Cell Reports, Volume *45*

Supplemental information

Rapid microbiome restructuring associated with medical exposure in remote Amazonian Indigenous communities

Haipeng Sun, Oscar Noya-Alarcon, Jincheng Wang, Scott G. Daniel, Kyle Bittinger, Wessam Mohamed, Pauline D. Scanlan, Charvi Shah, Tanima Kundu, Martin J. Blaser, Noel T. Mueller, Orlana Lander, Mariapia Bevilaqua, Daniela Vargas Robles, Monica Contreras, and Maria G. Dominguez-Bello

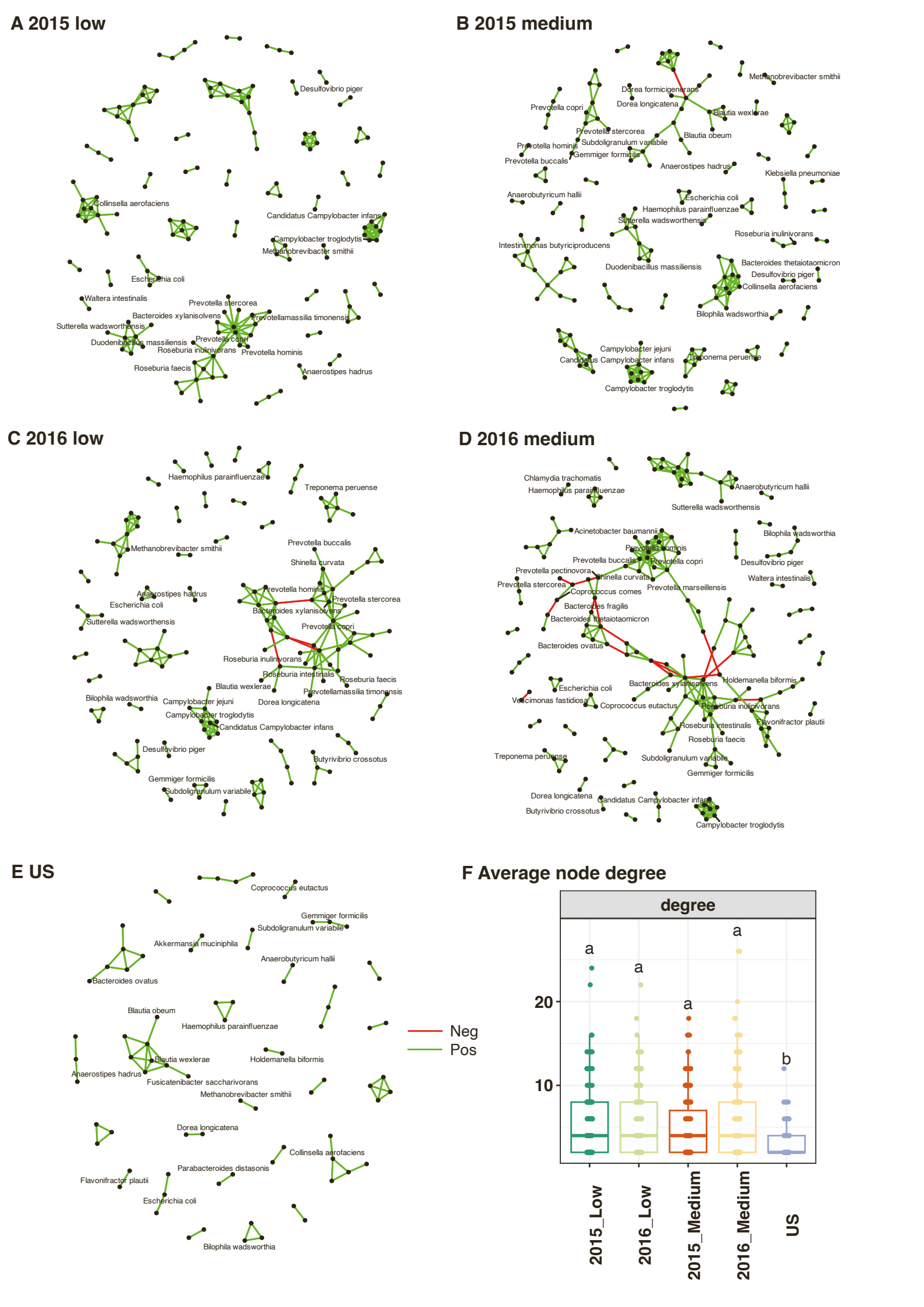

A 2015 low
Desulfovibrio piger
Collinsella aerofaciens
Candidatus Campylobacter infans
Campylobacter troglodytis
Methanobrevibacter smithii
Escherichia coli
Waltera intestinalis
Prevotella stercorea
Bacteroides xylanisolvens
Prevotellamassilia timonensis
Sutterella wadsworthensis
Duodenibacillus massiliensis
Prevotella copri
Roseburia inulinivorans
Prevotella hominis
Roseburia faecis
Anaerostipes hadrus
B 2015 medium
Methanobrevibacter smithii
Dorea formicigenerans
Dorea longicatena
Prevotella copri
Blautia wexlerae
Prevotella stercorea
Blautia obeum
Subdoligranulum variabile
Prevotella hominis
Gemmiger formicilis
Prevotella buccalis
Anaerostipes hadrus
Klebsiella pneumoniae
Anaerobutyricum hallii
Escherichia coli
Haemophilus parainfluenzae
Sutterella wadsworthensis
Roseburia inulinivorans
Intestinimonas butyriciproducens
Duodenibacillus massiliensis
Bacteroides thetaiotaomicron
Desulfovibrio piger
Collinsella aerofaciens
Bilophila wadsworthia
Campylobacter jejuni
Candidatus Campylobacter infans
Treponema peruense
Campylobacter troglodytis
C 2016 low
Haemophilus parainfluenzae
Treponema peruense
Methanobrevibacter smithii
Prevotella buccalis
Shinella curvata
Prevotella hominis
Prevotella stercorea
Anaerostipes hadrus
Escherichia coli
Bacteroides xylanisolvens
Sutterella wadsworthensis
Prevotella copri
Roseburia inulinivorans
Roseburia intestinalis
Roseburia faecis
Blautia wexlerae
Prevotellamassilia timonensis
Bilophila wadsworthia
Campylobacter jejuni
Dorea longicatena
Campylobacter troglodytis
Candidatus Campylobacter infans
Desulfovibrio piger
Butyrivibrio crossotus
Gemmiger formicilis
Subdoligranulum variabile
D 2016 medium
Chlamydia trachomatis
Haemophilus parainfluenzae
Anaerobutyricum hallii
Sutterella wadsworthensis
Acinetobacter baumannii
Bilophila wadsworthia
Prevotella hominis
Prevotella buccalis
Prevotella copri
Desulfovibrio piger
Prevotella pectinovora
Shinella curvata
Waltera intestinalis
Prevotella stercorea
Coprococcus comes
Prevotella marseillensis
Bacteroides fragilis
Bacteroides thetaiotaomicron
Bacteroides ovatus
Holdemanella biformis
Escherichia coli
Bacteroides xylanisolvens
Veillonella fastidiosa
Coprococcus eutactus
Roseburia inulinivorans
Flavonifractor plautii
Roseburia intestinalis
Roseburia faecis
Treponema peruense
Subdoligranulum variabile
Gemmiger formicilis
Dorea longicatena
Candidatus Campylobacter infans
Butyrivibrio crossotus
Campylobacter troglodytis
E US
Coprococcus eutactus
Gemmiger formicilis
Subdoligranulum variabile
Akkermansia muciniphila
Anaerobutyricum hallii
Bacteroides ovatus
Blautia obeum
Haemophilus parainfluenzae
Blautia wexlerae
Holdemanella biformis
Anaerostipes hadrus
Fusicatenibacter saccharivorans
Methanobrevibacter smithii
Dorea longicatena
Collinsella aerofaciens
Parabacteroides distasonis
Flavonifractor plautii
Escherichia coli
Bilophila wadsworthia
Neg
Pos
F Average node degree
degree
20
10
a
a
a
a
b
2015_Low
2016_Low
2015_Medium
2016_Medium
US

Figure S1. Co-occurrence network structure of the fecal microbiota; related to STAR Methods and Figure 2

Co-occurrence networks constructed from shotgun metagenomic data, stratified by urban exposure level and sampling year. Nodes represent bacterial taxa and edges represent significant co-occurrence relationships. Differences in network connectivity and node degree illustrate variation in community structure across exposure groups.

