

Spatial patterns of antimicrobial resistance genes in a cross-sectional sample of pig farms with indoor non-organic production of finishers

A. C. BIRKEGÅRD¹*, A. K. ERSBØLL², T. HALASA¹, J. CLASEN¹,
A. FOLKESSON¹, H. VIGRE³ AND N. TOFT¹

¹ *Technical University of Denmark, Frederiksberg, Denmark*

² *University of Southern Denmark, Copenhagen, Denmark*

³ *Technical University of Denmark, Lyngby, Denmark*

*Received 30 May 2016; Final revision 4 January 2017; Accepted 16 January 2017;
first published online 20 February 2017*

SUMMARY

Antimicrobial resistance (AMR) in pig populations is a public health concern. There is a lack of information of spatial distributions of AMR genes in pig populations at large scales. The objective of the study was to describe the spatial pattern of AMR genes in faecal samples from pig farms and to test if the AMR genes were spatially randomly distributed with respect to the geographic distribution of the pig farm population at risk. Faecal samples from 687 Danish pig farms were collected in February and March 2015. DNA was extracted and the levels of seven AMR genes (*ermB*, *ermF*, *sulI*, *sulII*, *tet(M)*, *tet(O)* and *tet(W)*) were quantified on a high-throughput real-time PCR array. Spatial differences for the levels of the AMR genes measured as relative quantities were evaluated by spatial cluster analysis and creating of risk maps using kriging analysis and kernel density estimation. Significant spatial clusters were identified for *ermB*, *ermF*, *sulII* and *tet(W)*. The broad spatial trends in AMR resistance evident in the risk maps were in agreement with the results of the cluster analysis. However, they also showed that there were only small scale spatial differences in the gene levels. We conclude that the geographical location of a pig farm is not a major determinant of the presence or high levels of AMR genes assessed in this study.

Key words: Antimicrobial resistance genes, Denmark, pig farms, spatial patterns.

INTRODUCTION

After the Swann Report was published in 1969 [1], antimicrobial resistant (AMR) bacteria and the use of antibiotics in animals have been under scrutiny for their potentially negative effects on human health. Every year in Europe, more than 25 000 people die of

diseases caused by AMR bacteria [2]. Management of this problem would benefit from an epidemiological approach to identify both direct and indirect causes of human infections arising from AMR bacteria.

Bacteria harbouring AMR genes are present in porcine faeces [3, 4], and it is generally accepted that AMR bacteria can be transferred from animals to humans through meat consumption [5] and via direct contact with pigs [6, 7]. Spreading slurry on farmland for the purposes of crop fertilisation might be a third way of transferring AMR genes from pigs to humans as fertilisation with porcine manure can increase the

* Author for correspondence: A. C. Birkegård, Section for Diagnostics and Scientific Advice, National Veterinary Institute, Technical University of Denmark, Bülowsvej 27, 1870 Frederiksberg C, Denmark.
(Email: acbir@vet.dtu.dk)

AMR levels in soil [8, 9]. The relative importance of transmission through meat compared with other transmission routes vary from gene to gene depending on the bacterial host of the gene.

Some bacteria are intrinsically resistance to AMR. Therefore, the bacterial composition of the porcine gut will affect the levels of the AMR genes. Previous studies have identified spatial patterns in the distribution of the different bacterial pathogens of livestock [10–13]. Thus, this study was planned with the hypothesis that AMR genes show a non-random spatial distribution. This hypothesis is supported by previous studies that have found spatial patterns of phenotypic AMR in enteric pathogens [14] and indicator bacteria [15]. In this study, we report the spatial patterns of the endemic levels of seven selected AMR genes in faecal total community DNA from pig farms. To the best of our knowledge, this is the first study of its kind. The pigs from which the samples were obtained had no clinical signs of disease. Therefore, the levels of AMR genes are assumed to reflect the background level of AMR in the Danish pig population, potentially acting as a reservoir for AMR in humans. The present study was designed to assess whether the spatial distribution of seven selected AMR genes was random with respect to the geographic distribution of the pig farm population at risk.

Seven genes, *ermB*, *ermF*, *sulI*, *sulII*, *tet(M)*, *tet(O)* and *tet(W)* were included in this study because they have previously been identified as being present on Danish pig farms and a validated qPCR assay was available for testing for the presence of these genes [4]. The genes included in the assay comprise genes coding for two of the three most commonly used antimicrobial classes in Danish pig production, tetracyclines and macrolides [16]. The *ermF* and *ermB* genes code for resistance against macrolides whereas the *tet(M)*, *tet(O)* and *tet(W)* genes encode resistance against tetracycline. These genes were included because they are expected to be found at high levels in some farms and could be used for detecting potential differences between farms. This might not be the case if the differences were below the sensitivity of the qPCR. Furthermore, the assay included two genes that are relatively rare in finisher pigs, i.e. the *sulI* and *sulII*. Sulphonamides, the antimicrobial class that these two genes encode resistance against, are rarely used in finisher in Danish pig production [16]. Evaluation of the levels of AMR genes using spatial statistical and geostatistical methods can be useful in generation of hypotheses regarding how the genes

might spread through pig populations. Identification of spatial clustering of farms according to a specific AMR gene would provide a foundation for further analyses to explain the presence of these clusters, aiding our understanding of determinants of AMR genetics among Danish pig farms. This would help in the introduction of surveillance and monitoring systems, as well as preventive initiatives to limit the extent of AMR genes in pig farms. Furthermore, the findings of risk areas for specific AMR genes would indicate that the AMR genes are spread from farm to farm.

The objectives of the study were to describe the spatial patterns of AMR genes in faecal samples from pig farms and to test if the AMR genes were spatially randomly distributed.

METHODS

Study design and sampling

This was a cross-sectional study with sampling carried out from 2 February 2015 to 3 March 2015. The sampling period was restricted to these months to avoid seasonal changes in the level of AMR in pig farms [17].

Sampling took place at five of the seven largest Danish-owned slaughterhouses for finisher pigs in Denmark to ensure spatial randomness. Previous investigations showed that these slaughterhouses primarily received pigs from local farms [unpublished data]. The remaining two slaughterhouses were excluded because one primarily slaughtered pigs from free-range and organic farms, and the other was located on Bornholm. This remote island was excluded for all analysis, and therefore in this study ‘Denmark’ refers to ‘Denmark excluding Bornholm’.

The number of farms to sample at each slaughterhouse was weighted according to the average number of farms sending pigs to slaughter during two 5-week periods starting from February and November 2014. The data used to plan the sampling were meat inspection data. These data were obtained from the Danish Classification Inspection and include details of individual pigs slaughtered at each of the study slaughterhouses. A total of 15 31 600 finishers were slaughtered in Denmark in February 2015 [18]. Of these, 13 65 963 (89%) were slaughtered at the seven major slaughterhouses.

