1 **Linking phenotype, genotype and ecology: antimicrobial resistance in**
2 **the zoonotic pathogen *Streptococcus suis***
3
4 Nazreen F. Hadjirin, Eric L. Miller, Gemma G. R. Murray, Phung L. K. Yen, Ho
5 D. Phuc, Thomas M. Wileman, Juan Hernandez-Garcia, Susanna M.
6 Williamson, Julian Parkhill, Duncan J. Maskell, Rui Zhou, Nahuel Fittipaldi,
7 Marcelo Gottschalk, A.W. (Dan) Tucker, Ngo Thi Hoa, John J. Welch, Lucy A.
8 Weinert

9 **Abstract**

10 Antimicrobial resistance (AMR) is a major problem for human and animal
11 health, and can co-vary with features of the bacterial genome, its ecological
12 context, and with other traits. Here, we present an integrated study of AMR in
13 the porcine zoonotic pathogen, *Streptococcus suis*. Collecting a diverse panel
14 of 678 isolates, we obtained replicated measures of Minimum Inhibitory
15 Concentration (MIC) for 16 widely used antimicrobials, from eight distinct
16 classes. Then using whole genome sequences for each isolate, we identified
17 42 candidate AMR alleles, of which 21 were novel for *S. suis*. The presence of
18 these alleles explains most of the observed variation in MIC but there was
19 also unexplained variation, partly due to alleles of small effect, and epistatic
20 interactions. We find that MIC varies with *S. suis* genetic background and
21 country, but less so with clinical phenotype, serotype or over time. These
22 ecological patterns are largely explained by allele prevalence. MIC values for
23 different antibiotic classes positively correlate with one another, but there is
24 limited evidence of direct pleiotropy, where single alleles confer resistance to
25 multiple drug classes. Instead, this covariation is due to alleles for separate
26 classes appearing together, but not necessarily in the same genomic region.
27 The result is a significant excess of multi-drug resistant isolates. Indeed, we
28 found more isolates that carried resistance alleles to five classes of antibiotics
29 than isolates that carried no resistance alleles. We discuss the implications of
30 our results for combatting AMR in *S. suis*, and for understanding multi-drug
31 resistance in bacteria.

32

33

34 **Importance**

35 *Streptococcus suis* is a bacterium that causes serious disease with significant
 36 welfare and economic impacts for the global pig industry, and meningitis in
 37 humans. Antibiotics are the principal treatment for these bacteria, and so the
 38 evolution of antibiotic resistance presents a major problem for global food
 39 security and human health. We tested susceptibility to 16 different antibiotics
 40 in a global sample of 678 *S. suis* isolates. Using comparative analysis of
 41 whole genome sequences, we identified existing and novel genetic variants
 42 that affect resistance. Results show that, in most cases, the
 43 presence/absence of these alleles is a good predictor of resistance levels,
 44 which could help facilitate resistance monitoring through genome sequencing.
 45 We also observe a tendency for AMR variants to co-occur together in some *S.*
 46 *suis* isolates. These multi-drug resistant isolates should be monitored as a
 47 priority.

48

49 **Introduction**

50 The ability of bacterial pathogens to evolve resistance to antimicrobials is one
 51 of the gravest threats to human health and food security worldwide.
 52 Antimicrobial resistance (AMR) to a given drug can be quantified via the
 53 minimum inhibitory concentration (MIC), i.e., the concentration of the drug that
 54 is sufficient to inhibit growth of a bacterial culture. While, for practical use,
 55 bacterial isolates are often categorised as either susceptible or resistant, MIC
 56 is, in principle, a continuously varying trait. Like any other such trait, it may co-
 57 vary with aspects of the population ecology, with features of their genomes,
 58 and with other phenotypic traits. AMR cannot be fully understood without
 59 considering all of these contexts. In particular, the prevalence of multi-drug
 60 resistance, implies that resistance cannot be understood on a drug-by-drug
 61 basis. In this study, we present a comprehensive portrait of AMR in the
 62 bacterium *Streptococcus suis*.

63

64 *S. suis* primarily exists as a commensal in pigs, colonising the nasopharynx,
 65 gut and vagina, but it also causes systemic and respiratory disease,
 66 particularly in young pigs (1). *S. suis* is also a serious zoonotic disease, being
 67 the leading cause of adult bacterial meningitis in Vietnam (2). While some
 68 autogenous vaccines are used in pig production, they are serotype-specific
 69 and give inconsistent cross protection against heterogenous *S. suis* (3).
 70 Antimicrobials therefore remain the standard treatment for *S. suis*, and as
 71 such, *S. suis* is a leading driver of antimicrobial usage in pig farms (4).

72

73 As well as being a serious problem in itself, *S. suis* also has unique benefits
 74 as a model for studying AMR. By weight, more pork is consumed globally than
 75 any other meat (5, 6), and *S. suis* is found in most, if not all pigs (7).
 76 Furthermore, antibiotic consumption is higher in pigs (172 mg per population
 77 corrected unit) than any other livestock (e.g. cattle (45 mg) and chicken (148
 78 mg)) (8). As a result, *S. suis* is extremely diverse, and yet most lineages
 79 experience antibiotics. These include not only antibiotics administered directly
 80 against *S. suis*, whether as a therapeutic, prophylaxis or metaphylaxis, but
 81 also, and perhaps more commonly, in response to many other bacterial
 82 infections.

83

84 The strong selection pressure caused by widespread use of antimicrobials in
 85 pig farming is expected to give rise to AMR in *S. suis*. Consistent with this,
 86 several phenotypic studies show high MICs for each of the major classes of
 87 antibiotics in one or more *S. suis* collections (9-13). In addition, there have
 88 been demonstrations of individual resistance alleles affecting MIC in *S. suis*
 89 (14-16), and also some mining of *S. suis* genome collections for known
 90 resistance alleles (17-19). However, to our knowledge, no study has
 91 combined complete genomic and phenotypic information in large numbers of
 92 isolates, collected from a diverse range of populations.

93

94 To this end, we obtained replicated measures of MIC for 16 antibiotics, each
 95 widely used in pigs, for six different *S. suis* collections, comprising 678
 96 isolates. These collections were chosen to include the three main pig-
 97 producing regions of the globe - namely the Americas, South East Asia and

98 Europe - and targeted both human and pig hosts, as well as a range of years,
99 serotypes and clinical phenotypes (20-22). We obtained whole genome
100 sequences for each of our isolates, allowing us to compare phenotype and
101 genotype on an unprecedented scale.

102 **Results**

103 **Measurement of MIC for 16 antibiotics in *S. suis***

104 We scored 678 isolates of *Streptococcus suis* for replicated measures of
 105 Minimum Inhibitory Concentration (MIC) for 16 antibiotics (Table S1). Four of
 106 these are beta-lactams (amoxicillin, cefquinome, ceftiofur and penicillin),
 107 which are typically used to treat *S. suis* infection in pigs. The other antibiotics
 108 are all widely used in the pig industry, but found in different drug classes.
 109 They comprise macrolide-lincosamide-streptogramin B (MLS_B: erythromycin,
 110 lincomycin, tilmicosin and tylosin), tetracyclines (doxycycline and tetracycline),
 111 and fluoroquinolones (enrofloxacin and marbofloxacin), plus one each of an
 112 aminoglycoside (spectinomycin), a pleuromutilin (tiamulin), trimethoprim
 113 (TMP), and a phenicol (florfenicol). Results, shown in the left hand plot of
 114 Figure 1, reveal wide variation among our isolates in MIC for most of the
 115 antibiotics. For a few antibiotics, especially the MLS_B and tetracyclines, the
 116 MIC values are clearly bimodal, suggesting a meaningful division between
 117 “resistant” and “wild-type” isolates. However, for most antibiotics, the
 118 distributions are either roughly lognormal (fluoroquinolones and phenicols) or
 119 positively skewed (spectinomycin, pleuromutilin, TMP and beta-lactams). It is
 120 striking that the distribution for beta-lactams, the most common treatment
 121 against *S. suis*, are not clearly outlying. Given these distributions, and the
 122 general lack of defined clinical breakpoints for these antibiotics in *S. suis* (23),
 123 we treated MIC as a quantitative trait in our subsequent analyses.

124

125 **Ecological and genomic predictors of MIC**

126 Our *S. suis* isolates were collected at different times, and from different
 127 countries, hosts and body sites, including sites associated with respiratory
 128 and systemic disease. Isolates were also genetically heterogeneous, and
 129 differed in their serotype; including serotypes with a known disease
 130 association (see methods and Table S1). We first used linear models, to ask
 131 whether MIC levels varied systematically with these ecological and genomic
 132 factors.

133
 134 *S. suis* is highly recombining, such that no single genealogy characterises its
 135 diversity. As such, to model overall genomic diversity in our sample, we
 136 identified 30 clusters of genetically-similar strains (see methods). The
 137 resulting clusters were a highly significant predictor of MIC ($p < 10^{-15}$; Table 1a)
 138 showing that genetically similar isolates tend to have similar MIC values. To
 139 account for this genetic structure, "cluster" was included as a random effect in
 140 subsequent analyses.

141
 142 A model including all potential predictors showed that country of isolation was
 143 also a highly significant predictor of MIC. In addition, there were weaker
 144 effects of year of collection (with a slight trend for increasing MIC over time);
 145 for serotype (with non-disease-associated serotypes tending to have higher
 146 MIC); and for host disease status (with non-clinical isolates having higher
 147 MIC) (Table 1b).

148
 149 Separate analyses for each antibiotic, showed that the effect of country was
 150 driven by consistently higher MIC in the samples from Canadian pigs. This

151 applied to 15/16 antibiotics, and to various models and subsets of the data
152 (Figure S1a; Table 1c). By contrast, the other predictors (year, serotype and
153 clinical status) had effects that were less consistent or robust (Table 1d-f). For
154 example, the trend for higher MIC in non-clinical isolates (Table 1d), applied
155 consistently only in tetracyclines and TMP (Figure S2a).

156

157

158 **Identification of candidate AMR alleles within *S. suis* genomes**

159 Using a candidate locus approach, we next identified 42 known and putative
160 AMR alleles that were present in our dataset: nine novel alleles of known
161 resistance genes, 12 known resistance alleles detected in *S. suis* for the first
162 time, and 21 previously reported in *S. suis* (Table 2 and Table S1). Our
163 candidate alleles come from three sources: the antimicrobial resistance
164 database CARD (24), previously published *S. suis* variants not present in the
165 CARD database (25-27) and novel alleles detected using a subset of our data
166 (see methods).

167

168 Of the nine novel alleles, five are variants in genes previously associated with
169 resistance in other bacteria (including a promoter variant). In particular, we
170 discovered novel haplotypes within the loop region of the central
171 transpeptidase domains of the *pbp2B* and *pbp2X* genes and within the signal
172 peptide region of the *mraY* gene (encoding the enzyme, acetylmuramoyl-
173 pentapeptide-transferase, essential to cell wall biosynthesis) that were
174 associated with variation in beta-lactam MIC (Table 2). In addition, a variant of
175 the chromosomal dihydrofolate reductase gene *dhfr* (allele frequency of 0.34)

176 and its promoter (allele frequency of 0.13) were associated with reduced
177 susceptibility against trimethoprim (Table 2). Furthermore, these alleles were
178 independently associated with changes in MIC in at least seven different
179 genetic clusters (Table S2) implying that the association between these
180 haplotypes and MIC is either causal or directly compensatory of the causal
181 mechanism (28).

