



OPEN

DATA DESCRIPTOR

Chromosome-level genome assembly and annotation of *Anthurium amnicola*

Haomin Lyu^{1,2}, Sheina B. Sim¹, Scott M. Geib¹, Joanne S. L. Imamura³, Briette L. Corpuz¹, Renee L. Corpuz¹, Angela N. Kauwe¹, Tyler J. Simmonds¹, Claire N. Arakawa¹, Roxana Y. Myers¹, Lisa M. Keith¹, Qingyi Yu¹, Tracie K. Matsumoto¹, Teresita D. Amore³ & Jon Y. Suzuki¹✉

Anthurium amnicola is in the monocot family Araceae, subfamily Pothoideae and is a contributing species in Hawaii floriculture industry hybrids. To support future genetic improvements to this commodity, we sequenced and assembled the *A. amnicola* genome to chromosome-scale using PacBio HiFi and short-read Hi-C sequencing. A total of 98.51% of the 4.79 Gb genome is anchored into 15 chromosomes, with 99.2% gene completeness and a high LTR assembly index (LAI) score of 21.73, indicative of a complete, high-quality assembly. Annotation reveals the presence of 20,380 protein-coding genes, with 78.52% of the genome composed of repetitive sequences, predominantly long terminal repeat retrotransposons (LTR-RT). Phylogenetic analysis identified evolutionary relationships between *A. amnicola* and representative species in the Araceae and other plant families. This study provides the first reference genome sequence for the neotropical genus *Anthurium* and insights into Araceae evolution, benefiting the floriculture industry and evolutionary studies.

Background & Summary

Anthurium amnicola, first described in 1978, is a distinctive species belonging to the genus section *Calomystrium*, subsection *Rupicola*^{1,2}. This species is endemic to the stream-side habitats of Panama, where it thrives in the unique ecological niches these environments provide^{3,4}. The floral stem of this species consists of a spathe, a structure akin to a modified leaf and a spadix, an organ from which the female and male flowers arise; together these floral structures are characteristic of the family Araceae (Fig. 1a,b). These floral structures are not only visually striking but also serve as the basis for the species' inclusion in ornamental breeding programs^{5,6}.

The integration of *A. amnicola* by intergeneric hybridization has been instrumental in enhancing the aesthetic and horticultural traits of cultivated anthurium varieties. This effort has a rich history, particularly in Hawaii, where anthurium breeding has been a focal point for improving the crop's commercial value and global appeal^{7–11}. Research endeavors have spanned a range of disciplines, from genetics and the study of trait-associated genes to investigations into color^{7,12–23}, physiology^{24–28}, productivity²⁹, form³⁰, scent^{31–33}, nematode susceptibility^{34–38} and bacterial blight susceptibility and resistance^{39–42}, among others. Recent advancements in tissue culture propagation^{43–48} and biotechnology⁴⁹ have further supported the commercial expansion of anthurium cultivation.

Nuclear and chloroplast DNA sequences have proven invaluable for elucidating phylogenetic relationships among species in *Anthurium*, the largest genus in the Araceae family, and for understanding its relationship with other family members^{50–53}. Molecular marker sequences have also been useful for assessing the variedness and relationships of cultivars and species within germplasm collections^{54–56} and linkage maps of *Anthurium* have been produced^{57,58} for use in the mapping of genetic traits. However, until now a reference genome to clarify genome structure and content has been absent, limiting the depth of phylogenetic and evolutionary studies and breeding efforts.

Most commercial anthurium cultivars are complex hybrids or have uncertain lineage. *A. amnicola* was selected for whole-genome sequencing due to its close relation to commercial cultivars and its known genome

¹USDA ARS Daniel K. Inouye, U.S. Pacific Basin Agricultural Research Center, Hilo, Hawaii, 96720, USA. ²Hawaii Agriculture Research Center, Kunia, Hawaii, 96759, USA. ³Department of Tropical Plant and Soil Sciences, University of Hawaii at Manoa, Honolulu, Hawaii, 96822, USA. ✉e-mail: jon.suzuki@usda.gov

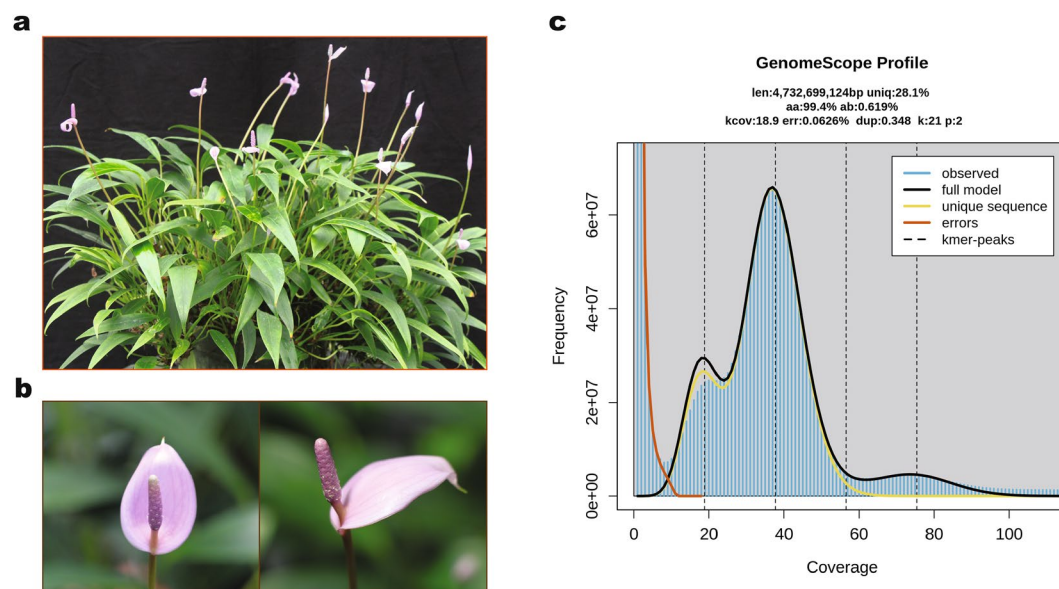


Fig. 1 *Anthurium amnicola* Dressler, physical and genomic features. **(a)** Whole plant. **(b)** Floral stem (spathe and spadix). **(c)** GenomeScope2 analysis based on K-mer data indicating an estimation of genome size and heterozygosity.

size and chromosome number^{5,59}. Unlike other species, *A. amnicola* produces numerous growing stems, from which flowers in sufficient quantities are produced for the study of floral traits⁶⁰. Furthermore, previous studies on tissue-specific expression were anticipated to enhance gene annotation¹⁶.

