A study of the correlation between phenotypic antimicrobial susceptibility testing methods and the associated genotypes determined by whole genome sequencing for a collection of *Escherichia coli* of bovine origin

Thomas J Maunsell^{1,2,3}, Scott Nguyen^{1,4}, Farid El Garach⁵, Christine Miossec⁵, Emmanuel Cuinet⁵,
Frédérique Woehrlé⁵, Séamus Fanning¹ and Dagmara A. Niedziela^{1,6}

¹UCD-Centre for Food Safety¹, UCD School of Public Health, Physiotherapy and Sports Science,
University College Dublin, Belfield, Dublin D04 N2E5, Ireland.

²Department of Biopharmaceutical and Medical Science, Galway-Mayo Institute of Technology,
Dublin Road, Galway, Ireland.

³Microbiology Laboratory, Pathology Department, Blackrock Clinic, Rock Road, Blackrock,
Dublin, Ireland.

⁴Public Health Laboratory Division, District of Columbia Department of Forensic Sciences,
Washington, District of Columbia, 20024, USA

⁵Vétoquinol SA, BP 189-7024 Lure, France.

⁶School of Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland

Corresponding author:

Dr Dagmara A. Niedziela, School of Veterinary Medicine, University College Dublin, Belfield,
Dublin 4, Ireland

Email: dagmara.niedziela@ucdconnect.ie

23 **Keywords**

24 Antimicrobial resistance; *Escherichia coli*; disk diffusion; minimum inhibitory concentration;
25 bovine isolates; genotype-phenotype correlation; whole genome sequencing.

26

27 **Highlights**

- 28 • Culture based antimicrobial susceptibility testing is used to identify therapeutics in the
29 treatment of clinical veterinary isolates
- 30 • Whole genome sequencing is increasingly adopted for surveillance, epidemiological
31 traceback investigations, and detection of antimicrobial resistance genes
- 32 • Little is known in correlations between antimicrobial resistance genotypes and disk
33 diffusion antimicrobial susceptibility testing
- 34 • This study finds that whole genome sequencing is a useful predictor for antimicrobial
35 susceptibility however, minimum inhibitory concentration measurements may still be
36 needed for intermediately resistant isolates

37

38 **Abstract**

39 Antimicrobial resistance (AMR) has increased at an alarming pace in the recent years. Molecular-
40 based methods such as whole genome sequencing (WGS) offer a potential alternative to the
41 conventional labour-intensive methods traditionally used to characterise AMR phenotypes. The aim
42 of this study was to investigate whether WGS could be used as a predictor of AMR in *Escherichia*
43 *coli* isolates of bovine origin.

44 Genomes of 143 *E. coli* cultured from cattle presenting with diarrhoea or mastitis were sequenced
45 on an Illumina MiSeq platform. AMR genes were identified using the ResFinder and AMRFinder
46 databases. Antimicrobial susceptibility testing by disk diffusion was performed on a panel of 10
47 antibiotics, covering 7 antimicrobial classes. Minimum inhibitory concentration (MIC)

48 measurements were made using the Sensititre plate with 6 antibiotics, covering 5 antimicrobial
49 classes. Correlation between genotype and phenotype was assessed statistically by means of a two-
50 by-two table analysis and Cohen's kappa (κ) test.

51 The overall κ correlation between WGS and disk diffusion was 0.81, indicating a near perfect
52 agreement, and the average positive predicted value was 77.4 %. Correlation for individual
53 antimicrobial compounds varied, with five yielding near perfect agreement ($\kappa = 0.81$ – 1.00 ;
54 amoxicillin, florfenicol, gentamicin, tetracycline and trimethoprim-sulfamethoxazole), one showing
55 substantial agreement ($\kappa = 0.65$; nalidixic acid), and four showing moderate agreement ($\kappa = 0.41$ –
56 0.60). The overall κ correlation between WGS and MIC was 0.55 indicating moderate agreement,
57 and the average positive predicted value was 68.6 %. Three antibiotics yielded near perfect
58 agreement (gentamicin, tetracycline and trimethoprim-sulfamethoxazole) and a further three
59 showed fair agreement ($\kappa = 0.21$ – 0.40).

60 WGS is a useful tool that can be used for the prediction of AMR phenotypes, and correlates well
61 with disk diffusion results. MIC measurements may be necessary for antimicrobial compounds with
62 a high proportion of intermediately resistant isolates recorded, such as cephalothin.

63 Introduction

64 Antimicrobial resistance (AMR) is a modern societal challenge with veterinary public health
65 dimensions. This situation has arisen owing to the widespread use of antimicrobial compounds in
66 various settings to treat human and animal infections (Zankari et al., 2013). Emergence of AMR in
67 bacteria of importance to veterinary public health has consequently limited the chemotherapeutic
68 options available for effective treatment of clinical infections. In the case of some multi-drug
69 resistant (MDR) bacteria this development has given rise to the fact that no treatments are available,
70 for instance for carbapenem-resistant *Enterobacteriaceae* (CRE) which are associated with human
71 nosocomial infections globally (Jasovský et al., 2016). In 2015, over 670,000 antibiotic resistant
72 infections were recorded in Europe with approximately 33,000 associated deaths thereby
73 highlighting the need for rapid and accurate antimicrobial susceptibility testing which offers clinical
74 utility (Cassini et al., 2019). AMR can evolve through several mechanisms including point
75 mutations in target housekeeping genes, or *via* horizontal gene transfer (HGT) of AMR-encoding
76 genes mapped to various mobile genetic elements (MGEs) such as plasmids, transposons and
77 prophages. Upregulation of gene expression for efflux pumps can also contribute to this resistance
78 phenotype in bacteria (Punina et al., 2015). The continual evolution of multidrug resistance
79 highlights the fact that AMR will continue to pose a challenge for veterinary public health in the
80 future (Feldgarden et al., 2019).

81 Antimicrobial susceptibility testing (AST) is an important laboratory method for determining which
82 antimicrobial compounds will exhibit therapeutic success if administered to treat a bacterial
83 infection. To reduce the inappropriate use of these drugs whilst attempting to limit the subsequent
84 development and spread of antibiotic resistant bacteria, AST should be performed to guide clinical
85 treatment decisions. Susceptibility testing is essential for detecting the emergence of new resistance
86 phenotypes, determining the extent of AMR in bacterial populations and assessing changes in
87 resistance patterns over time (Zankari et al., 2013). Traditional AST relies on phenotypic-based

88 methods that are cost effective and sparing in terms of resources. Molecular-based methods are also
 89 available and based on the amplification of a known gene or more recently the application of whole
 90 genome sequencing (WGS). Using WGS, it is possible to detect all AMR-associated genotypes in a
 91 given bacterium, rather than evaluating a selected panel of antimicrobial agents with traditional
 92 laboratory-based methods. In order to understand AMR fully and improve AMR diagnostics, its
 93 link to bacterial genotype must be further understood (Feldgarden et al., 2019). WGS also has the
 94 potential to extrapolate new information, e.g., the presence of novel antimicrobial resistance genes,
 95 from bacterial genomes, which may alter how future antimicrobial resistance testing is performed.

96 WGS is currently a valuable diagnostic tool applied to outbreak investigation, tracking and
 97 surveillance (Ribot et al., 2019). The potential of WGS to aid in the detection of AMR is
 98 substantial, allowing for detection of novel resistance variants (Feldgarden et al., 2019). Increased
 99 accessibility to genomic databases *via* electronic means along with a reduction in the cost of
 100 sequencing can, facilitate the replacement of those traditional methods, in time with WGS in the
 101 clinical veterinary laboratory (Tyson et al., 2015; Moran et al., 2017). WGS-based AMR genotype
 102 detection is a rapidly developing area with the potential to accurately and rapidly predict all known
 103 resistance genotypes for an individual isolate using a single assay (Su et al., 2019). This gives
 104 WGS-based methods a clear advantage over culture-based AST methods and PCR where the
 105 number of resistant phenotypes determined is dependent on the antimicrobial resistance-encoding
 106 genes selected for inclusion (Su et al., 2019). Once generated, WGS data can be used for other
 107 purposes, such as investigation of virulence or phylogeny determination. This approach aims to
 108 predict the antimicrobial resistance phenotype of a bacterium that would otherwise have been
 109 obtained if the isolate was tested phenotypically for the same trait. This is done by cross referencing
 110 the sequence obtained following WGS, using a suitable bioinformatics pipeline, with a database of
 111 known resistance genes such as ResFinder, CARD or AMRFinder (Zankari et al., 2012; McArthur
 112 et al., 2013; Feldgarden et al., 2019). AMR genotypes can be predicted based on the presence of

113 these AMR genes or mutations in highly conserved peptides which are known to play a key role in
114 bacterial cell physiology (Su et al., 2019). Mutations also occur in accessory genes, which code for
115 mechanisms such as efflux pumps, which results in antibiotic resistance (Vogwill and Maclean,
116 2015). Studies that compare phenotypic and genotypic resistance profiles are needed in order to
117 establish the reliability of WGS-based approaches as a stand-alone predictor of true AMR in
118 bacteria (Moran et al., 2017).

