# Studying Vertical Microbiome Transmission from Mothers to Infants by Strain-Level Metagenomic Profiling

Francesco Asnicar,ª Serena Manara,ª Moreno Zolfo,ª Duy Tin Truong,ª
Matthias Scholz,ª Federica Armanini,ª Pamela Ferretti,ª Valentina Gorfer,ᵇ
Anna Pedrotti,ᵇ Adrian Tett,ª Nicola Segataª

Centre for Integrative Biology, University of Trento, Trento, Italyª; Azienda Provinciale per i Servizi Sanitari, Trento, Italyᵇ

**ABSTRACT** The gut microbiome becomes shaped in the first days of life and continues to increase its diversity during the first months. Links between the configuration of the infant gut microbiome and infant health are being shown, but a comprehensive strain-level assessment of microbes vertically transmitted from mother to infant is still missing. We collected fecal and breast milk samples from multiple mother-infant pairs during the first year of life and applied shotgun metagenomic sequencing followed by computational strain-level profiling. We observed that several specific strains, including those of *Bifidobacterium bifidum*, *Coprococcus comes*, and *Ruminococcus bromii*, were present in samples from the same mother-infant pair, while being clearly distinct from those carried by other pairs, which is indicative of vertical transmission. We further applied metatranscriptomics to study the *in vivo* gene expression of vertically transmitted microbes and found that transmitted strains of *Bacteroides* and *Bifidobacterium* species were transcriptionally active in the guts of both adult and infant. By combining longitudinal microbiome sampling and newly developed computational tools for strain-level microbiome analysis, we demonstrated that it is possible to track the vertical transmission of microbial strains from mother to infants and to characterize their transcriptional activity. Our work provides the foundation for larger-scale surveys to identify the routes of vertical microbial transmission and its influence on postinfancy microbiome development.

**IMPORTANCE** Early infant exposure is important in the acquisition and ultimate development of a healthy infant microbiome. There is increasing support for the idea that the maternal microbial reservoir is a key route of microbial transmission, and yet much is inferred from the observation of shared species in mother and infant. The presence of common species, *per se*, does not necessarily equate to vertical transmission, as species exhibit considerable strain heterogeneity. It is therefore imperative to assess whether shared microbes belong to the same genetic variant (i.e., strain) to support the hypothesis of vertical transmission. Here we demonstrate the potential of shotgun metagenomics and strain-level profiling to identify vertical transmission events. Combining these data with metatranscriptomics, we show that it is possible not only to identify and track the fate of microbes in the early infant microbiome but also to investigate the actively transcribing members of the community. These approaches will ultimately provide important insights into the acquisition, development, and community dynamics of the infant microbiome.

**KEYWORDS** infant microbiome, metagenomics, microbial ecology, microbial genomics, vertical microbiome transmission

The community of microorganisms that dwell in the human gut has been shown to play an integral role in human health (1–4), facilitating, for instance, the harvesting of nutrients that would otherwise be inaccessible (5), modulating the host metabolism and immune system (6), and preventing infections by occupying the ecological niches that could otherwise be exploited by pathogens (7). The essential role of the intestinal microbiome is probably best exemplified by the successful treatment of dysbiotic states, such as chronic life-threatening *Clostridium difficile* infections, using microbiome transplantation therapies (8–10).

The gut microbiome is a dynamic community shaped by multiple factors throughout an individual's life, possibly including prebirth microbial exposure. The early development of the infant microbiome has been proposed to be particularly crucial for longer-term health (11–13), and a few studies have investigated the factors that are important in defining its early structure (14–17). In particular, gestational age at birth (17), mode of delivery (14, 15), and early antibiotic treatments (18) have all been shown to influence the gut microbial composition in the short term and the pace of its development in the longer term.

Vertical transmission of bacteria from the body and breast milk of the mother to her infant has gained attention as an important source of microbial colonization (14, 19–21) in addition to the microbial organisms obtained from the wider environment (22, 23), including the delivery room (24). Results from early cultivation-based and cultivation-free methods (16S rRNA community profiling and a single metagenomic study) have indeed suggested that the mother could transfer microbes to the infant by breastfeeding (25) and that a vaginal delivery has the potential of seeding the infant gut with members of the mother's vaginal community (11, 14, 26, 27) that would not be available via caesarean section. However, a more in-depth analysis is required to elucidate the role of vertical transmission in the acquisition and development of the infant gut microbiome.

Current knowledge of the vertical transmission of microbes from mothers to infants has hitherto focused on the cultivable fraction of the community (28) or lacked strain-level resolution (11). Many microbial species are common among unrelated individuals (29); therefore, in instances where a species is identified in both mother and infant (13, 30), it remains inconclusive if this is due to vertical transmission. Strain-level analysis has shown that different individuals are associated with different strains of common species (31, 32), and it is therefore crucial to profile microbes at the strain level to ascertain the most probable route of transfer. This has been performed only for specific microbes by cultivation methods (16, 28), but many vertically transmitted microorganisms remain hard to cultivate (16); thus, the true extent of microbial transmission remains unknown. A further crucial aspect, still largely unexplored, is the fate of vertically acquired strains: if they are transcriptionally active rather than merely transient, that may suggest possible colonization of the infant intestine. Although studies have described the transcriptional activity of intestinal microbes under different conditions (33–36), no studies have applied metatranscriptomics to characterize the activity of vertically transmitted microbes *in vivo*.

In this work, we present and validate a shotgun metagenomic pipeline to track mother-to-infant vertical transmission of microbes by applying strain-level profiling to members of the mother and infant microbiomes. Moreover, we assessed the transcriptional activity of vertically transmitted microbes to elucidate if transferred strains are not only present but also transcriptionally active in the infant gut.

## RESULTS AND DISCUSSION

We analyzed the vertical transmission of microbes from mother to infant by enrolling 5 mother-infant pairs and collecting fecal samples and breast milk (see Materials and Methods) when each infant was 3 months of age (time point 1). Two mother-infant pairs (pair 4 and pair 5) were additionally sampled at 10 months postbirth (time point 2), and one pair (pair 5) was sampled at 16 months postbirth (time point 3; see Fig. S1 in the supplemental material). We applied shotgun metagenomic sequencing to all 24 microbiome samples (8 mother fecal samples, 8 infant fecal samples, and 8 milk samples), generating 1.2 G reads (average, 39.6 M reads/sample; standard deviation [SD], 28.7 M reads/sample) (see Ta-
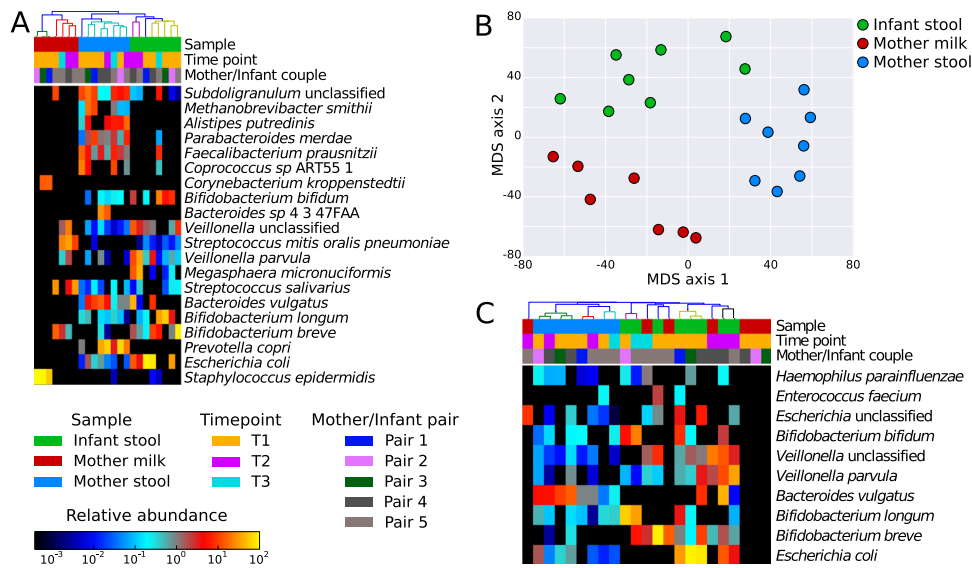
**FIG 1** Microbial composition of mother and infant samples and shared bacteria within mother-infant pairs. (A) Quantitative microbial taxonomic composition of the metagenomic samples from milk and fecal samples of mothers and infants as estimated by MetaPhlAn2 analysis (77) (only the 20 most abundant species are indicated). Milk samples present low microbial richness compared to fecal samples. (B) Ordination plot of microbiome composition showing clustering of the three different sample types: mother feces, infant feces, and breast milk samples. The two infant samples close to the cluster of mother feces and in between the clusters of mothers and infants are from later time points, denoting the convergence of the infant microbiome toward an adult-like one. (C) The abundances of the 10 microbial species detected (>0.1% abundance) in at least one infant and the respective mother (shared species have been identified on the basis of samples from time point 1 [T1] only).

ble S1 in the supplemental material). Metatranscriptomics (average, 90.55 M reads/sample; SD, 46.86 M reads/sample) was also applied on fecal samples of two pairs (pair 4 at time point 2 and pair 5 at time point 3) to investigate the differential expression profiles of the bacterial strains in the gut of mothers and their infants.

