ARTICLES
https://doi.org/10.1038/s41564-018-0192-9

nature microbiology

Corrected: Author Correction

Abundance and diversity of the faecal resistome in slaughter pigs and broilers in nine European countries

Patrick Munk[1], Berith Elkær Knudsen[1], Oksana Lukjancenko[1], Ana Sofia Ribeiro Duarte[1], Liese Van Gompel[2], Roosmarijn E. C. Luiken[2], Lidwien A. M. Smit[2], Heike Schmitt[2], Alejandro Dorado Garcia[2], Rasmus Borup Hansen[3], Thomas Nordahl Petersen[1], Alex Bossers[2,4], Etienne Ruppé[5], EFFORT Group[6], Ole Lund[1], Tine Hald[1], Sünje Johanna Pamp[1], Håkan Vigre[1], Dick Heederik[2], Jaap A. Wagenaar[4,7], Dik Mevius[4,7] and Frank M. Aarestrup[1*]

Antimicrobial resistance (AMR) in bacteria and associated human morbidity and mortality is increasing. The use of antimicrobials in livestock selects for AMR that can subsequently be transferred to humans. This flow of AMR between reservoirs demands surveillance in livestock and in humans. We quantified and characterized the acquired resistance gene pools (resistomes) of 181 pig and 178 poultry farms from nine European countries, sequencing more than 5,000 Gb of DNA using shotgun metagenomics. We quantified acquired AMR using the ResFinder database and a second database constructed for this study, consisting of AMR genes identified through screening environmental DNA. The pig and poultry resistomes were very different in abundance and composition. There was a significant country effect on the resistomes, more so in pigs than in poultry. We found higher AMR loads in pigs, whereas poultry resistomes were more diverse. We detected several recently described, critical AMR genes, including *mcr-1* and *optrA*, the abundance of which differed both between host species and between countries. We found that the total acquired AMR level was associated with the overall country-specific antimicrobial usage in livestock and that countries with comparable usage patterns had similar resistomes. However, functionally determined AMR genes were not associated with total drug use.

Antimicrobial resistance (AMR) is considered one of the largest threats to human health[1]. In addition to the use of antimicrobial agents in humans, livestock is considered an important source of AMR, potentially compromising human health[2]. Besides AMR in zoonotic pathogens, AMR in commensal bacteria is worrisome because of its ability to spread horizontally to pathogens.

Multiple studies have shown that the use of antimicrobials in livestock will lead to an increased occurrence of AMR and that the reduction of usage will eventually lead to reduced resistance[3–8]. Several national surveillance programmes have been implemented to monitor the occurrence of AMR in different reservoirs and follow trends over time[1,9–11]. There are major differences in antimicrobial consumption patterns between different countries globally and also within Europe[12]. Major differences in the occurrence of AMR have also been observed among indicator organisms (for example, *Escherichia coli*) isolated from different European countries[3,13]. Current monitoring efforts are mainly based on culturing indicator bacteria followed by phenotypic AMR determination[13,14]. This procedure only targets a limited number of species present in the gut microbiota and, therefore, probably represents only a fraction of its resistome (the collective pool of AMR genes). Metagenomic approaches have been used in several recent studies and have shown that metagenomic read mapping describes AMR abundance in bacterial communities more accurately than commonly used technologies on selected indicator organisms[15–17]. A recent study focused on sampling a diverse group of individual pigs from 11 farms in 3 countries and showed that genetics, age, diet and geography all probably influence the pig microbiota, but little information is available for the poultry microbiota[16].

As part of the European Union-funded EFFORT project (www.effort-against-amr.eu), we sampled >9,000 animals in 181 pig and 178 poultry herds in 9 European countries, generating herd-level composite samples as previously described[17]. Metagenomic sequencing of these samples gives us a unique insight into the abundance, diversity and structure of the acquired pig and broiler resistomes in Europe. An association between AMR gene abundance and national veterinary antimicrobial usage (AMU) was also analysed. The results and raw data presented here can be used as a baseline for future metagenomic AMR monitoring. To our knowledge, this study represents the single largest metagenomic AMR monitoring effort of livestock: both in terms of countries (9), herds included (359), individual animals sampled (>9,000) and sequencing effort (>5,000 Gb)[16].

Results

Acquired resistome characterization. The total AMR load varied significantly across samples, depending on both the host animal

[1]Research Group for Genomic Epidemiology, National Food Institute, Technical University of Denmark, Kongens Lyngby, Denmark. [2]Institute for Risk Assessment Sciences, Utrecht University, Utrecht, the Netherlands. [3]Intomics A/S. Diplomvej 377, Kongens Lyngby, Denmark. [4]Wageningen Bioveterinary Research, Lelystad, the Netherlands. [5]Genomic Research Laboratory, Hôpitaux Universitaires de Genève, Geneva, Switzerland. [7]Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, the Netherlands. [15]A list of participants and their affiliations appears at the end of the paper. *e-mail: fmaa@food.dtu.dk

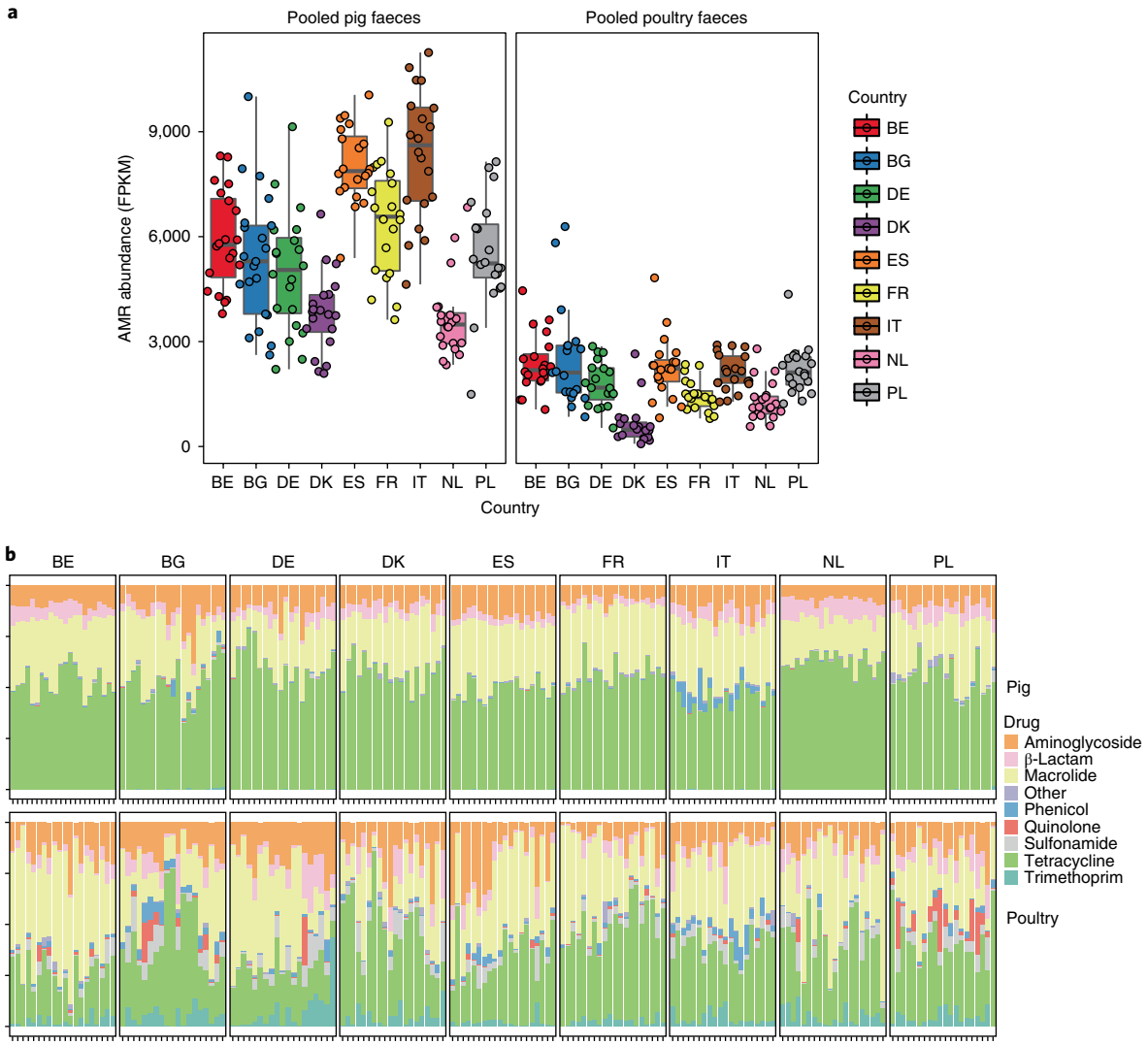

Fig. 1 | Overview of AMR abundance and composition. From read mapping to the ResFinder database, AMR abundance was calculated for each reference gene in each sample. **a**, Box plots showing the total AMR level per sample, stratified by host species and country. Each herd is also represented by a dot with sideways jitter to minimize overplotting. The horizontal box lines represent the first quartile, the median and the third quartile. Whiskers denote the range of points within the first quartile − 1.5× the interquartile range and the third quartile + 1.5× the interquartile range. $n = 359$ metagenomes from independent herds. Twenty metagenomes per livestock species per country were used, with the exception of Bulgarian pigs (21), Bulgarian poultry (19) and German poultry (19). **b**, Stacked bar chart of AMR abundance per type (colours) per sample (*x* axis), proportional to the total AMR within each sample. Note that the two-letter country code is used in all figures: BE, Belgium; BG, Bulgaria; DE, Germany; DK, Denmark; ES, Spain; FR, France; IT, Italy; NL, the Netherlands; PL, Poland.

and the country of origin. In general, pigs had a higher AMR level than poultry (Fig. 1a). The highest AMR levels were found in Italian pigs, where the top four resistance-scoring samples originated, all in excess of 10,000 fragments per kilobase reference per million bacterial fragments (FPKM) AMR. At the lower end of the spectrum were Danish poultry samples that occupied the 11 samples with the least AMR, all below 500 FPKM.