182

183 The remaining four novel alleles, were genes with homologies to known
184 resistance genes, three located on mobile genetic elements (MGEs) and one
185 chromosomal (Table 2). Three novel aminoglycoside genes, carried by
186 isolates with high spectinomycin MIC, were identified based on
187 aminoglycoside resistance determinant protein homologies. Although
188 previously undescribed, we found homologues in other bacteria using a blastn
189 search of the non-redundant database in Genbank (Table 2). We also
190 characterised a chromosomally-encoded *vgaC* homologue (37% protein
191 homology to *vgaC*), encoding an ABC-F ATP-binding cassette ribosomal
192 protection protein, that we designate *vgaF* (allele frequency of 0.25)
193 associated with reduced susceptibility to tiamulin.

194

195 The 12 AMR alleles detected in *S. suis* for the first time are mobile genetic
196 element (MGE)-linked genes that confer resistance to aminoglycosides ($n=2$),
197 MLS_B ($n=2$), tetracyclines ($n=5$), TMP ($n=2$) and phenicol ($n=1$). Based on the
198 CARD and NCBI databases, these genes are associated with other gram-
199 positive and gram-negative bacteria (Table 2).

200

Overall, most of our candidate alleles (28/42) were found in fewer than 5% of the isolates, but *ermB* (which confers resistance to the MLS_B class) and *tetO* (which confers resistance to tetracyclines) were present in a majority of the isolates (Table 2; Figure S3). Only about a quarter of isolates carry a candidate allele for the beta-lactams.

Candidate AMR alleles explain most of the variation in MIC

We next investigated how well our candidate AMR alleles explain the variation we observe in MIC. The right-hand panel of Figure 1 compares the distribution of MIC values for isolates carrying one or more of the candidate alleles for that antibiotic class (denoted A for ‘allele’), to the remaining isolates, which carry no such allele (N for ‘no allele’).

The proportion of the variation in MIC explained by our alleles varied between antibiotics. Explanatory power was strongest ($r^2 > 0.75$) for antibiotics where the distributions of MIC were most clearly bimodal (MLS_B and tetracyclines). By contrast, we found no association between our alleles and MIC for the fluoroquinolones and phenicol, where the distribution of MIC values was roughly lognormal, consistent with our isolates representing a “wild-type” population. For the remaining antibiotics, our alleles had intermediate explanatory powers (Figure 1), but there were very few isolates with high MIC that did not carry a candidate allele. This suggests that we have detected most of the causal alleles of large effect in our dataset, removing the need to perform additional genome-wide associations. Where our alleles show low

225 explanatory power ($r^2 < 0.5$), this appears to be caused mainly by isolates that
226 carry an allele but whose MIC is low (Figure 1).

227

228 We conjectured that the variable presence of our candidate alleles in different
229 ecological settings would explain the significant ecological predictors of MIC in
230 our linear models (Table 1).

231

232 Consistent with this hypothesis, we found that the fitted MIC in genetic
233 clusters was highly correlated with the candidate allele frequency in that
234 cluster (Figure S4), and that candidate allele frequencies are consistently
235 higher in Canada (Figure S3). Accordingly, when we include candidate allele
236 presence in our models, the effect size for country greatly decreases (Figure
237 S1b). Nevertheless, this cannot be the whole picture, because genetic cluster
238 and country remain significant predictors for all 16 antibiotics, even when
239 allele presence is included (Table S3). Including allele presence does
240 however remove the effects of the other weaker ecological predictors. For
241 example, disease status is no longer a significant predictor for 15/16
242 antibiotics (enrofloxacin remains significant but, as mentioned above, the
243 candidate alleles explain very little variation in this case) (Figure S2b; Table
244 S3).

245

246 A possible explanation for the unexplained variation in Figure 1, and the
247 continuing predictive power of genetic cluster and country, is the way in which
248 we have included our candidate alleles as predictor variables (i.e. an isolate is
249 scored as positive if it contains any one allele for that class).

250

251 In some cases, simple amendments to the model increase explanatory power.
 252 Spectinomycin MIC, for example, is unchanged by some AMR genes from the
 253 CARD database (e.g. *ant(6')-Ib* and *aph(3')-IIIa* in Figure S5). Including each
 254 of the candidate alleles as a separate predictor (i.e., a two-level factor:
 255 present/absent) removes the influence of these non-functional alleles, and
 256 improves the r^2 value from 0.37 (Figure 1) to 0.69 (Figure S6a).

257

258 A similar improvement is also seen for the beta-lactams (Figure S7), and in
 259 this case, it appears that resistance has a polygenic architecture with alleles
 260 acting additively on log MIC (Figure S7). This implies that isolates will have
 261 high MIC only if they carry multiple alleles. Accordingly, if we predict MIC from
 262 the total number of candidate alleles that are carried (treating "number of
 263 candidate alleles carried" as a continuous predictor variable), the beta-lactam
 264 r^2 values improve from 0.16-0.53 (Figure 1) to 0.35-0.63 (Figure S6b).

265

266 These simple explanations (ineffective or small-effect alleles) apply only to
 267 some of the antibiotics (Figure S6). In other cases, our data suggest other
 268 complexities (Figure S8-S12). Sometimes, as is well known, alleles act
 269 against only some of the antibiotics in a class. For example, while modification
 270 of the ribosomal target confers cross resistance to macrolides and
 271 lincosamides, alleles encoding mechanisms such as efflux and enzymatic
 272 inactivation affect only some of the antibiotics within the class (29). This is
 273 illustrated in Figure 2, where *mel* acts against erythromycin, but not against
 274 lincomycin, while *linB* inactivates lincomycin, but not erythromycin. We also

see that some alleles generally cause high MIC, but are non-functional in some isolates. For example, for isolates carrying *ermB*, the distributions of MIC values are clearly bimodal (Figure 2). Some of this variation appears to be allelic. Examining the 363 isolates with a complete *ermB* sequence in our assemblies, we found that 68/363 were associated with low level resistance (MICs ≤ 32 mg/ml). Of the 363, 304 isolates carried one or more rare amino acid variants, and these isolates had a lower mean MIC for erythromycin (Welch's t test $t = 4.1231$, $df = 148.35$, $p < 10^{-4}$). Nevertheless, some isolates carrying the wild-type *ermB* also had low MIC. This is likely an epistatic effect between *ermB* and its genetic background. One possible interaction of this kind involves another candidate allele: *InuB*, which alone has no effect on erythromycin, but decreases MIC in the presence of *ermB* (Figure 2). The non-linear effects of genotype, evident in Figure 2, are the most probable explanation of the remaining unexplained variation in MIC in our data.

289

290 **Covariation in MIC between antibiotic classes**

Results above consider each antibiotic independently, but our data also show strong positive covariation in MIC between different drug classes (Figure 3).

Of 120 pairwise correlations, 112 are positive, and all 8 exceptions involve enrofloxacin, where there is little variation in MIC (Figure 1a; Figure 3 inset).

Covariances are particularly strong for beta-

lactams/pleuromutilin/TMP/aminoglycoside, along with MLS_B/tetracyclines and MLS_B/phenicol (Figure 3).

298

299 The simplest hypothesis to explain this pattern would be that some alleles act
 300 against more than one antibiotic class (i.e. direct pleiotropy). To test for
 301 pleiotropy in our data, we asked whether isolates with and without candidate
 302 alleles for a different antibiotic class, tended to differ in their MICs (Figure
 303 S13). The strongest support for the hypothesis of direct pleiotropy was found
 304 in the tetracyclines. Here, isolates containing candidate alleles for non-
 305 tetracycline drugs showed higher MIC for tetracycline, whether or not they
 306 also carried tetracycline-resistance alleles (Figure S13). The same result was
 307 observed in doxycycline, but only for isolates carrying tetracycline-resistance
 308 alleles. Overall, however, these were isolated results, with modest effect
 309 sizes. No other class showed consistent evidence of pleiotropy (Figure S13),
 310 and there was little evidence of pleiotropic effects for the most common single
 311 alleles (*ermB*, *tetO*, and *dhfr102*; see Figure S14).

312

313 If direct pleiotropy cannot explain the observed covariation in MIC, an
 314 alternative explanation is that alleles of different classes tend to occur
 315 together in isolates. Consistent with this, we found that the co-occurrence
 316 between candidate alleles from different antibiotic classes strongly predicts
 317 the correlation in their MIC values (upper right graph in Figure 3, Mantel test
 318 $p < 10^{-4}$).

319

320 An obvious explanation for the co-occurrence of our alleles is their physical
 321 linkage, and consistent with this, many of our candidate alleles occur together
 322 on mobile genetic elements. One example is the integrative conjugative
 323 element (ICE) ICESsu584 (GenBank: KT716758), which carries two

324 aminoglycoside resistance genes *ant6-1a* and *ant9-1a*, a lincosamide
 325 resistance gene *InuB* and a tiamulin resistance gene *IsaE*. This may explain
 326 the observed covariation in MIC between spectinomycin, tiamulin and
 327 lincosamide ("spc", "tia" and "lcm" in Figure 3). In addition, the two most
 328 common alleles, *ermB* and *tetO* often co-occur on ICEs (19) and this may
 329 explain the co-variation between tetracycline and MLS_B.

330

331 We also see physical linkage of alleles within the chromosome. The *vgaF*
 332 homologue is often located upstream of the *pbp2B* gene (Figure S15), which
 333 may explain the covariation between beta-lactams and pleuromutilin (Figure
 334 3).

335

336 Furthermore, there is a general and significant trend in the data for co-
 337 occurring alleles to be closer together in the genome (see Figure S15; Mantel
 338 test $p < 10^{-4}$). Nevertheless, the strength of the correlation is weak, and linkage
 339 clearly cannot explain all of the co-occurrence. Some highly correlated alleles
 340 are very distant in our genomes. For example, the *vgaF* homologue co-occurs
 341 with the *dhfr102* variant ($\rho = 0.51$), yet when they assemble on the same
 342 contig, they are located a median of 367,404bp apart (Figure S15).

343

344 **AMR allele clustering**

345 A corollary of allelic co-occurrence is a high level of clustering in the number
 346 of candidate AMR alleles carried. Our data contained a class of "globally
 347 susceptible" isolates, which carry no candidate allele (60/678), and out of the
 348 many isolates (525/678) that carry multiple candidate alleles (Figure 4a),

almost all of these (508/525) were "multi-drug resistant" (MDR), carrying candidate alleles for two or more antibiotic classes. Indeed, more isolates carry alleles for five distinct classes of antibiotic than carry no candidate allele at all (Figure 4b).