In Clasen *et al.* [4], it was demonstrated that samples from five pigs were sufficient to obtain a representative sample of AMR genes at farm level. However, it

was not known in advance which farms were sending animals for slaughter on a given day or how many pigs they would send. Hence, a purposive sampling strategy was adopted: when a number of pigs from the same farm were identified at the slaughter line, five of the pigs were sampled. That this approach resulted in a random sample that was later verified using meat inspection data from the sampling period [unpublished data]. Slaughterhouse technicians, who were introduced to the sampling methods by the first author on the first sampling day, collected the samples. The samples were taken at the slaughter line after the gut was removed from the carcass by squeezing a small amount of faeces out of the rectum into an empty 12.5 ml sample glass. The samples were kept at room temperature until all samples were collected for the day, and were then placed in a Styrofoam box with cooling elements and mailed overnight to the laboratory. Some deliveries were delayed by one day, but the cooling element was still frozen at arrival and the samples were deemed to be valid.

Quantification of AMR gene levels

The five samples per farm were pooled into a single aliquot and AMR levels were quantified as described by Clasen *et al.* [4]. Pooling was performed by taking an amount fitting the eye of a 10 µl inoculation loop from each of the five samples and dissolving it in 3.5 ml phosphate buffered saline (PBS). The pooled samples were vortexed individually and 2 ml of them was stored at -20°C until further processing. DNA was extracted with the Maxwell 16 Blood DNA Purification Kit (Promega) and DNA concentrations were diluted to 40 ng/µl. Seven AMR genes (*ermB*, *ermF*, *sulI*, *sulIII*, *tet(M)*, *tet(O)* and *tet(W)*) were included in the study as a high-throughput real-time PCR (qPCR) assay was optimised and ready to use [4]. The genes were quantified using the high-capacity qPCR chip 'Gene Expression 192 × 24' (Fluidigm) with two technical replicates. The amplification efficiency of the primers was determined by standard curves and obtained results were normalised with 16S ribosomal DNA, which was used as the reference gene.

Data analysis

Raw quantification cycle (C_q) values generated by the qPCR were taken from the Fluidigm Real-Time PCR Analysis Software version 4.1.3 [19] and exported to R version 3.2.2 [20]. The mean of the C_q values for

technical replicates for each sample per gene was calculated. The C_q values were corrected with the interplate calibrators included in all runs, along with an efficiency calibration [21] calculated from standard curves generated for each of the primer sets [4]. The C_q value reflects the number of PCR cycles until a pre-defined threshold is reached. Therefore, high C_q values reflect a low-level presence of the gene. Values above gene-specific limits of detection [4] were coded as non-detects. Relative quantification (RQ) values indicate the quantity of genes in relation to the total amount of bacterial DNA found in the sample. The latter was measured by the reference gene 16S. The RQ values were calculated using the Livak method [22] as follows:

$$RQ_{\text{primer}_{\text{setX}}} = 2^{-(C_{q,\text{gene of interest}} \pm C_{q,\text{reference gene}})}$$

The RQ value was calculated for all genes except *sulI* and *sulIII*. Samples with non-detects were excluded before calculating the RQ values. Due to a large number of non-detects among the samples for *sulI* and *sulIII*, these genes were dichotomised as present or absent and analysed on a binary scale. The gene was deemed to be present if the qPCR assay resulted in a C_q value even though it was above limit of detection.

Genes with RQ values were also grouped according to the quantiles of the RQ values, as it is not known whether quantitative levels of AMR genes measured by the C_q values show a linear relation to the amount of the gene present in the sample.

Spatial analyses

To test the hypothesis that the distribution of the seven AMR genes were not randomly spatially distributed two sets of complimentary spatial analyses were conducted. First, spatial cluster analysis using scan statistic was performed to identify significant areas with significantly higher or lower risk (or higher or lower mean RQ values) of the seven AMR genes. Secondly, risk maps created using kriging and kernel density estimation were developed to allow us to visualise and describe the geographic distribution of AMR genes.

Cartesian coordinates given in UTM EUREF89 zone 32 format were obtained from the national Central Husbandry Register where all pig farms in Denmark are registered with a unique identification number [23].

Spatial cluster analysis

The spatial scan statistic is a non-parametric test for the presence of clustering of events, accounting for the geographically irregular distribution of (in this example) the Danish pig farm population at risk. The spatial scan statistic is a cluster detection test able both to identify and to test the significance of specific clusters while it simultaneously provides the location of the clusters. Purely spatial cluster analyses were performed to identify spatial clusters of low and high levels of the AMR genes. Briefly, the test sequentially centres a circle or an ellipse in each farm in the study population and compares the RQ values of the AMR genes inside the circle with the RQ values of the farms outside. This circle or ellipse is called the search window. The search window will be increased until it reaches a predefined maximum. The predefined maximum can either be a specified size the search window (i.e. radius of the circle) or a maximum proportion of the population at risk inside a cluster. Often the maximum is set by using existing epidemiological knowledge of the disease in question. However, in this study no such information was available and different settings were used. The likelihood function was computed for each search window. The cluster with the highest likelihood constitutes the most likely cluster. Spatial scan analysis was carried out for the seven AMR genes separately. Depending on the type of the variable used for the analysis different models (i.e. statistical distributions) can be selected. Three different models were used:

- (1) A normal model [24] for continuous RQ values for *ermB*, *ermF*, *tet(M)*, *tet(O)* and *tet(W)*. The model calculates the mean within and outside the search window and the level of significance is calculated for the difference between the two means. The normal model implemented in SaTScan can also handle non-normal data [24].
- (2) A multinomial model [25] for ordinal RQ values in quantiles for *ermB*, *ermF*, *tet(M)*, *tet(O)* and *tet(W)*. The model calculates the expected and observed number of observations within each category for each search window and thus results in a relative risk for each of the four categories of the genes in relation to the other categories.
- (3) A Bernoulli model [26] for binary values for *sulI* and *sulII*. Samples where the gene was present were defined as cases, and samples where the gene was not present were defined as controls.

For each model and gene, the cluster analysis was run with different parameter settings for the shape of the search window (elliptic or circular) and the maximum percentage of the population at risk was included in clusters (1, 5, 12.5, 25 or 50% of the population). For the Bernoulli and normal models, the search for high- and low-level clusters was carried out simultaneously. The test statistics are generated using a randomisation process based on Monte Carlo simulation. The number of iterations for all tests was set to 999. The most likely cluster and a number of secondary clusters will be identified. Only secondary clusters that did not overlap with the most likely cluster were requested. If a cluster is identified the test determines its significance and the cluster declared statistically significant if the *P*-value was less than the α level of 0.05.

Risk maps

Kriging and kernel density estimation were used to estimate values of a variable at an unmeasured location from observed values at surrounding locations. Kriging was used for continuous variables (RQ values of the AMR gene levels) and kernel density estimation was used for binary variables (presence of *sulI* and *sulII* genes). Kriging and kernel density estimation techniques were used to describe the first-order trends in the spatial distribution of AMR genes.

