

Comparison of lung microbiota between antineutrophil cytoplasmic antibody-associated vasculitis and sarcoidosis

Shoichi Fukui MD, PhD^{1,2,*}, Shimpei Morimoto PhD^{3,*}, Kunihiro Ichinose MD, PhD^{1,**}, Shota Nakashima MD, PhD⁴, Hiroshi Ishimoto MD, PhD⁴, Atsuko Hara MD, PhD⁴, Tomoyuki Kakugawa MD, PhD⁴, Noriho Sakamoto MD, PhD⁴, Yoshika Tsuji MD⁵, Toshiyuki Aramaki MD, PhD⁵, Tomohiro Koga MD, PhD^{1,6}, Shin-ya Kawashiri MD, PhD^{1,2}, Naoki Iwamoto MD, PhD¹, Mami Tamai MD, PhD¹, Hideki Nakamura MD, PhD¹, Tomoki Origuchi MD, PhD^{1,7}, Yuditaka Ueki MD, PhD⁵, Shino Suzuki PhD⁸, Hiroshi Mukae MD, PhD⁴, Atsushi Kawakami MD, PhD¹

¹Department of Immunology and Rheumatology, ²Department of Community Medicine,

³Innovation Platform & Office for Precision Medicine, ⁴Department of Respiratory Medicine,

⁶Center for Bioinformatics and Molecular Medicine, and ⁷Department of Rehabilitation Sciences, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan

⁵Rheumatic and Collagen Disease Center, Sasebo Chuo Hospital, Sasebo, Japan

⁸Kochi Institute for Core Sample Research, X-star, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Nankoku, Japan

*These authors contributed this work equally.

****Correspondence to:** Dr. Kunihiro Ichinose, Department of Immunology and Rheumatology,

Nagasaki University Graduate School of Biomedical Sciences, 1-7-1 Sakamoto, Nagasaki 852-

8501, Japan. Tel.: +81-095-819-7262; Fax: +81-095-849-7270

Email: kichinos@nagasaki-u.ac.jp

Key words: ANCA-associated vasculitis, bronchoalveolar lavage, lung, microbiota, sarcoidosis

