Skin, gut, and sand metagenomic data on placebo-controlled sandbox biodiversity intervention study

Roslund, Marja

2023-04

Roslund , M , Parajuli , A , Hui , N , Puhakka , R , Grönroos , M , Soininen , L , Nurminen , N , Oikarinen , S , Cinek , O , Kramna , L , Schroderus , A-M , Laitinen , O H , Kinnunen , T , Hyöty , H & Sinkkonen , A 2023 , ' Skin, gut, and sand metagenomic data on placebo-controlled sandbox biodiversity intervention study ' , Data in Brief , vol. 47 , 109003 . https://doi.org/10.1016/

http://hdl.handle.net/10138/356714
https://doi.org/10.1016/j.dib.2023.109003

Data in Brief 47 (2023) 109003

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Skin, gut, and sand metagenomic data on placebo-controlled sandbox biodiversity intervention study

Marja I. Roslund[a,b], Anirudra Parajuli[b,1], Nan Hui[b,2], Riikka Puhakka[b], Mira Grönroos[b], Laura Soininen[b], Noora Nurminen[c], Sami Oikarinen[c], Ondřej Cinek[d], Lenka Kramná[d], Anna-Mari Schroderus[e], Olli H. Laitinen[c], Tuure Kinnunen[e,f], Heikki Hyöty[c,g], Aki Sinkkonen[a,*]

[a] *Natural Resources Institute Finland (Luke), Helsinki and Turku, Finland*
[b] *Ecosystems and Environment Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, Niemenkatu 73, Lahti FI-15140, Finland*
[c] *Faculty of Medicine and Health Technology, Tampere University, Arvo Ylpön katu 34, Tampere FI-33520, Finland*
[d] *Department of Pediatrics, 2nd Faculty of Medicine, Charles University, V Úvalu 84, Praha 5, Prague 150 06, Czech Republic*
[e] *Department of Clinical Microbiology, Institute of Clinical Medicine, University of Eastern Finland, Kuopio, Finland*
[f] *Eastern Finland Laboratory Centre (ISLAB), Kuopio, Finland*
[g] *Fimlab Laboratories, Pirkanmaa Hospital District, Tampere, Finland*

ARTICLE INFO

Article history:
Received 30 January 2023
Accepted 15 February 2023
Available online 21 February 2023

Dataset link: Skin, gut, and sand metagenomic data tables on placebo-controlled sandbox biodiversity intervention study (Original data)
Dataset link: Sandbox sand soil, human skin and stool metagenome (Original data)

ABSTRACT

The metagenomic data presented in this article are related to the published research of "A Placebo-controlled double-blinded test of the biodiversity hypothesis of immune-mediated diseases: Environmental microbial diversity elicits changes in cytokines and increase in T regulatory cells in young children"
This database contains 16S ribosomal RNA (rRNA) metagenomics of sandbox sand and skin and gut microbiota of

* Corresponding author.
E-mail address: aki.sinkkonen@luke.fi (A. Sinkkonen).
Social media: @MarjaRoslund, @LukeFinlandInt (M.I. Roslund), @PuhakkaRiikka (R. Puhakka), @GronroosMira (M. Grönroos), @HEDIMED2020 (N. Nurminen), @biwe_stn (O.H. Laitinen), @KinnunenLab (T. Kinnunen), @HHyotyLab (H. Hyöty), @akisink, @LukeFinlandInt (A. Sinkkonen)
[1] The current affiliation for Parajuli A. is Department of Medicine, Karolinska Institutet, Huddinge, Sweden.
[2] The current affiliation for Hui N. is School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai, China.

https://doi.org/10.1016/j.dib.2023.109003

Keywords:
Human commensal microbiota
16s rRNA metagenome
Skin bacteria
Gut bacteria
Children
Biodiversity
Placebo

children in the intervention and placebo daycares. In intervention daycares, children aged 3–5 years were exposed to playground sand enriched with microbially diverse soil. In placebo daycares, children were exposed to visually similar as in intervention daycares, but microbially poor sand colored with peat. Sand, skin and gut metagenomics were analyzed at baseline and after 14 and 28 days of intervention by high throughput sequencing of bacterial 16S rRNA gene on the Illumina MiSeq platform. This dataset shows how skin bacterial community composition, including classes Gammaproteobacteria and Bacilli, changed, and how the relative abundance of over 30 bacterial genera shifted on the skin of children in the intervention treatment, while no shifts occurred in the placebo group.

