

Draft genome sequence of the peptaibol producer *Trichoderma asperellum* TRC-03

Pamela Alfaro-Vargas,¹ Daniela Wicki-Emmenegger,¹ Esteve Mesén-Porras,¹ Cristina Vargas-Chacón,² Oscar Bravo Bonilla,² Priscila Chaverri,^{1,3,4} Max Chavarria^{1,4,5}

AUTHOR AFFILIATIONS See affiliation list on p. 4.

ABSTRACT We present the genome sequence of *Trichoderma asperellum* TRC-03, an isolate obtained from an agricultural soil and whose capacity to produce peptaibols was previously demonstrated. The *Trichoderma* genome showed a size of 36,213,208 bp and a total of 55 metabolic genes related to peptaibol biosynthesis.

KEYWORDS *Trichoderma asperellum*, peptaibols, short-chain-length peptides, antifungal

Several fungal species, including those of the genus *Trichoderma*, have demonstrated the ability to produce peptaibols, short-chain-length peptides of 5–20 residues, most non-proteinogenic (1–3). The interest in these biomolecules lies in their different biological activities, including antifungal (4, 5), anticancer (6), anti-*Toxoplasma* (7), insecticidal (8), and antiplasmodial (9). Here, we report the genome of *Trichoderma asperellum* TRC-03, a strain whose peptaibol-rich extracts showed the ability to inhibit pathogens such as *Alternaria alternata*, *Botrytis cinerea*, *Colletotrichum gloeosporioides*, and *Fusarium oxysporum* that affect important crops (4). Strain TRC-03 was isolated from agricultural soil from Desamparados, San José, Costa Rica, using the procedure described by Cai and Druzhinina (10).

The isolate stored in glycerol 15% vol/vol at –80°C was grown in potato dextrose agar (7 days, 25°C). DNA extraction was performed using cetyltrimethylammonium bromide protocol (11) and sequenced through Oxford Nanopore (ONT) and Illumina NovaSeq technologies. For ONT sequencing, the library was prepared using the SQK-LSK109 Ligation Kit (ONT, UK). A total of 12 µL of the library was loaded onto a MinION flow cell and sequenced with a MinION MK1B device (ONT) for 72 hours, following the standard protocol (12). Raw data were base called with Dorado v0.8.1 (13) to convert POD5 to FASTQ. For Illumina sequencing, a paired-end library was prepared and loaded onto the NovaSeq 6000 (PE150). Raw data quality was assessed with FastQC v0.11.9 (14). Low-quality reads were filtered based on a Phred score of 20 for ONT and 28 for Illumina data. Adapter sequences were filtered using Trimmomatic v0.36C (15). A *de novo* hybrid assembly was performed with MaSuRCA v7.0.7 (16), and its quality was assessed with QUAST v5.2.1 (17). Genome annotations were obtained with BRAKER v3.06 (18) and compared to the *T. asperellum* reference CGF_020647865 using Companion v2.2.11 (19). Default parameters were used for all software. Genome assembly statistics and annotation features are presented in Table 1. A phylogenetic tree for translation elongation factor 1α (TEF-1α) was constructed with Sanger sequences using Phylosuite v1.2.3 (20). Sequences were aligned with MAFFT (21) and trimmed with TrimAl, and a concatenated tree was generated with IQ-TREE 2 v1.6.12 (22). Results (Fig. 1A) confirm that the TRC-03 strain is closely related to *T. asperellum* species. The biosynthetic gene clusters of TRC-03 and other peptaibol producers were analyzed with FunBGCeX v0.0.2 (23). A data frame was then constructed to identify genes involved in peptaibol

Editor André O. Hudson, Rochester Institute of Technology, Rochester, New York, USA

Address correspondence to Max Chavarria, max.chavarria@ucr.ac.cr.

The authors declare no conflict of interest.

See the funding table on p. 4.

