Science

Robust variation in infant gut microbiome assembly across a spectrum of lifestyles

MATTHEW R. OLM, DYLAN DAHAN, MATTHEW M. CARTER, BRYAN D. MERRILL, FEIQIAO B. YU, SUNIT JAIN, XIANDONG MENG, SURYA TRIPATHI, HANNAH WASTYK, NORMA NEFF, SUSAN HOLMES, ERICA D. SONNENBURG, AASHISH R. JHA, AND JUSTIN L. SONNENBURG fewer Authors Info & Affiliations

SCIENCE • 9 Jun 2022 • Vol 376, Issue 6598 • pp. 1220-1223 • DOI: 10.1126/science.abj2972

Led by Anthony Grigsby and Chance Roberts

MIP 505 2026-02-11

novel species. Industrialized infants even those who are breastfed have microbiomes characterized by a paucity of Bifidobacterium infantis and gene cassettes involved in human milk utilization. Strains within lifestyle-associated taxonomic groups are shared between mother-infant dyads, consistent with early life inheritance of lifestyle-shaped microbiomes. The population-specific differences in infant microbiome composition and function underscore the importance of studying microbiomes from people outside of wealthy, industrialized nations.

Olm et al. performed deep metagenomic sequencing on infant stool samples from the Hadza hunter-gatherers of Tanzania and combined them with a global meta-analysis of 1,900 samples across 18 populations, revealing that infant microbiomes develop along lifestyle-associated trajectories that diverge within the first six months of life.

Olm et al., Science 376, 1220-1223 (2022)

Authors

Corresponding Author:
Justin L. Sonnenburg
Stanford University,
Dept. of Microbiology and Immunology

Co-First Authors:
Matthew R. Olm & Dylan Dahan

Funding: NIH (DP1-AT009892, R01-DK085025), NSF GRFP, Bill & Melinda Gates Foundation

Paper Details

Journal:
Science (Impact Factor ~56)

Published:
Matthew R. Olm & Dylan Dahan

Data Availability:
BioProject PRJEB49206

Ethical Note:
MTA w/ Tanzania NIMR; verbal consent from Hadza w/ translator

Research Motivation

The infant gut microbiome undergoes a critical assembly process beginning at birth. Species acquired early may shape the final adult microbiome through niche partitioning and establishment.

Knowledge Gap

Microbiome assembly is well-characterized in industrialized infants, but largely unknown in non-industrialized populations, including hunter-gatherer societies like the Hadza of Tanzania.

Research Questions

1. When do infant microbiomes from different lifestyles diverge?
2. What microbes and functions characterize infants from different lifestyles?
3. Are there differences in taxa vertically transmitted from mothers to infants that seed the assembly process?

Olm et al., Science 376, 1220-1223 (2022)

Industrialization impacts the gut microbiome of infants, which can be seen both temporally in development and in the functions the microbes conduct.

Olm et al., Science 376, 1220-1223 (2022)

Experimental Design

16S rRNA

1,900 samples
across 18 global populations

Includes:
- **62 Hadza infant samples**
- **Industrialized populations**
(US, Sweden, Finland, etc.)
- **Transitional populations**
(Bangladesh, India, Peru)
- **Non-industrialized**
(Hadza, Bolivia, Tsimane)

"Deep" Metagenomics

39 Hadza infants
+ 23 matched maternal samples

Analyses:
- MAG assembly & binning
- Sub-species variation
- Functional potential
(HMO genes)
- Vertical transmission
(inStrain)

Comparative Framework

3 lifestyle categories

Non-industrialized
Hadza, Bolivia (Tismane)

Transitional
Bangladesh, Malawi, Peru, India, Nigeria

Industrialized
US, Sweden, Finland, Germany, Estonia, Russia

Key Methods

Amplicon Sequencing

- 16S rRNA gene sequencing (V4)
- UniFrac dissimilarity
- Co-abundance group (CAG) analysis
- Global meta-analysis of 18 populations

Metagenome Assembly

- Deep shotgun metagenomics on Hadza samples
- MAG binning → 745 species
- Integrated w/ UHGG & adult genomes
- 5,755 species-rep. genome database

