

Salivary microbiota and metabolome associated with celiac disease

Ruggiero Francavilla¹, Danilo Ercolini², Maria Piccolo³, Lucia Vannini^{4,5}, Sonya Siragusa³,
 Francesca De Filippis², Ilaria De Pasquale³, Raffaella Di Cagno³, Michele Di Toma¹, Giorgia
 Gozzi^{4,5}, Diana I. Serrazanetti^{4,5}, Maria De Angelis^{3*}, Marco Gobetti³

¹*Department of Interdisciplinary Medicine, University of Bari Aldo Moro, Bari, Italy;*

²*Department of Agricultural Sciences, Division of Microbiology, University of Naples Federico II,
 Portici, Italy;*

³*Department of Soil, Plant and Food Sciences, University of Bari Aldo Moro, Bari, Italy;*

⁴*Department of Agricultural and Food Sciences, University of Bologna, Bologna, Italy;*

⁵*Inter-departmental Centre for Industrial Agri-Food Research, University of Cesena, Cesena, Italy.*

*Corresponding author details: Maria De Angelis, Via G. Amendola 165/a 70126 Bari, Italy

Tel. +39 0805442946, Fax +39 080 5442911, E-mail: maria.deangelis@uniba.it

25 **ABSTRACT**

26 This study aimed at investigating the salivary microbiota and metabolome of 13 celiac disease (CD)
27 children under gluten-free diet (treated celiac disease, T-CD). The same number of healthy children
28 (HC) was used as controls. The salivary microbiota was analyzed through an integrated approach of
29 culture-dependent and -independent methods. The metabolome analysis was carried out through
30 gas-chromatography mass spectrometry-solid-phase microextraction. Compared to HC, the number
31 of some cultivable bacterial groups (e.g., total anaerobes) significantly ($P < 0.05$) differed in the
32 saliva samples of T-CD. As shown by Community Level Catabolic Profiles, the highest substrate
33 utilization pattern and substrate richness were found in HC. Pyrosequencing data showed the
34 highest richness estimator and diversity index in HC. *Lachnospiraceae*, *Gemellaceae* and
35 *Streptococcus sanguinis* were the highest in T-CD. *Streptococcus thermophilus* markedly decreased
36 in T-CD. The salivary of T-CD children showed the highest amount of Bacteroidetes (e.g.,
37 *Porphyromonas* sp., *Porphyromonas endodontalis*, and *Prevotella nanceiensis*), together with the
38 lowest of Actinobacteria. T-CD children were also characterized by a decreased level of some
39 *Actinomyces* sp., *Atopobium* sp. and *Corynebacterium durum*. *Rothia mucilaginosa* was the only
40 Actinobacteria species found at the highest level in T-CD. As shown by multivariate statistical
41 analyses, the levels of organic volatile compounds markedly differentiated T-CD children. Some
42 compounds (e.g., ethyl-acetate, nonanal and 2-hexanone) were found associated with T-CD.
43 Correlations ($FDR < 0.05$) were found between the relative abundance of bacteria and some VOC.
44 The findings of this study indicated that CD is associated with oral dysbiosis that could affect the
45 oral metabolome.

46

47 INTRODUCTION

48 Celiac disease (CD) is a chronic immune-mediated enteropathy, which affects the small
49 intestinal mucosa after the ingestion of gluten from wheat, rye and their cross-related varieties in
50 genetically susceptible individuals (1). CD is one of the most diffuse chronic diseases in Europe and
51 North America, as well as in the developing countries (South America, South Africa and South
52 Asia) (2). Nowadays, the incidence of CD is increasing and it is estimated to be 0.7-2.0% (3).
53 Overall, CD is characterized by steatorrhea and malnutrition coupled with multiple deficiency
54 states. Diverse problems such as dental anomalies, short stature, osteopenic bone disease, lactose
55 intolerance, infertility, and nonspecific abdominal pain are associated with CD (4). Besides, the
56 alteration of the oral ecosystem and of the saliva composition seemed to be determined by CD (2, 5,
57 6). Compared to healthy subjects, the saliva of CD patients contains lower amounts of amylase, and
58 secretory IgA and IgM (7–8), and has lower buffering capacity, speed of salivary flow, and
59 concentration of calcium and Ca/P ratio (5, 9).

60 The gluten-free diet (GFD) is effective and safe. At the present, the exclusion of gluten-
61 containing products is the only treatment available. Recently, some reports also suggested that the
62 gastrointestinal (GI) microbiota is somewhat affected during CD pathogenesis and treatment with
63 GFD (10–11). Early microbial infections (12–13) and imbalances in the composition of the GI
64 microbiota (14–21) were associated with CD. Compared to healthy individuals, GFD lasting two
65 years did not complete restore the GI microbiota and, consequently, the metabolome of CD children
66 (14).

67 Doubtless, oral and GI microbiota play a key role on health and disease (22–28). The human oral
68 cavity is a complex ecosystem populated by ca. 700 bacterial species that reach numbers of 10^{11} /g
69 of wet weight dental plaque and 10^8 - 10^9 CFU/g of saliva (29). Some of these bacterial species are
70 closely associated with the development of oral diseases, mainly dental caries and periodontitis
71 (25–26). Also several non-oral diseases such as bacterial endocarditis (30), heart disease (31),

72 obesity (28), pneumonia (32), atherosclerosis (33) and preterm low birth weight (34) seemed to be
73 somehow related to bacteria from the human oral cavity. Overall, the saliva is considered to be the
74 most suitable tool to get information on the microbiota of the oral cavity (35). Saliva composition
75 includes volatile organic compounds (VOC) (metabolome) deriving from various sources (e.g.,
76 serum, blood, microorganisms, environmental pollution). Metabolites are considered indicators of
77 physiological or pathological states (36). Recently, saliva metabolomics analysis opens new
78 possibilities for the identification of metabolite biomarkers representative of specific disorders
79 (36–37).

80 Previously, few studies have dealt with the salivary microbiota of CD patients. It was shown that
81 a strict adherence to GFD reduces the prevalence of dental caries (38). Compared to healthy
82 individuals, CD children were characterized by a lower prevalence of salivary mutans streptococci
83 and lactobacilli (9). On the contrary, no statistically differences ($P < 0.05$) were found for
84 streptococci and lactobacilli counts between the saliva of HC and CD children, who were subjected
85 to GFD (39). A more in depth characterization of the oral microbiota of CD patients during GFD is
86 needed and none of the previous studies combined the characterization of the salivary microbiota
87 and metabolome.

88 This study aimed at comparing the salivary microbiota and metabolome of CD children
89 subjected to GFD and healthy children.

90 MATERIALS AND METHODS

91 **Study design.** This study was approved by the Institutional Review Board of the Faculty of
92 Medicine and Surgery of the University of Bari Aldo Moro (Italy). The informed written consent
93 was obtained from parents. Two groups of children (median age ca. 10 ± 1.4 years) were included in
94 the study: 32 symptom-free CD patients, who were treated with GFD for at least 2 years (T-CD);
95 and 37 healthy children (HC), without CD and other known food intolerance, who referred as the
96 controls. CD diagnosis was based on the positivity of TTG-IgA and EMA in the presence of clinical

97 symptoms and positive histological evidence of villous atrophy with crypt hyperplasia and increase
 98 in intra-epithelial lymphocytes on a gluten-containing diet, and the disappearance of symptoms with
 99 the normalization of positive serum-specific antibody(ies) on GFD (40) HC were selected after
 100 negative results of serological tests against CD.

101 Based on the exclusion criteria, 43 children were excluded after the first visit (Fig. S1). The
 102 information and characteristics of the recruited children were reported in Table 1.

103 **Collection of the saliva samples.** Unstimulated whole saliva samples were collected in the
 104 morning, 2 h after tooth brushing, by direct spitting into a sterile plastic tube in a time span not
 105 exceeding 30 min. No intakes of food and drink were allowed in the morning before sampling. The
 106 samples were either immediately subjected to analysis (plate counts and Biolog system) or frozen at
 107 -20°C (DNA extraction and metabolome analysis).

108 **Enumeration of cultivable bacteria.** Salivary samples (1 g) were mixed with 9 mL of sterilized
 109 physiological solution and homogenized. Counts of viable bacterial cells were carried out as
 110 described by De Angelis and co-workers (41). The following selective media were used: Plate count
 111 agar (total facultative aerobes and anaerobes); MRS agar (lactobacilli and enterococci);
 112 *Bifidobacterium* agar modified (bifidobacteria) (Becton Dickinson France SA, Le Pont de Claix,
 113 France); Glucose-M17 (lactococci and streptococci); Baird Parker plus Egg Yolk Tellurite
 114 Emulsion (stafilococci and micrococci); Wilkins-Chalgren anaerobe agar (total anaerobes);
 115 Wilkins-Chalgren anaerobe agar plus GN selective supplements and sheep blood defibrinated
 116 (*Bacteroides*, *Porphyromonas* and *Prevotella*); Violet red bile agar (Enterobacteria); and Slanetz
 117 and Bartley (Enterococci). Except for *Bifidobacterium* agar modified all media were purchased by
 118 Oxoid Ltd (Hampshire, England).

