Metagenomics reveals fibre fermentation and AMR pathways in red grouse (*Lagopus scotica*) microbiota

Anum Ali Ahmad[1], Kathy Fletcher[2], Nicholas Hesford[2], Laura Glendinning[1]*

[1] The Roslin Institute, The University of Edinburgh, Easter Bush Campus, Midlothian, Edinburgh, EH25 9RG, UK.

[2] The Game and Wildlife Conservation Trust, Hopetoun Estates Office, Home Farm, Hopetoun, South Queensferry, EH30 9SL, UK.

*Corresponding author: Laura.glendinning@roslin.ed.ac.uk

Abstract:

Background: The avian caecal microbiota plays a vital role in host nutrition, enabling non-digestible, fibrous material to be converted into compounds that can be absorbed and used as an energy source by the host. The diet of adult red grouse (*Lagopus scotica*) is dominated by heather (*Calluna vulgaris*), which is particularly high in fibre. It is therefore likely that the caecal microbiota plays a key role in enabling grouse to thrive on this diet. In this study, we present the first characterisation of the caecal microbiota of red grouse using modern sequencing methods.

Results: We performed metagenomic sequencing on caecal content samples from fifteen red grouse from three upland estates in Scotland. From this data, we constructed and characterised twelve high-quality, species-level metagenome assembled genomes (MAGs). Eleven of these MAGs could not be assigned a taxonomic label at species level, indicating that they may be novel species. MAGs belonged to diverse taxa (5 phyla) and several encoded genes and pathways for the digestion of fibres, including cellulose, hemi-cellulose, xylooligosaccharides and pectin. Several MAGs also contained antimicrobial resistance genes, predominantly related to vancomycin resistance.

Conclusions: This study is the first to reconstruct commensal microbial genomes from red grouse. The caeca contain diverse, often novel, microbial taxa capable of fermenting various fibres, potentially aiding in the digestion of the red grouse's high-fibre diet. Further research is necessary to explore how these bacteria support red grouse nutrition and health.

Keywords: Microbiota; avian; metagenomics; gut; caeca; caecum; grouse; MAG; fibre; AMR.

Background:

The avian gut microbiota plays a vital role in maintaining host health and contributing to host nutrition (1). While the vast majority of avian microbiota studies have been conducted in farmed species, there are an increasing number of studies examining the microbiota of wild bird populations. These studies have shown that various environmental and genetic factors are important influencers of the composition and function of the gut microbiota (2).

An important challenge when studying wild species is acquiring samples that are truly representative of the gut microbiota. Studies that use wild birds are often limited to collecting non-invasive samples, such as faeces and caecal swabs. However, these samples do not quantitatively represent the microbiota of the large and small intestine (3, 4). Obtaining samples directly from the gastrointestinal section of interest is therefore advantageous, but is often not possible as it requires killing the bird. Game bird species offer a potential avenue for studying the microbiota in “wild” or minimally-managed avian populations, where gastrointestinal samples can be obtained in an ethical manner. Game birds are regularly shot during hunts, or to collect samples to monitor population health. The gastrointestinal tracts from these birds are usually discarded and can therefore be seen as a waste product of this industry.

Most game bird species consume plants as a large component of their diet. Diet has been reported as the most important factor affecting the wild avian microbiota (5). For wild avian species that consume a predominantly or partially plant-based diet, it is likely that the gut microbiota plays an important role in enabling the host to extract energy from plant fibres. Members of the gut microbiota can ferment fibre into short chain fatty acids that can be used as an energy source by the host (6). In avians the vast majority of fibre fermentation occurs in the caeca (7), blind pouches located near the end of the gastrointestinal tract. In general, herbivorous avian species have longer caeca than non-herbivorous species (8).

Red grouse (*Lagopus scotica*) represent a particularly interesting game bird species in which to study microbiota fibre fermentation. They sit within the same taxonomic family as chickens (Phasianidae), making them an interesting comparison to an important domestic species. The main component of the adult red grouse diet is heather (*Calluna vulgaris*), which is high in fibre (25%) and low in protein (7%) (9). Red grouse have long caeca (~75 cm), which increase in size when they are supplied with a higher fibre diet (10). From feed digestibility studies, there is some evidence that the red grouse microbiota is able to digest a wide variety of fibres, including those that are generally considered more resistant to fermentation (e.g. lignin) (9).

The gut microbiota has been studied in several grouse species using 16S rRNA gene sequencing and/or metagenomic sequencing, including in capercaillies (11, 12), greater sage grouse (13-15), wild prairie grouse (16) and ptarmigans (17-19). However, while bacteria have been cultured from the gastrointestinal tracts of red grouse (20, 21), their microbiota has not been characterised using modern sequencing methods. In this study, we perform metagenomic sequencing analysis of red grouse caecal content samples. We construct twelve high-quality bacterial metagenome assembled genomes, and characterise their functional potential to metabolise fibre and their antimicrobial resistance (AMR) potential.

Methods:

Sample collection and processing

Samples were collected from red grouse killed during legal game bird shoots within the open season in Scotland. Samples were collected in the 2023 Scottish red grouse (*Lagopus scotica*) shooting season (August-October), at three distinct areas of moorland managed for recreational hunting of red grouse (Latitude and longitude: 55.75 -3.01 (Samples RS1-5), 55.98 -3.33 (Samples RB1-5), and 57.02 -4.12 (Samples CA1-CA5)). Samples were collected from five birds per area. The intestines were removed from birds and stored at -80°C until DNA extraction. DNA was extracted from caecal contents using the QIAamp PowerFecal Pro DNA Kit according to the manufacturer's instructions. DNA quality and quantity were assessed using a NanoDrop spectrophotometer and a Qubit Fluorometer with a Qubit dsDNA Quantification Assay Kit (broad range). Shotgun metagenomic Ilumina sequencing was performed by Novogene, using an Illumina NovaSeq 6000 producing 150 bp paired-end reads. Samples RS1 and RS4 were produced on two separate sequencing runs in order to produce sufficient data quantity, then fastqs from the same sample were concatenated prior to analysis.

Bioinformatic analysis

Sequences were checked for quality using FastQC (v.0.12.1) (22) and MultiQC (v.1.6) (23). Quality control was performed using fastp (v.0.24) (24), with the options "--detect_adapter_for_pe -l 30 -q 20 --trim_poly_g --trim_poly_x". The host genome and feed genome were removed from adaptor trimmed read files. There is currently no RefSeq or GenBank genome available for *Lagopus scotica*, therefore the *Lagopus muta* (rock ptarmigan) genome was downloaded from RefSeq (GCF_023343835.1) and used instead of a host genome. Heather (*Calluna vulgari*) is the main food source for adult red grouse (25); therefore the *Calluna vulgari* genome was downloaded from GenBank (GCA_964145215.1).

Sample reads were mapped to host and feed reference genomes using BWA-MEM (v.0.7.18) (26), followed by SAMtools (v.1.21, samtools fastq -f 12) (27) to select reads where both paired-end reads were unmapped. Single-sample assemblies and coassemblies were constructed using Megahit (v.1.2.9) (28).

