METHODS

Chemical lysis was undertaken by adding 200 μL of guanidinium thiocyanate-EDTA-sarkosyl and 900 μL of PBS to the samples before they underwent bead-beating. Cell disruption was undertaken using Qiagen stainless steel beads and tungsten carbide beads on a Fastprep-24 Instrument (MP Biomedicals Europe, Illkirch, France) running at 6.5 m/s for 45 seconds. Two cycles of thermolysis followed by alternating incubation at 90°C and −20°C for 10 minutes each, before cell debris was pelleted by centrifugation at 13,000*g* for 10 minutes. Supernatant was transferred to a fresh microfuge tube, where it was inverted with 140 μL of 5 molar sodium chloride and 374 μL 40% polyethylene glycol and precipitated for 1 hour at 4°C. DNA was pelleted by centrifugation at 13,000*g* for 10 minutes and resuspended in 500 μL of sterile distilled water. Three hundred microliter of phenol/chloroform (1:1) was added and samples were vortexed before centrifugation at 13,000*g* for 5 minutes. The upper phase was then transferred to a fresh microfuge tube. Total DNA was then precipitated by the addition of an equal volume of isopropanol and 0.1 volume of 10 molar ammonium acetate and stored at −20°C for 1 hour. DNA was pelleted by centrifugation at 13,000*g* for 10 minutes. Pelleted DNA was washed in 70% ethanol, dried, and resuspended in 30 μL of sterile distilled water.[E1,E2]

DNA extracts were amplified using universal bacterial primers targeting the 16S ribosomal RNA gene at hypervariable region V4 (515F-806R), which are tailed with sequences to incorporate Illumina (San Diego, Calif) adapters and indexing barcodes.[E3] Sequencing was performed on the MiSeq instrument using version 2 chemistry and 250 cycles, stratifying the amplicon samples according to 3-, 6-, and 12-month time points between each plate.

REFERENCES

E1. Green GL, Brostoff J, Hudspith B, Michael M, Mylonaki M, Rayment N, et al. Molecular characterization of the bacteria adherent to human colorectal mucosa. J Appl Microbiol 2006;100:460-9.

E2. Rogers GB, Cuthbertson L, Hoffman LR, Wing PA, Pope C, Hooftman DA, et al. Reducing bias in bacterial community analysis of lower respiratory infections. ISME J 2013;7:697-706.

E3. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci U S A 2011;108:4516-22.

E4. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. Genome Biol 2011;12:R60.

