

4 Anum Ali Ahmad¹, Kathy Fletcher², Nicholas Hesford², Laura Glendinning^{1*}

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

10 *Corresponding author: Laura.glendinginning@roslin.ed.ac.uk

11 **Abstract:**

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

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

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

31

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

34 **Background:**

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

40 An important challenge when studying wild species is acquiring samples that are truly
41 representative of the gut microbiota. Studies that use wild birds are often limited to collecting
42 non-invasive samples, such as faeces and caecal swabs. However, these samples do not
43 quantitatively represent the microbiota of the large and small intestine (3, 4). Obtaining
44 samples directly from the gastrointestinal section of interest is therefore advantageous, but is
45 often not possible as it requires killing the bird. Game bird species offer a potential avenue
46 for studying the microbiota in “wild” or minimally-managed avian populations, where
47 gastrointestinal samples can be obtained in an ethical manner. Game birds are regularly
48 shot during hunts, or to collect samples to monitor population health. The gastrointestinal
49 tracts from these birds are usually discarded and can therefore be seen as a waste product
50 of this industry. This allows researchers to access gastrointestinal samples without requiring
51 additional animal sacrifice.

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

60 Red grouse (*Lagopus scotica*) represent a particularly interesting game bird species in which
 61 to study microbiota fibre fermentation. The main component of the adult red grouse diet is
 62 heather (*Calluna vulgaris*), which is high in fibre (25%) and low in protein (7%) (9). Red
 63 grouse have long caeca (~75 cm), which increase in size when they are supplied with a
 64 higher fibre diet (10). From feed digestibility studies, there is some evidence that the red
 65 grouse microbiota is able to digest a wide variety of fibres, including those that are generally
 66 considered more resistant to fermentation (e.g. lignin) (9).

67 There is also increasing interest in studying antimicrobial resistance in wild-living animal
 68 populations through a one health lens (11). As game birds are frequently consumed by
 69 humans as food, the presence of AMR in these birds is of interest from this perspective.
 70 Previous studies have identified AMR in gut-associated microbes in UK pheasants and
 71 partridges (12, 13), but this has not been examined in red grouse.

72 The gut microbiota has been studied in several grouse species using 16S rRNA gene
 73 sequencing and/or metagenomic sequencing, including in capercaillies (14, 15), greater
 74 sage grouse (16-18), wild prairie grouse (19) and ptarmigans (20-22). However, while
 75 bacteria have been cultured from the gastrointestinal tracts of red grouse (23, 24), their
 76 microbiota has not been characterised using these modern sequencing methods. In this
 77 study, we perform metagenomic sequencing analysis of red grouse caecal content samples.
 78 We construct twelve high-quality bacterial metagenome assembled genomes, and
 79 characterise their functional potential to metabolise fibre and their AMR potential.

80 **Methods:**

81 Sample collection and processing

82 Samples were collected from red grouse killed during legal game bird shoots within the open
 83 season in three upland estates in Scotland. Samples were collected in the 2023 Scottish red
 84 grouse (*Lagopus scotica*) shooting season (August-October), at three distinct areas of
 85 moorland managed for recreational hunting of red grouse (Latitude and longitude: 55.75 -
 86 3.01 (Samples RS1-5), 55.98 -3.33 (Samples RB1-5), and 57.02 -4.12 (Samples CA1-CA5)).
 87 Samples were collected from five birds per area. The intestines were removed from birds
 88 and frozen on site using a standard domestic freezer within 24 hours. Samples were then
 89 transported to the Roslin Institute within one month of sample collection, and stored at -80°C
 90 until DNA extraction. DNA was extracted from caecal contents using the QIAamp
 91 PowerFecal Pro DNA Kit according to the manufacturer's instructions, using 250-300 mg of
 92 caecal contents. DNA quality and quantity were assessed using a NanoDrop
 93 spectrophotometer and a Qubit Fluorometer with a Qubit dsDNA Quantification Assay Kit
 94 (broad range). Shotgun metagenomic library preparation and sequencing was performed by
 95 Novogene, using an Illumina NovaSeq 6000 producing 150 bp paired-end reads. Samples
 96 RS1 and RS4 were produced on two separate sequencing runs in order to produce sufficient
 97 data quantity, then fastq files from the same sample were concatenated prior to analysis.

98

99 Bioinformatic analysis

100 Sequences were checked for quality using FastQC (v.0.12.1) (25) and MultiQC (v.1.6) (26).
 101 Quality control was performed using fastp (v.0.24) (27), with the options "--
 102 detect_adapter_for_pe -l 30 -q 20 --trim_poly_g --trim_poly_x". The host genome and feed
 103 genome were removed from adaptor trimmed read files. There is currently no RefSeq or
 104 GenBank genome available for *Lagopus scotica*, therefore the *Lagopus muta* (rock
 105 ptarmigan) genome was downloaded from RefSeq (GCF_023343835.1) and used instead of

106 a host genome. Heather (*Calluna vulgaris*) is the main food source for adult red grouse (28);
 107 therefore the *Calluna vulgaris* genome was downloaded from GenBank (GCA_964145215.1).
 108 Sample reads were mapped to host and feed reference genomes using BWA-MEM
 109 (v.0.7.18) (29), followed by SAMtools (v.1.21, samtools fastq -f 12) (30) to select reads
 110 where both paired-end reads were unmapped. The unmapped reads were taxonomically
 111 classified using Kraken2 (v.2.3.1) (31) with a custom database built using *Lagopus muta* and
 112 bacterial, fungal, archaeal, viral, plant, plasmid, and protozoal RefSeq genomes (built on 23rd
 113 June 2025). Bracken (v.3.0.1) (32) was then used to estimate their relative abundances
 114 (options: -k 35 -l 150). Single-sample assemblies and coassemblies were constructed using
 115 Megahit (v.1.2.9) (33).