CoverM (v.0.7) (29) was used to create depth files, which were used as input for genome binning by MetaBAT 2 (v.2.17) (30). Contamination and completeness of bins were estimated using CheckM2 (v.1.0.2) (31). dRep (v.3.5.0) (32) was used to dereplicate genome bins into MAGs, with minimum completeness set at 80% and maximum contamination at 10%. Viruses were identified from megahit assemblies using VirSorter2 (v.2.2.4: database downloaded 3rd February 2025) (33) (options: --include-groups dsDNAphage,ssDNA --min-length 5000 --min-score 0.5 -j 28 all).

GTDB-Tk (v.2.4.0) (34) was used to assign taxonomy to MAGs (database downloaded 24th January 2025). METABOLIC was used to assign metabolic function to MAGs (v.4.0) (35). Antimicrobial resistance (AMR) genes present in MAGs and megahit assemblies were identified using RGI main (v. 6.0.3) (36) with the Comprehensive Antibiotic Resistance Database (database downloaded 17th March 2025). Loose hits were included, and the --low_quality option was used for megahit assemblies. The abundance of MAGs within samples was calculated using CoverM (options: -p bwa-mem -m mean relative_abundance trimmed_mean covered_bases variance length reads_per_base count rpkm tpm --min-read-aligned-percent 75 --min-read-percent-identity 95 --min-covered-fraction 0). Graphs and data summaries were produced in R using the packages reshape2 (v.1.4.4) (37), cowplot (v.1.1.3) (38), dplyr (v.1.1.4) (39), ggplot2 (v.3.5.1) (40) and tidyverse (v.2.0.0) (41). MAG taxonomic trees were constructed using PhyloPhlAn (v.3.1.1) (42) (options: --diversity high -f supermatrix_aa.cfg --subsample phylophlan –fast). Taxonomic trees were midpoint rerooted and visualised using figtree (v.1.4.3) (43).

Results:

Sequence quality

We performed metagenomic sequencing on caecal content samples from fifteen Scottish red grouse. After removal of poor-quality sequences, and trimming of adaptors, polyAs and polyGs, there were on average 102,686,250 ± 24,122,717 sequences per sample. After removal of host and feed sequences the average number of reads per sample dropped by 89%, to 10,902,109 ± 7,733,350. This was predominantly due to the removal of host reads, indicating that our samples had a large proportion of host contamination. From single sample assemblies and a coassembly, we constructed 161 putative genome bins. Quality-filtering and dereplication resulted in twelve high-quality (<10% contamination, >80% complete) species-level (95% ANI) MAGs (**Table 1**). We also constructed strain-level MAGs (99% ANI), but as we found only one more MAG than at species-level we decided to restrict our further analyses to species level. Eighteen high quality viral genomes were identified (**Table S1**), two of which were proviruses.

Table 1: **Genome stats for species-level MAGs**

MAG	**GTDB taxonomy**	**Contamination (%)**	**Completeness (%)**	**Contig N50**	**Genome size (bp)**	**Total coding sequences**	**GC (%)**
metabat2_RB1.3	d__Bacteria;p__Bacillota_A;c__Clostridia;o__Oscillospirales;f__Acutalibacteracea	100	0.17	167,262	1,021,943	930	33

	e;g__CASPLG01;s__						
metabat2_RB3.3	d__Bacteria;p__Bacillota_A;c__Clostridia;o__Christensenellales;f__Borkfalkiaceae;g__Coprophlasma;s__	97.91	0	54,893	1,424,430	1,267	49
metabat2_RS5.3	d__Bacteria;p__Campylobacterota;c__Campylobacteria;o__Campylobacterales;f__Helicobacteraceae;g__Helicobacter_B;s__	99.78	0.01	27,466	1,541,840	1,592	37
metabat2_coassembly.1	d__Bacteria;p__Bacillota_A;c__Clostridia;o__Lachnospirales;f__CA	98.34	0.27	28,654	1,333,034	1,241	28

	G-274;g__;s__						
metabat2_coassembly.10	d__Bacteria;p__Verrucomicrobiota;c__Verrucomicrobiae;o__Opitutales;f__LL51;g__JAIDVZ01;s__	83.77	0.37	8,043	804,428	898	32
metabat2_coassembly.14	d__Bacteria;p__Bacillota_A;c__Clostridia;o__Christensenellales;f__Borkfalkiaceae;g__Coprophlasma;s__	88.08	0.05	6,612	1,044,100	1,098	52
metabat2_coassembly.16	d__Bacteria;p__Actinomycetota;c__Actinomycetes;o__Actinomycetales;f__Actinomycetaceae;g__Scrofimicr	96.84	0	30,426	1,474,562	1,248	50

	obium;s__						
metabat2_coassembly.28	d__Bacteria;p__Bacillota_A;c__Clostridia;o__Lachnospirales;f__Lachnospiraceae;g__Eubacterium_I;s__	100	0.02	13,892	2,419,852	2,339	48
metabat2_coassembly.36	d__Bacteria;p__Bacillota_A;c__Clostridia;o__Oscillospirales;f__Oscillospiraceae;g__;s__	100	0.06	20,806	2,734,688	2,460	57
metabat2_coassembly.50	d__Bacteria;p__Actinomycetota;c__Coriobacteriia;o__Coriobacteriales;f__UMGS124;g__JAHZGC01;s__	92.48	3.21	11,155	2,307,052	1994	68

metabat2_coassembly.53	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Prevotella;s__	81.06	1.36	8,397	2,145,104	1,915	56
metabat2_coassembly.7	d__Bacteria;p__Bacillota_A;c__Clostridia;o__Oscillospirales;f__Ruminococcaceae;g__RGIG1955;s__RGIG1955 sp017400345	97.55	0.04	73,819	871,746	975	35

Taxonomy of MAGs

All MAGs belonged to the kingdom bacteria (**Figure 1**), with the most common phylum being Bacillota_A (n=7), all of the members of which belonged to the class Clostridia. Two members of this phylum were identified as members of the genus *Coproplasma,* within the family Christensenellales. The remaining MAGs in this phylum belonged to Lachnospirales (one identified to family CAG-274 and one identified to the genus *Eubacterium*), and Oscillospirales (one identified to family Oscillospiraceae, one identified to genus *CASPLG01*,

and one identified as *RGIG1955 sp017400345*). This latter MAG was the only one from our dataset to be identified to species level by GTDB-Tk. Two MAGs were identified as members of the phylum Actinomycetota, and were assigned to the genera *Scrofimicrobium* and *JAHZGC01*. One MAG belonged to the phylum Bacteroidota (genus *Prevotella*), one belonged to the phylum Campylobacterota (genus *Helicobacter_B*) and one belonged to the phylum Verrucomicrobiota (genus *JAIDVZ01*).

