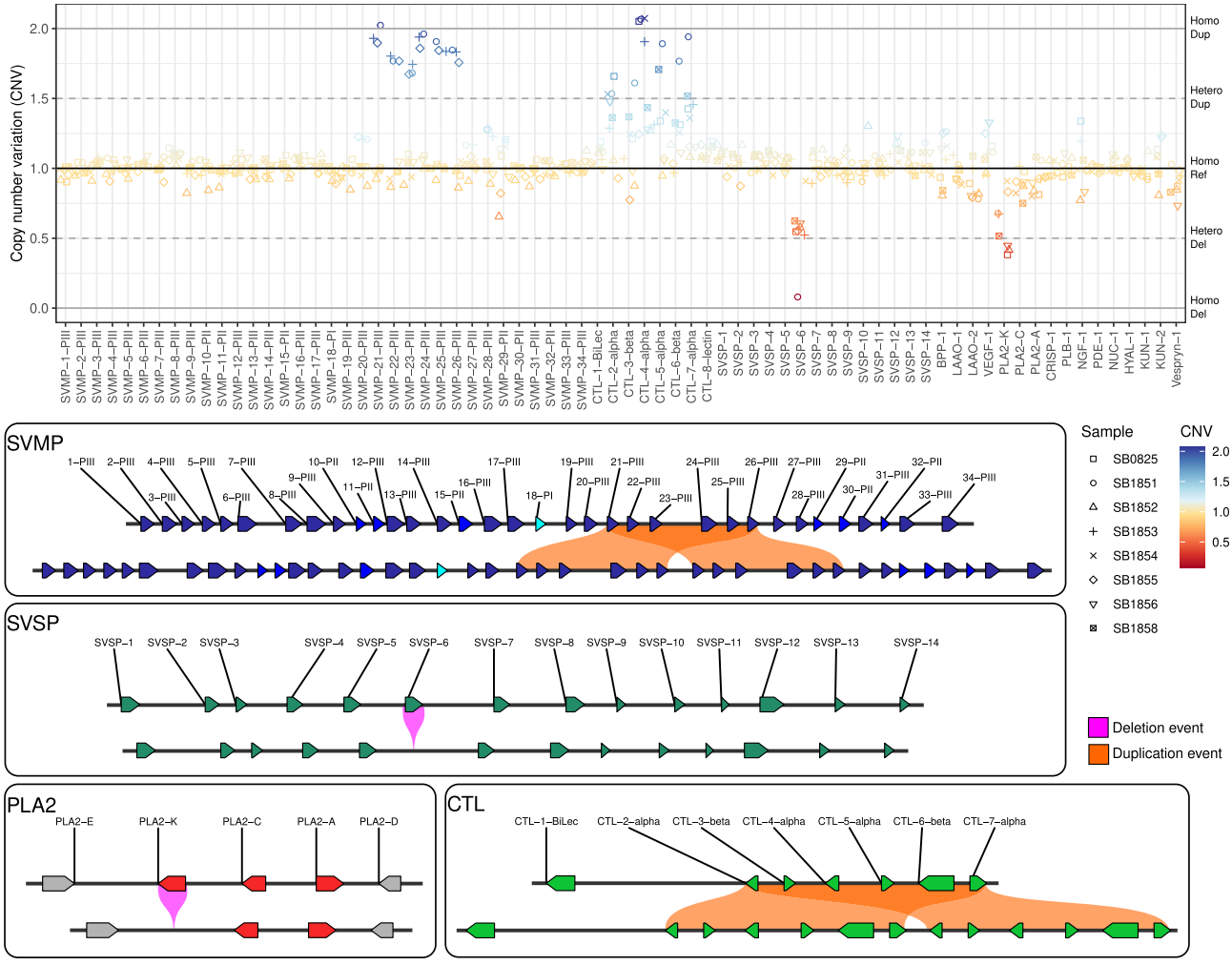

Fig. 7. Copy number variation (CNV) of toxin genes in *B. insularis*. At the top, the average coverage of each toxin gene compared to the average coverage for the chromosome it is located at for each sample analyzed. These values indicate duplications (CNV values higher than 1) and/or deletions (CNV values lower than 1) for specific toxin genes. HomoDup, homozygous for duplication (CNV around 2); HeteroDup, heterozygous for duplication (CNV around 1.5); HomoRef, homozygous for the reference (CNV around 1); HeteroDel, heterozygous for the deletion (CNV around 0.5); HomoDel, Homozygous for the deletion (CNV around 0). At the bottom, a schematic representation of haplotypes with the duplication and deletion events of the toxin genes detected in the analysis. Toxins are color coded following Fig. 3.

(Ne ~10,000). A gradual decline was inferred from 10 kya to 5 kya (Ne decreasing from ~10,000 to ~1,000). Around 2 kya, the species underwent a population growth (Ne ~15,000), followed by a subsequent decline, reaching a Ne of ~5,000 in the last millennium. This final estimate is comparable to the current population size of 2,000–4,000 individuals obtained from mark-recapture methods (Guimarães et al. 2014), indicating concordance between genetic Ne estimates and field-based census size (Nc) for *B. insularis*.

Discussion

We present a high-quality reference genome for the golden lancehead, providing an essential genomic resource to infer its evolutionary history, assess patterns of genetic diversity and guide science-based conservation strategies for this critically endangered species. In addition, our intra- and interspecific comparisons clarify macrosyntenic relationships among viper genomes, elucidate the composition and evolutionary basis of the genes underlying the distinct venom phenotypes and offer a more robust preliminary reconstruction of the demographic history of *B. insularis*. Our findings also provide compelling evidence of selection in toxin genes and contribute with new perspectives for future studies aimed at understanding the mechanisms driving phenotypic differentiation and adaptation in insular species.

Macrosyntenic Relationships in Viper Genomes

With the recent availability of more Viperidae genomes, the conservation of their key genetic elements remains to be

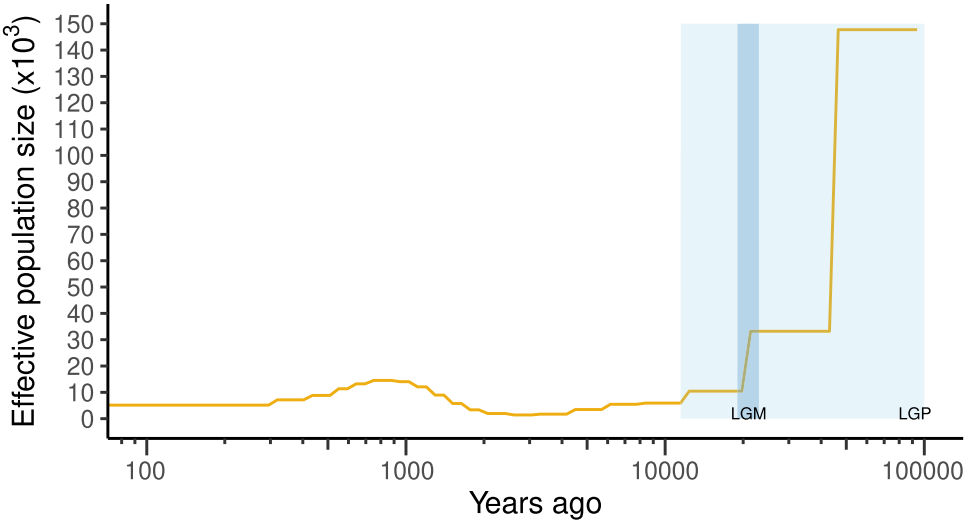

Fig. 8. Demographic history of the golden lancehead. The effective population size of *B. insularis* as estimated with SMC++ is shown on the *y*-axis and time in years ago on the *x*-axis. Light blue and blue regions show the timing of the Last Glacial Period (LGP) and the Last Glacial Maximum (LGM), respectively.

investigated. Despite the presence of syntenic blocks, a few rearrangements, like translocation and inversions, were observed within similar macrochromosomes, whereas microchromosomes presented a higher rate of rearrangements, including fission and fusions among distinct microchromosomes. Our data, together with previous analysis, also revealed that microchromosomes in vipers are GC-rich and gene-dense, have low repeat content, high interchromosomal contact frequency, and high recombination rates (Perry et al. 2020; Waters et al. 2021; Peng et al. 2023). Altogether, these findings suggest microchromosomes are "hot spots" for interchromosomal rearrangements while macrochromosomes are more conserved in vipers, which is a pattern also observed in other snakes and species containing macro- and microchromosomes (Waters et al. 2021; Ludington et al. 2023).

Although the comparative analysis revealed chromosomal rearrangements mainly affecting microchromosomes in accordance with previous studies, we emphasize the need to assess the quality of genomes used in such analysis (Rhie et al. 2021). Here, the assembled genomes used for comparison have high-quality statistics (Supplementary Table S4), with only *C. viridis* showing a gap percentage greater than 1%. Therefore, aside from the results related to this species, the rearrangements observed here represent reliable chromosomal structural variants among vipers. However, further experiments using cytogenetics techniques are required to confirm our results.

