

High-throughput detection of microeukaryotic parasites in insects using Nanopore sequencing

Edouard Bessette^{1,2,*†}, Sam Edwards^{1,2‡} and Bryony A. P. Williams^{1,*}

Abstract

The mass rearing of insects for food and feed is increasing significantly, but it presents challenges in detecting and characterizing pathogens which may cause severe economic losses. Much attention is placed on detecting viruses, bacteria and fungi; however, other significant entomopathogens remain overlooked, such as protist and microsporidian parasites. This study employed Nanopore sequencing to explore and phylogenetically analyse microsporidian and protist parasites in reared insects. Using group-specific primers and amplicon pooling strategies, we successfully amplified and sequenced a key protist group (Gregarinasina), microsporidia and fungal pathogens from insect host DNA, providing data for their future detection. Our approach generated novel molecular data, enabling the inference of small subunit and long subunit rRNA phylogenies which enhance the understanding of the evolutionary origin of these parasites. Furthermore, we highlight potential shortcomings of currently available primers and provide new Gregarine-specific primers to aid in further research of these overlooked entomopathogens. This research demonstrates the effectiveness of third-generation sequencing technologies in improving diagnostic tools and advances the study of pathogen biodiversity within economically important insect hosts.

Impact Statement

This study introduces an efficient, high-throughput method for detecting diverse microeukaryotic parasites in insect hosts using pooled amplicon Nanopore sequencing. Our approach enables sensitive pathogen identification from minimal DNA and reveals previously unknown parasite diversity. The validated primers and bioinformatics pipeline can support routine, cost-effective pathogen screening in insect farming, helping to safeguard production and advance food security.

DATA SUMMARY

This study generated Nanopore sequencing data from amplicon libraries targeting protist and microsporidian parasites in insect hosts. All raw sequencing data, consensus 18S/28S sequences, sample table (Samples_Seq_table.xlsx) and bioinformatics scripts are publicly available either through GitHub ([Amplicon-high-throughput-ONT-sequencing](#)) or Zenodo repository (DOI: 10.5281/zenodo.15129942). All supporting phylogenetic alignments and raw data files used for tree construction can be accessed through the Phylogeny-Alignments repository link ([Phylogeny-Alignments](#)).

Received 19 June 2025; Accepted 29 January 2026; Published 16 March 2026

Author affiliations: ¹Department of Biosciences, Faculty of Health and Life Sciences, University of Exeter, Exeter, EX4 4QD, UK; ²Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg C, Denmark.

***Correspondence:** Edouard Bessette, edouard.bessette@gmail.com

Keywords: Nanopore MinION; microeukaryotes; parasites; reared insects; phylogenetics; pathogen surveillance.

Abbreviations: GTR, general time reversible; LSU, large subunit; SSU, small subunit; TGS, third-generation sequencing.

†Present address: Natural History Museum Denmark, University of Copenhagen, Sølvgade 83, 1307 Copenhagen, K, Denmark

‡Present address: Laboratory of Virology, Plant Sciences Group, Wageningen University, Droevendaalsesteeg 1, 6708 PB, Wageningen, Netherlands.

All supporting data, code and protocols have been provided within the article or through supplementary data files. Two supplementary figures, two supplementary tables and supplementary files are available with the online version of this article.

001649 © 2026 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution License. This article was made open access via a Publish and Read agreement between the Microbiology Society and the corresponding author's institution.

DATA AVAILABILITY

Our research contributes to reproducibility and transparency by making all raw data and analysis scripts publicly available through a GitHub (Amplicon-high-throughput-ONT sequencing: ONT sequencing of protists and microsporidia 18S/28S amplicons) and a Zenodo repository (Zenodo.15129942).

INTRODUCTION

Insect mass rearing for food and feed requires robust diagnostic tools to detect pathogens beyond well-studied bacteria, viruses and fungi. Protist and microsporidian parasites can thrive in high-density conditions, causing severe economic losses and colony collapse [1, 2]. For example, the amoeba *Malamoeba locustae* is known to infect reared orthopterans, causing premature death [3, 4]. Similarly, the gregarine *Gregarina cuneata* affects the mealworm *Tenebrio molitor*, leading to higher mortality, weight loss and developmental delays [5, 6]. Microsporidian parasites are known to induce a general reduction in host fitness across multiple insect hosts, resulting in higher mortality, reduced longevity, weight loss and developmental delays [2, 7].

Despite their impact, many entomopathogenic protists lack molecular data for diagnostics [8]. Although metabarcoding has emerged as a powerful approach for assessing microbial diversity, these pathogens are often overlooked. Comprehensive molecular data collection could improve our understanding of their phylogenetic relationships, particularly in neglected clades such as Gregarinasina or Amoebozoa [8]. However, conventional metabarcoding approaches typically rely on short-read sequences, which restrict the phylogenetic resolution and can prevent accurate species delineation. Third-generation sequencing (TGS) allows the sequencing of longer fragments, improving taxonomic identification and phylogenetic accuracy [9, 10]. A current challenge for TGS compared to Illumina (second-generation sequencing) is the variable sequencing error rates, ranging from 1 to 15%, depending on the used platforms [11–14]. This challenge has been addressed by the creation of novel tools and approaches that can perform *de novo* clustering of long noisy ONT reads into operational taxonomic units, such as NGSpeciesID, reaching more than 99% accuracy [12, 15–17]. Until now, DNA barcoding analyses have predominantly relied on QIIME2, which is widely used for 16S amplicons [18, 19]. However, QIIME2's support for long-read amplicon analysis remains limited [12].

In this study, we used short- and long-read metabarcoding and cost-effective MinION sequencing to generate molecular data for poorly characterized protists and improve diagnostics. Targeting entomopathogenic protists and Microsporidia, we employed group-specific primers, as no universal protist primers exist due to host-parasite phylogenetic overlap [20]. Anti-metazoan primers [21] and a reverse large subunit (LSU) rRNA gene primer [9] enhanced detection and phylogenetic resolution.

A key strength of this study was pooling amplicons from different primer sets before sequencing, a strategy previously explored for mitochondrial and nuclear genes [22] and aquatic microbial communities [23]. This method improves efficiency compared to metagenomics while enabling accurate recovery of taxonomic markers using tools like Minibar and cutadapt [24, 25]. Using this approach, we identified protists and other pathogens in reared insects and generated novel genomic data for microsporidia, gregarines and fungi. Moreover, small subunit (SSU) rRNA gene and LSU phylogenies provided new insights into protist biodiversity. By integrating TGS with amplicon pooling, this study refines long-read metabarcoding and enhances diagnostic tools for parasites in insects of economic importance.

METHODS

Insect samples

Host insects investigated in this study were either commercially reared or wild-caught. Reared insects were purchased in four different pet shops: Exeter Exotics (Exeter, UK), Pets at Home (Exeter, UK), Live Foods (UK) and Monis (Copenhagen, Denmark). The following reared insect species were screened for microsporidia and protist parasites: *Zophobas morio*, *T. molitor*, *Schistocerca gregaria*, *Gryllus bimaculatus*, *Gryllus assimilis* and *Galleria mellonella*. In different locations around Exeter, UK, the following selected wild Lepidoptera were sampled and screened: *G. mellonella* from Streatham Campus, University of Exeter (50° 44' 14.0" N 3° 32' 03.5" W) and Cockington apiary, Torbay (50° 27' 49.7" N 3° 33' 50.5" W), where larvae or adults were collected from active beehives; *Aphomia sociella* from Streatham Campus, University of Exeter, collected from bumblebee colonies; and *Pieris rapae* from Shillingford Organics, Exeter (50° 41' 14.4" N 3° 33' 04.4" W), where caterpillars were hand-collected and adults were caught with butterfly nets from cabbage crops.