For both kriging and kernel density estimation analyses a regular grid comprised of individual cells 5 km length east to west and 5 km north to south was superimposed over the geographic boundaries of Denmark.

The ordinary kriging and the kernel density estimation analyses were done according to Bihmann *et al.* [10] where details on mathematical equations can be found. The methods are explained briefly in the following sections.

Kriging

Kriging is considered an optimal method of spatial prediction of variables representing a spatially continuous surface. It refers to a family of least-square linear regression algorithms that attempt to predict values of a variable at locations where data are not observed, based on the spatial pattern of the observed data. Ordinary kriging is a common method to use and it relies on the observations of the target variable and its corresponding spatial positions. Kriging has

the advantage that along with a smooth surface of predicted values, prediction variance is also estimated. Kriging is a weighted average of observed values, where the weight function is based on the spatial variation between measurements which is modelled by the semivariogram. Kriging can be used to estimate the spatial distribution of a disease measured at farm level (e.g. farm level incidence or prevalence of infected animals) [10, 27–29]. Although the disease variable is measured in particular farms it is assumed that the disease variable represents a spatially continuous surface of the disease level. This can be interpreted as the disease level we would expect at the location of a virtual (or new) farm. The method assumes a stationary rate, but it has also been effective on non-stationary rates [30].

Semivariograms were derived to obtain estimates of three parameters (range of influence, nugget and partial sill) that were then used to estimate the spatial variation and the weight function for kriging. Semivariograms measure the degree of dissimilarity between observations as a function of the distance. Typically, semi-variance, half the variance, increases as the distance between the locations grows until at some point the locations are considered independent of each other and the semi-variance no longer increases. If neighbouring data points resemble each other more closely than those further apart spatial dependence is assessed to be present. This would be indicated by a rising curve in the semivariogram, which plateaus as the similarities diminish with increasing distance. A semivariogram is characterised by three parameters the nugget effect, the sill and the range of influence. The nugget effect refers to the variability in the variable that cannot be explained by distance between the observations. Many factors influence the magnitude of the nugget effect including imprecision in sampling techniques, underlying variability of the attribute that is being measured, and the minimum spacing between observations. The latter is due to no observations sampled close to each other, it is impossible to estimate spatial dependence at small distances. The sill refers to the maximum observed variability in the data and corresponds to the variance of the data. The difference between the sill and the nugget effect (the partial sill) represents the amount of observed variation that can be explained by distance between observations. Finally, the range of influence is the point at which the semi-variance stops increasing and represents the distance at which two observations on average are not

correlated. Often a model is fit to the semivariogram to estimate the parameters and in order to make use of the spatial dependence in other statistical techniques, including the kriging analysis.

In the present study kriging has been used to estimate the spread patterns of AMR gene levels. The spatial dependence would be a result of neighbouring farms having more similar AMR gene levels than those that are further apart. For each gene, two semivariograms were created, one as a primary analysis and the other as a sensitivity analysis. The semivariograms were created in two ways, the first was chosen where possible and the second as an alternative:

- (1) Two models with different parameter settings were fitted to the same semivariogram.
- (2) Two semivariograms were fitted using different lag widths, to which models with equal settings were fitted.

An exponential semivariogram model was used and the best fitted model was chosen. The model semivariogram parameter estimates the partial sill, the nugget effect, and the practical range of influence (three times the range of influence reported by the fitted model) were reported.

Directional semivariograms in four directions (north, north-east, east and south-east) were estimated to visually evaluate anisotropy. Anisotropy exists if there are substantial differences between the semivariograms in different directions.

For each semivariogram model, ordinary kriging was performed using the grid and repeated with different numbers of nearest neighbours in the kriging estimation. The number of neighbours ranged from 15 to 50 farms with intervals of five farms. A smoothed map showing the distribution of AMR gene levels measured in RQ values across Denmark was then produced. Furthermore, the prediction variances were plotted as an estimate of the uncertainties in the maps.

Kernel density estimation

The first-order spatial trend in the distribution of pig farms with *sulI* and *sulIII* genes was described using kernel density estimation methods. Kernel density estimation gives weighted means for each location in the study region. Here a Gaussian, edge-corrected kernel smoothed map of gene-positive farms (showing the number of gene-positive farms per square kilometre) was computed as the numerator and a kernel

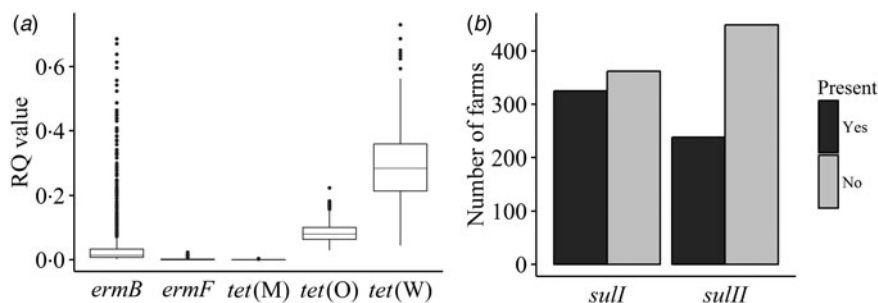


Fig. 1. Descriptive statistics of the genes. (a) The distribution of the RQ values for the *ermB*, *ermF*, *tet(M)*, *tet(O)* and *tet(W)* genes. (b) The distribution of *sulI* and *sulIII* genes; grey indicates the absence of the gene, while black indicates the presence of the gene.

smoothed map of all of the sampled pig farms computed as the denominator using the ‘spatialkernel’ package in R [20, 31]. A raster map showing the prevalence of *sulI*- and *sulIII*-positive farms (expressed as the number of gene-positive farms per 100 farms per square kilometre) was produced by dividing the numerator raster map by the denominator. Bandwidths for each of the kernel smoothed maps were calculated using the normal optimal method and an average of the bandwidths for the positive and negative farms were used [32].

Software for spatial analyses

All data were handled in R version 3.2.2 [20]. Spatial cluster analysis was performed in SaTScan version 9.4.1 [33]. Maps were derived using the ‘sp’ package [20, 34]. Semivariograms, ordinary kriging and kernel density estimation were performed using the ‘gstat’ package in R version 3.2.2 [20, 35]. Bandwidths for the kernel density estimations were computed using ‘sm’ package in R version 3.2.2 [20, 32].

RESULTS

Study population

The cross-sectional study comprised a study population of 687 Danish indoor non-organic pig farms with finishers sent to slaughter in Denmark. Samples were collected from 129, 253, 125, 104 and 76 farms, respectively, from the five slaughterhouses. More information regarding the farms can be found elsewhere [unpublished data]. The sampling technique resulted in an almost random spatial distribution of the study population with respect to the Danish finisher pig farms at risk with relative under-sampling in the western part of Jutland. The spatial distribution of AMR genes in this area should be evaluated carefully [unpublished data].