116 CoverM (v.0.7) (34) was used to create depth files, which were used as input for genome
 117 binning by MetaBAT 2 (v.2.17) (35). Contamination and completeness of bins were
 118 estimated using CheckM2 (v.1.0.2) (36). dRep (v.3.5.0) (37) was used to dereplicate
 119 genome bins into MAGs, with minimum completeness set at 80% and maximum
 120 contamination at 10%. Genomes were dereplicated at 95% ANI for species-level MAGs, and
 121 99% ANI for strain-level MAGs. Viruses were identified from megahit assemblies using
 122 VirSorter2 (v.2.2.4: database downloaded 3rd February 2025) (38) (options: --include-groups
 123 dsDNAphage,ssDNA --min-length 5000 --min-score 0.5 -j 28 all).

124 GTDB-Tk (v.2.4.0) (39) was used to assign taxonomy to MAGs (database downloaded 24th
 125 January 2025). METABOLIC was used to assign metabolic function to MAGs (v.4.0) (40).
 126 Antimicrobial resistance (AMR) genes present in MAGs and megahit assemblies were
 127 identified using RGI main (v. 6.0.3) (41) with the Comprehensive Antibiotic Resistance
 128 Database (database downloaded 17th March 2025). Loose hits were included, and the --
 129 low_quality option was used for megahit assemblies. The abundance of MAGs within
 130 samples was calculated using CoverM (options: -p bwa-mem -m mean relative_abundance
 131 trimmed_mean covered_bases variance length reads_per_base count rpkm tpm --min-read-
 132 aligned-percent 75 --min-read-percent-identity 95 --min-covered-fraction 0). Graphs and data

133 summaries were produced in R using the packages reshape2 (v.1.4.4) (42), cowplot
 134 (v.1.1.3) (43), dplyr (v.1.1.4) (44), ggplot2 (v.3.5.1) (45) and tidyverse (v.2.0.0) (46). MAG
 135 taxonomic trees were constructed using PhyloPhlAn (v.3.1.1) (47) (options: --diversity high -f
 136 supermatrix_aa.cfg --subsample phylophlan --fast). Taxonomic trees were midpoint rerooted
 137 and visualised using figtree (v.1.4.3) (48).

138

139 **Results:**

140 Sequence quality

141 We performed metagenomic sequencing on caecal content samples from fifteen Scottish red
142 grouse. After removal of poor-quality sequences, and trimming of adaptors, polyAs and
143 polyGs, there were on average 102,686,250 ± 24,122,717 reads per sample. After removal
144 of host and feed sequences the average number of reads per sample dropped by 89%, to
145 10,902,109 ± 7,733,350. This was predominantly due to the removal of host reads, indicating
146 that our samples had a large proportion of host contamination. From single sample
147 assemblies and a coassembly, we constructed 161 putative genome bins. Quality-filtering
148 and dereplication resulted in twelve high-quality (<10% contamination, >80% complete)
149 species-level MAGs (**Table 1**). We also constructed strain-level MAGs, but as we found only
150 one more MAG than at species-level we decided to restrict our further analyses to species
151 level. Eighteen high quality viral genomes were identified (**Table S1**), two of which were
152 proviruses.

153

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

MAG	GTDB taxonomy	Complete ness (%)	Contamin ation (%)	Contig N50	Genome size (bp)	Total coding sequences	GC (%)
metabat 2_RB1.3	d__Bacteria;p __Bacillota_A ;c__Clostridia ;o__Oscillospi rales;f__Acut alibacteracea	100	0.17	167,26 2	1,021,94 3	930	33

	e;g__CASPL G01;s__						
metabat 2_RB3.3	d__Bacteria;p __Bacillota_A ;c__Clostridia ;o__Christens enellales;f__ Borkfalkiaceae e;g__Copropl asma;s__	97.91	0	54,893	1,424,43 0	1,267	49
metabat 2_RS5.3	d__Bacteria;p __Campyloba cterota;c__Ca mpylobacteria ;o__Campylo bacterales;f_ __Helicobacter aceae;g__Hel icobacter_B;s __	99.78	0.01	27,466	1,541,84 0	1,592	37
metabat 2_coass embly.1	d__Bacteria;p __Bacillota_A ;c__Clostridia ;o__Lachnos pirales;f__CA	98.34	0.27	28,654	1,333,03 4	1,241	28

	G- 274;g__s__						
metabat 2_coass embly.10	d__Bacteria;p __Verrucomi robiota;c__Ve rrucomicrobia e;o__Opituta les;f__LL51;g __JAIDVZ01; s__	83.77	0.37	8,043	804,428	898	32
metabat 2_coass embly.14	d__Bacteria;p __Bacillota_A ;c__Clostridia ;o__Christens enellales;f__ Borkfalkiacea e;g__Copropl asma;s__	88.08	0.05	6,612	1,044,10 0	1,098	52
metabat 2_coass embly.16	d__Bacteria;p __Actinomy etota;c__Acti nomycetes;o_ _Actinomyces ales;f__Actin omycetaceae; g__Scrofimicr	96.84	0	30,426	1,474,56 2	1,248	50

	obium;s__						
metabat 2_coass embly.28	d__Bacteria;p __Bacillota_A ;c__Clostridia ;o__Lachnos pirales;f__Lac hnospiraceae ;g__Eubacteri um_l;s__	100	0.02	13,892	2,419,85 2	2,339	48
metabat 2_coass embly.36	d__Bacteria;p __Bacillota_A ;c__Clostridia ;o__Oscillospi rales;f__Oscil lospiraceae;g __;s__	100	0.06	20,806	2,734,68 8	2,460	57
metabat 2_coass embly.50	d__Bacteria;p __Actinomyce tota;c__Cori obacteriia;o_ __Coriobacteri ales;f__UMG S124;g__JAH ZGC01;s__	92.48	3.21	11,155	2,307,05 2	1994	68

metabat	d__Bacteria;p	81.06	1.36	8,397	2,145,10	1,915	56
2_coass	__Bacteroidot				4		
embly.53	a;c__Bacteroi						
	dia;o__Bacter						
	oidales;f__Ba						
	cteroidaceae;						
	g__Prevotella						
	;s__						
metabat	d__Bacteria;p	97.55	0.04	73,819	871,746	975	35
2_coass	__Bacillota_A						
embly.7	;c__Clostridia						
	;o__Oscillospi						
	rales;f__Rumi						
	nococcaceae;						
	g__RGIG195						
	5;s__RGIG19						
	55						
	sp017400345						

