Check for updates

Investigation of the microbial community structure and diversity in the environment surrounding a veterinary antibiotic production factory

Junjie Miao, Zhendong Yin, Yuqin Yang, Yiwen Liang, Hongmei Shi [iD] * and Xiangdong Xu [iD] *

The ecological toxicity caused by antibiotic residues and resistance genes in the environment affects the community structures and activities of environmental microorganisms; the ecological toxicity effects of a long-term exposure to low doses antibiotic residues on microorganisms have not however been well-studied. In this work, sequence analysis and species annotation of the full-length 16S rRNA gene were carried out on the extracted whole genome by a 3-generation sequencing method to analyze the diversity of the microbial populations and the population differences among different sampling sites in the environment surrounding a veterinary antibiotic production factory. A total of 1720 OTUs (Operational Taxonomic Unit, OTU) were found, of which 1055 OTUs were in the river samples and 993 OTUs were in the soil samples. 643 and 438 bacterial strains were identified in the river water and soil samples respectively. The bacterial populations are classified into 29 phylum, 612 genus, and 849 species. The dominant phylum of bacteria was Proteobacteria, which was also the absolutely dominant phylum. Shannon diversity index of the bacteria showed that the bacterial abundance in downstream river was significantly higher than that in an upstream non-polluted area ($P < 0.001$), but the difference of bacterial abundance between soil samples was not significant. There were 61 biomarkers in the river water samples from different sampling points, and 14 biomarkers in soil samples. It was found by the difference statistics of the microbial community that there are multiple biomarkers between this veterinary drug production site and the upstream non-polluted area. Significant differences between multiple functional genes were also found in metabolic pathways of the microorganisms. A similar trend was found for the distribution of antibiotic resistance genes (ARGs). It is concluded that the population composition of microorganisms and diversity are likely related to antibiotic residues and to the distributions of ARGs in the environment surrounding the antibiotic production factory.

1. Introduction

Antibiotic residues in the environment have not only caused continuous environmental pollution and a body exposure risk,[1–3] but also induced antibiotic resistance genes (ARGs) due to the excessive use and abuse of antibiotics in medical treatments, livestock industry, and in aquaculture. Studies have indicated that pollution by the antibiotic resistance genes has been spreading throughout rivers, soils, air particles and other ecological environments.[4,5] An international group of researchers tested 4000 kilometres of China's coastline and found that the sediment from estuaries where fresh water from rivers and streams meets the sea had levels of antibiotic resistance gene pollution ranging from 1 million to 100 million genes per gram.[6] Knapp and co-workers have shown that the

ARGs in all categories of soils tested have increased significantly since 1940.[7] Antibiotics can enter the environment through various pathways including discharges by municipal sewages, antibiotic production sites, and by animal husbandry, landfill leachates of antibiotic disposals, and runoffs from agricultural field containing livestock manure, aquaculture ponds, and urban centers.[8] The entered antibiotics will lead to an accumulation of antibiotic resistance genes in the environment, moreover, antibiotic residues even at low concentrations would exert a selective pressure on bacteria. They can promote the generation of drug-resistant bacteria and their pathogenic microorganisms,[9–11] and facilitate the transmissions among different microbial populations.[12–15]

Chemical pollutants, such as pesticides,[16] antibiotics,[17,18] disinfectants[19] and heavy metal[20] have a significant impact on the environmental microbial community and diversity. Ecological toxicity caused by antibiotic residues and drug-resistant gene pollution in the environment affects the community structures of environmental microorganisms and

Hebei Key Laboratory of Environment and Human Health, School of Public Health, Hebei Medical University, Shijiazhuang, 050017, PR China. E-mail: hongmshi@163.com; xuxd@hebmu.edu.cn

their activities.[9,21] Several studies have shown that microbial communities affect the generation and spread of antibiotic resistance genes.[22,23] Luo *et al.* reported a correlation between microbial communities and antibiotic resistance genes in polluted rivers, and the spread of resistance genes was only related to the specific bacterial population.[24] Liu *et al.*[25] have shown that the abundance of potential pathogen was reduced after animal manure composting. The abundances of antibiotic resistance genes of a few sulfonamides, tetracycline, and β-lactam were decreased, showing a significant correlation. In recent years, industrial discharges of antibiotics from their manufacturing sites have been recognized as a risk factor fostering the development and dissemination of ARGs and antibiotic-resistant bacterial pathogens.[26] Drug-resistant bacteria in the environment have a greater impact on the spread of different antibiotic resistance genes,[27] it is thus of great significance to study the microbial community structures and biodiversity in the environment surrounding an antibiotic manufacturing site.

