

Technical University of Denmark



## A sampling and metagenomic sequencing-based methodology for monitoring antimicrobial resistance in swine herds

**Munk, Patrick; Dalhoff Andersen, Vibe; de Knegt, Leonardo; Jensen, Marie Stengaard; Knudsen, Berith Elkær; Lukjancenko, Oksana; Mordhorst, Hanne; Clasen, Julie; Agersø, Yvonne; Folkesson, Anders; Pamp, Sünje Johanna; Vigre, Håkan; Aarestrup, Frank Møller**

*Published in:*  
Journal of Antimicrobial Chemotherapy

*Link to article, DOI:*  
[10.1093/jac/dkw415](https://doi.org/10.1093/jac/dkw415)

*Publication date:*  
2016

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Munk, P., Dalhoff Andersen, V., de Knegt, L., Jensen, M. S., Knudsen, B. E., Lukjancenko, O., ... Aarestrup, F. M. (2016). A sampling and metagenomic sequencing-based methodology for monitoring antimicrobial resistance in swine herds. *Journal of Antimicrobial Chemotherapy*, 72, 385-392. [dkw415]. DOI: 10.1093/jac/dkw415

**DTU Library**  
Technical Information Center of Denmark

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

## A sampling and metagenomic sequencing-based methodology for monitoring antimicrobial resistance in swine herds

Patrick Munk<sup>1</sup>, Vibe Dalhoff Andersen<sup>1</sup>, Leonardo de Knecht<sup>1</sup>, Marie Stengaard Jensen<sup>1</sup>, Berith Elkær Knudsen<sup>1</sup>, Oksana Lukjancenko<sup>1</sup>, Hanne Mordhorst<sup>1</sup>, Julie Clasen<sup>2</sup>, Yvonne Agersø<sup>1</sup>, Anders Folkesson<sup>2</sup>, Sünje Johanna Pamp<sup>1</sup>, Håkan Vigre<sup>1</sup> and Frank Møller Aarestrup<sup>1\*</sup>

<sup>1</sup>Research Group for Genomic Epidemiology, National Food Institute, Technical University of Denmark, Søtofts Plads, Building 221, 2800 Kgs Lyngby, Denmark; <sup>2</sup>Section for Bacteriology and Pathology, National Veterinary Institute, Technical University of Denmark, Bülowssvej 27, 1870 Frederiksberg C, Denmark

\*Corresponding author. Tel: +45-35-88-62-81; Fax: +45-35-88-63-41; E-mail: fmaa@food.dtu.dk

Received 27 April 2016; returned 29 June 2016; revised 28 August 2016; accepted 31 August 2016

**Objectives:** Reliable methods for monitoring antimicrobial resistance (AMR) in livestock and other reservoirs are essential to understand the trends, transmission and importance of agricultural resistance. Quantification of AMR is mostly done using culture-based techniques, but metagenomic read mapping shows promise for quantitative resistance monitoring.

**Methods:** We evaluated the ability of: (i) MIC determination for *Escherichia coli*; (ii) cfu counting of *E. coli*; (iii) cfu counting of aerobic bacteria; and (iv) metagenomic shotgun sequencing to predict expected tetracycline resistance based on known antimicrobial consumption in 10 Danish integrated slaughter pig herds. In addition, we evaluated whether fresh or manure floor samples constitute suitable proxies for intestinal sampling, using cfu counting, qPCR and metagenomic shotgun sequencing.

**Results:** Metagenomic read-mapping outperformed cultivation-based techniques in terms of predicting expected tetracycline resistance based on antimicrobial consumption. Our metagenomic approach had sufficient resolution to detect antimicrobial-induced changes to individual resistance gene abundances. Pen floor manure samples were found to represent rectal samples well when analysed using metagenomics, as they contain the same DNA with the exception of a few contaminating taxa that proliferate in the extraintestinal environment.

**Conclusions:** We present a workflow, from sampling to interpretation, showing how resistance monitoring can be carried out in swine herds using a metagenomic approach. We propose metagenomic sequencing should be part of routine livestock resistance monitoring programmes and potentially of integrated One Health monitoring in all reservoirs.

### Introduction

Antimicrobial resistance (AMR) is considered one of the greatest threats to human health.<sup>1</sup> The resistance epidemic has been attributed to the use of antimicrobials in clinical settings, and in livestock.<sup>2</sup> In addition to the abundance of resistance in animal reservoirs, animals are also important in the early evolution and emergence of novel resistance genes.<sup>3</sup>

Many research studies have shown that the use of antimicrobials in livestock will lead to increased occurrence of AMR.<sup>4–6</sup> In addition, research shows that reducing the use of antimicrobials can decrease the occurrence of resistance.<sup>7,8</sup> In order to identify

areas of priority and document effects of interventions, a robust monitoring system is essential. The first integrated routine monitoring of AMR was established in Denmark in 1995.<sup>9</sup> Since its inception, other nations have developed similar monitoring programmes that embrace the One Health approach.<sup>10,11</sup>

Current monitoring efforts are mainly based on culturing indicator bacteria and determining their resistance phenotypically.<sup>12</sup> The procedures for sampling and analysis are time consuming, expensive and only target a limited number of the bacteria and resistance mechanisms found in the different reservoirs.

Traditionally, the proportion of resistant bacteria has been assessed through cfu counting, with and without a selective

agent, and used as a measure of resistance.<sup>6</sup> Alternatively, phenotypic MIC determination provides high accuracy, reproducibility and the isolate-specific context to monitor multiresistance. Both methods are, however, blind to the cause of resistance and the epidemiology of the responsible genes. To alleviate this, isolated bacteria can be screened with PCR using primers targeting the suspected causal resistance genes.<sup>13,14</sup> More recently, WGS of indicator organisms has become economically feasible and pipelines for *in silico* resistance determination have become accurate enough to compete with phenotypic resistance determination and PCR.<sup>15</sup> For monitoring purposes, however, a prohibitively large number of isolates needs to be analysed to obtain reliable prevalence estimates for resistance.

