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

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 metagenomic sequencing study. Proteobacteria dominated the faecal microbiomes, within which Aeromonadales were strongly associated with the total ARG load. We detected putative ARGs conferring resistance to 13 distinct antibiotic classes, including carbapenem resistance genes in all countries. In Malawi, we linked $bla_{\text{OXA-181}}$ ($bla_{\text{OXA-48}}$-like), a putatively mobile carbapenemase gene, to *Shewanella xiamenensis*—the gene's known progenitor—in a metagenome-assembled genome. We detected another $bla_{\text{OXA-48}}$-like variant by read-mapping from a sample in the United Kingdom and identified mobile colistin resistance (*mcr*)-like genes at 11 of 15 locations. Our findings highlight freshwater snail microbiomes as One Health-relevant, environmental reservoirs of clinically important ARGs.

Keywords

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

58 **MAIN TEXT**

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

69 Outside the context of AMR, freshwater gastropod snails are already of interest in a One
70 Health context. Snails are globally distributed in aquatic environments, often playing critical
71 roles as obligate intermediate hosts in the transmission of neglected tropical diseases when
72 freshwater sources are contaminated with human and animal excreta. For instance, snails of
73 the *Bulinus*, *Biomphalaria*, and *Oncomelania* genera are key in the transmission of
74 schistosomiasis (5), while *Lymnaea* snails transmit fascioliasis (6). Recent efforts to
75 characterize the microbiomes of medically important freshwater snails are generally motivated
76 by the growing recognition that host-associated microorganisms significantly influence host-
77 parasite interactions (7). The microbiomes of *Biomphalaria* snails have been shown by 16S
78 taxonomic profiling to be diverse and distinct from those of their freshwater environments (8),
79 yet bacteriome transplant experiments have demonstrated the potential for these microbiomes
80 to change depending on the surrounding environmental bacteria (9). Freshwater snail
81 populations are highly mobile (10), capable of translocating due to changes in ecosystems
82 caused by flooding or the formation of artificial water bodies (11,12), and are distributed across
83 a range of human-influenced areas (13,14). Considering that the environments where
84 freshwater snails graze—typically on detritus—can be contaminated with AMR enteric bacteria
85 at sites of interaction with humans and animals, we wanted to explore the potential role of
86 snails as widespread and persistent environmental reservoirs of ARGs (15). This may be
87 particularly relevant in Low- and Middle-Income Countries (LMICs), where AMR is
88 disproportionately linked to increased mortality (16), and where poor environmental health
89 infrastructure is linked to AMR transmission across various One Health compartments (17,18).

90 The analysis of metagenomic sequence data is increasingly being used in the study of
91 complex microbial samples and AMR, made possible by advances in next-generation
92 sequencing and computational science technologies in recent years (19). Metagenomic

studies are not reliant on culture steps that alter microbial communities and are therefore able to yield non-selective data on sample composition (20). 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 (21,22). 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 (23). When combining read-based profiling with assembly-based approaches that uncover the genomic context of AMR—linking 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 metagenomic 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 (*Supplementary Table 1*). Acknowledging that freshwater snail microbiomes vary by organ (24), 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 utilize 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.

Results

Sample characteristics

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. 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 in Zanzibar (one each of *Bulinus*, *Cleopatra*, *Lanistes*, and *Pila*). In Europe, two locations in the United Kingdom were sampled (one *Anisus* and one *Lymnaea*). At each

128 location, groups of 50 snails from the identified genus were gathered ahead of faecal
129 harvesting. For downstream analysis, we pooled the faeces of each group of snails to obtain
130 enough biomass for DNA extraction and Illumina short-read sequencing, resulting in 15
131 metagenomes, each representing a specific snail genus at a specific location (*Supplementary*
132 *Table 1*). The number of 150 bp Illumina paired end reads per sample ranged from 9,340,457
133 to 35,991,622.

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

136 After read processing and quality control, reads classified as bacterial by Kraken2 underwent
137 species abundance estimation with Bracken (25). Proteobacteria was the most abundant
138 phylum observed, with a mean relative abundance of 66.27% (*Figure 1a*). Enterobacterales,
139 an order within the Proteobacteria containing multiple species of clinical concern due to their
140 ability to acquire and exchange ARGs, were well represented, with a mean relative abundance
141 of 4.79%. The Uk1 sample was a notable outlier in which the relative abundance of
142 Enterobacterales was estimated to be 41.70%. When excluding Uk1, the mean relative
143 abundance of Enterobacterales dropped to 2.15%. Four orders were more abundant than
144 Enterobacterales (Burkholderiales, Rhizobiales, Aeromonadales, and Pseudomonadales) all
145 of which had a mean relative abundance of greater than 5% (*Figure 1b*). All samples exhibited
146 high community (alpha) diversity. Shannon's diversity index ranged from 3.881 (M5) – 7.372
147 (Z2) (mean = 6.474, median = 6.888). Simpson's diversity index, in which scores from 0 – 1
148 are possible (1 = maximum possible diversity), ranged from 0.793 (M5) – 0.999 (Z2) (mean =
149 0.972, median = 0.994). We observed that alpha diversity (Shannon's diversity index) was
150 negatively corelated with the relative abundance of Proteobacteria (Pearson correlation
151 coefficient = -0.600; p = 0.0184).

152 Relative abundance data were aggregated at the order level, subjected to centred log-ratio
153 (CLR) transformation, and subsequently standardized (CLR-z) to facilitate visual comparisons
154 between taxa with varying absolute abundances (*Figure 1c*). From the 160 bacterial orders
155 identified across all samples, the 50 most abundant orders, accounting for 97.92% of the total
156 microbiome, were selected for hierarchical clustering and visualization. Samples from the
157 same administrative region clustered closely together, with the exception of Malawian
158 samples, which were split into two distinct groups. This split corresponded to the two districts
159 where the samples were collected: samples M1, M2, M3, and M4, collected in Mangochi
160 District on the southern shorelines of Lake Malawi, clustered together, while samples M5 and
161 M6, collected in Nsanje District, over 275 km from Mangochi District sampling sites, formed a
162 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. No clear trend in taxonomic profiles was observed by snail genus. 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.