Venom Composition and Evolution

Venom is a key feature in the Viperidae biology, and our results offer an opportunity to improve our understanding of venom evolution in *B. insularis* and within the *B. jararaca* complex. By analyzing the toxin repertoire of the golden lancehead, we identified homologous venom components previously characterized functionally in this species or in the mainland *B. jararaca* (Junqueira-de Azevedo and Ho 2002; Valente et al. 2009; Junqueira-de Azevedo et al. 2015; Almeida et al. 2021). Our analysis indicates that the venom has a potent hemorrhagic and anticoagulant effect, primarily caused by the high abundance of SVMPs, SVSPs, and CTLs (Junqueira-de Azevedo and Ho 2002; Valente et al. 2009). This venom toxicity may be exacerbated by the presence of PLA2s with myotoxic effects. The diverse toxin repertoire may reflect adaptation to the available prey in the island, with specialization on migrant birds potentially providing an evolutionary advantage (Martins et al. 2002). Indeed, some studies have suggested that the venom of *B. insularis* is highly toxic to birds compared with *B. jararaca* venom (e.g., Zelanis et al. 2008). However, this hypothesis needs to be tested in a population-level framework for the highly variable venom of the mainland species.

Despite the notable repertoire of coagulant toxins, we also identified anticoagulant SVSPs, which have strong potential for development of anticoagulant drugs (Serrano and Maroun 2005). The presence of the VEGF-F, which induces vascular permeability and hypotension, may contribute to the venom dispersion within the prey (Junqueira-de Azevedo et al. 2001). This venom component also induces drastic hypotensive effects, indicating its potential for therapeutic development. We detected the presence of the highly expressed BPP in *B. insularis*, a venom component that is cleaved into peptides with hypotensive effects that have been used in medicine for decades (Raudonat and Rocha e Silva 1962; Ferreira et al. 1970). Additionally,

a recent report revealed the potential of a SVMP from *B. cotiara* to be cleaved into a peptide named Bc-7a with hypotensive effects (Miyamoto et al. 2024). This SVMP (named SVMPIII-8 in Nachtigall et al. 2022) is highly similar to the SVMP-19 from *B. insularis* and to the bothropasin from *B. jararaca* (O93523) and presents the conserved amino acids of the peptide Bc-7a (Supplementary Fig. S3). This indicates that the potential of SVMP genes to generate peptides with medicinal effects may be broadly conserved in lanceheads and other vipers, and should be further explored by expanding species sampling in analyses integrating venom peptidomics with functional experiments.

From an evolutionary perspective, we found that most toxin genes are under balancing or positive selection (i.e., Tajima's D values are higher or lower than zero), whereas only two toxin families (VEGF-F and CRISP) present Tajima's D values close to zero. This result suggests that natural selection processes, rather than genetic drift, is driving the venom variation in *B. insularis*. Among the major venom components, PLA2 and CTL show high nucleotide diversity with signatures of balancing selection, whereas SVMP and SVSP exhibit low nucleotide diversity with signatures of positive selection. The balancing selection observed in PLA2 and CTL gene in *B. insularis* resembles the selective regime observed in SVMP, SVSP, and PLA2 genes in other vipers, including lanceheads (Nachtigall et al. 2022) and rattlesnakes (Rautsaw et al. 2019; Schield et al. 2022). In contrast, the positive selection observed in SVMP and SVSP genes in *B. insularis* shows a similar selection pattern previously observed in a few toxin genes in other vipers, including PLA2-B1 in *Crotalus viridis* (Schield et al. 2022), CTLs in *C. oreganus* (Schield et al. 2022), SVMP and CTL in *C. cerastes* (Rautsaw et al. 2019), and one SVMP in some species of *Cerrophidion* (Rosales-García et al. 2023).

The signatures of selection observed in our population analyses of major components of *B. insularis* venom may be associated with the adaptation to dietary specialization, although our sampling design lacks a comprehensive coverage of *B. jararaca* mainland diversity. Toxin paralogs may evolve through selection-driven functional diversification in response to the molecular targets present in their prey (Bernardoni et al. 2014; Modahl et al. 2018; Holding et al. 2021). Such processes may also affect the relative abundance of specific toxin paralogs, enhancing venom efficiency against distinct prey (Margres et al. 2017; Mason et al. 2022). Some studies suggested that *B. insularis* venom is more toxic to chickens than that of the mainland *B. jararaca*, which rarely preys upon birds (Cogo 1991; Cogo et al. 1993). However, as in our study, many of these pivotal toxicity analyses employed sampling designs that hinder the evaluation of the highly divergent venom composition among different populations of the mainland *B. jararaca* (Gonçalves-Machado et al. 2016).

The specificity of the *B. insularis* venom against birds is also challenged by the similarity of venom genes and expression profiles to those of *B. jararaca*, particularly regarding the predominance of the same toxin families (i.e., SVMP, SVSP, CTL, and PLA2). However, our analyses indicate that all these major components of *B. insularis* venom evolved under selection, possibly in response to adaptative pressures associated with the insularization process. Further comparative genomic analyses, coupled with functional assays using *B. insularis*, population-level sampling of mainland *B. jararaca*, and other closely related species, are necessary to better understand this pattern and test this hypothesis.

The major components of *B. insularis* venom consist of multi-loci toxin families that also present CNV among the individuals analyzed. Specifically, we detected a large duplication of six genes in tandem in each SVMP and CTL loci and a minor deletion of one gene in each SVSP and PLA2 loci. Indeed, CNVs are a documented source of genetic variation of venom components in viper species (Margres et al. 2017, 2019; Giorgianni et al. 2020; Margres et al. 2021; Nachtigall et al. 2025). The CNVs can be linked to the abundance of the specific proteins facing duplication or deletion events in the final venom phenotype, which may confer a fitness advantage. The CNVs detected in *B. insularis* can be under diversifying selection based on the assumptions that these CNVs were detected at homozygous and heterozygous states among individuals. Such duplication events may be favoring an increase in expression level, whereas the deletions may be favoring a decrease in expression level. Indeed, CNVs affecting venom phenotype were previously detected in rattlesnakes (Margres et al. 2017; Nachtigall et al. 2025) and sea anemones (Smith et al. 2023). Further analysis focusing on generating a genotype-phenotype map for *B. insularis* and comparing it to the mainland *B. jararaca* may help to confirm if such detected CNVs are a source of genetic variation of the *B. insularis* venom.

The nucleotide diversity and CNV observed among multi-loci toxin families suggest that distinct toxin families present different selective pressures and gene-family evolution in the golden lancehead. The positive selection and duplications observed in SVMPs suggest the permanent heterozygote model, which relies on the fixation and maintenance of duplicated genes in heterozygous states to decrease the reduced fitness that can be caused by homozygotes (Innan and Kondrashov 2010).

A particularly interesting case is the SVMP-19 gene, which harbors a majorly expressed allele coding for SVMP PIII (bothropasin), and a minorly expressed allele coding for an unknown SVMP PI-like protein. Bothropasin, sometimes named jararhagin, is perhaps the most important toxin of *B. jararaca* venom, a potent and well characterized hemorrhagic factor (Mandelbaum et al. 1982).

Several studies were devoted to reconstructing the evolutionary relationship between PIII, PII and PI SVMPs in Viperidae, with the general understanding that PII arose through exon losses from PIII gene in ancestral Viperidae (Giorgianni et al. 2020; Almeida et al. 2021). PI origin is more complex, with evidence for multiple independent origins from PII genes (Casewell 2012), indicating that different snakes followed a convergent pathway towards SVMP simplification. The PI-like allele found here represents a novel case in Viperidae, since the whole sequence of the coding sequence is identical to the PIII gene with the exception of a SNP inserting a premature stop codon that impairs the translation of disintegrin and the Cys-rich domain of the PIII protein. This direct transition from PIII to PI-like was also observed in a particular genus of another family of snakes (*Phalotris spp*, Dipsadidae), in that case through a gene rearrangement rather than through a SNP, and the PI-like was shown to be a major protein of its venom (Campos et al. 2016; Entiauspe-Neto et al. 2024). For *B. insularis*, this allele is present in six of the eight individuals sequenced. However, the bothropasin version appears to be dominant in these individuals, since its mRNA has a 98/2 ratio with respect to the PI-like mRNA. The functional role of this PI-like protein, if it makes up to the venom, remains unknown; however, the adjacent highly expressed gene (SVMP-18) is also a PI type. Moreover, PI expression has been associated with the ontogenetic phenotypic transition of venom in *B. jararaca* (Zelanis et al. 2008, 2010), suggesting that this PI-like allele may yield different expression levels in newborn individuals.