We summed the relative abundance of AMR to the corresponding drug class level for each sample to look for major trends across host species and countries (Fig. 1b). When considering the proportion of the total resistome by AMR phenotype, the pig samples were relatively homogenous: tetracycline AMR was by far the most common, followed by macrolide AMR. β-Lactam and aminoglycoside AMR genes followed by other kinds of AMR were rare. Italian pigs had a notably larger proportion of phenicol AMR than pigs of other countries and it seemed to be consistent across Italian farms. A subset of Bulgarian pig farms had a similar proportion of phenicol AMR.

Among the poultry farms, there was less consistency. Both within and between countries, the relative proportions of AMR per drug were more varied. As in pig samples, tetracycline, macrolide, β-lactam and aminoglycoside AMR made up the majority, but the two latter classes had very minimal contributions in a subset of herds. Sulfonamide and trimethoprim AMR was more abundant in poultry samples than in pig samples across all countries. In many Polish poultry herds, quinolone AMR made up a sizeable fraction of the combined resistome. This was also true for a few non-Polish herds, notably in Bulgaria. For non-proportional graphical representations of the AMR load stratified by sample and drug class, see the Supplementary Material for an unscaled, stacked bar chart (Supplementary Fig. 1) and a heatmap (Supplementary Fig. 2). Class-level AMR relative abundances can be found in Supplementary Table 1.

To characterize the individual components of the resistome, we summed the relative abundance to the gene level, as was done for

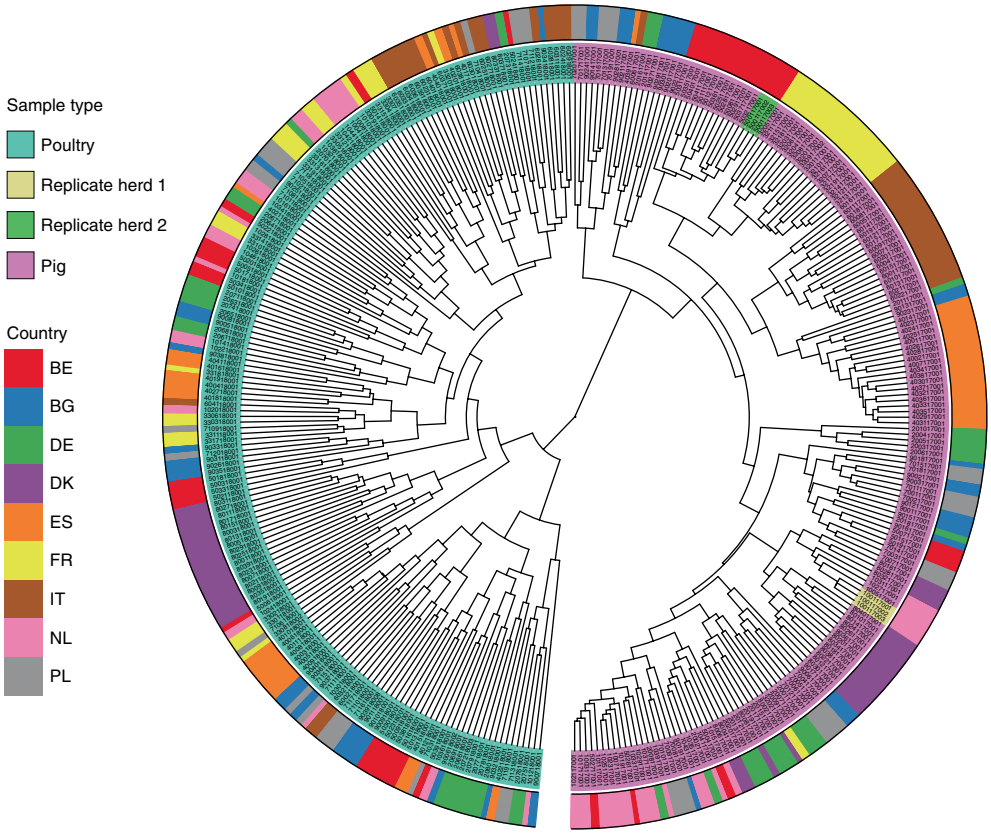

Fig. 2 | Resistome clustering is influenced by both host animal and country. A dendrogram showing the complete linkage clustering of Bray–Curtis dissimilarities between all pig and poultry resistomes. Triple-sampled pig herds are highlighted in separate colours. $n = 363$ metagenomes from 359 independent herds.

the phenotypic level. We found evidence for 407 different genes across all pig and poultry samples (Supplementary Table 2).

We calculated the dissimilarities between the gene-level resistomes of all samples and visualized it in a dendrogram (Fig. 2). There was a perfect host separation, with all pig samples clustering separately from all poultry, suggesting that pig and poultry resistomes are very distinct. In the pig cluster, the country separation was more pronounced than in the poultry cluster. An exception was Danish poultry, where 18 out of 20 farm resistomes clustered.

To assess the reproducibility of our protocol, from sampling through to sequencing, we evaluated the similarities between the resistomes of two triple-sampled swine herds. The Dutch triple-sampled herd had the highest similarities between composite samples, ranging from 93.6% to 93.7% Bray–Curtis similarity. The Belgian triple-sampled herd pools had values ranging from 91.5% to 93.3% similarity. No replicated sample pool had a higher similarity to other herds than to its own replicates, and the two sets of three samples can therefore be seen clustering separately, indicating reproducibility in both sampling and sequencing (Fig. 2). A farm resistome similarity heatmap is included in the Supplementary Material (Supplementary Fig. 3).

We ordinated the gene-level resistomes for all samples (Supplementary Fig. 4) and pig and poultry samples separately (Supplementary Fig. 5a,b). As with hierarchical clustering, there was a clear separation of pig and poultry samples, along the first principal coordinate, which explained 48% of the variation across all resistomes.

When analysing the two species separately, we observed clustering according to the country of origin in pigs (Supplementary Fig. 5a), whereas clustering was more diffuse for poultry (Supplementary Fig. 5b). We tested for the country effect and found it to be significant in both pigs (adonis2, $P < 0.001$) and poultry (adonis2, $P < 0.001$). However, in poultry, the country effect only explained 24% of the variation, whereas the country effect explained 41% of the variation in pigs. In the pig resistome ordination, the Danish and Dutch samples clustered closely together. The same could be seen for the French and Belgian resistomes and to a lesser degree, the Italian and Spanish samples. Bulgaria, Germany and Poland showed larger dispersions than the other countries. Beta-dispersion levels varied significantly between countries in both pigs (beta-dispersion $P < 0.001$; Supplementary Fig. 5c) and poultry (beta-dispersion $P < 0.001$; Supplementary Fig. 5d).

We visualized the AMR gene abundances in a heatmap to look at the overall structure and composition of the resistomes and the co-occurrence of AMR genes (Supplementary Fig. 6). Some AMR genes were more abundant in one species, whereas others, including *tet(W)* and *erm(B)*, were ubiquitous in all samples for both species. Among the pig samples, the Italian samples stood out: several chloramphenicol AMR genes, including *cat(pC194)*, *catP* and *cat_2*, were much more abundant in Italy than in the other countries, consistent with our inspection of AMR at the class level (Fig. 1). Several AMR genes known to be co-located indeed co-occurred across samples. The genes in the vancomycin AMR *VanA* cassette were co-located in several poultry samples. This was also true for the *VanB* cassette members, clustering together but separately from *VanA*, showing an ability to distinguish variants of homologous genes. As indicated earlier, the poultry samples showed less country-based clustering than pigs. An exception was the Danish poultry samples; these had a noticeably lower abundance of many AMR genes that were widespread in other countries.

Core resistome. To determine whether specific genes were unique to each of the host animals, we examined the set of AMR genes that was consistently observed within each animal species (evidence for it in 95% of samples). We identified 33 core AMR genes in pigs and 49 core AMR genes in poultry, with 24 being shared between the two hosts (Supplementary Fig. 7). Hence, only nine AMR genes were pig-core genes without also being poultry-core genes. These included the genes making up the Van-G vancomycin cassette, *tet(C)*, bla_{ACI} and *cfxA*. Twenty-five AMR genes were poultry-core genes without also being pig-core genes and include the Enterobacteriaceae-associated *strAB*, *sul2*, *blaTEM* and *tet(A)* genes.

Differential abundance analysis. To test which specific genes differed in abundance between countries, we carried out a differential abundance analysis for ResFinder gene cluster read counts. Heavy overrepresentation of low unadjusted *P* values indicated a large effect of country in both the pig and the poultry data sets (Supplementary Fig. 8). Of special interest was the newly characterized *Enterococcus*-associated linezolid-resistance gene *optrA*, which had a significantly higher abundance in Bulgarian poultry farms than in poultry farms in all other countries (false discovery rate (FDR) < 0.05) (Fig. 3). However, a single Spanish farm did have an even higher *optrA* abundance than any other farm. Among the pig herds, the *optrA* gene was more abundant in Bulgarian and Italian herds than anywhere else (except for two farms in Spain) (FDR < 0.05).

