Influence of pig farming on the human’s nasal microbiota: The key role of the airborne microbial communities

Julia G Kraemer¹,², Alban Ramette¹, Suzanne Aebi¹, Anne Oppliger² and Markus Hilty¹, *

¹ Institute for Infectious Diseases, University of Bern, Bern, Switzerland
² Institute for Work and Health (IST), University of Lausanne and University of Geneva, Epalinges, Switzerland

Running Head: Pig farm associated nasal microbiota

*Address correspondence to Markus Hilty, PhD, Institute for Infectious Diseases, University of Bern, Friedbühlstrasse 51, 3010 Bern, Switzerland. E-mail: markus.hilty@ifik.unibe.ch.

Word count Abstract: 241; Main text: 4073
ABSTRACT

It has been hypothesized that the environment can influence the composition of the nasal microbiota. However, the direct influence of pig farming on the anterior and posterior nasal microbiota is unknown. Using a cross-sectional design, pig farms (n=28) were visited in 2014-2015 and nasal swabs from 43 pig farmers and 56 pigs as well as 27 air samples taken in the vicinity of pig enclosure were collected. As controls, nasal swabs from 17 cow farmers and 26 non-animal exposed individuals were also included. Analyses of the microbiota were performed based on 16S rRNA amplicon sequencing and the DADA2 pipeline to define sequence variants (SVs). We found that pig farming is strongly associated with specific microbial signatures (including alpha- and beta-diversity), which are reflected in the microbiota of the human nose. Furthermore, the microbial communities were more similar within the same farm as compared to between the different farms, indicating a specific microbiota pattern for each pig farm. In total, there were 82 SVs that occurred significantly more abundantly in samples from pig farms than from cow farmers and non-exposed (i.e. the core pig farm microbiota). Of those, nine SVs were significantly associated with the posterior part of the humans’ nose. The results strongly indicate that pig farming is associated with a distinct human nose microbiota. Finally, the community structures derived by the DADA2 pipeline showed an excellent agreement with the outputs of the mothur pipeline which was revealed by procrustes analyses.

Importance

The knowledge about the influence of animal keeping on the human microbiome is important. Previous research shows that pets are significantly affecting the microbial communities of humans. However, the effect of animal farming on the human microbiome is less clear although it is known that the air in farms, and in particular pig farms, is charged with high amounts of dust, bacteria and fungi. In this study we have simultaneously investigated the nasal microbiota of pigs, humans and the environment in pig farms. We reveal an enormous impact of pig farming on the human nasal microbiota which is far more pronounced as compared to cow farming. In addition, we have analyzed the airborne microbiota and found significant associations suggesting an animal-human transmission of the microbiota within pig
farms. We also reveal that microbial patterns are farm-specific suggesting that the environment
influences animals and humans in a similar manner.
INTRODUCTION

The human nares are an important niche for bacterial colonization by both, pathogens and commensals and it is one of the main interfaces between the internal body and the external environment. Pig farmers are exposed to a complex and heterogeneous environment including large amounts of bacteria on a daily basis (1) and swine represent a potential reservoir for many pathogens which can be transmitted to humans, such as Streptococcus suis and Clostridium difficile (2). Also, there is a growing concern with the transmission of antibiotic resistant bacteria, such as methicillin-resistant Staphylococcus aureus (MRSA) in pig farms and other livestock-associated areas (3-7). A considerable number of studies has been published, showing transmission of these bacteria from pigs to humans (for reviews, see 2, 8, 9, 10). However, previous studies mainly focused on the investigation of only one or two bacterial species and were culture–dependent, but the overall impact on the entire human microbiota has never been investigated.

