

Modelling among-site compositional heterogeneity resolves ant backbone phylogeny: A reply to Boudinot & Lieberman (2025)

CHEN-YANG CAI

State Key Laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences, Nanjing 210008, China

cycail@nigpas.ac.cn; <https://orcid.org/0000-0002-9283-8323>

Abstract

Understanding the early evolution of ants has been hindered by conflicting phylogenetic hypotheses and methodological inconsistencies across studies. In Cai (2024), I reanalyzed both Sanger-sequencing and genome-scale datasets of ants using rigorous model comparison and methods that account for among-site compositional heterogeneity to identify the sources of phylogenetic conflict. The results showed that the 11-loci datasets in Borowiec *et al.* (2019) failed to resolve deep ant relationships and could not determine the position of *Martialis heureka*. Analyses of the genome-scale data further revealed that the placement of key lineages depends strongly on model fit. Bayesian cross-validation and posterior predictive assessments demonstrated that the infinite mixture CAT-GTR+G4 model substantially outperforms empirical finite mixture models, providing robust support for the Leptanillinae-sister hypothesis. Criticisms by Boudinot & Lieberman (2025) regarding the study design, model choice, and convergence assessments stem from misinterpretations of the analytical framework. The matrices in Cai (2024) were explicitly designed to test model performance under controlled subsampling and filtering schemes, and all analyses showed consistent results across datasets. The findings reaffirm that accurately modelling among-site compositional heterogeneity is essential for resolving the backbone phylogeny of ants, and that under the best-fitting models, *Martialis heureka* occupies a well-supported position as sister to all non-leptanilline ants.

Keywords: evolution, ants, systematic error, phylogenomics

Introduction

In Cai (2024) (C24), I reanalyzed the published Sanger-sequencing and genome-scale datasets of ants using model comparison and methods that model among-site compositional heterogeneity to understand the sources of conflict in phylogenetic studies. I showed evidence

that the 11-loci datasets from Borowiec *et al.* (2019) (BEA) failed to resolve the deeper phylogeny of ants, and the position of Martialinae cannot be determined. Moreover, my analyses of genome-scale datasets from Romiguier *et al.* (2022) (REA) identified contentious nodes in ant phylogeny whose resolution is modelling-dependent. Based on Bayesian cross-validation and posterior predictive model checking, I showed that the infinite mixture CAT model outperforms empirical finite mixture models (C20, C40, and C60) and that, under the best-fitting CAT-GTR+G4, *Martialis heureka* Rabeling & Verhaagh, 2008 is sister to all ants except Leptanillinae (the Leptanillinae-sister hypothesis), rejecting the Leptanillomorpha hypothesis supported under worse-fitting models (Romiguier *et al.*, 2022) (Fig. 1).

Boudinot & Lieberman (2025) (BL) criticised the study design and reporting in C24 without any further Bayesian analyses. Instead, they performed a single, overly simplistic maximum-likelihood analysis in IQ-TREE2 using automatic model selection, which recovered Leptanillinae as the earliest-diverging ant lineage (see BL supplementary material). First, they chose not to discuss my results regarding the inefficacy of the 11-loci datasets from BEA, arguing that these data ‘are known to be insufficient to confidently resolve the phylogenetic problem at hand’. The absence of discussion is misleading, given that it constitutes one of the two main components of my study design. My reanalyses of the Sanger-sequencing data indicated that such small matrices were insufficient for resolving the deep phylogeny of ants, disagreeing with the Leptanillomorpha hypothesis proposed in BEA. Second, BL criticized 1) the study design (matrix preparation and model fit), 2) the model testing, and 3) the reporting of Bayesian analyses in C24, and finally claimed that CAT-GTR does not meaningfully resolve the phylogeny of ants.

