

Group 3:

Alexander Kruhøffer Bloch Andersen (s135271), Wanli He (s195995),
Huayi Huang (s181285) & Andreea-Aiza Watzlawek (s172359)

INTRODUCTION

Co-habitation, an adaptive state of existence in which two or more people live in the same house, is widespread in human society. People are willing to choose this kind of lifestyle simply because it can bring a multitude of benefits. However, while considering the benefits, this kind of lifestyle could also have an influence on the bacterial communities comprising the household biofilms, but much is still unknown¹⁻⁴.

In this study, we are trying to investigate how the number of co-habitants influence the microbiome in the existing domestic biofilms. The effects on the microbial and gene diversity will be assessed by using metagenomics sequencing and analysis, as these currently represent the most appropriate and fastest research methods.

WORKFLOW

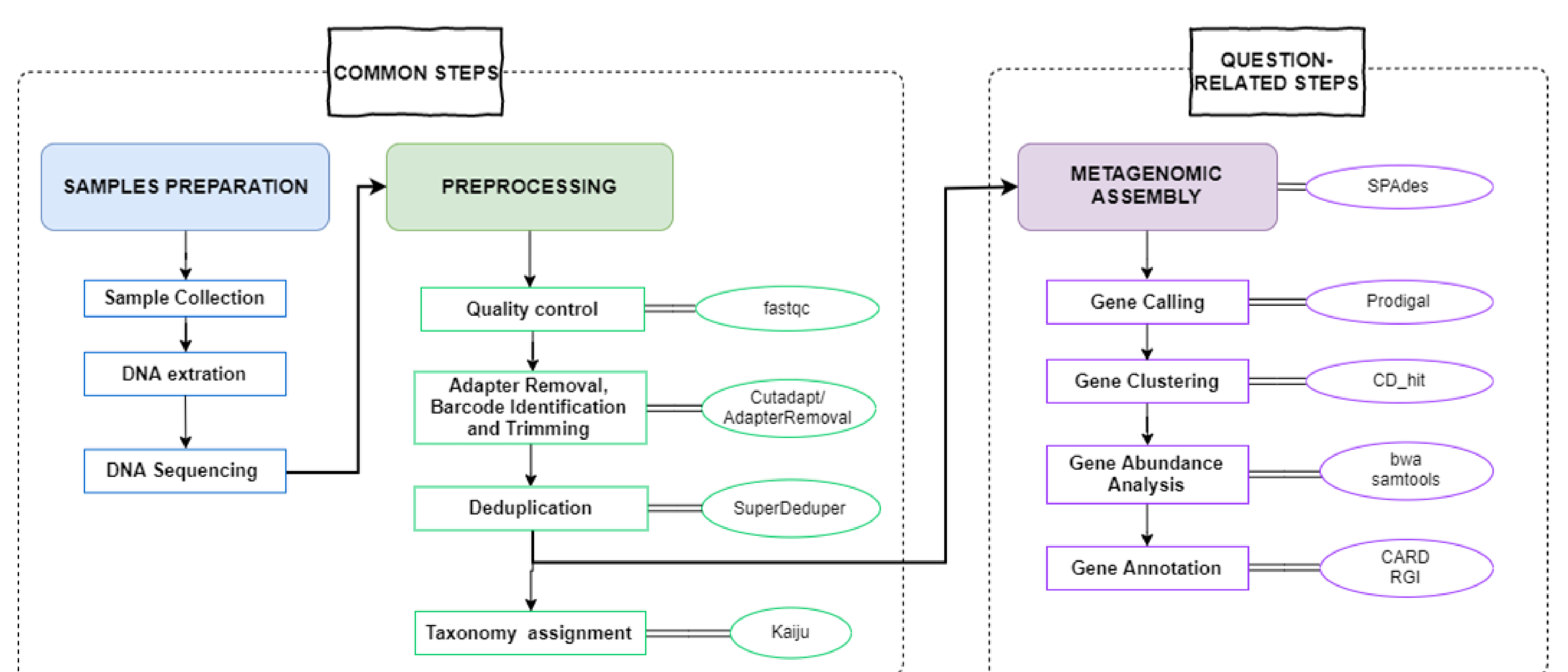


Fig. 1. Workflow of the project. The steps followed throughout the project are placed in two categories: the common steps followed in class followed by the steps required to reach the aim of this project. Samples preparation was carried out in laboratory with NEXTflex DNA-seq kit. DNA Sequencing was performed using an Illumina sequencer. The preprocessing and downstream analysis were performed using UNIX command lines coupled with R software.

