Supplementary Figures

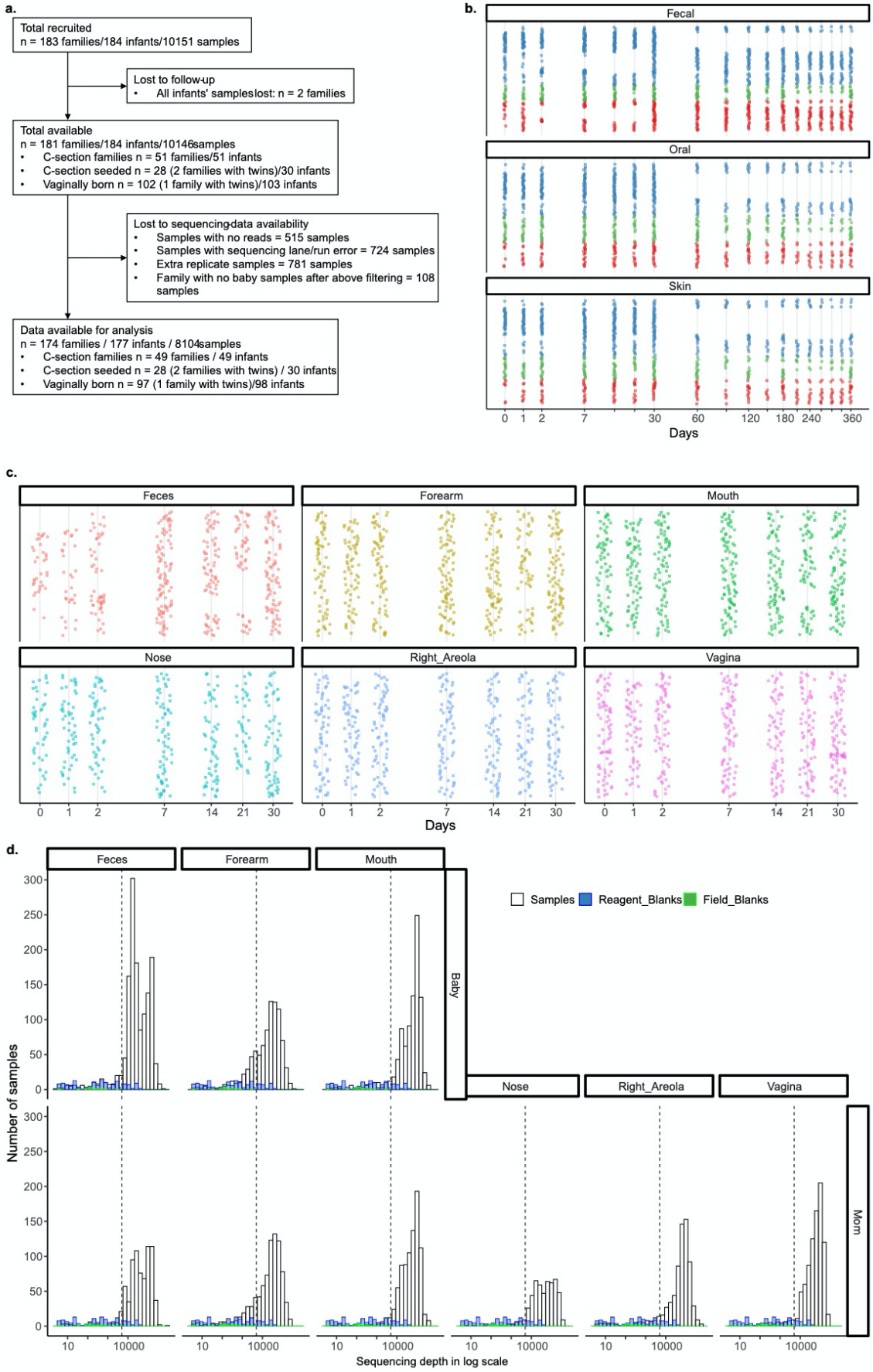

Figure S1. Longitudinal sampling of mother-infant pairs, Related to STAR Methods. (a) Number of families, infants and samples from the current study. (b) Longitudinal sampling of infant samples by birth modes and body