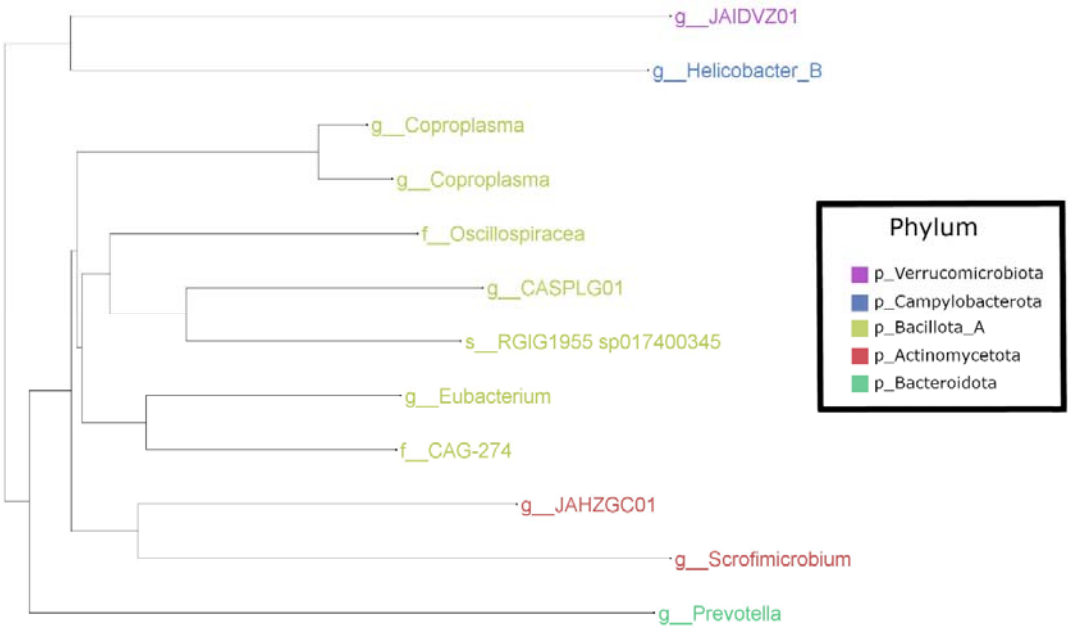


Figure 1: Taxonomic tree of species-level, high-quality metagenome assembled genomes constructed from red grouse caecal metagenomic data. Tip labels represent the lowest taxonomy assigned to each MAG by GTDB-Tk.

For most samples, the majority of reads did not map to MAGs (69.5 ± 14.3% relative abundance) (**Figure 2**). By far the most abundant MAG was metabat2_RS5.3 (genus *Helicobacter_B*) (26.3 ± 15.4%), followed by metabat2_RB1.3 (genus *CASPLG01*) (1.9 ± 1.3%) and metabat2_RB3.3 (genus *Coproplasma*) (1.4 ± 1.9%). Some samples contained particularly high levels of *Helicobacter_B*, including RB2 (64.5%), RB4 (46.7%) and RB5 (44.0%).

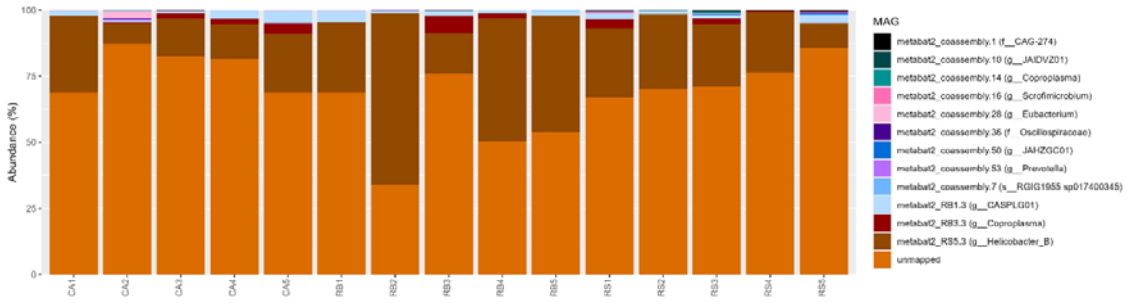


Figure 2: Relative abundance of species-level, high-quality metagenome assembled genomes in red grouse caecal samples. Taxonomies represent the lowest taxonomy assigned to each MAG by GTDB-Tk.

<u>Metabolism of MAGs</u>

Carbohydrate active enzymes (CAZymes) play a key role in carbohydrate fermentation by the microbiota, including the fermentation of fibrous compounds. Overall, 40 different CAZymes were identified in our MAGs. The most common CAZyme (18 genes) was GH13, followed by GH3 (16 genes), GH25 (11 genes), GH2 (9 genes), GH43 (8 genes), GH18, (7 genes), GH77 (7 genes), GH23 (5 genes), GH28 (5 genes). All other CAZymes had less than 5 genes present in our MAGs (**Table S2**). MAG metabat2_coassembly.53 (*Prevotella*) contained substantially more CAZyme genes than any other MAG: 55 genes belonging to 27 CAZyme families. Metabat2_coassembly.28 (*Eubacterium*) and metabat2_coassembly.36 (Oscillospiraceae) also contained a large number of CAZyme genes: 25 genes from 17 CAZyme families and 20 genes from 13 families respectively. The three most abundant MAGs in our samples contained far fewer CAZyme genes: metabat2_RS5.3 (*Helicobacter_B*) contained 3 genes from 2 families, metabat2_RB1.3 (*CASPLG01*) contained 7 genes from 4 families and metabat2_RB3.3 (*Coproplasma*) contained 9 genes from 5 families.

Several MAGs were predicted to be able to metabolise complex carbohydrates commonly found in plants. Cellulose degradation via beta-glucosidase was present in five of our MAGs, including the taxa *Coproplasma*, *Eubacterium*, Oscillospiracea and *Prevotella*. Cellulose degradation via cellobiosidase or cellulase was not present in our dataset. Hemicellulose-debranching enzymes were also present amongst our MAGs, including arabinosidase (*Prevotella*), beta-glucuronidase (*Coproplasma* and *JAHZGC01*) and alpha-L-rhamnosidase (Oscillospiraceae and *Prevotella*). One endohemicellulase gene (alpha-D-xyloside xylohydrolase) was also identified in our *Prevotella* MAG. Beta-xylosidase (targeting xylooligosaccharides) was also only identified in this *Prevotella* MAG.

Heather is known to contain various phenolic compounds. Enzymes for converting phenol into benzoyl-CoA were found in two MAGs (Oscillospiraceae and *Helicobacter_B*). Although starch is present in low concentrations in heather, amylolytic enzymes involved in starch digestion were also present in two MAGs: alpha-amylase in the *Prevotella* MAG and pullulanase in the *JAIDVZ01* MAG. Only two genes were identified related to sulphur metabolism, namely sulphite reduction in the *Eubacterium* MAG, and sulphur oxidation in the *CAG-274*.

<u>Antimicrobial resistance genes</u>

We next investigated whether our MAGs contained antimicrobial resistance (AMR) genes (**Table S3**). Whilst we do not expect any of the sampled red grouse to have been directly treated with antibiotics, AMR genes occur naturally in many bacteria, even without antibiotic exposure. The RGI “strict” algorithm is used to detect previously unknown variants of known AMR genes. Twenty-nine AMR genes were detected in our MAGs using this algorithm (**Table 2**), the vast majority of which encoded resistance to vancomycin (27 genes in 9 MAGs). A metronidazole resistance gene was detected in metabat2_RB3.3 (*Coproplasma*) and a tetracycline gene in metabat2_RS5.3 (*Helicobacter*).