Functional & Strain Analysis

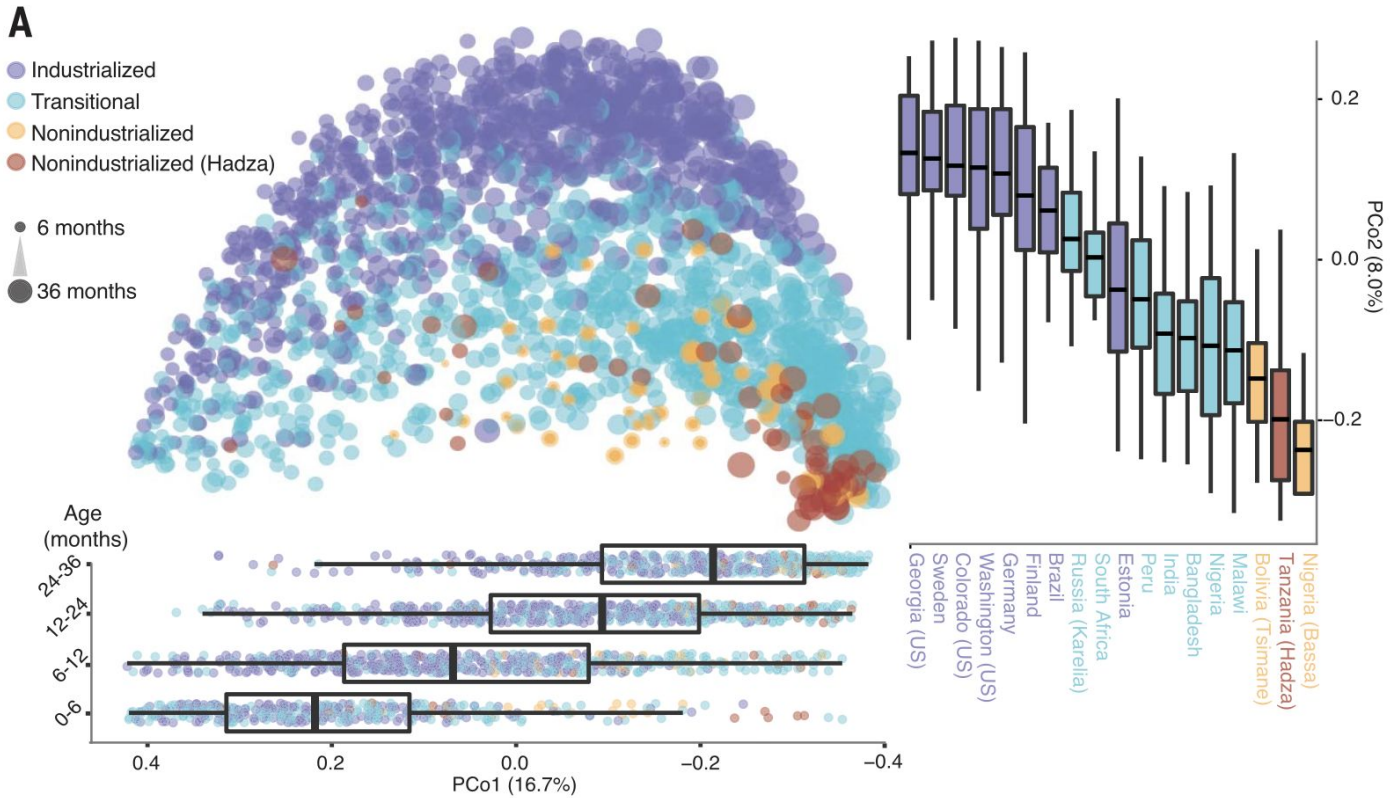
- HMO utilization gene cluster mapping
- *B. infantis* strain-level phylo.
- inStrain strain tracking
- 20 *B. infantis* isolates

Statistical Methods

- EnvFit for ordination axis assoc.
- Wilcoxon tests for comparisons
- Fisher's exact test w/ FDR
- *in silico* rarefaction for depth normalization