Running head: Lung microbiota in AAV and sarcoidosis

Figure S1

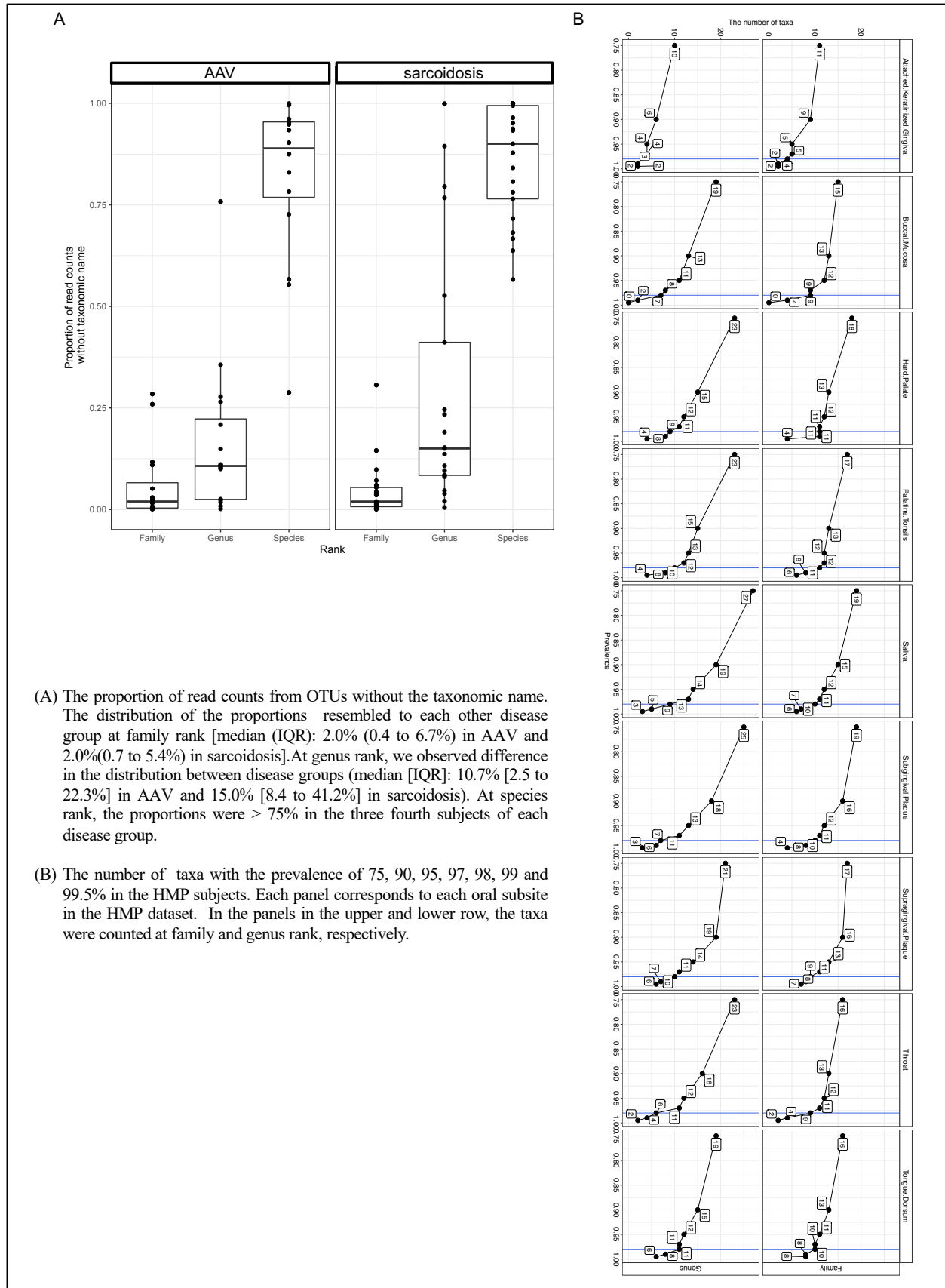


Figure S2

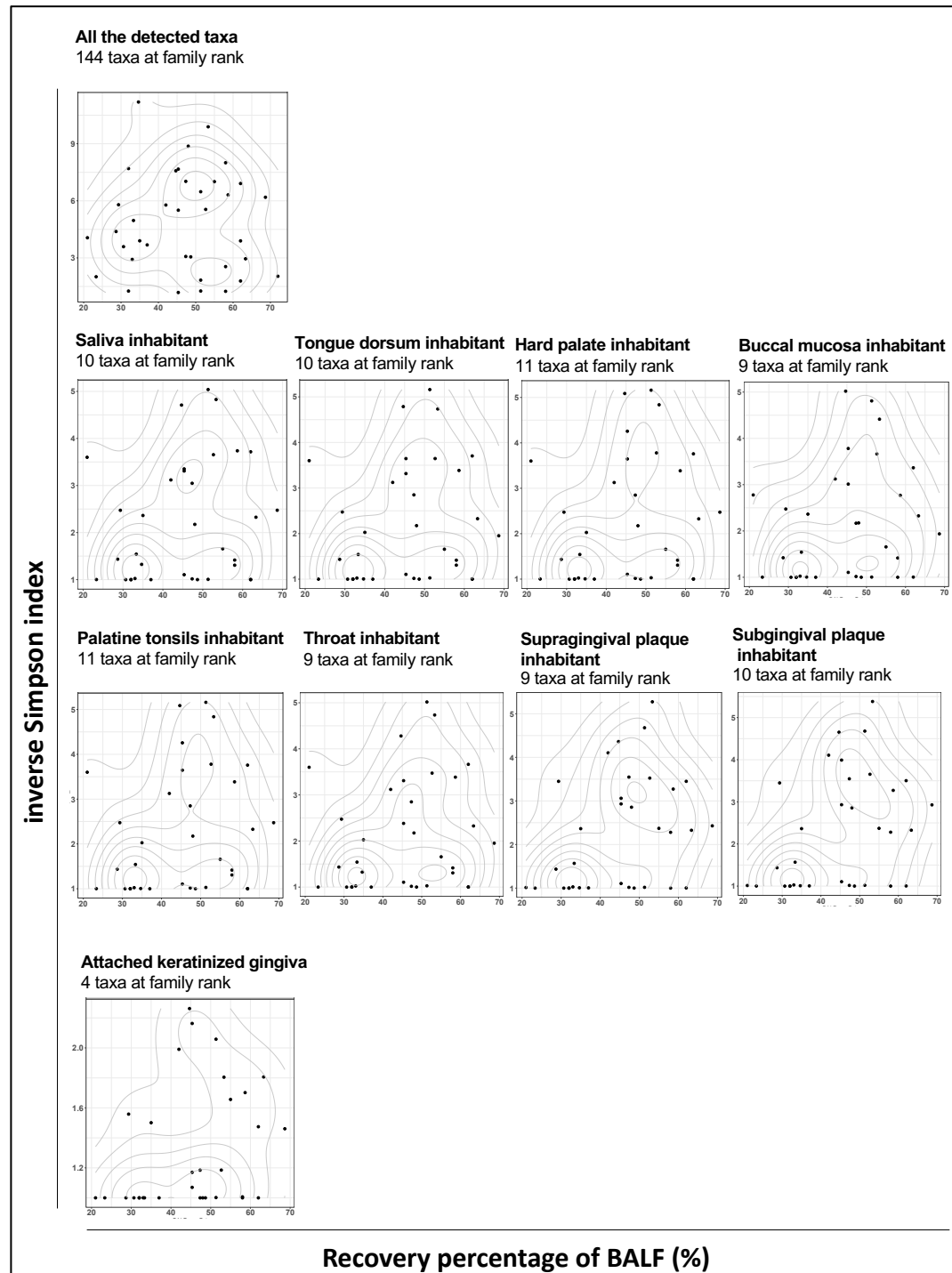


Figure S2. The bivariate relationship between the inverse Simpson index and the recovery percentage of BALF (%) evaluated for each set of taxa of inhabitant in the respective oral subsites. The contour lines are two-dimensional probability densities drawn based on corresponding kernel density estimate.