119 **Community-level catabolic profiles (CLCP).** Carbon source utilization patterns of the salivary
 120 microbiota were assessed using Biolog 96-well Eco micro-plates (Biolog, Inc., Hayward, CA) (42).
 121 Micro-plates contained 31 different carbon sources (carbohydrates, carboxylic acids, polymers,
 122 amino acids, amines, and miscellaneous substrates) in triplicate. Ten grams of saliva were

123 homogenized with 90 mL of sterile sodium chloride (0.9% [wt/vol]) solution (Classic Blender) and
 124 centrifuged at $12,500 \times g$ for 15 min at 4°C. The pellet was washed with 50 mM Tris-HCl (pH 7.0),
 125 then with sterile sodium chloride solution and centrifuged at $12,500 \times g$ for 15 min at 4°C. The
 126 cellular suspension was diluted (1:10) into the sterile sodium chloride solution and, subsequently,
 127 dispensed (150 μ L) into each of the 96 wells of the Biolog Eco micro-plates. The micro-plates were
 128 incubated at 30°C in the dark, and the color development was measured at 590 nm every 24 h with
 129 a micro-plate reader (Biolog Microstation). Three indices were determined (43–45). Shannon's
 130 diversity (H'), indicating the substrate utilization pattern, was calculated as follows: $H' = -\sum p_i \ln$
 131 (p_i), where p_i is the ratio of the activity of a particular substrate to the sums of activities of all
 132 substrates at 120 h. Substrate richness (S), measuring the number of different substrates used, was
 133 calculated as the number of wells with a corrected absorbance greater than 0.25. Substrate evenness
 134 (E) was defined as the equitability of activities across all utilized substrates: $E = H' / \log S$.

135 **16S rRNA gene amplicon library preparation and sequencing.** Total DNA extraction was
 136 carried out from the pellet of 2 mL of saliva, using a Biostic™ Bacteremia DNA isolation kit (Mo
 137 Bio Laboratories, Inc., Carlsbad, CA).

138 The microbial diversity was studied through pyrosequencing of the amplified V1–V3 region of
 139 the 16S rRNA gene. A fragment of 520 bp was amplified using primers and PCR conditions as
 140 previously described (46). 454-Adaptors were included in the forward primer followed by a 10 bp
 141 sample-specific Multiplex Identifier (MID). After agarose gel electrophoresis, PCR products were
 142 purified twice by Agencourt AMPure kit (Beckman Coulter, Milano, Italy), quantified using the
 143 QuantiFluor™ (Promega, Milano, Italy) and an equimolar pool was obtained prior to further
 144 processing. Duplicate PCR products were pooled for each sample. The amplicon pool was used for
 145 pyrosequencing on a GS Junior platform (454 Life Sciences, Roche, Italy), according to the
 146 manufacturer's instructions and using a Titanium chemistry.

147 **Bioinformatics and data analysis.** Raw reads were firstly filtered according to the 454
148 processing pipeline. Sequences were then analyzed using the QIIME 1.7.0 software (47). Raw reads
149 were demultiplexed and further filtered through the split_library.py script of QIIME. In order to
150 guarantee a higher level of accuracy, the reads were excluded from the analysis if they had an
151 average quality score lower than 25, if there were ambiguous base calls, if there were primer
152 mismatches and if they were shorter than 300 bp. Sequences that passed the quality filter were
153 denoised (48) and singletons were excluded. OTUs defined by a 97% of similarity were picked
154 using the uclust method (49) and the representative sequences, chosen as the most abundant in each
155 cluster, were submitted to the RDP-II classifier (50) to obtain the taxonomy assignment and the
156 relative abundance of each OTU using the Greengenes 16S rRNA gene database (51). Alpha and
157 beta diversity were evaluated through QIIME as recently described (52). Adonis and Anosim
158 statistical tests were performed through the compare_category.py script of QIIME, in order to verify
159 if there were differences among the two types of individuals. For categorical variables, ANOVA
160 and G test were carried out through the otu_category_significance.py script of QIIME in order to
161 test whether the presence/abundance of any OTUs was significantly associated with a specific
162 subject type or to the other variables, while Spearman correlations were computed between OTUs
163 and metabolite concentrations.

164 **Gas-chromatography mass spectrometry-solid-phase microextraction (GC-MS/SPME)**
165 **analysis of salivary volatile compounds.** After preconditioning, according to the manufacturer's
166 instructions, a carboxen/polydimethylsiloxane (CAR/PDMS) (85 µm) fiber and a manual solid
167 phase micro-extraction (SPME) holder (Supelco Inc., Bellefonte, PA, USA) were used. Before
168 headspace sampling, the fiber was exposed to GC inlet for 5 min for thermal desorption at 250°C.
169 Three grams of salivary sample were placed into 10 mL glass vials and added with 10 µL of 4-
170 methyl-2-pentanol (final concentration of 33 mg/L), as the internal standard. Samples were then
171 equilibrated for 10 min at 45°C. SPME fiber was exposed to each sample for 40 min. Both
172 equilibration and absorption phases were carried out with stirring. The fiber was then inserted into

the injection port of the gas chromatograph for 10 min of sample desorption. GC-MS analyses were carried out with an Agilent 7890A gas chromatograph (Agilent Technologies, Palo Alto, CA) coupled to an Agilent 5975C mass selective detector operating in an electron impact mode (ionization voltage, 70 eV). A Varian CP7773 Wax 52 CB capillary column (length, 50 m; inside diameter, 0.32 mm; Agilent Technologies) was used. The temperature program was 40°C for 1 min, followed by an increase, at a rate of 4.5°C/min, to 65°C, an increase, at a rate of 10°C/min, to 230°C, and then 230°C for 17 min. The injector, interface and ion source temperatures were 250, 250, and 230°C, respectively. The mass-to-charge ratio interval was 30 to 350 Da at a rate of 2.9 scans per sec. Injection was carried out in splitless mode, and helium (flow rate, 1 mL/min) was used as the carrier gas. Molecules were identified based on comparison of their retention times with those of pure compounds (Sigma-Aldrich, Milan, Italy). Identities were confirmed by searching mass spectra in the available databases (NIST, version 2005; Wiley, version 1996). All the GC-MS raw files were converted to netCDF format via Chemstation (Agilent Technologies) and subsequently processed with the XCMS toolbox (<http://metlin.scripps.edu/download/>). XCMS software allows automatic and simultaneous retention time alignment, matched filtration, peak detection and peak matching. The resulting table containing information such as peak index (retention time-m/z pair) and normalized peak area was exported into R (www.r-project.org) for subsequent statistical or multivariate analyses. Quantitative data for the compounds identified were obtained by the interpolation of the relative areas versus the internal standard area. GC-MS/SPME data were organized into matrix and analyzed by Canonical discriminant Analysis of Principal coordinates (41).

Statistical analysis. Culture dependent and metabolome data were obtained at least in triplicates. The analysis of variance (ANOVA) was carried out on transformed data followed by separation of means with Tukey's HSD, using the statistical software Statistica for Windows (Statistica 6.0 per Windows 1998, StatSoft). Letters indicate significant different groups ($P < 0.05$) by Tukey's test. Canonical discriminant Analysis of Principal Coordinates (CAP) analysis was also carried out for

GC-MS/SPME data. The hypothesis of non-significant differences in the multivariate location within groups was tested using the trace statistic based on 9999 permutations (53). The correlation between the concentration of metabolites and the number of predominant bacterial cells and genera was examined by linear regression analysis.

Nucleotide sequence accession numbers. The 16S rRNA sequences produced in this study are available at the sequence read archive of the NCBI (SRP0035361).

RESULTS

Enumeration of cultivable bacteria. Selective media were used to enumerate cultivable cells of the main microbial groups (Table 2). Compared to HC, the number of total anaerobes significantly ($P < 0.05$) decreased in the saliva of T-CD children. The other significant ($P < 0.05$) difference concerned the number of presumptive Enterobacteria. No significant ($P > 0.05$) differences were found between T-CD and HC for the other microbial groups.