Received 1 February 2026

Accepted 23 March 2026

Published 20 April 2026

Copyright © 2026 Alfaro-Vargas et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

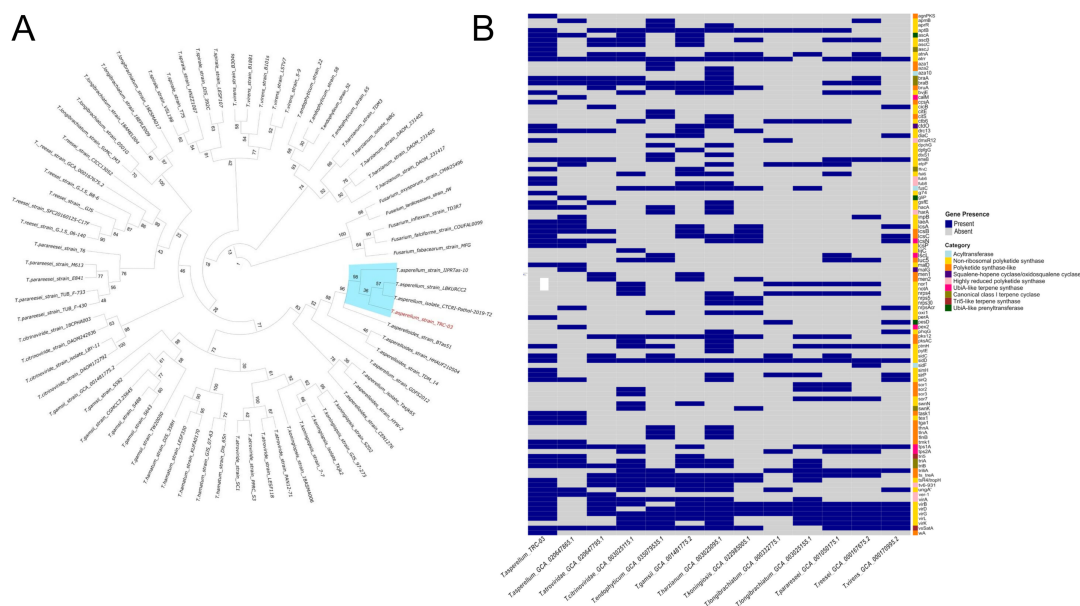


FIG 1 Phylogenetic tree and genes involved in peptaibol biosynthesis in *T. asperellum* TRC-03. (A) A fungal phylogenetic tree was constructed using the TEF-1 α gene, with the TRC-03 strain located within the *T. asperellum* clade (light blue shaded). The phylogenetic tree with ITS sequences (not shown) also groups strain TRC-03 within the *T. asperellum* clade. (B) Genes related to peptaibol biosynthesis identified in *T. asperellum* TRC-03 and other *Trichoderma* species. The color scale in the heatmap indicates the absence (gray) or presence (dark blue) of 144 distinct genes associated with peptaibol biosynthesis. The gene symbols (right Y-axis) are categorized into nine biosynthetic gene clusters (BGCs), each represented by a different color (legend on the right). Gene symbols and descriptions of the protein products were obtained from the manually curated fungal BGC database and the fungal genome mining tool, part of the FunBGCeX module for predicting novel BGCs. The following biosynthetic genes are part of or code for non-ribosomal peptide synthase (NRPS): *apmB* (NRPS L-2-aminopimelic acid B), *apirR* (short-chain dehydrogenase asperipin-2a biosynthesis cluster protein R), *aptB* (NRPS adenine phosphoribosyltransferase B), *ascB* (NRPS-like carboxylic acid reductase ascofuranone/ascochlorin B), *ascC* (NRPS-like carboxylic acid reductase ascofuranone/ascochlorin C), *atnA* (NRPS aspercryptin A1), *atrr* (NRPS-like glycine betaine reductase transcription factor ATRR), *bvjE* (NRPS-like brevijanazine E), *citE* (NRPS citrate synthase subunit E), *cicB* (NRPS-like carboxylic acid reductase cichorine B), *ctb6* (short-chain dehydrogenase cercosporin toxin biosynthesis 6), *drc13* (short-chain dehydrogenase DNA replication checkpoint protein 13), *diaC* (short-chain dehydrogenase dibenzodioxocinone C), *dpfgG* (NRPS DDP fungal gene cluster G), *dtxS1* (NRPS destruxin synthase 1), *emeB* (NRPS emericellamide B), *etpP* (NRPS epipolythiodioxopiperazine protein P), *fsl6* (NRPS-like carboxylic acid reductase fusarielin 6), *g74* (NRPS protein kinase G74), *gsfE* (short-chain dehydrogenase griseofulvin E), *hacA* (PKS harzianic acid A), *laeA* (uncharacterized NRPS putative methyltransferase), *lcsA* (NRPS leucinostatin A), *lcsP* (NRPS leucinostatin P), *lijC* (short-chain dehydrogenase lijiquinone C), *malD* (short-chain dehydrogenase malic enzyme D), *nor1* (short-chain dehydrogenase norsolorinic acid ketoreductase 1), *notA* (short-chain dehydrogenase notoamide A), *nrps4* (NRPS 4), *nrps5* (NRPS 5), *nrps30* (NRPS 30), *nrpsAcr* (NRPS acremopeptaibol), *oxi1* (short-chain dehydrogenase oxidative signal-inducible 1), *perA* (NRPS peramine synthesis), *phqG* (short-chain dehydrogenase paraherquamide G), *ptmH* (short-chain