**Shared mother-infant microbial species.** In our cohort, the infant intestinal microbiome was dominated by *Escherichia coli* and *Bifidobacterium* spp., such as *B. longum*, *B. breve*, and *B. bifidum* (Fig. 1A and S2). These species in some cases reached abundances higher than 75% (e.g., *E. coli* at 85.2% in infant pair 3 at time point 1 and *B. breve* at 78.8% in infant pair 5 at time point 1), which is consistent with previous observations (12, 37, 38). As expected, the intestines of the mothers had a greater microbial diversity than those of the infants, with high abundances of *Prevotella copri*, *Clostridiales* (e.g., *Coprococcus* spp. and *Faecalibacterium prausnitzii*), and *Bacteroidales* (e.g., *Parabacteroides merdae* and *Alistipes putredinis*). Interestingly, the postweaning microbiome of infant of pair 5 (time point 3, 16 months postbirth) had already shifted toward a more "mother-like" composition (Fig. 1B), with an increase in diversity and the appearance of *Parabacteroides merdae*, *Coprococcus* spp., and *Faecalibacterium prausnitzii* (13, 38). Nevertheless, this 16-month-old infant still retained some infant microbiome signatures, such as a high abundance of bifidobacteria that were present at only low levels in the mothers' samples (Fig. 1A and C).

We extracted and successfully sequenced microbial DNA from 7 of 8 milk samples. Microbial profiling of milk samples was hindered by a high abundance of interfering molecules (proteins, fats, proteases— e.g., plasmin—and calcium ions) (39–41) that affected the efficiency of the extraction and amplification steps. Even so, we obtained an average of 3.08 Gb (SD, 1.5 Gb) per sample, of which 26 Mb (SD, 56 Mb) were from nonhuman reads (a level higher than that seen in the only other metagenomic study) (42) (see Table S1).

Milk samples had limited microbial diversity at the first sampling time (time point 1, 3 months postbirth) and included skin-associated bacteria such as *Corynebacterium kroppenstedtii* and *Staphylococcus epidermidis*. Cutaneous taxa, however, were observed in only low abundances in the gut microbiome of infants, confirming that skin microbes

are not colonizers of the human gut (Fig. 1A). At later time points, the milk samples were enriched in *B. breve* and in bacteria usually found in the oral cavity, such as *Streptococcus* and *Veillonella* spp. The presence of oral taxa in milk has been previously observed by 16S rRNA sequencing (14, 20, 21, 25) and shotgun metagenomics (42). This could be caused by retrograde flux into the mammary gland during breastfeeding (43) whereby cutaneous microbes of the breast and from the infant oral cavity are transmitted to the breast glands (44). However, this remains a hypothesis because no oral samples were collected in this study. These observations are summarized in the ordination analysis (Fig. 1B), in which the different samples (infant feces, mother feces, and milk) clustered by type, with weaning representing a key factor in the shift from an infant to an adult-like microbiome structure (13, 38, 45).

Comparing the species present in both the mother and infant pairs (Fig. 1C), we observed that many shared species (e.g., *Escherichia*, *Bifidobacterium*, and *Veillonella* spp.) occurred at a much higher abundance in the infant than in the mother, possibly due to the lower level of species diversity and therefore to competition in the gut. *Bacteroides vulgatus* was found at relatively high abundance (average, 16.3%; SD, 13%) in both the infant and the mother of pair 4 at both time point 1 and time point 2. The presence of shared species in mother-infant pairs observed here and elsewhere (14, 16, 17, 25, 46) confirms that mothers are a potential reservoir of microbes vertically transmissible to infants, but it remains unproven whether the same strain is transmitted to the infant from the mother or if an alternative transmission route is involved.

**Strains shared between mothers and infants are indicative of vertical transmission.** While different individuals have a core of shared microbial species, it has been shown that these common species consist of distinct strains (31, 32). To analyze microbial transmission, it is therefore crucial to assess whether a mother and her infant harbor the same strain. To this end, we further analyzed the metagenomic samples at a finer strain-level resolution. This was achieved by applying a recent strain-specific pangenome-based method called PanPhlAn (31), as well as a genetics-based method called StrainPhlAn (D. T. Truong, A. Tett, E. Pasolli, C. Huttenhower, and N. Segata, submitted for publication) (see Materials and Methods), which identifies single-nucleotide variants (SNVs) in species-specific marker genes.

Using the SNV-based analysis, we observed considerable strain-level heterogeneity in the species present in the intestines of the mothers also with respect to available reference genomes (Fig. 2; see also Fig. S3 in the supplemental material). This heterogeneity was not observed within the mother-infant pairings, as in the case of *Bifidobacterium* spp., *Ruminococcus bromii*, and *Coprococcus comes*. The infant of pair 4 at time point 2, for example, harbored a strain of *B. bifidum* that matched his mother's at 99.96% sequence identity and yet was clearly distinct from the *B. bifidum* strains of other infants in the cohort (Fig. 2A), which differed by at least 0.6% of the nucleotides. The observation that the *B. bifidum* strains from the mother and the infant of pair 4 were too similar to be consistent with the observed strain-level variation across subjects in the cohorts was highly statistically significant ($P$ value, 4.7e−40) (see Fig. S4). This was also true for the *C. comes* ($P$ value, 1.9e−3) (99.87% intrapair similarity and 1.6% and 1.61% divergence compared to the closest strain and the average value, respectively) (Fig. 2B) and *R. bromii* ($P$ value, 4.9e−8) (99.93% similarity and 1.53% and 2.63% diversity—same as described above) (Fig. 2C) strains that were shared by pair 5. Mother-infant sharing of the same strain was also confirmed by strain-level pangenome analysis (31) that showed that the strains from the same pair carried the same unique gene repertoire (see Fig. S5). It is accepted that, while the possibility of independent acquisition of strains from a shared environmental source cannot be excluded, the finding that mother-infant pairs have shared strains represents strong evidence of vertical microbiome transmission. On average, we could reconstruct and observe vertical transmission from mother to infant for 14% of the species found to be shared within mother and infant pairings.

Strain transmission does not, however, exclude later replacement of the vertically acquired organisms, as we highlighted by looking at the postweaning time point in our
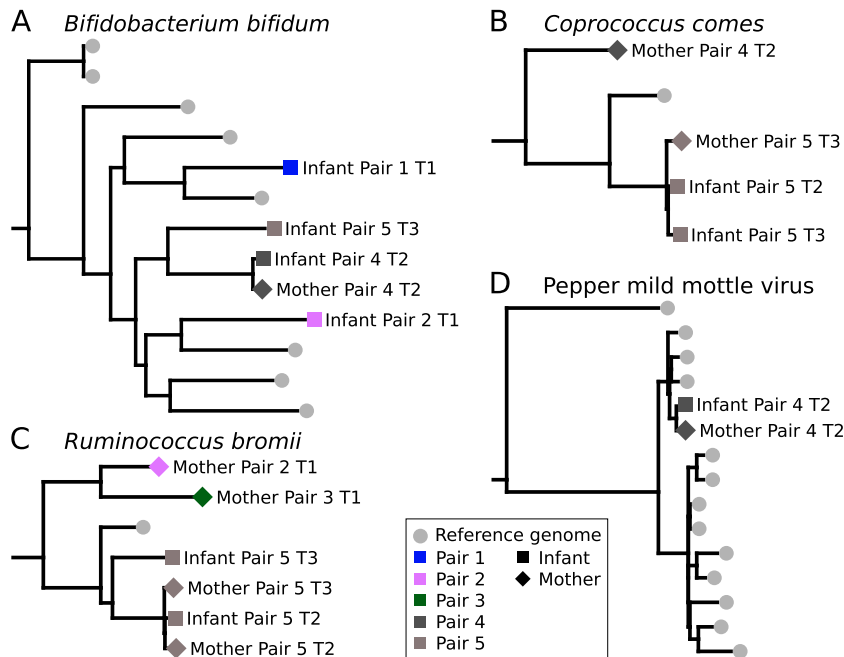
**FIG 2** Strain-level phylogenetic trees for microbes present in both the mother and infant. Phylogenetic trees were built by the StrainPhlAn method using species-specific markers confirming the presence of the same strain in the mother and infant intestinal microbiomes, thus suggesting vertical transmission. Available reference genomes were included in the phylogenetic trees. Here we report three bacterial species, namely, (A) *Bifidobacterium bifidum*, (B) *Coprococcus comes*, and (C) *Ruminococcus bromii*, and the most abundant viral species found in pair 4, (D) pepper mild mottle virus. Other species-specific phylogenetic trees (*B. adolescentis*, *B. breve*, and *B. longum*) are reported in Fig. S3.

cohort (pair 5 at time point 3) which harbored the highest number of shared species, with 70.4% present in the infant and mother (at a relative abundance of >0.1%, according to the MetaPhlAn2 profiles). A proportion (11%) of these common species were shown to be the same strain (Fig. 2; see also Fig. S3 in the supplemental material), according to both PanPhlAn and StrainPhlAn analyses (see, for example, the data from *B. adolescentis* and *C. comes*) (Fig. 2; see also Fig. S3 and S5). However, some strains that were shared at earlier time points were replaced at time point 3. Of note, the *R. bromii* strain found in an infant at time point 3 was different from that found at time point 2, and both strains were distinct from the strain observed in the mother at both time points (Fig. 2C). This was also observed for the latter infant time point for *B. breve* (see Fig. S3B) and *B. longum* (see Fig. S3C). Although it is not possible to generalize these results because of the small sample size, these replacement events suggest that originally acquired maternal strains can subsequently be replaced (47, 48).