As distinctly indicated in the title and abstract, C24 focused on the “ant backbone phylogeny”, that is, the

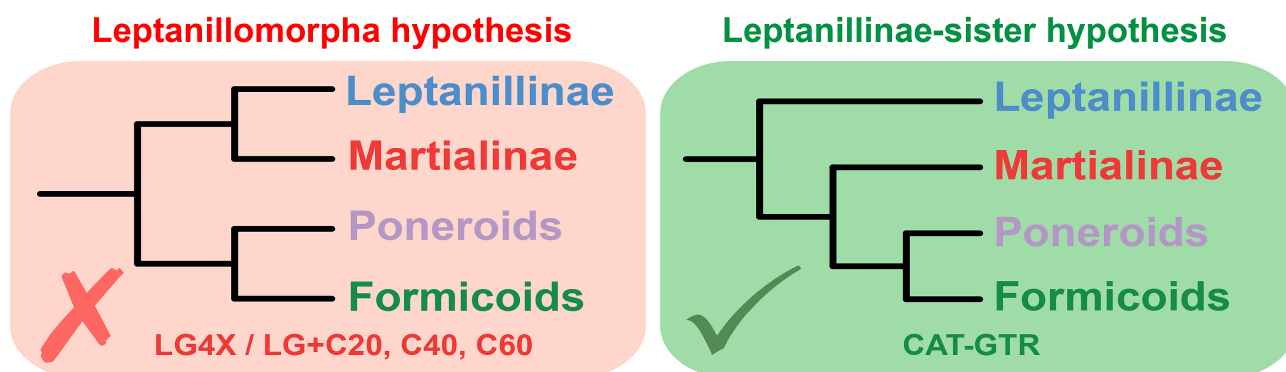


FIGURE 1. Competing hypotheses of ant backbone phylogeny. Simplified rejected topology under LG4X and empirical finite mixture models, showing the relationships of the four major lineages of ants (left) and simplified topology of the preferred tree in C24 under CAT-GTR (right).

resolution of the deepest nodes within the ant tree of life, rather than on the broader “ant phylogeny” which concerns the placement and relationships among more recently diverged subfamilies. C24 clearly stated that the remaining incongruences in poneroid relationships among different studies remain to be addressed by future studies. Despite the misleading title of BL, I reiterate here that ant backbone phylogeny can be resolved by modelling among-site compositional heterogeneity in genome-scale data.

Matrix preparation is key to the general conclusion of C24, so it was carefully considered and tested. It is common sense in modern phylogenomics that the infinite mixture CAT model can be too computationally demanding for huge phylogenomic datasets (Whelan & Halaných, 2017: table 4; Fleming *et al.*, 2023; Kapli *et al.*, 2021; Tihelka *et al.*, 2021), such as the two original supermatrices in REA. Although REA were clearly aware of the significance of modelling among-site compositional heterogeneity in suppressing long-branch-attraction artefacts (Kapli *et al.*, 2020, 2021; Tihelka *et al.*, 2021), they opted to only use the much faster site-heterogeneous LG+C20+F+G-PMSF (posterior mean site frequency) model and skipped computationally heavy model comparison. As stated in the Method of C24, that was also one of the reasons why the original supermatrices of REA were not computed under CAT in C24. C24 chose to test the impact of the application of the empirical finite mixture models (C20, C40, and C60) and the infinite mixture CAT model to carefully prepared datasets with more balanced sampling, focusing on the deeper phylogeny of ants (or the placement of *Martialis*). C24 considered the potential impact of subsampling and data filtering on tree inference and conducted sensitivity tests to rule out its influence. In C24, Matrix 1 (38 taxa, 647,114 amino acid [AA] sites) was formed after removing constant sites, which was a useful method to speed up heavy PhyloBayes runs, and

