The salivary microbiome shows a high prevalence of core bacterial members yet variability across human populations

Xinwei Ruan[1], Jiaqiang Luo[1], Pangzhen Zhang[1] and Kate Howell[1, *]

[1]School of Agriculture and Food, Faculty of Veterinary and Agricultural Sciences, University of Melbourne, Parkville 3010, Australia

Xinwei Ruan: ORCID: 0000-0002-1963-3807

Jiaqiang Luo: ORCID: 0000-0001-6459-3309

Pangzhen Zhang: ORCID: 0000-0002-9794-2269

* Corresponding author. khowell@unimelb.edu.au; ORCID: 0000-0001-6498-0472

Abstract

Human saliva contains diverse bacterial communities, reflecting human health status, dietary patterns and contributing to variability in the sensory perception of food. Many descriptions of salivary microbiome diversity compare commonalities and differences with reference to a diseased state, but the composition of healthy saliva has not been described. Here, we use a meta-analysis approach to define and explore the core membership of the human salivary microbial community by collecting and re-analysing raw 16S rRNA amplicon sequencing data from 47 studies with 2206 saliva samples. We found 68 core bacterial taxa that were consistently detected. Differences induced by various host intrinsic and behaviour factors, including gender, age, geographic location, tobacco usage, and alcohol consumption, were evident. The core of the salivary microbiome was verified by collecting and analysing saliva in an independent study. These results suggest that the methods used can effectively define a core microbial community in human saliva with high prevalence. The core salivary microbiome demonstrated both stability and variability among populations. Geographic location was identified as the host factor with the largest effect in shaping salivary microbiota. The independent analysis underlined that the impact of geographic variation is likely due to diet.

Keywords: Microbiome, Oral microbiome, Westernization, Dietary pattern

Introduction

Human saliva plays an essential role in influencing the sensory perception of foods and beverages and driving the purchase decisions of consumers. When food is taken into the mouth, mixing and mastication allow a semi-solid bolus to be formed, and at the same time, aroma and flavour to be released from the food [1]. The variation in perceived responses can be attributed to the inter-individual variability in salivary composition, especially salivary microorganisms. As a complex ecosystem, the human oral cavity hosts thousands of bacterial taxa, interacting with themselves, other microorganisms and the broader cavity [2]. It is an ecological system that contains many distinct sub-niches, including saliva, dental plaques, gingival sulcus, epithelial cells on the cheek, tongue, and teeth [3]. High heterogeneity has been reported between the composition of microbial communities that colonise on different sites [4]. Saliva is recognised as a reservoir of microorganisms from all ecological niches in the human mouth with long-term stability [5]. The ensemble of microorganisms and the expressed genetic material in human saliva; the 'salivary microbiome' is crucial to understanding the healthy state of humans, and indeed, what this means for sensory appreciation of foods and beverages.

The contribution of salivary microbiome to sensory perception of foods has been described by various studies [6-8]. However, the diverse conclusions suggest that the role of the salivary microbiome may be confounded by inter-individual variance. Characterising the microbial communities commonly found in most human saliva regardless of the study-specific variation could help establish the connection between salivary composition and food preference. Comparing results from different studies introduces significant technical and bioinformatic biases [9] especially when studies have targeted different 16S rRNA hypervariable regions for amplification [10]. On this basis, the shifts from a common salivary microbiome by diseases or host lifestyle factors will also be more prominent. A meta-analysis of both sequencing data and associated meta-data can summarise the existing knowledge and identify the commonalities and differences in salivary microbiota between people from various backgrounds.

Attempts to define the core oral microbiome of a healthy human has occurred during the last ten years [11]. The core microbiome is described as the common group of microbes that are important for host biological function [12]. It provides a foundation for prioritising members adapted to the host environment [13]. Although the shift in human salivary composition caused by diseases has been studied for decades [14-19], our understanding of the impact of host intrinsic and behaviour factors is still limited.

Many host characteristics have been shown to have an impact on the composition of the salivary microorganisms, including age [20], diet [21, 22], ethnicity [23], gender [24], smoking [25], alcohol use [26], circadian rhythm [27], body mass index [28], and the type of stimulation [29]. Some studies have correlated the diverse microbiome with the distinct sensory responses between consumer groups [30, 31] It has also been reported that people from different countries are colonised with distinct salivary bacterial communities [32]. Li et al. analysed the human oral microbiome from Africa, Alaska, and Germany and reported differences between the human groups living in various climate conditions [33]. However, a clear global pattern in salivary microbial composition is lacking, which considers both core microbial populations and variability amongst different groups.

In this paper, we collected raw 16S rRNA sequences of human salivary microbiota from 47 publicly available datasets spanning 15 different countries. These raw data were systematically re-analysed and pooled together to define a core salivary microbiome. We classified all sequences into operational taxonomic units (OTUs) at 97% identity against the Human Oral Microbiome Database (HOMD) to minimise the technical variation induced by comparing data from different hypervariable regions. It allowed us to make a comparison between studies and reduced the redundancy in the dataset for defining the “core” microbial members. Using the metadata acquired with raw data, we also investigated the influences of several host factors and technical factors on human salivary microbiota. Factors that showed a potentially strong impact on shaping microbial communities in saliva were selected, and the taxa as potential biomarkers were identified, and linked with functional predictions.

Finally, verification of the meta-analysis was sought through a targeted sampling effort. Here, saliva samples were collected from 30 healthy individuals and analysed for microbial composition to compare against the results found from the global dataset. These data confirmed the composition of the core microbiome members, but the verification of ethnic origin was not possible. Our study contributes to fundamental understandings of the stable and differential salivary microbiome across healthy adult populations. We have identified bacteria linked to particular identities of participants and points to salivary microbiome composition being linked to diet, rather than ethnic origin.

Materials and methods

1. Literature search and data collection

To acquire sufficient data from healthy human saliva, available public studies related to human salivary microbiota were systematically reviewed. A literature search was performed using the combination of relative terms in EMBASE, MEDLINE and Web of Science for the studies published before November 2020 using the terms described in supplementary data (Table S1). A supplementary dataset search in NCBI's Sequence Read Archive (SRA) was also performed using the search term 'salivary microbiome'. The resulting studies were screened according to the criteria described in supplementary material. Raw sequence data acquired from the healthy individuals of selected studies were downloaded from SRA and European Nucleotide Archive (ENA), using SRA Toolkit. The files were converted to the FASTQ formats if necessary. Sequence data from each selected study were processed separately using QIIME2 (version.2020.2) [34]. The resulting OTU table was exported into BIOM format. Further analyses were carried out in R (version 4.1.0) with custom scripts as detailed in supplementary material.

2. Defining the core microbiome

The core microbiome was determined based on the abundance-occupancy pattern. Two methods adapted from Shade and Stopnisek [35] and Wu et al. [13] were used. For the methods adapted from the study of Shade and Stopnisek [35], samples were rarefied to 5 000. Both Bray-Curtis similarity

121 and weighted Unifrac distance were used to determine the contribution in the percentage of the
122 prospective core set to the overall beta diversity. For the method adapted from Wu et al. [13], OTUs
123 were filtered out with the mean relative abundance (MRA) bigger than 0.1% and the presence in more
124 than 75% of samples or 100% occupancy in more than 10 studies. To investigate the bacteria-bacteria
125 interactions in salivary microbial communities, the co-occurrence network was constructed using
126 pairwise Spearman's correlation based on relative abundance. The Spearman's correlation was
127 calculated using the *rcorr* function in the *Hmisc* R package [36] and visualised by Cytoscape v3.8.2
128 [37]. A correlation with Spearman's correlation coefficient > 0.5 or < -0.5 and p-value < 0.01 is
129 considered as statistically robust and shown in the network.

130 **3. Functional prediction**

131 Microbial metagenomes were inferred from 16S rRNA gene-based bacterial profiles, and the
132 functional prediction were conducted based on Kyoto Encyclopedia of Gene and Genomes (KEGG)
133 database [38] using the default pipeline in Phylogenetic Investigation of Communities by
134 Reconstruction of Unobserved States 2 (PICRUSt2) (Douglas et al., 2020). The ANCOM-BC analysis
135 was used to identify the differential abundant KEGG pathways by geographic location, adjusting for
136 hypervariable regions. At the same time, a random forest model was established for distinguishing
137 Chinese and Western samples, and the importance of pathways was measured using mean decreased
138 accuracy. Spearman's correlation was performed to assess the relationship between the relative
139 abundance of differential pathways and genera. The significant correlations were visualised using the
140 *corrplot* package [39] in R.

141 **4. Comparison between Chinese and Western people on an independent cohort**

142 Saliva samples were collected from 26 participants (aged 20-60 years) recruited for a wine assessment
143 experiment and consisted of 13 Chinese and 13 Western wine experts (Table 1). The study was
144 approved by the Office for Research Ethics and Integrity of the University of Melbourne (Ethics ID:
145 1852616). Each group had six female panellists and seven male panellists. The Western panellists
146 were defined as people who have lived in Australia for more than ten years. Chinese panellists were
147 defined as people who were born in China and had lived in Australia for no more than 18 months.

148 Bacteria genomic DNA was extracted from human saliva using QIAGEN® MagAttract® PowerSoil®
149 DNA KF Kit [40] and subjected to 16S rRNA amplicon sequencing on the Illumina platform
150 following the Earth Microbiome Project protocols (https://earthmicrobiome.org/protocols-and-
151 standards/16s/). The raw data are available in NCBI Sequence Read Archive, with accession number
152 PRJNA786805.
153
154 The raw sequences were processed using the same pipeline in the meta-analysis as described in
155 section 2. Additionally, raw sequencing reads were denoised into zero-radios OTUs (zOTUs) by
156 UNOISE3 pipeline [41], and taxonomically classified by classifiers trained on the full-length 16S
157 rRNA gene SILVA v138 [42] database and eHOMD v15.22 [43], respectively. The affiliation
158 between each ZOTUs and the originating OTU was determined using a customised code adapted from
159 Stopnisek and Shade [44] and available at
160 https://github.com/XINWEIR/SalivaryMicrobiome_MetaAnalysis. The relative abundance of taxa at
161 the genus level in this cohort was used as the test set for the random forest model trained using the
162 genus-level OTU assignment information in the meta-analysis.

163 Results

164 1. Inclusion of studies and sequences

165 In this study, we extracted the 16S rRNA gene amplicon sequencing data of healthy human saliva
166 from 47 studies (Figure 1A) [15, 16, 48, 67-110]. We abstracted data from subjects who had no
167 diagnosed disease state, hereafter named "healthy". Of course, subjects could have had subclinical
168 diseases or may have altered health status for undisclosed reasons, but we considered that this would
169 be true of the wider human population and therefore able to be included in our study. A total of 107
170 005 868 high quality 16S rRNA sequences were obtained. After removing all samples below 2,000
171 reads, 2206 samples with 909 features were retained. The retained samples included studies from 15
172 countries (Table S2). Most studies were conducted in three geographic regions: North America,

173 Europe, and China (Figure 1B). Most sequences included in this meta-analysis were generated from
174 the hypervariable region "V3-V4" and "V4" (Figure 1C).

175 **2. Intrinsic and lifestyle factors have a significant effect on the host salivary microbiome**

176 Large variability between studies was observed in the number of reads, taxonomic profile, and alpha
177 diversity (Figure 2; Figure S1). Phylum *Bacteroidetes*, *Proteobacteria*, *Firmicutes*, *Fusobacteria* were
178 dominated among all studies, while their proportion varies (Figure 2A). When studies were grouped
179 by the geographic locations they originated from (coloured in Figure 2B, C, D), there is generally no
180 difference between their intra-community diversity, represented by Shannon, Chao1, and Simpson
181 indices. Only one study conducted in Qatar showed relatively lower Chao 1 index and higher
182 Simpson's diversity indices than studies from other locations. However, it is hard to decide whether
183 such variation is caused by the geographic location or other technical variations.
184
185 Because of the large disparity of methodologies amongst the studies used in our global analysis, we
186 applied several different strategies for normalisation as described in Supplementary Methods. When
187 investigating the influences of different categories using permutational multivariate analysis of
188 variance (PERMANOVA) tests, these normalising methods were combined as appropriate with
189 different distance metrics, including Bray-Curtis, weighted UniFrac, and Euclidean distance. Overall,
190 the effect of rarefaction (RAR), total-sum scaling (TSS), and the rarefied relative abundance
191 transformation (RRA) were very similar in the result of PERMANOVA (Figure 3A-D; Table S4).
192 Meanwhile, the centred log-ratio transformation (CLR) enlarged the variance induced by an unwanted
193 technical factor, namely the amplified hypervariable region, at all taxonomic levels.
194 The beta-diversity analyses showed that all metadata categories measured have a significant ($p<0.001$)
195 effect on the bacterial profile of human saliva at all taxonomic levels, adjusted for the study effect
196 (Table S4). However, only limited variation among samples has been explained by these factors ($R2 <$
197 10%). In contrast, "study" accounts for around 35% of the variability between samples. At the OTU
198 level, the combination of weighted UniFrac distance with two normalisations (TSS and RRA) best
199 minimised the variability raised by different hypervariable regions (Figure 3D). The results of the

200 unconstrained principal coordinate analysis (PCoA) are in agreement with the results of
201 PERMANOVA. When using Bray-Curtis dissimilarity and Euclidean distance, the samples separated
202 distinctly according to the hypervariable regions in PCoA plots, whereas the plot constructed using
203 the weighted UniFrac revealed the clusters formed by the samples from different geographic locations
204 (Figure S2). A distinct separation of samples from three main geographic locations (Figure 3F), with
205 more than half (58.0%) of the variance explained by the first two dimensions, using weighted UniFrac
206 distance. In contrast, the differences between locations were confounded by which hypervariable
207 regions were sequenced in the PCoA plot for Bray-Curtis dissimilarity (Figure 3E; Figure S2D-F).
208 The results suggest that host intrinsic and lifestyle factors significantly influence the microbial profile
209 in human saliva, regardless of the variation induced by technical factors.