Figure S3

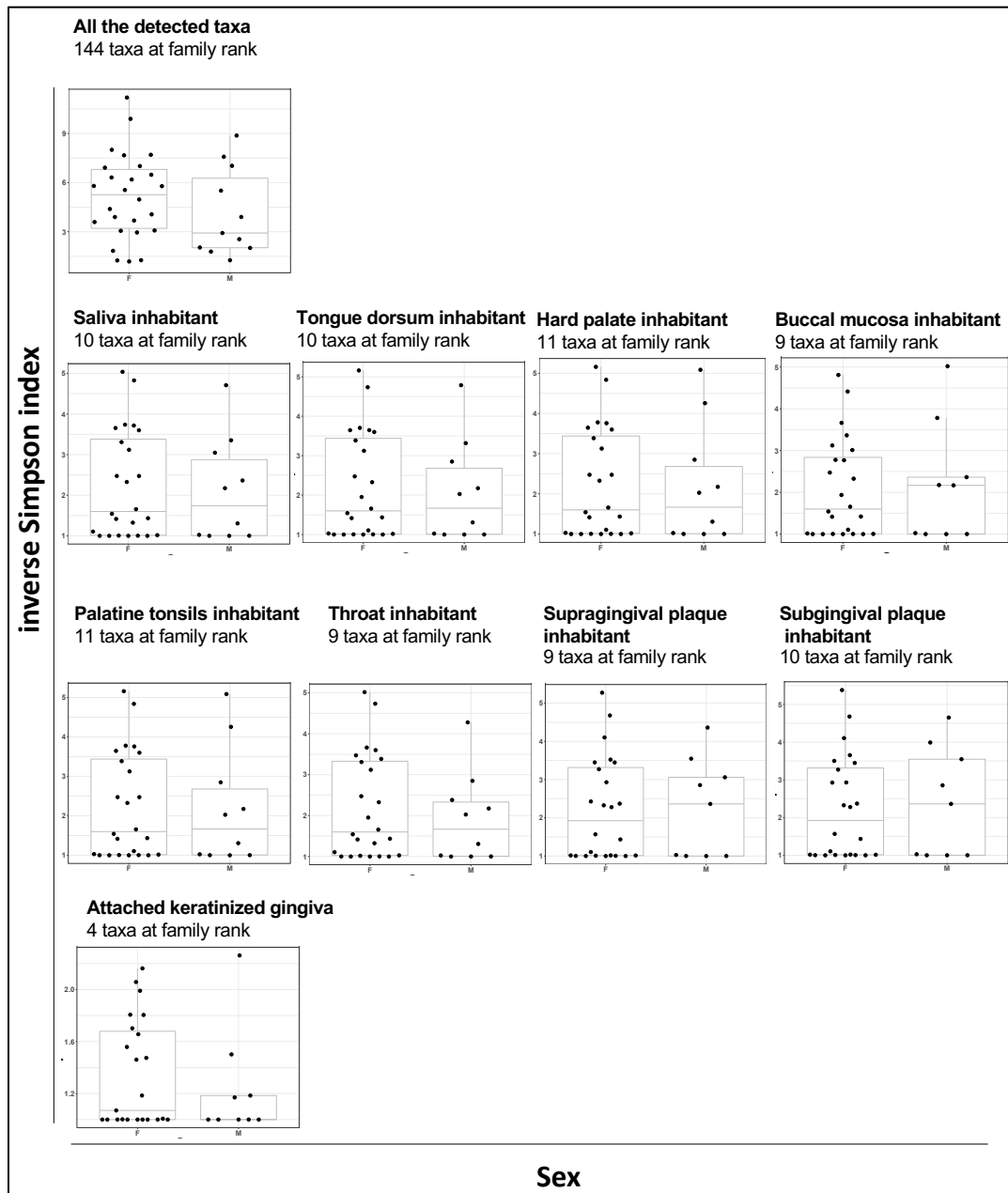


Figure S3. The difference in the inverse Simpson index between sexes evaluated for each set of taxa of inhabitant in the respective oral subsites. The width of the dots' distribution was based on corresponding kernel-density estimation with Gaussian-kernel. Each box denotes interquartile range and each horizontal line in the middle of each box denotes median of corresponding density estimates.