Community-level catabolic profiles (CLCP). The substrate utilization pattern (H' index) and substrate richness (S index) were calculated (Fig. 1). Compared to HC, the H' and S indices significantly ($P < 0.05$) decreased in the saliva of T-CD children. The E index, which measures the statistical significance (equitability) of the values of H' and S indices, confirmed the above significant ($P < 0.05$) differences.

Richness and diversity of the salivary microbiota based on 16S rRNA gene sequencing data analysis. A total of 235,481 raw sequences was obtained and analyzed; 179,276 reads passed the filters applied through the QIIME split_library.py script, with an average value of 7,171 reads/sample and an average length of 499 bp calculated after primer removal. The Good's estimated sample coverage (median value of 98%; $P > 0.05$), indicated that a satisfactory coverage was reached for all the samples analyzed. The average number of species (OTU) identified in the salivary samples of T-CD children (median value of 319) significantly ($P = 0.005$) differed from

223 HC (median value of 393). Overall, six phyla (Firmicutes, Bacteroidetes, Proteobacteria,
 224 Fusobacteria, Actinobacteria and Tenericutes) and two candidate division (TM7 and SR1) were
 225 identified (Fig. 2 and Supplementary Table S1). Compared to HC, the saliva of T-CD children only
 226 differed in the levels of Bacteroidetes (40 vs 33%; $P = 0.001$) and Actinobacteria (5 vs 7%; $P =$
 227 0.023) (Fig. 2). The bacterial diversity among the microbial communities of T-CD and HC children
 228 was estimated also using the richness estimator (Chao1) and diversity index (Shannon) (Fig. 3).
 229 Compared to HC, Chao1 and Shannon diversity indices were lower ($P > 0.05$) in the saliva of T-CD
 230 children. The difference for the community structure was further confirmed using three phylogeny-
 231 based beta-diversity measures. The salivary microbiota of HC and T-CD was clearly differentiated
 232 based on the bacterial lineage-specific principal coordinate analysis with weighted Unifrac distance
 233 matrix (Fig. 4). In addition, both Adonis and Anosim statistical tests indicated a significant
 234 influence of subject type (HC vs T-CD) on the microbial diversity.

235 **Distinctive salivary microbiome associated with CD.** The differences ($P < 0.05$) for the
 236 relative abundance of OTUs associated with the saliva of T-CD or HC children are shown in
 237 Figure 5. The abundance of some OTUs belonging to Firmicutes phylum significantly differed
 238 between T-CD and HC ($P < 0.05$). *Streptococcus sanguinis*, *Gemella* sp. and *Lachnospiraceae*
 239 were more abundant in the saliva of T-CD children, while some lactic acid bacteria (e.g.,
 240 *Streptococcus thermophilus*) were at the highest levels in HC. *Granulicatella adiacens*,
 241 *Mogibacterium* sp., *Selenomonas* sp. and *Veillonella parvula* were also positively associated ($P <$
 242 0.05) with HC. Within the phylum Bacteroidetes, *Porphyromonas* sp., *Porphyromonas*
 243 *endodontalis*, and *Prevotella nanceiensis* showed the highest abundance in the saliva of T-CD
 244 children. *Actinomyces oris*, *Atopobium* sp. and *Corynebacterium durum* were less abundant
 245 Actinobacteria in the saliva of HC. The candidate division SR1 was more abundant in T-CD ($P <$
 246 0.05).

247 **Volatile organic compounds (VOC) profiling of saliva.** Overall, the content of the various
 248 metabolites largely varied within the same group (T-CD or HC) and the total median values of some
 249 VOC significantly ($P < 0.05$) differ between T-CD and HC. GC-MS/SPME data were analyzed
 250 through Canonical discriminant Analysis of Principal Coordinates (CAP). Compounds with
 251 negative values were those significantly ($P < 0.05$) associated with T-CD, while those on the
 252 positive axis were significantly ($P < 0.05$) associated with HC (Fig. 6A). The median values for
 253 alcohols and phenols (e.g., 2-ethyl-1-hexanol, 4-(1,1,3,3-tetramethylbutyl)-phenol and ethyl
 254 alcohol) were significantly ($P < 0.05$) higher in HC children (Table S2, Fig. 6A). Within aldehydes,
 255 nonanal was found at the highest level in T-CD as well as octanal was the highest in HC. The levels
 256 of butanoic acid 2-methyloctyl ester and acetic acid ethyl ester were the highest in HC and T-CD,
 257 respectively. Hydrocarbons were the largest group of VOC. Overall, hydrocarbons (e.g., 1-
 258 octadecene) were found at the highest level in HC children. On the other hand, halogenated and
 259 aromatic hydrocarbons (e.g., 1-chlorodecane and trichloromethane) were significantly associated
 260 with T-CD. Carbene disulphide was also associated with T-CD. With few exceptions, ketones,
 261 terpenes, thiophenes were found at the highest levels in the saliva of HC. According to the main
 262 VOC composition, the two groups of children were separated through CAP analysis (Fig. 6B).

263 **Correlation between microbiome and metabolome data.** Spearman's correlation analysis run
 264 through QIIME allowed finding some correlations between the relative abundance of salivary
 265 bacteria and metabolites (data not shown). A positive correlation ($FDR < 0.05$) was found between
 266 *P. endodontalis* and *Prevotella* sp. and the levels of nonanal and 1-chloro-decane. *C. durum* was
 267 positively correlated with 1,2,3-trimethyl-benzene, 2,6-dimethyl-4-heptanone, 4-methyl-2-hexanone
 268 and 4-methyl-3-penten-2-one as well as *G. adiacens*, *Atopobium* sp. and Bacilli correlated ($FDR <$
 269 0.05) with the levels of 1(3H)-isobenzofuranone (γ -lactone) and 1-octadecene. Also γ -lactone and
 270 1-chloro decane were correlated to *V. parvula* and *S. thermophilus*, and the candidate division SR1.

271 **DISCUSSION**

272 According to the World Health Organization, the oral human microbiota plays a crucial role for
273 health or disease status of the human host (54–56). Previously, variations of the composition of the
274 oral microbiota were described for CD patients, but the trend and the importance were not already
275 defined (9, 39).

276 To best of our knowledge, this study represents the largest effort to characterize the diversity, the
277 population structure and metabolome of the oral microbiota of CD children under GFD.

278 Some cultivable microbial populations differed between the saliva of T-CD and HC children. In
279 agreement with this study, total anaerobes bacteria were found to be at a higher level in the fecal
280 samples of HC compared to T-CD (14). According to culture-dependent data, the Community Level
281 Catabolic Profiles (CLCP) also varied between T-CD and HC. As previously shown for other
282 ecosystems (45, 57–58), CLCP successfully described the global metabolic activities of the salivary
283 microbiota. The lowest utilization pattern substrate (H' index) was found in the saliva of T-CD. The
284 capacity of using multiple substrates (e.g., carbohydrates, carboxylic acids, polymers, amino acids,
285 amines and miscellaneous compounds) was usually linked to the microbial diversity of a certain
286 microbial ecosystem. As shown by the Shannon-Weaver index, CLCP showed the highest microbial
287 diversity in HC children.

288 The analysis of pyrosequencing data showed that bacterial groups from the saliva of HC children
289 were similar to those found in previous studies (55–56). In agreement with CLCP profiles,
290 sequencing data showed the highest richness estimator (Chao1) and diversity index (Shannon) in
291 the saliva of HC children. The gut and oral microbiota may have common structures (29) and the
292 bacterial species richness of both these ecosystems usually decreases during inflammatory bowel
293 disease (IBD) (59–61). It was hypothesized that high microbial richness and diversity, which
294 characterized the healthy microbiota, may have a protective effect for humans (59). The
295 composition of the main bacterial phyla differed between the salivary microbiota of T-CD and HC
296 children. Compared to HC children, *Lachnospiraceae*, *Gemellaceae* (genus *Gemella* sp.) and *S.*