Specifications Table

Subject	Microbiology: Microbiome
Specific subject area	16S Metagenomics of sandbox sand, skin and gut bacteriome of children.
Type of data	Table Figure
How the data were acquired	Metagenomics were determined by high throughput sequencing of bacterial 16S rRNA gene on the Illumina MiSeq platform. Raw sequence data was processed using Mothur version v1.35.1 [1]. All the statistical tests were done with R v3.6.1 [2] and with *vegan* [3] and *lme4* [4] packages.
Data format	Analyzed secondary data
Description of data collection	Sand, skin swab and stool samples were collected for microbial analyses at three time points; at baseline (day 0), after two weeks supervised period of guided activities in the sandbox (day 14), and after four weeks follow up period (day 28). A trained study nurse collected skin swab samples in daycare centers, and the children's parents collected stool samples at home.
Data source location	· City/Town/Region: Lahti · Country: Finland · Latitude and longitude for collected samples/data: 60°58′57.61″N, 25°39′41.44″E Raw data location: The National Center for Biotechnology Information: Sequence Read Archive, data identification number: BioProject ID: PRJNA746448
Data accessibility	Repository name: The National Center for Biotechnology Information: Sequence Read Archive Data identification number: BioProject ID: PRJNA746448 Direct URL to data: https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA746448 Repository name: Mendeley Data Data identification number: DOI:10.17632/5dnyrc3pnw.1. Direct URL to data: https://data.mendeley.com/datasets/5dnyrc3pnw/1 The sensitive data that support the findings of this study are available from University of Helsinki but restrictions defined in General Data Protection Regulation (EU 2016/679) and Finnish Data Protection Act 1050/2018 apply to the availability of these data, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission from the ethical committee of the local hospital district (Tampereen yliopistollisen sairaalan erityisvastuualue, Pirkanmaa, Finland, ethical number R18067).

(*continued on next page*)

Related research article: Marja I. Roslund, Anirudra Parajuli; Nan Hui; Riikka Puhakka; Mira Grönroos; Laura Soininen; Noora Nurminen; Sami Oikarinen; Ondřej Cinek; Lenka Kramná; Anna-Mari Schroderus; Olli H. Laitinen; Tuure Kinnunen; Heikki Hyöty; Aki Sinkkonen. 2022: A Placebo-controlled double-blinded test of the biodiversity hypothesis of immune-mediated diseases: Environmental microbial diversity elicits changes in cytokines and increase in T regulatory cells in young children. Ecotoxicology and Environmental safety. 10.1016/j.ecoenv.2022.113900.

Value of the Data

- According to the biodiversity hypothesis of immune-mediated diseases, lack of microbiological diversity in the everyday living environment is a core reason for dysregulation of immune tolerance. Unraveling the bacterial community using metagenomics provides a deeper understanding of microbiota that regulate immune system.
- The dataset unraveled bacterial community on the skin and in the gut of 3–5-year-old children playing with sandbox sand with high microbial diversity or with placebo sandbox sand with low microbial diversity.
- These data are beneficial for researchers and health professionals involved in environmental health, immunology and microbiology, as well as for decision makers and urban planners who design children's playgrounds and other public places.
- The data may be used to increase awareness or develop novel sustainable nature-based prophylactic practices to cope with the high incidence of immune-mediated diseases among urban dwellers.
- The data provides useful information on the microbiota of the skin and gut of healthy children that can be used by researchers to compare findings from other studies.

