HAL open science

Population genomics revealed cryptic species within host-specific zombie-ant fungi (Ophiocordyceps unilateralis)

Noppol Kobmoo, Suchada Mongkolsamrit, Nuntanat Arnamnart, Janet Jennifer Luangsa-Ard, Tatiana Giraud

To cite this version:

Noppol Kobmoo, Suchada Mongkolsamrit, Nuntanat Arnamnart, Janet Jennifer Luangsa-Ard, Tatiana Giraud. Population genomics revealed cryptic species within host-specific zombie-ant fungi (Ophiocordyceps unilateralis). Molecular Phylogenetics and Evolution, In press, 10.1016/j.ympev.2019.106580 . hal-02266424

HAL Id: hal-02266424

https://agroparistech.hal.science/hal-02266424v1

Submitted on 14 Aug 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Population genomics revealed cryptic species within host-specific zombie-ant fungi (*Ophiocordyceps unilateralis*)

Noppol KOBMOO[†,‡], Suchada MONGKOLSAMRIT[†], Nuntanat ARNAMNART[†], Janet Jennifer LUANGSA-ARD[†], Tatiana GIRAUD[‡]

[†] National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), 113 Thailand Science Park, Phahonyothin Rd., Khlong Nueng, Khlong Luang, 12120 THAILAND.

[‡] Ecologie Systématique Evolution, Univ. Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay 91400 Orsay, France.

Corresponding author: Noppol Kobmoo,

E-mail: noppol.kob@biotec.or.th

© 2019. This manuscript version is made available under CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

Abstract

The identification and delimitation of species boundaries are essential for understanding speciation and adaptation processes and for the management of biodiversity as well as development for applications. *Ophiocordyceps unilateralis* sensu lato is a complex of fungal pathogens parasitizing Formicine ants, inducing zombie behaviors in their hosts. Previous taxonomic works with limited numbers of samples and markers led to the "one ant-one fungus" paradigm, resulting in the use of ant species as a proxy for fungal identification. Here, a population genomic study with sampling on three ant species across Thailand supported the existence of host-specific species in *O. unilateralis* s.l. with no footprints of long term introgression despite occasional host shifts and first-generation hybrids. We further detected genetic clusters within the previously delimited fungal species, with each little footprints of recombination, suggesting high levels of inbreeding. The clusters within each of *O. camponoti-leonardi* and *O. camponoti-saundersi* were supported by differentiation throughout the genome, suggesting they may constitute further cryptic species parasitizing the same host, challenging the one ant-one fungus paradigm. These genetic clusters had different geographical ranges, supporting different biogeographic influences between the north/center and the south of Thailand, reinforcing the scenario in which Thailand endured compartmentation during the latest Pleistocene glacial cycles.

Key words: *Ophiocordyceps unilateralis*, cryptic species, phylogenetic species, recombination, population genomics.

1. Introduction

The identification and delimitation of species boundaries are essential for understanding speciation and adaptation processes and for the management of biodiversity and conservation. The existence of cryptic species, i.e. distinct species with hardly distinguishable morphologies, have been increasingly documented in all phyla and constitutes a challenge for the study of speciation, the assessment of biodiversity and the implementation of conservation plans (Fišer et al., 2018; Struck et al., 2018). Particularly for fungi, the recognition of species can be difficult as they are versatile organisms, with variable morphology and life styles (Khonsanit et al., 2019; Molina et al., 2011; Mongkolsamrit et al., 2018; Montarry et al., 2009; Vincenot et al., 2017). Fungi are frequently symbionts to other organisms and often evolve host specialization (Fournier and Giraud, 2008; Gerardo et al., 2006, 2004), which is a fertile background for host-specific, cryptic species. The use of the phylogenetic species criterion based on concordance between gene genealogies has been highly useful for delimiting fungal cryptic species (Taylor et al., 2000). With the advent of next-generation sequencing, the concordance among gene genealogies can be assessed throughout the genome and new phylogenetic species can be revealed with even more power and reliability (Matute and Sepúlveda, 2019; Weiss et al., 2018) or well-established species can be questioned (Lifjeld, 2015). Furthermore, a population genomics approach can identify gene flow and introgression at fine scale along the genome (Chan et al., 2017; Savary et al., 2018).

Ophiocordyceps unilateralis sensu lato is a ubiquitous species complex of ant-pathogenic fungi with a pan-tropical distribution (Araújo et al., 2018) and an outstanding biology; these fungi modify the behavior of their ant hosts to promote their own dispersal. Infected ants are called zombie ants because they develop erratic behaviors, excluding themselves from their nest before climbing into vegetation and biting vegetal substrates until death (Hughes et al., 2011). The taxon called *O. unilateralis* has historically been characterized by the presence of fungal stalks arising from dead ants with lateral cushions containing perithecia and whole ascospores. Recent studies

based on morphology and molecular markers have shown that *O. unilateralis* represented a complex of cryptic species, with allopatric species from distinct continents (Araújo et al., 2018) as well as host-specific sympatric species (Kobmoo et al., 2012). A single and host-specific fungal species appeared to be associated to each ant species, leading to the "one ant-one fungus" paradigm and to the use of ant host identity as proxy for fungal species identification. However, little attention has been paid to intra-specific diversity and to the possibility of differentiated populations or species occurring on a given ant species. Species in the *O. unilateralis* complex are closely related and many are sympatric, which could lead to introgression and hybridization, although these questions have never been comprehensively studied.

In Thailand, six cryptic species of the *O. unilateralis s.l.* complex have been identified based on molecular phylogenies and morphological traits (Kobmoo et al., 2015; Luangsa-ard et al., 2011): *O. polyrhachis-furcata*, *O. camponoti-leonardi*, *O. camponoti-saundersi*, *O. septa* and *O. rami* and *O. halabalaensis*. At the macro-morphological level, only *O. rami* and *O. halabalaensis* can be easily distinguished from the others, having one stalk arising from the junction of the ant head and thorax, and two others arising from the junction of the forelegs and thorax; *O. rami* has sinuous purple stromata while *O. halabalaensis* has bigger but less sinuous blackish-brown stromata. At the micro-morphological level, *O. septa* has distinctive swollen and multi-septate ascospores; *O. polyrhachis-furcata*, *O. camponoti-leonardi* and *O. camponoti-saundersi* are the most similar genetically and morphologically and can be found on the ants *Polyrhachis furcata*, *Colobopsis leonardi* and *Colobopsis saundersi,* respectively (Kobmoo et al., 2012). Although morphological traits can distinguish these species, host species is the most convenient proxy for species identification and appears reliable (Araújo et al., 2018; Kobmoo et al., 2015, 2012). The criterion of congruence among gene genealogies was consistent with the separation of these host-specialized cryptic species in *O. unilateralis s.l.* (Kobmoo et al., 2012); however, they were based on only two gene sequences (Elongation factor 1-alpha and Beta tubulin).

Due to frequent expansion and contraction events of the forest during the Pleistocene glacial cycles, Southeast Asia shows strong biogeographic compartmentation (Woodruff, 2010). Genetic differentiation between the north and the peninsular south have been reported in many organisms, including insects such as bees (Insuan et al., 2007; Sittipraneed et al., 2001) and flies (Pramual et al., 2005), trees (Yu et al., 2019), vertebrates including frogs (Chen et al., 2018), monkeys (Karanth et al., 2008) and bats (Tu et al., 2017); so far this geographical pattern has not been studied in fungi. Furthermore, an increase in genetic diversity from North to South has been reported in various organisms in the Northern hemisphere, including fungi (Miraldo et al., 2016; Smith et al., 2017; Vercken et al., 2010) although investigations in Southeast Asia are still lacking.

Furthermore, genetic diversity in fungi is also influenced by reproductive modes (sexual vs clonal) and mating systems (selfing vs outcrossing) (Nieuwenhuis and James, 2016). The sexual reproductive structures of *O. unilateralis* has been widely observed in nature; the fungi develop ascomata from diseased ants, which release ascospores into the environment. However, additional transmission via conidia (asexual spores) cannot be ruled out and sexual reproduction can occur between closely related individuals, and even between clonemates in case these fungi are homothallic, which is unknown. Which reproductive mode and mating system are predominant, and whether there are differences between species, are still open questions which can have major implications for genetic diversity, reproductive barrier and adaptation in these fungi.

In the present study, we therefore used a population genomics approach to investigate: i) whether the existence of host-specific cryptic species was supported by genome-wide genetic differentiation in *O. unilateralis*, ii) in this case, whether introgression occurred between species, and in which genomic locations, iii) whether cryptic genetic subdivision could be identified beyond host specific species, e.g. according to geographic regions, iv) whether genetic diversity increased towards the South within each species, and v) what the reproductive mode and mating system of the different species were. For these goals, we collected samples of *O. polyrhachis-furcata*, *O.*

camponoti-leonardi and *O. camponoti-saundersi* (identified based on host identity as a first proxy) across Thailand, and sequenced their genomes. We investigated their species status using the criterion of concordance between genealogies performed using multiple genomic regions. We also studied the genetic subdivision within each species, possibly related to geography, as well as possible clines in diversity, and the possibility of introgression between species through analyses of Bayesian clustering and allelic spectrum-based models. Finally, we also searched for footprint of recombination to gain insight into the mode of reproduction.

2. Materials and Methods

2.1 Sampling, isolation and sequencing

A total of 59 samples of *O. unilateralis s.l.* were collected in Thailand, 29 on *C. leonardi*, 16 on *C. saundersi* and 14 on *P. furcata* (Fig. 1a and Table 1S); fungal samples were named according to their host species (i.e., *O. polyrhachis-furcata*, *O. camponoti-leonardi* and *O. camponoti-saundersi,* respectively*)*. To conform to the Nagoya protocols on access and benefit-sharing, we obtained an authorization from the Department of National Parks, Wildlife and Plant Conservation (DNP) at the Ministry of Natural Resources and Environment of Thailand for all strain collections. Isolations and extractions of fungal DNA were conducted following Kobmoo et al. (2018). Basically, ascospores were harvested after being discharged from the sexual reproductive structures (perithecia) of each sample on potato dextrose agar; whenever possible, random single ascospores were picked up with a needle and transferred to a liquid medium for culturing haploid strains. A single monospore strain per sample was retained for analyses. However, initiating cultures from single spores was not consistently successful; some strains were thus obtained by pooling the discharged ascospores together and thus result from a mixture of products of meiosis from a diploid individual; the corresponding DNAs thus correspond to diploid genotypes. Ascospores are thought to constitute the principal mode of infection for *O. unilateralis* (Araújo and Hughes, 2016) but it is

still unknown whether infections by two mating-type compatible haploid individuals are required to allow for sexual reproduction or whether these fungi are homothallic.

DNA quality and quantity were checked with gel electrophoresis using 0.8 % agarose gel and a Nanodrop (ThermoFisher). Paired-end genomic libraries (150 bp) were constructed and sequenced with an Illumina HiSeq3000 system at the GenoToul platform (Toulouse, France).

2.2 Mapping and detection of single nucleotide polymorphisms

The raw reads were trimmed for adaptors and low-quality bases on both ends (q < 20) with Trim Galore! (Krueger, 2015). Reads shorter than 50 bp were discarded. Processed reads were then initially mapped to the latest version of *O. polyrhachis-furcata* reference (Kobmoo et al., 2018), thanks to its higher quality (43.25 Mb, 68 scaffolds, N50 ~ 2.9 Mb) compared to those of the two other species (*O. camponoti-leonardi*: 37.91 Mb, 531 scaffolds, N50 ~ 140 Kb; *O camponoti-saundersi*: 49.26 Mb, 1,700 scaffolds, N50 ~ 102 Kb), using bwa-mem (Li, 2013) with the options -k 15 -d 150 -c 2000 -D 0.3 -B 1 -O 3 -Y -M. SNPs were detected using GATK (McKenna et al., 2010) through the GATK HaplotypeCaller function. The ploidy parameter was set as haploid for single-spore strains, and as diploid for multiple-spore strains as, for a given strain, a bulk of spores from the sample represents a mix of products of multiple meioses from the same diploid individual. As there were no reference variants, the variant recalibration step was dismissed but the variants were filtered *a posteriori* with total depth between 700x and 5000x, depth for individual samples > 10, mapping quality > 10, quality normalized by depth > 20, genotyping quality > 50; various corrections for biases related to the difference of quality between reference and variant alleles were also applied. Any SNP overlapping with regions including repeats identified beforehand with RepeatMasker (Smit et al., http://www.repeatmasker.org) was excluded. Finally, only biallelic SNPs without indels and with at least 80% samples with reliable inferred variants were retained.

For the within-species analyses, the reads from *O. camponoti-leonardi* and *O. camponoti-saundersi* were mapped to their respective reference (Kobmoo et al., 2018). The same procedures of mapping and variant calling were applied as in *O. polyrhachis-furcata* above with slightly different filtering (for *O. camponoti-leonardi*, final SNPs with total depth between 50x and 3000x, mapping quality > 10, quality normalized by depth > 20; for *O. camponoti-saundersi*, final SNPs with total depth between 50x and 1000x, depth for individual samples > 20, mapping quality > 10, quality normalized by depth > 20).

2.3 Phylogenetic reconstruction and genealogical concordance phylogenetic species recognition (GCPSR)

Whole scaffold sequences for each sample were obtained using the *FastaAlternateReferenceMaker* tool in GATK. Sequences of a given scaffold from different samples altered at the SNPs positions were aligned using an in-house awk script. The alignments of all scaffolds were concatenated to generate a whole-genome alignment. The alignments were used to build maximum-likelihood phylogenetic trees with RaxML based on the model GTRCAT with 1000 bootstrap replicates. A tree for each scaffold with variants was inferred with the same parameters. All trees were rooted with *O. kimflemingiae* (De Bekker et al., 2017), by aligning this latest genome to that of *O. polyrhachis-furcata* using the nucmer command from MUMmer 3 (Delcher, 2002; Kurtz et al., 2004). The concordance between different genomic regions was also evaluated under the framework of Bayesian concordance analyses using BUCKy v.1.4.4 (Larget et al., 2010), after Bayesian inferences over all scaffolds with SNPs bigger than 100 Kb chunked by 100 Kb separately using MrBayes 3.2.2 (Ronquist et al., 2012). The MCMC chains were allowed to sample across the substitution model space. Chunks of 100 Kb were used due to the high linkage disequilibrium found (see the section 3.7). The posterior tree distributions from the Bayesian inferences were fed into BUCKy to estimate the concordance of the nodes among the tree topologies, given as a

concordance factor (CF) ranging from 0 to 1 (absence of concordance to complete concordance). The concordance factor thus represents the genome-wide genealogical support for the monophyly of clades.

2.4 Linkage disequilibrium, recombination footprints

The correlation coefficient r^2 between SNPs of less than 1M bp with minor allele frequency of at least 0.1 and per-sample inbreeding coefficients for diploid samples were calculated with plink (Purcell et al., 2007); r^2 is a correlation coefficient representing the association and thus linkage between SNPs. In case of regular recombination, physically close SNPs are expected to be linked thus to have higher r^2 than farther SNPs. In case of low frequency of sexual reproduction, linkage disequilibrium, and thus r^2, will be low across the whole genome. The r^2 coefficient was averaged over 100 bp intervals and plotted against the pairwise physical distance to obtain linkage disequilibrium (LD) decay plots. Pairwise Homoplasy Index (PHI) tests of recombination (Bruen et al., 2006) were performed for the whole dataset and for each species on each scaffold. The test is based on the detection of incompatibility between pairs of sites regarding whether there is a genealogical history that can be inferred parsimoniously that does not involve any convergent mutations; if any incompatibility exists, under an infinite-site model of evolution, it must come from recombination. PHI-test returns the probability of observing the data under the null hypothesis of no recombination. Bonferroni corrections were applied to the p-values. Population-scaled recombination rates (rho = 2Ne × r; Ne = effective population size, r = recombination per site per generation) were calculated for each species using the R package FastEPRR (Gao et al., 2016). The per-sample inbreeding coefficients were calculated for diploid individuals as method-of-moments estimate F coefficients based on observed and expected homozygosity for each diploid sample.

Neighbor-net networks were inferred using SplitsTree (Huson and Bryant, 2006) for the whole dataset and for each species to gain insights into the level of recombination between and within the various *O. unilateralis* species.

2.5 Principal component analysis (PCA), molecular diversity and genetic differentiation

The nucleotide diversity (Pi) of all species/populations and inter-specific/inter-population F_{ST} and d_{xy} values were calculated using PopGenome package in R (Pfeifer et al., 2014). These calculations were performed based on SNPs called using the *O. polyrhachis-furcata* reference to use a common set of SNPs. The correlation between Pi and the latitude of the populations was tested with a Spearman's correlation test. We investigated the pattern of isolation by distance (IBD) by testing the existence of a correlation between $F_{ST}/(1- F_{ST})$ and the log of geographic distance (Rousset, 1997), using a Mantel test with 10,000 permutations.

2.6 Genetic clustering, inference of gene flow and introgression

A Bayesian clustering was conducted via FastStructure (Raj et al., 2014) with admixture model for the whole dataset to infer genetic structures between species. At the intra-specific level, we conducted the same analysis using the SNPs generated from each species reference. The most relevant number of clusters (K) were determined based on the plots of marginal likelihood and visual assessment of the barplots. The introgression footprints along the genome were investigated with a clustering analysis of genetic similarity on 500 SNPs sliding windows among the samples based on Gower's distance, using Mclust package in R (Scrucca et al., 2016). Neighbor-joining trees based on the F84 distance (Kishino and Hasegawa, 1989)(Kishino and Hasegawa, 1989) were also constructed to illustrate the relationships among the different clusters.

To infer events of gene flow between species, we used TreeMix (Pickrell and Pritchard, 2012) which estimates the most likely topology representing the evolutionary history between

populations and mixing based on genome-wide allele frequency data. TreeMix compares the likelihood of models with different numbers of migration events between populations.

3. Results

3.1 Detection of single nucleotide polymorphisms (SNPs)

By using *O. polyrhachis-furcata* BCC54312 as the reference for mapping all genomes, we retained 688,264 SNPs in the whole dataset. Only a small proportion of them was shared between the three host species (1144 SNP ~ 0.166 %, Fig.S1). Based on the phylogenetic species recognition analysis (see the subsection 3.2), we excluded a sample from the host *C. leonardi* with an unclear placement (NK122ss). There were 443,670 SNPs in the remaining dataset. The samples classified as *O. camponoti-leonardi* had more SNPs (219,017) than the other species (92,133 for *O. camponoti-saundersi* and 53,552 for *O. polyrhachis-furcata*). By using the *O. polyrhachis-furcata* reference genome to detect SNPs in all species, we could expect to find higher diversity in *O. polyrhachis-furcata* due to a bias favoring detection of SNPs from this species. However, we observed in contrast fewer SNPs in *O. polyrhachis-furcata*, which is likely due to genuine lower genetic diversity in this species (see section 3.4). When using their respective references to call SNPs within each species, we found as expected more SNPs than when using the *O. polyrhachis-furcata* reference in the two other species (1,301,966 SNPs for *O. camponoti-leonardi* and 159,449 SNPs for *O. camponoti-saundersi)*.

