doi.org/10.1002/cbdv.202400611 Research Article CHEMISTRY & BIODIVERSITY

www.cb.wiley.com

Amazonins: New Peptaibol Sequences from an Endophytic Strain of *Trichoderma amazonicum*

Gleucinei dos Santos Castro,[a] Thiago Fernandes Sousa,[b] Ingride Jarline Santos da Silva,[b] Débora Sena Raposo,[a] José Carlos Ipuchima da Silva,[a] Evelyn Peñaloza,[c] Rafael Garrett,[c] Michel Eduardo Beleza Yamagishi,[d] Gilvan Ferreira da Silva,[e] and Hector Henrique Ferreira Koolen*[a]

Three new putative sequences of 14-residue peptaibols, named amazonins I–III were characterized from the endophytic fungus *Trichoderma amazonicum* via genome mining, high-performance liquid chromatography coupled to high-resolution tandem mass spectrometry (LC–MS/MS), and molecular networking. Bioinformatic analysis of the *T. amazonicum* genome assembly revealed 63 clusters of biosynthetic genes (BGCs) related to secondary metabolites, including a nonribosomal peptide synthetase accountable for the biosynthesis of the discovered peptide sequences. Analysis of the adenylation domains, along with manual interpretation of MS/MS spectra, allowed extensive annotation of the new peptaibol sequences. The combination of bioinformatic tools and LC–MS/MS provides a better opportunity to characterize and identify new peptaibol sequences. Thus, the importance of studies on the production and characterization of peptaibols produced by *Trichoderma* species from the Amazon region is highlighted.

Introduction

Endophytic fungi can be defined as those that exhibit a mutualistic interaction during the period they are in contact with the host plant, not causing direct or indirect adverse effects.[1] Among the endophytic fungi, species of *Trichoderma* stand out as they are predominant in different ecosystems across a wide range of climatic zones.[2]

Trichoderma has been widely associated with plant protection and growth promotion, but the importance of this genus goes far beyond, due to the vast possibilities of application in the industry, such as bioremediation,[3] industrial enzyme production,[4] food additives,[5] antibiotics,[6] biofuel production,[7] battery components,[8] and new metabolites of biotechnological interest.[9]

From a genomic standpoint, *Trichoderma* species present numerous biosynthetic gene clusters (BGCs) related to the synthesis of distinct secondary metabolites, aiding adaptation to different environments and hosts.[10] Although genomic studies reveal a vast biosynthetic potential in *Trichoderma*, under laboratory conditions, most BGCs are silenced, or their products are not detected by analytical spectrometric techniques because of the low quantity of the produced metabolite.[11]

The metabolic arsenal of *Trichoderma* mainly includes nonribosomal peptides (NRPs) synthesized by nonribosomal peptide synthetase (NRPS), which are megasynthases with different enzymatic modules, each of which adds an amino acid to the growing peptide chain.[12] In contrast to ribosomal peptide synthesis, the amino acids added along the chain can be proteinogenic or non-proteinogenic amino acids, such as α-aminoisobutyric acid (Aib) and isovaline (Iva), leading to a wide variety of peptaibols biosynthesized by a single NRPS.[13]

Among the NRPs produced, members of the peptaibol class have received much attention because of their diverse biological activities, including antibacterial properties,[14] antitumor effects,[15] and neuroleptic activities.[16] Peptaibols consist of a diverse class of biologically effective secondary metabolites, possessing linear and amphipathic structures, ranging from 5 to 20 amino acid residues, containing an acylated N-terminal (Ac) and an amino alcohol at the C-terminal portion, which can be Leuol, Valol, Pheol, Tirol, Ileol, Alaol, and Prool.[17]

The promising antimicrobial properties of these peptaibols are due to their membrane-disrupting abilities, as their helical structure, mainly formed by amino acid residues (Aib), generates ionic channels in the lipid bilayer, inducing an increase in the permeability of the cell membrane.[18] Peptaibols of the trikoningin KA V class, isolated from the fungus *Trichoderma koningii*, act on *Staphylococcus aureus* by disrupting the plasma membrane, leading to the loss of osmotic balance, making

[a] *G. dos Santos Castro, D. Sena Raposo, J. Carlos Ipuchima da Silva, H. Henrique Ferreira Koolen*
Grupo de Pesquisas em Metabolômica e Espectrometria de Massas, Universidade do Estado do Amazonas (UEA), 690065-130 Manaus, Brazil
E-mail: hkoolen@uea.edu.br

[b] *T. Fernandes Sousa, I. Jarline Santos da Silva*
Programa de Pós-graduação em Biotecnologia, Universidade Federal do Amazonas (UFAM), 69080-900 Manaus, Brazil

[c] *E. Peñaloza, R. Garrett*
Laboratório de Metabolômica (LabMeta-LADETEC), Instituto de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro 21941-598, Brazil

[d] *M. Eduardo Beleza Yamagishi*
Embrapa Agricultura Digital, 13083-970 Campinas, Brazil

[e] *G. Ferreira da Silva*
Embrapa Amazônia Ocidental, 69010-970 Manaus, Brazil

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cbdv.202400611

these peptides excellent weapons against multidrug-resistant bacteria.[19]

In addition to their antimicrobial potential, peptaibols such as pentadecaibin and its analogs To1-VX and To11-VX produced by *Trichoderma* spp. have promising activity against triple-negative breast cancer cell lines (MDA-MB-231) and ovarian cancer (SK-OV-3).[20] Extracts containing peptaibols also affected the proliferation of human breast cancer and human ovarian cancer cell lines in a 2D model, including multidrug-resistant sublines.[21] Peptaibol longibrachin-AI produced by the fungus *Trichoderma longibrachiatum* can potentiate the neurotoxicity of domoic acid, even at concentrations below the minimum effective dose.[16] These and many other examples demonstrate the potential of *Trichoderma* genus species to produce useful peptides. Additionally, numerous species of the genus *Trichoderma* that have not yet been studied represent an unexplored source of potential new secondary metabolites.

In this regard, we highlight the species *Trichoderma amazonicum*, which had its first described lineage isolated from *Hevea brasiliensis* rubber tree leaves from the Peruvian Amazon and described by Chaverri and colleagues.[22] Although not a recently described species, little is known about its genome and metabolome. In this context, this study performed genomic and metabolomic characterization of a strain of *T. amazonicum* isolated from açai fruits. Based on different approaches involving genomic mining and mass spectrometry (LC–MS/MS), we report the putative sequence of three new 14-residue peptaibols, named amazonins I–III.