The SVSP genes show signatures of positive selection and a gene gain/loss event, which suggests the birth-and-death model, a process that involves gene gains and losses following rounds of duplications (Innan and Kondrashov 2010). PLA2 has signatures of balancing selection and gain/loss of a gene, which also reflects the birth-and-death model (Innan and Kondrashov 2010). The balancing selection and duplications of CTLs suggest the multi-allelic diversifying model, which is similar to the permanent heterozygote model, but it also favors genetic variability (Innan and Kondrashov 2010). The classic birth-and-death model of gene-family evolution is widely accepted to explain multi-loci toxin family evolution and diversification (Fry et al. 2008; Casewell et al. 2011; Sachkova et al. 2019), but other studies have revealed distinct evolution models, such as the positive dosage model (Margres et al. 2017) and the concerted evolution model (Moran et al. 2008). It was shown that these multi-loci toxin families present distinct patterns in vipers (Holding et al. 2021; Mason et al. 2022) and that there is no single evolutionary model that can explain a general toxin family evolution and diversification (Jackson and Koludarov 2020).

Demography

Although our primary goal was to characterize the genomic diversity and the genetic basis of venom diversification in *B. insularis*, our analysis of highly resolved genome combined with additional low-coverage whole-genome data provides the most robust inference to date of the historical population dynamics of the species. The inferred demographic history suggests that the golden lancehead was likely isolated before the last sea-level rise, supporting previous hypotheses (Barbo et al. 2022), in which the process of isolation along the Brazilian coast resulted from a dynamic scenario involving three insularization events in the last 100,000 thousand years ago (kya). Our results indicate that the first population decline coincided with the earliest insularization of Queimada Grande Island, approximately 50–60 kya, and caused a pronounced population bottleneck that left a strong signature in the genomic diversity of the species. A second population decline was inferred during the second-earliest insularization period, approximately 10–15 kya. After that, although showing some demographic fluctuations around 1 kya, our results suggest that over the last 10 kyr the population remained relatively stable, with an Ne of 5,000 individuals, similar to the census size (Nc) estimated for the species (Guimarães et al. 2014). The similarity between Ne and Nc suggests the absence of severe fluctuations or bottlenecks over many generations and may indicate that, despite being classified as critically endangered, *B. insularis* is both genetically and demographically healthy. Based on the results provided here, conservation strategies can be developed to maintain this favorable relationship between Ne and Nc, including the planning of ex-situ conservation actions. Specifically, our genomic dataset will facilitate genomic surveillance by providing a comprehensive SNP database to monitor genetic diversity over time and assess potential risks from stochastic threats associated with deterministic threats like illegal trade and climatic change.

Finally, the framework presented, by combining chromosome-level, well-annotated genome with low-coverage genomic data from additional individuals, may serve as a model for future comparative analysis addressing open evolutionary questions in the *Bothrops jararaca* complex and other insular systems.

Material and Methods

Sampling

Eight specimens of *Bothrops insularis* were collected between 2019 and 2020 in Queimada Grande Island, SP, Brazil (Supplementary Fig. S12). Blood was extracted from the caudal vein, transferred to a tube containing 100% ethanol solution, and stored at −80°C until use. High-molecular-weight (HMW) genomic DNA (gDNA) was

extracted by using a pipette-free protocol as previously described (Margres et al. 2021). The specimens were milked and euthanized 4 days post-milking using a single-step sodium pentobarbital (100 mg/kg) injection following standard approved American Veterinary Medical Association (AVMA) guidelines. Venom glands and other tissues were sampled and transferred to RNAlater (Ambion) and stored at −80°C until use. The snakes were handled and collected under Protocol Number 4479020217 from the Ethics Committee on Animal Use of the Butantan Institute (CEUAIB).

DNA Sequencing

The gDNA from a male individual (SB1851) was used to construct the PacBio HiFi sequencing libraries with the SMRTbell Express Template Prep Kit 2.0 following the manufacturer's protocol. Sequencing was performed with two cells on the PacBio Sequel II system in CCS mode at the University of Delaware Sequencing and Genotyping Center. The sequencing resulted in 3,183,975 reads (with an average read size of 12.7 Kb and a total of 40,631,104,454 bp). Reads containing adapters were removed using cutadapt (v4.1; Martin 2011), and putative human and bacterial contaminants were removed using kraken2 (v2.1.2; Wood et al. 2019) to remove human or bacterial contaminants.

The blood sample from the same individual (SB1851) was used to construct a HiC library, following the protocol for nucleated blood cells for the Arima High Coverage HiC Kit (Arima Genomics) to crosslink DNA and generate the proximally ligated DNA. The proximally ligated DNA was used to construct the final Hi-C library using the Arima High Coverage Hi-C Library Preparation Kit (Arima Genomics) following the manufacturer's instructions. The HiC library was sequenced using the NovaSeq 6000 platform (Illumina) with paired-end reads layout (2×150 bp) at the Florida State University College of Medicine Translational Science Laboratory. Sequencing yielded 466,781,110 paired-end reads. Adapters were trimmed and reads with Phred scores lower than 25 and shorter than 75 bp were removed using Trim_Galore! (v0.4.4; https://github.com/FelixKrueger/TrimGalore), which returned a final dataset consisting of 466,781,110 HiC paired-end reads (total of 137,669,877,905 bp).

To estimate the genetic diversity in toxin genomic regions in *B. insularis*, we analyzed whole genome data from a total of eight individuals. Blood samples were used to extract DNA using a standard Phenol:Chloroform: Isoamyl protocol as previously described (Mathur and Gibbs 2025). Illumina short-read libraries were prepared using the TruSeq DNA PCR-free library prep kit following the manufacturer's instructions. The libraries were then sequenced on an Illumina HiSeq 6000 at the University of Illinois Roy J. Carver Biotechnology Center. Adapters were trimmed and reads with Phred scores lower than 25 and shorter than 75 bp were removed using Trim_Galore! (v0.4.4; https://github.com/FelixKrueger/TrimGalore). The final genomic datasets returned a mean 12.78 of depth (ranging from 10.05 to 22.62).

RNA Sequencing

Total RNA from the venom gland and other tissues (i.e., infralabial gland, harderian gland, brain, pancreas, spleen, kidney, skeletal muscle, heart, and ovary) were extracted using TRIzol Reagent (Invitrogen), following the manufacturer's protocol. The RNA concentration and contamination level were measured by ultraviolet absorbance using NanoDrop 1000 (Thermo Scientific), and RNA integrity was assessed with the equipment Agilent 2100 Bioanalyzer (Agilent Technologies). Next, the messenger RNA (mRNA) was purified from the total RNA by using the Dynabeads mRNA DIRECT kit (Ambion) and used to prepare independent cDNA libraries for each sample. The complementary DNA (cDNA) libraries were prepared following the protocol for TruSeq RNA Sample Preparation Kits v2 (Illumina), and sequenced using the HiSeq1500 platform (Illumina), generating strand-specific paired-end reads. Illumina adapters and low-quality reads were trimmed using Trim_Galore! (v0.4.4; https://github.com/FelixKrueger/TrimGalore), removing reads with Phred scores lower than 25 and a length shorter than 75 bp. The RNA-seq data were used for the genome annotation step and to estimate the expression level of toxins.

Genome Assembly

The HiFi long reads and paired-end HiC short reads were provided to hifiasm (v0.16.1; Cheng et al. 2021) to generate the primary assembly contig graphs with default parameters. The contigs of the primary assembly were used as reference to map the HiC reads using Chromap (v0.2.3; Zhang et al. 2021) and to scaffold using YaHS (v1.2a.2; Zhou et al. 2023). We then used PretextView (v0.2.5; https://github.com/sanger-tol/PretextView) to manually review the scaffolded genome following the "Rapid curation" manual (https://gitlab.com/wtsi-grit/rapid-curation). Then the unplaced contigs (i.e., contigs not integrated into scaffolds of chromosomes) were purged when being shorter than 20 Kb and representing spurious contigs or haplotig duplications as identified using purge_dups (v1.2.5; Guan et al. 2020). The genome assembly statistics were obtained using Inspector (v1.0.1; Chen et al. 2021), which also calculates a QV score to measure putative errors in the assembly, and the completeness was assessed using BUSCO (v5.2.2; Waterhouse et al. 2018) with the Tetrapoda gene set (odb10; a total of 5,310 genes). We confirmed the identity of chromosomes by performing BLAST searches against a set of chromosome-specific markers (NCBI accessions SAMN00177542 and SAMN00152474) of snakes

(Matsubara et al. 2006) and the chromosome-level assemblies of *Crotalus viridis* (Schield et al. 2019) and *C. adamanteus* (Hogan et al. 2024). We also confirmed the identity of the Z chromosome based on male–female read coverage ratio mapping whole-genome sequencing data of male and female individuals as previously described (Hogan et al. 2021). The mitochondrial genome was assembled using the long-read mode of MITGARD (v.1.1; Nachtigall et al. 2021a) with the HiFi reads and the *B. jararaca* mitogenome as reference (NC_030760.1). The assembled mitogenome was annotated using MitoZ (v3.6; Meng et al. 2019), followed by a manual inspection. The final genome assembly has been deposited at DDBJ/ENA/GenBank under the accession JBLSAJ000000000.