There are countless microorganisms in nature, and cultivable microorganisms account for only a very small portion of them. Traditional methods of microbial culture are difficult to operate, poorly controllable, time-consuming and labour-intensive, and can only cultivate a very small part of the known and common flora. In the 1970s and 1980s, the first-generation of sequencing technology could only read one read. The second-generation of sequencing technology emerged in 2005, which can read millions of reads typically restricted to 100–500 nucleotide regions, such as single variable regions and combinations of variable regions of the bacterial 16S rRNA gene.[28–30] However, the short read approach introduces biases and can't provide effective resolution below the bacterial genus level. Until recently, an alternative for full-length bacterial 16S rRNA gene sequencing is provided by the third-generation single-molecule real-time sequencing technology. By this technology, the study of microbial flora diversity has been greatly promoted.[31,32]

China is both a major producer and a major consumer of antibiotics. Shijiazhuang, the capital city of Hebei Province, is an antibiotic production base of China; many antibiotic drugs are produced in a great amount and are distributed here. A manufacturer located in the urban area of Shijiazhuang has produced a variety of veterinary antibiotics for more than 20 years. Effluents have been discharged into the nearby Minxin River by this manufacturer. Waste water and gas discharged from this manufacturer are one of the major sources for the environmental antibiotic residues. In general the low-dose antibiotic residues discharged from antibiotic production sites may affect the dynamic balance of bacterial populations.

In this work, the total DNA extracted from the environmental samples around the veterinary antibiotic manufactural site was sequenced for three generations to analyse the composition of microbial community and the diversity of the populations. The analysed data, together with some previous results obtained for the abundance of antibiotic resistant genes enabled us to deeply examine the relationships among the bacterial community composition, antibiotic resistant genes and antibiotic residues.

2. Experimental

2.1. Reagents and instrumentation

Genomic DNA extraction kits were obtained from OMEGA Bio-Tek, USA. Heal force ultrapure water machine (Likang Biomedical Technology) was used to make the ultrapure water used in this work. The third-generation sequencing machine was from Biomarker Technologies Co, Ltd (Beijing, China).

2.2. Sample collection processes

Water samples from upstream and downstream of the sewage outlets of Minxin River were collected in October, December, and March of 2019, respectively. Four sampling points were set for the water sample collections. As shown in Fig. 1, the P1 point was in the upstream of the river; the P2 point was near the sewage outlet and in the middle stream; the P3 point was set at the sewage outlet; the P4 point was in the downstream. The straight-line distance between P1 and P2 was about 1500 m, and P2 and P3 was about 50 m. P3 point was at the sewage outlet, and P3 and P4 was about 50 m. At each sampling point, water sample was collected continuously for 5 days in different sampling times. A total of 1 L of water (at 0.5 m below the water surface) was collected in a light-tight glass bottle, which was filtered through 0.22 μm membranes after the samples were transported to the laboratory.

Soil samples were collected in parallel with the sampling points of the river water samples. Two soil sampling points (points A and B in Fig. 1) were set up to observe the impact of the wastewater by the veterinary antibiotics factory on the surrounding environment. Point A was the control point, and was in parallel with water sample sampling point P1. Point B was located on the bank of the sewage outlet, and was in parallel with the water sample sampling point P3. Surface soil samples (at 10 cm below the soil surface) were collected in sterile bags, and stored at −20 °C. For a big of soil, the moisture content was measured; a certain amount of soil (about 50 g) was accurately weighed into a 50 mL centrifuge tube. The tube was put in a vacuum oven to dry the sample until a constant weight was reached.