3.2 Species delimitation using the phylogenetic criterion of species recognition

The maximum likelihood (ML) tree based on whole genomes supported the existence of the three host-specific species. Indeed, we found three clades supported by 100% bootstrap support corresponding to the three hosts sampled (Fig.2a). The only exceptions were NK604, isolated from a *C. saundersi* ant but placed within the *O. camponoti-leonardi* clade, and NK122ss, isolated from

C. leonardi but placed as a sister group to the *O. camponoti-saundersi* clade with full bootstrap support (Fig.2a). The trees inferred from the different scaffolds displayed congruent topologies (Fig.2b) except for the scaffolds 6.2, 16, 18, 20, 21, 25, 29, 36, 41, 53, 105 (Fig.S2). Some of these scaffolds however suffered from a lack of polymorphism and/or yielded trees with low bootstrap supports (Table 2S).

The Bayesian concordance analysis comparing topologies between genealogies built from 100 kb fragments showed high concordance factors (CFs above 0.9; Fig. 3a), further indicating that most genomic regions supported the three host-specific clades and thus the existence of distinct, host-specialized species. The Bayesian concordance analysis also confirmed the position of NK122ss as the sister group of *O. camponoti-saundersi* (Fig. 3a).

3.3 Inter and intra-specific genetic differentiation

Based on the SNPs inferred using *O. polyrhachis-furcata* as the reference, the Bayesian clustering also yielded clusters grouping samples according to host species, except again for NK604 (Fig.1c). The sample NK122ss, from the host *C. leonardi*, had admixed membership up to K=4 but gained a total membership to a distinct cluster from K=5, where two other clusters were distinguished in this host (Fig.S3a), a pattern corroborated by PCA (Fig.1b). The sample NK122ss may belong to a species different from *O. camponoti-leonardi*, as also suggested by the phylogenies, and was thus excluded from further analyses. No additional cluster was found at higher Ks (Fig.S3a). The marginal likelihood curve in fact reached a plateau at K = 5 (Fig.S3b). The sample NK604, isolated from the host *C. saundersi*, clustered with those from the host *C. leonardi* in the PCA (Fig.1b), in agreement with the results on species delimitation, and was thus considered to belong to *O. camponoti-leonardi* in the following analyses. It is likely a spill over, i.e. an occasional infection from a non-endemic species without sustainable persistence. Furthermore, two genetic clusters could be clearly distinguished within *O. camponoti-leonardi* (OCL1 and OCL2 in

Fig.S1b and 1c). The genetic differentiation between the three main species was high (F_{ST}> 0.5, Table 1).

Table 1. Genetic differentiation between species (F_{ST}: above the diagonal, d_{xy}: below the diagonal; OCL = *Ophiocordyceps camponoti-leonardi*, OCS = *O. camponoti-saundersi*, OPF = *O. polyrhachis-furcata*). The nucleotide diversity (Pi) within species is given in the diagonal in bold; dxy = absolute divergence between two species; F_{ST} = fixation index measuring relative divergence.

F_{ST} / d_{xy}	OCL	OCS	OPF	*Ophiocordyceps* sp. nov. (NK122ss)
OCL	**0.026**	0.696	0.588	0.865
OCS	0.224	**0.015**	0.771	0.902
OPF	0.127	0.162	**0.007**	0.959
Ophiocordyceps sp. nov. (NK122ss)	0.309	0.269	0.257	**0**

We inferred genetic clusters within each of the three main species based on SNPs called using specific reference per species. For *O. camponoti-leonardi*, the results were slightly different from those when the *O. polyrhachis-furcata* reference was used (see above), as we retrieved three clusters (OCL1, OCL2, OCL3) for most K values (Fig.4a and Fig.S4a). Only at K=6 were four clusters observed but the fourth disappeared at higher K values and the marginal likelihood was lowest at K=6 (Fig.S4b). We therefore considered three genetic clusters in *O. camponoti-leonardi*; two clusters were parapatric–one in the extreme North (OCL1) and another one extended to the center (OCL2). The samples NK140 and NK561, which appeared to belong to OCL2 using SNPs called with the *O. polyrhachis-furcata* reference (Fig.1b and 1c), were assigned differently when using SNPs called with the *O. camponoti-leonardi* reference; the sample NK140 displayed an

admixed genetic membership but still mostly belonged to OCL2 while the sample NK561 from the south of Thailand was assigned to another cluster (OCL3).

Regarding *O. camponoti-saundersi*, three sympatric clusters (OCS1, OCS2, OCS3) were found from K=3 and above, with a plateau reached at K=3 for the marginal likelihood (Fig. S5); the three clusters were localized mainly in the north-east of Thailand (Fig. 4b). For *O. polyrhachis-furcata*, the marginal likelihood oscillated between K=2 and K=10 (Fig. S6b); we identified three main genetic clusters, one in the north-east of Thailand (OPF1) and two in the South (OPF2 and OPF3), OPF3 being however represented by a single strain (NK157) (Fig. 4c and Fig. S6a). Another individual from the South of Thailand (NK151) had different cluster assignments across the various K values (Fig.4c and Fig.S6a). NJ trees and PCAs retrieved the same patterns of genetic subdivisions within each of the three species (Fig.S7 - S9). The intra-specific differentiation level (between clusters of a given named species) was overall slightly lower than inter-specific differentiation level (Table 2).

Table 2. Genetic differentiation between genetic clusters (F_{ST}: above the diagonal, d_{xy}: below the diagonal). The nucleotide diversity (Pi) within clusters are given in the diagonal in bold. OCL3 includes a single individual (NK561).

Fst / d_{xy}	OCL1	OCL2	OCL3	OCS1	OCS2	OCS3	OPF1	OPF2
OCL1	**0.011**	0.711	0.907	0.934	0.915	0.904	0.884	0.870
OCL2	0.144	**0.016**	0.781	0.877	0.853	0.841	0.749	0.736
OCL3	0.175	0.117	**0**	0.996	0.976	0.962	0.985	0.966
OCS1	0.262	0.216	0.246	**0.001**	0.913	0.834	0.979	0.962
OCS2	0.261	0.214	0.244	0.079	**0.010**	0.813	0.947	0.931
OCS3	0.268	0.221	0.250	0.063	0.081	**0.006**	0.927	0.912
OPF1	0.159	0.110	0.144	0.155	0.152	0.159	**0.001**	0.425
OPF2	0.164	0.116	0.149	0.161	0.158	0.165	0.013	**0.003**

3.4 Nucleotide diversity and isolation by distance

We used the SNPs inferred using the best assembled reference genome (*O. polyrhachis-furcata)* to compute nucleotide diversity (Pi) to compare between species (Table 1) as well as between intra-specific clusters (Table 2). The nucleotide diversity was highest for *O. camponoti-leonardi* (0.026), followed by *O. camponoti-saundersi* (0.015) and *O. polyrhachis-furcata* (0.007). Based on the Bayesian clustering, we considered *O. camponoti-leonardi* to be composed of two genetic clusters, OCL1 and OCL2 (excluding the strain NK561 that appeared to belong to a distinct cluster; Fig. S4a and S7), *O. camponoti-saundersi* to be composed of three genetic clusters (OCS1, OCS2 and OCS3), and *O. polyrhachis-furcata* to be composed of two clusters (OPF1 and OPF2), excluding the two genetically distant individuals NK151 and NK157. The genetic diversity was variable among genetic clusters within each species (Pi values in Table 2). OCS1 had much lower diversity compared to OCS2 and OCS3, as shown by the Pi values and the shorter branches on the NJ tree (Fig. S8a). The correlation between genetic diversity (Pi) and latitude for *O. camponoti-leonardi* was not significant (Spearman's correlation test: r = -0.3, p-value = 0.683; Table 3S and Fig. 10Sa). Isolation by distance was not significant in *O. camponoti-leonardi* (Mantel's test: r = 0.236, p-value = 0.25; Fig.10Sb). Not enough sampling sites were available for the other species to test for such correlations.

3.5 Concordance among genealogies for within-species clusters

We then investigated whether the different clusters found within each of the identified cryptic species so far could be considered as distinct species based on the criterion of concordance among multiple genomic genealogies. The Bayesian concordance analysis based on 100-kb DNA fragments showed that the three clusters in *O. camponoti-saundersi* formed respectively clades with high CFs (above 0.85; Fig. 3b). For *O. camponoti-leonardi*, the clade OCL1 was strongly supported with a high CF (0.993) while the clade OCL2 as defined previously was poorly supported (CF =

0.477; OCL2 *sensu lato* in Fig.3b). However, when excluding NK140 from OCL2, the clade became strongly supported (CF = 0.901), while NK140, like NK561 which was previously shown to be admixed, had unclear status. For *O. polyrhachis-furcata*, the clusters OPF1 and OPF2 formed each a monophyletic clade but with low CFs (0.433 and 0.485 respectively: Fig.3b). Neighbor-net networks using SplitsTree also recovered different species (Fig.5) and the same well differentiated genetic clusters within species (Fig.S11).

3.6 Introgression and gene flow between species

Previous analyses showed that a few individuals had admixed genetic membership. We therefore performed genetic clustering analyses along the genome to identify possible footprints of localized introgression. The general pattern of host-specific clusters was found throughout the genome (Fig.6). However, introgression footprints was found along the genome for two individuals, NK151 and NK157, that showed admixed or distinct membership in comparison to the other *O. polyrhachis-furcata* samples in Bayesian clustering. The *O. camponoti-leonardi* individual NK561 appeared here highly admixed between OCL1 and OCL2. We did not recover in this analysis the different genetic clusters in *O. camponoti-saundersi* (Fig.6). The Gower's genetic distance on which this clustering is based may not be able to detect finer genetic differences. Splitstree analyses recovered the same well differentiated genetic clusters within species and without footprints of extensive gene flow, except perhaps between OPF1 and OPF2 (Fig.S11).

Gene flow events were inferred using Treemix based on the allelic frequency spectra of the different species and clusters retrieved from the SNPs called on the *O. polyrhachis-furcata* reference genome. We excluded the individual with unclear status, i.e., NK122ss. To account for linkage disequilibrium, blocks of 500 SNPs were used. Setting the number of migration events to one, we inferred one gene flow event between OCL1 and OCS (Fig.7). However, the model with

gene flow was not significantly more likely than the model without migration (log likelihood ratio test: p-value 0.235).

3.7 Linkage disequilibrium and recombination footprints

We investigated recombination footprints within species and genetic clusters by looking at linkage disequilibrium (LD) decay (Fig.8). None of the species had their LD decaying to less than 0.2 at the distance of 1 Mb, and the LD of both *O. camponoti-leonardi* and *O. camponoti-saundersi* remained flat after an initial rapid decrease, resembling patterns in clonal species (Sepúlveda et al., 2017). Using half of the maximum LD as the criteria for linkage distance (Taylor et al., 2015), *O. unilateralis* species seemed to maintain linkage over more than 1 Mb. Nevertheless, the PHI-tests revealed significant deviations from the expectation of no recombination for most scaffolds in *O. camponoti-leonardi* and *O. saundersi,* and their respective intra-specific genetic clusters. Evidence of recombination was found on fewer scaffolds for *O. polyrhachis-furcata,* both when considering the whole species and considering separately its intra-specific clusters (Table S4). For all species and their clusters, no footprints of recombination could be recovered from small scaffolds with little polymorphism (Table S4). The population-scaled recombination rates (rho) were all of similar order of magnitude (1.88e-5 for *O. polyrhachis-furcata,* 1.49e-5 for *O. camponoti-leonardi* and 5.95e-5 for *O. camponoti-saundersi).* Splitstree analyses also suggested low levels of recombination within species (Fig.4) and within genetic clusters (Fig.S11); reticulations were observed but seemed overall limited compared to the whole network except for *O. polyrhachis-furcata,* in which the reticulations appeared to be restricted to the two admixed samples (NK151 and NK157) (Fig.S11b). The per-sample inbreeding coefficients of the diploid samples were extremely high in all species, approaching 1; this means that most sites were homozygous (Table.S5), suggesting a very high level of inbreeding.

4. Discussion

4.1 Host-specific species with occasional host-jump and gene flow

Our population genomics study on Thai populations of *O. unilateralis sensu lato* from three ant hosts showed that, apart from a few exceptions, fungal species matched ant species. The previous separation of host-specific species within *O. unilateralis s.l.* from Thailand (*O. camponoti-leonardi*, *O. camponoti-saundersi* and *O. polyrhachis-furcata*) had strong whole-genome support with different methods, i.e., Bayesian clustering, PCA, maximum likelihood bootstrap and Bayesian concordance factors (BCF). In particular, the tree topologies obtained based on different genomic regions were congruent and supported the monophyly of individuals parasitizing a given host species, indicating that the three species *O. camponoti-leonardi*, *O. camponoti-saundersi* and *O. polyrhachis-furcata* form distinct lineages under the phylogenetic species criterion.

Nevertheless, one individual was found in an ant species different from the typical host of its species, as previously documented (Kobmoo et al., 2012), and two others showed signals of genetic admixture in Bayesian clustering which were also reflected into ambiguous or poorly supported placements in phylogenetic trees. Using an allelic frequency spectrum-based inference, a single event of gene flow from *O. camponoti-leonardi* to *O. camponoti-saundersi* was detected; however, the model with gene flow was not statistically more likely than the model without gene flow. Occasional spill-over or hybridization thus did not seem to lead to long term introgression among species, potentially due to selection against hybrids.

4.2 Possible additional cryptic species and the one ant-one fungus paradigm

Our population genomics study further revealed the existence of differentiated genetic clusters within each host-specific species. The clusters within *O. camponoti-leonardi* and *O. camponoti-saundersi* were particularly well differentiated and supported through different analyses, with little footprints of gene flow. Multiple genealogies were congruent in separating the genetic

clusters within *camponoti-saundersi* and *O. camponoti-leonardi* indicating they should be considered as distinct species according to the GCPSR criterion (Taylor et al., 2000), and not just as different populations, and they would deserve further morphological examinations in future studies. The clusters within *O. polyrhachis-furcata* in contrast did not appear supported as distinct species.

This study thus revealed hidden diversity beyond well described species, which could not be detected previously (Kobmoo et al., 2015, 2012), probably due to the limited quantity of markers and examined samples (Balasundaram et al., 2015; Taylor et al., 2000). In addition, a sample (NK122ss) did not cluster with other individuals parasitizing the same host species and was instead placed as a sister taxon to *O. camponoti-saundersi*, without presenting admixture footprints, suggesting that it could belong to a different species or be admixed with another species not present in the sample.

The finding that there could be more than one *Ophiocordyceps* species parasitizing a given host challenges the well-accepted "one ant-one fungus" paradigm (Araújo et al., 2018). It should however be noted that the ant taxonomy is itself still debated. The ants *Colobopsis leonardi* and *C. saundersi* were previously classified in the genus *Camponotus* before a recent resurrection of the genus *Colobopsis* (Ward et al., 2016) and both are cryptic species with unclear taxonomy (Laciny et al., 2018). The genetic clusters found within described *Ophiocordyceps* species in our study could actually be associated to different ant species in agreement with the one ant-one fungus hypothesis. A taxonomic work with a strong population genomics aspect on both the hosts and the parasites will be necessary to elucidate species delimitation and cophylogeny analyses to assess whether co-divergence occurred (De Vienne et al., 2013; Morand et al., 2015).

4.3 Reproductive mode and mating system in **Ophicordyceps unilateralis**

Recombination likely occurs in *O. unilateralis* sensu lato, but probably very rarely. The rho values were of the same order of magnitude (-5) for all three host-specific species and much lower

than in fungi with regular outcrossing such as the heterothallic *Zymoseptoria tritici* and *Z. ardabilae* (rho values being respectively 0.0217 and 0.0045: Stukenbrock and Dutheil, 2018), or other fungi with diploid-selfing obligate sexual reproduction like *Microbotryum* species (order of magnitude -3 to -2: Badouin et al., 2017) and *Saccharomyces paradoxus* (0.001 – 0.003: Tsai et al., 2008). The per-sample inbreeding coefficients for the diploid individuals were actually all very high in all species and clusters, indicating that most SNPs were homozygous. As all the strains used in this study were isolated from sexual spores resulting from meiosis, *O. unilateralis* regularly undergoes sexual reproduction. Altogether, these findings suggest for the whole species complex a high level of inbreeding or even frequent intra-haploid mating, i.e. mating between genetically identical haploid cells (Billiard et al., 2012; Giraud et al., 2008). Intra-haploid mating is only possible in homothallic fungi as the existence of mating types in heterothallic fungi prevents such mating. Whether *O. unilateralis* is homothallic or heterothallic is still unclear; the genome of *O. polyrhachis-furcata* contains at least a mating type locus (MAT-1-2) and a putative mating type-switching gene (Wichadakul, unpublished data) which suggests that it may be homothallic but no experimental observation has been done.

Such high inbreeding is consistent with the life history traits of these fungi. Fungal development into the transmission stage inside ant nest was not observed and infection outside the nest is likely favored to avoid social immunity (Loreto et al., 2014). Field observations and a theoretical model established that effective infections were rare; only a weak proportion of samples observed on field went through sexual reproduction (Andersen et al., 2012; Mongkolsamrit et al., 2012). The probability to encounter compatible spores from distinct diploid fungal individuals may be low, thus promoting selfing in the fungi. Low probability of outcrossing can be found in other pathogenic fungi such as the plant castrating anther-smut fungi *Microbotryum* in which mating occurs between cells resulting from the same meiosis issued from a dispersing diploid spore (Giraud et al., 2008; Vercken et al., 2010). The infective spores of *O. unilateralis* are haploid which could

render the probability to find a compatible mate in an insect even lower, favoring intra-haploid mating.

4.4 The biogeography of **Ophiocordyceps unilateralis** *in Thailand*

While the genetic clusters detected within *O. camponoti-saundersi* were sympatric and all restricted to the Northeast of Thailand, *O. camponoti-leonardi* and *O. polyrhachis-furcata* had widely distributed genetic clusters, with different geographical ranges, mainly reflecting the separation between the north/center and the south of Thailand. Such genetic differentiation between the North and the South of Thailand has been documented in many organisms (Chen et al., 2018; Karanth et al., 2008; Pramual et al., 2005; Sittipraneed et al., 2001; Tu et al., 2017; Yu et al., 2019), which was suggested to reflect different biogeographic entities, one from the Indochina and another from the Malay peninsula (Woodruff, 2010). The existence of northern and southern distinct biogeographic compartments also explains that the genetic diversity does not increase toward south, contrary to what is found in the northern hemisphere for many organisms (Hewitt, 2012). Enlarging the sampling of both the hosts and the pathogens across Southeast Asia would give novel insights into the host-pathogen phylogeography of this region.

5. Conclusion

In this study, population genomics supported the host-specificity of previously described cryptic species in ant pathogenic fungi and further revealed genetic clusters sharing the same host but with contrasting geographical ranges as candidates to be erected as new species in the zombie ant fungal complex *O. unilateralis* sensu lato. Whole genome analyses also revealed individuals with admixed genetic membership but not leading to persistent gene flow among species or genetic clusters. The recognition of cryptic species has direct impacts on species diversity assessment and ecological research and population genomics can be powerful to reveal hidden diversity, as shown

in our study. More generally, population genomics can reveal the existence of cryptic species or different populations of pathogens, contributing to our understanding of host-parasite/pathogen specificity and phylogeography. Furthermore, we studied for the first time recombination footprints in *O. unilateralis* sensu lato which suggested limited recombinations and a high level of inbreeding.

Acknowledgements

We would like to thank Alodie Snirc for advice concerning DNA extraction, Antoine Branca for help with bioinformatic protocols; Kanoksri Tasanathai, Donnaya Thanakitpipattana, Wasana Noisripoom and Artit Khonsanit for help in samples collection; Navavit Ponganan for making the maps in this publication. We also thank the Department of National Parks, Wildlife and Plant Conservation of Thailand for their cooperation and support of our project.

