Freshwater snail faecal metagenomes reveal environmental reservoirs of antimicrobial resistance genes across two continents

Angus M. O'Ferrall [a], Alexandra Juhász [a, b], Sam Jones [a], Peter Makaula [c], Gladys Namacha [c], Shaali Ame [d], David Oguttu [e], Aidan Foo [a, f], Sekeleghe A. Kayuni [a, c], James E. LaCourse [a], Janelisa Musaya [c, g], J. Russell Stothard [a], Adam P. Roberts [a] *

[a] Department of Tropical Disease Biology, Liverpool School of Tropical Medicine, Liverpool, United Kingdom

[b] Institute of Medical Microbiology, Semmelweis University, Budapest, Hungary

[c] Malawi Liverpool Wellcome Programme, Queen Elizabeth Central Hospital, Blantyre, Malawi

[d] Neglected Tropical Diseases Program, Ministry of Health, Stone Town, Zanzibar, Tanzania

[e] Vector Control Division, Ministry of Health, Kampala, Uganda

[f] Department of Vector Biology, Liverpool School of Tropical Medicine, Liverpool, United Kingdom

[g] Department of Pathology, School of Medicine and Oral Health, Kamuzu University of Health Sciences, Blantyre, Malawi

* Corresponding author – E-mail: adam.roberts@lstmed.ac.uk

Keywords

AMR, Carbapenemase, Gastropoda, MCR, One Health, OXA-181

Abstract

The transfer of antimicrobial resistance genes (ARGs) from environmental microbes to pathogens is a critical but underexplored One Health driver of antimicrobial resistance (AMR). Here, we evaluate freshwater snails, which are geographically widespread aquatic invertebrates, as environmental reservoirs of ARGs. We collected faeces from eight gastropod genera at 15 freshwater locations across Malawi, Uganda, Zanzibar, and the United Kingdom, and conducted the first freshwater snail faecal metagenomics study. We detected putative ARGs predicted to confer resistance to 13 antibiotic classes, including carbapenems in all countries. All ARGs that could be assembled into metagenome-assembled genomes (MAGs) were found within Proteobacteria, which dominated the faecal microbiomes and were strongly associated with the total ARG load. In Malawi, we linked $bla_{OXA-181}$ (bla_{OXA-48}-like), a previously mobilised carbapenemase gene, to *Shewanella xiamenensis*, the gene's known progenitor. We detected another bla_{OXA-48}-like gene by read-mapping from a sample in the United Kingdom. We identified mobile colistin resistance (*mcr*)-like genes at 11 of 15 locations, with two *mcr-7*-like genes being found in an *Aeromonas jandaei* MAG in Uganda. Our findings highlight freshwater snail faeces as a One Health-relevant environmental reservoir of clinically important ARGs.

Data Summary

Short reads from all samples sequenced in this study have been deposited in the Sequence Read Archive (SRA) under BioProject PRJNA1211045, with accession numbers SRX27371064 – SRX27371078. Data and code used to carry out analyses in R are available at https://github.com/amoreo71/freshwater_snail_faecal.

59 **MAIN TEXT**

60 **Introduction**

61 Antimicrobial resistance (AMR) is a growing global health emergency driven by industrial scale
62 use of antimicrobials in medicine, veterinary practice, and agriculture since their discovery in
63 the twentieth century (1). The impact of AMR on human health is largely driven by the
64 acquisition of antimicrobial resistance genes (ARGs) in pathogenic bacteria (2). Consequently,
65 most data used to investigate AMR originate from disease-causing organisms, yet there is
66 increasing recognition that One Health frameworks accounting for the interdependence of
67 human, animal, and environmental health systems need consideration to better understand
68 the origins, drivers, and spread of AMR (3). This approach is essential for assessing how AMR
69 may be linked to crossing planetary boundaries, such as those related to freshwater use,
70 biodiversity loss, and climate change (4,5).

71 Outside the context of AMR, freshwater gastropod snails are already of interest in One Health
72 research. Snails are globally distributed in aquatic environments, often playing critical roles as
73 obligate intermediate hosts in the transmission of neglected tropical diseases when freshwater
74 sources are contaminated with human and animal excreta. For instance, snails of the *Bulinus*,
75 *Biomphalaria*, and *Oncomelania* genera are key in the transmission of schistosomiasis (6),
76 while *Lymnaea* snails transmit fascioliasis (7). Recent efforts to characterise the microbiomes
77 of medically important freshwater snails are generally motivated by the growing recognition
78 that host-associated microorganisms can significantly influence host-parasite interactions (8).
79 The microbiomes of *Biomphalaria* snails have been shown by 16S taxonomic profiling to be
80 diverse and distinct from those of their freshwater environments (9), yet bacteriome transplant
81 experiments have demonstrated the potential for these microbiomes to change depending on
82 the surrounding environmental bacteria (10). Freshwater snail populations are highly mobile
83 (11), capable of translocating due to changes in ecosystems caused by flooding or the
84 formation of artificial water bodies (12,13), and are distributed across a range of human-
85 influenced areas (14,15). Considering that the environments where freshwater snails graze—
86 typically on detritus—can be contaminated with AMR enteric bacteria at sites of interaction
87 with humans and animals, we wanted to explore the potential role of snails as widespread
88 environmental reservoirs of ARGs (16). This may be particularly relevant in Low- and Middle-
89 Income Countries (LMICs), where AMR is disproportionately linked to increased mortality (17),
90 and where poor environmental health infrastructure is linked to AMR transmission across
91 various One Health compartments (18,19).

92 The analysis of shotgun metagenomic sequence data is increasingly being used in the study
93 of complex microbial samples and AMR, made possible by advances in next-generation

sequencing and computational science technologies (20). Metagenomic studies are not reliant on culture steps that alter microbial communities and are therefore able to yield non-selective data on sample composition (21). Metagenomic analyses generally utilise short-read sequence data, which when coupled with the complexity of microbiome samples poses challenges in the assembly of metagenome-assembled genomes (MAGs) for low-abundance organisms without ultra-deep sequencing (22,23). Approaches that map reads directly to reference databases are therefore useful for taxonomic profiling and the detection of low-abundance target genes as they do not rely on metagenome assembly, and have proved to be more sensitive than assembly-based approaches for the detection of ARGs in complex samples such as faeces (24). When combining read-based profiling with assembly-based approaches that uncover the genomic context of AMR, allowing the linking of ARGs to specific hosts and mobile genetic elements (MGEs), metagenomics can enable a comprehensive understanding of the resistome and its potential impacts on microbial ecology and global health.

Here, we present a dataset characterising the first freshwater snail faecal metagenomes from 15 locations within Malawi, Uganda, Zanzibar (a self-governing region within the Republic of Tanzania), and the United Kingdom. Acknowledging that freshwater snail microbiomes vary by organ (25), and to mirror the natural mechanisms of faecal bacterial transfer from gastropods to the environment, we sampled faeces excreted by snails immediately after they were collected from their habitats, rather than extracting DNA directly from snail organs or rearing snails under laboratory conditions prior to DNA extraction. We use whole-genome metagenomics to investigate the role of freshwater gastropods as reservoirs, and potentially vectors, of AMR. Our goal is to identify clinically relevant ARGs within these common aquatic invertebrates and to examine their genomic context, providing insights and targets for future studies on ARG transfer to human and animal pathogens. We utilise direct read mapping for taxonomic profiling and ARG identification, followed by less sensitive but highly insightful assembly-based methods to explore the genomic context of ARGs.

Methods

Sampling, processing, and sequencing

In 2023 and 2024, we collected freshwater snails at 15 freshwater locations in four administrative regions across two continents, targeting snails of a single genus at each location (*Supplementary Table 1*). In Africa, collections were made from six locations in Malawi (five *Bulinus* and one *Cleopatra*), three locations in Uganda (one each of *Biomphalaria*, *Bulinus*, and *Physella*), and four locations on Unguja Island, Zanzibar (one each of *Bulinus*, *Cleopatra*, *Lanistes*, and *Pila*). In Europe, two locations in the United Kingdom were sampled

129 (one *Anisus* and one *Lymnaea*). Sampling was opportunistic in nature and was carried out in
130 conjunction with ongoing parasitological surveys that were not originally designed for balanced
131 ecological comparisons between snail genera. As such, the presence of snail genera across
132 the 15 sites reflects natural distributions at the time of sampling, rather than a standardised
133 genus-level survey.

134 At each location, groups of 50 snails from the identified genus were gathered, rinsed, and
135 temporarily pooled in sterile water. After 24 hours, faecal material deposited by each pooled
136 snail group was collected to obtain adequate biomass for DNA extraction and sequencing.
137 DNA was extracted from PBS-washed faecal samples with the QIAamp PowerFecal Pro DNA
138 Kit (QIAGEN, Germany), as per manufacturer's instructions. Concentration, purity and
139 integrity of extracted DNA was assessed using a Qubit 4 fluorometer (Thermo Fisher Scientific,
140 USA), NanoDrop One spectrophotometry (Thermo Fisher Scientific, USA), and the
141 TapeStation System (Agilent Technologies, USA). Purified DNA extracts were shipped to
142 Genewiz/AZENTA (UK) for library preparation, validation, and shotgun metagenomic
143 sequencing. Sequencing was carried out with the Illumina NovaSeq instrument using a 2 x
144 150 bp paired-end configuration. The number of 150 bp Illumina paired end reads per sample
145 ranged from 9,340,457 to 35,991,622 (*Supplementary Table 2*). The resulting 15
146 metagenomes proceeded to downstream analyses, each representing the faeces of 50 snails
147 from a specific genus at a specific location.

148 **Bioinformatics**

149 Read processing

150 Raw FASTQ sequence files were quality assessed using FASTQC (https://github.com/s-
151 andrews/FastQC) Version 0.11.9. Trimming with Trimmomatic (26) Version 0.39 was
152 performed to remove adaptor sequences and low-quality bases with a sliding window quality
153 cutoff of Q20 and a minimum read length of 50 bp. Standard practice in metagenomic studies
154 is to filter out host DNA by precisely aligning all reads to host reference genomes. As our study
155 involved the analysis of pooled DNA samples from hosts whose genomes are not all available
156 in databases, we applied the Kraken Software Suite protocol outlined by Lu et al. (27) to
157 identify and extract contaminant and host DNA in a multi-step approach. Trimmed reads were
158 classified using Kraken 2 (28) (Version 2.1.3). Kraken 2 works by matching read k-mers
159 against a database of k-mers from existing genomes, before assigning each read to the lowest
160 common ancestor (LCA) of all genomes that share those k-mers. We initially classified reads
161 using the Kraken 2 standard database, which contains RefSeq archaea, bacteria, viruses,
162 plasmid complete genomes, UniVec Core and the human reference genome, GRCh38. We
163 then filtered out reads identified as eukaryotic (mapping to the human reference genome,

164 either due to low-level human DNA contamination or because they were eukaryote-generic
165 reads originating from snails) using the KrakenTools "extract_kraken_reads.py" script
166 (https://github.com/jenniferlu717/KrakenTools). Retained reads were considered to have
167 passed quality control.