Results and Discussion

Based on the concatenated alignment of the translation elongation factor tef1-α and rpb2 from the closest species within the Harzianum clade, we found 1327 characters, including gaps (tef1-α: 451 and rpb2: 876), resulting in the identification of the isolate as the species *Trichoderma amazonicum* (Figure 1). The closest species to *T. amazonicum* are *T. pleuroticola* and *T. pleuroti*.

The species *T. amazonicum* was identified by Chaverri and colleagues[22] isolated from *Hevea brasiliensis*, known only as an endophyte in Neotropical trees, thus marking the first time it was isolated from the fruit of *E. precatoria*. It is one of the most common *Trichoderma* endophytes found in sapwood in *Theobroma* spp. and *Hevea* spp.[22] Studies show that this species is a proficient producer of enzymes that can be used in the conversion of lignocellulosic biomass,[23] an antagonist of *Claviceps paspali*.[24]

Trichoderma is one of the most widely studied and intensively explored genera because of its significant biosynthetic potential in producing biologically active molecules.[25] Although many *Trichoderma* species have been reported as prolific producers of bioactive metabolites, genomic sequencing has revealed that there are still many biosynthetic gene clusters (BGCs) with new and promising activities yet to be discovered.[26]

Genome Mining of *T. amazonicum*

The genome of *T. amazonicum* contains 63 BGCs (Biosynthetic Gene Clusters), including 17 hybrid BGCs, 8 fungal RiPP-like BGCs (Ribosomally synthesized and Post-Translationally modified Peptides), 10 terpene-related clusters, 6 NRPS-like clusters (Non-Ribosomal Peptide Synthetases), 14 T1PKS-related clusters (Type I Polyketide Synthases), and 8 NRPS clusters (Non-Ribosomal Peptide Synthetases) (Figure 2). Among these BGCs, three showed 100% similarity to clusters deposited in the MIBIG (Minimum Information about a Biosynthetic Gene Cluster) repository and are directly related to the biosynthesis of the

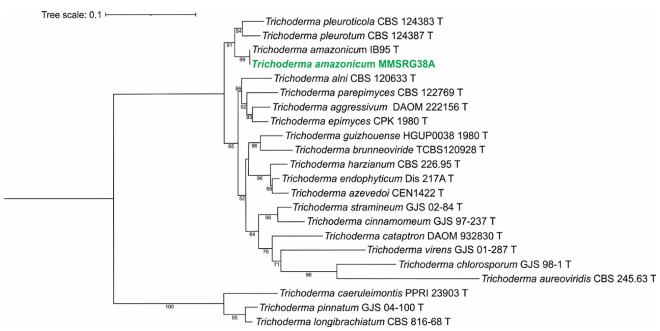

Figure 1. Dendrogram generated by concatenating the sequences of tef1-α and rpb2 from the *Trichoderma* species closest to *T. amazonicum*.

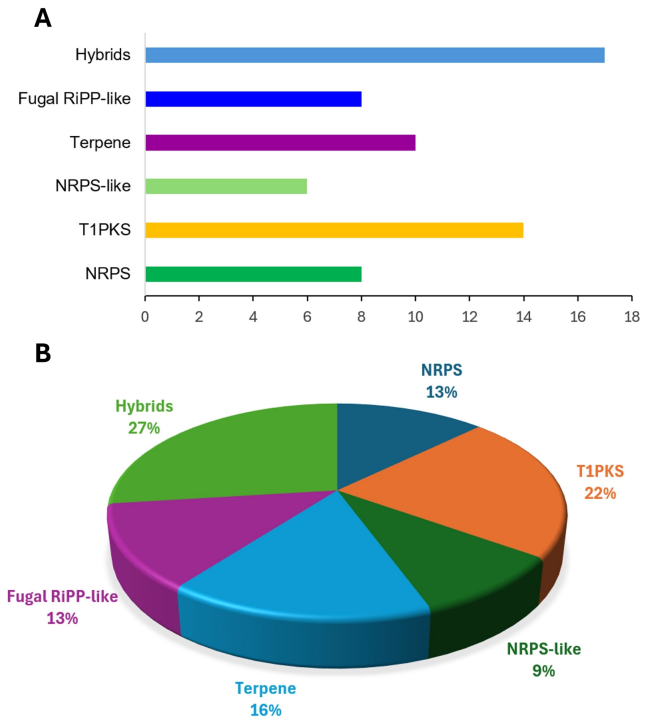

Figure 2. In (A), identification and cluster number of biosynthetic genes in *T. amazonicum*: NRPSs (nonribosomal peptide synthetases), T1PKS (type I polyketide synthetase), hybrids (BGCs containing genes encoding more than one enzyme synthesizing molecules of different classes), terpene, fungal-RiPP-like (clustering of fungal peptide products, unspecified, ribosomally-synthesized and post-translationally modified (RiPP)). In (B), distribution of BGCs based on similarity with BGCs deposited in the MIBiG database.

compounds: clavaric acid (BGC 1.2), choline (BGC 2.2) and harziphilone (BGC 82.1). Two clusters showed more than 80% similarity: harzianopyridone (BGC 20.2) and tricholignan A (BGC 36.1). The NRPS responsible for the synthesis of 14-residue peptaibols in the genome of *T. amazonicum* is in BGC (7.3).

The diversity of Biosynthetic Gene Clusters (BGCs) revealed an impressive variety of BGCs, with a total of 63 clusters identified. These BGCs encompass a wide range of classes, highlighting the potential of this species to produce a variety of secondary metabolites. Notably, 45% of the identified BGCs showed no similarity to clusters deposited in the MIBiG database, suggesting that they may be involved in the biosynthesis of new molecules or known molecules whose pathways have not yet been elucidated.

Comparative analysis of 38 *Trichoderma* genomes revealed that endophytic isolates contain, on average, more total BGCs than non-endophytic isolates. This observation suggests that the ability to create/modify a diversity of potential metabolites is beneficial or necessary for *Trichoderma*.[26] The high number of uncharacterized BGCs highlights the potential of T. amazonicum to produce novel secondary metabolites. These BGCs may be related to the biosynthesis of new molecules or known molecules whose biosynthetic pathways have not yet been deposited in the repository (Table S1, Supplementary Material).