Figure S4

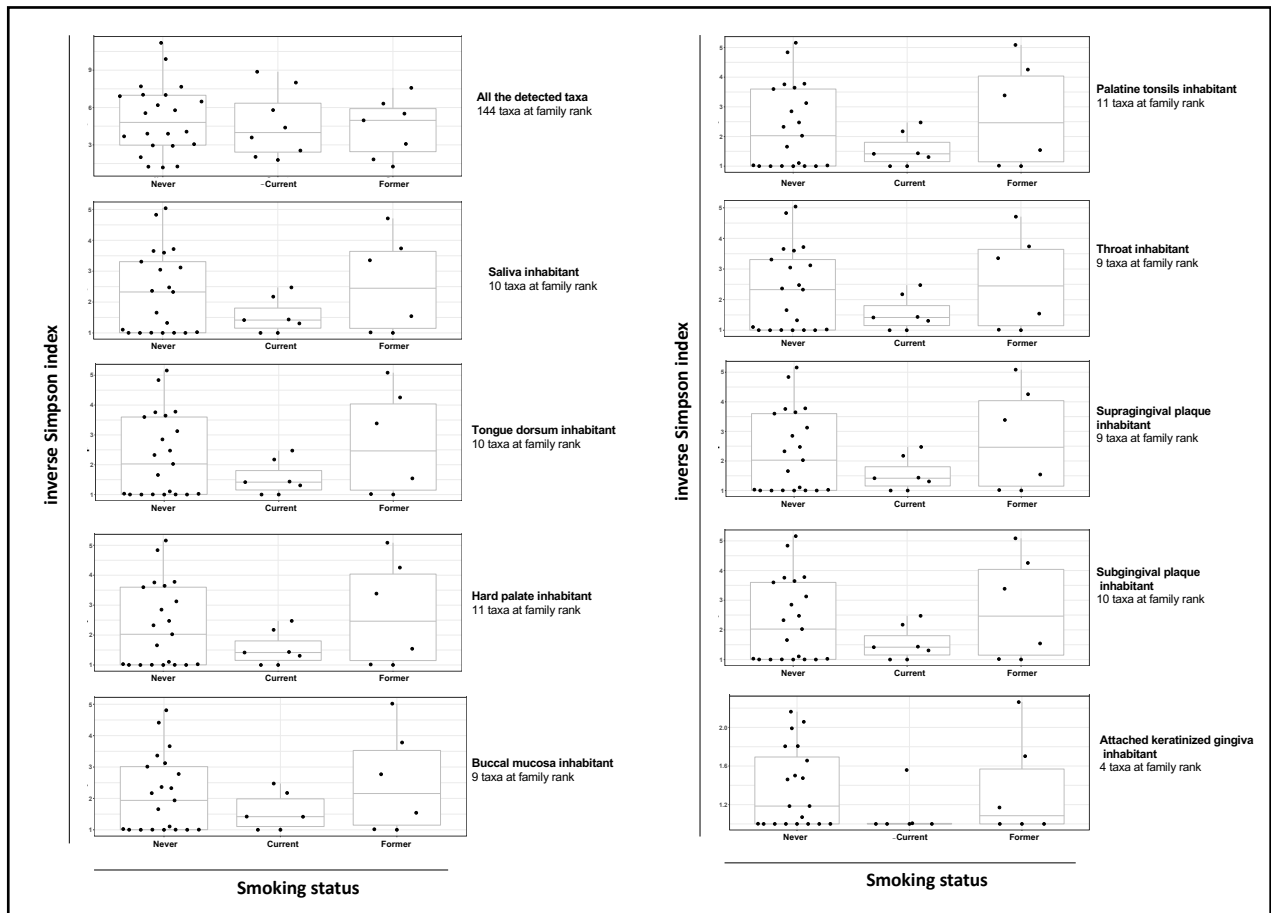


Figure S4. The difference in the inverse Simpson index between smoking status evaluated for each set of taxa of inhabitant in the respective oral subsites. The width of the dots' distribution was based on corresponding kernel-density estimation with Gaussian-kernel. Each box denotes interquartile range and each horizontal line in the middle of each box denotes median of corresponding density estimate.