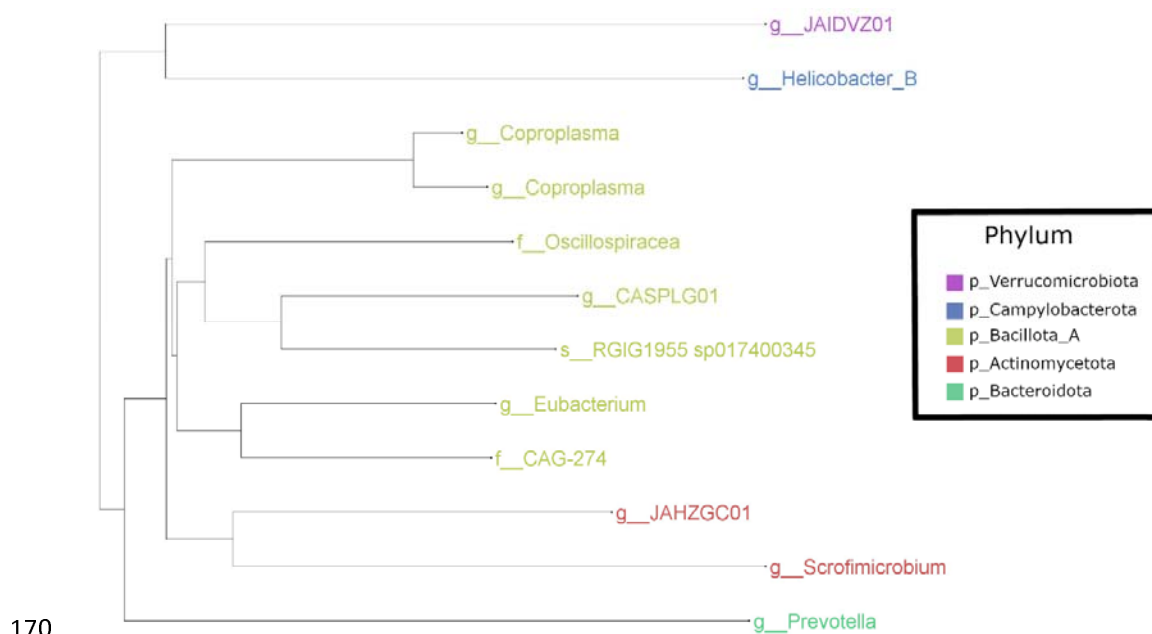
155

156

157 Taxonomy of MAGs

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

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



170
171 **Figure 1: Taxonomic tree of species-level, high-quality metagenome assembled**
172 **genomes constructed from red grouse caecal metagenomic data. Tip labels represent**
173 **the lowest taxonomy assigned to each MAG by GTDB-Tk.**

174
175 For most samples, the majority of reads did not map to MAGs ($69.5 \pm 14.3\%$ relative
176 abundance) (**Figure 2**). By far the most abundant MAG was metabat2_RS5.3 (genus
177 *Helicobacter_B*) ($26.3 \pm 15.4\%$), followed by metabat2_RB1.3 (genus *CASPLG01*) ($1.9 \pm$
178 1.3%) and metabat2_RB3.3 (genus *Coproplasma*) ($1.4 \pm 1.9\%$). Some samples contained
179 particularly high levels of *Helicobacter_B*, including RB2 (64.5%), RB4 (46.7%) and RB5
180 (44.0%).

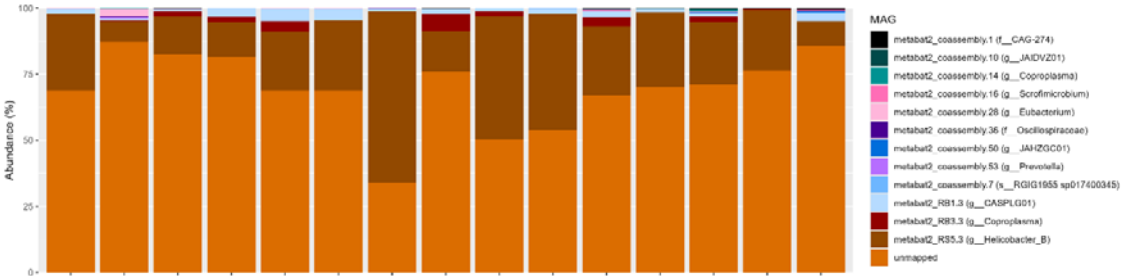


Figure 2: Relative abundance of species-level, high-quality metagenome assembled genomes in red grouse caecal samples. Taxonomies represent the lowest taxonomy assigned to each MAG by GTDB-Tk.

Metabolism of MAGs

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*, Oscillospiraceae 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*.

Antimicrobial resistance genes

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*).