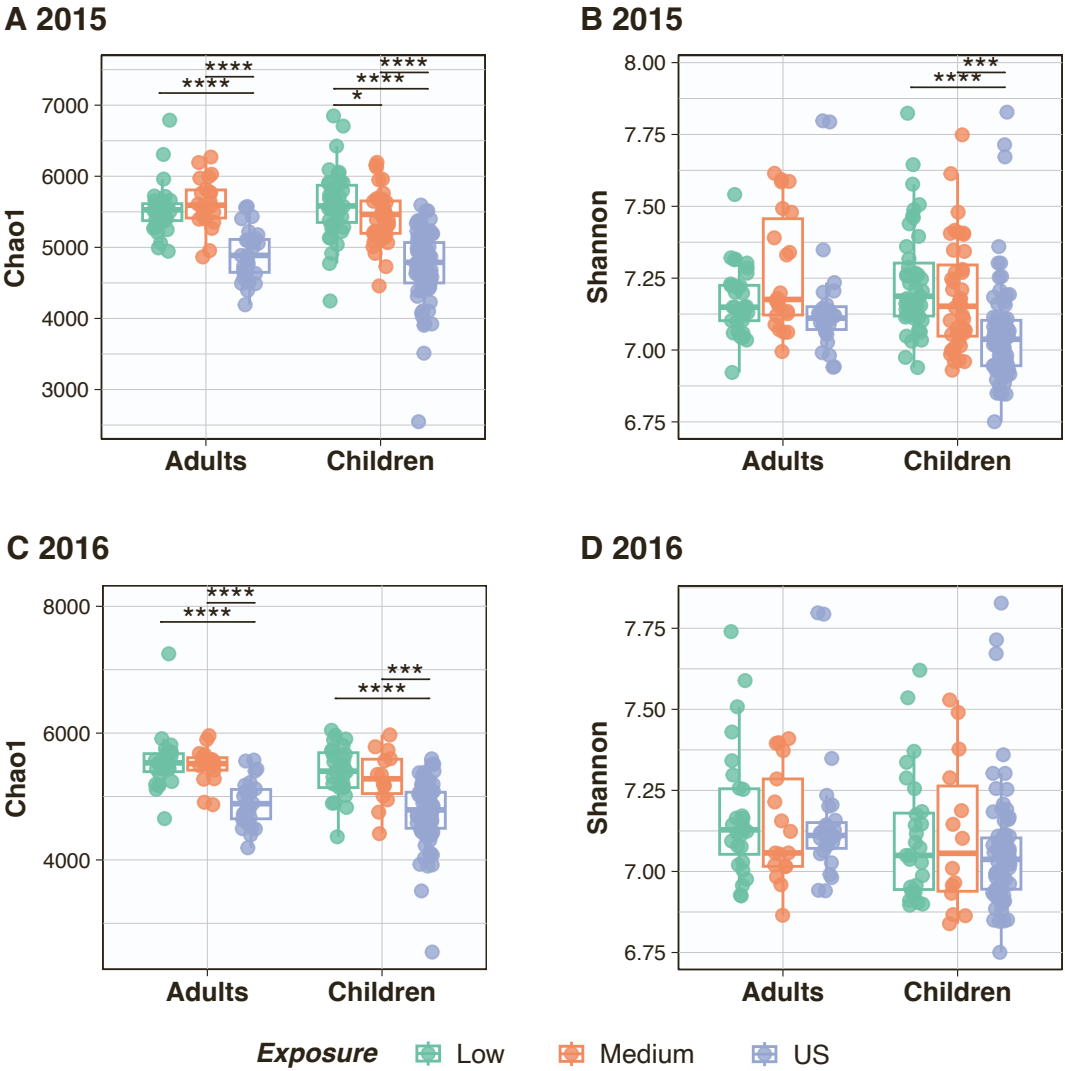
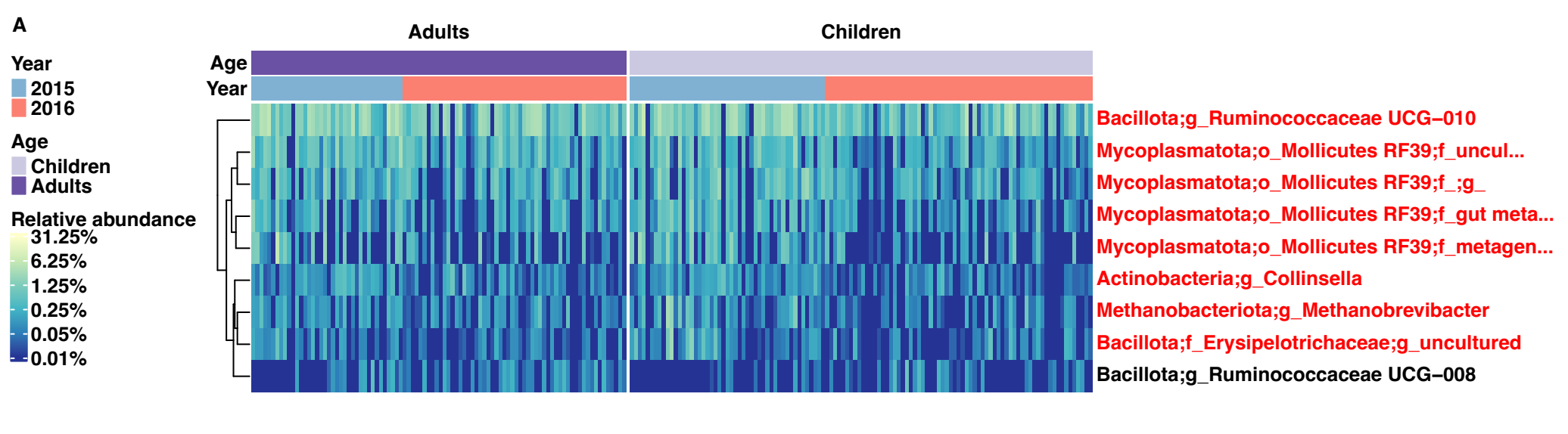

A 2015
B 2015
C 2016
D 2016
Chao1
Shannon
Adults
Children
****
***
*
Exposure
Low
Medium
US

Figure S2. Fecal α-diversity across urban exposure levels; related to STAR Methods and Figure 2

Chao1 richness estimates and Shannon diversity derived from shotgun metagenomic data across urban exposure levels. Asterisks indicate levels of statistical significance following the standard convention used throughout the figures: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$.