RESULTS

Data Description

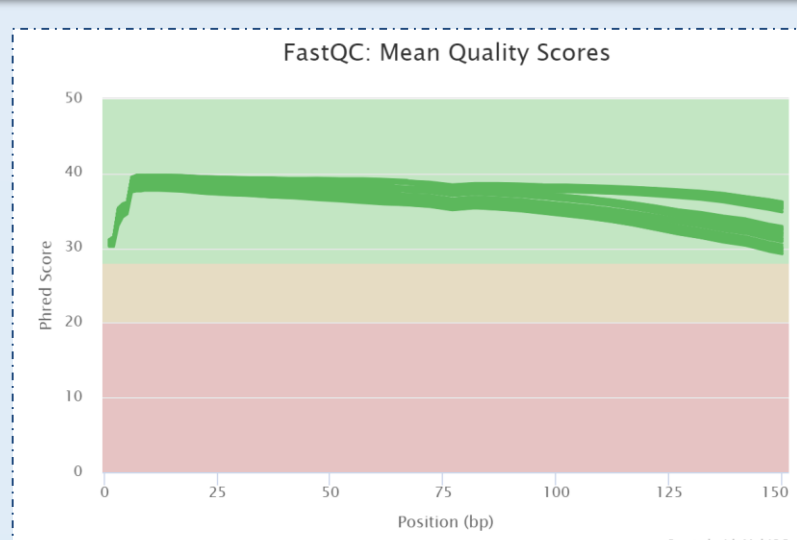


Fig. 2. Quality control of the reads, before and after pre-processing steps. Quality control was performed by fastqc and visualized by MultiQC. Per base sequence quality plot indicates high reads quality after trimming.

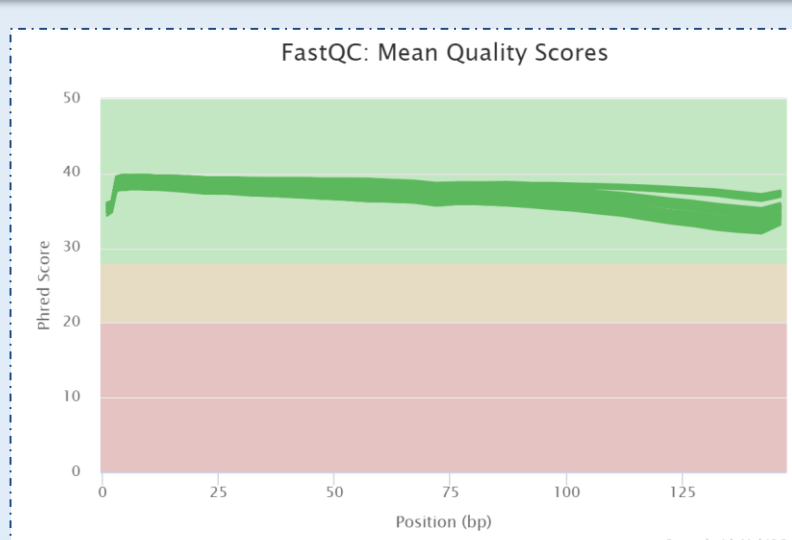


Fig. 3. Estimated average coverage. Nonpareil-analysis was performed to assess the sequence depth. Most of the deduplicated samples have an average coverage close to 100%.