210 **3. A core microbiome is defined from saliva from healthy humans**

211 Despite the large intra- and inter-study variability, many OTUs still showed a consistently high
212 presence and relative abundance across studies (Figure S3). These persistent OTUs detected across
213 studies with different protocols could be functionally important for the salivary microbiome of healthy
214 adults. We wanted to identify the most widespread microbial taxa within a specific population that
215 allows us to better understand the broad structure of microbiomes and their potential functional
216 consequences [12]. The abundance and occurrence frequency of taxa are two important criteria used
217 to define the "core salivary microbiome". Conventionally, thresholds on these two parameters filter all
218 taxa detected, and taxa that meet both criteria can be classified as the core salivary microbiome [13].
219 Recently, a more standardised procedure based on abundance-occupancy distribution was proposed
220 [35]. We employed both strategies to define the core salivary microbiome to identify the microbial
221 features with high persistence and robustness in human saliva. Considering the sequences involved in
222 this meta-analysis were collected from studies targeted at different hypervariable regions, close-
223 referenced clustering at 97% identity were used to cluster sequences into OTUs. In addition, taxa
224 defined at 100% sequence identity may increase the redundancy in the dataset [35]. Therefore, the
225 core salivary microbiome was defined using the clustered OTUs at 97% identity. To begin, the core
226 OTUs were determined by filtering all OTUs based on mean relative abundance and occurrence

227 frequency using the criteria described in the Methods [13]. In total, 11.6% of all OTUs (105 OTUs)
228 were included as Core 1 (Figure 4A: MRA + OCC; Table S5). Meanwhile, according to the method
229 proposed by Shade and Stopnisek [35], OTUs were ranked depending on their occupancy across
230 studies, and the contribution of top-ranked OTUs to beta-diversity was expressed by Bray-Curtis
231 similarity and weighted UniFrac distance. Two groups of the core microbiome were prioritised by
232 these two indices, to give different inclusions in the core, consisting of the top 69 OTUs (using
233 weighted UniFrac; Figure 4A: BC) and 94 OTUs (using Bray-Curtis; Figure 4A: wUF) OTUs (Figure
234 S4).
235
236 Overall, sixty-eight OTUs were shared across all three methods (Figure 4A), accounting for 7.5% of
237 all OTUs detected and 72.5% of all 16S rRNA gene sequences after clustering and filtering (Figure
238 4B, C). *Firmicutes* account for nearly half (46.4 %) of all core OTUs, while only one OTU belongs
239 to *Saccharibacteria*. The mean relative abundance (MRA) of each OTU in sub-groups classified by
240 different factors was also measured (Figure 4D). On average, the core OTUs were highly prevalent
241 (73.2% ± 3.4% of cumulative relative abundance) in saliva samples across different levels in
242 subgroups classified by age, gender, geographic locations, hypervariable regions, sample type,
243 smoking, and drinking habits. The core OTUs were clustered into four main groups based on their
244 distribution pattern in sub-groups (Figure 4E). The eight OTUs affiliated to Cluster 1 showed overall
245 high abundance in all sub-groups. Cluster 2 consists of core OTUs with a slightly lower mean relative
246 abundance than Cluster 1 and higher intra-group variability. Notably, although having a higher MRA
247 than some members of Cluster 1, "476_9291" was still classified as Cluster 2. The reason could be its
248 biased presence in sub-groups. For example, the relative abundance of "476_9291" is higher in
249 samples from China than other two locations. The other two clusters contain OTUs with lower MRA
250 than Cluster 1 and 2, while variations can still be observed within sub-groups.
251
252 We further applied a network analysis built by Spearman's correlations to investigate whether the
253 core OTUs defined were also important to the structure of a co-occurrence pattern. The resulting co-
254 occurrence network contains 293 nodes and 1,424 significant correlations (edges) (Figure S5).

255 Although most of the core OTUs have relatively low connectivity, they associate with each other
256 rather than with rare OTUs. Nine OTUs were identified as potential "hub" OTUs based on their
257 centrality and the number of links in the network (Figure S5, 6), and as such to community stability.
258 Two core OTUs were identified as hub taxa; "122BU057" (*Megasphaera micronuciformis*) and
259 "524_3631" (*Veillonella atypica*). Compared to other "hub" taxa, they showed lower connectivity and
260 relatively high betweenness centrality (Figure S6).

261 **4. Geographic location is the host factor with the largest impact on bacterial composition**

262 To investigate which metadata category has the largest impact on the salivary microbiome, we
263 established random forest models to link the seven categories described above and a new category,
264 study, with the salivary microbiota data at seven taxonomic levels (OTU, phylum, class, order, family,
265 genus, and species). The effect of four normalisation methods was compared using the error rate
266 generated by random forest classification. In total, 224 random forest models were constructed
267 (Figure 5A). Among four normalisation methods, total-sum scaling produced the models that were, on
268 average, the most accurate. Generally, the random forest models built with microbial communities at
269 OTU levels have the lowest error rate (mean = 13.0%), while the models constructed at phylum levels
270 have the highest (mean =26.8%). The model built with the hypervariable region used for sequencing
271 was also the category that showed the lowest error rate (Figure 3G). Geographic locations
272 demonstrated the second important impact on the bacterial communities, with the lowest error rate
273 among biological factors. The random forest model constructed by the other two categories, sample
274 type and tobacco usage, also showed a relatively lower error rate than other categories. Study
275 constructed the models with high error rates at phylum (44.3% ± 2.5%) and class level (29.3% ±
276 2.1%). However, the error rate of models built with study rapidly dropped with the increase of
277 taxonomic levels, reaching 10.5% ± 5.1% at OTU level. Gender and age range led to poorly
278 performing models.
279
280 We wanted to determine whether the defined core microbiome could be used as biomarkers to
281 differentiate people categorised by intrinsic and lifestyle factors. The random forest models showed

282 high accuracy at OTU level were used (i.e., geographic location and smoking factors). The differential
283 OTUs induced by hypervariable regions and sample types were also analysed to exclude the influence
284 of technical factors. We further performed ten-fold cross-validation five times to measure the
285 importance of OTUs used to train the model. All OTUs before the point that the cross-validation error
286 curve starts to stabilise were defined as important OTUs. In total, we defined 59, 57, 34 and 70
287 important OTUs as biomarkers to differentiate samples according to geographic location, smoking
288 habit, hypervariable region, and sample type, respectively (Figure S7). Of these, 28, 10, 13 and 22
289 biomarkers were also classified as "core" (Figure 5B; Figure S8). Although 31 core OTUs showed the
290 importance in discriminating samples according to geographic location and smoking, nearly half of
291 them (15 OTUs) had the possibility of being confounded by technical factors (Figure 5B). After
292 excluding the OTUs that could be influenced by other factors, core OTU "322AK152" (*Bergeyella*
293 *sp.HMT_322*) was the OTU with the highest contribution to the classification of samples from three
294 geographic locations. Meanwhile, "122BU057" (*Megasphaera micronuciformis*) showed the highest
295 importance among the core differential OTUs specific to smoking, followed by "524_3631"
296 (*Veillonella atypica*). We were surprised to find that these two OTUs were the only two OTUs that
297 were defined as hub taxa in the co-occurrence network analysis (Figure S6).

298 **5. The salivary microbiota as biomarkers to differentiate Chinese and Western**
299 **participants**

300 We further analysed the changes caused by geographic locations in higher taxonomic hierarchies,
301 where many differences have been revealed. Of particular interest were taxa under
302 phylum *Synergistetes* and *Spirochaetes*, Class *Mollicutes* and *Betaproteobacteria*,
303 Family *Clostridiales*, and genus *Prevotella*. Interestingly, many taxa showed a higher relative
304 abundance in the Chinese samples, both compared to North American samples and compared to
305 European samples (Figure 6A). It suggested that the variance induced by geographic locations may be
306 dominated by the differences between samples from Chinese and Western people. Therefore, we
307 combined samples from North America and Europe into a single group, "Western." Compared to the
308 Chinese grouping, the Western group has significantly lower within-sample diversity (alpha-diversity)

309 (Wilcoxon rank-sum test, $p < 0.001$; Figure 6B, C). Next, we examined the differences between
310 Chinese and Western in the salivary microbiota at the genus and species level (Table S6, S7). Besides
311 establishing a random forest model, we also identified differential taxa using ANCOM-BC, adjusting
312 for the hypervariable region. We found 48 genera identified as significantly different by both methods
313 (Figure 6D, Table S6).
314
315 Finally, we performed the functional prediction-based 16S rRNA gene profiles to investigate whether
316 differences in the salivary microbiota between Chinese and Western affect its function. Two methods,
317 ANCOM-BC, and random forest model were used to identify which pathways were differential
318 between Chinese and Western. The result of ANCOM-BC indicated that 69 pathways related to
319 metabolism were differentially abundant between the two groups. The random forest classification
320 model established using KEGG pathways demonstrated an error rate of 10.01% and revealed 46
321 differential pathways. Among them, thirty pathways belonging to nine upper pathways (level 2) were
322 simultaneously defined by two methods as differing in abundance between Chinese and Western
323 (Figure 6E, Table S8). A variety of pathways was in higher abundance in Chinese samples. The
324 enrichment of these pathways in Chinese samples was mainly associated with the increased
325 abundance of *Neisseria* and *Lautropia* and the depleted abundance of *Prevotella*, *Veillonella*,
326 and *Atopobium*. Notably, three lipid metabolism pathways enriched in Chinese samples, including
327 "Ether lipid metabolism" (ko00565), "alpha-Linolenic acid metabolism" (ko00592), and "Linoleic
328 acid metabolism" (ko00591), have the highest standardised effect size (W statistics, Table S8). The
329 enrichment of these pathways related to lipid metabolism has been positively associated with the
330 higher abundance of *Neisseri*a in Chinese. *Neisseria* may have also contributed to the pathway
331 "Carotenoid biosynthesis" (ko00906). Another metabolic pathway related to the metabolism of
332 terpenoids and polyketides, "Sesquiterpenoid and triterpenoid biosynthesis" (ko00909), showed a
333 positive correlation with a genus belong to *Peptostreptococcaceae*. In contrast, only one pathway
334 named "Flavone and flavonol biosynthesis" (ko00944) was enriched in the saliva samples from

335 Western. A strong positive correlation has been demonstrated between this pathway and the increased
336 abundance of *Veillonella* in the samples from the Western grouping.