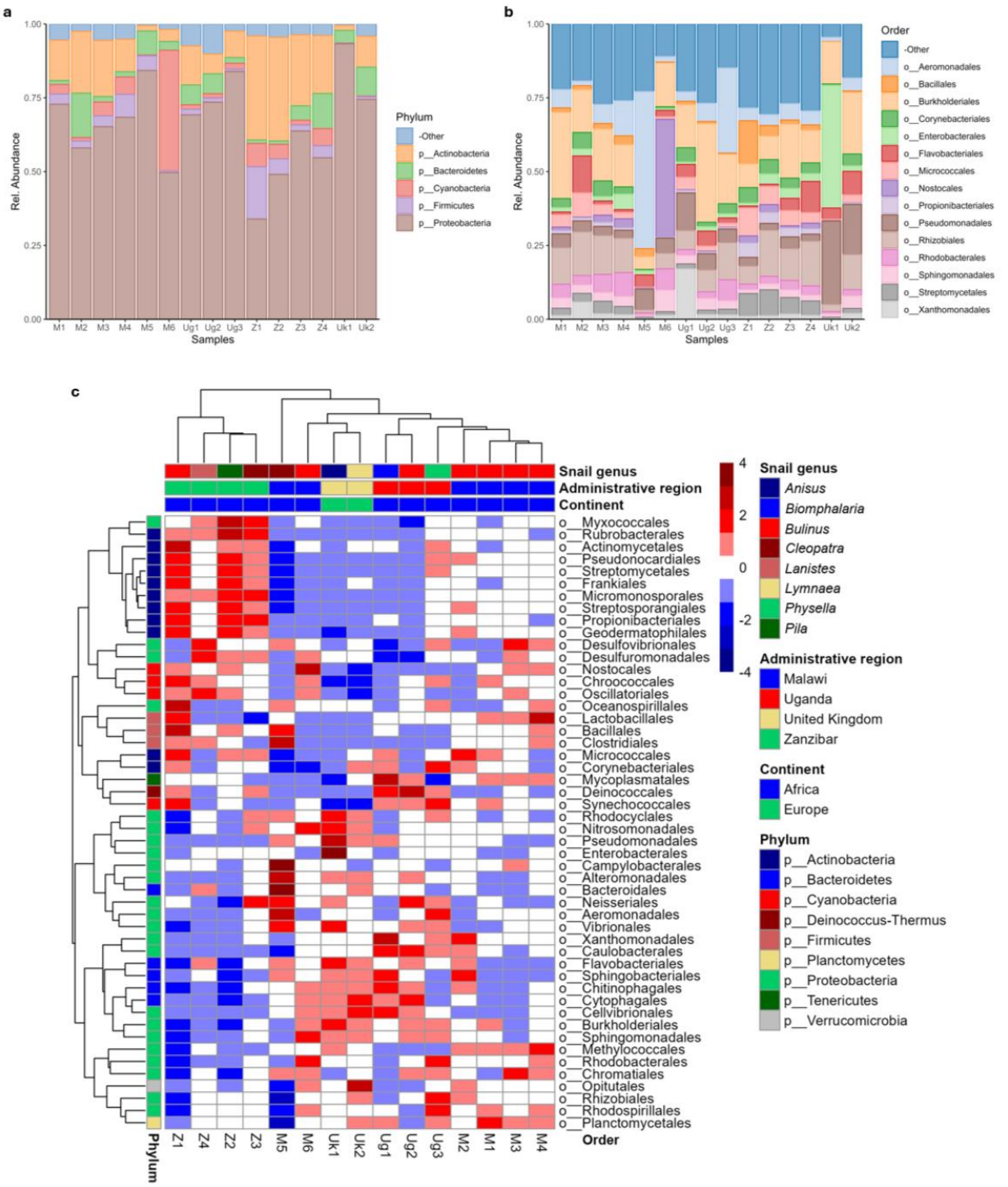


Figure 1. Exploration and comparison of bacterial taxa within freshwater snail faeces: a) Relative abundance of the top phyla; b) Relative abundance of the top orders; c) Clustered

172 *heatmap of standardised order abundances (CLR-z) across snail hosts and geographic*
173 *locations. Column and row trees cluster samples and taxa, respectively, based on a*
174 *hierarchical clustering approach.*

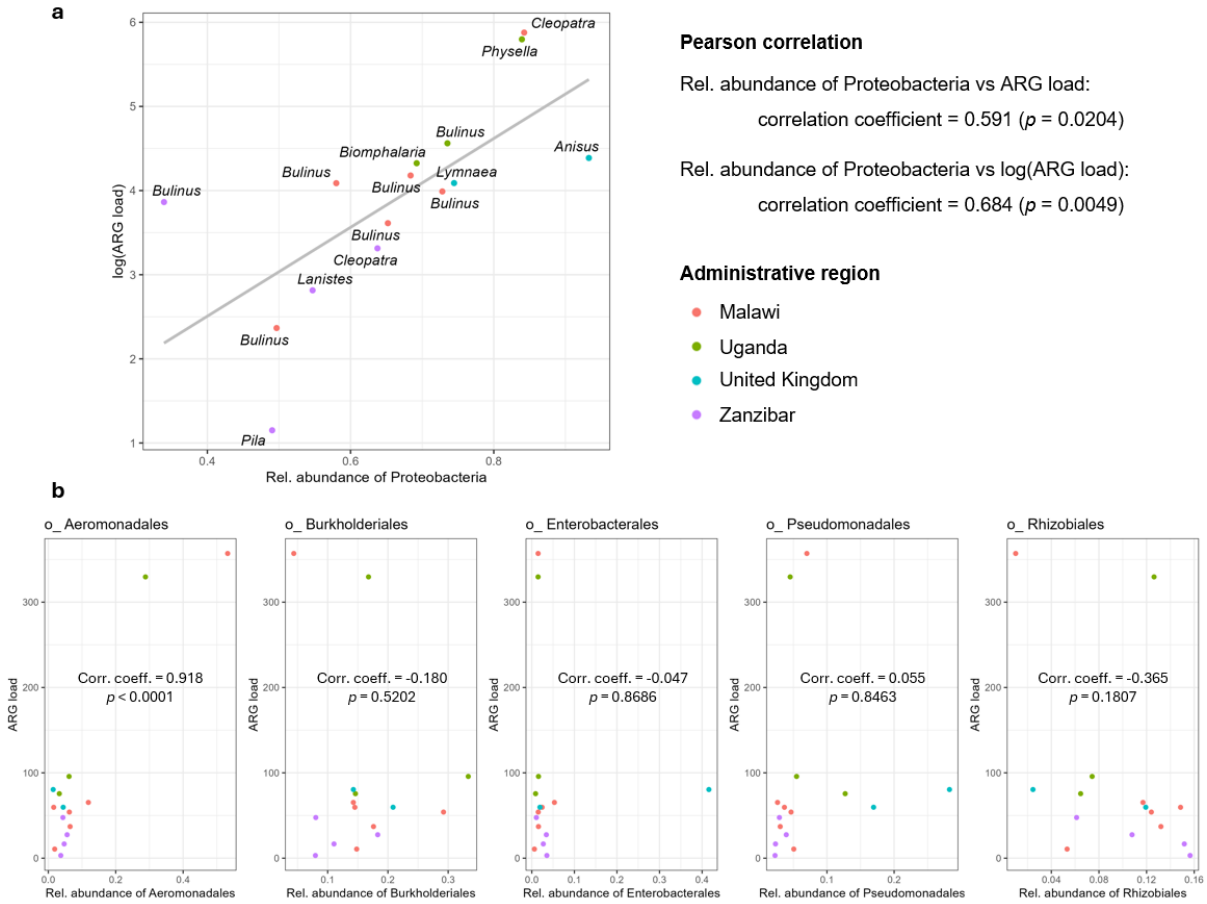
175 When investigating at species level, a Bray-Curtis dissimilarity (beta diversity) matrix
176 (*Supplementary Table 2*) showed a range of divergence between Malawian *Bulinus* snails. We
177 observed moderate-high divergence (0.649) between two samples (M2, M4) collected a little
178 over five kilometres from each other, both from stagnant water bodies close to the Lake Malawi
179 shoreline, suggesting that local environmental factors play a key role in influencing the
180 microbiome.

