CPBLUE, Volume *1*

Supplemental information

Strain transmission links human microbiomes along the oral-gut axis and across cohabiting individuals

Vitor Heidrich, Gloria Fackelmann, Liviana Ricci, Roan Spadazzi, Gabriel Baldanzi, Michal Punčochář, Giulia Catassi, Paolo Marchi, Monica Modesto, Gianmarco Piccinno, Serena Porcari, Debora Rondinella, Francesco Asnicar, Mireia Valles-Colomer, Paola Mattarelli, Gianluca Ianiro, and Nicola Segata

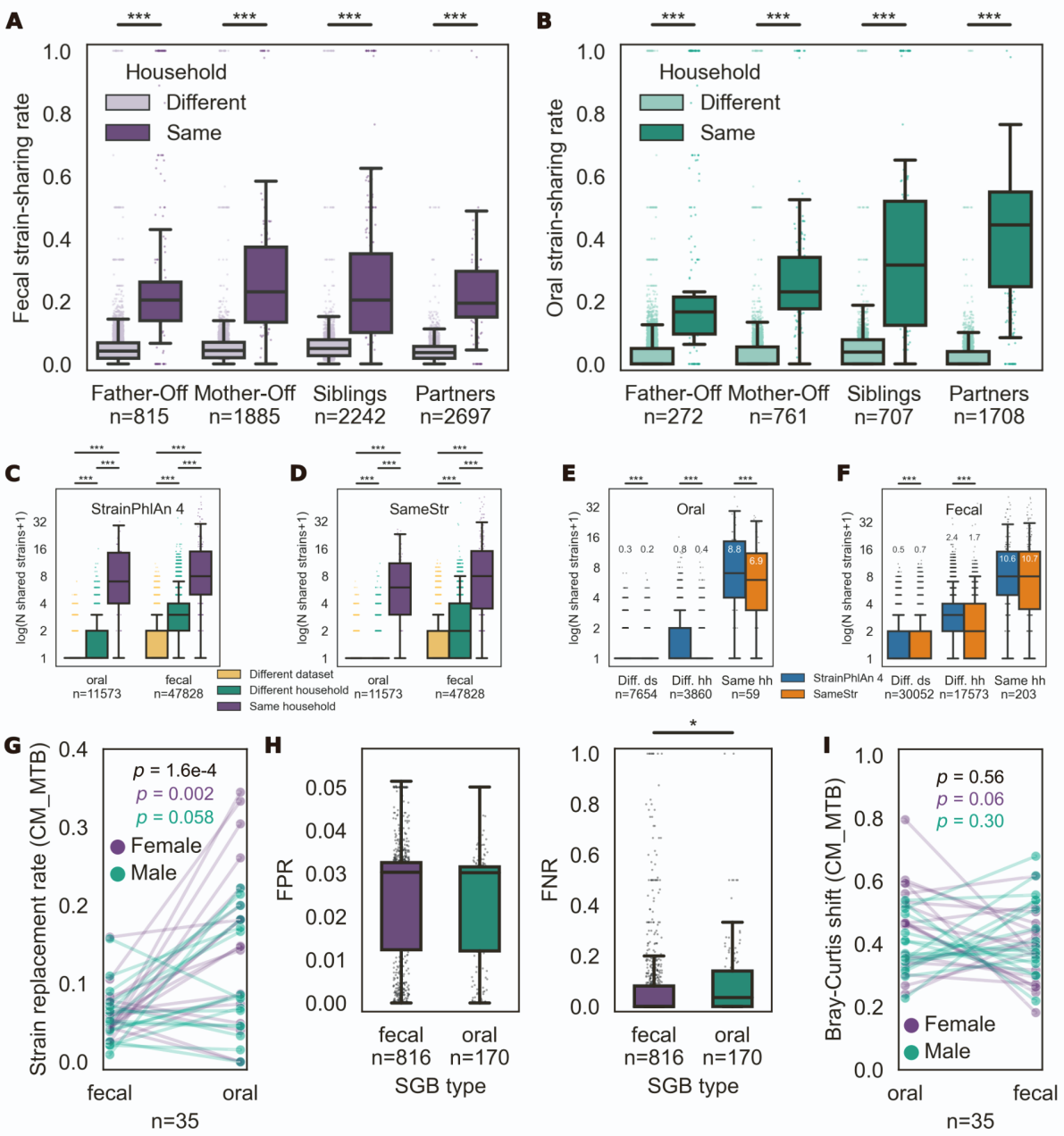

Figure S1. Additional validation of increased oral and gut microbiome strain-sharing rates among cohabitants.