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

236

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

MAG	Taxonomy	Antibiotic	Gene	% Identity	Resistance Mechanism	AMR Gene Family
metabat2_coassembly.1	f__CAG-274	vancomycin; teicoplanin	vanW	30.04	antibiotic target alteration	vanW; glycopeptide resistance gene cluster
metabat2_coassembly.1	f__CAG-274	vancomycin	vanG	41.88	antibiotic target alteration	glycopeptide resistance gene cluster; Van ligase

metabat2_ coassembly.1	f__CAG- 274	vancomycin	vanT	36.04	antibiotic target alteration	glycopeptide resistance gene cluster; vanT
metabat2_ coassembly.14	g__Copro plasma	vancomycin; teicoplanin	vanW	30.75	antibiotic target alteration	vanW; glycopeptide resistance gene cluster
metabat2_ coassembly.28	g__Eubact erium	vancomycin; teicoplanin	vanW	32.82	antibiotic target alteration	vanW; glycopeptide resistance gene cluster
metabat2_ coassembly.28	g__Eubact erium	vancomycin	vanT	38.24	antibiotic target alteration	glycopeptide resistance gene cluster; vanT
metabat2_ coassembly.36	f__Oscillo spiraceae	vancomycin; teicoplanin	vanW	39.9	antibiotic target alteration	vanW; glycopeptide

						resistance gene cluster
metabat2_ coassembl y.36	f__Oscillo spiraceae	vancomyci n	vanT	35.22	antibiotic target alteration	glycopepti de resistance gene cluster; vanT
metabat2_ coassembl y.36	f__Oscillo spiraceae	vancomyci n	vanY	61.62	antibiotic target alteration	vanY; glycopepti de resistance gene cluster
metabat2_ coassembl y.36	f__Oscillo spiraceae	vancomyci n; teicoplanin	vanW	39.59	antibiotic target alteration	vanW; glycopepti de resistance gene cluster
metabat2_ coassembl y.36	f__Oscillo spiraceae	vancomyci n	vanG	39.94	antibiotic target alteration	glycopepti de resistance gene cluster; Van ligase

metabat2_ coassembl y.36	f__Oscillo spiraceae	vancomyci n	vanT	48.37	antibiotic target alteration	glycopepti de resistance gene cluster; vanT
metabat2_ coassembl y.36	f__Oscillo spiraceae	vancomyci n	vanG	53.78	antibiotic target alteration	glycopepti de resistance gene cluster; Van ligase
metabat2_ coassembl y.36	f__Oscillo spiraceae	vancomyci n	vanW	56.25	antibiotic target alteration	vanW; glycopepti de resistance gene cluster
metabat2_ coassembl y.50	g__JAHZ GC01	vancomyci n	vanXY	41.85	antibiotic target alteration	glycopepti de resistance gene cluster; vanXY
metabat2_ coassembl y.50	g__JAHZ GC01	vancomyci n	vanG	40.51	antibiotic target alteration	glycopepti de resistance

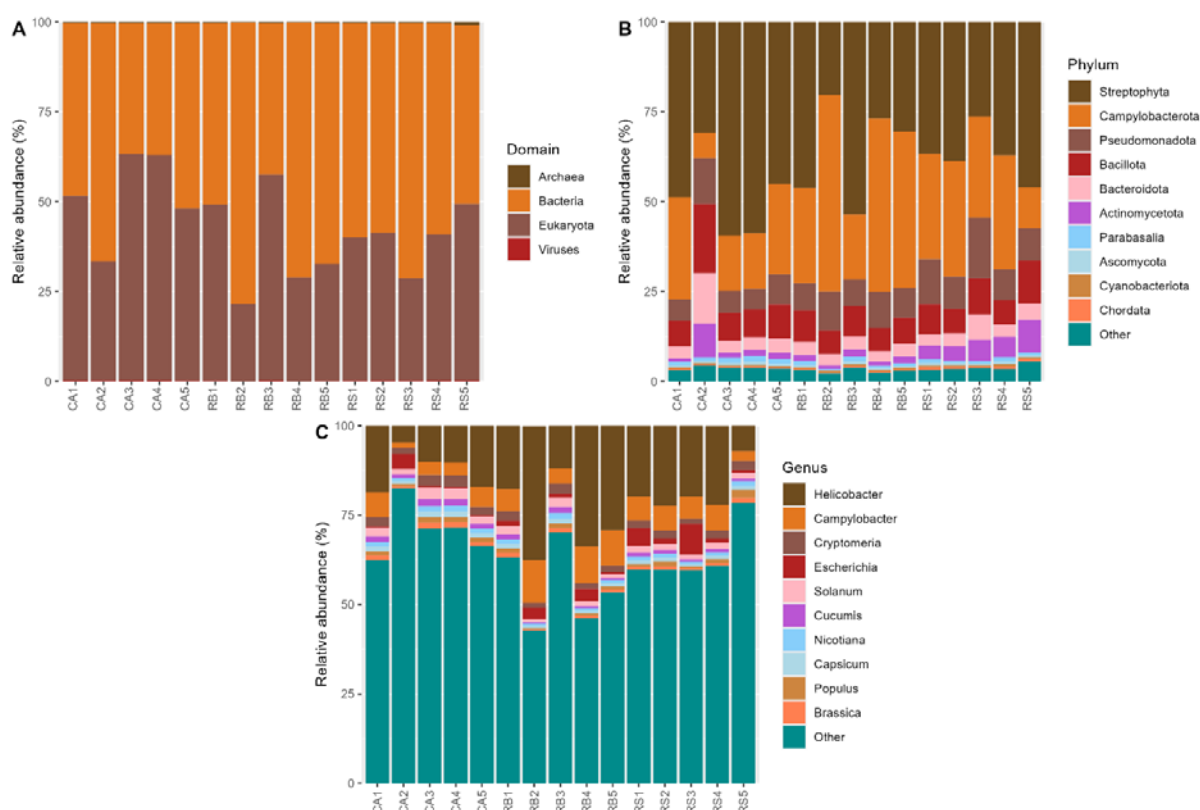
						gene cluster; Van ligase
metabat2_ coassembly.50	g__JAHZ GC01	vancomycin	vanT	40.51	antibiotic target alteration	glycopeptide resistance gene cluster; vanT
metabat2_ coassembly.50	g__JAHZ GC01	metronidazole	nimJ	49.07	antibiotic inactivation	nitroimidazole reductase
metabat2_ coassembly.50	g__JAHZ GC01	vancomycin	vanH	43.71	antibiotic target alteration	vanH; glycopeptide resistance gene cluster
metabat2_ coassembly.53	g__Prevotella	vancomycin	vanT	36.98	antibiotic target alteration	glycopeptide resistance gene cluster; vanT
metabat2_ coassembly.7	s__RGIG1 955 sp017400	vancomycin	vanY	37.74	antibiotic target alteration	vanY; glycopeptide