The newly identified colistin-resistance gene *mcr-1* was significantly more abundant in Bulgarian and Italian poultry farms than in most other countries (FDR < 0.05). France, Poland and Spain had intermediate levels, whereas Denmark, the Netherlands and Germany had the lowest levels (Fig. 3). The Bulgarian poultry farms enrolled in this study did not report any polymyxin usage, whereas Italian farmers reported the highest average treatment incidents.

As previously noted from visual inspection of heatmaps, multiple chloramphenicol AMR genes including *cat(pC194)* were much more abundant in Italian pigs than in other pigs. The extended-spectrum β-lactamase $bla_{CTX\text{-}M}$ gene cluster 1 also showed country dependency, being significantly more abundant in poultry samples from Spain, Poland, Italy, France and Belgium than in poultry samples from Germany (FDR < 0.05). Differential abundance analysis results can be found in Supplementary Tables 3 and 4 for pig and poultry, respectively.

Alpha diversity and richness. We calculated several alpha-diversity indexes for each farm resistome (Fig. 4 and Supplementary Table 5). The range of AMR diversity was generally much larger for poultry samples, having both lower and higher diversity, than for pig samples, which had a tighter spread of diversity. The poultry samples had a higher estimated number of different AMR genes (that is, a higher Chao1-estimated richness).

Interestingly, countries with higher estimates of unique AMR genes in pigs also tended to have a high AMR richness in poultry (Spearman's rho: 0.88, $P = 0.02$; Supplementary Fig. 9). Spain, Italy, Bulgaria and Poland had the highest estimated number of unique AMR genes in both pig and poultry. There was no such association for Pielou's evenness or Simpson diversity ($P > 0.05$). Rarefaction curves for pig and poultry resistomes can be found in Supplementary Fig. 10.

Association between the bacteriome and the resistome. To test the degree to which the bacterial composition of the microbiota dictates the resistomes, Procrustes analyses were performed. We found that for both pig and poultry, the resistome correlated significantly with the bacterial composition ($P = 0.001$; Fig. 5). Thus, samples with similar taxonomic compositions tended to have similar resistome compositions. In addition, most of the between-country differences in resistomes seem to be explained by systematic between-country differences in bacteriomes.

The correlation between AMR and taxonomy was similar in pigs (correlation: 0.86) and poultry (correlation: 0.88). Interestingly, in the pig samples, we saw a country effect on the strength of association between the bacteriome and the resistome. In the Dutch and Spanish pig herds, ordinations based on bacterial genera and AMR genes gave similar results (Fig. 5b). In German farms, in particular, the resistome and bacteriome ordinations yielded more dissimilar results. This was less evident for poultry, although a single Danish poultry herd had a very unusual resistome, considering its taxonomic composition (Fig. 5d).

AMR and drug use association. We found that the total country-level veterinary AMU from the European Medicines Agency's European Surveillance of Veterinary Antimicrobial Consumption (ESVAC) was positively associated with AMR in both pigs and poultry. The AMR abundance increased by 1,736–3,507 FPKM (95% CI, $\beta = 2{,}621$) in pigs when the AMU increased by 1 $\log_e$ unit (a 36.8% increase in AMU) (Fig. 6a) and to a lesser degree in poultry, where the AMR abundance increased by 68–1,330 FPKM (95% CI, $\beta = 700$) when the AMU increased by 1 $\log_e$ unit (Fig. 6b). For pigs, the variance between farms within-country was seven-times larger than the variance between countries in general, whereas in poultry, the variance was four-times larger within-country than between countries.

We repeated the regression with the treatment incidents data for the farms (Supplementary Fig. 11). Interestingly, these data were less associated with the AMR load than the national veterinary drug use data. The association remained for pigs, whereas poultry was not significant ($P > 0.05$). Bulgaria had low reported usage in both livestock species, whereas AMR was high. The Danish farms reported a higher average AMU in poultry than Bulgaria and the Netherlands, but the total AMR was far lower.

To test whether the AMU pattern across multiple antimicrobial classes was associated with AMR gene profiles, we compared the AMR gene cluster abundances for pig and poultry against both the ESVAC and the farm treatment incidents data (Supplementary Tables 6–8 and Supplementary Fig. 12). Using Procrustes analyses, all matrix–matrix correlations were significant ($P = 0.001$), although with low symmetric correlation coefficients (correlation: 0.34–0.45). As for the regression analysis, there was a better fit between pig AMR and the ESVAC data than between pig AMR and the farm treatment incidents data.

Functional AMR genes. In addition to using ResFinder, we also ran most analyses with the functional resistance database (FRD) to elucidate whether the functionally determined AMR genes behave similarly to the acquired AMR genes in ResFinder. If the FRD genes serve similar AMR functionality as the acquired ResFinder genes, we would expect similar results.

Using the FRD, we found both similar and different patterns than using ResFinder. There was still a perfect separation between pig and poultry samples, but the country separation in pigs was less distinct than when using ResFinder (Supplementary Fig. 13). Although less variation could be explained by two axes, the principal coordinates analysis (PCoA) plot of pig samples now clustered German and Spanish samples, with the remaining countries being more similar. The resistome richness showed similar patterns to ResFinder: Spanish, Italian, Polish and Bulgarian samples had a higher estimated richness in both pig and poultry than the other countries. The Procrustes correlation between the resistome and drug usage was lower (0.40 for pig and 0.25 for poultry). This result was echoed by the lack of association between the total AMR and the total AMU, for both pig and poultry ($P > 0.05$; Supplementary Fig. 14).

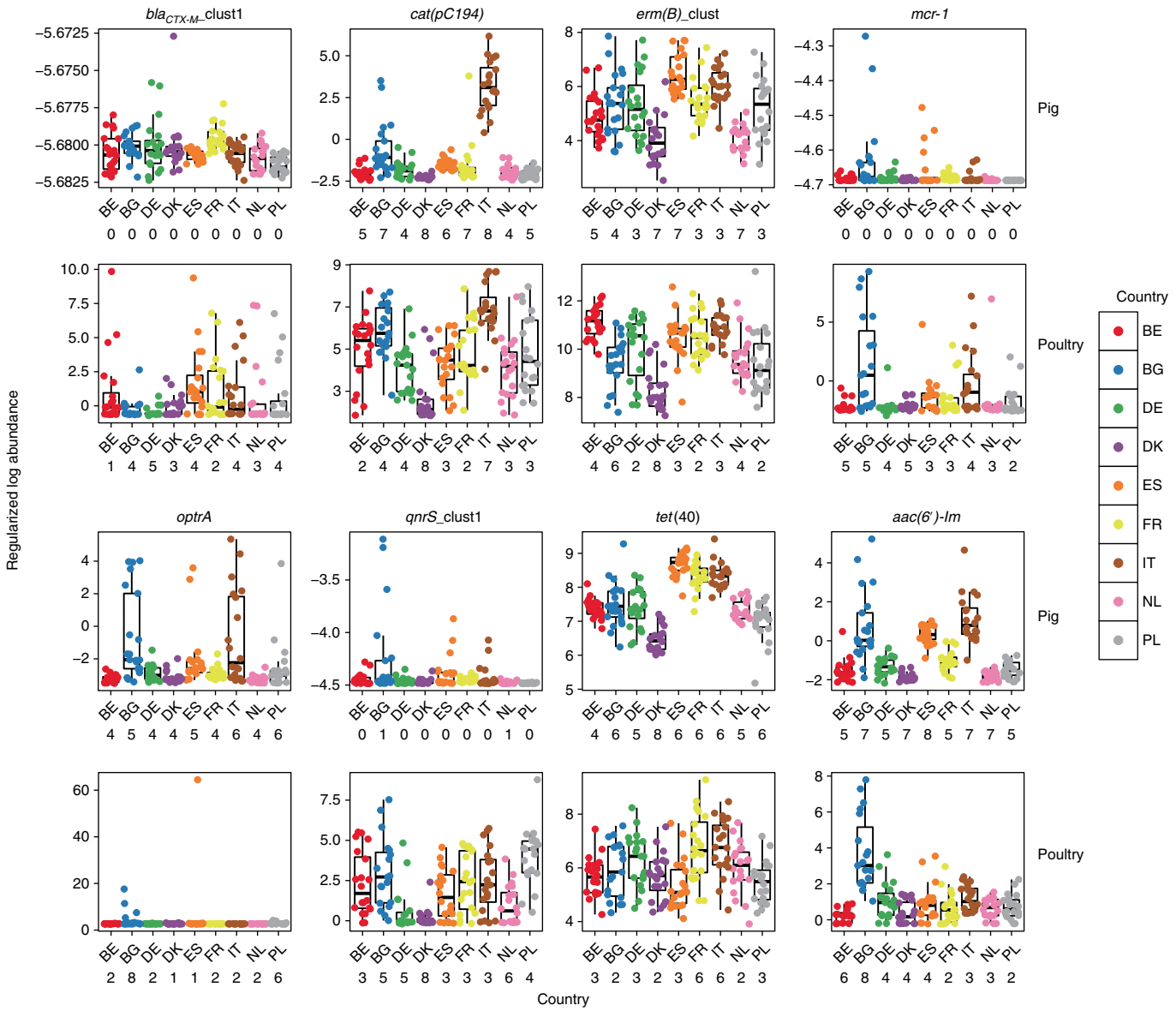

Fig. 3 | AMR genes differ in abundance between countries. A handpicked subset of genes that differed significantly in abundance between at least two countries in either pig or poultry farms is shown. The regularized log abundance (rlog) is shown on the *y* axis in box plots and points. Points were sideways jittered to reduce overplotting. The numbers along the *x* axis denote the number of countries with a significantly different mean abundance (DESeq2 Wald test, two-sided, $\alpha = 0.05$). Testing was done for all ResFinder genes, and *P* values were adjusted for multiple testing (FDR) for all country comparisons. The horizontal box lines represent the first quartile, the median and the third quartile. Whiskers denote the range of points within the first quartile − 1.5× the interquartile and the third quartile + 1.5× the interquartile. $n = 359$ metagenomes from independent herds. Twenty metagenomes per livestock species per country were used, with the exception of Bulgarian pigs (21), Bulgarian poultry (19) and German poultry (19).