Funding: This work was supported by a Marie-Skłodowska Curie Action [grant number 655278] to NK, a Biotec Fellow's Research Grant [grant number P1950231] to JJL, and by the National Science and Technology Development Agency, Cluster and Program Management Office [grant number P1300854] to SM.

Supplementary Materials

Supplementary Figures.

Table S1: Fungal isolates used in this study.

Table S2: Genetic diversity per scaffold in *Ophiocordyceps unilateralis sensu lato.*

Table S3: Genetic diversity by site and by species.

Research Data

The raw reads in this study were deposited at the NCBI Sequence Read Archive (SRA) under the accession PRJNA530395. The information regarding the fungal isolates used in this study were deposited at the NCBI BioSample under the accessions reported in the table S1 of the Supplementary Materials. The final sets of SNPs are available at…

References

Andersen, S.B., Ferrari, M., Evans, H.C., Elliot, S.L., Boomsma, J.J., Hughes, D.P., 2012. Disease dynamics in a specialized parasite of ant societies. PLoS One 7. https://doi.org/10.1371/journal.pone.0036352

Araújo, J.P.M., Evans, H.C., Kepler, R., Hughes, D.P., 2018. Zombie-ant fungi across continents: 15 new species and new combinations within *Ophiocordyceps*. I. Myrmecophilous hirsutelloid species. Stud. Mycol. 90, 119–160. https://doi.org/10.1016/j.simyco.2017.12.002

Araújo, J.P.M., Hughes, D.P., 2016. Diversity of Entomopathogenic Fungi. Which Groups Conquered the Insect Body? Adv. Genet. 94, 1–39. https://doi.org/10.1016/bs.adgen.2016.01.001

Badouin, H., Gladieux, P., Gouzy, J., Siguenza, S., Aguileta, G., Snirc, A., Le Prieur, S., Jeziorski, C., Branca, A., Giraud, T., 2017. Widespread selective sweeps throughout the genome of model plant pathogenic fungi and identification of effector candidates. Mol. Ecol. 26, 2041–2062. https://doi.org/10.1111/mec.13976

Balasundaram, S. V., Engh, I.B., Skrede, I., Kauserud, H., 2015. How many DNA markers are needed to reveal cryptic fungal species? Fungal Biol. 119, 940–945. https://doi.org/10.1016/j.funbio.2015.07.006

Billiard, S., López-Villavicencio, M., Hood, M.E., Giraud, T., 2012. Sex, outcrossing and mating types: Unsolved questions in fungi and beyond. J. Evol. Biol. https://doi.org/10.1111/j.1420-9101.2012.02495.x

Bruen, T.C., Philippe, H., Bryant, D., 2006. A simple and robust statistical test for detecting the presence of recombination. Genetics 172, 2665–2681. https://doi.org/10.1534/genetics.105.048975

Chan, K.O., Alexander, A.M., Grismer, L.L., Su, Y.-C., Grismer, J.L., Quah, E.S.H., Brown, R.M., 2017. Species delimitation with gene flow: A methodological comparison and population genomics approach to elucidate cryptic species boundaries in Malaysian Torrent Frogs. Mol. Ecol. 26, 5435–5450. https://doi.org/10.1111/mec.14296

Chen, J.M., Poyarkov, N.A., Suwannapoom, C., Lathrop, A., Wu, Y.H., Zhou, W.W., Yuan, Z.Y., Jin, J.Q., Chen, H.M., Liu, H.Q., Nguyen, T.Q., Nguyen, S.N., Duong, T. Van, Eto, K., Nishikawa, K., Matsui, M., Orlov, N.L., Stuart, B.L., Brown, R.M., Rowley, J.J.L., Murphy, R.W., Wang,

Y.Y., Che, J., 2018. Large-scale phylogenetic analyses provide insights into unrecognized diversity and historical biogeography of Asian leaf-litter frogs, genus *Leptolalax* (Anura: Megophryidae). Mol. Phylogenet. Evol. 124, 162–171. https://doi.org/10.1016/j.ympev.2018.02.020

De Bekker, C., Ohm, R.A., Evans, H.C., Brachmann, A., Hughes, D.P., 2017. Ant-infecting *Ophiocordyceps* genomes reveal a high diversity of potential behavioral manipulation genes and a possible major role for enterotoxins. Sci. Rep. 7, 1–13. https://doi.org/10.1038/s41598-017-12863-w

De Vienne, D.M., Refrégier, G., López-Villavicencio, M., Tellier, A., Hood, M.E., Giraud, T., 2013. Cospeciation vs host-shift speciation: Methods for testing, evidence from natural associations and relation to coevolution. New Phytol. 198, 347–385. https://doi.org/10.1111/nph.12150

Delcher, A.L., 2002. Fast algorithms for large-scale genome alignment and comparison. Nucleic Acids Res. 30, 2478–2483. https://doi.org/10.1093/nar/30.11.2478

Fišer, C., Robinson, C.T., Malard, F., 2018. Cryptic species as a window into the paradigm shift of the species concept. Mol. Ecol. 27, 613–635. https://doi.org/10.1111/mec.14486

Fournier, E., Giraud, T., 2008. Sympatric genetic differentiation of a generalist pathogenic fungus, *Botrytis cinerea*, on two different host plants, grapevine and bramble. J. Evol. Biol. 21, 122–132. https://doi.org/10.1111/j.1420-9101.2007.01462.x

Gao, F., Ming, C., Hu, W., Li, H., 2016. New Software for the Fast Estimation of Population Recombination Rates (FastEPRR) in the Genomic Era. G3 Genes|Genomes|Genetics 6, 1563–1571. https://doi.org/10.1534/g3.116.028233

Gerardo, N.M., Jacobs, S.R., Currie, C.R., Mueller, U.G., 2006. Ancient host-pathogen associations maintained by specificity of chemotaxis and antibiosis. PLoS Biol. 4, 1358–1363. https://doi.org/10.1371/journal.pbio.0040235

Gerardo, N.M., Mueller, U.G., Price, S.L., Currie, C.R., 2004. Exploiting a mutualism: Parasite specialization on cultivars within the fungus-growing ant symbiosis. Proc. R. Soc. B Biol. Sci. 271, 1791–1798. https://doi.org/10.1098/rspb.2004.2792

Giraud, T., Yockteng, R., López-Villavicencio, M., Refrégier, G., Hood, M.E., 2008. Mating system of the anther smut fungus *Microbotryum violaceum*: Selfing under heterothallism. Eukaryot. Cell. https://doi.org/10.1128/EC.00440-07

Hewitt, G.M., 2012. Quaternary Genetic consequences of climatic oscillations in the Genetic consequences of climatic oscillations in the Quaternary. https://doi.org/10.1098/rstb.2003.1388

Hughes, D.P., Andersen, S.B., Hywel-Jones, N.L., Himaman, W., Billen, J., Boomsma, J.J., 2011. Behavioral mechanisms and morphological symptoms of zombie ants dying from fungal infection. BMC Ecol. 11, 13. https://doi.org/10.1186/1472-6785-11-13

Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254–267. https://doi.org/10.1093/molbev/msj030

Insuan, S., Deowanish, S., Klinbunga, S., Sittipraneed, S., Sylvester, H.A., Wongsiri, S., 2007. Genetic Differentiation of the Giant Honey Bee (Apis dorsata) in Thailand Analyzed by Mitochondrial Genes and Microsatellites. Biochem. Genet. 45, 345–361. https://doi.org/10.1007/s10528-007-9079-9

Karanth, K.P., Singh, L., Collura, R. V., Stewart, C.B., 2008. Molecular phylogeny and biogeography of langurs and leaf monkeys of South Asia (Primates: Colobinae). Mol. Phylogenet. Evol. 46, 683–694. https://doi.org/10.1016/j.ympev.2007.11.026

Khonsanit, A., Luangsa-ard, J.J., Thanakitpipattana, D., Kobmoo, N., Piasai, O., 2019. Cryptic species within *Ophiocordyceps myrmecophila* complex on formicine ants from Thailand. Mycol. Prog. https://doi.org/10.1007/s11557-018-1412-7

Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. J. Mol. Evol. 29, 170–179. https://doi.org/10.1007/BF02100115

Kobmoo, N., Mongkolsamrit, S., Tasanathai, K., Thanakitpipattana, D., Luangsa-Ard, J.J., 2012. Molecular phylogenies reveal host-specific divergence of *Ophiocordyceps unilateralis* sensu lato following its host ants. Mol. Ecol. 21, 3022–3031. https://doi.org/10.1111/j.1365-294X.2012.05574.x

Kobmoo, N., Mongkolsamrit, S., Wutikhun, T., Tasanathai, K., Khonsanit, A., Thanakitpipattana, D., Luangsa-Ard, J.J., 2015. New species of *Ophiocordyceps unilateralis*, an ubiquitous pathogen of ants from Thailand. Fungal Biol. 119, 44–52. https://doi.org/10.1016/j.funbio.2014.10.008

Kobmoo, N., Wichadakul, D., Arnamnart, N., Rodríguez De La Vega, R.C., Luangsa-ard, J.J., Giraud, T., 2018. A genome scan of diversifying selection in *Ophiocordyceps* zombie-ant fungi suggests a role for enterotoxins in co-evolution and host specificity. Mol. Ecol. 27, 3582–3598. https://doi.org/10.1111/mec.14813

Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., Salzberg, S.L., 2004. Versatile and open software for comparing large genomes. Genome Biol. 5, R12. https://doi.org/10.1186/gb-2004-5-2-r12

Laciny, A., Druzhinina, I.S., Lim, L., Kopchinskiy, A., Zettel, H., Abu Salim, K., Pal, A., Hoenigsberger, M., Javad Rahimi, M., Jaitrong, W., Pretzer, C., 2018. Colobopsis explodens sp. n., model species for studies on “exploding ants” (Hymenoptera, Formicidae), with biological notes and first illustrations of males of the Colobopsis cylindrica group. Zookeys 751, 1–40. https://doi.org/10.3897/zookeys.751.22661

Larget, B.R., Kotha, S.K., Dewey, C.N., Ané, C., 2010. BUCKy: Gene tree/species tree reconciliation with Bayesian concordance analysis Bret. Bioinformatics 26, 2910–2911. https://doi.org/10.1007/s00285-010-0355-7

Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv: 1303.3997v2 [q-bio.GN]. https://doi.org/10.1186/s13756-018-0352-y

Lifjeld, J.T., 2015. When taxonomy meets genomiocs: lessons from a common songbird. Mol. Ecol. 2901–2903. https://doi.org/10.14344/IOC.ML.5.2

Loreto, R.G., Elliot, S.L., Freitas, M.L.R., Pereira, T.M., Hughes, D.P., 2014. Long-term disease dynamics for a specialized parasite of ant societies: A field study. PLoS One 9, e103516. https://doi.org/10.1371/journal.pone.0103516

Luangsa-ard, J.J., Ridkaew, R., Tasanathai, K., Thanakitpipattana, D., Hywel-Jones, N., 2011. Ophiocordyceps halabalaensis: A new species of *Ophiocordyceps* pathogenic to *Camponotus* gigas in Hala Bala Wildlife Sanctuary, Southern Thailand. Fungal Biol. 115, 608–614. https://doi.org/10.1016/j.funbio.2011.03.002

Matute, D.R., Sepúlveda, V.E., 2019. Fungal species boundaries in the genomics era. Fungal Genet. Biol. 103249. https://doi.org/10.1016/j.fgb.2019.103249

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303. https://doi.org/10.1101/gr.107524.110

Miraldo, A., Li, S., Borregaard, M.K., Flórez-Rodríguez, A., Gopalakrishnan, S., Rizvanovic, M., Wang, Z., Rahbek, C., Marske, K.A., Nogués-Bravo, D., 2016. An Anthropocene map of genetic diversity. Science. 353, 1532–1535. https://doi.org/10.1126/science.aaf4381

Molina, M.C., Del-Prado, R., Divakar, P.K., Sánchez-Mata, D., Crespo, A., 2011. Another example of cryptic diversity in lichen-forming fungi: The new species *Parmelia mayi* (Ascomycota: Parmeliaceae). Org. Divers. Evol. 11, 331–342. https://doi.org/10.1007/s13127-011-0060-4

Mongkolsamrit, S., Kobmoo, N., Tasanathai, K., Khonsanit, A., Noisripoom, W., Srikitikulchai, P., Somnuk, R., Luangsa-ard, J.J., 2012. Life cycle, host range and temporal variation of *Ophiocordyceps unilateralis*/*Hirsutella formicarum* on Formicine ants. J. Invertebr. Pathol. 111. https://doi.org/10.1016/j.jip.2012.08.007

Mongkolsamrit, S., Noisripoom, W., Thanakitpipattana, D., Wutikhun, T., Spatafora, J.W., Luangsa-ard, J., 2018. Disentangling cryptic species with isaria-like morphs in Cordycipitaceae. Mycologia. 110, 230–257. https://doi.org/10.1080/00275514.2018.1446651

Montarry, J., Cartolaro, P., Richard-Cervera, S., Delmotte, F., 2009. Spatio-temporal distribution of *Erysiphe necator* genetic groups and their relationship with disease levels in vineyards. Eur. J. Plant Pathol. https://doi.org/10.1007/s10658-008-9343-9

Morand, S., Krasnov, B.R., Littlewood, T.J. (Eds.), 2015. Parasite Diversity and Diversification: Evolutionary Ecology Meets Phylogenetics. Cambridge University Press. https://doi.org/10.1017/CBO9781139794749

Nieuwenhuis, B.P.S., James, T.Y., 2016. The frequency of sex in fungi. Philos. Trans. R. Soc. B Biol. Sci. 371, 20150540. https://doi.org/10.1098/rstb.2015.0540

Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S.E., Lercher, M.J., 2014. PopGenome: An efficient swiss army knife for population genomic analyses in R. Mol. Biol. Evol. https://doi.org/10.1093/molbev/msu136

Pickrell, J.K., Pritchard, J.K., 2012. Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. PLoS Genet. 8, e1002967. https://doi.org/10.1371/journal.pgen.1002967

Pramual, P., Kuvangkadilok, C., Baimai, V., Walton, C., 2005. Phylogeography of the black fly *Simulium tani* (Diptera: Simuliidae) from Thailand as inferred from mtDNA sequences. Mol. Ecol. 14, 3989–4001. https://doi.org/10.1111/j.1365-294X.2005.02639.x

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575. https://doi.org/10.1086/519795

Raj, A., Stephens, M., Pritchard, J.K., 2014. fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets. Genetics 197, 573–589. https://doi.org/10.1534/genetics.114.164350

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. Syst. Biol. 61, 539–542. https://doi.org/10.1093/sysbio/sys029

Rousset, F., 1997. Genetic Differentiation and Estimation of Gene Flow from F-Statistics Under Isolation by Distance. Genetics 145, 1219–1228.

Savary, R., Masclaux, F.G., Wyss, T., Droh, G., Cruz Corella, J., Machado, A.P., Morton, J.B., Sanders, I.R., 2018. A population genomics approach shows widespread geographical distribution of cryptic genomic forms of the symbiotic fungus Rhizophagus irregularis. ISME J. 12, 17–30. https://doi.org/10.1038/ismej.2017.153

Scrucca, L., Fop, M., Murphy, T.B., Raftery, A.E., 2016. mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. R J. 8, 289–317. https://doi.org/10.1177/2167702614534210

Sepúlveda, V.E., Marquez, R., Turissini, D.A., Goldman, W.E., Matute, D.R., 2017. Genome Sequences Reveal Cryptic Speciation in the Human Pathogen *Histoplasma capsulatum*. mBio 8, e01339-17.

Sittipraneed, S., Sihanuntavong, D., Klinbunga, S., 2001. Genetic differentiation of the honey bee (*Apis cerana*) in Thailand revealed by polymorphism of a large subunit of mitochondrial ribosomal DNA. Insectes Sociaux. 48, 266–272. https://doi.org/10.1007/PL00001776

Smit, A.F.A., Hubley, R., Green, P., n.d. RepeatMasker Open-4.0. [WWW Document]. URL http://www.repeatmasker.org

Smith, B.T., Seeholzer, G.F., Harvey, M.G., Cuervo, A.M., Brumfield, R.T., 2017. A latitudinal phylogeographic diversity gradient in birds. PLoS Biol. 15, 1–24. https://doi.org/10.1371/journal.pbio.2001073

Struck, T.H., Feder, J.L., Bendiksby, M., Birkeland, S., Cerca, J., Gusarov, V.I., Kistenich, S., Larsson, K.H., Liow, L.H., Nowak, M.D., Stedje, B., Bachmann, L., Dimitrov, D., 2018. Finding Evolutionary Processes Hidden in Cryptic Species. Trends Ecol. Evol. 33, 153–163. https://doi.org/10.1016/j.tree.2017.11.007

Stukenbrock, E.H., Dutheil, J.Y., 2018. Fine-Scale Recombination Maps of Fungal Plant 208, 1209–1229. https://doi.org/10.1534/genetics.117.300502/-/DC1.1

Taylor, J.W., Hann-Soden, C., Branco, S., Sylvain, I., Ellison, C.E., 2015. Clonal reproduction in fungi. Proc. Natl. Acad. Sci. 112, 8901–8908. https://doi.org/10.1073/pnas.1503159112

Taylor, J.W., Jacobson, D.J., Kroken, S., Kasuga, T., Geiser, D.M., Hibbett, D.S., Fisher, M.C., 2000. Phylogenetic Species Recognition and Species Concepts in Fungi. Fungal Genet. Biol. 31, 21–32. https://doi.org/10.1006/fgbi.2000.1228

Tsai, I.J., Bensasson, D., Burt, A., Koufopanou, V., 2008. Population genomics of the wild yeast *Saccharomyces paradoxus*: Quantifying the life cycle. Proc. Natl. Acad. Sci. 105, 4957–4962. https://doi.org/10.1073/pnas.0707314105

Tu, V.T., Csorba, G., Ruedi, M., Furey, N.M., Son, N.T., Thong, V.D., Bonillo, C., Hassanin, A., 2017. Comparative phylogeography of bamboo bats of the genus *Tylonycteris* (Chiroptera, Vespertilionidae) in Southeast Asia. Eur. J. Taxon. 1–38. https://doi.org/10.5852/ejt.2017.274

Vercken, E., Fontaine, M.C., Gladieux, P., Hood, M.E., Jonot, O., Giraud, T., 2010. Glacial refugia in pathogens: European genetic structure of anther smut pathogens on *Silene latifolia* and *Silene dioica*. PLoS Pathog. 6, e1001229. https://doi.org/10.1371/journal.ppat.1001229

Vincenot, L., Popa, F., Laso, F., Donges, K., Rexer, K.H., Kost, G., Yang, Z.L., Nara, K., Selosse, M.A., 2017. Out of Asia: Biogeography of fungal populations reveals Asian origin of diversification of the *Laccaria amethystina* complex, and two new species of violet *Laccaria*. Fungal Biol. 121, 939–955. https://doi.org/10.1016/j.funbio.2017.08.001

Ward, P.S., Blaimer, B.B., Fisher, B.L., 2016. A revised phylogenetic classification of the ant subfamily Formicinae (Hymenoptera: Formicidae), with resurrection of the genera *Colobopsis* and *Dinomyrmex*. Zootaxa 4072, 343–357. https://doi.org/10.11646/zootaxa.4072.3.4

Weiss, M., Weigand, H., Weigand, A.M., Leese, F., 2018. Genome-wide single-nucleotide polymorphism data reveal cryptic species within cryptic freshwater snail species—The case of the *Ancylus fluviatilis* species complex. Ecol. Evol. 8, 1063–1072. https://doi.org/10.1002/ece3.3706

Woodruff, D.S., 2010. Biogeography and conservation in Southeast Asia: How 2.7 million years of repeated environmental fluctuations affect today's patterns and the future of the remaining

refugial-phase biodiversity. Biodivers. Conserv. 19, 919–941. https://doi.org/10.1007/s10531-010-9783-3

Yu, H., Tian, E., Zheng, L., Deng, X., Cheng, Y., Chen, L., Wu, W., Tanming, W., Zhang, D., Compton, S.G., Kjellberg, F., 2019. Multiple parapatric pollinators have radiated across a continental fig tree displaying clinal genetic variation. Mol. Ecol. https://doi.org/10.1111/mec.15046

Figures Captions

Figure 1. a) Sampling sites of *Ophiocordyceps unilateralis sensu lato* across Thailand (1 = Ban Huai Baba, 2 = Doi Inthanon National Park, 3 = Nam Nao National Park, 4 = Phu Khiao Wildlife Sanctuary, 5 = Khao Yai National Park, 6 = Ko Chang Island, 7 = Khlong Nakha Wildlife Sanctuary, 8 = Khao Luang National Park, 9 = Hala Bala Wildlife Sanctuary). The pie charts represent the proportions of fungal samples collected on different ant species (CL red = *Colobopsis leonardi*, CS green = *Colobopsis saundersi*, PF blue= *Polyrhachis furcata*). b) Principal component analysis (PCA) based on whole genome single nucleotide polymorphisms (SNPs); colors correspond to the host species as in Figure 1a. c) Bayesian clustering based on whole genome single nucleotide polymorphisms (SNPs) (K=5). Vertical bars represent individuals and the colors their assignment probabilities to five genetic clusters. The arrows indicate the samples NK122ss and NK604. Numbers below the barplots represent the sample sites as in Figure 1a.