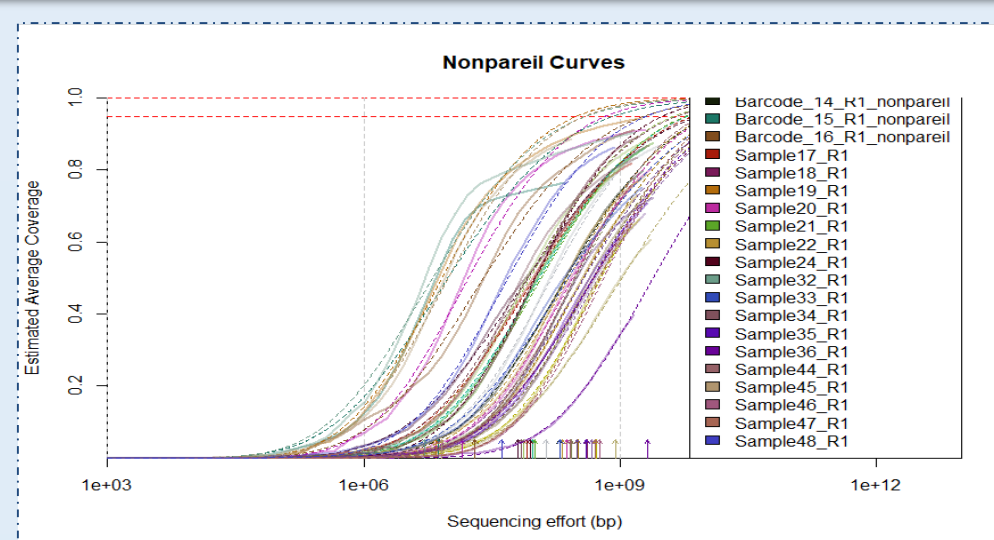


Fig. 4. Assessing the assembly quality by performing the scaffold length distribution. The histogram represents the number of scaffolds per scaffold length, ranging from 0-3000 with a general distribution.