181 **Proteobacteria, specifically Aeromonadales, drive the ARG load**

182 Using a read-based k-mer alignment (KMA) method to screen for ARGs against the ResFinder
183 database (26,27), reads mapped to 54 unique ARGs in total (minimum 90% query identity and
184 60% template coverage), belonging to 41 ARG Cluster90s (in which reference genes share
185 minimum 90% nucleotide identity) (23). ARG diversity (the number of unique ARGs identified
186 by read mapping in any given sample) ranged from 1 – 16 (mean = 7.07, median = 6). We
187 detected fewer ARGs by screening metagenome assemblies against the ResFinder database
188 (mean = 2.53, median = 2) (*Supplementary Figure 1*). Across all samples, 10 of the 38 ARGs
189 (26.3%) assembled on contiguous sequences (contigs) could be binned to high-quality (>90%
190 completeness, <5% contamination) or medium-quality (>50% completeness, <10%
191 contamination) metagenome-assembled genomes (MAGs). These ARGs were all binned to
192 MAGs classified as Proteobacteria (genera: *Aeromonas*, n = 7; *Shewanella*, n = 1;
193 *Stenotrophomonas*, n = 2) (*Figure 5*).

194 We calculated the normalised abundance of each ARG from read-mapping data, expressed
195 as RPKM (Reads Per Kilobase per Million mapped bacterial reads), then summed the RPKM
196 of all ARGs to calculate total ARG load per sample. ARG load ranged from 3.16 – 357.04
197 RPKM (mean = 87.99, median = 59.62). The relationship observed between relative
198 abundance of Proteobacteria (the predominant bacterial order) and ARG load appeared to be
199 non-linear and was linearised with a log transformation of ARG load (*Supplementary Figure*
200 *2*). Proteobacteria relative abundance was positively correlated with log-transformed ARG
201 load in faecal microbiomes (Pearson correlation coefficient = 0.684; p = 0.0049) (*Figure 2a*).
202 A significant association between Proteobacteria relative abundance and ARG diversity was
203 not detected (Pearson correlation coefficient = 0.122; p = 0.6654). Among the 20 orders of
204 bacteria with a mean relative abundance above 1%, Aeromonadales (belonging to the
205 Proteobacteria) were strongly and positively correlated with ARG load (Pearson correlation

206 coefficient = 0.918; $p < 0.0001$) (*Figure 2b*), while no other orders were associated with either
207 positive or negative correlation (*Supplementary Table 3*).



208

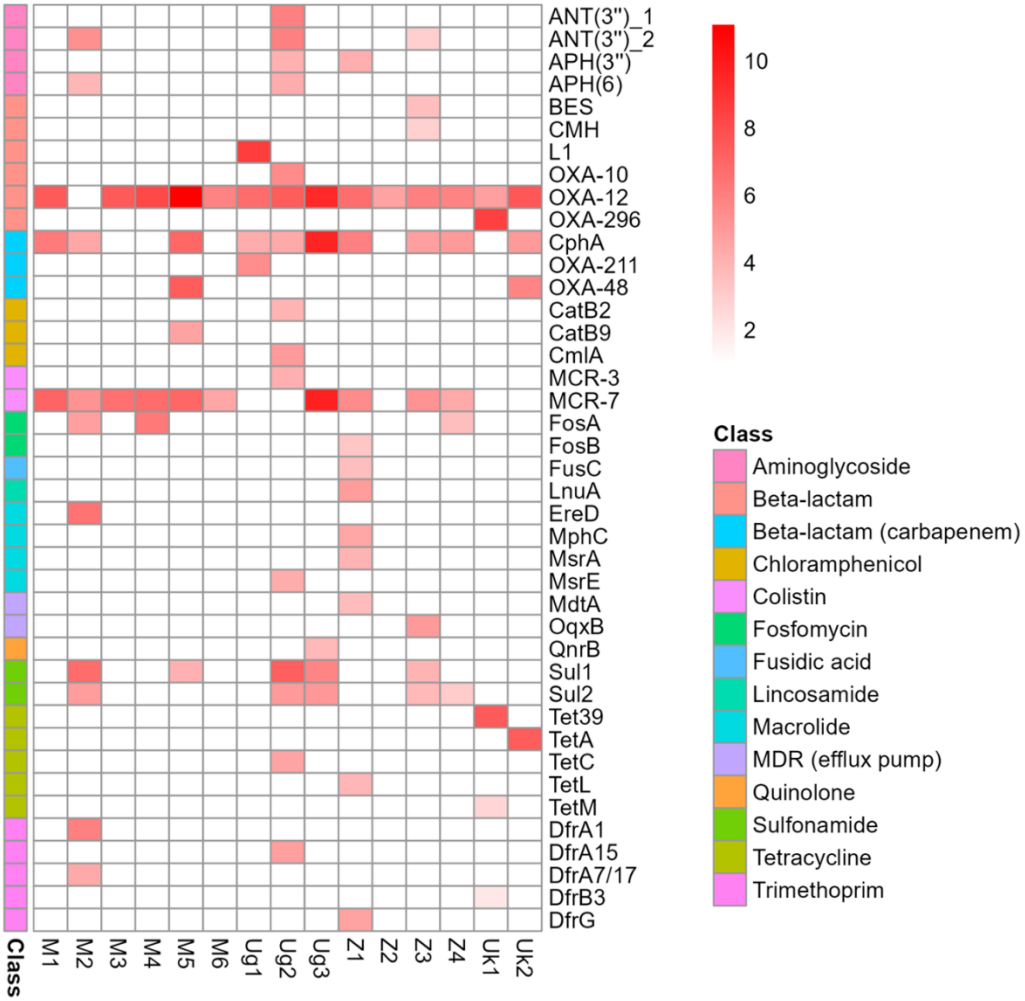
209 *Figure 2. Associations between Proteobacteria and ARG load: a) Relative abundance of*
210 *Proteobacteria versus log(ARG load) with a linear regression line and associated Pearson*
211 *correlation coefficients indicating improved fit of data following log-transformation of ARG load;*
212 *b) Relative abundance of the five most abundant orders within Proteobacteria versus ARG*
213 *load with associated Pearson correlation coefficients indicating that Aeromonadales contribute*
214 *significantly to ARG load within the Proteobacteria. Within-plot Labels in part (a) are snail*
215 *genera. Relationships between relative abundance and ARG load in part (b) either appeared*
216 *linear (Aeromonadales) or non-correlated (others), hence transformations were not applied.*

217 **ARGs conferring resistance to first- and last-line antibiotics detected**

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

224 and quinolones (0.03%) (*Supplementary Figure 3*). Efflux pump genes conferring multidrug
225 resistance contributed 1.16% of the total ARG load.