The uncharacterized BGCs may be related to the biosynthesis of new molecules or known molecules whose biosynthetic pathways have not yet been deposited in the repository (Table S1, Supplementary Material). BGCs showing 100% similarity to those deposited in the MIBIG database have pathways identified to produce clavaric acid (BGC 1.2), a triterpene that inhibits farnesyl transferase protein and, therefore, is a promising anticancer substance targeting the Ras oncogene (a gene mutated in various types of cancer).[27] Another BGC was related to the molecule harziphilone, which has inhibitory activity against the binding of the REV protein (Regulation of Virion Expression) to RRE RNA (Rev Responsive Element), as well as exerting cytotoxic activity against the murine tumor cell line M-109,[28] and t22 azaphilone, a potent antifungal produced by *Trichoderma harzianum* T22.[29]

BGCs with 80% and 90% similarity correspond to the compounds harzianopyridone (BGC 20.2), which exhibit significant biological activity, including antifungal properties and cytotoxic activity against cancer cells,[30] and tricholignan A (BGC 36.1), responsible for reducing Fe^{3+}, which plays a role in promoting plant growth under conditions of deficiency of this element.[31]

Analysis of 14-Module NRPS

The *T. amazonicum* isolate was cultivated on rice and subjected to an exploratory analysis of its chemical profile. Using LC–MS/MS in positive mode, the analysis revealed that the crude extract contained compounds with a molecular mass greater than 1.000 u. This observation suggests the presence of peptide-like molecules.

Once the peptaibol sequences were deduced by LC–MS/MS analyses, the BGC located on scaffold 8 (BGC 7.3) emerged as a candidate responsible for the production of the identified peptaibols. This 50 Kb BGC encodes an NRPS with a starter region featuring PKS domains (KS-AT-ACP), indicating it is a PKS-NRPS hybrid.[32] Phylogenetic analysis of all adenylation domains present in the PKSs-NRPSs of 14 modules from *T. amazonicum* were grouped with the NRPS2 domains of the peptaibol synthetase from *T. virens* Tv29–8 (Figure 3), as described in previous studies.[33] The assembly position of α-aminoisobutyric acid (Aib) can be inferred in the adenylation domains corresponding to modules 1, 4, 5, 10, 11, and 12, while modules 3, 6, and 7 show promiscuity for leucine/isoleucine or valine/isovaline due to substrate similarity. Glutamine amino acids are predicted for positions 2 and 12, and they are only grouped with the adenylation domain of module 2 of the *T. virens* NRPS, which occupies a distinct phylogenetic position compared with other adenylation domains that have affinity for these substrates.

In silico prediction of 14-residue peptaibols was performed using a phylogenetic approach combined with manual interpretation of spectra and molecular networks. This allowed for the precise positioning of b5 (Aib) and b6 (Gln) ions. Amino acid prediction was conducted through phylogenetic analyses using the adenylation domains of peptaibol synthetases, as previously described.[33] The primary use of this tool is to identify the position of amino acids based on the collinearity of the peptide and NRPS modules.

Annotation of the Peptaibols

In the molecular networks (Figure 4), several connected nodes were observed, with the largest network (15 nodes) being of interest for peptaibol identification, as it contained nodes corresponding to ions with high *m/z*. The molecular networking analysis revealed the presence of multiple peptaibol metabolites produced by *T. amazonicum*. The nodes in the largest network did not show connections to any characterized peptaibols in the GNPS[34] libraries, indicating the novelty of these compounds. Manual analysis (Figure 5) enabled their annotation

The characterization of the three amazonins demonstrates the ability of *T. amazonicum* to produce structurally similar peptaibols with variations in specific residues. This suggests that the NRPS responsible for their biosynthesis has some flexibility in amino acid incorporation, especially at positions 4, 7, 8, and 10. The preservation of the N-terminal and C-terminal motifs may indicate their importance for the biological activity of these molecules.

Future studies involving the isolation and complete structural characterization of these compounds may reveal the presence of additional peptaibols produced by *T. amazonicum*. Exploring different cultivation conditions and employing other analytical approaches, such as sequential tandem mass spectrometry (MSn), could contribute to identifying an even greater diversity of peptaibols in this fungus.

