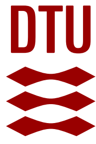
Downloaded from orbit.dtu.dk on: Apr 27, 2025



DTU Library

Robustness in quantifying the abundance of antimicrobial resistance genes in pooled faeces samples from batches of slaughter pigs using metagenomics analysis

Dalhoff Andersen, Vibe; Jensen, Marie Stengaard; Munk, Patrick; Vigre, Håkan

Published in:
Journal of Global Antimicrobial Resistance

Link to article, DOI:
10.1016/j.jgar.2021.02.005

Publication date:
2021

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

Link back to DTU Orbit

Citation (APA):
Dalhoff Andersen, V., Jensen, M. S., Munk, P., & Vigre, H. (2021). Robustness in quantifying the abundance of antimicrobial resistance genes in pooled faeces samples from batches of slaughter pigs using metagenomics analysis. *Journal of Global Antimicrobial Resistance*, *24*, 398-402. https://doi.org/10.1016/j.jgar.2021.02.005

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.

Journal of Global Antimicrobial Resistance 24 (2021) 398–402

Contents lists available at ScienceDirect

Journal of Global Antimicrobial Resistance

journal homepage: www.elsevier.com/locate/jgar

Short Communication

Robustness in quantifying the abundance of antimicrobial resistance genes in pooled faeces samples from batches of slaughter pigs using metagenomics analysis

Vibe Dalhoff Andersen*, Marie Stengaard Jensen, Patrick Munk, Håkan Vigre

The National Food Institute, Research Group for Genomic Epidemiology, Technical University of Denmark, Kemitorvet, Kgs. Lyngby, Denmark

A R T I C L E I N F O

Article history:
Received 7 September 2020
Received in revised form 15 January 2021
Accepted 3 February 2021
Available online 21 February 2021

Keywords:
Pigs
Antimicrobial resistance
Sampling
Pooled sample
Sequencing
Measurement error

A B S T R A C T

Objectives: With the continued spread of antimicrobial resistance (AMR) in animals, it is important to assess its occurrence throughout a microbiome quantitatively in order to evaluate significantly affecting factors, e.g. antimicrobial usage. Metagenomics methods make it possible to measure the abundance of AMR genes in complex samples such as pooled faeces samples from batches of slaughter pigs. This study was performed to determine the random error in pooled samples from batches of pigs at slaughter and the measurement error from the metagenomics processes.
Methods: In four farms, two pooled samples were obtained from a batch of slaughter pigs by two individual samplers, and each pooled sample was thereafter processed twice. Hierarchically clustered heatmaps were applied to evaluate dissimilarities between samples. The coefficient of variation was used to calculate the percentage difference between samples from the same farm.
Results: Results of the analysis revealed that it was not possible to quantitatively separate the variation arising from sampling and metagenomics processes. They both contributed to the overall measurement error in batches of slaughter pigs.
Conclusion: Sampling of single pigs in 30 randomly selected pig pens within the farms provides a composition representative for frequently occurring AMR genes present within the farms, while rare genes were not dispersed in a similar manner. Aggregating the resistance abundance at gene family or antimicrobial class level will reduce the apparent variation originating from errors in sampling and metagenomics processing.
© 2021 The Authors. Published by Elsevier Ltd on behalf of International Society for Antimicrobial Chemotherapy. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

A sound assessment of the relationship between antimicrobial use and antimicrobial resistance (AMR) requires accurate and precise estimates describing the actual antimicrobial use and the real occurrence of AMR, which in turn will enhance the power to detect significant effects of antimicrobials on AMR and increase the validity of the estimated size of these effects [1].

Resistance can be measured by phenotypic methods, e.g. cultivation and minimum inhibitory concentration (MIC) determination [2]. These methods detect the AMR occurrence to specific antimicrobials using indicator bacteria, e.g. *Escherichia coli* and *Enterococcus* spp., which are part of the normal gut flora and carry AMR genes easily [3]. The methods provide valuable information of resistance occurrence but do not quantify its amount or dispersion within an entire microbiome [4,5]. Samples from the pig gut hold more than 400 different bacterial species, with *E. coli* and *Enterococcus* spp. not comprising the majority of bacteria in the gut community [6,7]. Therefore, estimating resistance using indicator bacteria might provide misleading estimates.