Levels of the AMR genes

The distribution of the RQ values for each gene and the distributions of presence and absence of *sulI* and *sulIII* can be seen in Figure 1. For *tet(M)*, 43 samples were excluded from the analyses on the basis of non-detection, and for *ermF*, 19 samples were excluded from the analysis for the same reason. Of these samples, two were excluded from analysis for both *ermF* and *tet(M)*. No samples were excluded for *ermB*, *tet(O)* and *tet(W)*.

Spatial cluster analysis

Different parameter settings resulted in slightly different cluster locations and sizes. If two clusters were found in the same area the cluster including the highest number of farms was shown on the map (Fig. 2). The following significant spatial clusters were found: two high-risk clusters for *ermF*, one low-risk cluster for *ermF*, *ermB* and *tet(W)*, and one high-risk cluster for *sulIII*. For *ermB*, *ermF* and *tet(W)*, the reported clusters were found with the multinomial model. No significant spatial clusters were found for *sulI*, *tet(M)* and *tet(O)*.

Risk maps

Semivariograms for the *ermB*, *ermF*, *tet(M)*, *tet(O)* and *tet(W)* genes are shown in Figure 3. Table 1 shows the parameters for the chosen exponential semivariogram model. The model estimates for *tet(O)* were very similar, whereas the model estimates for *tet(W)*, *tet(M)*, *ermB* and *ermF* differed between the two models (results not shown). The directional semivariograms showed no indication of anisotropy for any of the genes (results not shown).

Using the two models from the semivariogram and different numbers of nearest neighbours in the

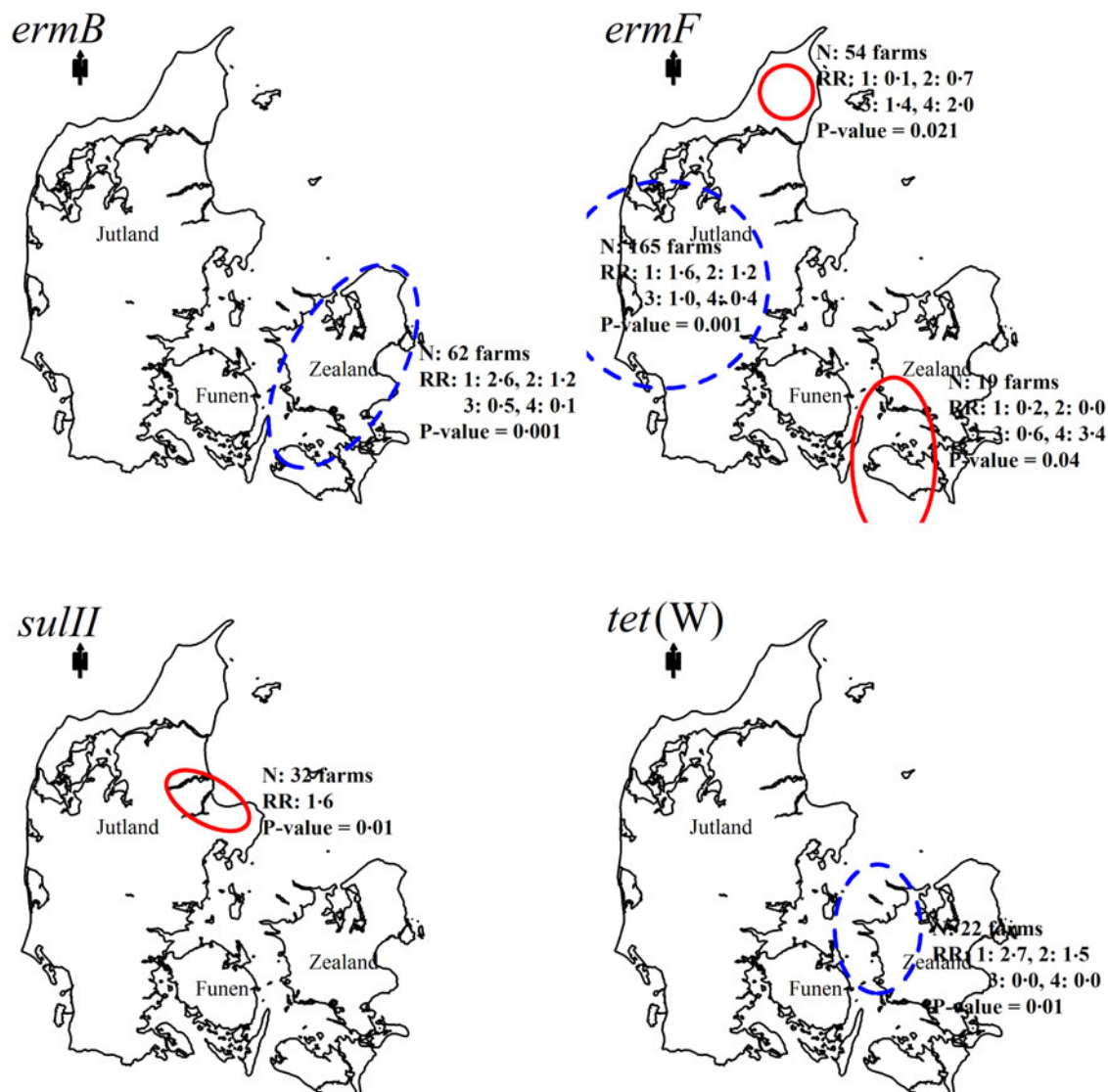


Fig. 2. Results of cluster analysis of AMR genes. Blue dashed lines indicate low-risk clusters, while red solid lines indicate high-risk clusters. Relative risk (RR) for multinomial models (i.e. *ermB*, *ermF* and *tet(W)*), the RR is indicated for each of the categories (1–4) in relation to the other models. For the Bernoulli model (i.e. *sulII*), the RR indicates the risk of being positive relative to the risk of being negative. *N*, number of farms in the cluster.

estimation of the RQ value only introduced minor changes to the estimated value. The visual patterns of high, medium and low levels for all genes did not change. A stable kriging map was produced with 40 nearest neighbours, so this number was chosen in the shown kriging maps (Fig. 4). Colours going from blue to increasing darker red on the maps indicate an increasing RQ value reflecting a higher level of the AMR gene.

Figure 5 shows the results of the kernel density estimation for *sulI* and *sulII*. Colours going from yellow to increasingly darker red indicate increasing population prevalence for the genes found in the area. The

common bandwidth used for both genes was (22773 m and 32971 m, respectively).

DISCUSSION

This study showed that some AMR genes found in faecal samples from pigs are not completely randomly spatially distributed. In the spatial cluster analysis, one low-risk cluster for *ermF*, *ermB* and *tet(W)* and two high-risk clusters for *ermF* were identified, together with one high-risk cluster for *sulII*. No clusters were found for *sulI*, *tet(O)* and *tet(M)*. The size and location of the clusters varied among the genes.