A
Year
2015
2016
Age
Children
Adults
Relative abundance
31.25%
6.25%
1.25%
0.25%
0.05%
0.01%
Adults
Children
Age
Year
Bacillota;g_Ruminococcaceae UCG−010
Mycoplasmatota;o_Mollicutes RF39;f_uncul...
Mycoplasmatota;o_Mollicutes RF39;f_;g_
Mycoplasmatota;o_Mollicutes RF39;f_gut meta...
Mycoplasmatota;o_Mollicutes RF39;f_metagen...
Actinobacteria;g_Collinsella
Methanobacteriota;g_Methanobrevibacter
Bacillota;f_Erysipelotrichaceae;g_uncultured
Bacillota;g_Ruminococcaceae UCG−008

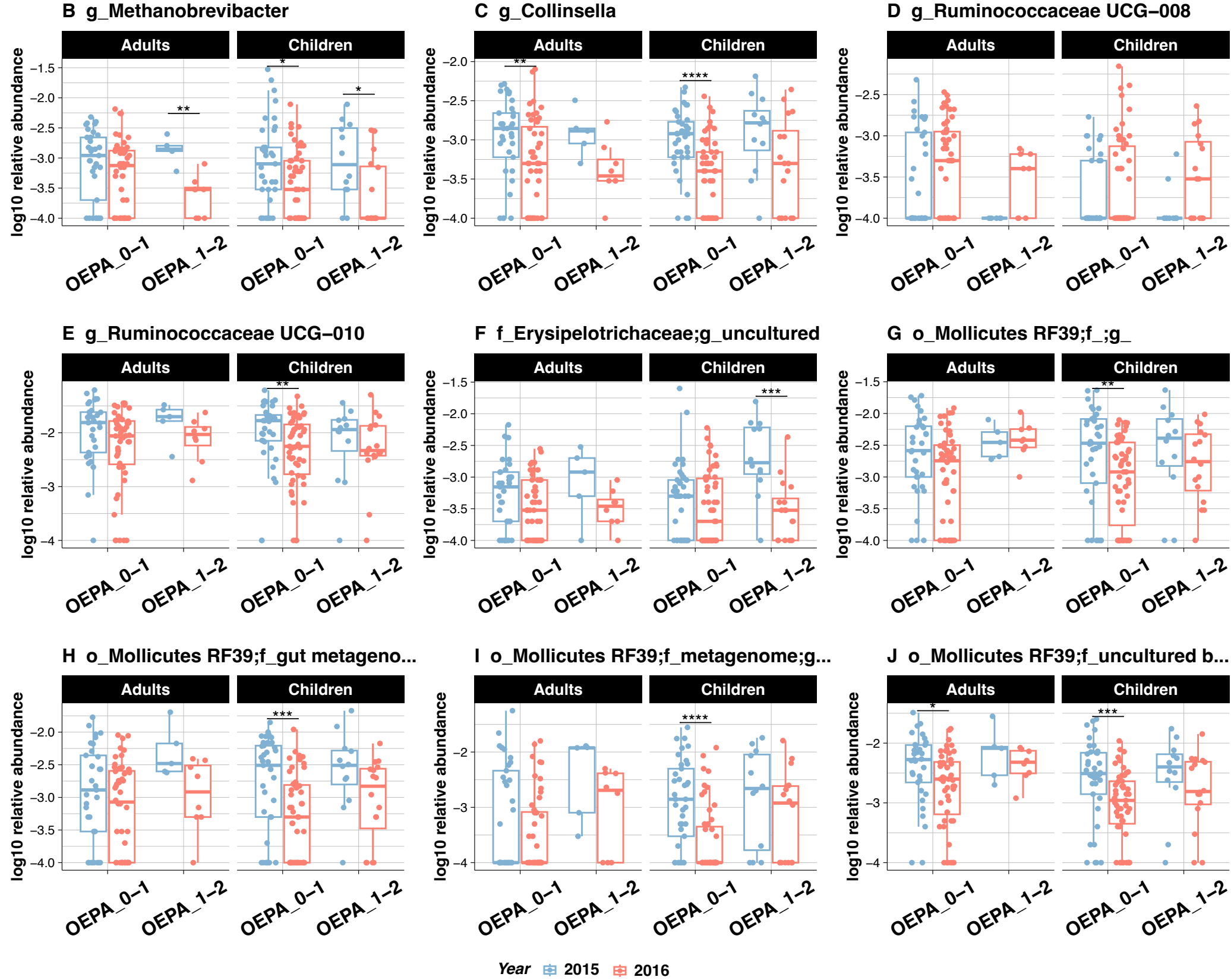
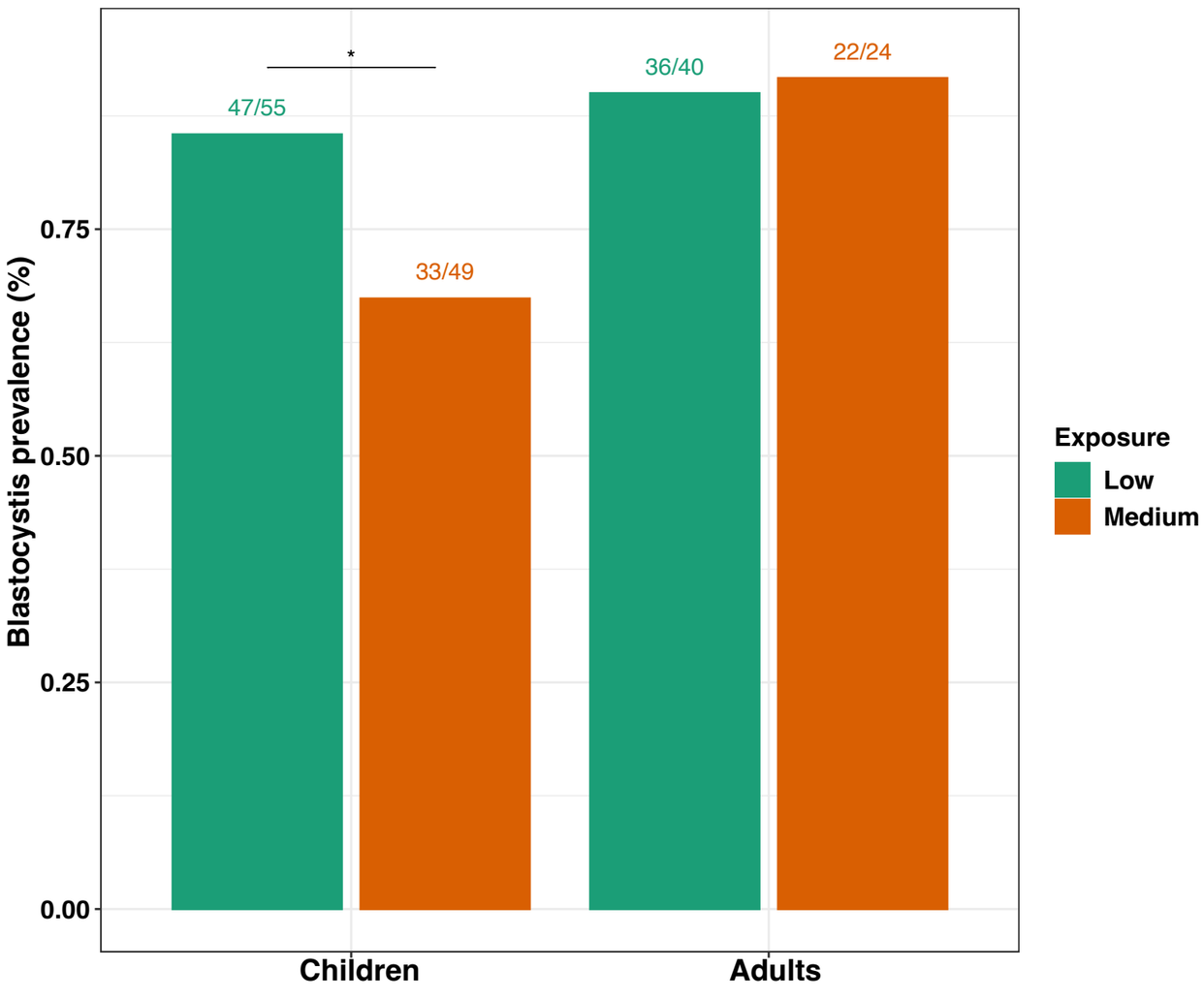