353

Without a well-defined null model, it is not clear whether this co-occurrence calls for any special explanation, (e.g. this level of MDR may simply reflect the high number of individual AMR alleles in our *S. suis* samples). To generate null expectations, we used three randomization procedures. First, we randomly permuted allelic states across the isolates. Results (Table 3; "over isolates") show that, compared to the real data, the randomized data sets had increased levels of resistance overall, with significantly fewer susceptible isolates, but significantly more MDR isolates (an illustrative randomisation of this kind is shown as the narrow red bars in Figure 4).

363

Randomisation across isolates implicitly assumes that all alleles evolve independently, so we next made the opposite assumption, assuming clonal evolution, and placing each of our resistance alleles, at random, on the consensus phylogeny of our isolates. As shown in Table 3 ("across phylogeny") this type of randomisation had the opposite effect to randomisation across isolates, leading to significantly more susceptible isolates, but significantly less MDR (an illustrative randomisation is shown as the narrow yellow bars in Figure 4).

372

373 Because these procedures make extreme assumptions of different kinds, we
 374 next combined them, applying each to a subset of our candidate alleles. For
 375 the 31/42 alleles that were found on MGEs (see Table 2), we randomised
 376 across isolates, while the remaining 11 alleles were placed at random on the
 377 phylogeny. These randomised data sets (Table 3; "both") were much closer to
 378 the true data. However, they were much less clustered, with significantly
 379 fewer susceptible isolates, and significant fewer MDR isolates than the true
 380 data. Together, then, we can conclude that the clustering of AMR in our data
 381 is greater than would be expected by chance.

382

383 We also examined our data for a "nested" pattern of resistance in which lower
 384 frequency AMR alleles tend to appear in genetic backgrounds that already
 385 contain an AMR allele (see (30) for details). In total 63% of our isolates were
 386 nested in this way (Figure S16), but this value was not clearly higher than
 387 expectations under the null model of evolution over the genealogy (Table 3,
 388 bottom row).

389

390

391 Discussion

392 We have examined antimicrobial susceptibility to 16 commonly-used
393 antibiotics in *Streptococcus suis*, using replicated measures of MIC. Data
394 suggest that, even when some isolates have very high MIC, there is often no
395 natural separation into "resistant" and "susceptible". This highlights the need
396 to define clinical breakpoints for these antibiotics in *S. suis* to understand the
397 relevance of these MIC measurements in treatment.

398
399 We have described 42 candidate resistance alleles, and showed that they
400 explain the majority of the observed variation in MIC for 11/16 antibiotics, and
401 a substantial fraction of the variation for another 2/16 (Figure 1; Figure S6).

402
403 While many of these alleles were previously described, and most were found
404 at low frequency (Table 2; Figure S3), we detected three novel candidates
405 that appear in many isolates with clinically-relevant changes in MIC: *vgaF*,
406 *dhfr102* and the *dhfr* promoter. If these resistance mechanisms are
407 functionally verified, they should be included in routine AMR gene testing for
408 *S. suis*.

409
410 Of particular interest are the novel haplotypes that are linked to changes in
411 beta-lactam MIC, because penicillin is the first line treatment for *S. suis*
412 infection. Functional studies will be needed to show whether the *S. suis* PBP
413 alleles uncovered here confer a similar mode of resistance to beta-lactams as
414 is seen in other streptococci (31-35). Studies have shown that cefotaxime, a
415 third-generation cephalosporin like ceftiofur, selectively inactivates PBP2X but

not PBP2B (36) and that mutations within the catalytic domain in PBP2X give rise to cefotaxime, but not to penicillin, resistance (32). Our results echo this pattern because our PBP2B haplotype has a smaller change on mean MIC of ceftiofur than penicillin (Figure S7), and the explanatory effects of our PBP haplotypes are different for the different beta-lactams (Figure 1). Previous variants in PBP2X thought to confer cefuroxime resistance (25), whilst validated by our data, were found at very low frequency in our collection (Table 2). We show that our PBP haplotypes have additive effects on mean log MIC such that some of our strains reach clinically relevant changes in beta-lactam MIC (e.g. MIC > 1 mg/L for Penicillin) but that no one allele is associated with this change. Additive effects have also previously been shown for *tet* alleles in *S. suis* (16). Our work illustrates a need to address how these additive effects can be incorporated in to predictive genome models of MIC.

Unlike the PBPs, the contribution of mutations in *mraY* to changes in beta-lactam MIC might not be directly causal. It has been suggested that altered PBPs, while conferring resistance, might be less active transpeptidases than their wild type counterparts, and that organisms are likely to evolve compensatory mechanisms to circumvent this problem (31). For example, the amidotransferase enzyme encoded by *murT* enables cross-linking of cell wall peptidoglycans by preparing lipid II used by pneumococcal PBPs (37, 38). Mutations in *mraY* are therefore suggestive of a possible compensatory mechanism for reduced enzymatic activity of altered PBPs in *S. suis*. In agreement with this, Figure S7 shows that the *mraY* haplotypes are only found in concert with an altered PBP. Polymorphisms in *mraY* were also

441 identified in a genome-wide association study of beta-lactam resistance in *S.*
 442 *pneumoniae* (39) and could make a promising drug target for beta-lactam
 443 resistant streptococci.

444

445 While our candidate alleles were explanatory, we also found several
 446 complexities in the genotype-phenotype map. These include epistatic
 447 interactions, where the effects of some candidate genes vary with their *S. suis*
 448 genetic background. Phenotypic reversion (i.e. mutations in transcriptional
 449 regulators or elsewhere in the genome, counteracting a resistance allele) can
 450 be common if resistance has a fitness cost (40), and this could explain why
 451 we see many isolates carrying candidate alleles but with low MIC (Figure 1).
 452 These complexities have consequences for diagnostic investigation of AMR of
 453 *S. suis* using whole genome sequences. While the predictive power is quite
 454 high for many antibiotics, we should expect many Type I errors. These results
 455 also caution against using AMR allele presence as a measure of phenotypic
 456 resistance more generally.

457

458 We have also shown suggestive links between antibiotic exposure and non-
 459 susceptibility. In particular, antibiotic usage is much higher in Canada than the
 460 UK (41, 42) and this is reflected in consistently higher MIC, and consistently
 461 higher frequency of the same candidate AMR alleles (Figures S1, S3).
 462 Results for Vietnam are more difficult to interpret, partly due to a lack of
 463 official records of antibiotic usage (although see (43)), and partly because our
 464 Vietnamese sample is relatively small and genomically homogeneous (see
 465 methods).

466

467 When considering results for all of the 16 antibiotics together, we have shown
 468 that our data contain a significant excess of MDR. This is evident both in
 469 highly correlated MIC values and, indirectly, via the co-occurrence of alleles.
 470 Notably, we have found that isolates with candidate resistance alleles to five
 471 different drug classes are more common than isolates with no resistance
 472 alleles at all (Figure 4b).

473

474 While linkage between resistance alleles plays some role in explaining this
 475 pattern (Figure S15), we argue that it cannot do so alone. One possible factor
 476 is resistance to one antibiotic somehow facilitates the evolution of resistance
 477 to a second. We found limited evidence of direct pleiotropy (where AMR
 478 alleles to a previously administered drug provide some action against a novel
 479 drug; Figure S13-S14), and so this might plausibly involve indirect pleiotropy,
 480 (e.g., mitigation of a cost to antibiotic resistance that applies to multiple drug
 481 classes).

482

483 However, a simpler explanation for the observed pattern in MDR is correlated
 484 selection pressures, where isolates are exposed to many antibiotics, or to
 485 none (30). Some forms of correlated selection lead to a characteristic pattern
 486 of nested MDR, where rarer alleles tend to be found on backgrounds
 487 containing more common alleles (30). But while we found some nestedness in
 488 our data, it was not clearly above chance levels (Figure S16; Table 3).
 489 Correlated selection pressures might also result if isolates have very different
 490 durations of carriage (44). It is notable, in this respect, that we have found

491 higher levels of non-susceptibility in isolates from the tonsils of pigs without
 492 clinical disease (Table 1; Figure S2).
 493
 494 Two final aspects of our results highlight the fact that AMR is not a problem
 495 that can be tackled one disease at a time. First, we have found high levels of
 496 resistance in *S. suis* even to antibiotics other than beta-lactams that are
 497 typically used to treat this infection. In fact, resistance alleles to four of the
 498 seven other classes were more common (Figure S16). This must reflect a
 499 "bystander selection effect", common in opportunistic pathogens that are part
 500 of the healthy microbiota and frequently exposed to antibiotics used as growth
 501 promoters or therapeutics. Second, many of the rarer AMR alleles in our
 502 sample were not previously known in *S. suis*, but are common in other
 503 bacteria, both gram negative and gram positive. This raises the possibility of
 504 *S. suis* acting as a reservoir for AMR alleles in other bacteria. Indeed *S. suis*
 505 shares conserved chromosomal insertion sites with many human pathogens,
 506 such as *S. pyogenes*, *S. pneumoniae*, and *S. agalactiae* (17). Given that *S.*
 507 *suis* causes human clinical disease, a direct exchange of AMR alleles
 508 between *S. suis* and other human pathogens is plausible.
 509

510 **Material and Methods**

511

512 ***Streptococcus suis* isolates**

513 We assembled a collection of 678 *S. suis* isolates, including isolates from the
514 UK ($n=423$), Canada ($n=205$) and Vietnam ($n=50$) (Table S1).

515

516 In the UK, isolates came from three different collections. The first collection in
517 2009-2011 sampled non-clinical and clinical isolates from pigs across England
518 and Wales (described in Weinert *et al.* 2015 (20)). The second collection in
519 2013-2014 sampled non-clinical isolates from five farms (described in Zou *et*
520 *al.* 2018 (22)). The third collection during 2013–2014 targeted clinical isolates
521 from pigs across England and Wales (described in Wileman *et al.* 2019 (21)).

522 In pigs that showed clinical symptoms consistent with *S. suis* infections (e.g.
523 meningitis, septicaemia and arthritis), the site of isolation was classified as
524 ‘systemic’ if recovered from systemic sites. The site of recovery was classified
525 as ‘respiratory’ if derived from lungs with gross lesions of pneumonia. *S. suis*
526 isolates from the tonsils or tracheo-bronchus of healthy pigs or dead pigs
527 without any typical signs of *S. suis* infections were defined as ‘non-clinical’.

528 Isolates that could not confidently be assigned to these categories (e.g. a
529 tonsil isolate from a pig with systemic signs) were classified as unknown.

530 Altogether, the UK isolates were classified by clinical status as ‘systemic’
531 ($n=94$), ‘respiratory’ ($n=50$), ‘non-clinical’ ($n=197$) or ‘unknown’ ($n=82$),
532 respectively.

533

534 The Canadian pig *S. suis* isolates from 1983-2016 were collected to target
535 similar numbers of clinical and non-clinical isolates and were also classified by
536 clinical status as 'systemic' ($n=81$), 'respiratory' ($n=30$), 'non-clinical' ($n=55$) or
537 'unknown' ($n=39$).