168 We also built a custom Kraken 2 Gastropoda database containing the available GenBank snail
169 genomes for all species within the genera represented in our samples: *Anisus vortex*
170 (GCA_949126835.1), *Anisus vorticulus* (GCA_964264155.1), *Biomphalaria glabrata*
171 (GCA_947242115.1), *Biomphalaria pfeifferi* (GCA_030265305.1), *Biomphalaria straminea*
172 (GCA_021533235.1), *Biomphalaria sudanica* (GCA_036873155.1), *Bulinus truncatus*
173 (GCA_021962125.1), *Lanistes nyassanus* (GCA_004794575.1), *Lymnaea stagnalis*
174 (GCA_964033795.1), and *Physella acuta* (GCA_028476545.3). Genomes from the *Cleopatra*
175 and *Pila* genera were not included, as they were not available within GenBank. Separately
176 from classifications using the Kraken 2 standard database, we mapped all trimmed reads
177 passing quality control to the custom-built Gastropoda database. We then extracted reads
178 identified to be of gastropod origin, that had not already been discarded after being labelled
179 as eukaryotic by the Kraken 2 standard database. This conservatively extracted subset of
180 remaining Gastropoda reads that passed quality control (0.09 – 0.44% per sample) also
181 underwent ARG detection steps to provide us with reassurance that our inability to precisely
182 align each sample's reads to a specific host reference genome did not bias ARG profiles.

183 Taxonomic profiling

184 Kraken 2 reports containing bacterial read taxonomic assignments were used to run species
185 abundance estimation with Bracken (29) (Version 2.9), with a minimum number of reads
186 required for a classification at the specified rank set at 10, to reduce noise from potential low-
187 abundance contaminant species, as per default recommendations.

188 ARG database and clustering

189 We used the ResFinder database (30) (Version 2.3.2) to screen for ARGs. This database was
190 selected because it focuses on acquired ARGs with known clinical relevance and mobile
191 potential, thereby aligning with our aim to identify resistance mechanisms that may contribute
192 to the spread of AMR from environmental reservoirs to clinical settings. Sequences in the
193 ResFinder database were downloaded, then clustered using CD-HIT (31) (Version 4.8.1) to
194 identify ARG clusters containing sequences with a minimum nucleotide identity of 90%,
195 referred to as Cluster90s (24).

196

197

198 Read-based ARG detection

199 To identify ARGs within each sample, reads were aligned to reference sequences in the
200 ResFinder database using KMA (32) (Version 1.4.14), a tool designed to map raw reads
201 directly against redundant databases such as ARG databases. Alignments were then filtered
202 to retain only those with a minimum 90% query identity and 60% template coverage—a
203 common approach in the analysis of complex environmental metagenomes that balances
204 accuracy and sensitivity by favouring reliable ARG identification while allowing for partial
205 matches, thereby reducing the risk of missing low-abundance or divergent ARGs (33). For all
206 reference sequences with alignment hits after filtering, the corresponding Cluster90 was
207 identified and annotated manually. We calculated the normalised abundance of each ARG
208 from read-mapping data, expressed as RPKM (Reads Per Kilobase per Million mapped
209 bacterial reads), then summed the RPKM of all ARGs to calculate total ARG load per sample.
210 We also performed read-based ARG detection on random sub-sets of reads (5 million, 10
211 million, 15 million, 20 million) from samples with > 20 million read pairs following quality control
212 (Uk1, Z1, Z3, Z4). The ARG detection data generated from these read sub-sets was used to
213 plot rarefaction curves and assess diversity saturation.

214 Metagenome assembly, binning, and classification

215 Reads identified as being bacterial by Kraken 2 were assembled with metaSPAdes (34)
216 (Version 3.11.1), using the default k-mer spectrum (21, 33, 55). Assembly statistics were
217 calculated with QUAST (35) (Version 5.0.2) using contigs ≥ 500 bp, to ignore low-quality
218 sequences and ensure compatibility with ARG-containing regions (*Supplementary Table 2*).
219 Contigs were indexed using the Burrows Wheeler Aligner (BWA) (Version 0.7.17), before
220 bacterial reads were aligned to the indexed contigs with "bwa-mem" (36). The contigs and
221 resulting sorted BAM files were parsed to the "jgi_summarize_bam_contig_depth" script from
222 MetaBAT2 (37) (Version 2.17), before the resulting depth files were used by MetaBAT2 to bin
223 assembled contigs of ≥ 2,500 bp to putative genomes (i.e., MAGs). CheckM (38) (Version
224 1.1.2) was used to assess the completeness and contamination of each MAG. CheckM
225 outputs were compared to standards set by the genome standards consortium to classify
226 MAGs as high-quality (>90% completeness, <5% contamination), medium-quality MAGs
227 (>50% completeness, <10% contamination) or low-quality (<50% completeness and/or >10%
228 contamination) (39). MAGs deemed as medium- or high-quality proceeded to further analysis.
229 MAGs underwent taxonomic classification using GTDB-Tk (40) (Version 2.1.1) against the
230 Genome Taxonomy Database (Version 09-RS220).

231

232

233 Assembly-based profiling of ARGs and MGEs

234 Assemblies were screened against the ResFinder (30) and ISfinder (41) databases using
235 ABRicate (https://github.com/tseemann/abricate) (Version 1.0.1) at default settings (minimum
236 80% identity and 80% query cover) to identify contigs containing ARGs and IS elements.
237 Compared to our read-mapping approach, in which we used a higher identity threshold to
238 identify hits (90%) in combination with a lower minimum query cover (60%) due to the
239 fragmented nature of using raw sequences to identify low-abundance genes, the assembly-
240 based approach allowed for the detection of ARGs and MGEs that could then be characterised
241 further within their genomic context following annotation. Annotations were performed using
242 Bakta (63) (Version 1.9.3; full database Version 5.1). To elucidate the potential origins and
243 dissemination pathways of ARGs, selected contigs were subjected to a sequence similarity
244 search using the Nucleotide Basic Local Alignment Search Tool (BLAST) webtool (megablast
245 algorithm) against the GenBank Nucleotide database. This step aimed to identify existing
246 homologous sequences and/or gene clusters. The Clinker webtool
247 (https://cagecat.bioinformatics.nl/tools/clinker) (64) was used to visualise gene cluster
248 comparisons.

249 **Statistical analysis and visualisation**

250 Species abundance and read-based ARG assay data were combined with
251 hierarchical/classification data and sample metadata to generate a
252 TreeSummarizedExperiment data container (65). Data associated with MAGs were handled
253 separately. Within the R environment, downstream data wrangling, analysis, and visualisation
254 took place predominantly using the R/Bioconductor *mia* and *miaViz* (Version 1.12.0)
255 packages, and the R/CRAN *ggplot2* (Version 3.5.1) and *pheatmap* (Version 1.0.12) packages.

256 We calculated alpha diversity statistics at the species level, leveraging the high taxonomic
257 resolution of shotgun metagenomics data. A Bray-Curtis dissimilarity matrix containing
258 pairwise dissimilarities among microbiome samples was generated at species level using
259 relative abundance data to investigate beta diversity and to perform principal coordinates
260 analysis (PCoA). Ahead of hierarchical clustering of samples and taxa at the order level,
261 centered log-ration (CLR) transformations were applied to microbial composition data (66) to
262 normalise for compositional effects, with a pseudocount added to manage zero values using
263 the *mia* package. Hierarchical clustering of both taxa and samples was applied to standardised
264 CLR-transformed (CLR-z) data using Euclidean distance and complete linkage, as
265 implemented by default in the *pheatmap* package. Before plotting ARG Cluster90 heatmaps,
266 a log-transformation was applied to ARG load to enable visualisation across a wide range of
267 abundances, with a pseudocount added to manage zero values using the *mia* package.

To assess monotonic relationships between numerical variables, such as the relative abundance of Proteobacteria and ARG load, we used Spearman's rank correlation coefficient testing. This non-parametric test was chosen due to the small number of locations sampled (n = 15), its suitability for non-normally distributed data, and its robustness to outliers. Significance was determined with a threshold of $p < 0.05$. These analyses were followed by analyses of ARGs and MGEs found in MAGs, which supplement read-based profiles with high-resolution, species-specific context. Here, we used Wilcoxon rank-sum tests (non-parametric) to assess differences in the counts of ARGs and IS elements detected per complete MAG (calculated using raw ARG and IS element counts alongside genome completeness values per MAG) between taxa. Again, the threshold for statistical significance was set at $p < 0.05$.

Results

The freshwater snail faecal microbiome is diverse, varied, yet consistently rich in Proteobacteria

Species abundance estimations showed that Proteobacteria was the most abundant phylum observed, with a mean relative abundance of 66.27% (*Figure 1a*). Enterobacterales, an order within the Proteobacteria containing multiple species of clinical concern due to their ability to acquire and exchange ARGs, were well represented, with a mean relative abundance of 4.79%. The Uk1 sample was a notable outlier in which the relative abundance of Enterobacterales was estimated to be 41.70%. When excluding Uk1, the mean relative abundance of Enterobacterales dropped to 2.15%. Four orders were more abundant than Enterobacterales (Burkholderiales, Rhizobiales, Aeromonadales, and Pseudomonadales) all of which had a mean relative abundance of greater than 5% (*Figure 1b*). All samples exhibited high community (alpha) diversity. Shannon's diversity index ranged from 3.881 (M5) – 7.372 (Z2) (mean = 6.474, median = 6.888). Simpson's diversity index, in which scores from 0 – 1 are possible (1 = maximum possible diversity), ranged from 0.793 (M5) – 0.999 (Z2) (mean = 0.972, median = 0.994). We observed that alpha diversity (Shannon's diversity index) was negatively corelated with the relative abundance of Proteobacteria (Spearman's rank correlation coefficient = -0.636; $p = 0.0129$).