1. Objective

Reduced contact to microbially rich environment has long been hypothesized to be a core reason for dysregulation of immune system [5–7]. To better understand how contact with microbially rich material affects children's health-associated commensal microbiota on the skin and in the gut, a high-throughput sequencing experiment was designed to screen the bacterial shifts. The dataset described in this manuscript was generated during the 28 days placebo-controlled double-blinded trial in which children were exposed to sandbox sand with high microbial diversity or visually similar sand with low microbial content. This paper goes beyond what was published in our previous publications about temporal shifts of biodiversity in urban environment [7,8,9,10].The dataset in the current paper extends the information by Roslund et al. [11] about bacterial metagenome. Here we show that the relative abundance of more than 30 bacterial genera shifted on the skin in the intervention treatment, while no shifts were observed in the placebo group. We also visualize how skin bacterial community, including classes Gammaproteobacteria and Bacilli, shifted among the intervention but not the placebo group during the intervention.

2. Data Description

We herein report the analyzed T test results of sand bacterial richness, diversity and relative abundance (Mendeley Data: Table 1), and permutational multivariate analysis of variance results i.e. Beta diversity differences between enriched intervention and placebo sandbox sand on day 0, 14 and 28 (Mendeley Data: Table 2). Reported bacterial relative abundance shifts on the skin and in the gut were analyzed with Linear Mixed Models (Mendeley Data: Table 3 and Table 4, respectively) between intervention and placebo groups. The reported differences in gut

bacterial diversity and richness were analyzed with paired T tests or in case of non-normally distributed data with the Wilcoxon-Signed Rank Test (Mendeley Data: Table 5). Principle coordinated analysis (PCoA) for skin bacterial communities between intervention and placebo groups are shown in Figs. 1 and 2. All bacterial raw sequence data were accessioned into the Sequence Read Archive (BioProject ID: PRJNA746448). We include R code for Linear Mixed Models as supplementary.

3. Experimental Design, Materials and Methods

Sand bacterial metagenomes were obtained from sandboxes in the yards of intervention and placebo daycares. The skin and gut bacterial metagenome was obtained from twenty-six children, aged between 3 and 5 at baseline, after 14 days of organized activities (exposure period) in the sandbox and after 28 days (follow-up period). Detailed experimental methods are in related research article.

The amplicon sequencing of V4 region of 16S rRNA was performed on an Illumina Miseq instrument using a v3 reagent kit. The forward and reverse sequence .fastq files were aligned into contigs and sequences that had any ambiguous bases, homopolymers longer than 8 base pairs, or an overlap shorter than 50 base pairs were removed in Mothur version v1.35.1 [1]. Almost identical unique sequences (over 99%) were preclustered to remove erroneous reads [12]. Chimeras were screened and removed using UCHIME [13]. Detailed sequence processing methods are in related research article.

Real-time quantitative PCR (qPCR) for total 16S rRNA was performed as in [14]. Skin bacterial 16S rRNA counts (copies µl-1) varied between 214 and 64,700. There were no differences between treatments nor time points (LMM: R2 = 0.07, $P > 0.1$) and the random variance between study subjects was high in 16S rRNA counts (R2 = 0.12).

Statistical tests included *t*-test, Wilcoxon signed-rank test, Permutational Multivariate Analysis of Covariance, principal coordinates analysis and linear mixed-effect models. Statistical analyzes were done with R v3.6.1 [2] and the detailed methods are in related research article.

Ethics Statements

The study was carried out in accordance with the recommendations of the "*Finnish Advisory Board on Research Integrity*" with an approval from the ethics committee of the local hospital district (Pirkanmaa Hospital District, Finland, ethical number R18067). Written informed consents were obtained from all guardians that were in accordance with the *Declaration of Helsinki*.

The sensitive data that support the findings of this study are available from University of Helsinki but restrictions defined in General Data Protection Regulation (EU 2016/679) and Finnish Data Protection Act 1050/2018 apply to the availability of these data, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission from the ethical committee of the local hospital district (Tampereen yliopistollisen sairaalan erityisvastuualue, Pirkanmaa, Finland).

Funding

This work was supported by Business Finland (grant numbers 40333/14 and 6766/31/2017) (grant to A.S. and H.H), Tampere Tuberculosis Foundation (grant to H.H) and Strategic Research Council (grant numbers 346136 to A.S., 346138 to O.H.L.). This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No: 874864.