119 A number of studies have been performed to assess the correlation between WGS and phenotypic
120 AMR (Card et al., 2013; Stoesser et al., 2013; Zankari et al., 2013; Tyson et al., 2015; Feldgarden et
121 al., 2019; Stubberfield et al., 2019). Reported correlation study data, however, can depend on the
122 nature of the antibiotic; the bacterial species tested and the associated mechanism of resistance. The
123 cited studies use MIC as the phenotypic AST method of choice for comparison. Little is known
124 about the correlation between disk diffusion AST and WGS antimicrobial resistance testing.
125 Therefore, further research evidence is required to determine whether WGS can accurately infer
126 antimicrobial susceptibility (Stubberfield et al., 2019).

127 *Escherichia coli* is found as part of normal gut flora of both humans and warm-blooded animals.
128 Most *E. coli* are non-pathogenic but can act as a reservoir of AMR genes, which can be transmitted
129 to other bacteria sharing the same ecological niche thereby posing a threat to human health
130 (Stubberfield et al., 2019). Transmission to humans occurs through consumption of undercooked or
131 contaminated animal products such as hamburgers or milk. Transmission can also occur through
132 consumption of plant-based foods contaminated with faecal matter, and direct contact between
133 livestock and farmers (Espinosa et al., 2018). This emphasises the need for a multisectoral One
134 Health approach to combating antimicrobial resistance.

135 The aim of this study is to investigate the use of WGS as a modern diagnostic tool to accurately
136 predict AMR phenotypes in *E. coli* of bovine origin. Genotypic and phenotypic AMR results of 143

137 *E. coli* of bovine origin were compared to determine the correlation between the resistance
138 genotype and the phenotypic expression of the determinant.

139 **Methodology**

140 **Bacterial Isolates and culture methods**

141 A total of 143 *E. coli* of bovine origin were isolated from faecal and milk samples taken from
142 animals presenting with mastitis and diarrhoea. These cattle were located in Germany and France
143 and samples were obtained between the years 2004 and 2010 (**Table 1**).

144 Isolates were placed in Lysogeny Broth (LB) (Sigma Aldrich) with 15 % [v/v] glycerol and stored at
145 -80 °C. Working cultures of each isolate were prepared by streaking the LB broth onto Mueller-
146 Hinton (MH) agar (Sigma Aldrich) and from there taking three single colonies and re-streaking onto
147 MH agar. Three single colonies from the second inoculum were used to inoculate 5 ml MH broth
148 (Sigma Aldrich). The MH broth was incubated overnight at 37 °C, with orbital shaking at 200 rpm
149 to promote growth.

150 **Whole Genome Sequencing**

151 **DNA isolation, purification and library preparation**

152 Total genomic DNA (gDNA) from overnight cultures of bacterial isolates grown in MH broth was
153 purified using the DNeasy UltraClean Microbial Kit (Qiagen). Quality and quantification of total
154 gDNA extract was assessed using the NanoDrop Microvolume Spectrophotometer (Thermo Fisher
155 Scientific). Following quantification, total gDNA was serially diluted 1:10 for each isolate to a
156 concentration of 2 ng/μl which was confirmed by fluorometric quantification using the Qubit
157 fluorometer and a dsDNA High Sensitivity assay kit (Thermo Fisher Scientific). Libraries were
158 generated using the Nextera XT Library Preparation Kit (Illumina). All libraries were then
159 subjected to 300 bp pair-end sequencing (V3 chemistry) on a MiSeq sequencer (Illumina) to
160 produce raw read data.

161 **Whole genome assembly and annotation**

162 FastQC (v0.11.7) (Wingett and Andrews, 2018) was used to assess raw read quality while
163 Trimmomatic (v0.39) (Bolger et al., 2014) was used to trim low-quality sequences. SPAdes
164 (v3.13.0) (Bankevich et al., 2012) was then used to *de novo* assembly trimmed paired reads and the
165 quality of the resulting contigs was determined using QUAST (v5.0.2) (Gurevich et al., 2013).

166 **Identification of AMR genes, serogroup, phylogroup and strain type**

167 Identification of AMR encoding genes was carried out using Abricate (v1.0.0;
168 <https://github.com/tseemann/abricate>) and the ResFinder databases (version last updated on 14 May
169 2020; (Zankari et al., 2012), as well as using NCBI AMRFinder Plus (v3.8.4; Feldgarden et al.,
170 2019). The criteria for defining AMR gene presence included a cut-off of 80 % for gene coverage
171 and 95 % for gene identity. A consensus from both ResFinder and AMRFinder databases was used
172 for final genotype to phenotype correlation. Multi-locus sequence typing (MLST) was performed on
173 the isolates using the mlst package (v2.18.1; <https://github.com/tseemann/mlst>). Serotyping was
174 carried out using the EcOH database within Abricate.

175 **Phylogenomic analysis**

176 Enterobase (<https://enterobase.warwick.ac.uk/>) was used to curate strains from the whole genome
177 sequencing project (BioProject Accession PRJEB32666) (Zhou et al., 2020). A SNP project in
178 Enterobase (Zhou et. al, 2020) was used to construct a phylogenetic tree. In brief, *E. coli* genomes
179 were aligned against reference CFS3246 using default parameters in the Enterobase SNP project.
180 313947 variant sites were called with a minimum of 95% sites present in all genomes. The tree was
181 visualised in GrapeTree in Enterobase (Zhou et. al., 2020). Phylogroup prediction was also
182 performed within Enterobase. Additional metadata and sample IDs are available in Table S1.

183 **Data availability**

184 Fastqc files are available in European Nucleotide Archive (ENA), project number PRJEB32666.

185 <https://www.ebi.ac.uk/ena/data/view/PRJEB32666>

186 Genome assemblies and associated metadata are available in Enterobase under Bio Project

187 ID: PRJEB32666. http://enterobase.warwick.ac.uk/species/ecoli/search_strains

188 **Antimicrobial Susceptibility Testing by Disk Diffusion (AST-DD)**

189 Antimicrobial Susceptibility testing was carried out on all 143 isolates in accordance with Clinical

190 and Laboratory Standards Institute (CLSI) guidelines for disk diffusion. A total of 10 antibiotics

191 were used covering 7 antimicrobial classes and included: amoxicillin (10 µg),

192 amoxicillin/clavulanic acid (20/10 µg), gentamicin (10 µg), cephalothin (30µg), ciprofloxacin

193 (5µg), florfenicol (30µg), marbofloxacin (5µg), nalidixic acid (30µg), tetracycline (30µg),

194 trimethoprim-sulfamethoxazole (1.25/3.75µg). Each antimicrobial was tested in triplicate and the

195 average zone size was used to determine resistance/sensitivity according to CLSI document

196 VET01S, 5th ed. *E. coli* ATCC™25922 was used as a control strain.

197 **Antimicrobial Susceptibility Testing by minimum inhibitory concentration (AST-MIC)**

198 Minimal inhibitory concentrations (MIC) of 22 antibiotics against 82 *E. coli* isolates was

199 determined using the ready to use Sensititre Gram Negative MIC Plates GN3F (ThermoFisher

200 Scientific) according to manufacturer's instructions. Briefly, a 0.5 McFarland suspension in

201 physiological water was prepared from an agar overnight culture of each isolate. Then the

202 suspension was diluted in cation adjusted Mueller Hinton broth to reach 10⁵ CFU/mL. 50 µL of

203 diluted inoculum was then used to inoculate each well of the Sensititre plate except negative control

204 wells and the plate was incubated at 34-36°C for 18-24 hours. MIC results were interpreted

205 according to CLSI document VET01S. Antibiotics were selected for correlation if 10 or more

206 phenotypically resistant isolates and AMR genes were present for a given compound, and included

207 ampicillin, cephalothin, ciprofloxacin, gentamicin, tetracycline and trimethoprim-sulfamethoxazole.

208

209 **Statistical analysis**

210 Correlations between genotypic and phenotypic data obtained in this study was assessed statistically
 211 by means of a two-by-two table analysis (Mackinnon, 2000). AMR gene presence/absence
 212 (obtained from WGS data) was compared with phenotypic disk diffusion/MIC AST data giving the
 213 actual resistance profile of the isolate. Sensitivity, specificity, positive predicted value (PPV) and
 214 negative predicted value (NPV) were calculated taking the following into consideration:

- 215 - True positive – gene present and disk diffusion/MIC resistant
- 216 - False positive – gene present and disk diffusion/MIC susceptible
- 217 - True negative – gene absent and disk diffusion/MIC susceptible
- 218 - False negative – gene absent and disk diffusion/MIC resistant

219 Based on recommendations for surveillance of antimicrobial resistance, intermediate isolates were
 220 treated as resistant and subsequently referred to as “intermediately resistant” (Cornaglia et al.,
 221 2004). Cohen’s kappa (κ) test was performed to measure the agreement of AST-WGS to Disk
 222 Diffusion/AST-MIC in the detection of AMR. A κ value of 1 would indicate perfect agreement.