Figure 2. Maximum likelihood-based tree inferred using a) the whole genome altered at the SNP positions, b) separate scaffolds. The colors at the tips represent the identities of ant hosts (red = *Colobopsis leonardi*, green = *C. saundersi*, blue = *Polyrhachis furcata*). Thick branches correspond to those leading to nodes with 100% bootstrap support.

Figure 3. Bayesian inference-based trees with genome-wide concordance factors and their confidence intervals, supporting monophyletic clades representing a) the three host-specific species

(OCL = *Ophiocordyceps camponoti-leonardi*, OCS = *O. camponoti-saundersi*, OPF = *O. polyrhachis-furcata*), and b) different genetic clusters as inferred from Bayesian clustering (OCL1 and OCL2 = *O. camponoti-leonardi* clusters 1 and 2; OCS1, OCS2 and OCS3 = *O. camponoti-saunders*i clusters 1, 2 and 3; OPF1 and OPF2 = *O. polyrhachis-furcata* clusters 1 and 2). Concordance factors (CF) can range from 0 to 1 (absence of concordance to complete concordance). In the panel a (left), the subclades identified in the panel b (right) were collapsed to better show the support for OCS, OPF and OCL.

Figure 4. Barplots of Bayesian clustering using single nucleotide polymorphisms (SNPs) inferred using conspecific genome references, with the geographical location of samples assigned to the different clusters, for each of the three species at K = 3: a) *Ophiocordyceps camponoti-leonardi*, b) *O. camponoti-saundersi* and c) *O. polyrhachis-furcata*.

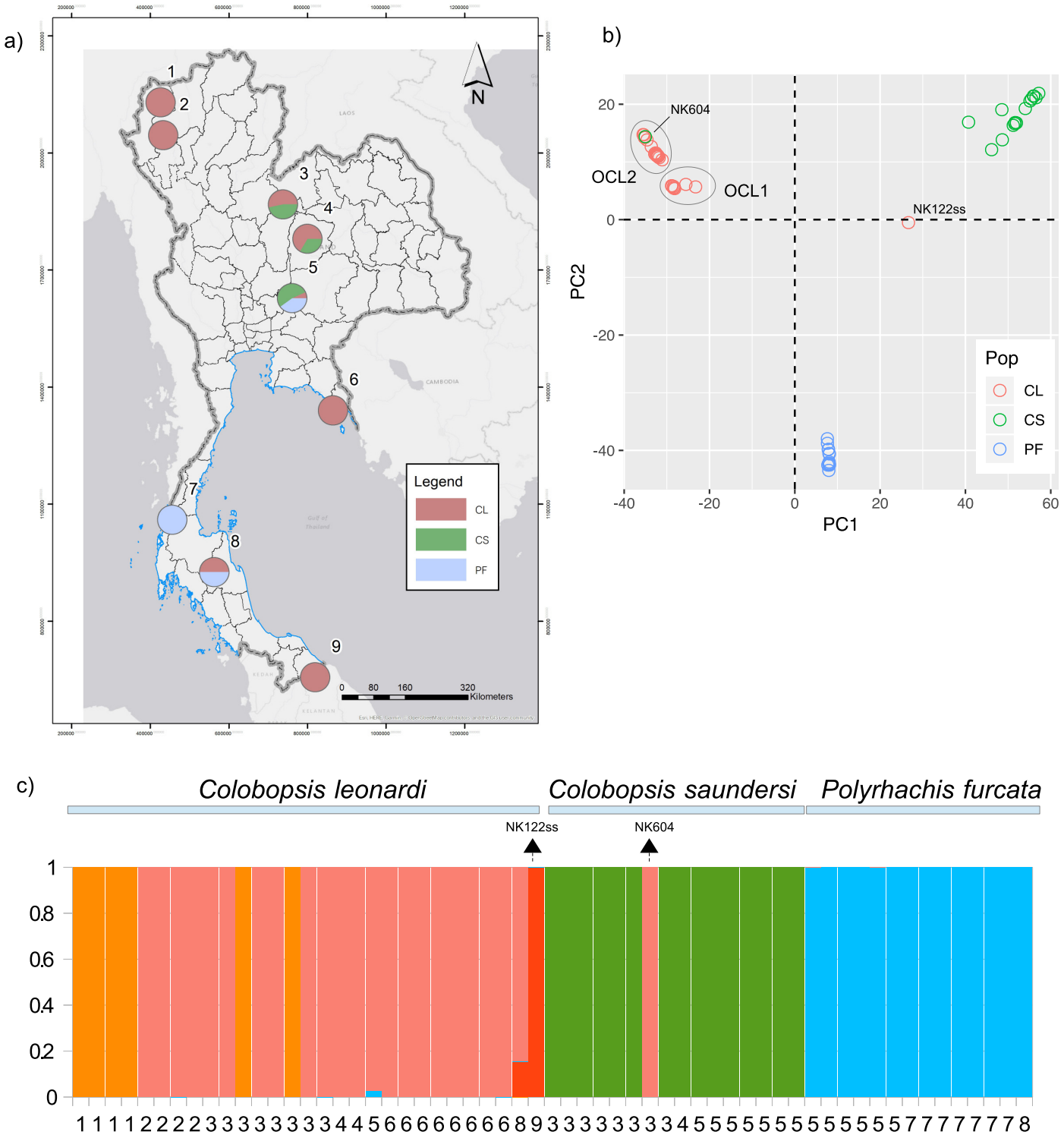
Figure 5. A neighbor-net phylogenetic network of *Ophiocordyceps unilateralis* (OCL = *O. camponoti-leonardi*, OCS = *O. camponoti-saundersi*, OPF = *O. polyrhachis-furcata*).

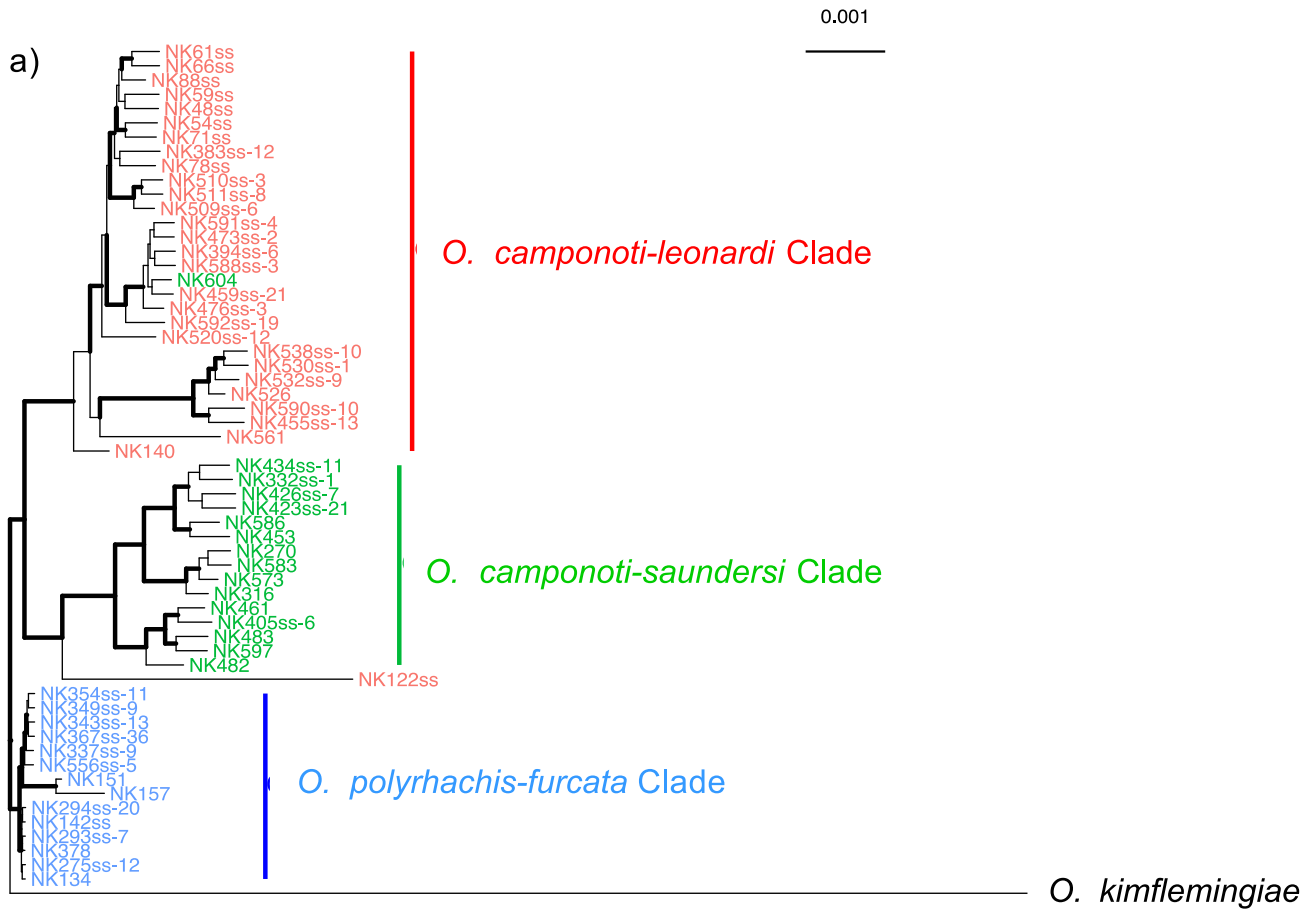
Figure 6. Clustering of *Ophiocordyceps unilateralis sensu lato*, based on Gower's distance, along the *O. polyrhachis-furcata* BCC54312 reference genome using 500 SNPs sliding windows (OCL = *O. camponoti-leonardi*, OCS = *O. camponoti-saundersi*, OPF = *O. polyrhachis-furcata*). The red asterix (★) indicates the sample NK561 with a footprint of admixture between OCL1 and OCL2. The red arrows (←) point to the samples NK151 and NK157 of OPF which have a clear signature of admixture.

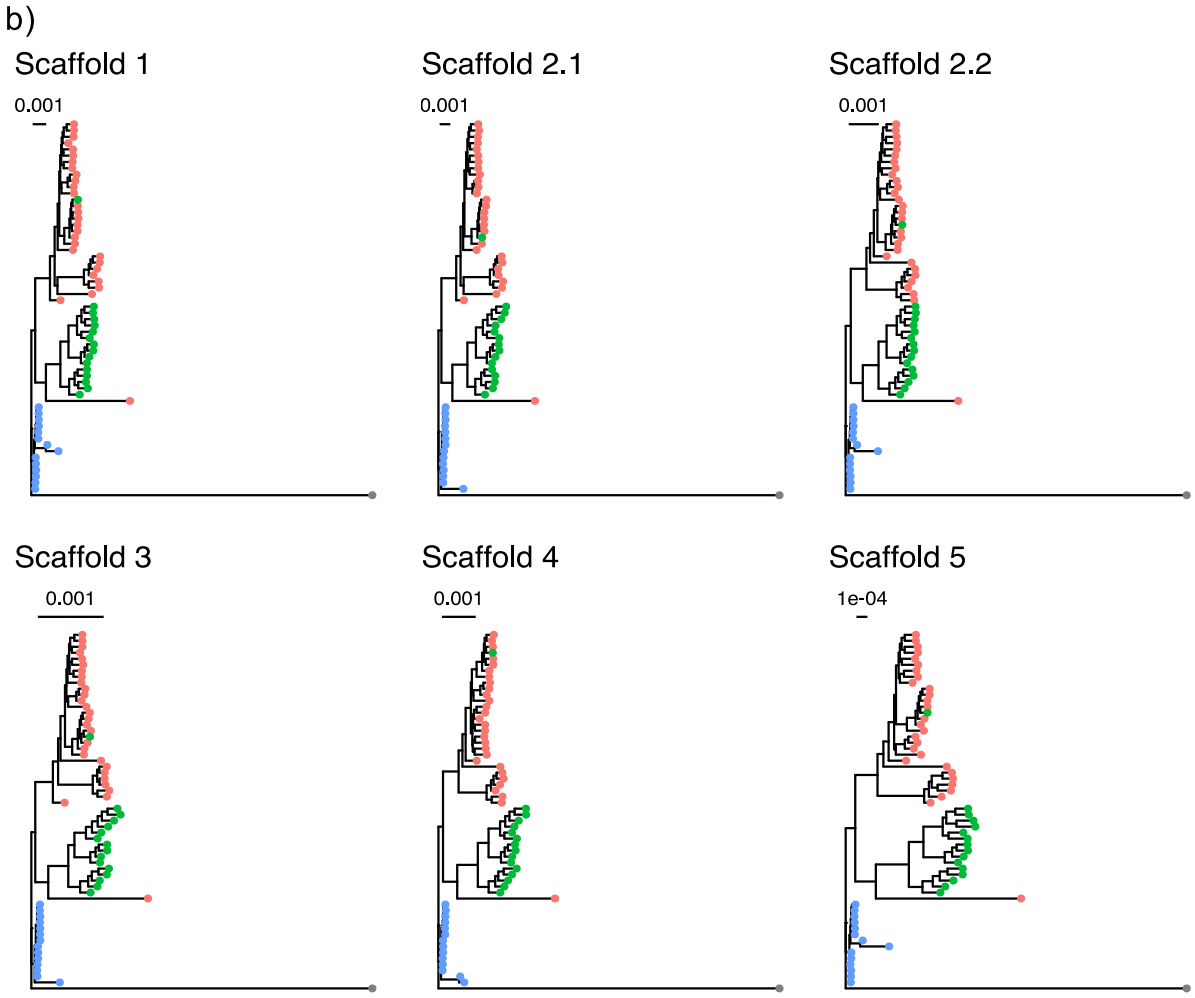
Figure 7. Results of TreeMix. a) Best-fitting genealogy without admixture (m=0) for different *Ophiocordyceps unilateralis* species and clusters as inferred from the variance-covariance matrix of

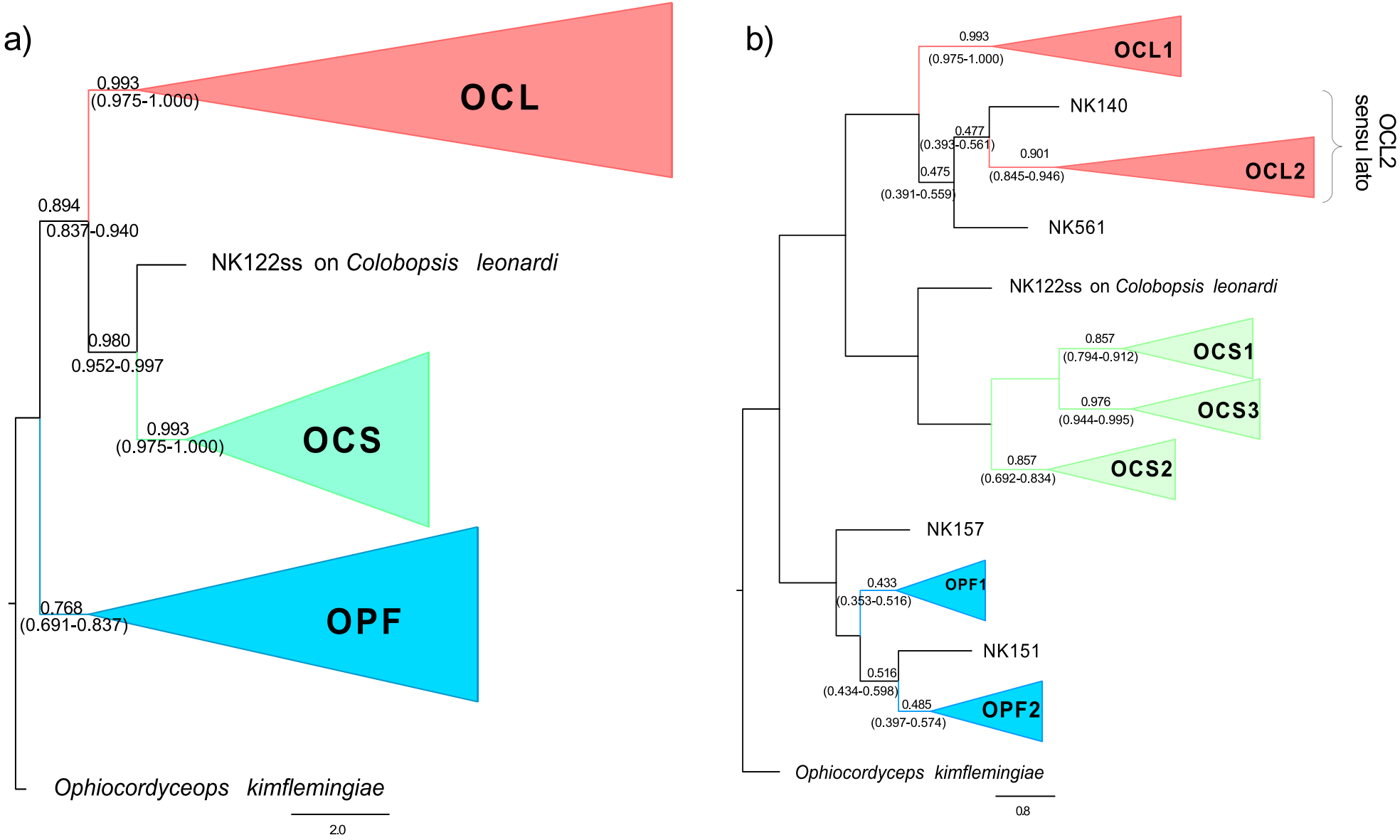
the genome-wide alleles frequencies based on the SNPs using *O. polyrhachis-furcata* as the reference. b) A scenario of one event of migration (m=1). The model a) is nested within b). We estimated the likelihood of both scenarios and used a log likelihood ratio test (LRT) to determine the best-fitting scenario.

