

**Supplementary information for:**

*The subway microbiome: Seasonal dynamics and direct comparison of air and surface bacterial communities.*

**Authors:**

Jostein Gohli<sup>1\*</sup>, Kari Oline Bøifot<sup>1,2</sup>, Line Victoria Moen<sup>1</sup>, Paulina Pastuszek<sup>3</sup>, Gunnar Skogan<sup>1</sup>,  
Klas I. Udekwu<sup>4</sup>, and Marius Dybwad<sup>1,2</sup>

- <sup>1)</sup> Norwegian Defence Research Establishment, P O Box 25, NO-2027 Kjeller, Norway
- <sup>2)</sup> Department of Analytics, Environmental & Forensic Sciences, King's College London,  
150 Stamford Street, London SE1 9NH, UK
- <sup>3)</sup> Department Of Molecular Biosciences, Stockholm University, TWGI Stockholm, SE  
10691
- <sup>4)</sup> SoS BIOs Sweden, Tiundagatan 41, Uppsala SE 752 30

**Table S1** – Type of environment, latitude and longitude for all sampled stations.

| Station            | Environment | Latitude | Longitude |
|--------------------|-------------|----------|-----------|
| Carl Berners plass | Indoor      | 59.92566 | 10.77959  |
| Ellingsrudåsen     | Indoor      | 59.93651 | 10.91664  |
| Grønland           | Indoor      | 59.91278 | 10.76009  |
| Helsfyr            | Indoor      | 59.91259 | 10.80139  |
| Jernbanetorget     | Indoor      | 59.91226 | 10.75027  |
| Lindeberg          | Indoor      | 59.93276 | 10.88167  |
| Løren              | Indoor      | 59.93010 | 10.78959  |
| Nationaltheateret  | Indoor      | 59.91546 | 10.73209  |
| Nydalen            | Indoor      | 59.94940 | 10.76592  |
| Romsås             | Indoor      | 59.96260 | 10.89248  |
| Stortinget         | Indoor      | 59.91343 | 10.74329  |
| Tøyen              | Indoor      | 59.91558 | 10.77580  |
| Vestli             | Indoor      | 59.97204 | 10.92944  |
| Forskningsparken   | Outdoor     | 59.94370 | 10.72088  |
| Majorstuen         | Outdoor     | 59.93016 | 10.71473  |
| Montebello         | Outdoor     | 59.93482 | 10.66640  |

**Table S2** - Overview of all samples included in the analyses.

| Season | Air samples (n) | Surface samples (n) | Collection date |
|--------|-----------------|---------------------|-----------------|
| Autumn | 16              | 44                  | 17.nov.16       |
| Winter | 15              | 43                  | 07.feb.17       |
| Spring | 14              | 42                  | 20.apr.17       |
| Summer | 24              | 48                  | 28.jun.17       |

**Table S3** - PCR program for 16S rRNA gene amplicon sequencing.

|                 | Temperature (°C) | Time   | Cycles |
|-----------------|------------------|--------|--------|
| Denaturation    | 98               | 30 sec | 1      |
| Denaturation    | 98               | 10 sec | 30     |
| Annealing       | 55               | 30 sec |        |
| Extension       | 72               | 30 sec |        |
| Final extension | 72               | 4 min  | 1      |
| Hold            | 4                |        |        |

**Table S4** – The best-fit models of qPCR 16S rRNA gene copies for A) air samples and B) surface samples.

| <b>A</b>         |           |               |                |          |                  |     |
|------------------|-----------|---------------|----------------|----------|------------------|-----|
| <i>Predictor</i> | <i>Df</i> | <i>Sum Sq</i> | <i>Mean Sq</i> | <i>F</i> | <i>P</i>         |     |
| Humidity SD      | 1         | 0.46          | 0.46           | 0.55     | 0.461            |     |
| Humidity         | 1         | 7.02          | 7.02           | 8.52     | <b>0.005</b>     | **  |
| Season           | 3         | 14.34         | 4.78           | 5.80     | <b>0.002</b>     | **  |
| Residuals        | 55        | 45.33         | 0.82           |          |                  |     |
| <b>B</b>         |           |               |                |          |                  |     |
| <i>Predictor</i> | <i>Df</i> | <i>Sum Sq</i> | <i>Mean Sq</i> | <i>F</i> | <i>P</i>         |     |
| Sequence run     | 2         | 49.51         | 24.75          | 27.25    | <b>&lt;0.001</b> | *** |
| Humidity SD      | 1         | 1.44          | 1.44           | 1.59     | 0.209            |     |
| Temperature      | 1         | 0.78          | 0.78           | 0.86     | 0.355            |     |
| Humidity         | 1         | 2.21          | 2.21           | 2.43     | 0.121            |     |
| Season           | 3         | 15.18         | 5.06           | 5.57     | <b>0.001</b>     | **  |
| Indoor/Outdoor   | 1         | 17.34         | 17.34          | 19.09    | <b>&lt;0.001</b> | *** |
| Surface type     | 2         | 48.91         | 24.46          | 26.92    | <b>&lt;0.001</b> | *** |
| Residuals        | 164       | 148.96        | 0.91           |          |                  |     |