Insect dissection and DNA extraction

All reared and wild insects were processed using the same protocol. Insects were anesthetized by freezing before dissection. Under a stereomicroscope, the gut and fat tissues of an individual insect were extracted in a Petri dish and triturated with fine forceps in ca. 250 µl of 1× PBS. Although gregarines were occasionally observed as large parasites during tissue dissection, microscopic detection of parasites was not a prerequisite for sample processing; all dissected insects were subjected to DNA extraction regardless of visible infection signs. The formed 'dissection pool' and tissues were collected in a 2-ml SafeLock tube for later DNA extraction. The DNA was extracted using a Cetyltrimethylammonium

bromide (CTAB) procedure including phenol/chloroform/isoamyl. Briefly, tissues were homogenized with a FastPrep Cell Disrupter (Qbiogene, Carlsbad, CA) in a mix of 500 µl CTAB buffer and 400 µl glass beads (0.17–0.18 mm Ø, Sigma-Aldrich). The mixture was then incubated with 10 µl of proteinase K (400 µg ml⁻¹) at 54 °C overnight. After cooling, 2 µl of RNaseA (10 µg ml⁻¹) was added, and the solution was incubated at 37 °C for 20 min. Thereafter, the genomic DNA was extracted in 700 µl of phenol/chloroform/isoamyl alcohol (25:24:1) followed by a single extraction in chloroform [26]. The DNA was finally precipitated with 3 M NaOAc (0.1× the sample volume) and absolute ethanol (2.5× the total volume). The DNA was eventually dissolved in 50 µl of sterile water and stored at –20 °C for later PCR. A total of 296 DNA samples were extracted and used for further analysis. Sample information can be accessed in the following: [Samples Seq table.xlsx](#).

Gregarine DNA extraction and primer design

Gregarines were occasionally observed during dissection, and individual gregarines from *Acheta domesticus* and *Alphitobius diaperinus* were isolated for DNA extraction. Gregarines were gently spread out by triturating the insect host gut tissues with fine forceps in 1× PBS and transferred with a 200 µl pipette (avoiding gut tissue remnants) to a new glass Petri dish with fresh PBS for washing. The gregarine cells were then transferred to a 1.5-ml Eppendorf tube, minimizing PBS volume, and centrifuged for 30 s in a microcentrifuge. The pelleted gregarine cells were then used for CTAB genomic extraction as described in Insect dissection and DNA extraction. The samples were identified as ‘T Ga’ and ‘T Lg’ ([Samples Seq table.xlsx](#)). The obtained pure gregarine DNA was used to validate the designed primers, as a positive control, and provide reliable genomic data for subsequent phylogenetic analysis (unpublished data).

Gregarines are an important protist group infecting reared insects, yet no primer set was available for this group in the PR2 primer database v2 [27], so new 18S primers were designed based on Gregarinoidea (Terrestrial Gregarine clade) sequences from GenBank. The forward primer targeted a conserved region (approximate positions 1,200–1,400 bp) in the reference sequences, as indicated by a low mean entropy (0.11±0.05). By contrast, the reverse primer targeted a more variable region (~1,500–1,700 bp) with the aim of retrieving a more diverse species set, as shown by a higher mean entropy (0.74±0.05). The primers were designed to amplify a short region of ca. 350 bp. Primer specificity was tested *in silico* using TestPrime (SILVA) (Table S1). The alignment of the gregarine reference sequences used for primer design and the corresponding primer set is available in the following: [Gregarine-Primer-design](#).

DNA pooling and PCR

DNA extracts were pooled by host species and used for PCR with universal and group-specific primers targeting protists and microsporidia. Group-specific primers were combined with the broad LSU primer (iLSUr) for long amplicons (Table 1, Fig. S1, available in the online Supplementary Material). Their purity was assessed using a NanoDrop™ Lite Spectrophotometer (Thermo Fisher Scientific), and their concentrations were determined with a Qubit 4™ Fluorometer (Thermo Fisher Scientific) (Fig. S2). When assessing the DNA concentration of the initial samples, if it was assessed ‘too low’ by Qubit, or below 10 ng µl⁻¹, the extract was concentrated by ethanol precipitation and elution in a smaller volume of water (10 µl). To form the host DNA pools, each DNA sample was diluted to 10 ng µl⁻¹ in a 10 µl solution (v/v, Milli-Q water), and 2 µl from each sample was pooled (i.e. 20 ng of DNA), yielding a final volume dependent on the number of individuals per species.

Initial ‘short’ PCRs amplified SSU regions using universal and group-specific primers to identify positive samples. These samples were then subjected to a second ‘long’ PCR targeting both SSU and LSU regions, using the same forward 18S primer with the iLSUr reverse primer. Short and long PCR products were purified (Qiagen Gel PCR Clean-up). If the gel showed only the expected band, the entire mixture was purified; otherwise, the correct band was excised and purified. Nested PCRs were performed using long PCR amplicons as templates. The reverse complement of the group-specific reverse primer previously used served as the new forward primer, paired with iLSUr, to improve amplification.

Each 20 µl PCR reaction contained 2 µl of DNA (~20 ng), 1 µM primers, 200 µM dNTPs, 1× Phusion HF Buffer and 0.02 U Phusion DNA polymerase (NEB). The thermocycler programme included an initial 98 °C denaturation (30 s), 30 cycles of 98 °C denaturation (10 s), annealing (30 s) and 72 °C extension (2 min), with a final 72 °C extension (10 min). Long and nested PCRs used similar conditions with a 2 min 30 s extension. Negative PCR reactions were always undertaken with water used as template and optimal annealing temperature for each primer used.

PCR products (5 µl) were run on a 1.3% agarose gel (pre-mixed with SafeView™ dye) at 100 V for 30 min in TAE buffer and visualized under UV. Expected bands were excised, purified (Qiagen Gel PCR Clean-up), pooled by insect species and quantified (Qubit dsDNA HS Assay Kit, Thermo Fisher).

All the performed PCRs and respective gels can be found in the supplementary file 1 ‘Laboratory notes and PCR gels’ (Supplementary Material).

Table 1. List of primer pairs used to perform the PCRs

Tm: Melting or annealing temperature used during the PCR programme with Phusion polymerase (see the different used programmes in File S1). When the reverse primer iLSUr [77] is present, it means that positive samples were found within the targeted group and further used to produce long amplicons.