To provide a high-quality reference genome for this valuable species, a chromosome-level genome of *A. amnicola* was generated through PacBio HiFi reads, coupled with the Hi-C technology for chromosome assembly. We determined a primary assembly of 4.79 Gb genome, aligning with prior estimates ranging from 4.765 Gb to 5.287 Gb for the haploid genome⁵⁹ with a contig N50 of 26.98 Mb. The integration of Hi-C data with PacBio HiFi reads enabled the anchoring of 98.15% of the contig length into 15 pseudo-chromosomes, reflecting the cytological observation of a haploid chromosome count of 15 for *A. amnicola*⁵. The genome was annotated to include 20,380 protein-coding genes, with 78.52% of the genome attributed to transposable elements.

Methods

Sample collection, library construction and sequencing. A spathe from the floral stem of *A. amnicola* accession UH A667 was divided into four 50 mg samples (Fig. 1a,b). Samples were flash frozen in liquid nitrogen and processed on a MP FastPrep-24 classic (MP Biomedicals, Irvine, California, USA) at 4.5 m/s for 20 seconds in a 2 mL disruption tube (MP FastPrep Tubes) containing a metal grinding bead (BC Percision 5/32" 440 Stainless Steel Ball Bearings G100) and DNA was isolated using the PacBio Nanobind® Plant Nuclei Kit TR (Pacific Biosciences, Menlo Park, California, USA) (HMW DNA extraction from plant nuclei) following the manufacturers' instructions. A whole genome sequencing PacBio HiFi library was prepared using the SMRTbell® prep kit 3.0 (Pacific Biosciences, Menlo Park, California, USA) following the manufacturer's instructions. The final library was size selected using 35% v/v dilution of AMPure PB beads following the manufacturer's instructions to remove fragments shorter than 5 kb. The library sample was quantified using the Invitrogen Qubit dsDNA HS Assay Kit, and the size distribution of the library was measured and visualized using the Agilent Femto Pulse System (Agilent Technologies, Santa Clara, CA, USA).

The HiFi library was then sequenced using the PacBio Revio System, obtaining a total of 181.0 Gb HiFi reads, representing approximately 40 × coverage of the *A. amnicola* genome. We performed GenomeScope2 (v2.0.1) analysis on the HiFi reads⁶¹, leading to an estimation of a genome size of 4.73 Gb and heterozygosity of 0.62% (Fig. 1c).

For the Hi-C dataset, tissue was cross-linked, digested using the enzymes DdeI and DpnII, biotinylated, and proximity ligated. After proximity ligation and nucleic acid purification, DNA was sheared using the Diagenode Bioruptor Pico, enriched for proximity ligations using streptavidin beads, and size-selected to enrich for biotin labeled DNA fragments 200–600 bp. Following fragmentation and sample purification using a room temperature polyethylene glycol solution containing solid-phase reversible immobilization (SPRI) beads⁶², library preparation was performed using the NEB Next Ultra II DNA Library Prep Kit and followed by the Element Biosciences Adept Library Prep Kit. The resulting Hi-C library was sequenced on a partial 2 × 150 bp flow cell of an Element Biosciences Aviti system (Element Biosciences, San Diego, California, USA).

De novo assembly and Hi-C scaffolding. The PacBio HiFi reads were assembled using the Hifiasm (v0.19.5-r587)⁶³ with the default parameters. Duplicate haplotigs and erroneous assemblies were removed using the Purge_Dup (v1.2.6) pipeline⁶⁴. Raw Hi-C reads were aligned to the contig-level genome assembly using BWA-mem (v2.2.1)⁶⁵, with subsequent removal of PCR-derived duplicates. The program, YAHS (v1.2a.1)⁶⁶

Assembly characteristics	Values
Main genome scaffold total number	984
Main genome scaffold sequence total length	4789.83 Mb
Main genome scaffold N50	6
Main genome scaffold L50	306.43 Mb
Main genome contig N50	56
Main genome contig L50	26.98 Mb
BUSCO	99.2%
LAI	21.73

Table 1. Statistics of the *Anthurium amnicola* genome assembly and scaffolding.

was then used for scaffolding with default parameters. Assembly errors were manually adjusted using Juicebox (v1.11.08)⁶⁷, culminating in the resolution of 15 chromosomes.

The total length of the contig-level genome was 4.79 Gb (Table 1), approximately matching the genome size estimated by the flow cytometry method for this species, ranging from 4.77 to 5.29 Gb⁵⁹. The assembly comprised 1,392 contigs, achieving an L50 of 26.98 Mb and an N50 of 56. After scaffolding, a total of 4.72 Gb of the assembled contigs were anchored into 15 pseudo-chromosomes, spanning 98.51% of the final genome assembly (Fig. 2a,b; Table 1).

Annotation of repetitive sequences. To delineate the repetitive sequences across the *A. amnicola* genome, the EDTA pipeline for transposon prediction and annotation was employed using default parameters⁶⁸. This pipeline combined three types based on their underlying approach: homology-based, structural-based, and *de novo* methods. Our analysis revealed a total of 3.76 Gb of repetitive sequences, constituting 78.52% of the *A. amnicola* genome assembly (Fig. 2b; Table 2). The predominant class was long terminal repeat retrotransposons (LTR-RT), which made up 82.06% of the total repetitive sequences (Table 2). DNA transposons were also identified, comprising 2.99% of the entire genome (Table 2).