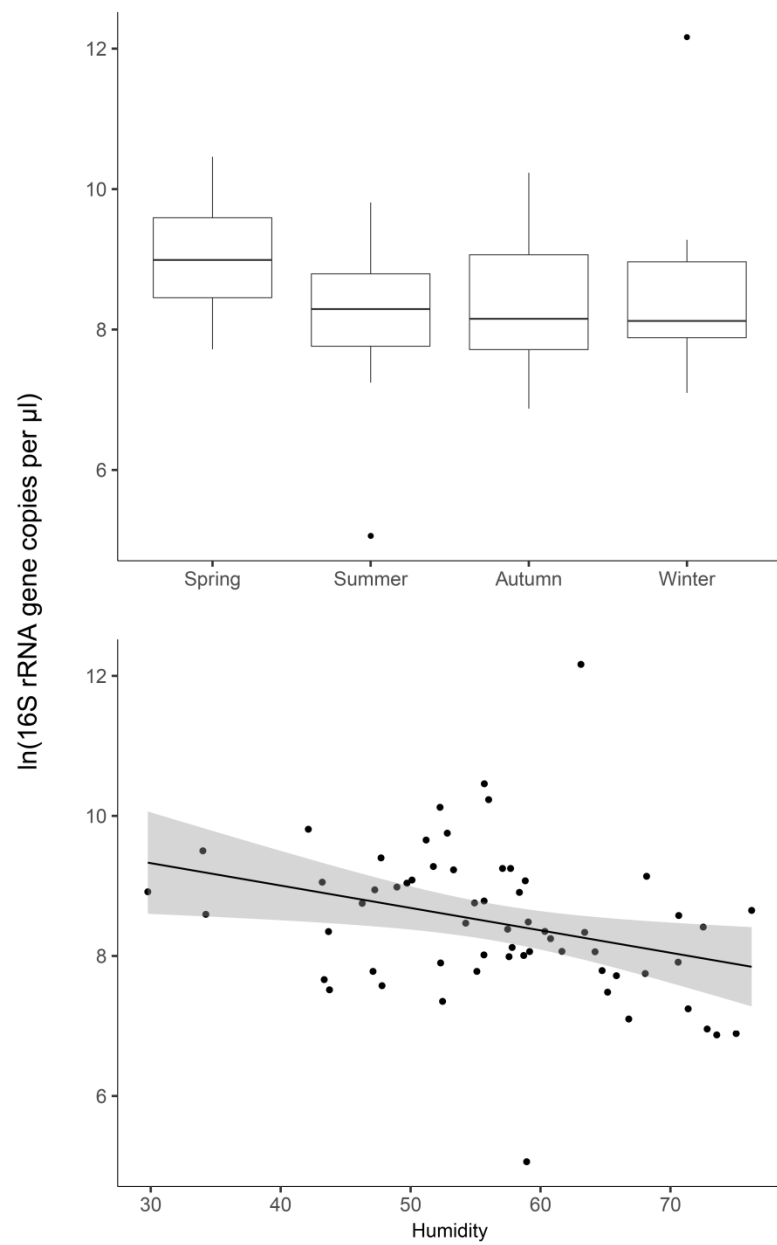
**Table S5** - Top 20 phyla, families, and genera and species in surface samples collected on kiosks (N=62), benches (N=59), and railings (N=56). Dots indicate that a group is also represented in the top 20 set from the other two surface types.