226 Beta-lactamase genes from the OXA-12 Cluster90 were the most prevalent in our samples
227 and were detected in fourteen of the fifteen metagenomes (*Figure 3*). We also detected ARGs
228 from the CphA Cluster90, a family of carbapenem-hydrolyzing metallo-beta-lactamase genes,
229 in ten of the fifteen samples. Three OXA-carbapenemase genes were identified, each from a
230 different country. These included two ARGs from the OXA-48 Cluster90 ($bla_{OXA-181}$ and bla_{OXA-}
231 $_{547}$, in Malawi and the United Kingdom, respectively) of concern globally due to plasmid-borne
232 dissemination amongst Enterobacterales (28), and $bla_{OXA-212}$ (Uganda) from the OXA-211
233 Cluster90 of carbapenemase genes that naturally occur in *Acinetobacter johnsonii* (29). We
234 also detected putative colistin resistance genes, predominantly variants of the *mcr-7.1* gene,
235 in eleven metagenomes.



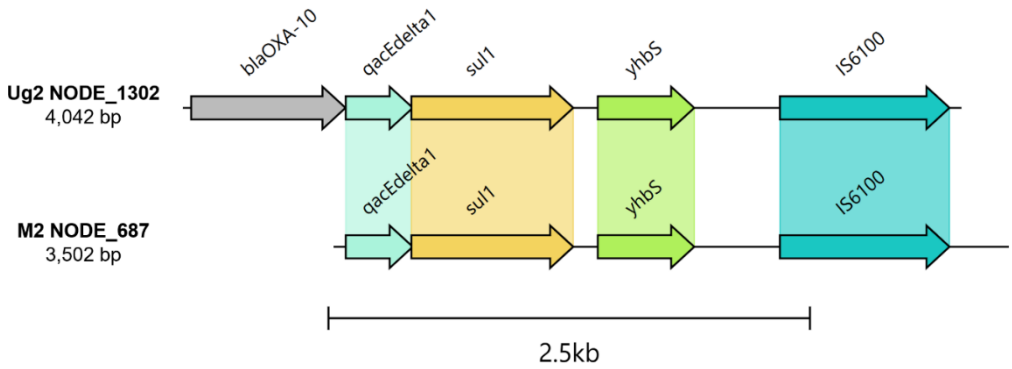
236

237 *Figure 3. Heatmap of normalized ARG Cluster90 abundances across samples, expressed as*
238 *log(RPKM). Cluster90s are colour-coded by the antibiotic class that they are considered to*
239 *confer resistance against.*

240 **MGEs in snail faecal microbiomes are associated with ARGs**

241 To investigate the relationship between MGEs and ARGs in metagenome assemblies, we
242 screened all contigs for insertion sequence (IS) elements (ISfinder database (30)) in addition
243 to ARGs (ResFinder database (27)). We identified two contigs containing IS elements and
244 ARGs, neither of which could be binned to a putative MAG. A 4042 bp contig (Ug2
245 NODE_1302: *Bulinus* spp. faeces, Lake Victoria, Uganda) contained *sul1* (sulfonamide
246 resistant dihydropteroate synthase) and bla_{OXA-10} (beta-lactamase) genes, on which both
247 ARGs were located upstream of IS*6100*. The gene cluster also contained *qacEdelta* (a gene
248 conferring resistance to antiseptics) and *yhbS* (putative N-acetyltransferase). The contig
249 displays 100% query cover with one complete genome in the GenBank Nucleotide database—
250 *Aeromonas hydrophila* (accession no. AP025277)—with 99.97% identity, but also maintains
251 high query cover (> 95%) and identity (> 99%) with chromosomal and plasmid sequences from
252 various other bacteria (*Pseudomonas aeruginosa*, *Vibrio cholerae*, *Morganella morganii*,
253 *Salmonella enterica*, *Klebsiella pneumoniae*). A second, shorter, 3505 bp contig (M2
254 NODE_687: *Bulinus* spp. faeces, Mangochi District, Malawi), aligned with 99.94% nucleotide
255 identity to Ug NODE_1302 over the region containing *qacEdelta, sul1* and *yhbS* upstream of
256 IS*6100* (*Figure 4*), with a single nucleotide polymorphism (SNP) in the *sul1* gene. The similarity
257 observed between the two contigs from Ug2 and M2, and with sequences from a range of
258 bacterial species in GenBank, demonstrates the widespread nature of certain ARGs
259 associated with MGEs, which we show can be found in freshwater snail faeces.

260



261 *Figure 4: Gene cluster comparison between Ug2 NODE_1302 and M2 NODE_687 containing*
262 *bla_{OXA-10} (Ug2 only), qacEdelta1, sul1 and yhbS upstream of IS6100. Regions coding for*
263 *hypothetical proteins are not displayed.*

264 We detected between 8 and 97 putative IS elements per sample (mean = 37.8, median = 29).
265 However, we could only link these directly to ARGs on contigs in the above two examples.
266 Searches for ARGs associated with IS elements in transposons should ideally capture 24.34
267 kbp of DNA (based on the length of composite and unit transposons in The Transposon

268 Registry) (31). Due to the limitations in assembling large contigs in diverse metagenomic
269 samples (22), and the associated inability to capture all ARGs and MGEs on contigs (the
270 median length of contigs containing ARGs in our dataset is 2.10 kbp), we constructed a
271 generalized linear model (*Supplementary Information 1*) with a negative binomial distribution
272 to model the relationship and show associations where MGEs and ARGs could not be directly
273 linked by assembly. The choice of negative binomial distribution was made to account for
274 overdispersion in the ARG count data, which is common in datasets involving count outcomes
275 such as this where the variance in the number of ARGs assembled per sample (4.12) exceeds
276 the mean (2.53).

277 *Table 1: Outputs of a negative binomial generalized linear model exploring associations*
278 *between counts of ARGs and IS elements in metagenome assemblies.*

Predictor	**Estimate**	**Standard Error**	**95% confidence interval**	***p*-value**
Intercept (β_0)	-17.3585	0.3934	-18.1757, -16.6292	<0.0001
IS elements assembled (β_1)	0.01393	0.0067	0.0007, 0.0271	0.0383

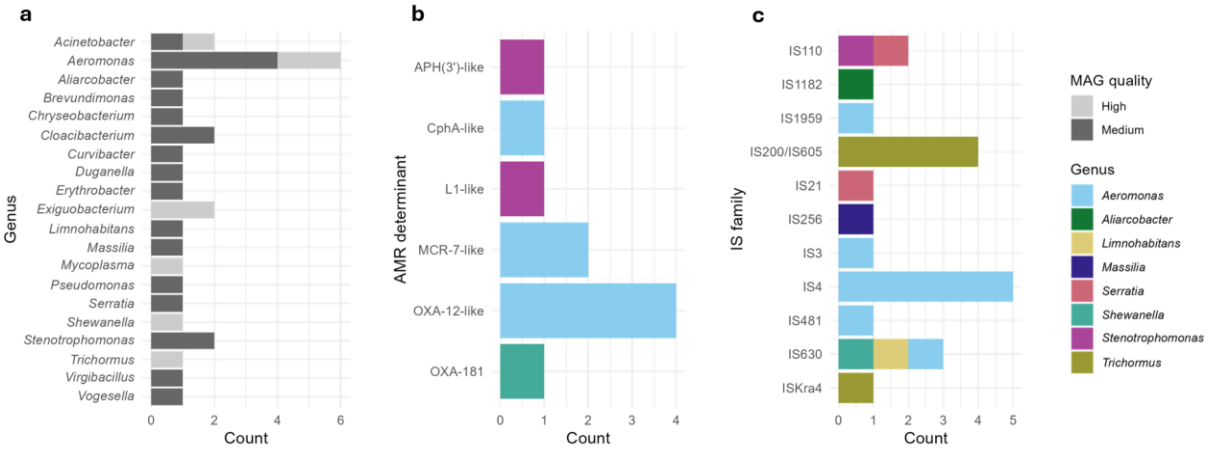
279

280 In tangible terms, the model outputs (*Table 1*) indicate that the addition of 10 IS elements to a
281 snail faecal microbiome is associated, on average, with a 14.9% increase in the number of
282 ARGs (assuming that the assembly length remains constant), while ARGs are expected to be
283 identified at a rate of 0.030 per Mb assembly (or 30 per Gb), in the absence of IS elements
284 (solutions available in *Supplementary Information 1*).