538

539 The Vietnamese isolates were collected to sample related populations from
540 human and pig (described in Weinert *et al.* 2015 (20)). These comprised
541 'systemic' isolates ($n=28$) from human clinical cases of meningitis from
542 provinces in southern and central Vietnam, and 'systemic' ($n=4$) or 'non-
543 clinical' isolates ($n=18$) from pigs, collected between 2000 and 2010. These
544 isolates were exclusively serotype 2 or serotype 14 and belong to one genetic
545 population (Table S1).

546

547 **Antimicrobial susceptibility testing**

548 The minimum inhibitory concentrations (MIC) for a range of antibiotics were
549 determined by the micro-broth dilution method, which was performed and the
550 results interpreted in accordance with CLSI Approved Standards, M100-S25
551 (2015), Vet01S 3rd Edition (2015) and de Jong *et al.* (2014) (9, 45). MIC
552 measurements for some of our isolates were previously published (9). For the
553 remaining MIC measurements, the MICs were determined for sixteen different
554 antimicrobial compounds, representing nine antimicrobial classes at LGC,
555 Fordham, UK (formerly Quotient Bioresearch, Fordham, UK) for the UK and
556 the Canadian isolates. MIC testing of antibiotics for the Vietnamese isolates
557 was performed at the Oxford University Clinical Research Unit, Ho Chi Minh

City, Vietnam, in collaboration with the Department of Veterinary Medicine,
University of Cambridge, UK.

Whole genome sequencing, assembly and inference of population structure

Genome DNA extractions and whole genome sequencing of newly sequenced
S suis in this study were as previously described by Weinert *et al.* (2015) (20).
Briefly, single colonies of strains were grown up in broth culture, DNA was
extracted using DNeasy kits (Qiagen), Illumina library preparations were
performed as described by Quail *et al.* 2008 (46) and the whole genomes
sequenced on the HiSeq2000 according to the manufacturer's instructions
(Illumina, San Diego, CA, USA) at the Wellcome Trust Sanger Institute,
Cambridge, UK. Sequencing generated 125bp paired end reads, which were
assessed for quality using Sickle (47) after the removal of adapter sequences.

Sequencing reads that passed the quality threshold were put forward for *de-*
novo assembly generation using Spades v.3.10.1 (48) utilising a variety of
parameters conditions. Our 678 isolates were combined with 401 additional
genomes given in Table S4 to increase robustness of *S. suis* population
structure estimation, although these isolates were unavailable for MIC testing.
Draft genomes were annotated using Prokka v1.12-beta (v2.8.2). (49) and
bacterial species assignment was performed by a combination of MLST
assignment and FastQ Screen v.0.11.1 using a custom database (50). We
mapped the Illumina reads back to the *de-novo* assembly to investigate
polymorphic reads in the samples (indicative of mixed cultures) using BWA

v.0.7.16a (51). Genomes that exhibited poor sequencing quality (i.e poor assembly as indicated by a large number of contigs, low N50 values or a high number of polymorphic reads) or that which were inconsistent with an *S. suis* species assignment were excluded from the analysis.

587

In order to group our *S. suis* isolates in to different genetic clusters, we used the program hierBAPs in R (52). First, we inferred core genes from our isolates using Roary (v2.8.2) (53), aligned them using MACSE (54), and stripped regions that could not be aligned unambiguously due to high divergence, indels or missing data. This conserved region of the core genome was used as input in hierBAPS. We also constructed a neighbour joining tree of the core genome by importing the fasta alignment in to R v.3.4.1 (55). We computed a distance matrix using the function *dist.dna* with Kimura's 2-parameter distance model of sequence evolution ("K80") (56) and created a neighbour-joining tree using the function *nj*, both from the R package *ape* (57, 58). We midpoint rooted the tree using the *midpoint* function in the R package *Phangorn* (59).

600

601 **Known AMR allele detection**

ARIBA (60) identifies AMR determinants (or any sequence of interest) directly from paired sequencing reads using a public or custom reference database and relies on mapping the reads to reference sequence clusters followed by the formation of local assemblies of the mapped reads. The tool is further able to confirm the intactness of resistance genes as well to identify known or novel SNPs within a gene of interest.

608

609 The AMR alleles in *S. suis* were identified using ARIBA v2.10.0 based on the
 610 public AMR database, CARD (24), which was further supplemented with a
 611 custom database (given in the Supplement). The custom database contained
 612 gene sequences not present within the CARD database at the time of testing.
 613 These were previously published *S. suis* AMR genes, single nucleotide
 614 polymorphisms (SNPs) in genes known to confer antibiotic resistance and
 615 AMR genes found in other bacteria. In addition, we investigated whether there
 616 were additional novel resistance alleles in known resistance genes (described
 617 below). The sequence identity threshold against a reference was set at 90%.
 618 Only paired-end sequence reads that passed the quality control thresholds
 619 were used as input for ARIBA. To identify genes falling under the 90%
 620 identity threshold, we also performed blast searches of the draft assemblies
 621 against the non-redundant NCBI protein or nucleotide databases to identify
 622 variants or chimeric alleles of known resistance determinants that might not
 623 be present in the CARD database. Examples of allelic variants identified this
 624 way include *tet*, *aad*, *ant* and *cat* (Table 2).

625

626 Novel AMR candidate allele detection

627 We identified novel resistance alleles by a range of methods. Candidate
 628 alleles that might encode novel resistance mechanisms were identified by
 629 scanning literature describing experimental studies and others found by
 630 Genome Wide Association Studies (GWAS) in related bacterial species, for
 631 example, alleles of *pbps*, *mraY*, *dhfr* and *folP* (33, 61-63) along with their
 632 promoter regions. To avoid over-fitting our generalised linear models, gene or

633 promoters of interest in a “training” subset of the collection ($n=205$) were then
 634 extracted, aligned and ranked from the highest to lowest MIC values, using
 635 MUSCLE (64) in SEAVIEW (65). Manual sequence analysis was then
 636 performed to identify either amino acid or nucleotide variations that associated
 637 with high MIC.

638

639 Kinetic and structural studies have previously established that beta-lactam
 640 resistance is conferred by substitutions within PBPs in *Streptococcus*
 641 *pneumoniae* (33, 66) and other streptococci (34, 35, 67). While mutations are
 642 present throughout the entire PBPs, we noted statistically significant
 643 mutations (Table S2) in altered PBP2B and PBP2X expressing alleles in
 644 strains exhibiting high penicillin ($\geq 1\text{mg/L}$) and ceftiofur MICs ($\geq 2\text{mg/L}$),
 645 respectively. In PBP2B, the residue variations were present in loop regions,
 646 K479T/A, D512E/Q/K/A, K513E/D and T515S, within the *S. suis*
 647 transpeptidase domain (residues 351-681). In PBP2X, the mutations were
 648 located at positions M437L, S445T, T467S and Y525F, also in loop regions
 649 within its catalytic domain (residues 265–619). As multiple amino acid variants
 650 were identified within the *pbp2X* and *pbp2B*, many of which were in the same
 651 genomic region (Table 2), we grouped the variants in to haplotypes. We did
 652 this by scoring haplotype presence if the isolate had all of the amino acid
 653 variants associated with high MIC in a given gene.

654

655 In addition to allelic differences in PBP2B and PBP2X, we also detected
 656 mutations in *mraY*. Found immediately upstream of *pbp2X*, *mraY* encodes an
 657 acetylmuramoyl-pentapeptide-transferase enzyme which is essential to lipid

658 cycle reactions in the peptidoglycan cell wall biosynthesis pathway. Residue
659 substitutions were present at either A4S/T together with M6I/L or M6I/L
660 together with G8S within the signal peptide regions of the protein and were
661 associated with variation in both penicillin and ceftiofur MIC. These were
662 scored as two separate haplotypes in further analyses. The distribution of
663 these mutations was statistically significant at ≥ 1 mg/L for both ceftiofur and
664 penicillin (Table S2).

665

666 Not all isolates with high tiamulin MICs possessed a *IsaE* gene, suggesting an
667 additional mechanism in *S. suis* conferring resistance to pleuromutilins. Using
668 the CARD RGI (resistance gene identifier) protein homology models, a *vgaC*
669 homologue was identified as a candidate gene associated with variation in
670 tiamulin MIC. The gene, designated as *vgaF* in this work, encodes an ABC-F
671 ATP-binding cassette ribosomal protection protein and shares 37.14%
672 homology to the reference. Located within the chromosome, *vgaF* is found
673 intact (~1386bp) in statistically significant numbers in isolates with MIC $\geq$
674 8mg/L (Table S2) but is truncated (<750bp) in others. Plasmid borne *vga*
675 homologues including *vgaC* are frequently detected in staphylococci and have
676 been shown to confer cross resistance to pleuromutilins, lincosamides and
677 streptogramin A antibiotics. A chromosome based *vga(A)* gene variant,
678 encoding an ATP-binding cassette protein conferring resistance to
679 streptogramin A and related antibiotics in *S. aureus* has also been described
680 (68).

681

682 An amino acid substitution (I102-L) was detected in the dihydrofolate
 683 reductase (DHFR) gene, *dhfr*, in the majority of isolates with trimethoprim
 684 MICs of 0.12mg/L or greater. Similar substitutions of isoleucine to leucine at
 685 position 100 in *S. pneumoniae* (62, 69) and *Streptococcus pyogenes* (70) is
 686 known to cause resistance to TMP. Additionally, we also identified
 687 polymorphisms within the promoter region (0-30bp upstream) of the *dhfr*
 688 gene; an A5G substitution and insertions, in isolates exhibiting MICs $\geq$ 1mg/L.

689

690 All of these alleles might be causal, compensatory or linked to high MIC
 691 because of population structure. We therefore examined our alleles and
 692 showed that susceptible and resistance alleles were found repeatedly in
 693 different BAPs clusters, which were significant using binomial sign tests when
 694 defining a cutoff (Table S2).

695

696 The *pbp*, *mraY* and *vgaF* genes were unable to be included into the custom
 697 database in ARIBA because of the sequence divergence. Therefore, we
 698 manually aligned the genes in all 678 isolates using MUSCLE (64) in
 699 SEAVIEW (65).

700

701 Similarly, analysis for known variants in *gyrA* and *parC* were also performed
 702 by manually aligning the alleles. *tet(O/W/32/O)* was scored by using blastn as
 703 ARIBA was unable to differentiate mosaic sequence patterns

704

705

706 **Serotype inference from whole genomes**

707 Serotypes were determined *in silico* using the Athey *et al.* (2016) (71)
 708 serotype database, which we implemented in ARIBA. Failed ARIBA runs and
 709 sequence non-matches were designated as not available (NA).

710

711 **Statistical analyses**

712 All general linear models were fit in R v. 3.3 (55) using the built-in function *lm*
 713 for models with solely fixed effects, or via Reduced Maximum Likelihood,
 714 using the function *lme* from the package nlme v. 3.1-141 (72), when BAPs
 715 cluster was included as a random effect. The response MIC values were log
 716 transformed. Our data for host and country is confounded because all of our
 717 human samples came from Vietnam (human *S. suis* has a higher prevalence
 718 in South East Asia (73)). Therefore, we coded “Country” as separate
 719 populations – Canada, UK, Vietnam-pig and Vietnam-human. We classified
 720 “serotypes” into either disease-associated or non-disease associated
 721 according to Wileman *et al.* (2019) (21).