Discussion

Using a metagenomic shotgun sequencing strategy, we were able to detect and quantify >400 AMR genes across 181 pig and 178 poultry herds in 9 European countries.

A recent study including Chinese, Danish and French pigs showed that the Chinese pig resistomes clustered separately, whereas the Danish and French pig resistomes overlapped[16]. Here, we demonstrate that even among European countries, the livestock resistomes differ in a country-specific manner that might be explained by differential AMU so that countries with similarly high and diverse AMU (Spain and Italy) have similar resistomes, the same way as countries with similarly low AMU (Denmark and the Netherlands) also have similar pig resistomes.

We found that within-country resistome dispersion is country dependent, particularly in pigs, with Bulgarian, German and Polish pig herds having more dispersed AMR. Although we cannot currently explain this, we consider the possible causes as differences in trade and management, among others.

We found the recently discovered plasmid-borne colistin resistance gene *mcr-1* in numerous poultry herds, especially in Bulgaria, Spain and Italy. Spain and Italy had the highest reported veterinary colistin usage among the surveyed countries, whereas Bulgaria has a low reported usage, which is uncharacteristic for the high *mcr-1* level found here[13]. This gene was recently discovered in China and identified throughout the world and has been identified in pigs, poultry and human clinical infections alike[18].

A newly characterized enterococcal linezolid-resistance gene, *optrA*, was detected in a subset of pig samples, with Bulgaria, Italy and Spain having the highest abundances. The *optrA* gene provides AMR to both oxazolidinone and amphenicols, including the

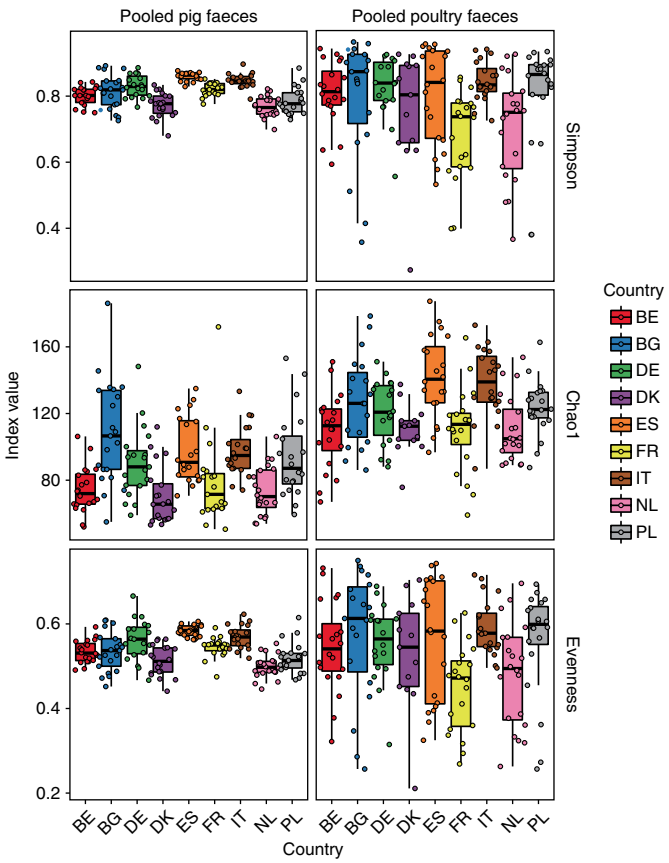

Fig. 4 | Resistome alpha diversity and richness differ between animal host and countries. From the read count pair matrix, several indexes were calculated: Simpson diversity index, Chao1-estimated richness and Pielou's evenness. The horizontal box lines represent the first quartile, the median and the third quartile. Whiskers denote the range of points within the first quartile − 1.5× the interquartile range and the third quartile + 1.5× the interquartile range. $n = 359$ metagenomes from independent herds. Twenty metagenomes per livestock species per country were used, with the exception of Bulgarian pigs (21), Bulgarian poultry (19) and German poultry (19).

veterinary-used florfenicol[13,19]. The high abundance of this gene in these countries can probably be explained by the fact that they have the highest veterinary amphenicol usage among the nine countries investigated. This explanation fits well with the fact that Bulgaria, Italy and Spain also had the highest abundances of chloramphenicol AMR genes, such as *cat(pC194)*, in poultry.

Another AMR gene of special interest, the $bla_{\text{CTX-M}}$, was also observed in the poultry herds. The higher abundance of $bla_{CTX\text{-}M}$ cluster 1 in Spain, Italy, Poland and Belgium could possibly be explained by co-selection by fluoroquinolones, which is used more in Spain, Poland, Italy and Belgium than in other sampled countries. *qnr* and $bla_{\text{CTX-M}}$ genes are frequently co-located on large extended-spectrum β-lactamase plasmids. Veterinary cephalosporin usage did not seem to explain the observed levels.

Poland and Spain use far more veterinary fluoroquinolones than other countries included in this study. We found that plasmid-mediated quinolone AMR (*qnr* genes) was frequently abundant in Polish, but not in Spanish, poultry. In Bulgaria, quinolone AMR was also frequently observed, although their reported AMU did not follow the same trend.

Interestingly, we observed that the number of unique AMR genes predicted (Chao1) significantly correlated between pig and poultry farms across countries. In addition, countries with a high estimated number of unique AMR genes also have a high AMR abundance (Italy, Spain, Bulgaria and Poland). The fact that countries' AMR abundance and the predicted number of unique AMR genes in pig and poultry tend to follow each other, could be explained with policy: if a country has strict AMU regulations in one livestock species, the chances are that similar regulations are in place for other livestock species. Indeed, the treatment incidents data showed that countries with higher AMU in pigs, had higher AMU in poultry. It might also be speculated that an ecological country effect plays a role; for example, the total country AMU might influence AMR abundance in all reservoirs. This, might explain why the ESVAC data correlate better with the observed AMR than the treatment incidents data. Better AMU data, at the herd and country level and over time, are needed to further explore the specific AMU–AMR associations. It has previously been reported that the composition of the bacterial community structures the resistome[20,21]. We found the same to be true for pig and poultry resistomes; in addition, we showed that the taxa–AMR association strength differs between countries. Horizontal gene transfer could explain this phenomenon, if a larger proportion of certain countries' resistome is mobile and AMR genes are more frequently introduced and re-introduced to genera. Conversely, vertical AMR transmission can also play a role, if, for example, one country's livestock is more isolated from trade. As we found that a large part of the observed resistome is dictated by the taxonomic composition, we expect much of the country resistome differences to be explained by systematic differences in feed and management.

In contrast to ResFinder, when using FRD, we found no relationship between the ESVAC total drug use and the total functional AMR abundance. This suggests that, although many genes can provide AMR when cloned into, for example, *E. coli*, in functional metagenomic assays, they might not provide AMR functionality in their natural hosts with natural expression levels. If most of them did, we could expect to see antimicrobial-based selection and an association to AMU, as it is observed for the AMR genes in ResFinder. This finding echoes previous sentiments that one should carefully consider the risk to human health imposed by functionally determined AMR genes[22]. Some FRD genes might represent high-risk AMR genes, but we currently do not know what subset that is. Creating the FRD is a first step in trying to catalogue the many AMR genes found in functional metagenomic studies. Screening sequenced pathogenic isolates and metagenomic assemblies for FRD genes would be a good start for assessing their host range and risk potential.