Metagenomics methods quantify the abundance of resistance genes in complex samples from animals, e.g. a pooled faeces sample representing a pig population resistome [4], thereby making it possible to assess the quantitative effect of altered antimicrobial use on that resistome, including assessment of potential co- and cross-resistance [8].

Combined investigation of the performance of the sampling scheme to obtain a population representative pooled sample and metagenomics processes in quantifying the abundance of AMR genes in a population of slaughter pigs is, to the best of our knowledge, less well elucidated. However, when quantifying the

* Corresponding author at: The National Food Institute, Research Group for Genomic Epidemiology, Technical University of Denmark, Kemitorvet, Kgs. Lyngby, Denmark.

E-mail address: vdan@food.dtu.dk (V.D. Andersen).

http://dx.doi.org/10.1016/j.jgar.2021.02.005
2213-7165/© 2021 The Authors. Published by Elsevier Ltd on behalf of International Society for Antimicrobial Chemotherapy. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

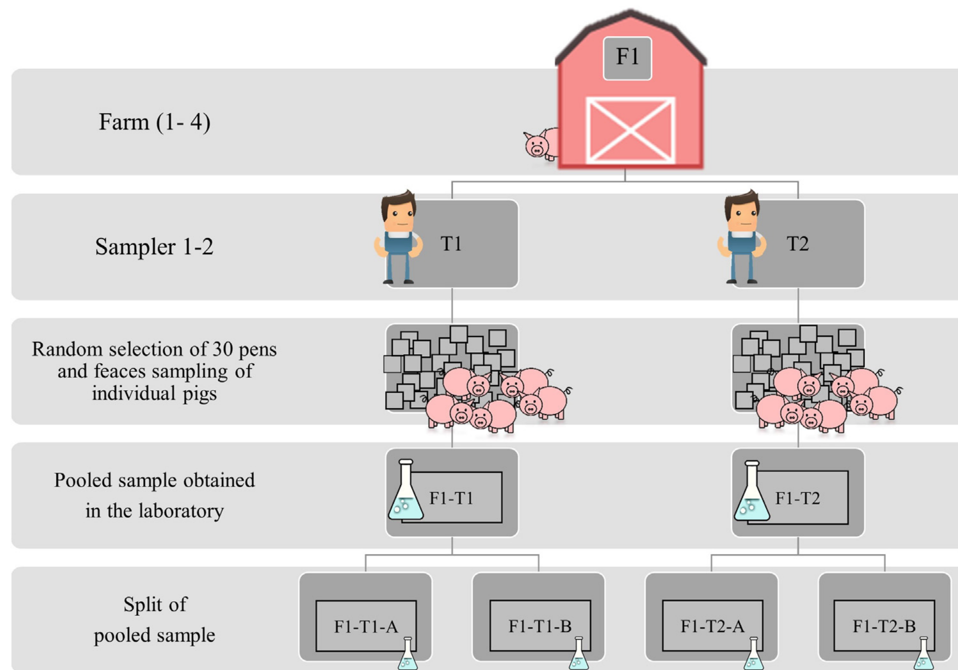


Fig. 1. In four farms, the conduct of obtaining four pooled samples from a batch of slaughter pigs.

abundance of AMR genes in pig populations, it is equally key to assess the performance of the sampling scheme.

The objective of this study was to assess the robustness of the entire sampling procedure. Consequently, to investigate (i) the error from the sampling scheme originating from natural variation in resistance between pigs included in a pooled sample and (ii) the metagenomics process error originating from random variation in the laboratory methods and the subsequent assignment of sequencing fragments to AMR genes.

2. Materials and methods

2.1. Study design

This cross-sectional study was conducted in four Danish conventional pig farms in which faecal samples were collected from one batch of slaughter pigs. In each farm, two samplers obtained 30 samples individually and the samples from each sampler were pooled into a single sample (Fig. 1). The two pooled samples from each farm were thereafter processed for metagenomic sequencing individually twice (Fig. 1).