	345					resistance gene cluster
metabat2_ coassembl y.7	s__RGIG1 955 sp017400 345	vancomyci n	vanG	61.21	antibiotic target alteration	glycopepti de resistance gene cluster; Van ligase
metabat2_ coassembl y.7	s__RGIG1 955 sp017400 345	vancomyci n	vanXY	53.66	antibiotic target alteration	glycopepti de resistance gene cluster; vanXY
metabat2_ coassembl y.7	s__RGIG1 955 sp017400 345	vancomyci n	vanT	53.12	antibiotic target alteration	glycopepti de resistance gene cluster; vanT
metabat2_ RB1.3	g__CASP LG01	vancomyci n	vanT	37.67	antibiotic target alteration	glycopepti de resistance gene cluster; vanT

metabat2_ RB1.3	g__CASP LG01	vancomycin	vanG	41.13	antibiotic target alteration	glycopeptide resistance gene cluster; Van ligase
metabat2_ RB1.3	g__CASP LG01	vancomycin	vanY	36.21	antibiotic target alteration	vanY; glycopeptide resistance gene cluster
metabat2_ RB3.3	g__Copro plasma	vancomycin; teicoplanin	vanW	28.37	antibiotic target alteration	vanW; glycopeptide resistance gene cluster
metabat2_ RS5.3	g__Helico bacter	Tetracycline	adeF	43.68	antibiotic efflux	resistance - nodulation -cell division (RND) antibiotic efflux pump

Taxonomic classification of reads

As MAGs normally only capture the most abundant bacterial and archaeal taxa in samples, we used Kraken2/Bracken to taxonomically classify our metagenomic reads and characterise taxonomies that may not have been captured by MAGs (**Figure 3**). On average, $85.0 \pm 1.63\%$ metagenomic reads remained unclassified in our data. The classified reads were assigned to 4 domains, 82 phyla, and 2,632 genera across all samples. At the domain level, the reads were dominated by bacteria ($56.3 \pm 12.8\%$, relative abundance) and Eukaryota ($43.2 \pm 12.7\%$) across all samples (**Figure 3A**). Streptophyta ($40.3 \pm 12.0\%$) and Campylobacterota ($27.7 \pm 13.4\%$) were dominant phyla (**Figure 3B**). Other minor phyla included Pseudomonadota ($9.27 \pm 3.01\%$), Bacillota ($8.87 \pm 3.25\%$), Bacteroidota ($4.41 \pm 2.84\%$), and Actinomycetota ($3.49 \pm 2.87\%$). At the genus level, *Helicobacter* ($18.8 \pm 9.42\%$) and *Campylobacter* ($6.21 \pm 2.89\%$) were highly abundant across all samples (**Figure 3C**). Other minor genera included *Cryptomeria* ($2.28 \pm 0.57\%$), *Escherichia* ($2.18 \pm 2.35\%$), *Solanum* ($1.88 \pm 0.69\%$), and *Cucumis* ($1.14 \pm 0.52\%$).



253 **Figure 3: Barplot showing Kracken2/Bracken classified metagenomic reads at the**
254 **domain (A), phylum (B), and genus (C) levels. The top 10 highly abundant taxa are**
255 **displayed, while the remaining taxa are grouped as “Other”.**

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 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.

Taxonomic profiling using Kraken2/Bracken revealed that the majority of our metagenomic reads remained unclassified. This is likely due to the fact that the gut microbiota of red grouse has not previously been examined using metagenomic sequencing methods, so red grouse-specific members of the gut microbiota will not currently be included in genomic databases. Similar results have been reported for other wild bird species (49-51), demonstrating the understudied nature of the microbiota of these species. A large proportion of our reads originated from plants, likely from the bird's diet. Due to our quality control methods these reads are unlikely to have impacted our microbiota analysis, but these results highlight the possibility of using gut metagenomics to study the diets of wild birds. The results from our Kraken2/Bracken analysis should be interpreted cautiously. As members of the red grouse microbiota are unlikely to be represented in the RefSeq database, they are unlikely to be present in the Kraken2 database used in our analysis. Misassignment of taxonomies is therefore likely, particularly at lower taxonomic levels.

We constructed MAGs to gain a deeper understanding of the microbial community composition and its functional potential, and to begin to address the lack of red grouse microbiota representatives in public sequence databases. 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 (52, 53). A MAG from the genus *Helicobacter*_B was the most abundant MAG in our samples, comprising on average 26% of the microbiota. Kraken2/Bracken analysis also identified this genus to be highly abundant in our metagenomic reads. *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 (16, 21). In ptarmigan intestines, Bjørnsen *et al.* found similar average abundances of *Helicobacter* as observed in our study ($24\% \pm 3$) (54).