A recent study investigated 25 households containing 56 pets and 30 humans and revealed that household membership was strongly associated with microbial communities, in both humans and pets, using culture-independent, next generation sequencing methods (11). In another, longitudinal study, evidence for substantial exchanges among human, home, and pet microbiota were shown as well (12). The authors concluded that such interactions could have considerable human and animal health implications. Some studies have also shown that living or working with animals can protect against asthma and atopic diseases due to the exposure to specific animal microorganisms (13, 14). However, despite the relevance, the pattern of the microbiota exchange among animals, humans and environment in pig farms has never been investigated. The aims of our study were to 1) describe the influence of pig farming on the human nasal microbiota, 2) identify the sequence variants (SVs) predominantly shared between pigs, air from the pig enclosure and pig farmers, 3) identify which of the latter were significantly associated with either the posterior or anterior nasal cavities of pig farmers and 4) to compare the findings derived by DADA2 with the outputs of the more traditionally used mothur pipeline.
RESULTS

Characterization of study cohort and sequence analysis

Details of the sampling population can be found in Table 1. In total, 28 pig farms were visited, on which one to three pigs (total n=56), one air sample (total n=27) and one to four pig farmers (total n=43) were sampled (Table 1). As control, individuals with contact to cows but no contact to pigs (cow farmers, n=17), and individuals without contact to any type of farm animal (non-exposed persons, n=26) working in offices were chosen to assess the effect of pig exposure on the human nasal microbiota. All individuals were recruited in the same geographical area and were roughly age-matched. After exclusion of 17 samples due to PCR amplification issues, 255 samples with a total of 9,692,391 reads were included in our study. The mean number of reads per sample was 38,009 (± standard deviation 19,412) ranging from 2,243 to 120,642 reads. Reads were clustered into a total of 13,585 SVs (Sequence Variants). Sequencing depth was sufficient, as determined by the low slope of the rarefaction curves (Supplementary Fig. S1).

Pig farming is associated with increased diversity

All 13,585 SVs were grouped into 43 phyla and 310 families and the phyla Actinobacteria, Bacteroidetes, Firmicutes and Proteobacteria included the majority of all SVs (at least 97% mean relative abundance for all sample groups). Pig-farmer nasal samples showed the highest Shannon diversity indices and richness, while non-exposed samples displayed the lowest SDI and richness values (Fig. 1, A and B). To take into account that multiple samples were collected on the same farms, we additionally performed a linear mixed regression with the location ID as random effect to compare the differences between groups. The overall model was significant (analysis of variance; SDI P value<0.001, richness P value<0.001), and revealed that the bacterial richness in nasal samples from pig farmers was significantly higher than those of non-exposed individuals (SDI: P < 0.001; richness: P < 0.001), air samples (SDI: P = 0.03; richness: P = 0.001), and pig nasal samples (SDI: P < 0.001; richness: P < 0.001). The alpha diversity in cow farmers was nearly as high as in pig farmers and the differences were also significant when compared to non-exposed (SDI: P < 0.001; richness: P < 0.001) and pig nasal samples (SDI: P = 0.003; richness: P = 0.002).
Pig farming influences the microbial community composition

The ordination method based NMDS plots with weighted and unweighted input (Fig. 1, C and D) showed a distinct clustering of pigs, air, pig farmers, cow farmers and non-exposed and was confirmed by permutational multivariate analysis of variance (PERMANOVA, unweighted: $F$-value: 0.15, $P < 0.001$, weighted: $F$-value: 0.18, $P < 0.001$). Profiles of cow farmers were more similar to non-exposed controls as compared to pig farmers, indicating a very strong effect of pig farming on the human microbiota.

Analysis of similarity (ANOSIM) further confirmed the strong differences between pig farmer and cow farmer/non-exposed samples (Supplementary table S1). Interestingly, pig farmers seemed to display a significantly lower beta-diversity dispersion as compared to cow farmers and non-exposed individuals (weighted distances from the centroid; Tukey’s HSD test; $P < 0.001$; Fig. 1, E), indicating that pig farming leads to a more homogenous microbial community structure. All comparisons of unweighted distances from the centroid were non-significant (Tukey’s HSD tests; $P < 0.05$; Fig. 1, F), suggesting more of an effect of community structure than community composition on variation in beta-diversity across groups of samples.