722

723 For the Mantel tests using the data plotted in Figure 3 and Figure S15, the
 724 test statistic was the Pearson correlation coefficients of the off-diagonal matrix
 725 elements (excluding the diagonals for correlation matrices, and, with missing
 726 data, including all observations that were pairwise complete), and generated a
 727 null distribution from 10,000 random permutations of the row/column ordering
 728 for one of the matrices. Given the hypotheses, these tests were one-tailed.

729

730 **Inferring physical linkage of candidate alleles**

731 The coordinates of candidate alleles in *S. suis* draft genomes were obtained
 732 using a custom python script (given in the Supplement). This script identifies
 733 the allele locations by performing blastn of the alleles against local databases
 734 of each draft genome. For each allele pairing, the number of genomes that
 735 contained both alleles were computed and whether these alleles were found
 736 on the same contig. For the isolates where the alleles were found on the
 737 same contig, a distribution of base pair distances was computed. The median
 738 base pairwise distance between two alleles was used to generate the
 739 heatmap in Figure S15.

740

741 **Description of randomisation tests**

742 For the permutation tests shown in Table 3, and illustrated in Figure 4, we
 743 used two approaches. For permutation "over isolates", we randomly permuted
 744 the present/absent records for each allele, independently, over the strains.
 745 This preserves the overall frequency of each allele, and the mean number of
 746 alleles per strain, but effectively assumes free recombination among all sites.
 747 We then assigned each of the candidate alleles, at random, to one of the 679
 748 branches of our core genome phylogeny (described above), choosing each
 749 branch with a probability proportional to its molecular branch length (thereby
 750 accounting for the fact that the substitutions are more likely to occur along
 751 lineages that have undergone more molecular evolution). The resistance
 752 allele was then treated as present in all strains that descend from the
 753 randomly chose branch. For the combined approach (labelled "both"),
 754 permutation over isolates was applied to candidate alleles found on Mobile
 755 Genetic Elements (Table 2), and permutation over phylogeny applied to the

756 remaining alleles. To assign p -values, we report the proportion of the
 757 randomly permuted data sets whose value of the test statistic was more
 758 extreme (i.e., closer to the tails of the distribution) than the value for the true
 759 data, doubled for a two-tailed test.

760

761 **Data availability**

762 Whole genome sequence assemblies are available from NCBI under Project ID
 763 PRJNA628943. The gene *cat* from isolate FX419 has accession MT367165. The
 764 gene *aadE1* from isolate CF2D3-1A has accession MT383663. The gene *aadE2*
 765 from isolate BH3D7-4E has accession MT383664. The gene *ant1* from isolate
 766 CF2D3-1C has accession MT383665. The gene *tetM1* from isolate 1129711 has
 767 accession MT383666. The gene *tetM2* from isolate LSS85 has
 768 accession MT383667. The gene *tetM3* from isolate SS981 has
 769 accession MT383668. The gene *tetO1* from isolate D16-010262 has
 770 accession MT383669. The gene *vgaF* from isolate 1230091 has
 771 accession MT₄₃1628. All R and Python scripts for conducting statistical tests and
 772 figures are given in the supplement. The ARIBA custom database is given in the
 773 supplement.

774

775 **Acknowledgements**

776 This work was primarily supported by a Challenge grant from the Royal
777 Society (reference: CH160114) and an Isaac Newton Trust Research Grant
778 (reference: 17.24(u)). In addition, LAW and GGRM are supported by a Sir
779 Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal
780 Society (109385/Z/15/Z). We also acknowledge a BBSRC grant to AWT
781 (BBSRC grant BB/L003902/1), a LoLa grant to DJM and AWT
782 (BB/G019274/1) and a China MoST grant to RZ (2013DFG32360) for sample
783 collections. In addition, we wish to thank Abiyad Baig, Andrew Balmer and
784 Olivier Restif for helpful discussion on antimicrobial resistance in *S. suis*.

785

786

787 **Contributions**

788 LAW and NTH conceived the study. AWT, SMW, RZ, MG, NF, TMW, JP, DJM
789 contributed isolates. NFH, ELM, PLKY, HDP, NTH collected new data. NFH,
790 ELM, GGRM, JJW and LAW analysed data. JJW produced all figures. LAW,
791 NFH, ELM and JJW wrote the manuscript.

792

793

794

Table 1. Ecological and genomic predictors of MIC

Analysis	#strains	Fixed effects	Df	F	p
(a)	678	Antibiotic	15	1404.125	$<10^{-15}$
		genetic cluster	29	40.644	$<10^{-15}$
(b)	450	Antibiotic	15	972.3107	$<10^{-4}$
		Year	1	6.6078	0.0102
		Serotype	1	6.6261	0.0101
		Disease Status	2	2.9760	0.0511
		Country	3	92.5155	$<10^{-4}$
(c)	678	Antibiotic	15	1457.4043	$<10^{-4}$
		Country	3	139.5902	$<10^{-4}$
(d)	557	Antibiotic	15	1144.3846	$<10^{-4}$
		Disease Status	2	9.9619	$<10^{-4}$
(e)	542	Antibiotic	15	1131.6347	$<10^{-4}$
		Serotype	1	8.7216	0.0032
(f)	652	Antibiotic	15	1350.8781	$<10^{-4}$
		Year	1	0.1591	0.69

Table 2. AMR determinants of *Streptococcus suis*

AMR gene	Name in figures	Resistance conferred (Antibiotic Class)	No. of isolates	Reported in <i>S. suis</i>	Reported in other species (CARD and NCBI)	Reported genetic location in <i>S. suis</i> or other species
<i>ant(4')-Ib</i>	ant4Ib	Aminoglycoside	1	This report	<i>Staphylococcus aureus</i> , <i>Staphylococcus epidermidis</i> , <i>Enterococcus</i> spp.	Plasmid
<i>ant(6')-Ia/aadE</i>	ant6Ia	Aminoglycoside	56	(19)	<i>Staphylococcus epidermidis</i> , <i>Streptococcus</i> spp., <i>Enterococcus faecalis</i>	Plasmid, ICE
<i>ant(6')-Ib</i>	ant6Ib	Aminoglycoside	12	This report	<i>Campylobacter</i> spp., <i>Bacillus subtilis</i> , <i>Clostridium</i> spp.	Pathogenicity Island
<i>ant(9')-Ia/spw/aad9</i>	ant9Ia	Aminoglycoside	45	(19)	<i>Staphylococcus aureus</i> , <i>Enterococcus</i> spp., <i>Staphylococcus</i> spp.	Plasmid, ICE
<i>aph(3')-IIIa</i>	aph3IIIa	Aminoglycoside	46	(74)	<i>Campylobacter</i> spp., <i>Staphylococcus</i> spp., <i>Streptococcus</i> spp., <i>Enterococcus</i> spp.	Plasmid, transposon
<i>ant1*</i>	ant1	Aminoglycoside	12	This report	<i>Campylobacter</i> spp., <i>Clostridium difficile</i> , <i>Lactobacillus amylovorus</i>	Plasmid, ICE
<i>aadE1*</i>	aadE1	Aminoglycoside	2	This report	<i>Enterococcus faecium</i> , <i>Staphylococcus aureus</i>	Plasmid
<i>aadE2*</i>	aadE2	Aminoglycoside	5	This report	<i>Clostridium</i> spp., <i>Campylobacter fetus</i>	ICE
<i>ermB</i>	ermB	MLS _B (Macrolide)	370	(75)	<i>Campylobacter</i> spp., <i>Clostridium</i> spp., <i>Enterococcus</i> spp., <i>Escherichia coli</i> , <i>Klebsiella pneumoniae</i> , <i>Proteus mirabilis</i> , <i>Shigella</i> spp., <i>Staphylococcus</i> spp., <i>Streptococcus</i> spp., <i>Vibrio vulnificus</i> .	Plasmid, ICE
<i>ermG</i>	ermG	MLS _B (Macrolide)	1	This report	<i>Clostridium difficile</i> , <i>Enterococcus faecium</i>	Conjugative transposon
<i>ermT</i>	ermT	MLS _B (Macrolide)	3	This report	<i>Enterococcus faecium</i> , <i>Staphylococcus aureus</i> , <i>Streptococcus</i> spp.	Plasmid
<i>mefA</i>	mefA	MLS _B (Macrolide)	11	(75)	<i>Streptococcus</i> spp., <i>Enterococcus faecium</i> , <i>Haemophilus parainfluenzae</i> , <i>Proteus mirabilis</i>	Transposon, ICE
<i>mel/msrD</i>	mel	MLS _B (Macrolide)	9	(75)	<i>Streptococcus</i> spp., <i>Enterococcus faecium</i> , <i>Haemophilus parainfluenzae</i> , <i>Proteus mirabilis</i> .	Transposon, ICE
<i>lnuB</i>	lnuB	MLS _B (Lincosamide)	3	(17)	<i>Enterococcus</i> spp., <i>Staphylococcus</i> spp., <i>Streptococcus agalactiae</i>	Plasmid
<i>linB</i>	linB	MLS _B (Lincosamide)	42	(76)	<i>Streptococcus suis</i>	
<i>lnuC</i>	lnuC	MLS _B (Lincosamide)	9	(77)	<i>Campylobacter</i> spp., <i>Clostridium difficile</i> , <i>Streptococcus agalactiae</i> .	Plasmid
<i>lsaE</i>	lsaE	Pleuromutilin	43	(17)	<i>Enterococcus</i> spp., <i>Staphylococcus</i> spp., <i>Streptococcus agalactiae</i> .	Plasmid, ICE
<i>tet44</i>	tet44	Tetracycline	4	This report	<i>Clostridium</i> spp.	ICE
<i>tetM</i>	tetM	Tetracycline	81	(19)	<i>Streptococcus</i> spp., <i>Escherichia coli</i> , <i>Staphylococcus</i> spp., <i>Enterococcus</i> spp., <i>Klebsiella pneumoniae</i> , <i>Neisseria meningitidis</i> , <i>Clostridium difficile</i> , <i>Salmonella enterica</i>	Plasmid, ICE
<i>tetM1*</i>	tetM1	Tetracycline	34	This report	<i>Staphylococcus</i> spp., <i>Streptococcus</i> spp., <i>Enterococcus</i> spp., <i>Bacillus</i> spp., <i>Clostridium</i> spp.	Plasmid, transposon
<i>tetM2*</i>	tetM2	Tetracycline	43	This report	<i>Enterococcus</i> spp., <i>Clostridium difficile</i> .	
<i>tetM3*</i>	tetM3	Tetracycline	13	This report	<i>Streptococcus pneumoniae</i> , <i>Streptococcus pyogenes</i> , <i>Enterococcus</i> spp., <i>Clostridium</i> spp.	Transposon, ICE
<i>tetO</i>	tetO	Tetracycline	465	(19)	<i>Streptococcus suis</i> , <i>Campylobacter</i> spp.	ICE, plasmid
<i>tetO1*</i>	tetO1	Tetracycline	2	This report	<i>Dorea</i> spp., uncultured bacterium	Conjugative transposon
<i>tetW</i>	tetW	Tetracycline	8	(16)	<i>Streptococcus agalactiae</i> , <i>Clostridium difficile</i> , <i>Chlamydia trachomatis</i> .	Plasmid