The majority of our MAGs were identified as Clostridia, which is a commonly identified taxon in the guts of other grouse species (16, 17, 20). Of these MAGs, most are from poorly described taxa identified from metagenomic data: the genus *Coproplasma* has been identified in the chicken gut (55), the genus *CASPLG01* has been identified in mice (56-58), and the family CAG-274 has been identified in metagenomic datasets from diverse species (59-61). 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 (62). 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 (63). The other MAG belonged to the genus *Scrofimicrobium*, which has been isolated from the gastrointestinal tracts of diverse species (64, 65). 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 (66, 67).
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 (68) and is commonly found in the microbiota of avians, including
grouse species (16, 17).

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 (14, 20). Various phenolic
compounds are also found in heather, and have been shown to affect fermentation in the
rumen microbiota (69). Pathways for the degradation of phenols were found in both our
MAGs and other grouse species (17, 20), as were pathways for starch degradation (20).
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. It is likely that many of the
hits found using the “loose” algorithm from RGI are false-positives, due to their high number.
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 (70) and have also
been identified in samples from wild western capercaillies (14), but not amongst wild prairie
grouse (19). Many of these resistance genes are not well characterised, with limited
information on their molecular structures (70). 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* (71). A nitroimidazole reductase gene was also identified in our JAHZGC01 MAG. Similar genes can be found in a wide variety of bacterial taxa (72), but to our knowledge this is the first time this gene has been identified in this taxon. The ecological and one health relevance of these AMR genes is currently unknown, and is worth further investigation.

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 (14, 51, 73). 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 (74). 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 (75), 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 carcasses were privately owned by the estates on which the birds were shot. The Game Wildlife and Conservation Trust accessed and collected the gastrointestinal samples for this study. Informed consent was obtained from the owners to use these samples in our study. Field studies were conducted in accordance with local legislation. In the United Kingdom the birds listed in Schedule 2 of the Wildlife and Countryside Act 1981 (which include red grouse) are exempt from the general provision of protection. Any person is able to kill or 'take' a species listed outside of the close season for that bird. Therefore, red grouse can be legally taken between 12th August and 10th December within the United Kingdom. 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 Welfare and Ethical Review Body 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.

406

407 **Competing interests:**

408 The authors declare that they have no competing interests.

409

410 **Funding:**

411 This work was supported by the Biotechnology and Biological Sciences Research Council
412 Institute through Strategic Program Programme Grant funding (BBS/E/RL/230001A). Laura
413 Glendinning is supported by a University of Edinburgh Chancellor's Fellowship. As a UK
414 registered charity, the Game and Wildlife Conservation Trust relies mainly on fundraising
415 from the private sector. For the purpose of open access, the author has applied a Creative
416 Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising
417 from this submission.

418

419 **Authors' contributions:**

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

424

425 **Acknowledgements:**

426 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. Larsson DGJ, Gaze WH, Laxminarayan R, Topp E. AMR, one health and the environment. *Nat Microbiol.* 2023;**8**: 754-5 doi: 10.1038/s41564-023-01351-9

- 453 12. Bain C, Rathor G, Jones BP, Hassan MM, Papke P, et al. β -Lactam resistance genes
454 present in UK pheasants and red-legged partridges. Vet Rec. 2023;**192**: e2540 doi:
455 <https://doi.org/10.1002/vetr.2540>
- 456 13. Chintoan-Uta C, Seguino A, editors. Antimicrobial resistance of *Campylobacter* spp.
457 in wild game pheasants (*Phasianus colchicus*) in Scotland. Microbiology Society Annual
458 Conference; 2018.
- 459 14. Rasmussen JA, Chua PYS. Genome-resolving metagenomics reveals wild western
460 capercaillies (*Tetrao urogallus*) as avian hosts for antibiotic-resistance bacteria and their
461 interactions with the gut-virome community. Microbiol Res. 2023;**271**: 127372 doi:
462 10.1016/j.micres.2023.127372
- 463 15. Wienemann T, Schmitt-Wagner D, Meuser K, Segelbacher G, Schink B, et al. The
464 bacterial microbiota in the ceca of Capercaillie (*Tetrao urogallus*) differs between wild and
465 captive birds. Syst Appl Microbiol. 2011;**34**: 542-51 doi: 10.1016/j.syapm.2011.06.003
- 466 16. Drovetski SV, O'Mahoney MJV, Matterson KO, Schmidt BK, Graves GR. Distinct
467 microbiotas of anatomical gut regions display idiosyncratic seasonal variation in an avian
468 folivore. Anim microbiome. 2019;**1**: 2 doi: 10.1186/s42523-019-0002-6
- 469 17. Kohl KD, Connelly JW, Dearing MD, Forbey JS. Microbial detoxification in the gut of
470 a specialist avian herbivore, the Greater Sage-Grouse. FEMS Microbiol Lett. 2016;**363**:
471 fnw144 doi: 10.1093/femsle/fnw144
- 472 18. Vaasjo E, Stothart MR, Black SR, Poissant J, Whiteside DP. The impact of
473 management on the fecal microbiome of endangered greater sage-grouse (*Centrocercus*
474 *urophasianus*) in a zoo-based conservation program. Conserv Physiol. 2024;**12**: coae052
475 doi: 10.1093/conphys/coae052
- 476 19. Drovetski SV, Schmidt BK, Lai JE, Gross MS, Hladik ML, et al. Exposure to crop
477 production alters cecal prokaryotic microbiota, inflates virulome and resistome in wild prairie
478 grouse. Environ Pollut. 2022;**306**: 119418 doi: 10.1016/j.envpol.2022.119418