Sampling principles of water and soil samples are based on HJ/T 91-2002 and GB/T 36179-2018 of China National Standard,
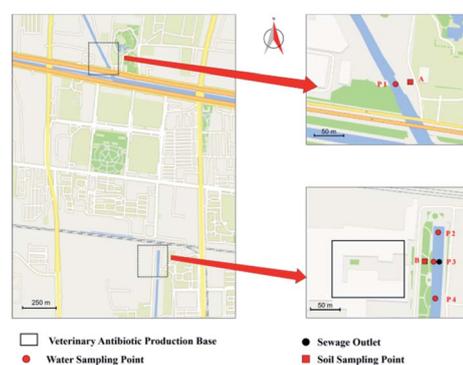

Fig. 1 Geographical illustration of the water and soil sampling site.

respectively. Water and soil samples after collection were all stored in a car refrigerator and then transported back to laboratory by driving within two hours.

2.3. DNA extraction, PCR amplification and the third-generation sequencing

DNA Kit was used in DNA extraction from water and soil. Filter membrane of water sample was removed to centrifuge tube and extracted DNA by an E.Z.N.A.™ Water DNA Kit. The soil sample (1.0 g) was placed in centrifuge tube, subjected to DNA extraction using an E.Z.N.A.™ soil DNA Kit. The extracted DNA was placed in a 1.5 mL EP tube and stored at −20 °C.

In PCR amplification, the primers (27F-5′-GTGCCAG-CAGCCGCGGTAA-3′, 1492-R-5′-GGACTACAAGGGTATCTAAT-3′) were used to amplify the full length of 16S rRNA, and the totally extracted DNA.[33] The best process of PCR amplification was optimized to be: pre-denaturation at 95 °C for 5 min; denaturation at 95 °C for 15 s; denaturation at 50 °C for 30 s; annealing for 1 min; extension at 72 °C for 90 s. The cycle repeats thirty times and ends in a 7 min extension at 72 °C, save it at 4 °C. Agarose gel (2%) electrophoresis was used for detections.

The DNA samples that had passed the screening test were sent to Beijing Biomarker Biotechnology Co., Ltd. The marker genes were sequenced by use of the method of single molecule real-time sequencing (SMRL Cell) which is based on the PacBio sequencing platform.

2.4. Data analysis

The data of all samples after single-molecule real-time sequencing were identified, filtered, and compared for obtaining the final valid data. Smartlink tool provided by the Pacbio platform was used to obtain the CCS (Circular Consensus Sequencing) sequence from the valid data. Lima v1.7.0 software was use to obtain Barcode-CCS sequence data after identifying the CCS through barcode, then filtered it by length and used UCHIME v4.2 software to identify and remove the chimera sequence. Finally, the optimization-CCS sequence was obtained. Usearch software[34] was used to cluster them at a similarity level of 97%, to divide OTU (Operational Taxonomic Unit), and to get its species classification based on the sequence composition of OTU. Taxonomic analysis was performed on the samples at each classification level to obtain the community structure diagram of each sample at the taxonomic level of phylum, class, order, family, genus, and species.[35] Through the alpha diversity analysis,[36] the diversity of species was studied within a single sample. By the beta diversity analysis, the species diversity of different samples was compared to find significant differences in species between different groups, and to look for biomarkers with statistical differences between different groups. LEfSe software was used to draw LEfSe graphs. The default setting of LDA score was four. species difference between groups was analyzed and predicted for 16S functional genes.[37,38] Finally, gene function of samples was predicted and the functional gene abundance was calculated.

3. Results and discussion

3.1. Population classification and identification

The raw data from PacBio sequencing platform was bam file. The CCS file was obtained by Smartlink tool, and then identified through barcode. After the barcode-CCS was filtered and the chimera was removed, filtered-CCS and optimization-CCS were obtained respectively. The effective (%) that the percentage of optimization-CCS in barcode-CCS, was used to evaluate data quality. The effective (%) of all samples had a mean ± SD of 93.09 ± 4.32%.

In this study, the compositions of bacterial communities of the environmental samples were different among sampling times and locations. As shown in Fig. 2, a total of 1720 OTUs were found, of which 1055 OTUs were in the river samples and 993 OTUs were in the soil samples. The OTU number in the river samples was less at P1 point than those at P2, P3, and P4 points, but the difference was very small. The OTU number at point B in the soil sample was significantly bigger than that at point A. The OTU numbers in the samples collected from river and soil sampling points in different times differed only slightly, but they all showed highest content in December, followed by October, and the least in March. The difference in the number of OTUs of water samples between different sampling times was greater than that of soil samples. A total of 29 bacterial phyla were found through the identification of these OTUs, of which 612 bacterial genera and 849 bacterial species were identified. The bacterial colonies in samples from different sampling points were distributed differently at the phylum, genus, and species level.