it has been widely used for inferring deep phylogenies of various life forms (Cunha *et al.*, 2022; Laumer *et al.*, 2018; Mulhair *et al.*, 2022; Janouškovec *et al.*, 2017; Strassert & Monaghan, 2022; Ochoa de Alda *et al.*, 2014; Philippe *et al.*, 2019). Matrix 2 was generated with the default setting in BMGE (-m BLOSUM62, -h 0.5) to remove potential saturated amino acid sites (Crisuolo & Gribaldo, 2010). This filtering setting has also been widely used as a default step in phylogenomic analyses of deeper relationships (Li *et al.*, 2023; Martijn *et al.*, 2018; Zaremba-Niedzwiedzka *et al.*, 2017). Matrix 3 (47 taxa, 95,201 sites), formed with a stringent setting to select slow-evolving sites, was designed to compare with Matrix 2 (47 taxa, 623,908 AA sites), as both have the same taxon sampling but different trimming parameters. In the phylogenetic analyses of C24, both matrices behaved in a similar way: under simpler models (LG4X and LG+C20, C40 and C60), the Leptanillomorpha hypothesis was strongly supported, but under CAT-GTR+G4, the Leptanillinae-sister hypothesis was maximally supported. Additionally, C24 designed a subsampled (but not trimmed) matrix to test the effect of subsampling on phylogenomic analyses. Matrix 4 (17 taxa, 1,692,050 AA sites) was a subset of the original 4,151-gene supermatrix. Like Matrix 2 and Matrix 3, phylogenetic analyses based on Matrix 4 behaved consistently. For Matrix 5, all models pointed to a consistent tree supporting the Leptanillinae-sister hypothesis. Collectively, it is obvious that subsampling of taxa and sites has no negative effects on phylogenetic inferences, and the inconsistent placement of *Martialis* among analyses in C24 was caused by adequately modelling among-site compositional heterogeneity. BL’s assertions regarding the impact of subsampling on the phylogenomic analyses are not true.

Modelling of amino acid replacement is central to phylogenomic inference, particularly so when dealing with deeper relationships and rapid radiations (Kapli *et al.*,

2021). Model comparison is a challenging yet crucial step in phylogenomic analyses. In the original study by REA, they arbitrarily selected the LG+C20+F+G-PMSF model as so to mitigate systematic error and used the result as their main tree, but they did not compare it with other better-fitting models such as the C60 and CAT. As mentioned in C24, the selection of this particular model was apparently a compromise since runs of supermatrices under the C40 and C60 models are computationally expensive in terms of both running time and memory requirements. The partition model in REA (advocated by BL) was selected from a limited collection of site-homogeneous models, and this strategy has been proven to behave worse than the CAT model (Feuda *et al.*, 2017). BL correctly cited the efficacy of CAT-GTR that better approximates biological reality and fits well (Giacomelli *et al.*, 2022), but they failed to correctly understand how the model works. As convincingly demonstrated in Giacomelli *et al.* (2022), CAT uses more frequency categories but the advantage of Bayesian methods (Fabreti & Höhna, 2022) is that they allow the complexity of the model to be tuned to the analyzed dataset and more adequately describe it. Even if all matrices generated in C24 were across-site compositionally homogeneous ones, PhyloBayes would effectively analyze the data under a GTR model, rather than using an over-parameterized model with hundreds of site-frequency categories (Cai *et al.*, 2024; Giacomelli *et al.*, 2022). This property of CAT-GTR is critical, but it was often misunderstood by critics of CAT-GTR, including Boudinot and colleagues (Boudinot *et al.*, 2022; Boudinot & Lieberman, 2025). As suggested in previous studies (Feuda *et al.*, 2017; Giacomelli *et al.*, 2022), absolute goodness-of-fit tests (such as posterior predictive analysis [PPA]) were used by C24 to discriminate models. The analyses showed that LG+C20 (and C40, C60 models) describes the across-site compositional heterogeneity of Matrix 3 poorly, and worse than CAT. This result is congruent with those of Bujaki & Rodrigue (2022), who demonstrated in simulations that free finite mixtures (CAT-GTR+G4) consistently outperform empirical finite mixtures.