297 *sanguinis* were the most abundant in the saliva of T-CD. On the contrary, *S. thermophilus* (the main
 298 *Streptococcus* species), markedly decreased in T-CD. *S. sanguinis* was found associated to
 299 endocarditis and other distant site infections (30). *Streptococcus* is the main genus of the healthy
 300 oral microbiota (25, 59, 62). Similar results were found in the saliva of individuals affected by IBD
 301 (59). Besides, duodenal and fecal samples of CD patients (both at diagnosis and during GFD) were
 302 associated with the decrease of *Streptococcaceae* (63). Other Firmicutes (e.g., *V. parvula*), also
 303 associated with oral healthy (64), were found at the highest levels in the saliva of HC children. The
 304 saliva of T-CD children harbored the highest level of Bacteroidetes (e.g., *Porphyromonas* sp., *P.*
 305 *endodontalis*, and *Pr. nanceiensis*) and the lowest of Actinobacteria. The fecal and duodenal
 306 microbiota of CD patients was characterized by the highest numbers of Gram-negative bacteria
 307 (Bacteroidetes and Enterobacteria) and the lowest numbers of Gram-positive bacteria such as
 308 Firmicutes (e.g., lactic acid bacteria) and Actinobacteria (e.g., *Bifidobacterium*) (14, 20–21). Some
 309 *Porphyromonas* (e.g. *P. endodontalis*) were previously associated with periodontal diseases (65).
 310 Within the Actinobacteria phylum, the saliva of T-CD children was characterized by a decreased
 311 level of some *Actinomyces* (including *A. oris*), *Atopobium* sp. and *C. durum*. *Actinomyces* spp. are
 312 predominant Gram-positive members of the human oral commensal microbiota and are known as
 313 the initial colonizers of the tooth surfaces (66). *Rothia mucilaginosa* was the only Actinobacteria
 314 species found at the highest level in T-CD. *Rothia* species are components of the oral microbiota,
 315 and were also identified from duodenal biopsies (67). Although some infections by *R. mucilaginosa*
 316 were described, this species is considered a harmless colonizer of the oral cavity (68). Recently, it
 317 was hypothesized that some species of the *Rothia* genus, including *R. mucilaginosa*, were involved
 318 in the gluten degradation (69–70).

319 Overall, the human-associated microbiota interacts directly with the host by means of metabolic
 320 products (71–72). This study combined the characterization of the salivary microbiota and the
 321 related metabolome. The analysis of the saliva metabolome was successfully used in several fields

of physiology, diagnostics, functional genomics, pharmacology, toxicology and nutrition (36, 73). CAP analysis of the GC-MS/SPME metabolic profiles allowed the identification of VOC, which changed in the saliva of T-CD children. Many of the VOC described in this study corresponded to those previously reported in the saliva of healthy individuals (37,74). These compounds may have various origin, as there are many possible routes of entry into the salivary flow, which include environmental exposure through inhalation of air and/or water vapor through the lungs, ingestion through the mouth, food intake and transdermal absorption through the skin. Besides, the microbial metabolic activities at the level of oral cavity may also affect the synthesis of VOC (37, 74). First, this study showed some correlations between the microbiome and metabolome of CD children. It was hypothesized that the salivary microbiota affects the development of the gut microbiota (59). According to this hypothesis, high levels of nonanal, 4-methyl-2-hexanone (2-hexanone), ethyl-acetate (acetic acid ethyl ester) were found in both saliva (this study) and fecal samples (14, 75) of T-CD.

Some microbial indices (e.g., ratio between some Firmicutes and Actinobacteria vs. Bacteroidetes) and the levels of some metabolites (e.g., nonanal, ethyl-acetate, 2-hexanone) are signatures of CD patients. GFD lasting at least two years did not completely restore the salivary microbiota and, consequently, the salivary metabolome of T-CD children. Similar data were also found for duodenal (63) and fecal samples (14, 75). Further screenings on a wider sample size could provide informative sources for discovering CD-specific non-invasive salivary biomarkers. The limitation of this study was related to the low number of T-CD children analyzed. Further studies dealing with different severity of CD may be useful to highlight the correlations between the severity of CD and salivary microbiome. Different GFD could be analyzed to determine the effects of specific food on salivary microbiota of T-CD children. In addition, dietary implementation with probiotics could be regarded as alternative strategies to correct oral dysbiosis.

REFERENCES

- 347 1. **Tye-Din J, Anderson R.** 2008. Immunopathogenesis of celiac disease. *Curr. Gastroenterol.*
348 *Rep.* **10**:458–465.
- 349 2. **Mina S, Riga C, Azcurra AI, Brunotto M.** 2012. Oral ecosystem alterations in celiac
350 children: a follow-up study. *Arch. Oral Biol.* **57**:154–160.
- 351 3. **Lamireau T, Clouzeau H.** 2011. Epidemiology of celiac disease. *Pathol. Biol.* **61**(2):1–4.
- 352 4. **Murray JA.** 1999. The widening spectrum of celiac disease. *Am. J. Clin. Nutr.* **69**(3):354–
353 365.
- 354 5. **Mina SS, Azcurra AI, Dorronsoro S, Brunotto MN.** 2008. Alterations of the oral
355 ecosystem in children with celiac disease. *Acta Odontol. Latinoam.* **21**(2):121–126.
- 356 6. **Toscano V, Conti FG, Anastasi E, Mariani P, Tiberti C, Poggi M, Montuori M, Monti**
357 **S, Laureti S, Cipolletta E, Gemme G, Caiola S, Di Mario U, Bonamico M.** 2000.
358 Importance of gluten in the induction of endocrine autoantibodies and organ dysfunction in
359 adolescent celiac patients. *Am. J. Gastroenterol.* **95**:1742–1748.
- 360 7. **Lenander-Lumikari M, Ihalin R, Lähteenoja H.** 2000. Changes in whole saliva in
361 patients with coeliac disease. *Arch. Oral Biol.* **45**(5):347–354.
- 362 8. **Samaşca G, Iancu M, Farcău D, Butnariu A, Pop T, Pîrvan A, Andreica M, Miu N,**
363 **Cristea V, Dejica D.** 2011. IgA anti-tissue transglutaminase antibodies, first line in the
364 diagnosis of celiac disease. *Clin. Lab.* **57**(9-10):695–701.
- 365 9. **Acar S, Yetkiner AA, Ersin N, Oncag O, Aydogdu S, Arikan C.** 2012. Oral findings and
366 salivary parameters in children with celiac disease: a preliminary study. *Med. Princ. Pract.*
367 **21**:129–133.
- 368 10. **Sanz Y.** 2009. Novel perspectives in celiac disease therapy. *Mini Rev. Med. Chem.* **9**:359–
369 367.
- 370 11. **Niewinski MM.** 2008. Advances in celiac disease and gluten-free diet. *J. Am. Diet. Assoc.*
371 **108**:661–672.
- 372 12. **Stene LC, Honeyman MC, Hoffenberg EJ, Haas JE, Sokol RJ, Emery L, Taki I,**

- 373 **Norris JM, Erlich HA, Eisenbarth GS, Rewers M.** 2006. Rotavirus infection frequency
374 and risk of celiac disease autoimmunity in early childhood: a longitudinal study. *Am. J.*
375 *Gastroenterol.* **101**:2333–2340.
- 376 13. **Walters JR, Bamford KB, Ghosh S.** 2008. Coeliac disease and the risk of infections. *Gut.*
377 **57**:1034–1035.
- 378 14. **Di Cagno R, De Angelis M, De Pasquale I, Ndagijimana M, Vernocchi P, Ricciuti P,**
379 **Gagliardi F, Laghi L, Crecchio C, Guerzoni ME, Gobetti M, Francavilla R.** 2011.
380 Duodenal and faecal microbiota of celiac children: molecular, phenotype and metabolome
381 characterization. *BMC Microbiol.* **11**:219.
- 382 15. **Sánchez E, Nadal I, Donat E, Ribes-Koninckx C, Calabuig M, Sanz Y.** 2008. Reduced
383 diversity and increased virulence-gene carriage in intestinal enterobacteria of coeliac
384 children. *BMC Gastroenterol.* **8**:50. doi:10.1186/1471-230X-8-50.
- 385 16. **Sánchez E, Donat E, Ribes-Koninckx C, Calabuig M, Sanz Y.** 2010. Intestinal
386 *Bacteroides* species associated with coeliac disease. *J. Clin. Pathol.* **63**:1105–1111.
- 387 17. **Sánchez E, Ribes-Koninckx C, Calabuig M, Sanz Y.** 2012. Intestinal *Staphylococcus*
388 spp. and virulent features associated with coeliac disease. *J. Clin. Pathol.* **65**:830–834.
- 389 18. **Sanz Y, Sanchez E, Marzotto M, Calabuig M, Torriani S, Dellaglio F.** 2007.
390 Differences in faecal bacterial communities in coeliac and healthy children as detected by
391 PCR and denaturing gradient gel electrophoresis. *FEMS Immunol. Med. Microbiol.*
392 **51**:562–568.
- 393 19. **Schippa S, Iebba V, Barbato M, Di NG, Totino V, Checchi MP, Longhi C, Maiella G,**
394 **Cucchiara S, Conte MP.** 2010. A distinctive ‘microbial signature’ in celiac pediatric
395 patients. *BMC Microbiol.* **10**:175. doi:10.1186/1471-2180-10-175.
- 396 20. **Collado MC, Donat E, Ribes-Koninckx C, Calabuig M, Sanz Y.** 2009. Specific
397 duodenal and faecal bacterial groups associated with pediatric coeliac disease. *J. Clin.*
398 *Pathol.* **62**:264–269.