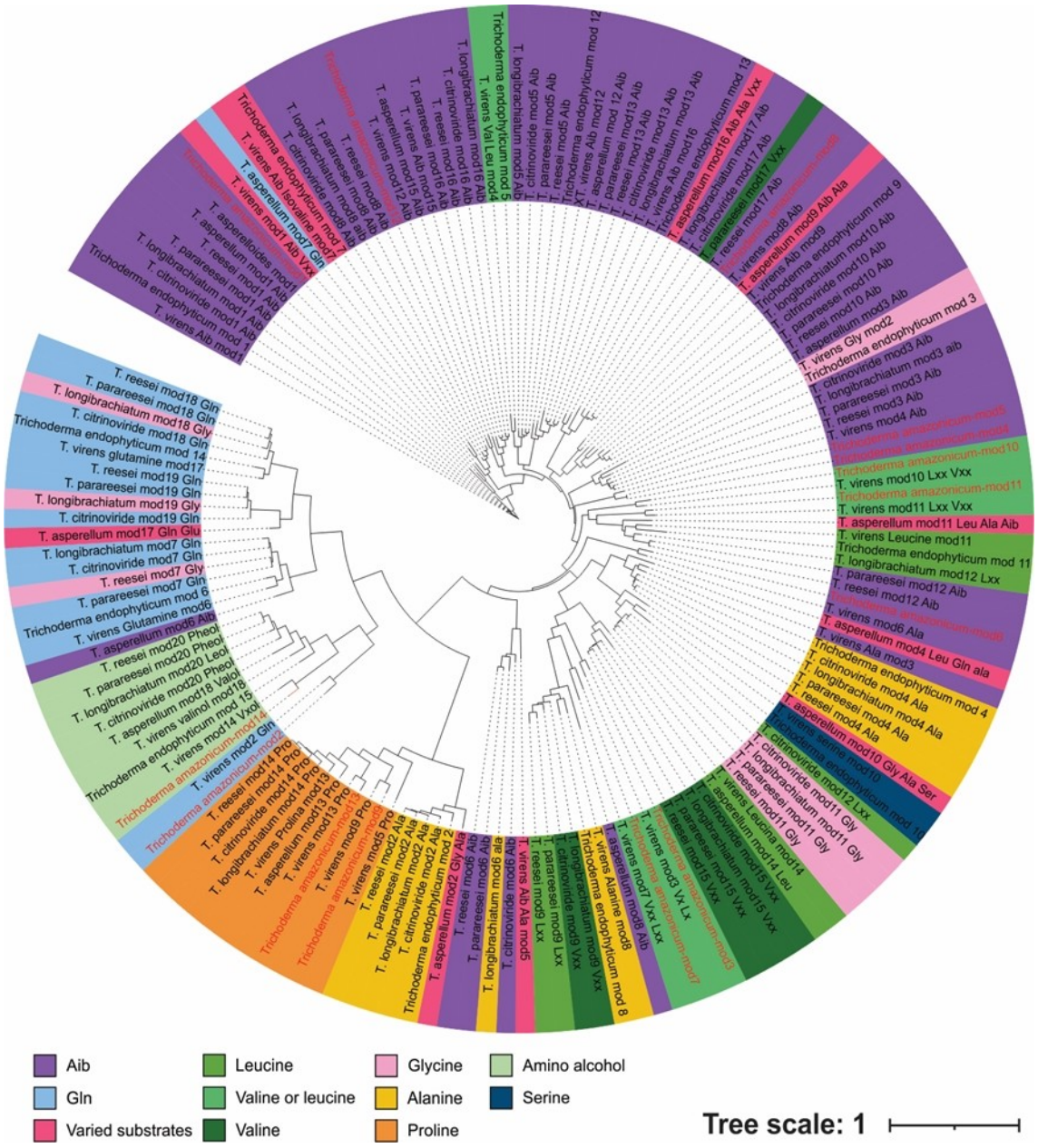

Figure 3. Phylogeny of NRPS adenylation modules from 14 modules.

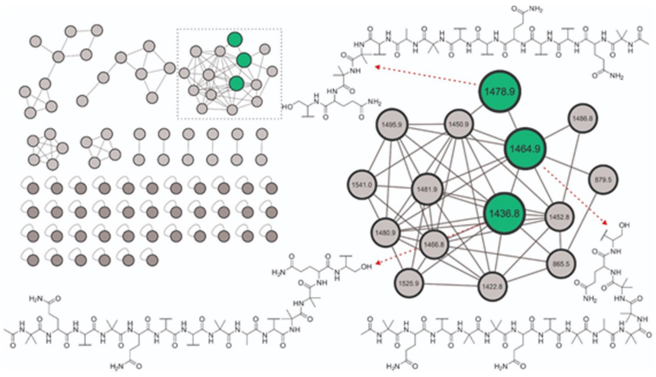

Figure 4. Annotation of the molecular network of peptaibols. Green nodes: putative 14-residue sequences.

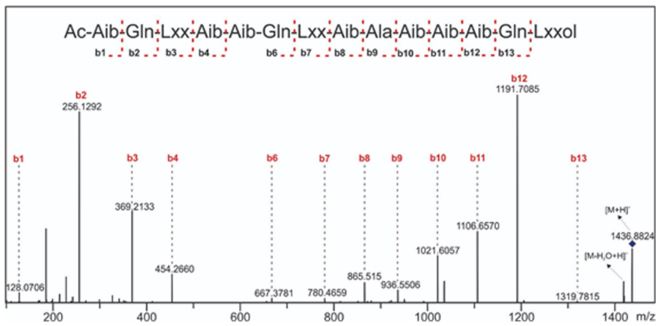

Figure 5. MS/MS sequencing of peptaibol. Sequences are provided in standard orthographic code (Vxx, Val/Iva, Lxx, Leu/Ile, and Lxxol represent the C-terminal amino alcohol).

Table 1. New putative peptaibols sequences produced by *T. amazonicum*.

Compound	*m/z* $[M+H]^+$	t_R/min	Chemical formula	Error/ppm	Sequence
Amazonin I	1436.8812	11.4	$C_{66}H_{118}N_{17}O_{18}$	1.94	Ac-Aib-Gln-**Lxx-Aib**-Aib-Gln-**Lxx-Aib**-Ala-**Aib**-Aib-Aib-Gln-lxxol
Amazonin II	1464.9156	11.6	$C_{68}H_{121}N_{17}O_{18}$	0.20	Ac-Aib-Gln-**Lxx-Aib**-Aib-Gln-**Vxx-Aib**-Ala-**Vxx**-Aib-Aib-Gln-Lxxol
Amazonin III	1478.9308	11.7	$C_{69}H_{123}N_{17}O_{18}$	0.13	Ac-Aib-Gln-**Lxx-Vxx**-Aib-Gln-**Vxx-Lxx**-Ala-**Vxx**-Aib-Aib-Gln-Lxxol

Thus, the annotation of putative sequences was only possible through manual interpretation of MS/MS spectra from the different chromatographic peaks in the LC–MS/MS analysis, which enabled the characterization of three putative peptaibol sequences containing 14 amino acid residues that were structurally similar.

Initially, the fragmentation mechanism of compound **1** (*m/z* 1436.8812 $[M+H]^+$, $C_{66}H_{118}N_{17}O_{18}$, 1.94 ppm) was analyzed as a model for the fragmentation of the other two present peptaibols (Figure 5). In the MS/MS spectrum, ions from the b series are observed, comprising ions from b1 to b14, moving from the C-terminal to the N-terminal portion. In the C-terminal portion, fragments *m/z* 1436→*m/z* 1418 ($-H_2O$) and *m/z* 1319→(b14, −117 u) confirmed that this portion is the amino alcohol leucinol/isoleucinol (Lxxol). Subsequently, characteristic losses of amino acid residues such as glutamine (Gln, −128 u), three residues of Aib (−85 u), alanine (71 u), Aib (−85 u), leucine/isoleucine (Lxx, −113 u), glutamine and Aib (Gln, −128 u + 85 u), Aib (−85 u), leucine/isoleucine (Lxx, −113 u), glutamine (Gln, −128 u), and Aib + acetyl group (−127 u) were observed. Although good coverage of fragments in terms of amino acid sequence was observed, the b5 fragment could not be seen in the MS/MS spectrum. To address this issue, we combined genome mining information with a phylogenetic analysis of the peptaibol synthetase adenylation modules to determine the sequence of this peptide portion. Upon analyzing the phylogenetic data, it was observed that modules 4 and 5 incorporate Aib and Gln, respectively, based on high bootstrap support (>92%). Furthermore, this is reinforced by the phylogeny of all modules, which showed a bootstrap of 100% for the incorporation of the amino acids composing the peptaibol sequence. Thus, the sequence of compound 1 was determined as Ac-Aib-Gln-Lxx-Aib-Aib-Gln-Lxx-Aib-Ala-Aib-Aib-Aib-Gln-Lxxol, named amazonin I.