Indicator-organism-centric approaches only provide insight into a minor fraction of naturally occurring bacteria in which the bulk of resistance genes may be present.<sup>16</sup> Metagenomic read mapping has recently been used to quantify resistance genes directly in multiple reservoirs, including waste water, humans, livestock, drinking water, activated sludge and aircraft septic tanks.<sup>17–21</sup> The approach has furthermore proved useful in quantifying the selection pressure induced by ciprofloxacin in infants.<sup>22</sup> Using metagenomics, one can quantify thousands of targets in a sample without requiring *a priori* knowledge of which genes or bacteria are present. A single dataset of DNA sequences can therefore be used to quantify all known resistance genes.

The aim of our study is to compare methods for measuring the occurrence of AMR in swine herds. In particular, we focus on sampling and analytic procedures, and we test whether pen floor sampling provides a valid approximation to sampling individual pigs. We compare metagenomics with traditional culture-based methods for quantifying AMR in swine herds. The study was carried out in 10 herds.

## Materials and methods

### Choice of sample herds

In Denmark, detailed records are kept on all antimicrobial prescriptions for livestock herds. The farmer obtains drug prescriptions from veterinarians and the information is recorded in the VETSTAT database.<sup>23</sup> Another database, the Danish Central Husbandry Register (CHR), collects mandatory information from all livestock herds including ownership, animal counts, animal species and age groups (e.g. sows/piglets, weaners, finishers). Using these databases, we generated a list of herds fitting our study population criteria: conventional Danish integrated pig herds with more than 500 sows and producing at least 5000 slaughter pigs annually. For these herds, we calculated the defined animal daily doses (ADD) for antimicrobial drug groups and adjusted for the CHR-based herd sizes.<sup>24</sup> To include herds covering a wide spectrum of antimicrobial consumption, we invited herd owners from the top and bottom 10% tetracycline consumption quantiles to participate in the study, until five herds within each quantile had accepted. For information on which active compounds make up each drug class category, see Table S1 (available as Supplementary data at JAC Online).

### Sampling and pooling

For an overview of sampling, pooling and downstream analysis, see Figure S1. Briefly, from each of the 10 herds, herd-level floor (HL-F) samples were obtained from 30 pens and paired pen-level floor (PL-F) and pig (PL-P) samples were obtained from four selected pens. Within

each pen, the sampled pigs were randomly chosen. For details, see the Supplementary data.

### Cultivation of faecal bacteria

For each pooled sample, 1 g of faeces was suspended in 9 mL of isotonic saline and serially diluted. Aliquots of 100  $\mu$ L dilutions were plated onto selective and non-selective LB and MacConkey agar plates to quantify aerobic bacteria and *Escherichia coli*, respectively. Selective plates contained 8 mg/L tetracycline (T3383 tetracycline hydrochloride, Sigma–Aldrich) in LB and 16 mg/L tetracycline or ampicillin (A9393 Ampicillin, Sigma–Aldrich) in MacConkey. All assays were performed in triplicate and plates were incubated at 37°C overnight. For each triplicate set, a weighted average resistance proportion was calculated. For details, see the Supplementary data.

### MIC

We determined MIC values for a subset of *E. coli* from the cultivation experiments using the previously described DKMVN4 panel.<sup>25</sup> For details, see the Supplementary data.

### DNA extraction

For DNA extraction, a modified QIAamp Fast DNA Stool Mini Kit protocol was employed (Qiagen, Valencia, CA, USA). Cryotubes with pooled faeces were gently thawed on ice. Prior to the protocol, 0.2 g of sample was mixed with 1 mL of InhibitEX buffer in a Lysing Matrix A tube (MP Biomedicals). Samples were treated with a TissueLyser (3  $\times$  30 s, 30 Hz) and were chilled on ice between repetitions. Following bead beating, samples were heated to 95°C for 7 min and centrifuged to eliminate stool particles. During the extraction protocol, volumes of sample, proteinase K, buffer AL and ethanol were doubled. DNA was eluted in 100  $\mu$ L of elution buffer.

### Shotgun DNA sequencing

PCR-free DNA libraries were generated and sequenced on the HiSeq2500 (Illumina) to generate roughly 7 gigabases of paired-end reads per sample, enough to get 20 $\times$  coverage of bacteria with 1% abundance.<sup>26</sup> Cutadapt was used to trim the reads to a mean Phred score of 20.<sup>27</sup> For details, see the Supplementary data.

### Read mapping (resistance)

Resistance was quantified using the ResFinder database, a collection of 2130 annotated resistance genes.<sup>15</sup> Paired-end mapping was done with a BWA-mem-based script as described elsewhere.<sup>20,28</sup> To avoid favouring poorer-quality samples (those that on average get shorter reads post-trimming), we required each read to align with a static 50 bp instead of the default setting that is relative to the trimmed read length (80%).

Because the ResFinder database contains many highly identical sequences, unspecific mapping occurs when reads map to identical parts of homologous gene variants. Read counts from variants of the same gene were aggregated to gene levels according to common gene names.<sup>20</sup> For the variant-to-common name conversion table and the final abundance matrix, see Tables S2 and S3, respectively.

### Read mapping (bacteria)

For the bacterial count data, we employed the EBI Metagenomics resource.<sup>29,30</sup> Briefly, the pipeline identifies DNA reads matching the bacterial 16S gene. It then uses QIIME to map those to the Greengenes 16S database and to identify the lowest common ancestor of each 16S sequence.<sup>31,32</sup> The taxonomic abundance matrices can be downloaded from EBI Metagenomics and are associated with the accession number (ERP013942).