- 399 21. **Nadal I, Donat E, Ribes-Koninckx C, Calabuig M, Sanz Y.** 2007. Imbalance in the
400 composition of the duodenal microbiota of children with coeliac disease. *J. Med.*
401 *Microbiol.* **56**:1669–1674.
- 402 22. **Petrof EO, Claud EC, Gloor GB, Allen-Vercoe E.** 2013. Microbial ecosystems
403 therapeutics: a new paradigm in medicine? *Benef. Microbes.* **4**:53–61.
- 404 23. **Blaser MJ.** 2010. Harnessing the power of the human microbiome. *Proc. Natl. Acad. Sci.*
405 *U S A.* **107**:6125–6126.
- 406 24. **Sekirov I, Russel SL, Antunes LCM, Finlay BB.** 2010. Gut microbiota in health and
407 disease. *Physiol. Rev.* **90**:859–904.
- 408 25. **Bik ME, Long CD, Armitage GC, Loomer P, Emerson J, Mongodin EF, Nelson KE,**
409 **Gill SR, Fraser-Liggett CM, Relman DA.** 2010. Bacterial diversity in the oral cavity of
410 ten healthy individuals. *ISME J.* **4**:962–974.
- 411 26. **Ling Z, Kong J, Jia P, Wei C, Wang Y, Pan Z, Huang W, Li L, Chen H, Xiang C.**
412 2010. Analysis of oral microbiota in children with dental caries by PCR-DGGE and
413 barcoded pyrosequencing. *Microb. Ecol.* **60**(3):677–690.
- 414 27. **Francavilla R, Calasso M, Calace L, Siragusa S, Ndagijimana M, Vernocchi P,**
415 **Brunetti L, Mancino G, Tedeschi G, Guerzoni E, Indrio F, Laghi L, Miniello V,**
416 **Gobbetti M, De Angelis M.** 2012. Effect of lactose on gut microbiota and metabolome of
417 infants with cow's milk allergy. *Pediatr. Allergy Immu.* **23**:420–427.
- 418 28. **Piombino P, Genevose A, Esposito S, Moio L, Cutolo PP, Chambery A, Severino V,**
419 **Moneta E, Smith DP, Owens SM, Gilbert JA, Ercolini D.** 2014. Saliva from obese
420 individuals suppresses the release of aroma compounds from wine. *PLoS ONE.* **9**:e85611.
- 421 29. **Maukonen J, Matto J, Suihko ML, Saarela M.** 2008. Intra-individual diversity and
422 similarity of salivary and faecal microbiota. *J. Med. Microbiol.* **57**:1560–1568.
- 423 30. **Lockhart PB, Durak DT.** 1999. Oral microflora as a cause of endocarditis and other
424 distant site infections. *Infect Dis. Clin. North Am.* **13**:833–850.

- 425 31. **Beck JD, Eke P, Heiss G, Manadios P, Couper D, Lin D, Moss K, Elter J, Offenbacher**
426 **S.** 2005. Periodontal disease and coronary heart disease: a reappraisal of the exposure.
427 *Circulation*. **112**:19–24.
- 428 32. **Paju S, Scannapeico FA.** 2007. Oral biofilm, periodontitis, and pulmonary infections. *Oral*
429 *Dis*. **13**:508–512.
- 430 33. **Koren O, Spor A, Felin J, Fak F, Stombaugh J, Tremaroli V, Behre CJ, Knight R,**
431 **Fagerberg B, Ley RE, Backhed F.** 2011. Human oral, gut, and plaque microbiota in
432 patients with atherosclerosis. *Proc. Natl. Acad. Sci. USA*. **108**:4592–4598.
- 433 34. **Bogges KA, Beck JD, Murtha AP, Moss K, Offenbacher S.** 2006. Maternal periodontal
434 disease in early pregnancy and risk for a small-for-gestational-age infant. *Am. J. Obstet.*
435 *Gynecol*. **194**:1316–1322.
- 436 35. **Takeshita T, Nakano Y, Kumagai T, Yasui M, Kamio N, Shibata Y, Shiota S,**
437 **Yamashita Y.** 2009. The ecological proportion of indigenous bacterial populations in
438 saliva is correlated with oral health status. *ISME J*. **3**:65–78.
- 439 36. **Zhang A, Sun H, Wang X.** 2012. Saliva metabolomics opens door to biomarker discovery,
440 disease diagnosis, and treatment. *Appl. Biochem. Biotechnol*. **168**:1718–1727.
- 441 37. **AL-Kateb H, de Lacy Costello B, Ratcliffe N.** 2013. An investigation of volatile organic
442 compounds from the saliva of healthy individuals using headspace-trap/GC-MS. *J. Breath*
443 *Res*. **7**:1–13.
- 444 38. **Avsar A, Kalayci AG.** 2008. The presence and distribution of dental enamel defects and
445 caries in children with celiac disease. *Turk. J. Pediatr*. **50**:45–50.
- 446 39. **Shteyer E, Berson T, Lachmanovitz O, Hidas A, Wilschanski M, Menachem M,**
447 **Shachar E, Shapira J, Steinberg D, Moskovitz M.** 2013. Oral health status and salivary
448 properties in relation to gluten free diet in children with celiac disease. *JPGN*.
449 doi:10.1097/MPG.0b013e31828b3705.
- 450 40. **Husby S, Koletzko S, Korponay-Szabó IR, Mearin ML, Phillips A, Shamir R,**

- Troncone R, Giersiepen K, Branski D, Catassi C, Lelgeman M, Mäki M, Ribes-Koninckx C, Ventura A, Zimmer KP; ESPGHAN Working Group on Coeliac Disease Diagnosis; ESPGHAN Gastroenterology Committee; European Society for Pediatric Gastroenterology, Hepatology, and Nutrition. 2012. European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. *J Pediatr Gastroenterol Nutr.* **54**:136–160.
41. De Angelis M, Piccolo M, Vannini L, Siragusa S, De Giacomo A, Serrazzanetti DI, Cristofori F, Guerzoni ME, Gobbetti M, Francavilla R. 2013. Fecal microbiota and metabolome of children with autism and pervasive developmental disorder not otherwise specified. *PLoS ONE.* **8**(10):e76993.
42. Crecchio C, Gelsomino A, Ambrosoli R, Minati JL, Ruggiero P. 2004. Functional and molecular responses of soil microbial communities under differing soil management practices. *Soil Biol. Biochem.* **36**:1873–1883.
43. Shannon, CE. 1948. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**:379–423.
44. Shannon, CE. 1948. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**:623–656.
45. Zak JC, Willing MR, Moorhead DL, Wildman HG. 1994. Functional diversity of microbial communities: a quantitative approach. *Soil Biol. Biochem.* **26**:1101–1108.
46. Ercolini D, De Filippis F, La Stora A, Iacono M. 2012. “Remake” by high throughput sequencing of the microbiota involved in the production of water buffalo mozzarella cheese. *Appl. Environ. Microbiol.* **78**:8142–8145.
47. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J,

- 477 **Knight R.** 2010. QIIME allows analysis of high-throughput community sequencing data.
478 *Nat. Methods.* **7**:335–336.
- 479 48. **Reeder J, Knight R.** 2010. Rapidly denoising pyrosequencing amplicon reads by
480 exploiting rank-abundance distributions. *Nat. Methods.* **7**:668–669.
- 481 49. **Edgar RC.** 2010. Search and clustering orders of magnitude faster than BLAST.
482 *Bioinformatics.* **26**:2460–2461.
- 483 50. **Wang Q, Garrity GM, Tiedje JM, Cole JR.** 2007. Naïve Bayesian classifier for rapid
484 assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.*
485 **73**:5261–5267.
- 486 51. **Mc Donald D, Price MN, Goodrich J, Nawrocki EP, De Santis TZ, Probst A,**
487 **Andersen GL, Knight R, Hugenholtz P.** 2012. An improved Greengenes taxonomy with
488 explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.*
489 **6**:610–618.
- 490 52. **De Filippis F, La Storia A, Villani F, Ercolini D.** 2013. Exploring the sources of
491 beefsteaks contamination by culture-independent high-throughput sequencing. *PLoS ONE.*
492 **8**:e70222.
- 493 53. **Ndagijimana M, Laghi L, Vitali B, Placucci G, Brigidi P, Guerzoni ME.** 2009. Effect
494 of a synbiotic food consumption on human gut metabolic profiles evaluated by ¹H Nuclear
495 Magnetic Resonance spectroscopy. *Int. J. Food Microbiol.* **134**:147–153.
- 496 54. **Petersen PE.** 2003. The World Oral Health Report 2003: Continuous improvement of oral
497 health in the 21st century, The approach of the WHO global oral health programme.
498 *Community Dent. Oral Epidemiol.* **31**:3–23.
- 499 55. **Ling Z, Xia L, Wang Y, Li L, Luo Y, Yuan L, Nelson KE, Xiang C.** 2013.
500 Pyrosequencing analysis of the human microbiota of healthy Chinese undergraduates.
501 *BMG Genomics.* **14**:390.
- 502 56. **Ling Z, Liu X, Wang Y, Li L, Xiang C.** 2012. Pyrosequencing analysis of the salivary