It is absolutely true that convergence and mixing checking are important in PhyloBayes analyses, and that is why I generated smaller matrices to understand the ant backbone phylogeny. However, regarding the convergence of Bayesian runs in C24, the claims of BL are all subjective speculations that stemmed from their misunderstanding of PhyloBayes. C24 clearly reported in supplementary figure captions (see figs S1–S5 in Cai, 2024) the required statistics for an acceptable PhyloBayes run (Lartillot, 2020; Lartillot *et al.*, 2013). As in similar studies using PhyloBayes (Feuda *et al.*, 2017; Lozano-Fernandez *et al.*, 2019a, b), C24 showed all convergence statistics, including total number of

cycles, burnin, maxdiff (using *bpcomp*), and minimal overall effsize (using *tracecomp*), except for Matrix 2 due to its large size (*bpcomp* maxdiff/meandiff = 0). If BL referred to the manual of PhyloBayes, they would have easily understood that *tracecomp* generated the effective sizes for all parameters recorded in the trace files, and *bpcomp* produced the largest discrepancy observed across all bipartitions. C24 showed all required convergence and mixing statistics for focal analyses, and this is in accordance with the practice of PhyloBayes runs. Rather than engaging in criticisms of others' analyses, BL should have rerun even just one PhyloBayes analysis using the freely available matrices in C24.

Regarding cross-validation of C24, Boudinot & Lieberman (2025) again misunderstood the method C24 used. C24 mentioned many times that, leave-one-out cross-validation and the widely applicable information criterion, a new method recently implemented in PhyloBayes v. 1.9, were used (Lartillot, 2023), but not cross-validation as in older versions of PhyloBayes. The quality of C24's estimation was evaluated based on ESS statistics, including %(ess<10) and f(ess<10). Both values in C24 were satisfactory (less than 0.1) and were freely accessible in the Dryad Digital Repository (Cai, 2024).

Besides the misunderstandings mentioned above, BL made numerous mistakes in understanding the models used in C24: 1) LG4X+R is not a site-homogeneous model, because it uses four amino acid replacement matrices summarizing the biochemical properties of amino acids (Le *et al.*, 2012); 2) For empirical finite mixture models, C24 used the PMSF model (Wang *et al.*, 2018) in many cases; and 3) GHOST (LG+FO*H4) is defined as a heterotachous model (Crotty *et al.*, 2020), but not site-homogeneous one as in an earlier version of BL.

To my surprise, BL incorrectly adopted the reviewer's misunderstanding regarding the distinction between *among-site compositional heterogeneity* and *among-lineage compositional heterogeneity* (the same reviewer who had recommended rejection of my Reply to BL during the review process in *Communications Biology*). Throughout Cai (2024), including the title, abstract, and main text, I consistently referred to among-site compositional heterogeneity and never suggested that CAT-GTR can model among-lineage compositional heterogeneity (differences in the nucleotide/amino acid composition across taxa caused by adaptation to local habitats or lineage specific substitutional biases; Tihelka *et al.*, 2021). The term *among-site compositional heterogeneity*, rather than the ambiguous term *site heterogeneity* used in BL (as advocated by the reviewer), is the one commonly used in the phylogenomic literature (e.g., Feuda *et al.*, 2017; Kapli *et al.*, 2020, 2021; Liu *et al.*, 2014; Schrepf *et al.*, 2020; Tihelka *et al.*, 2021).

Simply put, *site heterogeneity* used in BL is an imprecise term in phylogenetics, since it may refer to among-site compositional heterogeneity (commonly modeled with CAT models) or among-site rate heterogeneity (commonly modeled with a Gamma distribution) (detailed in Tihelka *et al.*, 2021). Regrettably, it appears that many contemporary taxonomists (*e.g.*, Boudinot & Lieberman, 2025) lack a firm grasp of basic phylogenomic principles—for example, the distinction between among-site and among-lineage compositional heterogeneity. This shortfall leads to terminological conflation and confusion and methodological misapplication in the literature, and highlights the need for improved training and more rigorous peer review in phylogenomics.

In summary, the conclusion in Cai (2024) stands firm and clear: *Martialis heureka* is sister to all extant ants except Leptanillinae, as supported by the best-fitting CAT model. The synapomorphies of Leptanillomorpha claimed in Boudinot *et al.* (2022) could be a consequence of convergent evolution.