- 479 20. Salgado-Flores A, Tveit AT, Wright A-D, Pope PB, Sundset MA. Characterization of
480 the cecum microbiome from wild and captive rock ptarmigans indigenous to Arctic Norway.
481 PLOS ONE. 2019;**14**: e0213503 doi: 10.1371/journal.pone.0213503
- 482 21. Durnova E, Karmanova E, Sorokina T, Mayorova K, Aksenov A. Bacterial diversity
483 and composition in the internal organs of Taiga Bean Goose, Greater White-Fronted Goose
484 and Willow Ptarmigan as a new tools in the Arctic biomonitoring system. Diversity. 2025;**17**:
485 101 doi: 10.3390/d17020101
- 486 22. Ingvaldsen EW, Østnes JE, Kleven O, Davey M, Fossøy F, Nilsen EB. Fecal DNA
487 metabarcoding reveals seasonal and annual variation in willow ptarmigan diet. R Soc Open
488 Sci. 2024;**11**: 231518 doi: 10.1098/rsos.231518
- 489 23. Dunn JC, Clegg SR. High prevalence of multi-drug resistant bacteria in faecal
490 samples from UK passerine birds. bioRxiv. 2024: 2024.07.24.604896 doi:
491 10.1101/2024.07.24.604896
- 492 24. Fantham HB. Observations on the parasitic protozoa of the Red Grouse (*Lagopus*
493 *scoticus*), with a note on the Grouse Fly. Proc Zool Soc Lond. 1910;**80**: 692-708 doi:
494 10.1111/j.1096-3642.1910.tb01911.x
- 495 25. FastQC: A quality control analysis tool for high throughput sequencing data.
496 <https://github.com/s-andrews/FastQC>.
- 497 26. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: Summarize analysis results for
498 multiple tools and samples in a single report. Bioinformatics. 2016;**32**: 3047-8 doi:
499 10.1093/bioinformatics/btw354
- 500 27. Chen S, Zhou Y, Chen Y, Gu J. fastp: An ultra-fast all-in-one FASTQ preprocessor.
501 Bioinformatics. 2018;**34**: i884-i90 doi: 10.1093/bioinformatics/bty560
- 502 28. Eastman DS, Jenkins D. Comparative food habits of red grouse in northeast
503 Scotland, using fecal analysis. J Wildl Manag. 1970;**34**: 612-20 doi: 10.2307/3798870
- 504 29. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-
505 MEM. arXiv 2013;**1303**: 3997 doi: 10.48550/arXiv.1303.3997

506 30. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. The Sequence
507 Alignment/Map format and SAMtools. *Bioinformatics*. 2009;**25**: 2078-9 doi:
508 10.1093/bioinformatics/btp352

509 31. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2.
510 *Genome Biol*. 2019;**20**: 257 doi: 10.1186/s13059-019-1891-0

511 32. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: Estimating species
512 abundance in metagenomics data. *PeerJ Comput Sci*. 2017;**3**: e104 doi: 10.7717/peerj-
513 cs.104

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

517 34. Aroney ST, Newell RJ, Nissen JN, Camargo AP, Tyson GW, Woodcroft BJ. CoverM:
518 Read alignment statistics for metagenomics. *arXiv*. 2025;**preprint**: 2501.11217 doi:
519 2501.11217

520 35. Kang DD, Li F, Kirton E, Thomas A, Egan R, et al. MetaBAT 2: An adaptive binning
521 algorithm for robust and efficient genome reconstruction from metagenome assemblies.
522 *PeerJ*. 2019;**7**: e7359 doi: 10.7717/peerj.7359

523 36. Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. CheckM2: A rapid, scalable and
524 accurate tool for assessing microbial genome quality using machine learning. *Nat Methods*.
525 2023;**20**: 1203-12 doi: 10.1038/s41592-023-01940-w

526 37. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: A tool for fast and accurate
527 genomic comparisons that enables improved genome recovery from metagenomes through
528 de-replication. 2017;**11**: 2864-8 doi: 10.1038/ismej.2017.126

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

- 532 39. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: A toolkit to classify
533 genomes with the Genome Taxonomy Database. *Bioinformatics*. 2019;**36**: 1925-7 doi:
534 10.1093/bioinformatics/btz848
- 535 40. Zhou Z, Tran PQ, Breister AM, Liu Y, Kieft K, et al. METABOLIC: High-throughput
536 profiling of microbial genomes for functional traits, metabolism, biogeochemistry, and
537 community-scale functional networks. *Microbiome*. 2022;**10**: 33 doi: 10.1186/s40168-021-
538 01213-8
- 539 41. Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya Amogelang R, et al. CARD
540 2023: Expanded curation, support for machine learning, and resistome prediction at the
541 Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res*. 2022;**51**: D690-D9 doi:
542 10.1093/nar/gkac920
- 543 42. Wickham H. Reshaping data with the reshape package. *J Stat Softw*. 2007;**21**: 1 - 20
544 doi: 10.18637/jss.v021.i12
- 545 43. Wilke CO. cowplot: Streamlined plot theme and plot annotations for 'ggplot2'. 2024.
- 546 44. Wickham H, François R, Henry L, Müller K, Vaughan D. dplyr: A grammar of data
547 manipulation. 2023.
- 548 45. Wickham H. ggplot2: Elegant graphics for data analysis: Springer-Verlag New York;
549 2016.
- 550 46. Wickham H, Averick M, Bryan J, Chang W, McGowan LDA, et al. Welcome to the
551 Tidyverse. *J Open Source Softw*. 2019;**4**: 1686 doi: 10.21105/joss.01686
- 552 47. Asnicar F, Thomas AM, Beghini F, Mengoni C, Manara S, et al. Precise phylogenetic
553 analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. *Nat*
554 *Commun*. 2020;**11**: 2500 doi: 10.1038/s41467-020-16366-7
- 555 48. Figtree. <https://github.com/rambaut/figtree>.
- 556 49. Pallen MJ, Ponsero AJ, Telatin A, Moss CJ, Baker D, et al. Faecal metagenomes of
557 great tits and blue tits provide insights into host, diet, pathogens and microbial biodiversity.
558 *Access Microbiol*. 2025;**7**: doi: 10.1099/acmi.0.000910.v3