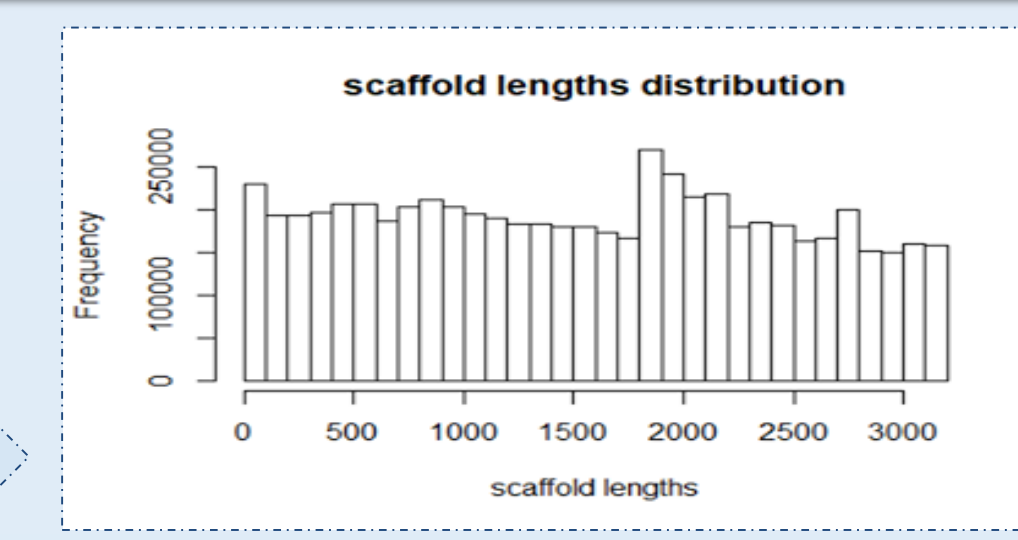
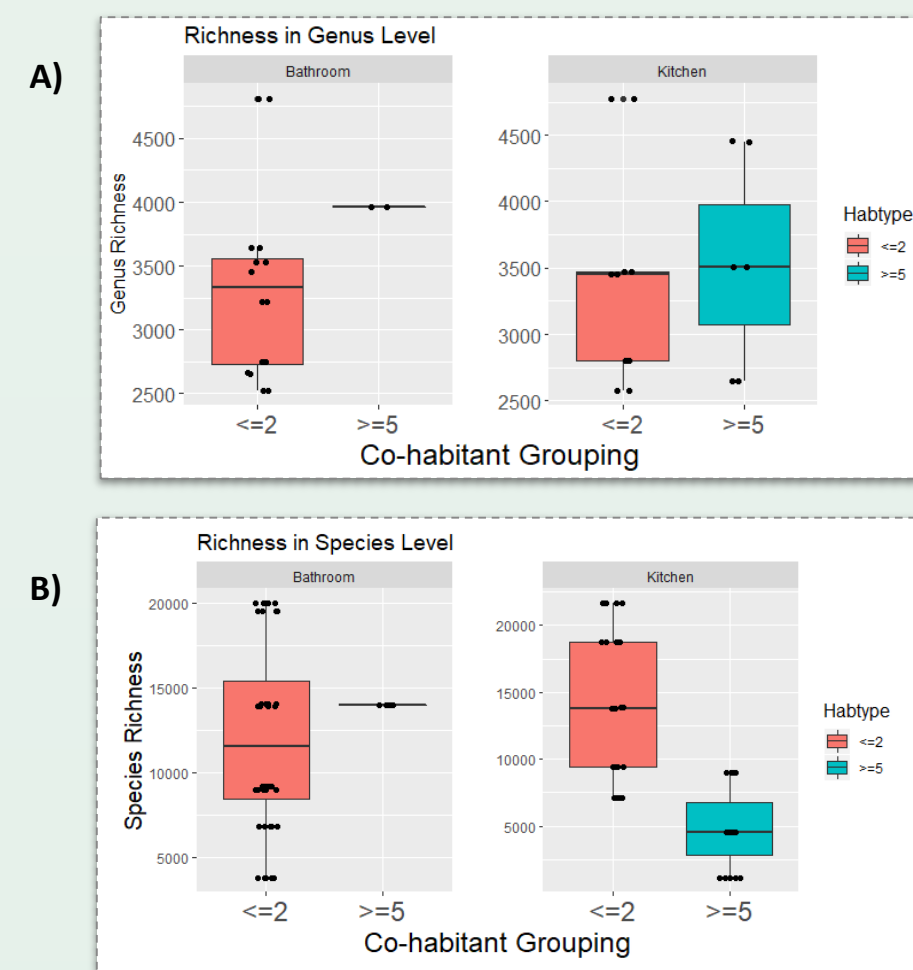


Fig. 5. Sample distribution. Samples were grouped based on the number of inhabitants and of the site of sampling.

Richness



P-value(pairwise t-test, p = 0.95):

	Bathroom	Kitchen
Genus	8.488e-06	0.0001356
Species	0.0002631	0.0003342

Fig. 6. Boxplot showing the comparison of richness between different co-habitant groups in species level and genus level, respectively.

T-test was performed to determine the significant difference level between these two habitat groups. The P-value depict a significant difference between these groups.