3.2. Species diversity index analysis

Bacterial Shannon index was used to indicate the diversity of bacterial community, the greater the value, the higher the diversity and the more uniform the individual distribution. Shannon index of the river water and soil samples was between 2.64 and 4.88 and between 5.04 and 6.02, respectively. A total of
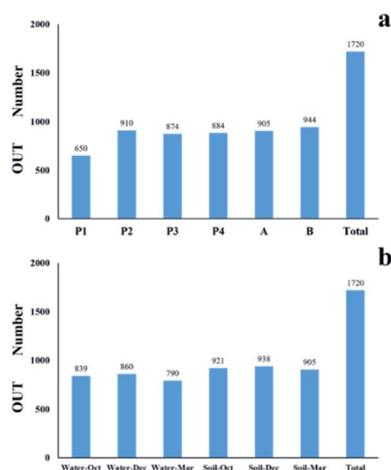

Fig. 2 OTU numbers acquired from the samples at different sampling sites (a) and at different sampling times (b).

643 bacterial strains were identified in the river water samples. Variance analysis showed that there were significant differences in the types of bacteria in river water samples between different sampling points ($F = 22.784$, $P < 0.001$). The number of strains in the upstream non-polluted area was significantly lower than that near the sewage outlet. Species diversity results showed that OTU, ACE and Chao1 were significantly different in among the sampling points ($P < 0.05$), namely, the species abundance of river samples in the downstream was higher than that in the upstream. A total of 438 bacterial strains were identified in the soil samples. Species abundances and diversity indexes between A and B points are all similar.

3.3. LEfSe analysis of microbial community

LEfSe analysis is an algorithm for the discovery of high-dimensional biomarkers and for the explanation that identifies genomic features characterizing the differences between biological conditions. In this work, LEfSe difference statistical analysis was performed on the above microbial communities, in order to find the possible biomarkers of microorganisms at different classification levels in samples between different groups. LEfSe analysis ($P < 0.05$, LDA > 4) showed that there were 61 biomarkers in the river water samples from different sampling points, including 8 phyla, 11 classes, 15 orders, 12 families, 10 genera and 10 species. As shown in Fig. 3, there were 31, 2, 9, and 19 biomarkers at P1, P2, P3, and P4 point, respectively. Fourteen biomarkers were found in soil samples, including 4 phyla, 2 classes, 3 orders, 2 families, 2 genera and 1

species. Among them, point A and point B had 11 and 3 biomarkers, respectively. The bacterial communities in the upstream non-polluted areas and other sampling points near the sewage outlet were quite different. To a certain extent, this indicates that there is a great difference between the upstream of pollution free and the surrounding environment of the veterinary drug production site. It is also noticeable from Fig. 4 that there are many red nodes representing samples in the non-pollution area; this suggests that these differences have a great contribution.

3.4. COG function notes

As shown in Fig. 5, all the soil and river water samples were annotated to 25 functionally classified genes, and among them, the relative abundances of genes classified into amino acid metabolism, carbohydrate metabolism, energy metabolism, global and overview maps, membrane transport, metabolism of cofactors and vitamins, were relatively higher.