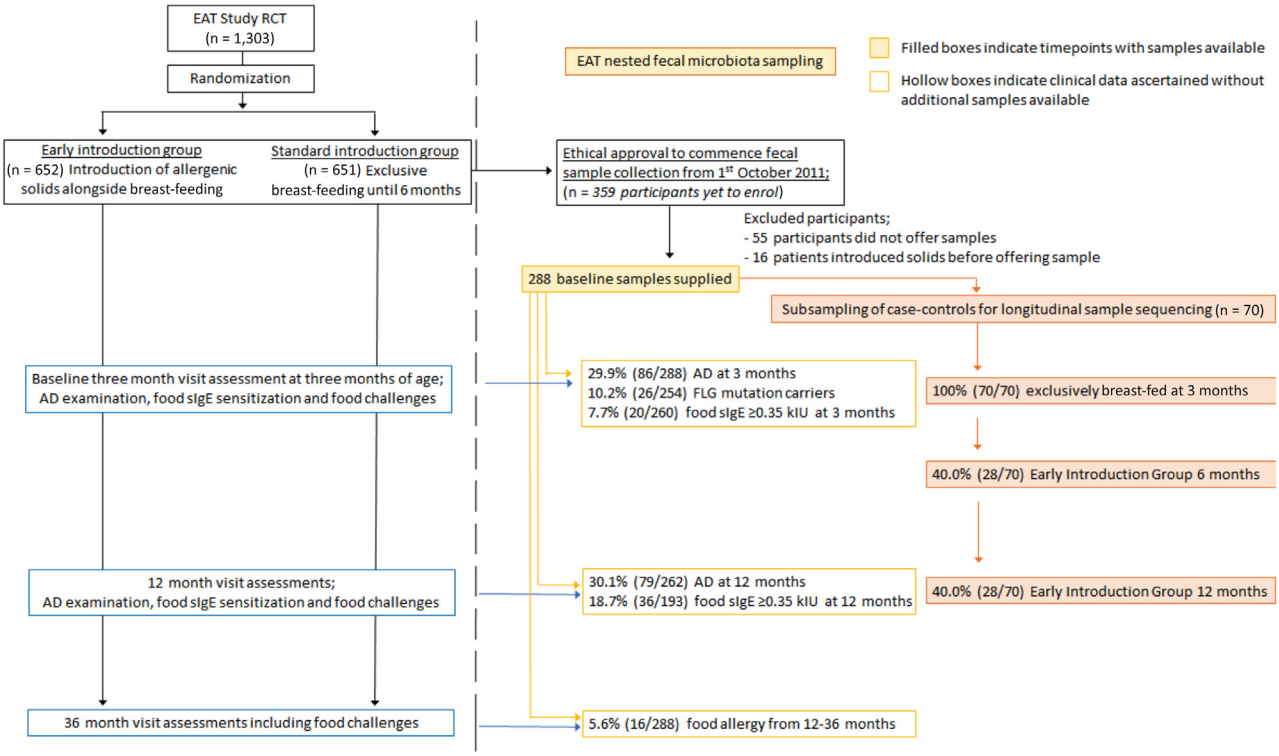

FIG E1. Participant flow diagram. Of the potential participants (n = 359) in the nested study, 288 participants provided a baseline (3-month) fecal sample. For the subset of case-control longitudinal sampling (n = 70), participants provided additional fecal samples at 6 and 12 months.

