



Uncertainties in the phylogeny and biogeography of cave crickets



Yehao Wang¹, Shiyu Du², Michael S. Engel^{3,4,5} & Chenyang Cai¹✉

ARISING FROM D.-Y. Kim et al. *Communications Biology* <https://doi.org/10.1038/s42003-024-06068-x> (2024)

Rhaphidophoridae (cave crickets) are a captivating group of wingless crickets that comprises nine extant subfamilies, with a distribution spanning the globe, excluding Antarctica¹. Due to their limited mobility, each rhaphidophorid subfamily exhibits a high degree of endemism¹. Elucidating robust phylogenetic relationships among these subfamilies will greatly enhance our understanding of the evolutionary and biogeographic history of this intriguing group. In a recent study by Kim et al. [KIM24]², a Sanger sequencing-based phylogeny of Rhaphidophoridae was presented, encompassing all extant subfamilies. Notably, the enigmatic subfamily Anoplophilinae was included for the first time, alongside Tropidischiniae and Gammarotettiginae². Based on their phylogenetic analysis, divergence-time estimation, and deep-time biogeographic reconstruction, two key insights emerged: (1) Anoplophilinae are the sister group to Gammarotettiginae; and (2) Geological events coincide with lineage diversification events and potentially promoted rhaphidophorid divergence, such as the origin of Gammarotettiginae in East Asia and its dispersal to western North America via the Beringia land bridge, and the separation of Ceuthophilinae from Tropidischiniae in North America influenced by the opening of the Western Interior Seaway². However, upon reanalyzing the original datasets from KIM24 using both maximum likelihood (ML) and Bayesian inference (BI) methods, we find that the reproducibility of a consistent topology is not guaranteed based on repeated experiments, due to weak phylogenetic signals. We also show that the tree topology may change significantly based on a better-fitting site-heterogeneous model. Consequently, the phylogenetic relationships within Rhaphidophoridae proposed by KIM24 are not robust, and the downstream timetree and biogeographic reconstruction are challenged by the topological uncertainty.

In the analysis of the 110-taxon matrix, we observed the same topology at the subfamilial level as produced by KIM24 in repeats 4 and 5 under the partitioned method, and repeats 1 to 3 under the GTR-restricted partitioned method (Supplementary Figs. 8, 9, and 11–13). However, in most cases, regardless of the partitioned or other methods, the reproducibility of KIM24's results were limited (Supplementary Figs. 1–7, 10, 14, and 15). In our results, the monophyly of rhaphidophorid subfamilies was well supported in both ML and BI analyses, except for the undersampled subfamilies, such as Gammarotettiginae, which included only one species, and

the controversial Anoplophilinae, which excluded the genus *Alpinoplophilus* (Fig. 1a).

KIM24 suggested that the most recent common ancestor of the Rhaphidophoridae diverged into two main lineages². Furthermore, the study revealed that the three Asian subfamilies did not form a monophyletic group². Instead, the Asian subfamily Anoplophilinae was found to be grouped with Gammarotettiginae from the west coast of North America, indicating a more complex evolutionary relationship². However, this relationship was unstable, as it could not be reproduced by their molecular clock dating tree, and which was why they used a manipulated topology with the node constrained². In our analysis, the clade (Anoplophilinae + Gammarotettiginae) occurred under partitioned methods but consistently with low support (Supplementary Figs. 4–15). Surprisingly, the mysterious Gammarotettiginae emerged as the first-diverging lineage, strongly supported by BI analyses under both CAT-GTR + G4 and GTR models, albeit with relatively weak support from the ML analysis under the GTR + F + I + G4 model (Supplementary Figs. 1–3). Meanwhile, the three Asian subfamilies were clustered into a monophyletic group (Fig. 1a). Within the Asian lineages, Anoplophilinae diverged first as the sister group to (Rhaphidophorinae + Aemodogryllinae) (Fig. 1a). Our results suggest that the straight dorsal profile and denticles at the apex of the upper margin of the upper valve of the ovipositor could be symplesiomorphies.