- microbiota of healthy chinese children and adults. *Microb. Ecol.* **65**:487–495.
57. **Siragusa S, Di Cagno R, Ercolini E, Minervini F, Gobbetti M, De Angelis M.** 2009. Taxonomic structure and monitoring of the dominant population of lactic acid bacteria during wheat flour sourdough type I propagation using *Lactobacillus sanfranciscensis* starters. *Appl. Environ. Microbiol.* **75**(4):1099–1109.
58. **Minervini F, De Angelis M, Di Cagno R, Pinto D, Siragusa S, Rizzello CG, Gobbetti M.** 2010. Robustness of *Lactobacillus plantarum* starters during daily propagation of wheat flour sourdough type I. *Food Microbiol.* **27**:897–908.
59. **Said SH, Susa W, Nakagome S, Chinen H, Oshima K, Kim S, Kimura R, Iraha A, Fujita J, Mano S, Morita H, Dohi T, Oota H, Hattori M.** 2013. Dysbiosis of salivary microbiota in inflammatory bowel disease and its association with oral immunological biomarkers. *DNA Res.* 1–11 doi:10.1093/dnares/dst037.
60. **Manichanh C, Rigottier-Gois L, Bonnaud E, Gloux K, Pelletier E, Frangeul L, Nalin R, Jarrin C, Chardon P, Marteau P, Roca J, Dore J.** 2006. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut.* **55**:205–211.
61. **Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermúdez-Humarán LG, Gratadoux JJ, Blugeon S, Bridonneau C, Furet JP, Corthier G, Granette C, Vasquez N, Pochart P, Trugnan G, Thomas G, Blottière HM, Doré J, Marteau P, Seksik P, Langella P.** 2008. *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc. Natl Acad. Sci.* **105**:16731–16736.
62. **Kang JG, Kim SH, Ahn TY.** 2006. Bacterial diversity in the human saliva from different ages. *J. Microbiol.* **44**:572–576.
63. **Sánchez E, Donat E, Ribes-Koninckx C, Fernández-Murga ML, Sanz Y.** 2013. Duodenal-mucosal bacteria associated with celiac disease in children. *Appl. Environ. Microbiol.* **79**:5472–5479.

- 529 64. **Kumar PS, Griffen AL, Moeschberger ML, Leys EJ.** 2005. Identification of candidate
530 periodontal pathogens and beneficial species by quantitative 16S clonal analysis. *J. Clin.*
531 *Microbiol.* **43**:3944–3955.
- 532 65. **Wade WG.** 2013. The oral microbiome in health and disease. *Pharm. Res.* **69**:137–143.
- 533 66. **Kolenbrander PE, Eglund PG, Diaz PI, Palmer RJ Jr.** 2005. Genome-genome
534 interactions: bacterial communities in initial dental plaque. *Trends Microbiol.* **13**:11–15.
- 535 67. **Ou G, Hedberg M, Horstedt P, Baranov V, Forsberg G, Drobni M, Sandström O, Wai**
536 **SN, Johansson I, Hammarström ML, Hernell O, Hammarström S.** 2009. Proximal
537 small intestinal microbiota and identification of rod-shaped bacteria associated with
538 childhood celiac disease. *Am. J. Gastroenterol.* **104**:3058–3067.
- 539 68. **Yamane K, Nambu T, Yamanaka T, Mashimo C, Sugimori C, Leung KP, Fukushima**
540 **H.** 2010. Complete genome sequence of *Rothia mucilaginosa* DY-18: a clinical isolate with
541 dense meshworklike structures from a persistent apical periodontitis lesion. Sequencing.
542 doi:10.1155/2010/457236.
- 543 69. **Zamakhchari M, Wei G, Dewhirst F, Lee J, Schuppan D, Oppenheim FG,**
544 **Helmerhorst EJ.** 2011. Identification of *Rothia* bacteria as gluten-degrading natural
545 colonizers of the upper gastro-intestinal tract. *PLoS ONE.* **6**(9):e24455.
546 doi:10.1371/journal.pone.0024455.
- 547 70. **Fernandez-Feo M, Ewi G, Blumenjranz G, Dewhirst FE, Schuppan D, Oppenheim**
548 **FG, Helmerhorst EF.** 2013. The cultivable human oral gluten-degrading microbiome and
549 its potential implications in celiac disease and gluten sensitivity. *Clin. Microbiol. Infec.*
550 **19**:e386–e394.
- 551 71. **The Human Microbiome Consortium.** 2012. Structure, function and diversity of the
552 human microbiome in an adult reference population. *Nature.* doi:10.1038/nature11234.
- 553 72. **Want EJ, Nordstrom A, Morita H, Siuzdak G.** 2007. From exogenous to endogenous:
554 the inevitable imprint of mass spectrometry in metabolomics. *J. Proteome Res.* **6**:459–468.

- 555 73. **Winder CL, Cornmell R, Schuler S, Jarvis RM, Stephens GM, Goodacre R.** 2011.
556 Metabolic fingerprinting as a tool to monitor whole-cell biotransformations. *Anal. Bioanal.*
557 *Chem.* **399**:387–401.
- 558 74. **Kusano M, Mendez E, Furton KG.** 2013. Comparison of the volatile organic compounds
559 from different biological specimens for profiling potential. *J. Forensic Sci.* **58**(1):29–39.
- 560 75. **Di Cagno R, Rizzello CG, Gagliano F, Ricciuti P, Ndagijimana M, Francavilla R,**
561 **Guerzoni ME, Crecchio C, Gobbetti M, De Angelis M.** 2009. Different fecal
562 microbiotas and volatile organic compounds in treated and untreated children with celiac
563 disease. *Appl. Environ. Microbiol.* **75**:3963–3971.
- 564
- 565

566 **Legend to figures**

567 **Figure 1.** Community Level Catabolic Profile (CLCP) indices (utilization pattern substrate, H' ;
568 substrate richness, S ; and equitability, E) of the salivary microbiota of celiac, treated with the
569 gluten-free diet for at least 2 years (T-CD), and healthy children (HC). Data are the means of three
570 independent experiments ($n=3$). The center line of each box represents the median ($\square$), the top and
571 bottom of the box represent the 75th and 25th percentile of the data, respectively. The top and bottom
572 of the error bars represent the 5th and 95th percentile of the data, respectively. The circles in each
573 box plot extend to the outliers of the data.

574 **Figure 2.** Relative abundance (%) of total bacteria, which were found at the phylum level in the
575 saliva of celiac, treated with the gluten-free diet for at least 2 years (T-CD), and healthy children
576 (HC).

577 **Figure 3.** Average number of species (OTU), richness (Chao 1) and diversity (Shannon index)
578 values in the saliva of celiac, treated with the gluten-free diet for at least 2 years (T-CD), and
579 healthy children (HC). The center line of each box represents the median ($\square$), the top and bottom of
580 the box represent the 75th and 25th percentile of the data, respectively. The top and bottom of the
581 error bars represent the 5th and 95th percentile of the data, respectively. The circles and asterisks in
582 each box plot extend to the outliers and extremes of the data, respectively.

583 **Figure 4.** Principle Coordinate Analysis (PCoA) based on weighted UniFrac analysis of all 16S
584 rRNA gene sequences found in the saliva of celiac, treated with the gluten-free diet for at least 2
585 years (T-CD), and healthy children (HC).

586 **Figure 5.** Relative proportions (%) of predominant bacteria, showing significant ($P < 0.05$)
587 differences, between the saliva of celiac, treated with the gluten-free diet for at least 2 years (T-
588 CD), and healthy children (HC).

589 **Figure 6.** Canonical discriminant analysis of principal coordinates (CAP) loading coefficient plot
590 (A) and score plot (B) of the volatile organic compounds found in saliva of celiac, treated with the
591 gluten-free diet for at least 2 years (T-CD), and healthy children (HC).

