A genome-scale phylogeny of the superfamily Entomobryoidea (Entomobryomorpha: Collembola)†

Nerivania Nunes Godeiro[1,2,*], Yinhuan Ding[2], Bruno Cavalcante Bellini[3], Nikolas Gioia Cipola[4], Sopark Jantarit[5] and Feng Zhang[2]

[1]Shanghai Natural History Museum, Shanghai Science & Technology Museum, Shanghai, China; [2]Department of Entomology, College of Plant Protection, Nanjing Agricultural University, Nanjing, China; [3]Department of Botany and Zoology, Biosciences Center, Federal University of Rio Grande do Norte, Natal, Rio Grande do Norte, Brazil; [4]Laboratório de Sistemática e Ecologia de Invertebrados do Solo, Instituto Nacional de Pesquisas da Amazônia—INPA, CPEN, Manaus, AM, Brazil; [5]Excellence Center for Biodiversity of Peninsular Thailand, Faculty of Science, Prince of Songkla University, Hat Yai, Songkhla, Thailand. † Presented the 2nd International Electronic Conference on Diversity - Animals, Plants and Microbes, online,15–31 March 2022. Correspondence: nerivania@gmail.com

INTRODUCTION

Molecular studies related to the Entomobryoidea superfamily are still incipient, and no study in the literature analyzed its internal relationships with high-resolution molecular data, while just a few studies used mitogenomes or instead have focused on only a few species [1, 2]. Currently, the superfamily is divided into three families: Entomobryidae, Paronellidae, and Orchesellidae [3], but during the past years many changes were made in its internal organization. While the relationships within some entomobryid clades are robust, like within the Seirinae, the resolution of the paraphyly between the families Paronellidae and Entomobryidae remains unsolved and have not been assessed genome-wide to determine the specific points of discordance suggested in previous studies.

Here we tested a dataset of universal molecular markers for Entomobryoidea using 33 representatives of the superfamily and two outgroups assembled De Novo from low-coverage whole-genome data (10x). Employing a series of computationally efficient bioinformatic tools [4] we extracted thousands of genes used to create the first genome-scale phylogeny of the superfamily Entomobryoidea.

MATERIALS AND METHODS

From the current nine subfamilies of Entomobryoidea, seven were sampled here: Heteromurinae, Orchesellinae, Entomobryinae, Lepidocyrtinae, Seirinae, Paronellinae, and Salininae. Twenty-one genomic data were newly generated for this study, while other data were previously published or are in the publication process.

MGIseq2000 platform was used for sequencing paired-end reads with 150 bp length. Approximately 10G of low-coverage data were produced for each sample.

