

# Faecal metagenomes of great tits and blue tits provide insights into host, diet, pathogens and microbial biodiversity

Mark J. Pallen<sup>1,2,\*†</sup>, Alise Jany Ponsero<sup>1,†</sup>, Andrea Telatin<sup>1</sup>, Cara-Jane Moss<sup>1</sup>, David Baker<sup>1</sup>, Darren Heavens<sup>3</sup> and Gabrielle L. Davidson<sup>2,4,\*</sup>

## Abstract

**Background.** The vertebrate gut microbiome plays crucial roles in host health and disease. However, there is limited information on the microbiomes of wild birds, most of which is restricted to barcode sequences. We therefore explored the use of shotgun metagenomics on the faecal microbiomes of two wild bird species widely used as model organisms in ecological studies: the great tit (*Parus major*) and the Eurasian blue tit (*Cyanistes caeruleus*).

**Results.** Short-read sequencing of five faecal samples generated a metagenomic dataset, revealing substantial variation in composition between samples. Reference-based profiling with Kraken2 identified key differences in the ratios of reads assigned to host, diet and microbes. Some samples showed high abundance of potential pathogens, including siadenoviruses, coccidian parasites and the antimicrobial-resistant bacterial species *Serratia fonticola*. From metagenome assemblies, we obtained complete mitochondrial genomes from the host species and from *Isospora* spp., while metagenome-assembled genomes documented new prokaryotic species.

**Conclusions.** Here, we have shown the utility of shotgun metagenomics in uncovering microbial diversity beyond what is possible with 16S rRNA gene sequencing. These findings provide a foundation for future hypothesis testing and microbiome manipulation to improve fitness in wild bird populations. The study also highlights the potential role of wild birds in the dissemination of antimicrobial resistance.

## Impact Statement

The community of micro-organisms that lives in an animal's gut is crucial for the host's health. However, in wild birds, such as the great tit and Eurasian blue tit, the role of this community is poorly understood. Here, we have applied advanced DNA sequencing and computational techniques to faecal samples from these birds, revealing a rich diversity of microbes, including previously unknown species and potential pathogens. We discovered antibiotic resistance genes, raising concerns about the potential spread of resistance from wild birds to humans. This research highlights the importance of understanding gut microbes in wild birds and lays the groundwork for future efforts to manipulate their gut microbial communities to enhance their health and resilience in changing environments.

*Access Microbiology* is an open research platform. Pre-prints, peer review reports, and editorial decisions can be found with the online version of this article. Received 09 September 2024; Accepted 14 April 2025; Published 28 April 2025

**Author affiliations:** <sup>1</sup>Quadram Institute Bioscience, Norwich Research Park, Norwich, UK; <sup>2</sup>University of East Anglia, Norwich Research Park, Norwich, UK; <sup>3</sup>Earlham Institute, Norwich Research Park, Norwich, Norfolk, NR4 7UZ UK; <sup>4</sup>University of Cambridge, Downing Street, Cambridge, CB2 3EB, UK.

**\*Correspondence:** Mark J. Pallen, mark.pallen@quadram.ac.uk; Gabrielle L. Davidson, gabrielle.davidson@uea.ac.uk

**Keywords:** antimicrobial resistance; *Cyanistes caeruleus*; FONA beta-lactamase; gut microbiome; *Isospora*; metagenomics; *Parus major*; *Serratia fonticola*; siadenovirus.

**Abbreviations:** ANI, average NT identity; ARG, antibiotic resistance genes; GTDB, Genome Taxonomy Database; MAG, metagenome-assembled genome; vOTU, viral operational taxonomic unit.

†These authors contributed equally to this work

000910.v3 © 2025 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 SUMMARY

The authors confirm that all supporting data, code and protocols have been provided within the article or through supplementary data files. Metagenome sequences are available in the International Nucleotide Sequence Database Collaboration under BioProject ID PRJEB79635, with accession numbers ERX13007473, ERX13007474, ERX13007475, ERX13007476 and ERX13007477. We have made the MAG catalogue available in Zenodo archive (<https://doi.org/10.5281/zenodo.13643296>), the viral sequence catalogue available at <https://doi.org/10.5281/zenodo.13643307> and the mitochondrial sequences available at <https://doi.org/10.5281/zenodo.13734416>. Supplementary tables are available from <https://doi.org/10.6084/m9.figshare.26968957> [1].

## INTRODUCTION

The vertebrate gut microbiome plays crucial roles in animal health and host fitness, including nutrient recovery, protection from infection and priming normal development of the gut and immune system [2, 3]. Additionally, this key microbial community can act as a reservoir of pathogens and antimicrobial resistance [4, 5], as well as a source of biotechnological resources [6]. As a result, the gut microbiome has been a major focus of intensive research in humans and domesticated animals [7–9]. Variation in the gut microbiome is also likely to affect host health fitness in wild animals, including birds, although few experimental studies have tested this hypothesis [10, 11].

Wild birds face numerous threats associated with the Anthropocene, including climate change, habitat loss and reduction in genetic diversity [12]. Endangered populations, whether in the wild or in captivity, are also likely to experience a loss of microbial diversity in the gut, which can negatively impact fitness [13, 14]. In addition, variation in parasite load, interacting with the gut microbiome, is likely to affect survival [15–18]. In these circumstances, microbiome manipulation offers a potential path to enhance fitness for individual birds and for endangered populations [19]. Understanding how the gut microbiome impacts wildlife hosts is an important goal for conservation and evolutionary biology.

The great tit (*Parus major*) and the Eurasian blue tit (*Cyanistes caeruleus*) are small passerines from the family *Paridae*, widely used as model species in ecological studies [20–22]. These species are experimentally tractable because they are altricial (growing quickly, while remaining in their nests until near independence) and so are easily monitored in nest boxes. They are also widespread and common species across Eurasia, where they have adapted to human-modified habitats, dominating garden bird feeders. They also represent potentially important vectors of zoonotic pathogens, such as *Campylobacter* and *Borrelia* [23, 24]. Although not endangered, these species have experienced unusual mortality incidents in recent decades in the UK and Germany, linked to bacterial and viral pathogens [25–27]. Consequently, they provide an attractive test case for investigating microbial biodiversity within the gut microbiomes of wild birds.

Previous studies of parid gut microbiomes have documented a rich diversity of micro-organisms and variation in taxonomic composition based on age and diet [28–33]. However, these efforts have largely relied on metabarcoding with incomplete 16S rRNA gene sequences. Unfortunately, this approach fails to provide species-level resolution, cannot detect viruses or eukaryotes and provides no insights into the genome sequences, population structures or functional repertoires of resident microbial species [34]. Thus, despite previous efforts, the great tit and blue tit gut microbiomes present us with a largely unexplored landscape of taxonomic, ecological and functional diversity, with potentially important, undiscovered roles in fitness, while also encompassing known and newly discovered parasites.

The application of shotgun metagenomics to gut microbiomes has proven capable of linking host, diet and microbes, while also documenting microbial diversity through reference-based phylogenetic profiling and recovery of genome sequences through assembly and binning [9, 35–38]. Here, we apply shotgun metagenomics to five parid faecal samples, providing proof of principle for sampling, sequencing and bioinformatic methods, while also laying the foundations for hypothesis testing and microbiome manipulation in wild birds.

## METHODS

### Sample collection and analysis

This study was conducted with ethical approval from the University of East Anglia Animal Welfare and Ethical Review Body (AWERB) (Application ID: ETH2223-2022) and the University of Cambridge AWERB (Application ID: Z0067/19 and Z0068/19). Faecal samples (Table S1, available in the online Supplementary Material) were collected from

- nest boxes in Madingley Wood, Cambridge, UK, from one full-grown great tit in October 2019 and from five great tit nestlings in May–June 2020. Nests were monitored for breeding progress, and samples were collected in sterile bags [39] from the adult and from nestlings when chicks were 15 days old, following our established protocol [28].
- six full-grown blue tits, using mist nets at the University of East Anglia, Norwich, UK, in November 2023.

Samples from great tits were stored in 100% ethanol, while those from blue tits were stored in 1:1 DNA/RNA Shield Solution (Zymo Research). Samples were stored at  $-80^{\circ}\text{C}$  within a few hours of collection and until DNA extraction. DNA was extracted using the QIAamp PowerFecal Pro DNA Kit (Qiagen; 51804) following the supplier's protocol, with modifications to increase DNA yield.

During DNA extractions on the great tit samples performed in 2021, the bead tube was vortexed for 10 min and samples were heated in a heat block at  $65^{\circ}\text{C}$  for 15 min and this step was repeated three times. The CD2 solution step was repeated to remove any potential inhibitors, and samples were washed with C5 twice. DNA was eluted in  $50\ \mu\text{l}$  10 mM Tris-HCl and stored at  $-80^{\circ}\text{C}$  until further analysis.