Genome Annotation

We annotated repetitive regions and transposable elements (TEs) using RepeatModeler2 and RepeatMasker as previously described (Nachtigall et al. 2025). We used the RepeatModeler2 (v2.0.1; Flynn et al. 2020) to generate a de novo species-specific repeat library (TE library). We split the library into "known" and "unknown" sets as output by RepeatModeler2. The "unknown" set was classified using DeepTE (v1.0; Yan et al. 2020) with the model designed for metazoans. To reduce false-positives, we removed any sequence classified as "NonTE" using TERL (v1.0; da Cruz et al. 2021). Then, the species-specific TE library (i.e., the "known" set and the "unknown" reclassified set) was merged into a curated TE library designed for snakes (Castoe et al. 2013), and the final TE library was used to perform the repetitive annotation using RepeatMasker (v4.1.1; https://www.repeatmasker.org/). The divergence between the individual TE copies versus their consensus sequences based on CpG-adjusted Kimura distance was estimated using built-in scripts from RepeatMasker. We also searched for telomeric sequences occurring at chromosome terminals using tidk-search (v0.2.0; Brown et al. 2025) with the conserved vertebrate telomeric repeat sequence TTAGGG as a query.

The soft-masked genome was used as input for gene annotation using GALBA (v1.0.11; Bruna et al. 2023), with default parameters, and the proteins annotated in the *Crotalus adamanteus* genome (Hogan et al. 2024) as the protein evidence. Then, we mapped the transcriptomic reads of several tissues against the *B. insularis* genome using STAR (v2.7.11; Dobin et al. 2013) and retrieved the counts using featureCounts (v2.0.6; Liao et al. 2014) to use as the transcript evidence and removed any gene with no counts in any tissue from the annotation set. We compared the transcriptome-filtered genes to the annotations available for *Mus musculus*, *Gallus gallus*, *Anolis carolinensis*, *Pogona vitticeps*, *Varanus komodoensis*, *Podarcis muralis*, *Notechis scutatus*, and *Pseudonaja textilis* in the ENSEMBL (release 112) using DIAMOND (v2.1.8; Buchfink et al. 2021). Finally, we assigned functional annotation through orthology and homology inferences using the gene sets available for *C. adamanteus* (Hogan et al. 2024), *Naja naja* (Suryamohan et al. 2020), *G. gallus* (GCA_016699485.1), and *M. musculus* (GCA_000001635.9) using Orthofinder (v2.5.4; Emms and Kelly 2019). To annotate toxins, we used ToxCodAn-Genome (v1.0; Nachtigall et al. 2024) with default parameters and followed their guide to ensure a confident toxin annotation set. Briefly, the venom-gland transcriptomic data was assembled and annotated using ToxCodAn (v1.0; Nachtigall et al. 2021b) with default parameters to generate a species-specific toxin database. The species-specific and the Viperidae toxin databases were used as database sources to annotate the toxins in the genome using ToxCodAn-Genome (v1.0; Nachtigall et al. 2024). We then generated a final annotation set by merging the toxin and nontoxin annotations. The final gene annotation set was used as input to estimate the toxin expression level using featureCounts and the expression data was normalized through the trimmed mean of M-values normalization method (TMM) using the edgeR package (v3.40.2; Robinson et al. 2010).

Additionally, we compared our toxin-annotated sequences to the ToxProt database (Jungo et al. 2012), which comprises a curated and reviewed set of toxin sequences, and to the expressed sequence tag (EST) of the venom gland of *B. insularis* (NCBI accessions from BM401391 to BM402067; Junqueira-de Azevedo and Ho 2002) using BLAST. We set the BLAST search to return the best hit ("-max_target_seqs 1") in each database to check for: (1) the presence of toxins previously identified through EST approaches; and (2) the homologous toxins in a reviewed set of toxin sequences.

Comparative Genomics

To better understand the genomic context of toxins annotated in the *B. insularis* genome, we compared them to the toxins annotated in the *B. jararaca* genome (Almeida et al. 2021). To assign orthology and homology among toxin genes, we first used Orthofinder to detect orthogroups. Then, we aligned the coding sequences (CDSs) of toxins within the same family using MAFFT (v7.453; Rozewicki et al. 2019) and inferred phylogenetic trees using IQ-TREE (v2.0.3; Nguyen et al. 2015) with the parameters "-m TEST -bb 1000 -alrt 1000". We also aligned the genomic regions through BLAST search to identify similar genomic regions longer than 5 Kb to help assign the toxins orthology. We used all of these results to properly infer the orthology and homology of toxin genes between both species. Additionally, we measured the pairwise synonymous substitution rates (dS), nonsynonymous substitution rates (dN), and dN/dS ratios for the orthologous toxin genes. We performed a codon-based alignment for each orthologous pair using PRANK (v.170427) and used it as input to

codeml from paml package (v4.9) to estimate the dS, dN, and dN/dS ratios.

We also performed a macrosyntenic analysis among viper species with a near chromosome-level genome assembly available in the literature. Specifically, we used the newly assembled *B. insularis* genome and the ones available for *Crotalus adamanteus* (Hogan et al. 2024; Nachtigall et al. 2025), *Crotalus viridis* (Schield et al. 2019), *Sistrurus catenatus* (Mathur and Gibbs 2025), *Gloydius shedaoensis* (Peng et al. 2023), *Cerastes gasperettii* (Mochales-Riaño et al. 2025), and *Vipera ursinii* (Halpern et al. 2025). To do this, we used GENESPACE (v1.2.3; Lovell et al. 2022) and inverted chromosomes using the built-in function "invertTheseChrs" when needed to improve the visualization of the final figure. For the sex chromosomes, we excluded the W from the analysis because only two species had it fully assembled (i.e., *C. adamanteus* and *C. gasperettii*). Additionally, we calculated the GC content and gene density in 1-Mb sliding windows for each chromosome and compared macro- and microchromosomes using the Wilcoxon Rank Sum test in R base.

Analysis of Genetic Variation in Toxin Genes

We used whole genome sequencing data from eight specimens of *B. insularis* to check for the genetic variation of toxin genes. Briefly, we called variants to check for SNPs in toxin genes and used the read coverage to check for copy number variation (CNV) of toxin genes. To identify SNPs, we mapped the reads using bwa (v0.7.17; Li and Durbin 2009) and removed read alignments with MAPQ <30 using samtools (v1.17; Danecek et al. 2021) and PCR duplicates using picard (v3.3.0; http://broadinstitute.github.io/picard/). Then, we called variants using BCFtools (v1.20; Danecek et al. 2021). We removed variants with quality lower than 20 and read depth lower than 7. Additionally, we removed variants with mean read depths above the 97.5th quantile to avoid spurious SNP calls based on paralogous mappings. Variants were phased using WhatsHap (v2.3; Patterson et al. 2015). Finally, we selected only biallelic SNPs to perform the genetic variation analysis of toxin genes.

To ensure we obtained high-quality SNPs that may not reflect artifacts, we analyzed depth used for variant calling (DP), missing data, minimum allele frequency (MAF), and singletons using VCFtools (v0.1.16; Danecek et al. 2011). First, we analyzed those metrics using the entire SNP set, which comprises the whole genome (Supplementary Figs. S13–S16). This quality control revealed that the entire SNP set have high quality for use in downstream analysis. When analyzing the SNP set comprising the genomic regions of the toxin genes, we observed patterns similar to those of the entire genome for all metrics (Supplementary Figs. S17–S19), suggesting that downstream analyses focusing on toxin genes are not affected by artifacts.

We estimated within-population nucleotide diversity (π), to measure the degree of polymorphism, and Tajima's D statistic (Tajima 1989), to test for deviations from neutral expectations. Both measures were calculated using VCFtools (v0.1.16; Danecek et al. 2011). We performed multiple runs to calculate statistics in 50-bp sliding windows. We compared the π and Tajima's D obtained in the coding sequence (CDS), intron, and flanking regions (i.e., 500 bp upstream and downstream of the CDS) to check if SNPs were mainly affecting CDSs. We also measured the heterozygous sites in each feature by calculating the number of SNPs divided by the size of the feature (i.e., the size of the CDS, intron, or flanking region). We estimated correlations between those measures (i.e., π, Tajima's D, and number of SNPs) using the Pearson correlation test in R. Finally, we analyzed variants with functional impact in CDSs of toxin genes by inspecting when variants resulted in synonymous, nonsynonymous or nonsense mutations.