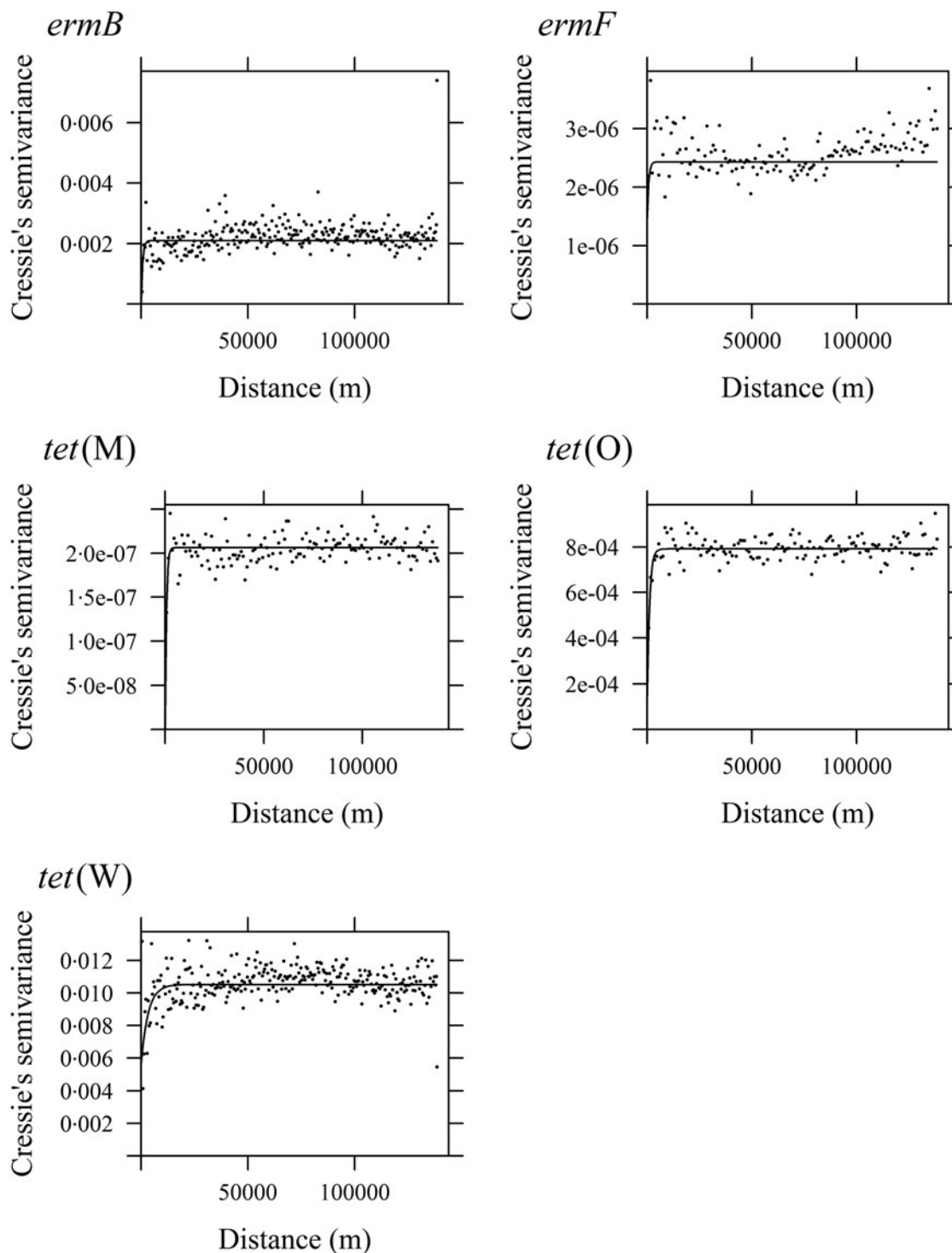


Fig. 3. Semivariograms. On each semivariogram, the fitted model is shown as black line. Each dot in the semivariogram cloud represents a point-pair of farms. Point-pairs comprised by farms within the distance of a specified lag width are plotted against the half of the variation (semi-variance) in the RQ values for the gene on the y-axis. When the cloud flattens out the relationship between the pairs of locations beyond this distance is no longer correlated. This distance is defined as the range of influence. However, when an exponential model is used the range of influence is multiplied by three to get the practical range of influence. The sill is defined as the semi-variance at the point where the semi-variance model flattens and the nugget effect is the intersection of the model and the y-axis. The partial sill is the sill minus the nugget.

Table 1. *Semivariogram settings and parameter estimates*

Gene	Parameter estimates in fitted exponential semivariogram model			Model setting* Weighing	Semivariogram settings*	
	Nugget	Partial sill	Practical range of influence, km		Lag width, m	Cut-off, m
<i>ermB</i>	0	0.0021	2.0	N_j/h_j^2	500	80 000
<i>ermF</i>	$1.47e^{-6}$	$9.5e^{-7}$	2.1	N_j/h_j^2	1000	None
<i>tet(M)</i>	0	$2.06e^{-7}$	1.9	$N_j/\{\gamma(h_j)\}^2$	1000	None
<i>tet(O)</i>	0.00015	0.000640	3.4	N_j/h_j^2	1000	None
<i>tet(W)</i>	0.0057	0.0048	9.9	$N_j/\{\gamma(h_j)\}^2$	500	50 000

* Refers to settings in the programming in R. h_j represents the distance in metres; N_j represents the number of point-pairs; lag width represents the step size of distance intervals for creating the semivariogram; and cut-off represents the maximum distance at which pairs of data points will be considered for inclusion in the semivariogram.

The clusters on Zealand include fewer farms than clusters of a similar size in Jutland. This is due to an uneven distribution of farm locations in Denmark [36]. The clusters found by the multinomial model have a relative risk above one in either category one and two, or in category three and four, meaning that they are either high-risk clusters or low-risk clusters. No mixed clusters were found. It is possible that the current sample size is insufficient to show clustering for the three genes where no clusters were found, if they truly exist. The risk maps created with kriging analysis and the kernel density estimation were consistent with the results of the spatial cluster analysis. Both interpolation methods and the spatial cluster analysis showed consistent results with different parameter settings, indicating that the findings regarding the absence and presence of spatial differences for the genes could be considered reliable. The spatial scan statistic provides the location, size and significance of any clusters identified. Because its approach is circular or elliptic in nature, the assessment of clusters along natural or artificial borders may be biased to some extent. However, the spatial scan statistics can account for irregular dispersal of the farms over space which is the case for the distribution of pig farms in most countries including Denmark. On the other hand, this irregular dispersal of farms can lead to unreliable estimates of interpolation. The kriging analysis provide an error map and this show that for most parts of the country the predictive values are provided with the same error level. In the north-eastern part of Zealand very few pig farms are located why this area is associated with a higher prediction variance and thus a higher uncertainty is associated with the RQ values predicted in those areas.