Beta-Diversity

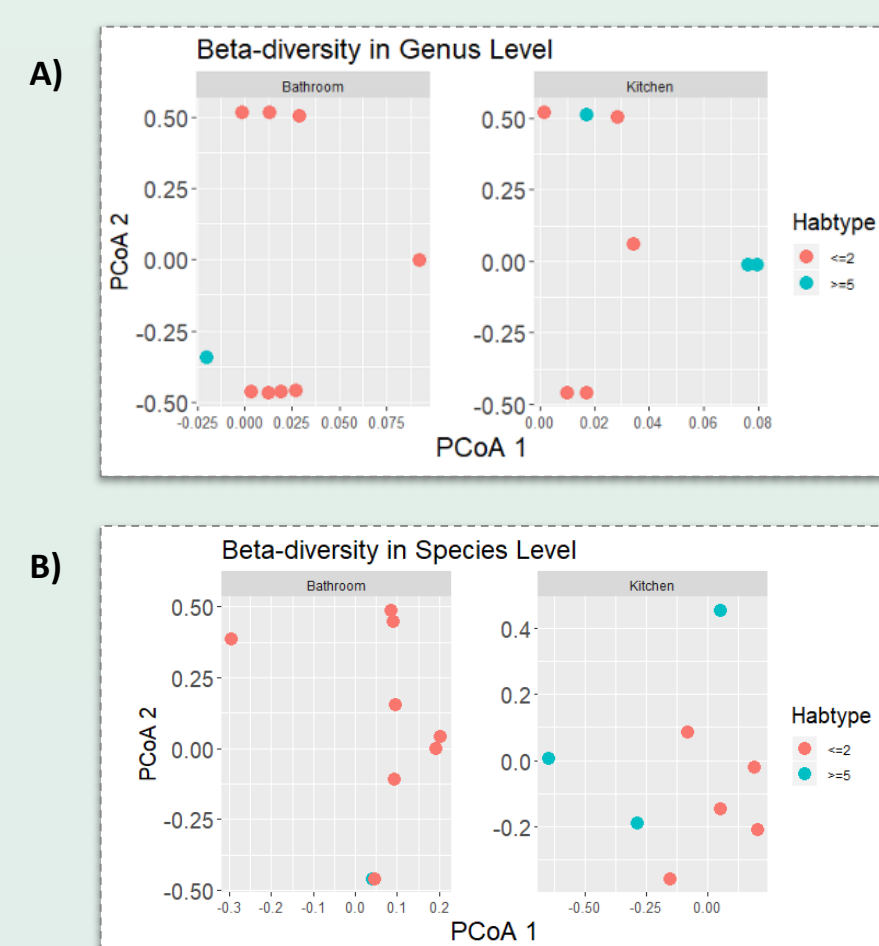


Fig. 7. Principal coordinate Analysis (PCoA) scatter plot for Beta-diversity analysis. The right hand plots represent different sample sites groups different habitat types group.

Beta-diversity was calculated by Bray-Curtis dissimilarity and illustrated on PCoA scatter plots. However, there is no obvious separation between these groups either for PCoA1 or PCoA2 either within kitchen samples or bathroom samples.