The RGI “loose” algorithm aims to identify more distant homologs of AMR genes. While this method is more prone to false positives, we decided to cautiously employ it as members of the grouse microbiota are likely to be poorly represented in the CARD database. This method identified 13,159 potential AMR gene homologs amongst our MAGs, with the top three targeted antibiotics as oxacillin, tetracycline and vancomycin. The MAGs with the most putative AMR gene homologs were metabat2_coassembly.36 (n=1972, Oscillospiraceae), metabat2_coassembly.28 (n=1882, *Eubacterium*), and metabat2_coassembly.50 (n=1525, *JAHZGC01*).

Table 2: AMR genes present in MAGs (strict RGI hits)

MAG	Taxonomy	Antibiotic	Resistance Mechanism	AMR Gene Family
metabat2_coassembly.1	f__CAG-274	vancomycin; teicoplanin	antibiotic target alteration	vanW; glycopeptide resistance gene cluster
metabat2_coassembly.1	f__CAG-274	vancomycin	antibiotic target alteration	glycopeptide resistance gene cluster; Van ligase
metabat2_coassembly.1	f__CAG-274	vancomycin	antibiotic target alteration	glycopeptide resistance gene cluster; vanT
metabat2_coassembly.14	g__Coprophlasma	vancomycin; teicoplanin	antibiotic target alteration	vanW; glycopeptide resistance gene cluster
metabat2_coassembly.28	g__Eubacterium	vancomycin;	antibiotic target	vanW; glycopeptide resistance gene cluster

		teicoplanin	alteration	
metabat2_co assembly.28	g__Eubacteri um	vancomyci n	antibiotic target alteration	glycopeptide resistance gene cluster; vanT
metabat2_co assembly.36	f__Oscillospir aceae	vancomyci n; teicoplanin	antibiotic target alteration	vanW; glycopeptide resistance gene cluster
metabat2_co assembly.36	f__Oscillospir aceae	vancomyci n	antibiotic target alteration	glycopeptide resistance gene cluster; vanT
metabat2_co assembly.36	f__Oscillospir aceae	vancomyci n	antibiotic target alteration	vanY; glycopeptide resistance gene cluster
metabat2_co assembly.36	f__Oscillospir aceae	vancomyci n; teicoplanin	antibiotic target alteration	vanW; glycopeptide resistance gene cluster
metabat2_co assembly.36	f__Oscillospir aceae	vancomyci n	antibiotic target alteration	glycopeptide resistance gene cluster; Van ligase
metabat2_co assembly.36	f__Oscillospir aceae	vancomyci n	antibiotic target alteration	glycopeptide resistance gene cluster; vanT
metabat2_co assembly.36	f__Oscillospir aceae	vancomyci n	antibiotic target alteration	glycopeptide resistance gene cluster; Van ligase
metabat2_co assembly.36	f__Oscillospir aceae	vancomyci n	antibiotic target	vanW; glycopeptide resistance gene cluster

			alteration	
metabat2_co assembly.50	g__JAHZGC 01	vancomycin	antibiotic target alteration	glycopeptide resistance gene cluster; vanXY
metabat2_co assembly.50	g__JAHZGC 01	vancomycin	antibiotic target alteration	glycopeptide resistance gene cluster; Van ligase
metabat2_co assembly.50	g__JAHZGC 01	vancomycin	antibiotic target alteration	glycopeptide resistance gene cluster; vanT
metabat2_co assembly.50	g__JAHZGC 01	metronidazole	antibiotic inactivation	nitroimidazole reductase
metabat2_co assembly.50	g__JAHZGC 01	vancomycin	antibiotic target alteration	vanH; glycopeptide resistance gene cluster
metabat2_co assembly.53	g__Prevotella	vancomycin	antibiotic target alteration	glycopeptide resistance gene cluster; vanT
metabat2_co assembly.7	s__RGIG1955 sp017400345	vancomycin	antibiotic target alteration	vanY; glycopeptide resistance gene cluster
metabat2_co assembly.7	s__RGIG1955 sp017400345	vancomycin	antibiotic target alteration	glycopeptide resistance gene cluster; Van ligase
metabat2_co	s__RGIG195	vancomyci	antibiotic	glycopeptide resistance gene

assembly.7	5 sp017400345	n	target alteration	cluster; vanXY
metabat2_coassembly.7	s__RGIG1955 sp017400345	vancomycin	antibiotic target alteration	glycopeptide resistance gene cluster; vanT
metabat2_RB1.3	g__CASPLG01	vancomycin	antibiotic target alteration	glycopeptide resistance gene cluster; vanT
metabat2_RB1.3	g__CASPLG01	vancomycin	antibiotic target alteration	glycopeptide resistance gene cluster; Van ligase
metabat2_RB1.3	g__CASPLG01	vancomycin	antibiotic target alteration	vanY; glycopeptide resistance gene cluster
metabat2_RB3.3	g__Coprophasma	vancomycin; teicoplanin	antibiotic target alteration	vanW; glycopeptide resistance gene cluster
metabat2_RS5.3	g__Helicobacter	tetracycline	antibiotic efflux	resistance-nodulation-cell division (RND) antibiotic efflux pump

227

Discussion:

The red grouse represents a particularly interesting species in which to study the microbiota, due to its reliance on a single, high-fibre plant species (heather - *Calluna vulgaris*) for the majority of its diet (9). Avians do not produce fibre fermenting enzymes, and therefore in order to acquire nutrition from fibre they are reliant on their gut microbiota to ferment fibre into compounds that they can absorb and use as an energy source (e.g. short chain fatty acids) (6). In this study, we examined the gut microbiota of red grouse for the first time using modern sequencing methods, using metagenomics to construct metagenome assembled genomes. Unlike metabarcoding methods (e.g. 16S rRNA gene sequencing), metagenomics allows us to not only study the taxa that are present in a community, but also provides functional information on their metabolic and virulence potential. By constructing MAGs, we can link genes to their species of origin. In this study we identified twelve high-quality, species-level MAGs.

Of our twelve MAGs, eleven could not be assigned to a species by GDTB-Tk. The GTDB is constructed from high-quality genomes obtained from RefSeq and GenBank, and the release used in this study contains 596,859 microbial genomes in 113,104 species clusters, including genomes constructed from metagenomic data. If genomes cannot be assigned a taxonomy by this tool it is therefore a good indication that they are likely to be novel. Only one MAG was assigned a species label, RGIG1955 sp017400345 within the Ruminococcaceae family. This is a poorly characterised species, which has previously only been identified in metagenomics data from ruminants (44, 45). A MAG from the genus *Helicobacter_B* was the most abundant MAG in our samples, comprising on average 26% of the microbiota. *Helicobacter* is a highly diverse genus, which while commonly associated with gastrointestinal disease (e.g. *H.pylori* infection in humans) is also a common member of the gut microbiota in many species, including grouse (13, 18). In ptarmigan intestines, Bjørnsen *et al.* found similar average abundances of *Helicobacter* as observed in our study (24% ± 3) (46).

