



## ARTICLE OPEN



# The sources and transmission routes of microbial populations throughout a meat processing facility

Benjamin Zwirzitz<sup>1,2</sup><sup>✉</sup>, Stefanie U. Wetzels<sup>1,2</sup>, Emmanuel D. Dixon<sup>1</sup>, Beatrix Stessl<sup>1</sup>, Andreas Zaiser<sup>1</sup>, Isabel Rabanser<sup>1</sup>, Sarah Thalgueter<sup>1</sup>, Beate Piniör<sup>1</sup>, Franz-Ferdinand Roch<sup>1</sup>, Cameron Strachan<sup>1,2</sup>, Jürgen Zanghellini<sup>3</sup>, Monika Dzieciol<sup>1</sup>, Martin Wagner<sup>1,2</sup> and Evelyn Selberherr<sup>1</sup>

Microbial food spoilage is responsible for a considerable amount of waste and can cause food-borne diseases in humans, particularly in immunocompromised individuals and children. Therefore, preventing microbial food spoilage is a major concern for health authorities, regulators, consumers, and the food industry. However, the contamination of food products is difficult to control because there are several potential sources during production, processing, storage, distribution, and consumption, where microorganisms come in contact with the product. Here, we use high-throughput full-length 16S rRNA gene sequencing to provide insights into bacterial community structure throughout a pork-processing plant. Specifically, we investigated what proportion of bacteria on meat are presumptively not animal-associated and are therefore transferred during cutting via personnel, equipment, machines, or the slaughter environment. We then created a facility-specific transmission map of bacterial flow, which predicted previously unknown sources of bacterial contamination. This allowed us to pinpoint specific taxa to particular environmental sources and provide the facility with essential information for targeted disinfection. For example, *Moraxella* spp., a prominent meat spoilage organism, which was one of the most abundant amplicon sequence variants (ASVs) detected on the meat, was most likely transferred from the gloves of employees, a railing at the classification step, and the polishing tunnel whips. Our results suggest that high-throughput full-length 16S rRNA gene sequencing has great potential in food monitoring applications.

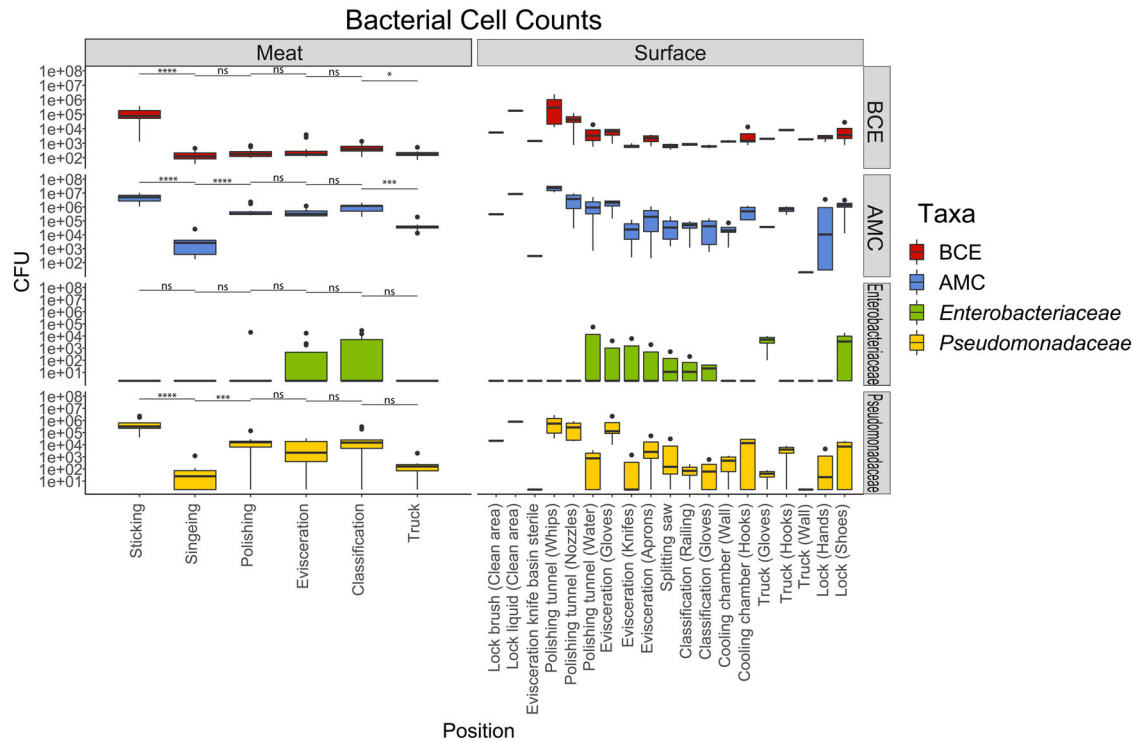
npj Biofilms and Microbiomes (2020)6:26; <https://doi.org/10.1038/s41522-020-0136-z>

## INTRODUCTION

Minimizing food loss and food waste are one of the major sustainable development goals of the United Nations. Reducing food losses along the production and supply chains are essential aspects, by virtue of estimations that about 31% of the food produced in the US is lost before it even reaches the consumers<sup>1</sup>. A key component in the reduction of food loss is the prevention of microbial spoilage of food, which is estimated to account for one-fourth of global food waste<sup>2</sup>. In addition, microbial spoilage of food also poses a key public health concern. The World Health Organization (WHO) reported 600 million foodborne illnesses causing 420,000 deaths worldwide in 2010 alone, indicating that the global burden of food-borne diseases is comparable to those of major infectious diseases<sup>3</sup>. Indeed, the food industry faces major and continuing challenges in trying to lower the extent to which food products become contaminated with pathogenic or spoilage bacteria during primary processing. This is especially true for animal-derived products like poultry, eggs, milk, and pork, which are main vehicles of food-borne diseases<sup>4</sup>. Thus, microbial meat spoilage is a global health and economic challenge, yet little is known about the microbial diversity in slaughterhouses and meat-cutting plants. Pork provides an ecosystem with high water content and nutrient availability for diverse microorganisms, particularly psychrotolerant organisms, which can then grow during storage and lead to spoilage and a reduced shelf life<sup>5</sup>. The formation of biofilms on processing equipment is also of great concern and can lead to continuous dispersal of microorganisms on the meat<sup>6</sup>. To date, transmission routes of microorganisms during meat processing have been difficult to track and monitor

and therefore remained largely elusive<sup>7,8</sup>. The main reason for this is that the meat industry is still relying on ISO reference methods applying microbiological techniques to monitor hygiene aerobic colony counts (ISO 4833) and *Enterobacteriaceae* (ISO 21528-2)<sup>9</sup>. Indeed, culture-dependent techniques can be useful in determining the overall hygiene status of a facility, but they fail in describing complex microbial communities and population flows<sup>10</sup>. Just recently, first efforts have been made to characterize the microbiome of food processing environments with next-generation sequencing techniques, demonstrating the suitability of these tools to map microbial ecosystems in different sectors of food industry<sup>11–14</sup>. In a previous study, we found that many bacteria in musculature samples of slaughter pigs are not animal-associated<sup>15</sup>. Thus, we hypothesize, that a large part of the microbial community is transmitted during cutting via personnel, from the equipment, or from the machine and slaughter environment. Our approach was to investigate the abundances as well as the types of microorganisms present at different stages along the meat-processing chain and explore the possible sites of microbial transmission in a slaughterhouse to effectively reduce the risk of cross-contamination. Specifically, we took samples from 12 pigs at different stages during processing as well as from multiple sites throughout the facility environment. Using full-length 16S rRNA gene sequencing, we were able to identify key carrier points of bacterial contamination and create a facility-specific transmission map of bacterial flows that exposed unique transmission patterns for individual taxa. Although this study primarily aids to optimize slaughter processes to systematically avoid contamination of microbes throughout meat processing,

<sup>1</sup>Institute of Food Safety, Food Technology and Veterinary Public Health, University of Veterinary Medicine, Veterinärplatz 1, 1210 Vienna, Austria. <sup>2</sup>Austrian Competence Centre for Feed and Food Quality, Safety and Innovation FFOQSI GmbH, Technopark 1C, 3430 Tulln, Austria. <sup>3</sup>Department of Analytical Chemistry, University of Vienna, 1010 Vienna, Austria. ✉email: Benjamin.Zwirzitz@vetmeduni.ac.at



**Fig. 1 Microbial numbers at the different sampling positions.** Bacterial cell equivalents (BCE) and colony forming units (CFU) of aerobic mesophilic counts (AMC), *Enterobacteriaceae* (EB), and *Pseudomonadaceae* (PS) for the different sampling positions. Boxes indicate the interquartile range (75th–25th) of the data. The median value is shown as a line within the box. Whiskers extend to the most extreme value within 1.5 \* interquartile range. Levels of significance: ns:  $p > 0.05$ , \* $p \leq 0.05$ , \*\* $p \leq 0.01$ , \*\*\* $p \leq 0.001$ .

the techniques we applied can also be extended to other food-processing environments. In this way, we can expand our knowledge about microbial transmission routes, thereby improving hygiene standards in food-related industry to increase food safety while minimizing food waste.

