High Prevalence of Colistin Resistance Genes in German Municipal Wastewater

Kneis, David[a*,b], Berendonk, Thomas U.[a], Heß, Stefanie[a,c]

[a] Institute of Hydrobiology, TU Dresden, Germany

[b] Helmholtz Centre for Environmental Research – UFZ, Magdeburg, Germany

[c] Dept. of Microbiology, University of Helsinki, Finland

* Corresponding author: david.kneis@tu-dresden.de

Abstract

Bacterial resistance against the last-resort antibiotic colistin is of increasing concern on a global scale. Wastewater is suspected to be one of the pathways by which resistant bacteria and the respective genes are disseminated. We employed a metagenomics approach to detect and quantify colistin resistance genes in raw municipal wastewater sampled at 9 locations all over Germany (14 samples in total, collected in 2016/2017). Our data support the findings of earlier studies according to which the prevalence of the colistin resistance gene *mcr*-1 is still low. However, we were able to demonstrate that the total prevalence of colistin resistance genes is dramatically underestimated if the focus is put on that specific gene alone. In comparison to *mcr*-1, other gene variants like *mcr*-3 and *mcr*-7 proved to be 10 to 100 times more abundant in samples of untreated wastewater. The average relative abundances expressed as copies per 16S rRNA gene copies were 2.3×10^{-3} for *mcr*-3, 2.2×10^{-4} for *mcr*-4, 3.0×10^{-4} for *mcr*-5, and 4.4×10^{-4} for *mcr*-7. While these four gene variants were ubiquitous in all 14 samples, *mcr*-1 was detected only once at a relative abundance of 1.4×10^{-5}. Our results suggest a high risk of increasing incidence of colistin resistance as large amounts of *mcr* genes are continuously disseminated to diverse microbial communities via the wastewater path.

Keywords: antibiotic resistance genes, colistin, *mcr*, wastewater, Germany, metagenomics

1 Introduction

Polymyxin E (aka colistin) is a polypeptide antibiotic with bactericidal effects on Gram negatives. Although being available for about 60 years, this drug has never been a first-choice antibiotic for human therapy owing to its neuro- and nephrotoxicity (Spapen et al., 2011). In spite of such side effects, the rapid proliferation of bacterial resistances against other classes of antimicrobials has pushed colistin on the current list of last-resort antibiotics (World Health Organization, 2017). Reports on the frequent detection of *mcr*-1, a gene conveying resistance against colistin, in municipal wastewater are thus alarming (Hembach et al., 2017; Lekunberri et al., 2017). These and other studies that screened wastewater for resistance genes, e.g. Pärnänen et al. (2019), relied on real-time PCR (qPCR) which is highly sensitive but has a narrow focus dictated by the chosen primers. In this work, we employed a metagenomics approach to search for colistin resistance genes in raw municipal wastewater. Although being less sensitive than qPCR, metagenomics allows samples to be scanned for a multitude of antibiotic resistance genes (ARG) including those *mcr* gene variants that were first described only recently (e.g. Yin et al., 2017; Yang et al., 2018).

2 Material and methods

Raw wastewater was collected in 2016/2017 at nine locations distributed over Germany. Four of the sites were visited multiple times with a delay of several weeks (total $n = 14$; see Table 1 for details). Samples were stored in 1 L sterile glass bottles at 4 °C and DNA was extracted within 24 h using the PowerWaterKit (MoBio, Vancouver, Canada) from volumes of 100 mL (highly turbid samples) or 200 mL (less turbid samples). DNA samples were shotgun-sequenced (2×150 bp) on a MiSeq device (GATC Biotech AG, Konstanz, Germany) and the quality checked forward reads were aligned to the latest (2019-03-05) Resfinder data base (Zankari et al., 2012). To make ARG prevalences comparable across samples and studies, we report relative gene abundances instead of absolute counts. Hence, we scaled the number of ARG copies in a sample by the corresponding number of 16S rRNA gene copies, the latter being a proxy for total bacterial abundance. We employed METAXA2 (Bengtsson-Palme et al., 2015, version 2.1.3) to quantify 16S

rRNA genes. Data of replicate samples taken at different points in time (Table 1) were pooled.

qPCR-based estimates of resistance gene abundances were obtained according to the protocol of Heß et al. (2018). Relevant references for additional primers are Hembach et al. (2017) (*mcr*-1) and Peak et al. (2007) (*tet*M).

Data analysis was conducted with R version 3.4.4 (R Core Team, 2017). Confidence intervals for observed relative gene abundances were computed with `binom.test`. Quantiles of the binomial distribution forming the basis of Fig. 2 were computed with `qbinom`. Pearson and rank-based correlation coefficients were calculated with `cor` choosing method "pearson" or "spearman", respectively.