Prediction of gene structure and functional annotation. The protein-coding genes of the *A. amnicola* genome were predicted using a combination of three methods: *ab initio* gene prediction, homology-based gene prediction, and RNA-seq data gene prediction. Initially, RepeatMasker (v4.1.0)⁶⁹ was applied to annotate the genome-wide transposons, providing the masked genome sequences for gene annotation. For homology-based predictions, we used GeMoMa (v1.9)⁷⁰, aligning protein sequences from *Spirodela polyrhiza*⁷¹, *Lemna minor*⁷², *Zantedeschia elliottiana*⁷³, and *Colocasia esculenta*⁷⁴ to our genome assemblies. These four genomes exhibited high assembly quality and were publicly available to enable genome-wide comparisons, and thus chosen to represent main lineages of the Araceae (Table 3). Coding genes were predicted with GeMoMa (v1.9) using⁷⁰ default parameters. Transcriptome evidence was derived from genome-guided assembly of RNA-seq data accessible in the NCBI database (PRJNA288827)¹⁶. Hisat2 (v2.2.0)⁷⁵ was used to align the RNA-seq reads to the reference genomes, and TACO (v0.7.3)⁷⁶ was employed to reconstruct the full-length transcripts, and Transdecoder (v5.7.0) (<https://github.com/TransDecoder/TransDecoder>) identified the candidate coding regions. Augustus (v3.3.2)⁷⁷, GeneMark-ET (v4.59)⁷⁸ and Braker2 (v2.1.2)⁷⁹ were used to perform *ab initio* gene prediction with training based on mapping files of the RNA-seq data. Finally, EVidenceModeler (v2.0.0)⁸⁰ was employed to integrate results from all three methods to produce a final gene prediction model. After filtering out low quality predictions, a total of 20,380 predicted protein-coding genes in the *A. amnicola* genome was obtained (Fig. 2b).

Functional annotation was conducted using eggNOG-mapper (v5.0)⁸¹. The protein-coding genes were used to compare with the eggNOG database and generated functional annotations for gene sets of *A. amnicola*. According to the results, the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were assigned to 7,490 and 7,122 genes, respectively.

Phylogenetic tree construction. To trace the evolutionary relationships among Alismatales species, we collected 19 high-quality genomes from public databases, including Phytozome. Our dataset comprised eight Araceae species, five non-Araceae Alismatales species, and five additional eudicots and monocots, with *Nymphaea colorata* serving as the outgroup^{71–74,82–84}. The gene families across these 19 genomes were inferred using the OrthoFinder program (v2.5.5) with default parameters⁸⁵. The entire gene set from all 19 genomes were compared with the publicly available databases of Swiss-Prot and RefSeq non-redundant proteins using BLASTP program, and against the Dfam database using HMMER (v3.4)^{86,87}. When necessary, transposon protein genes could be removed from the gene set of any one of the 19 genomes in cases where they were not already removed prior to publication. Finally, the 20,380 protein-coding genes of *A. amnicola*, as well as those of the other 18 species, were assigned to 19,543 families, of which 305 were identified as single-copy gene families across all 19 species.

The genes representing the 305 single-copy gene families were then used to reconstruct the phylogenetic relationship among these species. First, the CDS sequences in each family were aligned based on protein sequence alignment using the program MUSCLE (v5.1.0) with default parameters⁸⁸. The jmodeltest2 (v2.1.10) was then used to detect the best substitution model for the 305 genes across 19 species⁸⁹, determining the optimal model of “GTR + I + G”. Subsequently, RAXML (v8.1.12) software was used to infer the phylogenetic tree with the parameters “-m GTRGAMMAI -f a” setting 1000 bootstraps and *N. colorata* as an outgroup⁹⁰.

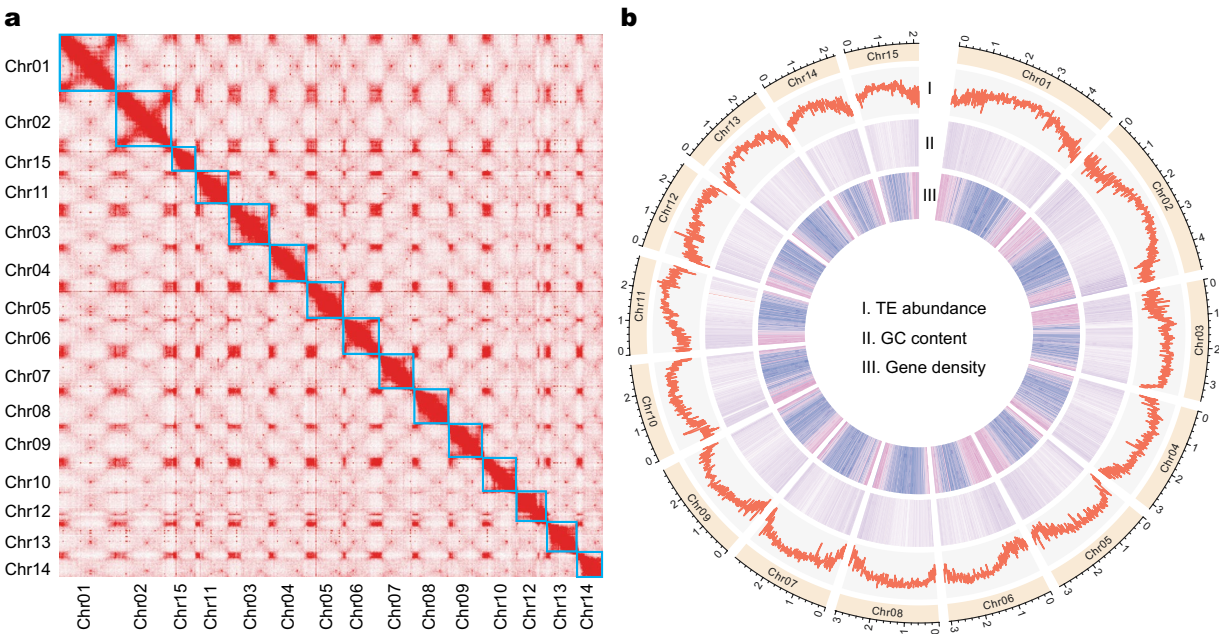


Fig. 2 Chromosome-level assembly and characteristics of the *Anthurium amnicola* genome. **(a)** Hi-C chromosomal scaffolding heatmap. **(b)** The landscape of the *A. amnicola* genome. The circus plot depicts the 15 identified chromosomes and their genomic features: (I) abundance of transposable elements, (II) GC content, and (III) gene density.