## RESULTS

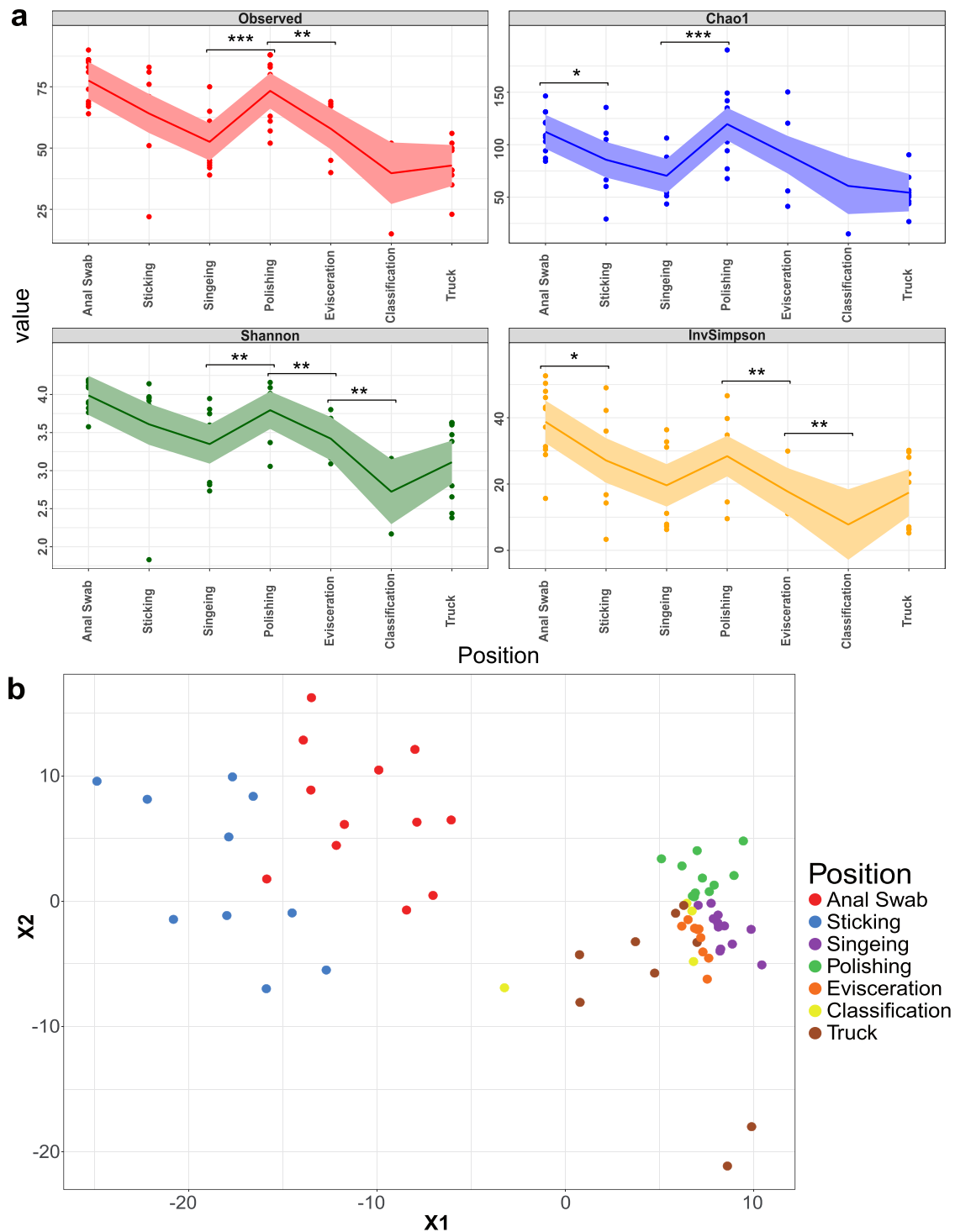
### Bacterial cell counts strongly differ between sampling positions and surface locations

In order to get an initial basic understanding of the microbiological status along the processing line of the facility we determined total bacterial cell equivalents (BCE) using 16S rRNA gene qPCR, as well as aerobic mesophilic (AMC), *Enterobacteriaceae* (EB), and *Pseudomonadaceae* (PS) counts by applying ISO reference methods (ISO 4833, ISO 21528-2) (Fig. 1). The animals entered the facility with a high microbial load on skin (Fig. 1, “Sticking”), which was reduced significantly after singeing. In the next step (“Polishing”), AMC and PS counts were increased significantly, while EB counts were not detected until that point, but were found at low levels at the evisceration and classification step and in environmental samples (Gloves, knives, aprons, etc.). Total aerobic mesophilic counts along the slaughter line ranged from  $4.14 \times 10^3$  CFU/cm<sup>2</sup> after singeing to  $5.21 \times 10^6$  CFU/cm<sup>2</sup> at sticking. The highest levels of bacteria in the environment were found at the whips of the polishing tunnel ( $2.19 \times 10^7$  CFU/cm<sup>2</sup>). BCE counts ranged from  $1.64 \times 10^2$  BCE/cm<sup>2</sup> after singeing to  $1.20 \times 10^5$  BCE/cm<sup>2</sup> at sticking and were not significantly higher after polishing. We found that BCE and AMC numbers correlate in skin samples taken at the beginning of the processing line ( $R: 0.76$ ,  $p < 0.01$ ), but not on meat samples from the end of the processing line ( $R: 0.38$ ,  $p = 0.32$ ) and from environmental samples taken in the facility (e.g., Polishing tunnel whips  $R: 0.74$ ,  $p = 0.26$ ).

Microbial community structure changes throughout the processing line

On average, 1186 high quality full-length 16S rRNA gene sequences per sample remained after stringent quality filtering. A rarefaction curve indicated that one third, but not all samples, were sequenced deep enough to infer the full diversity of microorganisms in the samples (Supplementary Fig. 1). All analysed alpha diversity indices showed a similar pattern, where there is an overall decrease of microbial species diversity from start to end of the processing line, but with a transient increase after the polishing step (Fig. 2a). Beta diversity analysis revealed two major shifts in the microbial community structure (Fig. 2b). The differences in microbial composition between samples from the skin and rectum (Positions “Sticking” and “Anal Swab”) of the pigs was higher compared to samples later in the processing line. A shift occurred during the singeing step, which is also reflected by significantly reduced microbial numbers as well as species diversity. From that point on the microbial community stayed relatively constant until the end of the processing line between the stations “Classification” and “Truck” at which the communities dispersed from each other.

A majority (91%) of the full-length 16S rRNA gene sequences could be assigned to specific genera and 74% to the species level. The 50 most abundant ASVs show heterogeneous distributions and relative abundances across all meat samples (Fig. 3). The shift in the microbial community structure that was observed in beta diversity analysis can also be discerned by the relative abundances of ASVs. At the beginning of the slaughterline (Anal Swab and Sticking station), the 50 most abundant ASVs make up less than 20% of the total community and have higher levels of *Helicobacter*, and *Curvibacter*, compared to samples from the other positions. In contrast, the genera *Anoxybacillus*, *Chryseobacterium*, and *Moraxella* were the most abundant ASVs in the meat samples after the singeing step, making up more than 50% of the

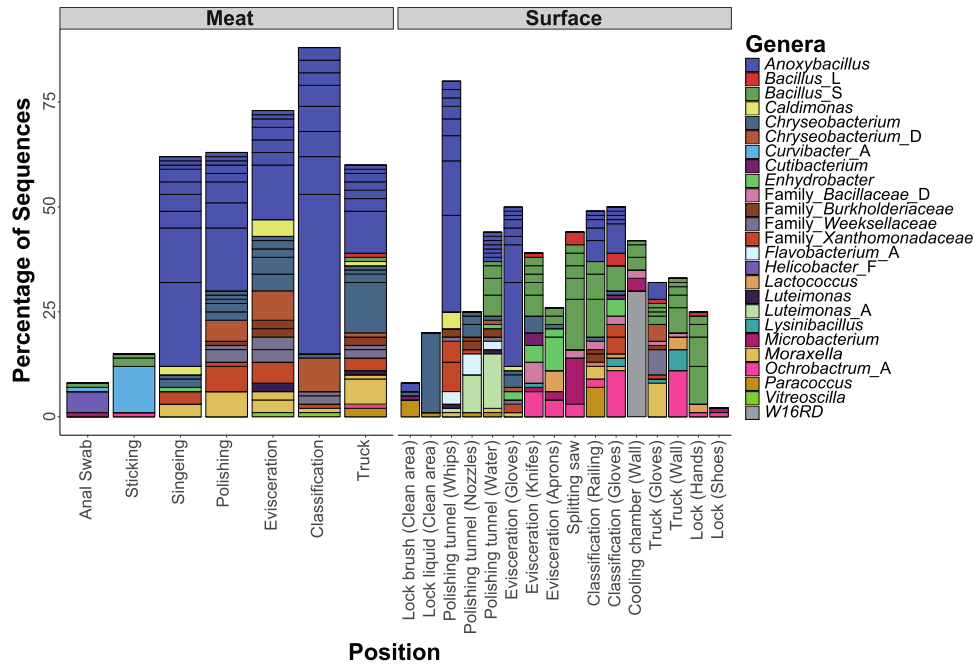


**Fig. 2 Diversity of microbial communities associated with meat samples. a** Change in alpha diversity indices of meat samples over time. Points represent individual samples, the trend lines connect the means, and shaded regions indicate the standard error. Levels of significance:  $*p \leq 0.05$ ,  $**p \leq 0.01$ ,  $***p \leq 0.005$ . **b** t-SNE plot of Bray–Curtis distances based on 16S rRNA gene libraries. Each point represents values from individual libraries with colors expressing meat samples from different positions along the processing line.

sequences. The same ASVs that were highly abundant on the meat samples were also frequently detected in the surface samples of the facility environment. However, the distribution of these ASVs throughout the facility was varied in terms of their location specificity and prevalence. For instance, ASVs belonging to the genus *Bacillus\_S* or *Moraxella* were homogeneously distributed across many different positions, whereas others were

detected only at very specific locations e.g., *Luteimonas\_A* and *Helicobacter\_F* at the polishing tunnel or W16RD (NCBI taxonomy: *Shingomonas*) at the wall of the cooling chamber (Fig. 3 and Supplementary Table 1).