The amazonins II and III differ by the replacement of amino acids in positions 4, 7, 8 and 10 (Table 1, Figures S2 and S3, Supplementary material). Amazonins preserve the motifs Ac-Aib-Gln (residues 1 and 2) and Aib-Aib-Gln and Lxxol (residues 11, 12, 13, and 14, respectively). These data were confirmed based on in silico predictions of NRPS adenylation domains located in the cluster 7.3 (Figure 3). The characterized peptaibols here present leucinol/isoleucinol as a C-terminal amino alcohol.

The 14-residue peptaibols produced by the PKSs-NRPSs of cluster 7.3 represent novel peptaibols. These peptaibols show similarities to the endofitins from series A (Table 2), as characterized in *T. endophyticum*.[33] However, they exhibit sequence patterns distinct from the endofitins, with an Aib-Gln sequence at the N-terminal portion and Aib-Aib-Gln-Lxxol at the C-terminal portion. The unique

Table 2. Comparison of putative 14-residue peptaibol sequences with endophytins and harzianins peptaibols.

Compound	Sequence
Amazonin I	Ac-Aib-Gln-Lxx-Aib-Aib-Gln-Lxx-Aib-Ala-Aib-Aib-Aib-Gln-Lxxol
Amazonin II	Ac-Aib-Gln-Lxx-Aib-Aib-Gln-Vxx-Aib-Ala-Vxx-Aib-Aib-Gln-Lxxol
Amazonin III	Ac-Aib-Gln-Lxx-Vxx-Aib-Gln-Vxx-Lxx-Ala-Vxx-Aib-Aib-Gln-Lxxol
Endophytin A1	Ac-Aib-Ala-Aib-Ala-Ala-Gln-Aib-Lxx-Aib-Ala-Aib-Aib-Ala-Gln-Lxxol
Endophytin A2	Ac-Aib-Ala-Aib-Ala-Aib-Gln-Aib-Vxx-Aib-Ala-Aib-Aib-Ala-Gln-Lxxo
Endophytin A3	Ac-Aib-Ala-Aib-Ala-Aib-Gln-Aib-Vxx-Aib-Ala-Aib-Aib-Aib-Gln-Lxxol
Endophytin A4	Ac-Aib-Ala-Aib-Ala-Aib-Gln-Vxx-Vxx-Aib-Ala-Aib-Aib-Ala-Gln-Lxxol
Harzianin HC I	Ac-Aib-Asn-Leu-Aib-Pro-Ser-Val-Aib- Pro-Aib-Leu-Aib-Pro-Leuol
Harzianin HCXIII	Ac-Aib-Gln-Leu-Aib-Pro-Ser-Ile-Aib-Pro-Iva-Leu-Aib-Pro-Leuol
Harzianin HCXIV	Ac-Aib-Asn-Leu-Aib-Pro-Ala-Ile-Aib-Pro-Aib-Leu-Aib-Pro-Leuol
Harzianin HCXV	Ac-Aib-Gln-Leu-Aib-Pro-Ala-Ile-Aib-Pro-Iva-Leu-Aib-Pro-Leuol
Harzianin HC II	Ac-Aib-Asn-Leu-Aib-Pro-Ser-Ile-Aib-Pro-Iva-Leu-Aib-Pro-Leuol

structural features of the amazonins suggest they may have different biological activities compared to known peptaibols. The stereochemistry of these compounds has not been determined, as they have not yet been isolated. The new peptaibols from cluster 7.3, despite their unique sequence patterns, are likely to share the same configuration at Cα as other peptaibols produced by similar NRPS machinery, assuming no epimerization domains alter this configuration. Future studies will be crucial in confirming these stereochemical details and fully understanding the biological significance of these novel compounds.[35] Interestingly, the amazonins do not show significant similarity to the 14-residue harzianin HC series peptaibols (Table 2) produced by other Trichoderma species. Some harzianins, such as harzianin HC I, III, XII, XIV, and XV, have been reported to exhibit antiplasmodial activity.[36,37] The distinct sequences of the amazonins compared to the harzianins indicate they may have different target specificities and biological effects.

The discovery of novel peptaibols in *T. amazonicum* highlights the importance of exploring the chemical diversity of endophytic fungi from unique ecological niches, such as the Amazon rainforest. The integration of genomic analysis, molecular networking, and detailed MS/MS interpretation proved to be a powerful approach for identifying new peptide natural products.

Conclusions

In this study, we investigated the ability of *T. amazonicum*, isolated for the first time from *E. precatoria*, to produce new secondary metabolites. Genomic analysis indicated a vast repertoire of metabolites through the presence of uncharacterized BGCs (biosynthetic gene clusters). The isolate proved to be a prolific producer of peptides belonging to the class of peptaibols. Through an integrated approach involving chemistry, molecular biology, and bioinformatics, three new putative sequences of 14-residue peptaibols were sequenced and named amazonins I–III. These molecules, with their unique properties, have the potential to develop new drugs and therapies that are less harmful to humans. Future studies could incorporate the determination of effective application levels in the field or in less invasive therapies. This result demonstrates the potential of combining molecular and chemical approaches for the discovery of new natural products and highlights the potential of Amazonian fungi in the production of new chemical entities.