Relative abundance data were aggregated at the order level, subjected to centred log-ratio (CLR) transformation, and subsequently standardised (CLR-z) to facilitate visual comparisons between taxa with varying absolute abundances (*Figure 1c*). From the 160 bacterial orders identified across all samples, the 50 most abundant orders, accounting for 97.92% of the total microbiome, were selected for hierarchical clustering and visualisation. Samples from the same administrative region clustered closely together, with the exception of Malawian samples, which were split into two distinct groups. This split corresponded to the two districts

where the samples were collected: samples M1, M2, M3, and M4, collected in Mangochi District on the southern shorelines of Lake Malawi, clustered together, while samples M5 and M6, collected in Nsanje District, over 275 km from Mangochi District sampling sites, formed a separate cluster. Although they clustered together, northern hemisphere samples (United Kingdom) did not form a highly divergent cluster despite likely differences in climate-related and environmental conditions. Microbiomes from *Bulinus* snails (the most-represented genus, obtained from seven locations) showed notable divergence: the sample collected in Zanzibar clustered at one extreme of the dendrogram alongside samples from other snail genera, distinct from *Bulinus* samples collected in Uganda and the two Malawian districts.

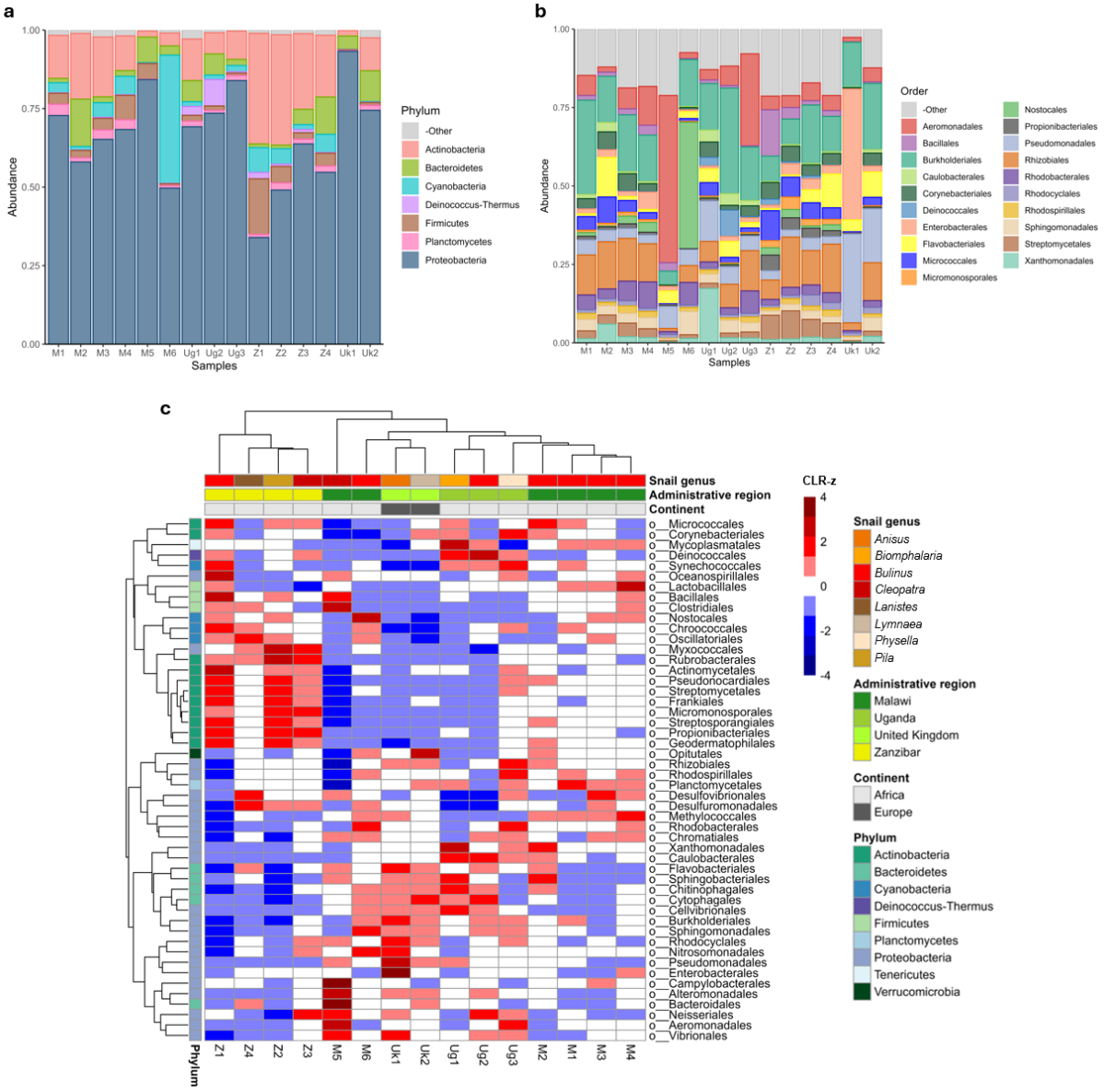


Figure 1. Exploration and comparison of bacterial taxa within freshwater snail faeces: a) Relative abundance of the top phyla; b) Relative abundance of the top orders; c) Clustered heatmap of standardised order abundances (CLR-z) across snail hosts and geographic

316 *locations. Column and row trees cluster samples and taxa, respectively, based on a*
317 *hierarchical clustering approach. Taxa with a mean relative abundance < 1% are classified as*
318 *'Other' in parts a) and b).*

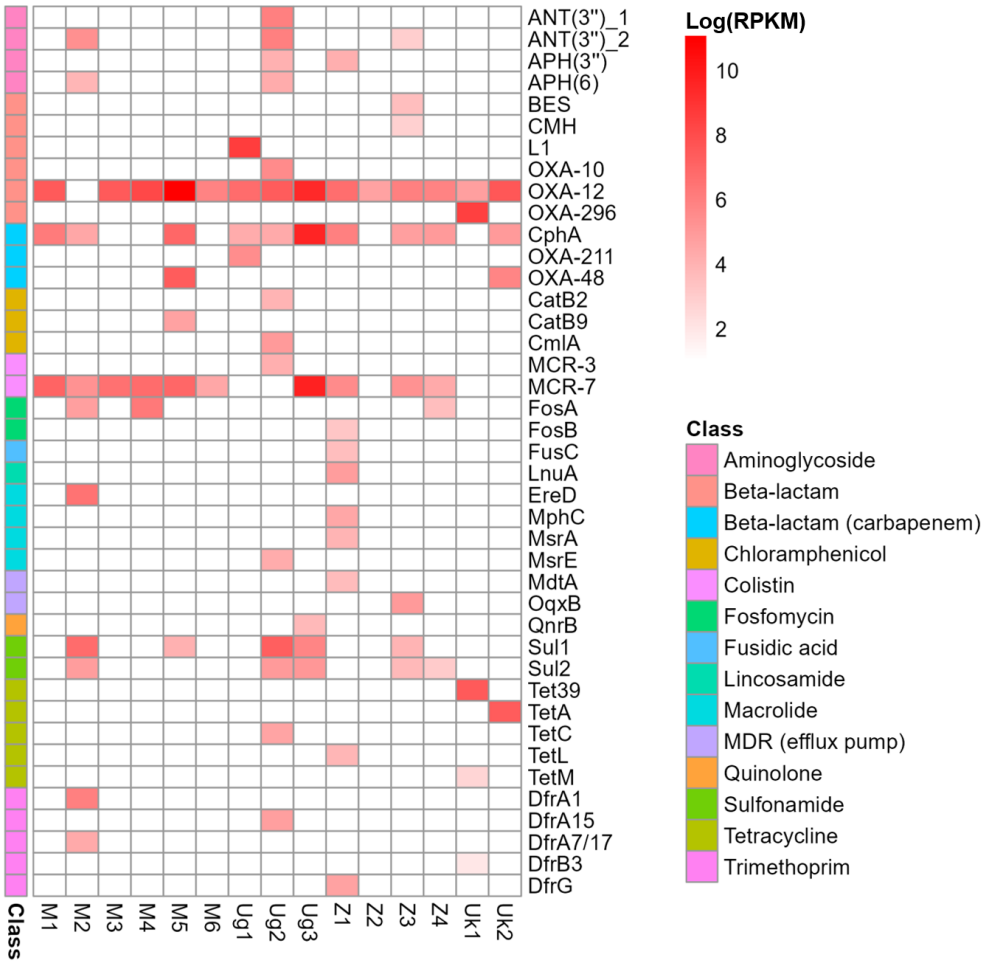
319 When investigating diversity between samples at the species level, a Bray-Curtis dissimilarity
320 matrix *(Supplementary Figure 1a)* showed a range of divergence between Malawian *Bulinus*
321 snails. We observed moderate divergence (0.455) between two samples (M2, M4) collected
322 a little over five kilometres from each other, both from stagnant water bodies close to the Lake
323 Malawi shoreline, suggesting that local factors play a key role in influencing the microbiome.
324 Despite these local variations, PCoA performed using Bray-Curtis distances at the species
325 level *(Supplementary Figure 1b)* revealed similar locational trends to our order-level
326 hierarchical clustering visualisation, in particular showing that the four samples from Mangochi
327 District, Malawi (M1 – M4), as well as the four from Unguja Island, Zanzibar (Z1 – Z4),
328 clustered closely together within their respective groups.

329 **ARGs conferring resistance to first- and last-line antibiotics detected by read mapping**

330 Using read-based k-mer alignment (KMA) to screen for ARGs against the ResFinder database
331 (30,32), reads mapped to 54 unique ARGs in total (minimum 90% query identity and 60%
332 template coverage), belonging to 41 ARG Cluster90s (in which reference genes share
333 minimum 90% nucleotide identity) (24). ARG richness (the number of unique ARGs identified
334 by read mapping in any given sample) ranged from 1 – 16 (mean = 7.07, median = 6). ARG
335 load ranged from 3.16 – 357.04 RPKM (mean = 87.99, median = 59.62). Although the depth
336 of sequencing and number of unique ARGs detected per sample were not correlated
337 (Spearman's rank correlation coefficient = 0.283; p = 0.3065) (*Supplementary Figure 2*), we
338 used random sub-sets of 5 million, 10 million, 15 million, and 20 million read pairs from
339 samples with more than 20 million pairs passing quality control (Uk1, Z1, Z3, Z4) to further
340 explore ARG saturation. Rarefaction curves showed that while ARG saturation can be reached
341 by 10 million reads (e.g. Uk1), other samples (Z1, Z3, Z4) required up to 20 million reads to
342 achieve saturation (*Supplementary Figure 3*). The ARGs that were not detected in smaller
343 read sub-sets were only detected at very low in abundance in full read sets, therefore having
344 a minimal effect on ARG load.