sites. Sampling with sterile swabs in different body sites took place within the first hour after birth in all babies, (including the vaginal gauze used to expose CS-born neonates) who were sampled within the hour after birth (after gauze swabbing of the CS-seeded babies), then at day 1-3, weekly for the first month, and monthly for up to the first year. Each row along y-axis is an individual baby. Each point represent a sample for one baby. The points are colored by birth modes, vaginal (blue), cesarean-seeded (green), and cesarean (red). On average, each baby contributed 18, 17, and 21 samples (across three body sites and multiple time points for the first year) for vaginal, cesarean, and cesarean-seed groups. (c) Longitudinal sampling of maternal samples by body sites within the first month after delivery. Each row along y-axis is an individual mom. On average each mom contributed 17 samples (across six body sites and multiple time points for the first month). (d) Distribution of number of reads per sample by different body sites in moms or babies Reagent blanks (blue), and field blanks (green) presentation were overplayed on each panel, and show much lower depth than the samples, indicating good overall quality of the sequences and lack of contamination. Dashed line marked the 5000 reads per sample position.

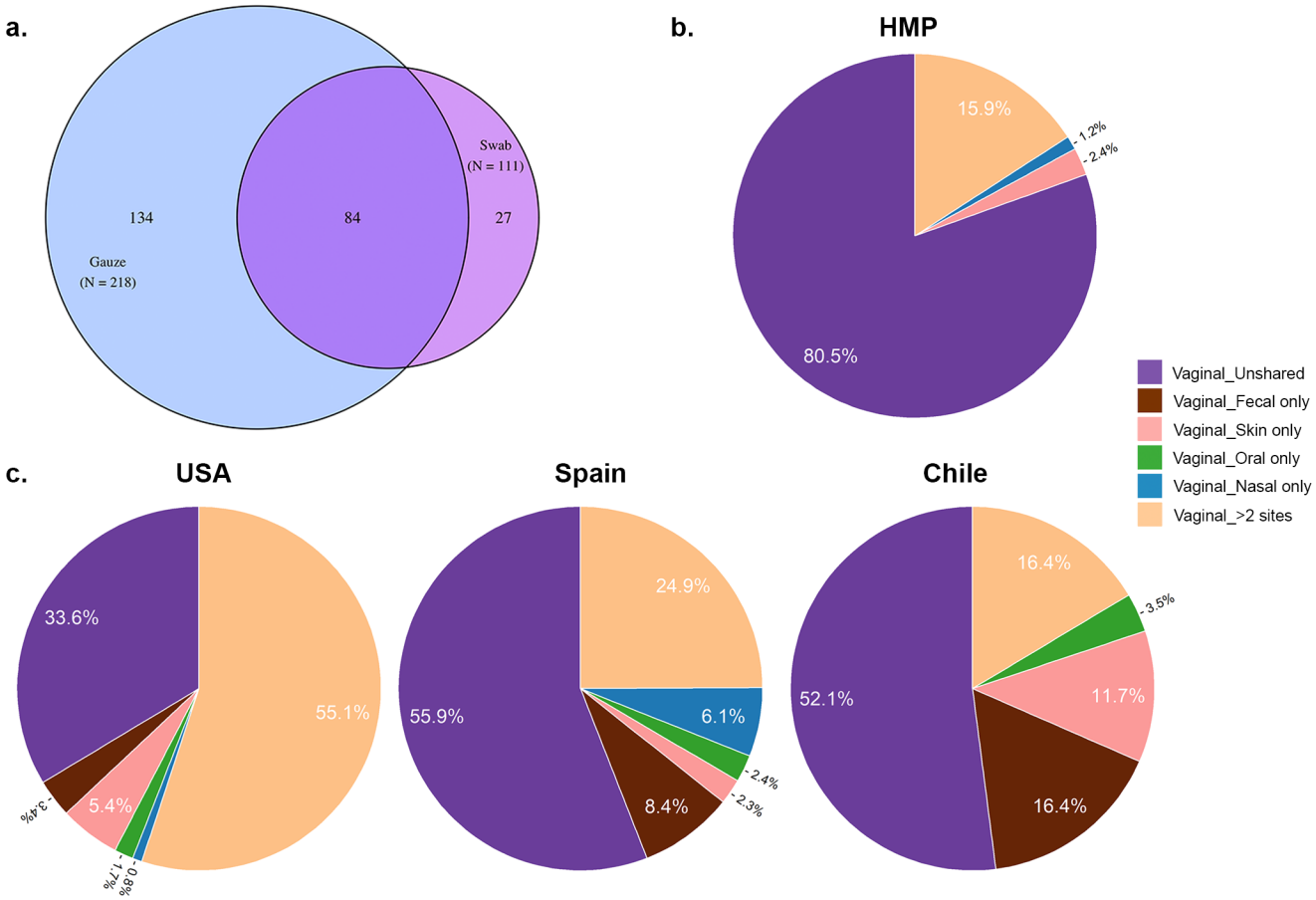

Figure S2. Pluripotential nature of perinatal vaginal microbiome (related to Figure 5, Supplementary Methods S13-S14). (a) Number of ASVs shared between perinatal vaginal swabs and vaginal gauzes. Gauzes show higher ASVs richness than vaginal swabs. Proportions of bacterial vaginal ASVs shared with other body sites. (b) in HMP data of non-pregnant women (105 women), (c) in parturient mothers at the day of delivery from USA (53 mothers), Spain (24 mothers) and Chile (20 mothers). HMP data was reprocessed by extracting V4 sequences and analyzed using the same pipeline.

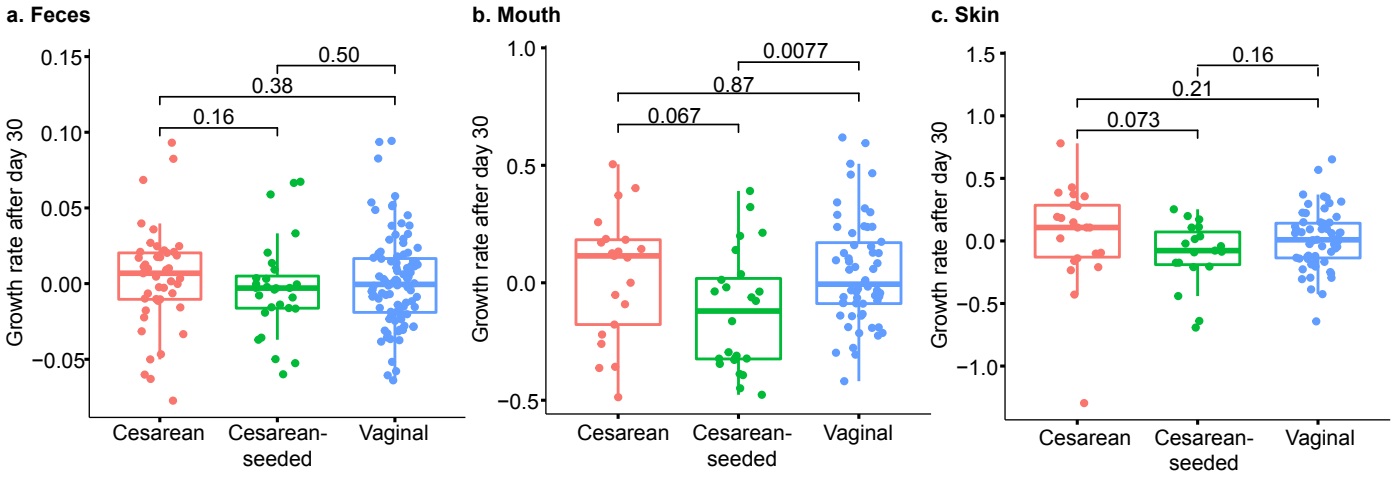

Figure S3. Bayesian Sparce Functional PCA (SFPCA) analyses on Shannon alpha diversity from 1 to 12 months of age, Related to STAR Methods. Bayesian Sparse Functional Principal Components Analysis (SFPCA) performed on Shannon alpha diversity across time did not differ by birth mode using Wilcoxon rank-sum test. The rate of growth of the Shannon diversity after day 30 (y-axes) is shown across birth modes (x-axis) for fecal (left), oral (middle), and skin (right) samples.