Class		bpMasked	%masked
LTR	Copia	375732402	7.84%
	Gypsy	1793785252	37.43%
	unknown	918175271	19.16%
TIR	CACTA	136682265	2.85%
	Mutator	113434220	2.37%
	PIF_Harbinger	35238676	0.74%
	Tc1_Mariner	12498453	0.26%
	hAT	64752156	1.35%
Non-LTR	LINE_element	9991839	0.21%
	unknown	404539	0.01%
Non-TIR	helitron	76003625	1.59%
repeat_region		225997332	4.72%
Total		3762696030	78.52%

Table 2. Information on the repeat annotation in the *Anthurium amnicola* genome assembly.

	<i>A. amnicola</i>	<i>C. esculenta</i>	<i>L. minor</i>	<i>S. polyrhiza</i>	<i>Z. Elliottiana</i>
Subfamily	Pothoideae	Aroideae	Lemnoideae	Lemnoideae	Philodendroideae
Genome size	4.79 Gb	2.41 Gb	472 Mb	158 Mb	1.15 Gb
Karyotype	2n = 2x = 30	2n = 2x = 28	2n = 2x = 40	2n = 2x = 64	2n = 2x = 32
Gene number	20,380	28,695	22,382	19,623	36,165
Repeat	78.52%	88.43%	61.50%	15.79%	60.18%

Table 3. Genomic information of select Araceae species.

Finally, the resulting phylogenetic relationship data together with the 305 gene sequences were used to infer the divergent time across the tree using MCMCtree implemented in PAML (v4.10.5)⁹¹. The divergent ages were calibrated using two time points, including the ancestor of *N. colorata* at 168~192 million years ago (MYA) and the divergent time between monocots and eudicots at 142~164 MYA⁹². Our analysis revealed the divergence age among different lineages of Alismatales (Fig. 3). Within the Araceae family, the aquatic lineage of

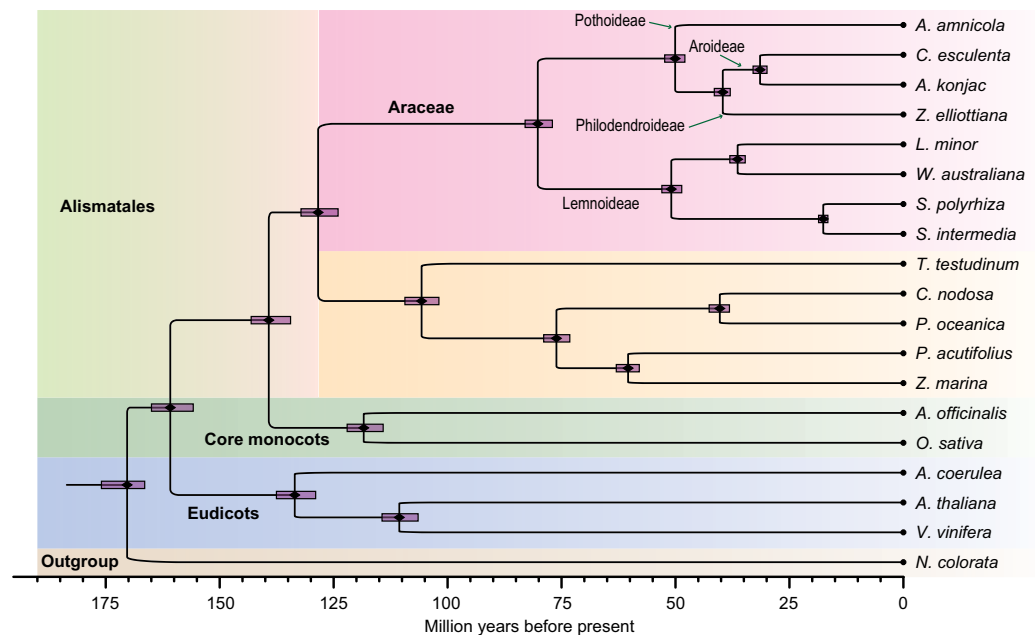


Fig. 3 Phylogenetic relationships across select species in the monocotyledon order Alismatales. The genomes of *Anthurium amnicola* and 18 other species were used to infer the origin and divergent ages of Alismatales species including those representing four subfamilies of the family Araceae. The boxes at each node represent 95% confidence intervals.

Lemnoideae originated around 80 MYA, while the first investigated subfamily Pothoideae, represented by *A. amnicola*, diverged around 50 MYA (Fig. 3).

Data Records

The PacBio HiFi long reads and Hi-C reads were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database under the SRA accession number SRP525474⁹³. The genome assembly has been deposited in GenBank with the accessions GCA_041870085.1⁹⁴. The genome annotation results have been deposited in the figshare database⁹⁵.

Technical Validation

The genome assembly quality was first evaluated using BUSCO (v5.2.2) and the orthologous database of embryophyta_odb10 (<https://busco-data.ezlab.org/v5/data/lineages/>), representing 1,614 single-copy genes conserved in embryophytes⁹⁶. The BUSCO analysis revealed that the *A. amnicola* genome assembly achieved a completeness of 99.2% at the chromosome level, with 86.3% being single-copy and 12.9% duplicated genes (Table 1). This assessment indicated a highly complete genome assembly of *A. amnicola* in this study.