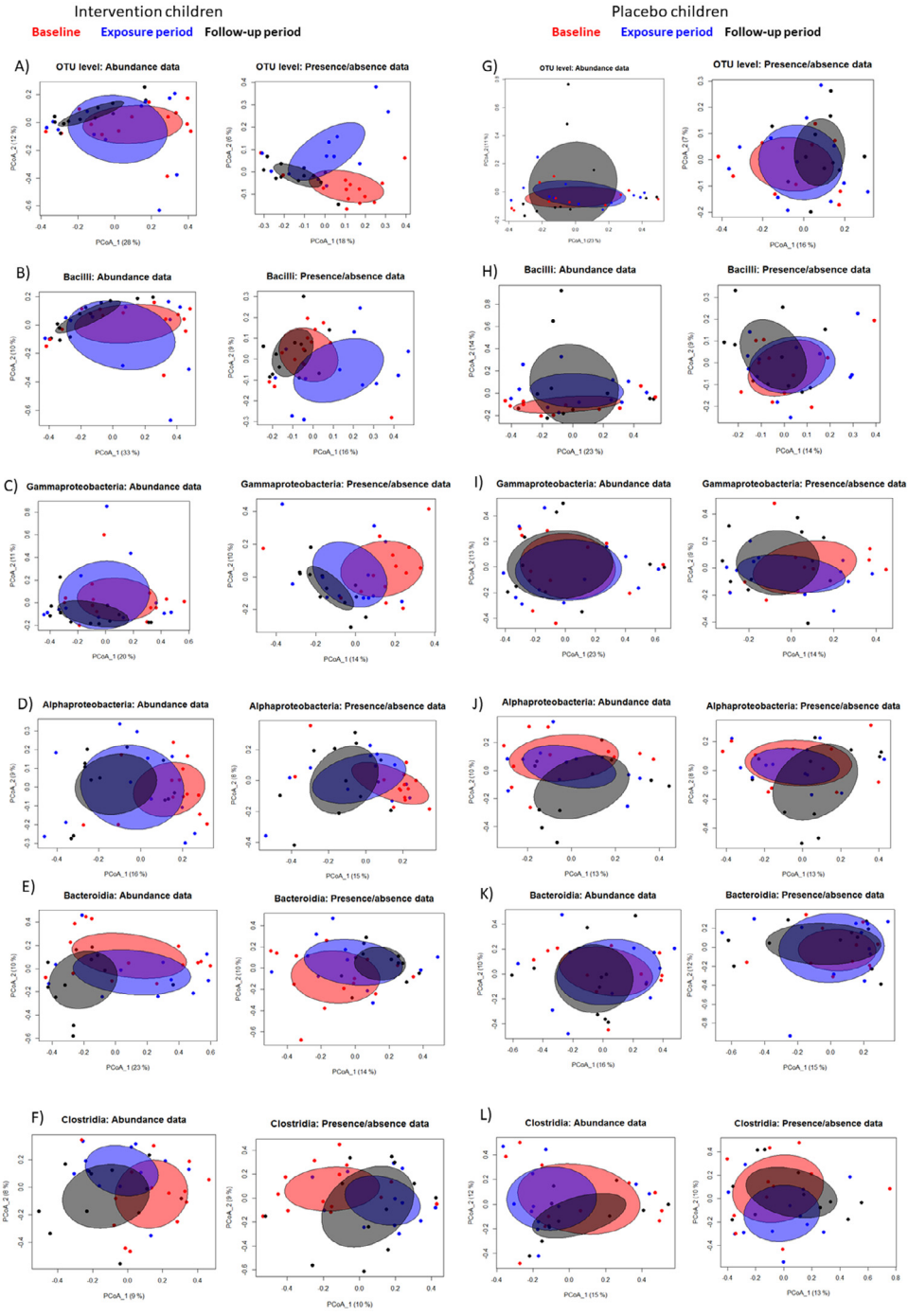

Fig. 1. Bacterial community composition shifts on the skin in intervention (A-F) and placebo groups (G-L) (Statistic are shown in related research article in table S5). Principle coordinated analysis plots are calculated with Bray-Curtis metric.