345 When we aggregated ARG relative abundance by class (i.e., the antibiotic that the ARGs
346 confer resistance to), beta-lactamase genes (excluding carbapenemases) were the most
347 abundant, contributing to 53.58% of the total ARG load across samples, followed by genes
348 conferring resistance to colistin (14.38%), carbapenems (10.27%), sulfonamides (5.32%),
349 tetracyclines (5.12%), aminoglycosides (2.77%), fosfomycin (2.15%), macrolides (2.14%),
350 trimethoprim (1.90%), lincosamides (0.51%) chloramphenicol (0.47%), fusidic acid (0.19%),

351 and quinolones (0.03%) (*Supplementary Figure 4*). Efflux pump genes conferring multidrug
352 resistance contributed 1.16% of the total ARG load. Beta-lactamase genes from the OXA-12
353 Cluster90 were the most prevalent in our samples and were detected in 14 of the 15
354 metagenomes (*Figure 2*). We also detected ARGs from the CphA Cluster90, a family of
355 carbapenem-hydrolyzing metallo-beta-lactamase genes, in 10 of the 15 samples. Three OXA-
356 carbapenemase genes were identified, each from a different country. These included two
357 ARGs from the OXA-48 Cluster90 ($bla_{OXA-181}$ and $bla_{OXA-547}$, in Malawi and the United Kingdom,
358 respectively), of concern globally due to plasmid-borne dissemination amongst
359 Enterobacterales (46), and $bla_{OXA-212}$ (Uganda) from the OXA-211 Cluster90 of
360 carbapenemase genes that naturally occur in *Acinetobacter johnsonii* (47). We also detected
361 putative colistin resistance genes, predominantly variants of the *mcr-7.1* gene, in eleven
362 metagenomes.



363

364 *Figure 2. Heatmap of normalised ARG Cluster90 abundances across samples, expressed as*
365 *log(RPKM). Cluster90s are colour-coded by the antibiotic class that they are predicted to*
366 *confer resistance against.*

We observed a positive correlation between the relative abundance of Proteobacteria (the predominant bacterial order) and ARG load (Spearman's rank correlation coefficient = 0.846; $p < 0.0001$) (*Figure 3a*). A significant association between Proteobacteria relative abundance and ARG richness was not detected (Spearman's rank correlation coefficient = 0.294; $p = 0.2877$). Among the 20 orders of bacteria with a mean relative abundance above 1% (*Supplementary Table 3*), only Pseudomonadales had a significant positive correlation with ARG load (Spearman's rank correlation coefficient = 0.604; $p = 0.0195$), although this was not a clear linear trend (*Figure 3b*). Meanwhile, the relative abundance of Aeromonadales was particularly high in two samples (M5, Ug3), within which ARG load was also markedly higher (> 300 RPKM) than all other samples (< 100 RPKM). However, the overall correlation between samples was deemed non-significant (Spearman's rank correlation coefficient = 0.357; $p = 0.1916$). These results indicate that different orders within the Proteobacteria are involved in AMR to varying extents in different snail microbiomes, and that assembly-based methods that elucidate the genomic context of AMR are needed to explore this with greater resolution.

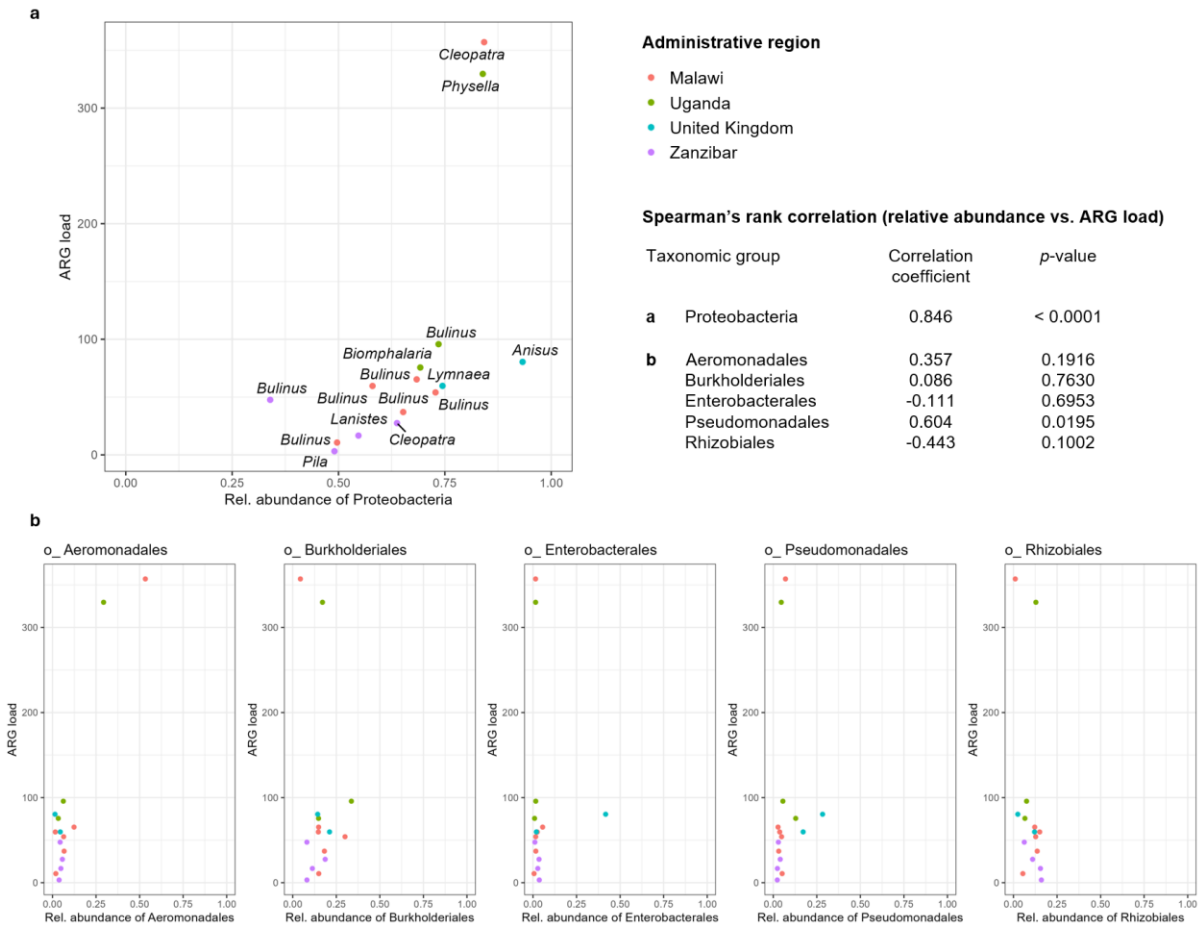
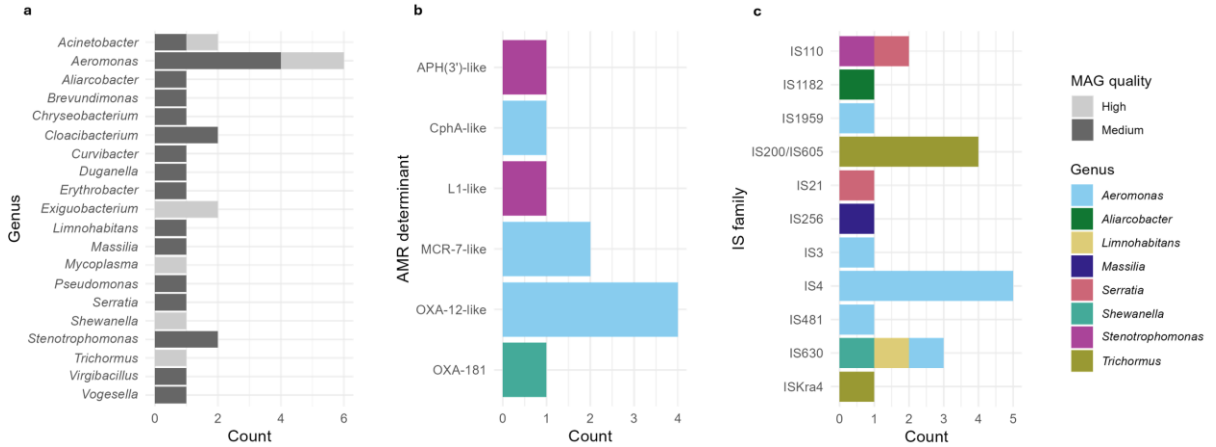


Figure 3. Associations between Proteobacteria and ARG load: a) Relative abundance of Proteobacteria versus ARG load; b) Relative abundance of the five most abundant orders within the Proteobacteria versus ARG load. Within-plot Labels in part a) are snail genera.