B g_Methanobrevibacter
C g_Collinsella
D g_Ruminococcaceae UCG−008
E g_Ruminococcaceae UCG−010
F f_Erysipelotrichaceae;g_uncultured
G o_Mollicutes RF39;f_;g_
H o_Mollicutes RF39;f_gut metageno...
I o_Mollicutes RF39;f_metagenome;g...
J o_Mollicutes RF39;f_uncultured b...
Adults
Children
log10 relative abundance
OEPA_0−1
OEPA_1−2
Year 2015 2016

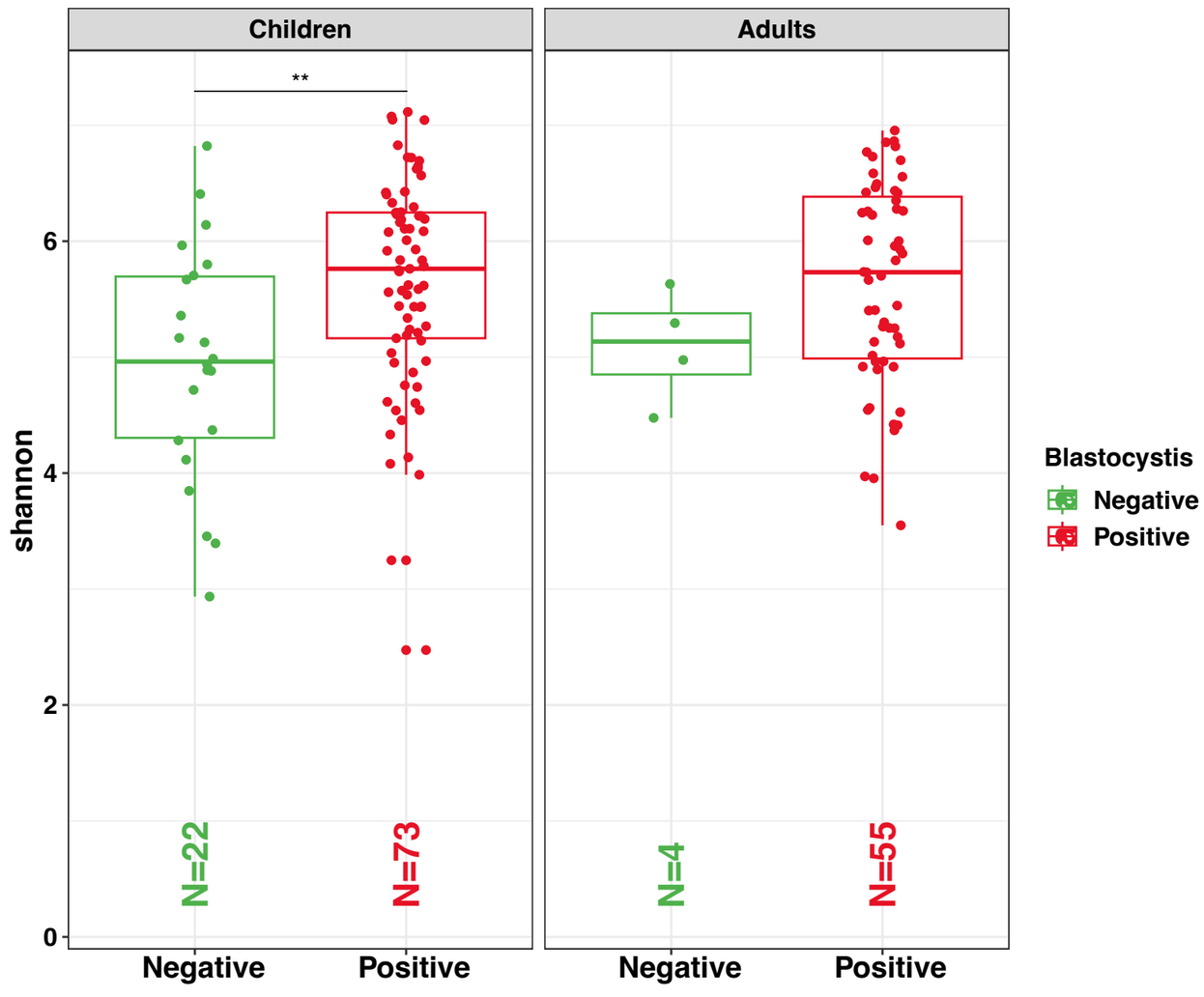
Figure S3. Differentially abundant fecal bacterial taxa between sampling years; related to STAR Methods and Figure 3

Differentially abundant fecal bacterial taxa identified by *16S rRNA* gene sequencing between the 2015 and 2016 sampling visits in low–urban-exposure villages. Taxa shown meet the statistical significance threshold (FDR-corrected $p < 0.05$). (B-J) Boxplots showing $\log_{10}$-transformed relative abundances of representative genera that changed significantly between years. Comparisons were performed using Wilcoxon rank-sum tests; FDR-adjusted *p* values are indicated:*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

A


*
47/55
33/49
36/40
22/24
Blastocystis prevalence (%)
0.75
0.50
0.25
0.00
Children
Adults
Exposure
Low
Medium

B


Children
Adults
**
shannon
6
4
2
0
N=22
N=73
N=4
N=55
Negative
Positive
Negative
Positive
Blastocystis
Negative
Positive