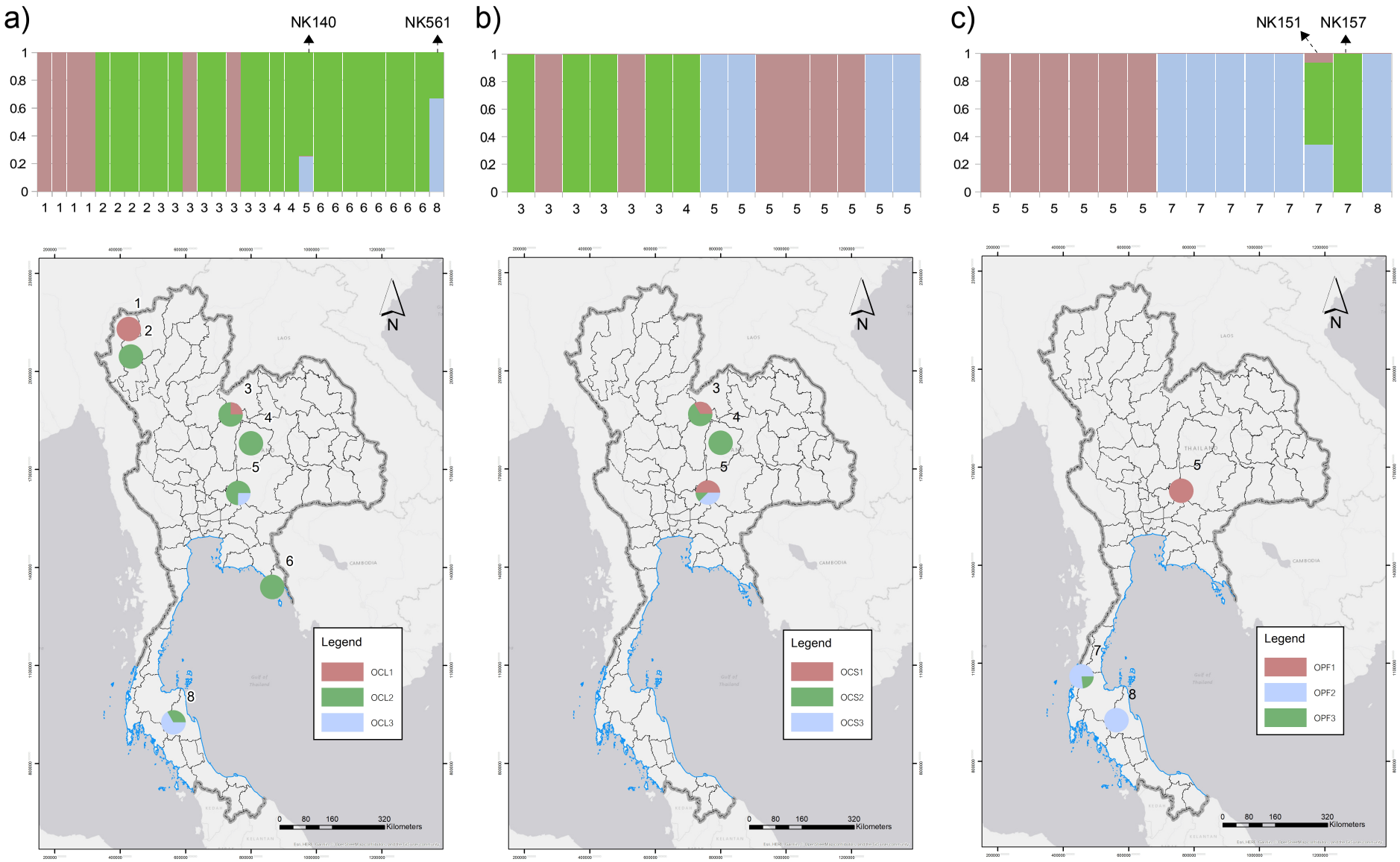
Figure 8. Linkage disequilibrium (LD) decay of *Ophiocordyceps camponoti-leonardi* (red), *O. camponoti-saundersi* (green) and *O. polyrhachis-furcata* (blue).

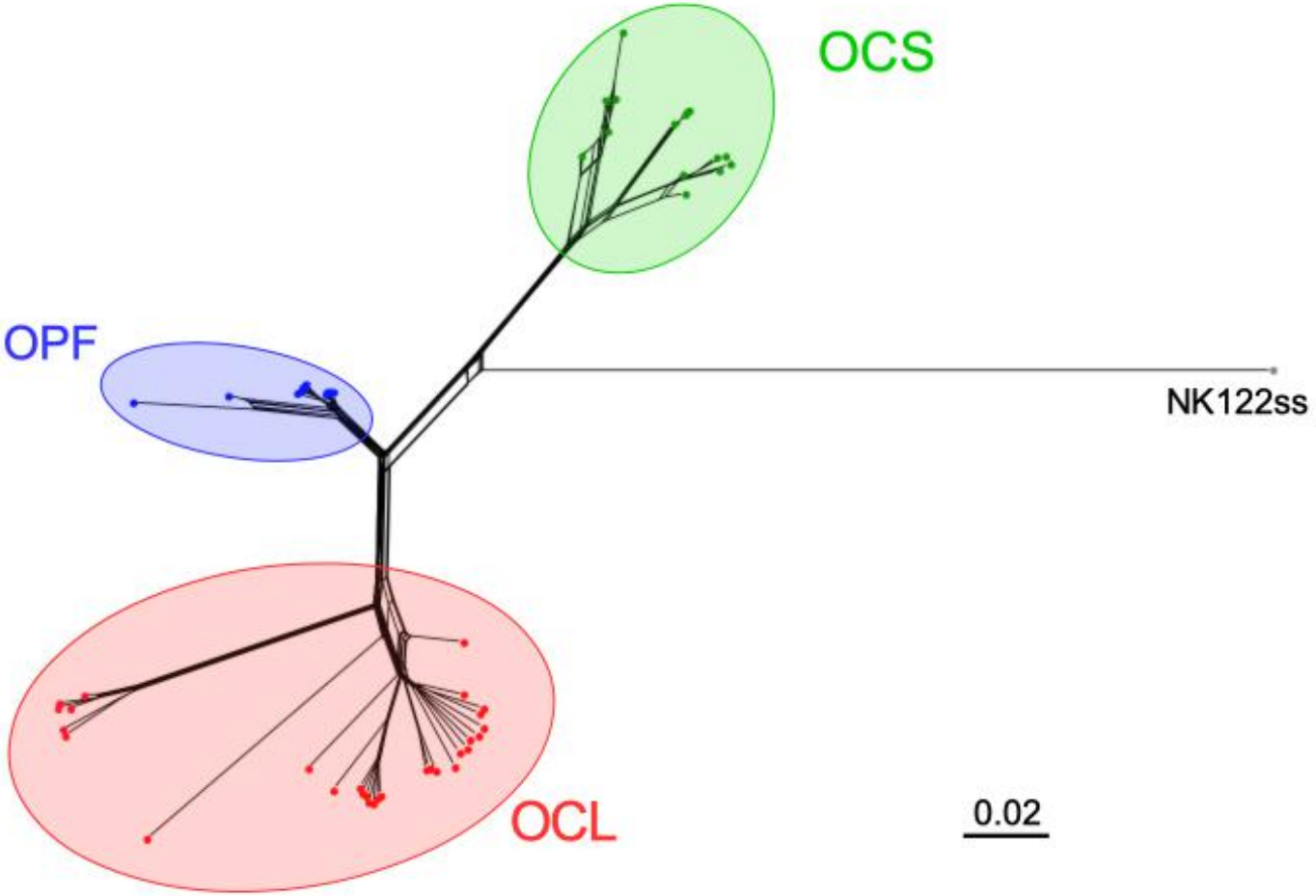

a)
1
2
3
4
5
6
7
8
9
N
LAOS
CAMBODIA
Gulf of Thailand
KEDAH
KELANTAN
Legend
CL
CS
PF
0
80
160
320
Kilometers
b)
NK604
OCL2
OCL1
NK122ss
PC2
PC1
20
0
-20
-40
-40
-20
0
20
40
60
Pop
CL
CS
PF
c)
Colobopsis leonardi
Colobopsis saundersi
Polyrhachis furcata
NK122ss
NK604
1
0.8
0.6
0.4
0.2
0
1 1 1 1 2 2 2 2 3 3 3 3 3 3 3 3 4 4 5 6 6 6 6 6 6 6 6 8 9 3 3 3 3 3 3 3 4 5 5 5 5 5 5 5 5 5 5 5 5 7 7 7 7 7 7 8

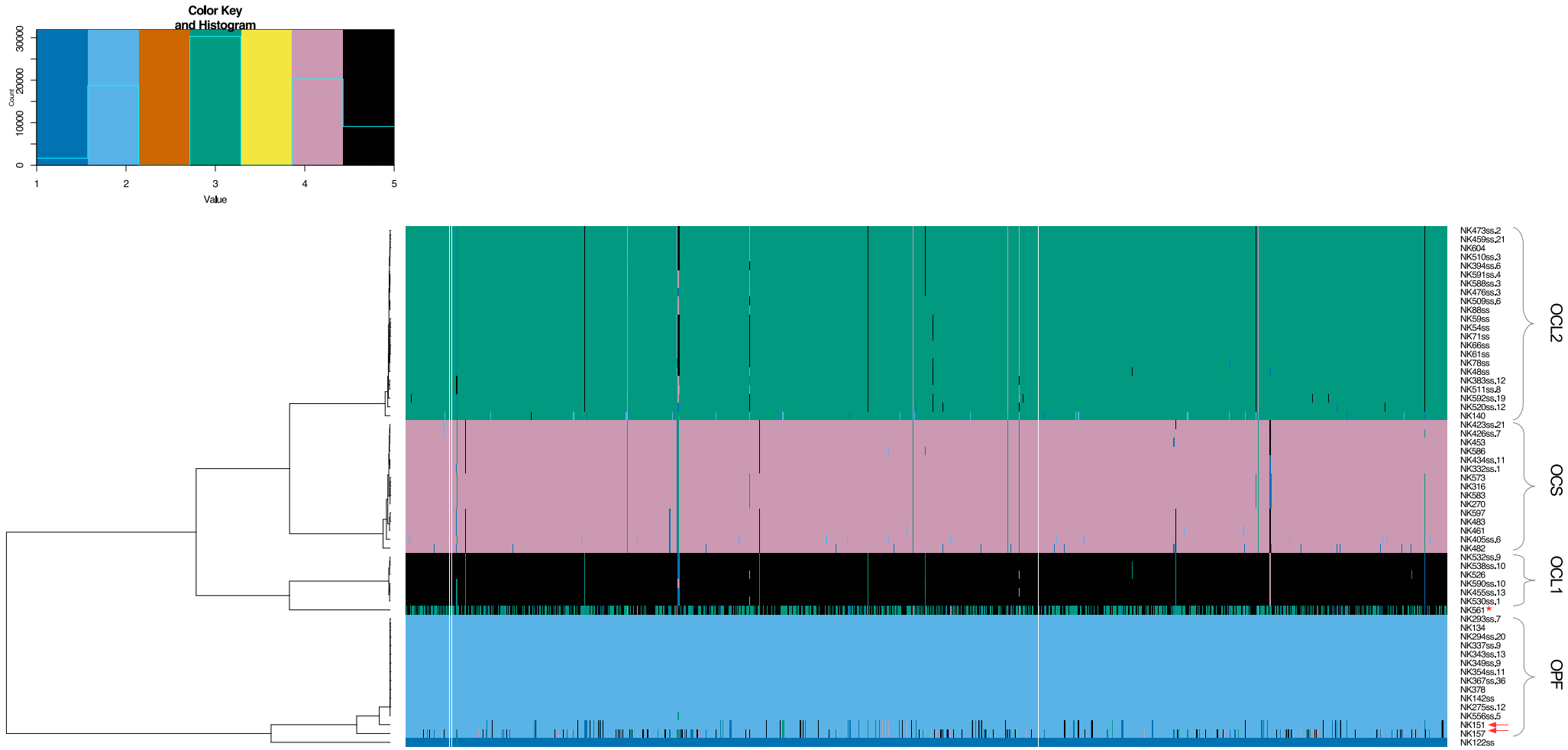

a)
0.001
NK61ss
NK66ss
NK88ss
NK59ss
NK48ss
NK54ss
NK71ss
NK383ss-12
NK78ss
NK510ss-3
NK511ss-8
NK509ss-6
NK591ss-4
NK473ss-2
NK394ss-6
NK588ss-3
NK604
NK459ss-21
NK476ss-3
NK592ss-19
NK520ss-12
NK538ss-10
NK530ss-1
NK532ss-9
NK526
NK590ss-10
NK455ss-13
NK561
NK140
O. camponoti-leonardi Clade
NK434ss-11
NK332ss-1
NK426ss-7
NK423ss-21
NK586
NK453
NK270
NK583
NK573
NK316
NK461
NK405ss-6
NK483
NK597
NK482
NK122ss
O. camponoti-saundersi Clade
NK354ss-11
NK349ss-9
NK343ss-13
NK367ss-36
NK337ss-9
NK556ss-5
NK151
NK157
NK294ss-20
NK142ss
NK293ss-7
NK378
NK275ss-12
NK134
O. polyrhachis-furcata Clade
O. kimflemingiae


b)
Scaffold 1
0.001
Scaffold 2.1
0.001
Scaffold 2.2
0.001
Scaffold 3
0.001
Scaffold 4
0.001
Scaffold 5
1e-04

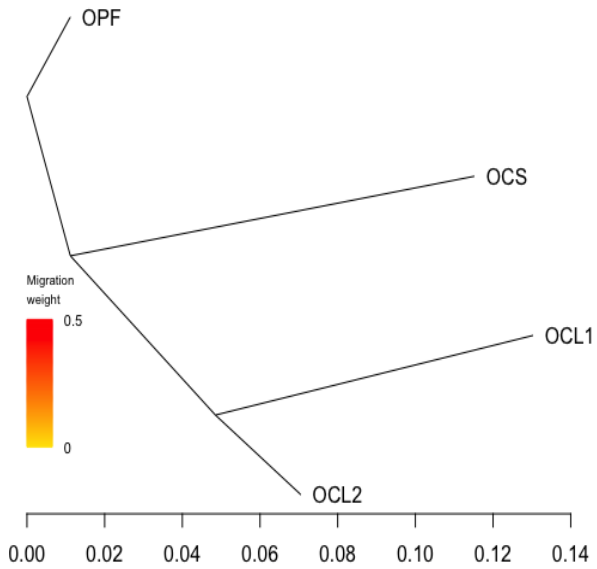

a)
0.993
(0.975-1.000)
OCL
0.894
0.837-0.940
NK122ss on Colobopsis leonardi
0.980
0.952-0.997
0.993
(0.975-1.000)
OCS
0.768
(0.691-0.837)
OPF
Ophiocordyceops kimflemingiae
2.0
b)
0.993
(0.975-1.000)
OCL1
NK140
0.477
(0.393-0.561)
0.901
(0.845-0.946)
OCL2
OCL2 sensu lato
0.475
(0.391-0.559)
NK561
NK122ss on Colobopsis leonardi
0.857
(0.794-0.912)
OCS1
0.976
(0.944-0.995)
OCS3
0.857
(0.692-0.834)
OCS2
NK157
0.433
(0.353-0.516)
OPF1
NK151
0.516
(0.434-0.598)
0.485
(0.397-0.574)
OPF2
Ophiocordyceps kimflemingiae
0.8


a)
NK140
NK561
1 0.8 0.6 0.4 0.2 0
1 1 1 1 2 2 2 2 3 3 3 3 3 3 3 3 4 4 5 6 6 6 6 6 6 6 6 8
b)
3 3 3 3 3 3 4 5 5 5 5 5 5 5 5
c)
NK151
NK157
5 5 5 5 5 5 7 7 7 7 7 7 7 8
N
LAOS
CAMBODIA
THAILAND
Gulf of Thailand
Legend
OCL1
OCL2
OCL3
OCS1
OCS2
OCS3
OPF1
OPF2
OPF3
0 80 160 320
Kilometers