The AMU data used in this study are not optimal. There is variation in drug use within each country's farms that we did not account for by using the available country-wide averages per drug class. Moreover, the population correction unit (PCU) denominator used by the ESVAC may vary greatly between countries, and no independent validation of the data reported by the national competent authorities have been performed. Furthermore, the integrated herds enrolled in this study might represent only a limited subset of the overall livestock production in some countries. However, even with the crude ESVAC-based total veterinary AMU, we found significant associations with the total AMR abundance. The similar conclusion when considering the specific drug usage profile of each country indicates that the resistome is responding to AMU. The AMR–AMU association is well documented for specific cultured indicator species and certain antimicrobial drugs, but is relatively unknown when considering the whole microbiota and resistome and the newer approach of metagenomic shotgun sequencing[3,8]. We do not know why the pig samples had a larger within-country spread of total AMR, but perhaps the more heterogeneous production system and production management are responsible. Curiously, the treatment incidents data, which are specific to the sampled farms, was less associated with the resistomes and total AMR than the ESVAC

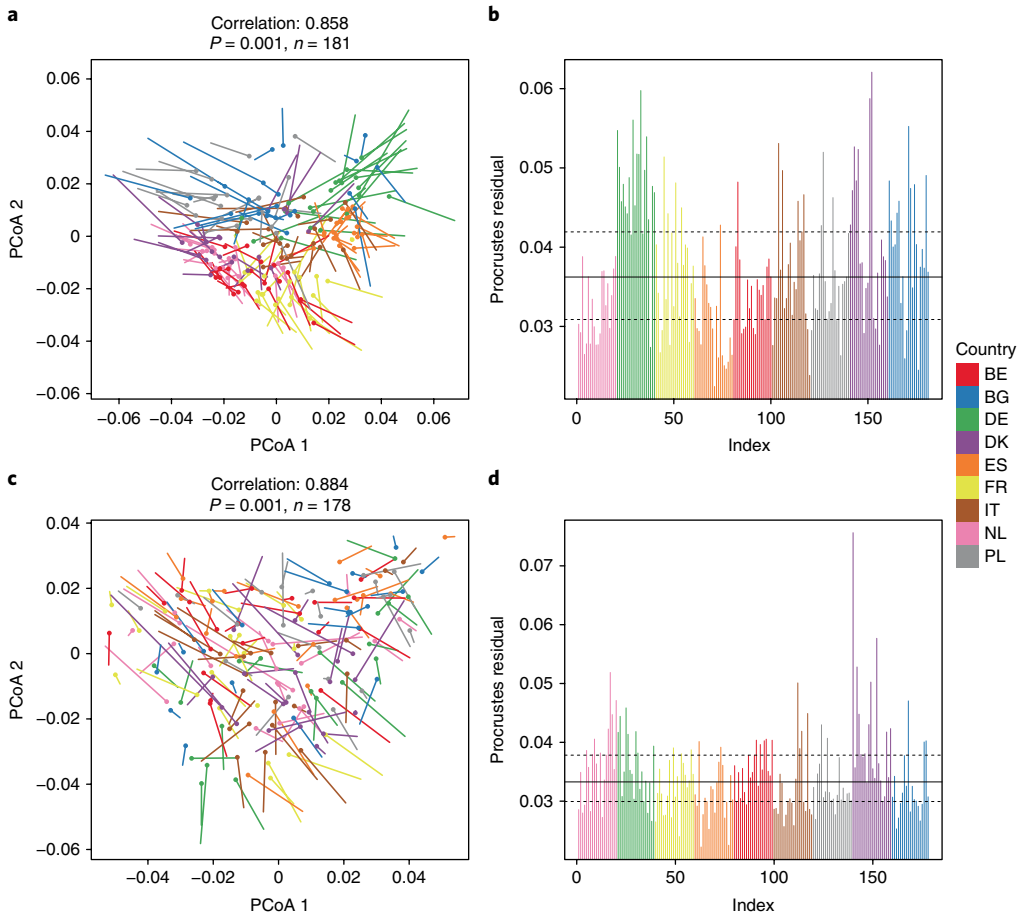

Fig. 5 | Taxonomic variation explains resistome variation. Bacterial and AMR abundance profiles were correlated with each other using Procrustes analyses for pig (**a**) and poultry (**c**) herds, thus comparing the two multivariate data sets. The lines show the Procrustes residuals; the change in the ordination position when using the resistome (dotted ends) compared to the bacteriome (non-dotted ends) is displayed. The correlation coefficients and significance were derived using the protest function in vegan. The residual line plot for pig (**b**) and poultry (**d**) farms enables easier residual size comparison, showing the difference in the bacteriome–resistome association between farms. Horizontal lines denote the median (solid), 25% and 75% quantiles (dashed). $n = 359$ metagenomes from independent herds. Twenty metagenomes per livestock species per country were used, with the exception of Bulgarian pigs (21), Bulgarian poultry (19) and German poultry (19).

data. Instead of reporting that AMU does not affect AMR after all, we think that it is worthwhile considering whether there are some AMU reporting biases between countries.

DNA extractions from the pooled poultry samples resulted in relatively low DNA yields. The protocol used was optimized for pig faeces, human faeces and sewage, but not for poultry faeces[23]. The lower yields necessitated the use of a PCR-based library preparation kit that can influence downstream analysis of shotgun sequencing[24]. Although the large difference between pig and poultry resistomes in our study is probably real, we caution the use of sensitive, quantitative analyses when comparing between samples prepared using different library preparation kits. For this reason, we have mostly tested within each reservoir.

The sensitivity of metagenomic approaches does not yet rival phenotypic alternatives such as selective enrichment. There are AMR genes in important pathogens that we know are probably present but are below our detection limit. For example, we only found evidence for $bla_{\text{CTX-M}}$ in three of the pig herds, whereas in phenotypic studies, the prevalence is high even among farms with no cephalosporin usage[25].

The primary concern with read-mapping techniques is the lack of genomic context, which can be solved using metagenomic assembly and binning approaches[16,26,27]. In this way, AMR alleles in full length, their genomic context and their associated taxa have been identified in both pig, poultry and human faecal samples[28]. As shown previously, the association between AMR and AMU is similar for metagenomics and traditional phenotypic methods, but several aspects make metagenomics an intriguing monitoring tool[17]. The fact that both types of analyses (quantitative, sensitive read mapping and qualitative, context-giving binning) use the same raw data makes metagenomics an attractive tool. In addition, the digital nature of sequence data would also allow future re-use and form the basis of an invaluable historical archive, potentially usable for both AMR and pathogen-tracking worldwide.

We found that the metagenomic resistome varied significantly between the pig and poultry reservoirs, but also within each species, in a country-dependent manner. Within each country, we found different levels of variation, with some countries having more homogenous herds than others. Differences were seen both in the total AMR abundance, but also the abundances of AMR types and specific genes, including clinically relevant AMR genes. Some of this variation we attributed to differential drug usage between the countries. We also identified the microbiome background as an important factor in determining the resistome in livestock, but found that the strength of the association was country dependent, at least in pigs. Interestingly, we found that the AMR richness in one livestock

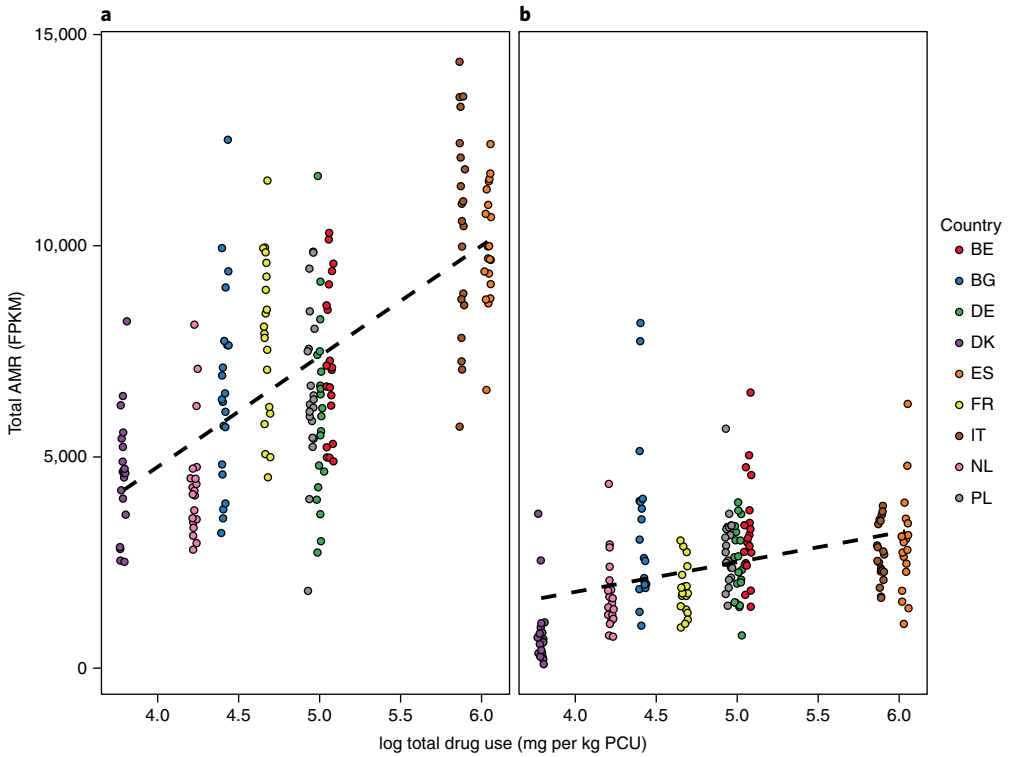

Fig. 6 | National veterinary AMU is associated with total metagenomic AMR. Scatter plots of the average total veterinary AMU (ESVAC) and the pooled sample total AMR. A slight sideways jitter was added to the points to minimize overplotting. **a**, The association between the average veterinary drug use and the total AMR load in pig farms. $n=181$ independent herds. **b**, The association between the average veterinary drug use and the total AMR load in poultry farms. $n=178$ independent herds. See the main text for a description of the trend lines.

species in a country is linked to the abundance in another livestock species. Finally, we observed some indications that newly described AMR genes from functionally metagenomic studies might not provide AMR functionality when expressed in their natural host, even though they have the potential at the right expression levels in the right organism.

Methods

Farm selection and sampling. The sampling protocol for pig and broiler farms that has been agreed on by the EFFORT consortium is described below. The selection of farms and the sampling procedure followed these guidelines to the extent possible, but some deviations from the protocol were occasionally necessary. The selection and sampling goals are described below, whereas a detailed description of the sampling conducted in the individual countries and exceptions is provided in the Supplementary Material.