Figure 1: Age and lifestyle are associated with infant microbiome composition

Figure 1A/1B Esther Alorkpa



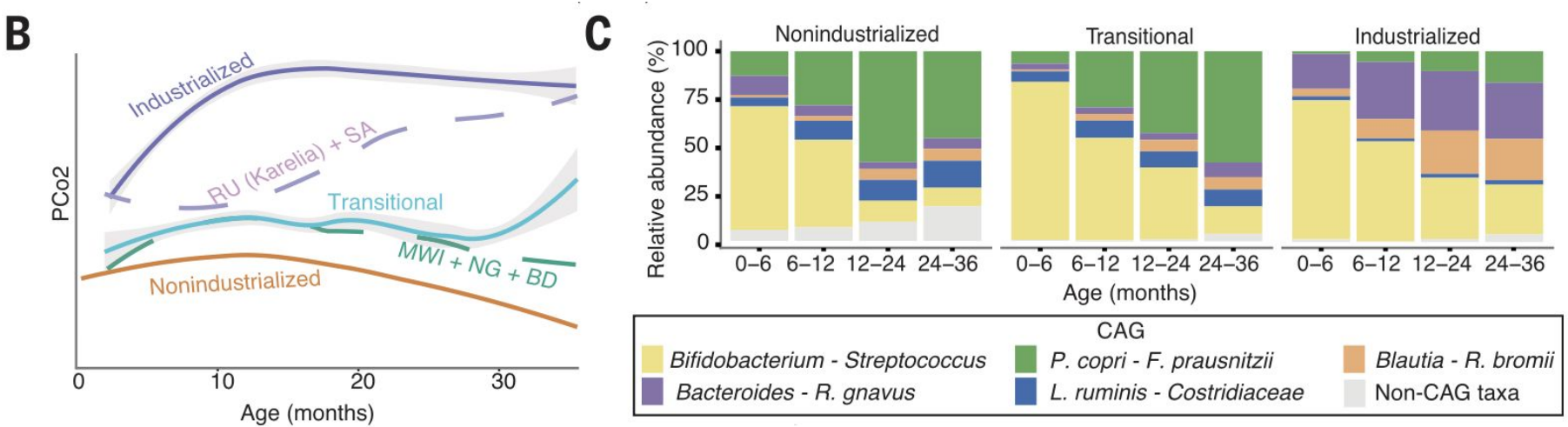
(A) Unweighted UniFrac dissimilarity Principal Coordinates Analysis (PCoA) (top left panel) of 1900 fecal samples from infants (<3 years old) across 18 populations based on amplicon sequence variant abundance.

Point color indicates lifestyle and point size is proportional to age in months. Boxplots show the distribution of indicated age groups along PCo1 (bottom) and cohorts along PCo2 (right).