dehydrogenase platensimycin H), *pytE* (short-chain dehydrogenase pyranterreone E), *sidD* (NRPS siderophore D), *simH* (NRPS simvastatin H), *sirP* (NRPS sirodesmin P), *sirQ* (short-chain dehydrogenase sirodesmin Q), *sor7* (short-chain dehydrogenase sorbicillinol 7), *tex1* (short-chain dehydrogenase TRE X complex subunit 1), *tsR4/tropH* (NRPS aldehyde dehydrogenase temperature-sensitive ribosomal protein four and gene cluster for tropolone H), *ungA'* (NRPS unguisin A'), *virB* (short-chain dehydrogenase virulence factor B), *virD* (short-chain dehydrogenase virulence factor D), *virG* (short-chain dehydrogenase virulence factor G), *virK* (short-chain dehydrogenase virulence factor K), and *virL* (short-chain dehydrogenase virulence factor L). The following genes are part of or code for canonical class I terpene synthases: *ascJ* (short-chain dehydrogenase ascJ), *braA* (brasilane glycoside A synthase), *braB* (brasilane B glycosyltransferase), *ffnC* (terpene synthase fosfonochlorin C), *swnK* (swainsonine K synthase), *triA* (trichothecene A synthase), and *triB* (trichothecene B synthase). *atrr* codes for NRPS-like glycine betaine reductase ATRR. The following genes code for acyltransferases: *aza10* (acyltransferase azasperpyranone 10), *fusC* (acyltransferase fusidic acid C), and *sidF* (acyltransferase siderophore F). The following genes code for polyketide synthase-like enzymes (PKS-like): *agnPKS* (agnestin PKS), *aza1* (PKS azasperpyranone 1), *aza2* (non-reducing PKS azasperpyranone 2), *bruA* (PKS brunneoxanthone A), *ccsA* (PKS copper chaperone for superoxide dismutase A), *citS* (PKS citrate synthase subunit S), *lcsB* (PKS leucinostatin B), *lcsC* (PKS leucinostatin C), *luc5* (PKS lucilactaene 5), *men1* (short-chain dehydrogenase menisporopsin A), *men2* (PKS menisporopsin A), *pkS12* (PKS 12), *pkSAC* (PKS ascoclitine), *sor1* (PKS sorbicillinol 1), *sor2* (PKS sorbicillinol 2), *sor3* (PKS sorbicillinol 3), *task1* (PKS derived from *Trichoderma asperellum* specific mitogen-activated protein kinase 1), *tga1* (PKS derived from *Trichoderma viride* F-534 G alpha subunit gene), *thnA* (PKS tetrahydroxynaphthalene reductase A), *tnA* (PKS tricholignan A), *tnB* (PKS tricholignan B), *tmk1* (PKS derived from *Trichoderma* sp. mitogen-activated protein kinase 1), *triliA* (PKS *Trichoderma* sp. biosynthetic cluster tenellin synthase A), *ts_treA* (PKS derived from *Trichoderma* sp. periplasmic trehalase A), and *wA* (PKS for wall pigment secretion). The following genes are part of or code for HR PKS: *cmtA* (HR PKS copper metallothionein A), *dmxR12* (short-chain dehydrogenase dimeric xanthone biosynthesis cluster protein R12), *fub6* (HR PKS *Fusarium* sp. fusaric acid 6), *fub8* (HR PKS carboxylic acid reductase acid 8), *harA* (HR PKS harzianopyridone A), *scyPKS* (HR PKS scytalidin polyketide synthase), *tv6-931* (HR PKS from *Trichoderma virens* PKS cAT domain 6-931), *ver-1* (Continued on next page)

