# Introduction

# Materials and methods

## Sample collection and processing

For this study, we chose two different cooling towers located on the rooftop of two different hospitals in Vienna for sampling. The first hospital is located near the center of Vienna whereas the second hospital is located at Vienna’s periphery. The two locations are hereafter referred to as “Cooling-Tower 1” (CT-01) and “Cooling-Tower 2” (CT-02). After a *Legionella* outbreak in 2007, CT-01 was demolished and renewed. For both sites, the water temperature were recorded throughout the study, whereas pH-value were only obtained from CT-02.

During the sampling period between September 2013 and September 2014, water samples were retrieved biweekly. This study included water samples from two different cooling towers located on the rooftop of two hospitals in Vienna. 3 L water was taken from the basin of the tank and stored at 4°C for a day before further processing. 100 ml water samples were analysed for *Legionella* spp. CFU/100 ml after centrifugation and filtration according to ISO 11731-2:2004 (International Organization for Standardization, 2004) and the *mip*-gene sequencing was used for the identification of the *Legionella* species. *Pseudomonas aeruginosa* was evaluated based on ISO 16266:2008 and the total heterotrophic bacteria were counted as CFU in 1 ml at 36°C (ISO 6222:1999).

2 L water samples were stirred first and then vacuum-filtered onto a cellulose nitrate filter (12.5 cm2; 11407-50-ACN, Sartorius stedim) in 500 µl steps. By default, filter with 0.2 µm pore size were used. If due to filter clogging, the filter-duration of 500 µl water sample exceeds 1h, a new filter (0.45 µm ø) was used for further filtration. The filters were removed from the filter tower with a sterile tweezer and stored in a sterile petridish before further processing. For water samples which contains a lot of micro particles or biofilms, up to four filter were used for one single sample. Those filters were processed and treated equally in the subsequent DNA-isolation.

### DNA extraction, PCR amplification, and sequencing of 16S V3-V4 and 18S V9 amplicons

We used the PowerWater® DNA Isolation Kit (14900-100-NF; Mo Bio Laboratories, Germany) for the DNA extraction from the filters. All steps were performed according to standard centrifuge-based protocol recommended by the manufacturer. Briefly, the filters were inserted in a lysis buffer containing bead beating tube where the cells were mechanically and chemically lysed. Afterwards the lysate were transferred to a DNA-retaining spin column. Lysate of filters belonging to the sample were loaded onto the same spin column to pool the DNA. After washing with provided buffers to purify the bound DNA, it is subsequently eluted in DNA elution buffer and are ready for downstream PCR. We used a barcoded amplicon sequencing approach described in Herbold et al (Herbold et al., 2015). Fragments of prokaryotic and eukaryotic rRNA of each timepoint were amplified for this analysis, resulting in 26 samples for each sampling location, a total of 54 samples including negative controls. V3 and V4 regions of the bacterial 16S rRNA were amplified with barcoded versions of the primers Bakt\_341F (5’-CCTACGGGNGGCWGCAG-3’) and Bakt\_805R (5’-GACTACHVGGGTATCTAATCC-3’) (Herlemann et al., 2011). Further we used barcoded eukaryotic 18S rRNA primer-pair EUK\_1391F (5’-GTACACACCGCCCGTC-3’) and EUK\_1510R (5’-CCTTCYGCAGGTTCACCTAC-3’) based on those of Amaral-Zettler et al. (Amaral-Zettler et al., 2009) and the Earth Microbiome Project (Version 5 2012 Gilbert et al., 2014) to amplify the 18S-V9 region. Each PCR reaction included 1x DreamTaq Green Buffer (Fermentas, Thermo Fisher Scientific, Vienna, Austria), 2 mM MgCl2, 0.2 mM dNTP mix (Fermentas), 0.1 mg mL-1 bovine serum albumin, 1 µM of each of the forward and reverse primers, 0.025 U DreamTaq polymerase (Fermentas) and 1 µL of template. The products of the first PCR amplification (94°C for 3 min; 25 cycles of 45 sec at 94°C, 30 sec at 52°C for 16S and 57°C for 18S, 90 sec at 72°C; and 72°C for 10 min) was done in triplicates and pooled together as the template for the second barcoding PCR (95°C for 3 min; 10 cycles of 30 sec at 95°C, 30 sec at 55°C, 60 sec at 72°C; and 72°C for 7 min). The purification, quantification, and sequencing was done as described in Herbold et al.

### Data compilation, filtering, and clustering

Based on the primer and barcode of the read, the sequencing data was divided into two primer-datasets for each sample, resulting in a total of 72 different datasets. After sequencing, reads assigned to 18S-dataset were trimmed by 88 nucleotides prior to formation of operational taxonomic units (OTUs) because the expected amplicon length of 178 bp is shorter than the read length of 300 bp. We also did an additional Q-score end-trimming of the reads, where the read was trimmed to the first instance of a Q-score of 37. This approach was chosen to maximize the number of contigs produced for each dataset. 16S datasets did not undergo initial nucleotide trimming, since the expected amplicon length exceeds the read length, and were directly end-trimmed based on the determined Q-score. The Q-score determination and further clustering into OTUs was performed as described previously (Craig W. Herbold et al., 2015). The OTU taxonomic classification with an identity threshold of 97% was carried out using the mothur classify.seqs function (Schloss et al., 2009) and the Silva 1.19 SSU database as reference (Quast et al., 2013) for both 16S and 18S OTUs. Further the protist ribosomal database (Guillou et al., 2013) has been used for more accurate taxonomic characterization of the eukaryotic 18S. Every OTU with a classification confidence score lower than 80 for 16S and 60 for 18S were changed to ‘unclassified’.

### OTU correlation analysis

The program SparCC (Friedman and Alm, 2012) was used for the correlation analysis between OTUs in the two sampled locations.

Amaral-Zettler, L. a, McCliment, E. a, Ducklow, H.W., and Huse, S.M. (2009) A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS One* **4**: e6372.

Friedman, J. and Alm, E.J. (2012) Inferring Correlation Networks from Genomic Survey Data. *PLoS Comput. Biol.* **8**: 1–11.

Gilbert, J.A., Jansson, J.K., and Knight, R. (2014) The Earth Microbiome project: successes and aspirations. *BMC Biol.* **12**: 69.

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., et al. (2013) The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* **41**: D597–604.

Herbold, C.W., Pelikan, C., Kuzyk, O., Hausmann, B., Angel, R., Berry, D., and Loy, A. (2015) A flexible and economical barcoding approach for highly multiplexed amplicon sequencing of diverse target genes. *Front. Microbiol.* **6**: 1–8.

Herbold, C.W., Pelikan, C., Kuzyk, O., Hausmann, B., Angel, R., Berry, D., and Loy, A. (2015) A flexible and economical barcoding approach for highly multiplexed amplicon sequencing of diverse target genes A flexible and economical barcoding approach for highly multiplexed amplicon sequencing of diverse target genes. **6**: 1–8.

Herlemann, D.P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J.J., and Andersson, A.F. (2011) Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J.* **5**: 1571–1579.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**: 590–596.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al. (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75**: 7537–7541.