Parasite group	Primer name	Sequence (5'–3')	Product size (bp)	Tm (°C)	Primer reference
Microsporidia	V1F/CM-V5R	F: CACCAGGTTGATTCTGCCTGAC R: TAANCAGCACAMTCCACTC	ca. 780	59	[77–81]
	V1F/530R	F: CACCAGGTTGATTCTGCCTGAC R: CCGCGGCKGCTGGCAC	ca. 400	69	
	V1F/iLSUr	F: CACCAGGTTGATTCTGCCTGAC R: CCGCGGCKGCTGGCAC	ca. 3,500	67	
	CM-V5F/iLSUr	F: GATTAGANACNNNGTAGTTC R: ACCTGTCTCAGCAGGTCTAAAC	2,800–3,000	55	
	CM-V5R(rc)/iLSUr	F: GAGTGGAKTGTGCTGNTTA R: ACCTGTCTCAGCAGGTCTAAAC	2,700–2,900	58	
Gregarinasina	WL1/EukP3	F: GCGCTACCTGGTTGATCCTGCC R: GACGGGCGGTGTGTAC	ca. 1,600	55	[82]
	1-18S-GRF/1222-18S-GRR	F: CGGTAATTCCAGCTCCAAAT R: TGACTTGCGCTTACTAGGG	ca. 950	55	[83]
	GregF/GregR	F: CCCTTAGATRRYCTGGGCTGC R: CGTGTTACGACTTCTTC	ca. 330	55	Designed; [77]
	GregF/iLSUr	F: CCCTTAGATRRYCTGGGCTGC R: ACCTGTCTCAGCAGGTCTAAAC	3,300–4,000	64	
	GregR_RC/iLSUr	F: GAAGAAGTCGTAACACG R: ACCTGTCTCAGCAGGTCTAAAC	3,000–3,700	56	
Cryptosporidia	SCL2/SCR2	F: CAGTTATAGTTTACTTGATAATC R: CAATACCCTACCGTCTAAAG	200	58	[84]
Chlorophyta	MGF/MGR	F: AGGATAGAGGCCTACCATGGTTTCAA R: CCCCAGCTGTCCCTCTCCAT	569	67	[85]
Amoebozoa	AmoebozoaF/AmoebozoaR	F: GAATTGACGGAAGGGCACAC R: CCAAGAYRTCTAAGGGCATCAC	360–420	60	[77, 86]
	AmoebozoaF/iLSUr	F: GAATTGACGGAAGGGCACAC R: ACCTGTCTCAGCAGGTCTAAAC	ca. 4,500	66	
	AmoebozoaR(rc)/iLSUr	F: GTGATGCCCTTAGAYRTCTTGG R: ACCTGTCTCAGCAGGTCTAAAC	ca. 4,200	64	
Ichthyosporea	500F/900R	F: CGGCTACCACTTCTACGGAGG R: ATTAACGCCCCCACTATCCC	470–509	66	[87]
Apicomplexa	Apicomplexa B F/ Apicomplexa B R	F: TGYGTTTGAATACTAYAGCATGG R: TCTGATCGTCTTCACTCCCTT	228–244	62	[86]
Kinetoplastea	Kineto14F/kineto2026R	F: CTGCCAGTAGTCATATGCTTGTTCAGGA R: GATCCTTCTGCAGGTTACCTACAGCT	1,900–2,200	68	[77, 88]
	Kineto14F/iLSUr	F: CTGCCAGTAGTCATATGCTTGTTCAGGA R: ACCTGTCTCAGCAGGTCTAAAC	6,000–7,000	67	
	kineto2026R(rc)/iLSUr	F: AGCTGTAGGTGAACCTGCAGAAGGATC R: ACCTGTCTCAGCAGGTCTAAAC	ca. 5,000	67	
Universal non-metazoan primers	574*f/UNonMet_DB	F: CGGTAAYTCCAGCTCYV R: CTTTAARTTTCASYCTTGCG	ca. 574 (median length)	56	[21, 89]

Oxford Nanopore sequencing

Nanopore sequencing followed the ONT native barcoding protocol vNBA_9093_v109_revJ_12Nov2019 using EXP-NBD104 and SQK-LSK109 kits [28]. Library preparation involved native barcode and sequencing adapter ligation. The final library, prepared as recommended, was purified (AMPure XP beads), quantified (Qubit) and sequenced on a MinION (Oxford Nanopore Technologies) using an R9 flow cell. ONT MinKNOW software collected raw sequencing data in fasta5 format. Three MinION runs were performed (Table S2).

Bioinformatic analysis

Raw fast5 reads were basecalled using Guppy v6.4.8 (*guppy_basecaller*) with a minimum quality score of 7 and the *dna_r9.4.1_450bps_hac.cfg* configuration. Reads were then demultiplexed (*guppy_barcode*) with *--detect_mid_strand_barcode*s to improve barcode detection. Fastq files were merged and processed with an in-house bash script.

Briefly, reads were trimmed using cutadapt v3.5 [24], with the non-metazoan and group-specific primers for short SSU amplicons and the reverse iLSUr for long amplicons (Table 1). For long reads, only the nested primer was searched to avoid repetitive sequences. A 0.1 error rate was applied for primer alignment [29]. Quality filtering was performed using fastq_quality_trimmer (FASTX Toolkit v0.0.14) [30] and seqkit [31] based on the expected amplicon size. Consensus sequences were generated with NGSspeciesID [17], using the options *ont*, *consensus*, *medaka* and *sample size* 300, as recommended for ONT amplicon sequencing data [25, 32, 33]. If more than one consensus was formed, due to co-amplifications, only the consensus classified as an expected taxon (see below) were further kept for phylogenetic analysis [22] [[Amplicons Table](#), see the column 'Included in phylo. Analysis'].

The general reproducible bash script is available via the following link, [amplicon.sh](#), and can be adapted with custom primers and expected amplicon length. Demultiplexed ONT fastq reads, cutadapt intermediate files, NGSspeciesID outputs and specific scripts are accessible at [Zenodo.15129942](#).

Amplicon consensus sequences were assigned taxonomically via MEGABLAST against the NCBI nt database, using the hit with the highest bit score to select the best taxonomic match. We did not apply a fixed sequence identity cutoff, as our aim was to retain the closest available matches. All MEGABLAST searches were run online against the NCBI nt database [34] accessed in November 2024. Validated consensus sequences that showed an expected taxon and had more than one supporting read were gathered by group (Gregarinasina, Microsporidia and Fungi) and clustered with CD-HIT [35] at a 98% threshold, commonly used for species delimitation with 18S markers [36–39], before phylogenetic analysis.

The specificity of each primer set was eventually assessed by dividing the number of consensus sequences corresponding to the expected target taxon by the total number of recovered consensus sequences:

$$\text{Specificity (\%)} = \frac{\text{Number of expected taxon consensus sequences}}{\text{Total recovered consensus sequences}} \times 100$$

Phylogenetic analysis

Standard BLASTN search (*E*-value cutoff of 1e-10) was used to identify homologous sequences for each Nanopore amplicon consensus against the NCBI non-redundant database [40, 41]. The top hits (lowest *e*-value) were downloaded and aligned with the amplicon consensus sequences using MAFFT v7 [42] and then trimmed with BMGE -g 0.3 [43]. Phylogenies for the different recovered taxa were constructed separately for the SSU and LSU genes, when possible. Maximum-likelihood trees were built with FastTree v2.1.11 and the general time reversible (GTR) model (gamma option, 1,000 bootstraps) [44, 45]. Trees were visualized with the *ape* and *ggtree* R packages and eventually edited in Adobe Illustrator [46–48].

One concatenated tree was constructed by creating a matrix from gregarine consensus sequences recovered from the same host that had both 18S and 28S data. For this, trimmed and curated sequences were concatenated with FASconCAT-G [49], and the tree was made with the same method as above. For the gregarine–host co-phylogeny, one representative nucleotide sequence of the cytochrome c oxidase subunit I (COI) gene for each host species was downloaded from GenBank. A maximum-likelihood tree with 1,000 bootstraps was constructed as above using MAFFT and BMGE, with the exception of using iqtree2 and the GTR+G5+I model [50–53]. The different raw files used for the phylogenetic analyses can be found in the following: [Phylogeny-Alignments](#).

RESULTS

Sequencing output and primer performance

Three Nanopore MinION runs generated a total of 1,995,585 reads after basecalling and demultiplexing. Nanopore barcode scores ranged from 82.5 to 96.3, indicating successful ONT barcode assignment and ligation (Table S1). The 'Run 2' showed substantially lower read counts due to low ONT library concentration. However, this run still yielded sufficient coverage for consensus generation.