337 **6. Validation of the core in an independent Australian cohort**

338 To validate the prevalence of the core OTUs in human saliva, we collected saliva samples from 13
339 Chinese and 13 Western participants in Melbourne and sequenced the extracted DNA with 515F-
340 806R primers. In total, 841,188 high-quality 16S rRNA sequences were obtained, which clustered into
341 397 OTUs with 97% identity to the HOMD database. Among them, the core OTUs we defined in the
342 meta-analysis showed high relative abundance (78.3 ± 6.9%) in all collected samples. To increase the
343 accuracy of the OTU assignment, we denoised the sequencing reads using the UNOISE3 pipeline and
344 generated ZOTUs with 100% sequence identity. After re-clustering the core OTUs defined in the
345 meta-analysis to ZOTUs, we observed that 59 of the identified OTUs in this independent dataset
346 consisted of 87 ZOTUs, and thus made up close to 80% of the relative abundance (Figure 7A).
347 Although some sequences belonging to the same ZOTU are clustered to different OTUs, all the core
348 OTUs contain at least one highly abundant ZOTU. The taxonomic profiles of the global core
349 annotated by the HOMD database and the ZOTUs annotated by the SILVA database were very
350 similar at the genus level (Figure 7A).
351
352 We wanted to verify the observed differences between the OTUs in saliva samples from Chinese and
353 Western people in this independent dataset. Although two groups did not differ significantly when
354 considering the Shannon diversity index (Wilcoxon rank-sum test, $p = 0.073$; Figure 7B), the Chinese
355 group showed a higher Chao 1 Index than the Western group (Wilcoxon rank-sum test, $p < 0.001$;
356 Figure 7C), which is in agreeance with the result of the meta-analysis. Meanwhile, no significant
357 differences were observed between the two groups by beta-diversity analyses (Bray-Curtis and
358 weighted uniFrac distance, Table S9). We used the random forest classification model constructed
359 using the genus level profile of the large-scale dataset to predict the Chinese and Western samples in
360 this independent study. This dataset's genus-level relative abundance table was prepared from both the
361 OTU table with 97% identity to HOMD v15.1 database (Figure 7D) and the ZOTU table annotated by

362 the HOMD v15.22 database (Figure 7E). The accuracy of both predictions was relatively low, with
363 57.7% for the OTU table and 50% for the ZOTU table. Interestingly, most Western samples were
364 correctly classified, while most samples from Chinese participants were classified as being 'Western'
365 in this analysis.
366

367 **Discussion**

368 There is ample knowledge on the disease-affected salivary microbiota, yet our perspectives to the
369 bacteria present in healthy humans remains limited. Our systematic selection of studies, together with
370 the re-analysis of the 16S rRNA amplicon sequencing data from 47 studies offers a comprehensive
371 description of the salivary microbiome presented in adults without diagnosed disease. Our study has
372 defined the core members of salivary bacterial communities across 2211 samples from 47 studies and
373 has used metadata captured in these studies to investigate the role of different intrinsic and extrinsic
374 factors on the occurrence of these core. It is clear that core members differ between geographic
375 locations of collected saliva, and our analysis shows that Chinese participants are different from
376 Western participants (encompassing European and North American studies). A prediction of the
377 pathways enriched in each collective indicates that bacterial metabolic pathways are likely to
378 influence the aroma and flavour perception of foods. These results show that despite the core
379 microbial members of saliva being common across humans, there are differences, likely due to diet.
380 We suggest that the aroma and flavour of foods and beverages are likely to be differently affected in
381 healthy humans across the globe, meaning that preference and consumption of different foods is likely
382 to be prioritised. These results have important consequences for food and beverage design,
383 composition, and dietary advice across the globe.
384

385 Based on the abundance-occupancy pattern, the definition of core microbiome highlighted the
386 persistent and conserved microbial communities in human saliva across the globe. Here, we compared
387 two approaches adapted from two studies (references) to defining the core. The method adapted from
388 the study of Wu et al. [13] is a relatively conventional strategy that has been chosen by many other

389 studies with customised thresholds [45-47]. For another approach we applied in this study [35], we
390 determined the occupancy of OTUs according to their detection over study. When evaluating the
391 contribution of top-ranked OTUs in occupancy to the beta-diversity of the community, we further
392 used weighted Unifrac due to its effectiveness in minimising the biases induced by the selection of
393 hypervariable regions.
394
395 The resulting core members identified by these methods have a lot in common. A majority of the core
396 OTUs defined by Shade's method is also included in the cores defined by Wu's method, suggesting
397 the recently developed multi-step approach is effective in determining taxa with high prevalence. The
398 general high relative abundance of the core across different sub-groups emphasised the utility of this
399 pipeline in identifying the persistent members across diverse datasets. Most of the core salivary
400 microbiota we defined had been proposed in previous studies as prevalent bacteria in the human oral
401 cavity that persistently span across different individuals [11, 48-50]. The dominant genus of the core
402 we defined, *Streptococcus*, *Neisseria*, and *Prevotella*, were concluded as core human salivary
403 microbiome by a recent study based on the MG-RAST data [51]. Ten OTUs belonging to genus
404 *Streptococcus* were included in the "core", two of which were classified to the cluster with overall the
405 highest relative abundance across all dimensions. The prevalence of *Streptococcus* we observed is
406 consistent with a previous study defining the healthy core from the 454 pyrosequencing results of
407 three individuals [11]. The most abundant core OTU we found, *Streptococcus*
408 *oralis* subspecies *dentisani*, has been documented in previous studies as potential oral health-
409 promoting organisms and being highly abundant at various oral niches of healthy humans [52, 53].
410
411 We further conducted a co-occurrence network analysis to investigate the role of these core
412 microbiota in shaping the microbial community and found the co-existence between many members
413 of the core (Figure S5). The presence of the rare OTU that became the hub suggests that although
414 some taxa are not persistently detected across the community, they may still be important for the
415 overall structure of the salivary microbiome. It has been proposed that the oral microbiota of healthy
416 individuals is both homeostatic and dynamic [48]. The core microbial members consistently present in

417 human saliva identified here may explain the stability of oral microbiota to some extent. Moreover,
418 we conducted an independent study to verify the prevalence of the core defined from the published
419 studies. The high relative abundance and occurrence of the original core OTUs in this independent
420 cohort suggest that the core human salivary microbiome we defined can be applied to different
421 datasets. After re-clustering the core OTUs to ZOTUs with 100% sequence identity, we may conclude
422 that the same members constitute the core, even if different taxonomic resolution is applied.
423
424 Besides the core microbiome, there are "variable" microbiota in human microbial communities, which
425 vary among individuals because of unique lifestyle and genetic factors [54]. We performed analyses
426 for beta-diversity of samples (Figure 3) and random forest classifications (Figure 5A), demonstrating
427 several factors-both technical and physiological-significantly discriminated between sub-populations.
428 Because of the high heterogeneity between studies in their methodology, large inter-study variability
429 was the main factor that affected the observed salivary microbiota [55, 56]. For example, the impact
430 of chosen hypervariable regions for sequencing on driving microbial community structures has been
431 confirmed by our study. Inter-study variation may also be attributed to the criteria of recruiting
432 participants. Although all samples included in this meta-analysis were collected from the control
433 groups and population without specific diagnosed disease, the definition of 'healthy' varies. For
434 example, the use of antibiotics was not always considered as an exclusion criterion, and when
435 included, different time intervals were adopted. However, we aimed to construct a microbial
436 community that reflects the salivary microbiota of real-life consumers.
437
438 The importance of various host intrinsic and lifestyle factors, especially smoking, changed the
439 salivary microbial composition [57]. We identified a core OTU belonging to *Megasphaera*
440 *micronuciformis* as a biomarker for reported smoking [58, 59]. *Veillonella atypica* OTU's vary
441 between smokers and non-smokers, agreeing with a previous study [60]. Another strong determinant
442 of the salivary microbiome we defined was "sample type". As we found in this study, the difference
443 between the mouthwash sample and the other two collection methods is greater than the difference

444 between the stimulated and unstimulated saliva. Contrary to the result of Jo et al. [29], the OTU
445 belonging to *Neisseria flava* has not been identified as differential taxa for the type of saliva. The
446 microbial composition of alcohol drinkers was different to non-drinkers, but small sample size hinders
447 other conclusions.
448
449 Geographic location has been identified as the host physiological factor with the largest impact on
450 salivary microbiota (Figure 5A). Although it only explained limited variability between samples'
451 microbial profiles, the observed variations were robust to the heterogeneity induced by different
452 hypervariable regions used (Figure 3E, F). To date, little is known regarding the influence of
453 geographic locations on the human salivary microbiome. A comparative study reported the
454 differences in saliva microbial composition in Alaskans, Africans, and Germans [33]. Our result
455 extends this study to demonstrate that core OTUs may differentiate saliva samples from North
456 America, Europe, and China. Given the high abundance and occupancy frequency of the core
457 microbiome, we would expect these taxa to be effective indicators to predict the geographic
458 background of saliva donors.
459
460 Due to the sometimes large differences in culture and lifestyle between Western and non-Western
461 populations, we further grouped our data into Western and Chinese samples. The comparison between
462 Western and non-Western populations has already been applied to the human gut microbiota, whereas
463 less is known about the saliva microbiota [61]. Our study found the difference between saliva fromcc
464 Western and Chinese people in the abundance of *Veillonella* spp. Here, we showed that *Veillonella*
465 was generally higher in Western samples. Such differences may further influence the flavone and
466 flavonol biosynthesis pathways in the oral cavity. Our previous study revealed the Western-born and
467 Chinese-born wine experts had different responses to the flavour of wine (unpublished results:
468 10.13140/RG.2.2.11263.12965) [62]. Since flavonol is a well-known constituent of wine-related to
469 the bitterness and astringency perception [63], we would hypothesis that the enrichment of *Veillonella*
470 in Western may affect their sensitivity to the phenolic compounds in wine. Regular consumption of

471 flavonoid-rich foods, such as oolong tea, may increase the abundance of *Veillonella* spp. in human
472 saliva [64]. These could link the differences between Chinese and Western groups in sensory
473 evaluation and their salivary microbiota together.
474
475 Although the independent dataset we collected did not show differences between samples taken from
476 Chinese and Western participants, the results of random forest classification may lead to some
477 interesting hypotheses. The prediction of models revealed that most of the Chinese samples in this
478 cohort were classified as Western. The donors of these samples were wine experts, of Chinese
479 ethnicity, born in China and living in Australia for no more than 18 months. Recent studies reported
480 that immigrants from Asia experience a "Westernization" of gut microbiota induced by dietary
481 acculturation [65, 66]. We hypothesise that such a phenomenon may also happen in salivary
482 microbiota. It may suggest that dietary pattern is a more important determinant than ethnicity in
483 shaping the salivary microbiota of the participants, leading to variation among different geographic
484 locations that we observed in the meta-analysis.
485 In summary, we have defined a core bacterial community in saliva from healthy humans, and this core
486 demonstrated both stability and variability among populations. The prevalence of the core members of
487 the saliva microbiome has been confirmed in an independent cohort. We have revealed the influence
488 of various host factors, such as geographic locations, incidence of smoking and drinking, on the
489 salivary microbiome. We also identified microbial and functional biomarkers to differentiate the
490 Chinese and Western people, underlying the potential relationship between salivary microbiota and
491 sensory perception. Results in this work will provide foundational information to inform future
492 studies to understand the similarities and differences in saliva microbial composition, potentially
493 associating oral to aroma and flavour perception of foods.

494 # Declarations

495 **Ethics approval and consent to participate**

496 The study was approved by the Office for Research Ethics and Integrity of the University of
497 Melbourne (Ethics ID: 1852616), and all participants provided written informed consent.

498 **Data availability**

499 The sequencing data supporting the conclusion of the meta-analysis in this article are available in
500 publicly accessible databases (full details can be found in Table S2). The sequencing data generated
501 and/or analysed during the current study are available in the NCBI Bioproject repository,
502 PRJNA786805 (https://www.ncbi.nlm.nih.gov/bioproject/786805). Original scripts generated during
503 the current study are available in Github
504 (https://github.com/XINWEIR/SalivaryMicrobiome_MetaAnalysis).

505 **Competing interests**

506 The authors declare that they have no competing interests.

507 **Authors' contributions**

508 The research topic was developed by XR and KH. XR conducted the meta-analysis with the guidance
509 of KH. XR and JL collected and analysed the local samples. XR wrote the first draft of the manuscript
510 and all authors contributed to revise it. All authors read and approved the final manuscript for
511 submission.

512 **Acknowledgements**

513 This study was funded by the Faculty of Veterinary and Agricultural Sciences at the University of
514 Melbourne. JL and XR gratefully acknowledge a Melbourne Research Scholarship administered by
515 the University of Melbourne.