592

593

594 **Figure S1.** Participants involved in the trial from the assessment for eligibility through follow-up.

595

596

597 **Table 1.** Baseline demographic and clinical characteristics of children.

	Celiac children (n = 13)	Healthy controls (n = 13)
Mean age $\pm$ SD (years)	9.7 $\pm$ 1.4	10.2 $\pm$ 1.4
Male/female	4/9	5/8
Feeding habits	Gluten free diet	Unrestricted
Duration of gluten free diet (years)	4.6 $\pm$ 2.2	/
BMI centile (%)	63.5 $\pm$ 25.2	58.3 $\pm$ 31.24
Tooth brushing habits	2 times a day	2 times a day
Hemoglobin (g/dl)	13.6 $\pm$ 0.7	13.1 $\pm$ 1.2
Iron (ng/ml)	76.3 $\pm$ 25.1	65.2 $\pm$ 31.4
Ferritin (ng/ml)	33.6 $\pm$ 14.5	28.8 $\pm$ 15.4

598

599

600 **Table 2.** Median values and range of cultivable cells (log CFU/mL) of the main microbial groups in
601 the saliva of celiac, treated with the gluten-free diet for at least 2 years (T-CD), and healthy children
602 (HC).

Microbial group	T-CD	HC	P
Total aerobic bacteria	7.29 (6.22 – 8.02)	7.55 (6.92 – 7.89)	0.057
Total anaerobes	7.62 (6.00 – 8.36)	7.81 (7.00 – 8.93)	0.029
<i>Enterococcus</i> and <i>Lactobacillus</i>	7.23 (5.85 – 8.51)	7.42 (6.06 – 8.02)	0.27
<i>Lactococcus</i> and <i>Streptococcus</i>	7.02 (6.00 – 7.90)	7.05 (6.14 – 7.88)	0.156
<i>Staphylococcus</i> and <i>Micrococcus</i>	5.70 (5.10 – 6.13)	5.50 ^a (4.87 – 7.18)	0.178
<i>Bacteroides</i> , <i>Porphyromonas</i> and <i>Prevotella</i>	6.10 ^a (5.09 – 7.01)	6.07 ^a (5.02 – 6.90)	0.074
Enterobacteriaceae	3.66 (1.50 – 4.59)	1.00 (0 – 5.59)	0.042
<i>Bifidobacterium</i>	7.32 (6.04 – 8.43)	7.18 (6.90 – 8.03)	0.285