The newly designed GregF/GregR primer set, targeting Gregarinasina, generated 45,758 reads [as determined by cutadapt; see the [Amplicons Table](#)]. These reads produced 15 gregarine consensus sequences (each supported by at least two ONT reads) using NGSspeciesID and BLAST search, yielding a gregarine specificity of 75% (i.e. the number of expected taxa divided by all taxa recovered from NGSspeciesID consensus sequences). By contrast, the 1-18S-GRF/1222-18S-GRR primer set generated 18,965 reads but did not produce any gregarine consensus sequences, indicating no gregarine specificity (Table 2). When used with a broad LSU reverse primer, the nested PCR primer set GregRC/iLSUr, where GregRC (the reverse complement of GregR) is applied as a nested primer following a long PCR with GregF and iLSUr, generated 54 reads and produced 11 gregarine consensus sequences. The AmoebozoaF/AmoebozoaR pair, expected to amplify amoebas, produced 2,182 reads

Table 2. Different primer pairs used and the taxonomy assignment of the obtained amplicon consensus (results shown for consensus supported by at least two reads)

The taxonomy was determined via BLASTN (NCBI nt database). Primer reads correspond to the number of reads successfully recovered and trimmed by cutadapt [24]. The specificity was determined, for a given primer pair, by dividing the number of expected consensus by the total of produced consensus. Data used to construct this table are available in the following: [Amplicons Table](#). Greg, gregarine; Microsp., microsporidia.

Targeted group	Primer set	Primer reads	NGSpeciesID consensus							Specificity (%)
			Fungi	Fungi incertae sedis	Greg.	Microsp.	Bacteria	Virus	Insect	
Gregarinasina	1-18S-GRF/1222-18S-GRR	18,965	5	-	-	1	-	-	2	0
	WL1/EukP3	289	-	-	1	-	-	-	1	50
	GregF/GregR	45,758	-	-	15	-	3	1	1	75
	GregR_RC/iLSUr	54	-	-	11	-	-	-	-	100
Amoebozoa	AmoebozoaF/AmoebozoaR	2,182	-	-	-	-	-	-	15	0
	AmoebozoaR_RC/iLSUr	3	2	-	-	-	-	-	-	0
Kinetoplastea	Kineto14F/kineto2026R	72	9	1	4	-	-	-	-	0
	Kineto2026R_RC/iLSUr	12	-	-	4	-	-	-	-	0
Microsporidia	V1F/530R	18,824	-	-	1	6	1	-	-	75
Non-metazoa	574*f/UNonMet_DB	1,809	7	2	3	-	-	-	1	92

but only generated insect consensus sequences, indicating low specificity (Table 2). For Microsporidia, the V1F/530R set generated 18,824 reads and allowed to recover 6 microsporidia consensus sequences, showing high specificity (Table 2). The non-metazoan 574*f/UNonMet_DB primer set retrieved 1,809 reads from which 3 gregarine and 9 fungal consensus were created (Table 2).

PHYLOGENETIC RESULTS

Microsporidia

Four microsporidian consensus sequences were recovered using V1F/530R, supported by 69–205 reads ([Amplicons Table](#)). Most clustered with *Albopleistophora grylli* (Pleistophoridae sp. YST), a known microsporidian species infecting *G. bimaculatus* and *G.*

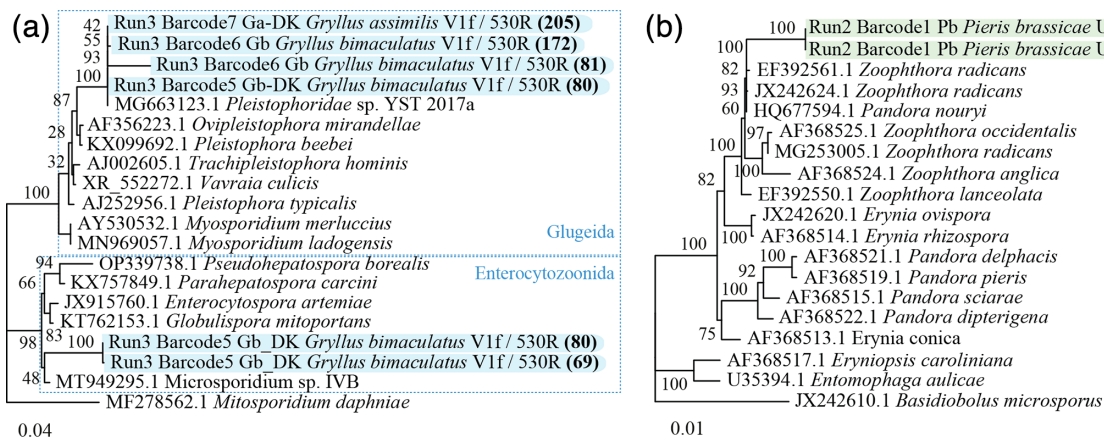


Fig. 1. (a) Microsporidia 18S phylogenetic tree. When the same species is represented several times, its sampling location differed. 'DK' was sampled in Denmark and the rest in the UK. Tips depicting amplicons from the same run for *G. bimaculatus* barcode 5 indicate a potential co-infection. (b) Fungi 18S phylogenetic tree. *Furia* have been renamed to *Pandora* as these were recently described to be synonymous [72]. The trees were constructed using the GTR model, and a bootstrap analysis was performed with 1,000 replicates. The number of reads supporting the different consensus sequences used for phylogenetic analyses is shown in bold and in parentheses.

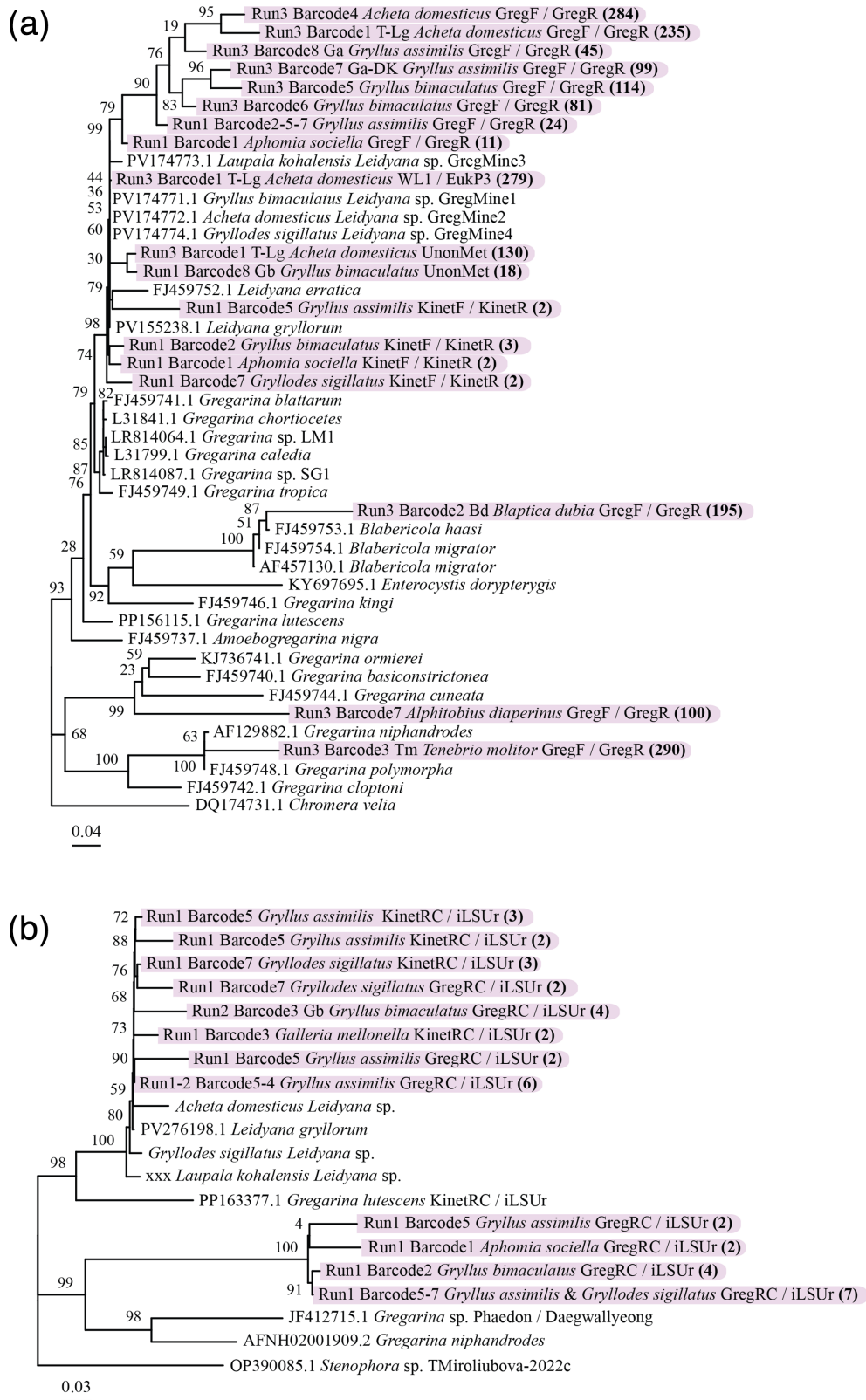


Fig. 2. (a) Gregarine 18S and (b) 28S phylogenetic tree. The trees were generated with the GTR model, and a bootstrap analysis was performed with 1,000 replicates. The number of reads supporting the different consensus sequences used for phylogenetic analyses is shown in bold and in parentheses.

