

OPEN

# Antimicrobial resistance genes in raw milk for human consumption

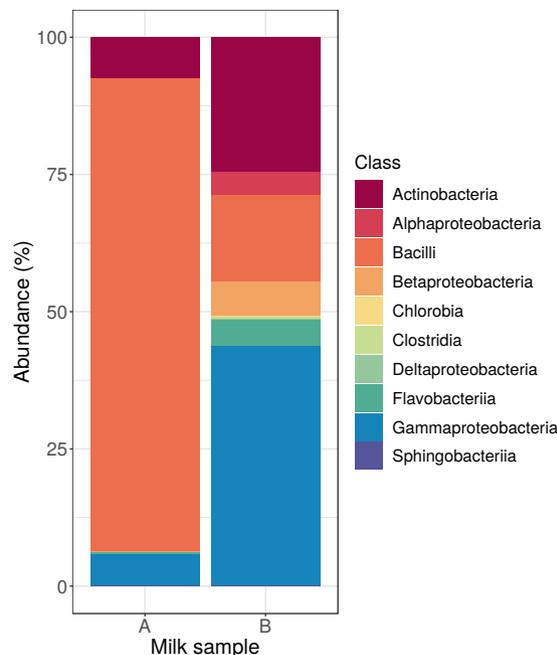
Adrienn Gréta Tóth<sup>1,2</sup>, István Csabai<sup>3</sup>, Eszter Krikó<sup>1</sup>, Dóra Tőzsér<sup>1,4</sup>, Gergely Maróti<sup>5</sup>, Árpád V. Patai<sup>6,7</sup>, László Makrai<sup>2</sup>, Géza Szita<sup>4</sup> & Norbert Solymosi<sup>1,3</sup> ✉

The increasing prevalence of antimicrobial resistance (AMR) is a significant threat to global health. More and more multi-drug-resistant bacterial strains cause life-threatening infections and the death of thousands of people each year. Beyond disease control animals are often given antibiotics for growth promotion or increased feed efficiency, which further increase the chance of the development of multi-resistant strains. After the consumption of unprocessed animal products, these strains may meet the human bacteriota. Among the foodborne and the human populations, antimicrobial resistance genes (ARGs) may be shared by horizontal gene transfer. This study aims to test the presence of antimicrobial resistance genes in milk metagenome, investigate their genetic position and their linkage to mobile genetic elements. We have analyzed raw milk samples from public markets sold for human consumption. The milk samples contained genetic material from various bacterial species and the in-depth analysis uncovered the presence of several antimicrobial resistance genes. The samples contained complete ARGs influencing the effectiveness of acridine dye, cephalosporin, cephamycin, fluoroquinolone, penam, peptide antibiotics and tetracycline. One of the ARGs, PC1 beta-lactamase may also be a mobile element that facilitates the transfer of resistance genes to other bacteria, e.g. to the ones living in the human gut.

The increasing prevalence of antimicrobial resistance (AMR) is a significant threat to global health. The widespread use of antibiotics, both in human healthcare and animal disease control<sup>1–3</sup>, is increasingly shortening the time it takes for resistant strains to develop and more and more multi-drug-resistant bacterial strains cause life-threatening infections. There is increasing evidence that antimicrobial resistance genes (ARGs) responsible for the occurrence of phenotypically expressed antimicrobial resistance are widespread in various environmental samples<sup>4–7</sup>. The pool of ARGs being present in a particular environmental sample is called the resistome<sup>8</sup>. In samples where the medical use of antibiotics can be excluded, normally, only a few ARGs are present<sup>9,10</sup>. When antibiotics are extensively used for preventive and therapeutic purposes, bacterial strains respond to this selective pressure and will become increasingly resistant what finally leads to these agents' elevated prevalence. ARGs hosted by non-pathogenic bacteria can be transferred to pathogens with horizontal gene transfer (HGT) what elevates the latter group's resistance against antibiotics. The execution of HGT depends on several factors, albeit physical closeness of bacteria always increases the chances<sup>11</sup>. The likelihood of HGT is even higher if ARGs are carried on mobile genetic elements (e.g. plasmids). In order to understand the chance of an ARG's horizontal gene transfer derived spread, studies aiming to assess the bacteria's resistome and the specific position of the identified ARGs<sup>12</sup> are very well-reasoned and necessary.