Figure 1: Age and lifestyle are associated with infant microbiome composition

Figure 1A/1B Esther Alorkpa

Figure 1B/1C Ghazaleh Aminiershad

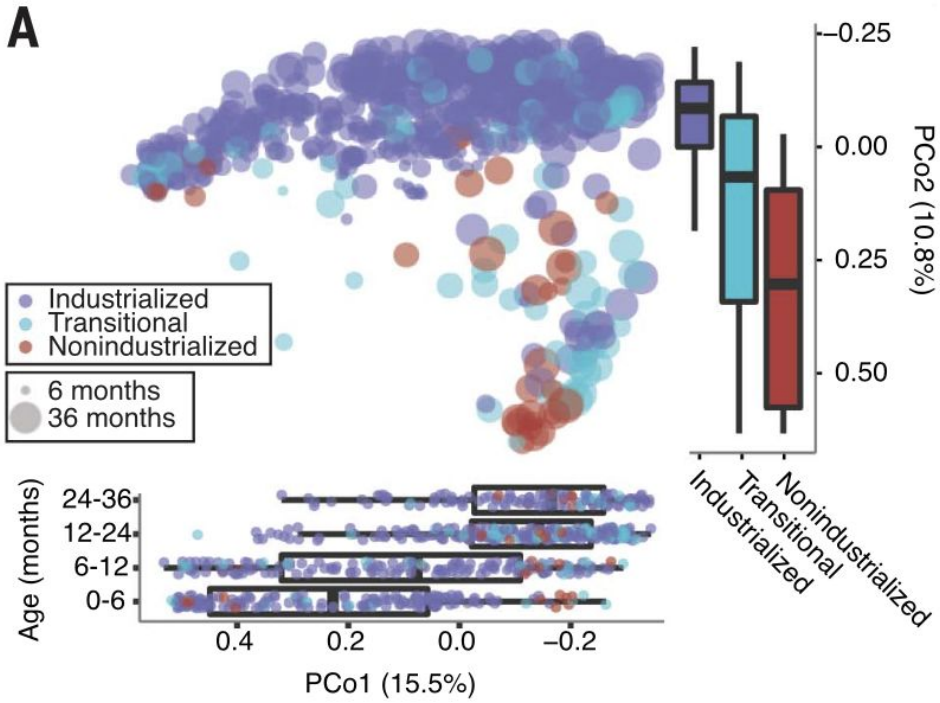


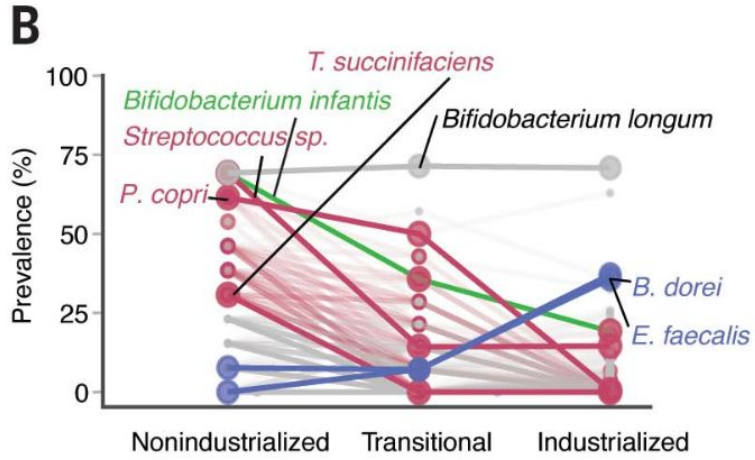
(B) PCo2 versus sample age for the three lifestyle categories (solid lines) and specific indicated subpopulations (dashed lines). The purple dashed line includes Russia (Karelia) and South Africa [RU (Karelia) + SA] and the green dashed line includes Malawi, Nigeria (Urban), and Bangladesh (MWI + NG + BD). The middle transitional line (blue) contains all transitional samples. Lines are the smoothed conditional mean of PCo2 loadings (loess fit).

(C) Relative abundance of CAGs by age group and lifestyle. Taxa in annotation are the most abundant taxa in a CAG.