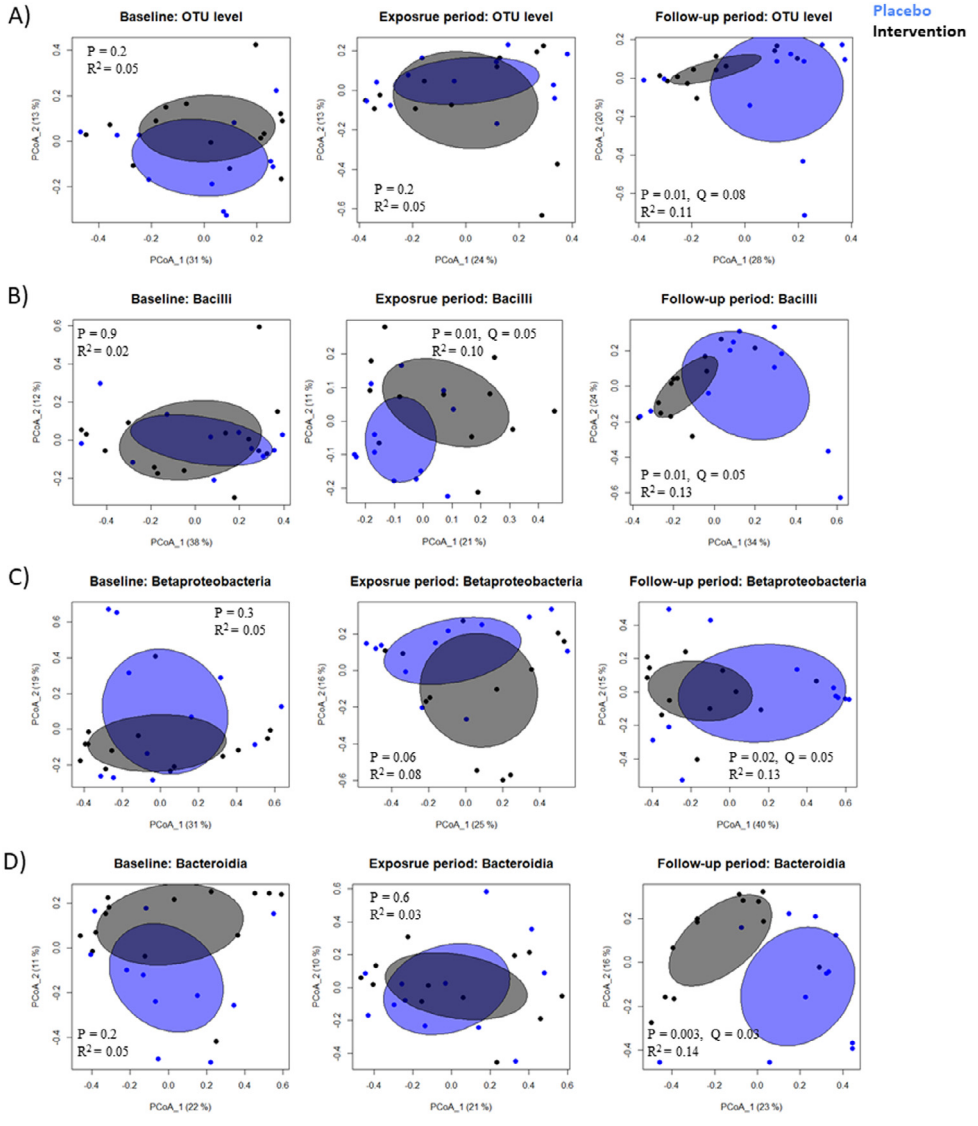

Fig. 2. Principle coordinated analysis (PCoA) for skin bacterial communities between intervention and placebo group. PCoA plots are calculated with Bray-Curtis metric for abundance data A) at OTU level, and for B) Bacilli, C) Betaproteobacteria, and C) Bacteroidia community.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:A.S., H.H., O.H.L., M.G., N.N. and S.O. have been named as inventors in a patent application 'immunomodulatory compositions' submitted by University of Helsinki (Patent application number 20165932 at Finnish Patent and Registration Office). M.G., A.P., M.I.R. and A.S., have been named as inventors in a patent application 'Immunomodulatory gardening and landscaping material' submitted by University of Helsinki

(Patent application number 175196 at Finnish Patent and Registration Office). A.S., H.H. and O.H.L are members of the board of Uute scientific LtD which develops immunomodulatory treatments.