During DNA extractions on the blue tit samples performed in 2023, samples were homogenized with zirconium beads in CD1 solution and beaten in the portable homogenizer SuperFast Prep-2 (MP Biomedicals, USA) for 25 s at speed code 20. The supernatant was recovered and subjected to  $1\times$  bead clean-up using Kapa pure beads (Roche Catalogue No. 07983298001), followed by two 70% ethanol washes, and DNA was eluted in  $20\ \mu\text{l}$  10 mM Tris-HCl. Extracted DNA was stored at  $-80^{\circ}\text{C}$  before further analysis.

## Metagenomic sequencing

DNA quantification was performed using a Qubit™ 3.0 fluorometer (Invitrogen, CA) and dsDNA HS assay kit. Samples were concentrated through an AMPure bead clean-up, using a  $1.5\times$  bead volume. DNA was resuspended in  $6\ \mu\text{l}$  PCR grade water, the full volume of which was used in the tagmentation reaction of library preparation.

Illumina sequencing libraries were constructed using the Nextera XT library preparation kit according to the manufacturer's recommendations.

The quality of the libraries was assessed using the Agilent 2200 TapeStation system. Seven samples that failed to yield DNA libraries of sufficient quality for sequencing were excluded from further analysis, leaving (Table 1):

- two samples from the great tit *P. major* (TV34554 and TV35040) from 15-day-old fledglings
- three samples from the blue tit, *Cyanistes caeruleus* (one adult ABR6151 and two juveniles, ABR6116 and ABR6154).

Successful libraries were equimolar pooled, and the final pool was double size selected by solid phase reversible immobilisation between  $0.5\times$  and  $0.7\times$  bead volumes using AMPure beads. The final pool was quantified using a Qubit™ 3.0 fluorometer (Invitrogen) and run on the Agilent 2200 TapeStation system to calculate the molarity.

Paired-end metagenomic sequencing was performed on the Illumina NextSeq 2000 platform at the Quadram Institute Norwich, yielding  $2\times 150$  bp paired-end sequencing reads.

## Metagenomic analysis

Metagenomic reads were trimmed and quality controlled using FastP configured to a minimum phred score of 20 and minimum length of 50 bp [40]. Taxonomic profiling of sequencing reads was performed using Kraken2 v2.1.0 [41], with a microbial database built from archaeal, bacterial, fungal, protozoan, viral and univec\_core sequences in RefSeq in 2024.

Individual assemblies were performed on each metagenomic sample using MEGAHIT v1.2.9 [42]. The BLAST 2.16 suite of programmes [43] was downloaded from <ftp.ncbi.nlm.nih.gov/blast> and used to perform homology searches of the assemblies, using the *makeblastdb* utility to generate libraries, the *blastn* and *tblastx* utilities to search contigs with query sequences under high stringency ( $e$  value  $\geq 1\times 10^{-200}$ ) and the *blastdbcmd* utility to retrieve hits from databases.

Antibiotic and metal resistance genes were identified from the assembled contigs using Abricate v1.0.1 (<https://github.com/tseemann/abricate>) against the PanRes database v1.0.1 (available at <https://zenodo.org/records/8055116>). The potential host of these ARGs (antibiotic resistance genes) was inferred by taxonomic classification of the contigs carrying ARGs using Kraken2 against the RefSeq database.

**Table 1.** Subject and sample characteristics

Bird ID	Collection date	Site code	Site	Coordinate	Species	Age	Sex	Capture method
ABR6116	11 November 2023	UEARE	UEA Norwich	52.616424, 1.234345	Blue tit	Juvenile	Female	Mist net
ABR6151	11 November 2023	UEARE	UEA Norwich	52.616424, 1.234345	Blue tit	Adult	Male	Mist net
ABR6154	11 November 2023	UEARE	UEA Norwich	52.616424, 1.234345	Blue tit	Juvenile	Unknown	Mist net
TV34554	16 May 2020	MW73	Madingley Cambridge	52.21788, 0.04674	Great tit	15 days	Unknown	Nest box
TV35040	25 May 2020	MW89	Madingley Cambridge	52.2171, 0.04985	Great tit	15 days	Unknown	Nest box

To avoid contamination of the bins by eukaryotic sequences, Tiara v1.0.3 [44] was used to classify contigs longer than 3,000 kb into their high-level kingdoms, allowing to exclude sequences of a eukaryotic or of an organelle origin and only retaining all unclassified contigs and prokaryotic contigs for the binning step.

Contigs were binned using MaxBin2 v2.2.7, SemiBin2 v2.1.0 and Metabat2 v2.15 independently [45–47]. The bins were refined using DasTool v1.1.7 [48] using a min score threshold of 0.3. The quality of the refined bins was obtained using CheckM2 v1.0.2 [49], and any bins with contamination above 10% were excluded. The final metagenome-assembled genomes (MAGs) were classified as low quality (<50% completeness and <10% contamination), medium quality (>50% completeness and <10% contamination) and high quality (>90% completeness and <5% contamination), as recommended by the MIMAG specification [50]. Finally, the MAGs were dereplicated using dRep v3.4.3 [51] with an average NT identity (ANI) of 95% and classified using gtdb-tk v2.4.0 on the Genome Taxonomy Database (GTDB) Release 220 [52–54].

DRAM v1.5 [55] was run against the DRAM database downloaded on July 2024 on the MAGs. Briefly, DRAM first annotates the genes of each MAG using the UniRef90 [56], Pfam [57] and dbCan [58] databases. Then, a secondary step (distillation) is run to curate these annotations into useful functional categories, allowing to infer the main metabolic functions detected in the MAG.

The potential viral and plasmid sequences were retrieved from the assembled contigs using GeNomad v1.8.0 [59], and the quality of these viral sequences was then assessed using CheckV v1.0.3 [60] using the CheckV database v1.5. Putative viral sequences longer than 1,000 bp carrying at least one hallmark viral gene or no detectable cellular genes were considered in this analysis. The viral sequences were clustered into species-level viral operational taxonomic unit (vOTU) clusters using mmseqs2 v2.13 [61] using an identity threshold of 95% over 75% of the longest sequence. For viral species identified as belonging to the Caudovirales, the potential host was inferred using iPhop v1.3.3 [62].

To estimate the relative abundance of the vOTUs in the samples, Bowtie2 v2.5.4 [63] was used to map the short reads after quality control to the vOTU representative sequences, generating a count table that represents the number of reads mapped to each contig in the original sample. Finally, bbmap v39.01 (<https://jgi.doe.gov/data-and-tools/software-tools/bbtools/>) was used to normalize the read counts in read counts per kb million reads, corresponding to the read counts normalized by the vOTU reference sequence length in kb and by the total of reads mapped to the viral catalogue.

## RESULTS AND DISCUSSION

### Reference-based profiling provides insights into host diet, gut microbiota and pathogens

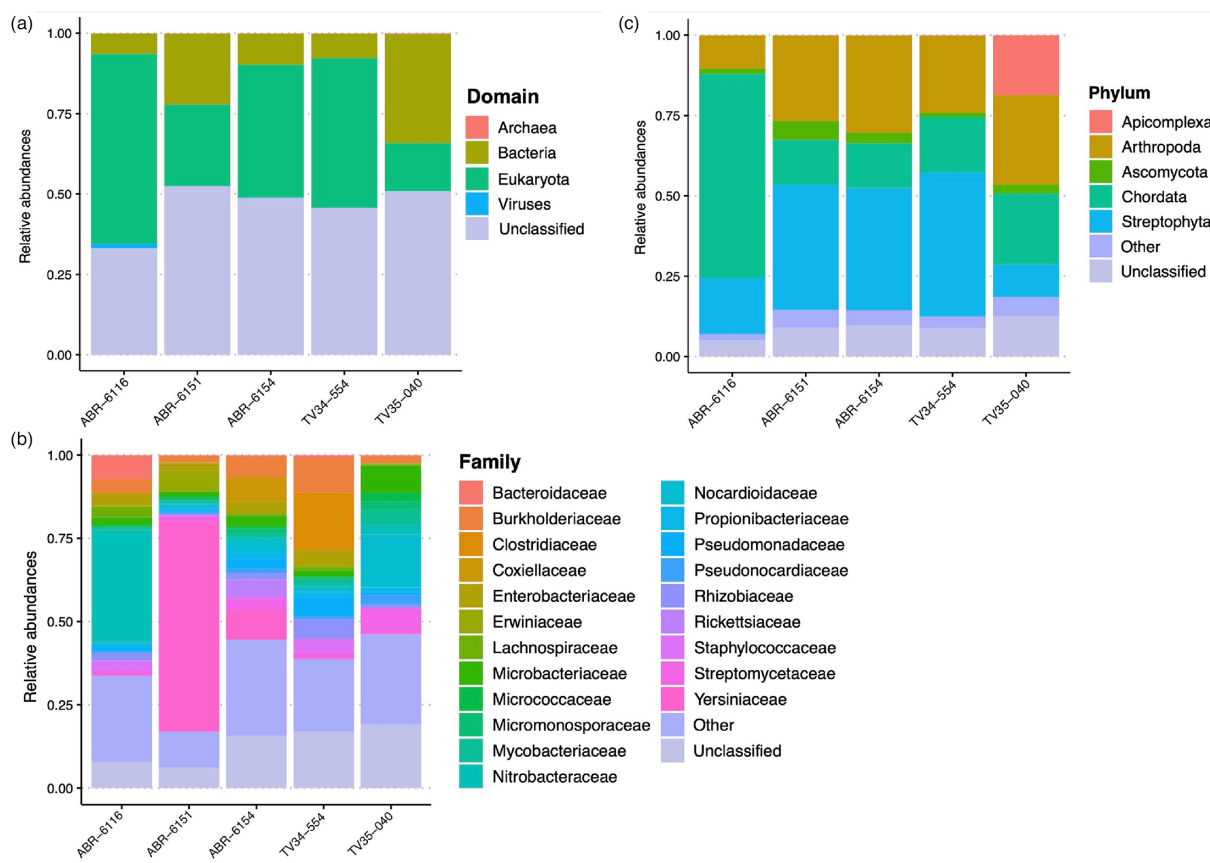
Short-read sequencing of faecal samples generated a metagenomic dataset of 258 million paired-end reads, with an average of 52 million paired reads per sample and >98% of reads from each sample surviving trimming (Table S1, available from <https://doi.org/10.6084/m9.figshare.26968957>).