C

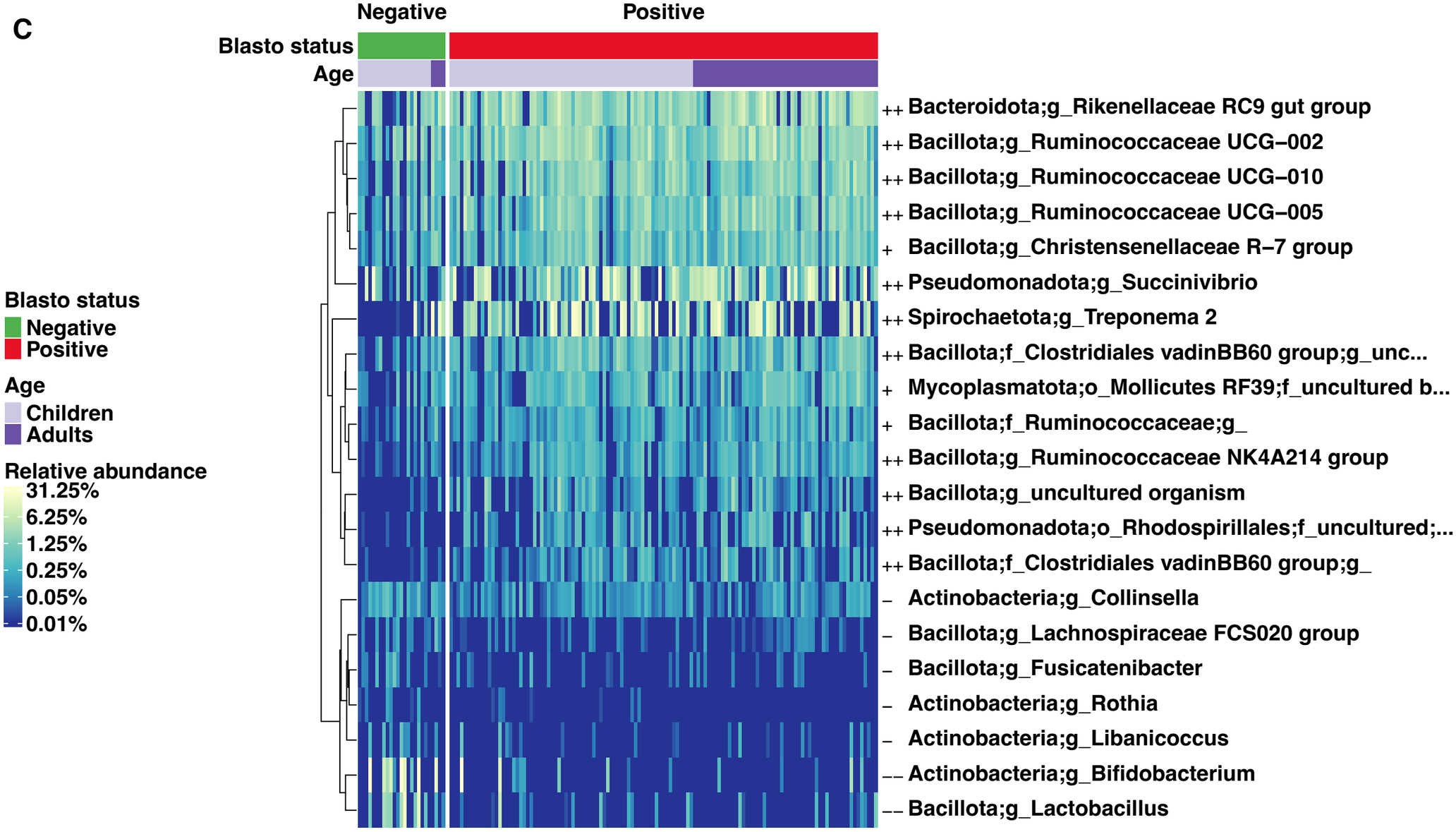

Negative
Positive
Blasto status
Age
++ Bacteroidota;g_Rikenellaceae RC9 gut group
++ Bacillota;g_Ruminococcaceae UCG–002
++ Bacillota;g_Ruminococcaceae UCG–010
++ Bacillota;g_Ruminococcaceae UCG–005
+ Bacillota;g_Christensenellaceae R–7 group
++ Pseudomonadota;g_Succinivibrio
++ Spirochaetota;g_Treponema 2
++ Bacillota;f_Clostridiales vadinBB60 group;g_unc...
+ Mycoplasmatota;o_Mollicutes RF39;f_uncultured b...
+ Bacillota;f_Ruminococcaceae;g_
++ Bacillota;g_Ruminococcaceae NK4A214 group
++ Bacillota;g_uncultured organism
++ Pseudomonadota;o_Rhodospirillales;f_uncultured;...
++ Bacillota;f_Clostridiales vadinBB60 group;g_
– Actinobacteria;g_Collinsella
– Bacillota;g_Lachnospiraceae FCS020 group
– Bacillota;g_Fusicatenibacter
– Actinobacteria;g_Rothia
– Actinobacteria;g_Libanicoccus
–– Actinobacteria;g_Bifidobacterium
–– Bacillota;g_Lactobacillus
Blasto status
Negative
Positive
Age
Children
Adults
Relative abundance
31.25%
6.25%
1.25%
0.25%
0.05%
0.01%

Figure S4. Association of fecal microbiota composition with Blastocystis colonization

(A) Prevalence of *Blastocystis* detected by PCR in Amerindian villagers with low or medium urban exposure. Children were defined as ≤17 years old. The numbers of positive and total subjects are shown above each bar. *Blastocystis* prevalence was lower among children from medium–urban-exposure villages compared with those from low-exposure villages, whereas positivity was nearly universal among adults.
(B) Alpha diversity of the fecal microbiota according to *Blastocystis* status and age group. Diversity was significantly higher in *Blastocystis*-positive children, with a similar trend in adults.
(C) Differentially abundant bacterial taxa associated with *Blastocystis* status based on *16S* rRNA sequencing. Discordant taxa were identified using the ANCOM log-ratio test and are clustered by relative abundance. "+" and "–" indicate enrichment and depletion in the *Blastocystis*-positive group, respectively. Double signs "++" or "– –" indicate ANCOM $|\log_2 \text{ratio}| > 2$.