<i>tet40</i>	tet40	Tetracycline	7	(17, 26)	<i>Clostridium difficile</i>	Composite mobile genetic element, ICE
<i>tetL</i>	tetL	Tetracycline	12	(19)	<i>Streptococcus</i> spp., <i>Escherichia coli</i> , <i>Staphylococcus</i> spp., <i>Enterococcus</i> spp., <i>Klebsiella pneumoniae</i> .	Plasmid
<i>tet(O/W/32/O) †</i>	tetO:W:32:O	Tetracycline	20	(73)	<i>Streptococcus suis</i>	ICE
<i>dfrF</i>	dfrF	TMP	11	This report	<i>Clostridium difficile</i> , <i>Enterococcus</i> spp., <i>Staphylococcus aureus</i> , <i>Streptococcus agalactiae</i>	plasmid
<i>dfrK</i>	dfrK	TMP	7	This report	<i>Staphylococcus aureus</i> , <i>Staphylococcus epidermidis</i> .	plasmid
<i>vgaF*</i>	vgaF	Pleuromutilin	173	This report	Not known	Chromosome. Found as an intact gene (~1386bp) or truncated (<750bp)
<i>cat*</i>	cat	Florfenicol	3	This report	<i>Enterococcus faecalis</i> , <i>Streptococcus iniae</i> , <i>Listeria monocytogenes</i> , <i>Staphylococcus aureus</i>	Plasmid, Transposon
<i>dhfr*</i>	DHFR102	TMP	235	This report	Mutations in I100L in <i>Streptococcus</i> spp. and <i>Staphylococcus aureus</i>	<i>dhfr</i> : I102L
<i>dhfr promoter*</i>	DHFRPromoter	TMP	86	This report	None reported	<i>dhfr</i> : A5G upstream/indels 0-30bp upstream
<i>pbp2B/penA*</i>	PBP2B:hap	Penicillin	181	This report	Mutations in other locations in <i>Streptococcus pneumonia</i> and Group B <i>Streptococcus</i> spp	<i>pbp2B</i> : K479T/A, D512E/Q/K/A, K513E/D, T515S
<i>pbp2X*</i>	PBP2X:hap	Ceftiofur	93	This report	None reported	<i>pbp2X</i> : M437L, S445T, T467S Y525F
<i>pbp2X</i>	PBP2X594:596	Penicillin	15	(25)	<i>Streptococcus suis</i> . Mutations in other locations in <i>Streptococcus pneumonia</i> , Group B <i>Streptococcus</i> sp. and <i>Streptococcus uberis</i>	<i>pbp2X</i> : L594Y/F, V596G
<i>pbp2X</i>	PBP2X569	Penicillin	10	(25)	<i>Streptococcus suis</i>	<i>pbp2X</i> : N569Q
<i>mraY*</i>	MraY:hap	Penicillin	62	This report	None reported	<i>mraY</i> : A4S/T, M6I/L G8S
<i>parC†</i>	ParC79	Fluoroquinolone	4	(27)	<i>Escherichia coli</i> , <i>Pseudomonas aeruginosa</i> , <i>Streptococcus</i> spp., <i>Staphylococcus aureus</i>	<i>parC</i> : S79
<i>gyrA†</i>	GyrA81	Fluoroquinolone	13	(27)	<i>Escherichia coli</i> , <i>Klebsiella</i> spp., <i>Streptococcus</i> spp., <i>Staphylococcus aureus</i>	<i>gyrA</i> : S81
<i>gyrA†</i>	GyrA85	Fluoroquinolone	1	(27)	<i>Escherichia coli</i> , <i>Klebsiella</i> spp., <i>Streptococcus</i> spp., <i>Staphylococcus aureus</i>	<i>gyrA</i> : E85

*n novel determinants reported in this study
† Not documented in the CARD database
MLS_B Macrolide Linocosome Streptogramin B
TMP Trimethoprim

836
837
838
839
840
841
842

statistic	value	over isolates		over genealogy		both	
		null: mean (range)	<i>p</i>	null: mean (range)	<i>p</i>	null: mean (range)	<i>p</i>
proportion susceptible	0.088	0.014 (0.001-0.032)	<10 ⁻⁴	0.621 (0.001-0.855)	0.033	0.054 (0.001-0.086)	<10 ⁻⁴
proportion MDR	0.748	0.881 (0.842-0.928)	<10 ⁻⁴	0.109 (0.006-0.900)	8×10 ⁻⁴	0.667 (0.609-0.928)	0.020
proportion nested	0.631	0.272 (0.226-0.326)	<10 ⁻⁴	0.753 (0.487-0.923)	0.065	0.708 (0.655-0.860)	<10 ⁻⁴

Table 3: Evidence of clustering of AMR in *S. suis*

843
844
845
846
847
848
849
850
851

852 **Figure legends**

853

854 **Figure 1. Candidate AMR alleles explain most of the variation in MIC. a)**

855 Histograms of log transformed MIC measures for each of our 16 different
 856 antibiotics, across our panel of 678 *S. suis* isolates. Antibiotics are coloured
 857 by their class (Beta: beta-lactams; MLS_B: macrolide-lincosamide-
 858 streptogramin B; Tetra: tetracyclines; Fluoro: fluoroquinolones; Amino:
 859 aminoglycoside; Pleuro: pleuromutilin; TMP: trimethoprim and Phen:
 860 phenicol). **b)** Histograms showing the predictive power of our candidate AMR
 861 alleles. The left-hand histograms (labelled N) show the MIC values for isolates
 862 that carry no allele for that antibiotic class, while the right-hand histograms
 863 (labelled A) show the MIC values for isolates carrying one or more such allele.
 864 If all resistance alleles perfectly explain MIC, then we expect to see histogram
 865 distributions on the bottom left (low MIC, no allele) and the top right (high MIC,
 866 presence of candidate allele(s)). For the first antimicrobial in each class, we
 867 show the number of candidate AMR alleles in square brackets, along with the
 868 number of isolates where candidate alleles are absent or present. r^2 values
 869 show the proportion of the variance explained by the binary
 870 presence/absence of these alleles.

871

872 **Figure 2. Genotype-specific MIC values for lincomycin and erythromycin**

873 **show epistatic interactions.** The middle panels show the MIC values for
 874 lincomycin and erythromycin (both in the MLS_B class), for a range of
 875 genotypes with mean MIC values shown as black empty circles. Genotypes
 876 are classified in the lower panel, according to the subset of the 8 candidate

alleles for the MLS_B class that they contain, and the number of isolates with each genotype are shown in the upper panel. The plots illustrate several complexities of the genotype-phenotype map, including alleles that confer resistance to only one of the two drugs, and epistatic interactions, i.e., resistance effects that vary with the genetic background.

Figure 3. Covariation in MIC values for different antibiotic classes. The matrix represents correlation coefficients between log(MIC) values for pairs of antibiotics. The greyscale runs from white (-0.14) to black (1). Correlations are strongest for most antibiotics in the same class (as indicated by colour), but there are also strong correlations between classes. All values are positive, except for some values involving enrofloxacin (enr), which does not even correlate strongly with the MIC for the other Fluoroquinolone, marbofloxacin (mar). The inset figure plots the correlation coefficients on the y-axis (with values involving enrofloxacin shown in yellow). The x-axis shows the co-occurrence of candidate alleles conferring resistance to each pair of antibiotics. These were calculated as

$\rho_{\text{allele pres/abs}} = (p_{ij} - p_i p_j) / \sqrt{p_i(1 - p_i)p_j(1 - p_j)}$, where p_i and p_j are the proportions of the isolates carrying candidate resistance alleles against each of the classes, and p_{ij} is the proportion of isolates carrying alleles against both classes. The trend suggests that much of the covariance in MIC values is due to co-occurring resistance alleles.

Figure 4 Clustering of AMR alleles in *S. suis*. Histograms of the number of candidate resistance alleles carried by each isolate (upper panel), or the

902 number of antibiotic classes to which these alleles confer resistance (lower
903 panel). The true data (thicker blue bars), are compared to two comparison
904 data sets, in which the presence/absence records of each allele was
905 randomized over isolates (thinner red bars), or where the appearance of each
906 allele was placed at random on a consensus genealogy of the isolates
907 (thinner yellow bars). These randomization procedures suggest that the co-
908 occurrence of candidate resistance alleles, and especially alleles conferring
909 resistance to different drug classes, is more clustered than would be expected
910 by chance.

911

912

- 913 1. de Moor CE. 1963. Septicaemic infections in pigs, caused by haemolytic
914 streptococci of new Lancefield groups designated R, S and T. *Antonie van*
915 *Leeuwenhoek* 29:272-280.
- 916 2. Wertheim HFL, Nguyen HN, Taylor W, Lien TTM, Ngo HT, Nguyen TQ,
917 Nguyen BNT, Nguyen HH, Nguyen HM, Nguyen CT, Dao TT, Nguyen TV,
918 Fox A, Farrar J, Schultz C, Nguyen HD, Nguyen KV, Horby P. 2009.
919 *Streptococcus suis*, an Important Cause of Adult Bacterial Meningitis in
920 Northern Vietnam. *PLOS ONE* 4:e5973.
- 921 3. Baums CG, Kock C, Beineke A, Bennecke K, Goethe R, Schröder C,
922 Waldmann K-H, Valentin-Weigand P. 2009. *Streptococcus suis* bacterin
923 and subunit vaccine immunogenicities and protective efficacies against
924 serotypes 2 and 9. *Clinical and vaccine immunology* 16:200-208.
- 925 4. van Rennings L, von Münchhausen C, Ottilie H, Hartmann M, Merle R,
926 Honscha W, Käsbohrer A, Kreienbrock L. 2015. Cross-sectional study on
927 antibiotic usage in pigs in Germany. *PloS one* 10:e0119114.
- 928 5. McGlone JJ. 2013. The Future of Pork Production in the World: Towards
929 Sustainable, Welfare-Positive Systems. *Animals* 3:401-415.
- 930 6. Steinfeld H, Gerber PJ, Wassenaar T, Castel V, Rosales M, De haan C. 2006.
931 *Livestock's Long Shadow: Environmental Issues and Options*, vol 24.
- 932 7. Goyette-Desjardins G, Auger J-P, Xu J, Segura M, Gottschalk M. 2014.
933 *Streptococcus suis*, an important pig pathogen and emerging zoonotic
934 agent—an update on the worldwide distribution based on serotyping and
935 sequence typing. *Emerging Microbes and Infections* 3:1-20.
- 936 8. Van Boeckel TP, Brower C, Gilbert M, Grenfell BT, Levin SA, Robinson TP,
937 Teillant A, Laxminarayan R. 2015. Global trends in antimicrobial use in
938 food animals. *Proceedings of the National Academy of Sciences* 112:5649-
939 5654.