We then extended our analysis to the viral organisms detectable from metagenomes and metatranscriptomes, as viruses have the potential to be vertically transmitted also. The DNA viruses identified from our metagenome samples largely consisted of bacteriophages of the *Caudovirales* order, a common order of tailed bacteriophages found in the intestine (3, 49). We identified *Enterobacter* and *Shigella* phages as the most prevalent phages among the tested samples, in agreement with the high prevalence of members of the *Enterobacteriaceae* family and particularly of members of the *Escherichia* genus (see Fig. 1A and Table S3). We also identified crAssphage at high breadth of coverage (50) and provided further evidence for the hypothesis that the *Bacteroides* genus is the host for this virus (50), as the microbiome of crAssphage-positive mothers was enriched in *B. vulgatus* (see Fig. 1A and Table S3). However, the low breadth of coverage for many of the DNA viruses made it difficult to identify pair-specific phage variants (see Table S3). Analysis of the RNA viruses from the metatranscriptomic samples identified instead the presence of an abundant pepper mild mottle virus

(PMMoV), a single-stranded positive-sense RNA virus of the genus *Tobamovirus*, in all of the four metatranscriptomes from pairs 4 and 5. Surprisingly, transcripts from the PMMoV were found in greater abundance than all the other microbial transcripts found for the mother of pair 4. PMMoV has already been reported in the gut microbiome (51–53), and other related viruses of the same family have been shown to be able to enter and persist in eukaryotic cells (54, 55). The high abundance of PMMoV in mother-infant pair 4 allowed us to reconstruct its full genome (99.9%) and to perform a phylogenetic analysis demonstrating that the mother and the infant shared identical PMMoV strains, which were clearly distinct from the PMMoV reference genomes (27 SNVs in total; Fig. 2D). Although the coverage was lower, the same evidence of a shared PMMoV strain was observed within pair 5. The analysis of PMMoV polymorphisms within each sample also suggests the coexistence of different PMMoV haplotypes in the same host (Fig. S6). Although vertical transmission of RNA viruses and PMMoV specifically would be intriguing, because of the age and dietary habits of the infants (see Table S1) this finding could be related to the exposure to a common food source (56). Our analysis of the virome characterized directly from shotgun metagenomics thus highlighted that viruses can be tracked across mother-infant microbiomes also and that experimental virome enrichment protocols (57, 58) have the potential to provide an even clearer snapshot of viral vertical transmission.

**Differences in the overall levels of functional potential and expression in mothers and infants.** The physiology of the mammary gland (milk) as well as the adult and infant intestine is reflected by niche-specific microbial communities as reported above and in previous studies (11, 13, 15, 20, 21, 44, 45). To characterize the overall functional potential of the microbial communities inhabiting these niches, we complemented the taxonomic analysis above by employing HUMAnN2 (see Materials and Methods). As expected, there was considerable overlap in the functionality of the gut microbiomes of the mothers and infants (Fig. 3A), with 87% of pathways present in mother and infant, 50% of which were significantly different in abundance (at an alpha value of 0.05). Nevertheless, there were notable differences. For instance, the microbiomes of the infants showed a higher potential for utilization of intestinal mucin as a carbon source (*P* value, 0.016) and for folate biosynthesis (*P* value, 1.8e−6) while displaying a lower potential for starch degradation (*P* value, 9.8e−6), consistent with previous observations (12, 59–62). Mucin utilization, specifically by infant gut microbial communities, is reflective of the higher abundance of mucin-degrading bifidobacteria observed from the taxonomic analyses described above (12, 59–61), whereas increased folate biosynthesis (12, 59, 60, 62) and decreased starch degradation (5) have been purported to represent responses to the limited dietary intake in infants compared to adults. Interestingly, the intestinal samples from the postweaning infant of pair 5 (16 months postbirth) clustered together with the adults' intestinal samples (Fig. 3B), suggesting that the shift toward an adult-like microbiome observed in the taxonomic profiling (Fig. 1B) is also reflected by or is a consequence of a change in community functioning. Among the most prevalent pathways in the milk microbiomes that we observed were those involved in galactose and lactose degradation (63), as well as in biosynthesis of aromatic compounds (Fig. S7A). This was specifically true for production of chorismate, a key intermediate for the biosynthesis of essential amino acids and vitamins found in milk (62) (Fig. 3C and S7A).

To further evaluate the functional capacity of the gut-associated microbiomes and analyze the *in vivo* transcription, we performed metatranscriptomics analyses of the feces of two mother-infant pairs (see Materials and Methods). HUMAnN2 was used to identify differences in the transcriptional levels of pathways in the gut of the mothers and infants. The most notable global difference was that fermentation pathways were highly transcribed in the mother compared to that of the infant. This reflects the transition of the gut from an aerobic to an anaerobic state and the associated shift from facultative anaerobes to obligate anaerobes over the first few months of life (64, 65). The same is true for pathways involved in starch degradation, which were not only poorly represented in the metagenomes but also negligibly expressed in the infants'
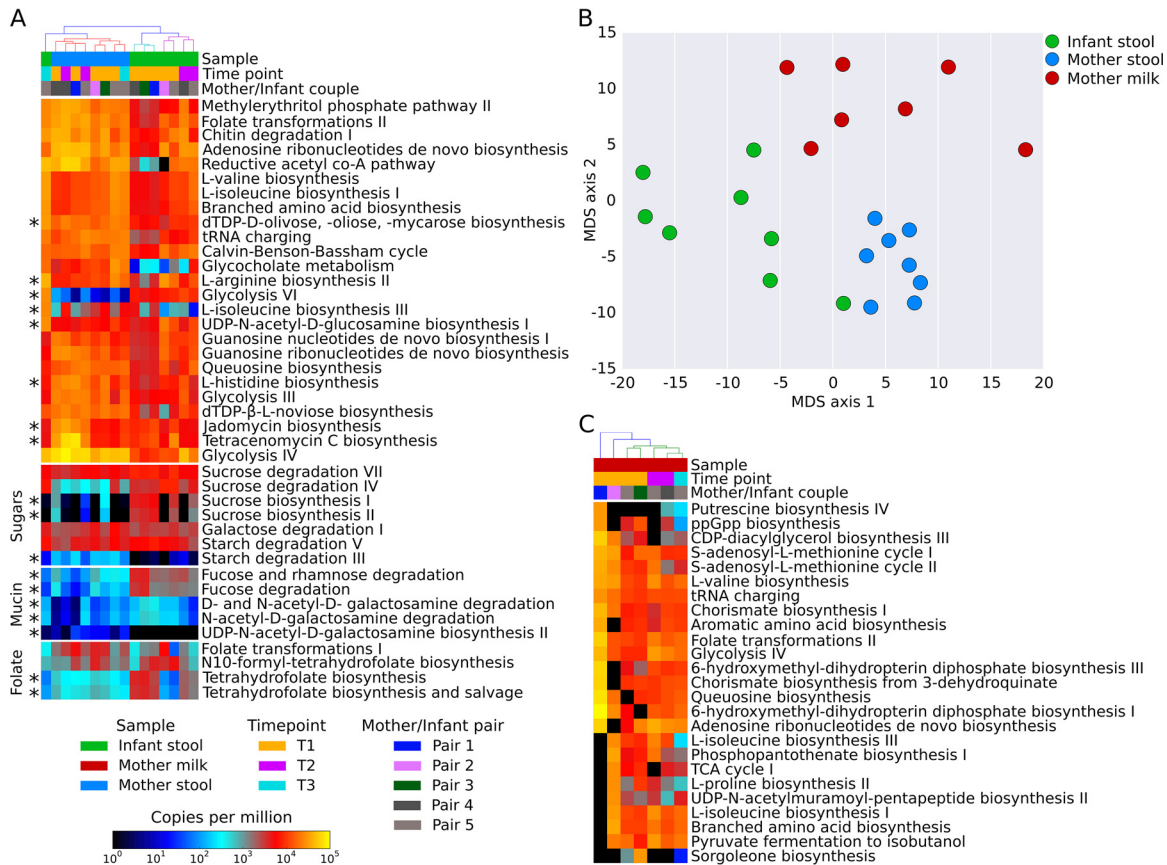
**FIG 3** Functional potential analyses. (A) HUMAnN2 heat map reporting the 25 most abundant pathways in the fecal samples of mothers and infants. Specific pathways of interest (sugars, mucin, and folate metabolism) are added at the bottom. The asterisk (*) near the heat map highlights statistically significant pathways. (B) Multidimensional scaling (MDS) result from functional potential profiles, showing the differences between fecal samples of mothers and infants and milk samples. In particular, the infant feces point in the mother feces cluster corresponds to time point 3 of pair 5, showing a shift from the infant microbiome toward an adult-like microbiome. (C) HUMAnN2 results for the 25 most abundant pathways found only in the milk samples. TCA, tricarboxylic acid.