Fig. 6 shows the functional genes whose relative content reached significant differences in the water samples from different sampling points. They were chromatin structure and dynamics, cell motility, secondary metabolites biosynthesis, transport and catabolism, RNA processing and modification, general function prediction only, signal transduction mechanisms, cell wall/membrane/envelope biogenesis, inorganic ion transport and metabolism. The functional genes with significant differences in relative content in soil samples from different locations were carbohydrate transport and metabolism, energy production and conversion, chromatin structure and dynamics, signal transduction mechanisms, nucleotide transport and metabolism, inorganic ion transport and metabolism, posttranslational modification, protein turnover, chaperones, coenzyme transport and metabolism, cell motility, amino acid transport and metabolism, and cell wall/membrane/envelope biogenesis.
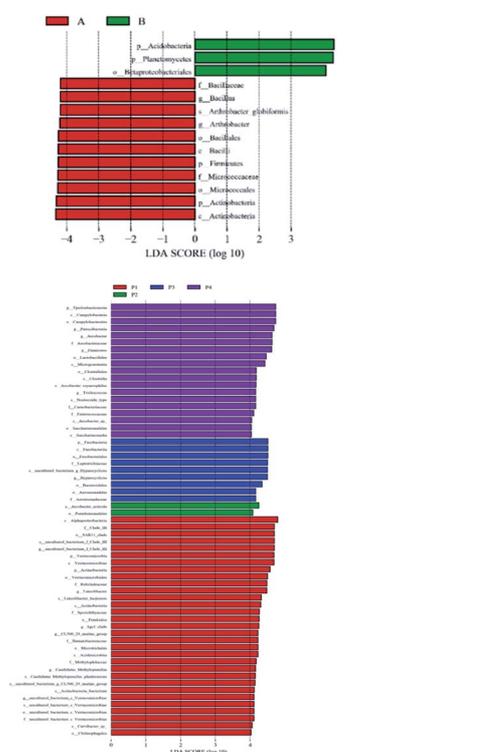

Fig. 3 Histogram of taxon with statistical difference at different sampling sites.
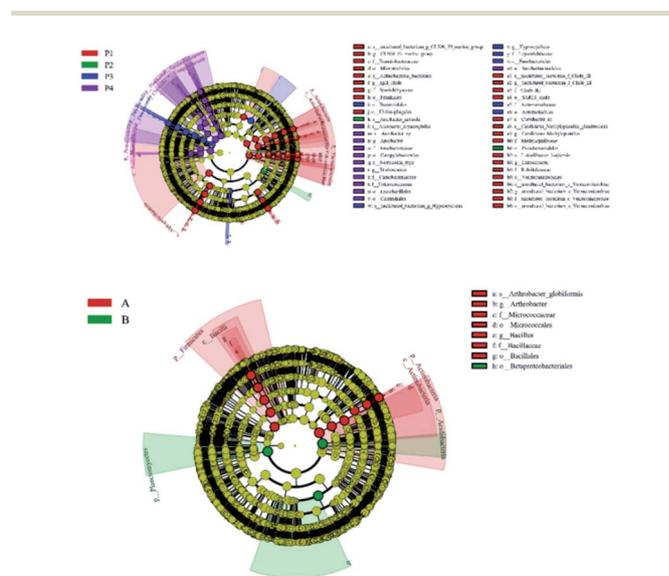

Fig. 4 Cladograms of taxon with statistical difference at different sampling sites.
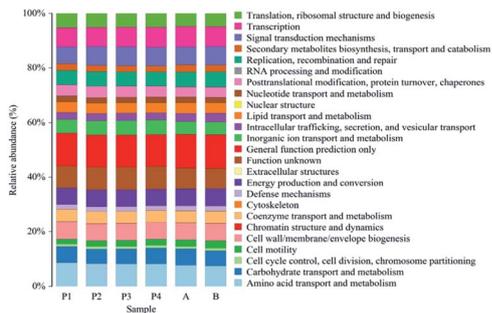
Fig. 5 The relative abundance of different functional classification genes.
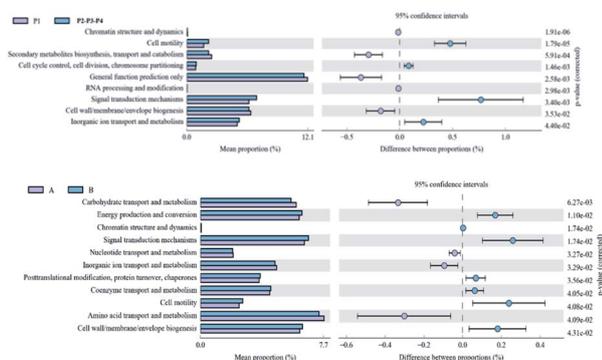

Fig. 6 Abundance of differential protein functions at different sampling sites.