The shorter sequences obtained with the Illumina Miseq had a lower classification rate (Genus: 80.6%, Species: 56.8%) compared to the dataset with the full-length 16S rRNA gene sequences. This



**Fig. 3 Genus-level classification of the 50 most abundant ASVs parted by type (meat or surface).** Data represents average of ASV counts from replicate libraries for each category. Individual ASVs are separated by a black line within the bar graph. ASVs assigned to candidate genera that do not have a name assignment yet are indicated with "Family\_". Genera names with an alphabetic suffix indicate that are polyphyletic and were therefore subdivided in the genome taxonomy database.

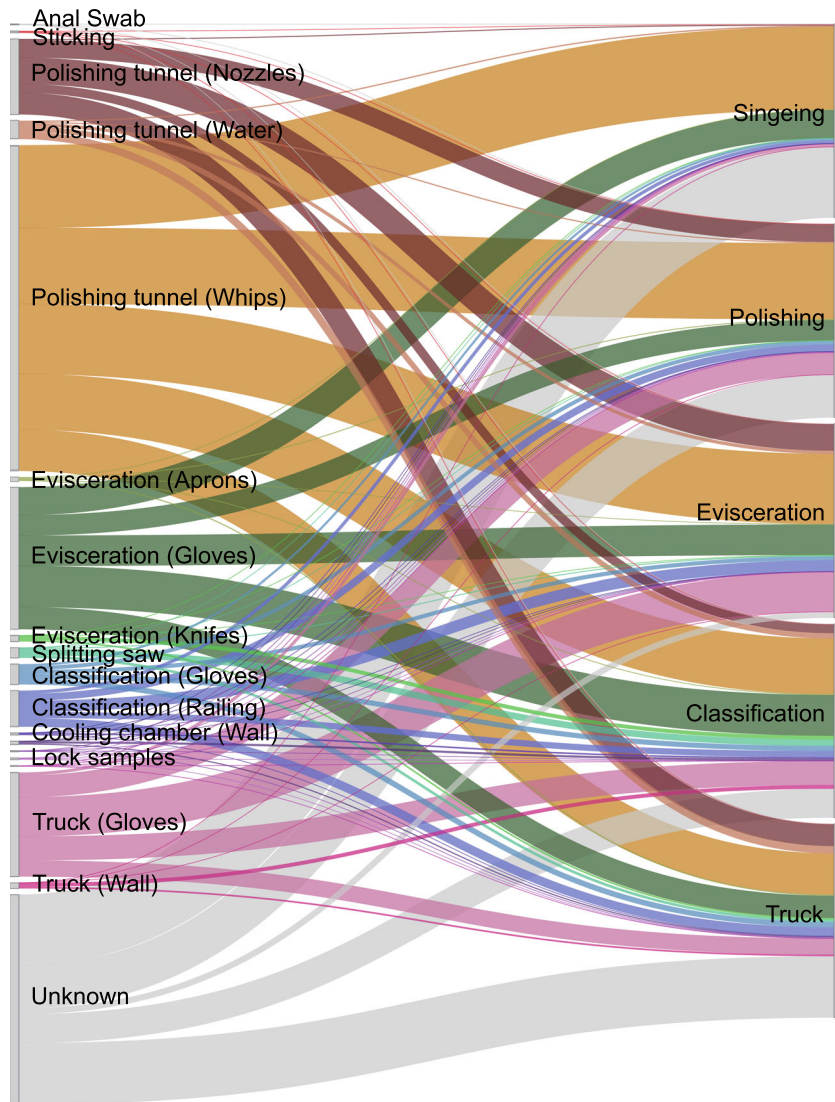
constitutes an increase in species level taxonomic assignment of 30.3% with the long read sequencing technology. In terms of capturing the overall microbial community structure, both sequencing technologies revealed consistent transitions between the individual processing steps. The same genera were found to be highly abundant across the processing line although their relative abundances differed significantly between datasets (Fig. 3 and Supplementary Fig. 2).

Variances in spatial distribution of microorganisms throughout the facility result in distinct transmission routes for specific taxa

Presumptive transfer of microorganisms from environmental samples to meat samples was inferred using the software SourceTracker. Samples collected from pork carcasses were designated as sinks for testing against the communities of samples from the anal region, the skin, and from the equipment and material surface (source). Alpha diversity indices had already indicated transfer of new species onto the meat at the polishing step (Fig. 2a). This was confirmed by SourceTracker analysis showing a high contribution of polishing tunnel samples (nozzles, water, and especially whips) to the microbial community of meat samples (Fig. 4, Supplementary Fig. 3). Anal swab and sticking positions were not identified as one of the major sources of bacterial contamination despite having shown high alpha diversity and bacterial cell count values. Hence, the decontamination processes, e.g., singeing, likely eliminated most of the bacteria that were initially on the pork carcass when pigs entered the facility. Furthermore, the small contribution of anal swab samples validates the effectiveness of the general practice in the analyzed facility to seal off the rectum with a small plastic bag before evisceration to avoid contamination with fecal matter. The evisceration step is usually considered to be a critical point of re-contamination if cutting processes are not executed with high hygiene standards. In this case, transfer of microorganisms from aprons and knives was low but a high proportion (11.4%) of microorganisms detected on meat samples taken from the last position (Truck) originated from the gloves of employees

performing the evisceration. Gloves of employees were also identified as major contamination sources at other positions (Truck (8.3%), classification (2.8%)). Furthermore, a railing that all carcasses touched while passing the classification contributed to the microbial composition of the meat samples (4.4%). Environmental samples without direct contact to the carcasses (Lock and wall samples) were also not identified as major contamination sources, suggesting that most bacteria are transferred through direct contact with the surface and that air transfer is marginal. One third (31.6%) of the bacteria on meat could not be linked to a specific source and was therefore attributed to an unknown source. Possible reasons for this are that the sequencing depth for some samples was too low or that the primary source of these bacteria was missed during sampling. This could potentially lead to overfitting of the SourceTracker analysis. Thus, future studies should also consider soil and human skin samples as possible sources.

Finally, we were interested if certain microbial species were transmitted from specific sources or if they were spread throughout the whole facility. We investigated the distribution of all detected genera by determining their relative abundances across all meat and environmental positions. This was done for each individual genus (Supplementary Figs. 4 and 5). Many genera were unique to specific sites demonstrating that they occupy particular environmental niches in the facility. In addition, we predicted the relative contribution of genera that are associated with meat spoilage or that include relevant pathogens based on the SourceTracker results (Fig. 5). Indeed, some of the microorganisms could be pinpointed to a specific location from where they disseminated, whereas others were found to have multiple possible origins. For example, the genus *Escherichia* originated almost exclusively from the anal swab samples. On the other hand, *Lactococcus*, *Staphylococcus*, *Chryseobacterium*, and *Moraxella* species were predicted to be transmitted from various different positions. In general, several taxa clustered together based on shared-source similarity in the heatmap, suggesting that they have the same presumed source, e.g., *Flavobacterium* and *Lactobacillus\_H* from the gloves at the evisceration step or



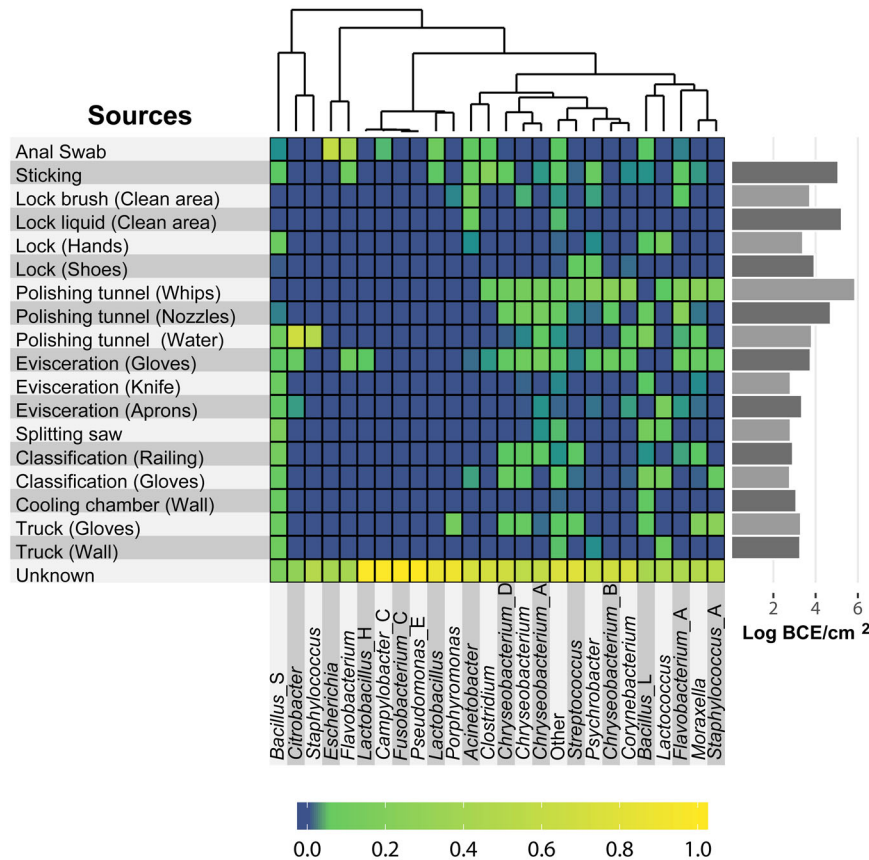
**Fig. 4 Source environment proportions for meat samples.** Source environment proportions for meat samples estimated using SourceTracker and visualized as a Sankey flow diagram. Environmental source samples are represented on the left and meat samples, as sinks, are shown on the right. The line width of individual flows between them illustrates the average predicted contribution/proportion of microorganisms from source samples to the microbial community of respective sink samples. The height of the individual bars of sink samples on the right (Singeing to Truck) sums up to 100%. The height of individual bars of source samples on the left represents the sum of proportions to each of the five sink samples.