In KIM24, Macropathinae were identified as the sister group to ((Gammarotettiginae + Anoplophilinae) + (Rhaphidophorinae + Aemodogryllinae))². However, our analysis supported another topology, suggesting that Macropathinae may also be the sister group to the clade containing all four lineages from North America and the Mediterranean Region, albeit with generally low support (Fig. 1a). Consequently, this underscores the need to maintain skepticism regarding the precise phylogenetic position of Macropathinae. Within Macropathinae, taxa from Australia were not monophyletic, while the phylogenetic placement of *Micropathus* remains unresolved, consistent with the findings of KIM24. However, diverging from the conclusions of KIM24, taxa from South America were also not monophyletic under the CAT-GTR + G4 model (Fig. 1a).

In KIM24, Troglophilinae were inferred as the sister group to Dolichopodainae from the Mediterranean region in the ML analysis, while in the

¹State Key Laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences, Nanjing, China. ²College of Life Sciences, Nanjing Normal University, Nanjing, China. ³Division of Invertebrate Zoology, American Museum of Natural History, New York, NY, USA. ⁴Facultad de Ciencias Biológicas, Universidad Nacional Mayor de San Marcos, Lima, Peru. ⁵Departamento de Entomología, Museo de Historia Natural, Universidad Nacional Mayor de San Marcos, Lima, Peru. ✉e-mail: cyc@nigpas.ac.cn