285 **Progenitors of mobile carbapenemase and colistin resistance genes identified**

286 One of the two *bla*$_{\text{OXA-48}}$-like carbapenemase genes identified by read-mapping was also
287 detected in metagenome assemblies. NODE_173 from the M5 metagenome (*Cleopatra* spp.
288 faeces, Nsanje District, Malawi) contained *bla*$_{\text{OXA-181}}$ on a 20996 bp contig with 17 open-
289 reading frames (ORFs). The contig was binned to a high-quality MAG classified as *Shewanella*
290 *xiamenensis* (*Figure 5*). When the *bla*$_{\text{OXA-181}}$-containing contig was queried against the
291 GenBank Nucleotide database, the closest alignment was to chromosomal DNA from an *S.*
292 *xiamenensis* isolate (accession no. AP026732) obtained from urban drain water from Vietnam
293 (100% query cover, 98.51% identity) (32). With only one SNP in a non-coding region, the
294 contig also aligns with 99.92% identity over a 1309 bp region (8073 – 9381) to various
295 plasmids from the Enterobacterales containing the internationally disseminated *bla*$_{\text{OXA-181}}$ gene
296 and a truncated *lysR* (Δ*lysR*) gene, including the KP3-A plasmid from which *bla*$_{\text{OXA-181}}$ was first
297 characterised in a clinical *K. pneumoniae* isolate (accession no. JN205800) (33). The region
298 containing *bla*$_{\text{OXA-181}}$ and Δ*lysR* within the KP3-A (and other) plasmid(s) can be traced back to
299 its origin, *S. xiamenensis*, where *bla*$_{\text{OXA-181}}$ is found adjacent to a complete *lysR* gene, as

300 previously shown (34) and also visible in our annotated sequences (*Supplementary Figure 5*).
301 The ability to detect, assemble, and bin this critical priority ARG is aided by the high relative
302 abundance of *Shewanella* in the M5 sample (5.07%). We also observed a strong *Shewanella*
303 signal in the Uk2 sample (relative abundance = 1.33%), from which another bla_{OXA-48}-like
304 variant ($bla_{OXA-547}$) was detected by read-mapping, when compared to the remaining 13
305 samples in which no bla_{OXA-48}-like genes were detectable (mean *Shewanella* relative
306 abundance = 0.11%).



307

308 *Figure 5. Summary of medium- and high-quality MAGs – a) Genera; b) AMR determinants*
309 *(annotated as '-like' if nucleotide identity is <100% when compared to reference); c) IS families.*

310 Two *mcr-7*-like lipid A phosphoethanolamine transferase genes were detected in a high-quality
311 *Aeromonas jandaei* MAG recovered from the Ug3 metagenome (*Physella* spp. faeces, Kimi
312 Island, Uganda) on separate contigs (NODE_28 and NODE_32: 86.88% and 81.93%
313 nucleotide identity to *mcr-7.1*, respectively). Lipid A phosphoethanolamine transferase
314 enzymes cause a reduction in overall net-negative charge of the outer membrane, conferring
315 resistance to colistin among some gram-negative bacteria, and are a major public health
316 concern due to their propensity to be mobilised from progenitor species chromosomes to
317 plasmids, limiting last-resort antibiotic treatment options for infections caused by multi-drug-
318 resistant pathogens (35). *Aeromonas* spp. are the known progenitors of *mcr-3* and *mcr-7* (36).
319 In our 15 samples, reads mapped to the MCR-3 Cluster90 in one and the MCR-7 Cluster90 in
320 ten, underscoring the widespread prevalence of putative *Aeromonas*-associated colistin
321 resistance genes in the sampled environment.

322 **Discussion**

323 Our findings provide novel insights into the role that freshwater snails play as AMR reservoirs,
324 revealing a potential widespread environmental source of clinically relevant ARGs. Through
325 analysis of snail faecal metagenomes from diverse geographic locations, we demonstrate that

326 their microbiomes harbour a wide array of bacterial species and ARGs, including those
327 conferring resistance to both first- and last-line antibiotics. Identification of environmental
328 progenitors of potentially mobile ARGs conferring resistance to carbapenems and colistin is
329 particularly noteworthy in sub-Saharan Africa, where such antibiotics are not commonly
330 available or prescribed for patients with severe infections caused by Enterobacterales, but
331 where resistance among Enterobacterales to other treatment options such as penicillins, third-
332 generation cephalosporins, and gentamicin has increased (37). A likely increase in
333 carbapenem prescribing as a result of this may introduce additional selective pressures to
334 aquatic habitats, driving selection and transfer of ARGs from environmental reservoirs to
335 clinically relevant bacteria (38). Our results emphasize the potential for snails to contribute to
336 this dissemination in freshwater ecosystems, particularly in those modified by anthropogenic
337 change with deficits in local WASH (water, sanitation, and hygiene) infrastructures (3,18).

338 Proteobacteria dominated the snail faecal microbiomes, reinforcing results from previous
339 studies that found Proteobacteria to be highly abundant in other freshwater snail microbiomes
340 such as the gut and haemolymph (8,39–41). The significant correlations observed between
341 Proteobacteria abundance and both alpha diversity and ARG load suggest that an abundance
342 of Proteobacteria reduces overall microbiome diversity while increasing the relative
343 abundance of ARGs within the microbial community. Specifically, we show that this effect is
344 associated with Aeromonadales, within which we identified geographically widespread
345 reservoirs of three ARG Cluster90s (CphA, MCR-7, OXA-12). The Proteobacteria order and
346 *Aeromonas* genus have both recently been shown to be core taxa in gut samples extracted
347 from laboratory-reared *Biomphalaria* snails (41); as well as signalling that snail faecal
348 microbiomes are a reflection the intestinal tract, our results indicate that these core features
349 are preserved across various snail genera and within their natural habitats. Meanwhile, our
350 visualisation of hierarchical clusters at the order level demonstrates the variation in numerous
351 other bacterial taxa observed between countries (and between districts in the case of Malawi).
352 The differences observed between proximal samples from Malawian *Bulinus* snails at the
353 species level highlight the potential role of local environmental conditions, such as water
354 pollution, climate, and antibiotic exposure, in addition to host-specific characteristics, in
355 shaping faecal microbiomes beyond the core taxa.