*Lactococcus* and *Bacillus\_L*, which shared the splitting saw as a common source. A closer look at the phylogenetic relationship of ASVs assigned to the polyphyletic genus *Chryseobacterium* revealed that different populations were transferred from respective sources (Supplementary Fig. 6). An ASV classified as *Chryseobacterium* sp. Leaf405 in NCBI taxonomy (GCF\_001425355.1 in GTDB), was spread across the polishing tunnel and evisceration positions, whereas others have been found exclusively at the polishing tunnel (*Chryseobacterium indoltheticum*) or at the locks (genus *Chryseobacterium*; no species verification). Other prominent meat spoilage microorganisms, for example *Moraxella* spp., which was also one of the most abundant ASVs on the meat, was most likely transferred from the polishing tunnel whips, gloves of employees as well as from the railing at the classification step. Overall, we were able to identify transmission routes for the majority of the genera that include relevant meat spoilage organisms or pathogens. However, the origin of some taxa (e.g., *Fusobacterium\_C* and *Pseudomonas\_E*) remains

elusive since their contribution to the meat microbiota was attributed to an unknown source.

Sequencing depth has little effect on the performance of sourcetracker

To test whether the shallow sequencing depth (avg.: 1186 sequences/sample) of the “Pacbio” dataset was sufficient for a comprehensive analysis with SourceTracker, we additionally sequenced a subset of the samples on an Illumina Miseq machine. This dataset was then rarefied to different sizes and SourceTracker analysis was performed several times in order to simulate the effect of sequencing depth. Results from all randomly generated datasets correlated with the original dataset (spearman’s rho ranged from 0.95 to 0.99). The variation around the line of the best fitting original dataset increased with smaller dataset sizes (Fig. 6a) and thus, the mean squared differences decreased with dataset size (Supplementary Table 2). The Kruskal–Wallis rank sum test shows significant differences between the different rarefied datasets regarding to the squared differences ( $\chi^2 = 125.9$ ,  $p$ -value



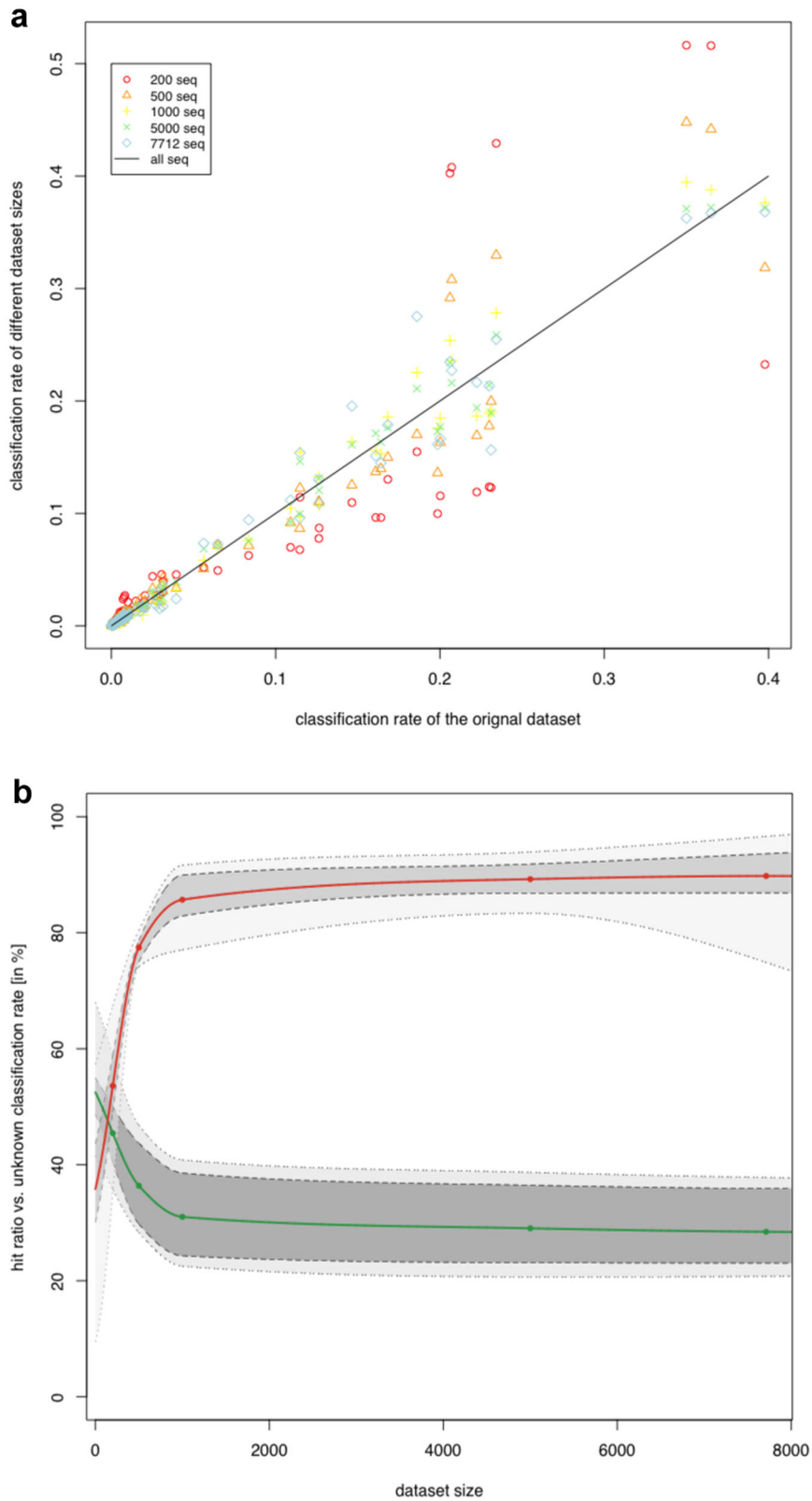
**Fig. 5** Heatmap showing the predicted relative contribution of specific genera from different source environments. Only taxa that are associated with meat spoilage or include relevant pathogens are shown. Non-relevant taxa were grouped into a single group called “Other”. The relative contribution of each genus (columns) from each source (rows) is indicated by the color of the intersecting tile. Taxa are clustered by shared-source similarity. BCE: bacterial cell equivalents as determined by 16S rRNA gene qPCR.

<0.0001). Supplementary Table 3 shows that the datasets with 200 and 500 random sequences were significantly different to the larger dataset sizes with 5000 and 7712 sequences ( $p$ -value  $\leq 0.0034$ ), whereas datasets with 1000, 5000, and 7712 showed no significant differences to each other (Supplementary Table 3). The lowest average hit ratio of 53.6% was reached in the datasets with 200 random sequences and the highest average hit ratio of 89.8% was determined for the datasets with 7712 random sequences (Supplementary Table 2; Fig. 6b). A significant difference between all the datasets with regards to the hit ratio was identified ( $\chi^2 = 90.9$ ,  $p$ -value <0.0001). The hit ratio for the dataset with 200 and 500 random sequences is significantly different to the larger dataset sizes with 5000 and 7712 sequences ( $p$ -value  $\leq 0.0079$ ). The average unknown classification rate of contamination decreased with increasing number of sequences used i.e., from 45.4% for the dataset with 200 sequences to 28.4% with 7712 sequences (Fig. 6b). The average difference of unknown classified contamination compared to the original dataset ranged from 18.2% (dataset size 200) to 1.5% (dataset size 7712; Supplementary Table 2). The Kruskal–Wallis test detected significant differences between the variations of unknown classified contaminations among the dataset sizes ( $\chi^2 = 93.9$ ,  $p$ -value < 0.0001), which can be mainly assigned to the difference between datasets with 200 and 500 sequences vs. 5000 and 7712 sequences (Supplementary Table 3). In addition to a similar classification rate, the datasets with less sequences per sample also revealed comparable general shifts in microbial community structure (Supplementary Fig. 7).