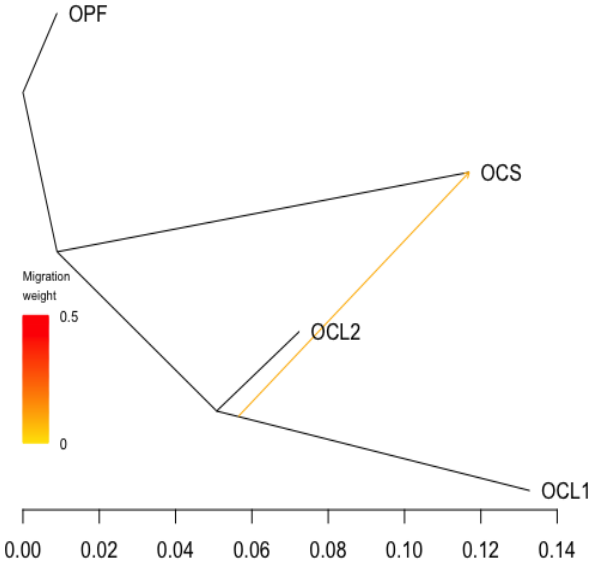

OCS
OPF
NK122ss
OCL
0.02

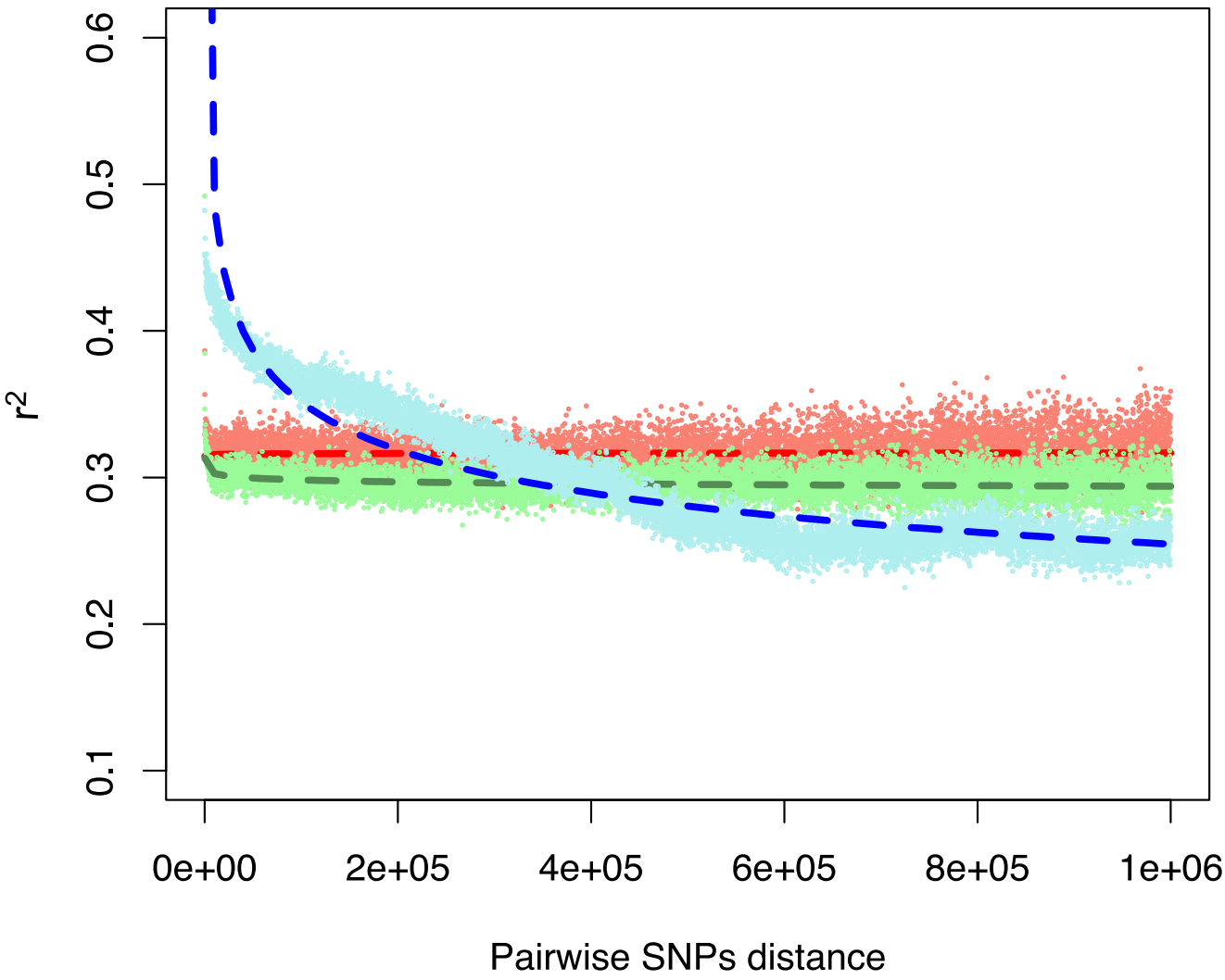

Color Key
and Histogram
Count
0
10000
20000
30000
1
2
3
4
5
Value
NK473ss.2
NK459ss.21
NK604
NK510ss.3
NK394ss.6
NK591ss.4
NK588ss.3
NK476ss.3
NK509ss.6
NK88ss
NK59ss
NK54ss
NK71ss
NK66ss
NK61ss
NK78ss
NK48ss
NK383ss.12
NK511ss.8
NK592ss.19
NK520ss.12
NK140
NK423ss.21
NK426ss.7
NK453
NK586
NK434ss.11
NK332ss.1
NK573
NK316
NK583
NK270
NK597
NK483
NK461
NK405ss.6
NK482
NK532ss.9
NK538ss.10
NK526
NK590ss.10
NK455ss.13
NK530ss.1
NK561
NK293ss.7
NK134
NK294ss.20
NK337ss.9
NK343ss.13
NK349ss.9
NK354ss.11
NK367ss.36
NK378
NK142ss
NK275ss.12
NK556ss.5
NK151
NK157
NK122ss
OCL2
OCS
OCL1
OPF

a. m=0, ln(likelihood) = 73.50



b. m=1, ln(likelihood) = 74.95, p-value = 0.2




0.6
0.5
0.4
0.3
0.2
0.1
r^2
0e+00
2e+05
4e+05
6e+05
8e+05
1e+06
Pairwise SNPs distance

Appendix A: Supplementary Figures

Cryptic species within *Ophiocordyceps unilateralis* as revealed by population genomicsThe one ant-one fungus paradigm in *Ophiocordyceps unilateralis sensu lato* challenged by population genomics

Noppol KOBMOO[†,‡], Suchada MONGKOLSAMRIT[†], Nuntanat ARNAMNART[†], Janet Jennifer LUANGSA-ARD[†], Tatiana GIRAUD[‡]

[†] National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), 113 Thailand Science Park, Phahonyothin Rd., Khlong Nueng, Khlong Luang, 12120 THAILAND.

[‡] Ecologie Systématique Evolution, Univ. Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay 91400 Orsay, France.

Corresponding author: Noppol Kobmoo,

E-mail: noppol.kob@biotec.or.th

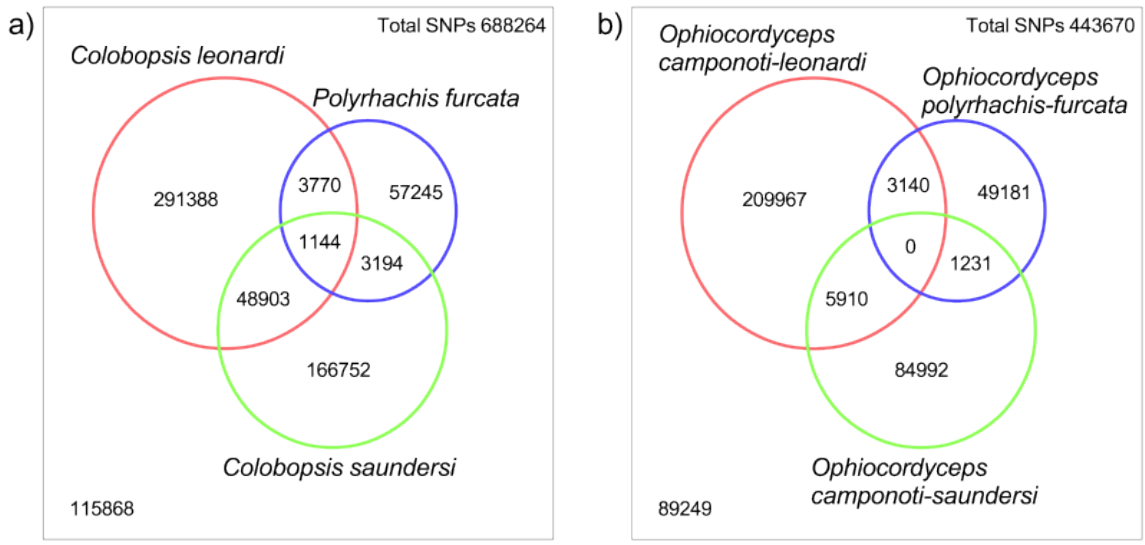
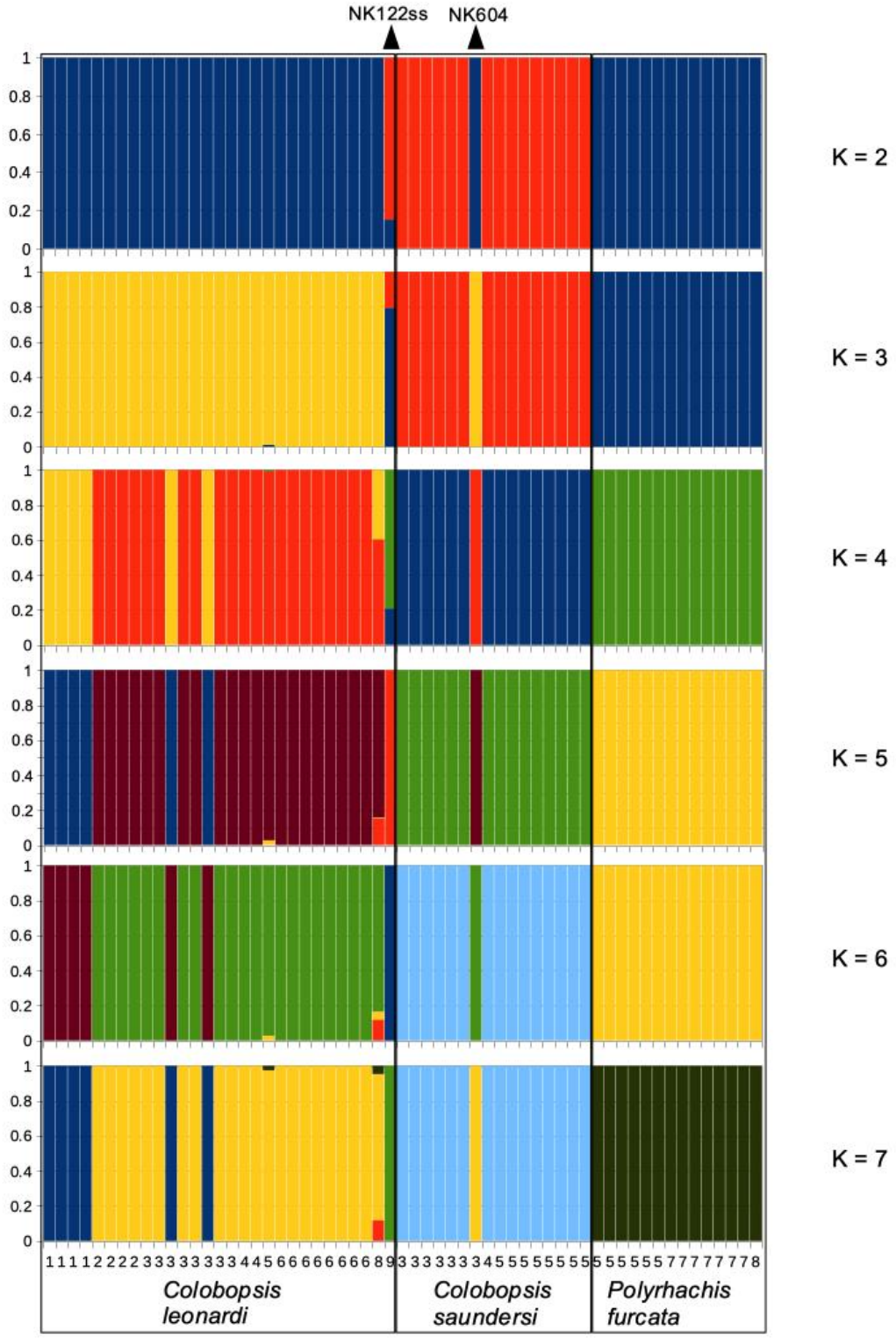


Fig.S1 The number of single nucleotide polymorphisms (SNPs) shared between a) samples from different ant hosts, and b) samples of different fungal species.

The figure S2 is given in a separate pdf file (FigS2.pdf) due to its big size.
Fig.S2 Maximum likelihood-based trees inferred using separate scaffolds. The colors at the tips represent the identities of ant hosts (red = *Colobopsis leonardi*, green = *C. saundersi*, blue = *Polyrhachis furcata*). The black circles on nodes represent 100% bootstrap support.

Fig.S3 Bayesian clustering based on whole-genome single nucleotide polymorphisms (SNPs). a) Barplots of assignment probabilities to different clusters from K = 1 to K = 7. b) The marginal likelihood for different K values. 1 = Ban Huai Baba, 2 = Doi Inthanon National Park, 3 = Nam Nao National Park, 4 = Phu Khiao Wildlife Sanctuary, 5 = Khao Yai National Park, 6 = Ko Chang Island, 7 = Khlong Nakha Wildlife Sanctuary, 8 = Khao Luang National Park, 9 = Hala Bala Wildlife Sanctuary.

a)


NK122ss
NK604
K = 2
K = 3
K = 4
K = 5
K = 6
K = 7
Colobopsis leonardi
Colobopsis saundersi
Polyrhachis furcata

b)

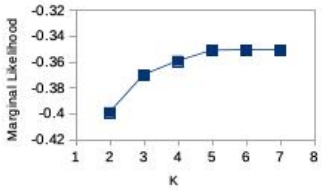

Marginal Likelihood
K

Fig.S4 Bayesian clustering based on single nucleotide polymorphisms (SNPs) inferred from the reference genome of *Ophiocordyceps camponoti-leonardi* for the samples belonging to this species. a) Barplots of assignment probabilities to different clusters from K = 1 to K = 8 (The labels on the x axis correspond to the sampling sites). b) The marginal likelihood for different K values. 1 = Ban Huai Baba, 2 = Doi Inthanon National Park, 3 = Nam Nao National Park, 4 = Phu Khiao Wildlife Sanctuary, 5 = Khao Yai National Park, 6 = Ko Chang Island, 8 = Khao Luang National Park.

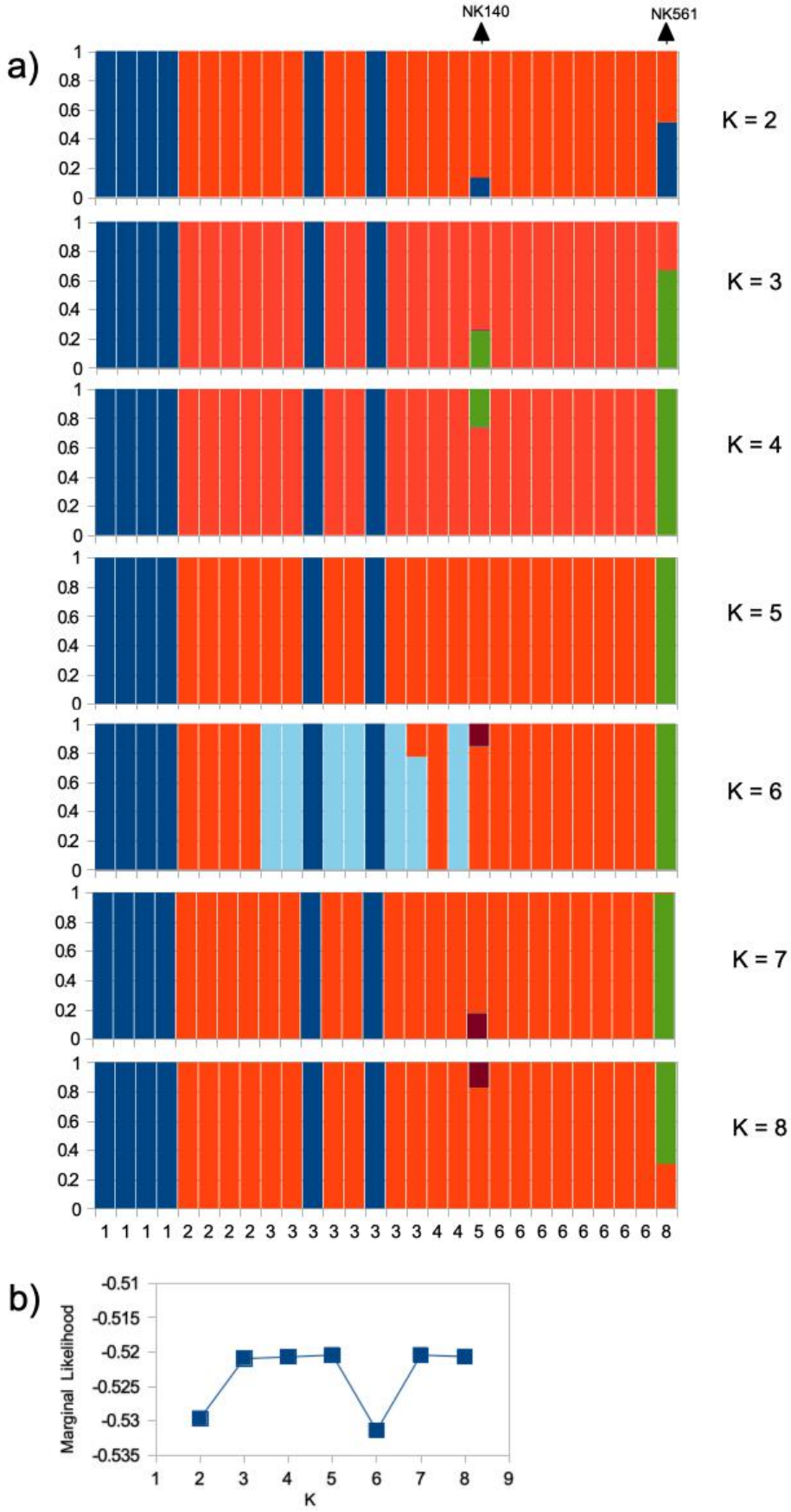
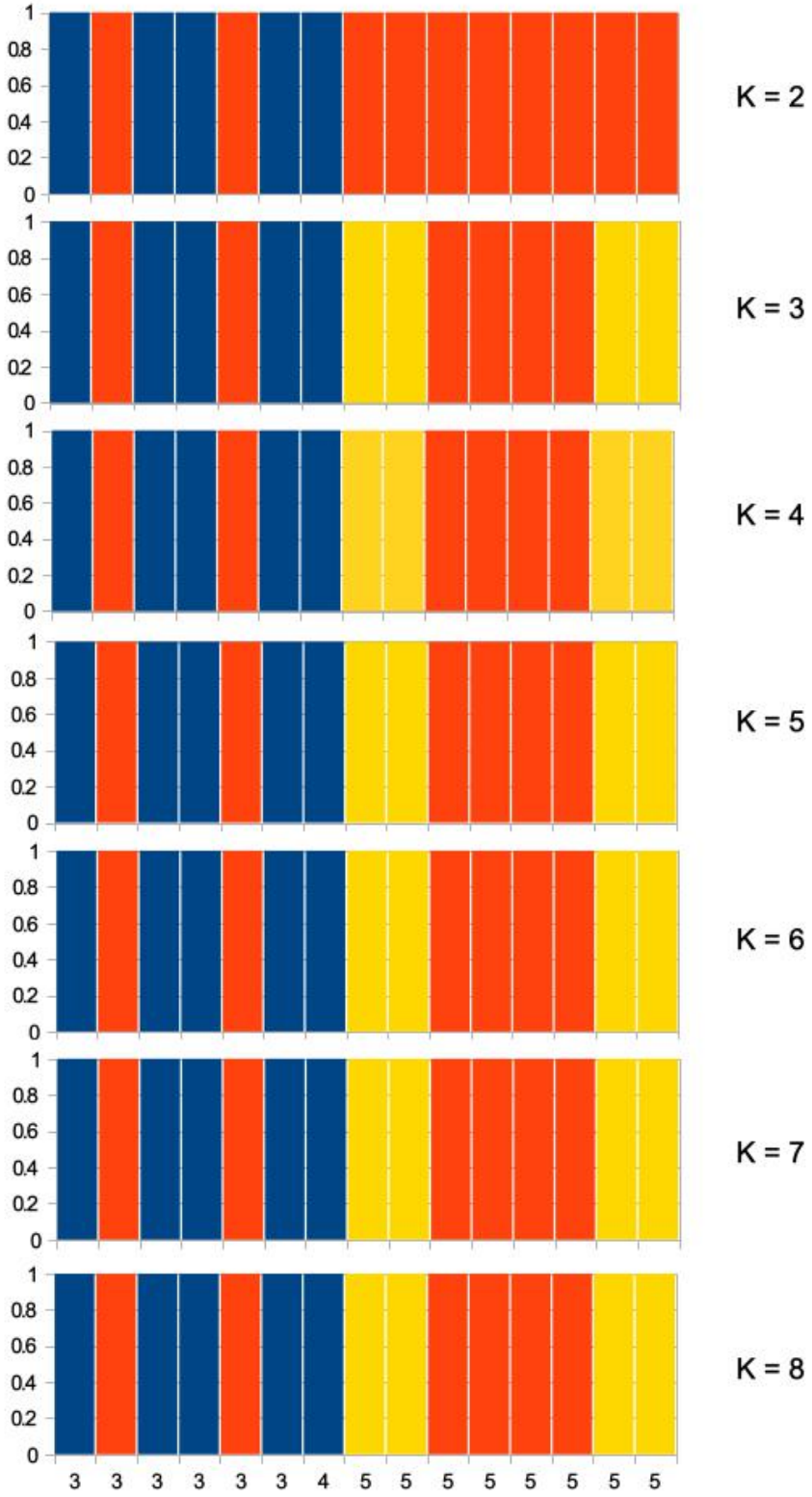
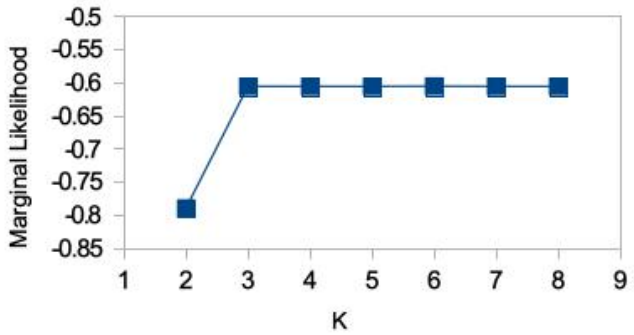

a)
NK140
NK561
K = 2
K = 3
K = 4
K = 5
K = 6
K = 7
K = 8
1 1 1 1 2 2 2 2 3 3 3 3 3 3 3 4 4 5 6 6 6 6 6 6 6 6 8
b)
Marginal Likelihood
K