385 *Spearman's rank correlation coefficients and associated p-values are presented for each*
386 *taxonomic group for relative abundance versus ARG load.*

387 **Assembly-based methods link ARGs and MGEs to various bacterial genera and identify**
388 **known progenitors of mobile carbapenemase and colistin resistance**

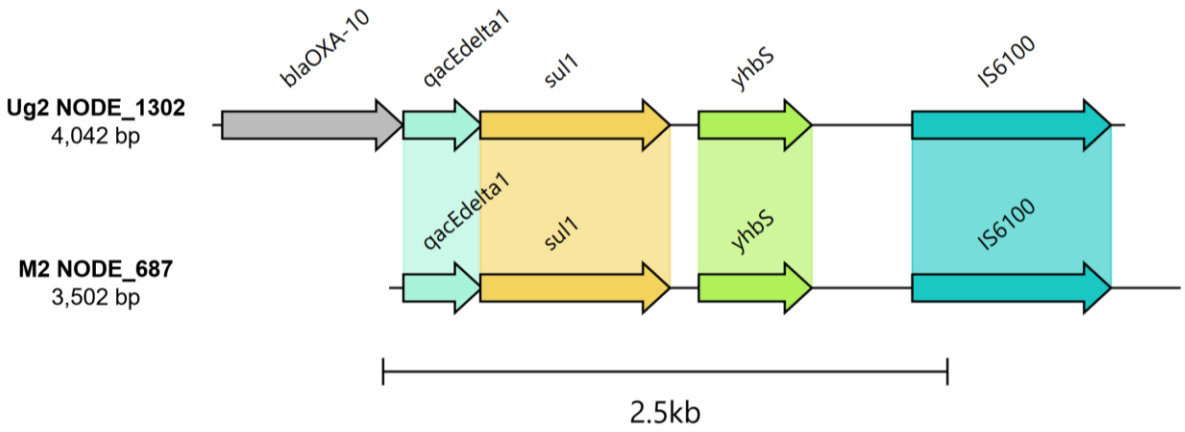
389 We detected fewer ARGs by screening metagenome assemblies against the ResFinder
390 database using ABRicate (mean = 2.53, median = 2) than we did by read mapping
391 (*Supplementary Figure 5*). However, we were able to acquire information on the carriage of
392 ARGs by specific taxonomic groups following assembly and binning of contigs. Across all
393 samples, 10 of the 38 ARGs (26.3%) assembled on contiguous sequences (contigs) could be
394 binned to high-quality (>90% completeness, <5% contamination) or medium-quality (>50%
395 completeness, <10% contamination) MAGs. These ARGs were all binned to MAGs classified
396 as Proteobacteria (genera: *Aeromonas*, n = 7; *Shewanella*, n = 1; *Stenotrophomonas*, n = 2)
397 (*Figure 4*). Specifically, these MAGs all belonged to the Gammaproteobacteria class, from
398 which ARGs were identified at a rate of 0.62 per complete genome. Even within our small
399 dataset of 29 medium- and high-quality MAGs, the difference in ARG detection rates in
400 Gammaproteobacteria, which made up 85.7% (18/21) of Proteobacteria MAGs and 62.1%
401 (18/29) of all MAGs, was significantly higher than in all other bacterial classes (p = 0.0391).
402 We detected between 8 and 97 putative IS elements per sample (mean = 37.8, median = 29).
403 IS elements were more commonly identified in Gammaproteobacteria MAGs (1.03 per
404 complete genome) than in other MAGs (0.57 per complete genome). Although the difference
405 observed did not meet the threshold for statistical significance within this relatively small
406 dataset (p = 0.1223), the trend likely reflects underlying biological differences in MGE content
407 between taxonomic groups.



408

409 *Figure 4. Summary of medium- and high-quality MAGs – a) Genera; b) AMR determinants*
410 *(annotated as '-like' if nucleotide identity is <100% when compared to reference); c) IS families.*

411 We identified two contigs containing both IS elements and ARGs, neither of which could be
412 binned to a MAG. A 4042 bp contig (Ug2 NODE_1302: *Bulinus* spp. faeces, Lake Victoria,
413 Uganda) contained *sul1* (sulfonamide resistant dihydropteroate synthase) and bla_{OXA-10} (beta-
414 lactamase) genes, on which both ARGs were located upstream of IS*6100*. The gene cluster
415 also contained *qacEdelta* (a gene conferring resistance to antiseptics) and *yhbS* (putative N-
416 acetyltransferase). The contig displays 100% query cover with one complete genome in the
417 GenBank Nucleotide database—*Aeromonas hydrophila* (accession no. AP025277)—with
418 99.97% identity, but also maintains high query cover (> 95%) and identity (> 99%) with
419 chromosomal and plasmid sequences from various other bacteria (*Pseudomonas aeruginosa*,
420 *Vibrio cholerae*, *Morganella morganii*, *Salmonella enterica*, *Klebsiella pneumoniae*). A second,
421 shorter, 3505 bp contig (M2 NODE_687: *Bulinus* spp. faeces, Mangochi District, Malawi),
422 aligned with 99.94% nucleotide identity to Ug NODE_1302 over the region containing
423 *qacEdelta, sul1* and *yhbS* upstream of IS*6100* (*Figure 5*), with a single nucleotide
424 polymorphism (SNP) in the *sul1* gene. The similarity observed between the two contigs from
425 Ug2 and M2, and with sequences from a range of bacterial species in GenBank, demonstrates
426 the widespread nature of certain ARGs associated with MGEs, which we show can be found
427 in freshwater snail faeces. However, we could only link MGEs directly to ARGs on contigs in
428 the above two examples. Searches for ARGs associated with IS elements in transposons
429 should ideally capture 24.34 kbp of DNA (based on the length of composite and unit
430 transposons in The Transposon Registry) (48). The median length of contigs containing ARGs
431 in our dataset was 2.10 kbp, demonstrating the limitations in assembling large contigs from
432 short reads in diverse metagenomic samples (23), and the associated inability to capture all
433 co-localised ARGs and MGEs.



434

435 *Figure 5: Gene cluster comparison between Ug2 NODE_1302 and M2 NODE_687 containing*
436 *bla_{OXA-10} (Ug2 only), qacEdelta1, sul1 and yhbS upstream of IS6100. Regions coding for*
437 *hypothetical proteins are not displayed.*

438 One of the two *bla*$_{OXA-48}$-like carbapenemase genes identified by read-mapping was also
439 detected in metagenome assemblies. NODE_173 from the M5 metagenome (*Cleopatra* spp.
440 faeces, Nsanje District, Malawi) contained *bla*$_{OXA-181}$ on a 20,996 bp contig with 17 open-
441 reading frames (ORFs). The contig was binned to a high-quality MAG classified as *Shewanella*
442 *xiamenensis* (*Table 1*). When the *bla*$_{OXA-181}$-containing contig was queried against the
443 GenBank Nucleotide database, the closest alignment was to chromosomal DNA from an *S.*
444 *xiamenensis* isolate (accession no. AP026732) obtained from urban drain water from Vietnam
445 (100% query cover, 98.51% identity) (49). With only one SNP in a non-coding region, the
446 contig also aligns with 99.92% identity over a 1309 bp region (8073 – 9381) to various
447 plasmids from the Enterobacterales containing the internationally disseminated *bla*$_{OXA-181}$ gene
448 and a truncated *lysR* (Δ*lysR*) gene, including the KP3-A plasmid from which *bla*$_{OXA-181}$ was first
449 characterised in a clinical *K. pneumoniae* isolate (accession no. JN205800) (50). The region
450 containing *bla*$_{OXA-181}$ and Δ*lysR* within the KP3-A (and other) plasmid(s) has previously be
451 traced back to its origin, *S. xiamenensis*, where *bla*$_{OXA-181}$ is found adjacent to a complete *lysR*
452 gene (51), as is also seen in our annotated sequences (*Supplementary Figure 6*). The ability
453 to detect, assemble, and bin this critical priority ARG is aided by the high relative abundance
454 of *Shewanella* in the M5 sample (5.07%). We also observed a strong *Shewanella* signal in the
455 Uk2 sample (relative abundance = 1.33%), from which another *bla*$_{OXA-48}$-like variant (*bla*$_{OXA-}$
456 $_{547}$) was detected by read-mapping, when compared to the remaining 13 samples in which no
457 *bla*$_{OXA-48}$-like genes were detectable (mean *Shewanella* relative abundance = 0.11%).

458 Two *mcr-7*-like lipid A phosphoethanolamine transferase genes were detected in a high-quality
459 *Aeromonas jandaei* MAG (*Table 1*) recovered from the Ug3 metagenome (*Physella* spp.
460 faeces, Kimi Island, Uganda) on separate contigs (NODE_28 and NODE_32: 86.88% and
461 81.93% nucleotide identity to *mcr-7.1*, respectively). Lipid A phosphoethanolamine transferase
462 enzymes cause a reduction in overall net-negative charge of the outer membrane, conferring
463 resistance to colistin among some gram-negative bacteria, and are a major public health
464 concern due to their propensity to be mobilised from progenitor species chromosomes to
465 plasmids, limiting last-resort antibiotic treatment options for infections caused by multi-drug-
466 resistant pathogens (52). *Aeromonas* spp. are the known progenitors of *mcr-3* and *mcr-7* (53).
467 From our 15 samples, reads mapped to the MCR-3 Cluster90 in one and the MCR-7 Cluster90
468 in 10, underscoring the widespread prevalence of *Aeromonas*-associated *mcr*-like genes in
469 the sampled environment.

470

471

472

473 *Table 1: Classification of MAGs recovered from freshwater snail faecal microbiomes, with*
474 *quality metrics, ARGs and IS elements detected, and GenBank Accession numbers for high-*
475 *quality MAGs.*

Phylum	Species	Sample	Quality (Completeness, Contamination)	ARGs detected	IS families detected	GenBank Accession
Bacteroidota	*Chryseobacterium aquaticum*	Ug3	Medium (51.7%, 0.3%)	---	---	---
	Cloacibacterium spp.	Z3	Medium (52.3%, 0.0%)	---	---	---
	Cloacibacterium spp.	Z4	Medium (63.8%, 0.0%)	---	---	---
Cyanobacteria	*Trichormus* spp.	M6	High (97.6%, 0.2%)	---	IS200/IS605, IS200/IS605, IS200/IS605, IS200/IS605, ISKra4	JBODRP000000000
Firmicutes	*Exiguobacterium indicum*	Z1	High (90.4%, 2.0%)	---	---	JBODRM000000000
	Exiguobacterium profundum	Z1	High (90.7%, 0.7%)	---	---	JBODRL000000000
	Mycoplasma spp.	Ug1	High (90.2%, 0.4%)	---	---	JBODRO000000000
	Virgibacillus spp.	M5	Medium (55.1%, 0.6%)	---	---	---
Proteobacteria	*Acinetobacter bohemicus*	Uk1	Medium (51.7%, 3.4%)	---	---	---
	Acinetobacter tandoii	M5	High (97.9%, 1.2%)	---	---	JBODRR000000000
	Aeromonas jandaei	Ug3	High (95.3%, 0.6%)	bla_{cphA}-like, bla_{OXA-12}-like, *mcr-7*-like, *mcr-7*-like	IS630, IS1595	JBODRN000000000
	Aeromonas jandaei	Z1	Medium (85.1%, 1.2%)	bla_{OXA-12}-like	---	---
	Aeromonas sobria	Uk2	Medium (82.1%, 0.6%)	bla_{OXA-12}-like	---	---
	Aeromonas spp.	M5	High (93.3%, 0.6%)	bla_{OXA-12}-like	IS4, IS4, IS4, IS4	JBODRQ000000000
	Aeromonas spp.	Z4	Medium (74.8%, 1.2%)	---	IS4	---
	Aeromonas veronii	Ug2	Medium (76.8%, 0.8%)	---	IS3, IS481	---
	Aliarcobacter spp.	M5	Medium (89.0%, 2.4%)	---	IS1182	---
	Brevundimonas nasdae	Ug2	Medium (73.4%, 1.5%)	---	---	---
	Curvibacter spp.	M1	Medium (70.1%, 0.5%)	---	---	---
	Duganella spp.	M1	Medium (76.6%, 0.5%)	---	---	---
	Erythrobacter spp.	M6	Medium (57.8%, 0.0%)	---	---	---
	Limnohabitans spp.	M2	Medium (56.5%, 0.2%)	---	IS630	---
	Massilia spp.	Ug2	Medium (68.9%, 0.7%)	---	IS256	---
	Pseudomonas sivasensis	Uk1	Medium (51.8%, 1.8%)	---		---
	Serratia fonticola	Uk1	Medium (71.9%, 0.0%)	---	IS21, IS110	---
	Shewanella xiamenensis	M5	High (96.0%, 0.8%)	$bla_{OXA-181}$	IS630	JBODRS000000000
	Stenotrophomonas bentonitica	Ug1	Medium (55.2%, 0.0%)	---	---	---
	Stenotrophomonas maltophilia	Ug1	Medium (84.1%, 1.7%)	*aph(3')*-like, bla_{L1}-like	IS110	---
	Vogesella spp.	Z3	Medium (65.5%, 0.0%)	---	---	---