2.2. Sample procedure

2.2.1. Sampling scheme

Four farms (F1–F4) containing slaughter pigs >80 kg and a yearly production above 5000 slaughter pigs were visited during 2015 and 2016 (Fig. 1). At each farm, two samplers (T1 and T2) carried out a sampling scheme individually (Fig. 1). A sampling scheme consisted of collection of 30 individual faecal samples from 30 randomly selected pig pens (round). An online randomisation tool was applied to select 30 pens [9]. Samplers T1 and T2 were the same two persons throughout the study (Fig. 1).

Faeces were collected during defecation, which assured that neither the environment nor the time affected the bacterial content or composition. Samples were kept cold at ~5 °C until arrival at the laboratory on the same day. At the laboratory, each round of samples was pooled into a single sample by weighing out a 5 g portion of each of the 30 samples and gathering them into one pool. Preceding DNA processing of the pooled samples, these were split by obtaining 0.5 g from each pooled sample twice (A and B) (Fig. 1).

2.2.2. Metagenomics processes

DNA extraction was performed with a modified QIAamp Fast DNA Stool Mini Kit protocol (QIAGEN, Valencia, CA, USA) [10]. Subsequent sequencing of all DNA extracted from the pooled faeces sample has been described by Munk et al. [4]. Resistance was quantified using the KMA tool to map the paired-end fragments against the ResFinder database (December 2019) [11]. The abundance of AMR genes was normalised as fragments per kilobase reference per million fragments (FPKM) using the following formula:

$$AMR\ gene_p\ abundance = \left(\frac{n}{(N*l)}\right) * 10^6 F * 1000bp \quad (1)$$

where n = number of mapped fragments, N = total number of fragments, F = fragments, bp = base pair and p = gene [12].

Because the ResFinder database contains multiple similar resistance gene sequences, identical parts of homologous gene variants causes unspecific mapping of the fragments, therefore the AMR gene abundances were also aggregated at the gene family and the antimicrobial class level (Supplementary Table S1) [4].

2.3. Data analysis

The variation between the pooled faecal samples originating from the sampling scheme and the metagenomic processing was assessed in scatterplots. To evaluate minor variations in abundance of AMR genes, the FPKMs were also plotted with a $\log_{10}$ transformation, with 0.01 added to all FPKM values to address zeros in the data set.

The similarities and dissimilarities in abundance of AMR genes of the pooled samples were assessed using heatmaps based on hierarchical clustering analysis at gene variant, gene family and antimicrobial class level. The Euclidean distance and Ward's

Table 1
Sampling time and population and production characteristics of the selected farms.

Farm	Time of sampling	No. of pens with pigs >80 kg	Slaughter pig production[a]
F1	29-06-2015	41	5185
F2	06-07-2015	46	5366
F3	18-07-2016	60	9216
F4	26-07-2016	142	9485

[a] Number of delivered slaughter pigs during the year of sampling.

minimum-variance method were applied to calculate the length between observations and for the clustering, respectively.

The sampling scheme reproducibility of sampler T1 versus T2 within farms and the metagenomic process reproducibility of sample A versus B within farms was estimated as the coefficient of variation (CV) using the following formula:

$$CV = 100 * \sqrt{\left(\sum \frac{(d-m)^2}{2n}\right)} \quad (2)$$

where n = number of AMR genes, d = difference between two estimates of AMR abundance, and m = mean of two estimates of AMR gene abundance [13]. It was assumed that there was no systematic difference between samples A and B, and sampler T1 and T2 originating from the same farm. When comparing sampler T1 to sampler T2, their metagenomic process A samples and metagenomic process B samples were assessed. The CVs were calculated taking into account all genes and then only genes occurring in both samples, the latter to exclude rare and very rare genes from the analysis. Finally, CVs were estimated for the AMR gene abundances aggregated to gene family and antimicrobial class level.

Microsoft Excel 2010 (Microsoft Corp., Redmond, WA, USA) and R v.3.3.3 with the R packages tidyverse for data management [14] and pheatmap for heatmaps [15] were applied.