(A) Gut microbiome strain-sharing rates for same vs. different households for different relationship types within households.

(B) Oral microbiome strain-sharing rates for same vs. different households for different relationship types within households.

(C) Number of strains shared between individuals from same versus different households versus different datasets for each body site according to our StrainPhlAn 4-based methodology.

(D) Number of strains shared between individuals from same versus different households versus different datasets for each body site according to our SameStr (**Methods**). Mann-Whitney U test was used.

(E) Comparison between the number of oral strains shared as reported by our StrainPhlAn 4-based methodology versus SameStr for each comparison category.

(F) Comparison between the number of fecal strains shared as reported by our StrainPhlAn 4-based methodology versus SameStr for each comparison category.

(G) Gut vs. oral microbiome strain replacement rate (**Methods**) among females (purple), males (green), or overall (black) in the CM_MTB dataset. Individuals undergoing antibiotic treatment

between sampling timepoints were excluded.
(H) False positive rate (FPR, left) and false negative rate for fecal vs. oral SGBs, defined based on the body site in which they have higher prevalence.
In A–F and H, the Mann-Whitney U test was used with statistical significance represented as: *, $p < 0.05$; ***, $p < 0.001$. The *n* below x-tick labels indicates the number of individual pairs (A–F) or SGBs (H).
(I) Gut vs. oral microbiome Bray-Curtis dissimilarity between consecutive timepoints among females (purple), males (green), or overall (black) in the CM_MTB dataset. Individuals undergoing antibiotic treatment between sampling timepoints were excluded.
In G and I, the Wilcoxon signed-rank test was used. The *n* below x-tick labels indicates the number of individuals.

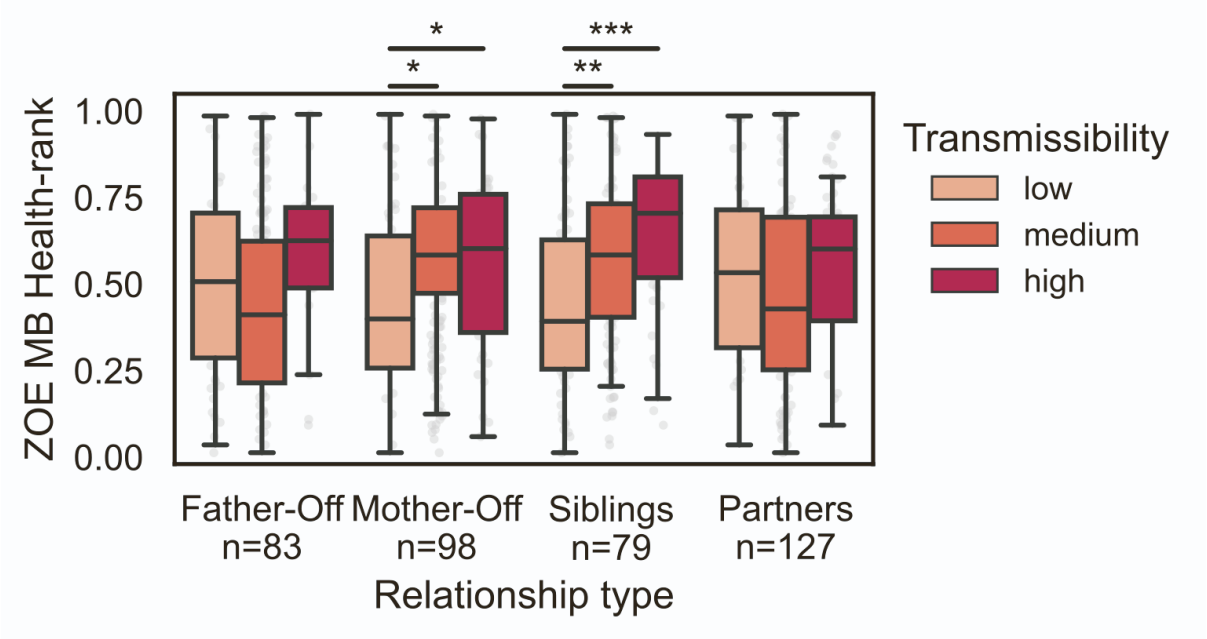

Figure S2. Association between ZOE MB Health-ranks and household transmissibility. ZOE MB health-ranks of SGBs with low vs. medium vs. high transmissibility (Methods) for different relationship types within households. The Mann-Whitney U test was used with statistical significance represented as: *, $p < 0.05$; **, $p < 0.05$, ***, $p < 0.001$. The *n* below x-tick labels indicates the number of SGBs.