| <i>Kiosk samples</i>        |   |           | <i>Bench samples</i>        |   |           | <i>Railing samples</i>      |   |           |
|-----------------------------|---|-----------|-----------------------------|---|-----------|-----------------------------|---|-----------|
| Phylum                      |   | abundance | Phylum                      |   | abundance | Phylum                      |   | abundance |
| <i>Actinobacteria</i>       | • | 32.40 %   | <i>Actinobacteria</i>       | • | 31.94 %   | <i>Actinobacteria</i>       | • | 29.51 %   |
| <i>Proteobacteria</i>       | • | 26.73 %   | <i>Proteobacteria</i>       | • | 28.28 %   | <i>Proteobacteria</i>       | • | 27.35 %   |
| <i>Firmicutes</i>           | • | 18.26 %   | <i>Firmicutes</i>           | • | 13.98 %   | <i>Firmicutes</i>           | • | 17.91 %   |
| <i>Bacteroidetes</i>        | • | 9.96 %    | <i>Bacteroidetes</i>        | • | 11.04 %   | <i>Bacteroidetes</i>        | • | 12.02 %   |
| <i>Cyanobacteria</i>        | • | 6.74 %    | <i>Cyanobacteria</i>        | • | 9.60 %    | <i>Cyanobacteria</i>        | • | 7.19 %    |
| <i>Deinococcus-Thermus</i>  | • | 1.60 %    | <i>Deinococcus-Thermus</i>  | • | 1.31 %    | <i>Deinococcus-Thermus</i>  | • | 1.19 %    |
| <i>Acidobacteria</i>        | • | 1.00 %    | <i>Acidobacteria</i>        | • | 1.05 %    | <i>Acidobacteria</i>        | • | 1.11 %    |
| <i>Chloroflexi</i>          | • | 0.66 %    | <i>Chloroflexi</i>          | • | 0.67 %    | <i>Fusobacteria</i>         | • | 0.80 %    |
| <i>Fusobacteria</i>         | • | 0.55 %    | <i>Planctomycetes</i>       | • | 0.46 %    | <i>Chloroflexi</i>          | • | 0.68 %    |
| <i>Planctomycetes</i>       | • | 0.50 %    | <i>Fusobacteria</i>         | • | 0.44 %    | <i>Planctomycetes</i>       | • | 0.56 %    |
| <i>Verrucomicrobia</i>      | • | 0.40 %    | <i>Verrucomicrobia</i>      | • | 0.32 %    | <i>Patescibacteria</i>      | • | 0.39 %    |
| <i>Patescibacteria</i>      | • | 0.32 %    | <i>Patescibacteria</i>      | • | 0.29 %    | <i>Verrucomicrobia</i>      | • | 0.39 %    |
| <i>Gemmatimonadetes</i>     | • | 0.29 %    | <i>Gemmatimonadetes</i>     | • | 0.22 %    | <i>Gemmatimonadetes</i>     | • | 0.27 %    |
| <i>Euryarchaeota</i>        | • | 0.19 %    | <i>FBP</i>                  | • | 0.17 %    | <i>FBP</i>                  | • | 0.20 %    |
| <i>FBP</i>                  | • | 0.18 %    | <i>Armatimonadetes</i>      | • | 0.11 %    | <i>Armatimonadetes</i>      | • | 0.14 %    |
| <i>Armatimonadetes</i>      | • | 0.13 %    | <i>Epsilonbacteraeota</i>   | • | 0.03 %    | <i>Euryarchaeota</i>        | • | 0.13 %    |
| <i>Epsilonbacteraeota</i>   | • | 0.02 %    | <i>Euryarchaeota</i>        | • | 0.02 %    | <i>Epsilonbacteraeota</i>   | • | 0.04 %    |
| <i>Chlamydiae</i>           | • | 0.02 %    | <i>Chlamydiae</i>           | • | 0.02 %    | <i>Chlamydiae</i>           | • | 0.03 %    |
| <i>Spirochaetes</i>         | • | 0.01 %    | <i>Spirochaetes</i>         | • | 0.01 %    | <i>Spirochaetes</i>         | • | 0.02 %    |
| <i>Fibrobacteres</i>        |   | 0.01 %    | <i>WPS-2</i>                |   | 0.01 %    | <i>Nitrospirae</i>          |   | 0.01 %    |
| Family                      |   | abundance | Family                      |   | abundance | Family                      |   | abundance |
| <i>Unassigned</i>           | • | 12.72 %   | <i>Unassigned</i>           | • | 16.64 %   | <i>Unassigned</i>           | • | 13.16 %   |
| <i>Micrococcaceae</i>       | • | 7.75 %    | <i>Micrococcaceae</i>       | • | 7.33 %    | <i>Micrococcaceae</i>       | • | 6.51 %    |
| <i>Staphylococcaceae</i>    | • | 5.88 %    | <i>Sphingomonadaceae</i>    | • | 6.32 %    | <i>Staphylococcaceae</i>    | • | 5.47 %    |
| <i>Sphingomonadaceae</i>    | • | 5.80 %    | <i>Burkholderiaceae</i>     | • | 4.95 %    | <i>Streptococcaceae</i>     | • | 5.37 %    |
| <i>Corynebacteriaceae</i>   | • | 4.17 %    | <i>Streptococcaceae</i>     | • | 4.49 %    | <i>Sphingomonadaceae</i>    | • | 5.34 %    |
| <i>Burkholderiaceae</i>     | • | 4.05 %    | <i>Hymenobacteraceae</i>    | • | 4.33 %    | <i>Burkholderiaceae</i>     | • | 4.23 %    |
| <i>Streptococcaceae</i>     | • | 3.77 %    | <i>Staphylococcaceae</i>    | • | 3.75 %    | <i>Moraxellaceae</i>        | • | 4.06 %    |
| <i>Hymenobacteraceae</i>    | • | 3.39 %    | <i>Corynebacteriaceae</i>   | • | 2.76 %    | <i>Hymenobacteraceae</i>    | • | 3.78 %    |
| <i>Moraxellaceae</i>        | • | 3.32 %    | <i>Moraxellaceae</i>        | • | 2.75 %    | <i>Corynebacteriaceae</i>   | • | 3.12 %    |
| <i>Acetobacteraceae</i>     | • | 2.71 %    | <i>Acetobacteraceae</i>     | • | 2.73 %    | <i>Acetobacteraceae</i>     | • | 2.46 %    |
| <i>Propionibacteriaceae</i> | • | 2.37 %    | <i>Nocardiodaceae</i>       | • | 2.38 %    | <i>Flavobacteriaceae</i>    | • | 2.37 %    |
| <i>Nocardiodaceae</i>       | • | 2.14 %    | <i>Beijerinckiaceae</i>     | • | 2.29 %    | <i>Nocardiodaceae</i>       | • | 2.37 %    |
| <i>Beijerinckiaceae</i>     | • | 2.02 %    | <i>Propionibacteriaceae</i> | • | 2.24 %    | <i>Propionibacteriaceae</i> | • | 1.98 %    |
| <i>Microbacteriaceae</i>    | • | 1.69 %    | <i>Microbacteriaceae</i>    | • | 2.21 %    | <i>Microbacteriaceae</i>    | • | 1.89 %    |
| <i>Lactobacillaceae</i>     |   | 1.58 %    | <i>Flavobacteriaceae</i>    | • | 1.92 %    | <i>Beijerinckiaceae</i>     | • | 1.84 %    |
| <i>Deinococcaceae</i>       |   | 1.42 %    | <i>Intrasporangiaceae</i>   | • | 1.73 %    | <i>Prevotellaceae</i>       |   | 1.52 %    |
| <i>Flavobacteriaceae</i>    | • | 1.39 %    | <i>Geodermatophilaceae</i>  |   | 1.72 %    | <i>Intrasporangiaceae</i>   | • | 1.48 %    |
| <i>Family_XI</i>            |   | 1.37 %    | <i>Rhodobacteraceae</i>     |   | 1.46 %    | <i>Geodermatophilaceae</i>  |   | 1.34 %    |
| <i>Intrasporangiaceae</i>   | • | 1.35 %    | <i>Deinococcaceae</i>       |   | 1.14 %    | <i>Rhodobacteraceae</i>     |   | 1.21 %    |
| <i>Pseudomonadaceae</i>     |   | 1.34 %    | <i>Pseudomonadaceae</i>     |   | 1.03 %    | <i>Lactobacillaceae</i>     |   | 1.19 %    |
| Genus                       |   | abundance | Genus                       |   | abundance | Genus                       |   | abundance |
| <i>Unassigned</i>           | • | 22.58 %   | <i>Unassigned</i>           | • | 28.00 %   | <i>Unassigned</i>           | • | 23.36 %   |
| <i>Staphylococcus</i>       | • | 5.56 %    | <i>Sphingomonas</i>         | • | 4.58 %    | <i>Streptococcus</i>        | • | 4.90 %    |
| <i>Sphingomonas</i>         | • | 4.31 %    | <i>Hymenobacter</i>         | • | 4.21 %    | <i>Staphylococcus</i>       | • | 4.80 %    |
| <i>Corynebacterium</i>      | • | 3.43 %    | <i>Streptococcus</i>        | • | 3.87 %    | <i>Sphingomonas</i>         | • | 3.76 %    |
| <i>Hymenobacter</i>         | • | 3.32 %    | <i>Staphylococcus</i>       | • | 3.43 %    | <i>Hymenobacter</i>         | • | 3.68 %    |
| <i>Streptococcus</i>        | • | 3.16 %    | <i>Arthrobacter</i>         | • | 2.54 %    | <i>Corynebacterium</i>      | • | 2.56 %    |
| <i>Kocuria</i>              | • | 2.65 %    | <i>Corynebacterium</i>      | • | 2.13 %    | <i>Flavobacterium</i>       | • | 2.00 %    |
| <i>Micrococcus</i>          | • | 2.22 %    | <i>Kocuria</i>              | • | 1.85 %    | <i>Psychrobacter</i>        | • | 1.83 %    |
| <i>Arthrobacter</i>         | • | 1.81 %    | <i>Psychrobacter</i>        | • | 1.52 %    | <i>Micrococcus</i>          | • | 1.82 %    |
| <i>Psychrobacter</i>        | • | 1.60 %    | <i>Flavobacterium</i>       | • | 1.48 %    | <i>Arthrobacter</i>         | • | 1.81 %    |
| <i>Lactobacillus</i>        |   | 1.57 %    | <i>Nocardioidea</i>         | • | 1.47 %    | <i>Nocardioidea</i>         | • | 1.49 %    |
| <i>Cutibacterium</i>        | • | 1.43 %    | <i>Micrococcus</i>          | • | 1.43 %    | <i>Kocuria</i>              | • | 1.46 %    |
| <i>Deinococcus</i>          | • | 1.41 %    | <i>Massilia</i>             | • | 1.22 %    | <i>Lactobacillus</i>        |   | 1.19 %    |
| <i>Nocardioidea</i>         | • | 1.33 %    | <i>Deinococcus</i>          | • | 1.13 %    | <i>Cutibacterium</i>        | • | 1.15 %    |
| <i>Pseudomonas</i>          |   | 1.20 %    | <i>Cutibacterium</i>        | • | 1.12 %    | <i>Acinetobacter</i>        |   | 1.14 %    |
| <i>Flavobacterium</i>       | • | 1.17 %    | <i>1174-901-12</i>          |   | 1.03 %    | <i>Massilia</i>             | • | 0.91 %    |
| <i>Massilia</i>             | • | 0.89 %    | <i>Pseudomonas</i>          |   | 0.91 %    | <i>Haemophilus</i>          |   | 0.91 %    |
| <i>1174-901-12</i>          |   | 0.79 %    | <i>Roseomonas</i>           |   | 0.86 %    | <i>Deinococcus</i>          | • | 0.88 %    |
| <i>Methylobacterium</i>     |   | 0.79 %    | <i>Blastococcus</i>         |   | 0.83 %    | <i>Acidiphilium</i>         |   | 0.81 %    |
| <i>Pedobacter</i>           |   | 0.76 %    | <i>Nakamurella</i>          |   | 0.82 %    | <i>Veillonella</i>          |   | 0.75 %    |