603

604 Data are the means of three independent experiments (n = 3) for each child.

Figure 1

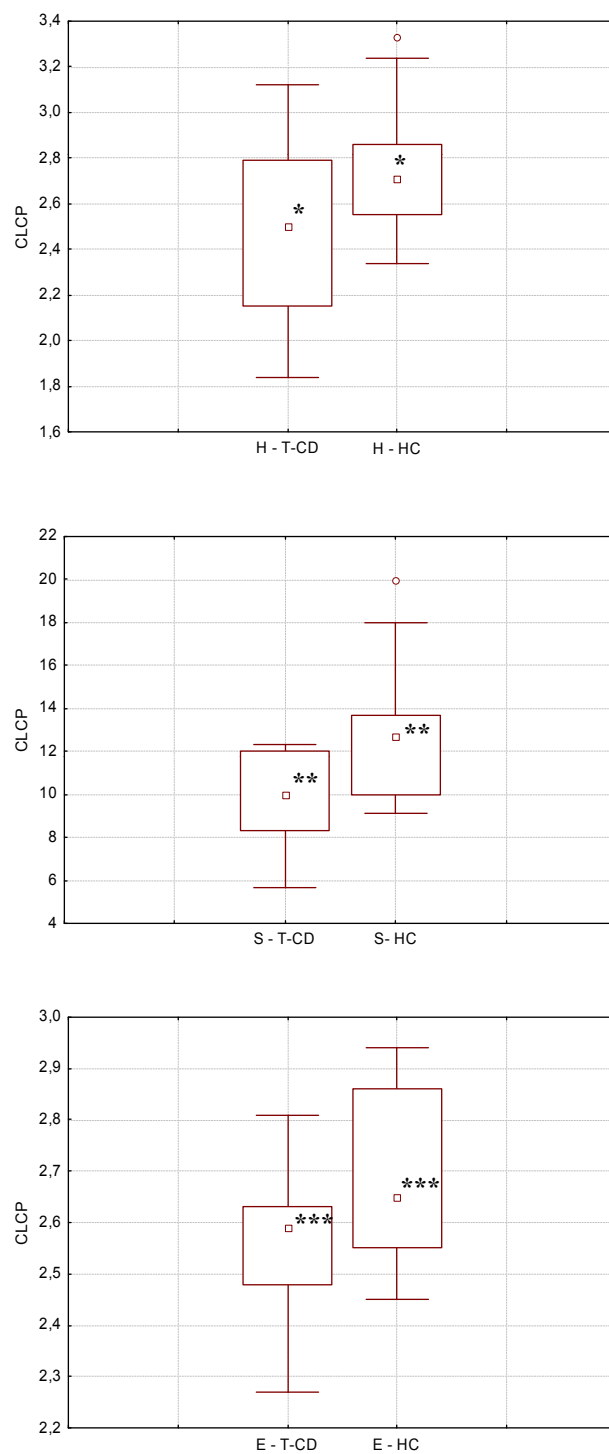


Figure 2

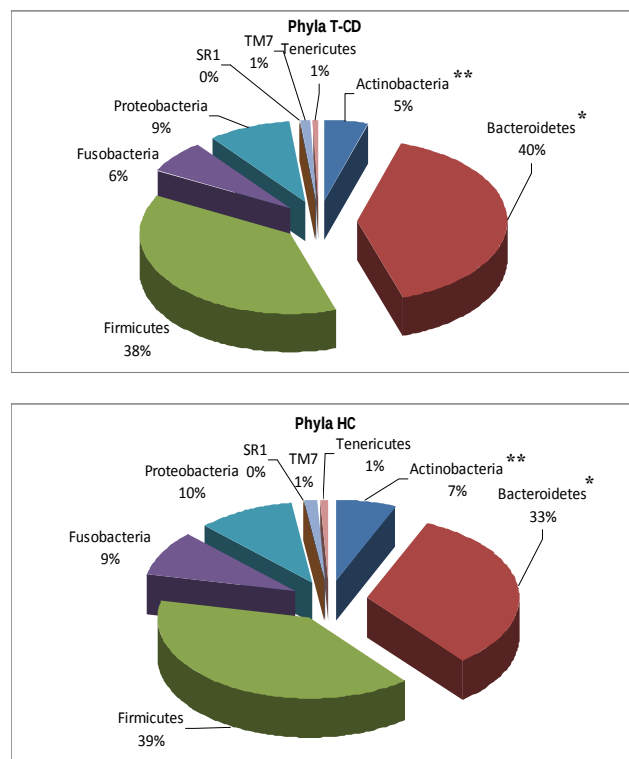


Figure 3

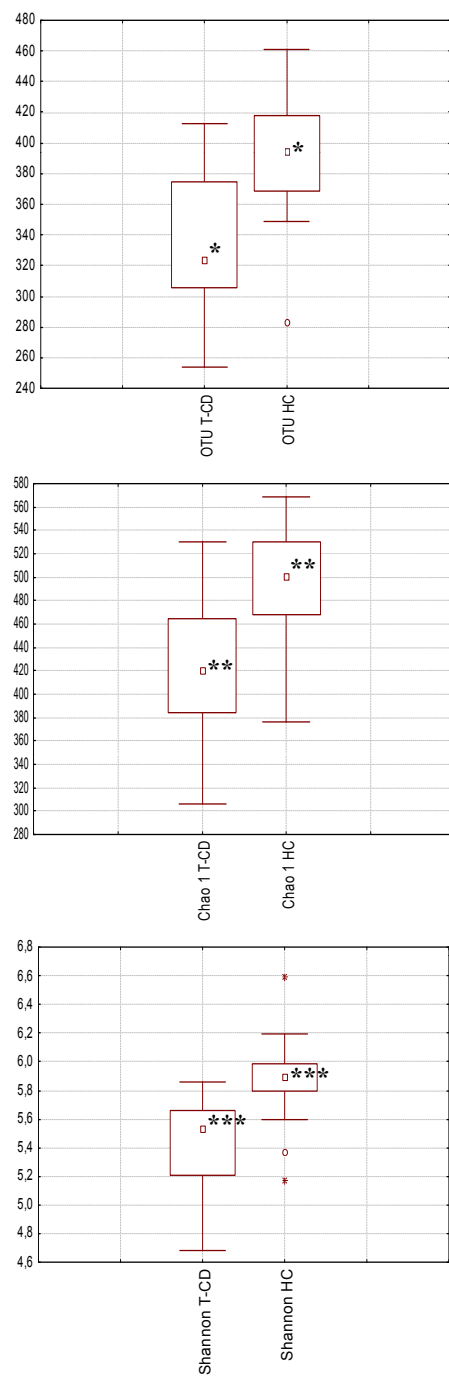


Figure 4

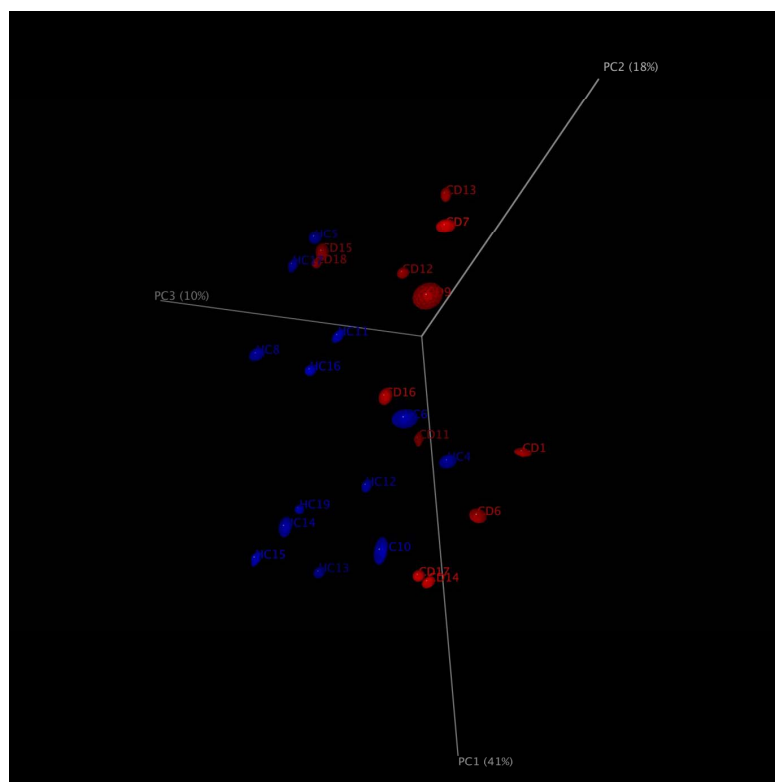


Figure 5

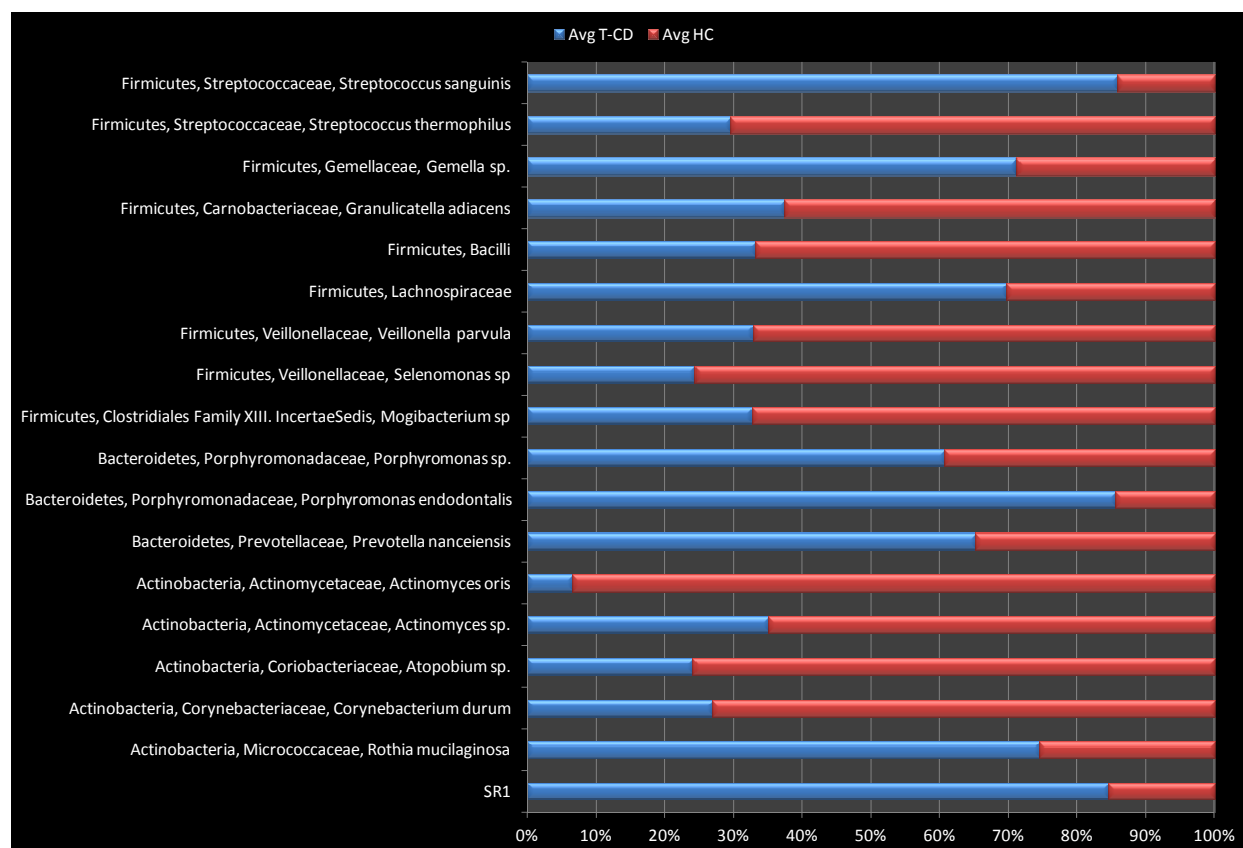


Figure 6A

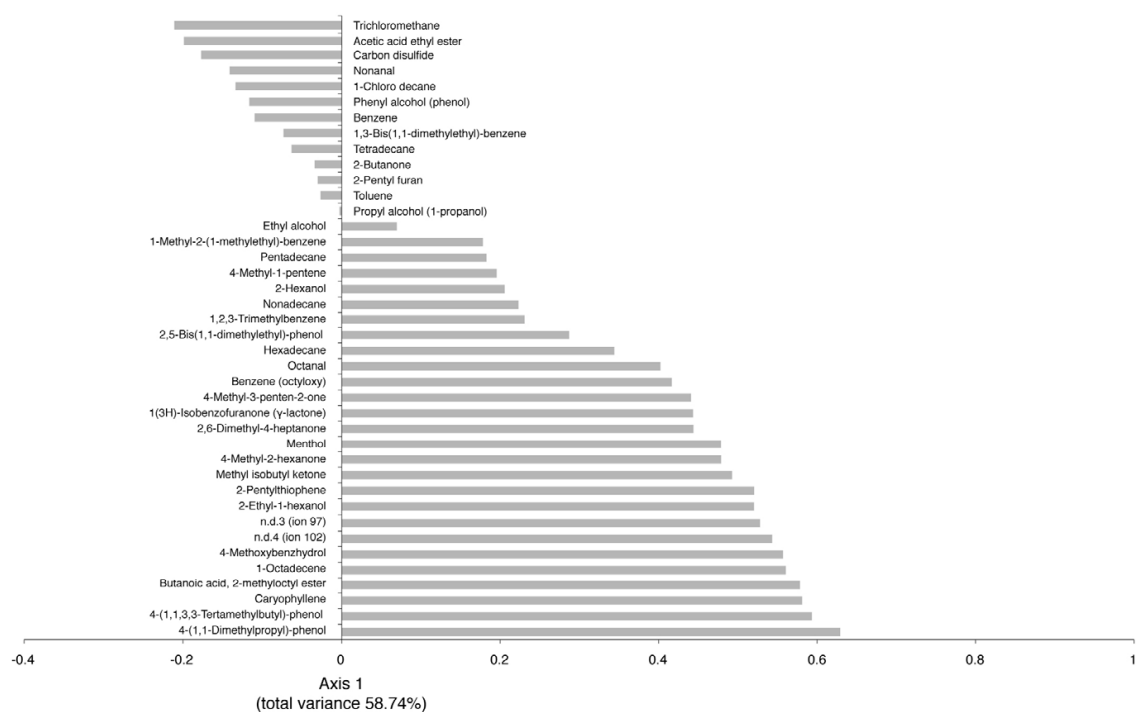


Figure 6B