assimilis [54, 55], with 100% bootstrap support (Fig. 1a). Another sequence clustered with *Microsporidium* sp. IVB, which was derived from *G. bimaculatus* tissue (98% bootstrap support), previously isolated from *Gammarus* spp., suggesting a potential co-infection or incidental ingestion of spores through contaminated feed.

Fungi

Multiple yeast and fungal taxa were recovered, including genera such as *Debaryomyces*, *Hyphopichia*, *Tetrapispora*, *Kodamaea* and *Penicillium*. While most of these are common organisms and not typically associated with insect pathology, two noteworthy fungal amplicon consensus sequences, belonging to the Eryniaceae (Entomophthoromycotina), were recovered from *Pieris brassicae* using the 574*f/UNonMet_DB primer pair (Fig. 1b) [56]. The SSU amplicon consensus sequences from this group matched either *Pandora nouri* or *Zoophthora radicans* (>99% identity), both of which are well documented as entomopathogenic fungi [see BLAST results in [Amplicons Table](#)]. The phylogenetic placement of *P. nouri* in the SSU tree shows that the species clusters with *Zoophthora* spp., suggesting that the amplicon sequences likely belong to a single *Zoophthora* species infecting *P. brassicae*.

Gregarines

With regard to gregarines, SSU phylogenetic analysis confirmed expected placements. A consensus sequence from *Blaptica dubia* clustered with *Blabercola migrator*, a gregarine infecting blaberid cockroaches. Another sequence from *T. molitor* had 98.28% similarity to *Gregarina polymorpha*, its known gregarine parasite (Fig. 2) [57].

Amplicons from *A. diaperinus* clustered within a clade comprising *Gregarina basiconstrictonea* and *Gregarina ormierei*, parasites of darkling beetles (Tenebrionidae) [58]. The tree topology and sequence similarities suggest that *A. diaperinus* consensus sequences are representing a species lacking molecular data to date. It could be *Gregarina alphetobii*, previously described, solely based on morphology [59].

Another interesting clade is the one encompassing all the cricket (Gryllidae) amplicon consensus sequences (Fig. 2a). This clade includes the references of *Leidyana erratica* and *Leidyana gryllorum*, the latter having been published recently (GenBank accessions FJ459752 and PV155238, respectively). An amplicon consensus from *A. sociella* falling within this clade suggests possible contamination (i.e. from the host during handling in the lab, during sample pooling or barcode demultiplexing). Past work has also shown that gregarine oocysts can be volatile and could easily contaminate a lab environment [60, 61].

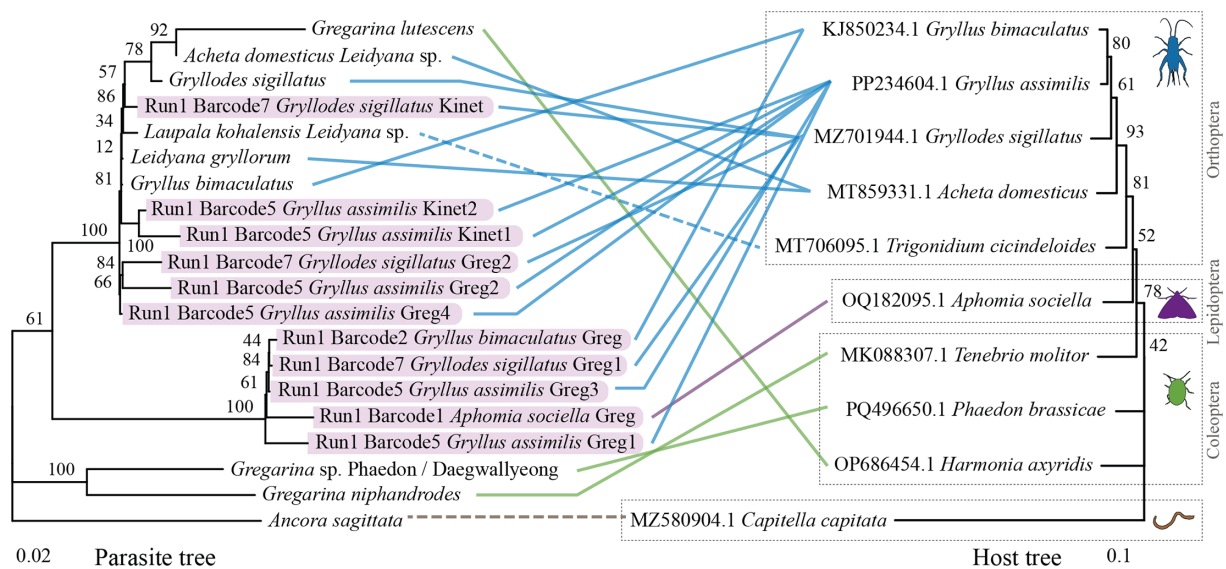


Fig. 3. Gregarine–host co-phylogenetic relationship. The left phylogeny depicts the gregarine phylogenetic tree performed with concatenated 18S and 28S sequences. The phylogeny on the right is a rooted maximum likelihood tree of the gregarine–host cytochrome c oxidase subunit I (COI) gene based on nucleotide sequences recovered from GenBank. Hosts are highlighted with boxes and silhouette images (modified from <https://www.phylopic.org/>). The outgroup *Capitella capitata* was selected as it is the host of the gregarine phylogeny outgroup (*Ancora sagittata*). As there is no available COI gene for *Laupala kohalensis*, the COI gene of the close relative *Trigonidium cicindeloides* was used. The lines between the two phylogenies depict the gregarine–host relationship, with the colours indicating host orders.

Long-read sequencing of gregarine SSU-LSU regions yielded 11 consensus sequences. Two clusters were identified (Fig. 2b). The first is formed with *L. gryllorum*, while the other lacks an available reference. Gregarine and host co-phylogenetic trees support the divergence of a potential new gregarine clade infecting cricket hosts (Fig. 3).

DISCUSSION

Primer design for metabarcoding eukaryotic endosymbionts is challenging, requiring specificity for target parasites while minimizing host DNA amplification [21, 29]. The V4 region of the 18S gene is often selected to reduce off-target amplification [21, 62, 63]. *In silico* tools like TestPrime [64] assess primer specificity, but our results emphasize that *in situ* validation through PCR remains essential for reliability and to demonstrate how they perform on samples that contain a complex mix of organisms. For example, Amoebozoan primers only retrieved insect DNA in this study.

