Received: 10 May 2023 | Revised: 11 October 2023 | Accepted: 16 October 2023

DOI: 10.1111/1755-0998.13885

RESOURCE ARTICLE

MOLECULAR ECOLOGY RESOURCES WILEY

PacBio high-throughput multi-locus sequencing reveals high genetic diversity in mushroom-forming fungi

Rachel A. Swenie[1,2] | Brian P. Looney[3] | Yi-Hong Ke[3,4] | J. Alejandro Rojas[3,5] | Marc A. Cubeta[6] | Gitta J. Langer[7] | Rytas Vilgalys[3] | P. Brandon Matheny[1]

[1]Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, Tennessee, USA

[2]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts, USA

[3]Department of Biology, Duke University, Durham, North Carolina, USA

[4]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan, USA

[5]Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, Michigan, USA

[6]Department of Entomology and Plant Pathology, Center for Integrated Fungal Research, North Carolina State University, Raleigh, North Carolina, USA

[7]Department of Forest Protection, Northwest German Forest Research Institute, Göttingen, Germany

Correspondence
Rachel A. Swenie, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA.
Email: rswenie@fas.harvard.edu

Funding information
University of Tennessee; National Science Foundation, Grant/Award Number: DEB-2030779 and 1452154

Handling Editor: Isheng J. Tsai

Abstract

Multi-locus sequence data are widely used in fungal systematic and taxonomic studies to delimit species and infer evolutionary relationships. We developed and assessed the efficacy of a multi-locus pooled sequencing method using PacBio long-read high-throughput sequencing. Samples included fresh and dried voucher specimens, cultures and archival DNA extracts of Agaricomycetes with an emphasis on the order Cantharellales. Of the 283 specimens sequenced, 93.6% successfully amplified at one or more loci with a mean of 3.3 loci amplified. Our method recovered multiple sequence variants representing alleles of rDNA loci and single copy protein-coding genes *rpb1*, *rpb2* and *tef1*. Within-sample genetic variation differed by locus and taxonomic group, with the greatest genetic divergence observed among sequence variants of *rpb2* and *tef1* from corticioid Cantharellales. Our method is a cost-effective approach for generating accurate multi-locus sequence data coupled with recovery of alleles from polymorphic samples and multi-organism specimens. These results have important implications for understanding intra-individual genomic variation among genetic loci commonly used in species delimitation of fungi.

KEYWORDS
allelic diversity, DNA barcoding, fungal systematics, haplotype, long-read sequencing, sequence variant

1 | INTRODUCTION

High-throughput sequencing (HTS) methods have grown in popularity due to increased reliability and decreased cost. Single molecule sequencing, also referred to as 'third generation' HTS, has been widely adopted in genome shotgun sequencing and transcriptome sequencing for generating accurate DNA sequences up to or exceeding 10 kb in length (Heather & Chain, 2016). The PacBio Sequel third-generation HTS platform has also shown promise for metabarcoding of microbial soil samples and amplicon libraries (Bourne et al., 2018; Runnel et al., 2022; Tedersoo et al., 2018).

Multi-locus data from fungal specimens, conventionally produced by Sanger sequencing, are widely used in systematic and taxonomic studies. While Sanger sequencing remains popular due to accuracy and ease of use, it is a costly and time-consuming method for producing multi-locus data that is prone to generating heterogeneous reads when multiple alleles or contamination are present in a sample. Sanger sequencing is also limited to reads 1 kb or less in