transcriptomes. What is evident is that the transcriptional patterns for different members differed considerably, as illustrated for pair 4 and pair 5 (Fig. 4A and S7B, respectively). For example, we observed in the infant of pair 4 that *B. vulgatus* was more transcriptionally active (average of 2.7 [SD, 2.5] normalized transcript abundance [NTA]; see Materials and Methods) than both *E. coli* (245-fold change [average, 0.4 SD and 0.6 NTA]) and *Bifidobacterium* spp. (6.6-fold change [average, 0.01 SD and 0.01 NTA]). Although these differences were statistically significant (*P* values were lower than 1e−50 in both cases), their physiological significance remains unclear.

**Strain-specific transcriptional differences in mothers and infants.** To further explore the transcriptional activity of the intestinal microbiomes and, more specifically, to ascertain which individual microbial members are transcriptionally active in the gut, we employed the strain-specific metatranscriptomic approach implemented in PanPhlAn (31) (see Materials and Methods). Of particular interest is the transcriptional activity of the shared mother-infant strains that, based on our strain-level analyses, are likely to have been vertically acquired by the infant by the maternal route. Such transcriptional analyses can clarify whether these transmitted strains were not only present in the infant gut but also functioning, therefore suggesting that the transmitted strains could have potentially colonized. For three transmitted species in pair 4 (*B. vulgatus*, *E. coli*, and *B. bifidum*), we show that they were active in both the mother intestine and the infant intestine (Fig. 4B). Of note is that *B. bifidum* was more active in the infant than in the mother (2.5-fold change; Fig. 4B), which was expected as this species is a known early colonizer of the infant gut (12, 37, 38). Interestingly, the *B. bifidum* strain of pair 5 showed the opposite behavior (Fig. S7C). We
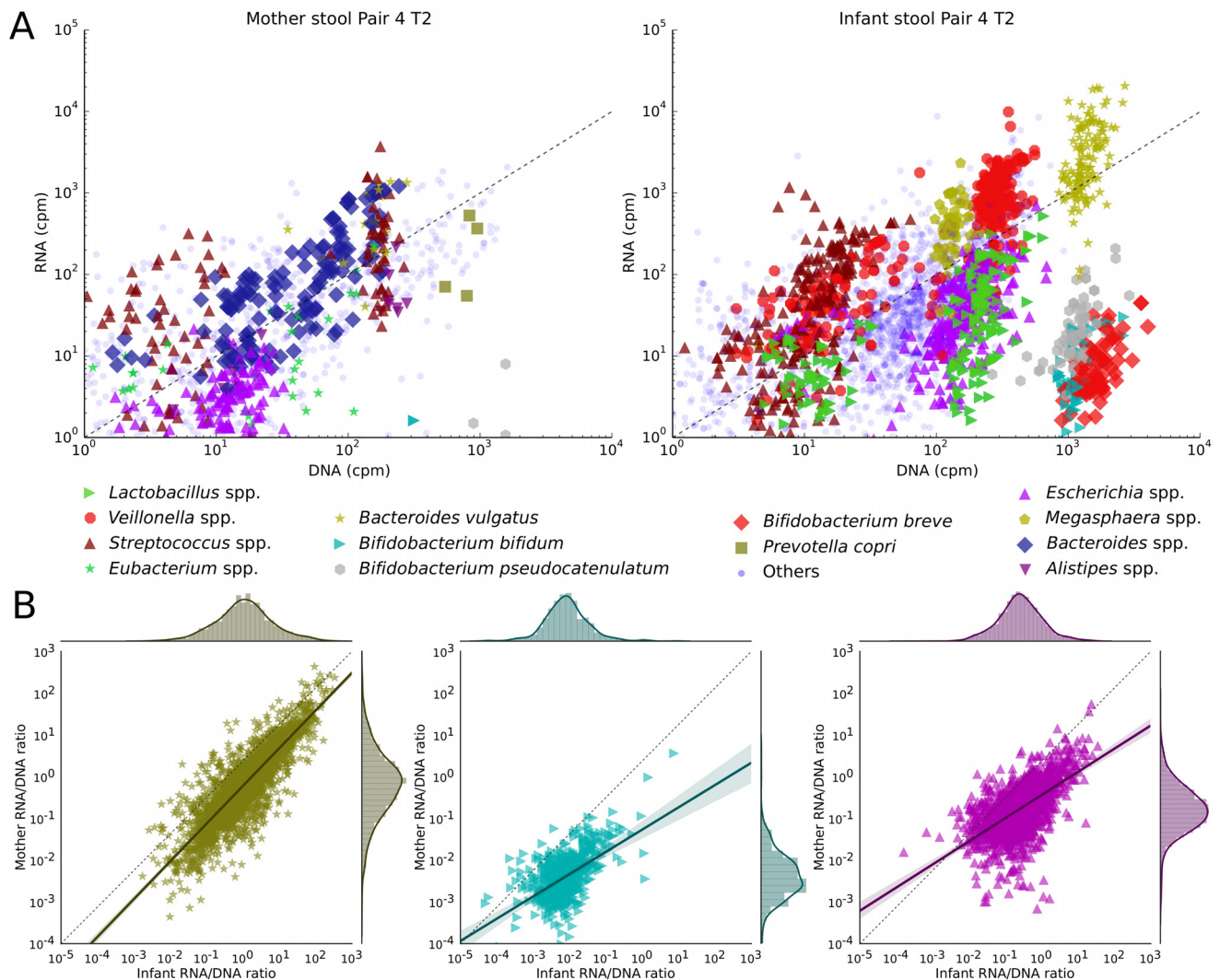
**FIG 4** Transcription levels of metabolic pathways and genes in mother and infant pair 4 at time point 2. (A) Scatterplots showing the transcription rates of metabolic pathways of shared and nonshared species and genera of interest for both the mother and infant of pair 4 at time point 2. (B) Comparison between transcription rates of gene families in mother and infant gut microbiomes.

postulate that this was because the infant of pair 5 was of postweaning age (10% breast milk diet) compared to the infant of pair 4 (90% breast milk diet) and that the difference reflects the change in substrate availability from breast milk to solid food, which might have a detrimental effect on the bifidobacterial population (38, 66, 67). Moreover, in support of our metagenomics analyses indicating that the microbiome of infant of pair 5 was shifting toward a more adult-like structure (Fig. 1B), we observed high transcriptional activity for *R. bromii*, a species commonly associated with adults, which could be seen as a hallmark of this transition (68, 69).

It is well established that metatranscriptomic profiling provides a more accurate account of the actual community functioning than metagenomics alone. Here we show that the combination of the two approaches affords the exploration of which members not only are transmitted but also are actively participating in the community and therefore offers a more detailed account of the microbial community dynamics.

**Conclusions.** Human-associated microbiomes are complex and dynamic communities that are continuously interacting with the host and are under the influence of environmental sources of microbial diversity. Identifying and understanding the transmission from these external sources are crucial to understanding how the infant gut is

colonized and ultimately develops an adult-like composition. However, detecting direct transmission is not a trivial task: many species are ubiquitous in host-associated environments and in the wider environment alike, and yet they comprise a myriad of different strains and phenotypic capabilities. Therefore, detection of microbial transmission events requires the ability to characterize microbes at the strain level. The epidemiological tracking of pathogens by cultivation-based isolate sequencing has proven successful (70, 71), but it relies on time-consuming protocols and can focus on only a limited number of species. In contrast, while there have been some examples of strain-level tracking from metagenomic data (71, 72), this remains challenging. In this study, we developed methods for identifying the vertical flow of microorganisms from mothers to their infants and showed that mothers are sources of microbes that might be important in the development of the infant gut microbiome.