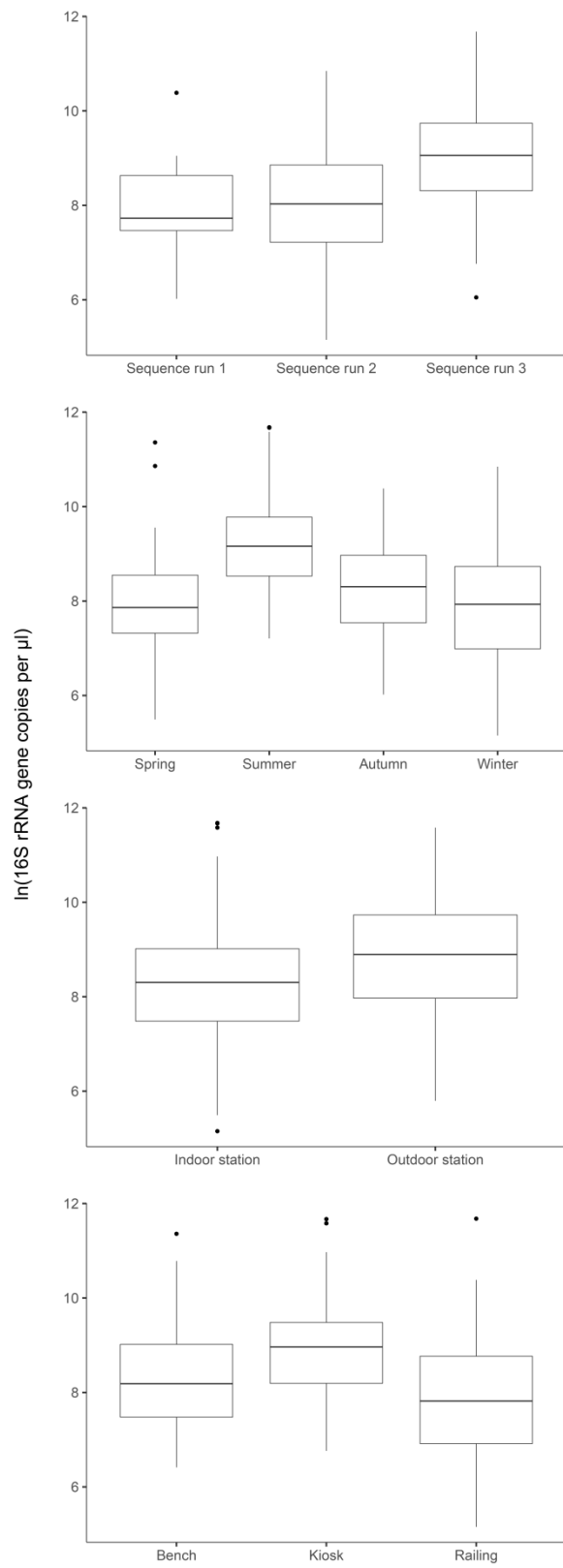
**Table S6** - Random forest classification models of samples collected from different surface types. Confusion matrices show the classification of samples and the associated class error. The mean decrease in model accuracy (MDA; from removing the genera in question) and Z-scores are given for the 20 most important genera for classifying samples.