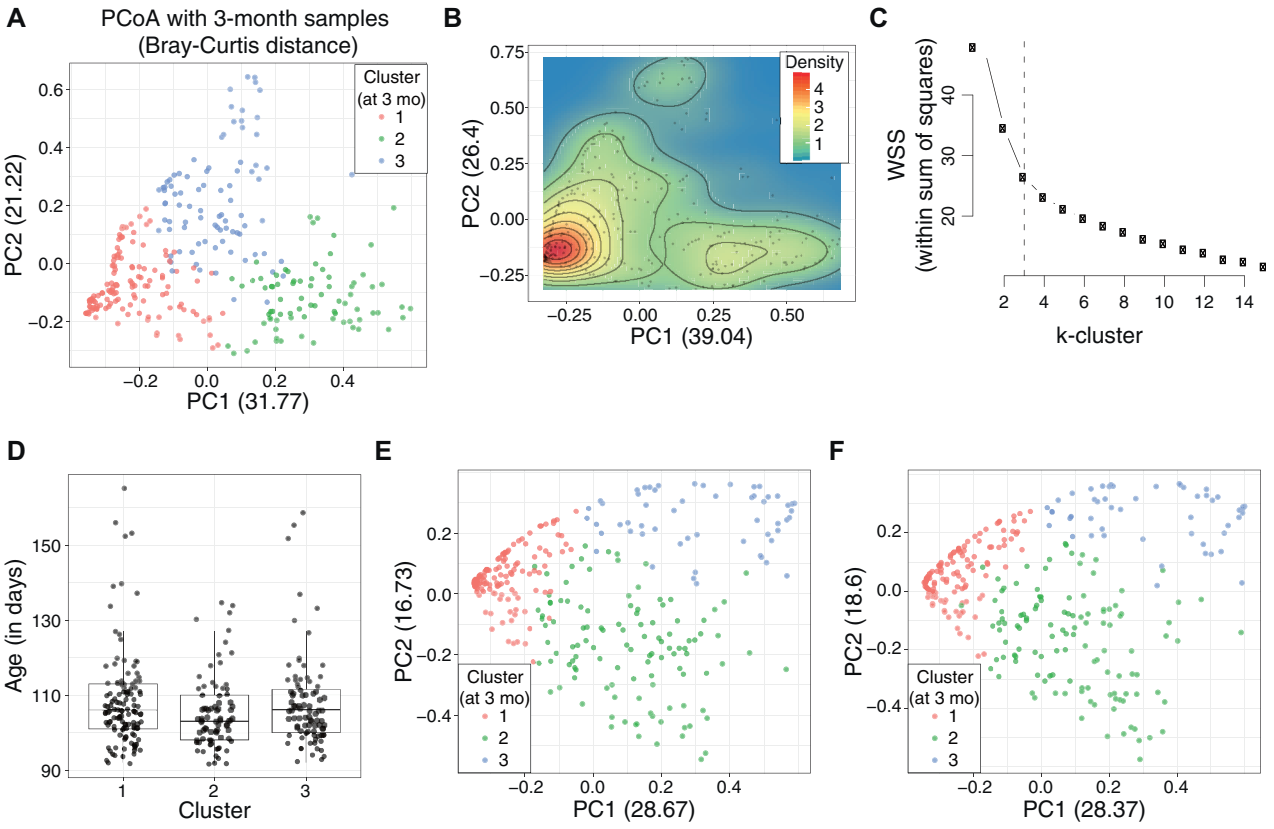

FIG E2. Patterns of gut microbiota of baseline fecal samples. **A,** PCoA of the gut microbiome at 3 months. Pair-wise distances (Bray-Curtis distance) among all samples were calculated and 2 major axes (PC1 and PC2) from the multidimensional distance space were calculated and depicted on a scatter plot. Colors indicate different clusters, according to *k*-means clustering. **B,** Density distribution of 3-month-old infants' microbiome based on PCoA (same as Fig 1, *A*). There are 3 high-density peaks pronounced, suggesting 3 clusters of microbial community populations. **C,** Evaluation of optimal clustering using "within sum of squares" by each *k* (number of clusters). Based on the graphical visualization of within sum of squares or the "elbow method," $k = 3$ (3 clusters) is the optimal number of clusters. Increasing the number of clusters beyond 3 ($k > 3$) results in overfitting, whereas *k* less than 3 is not as effective in minimizing intracluster variation. **D,** Distribution of age (in days) of the infants when their 3-month samples were acquired, grouped by each cluster (P = NS; Kruskal-Wallis test). **E,** PCoA of microbiome communities at 3 months with species-level classification. The colors indicate clusters. **F,** PCoA of microbiome communities at 3 months with 97% *de novo* OTU-level classification. The colors indicate clusters. *NS*, Nonsignificant; *OTU*, operational taxonomic unit.

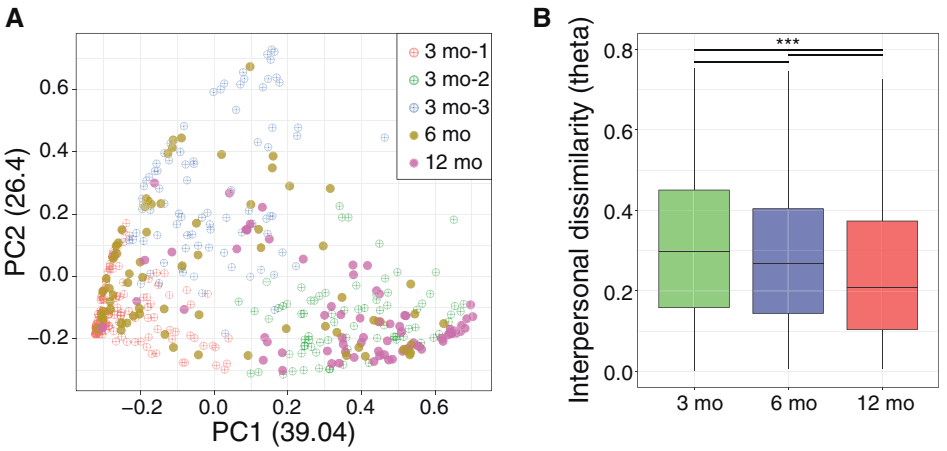

FIG E3. Longitudinal changes in the gut microbiome in the EAT study. **A,** Gut microbiota for 6- and 12-month samples was overlaid on the PCoA ordination of the microbiota of 3-month samples (same as Fig 1, *A*). The colors indicate clustering and age groups. At age 12 months, the microbiota largely converges to cluster 2 of the 3-month gut microbiota. **B,** Boxplot showing changes in interpersonal dissimilarity at different time points. All-to-all theta distances within each age group were calculated and plotted. Lower value (close to 0) indicates communities are more similar.