The microbiota of livestock products may come to direct contact with the human bacteriota, either during the processing steps or during the consumption of these products. The antibiotics used for farm animal disease control often share active substances with human medicines. Consequently, there is a risk that ARGs accumulated as a response to the high amount of antibiotics used in livestock farming may be transmitted to the human microbiota through animal products. Such spread of ARGs may reduce the efficacy of antibiotic therapies even more and may lead to the development of new multidrug-resistant strains. Fortunately, food processing typically

<sup>1</sup>University of Veterinary Medicine Budapest, Centre for Bioinformatics, Budapest, 1078, Hungary. <sup>2</sup>University of Veterinary Medicine Budapest, Department of Microbiology and Infectious Diseases, Budapest, 1143, Hungary. <sup>3</sup>Eötvös Loránd University, Department of Physics of Complex Systems, Budapest, 1117, Hungary. <sup>4</sup>University of Veterinary Medicine Budapest, Department of Food Hygiene, Budapest, 1078, Hungary. <sup>5</sup>Biological Research Center, Institute of Plant Biology, Szeged, 6726, Hungary. <sup>6</sup>Semmelweis University, 2nd Department of Internal Medicine, Budapest, 1088, Hungary. <sup>7</sup>Mayo Clinic, Gastroenterology Research Unit, Division of Gastroenterology and Hepatology, Department of Medicine, Rochester, MN, 55905, USA. ✉e-mail: [solymosi.norbert@gmail.com](mailto:solymosi.norbert@gmail.com)



**Figure 1.** Core bacteriome composition. Relative abundance of bacteria classes by milk samples.

contains heat treatment steps that kill the majority of bacteria. Thus the role of active DNA-export mechanisms between the intestinal and the nutriment's bacteriome is lower<sup>13</sup>.

Raw milk is a product sold unprocessed; thus the presence or the grade of heat-treatment steps are upon the decision of consumers. In addition to this, the consumption of non-heat-treated raw milk justified by its favourable health effects is nowadays commonly set as a trend in the developed countries<sup>14,15</sup>.

To our best knowledge, no previous study has investigated the possible presence of ARGs in raw milk, and we have found no data on the raw milk's resistome. This study aims to test the presence of ARGs in milk metagenome, investigate their genetic position and their linkage to mobile genetic elements<sup>16,17</sup>.

## Results

**Metagenome.** After DNA extraction and sequencing (see Methods), from sample A 17,773,004 while from sample B 8,425,326 paired-end reads were recovered. By the quality filtering, 0.20% and 0.80% of the reads were discarded from sample A and B, respectively. The reads were aligned to the host (*Bos taurus*) genome. As expected, most of the genetic material originated from the milking cow, from sample A 96.41% and from sample B 97.01% of the cleaned reads were filtered out due to host origin.

Of the reads, not aligning to the cow genome, we were able to classify 42.11% in sample A and 52.96% in sample B to known taxa. 185,982 reads of sample A and 11,437 reads of sample B were identified to belong to the kingdom of Bacteria. In sample A 93.54% of the reads were classified as Gram-positive bacteria, while in sample B this proportion was only 40.54%. The detailed composition of the core bacteriomes at class level is shown in Fig. 1.

**ARGs and MGEs.** Reads with overlapping pieces were assembled into longer DNA contigs by the metaSPAdes tool. The assembled contigs having a shorter length than 162 bp (sample A: 0.68%, sample B: 0.33% of all contigs) were excluded. The remaining contig's median length was 268 (IQR: 126.5) and 244 (IQR: 42) in sample A and B, respectively. Then, the contigs having any open reading frame (ORF) matched with an ARG in the Comprehensive Antibiotic Resistance Database (CARD) were collected. The detected ARGs and particular properties are presented in Table 1.

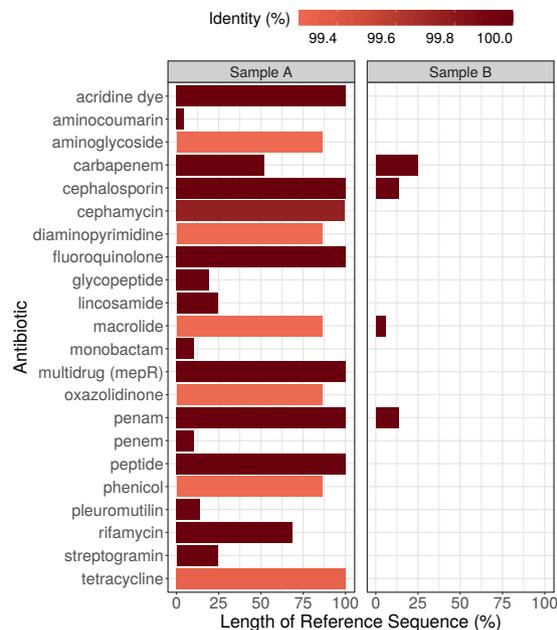
The identified ARGs were classified with the Resistance Gene Identifier (RGI) tool according to the ratio of their coverage in the samples and to the identity between the contigs assembled from the sequenced reads and the CARD ARG reference sequences. In Table 1 we list each ARG with perfect or strict hits predicted by RGI. We were able to identify three perfect ARG matches in sample A, *mepR*, *mgrA* and *Staphylococcus aureus* *norA*. According to the taxonomical classification of the contigs harbouring these ORFs their most likely origin is bacteria from *Staphylococcus* genus. The MGE analysis showed that none of these ORFs is mobile.