476

477 **Discussion**

478 Our findings provide novel insights into the role that freshwater snails play as AMR reservoirs,
479 revealing a potential widespread environmental source of clinically relevant ARGs. Through
480 analysis of snail faecal metagenomes from diverse geographic locations, we demonstrate that
481 their microbiomes harbour a wide array of bacterial species and ARGs, including those
482 conferring resistance to both first- and last-line antibiotics. Identification of environmental
483 progenitors of potentially mobile ARGs conferring resistance to carbapenems and colistin is
484 particularly noteworthy in sub-Saharan Africa, where such antibiotics are not commonly
485 available or prescribed for patients with severe infections caused by Enterobacterales, but
486 where resistance among Enterobacterales to other treatment options such as penicillins, third-
487 generation cephalosporins, and gentamicin has increased (54). A likely increase in
488 carbapenem prescribing as a result of this may introduce additional selective pressures to
489 aquatic habitats, driving selection and transfer of ARGs from environmental reservoirs to
490 clinically relevant bacteria (55). Our results emphasise the potential for snails to contribute to
491 this dissemination in freshwater ecosystems, particularly in those modified by anthropogenic
492 change with deficits in local WASH (water, sanitation, and hygiene) infrastructures (3,19),
493 although we are not able to confirm ARG mobility from these sequencing data alone.

494 Proteobacteria dominated the snail faecal microbiomes, reinforcing results from previous
495 studies that found Proteobacteria to be highly abundant in other freshwater snail microbiomes
496 such as the gut and haemolymph (9,56–58). The significant correlations observed between
497 Proteobacteria abundance and both alpha diversity and ARG load suggest that an abundance
498 of Proteobacteria reduces overall microbiome diversity while increasing the relative
499 abundance of ARGs within the microbial community. We binned genes from the three most
500 prevalent ARG Cluster90s (CphA, MCR-7, OXA-12) to *Aeromonas* spp. MAGs, highlighting
501 this group of bacteria as geographically widespread AMR reservoirs. The Proteobacteria order
502 and *Aeromonas* genus have both recently been shown to be core taxa in gut samples
503 extracted from laboratory-reared *Biomphalaria* snails (58); as well as signalling that snail
504 faecal microbiomes are a reflection of the intestinal tract, our results indicate that these core
505 features are preserved across various snail genera and within their natural habitats.
506 Meanwhile, our hierarchical clustering and PCoA visualisations demonstrate the variation in
507 numerous other bacterial taxa between snails of the same genus (*Bulinus*) observed between
508 countries, and between districts in the case of Malawi. The moderate diversity observed
509 between proximal samples from Malawian *Bulinus* snails at the species level highlight the
510 potential role of local environmental conditions, such as water pollution, climate, and antibiotic
511 exposure, in addition to host-specific characteristics, in shaping faecal microbiomes beyond
512 the core taxa.

513 We recovered eight high-quality MAGs and 21 medium-quality MAGs from 20 bacterial genera
514 in total, yet 1246 unique genera were detected in total by Kraken 2 read profiling,
515 demonstrating the limitations of assembly-based approaches in complex samples such as
516 these. It is therefore expected that fewer ARGs were detectable in MAGs, which are generally
517 recovered from high-abundance organisms, than by read mapping or by screening all contigs,
518 including those that could not be binned to MAGs. The inability to recover medium- to high-
519 quality MAGs (with >50% completeness) in 1226 of the 1246 genera detected by read profiling
520 (98.39%) suggests that some ARGs from low-abundance species are likely to have been
521 undetectable even by our read-based approach, particularly in samples with lower sequencing
522 depths. We specified that 60% template coverage was needed to call a hit to reduce the
523 chance of false positives, and hence the ARG profiles that we present here are likely biased
524 towards those found in high-abundance genera such as *Aeromonas*. This idea is supported
525 by our rarefaction analysis, in which varying sequencing depths were required to achieve ARG
526 saturation. Nonetheless, ARGs associated with abundant bacterial taxa were commonly
527 detected across the four samples with fewer than 10 million reads.

528 The detection of ARGs from the CphA, MCR-7 and OXA-12 Cluster90s in 66.7%, 66.7% and
529 93.3% of samples, respectively, all of which originate from *Aeromonas* species (59,60), is a
530 reflection of the skew towards the detection of ARGs associated with abundant taxa. The *mcr-*
531 *7.1* colistin resistance gene was first identified on plasmids in *K. pneumoniae* isolates from
532 chickens in China (61), but chromosome-borne variants of the gene have more recently been
533 identified in *Aeromonas* isolated from poultry in the United States (62), and in global
534 environmental sources (water) (63). The suggestion that *Aeromonas* species are the natural
535 reservoirs of *mcr-3* and *mcr*-7 was since proven by robust evolutionary models (53), and we
536 were able to assemble and bin *mcr*-7-like variants to a putative *Aeromonas jandaei* MAG. Our
537 results signpost to the potential for emergence of new variants of these ARGs from freshwater
538 gastropod faeces, and specifically, the *Aeromonas* reservoirs within.

539 Similar reservoirs of ARGs in low-abundance species may be more widespread than our data
540 indicate. We demonstrate the presence of $bla_{\text{OXA-181}}$ and $bla_{\text{OXA-547}}$ in samples from Malawi and
541 the United Kingdom, respectively, both of which are $bla_{\text{OXA-48}}$-like genes. In the case of $bla_{\text{OXA-}}$
542 $_{181}$, we could bin the > 20 kbp contig containing the gene to a high-quality MAG of *S.*
543 *xiamenensis*, the species known to be the gene's progenitor (51). *Shewanella* species have
544 been identified as the progenitors of numerous $bla_{\text{OXA-48}}$ variants (64), and the enzymes coded
545 for by these genes hydrolyse carbapenems and confer resistance to most beta-lactam, beta-
546 lactamase inhibitor combinations. They emerged and proliferated over the last two decades
547 to become the most prevalent carbapenemases amongst Enterobacterales in many parts of
548 Europe, Northern Africa, and the Middle East (46). Consequently, they are a major public

549 health concern, with carbapenemase-producing Enterobacterales (CPE) being listed by the
550 World Health Organisation as critical priority pathogens for which new antimicrobial agents
551 are urgently needed (65). While we cannot infer mobilisation of *bla*$_{OXA-48}$- or *mcr*-like genes
552 from environmental species to Enterobacterales in our study, the rapid dissemination of ARGs
553 co-localised with IS elements among various species within the Enterobacterales often occurs
554 as a result of one initial mobilisation event, after which transposable units can move between
555 replicons and cells (64,66).

556 To our knowledge, this is the first report of *bla*$_{OXA-181}$ in Malawi, a country where CPE are
557 uncommon in the clinical setting, but extended-spectrum beta-lactamase (ESBL)-producing
558 Enterobacterales resistant to a range of other beta-lactam antibiotics are, and contribute to
559 increased mortality in hospital patients (67,68). ESBL-producing Enterobacterales have also
560 been widely detected in water systems as well as human and animal faeces in Malawi (68,69).
561 Efforts to characterise the selective pressures behind the spread of ESBL-producing
562 Enterobacterales in Malawian wastewater and freshwater are ongoing, but low environmental
563 antimicrobial concentrations are known to drive the transfer and persistence of ARGs in water
564 systems (55). This selective pressure may also facilitate the emergence of *bla*$_{OXA-48}$-like genes
565 from environmental reservoirs where environmental and human health intersect. Meanwhile,
566 such reservoirs of carbapenem resistance genes could allow bacteria to react rapidly if
567 carbapenems are introduced more routinely into clinical practice, limiting their efficacy. The
568 emergence and spread of ARGs is often driven by horizontal gene transfer (HGT) associated
569 with MGEs, such as conjugative plasmids, which regularly carry multiple IS elements (3,70).
570 We show that IS elements are abundant within snail faecal microbiomes. However, the
571 suggestion of HGT is purely hypothetical in our context, without capturing widespread co-
572 localisation of ARGs on MGEs and performing functional assays. We could only demonstrate
573 co-localisation of ARGs and IS elements in two instances where *sul1*-containing gene clusters
574 were identified upstream of IS*6100* in samples from Malawi and Uganda. This is a common
575 issue in short-read metagenomic due to the fragmented nature of assemblies (71).

576 Our study does have several limitations. For example, the use of pooled snail faeces for
577 metagenomic analysis, while necessary to obtain sufficient biomass for DNA extraction and
578 sequencing, limits the ability to assess individual snail variability and the potential
579 heterogeneity in microbiome composition and ARG content in snails from the same location.
580 Meanwhile, the use of short-read sequencing, although effective for taxonomic profiling and
581 ARG identification, poses challenges in assembling complete genes and genomes,
582 particularly for low-abundance species, potentially leading to an underestimation of the full
583 richness of ARGs. This limitation is compounded by the inherent bias toward detecting ARGs
584 in more abundant taxa which may skew the representation of ARGs in less abundant

585 microbes. Ultra-seep sequencing may be required to detect such ARGs, even by read
586 mapping methods that do not rely on metagenome assembly. Additionally, while read-based
587 ARG detection approaches are widely used in the analysis of complex metagenomes to
588 enable the detection of low-abundance ARGs in the absence of assembly (24,33), partial hits
589 may represent truncated or non-functional genes. Inherently, short-read metagenome
590 assemblies from diverse samples are highly fragmented. This limited our ability to explore co-
591 localisation of ARGs and on plasmids and composite transposons. The accuracy of, and
592 access to, long-read metagenomics sequencing technologies is rapidly improving (71), and
593 these technologies could be used in future metagenomic studies to address the issues
594 associated with fragmented assemblies. Proximity ligation methods such as Hi-C could also
595 assist by linking MGEs to host genomes (72), though this was not a feasible addition to our
596 study due to limited sample biomass and the need for dedicated extraction protocols.
597 Furthermore, while this study highlights the potential role of freshwater snails as AMR
598 reservoirs, the cross-sectional nature of the sampling limits our ability to infer causality or the
599 dynamics of ARG transmission over time. Longitudinal studies are needed to better
600 understand temporal changes in ARG prevalence and the environmental factors influencing
601 these patterns. Finally, while the study identifies ARGs predicted or known to be of clinical
602 relevance, it does not explore the functional impact of these ARGs, nor does it confirm their
603 potential for HGT to human pathogens, a critical future step to quantify the risks associated
604 with our findings.