To the best of our knowledge, this is the first study to report spatial patterns of AMR genes of total community DNA from porcine faeces. However, it is not the first to evaluate spatial patterns of AMR in Danish pig farms. A previous study in Denmark [15] evaluated spatial patterns in ampicillin resistance in *Escherichia coli*. However, *E. coli* only constitutes a small part of the porcine gut microbiota. The present study evaluates AMR genes in total community DNA, thereby taking into account all bacteria in porcine faeces and for several AMR genes. This means that there is no indication of which bacteria are found in the samples and in which bacteria the AMR genes are harboured. The AMR genes included in the study might be harboured in different bacterial species. The spatial distribution of the bacteria species would therefore affect the spatial distribution of the AMR genes. This could be the reason why spatial autocorrelation is found for some of the genes and not for other genes.

Many factors contribute to the occurrence of AMR in farms and in the environment. The different patterns in the semivariogram might suggest that the genes are spread by different mechanisms. Local differences in antimicrobial usage or in the presence of bacterial species are two factors that might explain the spatial patterns of AMR genes [37]. There is evidence for local variation in the prevalence of different bacteria in Danish pigs [38]. The local distribution of bacteria might be affected by introducing live pigs into the farm, as these movements of live pigs for meat production occur very locally in Denmark [39].

Antimicrobial usage has been shown to be spatially clustered in Denmark [36] and therefore AMR genes could be expected to cluster accordingly. It is

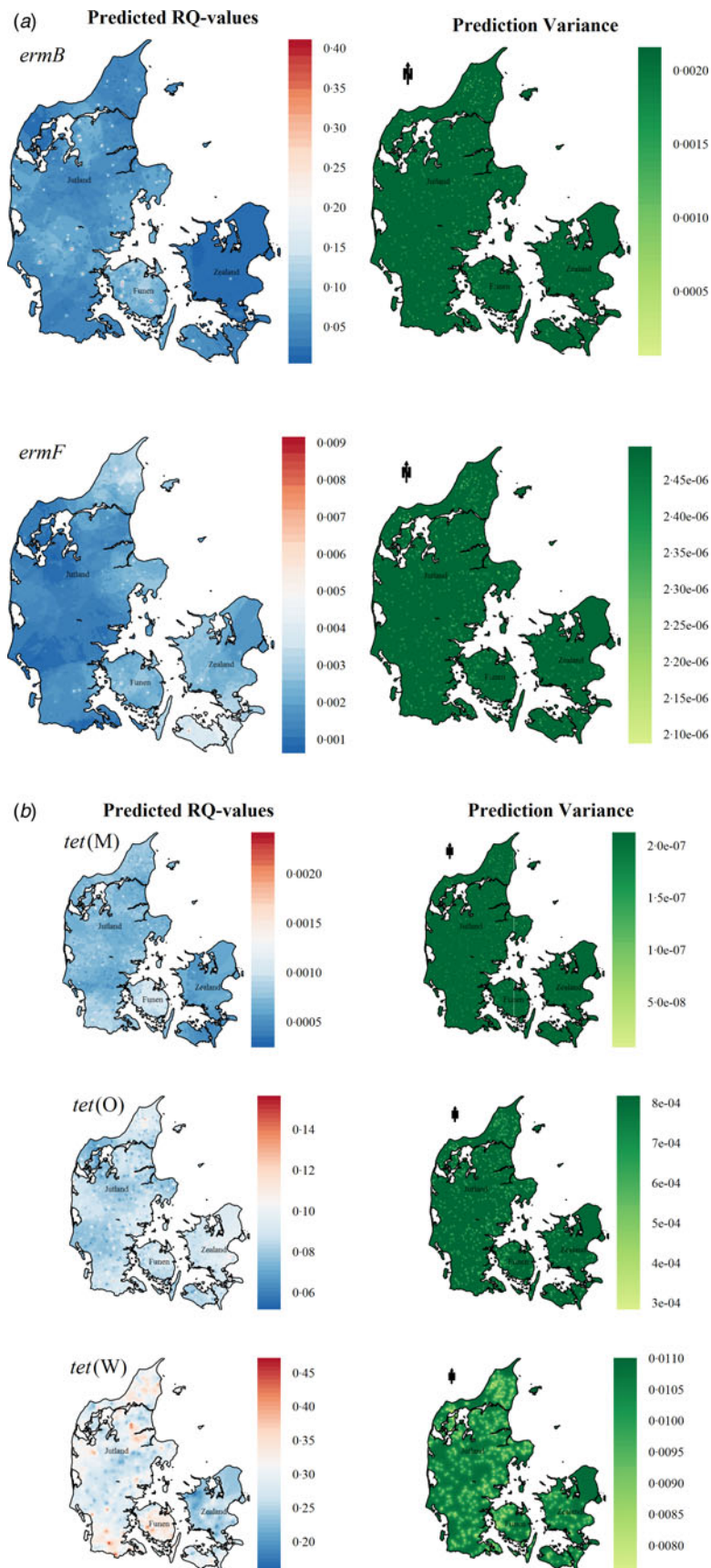


Fig. 4. (a) Risk maps for the levels of *ermB* and *ermF* genes produced by ordinary kriging. Each panel shows the distribution of predicted RQ values and the corresponding map for the prediction variance. The legends are unique for each gene due to the heterogeneous distributions of the genes even though same colour scale is used to produce the maps. (b) Risk maps for the levels of *tet(M)*, *tet(O)* and *tet(W)* genes produced by ordinary kriging. Each panel shows the distribution of predicted RQ values and the corresponding map for the prediction variance. The legends are unique for each gene due to the heterogeneous distributions of the genes even though same colour scale is used to produce the maps.