The sequence coverages of the strict matches in sample A ranged between 2.12% and 100%, with a mean of 36.61%. The identity of ORFs and CARD ARG reference sequences ranged between 95.02% and 100%, with a mean of 99.59%. Contigs containing ARG were classified on genus level and *Acinetobacter* (2.86%), *Carnobacterium* (11.43%), *Chryseobacterium* (2.86%), *Corynebacterium* (2.86%), *Kocuria* (11.43%), *Lactococcus* (8.57%), *Leuconostoc* (2.86%), *Macrococcus* (2.86%), *Moraxella* (5.71%), *Staphylococcus* (37.14%) and *Streptococcus* (11.43%) genera were identified.

In the bacterial genome, ARGs may be located on chromosomes or on plasmids, the latter ones being more likely to translocate between bacteria. With the PlasFlow tool, we identified contigs harbouring chloramphenicol acetyltransferase, PC1 beta-lactamase (*blaZ*) and *tet(38)* ARGs that may be encoded on plasmids. The results of

ARG	Coverage %	Identity %	Genus	Species	Localization	MGE
<b>Perfect RGI match in sample A</b>						
mepR	100.00	100.00	<i>Staphylococcus</i>	<i>aureus</i>	chromosome	
mgrA	100.00	100.00	<i>Staphylococcus</i>	<i>aureus</i>	unclassified	
<i>Staphylococcus aureus</i> norA	100.00	100.00	<i>Staphylococcus</i>	<i>aureus</i>	chromosome	
<b>Strict RGI match in sample A</b>						
AAC(6')-IIc	30.05	100.00	<i>Carnobacterium</i>	<i>maltaromaticum</i>		
<i>Acinetobacter baumannii</i> AbaQ	17.97	100.00	<i>Leuconostoc</i>	<i>mesenteroides</i>		
APH(2'')-Ig	29.74	100.00	<i>Chryseobacterium</i>			
APH(3'')-Ia	7.35	100.00	<i>Acinetobacter</i>	sp. TTH0-4		
APH(3')-Ia	8.12	100.00				
APH(7'')-Ia	13.25	100.00	<i>Lactococcus</i>	<i>raffinolactis</i>		
arlR	74.89	95.12	<i>Staphylococcus</i>	<i>aureus</i>		
arlR	30.14	98.48	<i>Staphylococcus</i>	<i>aureus</i>		
arlS	29.93	100.00	<i>Staphylococcus</i>	<i>aureus</i>		
arlS	70.95	99.69	<i>Staphylococcus</i>	<i>aureus</i>	chromosome	
baeS	4.71	100.00				
BUT-1	11.59	100.00	<i>Moraxella</i>	<i>osloensis</i>		
<i>Campylobacter coli</i>						
chloramphenicol acetyltransferase	52.17	100.00	<i>Lactococcus</i>	<i>raffinolactis</i>	plasmid	
CatU	11.98	100.00	<i>Streptococcus</i>	<i>thermophilus</i>		
cfr(B)	24.36	100.00	<i>Streptococcus</i>	<i>urinalis</i>		
DHA-1	99.75	99.75	<i>Staphylococcus</i>	<i>aureus</i>	chromosome	
ErmW	10.61	100.00				
ICR-Mo	28.32	98.10	<i>Moraxella</i>	<i>osloensis</i>		
<i>Klebsiella pneumoniae</i> KpnF	68.81	100.00	<i>Corynebacterium</i>	<i>provincense</i>		
MCR-3.2	12.75	100.00	<i>Kocuria</i>	sp. BT304		
mecD	11.80	100.00	<i>Macroccoccus</i>	<i>caseolyticus</i>		
mepA	100.00	99.78	<i>Staphylococcus</i>	<i>aureus</i>	chromosome	
mphM	27.09	100.00	<i>Carnobacterium</i>	<i>maltaromaticum</i>		
mphO	11.36	100.00	<i>Kocuria</i>			
MuxC	2.12	100.00				
norB	9.23	100.00				
OCH-2	10.26	100.00		<i>Brevibacterium</i>		
				phage Cantare		
PC1 beta-lactamase (blaZ)	100.00	95.02	<i>Staphylococcus</i>		plasmid	phage integrase
PEDO-1	20.98	100.00	<i>Lactococcus</i>	<i>lactis</i>		
PEDO-3	51.90	100.00	<i>Carnobacterium</i>	<i>maltaromaticum</i>		
QnrB42	20.09	100.00				
srmB	14.18	100.00	<i>Carnobacterium</i>	<i>maltaromaticum</i>		
<i>Staphylococcus aureus</i> LmrS	12.71	100.00	<i>Staphylococcus</i>	<i>aureus</i>	chromosome	
<i>Staphylococcus aureus</i> LmrS	86.25	99.27	<i>Staphylococcus</i>	<i>aureus</i>	unclassified	
tet(38)	100.44	99.33	<i>Staphylococcus</i>	<i>aureus</i>	plasmid	
tetS	13.42	100.00				
tetS	11.23	97.22				
vanJ	18.48	100.00	<i>Streptococcus</i>	<i>thermophilus</i>	unclassified	
vanRG	19.57	100.00	<i>Streptococcus</i>	<i>thermophilus</i>		
vanTN	6.43	100.00	<i>Kocuria</i>			
ykkC	25.89	100.00	<i>Kocuria</i>			
<b>Strict RGI match in sample B</b>						
mefE	5.71	100.00	<i>Delftia</i>	<i>tsuruhatensis</i>		
OXA-269	13.55	100.00				
OXA-442	9.12	100.00				
PEDO-1	24.83	100.00	<i>Acinetobacter</i>	sp. TTH0-4		