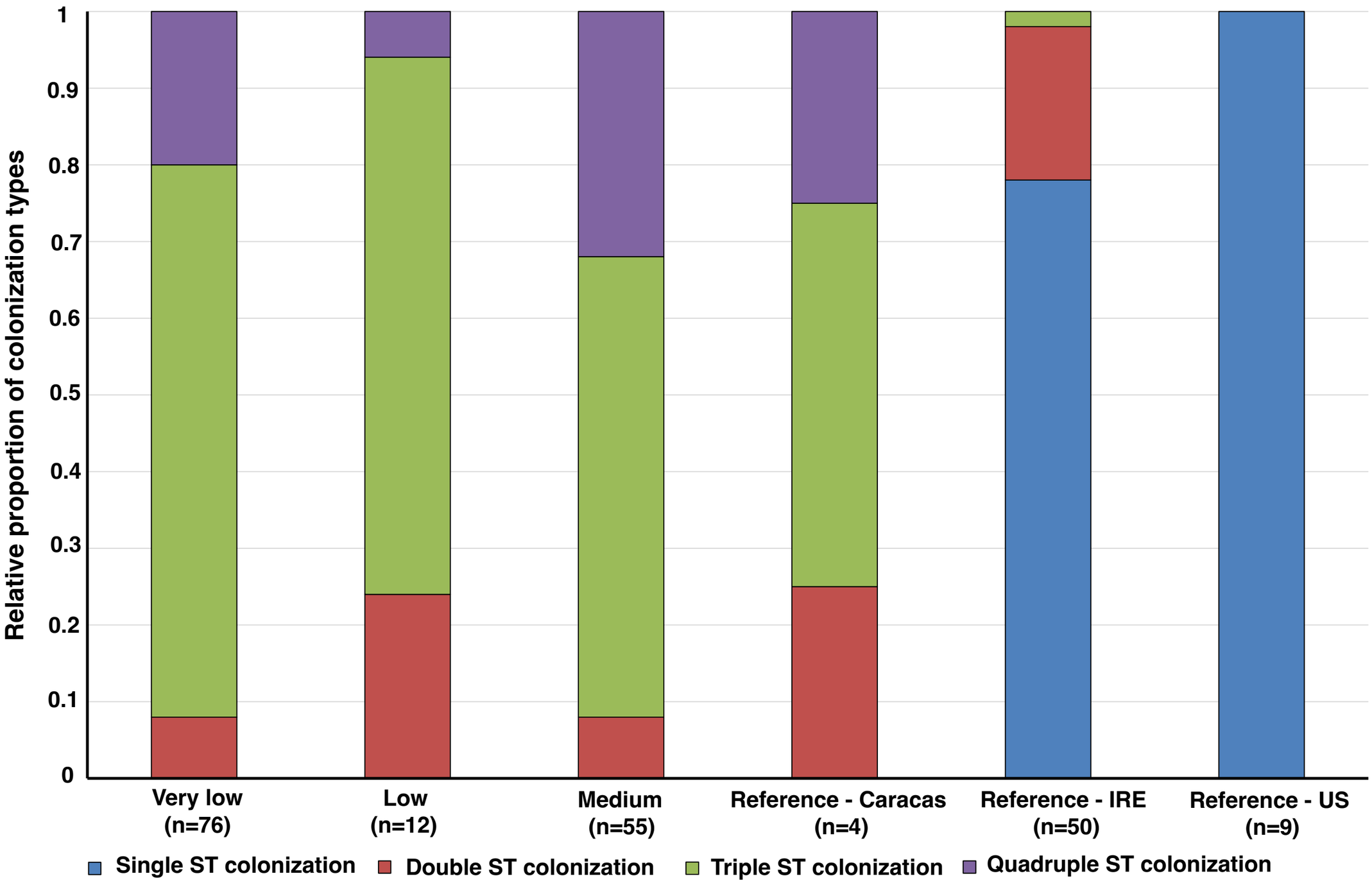

Relative proportion of colonization types
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Very low (n=76)
Low (n=12)
Medium (n=55)
Reference - Caracas (n=4)
Reference - IRE (n=50)
Reference - US (n=9)
Single ST colonization
Double ST colonization
Triple ST colonization
Quadruple ST colonization

Figure S5. Distribution of Blastocystis subtype colonization patterns; related to STAR Methods

Bars show the relative frequency of single- versus multi-subtype *Blastocystis* colonization within each population group.

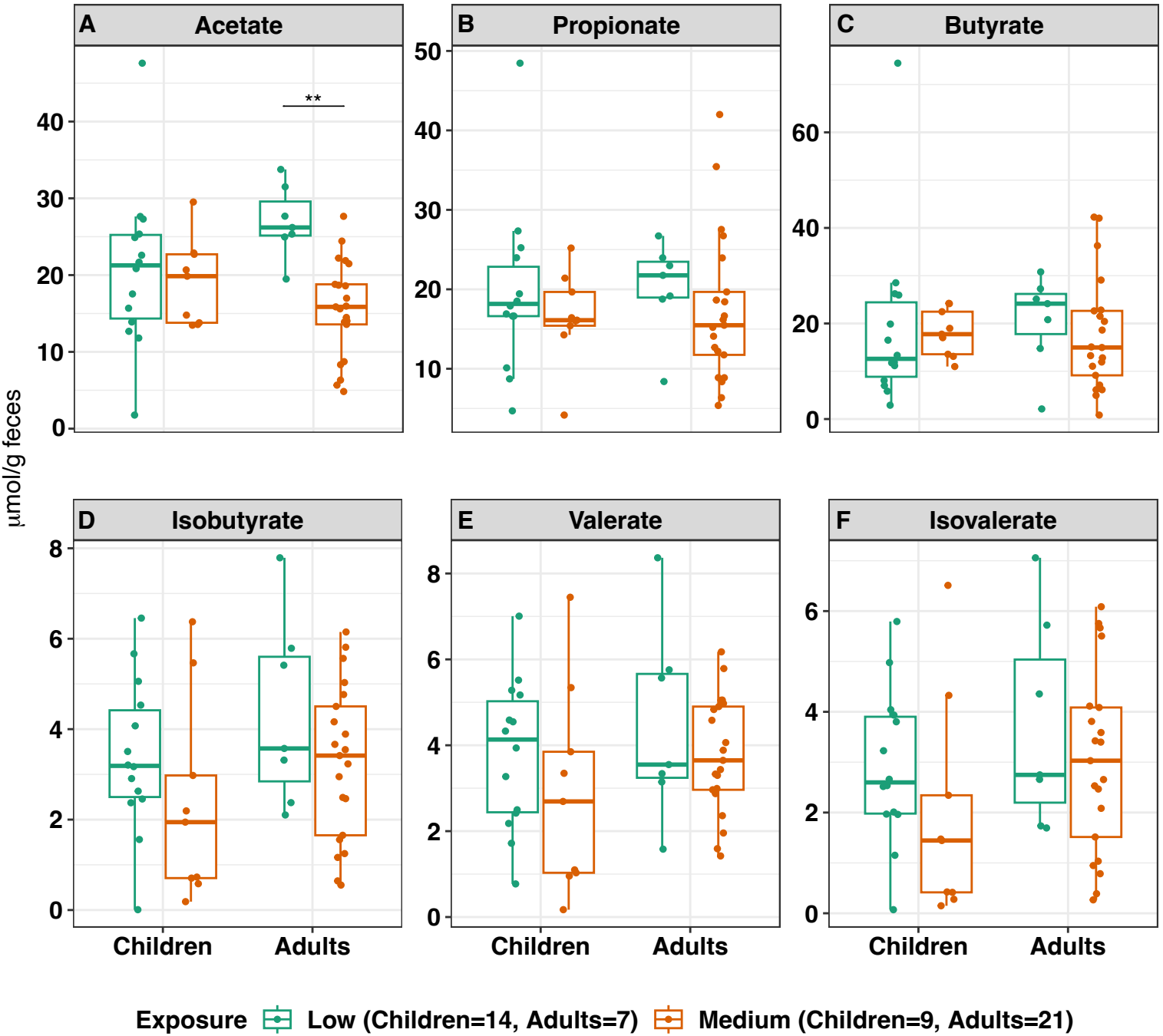

A
Acetate
B
Propionate
C
Butyrate
D
Isobutyrate
E
Valerate
F
Isovalerate
**
µmol/g feces
Children
Adults
Exposure
Low (Children=14, Adults=7)
Medium (Children=9, Adults=21)

Figure S6. Fecal short-chain fatty acid concentrations by urban exposure level; related to STAR Methods

Short-chain fatty acid concentrations measured in fecal samples collected in 2016, stratified by urban exposure level. Statistically significant differences are indicated where applicable.