Figure 2: Age and lifestyle are associated with infant microbiome functions

Figure 2A/2B Dominikus Atmaka





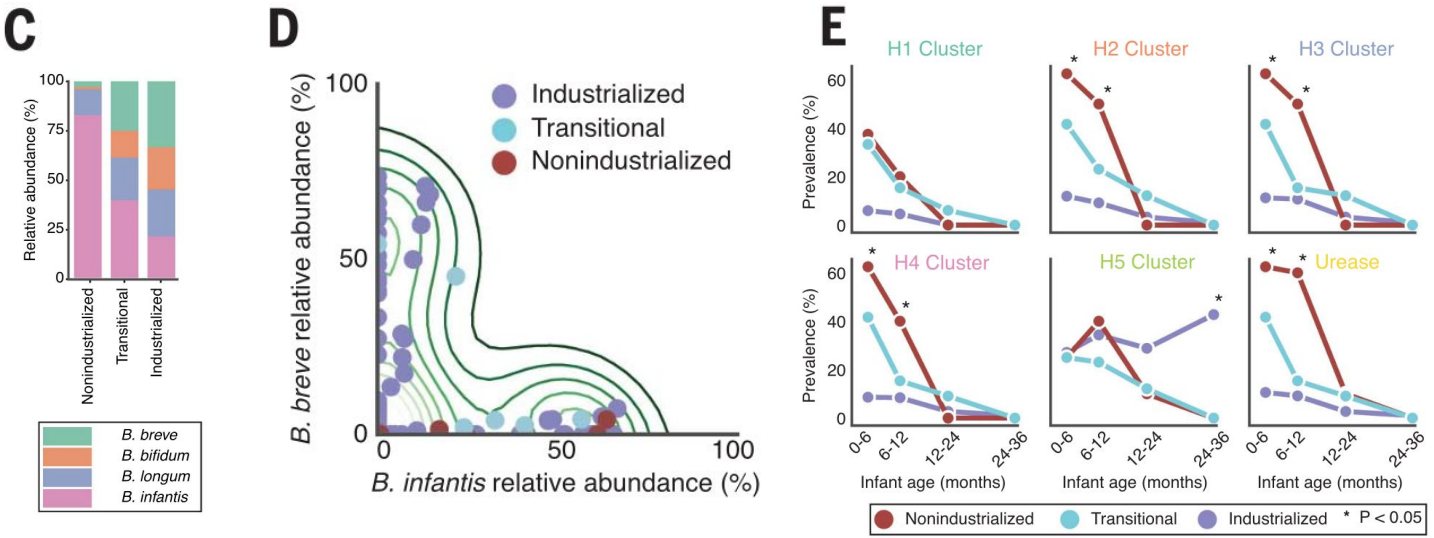
(A) PCoA on the basis of on 682 infant fecal metagenomes described at the gene abundance level in reads per kilobase million (RPKM). Points are colored by lifestyle and point size indicates infant age in months. Boxplots (bottom) show the distribution of indicated age groups in months along PCo1. Boxplots (right) show the distribution of each lifestyle along PCo2. The main axis of variation in this gene-based ordination is significantly associated with age (EnvFit; R^2 = 0.30; n = 679; P = 0.001) and the second axis of variation is significantly associated with lifestyle (EnvFit; R^2 = 0.35; n = 679; P = 0.001).