Fig.S5 Bayesian clustering based on single nucleotide polymorphisms (SNPs) inferred from the reference genome of *Ophiocordyceps camponoti-saundersi* for the samples belonging to this species. a) Barplots of assignment probabilities to different clusters from K = 1 to K = 8 (The labels on the x axis correspond to the sampling sites). b) The marginal likelihood for different K values. 3 = Nam Nao National Park, 4 = Phu Khiao Wildlife Sanctuary, 5 = Khao Yai National Park.

a)


K = 2
K = 3
K = 4
K = 5
K = 6
K = 7
K = 8
0
0.2
0.4
0.6
0.8
1
3 3 3 3 3 3 4 5 5 5 5 5 5 5 5

b)


Marginal Likelihood
-0.5
-0.55
-0.6
-0.65
-0.7
-0.75
-0.8
-0.85
1 2 3 4 5 6 7 8 9
K

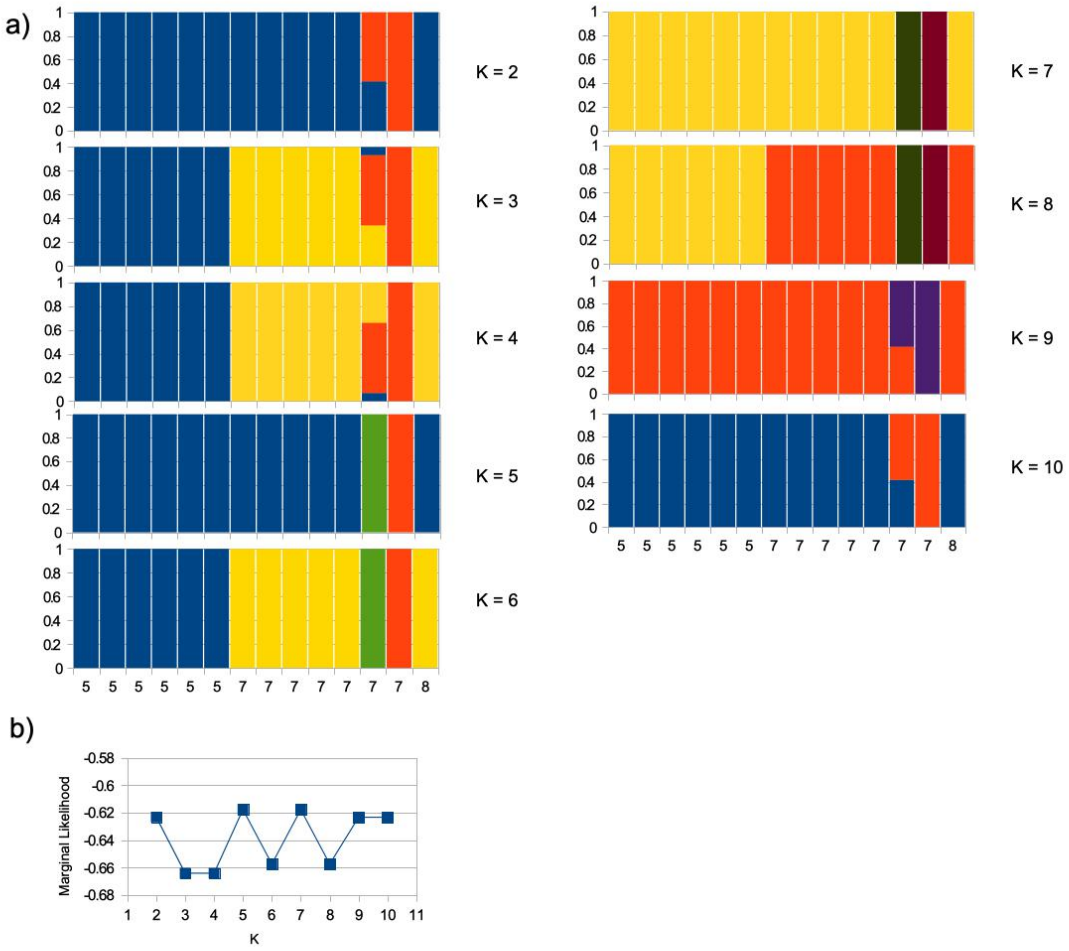


Fig.S6 Bayesian clustering based on single nucleotide polymorphisms (SNPs) inferred from the reference genome of *Ophiocordycep polyrhachis-furcata* for the samples belonging to this species. a) Barplots of assignment probabilities to different clusters from K = 1 to K = 8 (The labels on the x axis correspond to the sampling sites). b) The marginal likelihood for different K values. 5 = Khao Yai National Park, 7 = Khlong Nakha Wildlife Sanctuary, 8 = Khao Luang National Park.

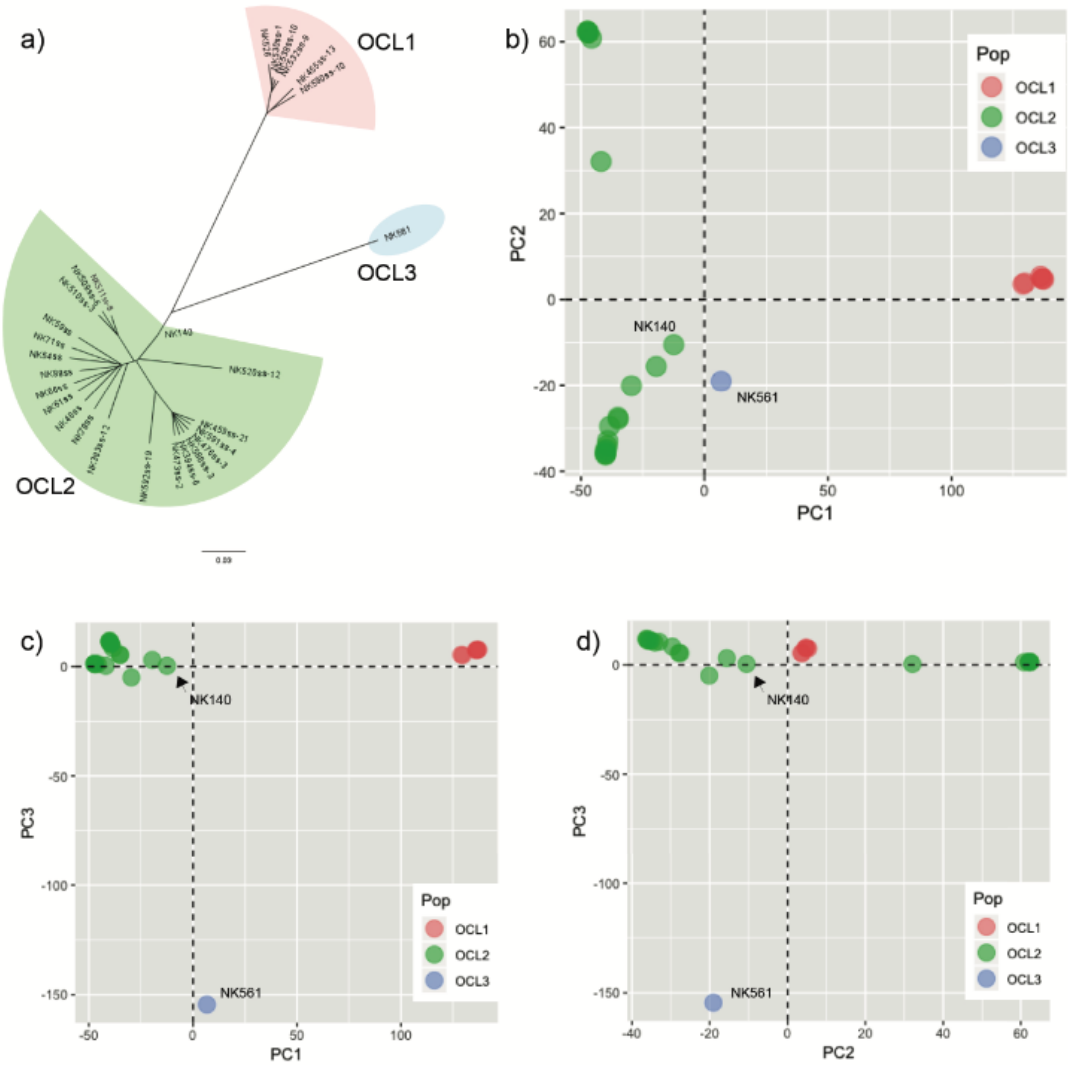


Fig.S7 Population structure among the samples belonging to *Ophiocordyceps camponoti-leonardi*. a) Neighbor-joining tree based on an alignment of single nucleotide polymorphisms (SNPs) and using the F84 distance. b) – d) Principal component analyses (PCA) based on SNPs.

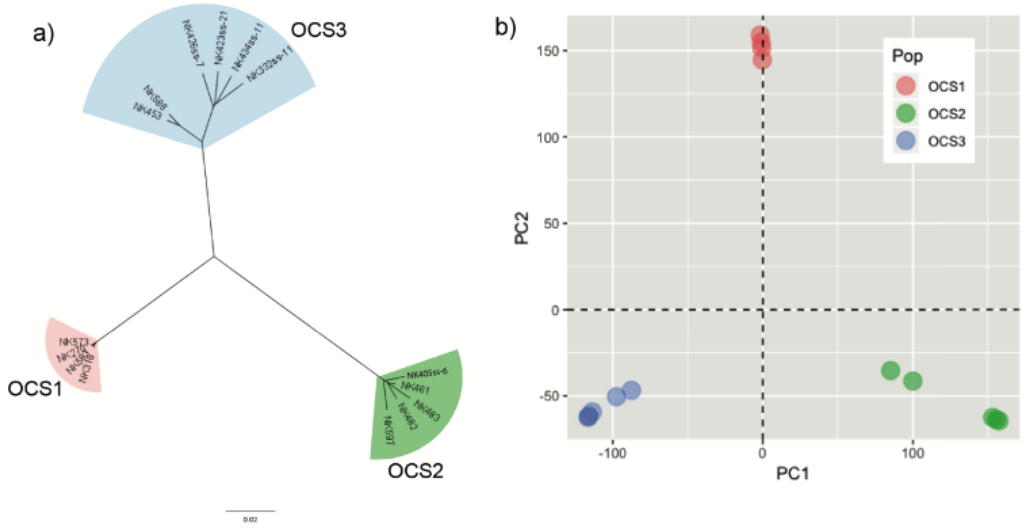


Fig.S8 Population structure among the samples belonging to *Ophiocordyceps camponoti-saundersi*. a) Neighbor-joining tree based on an alignment of single nucleotide polymorphisms (SNPs) using the F84 distance. b) A principal component analysis (PCA) based on SNPs.

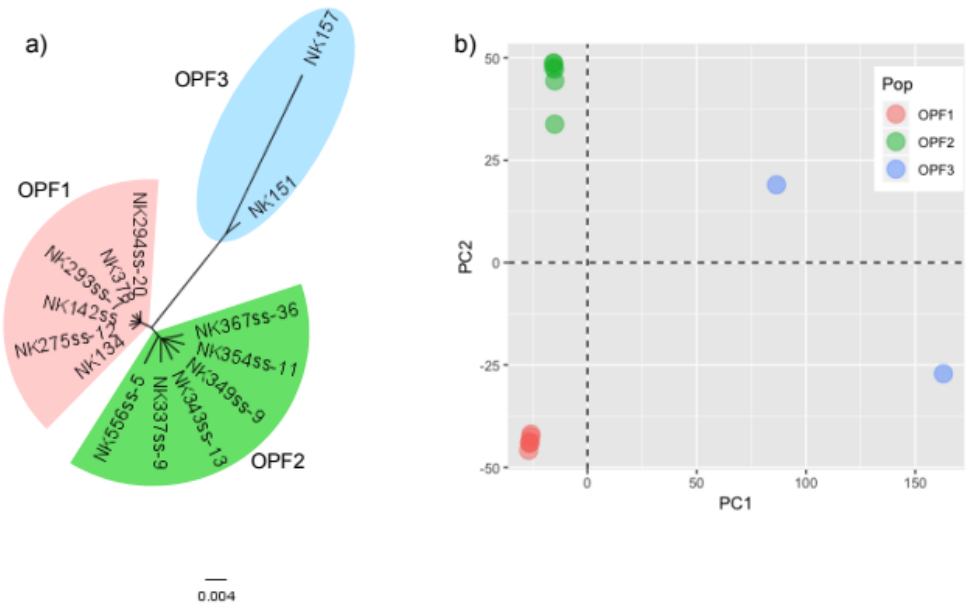


Fig.S9 Population structure among the samples belonging to *Ophiocordyceps polyrhachis-furcata*. a) Neighbor-joining tree from based on an alignment of single nucleotide polymorphisms (SNPs) and using the F84 distance. b) A principal component analysis (PCA) based on SNPs.

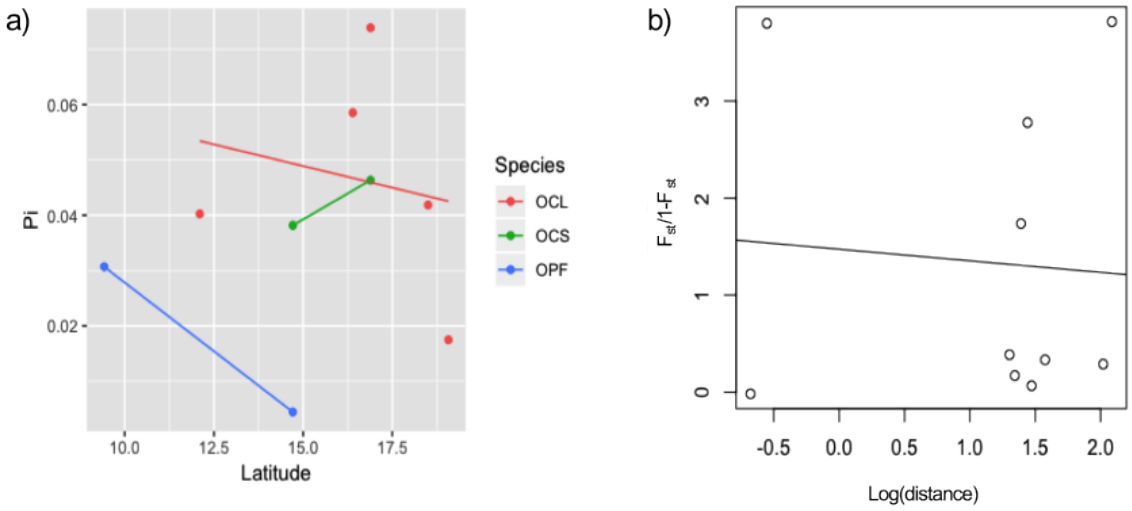


Fig.S10 a) Correlation between nucleotide diversity (Pi) per site and per species and the sampling site latitude (OCL = *Ophiocordyceps camponoti-leonardi*, OCS = *O. camponoti-saundersi*, OPF = *O. polyrhachis-furcata*). b) Correlation between genetic dissimilarity between sampling sites of *O. camponoti-leonardi* ($F_{st}/1-F_{st}$) and the log distance between sites.

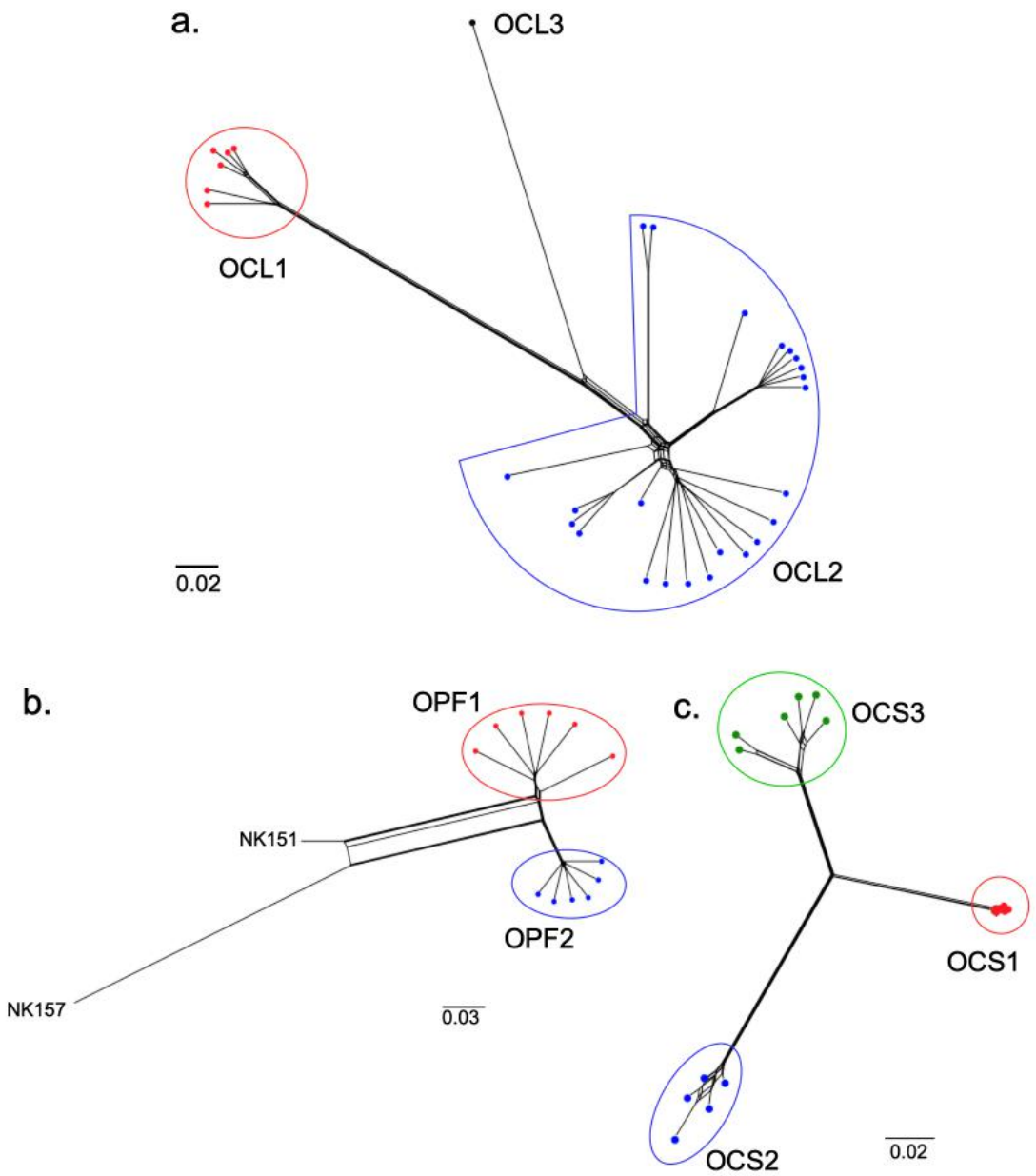


Fig.S11 Neighbor-net phylogenetic networks for a) *Ophiocordyceps camponoti-leonardi*, b) *O. polyrhachis-furcata* and c) *O. camponoti-saundersi*.