516

References

1. Mosca AC, Chen J. Food-saliva interactions: Mechanisms and implications. Trends Food Sci Technol. 2017;66:125-34. https://doi.org/https://doi.org/10.1016/j.tifs.2017.06.005.
2. Liu D, Jiang X, Zheng HJ, Xie B, Wang H, He T, et al., editors. The modularity of microbial interaction network in healthy human saliva: Stability and specificity. 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); 2017 13-16 Nov. 2017.
3. Acharya A, Chan Y, Kheur S, Jin LJ, Watt RM, Mattheos N. Salivary microbiome in non-oral disease: A summary of evidence and commentary. Arch Oral Biol. 2017;83:169-73. https://doi.org/https://doi.org/10.1016/j.archoralbio.2017.07.019.
4. Mark Welch JL, Rossetti BJ, Rieken CW, Dewhirst FE, Borisy GG. Biogeography of a human oral microbiome at the micron scale. PNAS. 2016;113(6):E791-E800. https://doi.org/10.1073/pnas.1522149113.
5. Shaw L, Ribeiro ALR, Levine AP, Pontikos N, Balloux F, Segal AW, et al. The Human Salivary Microbiome Is Shaped by Shared Environment Rather than Genetics: Evidence from a Large Family of Closely Related Individuals. mBio. 2017;8(5):e01237-17. https://doi.org/doi:10.1128/mBio.01237-17.
6. Muñoz-González C, Cueva C, Ángeles Pozo-Bayón M, Victoria Moreno-Arribas M. Ability of human oral microbiota to produce wine odorant aglycones from odourless grape glycosidic aroma precursors. Food Chem. 2015;187:112-9. https://doi.org/https://doi.org/10.1016/j.foodchem.2015.04.068.
7. Parker M, Onetto C, Hixson J, Bilogrevic E, Schueth L, Pisaniello L, et al. Factors Contributing to Interindividual Variation in Retronasal Odor Perception from Aroma Glycosides: The Role of Odorant Sensory Detection Threshold, Oral Microbiota, and Hydrolysis in Saliva. J Agric Food Chem. 2020;68(38):10299-309. https://doi.org/10.1021/acs.jafc.9b05450.
8. Piombino P, Genovese A, Esposito S, Moio L, Cutolo PP, Chambery A, et al. Saliva from Obese Individuals Suppresses the Release of Aroma Compounds from Wine. PLoS One. 2014;9(1):e85611. https://doi.org/10.1371/journal.pone.0085611.
9. De Filippis F, Parente E, Zotta T, Ercolini D. A comparison of bioinformatic approaches for 16S rRNA gene profiling of food bacterial microbiota. Int J Food Microbiol. 2018;265:9-17. https://doi.org/https://doi.org/10.1016/j.ijfoodmicro.2017.10.028.
10. Soriano-Lerma A, Pérez-Carrasco V, Sánchez-Marañón M, Ortiz-González M, Sánchez-Martín V, Gijón J, et al. Influence of 16S rRNA target region on the outcome of microbiome studies in soil and saliva samples. Sci Rep. 2020;10(1):13637. https://doi.org/10.1038/s41598-020-70141-8.
11. Zaura E, Keijser BJF, Huse SM, Crielaard W. Defining the healthy "core microbiome" of oral microbial communities. BMC Microbiol. 2009;9(1):259. https://doi.org/10.1186/1471-2180-9-259.
12. Risely A. Applying the core microbiome to understand host–microbe systems. J Anim Ecol. 2020;89(7):1549-58. https://doi.org/https://doi.org/10.1111/1365-2656.13229.
13. Wu L, Ning D, Zhang B, Li Y, Zhang P, Shan X, et al. Global diversity and biogeography of bacterial communities in wastewater treatment plants. Nat Microbiol. 2019;4(7):1183-95. https://doi.org/10.1038/s41564-019-0426-5.
14. Guo R, Zheng Y, Zhang L, Shi J, Li W. Salivary microbiome and periodontal status of patients with periodontitis during the initial stage of orthodontic treatment. Am J Orthod Dentofacial Orthop. 2021;159(5):644-52. https://doi.org/https://doi.org/10.1016/j.ajodo.2019.11.026.
15. Relvas M, Regueira-Iglesias A, Balsa-Castro C, Salazar F, Pacheco JJ, Cabral C, et al. Relationship between dental and periodontal health status and the salivary microbiome: bacterial diversity, co-occurrence networks and predictive models. Sci Rep. 2021;11(1):929. https://doi.org/10.1038/s41598-020-79875-x.
16. Wang K, Lu W, Tu Q, Ge Y, He J, Zhou Y, et al. Preliminary analysis of salivary microbiome and their potential roles in oral lichen planus. Sci Rep. 2016;6(1):22943. https://doi.org/10.1038/srep22943.
17. Sabharwal A, Ganley K, Miecznikowski JC, Haase EM, Barnes V, Scannapieco FA. The salivary microbiome of diabetic and non-diabetic adults with periodontal disease. J Periodontol. 2019;90(1):26-34. https://doi.org/https://doi.org/10.1002/JPER.18-0167.

18. Torres PJ, Fletcher EM, Gibbons SM, Bouvet M, Doran KS, Kelley ST. Characterization of the salivary microbiome in patients with pancreatic cancer. PeerJ. 2015;3:e1373. https://doi.org/10.7717/peerj.1373.
19. Lindheim L, Bashir M, Münzker J, Trummer C, Zachhuber V, Pieber TR, et al. The Salivary Microbiome in Polycystic Ovary Syndrome (PCOS) and Its Association with Disease-Related Parameters: A Pilot Study. Front Microbiol. 2016;7(1270). https://doi.org/10.3389/fmicb.2016.01270.
20. Nomura Y, Kakuta E, Okada A, Otsuka R, Shimada M, Tomizawa Y, et al. Oral Microbiome in Four Female Centenarians. Appl Sci. 2020;10(15):5312. https://doi.org/10.3390/app10155312.
21. Hansen TH, Kern T, Bak EG, Kashani A, Allin KH, Nielsen T, et al. Impact of a vegan diet on the human salivary microbiota. Sci Rep. 2018;8(1):5847. https://doi.org/10.1038/s41598-018-24207-3.
22. Kato I, Vasquez A, Moyerbrailean G, Land S, Djuric Z, Sun J, et al. Nutritional Correlates of Human Oral Microbiome. J Am Coll Nutr. 2017;36(2):88-98. https://doi.org/10.1080/07315724.2016.1185386.
23. Liu K, Chen S, Huang J, Ren F, Yang T, Long D, et al. Oral Microbiota of Children Is Conserved across Han, Tibetan and Hui Groups and Is Correlated with Diet and Gut Microbiota. Microorganisms. 2021;9(5):1030. https://doi.org/10.3390/microorganisms9051030.
24. Renson A, Jones HE, Beghini F, Segata N, Zolnik CP, Usyk M, et al. Sociodemographic variation in the oral microbiome. Ann Epidemiol. 2019;35:73-80.e2. https://doi.org/https://doi.org/10.1016/j.annepidem.2019.03.006.
25. Wu J, Peters BA, Dominianni C, Zhang Y, Pei Z, Yang L, et al. Cigarette smoking and the oral microbiome in a large study of American adults. ISME J. 2016;10(10):2435-46. https://doi.org/10.1038/ismej.2016.37.
26. Fan X, Peters BA, Jacobs EJ, Gapstur SM, Purdue MP, Freedman ND, et al. Drinking alcohol is associated with variation in the human oral microbiome in a large study of American adults. Microbiome. 2018;6(1):59. https://doi.org/10.1186/s40168-018-0448-x.
27. Takayasu L, Suda W, Takanashi K, Iioka E, Kurokawa R, Shindo C, et al. Circadian oscillations of microbial and functional composition in the human salivary microbiome. DNA Res. 2017;24(3):261-70. https://doi.org/10.1093/dnares/dsx001.
28. Wu Y, Chi X, Zhang Q, Chen F, Deng X. Characterization of the salivary microbiome in people with obesity. PeerJ. 2018;6:e4458.
29. Jo R, Nishimoto Y, Umezawa K, Yama K, Aita Y, Ichiba Y, et al. Comparison of oral microbiome profiles in stimulated and unstimulated saliva, tongue, and mouth-rinsed water. Sci Rep. 2019;9(1):16124. https://doi.org/10.1038/s41598-019-52445-6.
30. Cattaneo C, Riso P, Laureati M, Gargari G, Pagliarini E. Exploring Associations between Interindividual Differences in Taste Perception, Oral Microbiota Composition, and Reported Food Intake. Nutrients. 2019;11(5):1167. https://doi.org/10.3390/nu11051167.
31. Gardner A, So PW, Carpenter GH. Intraoral Microbial Metabolism and Association with Host Taste Perception. J Dent Res. 2020;99(6):739-45. https://doi.org/10.1177/0022034520917142.
32. Henne K, Li J, Stoneking M, Kessler O, Schilling H, Sonanini A, et al. Global analysis of saliva as a source of bacterial genes for insights into human population structure and migration studies. BMC Evol Biol. 2014;14(1):190. https://doi.org/10.1186/s12862-014-0190-3.
33. Li J, Quinque D, Horz H-P, Li M, Rzhetskaya M, Raff JA, et al. Comparative analysis of the human saliva microbiome from different climate zones: Alaska, Germany, and Africa. BMC Microbiol. 2014;14(1):316. https://doi.org/10.1186/s12866-014-0316-1.
34. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37(8):852-7. https://doi.org/10.1038/s41587-019-0209-9.
35. Shade A, Stopnisek N. Abundance-occupancy distributions to prioritize plant core microbiome membership. Curr Opin Microbiol. 2019;49:50-8. https://doi.org/https://doi.org/10.1016/j.mib.2019.09.008.
36. Frank E, Jr. H with contributions from Charles Dupont and many others. Hmisc: Harrell Miscellaneous. R package version 4.5-0. 2021. https://CRAN.R-project.org/package=Hmisc.

37. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498-504.
38. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 2013;42(D1):D199-D205. https://doi.org/10.1093/nar/gkt1076.
39. Wei T, Simko V. R package 'corrplot': Visualization of a Correlation Matrix (Version 0.90). 2021. https://github.com/taiyun/corrplot.
40. Marotz C, Amir A, Humphrey G, Gaffney J, Gogul G, Knight R. DNA extraction for streamlined metagenomics of diverse environmental samples. BioTechniques. 2017;62(6):290-3. https://doi.org/10.2144/000114559.
41. Edgar RC. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. bioRxiv. 2016:081257. https://doi.org/10.1101/081257.
42. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013;41(Database issue):D590-D6. https://doi.org/10.1093/nar/gks1219.
43. Chen T, Yu W-H, Izard J, Baranova OV, Lakshmanan A, Dewhirst FE. The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. Database. 2010;2010. https://doi.org/10.1093/database/baq013.
44. Stopnisek N, Shade A. Persistent microbiome members in the common bean rhizosphere: an integrated analysis of space, time, and plant genotype. ISME J. 2021;15(9):2708-22. https://doi.org/10.1038/s41396-021-00955-5.
45. Wang Y, Gong J, Li J, Xin Y, Hao Z, Chen C, et al. Insights into bacterial diversity in compost: Core microbiome and prevalence of potential pathogenic bacteria. Sci Total Environ. 2020;718:137304. https://doi.org/https://doi.org/10.1016/j.scitotenv.2020.137304.
46. Giordano C, Boscaro V, Munz G, Mori G, Vannini C. Summer holidays as break-point in shaping a tannery sludge microbial community around a stable core microbiota. Sci Rep. 2016;6(1):30376. https://doi.org/10.1038/srep30376.
47. Salonen A, Salojärvi J, Lahti L, de Vos WM. The adult intestinal core microbiota is determined by analysis depth and health status. Clin Microbiol Infect. 2012;18(s4):16-20. https://doi.org/https://doi.org/10.1111/j.1469-0691.2012.03855.x.
48. Hall MW, Singh N, Ng KF, Lam DK, Goldberg MB, Tenenbaum HC, et al. Inter-personal diversity and temporal dynamics of dental, tongue, and salivary microbiota in the healthy oral cavity. NPJ Biofilms Microbiomes. 2017;3(1):2. https://doi.org/10.1038/s41522-016-0011-0.
49. Simón-Soro Á, Tomás I, Cabrera-Rubio R, Catalan MD, Nyvad B, Mira A. Microbial Geography of the Oral Cavity. J Dent Res. 2013;92(7):616-21. https://doi.org/10.1177/0022034513488119.
50. Yao T, Han X, Guan T, Zhai C, Liu C, Liu C, et al. Exploration of the microbiome community for saliva, skin, and a mixture of both from a population living in Guangdong. Int J Legal Med. 2021;135(1):53-62. https://doi.org/10.1007/s00414-020-02329-6.
51. Oliveira SG, Nishiyama RR, Trigo CAC, Mattos-Guaraldi AL, Dávila AMR, Jardim R, et al. Core of the saliva microbiome: an analysis of the MG-RAST data. BMC Oral Health. 2021;21(1):351. https://doi.org/10.1186/s12903-021-01719-5.
52. Conrads G, Westenberger J, Lürkens M, Abdelbary MMH. Isolation and Bacteriocin-Related Typing of Streptococcus dentisani. Front Cell Infect Microbiol. 2019;9(110). https://doi.org/10.3389/fcimb.2019.00110.
53. López-López A, Camelo-Castillo A, Ferrer MD, Simon-Soro Á, Mira A. Health-Associated Niche Inhabitants as Oral Probiotics: The Case of Streptococcus dentisani. Front Microbiol. 2017;8(379). https://doi.org/10.3389/fmicb.2017.00379.
54. Zarco M, Vess T, Ginsburg G. The oral microbiome in health and disease and the potential impact on personalized dental medicine. Oral Dis. 2012;18(2):109-20. https://doi.org/https://doi.org/10.1111/j.1601-0825.2011.01851.x.
55. Cornejo-Granados F, Gallardo-Becerra L, Leonardo-Reza M, Ochoa-Romo JP, Ochoa-Leyva A. A meta-analysis reveals the environmental and host factors shaping the structure and function of the shrimp microbiota. PeerJ. 2018;6:e5382. https://doi.org/10.7717/peerj.5382.