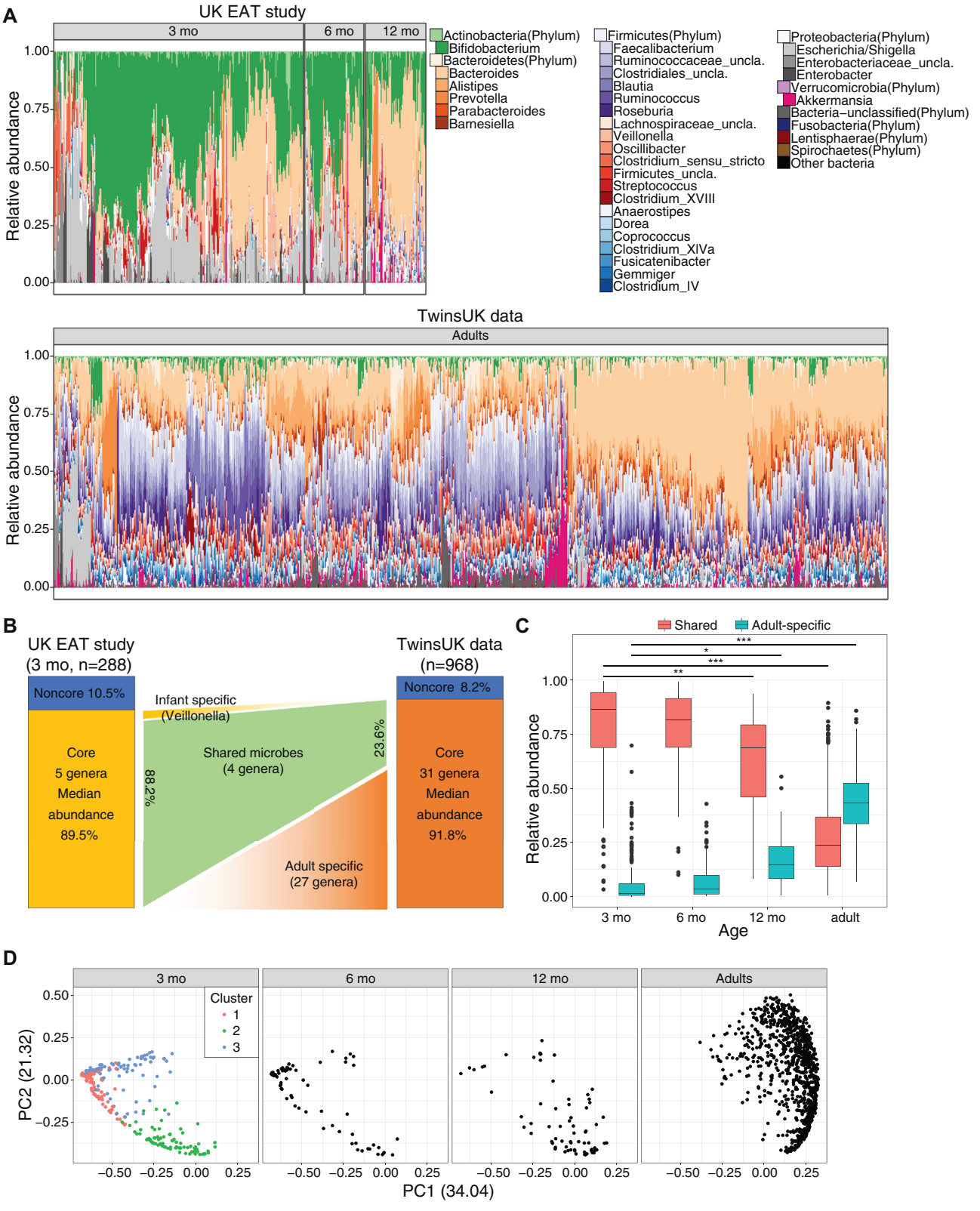

FIG E4. Microbiota differences between infants and adults. **A,** Stacked bar charts for relative abundances of major genera in all samples, from both EAT (3, 6, and 12 months) and TwinsUK (adults) cohorts. Each bar indicates an individual. **B,** Schematic diagram describing differences in core/noncore genera between 3-month-old infants and adults. There are 5 core genera (genera exist in more than 95% of individuals, *Bifidobacterium*, *Bacteroides*, *Streptococcus*, *Escherichia/Shigella*, and *Veillonella*) from 3-month-old infants. Among them, 4 genera (except *Veillonella*) are also core in adults. **C,** Boxplot showing relative changes in shared and adult-specific core microbes (**P* < .05, ***P* < .01, ****P* < .001; Wilcoxon rank-sum test after Kruskal-Wallis test). At 12 months, relative abundances of adult-specific core genera are increased and shared cores are decreased, compared with 3 months. **D,** PCoA plot showing relative changes in gut microbiome during maturation. Twelve-month samples clustered more closely with adults.