3. Results

3.1. Study sample

Two farms were sampled in 2015 and two farms in 2016 (Table 1). The number of pens with pigs >80 kg in the four farms ranged from 41 to 142 (Table 1). On average, each pen contained between 12 and 16 pigs. The yearly production of slaughter pigs of the farms sampled in 2016 was approximately twice the size of the farms sampled in 2015 (Table 1).

3.2. Data analysis

In total, 181 different AMR genes were found across the samples (Supplementary Table S1).

The scatterplots illustrating the variation between estimated FPKM of each gene of the metagenomic process A sample to the corresponding B sample at sampler and farm level showed little variation between the results of the two (Supplementary Fig. S2). However, the scatterplots with the $\log_{10}$ transformation FPKM provided a good illustration of the variation of the scarcely found genes occurring in the metagenomic process A or B only (Supplementary Fig. S3). In particular, farm F1 had numerous genes only found in one of the two processes (Supplementary Fig. S3). The $\log_{10}$ transformation also illustrated larger variations between genes of sampler T1 compared with sampler T2 in farm F4 and reversely in farm F3 (Supplementary Fig. S3).

The heatmaps show that the dissimilarity between farms was larger than the dissimilarity within farms, i.e. samples from the same farm clustered, independent of the level (gene variant, gene family and antimicrobial class) the data were analysed at (Fig. 2). Furthermore, independent of the level, farms F2 and F4 were more similar to each other compared with farms F1 and F3 (Fig. 2).

At gene variant level, the dissimilarity between slaughter pigs from the same batch (T1 and T2) was shown to be larger than the dissimilarity from the metagenomic process A and B of the farms F1 and F2. In contrast, the dissimilarity from the process A and B was shown to be larger than the dissimilarity between slaughter pigs from the same batch (T1 and T2) of the farms F3 and F4 (Fig. 2a). At the same level, the scarcely found AMR genes were not necessarily farm-specific, e.g. an AMR gene found in two samples only was seldom farm-specific (result not shown).

At gene family level, the dissimilarity between samples from the same farm decreased for all the farms compared with the gene variant level (Fig. 2a,b). One distinct difference could be observed; only one of the two samplers metagenomic processed A and B sample clustered (Fig. 2a,b). Subsequently, within farms the dissimilarity from the metagenomic process A and B was both smaller and larger compared with the dissimilarity between pigs from the same batch (T1 and T2) (Fig. 2a,b).

The dissimilarity at antimicrobial class level decreased for all the farms compared with the gene variant level, although it remained the same compared with the gene family level (Fig. 2). Overall, none of the within-farm clustering remained the same at all three levels (Fig. 2).

The CV of the abundance of AMR genes estimated at gene variant level between metagenomic process A and B within farm and sampler ranged from 65% to 89% with all AMR genes included and from 17% to 36% with the AMR genes occurring in both samples (Supplementary Table S4). At the gene family and antimicrobial class level, the CV of the abundance of AMR genes decreased overall, thus with all AMR genes included it ranged from 58% to 85% and 5% to 77%, respectively, and with AMR genes occurring in both samples it ranged from 11% to 23% and 5% to 22%, respectively (Supplementary Table S4).

Within farms, the estimated CV of the abundance of AMR genes at gene variant level between sampler T1 and sampler T2 ranged from 69% to 88% with all AMR genes included and from 19% to 33% with the AMR genes occurring in both samples (Supplementary Table S5). At gene family level, the estimated CV between samplers ranged from 60% to 78% and 9% to 20% with all AMR genes and with genes occurring in both samples, respectively (Supplementary Table S5). In contrast, at antimicrobial class level, the CV decreased to the range 38% to 77% taking all the AMR genes into account, but the CV increased to the range 14% to 33% taking only AMR genes occurring in both samples into account (Supplementary Table S5).

The metagenomic sequences have been deposited in the European Nucleotide Archive (ENA) (Supplementary Table S6) (EMBL-EBI, 2016).

