

# Effect of cleaning on virulence factors in domestic biofilms

Group 5: Ilse Pardijs, Aileen Geers and Tor Noema Nørgaard

## Introduction

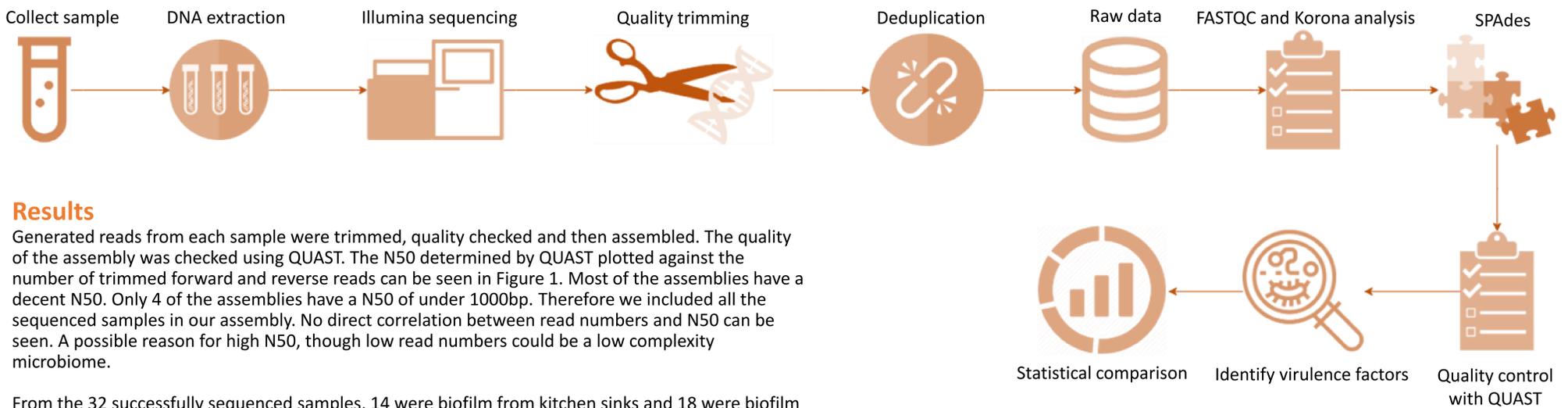
Many micro-organisms are able to form biofilms by adhering to a surface and each other.

Not only in natural environments this occurs, but also in our own homes, for example in dishwashers, shower drains and kitchen sinks<sup>1</sup>. In previous research, it has been found that infrequent cleaning correlates with higher bacterial diversity<sup>2</sup>. This gives reason to believe that this is also true for virulence genes.

Here, we investigate the abundance of bacterial species harboring pathogenic and virulent genes are prevalent in kitchen and bathroom biofilms, and the effect of cleaning on their abundance and diversity. Hypothesize the following:

**“The amount of pathogens and virulence genes in domestic biofilms negatively correlates with the amount of cleaning.”**

## Project Pipeline



## Results

Generated reads from each sample were trimmed, quality checked and then assembled. The quality of the assembly was checked using QUAST. The N50 determined by QUAST plotted against the number of trimmed forward and reverse reads can be seen in Figure 1. Most of the assemblies have a decent N50. Only 4 of the assemblies have a N50 of under 1000bp. Therefore we included all the sequenced samples in our assembly. No direct correlation between read numbers and N50 can be seen. A possible reason for high N50, though low read numbers could be a low complexity microbiome.

From the 32 successfully sequenced samples, 14 were biofilm from kitchen sinks and 18 were biofilm from the shower drain. We separated the samples into three categories based on the frequency of cleaning: less than 2 times per month, 2 times per month or more than 2 times per month. Additionally we recorded the time of the last cleaning in month.

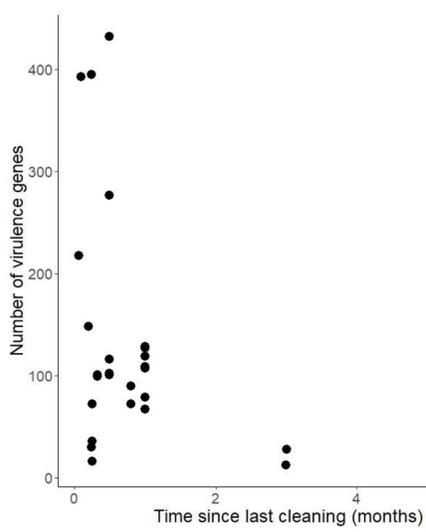


Figure 2: The amount of virulence genes based on the time since last cleaning the drain. N = 27.

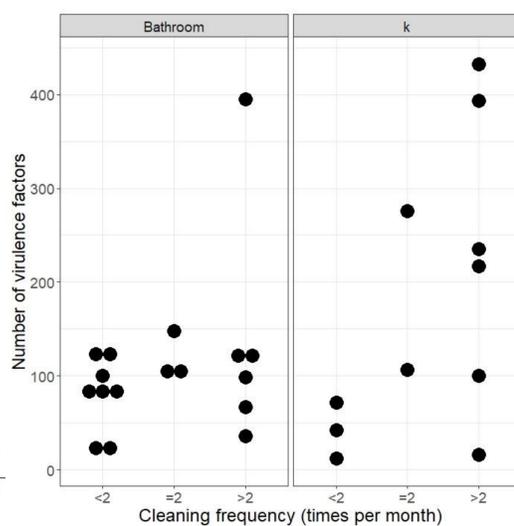


Figure 3: Impact of cleaning frequency on the amount of virulence genes found. Data is grouped into times pr. month. N = 28.