356 We recovered eight high-quality MAGs and 21 medium-quality MAGs from 20 bacterial genera
357 in total, yet 4773 unique species from 1246 genera were detected in total by Kraken read
358 profiling, demonstrating the limitations of assembly-based approaches in complex samples
359 such as these. The inability to recover medium- to high-quality MAGs (with >50%
360 completeness) in 1226 of the 1246 genera detected by read profiling (98.39%) suggests that
361 some ARGs from low-abundance species are likely to have been undetectable even by our

362 read-based approach in which 60% template coverage was needed to call a hit, and hence
363 the ARG profiles that we present here are likely biased towards those found in high-abundance
364 genera such as *Aeromonas*. our The detection of ARGs from the CphA, MCR-7 and OXA-12
365 Cluster90s in 66.7%, 66.7% and 93.3% of samples, respectively, all of which originate from
366 *Aeromonas* species (42,43), is a reflection of this. The *mcr-7.1* colistin resistance gene was
367 first identified on plasmids in *K. pneumoniae* isolates from chickens in China (44), but
368 chromosome-borne variants of the genes have more recently been identified in *Aeromonas*
369 isolated from poultry in the United States (45), and in global environmental sources (water)
370 (46). The suggestion that *Aeromonas* species are the natural reservoirs of *mcr-3* and *mcr-*7
371 was since proven by robust evolutionary models (36), and we were able to assemble and bin
372 *mcr-*7-like variants to a putative *Aeromonas jandaei* MAG. Our results signpost to the potential
373 for emergence of new variants of these ARGs from freshwater gastropod faeces, and
374 specifically, the *Aeromonas* reservoirs within.

375 Similar reservoirs of ARGs in low-abundance species may be more widespread than our data
376 indicate. We demonstrate the presence of $bla_{\text{OXA-181}}$ and $bla_{\text{OXA-547}}$ in samples from Malawi and
377 the United Kingdom, respectively, both of which are $bla_{\text{OXA-48}}$-like genes. In the case of $bla_{\text{OXA-}}$
378 $_{181}$, we could bin the >20 kbp contig containing the gene to a high-quality MAG of *S.*
379 *xiamenensis*, the species known to be the gene's progenitor (34). *Shewanella* species have
380 been identified as the progenitors of numerous bla_{OXA-48} variants (47), and the enzymes coded
381 for by these genes hydrolyze carbapenems and confer resistance to most beta-lactam, beta-
382 lactamase inhibitor combinations. They emerged and proliferated over the last two decades
383 to become the most prevalent carbapenemases amongst Enterobacterales in many parts of
384 Europe, Northern Africa, and the Middle East (28). Consequently, they are a major public
385 health concern, with carbapenemase-producing Enterobacterales (CPE) being listed by the
386 World Health Organisation as critical priority pathogens for which new antimicrobial agents
387 are urgently needed (48).

388 To our knowledge, this is the first report of the $bla_{\text{OXA-181}}$ variant in Malawi, a country where
389 CPE are uncommon in the clinical setting, but extended-spectrum beta-lactamase (ESBL)-
390 producing Enterobacterales resistant to a range of other beta-lactam antibiotics are, and
391 contribute to increased mortality in hospital patients with bloodstream infections (49,50).
392 ESBL-producing Enterobacterales have also been widely detected in water systems as well
393 as human and animal faeces in Malawi (50,51). Efforts to characterize the selective pressures
394 behind the spread of ESBL-producing Enterobacterales in Malawian wastewater and
395 freshwater are ongoing, but low environmental antimicrobial concentrations are known to drive
396 the transfer and persistence of ARGs in water systems (38). This selective pressure may also
397 facilitate the emergence of $bla_{\text{OXA-48}}$-like genes from environmental reservoirs where

398 environmental and human health intersect. Meanwhile, such reservoirs of carbapenem
399 resistance genes could allow bacteria to react rapidly if carbapenems are introduced more
400 routinely into clinical practice, limiting their efficacy. Events like this are often driven by
401 horizontal gene transfer (HGT) associated with MGEs such as ISs (3), which we show are
402 abundant within snail faeces. We were able to demonstrate in general via a statistical model
403 that IS elements are associated with an increase in the number of different ARGs assembled
404 in freshwater snail faecal metagenomes. An example of this is shown by the identification of a
405 *sul1*-containing gene cluster upstream of IS*6100* in samples from Malawi and Uganda.

406 Our study does have several limitations. For example, the use of pooled snail faeces for
407 metagenomic analysis, while necessary to obtain sufficient biomass for DNA extraction and
408 sequencing, limits the ability to assess individual snail variability and the potential
409 heterogeneity in microbiome composition and ARG content in snails from the same location.
410 Meanwhile, the reliance on short-read sequencing, although effective for taxonomic profiling
411 and ARG identification, poses challenges in assembling complete genes and genomes,
412 particularly for low-abundance species, potentially leading to an underestimation of the full
413 diversity of ARGs. This limitation is compounded by the inherent bias toward detecting ARGs
414 in more abundant taxa which may skew the representation of ARGs in less abundant microbial
415 communities. Inherently, short-read metagenome assemblies from diverse samples are highly
416 fragmented. This limited our ability to uncover various examples of MGEs linked to ARGs on
417 composite transposons, despite our ability to show a statistical association with regression
418 modelling. Furthermore, while this study highlights the potential role of freshwater snails as
419 AMR reservoirs, the cross-sectional nature of the sampling limits our ability to infer causality
420 or the dynamics of ARG transmission over time. Longitudinal studies are needed to better
421 understand temporal changes in ARG prevalence and the environmental factors influencing
422 these patterns. Finally, while the study identifies ARGs of clinical relevance, it does not explore
423 the functional impact of these ARGs, nor does it confirm their potential for HGT to human
424 pathogens, a critical future step to quantify the risks associated with our findings.

425 Our study represents the first metagenomic analysis of AMR in freshwater gastropod faeces,
426 providing important insights into the potential role of these invertebrates as reservoirs of
427 clinically relevant ARGs. These findings have significant implications for AMR surveillance
428 within One Health frameworks, particularly in areas where agricultural animals, wild animals
429 and humans utilise the same open water sources. This reservoir of ARGs poses a threat to
430 public health in an era of industrial-scale antibiotic use, particularly where access to clean
431 water and sanitation is limited. Future research should build on these results through
432 longitudinal sampling and by investigating how environmental factors, such as antibiotic
433 contamination, influence ARG diversity and abundance in snail microbiomes. Additional

434 sampling in nearby human and animal populations to map potential transmission of ARGs
435 from snail-associated bacteria to human pathogens will be essential for designing targeted
436 interventions to curb AMR spread from and into shared aquatic environments.