Taxonomic profiling of the reads using Kraken2 [41] revealed that 45–52% of the sequencing reads remained unclassified, confirming the largely unknown nature of the great tit and blue tit faecal microbiomes (Fig. 1 and Table S2). These analyses revealed substantial differences between samples in terms of ratios of reads assigned to host, food, gut microbiome and potential pathogens. Such variation – which we know from experience is not uncommon when analysing such microbiomes [5, 9, 38, 64, 65] – is likely to reflect differences in the shedding of host DNA from the gut, release of DNA from food components and recent dietary intake.

In ABR-6116, over a third of the reads were assigned to *Aves* (i.e. to the host), while in the other samples, this was <5%. The relative abundance of reads assigned to phylum *Arthropoda* (home to most potential food animals, including spiders and insect larvae) was consistent across our samples, ranging from 4.17 to 12.5% of reads (Fig. 1). Great tits and blue tit nestlings are highly dependent on lepidopteran larvae (caterpillars) for a high-quality diet during development, but will provision young with lower quality invertebrates in urban environments, where caterpillar abundance is low [66, 67]. As the nests in our study were in high-quality oak woodland habitat, within a rural area, caterpillar abundance is expected to be high.

The phylum *Streptophyta* (potential food plants) accounted for ≥10% of reads in most samples, but for only 1.5% in great tit TV34-554. Two plant species featured a high relative abundance in the Kraken assignments. The sunflower *Helianthus annuus* accounted for 11.4% and 7.6% of reads from blue tits ABR6151 and ABR6154, while the oak *Quercus robur* accounted for 3.2% of reads in the remaining blue tit ABR6116 and for 3.9% in great tit TV35040. Although these results are generally consistent with a similar previous analysis using DNA metabarcoding [68], the presence of oak reads is best explained by the consumption of caterpillars that eat oak (e.g. oak tortrix moth caterpillars) rather than direct consumption of oak, while high abundance of sunflower reads suggests provisioning of garden bird seed by adults to their young. These findings highlight the wide-ranging human-influenced habitat of these wild birds.

Bacterial reads were predominantly assigned to the four phyla in the NCBI taxonomy most associated with animal gut microbiomes – *Proteobacteria*, *Firmicutes*, *Bacteroidetes* and *Actinobacteria* (Fig. 1), but also suggested the presence of potential pathogens. In one juvenile blue tit, ABR6154, a single bacterial species, *Serratia fonticola*, accounted for 8.8% of all sequence reads.



**Fig. 1.** Read-based profiling of the parid faecal microbiota. (a) Relative abundance of the global community composition aggregated at the kingdom level. The read classification counts from Kraken2 were normalized using Total sum scaling normalization, and the counts were aggregated at the kingdom level. Counts classified as ‘root’ by Kraken2 were considered unclassified. (b) Relative abundance of the bacterial fraction. The counts were sub-setted for the bacterial fraction, transformed using TSS and aggregated at the family level. Families with a relative abundance below 2% in the samples were grouped as ‘other’ in this plot. (c) Relative abundance of the eukaryotic fraction. The counts were sub-setted for the eukaryotic fraction, transformed using TSS and aggregated at the phylum level. Phyla with a relative abundance below 2% in the samples were grouped as ‘other’ in this plot.

Such monodominance of the gut microbiome by a single organism has been seen in other contexts, particularly in human neonates and adults who have received broad-spectrum antibiotics [5, 69–71]. However, its significance in this context remains uncertain: given that the avian host appeared healthy, it might represent a normal step in the ecological succession required to build a gut microbiome in this species, or it might be associated with an inapparent decrease in fitness.

In the other juvenile blue tit, ABR6116, just over 1% of reads were assigned to a single viral species ‘Great tit siadenovirus A’. Siadenoviruses have been detected in various avian species, including great tits and blue tits, where they have been associated with renal pathology and mortality [72, 73].

In the fledgling great tit TV34554, sequences assigned to the phylum *Apicomplexa* (obligate endoparasites of birds and other animals) accounted for 2.80% of reads. Most of these reads were assigned to the subclass *Coccidia* and to the family *Eimeriidae*. Importantly, it has been reported that the classification of morphologically defined genera among the eimeriid coccidia conflicts with sequence-based phylogenies [74].

These results, while providing a starting point for further investigations, may be confounded by a high prevalence of false positive detection. Indeed, read-based profiling is known to provide assignments at the level of genus and species that can be untrustworthy [75]. This is evident from our dataset in that many reads are quite implausibly assigned to bird species other than the great tit and blue tit. For example, in blue tit ABR6116, 11.4% of reads were assigned to the common reed warbler *Acrocephalus scirpaceus* and 7.2% to the European robin *Erithacus rubecula*. Taxonomic profiling also fails to provide whole-genome data or insights into functional diversity or population structure of the species they report. In addition, as this approach relies on a reference

database, it can only report previously known organisms, which can limit the exploration of lesser-studied ecosystems. Given these limitations, we next explored the microbial communities associated with these birds through a sequence assembly approach.

### Recovery of mitochondrial genomes from faecal microbiome allows the genetic confirmation of the host species

Assemblies from the five faecal samples generated 504,238 contigs longer than 1,000 bp (Table S1). To confirm the identity of the host species, the great tit mitochondrial reference genome (GenBank ID: NC\_040875.1) [76] was used as a query in high-stringency BLASTN searches of the metagenomic assemblies. This allowed us to retrieve 16 kb mitochondrial genomes from the great tit samples (available from <https://doi.org/10.5281/zenodo.13734416>) that both showed >99.9% sequence identity with the closest of the five available great tit mitochondrial genomes (GenBank ID: MN356395.1). We also identified three mitochondrial genomes from the blue tit samples, each of which showed 97% sequence identity to the only available complete *Cyanistes* mitochondrial genome, from the azure tit *Cyanistes cyanus* (GenBank ID: KX388472.1) [77], and >99.9% identity to *Cyanistes caeruleus* mitochondrial sequences from the control region and cytochrome b gene [78]. These sequences, which together share >99.9% NT identity, represent the first full mitochondrial genomes from *Cyanistes caeruleus* (available from <https://doi.org/10.5281/zenodo.13734416>).

Next, we sought to confirm the presence of apicomplexan parasites in the great tit samples. We used as a query in high-stringency TBLASTN searches of the metagenomic assemblies the cytochrome c oxidase subunit I protein sequence from *Isospora picoflavae* (GenBank ID: UQT67988). This species has been described from a woodpecker – the northern yellow-shafted flicker *Colaptes auratus luteus* – but sits within a clade containing *Isospora* species from various passerine hosts [79]. This search allowed us to retrieve a single complete 6 kb *Isospora* mitochondrial genome from each of the great tit samples (available from <https://doi.org/10.5281/zenodo.13734416>). The *Isospora* mitochondrial genome retrieved from the great tit TV34-554 showed 99.4% NT identity to the mitochondrial genome from *Isospora* sp. JRBarta-2021c from the American robin, *Turdus migratorius* (GenBank ID MW645338;), 97.1% NT identity to the mitochondrial genome of *I. picoflavae* (GenBank: MW599989.1) and 97.7% NT identity to a partial cytochrome c oxidase subunit I sequence from a great tit in the Czech Republic (GenBank: MK573833.1) [80].