### Metagenomic data analysis

A resistance gene count matrix was constructed from the DNA read mapping output, with genes (2130) and samples (20) constituting rows and columns, respectively (Table S3). Values  $>0$  were divided by 2 so each count reflected a single mapping fragment. Read counts were aggregated according to the common gene names specified in Table S2, resulting in the final abundance matrix. To avoid inappropriate statistical models and losing power by rarefying, metagenomic data such as RNAseq data, should be analysed using zero-inflated negative binomial models as suggested by McMurdie and Holmes.<sup>33</sup> Therefore, gene-level differential abundance analysis was carried out using the R-package DESeq2.<sup>34</sup> For details on differential abundance analysis, see the Supplementary data.

### Expected resistance

The expected tetracycline resistance was derived by ranking the farms from 1 to 10 in terms of VETSTAT-derived, herd-size-adjusted tetracycline consumption 1 year prior to sampling. Ranks based on resistance measurements can thus be compared with the expected resistance, based on the assumption that tetracycline resistance occurrence follows tetracycline consumption.

### qPCR

Primers and probes have been presented previously by Schmidt *et al.*<sup>35</sup> The inclusion criteria for the AMR genes were: (i) the use of the antimicrobial class linked to the AMR gene in question in the Danish pig production; (ii) the occurrence of the gene in a wide bacterial population; and (iii) the possibility of designing a qPCR assay for the chosen genes utilizing the same temperature profile.<sup>35</sup> Quantification of the four AMR genes [*erm*(B), *sul*1, *tet*(B) and *tet*(W)] was done using the high-capacity qPCR chip '192.24 Gene Expression IFC' (Fluidigm®) using two technical replicates as previously described by Clasen *et al.*<sup>36</sup> The efficiency of the primers was determined using standard curves and obtained results were normalized with 16S ribosomal DNA (reference gene). Negative  $\Delta\Delta Cq$  values, indicating no observed gene, were treated as missing values. Subsequently,  $\Delta\Delta Cq$  values were unlogged ( $2^{-\Delta\Delta Cq}$ ) to directly compare with alternative resistance measurements.

### Comparison of metagenomics, phenotypic methods and expected resistance

The four resistance estimates, as well as the expected tetracycline resistance were correlated to each other (Spearman rank correlation). Using a rank-based statistic, we avoid making assumptions about the shape of any relationship. Since we found that floor and rectal samples were comparable (see the Results and discussion section), both pig (PL-P) and floor (PL-F/HL-F) isolates were used for the MIC estimate, as too few isolates were obtained from the individual HL-F samples. The remaining methods were based on the HL-F samples alone. All herd-level tetracycline resistance measurements can be found in Table S4. The bootstrapped hierarchical dendrogram was produced with the R package 'pvclust', using 10000 bootstrap iterations (with default other settings), and shows node certainty percent based on the approximately unbiased *P* value (AUP).<sup>37</sup>

### Accession numbers

The DNA sequences (reads) from the 20 metagenomic samples are deposited in the European Nucleotide Archive (ENA) under the project accession number ERP013942.

### Visualization

The R package RcolorBrewer was used to generate the colour palettes used in several figures. The package is based on work by Cynthia A. Brewer (<http://www.ColorBrewer.org>).

### Results and discussion

Ten Danish swine herds with historical high and low antimicrobial consumption were enrolled in the study. To make sure their usage was also varied in the period leading up to the sampling, we recalculated their consumption for a 1 year window prior to the individual sampling dates. As expected from our selection, a large variation in overall drug use was observed between the herds (Figure 1a). This was also the case for most individual drug classes that were not considered in the initial selection of the study herds (Figure 1b and Table S5). By dosage, tetracycline was the most used drug class, followed by  $\beta$ -lactams and macrolides. This pattern among the study herds is consistent with Denmark's antimicrobial usage in all swine the same year.<sup>38</sup>

### Comparison of sampling strategies

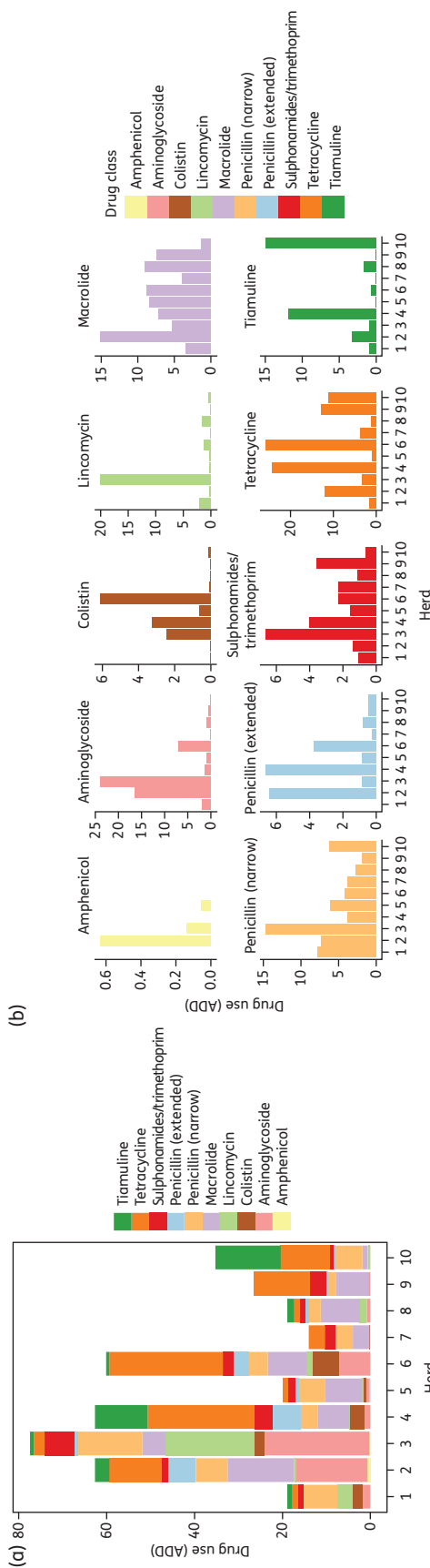
The pig intestine is an important source of zoonotic bacteria contaminating meat during slaughter. In this study, we obtained three kinds of samples. Those were: (i) rectal pig faecal samples; (ii) undisturbed pen floor faecal samples; and (iii) pen floor manure samples. Rectal samples and manure samples are the most and least time consuming to collect, respectively. For an overview of sampling and pooling strategy, see Figure S1.