Table 1: Characteristics of the analyzed samples of raw wastewater collected throughout Germany. Subscripts in site codes denote individual sewers serving the same treatment plant. All plants perform secondary treatment combined with nitrogen and phosphorus removal.

Site code	Region	Plant capacity (popul. eq.)	Treated vol. (m^3 d^{-1})	Temporal replicates	Analyzed reads ($\approx$ 150 bp each)
A	South	70,000	12,000	2	37–60$\times 10^6$
B	East	100,000	26,000	3	45–89$\times 10^6$
C	South	180,000	22,000	2	41–44$\times 10^6$
D_1	North	2,220,000	350,000	1	37$\times 10^6$
D_2	North	2,220,000	350,000	1	59$\times 10^6$
E	West	470,000	50,000	1	58$\times 10^6$
F_1	West	1,300,000	220,000	1	45$\times 10^6$
F_2	West	1,300,000	220,000	1	48$\times 10^6$
G	East	300,000	80,000	2	38–68$\times 10^6$

3 Results and discussion

In accordance with expectation from earlier qPCR-based studies (Hembach et al., 2017; Lekunberri et al., 2017), the relative abundance of the *mcr*-1 gene in untreated municipal wastewater was low. In fact, we detected *mcr*-1 in only one out of 14 samples (site B) and the estimated relative abundance was less than 10^{-4} copies (16S rRNA gene copies)$^{-1}$ with 95% confidence (Fig. 1). However, we observed unexpectedly high relative abundances of other recently discovered colistin resistance gene variants. These include *mcr*-3 (Yin et al., 2017), *mcr*-4 (Carattoli et al., 2017), *mcr*-5 (Borowiak

et al., 2017), and *mcr*-7 (Yang et al., 2018). Each of these gene variants was consistently detected in all samples including those taken on different dates at the sites A, B, C, and G (Table 2). Most notably, we found relative abundances of *mcr*-3 genes in the range of 10^{-3} to 10^{-2} copies (16S rRNA gene copies)$^{-1}$ at seven of nine sampling sites. The overall relative abundance of colistin resistance genes (3.2×10^{-3} copies (16S rRNA gene copies)$^{-1}$ on average) was on par with genotypic resistances against more common classes of antibiotics like phenicols (2.4×10^{-3}) or quinolones (3.3×10^{-3}).

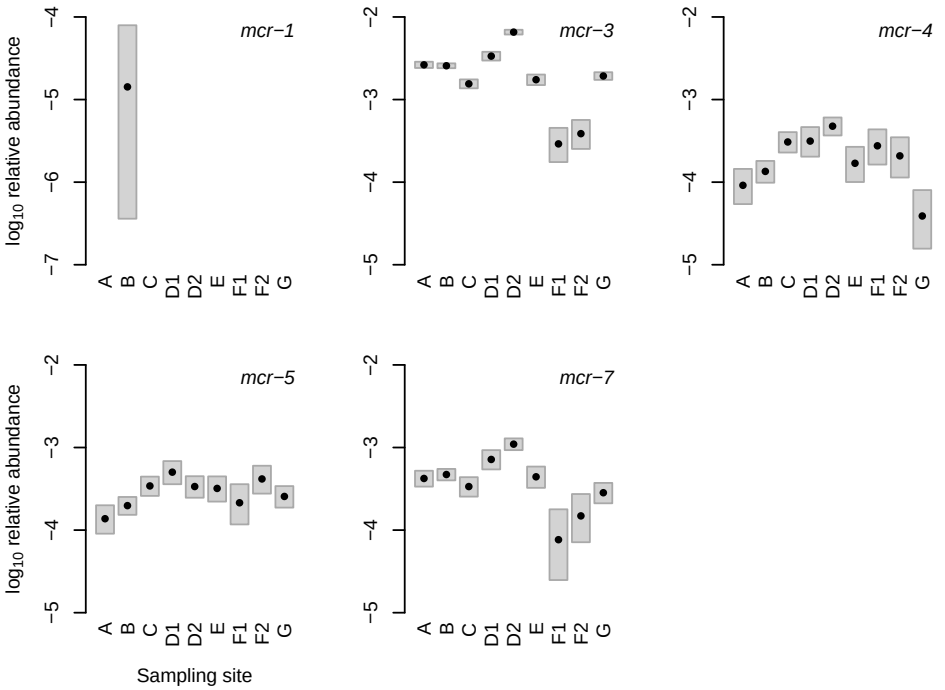

Figure 1: Prevalence of *mcr* genes in raw wastewater sampled at sites A–G (Tab. 1). Dots indicate observed values, boxes represent 95% confidence intervals. Note that scales differ between *mcr*-1 and the other gene variants.