We next examined how many SVs were shared between sample types; 54% of all SVs occurring in pig farmers also occurred in pigs and/or air, whereas only 25% of the SVs from pig farmers were shared with cow farmers and/or non-exposed (Fig. 2, A and B). This illustrates that more SVs are shared within the same environment (pig farms) of the different sample types (pigs, air and pig farmers) than within the same sample type (humans) of the different environments (pig farms, cow farms and offices).

Within-farm as compared to between-farm dissimilarity is reduced

In order to investigate if the microbiota in pig farm samples is influenced by the farm (location ID), we compared pairwise distances between samples originated from the same farm (within farm) and between samples originating from different farms (between farms) (Fig 3, A and B). All “within” distances were significantly lower than the “between” distances (Kruskal-Wallis rank sum tests with BH correction (15); all $P < 0.001$), strongly indicating the existence of an effect of farm location on the microbiota. This was true within (Fig. 3, left side of dotted line) but also between (Fig. 3, right side of dotted line) different
sample types. However, as expected, the values for the within-dissimilarities for a given host (Fig. 3; pigs vs pigs and pig farmers vs pig farmers) were generally smaller than values observed between sample types.

**Identification of SVs significantly associated to pig farming (core pig farm microbiota)**

Performing an omnibus test (PERMANOVA) with all factors and all samples (n=255) revealed overall significant factor effects on community variation ($P = 0.001$, with and without stratifying for farm ID). Thus we next analyzed the SVs which were associated with the changes performing three different analyses.

First, SVs that where significantly associated with samples from pig farms were identified by screening all SVs for significantly higher abundance in pigs, air and pig farmers as compared to cow farmers, by applying pairwise Mann-Whitney-Wilcoxon Tests followed by BH correction for multiple testing. A total of 82 SVs were identified with significantly higher abundance in samples from pig farms compared to cow farmers and this low abundance was also present in non-exposed (Fig. 4, A). Second, we conducted a similar approach using frequency (presence-absence) data as input, using Fisher’s exact test with BH correction. Eighty-one SVs were identified in both approaches and one SV (SV125) was identified only by the approach based on relative abundances (Supplementary table S2). Finally we performed an ANOVA-Like Differential Expression (ALDEx) analysis for the analysis of the proportional data (16, 17). Effect size plots showing the within and between differences of SVs between the respective groups are shown in the Supplementary Fig. S2 A-C. Overall 41 SVs (50%) were significant for all three analyses and 9SVs were newly identified with ALDEx (Supplementary Fig. S3 and Supplementary table S2).

**Differences and similarities of the microbiota between anterior and posterior nasal samples**

After having identified large microbiota differences in the anterior nasal cavities associated with pig farming, we subsequently analyzed if there were also associations with the posterior part of the nose. For this, we again first performed an omnibus test (PERMANOVA, nested per individual) with all SVs from pig farmer samples (n=86) which showed an overall significance ($P = 0.001$) between anterior and posterior in each individual. We next analyzed all the 82 SVs that were identified as being specific for pig farming.
farming. In total, 9 out of 82 SVs were significantly more abundant in the posterior than in the anterior part of the nose, and included SVs from the bacterial families of Prevotellaceae and Veillonellaceae (Wilcoxon signed rank tests with BH correction (15) (Fig. 4B; \( P < 0.05 \)). We then analyzed the ten most abundant SVs (Supplementary Fig. S4), and mainly SVs from Corynebacteriaceae and Staphylococcaceae were more frequently found in the anterior as compared to the posterior part of the nose (Supplementary Fig. S4).