To check for copy number variation (CNV) of toxin genes, we analyzed the read coverage of toxin genes and compared them to the average read coverage for the chromosome in which they are located. To do this, we used the mapped reads from the previous step to estimate an average read coverage for the gene (i.e., CDS and introns) and divide it against the average read coverage for the chromosome. The CNV values could suggest homozygous or heterozygous states for the reference and deletion or duplication events. Specifically, values around 1 suggest a homozygous state similar to the reference, values around 0 suggest a deletion event in the homozygous state, values around 0.5 suggest a deletion event in the heterozygous state, values around 1.5 suggest a duplication event in the heterozygous state, and values around 2 suggest a duplication event in the homozygous state.

Demographic Inference

We used SMC++ (v1.15.2; Terhorst et al. 2017) to estimate effective population size through time for the *B. insularis* population. We used the filtered and biallelic SNPs as described in the previous step. The VCF file was split per chromosome and converted to the SMC++ input file format using the "vcf2smc" command. We excluded the sex chromosome from this analysis to avoid bias in the demographic inference (Ludington and Sanders 2021). We used a mutation rate of 1.25 × 10e−8 per site per generation and a generation time of 2 years as previously described for *B. jararaca* and other snake species (Grazziotin et al. 2006; Ludington and Sanders 2021; Bergeron et al. 2023).

Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

Acknowledgments

The authors thank the anonymous reviewers for their valuable suggestions. Funding for this work was provided by Fundação de Amparo á Pesquisa do Estado de São Paulo (FAPESP; processes number 2013/07467-1, 2016/50127-5, 2018/26520-4, 2022/04988-0, 2022/12660-4) and the National Science Foundation (NSF DEB 1638902). I.L.M.J.A. and F.G.G. were also supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq; processes numbers 304532/2013-4 and 312016/2021-2).

Author Contributions

P.G.N., I.L.M.J.A., and F.G.G. conceived and designed the experiments. I.L.M.J.A., F.G.G., D.R.R., and H.L.G. contributed to the biological data and reagents. P.G.N. performed the experiments. PGN, ILMJA, FGG, and AJM analyzed the data. P.G.N., I.L.M.J.A., and F.G.G. wrote the manuscript. All authors read and approved the final manuscript.

Conflict of Interest

No competing interest is declared.

Data Availability

The genome assembly and the data generated in the present study are available in the NCBI under the project number PRJNA679826. Additionally, the assembled genome and annotations are available in the figshare database (https://figshare.com/projects/Bothrops_insularis_genome/237995). A list with all datasets used in the present study is available in Supplementary Table S5. The code to perform the analysis presented here is available at https://github.com/pedronachtigall/Binsularis_genome.

Ethical Approval and Collecting Permits

The specimens were collected in accordance with Brazilian laws, under the permit provided by Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio; SISBIO permit number 63300). The snake specimens were handled and processed under approval of the Ethic Committee on Animal Use of the Butantan Institute (CEUAIB; Protocol Number 4479020217).

Literature cited

Alencar LR et al. Diversification in vipers: phylogenetic relationships, time of divergence and shifts in speciation rates. Mol Phylogenet Evol. 2016:105:50–62. https://doi.org/10.1016/j.ympev.2016.07.029.

Almeida DD et al. The complete mitochondrial genome of Bothrops jararaca (reptilia, serpentes, viperidae). Mitochondrial DNA B Resour. 2016:1:907–908. https://doi.org/10.1080/23802359.2016.1149783.

Almeida DD et al. Tracking the recruitment and evolution of snake toxins using the evolutionary context provided by the *Bothrops jararaca* genome. Proc Natl Acad Sci U S A. 2021:118:e2015159118. https://doi.org/10.1073/pnas.2015159118.

Barbo FE et al. Another new and threatened species of lancehead genus bothrops (serpentes, viperidae) from Ilha dos Franceses, southeastern Brazil. Zootaxa. 2016:4097:511–529. https://doi.org/10.11646/zootaxa.4097.4.

Barbo FE et al. Speciation process on Brazilian continental islands, with the description of a new insular lancehead of the genus bothrops (serpentes, viperidae). Syst Biodivers. 2022:20:1–25. https://doi.org/10.1080/14772000.2021.2017059.

Barbo FE, Grazziotin FG, Sazima I, Martins M, Sawaya RJ. A new and threatened insular species of lancehead from Southeastern Brazil. Herpetologica. 2012:68:418–429. https://doi.org/10.1655/HERPETOLOGICA-D-12-00059.1.

Barbosa AM, do Amaral RO, Teixeira CF, Hyslop S, Cogo JC. Pharmacological characterization of mouse hind paw oedema induced by bothrops insularis (jararaca ilhoa) snake venom. Toxicon. 2003:42:515–523. https://doi.org/10.1016/S0041-0101(03)00230-7.

Beçak W, Beçak M, Nazareth H. Karyotypic studies of two species of South American snakes (boa constrictor Amarali and bothrops jararaca). Cytogenetics. 2008:1:305–313. https://doi.org/10.1159/000129740.

Bellairs A, Underwood G. The origin of snakes. Biol Rev. 1951:26:193–237. https://doi.org/10.1111/brv.1951.26.issue-2.

Bergeron LA et al. Evolution of the germline mutation rate across vertebrates. Nature. 2023:615:285–291. https://doi.org/10.1038/s41586-023-05752-y.

Bernardoni JL et al. Functional variability of snake venom metalloproteinases: adaptive advantages in targeting different prey and implications for human envenomation. PLoS One. 2014:9:e109651. https://doi.org/10.1371/journal.pone.0109651.

Braga MDM et al. Purification and biological effects of c-type lectin isolated from bothrops insularis venom. Toxicon. 2006:47:859–867. https://doi.org/10.1016/j.toxicon.2006.02.016.

Brown MR, Manuel Gonzalez de La Rosa P, Blaxter M. tidk: a toolkit to rapidly identify telomeric repeats from genomic datasets. Bioinformatics. 2025:41:btaf049. https://doi.org/10.1093/bioinformatics/btaf049.

Bruna T et al. Galba: genome annotation with miniprot and augustus. BMC Bioinformatics. 2023:24:327. https://doi.org/10.1186/s12859-023-05449-z.

Buchfink B, Reuter K, Drost HG. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods. 2021:18:366–368. https://doi.org/10.1038/s41592-021-01101-x.

Campos PF et al. Trends in the evolution of snake toxins underscored by an integrative omics approach to profile the venom of the colubrid phalotris mertensi. Genome Biol Evol. 2016:8:2266–2287. https://doi.org/10.1093/gbe/evw149.

Card DC et al. Genomic basis of convergent island phenotypes in boa constrictors. Genome Biol Evol. 2019:11:3123–3143. https://doi.org/10.1093/gbe/evz226.

Carone SE et al. Bjsp, a novel serine protease from bothrops jararaca snake venom that degrades fibrinogen without forming fibrin clots. Toxicol Appl Pharmacol. 2018:357:50–61. https://doi.org/10.1016/j.taap.2018.08.018.

Casewell NR. On the ancestral recruitment of metalloproteinases into the venom of snakes. Toxicon. 2012:60:449–454. https://doi.org/10.1016/j.toxicon.2012.02.006.

Casewell NR, Wagstaff SC, Harrison RA, Renjifo C, Wüster W. Domain loss facilitates accelerated evolution and neofunctionalization of