Fig 1 (Continued)

(short-chain dehydrogenase versicolorin A), and *virA* (HR PKS virulence factor A). The following three genes code for squalene-hopene cyclase/oxidosqualene cyclase: *ctdO* (short-chain dehydrogenase cycloaddition to form the bicyclo[2.2.2]diazaoctane O), *malG* (squalene-hopene cyclase transport of maltose G). These four genes code for ubiA-like prenyltransferase: *ascA* (Ubi-like prenyltransferase ascofuranone/ascochlorin A), *gliP* (ubiA-like gliotoxin biosynthesis protein P), *pesD* (UbiA-like prenyltransferase peptide synthase D). The following genes code for UbiA-like terpene synthases: *calM* (UbiA-like prenyltransferase calmodulin M), *lcsL* (UbiA-like terpene synthase leucinostatin L), *lcsN* (UbiA-like terpene synthase leucinostatin N), *pex2* (*Penicillium expansum* short-chain dehydrogenase peroxisomal membrane protein 2), *tps1A* (UbiA-like terpene prenyltransferase synthase 1A), and *tps2A* (UbiA-like terpene prenyltransferase synthase 2A). These two genes code for tri5-like terpene synthases: *tri5* (Tri5-like terpene synthase), *vsSatA*: (tri5-like terpene *Verticillium sakurae* streptothricin acetyltransferase A).

biosynthesis. A matrix for plotting a heatmap was built using *dplyr* v1.1.4 (24), *reshape* v1.4.4 (25), *tidyverse* 2.0.0 (26), and *pheatmap* v1.0.12 (27). Figure 1B presents a heatmap of genes involved in peptaibol synthesis across *Trichoderma* species. In the TRC-03 genome, 55 metabolic genes were predicted.

TABLE 1 Annotation and genome assembly statistics of the peptaibol producer *T. asperellum* TRC-03 isolated from a Costa Rican agricultural soil

Isolate feature	<i>Trichoderma asperellum</i> TRC-03
BioProject accession no.	PRJNA1172749
Assembly accession no.	JBLGTG000000000
No. of raw reads (Illumina)	1,403,177,100
No. of paired-end reads (Illumina)	4,677,257
No. of raw reads (ONT)	1,734,967,438
Genome size (bp)	36,213,208
No. of contigs	381
No. of predicted genes	10,529
Coverage (1x) (bp)	47.90
G + C content (%)	48.79
N ₅₀ (bp)	186,091
L ₅₀ (bp)	48
No. of tRNAs	171
No. of rRNAs	57
No. of ncRNAs	3
No. of snRNAs	12

AUTHOR AFFILIATIONS

¹Centro Nacional de Innovaciones Biotecnológicas (CENIBiot), CeNAT-CONARE, San José, Costa Rica