Figure S5

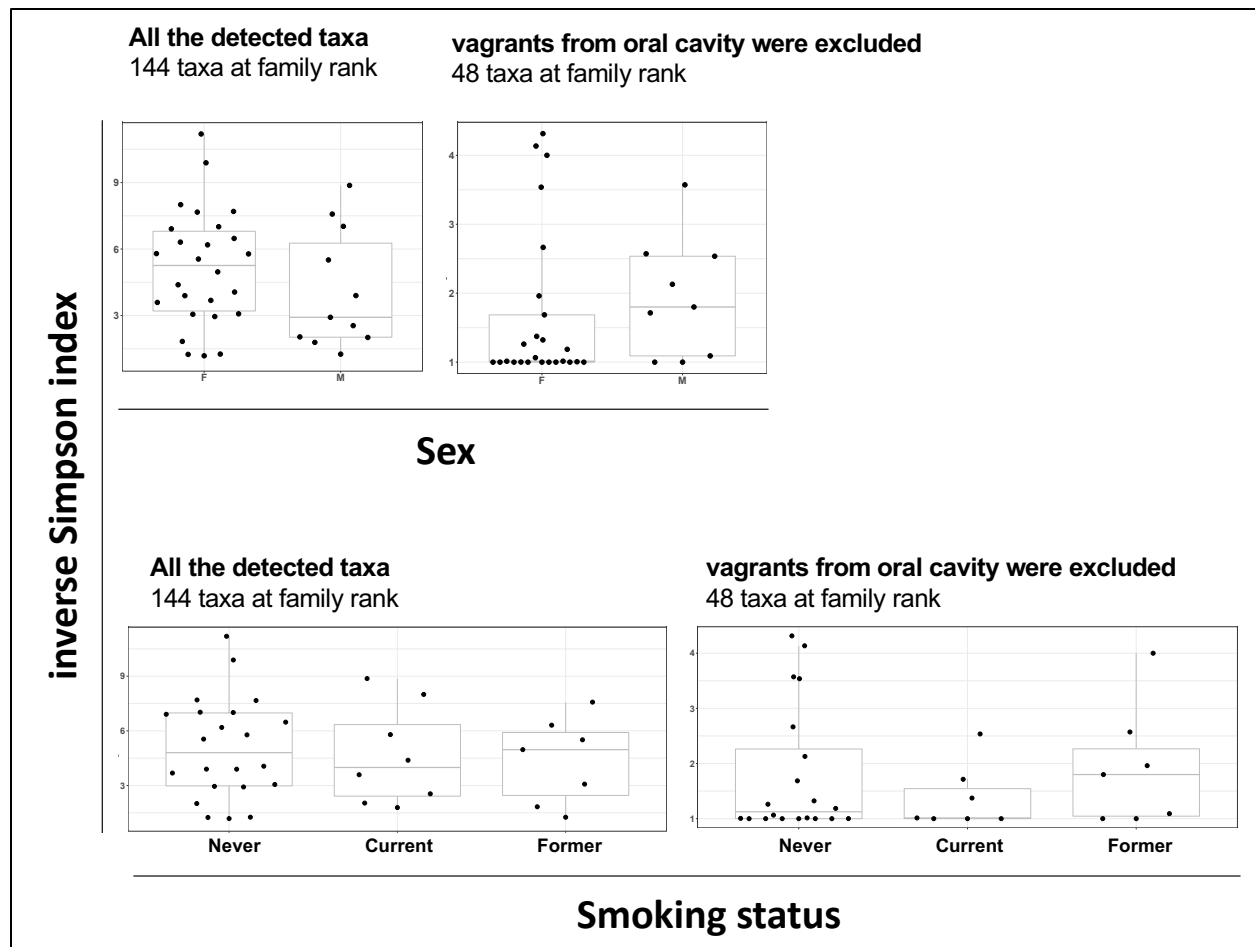


Figure S5. The difference in the inverse Simpson index between sexes and smoking status evaluated for a set of taxa from which the “vagrants” was excluded. The width of the dots’ distribution was based on corresponding kernel-density estimation with Gaussian-kernel. Each box denotes interquartile range and each horizontal line in the middle of each box denotes median of corresponding density estimate.

Figure S6. To depict the branching of the clusters more clearly, the dendrograms of the subjects in the Figure 3, 5 and 7 were represented in polar shape.