To determine whether sampling undisturbed faeces on the pen floor is an acceptable proxy for sampling recta, we compared paired PL-F and PL-P faeces pools from four pens per herd, for a total of 40 pairs. When using cfu counting, we found that tetracycline resistance correlated significantly (Spearman rank correlation,  $P < 0.05$ ) between aerobic bacteria in floor and rectal samples (Figure S2a). We also found significant relationships for tetracycline-resistant (Figure S2b) and ampicillin-resistant *E. coli* (Figure S2c).

To confirm that pig and fresh floor faecal samples were comparable, we attempted to assess the association using a different method (qPCR) and genes representing more resistance types. We quantified *erm*(B), *sul*1, *tet*(B) and *tet*(W) in the same 40 paired sets of PL-F and PL-P samples. The *erm*(B) and *tet*(W) abundances correlated well between floor and pig samples ( $P < 2.2 \times 10^{-16}$ ) (Figure S2d and e). The genes *tet*(B) and *sul*1 were below the detection limit in several samples, yielding too few pairs for comparison. The qPCR abundances for the four genes are included in Figure S3.

While the earlier described PL-F samples were derived from fresh, undisturbed droppings, the herd-level floor (HL-F) samples originated from floor manure of unknown age, which is readily available and requires no animal handling. We sequenced DNA from the 20 herd-level samples, for a total of 690 million pairs (1380 million reads), corresponding to 139 billion bp (Table S6). This yielded an average of 7 billion bp/sample, the recommended amount to get 20x coverage on bacteria with a 1% relative abundance.<sup>26</sup> We then used those DNA sequences to quantify known resistance genes.

Principal component analysis (PCA) showed herd-wise clustering, suggesting that herd is a stronger resistance predictor than



**Figure 1.** Antimicrobial consumption in the study herds. Information about antimicrobial drug usage, prescribed in a 1 year window prior to sampling, was extracted from the VETSTAT database. For each combination of herd and drug type, ADD were calculated. (a) Stacked bar chart showing total antimicrobial consumption per herd (ADD). (b) Multiple bar charts showing ADD per herd stratified by antimicrobial group.

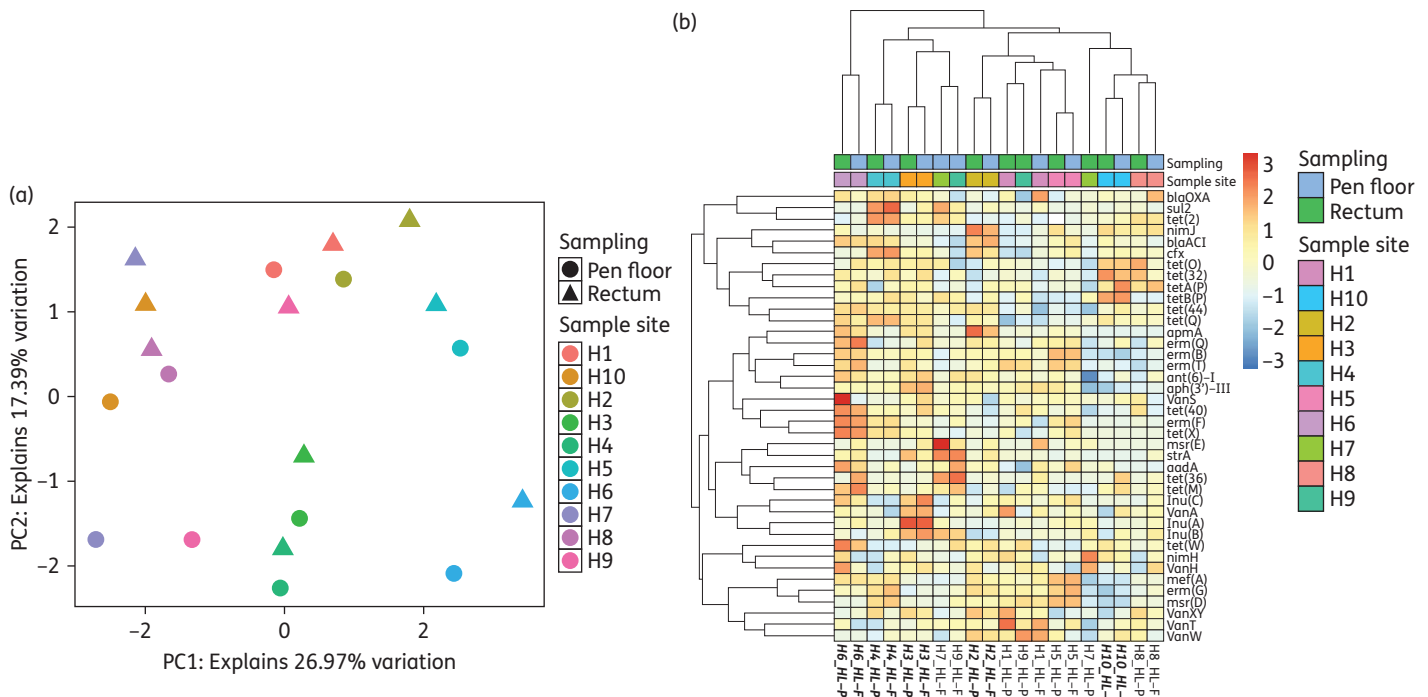
sampling location (floor versus pig) (Figure 2a). This was confirmed by hierarchical clustering based on the most common genes, where we found 7 of the 10 sample pairs clustered according to herd of origin (Figure 2b). Despite the high herd-wise similarity, we noticed a consistent pen floor shift along the second principal component. To investigate this systematic change in the resistome, we tested for differential abundance between the paired HL-P and HL-F samples. Only two resistance genes differed significantly between the sample types: *tet(M)* and *tet(36)*, both overrepresented in floor manure samples (false discovery rate-adjusted *P* value,  $FDR < 0.1$ ). *tet(M)* has the broadest host range of all tetracycline resistance genes, being found in both Gram-positive and Gram-negative bacteria.<sup>39</sup> *tet(36)*, on the other hand, is quite limited and primarily found in *Bacteroides coprosuis*, where it was first isolated from swine manure.<sup>40</sup> Consistent with this, the gene was not detected in any pig samples, whereas most manure samples contained the gene at low abundance. As shown by Forsberg et al.,<sup>41</sup> different samples' resistomes are structured according to their microbial taxonomic composition. In order to check if the change in resistance was explained by a taxonomic shift, we performed differential abundance analysis on bacterial genera quantified in the same datasets. Five genera were differentially abundant ( $FDR < 0.1$ ): *Succinivibrio*, *Fusobacterium*, *Rummeliibacillus*, *Acinetobacter* and *Ignatzschineria* (Figure S4). Like for resistance genes, all significant genera were overrepresented in the floor samples. Both *Fusobacterium* and *Acinetobacter* isolates are known to harbour *tet(M)* genes.<sup>39,42</sup> *Bacteroides* was not differentially abundant between the sample types, but this genus also spans many species. Concordant with finding *tet(36)* only in floor samples, *B. coprosuis*-specific reads were also only identified in the floor samples (data not shown, based on EBI Metagenomics).