**Analysis of sequencing data using the mothur pipeline**

Finally, we compared our findings from the DADA2 with the mothur pipeline. As for mothur, the final mean number of reads per sample was 34232 (95% Confidence Interval: ±2117) ranging from 3340 to 109182 reads and the sequences were clustered into a total of 31951 OTUs (Operational taxonomic units). After rarefying, 10553 OTUs were left with 3340 reads per sample. These OTUs clustered into 41 phyla and 310 families respectively. The taxonomic profiles were very similar to the profiles obtained with DADA2 (Supplementary Fig. S5, A-D), except for a slightly higher abundance of “Others” for the samples analysed using mothur. We also noted a very high correlation between DADA2 and mothur in case of alpha- and beta-diversity. Richness (\( R^2 = 0.68 \)) and SDI (\( R^2 = 0.92 \)) showed strong positive linear relationships between values based on DADA2 and mothur (Figure 5, A and B). The Procrustes analysis comparing beta-diversity values from these two pipelines (Figure 5, C-F) also showed a strong correspondence between these two datasets for both Jaccard and Ružička dissimilarity (Procrustes symmetric correlation: Jaccard: 0.95, \( P = 0.001 \); Ružička: 0.91, \( P = 0.001 \)). The number of procrustes residuals were evenly distributed between the investigated sample types (Figure 5, D and F).
DISCUSSION

This cross-sectional study investigated the relationships between pig farming and the composition of the nasal microbiota of farmers. We revealed an increased bacterial richness and diversity in the anterior nose of pig farmers as compared to cow farmers and non-exposed control group. In addition, beta-diversity analyses revealed significant differences in the composition of the nasal microbiota of these human groups. Samples from within the pig farms shared more of their microbiota as compared to the samples from between farms. We were also able to identify the SVs that were significantly associated with pig farming and the SVs which were predominantly more abundant in posterior than in anterior nasal cavities of the pig farmers.

The found differences in alpha-diversity suggest that farmers raising pigs have an increased bacterial diversity in their nose as compared to non-exposed individuals and to farmers working on a cow farm. A possible explanation is that the high concentration of diverse aerosolized bacteria present in pig barns leads to a modification and an enrichment of the “natural” farmer’s nasal microbiota. Therefore, it appears that the establishment of this modified microbial community could be a “finger print” of the nasal microbiota of pig farmers. As for changes in community structure (beta diversity), we revealed that samples from pigs, air and pig farmers form distinct, yet related clusters, which are all clearly separated from samples from cow farmers and non-exposed office workers. It becomes obvious from our data that pig farming is associated with stronger divergence of the human nasal microbiota as compared to cow farming. These findings could be explained by the fact that pig farmers spend more time in a confined environment with the animals than cow farmers and by the fact that airborne dust concentration are higher in pig than in cow farms (18). It has been shown that pets can share a small part of their microbiota with their owners by hypothesized, frequent direct contacts (11, 19). However, our study data strikingly points out that airborne microbiota may indeed play an important role in this microbial transfer. Moreover, we also show that the extent of microbiota sharing between pigs and farmers is remarkable.

We additionally found that samples from the farmers working on the same farm shared more of their microbiota than they do with individuals from different farms. This was true not only for pig farmers, but also when comparing air samples and pigs from the same farm, hinting at the existence of an even more...
pronounced farm-specific microbiome. Similarly, it has been shown that household members shared more of their microbiota than they do with individuals from different households (12, 19). In our study, the degree of shared microbiota was again large and the type of farm management practices could be influential. Indeed, it has been already shown that farm management (diet and antimicrobial use) influenced the nasal microbiota of pigs (20). Therefore, we can suppose that the management and the farm characteristics can also have an influence on the air quality of the barn. Humans inhale 10,000L of air per day and airborne bacteria may have a direct effect on the nasal bacterial communities of humans (21). Furthermore, it is known that the air in farms, and in particular pig farms, is charged with high amounts of dust, bacteria and fungi (as shown in this and other studies (22, 23)) and that the concentration of airborne bacteria can be $2 \times 10^7$ times higher than the level usually measured in indoor air (24).