Experimental Section

Fungal Collection and Acquisition

Fungal isolation from the fruit of *Euterpe precatoria* was performed using a methodology previously described.[38] Initially, the isolation procedure involved superficial disinfection of the açai fruit with 70% ethanol for 45 s, followed by washing with sterile water. Subsequently, square pieces of the açai fruit peel were inoculated onto Petri dishes containing PDA culture medium. After 5 days, the isolate emerged, which was then transferred to other Petri dishes containing the same culture medium and subsequently purified using monosporium cultures. Access to the genetic heritage was authorized by SISGEN N° A5F89AB.

Fungal Identification

To obtain the mycelial mass of the isolate, the mycelium was cultivated in potato dextrose broth (PD, 20 g/L potato; 20 g/L dextrose) at 28 °C with rotation at 125 rpm for 2 days. The mycelial mass obtained was filtered, and DNA extraction was performed using 2% cationic detergent CTAB, according to Doyle and Doyle.[39]

PCR amplifications of the three loci (ITS, tef1-α, and rpb2) were performed using an Easytaq® kit (Synapse Biotechnology) with three pairs of primers: ITS1 (TCCGTAGGTGAACCTGCGG) and ITS4 (TCCTCCGCTTATTGATATG) were used to amplify the internal transcribed spacer[40]; primers EF-1αF (ATGGGTAAGGARGACAAGAC) and EF-1αR (GGARGTACCAGTSATCATGTT) were used to amplify the partial sequence of the translation elongation factor gene; primers fRPB2–5f (GAYGAYMGWGATCAYTTYGG) and fRPB2–7cr (CCCATRGCTbTGTYYY) were used to amplify the partial sequence of the second largest subunit of RNA polymerase II.[41] The PCR conditions for amplification of all loci was: initial denaturation at 95 °C for 3 min, 35 cycles of denaturation at 95 °C for 45 seconds, annealing temperature at 57 °C for 45 seconds, extension at 72 °C for 1 min, and final extension at 72 °C for 5 min. The PCR products were subjected to agarose gel electrophoresis (1.5%) to confirm the amplicon length using a 1 kb marker (Invitrogen).

Phylogenetic Analyses

The tef1-α and rpb2 sequences were individually aligned using the MAFFT tool in the UGENE software.[42] The alignments were then input into the IQ-TREE platform,[43] and phylogenetic analysis using maximum likelihood (ML) was performed with the concatenated sequences of tef1-α and rpb2.

The ML analysis included 1,000 replicates (bootstrap) conducted by IQ-TREE. All sites were considered, and the analysis was run for ten million generations, with the first 25% of trees discarded and burned using the MrBayes tool (v. 3.7). The posterior probability (PP) and tree topology were visualized using the iTOL platform (itol.embl.de).

Fermentation and Chemical Extraction

After reactivation of the isolate, three square agar plugs (3 cm^2) containing mycelium and spores were inoculated into three Erlenmeyer flasks (three plugs per flask) containing 20 g of parboiled rice and 70 mL of a solution containing sodium nitrate ($NaNO_3$, 0.3 g/L), monopotassium phosphate (KH_2PO_4, 0.1 g/L), magnesium sulfate ($MgSO_4 .7H_2O$, 0.05 g/L), and potassium chloride (KCl, 0.05 g/L). Subsequently, the flasks were incubated for a period of 14 days at room temperature (25 °C). After this period, the fungal material was extracted once by cold maceration with 100 mL of ethyl acetate, and the organic phases from the three flasks were combined and concentrated under reduced pressure to yield the crude extract.[33]

The crude extract was subjected to chromatographic analysis using a Thermo Scientific Dionex Ultimate 3000 liquid chromatography system coupled to a Thermo Scientific™ Q Exactive™ Plus high-resolution mass spectrometer (Waltham, MA, USA). Liquid chromatography analyses were performed using an Ascentis C_{18} Express column (100×4.6 mm; 2.7 μm) (with a guard column) (Supelco, Bellefonte, PA, USA) with ammonium formate (0.1% w/v) (mobile phase A): acetonitrile/formic acid (0.1% w/v) (mobile phase B) at a flow rate of 0.5 mL/min in gradient elution mode as follows: B 15% (0–1 min); B 15–95% (1–16 min); B 95% (16–21 min); B 95–15% (21–22 min); and B 15% (22–30 min). The oven temperature was maintained at 40 °C. The source ionization parameters were: spray voltage 3.9 kV; capillary temperature 300 °C; S-Lens level 50, sheath gas 50, auxiliary gas 15. Samples were analyzed in the mass scan range of *m/z* 150 to 1600 at a resolution of 35,000 (full scan positive and negative) followed by data-dependent MS2 (ddMS2 Top3 experiments) using a resolution of 17,500 and collision energy normalized (NCE) scaled to 35–50%.

Genome Mining

The entire genome was sequenced using the Illumina platform with 150 bp paired-end reads. De novo assembly was performed using the SPAdes assembler.[44] The genome assembly resulted in 41,106,830 bp with coverage of 294X and 175 scaffolds with an N50 of 3,103,555 bp. The longest scaffold had a size of 7.46 Mbp.

The genome of *T. amazonicum* was subjected to the FungiSMASH platform (https://fungismash.secondarymetabolites.org/), accessed on January 15, 2024, for the prediction of biosynthetic gene clusters (BGCs). The amino acid sequence of the TEX1 protein from *Trichoderma virens* (XP 013953110.1) was used to locate scaffolds containing BGCs related to peptaibol biosynthesis using the BLASTp tool on Web-Server 2.0.

The BGCs related to peptaibol production were reannotated using the MAKER software available on the Galaxy Australia platform

(usegalaxy.org.au) using *Fusarium graminearum*. Adenylation domains of NRPSs characterized in the literature were used to generate a dataset.[33] For this, prediction of adenylation domain binding sites was performed through the PKS/NRPS analysis site (//nrps.igs.umaryland.edu/ (accessed on January 15, 2024)). The constructed dataset was used for substrate affinity prediction through phylogenetic inference using the ML method with 1000 replicates on the IQ-TREE platform. Using the iTOL site, a dendrogram was generated which was visualized and edited with the assistance of CorelDraw software, version 2020.

Supporting Information Summary

Supplementary material from MS/MS spectra and Genome Mining can be accessed through the following link: https://doi.org/10.1002/cbdv.202400611.