Despite these systematic differences between the sample types, the samples clustered according to herd of origin in 9 out of 10 cases based on bacterial composition (Figure S4).

Together these findings suggest the increased resistance in floor samples is explained by an expansion of certain bacterial taxa, either through contamination or proliferation. In summary, both fresh floor droppings and manure samples represent the intestinal resistome well and can provide acceptable proxies for sampling individual pigs. Pig pen floors have earlier been sampled for risk assessment, but until recently very little was known about how well these samples represent the intestines. Multiple sampling strategies have recently been evaluated for pig herd resistance monitoring with a qPCR quantification protocol. Consistent with our results, the authors found pen floor sampling appropriate for monitoring intestinal resistance.<sup>35</sup>

### Comparison of resistance monitoring methods

In order to benchmark the techniques, we assumed antimicrobial use increases AMR—a frequently made observation.<sup>5,6,18</sup> The tetracycline ADD were used to rank the expected tetracycline resistance levels in the 10 herds relative to one another. We then used: (i) cfu counting of aerobic bacteria; (ii) cfu counting of *E. coli*; (iii) MIC determination for *E. coli*; and (iv) metagenomic read mapping to quantify the overall tetracycline resistance in the same 10 HL-F samples. We compared the methods to each other and to the expected resistance using correlation analysis and hierarchical clustering. The two phenotypic resistance measurements



**Figure 2.** Metagenomic associations between pig and manure floor samples. The resistance gene abundances were rlog transformed using DESeq2 prior to visualization. (a) Principal component analysis plot visualizing resistome differences, annotated according to herd of origin (colour) and sampling strategy (shape). (b) Heatmap of the resistance genes with highest average abundance across the 20 datasets. Rlog abundances are scaled to each gene (Z-score) allowing easy inter-sample comparison. Bold italic sample names indicate highest overall antimicrobial consumption.

for *E. coli* (cfu- and MIC-based) correlated significantly (Spearman rank,  $\alpha < 0.05$ ) (Figure S5). This is meaningful because they both describe the same species phenotypically. See Figure S6 for all MIC panel results. The only tetracycline resistance measurement to significantly correlate with the expected resistance was the metagenomic measurement. For the hierarchical clustering of methods, the *E. coli* resistance estimates group together, whereas the two species-independent estimates (aerobic bacteria and metagenomics) were both closer to the expected resistance, with metagenomics being the closest (Figure S5b). Bootstrapping indicated that clustering was robust. All herd-level tetracycline resistance numbers are presented in Table S4.

The lack of negative correlations suggests that tetracycline usage increases tetracycline resistance and that this observation is independent of measurement method. However, only the metagenomic measurement significantly correlated with expected resistance. The specific correlation coefficients suggest the bulk of tetracycline resistance genes, found using metagenomic read mapping, follow tetracycline consumption closer than *E. coli* alone would suggest. Other studies have found significant associations between antimicrobial consumption and phenotypic resistance prevalence estimates.<sup>5,6</sup> We stress that our study, which has a low number of truly independent samples, does not disqualify phenotypic association studies. We simply find that the drug-resistance association is strongest with our metagenomic approach.

### Gene-level associations with antimicrobial consumption

As our metagenomic measurement for tetracycline resistance was the only measurement to significantly correlate with the