We demonstrated that high-resolution computational methods applied to shotgun metagenomic and metatranscriptomic data enable the tracking of strains and strain-specific transcriptional patterns across mother-infant pairs. In our cohort of five mother-infant pairs, we detected several species with substantial genetic diversity between different pairs but identical genetic profiles in the mother and her infant, indicative of vertical transmission. These include some bifidobacteria typical of the infant gut (i.e., *B. longum*, *B. breve*, *B. bifidum*, and *B. adolescentis*) but also *Clostridiales* species usually found in the adult intestine (i.e., *R. bromii* and *C. comes*) and viral organisms. These results confirm that the infant receives a maternal microbial imprinting that might play an important role in the development of the gut microbiome in the first years of life.

The strain-level investigation of vertically transmitted microbes was followed by characterization of the transcriptional activity of the transmitted strains in the mother and infant environments. We found that the transcriptional patterns of strains shared within the single pairs were different between mother and infant, suggesting successful adaptation of maternally transmitted microbes to the infant gut.

Taking the results together, our work provides preliminary results and methodology to expand our knowledge of how microbial strains are transmitted across microbiomes. Expanding the cohort size and considering other potential microbial sources of transmission, such as additional mother and infant body sites, as well as other family members (i.e., fathers and siblings) and environments (hospital and house surfaces), will likely shed light on the key determinants in early infant exposure and the seeding and development of the infant gut microbiome.

## MATERIALS AND METHODS

**Sample collection and storage.** In total, five mother-infant pairs were enrolled. Fecal samples and breast milk were collected for all pairs at 3 months (time point 1); additional samples were collected for pair 4 and pair 5 at 10 months (time point 2) and for pair 5 only at 16 months (time point 3) (see Table S1 and Fig. S1 in the supplemental material). All aspects of recruitment and sample and data processing were approved by the local ethics committee. Fecal samples were collected from mothers and infants in sterile feces tubes (Sarstedt, Nümbrecht, Germany) and immediately stored at −20°C. In those cases where metatranscriptomics was applied, a fecal aliquot was removed prior to freezing the remaining feces. This aliquot was stored at 4°C, and the RNA was extracted within 2 h of sampling to preserve RNA integrity. Milk was expressed and collected midflow by mothers into 15-ml centrifuge tubes (VWR, Milan, Italy) and immediately stored at −20°C. Within 48 h of collection, all milk samples and feces samples were moved to storage at −80°C until processed.

**Extraction of nucleic acids for metagenomic analysis.** DNA was extracted from feces using a QIAamp DNA stool minikit (Qiagen, Netherlands). Milk DNA was extracted using a PowerFood microbial DNA isolation kit (Mo Bio, Inc., CA). Both procedures were performed according to the specifications of the manufacturers. Extracted DNA was purified using an Agencourt AMPure XP kit (Beckman Coulter, Inc., CA). Metagenomic libraries were constructed using a Nextera XT DNA library preparation kit (Illumina, CA, USA) according to manufacturer instructions and were sequenced on a HiSeq 2500 platform (Illumina, CA, USA) at an expected sequencing depth of 6 Gb/library.

**Extraction of nucleic acids for metatranscriptomic analysis.** Fecal samples for metatranscriptomic profiling were pretreated as described previously (73). Briefly, 110 $\mu$l of lysis buffer (30 mM Tris-Cl, 1 mM EDTA [pH 8.0], 1.5 mg/ml of proteinase K, and 15 mg/ml of lysozyme) was added to 100 mg of feces and incubated at room temperature for 10 min. After pretreatment, samples were treated with 1,200 $\mu$l of Qiagen RLT Plus buffer (from an AllPrep DNA/RNA minikit [Qiagen, Netherlands]) containing 1% (vol) beta-mercaptoethanol and were transferred into 2-ml sterile screw-cap tubes (Starstedt, Germany) filled with 1 ml of zirconia-silica beads (BioSpec Products, OK, USA) (<0.1 mm in diameter). Tubes were placed on a Vortex-Genie 2 mixer with

a 13000-V1-24 Vortex adapter (Mo Bio, Inc., CA) and shaken at maximum speed for 15 min. Lysed fecal samples were homogenized using QIAshredder spin columns (Qiagen, Netherlands), and homogenized sample lysates were then extracted with an AllPrep DNA/RNA minikit (Qiagen, Netherlands) according to the manufacturer's specifications. Extracted RNA and DNA were purified using Agencourt RNAClean XP and Agencourt AMPure XP (Beckman Coulter, Inc., CA) kits, respectively. Total RNA samples were subjected to rRNA depletion, and metatranscriptomic libraries were prepared using a ScriptSeq Complete Gold kit (epidemiology)-low input (Illumina, CA, USA). Metagenomic libraries were prepared with a Nextera XT DNA library preparation kit (Illumina, CA, USA). All libraries were sequenced on a HiSeq 2500 platform (Illumina, CA, USA) at an expected depth of 6 Gb/library.

**Sequencing data preprocessing.** The metagenomes and metatranscriptomes were preprocessed by removing low-quality reads (mean quality value of less than 25), trimming low-quality positions (quality less than 15), and removing reads less than 90 nucleotides in length using FastqMcf (74). Further quality control steps involved the removal of human reads and the reads from the Illumina spike-in (bacteriophage Phi-X174) by mapping the reads against the corresponding genomes with Bowtie 2 (75). Metatranscriptomes were additionally processed to remove rRNA by mapping the reads against 16S and 23S rRNA gene databases (SILVA_119.1_SSURef_Nr99_tax_silva and SILVA_119_LSURef_tax_silva [76]) and to remove contaminant adapters using trim_galore (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) with the following parameters: -q 0, –nextera, and –stringency 5. The milk sample of mother-infant pair 4 at time point 1 was discarded from further analyses because of the low number of microbial reads (less than 400,000 bp) obtained after the quality control steps (see Table S1). All metagenomes and metatranscriptomes have been deposited in and are available at the NCBI Sequence Read Archive.

**Taxonomic and strain-level analysis.** Taxonomic profiling was performed with MetaPhlAn2 (77) (with default parameters) on the 23 metagenomic samples that passed the quality control. MetaPhlAn2 uses clade-specific markers for taxonomically profiling shotgun metagenomic data and to quantify the clades present in the microbiome with species-level resolution.

Strain-level profiling was performed with PanPhlAn (31) and a novel strain-level profiling method called StrainPhlAn (Truong et al., submitted). PanPhlAn is a pangenome-based approach that profiles the presence/absence pattern of species-specific genes in the metagenomes. The presence/absence profiles of the genes are then used to characterize the strain-specific gene repertoire of the members of the microbiome. PanPhlAn has been executed using the following parameters: –min_coverage 1, –left_max 1.70, and –right_min 0.30. PanPhlAn is available with supporting documentation at http://segatalab.cibio.unitn.it/tools/panphlan. StrainPhlAn is a complementary method based on analysis of SNVs that reconstructs the genomic sequence of species-specific markers. StrainPhlAn builds the strain-level phylogeny of microbial species by reconstructing the consensus marker sequences of the dominant strain for each detected species. The extracted consensus sequences are multiply aligned using MUSCLE version v3.8.1551 (78) (default parameters), and the phylogeny is reconstructed using RAxML version 8.1.15 (79) (parameters: -m GTRCAT and -p 1234). StrainPhlAn is available with supporting documentation at http://segatalab.cibio.unitn.it/tools/strainphlan.

**Functional profiling from metagenomes and metatranscriptomes.** The functional potential and transcriptomic analyses were performed with both HUMAnN2 (80) and PanPhlAn (31). HUMAnN2 selects the most representative species from a metagenome and then builds a custom database of pathways and genes that is used as a mapping reference for the coupled metatranscriptomic sample to quantify transcript abundances. We computed the normalized transcript abundance (NTA), which we define as the average coverage of a genomic region in the metatranscriptomic versus that in the corresponding metagenomic sample normalized by the total number of reads in each sample. PanPhlAn infers the expression of the strain-specific gene families by extracting them from the metagenome and matching them in the metatranscriptome. PanPhlAn has been executed using the following parameters: –rna_norm_percentile 90 and –rna_max_zeros 90.

**Profiling of DNA and RNA viruses.** We investigated the presence of viral and phage genomes by mapping the reads present in the metagenomes and metatranscriptomes against 7,194 viral genomes available in RefSeq (release 77). The average coverage and average sequencing depth were computed with SAMtools (81) and BEDTools (82).

