

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | | |
|-------------------------------------|--|
| n/a | Confirmed |
| ☐ | ☒ The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☒ | ☐ The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☒ | ☐ A description of all covariates tested |
| ☒ | ☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☒ | ☐ For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted
<i>Give P values as exact values whenever suitable.</i> |
| ☐ | ☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	The genome sequencing was then performed by Personal Biotechnology Co., Ltd (Shanghai, China) using the Illumina Novaseq platform. An average of 10 Gb of raw reads (paired-end, 150 bp each) was produced from each DNA sample through metagenomic sequencing or 30 Gb (5x coverage) through whole-genome resequencing.
Data analysis	AdapterRemoval v2.3.1, SOAPec v2.01, SPAdes v3.7.1, Megahit v1.2.9, NOVOPlasty v4.3.1, blastn 2.15.0+, tblastn 2.15.0+, BBTools v38.78, Tablet v1.20.12.24, Prokka v1.14.6, Pseudofinder v1.1.0, BUSCO v5.4.7, MAFFT v7.450, TrimAl v1.4.1, FASConCAT-g v1.04, PhyKIT v1.0.0, ModelFinder v2.2.2, IQ-TREE v1.6.8, MrBayes v3.2.6, PhyloSuite v1.2.3, iTOL v6, MCMCTree (implemented in paml v4.10.7), Tracer v1.7.2, FigTree v1.4.4, IGV v2.19.2 and Mauve (available at: http://www.darlinglab.org/mauve/mauve.html).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The data supporting the findings of this work are available within the paper and its Supplementary Information files. The sequences obtained in this study have been deposited in the NCBI Nucleotide database under the following accession numbers: host cicadas: (cox1: PQ166233–PQ166291; cox2: PQ189168–PQ189227; cob: PQ189228–PQ189282; mitogenomes: PQ328616, PQ619044–PQ619059 and PQ619061–PQ619081); Karelulcia: (16S rRNA: PQ165220–PQ165279; 23S rRNA: PQ165176–PQ165218; genomes: JBHUP001, JBHUP001, CP168741, CP169620–CP169629 and CP170724–CP170733); Hodgkinia: (16S rRNA: PQ169576–PQ169607 and PV590188–PV590190; rpoB: PQ278174–PQ278192; genomes: CP178246, CP178247, JBHERG01, JBHERH01 and JBJOXD01–JBJOXK01); YLS: (18S rRNA: PQ203679–PQ203715; ITS: PQ219292–PQ219308; 28S rRNA: PQ165280–PQ165296; RPB1: PQ202053–PQ202094; RPB2: PQ206318–PQ206351; Ef-1 α : PQ202095–PQ202135; mitogenomes: PV563784–PV563790, PV550736–PV550738, PV562129–PV562134, PV551120–PV551124 and PV591855); Arsenophonus genomes: JBHEQS01–JBHEQW01; Wolbachia genomes: JBHEQX01–JBHERD01; Rickettsia genome: JBHEQR01 and Tisiphia genomes: JBHERE01 and JBHERF01. The other molecular sequences were downloaded from NCBI database. The Illumina DNA-seq data and genome assemblies generated in this study have been deposited in the NCBI database under BioProject accession numbers PRJNA1146783 and PRJNA1164630. Source data are provided with this paper.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	This study does not involve human participants or human data.
Reporting on race, ethnicity, or other socially relevant groupings	This study does not involve human participants or human data.
Population characteristics	This study does not involve human participants or human data.
Recruitment	This study does not involve human participants or human data.
Ethics oversight	This study does not involve human participants, their data, or biological material.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We reconstructed patterns and consequences of bacteriome- and fat-body-associated obligate endosymbiont loss, as well as concordant gains of YLS in 133 cicada species using microscopy, comparative genomics and cophylogenetic analyses.
Research sample	We collected 65 cicada species (a total of 236 individuals used in this study) during the adult emergence period between 2021 and 2024, and detailed sample information can be found in Supplementary Table 1. Species identification was performed using morphological characteristics and DNA barcoding (cox1 gene). The remaining data were obtained from previous studies.
Sampling strategy	Among the 65 cicada species collected by authors of this study, all species with fewer than three individuals were fully sampled, and for those with more than three individuals, three to eight individuals were randomly selected for this study (except for the cicada <i>Chremistica ochracea</i>). For <i>Chremistica ochracea</i> , to determine the retention rate of <i>Hodgkinia</i> at the population level, 30 individuals were randomly selected for study. Typically, the types and distribution of symbionts are consistent across different individuals of the same host cicada species, so the sample size is sufficient.
Data collection	Jinrui Zhou, Qiong Guo, Xiaohong Han, Wenzhe Zhang and Zhi Huang collected 65 cicada species during their adult emergence period between 2021 and 2024. At the same time, Jinrui Zhou, Qiong Guo and Xiaohong Han collected molecular data of these cicadas and their symbionts using PCR amplification and RFLP-based techniques. They also collected images of the symbiont distribution using histological and fluorescence microscopy. Jinrui Zhou used SPAdes v3.7.1, Megahit v1.2.9 and/or NOVOPlasty v4.3.1 to assemble the genomes of cicadas and their symbionts between 2023 and 2024. Jinrui Zhou used SPAdes v3.7.1 to assemble the mitogenomes of fungal symbionts in April 2025. All the authors collected relevant data from previous studies between 2023 and 2024.

