APPENDIX A

Mathematical representations of the three models discussed are given below.

Model 1
For each antimicrobial, $a \in \{1, 2, 3, 4, 5, 6\}$ (where $1 = \text{ampicillin}$, $2 = \text{chloramphenicol}$, $3 = \text{spectinomycin}$, $4 = \text{streptomycin}$, $5 = \text{sulphamethoxazole}$ and $6 = \text{tetracycline}$), tested for each isolate, $i$, the probability of observing a resistance to antimicrobial $a$ in isolate $i$, $p(R_{ai})$, is given by:

$$P(R_{ai}) = 1 - (1 - P(G_a))(1 - P(S_{ai}))$$

where $S_{ai}$ refers to observation of resistance to $a$ when it is encoded by SGI1, and $G_a$ refers to the observation of resistance to $a$ when it is encoded elsewhere on the genome. The latter event represents the combined probability of one or more such genes existing and the probability of phenotypic expression of these genes given that they exist.

The probability of observing a resistance encoded by SGI1, $p(S_{ai})$, is dependent on the probability of expression of a genetic element, $p(E_a)$, conditional on the presence of SGI1, $L$. For each isolate, $i$,

$$P(S_{ai}) = \begin{cases} P(E_a), & \text{if } L = 1 \\ 0, & \text{if } L = 0 \end{cases}$$

The probabilities $p(G_a)$, $p(E_a)$ and $p(L)$ are estimated using minimally informative Beta priors.

Model 2
The probability of observing resistance to antimicrobial $a$ in isolate $i$ is given as for Model 1:

$$P(R_{ai}) = 1 - (1 - P(G_a))(1 - P(S_{ai}))$$

The probability of observing a resistance encoded by SGI1 for each antimicrobial is now conditional on the type of SGI1 variant, $V \in \{1, 2, 3, 4, 5, 6\}$ (table 1), as well as depending on the probability of expression for that antimicrobial, $E_a$. For each isolate, $i$,

$$P(S_{ai}) = \begin{cases} P(E_1), & \text{if } V \in \{1, 3, 4, 5\} \\ 0, & \text{if } V \in \{2, 6\} \end{cases}$$

$$P(S_{2ai}) = \begin{cases} P(E_2), & \text{if } V \in \{4, 5\} \\ 0, & \text{if } V \in \{1, 2, 3, 6\} \end{cases}$$

$$P(S_{3ai}) = \begin{cases} P(E_3), & \text{if } V \in \{2, 3, 5\} \\ 0, & \text{if } V \in \{1, 4, 6\} \end{cases}$$

$$P(S_{4ai}) = \begin{cases} P(E_4), & \text{if } V \in \{2, 3, 5\} \\ 0, & \text{if } V \in \{1, 4, 6\} \end{cases}$$

$$P(S_{5ai}) = \begin{cases} P(E_5), & \text{if } V \in \{1, 2, 3, 4, 5\} \\ 0, & \text{if } V \in \{6\} \end{cases}$$

and

$$P(S_{6ai}) = \begin{cases} P(E_6), & \text{if } V \in \{1, 2, 3, 4, 5\} \\ 0, & \text{if } V \in \{6\} \end{cases}$$

The probabilities $p(G_a)$ and $p(E_a)$ are estimated as for model 1, and the multinomial probability vector $p(V = v)$ for $v = 1, 2, 3, 4, 5, 6$ estimated using a minimally informative Dirichlet prior. The probability corresponding to the full SGI1 variant ($V = 5$) is conceptually equivalent to the probability in model 1 in the presence of SGI1.

Model 3
The probability of observing resistance to antimicrobial $a$ in isolate $i$ is given as for model 1:

$$P(R_{ai}) = 1 - (1 - P(G_a))(1 - P(S_{ai}))$$

The probability of observing a resistance encoded by SGI1 is dependent on the probability of observing a resistance to each given antimicrobial originating from a gene within SGI1, $T_a$, conditional on the presence of SGI1, $L$. For each isolate, $i$,

$$P(S_{ai}) = \begin{cases} P(T_a), & \text{if } L = 1 \\ 0, & \text{if } L = 0 \end{cases}$$

The probability of observing a resistance originating from with SGI1 depends on the probability of expression of each antimicrobial resistance, conditional on the existence of a gene within SGI1 that confers resistance to that antimicrobial. Four genetic elements, their presence denoted by $W_1$, $W_2$, $W_3$, $W_4$, exist independently of each other, and provide resistance genes for antimicrobials 1&5 (A&Sx), 3&4 (Sp&St), 2 (C) and 6 (T), respectively. For each isolate, $i$,

$$P(T_{1i}) = \begin{cases} P(E_1), & \text{if } W_1 = 1 \\ 0, & \text{if } W_1 = 0 \end{cases}$$

$$P(T_{3i}) = \begin{cases} P(E_3), & \text{if } W_1 = 1 \\ 0, & \text{if } W_1 = 0 \end{cases}$$

$$P(T_{4i}) = \begin{cases} P(E_4), & \text{if } W_1 = 1 \\ 0, & \text{if } W_1 = 0 \end{cases}$$

$$P(T_{2i}) = \begin{cases} P(E_2), & \text{if } W_1 = 1 \\ 0, & \text{if } W_1 = 0 \end{cases}$$

and

$$P(T_{6i}) = \begin{cases} P(E_6), & \text{if } W_1 = 1 \\ 0, & \text{if } W_1 = 0 \end{cases}$$
The probabilities \(p(G_i)\), \(p(E_i)\), and \(p(L)\) are estimated as for Model 1, and the additional probabilities \(p(W_i)\) for \(i = 1, 2, 3, 4\) are also estimated using minimally informative Beta priors.