678 56. Holman DB, Brunelle BW, Trachsel J, Allen HK, Bik H. Meta-analysis To Define a Core
679 Microbiota in the Swine Gut. mSystems. 2017;2(3):e00004-17.
680 https://doi.org/doi:10.1128/mSystems.00004-17.
681 57. Roberts FA, Darveau RP. Microbial protection and virulence in periodontal tissue as a
682 function of polymicrobial communities: symbiosis and dysbiosis. Periodontol 2000. 2015;69(1):18-
683 27. https://doi.org/https://doi.org/10.1111/prd.12087.
684 58. Sato N, Kakuta M, Hasegawa T, Yamaguchi R, Uchino E, Kobayashi W, et al. Metagenomic
685 analysis of bacterial species in tongue microbiome of current and never smokers. NPJ Biofilms
686 Microbiomes. 2020;6(1):11. https://doi.org/10.1038/s41522-020-0121-6.
687 59. Vogtmann E, Flores R, Yu G, Freedman ND, Shi J, Gail MH, et al. Association between
688 tobacco use and the upper gastrointestinal microbiome among Chinese men. Cancer Causes Control.
689 2015;26(4):581-8. https://doi.org/10.1007/s10552-015-0535-2.
690 60. Karabudak S, Ari O, Durmaz B, Dal T, Basyigit T, Kalcioglu MT, et al. Analysis of the effect
691 of smoking on the buccal microbiome using next-generation sequencing technology. J Med
692 Microbiol. 2019;68(8):1148-58. https://doi.org/https://doi.org/10.1099/jmm.0.001003.
693 61. Prasoodanan P. K V, Sharma AK, Mahajan S, Dhakan DB, Maji A, Scaria J, et al. Western
694 and non-western gut microbiomes reveal new roles of Prevotella in carbohydrate metabolism and
695 mouth–gut axis. NPJ Biofilms Microbiomes. 2021;7(1):77. https://doi.org/10.1038/s41522-021-
696 00248-x.
697 62. Luo J, Ruan X, Ang CS, Marriott P, Zhang P, Howell K. Can variation in wine preference
698 amongst consumers be explained by salivary protein composition?
699 http://hdl.handle.net/11343/294188. Accessed 20 Dec 2021.
700 63. García-Estévez I, Ramos-Pineda AM, Escribano-Bailón MT. Interactions between wine
701 phenolic compounds and human saliva in astringency perception. Food Funct. 2018;9(3):1294-309.
702 64. Liu Z, Guo H, Zhang W, Ni L. Salivary Microbiota Shifts under Sustained Consumption of
703 Oolong Tea in Healthy Adults. Nutrients. 2020;12(4):966. https://doi.org/10.3390/nu12040966.
704 65. Peters BA, Yi SS, Beasley JM, Cobbs EN, Choi HS, Beggs DB, et al. US nativity and dietary
705 acculturation impact the gut microbiome in a diverse US population. ISME J. 2020;14(7):1639-50.
706 https://doi.org/10.1038/s41396-020-0630-6.
707 66. Vangay P, Johnson AJ, Ward TL, Al-Ghalith GA, Shields-Cutler RR, Hillmann BM, et al. US
708 Immigration Westernizes the Human Gut Microbiome. Cell. 2018;175(4):962-72.e10.
709 https://doi.org/https://doi.org/10.1016/j.cell.2018.10.029.
710 67. Acharya A, Chan Y, Kheur S, Kheur M, Gopalakrishnan D, Watt R, et al. Salivary
711 microbiome of an urban Indian cohort and patterns linked to subclinical inflammation. Oral Dis.
712 2017;23(7):926-40. https://doi.org/10.1111/odi.12676.
713 68. Anderson A, Al-Ahmad A, Schlueter N, Frese C, Hellwig E, Binder N. Influence of the long-
714 term use of oral hygiene products containing stannous ions on the salivary microbiome–a randomized
715 controlled trial. Sci Rep. 2020;10(1):1-8. https://doi.org/10.1038/s41598-020-66412-z.
716 69. Bhushan B, Yadav A, Singh S, Ganju L. Diversity and functional analysis of salivary
717 microflora of Indian Antarctic expeditionaries. J Oral Microbiol. 2019;11(1):1581513.
718 https://doi.org/10.1080/20002297.2019.1581513.
719 70. Cabral DJ, Wurster JI, Flokas ME, Alevizakos M, Zabat M, Korry BJ, et al. The salivary
720 microbiome is consistent between subjects and resistant to impacts of short-term hospitalization. Sci
721 Rep. 2017;7(1):11040. https://doi.org/10.1038/s41598-017-11427-2.
722 71. Cameron SJS, Huws SA, Hegarty MJ, Smith DPM, Mur LAJ. The human salivary
723 microbiome exhibits temporal stability in bacterial diversity. FEMS Microbiol Ecol. 2015;91(9).
724 https://doi.org/10.1093/femsec/fiv091.
725 72. Chaudhari DS, Dhotre DP, Agarwal DM, Gaike AH, Bhalerao D, Jadhav P, et al. Gut, oral
726 and skin microbiome of Indian patrilineal families reveal perceptible association with age. Sci Rep.
727 2020;10(1):5685. https://doi.org/10.1038/s41598-020-62195-5.
728 73. Chen C, Hemme C, Beleno J, Shi ZJ, Ning D, Qin Y, et al. Oral microbiota of periodontal
729 health and disease and their changes after nonsurgical periodontal therapy. ISME J. 2018;12(5):1210-
730 24. https://doi.org/10.1038/s41396-017-0037-1.

731 74. Collado MC, Engen PA, Bandín C, Cabrera-Rubio R, Voigt RM, Green SJ, et al. Timing of
732 food intake impacts daily rhythms of human salivary microbiota: a randomized, crossover study.
733 FASEB J. 2018;32(4):2060-72. https://doi.org/10.1096/fj.201700697RR.
734 75. Debelius JW, Huang T, Cai Y, Ploner A, Barrett D, Zhou X, et al. Subspecies niche
735 specialization in the oral microbiome is associated with nasopharyngeal carcinoma risk. Msystems.
736 2020;5(4):e00065-20. https://doi.org/10.1128/mSystems.00065-20.
737 76. Hijazi K, Morrison RW, Mukhopadhya I, Martin B, Gemmell M, Shaw S, et al. Oral bacterial
738 diversity is inversely correlated with mucosal inflammation. Oral Dis. 2020;26(7):1566-75.
739 https://doi.org/https://doi.org/10.1111/odi.13420.
740 77. Ji Y, Liang X, Lu H. Analysis of by high-throughput sequencing: Helicobacter pylori
741 infection and salivary microbiome. BMC Oral Health. 2020;20(1):84. https://doi.org/10.1186/s12903-
742 020-01070-1.
743 78. Jiang Q, Liu J, Chen L, Gan N, Yang D. The Oral Microbiome in the Elderly With Dental
744 Caries and Health. Front Cell Infect Microbiol. 2019;8(442).
745 https://doi.org/10.3389/fcimb.2018.00442.
746 79. Kumpitsch C, Moissl-Eichinger C, Pock J, Thurnher D, Wolf A. Preliminary insights into the
747 impact of primary radiochemotherapy on the salivary microbiome in head and neck squamous cell
748 carcinoma. Sci Rep. 2020;10(1):1-12. https://doi.org/10.1038/s41598-020-73515-0.
749 80. Lee W-H, Chen H-M, Yang S-F, Liang C, Peng C-Y, Lin F-M, et al. Bacterial alterations in
750 salivary microbiota and their association in oral cancer. Sci Rep. 2017;7(1):1-11.
751 https://doi.org/10.1038/s41598-017-16418-x.
752 81. Lin M, Li X, Wang J, Cheng C, Zhang T, Han X, et al. Saliva microbiome changes in patients
753 with periodontitis with and without chronic obstructive pulmonary disease. Front Cell Infect
754 Microbiol. 2020;10:124. https://doi.org/10.3389/fcimb.2020.00124.
755 82. Lindheim L, Bashir M, Münzker J, Trummer C, Zachhuber V, Pieber TR, et al. The salivary
756 microbiome in polycystic ovary syndrome (pcos) and its association with disease-related parameters:
757 a pilot study. Front Microbiol. 2016;7:1270. https://doi.org/10.3389/fmicb.2016.01270.
758 83. Liu Y, Zhang Q, Hu X, Chen F, Hua H. Characteristics of the salivary microbiota in cheilitis
759 granulomatosa. Med Oral Patol Oral Cir Bucal. 2019;24(6):e719-e25.
760 https://doi.org/10.4317/medoral.23041.
761 84. Lundmark A, Hu YO, Huss M, Johannsen G, Andersson AF, Yucel-Lindberg T.
762 Identification of salivary microbiota and its association with host inflammatory mediators in
763 periodontitis. Front Cell Infect Microbiol. 2019;9:216. https://doi.org/10.3389/fcimb.2019.00216.
764 85. Menon R, Gomez A, Brandt B, Leung Y, Gopinath D, Watt R, et al. Long-term impact of oral
765 surgery with or without amoxicillin on the oral microbiome-A prospective cohort study. Sci Rep.
766 2019;9(1):1-10. https://doi.org/10.1038/s41598-019-55056-3.
767 86. Niu C, Dong T, Jiang W, Gao L, Yuan K, Hu X, et al. Pregnancy-Related Ecological Shifts of
768 Salivary Microbiota and its Association with Salivary Sex Hormones. 2020.
769 https://doi.org/10.21203/rs.3.rs-54822/v1.
770 87. Ozga AT, Sankaranarayanan K, Tito RY, Obregon-Tito AJ, Foster MW, Tallbull G, et al.
771 Oral microbiome diversity among Cheyenne and Arapaho individuals from Oklahoma. Am J Phys
772 Anthropol. 2016;161(2):321-7. https://doi.org/10.1002/ajpa.23033.
773 88. PRJNA421776. Exercise frequency is associated with the oral microbiota of student athletes
774 and non-athletes. 2017. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA421776.
775 89. Pushalkar S, Paul B, Li Q, Yang J, Vasconcelos R, Makwana S, et al. Electronic Cigarette
776 Aerosol Modulates the Oral Microbiome and Increases Risk of Infection. iScience.
777 2020;23(3):100884. https://doi.org/https://doi.org/10.1016/j.isci.2020.100884.
778 90. Russo E, Bacci G, Chiellini C, Fagorzi C, Niccolai E, Taddei A, et al. Preliminary
779 comparison of oral and intestinal human microbiota in patients with colorectal cancer: a pilot study.
780 Front Microbiol. 2018;8:2699. https://doi.org/10.3389/fmicb.2017.02699.
781 91. Sohail MU, Elrayess MA, Al Thani AA, Al-Asmakh M, Yassine HM. Profiling the oral
782 microbiome and plasma biochemistry of obese hyperglycemic subjects in Qatar. Microorganisms.
783 2019;7(12):645. https://doi.org/10.3390/microorganisms7120645.