Acknowledgements

We are grateful to two reviewers for their valuable comments on the manuscript. This work has been supported by the National Natural Science Foundation of China (42222201).

References

- Borowiec, M.L., Rabeling, C., Brady, S.G., Fisher, B.L., Schultz, T.R. & Ward, P.S. (2019) Compositional heterogeneity and outgroup choice influence the internal phylogeny of the ants. *Molecular Phylogenetics and Evolution*, 134, 111–121. <https://doi.org/10.1016/j.ympev.2019.01.024>
- Boudinot, B.E. & Lieberman, Z.E. (2025) Ant phylogeny is not resolved by the application of site heterogeneous models. *Communications Biology*, 8, 1458. <https://doi.org/10.1038/s42003-025-07849-8>
- Boudinot, B.E., Khouri, Z., Richter, A., Griebenow, Z.H., van de Kamp, T., Perrichot, V. & Barden, P. (2022) Evolution and systematics of the Aculeata and kin (Hymenoptera), with emphasis on the ants (Formicoidea: †@@@idae fam. nov., Formicidae). *bioRxiv*, preprint. <https://doi.org/10.1101/2022.02.20.480183>
- Boudinot, B.E., Fikáček, M., Lieberman, Z.E., Kusy, D., Bocak, L., McKenna, D.D. & Beutel, R.G. (2022) Systematic bias and the phylogeny of Coleoptera—A response to Cai *et al.* (2022) following the responses to Cai *et al.* (2020). *Systematic Entomology*, 48, 223–232. <https://doi.org/10.1111/syen.12570>
- Bujaki, T. & Rodrigue, N. (2022) Bayesian cross-validation comparison of amino acid replacement models: contrasting profile mixtures, pairwise exchangeabilities, and gamma-distributed rates-across-sites. *Journal of Molecular Evolution*, 90, 468–475. <https://doi.org/10.1007/s00239-022-10076-y>
- Cai, C.Y. (2024) Ant backbone phylogeny resolved by modelling compositional heterogeneity among sites in genomic data. *Communications Biology*, 7, 106. <https://doi.org/10.1038/s42003-024-05793-7>
- Cai, C.Y., Tihelka, E., Pisani, D. & Donoghue, P.C. (2024) Resolving incongruences in insect phylogenomics: a reply to Boudinot *et al.* (2023). *Palaeoentomology*, 7 (2), 176–183. <https://doi.org/10.11646/palaeoentomology.7.2.2>
- Criscuolo, A. & Gribaldo, S. (2010) BMGE (Block Mapping and Gathering with Entropy): selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evolutionary Biology*, 10, 210. <https://doi.org/10.1186/1471-2148-10-210>
- Crotty, S.M., Minh, B.Q., Bean, N.G., Holland, B.R., Tuke, J., Jermin, L.S. & Haeseler, A.V. (2020) GHOST: recovering historical signal from heterotachously evolved sequence alignments. *Systematic Biology*, 69, 249–264. <https://doi.org/10.1093/sysbio/syz051>
- Cunha, T.J., Reimer, J.D. & Giribet, G. (2022) Investigating sources of conflict in deep phylogenomics of vetigastropod snails. *Systematic Biology*, 71, 1009–1022. <https://doi.org/10.1093/sysbio/syab071>
- Fabreti, L.G. & Höhna, S. (2022) Bayesian inference of phylogeny is robust to substitution model over-parameterization. *bioRxiv*, preprint. <https://doi.org/10.1101/2022.02.17.480861>
- Feuda, R., Dohrmann, M., Pett, W., Philippe, H., Rota-Stabelli, O., Lartillot, N., Wörheide, G. & Pisani, D. (2017) Improved modeling of compositional heterogeneity supports sponges as sister to all other animals. *Current Biology*, 27, 3864–3870. <https://doi.org/10.1016/j.cub.2017.11.008>
- Fleming, J.F., Valero-Gracia, A. & Struck, T.H. (2023) Identifying and addressing methodological incongruence in phylogenomics: A review. *Evolutionary Applications*, 16, 1087–1104. <https://doi.org/10.1111/eva.13565>
- Giacomelli, M., Rossi, M.E., Lozano-Fernandez, J., Feuda, R. & Pisani, D. (2022) Resolving tricky nodes in the tree of life through amino acid recoding. *iScience*, 25, 105594. <https://doi.org/10.1016/j.isci.2022.105594>
- Janouškovec, J., Tikhonenkov, D.V., Burki, F., Howe, A.T., Rohwer, F.L., Mylnikov, A.P. & Keeling, P.J. (2017) A new lineage of eukaryotes illuminates early mitochondrial genome reduction. *Current Biology*, 27, 3717–3724. <https://doi.org/10.1016/j.cub.2017.10.051>
- Kapli, P., Flouri, T. & Telford, M.J. (2021) Systematic errors in phylogenetic trees. *Current Biology*, 31, R59–R64. <https://doi.org/10.1016/j.cub.2020.11.043>
- Kapli, P., Yang, Z. & Telford, M.J. (2020) Phylogenetic tree building in the genomic age. *Nature Reviews Genetics*, 21 (7), 428–444. <https://doi.org/10.1038/s41576-020-0233-0>