223 **Results and Discussion**

224 The bovine *E. coli* cultured from diarrhoea and mastitis cases and used in this study were sequenced
 225 using short-read sequencing on an Illumina MiSeq platform and antimicrobial resistance genes were
 226 identified by bioinformatic analysis. Phenotypic AST by disk diffusion (AST-DD), including 10
 227 antimicrobial agents, covering 8 classes, was carried out and standardised to CLSI guidelines
 228 (CLSI, 2021). A subset of 82 isolates with more than one AMR gene was subsequently selected for
 229 MIC analysis using Sensititre plates (AST-MIC). A total of 6 antibiotics belonging to 5
 230 antimicrobial classes were used for correlation between genotype and MIC.

231 **Population structure of French and German bovine-associated *E. coli***

232 During 2004-2010, 147 *E. coli* were isolated from animals presenting with diarrhoea or mastitis
233 symptoms and subjected to whole genome sequencing, with 4 genomes excluded from the analysis
234 (CFS3241, CFS3306, CFS3348, and CFS3367) due to failed genome assembly QC. The genomic
235 features of these isolates have been described elsewhere (manuscript in submission, 2021).

236 The isolates belonged to phylogroups A (n = 45), B1 (n = 42), B2 (n = 5), C (n = 18), D (n = 20), E
237 (n = 11) and G (n = 2), as determined by ClermontTyping in Enterobase (Clermont et al., 2000,
238 2013; Beghain et al., 2018) (**Fig. 1A**). Bovine mastitis isolates were broadly distributed throughout
239 all phylogroups, supporting evidence that there are no clear related genotypes for a proposed bovine
240 mastitis pathotype (Leimbach et al., 2017) (**Fig. 1B**).

241 **AMR genotypes identified by database comparisons**

242 AMR genes were identified using AMRFinder and ResFinder databases. All ResFinder database
243 genes were also detected by AMRFinder, with AMRFinder providing additional genes. All isolates
244 encoded multidrug efflux RND transporter permease subunit gene *acrF* and a version of *blaEC*
245 gene encoding beta-lactam resistance, and 134 isolates encoded multidrug efflux MFS transporter
246 *mdtM* (**Figure 2**). When the efflux pump genes were excluded, 82 isolates encoded more than 1
247 AMR gene responsible for resistance to a specific antibiotic. The most common gene identified was
248 *bla_{TEM-1}* which is associated with beta-lactam resistance, being detected in 53 isolates (37.06%)
249 (**Table 2**). This gene confers resistance to cephalothin and amoxicillin. Other genes of high
250 frequency include *tet(A)* (30.77%), and *sul2* (35.66%) which are associated with tetracycline and
251 sulfonamide resistance, respectively. There was a high prevalence of resistance to tetracycline: five
252 different genes conferred resistance to tetracycline with *tet(A)* and *tet(B)* (16.08 %) being the most
253 prevalent. Trimethoprim-sulfamethoxazole resistance requires the presence of genes from both *sul*
254 and *dfr* gene groups. There was relatively low resistance detected for the aminoglycoside antibiotic

gentamicin with 15 of the 143 genomes harbouring gentamicin resistance-encoding genes. The most frequent AMR gene promoting gentamicin resistance was *ant(2'')-Ia*, which was detected in 8 of the 143 isolates. Florfenicol resistance marked by the presence of the *floR* gene was detected in 14 isolates (9.79%).

There was a high level of antimicrobial resistance present in the cohort of isolates tested, 51.04% (n=73) of which had an MDR genotype harbouring resistance to between 3 and 10 of the antibiotics investigated for the correlation in this study. The antimicrobial therapy options against Gram-negative MDR bacteria are limited. This poses a risk to veterinary public health and a challenge in the fight against nosocomial infections (Giamarellou and Poulakou, 2009). The AMR genes most frequently detected in the isolates (*bla_{TEM-1}*, *tet(A)* and *sul2*) agreed with previous studies investigating AMR genes in MDR *E. coli* of bovine origin, which detailed a high level of tetracycline, sulfonamide and beta-lactam resistance among isolates sampled from cattle in different European countries (Karczmarczyk et al., 2011).

AST by disk diffusion (AST-DD)

A total of 10 antibiotics were investigated by conventional disk diffusion testing for correlation with genotype determination. These compounds belonged to 7 classes of antimicrobial agent which included aminoglycoside, beta-lactam, tetracycline, phenicol, fluoroquinolone, quinolone and sulfonamide-trimethoprim compounds. Of the 143 isolates tested, 46 (32.17%) *E. coli* were fully susceptible to all 10 antibiotics tested for by disk diffusion. All other isolates were resistant or intermediately resistant to between 1 and 10 antibiotics. AST-DD determined that 45.5 % (65/143) of isolates were multi-drug resistant (MDR). The most frequent resistance recorded was seen for the antibiotics amoxicillin and tetracycline with 66 and 65 resistant isolates (46% and 45%), respectively (**Figure 3A**). The antibiotics demonstrating least resistance among isolates recorded were cephalothin (7.69%) and amoxicillin/clavulanic acid (8.39%).

Resistance patterns of the isolates were compared based on disease type (diarrhoea, **Figure 3B** versus mastitis, **Figure 3C**). There was a higher prevalence of antimicrobial resistance in the cohort of diarrhoeal samples in comparison to mastitis samples. The highest percentage resistance recorded for diarrhoeal samples was 76.7 % for amoxicillin, while for mastitis isolates it was 15.7 % for tetracycline. No resistance to amoxicillin/clavulanic acid was recorded for mastitis samples, and low resistance to ciprofloxacin, florfenicol, gentamicin, marbofloxacin and nalidixic acid (1.4 %) was detected. Overall, there was a higher prevalence of resistant isolates among samples recovered from cases of diarrhoea.

AST by minimum inhibitory concentration (AST-MIC)

A subset of 82 isolates with more than one AMR-encoding gene was selected for MIC determination. Of these isolates, 23.17 % were sourced from mastitis cases, while 76.83 % were sourced from cases of diarrhoea. A total of 6 antibiotics were investigated for correlation between both methods. These antibiotics belonged to 5 classes of antimicrobial agent. Of the 82 isolates tested, 11 (13.41 %) *E. coli* were fully susceptible to all 6 compounds tested. All other isolates were resistant or intermediately resistant to between 1 and 6 antibiotics. The most frequent resistance was recorded for the antimicrobial compounds tetracycline and trimethoprim-sulfamethoxazole with 60 and 33 resistant isolates (73.17 % and 40.24 %), respectively (**Figure 4**). The antibiotics showing least resistant isolates recorded were cephalothin (12.20 %) and gentamicin (14.63 %). As with AST-DD, the antimicrobial resistance was higher among those isolates recovered from cases of diarrhoea (n = 63), rather than mastitis (n = 19).

As with AST-DD, when comparing resistance profiles based on disease type for AST-MIC, there was a higher prevalence of antimicrobial resistance in the cohort of diarrhoeal samples in comparison to mastitis samples. The highest percentage resistance recorded for diarrhoeal samples was 79.37 % for tetracycline while for mastitis isolates it was 52.63 % for tetracycline. No resistance to gentamicin was recorded for mastitis samples while low resistance to ciprofloxacin

(5.26 %), marbofloxacin (5.26 %) and cephalothin (5.26 %) was noted. Overall, there was a higher prevalence of resistant isolates among diarrhoeal samples. Overall, both when determined by AST-DD and AST-MIC, the resistance phenotypes of study isolates tended to differ according to the disease they were associated with. Interestingly, the isolates did not cluster on the phylogenetic tree according to disease type, with mastitis and diarrhoeal samples spread evenly across phylogroups and wgMLST types. This suggests that the difference in resistance phenotype could be due to a selective pressure due to the antibiotics used to treat these infections. Mastitis is predominantly treated with amoxicillin and other beta-lactam antibiotics which correlates well with the resistance detected in mastitis isolates. Broader spectrum antibiotics such as tetracycline are usually prescribed to treat bovine diarrhoea (Smith, 2015). However, digestive pathogens also tend to be more exposed to antibiotics than mastitis pathogens, which could also explain the higher rate of resistance.

Correlation of genotypic and phenotypic methods (Disk Diffusion)

Correlation between genotypic and phenotypic data in this study was assessed by means of a two-by-two table analysis. Correlation of AST-WGS with AST-DD for all antibiotics overall was 92.5% for test specificity and 98.8% for test sensitivity (**Table 3**). The kappa score ($\kappa = 0.81$) indicates almost perfect agreement. Correlation for individual antibiotics varied with five compounds yielding almost perfect agreement ($\kappa > 0.90$), one showing substantial agreement and four showing moderate agreement.