## DISCUSSION

High-throughput sequencing technologies have proven to be a powerful approach to explore microbial communities in a large variety of natural habitats as well as in built environments. The decreasing costs of these tools now also offer new perspectives to implement them in food production in order to investigate the impact of a given shift on the microbiota and their roles in a food system, which is directly correlated to food safety, food shelf life, flavor, and many other aspects<sup>16</sup>. For example, Bridier et al. investigated the impact of disinfection procedures on the microbial ecology and *Salmonella* prevalence in a pig slaughterhouse using a combination of traditional culturing techniques and 16S rRNA gene sequencing<sup>17</sup>. Metagenomic and 16S rRNA gene profiling has been used in a variety of other meat and food-processing facilities to describe the spoilage associated microbiota and resistome in different facilities<sup>18–20</sup>. Here, we advance our knowledge about microbial diversity and biogeography along the pig meat processing chain by utilizing the high-throughput, long-read sequencing capability of the PacBio technology to obtain thousands of full-length 16S rRNA gene sequences. Although PacBio sequencing is not as cost-effective as some of the available short-read platforms such as Illumina, it is able to produce longer read lengths, resulting in higher resolution for taxonomic classification and microbial source tracking<sup>21</sup>.

First, we determined bacterial cell numbers with current standard techniques (Aerobic mesophilic and *Enterobacteriaceae* counts) and with additional methods (*Pseudomonadaceae* and BCE counts) to get a framework of the overall microbiological status of the slaughterhouse. This initial assessment revealed similar trends along the processing line as past investigations, and showed high



**Fig. 6 Simulation of the effect of sequencing depth on SourceTracker performance.** **a** The correlation of the different dataset sizes compared to the original dataset. **b** The hit ratio (red points) and unknown classified sequences (green points) for the different used dataset sizes, with extrapolation values for not tested dataset sizes shown as solid lines. Dotted lines presented the min–max values and the dashed lines indicated the 25th and 75th percentile of hit ratio and unknown classified sequences.

microbial numbers for several surface samples<sup>22,23</sup>. The high bacterial numbers on the polishing tunnel equipment and the significant increase of AMC, and *Pseudomonadaceae* counts on the meat from singeing to after polishing already indicate transfer of bacteria at this step. A significant increase in bacterial contamination after polishing was also reported by Wheatley et al.<sup>24</sup>. It is likely that the bacteria that survived the singeing treatment get spread over the carcass during polishing in addition of transfer of bacteria that persist in the polishing equipment<sup>25</sup>. Basic community analysis of the 16S rRNA gene sequencing data also exposed a higher microbial diversity after the polishing step when compared to after singeing, further indicating the same step as a possible transmission event.

When the pigs entered the facility, they had a high variation of the microbial community composition on the skin, which points out large differences of the individual pig skin microbiota. However, this variation is greatly reduced after the singeing step. Thus, the initial microbial community structure on the carcass surface has little impact on the effectiveness of decontamination measures like singeing, resulting in the establishment of a similar community on each carcass. The community then remains relatively stable until the loading station, where it starts to diversify again. This can be explained by the cooling period (16 h at 7 °C) between the classification and truck step, which could lead to higher attachment and differences in recovery of bacteria. Small variations in the psychrotolerant part of the microbial community could also cause disparate alterations during the cooling period.

The 50 most abundant ASVs that were detected are associated to genera that have been previously found within the meat processing environment and some of them, e.g. *Bacillus*, *Chryseobacterium*, *Flavobacterium*, *Lactococcus*, *Microbacterium*, and *Moraxella* are considered to have spoilage potential<sup>11,26–28</sup>. Interestingly, the meat samples were dominated by a number of ASVs associated to the genus *Anoxybacillus* after the singeing step. This genus consists of 22 species that were isolated from hot springs or manure, but were also regularly detected in dairy and meat-processing environments<sup>11,29–31</sup>. *Anoxybacillus* are aerobic or facultative anaerobic spore-formers, with an optimum growth at 50–65 °C and neutral pH, are alkali-tolerant, and are able to form biofilms<sup>32,33</sup>. Thus, they are able to survive the heat treatment (Scalding and singeing), explaining their high relative abundance on the meat after the singeing step. Overall, the microbial community on the meat samples (Singeing–Truck) was vastly different from the community on the skin of the animals when they entered the facility (Sticking). Hence, we hypothesized that the majority of the microorganisms were transferred onto the meat during processing and that only a smaller portion of the skin microbiota persists on the meat surface.

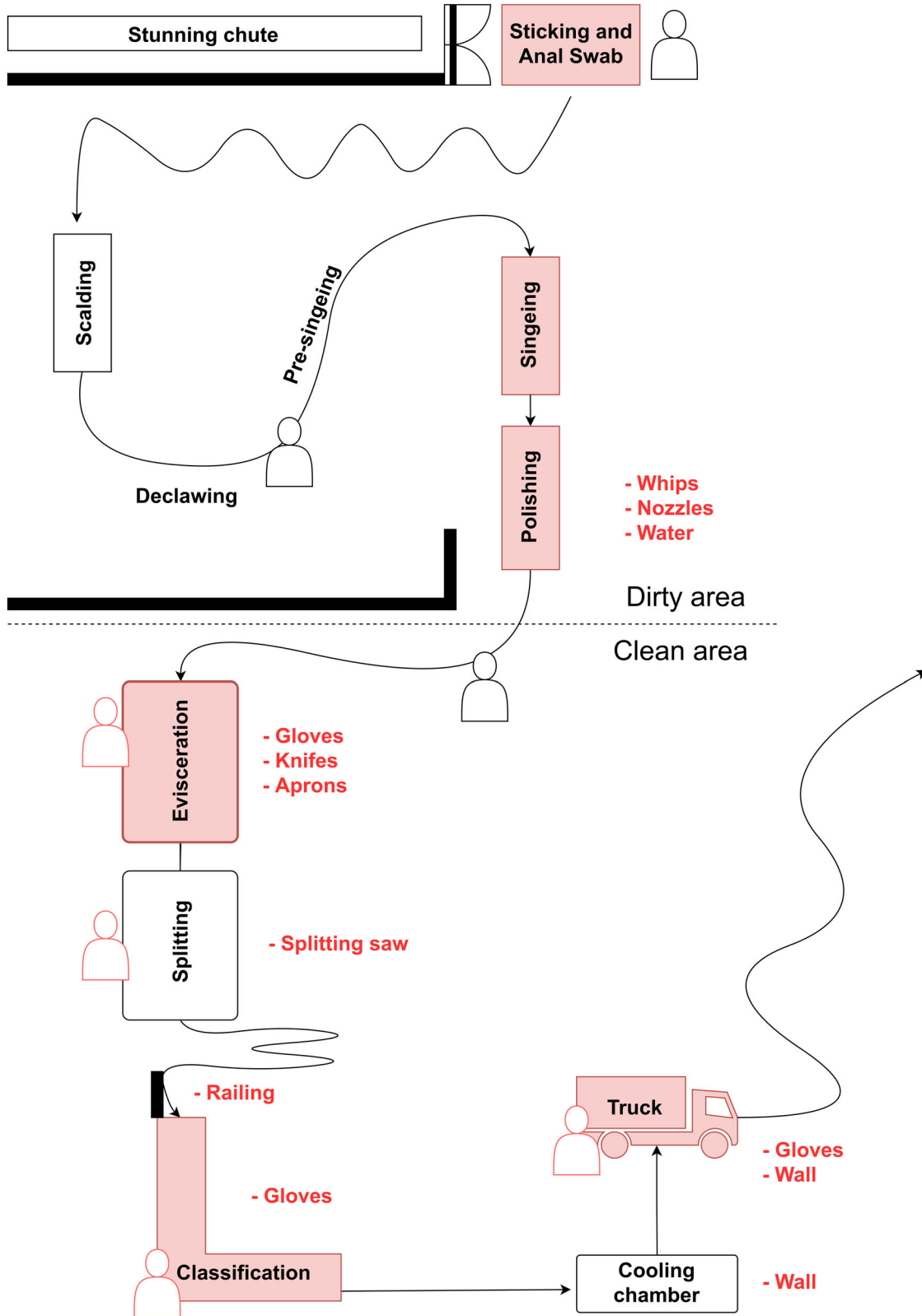
We then used the software SourceTracker, which applies a Bayesian framework to estimate the proportion of each source contributing to a designated sink sample, to trace transmission routes and track down the sources of these microorganisms<sup>34</sup>. This tool has been widely used to map microbial populations or gene flows in a variety of ecosystems, from coastal waters and drinking water systems to ATM keypads in New York, neonatal intensive care units, and to global antibiotic resistance gene pollution over diverse environmental types<sup>35–39</sup>. Surprisingly, only one study applied SourceTracker to a dataset obtained from a food environment (Brewery) so far<sup>14</sup>. Here, we use SourceTracker on full-length 16S rRNA gene sequencing data in the scope of a meat processing environment. Our analysis indicates key microbial transmission sites throughout the facility that were not identified with the current standard techniques. The main contamination sources contributing to the microbial community found on meat were the polishing tunnel equipment, gloves of employees, and a railing at the classification step. The polishing tunnel was identified as a critical operation during pork slaughtering in the

past<sup>40,41</sup>. However, the gloves and railings are generally not considered as such and had low microbial levels indicating good hygiene practices. Still a lot of microorganisms were transferred from these positions, albeit it is also possible that the transfer happened the other way around. It is important to note that we presume source/sink relationships in the SourceTracker analysis and can therefore not account for directionality, meaning that transfer of microorganisms could occur in both ways. A closer look at the taxonomy of the transferred bacteria exposed unique transmission patterns for individual taxa. Noticeably, particular species occupy different environmental niches across the facility showing the importance of high taxonomic resolution for microbial source tracking in food processing plants. In fact, whole genome sequencing (WGS) or metagenomic shotgun sequencing have been proposed to achieve strain-level resolution and are thought to be necessary to track microbes during food processing<sup>42–44</sup>. WGS has been widely adopted as a surveillance tool for specific pathogens and outbreak investigations across the world in recent years<sup>45,46</sup>. Related to our study, Nastasišević et al. have successfully applied WGS to track the main entry routes of *L. monocytogenes* in a meat establishment<sup>47</sup>. However, all these studies state costs as one of the major drawbacks of the use of WGS, especially for developing countries. While we were not able to identify potential pathogens or specific spoilage organisms on a strain level, our results show that full-length 16S rRNA gene sequencing delivers a deep enough resolution for environmental monitoring within a facility. The source/sink relationships predicted by SourceTracker can be used to infer consistent contamination events. Thus, strain-level tracking using metagenomics is still essential for molecular epidemiology and diagnostics, but our approach could be a more cost effective solution for regular monitoring systems and the identification of general transmission routes.