On the other hand, the *Isospora* mitochondrial genome from great tit TV35-040 showed 96.9% NT identity to the mitochondrial genome from *Isospora* sp. JRBarta-2021c from *T. migratorius*, 97.6% NT identity to the *I. picoflavae* mitochondrial genome and 95.4% NT identity to the partial cytochrome c oxidase from the great tit from the Czech Republic. This sequence had a 96.0% nt identity to the *Isospora* mitochondrial genome from great tit TV34-554.

Additionally, a BLASTN search of contigs with the small subunit ribosomal RNA gene from *I. picoflavae* (GenBank: MW618926.1) retrieved two sequences from great tit TV34-554 and TV35-040 that showed a 99% NT identity across the entire small subunit ribosomal RNA gene from *I. picoflavae* sequence (File S4).

Classically, *Isospora* species have been described based on the morphology of their sporulated oocysts, pathology and host identity, and no guidelines are available for the description of new species from sequence data alone [74, 79]. Nonetheless, given the consistent results obtained with mitochondrial and rRNA sequences, it seems likely that the two great tits in our survey carry different species of *Isospora*, one shared with the American robin and the other apparently novel.

### MAGs document taxonomic and functional diversity

From the parid faecal samples, we generated 26 non-redundant bins from single-sample assemblies using three different approaches to binning (Table S3). A total of 18 bins represent medium- or high-quality MAGs. De-replication at 95% ANI clustered MAGs into 17 species clusters, spanning 7 phyla. Notably, four of these species have validly published Latin binomials, while one has been assigned alphanumerical placeholders by GTDB [81].

Among our MAG species clusters, 12 represent new candidate species within bacterial genera previously delineated by GTDB. Building on our recent efforts with the creation of arbitrary well-formed Latin names [82] and with the oesophageal microbiome [9], we have assigned *Candidatus* names (abbreviated as *Ca.*) to all the unnamed taxa revealed by our metagenomic analyses (Table 2).

The metabolic capacity of each MAG was characterized using DRAM, predicting the high-level metabolic characteristics of these uncharacterized genomes (Fig. 2). Notably, MAGs representing the novel species *Cedopifca cyanistis* carried carbohydrate-active enzymes involved in the cleavage of amorphous cellulose (GH5), suggesting a potential role of this species in the passerine digestive function.

Consistent with the read-based profiling results that showed the presence of *Serratia fonticola* at high relative abundance in the blue tit ABR-6154, we assembled a high-quality MAG for this species. *Serratia fonticola* is a recognized pathogen of humans associated with gastrointestinal, respiratory, urinary, bloodstream, soft tissue and joint infections [83] and has been isolated from a range of environmental sources, including wild and domesticated birds. However, it has not previously been seen in wild parids [84–86].

**Table 2.** Protologues for newly named species. Protologues for new *species* identified by the analysis of MAGs from the faecal microbiomes of the great tit and blue tit

Species	Etymology	Description
<i>Ca. Uzinuria cyanistis</i> sp. nov.	<i>cyanistis</i> . N.L. gen. masc n. cyanistis, of <i>Cyanistes</i> , genus name of the host bird	A bacterial species identified by metagenomic analysis of a faecal sample from the Eurasian blue tit <i>Cyanistes caeruleus</i> and assigned to this genus according to the algorithms of the GTDB Toolkit operating on GTDB Release 220; this species includes all bacteria with genomes that show $\geq 95\%$ ANI to the type genome for the species to which we have assigned the MAG ID ABR-6151_bin.14_sub and which is available via Zenodo archive 10.5281/zenodo.13643296. The G+C content in mol% of the type genome is 30%, and the genome length is 211,505 bp
<i>Ca. Cedopifica cyanistis</i> sp. nov.	<i>cyanistis</i> . N.L. gen. masc n. cyanistis, of <i>Cyanistes</i> , genus name of the host bird	A bacterial species identified by metagenomic analysis of a faecal sample from the Eurasian blue tit <i>Cyanistes caeruleus</i> and assigned to this genus according to the algorithms of the GTDB Toolkit operating on GTDB Release 220; this species includes all bacteria with genomes that show $\geq 95\%$ ANI to the type genome for the species to which we have assigned the MAG ID ABR-6151_SemiBin_15 and which is available via Zenodo archive 10.5281/zenodo.13643296. The G+C content in mol% of the type genome is 42%, and the genome length is 1,442,697 bp
<i>Ca. Cedopifica</i> gen. nov.	<i>Cedopifica</i> . N.L. fem. n. Cedopifica, a name created arbitrarily, while preserving the phonotactics and morphology of Latin	A bacterial genus identified and delineated according to the algorithms of the GTDB Release 09-RS220; the GTDB alphanumeric placeholder designation for this genus is JAHHUI01; the type species for the genus is <i>Ca. Cedopifica cyanistis</i> ; according to GTDB Release 09-RS220, this genus belongs to the family WRBN01
<i>Ca. Cimetritia cyanistis</i> sp. nov.	<i>cyanistis</i> . N.L. gen. masc n. cyanistis, of <i>Cyanistes</i> , genus name of the host bird	A bacterial species identified by metagenomic analysis of a faecal sample from the Eurasian blue tit <i>Cyanistes caeruleus</i> and assigned to this genus according to the algorithms of the GTDB Toolkit operating on GTDB Release 220; the GTDB alphanumeric placeholder designation for this genus is GCF-002259525; the arbitrary name <i>Ca. Cimetritia</i> has been published elsewhere [84]; this species includes all bacteria with genomes that show $\geq 95\%$ ANI to the type genome for the species to which we have assigned the MAG ID ABR-6151_SemiBin_32 and which is available via Zenodo archive 10.5281/zenodo.13643296; the G+C content in mol% of the type genome is 33%, and the genome length is 1,507,853 bp
<i>Ca. Udofia cyanistis</i> sp. nov.	<i>cyanistis</i> . N.L. gen. masc n. cyanistis, of <i>Cyanistes</i> , genus name of the host bird	A bacterial species identified by metagenomic analysis of a faecal sample from the Eurasian blue tit <i>Cyanistes caeruleus</i> and assigned to this genus according to the algorithms of the GTDB Toolkit operating on GTDB Release 220; this species includes all bacteria with genomes that show $\geq 95\%$ ANI to the type genome for the species to which we have assigned the MAG ID ABR-6151_SemiBin_38_sub and which is available via Zenodo archive 10.5281/zenodo.13643296; the G+C content in mol% of the type genome is 35%, and the genome length is 648,929 bp
<i>Ca. Udofia</i> gen. nov.	<i>Udofia</i> , N.L. fem. n. Udofia, a name created arbitrarily, while preserving the phonotactics and morphology of Latin	A bacterial genus identified and delineated according to the algorithms of the GTDB Release 09-RS220; the GTDB alphanumeric placeholder designation for this genus is CALTSX01; the type species for the genus is <i>Ca. Udofia cyanistis</i> ; according to GTDB Release 09-RS220, this genus belongs to the family f__ CALTSX01
<i>Ca. Pugiria cyanistis</i> sp. nov.	<i>cyanistis</i> . N.L. gen. masc n. cyanistis, of <i>Cyanistes</i> , genus name of the host bird	A bacterial species identified by metagenomic analysis of a faecal sample from the Eurasian blue tit <i>Cyanistes caeruleus</i> and assigned to this genus according to the algorithms of the GTDB Toolkit operating on GTDB Release 220; this species includes all bacteria with genomes that show $\geq 95\%$ ANI to the type genome for the species to which we have assigned the MAG ID ABR-6154_bin.9 and which is available via Zenodo archive 10.5281/zenodo.13643296; the G+C content in mol% of the type genome is 38%, and the genome length is 1,811,547 bp
<i>Ca. Pugiria</i> gen. nov.	<i>Pugiria</i> , N.L. fem. n. Pugiria, a name created arbitrarily, while preserving the phonotactics and morphology of Latin	A bacterial genus identified and delineated according to the algorithms of the GTDB Release 09-RS220; the GTDB alphanumeric placeholder designation for this genus is JAVHYG01; the type species for the genus is <i>Ca. Pugiria cyanistis</i> ; according to GTDB Release 09-RS220, this genus belongs to the family <i>Diplorickettsiaceae</i>
<i>Ca. Paridicola</i> gen. nov.	<i>Paridicola</i> . N. L. gen. fem. n. parida, a bird from the family <i>Paridae</i> ; L. masc./fem. n. suff. -cola, inhabitant of; N.L. masc./fem. n. paridicola, inhabitant of parids	A bacterial genus identified by metagenomic analyses; the genus includes all bacteria with genomes that show $\geq 60\%$ average AA identity (AAI) to the genome of the type strain from the type species <i>Ca. Paridicola cyanistis</i> This genus has been assigned by GTDB-Tk v1.3.0 working on GTDB Release 220 to the order <i>Legionellales</i> and to the family <i>Legionellaceae</i>