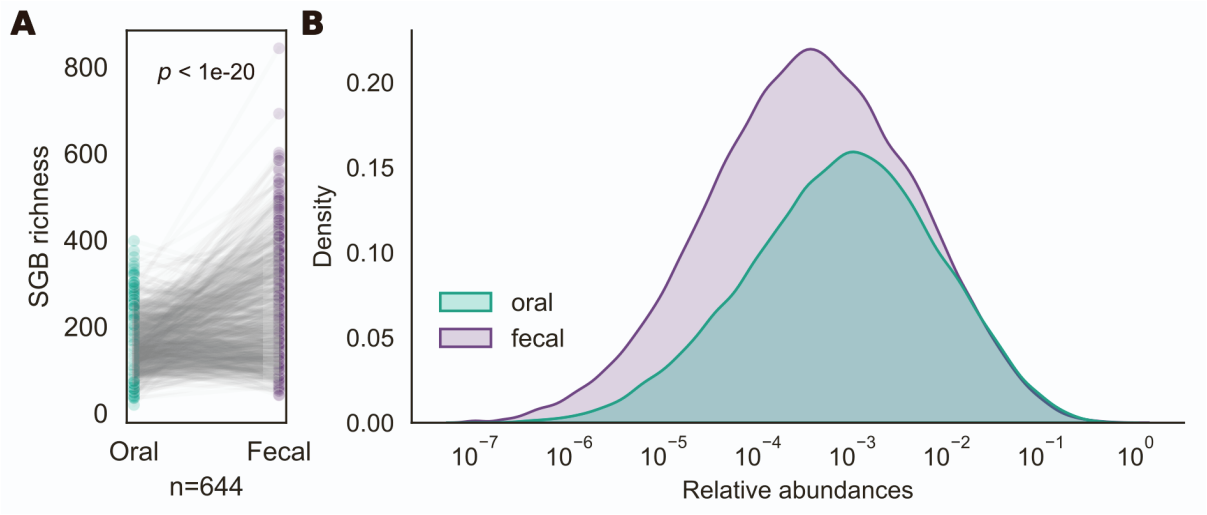

Figure S3. Comparison between oral and gut microbiome diversities.
(A) Oral vs. gut microbiome SGB richness for 644 individuals with paired samples available. The Wilcoxon signed-rank test was used.
(B) Distribution of SGB relative abundance values for oral vs. gut microbiomes.