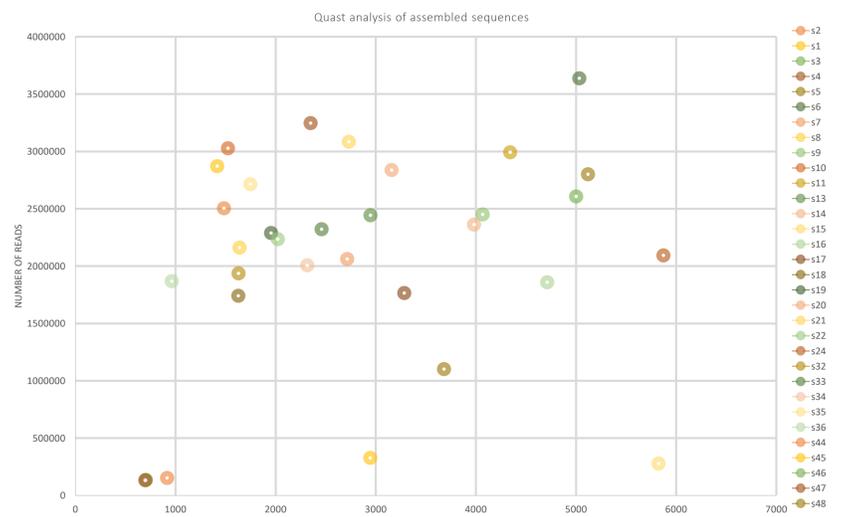


Figure 1: Diagram showing the N50 against the number of reads from each sample.

For further analysis we determined the number of virulence factors in each sample, by blasting the assembly against the pathogenic bacterial virulence factor data base<sup>3</sup> (VFDB) the core dataset. The hits were filtered by an e-value cut-off of  $10^{-10}$ , culling limit 1, percent identity of 75, and a length of over 500bp, which resulted to 4251 hit for all samples.

Unfortunately, there was no metadata available on the frequency of cleaning for 4 samples, and they were discarded. For the remaining 28 samples the number of virulence factors is depicted in figure 3. For the bathroom samples no effect of cleaning can be seen. For the kitchen samples there is a slight tendency of more detected virulence factors for frequently cleaned sinks. However, the tendency is only slight and not statistically significant. This was also seen when plotting the mean and standard deviation (not shown here.)

Additionally we investigated the number of virulence factors against the time since the last cleaning, which can be seen in figure 2. 27 samples are depicted where the corresponding metadata was available, and which were cleaned in the last 30 months. Like the frequency of cleaning, no effect of the time of the last cleaning on the number of virulence factors can be seen.

## Conclusion

Both the effect of frequency of cleaning and the last time cleaning the drain, on the number of virulence genes annotated in VFDB have been investigated. Multiple analysis were performed, but none showed significant correlations. This could be due to small sample size, inaccurate metadata, database bias or the absence of any actual correlation. As change in diversity of bacteria due to higher frequency of use has been shown in previous research, further research into cleaning effects could prove interesting. However, the conclusion of this research is:

**Based on the available samples and metadata, no correlation between cleaning and presence of virulence genes could be found.**

## References

- Hemdan, B. A., El-Liethy, M. A., ElMahdy, M. and EL-Taweel, G. E. Hemdan, B. et al. (2019) "Metagenomics analysis of bacterial structure communities within natural biofilm", *Heliyon*, 5(8), p. e02271. doi: 10.1016/j.heliyon.2019.e02271.
- Flores, G. E., Bates, S. T., Caporaso, J. G., Lauber, C. L., Leff, J. W., Knight, R. and Fierer, N. Flores, G. et al. (2012) "Diversity, distribution and sources of bacteria in residential kitchens", *Environmental Microbiology*, 15(2), pp. 588-596. doi: 10.1111/1462-2920.12036.
- Liu, B., Zheng, D., Jin, Q., Chen, L. and Yang, J. Liu, B. et al. (2018) "VFDB 2019: a comparative pathogenomic platform with an interactive web interface", *Nucleic Acids Research*, 47(D1), pp. D687-D692. doi: 10.1093/nar/gky1080.
- Kang, D. D., Froula, J., Egan, R. and Wang, Z. Kang, D. et al. (2015) "MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities", *PeerJ*, 3, p. e1165. doi: 10.7717/peerj.1165.

## Discussion

In this study we relied on VFDB database for identifying virulence genes. This means that our results are based on a singular source and could cause the data to have a bias. An easy way to address this bias would be to include additional databases such as Victors to be double validate the results from VFDB. Another concern would still be the reliance on existing data.

From the reads we have it is not possible to determine with greater accuracy the origins of individual virulence genes and link those to specific organism. To determine origin binning via Metabat. Using this method we would be able to tie specific genes to specific organism.

The collected metadata used in this study should be viewed with caution, since the reliability of the data is questionable. One example of this is the frequency of cleaning the drains. Based on the metadata, it can be seen that there is no pattern to cleaning frequency, also it is hard to separate usage from cleaning, since both actions interacts and removes biofilm.