- 940 9. Hernandez-Garcia J, Wang J, Restif O, Holmes MA, Mather AE, Weinert LA,
941 Wileman TM, Thomson JR, Langford PR, Wren BW, Rycroft A, Maskell DJ,
942 Tucker AW, Consortium BT. 2017. Patterns of antimicrobial resistance in
943 *Streptococcus suis* isolates from pigs with or without streptococcal disease
944 in England between 2009 and 2014. *Veterinary microbiology* 207:117-
945 124.
- 946 10. Yongkiettrakul S, Maneerat K, Arechanajan B, Malila Y, Srimanote P,
947 Gottschalk M, Visessanguan W. 2019. Antimicrobial susceptibility of
948 *Streptococcus suis* isolated from diseased pigs, asymptomatic pigs, and
949 human patients in Thailand. *BMC Veterinary Research* 15:5.
- 950 11. Marie J, Morvan H, Berthelot-Hérault F, Sanders P, Kempf I, Gautier-
951 Bouchardon AV, Jouy E, Kobisch M. 2002. Antimicrobial susceptibility of
952 *Streptococcus suis* isolated from swine in France and from humans in
953 different countries between 1996 and 2000. *Journal of Antimicrobial*
954 *Chemotherapy* 50:201-209.
- 955 12. Šeol B, Kelnerić Ž, Hajsig D, Madić J, Naglić T. 1996. Susceptibility to
956 Antimicrobial Agents of *Streptococcus suis* Capsular Type 2 Strains
957 Isolated from Pigs. *Zentralblatt für Bakteriologie* 283:328-331.
- 958 13. Turgeon PL, Higgins R, Gottschalk M, Beaudoin M. 1994. Antimicrobial
959 susceptibility of *Streptococcus suis* isolates. *British Veterinary Journal*
960 150:263-269.
- 961 14. Gurung M, Tamang MD, Moon DC, Kim S-R, Jeong J-H, Jang G-C, Jung S-C,
962 Park Y-H, Lim S-K. 2015. Molecular Basis of Resistance to Selected
963 Antimicrobial Agents in the Emerging Zoonotic Pathogen *Streptococcus*
964 *suis*. *Journal of Clinical Microbiology* 53:2332-2336.
- 965 15. Yao J, Shang K, Huang J, Ran W, Kashif J, Wang L. 2014. Overexpression of
966 an ABC transporter and mutations of GyrA, GyrB, and ParC in contributing
967 to high-level ciprofloxacin resistance in *Streptococcus suis* type 2.
968 *BioScience Trends* 8:84-92.
- 969 16. Hoa NT, Chieu TT, Nghia HD, Mai NT, Anh PH, Wolbers M, Baker S,
970 Campbell JI, Chau NV, Hien TT, Farrar J, Schultz C. 2011. The
971 antimicrobial resistance patterns and associated determinants in
972 *Streptococcus suis* isolated from humans in southern Vietnam, 1997-2008.
973 *BMC Infectious Diseases* 11:6.
- 974 17. Huang J, Ma J, Shang K, Hu X, Liang Y, Li D, Wu Z, Dai L, Chen L, Wang L.
975 2016. Evolution and Diversity of the Antimicrobial Resistance Associated
976 Mobilome in *Streptococcus suis*: A Probable Mobile Genetic Elements
977 Reservoir for Other Streptococci. *Frontiers in Cellular and Infection*
978 *Microbiology* 6.
- 979 18. Hu P, Yang M, Zhang A, Wu J, Chen B, Hua Y, Yu J, Chen H, Xiao J, Jin M.
980 2011. Comparative genomics study of multi-drug-resistance mechanisms
981 in the antibiotic-resistant *Streptococcus suis* R61 strain. *PloS one*
982 6:e24988.
- 983 19. Holden MT, Hauser H, Sanders M, Ngo TH, Cherevach I, Cronin A,
984 Goodhead I, Mungall K, Quail MA, Price C, Rabinowitsch E, Sharp S,
985 Croucher NJ, Chieu TB, Mai NT, Diep TS, Chinh NT, Kehoe M, Leigh JA,
986 Ward PN, Dowson CG, Whatmore AM, Chanter N, Iversen P, Gottschalk M,
987 Slater JD, Smith HE, Spratt BG, Xu J, Ye C, Bentley S, Barrell BG, Schultz C,
988 Maskell DJ, Parkhill J. 2009. Rapid evolution of virulence and drug

- 1198 resistance in the emerging zoonotic pathogen *Streptococcus suis*. PLoS
- 1199 One 4:e6072.
- 1200 20. Weinert LA, Chaudhuri RR, Wang J, Peters SE, Corander J, Jombart T, Baig
- 1201 A, Howell KJ, Vehkala M, Välimäki N, Harris D, Chieu TTB, Van Vinh Chau
- 1202 N, Campbell J, Schultsz C, Parkhill J, Bentley SD, Langford PR, Rycroft AN,
- 1203 Wren BW, Farrar J, Baker S, Hoa NT, Holden MTG, Tucker AW, Maskell DJ,
- 1204 Bossé JT, Li Y, Maglennon GA, Matthews D, Cuccui J, Terra V, Consortium
- 1205 BRT. 2015. Genomic signatures of human and animal disease in the
- 1206 zoonotic pathogen *Streptococcus suis*. Nature Communications 6:6740.
- 1207 21. Wileman TM, Weinert LA, Howell KJ, Wang J, Peters SE, Williamson SM,
- 1208 Wells JM, Langford PR, Rycroft AN, Wren BW, Maskell DJ, Tucker AW.
- 1209 2019. Pathotyping the Zoonotic Pathogen *Streptococcus suis*: Novel
- 1210 Genetic Markers To Differentiate Invasive Disease-Associated Isolates
- 1211 from Non-Disease-Associated Isolates from England and Wales. Journal of
- 1212 Clinical Microbiology 57:e01712-18.
- 1213 22. Zou G, Zhou J, Xiao R, Zhang L, Cheng Y, Jin H, Li L, Zhang L, Wu B, Qian P,
- 1214 Li S, Ren L, Wang J, Oshota O, Hernandez-Garcia J, Wileman TM, Bentley S,
- 1215 Weinert L, Maskell DJ, Tucker AW, Zhou R. 2018. Effects of Environmental
- 1216 and Management-Associated Factors on Prevalence and Diversity of
- 1217 *Streptococcus suis* in Clinically Healthy Pig Herds in China and the United
- 1218 Kingdom. Applied and Environmental Microbiology 84:e02590-17.
- 1219 23. van Hout J, Heuvelink A, Gonggrijp M. 2016. Monitoring of antimicrobial
- 1220 susceptibility of *Streptococcus suis* in the Netherlands, 2013–2015.
- 1221 Veterinary Microbiology 194:5-10.
- 1222 24. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, Lago BA,
- 1223 Dave BM, Pereira S, Sharma AN, Doshi S, Courtot M, Lo R, Williams LE,
- 1224 Frye JG, Elsayegh T, Sardar D, Westman EL, Pawlowski AC, Johnson TA,
- 1225 Brinkman FSL, Wright GD, McArthur AG. 2017. CARD 2017: expansion
- 1226 and model-centric curation of the comprehensive antibiotic resistance
- 1227 database. Nucleic acids research 45:D566-D573.
- 1228 25. Ge Y, Wu J, Xia Y, Yang M, Xiao J, Yu J. 2012. Molecular dynamics
- 1229 simulation of the complex PBP-2x with drug cefuroxime to explore the
- 1230 drug resistance mechanism of *Streptococcus suis* R61. PLoS One 7:e35941.
- 1231 26. Palmieri C, Magi G, Mingoia M, Bagnarelli P, Ripa S, Varaldo PE, Facinelli B.
- 1232 2012. Characterization of a *Streptococcus suis* tet(O/W/32/O)-carrying
- 1233 element transferable to major streptococcal pathogens. Antimicrobial
- 1234 Agents and Chemotherapy 56:4697-702.
- 1235 27. Escudero JA, San Millan A, Catalan A, de la Campa AG, Rivero E, Lopez G,
- 1236 Dominguez L, Moreno MA, Gonzalez-Zorn B. 2007. First characterization
- 1237 of fluoroquinolone resistance in *Streptococcus suis*. Antimicrobial Agents
- 1238 and Chemotherapy 51:777-82.
- 1239 28. Albarracin Orio AG, Pinas GE, Cortes PR, Cian MB, Echenique J. 2011.
- 1240 Compensatory evolution of pbp mutations restores the fitness cost
- 1241 imposed by beta-lactam resistance in *Streptococcus pneumoniae*. PLoS
- 1242 Pathogens 7:e1002000.
- 1243 29. Leclercq R. 2002. Mechanisms of Resistance to Macrolides and
- 1244 Lincosamides: Nature of the Resistance Elements and Their Clinical
- 1245 Implications. Clinical Infectious Diseases 34:482-492.