Our results strikingly revealed a very high number of SVs shared between the pigs, air and the pig farmers, indicating a frequent exchange of members of the microbiota and suggesting that air could play an important role in the transmission of animal-associated bacteria to the farmers, too. Among these SVs, Veillonellaceae and Lactobacillaceae were the most abundant groups in pigs, air, and pig-farmers. Lactobacillaceae and Veillonellaceae have been found in the nares of both healthy pigs and humans (20, 25-27). We also simultaneously sampled the posterior and anterior nasal cavities of the human participants. This is important as spatial variation in nasal microbial communities has been highlighted (28), although in another study, the bacterial composition did not significantly change along the nasal passage (29). In addition, the microbiota of the posterior cavity should reflect a more persistent (vs transient) colonization than that of anterior cavity. Our data shows, that the microbiota between anterior and posterior differ and that there are some SVs which are associated with either of the two sites.

The farmer’s respiratory tract also receives a lot of attention due to the hygiene hypothesis demonstrating that growing up on a farm may be protective against allergies and asthma (13). This has been shown to be very significant in the case of pig farming (30). Therefore, SVs associated with pig farming identified in this study which were found in higher abundance in the posterior region of the nose, could hypothetically be protective against asthma development. Indeed, many SVs found in our study
have been associated with respiratory health rather than disease, like asthma (31-34). Therefore, these
SVs could have potential protective implication for allergic and atopic diseases. However, as we only
included healthy adult subjects, differences in the nasal microbiota that were associated to certain
occupational health problems and/or health benefits (e.g. atopic and allergic diseases) were not
investigated.

Within this study, we decided to use the DADA2 algorithm rather than the more known 97% Operational
Taxonomic Units (OTU) approach. The DADA2 algorithm has been shown to produce a higher resolution
of microbial populations when applied to 16S rRNA gene sequences as compared to the popular
clustering into OTU as implemented in *mothur* or *QIIME* pipelines (35). The resulting SVs only contain
one read per SV, making additional analysis steps, such as oligotyping, unnecessary (36). Even though
DADA2 leads to a decrease in alpha-diversity, it does not lead to changes in the community structure,
which makes the approach comparable to results produced by other clustering algorithms (35, 37). By
comparing DADA2 with mothur in our study, we can clearly confirm the later statement as shown in our
procrustes analyses.

This study has some major strengths: By taking into account all potential confounding factors (season,
age, geographical region); we reliably demonstrated, that pig farming has an extensive effect on the
human nasal microbiota and we were able to reveal the specific SVs which were associated to these
changes. Moreover, recruiting cow farmers as control group allowed ascertaining that the observed
differences are linked to the close contact to pigs and not simply by the lifestyle associated with living on
a farm. By including multiple samples from identical farms, we additionally were able to reveal the
existence of a pronounced farm-specific microbiome by observing more similarity between the
microbiota within a same farm than between the different farms. Finally we also included microbiota
analysis of posterior nasal samples, and bacteria from this region of the nose are more likely to be
relevant for the respiratory tract microbiota and community disturbance which could lead to potential
dysbiosis (34, 38).

There are limitations to this study, too. We only included healthy adult subjects. We were thus not able to
investigate differences in the nasal microbiota that were associated to certain occupational health
problems (e.g. atopic and allergic diseases). Therefore the relevance of the distinct microbiota needs to be studied in the future with different experimental designs. In addition, we did not perform longitudinal sampling, and, therefore, were not able to investigate the temporal stability of the different microbiota. Finally, as this was a ‘field study’, we did not perform some additional upper or even lower respiratory tract sampling. This would more clearly reveal the composition of the respiratory tract microbiota as shown before (32, 33).

In conclusion, we have identified that pig farming has an extensive effect on the human nasal microbiota and we were able to reveal the specific SVs associated to these changes. The relevance and stability of these changes need to be investigated in the future.