Based on these calculations, an almost perfect agreement was observed between AST-WGS and AST-DD for the antibiotics florfenicol, tetracycline, gentamicin, trimethoprim-sulfamethoxazole and amoxicillin. These had a $\kappa > 0.95$ and all had a correlation specificity $> 98.7\%$ and sensitivity $> 97.0\%$, which suggests AST-WGS genotype calling could be used as an accurate predictor of AMR phenotype for these compounds. Both amoxicillin and trimethoprim-sulfamethoxazole had a test specificity of 100% with no false negative results. A large number of genotypes conferred resistance to amoxicillin (**Figure 5A**), all of which can be considered good predictors of resistance

due to the absence of false positive results. There was also an almost perfect agreement for the antibiotic florfenicol (**Figure 6A**). Only one genotype (*floR*) conferred resistance to florfenicol and there was a clear distinction between resistant and susceptible isolates. The gene that seems to most reliably predict resistance to gentamicin (**Figure 6B**) is *aac(3)-Ile*. The least reliable gene seems to be *ant(2'')-Ia* as it was found to be present in a susceptible isolate. The prediction of tetracycline resistance by AST-WGS also had an almost perfect agreement between the methods. There were a large number of genotypes that confer resistance and susceptible isolates that had no associated genotype (**Figure 6C**). In the case of trimethoprim-sulfamethoxazole both *dfr* and *sul* genes were required to be present to confer resistance (**Figure 6D**). There were no false positive results recorded for this antibiotic combination and differences between resistant and susceptible were clear due to the absence of intermediately resistant isolates.

Substantial agreement ($\kappa > 0.6$) was observed for nalidixic acid with a κ value of 0.65. Similarly to the antimicrobial agents discussed previously, nalidixic acid had no false negative results, giving the compound 100% NPV in the correlation. The correlation for this antibiotic could be improved by investigating the genes causing false positive results for 18 of the isolates (**Figure 7B**).

Moderate agreement was observed for amoxicillin/clavulanic acid, cephalothin, ciprofloxacin and marbofloxacin. In the case of amoxicillin/clavulanic acid, which had a test specificity of 99.2%, test sensitivity (45.5 %) negatively impacted correlation resulting in a Cohen's kappa score of 0.56. For amoxicillin/clavulanic acid, *bla_{OXA-2}* was not an accurate predictor of resistance resulting in a false positive (**Figure 5B**). There were also a large number of false negative results for this antibiotic which suggests there are genotypes conferring amoxicillin/clavulanic acid resistance that are undiscovered or not included in the AMRFinder database. For the correlation with cephalothin, high levels of false positive and false negative results were recorded, as 11 susceptible isolates with AMR genotypes were present and 32 isolates were intermediately resistant with others being recorded as resistant with no AMR genotype respectively (**Figure 5C**). It is therefore apparent that

354 further consideration should be given to those genes that confer resistance to beta-lactam antibiotics
355 due to the presence of resistant isolates with no genotype. The fluoroquinolones ciprofloxacin and
356 marbofloxacin also were moderately correlated, both with a κ value of 0.47. Both antimicrobial
357 agents also had the same test performances in terms of test specificity (79.5%) and test sensitivity
358 (100.0%). There were no false negative results for these antibiotics but the poorer correlation is
359 likely due to the high number of false positive results i.e., isolates that are phenotypically
360 susceptible but have an associated genotype (**Figure 7**, isolates in the blue/green shaded area with
361 an associated genotype). Genes that were poor predictors of AMR in this class of antibiotic, causing
362 false positive results in the correlation, included *gyrA_S83L* and *parC_S57T* (**Figure 7A & 7C**).
363 Mutations in the gene encoding the A subunit of DNA gyrase (*gyrA*) and topoisomerase IV subunit
364 ParC gene (*parC*) are the major cause of quinolone resistance encountered in clinical isolates.
365 Studies have revealed that these mutations predominate in conserved regions known as the
366 quinolone-resistance-determining region or QRDR, involved in DNA binding. In GyrA, the QRDR
367 comprise amino acids between 67 and 106 (Varughese et al., 2018). The most common mutations
368 discovered in gram-negative bacteria such as *E.coli*, *Shigella*, *Citrobacter* and *Pseudomonas* were
369 at codon 83 and 87 in GyrA (Ruiz, 2003), which were the only mutations found in this study. In *E.*
370 *coli*, *gyrA* mutations usually occur at serine-83, which is substituted by leucine or tryptophan and
371 causes high resistance to quinolones. However, the replacement of serine by alanine causes a lower
372 resistance level (Ruiz, 2003; Krishnan et al., 2012). In our study, the *gyrA_S83A* gene resulted in a
373 strain which was phenotypically susceptible to ciprofloxacin and marbofloxacin and intermediately
374 resistant to nalidixic acid (**Figure 7**), therefore confirming the findings of previous studies.
375 Mutations in *parC* usually occur at 78, 80 and 84 positions (Hiasa, 2002; Ito et al., 2008), which
376 calls into question the point mutations we observed in positions 56 and 57, with all strains
377 harbouring *parC_A56T* and *parC_S57T* susceptible to all 3 quinolones in the study. Finally, studies
378 show that multiple mutations are required for high level of resistance (Fàbrega et al., 2008;
379 Shigemura et al., 2012), which is also consistent with our findings.

Most of the discrepancies in correlation were due to false negative results, i.e., isolates that were phenotypically resistant but genotypically susceptible. This is most likely due to the presence of new AMR gene variants that have not yet been described. False negatives may also be caused by other resistance mechanisms such as the action of efflux pumps (Stubberfield et al., 2019). Two efflux pump genes, *acrF* and *mdtM*, which were identified in all study isolates, may partially explain the discrepancy between the 81 genotypically resistant isolates and the 135 phenotypically resistant isolates. In *Escherichia coli*, the AcrAB-TolC efflux pump plays an intrinsic role in resistance to hydrophobic solvents. The *acrEF* operon encodes the components of a second efflux pump with AcrE and AcrF highly homologous to AcrA and AcrB, and can confer the same solvent resistance (Kobayashi et al., 2001). However, both *acrE* and *acrF* would be necessary for the resistance mechanism, making the role of *acrF* only in the resistance of the isolates in our study unclear. The multidrug resistance protein MdtM, a recently characterised homologue of MdfA and a member of the major facilitator superfamily, functions in alkaline pH homeostasis (Holdsworth and Law, 2013). Both *acrEF* and *mdtM* (also known as *yjiO*) have been found to confer resistance to some antibiotics as well as other compounds such as ethidium bromide (Nishino and Yamaguchi, 2001).

False positives i.e., isolates that are genotypically resistant and phenotypically susceptible to certain antimicrobial compounds, account for the remaining discrepancies. False positives may be caused by detection of silenced genes that do not express resistance (Enne et al., 2006). Genotypically resistant, phenotypically susceptible isolates may also arise if the wrong AMR genes are selected to correlate genotype to phenotype. This highlights the need for careful curation of AMR gene alleles conferring resistance to each antimicrobial compound in order to ensure reliable and accurate prediction of AMR.

Poor correlation between genotype and phenotype was often accompanied by a high proportion of intermediately resistant isolates (**Table 4**). Higher percentages of intermediately resistant strains

were found for antibiotics with $\kappa < 0.6$. Notably, 44.1% isolates were intermediately resistant to cephalothin which had a κ correlation of 0.40. Cephalothin had the highest proportion of intermediately resistant isolates in this study, and was the likely cause of the poor correlation. The high number of isolates which were intermediately resistant to cephalothin may be due to issues with the method. Previous studies comparing reference broth dilution MIC and disk diffusion methods for cephalothin AST determination highlighted discrepancies in results (Nguyen and Graber, 2013). This discrepancy was also highlighted in our study. Of the 63 isolates which were intermediately resistant to cephalothin by disk diffusion, 44 were captured in the sub-selection of 82 isolates tested by AST-MIC. Only 21 of this sub-selection were then intermediately resistant when tested by AST-MIC.

Correlation of genotypic and phenotypic methods (MIC)

Correlation of AST-WGS with AST-MIC for all antibiotics overall was 75 % for test specificity and 95.9 % for test sensitivity (**Table 5**). Two antibiotics had a $\kappa > 0.90$ which indicated a near perfect agreement between the methods, and notably, the κ correlation for trimethoprim-sulfamethoxazole was 1, with no false positives or negatives for this antibiotic. Fair agreement was observed for some compounds such as cephalothin, ciprofloxacin and marbofloxacin.

Almost perfect agreement was observed between AST-WGS and AST-MIC for the antibiotics gentamicin, tetracycline and trimethoprim-sulfamethoxazole. These had $\kappa > 0.80$ values and a correlation specificity > 84 % and sensitivity > 90 %, which suggests AST-WGS could be used as an accurate predictor of AMR for these antibiotics. For gentamicin (**Figure 8B**), the gene that seems to most reliably predict resistance is *aac(3)-IIe*. The least reliable gene seems to be *ant(2'')-Ia* as it is present in a susceptible isolate. This finding agrees with the results for gentamicin AST-DD. Tetracycline resistance was predicted well by AST-WGS, with almost perfect agreement between the methods. There were a large number of genotypes that confer resistance and susceptible isolates that had no associated genotype (**Figure 8C**). Once again, both *dfr* and *sul*

430 genes were both required to confer resistance to trimethoprim-sulfamethoxazole (**Figure 8D**).
 431 Differences between resistant and susceptible were clear due to the absence of isolates
 432 intermediately resistant to trimethoprim-sulfamethoxazole.