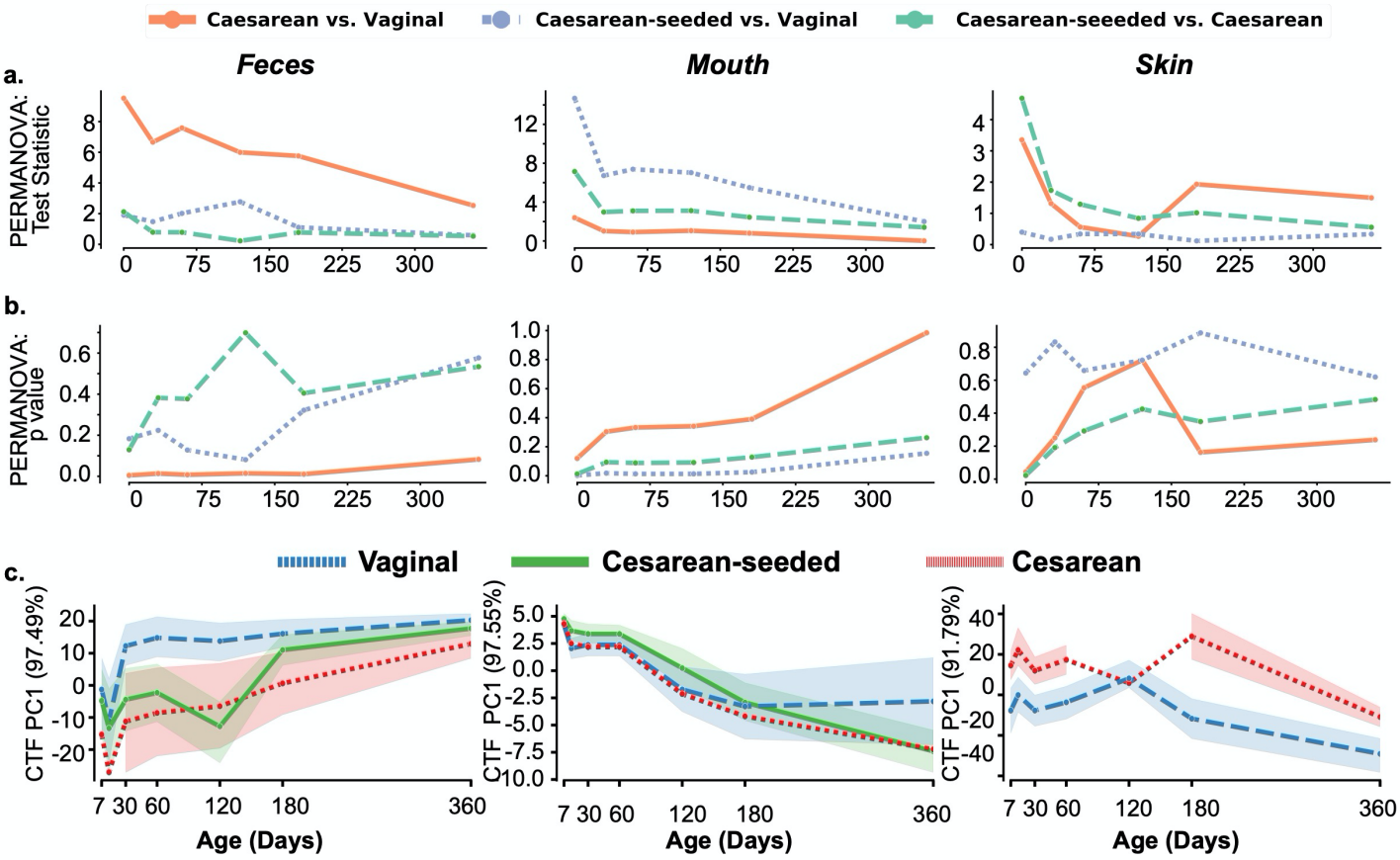

Figure S4. Compositional Tensor Factorization identifies the partial restoration of microbiome among cesarean-seeded babies, Related to Figure 1,2, 3. PERMANOVA of Aitchison distances from Compositional Tensor Factorization (CTF) during the first year of life. (a) PERMANOVA test statistics and (b) Bonferroni corrected p-values are plotted across age in days (x-axes). PERMANOVA plots are colored by compared pairs, Caesarean vs. Vaginal, Caesarean-seeded vs. Vaginal, and Caesarean-seeded vs. Caesarean. (c) Compositional Tensor Factorization (CTF) in the USA cohort. CTF ordination plot as in Figure 1a but only with the 101 US infants shows the same trends as the whole dataset, with vaginally born and seeded babies clustering together and separately from Cesarean-born infants. Comparison of Vaginal (blue; n=62), Cesarean (red; n=23), Cesarean-seeded (green; n=16) with CTF first principal component (y-axes) of infant samples over age in days (x-axes); error bars show the standard error of the mean. There were not enough sequences after filtering the samples in the skin of Cesarean-seeded babies.