The presence of the pepper mild mottle virus (PMMoV) was confirmed by mapping the reference genome (NC_003630) against the metatranscriptomic samples from the mother and infant of pair 4 and pair 5. In the mother and infant of pair 4, 424,510 and 119 reads were mapped, respectively, while in the mother and infant of pair 5, 1,444 and 61 of the reads were mapped, respectively. In the two mothers (pair 4 and pair 5), the values for breadth of coverage were 0.99 and 0.98 and for average coverage were 6,562 and 22, respectively. In the two infants (pair 4 and pair 5), the values for breadth of coverage were 0.6 and 0.5 and for average coverage were 1.81 and 0.95, respectively. Additionally, we extracted the shared fractions of the PMMoV genome present in both the mother and the infant of pair 4, together with the same regions of all the available reference genomes ($n = 13$ [specifically, accession no. LC082100.1, KJ631123.1, AB550911.1, AY859497.1, KU312319.1, KP345899.1, NC_003630.1, M81413.1, KR108207.1, KR108206.1, AB276030.1, AB254821.1, and LC082099.1]). The resulting sequences were aligned using MUSCLE version v3.8.1551 (default parameters), and the resulting alignment was used to build a phylogenetic tree with RAxML v. 8.1.15 (parameters: -m GTRCAT and -p 1234).

**Statistical analyses and data visualization.** The taxonomic and functional heat maps were generated using hclust2 (parameters: –f_dist_f Euclidean, –s_dist_f braycurtis, and –l) available at https://bitbucket.org/nsegata/hclust2. The multidimensional scaling plots were computed with the sklearn Python package (83).

Biomarker discovery (Fig. S7A) was performed by applying the linear discriminant analysis effect size (LEfSe) algorithm (84) (parameter: -l 3.0) on HUMAnN2 profiles. The two functional trees (Fig. S7A) have been automatically annotated with export2graphlan.py (GraPhlAn package) and displayed with GraPhlAn (85) using default parameters.

**Accession number(s).** All metagenomes and metatranscriptomes have been deposited and are available at the NCBI Sequence Read Archive under BioProject accession number PRJNA339914.

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/mSystems.00164-16.

**FIG S1,** EPS file, 0.5 MB.
**FIG S2,** EPS file, 1.8 MB.
**FIG S3,** EPS file, 0.1 MB.
**FIG S4,** EPS file, 0.2 MB.
**FIG S5,** EPS file, 0.1 MB.
**FIG S6,** EPS file, 0.1 MB.
**FIG S7,** PDF file, 2.5 MB.
**TABLE S1,** XLSX file, 0.01 MB.
**TABLE S2,** XLSX file, 0.1 MB.
**TABLE S3,** XLSX file, 0.01 MB.

## ACKNOWLEDGMENTS

## REFERENCES

1. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto JM, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Li S, Jian M, Zhou Y, Li Y, Zhang X, Li S, Qin N, Yang H, Wang J, Brunak S, Dore J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J; Meta HIT Consortium, Bork P, Ehrlich SD, Wang J. 2010. A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464:59–65. https://doi.org/10.1038/nature08821.

2. Clemente JC, Ursell LK, Parfrey LW, Knight R. 2012. The impact of the gut microbiota on human health: an integrative view. Cell 148:1258–1270. https://doi.org/10.1016/j.cell.2012.01.035.

3. Tamburini S, Shen N, Wu HC, Clemente JC. 2016. The microbiome in early life: implications for health outcomes. Nat Med 22:713–722. https://doi.org/10.1038/nm.4142.

4. Human Microbiome Project Consortium. 2012. Structure, function and diversity of the healthy human microbiome. Nature 486:207–214. https://doi.org/10.1038/nature11234.

5. Bäckhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI. 2005. Host-bacterial mutualism in the human intestine. Science 307:1915–1920. https://doi.org/10.1126/science.1104816.

6. Palm NW, de Zoete MR, Flavell RA. 2015. Immune-microbiota interactions in health and disease. Clin Immunol 159:122–127. https://doi.org/10.1016/j.clim.2015.05.014.

7. Stecher B, Hardt WD. 2011. Mechanisms controlling pathogen colonization of the gut. Curr Opin Microbiol 14:82–91. https://doi.org/10.1016/j.mib.2010.10.003.

8. Fuentes S, van Nood E, Tims S, Heikamp-de Jong I, ter Braak CJ, Keller JJ, Zoetendal EG, de Vos WM. 2014. Reset of a critically disturbed microbial ecosystem: faecal transplant in recurrent Clostridium diffi-cile infection. ISME J 8:1621–1633. https://doi.org/10.1038/ismej.2014.13.

9. Khoruts A, Dicksved J, Jansson JK, Sadowsky MJ. 2010. Changes in the composition of the human fecal microbiome after bacteriotherapy for recurrent Clostridium difficile-associated diarrhea. J Clin Gastroenterol 44:354–360. https://doi.org/10.1097/MCG.0b013e3181c87e02.

10. Britton RA, Young VB. 2014. Role of the intestinal microbiota in resistance to colonization by Clostridium difficile. Gastroenterology 146:1547–1553. https://doi.org/10.1053/j.gastro.2014.01.059.

11. Bäckhed F, Roswall J, Peng Y, Feng Q, Jia H, Kovatcheva-Datchary P, Li Y, Xia Y, Xie H, Zhong H, Khan MT, Zhang J, Li J, Xiao L, Al-Aama J, Zhang D, Lee YS, Kotowska D, Colding C, Tremaroli V, Yin Y, Bergman S, Xu X, Madsen L, Kristiansen K, Dahlgren J, Wang J, Jun W. 2015. Dynamics and stabilization of the human gut microbiome during the first year of life. Cell Host Microbe 17:690–703. https://doi.org/10.1016/j.chom.2015.04.004.

12. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP, Heath AC, Warner B, Reeder J, Kuczynski J, Caporaso JG, Lozupone CA, Lauber C, Clemente JC, Knights D, Knight R, Gordon JI. 2012. Human gut microbiome viewed across age and geography. Nature 486:222–227. https://doi.org/10.1038/nature11053.

13. Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO. 2007. Development of the human infant intestinal microbiota. PLoS Biol 5:e177. https://doi.org/10.1371/journal.pbio.0050177.

14. Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, Knight R. 2010. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proc Natl Acad Sci U S A 107:11971–11975. https://doi.org/10.1073/pnas.1002601107.

15. Azad MB, Konya T, Maughan H, Guttman DS, Field CJ, Chari RS, Sears MR, Becker AB, Scott JA, Kozyrskyj AL, CHILD Study Investigators. 2013. Gut

microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months. CMAJ 185:385–394. https://doi.org/10.1503/cmaj.121189.

16. Milani C, Mancabelli L, Lugli GA, Duranti S, Turroni F, Ferrario C, Mangifesta M, Viappiani A, Ferretti P, Gorfer V, Tett A, Segata N, van Sinderen D, Ventura M. 2015. Exploring vertical transmission of bifidobacteria from mother to child. Appl Environ Microbiol 81:7078–7087. https://doi.org/10.1128/AEM.02037-15.

17. La Rosa PS, Warner BB, Zhou Y, Weinstock GM, Sodergren E, Hall-Moore CM, Stevens HJ, Bennett WE, Jr., Shaikh N, Linneman LA, Hoffmann JA, Hamvas A, Deych E, Shands BA, Shannon WD, Tarr PI. 2014. Patterned progression of bacterial populations in the premature infant gut. Proc Natl Acad Sci U S A 111:12522–12527. https://doi.org/10.1073/pnas.1409497111.

18. Greenwood C, Morrow AL, Lagomarcino AJ, Altaye M, Taft DH, Yu Z, Newburg DS, Ward DV, Schibler KR. 2014. Early empiric antibiotic use in preterm infants is associated with lower bacterial diversity and higher relative abundance of Enterobacter. J Pediatr 165:23–29. https://doi.org/10.1016/j.jpeds.2014.01.010.

19. Aagaard K, Riehle K, Ma J, Segata N, Mistretta TA, Coarfa C, Raza S, Rosenbaum S, Van den Veyver I, Milosavljevic A, Gevers D, Huttenhower C, Petrosino J, Versalovic J. 2012. A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. PLoS One 7:e36466. https://doi.org/10.1371/journal.pone.0036466.

20. Hunt KM, Foster JA, Forney LJ, Schütte UM, Beck DL, Abdo Z, Fox LK, Williams JE, McGuire MK, McGuire MA. 2011. Characterization of the diversity and temporal stability of bacterial communities in human milk. PLoS One 6:e21313. https://doi.org/10.1371/journal.pone.0021313.

21. Cabrera-Rubio R, Collado MC, Laitinen K, Salminen S, Isolauri E, Mira A. 2012. The human milk microbiome changes over lactation and is shaped by maternal weight and mode of delivery. Am J Clin Nutr 96:544–551. https://doi.org/10.3945/ajcn.112.037382.

22. Flores GE, Caporaso JG, Henley JB, Rideout JR, Domogala D, Chase J, Leff JW, Vázquez-Baeza Y, Gonzalez A, Knight R, Dunn RR, Fierer N. 2014. Temporal variability is a personalized feature of the human microbiome. Genome Biol 15:531. https://doi.org/10.1186/s13059-014-0531-y.