The completeness of genome assembly was also assessed employing the LTR Assembly Index (LAI)⁹⁷, implemented in the LTR_retriever pipeline (v3.0.1)⁹⁸. Following the detection of LTR-RT candidate elements using LTR_finder (v1.5.1045) and LTR_harvest (v1.0646)^{99,100}, the LTR_retriever pipeline was applied to filter the confidential LTR-RTs, and assessing genome assembly quality. These analyses showed a general LAI value of 21.73, indicative of a high continuity of the chromosomal genome assembly for *A. amnicola* (Table 1).

Code availability

All software and pipelines were executed according to the manual and protocols of the published bioinformatics tools. The version and parameters of the software have been described in Methods. The pipelines for genome assembly and gene annotation were accessible in the GitHub repository (<https://github.com/HaominLyu/HiGenAA>).

Received: 3 October 2024; Accepted: 1 April 2025;

Published online: 10 April 2025

References

- Dressler, R. L. Two noteworthy ornamentals from the Caribbean coast of Panama. *Selbyana* **2**, 300–302 (1978).
- Dressler, R. L. A new name for the dwarf purple *Anthurium*. *Aroideana* **3**, 55 (1980).
- Croat, T. B., Whitehill, J. & Yates, E. D. A new subsection of *Anthurium* section *Calomystrum* (Araceae) and five new species from Colombia and Ecuador. *Aroideana* **30**, 23–37 (2007).
- Martel, C. & Croat, T. B. A new Species of *Anthurium* section *Calomystrum*, subsection *Rupicola* (Araceae) from Peru. *Phytotaxa* **585**, 225–232 (2023).
- Marutani, M., Wannakrairoj, S. & Kamemoto, H. Chromosome studies on *Anthurium amnicola* and its hybrids. *Aroideana* **11**, 7–14 (1988).