Moreover, we were able to show that it is not necessary to deeply sequence the amplicon libraries, but that datasets with 1000 sequences/sample provide a comparable result to deeper sequenced datasets in terms of beta diversity assessment and SourceTracker classification rate in the analysed facility. Additional studies are necessary to investigate whether this is generally true or is specific only to the analysed facility. Thus, it is feasible to multiplex many samples on a single sequencing run substantially decreasing costs, which essentially accomplishes affordability for regular monitoring checks. However, several challenges and limitations remain before we can realize the full potential of NGS techniques as food safety applications. Currently, bioinformatics workflows are implemented and executed based on ad hoc lab specific experiences. Each lab uses different protocols and the documentation of the actual process is rarely well recorded or presented. Hence, it is paramount to expose, formalize, and standardize sampling techniques, as well as workflows for bioinformatics pipelines, processing, and data management<sup>48–50</sup>. In that way, it would be possible to compare datasets and leverage systematic authentication of the microbiome and its variation throughout the supply chain to understand microbial contamination during food production on a broader scale.

Since we observed only one significant increase in microbial numbers along the slaughter line, we consider AMC as good general hygiene indicators that can reflect substantial contamination incidents, but they fail in describing more complex population flows. Thus, AMC and EB determination by microbiological reference methods is not sufficient to detect the full extent of microbial transmission events. This study showcases that high-throughput full-length 16S rRNA gene sequencing can reveal valuable information about the microbial communities in pork production plants and expose critical contamination steps during slaughtering. We were able to pinpoint many taxa to specific sources, facilitating targeted combat of potential pathogens or meat-spoilage organisms in the





**Fig. 7 Schematic map of the slaughterhouse depicting the whole processing line.** Red shaded boxes indicate locations where skin/meat was sampled. Red text specifies environmental samples taken at corresponding positions. Human sketches point out non-automated locations with working employees.

analysed facility. Our findings contribute to improve or optimize hygiene standards in the meat industry to further minimize the risk of microbial cross-contamination. Furthermore, the methods used in this study can be applied to any other food-related industry to universally promote our knowledge about microbial transfer during food processing. Continuing advances in long-read sequencing technologies like the release of the Sequel II system and the development of full rRNA operon sequencing strategies will further increase the throughput and taxonomic resolution, offering great potential to implement them in monitoring systems<sup>51</sup>.

## METHODS

### Facility structure and sampling

Samples for this study were taken from an Austrian slaughterhouse with a capacity of 200–250 pigs per hour. Figure 7 illustrates the structural design of the production chain and the sampling positions. Antemortem inspection took place during the offload and holding. The pigs were herded through a series of gates in small groups until they were finally presented for stunning one by one. After electrocution to the head, each pig's main throat vasculature was cut with a knife and the bodies hanged up to exsanguinate (Position Sticking). Then the carcasses were scalded for 3 min by steam condensation, after which the claw shoes were manually removed. Before the carcasses entered the singeing tunnel for approximately 6 seconds, they were dehaired by rotating whips and manually pre-singed. Polishing was then performed in another tunnel with rotating whips and water spray (Position Polishing). Eyes and external ear canal were removed from the bodies as they were moved from the "dirty area" to the "clean area". Operators were not allowed to move from one area to the other. To avoid contaminations from feces, the rectum was sealed off with a small plastic bag. Evisceration was the first step in the clean area, followed by splitting the carcasses in halves with a saw, post mortem inspection, removal of the spinal cord with an aspirator, and classification. After classification, the carcasses passed a shock shower with 4 °C cold water and entered the cooling chamber, where they were held for up to 16 h at 7 °C. The chilled carcasses were then transported on a refrigerated truck to a meat cutting facility.

Twelve pigs from three different farms (four pigs each) were used for this study. Carcass and environmental samples were taken with sterile polyurethane sponges (SampleRight™ Sponge Sampler, World Bioproducts, Woodinville, USA), which recover significantly higher amounts of bacteria compared to sponges made from cellulose, reaching a similar recovery rate as excision methods<sup>52</sup>. The sampling area was 100 cm<sup>2</sup> at the back of the carcasses. The back was chosen based on the results of a preliminary test already conducted prior the start of this study, which showed that samples taken from the back harbor similar levels of relevant microorganisms compared to samples taken from the belly or the musculature along the cutting area<sup>53</sup>. In total, 84 swabs were taken from carcass surfaces at different processing positions. When the pigs entered the facility (Position Sticking), their skin was sampled, while at the other positions their meat was sampled. The carcasses chosen for sampling were ear tagged and followed throughout the entire processing chain so that the same carcasses could be sampled at each position. In addition, 75 swabs were taken from equipment, staff, and infrastructure of the facility, resulting in 159 samples overall. The sampling was done over a period of one day. For a detailed list of samples and sampling positions see supplementary Table 4. For sampling, the sponge was swabbed for 10 s horizontally, then flipped and swabbed again for 10 s in vertical direction. Then, the sponge was placed back into the sterile plastic bag, which was sealed and chilled in a container placed in the cooling chamber of the facility (4 °C) until sampling was finished. A fresh polyurethane sponge, a new sterile template of 100 cm<sup>2</sup> and new gloves were used for each new sample. All the samples were transported back to the lab on ice (Transport time: 2 h). Back in the laboratory, each sponge was squeezed thoroughly, and the obtained liquid was split into two 15 ml falcon tubes, one of which was directly used for cultivation experiments, while the other one was stored at –20 °C until further processing (Molecular analysis).

### Microbiological investigation

The enumeration of aerobic mesophilic counts (AMC) (ISO 4833-2:2013), *Enterobacteriaceae* (EB) (ISO 21528-2:2017) and *Pseudomonadaceae* (PS) was performed after preparing a ten-fold serial dilution in buffered peptone water (BPW) (Thermo Fisher Scientific Inc., Oxoid Ltd.,

Basingstoke, UK) up to dilution –10<sup>8</sup>. The dilutions were plated in duplicates on Plate Count Agar (PCA, Thermo Fisher Scientific Inc., Oxoid Ltd.), Violet Red Bile Glucose (VRBG, Thermo Fisher Scientific Inc., Oxoid Ltd.) and Glutamate Starch Phenol Red Agar (GSP, Merck KGaA; Darmstadt, Germany) by surface plating technique. GSP and VRBG agar were incubated at 25 and 37 °C for 24–48 h. PCA was incubated at 30 °C for a maximum of 72 h. To determine the AMC/EB and PS counts/cm<sup>2</sup>, microbial colonies between 10 and 300 colony forming units (CFU) were included in the calculation. Presumptive EB and PS isolates ( $n=5$  each) were confirmed by Oxidase reaction and biochemical profiling up to genus level for PS and up to species level for EB (API 20E, bioMérieux Marcy-l'Étoile, France).

### DNA-extraction, qPCR, and 16S rRNA gene sequencing

In order to increase microbial cell density, samples were centrifuged at 3220× rcf for 20 min and the pellet was resuspended in 400 µl of 1× phosphate buffered saline (PBS). The DNA was then extracted from 200 µl with the QIAamp DNA Stool Mini Kit (Qiagen GmbH, Hilden, Germany) according to manufacturer instructions. The elution step of the protocol was modified; instead of 200 µl AE buffer, two times 25 µl DEPC treated water was used. Negative controls (DEPC treated water), one for each used kit, were also extracted together with the regular samples. The DNA concentration of the samples was measured with the Qubit dsDNA HS Assay Kit and Qubit 2.0 Fluorometer (Invitrogen, Thermo Fisher Scientific, Oregon, USA).

The 16S rRNA gene was amplified by qPCR to enumerate total bacterial cell equivalents (BCE) as previously described (Supplementary Table 5<sup>53,54</sup>). All qPCR samples and standards were run in duplicates and negative controls were included in each run (Mx3000P qPCR thermocycler, analyzed with MxPro v.4.10 (Stratagene, San Diego, USA)). Total BCE were extrapolated with an average of four 16S rRNA gene copies as estimated by rrnDB, a database for ribosomal RNA operon variation in bacteria and archaea<sup>55,56</sup>. Statistics for qPCR and plate count data were tested for normal distribution using the Shapiro–Wilk normality test and with visual assessment of qqplots and histograms. The not normal distributed groups were tested using the Wilcoxon-test for connected samples. We applied Benjamini–Hochberg procedure to reduce the false discovery rate. To investigate correlations between BCE and AMC counts, spearman rank correlation coefficients were calculated for individual sampling positions. Data are considered significant at  $p \leq 0.05$ .