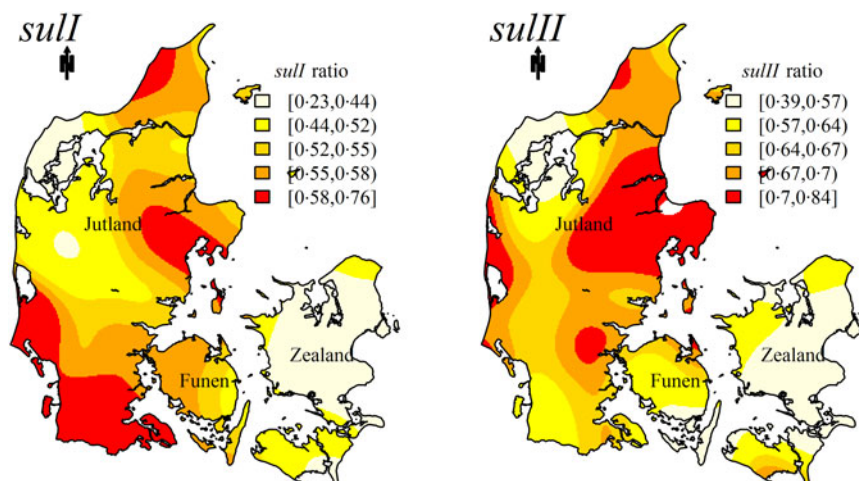


Fig. 5. Risk maps for the prevalence of the *sulI* and *sulII* genes. The maps are created using kernel density estimation.

interesting to note that for genes coding for resistance to the same antimicrobial class, significant spatial clusters were found for some of the genes, such as *sulII* and *tet(W)*, but not for others such as *sulI* and *tet(M)* or *tet(O)*, indicating that local patterns of antimicrobial usage cannot explain the spatial clusters alone. Another explanation for spatial patterns of AMR could be local differences in feeding strategies. It was not possible to assess this factor in the current study, as feeding practice is not recorded in any nationwide Danish register. Feeding strategies can alter the composition of the pig gut microbiome, leading to an increase in *Bifidobacterium* spp., which constitute a large part of the animal gut microbiota and promotes gut health [40]. The *tet(W)* or *tet(M)* genes are highly prevalent among *Bifidobacterium* spp. [41]. Furthermore, *tet(M)*, *tet(O)* and *tet(W)* have been found in different types of swine feed [42] and might be present in probiotic microorganisms also used in some feeding schedules [43] which could significantly affect the distribution. The distribution of the *tet* genes in particular might be caused by differences in feeding practices or gut microbiota of the pigs. This study has shown that the *tet(W)* gene is present at the highest levels, and that there is a large variation in these levels among the sampled farms.

Within the practical range of influence, the farms are correlated, but due to the methods used in this study, it is not possible to estimate the size of the auto-correlations by for example a correlation coefficient. We deemed this to be beyond the scope of the paper as it is a purely descriptive study. Furthermore, there is no available method for assessing the adequacy of a fitted semivariogram model, and these results should

therefore be treated with caution. In addition the accuracy of the semivariogram at small scale is weak because it is not possible to sample within a smaller distance than the distance between two pig farms. It is also important to note that only indoor, non-organic finisher farms were included in the study, and the spatial relationship might not be applicable for all Danish pig farms.

The risk maps showed results consistent of the cluster analysis, but also that the spatial difference were at small scale. Spatial clusters were found for specific AMR genes in Danish pig farms. However, the spatial distribution does not reveal major cold or hotspots in Denmark for the genes in question.

The conclusion was that the geographical location of a pig farm is not the major risk factor for presence or high levels of the AMR genes included in the study. Further analyses are needed to explain the clusters found in this study.

ACKNOWLEDGEMENTS

This research was carried out as part of a grant from Danish Veterinary and Food administration regarding science-based advice on AMR. The authors would like to thank the employees of Danish Crown and Tican, especially Marie B. Hansen, Vagner Bøge and Henrik Bækstrøm, for their cooperation, assistance with sampling, and help in setting up the study. They thank Jesper Larsen for support with the meat inspection data, Margarida Arede for help with sampling, Joanna Zeitman Amenuvor for help with sample pooling, and Karin Tarp and Sophia Rasmussen for their technical assistance. They would also like

to thank Kerstin Skovgaard for advice on handling qPCR data.

DECLARATION OF INTEREST

None.

REFERENCES

1. **HM Stationary Office.** *Report of the Joint Committee on the Use of Antibiotics in Animal Husbandry and Veterinary Medicine (1969)*. London: 1969.
2. **ECDC/EMA Joint Working Group.** *The Bacterial Challenge: Time to React – A Call to Narrow the Gap Between Multidrug-Resistant Bacteria in the EU and the Development of New Antibacterial Agents*. Stockholm: European Centre for Disease Prevention and Control, 2009.
3. **Tremblay C-L, et al.** Antibiotic-resistant *Enterococcus faecalis* in abattoir pigs and plasmid colocalization and cotransfer of *tet(M)* and *erm(B)* genes. *Journal of Food Protection* 2012; **75**(9): 1595–1602. DOI: 10.4315/0362-028X.JFP-12-047.
4. **Clasen J, et al.** Determining the optimal number of individual samples to pool for quantification of average herd levels of antimicrobial resistance genes in Danish pig herds using high-throughput qPCR. *Veterinary Microbiology* 2016; **189**: 46–51. DOI: 10.1016/j.vetmic.2016.04.017.
5. **World Health Organization.** *Tackling antibiotic resistance from a food safety perspective in Europe*. 2011. <http://www.euro.who.int/en/publications/abstracts/tackling-antibiotic-resistance-from-a-food-safety-perspective-in-europe>.
6. **Moodley A, Guardabassi L.** Transmission of IncN plasmids carrying blaCTX-M-1 between commensal *Escherichia coli* in pigs and farm workers. *Antimicrobial Agents and Chemotherapy* 2009; **53**(4): 1709–1711. DOI: 10.1128/AAC.01014-08.
7. **Nijsten R, et al.** Resistance in faecal *Escherichia coli* isolated from pig farmers and abattoir workers. *Epidemiology and Infection* 1994; **113**(1): 45–52.
8. **Sengeløv G, et al.** Bacterial antibiotic resistance levels in Danish farmland as a result of treatment with pig manure slurry. *Environment International* 2003; **28**(7): 587–595.
9. **Agersø Y, Sengeløv G, Jensen LB.** Development of a rapid method for direct detection of *tet(M)* genes in soil from Danish farmland. *Environment International* 2004; **30**(1): 117–122. DOI: 10.1016/S0160-4120(03)00156-9.
10. **Bihmann K, et al.** Spatial differences in occurrence of paratuberculosis in Danish dairy herds and in control programme participation. *Preventive Veterinary Medicine* 2012; **103**(2–3): 112–119. DOI: 10.1016/j.prevetmed.2011.09.015.
11. **Ersbøll AK, Nielsen LR.** Spatial patterns in surveillance data during control of *Salmonella* Dublin in bovine dairy herds in Jutland, Denmark 2003–2009. *Spatial and Spatio-temporal Epidemiology* 2011; **2**(3): 195–204. DOI: 10.1016/j.sste.2011.07.003.
12. **Vigre H, et al.** Spatial and temporal patterns of pig herds diagnosed with Postweaning Multisystemic Wasting Syndrome (PMWS) during the first two years of its occurrence in Denmark. *Veterinary Microbiology* 2005; **110**(1–2): 17–26. DOI: 10.1016/j.vetmic.2005.07.001.
13. **Chowdhury S, et al.** The effect of presence of infected neighbouring farms for the *Campylobacter* infection status in Danish broiler farms. *Spatial and Spatio-temporal Epidemiology* 2012; **3**(4): 311–322. DOI: 10.1016/j.sste.2012.06.001.
14. **Cox R, et al.** Spatial and temporal patterns in antimicrobial resistance of *Salmonella typhimurium* in cattle in England and Wales. *Epidemiology and Infection* 2012; **140**(11): 2062–2073. DOI: 10.1017/S0950268811002755.
15. **Abatih EN, et al.** Space–time clustering of ampicillin resistant *Escherichia coli* isolated from Danish pigs at slaughter between 1997 and 2005. *Preventive Veterinary Medicine* 2009; **89**(1–2): 90–101. DOI: 10.1016/j.prevetmed.2009.02.002.
16. **Anonymous.** DANMAP 2014 - use of antimicrobial agents and occurrence of antimicrobial resistance in bacteria from food animals, food and humans in Denmark. <http://www.danmap.org/2015>.
17. **Abatih EN, Emborg HD, Jensen VF, et al.** Regional, seasonal, and temporal variations in the prevalence of antimicrobial-resistant *Escherichia coli* isolated from pigs at slaughter in Denmark (1997–2005). *Foodborne Pathogens and Disease* 2009; **6**(3): 305–319. DOI: 10.1089/fpd.2008.0168.
18. **Statistics Denmark.** StatBank Denmark. <http://www.statistikbanken.dk/statbank5a/SelectVarVal/Define.asp?Maintable=SVIN&PLanguage=12015>.
19. **Fluidigm.** Fluidigm Real-Time PCR analysis, v 4.1.3. 2014.
20. **R Core Team.** R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria. 2015. <https://www.R-project.org/>.
21. **Ståhlberg A, et al.** RT-qPCR work-flow for single-cell data analysis. *Methods* 2013; **59**(1): 80–88. DOI: 10.1016/j.ymeth.2012.09.007.
22. **Livak KJ, Schmittgen TD.** Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻ $\Delta\Delta$ CT method. *Methods* 2001; **25**(4): 402–408. DOI: 10.1006/meth.2001.1262.
23. **Anonymous.** Order (DK) 1237/2013 Regarding registration of herds in CHR [in Danish]. 2013. <https://www.retsinformation.dk/Forms/R0710.aspx?id=185121>
24. **Kulldorff M, Huang L, Konty K.** A scan statistic for continuous data based on the normal probability model. *International Journal of Health Geographics* 2009; **8**(1): 58. DOI: 10.1186/1476-072X-8-58.
25. **Jung I, Kulldorff M, Richard OJ.** A spatial scan statistic for multinomial data. *Statistics in Medicine* 2010; **29**(18): 1910–1918. DOI: 10.1002/sim.3951.
26. **Kulldorff M.** A spatial scan statistic. *Communications in Statistics – Theory and Methods* 1997; **26**(6): 1481–1496. DOI: 10.1080/03610929708831995.