Materials and Methods

Study design and sampling

Ethical clearance for this study was sought and obtained from Human Research Ethics Committee of the Canton Vaud (243/14 and P_2017-00265) and the Veterinary Ethics Committee of the Canton Vaud (VD2903). Sample collection was conducted between October 2014 and March 2015 in the western part of Switzerland. We focused on the winter season as hypothesized that, the doors etc. may be more likely to be closed and, therefore, the pig farmers are more exposed to the indoor bacterial communities. Related to this, it has been described that there is a decrease in some of the air contaminants during summer of swine confinement buildings. In total, 28 pig farms were visited and nasal swabs from suckling or weaning pigs were obtained by swabbing their noses using sterile cotton swabs. The piglets rather than pigs were chosen for ease of handling and sampling. The pig farmers collected two swabs, outside the pig barn, from their left nares (anterior and posterior) themselves under supervision of the study personnel. In addition, personal information was collected in a questionnaire. Airborne bacteria were sampled with a Coriolis μ air sampler (Bertin Technologies, Montigny-le-Bretonneux, France), positioned approximately one meter above ground in the middle of the pig house and the airborne particles of a total of 3 m³ air (0.3 m³/min for 10 minutes) were collected into a sterile cone containing 15 ml 0.005% Triton X-100 solution. As controls, 17 cow farmers and 26 non-farming individuals, having no
contact with any type of farm animal, were included. All samples were immediately transported to the laboratory in a cold box (4 °C), and stored at -20 °C until further analysis. DNA extraction, amplification and sequencing were done as outlined in the supplementary material. In brief, the V4 region of the 16S rRNA gene was amplified using forward (5’-GTGCCAGCMGCCGCGGTAA-3’) and reverse (5’-GGACTACHVGGGTWTCTAAT-3’) primers previously described (39) and modified with an Illumina adaptor sequence at the 5’ end. Samples were submitted to the Next Generation Sequencing Platform for indexing and pair-end 2x250 bp sequencing (Reagent Kit v2) on the Illumina MiSeq platform (San Diego, USA). The reads were deposited at the National Center for Biotechnology Information Sequence Read Archive (accession no. PRJEB21578). Reads were analyzed using the DADA2 package version 1.5.0 and workflow (35) in R version 3.1.2 (http://www.R-project.org) as illustrated in the supplementary material. The output of DADA2 consist of exact SVs which replace the traditional OTUs received by more ‘traditional’ pipelines like Mothur. Using DADA2, no rarefying of sequence reads was necessary.

Alpha-, beta-diversity analyses and identification of SVs associated with pig farming

If not otherwise stated, all calculations were performed in R utilizing functions from R base or the vegan package. We did not rarefy our sequences for downstream analyses as the DADA2 algorithm drastically reduces the issues of having different sequencing depths for the samples being compared, which is the main reason for rarefying. Alpha-diversity (within-sample diversity) was assessed by calculating richness and Shannon Diversity Indices (SDIs), using the functions estimateR and diversity. Linear regression models with a random effect to correct for clustering on the location level was used to test for statistical significances between sample types (lmer function from the lmeTest package) and overall significance of these models was confirmed with analyses of variance (anova function).

Beta-diversity (between-sample diversity) was measured by the weighted Ružička index (abundance-based) and the unweighted Jaccard index (presence/absence-based) of dissimilarity. Ružička is also called the quantitative version of Jaccard and unlike Bray-Curtis which is semimetric, is metric and probably should be preferred (http://cc.oulu.fi/~jarioksa/softhelp/vegan/html/vegdist.html).

Pairwise distances between samples were calculated using the vegdist function and the resulting matrices were used to generate non-metric multidimensional scaling (NMDS) plots (metaMDS function).
and dissimilarity boxplots. Significant groupings between samples were assessed by a permutational multivariate analysis of variance using 1000 Monte Carlo permutation tests (PERMANOVA; adonis function). Analyses of similarities were performed to test for significant differences between groups of samples using 1000 Monte Carlo permutation tests (ANOSIM; anosim function), followed by Bonferroni correction for multiple testing. Both PERMANOVA and ANOSIM were performed as hierarchical models with nesting at the farm level to address the fact that several samples originated from the same farm.

The extent of beta-diversity dispersion for each sample group was determined by calculated as the average distance (based on Jaccard and Ružička index) to the sample type’s centroid using the betadisper function (40), and significant differences were assessed with Tukey’s Honest Significant Difference Test (TukeyHSD function). Significant differences between the groups in the dissimilarity boxplots were assessed by Kruskal-Wallis rank sum tests with Benjamini-Hochberg (BH) correction for multiple testing (15). Boxplots and NMDS plots were generated in R utilizing the ggplot2 package and Venn diagrams were created with help of the VennDiagram package.