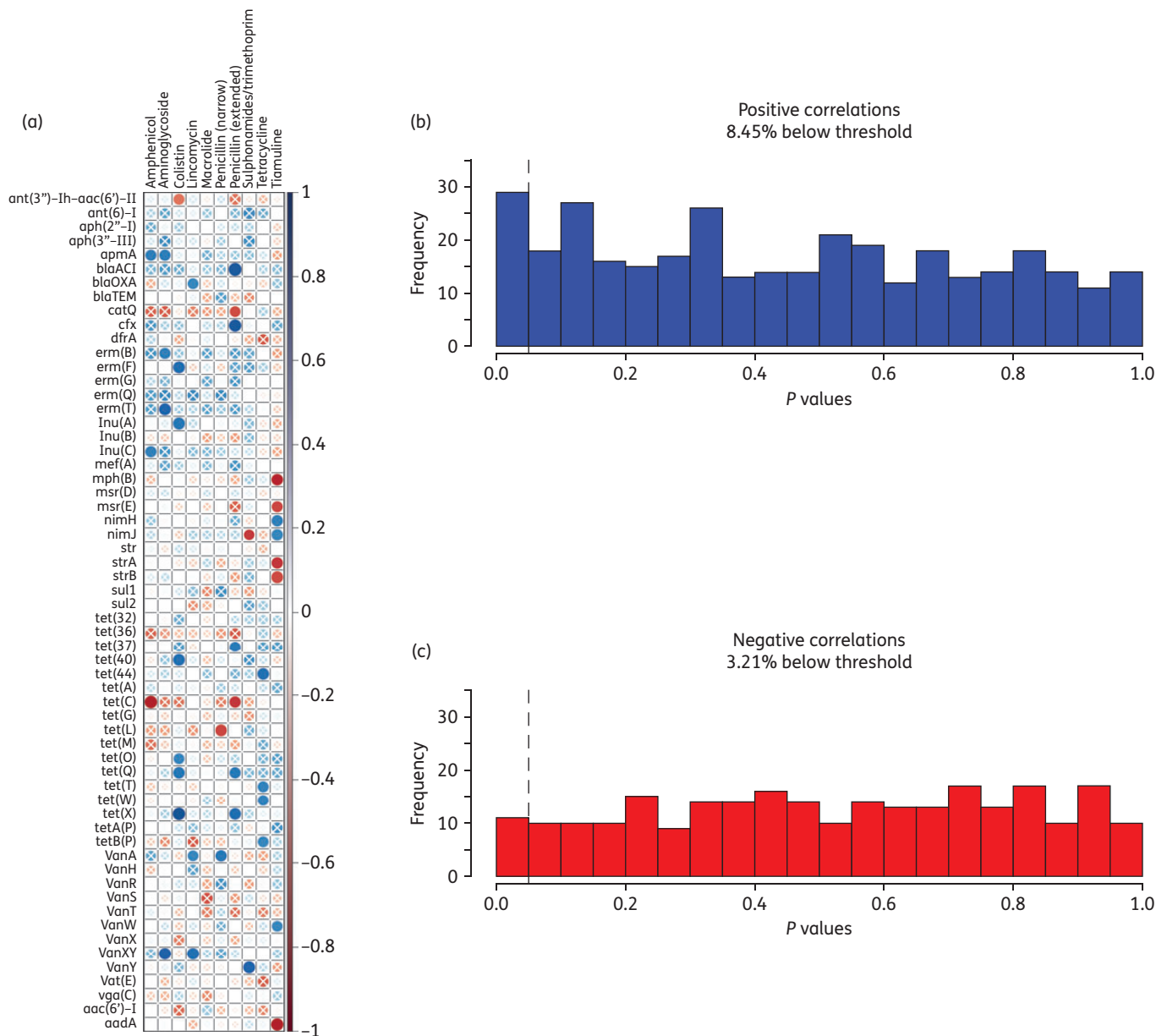
expected tetracycline resistance, we chose to perform regression analyses on metagenomic data to assess whether specific genes were associated with consumption of the most commonly used drug classes: tetracyclines, macrolides, narrow-spectrum penicillins and aminoglycosides.

Tetracycline use was associated with the tetracycline resistance gene *tet(44)* in pig samples ( $FDR < 0.1$ ). Macrolide usage was associated with macrolide resistance genes *erm(B)* and *erm(G)* in floor and pig samples, respectively.

*apmA*, a gene providing resistance to the veterinary aminoglycoside apramycin, followed the aminoglycoside consumption, both in manure and pig samples. In the pig samples, aminoglycoside use was associated with *tet(44)* and two additional aminoglycoside resistance genes: *ant(6)-I* and *str(B)*. In summary, associations were between resistance genes and the antibiotic drug classes they protect against. The exception was *tet(44)*, which followed aminoglycoside consumption in pigs. *tet(44)* is co-located with the *ant(6)-I* variant *ant(6)-Ib* in a mobile *Campylobacter* pathogenicity island.<sup>43</sup> Since these both followed aminoglycoside usage, it seems likely the association is explained by aminoglycoside co-selection of *tet(44)*. Regression analysis results are included in Table S7.

We noticed an overrepresentation of positive associations among genes with low unadjusted *P* values, suggesting even more genes were positively associated with drug use, but outside our resolution.

To test if this was the case, while allowing for both co- and cross-resistance, we correlated the most common resistance gene abundances in the 20 metagenomes with the drug consumption values (Figure 3a). There were many more positive



**Figure 3.** Association between resistance gene abundance and drug class consumptions. Resistance gene abundances (FPKM) were Spearman correlated against consumption estimates for the most used drug classes. (a) Correlation plot where colour denotes the Spearman correlation coefficient ( $R$ ) of a given correlation and the circle size is proportional to the coefficient of determination ( $R^2$ ). White crosses indicate  $P$  values above 0.05. (b) Histogram of  $P$  values for the positive correlation coefficients. (c) Histogram of  $P$  values for the negative correlation coefficients. (b and c) A broken line shows the  $P$  value threshold of 0.05 and, above the plots, the proportion of correlations below that threshold is indicated.

correlations than negative correlations. In addition, among the positive correlations,  $P$  values  $<0.05$  were overrepresented (8.45%) (Figure 3b). This was in contrast to the negative correlations, where low  $P$  values were underrepresented (3.21%) (Figure 3c). While no single correlation should be trusted without correcting for multiple testing, this confirms that drug use is associated with AMR, and indicates our metagenomic approach is sensitive enough to detect drug-induced changes in the swine herd resistome.