Author Contributions

Conceptualization and supervision, **Hector Henrique Ferreira Koolen** and **Gilvan Ferreira da Silva**; data curation and fungal acquisition, **Débora Sena Raposo**; fungal identification and phylogenetic analysis, **Thiago Fernandes Sousa**; fungal fermentation and extraction, **Ingride Jarline Santos da Silva**; genome mining and modular peptaibol assembly prediction, **Thiago Fernandes Sousa, Michel Eduardo Beleza Yamagishi**; LC–MS/MS analysis, **Gleucinei dos Santos Castro, Evelyn Peñaloza, Rafael Garrett**; molecular networking, **José Carlos Ipuchima da Silva**; writing – original draft preparation, **Gleucinei dos Santos Castro, Hector Henrique Ferreira Koolen**, and **Gilvan Ferreira da Silva**; writing – review and editing, all co-authors; project administration and funding acquisition, **Hector Henrique Ferreira Koolen** and **Gilvan Ferreira da Silva**; All authors have read and agreed to the published version of the manuscript.

Acknowledgements

The authors would like to thank the funding agencies Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES – Finance code 001), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM) and Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ). Regarding FAPEAM, the authors would like to mention the following grants received: POSGRAD 2023/FAPEAM, 008/2021 PROSPAM/FAPEAM and 010/2021 CTI AREAS PRIORITÁRIAS/FAPEAM for H.H.F.K. The author H.H.F.K. also acknowledges CNPq for the productivity scholarship (No. 305602/2023-3).

Conflict of Interests

The authors declare no conflict of interest.

Keywords: Endophytic · Peptaibols · LC–MS/MS

[1] Z. Wang, L. Wang, Y. Pan, X. Zheng, X. Liang, L. Sheng, D. Zhang, Q. Sun, Q. Wang, *Bioprocess Biosyst. Eng.* **2023**, *46*, 165–170. https://doi.org/10.1007/s00449-022-02840-7.
[2] V. N. Brito, J. L. Alves, K. S. Araújo, T. S. Leite, C. B. Queiroz, O. L. Pereira, M. V. Queiroz, *Front. Microbiol.* **2023**, *14*, 1095199. https://doi.org/10.3389/fmicb.2023.1095199.
[3] C. Daccò, N. Lidia, M. E. E. Temporiti, M. Barbara, F. Corana, C. Giovanna, S. Tosi, *Appl. Sci.* **2020**, *10*, 3152. https://doi.org/10.3390/app10093152.
[4] R. Kalsoom, S. Ahmed, M. Nadeem, S. Chohan, M. Abid, *Int. J. Environ. Sci. Technol.* **2019**, *16*, 921–928. https://doi.org/10.1007/s13762-018-1717-8.
[5] Y. Cui, J. Li, D. Deng, H. Lu, Z. Tian, Z. Liu, X. Ma, *Plos One* **2021**, *16*, e0260045. https://doi.org/10.1371/journal.pone.0260045.
[6] E. L. Ghisalberti, K. Sivasithamparam, *Soil Biol. Biochem.* **1991**, *23*, 1011–1020. https://doi.org/10.1016/0038-0717(91)90036-J.
[7] V. Gusakov, *Trends Biotechnol.* **2011**, *29*, 419–425. https://doi.org/10.1016/j.tibtech.2011.04.004.
[8] S. Shen, X. Xia, Y. Zhong, S. Deng, D. Xie, B. Liu, Y. Zhang, G. Pan, X. Wang, J. Tu, *Adv. Mater.* **2019**, *31*, 1900009. https://doi.org/10.1002/adma.201900009.
[9] K. R. P. Fernandes, A. Q. L. Souza, R. Al. Santos, F. C. S. Nogueira, J. A. M. Evaristo, G. R. A. Carneiro, G. F. Silva, J. C. Cruz, T. F. Sousa, S. R. S. Silva, I. S. Bastos, P. P. Orlandi, A. D. L. Souza, *Rev. Bras. Farmacogn.* **2021**, *31*, 667–675. https://doi.org/10.1007/s43450-021-00215-9.
[10] C. Derntl, F. Guzmán-Chávez, T. M. M. Sousa, H. J. Busse, A. J. M. Driessen, R. L. Mach, A. R. M. Aigner, *Front. Microbiol.* **2017**, *8*, 2037. https://doi.org/10.3389/fmicb.2017.02037.
[11] Q. Guo, I. Shi, X. Wang, D. Li, Z. Yin, J. Zhang, G. Ding, L. Chen, *J. Agric. Food Chem.* **2023**, *71*, 13612–13632. https://doi.org/10.1021/acs.jafc.3c04540.
[12] R. D. Süssmuth, A. Mainz, *Angew. Chem.* **2017**, *129*, 3824–3878. https://doi.org/10.1002/ange.201609079.
[13] J. Grünewald, M. A. Marahiel, *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 121–146. https://doi.org/10.1128/mmbr.70.1.121-146.2006.
[14] Y. Sang, F. Blecha, *Anim. Health Res. Rev.* **2008**, *9*, 227–235. https://doi.org/10.1017/S1466252308001497.
[15] N. Casagrande, C. Borghese, L. Gabbatore, L. Morbiato, M. Zotti, D. Aldinucci, *Int. J. Mol. Sci.* **2021**, *22*(16), 8362. https://doi.org/10.3390/ijms22168362.
[16] N. Ruiz, K. Petit, M. Vansteelandt, I. Kerzaon, J. Baudet, Z. Amzil, J–F. Biard, O. Grovel, Y. F. Pouchus, *Toxicon* **2010**, *55*, 805–810. https://doi.org/10.1016/j.toxicon.2009.11.015.
[17] N. Stoppacher, N. K. N. Neumann, L. Burgstaller, S. Zeilinger, T. Degenkolb, H. Brückner, R. Schuhmacher, *Chem. Biodivers.* **2013**, *10*, 734–743. https://doi.org/10.1002/cbdv.201200427.
[18] P. K. Bjørk, N. T. Johansen, N. W. Havshøi, S. A. Rasmussen, J. Ø. Ipsen, T. Isbrandt, T. O. Larsen, A. T. Fuglsang, *ACS Omega* **2023**, *8*, 34928–34937. https://doi.org/10.1021/acsomega.3c04299.
[19] C. A. Guette, S. Rebuffat, I. Vuidepot, M. Massias, B. Bodo, *Chem. Soc., Perkin Trans.* **1993**, *2*, 249–255. https://doi.org/10.1039/P19930000249.
[20] L. Morbiato, C. Quaggia, L. Menilli, C. D. Torre, A. Barbon, M. Zotti, *Int. J. Mol. Sci.* **2023**, *24*, 13396. https://doi.org/10.3390/ijms241713396.
[21] J. Víglaš, S. Dobiasová, J. Viktorová, T. Ruml, V. Repiská, P. Olejníková, H. Gbelcová, *Molecules* **2021**, *26*, 6025. https://doi.org/10.3390/molecules26196025.
[22] P. Chaverri, R. O. Gazis, G. J. Samuels, *Mycologia* **2011**, *103*, 139–151. https://doi.org/10.3852/10-078.
[23] B. C. Okeket, R. W. Hall, A. Nanjundaswamy, M. S. Thomson, Y. Deravi, L. Sawyer, A. Prescott, *Microbiol. Res.* **2015**, *175*, 24–33. https://doi.org/10.1016/j.micres.2015.03.001.
[24] M. B. Depetris, S. A. Gutiérrez, V. B. Skuletich, D. A. Zandoná, C. A. Acuña, S. A. Felitti, *Crop Sci.* **2023**, *6*, 3498–3508. https://doi.org/10.1002/csc2.21113.
[25] J. W. Lee, J. E. Collins, K. L. Wendt, D. Chakrabarti, R. H. Cichewicz, *J. Nat. Prod.* **2021**, *84*, 503–517. https://doi.org/10.1021/acs.jnatprod.0c01370.
[26] K. Scott, Z. Konkel, E. G. Thaler, G. E. V. David, C. F. Simmt, D. Grootmyers, P. Chaverri, J. Slot, *Plos one* **2023**, *18*, e0289280. https://doi.org/10.1371/journal.pone.02892.
[27] R. B. Lingham, K. C. Silverman, H. Jayasuriya, B. M. Kim, S. E. Amo, F. R. Wilson, D. J. Rew, M. D. Schaber, J. D. Bergstrom, K. S. Koblan, S. L. Graham, N. E. Kohl, J. B. Gibbs, S. B. Singh, *J. Med. Chem.* **1998**, *41*, 4492–4501.