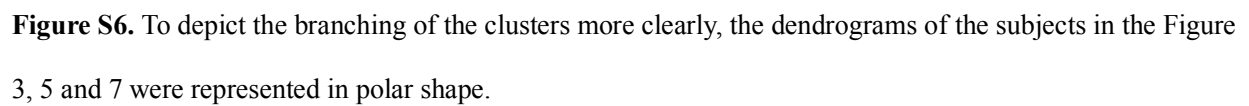


Figure S7

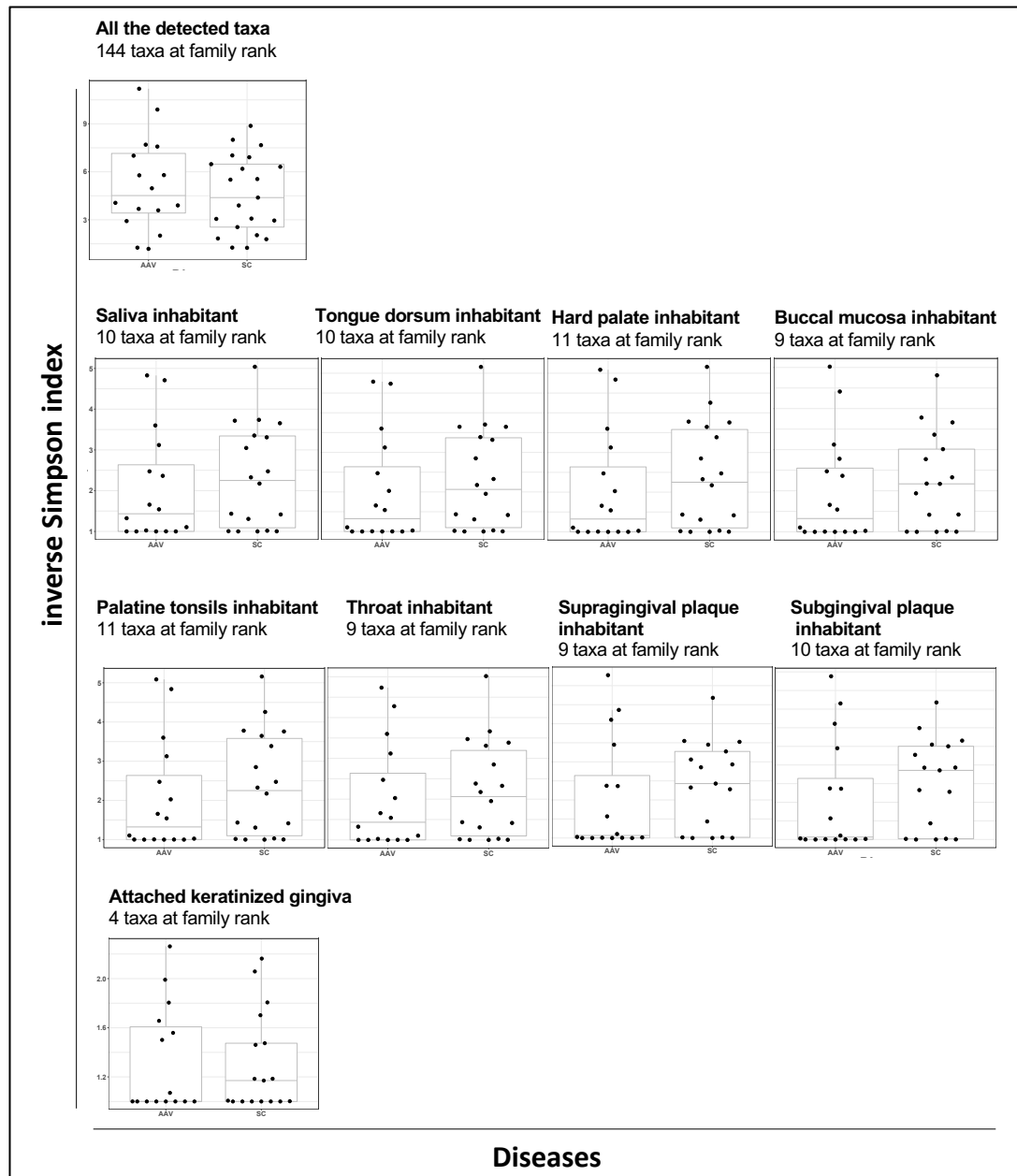


Figure S7. The difference in the inverse Simpson index between diseases evaluated for each set of taxa of inhabitant in the respective oral subsites. The width of the dots' distribution was based on corresponding kernel-density estimation with Gaussian-kernel. Each box denotes interquartile range and each horizontal line in the middle of each box denotes median of corresponding density estimate.