92. Soriano-Lerma A, Pérez-Carrasco V, Sánchez-Marañón M, Ortiz-González M, Sánchez-Martín V, Gijón J, et al. Influence of 16S rRNA target region on the outcome of microbiome studies in soil and saliva samples. Sci Rep. 2020;10(1):1-13. https://doi.org/10.1038/s41598-020-70141-8.
93. Stewart CJ, Auchtung TA, Ajami NJ, Velasquez K, Smith DP, De La Garza II R, et al. Effects of tobacco smoke and electronic cigarette vapor exposure on the oral and gut microbiota in humans: a pilot study. PeerJ. 2018;6:e4693. https://doi.org/10.7717/peerj.4693.
94. Sun X, Li M, Xia L, Fang Z, Yu S, Gao J, et al. Alteration of salivary microbiome in periodontitis with or without type-2 diabetes mellitus and metformin treatment. Sci Rep. 2020;10(1):1-14. https://doi.org/10.1038/s41598-020-72035-1.
95. Tian N, Faller L, Leffler DA, Kelly CP, Hansen J, Bosch JA, et al. Salivary gluten degradation and oral microbial profiles in healthy individuals and celiac disease patients. Appl Environ Microbiol. 2017;83(6):e03330-16. https://doi.org/10.1128/AEM.03330-16.
96. Tong Y, Zheng L, Qing P, Zhao H, Li Y, Su L, et al. Oral microbiota perturbations are linked to high risk for rheumatoid arthritis. Front Cell Infect Microbiol. 2020;9:475. https://doi.org/10.3389/fcimb.2019.00475.
97. Urbaniak C, Lorenzi H, Thissen J, Jaing C, Crucian B, Sams C, et al. The influence of spaceflight on the astronaut salivary microbiome and the search for a microbiome biomarker for viral reactivation. Microbiome. 2020;8(1):1-14. https://doi.org/10.1186/s40168-020-00830-z.
98. Vesty A, Gear K, Biswas K, Radcliff FJ, Taylor MW, Douglas RG. Microbial and inflammatory-based salivary biomarkers of head and neck squamous cell carcinoma. Clin Exp Dent Res. 2018;4(6):255-62. https://doi.org/https://doi.org/10.1002/cre2.139.
99. Vogtmann E, Chen J, Kibriya MG, Amir A, Shi J, Chen Y, et al. Comparison of oral collection methods for studies of microbiota. Cancer Epidemiol Biomark Prev. 2019;28(1):137-43. https://doi.org/10.1158/1055-9965.EPI-18-0312.
100. Wang T, Yu L, Xu C, Pan K, Mo M, Duan M, et al. Chronic fatigue syndrome patients have alterations in their oral microbiome composition and function. PLoS One. 2018;13(9):e0203503. https://doi.org/10.1371/journal.pone.0203503.
101. Wang Q, Rao Y, Guo X, Liu N, Liu S, Wen P, et al. Oral microbiome in patients with oesophageal squamous cell carcinoma. Sci Rep. 2019;9(1):1-9. https://doi.org/10.1038/s41598-019-55667-w.
102. Wang X, Zhao Z, Tang N, Zhao Y, Xu J, Li L, et al. Microbial community analysis of saliva and biopsies in patients with oral lichen planus. Front Microbiol. 2020;11:629. https://doi.org/10.3389/fmicb.2020.00629.
103. Wolf A, Moissl-Eichinger C, Perras A, Koskinen K, Tomazic PV, Thurnher D. The salivary microbiome as an indicator of carcinogenesis in patients with oropharyngeal squamous cell carcinoma: A pilot study. Sci Rep. 2017;7(1):1-10. https://doi.org/10.1038/s41598-017-06361-2.
104. Xun Z, Zhang Q, Xu T, Chen N, Chen F. Dysbiosis and Ecotypes of the Salivary Microbiome Associated With Inflammatory Bowel Diseases and the Assistance in Diagnosis of Diseases Using Oral Bacterial Profiles. Front Microbiol. 2018;9(1136). https://doi.org/10.3389/fmicb.2018.01136.
105. Yano Y, Hua X, Wan Y, Suman S, Zhu B, Dagnall CL, et al. Comparison of Oral Microbiota Collected Using Multiple Methods and Recommendations for New Epidemiologic Studies. Msystems. 2020;5(4):e00156-20. https://doi.org/10.1128/mSystems.00156-20.
106. Yeo L-F, Aghakhanian FF, Tan JS, Gan HM, Phipps ME. Health and saliva microbiomes of a semi-urbanized indigenous tribe in Peninsular Malaysia. F1000Research. 2019;8. https://doi.org/10.12688/f1000research.17706.2.
107. Yu FY, Wang QQ, Li M, Cheng Y-H, Cheng Y-SL, Zhou Y, et al. Dysbiosis of saliva microbiome in patients with oral lichen planus. BMC Microbiol. 2020;20(1):1-12. https://doi.org/10.1186/s12866-020-01733-7.
108. Zhu C, Yuan C, Wei F-Q, Sun X-Y, Zheng S-G. Comparative evaluation of peptidome and microbiota in different types of saliva samples. Ann Transl Med. 2020;8(11):686-. https://doi.org/10.21037/atm-20-393.
109. Zhu C, Yuan C, Wei FQ, Sun XY, Zheng SG. Intraindividual Variation and Personal Specificity of Salivary Microbiota. J Dent Res. 2020;99(9):1062-71. https://doi.org/10.1177/0022034520917155.

110. Ziganshina EE, Sagitov II, Akhmetova RF, Saleeva GT, Kiassov AP, Gogoleva NE, et al. Comparison of the Microbiota and Inorganic Anion Content in the Saliva of Patients with Gastroesophageal Reflux Disease and Gastroesophageal Reflux Disease-Free Individuals. Biomed Res Int. 2020;2020:2681791. https://doi.org/10.1155/2020/2681791.

865 # Figures

866 **Figure 1. Overview of literature search procedure and metadata of included studies. A)** Large-
867 scale literature searching and data filtering process, followed by the number of samples submitted to
868 the bioinformatic analyses; **B)** The locations of studies, the scale of symbols that reflect the number of
869 samples of each study; **C)** Distribution of metadata categories.
870

871 **Figure 2. Summary of taxonomic composition and alpha diversity of included studies. A)** The
872 mean community composition of each study at the phylum level; The alpha-diversity measured by **B)**
873 Shannon index; **C)** Chao 1 index; **D)** Simpson's index, the colour of boxes stands for the geographic
874 location of the studies. The horizontal bars within boxes represent medians. The tops and bottoms of
875 boxes represent the 75th and 25th percentiles, respectively.

876 **Figure 3. The variability in human salivary microbiota have been explained by different factors.**
877 Among them, hypervariable regions and geographic locations have the largest impact. The effect of
878 the categories on the clustering of the sample was measured using PERMANOVA at four taxonomic
879 levels: family **(A),** Genus **(B),** species **(C)** and OTU level **(D).** The colour indicates the different
880 combinations of normalisation (TSS, Total-sum scaling; RRA, Rarefied relative abundance; CLR,
881 Centred log ratio) and indices (BC, Bray-Curtis; EUC, Euclidean; wUF, weighted uniFrac). Because
882 the results of rarefication (RAR) were very close to TSS and RRA, they were not displayed in the
883 figures. Principal coordinate analysis (PCoA) with Bray-Curtis **(E)** and weighted uniFrac **(F)** showing
884 the differences between samples from North America, Europe, and China.

885 **Figure 4. The core OTUs defined by abundance-occupancy pattern. A)** Venn diagram showing
886 the interaction between three methods used to define the core. Sixty-eight OTUs were defined as the
887 core for all methods. (MRA+OCC: The thresholds were setting on mean relative abundance and
888 occupancy to define the core; BC: The method adapted from Shade and Stopnisek using Bray-Curtis
889 similarity; wUF: The method adapted from Shade and Stopnisek using weighted uniFrac distance). **B)**
890 Pie chart showing the number of the core (pink) versus other OTUs (blue) identified in percentage. **C)**

891 Pie chart showing the relative abundance of the core and other OTUs across all samples. **D)** Relative
892 abundance of 68 core OTUs across subgroups classified by seven categories. **E)** Heatmap showing the
893 log-transformed mean relative abundance of each core OTU at each level of different categories.
894

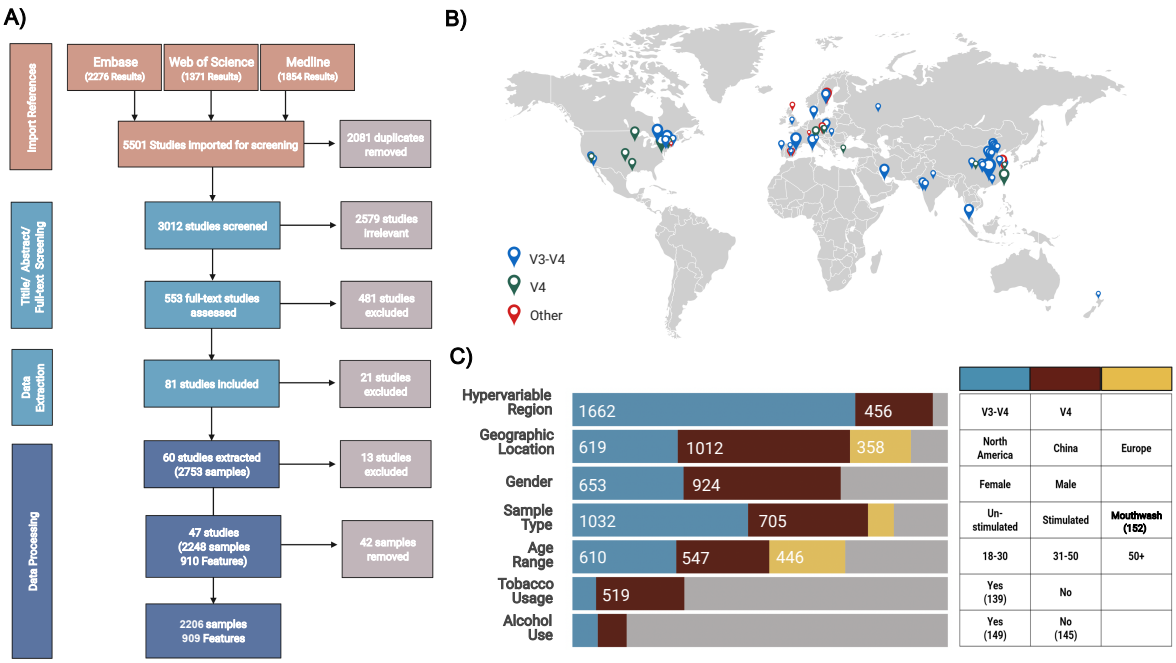
895 **Figure 5. Salivary microbiome members which significantly contribute to categorisation of**
896 **metadata.** Random Forest models showed the impact of categories on salivary microbiome and the
897 core OTUs contributing to accuracy of these models. **A)** Error rate (%) for the random forest
898 classifications conducted with samples grouped by eight different categories. **B)** Phylogenetic tree
899 indicates the taxonomic information of 68 core OTUs. The coloured squares between the tree and the
900 annotation of phylum indicate the OTUs that were defined by the Random Forest model as
901 "important" for distinguishing between different levels in each category. The bars on the outmost ring
902 showing the mean relative abundance of each OTU.

903 **Figure 6. Distinct microbial profiles are evident in the saliva samples from Chinese and Western**
904 **adults. A)** Taxonomic hierarchies show the relative enrichment of taxa in three geographic locations
905 at phylum through species level. Coloured nodes indicate log2-fold increase in median abundance of
906 the group in x-axis (pink) or y-axis (blue). Only taxa showed significant changes (false discovery rate-
907 adjusted Wilcoxon rank sum $q < 0.05$) are displayed. **B)** and **C)** Comparison of salivary microbial
908 alpha diversity between the Chinese and Western samples, calculated by Chao1 (B: $p < 0.001$,
909 Wilcoxon rank-sum test) and Shannon index (C: $p < 0.001$, Wilcoxon rank-sum test). **D)** Differential
910 abundant genera identified between saliva from Chinese and Western samples. The panel on the left
911 indicates the standardised effect sizes (W statistic) estimated via the difference on relative abundance
912 using ANCOM-BC (taxa enriched in Western samples have a value shifted to right, whereas taxa
913 enriched in Chinese samples have a value shifted to left); The panel in the middle shows the relative
914 abundance of selected genera; the panel on the right indicates the Mean Decrease Accuracy of the
915 random forest model established. **E)** Spearman's correlation coefficients were calculated between
916 each pairwise comparison of differential genus and KEGG pathway. Only significantly correlated

917 comparisons (p<0.01, FDR adjusted Spearman's rank correlation) are displayed. The only Western-
918 enriched pathway is marked in pink.