The majority of our MAGs were identified as Clostridia, which is a commonly identified taxon in the guts of other grouse species (13, 14, 17). Of these MAGs, most are from poorly described taxa identified from metagenomic data: the genus *Coproplasma* has been identified in the chicken gut (47), the genus *CASPLG01* has been identified in mice (48-50), and the family CAG-274 has been identified in metagenomic datasets from diverse species (51-53). In contrast, the diverse genus *Eubacterium* is relatively well studied, and has been proposed as a potential probiotic in humans due to the ability of some species to produce butyrate (54). Of the two MAGs assigned to the phylum Actinomycetota, one belonged to the genus *JAHZGC01* which has only one representative in the GTDB, obtained from human metagenomic data (55). The other MAG belonged to the genus *Scrofimicrobium*, which has been isolated from the gastrointestinal tracts of diverse species (56, 57).Only one MAG was identified as belonging to the phylum Verrucomicrobiota, a member of the genus *JAIDVZ01*, which is poorly characterised and has only been identified in metagenomic data (58, 59). Conversely, the only MAG identified in the phylum Bacteroidota belonged to the very well characterised genus *Prevotella*. This genus is known to play a key role in the fermentation of fibre in various herbivores (60) and is commonly found in the microbiota of avians, including grouse species (13, 14).

It is therefore unsurprising that our *Prevotella* MAG contained the largest quantity of CAZyme genes amongst our MAGs, and was predicted to degrade fibrous compounds including cellulose, hemi-celluloses and xylooligosaccharides. However, several MAGs from less well characterised taxa were also found to encode for diverse CAZyme genes and to be capable of degrading fibres, highlighting the potential role of poorly characterised taxa in fibre fermentation in the grouse caeca. Several CAZymes families with xylooligosaccharide, pectin and hemicellulose degrading activity (e.g. GH36, GH42, GH73) were identified both in our dataset and in the microbiota of other grouse species (11, 17). Various phenolic compounds are also found in heather, and have been shown to affect fermentation in the rumen microbiota (61). Pathways for the degradation of phenols were found in both our

MAGs and other grouse species (14, 17), as were pathways for starch degradation (17). These findings indicate some degree of shared function in the grouse gut microbiota, in relation to digestion of plant-derived components of the diet.

Our MAGs also contained several antimicrobial resistance genes. Using the strict RGI algorithm, the most commonly targeted antibiotic by AMR genes was vancomycin. This included the gene families: vanW, Van ligase, vanT, vanY, vanXY, and vanH. These genes are present in a wide range of taxa bacterial taxa (62) and have also been identified in samples from wild western capercaillies (11), but not amongst wild prairie grouse (16). Many of these resistance genes are not well characterised, with limited information on their molecular structures (62). vanT is found within resistance cassettes and encodes a membrane-bound serine racemase, while vanH encodes a dehydrogenase that converts pyruvate to d-lactate. The mechanism of action of vanW is currently unknown. vanY and vanXY are not required for vancomycin resistance but may contribute to increased resistance. A resistance-nodulation-cell division antibiotic efflux pump, predicted to target tetracycline was found in our *Helicobacter* MAG. Similar efflux pumps have been found in other species of *Helicobacter* (63). A nitroimidazole reductase gene was also identified in our *JAHZGC01* MAG. Similar genes can be found in a wide variety of bacterial taxa (64), but to our knowledge this is the first time this gene has been identified in this taxon.

The main weakness of our study is the relatively small number of MAGs which we were able to construct in comparison to studies in other wild avians (11, 65, 66). However, it is worth noting that the MAG quality cut-offs used for these studies were often substantially lower than in our study (e.g. minimum 50% completeness). We chose to use a higher cut-off of 80% completeness as we were predominantly interested in predicting the metabolism of these bacterial species, and lower genome completeness can have a significant effect on functional inference (67). The small number of taxa observed in our samples may be truly representative of the microbiota, with the community being dominated by a small number of more abundant bacteria. In general, the more abundant a species is in a sample then the

more likely it will be able to be constructed into a high-quality MAG. The high abundance of *Helicobacter* in our samples may have led to other species not being present in a high enough relative abundance to be constructed into MAGs. The low number of MAGs may also be due to reduced sequencing depth, due to the large amount of reads that belonged to host (89%).This percentage is similar to that found in saliva, throat, and vaginal swabs (68), but is unusual for the caeca or large intestine. One possible methodological explanation for this high degree of contamination could be how samples were stored. In order to enhance ease of sampling in the field and prevent spoiling of the carcass, intestines were initially stored whole, then defrosted and the contents removed for DNA extraction. This may have led to increased lysis of host cells, and thereby increased host DNA in our samples.

Conclusion:

This is the first study to reconstruct microbial genomes from the caeca of red grouse. The red grouse caeca harbours diverse microbial taxa, many of which are novel. These taxa have the potential to ferment various forms of fibre, thereby contributing to the digestion of the red grouse's high fibre diet. Further studies are needed to understand how these bacteria contribute to grouse nutrition and health.

List of abbreviations:

AMR: Antimicrobial resistance

ANI: Average nucleotide identity

Bp: Base pairs

CAZymes Carbohydrate active enzymes

GH: Glycoside hydrolases

GTDB: Genome taxonomy database

MAG: Metagenome assembled genome

rRNA: ribosomal RNA

Declarations

Ethics approval and consent to participate:

Red grouse were not killed for the purpose of providing samples for this study. Samples were collected from red grouse killed during legal game bird shoots within the open season in Scotland, from gastrointestinal samples that would normally be discarded as waste material. Therefore, approval for this study from an animal ethics committee was not required.

Consent for publication:

Not applicable

Availability of data and materials:

The datasets generated and analysed during the current study (fastq files, assemblies, genome bins and metagenome assembled genomes) are available in the European Nucleotide Archive under the project number PRJEB85541.

Competing interests:

The authors declare that they have no competing interests.

Funding:

This work was supported by the Biotechnology and Biological Sciences Research Council Institute through Strategic Program Programme Grant funding (BBS/E/RL/230001A). Laura

Glendinning is supported by a University of Edinburgh Chancellor's Fellowship. As a UK registered charity, the Game and Wildlife Conservation Trust relies mainly on fundraising from the private sector. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

Authors' contributions:

The authors confirm contribution to the paper as follows: study conception and design - LG; data collection - AA, NH and KF; analysis and interpretation of results - AA and LG; writing – original draft - LG; writing – editing and review - AA, LG, NH and KF. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgements:

We would like to acknowledge the upland estate staff who provided the red grouse samples.

References:

1. Sun F, Chen J, Liu K, Tang M, Yang Y. The avian gut microbiota: Diversity, influencing factors, and future directions. Front Microbiol. 2022;**13**: 934272 doi: 10.3389/fmicb.2022.934272
2. Bodawatta KH, Hird SM, Grond K, Poulsen M, Jønsson KA. Avian gut microbiomes taking flight. Trends Microbiol. 2022;**30**: 268-80 doi: 10.1016/j.tim.2021.07.003
3. Videvall E, Strandh M, Engelbrecht A, Cloete S, Cornwallis CK. Measuring the gut microbiome in birds: Comparison of faecal and cloacal sampling. Mol Ecol Resour. 2018;**18**: 424-34 doi: 10.1111/1755-0998.12744
4. Stanley D, Geier MS, Chen H, Hughes RJ, Moore RJ. Comparison of fecal and cecal microbiotas reveals qualitative similarities but quantitative differences. BMC Microbiol. 2015;**15**: 51 doi: 10.1186/s12866-015-0388-6
5. Matheen MIA, R. GM, and Dudaniec RY. Dominant factors shaping the gut microbiota of wild birds. Emu. 2022;**122**: 255-68 doi: 10.1080/01584197.2022.2114088
6. Mahmood T, Guo Y. Dietary fiber and chicken microbiome interaction: Where will it lead to? Anim Nutr. 2020;**6**: 1-8 doi: 10.1016/j.aninu.2019.11.004
7. Józefiak D, Rutkowski A, Martin SA. Carbohydrate fermentation in the avian ceca: A review. Anim Feed Sci Technol. 2004;**113**: 1-15 doi: 10.1016/j.anifeedsci.2003.09.007
8. Hunt A, Al-Nakkash L, Lee AH, Smith HF. Phylogeny and herbivory are related to avian cecal size. Sci Rep. 2019;**9**: 4243 doi: 10.1038/s41598-019-40822-0
9. Moss R, Parkinson JA. The digestion of heather (*Calluna vulgaris*) by red grouse (*Lagopus lagopus scoticus*). Br J Nutr. 1972;**27**: 285-98 doi: 10.1079/BJN19720093
10. Moss R, and Trenholm IB. Food intake, digestibility and gut size in red grouse. Br Poult Sci. 1987;**28**: 81-9 doi: 10.1080/00071668708416939
11. Rasmussen JA, Chua PYS. Genome-resolving metagenomics reveals wild western capercaillies (*Tetrao urogallus*) as avian hosts for antibiotic-resistance bacteria and their

interactions with the gut-virome community. Microbiol Res. 2023;**271**: 127372 doi: 10.1016/j.micres.2023.127372

12. Wienemann T, Schmitt-Wagner D, Meuser K, Segelbacher G, Schink B, et al. The bacterial microbiota in the ceca of Capercaillie (*Tetrao urogallus*) differs between wild and captive birds. Syst Appl Microbiol. 2011;**34**: 542-51 doi: 10.1016/j.syapm.2011.06.003

13. Drovetski SV, O'Mahoney MJV, Matterson KO, Schmidt BK, Graves GR. Distinct microbiotas of anatomical gut regions display idiosyncratic seasonal variation in an avian folivore. Anim microbiome. 2019;**1**: 2 doi: 10.1186/s42523-019-0002-6

14. Kohl KD, Connelly JW, Dearing MD, Forbey JS. Microbial detoxification in the gut of a specialist avian herbivore, the Greater Sage-Grouse. FEMS Microbiol Lett. 2016;**363**: fnw144 doi: 10.1093/femsle/fnw144

15. Vaasjo E, Stothart MR, Black SR, Poissant J, Whiteside DP. The impact of management on the fecal microbiome of endangered greater sage-grouse (*Centrocercus urophasianus*) in a zoo-based conservation program. Conserv Physiol. 2024;**12**: coae052 doi: 10.1093/conphys/coae052

16. Drovetski SV, Schmidt BK, Lai JE, Gross MS, Hladik ML, et al. Exposure to crop production alters cecal prokaryotic microbiota, inflates virulome and resistome in wild prairie grouse. Environ Pollut. 2022;**306**: 119418 doi: 10.1016/j.envpol.2022.119418

17. Salgado-Flores A, Tveit AT, Wright A-D, Pope PB, Sundset MA. Characterization of the cecum microbiome from wild and captive rock ptarmigans indigenous to Arctic Norway. PLOS ONE. 2019;**14**: e0213503 doi: 10.1371/journal.pone.0213503

18. Durnova E, Karmanova E, Sorokina T, Mayorova K, Aksenov A. Bacterial diversity and composition in the internal organs of Taiga Bean Goose, Greater White-Fronted Goose and Willow Ptarmigan as a new tools in the Arctic biomonitoring system. Diversity. 2025;**17**: 101 doi: 10.3390/d17020101

19. Ingvaldsen EW, Østnes JE, Kleven O, Davey M, Fossøy F, Nilsen EB. Fecal DNA metabarcoding reveals seasonal and annual variation in willow ptarmigan diet. R Soc Open Sci. 2024;**11**: 231518 doi: 10.1098/rsos.231518

20. Dunn JC, Clegg SR. High prevalence of multi-drug resistant bacteria in faecal samples from UK passerine birds. bioRxiv. 2024: 2024.07.24.604896 doi: 10.1101/2024.07.24.604896

21. Fantham HB. Observations on the parasitic protozoa of the Red Grouse (*Lagopus scoticus*), with a note on the Grouse Fly. Proc Zool Soc Lond. 1910;**80**: 692-708 doi: 10.1111/j.1096-3642.1910.tb01911.x

22. FastQC: A quality control analysis tool for high throughput sequencing data. https://github.com/s-andrews/FastQC.

23. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;**32**: 3047-8 doi: 10.1093/bioinformatics/btw354

24. Chen S, Zhou Y, Chen Y, Gu J. fastp: An ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;**34**: i884-i90 doi: 10.1093/bioinformatics/bty560

25. Eastman DS, Jenkins D. Comparative food habits of red grouse in northeast Scotland, using fecal analysis. J Wildl Manag. 1970;**34**: 612-20 doi: 10.2307/3798870

26. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv 2013;**1303**: 3997 doi: 10.48550/arXiv.1303.3997

27. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;**25**: 2078-9 doi: 10.1093/bioinformatics/btp352

28. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 2015;**31**: 1674-6 doi: 10.1093/bioinformatics/btv033

29. Aroney ST, Newell RJ, Nissen JN, Camargo AP, Tyson GW, Woodcroft BJ. CoverM: Read alignment statistics for metagenomics. arXiv. 2025;**preprint**: 2501.11217 doi: 2501.11217

30. 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 doi: 10.7717/peerj.7359

31. 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**: 1203-12 doi: 10.1038/s41592-023-01940-w

32. 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. 2017;**11**: 2864-8 doi: 10.1038/ismej.2017.126

33. Guo J, Bolduc B, Zayed AA, Varsani A, Dominguez-Huerta G, et al. VirSorter2: A multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. Microbiome. 2021;**9**: 37 doi: 10.1186/s40168-020-00990-y

34. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: A toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics. 2019;**36**: 1925-7 doi: 10.1093/bioinformatics/btz848

35. Zhou Z, Tran PQ, Breister AM, Liu Y, Kieft K, et al. METABOLIC: High-throughput profiling of microbial genomes for functional traits, metabolism, biogeochemistry, and community-scale functional networks. Microbiome. 2022;**10**: 33 doi: 10.1186/s40168-021-01213-8

36. Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya Amogelang R, et al. CARD 2023: Expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. Nucleic Acids Res. 2022;**51**: D690-D9 doi: 10.1093/nar/gkac920

37. Wickham H. Reshaping data with the reshape package. J Stat Softw. 2007;**21**: 1 - 20 doi: 10.18637/jss.v021.i12

38. Wilke CO. cowplot: Streamlined plot theme and plot annotations for 'ggplot2'. 2024.

39. Wickham H, François R, Henry L, Müller K, Vaughan D. dplyr: A grammar of data manipulation. 2023.

40. Wickham H. ggplot2: Elegant graphics for data analysis: Springer-Verlag New York; 2016.

41. Wickham H, Averick M, Bryan J, Chang W, McGowan LDA, et al. Welcome to the Tidyverse. J Open Source Softw. 2019;**4**: 1686 doi: 10.21105/joss.01686