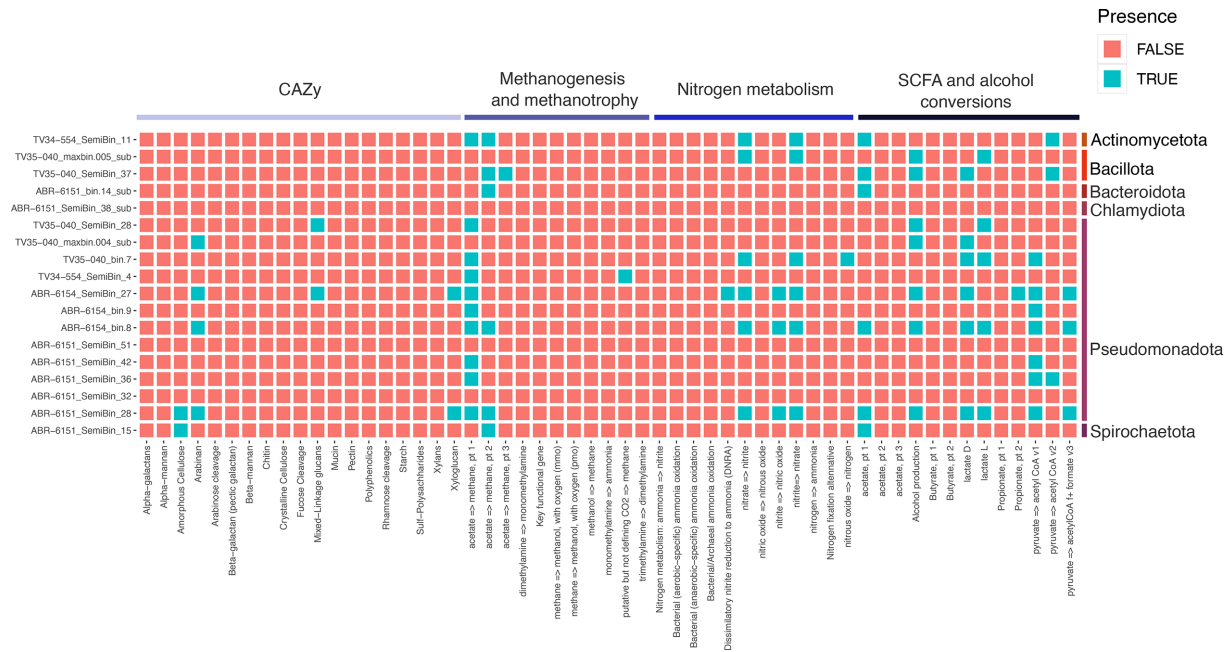
Continued

Table 2. Continued

Species	Etymology	Description
<i>Ca. Paridicola cyanistis</i> sp. nov.	<i>cyanistis</i> . <i>N.L. gen. masc n.</i> <i>cyanistis</i> , of <i>Cyanistes</i> , genus name of the host bird	A bacterial species identified by metagenomic analysis of a faecal sample from the Eurasian blue tit <i>Cyanistes caeruleus</i> and assigned to this genus according to the algorithms of the GTDB Toolkit operating on GTDB Release 220; this species includes all bacteria with genomes that show $\geq 95\%$ ANI to the type genome for the species to which we have assigned the MAG ID ABR-6151_SemiBin_36 and which is available via Zenodo archive 10.5281/zenodo.13643296; the G+C content in mol% of the type genome is 40%, and the genome length is 2,370,451 bp
<i>Ca. Rosutilega pari</i> sp. nov.	<i>pari</i> . <i>L. gen. masc. n. pari</i> , of <i>Parus</i> , genus name of the host bird	A bacterial species identified by metagenomic analysis of a faecal sample from the great tit <i>P. major</i> and assigned to this genus according to the algorithms of the GTDB Toolkit operating on GTDB Release 220; this genus has been assigned by GTDB the alphanumeric designation JAVHYG01; the arbitrary name has been published elsewhere [84]; this species includes all bacteria with genomes that show $\geq 95\%$ ANI to the type genome for the species to which we have assigned the MAG ID TV34-554_SemiBin_11 and which is available via Zenodo archive 10.5281/zenodo.13643296; the G+C content in mol% of the type genome is 68%, and the genome length is 1,776,244 bp
<i>Ca. Methyloceanibacter pari</i> sp. nov.	<i>pari</i> . <i>L. gen. masc. n. pari</i> , of <i>Parus</i> , genus name of the host bird	A bacterial species identified by metagenomic analysis of a faecal sample from the great tit <i>Parus major</i> and assigned to this genus according to the algorithms of the GTDB Toolkit operating on GTDB Release 220; this species includes all bacteria with genomes that show $\geq 95\%$ ANI to the type genome for the species to which we have assigned the MAG ID TV34-554_SemiBin_4 and which is available via Zenodo archive 10.5281/zenodo.13643296; the G+C content in mol% of the type genome is 63%, and the genome length is 1,837,002 bp
<i>Ca. Scandinavium pari</i> sp. nov.	<i>pari</i> . <i>L. gen. masc. n. pari</i> , of <i>Parus</i> , genus name of the host bird	A bacterial species identified by metagenomic analysis of a faecal sample from the great tit <i>Parus major</i> and assigned to this genus according to the algorithms of the GTDB Toolkit operating on GTDB Release 220; this species includes all bacteria with genomes that show $\geq 95\%$ ANI to the type genome for the species to which we have assigned the MAG ID TV35-040_maxbin.004_sub and which is available via Zenodo archive 10.5281/zenodo.13643296; the G+C content in mol% of the type genome is 50%, and the genome length is 2,647,412 bp
<i>Ca. Mammaliococcus pari</i> sp. nov.	<i>pari</i> . <i>L. gen. masc. n. pari</i> , of <i>Parus</i> , genus name of the host bird	A bacterial species identified by metagenomic analysis of a faecal sample from the great tit <i>Parus major</i> and assigned to this genus according to the algorithms of the GTDB Toolkit operating on GTDB Release 220; this species includes all bacteria with genomes that show $\geq 95\%$ ANI to the type genome for the species to which we have assigned the MAG ID TV35-040_maxbin.005_sub and which is available via Zenodo archive 10.5281/zenodo.13643296; the G+C content in mol% of the type genome is 34%, and the genome length is 8,553,045 bp
<i>Ca. Agrobacterium pari</i> sp. nov.	<i>pari</i> . <i>L. gen. masc. n. pari</i> , of <i>Parus</i> , genus name of the host bird	A bacterial species identified by metagenomic analysis of a faecal sample from the great tit <i>Parus major</i> and assigned to this genus according to the algorithms of the GTDB Toolkit operating on GTDB Release 220; this species includes all bacteria with genomes that show $\geq 95\%$ ANI to the type genome for the species to which we have assigned the MAG ID TV35-040_SemiBin_28 and which is available via Zenodo archive 10.5281/zenodo.13643296; the G+C content in mol% of the type genome is 59%, and the genome length is 4,361,088 bp
<i>Ca. Dwaynesavagella pari</i> sp. nov.	<i>pari</i> . <i>L. gen. masc. n. pari</i> , of <i>Parus</i> , genus name of the host bird	A bacterial species identified by metagenomic analysis of a faecal sample from the great tit <i>Parus major</i> and assigned to this genus according to the algorithms of the GTDB Toolkit operating on GTDB Release 220; this species includes all bacteria with genomes that show $\geq 95\%$ ANI to the type genome for the species to which we have assigned the MAG ID TV35-040_SemiBin_37 and which is available via Zenodo archive 10.5281/zenodo.13643296; the G+C content in mol% of the type genome is 30%, and the genome length is 1,518,404 bp

### Antibiotic resistance genes in the parid gut

Next, we surveyed the repertoire of antibiotic and metal resistance genes found in the passerine faecal microbiota. Interestingly, we identified a large repertoire of antibiotic resistance genes, covering genes potentially involved in the resistance to beta-lactams, tetracycline, fluoroquinolones and polymyxins. Notably, 43 (17%) of these resistance genes were detected on a putative plasmid and so are potentially mobile (File S4). Resistance genes were associated with a variety of bacterial hosts, in particular the genera *Rahnella*, *Pseudomonas* and *Serratia*.



**Fig. 2.** Predicted metabolic capability of the MAGs from the parid gut. Predicted coding sequences in each high- and medium-quality MAG were annotated using DRAM against the UniRef90, PFAM and dbCAN databases. The metabolic capacities of each MAG were inferred using the ‘distilled’ function from DRAM, and the presence/absence of metabolism of interest was plotted as a tile plot.

Resistance to beta-lactams in *Serratia fonticola* is mediated by chromosomal class A extended-spectrum beta-lactamases belonging to the FONA family, encoded by the *blaFONA* genes [85], and we confirmed the presence of this gene in the gut microbiome of blue tit ABR-6154, carried on a contig sequence classified as *Serratia fonticola*. Concerns have been raised that wild birds can harbour clinically significant resistance genes and could potentially play a role in the global dissemination of antibacterial resistance [87–91]. However, whether the presence of this antimicrobial-resistant pathogen in wild birds in the UK represents a risk to human health remains unclear without an in-depth analysis of the population structure of *Serratia fonticola*.