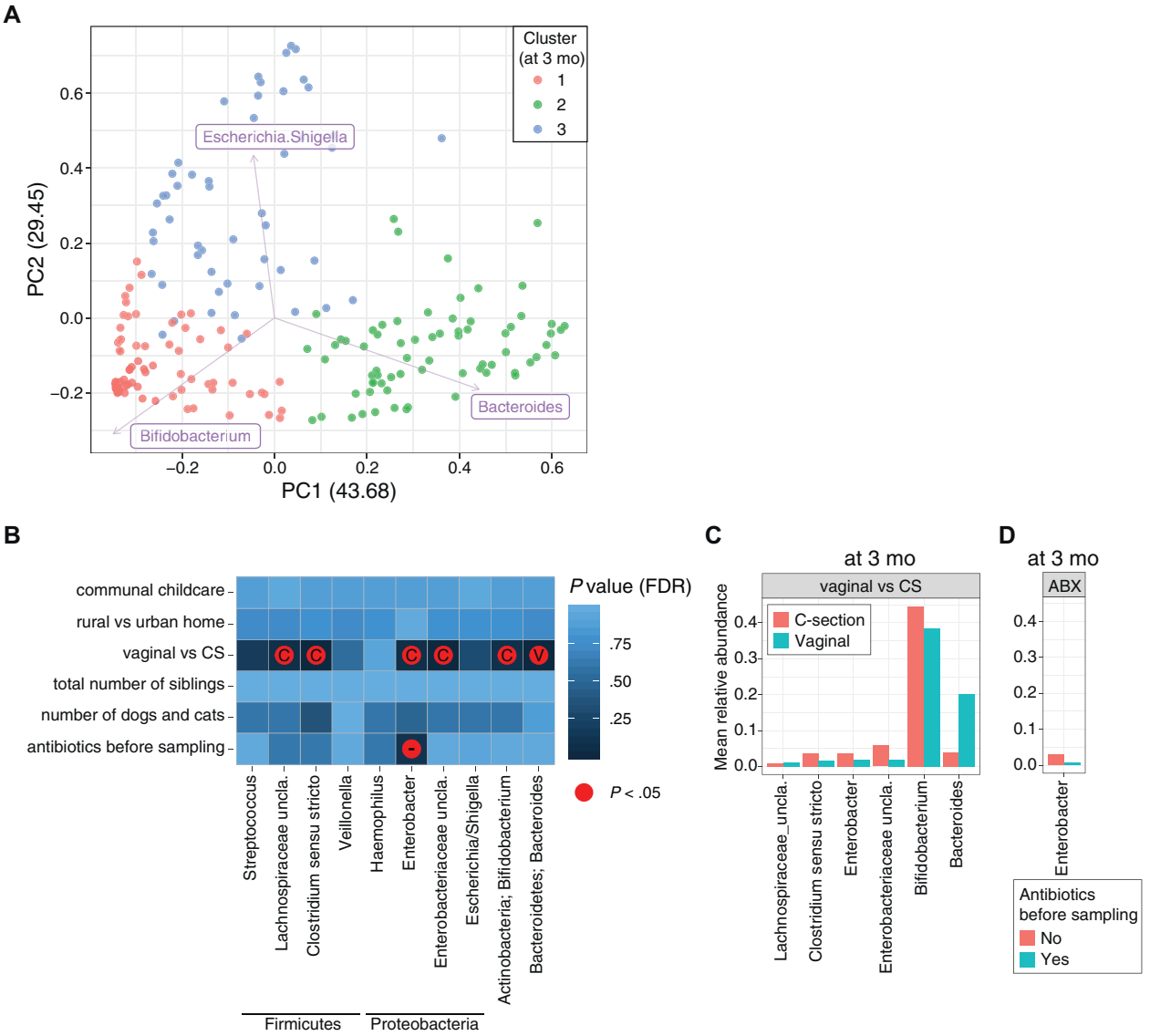

FIG E5. A, PCoA of baseline gut microbiota of vaginally born infants. Different colors indicate clusters, and arrows indicate specific genera significantly correlated with PCoA ordination ($P < .05$, lengths of arrows are proportional to R^2 (calculated by EnvFit in R)). **B,** Heatmap showing associations between environmental exposures and microbial composition. Darker colors indicate lower P values, and red marks indicate associations with FDR-corrected P values lower than .05. **C,** Mean relative abundance differences of selected taxa from B (vaginal vs CS). **D,** Mean relative abundance differences of selected taxa from B (antibiotics before sampling). *CS*, Cesarean section; *FDR*, false-discovery rate.