**Table 1.** ARGs identified in milk samples. The coverage column shows the fraction of CARD ARG reference sequence covered by the most similar ORF sequence. Identity represents the proportion of the identical nucleotides between the detected ORF and CARD ARG reference sequence. Species column shows the most likely species related to the ARG harbouring contig classified by Kraken2. For some contigs, the species level classification was ambiguous, genus reported only. The localization of contigs with ARG and longer than 1000 bp predicted by PlasFlow. Mobile genetic element domains coexisting with ARG are listed in column MGE.



**Figure 2.** Maximal coverage and identity of detected ORFs by antibiotics. The ORF covered proportion of the reference ARG sequence (x axis) and the identity % of predicted protein (color).

MGE domain coexisting analysis showed that PC1 beta-lactamase (*blaZ*) ARG might be mobile since the contig had a phage integrase ORF within the distance of 10 ORFs.

There were no perfect matches in sample B that is not surprising since its overall bacterial nucleic acid content was less than 10% of that of sample A. The sequence coverages of the strict matches in this sample ranged between 5.71% and 24.83%, with the mean of 13.31%. The identity between ORFs and CARD ARG reference sequences was 100.00% in each detected ARG. Contigs containing ARGs were classified and genera *Acinetobacter* (50%) and *Delftia* (50%) were identified. None of the identified ARGs could be related to any mobile genetic elements.

The detected ARGs in both samples were matched to their corresponding antibiotics. Since one antibiotic compound may be related to more than one ORFs, we decided to select those to which we could link the ORFs with the broadest coverage and the highest identity to the reference ARG sequence. The maximal coverage and identity of detected ORFs are shown in Fig. 2. In sample A ARGs known to be decreasing the effectiveness of acridine dye, cephalosporin, fluoroquinolone, penam and peptide antibiotics were found in full length and with 100% identity. There were two other ARGs identified in sample A in full length and with identity above 99% that encoded resistance against further antibiotics (cephamycin and tetracycline).