The GregF/GregR primer set recovered more gregarine sequences than previously published primers, demonstrating its potential efficiency and high specificity. The non-metazoan primer 574*f/UNonMet_DB successfully amplified non-metazoan DNA but resulted in few protistan sequences and lower diversity of gregarines than the GregF/GregR primer set. One possible optimization for this primer pair is adjusting the annealing temperature, which can influence which organisms are successfully amplified, but this requires specific optimization for each taxon, which is labour intensive, and results can be unpredictable when working on taxa with non-previous molecular data [21]. The AmoebozoaF/AmoebozoaR and V1F/CM-V5R primers had poor specificity, amplifying insect, bacterial and fungal DNA [see Amplicons Table]. Balancing amplification efficiency and specificity remains a key challenge.

Primers yielding shorter reads, like GregF/GregR, could still recover diverse and novel sequences even in the presence of host DNA. This study highlights a scarcity of gregarine data in reference databases and phylogenies [8], especially from crickets. Previous studies reported potential cross-infections among gregarines infecting Gryllidae [65]. However, more recent investigations that integrate both morphological and molecular evidence suggest that gregarines may in fact exhibit a narrower host range, being limited to a single host species. In these newer studies, what was once interpreted as cross-infection may instead represent phenotypic plasticity, where variations in gregarine morphology are driven by host subspecies differences and geographical factors, which likely biased earlier morphology-based assessments [66, 67]. Our co-phylogenetic analysis (Fig. 3) revealed a potential novel Gryllidae-infecting clade that will need further research to fully describe these gregarines. Our study also provides one of the first molecular identifications of a gregarine infecting *A. diaperinus*, closely related to *G. basiconstrictonea* and *G. ormierei*, known from tenebrionid and scarabaeid beetles [67, 68]. However, the 84.75% identity with our sequence suggests that a different species, possibly *G. alphetobii* as described by Bala *et al.* [59], might be infecting *A. diaperinus*.

Regarding fungal organisms, a potential *Zoophthora* sp. was detected in *P. brassicae* using the 574*f/UNonMet_DB primer pair, suggesting a natural or accidental infection. The SSU phylogenetic analysis clustered the consensus sequences with *Z. radicans*, a well-known entomopathogen associated with insect epizootics [69]. Such fungi could potentially help control *P. brassicae* populations in organic farms, reducing economic losses to Brassica crops. Interestingly, although *P. nouryi* was initially described and isolated from aphid hosts [70, 71], systematic studies on its taxonomic status are scarce. This may explain its unexpected clustering and sequence similarity with *Z. radicans* (99.7% identity between *P. nouryi* accession HQ677594 and *Z. radicans* accession MG252997). The Erynioideae subfamily, which includes the genera *Zoophthora*, *Erynina*, *Furia* (recently synonymized with *Pandora* [72]) and *Pandora*, shows taxonomic inconsistencies; while *Zoophthora* is monophyletic, the other three genera are polyphyletic and in need of revision based on molecular data [73–75]. Regarding the fungi sister clade Microsporidia, the V1F/530R primer detected a potential co-infection in *G. bimaculatus*, reinforcing its effectiveness. The detection of *A. grylli* (Pleistophoridae sp.) in this cricket host species has been recently described [55], but a potential infection by an Enterocytozoonida will require further investigation.

This study highlights the importance of primer validation for accurate microeukaryote metabarcoding in reared insects and other hosts. By combining a cocktail of group-specific primers with an amplicon pooling strategy and Nanopore barcoding, we achieved a high-throughput, multiplexed approach that overcomes the limitations of traditional PCR assays targeting microeukaryote pathogens. This integrated method not only reduces time and cost compared to conventional techniques and Illumina-based approaches [29], but it also enhances the resolution of parasite systematics and improves diagnostic accuracy for detecting protist and microsporidian pathogens. Although similar high-throughput amplicon sequencing methods have been applied to broader eukaryotic diversity surveys [76], our application to reared insect systems is relatively unique. Moreover, the ongoing need for more accurate DNA barcodes to generate well-supported phylogenies makes the development of gregarine-specific primers (GregF/GregR) a significant advancement, as this economically important group was previously absent from major primer databases. [10, 22]. Our detection of novel gregarine lineages in crickets and the first molecular characterization of gregarines in *A. diaperinus* demonstrates the capacity of this approach to uncover hidden pathogen diversity with direct applications for pathogen surveillance in the growing insect farming industry.

The amplicon pooling strategy presented here can be readily adapted to other host–parasite systems and expanded to include additional primer sets targeting neglected groups such as Amoebozoa or Kinetoplastea. By combining practical diagnostics with phylogenetically informative data, this Nanopore-based approach provides a valuable tool for both improved biosecurity measures in insect mass-rearing systems and advancing our understanding of protist evolution and host–parasite relationships.

Funding information

This work was done within the project 'Insect Doctors' which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 859850.

Acknowledgements

We would like to thank Thea Kristensen who proofread the manuscript and Fang Shiang Lim who gave us precious advice on ONT methods and analyses.

Author contributions

E.B.: methodology, investigation, formal analysis, data curation, visualization, validation and writing – original draft. S.E.: investigation, formal analysis, visualization and writing – review and editing. B.A.P.W.: conceptualization, methodology, validation, supervision, funding acquisition and writing – review and editing.

Conflicts of interest

The authors declare that there are no conflicts of interest.

Disclaimer

During the preparation of this work, the author Edouard Bessette utilized GitHub Copilot (Visual Studio Code extension) to aid in the creation and refinement of the script amplicon.sh (<https://github.com/Edouard94>). Following the assistance provided by this tool, the author thoroughly reviewed and edited the content as necessary, assuming full responsibility for it.