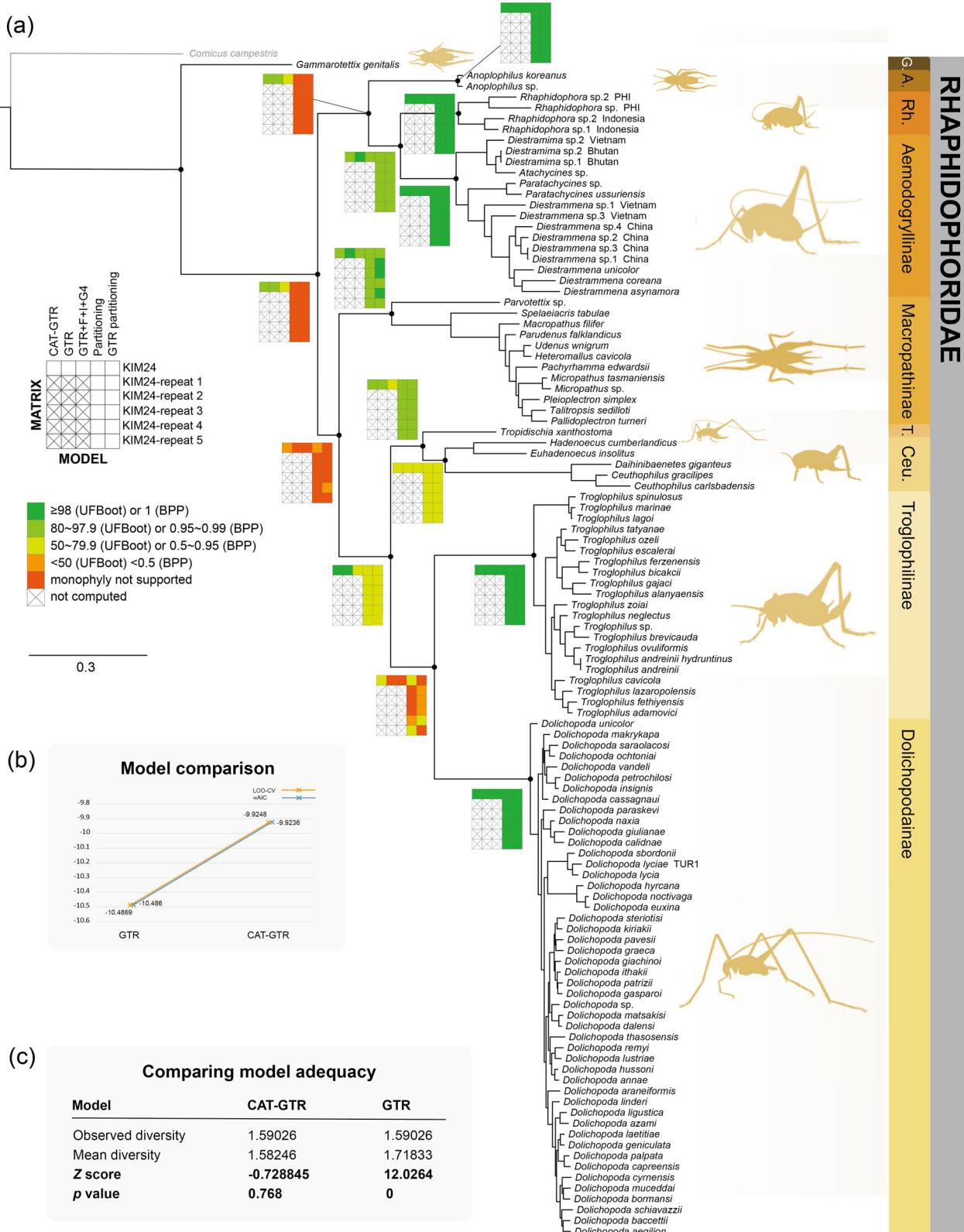


Fig. 1 | Phylogeny of Rhaphidophoridae. **a** Combined results of our phylogenetic analyses based on the 110-taxon matrix from KIM24 under both ML and BI methods. The topology is derived from the Bayesian inference under CAT-GTR + G4 in PhyloBayes. Square represents the corresponding analysis and coloured for indicating node support (support values of each phylogenetic tree are shown in

Supplementary Appendix). **b** LOO-CV and the wAIC scores of CAT-GTR and GTR models. **c** PPC results under CAT-GTR and GTR models. Both p and Z values indicated that CAT-GTR fit the dataset better. AAAnoplophilinae, Ceu. Ceuthophilinae, GGammarotettiginae, Rh Rhaphidophorinae, T Tropidischinae.

BI analysis, it was resolved as the sister group to (Ceuthophilinae + Tropidischiiinae)². However, in our BI analysis, the former relationship was relatively robustly supported under the CAT-GTR + G4 model, whereas the latter was upheld under the GTR model (Supplementary Figs. 1 and 2). Additionally, the relationship placing Troglophilinae as sister to (Ceuthophilinae + Tropidischiiinae) was corroborated in the ML analyses under GTR + F + I + G4 model and in some cases of partitioned analysis (Supplementary Figs. 3, 5–7, 10, and 15). Notably, a peculiar finding concerning the genus *Ceuthophilus*, which was not mentioned in KIM24, is that the species *C. gracilipes* consistently clustered with *Daihinibaenetes giganteus*, rendering *Ceuthophilus* paraphyletic. This situation aligns with our analysis, except for the CAT-GTR + G4 model, which supports the monophyly of *Ceuthophilus* (Fig. 1a).

Collectively, we observe that the inter-subfamilial relationships of Rhaphidophoridae remain uncertain. The results from ML methods are still unstable even after multiple repetitions, and they are difficult to reconcile with the Bayesian results. By utilizing the site-heterogeneous CAT-GTR + G4 model, it is expected to mitigate systematic errors, well represented by the monophyly of *Ceuthophilus*. Additionally, our model comparison reveals that CAT-GTR significantly outperformed the site-homogeneous GTR model, a finding that is also confirmed in practical applications^{3,4}. Our analyses highlight the significance of model comparison and modeling among-site heterogeneity in small nucleotide datasets.

Regarding the reanalysis of the 111-taxon matrix, the phylogenetic relationships within Rhaphidophoridae, as revealed by KIM24, were consistently recovered under the partitioned model (Supplementary Figs. 17–22). However, the topology inferred from the CAT-GTR + G4 model exhibited significant changes. Similar to the results based on the 110-taxon matrix, Gammarotettiginae did not cluster with Anoplophilinae but instead occupied a more basal position, attracted by *Comicus campestris* and placed within the outgroup (Supplementary Fig. 16). Although the 111-taxon matrix was not the original dataset used for phylogenetic analysis in KIM24, these results suggest that the CAT-GTR + G4-based findings from the 110-taxon matrix might also be influenced by long-branch attraction (Fig. 1a). In any case, the results of this analysis further increase the uncertainty in the phylogeny of Rhaphidophoridae.

Consequently, the current data are far from sufficient to resolve the phylogeny of Rhaphidophoridae. The Sanger-sequencing dataset contains a high proportion of missing data, with key species such as *Gammaerotettix genitalis* represented by only a single gene (601 sites). This lack of sufficient phylogenetic signal leads to inconsistent topologies across different models, and even within the same model under repeated analyses. Reconstructing the biogeographical history is crucial for understanding the evolutionary dynamics of Rhaphidophoridae. However, without a robust and well-supported phylogeny, such reconstructions lack a solid foundation. Future studies should prioritize the integration of genome-scale data, exploring complementary models and analytical methods, to address the outstanding evolutionary questions surrounding Rhaphidophoridae.