## Discussion

Antimicrobial resistance (AMR) is a natural feature of microorganisms that have originally occurred as a means of defence in the rivalry amongst the members of the microbiotas. The ubiquity of antimicrobial resistance genes (ARGs) is beyond question. Genes against antibiotics are present both in non-pathogenic and pathogenic bacteria. With the extended agricultural and clinical use of antibiotics, the number of ARGs are on the rise, and the growing number and spread of multi-resistant bacteria strains pose a global threat to global health. According to the CDC's Antibiotic Resistance Threats in the United States, 2019 report<sup>3</sup>, more than 2.8 million antibiotic-resistant infections occur in the U.S. each year, and more than 35,000 people die as a result. In addition, 223,900 cases of *Clostridioides difficile* occurred in 2017 and at least 12,800 people died. Dedicated prevention and infection control efforts in the U.S. and around the world are working to reduce the number of infections and deaths caused by antibiotic-resistant germs. However, the number of people facing antibiotic resistance is still too high. The AR Threats Report warns that not only people but also animals carry bacteria in their guts which may include antibiotic-resistant bacteria either with intrinsic or with acquired ARGs<sup>12</sup>. Beyond disease control, animals may be given antibiotics for growth promotion or increased feed efficiency. Since bacteria are exposed to low doses of the drugs over a long period, this inappropriate antibiotic use can lead to the development of resistant bacteria<sup>3</sup>. The CDC report notes that when animals are slaughtered and processed for food, resistant germs in the animal gut can contaminate meat or other animal products, but do not mention the possible contamination of milk.

Detected ARGs in raw milk (Table 1) can be transferred from non-pathogens to pathogens via HGT. The over-expression of such genes, e.g. *norA* (regulated by *mgrA*) and *mepA* (regulated by *mepR*) coding multidrug efflux pumps confer resistance to fluoroquinolones (including norfloxacin or ciprofloxacin) and even disinfectants<sup>18–21</sup>. Ciprofloxacin is a broad-spectrum antibiotic used to treat a variety of bacterial infections, including intra-abdominal, respiratory tract, skin, urinary tract, and bone and joint infections. Norfloxacin might be used for uncomplicated urinary tract infections (including cystitis) or the prevention of spontaneous bacterial peritonitis in cirrhotic patients, among others. *MepA* was also shown to result in resistance to tigecycline<sup>22</sup>, an antibiotic

that was developed to tackle complicated infections caused by multiresistant bacteria such as *Staphylococcus aureus*, *Acinetobacter baumannii*, and *E. coli*.

The two samples differ both in the composition of core bacteriome and the ARG abundance. In sample A the bacteriome was dominated by Gram-positive bacteria. Furthermore, most of the contigs harbouring ARG were classified taxonomically belonging to Gram-positive bacteria. In sample B, the Gram-negative bacteria governed the bacteriome. So the lower ARG abundance in sample B might come from the lower proportion of Gram-positives. Nevertheless, in sample B, not just the number of detected ARGs was lower, but the maximal coverage of the ARGs as well. One may find the reason for this phenomenon, the lower sequencing depth of sample B. The identity of these ORFs with the reference ARGs are very high so we may assume the assembled ORFs originated from ARGs. Accordingly, the possible reason of the lower coverage of ARGs may be caused by the insufficient read counts for assembly the complete ORFs. One possible argumentation of the ARGs' difference between sample A and B may be derived from the fact that health issues (e.g. mastitis) are relatively more common in large scale farms. Since the use of antibiotics is more permissive in veterinary practice - compared to human medicine - in the treatment of bacterial infections, it places a selective pressure on the bacteria of herds, what might increase the frequency and the diversity of ARGs.

Our results show that indeed ARGs can be present in raw milk. However, it should be the subject of further research to identify how resistant bacterial DNA gets into the milk, is it already there in the cow's udder or does it only mixed into the milk as contamination during or after milking.

At raw milk's environment of origin (dairy farms), the use of antimicrobial agents is widespread. Consequently, the microbiome of this product may show relatively high levels of resistance. Without heat-treatment, bacteria that are present in raw milk are not hindered from further multiplication what results in the amplification of their resistance genes either. Such a rise in the number of ARGs may increase the risk of horizontal gene transfer (HGT) events. This risk may even be higher in case of mobile ARGs (e.g. blaZ, which was detected on a plasmid and near to a phage integrase ORF).

Beyond human intervention, there are natural mechanisms that limit ARG-transfer<sup>11</sup>. First of all, donor and recipient populations need to be present at the same physical space<sup>23</sup> and reach a specific critical density to ensure proper connectivity for a successful gene transfer event. Chances for a series of HGT events amongst two physically distant populations are relatively low except for the case when there is positive selection driven by any factors (e.g. selection by antibiotics). The second factor arises from the fact that genes encoding resistance against the same compounds may limit each other's spread. A population owning genes against a particular antibiotic is not under selective pressure to gain any other ARGs with the same effect. As a conclusion of earlier evolutionary steps, possession of resistance determinants of the same substrate profiles are possible. However, in a population where the distribution of these genes is stable, the chances of new recruitments are lower. Tertiary, acquisition of resistance genes sets metabolic costs deriving from the transfer and integration mechanisms needed. These costs vary by each ARG, and only affordable genes are spread<sup>11</sup>.