### A diverse community of phages and eukaryotic viruses

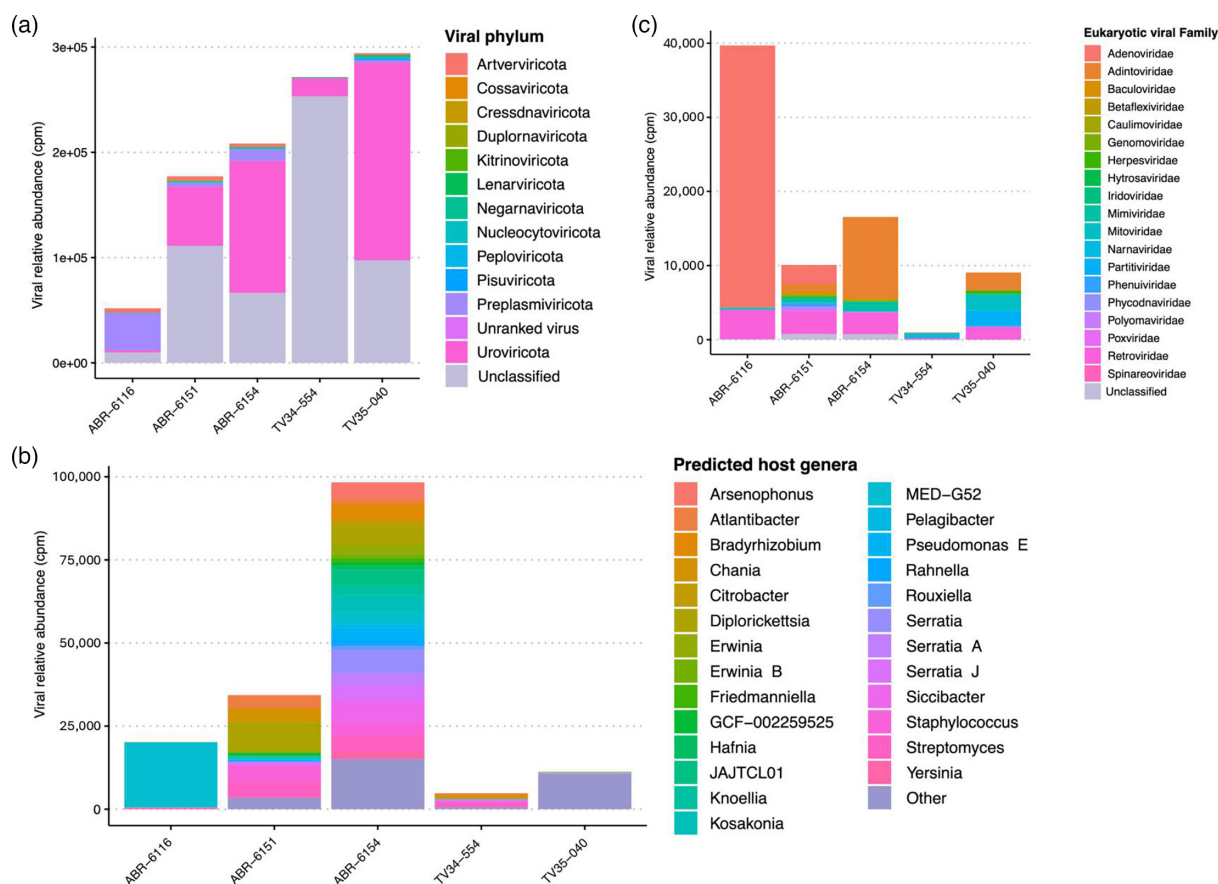
Finally, we explored viral diversity in these samples. A total of 1,870 viral sequences longer than 1 kb were identified, corresponding to 1,843 species-level vOTUs. We assessed the relative abundance of these viral species in the samples by read mapping against the representative genome of each vOTU, showing a very varied viral load and global composition between the samples (Fig. 3). The phage community was dominated by viruses from the phylum *Caudovirales* and accounted for 5–63% of the viral content of the samples. Putative bacterial hosts were predicted for 360 *Caudovirales* vOTUs, with a close correspondence between predicted bacterial hosts and the main bacterial taxa identified in the samples (Fig. 3).

We identified 173 sequences classified as eukaryotic viruses, from a wide variety of viral families (Fig. 3). As expected from read-based profiling, the viral community in the juvenile blue tit ABR6116 was dominated by two Adenovirus vOTUs [73]. One of these, represented by the 26 kbp contig MEGAHIT-ABR-6151\_935 (available from <https://doi.org/10.5281/zenodo.13643307>), showed 99.16% NT sequence identity with the complete genome of great tit adenovirus 3 (GenBank ID: MW508338.2), derived from a great tit found dead in Germany. This high degree of sequence identity, which includes 99% identity at the AA level for the viral DNA polymerase, indicates that this blue tit siadenovirus belongs to the species *Siadenovirus carbocapituli*, which so far has been found only in great tits [73].

The juvenile blue tit ABR6154 had a high relative abundance of five vOTUs belonging to the Adintoviridae family, a viral family identified in several wild animals and in humans, but not previously reported in wild birds [92].

## CONCLUSIONS

Compared to the guts of humans and domesticated animals, the microbiology of the wild bird gut remains largely unexplored. Here, we deliver new insights into this important ecosystem while also showcasing the advantages of shotgun metagenomics in



**Fig. 3.** Overview of the parid faecal virome. (a) Relative abundance of the global viral community composition aggregated at the phylum level. The reads were mapped against the non-redundant vOTU catalogue, and the counts were normalized as read count per kb and million reads (cpm) before being aggregated at the phylum level. (b) Predicted bacterial hosts of the bacteriophage community. The viral community read counts were sub-setted for the Caudovirales fraction with a putative bacterial host identified. The cpm counts were aggregated at the host genus level, and the genera accounting for less than 200 cpm in the samples were grouped as ‘other’. (c) Relative abundance of the eukaryotic viral fraction. The read counts were sub-setted for the eukaryotic viral phylum and aggregated at the viral family level.

providing catalogues of genes and genome sequences that take us well beyond what can be achieved using 16S ribosomal RNA gene sequences. Exploration of just five faecal samples allowed the recovery of gene and genome sequences from viruses, coccidian parasites and new bacterial species, substantially increasing the known microbial diversity of these habitats. In addition, we have been able to recover mitochondrial genomes from the host species, including the first mitochondrial genomes from the Eurasian blue tit. Deposition of such sequences into publicly available databases will strengthen all future studies, improving the quality of reference-based taxonomic assignments.

Although our focus here is on microbial ecology, phylogenetic profiling of metagenomic reads has provided informative but incomplete insights into host diet, paving the way for the delivery of more definitive insights using sophisticated bioinformatic approaches twinned with extensive genome sequencing of the UK flora and fauna [93, 94]. Similarly, attempts to quantify parasite burden from metagenome sequences are likely to benefit from previous efforts using DNA barcodes [95].

Given the small number of samples analysed in this pilot study, we cannot hope to provide a comprehensive view of taxonomic diversity within the great tit and blue tit gut microbiomes. Nonetheless, this work gives us a tantalizing glimpse of the richness that awaits us when such approaches are fully optimized and rolled out more widely and established shotgun metagenomics as a potentially important tool that could help humans help birds address the challenges of the Anthropocene.

#### Funding information

G.L.D. was funded by the Isaac Newton Trust and Leverhulme Early Career Fellowship (ECF-2018-700) and the University of East Anglia.

**Author contributions**

Conceptualization: G.L.D. and M.J.P. Data curation: A.T. and M.J.P. Formal analysis: M.J.P., A.T. and A.J.P. Funding acquisition: G.L.D. Investigation: G.L.D., M.J.P., C.-J.M., D.H. Methodology: G.L.D., A.T., D.B., M.J.P. and A.J.P. Project administration: A.T. and M.J.P. Resources: G.L.D., A.T., D.B. and M.J.P. Software: A.T., M.J.P. and A.J.P. Writing – original draft: M.J.P. Writing – review and editing: G.L.D., A.T., D.B. and M.J.P.

**Conflicts of interest**

The authors declare that there are no conflicts of interest.

**Ethical statement**

This study was conducted with ethical approval from the University of East Anglia Animal Welfare and Ethical Review Body (AWERB) (Application ID: ETH2223-2022) and the University of Cambridge AWERB (Application ID: Z0067/19 and Z0068/19).