The main steps of our methodology are described bellow:
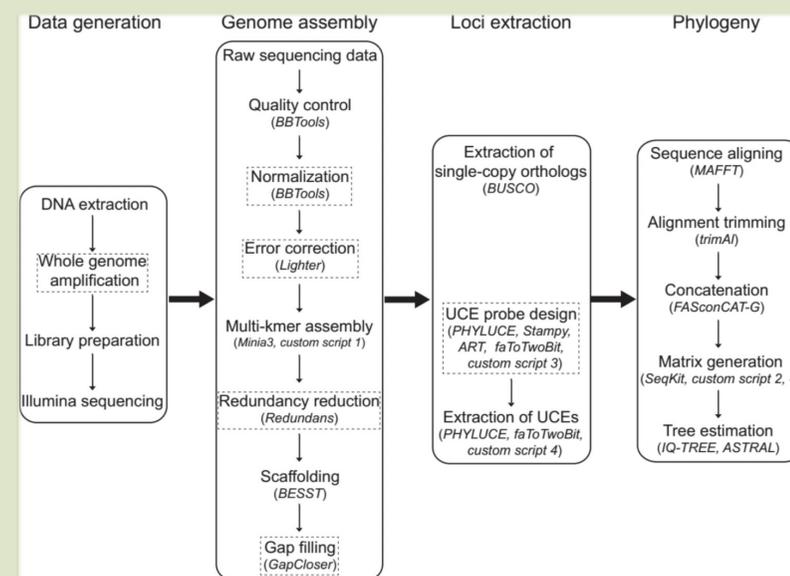
RESULTS
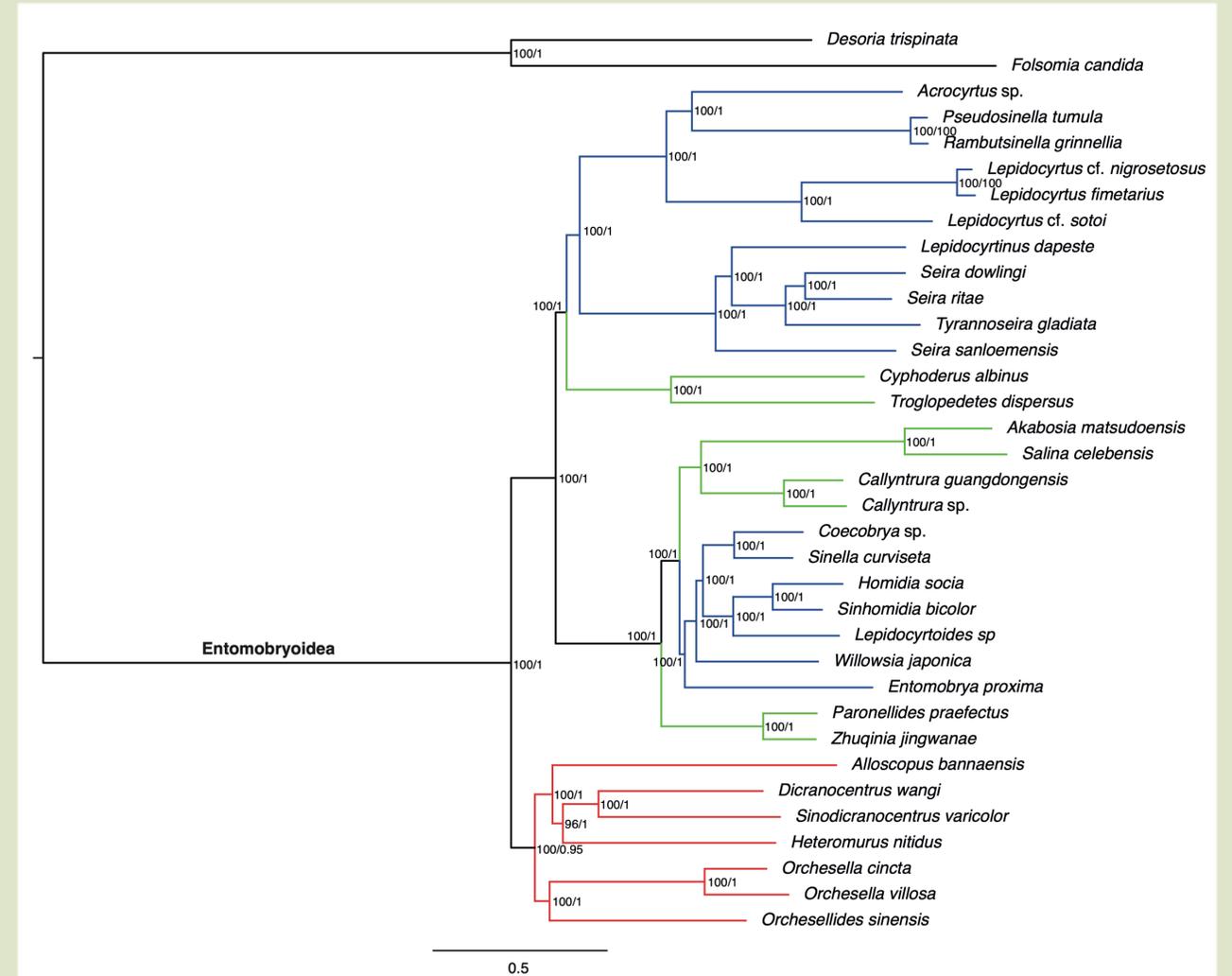
Based on an expanded set of molecular markers we were able to reaffirm the Orchesellidae (in red) as a family, holding at least two independent subfamilies: Heteromurinae and Orchesellinae; Entomobryidae (in blue) and Paronellidae (in green) remains paraphyletic with the Entomobryinae as group gathering unscaled and scaled taxa with different furcal morphologies; and the Seirinae as the sister-group of the Lepidocyrtinae, with the Paronellinae s. str. as a closely related group to both.

1. Godeiro, N.N.; Bellini, B.C.; Ding, N.; Xu, C.; Ding, Y.; Zhang, F. A Mitogenomic Phylogeny of the Entomobryoidea (Collembola): A Comparative Perspective. Zool. Scr. 2021, 50, 658–666;
2. Sun, X.; Ding, Y.; Orr, M.C.; Zhang, F. Streamlining Universal Single-Copy Orthologue and Ultraconserved Element Design: A Case Study in Collembola. Mol. Ecol. Resour. 2020, n/a, 1–12;
3. Zhang, F.; Bellini, B.C.; Soto-adames, F.N. New Insights into the Systematics of Entomobryoidea (Collembola: Entomobryomorpha): First Instar Chaetotaxy, Homology and Classification. Zoo. System. 2019, 44, 249–278;
4. Zhang, F.; Ding, Y.; Zhu, C.-D.; Zhou, X.; Orr, M.; Scheu, S.; Luan, Y.-X. Phylogenomics from Low-Coverage Whole-Genome Sequencing. Methods Ecol. Evol. 2019, 1–11.