A valid concern with using metagenomic-based read mapping for measuring resistance is the inability to distinguish between zoonotic and immobile genes in an environment.<sup>44</sup> This drawback of metagenomics could, however, be reduced. Metagenomic assembly techniques can give contextual information on genes' locations and is an area of active development.<sup>45</sup> Novel approaches such as using gene co-abundance profiles across many samples, metagenomic binning or molecularly linking adjacent DNA prior to lysis can also help elucidate the taxonomic

affiliation of accessory genes.<sup>46,47</sup> Metagenomic assembly would complement our read-based quantification well, as our proposed method is limited to quantifying resistance gene groups and is unable to identify specific alleles of interest. Rare resistance genes would not, however, get sufficient coverage for assembly at currently feasible sequencing depths, leaving an important role for read-mapping. Also, it is important to note that our chosen database ResFinder focuses on resistance genes found in culturable pathogens, meaning one might miss a large fraction of functional resistance genes. In order to determine which genes are ultimately important to human and animal health, further research is needed, especially into the host range and mobility of AMR genes, their disseminating vectors and taxonomic background. Using functional metagenomics one can detect novel resistance genes, while the aforementioned methods can help elucidate whether a discovered resistance gene is likely to transfer to potential human pathogens.<sup>44,48</sup> These methods should together inform which genes we need actively to screen for and add to databases. Our observational approach and more controlled experiments can then be used to determine drug resistance associations, which should inform future drug use.<sup>22</sup>

### Conclusions

We have developed an approach for measuring AMR in swine herds that can be applicable elsewhere. Key features are: (i) sampling multiple pen floors for manure, followed by pooling to a single composite sample; (ii) use of a DNA extraction kit that retains a high microbial diversity; (iii) use of PCR-free library preparation to avoid unnecessary amplification bias; (iv) paired-end sequencing for improved accuracy; and (v) quantification of DNA sequence mapping to a database of known resistance genes. Previous experimental studies have shown, using metagenomics, that therapeutic or sub-therapeutic additions of antibiotics cause increases in resistance.<sup>49,50</sup> To our knowledge, this is the first observational study showing that metagenomics can quantify and elucidate the effects of antimicrobial usage in livestock animals. The fact that the technology can detect these changes in natural populations makes us confident that metagenomics should be used in future AMR monitoring programmes.

### Acknowledgements

We would like to thank the 10 farmers who participated in the study without any form of compensation.

### Funding

This work was supported by the Danish Veterinary and Food Administration and the EU FP7 programme (grant number 613754, Effort against antimicrobial resistance, <http://www.effort-against-amr.eu/>). S. J. P. was supported by Carlsbergfondet (grant number 2013\_01\_0377). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

### Transparency declarations

None to declare.

### Author contributions

F. M. A., Y. A., H. V. and A. F. conceived the study. V. D. A., L. D. K., M. S. J. and H. V. designed the sampling strategy, selected study participants and performed sampling. H. M., J. C., B. E. K. and M. S. J. performed the laboratory work. P. M., H. V., F. M. A., S. J. P., Y. A. and O. L. analysed the biological and epidemiological data. P. M. prepared the manuscript. All authors reviewed and edited the manuscript.

### Supplementary data

Supplementary data, including Figures S1–S6 and Tables S1–S7, are available at JAC Online (<http://jac.oxfordjournals.org/>).

### References

- 1 WHO. *Antimicrobial Resistance: Global Report on Surveillance 2014*. <http://www.who.int/drugresistance/documents/surveillance-report/en/>.
- 2 Aarestrup FM. The livestock reservoir for antimicrobial resistance: a personal view on changing patterns of risks, effects of interventions and the way forward. *Philos Trans R Soc Lond B Biol Sci* 2015; **370**: 20140085.
- 3 Smith DL, Harris AD, Johnson J *et al*. Animal antibiotic use has an early but important impact on the emergence of antibiotic resistance in human commensal bacteria. *Proc Natl Acad Sci USA* 2002; **99**: 6434–9.
- 4 Aarestrup FM. Veterinary drug usage and antimicrobial resistance in bacteria of animal origin. *Basic Clin Pharmacol Toxicol* 2005; **96**: 271–81.
- 5 Chantziaras I, Boyen F, Callens B *et al*. Correlation between veterinary antimicrobial use and antimicrobial resistance in food-producing animals: a report on seven countries. *J Antimicrob Chemother* 2014; **69**: 827–34.
- 6 Dunlop RH, McEwen SA, Meek AH *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* 1998; **34**: 283–305.
- 7 Agersø Y, Aarestrup FM. Voluntary ban on cephalosporin use in Danish pig production has effectively reduced extended-spectrum cephalosporinase-producing *Escherichia coli* in slaughter pigs. *J Antimicrob Chemother* 2013; **68**: 569–72.
- 8 Aarestrup FM, Seyfarth AM, Emborg H *et al*. 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* 2001; **45**: 2054–9.
- 9 Aarestrup FM, Bager F, Jensen NE *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* 1998; **106**: 745–70.
- 10 CDC. *National Antimicrobial Resistance Monitoring System for Enteric Bacteria (NARMS): Human Isolates Final Report, 2012*. 2014. <http://www.cdc.gov/narms/pdf/2012-annual-report-narms-508c.pdf>.
- 11 Public Health Agency of Canada. *Canadian Integrated Program for Antimicrobial Resistance Surveillance (CIPARS) Annual Report*. 2015; Chapter 2. [http://publications.gc.ca/collections/collection\\_2015/asp-phac/HP2-4-2013-2-eng.pdf](http://publications.gc.ca/collections/collection_2015/asp-phac/HP2-4-2013-2-eng.pdf).
- 12 European Food Safety Authority. Harmonised monitoring of antimicrobial resistance in *Salmonella* and *Campylobacter* isolates from food animals in the European Union. *Clin Microbiol Infect* 2008; **14**: 522–33.
- 13 Kanwar N, Scott HM, Norby B *et al*. Impact of treatment strategies on cephalosporin and tetracycline resistance gene quantities in the bovine fecal metagenome. *Sci Rep* 2014; **4**: 5100.