Even though the bacterial compositions of milk are affected by the heat treatment<sup>24,25</sup>, the question may arise whether the ARG content of raw and pasteurized milk are different? In water, DNA degradation starts by 90 °C<sup>26</sup>. The HTST pasteurization (high temperature/short time) is performed at 72 °C for 15–40 seconds, while ultra-pasteurization (UHT) is at 135 °C for 1–2 seconds. Summarizing this information, one may conclude that the resistomes do not differ significantly in HTST and raw milk. On the other hand in UHT milk some DNA degradation might be suspected. Nevertheless, some aspects are broadening the picture, that are worth taking into consideration. First of all, in raw milk, the members of the bacteriota remain viable and may multiply depending on the storing temperature. The proliferation of bacterial cells increases the amount of the sample's extractable bacterial DNA content what appears in the results of the sequencing as raised bacterial read rates. Consequently, after the assembly of the reads, the likelihood of having contigs containing ARGs is higher. Pasteurization kills 99.99% of bacteria; thus, their multiplication has a low significance. Secondly, the bacteriome of milk consumers (humans and animals) may gain the ARGs of the milk-resistome by transformation and transduction only<sup>13</sup>, as pasteurization decreases the number of viable bacteria. In contrast, raw milk's higher viable bacterial cell count facilitates conjugation to the consumers' bacteriome while the above-mentioned horizontal gene transfer mechanisms<sup>13</sup> are also kept. Of course, this phenomenon rather has an impact on the risk of HGT than on the resistome of raw or pasteurized milk.

Nevertheless, heat-treatment of raw milk seems to be an advantageous and a more than considerable step that besides inhibiting the amplification of genes having a potential risk, makes active gene transfer mechanisms lose their significance. On the other hand, even though it reduces the number of multiplication cycles, after the lysis of cells free DNA fragments appear in the sample that may still be uptaken by newly arriving bacteria.

However, the interpretation of resistome studies is yet to be deepened. The combination of next-generation sequencing, metagenomic and computational methods provides valuable data on the presence of ARGs. Moreover, it makes it possible to find genes in full coverage and length, and to identify their taxonomical classes of origin and their exact sequential surroundings. Synteny with mobile genetic elements is a fact to be taken into consideration when examining the risks meant by an ARG. Thus, the combination of methods mentioned above serves as a core component of today's necessarily expanded antimicrobial resistance research.

As a means of evolutionary pressure, the use of antibiotics selects bacterial strains that have antimicrobial resistance genes. Moreover, in the production animal sector, the application of such compounds increases not only the number of antibiotic-resistant bacterial strains but also the frequency of their appearance. After the consumption of animal products, these strains may meet the human microbiota, and the circumstances may be appropriate for the horizontal gene transfer derived spread of antimicrobial resistance genes amongst these populations. This phenomenon unfolds a possible source of acquisition of human pathogens' antimicrobial resistance other than the direct presence of antibiotic residuals in animal products.

Our findings suggest the antimicrobial resistance gene content of unprocessed animal products may play a role in the development of antimicrobial resistance of human pathogens. Nevertheless, the generalization of these findings requires more comprehensive studies to transcend our results that are based on a limited sample size.

## Methods

**Milk samples.** Two times one litre of raw milk was purchased at a public market in the city of Budapest (sample A) and in the city of Szeged (sample B). Sample A and B originated from a large (with more than 250 dairy cattle) and a small scale (below 50 dairy cattle) dairy farm, respectively.

**DNA extraction and metagenomics library preparation.** Before the laboratory procedures, the milk samples were stored frozen. 120 mL of raw milk was centrifuged at 10,000 g for 10 min. Total DNA was extracted from the pellet using the ZR Fecal DNA Kit from Zymo Research. Paired-end fragment reads ( $2 \times 150$  nucleotides) were generated using the TG NextSeq 500/550 Mid Output Kits v2 sequencing kit with an Illumina NextSeq sequencer. Primary data analysis (base-calling) was carried out with bcl2fastq software (v.2.17.1.14, Illumina).