4. Discussion

Although larger farms were selected to avoid an overestimation of robustness, the four farms differed substantially in size. In farm F1 and F2, both samplers sampled several of the same pens, whereas only a small number of pig pens were sampled by both

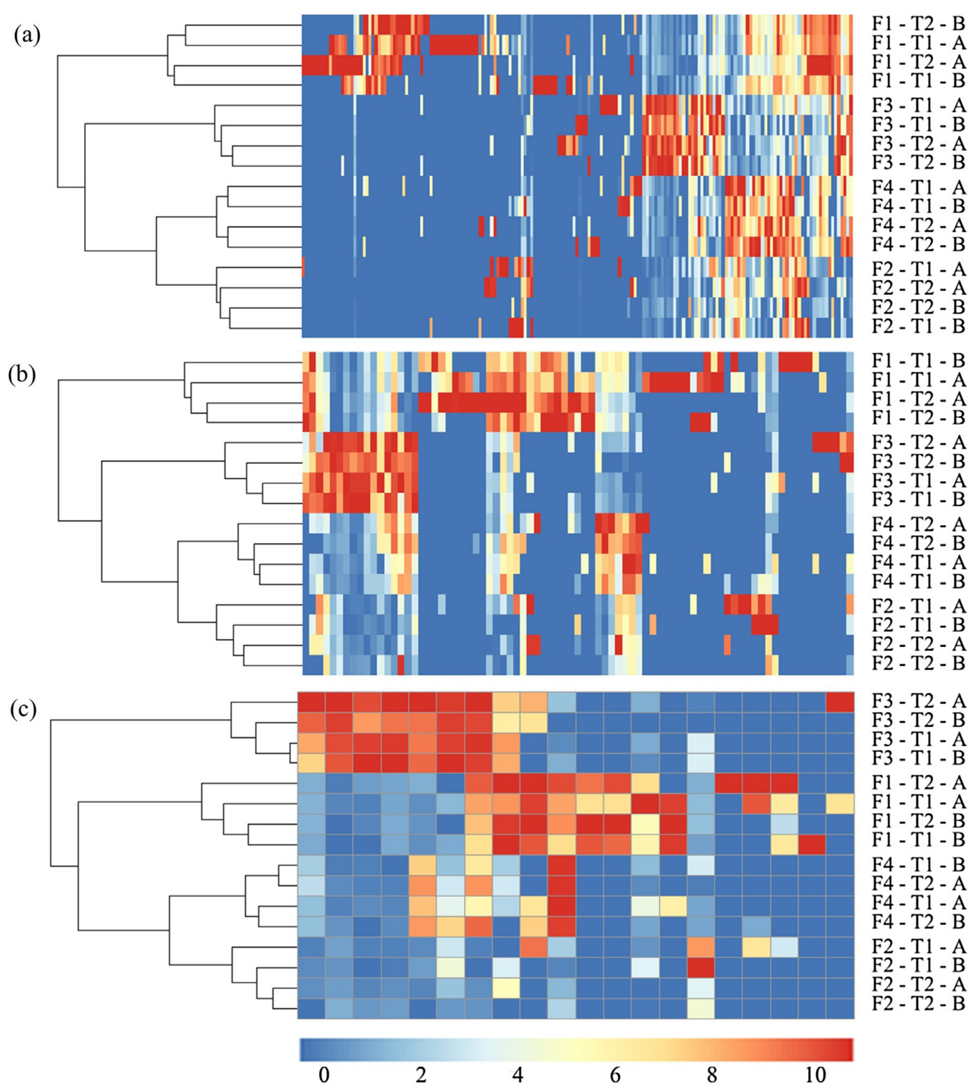


Fig. 2. Heatmaps with dendrograms of the hierarchical Ward D2 clustering on Euclidean distances of min–max scaled FPKM of the 16 pooled faeces samples and the abundance of resistance found in two or more samples at (a) gene variant level, (b) gene family level and (c) antimicrobial class level. FPKM, fragments per kilobase reference per million fragments.

samplers in farms F3 and F4. In theory, given a fixed number of samples contributing to a pooled sample, the sampling error will increase by an increased number of pens and pigs in the sections, provided there is significant variation in AMR between pens and pigs. The results of cluster analysis and estimation of CV regarding differences between samplers indicates a homogenous abundance of common AMR genes in batches of pigs. In consequence, the results obtained from the sampling schedule seem to provide reliable results independent of farm size, which is in agreement with previous studies [16,17].