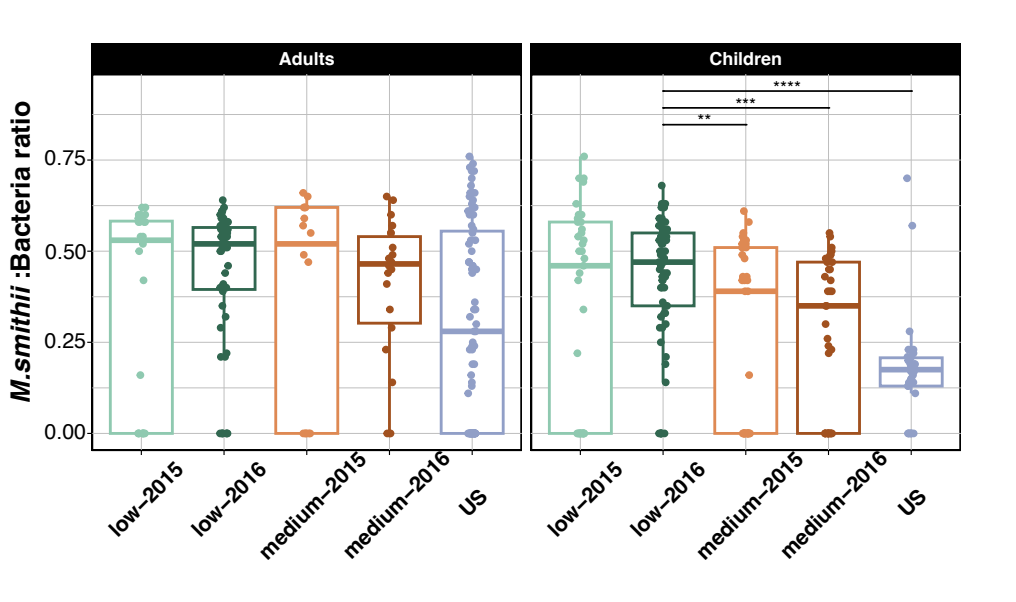

Adults
Children
M.smithii :Bacteria ratio
0.75
0.50
0.25
0.00
**
***
****
low−2015
low−2016
medium−2015
medium−2016
US

Figure S7. *Methanobrevibacter smithii* abundance across age groups and exposure levels; related to STAR Methods and Figure 2

Relative abundance of *Methanobrevibacter smithii* measured by quantitative PCR in fecal samples from children and adults collected in 2016, stratified by urban exposure level. Statistically significant differences are indicated.

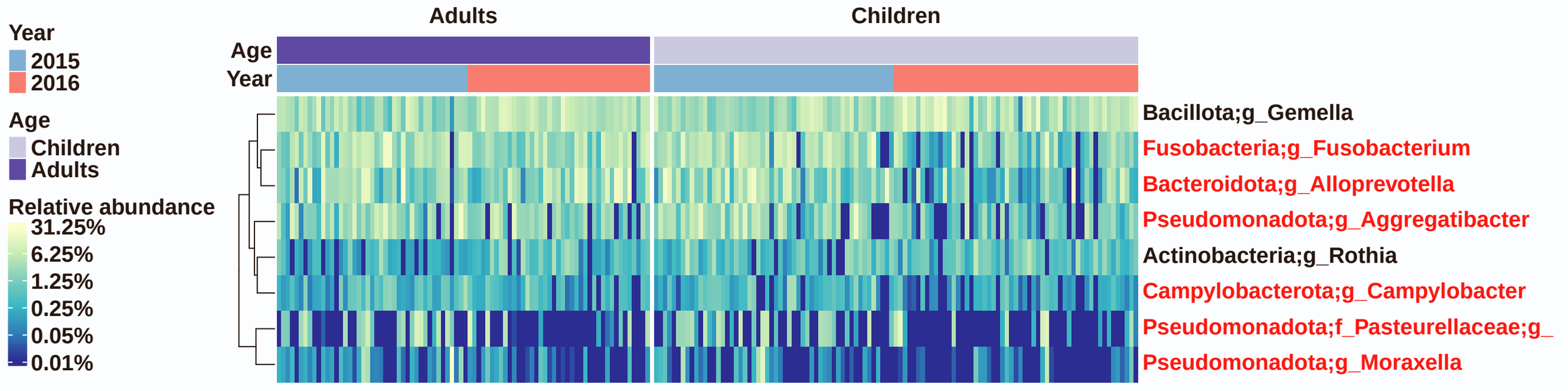

Year
2015
2016
Age
Children
Adults
Relative abundance
31.25%
6.25%
1.25%
0.25%
0.05%
0.01%
Adults
Children
Age
Year
Bacillota;g_Gemella
Fusobacteria;g_Fusobacterium
Bacteroidota;g_Alloprevotella
Pseudomonadota;g_Aggregatibacter
Actinobacteria;g_Rothia
Campylobacterota;g_Campylobacter
Pseudomonadota;f_Pasteurellaceae;g_
Pseudomonadota;g_Moraxella

Figure S8. Differentially abundant oral bacterial taxa between sampling years; related to STAR Methods and Figure 3

Differentially abundant oral bacterial taxa identified by *16S rRNA* gene sequencing between the 2015 and 2016 sampling visits in low–urban-exposure villages. Taxa shown meet the statistical significance threshold (FDR-corrected $p < 0.05$).

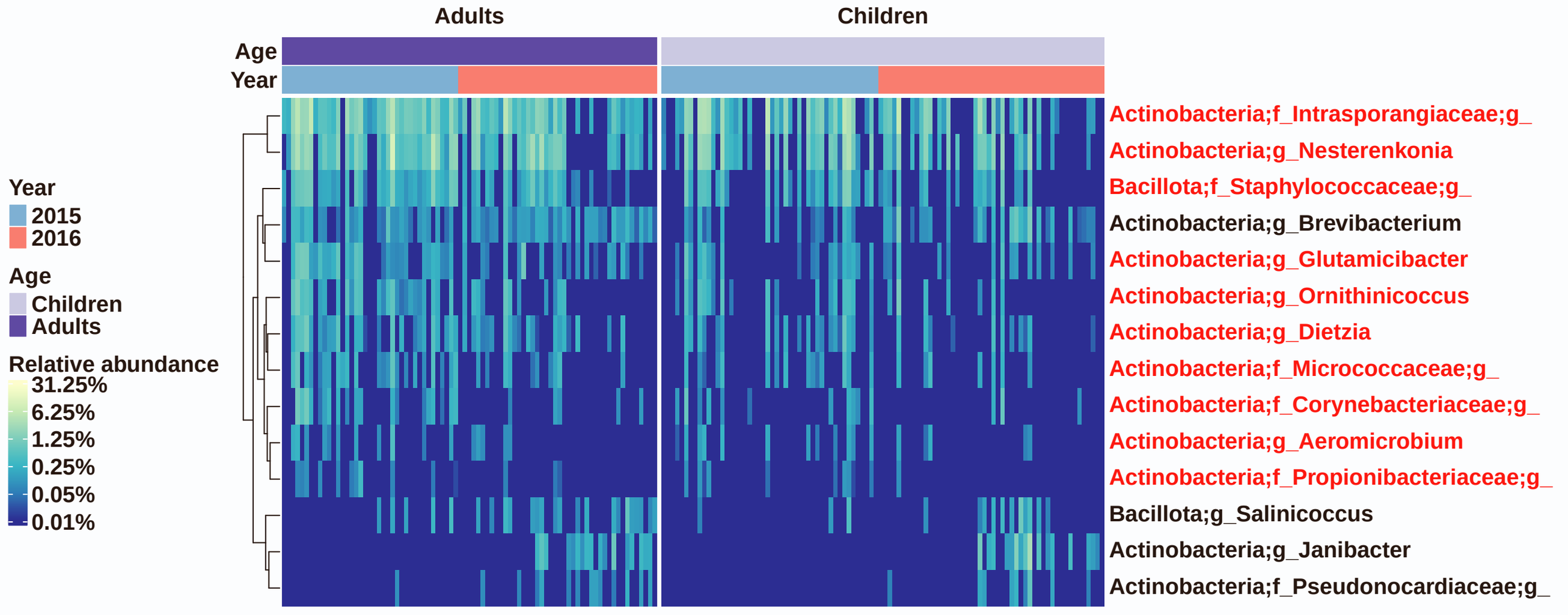

Adults
Children
Age
Year
Year
2015
2016
Age
Children
Adults
Relative abundance
31.25%
6.25%
1.25%
0.25%
0.05%
0.01%
Actinobacteria;f_Intrasporangiaceae;g_
Actinobacteria;g_Nesterenkonia
Bacillota;f_Staphylococcaceae;g_
Actinobacteria;g_Brevibacterium
Actinobacteria;g_Glutamicibacter
Actinobacteria;g_Ornithinicoccus
Actinobacteria;g_Dietzia
Actinobacteria;f_Micrococcaceae;g_
Actinobacteria;f_Corynebacteriaceae;g_
Actinobacteria;g_Aeromicrobium
Actinobacteria;f_Propionibacteriaceae;g_
Bacillota;g_Salinicoccus
Actinobacteria;g_Janibacter
Actinobacteria;f_Pseudonocardiaceae;g_