605 In conclusion, our study represents the first metagenomic analysis of AMR in freshwater
606 gastropod faeces, providing important insights into the potential roles of these invertebrates
607 as reservoirs of clinically relevant ARGs. Our findings have significant implications for AMR
608 surveillance within One Health frameworks, particularly in areas where agricultural animals,
609 wild animals, and humans utilise the same open water sources. This reservoir of ARGs poses
610 a threat to public health in an era of industrial-scale antibiotic use, particularly where access
611 to clean water and sanitation is limited. Future research should build on these results through
612 longitudinal sampling and by investigating how environmental factors, such as the presence
613 and concentration of antibiotics, influence ARG richness and abundance in snail microbiomes.
614 Meanwhile, incorporating long-read sequencing, proximity ligation methods, and functional
615 assays in future work, where feasible, will improve our understanding of ARG mobility.
616 Additional sampling in nearby human and animal populations to map potential transmission of
617 ARGs from snail-associated bacteria to human pathogens will be essential for designing
618 targeted interventions to curb AMR spread from and into shared aquatic environments.

619

620 **AUTHOR STATEMENTS**

621 **Author contributions**

622 The study was conceived and designed by AMO'F, JM, JRS and APR. Sampling was carried
623 out by AMO'F, AJ, SJ, PM, GM, SA, DO, SAK, JEL and JRS. Formal analysis was performed
624 by AMO'F, AF, JRS and APR. The original draft was prepared by AMO'F, JRS and APR, then
625 reviewed and edited by all authors. All authors approved the final version of the manuscript.

626 **Conflicts of interest**

627 None to declare.

628 **Funding**

629 AMO'F is funded by the Medical Research Council via the Liverpool School of Tropical
630 Medicine / Lancaster University doctoral training partnership (grant no.
631 MR/W007037/1). Fieldwork was supported by the Wellcome Trust within a Joint Investigator
632 Award (grant no. 220818/Z/20/Z) and by MÁEÖ (grant no. 2023-2024/183544). APR
633 acknowledges funding from the Medical Research Council, Biotechnology and Biological
634 Sciences Research Council and Natural Environmental Research Council which are all
635 Councils of UK Research and Innovation (grant no. MR/W030578/1) under the umbrella of the
636 JPIAMR (Joint Programming Initiative on Antimicrobial Resistance), and UKRI through the
637 Strength in Places Fund (grant no. SIPF 36348). The funders had no role in study design, data
638 collection and analysis, decision to publish, or preparation of the manuscript.

639 **Acknowledgements**

640 The authors thank Dr Simon Wagstaff and Mr Andrew Bennet for provision and maintenance
641 of the scientific computing environment at LSTM used to carry out bioinformatic analysis. The
642 authors acknowledge staff and local people in each of the sampling areas who aided the
643 sampling teams.

644

645

646

647

648

649

650 **REFERENCES**

651 1. Hutchings MI, Truman AW, Wilkinson B. Antibiotics: past, present and future. Curr
652 Opin Microbiol. 2019 Oct 1;51:72–80.

653 2. Tenover FC. Mechanisms of antimicrobial resistance in bacteria. Am J Med. 2006
654 Jun;119(6 Suppl 1):S3-10; discussion S62-70.

655 3. Despotovic M, de Nies L, Busi SB, Wilmes P. Reservoirs of antimicrobial resistance
656 in the context of One Health. Curr Opin Microbiol. 2023 Jun;73:102291.

657 4. Rockström J, Steffen W, Noone K, Persson Å, Chapin FS, Lambin EF, et al. A safe
658 operating space for humanity. Nature. 2009 Sep;461(7263):472–5.

659 5. Kirchhelle C, Roberts AP. Beyond breakpoint – reconceptualising AMR as a symptom
660 of planetary stress. Npj Antimicrob Resist. 2025 Jun 19;3(1):57.

661 6. Colley DG, Bustinduy AL, Secor WE, King CH. Human schistosomiasis. Lancet. 2014
662 Jun 28;383(9936):2253–64.

663 7. Caravedo MA, Cabada MM. Human Fascioliasis: Current Epidemiological Status and
664 Strategies for Diagnosis, Treatment, and Control. Res Rep Trop Med. 2020 Nov 26;11:149–
665 58.

666 8. Liu H, Macdonald CA, Cook J, Anderson IC, Singh BK. An Ecological Loop: Host
667 Microbiomes across Multitrophic Interactions. Trends Ecol Evol. 2019 Dec;34(12):1118–30.

668 9. Chevalier FD, Diaz R, McDew-White M, Anderson TJC, Le Clec'h W. The hemolymph
669 of *Biomphalaria* snail vectors of schistosomiasis supports a diverse microbiome. Environ
670 Microbiol. 2020 Dec;22(12):5450–66.

671 10. Schols R, Vanoverberghe I, Huyse T, Decaestecker E. Host-bacteriome transplants
672 of the schistosome snail host *Biomphalaria glabrata* reflect species-specific associations.
673 FEMS Microbiol Ecol. 2023 Sep 1;99(9):fiad101.

674 11. Madsen H, Frandsen F. The spread of freshwater snails including those of medical
675 and veterinary importance. Acta Trop. 1989 May;46(3):139–46.

676 12. Lv SB, He TT, Hu F, Li YF, Yuan M, Xie JZ, et al. The Impact of Flooding on Snail
677 Spread: The Case of Endemic Schistosomiasis Areas in Jiangxi Province, China. Trop Med
678 Infect Dis. 2023 Apr 30;8(5):259.

679 13. Alharbi MH, Condemine C, Christiansen R, LaCourse EJ, Makaula P, Stanton MC, et
680 al. *Biomphalaria pfeifferi* Snails and Intestinal Schistosomiasis, Lake Malawi, Africa, 2017–
681 2018. Emerg Infect Dis. 2019 Mar;25(3):613–5.

682 14. Pedersen UB, Stendel M, Midzi N, Mduluza T, Soko W, Stensgaard AS, et al.
683 Modelling climate change impact on the spatial distribution of fresh water snails hosting
684 trematodes in Zimbabwe. Parasit Vectors. 2014 Dec;7(1):1–12.

685 15. Kock KD, Wolmarans CT, Bornman M. Distribution and habitats of *Biomphalaria*
686 *pfeifferi*, snail intermediate host of *Schistosoma mansoni*, in South Africa. Water SA. 2004
687 Aug 3;30(1):29–36.

688 16. O'Ferrall AM, Musaya J, Stothard JR, Roberts AP. Aligning antimicrobial resistance
689 surveillance with schistosomiasis research: an interlinked One Health approach. Trans R
690 Soc Trop Med Hyg. 2024 Aug 1;118(8):498–504.

691 17. Murray CJ, Ikuta KS, Sharara F, Swetschinski L, Aguilar GR, Gray A, et al. Global
692 burden of bacterial antimicrobial resistance in 2019: a systematic analysis. Lancet. 2022 Feb
693 12;399(10325):629–55.

694 18. Olaru ID, Walther B, Schaumburg F. Zoonotic sources and the spread of
695 antimicrobial resistance from the perspective of low and middle-income countries. Infect Dis
696 Poverty. 2023 Jun 14;12(1):59.

697 19. Musicha P, Morse T, Cocker D, Mugisha L, Jewell CP, Feasey NA. Time to define
698 One Health approaches to tackling antimicrobial resistance. Nat Commun. 2024 Oct
699 10;15(1):8782.

700 20. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics,
701 from sampling to analysis. Nat Biotechnol. 2017 Sep;35(9):833–44.

702 21. Almeida A, Mitchell AL, Boland M, Forster SC, Gloor GB, Tarkowska A, et al. A new
703 genomic blueprint of the human gut microbiota. Nature. 2019 Apr;568(7753):499–504.

704 22. Kim CY, Lee M, Yang S, Kim K, Yong D, Kim HR, et al. Human reference gut
705 microbiome catalog including newly assembled genomes from under-represented Asian
706 metagenomes. Genome Med. 2021 Aug 27;13(1):134.

707 23. Howe A, Chain PSG. Challenges and opportunities in understanding microbial
708 communities with metagenome assembly (accompanied by IPython Notebook tutorial). Front
709 Microbiol. 2015 Jul 9;6:678.

710 24. Lee K, Raguideau S, Sirén K, Asnicar F, Cumbo F, Hildebrand F, et al. Population-
711 level impacts of antibiotic usage on the human gut microbiome. Nat Commun. 2023 Mar
712 2;14(1):1191.

713 25. Carruthers LV, Nordmeyer SC, Anderson TJc, Chevalier FD, Le Clec'h W. Organ-
714 Specific Microbiomes of *Biomphalaria* Snails [Internet]. 2024 [cited 2024 Aug 19]. Available
715 from: http://biorxiv.org/lookup/doi/10.1101/2024.06.11.598555

716 26. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina
717 sequence data. Bioinformatics. 2014 Aug 1;30(15):2114–20.

718 27. Lu J, Rincon N, Wood DE, Breitwieser FP, Pockrandt C, Langmead B, et al.
719 Metagenome analysis using the Kraken software suite. Nat Protoc. 2022 Dec;17(12):2815–
720 39.

721 28. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2.
722 Genome Biol. 2019 Nov 28;20(1):257.

723 29. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance
724 in metagenomics data. PeerJ Comput Sci. 2017 Jan 2;3:e104.

725 30. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, et al. ResFinder
726 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother. 2020 Dec
727 1;75(12):3491–500.