**Bioinformatic analysis.** Quality based filtering and trimming was performed by Adapterremoval<sup>27</sup>, using 15 as a quality threshold. Only reads longer than 50 bp were retained. *Bos taurus* genome (ARS-UCD1.2) sequences as host contaminants were filtered out by Bowtie2<sup>28</sup> with *very-sensitive-local* setting minimizing the false positive match level<sup>29</sup>. The remaining reads were taxonomically classified using Kraken2 ( $k = 35$ )<sup>30</sup> with the NCBI non-redundant nucleotide database<sup>31</sup>. The taxon classification data was managed in R<sup>32</sup> using functions of package phyloseq<sup>33</sup> and microbiome<sup>34</sup>. For further analysis, the reads assigned to Bacteria was used only<sup>35</sup>. Core bacteria was defined as the relative abundance of agglomerated counts at class level above 0.1% at least one of the samples. By metaSPAdes<sup>36</sup> the preprocessed reads were assembled to contigs, with the automatically estimated maximal  $k = 55$ . From these contigs having a shorter length than the shortest resistance gene of the Comprehensive Antibiotic Resistance Database (CARD) were discarded<sup>37,38</sup>. The ARG content of filtered contigs was analyzed with Resistance Gene Identifier (RGI) v5.1.0 and CARD v.3.0.6<sup>37,38</sup>. Contigs harbouring ARG identified by RGI with perfect or strict cut-off were preserved and classified by Kraken2 on the same way as was described above. The plasmid origin probability of the contigs was estimated by PlasFlow v.1.1<sup>39</sup>. To identify possible further mobile genetic element (MGE) homologs the predicted protein sequences of contigs were scanned by HMMER<sup>40</sup> against data of PFAM v32<sup>41</sup> and TnpPred<sup>42</sup>. Following Saenz *et al.*<sup>35</sup> from the hits with lower than  $E 10^{-5}$  the best was assigned to each predicted protein within the distance of 10 ORFs. The MGE domains coexisting with ARGs were categorized as phage integrase, resolvase, transposase or transposon.

## Data availability

All data are publicly available and can be accessed through the PRJNA591315 from the NCBI Sequence Read Archive (SRA).

Received: 31 December 2019; Accepted: 3 April 2020;

Published online: 04 May 2020

## References

- World Health Organization. *Antimicrobial resistance: global report on surveillance* (World Health Organization, Geneva, Switzerland, 2014).
- Aarestrup, F. M. The livestock reservoir for antimicrobial resistance: a personal view on changing patterns of risks, effects of interventions and the way forward. *Philos. Transactions Royal Soc. B: Biol. Sci.* **370**, 20140085 (2015).
- Centers for Disease Control and Prevention. *Antibiotic Resistance Threats in the United States* (U.S. Department of Health and Human Services, CDC, Atlanta, GA, USA, 2019).
- Surette, M. D. & Wright, G. D. Lessons from the environmental antibiotic resistome. *Annu. Rev. Microbiol.* **71**, 309–329 (2017).
- Li, J. *et al.* Global survey of antibiotic resistance genes in air. *Environ. Sci. & Technol.* **52**, 10975–10984 (2018).
- Che, Y. *et al.* Mobile antibiotic resistome in wastewater treatment plants revealed by Nanopore metagenomic sequencing. *Microbiome* **7**, 44 (2019).
- Hendriksen, R. S. *et al.* Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nat. Commun.* **10**, 1124 (2019).
- D'Costa, V. M., McGrann, K. M., Hughes, D. W. & Wright, G. D. Sampling the antibiotic resistome. *Science* **311**, 374–377 (2006).
- D'Costa, V. M. *et al.* Antibiotic resistance is ancient. *Nature* **477**, 457–461 (2011).
- Bhullar, K. *et al.* Antibiotic resistance is prevalent in an isolated cave microbiome. *PLoS One* **7**, e34953 (2012).
- Martínez, J. L., Coque, T. M. & Baquero, F. What is a resistance gene? Ranking risk in resistomes. *Nat. Rev. Microbiol.* **13**, 116–123 (2015).
- Crofts, T. S., Gasparrini, A. J. & Dantas, G. Next-generation approaches to understand and combat the antibiotic resistome. *Nat. Rev. Microbiol.* **15**, 422–434 (2017).
- Thomas, C. M. & Nielsen, K. M. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* **3**, 711–721 (2005).
- Claeys, W. L. *et al.* Raw or heated cow milk consumption: Review of risks and benefits. *Food Control.* **31**, 251–262 (2013).
- Lucey, J. A. Raw Milk Consumption: Risks and Benefits. *Nutr. Today* **50**, 189–193 (2015).
- Frost, L. S., Leplae, R., Summers, A. O. & Toussaint, A. Mobile genetic elements: the agents of open source evolution. *Nat. Rev. Microbiol.* **3**, 722–732 (2005).
- Partridge, S. R., Kwong, S. M., Firth, N. & Jensen, S. O. Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clin. Microbiol. Rev.* **31**, e00088–17 (2018).
- Kaatz, G. W., McAleese, F. & Seo, S. M. Multidrug resistance in staphylococcus aureus due to overexpression of a novel multidrug and toxin extrusion (mate) transport protein. *Antimicrob. Agents Chemother.* **49**, 1857–1864 (2005).
- Truong-Bolduc, Q. C. & Hooper, D. C. The transcriptional regulators norg and mgra modulate resistance to both quinolones and  $\beta$ -lactams in staphylococcus aureus. *J. Bacteriol.* **189**, 2996–3005 (2007).
- Schindler, B. D. & Kaatz, G. W. Multidrug efflux pumps of gram-positive bacteria. *Drug Resist. Updat.* **27**, 1–13 (2016).