Compositional analysis

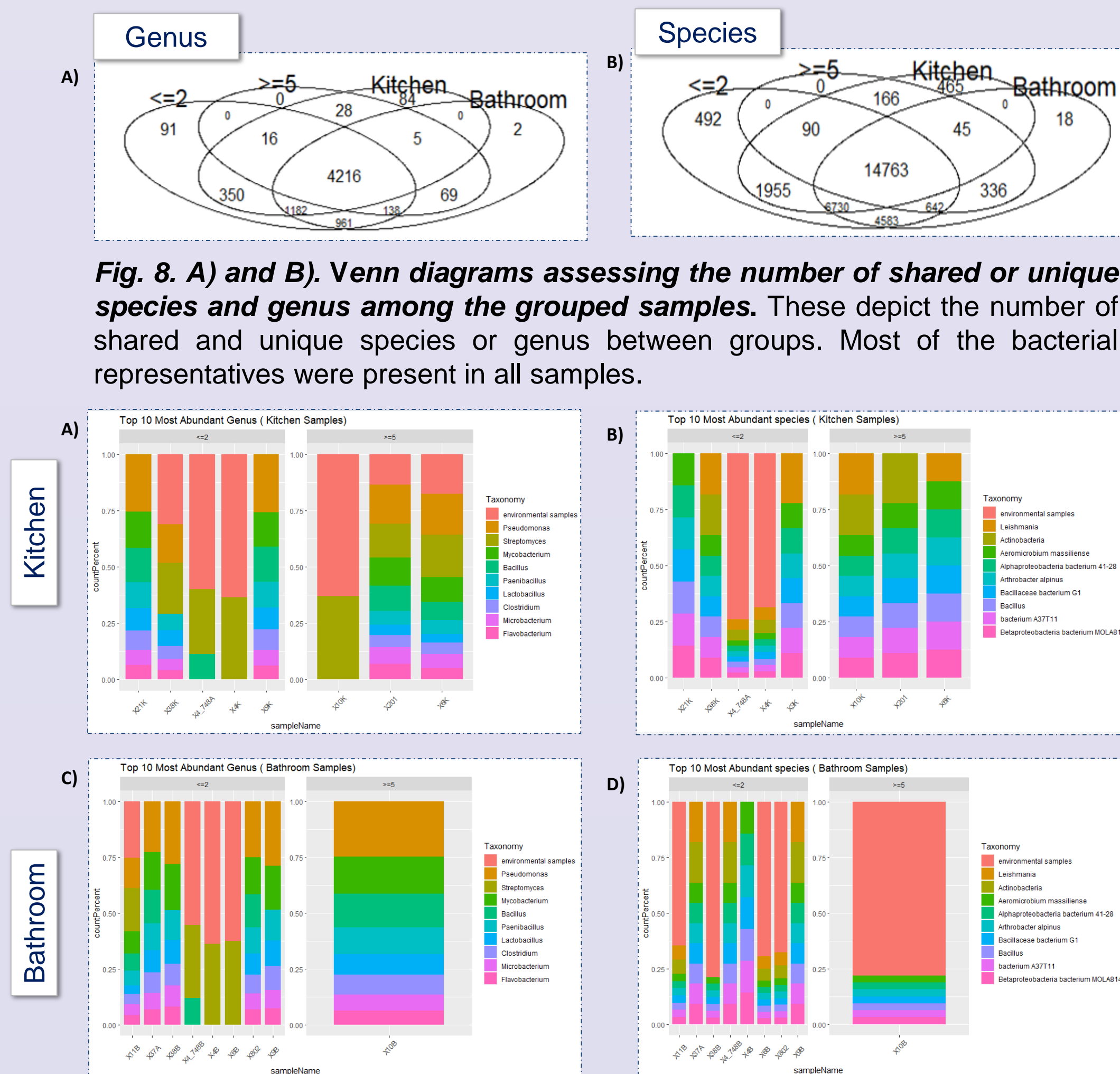


Fig. 8. A) and B). Venn diagrams assessing the number of shared or unique species and genus among the grouped samples. These depict the number of shared and unique species or genus between groups. Most of the bacterial representatives were present in all samples.