728 31. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-
729 generation sequencing data. Bioinforma Oxf Engl. 2012 Dec 1;28(23):3150–2.

730 32. Clausen PTLC, Aarestrup FM, Lund O. Rapid and precise alignment of raw reads
731 against redundant databases with KMA. BMC Bioinformatics. 2018 Aug 29;19(1):307.

732 33. Lal Gupta C, Kumar Tiwari R, Cytryn E. Platforms for elucidating antibiotic resistance
733 in single genomes and complex metagenomes. Environ Int. 2020 May 1;138:105667.

734 34. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile
735 metagenomic assembler. Genome Res. 2017 May;27(5):824–34.

736 35. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for
737 genome assemblies. Bioinforma Oxf Engl. 2013 Apr 15;29(8):1072–5.

738 36. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-
739 MEM. ArXiv. 2013 Mar 16;1303.

740 37. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: an adaptive
741 binning algorithm for robust and efficient genome reconstruction from metagenome
742 assemblies. PeerJ. 2019 Jul 26;7:e7359.

743 38. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing
744 the quality of microbial genomes recovered from isolates, single cells, and metagenomes.
745 Genome Res. 2015 Jul;25(7):1043.

746 39. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK,
747 et al. Minimum information about a single amplified genome (MISAG) and a metagenome-
748 assembled genome (MIMAG) of bacteria and archaea. Nat Biotechnol. 2017 Aug
749 8;35(8):725–31.

750 40. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify
751 genomes with the Genome Taxonomy Database. Bioinformatics. 2020 Mar 15;36(6):1925–7.

752 41. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. ISfinder: the reference
753 centre for bacterial insertion sequences. Nucleic Acids Res. 2006 Jan 1;34(suppl_1):D32–6.

754 42. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. Bakta:
755 rapid and standardized annotation of bacterial genomes via alignment-free sequence
756 identification. Microb Genomics. 2021;7(11):000685.

757 43. Gilchrist CLM, Chooi YH. clinker & clustermap.js: automatic generation of gene
758 cluster comparison figures. Bioinforma Oxf Engl. 2021 Aug 25;37(16):2473–5.

759 44. Huang R, Soneson C, Ernst FGM, Rue-Albrecht KC, Yu G, Hicks SC, et al.
760 TreeSummarizedExperiment: a S4 class for data with hierarchical structure.
761 F1000Research. 2021 Mar 2;9:1246.

762 45. Calle ML. Statistical Analysis of Metagenomics Data. Genomics Inform. 2019 Mar
763 31;17(1):e6.

764 46. Boyd SE, Holmes A, Peck R, Livermore DM, Hope W. OXA-48-Like β-Lactamases:
765 Global Epidemiology, Treatment Options, and Development Pipeline. Antimicrob Agents
766 Chemother. 2022 Jul 20;66(8):e00216-22.

767 47. Figueiredo S, Bonnin RA, Poirel L, Duranteau J, Nordmann P. Identification of the
768 naturally occurring genes encoding carbapenem-hydrolysing oxacillinases from
769 *Acinetobacter haemolyticus*, *Acinetobacter johnsonii*, and *Acinetobacter calcoaceticus*. Clin
770 Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis. 2012 Sep;18(9):907–13.

771 48. Nielsen TK, Browne PD, Hansen LH. Antibiotic resistance genes are differentially
772 mobilized according to resistance mechanism. GigaScience. 2022 Jan 1;11:giac072.

773 49. Dao TD, Takemura T, Kasuga I, Hirabayashi A, Nga NT, Anh PHQ, et al. Emergence
774 of plasmid-mediated RND-type efflux pump gene cluster *tmexCD-toprJ* in *Shewanella*
775 *xiamenensis* in a water environment [Internet]. 2022 [cited 2024 Jul 16]. Available from:
776 http://biorxiv.org/lookup/doi/10.1101/2022.09.15.508154

777 50. Potron A, Nordmann P, Lafeuille E, Maskari Z, Rashdi F, Poirel L. Characterization of
778 OXA-181, a Carbapenem-Hydrolyzing Class D -Lactamase from *Klebsiella pneumoniae*.
779 Antimicrob Agents Chemother. 2011 Sep 15;55:4896–9.

780 51. Potron A, Poirel L, Nordmann P. Origin of OXA-181, an Emerging Carbapenem-
781 Hydrolyzing Oxacillinase, as a Chromosomal Gene in *Shewanella xiamenensis*. Antimicrob
782 Agents Chemother. 2011 Sep;55(9):4405–7.

783 52. Samantha A, Vrielink A. Lipid A Phosphoethanolamine Transferase: Regulation,
784 Structure and Immune Response. J Mol Biol. 2020 Aug 21;432(18):5184–96.

785 53. Guo Y, Zou G, Kerdsin A, Schultsz C, Hu C, Bei W, et al. Characterization of NMCR-
786 3, NMCR-4 and NMCR-5, three novel non-mobile colistin resistance determinants:
787 Implications for MCR-3, MCR-7, and MCR-5 progenitors, respectively. Drug Resist Updat.
788 2024 Jul 1;75:101088.

789 54. Kowalski M, Obama BM, Catho G, Dewez JE, Merglen A, Ruef M, et al. Antimicrobial
790 resistance in Enterobacterales infections among children in sub-Saharan Africa: a systematic
791 review and meta-analysis. eClinicalMedicine [Internet]. 2024 Apr 1 [cited 2024 Dec 12];70.
792 Available from: https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370(24)00091-
793 9/fulltext

794 55. Murray AK, Stanton IC, Tipper HJ, Wilkinson H, Schmidt W, Hart A, et al. A critical
795 meta-analysis of predicted no effect concentrations for antimicrobial resistance selection in
796 the environment. Water Res. 2024 Nov 15;266:122310.

797 56. Silva TM, Melo ES, Lopes ACS, Veras DL, Duarte CR, Alves LC, et al.
798 Characterization of the bacterial microbiota of *Biomphalaria glabrata* (Say, 1818) (Mollusca:
799 Gastropoda) from Brazil. Lett Appl Microbiol. 2013 Jul;57(1):19–25.

800 57. Hu Z, Tong Q, Chang J, Yu J, Li S, Niu H, et al. Gut bacterial communities in the
801 freshwater snail *Planorbella trivolvis* and their modification by a non-herbivorous diet. PeerJ.
802 2021 Feb 12;9:e10716.

803 58. Li P, Hong J, Wu M, Yuan Z, Li D, Wu Z, et al. Metagenomic Analysis Reveals
804 Variations in Gut Microbiomes of the *Schistosoma mansoni*-Transmitting Snails *Biomphalaria*
805 *straminea* and *Biomphalaria glabrata*. Microorganisms. 2023 Sep 28;11(10):2419.

806 59. Chen PL, Ko WC, Wu CJ. Complexity of β-lactamases among clinical *Aeromonas*
807 isolates and its clinical implications. J Microbiol Immunol Infect. 2012 Dec 1;45(6):398–403.

808 60. Alksne LE, Rasmussen BA. Expression of the AsbA1, OXA-12, and AsbM1 beta-
809 lactamases in *Aeromonas jandaei* AER 14 is coordinated by a two-component regulon. J
810 Bacteriol. 1997 Mar;179(6):2006–13.

811 61. Yang YQ, Li YX, Lei CW, Zhang AY, Wang HN. Novel plasmid-mediated colistin
812 resistance gene *mcr-7.1* in *Klebsiella pneumoniae*. J Antimicrob Chemother. 2018 Jul
813 1;73(7):1791–5.

62. Wang Y, Hou N, Rasooly R, Gu Y, He X. Prevalence and Genetic Analysis of Chromosomal *mcr-3/7* in *Aeromonas* From U.S. Animal-Derived Samples. Front Microbiol. 2021 Apr 30;12:667406.

63. Khedher MB, Baron SA, Riziki T, Ruimy R, Raoult D, Diene SM, et al. Massive analysis of 64,628 bacterial genomes to decipher water reservoir and origin of mobile colistin resistance genes: is there another role for these enzymes? Sci Rep. 2020 Apr 6;10(1):5970.

64. Tacão M, Araújo S, Vendas M, Alves A, Henriques I. *Shewanella* species as the origin of *bla*OXA-48 genes: insights into gene diversity, associated phenotypes and possible transfer mechanisms. Int J Antimicrob Agents. 2018 Mar 1;51(3):340–8.

65. World Health Organisation. WHO publishes list of bacteria for which new antibiotics are urgently needed [Internet]. 2017 [cited 2023 Jun 15]. Available from: https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed

66. Wang R, van Dorp L, Shaw LP, Bradley P, Wang Q, Wang X, et al. The global distribution and spread of the mobilized colistin resistance gene mcr-1. Nat Commun. 2018 Mar 21;9(1):1179.

67. Lewis JM, Mphasa M, Banda R, Beale MA, Heinz E, Mallewa J, et al. Colonization dynamics of extended-spectrum beta-lactamase-producing Enterobacterales in the gut of Malawian adults. Nat Microbiol. 2022 Oct;7(10):1593–604.

68. Lester R, Musicha P, Kawaza K, Langton J, Mango J, Mangochi H, et al. Effect of resistance to third-generation cephalosporins on morbidity and mortality from bloodstream infections in Blantyre, Malawi: a prospective cohort study. Lancet Microbe. 2022 Dec 1;3(12):e922–30.

69. Cocker D, Chidziwisano K, Mphasa M, Mwapasa T, Lewis JM, Rowlingson B, et al. Investigating One Health risks for human colonisation with extended spectrum β-lactamase-producing *Escherichia coli* and *Klebsiella pneumoniae* in Malawian households: a longitudinal cohort study. Lancet Microbe. 2023 Jul 1;4(7):e534–43.

70. Che Y, Yang Y, Xu X, Břinda K, Polz MF, Hanage WP, et al. Conjugative plasmids interact with insertion sequences to shape the horizontal transfer of antimicrobial resistance genes. Proc Natl Acad Sci. 2021 Feb 9;118(6):e2008731118.

71. Agustinho DP, Fu Y, Menon VK, Metcalf GA, Treangen TJ, Sedlazeck FJ. Unveiling microbial diversity: harnessing long-read sequencing technology. Nat Methods. 2024 Jun;21(6):954–66.

72. Beitel CW, Froenicke L, Lang JM, Korf IF, Michelmore RW, Eisen JA, et al. Strain- and plasmid-level deconvolution of a synthetic metagenome by sequencing proximity ligation products. PeerJ. 2014 May 27;2:e415.