Timing and spatial scale	We continuously collected data since 2021. All the 65 cicada species collected by the authors occurred mainly in China and adjacent countries. We also obtained relevant data of the other cicada species from previous studies, which occurred in various countries. The data sources for this study are worldwide.
Data exclusions	No data were excluded from the analyses.
Reproducibility	The results of each experiment were reproducible. For example, we tested different individuals of the same cicada species using various methods and obtained consistent, reliable and reproducible results. Though the retention rate of <i>Hodgkinia</i> was low in <i>Chremistica ochracea</i> at the population level, we simultaneously detected the existence of <i>Hodgkinia</i> using different methods.
Randomization	The samples for the study were randomly selected.
Blinding	In this study, random sampling was adopted. Different methods were conducted independently.
Did the study involve field work?	☒ Yes ☐ No

Field work, collection and transport

Field conditions	Samples were collected from various natural habitats, including mountains, basins, plains, and plateaus. Sampling was conducted since 2021 to 2024, with the majority of samples gathered from February to September, 2023. The temperature at sampling sites ranged from approximately 10 to 40 degrees Celsius.
Location	Samples were collected from Shaanxi, Hubei, Sichuan, Guizhou, Yunnan, Guangdong, Hainan, Fujian provinces, Guangxi Zhuang Autonomous Region, and Xinjiang Autonomous Region, China; details are provided in Supplementary Table 1.
Access & import/export	No specimens collected are listed as "three-values animals" (i.e., terrestrial wild animals of important ecological, scientific, and social value) or nationally protected species.
Disturbance	All collected specimens are not included in the list of terrestrial wild animals under national protection with important ecological, scientific, and social value. Additionally, by adhering to the principle of "minimal necessary sampling", the samples collected in this study had no impact on natural populations or ecosystems.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

Methods

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other research organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals	The study did not involve laboratory animals.
Wild animals	Cicadas were collected using daytime netting and nighttime light trapping methods. Some of the collected individuals were dissected to obtain fresh tissues in the field, while the remaining specimens were separately preserved in absolute ethanol and/or made into dry specimens following standard entomological protocols.
Reporting on sex	The results of this study are applicable to both female and male cicadas. The gender of the cicadas did not affect the outcomes of this study. The sampling in this study was randomly selected, and both female and male cicadas were included in the samples.
Field-collected samples	Before dissection, cicadas were anesthetized at 0 °C for 5 min (anesthesia was omitted when conditions did not permit), externally sterilized with 75% ethanol, and then rinsed with sterile water 5 times. Dissection was performed under sterile conditions using a stereoscopic zoom microscope (Motic SMZ168, China) with 0.1 M phosphate-buffered saline (PBS, pH 7.2). Sterile forceps were used

to isolate intact bacteriomes, fat bodies, testes and ovaries from adult cicadas. Bacteriomes and fat bodies were observed using a stereomicroscope (Olympus SZX10, Japan) and photographed with a digital camera (Canon 550D, Japan). The dissected tissues were preserved in the corresponding reagents (2.5% glutaraldehyde fixative for transmission electron microscope (TEM) sample preparation; 4% paraformaldehyde or absolute ethanol for fluorescence in situ hybridization (FISH); absolute ethanol for DNA extraction) at 4 °C for subsequent TEM, FISH or molecular analyses.

Ethics oversight

None were required, because all collected specimens are not on the list of terrestrial wild animals under national protection that have important ecological, scientific, and social value.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Plants

Seed stocks

This study does not involve seed stocks.

Novel plant genotypes

This study does not involve novel plant genotypes.

Authentication

Not applicable. This study does not involve plant-related experiments.