References

1. Lange CE, Lord JC. Protistan Entomopathogens. In: *Insect Pathology*. Elsevier, 2012. pp. 367–394.
2. Stentiford GD, Becnel JJ, Weiss LM, Keeling PJ, Didier ES, *et al.* Microsporidia – Emergent Pathogens in the Global Food Chain. *Trends Parasitol* 2016;32:336–348.
3. Braun L. The life history of *Malameba locustae* (King and Taylor) in the migratory grasshopper *Melanoplus sanguinipes* (F.). University of Saskatchewan 1987.
4. Kleespies RG, Huger AM, Zimmermann G. Database on Arthropod Diseases; 2008. <https://arthropodenkrankheiten.julius-kuehn.de> [accessed 19 February 2021].
5. Giglioli AAS, Julio AHF, Conte H. The life cycle of *Gregarina cuneata* in the midgut of *Tribolium castaneum* and the effects of parasitism on the development of insects. *Bull Entomol Res* 2016;106:258–267.
6. Rueckert S, Betts EL, Tsousis AD. The symbiotic spectrum: where do the gregarines fit? *Trends Parasitol* 2019;35:687–694.
7. Becnel JJ, Andreadis TG. Microsporidia in Insects. In: *Microsporidia*. John Wiley & Sons, Ltd, 2014. pp. 521–570.
8. Bessette E, Williams B. Protists in the insect rearing industry: benign passengers or potential risk? *Insects* 2022;13:482.
9. Bojko J, Reinke AW, Stentiford GD, Williams B, Rogers MSJ, *et al.* Microsporidia: a new taxonomic, evolutionary, and ecological synthesis. *Trends Parasitol* 2022;38:642–659.
10. Gaonkar CC, Campbell L. A full-length 18S ribosomal DNA meta-barcoding approach for determining protist community diversity using Nanopore sequencing. *Ecol Evol* 2024;14:e11232.
11. Rhoads A, Au KF. PacBio Sequencing and its Applications. *Genomics, Proteomics & Bioinformatics* 2015;13:278–289.
12. Hui Y, Nielsen DS, Krych L. De novo clustering of long-read amplicons improves phylogenetic insight into microbiome data; 2023. <http://biorxiv.org/lookup/doi/10.1101/2023.11.26.568539> [accessed 27 May 2024].
13. Jeon M-S, Jeong DM, Doh H, Kang HA, Jung H, *et al.* A practical comparison of the next-generation sequencing platform and assemblers using yeast genome. *Life Sci Alliance* 2023;6:e202201744.
14. Zhang T, Li H, Ma S, Cao J, Liao H, *et al.* The newest Oxford Nanopore R10.4.1 full-length 16S rRNA sequencing enables the accurate resolution of species-level microbial community profiling. *Appl Environ Microbiol* 2023;89:e0060523.
15. Beaulaurier J, Luo E, Eppley JM, Uyl PD, Dai X, *et al.* Assembly-free single-molecule sequencing recovers complete virus genomes from natural microbial communities. *Genome Res* 2020;30:437–446.
16. Sahlin K, Medvedev P. De novo clustering of long-read transcriptome data using a greedy, quality value-based algorithm. *J Comput Biol* 2020;27:472–484.
17. Sahlin K, Lim MCW, Prost S. NGSspeciesID: DNA barcode and amplicon consensus generation from long-read sequencing data. *Ecol Evol* 2021;11:1392–1398.
18. Hall M, Beiko RG. 16S rRNA Gene Analysis with QIIME2. In: Beiko RG, Hsiao W and Parkinson J (eds). *Microbiome Analysis*. New York, NY: Springer New York; 2018. pp. 113–129.
19. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 2019;37:852–857.
20. Kounosu A, Murase K, Yoshida A, Maruyama H, Kikuchi T. Improved 18S and 28S rDNA primer sets for NGS-based parasite detection. *Sci Rep* 2019;9:15789.
21. Minardi D, Ryder D, Del Campo J, Garcia Fonseca V, Kerr R, *et al.* Improved high throughput protocol for targeting eukaryotic symbionts in metazoan and eDNA samples. *Mol Ecol Resour* 2022;22:664–678.
22. Gajski D, Wolff JO, Melcher A, Weber S, Prost S, *et al.* Facilitating taxonomy and phylogenetics: An informative and cost-effective protocol integrating long amplicon PCRs and third-generation sequencing. *Mol Phylogenet Evol* 2024;192:107988.
23. Vass M, Ramasamy KP, Andersson A. Microbial hitchhikers on microplastics: the exchange of aquatic microbes across distinct aquatic habitats. *Environ Microbiol* 2024;26:e16618.
24. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet j* 2011;17:10.
25. Krehenwinkel H, Pomerantz A, Henderson JB, Kennedy SR, Lim JY, *et al.* Nanopore sequencing of long ribosomal DNA amplicons enables portable and simple biodiversity assessments with high phylogenetic resolution across broad taxonomic scale. *GigaScience* 2019;8:giz006.
26. Chen Y ping, Pettis JS, Zhao Y, Liu X, Tallon LJ, *et al.* Genome sequencing and comparative genomics of honey bee microsporidia,

- Nosema apis reveal novel insights into host-parasite interactions. *BMC Genomics* 2013;14:451.
27. Vaulot D. pr2-primers version 2.0.0; 2023. <https://zenodo.org/record/7835730> [accessed 12 April 2024].
 28. Oxford Nanopore. Ligation sequencing amplicons - native barcoding (SQK-LSK109 with EXP-NBD104 and EXP-NBD114); 2022. https://community.nanoporetech.com/docs/prepare/library_prep_protocols/native-barcoding-amplicons/v/nba_9093_v109_rev_nov_12nov2019 [accessed 6 December 2022].
 29. Reis AL, Beckley LE, Olivar MP, Jeffs AG. Nanopore short-read sequencing: A quick, cost-effective and accurate method for DNA metabarcoding. *Environmental DNA* 2023;5:282–296.
 30. Gordon A. FASTQ/a short-reads preprocessing tools. 2010.
 31. Shen W, Le S, Li Y, Hu F. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS ONE* 2016;11:e0163962.
 32. Vasiljevic N, Lim M, Humble E, Seah A, Kratzer A, et al. Developmental validation of Oxford nanopore technology MinION sequence data and the NGSspeciesID bioinformatic pipeline for forensic genetic species identification. *Forensic Sci Int Genet* 2021;53:102493.
 33. Pomerantz A, Sahlin K, Vasiljevic N, Seah A, Lim M, et al. Rapid in situ identification of biological specimens via DNA amplicon sequencing using miniaturized laboratory equipment. *Nat Protoc* 2022;17:1415–1443.
 34. NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 2016;44:D7–D19.
 35. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 2012;28:3150–3152.
 36. Rodriguez Y, Omoto CK, Gomulkiewicz R. Individual and population effects of eugregarine, *Gregarina niphandrodes* (Eugregarinida: Gregarinidae), on *Tenebrio molitor* (Coleoptera: Tenebrionidae). *Environ Entomol* 2007;36:689–693.
 37. de Vargas C, Audic S, Henry N, Decelle J, Mahé F, et al. Eukaryotic plankton diversity in the sunlit ocean. *Science* 2015;348:1261605.
 38. Smith KF, Kohli GS, Murray SA, Rhodes LL. Assessment of the metabarcoding approach for community analysis of benthic-epiphytic dinoflagellates using mock communities. *New Zealand J Marine Freshwater Res* 2017;51:555–576.
 39. Massana R, Gobet A, Audic S, Bass D, Bittner L, et al. Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. *Environ Microbiol* 2015;17:4035–4049.
 40. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997;25:3389–3402.
 41. NCBI. Nucleotide [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information. [Internet]; 1988. <https://www.ncbi.nlm.nih.gov/nucleotide/> [accessed 9 June 2021].
 42. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013;30:772–780.
 43. Criscuolo A, Gribaldo S. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol Biol* 2010;10:210.
 44. Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 2009;26:1641–1650.
 45. Abadi S, Azouri D, Pupko T, Mayrose I. Model selection may not be a mandatory step for phylogeny reconstruction. *Nat Commun* 2019;10:934.
 46. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 2019;35:526–528.
 47. Yu G, Smith DK, Zhu H, Guan Y, Lam TT. Ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol* 2017;8:28–36.
 48. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2022. <https://www.R-project.org/>
 49. Kück P, Longo GC. FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. *Front Zool* 2014;11:81.
 50. Pentinsaari M, Salmela H, Mutanen M, Roslin T. Molecular evolution of a widely-adopted taxonomic marker (COI) across the animal tree of life. *Sci Rep* 2016;6:35275.
 51. Goh KS, Wang LJ, Ni JH, Wang TY. Luminescent characteristics and mitochondrial COI barcodes of nine cohabitated Taiwanese fireflies. *PeerJ* 2022;10:e14195.
 52. Hendrich L, Pons J, Ribera I, Balke M. Mitochondrial cox1 sequence data reliably uncover patterns of insect diversity but suffer from high lineage-idiosyncratic error rates. *PLoS ONE* 2010;5:e14448.
 53. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 2020;37:1530–1534.
 54. Tokarev YS, Peat KM, Malysh JM, Senderskiy IV. Discovery of a novel microsporidium in laboratory colonies of Mediterranean cricket *Gryllus bimaculatus*. 2018;117:2823–2829. DOI: 10.1007/s00436-018-5970-z.
 55. Bessette E, Bojko J, Bateman KS, Ross S, Meyling NV, et al. Identification of *Alboplectrophora grylli* n. gen. n. sp. (Microsporidia) and its impact on crickets (*Gryllus* spp.) in food-and-feed culture systems. *J Invertebr Pathol* 2025;208:108229.
 56. Gryganskyi AP, Hajek AE, Voloshchuk N, Idnurm A, Eilenberg J, et al. Potential for use of species in the subfamily erynioidea for biological control and biotechnology. *Microorganisms* 2024;12:168.
 57. Harry OG. The effect of a eugregarine *Gregarina polymorpha* (Hammerschmidt) on the mealworm larva of *Tenebrio molitor* (L.). *J Protozool* 1967;14:539–547.
 58. Théodoridès J. Les Eugregarines du genre *Gregarina* parasites de Coléoptères Ténébrionides. *Ann Parasitol Hum Comp* 1955;30:5–21.
 59. Bala P, Kaur D, Lipa J, Bhagat R, et al. *Gregarina alphetobii* sp. n. and *matteia alphetobii* sp. n., parasitizing alphetobius diaperinus panz. (tenebrionidae, coleoptera). 1990.
 60. Bessette E, Williams BAP, Meyling NV. Preparing eugregarine parasites and their cricket host *Acheta domesticus* as a model for gregarine infection studies. *MethodsX* 2024;13:102888.
 61. Harry OG. A jar for maintaining parasite-free insects and for collecting infected faeces. *Bull Entomol Res* 1969;58:833–834.
 62. Bukin YS, Mikhailov IS, Petrova DP, Galachyants YP, Zakharova YR, et al. The effect of metabarcoding 18S rRNA region choice on diversity of microeukaryotes including phytoplankton. *World J Microbiol Biotechnol* 2023;39:229.
 63. Owens LA, Friant S, Martorelli Di Genova B, Knoll LJ, Contreras M, et al. VESPA: an optimized protocol for accurate metabarcoding-based characterization of vertebrate eukaryotic endosymbiont and parasite assemblages. *Nat Commun* 2024;15:402.
 64. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2012;41:D590–D596.
 65. Corbel JC. La spécificité parasitaire des Grégariens d'Orthoptères. *Ann Parasitol Hum Comp* 1968;43:25–32.
 66. Florent I, Chapuis MP, Labat A, Boisard J, Leménager N, et al. Integrative taxonomy confirms that *Gregarina garhami* and *G. acridiorum* (Apicomplexa, Gregarinidae), parasites of *Schistocerca gregaria* and *Locusta migratoria* (Insecta, Orthoptera), are distinct species. *Parasite* 2021;28:12.
 67. Criado-Fornelio A, Verdú-Expósito C, Martín-Pérez T, Heredero-Bermejo I, Pérez-Serrano J, et al. A survey for gregarines (Protozoa: Apicomplexa) in arthropods in Spain. *Parasitol Res* 2017;116:99–110.