433 For cephalothin, there was an increase in false positive results and a decrease in false negative
 434 results in comparison to AST-DD (**Figure 8A**). A total of 27 susceptible isolates with AMR
 435 genotypes were present (*versus* 11 as determined by AST-DD) and 5 isolates were intermediately
 436 resistant without an AMR genotype respectively (*versus* 32 as determined by AST-DD). A better
 437 appreciation of the linkages between genotypes and phenotypes that confer resistance to beta-lactam
 438 antibiotics is required.

439 Fluoroquinolones ciprofloxacin and marbofloxacin returned a fair agreement between the methods
 440 also, with a κ value of 0.38 and 0.35 respectively. Both antimicrobial compounds also had the same
 441 test performances in terms of test sensitivity (100.0 %) but differed in terms of test specificity. No
 442 false negative results were recorded for these compounds; therefore, the poorer correlation is likely
 443 owed to the high number of false positive results (**Figures 8E & 8F**, isolates in the blue/green
 444 shaded area with an associated genotype).

445 **General discussion – comparison between disk diffusion and MIC correlations**

446 In this study, WGS was explored as an alternative means of predicting AMR in place of
 447 conventional disk diffusion and MIC determined by microbroth dilution, which are routinely used
 448 phenotypic methods of AST. WGS was carried out on 143 *Escherichia coli* of bovine origin and
 449 AMR genes detected using bioinformatic analysis. The same isolates were then subjected to AST-
 450 DD for a total of 10 antibiotics of clinical and veterinary importance. A subset of these same
 451 isolates were also subjected to AST-MIC analysis using Sensititre plates. Genotypic and phenotypic
 452 AMR results were correlated and statistical agreement between the methods was determined.

453 Previous studies reported on the correlation between AST-MIC and genotypic AMR patterns (Card
454 et al., 2013; Stoesser et al., 2013; Zankari et al., 2013; Tyson et al., 2015; Feldgarden et al., 2019;
455 Stubberfield et al., 2019) but no study to date examined the correlation between AST-DD and
456 genotypic AMR patterns. While broth microdilution is the clinical standard for AST, disk diffusion
457 is a cheaper, and a more accessible alternative. AST-DD correlated better with AST-WGS than
458 AST-MIC. The overall κ value for correlation of AST-DD with AST-WGS was 0.81, in comparison
459 to 0.55 for AST-MIC by Sensititre with AST-WGS. This study highlights AST-DD is equivalently
460 as useful a method of AST as AST-MIC.

461 The accuracy of genomic methods such as AST-WGS was determined by the reliability of the
462 software that is used to determine AMR. The ResFinder and AMRFinder databases which were
463 used to predict AMR genes, as previously discussed, may not always correctly identify AMR genes.
464 Previous studies implementing ResFinder have highlighted this feature (Feldgarden et al., 2019).
465 An isolate which is wrongly classified as genotypically susceptible due to lack of known genetic
466 determinant of resistance may negatively affect treatment decisions. Although AST-WGS can offer
467 more information on the test isolates, a disadvantage is that genomic approaches cannot always
468 predict a phenotype, as level of gene expression and protein production from identified genes may
469 differ between strains. Bacteria have gene silencing mechanisms, and mutations may create stop
470 codons in the data (Dryselius et al., 2003).

471 Similar studies that correlated WGS with phenotypic AST found that WGS could be used as an
472 accurate predictor of AMR in human, bovine or porcine isolates (Zankari et al., 2012; Card et al.,
473 2013; Stoesser et al., 2013; Tyson et al., 2015; Feldgarden et al., 2019; Stubberfield et al., 2019).
474 Our data further confirms the same correlation and usefulness of WGS in a collection of bovine
475 isolates. However, extended analysis with a larger number of test isolates would be desirable.
476 Furthermore, all the above studies differed with respect to WGS databases used, AST clinical
477 guidelines (CLSI/EUCAST), sample size and types and numbers of antibiotics examined. For

instance, while our study examined 10 antibiotics from 8 antibiotic classes, another recent veterinary study on *E. coli* isolates from pig samples by Stubberfield et al. (2019) tested antibiotics from 6 classes, with a larger sample size of 515. Even though the study by Stubberfield et al. (2019) may be more accurate due to the larger sample size, our study provides a novel insight into AMR correlation in bovine isolates, which often differ from isolates found in pigs with regards to strain type. A study by Feldgarden et al. (2019) examined 6,242 isolates including 294 *Campylobacter coli*, 476 *Campylobacter jejuni*, 47 *Escherichia coli*, and 5,425 *Salmonella enterica* isolates from human, animal and food sources, which is the largest sample size among the studies reported. However, the number of *E. coli* isolates was relatively low, and the animal sources were not specified. The only other study on bovine isolates was performed by Tyson et al. (2015), with only 76 *E. coli* isolates tested. Our study therefore provides an important insight on the agreement between genotypic and phenotypic AMR in bovine *E. coli*.

Several future approaches to AMR correlation analysis could be suggested. To improve correlation, isolates with three or more false positive/false negative results across three or more antimicrobial classes could be removed from the isolate group to exclude isolates where potential testing errors, sample discrepancies or other confounding factors may be at play. This approach was utilised previously to improve correlation in a similar study by Feldgarden et al. (2019). Isolates which are resistant to certain antimicrobial compounds but lack related AMR genes, i.e., false negative isolates, could be further examined and the reason for their resistance, such as encoding a novel AMR gene or an efflux pump, could be determined. Alternative methods of AST such as MIC by VITEK may yield AST results which would result in a better correlation with AST-WGS due to the reduction of unavoidable human error associated with manual methods.

To implement AST-WGS as a routine diagnostic test, standardisation of the DNA extraction, library preparation and bioinformatics analysis process would be necessary. There would be a need for harmonisation between laboratories and internationally, a huge task given the diversity in methods

of bioinformatics analysis currently (Moran-Gilad, 2017). Although WGS may not be a faster method, the main advantage of this method over phenotypic methods of AST is that WGS provides much more information on each isolate tested. Not only can the exact gene(s) that confer resistance in isolates be identified, other useful information can be extracted from the sequence, such as the potential virulence of the isolates. Their relatedness to other isolates that were previously sequenced can be useful in public health situations where outbreaks occur and for outbreak investigations and epidemiological tracebacks. In hospital settings, WGS could be used to differentiate between true infection and environmental contamination, which is not currently possible using phenotypic methods (Chapman et al., 2020).

Conclusion

The aim of this project was to determine if WGS could be used as an accurate predictor of AMR in bovine *E. coli* isolates. Results show promising agreement between AST-WGS and phenotypic AST. Resistance to tetracycline, florfenicol, gentamicin, trimethoprim-sulfamethoxazole and amoxicillin can be predicted using a genomic approach. However, more research is needed to study specific genes conferring resistance to amoxicillin-clavulanic acid, cephalothin, ciprofloxacin and marbofloxacin to improve correlation. Prediction of AMR phenotype requires careful curation of gene alleles conferring resistance, based on available literature.

Poor statistical correlation values may arise due to several factors such as 1) low numbers of resistant isolates in the sample set; 2) bovine *E. coli* having different genetic mechanisms of resistance when compared to human isolates; and 3) the large number of intermediately resistant isolates.

Acknowledgements

526 **Conflict of interest**

527 The authors declare no conflict of interest. Farid El Garach, Emmanuel Cuinet, Christine Miossec
528 and Frédérique Woehrlé are full-time employees of Vétquinol SA.

529

530 **Funding**

531 Vétquinol SA provided the isolates and financed the Illumina sequencing. DAN was funded by a
532 Department of Agriculture, Food and the Marine grant VTEC One Health.