23. Song SJ, Lauber C, Costello EK, Lozupone CA, Humphrey G, Berg-Lyons D, Caporaso JG, Knights D, Clemente JC, Nakielny S, Gordon JI, Fierer N, Knight R. 2013. Cohabiting family members share microbiota with one another and with their dogs. Elife 2:e00458. https://doi.org/10.7554/eLife.00458.

24. Shin H, Pei Z, Martinez KA, II, Rivera-Vinas JI, Mendez K, Cavallin H, Dominguez-Bello MG. 2015. The first microbial environment of infants born by C-section: the operating room microbes. Microbiome 3:59. https://doi.org/10.1186/s40168-015-0126-1.

25. Jost T, Lacroix C, Braegger CP, Rochat F, Chassard C. 2014. Vertical mother-neonate transfer of maternal gut bacteria via breastfeeding. Environ Microbiol 16:2891–2904. https://doi.org/10.1111/1462-2920.12238.

26. Biasucci G, Rubini M, Riboni S, Morelli L, Bessi E, Retetangos C. 2010. Mode of delivery affects the bacterial community in the newborn gut. Early Hum Dev 86(Suppl 1):13–15. https://doi.org/10.1016/j.earlhumdev.2010.01.004.

27. Dominguez-Bello MG, De Jesus-Laboy KM, Shen N, Cox LM, Amir A, Gonzalez A, Bokulich NA, Song SJ, Hoashi M, Rivera-Vinas JI, Mendez K, Knight R, Clemente JC. 2016. Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer. Nat Med 22:250–253. https://doi.org/10.1038/nm.4039.

28. Makino H, Kushiro A, Ishikawa E, Muylaert D, Kubota H, Sakai T, Oishi K, Martin R, Ben Amor K, Oozeer R, Knol J, Tanaka R. 2011. Transmission of intestinal Bifidobacterium longum subsp. longum strains from mother to infant, determined by multilocus sequencing typing and amplified fragment length polymorphism. Appl Environ Microbiol 77:6788–6793. https://doi.org/10.1128/AEM.05346-11.

29. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. 2012. Diversity, stability and resilience of the human gut microbiota. Nature 489:220–230. https://doi.org/10.1038/nature11550.

30. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP, Egholm M, Henrissat B, Heath AC, Knight R, Gordon JI. 2009. A core gut microbiome in obese and lean twins. Nature 457:480–484. https://doi.org/10.1038/nature07540.

31. Scholz M, Ward DV, Pasolli E, Tolio T, Zolfo M, Asnicar F, Truong DT, Tett A, Morrow AL, Segata N. 2016. Strain-level microbial epidemiology and population genomics from shotgun metagenomics. Nat Methods 13:435–438. https://doi.org/10.1038/nmeth.3802.

32. Schloissnig S, Arumugam M, Sunagawa S, Mitreva M, Tap J, Zhu A, Waller A, Mende DR, Kultima JR, Martin J, Kota K, Sunyaev SR, Weinstock GM, Bork P. 2013. Genomic variation landscape of the human gut microbiome. Nature 493:45–50. https://doi.org/10.1038/nature11711.

33. Turnbaugh PJ, Quince C, Faith JJ, McHardy AC, Yatsunenko T, Niazi F, Affourtit J, Egholm M, Henrissat B, Knight R, Gordon JI. 2010. Organismal, genetic, and transcriptional variation in the deeply sequenced gut microbiomes of identical twins. Proc Natl Acad Sci U S A 107:7503–7508. https://doi.org/10.1073/pnas.1002355107.

34. Maurice CF, Haiser HJ, Turnbaugh PJ. 2013. Xenobiotics shape the physiology and gene expression of the active human gut microbiome. Cell 152:39–50. https://doi.org/10.1016/j.cell.2012.10.052.

35. Gosalbes MJ, Abellan JJ, Durbán A, Pérez-Cobas AE, Latorre A, Moya A. 2012. Metagenomics of human microbiome: beyond 16S rDNA. Clin Microbiol Infect 18(Suppl 4):47–49. https://doi.org/10.1111/j.1469-0691.2012.03865.x.

36. Bao G, Wang M, Doak TG, Ye Y. 2015. Strand-specific community RNA-seq reveals prevalent and dynamic antisense transcription in human gut microbiota. Front Microbiol 6:896. https://doi.org/10.3389/fmicb.2015.00896.

37. Kurokawa K, Itoh T, Kuwahara T, Oshima K, Toh H, Toyoda A, Takami H, Morita H, Sharma VK, Srivastava TP, Taylor TD, Noguchi H, Mori H, Ogura Y, Ehrlich DS, Itoh K, Takagi T, Sakaki Y, Hayashi T, Hattori M. 2007. Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. DNA Res 14:169–181. https://doi.org/10.1093/dnares/dsm018.

38. Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, Angenent LT, Ley RE. 2011. Succession of microbial consortia in the developing infant gut microbiome. Proc Natl Acad Sci U S A 108(Suppl 1):4578–4585. https://doi.org/10.1073/pnas.1000081107.

39. Bickley J, Short JK, McDowell DG, Parkes HC. 1996. Polymerase chain reaction (PCR) detection of Listeria monocytogenes in diluted milk and reversal of PCR inhibition caused by calcium ions. Lett Appl Microbiol 22:153–158. https://doi.org/10.1111/j.1472-765X.1996.tb01131.x.

40. Cremonesi P, Castiglioni B, Malferrari G, Biunno I, Vimercati C, Moroni P, Morandi S, Luzzana M. 2006. Technical note: improved method for rapid DNA extraction of mastitis pathogens directly from milk. J Dairy Sci 89:163–169. https://doi.org/10.3168/jds.S0022-0302(06)72080-X.

41. Schrader C, Schielke A, Ellerbroek L, Johne R. 2012. PCR inhibitors—occurrence, properties and removal. J Appl Microbiol 113:1014–1026. https://doi.org/10.1111/j.1365-2672.2012.05384.x.

42. Ward TL, Hosid S, Ioshikhes I, Altosaar I. 2013. Human milk metagenome: a functional capacity analysis. BMC Microbiol 13:116. https://doi.org/10.1186/1471-2180-13-116.

43. Ramsay DT, Kent JC, Owens RA, Hartmann PE. 2004. Ultrasound imaging of milk ejection in the breast of lactating women. Pediatrics 113:361–367. https://doi.org/10.1542/peds.113.2.361.

44. Jeurink PV, van Bergenhenegouwen J, Jiménez E, Knippels LM, Fernández L, Garssen J, Knol J, Rodríguez JM, Martín R. 2013. Human milk: a source of more life than we imagine. Benef Microbes 4:17–30. https://doi.org/10.3920/BM2012.0040.

45. Costello EK, Stagaman K, Dethlefsen L, Bohannan BJ, Relman DA. 2012. The application of ecological theory toward an understanding of the human microbiome. Science 336:1255–1262. https://doi.org/10.1126/science.1224203.

46. Faith JJ, Guruge JL, Charbonneau M, Subramanian S, Seedorf H, Goodman AL, Clemente JC, Knight R, Heath AC, Leibel RL, Rosenbaum M, Gordon JI. 2013. The long-term stability of the human gut microbiota. Science 341:1237439. https://doi.org/10.1126/science.1237439.

47. Morowitz MJ, Denef VJ, Costello EK, Thomas BC, Poroyko V, Relman DA, Banfield JF. 2011. Strain-resolved community genomic analysis of gut microbial colonization in a premature infant. Proc Natl Acad Sci U S A 108:1128–1133. https://doi.org/10.1073/pnas.1010992108.

48. Sharon I, Morowitz MJ, Thomas BC, Costello EK, Relman DA, Banfield JF. 2013. Time series community genomics analysis reveals rapid shifts in bacterial species, strains, and phage during infant gut colonization. Genome Res 23:111–120. https://doi.org/10.1101/gr.142315.112.

49. Ogilvie LA, Jones BV. 2015. The human gut virome: a multifaceted majority. Front Microbiol 6:918. https://doi.org/10.3389/fmicb.2015.00918.

50. Dutilh BE, Cassman N, McNair K, Sanchez SE, Silva GG, Boling L, Barr JJ, Speth DR, Seguritan V, Aziz RK, Felts B, Dinsdale EA, Mokili JL, Edwards RA. 2014. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. Nat Commun 5:4498. https://doi.org/10.1038/ncomms5498.

51. Victoria JG, Kapoor A, Li L, Blinkova O, Slikas B, Wang C, Naeem A, Zaidi S, Delwart E. 2009. Metagenomic analyses of viruses in stool samples from children with acute flaccid paralysis. J Virol 83:4642–4651. https://doi.org/10.1128/JVI.02301-08.