Selection of pig and poultry farms. In each participating country, 20 conventional integrated pig farrow-to-finisher non-mixed farms were selected. The farms needed to have a minimum of 150 sows and 600 fatteners and use batch production to ensure that most of the animals of the sampled group originated from the same birth cohort. All-in all-out production at compartment level was preferred, and all fatteners sampled were required to have been on the same site during their entire life. Selected farms should have no contact through livestock trade and should have a random regional distribution.

In each country, 20 conventional broiler farms (no breeders) were selected. The farms had all-in all-out production, with a thinning procedure from day 30 onwards allowed. All selected farms should have no intended slaughter age higher than 50 days, no slow-growing breeds (intended growth rate of less than 55 g per day) and no stocking density lower than 10 birds per m^2. Only one flock per house per holding should be sampled and each flock should be between 20,000 and 40,000 birds. If possible, the selected farms should have a random regional distribution.

Procedure for sampling. We sampled pig farms between May 2014 and December 2015 and tried to minimize seasonal influences. The sampled fatteners were as close to slaughter as possible (that is, within the last week). A total of 25 fresh, still-warm and undisturbed faecal droppings were sampled from pen floors (a minimum of 10 g faeces per sample) randomly divided over all eligible compartments or stables of fatteners close to slaughter.

Broilers were sampled between May 2014 and June 2016 and we tried to minimize seasonal influences. On each farm, 25 undisturbed, fresh main bowel droppings were collected from the floor of the house (a minimum of 3 g faeces per sample). The flocks were sampled as close to slaughter as possible (the last week before the final depopulation).

All samples were collected aseptically in plastic containers and were stored at 4 °C and transported to the laboratory within 24 hours after sampling.

Pooling and handling of samples. Upon laboratory arrival, individual faecal samples were homogenized by stirring thoroughly with a sterile tongue depressor or spoon for a few minutes. From each pig sample, two 2-ml cryotubes were filled and frozen immediately at −80 °C (alternatively at −20 °C for a maximum of 4 days, before transferring to −80 °C). For broiler samples, two cryotubes were prepared with at least 0.5 g faeces each. Sample pooling was either done immediately or the frozen tubes were shipped to the Technical University of Denmark (DTU) on dry ice for pooling. Individual samples from the same herd were defrosted and placed on ice briefly before weighing. Following weighing, they were pooled with 0.5 g faeces from each sample and stirred for a few minutes with a sterile device (for example, a disposable wooden tongue depressor). All samples were only thawed once shortly before DNA extraction.

After the removal of two mislabelled samples, composite samples from a total of 178 broiler flocks and 181 pig herds remained.

Sampling to estimate the effect of random sampling. To study the potential effect of sampling randomness and the reproducibility of our sampling protocol, a Belgian and a Dutch pig herd were chosen for triplicate sampling. These two herds were sampled three times on the same day (25 samples × 3 sampling rounds), resulting in 6 pooled samples (2 herds × 3 sampling rounds), from which the within-farm resistome variation was assessed. A table with all the samples and the associated data is included as Supplementary Table 9.

DNA extraction and sequencing. From each of the pooled, herd-level faecal samples, DNA was extracted using a modified QIAamp Fast DNA Stool Mini Kit protocol (51604, Qiagen), as previously described[23]. One major modification is the addition of a bead-beating step at the beginning of DNA extraction. The protocol can be found at figshare.com/articles/SOP_-_DNA_Isolation_QIAamp_Fast_DNA_Stool_Modified/3475406. DNA purification of all the pooled samples was

processed centrally at the DTU, and the DNA was stored in duplicates at −20 °C until further use.

DNA was shipped on dry ice for library preparation and sequencing at the Oklahoma Medical Research Foundation (OMRF; Oklahoma City, OK, USA). At the OMRF, DNA from all samples was mechanically sheared using ultrasonication to a targeted fragment size of 300 bp (Covaris E220evolution). For pooled pig samples, library preparation was performed with the NEXTflex PCR-Free library preparation kit (Bioo Scientific). For poultry samples, owing to a lower DNA availability, the minimal amplification-based KAPA Hyper kit (Kapa Biosystems) was used. For all samples, the Bioo NEXTflex-96 adapter set (Bioo Scientific) was used. In batches of roughly 60 samples, the libraries were multiplexed and sequenced on the HiSeq3000 platform (Illumina), using 2 × 150-bp paired-end sequencing per flow cell. A total of 17 Belgian, Danish and Dutch pig faecal samples were sequenced on the HiSeq2500 platform (Illumina), using 2 × 100-bp paired-end sequencing before it became unavailable at the OMRF (see Supplementary Table 9).

In total, DNA from 365 pooled samples was extracted and shotgun sequenced, resulting in >36 billion sequences (18 billion paired-end reads), comprising > 5,000 Gb of DNA. The sequencing yielded an average of 50 million (s.d.: 18×10^6) paired-end reads per pooled sample. This was similar for pig and poultry samples, although the sampling depth was more varied in pig samples.

Bioinformatics processing. The DNA sequences (FASTQ reads) from each sample were analysed following the principles from the previously described MGmapper tool[15]. To avoid PCR copies in the poultry data, identical read pairs were removed using 'MarkDuplicates' from the Picard software (v2.8.3; broadinstitute.github.io/picard). Adaptor sequences were removed using BBduk (BBMap software)[29]. Sequences from phiX174, which is an internal sequencing control, were removed using the BWA-MEM algorithm[30]. Trimmed read pairs were aligned using the BWA-MEM algorithm (Burrows-Wheeler aligner) to the prokaryotic RefSeq genomes from the NCBI GenBank with 'reference' and 'representative' tags (downloaded on 18 November 2016). The BWA-MEM algorithm (v0.7.15) normally estimates the insert size individually per computer CPU core. We used a per-sample estimate to increase the robustness of the estimated insert sizes and, therefore, the acceptable mapping distances for read pairs.

The read pairs were aligned to the prokaryotic genomes again and to the AMR genes present in the ResFinder database (accessed 17 November 2016) using the robust insert size estimates[31]. ResFinder is a manually curated database of acquired AMR genes and, therefore, does not include intrinsic AMR genes and mutated housekeeping genes.

Properly paired read pairs, with at least a 50-bp alignment in each read were accepted. ResFinder mapping counts were adjusted for differences in both gene lengths and bacterial sequence abundances by computing FPKM values for each ResFinder reference sequence[32]. Raw mapping count data and their associated FPKM values can be found in Supplementary Tables 2 and 10.

Genes with many alleles in ResFinder result in unspecific mapping and randomly assigned read pairs. To avoid sensitivity loss and wrong assignments, we kept ambiguous hits, but aggregated their abundances to higher levels, corresponding to 90% gene identity clusters. To determine these clusters, we used CD-HIT-EST (v4.6.6) at a 90% identity level and otherwise default settings[33]. The resulting gene clusters were manually inspected and named to reflect their gene members (Supplementary Table 11). In addition to this 'gene cluster' level, we summed the FPKMs to resistance phenotype levels, as annotated in the ResFinder database.

FRD. Previous studies have identified a wide array of AMR genes in various reservoirs using functional metagenomics, referred to as functional AMR genes[20,34–36]. By cloning random DNA fragments from complex microbiomes into an expression vector expressed in a host (typically *E. coli*) and selecting for growth in the presence of certain antibiotics, they have been found to provide AMR to many antibiotics[20,34–36]. We constructed a FRD from 3,416 AMR gene variants identified in four major studies using 23 different antimicrobials for selection[20,34–36].

Briefly, in each of these studies, DNA was extracted from environmental and human faecal samples, fragmented and cloned into a plasmid vector and screened for AMR functionality in *E. coli* cultured with one of several antimicrobials. Cloned fragments in plasmids that were found to confer AMR were sequenced and the AMR genes were identified. The protocol for the database construction can be found at cge.cbs.dtu.dk/services/ResFinderFG. Genes were quantified using MGmapper, as was done for ResFinder. Genes with >90% identity to ResFinder genes were removed post-mapping to obtain the set of FRD genes that was absent from ResFinder. The reference gene abundances were summed to 90% gene clusters, using CD-HIT-EST, as was done for ResFinder[33]. The most frequent gene clusters remaining were derived from genes selected using trimethoprim, chloramphenicol, co-trimoxazole, cycloserine, amoxicillin, gentamicin, penicillin and tetracycline.

PCoA and resistome clustering. For PCoA, the gene cluster-level FPKM matrix was Hellinger transformed and the Bray–Curtis dissimilarities between all samples were calculated using the R package vegan[37]. PCoA was carried out for both pigs and poultry, combined and separately, using the vegan function 'betadisper'. The same analysis was used to test whether host animal and country were significant predictors of within-group dispersion. The effects of country on sample dissimilarities were determined using 'permutational multivariate analysis of variance using distance matrices' (the 'adonis2' function in the vegan package), separately for pig and poultry.

AMU in livestock. Data for the national livestock AMU were obtained from the ESVAC report and were stratified by major drug family[12]. The mass of active compound sold for use in animals in 2014 was divided by the PCU in 10^6 kg, approximating the biomass. The PCU is a unit that allows inter-species integration by adjusting for import/export and differences in the average weight between species when they are most likely to receive antimicrobial treatment. The estimate was multiplied by 1,000 to obtain drug per mg per PCU livestock. The country-specific veterinary drug use can be found in Supplementary Table 6 and Supplementary Fig. 15.