duplicate snake venom metalloproteinase toxin genes. Mol Biol Evol. 2011:28:2637–2649. https://doi.org/10.1093/molbev/msr091.
Castoe TA et al. The Burmese python genome reveals the molecular basis for extreme adaptation in snakes. Proc Natl Acad of Sci USA. 2013:110:20645–20650. https://doi.org/10.1073/pnas.1314475110.
Chen Y, Zhang Y, Wang AY, Gao M, Chong Z. Accurate long-read *de novo* assembly evaluation with inspector. Genome Biol. 2021:22:312. https://doi.org/10.1186/s13059-021-02527-4.
Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nat Methods. 2021:18:170–175. https://doi.org/10.1038/s41592-020-01056-5.
Cogo J, Prado-Franceschi J, Cruz-Hofling M, Corrado A, Rodrigues-Simioni L. Effect of bothrops insularis venom on the mouse and chick nerve-muscle preparation. Toxicon. 1993:31:1237–1247. https://doi.org/10.1016/0041-0101(93)90397-2.
Cogo JC. *Estudo dos efeitos da peçonha de Bothrops insularis (Jararaca ilhoa) sobre o musculo de camundongo e pintainho*. [Ph.D. thesis]. [sn]; 1991.
Cogo JC, Lilla S, Souza GH, Hyslop S, de Nucci G. Purification, sequencing and structural analysis of two acidic phospholipases a2 from the venom of bothrops insularis (jararaca ilhoa). Biochimie. 2006:88:1947–1959. https://doi.org/10.1016/j.biochi.2006.10.006.
Consortium ICGS. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature. 2004:432:695–716. https://doi.org/10.1038/nature03154.
Cox N et al. A global reptile assessment highlights shared conservation needs of tetrapods. Nature. 2022:605:285–290. https://doi.org/10.1038/s41586-022-04664-7.
da Cruz MHP, Domingues DS, Saito PTM, Paschoal AR, Bugatti PH. TERL: classification of transposable elements by convolutional neural networks. Brief Bioinform. 2021:22:bbaa185. https://doi.org/10.1093/bib/bbaa185.
Danecek P et al. The variant call format and vcftools. Bioinformatics. 2011:27:2156–2158. https://doi.org/10.1093/bioinformatics/btr330.
Danecek P et al. Twelve years of samtools and bcftools. Gigascience. 2021:10:giab008. https://doi.org/10.1093/gigascience/giab008.
Da Silva FO et al. The ecological origins of snakes as revealed by skull evolution. Nat Commun. 2018:9:376. https://doi.org/10.1038/s41467-017-02788-3.
de Albuquerque Modesto JC et al. Insularinase a, a prothrombin activator from bothrops insularis venom, is a metalloprotease derived from a gene encoding protease and disintegrin domains. Biol Chem. 2005:386:589–600. https://doi.org/10.1515/BC.2005.069.
Della-Casa MS et al. "Insularin, a disintegrin from bothrops insularis venom: inhibition of platelet aggregation and endothelial cell adhesion by the native and recombinant gst-insularin proteins". Toxicon. 2011:57:125–133. https://doi.org/10.1016/j.toxicon.2010.10.013.
Dobin A et al. Star: ultrafast universal rna-seq aligner. Bioinformatics. 2013:29:15–21. https://doi.org/10.1093/bioinformatics/bts635.
Dowell NL et al. The deep origin and recent loss of venom toxin genes in rattlesnakes. Curr Biol. 2016:26:2434–2445. https://doi.org/10.1016/j.cub.2016.07.038.
Emms DM, Kelly S. Orthofinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019:20:1–14. https://doi.org/10.1186/s13059-019-1832-y.
Entiauspe-Neto OM, Nachtigall PG, Borges-Martins M, de Azevedo ILJ, Grazziotin FG. Highly conserved and extremely variable: the paradoxical pattern of toxin expression revealed by comparative venom-gland transcriptomics of phalotris (serpentes: dipsadidae). Toxicon. 2024:244:107740. https://doi.org/10.1016/j.toxicon.2024.107740.
Ferreira SH, Bartelt DC, Greene LJ. Isolation of bradykinin-potentiating peptides from bothrops jararaca venom. Biochemistry. 1970:9:2583–2593. https://doi.org/10.1021/bi00815a005.
Flynn JM et al. RepeatModeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci U S A. 2020:117:9451–9457. https://doi.org/10.1073/pnas.1921046117.
Freitas-de Sousa LA et al. Size matters: an evaluation of the molecular basis of ontogenetic modifications in the composition of bothrops jararacussu snake venom. Toxins (Basel). 2020:12:791. https://doi.org/10.3390/toxins12120791.
Fry BG et al. Evolution of an arsenal: structural and functional diversification of the venom system in the advanced snakes (caenophidia)*. Mol Cell Proteomics. 2008:7:215–246. https://doi.org/10.1074/mcp.M700094-MCP200.
Furtado VV, Bonetti Filho J, Rodrigues M, Barcellos RL. Aspectos da sedimentação no canal de são sebastião. Relat Téc Insto Oceanogr. 1998:43:15–31.
Giani AM, Gallo GR, Gianfranceschi L, Formenti G. Long walk to genomics: history and current approaches to genome sequencing and assembly. Comput Struct Biotechnol J. 2020:18:9–19. https://doi.org/10.1016/j.csbj.2019.11.002.
Giorgianni MW et al. The origin and diversification of a novel protein family in venomous snakes. Proc Natl Acad Sci U S A. 2020:117:10911–10920. https://doi.org/10.1073/pnas.1920011117.
Gonçalves-Machado L et al. Combined venomics, venom gland transcriptomics, bioactivities, and antivenomics of two bothrops jararaca populations from geographic isolated regions within the Brazilian Atlantic rainforest. J Proteomics. 2016:135:73–89. https://doi.org/10.1016/j.jprot.2015.04.029.
Grazziotin FG, Monzel M, Echeverrigaray S, Bonatto SL. Phylogeography of the bothrops jararaca complex (serpentes: viperidae): past fragmentation and island colonization in the Brazilian Atlantic forest. Mol Ecol. 2006:15:3969–3982. https://doi.org/10.1111/mec.2006.15.issue-13.
Greene HW. Snakes: the evolution of mystery in nature. University of California Press; 1997.
Grundler MC, Rabosky DL. Rapid increase in snake dietary diversity and complexity following the end-cretaceous mass extinction. PLoS Biol. 2021:19:e3001414. https://doi.org/10.1371/journal.pbio.3001414.
Guan D et al. Identifying and removing haplotypic duplication in primary genome assemblies. Bioinformatics. 2020:36:2896–2898. https://doi.org/10.1093/bioinformatics/btaa025.
Guimarães M, Munguía-Steyer R, Doherty Jr PF, Martins M, Sawaya RJ. Population dynamics of the critically endangered golden lancehead pitviper, bothrops insularis: stability or decline? PLoS One. 2014:9:1–7. https://doi.org/10.1371/journal.pone.0095203.
Guimarães-Gomes V et al. Cloning, characterization, and structural analysis of a c-type lectin from bothrops insularis (bil) venom. Arch Biochem Biophys. 2004:432:1–11. https://doi.org/10.1016/j.abb.2004.08.018.
Halpern B, Vörös J, Mc Cartney AM, Formenti G, Mouton A. The genome sequence of the hungarian meadow viper, vipera ursinii rakosiensis (méhely, 1893. Wellcome Open Res. 2025:9:404. https://doi.org/10.12688/wellcomeopenres.
Hogan MP et al. The chemosensory repertoire of the eastern diamondback rattlesnake (*Crotalus adamanteus*) reveals complementary genetics of olfactory and vomeronasal-type receptors. J Mol Evol. 2021:89:313–328. https://doi.org/10.1007/s00239-021-10007-3.