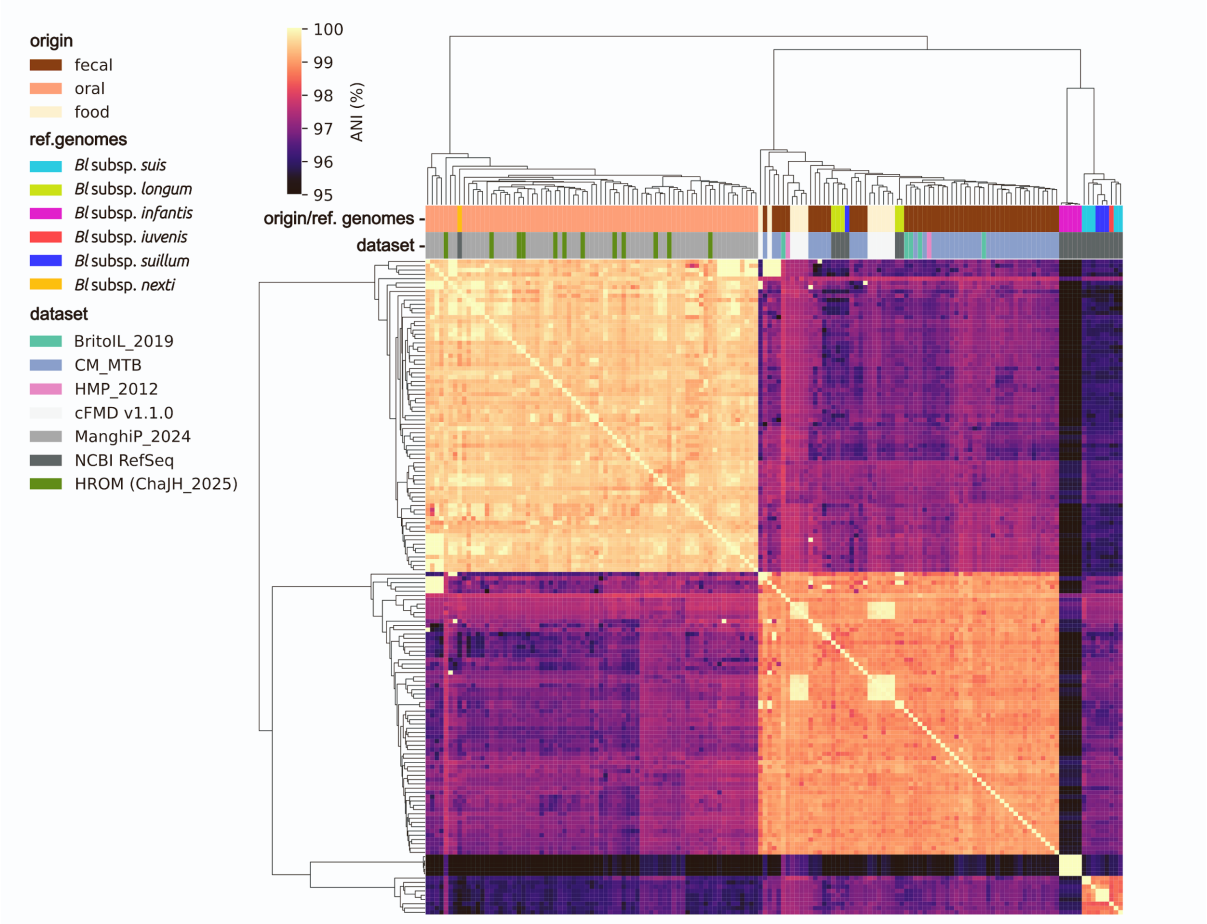

Figure S4. Genome analysis of *Bifidobacterium longum*.
All-vs-all clustermap representation of average nucleotide identities (ANIs) between *Bifidobacterium longum* (SGB17248) genomes reconstructed from metagenomes (i.e. metagenome-assembled genomes) and isolate genomes representative of each *B. longum* subspecies.

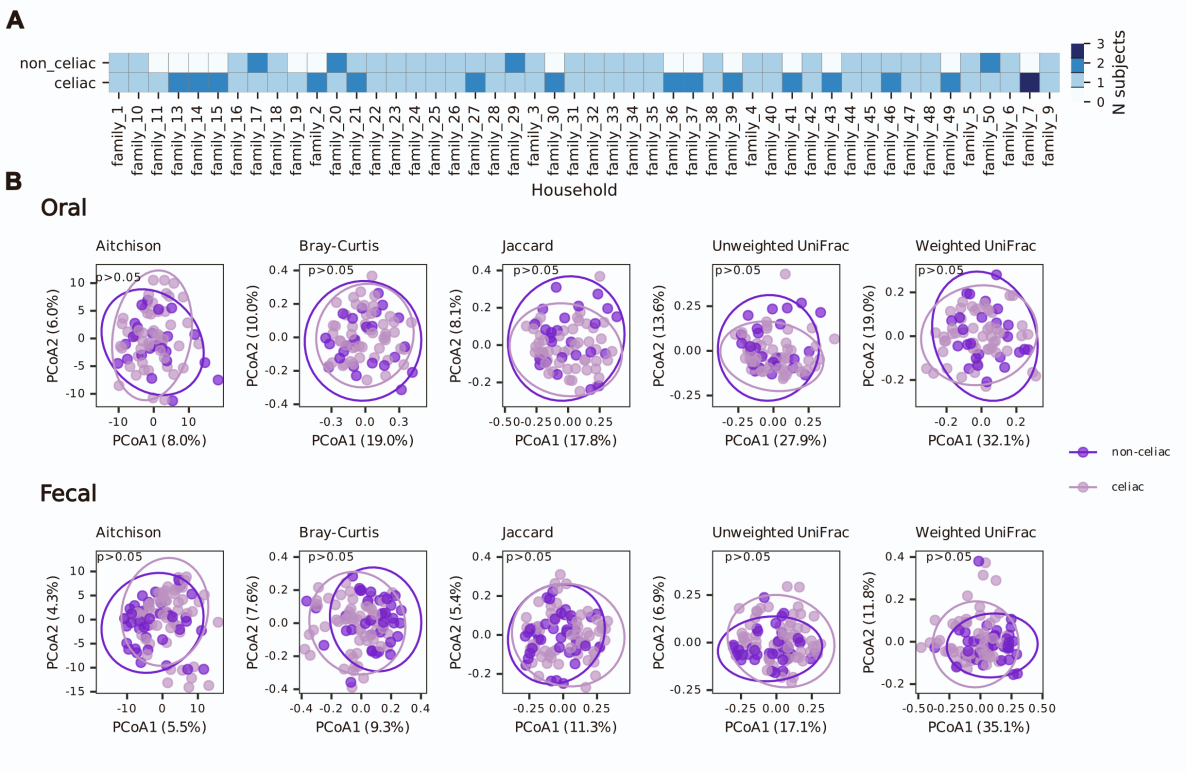

Figure S5. Evaluation of differences between celiac and non-celiac individuals from the MTC cohort.