In Danish pig farms, pigs are moved between sections and pens during the rearing period owing to uneven growth. This could explain the similarities observed between samplers, thus the slaughter pig batch has acquired similar diversity and abundance of AMR genes. On the other hand, the dispersion and persistence of the abundance of AMR genes within a farm could have become stably integrated into the faecal microbiome because of previous antimicrobial use [18].

The samples clustered at farm level in all analyses. It was expected that the abundance of AMR genes would be similar between the metagenomic processed samples from the same pooled sample than the samples obtained by the two samplers within a farm. However, the heatmaps clustering at gene variant, gene family and antimicrobial class level showed that both the sampling schedule and the metagenomic process contributed equally to the uncertainty in the measured abundance of AMR genes in the farms.

The deeper a sample is sequenced, the higher likelihood to detect rare and very rare AMR genes. The abundance of gene families will also be more accurately estimated, making it easier to determine the effect of interventions, e.g. antimicrobial use alterations. Therefore, sequencing depth should be chosen based on the overall aim of a study [19,20]. Rare AMR genes are important to evaluate the diversity within a resistome. To get a better representation of AMR genes would require a substantial increase in the number of fragments by increasing the sequencing depth to detect the rarer AMR genes [20]. Furthermore, in order to detect rare genes that are not spread throughout the entire batch, a more comprehensive sampling schedule has to be applied, where a very large number of pigs is represented in the pooled sample.

By aggregating the abundance of AMR genes to gene family or antimicrobial class level, the CV declined at both the metagenomic

processing and sampler level. With the exclusion of rarely occurring AMR genes in the calculation of CV at all three levels, the CV declined substantially in particular at the gene family level. Therefore, assessing the relationship between antimicrobial use and AMR, obtaining accurate and precise measurements of the occurrence of AMR could beneficially be performed at gene family level with the predominant abundant AMR genes to enhance the power to detect significant effects [8].

5. Conclusion

With the progress in metagenomics testing to describe AMR, it becomes paramount to evaluate the reproducibility of the test itself and the reproducibility in relation to the test material, in our case pooled faeces samples. This study was not able to quantitatively separate the variation arising from sampling and the metagenomic process, but the results show that both contributed to the overall measurement error of AMR in batches of slaughter pigs using metagenomics on pooled faecal samples. Sampling of single pigs in 30 randomly selected pens within the farms provides a composition representative of the frequently occurring AMR genes present within the farms, while rare genes were not dispersed in a similar manner. Finally, aggregating the abundance of AMR genes at gene family or antimicrobial class level will reduce the apparent variation originating from errors in sampling and metagenomics processing.

Funding

This work was supported by the Danish Ministry of Food, Agriculture and Fisheries.

Competing interests

None declared.

Ethical approval

Not required.

Acknowledgments

The authors would like to thank all of the farmers for participating in this study, whose help is greatly appreciated. The authors also thank their colleagues from the National Food Institute of the Technical University of Denmark for their dedicated technical assistance with DNA extraction.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jgar.2021.02.005.