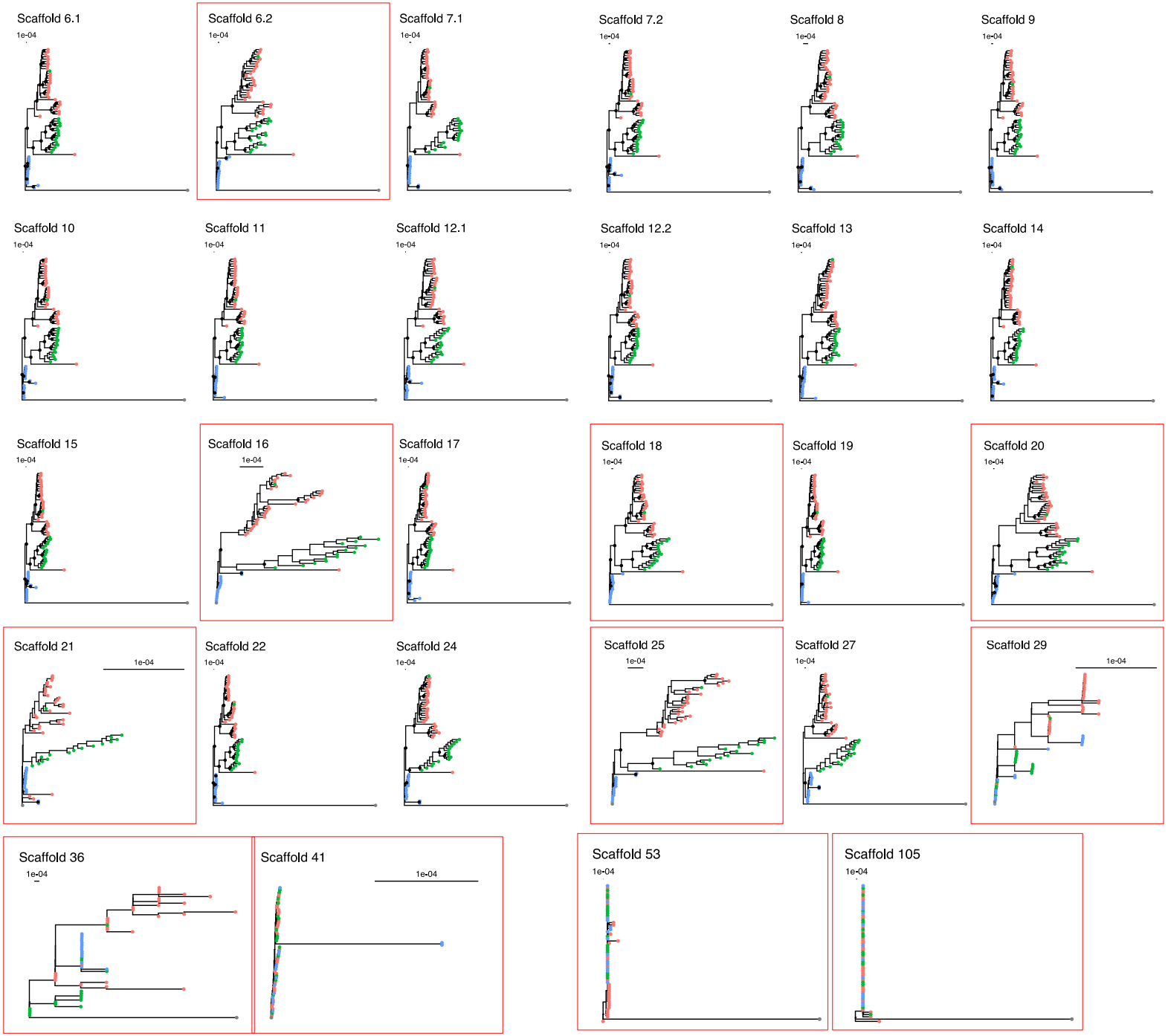

Scaffold 6.1
1e-04
Scaffold 6.2
1e-04
Scaffold 7.1
1e-04
Scaffold 7.2
1e-04
Scaffold 8
1e-04
Scaffold 9
1e-04
Scaffold 10
1e-04
Scaffold 11
1e-04
Scaffold 12.1
1e-04
Scaffold 12.2
1e-04
Scaffold 13
1e-04
Scaffold 14
1e-04
Scaffold 15
1e-04
Scaffold 16
1e-04
Scaffold 17
1e-04
Scaffold 18
1e-04
Scaffold 19
1e-04
Scaffold 20
1e-04
Scaffold 21
1e-04
Scaffold 22
1e-04
Scaffold 24
1e-04
Scaffold 25
1e-04
Scaffold 27
1e-04
Scaffold 29
1e-04
Scaffold 36
1e-04
Scaffold 41
1e-04
Scaffold 53
1e-04
Scaffold 105
1e-04

Table S1: Fungal isolates used in this study (BCC = Biotec Culture Collection Accession Number, BBH = Biotec Bangkok Herbarium accession number).

Isolate	Isolation method	Ant host	Site	Province	BCC	BBH	BioSample
NK122ss	single spore	*Colobopsis leonardi*	Halabala Wildlife Sanctuary	Narathiwat	69099	37954	SAMN11321611
NK134	multiple spores	*Polyrhachis furcata*	Khao Yai National Park	Nakhon Ratchasima	70388	38294	SAMN11321612
NK140	multiple spores	*Colobopsis leonardi*	Khao Yai National Park	Nakhon Ratchasima	70397	38300	SAMN11321613
NK270	multiple spores	*Colobopsis saundersi*	Khao Yai National Park	Nakhon Ratchasima	75911	39875	SAMN11321614
NK293ss-7	single spore	*Polyrhachis furcata*	Khao Yai National Park	Nakhon Ratchasima	75920	39883	SAMN11321615
NK294ss-20	single spore	*Polyrhachis furcata*	Khao Yai National Park	Nakhon Ratchasima	-	41439	SAMN11321616
NK332ss-1	single spore	*Colobopsis saundersi*	Khao Yai National Park	Nakhon Ratchasima	78437	39890	SAMN11321617
NK337ss-9	single spore	*Polyrhachis furcata*	Khlong Nakha Wildlife Sanctuary	Ranong	79284	41466	SAMN11321618
NK343ss-13	single spore	*Polyrhachis furcata*	Khlong Nakha Wildlife Sanctuary	Ranong	78443	41472	SAMN11321619
NK349ss-9	single spore	*Polyrhachis furcata*	Khlong Nakha Wildlife Sanctuary	Ranong	78445	41478	SAMN11321620
NK354ss-11	single spore	*Polyrhachis furcata*	Khlong Nakha Wildlife Sanctuary	Ranong	78446	41483	SAMN11321621
NK367ss-36	single spore	*Polyrhachis furcata*	Khlong Nakha Wildlife Sanctuary	Ranong	78449	41496	SAMN11321622
NK378	multiple spores	*Polyrhachis furcata*	Khao Yai National Park	Nakhon Ratchasima	79295	41503	SAMN11321623
NK383ss-12	single spore	*Colobopsis leonardi*	Phu Khiao Wildlife Sanctuary	Chaiyaphum	79300	41508	SAMN11321624
NK394ss-6	single spore	*Colobopsis leonardi*	Phu Khiao Wildlife Sanctuary	Chaiyaphum	79306	41519	SAMN11321625
NK48ss	single spore	*Colobopsis leonardi*	Ko Chang Island	Trat	68537	37537	SAMN11321626
NK54ss	single spore	*Colobopsis leonardi*	Ko Chang Island	Trat	68541	37541	SAMN11321627
NK59ss	single spore	*Colobopsis leonardi*	Ko Chang Island	Trat	68545	37543	SAMN11321628
NK61ss	single spore	*Colobopsis leonardi*	Ko Chang Island	Trat	68548	37545	SAMN11321629
NK66ss	single spore	*Colobopsis leonardi*	Ko Chang Island	Trat	68551	37548	SAMN11321630
NK71ss	single spore	*Colobopsis leonardi*	Ko Chang Island	Trat	69073	38283	SAMN11321631
NK78ss	single spore	*Colobopsis leonardi*	Ko Chang Island	Trat	69077	37944	SAMN11321632
NK88ss	single spore	*Colobopsis leonardi*	Ko Chang Island	Trat	69080	37555	SAMN11321633
NK459ss-21	single spore	*Colobopsis leonardi*	Nam Nao National Park	Petchabun	80028	41583	SAMN11321634
NK476ss-3	single spore	*Colobopsis leonardi*	Nam Nao National Park	Petchabun	80038	41600	SAMN11321635
NK509ss-6	single spore	*Colobopsis leonardi*	Doi Inthanon National Park	Chiang Mai	80043	41633	SAMN11321636

NK510ss-3	single spore	*Colobopsis leonardi*	Doi Inthanon National Park	Chiang Mai	80366	41634	SAMN11321637
NK511ss-8	single spore	*Colobopsis leonardi*	Doi Inthanon National Park	Chiang Mai	80369	41635	SAMN07662903
NK526	multiple spores	*Colobopsis leonardi*	Ban Huay Baba	Chiang Mai	80379	42812	SAMN11321638
NK532ss-9	single spore	*Colobopsis leonardi*	Ban Huay Baba	Chiang Mai	80044	42818	SAMN11321639
NK455ss-8	single spore	*Colobopsis leonardi*	Nam Nao National Park	Petchabun	80360	41580	SAMN11321640
NK473ss-2	single spore	*Colobopsis leonardi*	Nam Nao National Park	Petchabun	80614	41597	SAMN11321641
NK520ss-12	single spore	*Colobopsis leonardi*	Doi Inthanon National Park	Chiang Mai	80377	41644	SAMN11321642
NK530ss-1	single spore	*Colobopsis leonardi*	Ban Huay Baba	Chiang Mai	80380	42816	SAMN11321643
NK538ss-10	single spore	*Colobopsis leonardi*	Ban Huay Baba	Chiang Mai	80045	42824	SAMN11321644
NK561	multiple spores	*Colobopsis leonardi*	Khao Luang National Park	Nakhonsithammarat	80621	41657	SAMN11321645
NK588ss-3	single spore	*Colobopsis leonardi*	Khao Luang National Park	Petchabun	84568	42855	SAMN11321646
NK590ss-10	single spore	*Colobopsis leonardi*	Khao Luang National Park	Petchabun	84571	42857	SAMN11321647
NK591ss-4	single spore	*Colobopsis leonardi*	Khao Luang National Park	Petchabun	84581	42858	SAMN11321648
NK592ss-19	single spore	*Colobopsis leonardi*	Khao Luang National Park	Petchabun	-	42859	SAMN11321649
NK316	multiple spores	*Colobopsis saundersi*	Khao Yai National Park	Nakhon Ratchasima	-	39888	SAMN11321650
NK405ss-6	single spore	*Colobopsis saundersi*	Phu Khiao Wildlife Sanctuary	Chaiyaphum	79314	41530	SAMN07662932
NK423ss-21	single spore	*Colobopsis saundersi*	Khao Yai National Park	Nakhon Ratchasima	79321	41548	SAMN11321651
NK426ss-7	single spore	*Colobopsis saundersi*	Khao Yai National Park	Nakhon Ratchasima	79326	41551	SAMN11321652
NK434ss-11	single spore	*Colobopsis saundersi*	Khao Yai National Park	Nakhon Ratchasima	79331	41559	SAMN11321653
NK461	multiple spores	*Colobopsis saundersi*	Nam Nao National Park	Petchabun	-	41585	SAMN11321654
NK453	multiple spores	*Colobopsis saundersi*	Nam Nao National Park	Petchabun	-	41578	SAMN11321655
NK482	multiple spores	*Colobopsis saundersi*	Nam Nao National Park	Petchabun	80039	41606	SAMN11321656
NK483	multiple spores	*Colobopsis saundersi*	Nam Nao National Park	Petchabun	80040	41607	SAMN11321657
NK573	multiple spores	*Colobopsis saundersi*	Khao Yai National Park	Nakhon Ratchasima	82852	42840	SAMN11321658
NK583	multiple spores	*Colobopsis saundersi*	Khao Yai National Park	Nakhon Ratchasima	84565	42850	SAMN11321659
NK586	multiple spores	*Colobopsis saundersi*	Nam Nao National Park	Petchabun	-	42853	SAMN11321660
NK597	multiple spores	*Colobopsis saundersi*	Nam Nao National Park	Petchabun	-	42864	SAMN11321661
NK604	multiple spores	*Colobopsis saundersi*	Nam Nao National Park	Petchabun	84591	42871	SAMN11321662
NK142ss	single spore	*Polyrhachis furcata*	Khao Yai National Park	Nakhon Ratchasima	-	38301	SAMN11321663
NK275ss-12	single spore	*Polyrhachis furcata*	Khao Yai National Park	Nakhon Ratchasima	75913	39877	SAMN11321664

NK151	multiple spores	*Polyrhachis furcata*	Khlong Naka Wildlife Sanctuary	Ranong	71355	38599	SAMN11321665
NK157	multiple spores	*Polyrhachis furcata*	Khlong Naka Wildlife Sanctuary	Ranong	70665	38304	SAMN11321666
NK556ss-5	single spore	*Polyrhachis furcata*	Khao Luang National Park	Nakhonsithammarat	-	41652	SAMN11321667

Table S2: Genetic diversity per scaffold in *Ophiocordyceps unilateralis sensu lato*

Scaffold	Length	SNPs Number	Pi
1	5280523	136209	0.034
2.1	818828	24837	0.035
2.2	3902607	55058	0.032
3	2977708	19974	0.037
4	2824012	33938	0.038
5	2463813	14234	0.035
6.1	4344696	102669	0.032
6.2	247174	4663	0.046
7.1	407820	3955	0.000
7.2	4080701	49613	0.032
8	3328657	15513	0.033
9	1794465	17459	0.040
10	1211063	33165	0.034
11	1131199	32372	0.032
12.1	386211	10282	0.000
12.2	604937	16862	0.032
13	939609	18870	0.035
14	963256	24241	0.034
15	795048	16431	0.033
16	630955	1058	0.043
17	583368	12928	0.034
18	594624	8781	0.031
19	433702	9584	0.033
20	376207	9947	0.032
21	373426	113	0.000
22	342176	9096	0.043
24	164406	4377	0.036
25	156640	500	0.066
27	99302	1468	0.031
29	50804	17	0.015
36	7493	19	0.000
41	11136	1	0.000
53	1236	23	0.000
105	8111	7	0.000

Table S3: Genetic diversity by site and by species: OCL = *Ophiocordyceps camponoti-leonardi*, OCS = *O. camponoti-saundersi*, OPF = *O. polyrhachis-furcata*. Pi = nucleotide diversity.

Sampling Site	**Province**	**Latitude**	**Species**		
			OCL	**OCS**	**OPF**
Ban Huai Baba	Chiang Mai	19.070	0.0056	-	-
Doi Inthanon National Park	Chiang Mai	18.498	0.0138	-	-
Nam Nao National Park	Phetchabul	16.887	0.0225	0.0121	-
Phu Khiao Wildlife Sanctuary	Chaiyaphum	16.390	0.0195	- (one sample)	-
Khao Yai National Park	Nakhon Ratchasima	14.712	- (one sample)	0.0108	0.0014
Ko Chang Island	Trad	12.103	0.0129	-	-
Khlong Nakha Wildlife Sanctuary	Chumphon	9.423	-	-	0.0097
Khao Luang National Park	Nakhon Sithammarat	8.373	- (one sample)	-	- (one sample)
Hala Bala Wildlife Sanctuary	Narathiwat	5.928	-	-	-

Table S4: Results of PHI tests of signature of recombination. (OCL = *Ophiocordyceps camponoti-leonardi*, OCS = *O. camponoti-saundersi*, OPF = *O. polyrhachis-furcata*) NAs indicate lack of sufficient informative sites.

Scaffold	**All**	**OCL**	**OCL1**	**OCL2**	**OCS**	**OCS1**	**OCS2**	**OCS3**	**OPF**	**OPF1**	**OPF2**
scaffold1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
scaffold2.1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
scaffold2.2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
scaffold3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
scaffold4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
scaffold5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.658	1.000
scaffold6.1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
scaffold6.2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000
scaffold7.1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000
scaffold7.2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
scaffold8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
scaffold9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
scaffold10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
scaffold11	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
scaffold12.1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
scaffold12.2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.658	0.000	0.000
scaffold13	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
scaffold14	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
scaffold15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
scaffold16	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	NA	NA
scaffold17	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
scaffold18	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
scaffold19	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000
scaffold20	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.987
scaffold21	0.987	0.000	1.000	0.000	1.000	1.000	1.000	1.000	1.000	NA	NA
scaffold22	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.329
scaffold24	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000
scaffold25	0.000	0.000	1.000	0.000	0.000	0.000	1.000	0.000	1.000	NA	1.000
scaffold27	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000

scaffold29	1.000	1.000	1.000	1.000	1.000	NA	NA	NA	1.000	NA	NA
scaffold36	1.000	1.000	1.000	1.000	1.000	NA	NA	NA	NA	NA	NA
scaffold41	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
scaffold53	1.000	1.000	1.000	NA	NA	NA	NA	NA	1.000	NA	NA
scaffold109	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table S5: The observed, expected number of homozygous sites, total number of SNPs and inbreeding coefficient for the diploid samples.

Samples	Species	Observed	Expected	Total	F coefficient
NK134	*O. polyrhachis-furcata*	687858	596600	687858	1
NK140	*O. camponoti-leonardi*	640930	560000	642118	0.9855
NK270	*O. camponoti-saundersi*	645429	561400	645447	0.9998
NK378	*O. polyrhachis-furcata*	687619	596400	687620	1
NK526	*O. camponoti-leonardi*	663027	575500	663042	0.9998
NK561	*O. camponoti-leonardi*	663418	576800	663422	1
NK316	*O. camponoti-saundersi*	646450	563200	646468	0.9998
NK461	*O. camponoti-saundersi*	640107	558300	640151	0.9995
NK453	*O. camponoti-saundersi*	561257	487900	561296	0.9995
NK482	*O. camponoti-saundersi*	453036	391900	453065	0.9995
NK483	*O. camponoti-saundersi*	643876	560500	643892	0.9998
NK573	*O. camponoti-saundersi*	647183	562900	647204	0.9998
NK583	*O. camponoti-saundersi*	646163	562200	646175	0.9999
NK586	*O. camponoti-saundersi*	646587	562700	646629	0.9995
NK597	*O. camponoti-saundersi*	644848	561400	644867	0.9998
NK604	*O. camponoti-leonardi*	670466	582300	670481	0.9998
NK151	*O. polyrhachis-furcata*	686833	595700	686833	1
NK157	*O. polyrhachis-furcata*	685814	594800	685816	1