- Lartillot, N. (2020) PhyloBayes: Bayesian phylogenetics using site-heterogeneous models. In: Scornavacca, C., Delsuc, F. & Galtier, N. (Eds), *Phylogenetics in the genomic era*, pp. 1.5:1–1.5:16. No commercial publisher, author's open access book. <https://hal.science/PGE>
- Lartillot, N. (2023) Identifying the best approximating model in Bayesian phylogenetics: Bayes factors, cross-validation or wAIC? *Systematic Biology*, 72, 616–638. <https://doi.org/10.1093/sysbio/syad004>
- Lartillot, N., Rodrigue, N., Stubbs, D. & Richer, J. (2013) PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Systematic Biology*, 62, 611–615. <https://doi.org/10.1093/sysbio/syt022>
- Laumer, C.E., Gruber-Vodicka, H.R., Hadfield, M.G., Pearse, V.B., Riesgo, A., Marioni, J.C. & Giribet, G. (2018) Support for a clade of Placozoa and Cnidaria in genes with minimal compositional bias. *eLife*, 7, e36278. <https://doi.org/10.7554/eLife.36278>
- Le, S.Q., Dang, C.C. & Gascuel, O. (2012) Modeling protein evolution with several amino acid replacement matrices depending on site rates. *Molecular Biology and Evolution*, 29, 2921–2936. <https://doi.org/10.1093/molbev/mss112>
- Li, Y.D., Engel, M.S., Tihelka, E. & Cai, C.Y. (2023) Phylogenomics of weevils revisited: data curation and modelling compositional heterogeneity. *Biology Letters*, 19, 20230307. <https://doi.org/10.1098/rsbl.2023.0307>
- Liu, Y., Cox, C.J., Wang, W. & Goffinet, B. (2014) Mitochondrial phylogenomics of early land plants: mitigating the effects of saturation, compositional heterogeneity, and codon-usage bias. *Systematic Biology*, 63 (6), 862–878. <https://doi.org/10.1093/sysbio/syu049>
- Lozano-Fernandez, J., Tanner, A.R., Giacomelli, M., Carton, R., Vinther, J., Edgecombe, G.D. & Pisani, D. (2019a) Increasing species sampling in chelicerate genomic-scale datasets provides support for monophyly of Acari and Arachnida. *Nature Communications*, 10, 2295. <https://doi.org/10.1038/s41467-019-10244-7>
- Lozano-Fernandez, J., Giacomelli, M., Fleming, J.F., Chen, A., Vinther, J., Thomsen, P.F., Glenner, H., Palero, F., Legg, D.A., Iliffe, T.M. & Pisani, D. (2019b) Pancrustacean evolution illuminated by taxon-rich genomic-scale data sets with an expanded remipede sampling. *Genome Biology and Evolution*, 11, 2055–2070. <https://doi.org/10.1093/gbe/evz097>
- Martijn, J., Vosseberg, J., Guy, L., Offre, P. & Ettema, T.J. (2018) Deep mitochondrial origin outside the sampled alphaproteobacteria. *Nature*, 557, 101–105. <https://doi.org/10.1038/s41586-018-0059-5>
- Mulhair, P.O., McCarthy, C.G., Siu-Ting, K., Creevey, C.J. & O'Connell, M.J. (2022) Filtering artifactual signal increases support for Xenacoelomorpha and Ambulacraria sister relationship in the animal tree of life. *Current Biology*, 32, 5180–5188. <https://doi.org/10.1016/j.cub.2022.10.036>