Figure S8

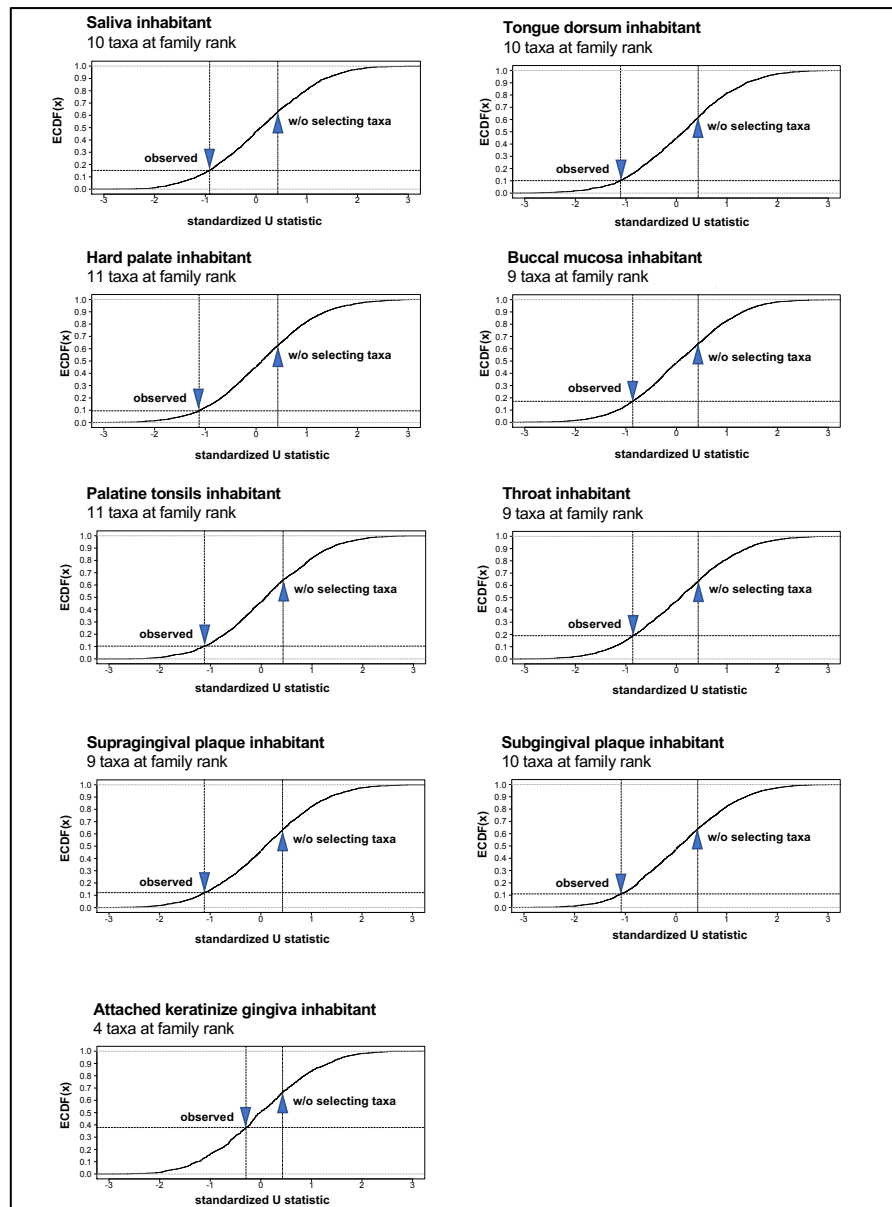


Figure S8. The effect of selecting respective sets of taxa of “inhabitant” contrasted with selecting them in random drawing on the association between the α -diversity and the diseases. The number of the taxa in random drawing were set to the number of taxa in corresponding “inhabitant” for respective bodysites. The standardized U statistics were calculated for the difference in the inverse Simpson index between diseases. The two blue triangles depict the standardized U statistics calculated from the inverse Simpson index for the taxa of “hard palate inhabitant” (“**observed**”) and all of the detected taxa (“**w/o selecting taxa**”), respectively.