533

534 **References**

535 Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M.,
536 Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler,
537 G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: A new genome assembly algorithm and its
538 applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477.
539 doi:10.1089/cmb.2012.0021

540 Beghain, J., Bridier-Nahmias, A., Nagard, H. Le, Denamur, E., Clermont, O., 2018.
541 ClermonTyping: An easy-to-use and accurate in silico method for Escherichia genus strain
542 phylotyping. *Microb. Genomics* 4. doi:10.1099/mgen.0.000192

543 Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: A flexible trimmer for Illumina sequence
544 data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu170

545 Card, R., Zhang, J., Das, P., Cook, C., Woodford, N., Anjum, M.F., 2013. Evaluation of an
546 expanded microarray for detecting antibiotic resistance genes in a broad range of gram-
547 negative bacterial pathogens. *Antimicrob. Agents Chemother.* 57, 458–465.
548 doi:10.1128/AAC.01223-12

549 Cassini, A., Högberg, L.D., Plachouras, D., Quattrocchi, A., Hoxha, A., Simonsen, G.S., Colomb-
550 Cotinat, M., Kretzschmar, M.E., Devleeschauwer, B., Cecchini, M., Ouakrim, D.A., Oliveira,
551 T.C., Struelens, M.J., Suetens, C., Monnet, D.L., Strauss, R., Mertens, K., Struyf, T., Catry, B.,
552 Latour, K., Ivanov, I.N., Dobrova, E.G., Tambic Andrašević, A., Soprek, S., Budimir, A.,
553 Paphitou, N., Žemlicková, H., Schytte Olsen, S., Wolff Sönksen, U., Märtin, P., Ivanova, M.,
554 Lyytikäinen, O., Jalava, J., Coignard, B., Eckmanns, T., Abu Sin, M., Haller, S., Daikos, G.L.,

555 Gikas, A., Tsiodras, S., Kontopidou, F., Tóth, Á., Hajdu, Á., Guólaugsson, Ó., Kristinsson,
556 K.G., Murchan, S., Burns, K., Pezzotti, P., Gagliotti, C., Dumpis, U., Liuimiene, A., Perrin,
557 M., Borg, M.A., de Greeff, S.C., Monen, J.C., Koek, M.B., Elstrøm, P., Zabicka, D., Deptula,
558 A., Hryniewicz, W., Caniça, M., Nogueira, P.J., Fernandes, P.A., Manageiro, V., Popescu,
559 G.A., Serban, R.I., Schréterová, E., Litvová, S., Štefkovicová, M., Kolman, J., Klavs, I.,
560 Korošec, A., Aracil, B., Asensio, A., Pérez-Vázquez, M., Billström, H., Larsson, S., Reilly,
561 J.S., Johnson, A., Hopkins, S., 2019. Attributable deaths and disability-adjusted life-years
562 caused by infections with antibiotic-resistant bacteria in the EU and the European Economic
563 Area in 2015: a population-level modelling analysis. *Lancet Infect. Dis.* 19, 56–66.
564 doi:10.1016/S1473-3099(18)30605-4

565 Chapman, P., Forde, B.M., Roberts, L.W., Bergh, H., Vesey, D., Jennison, A. V., Moss, S.,
566 Paterson, D.L., Beatson, S.A., Harris, P.N.A., 2020. Genomic investigation reveals
567 contaminated detergent as the source of an extended-spectrum- β -lactamase-producing
568 *Klebsiella michiganensis* outbreak in a neonatal unit. *J. Clin. Microbiol.* 58.
569 doi:10.1128/JCM.01980-19

570 Clermont, O., Bonacorsi, S., Bingen, E., 2000. Rapid and simple determination of the *Escherichia*
571 *coli* phylogenetic group. *Appl. Environ. Microbiol.* 66, 4555–4558.
572 doi:10.1128/AEM.66.10.4555-4558.2000

573 Clermont, O., Christenson, J.K., Denamur, E., Gordon, D.M., 2013. The Clermont *Escherichia coli*
574 phylo-typing method revisited: Improvement of specificity and detection of new phylo-groups.
575 *Environ. Microbiol. Rep.* 5, 58–65. doi:10.1111/1758-2229.12019

576 CLSI, 2021. CLSI. Performance Standards for Antimicrobial Susceptibility Testing. 31st ed. CLSI
577 guideline M100. Wayne, PA.

578 Cornaglia, G., Hryniewicz, W., Jarlier, V., Kahlmeter, G., Mittermayer, H., Stratchounski, L.,
579 Baquero, F., 2004. European recommendations for antimicrobial resistance surveillance. *Clin.*
580 *Microbiol. Infect.* 10, 349–383. doi:10.1111/j.1198-743X.2004.00887.x

581 Dryselius, R., Aswasti, S.K., Rajarao, G.K., Nielsen, P.E., Good, L., 2003. The Translation Start
582 Codon Region Is Sensitive to Antisense PNA Inhibition in *Escherichia coli*. *Oligonucleotides*
583 13, 427–433. doi:10.1089/154545703322860753

584 Enne, V.I., Delsol, A.A., Roe, J.M., Bennett, P.M., 2006. Evidence of Antibiotic Resistance Gene
585 Silencing in *Escherichia coli*. *Antimicrob. Agents Chemother.* 50, 3003–3010.

586 doi:10.1128/AAC.00137-06

587 Espinosa, L., Gray, A., Duffy, G., Fanning, S., McMahon, B.J., 2018. A scoping review on the
588 prevalence of Shiga-toxigenic *Escherichia coli* in wild animal species. *Zoonoses Public Health*
589 65, 911–920. doi:10.1111/zph.12508

590 Fàbrega, A., Madurga, S., Giralt, E., Vila, J., 2008. Mechanism of action of and resistance to
591 quinolones. doi:10.1111/j.1751-7915.2008.00063.x

592 Feldgarden, M., Brover, V., Haft, D.H., Prasad, A.B., Slotta, D.J., Tolstoy, I., Tyson, G.H., Zhao,
593 S., Hsu, C.H., McDermott, P.F., Tadesse, D.A., Morales, C., Simmons, M., Tillman, G.,
594 Wasilenko, J., Folster, J.P., Klimke, W., 2019. Validating the AMRFINDER tool and resistance
595 gene database by using antimicrobial resistance genotype-phenotype correlations in a
596 collection of isolates. *Antimicrob. Agents Chemother.* 63. doi:10.1128/AAC.00483-19

597 Giamarellou, H., Poulakou, G., 2009. Multidrug-resistant gram-negative infections: What are the
598 treatment options? *Drugs*. doi:10.2165/11315690-000000000-00000

599 Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: Quality assessment tool for
600 genome assemblies. *Bioinformatics* 29, 1072–1075. doi:10.1093/bioinformatics/btt086

601 Hiasa, H., 2002. The Glu-84 of the ParC subunit plays critical roles in both topoisomerase IV-
602 quinolone and topoisomerase IV-DNA interactions. *Biochemistry* 41, 11779–11785.
603 doi:10.1021/BI026352V

604 Holdsworth, S.R., Law, C.J., 2013. Multidrug resistance protein MdtM adds to the repertoire of
605 antiporters involved in alkaline pH homeostasis in *Escherichia coli*. *BMC Microbiol.* 2013 131
606 13, 1–19. doi:10.1186/1471-2180-13-113

607 Ito, C.A.S., Gales, A.C., Tognim, M.C.B., Munerato, P., Dalla Costa, L.M., 2008. Quinolone-
608 resistant *Escherichia coli*. *Brazilian J. Infect. Dis.* 12, 5–9. doi:10.1590/S1413-
609 86702008000100003

610 Jasovský, D., Littmann, J., Zorzet, A., Cars, O., 2016. Antimicrobial resistance—a threat to the
611 world’s sustainable development. *Ups. J. Med. Sci.* 121, 159–164.
612 doi:10.1080/03009734.2016.1195900

613 Karczmarczyk, M., Walsh, C., Slowey, R., Leonard, N., Fanning, S., 2011. Molecular
614 characterization of multidrug-resistant *Escherichia coli* isolates from Irish cattle farms. *Appl.*
615 *Environ. Microbiol.* 77, 7121–7127. doi:10.1128/AEM.00601-11

616 Kobayashi, K., Tsukagoshi, N., Aono, R., 2001. Suppression of hypersensitivity of *Escherichia coli*
617 *acrB* mutant to organic solvents by integrational activation of the *acrEF* operon with the IS1 or
618 IS2 element. *J. Bacteriol.* 183, 2646–2653. doi:10.1128/JB.183.8.2646-2653.2001

619 Krishnan, S., Balasubramanian, D., Raju, B.A., Lakshmi, B.S., Krishnan, S., Balasubramanian, D.,
620 Raju, B.A., Lakshmi, B.S., 2012. Use of a naturally occurring codon bias for identifying
621 topoisomerase mutations in ciprofloxacin resistant *Escherichia coli* using PCR and future
622 prospects with other bacterial genera: A pilot study. *Adv. Biol. Chem.* 2, 366–371.
623 doi:10.4236/ABC.2012.24045

624 Leimbach, A., Poehlein, A., Vollmers, J., Görlich, D., Daniel, R., Dobrindt, U., 2017. No evidence
625 for a bovine mastitis *Escherichia coli* pathotype. *BMC Genomics* 18, 1–22.
626 doi:10.1186/s12864-017-3739-x

627 Mackinnon, A., 2000. A spreadsheet for the calculation of comprehensive statistics for the
628 assessment of diagnostic tests and inter-rater agreement. *Comput. Biol. Med.* 30, 127–134.
629 doi:10.1016/S0010-4825(00)00006-8