**APPENDIX B**

MCMC model code (in the JAGS/BUGS syntax) follows for the three models discussed. For each, ‘Resistance’ is a Bernoulli observation of the phenotypic resistance for each antimicrobial and isolate, ‘ind.prob’ is the probability of resistance arising from elsewhere on the genome (non-linked) for each antimicrobial, and ‘linked.prob’ is the probability of resistance arising from the SGI1 for each antimicrobial. For models 2 and 3, Link.table and Sublink.table represent the active genes in each SGI1 variant (given in table 1), or genetic element within SGI1, respectively.

**Model 1**

```plaintext
model{

for(i in 1:N.isolates){
      for(a in 1:N.mics){
        Resistance[i,a] ~ dbern(1-prob.absent[i,a])
        prob.absent[i,a] <- (1 - ind.prob[a]) * (1-linked.prob[i,a])
        linked.prob[i,a] <- link.status[i] * link.prob[a]
      }

      link.status[i] ~ dbern(mainlink.prob)
}

mainlink.prob ~ dbeta(1,1)
}

for(a in 1:N.mics){
  ind.prob[a] ~ dbeta(1,1)
  link.prob[a] ~ dbeta(1,1)
}

#data# N.isolates, N.mics, Resistance
#monitor# ind.prob, link.prob, mainlink.prob, linked.prob
}
```

**Model 2**

```plaintext
model{

for(i in 1:N.isolates){
      for(a in 1:N.mics){
        Resistance[i,a] ~ dbern(combined.prob[i,a])
        combined.prob[i,a] <- 1 - ((1 - ind.prob[a]) * (1 - (effective.prob[a, link.type[i]])))
      }

      link.type[i] ~ dcat(link.probs[])
}

link.probs[1:N.links] ~ ddirch(alpha[])
}

for(l in 1:N.links){
  alpha[l] <- 1
  for(a in 1:N.mics){
    effective.prob[a,l] <- Link.table[a,l] * linked.prob[a]
  }
}

for(a in 1:N.mics){
  ind.prob[a] ~ dbeta(1,1)
  linked.prob[a] ~ dbeta(1,1)
}

#data# N.isolates, N.mics, N.links, Resistance, Link.table
#monitor# ind.prob, linked.prob, link.probs,
```
Model 3

model{

for(i in 1:N.isolates){
  for(a in 1:N.mics){
    Resistance[i,a] ~ dbern(1-prob.absent[i,a])
    prob.absent[i,a] <- (1 - ind.prob[a]) * prod(prob.absent.sep[i,a,1:N.linkages])
    for(l in 1:N.linkages){
      prob.absent.sep[i,a,l] <- 1 - (effective.prob[a,l] * link.status[i,l] * mainlink.active[i])
    }
  }
  for(l in 1:N.linkages){
    link.status[i,l] ~ dbern(link.prob[l])
  }
  mainlink.active[i] ~ dbern(mainlink.prob)
}

mainlink.prob ~ dbeta(1,1)

for(l in 1:N.linkages){
  link.prob[l] ~ dbeta(1,1)
  for(a in 1:N.mics){
    effective.prob[a,l] <- Sublink.table[a,l] * linked.prob[a]
  }
}

for(a in 1:N.mics){
  ind.prob[a] ~ dbeta(1,1)
  linked.prob[a] ~ dbeta(1,1)
}

#data# N.isolates, N.mics, N.linkages, Resistance, Sublink.table
#monitor# ind.prob, linked.prob, mainlink.prob, link.prob
#inits# ind.prob, linked.prob, mainlink.prob, link.prob
}

An example dataset for use as ‘Resistance’ in the models given above is included with the online electronic supplementary material. The file is in comma separated value (csv) format, with 1000 isolates represented by rows, and resistance (1) or non-resistance (0) to six antimicrobials represented in columns. The data were generated under the assumptions of the preferred model 2, with similar prevalences for each resistance and SGI1 variant as reported for the human data in the manuscript. Note: this is an example dataset only. In order to run the BUGS code provided in WinBUGS or JAGS, additional variables relating to the number of isolates (1000 for this example dataset) and antimicrobials (six) will need to be provided, along with a representation of the possible SGI1 variants (provided in table 1 and appendix A) for model 2, and the sub-link relationships (provided in appendix A) for model 3. We would be happy to provide, on request, the additional code required to generate additional simulated datasets, via email to: m.denwood@vet.gla.ac.uk.