In addition to the national veterinary AMU, we obtained data from collaborating researchers on the average treatment incidents in the sampled farms, stratified by antimicrobial class, country and livestock species. The treatment incidence was calculated as the antimicrobial dose per defined daily animal doses (DDDvet) per 1,000 animals at risk, adjusting for 200-day and 40-day production cycles for pigs and poultry, respectively[38]. These average AMU values for pigs (S. Sarrazin et al., manuscript in preparation) and poultry (P. Joosten et al., manuscript in preparation), stratified by drug group, are presented in Supplementary Tables 7 and 8, respectively. Data are visualized in Supplementary Figs. 16 and 17.

Procrustes analyses. To determine the effect of the underlying microbiota on the resistome, we used Procrustes analysis. The gene cluster FPKM ResFinder matrix and the genus-level FPKM taxonomy matrix were Hellinger transformed and Bray–Curtis dissimilarities were calculated. Each dissimilarity matrix was ordinated using PCoA. The symmetric Procrustes correlation coefficients between the bacteriome and the resistome ordinations, *P* values and plots were obtained using the 'protest' function in vegan[39].

To test the association between AMU patterns and the resistomes, we also used Procrustes analysis as follows. A PCoA was generated from Euclidean distances between the samples in the AMU data. The AMU PCoA configuration was tested against the AMR gene cluster PCoA configuration using the 'protest' function with the default 999 permutations. This was done separately for pig and poultry samples.

Alpha diversity. For all samples, we computed the within-herd resistome diversity using the Simpson's Index of Diversity (1-D), the Chao1 richness estimate and Pielou's evenness[40]. The gene cluster count matrix was rarified to 10,000 hits per sample for alpha diversity estimation, leading to the exclusion of samples with fewer hits to the AMR database.

Visualization. Heatmaps were produced using the pheatmap R package. For heatmaps showing individual-gene abundances, the Bray–Curtis dissimilarities between samples were used for hierarchical clustering. For all other dendrograms, the Pearson product-moment correlation coefficients were used. Complete-linkage clustering was used for all hierarchical sample clustering. For sample similarities, Bray–Curtis dissimilarity was converted to a similarity percentage, that is, 100 × (1 − Bray–Curtis).

The circular Bray–Curtis resistome dendrogram was constructed by exporting the dendrogram in Newick format using the ape package and further annotating it using the Interactive Tree of Life tool[41,42]. Bar plots, box plots and scatter plots were produced using the ggplot2 R library[43]. The R library RcolorBrewer was used to generate the colour palettes used for the figures. This library is based on work by C. A. Brewer (www.ColorBrewer.org).

Statistical analyses. All statistics were done in Microsoft R Open 3.3.2, using the libraries and the procedures detailed below. The exact package versions can be found here: mran.revolutionanalytics.com/snapshot/2016-11-01/bin/windows/contrib/3.3. For statistical tests, only samples from the first visit to the triple-sampled herds were included (see Supplementary Table 9), meaning $n_{pig} = 181$ and $n_{poultry} = 178$. Unless otherwise mentioned, all statistical analyses were performed on pigs and poultry separately.

Association between AMU and AMR. To test the effect of the total AMU on the total metagenomic AMR abundance (the sum of all genes), we used the lme4 1.1-12 package to make linear mixed-effects regression models, with the total livestock drug usage (the sum of the ESVAC PCUs) as the independent variable, the total AMR abundance (the sum of FPKM) as the dependent variable and country as a mixed-effect intercept[44]. The total AMU was log transformed, which resulted in lower Akaike's information criteria. Pig sample residuals and country residuals showed normality and so did the poultry country residuals. Poultry sample residuals had a longer right tail, but square-root transformation of the poultry AMR data gave more-normal residuals and a similar conclusion ($P < 0.05$). The effect and significance of drug usage were assessed using likelihood-ratio tests,

comparing the random-effect models with and without the AMU effect. The ResFinder-treatment incidents and FRD-PCU tests were done in the same way as the ResFinder-PCU tests.

Differential abundance analysis. To identify AMR genes that differ in abundance between countries, we analysed the gene cluster count matrix using the DESeq2 package as previously recommended for metagenomic read count data[45,46]. This was done on the full count matrix, based on recommendations that rarefying is not warranted in metagenomic studies[46]. The read-pair count matrices for pigs and poultry were analysed separately. The number of mapped bacterial pairs was divided by the minimum number of mapped bacterial pairs and was used as the size factor. For each gene, we used a two-sided Wald test to determine whether the fold change between countries differed from zero and extracted all the country-versus-country results. *P* values were adjusted for the FDR using the Benjamini–Hochberg approach and we used a significance threshold of alpha: 0.05 (ref. [47]).

Core resistome. To determine the set of AMR genes consistently found within each livestock species, we used a soft threshold. AMR gene clusters with mapping read pairs in at least 95% of samples from a livestock species were considered part of the core resistome.

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability. Drug use data are attached as Supplementary Tables. The DNA sequences (reads) from the 363 metagenomic samples from the 359 herds are deposited in the European Nucleotide Archive under the project accession number PRJEB22062.

Received: 20 September 2017; Accepted: 8 June 2018;
Published online: 23 July 2018

References

1. *Antimicrobial Resistance Global Report on Surveillance 2014* (WHO, 2014).
2. Aarestrup, F. M. The livestock reservoir for antimicrobial resistance: a personal view on changing patterns of risks, effects of interventions and the way forward. *Phil. Trans. R. Soc. B* **370**, 20140085 (2015).
3. Chantziaras, I., Boyen, F., Callens, B. & Dewulf, J. Correlation between veterinary antimicrobial use and antimicrobial resistance in food-producing animals: a report on seven countries. *J. Antimicrob. Chemother.* **69**, 827–834 (2014).
4. Dunlop, R. H. et al. Associations among antimicrobial drug treatments and antimicrobial resistance of fecal *Escherichia coli* of swine on 34 farrow-to-finish farms in Ontario, Canada. *Prev. Vet. Med.* **34**, 283–305 (1998).
5. Aarestrup, F. M., Seyfarth, A. M., Emborg, H., Pedersen, K. & Bager, F. Effect of abolishment of the use of antimicrobial agents for growth promotion on occurrence of antimicrobial resistance in fecal enterococci from food animals in Denmark. *Antimicrob. Agents Chemother.* **45**, 2054–2059 (2001).
6. Agersø, Y. & Aarestrup, F. M. Voluntary ban on cephalosporin use in Danish pig production has effectively reduced extended-spectrum cephalosporinase-producing *Escherichia coli* in slaughter pigs. *J. Antimicrob. Chemother.* **68**, 569–572 (2013).
7. Dutil, L. et al. Ceftiofur resistance in *Salmonella enterica* serovar Heidelberg from chicken meat and humans, Canada. *Emerg. Infect. Dis.* **16**, 48–54 (2010).
8. Dorado-García, A. et al. Quantitative assessment of antimicrobial resistance in livestock during the course of a nationwide antimicrobial use reduction in the Netherlands. *J. Antimicrob. Chemother.* **71**, 3607–3619 (2016).
9. Aarestrup, F. M. et al. Resistance to antimicrobial agents used for animal therapy in pathogenic-, zoonotic- and indicator bacteria isolated from different food animals in Denmark: a baseline study for the Danish Integrated Antimicrobial Resistance Monitoring Programme (DANMAP). *APMIS* **106**, 745–770 (1998).
10. European Food Safety Authority Harmonised monitoring of antimicrobial resistance in *Salmonella* and *Campylobacter* isolates from food animals in the European Union. *Clin. Microbiol. Infect.* ***14***, 522–533 (2008).
11. McEwan, S. A., Aarestrup, F. M. & Jordan, D. in *Antimicrobial Resistance in Bacteria of Animal Origin* (ed. Aarestrup, F. M.) 397–413 (ASM Press, Washington DC, 2006).
12. Veterinary Medicines Division *European Medicines Agency, European Surveillance of Veterinary Antimicrobial Consumption, 2016* (EMA, 2016).
13. European Centre for Disease Prevention and Control, European Food Safety Authority & European Medicines Agency ECDC/EFSA/EMA first joint report on the integrated analysis of the consumption of antimicrobial agents and occurrence of antimicrobial resistance in bacteria from humans and food-producing animals. *EFSA J.* **13**, 4006 (2015).
14. European Food Safety Authority Technical specifications on the harmonised monitoring and reporting of antimicrobial resistance in *Salmonella*, *Campylobacter* and indicator *Escherichia coli* and *Enterococcus* spp. bacteria transmitted through food. *EFSA J.* **10**, 64 (2012).
15. Nordahl Petersen, T. et al. Meta-genomic analysis of toilet waste from long distance flights; a step towards global surveillance of infectious diseases and antimicrobial resistance. *Sci. Rep.* **5**, 11444 (2015).
16. Xiao, L. et al. A reference gene catalogue of the pig gut microbiome. *Nat. Microbiol.* **1**, 16161 (2016).
17. Munk, P. et al. A sampling and metagenomic sequencing-based methodology for monitoring antimicrobial resistance in swine herds. *J. Antimicrob. Chemother.* **72**, 385–392 (2016).
18. Hasman, H. et al. Detection of *mcr-1* encoding plasmid-mediated colistin-resistant *Escherichia coli* isolates from human bloodstream infection and imported chicken meat, Denmark 2015. *Euro Suveill.* **20**, 30085 (2015).
19. Wang, Y. et al. A novel gene, *optrA*, that confers transferable resistance to oxazolidinones and phenicols and its presence in *Enterococcus faecalis* and *Enterococcus faecium* of human and animal origin. *J. Antimicrob. Chemother.* **70**, 2182–2190 (2015).
20. Forsberg, K. J. et al. Bacterial phylogeny structures soil resistomes across habitats. *Nature* **509**, 612–616 (2014).
21. Pehrsson, E. C. et al. Interconnected microbiomes and resistomes in low-income human habitats. *Nature* **533**, 212–216 (2016).
22. Martínez, J. L., Coque, T. M. & Baquero, F. What is a resistance gene? Ranking risk in resistomes. *Nat. Rev. Microbiol.* **13**, 116–123 (2014).
23. Knudsen, B. E. et al. Impact of sample type and DNA isolation procedure on genomic inference of microbiome composition. *mSystems* **1**, e00095-16 (2016).
24. Bowers, R. M. et al. Impact of library preparation protocols and template quantity on the metagenomic reconstruction of a mock microbial community. *BMC Genomics* **16**, 856 (2015).
25. Hammerum, A. M. et al. Characterization of extended-spectrum β-lactamase (ESBL)-producing *Escherichia coli* obtained from Danish pigs, pig farmers and their families from farms with high or no consumption of third- or fourth-generation cephalosporins. *J. Antimicrob. Chemother.* **69**, 2650–2657 (2014).
26. Nielsen, H. B. et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nat. Biotechnol.* **32**, 822–828 (2014).
27. Anantharaman, K. et al. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat. Commun.* **7**, 13219 (2016).
28. Ma, L. et al. Metagenomic assembly reveals hosts of antibiotic resistance genes and the shared resistome in pig, chicken, and human feces. *Environ. Sci. Technol.* **50**, 420–427 (2016).
29. BBduk v39.92 (Bushnell, B., 2018); https://sourceforge.net/projects/bbmap/
30. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
31. Zankari, E. et al. Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* **67**, 2640–2644 (2012).
32. Trapnell, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).
33. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
34. Pehrsson, E. C., Forsberg, K. J., Gibson, M. K., Ahmadi, S. & Dantas, G. Novel resistance functions uncovered using functional metagenomic investigations of resistance reservoirs. *Front. Microbiol.* **4**, 145 (2013).
35. Moore, A. M. et al. Pediatric fecal microbiota harbor diverse and novel antibiotic resistance genes. *PLoS ONE* **8**, e78822 (2013).
36. Sommer, M. O. A., Dantas, G. & Church, G. M. Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science* **325**, 1128–1131 (2009).
37. Legendre, P. & Gallagher, E. D. Ecologically meaningful transformations for ordination of species data. *Oecologia* **129**, 271–280 (2001).
38. Timmerman, T. et al. Quantification and evaluation of antimicrobial drug use in group treatments for fattening pigs in Belgium. *Prev. Vet. Med.* **74**, 251–263 (2006).
39. Oksanen, J. et al. vegan: Community Ecology Package v2.4-1 (CRAN, 2016); https://CRAN.R-project.org/package=vegan
40. Pielou, E. C. The measurement of diversity in different types of biological collections. *J. Theor. Biol.* **13**, 131–144 (1966).
41. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).
42. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127–128 (2007).
43. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, New York, NY, 2009).
44. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
45. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
46. McMurdie, P. J. & Holmes, S. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput. Biol.* **10**, e1003531 (2014).

47. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).

Acknowledgements
We thank all of the anonymous pig and poultry herd owners who agreed to participate in the study and especially everyone involved in sampling and laboratory work: M. Schlepers (Belgium); T. Ivanova, N. Cholakov, E. Gurova-Mehmedova and K. Penchev (Bulgaria); F. Nienhaus (Germany); C. L. Nielsen, P. Ryt-Hansen, B. Hølstad, B. Rasmussen and K. Nielsen (Denmark); C. Jenna, D. Virginie, E. Florent, E. Eric, S. Le Bouquin, L. Denis and T. Rodolphe (France); M. Gherpelli, M. Pegoraro and V. Carfora (Italy); D. de Vries (the Netherlands); B. Gawlik, D. Krasucka and A. Hoszowski (Poland). The EFFORT project (www.effort-against-amr.eu) and the work presented here is supported by the European Union, FP7-KBBE-2013-7, grant agreement 613754.

Author contributions
F.M.A., D.H., J.A.W., T.H., D.M. and the EFFORT group designed the study. F.M.A. and B.E.K. detailed the sampling and sequencing. A.S.R.D. and the EFFORT group carried out the sampling. B.E.K. and S.J.P. conducted the DNA purification and, with P.M., organized the sequencing. R.B.H., O.Lund. and T.N.P. created the read-mapping pipeline. E.R. created the FRD. P.M., O.Lukjacenko., R.E.C.L., L.V.G., L.A.M.S., H.S., A.B., A.D.G. and H.V. carried out the bioinformatics and statistical analysis. P.M. created the figures and drafted the manuscript. All authors helped review, edit and complete the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41564-018-0192-9.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to F.M.A.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

EFFORT Group

Haitske Graveland[2], Alieda van Essen[4], Bruno Gonzalez-Zorn[8], Gabriel Moyano[8], Pascal Sanders[9], Claire Chauvin[9], Julie David[9], Antonio Battisti[10], Andrea Caprioli[10], Jeroen Dewulf[11], Thomas Blaha[12], Katharina Wadepohl[12], Maximiliane Brandt[12], Dariusz Wasyl[13], Magdalena Skarzyńska[13], Magdalena Zajac[13], Hristo Daskalov[14], Helmut W. Saatkamp[4] and Katharina D. C. Stärk[15]

[8]Department of Animal Health and Health Surveillance Center (VISAVET), Complutense University of Madrid, Madrid, Spain. [9]Fougeres Laboratory, French Agency for Food, Environmental and Occupational Health & Safety, Fougères, France. [10]Istituto Zooprofilattico Sperimentale del Lazio e della Toscana, Department of General Diagnostics, National Reference Laboratory for Antimicrobial Resistance, Rome, Italy. [11]Department of Reproduction, Obstetrics and Herd Health, Ghent University, Merelbeke, Belgium. [12]University of Veterinary Medicine Hannover, Bakum, Germany. [13]National Veterinary Research Institute, Pulawy, Poland. [14]National Diagnostic Research Veterinary Institute, Sofia, Bulgaria. [15]SAFOSO AG, Liebefeld, Switzerland.

nature research

Corresponding author(s): Frank M. Aarestrup

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

   Describe how sample size was determined.

   For each host species, 25 fecal samples per farm were taken on 20 farms in 9 countries to ensure a sufficiently precise estimate of the resistome at the herd level and to obtain sufficient power to describe associations between determinants or modifying factors and AMR that are present across countries in a meta-analysis.

2. Data exclusions

   Describe any data exclusions.

   Two metagenomic datasets were excluded from all analysis. We had no pre-defined exclusion criteria, but during data exploration, it became obvious that two samples were mis-labeled and originated from from the wrong animal species. PCoA revealed a pig-labeled sample clustering with poultry and a poultry-labeled sample clustering with pigs. Follow-up investigation confirmed the labeling mistake. Since DNA library preparation was performed differently for the two species, we opted to exclude the data, rather than re-annotate it.

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   No replication was performed in the classical sense, but two pig herds were sampled three times each (2 farms x 3 sampling rounds x 25 fecal samples) to verify our procedure. Resistomes of the triple-sampled herds were similar and clustered together as expected.

4. Randomization

   Describe how samples/organisms/participants were allocated into experimental groups.

   Fecal samples were grouped according to the host animal species (pig/poultry) and country of origin (one of nine European countries).

5. Blinding

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   Since the "groups" in question were host animal (pig/poultry) and country, it was not feasible to blind samplers to either. No attempts were made at blinding during DNA extraction. Library preparation and shotgun sequencing was performed by OMRF, who were blind to "country" but not "host species", since different protocols were used for the two species, following discussion between us and OMRF. No blinding was attempted during data analysis.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☒	☐ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

Microsoft R Open 3.3.2 (R distribution with library version control and multithreading support)
RStudio 1.0.136 (GUI for R)
DESeq2 1.14.1 (differential abundance analysis R package)
lme4 1.1-12 (Linear, generalized linear, and nonlinear mixed models)
vegan 2.4.1 (Community Ecology R Package)
CD-HIT 4.4.6 (clustering of nucleotides based on sequence identity)
BBmap/BBduk2 39.92 (adapter- and quality-trimming of DNA read data)
BWA-MEM v0.7.15 (alignment of DNA reads to reference sequences)
Picard 2.8.3 (Removal of duplicate reads which can during PCR amplification)

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

All NGS data used in the study has been made publicly available on ENA (30th of May, 2018).

From some samples, all fecal material has been exhausted. The remaining faecal samples were collected for this study only and can not be used for other studies. For some samples with remaining material, fecal matter could be made available for the purpose of verification.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

No antibodies were used.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No eukaryotic cell lines were used.

b. Describe the method of cell line authentication used.

No eukaryotic cell lines were used.

c. Report whether the cell lines were tested for mycoplasma contamination.

No eukaryotic cell lines were used.

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

No commonly misidentified eukaryotic cell lines were used.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

Observational study of livestock animals in farms.
Pigs (Sus scrofa): Mean age of 179 days (min: 95 max: 320 days). Mixed sexes.
Poultry (Gallus gallus): Mean age of 34 days (min: 16, max: 54).

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

The study did not involve human research participants.