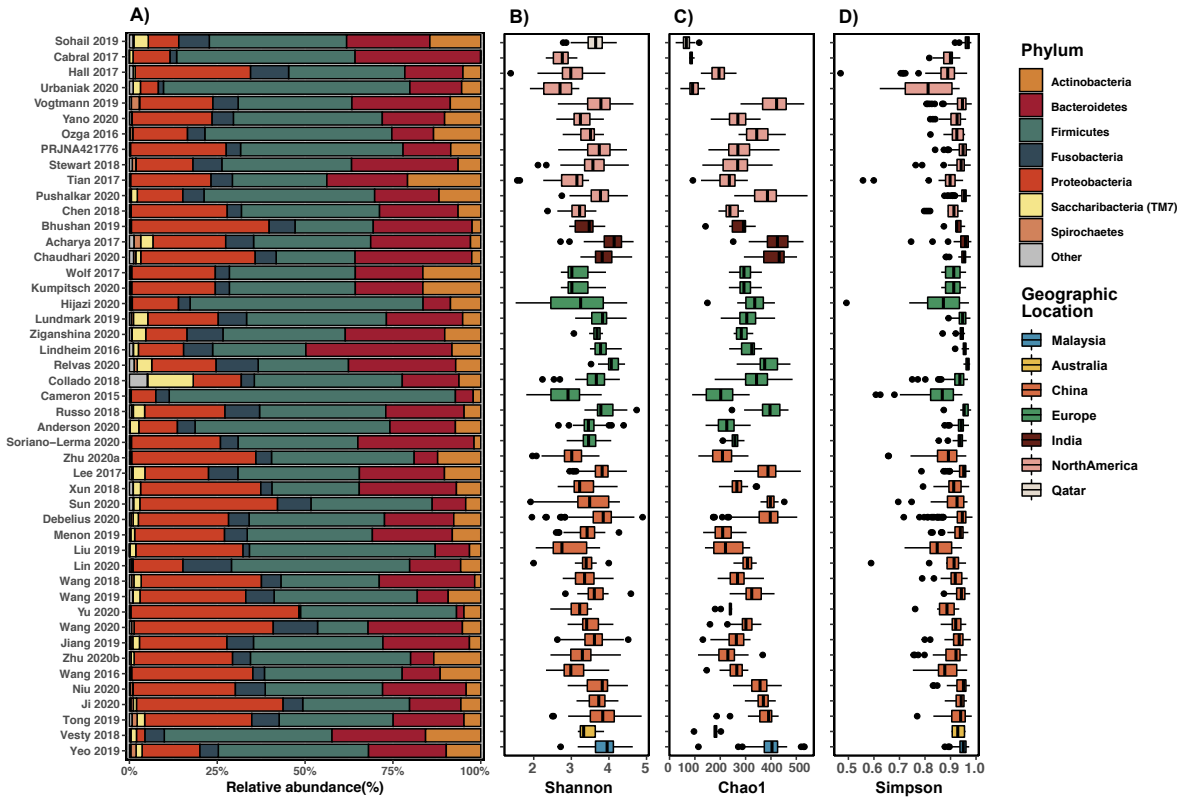
919 **Figure 7. An independent cohort verifies the definition of core microbiome membership**
920 **but cannot classify based on cultural background. A)** Alluvial plot showing the affiliation of
921 ZOTUs to their originating core OTUs defined in the meta-analysis. **B)** and **C)** Comparison of salivary
922 microbial alpha diversity between the Chinese and Western samples, calculated by Shannon (B: p =
923 0.073, Wilcoxon rank-sum test) and Chao1 index (C: p< 0.001, Wilcoxon rank-sum test). **D)** and **E)**
924 The prediction of the cultural backgrounds of the samples according to the random forest
925 classification model constructed using the genus profiles of samples in the meta-analysis. The genus
926 level profiles of samples processed by **D)** closed-reference clustering with 97% sequence identity and
927 **E)** UNOISE3 denoising with 100% sequence identity were used as the test set.

928


A)
Import References
Embase (2276 Results)
Web of Science (1371 Results)
Medline (1854 Results)
5501 Studies imported for screening
2081 duplicates removed
Title/ Abstract/ Full-text Screening
3012 studies screened
2579 studies irrelevant
553 full-text studies assessed
481 studies excluded
Data Extraction
81 studies included
21 studies excluded
Data Processing
60 studies extracted (2753 samples)
13 studies excluded
47 studies (2248 samples 910 Features)
42 samples removed
2206 samples 909 Feetures
B)
V3-V4
V4
Other
C)
Hypervariable Region
1662
456
Geographic Location
619
1012
358
Gender
653
924
Sample Type
1032
705
Age Range
610
547
446
Tobacco Usage
519
Alcohol Use
V3-V4
V4
North America
China
Europe
Female
Male
Un-stimulated
Stimulated
Mouthwash (152)
18-30
31-50
50+
Yes (139)
No
Yes (149)
No (145)

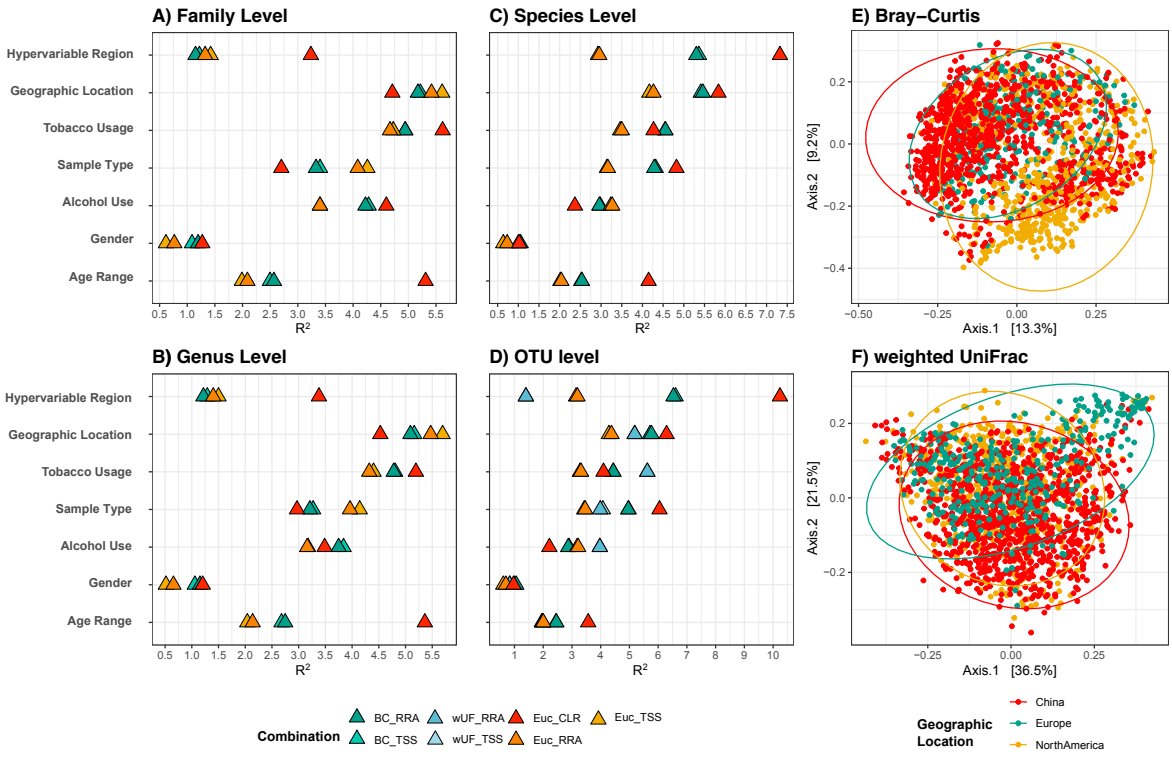
929

930


A)
B)
C)
D)
Sohail 2019
Cabral 2017
Hall 2017
Urbaniak 2020
Vogtmann 2019
Yano 2020
Ozga 2016
PRJNA421776
Stewart 2018
Tian 2017
Pushalkar 2020
Chen 2018
Bhushan 2019
Acharya 2017
Chaudhari 2020
Wolf 2017
Kumpitsch 2020
Hijazi 2020
Lundmark 2019
Ziganshina 2020
Lindheim 2016
Relvas 2020
Collado 2018
Cameron 2015
Russo 2018
Anderson 2020
Soriano−Lerma 2020
Zhu 2020a
Lee 2017
Xun 2018
Sun 2020
Debelius 2020
Menon 2019
Liu 2019
Lin 2020
Wang 2018
Wang 2019
Yu 2020
Wang 2020
Jiang 2019
Zhu 2020b
Wang 2016
Niu 2020
Ji 2020
Tong 2019
Vesty 2018
Yeo 2019
0%
25%
50%
75%
100%
Relative abundance(%)
2
3
4
5
Shannon
0
100
200
300
400
500
Chao1
0.5
0.6
0.7
0.8
0.9
1.0
Simpson
Phylum
Actinobacteria
Bacteroidetes
Firmicutes
Fusobacteria
Proteobacteria
Saccharibacteria (TM7)
Spirochaetes
Other
Geographic Location
Malaysia
Australia
China
Europe
India
NorthAmerica
Qatar

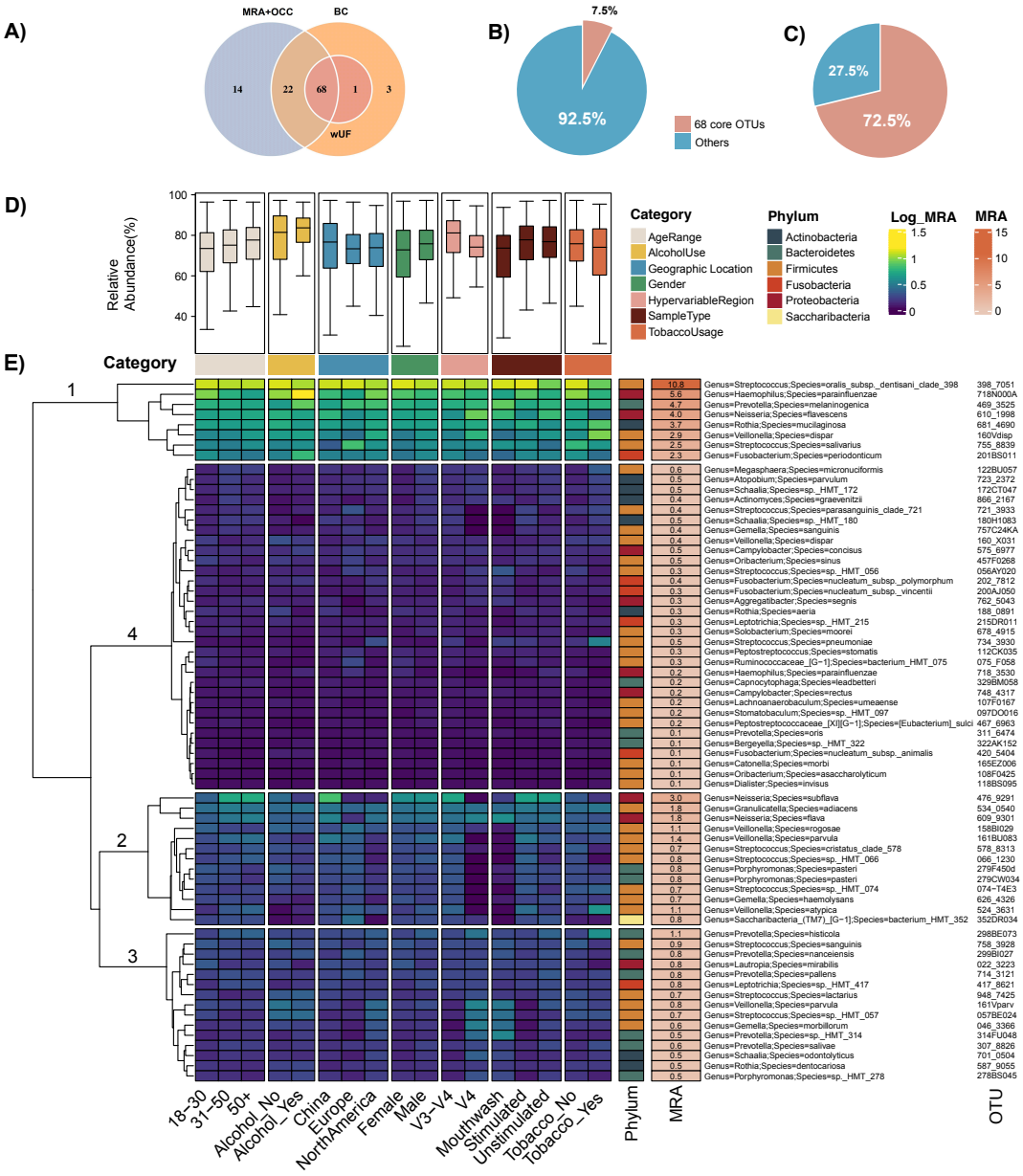
931

932


A) Family Level
C) Species Level
E) Bray−Curtis
B) Genus Level
D) OTU level
F) weighted UniFrac
Hypervariable Region
Geographic Location
Tobacco Usage
Sample Type
Alcohol Use
Gender
Age Range
R²
Axis.1 [13.3%]
Axis.2 [9.2%]
Axis.1 [36.5%]
Axis.2 [21.5%]
Combination
BC_RRA
BC_TSS
wUF_RRA
wUF_TSS
Euc_CLR
Euc_RRA
Euc_TSS
Geographic Location
China
Europe
NorthAmerica