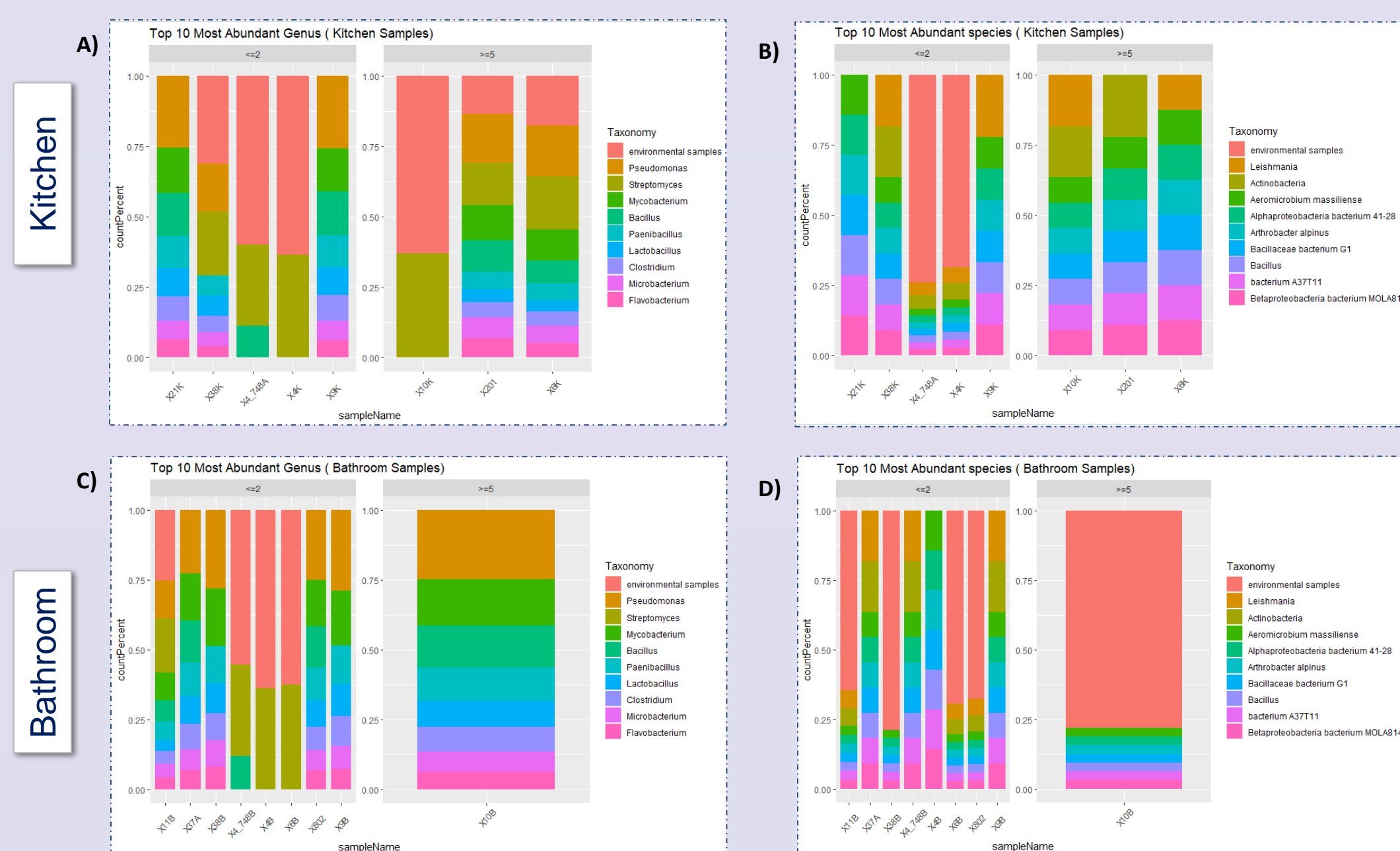


Fig. 9. Distribution of top 10 abundant genus and species. The figures show the relative abundance of the top 10 most abundant genus and species encountered. All the samples have been normalized without being re-sampled in order to avoid data loss.