The identification of SVs associated with pig farming and of SVs associated with either the anterior or posterior nasal cavities is described in the supplementary file. This includes the ANOVA-Like Differential Expression (ALDEx) Analysis in R to analyses proportional data using the aldex2 package as described (17).

**Comparison of the pipelines DADA2 and mothur**

We also compared the findings from the DADA2 with the mothur pipeline as illustrated in the supplementary material. In brief, reads of all samples were additionally analyzed using the mothur software (version 1.36.1) (41) as indicated in the MiSeq standard operating procedure (42). Unlike with DADA2, the data was normalized by random subsampling of sequences resulting in 3340 reads per sample. Beta-diversity comparison was accomplished by using Procrustes transformations with non-metric multidimensional scaling (NMDS) ordinations (based on Jaccard and Ružička indeces of dissimilarity) as input. The plots were obtained by using the protest function and the significance between the two configurations was confirmed with the protest function.
Funding: This work was supported by Swiss National Science Foundation (SNF) grant 310030_152880 to MH and AO.

Acknowledgements

We are grateful to all the participants in this study and we thank Pascal Wild for his help with statistical analyses.
References


Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH,
source, platform-independent, community-supported software for describing and comparing microbial

sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina
### Table 1: Characteristics of the study population

<table>
<thead>
<tr>
<th>location number</th>
<th>No. of sampled individuals</th>
<th>No. of sampled male individuals</th>
<th>No. of sampled smokers</th>
<th>Mean age of sampled individuals(^a) (s.d.)</th>
<th>No. of sampled pigs</th>
<th>Total No. of pigs on farm</th>
<th>No. of air samples (pig barn)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pig farm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>34 (11)</td>
<td>3</td>
<td>920</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>65</td>
<td>2</td>
<td>80</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>40 (1)</td>
<td>2</td>
<td>350</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>33</td>
<td>2</td>
<td>435</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>71</td>
<td>2</td>
<td>240</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>45</td>
<td>2</td>
<td>80</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>54</td>
<td>2</td>
<td>90</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>50</td>
<td>2</td>
<td>120</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>48</td>
<td>2</td>
<td>590</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>32</td>
<td>2</td>
<td>520</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>57</td>
<td>2</td>
<td>90</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>32 (5)</td>
<td>2</td>
<td>950</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>30 (14)</td>
<td>2</td>
<td>600</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>49</td>
<td>2</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>63</td>
<td>2</td>
<td>520</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>42</td>
<td>2</td>
<td>280</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>53 (1)</td>
<td>2</td>
<td>290</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>62</td>
<td>2</td>
<td>320</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>44 (30)</td>
<td>2</td>
<td>250</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>53</td>
<td>1</td>
<td>130</td>
<td>1</td>
</tr>
<tr>
<td>21</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>48 (11)</td>
<td>2</td>
<td>750</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>53 (4)</td>
<td>2</td>
<td>250</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>39 (14)</td>
<td>2</td>
<td>1950</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>58 (4)</td>
<td>2</td>
<td>2700</td>
<td>1</td>
</tr>
<tr>
<td>25</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>49</td>
<td>2</td>
<td>205</td>
<td>1</td>
</tr>
<tr>
<td>26</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>46 (8)</td>
<td>2</td>
<td>270</td>
<td>1</td>
</tr>
<tr>
<td>27</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>62</td>
<td>2</td>
<td>105</td>
<td>1</td>
</tr>
<tr>
<td>28</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>50</td>
<td>2</td>
<td>350</td>
<td>1</td>
</tr>
<tr>
<td>cow farm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>37 (14)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>9</td>
<td>8</td>
<td>0</td>
<td>38 (14)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>60 (0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-exposed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-65</td>
<td>26 (one per workplace)</td>
<td>24</td>
<td>9</td>
<td>39 (14)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\)Mean age and s.d. (standard deviation are shown if more than one individual was sampled
FIGURE LEGENDS