Amplicon library generation, quality control, and sequencing were performed at the Vienna Biocenter Core Facilities NGS Unit ([www.vbcf.ac.at](http://www.vbcf.ac.at)). Full-length 16S rRNA gene libraries of 133 samples (including three negative controls) were prepared using bacteria specific primers 27F (5'-AGRGTTYGATYMTGGCTCAG-3') and 1492R (5'-RGYTACCTGTTCAGACTT-3'). Barcodes were added during a second round of amplification with Pacbio Barcoded Universal primers, so that the amplicons could be multiplexed on three SMRT cells. Sequencing was carried out on a Pacbio Sequel machine with 2.1 chemistry. Detailed library preparation and sequencing procedure is available online <https://www.pacb.com/wp-content/uploads/Procedure-Checklist-Full-Length-16S-Amplification-SMRTbell-Library-Preparation-and-Sequencing.pdf>. Each SMRT cell generated approximately 50 GB of raw data producing 641,939 sequences/cell on average.

In addition to the full-length 16S rRNA gene of 133 samples sequenced on a Pacbio Sequel machine, the V3–V4 region of the 16S rRNA gene from 52 of these samples was also sequenced on an Illumina MiSeq sequencing platform with a 300 bp paired-end read protocol. The PCR reactions were performed as described in Klindworth et al. using the forward primer 341f (5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG) and the reverse primer 785r (5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG)<sup>57</sup>.

### Sequence processing and analysis

Accurate full-length 16S rRNA gene sequences were generated using Pacbio's single-molecule circular consensus sequencing. The circular consensus reads (ccs) were determined with a minimum predicted accuracy of 0.99 and the minimum number of passes set to three in the SMRT Link software package 5.1<sup>58</sup>. After demultiplexing, the ccs were further processed with DADA2 (version 1.9.1) to obtain amplicons with single-nucleotide resolution<sup>59,60</sup>. Similarly, sequences generated on Illumina's MiSeq platform were also processed with DADA2 and equivalent parameters in order to achieve a maximum comparability between the two datasets. Amplicon sequence variants (ASVs) were assigned a taxonomy using a DADA2 formatted version of the genome taxonomy database release 03-RS86<sup>61,62</sup>.

After initial quality filtering, samples with less than 200 reads and ASVs with less than five reads were removed. Additionally, contaminant ASVs were detected and removed with the R package “decontam” using a prevalence-based contaminant identification with a threshold value cutoff of 0.5<sup>63</sup>. Microbial community analysis was performed within the “phyloseq” and “tsnemiobiota” packages and visualized with ggplot2 in R<sup>64–66</sup>. Alpha diversity indices were calculated with a dataset rarefied to the minimum sample size. Normal distribution of individual alpha diversity indices was tested with the Shapiro–Wilk normality test and with visual assessment of qqplots and histograms. The not normal distributed groups were tested using the Wilcoxon-test and normal distributed groups were tested using a *t*-test for connected samples. We applied Benjamini–Hochberg procedure to reduce the false discovery rate.

Microbial source tracking was achieved with the software SourceTracker (version 1.0.0) and default parameters<sup>34</sup>. Samples taken from the facility environment and from the skin of the animals were assigned as sources whereas meat samples were assigned as sinks. The Illumina dataset (min: 7712; mean: 19,119; max: 30,340 sequences per sample) was rarefied to 7712, 5000, 1000, 500, and 200 reads per sample prior to SourceTracker analysis in order to infer the influence of sequencing depth on the performance of SourceTracker. To investigate the variation of same sized datasets this procedure was repeated three times for the larger datasets and ten times for the dataset with 200 sequences, because we expected a higher variance when using 200 random sequences. For each dataset size the goodness of fit was determined. In detail, the squared difference between randomly generated datasets and the original dataset (i.e., including all sequences) was calculated and the match rate for both, rarefied and original dataset were checked for correlation with spearman correlation coefficient and correlation plots. The normal distribution of the squared differences for each dataset size was investigated by using the Shapiro–Wilk test and the homoscedasticity was analysed by applying the Levene test. Due to non-normal distribution of the data, a Kruskal–Wallis rank sum test, followed by a Dunn test with Benjamini–Hochberg alpha-adjustment for post-hoc analysis was applied in order to compare the goodness of fit of the different dataset sizes. Additionally, the hit ratio of each dataset size was calculated. The hit ratio is defined as percentage of correctly assigned contamination sources (including ASVs with an unknown source), whereby the original dataset was used as reference. Further a comparison of amounts of unknown classified ASVs (unknown classification rate) was accomplished. Due to differences between the production steps regarding to the correctly assigned samples in the unknown sources, the differences between the amounts of unknown classified ASVs for each randomly generated dataset and the original dataset were separately calculated. The significant differences between the hit ratios and the rate of unknown classified ASVs were also analysed with Kruskal–Wallis rank sum test and Dunn test for post-hoc analysis. To further test the consistency of results at different rarefaction depths, beta diversity was calculated for each of the rarefied datasets. In addition, the R package “BAT” was used to check the reliability of beta diversity results according to sequencing depth<sup>67</sup>. Detailed parameters of our sequence processing workflow and all statistical analysis can be found in the supplements (Supplementary Note 1) and are freely available as an R markdown script at GitHub (github.com/FFoQS90).

## Reporting summary

Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

## DATA AVAILABILITY

Raw sequence reads are available on the European Nucleotide Archive under the accession number PRJEB37434.

Received: 25 November 2019; Accepted: 11 June 2020;

Published online: 10 July 2020

## REFERENCES

- Buzby, J. C., Wells, H. F. & Hyman, J. *The Estimated Amount, Value, and Calories of Postharvest Food Losses at the Retail and Consumer Levels in the United States*. (EIB-121, U.S. Department of Agriculture, Economic Research Service, Washington, 2014).
- Huis In't Veld, J. H. J. Microbial and biochemical spoilage of foods: an overview. *Int. J. Food Microbiol.* **33**, 1–18 (1996).
- Havelaar, A. H. et al. World Health Organization global estimates and regional comparisons of the burden of foodborne disease in 2010. *PLoS Med.* **12**, e1001923 (2015).
- EFSA (European Food Safety Authority) and ECDC (European Centre for Disease Prevention and Control), The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2015. *EFSA J.* **14**(12): 4634, 231, (2016).
- Gill, C. O. Meat spoilage and evaluation of the potential storage life of fresh meat. *J. Food Prot.* **46**, 444–452 (1983).
- Giaouris, E. et al. Attachment and biofilm formation by foodborne bacteria in meat processing environments: causes, implications, role of bacterial interactions and control by alternative novel methods. *Meat Sci.* **97**, 289–309 (2014).
- Choi, Y. M. et al. Changes in microbial contamination levels of porcine carcasses and fresh pork in slaughterhouses, processing lines, retail outlets, and local markets by commercial distribution. *Res. Vet. Sci.* **94**, 413–418 (2013).
- Sheridan, J. J. Sources of contamination during slaughter and measures of control. *J. Food Saf.* **18**, 321–339 (1998).
- International Organization for Standardization. *Microbiology of the Food Chain—Carcass Sampling for Microbiological Analysis*. (2015). ISO 17604:2015, Retrieved from <https://www.iso.org/standard/62769.html>
- Nocker, A., Burr, M. & Camper, A. K. Genotypic microbial community profiling: a critical technical review. *Microb. Ecol.* **54**, 276–289 (2007).
- Hultman, J., Rahkila, R., Ali, J., Rousu, J. & Björkroth, K. J. Meat processing plant microbiome and contamination patterns of cold-tolerant bacteria causing food safety and spoilage risks in the manufacture of vacuum-packaged cooked sausages. *Appl. Environ. Microbiol.* **81**, 7088–7097 (2015).
- Chaillou, S. et al. Origin and ecological selection of core and food-specific bacterial communities associated with meat and seafood spoilage. *ISME J.* **9**, 1105–1118 (2015).
- Yang, H. et al. Uncovering the composition of microbial community structure and metagenomics among three gut locations in pigs with distinct fatness. *Sci. Rep.* **6**, 27427 (2016).
- Bokulich, N. A., Bergsveinson, J., Ziola, B. & Mills, D. A. Mapping microbial ecosystems and spoilage-gene flow in breweries highlights patterns of contamination and resistance. *Elife* **4**, e04634 (2015).
- Mann, E. et al. Psychrophile spoilers dominate the bacterial microbiome in musculature samples of slaughter pigs. *Meat Sci.* **117**, 36–40 (2016).
- Bokulich, N. A., Lewis, Z. T., Boundy-Mills, K. & Mills, D. A. A new perspective on microbial landscapes within food production. *Curr. Opin. Biotechnol.* **37**, 182–189 (2016).
- Bridier, A. et al. Impact of cleaning and disinfection procedures on microbial ecology and *Salmonella* antimicrobial resistance in a pig slaughterhouse. *Sci. Rep.* **9**, 12947 (2019).
- Kang, S., Ravensdale, J., Coorey, R., Dykes, G. A. & Barlow, R. A comparison of 16S rRNA profiles through slaughter in Australian export beef abattoirs. *Front. Microbiol.* **10**, 2747 (2019).
- Stellato, G. et al. Overlap of spoilage microbiota between meat and meat processing environment in small-scale 2 vs. large-scale retail distribution. *Appl. Environ. Microbiol.* **82**, 4045–4054 (2016).
- Campos Calero, G. et al. Deciphering resistome and virulome diversity in a porcine slaughterhouse and pork products through its production chain. *Front. Microbiol.* **9**, 2099 (2018).
- Johnson, J. S. et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat. Commun.* **10**, 5029 (2019).
- Spescha, C., Stephan, R. & Zweifel, C. Microbiological contamination of pig carcasses at different stages of slaughter in two European Union—approved abattoirs. *J. Food Prot.* **69**, 2568–2575 (2006).
- Warriner, K., Aldsworth, T. G., Kaur, S. & Dodd, C. E. R. Cross-contamination of carcasses and equipment during pork processing. *J. Appl. Microbiol.* **93**, 169–177 (2002).
- Wheatley, P., Giotis, E. S. & McKeivitt, A. I. Effects of slaughtering operations on carcass contamination in an Irish pork production plant. *Ir. Vet. J.* **67**, 1 (2014).
- Gill, C. O. in *Woodhead Publishing Series in Food Science, Technology and Nutrition* (ed. Sofos, J. N. et al.) 630–672 (Woodhead Publishing, Sawston, 2005). <https://doi.org/10.1533/9781845691028.2.630>
- de Filippis, F., La Stora, A., Villani, F. & Ercolini, D. Exploring the sources of bacterial spoilers in beefsteaks by culture-independent high-throughput sequencing. *PLoS ONE* **8**, e70222 (2013).
- de Smidt, O. The use of PCR-DGGE to determine bacterial fingerprints for poultry and red meat abattoir effluent. *Lett. Appl. Microbiol.* **62**, 1–8 (2016).
- Andrew, D. & Board, R. *Microbiology of Meat and Poultry*. (Blackie Academic & Professional, Glasgow, 1998).