Hogan MP et al. The genetic regulatory architecture and epigenomic basis for age-related changes in rattlesnake venom. Proc Natl Acad Sci U S A. 2024:121:e2313440121. https://doi.org/10.1073/pnas.2313440121.
Hogg CJ. Translating genomic advances into biodiversity conservation. Nat Rev Genet. 2024:25:362–373. https://doi.org/10.1038/s41576-023-00671-0.
Hohenlohe PA, Funk WC, Rajora OP. Population genomics for wildlife conservation and management. Mol Ecol. 2021:30:62–82. https://doi.org/10.1111/mec.v30.1.
Holding ML et al. Phylogenetically diverse diets favor more complex venoms in North American pitvipers. Proc Natl Acad Sci U S A. 2021:118:e2015579118. https://doi.org/10.1073/pnas.2015579118.
Homsi-Brandeburgo M, Queiroz L, Santo-Neto H, Rodrigues-Simioni L, Giglio J. Fractionation of bothrops jararacussu snake venom: partial chemical characterization and biological activity of bothropstoxin. Toxicon. 1988:26:615–627. https://doi.org/10.1016/0041-0101(88)90244-9.
Innan H, Kondrashov F. The evolution of gene duplications: classifying and distinguishing between models. Nat Rev Genet. 2010:11:97–108. https://doi.org/10.1038/nrg2689.
Jackson TNW, Koludarov I. How the toxin got its toxicity. Front Pharmacol. 2020:11:574925. https://doi.org/10.3389/fphar.2020.574925.
Jungo F, Bougueleret L, Xenarios I, Poux S. The uniprotkb/swiss-prot toxprot program: a central hub of integrated venom protein data. Toxicon. 2012:60:551–557. https://doi.org/10.1016/j.toxicon.2012.03.010.
Junqueira-de Azevedo IL et al. Cloning and expression of calglandulin, a new ef-hand protein from the venom glands of bothrops insularis snake in E. coli. Biochim Biophys Acta (BBA)-Proteins Proteomics. 2003:1648:90–98. https://doi.org/10.1016/S1570-9639(03)00111-0.
Junqueira-de Azevedo IL et al. Venom-related transcripts from bothrops jararaca tissues provide novel molecular insights into the production and evolution of snake venom. Mol Biol Evol. 2015:32:754–766. https://doi.org/10.1093/molbev/msu337.
Junqueira-de Azevedo IL, Farsky SHP, Oliveira MLS, Ho PL. Molecular cloning and expression of a functional snake venom vascular endothelium growth factor (vegf) from thebothrops insularis pit viper: a new member of the vegf family of proteins* 210. J Biol Chem. 2001:276:39836–39842. https://doi.org/10.1074/jbc.M106531200.
Junqueira-de Azevedo IL, Ho PL. A survey of gene expression and diversity in the venom glands of the pitviper snake bothrops insularis through the generation of expressed sequence tags (ests). Gene. 2002:299:279–291. https://doi.org/10.1016/S0378-1119(02)01080-6.
Leite YL et al. Neotropical forest expansion during the last glacial period challenges refuge hypothesis. Proc Natl Acad Sci U S A. 2016:113:1008–1013. https://doi.org/10.1073/pnas.1513062113.
Li H, Durbin R. Fast and accurate short read alignment with burrows–wheeler transform. Bioinformatics. 2009:25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
Liao Y, Smyth GK, Shi W. Featurecounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014:30:923–930. https://doi.org/10.1093/bioinformatics/btt656.
Lovell JT et al. GENESPACE tracks regions of interest and gene copy number variation across multiple genomes. Elife. 2022:11:e78526. https://doi.org/10.7554/eLife.78526.
Ludington AJ, Hammond JM, Breen J, Deveson IW, Sanders KL. New chromosome-scale genomes provide insights into marine adaptations of sea snakes (hydrophis: Elapidae). BMC Biol. 2023:21:284. https://doi.org/10.1186/s12915-023-01772-2.
Ludington AJ, Sanders KL. Demographic analyses of marine and terrestrial snakes (elapidae) using whole genome sequences. Mol Ecol. 2021:30:545–554. https://doi.org/10.1111/mec.v30.2.
Machado Braga MD et al. Purification and renal effects of phospholipase a2 isolated from bothrops insularis venom. Toxicon. 2008:51:181–190. https://doi.org/10.1016/j.toxicon.2007.08.017.
Magalhães SFV, Peixoto HM, Moura N, Monteiro WM, de Oliveira MRF. Snakebite envenomation in the Brazilian Amazon: a descriptive study. Trans R Soc Trop Med Hyg. 2019:113:143–151. https://doi.org/10.1093/trstmh/try121.
Mandelbaum FR, Reichel AP, Assakura MT. Isolation and characterization of a proteolytic enzyme from the venom of the snake bothrops jararaca (jararaca). Toxicon. 1982:20:955–972. https://doi.org/10.1016/0041-0101(82)90098-8.
Margres MJ et al. Tipping the scales: the migration–selection balance leans toward selection in snake venoms. Mol Biol Evol. 2019:36:271–282. https://doi.org/10.1093/molbev/msy207.
Margres MJ et al. The Tiger Rattlesnake genome reveals a complex genotype underlying a simple venom phenotype. Proc Natl Acad Sci U S A. 2021:118:e2014634118. https://doi.org/10.1073/pnas.2014634118.
Margres MJ, Bigelow AT, Lemmon EM, Lemmon AR, Rokyta DR. Selection to increase expression, not sequence diversity, precedes gene family origin and expansion in rattlesnake venom. Genetics. 2017:206:1569–1580. https://doi.org/10.1534/genetics.117.202655.
Marques OA, Martins M, Sazima I. A new insular species of pitviper from Brazil, with comments on evolutionary biology and conservation of the bothrops jararaca group (serpentes, viperidae). Herpetologica. 2002:58:303–312. https://doi.org/10.1655/0018-0831(2002)058[0303:ANISOP]2.0.CO;2.
Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal. 2011:17:10–12. https://doi.org/10.14806/ej.17.1.200.
Martins M, Marques OA, Sazima I. Ecological and phylogenetic correlates of feeding habits in neotropical pitvipers of the genus bothrops. Biol Vipers. 2002:307:328.
Martins M, Sawaya RJ, Marques OAV. A first estimate of the population size of the critically endangered lancehead, bothrops insularis. South Am J Herpetol. 2008:3:168–174. https://doi.org/10.2994/1808-9798(2008)3[168:AFEOTP]2.0.CO;2.
Mason AJ et al. Venom gene sequence diversity and expression jointly shape diet adaptation in pitvipers. Mol Biol Evol. 2022:39:msac082. https://doi.org/10.1093/molbev/msac082.
Mathur S, Gibbs HL. Genomic evaluation of assisted gene flow options in an endangered rattlesnake. Mol Ecol. 2025:34:e70014. https://doi.org/10.1111/mec.v34.16.
Matsubara K et al. Evidence for different origin of sex chromosomes in snakes, birds, and mammals and step-wise differentiation of snake sex chromosomes. Proc Natl Acad Sci. 2006:103:18190–18195. https://doi.org/10.1073/pnas.0605274103.
Menezes MC et al. Activation of leukocyte rolling by the cysteine-rich domain and the hyper-variable region of hf3, a snake venom hemorrhagic metalloproteinase. FEBS Lett. 2008:582:3915–3921. https://doi.org/10.1016/j.febslet.2008.10.034.
Meng G, Li Y, Yang C, Liu S. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. Nucleic Acids Res. 2019:47:e63. https://doi.org/10.1093/nar/gkz173.
Miyamoto JG et al. A novel metalloproteinase-derived cryptide from bothrops cotiara venom inhibits angiotensin-converting enzyme activity. Biochimie. 2024:216:90–98. https://doi.org/10.1016/j.biochi.2023.10.010.
Mochales-Riaño G et al. Chromosome-level reference genome for the medically important arabian horned viper (cerastes gasperettii). Gigascience. 2025:14:giaf030. https://doi.org/10.1093/gigascience/giaf030.
Modahl CM, Mrinalini, Frietze S, Mackessy SP. Adaptive evolution of distinct prey-specific toxin genes in rear-fanged snake venom.