- Ochoa de Alda, J.A., Esteban, R., Diago, M.L. & Houmard, J. (2014) The plastid ancestor originated among one of the major cyanobacterial lineages. *Nature Communications*, 5, 4937. <https://doi.org/10.1038/ncomms5937>
- Philippe, H., Poustka, A.J., Chiodin, M., Hoff, K.J., Dessimoz, C., Tomiczek, B., Schiffer, P.H., Müller, S., Domman, D., Horn, M., Kuhl, H., Timmermann, B., Satoh, N., Hikosaka-Katayama, T., Nakano, H., Rowe, M.L., Elphick, M.R., Thomas-Chollier, M., Hankeln, T., Mertes, F., Wallberg, A., Rast, J.P., Copley, R.R., Martinez, P. & Telford, M.J. (2019) Mitigating anticipated effects of systematic errors supports sister-group relationship between Xenacoelomorpha and Ambulacraria. *Current Biology*, 29, 1818–1826. <https://doi.org/10.1016/j.cub.2019.04.009>
- Romiguier, J., Borowiec, M.L., Weyna, A., Helleu, Q., Loire, E., La Mendola, C., Rabeling, C., Fisher, B.L., Ward, P.S. & Keller, L. (2022) Ant phylogenomics reveals a natural selection hotspot preceding the origin of complex eusociality. *Current Biology*, 32, 2942–2947. <https://doi.org/10.1016/j.cub.2022.05.001>
- Schrempf, D., Lartillot, N. & Szöllösi, G. (2020) Scalable empirical mixture models that account for across-site compositional heterogeneity. *Molecular Biology and Evolution*, 37 (12), 3616–3631. <https://doi.org/10.1093/molbev/msaa145>
- Strasser, J.F. & Monaghan, M.T. (2022) Phylogenomic insights into the early diversification of fungi. *Current Biology*, 32, 3628–3635. <https://doi.org/10.1016/j.cub.2021.08.057>
- Tihelka, E., Cai, C.Y., Giacomelli, M., Lozano-Fernandez, J., Rota-Stabelli, O., Huang, D.Y., Engel, M.S., Donoghue, P.C. & Pisani, D. (2021) The evolution of insect biodiversity. *Current Biology*, 31, R1299–R1311. <https://doi.org/10.1016/j.cub.2021.08.057>
- Wang, H.C., Minh, B.Q., Susko, E. & Roger, A.J. (2018) Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate phylogenomic estimation. *Systematic Biology*, 67, 216–235. <https://doi.org/10.1093/sysbio/syx068>
- Whelan, N.V. & Halanich, K.M. (2017) Who let the CAT out of the bag? Accurately dealing with substitutional heterogeneity in phylogenomic analyses. *Systematic Biology*, 66 (2), 232–255. <https://doi.org/10.1093/sysbio/syw084>
- Zaremba-Niedzwiedzka, K., Caceres, E.F., Saw, J.H., Bäckström, D., Juzokaite, L., Vancaester, E., Seitz, K.W., Anantharaman, K., Starnawski, P., Kjeldsen, K.U., Stott, M.B., Nunoura, T., Banfield, J.F., Schramm, A., Baker, B.J., Spang, A. & Ettema, T.J.G. (2017) Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature*, 541, 353–358. <https://doi.org/10.1038/nature21031>