| <b>OOB estimate of error rate: 42.37 %</b>   |       |       |         |             |
|----------------------------------------------|-------|-------|---------|-------------|
| Confusion matrix:                            | bench | kiosk | railing | class error |
| bench                                        | 39    | 13    | 7       | 33.9 %      |
| kiosk                                        | 10    | 48    | 4       | 22.6 %      |
| railing                                      | 19    | 22    | 15      | 73.2 %      |
| <b>Most important genera in sample</b>       |       |       |         |             |
| Genera                                       | bench | kiosk | railing | MDA         |
| <i>Prevotellaceae-Prevotella_9</i>           | 0.005 | 0.004 | 0.001   | 0.003       |
| <i>Micrococcaceae-Arthrobacter</i>           | 0.001 | 0.005 | 0.000   | 0.002       |
| <i>Moraxellaceae-Enhydrobacter</i>           | 0.000 | 0.003 | 0.003   | 0.002       |
| <i>Lachnospiraceae-Agathobacter</i>          | 0.003 | 0.000 | 0.002   | 0.002       |
| <i>Intrasporangiaceae-Ornithinimicrobium</i> | 0.004 | 0.001 | 0.000   | 0.002       |
| <i>Ruminococcaceae-Faecalibacterium</i>      | 0.002 | 0.001 | 0.001   | 0.001       |
| <i>Lachnospiraceae-Coproccoccus_3</i>        | 0.002 | 0.002 | 0.000   | 0.001       |
| <i>Peptostreptococcaceae-Romboutsia</i>      | 0.002 | 0.001 | 0.001   | 0.001       |
| <i>Rhodobacteraceae-Rubellimicrobium</i>     | 0.001 | 0.002 | 0.001   | 0.001       |
| <i>Corynebacteriaceae-Corynebacterium_1</i>  | 0.001 | 0.002 | 0.000   | 0.001       |
| <i>Burkholderiaceae-Variovorax</i>           | 0.000 | 0.002 | 0.001   | 0.001       |
| <i>Microbacteriaceae-Rathayibacter</i>       | 0.001 | 0.003 | 0.000   | 0.001       |
| <i>Geodermatophilaceae-Geodermatophilus</i>  | 0.000 | 0.002 | 0.001   | 0.001       |
| <i>Micrococcaceae-Rothia</i>                 | 0.001 | 0.001 | 0.001   | 0.001       |
| <i>Actinomycetaceae-Actinomyces</i>          | 0.001 | 0.002 | 0.001   | 0.001       |
| <i>Family_XI-Peptoniphilus</i>               | 0.002 | 0.002 | -0.001  | 0.001       |
| <i>Streptococcaceae-Anthococcus</i>          | 0.000 | 0.002 | 0.000   | 0.001       |
| <i>Prevotellaceae-Prevotella_7</i>           | 0.001 | 0.000 | 0.002   | 0.001       |
| <i>Rhodobacteraceae-Paracoccus</i>           | 0.000 | 0.003 | 0.000   | 0.001       |
| <i>Bifidobacteriaceae-Gardnerella</i>        | 0.001 | 0.001 | 0.001   | 0.001       |

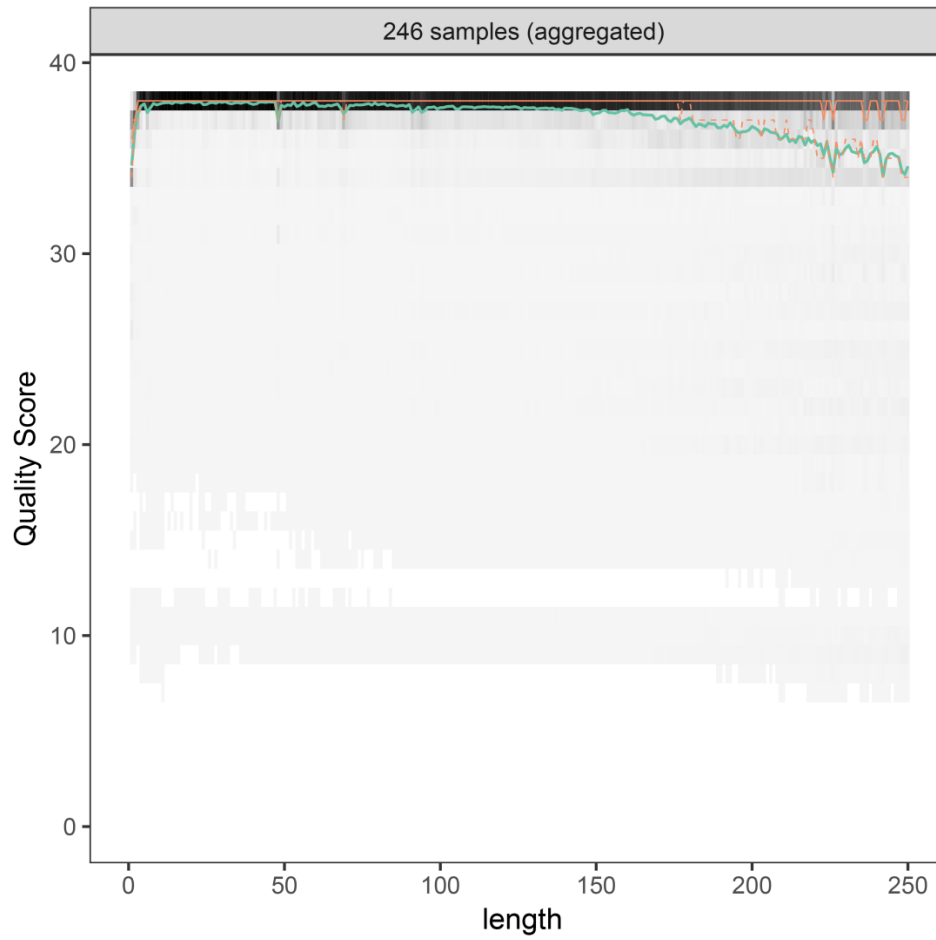
**Figure S1** – The significant predictors of qPCR 16S rRNA gene copy yields in air samples.