3.5. Discussion

The amount of OTU in Minxin river water and soil nearby this veterinary antibiotic production site was significantly greater than that in the upstream non-polluted area; the Patescibacteria also had the same trend. The dominant bacteria phyla in the river water and soil samples collected in different times and in different sampling points were all Proteobacteria, and Proteobacteria was absolutely dominant the bacteria. This is consistent with other reports.[39–42] The diversity of bacterial communities in the river surrounding the veterinary drug production factory was significantly higher than that in the upstream non-polluted areas, which is consistent with the trend of changes in the abundance of antibiotic resistance genes.[43] Several studies[44–46] proved that the transportation of antibiotic-resistant bacteria (ARB) and ARGs in rivers increased the abundance of downstream ARGs. The higher the number of bacteria and the diversity of bacterial communities, the higher the abundance of antibiotic resistance genes was likely to be.

The upstream of the river surrounding the production factory contained the least amount of biomarkers, and the downstream contained the most, and an increasing trend was shown in turn. The abundances of ARGs from upstream to downstream also increased sequentially.[43] It is speculated that discharge from this veterinary drug production site has a certain impact on the distribution of bacterial flora in the surrounding river and soil, which in turn affect the generation

and enrichment of antibiotic resistance genes. Although the wastewater discharged from the antibiotic production facilities has been treated, a higher level of antibiotics is still remained. Improper treatments can lead to the generation of a large number of antibiotic-resistant bacteria.[10,47] If a large number of drug-resistant pathogens are not effectively treated, once they enter the food chain, they will pose a serious threat to human health.

Through the difference analysis, the differences and changes of the functional genes of microbial community in the metabolic pathways between sampling points near the veterinary drug production site and in the upstream non-contaminated area could be observed. The microbial community had changed its metabolic function in order to adapt to the environmental changes. The content of two functional genes, cell motility and signal transduction mechanisms in the river and soil samples had the same trend, namely, the environmental media around the veterinary drug production factory > upstream non-pollution area. This phenomenon was the same as that of multiple antibiotics. High signal conductivity means a high cell activity, *i.e.*, a high activity of microbial populations have a strongly reproductive ability and high species abundance. The high cell motility indicates that the microorganism has a strong metabolic function and high energy exchange with the external environment, which enables the microbial population multiplication. It further illustrates that the veterinary drug production site promote the horizontal transfer of antibiotic resistance genes, thereby promoting their spread, and microbial populations in the surrounding environment of the increasing the potential harm to the environment.

4. Conclusion

The dominant bacteria phyla in the river water and soil sample are proteobacteria, absolutely dominating bacteria phyla. The abundance and diversity of microorganisms near the sewage outlet was significantly higher than that of the upstream non-polluted area. A variety of different biomarkers were found in the samples from the two regions, and the differences in the abundance of two functional genes, cell motility and signal transduction mechanism, were also found in the metabolic pathways of microbial communities. This is similar to the abundance of multiple antibiotic resistance genes detected before,[43] which implies that wastewater discharged from the sewage outlet of the veterinary antibiotic production factory, to a certain degree, is the pollution source of antibiotic residues in the local environment, and has a certain ecological toxicity to environmental microorganisms.

The abundance and diversity of bacteria in the downstream area of the river water is also significantly higher than those in the upstream area; significant differences between multiple functional genes were also found in the metabolic pathway. The microbial populations in the surrounding environment of the veterinary drug factory are probably related to the antibiotic resistant gene and antibiotic residues, and may promote the horizontal transfer of the antibiotic resistant genes.

Author contributions

Junjie Miao and Zhendong Yin were responsible for methodology; Yuqin Yang and Yiwen Liang were in charge of validation; Xiangdong Xu performed analysis; Junjie Miao was responsible for writing the original draft of the manuscript; and Hongmei Shi was in charge of the research process and of the manuscript revisions.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 D. G. Joakim Larsson, *Upsala J. Med. Sci.*, 2014, **119**(2), 108–112.

2 L. H. Gao, Y. L. Shi, W. H. Li, J. M. Liu and Y. Q. Cai, *Environ. Chem.*, 2013, **32**(9), 1619–1633.

3 Y. J. Ben, C. X. Fu, M. Hu, L. Liu, M. H. Wong and C. M. Zheng, *Environ. Res.*, 2019, **169**, 483–493.

4 Y. Luo and Q. X. Zhou, *Acta Sci. Circumstantiae*, 2008, **28**(8), 1499–1505.

5 J. L. Martinez, *Environ. Pollut.*, 2009, **157**(11), 2893–2902.

6 Y. G. Zhu, Y. Zhao, B. Li, C. L. Huang, S. Y. Zhang, S. Yu, Y. S. Chen, T. Zhang, M. R. Gillings and J. Q. Su, *Nat. Microbiol.*, 2017, **2**, 16270.