21. Hassanzadeh, S. *et al.* Epidemiology of efflux pumps genes mediating resistance among staphylococcus aureus; a systematic review. *Microb. Pathog.* 103850 (2019).
22. McAleese, F. *et al.* A novel mate family efflux pump contributes to the reduced susceptibility of laboratory-derived staphylococcus aureus mutants to tigecycline. *Antimicrob. Agents Chemother.* **49**, 1865–1871 (2005).
23. Hankin, L., Lacy, G. H., Stephens, G. R. & Dillman, W. F. Antibiotic-resistant bacteria in raw milk and ability of some to transfer antibiotic resistance to *Escherichia coli*. *J. Food Protect.* **42**, 950–953 (1979).
24. Quigley, L. *et al.* The microbial content of raw and pasteurized cow milk as determined by molecular approaches. *J. Dairy Sci.* **96**, 4928–4937 (2013).
25. Park, W. *et al.* Microbiological characteristics of gouda cheese manufactured with pasteurized and raw milk during ripening using next generation sequencing. *Food Sci. Anim. Resour.* **39**, 585 (2019).
26. Karni, M., Zidon, D., Polak, P., Zalevsky, Z. & Shefi, O. Thermal degradation of dna. *DNA Cell Biol.* **32**, 298–301 (2013).
27. Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **9**, 88 (2016).
28. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
29. Czajkowski, M. D., Vance, D. P., Frese, S. A. & Casaburi, G. GenCoF: a graphical user interface to rapidly remove human genome contaminants from metagenomic datasets. *Bioinformatics* **35**, 2318–2319 (2019).
30. Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257 (2019).
31. Pruitt, K. D., Tatusova, T. & Maglott, D. R. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucl. Acids Res.* **33**, D501–4 (2005).
32. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2019).
33. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* **8**, 1–11 (2013).
34. Lahti, L. & Shetty, S. microbiome R package (2012–2019).
35. Sáenz, J. S. *et al.* Oral administration of antibiotics increased the potential mobility of bacterial resistance genes in the gut of the fish *Piaractus mesopotamicus*. *Microbiome* **7**, 24 (2019).
36. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* **27**, 824–834 (2017).
37. McArthur, A. G. *et al.* The comprehensive antibiotic resistance database. *Antimicrob. Agents Chemother.* **57**, 3348–57 (2013).
38. Jia, B. *et al.* CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **45**, D566–D573 (2017).
39. Krawczyk, P. S., Lipinski, L. & Dziembowski, A. PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Res.* **46**, e35–e35 (2018).
40. Mistry, J., Finn, R. D., Eddy, S. R., Bateman, A. & Punta, M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.* **41**, e121–e121 (2013).
41. El-Gebali, S. *et al.* The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**, D427–D432 (2019).
42. Riadi, G., Medina-Moene, C. & Holmes, D. S. TnpPred: A Web Service for the Robust Prediction of Prokaryotic Transposases. *Int. J. Genomics* **2012**, Article ID 678761 (2012).

## Acknowledgements

The project is supported by the European Union and co-financed by the European Social Fund (No. EFOP-3.6.3-VEKOP-16-2017-00005), ÚNKP-19-2 New National Excellence Program of the Ministry for Innovation and Technology and has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 643476 (COMPARE).

## Author contributions

N.S. had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. L.M., I.C. and N.S. conceived of the study concept. G.M., L.M. and N.S. performed the sample collection and procedures. A.G.T., D.T., E.K. and N.S. participated in the bioinformatic and statistical analysis. A.G.T., A.V.P., L.M. and N.S. participated in the drafting of the manuscript. A.G.T., D.T., G.S., I.C. and N.S. carried out the critical revision of the manuscript for important intellectual content. All authors read and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to N.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020