42. Asnicar F, Thomas AM, Beghini F, Mengoni C, Manara S, et al. Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. Nat Commun. 2020;**11**: 2500 doi: 10.1038/s41467-020-16366-7

43. Figtree. https://github.com/rambaut/figtree.

44. Mitchell AL, Scheremetjew M, Denise H, Potter S, Tarkowska A, et al. EBI Metagenomics in 2017: enriching the analysis of microbial communities, from sequence reads to assemblies. Nucleic Acids Res. 2017;**46**: D726-D35 doi: 10.1093/nar/gkx967

45. Xie F, Jin W, Si H, Yuan Y, Tao Y, et al. An integrated gene catalog and over 10,000 metagenome-assembled genomes from the gastrointestinal microbiome of ruminants. Microbiome. 2021;**9**: 137 doi: 10.1186/s40168-021-01078-x

46. Bjørnsen MB, Valerón NR, Vásquez DP, Velasco EM, Hansen AJ, Hauptmann AL. Microbiota in the ptarmigan intestine—An Inuit delicacy and its potential in popular cuisine. PLOS ONE. 2024;**19**: e0305317 doi: 10.1371/journal.pone.0305317

47. Gilroy R, Ravi A, Getino M, Pursley I, Horton DL, et al. Extensive microbial diversity within the chicken gut microbiome revealed by metagenomics and culture. PeerJ. 2021;**9**: e10941 doi: 10.7717/peerj.10941

48. Beresford-Jones BS, Forster SC, Stares MD, Notley G, Viciani E, et al. The Mouse Gastrointestinal Bacteria Catalogue enables translation between the mouse and human gut microbiotas via functional mapping. Cell Host Microbe. 2022;**30**: 124-38.e8 doi: 10.1016/j.chom.2021.12.003

49. Rosshart SP, Herz J, Vassallo BG, Hunter A, Wall MK, et al. Laboratory mice born to wild mice have natural microbiota and model human immune responses. Science. 2019;**365**: eaaw4361 doi: doi:10.1126/science.aaw4361

50. Lesker TR, Durairaj AC, Gálvez EJC, Lagkouvardos I, Baines JF, et al. An integrated metagenome catalog reveals new insights into the murine gut microbiome. Cell Rep. 2020;**30**: 2909-22.e6 doi: 10.1016/j.celrep.2020.02.036

51. Riva A, Sahin E, Volpedo G, Petretto A, Lavarello C, et al. Identification of an epilepsy-linked gut microbiota signature in a pediatric rat model of acquired epilepsy. Neurobiol Dis. 2024;**194**: 106469 doi: 10.1016/j.nbd.2024.106469

52. Zheng J, Sun Q, Zhang M, Liu C, Su Q, et al. Noninvasive, microbiome-based diagnosis of inflammatory bowel disease. Nat Med. 2024;**30**: 3555-67 doi: 10.1038/s41591-024-03280-4

53. Blyton MDJ, Soo RM, Hugenholtz P, Moore BD. Maternal inheritance of the koala gut microbiome and its compositional and functional maturation during juvenile development. Environ Microbiol. 2022;**24**: 475-93 doi: 10.1111/1462-2920.15858

54. Mukherjee A, Lordan C, Ross RP, Cotter PD. Gut microbes from the phylogenetically diverse genus *Eubacterium* and their various contributions to gut health. Gut Microbes. 2020;**12**: 1802866 doi: 10.1080/19490976.2020.1802866

55. Kim CY, Lee M, Yang S, Kim K, Yong D, et al. Human reference gut microbiome catalog including newly assembled genomes from under-represented Asian metagenomes. Genome Med. 2021;**13**: 134 doi: 10.1186/s13073-021-00950-7

56. Wylensek D, Hitch TCA, Riedel T, Afrizal A, Kumar N, et al. A collection of bacterial isolates from the pig intestine reveals functional and taxonomic diversity. Nat Commun. 2020;**11**: 6389 doi: 10.1038/s41467-020-19929-w

57. IJdema F, De Smet J, Crauwels S, Lievens B, Van Campenhout L. Meta-analysis of larvae of the black soldier fly (*Hermetia illucens*) microbiota based on 16S rRNA gene amplicon sequencing. FEMS Microbiol Ecol. 2022;**98**: fiac094 doi: 10.1093/femsec/fiac094

58. Salgado JFM, Hervé V, Vera MAG, Tokuda G, Brune A. Unveiling lignocellulolytic potential: a genomic exploration of bacterial lineages within the termite gut. Microbiome. 2024;**12**: 201 doi: 10.1186/s40168-024-01917-7

59. Bowerman KL, Knowles SCL, Bradley JE, Baltrūnaitė L, Lynch MDJ, et al. Effects of laboratory domestication on the rodent gut microbiome. ISME Commun. 2021;**1**: 49 doi: 10.1038/s43705-021-00053-9

60. Betancur-Murillo CL, Aguilar-Marín SB, Jovel J. *Prevotella*: A key player in ruminal metabolism. 2022;**11**: 1 doi: 10.3390/microorganisms11010001

61. Murray AH, Iason GR, Stewart C. Effect of simple phenolic compounds of heather (*Calluna vulgaris*) on rumen microbial activity *in vitro*. J Chem Ecol. 1996;**22**: 1493-504 doi: 10.1007/bf02027727

62. Stogios PJ, Savchenko A. Molecular mechanisms of vancomycin resistance. Protein Sci. 2020;**29**: 654-69 doi: 10.1002/pro.3819

63. Li J, Li X-Z. Antimicrobial resistance and drug efflux pumps in *Helicobacter*. In: Li X-Z, Elkins CA, Zgurskaya HI, editors. Efflux-mediated antimicrobial resistance in bacteria: Mechanisms, regulation and clinical implications: Springer International Publishing; 2016. p. 489-513.