[28] J. Qian-Cutrone, S. Huang, L. P. Chang, D. M. Pirnik, S. E. Klohr, R. A. Dalterio, R. Hugill, S. Lowe, M. Alam, K. F. Kadow, *J. Antibiot.* **1996**, *49*, 990–997, https://doi.org/10.7164/antibiotics.49.990.
[29] F. Vinale, E. L. Ghisalberti, K. Sivasithamparam, R. Marra, A. Ritieni, R. Ferracane, S. Woo, M. Lorito, *Lett. Appl. Microbiol.* **2009**, *48*, 705–711. https://doi.org/10.1111/j.1472-765X.2009.02599.x.
[30] H. G. Cutler, J. M. Jacyno, *Agric. Biol. Chem.* **1991**, *55*, 2629–2631. https://doi.org/10.1271/bbb1961.55.2629.
[31] M. Chen, Q. Liu, S. S. Gao, Y. Tang, *PNAS* **2019**, *116*, 5499–5504. https://doi.org/10.1073/pnas.1819998116.
[32] G. Wu, B. T. M. Dentinger, J. R. Nielson, R. T. Peterson, J. M. Winter, *J. Nat. Prod.* **2021**, *84*, 1113–1126. https://doi.org/10.1021/acs.jnatprod.0c01186.
[33] G. S. Castro, T. F. Sousa, G. F. Silva, R. C. N. Pedroso, K. S. Menezes, M. A. Soares, G. M. Dias, A. O. Santos, M. E. B. Yamagishi, J. V. Faria, A. H. Januário, H. H. F. Koolen, *Metabolites* **2023**, *13*, 221. https://doi.org/10.3390/metabo13020221.
[34] Global Natural Product Social Molecular Networking (GNPS), https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=06e87640023b4548a2ad7c670dc31caa, accessed in July 2024.
[35] I. Pérez-Victoria, *ChemBioChem* **2021**, *22*, 2087–2092. https://doi.org/10.1002/cbic.202000805.
[36] I. Ségalas, Y. Prigent, D. Davoust, B. Bodo, S. Rebuffat, *Biopolymers* **1999**, *50*, 71–85. https://doi.org/10.1002/(SICI)1097-0282(199907)50:1<71::AID-BIP7>3.0.CO;2-5.
[37] I. Augeven-Bour, S. Rebuffat, C. Auvin, C. Goulard, Y. Prigent, B. Bodo, *J. Chem. Soc., Perkin Trans.* **1997**, *10*, 1587–1594. https://10.1039/A605629F.
[38] P. Toppo, R. Subba, K. Roy, S. Mukherjee, P. Mathur, *J. Plant Growth Regul.* **2023**, *42*, 1342–1363. https://doi.org/10.1007/s00344-022-10638-w.
[39] J. J. Doyle, J. L. Doyle, *Phytochem. Bull.* **1987**, *19*, 11–15.
[40] T. J. White, T. Bruns, S. Lee, J. W. Taylor, T. J. White, T. Bruns, S. Lee, J. W. Taylor, *PCR Protocols: A Guide to Methods and Applications*, Ed. M. A. Innis, D. H. Gelfand, J. J. Sninsky, Thomas J. White, Academic press, p. 315–322.
[41] Y. J. Liu, S. Whelen, B. D. Hall, *Mol. Biol. Evol.* **1999**, *16*, 1799–1808. https://doi.org/10.1093/oxfordjournals.molbev.a026092.
[42] K. Okonechnikov, O. Golosova, M. Fursov, U. team, *Bioinformatics* **2012**, *28*(8), 1166–1167. https://doi.org/10.1093/bioinformatics/bts091.
[43] S. Kumar, G. Stecher, K. Tamura, *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. https://doi.org/10.1093/molbev/msw054.
[44] A. Prjibelsk, D. Antipov, D. Meleshko, A. Lapidus, A. Korobeynikov, *Curr. Protoc. Bioinform.* **2020**, *70*, e102. https://doi.org/10.1002/cpbi.102.

Manuscript received: March 12, 2024
Version of record online: November 4, 2024