- 14 Strommenger B, Kettlitz C, Werner G et al. Multiplex PCR assay for simultaneous detection of nine clinically relevant antibiotic resistance genes in *Staphylococcus aureus*. *J Clin Microbiol* 2003; **41**: 4089–94.
- 15 Zankari E, Hasman H, Kaas RS et al. Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. *J Antimicrob Chemother* 2013; **68**: 771–7.
- 16 Wright GD. The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat Rev Microbiol* 2007; **5**: 175–86.
- 17 Ghosh TS, Gupta S, Sen, Nair GB et al. In silico analysis of antibiotic resistance genes in the gut microflora of individuals from diverse geographies and age-groups. *PLoS ONE* 2013; **8**: e83823.
- 18 Forslund K, Sunagawa S, Kultima JR et al. Country-specific antibiotic use practices impact the human gut resistome. *Genome Res* 2013; **23**: 1163–9.
- 19 Fang H, Wang H, Cai L et al. Prevalence of antibiotic resistance genes and bacterial pathogens in long-term manured greenhouse soils as revealed by metagenomic survey. *Environ Sci Technol* 2015; **49**: 1095–104.
- 20 Nordahl Petersen T, Rasmussen S, Hasman H 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* 2015; **5**: 11444.
- 21 Huang K, Tang J, Zhang X-X et al. A comprehensive insight into tetracycline resistant bacteria and antibiotic resistance genes in activated sludge using next-generation sequencing. *Int J Mol Sci* 2014; **15**: 10083–100.
- 22 Willmann M, El-Hadidi M, Huson DH et al. Antibiotic selection pressure determination through sequence-based metagenomics. 2015; **59**: 7335–45.
- 23 Stege H, Bager F, Jacobsen E et al. VETSTAT - The Danish system for surveillance of the veterinary use of drugs for production animals. *Prev Vet Med* 2003; **57**: 105–15.
- 24 Jensen VF, Jacobsen E, Bager F. Veterinary antimicrobial-usage statistics based on standardized measures of dosage. *Prev Vet Med* 2004; **64**: 201–15.
- 25 Nielsen KL, Dynesen P, Larsen P et al. Faecal *Escherichia coli* from patients with *E. coli* urinary tract infection and healthy controls who have never had a urinary tract infection. *J Med Microbiol* 2014; **63**: 582–9.
- 26 Ni J, Yan Q, Yu Y. How much metagenomic sequencing is enough to achieve a given goal? *Sci Rep* 2013; **3**: 2013.
- 27 Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 2011; **17**: 10–2.
- 28 Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009; **25**: 1754–60.
- 29 Hunter S, Corbett M, Denise H et al. EBI metagenomics—a new resource for the analysis and archiving of metagenomic data. *Nucleic Acids Res* 2014; **42**: 600–6.
- 30 Mitchell A, Bucchini F, Cochrane G et al. EBI metagenomics in 2016 - an expanding and evolving resource for the analysis and archiving of metagenomic data. *Nucleic Acids Res* 2016; **44**: D595–603.
- 31 Caporaso JG, Kuczynski J, Stombaugh J et al. QIIME allows analysis of high-throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing. *Nat Methods* 2010; **7**: 335–6.
- 32 DeSantis TZ, Hugenholtz P, Larsen N et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 2006; **72**: 5069–72.
- 33 McMurdie PJ, Holmes S. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol* 2014; **10**: e1003531.
- 34 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; **15**: 550.
- 35 Schmidt GV, Møllerup A, Christiansen LE et al. Sampling and pooling methods for capturing herd level antibiotic resistance in swine feces using qPCR and CFU approaches. *PLoS ONE* 2015; **10**: e0131672.
- 36 Clasen J, Møllerup A, Olsen JE et al. Determining the optimal number of individual samples to pool for quantification of average herd levels of antimicrobial resistance genes in Danish pig herd using high-throughput qPCR. *Vet Microbiol* 2016; **189**: 46–51.
- 37 Shimodaira H. Approximately unbiased tests of regions using multistep-multiscale bootstrap resampling. *Ann Stat* 2004; **32**: 2616–41.
- 38 DANMAP 2014. 2015. <http://www.danmap.org/Downloads/Reports.aspx>.
- 39 van Hoek AHAM, Mevius D, Guerra B et al. Acquired antibiotic resistance genes: an overview. *Front Microbiol* 2011; **2**: 1–27.
- 40 Whittle G, Whitehead TR, Hamburger N et al. Identification of a new ribosomal protection type of tetracycline resistance gene, *tet(36)*, from swine manure pits. *Appl Environ Microbiol* 2003; **69**: 4151–8.
- 41 Forsberg KJ, Patel S, Gibson MK et al. Bacterial phylogeny structures soil resistomes across habitats. *Nature* 2014; **509**: 612–6.
- 42 Roberts MC, Lanciardi J. Transferable Tet M in *Fusobacterium nucleatum*. *Antimicrob Agents Chemother* 1990; **34**: 1836–8.
- 43 Abril C, Brodard I, Perreten V. Two novel antibiotic resistance genes, *tet(44)* and *ant(6)-Ib*, are located within a transferable pathogenicity island in *Campylobacter fetus* subsp. *fetus*. *Antimicrob Agents Chemother* 2010; **54**: 3052–5.
- 44 Martínez JL, Coque TM, Baquero F. What is a resistance gene? Ranking risk in resistomes. *Nat Rev Microbiol* 2015; **13**: 116–23.
- 45 Peng Y, Leung HCM, Yiu SM et al. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 2012; **28**: 1420–8.
- 46 Nielsen HB, Almeida M, Juncker AS et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nat Biotechnol* 2014; **32**.
- 47 Korbel JO, Lee C. Genome assembly and haplotyping with Hi-C. *Nat Biotechnol* 2013; **31**: 1099–101.
- 48 Sommer MOA, Dantas G, Church GM. Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science* 2009; **325**: 1128–31.
- 49 Looft T, Johnson T, Allen HK et al. In-feed antibiotic effects on the swine intestinal microbiome. *Proc Natl Acad Sci USA* 2012; **109**: 1691–6.
- 50 Yin J, Zhang X-X, Wu B et al. Metagenomic insights into tetracycline effects on microbial community and antibiotic resistance of mouse gut. *Ecotoxicology* 2015; **24**: 2125–32.