Figure S9

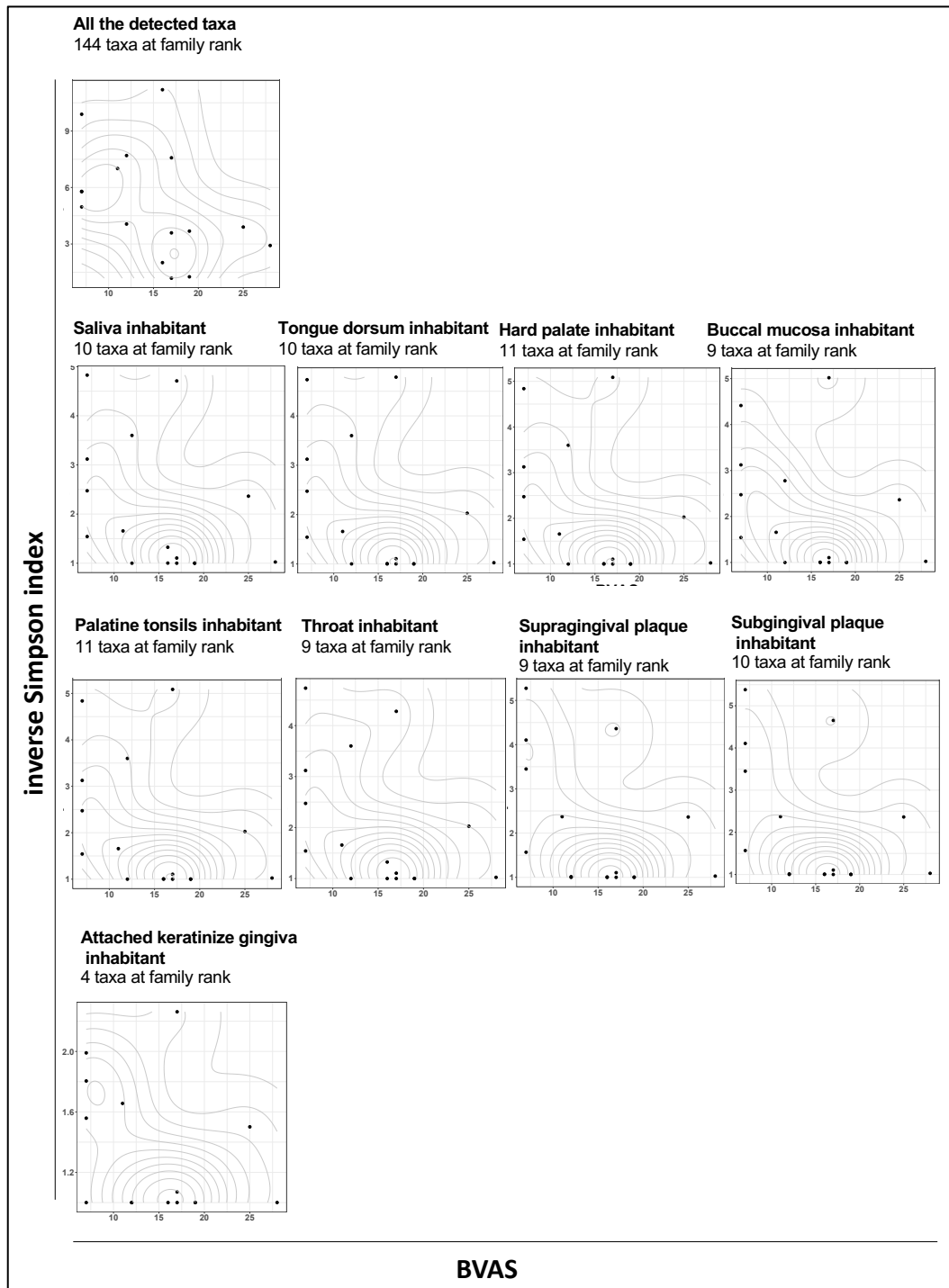


Figure S9. The bivariate relationship between the inverse Simpson index and the BVAS for each set of taxa of inhabitant in the respective oral subsites. The contour lines are two-dimensional probability densities drawn based on corresponding kernel density estimate.

Figure S10

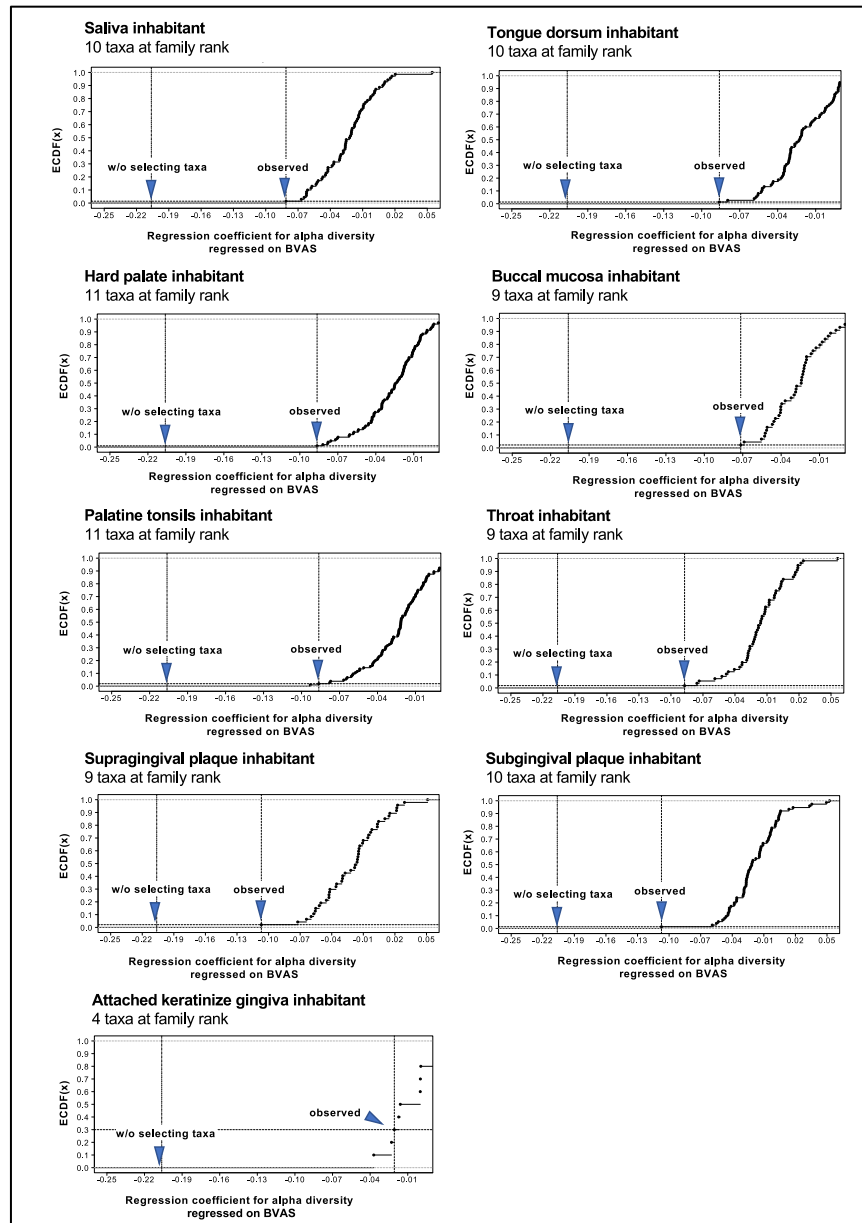


Figure S10. The effect of selecting respective sets of taxa of “inhabitant” contrasted with selecting them in random drawing on the association between the α -diversity and the BVAS. The number of the taxa in random drawing were set to the number of taxa in corresponding “inhabitant”. The standardized U statistics were calculated for the difference in the inverse Simpson index between diseases. The two blue triangles depict the standardized U statistics calculated from the inverse Simpson index for the taxa of “hard palate inhabitant” (“observed”) and all of the detected taxa (“w/o selecting taxa”), respectively.