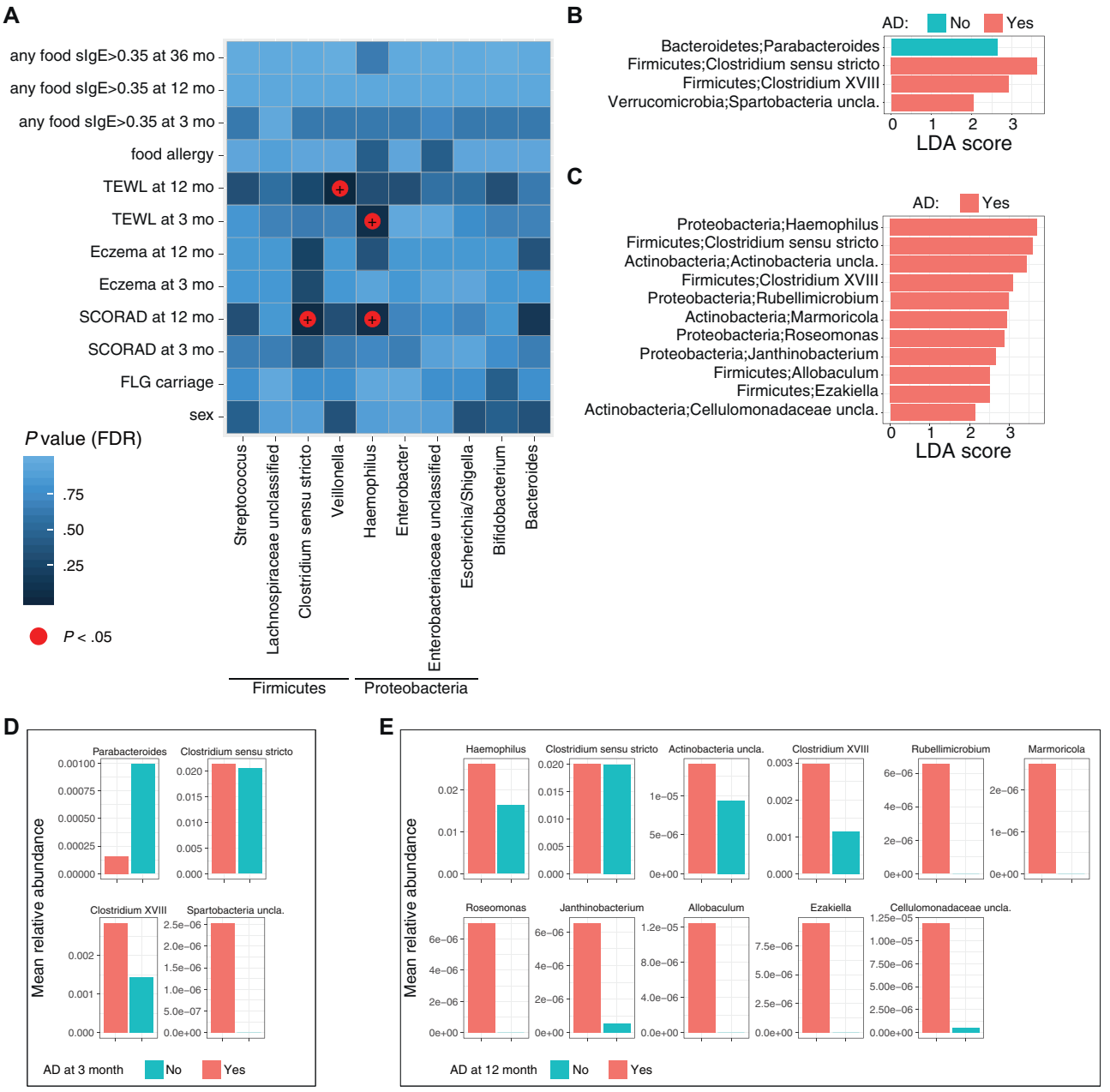

FIG E6. Association of clinical measurements with microbial compositions. **A,** Heatmap showing associations between various clinical measurements and microbial composition. Darker colors indicate lower *P* values, and red marks indicate associations with FDR-corrected *P* values lower than .05. **B,** Discriminatory taxa from 3-month gut microbiota in infants with or without eczema at age 3 months (LDA score). The LDA score is calculated by LEfSe.[E4] **C,** Discriminatory taxa from 3-month gut microbiota in infants with or without eczema at age 12 months. The LDA score is calculated by LEfSe. **D,** Mean relative abundance differences of selected taxa from *B* (with or without eczema at 3 months). **E,** Mean relative abundance differences of selected taxa from *C* (with or without eczema at 12 months). *FDR*, False-discovery rate; *LDA*, linear discriminant analysis; *SCORAD*, SCORing Atopic Dermatitis; *TEWL*, transepidermal water loss.