630 McArthur, A.G., Waglechner, N., Nizam, F., Yan, A., Azad, M.A., Baylay, A.J., Bhullar, K.,
631 Canova, M.J., De Pascale, G., Ejim, L., Kalan, L., King, A.M., Koteva, K., Morar, M.,
632 Mulvey, M.R., O'Brien, J.S., Pawlowski, A.C., Piddock, L.J.V., Spanogiannopoulos, P.,
633 Sutherland, A.D., Tang, I., Taylor, P.L., Thaker, M., Wang, W., Yan, M., Yu, T., Wright,
634 G.D., 2013. The comprehensive antibiotic resistance database. *Antimicrob. Agents Chemother.*
635 57, 3348–3357. doi:10.1128/AAC.00419-13

636 Moran-Gilad, J., 2017. Whole genome sequencing (WGS) for food-borne pathogen surveillance and
637 control-taking the pulse. *Eurosurveillance* 1. doi:10.2807/1560-7917.ES.2017.22.23.30547

638 Moran, R.A., Anantham, S., Holt, K.E., Hall, R.M., 2017. Prediction of antibiotic resistance from
639 antibiotic resistance genes detected in antibiotic-resistant commensal *Escherichia coli* using
640 PCR or WGS. *J. Antimicrob. Chemother.* 72, 700–704. doi:10.1093/jac/dkw511

641 Nguyen, H.M., Graber, C.J., 2013. Cephalothin susceptibility testing as class representative for oral
642 cephalosporins: Is it time to move on? *Diagn. Microbiol. Infect. Dis.* 76, 483–485.
643 doi:10.1016/j.diagmicrobio.2013.03.022

644 Nishino, K., Yamaguchi, A., 2001. Analysis of a complete library of putative drug transporter genes
645 in *Escherichia coli*. *J. Bacteriol.* 183, 5803–5812. doi:10.1128/JB.183.20.5803-5812.2001

Punina, N. V., Makridakis, N.M., Remnev, M.A., Topunov, A.F., 2015. Whole-genome sequencing targets drug-resistant bacterial infections. *Hum. Genomics*. doi:10.1186/s40246-015-0037-z

Ribot, E.M., Freeman, M., Hise, K.B., Gerner-Smidt, P., 2019. PulseNet: Entering the Age of Next-Generation Sequencing. *Foodborne Pathog. Dis.* 16, 451–456. doi:10.1089/fpd.2019.2634

Ruiz, J., 2003. Mechanisms of resistance to quinolones: target alterations, decreased accumulation and DNA gyrase protection. *J. Antimicrob. Chemother.* 51, 1109–1117. doi:10.1093/jac/dkg222

Shigemura, K., Tanaka, K., Yamamichi, F., Shirakawa, T., Miyake, H., Fujisawa, M., 2012. Does mutation in *gyrA* and/or *parC* or efflux pump expression play the main role in fluoroquinolone resistance in *Escherichia coli* urinary tract infections?: A statistical analysis study. *Int. J. Antimicrob. Agents* 40, 516–520. doi:10.1016/J.IJANTIMICAG.2012.07.019

Smith, G., 2015. Antimicrobial decision making for enteric diseases of cattle. *Vet. Clin. North Am. - Food Anim. Pract.* doi:10.1016/j.cvfa.2014.11.004

Stoesser, N., Batty, E.M., Eyre, D.W., Morgan, M., Wyllie, D.H., Del Ojo Elias, C., Johnson, J.R., Walker, A.S., Peto, T.E.A., Crook, D.W., 2013. Predicting antimicrobial susceptibilities for *Escherichia coli* and *Klebsiella pneumoniae* isolates using whole genomic sequence data. *J. Antimicrob. Chemother.* 68, 2234–2244. doi:10.1093/jac/dkt180

Stubberfield, E., AbuOun, M., Sayers, E., O'Connor, H.M., Card, R.M., Anjum, M.F., 2019. Use of whole genome sequencing of commensal *Escherichia coli* in pigs for antimicrobial resistance surveillance, United Kingdom, 2018. *Eurosurveillance*. doi:10.2807/1560-7917.ES.2019.24.50.1900136

Su, M., Satola, S.W., Read, T.D., 2019. Genome-based prediction of bacterial antibiotic resistance. *J. Clin. Microbiol.* doi:10.1128/JCM.01405-18

Tyson, G.H., McDermott, P.F., Li, C., Chen, Y., Tadesse, D.A., Mukherjee, S., Bodeis-Jones, S., Kabera, C., Gaines, S.A., Loneragan, G.H., Edrington, T.S., Torrence, M., Harhay, D.M., Zhao, S., 2015. WGS accurately predicts antimicrobial resistance in *Escherichia coli*. *J. Antimicrob. Chemother.* 70, 2763–2769. doi:10.1093/jac/dkv186

Varughese, L.R., Rajpoot, M., Goyal, S., Mehra, R., Chhokar, V., Beniwal, V., 2018. Analytical profiling of mutations in quinolone resistance determining region of *gyrA* gene among UPEC. doi:10.1371/journal.pone.0190729

676 Vogwill, T., Maclean, R.C., 2015. The genetic basis of the fitness costs of antimicrobial resistance:
677 A meta-analysis approach. *Evol. Appl.* 8, 284–295. doi:10.1111/eva.12202

678 Wingett, S.W., Andrews, S., 2018. FastQ Screen: A tool for multi-genome mapping and quality
679 control. *F1000Research* 7, 1338. doi:10.12688/f1000research.15931.2

680 Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F.M.,
681 Larsen, M.V., 2012. Identification of acquired antimicrobial resistance genes. *J. Antimicrob.*
682 *Chemother.* doi:10.1093/jac/dks261

683 Zankari, E., Hasman, H., Kaas, R.S., Seyfarth, A.M., Agersø, Y., Lund, O., Larsen, M.V.,
684 Aarestrup, F.M., 2013. Genotyping using whole-genome sequencing is a realistic alternative to
685 surveillance based on phenotypic antimicrobial susceptibility testing. *J. Antimicrob.*
686 *Chemother.* 68, 771–777. doi:10.1093/jac/dks496

687 Zhou, Z., Alikhan, N.F., Mohamed, K., Fan, Y., Achtman, M., 2020. The Enterobase user's guide,
688 with case studies on Salmonella transmissions, Yersinia pestis phylogeny, and Escherichia
689 core genomic diversity. *Genome Res.* 30, 138–152. doi:10.1101/gr.251678.119

690

691

692 **Figures**

693 **Figure 1:** SNP tree generated in Enterobase, using CFS3246 as the reference, and visualised using
694 GrapeTree. *E. coli* isolates were coloured in based on the Clermont Type phylogrouping scheme
695 (A) and disease that the sample was recovered from (B).

Figure 2: A heatmap showing an absence/presence matrix detailing AMR genotypes and associated phylogroups for 143 *Escherichia coli* isolates of bovine origin. A colour key is shown to the right indicating the classes of compounds tested.

Figure 3: A bar graph showing the disk diffusion susceptibility profile for all isolates (**A**; n = 143), isolates of diarrhoeal origin (**B**; n = 73) and mastitis origin (**C**; n = 70). AML - Amoxicillin, AMC - Amoxicillin/Clavulanic acid, KF - Cephalothin, CIP - Ciprofloxacin, FFC - Florfenicol, CN – Gentamicin, MAR - Marbofloxacin, NA - Nalidixic acid, TE – Tetracycline, SXT - Trimethoprim-Sulfamethoxazole.

708

709 **Figure 4:** MIC susceptibility profile for all isolates (n=82). CN – Gentamicin, KF – Cephalothin,
710 TE – Tetracycline, MAR - Marbofloxacin, CIP – Ciprofloxacin and SXT - Trimethoprim-
711 Sulfamethoxazole.

Figure 5: Bar charts showing the distribution of isolates with specific resistant genotypes at each zone of inhibition for beta-lactam antibiotics: amoxicillin (A), amoxicillin-clavulanic acid (B), cephalothin (C). Numbers shown at the top of each column represents the number of isolates with the specific genotype, at that zone of inhibition. Zones of inhibition represent an average of 3 independent measurements and were rounded up to integer level. Red shaded background denotes the zone of inhibition for which the isolates are considered to be resistant, while blue/green shaded background denotes the zone of inhibition for which the isolates are considered to be susceptible. White background represents zones of inhibition of intermediately resistant isolates.

Figure 6: Bar charts showing the distribution of isolates with specific resistant genotypes at each zone of inhibition for florfenicol (A), gentamicin (B), tetracycline (C) and trimethoprim-sulfamethoxazole (D). Number at each column represents the number of isolates with the specific genotype, at that zone of inhibition. Zones of inhibition represent an average of 3 independent measurements and were rounded up to integer level. Red background denotes the zone of inhibition for which the isolates are considered to be resistant, while blue/green background denotes the zone of inhibition for which the isolates are considered to be susceptible. White background represents zones of inhibition of intermediately resistant isolates.