.

For the purpose of validation, we checked the BLAST e-values serving as an indicator for the quality of sequence alignments. We found that over 90% of the hits were associated with e-values $< 10^{-29}$ for *mcr*-3, $< 10^{-68}$ for *mcr*-4, $< 10^{-59}$ for *mcr*-5, and $< 10^{-7}$ for *mcr*-7, indicating low probabilities of false positives. Additionally, we compared ARG counts delivered by the metagenomics approach with those obtained by qPCR for a set of 14 resistance genes. In the single instance where *mcr*-1 was detected with the

Table 2: Frequency of detection and average relative abundance of *mcr* gene variants in raw wastewater sampled at sites A–G.

Gene variant	Positive samples	Average relative abundance (copies per 16S rRNA gene copies)
mcr-1	1/14	$<1.4\times10^{-5}$
mcr-3	14/14	2.3×10^{-3}
mcr-4	14/14	2.2×10^{-4}
mcr-5	14/14	3.0×10^{-4}
mcr-7	14/14	4.4×10^{-4}

metagenomics approach, the estimated relative abundance was compatible with the qPCR measurement of 6.6 $\times10^{-6}$ copies (16S rRNA gene copies)$^{-1}$. For more common ARG, the metagenomics-based relative abundances correlated very well with their qPCR-based equivalents. For example, the Pearson correlation coefficients were > 0.98 for *sul*1 and *tet*M and > 0.93 for *erm*B. The metagenomics- and qPCR-based estimates generally matched well when the relative abundance of the gene of interest was 10^{-3} or above. In those cases, the deviation between the two methods' results hardly exceeded factor 5. For rare genes with relative abundances of 10^{-4} to 10^{-5} we still observed good matches but, at the same time, the metagenomics approach produced an increasing amount zero values resulting in an overall negative bias. This is in accordance with probability theory: Considering $\approx10^5$ 16S rRNA gene copies per sample, the binomial distribution model predicts a notable tendency towards underestimation as the relative abundance of the gene of interest falls below a threshold of about 5×10^{-5} (Fig. 2).

Importantly, a marked overestimation of ARG abundances by the metagenomics approach was not observed for any of the 14 genes measured with qPCR. At the same time, theory suggests that metagenomics-based estimates are rather prone to negative than to positive bias (Fig. 2). Therefore, the relative gene abundances reported in Fig. 1 represent conservative estimates in the sense that the true prevalences could be even higher.

In spite of the fact that most sites were sampled only once, Fig. 1 shows a largely consistent ranking of the different *mcr* gene variants in terms of their relative abundance. Namely the pattern *mcr*-3 > *mcr*-7 > *mcr*-1 holds for all sites. Differences in the relative abundance of particular gene variants might reflect contrasting prevalences in the bacterial community as a whole.

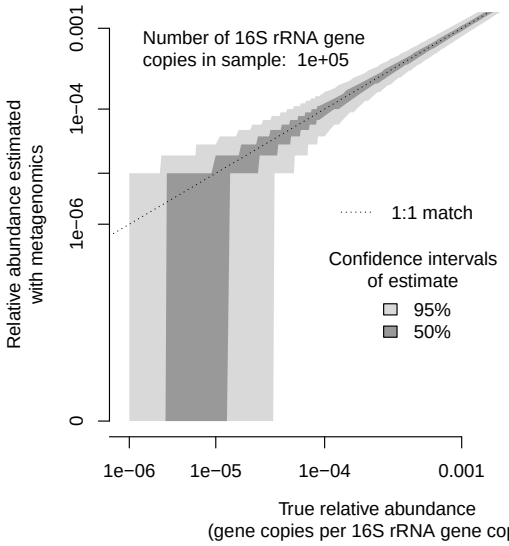

Figure 2: Expected accuracy of the metagenomics approach. The assumed number of 16S rRNA gene copies (10^5) is representative for the samples listed in Table 1.

.

Alternatively, the ranking could be explained by contrasting abundances of particular strains harboring different variants of *mcr*. A rank-based correlation analysis revealed that the occurrence of *mcr*-7 is closely linked to that of *mcr*-3 (Spearman's $\rho = 0.88$, $p < 0.01$). A weaker correlation was found between *mcr*-4 and *mcr*-5 ($\rho = 0.67$, $p = 0.06$). The co-occurrence of the respective *mcr* gene variants in the genome of a specific strain provides a possible explanation. One would expect a particularly close correlation if the respective gene variants co-occur on mobile genetic element that intrude a substantial proportion of the bacterial community via horizontal gene transfer.