Data Availability

Skin, gut, and sand metagenomic data tables on placebo-controlled sandbox biodiversity intervention study (Original data) (Mendeley Data)

Sandbox sand soil, human skin and stool metagenome (Original data) (Sequence Read ARchive)

CRediT Author Statement

Marja I. Roslund: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft, Visualization; **Anirudra Parajuli:** Formal analysis, Investigation; **Nan Hui:** Investigation, Writing – review & editing; **Riikka Puhakka:** Conceptualization, Methodology, Writing – review & editing; **Mira Grönroos:** Formal analysis; **Laura Soininen:** Investigation; **Noora Nurminen:** Investigation, Data curation, Writing – review & editing; **Sami Oikarinen:** Investigation, Data curation; **Ondřej Cinek:** Investigation, Writing – review & editing; **Lenka Kramná:** Investigation; **Anna-Mari Schroderus:** Formal analysis, Investigation; **Olli H. Laitinen:** Conceptualization, Writing – review & editing, Funding acquisition; **Tuure Kinnunen:** Formal analysis, Writing – review & editing; **Heikki Hyöty:** Conceptualization, Supervision, Writing – review & editing, Funding acquisition; **Aki Sinkkonen:** Conceptualization, Methodology, Validation, Supervision, Writing – review & editing, Project administration, Funding acquisition.

Acknowledgments

We thank all ADELE team members and Environmental Laboratory at University of Helsinki for their support. A special thank-you is owed to the children and families for their participation in this study. We are also very grateful to the participating daycare centers and their staff, and the city of Lahti, European Green Capital of 2021. The authors wish to acknowledge CSC – IT Center for Science, Finland, for computational resources, and FIMM Genomics NGS Sequencing unit at University of Helsinki supported by HiLIFE, and Biocenter Finland (BF) for the sequencing resources at Tampere Genomics Facility.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2023.109003.

References

[1] Schloss P.D., Westcott S.L., Ryabin T., Hall J.R., Hartmann M., Hollister E.B., Lesniewski R.A., Oakley B.B., Parks D.H., Robinson C.J., Sahl J.W., Stres B., Thallinger G.G., Horn D.J.V., Weber C.F., 2009. Introducing mothur : open-source, platform-independent, community-supported software for describing and comparing microbial communities 75, 7537–7541. doi:10.1128/AEM.01541-09.

[2] R, D. core teamR: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2018.

[3] Oksanen J., Blanchet F.G., Kindt R., Legendre P., Minchin P.R., O'Hara R.B., Simpson G.L., Solymos P., Stevens M.H.M., Wagner H., 2019. Package "vegan": community ecology package. ISBN 0-387-95457-0.

[4] D. Bates, M. Mächler, B.M. Bolker, S.C. Walker, Fitting linear mixed-effects models using lme4, J. Stat. Softw. 67 (2015), doi:10.18637/jss.v067.i01.

[5] T. Haahtela, H. Alenius, J. Lehtimäki, A. Sinkkonen, N. Fyhrquist, H. Hyöty, L. Ruokolainen, M.J. Mäkelä, Immunological resilience and biodiversity for prevention of allergic diseases and asthma, Allergy 76 (2021) 3613–3626, doi:10.1111/all.14895.