6. Wannakraioj, S. & Kamemoto, H. Inheritance of purple spathe in anthurium. *J. Am. Soc. Hort. Sci.* **115**, 169–171, <https://doi.org/10.21273/jashs.115.1.169> (1990).
7. Kamemoto, H., Iwata, R. Y. & Marutani, M. Genetics of the major spathe colors in anthuriums. *College of Tropical Agriculture and Human Resources Research Series* **56** (Univ. of Hawaii, 1988).
8. Kamemoto, H. & Kuehnle, A. *Breeding anthuriums in Hawaii*. (University of Hawaii Press, 1996).
9. Kamemoto, H. & Nakasone, H. Y. Evaluation and improvement of anthurium clones. *Hawaii Agricultural Experiment Station, TB* **58** (Univ. of Hawaii, 1963).
10. Sheffer, R. D. & Kamemoto, H. Chromosome numbers in the genus *Anthurium*. *Am. J. Bot.* **63**, 74–81, <https://doi.org/10.1002/j.1537-2197.1976.tb11787.x> (1976).
11. Higaki T., Lichty J. S. & Moniz D. Anthurium culture in Hawaii. *HITAHR Research Extension Series* **152** (Univ. of Hawaii, 1994).
12. Elibox, W. & Umaharan, P. Inheritance of major spathe colors in *Anthurium andraeanum* Hort. is determined by three major genes. *HortScience* **43**, 787–791 (2008).
13. Iwata, R. Y., Tang, C. S. & Kamemoto, H. Anthocyanins of *Anthurium andraeanum* Lind. *J. Am. Soc. Hort. Sci.* **104**, 464–466 (1979).
14. Clark, B. R., Bliss, B. J., Suzuki, J. Y. & Borris, R. P. Chemotaxonomy of Hawaiian *Anthurium* cultivars based on multivariate analysis of phenolic metabolites. *J. Agric. Food Chem.* **62**, 11323–11334, <https://doi.org/10.1021/jf502187c> (2014).
15. Collette, V. E., Jameson, P. E., Schwinn, K. E., Umaharan, P. & Davies, K. M. Temporal and spatial expression of flavonoid biosynthetic genes in flowers of *Anthurium andraeanum*. *Physiol. Plant.* **122**, 297–304, <https://doi.org/10.1111/j.1399-3054.2004.00402.x> (2004).
16. Suzuki, J. Y. *et al.* Organ-specific transcriptome profiling of metabolic and pigment biosynthesis pathways in the floral ornamental progenitor species *Anthurium amnicola* Dressler. *Sci. Rep.* **7**, 1596, <https://doi.org/10.1038/s41598-017-00808-2> (2017).
17. Li, J. *et al.* Identification of key genes responsible for green and white colored spathes in *Anthurium andraeanum* (Hort.). *Front. Plant Sci.* **14** (2023).
18. Wannakraioj, S. & Kamemoto, H. Histological distribution of anthocyanins in *Anthurium* spathes. *HortScience* **25**, 809–809, <https://doi.org/10.21273/hortsci.25.7.809> (1990).
19. Gopaulchan, D., Umaharan, P. & Lennon, A. M. A molecular assessment of the genetic model of spathe color inheritance in *Anthurium andraeanum* (Hort.). *Planta* **239**, 695–705, <https://doi.org/10.1007/s00425-013-2007-9> (2014).
20. Osorio-Guarín, J. A. *et al.* Comparative transcriptomic analysis reveals key components controlling spathe color in *Anthurium andraeanum* (Hort.). *PLOS ONE* **16**, e0261364, <https://doi.org/10.1371/journal.pone.0261364> (2021).
21. Li, C., Qiu, J., Huang, S., Yin, J. & Yang, G. AaMYB3 interacts with AaHLH1 to regulate proanthocyanidin accumulation in *Anthurium andraeanum* (Hort.) another strategy to modulate pigmentation. *Hort. Res.* **6**, 14, <https://doi.org/10.1038/s41438-018-0102-6> (2019).
22. Li, Z. *et al.* Transcriptome profiling in the spathe of *Anthurium andraeanum* ‘Alabama’ and its anthocyanin-loss mutant ‘Xueyu’. *Sci Data* **5**, 180247, <https://doi.org/10.1038/sdata.2018.247> (2018).
23. Iwata, R. Y., Tang, C. S. & Kamemoto, H. Concentration of anthocyanins affecting spathe color in anthuriums. *J. Amer. Soc. Hort. Sci.* **110**, 383–385 (1985).
24. Dou, N., Li, L., Fang, Y., Fan, S. & Wu, C. Comparative physiological and transcriptome analyses of tolerant and susceptible cultivars reveal the molecular mechanism of cold tolerance in *Anthurium andraeanum*. *Int. J. Mol. Sci.* **25**, 250 (2024).
25. Tian, D.-Q. *et al.* De novo characterization of the *Anthurium* transcriptome and analysis of its digital gene expression under cold stress. *BMC Genomics* **14**, 827, <https://doi.org/10.1186/1471-2164-14-827> (2013).
26. Dai, J. & Paull, R. E. Leaf development and anthurium flower growth. *HortScience* **25**, 1138e–1138, <https://doi.org/10.21273/HORTSCI.25.9.1138e> (1990).
27. Hayden, D. M. & Christopher, D. A. Characterization of senescence-associated gene expression and senescence-dependent and -independent cysteine proteases differing in microsomal processing in *Anthurium*. *Plant Sci.* **166**, 779–790, <https://doi.org/10.1016/j.plantsci.2003.11.019> (2004).
28. Perez, P. A. *Molecular genetic studies of senescence in anthurium* Ph.D. thesis, University of Hawaii at Manoa, (2012).
29. Elibox, W. & Umaharan, P. Genetic basis for productivity in *Anthurium andraeanum* Hort. *HortScience* **49**, 859–863, <https://doi.org/10.21273/hortsci.49.7.859> (2014).
30. Elibox, W. & Umaharan, P. A study of morphophysiological descriptors of cultivated *Anthurium andraeanum* Hort. *HortScience* **47**, 1234–1240, <https://doi.org/10.21273/hortsci.47.9.1234> (2012).
31. Kuanprasert, N., Kuehnle, A. R. & Tang, C. S. Floral fragrance compounds of some anthurium (Araceae) species and hybrids. *Phytochemistry* **49**, 521–528 (1998).
32. Kuanprasert, N. & Kuehnle, A. R. Fragrance quality, emission, and inheritance in *Anthurium* species and hybrids. *Aroideana* **22**, 48–62 (1999).
33. Wei, Q. *et al.* Profiling of volatile compounds and associated gene expression in two anthurium cultivars and their F1 hybrid progenies. *Molecules* **26** (2021).
34. Aragaki, M., Apt, W. J., Kunimoto, R. K., Ko, W. H. & Uchida, J. Y. Nature and control of Anthurium Decline. *Plant Dis.* **68**, 509–511, <https://doi.org/10.1094/PD-68-509> (1984).
35. Uchida, J. Y., Sipes, B. S., & Kadooka, C. Y. Burrowing nematode on anthurium: recognizing symptoms, understanding the pathogen and preventing disease. *University of Hawaii at Manoa, CTAHR Plant Dis.* PD 24. <http://hdl.handle.net/10125/12383> (Univ. of Hawaii, 2003).
36. Myers, R. *et al.* Yield Increases in burrowing nematode-infested anthurium with Fluopyram and Trifloxystrobin applications. *HortTechnology* **30**, 603–607, <https://doi.org/10.21273/horttech04648-20> (2020).
37. Wang, K. H., Kuehnle, A. R. & Sipes, B. S. *In vitro* tolerance and resistance to burrowing nematode, *Radopholus similis*, in *Anthurium* species. *Euphytica* **103**, 23–28, <https://doi.org/10.1023/A:1018384019938> (1998).
38. Wang, K.-H., Kuehnle, A. R. & Sipes, B. S. *In vitro* screening for burrowing nematode, *Radopholus citrophilus*, tolerance and resistance in commercial *Anthurium* hybrids. *In Vitro Cell. Dev. Biol. Plant* **33**, 205–208, <https://doi.org/10.1007/s11627-997-0023-9> (1997).
39. Alvarez, A. M., Toves, P. J., & Vowell, T. S. Bacterial blight of anthuriums: Hawaii’s experience with a global disease (APSnet Features). <http://172.16.1.61/online/feature/anthurium/print.asp> (apsnet.org) (2006).
40. Elibox, W. & Umaharan, P. Genetic basis of resistance to systemic infection by *Xanthomonas axonopodis* pv. *dieffenbachiae* in *Anthurium*. *Phytopathology* **98**, 421–426 (2008).
41. Elibox, W. & Umaharan, P. Inheritance of resistance to foliar infection by *Xanthomonas axonopodis* pv. *dieffenbachiae* in *Anthurium*. *Plant Dis.* **94**, 1243–1247 (2010).
42. Alvarez, A. M. in *Handbook of Florists’ Crops Diseases* (eds McGovern, R. J. & Elmer, W. H.) 283–316 (Springer International Publishing, 2018).
43. Uy, J. & Amore, T. D. A comparative analysis of *in vitro* responses of anthurium under temporary immersion (RITA®). *Acta Hort.* (2022).
44. Wang, G., Xu, C., Yan, S. & Xu, B. An efficient somatic embryo liquid culture system for potential use in large-scale and synchronic production of *Anthurium andraeanum* seedlings. *Front. Plant Sci.* **10** (2019).
45. Ramírez-Mosqueda, M. A., Cruz-Cruz, C. A., Cano-Ricárdez, A. & Bello-Bello, J. J. Assessment of different temporary immersion systems in the micropropagation of anthurium (*Anthurium andraeanum*). *3 Biotech* **9**, 307, <https://doi.org/10.1007/s13205-019-1833-2> (2019).