(B) Prevalence of species across lifestyles among infants 0 to 6 months old. VANISH (red and green) and BloSSUM (blue) species with the lowest adj-P values have text annotations. *B. infantis* is shown in orange. "Other" taxa (gray) are those that do not significantly differ according to lifestyle.

Figure 2: Age and lifestyle are associated with infant microbiome functions

Figure 2C/2D Cristina Blanco

Figure 2E Chayanee Chanpanich



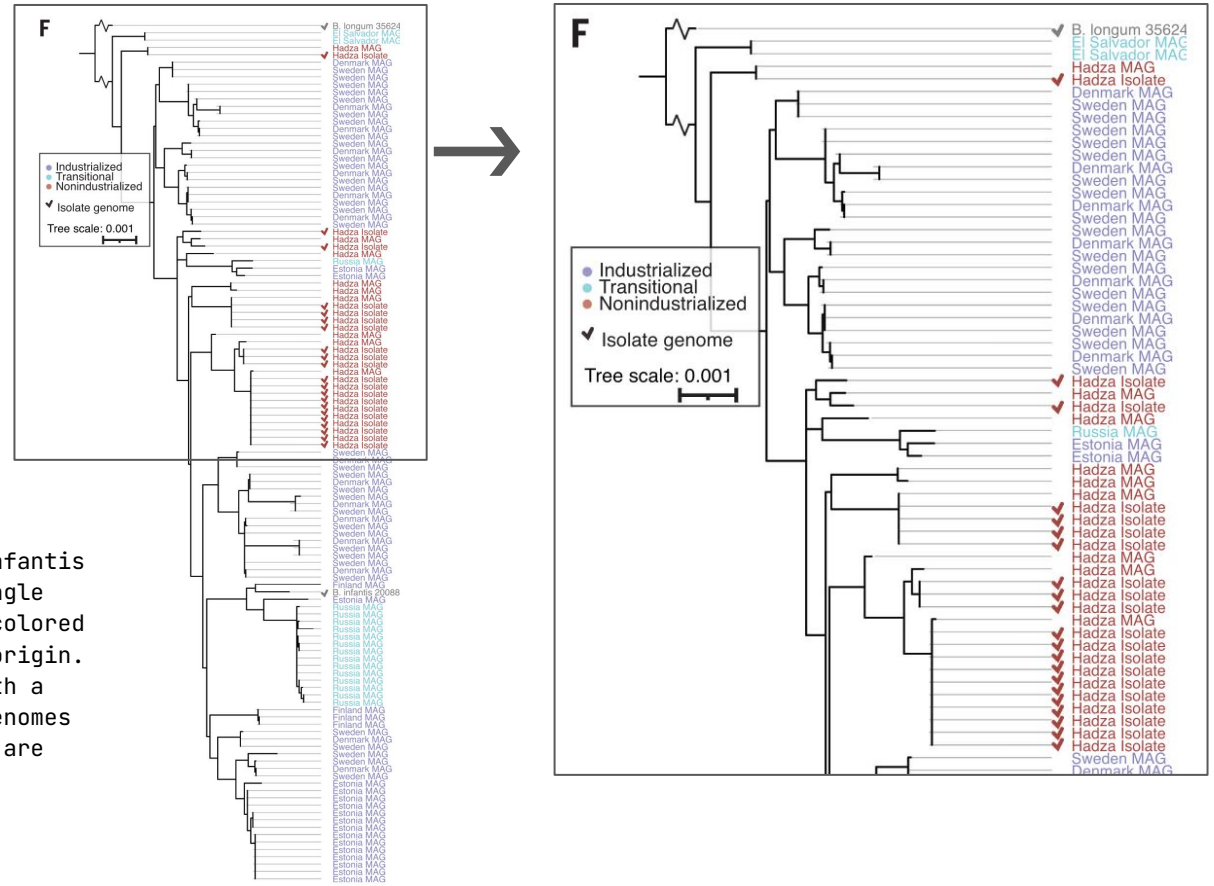
(C) Relative representation of four common Bifidobacterium species in infants 0 to 6 months old by lifestyle.