Proc Biol Sci. 2018:285:20181003. https://doi.org/10.1098/rspb.2018.1003.
Moran Y et al. Concerted evolution of sea anemone neurotoxin genes is revealed through analysis of the nematostella vectensis genome. Mol Biol Evol. 2008:25:737–747. https://doi.org/10.1093/molbev/msn021.
Nachtigall PG et al. Toxcodan: a new toxin annotator and guide to venom gland transcriptomics. Brief Bioinform. 2021b:22:bbab095. https://doi.org/10.1093/bib/bbab095.
Nachtigall PG et al. Differences in PLA2 constitution distinguish the venom of two endemic Brazilian mountain lanceheads, *Bothrops cotiara* and *Bothrops fonsecai*. Toxins (Basel). 2022:14:237. https://doi.org/10.3390/toxins14040237.
Nachtigall PG et al. A segregating structural variant defines novel venom phenotypes in the eastern diamondback rattlesnake. Mol Biol Evol. 2025:42:msaf058. https://doi.org/10.1093/molbev/msaf058.
Nachtigall PG, Durham AM, Rokyta DR, Junqueira-de Azevedo ILM. ToxCodAn-genome: an automated pipeline for toxin-gene annotation in genome assembly of venomous lineages. Gigascience. 2024:13:giad116. https://doi.org/10.1093/gigascience/giad116.
Nachtigall PG, Grazziotin FG, de Azevedo ILMJ. MITGARD: an automated pipeline for mitochondrial genome assembly in eukaryotic species using RNA-seq data. Brief Bioinform. 2021a:22:bbaa429. https://doi.org/10.1093/bib/bbaa429.
Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015:32:268–274. https://doi.org/10.1093/molbev/msu300.
Oliveira AL et al. The chemistry of snake venom and its medicinal potential. Nat Rev Chem. 2022:6:451–469. https://doi.org/10.1038/s41570-022-00393-7.
Paes Leme A et al. Bothrops protease-a, a unique highly glycosylated serine proteinase, is a potent, specific fibrinogenolytic agent. J Thromb Haemost. 2008:6:1363–1372. https://doi.org/10.1111/j.1538-7836.2008.02995.x.
Pasquesi GI et al. Squamate reptiles challenge paradigms of genomic repeat element evolution set by birds and mammals. Nat Commun. 2018:9:2774. https://doi.org/10.1038/s41467-018-05279-1.
Patterson M et al. Whatshap: weighted haplotype assembly for future-generation sequencing reads. J Comput Biol. 2015:22:498–509. https://doi.org/10.1089/cmb.2014.0157.
Peng C et al. The genome of Shaw's sea snake (*Hydrophis curtus*) reveals secondary adaptation to its marine environment. Mol Biol Evol. 2020:37:1744–1760. https://doi.org/10.1093/molbev/msaa043.
Peng C et al. Large-scale snake genome analyses provide insights into vertebrate development. Cell. 2023:186:2959–2976. https://doi.org/10.1016/j.cell.2023.05.030.
Perry BW, Schield DR, Adams RH, Castoe TA. Microchromosomes exhibit distinct features of vertebrate chromosome structure and function with underappreciated ramifications for genome evolution. Mol Biol Evol. 2020:38:904–910. https://doi.org/10.1093/molbev/msaa253.
Raudonat HW, Rocha e Silva M. Separation of the bradykinin releasing enzyme from the clotting factor in venom from bothrops jararaca. Naunyn Schmiedebergs Arch Exp Pathol Pharmakol. 1962:243:232–236. https://doi.org/10.1007/BF00247136.
Rautsaw RM et al. Intraspecific sequence and gene expression variation contribute little to venom diversity in sidewinder rattlesnakes (Crotalus cerastes). Proc R Soc Lond B Biol Sci. 2019:286:20190810. https://doi.org/10.1098/rspb.2019.0810.
Rhie A et al. Towards complete and error-free genome assemblies of all vertebrate species. Nature. 2021:592:737–746. https://doi.org/10.1038/s41586-021-03451-0.
Rieppel O. A review of the origin of snakes. Evol Biol. 1988:22:37–130. https://doi.org/10.1007/978-1-4613-0931-4.
Robinson MD, McCarthy DJ, Smyth GK. Edger: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010:26:139–140. https://doi.org/10.1093/bioinformatics/btp616.
Roll U et al. The global distribution of tetrapods reveals a need for targeted reptile conservation. Nat Ecol Evol. 2017:1:1677–1682. https://doi.org/10.1038/s41559-017-0332-2.
Rosales-García RA et al. Sequence divergence in venom genes within and between montane pitviper (viperidae: Crotalinae: Cerrophidion) species is driven by mutation–drift equilibrium. J Mol Evol. 2023:91:514–535. https://doi.org/10.1007/s00239-023-10115-2.
Rozewicki J, Li S, Amada KM, Standley DM, Katoh K. MAFFT-DASH: integrated protein sequence and structural alignment. Nucleic Acids Res. 2019:47:W5–W10. https://doi.org/10.1093/nar/gky874.
Sachkova MY et al. The birth and death of toxins with distinct functions: a case study in the sea anemone nematostella. Mol Biol Evol. 2019:36:2001–2012. https://doi.org/10.1093/molbev/msz132.
Schendel V, Rash LD, Jenner RA, Undheim EA. The diversity of venom: the importance of behavior and venom system morphology in understanding its ecology and evolution. Toxins (Basel). 2019:11:666. https://doi.org/10.3390/toxins11110666.
Schield DR et al. The origins and evolution of chromosomes, dosage compensation, and mechanisms underlying venom regulation in snakes. Genome Res. 2019:29:590–601. https://doi.org/10.1101/gr.240952.118.
Schield DR et al. The rattlesnake W chromosome: A GC-rich retroelement refugium with retained gene function across ancient evolutionary strata. Genome Biol Evol. 2022:14:evac116. https://doi.org/10.1093/gbe/evac116.
Selistre H, Queiroz L, Cunha O, De Souza G, Giglio J. Isolation and characterization of hemorrhagic, myonecrotic and edema-inducing toxins from bothrops insularis (jararaca ilhoa) snake venom. Toxicon. 1990:28:261–273. https://doi.org/10.1016/0041-0101(90)90062-C.
Serrano SM, Maroun RC. Snake venom serine proteinases: sequence homology vs. substrate specificity, a paradox to be solved. Toxicon. 2005:45:1115–1132. https://doi.org/10.1016/j.toxicon.2005.02.020.
Serrano SM, Mentele R, Sampaio CA, Fink E. Purification, characterization, and amino acid sequence of a serine proteinase, PA-BJ, with platelet-aggregating activity from the venom of Bothrops jararaca. Biochemistry. 1995:34:7186–7193. https://doi.org/10.1021/bi00021a033.
Silva MB et al. A prothrombin activator from bothrops erythromelas (jararaca-da-seca) snake venom: characterization and molecular cloning. Biochem J. 2003:369:129–139. https://doi.org/10.1042/bj20020449.
Silveira Aea. Bothrops insularis. the iucn red list of threatened species 2021: e.t2917a123180264. https://dx.doi.org/10.2305/IUCN.UK.2021-3.RLTS.T2917A123180264.en; 2025.
Smith CF et al. Snakes on a plain: biotic and abiotic factors determine venom compositional variation in a wide-ranging generalist rattlesnake. BMC Biol. 2023:21:136. https://doi.org/10.1186/s12915-023-01626-x.
Sousa LF et al. Individual variability in bothrops atrox snakes collected from different habitats in the Brazilian amazon: new findings on venom composition and functionality. Toxins (Basel). 2021:13:814. https://doi.org/10.3390/toxins13110814.
Stanley SM. Rates of evolution. Paleobiology. 1985:11:13–26. https://doi.org/10.1017/S0094837300011362.

Supple MA, Shapiro B. Conservation of biodiversity in the genomics era. Genome Biol. 2018:19:131. https://doi.org/10.1186/s13059-018-1520-3.
Suryamohan K et al. The Indian cobra reference genome and transcriptome enables comprehensive identification of venom toxins. Nat Genet. 2020:52:106–117. https://doi.org/10.1038/s41588-019-0559-8.
Tajima F. Statistical method for testing the neutral mutation hypothesis by dna polymorphism. Genetics. 1989:123:585–595. https://doi.org/10.1093/genetics/123.3.585.
Terhorst J, Kamm J, Song Y. Robust and scalable inference of population history from hundreds of unphased whole genomes. Nat Genet. 2017:49:303–309. https://doi.org/10.1038/ng.3748.
Title PO et al. The macroevolutionary singularity of snakes. Science. 2024:383:918–923. https://doi.org/10.1126/science.adh2449.
Uetz P et al. A quarter century of reptile and amphibian databases. Herpetol Rev. 2021:52:246–255.
Valente RH et al. Bothrops insularis venomics: a proteomic analysis supported by transcriptomic-generated sequence data. J Proteomics. 2009:72:241–255. https://doi.org/10.1016/j.jprot.2009.01.001.
Vanzolini P. Distribution and differentiation of animals along the coast and in continental islands of the state of S. Paulo, Brasil. i. Introduction of the area and problems. Pap Avulsos Zool. 1973:26:281–294. https://doi.org/10.11606/0031-1049.1973.26.p281-294.
Wang YM, Peng HF, Tsai IH. Unusual venom phospholipases a2 of two primitive tree vipers trimeresurus puniceus and trimeresurus borneensis. FEBS J. 2005:272:3015–3025. https://doi.org/10.1111/ejb.2005.272.issue-12.
Waterhouse RM et al. BUSCO applications from quality assessments to gene prediction and phylogenomics. Mol Biol Evol. 2018:35:msx319. https://doi.org/10.1093/molbev/msx319.
Waters PD et al. Microchromosomes are building blocks of bird, reptile, and mammal chromosomes. Proc Natl Acad Sci U S A. 2021:118:e2112494118. https://doi.org/10.1073/pnas.2112494118.
Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol. 2019:20:257. https://doi.org/10.1186/s13059-019-1891-0.
Wüster W, Duarte MR, Salomão MG. Morphological correlates of incipient arboreality and ornithophagy in island pitvipers, and the phylogenetic position of bothrops insularis. J Zool. 2005:266:1–10. https://doi.org/10.1017/S0952836904006247.
Yan H, Bombarely A, Li S. DeepTE: a computational method for de novo classification of transposons with convolutional neural network. Bioinformatics. 2020:36:4269–4275. https://doi.org/10.1093/bioinformatics/btaa519.
Yan J, Li H, Zhou K. Evolution of the mitochondrial genome in snakes: gene rearrangements and phylogenetic relationships. BMC Genomics. 2008:9:569. https://doi.org/10.1186/1471-2164-9-569.
Zelanis A et al. Analysis of the ontogenetic variation in the venom proteome/peptidome of bothrops jararaca reveals different strategies to deal with prey. J Proteome Res. 2010:9:2278–2291. https://doi.org/10.1021/pr901027r.
Zelanis A, Travaglia-Cardoso SR, Furtado MDFD. Ontogenetic changes in the venom of Bothrops insularis (Serpentes: Viperidae)and its biological implication. South Am J Herpetol. 2008:3:43–50. https://doi.org/10.2994/1808-9798(2008)3[43:OCITVO]2.0.CO;2.
Zhang H et al. Fast alignment and preprocessing of chromatin profiles with Chromap. Nat Commun. 2021:12:6566. https://doi.org/10.1038/s41467-021-26865-w.
Zhou C, McCarthy SA, Durbin R. YaHS: yet another Hi-C scaffolding tool. Bioinformatics. 2023:39:btac808. https://doi.org/10.1093/bioinformatics/btac808.

Associate editor: Rachel Mueller