46. Martínez-Estrada, E., Islas-Luna, B., Pérez-Sato, J. A. & Bello-Bello, J. J. Temporary immersion improves *in vitro* multiplication and acclimatization of *Anthurium andreanum* Lind. *Sci. Hortic.* **249**, 185–191, <https://doi.org/10.1016/j.scienta.2019.01.053> (2019).
47. Uy, J. N. *Rapid in vitro multiplication of anthurium using temporary Immersion system (RITA®)* M.S. thesis, University of Hawaii at Manoa, (2020).
48. Tanouye, M. A. R. *Improved multiplication of anthurium using the WeVitro system* M.S. thesis, University of Hawaii at Manoa, (2022).
49. Teixeira da Silva, J. A. *et al.* Genetic transformation and molecular research in *Anthurium*: progress and prospects. *Plant Cell Tiss. Org. Cult. (PCTOC)* **123**, 205–219, <https://doi.org/10.1007/s11240-015-0832-1> (2015).
50. Carlsen, M. M. & Croat, T. B. A molecular phylogeny of the species-rich neotropical genus *Anthurium* (Araceae) based on Combined chloroplast and nuclear DNA. *Syst. Bot.* **38**, 576–588, <https://doi.org/10.1600/036364413X670287> (2013).
51. Henriquez, C. L., Arias, T., Pires, J. C., Croat, T. B. & Schaal, B. A. Phylogenomics of the plant family Araceae. *Mol. Phylogen. Evol.* **75**, 91–102, <https://doi.org/10.1016/j.ympev.2014.02.017> (2014).
52. Carlsen, M. M. & Croat, T. B. An analysis of the sectional classification of *Anthurium* (Araceae): Comparing infrageneric groupings and their diagnostic morphology with a molecular phylogeny of the genus. *Ann. Mo. Bot. Gard.* **104**, 69–82, <https://doi.org/10.3417/2018215> (2019).
53. Abdullah *et al.* Complete chloroplast genomes of *Anthurium huixtlense* and *Pothos scandens* (Pothoideae, Araceae): Unique inverted repeat expansion and contraction affect rate of evolution. *J. Mol. Evol.* **88**, 562–574, <https://doi.org/10.1007/s00239-020-09958-w> (2020).
54. Saidi, A., Daneshvar, Z. & Hajibarat, Z. Comparison of Genetic Variation of Anthurium (*Anthurium andraeanum*) Cultivars Using SCOT, CDDP and RAPD Markers. *Plant Tissue Cult. Biotechnol.* **28**, 171–182, <https://doi.org/10.3329/ptcb.v28i2.39676> (2018).
55. Ge, Y. *et al.* Genetic variations within a collection of anthuriums unraveled by morphological traits and AFLP markers. *Biochem. Syst. Ecol.* **45**, 34–40, <https://doi.org/10.1016/j.bse.2012.07.023> (2012).
56. Suzuki, J. Y. *et al.* Development of chloroplast single nucleotide polymorphisms (SNPs) as a tool towards interspecies typing of *Anthurium* germplasm. *ActaHortic.* **1167**, 257–270, <https://doi.org/10.17660/ActaHortic.2017.1167.39> (2017).
57. Zhang, L. *et al.* A high-resolution linkage map construction and QTL analysis for morphological traits in anthurium (*Anthurium andraeanum* Linden). *Plants* **12**, 4185 (2023).
58. Venkat, S. K., Bommisetty, P., Patil, M. S., Reddy, L. & Chennareddy, A. The genetic linkage maps of Anthurium species based on RAPD, ISSR and SRAP markers. *Sci. Hortic.* **178**, 132–137, <https://doi.org/10.1016/j.scienta.2014.08.017> (2014).
59. Bliss, B. J. & Suzuki, J. Y. Genome size in *Anthurium* evaluated in the context of karyotypes and phenotypes. *AoB PLANTS* **2012** <https://doi.org/10.1093/aobpla/pls006> (2012).
60. Suzuki, J. Y., Horrell, J. R., Imamura, J. L., Amore, T. D. & Matsumoto, T. K. Floral developmental timing in the ornamental progenitor species *Anthurium amnicola* Dressler. *Aroidiana* **40**, 40–48 (2017).
61. Ranallo-Benavidez, T. R., Jaron, K. S. & Schatz, M. C. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat. Commun.* **11**, 1432, <https://doi.org/10.1038/s41467-020-14998-3> (2020).
62. DeAngelis, M. M., Wang, D. G. & Hawkins, T. L. Solid-phase reversible immobilization for the isolation of PCR products. *Nucleic Acids Res.* **23**, 4742–4743, <https://doi.org/10.1093/nar/23.22.4742> (1995).
63. Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. & Li, H. Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170–175, <https://doi.org/10.1038/s41592-020-01056-5> (2021).
64. Guan, D. *et al.* Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics* **36**, 2896–2898, <https://doi.org/10.1093/bioinformatics/btaa025> (2020).
65. Vasimuddin, M., Misra, S., Li, H. & Aluru, S. Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. IEEE Parallel and Distributed Processing Symposium (IPDPS) <https://doi.org/10.1109/IPDPS.2019.00041> (2019).
66. Zhou, C., McCarthy, S. A. & Durbin, R. YaHS: yet another Hi-C scaffolding tool. *Bioinformatics* **39** <https://doi.org/10.1093/bioinformatics/btac808> (2022).
67. Durand, N. C. *et al.* Juicer provides a one-click system for analyzing Loop-Resolution Hi-C experiments. *Cell Syst.* **3**, 95–98, <https://doi.org/10.1016/j.cels.2016.07.002> (2016).
68. Ou, S. *et al.* Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol.* **20**, 275–275, <https://doi.org/10.1186/s13059-019-1905-y> (2019).
69. Tarailo-Graovac, M. & Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinformatics* **4**, 4.10–14.10 (2009).
70. Keilwagen, J., Hartung, F. & Grau, J. in *Gene Prediction: Methods and Protocols* (ed Kollmar, M.) 161–177 (Springer New York, 2019).
71. Wang, W. *et al.* The *Spirodela polyrrhiza* genome reveals insights into its neotenus reduction fast growth and aquatic lifestyle. *Nat. Commun.* **5**, 3311, <https://doi.org/10.1038/ncomms4311> (2014).
72. Van Hoeck, A. *et al.* The first draft genome of the aquatic model plant *Lemna minor* opens the route for future stress physiology research and biotechnological applications. *Biotechnol. Biofuels* **8**, 188, <https://doi.org/10.1186/s13068-015-0381-1> (2015).
73. Wang, Y. *et al.* Chromosome level genome assembly of colored calla lily (*Zantedeschia elliottiana*). *Sci. Data* **10**, 605, <https://doi.org/10.1038/s41597-023-02516-1> (2023).
74. Yin, J. *et al.* A high-quality genome of taro (*Colocasia esculenta* (L.) Schott), one of the world's oldest crops. *Mol. Ecol. Resour.* **21**, 68–77, <https://doi.org/10.1111/1755-0998.13239> (2021).
75. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
76. Niknafs, Y. S., Pandian, B., Iyer, H. K., Chinnaiyan, A. M. & Iyer, M. K. TACO produces robust multisample transcriptome assemblies from RNA-seq. *Nat. Methods* **14**, 68–70, <https://doi.org/10.1038/nmeth.4078> (2017).
77. Stanke, M. *et al.* AUGUSTUS: *ab initio* prediction of alternative transcripts. *Nucleic Acids Res.* **34**, W435–439 (2006).
78. Besemer, J. & Borodovsky, M. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gki487> (2005).
79. Brůna, T., Hoff, K. J., Lomsadze, A., Stanke, M. & Borodovsky, M. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genom. Bioinform.* **3**, lqaa108, <https://doi.org/10.1093/nargab/lqaa108> (2021).
80. Haas, B. J. *et al.* Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biol.* **9**, R7, <https://doi.org/10.1186/gb-2008-9-1-r7> (2008).
81. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas, J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol. Biol. Evol.* **38**, 5825–5829, <https://doi.org/10.1093/molbev/msab293> (2021).
82. Ma, X. *et al.* Seagrass genomes reveal ancient polyploidy and adaptations to the marine environment. *Nat. Plants* **10**, 240–255, <https://doi.org/10.1038/s41477-023-01608-5> (2024).
83. Zhang, L. *et al.* The water lily genome and the early evolution of flowering plants. *Nature* **577**, 79–84, <https://doi.org/10.1038/s41586-019-1852-5> (2020).
84. Gao, Y. *et al.* A chromosome-level genome assembly of *Amorphophallus konjac* provides insights into konjac glucomannan biosynthesis. *Comput. Struct. Biotechnol. J.* **20**, 1002–1011, <https://doi.org/10.1016/j.csbj.2022.02.009> (2022).

85. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238, <https://doi.org/10.1186/s13059-019-1832-y> (2019).
86. Potter, S. C. *et al.* HMMER web server: 2018 update. *Nucleic Acids Res.* **46**, W200–W204, <https://doi.org/10.1093/nar/gky448> (2018).
87. Hubley, R. *et al.* The Dfam database of repetitive DNA families. *Nucleic Acids Res.* **44**, D81–D89, <https://doi.org/10.1093/nar/gkv1272> (2016).
88. Edgar, R. C. Muscle5: high-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nat. Commun.* **13**, 6968, <https://doi.org/10.1038/s41467-022-34630-w> (2022).
89. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* **9**, 772–772, <https://doi.org/10.1038/nmeth.2109> (2012).
90. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313, <https://doi.org/10.1093/bioinformatics/btu033> (2014).
91. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591, <https://doi.org/10.1093/molbev/msm088> (2007).
92. Kumar, S. *et al.* TimeTree 5: an expanded resource for species divergence times. *Mol. Biol. Evol.* **39**, msac174, <https://doi.org/10.1093/molbev/msac174> (2022).
93. NCBI Sequence Read Archive <https://identifiers.org/ncbi/insdc.sra:SRP525474> (2024).
94. Lyu, H. *et al.* Genome Assembly of Anthurium amnicola. *GenBank* <https://identifiers.org/ncbi/insdc:JBGNU1000000000.1> (2024).
95. Lyu, H. Chromosome-level genome assembly and annotation of Anthurium amnicola. *figshare. Dataset.* <https://doi.org/10.6084/m9.figshare.26778256.v1> (2024).
96. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212, <https://doi.org/10.1093/bioinformatics/btv351> (2015).
97. Ou, S., Chen, J. & Jiang, N. Assessing genome assembly quality using the LTR Assembly Index (LAI). *Nucleic Acids Res.* **46**, e126–e126, <https://doi.org/10.1093/nar/gky730> (2018).
98. Ou, S. & Jiang, N. LTR_retriever: a highly accurate and sensitive program for identification of long terminal repeat retrotransposons. *Plant Physiol.* **176**, 1410–1422, <https://doi.org/10.1104/pp.17.01310> (2018).
99. Ellinghaus, D., Kurtz, S. & Willhoeft, U. LTRharvest, an efficient and flexible software for *de novo* detection of LTR retrotransposons. *BMC Bioinform.* **9**, 18, <https://doi.org/10.1186/1471-2105-9-18> (2008).
100. Xu, Z. & Wang, H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **35**, W265–W268, <https://doi.org/10.1093/nar/gkm28> (2007).

Acknowledgements

This work was supported by the USDA Agricultural Research Service CRIS Project No. 2040-21000-019-000D.

Author contributions

H.L. performed annotation, comparative genome analysis and interpretation, S.B.S. and S.M.G. performed data curation and chromosome level assembly, S.M.G., R.L.C., T.J.S. and A.N.K. produced PacBio HiFi and Hi C sequences, B.L.C. performed DNA isolation, J.S.L.I., T.D.A. and C.N.A. provided tissue through clonal propagation of *A. amnicola* accession A667 in the University of Hawaii and USDA ARS DKU U.S. PBARC anthurium germplasm collections, Q.Y. provided expertise on plant whole genome analysis, H.L. wrote the original manuscript with contributions from S.B.S. and J.Y.S.; H.L., S.B.S., J.Y.S., T.D.A., T.K.M., Q.Y., L.M.K., R.Y.M., B.L.C., R.L.C., A.N.K., T.J.S. and J.S.L.I. contributed to and edited drafts, J.Y.S. and T.D.A. conceived the project, J.Y.S. supervised the project.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.Y.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2025