(A) Tilemap representing the number of individuals with and without celiac disease for each household of the CM_MTC dataset included in this study.

(B) Principal Coordinate Analysis (PCoA) of oral (top row) and gut microbiomes (bottom row) in the CM_MTC dataset. Multiple beta-diversity metrics (indicated as subtitles) are shown. A multivariate PERMANOVA test was used (**Methods**) to assess differences between groups.

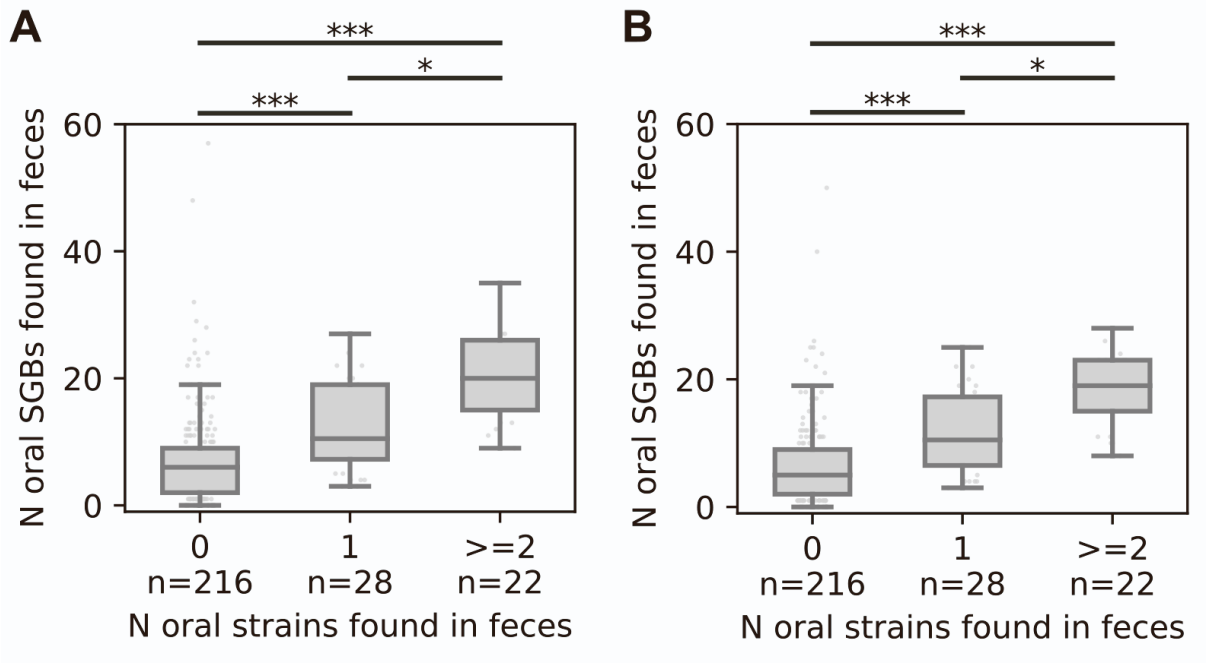

Figure S6. Sensitivity analysis of the oral-typical SGBs list definition.
(A) Number of oral-typical SGBs found in feces in cases in which 0 versus 1 versus ≥ 2 oral-derived strains are found in feces, using an alternatively generated oral-typical SGBs list considering a 10% prevalence cutoff for oral samples (**Methods**).
(B) Number of oral-typical SGBs found in feces in cases in which 0 versus 1 versus ≥ 2 oral-derived strains are found in feces, using an alternatively generated oral-typical SGBs list considering a 30% prevalence cutoff for oral samples (**Methods**).
In A and B, the post-hoc Dunn test for Kruskal-Wallis was used, with statistical significance represented as: **, $p < 0.01$; ***, $p < 0.001$. The *n* below x-tick labels indicates the number of individuals, restricting the analysis to those not used in the definition of oral-typical SGBs.