²Instituto Nacional de Innovación y Transferencia en Tecnología Agropecuaria (INTA), San José, Costa Rica

³Department of Natural Sciences, Bowie State University, Bowie, Maryland, USA

⁴Centro de Investigaciones en Productos Naturales (CIPRONA), Universidad de Costa Rica, San José, Costa Rica

⁵Escuela de Química, Universidad de Costa Rica, San José, Costa Rica

AUTHOR ORCID*s*

Max Chavarría  <http://orcid.org/0000-0001-5901-3576>

FUNDING

Funder	Grant(s)	Author(s)
Consejo Nacional para Investigaciones Científicas y Tecnológicas	FI-048B-19	Pamela Alfaro-Vargas
Centro Nacional de Innovaciones Biotecnológicas (CENIBiot, CeNAT-CONARE)		Max Chavarría

AUTHOR CONTRIBUTIONS

Pamela Alfaro-Vargas, Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Writing - original draft | Daniela Wicki-Emmenegger, Investigation, Methodology | Esteve Mesén-Porras, Data curation, Investigation, Software, Writing - original draft | Cristina Vargas-Chacón, Investigation | Oscar Bravo Bonilla, Investigation | Priscila Chaverri, Supervision, Validation, Writing – review and editing | Max Chavarría, Resources, Supervision, Writing – review and editing

DATA AVAILABILITY

The whole-genome sequencing of the TRC-03 strain was submitted to NCBI/GenBank under accession no. [JBLGTG0000000000](#), Bioproject no. [PRJNA1172749](#), and BioSample no. [SAMN44845004](#). The raw reads were deposited with SRA numbers [SRX26384498](#) (ONT) and [SRX26384497](#) (Illumina). The current version referred to in this paper is identified as version 1.0.

REFERENCES

- Komon-Zelazowska M, Neuhoof T, Dieckmann R, von Döhren H, Herrera-Estrella A, Kubicek CP, Druzhinina IS. 2007. Formation of atroviridin by *Hypocrea atroviridis* is conidiation associated and positively regulated by blue light and the G protein GNA3. *Eukaryot Cell* 6:2332–2342. <https://doi.org/10.1128/EC.00143-07>
- Du L, Risinger AL, Mitchell CA, You J, Stamps BW, Pan N, King JB, Bopassa JC, Judge SIV, Yang Z, Stevenson BS, Cichewicz RH. 2017. Unique amalgamation of primary and secondary structural elements transform peptaibols into potent bioactive cell-penetrating peptides. *Proc Natl Acad Sci USA* 114:E8957–E8966. <https://doi.org/10.1073/pnas.1707565114>
- Daniel JF de S, Rodrigues Filho E. 2007. Peptaibols of *Trichoderma*. *Nat Prod Rep* 24:1128. <https://doi.org/10.1039/b618086h>
- Alfaro-Vargas P, Bastos-Salas A, Muñoz-Arrieta R, Pereira-Reyes R, Redondo-Solano M, Fernández J, Mora-Villalobos A, López-Gómez JP. 2022. Peptaibol production and characterization from *Trichoderma asperellum* and their action as biofungicide. *JoF* 8:1037. <https://doi.org/10.3390/jof8101037>
- Bills G, Li Y, Chen L, Yue Q, Niu X-M, An Z. 2014. New insights into the echinocandins and other fungal non-ribosomal peptides and peptaibiotics. *Nat Prod Rep* 31:1348–1375. <https://doi.org/10.1039/c4np00046c>
- Gavryushina IA, Georgieva ML, Kuvarina AE, Sadykova VS. 2021. Peptaibols as potential antifungal and anticancer antibiotics: current and foreseeable development (review). *Appl Biochem Microbiol* 57:556–563. <https://doi.org/10.1134/S0003683821050070>
- Jiang T, Godinez-Macias KP, Collins JE, Lee JW, Wendt KL, Carolino K, Chakrabarti D, Cichewicz RH, Winzeler EA. 2024. Identification of fungal natural products with potent inhibition in *Toxoplasma gondii*. *Microbiol Spectr* 12:e0414223. <https://doi.org/10.1128/spectrum.04142-23>
- Bandani AR, Khambay BPS, Faull JL, Newton R, Deadman M, Butt TM. 2000. Production of efrapeptins by *Tolypocladium* species and evaluation of their insecticidal and antimicrobial properties. *Mycol Res* 104:537–544. <https://doi.org/10.1017/S0953756299001859>
- Lee JW, Collins JE, Wendt KL, Chakrabarti D, Cichewicz RH. 2021. Leveraging peptaibol biosynthetic promiscuity for next-generation antiparasitic therapeutics. *J Nat Prod* 84:503–517. <https://doi.org/10.1021/acs.jnatprod.0c01370>
- Cai F, Druzhinina IS. 2021. In honor of John Bissett: authoritative guidelines on molecular identification of *Trichoderma*. *Fungal Divers* 107:1–69. <https://doi.org/10.1007/s13225-020-00464-4>
- Anderson KL, Lebepe-Mazur S. 2003. Comparison of rapid methods for the extraction of bacterial DNA from colonic and caecal lumen contents