Figure 7: Bar charts showing the distribution of isolates with specific resistant genotypes at each zone of inhibition for quinolones: ciprofloxacin (A), nalidixic acid (B) and marbofloxacin (C). Number at each column represents the number of isolates with the specific genotype, at that zone of inhibition. Zones of inhibition represent an average of 3 independent measurements and were rounded up to integer level. Red background denotes the zone of inhibition for which the isolates are considered to be resistant, while blue/green background denotes the zone of inhibition for which the isolates are considered to be susceptible. White background represents zones of inhibition of intermediately resistant isolates.

Figure 8: Bar charts showing the distribution of isolates with specific resistant genotypes at each MIC for cephalothin (A), gentamicin (B), tetracycline (C), trimethoprim-sulfamethoxazole (D), ciprofloxacin (E) and marbofloxacin (F). Number at each column represents the number of isolates with the specific genotype, at that MIC. Red background denotes the MIC for which the isolates are considered to be resistant, while blue/green background denotes the MIC for which the isolates are considered to be susceptible. White background represents MIC of intermediately resistant isolates.

748 **Tables**

749 **Table 1:** A table showing the breakdown of test isolates (n=143) by country of origin, disease and
750 year isolated

Country of origin		Disease		Year of isolation	
France	67	Mastitis	70	2004	48
Germany	76	Diarrhoea	73	2007	48
				2010	47

751

752

Table 2: Antimicrobial resistance genes identified in *Escherichia coli* isolates of bovine origin isolated from animals suffering from diarrhoea and mastitis (n=143)

Antimicrobial class	Antimicrobial	AMR gene	Isolates positive for each gene (n)	Percentage of all isolates (%)
Aminoglycoside	Gentamicin	<i>aac(3)-Ile</i>	5	3.5
		<i>aac(3)-IVa</i>	4	2.8
		<i>ant(2'')-Ia</i>	8	5.59
Beta-lactam	Amoxicillin/Clavulanic acid	<i>ampC_C-42T</i>	3	2.1
		<i>bla_{CMY-2}</i>	1	0.7
		<i>bla_{OXA-1}</i>	5	3.5
		<i>bla_{OXA-2}</i>	2	1.4
		<i>bla_{TEM-30}</i>	2	1.4
		<i>bla_{TEM-1}</i>	53	37.06
		<i>bla_{CTX-M-1}</i>	2	1.4
	Cephalothin	<i>bla_{OXA-1}</i>	5	3.5
		<i>bla_{OXA-2}</i>	2	1.4
	Amoxicillin	<i>bla_{TEM}</i>	3	2.1
		<i>bla_{TEM-1}</i>	53	37.06
		<i>bla_{TEM-135}</i>	1	0.7
		<i>bla_{TEM-30}</i>	2	1.4
		<i>bla_{TEM-33}</i>	1	0.7
		<i>tet(A)</i>	44	30.77
		<i>tet(B)</i>	23	16.08
		<i>tet(C)</i>	2	1.4
		<i>tet(M)</i>	3	2.1
		<i>tet(Y)</i>	2	1.4
Fenicol	Florfenicol	<i>floR</i>	14	9.79
Fluoroquinolone	Marbofloxacin & Ciprofloxacin	<i>gyrA_D87G</i>	1	0.7
		<i>gyrA_D87N</i>	13	9.09
		<i>gyrA_D87Y</i>	3	2.1
		<i>gyrA_S83A</i>	1	0.7
		<i>gyrA_S83L</i>	22	15.38
		<i>parC_A56T</i>	1	0.7
		<i>parC_S57T</i>	15	10.49
		<i>parC_S80I</i>	13	9.09
		<i>parC_S80R</i>	3	2.1
		<i>parE_E460D</i>	1	0.7
		<i>parE_L416F</i>	1	0.7
		<i>parE_S458A</i>	2	1.4
		<i>parE_S458T</i>	2	1.4
		<i>marR_S3N</i>	2	1.4
		<i>gyrA_D87G</i>	1	0.7
Quinolone	Nalidixic acid	<i>gyrA_D87N</i>	13	9.09
		<i>gyrA_D87Y</i>	3	2.1
		<i>gyrA_S83A</i>	1	0.7
		<i>gyrA_S83L</i>	22	15.38
		<i>parC_A56T</i>	1	0.7
		<i>parC_S57T</i>	15	10.49
		<i>parC_S80I</i>	13	9.09
		<i>parC_S80R</i>	3	2.1

		<i>parE_E460D</i>	1	0.7
		<i>parE_L416F</i>	1	0.7
		<i>parE_S458A</i>	2	1.4
		<i>parE_S458T</i>	2	1.4
		<i>marR_S3N</i>	2	1.4
Sulfonamide/ trimethoprim	Trimethoprim- Sulfamethoxazole	<i>sul1</i>	43	30.07
		<i>sul2</i>	51	35.66
		<i>sul3</i>	1	0.7
		<i>dfrA1</i>	18	12.59
		<i>dfrA14</i>	1	0.7
		<i>dfrA17</i>	1	0.7
		<i>dfrA36</i>	7	4.9
		<i>dfrA5</i>	5	3.5
		<i>dfrA7</i>	1	0.7

755 **Legend:** AMR: Antimicrobial resistance

Table 3: A table showing the statistical correlation between genotypic (based on whole genome sequencing) and phenotypic (as determined by disk diffusion) methods of AMR detection, test performances and kappa correlations of *Escherichia coli* isolates of bovine origin (n=143).

Antibiotic	P+/G+	P-/G-	G+/P-	G-/P+	Specificity	Sensitivity	PPV	NPV	κ
Amoxicillin	65	77	0	1	100	98.5	100	98.7	0.99
Amoxicillin/ Clavulanic acid	10	120	1	12	99.2	45.5	90.9	90.9	0.56
Cephalothin	42	58	11	32	84.1	56.8	79.2	64.4	0.4
Ciprofloxacin	16	101	26	0	79.5	100	38.1	100	0.47
Florfenicol	13	129	1	0	99.2	100	92.9	100	0.96
Gentamicin	14	128	1	0	99.2	100	93.3	100	0.96
Marbofloxacin	16	101	26	0	79.5	100	38.1	100	0.47
Nalidixic acid	24	101	18	0	84.9	100	57.1	100	0.65
Tetracycline	64	76	1	2	98.7	97	98.5	97.4	0.96
Trimethoprim- Sulfamethoxazole	32	110	0	1	100	97	100	99.1	0.98
Overall	150	516	46	3	92.5	98.8	77.4	99.3	0.81

Footnote to Table 2: P+: Phenotype resistant; P-: Phenotype susceptible; G+: AMR gene detected; G-: AMR gene not detected; PPV: Positive Predicted Value; NPV: Negative Predicted Value; κ : Kappa correlation

Kappa correlation (κ): 0.01–0.20 (none to slight agreement), 0.21–0.40 (fair agreement), 0.41– 0.60 (moderate agreement), 0.61–0.80 (substantial agreement), 0.81–1.00 (almost perfect agreement).

764 **Table 4:** Number and percentage of intermediately resistant strains for each antibiotic

Antibiotic	No. of intermediately resistant strains	% of Total isolates	κ
Amoxicillin	0	0	0.99
Amoxicillin/ Clavulanic acid	10	7	0.56
Cephalothin	63	44.1	0.4
Ciprofloxacin	0	0	0.47
Florfenicol	0	0	0.96
Gentamicin	1	0.7	0.96
Marbofloxacin	1	0.7	0.47
Nalidixic acid	1	0.7	0.65
Tetracycline	1	0.7	0.96
Trimethoprim-Sulfamethoxazole	0	0	0.98

765

Table 5: Correlation between genotypic (whole genome sequencing) and phenotypic (MIC) methods of AMR detection, test performances and kappa correlations of *Escherichia coli* isolates of bovine origin (n=82)

Antibiotic	P+/G+	P-/G-	G+/P-	G-/P+	Specificity	Sensitivity	PPV	NPV	κ
Gentamicin	14	66	1	1	98.5	93.3	93.3	98.5	0.92
Cephalothin	26	24	27	5	47.1	83.9	49.1	82.8	0.27
Tetracycline	62	16	3	1	84.2	98.4	95.4	94.1	0.86
Marbofloxacin	15	40	27	0	59.7	100	35.7	100	0.35
Ciprofloxacin	16	40	26	0	60.6	100	38.1	100	0.38
Trimethoprim-Sulfamethoxazole	32	50	0	0	100	100	100	100	1
Overall	165	236	84	7	75	95.9	68.6	95.9	0.55

Footnote to Table 4: P+: Phenotype resistant; P-: Phenotype susceptible; G+: AMR gene detected; G-: AMR gene not detected; Spec: Specificity; Sens: Sensitivity; PPV: Positive Predicted Value; NPV: Negative Predicted Value; κ : Kappa correlation. **Kappa correlation (κ):** 0.01–0.20 (none to slight agreement), 0.21–0.40 (fair agreement), 0.41– 0.60 (moderate agreement), 0.61–0.80 (substantial agreement), 0.81–1.00 (almost perfect agreement).