- 1037 30. Lehtinen S, Blanquart F, Lipsitch M, Fraser C, with the Maela
1038 Pneumococcal C. 2019. On the evolutionary ecology of multidrug
1039 resistance in bacteria. *PLoS Pathogens* 15:e1007763.
- 1040 31. Calvez P, Breukink E, Roper DI, Dib M, Contreras-Martel C, Zapun A. 2017.
1041 Substitutions in PBP2b from beta-Lactam-resistant *Streptococcus*
1042 *pneumoniae* Have Different Effects on Enzymatic Activity and Drug
1043 Reactivity. *Journal of Biological Chemistry* 292:2854-2865.
- 1044 32. Coffey TJ, Daniels M, McDougal LK, Dowson CG, Tenover FC, Spratt BG.
1045 1995. Genetic analysis of clinical isolates of *Streptococcus pneumoniae*
1046 with high-level resistance to expanded-spectrum cephalosporins.
1047 *Antimicrobial Agents and Chemotherapy* 39:1306-13.
- 1048 33. Grebe T, Hakenbeck R. 1996. Penicillin-binding proteins 2b and 2x of
1049 *Streptococcus pneumoniae* are primary resistance determinants for
1050 different classes of beta-lactam antibiotics. *Antimicrobial Agents and*
1051 *Chemotherapy* 40:829-34.
- 1052 34. Haenni M, Galofaro L, Ythier M, Giddey M, Majcherczyk P, Moreillon P,
1053 Madec JY. 2010. Penicillin-binding protein gene alterations in
1054 *Streptococcus uberis* isolates presenting decreased susceptibility to
1055 penicillin. *Antimicrobial Agents and Chemotherapy* 54:1140-5.
- 1056 35. Piccinelli G, Carlentini G, Gargiulo F, Caruso A, De Francesco MA. 2017.
1057 Analysis of Point Mutations in the pbp2x, pbp2b, and pbp1a Genes of
1058 *Streptococcus agalactiae* and Their Relation with a Reduced Susceptibility
1059 to Cephalosporins. *Microbial Drug Resistance* 23:1019-1024.
- 1060 36. Kocaoglu O, Tsui HC, Winkler ME, Carlson EE. 2015. Profiling of beta-
1061 lactam selectivity for penicillin-binding proteins in *Streptococcus*
1062 *pneumoniae* D39. *Antimicrobial Agents and Chemotherapy* 59:3548-55.
- 1063 37. Figueiredo TA, Sobral RG, Ludovice AM, Almeida JM, Bui NK, Vollmer W,
1064 de Lencastre H, Tomasz A. 2012. Identification of genetic determinants
1065 and enzymes involved with the amidation of glutamic acid residues in the
1066 peptidoglycan of *Staphylococcus aureus*. *PLoS Pathogens* 8:e1002508.
- 1067 38. Zapun A, Philippe J, Abrahams KA, Signor L, Roper DI, Breukink E, Vernet
1068 T. 2013. In vitro reconstitution of peptidoglycan assembly from the Gram-
1069 positive pathogen *Streptococcus pneumoniae*. *ACS Chemical Biology*
1070 8:2688-96.
- 1071 39. Chewapreecha C, Marttinen P, Croucher NJ, Salter SJ, Harris SR, Mather
1072 AE, Hanage WP, Goldblatt D, Nosten FH, Turner C, Turner P, Bentley SD,
1073 Parkhill J. 2014. Comprehensive identification of single nucleotide
1074 polymorphisms associated with beta-lactam resistance within
1075 pneumococcal mosaic genes. *PLoS Genetics* 10:e1004547.
- 1076 40. Dunai A, Spohn R, Farkas Z, Lázár V, Györkei Á, Apjok G, Boross G,
1077 Szappanos B, Grézal G, Faragó A, Bodai L, Papp B, Pál C. 2019. Rapid
1078 decline of bacterial drug-resistance in an antibiotic-free environment
1079 through phenotypic reversion. *eLife* 8:e47088.
- 1080 41. Directorate VM. 2018. UK veterinary antibiotic resistance and sales
1081 surveillance report.
- 1082 42. Dunlop RH, McEwen SA, Meek AH, Black WD, Clarke RC, Friendship RM.
1083 1998. Individual and group antimicrobial usage rates on 34 farrow-to-
1084 finish swine farms in Ontario, Canada. *Preventive Veterinary Medicine*
1085 34:247-264.

- 1086 43. Carrique-Mas JJ, Choisy M, Van Cuong N, Thwaites G, Baker S. 2020. An
1087 estimation of total antimicrobial usage in humans and animals in Vietnam.
1088 *Antimicrobial Resistance and Infection Control* 9:16.
- 1089 44. Lehtinen S, Blanquart F, Croucher NJ, Turner P, Lipsitch M, Fraser C. 2017.
1090 Evolution of antibiotic resistance is linked to any genetic mechanism
1091 affecting bacterial duration of carriage. *Proceedings of the National*
1092 *Academy of Sciences* 114:1075-1080.
- 1093 45. de Jong A, Thomas V, Simjee S, Moyaert H, El Garch F, Maher K, Morrissey
1094 I, Butty P, Klein U, Marion H, Rigaut D, Vallé M. 2014. Antimicrobial
1095 susceptibility monitoring of respiratory tract pathogens isolated from
1096 diseased cattle and pigs across Europe: The VetPath study. *Veterinary*
1097 *Microbiology* 172:202-215.
- 1098 46. Quail MA, Kozarewa I, Smith F, Scally A, Stephens PJ, Durbin R, Swerdlow
1099 H, Turner DJ. 2008. A large genome center's improvements to the Illumina
1100 sequencing system. *Nature Methods* 5:1005-1010.
- 1101 47. Joshi N, Fass J. 2011. Sickle: A sliding-window, adaptive, quality-based
1102 trimming tool for FastQ files.
- 1103 48. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS,
1104 Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV,
1105 Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new
1106 genome assembly algorithm and its applications to single-cell sequencing.
1107 *Journal of Computational Biology* 19:455-477.
- 1108 49. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation.
1109 *Bioinformatics* 30:2068-2069.
- 1110 50. Wingett SW, Andrews S. 2018. FastQ Screen: A tool for multi-genome
1111 mapping and quality control. *F1000Research* 7:1338-1338.
- 1112 51. Li H, Durbin R. 2009. Fast and accurate short read alignment with
1113 Burrows-Wheeler transform. *Bioinformatics* 25:1754-1760.
- 1114 52. Cheng L, Connor TR, Sirén J, Aanensen DM, Corander J. 2013. Hierarchical
1115 and Spatially Explicit Clustering of DNA Sequences with BAPS Software.
1116 *Molecular Biology and Evolution* 30:1224-1228.
- 1117 53. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M,
1118 Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote
1119 pan genome analysis. *Bioinformatics* 31:3691-3693.
- 1120 54. Ranwez V, Harispe S, Delsuc F, Douzery EJP. 2011. MACSE: Multiple
1121 Alignment of Coding SEquences Accounting for Frameshifts and Stop
1122 Codons. *PLOS ONE* 6:e22594.
- 1123 55. Team RC. 2013. R: A language and environment for statistical computing.
- 1124 56. Kimura M. 1980. A simple method for estimating evolutionary rates of
1125 base substitutions through comparative studies of nucleotide sequences.
1126 *Journal of Molecular Evolution* 16:111-120.
- 1127 57. Paradis E. 2011. *Analysis of Phylogenetics and Evolution with R*. Springer
1128 Science & Business Media.
- 1129 58. Paradis E, Schliep K. 2018. ape 5.0: an environment for modern
1130 phylogenetics and evolutionary analyses in R. *Bioinformatics* 35:526-528.
- 1131 59. Schliep KP. 2011. phangorn: phylogenetic analysis in R. *Bioinformatics*
1132 27:592-593.

- 1133 60. Hunt M, Mather AE, Sánchez-Busó L, Page AJ, Parkhill J, Keane JA, Harris
1134 SR. 2017. ARIBA: rapid antimicrobial resistance genotyping directly from
1135 sequencing reads. *Microbial genomics* 3:e000131.
- 1136 61. Lopez P, Espinosa M, Greenberg B, Lacks St. 1987. Sulfonamide resistance
1137 in *Streptococcus pneumoniae*: DNA sequence of the gene encoding
1138 dihydropteroate synthase and characterization of the enzyme. *Journal of*
1139 *Bacteriology* 169:4320-4326.
- 1140 62. Adrian PV, Klugman KP. 1997. Mutations in the dihydrofolate reductase
1141 gene of trimethoprim-resistant isolates of *Streptococcus pneumoniae*.
1142 *Antimicrobial Agents and Chemotherapy* 41:2406-13.
- 1143 63. Hakenbeck R, Grebe T, Zähner D, Stock JB. 1999. β -Lactam resistance in
1144 *Streptococcus pneumoniae*: penicillin-binding proteins and non-penicillin-
1145 binding proteins. *Molecular Microbiology* 33:673-678.
- 1146 64. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high
1147 accuracy and high throughput. *Nucleic acids research* 32:1792-1797.
- 1148 65. Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: A multiplatform
1149 graphical user interface for sequence alignment and phylogenetic tree
1150 building. *Molecular Biology and Evolution* 27:221-4.
- 1151 66. Contreras-Martel C, Dahout-Gonzalez C, Martins Ados S, Kotnik M, Dessen
1152 A. 2009. PBP active site flexibility as the key mechanism for beta-lactam
1153 resistance in pneumococci. *Journal of Molecular Biology* 387:899-909.
- 1154 67. Fuursted K, Stegger M, Hoffmann S, Lambertsen L, Andersen PS, Deleuran
1155 M, Thomsen MK. 2016. Description and characterization of a penicillin-
1156 resistant *Streptococcus dysgalactiae* subsp. *equisimilis* clone isolated from
1157 blood in three epidemiologically linked patients. *Journal of Antimicrobial*
1158 *Chemotherapy* 71:3376-3380.
- 1159 68. Haroche J, Allignet J, Buchrieser C, El Solh N. 2000. Characterization of a
1160 Variant of vga (A) Conferring Resistance to Streptogramin A and Related
1161 Compounds. *Antimicrobial Agents and Chemotherapy* 44:2271.
- 1162 69. Pikis A, Donkersloot JA, Rodriguez WJ, Keith JM. 1998. A conservative
1163 amino acid mutation in the chromosome-encoded dihydrofolate
1164 reductase confers trimethoprim resistance in *Streptococcus pneumoniae*.
1165 *Journal of Infectious Diseases* 178:700-6.
- 1166 70. Bergmann R, van der Linden M, Chhatwal GS, Nitsche-Schmitz DP. 2014.
1167 Factors that cause trimethoprim resistance in *Streptococcus pyogenes*.
1168 *Antimicrobial Agents and Chemotherapy* 58:2281-8.
- 1169 71. Athey TBT, Teatero S, Lacouture S, Takamatsu D, Gottschalk M, Fittipaldi
1170 N. 2016. Determining *Streptococcus suis* serotype from short-read whole-
1171 genome sequencing data. *BMC microbiology* 16:162-162.
- 1172 72. Pinheiro J, Bates D, DebRoy S, Sarkar D, Heisterkamp S, Van Willigen B,
1173 Maintainer R. 2019. Package 'nlme'.
- 1174 73. Palmieri C, Varaldo PE, Facinelli B. 2011. *Streptococcus suis*, an Emerging
1175 Drug-Resistant Animal and Human Pathogen. *Frontiers in Microbiology*
1176 2:235.
- 1177 74. Huang K, Zhang Q, Song Y, Zhang Z, Zhang A, Xiao J, Jin M. 2016.
1178 Characterization of Spectinomycin Resistance in *Streptococcus suis* Leads
1179 to Two Novel Insights into Drug Resistance Formation and Dissemination
1180 Mechanism. *Antimicrobial Agents and Chemotherapy* 60:6390-2.

- 1181 75. Chang B, Wada A, Ikebe T, Ohnishi M, Mita K, Endo M, Matsuo H, Asatuma
1182 Y, Kuramoto S, Sekiguchi H, Yamazaki M, Yoshikawa H, Watabe N, Yamada
1183 H, Kurita S, Imai Y, Watanabe H. 2006. Characteristics of *Streptococcus*
1184 *suis* isolated from patients in Japan. Japanese Journal of Infectious
1185 Diseases 59:397-9.
- 1186 76. Bojarska A, Molska E, Janas K, Skoczynska A, Stefaniuk E, Hryniewicz W,
1187 Sadowy E. 2016. *Streptococcus suis* in invasive human infections in
1188 Poland: clonality and determinants of virulence and antimicrobial
1189 resistance. European Journal of Clinical Microbiology and Infectious
1190 Diseases 35:917-25.
- 1191 77. Athey TB, Auger JP, Teatero S, Dumesnil A, Takamatsu D, Wasserscheid J,
1192 Dewar K, Gottschalk M, Fittipaldi N. 2015. Complex Population Structure
1193 and Virulence Differences among Serotype 2 *Streptococcus suis* Strains
1194 Belonging to Sequence Type 28. PLoS One 10:e0137760.
1195