References

[1] Menditto A, Patriarca M, Magnusson B. Understanding the meaning of accuracy, trueness and precision. Accredit Qual Assur 2007;12:45–7, doi: http://dx.doi.org/10.1007/s00769-006-0191-z.
[2] Davison HC, Woolhouse MEJ, Low JC. What is antibiotic resistance and how can we measure it? Trends Microbiol 2000;8:554–9, doi:http://dx.doi.org/10.1016/S0966-842X(00)01873-4.
[3] Franklin A, Acar J, Anthony F, Gupta R, Nicholls T, Tamura Y, et al. Antimicrobial resistance: harmonisation of national antimicrobial resistance monitoring and surveillance programmes in animals and in animal-derived food. Rev Sci Tech 2001;20:859–70, doi:http://dx.doi.org/10.20506/rst.20.3.1315.
[4] Munk P, Andersen VD, de Knegt L, Jensen MS, Knudsen BE, Lukjancenko O, et al. A sampling and metagenomic sequencing-based methodology for monitoring antimicrobial resistance in swine herds. J Antimicrob Chemother 2017;72:385–92, doi:http://dx.doi.org/10.1093/jac/dkw415.
[5] Andersen VD, de Knegt LV, Munk P, Jensen MS, Agersø Y, Aarestrup FM, et al. The association between measurements of antimicrobial use and resistance in the faeces microbiota of finisher batches. Epidemiol Infect 2017;145:2827–37, doi:http://dx.doi.org/10.1017/S0950268817001285.
[6] Lim JS, Yang SH, Kim BS, Lee EY. Comparison of microbial communities in swine manure at various temperatures and storage times. Asian-Australas J Anim Sci 2018;31:1373–80, doi:http://dx.doi.org/10.5713/ajas.17.0704.
[7] Berg RD. The indigenous gastrointestinal microflora. Trends Microbiol 1996;4:430–5, doi:http://dx.doi.org/10.1016/0966-842X(96)10057-3.
[8] Andersen VD, Aarestrup FM, Munk P, Jensen MS, de Knegt LV, Bortolaia V, et al. Predicting effects of changed antimicrobial usage on the abundance of antimicrobial resistance genes in finisher' gut microbiomes. Prev Vet Med 2020;174:104853, doi:http://dx.doi.org/10.1016/j.prevetmed.2019.104853.
[9] Haahr M. True random number service. RANDOM.ORG; 2015. https://www.random.org.
[10] Knudsen BE, Bergmark L, Munk P, Lukjancenko O, Priemé A, Aarestrup FM, et al. Impact of sample type and DNA isolation procedure on genomic inference of microbiome composition. mSystems 2016;1:e00095–16, doi:http://dx.doi.org/10.1128/mSystems.00095-16.
[11] Clausen PTLC, Aarestrup FM, Lund O. Rapid and precise alignment of raw reads against redundant databases with KMA. BMC Bioinformatics 2018;19:307, doi: http://dx.doi.org/10.1186/s12859-018-2336-6.
[12] Hendriksen RS, Munk P, Njage P, van Bunnik B, McNally L, Lukjancenko O, et al. Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. Nat Commun 2019;10:1124, doi:http://dx.doi.org/10.1038/s41467-019-08853-3.
[13] Hyslop NP, White WH. Estimating precision using duplicate measurements. J Air Waste Manag Assoc 2009;59:1032–9, doi:http://dx.doi.org/10.3155/1047-3289.59.9.1032.
[14] Wickham H. Easily install and load' tidyverse' packages. 2020. https://tidyverse.tidyverse.org.
[15] Kolde R. pheatmap: pretty heatmaps. R package v.1.0.12. 2020. https://CRAN.R-project.org/package=pheatmap.
[16] Birkegård AC, Halasa T, Græsbøll K, Clasen J, Folkesson A, Toft N. Association between selected antimicrobial resistance genes and antimicrobial exposure in Danish pig farms. Sci Rep 2017;7:9683, doi:http://dx.doi.org/10.1038/s41598-017-10092-9.
[17] Clasen J, Mellerup A, Olsen JE, Angen Ø, Folkesson A, Halasa T, et al. Determining the optimal number of individual samples to pool for quantification of average herd levels of antimicrobial resistance genes in Danish pig herds using high-throughput qPCR. Vet Microbiol 2016;189:46–51, doi:http://dx.doi.org/10.1016/j.vetmic.2016.04.017.
[18] Pollock J, Muwonge A, Hutchings MR, Mainda G, Bronsvoort BM, Gally DL, et al. Resistance to change: AMR gene dynamics on a commercial pig farm with high antimicrobial usage. Sci Rep 2020;10:1708, doi:http://dx.doi.org/10.1038/s41598-020-58659-3.
[19] Sims D, Sudbery I, Ilott NE, Heger A, Ponting CP. Sequencing depth and coverage: key considerations in genomic analyses. Nat Rev Genet 2014;15:121–32, doi:http://dx.doi.org/10.1038/nrg3642.
[20] Zaheer R, Noyes N, Ortega Polo R, Cook SR, Marinier E, Van Domselaar G, et al. Impact of sequencing depth on the characterization of the microbiome and resistome. Sci Rep 2018;8:5890, doi:http://dx.doi.org/10.1038/s41598-018-24280-8.