559 50. Liu J, Li XY, Song WT, Zeng XX, Li H, et al. Metagenomic analysis of environmental
560 samples from wildlife rescue station at Poyang Lake, China. *Biomed Environ Sci*. 2023;**36**:
561 595-603 doi: 10.3967/bes2023.087

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

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

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

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

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

577 56. Beresford-Jones BS, Forster SC, Stares MD, Notley G, Viciani E, et al. The Mouse
578 Gastrointestinal Bacteria Catalogue enables translation between the mouse and human gut
579 microbiotas via functional mapping. *Cell Host Microbe*. 2022;**30**: 124-38.e8 doi:
580 10.1016/j.chom.2021.12.003

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

584 58. Lesker TR, Durairaj AC, Gálvez EJC, Lagkourados I, Baines JF, et al. An integrated
585 metagenome catalog reveals new insights into the murine gut microbiome. *Cell Rep*.
586 2020;**30**: 2909-22.e6 doi: 10.1016/j.celrep.2020.02.036

- 587 59. Riva A, Sahin E, Volpedo G, Petretto A, Lavarello C, et al. Identification of an
588 epilepsy-linked gut microbiota signature in a pediatric rat model of acquired epilepsy.
589 Neurobiol Dis. 2024;**194**: 106469 doi: 10.1016/j.nbd.2024.106469
- 590 60. Zheng J, Sun Q, Zhang M, Liu C, Su Q, et al. Noninvasive, microbiome-based
591 diagnosis of inflammatory bowel disease. Nat Med. 2024;**30**: 3555-67 doi: 10.1038/s41591-
592 024-03280-4
- 593 61. Blyton MDJ, Soo RM, Hugenholtz P, Moore BD. Maternal inheritance of the koala gut
594 microbiome and its compositional and functional maturation during juvenile development.
595 Environ Microbiol. 2022;**24**: 475-93 doi: 10.1111/1462-2920.15858
- 596 62. Mukherjee A, Lordan C, Ross RP, Cotter PD. Gut microbes from the phylogenetically
597 diverse genus *Eubacterium* and their various contributions to gut health. Gut Microbes.
598 2020;**12**: 1802866 doi: 10.1080/19490976.2020.1802866
- 599 63. Kim CY, Lee M, Yang S, Kim K, Yong D, et al. Human reference gut microbiome
600 catalog including newly assembled genomes from under-represented Asian metagenomes.
601 Genome Med. 2021;**13**: 134 doi: 10.1186/s13073-021-00950-7
- 602 64. Wylensek D, Hitch TCA, Riedel T, Afrizal A, Kumar N, et al. A collection of bacterial
603 isolates from the pig intestine reveals functional and taxonomic diversity. Nat Commun.
604 2020;**11**: 6389 doi: 10.1038/s41467-020-19929-w
- 605 65. IJdema F, De Smet J, Crauwels S, Lievens B, Van Campenhout L. Meta-analysis of
606 larvae of the black soldier fly (*Hermetia illucens*) microbiota based on 16S rRNA gene
607 amplicon sequencing. FEMS Microbiol Ecol. 2022;**98**: fiac094 doi: 10.1093/femsec/fiac094
- 608 66. Salgado JFM, Hervé V, Vera MAG, Tokuda G, Brune A. Unveiling lignocellulolytic
609 potential: a genomic exploration of bacterial lineages within the termite gut. Microbiome.
610 2024;**12**: 201 doi: 10.1186/s40168-024-01917-7
- 611 67. Bowerman KL, Knowles SCL, Bradley JE, Baltrūnaitė L, Lynch MDJ, et al. Effects of
612 laboratory domestication on the rodent gut microbiome. ISME Commun. 2021;**1**: 49 doi:
613 10.1038/s43705-021-00053-9

- 614 68. Betancur-Murillo CL, Aguilar-Marín SB, Jovel J. *Prevotella*: A key player in ruminal
615 metabolism. 2022;**11**: 1 doi: 10.3390/microorganisms11010001
- 616 69. Murray AH, Iason GR, Stewart C. Effect of simple phenolic compounds of heather
617 (*Calluna vulgaris*) on rumen microbial activity *in vitro*. J Chem Ecol. 1996;**22**: 1493-504 doi:
618 10.1007/bf02027727
- 619 70. Stogios PJ, Savchenko A. Molecular mechanisms of vancomycin resistance. Protein
620 Sci. 2020;**29**: 654-69 doi: 10.1002/pro.3819
- 621 71. Li J, Li X-Z. Antimicrobial resistance and drug efflux pumps in *Helicobacter*. In: Li X-
622 Z, Elkins CA, Zgurskaya HI, editors. Efflux-mediated antimicrobial resistance in bacteria:
623 Mechanisms, regulation and clinical implications: Springer International Publishing; 2016. p.
624 489-513.
- 625 72. Alauzet C, Lozniewski A, Marchandin H. Metronidazole resistance and nim genes in
626 anaerobes: A review. Anaerobe. 2019;**55**: 40-53 doi: 10.1016/j.anaerobe.2018.10.004
- 627 73. Tang B, Wang Y, Dong Y, Cui Q, Zeng Z, et al. The catalog of microbial genes and
628 metagenome-assembled genomes from the gut microbiomes of five typical crow species on
629 the Qinghai–Tibetan Plateau. Microorganisms. 2024;**12**: 2033 doi:
630 10.3390/microorganisms12102033
- 631 74. Eisenhofer R, Odriozola I, Alberdi A. Impact of microbial genome completeness on
632 metagenomic functional inference. ISME Commun. 2023;**3**: 12 doi: 10.1038/s43705-023-
633 00221-z
- 634 75. Gao Y, Luo H, Lyu H, Yang H, Yousuf S, et al. Benchmarking short-read
635 metagenomics tools for removing host contamination. GigaScience. 2025;**14**: giaf004 doi:
636 10.1093/gigascience/giaf004

637