437 **Methods**

438 **Sampling collection, processing, and sequencing**

439 In 2023 and 2024, freshwater snails were collected from 15 freshwater locations in four
440 administrative regions: Malawi, n = 6; Uganda, n = 3, Zanzibar (Republic of Tanzania), n = 4;
441 United Kingdom, n = 2. From each location, 50 morphologically identical snails of the same
442 genus were identified by a competent medical malacologist (JRS), rinsed and then pooled in
443 sterile water. The faecal material from each snail group was collected after 24 hours. DNA was
444 extracted from PBS-washed faecal samples with the QIAamp PowerFecal Pro DNA Kit
445 (QIAGEN, Germany), as per manufacturer's instructions. Concentration, purity and integrity
446 of extracted DNA was assessed using a Qubit 4 fluorometer (Thermo Fisher Scientific, USA),
447 NanoDrop One spectrophotometry (Thermo Fisher Scientific, USA), and the TapeStation
448 System (Agilent Technologies, USA). Purified DNA extracts were shipped to Genewiz/AZENTA
449 (UK) for library preparation, validation, and shotgun metagenomic sequencing. Sequencing
450 was carried out with the Illumina NovaSeq instrument using a 2 x 150 bp paired-end
451 configuration.

452 **Bioinformatics**

453 Read processing

454 Raw FASTQ sequence files were quality assessed using FASTQC (https://github.com/s-
455 andrews/FastQC) Version 0.11.9. Trimming with Trimmomatic (52) Version 0.39 was
456 performed to remove adaptor sequences and low-quality bases with a sliding window quality
457 cutoff of Q20 and a minimum read length of 50 bp. Usual practice in metagenomic studies is
458 to filter out host DNA before progressing to analysis of microbial reads by aligning all reads to
459 host reference genomes. As our study involves analysis of pooled DNA samples from hosts
460 whose genomes are not available in databases, we adapted the protocol outlined by Lu et al.
461 (25) to remove contaminant and host DNA. Trimmed reads were classified using Kraken 2
462 (53) Version 2.1.3. Kraken 2 works by matching read k-mers against a database of k-mers
463 from existing genomes. We aligned reads to the standard Kraken 2 database, which includes
464 RefSeq archaea, bacteria, viruses, plasmid complete genomes, UniVec Core and the human
465 reference genome, GRCh38. In a two-step approach using the KrakenTools
466 "extract_kraken_reads.py" script, contaminant reads identified as eukaryotic were first filtered
467 out, before reads identified as bacterial in origin were extracted. The larger read set obtained

468 after the first filtering step was used for read-based ARG profiling, as including only the reads
469 identified as bacterial by Kraken 2 may potentially have removed some ARG variants. The
470 twice filtered bacterial-only reads were used for taxonomic profiling and assembly.

471 Taxonomic profiling

472 Kraken 2 reports containing bacterial read taxonomic assignments were used to run species
473 abundance estimation with Bracken (54), with a minimum number of reads required for a
474 classification at the specified rank set at 10 to reduce noise from potential low-abundance
475 contaminant species, as per default recommendations. A Bray-Curtis dissimilarity matrix
476 containing pairwise dissimilarities among microbiome samples was generated by running the
477 KrakenTools "beta_diversity.py" script with Bracken outputs, to assess bacterial beta diversity.

478 ARG database and clustering

479 Sequences in the ResFinder database (27) Version 2.3.2 were downloaded, then clustered
480 using CD-HIT (55) Version 4.8.1 to identify ARG clusters containing sequences with a
481 minimum nucleotide identity of 90%, referred to as Cluster90s.

482 Read-based ARG detection

483 To identify ARGs within each sample, reads were aligned to reference sequences in the
484 ResFinder database using KMA (26) Version 1.4.14, a tool designed to map raw reads directly
485 against redundant databases such as ARG databases. Alignments were then filtered to retain
486 only those with a minimum 90% query identity and 60% template coverage. For all reference
487 sequences with alignment hits after filtering, the corresponding Cluster90 was identified and
488 annotated manually.

489 Metagenome assembly, binning, and classification

490 Reads identified as being bacterial by Kraken 2 were assembled with metaSPAdes (56)
491 Version 3.11.1. QUAST (57) Version 5.0.2 was used to assess the quality of metagenome
492 assemblies. Contigs were indexed using the Burrows Wheeler Aligner (BWA) Version 0.7.17
493 before bacterial reads were aligned to the indexed contigs with "bwa-mem" (58). The contigs
494 and resulting sorted BAM files were parsed to the "jgi_summarize_bam_contig_depth" script
495 from MetaBAT2 (59) Version 2.17, before the resulting depth files were used by MetaBAT2 to
496 bin assembled contigs of ≥ 2,500 bp to putative genomes (MAGs). CheckM (60) Version 1.1.2
497 was used to assess the completeness and contamination of each MAG. CheckM outputs were
498 compared to standards set by the genome standards consortium to classify MAGs as high-
499 quality (>90% completeness, <5% contamination), medium-quality MAGs (>50%
500 completeness, <10% contamination) or low-quality (<50% completeness and/or >10%

contamination) (61). MAGs deemed as medium- or high-quality proceeded to further analysis. MAGs underwent taxonomic classification using GTDB-Tk (62) Version 2.1.1 against the Genome Taxonomy Database Version 09-RS220 (24th April 2024).

Assembly-based profiling of ARGs and MGEs

Metagenome assemblies were screened against the ResFinder (27) and ISfinder (30) databases using ABRicate (https://github.com/tseemann/abricate) Version 1.0.1 at default settings (minimum 80% identity and 80% query cover) to identify contigs containing ARGs and ISs. Compared to our read-mapping approach, in which we used a higher identity threshold to identify hits (90%) in combination with a lower minimum query cover (60%) due to the more fragmented nature of using raw sequences, the assembly-based approach allowed for a broader detection of putative AMR determinants and MGEs, which could then be characterised further within their genomic context following annotation. Annotations were performed using Bakta (63) Version 1.9.3 (full database version 5.1). To elucidate the potential origins and dissemination pathways of ARGs, selected contigs were subjected to a sequence similarity search using the Nucleotide Basic Local Alignment Search Tool (BLAST) webtool (megablast algorithm) against the GenBank Nucleotide database. This step aimed to identify existing homologous sequences and/or gene clusters, thereby pinpointing potential source organisms. The Clinker webtool (https://cagecat.bioinformatics.nl/tools/clinker) (64) was used to visualise gene cluster comparisons.

Statistical analysis and visualisation

Species abundance and read-based ARG assay data were combined with hierarchical/classification data and sample metadata to generate a TreeSummarizedExperiment data container (65). Data associated with MAGs were handled separately. Within the R environment, downstream data wrangling, analysis, and visualisation took place predominantly using the R/Bioconductor *mia* and *miaViz* (Version 1.12.0) packages, and the R/CRAN *ggplot2* (Version 3.5.1) and *pheatmap* (Version 1.0.12) packages. CLR transformation was applied to microbial composition data (66) to normalize for compositional effects, with a pseudocount added to manage zero values using the *mia* package. Hierarchical clustering of both taxa and samples was applied to standardised CLR-transformed data (CLR-z) using Euclidean distance and complete linkage, as implemented by default in the *pheatmap* package. Pearson correlation was used to explore the relationships between numerical variables and ARG load, assuming linearity in the data. When exploratory analysis revealed non-linear trends in the data, an appropriate transformation (e.g. log transformation) was applied. Generalized linear regression with a negative binomial

535 distribution was carried out with the R/CRAN *MASS* package (Version 7.3.60.2) to assess the
536 association between IS element counts and ARG counts in metagenome assemblies.