Resistance genes analysis

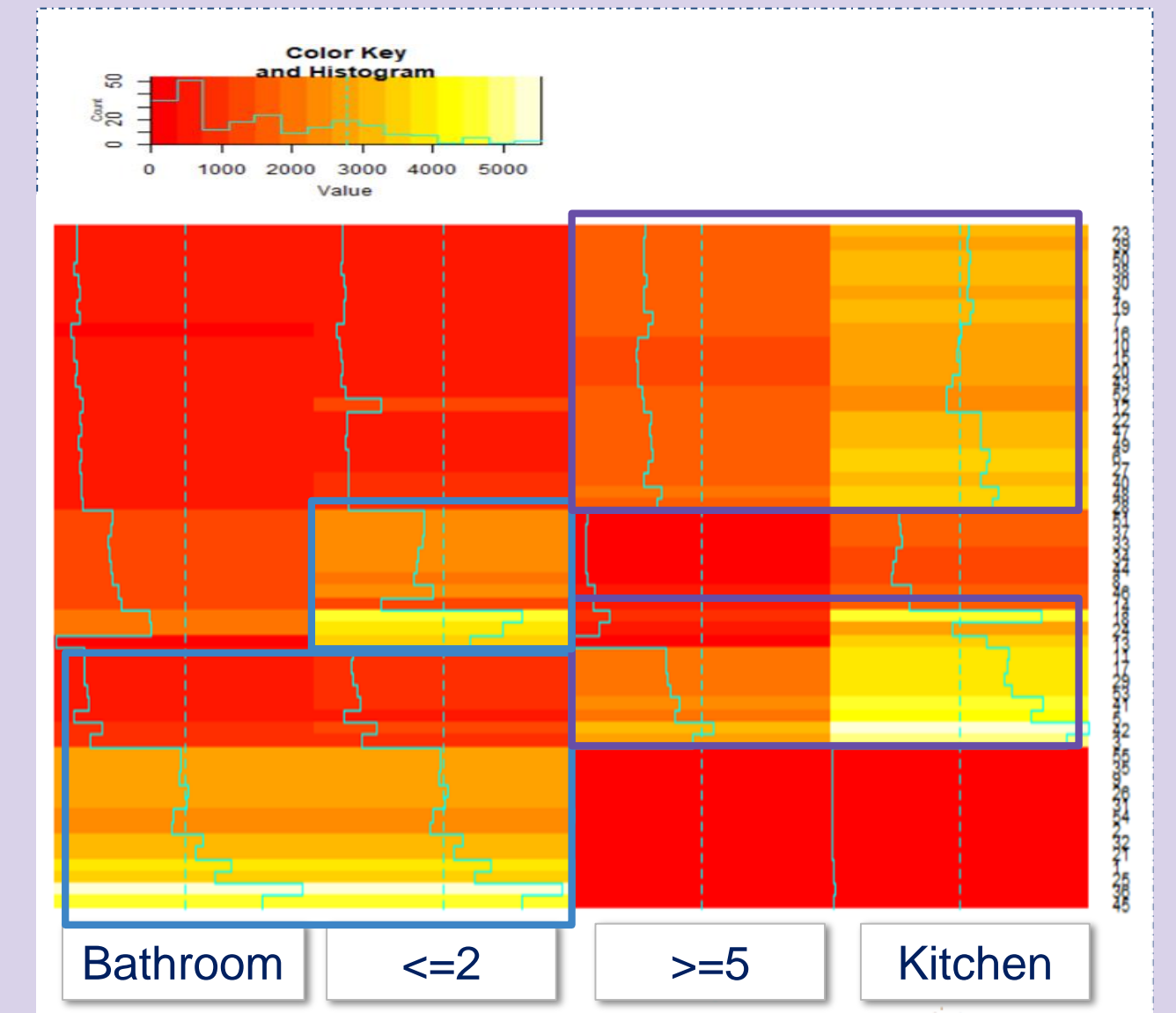


Fig. 10. Heatmap of the 55 most abundant resistance genes present in the samples grouped according to their sampling site and number of co-habitants.

There is a high correlation between the bottom 10-resistance genes in "Bathroom" and "<=2" groups, but a question arises: if there is a lot of the samples that are present in both groups.

There also seems to be a slight pattern between the abundance of the resistance genes seen in both the "Kitchen", as well as the ">=5" groups.

There seems to be a slightly higher abundance of certain resistance genes in the "<=2" group that is unique, in comparison to the other groups (the orange-yellow field in the middle of that column).

DISCUSSION

The aim of this study was to investigate the effect of household size on domestic biofilms. The more people living together, the higher the genus richness, which is correlated with a higher bacterial diversity. However, species richness of kitchen samples did not follow the same trend. Moreover, there were no significant dissimilarities between these groups according to beta-diversity and compositional analysis.

There is also an obvious difference in the distribution of the most abundant resistance genes in the groups having various numbers of co-habitants.

A relevant conclusion to the raised question could be reached if certain factors are considered. There are elements that hindered the analysis, such as the limited number of samples that fell in the considered categories, as well as partial meta-data.

PERSPECTIVES

A higher number of samples to collect, in order to make a more proper assessment.

More time to be assigned for the project in order to reach the desired details and bring relevant results and discussion.