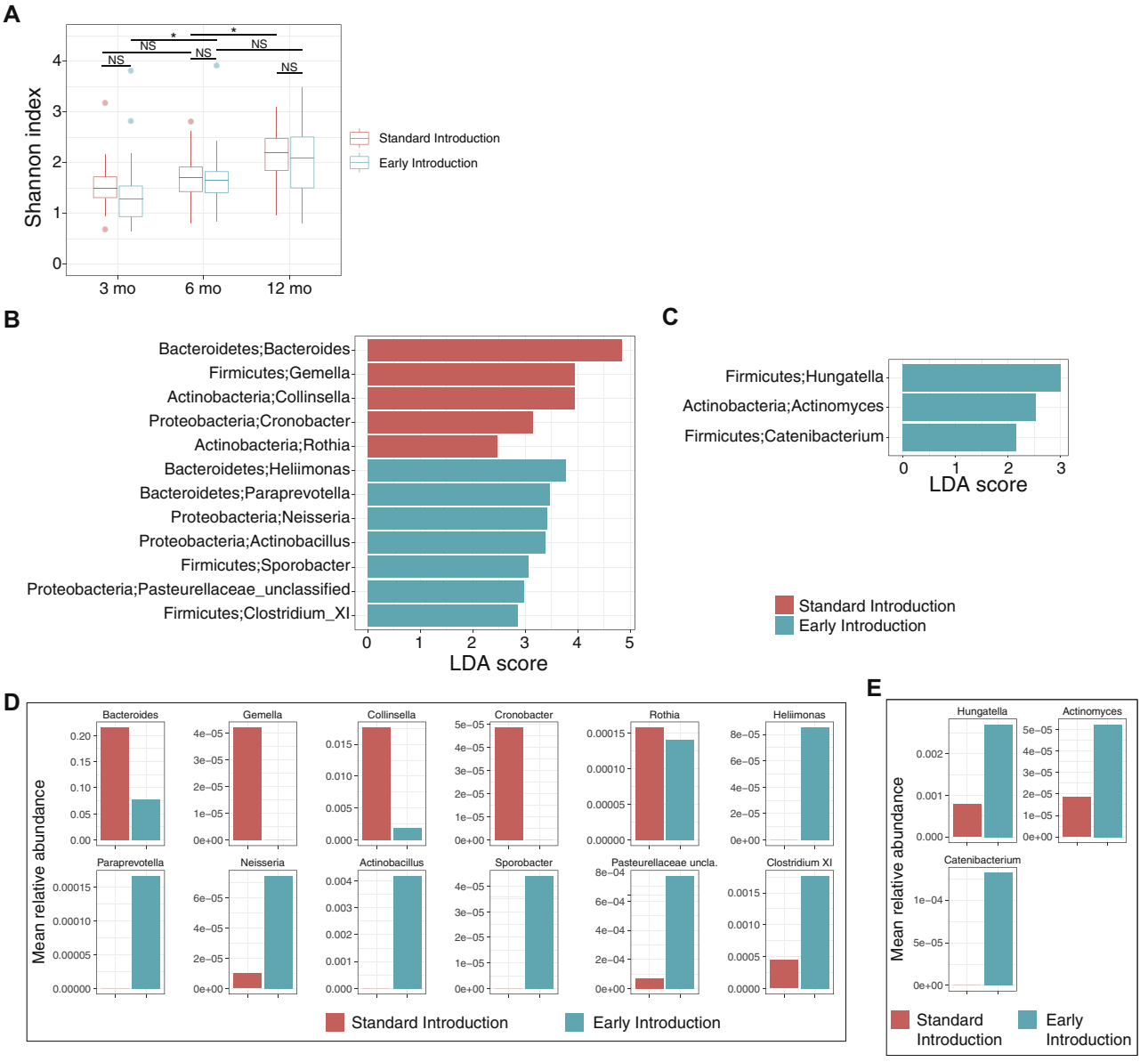

FIG E7. Differential microbiota dynamics by early food introduction in the EAT cohort. **A,** Boxplot comparing Shannon diversity changes among participants' longitudinal samples according to randomized allocation to continued exclusive breast-feeding (standard introduction group) or the introduction of allergenic solids (early introduction group). (*$P < .05$; paired Wilcoxon rank-sum test). **B,** Discriminatory taxa from 6-month gut microbiota of infants in standard vs early introduction (LDA score). The LDA score is calculated by LEfSe. **C,** Discriminatory taxa from 12-month gut microbiota of infants in standard vs early introduction (LDA score). The LDA score is calculated by LEfSe. **D,** Mean relative abundance differences of selected taxa from *B* (standrard vs early introduction). **E,** Mean relative abundance differences of selected taxa from *C* (standard vs early introduction). *LDA,* Linear discriminant analysis; *NS,* nonsignificant.

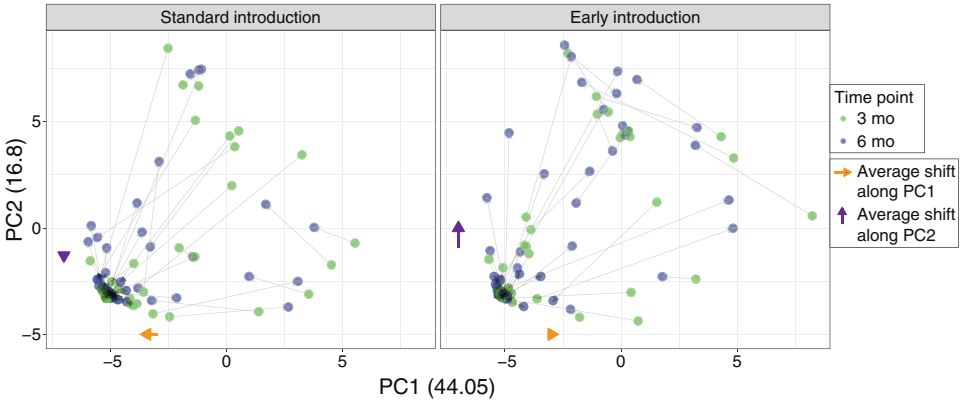

FIG E8. Differential microbiota dynamics by early food introduction in the TEDDY cohort. PCoA scatter plot demonstrating longitudinal transition from age 3 to 6 months of defined samples from the TEDDY cohort. From the TEDDY data set, samples from infants who were (i) exclusively breast-fed before solid food introduction and (ii) longitudinal data (>2 sampling before age 6 months) were selected. Gray lines connect samples from the same individual. Yellow and purple arrows on the sides indicate the average shift of the microbiota in the PCoA axes.