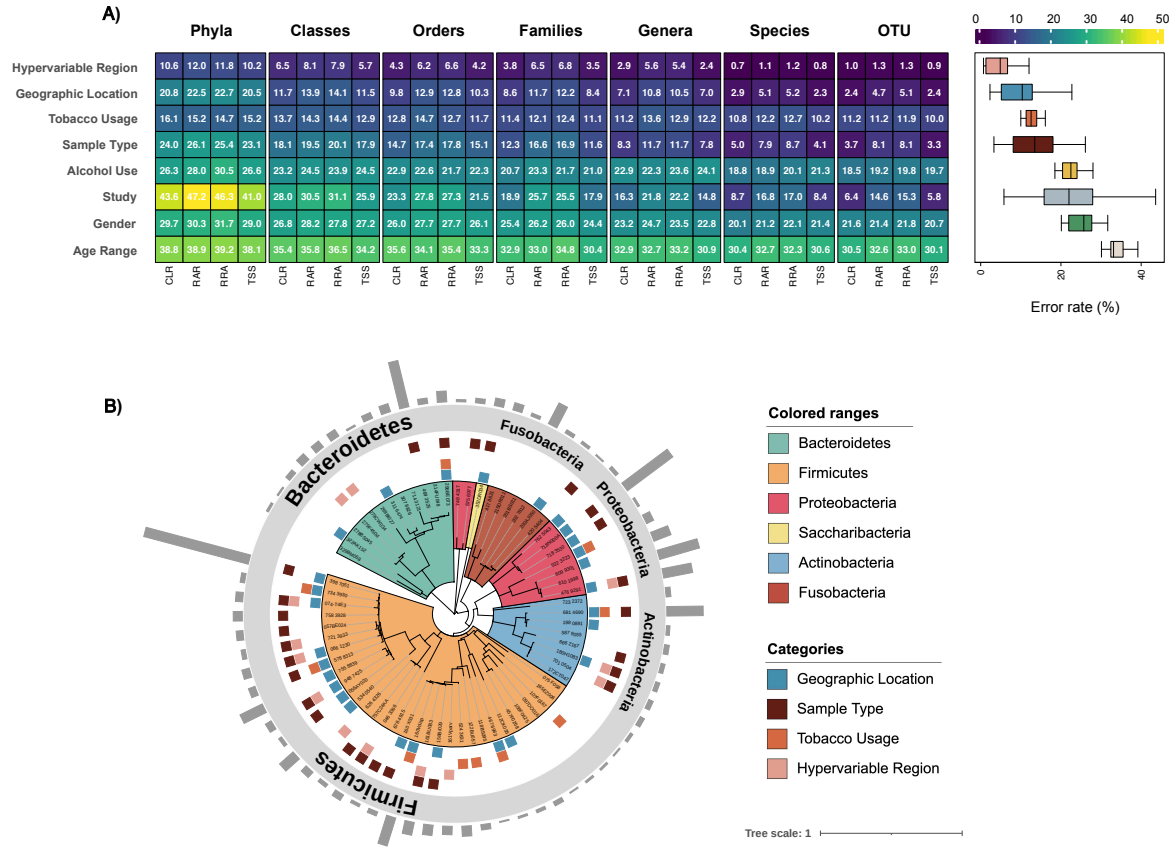
933

934


A)
MRA+OCC
BC
14
22
68
1
3
wUF
B)
7.5%
92.5%
68 core OTUs
Others
C)
27.5%
72.5%
D)
Relative Abundance(%)
100
80
60
40
Category
AgeRange
AlcoholUse
Geographic Location
Gender
HypervariableRegion
SampleType
TobaccoUsage
Phylum
Actinobacteria
Bacteroidetes
Firmicutes
Fusobacteria
Proteobacteria
Saccharibacteria
Log_MRA
1.5
1
0.5
0
MRA
15
10
5
0
E)
Category
1
4
2
3
18~30
31~50
50+
Alcohol_No
Alcohol_Yes
China
Europe
NorthAmerica
Female
Male
V3~V4
V4
Mouthwash
Stimulated
Unstimulated
Tobacco_No
Tobacco_Yes
Phylum
MRA
OTU

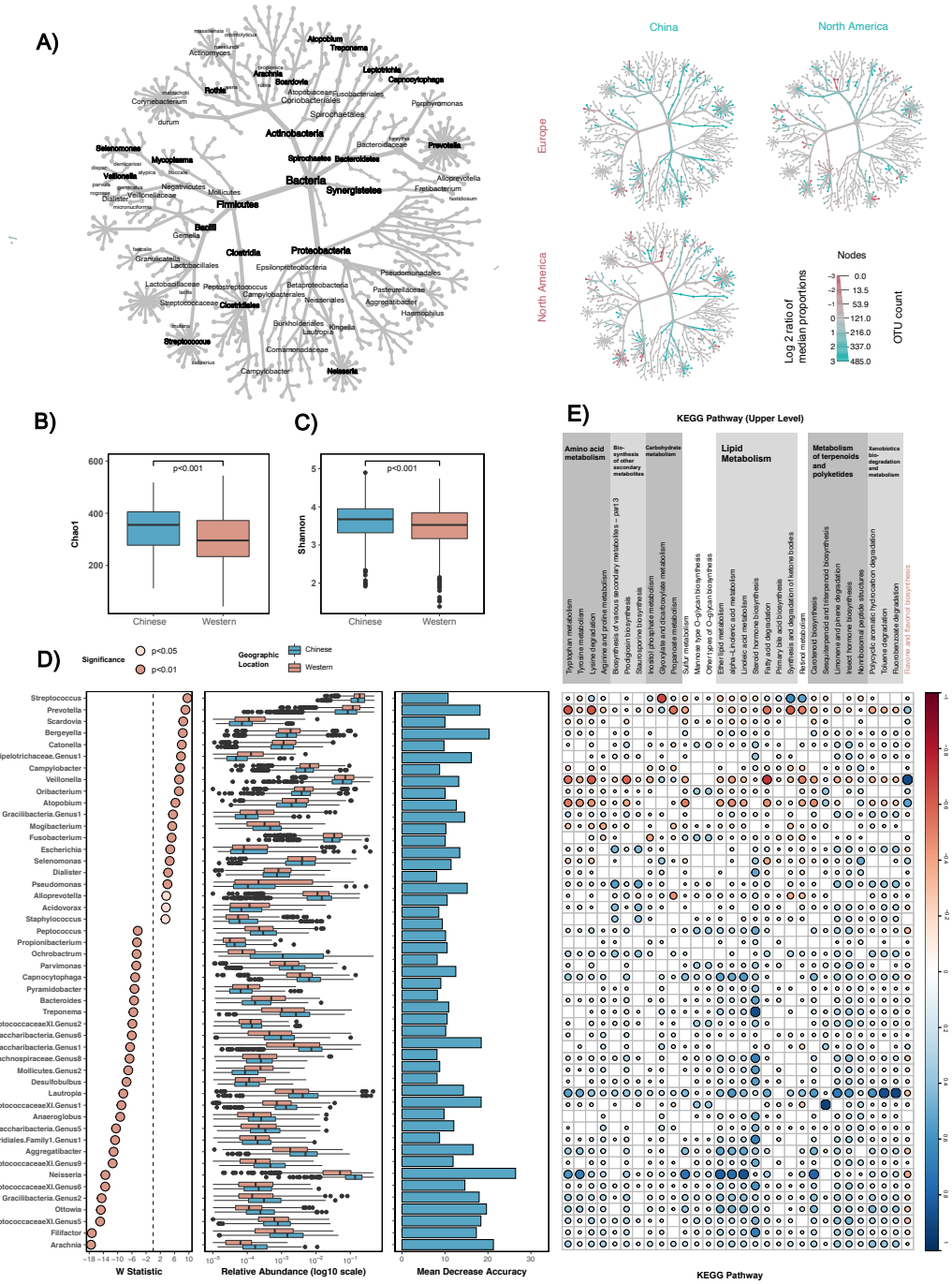
935

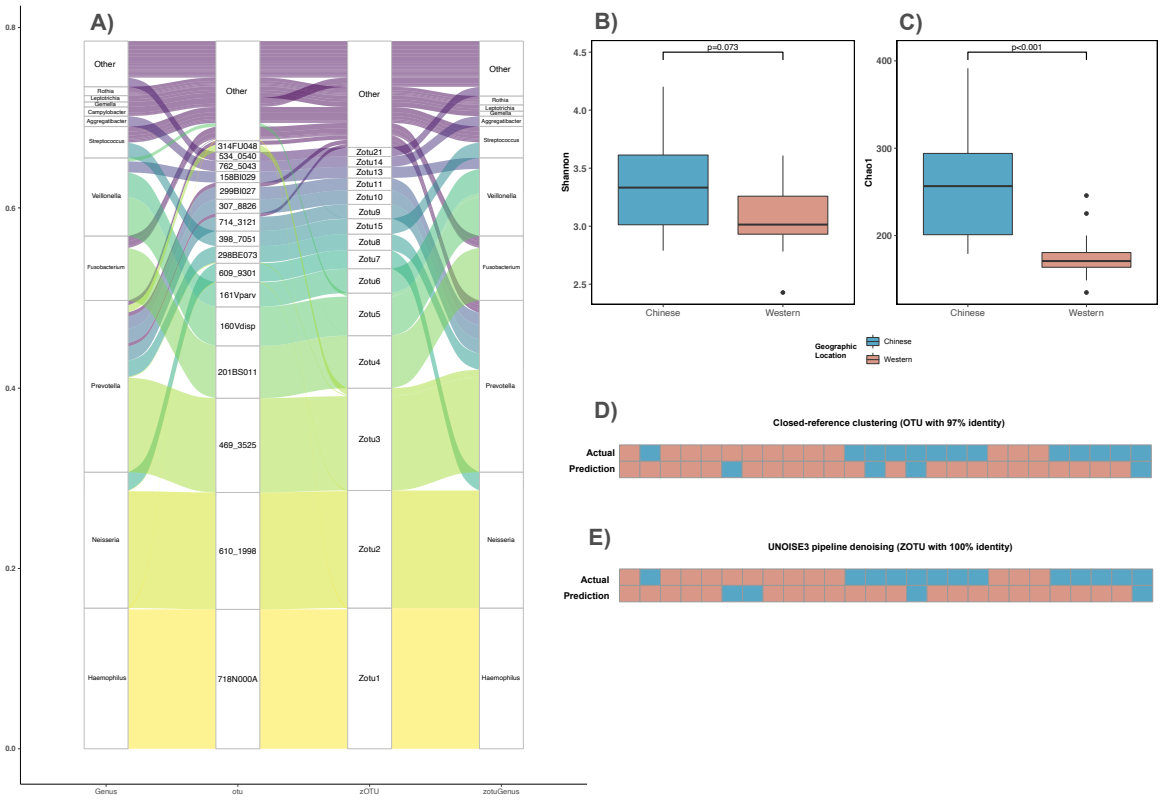
936


A)
Phyla
Classes
Orders
Families
Genera
Species
OTU
Hypervariable Region
Geographic Location
Tobacco Usage
Sample Type
Alcohol Use
Study
Gender
Age Range
CLR
RAR
RRA
TSS
0
10
20
30
40
50
Error rate (%)
B)
Bacteroidetes
Fusobacteria
Proteobacteria
Actinobacteria
Firmicutes
Colored ranges
Bacteroidetes
Firmicutes
Proteobacteria
Saccharibacteria
Actinobacteria
Fusobacteria
Categories
Geographic Location
Sample Type
Tobacco Usage
Hypervariable Region
Tree scale: 1

937

938


A)
China
North America
Europe
North America
Nodes
Log 2 ratio of median proportions
OTU count
B)
C)
D)
E)
Chao1
Shannon
Chinese
Western
p<0.001
Significance
p<0.05
p<0.01
Geographic Location
Chinese
Western
KEGG Pathway (Upper Level)
Amino acid metabolism
Lipid Metabolism
Metabolism of terpenoids and polyketides
W Statistic
Relative Abundance (log10 scale)
Mean Decrease Accuracy
KEGG Pathway


A)
B)
C)
D)
E)
Genus
otu
zOTU
zotuGenus
Other
Rothia
Leptotrichia
Gemella
Campylobacter
Aggregatibacter
Streptococcus
Veillonella
Fusobacterium
Prevotella
Neisseria
Haemophilus
314FU048
534_0540
762_5043
158BI029
299BI027
307_8826
714_3121
398_7051
298BE073
609_9301
161Vparv
160Vdisp
201BS011
469_3525
610_1998
718N000A
Zotu21
Zotu14
Zotu13
Zotu11
Zotu10
Zotu9
Zotu15
Zotu8
Zotu7
Zotu6
Zotu5
Zotu4
Zotu3
Zotu2
Zotu1
p=0.073
p<0.001
Shannon
Chao1
Chinese
Western
Geographic Location
Closed-reference clustering (OTU with 97% identity)
UNOISE3 pipeline denoising (ZOTU with 100% identity)
Actual
Prediction

941