68. Clopton RE. Phylogenetic relationships, evolution, and systematic revision of the septate gregarines (Apicomplexa: Eugregarinorida: Septatorina). *Comparative Parasitology* 2009;76:167–190.
69. Torres Acosta RI, Humber RA, Sánchez-Peña SR. Zoophthora radicans (Entomophthorales), a fungal pathogen of *Begrada hiliaris* and *Bactericera cockerelli* (Hemiptera: Pentatomidae and Triozidae): prevalence, pathogenicity, and interplay of environmental influence, morphology, and sequence data on fungal identification. *J Invertebr Pathol* 2016;139:82–91.
70. Scorsetti AC, Jensen AB, López Lastra C, Humber RA. First report of *Pandora neoaphidis* resting spore formation in vivo in aphid hosts. *Fungal Biol* 2012;116:196–203.
71. Toledo AV, Humber RA, Lastra CCL. First and southernmost records of *Hirsutella* (Ascomycota: Hypocreales) and *Pandora* (Zygomycota: Entomophthorales) species infecting dermaptera and psocodea. *J Invertebr Pathol* 2008;97:193–196.
72. Hajek AE, Gryganskyi AP, Gouli SY, Bittner TD, Sullivan CF, et al. Reassessing the diversity of the arthropod-pathogenic genus *Pandora* batko (Entomophthoromycotina; Erynioideae). *Diversity* 2024;16:603.
73. Gryganskyi AP, Humber RA, Smith ME, Miadlikowska J, Wu S, et al. Molecular phylogeny of the Entomophthoromycota. *Mol Phylogenet Evol* 2012;65:682–694.
74. Gryganskyi AP, Humber RA, Smith ME, Hodge K, Huang B, et al. Phylogenetic lineages in *Entomophthoromycota*. *Pers Int Mycol J* 2013;94–105.
75. Manfrino RG, Castrillo LA, López Lastra CC, Toledo AV, Ferrari W, et al. Morphological and Molecular Characterization of Entomophthorales (Entomophthoromycota: Entomophthoromycotina) from Argentina; 2020. <https://pbsociety.org.pl/journals/index.php/am/article/view/am.5522> [accessed 29 May 2024].
76. Tay WT, Court LN, Macfadyen S, Jacomb F, Vyskočilová S, et al. A high-throughput amplicon sequencing approach for population-wide species diversity and composition survey. *Mol Ecol Resour* 2022;22:1706–1724.
77. Zhu F, Shen Z, Xu X, Tao H, Dong S, et al. Phylogenetic analysis of complete rRNA gene sequence of *Nosema philosamia* isolated from the lepidopteran *Philosamia cynthia* ricini. *J Eukaryot Microbiol* 2010;57:294–296.
78. Zhu X, Wittner M, Tanowitz HB, Kotler D, Cali A, et al. Small subunit rRNA sequence of *Enterocytozoon bienersi* and its potential diagnostic role with use of the polymerase chain reaction. *J Infect Dis* 1993;168:1570–1575.
79. Trzebný A, Słodkiewicz-Kowalska A, Becnel JJ, Sanscrainte N, Dabert M. A new method of metabarcoding microsporidia and their hosts reveals high levels of microsporidian infections in mosquitoes (culicidae). *Mol Ecol Resour* 2020;20:1486–1504.
80. Vossbrinck CR, Maddox JV, Friedman S, Debrunner-Vossbrinck BA, Woese CR. Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. *Nature* 1987;326:411–414.
81. Williams BAP, Hamilton KM, Jones MD, Bass D. Group-specific environmental sequencing reveals high levels of ecological heterogeneity across the microsporidian radiation. *Environ Microbiol Rep* 2018;10:328–336.
82. Schrével J, Valigurová A, Prensier G, Chambouvet A, Florent I, et al. Ultrastructure of *Selenidium pendula*, the type species of archigregarines, and phylogenetic relations to other marine apicomplexa. *Protist* 2016;167:339–368.
83. Nocciolini C, Cucini C, Leo C, Francardi V, Dreassi E, et al. Assessing the efficiency of molecular markers for the species identification of gregarines isolated from the mealworm and super worm midgut. *Microorganisms* 2018;6:119.
84. Coupe S, Sarfati C, Hamane S, Derouin F. Detection of *Cryptosporidium* and identification to the species level by nested PCR and restriction fragment length polymorphism. *J Clin Microbiol* 2005;43:1017–1023.
85. Mancera N, Douma LG, James S, Liu S, Van A, et al. Detection of *Helicospiridium* spp. in metagenomic DNA. *J Invertebr Pathol* 2012;111:13–19.
86. Cannon MV, Bogale H, Rutt L, Humphrys M, Korpe P, et al. A high-throughput sequencing assay to comprehensively detect and characterize unicellular eukaryotes and helminths from biological and environmental samples. *Microbiome* 2018;6:195.
87. Lord JC, Hartzer KL, Kambhampati S. A nuptially transmitted ichthyosporean symbiont of *Tenebrio molitor* (Coleoptera: Tenebrionidae). *J Eukaryot Microbiol* 2012;59:246–250.
88. von der Heyden S, Cavalier-Smith T. Culturing and environmental DNA sequencing uncover hidden kinetoplastid biodiversity and a major marine clade within ancestrally freshwater *Neobodo* designis. *Int J Syst Evol Microbiol* 2005;55:2605–2621.
89. Bass D, Campo J. Microeukaryotes in animal and plant microbiomes: ecologies of disease?. *Eur J Protistol* 2020;76:125719.

The Microbiology Society is a membership charity and not-for-profit publisher.

Your submissions to our titles support the community – ensuring that we continue to provide events, grants and professional development for microbiologists at all career stages.

Find out more and submit your article at microbiologyresearch.org