[6] I. Hanski, L. von Hertzen, N. Fyhrquist, K. Koskinen, K. Torppa, T. Laatikainen, P. Karisola, P. Auvinen, L. Paulin, M.J. Mäkelä, E. Vartiainen, T.U. Kosunen, H. Alenius, T. Haahtela, Environmental biodiversity, human microbiota, and allergy are interrelated, Proc. Natl. Acad. Sci. 109 (2012) 8334–8339, doi:10.1073/pnas.1205624109.
[7] M.I. Roslund, R. Puhakka, M. Grönroos, N. Nurminen, S. Oikarinen, A.M. Gazali, O. Cinek, L. Kramná, N. Siter, H.K. Vari, L. Soininen, A. Parajuli, J. Rajaniemi, T. Kinnunen, O.H. Laitinen, H. Hyöty, A. Sinkkonen, ADELE research group, Biodiversity intervention enhances immune regulation and health-associated commensal microbiota among daycare children, Sci. Adv. 6 (2020) eaba2578, doi:10.1126/sciadv.aba2578.
[8] N. Hui, M. Grönroos, M.I. Roslund, A. Parajuli, H.K. Vari, L. Soininen, O.H. Laitinen, A. Sinkkonen, The ADELE Research Group, Diverse environmental microbiota as a tool to augment biodiversity in urban landscaping materials, Front. Microbiol. 10 (2019) 536, doi:10.3389/fmicb.2019.00536.
[9] N. Hui, A. Parajuli, R. Puhakka, M. Grönroos, M.I. Roslund, H.K. Vari, V.A.O. Selonen, G. Yan, N. Siter, N. Nurminen, S. Oikarinen, O.H. Laitinen, J. Rajaniemi, H. Hyöty, A. Sinkkonen, Temporal variation in indoor transfer of dirt-associated environmental bacteria in agricultural and urban areas, Environ. Int. 132 (2019) 105069, doi:10.1016/j.envint.2019.105069.
[10] M.I. Roslund, R. Puhakka, N. Nurminen, S. Oikarinen, N. Siter, M. Grönroos, O. Cinek, L. Kramná, A. Jumpponen, O.H. Laitinen, J. Rajaniemi, H. Hyöty, A. Sinkkonen, D. Cerrone, M. Grönroos, N. Hui, I. Mäkelä, N. Nurminen, S. Oikarinen, A. Parajuli, R. Puhakka, M.I. Roslund, M. Saarenpää, L. Soininen, Y. Sun, H.K. Vari, O.H. Laitinen, J. Rajaniemi, H. Hyöty, A. Sinkkonen, Long-term biodiversity intervention shapes health-associated commensal microbiota among urban day-care children, Environ. Int. 157 (2021) 106811, doi:10.1016/j.envint.2021.106811.
[11] M.I. Roslund, A. Parajuli, N. Hui, R. Puhakka, M. Grönroos, L. Soininen, N. Nurminen, S. Oikarinen, O. Cinek, L. Kramná, A.-.M. Schroderus, O.H. Laitinen, T. Kinnunen, H. Hyöty, A. Sinkkonen, D. Cerrone, M. Grönroos, O.H. Laitinen, A. Luukkonen, I. Mäkelä, N. Nurminen, S. Oikarinen, A. Parajuli, R. Puhakka, M.I. Roslund, M. Saarenpää, L. Soininen, J. Rajaniemi, H. Hyöty, A. Sinkkonen, A Placebo-controlled double-blinded test of the biodiversity hypothesis of immune-mediated diseases: environmental microbial diversity elicits changes in cytokines and increase in T regulatory cells in young children, Ecotoxicol. Environ. Saf. 242 (2022) 113900, doi:10.1016/j.ecoenv.2022.113900.
[12] S.M. Huse, D.M. Welch, H.G. Morrison, M.L. Sogin, Ironing out the wrinkles in the rare biosphere through improved OTU clustering: ironing out the wrinkles in the rare biosphere, Environ. Microbiol. 12 (2010) 1889–1898, doi:10.1111/j.1462-2920.2010.02193.x.
[13] R.C. Edgar, B.J. Haas, J.C. Clemente, C. Quince, R. Knight, UCHIME improves sensitivity and speed of chimera detection, Bioinformatics 27 (2011) 2194–2200, doi:10.1093/bioinformatics/btr381.
[14] M. Grönroos, A. Parajuli, O.H. Laitinen, M.I. Roslund, H.K. Vari, H. Hyöty, R. Puhakka, A. Sinkkonen, Short-term direct contact with soil and plant materials leads to an immediate increase in diversity of skin microbiota, Microbiologyopen 8 (2019) e00645, doi:10.1002/mbo3.645.