27. **Graham SL, et al.** Spatial distribution of antibodies to *Salmonella enterica serovar* Typhimurium O antigens in bulk milk from Texas dairy herds. *Preventive Veterinary Medicine* 2005; **69**(1–2): 53–61. DOI: 10.1016/j.prevetmed.2005.01.018.
28. **Kedmi M, et al.** Assessment of the productivity effects associated with epizootic hemorrhagic disease in dairy herds. *Journal of Dairy Science* 2010; **93**(6): 2486–2495. DOI: 10.3168/jds.2009-2850.
29. **Phiri BJ, et al.** Spatiosurvival analysis of mortality on smallholder dairy farms in Tanga and Iringa regions of Tanzania. *Tropical Animal Health and Production* 2012; **44**(4): 827–834. DOI: 10.1007/s11250-011-9974-2.
30. **Gotway CA, Wolfinger RD.** Spatial prediction of counts and rates. *Statistics in Medicine* 2003; **22**(9): 1415–1432. DOI: 10.1002/sim.1523.
31. **Zheng P, Diggle P.** Spatialkernel: Nonparametric estimation of spatial segregation in a multivariate point process. R package, v 0.4-19. 2013. <http://CRAN.R-project.org/package=satialkernel>
32. **Bowman AW, Azzalini A.** R package “sm”: non-parametric smoothing methods, v 2.2-5.4. 2014. <http://www.stats.gla.ac.uk/~adrian/sm>, http://azzalini.stat.unipd.it/Book_sm
33. **Kuldorff M.** and Information Management Services Inc. SaTScan™ ver. 9-4-1: Software for the spatial and space-time scan statistics. 2015.
34. **Pebesma EJ, Bivand RS.** Classes and methods for spatial data in R. *R News* 2005; **5**(2). <http://cran.r-project.org/doc/Rnews/>.
35. **Pebesma EJ.** Multivariable geostatistics in S: the gstat package. *Computers & Geosciences* 2004; **30**: 683–691.
36. **Fertner M, et al.** Persistent spatial clusters of prescribed antimicrobials among Danish pig farms – a register-based study. Kimman T (ed). *PLoS ONE* 2015; **10**(8): e0136834. DOI: 10.1371/journal.pone.0136834.
37. **Singer RS, Reid-Smith R, Sischo WM.** Stakeholder position paper: epidemiological perspectives on antibiotic use in animals. *Preventive Veterinary Medicine* 2006; **73**(2–3): 153–161. DOI: 10.1016/j.prevetmed.2005.09.019.
38. **Benschop J, et al.** Informing surveillance programmes by investigating spatial dependency of subclinical *Salmonella* infection. *Epidemiology and Infection* 2009; **137**(9): 1348–1359. DOI: 10.1017/S0950268809002106.
39. **Bigras-Poulin M, et al.** Relationship of trade patterns of the Danish swine industry animal movements network to potential disease spread. *Preventive Veterinary Medicine* 2007; **80**(2–3): 143–165. DOI: 10.1016/j.prevetmed.2007.02.004.
40. **Modler HW.** Bifidogenic factors – sources, metabolism and applications. *International Dairy Journal* 1994; **4**(5): 383–407. DOI: 10.1016/0958-6946(94)90055-8.
41. **Aires J, Doucet-Populaire F, Butel MJ.** Tetracycline resistance mediated by *tet*(W), *tet*(M), and *tet*(O) genes of *Bifidobacterium* Isolates from humans. *Applied and Environmental Microbiology* 2007; **73**(8): 2751–2754. DOI: 10.1128/AEM.02459-06.
42. **Aminov RI, Garrigues-Jeanjean N, Mackie RI.** Molecular ecology of tetracycline resistance: development and validation of primers for detection of tetracycline resistance genes encoding ribosomal protection proteins. *Applied and Environmental Microbiology* 2001; **67**(1): 22–32. DOI: 10.1128/AEM.67.1.22-32.2001.
43. **Sharma P, et al.** Antibiotic resistance among commercially available probiotics. *Food Research International* 2014; **57**: 176–195. DOI: 10.1016/j.foodres.2014.01.025.