52. Reyes A, Haynes M, Hanson N, Angly FE, Heath AC, Rohwer F, Gordon JI. 2010. Viruses in the faecal microbiota of monozygotic twins and their mothers. Nature 466:334–338. https://doi.org/10.1038/nature09199.

53. Zhang T, Breitbart M, Lee WH, Run JQ, Wei CL, Soh SW, Hibberd ML, Liu ET, Rohwer F, Ruan Y. 2006. RNA viral community in human feces: prevalence of plant pathogenic viruses. PLoS Biol 4:e3. https://doi.org/10.1371/journal.pbio.0040003.

54. de Medeiros RB, Figueiredo J, Resende RDO, De Avila AC. 2005. Expression of a viral polymerase-bound host factor turns human cell lines permissive to a plant- and insect-infecting virus. Proc Natl Acad Sci U S A 102:1175–1180. https://doi.org/10.1073/pnas.0406668102.

55. Balique F, Colson P, Barry AO, Nappez C, Ferretti A, Moussawi KA, Ngounga T, Lepidi H, Ghigo E, Mege JL, Lecoq H, Raoult D. 2013. Tobacco mosaic virus in the lungs of mice following intra-tracheal inoculation. PLoS One 8:e54993. https://doi.org/10.1371/journal.pone.0054993.

56. Colson P, Richet H, Desnues C, Balique F, Moal V, Grob JJ, Berbis P, Lecoq H, Harlé JR, Berland Y, Raoult D. 2010. Pepper mild mottle virus, a plant virus associated with specific immune responses, fever, abdominal pains, and pruritus in humans. PLoS One 5:e10041. https://doi.org/10.1371/journal.pone.0010041.

57. Reyes A, Semenkovich NP, Whiteson K, Rohwer F, Gordon JI. 2012. Going viral: next-generation sequencing applied to phage populations in the human gut. Nat Rev Microbiol 10:607–617. https://doi.org/10.1038/nrmicro2853.

58. Thurber RV, Haynes M, Breitbart M, Wegley L, Rohwer F. 2009. Laboratory procedures to generate viral metagenomes. Nat Protoc 4:470–483. https://doi.org/10.1038/nprot.2009.10.

59. Marcobal A, Barboza M, Sonnenburg ED, Pudlo N, Martens EC, Desai P, Lebrilla CB, Weimer BC, Mills DA, German JB, Sonnenburg JL. 2011. Bacteroides in the infant gut consume milk oligosaccharides via mucus-utilization pathways. Cell Host Microbe 10:507–514. https://doi.org/10.1016/j.chom.2011.10.007.

60. Tailford LE, Crost EH, Kavanaugh D, Juge N. 2015. Mucin glycan foraging in the human gut microbiome. Front Genet 6:81. https://doi.org/10.3389/fgene.2015.00081.

61. Turroni F, Milani C, van Sinderen D, Ventura M. 2011. Genetic strategies for mucin metabolism in Bifidobacterium bifidum PRL2010: an example of possible human-microbe co-evolution. Gut Microbes 2:183–189. https://doi.org/10.4161/gmic.2.3.16105.

62. LeBlanc JG, Milani C, de Giori GS, Sesma F, van Sinderen D, Ventura M. 2013. Bacteria as vitamin suppliers to their host: a gut microbiota perspective. Curr Opin Biotechnol 24:160–168. https://doi.org/10.1016/j.copbio.2012.08.005.

63. Flint HJ, Scott KP, Duncan SH, Louis P, Forano E. 2012. Microbial degradation of complex carbohydrates in the gut. Gut Microbes 3:289–306. https://doi.org/10.4161/gmic.19897.

64. Brown CT, Sharon I, Thomas BC, Castelle CJ, Morowitz MJ, Banfield JF. 2013. Genome resolved analysis of a premature infant gut microbial community reveals a Varibaculum cambriense genome and a shift towards fermentation-based metabolism during the third week of life. Microbiome 1:30. https://doi.org/10.1186/2049-2618-1-30.

65. Houghteling PD, Walker WA. 2015. Why is initial bacterial colonization of the intestine important to infants' and children's health? J Pediatr Gastroenterol Nutr 60:294–307. https://doi.org/10.1097/MPG.0000000000000597.

66. Turroni F, Peano C, Pass DA, Foroni E, Severgnini M, Claesson MJ, Kerr C, Hourihane J, Murray D, Fuligni F, Gueimonde M, Margolles A, De Bellis G, O'Toole PW, van Sinderen D, Marchesi JR, Ventura M. 2012. Diversity of bifidobacteria within the infant gut microbiota. PLoS One 7:e36957. https://doi.org/10.1371/journal.pone.0036957.

67. Turroni F, Foroni E, Serafini F, Viappiani A, Montanini B, Bottacini F, Ferrarini A, Bacchini PL, Rota C, Delledonne M, Ottonello S, van Sinderen D, Ventura M. 2011. Ability of Bifidobacterium breve to grow on different types of milk: exploring the metabolism of milk through genome analysis. Appl Environ Microbiol 77:7408–7417. https://doi.org/10.1128/AEM.05336-11.

68. Walker AW, Ince J, Duncan SH, Webster LM, Holtrop G, Ze XL, Brown D, Stares MD, Scott P, Bergerat A, Louis P, McIntosh F, Johnstone AM, Lobley GE, Parkhill J, Flint HJ. 2011. Dominant and diet-responsive groups of bacteria within the human colonic microbiota. ISME J 5:220–230. https://doi.org/10.1038/ismej.2010.118.

69. Scott KP, Antoine JM, Midtvedt T, van Hemert S. 2015. Manipulating the gut microbiota to maintain health and treat disease. Microb Ecol Health Dis 26:25877. https://doi.org/10.3402/mehd.v26.25877.

70. Gardy JL, Johnston JC, Ho Sui SJ, Cook VJ, Shah L, Brodkin E, Rempel S, Moore R, Zhao Y, Holt R, Varhol R, Birol I, Lem M, Sharma MK, Elwood K, Jones SJ, Brinkman FS, Brunham RC, Tang P. 2011. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. N Engl J Med 364:730–739. https://doi.org/10.1056/NEJMoa1003176.

71. Loman NJ, Constantinidou C, Christner M, Rohde H, Chan JZ, Quick J, Weir JC, Quince C, Smith GP, Betley JR, Aepfelbacher M, Pallen MJ. 2013. A culture-independent sequence-based metagenomics approach to the investigation of an outbreak of Shiga-toxigenic Escherichia coli O104:H4. JAMA 309:1502–1510. https://doi.org/10.1001/jama.2013.3231.

72. Li SS, Zhu A, Benes V, Costea PI, Hercog R, Hildebrand F, Huerta-Cepas J, Nieuwdorp M, Salojärvi J, Voigt AY, Zeller G, Sunagawa S, de Vos WM, Bork P. 2016. Durable coexistence of donor and recipient strains after fecal microbiota transplantation. Science 352:586–589. https://doi.org/10.1126/science.aad8852.

73. Giannoukos G, Ciulla DM, Huang K, Haas BJ, Izard J, Levin JZ, Livny J, Earl AM, Gevers D, Ward DV, Nusbaum C, Birren BW, Gnirke A. 2012. Efficient and robust RNA-seq process for cultured bacteria and complex community transcriptomes. Genome Biol 13:R23. https://doi.org/10.1186/gb-2012-13-3-r23.

74. Aronesty E. 2013. Comparison of sequencing utility programs. Open Bioinforma J 7:1–8. https://doi.org/10.2174/1875036201307010001.

75. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.

76. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and Web-based tools. Nucleic Acids Res 41:D590–D596. https://doi.org/10.1093/nar/gks1219.

77. Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, Tett A, Huttenhower C, Segata N. 2015. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat Methods 12:902–903. https://doi.org/10.1038/nmeth.3589.

78. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340.

79. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033.

80. Franzosa EA, Morgan XC, Segata N, Waldron L, Reyes J, Earl AM, Giannoukos G, Boylan MR, Ciulla D, Gevers D, Izard J, Garrett WS, Chan AT, Huttenhower C. 2014. Relating the metatranscriptome and metagenome of the human gut. Proc Natl Acad Sci U S A 111:E2329–E2338. https://doi.org/10.1073/pnas.1319284111.

81. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

82. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842. https://doi.org/10.1093/bioinformatics/btq033.

83. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E. 2011. Scikit-learn: machine learning in Python. J Mach Learn Res 12:2825–2830.

84. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. 2011. Metagenomic biomarker discovery and explanation. Genome Biol 12:R60. https://doi.org/10.1186/gb-2011-12-6-r60.

85. Asnicar F, Weingart G, Tickle TL, Huttenhower C, Segata N. 2015. Compact graphical representation of phylogenetic data and metadata with GraPhlAn. PeerJ 3:e1029. https://doi.org/10.7717/peerj.1029.