**Figure S2** – The significant predictors of qPCR 16S rRNA gene copy yields in surface samples.

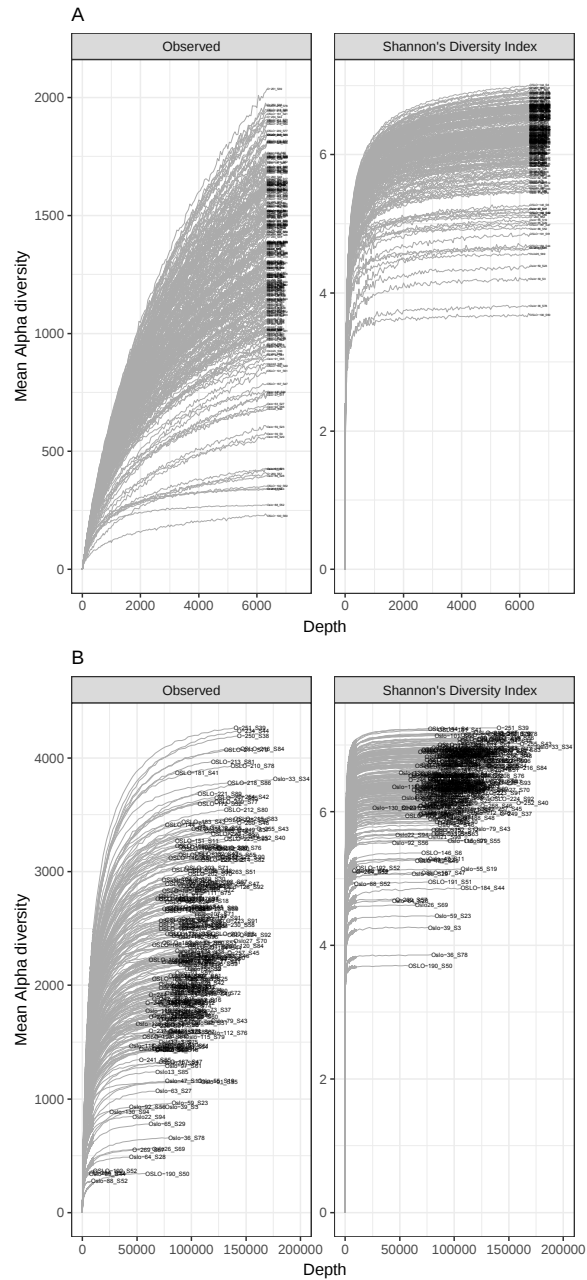


**Figure S3** – Quality profile of filtered reads (forward reads only). The quality score (QS) distribution is shown as a grey-scale heat map, where dark colors indicating higher frequency. The green is the mean QS, orange the median, and dashed orange lines give the 25<sup>th</sup> and 75<sup>th</sup> quantiles.