29. Khan, I. U. et al. *Anoxybacillus sediminis* sp. nov., a novel moderately thermophilic bacterium isolated from a hot spring. *Antonie Van. Leeuwenhoek* **111**, 2275–2282 (2018).
30. Pikuta, E. et al. *Anoxybacillus pushchinensis* gen. nov., sp. nov., a novel anaerobic, alkaliphilic, moderately thermophilic bacterium from manure, and description of *Anoxybacillus flavithermus* comb. nov. *Int. J. Syst. Evol. Microbiol.* **50**, 2109–2117 (2000).
31. Burgess, S. A., Lindsay, D. & Flint, S. H. Thermophilic bacilli and their importance in dairy processing. *Int. J. Food Microbiol.* **144**, 215–225 (2010).
32. Burgess, S. A., Brooks, J. D., Rakonjac, J., Walker, K. M. & Flint, S. H. The formation of spores in biofilms of *Anoxybacillus flavithermus*. *J. Appl. Microbiol.* **107**, 1012–1018 (2009).
33. Goh, K. M. et al. Recent discoveries and applications of *Anoxybacillus*. *Appl. Microbiol. Biotechnol.* **97**, 1475–1488 (2013).
34. Knights, D. et al. Bayesian community-wide culture-independent microbial source tracking. *Nat. Methods* **8**, 761–763 (2011).
35. Henry, R. et al. Into the deep: evaluation of sourcetracker for assessment of faecal contamination of coastal waters. *Water Res.* **93**, 242–253 (2016).
36. Liu, G. et al. Assessing the origin of bacteria in tap water and distribution system in an unchlorinated drinking water system by SourceTracker using microbial community fingerprints. *Water Res.* **138**, 86–96 (2018).
37. Bik, H. M. et al. Microbial community patterns associated with automated teller machine keypads in New York City. *mSphere* **1**, e00226–16 (2016).
38. Hewitt, K. M. et al. Bacterial diversity in two neonatal intensive care units (NICUs). *PLoS ONE* **8**, e54703 (2013).
39. Li, L.-G., Yin, X. & Zhang, T. Tracking antibiotic resistance gene pollution from different sources using machine-learning classification. *Microbiome* **6**, 93 (2018).
40. Bolton, D. J. et al. Washing and chilling as critical control points in pork slaughter hazard analysis and critical control point (HACCP) systems. *J. Appl. Microbiol.* **92**, 893–902 (2002).
41. Yu, S. L. et al. Effect of dehairing operations on microbiological quality of swine carcasses. *J. Food Prot.* **62**, 1478–1481 (1999).
42. Jagadeesan, B. et al. The use of next generation sequencing for improving food safety: translation into practice. *Food Microbiol.* **79**, 96–115 (2019).
43. Bergholz, T. M., Moreno Switt, A. I. & Wiedmann, M. Omics approaches in food safety: fulfilling the promise? *Trends Microbiol.* **22**, 275–281 (2014).
44. Leonard, S. R., Mammel, M. K., Lacher, D. W. & Elkins, C. A. Application of metagenomic sequencing to food safety: detection of shiga toxin-producing *Escherichia coli* on fresh bagged spinach. *Appl. Environ. Microbiol.* **81**, 8183–8191 (2015).
45. Moura, A. et al. Real-time whole-genome sequencing for surveillance of listeria monocytogenes, France. *Emerg. Infect. Dis.* **23**, 1462–1470 (2017).
46. Wang, S. et al. Food safety trends: from globalization of whole genome sequencing to application of new tools to prevent foodborne diseases. *Trends Food Sci. Technol.* **57**, 188–198 (2016).
47. Nastasijevic, I. et al. Tracking of listeria monocytogenes in meat establishment using whole genome sequencing as a food safety management tool: a proof of concept. *Int. J. Food Microbiol.* **257**, 157–164 (2017).
48. Weimer, B. C. et al. Defining the food microbiome for authentication, safety, and process management. *IBM J. Res. Dev.* **60**, 1:1–1:13 (2016).
49. Köster, J. & Rahmann, S. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* **28**, 2520–2522 (2012).
50. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
51. Martijn, J. et al. Confident phylogenetic identification of uncultured prokaryotes through long read amplicon sequencing of the 16S-ITS-23S rRNA operon. *Environ. Microbiol.* **21**, 2485–2498 (2019).
52. Pearce, R. A. & Bolton, D. J. Excision vs sponge swabbing—a comparison of methods for the microbiological sampling of beef, pork and lamb carcasses. *J. Appl. Microbiol.* **98**, 896–900 (2005).
53. Zwirzitz, B. et al. Culture-independent evaluation of bacterial contamination patterns on pig carcasses at a commercial slaughter facility. *J. Food Prot.* **82**, 1677–1682 (2019).
54. Muyzer, G., De Waal, E. C. & Uitterlinden, A. G. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl. Environ. Microbiol.* **59**, 695–700 (1993).
55. Stoddard, S. F., Smith, B. J., Hein, R., Roller, B. R. K. & Schmidt, T. M. rnmDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Res.* **43**, D593–D598 (2015).
56. Větrovský, T. & Baldrian, P. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS ONE* **8**, 1–10 (2013).
57. Klindworth, A. et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**, 1–11 (2013).
58. Pacific Biosciences SMRT® Tools Reference Guide. (2018).
59. Callahan, B. J. et al. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
60. Callahan, B. J. et al. High-throughput amplicon sequencing of the full-length 16S rRNA gene with single-nucleotide resolution. *Nucleic Acids Res.* **47**, e103–e103 (2019).
61. Alishum, A. et al. DADA2 formatted 16S rRNA gene sequences for both bacteria & archaea. <https://doi.org/10.5281/zenodo.2541239> (2019).
62. Parks, D. H. et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996–1004 (2018).
63. Davis, N. M., Proctor, D., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *bioRxiv* **221499**, (2017).
64. McMurdie, P. J. & Holmes, S. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217 (2013).
65. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer, New York, 2016).
66. Lindstrom, J. C. Tsnemicrobiota: T-distributed stochastic neighbor embedding for microbiota data. (2017). Github Repository, <https://github.com/opisthokonta/tsnemicrobiota>
67. Cardoso, P., Rigal, F. & Carvalho, J. C. BAT—biodiversity Assessment Tools, an R package for the measurement and estimation of alpha and beta taxon, phylogenetic and functional diversity. *Methods Ecol. Evol.* **6**, 232–236 (2015).

## ACKNOWLEDGEMENTS

We thank Viktoria Neubauer for fruitful discussions and proofreading of the manuscript and Nikolaus Pfisterer and Christian Mattes for their assistance during sampling. Additionally, we thank Lauren Alteio for R package recommendations and long-distance emotional support. The competence centre FFOQSI is funded by the Austrian ministries BMVIT, BMDW and the Austrian provinces Niederösterreich, Upper Austria and Vienna within the scope of COMET –Competence Centers for Excellent Technologies. The programme COMET is handled by the Austrian Research Promotion Agency FFG.

## AUTHOR CONTRIBUTIONS

B.Z., S.W., M.W., and E.M. conceived and designed the study. Sampling was performed by B.Z., S.W., I.R., and B.S. B.Z., I.R., S.T., M.D., and B.S. performed the experiments and B.Z., E.D., C.S., and A.Z. developed bioinformatics pipelines. Data analysis and statistics were performed by B.Z., E.M., B.P., J.Z., and F.F.R.; B.Z. and E.M. wrote the manuscript. All authors read and approved the final manuscript.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41522-020-0136-z>.

**Correspondence** and requests for materials should be addressed to B.Z.

**Reprints and permission information** is available at <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