7 C. W. Knapp, J. Dolfing, P. A. I. Ehlert and D. W. Graham, *Environ. Sci. Technol.*, 2010, **44**(2), 580–587.

8 N. J. Ashbolt, A. Amezquita, T. Bachhaus, P. Borriello, K. K. Brandt, P. Collignon, A. Coors, R. Finley, W. H. Gaze and T. Heberer, *Environ. Health Perspect.*, 2013, **121**, 993.

9 Y. G. Xu, W. T. Yu, Q. Ma, H. Zhou and C. M. Jiang, *Asian J. Ecotoxicol.*, 2015, **10**, 11–27.

10 H. Fang, L. X. Han, H. P. Zhang, Z. N. Long, L. Cai and Y. L. Yu, *J. Hazard. Mater.*, 2018, **357**, 53–62.

11 A. Pruden, *Environ. Sci. Technol.*, 2014, **48**, 5–14.

12 I. Lekunberri, M. Villagrasa, J. L. Balcázar and C. M. Borrego, *Sci. Total Environ.*, 2017, **601**, 206–209.

13 Y. W. Yang, Z. X. Liu, S. C. Xing and X. D. Liao, *Ecotoxicol. Environ. Saf.*, 2019, **182**, 109452.

14 R. N. Wang, Y. Zhang, Z. H. Cao, X. Y. Wang, B. Ma, W. B. Wu, N. Hu, Z. Y. Huo and Q. B. Yuan, *Sci. Total Environ.*, 2019, **651**, 1946–1957.

15 Y. H. Huang, Y. Liu, P. P. Du, L. J. Zeng, C. H. Mo, Y. W. Li, H. X. Lu and Q. Y. Cai, *Sci. Total Environ.*, 2019, **670**, 170–180.

16 C. Christiana Egbe, G. Oladunjoye Oyetibo and M. Olusoji Ilori, *Ecotoxicol. Environ. Saf.*, 2021, **207**, 111319.

17 L. L. Zhang, C. Zhang, K. T. Lian and C. X. Liu, *J. Hazard. Mater.*, 2021, **416**, 126141.

18 L. L. Qiu, T. J. Daniell, S. A. Banwart, M. Nafees, J. J. Wu, W. C. Du, Y. Yin and H. Y. Guo, *J. Hazard. Mater.*, 2021, **419**, 126388.

19 M. Tandukar, S. Oh, U. Tezel, K. T. Konstantinidis and S. G. Pavlostathis, *Environ. Sci. Technol.*, 2013, **47**, 9730–9738.

20 F. Wang, W. Y. Dong, Z. L. Zhao, H. J. Wang, W. Z. Li, G. H. Chen, F. F. Wang, Y. Zhao, J. Huang and T. Zhou, *Sci. Total Environ.*, 2021, **778**, 146383.

21 Z. Y. Chen, W. Zhang, L. X. Yang, R. D. Stedtfeld, A. P. Peng, C. Gu, S. A. Boyd and L. Lui, *Environ. Pollut.*, 2019, **248**, 947–957.

22 Y. Yang, B. Li, S. Zou, S. C. Zou, H. P. Fang Herbert and T. Zhang, *Water Res.*, 2014, **62**, 97–106.

23 F. Z. Gao, L. Y. He, L. X. He, H. Y. Zou, M. Zhang, D. L. Wu, Y. S. Liu, Y. J. Shi, H. Bai and G. G. Ying, *Sci. Total Environ.*, 2020, **741**, 140482.

24 X. Luo, W. L. Zhang, L. X. Yuan, M. Xu, L. He, Y. F. Jiang, W. Z. Zhong and F. Zhang, *China Environ. Sci.*, 2019, **39**, 2606–2613.

25 Y. W. Liu, D. M. Cheng, J. M. Xue, L. Weaver, S. A. Wakelin, Y. Feng and Z. J. Li, *J. Hazard. Mater.*, 2020, **389**, 122082.