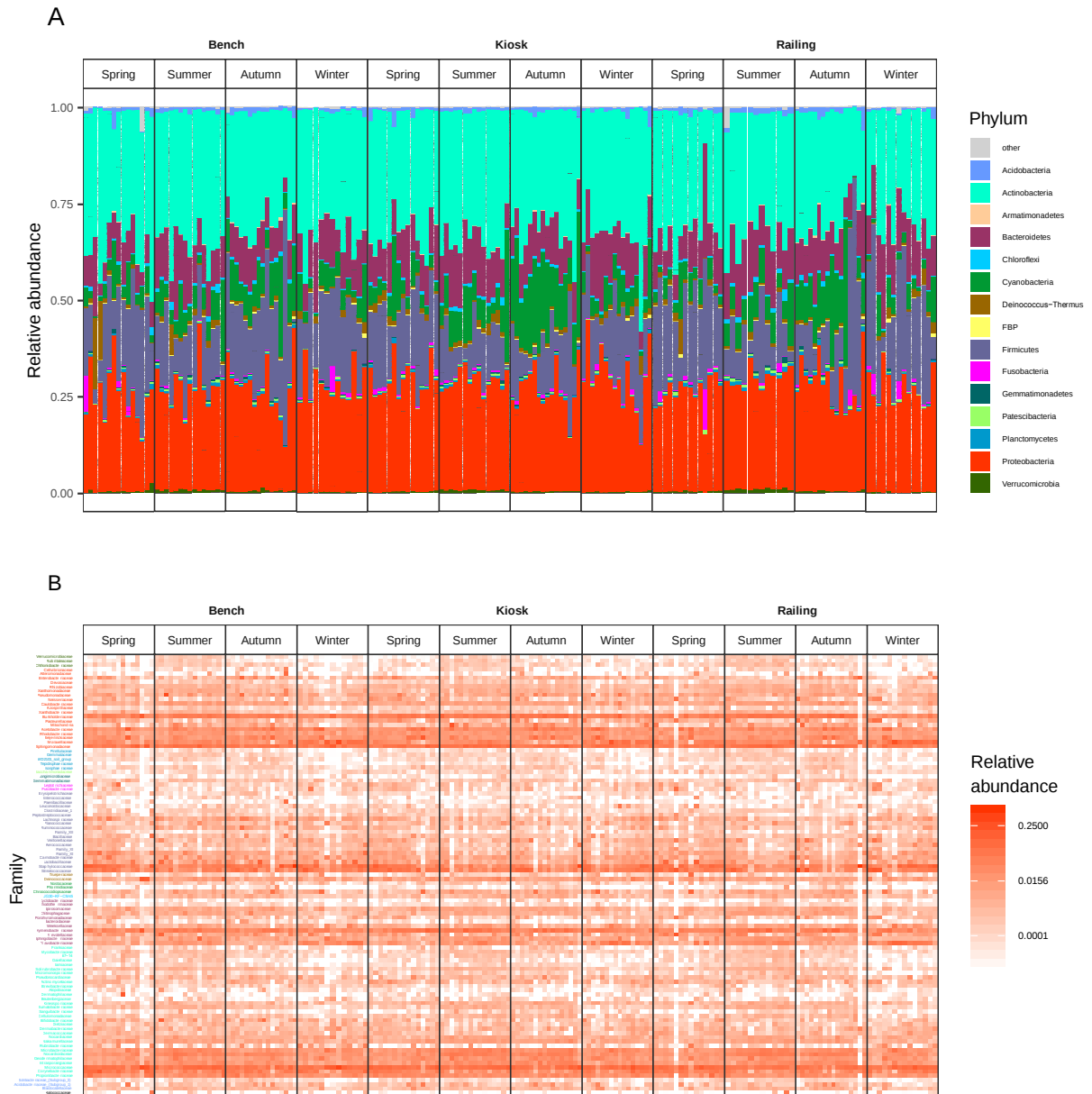


**Figure S4** - Rarefaction curves with observed diversity and Shannon's Diversity Index.

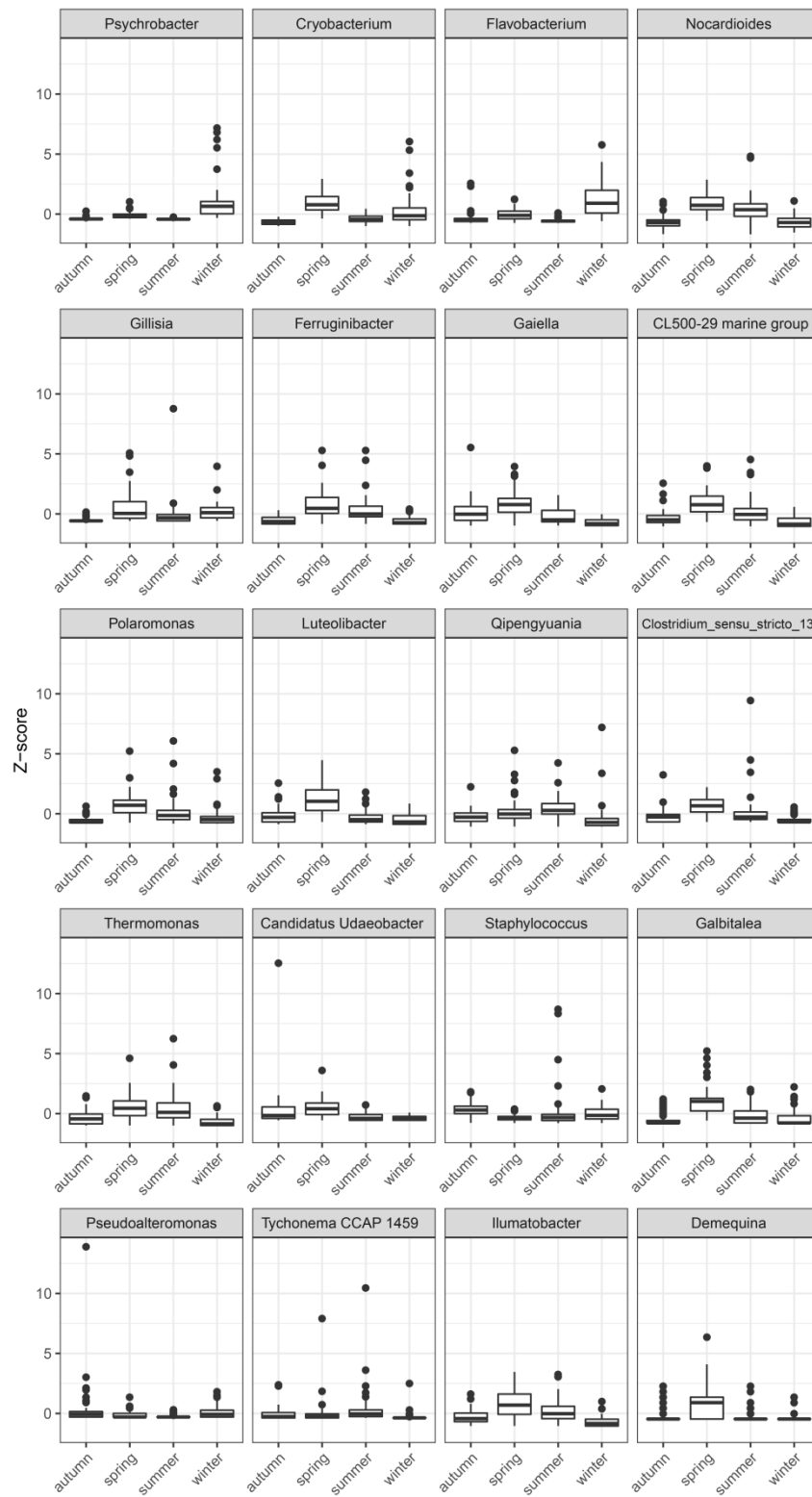
Diversity scores were obtained/calculated for sample sizes (reads) from 1 to 176394 (maximum sample size), with increasing interval size, using 5 iterations for each sample (1420 calculations \* two alpha diversity metrics). Mean values from iterations are shown. Panel A shows rarefaction curves to the rarefaction level used (smallest sample size; 6358), and B shows rarefaction plots for all available sample sizes.