- of the pig. *J Appl Microbiol* 94:988–993. <https://doi.org/10.1046/j.1365-2672.2003.01917.x>
12. Lu H, Giordano F, Ning Z. 2016. Oxford Nanopore MinION sequencing and genome assembly. *Genomics Proteomics Bioinformatics* 14:265–279. <https://doi.org/10.1016/j.gpb.2016.05.004>
 13. Ahsan MU, Gouru A, Chan J, Zhou W, Wang K. 2024. A signal processing and deep learning framework for methylation detection using Oxford Nanopore sequencing. *Nat Commun* 15:1448. <https://doi.org/10.1038/s41467-024-45778-y>
 14. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Babraham Inst. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
 15. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
 16. Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. 2013. The MaSuRCA genome assembler. *Bioinformatics* 29:2669–2677. <https://doi.org/10.1093/bioinformatics/btt476>
 17. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>
 18. Hoff KJ, Lomsadze A, Borodovsky M, Stanke M. 2019. Whole-genome annotation with BRAKER, p 65–95. *In* Walker JM (ed), *Gene prediction. Methods in molecular biology*, 1962nd ed. Humana Press, New York, USA.
 19. Haese-Hill W, Crouch K, Otto TD. 2024. Annotation and visualization of parasite, fungi and arthropod genomes with Companion. *Nucleic Acids Res* 52:W39–W44. <https://doi.org/10.1093/nar/gkae378>
 20. Xiang C, Gao F, Jakovlić I, Lei H, Hu Y, Zhang H, Zou H, Wang G, Zhang D. 2023. Using PhyloSuite for molecular phylogeny and tree - based analyses. *iMeta* 2:e87. <https://doi.org/10.1002/imt2.87>
 21. Katoh K, Rozewicki J, Yamada KD. 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* 20:1160–1166. <https://doi.org/10.1093/bib/bbx108>
 22. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 37:1530–1534. <https://doi.org/10.1093/molbev/msaa015>
 23. Tang J, Matsuda Y. 2024. Discovery of fungal onoceroide triterpenoids through domainless enzyme-targeted global genome mining. *Nat Commun* 15:4312. <https://doi.org/10.1038/s41467-024-48771-7>
 24. Wickham H, François R, Henry L, Müller K. 2021. Dplyr: a grammar of data manipulation. 1.0.7. R package
 25. Wickham H. 2007. Reshaping data with the reshape package. *J Stat Softw* 21:20. <https://doi.org/10.18637/jss.v021.i12>
 26. Henry L, Wickham H. 2022. Rlang: functions for base types and core R and “tidyverse” features. 1.0.1. R package
 27. Kolde R, Kolde MR. 2019. Package “pheatmap”. 1.0.12. R package