(D) Scatterplot of B. infantis versus B. breve abundance among infants 0 to 6 months old. Contour lines display the kernel density estimation.

(E) Prevalence of HMO-utilization clusters across ages and lifestyles. Clusters are considered present if all genes in the cluster are detected above a variable coverage threshold (to ensure that results are robust to differences in sequencing depth; see methods for details). * indicates adj-P < 0.05; Fisher's exact test with false discovery rate correction; nonindustrialized versus industrialized.

Figure 2: Age and lifestyle are associated with infant microbiome functions

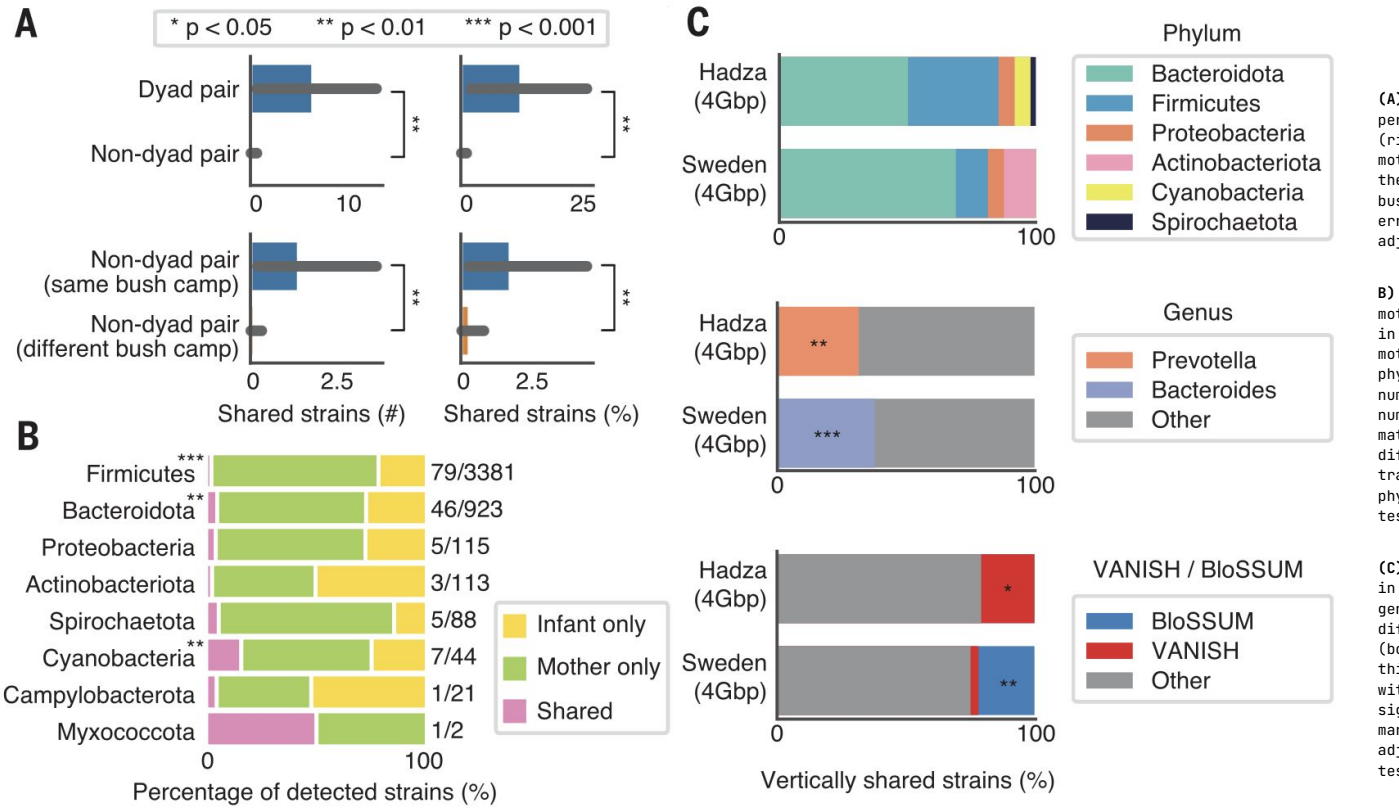
Figure 2F Amara Onwunzo



(F) Phylogenetic tree of B. infantis genomes based on universal single copy genes. Genome names are colored on the basis of lifestyle of origin. Isolate genomes are marked with a checkmark. Public reference genomes for B. longum and B. infantis are included (gray text).

Figure 3: Strain sharing between mother-infant dyads and non-dyads is lifestyle-specific

Figure 3A	Santiago G. Gonzalez
Figure 3B	Melanie Wilkinson
Figure 3C	Patricia Garcia



(A) The mean strains shared (left) and the percentage of infant strains found in mothers (right) in mother-infant dyads versus mother-infant nondyads (top) and nondyads from the same bushcamp versus nondyads from different bushcamps (bottom). Error bars represent standard error (*, adj-P < 0.05; **, adj-P < 0.01; ***, adj-P < 0.001; Wilcoxon rank-sum test). **(**

B) Percentage of strains detected in all Hadza mothers and infants and whether they are detected in infants only, mothers only, or shared within a mother-infant dyad ("shared") categorized by phylum. Numbers to the right of bars indicate the number of vertically shared strains over the number of strains detected in either infant or maternal samples. Phyla with a significant difference in the percentage of vertically transmitted strains as compared with all other phyla are marked with asterisks (Fisher's exact test with P value correction).

(C) Percentage of vertically transmitted strains in Hadza and Swedish cohorts by phylum (top), genus (middle; only genera with significant differences shown), and VANISH / BloSSUM (bottom). All metagenomes were subset to 4Gbp for this analysis to remove any biases associated with sequencing depth. Taxa that are significantly enriched in either cohort are marked with an asterisk (*, adj-P < 0.05; **, adj-P < 0.01; ***, adj-P < 0.001; Fisher's exact test)

Conclusions

1. Infant microbiomes develop along lifestyle-associated trajectories, diverging w/i the first 6 months of life in industrialized populations

2. 23.4% of species in Hadza infant guts are novel, and the infant gut harbors extensive uncharacterized diversity

3. *B. infantis* and HMO-utilization gene clusters are depleted in industrialized infants, even breastfed ones; and this could impact immune development

4. 310 VANISH species lost vs. only 12 BloSSUM species gained, so more diversity is lost than gained with industrialization

5. Vertical transmission is lifestyle-specific, as taxa that differentiate lifestyles are those most commonly shared between mother-infant dyads

6. Transitional populations show intermediate phenotypes, supporting lifestyle (not just geography) as a causative factor in microbiome assembly

Olm et al., Science 376, 1220-1223 (2022)

Question 1 Bianca Goncalves da Costa

Question 2 Alexander Jones

Question 3 Ally Mujica

Methods Critique Eliud Rivas Hernandez

Additional Analysis Ally Mujica

“Reviewer 3” Catie Wharton

Olm et al., Science 376, 1220-1223 (2022)