**References**

- Pallen MJ, Ponsero A, Telatin A, Moss C-J, Baker D, et al. Supplementary material to "Faecal metagenomes of great tits and blue tits provide insights into host, diet, pathogens and microbial biodiversity." *Microbiology Society Figshare*. 2024. . DOI: 10.6084/m9.figshare.26968957.v1 <https://doi.org/10.6084/m9.figshare.26968957.v1>
- Barko PC, McMichael MA, Swanson KS, Williams DA. The gastrointestinal microbiome: a review. *J Vet Intern Med* 2018;32:9–25.
- Bornbusch SL, Power ML, Schulkin J, Drea CM, Maslanka MT, et al. Integrating microbiome science and evolutionary medicine into animal health and conservation. *Biol Rev Camb Philos Soc* 2024;99:458–477.
- Anthony WE, Burnham CD, Dantas G, Kwon JH. The gut microbiome as a reservoir for antimicrobial resistance. *J Infect Dis* 2021;223:S209–S213.
- Ravi A, Halstead FD, Bamford A, et al. Loss of microbial diversity and pathogen domination of the gut microbiota in critically ill patients. *Microb Genom* 2019;5:e000293.
- Levin D, Raab N, Pinto Y, et al. Diversity and functional landscapes in the microbiota of animals in the wild. *Science* 2021;372.
- Salvadori M, Rosso G. Update on the gut microbiome in health and diseases. *World J Methodol* 2024;14:89196.
- Sergeant MJ, Constantinidou C, Cogan TA, Bedford MR, Penn CW, et al. Extensive microbial and functional diversity within the chicken cecal microbiome. *PLoS One* 2014;9:e91941.
- Gilroy R, Ravi A, Getino M, et al. Extensive microbial diversity within the chicken gut microbiome revealed by metagenomics and culture. *PeerJ* 2021;9:e10941.
- Suzuki TA. Links between natural variation in the microbiome and host fitness in wild mammals. *Integr Comp Biol* 2017;57:756–769.
- Ottinger MA, Mani S, Krüger S, Coverdale B, Willows-Munro S, et al. Microbiomes in birds: a review of links to health and reproduction. *Reprod Fertil* 2024;5:2024RAF–23.
- Lees AC, Haskell L, Allinson T, Bezeng SB, Burfield IJ, et al. State of the world's birds. *Annu Rev Environ Resour* 2022;47:231–260.
- Wei F, Wu Q, Hu Y, Huang G, Nie Y, et al. Conservation metagenomics: a new branch of conservation biology. *Sci China Life Sci* 2019;62:168–178.
- Jiang J, Hu D, Pei E. Integrated omics analysis reveals a correlation between gut microbiota and egg production in captive African penguins (*Spheniscus demersus*). *Anim Reprod Sci* 2024;263:107448.
- Klukowski N, Eden P, Uddin MJ, Sarker S. Virome of Australia's most endangered parrot in captivity evidenced of harboring hitherto unknown viruses. *Microbial Spectr* 2024;12:e0305223.
- Voegelé B, Saladin V, Wegmann M, Richner H. Parasites as mediators of heterozygosity-fitness correlations in the great tit (*Parus major*). *J Evol Biol* 2012;25:584–590.
- Rooyen J van, Lalubin F, Glaizot O, Christe P. Altitudinal variation in haemosporidian parasite distribution in great tit populations. *Parasit Vectors* 2013;6:139.
- Solomon G, Love AC, Vaziri GJ, Harvey J, Verrett T, et al. Effect of urbanization and parasitism on the gut microbiota of Darwin's finch nestlings. *Mol Ecol* 2023;32:6059–6069.
- Diez-Méndez D, Bodawatta KH, Freiberga I, Klečková I, Jönsson KA, et al. Indirect maternal effects via nest microbiome composition drive gut colonization in altricial chicks. *Mol Ecol* 2023;32:3657–3671.
- Davidson GL, Wiley N, Cooke AC, Johnson CN, Fouhy F, et al. Diet induces parallel changes to the gut microbiota and problem solving performance in a wild bird. *Sci Rep* 2020;10:20783.
- Davidson GL, Somers SE, Wiley N, Johnson CN, Reichert MS, et al. A time-lagged association between the gut microbiome, nestling weight and nestling survival in wild great tits. *J Anim Ecol* 2021;90:989–1003.
- Mark CM, Ian RH. Blue Tits. In: Jae CC (eds). *Encyclopedia of Animal Behavior*, Second edition. Oxford: Academic Press; 2019. pp. 11–22.
- Tryjanowski P, Nowakowski JJ, Indykiewicz P, Andrzejewska M, Śpica D, et al. *Campylobacter* in wintering great tits *Parus major* in Poland. *Environ Sci Pollut Res Int* 2020;27:7570–7577.
- Heylen D, Tijssse E, Fonville M, Matthyssen E, Sprong H. Transmission dynamics of *Borrelia burgdorferi* s.l. in a bird tick community. *Environ Microbiol* 2013;15:663–673.
- Leitzen E, Peters M, Merbach S, Wohlsein P, Baumgärtner W. *Suttonella ornithocola* detected within lesions of tit birds (*Paridae*) from epidemic death episodes in Germany, 2018–2020. *Front Vet Sci* 2022;9:977570.
- Lawson B, Lachish S, Colville KM, Durrant C, Peck KM, et al. Emergence of a novel avian pox disease in British tit species. *PLoS One* 2012;7:e40176.
- Lawson B, Malnick H, Pennycott TW, Macgregor SK, John SK, et al. Acute necrotising pneumonitis associated with *Suttonella ornithocola* infection in tits (*paridae*). *Vet J* 2011;188:96–100.
- Somers SE, Davidson GL, Johnson CN, Reichert MS, Crane JMS, et al. Individual variation in the avian gut microbiota: the influence of host state and environmental heterogeneity. *Mol Ecol* 2023;32:3322–3339.
- Benskin CMH, Rhodes G, Pickup RW, Mainwaring MC, Wilson K, et al. Life history correlates of fecal bacterial species richness in a wild population of the blue tit *Cyanistes caeruleus*. *Ecol Evol* 2015;5:821–835.
- Teyssier A, Lens L, Matthyssen E, White J. Dynamics of gut microbiota diversity during the early development of an avian host: evidence from a cross-foster experiment. *Front Microbiol* 2018;9:1524.
- Dion-Phénix H, Charmantier A, de Franceschi C, Bourret G, Kembel SW, et al. Bacterial microbiota similarity between predators and prey in a blue tit trophic network. *ISME J* 2021;15:1098–1107.
- Bodawatta KH, Freiberga I, Puzejova K, Sam K, Poulsen M, et al. Flexibility and resilience of great tit (*Parus major*) gut microbiomes to changing diets. *Anim Microbiome* 2021;3:20.
- Liukkonen M, Hukkanen M, Cossin-Sevrin N, Stier A, Vesterinen E, et al. No evidence for associations between brood size, gut microbiome diversity and survival in great tit (*Parus major*) nestlings. *Anim Microbiome* 2023;5:19.
- Hillmann B, Al-Ghalith GA, Shields-Cutler RR, Zhu Q, Gohl DM, et al. Evaluating the information content of shallow shotgun metagenomics. *mSystems* 2018;3:e00069-18.
- Yen S, Johnson JS. Metagenomics: a path to understanding the gut microbiome. *Mamm Genome* 2021;35:282–296.

36. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis. *Nat Biotechnol* 2017;35:833–844.
37. Gilroy R, Leng J, Ravi A, Adriaenssens EM, Oren A, et al. Metagenomic investigation of the equine faecal microbiome reveals extensive taxonomic diversity. *PeerJ* 2022;10:e13084.
38. Ravi A, Ereqat S, Al-Jawabreh A, Abdeen Z, Abu Shamma O, et al. Metagenomic profiling of ticks: Identification of novel rickettsial genomes and detection of tick-borne canine parvovirus. *PLoS Negl Trop Dis* 2019;13:e0006805.
39. Knutie SA, Gotanda KM. A non-invasive method to collect fecal samples from wild birds for microbiome studies. *Microb Ecol* 2018;34:851–855.
40. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018;34:i884–i890.
41. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol* 2019;20:257.
42. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 2015;31:1674–1676.
43. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990;215:403–410.
44. Karlicki M, Antonowicz S, Karnkowska A. Tiara: deep learning-based classification system for eukaryotic sequences. *Bioinformatics* 2022;38:344–350.
45. Wu YW, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 2016;32:605–607.
46. Kang DD, Li F, Kirton E, Thomas A, Egan R, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 2019;7:e7359.
47. Pan S, Zhao XM, Coelho LP. SemiBin2: self-supervised contrastive learning leads to better MAGs for short- and long-read sequencing. *Bioinformatics* 2023;39:i21–i29.
48. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, et al. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat Microbiol* 2018;3:836–843.
49. Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nat Methods* 2023;20:1203–1212.
50. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol* 2017;35:725–731.
51. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J* 2017;11:2864–2868.
52. Eren AM, Kiehl E, Shaiber A, Veseli I, Miller SE, et al. Community-led, integrated, reproducible multi-omics with anvio. *Nat Microbiol* 2021;6:3–6.
53. Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil PA, et al. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res* 2022;50:D785–D794.
54. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* 2022;38:5315–5316.
55. Shaffer M, Borton MA, McGivern BB, Zayed AA, La Rosa SL, et al. DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res* 2020;48:8883–8900.
56. Suzek BE, Huang H, McGarvey P, Mazumder R, Wu CH. UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics* 2007;23:1282–1288.
57. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, et al. The Pfam protein families database. *Nucleic Acids Res* 2004;32:D138–41.
58. Yin Y, Mao X, Yang J, Chen X, Mao F, et al. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 2012;40:W445–51.
59. Camargo AP, Roux S, Schulz F, Babinski M, Xu Y, et al. Identification of mobile genetic elements with geNomad. *Nat Biotechnol* 2024;42:1303–1312.
60. Nayfach S, Camargo AP, Schulz F, Eloe-Fadrosh E, Roux S, et al. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nat Biotechnol* 2021;39:578–585.
61. Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol* 2017;35:1026–1028.
62. Roux S, Camargo AP, Coutinho FH, Dabdoub SM, Dutilh BE, et al. iPHoP: An integrated machine learning framework to maximize host prediction for metagenome-derived viruses of archaea and bacteria. *PLoS Biol* 2023;21:e3002083.
63. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009;10:R25.
64. Loman NJ, Constantinidou C, Christner M, Rohde H, Chan JZ-M, et al. A culture-independent sequence-based metagenomics approach to the investigation of an outbreak of Shiga-toxicogenic *Escherichia coli* O104:H4. *JAMA* 2013;309:1502–1510.
65. Gilroy R, Adam ME, Kumar B, Pallen MJ. An initial genomic blueprint of the healthy human oesophageal microbiome. *Access Microbiol* 2023;5:v3.
66. Perrins CM. Tits and their caterpillar food supply. *Ibis* 1991;133:49–54.
67. Pollock CJ, Capilla-Lasheras P, McGill RAR, Helm B, Dominoni DM. Integrated behavioural and stable isotope data reveal altered diet linked to low breeding success in urban-dwelling blue tits (*Cyanistes caeruleus*). *Sci Rep* 2017;7:5014.
68. Coomes JR. Cognition and personality when deciding what to eat: foraging plasticity and diet in the great tit. PhD Thesis, University College Cork; 2021. <https://cora.ucc.ie/items/b4a9401c-5971-4219-b1f2-5cb2fc8d8446>
69. Dalby MJ, Hall LJ. Recent advances in understanding the neonatal microbiome. *F1000Res* 2020;9:422.
70. Leggett RM, Alcon-Giner C, Heavens D, Caim S, Brook TC, et al. Rapid MinION profiling of preterm microbiota and antimicrobial-resistant pathogens. *Nat Microbiol* 2020;5:430–442.
71. Hildebrand F, Moitinho-Silva L, Blasche S, Jahn MT, Gossmann TI, et al. Antibiotics-induced monodominance of a novel gut bacterial order. *Gut* 2019;68:1781–1790.
72. Kovács ER, Jánoska M, Dán A, Harrach B, Benko M. Recognition and partial genome characterization by non-specific DNA amplification and PCR of a new siadenovirus species in a sample originating from *Parus major*, a great tit. *J Virol Methods* 2010;163:262–268.
73. Gellért Á, Benkő M, Harrach B, Peters M, Kaján GL. The genome and phylogenetic analyses of tit siadenoviruses reveal both a novel avian host and viral species. *Infect Genet Evol* 2022;103:105326.
74. Ogedengbe ME, El-Sherry S, Ogedengbe JD, Chapman HD, Barta JR. Phylogenies based on combined mitochondrial and nuclear sequences conflict with morphologically defined genera in the eimeriid coccidia (Apicomplexa). *Int J Parasitol* 2018;48:59–69.
75. Gonzalez A, Vázquez-Baeza Y, Pettengill JB, Ottesen A, McDonald D, et al. Avoiding pandemic fears in the subway and conquering the platypus. *mSystems* 2016;1:00050–16.
76. Laine VN, Gossmann TI, van Oers K, Visser ME, Groenen MAM. Exploring the unmapped DNA and RNA reads in a songbird genome. *BMC Genomics* 2019;20:10.
77. Li X, Lin L, Cui A, Bai J, Wang X, et al. Taxonomic status and phylogenetic relationship of tits based on mitogenomes and nuclear segments. *Mol Phylogenet Evol* 2016;104:14–20.

78. Taberlet P, Meyer A, Bouvet J. Unusual mitochondrial DNA polymorphism in two local populations of blue tit *Parus caeruleus*. *Mol Ecol* 1992;1:27–36.
79. Rejman EE, Hák-Kovács K, Barta JR. The first isospora species (*Apicomplexa: Eimeriidae*) described from the northern yellow-shafted flicker (*Colaptes auratus luteus*) in Ontario, Canada. *Acta Parasitol* 2022;67:1162–1171.
80. Trefančová A, Kvičero J. *Isospora svecica* sp. n. (*Apicomplexa: Eimeriidae*), a new species of coccidium from the white-spotted bluethroat *Luscinia svecica cyanecula* (*Aves: Passeriformes: Muscicapidae*). *Parasitol Res* 2019;118:3043–3051.
81. Goodenough AE, Stallwood B. Differences in culturable microbial communities in bird nestboxes according to orientation and influences on offspring quality in great tits (*Parus major*). *Microb Ecol* 2012;63:986–995.
82. Pallen MJ, Rodriguez-R LM, Alikhan NF. Naming the unnamed: over 65,000 candidatus names for unnamed archaea and bacteria in the genome taxonomy database. *Int J Syst Evol Microbiol* 2022;72.
83. Aljorayid A, Viau R, Castellino L, Jump RLP. *Serratia fonticola*, pathogen or bystander? a case series and review of the literature. *IDCases* 2016;5:6–8.
84. Müller HE, Steigerwalt AG, Brenner DJ. Isolation of *Serratia fonticola* from birds. *Zentralbl Bakteriol Mikrobiol Hyg A* 1986;261:212–218.
85. Fuentes-Castillo D, Power P, Cerdeira L, Cardenas-Arias A, Moura Q, et al. FONA-7, a novel extended-spectrum  $\beta$ -lactamase variant of the FONA family identified in *Serratia fonticola*. *Microb Drug Resist* 2021;27:585–589.
86. Uyanik T, Gülel GT, Alişarlı M. Characterization of extended-spectrum beta-lactamase-producing *Enterobacterales* from organic and conventional chicken meats. *Lett Appl Microbiol* 2021;72:783–790.
87. Kenzaka T, Tani K. Draft genome sequence of extended-spectrum beta-lactamase-producing *Serratia fonticola* BWK15 isolated from feces of *Anas penelope*. *Genome Announc* 2017;5:e01102-17.
88. Mohsin M, Raza S, Schaufler K, Roschanski N, Sarwar F, et al. High prevalence of CTX-M-15-type ESBL-producing *E. coli* from migratory avian species in Pakistan. *Front Microbiol* 2017;8:2476.
89. Liakopoulos A, Olsen B, Geurts Y, Artursson K, Berg C, et al. Molecular characterization of extended-spectrum-cephalosporin-resistant *Enterobacteriaceae* from wild kelp gulls in South America. *Antimicrob Agents Chemother* 2016;60:6924–6927.
90. Dolejska M, Literak I. Wildlife is overlooked in the epidemiology of medically important antibiotic-resistant bacteria. *Antimicrob Agents Chemother* 2019;63:e01167-19.
91. Guenther S, Ewers C, Wieler LH. Extended-spectrum beta-lactamases producing *E. coli* in wildlife, yet another form of environmental pollution? *Front Microbiol* 2011;2:246.
92. Starrett GJ, Tisza MJ, Welch NL, Belford AK, Peretti A, et al. Adintoviruses: a proposed animal-tropic family of midsize eukaryotic linear dsDNA (MELD) viruses. *Virus Evol* 2021;7:veaa055.
93. Worth JRP, Kikuchi S, Kanetani S, Takahashi D, Aizawa M, et al. Chloroplast genome-based genetic resources via genome skimming for the subalpine forests of Japan and adjacent regions. *Ecol Evol* 2024;14:e11584.
94. Lewin HA, Robinson GE, Kress WJ, Baker WJ, Coddington J, et al. Earth biogenome project: sequencing life for the future of life. *Proc Natl Acad Sci USA* 2018;115:4325–4333.
95. Abdu S, Chimento M, Alarcón-Nieto G, Zúñiga D, Aplin LM, et al. The performance of field sampling for parasite detection in a wild passerine. *Ecol Evol* 2022;12:e9242.

**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](https://microbiologyresearch.org)**