The differences between sampling locations observed for specific gene variants (Fig. 1) cannot be sufficiently explained out of the current data set. For example, a significant relationship between the relative abundances of *mcr* genes and the capacity of the receiving plant could not be established. Neither is the geographic region alone a suitable predictor: E.g. sampling sites A and B exhibit similar values for all gene variants although they are about 350 km away from each other. Nevertheless, Pärnänen et al. (2019) have demonstrated on a much larger data set that the abundance of resis-

tance genes in wastewater mirrors the situation in the source area in terms of, e.g. antibiotic consumption and the prevalence of resistant bacteria. From that point of view, a hot spot of colistin resistance would be expected in the source area of the sampling sites D1 and D2. Among the visited location, these two sites exhibit the highest relative abundances for most *mcr* gene variants.

Our study supports the opinion of Bardet and Rolain (2018) according to which combined methodologies that include molecular and genomic techniques are needed to track the prevalence and proliferation of colistin resistances. The specific advantage of metagenomics lies in its capability to detect and quantify the full spectrum of ARG with known signatures. This is in contrast to PCR-based approaches (see Rebelo et al., 2018) which are superior in terms of sensitivity but limited with respect to the set of detectable ARG.

Our data suggest that the overall prevalence of colistin resistance genes in raw wastewater is higher than previously known. Seen from a one-health perspective, our results call for immediate action since a massive, continuous dissemination of *mcr* genes is very likely to promote the emergence of phenotypic colistin resistance in potential pathogens. This is mainly because all currently known *mcr* gene variants have been found on mobile genetic elements (Borowiak et al., 2017; Carattoli et al., 2017; Lekunberri et al., 2017; Yin et al., 2017; Yang et al., 2018). Since we analyzed samples of raw wastewater only, we cannot make reliable statements about the impact on receiving water bodies. However, the elimination of bacteria and ARG in wastewater through conventional activated sludge treatment is known to be incomplete (Lüddeke et al., 2015; Pärnänen et al., 2019) and a significant discharge of *mcr* genes into natural waters is thus likely.

First of all, we propose *mcr* genes to be traced back along sewer systems to identify the relative contribution of sources, including households, health care facilities, livestock farming sites (see, e.g., Apostolakos and Piccirillo, 2018), and slaughterhouses. Secondly, we suggest to also study the prevalence of *mcr* in treated wastewater so as to quantify the gene loads that the receiving water bodies are confronted with. Finally, an extension of the study region appears to be a valuable complement to current European activities targeted at elucidating the occurrence, geographic distribution and population dynamics of colistin resistant isolates with genomic methods (see

European Centre for Disease Prevention and Control, 2018).

Acknowledgements

Funding for this study has been provided by the German Federal Ministry of Education and Research [grant number 02WRS1377D] and the German Research Foundation ("Support the Best" pool of TU Dresden). S.H. also acknowledges financial support by the German Research Foundation [grant number HE8047/1].

References

Apostolakos, I. and Piccirillo, A. (2018). A review on the current situation and challenges of colistin resistance in poultry production. *Avian Pathology*, 47(6):546–558.

Bardet, L. and Rolain, J.-M. (2018). Development of new tools to detect colistin-resistance among enterobacteriaceae strains. *Canadian Journal of Infectious Diseases and Medical Microbiology*, 2018:Article ID 3095249.

Bengtsson-Palme, J., Hartmann, M., Eriksson, K. M., Pal, C., Thorell, K., Larsson, D. G. J., and Nilsson, R. H. (2015). metaxa2: improved identification and taxonomic classification of small and large subunit rRNA in metagenomic data. *Molecular Ecology Resources*, 15(6):1403–1414.

Borowiak, M., Fischer, J., Hammerl, J. A., Hendriksen, R. S., Szabo, I., and Malorny, B. (2017). Identification of a novel transposon-associated phosphoethanolamine transferase gene, mcr-5, conferring colistin resistance in d-tartrate fermenting Salmonella enterica subsp. enterica serovar Paratyphi B. *Journal of Antimicrobial Chemotherapy*, 72(12):3317–3324.

Carattoli, A., Villa, L., Feudi, C., Curcio, L., Orsini, S., Luppi, A., Pezzotti, G., and Magistrali, C. (2017). Novel plasmid-mediated colistin resistance mcr-4 gene in Salmonella and Escherichia coli, Italy 2013, Spain and Belgium, 2015 to 2016. *Eurosurveillance*, 22(31):pii=30589.