Figure 1. Alpha and Beta-diversity analyses of samples of pigs, air, pig farmers, cow farmers and non-animal exposed individuals. As for the alpha diversity, illustrated are A) the differences of richness (observed sequence variants (SVs)) and B) Shannon diversity indices based on sample types. As for the Beta diversity, illustrated are C) Unweighted (Jaccard) and D) Weighted (Ružička) distances in microbiota composition, reduced in a 2D-space by using NMDS, 95% confidence ellipse for the group centroid shown. In addition, shown are the beta-dispersion based on E) Ružička and F) Jaccard dissimilarity indices in each sample type. The boxplots represent median (midline), interquartile ranges (shaded boxes), and ranges (whiskers). Different colours are indicated: orange (pig), blue (air), red (pig farmer), green (cow farmer) and purple (non-exposed). Significant differences within A), B), E) and F) are displayed with either * (p < 0.05), ** (p < 0.01) or *** (p < 0.001).

Figure 2: Venn diagram showing unique and shared Sequence variants (SVs). Illustrated is A) the Venn Diagram showing the number of shared SVs between pig farmers, pigs and air and B) the Venn Diagram showing the number of shared SVs between pig farmers, cow farmers and non-exposed. Shared SVs were determined by identifying the total number of shared SVs between pig farmer, pig and air samples and between pig farmer, cow farmer and non-exposed samples. Pig farmers share more SVs with pigs and air than with cow farmers and non-exposed.

Figure 3. Within and between pig farms dissimilarities measurements. Shown are A) the Unweighted (Jaccard) and B) Weighted (Ružička) distances in microbiota composition within farms (pairwise distances between sample types originated from the same farm) and in between farms- (pairwise distances between samples originating from different farms) dissimilarities. The boxplot represents median (midline), interquartile ranges (shaded boxes), and ranges (whiskers). Significant differences are displayed with either * (p < 0.05), ** (p < 0.01) or *** (p < 0.001). The dotted line separates the comparisons within and between sample types.
Figure 4. Sequence variants (SVs) associated with pig farming and differential SVs between anterior and posterior nasal samples. Illustrated are the 82 SVs which were significantly associated for pig farming (see text for details). Shown is A) a phylogenetic tree based on differences in their sequence reads (distance displayed at substitutions per site) and heat maps depicting relative abundances and frequencies for pig (n=56), air (n=27), pig farmer (n=56), cow farmer (n=17) and non-exposed (n=26). Assigned taxonomy (bacterial genus, order or family) for each SV is shown, too. The B) Forest plot displays the coefficients of pairwise differences between anterior and posterior nasal samples from pig farmers derived by Wilcoxon signed rank tests followed by Benjamini-Hochberg correction. Significant differences after multiple testing are illustrated (*).

Figure 5: Alpha- and beta-diversity comparison calculated with DADA2 and mothur. Shown are A) the correlation analysis of richness values based on DADA2 and mothur, B) the correlation analysis of SDI values based on DADA2 and mothur, C) the Procrustes analysis of Jaccard dissimilarity calculated with DADA2 and mothur. Significant (P=0.001) correlation value indicated in figure, D) the bar chart of Procrustes residuals based on Jaccard dissimilarity calculated with DADA2 and mothur, E) the procrustes analysis of Ružička dissimilarity calculated with DADA2 and mothur. Significant (P=0.001) correlation value indicated in figure and F) the bar chart of Procrustes residuals based on Ružička dissimilarity calculated with DADA2 and mothur.
Figure 2

A

**pig farmer**

- 2384 (56% of total SVs)
- 291
- 1183 (29% of total SVs)
- 715
- 376
- 165

**pig**

- 708 (30% of total SVs)

**air**

B

**pig farmer**

- 3161 (75% of total SVs)
- 578
- 330
- 165
- 116
- 708 (54% of total SVs)

**cow farmer**

- 1892 (65% of total SVs)

**non-exposed**