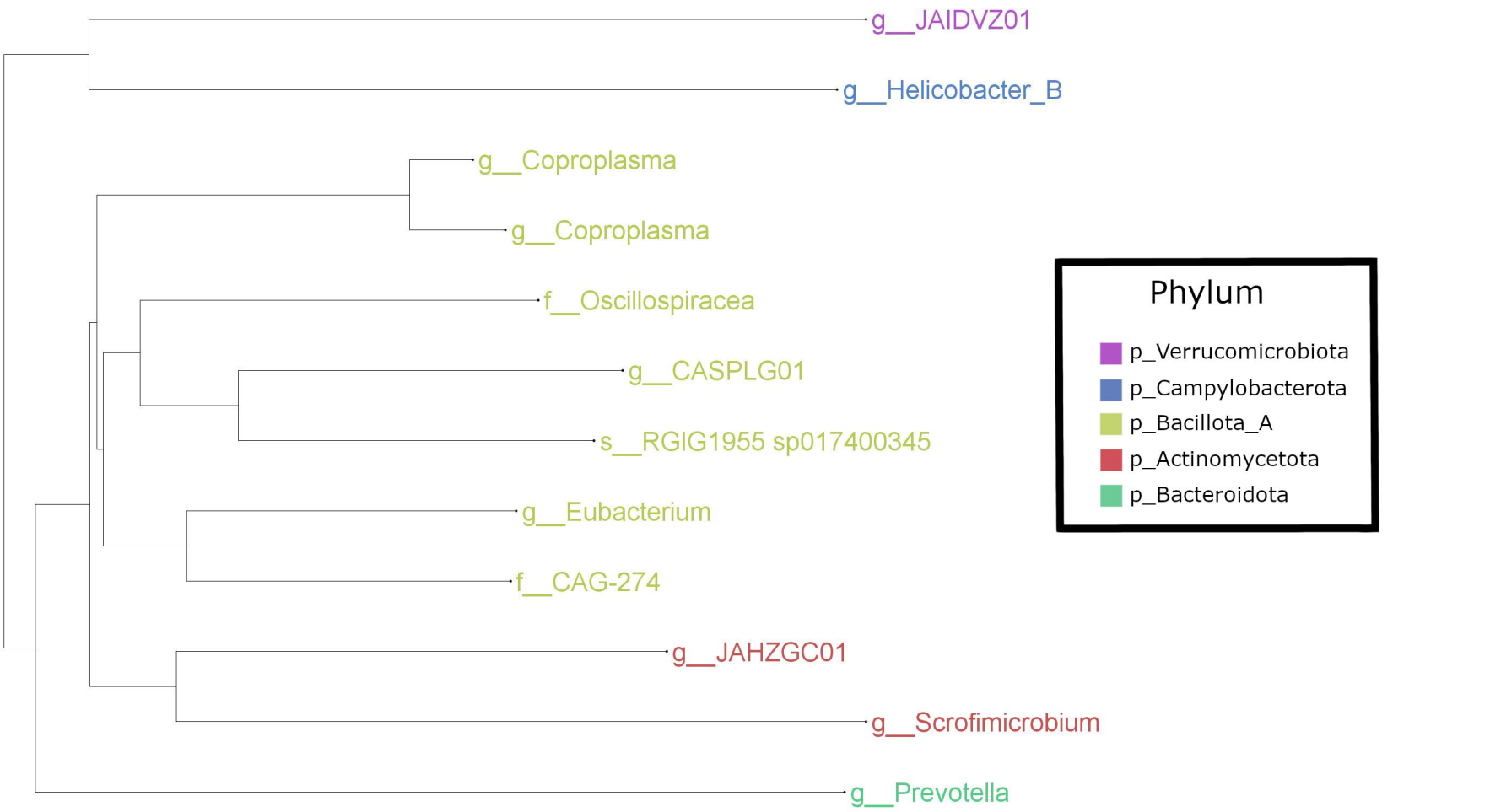
64. Alauzet C, Lozniewski A, Marchandin H. Metronidazole resistance and nim genes in anaerobes: A review. Anaerobe. 2019;**55**: 40-53 doi: 10.1016/j.anaerobe.2018.10.004

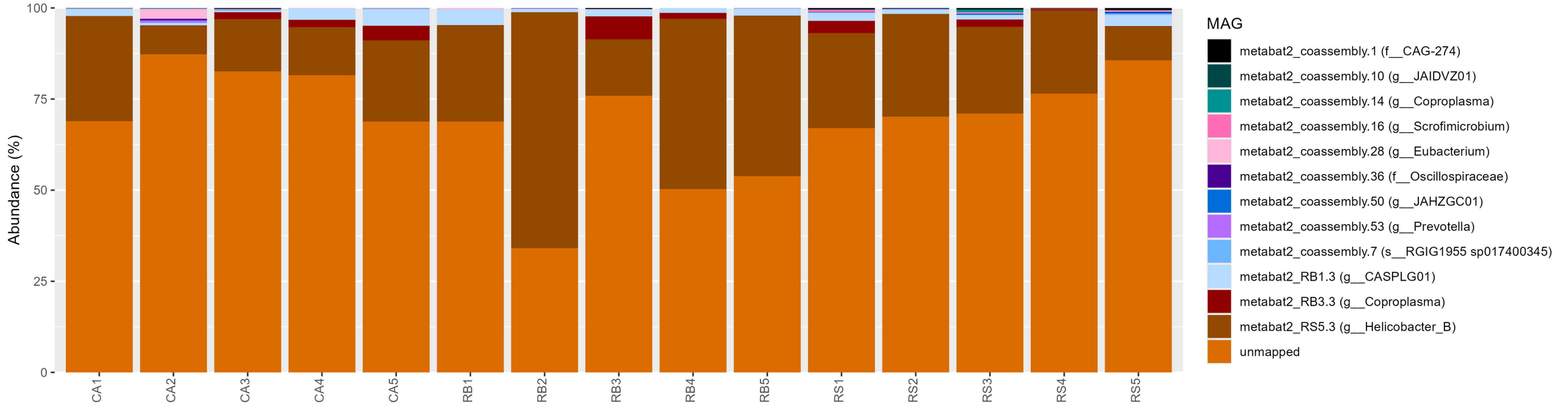
65. Tang B, Wang Y, Dong Y, Cui Q, Zeng Z, et al. The catalog of microbial genes and metagenome-assembled genomes from the gut microbiomes of five typical crow species on the Qinghai–Tibetan Plateau. Microorganisms. 2024;**12**: 2033 doi: 10.3390/microorganisms12102033

66. Wang Y, Zhai J, Tang B, Dong Y, Sun S, et al. Metagenomic comparison of gut communities between wild and captive Himalayan griffons. Front vet sci. 2024;**11**: 1403932 doi: 10.3389/fvets.2024.1403932

67. Eisenhofer R, Odriozola I, Alberdi A. Impact of microbial genome completeness on metagenomic functional inference. ISME Commun. 2023;**3**: 12 doi: 10.1038/s43705-023-00221-z

68. Gao Y, Luo H, Lyu H, Yang H, Yousuf S, et al. Benchmarking short-read metagenomics tools for removing host contamination. GigaScience. 2025;**14**: giaf004 doi: 10.1093/gigascience/giaf004


g__JAIDVZ01
g__Helicobacter_B
g__Coproplasma
g__Coproplasma
f__Oscillospiracea
g__CASPLG01
s__RGIG1955 sp017400345
g__Eubacterium
f__CAG-274
g__JAHZGC01
g__Scrofimicrobium
g__Prevotella
Phylum
p_Verrucomicrobiota
p_Campylobacterota
p_Bacillota_A
p_Actinomycetota
p_Bacteroidota


Abundance (%)
100
75
50
25
0
CA1
CA2
CA3
CA4
CA5
RB1
RB2
RB3
RB4
RB5
RS1
RS2
RS3
RS4
RS5
MAG
metabat2_coassembly.1 (f__CAG-274)
metabat2_coassembly.10 (g__JAIDVZ01)
metabat2_coassembly.14 (g__Coproplasma)
metabat2_coassembly.16 (g__Scrofimicrobium)
metabat2_coassembly.28 (g__Eubacterium)
metabat2_coassembly.36 (f__Oscillospiraceae)
metabat2_coassembly.50 (g__JAHZGC01)
metabat2_coassembly.53 (g__Prevotella)
metabat2_coassembly.7 (s__RGIG1955 sp017400345)
metabat2_RB1.3 (g__CASPLG01)
metabat2_RB3.3 (g__Coproplasma)
metabat2_RS5.3 (g__Helicobacter_B)
unmapped