Figure S9. Differentially abundant nasal bacterial taxa between sampling years; related to STAR Methods and Figure 3

Differentially abundant nasal bacterial taxa identified by *16S rRNA* gene sequencing between the 2015 and 2016 sampling visits in low–urban-exposure villages. Taxa shown meet the statistical significance threshold (FDR-corrected $p < 0.05$).

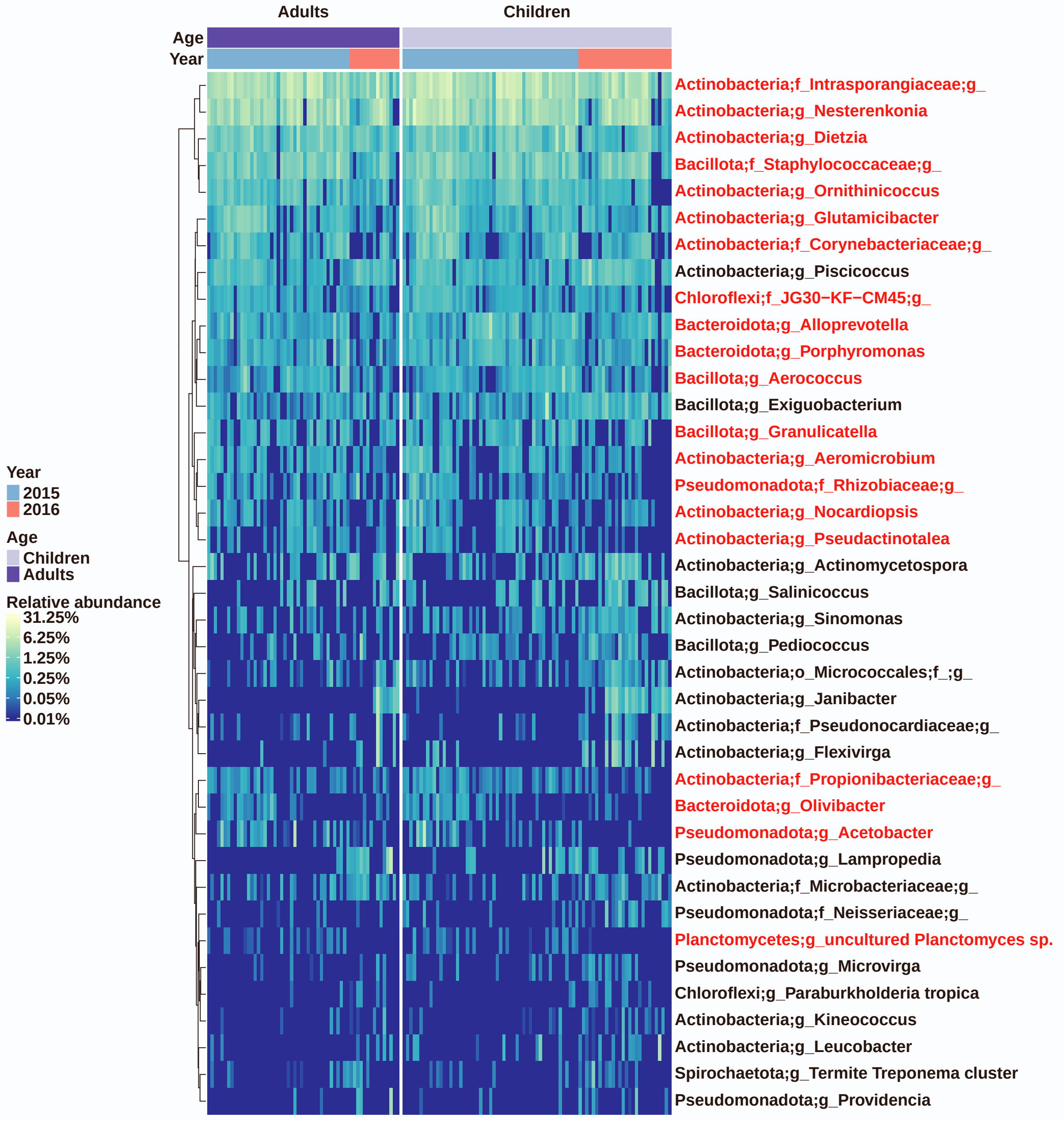

Adults
Children
Age
Year
Actinobacteria;f_Intrasporangiaceae;g_
Actinobacteria;g_Nesterenkonia
Actinobacteria;g_Dietzia
Bacillota;f_Staphylococcaceae;g_
Actinobacteria;g_Ornithinicoccus
Actinobacteria;g_Glutamicibacter
Actinobacteria;f_Corynebacteriaceae;g_
Actinobacteria;g_Piscicoccus
Chloroflexi;f_JG30−KF−CM45;g_
Bacteroidota;g_Alloprevotella
Bacteroidota;g_Porphyromonas
Bacillota;g_Aerococcus
Bacillota;g_Exiguobacterium
Bacillota;g_Granulicatella
Actinobacteria;g_Aeromicrobium
Pseudomonadota;f_Rhizobiaceae;g_
Actinobacteria;g_Nocardiopsis
Actinobacteria;g_Pseudactinotalea
Actinobacteria;g_Actinomycetospora
Bacillota;g_Salinicoccus
Actinobacteria;g_Sinomonas
Bacillota;g_Pediococcus
Actinobacteria;o_Micrococcales;f_;g_
Actinobacteria;g_Janibacter
Actinobacteria;f_Pseudonocardiaceae;g_
Actinobacteria;g_Flexivirga
Actinobacteria;f_Propionibacteriaceae;g_
Bacteroidota;g_Olivibacter
Pseudomonadota;g_Acetobacter
Pseudomonadota;g_Lampropedia
Actinobacteria;f_Microbacteriaceae;g_
Pseudomonadota;f_Neisseriaceae;g_
Planctomycetes;g_uncultured Planctomyces sp.
Pseudomonadota;g_Microvirga
Chloroflexi;g_Paraburkholderia tropica
Actinobacteria;g_Kineococcus
Actinobacteria;g_Leucobacter
Spirochaetota;g_Termite Treponema cluster
Pseudomonadota;g_Providencia
Year
2015
2016
Age
Children
Adults
Relative abundance
31.25%
6.25%
1.25%
0.25%
0.05%
0.01%

Figure S10. Differentially abundant skin bacterial taxa between sampling years; related to STAR Methods and Figure 3

Differentially abundant right arm skin bacterial taxa identified by *16S rRNA* gene sequencing between the 2015 and 2016 sampling visits in low–urban-exposure villages. Taxa shown meet the statistical significance threshold (FDR-corrected $p < 0.05$).