26 J. J. González-Plaza, K. Blau, M. Milaković, T. Jurina, K. Smalla and N. Udiković-Kolić, *Environ. Int.*, 2019, **130**(9), 104735.

27 B. Zhang, L. W. Wu and X. H. Wen, *Environ. Sci.*, 2019, **40**(8), 3699–3705.

28 B. Reddy and S. K. Dubey, *Environ. Pollut.*, 2018, **246**, 443–451.

29 T. M. Yuan, A. J. McCarthy, Y. X. Zhang and R. Sekar, *Curr. Microbiol.*, 2020, **77**, 3512–3525.

30 L. Song, Z. Z. Pan, Y. Dai, L. Chen, L. Zhang, Q. L. Liao, X. Z. Yu, H. Y. Guo and G. S. Zhou, *Environ. Sci. Pollut. Res.*, 2021, **28**, 47086–47098.

31 L. Tedersoo, S. Anslan, M. Bahram, U. Kõljalg and K. Abarenkov, *Fungal Divers.*, 2020, **103**, 1–21.

32 K. Balázs, T. Dóra, B. Zsolt, M. Norbert, C. Zsolt, T. Gábor, M. Klára, S. Michael and B. Zsolt, *Sci. Rep.*, 2021, **11**, 14487.

33 L. P. Ma, B. Li, X. T. Jiang, Y. L. Wang, Y. Xia, A. D. Li and T. Zhang, *Microbiome*, 2017, **5**(1), 154.

34 R. C. Edgar, *Nat. Methods*, 2013, **10**(10), 996–998.

35 D. H. Huson, A. F. Auch, J. Qi and S. C. Scuuster, *Genome Res.*, 2007, **17**(3), 377–386.

36 A. Grice Elizabeth, H. Kong Heidi, S. Conlan, B. Deming Clayton, J. Davis, C. Young Alice, G. Bouffard Gerard, W. Blakesley Robert, R. Murray Patrick, D. Green Eric, L. Turner Maria and A. Segre Julia, *Science*, 2009, **324**, 1190–1192.

37 J. R. White, N. Nagarajan and M. Pop, *PLoS Comput. Biol.*, 2009, **5**(4), e1000352.

38 D. H. Parks, G. W. Tyson, H. Philip and R. G. Beiko, *Bioinformatics*, 2014, **30**, 3123–3124.

39 J. H. Guo, J. Li, H. Chen, P. L. Bond and Z. G. Yuan, *Water Res.*, 2017, **123**, 468–478.

40 A. Behnami, D. Farajzadeh, S. Isazadeh, K. Z. Benis, M. Shakerkhatibi, Z. Shiri, S. Ghorghanlu and A. Yadeghari, *J. Water Process. Eng.*, 2018, **23**, 285–291.

41 A. A. Elokil, M. Magdy, S. Melak, H. Ishfaq, A. Bhuiyan, L. Cui, M. Jamil, S. Zhao and S. Li, *Animal*, 2019, **14**, 706–715.

42 B. J. Phiri, T. S. Hayman David, J. Biggs Patrick, P. French Nigel and R. J. C. Garcia, *N. Z. J. Zool.*, 2021, **48**, 188–201.

43 J. J. Miao, Z. D. Yin, Y. Q. Yang, Y. W. Liang, X. D. Xu and H. M. Shi, *Antibiotics*, 2021, **10**, 1361.

44 J. P. Brooks, A. Adeli and M. R. McLaughlin, *Water Res.*, 2014, **57**, 96–103.

45 H. Y. Chen, X. M. Bai, Y. Z. Li, L. Y. Jing and R. H. Chen, *Sci. Total Environ.*, 2019, **679**, 88–96.

46 H. Y. Chen, X. M. Bai, Y. Z. Li, L. J. Jing, R. H. Chen and Y. G. Teng, *Water Res.*, 2019, **167**, 115098.

47 A. Bielen, A. Šimatović, J. Kosić-Vukšić, I. Senta, M. Ahel, S. Babić, T. Jurina, J. J. G. Plaza, M. Milaković and N. Udiković-Kolić, *Water Res.*, 2017, **126**, 79–87.