European Centre for Disease Prevention and Control (2018). ECDC study protocol for genomic-based surveillance of carbapenem-resistant and/or

colistin-resistant Enterobacteriaceae at the EU level. Version 2.0. ECDC, Stockholm.

Hembach, N., Schmid, F., Alexander, J., Hiller, C., Rogall, E. T., and Schwartz, T. (2017). Occurrence of the mcr-1 Colistin Resistance Gene and other Clinically Relevant Antibiotic Resistance Genes in Microbial Populations at Different Municipal Wastewater Treatment Plants in Germany. *Frontiers in Microbiology*, 8:1282.

Heß, S., Berendonk, T., and Kneis, D. (2018). Antibiotic resistant bacteria and resistance genes in the bottom sediment of a small stream and the potential impact of remobilization. *FEMS Microbiology Ecology*, 94:fiy128.

Lekunberri, I., Balcázar, J. L., and Borrego, C. M. (2017). Detection and quantification of the plasmid-mediated mcr-1 gene conferring colistin resistance in wastewater. *International Journal of Antimicrobial Agents*, 50(6):734 – 736.

Lüddeke, F., Heß, S., Gallert, C., Winter, J., Güde, H., and Löffler, H. (2015). Removal of total and antibiotic resistant bacteria in advanced wastewater treatment by ozonation in combination with different filtering techniques. *Water Research*, 69:243–251.

Peak, N., Knapp, C. W., Yang, R. K., Hanfelt, M. M., Smith, M. S., Aga, D. S., and Graham, D. W. (2007). Abundance of six tetracycline resistance genes in wastewater lagoons at cattle feedlots with different antibiotic use strategies. *Environmental Microbiology*, 9(1):143–151.

Pärnänen, K., Narciso-da Rocha, C., Kneis, D., Berendonk, T., Cacace, D., Do, T., Elpers, C., Fatta-Kassinos, D., Henriques, I., Jaeger, T., Karkman, A., Martinez, J., Michael, S., Michael-Kordatou, I., O'Sullivan, K., Rodriguez-Mozaz, S., Schwartz, T., Sheng, H., Sørum, H., Stedtfeld, R., Tiedje, J., S.V., D., Walsh, F., Vaz-Moreira, I., Virta, M., and Manaia, C. (2019). Antibiotic resistance in european wastewater treatment plants mirrors the pattern of clinical antibiotic resistance prevalence. *Science Advances*, 5(3):eaau9124.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

231 Rebelo, A., Bortolaia, V., Kjeldgaard, J., Pedersen, S., Leekitcharoenphon,
232 P., Hansen, I., Guerra, B., Malorny, B., Borowiak, M., Hammerl, J., Bat-
233 tisti, A., Franco, A., Alba, P., Perrin-Guyomard, A., Granier, S., De Fru-
234 tos, E., Malhotra-Kumar, S., Villa, L., Carattoli, A., and Hendriksen, R.
235 (2018). Multiplex PCR for detection of plasmid-mediated colistin resis-
236 tance determinants, mcr-1, mcr-2, mcr-3, mcr-4 and mcr-5 for surveillance
237 purposes. *Eurosurveillance*, 23(6):pii=17–00672.

238 Spapen, H., Jacobs, R., Van Gorp, V., Troubleyn, J., and Honoré, P. M.
239 (2011). Renal and neurological side effects of colistin in critically ill pa-
240 tients. *Annals of Intensive Care*, 1(1):14.

241 World Health Organization (2017). WHO Model List of Essential
242 Medicines, 20th List. https://www.who.int/medicines/publications/
243 essentialmedicines/en/.

244 Yang, Y., Li, Y., Lei, C., Zhang, A., and Wang, H. (2018). Novel plasmid-
245 mediated colistin resistance gene mcr-7.1 in Klebsiella pneumoniae. *Jour-*
246 *nal of Antimicrobial Chemotherapy*, 73(7):1791–1795.

247 Yin, W., Li, H., Shen, Y., Liu, Z., Wang, S., Shen, Z., Zhang, R., Walsh,
248 T., Shen, J., and Wang, Y. (2017). Novel Plasmid-Mediated Colistin
249 Resistance Gene mcr-3 in Escherichia coli. *mBio*, 8(3):e00543–17.

250 Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S.,
251 Lund, O., Aarestrup, F. M., and Larsen, M. V. (2012). Identifica-
252 tion of acquired antimicrobial resistance genes. *Journal of Antimicrobial*
253 *Chemotherapy*, 67(11):2640–2644.