537

538 **SEQUENCE AND DATA AVAILABILITY**

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

542

543 **FUNDING**

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

554

555 **AUTHOR CONTRIBUTIONS**

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

560

561 **ACKNOWLEDGEMENTS**

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

566 **COMPETING INTERESTS**

567 None to declare.

568

569 **REFERENCES**

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

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

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

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

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

580 6. Caravedo MA, Cabada MM. Human Fascioliasis: Current epidemiological status and
581 strategies for diagnosis, treatment, and control. Res Rep Trop Med. 2020 Nov 26;11:149–58.

582 7. Liu H, Macdonald CA, Cook J, Anderson IC, Singh BK. An ecological loop: host
583 microbiomes across multitrophic interactions. Trends Ecol Evol. 2019 Dec;34(12):1118–30.

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

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

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

592 11. Lv SB, He TT, Hu F, Li YF, Yuan M, Xie JZ, et al. The impact of flooding on snail
593 spread: the case of endemic schistosomiasis areas in Jiangxi Province, China. Trop Med
594 Infect Dis. 2023 Apr 30;8(5):259.

595 12. Alharbi MH, Condemine C, Christiansen R, LaCourse EJ, Makaula P, Stanton MC, et
596 al. *Biomphalaria pfeifferi* snails and intestinal schistosomiasis, Lake Malawi, Africa, 2017–
597 2018. Emerg Infect Dis. 2019 Mar;25(3):613–5.

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

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

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

607 16. Murray CJ, Ikuta KS, Sharara F, Swetschinski L, Aguilar GR, Gray A, et al. Global
608 burden of bacterial antimicrobial resistance in 2019: a systematic analysis. Lancet. 2022 Feb
609 12;399(10325):629–55.

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

613 18. Musicha P, Morse T, Cocker D, Mugisha L, Jewell CP, Feasey NA. Time to define
614 One Health approaches to tackling antimicrobial resistance. Nat Commun. 2024 Oct
615 10;15(1):8782.

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

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

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

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

626 23. Lee K, Raguideau S, Sirén K, Asnicar F, Cumbo F, Hildebrand F, et al. Population-
627 level impacts of antibiotic usage on the human gut microbiome. Nat Commun. 2023 Mar
628 2;14(1):1191.

629 24. Carruthers LV, Nordmeyer SC, Anderson TJc, Chevalier FD, Le Clec'h W. Organ-
630 specific microbiomes of *Biomphalaria snails* [Preprint]. 2024. Available from:
631 http://biorxiv.org/lookup/doi/10.1101/2024.06.11.598555

632 25. Lu J, Rincon N, Wood DE, Breitwieser FP, Pockrandt C, Langmead B, et al.
633 Metagenome analysis using the Kraken software suite. Nat Protoc. 2022 Dec;17(12):2815–
634 39.

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

637 27. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, et al. ResFinder
638 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother. 2020 Dec
639 1;75(12):3491–500.

640 28. Boyd SE, Holmes A, Peck R, Livermore DM, Hope W. OXA-48-Like β-Lactamases:
641 global epidemiology, treatment options, and development pipeline. Antimicrob Agents
642 Chemother. 2022 Jul 20;66(8):e00216-22.

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

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

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

651 32. Dao TD, Takemura T, Kasuga I, Hirabayashi A, Nga NT, Anh PHQ, et al. Emergence
652 of plasmid-mediated RND-type efflux pump gene cluster *tmexCD-toprJ* in *Shewanella*
653 *xiamenensis* in a water environment [Preprint]. 2022. Available from:
654 http://biorxiv.org/lookup/doi/10.1101/2022.09.15.508154

655 33. Potron A, Nordmann P, Lafeuille E, Maskari Z, Rashdi F, Poirel L. Characterization of
656 OXA-181, a carbapenem-hydrolyzing class D-lactamase from *Klebsiella pneumoniae*.
657 Antimicrob Agents Chemother. 2011 Sep 15;55:4896–9.

658 34. Potron A, Poirel L, Nordmann P. Origin of OXA-181, an emerging carbapenem-
659 hydrolyzing oxacillinase, as a chromosomal gene in *Shewanella xiamenensis*. Antimicrob
660 Agents Chemother. 2011 Sep;55(9):4405–7.

661 35. Samantha A, Vrielink A. Lipid A phosphoethanolamine transferase: regulation,
662 structure and immune response. J Mol Biol. 2020 Aug 21;432(18):5184–96.

663 36. Guo Y, Zou G, Kerdsin A, Schultsz C, Hu C, Bei W, et al. Characterization of NMCR-
664 3, NMCR-4 and NMCR-5, three novel non-mobile colistin resistance determinants:
665 Implications for MCR-3, MCR-7, and MCR-5 progenitors, respectively. Drug Resist Updat.
666 2024 Jul 1;75:101088.

667 37. Kowalski M, Obama BM, Catho G, Dewez JE, Merglen A, Ruef M, et al. Antimicrobial
668 resistance in Enterobacterales infections among children in sub-Saharan Africa: a systematic
669 review and meta-analysis. eClinicalMedicine. 2024 Apr; 70:102512.

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

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

676 40. Hu Z, Tong Q, Chang J, Yu J, Li S, Niu H, et al. Gut bacterial communities in the
677 freshwater snail *Planorbella trivolvis* and their modification by a non-herbivorous diet. PeerJ.
678 2021 Feb 12;9:e10716.

679 41. Li P, Hong J, Wu M, Yuan Z, Li D, Wu Z, et al. Metagenomic analysis reveals
680 variations in gut microbiomes of the *Schistosoma mansoni*-transmitting snails *Biomphalaria*
681 *straminea* and *Biomphalaria glabrata*. Microorganisms. 2023 Sep 28;11(10):2419.

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

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

687 44. Yang YQ, Li YX, Lei CW, Zhang AY, Wang HN. Novel plasmid-mediated colistin
688 resistance gene mcr-7.1 in *Klebsiella pneumoniae*. J Antimicrob Chemother. 2018 Jul
689 1;73(7):1791–5.

690 45. Wang Y, Hou N, Rasooly R, Gu Y, He X. Prevalence and genetic analysis of
691 chromosomal mcr-3/7 in *Aeromonas* from U.S. animal-derived samples. Front Microbiol.
692 2021 Apr 30;12:667406.

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

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

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

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

706 50. Lester R, Musicha P, Kawaza K, Langton J, Mango J, Mangochi H, et al. Effect of
707 resistance to third-generation cephalosporins on morbidity and mortality from bloodstream
708 infections in Blantyre, Malawi: a prospective cohort study. Lancet Microbe. 2022 Dec
709 1;3(12):e922–30.

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

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

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

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

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

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

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

726 58. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-
727 MEM [PREPRINT]. 2013. Available from: https://arxiv.org/abs/1303.3997

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

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

734 61. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK,
735 et al. Minimum information about a single amplified genome (MISAG) and a metagenome-
736 assembled genome (MIMAG) of bacteria and archaea. Nat Biotechnol. 2017 Aug
737 8;35(8):725–31.

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

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

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

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

748 66. Calle ML. Statistical analysis of metagenomics data. Genomics Inform. 2019 Mar
749 31;17(1):e6.

750