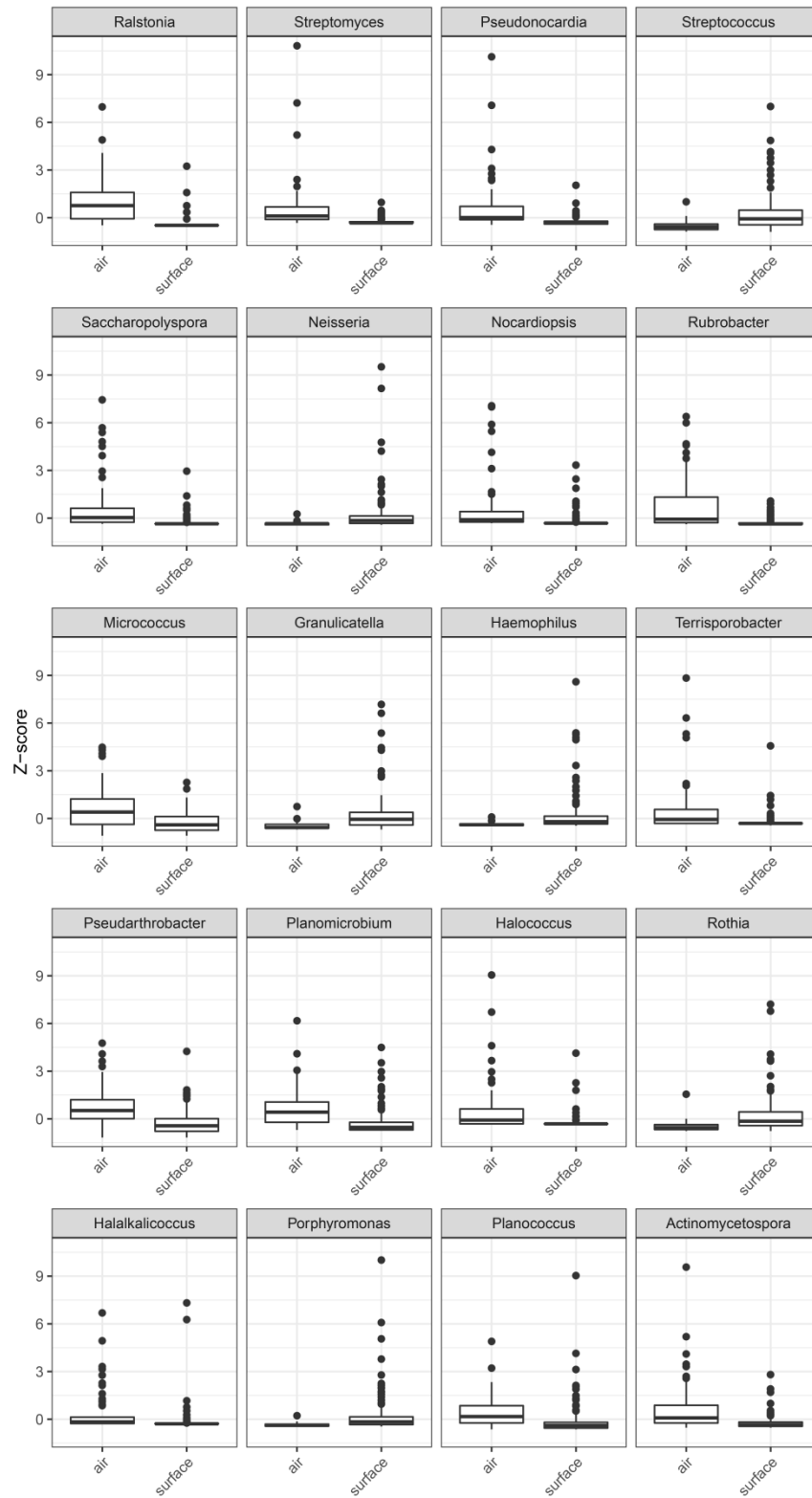
**Figure S5 - A:** Relative abundances of the top 15 phyla across the three surface types and seasons. **B:** Heatmap of most abundant families (relative abundance  $\geq 0.01$ ), color coded by phylum following the legend in panel A.



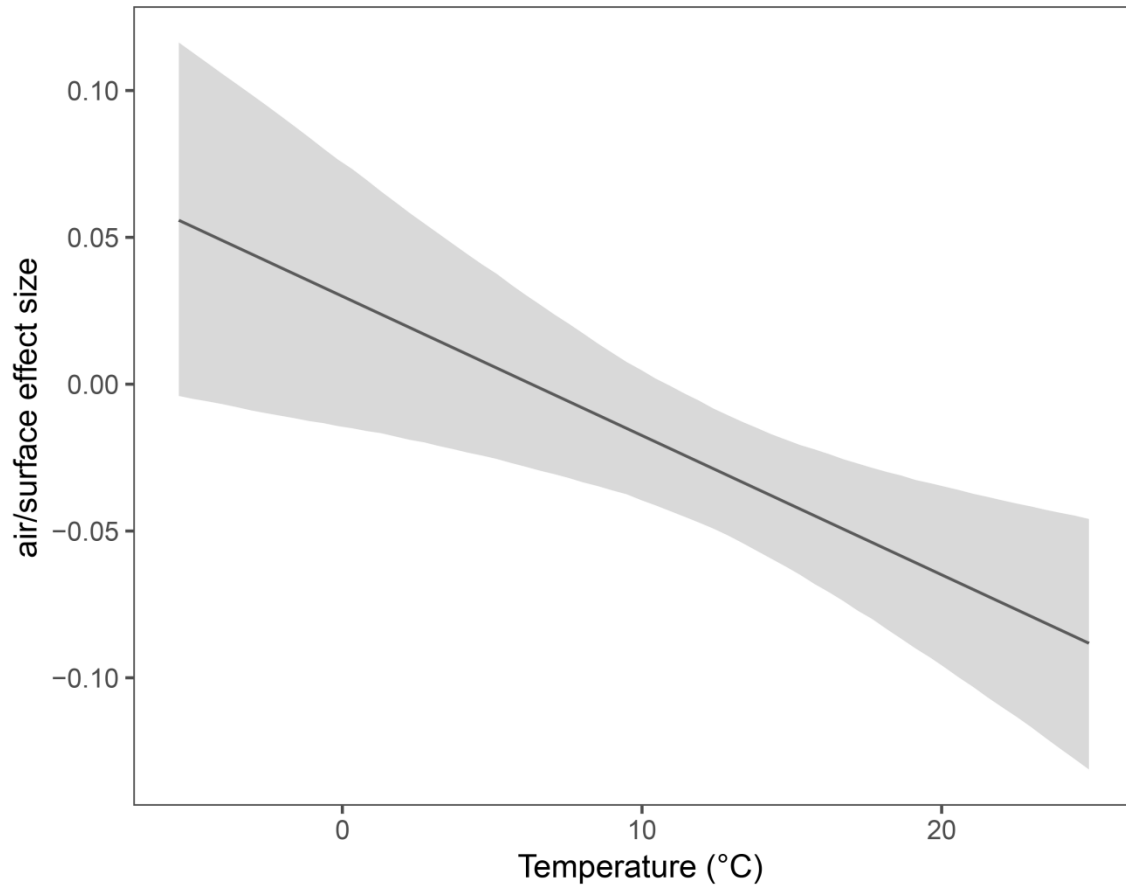
**Figure S6** - Top 20 most important genera in random forest classification analysis of samples collected in different seasons.



**Figure S7** - Top 20 most important genera in random forest classification analysis of air and surface samples.



**Figure S8** – Interaction effect between temperature (°C) and air/surface in the linear model of Shannon’s diversity index. The reference level of air/surface is air; hence, the effect size refers to how much larger or smaller the Shannon’s score is for surface samples, compared to air samples, at different temperatures.



**Figure S9** – PCoA plot of Bray Curtis dissimilarity distances with the only significant predictor (surface type) from the PERMANOVA model that included only surface-specific predictors (surface type, material and coating). Dashed circles represent 95 % CI for each cluster.