Method

KIM24 initially sampled a total of 109 ingroup taxa and 3 outgroup taxa for their phylogenetic analysis². However, in their divergence time estimation using MrBayes 3.2.62, a program limitation restricted the retention of only a single outgroup². As a result, *Camptonotus carolinensis* and *Tettigonia viridissima* were excluded from the original matrix used for phylogenetic inference². To ensure scientific rigor and reproducibility, the most appropriate matrix for our phylogenetic reanalysis would have been the 112-taxon matrix employed by KIM24. However, upon careful examination of the publicly available matrices prior to our reanalysis, we identified that the 112-taxon matrix⁵ (designated as FcC_supermatrix.fas, from Zenodo Digital Repository, <https://doi.org/10.5281/zenodo.8026258>) does not include the ingroup taxon *Parvotettix* sp. (Macropathinae), resulting in an actual count of 111 taxa. In contrast, the MrBayes-calibrated dataset⁵ (designated as MrBayes_Rhap_calibrated.nex, from Zenodo Digital Repository, <https://doi.org/10.5281/zenodo.8026258>) retained consistency in ingroup

composition with the original 112-taxon matrix. Given our objective to evaluate the reliability of evolutionary relationships among the subfamilies within Rhaphidophoridae and to maintain stringent control over variables, we proceeded with ML and BI analyses using the 110-taxon dataset, which includes *Comicus campestris* as the sole outgroup². The 110-taxon matrix comprises 3151 base pairs of nucleotide sequences, with 888 informative sites.

To reexamine the ML analysis results from KIM24, we performed an ML analysis using IQ-TREE v2.2.2⁶, employing partitioned analysis for the dataset⁷. The optimal partitioning scheme and recommended models for each partition were provided by ModelFinder⁸. Additionally, we tested a partitioned analysis under the GTR-restricted model (-mset GTR) and an unpartitioned analysis under the GTR + F + I + G4 model in IQ-TREE. Each partitioned analysis was repeated five more times to assess the reproducibility of the ML results. Branch supports were evaluated using 1000 replicates for ultrafast bootstrap approximation (UFBoot) with the -bnni option to mitigate potential overestimation of branch supports due to severe model violations⁹. The BI analysis was conducted under both the CAT-GTR + G4 and GTR models, with two independent runs for each in PhyloBayes MPI v1.9¹⁰. The CAT-GTR + G4 model, a free finite mixture model, is a better-fitting model of nucleotide substitution for addressing compositional heterogeneity¹¹. When performing a CAT-GTR analysis, the across-site compositional heterogeneity within the dataset is estimated, and the number of site-frequency categories required to adequately describe this heterogeneity is inferred^{12,13}. This demonstrates that, theoretically, CAT-GTR does not overfit the analyzed datasets, even when applied to compositionally homogeneous datasets¹². Convergence parameters for the runs were assessed using the bpcomp (maxdiff < 0.3) and tracecomp (reldiff < 0.1 and minimum effsize > 300) programs implemented in PhyloBayes^{10,14}. Additionally, we conducted ML and BI analyses on the publicly available 111-taxon matrix using partitioned models (with 6 replicates) and the CAT-GTR + G4 model, respectively.

For model comparison, the leave-one-out cross-validation (LOO-CV) score and the widely applicable information criterion (wAIC) were computed¹⁵. The scores of LOO-CV and wAIC were closely aligned, suggesting that wAIC serves as a reliable approximation of LOO-CV. Δ CV and Δ wAIC were calculated as the difference in the estimated predictive performance between the two tested models. As illustrated in Fig. 1b, the CAT-GTR model exhibited a superior fit to the dataset compared to the GTR model, as indicated by both LOO-CV (Δ CV = $-9.9248 + 10.4869 = 0.5621$) and wAIC (Δ wAIC = $-9.9236 + 10.4860 = 0.5624$). Consequently, topologies reconstructed with the CAT-GTR + G4 model were adopted as the preferred tree for elucidating relationships of rhaphidophorid subfamilies.

Posterior predictive checking implemented in PhyloBayes was employed to evaluate the goodness-of-fit of the models utilized in our phylogenetic analyses. The *p* values indicate the extent to which the model's predictions align with the observed data. The *p* value for the CAT-GTR model was 0.768, suggesting that the model's predictions correspond reasonably well with the observed data. Conversely, the *p* value for the GTR model was 0, indicating a poor fit between the model's predictions and the observed data (Fig. 1c). Furthermore, the *Z*-score quantifies the number of standard deviations the observed diversity deviates from the mean predicted diversity. For the CAT-GTR model, the *Z*-score was -0.728845 , suggesting that the observed diversity is slightly lower than the mean predicted diversity. In stark contrast, the GTR model yielded a *Z*-score of 12.0264 , signifying a substantial discrepancy between the observed diversity and the mean predicted diversity (Fig. 1c). Overall, these results demonstrate that CAT-GTR provides a significantly better fit to the data.

Data availability

The experimental data and results that support the findings of this study are available in GitHub repository with the identifier https://github.com/wyhhexa/Uncertainties_in_the_phylogeny_and_biogeography_of_cave_crickets.git.

Received: 15 June 2024; Accepted: 22 November 2025;

Published online: 22 December 2025

References

1. Cigliano, M. M., Braun, H., Eades, D. C. & Otte, D. Rhaphidophoridae Walker, 1869. Orthoptera Species File. Retrieved on 2024-05-25 at <http://orthoptera.speciesfile.org/otus/838181/overview>.
2. Kim, D. Y., Kim, S., Song, H. & Shin, S. Phylogeny and biogeography of the wingless orthopteran family Rhaphidophoridae. *Commun. Biol.* **7**, 401 (2024).
3. Li, H. et al. Mitochondrial phylogenomics of Hemiptera reveals adaptive innovations driving the diversification of true bugs. *Proc. R. Soc. B* **284**, 20171223 (2017).
4. Wang, Y. et al. Mitochondrial phylogenomics illuminates the evolutionary history of Neuropterida. *Cladistics* **33**, 617–636 (2017).
5. Kim, D. Y., Kim, S., Song, H. & Shin, S. 140 million years of evolutionary history without wings: Phylogeny and biogeography of cave crickets (Orthoptera: Rhaphidophoridae) [Data set]. Zenodo <https://doi.org/10.5281/zenodo.8026258> (2024).
6. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
7. Chernomor, O., von Haeseler, A. & Minh, B. Q. Terrace aware data structure for phylogenomic inference from supermatrices. *Syst. Biol.* **65**, 997–1008 (2016).
8. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
9. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
10. Lartillot, N., Rodrigue, N., Stubbs, D. & Richer, J. Phylobayes mpi: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Syst. Biol.* **62**, 611–615 (2013).
11. Bujaki, T., Van Looyen, K. & Rodrigue, N. Measuring the relative contribution to predictive power of modern nucleotide substitution modeling approaches. *Bioinform. Adv.* **3**, vbad091 (2023).
12. Cai, C., Tihelka, E., Pisani, D. & Donoghue, P. C. J. Resolving incongruences in insect phylogenomics: a reply to Boudinot et al. (2023). *Palaeoentomology* **007**, 176–183 (2024).
13. Giacomelli, M., Rossi, M. E., Lozano-Fernandez, J., Feuda, R. & Pisani, D. Resolving tricky nodes in the tree of life through amino acid recoding. *iScience* **25**, 105594 (2022).
14. Lartillot, N. PhyloBayes: Bayesian phylogenetics using site-heterogeneous models. in (eds Scornavacca, C., Delsuc, F. & Galtier, N.) *Phylogenetics in the Genomic Era*, 1.5:1–1.5:16 (2020).
15. Lartillot, N. Identifying the best approximating model in Bayesian phylogenetics: Bayes factors, cross-validation or wAIC? *Syst. Biol.* **72**, 616–638 (2023).

Acknowledgements

This work was supported by the National Key Research and Development Program of China (2024YFF0807601) and the National Natural Science Foundation of China (42222201).

Author contributions

C.C. and Y.W.: conceptualization, formal analysis, investigation, and writing; S.D.: writing and visualization; M.S.E.: writing.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-025-09324-w>.

Correspondence and requests for materials should be addressed to Chenyang Cai.

Reprints and permissions information is available at <http://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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2025