Factors associated with the diversification of the gut microbial communities within chimpanzees from Gombe National Park

Patrick H. Degnan, Anne E. Pusey, Elizabeth V. Lonsdorf, Jane Goodall, Emily E. Wroblewski, Michael L. Wilson, Rebecca S. Rudicell, Beatrice H. Hahn, and Howard Ochman

Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520; Department of Evolutionary Anthropology, Duke University, Durham, NC 27708; Department of Psychology, Franklin and Marshall College, Lancaster, PA 17604; GoMbe Stream Research Centre, The Jane Goodall Institute, Kigoma, Tanzania; Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94305; Department of Anthropology, University of Minnesota, Minneapolis, MN 55455; Department of Ecology, Evolution and Behavior, University of Minnesota, St. Paul, MN 55108; Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892; and Departments of Medicine and Microbiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104

Edited by Edward F. DeLong, Massachusetts Institute of Technology, Cambridge, MA, and approved June 19, 2012 (received for review July 12, 2011)

The gastrointestinal tract harbors large and diverse populations of bacteria that vary among individuals and within individuals over time. Numerous internal and external factors can influence the contents of these microbial communities, including diet, geography, physiology, and the extent of contact among hosts. To investigate the contributions of such factors to the variation and changes in gut microbial communities, we analyzed the distal gut microbiota of individual chimpanzees from two communities in Gombe National Park, Tanzania. These samples, which were derived from 35 chimpanzees, many of whom have been monitored for multiple years, provide an unusually comprehensive resource for understanding the interplay of multiple factors that contribute to the variation in the microbiota within and among hosts. Such questions might best be resolved through the long-term analysis of populations of hosts of known genealogical relationships and demographic histories. In this regard, the samples collected from chimpanzees in Gombe National Park provide an extraordinary resource for understanding the interplay of multiple factors on the variation of gut microbial communities within a species. These chimpanzee hosts are represented by feral specimens sampled for more than a decade from individuals of known identity, genealogy, provenance, health, and community status. These samples combine the host diversity and temporal depth to allow us address questions about the dynamics and persistence of gut microbes, many of which have never been addressed in any species. And because chimpanzees are our closest relatives, the dynamics of their gut microbial communities can serve as models for understanding the evolution of our own microbiota.

Results

We examined the gut microbial communities of chimpanzees (Pan troglodytes schweinfurthii) of known identity and kinship from Gombe National Park in Tanzania (Fig. 1 and SI Appendix, Table S1) (20–22). The ability to assign samples to specific hosts


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

1To whom correspondence should be addressed. E-mail: howard.ochman@yale.edu.

This article contains supporting information online at www.pnas.orglookup/suppl/doi:10.1073/pnas.1110994109/DCSupplemental.

The gastrointestinal tract harbors large and diverse populations of bacteria that vary among individuals and within individuals over time. Numerous internal and external factors can influence the contents of these microbial communities, including diet, geography, physiology, and the extent of contact among hosts. To investigate the contributions of such factors to the variation and changes in gut microbial communities, we analyzed the distal gut microbiota of individual chimpanzees from two communities in Gombe National Park, Tanzania. These samples, which were derived from 35 chimpanzees, many of whom have been monitored for multiple years, provide an unusually comprehensive resource for understanding the interplay of multiple factors that contribute to the variation in the microbiota within and among hosts. Such questions might best be resolved through the long-term analysis of populations of hosts of known genealogical relationships and demographic histories. In this regard, the samples collected from chimpanzees in Gombe National Park provide an extraordinary resource for understanding the interplay of multiple factors on the variation of gut microbial communities within a species. These chimpanzee hosts are represented by feral specimens sampled for more than a decade from individuals of known identity, genealogy, provenance, health, and community status (20–22). These samples combine the host diversity and temporal depth to allow us address questions about the dynamics and persistence of gut microbes, many of which have never been addressed in any species. And because chimpanzees are our closest relatives, the dynamics of their gut microbial communities can serve as models for understanding the evolution of our own microbiota.

Results

We examined the gut microbial communities of chimpanzees (Pan troglodytes schweinfurthii) of known identity and kinship from Gombe National Park in Tanzania (Fig. 1 and SI Appendix, Table S1) (20–22). The ability to assign samples to specific hosts


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

1To whom correspondence should be addressed. E-mail: howard.ochman@yale.edu.

This article contains supporting information online at www.pnas.orglookup/suppl/doi:10.1073/pnas.1110994109/DCSupplemental.
allowed us both to distinguish among factors that contribute variation among individuals and to investigate the long-term stability of gut microbiota within individuals.

**Placement of Gombe Chimpanzee Microbiota Among Great Ape Species.** Our previous analysis of distal gut microbiota within great apes established that the branching order of the phylogeny based on the composition of gut microbial communities mirrored the established relationships of the great ape hosts (10). To ascertain the distribution of samples from Gombe chimpanzees within this great ape phylogeny, we integrated the phylotype information for the Gombe chimpanzees along with that recovered previously for 24 individuals of five species of great apes (SI Appendix, Table S2). Using both phylogenetic and sample ordination approaches, we find that the topology and relatedness among samples generally recapitulates what is known about the species-level relationships among hosts (Fig. 2). (i) The samples from the 34 Gombe chimpanzees form a single clade that includes the Gombe sample previously examined but excludes all other great apes. (ii) The gut microbial communities within Gombe chimpanzees are most similar to those of other chimpanzees. (iii) Although Gombe chimpanzees are classified as subspecies *P. t. schweinfurthii*, their microbial communities do not cluster with those of *P. t. schweinfurthii* from the Democratic Republic of the Congo (DRC). This incongruity has been noted (10) and could be caused by the relatively low sampling of the microbial communities available for the DRC apes. (The mean number of pyrotags for DRC *P. t. schweinfurthii* samples was 4,217, whereas for every clade of *Pan* spp., there were at least 12,000.) This analysis reinforces the presence of a strong signal of host phylogeny on the composition of the great ape gut microbiota. Therefore, our analyses of factors influencing the microbial communities in Gombe chimpanzees examine the variation that

---

**Fig. 1.** Pedigree of sampled chimpanzees from Gombe National Park. Genealogies of chimpanzees (circles, females; squares, males) from the Kasekela (KK, blue) and Mitumba (MT, red) communities whose gut microbiota were analyzed. Individual chimpanzees are categorized by their current community affiliation. Individuals were sampled in 2008 (green), 2001 (orange), and/or 2000 (yellow). Those left uncolored are individuals whose gut microbiota were not surveyed. For all sampled individuals, ages in 2008 are provided next to each name. Question marks indicate fathers of unknown identity, and asterisks denote individuals shown more than once.

**Fig. 2.** Relationship of gut microbial communities in Gombe chimpanzees to those of other great apes. (A) Neighbor-joining phylogenetic reconstruction of log abundances of microbial phylotypes (99% OTUs) present in Gombe chimpanzees and other great apes. Phylogenies and the bootstrap support for branching orders in these phylogenies were generated using default parsimony and neighbor-joining parameters in PAUP 4b10 (30) as in ref. 10. Colored triangles delineate species or subspecies boundaries, and bootstrap values ≥0.75 are marked with asterisks. (B) Nonmetric multidimensional scaling ordination of Sørensen distances of the microbial phylotypes detected in great ape guts. Colors match the species and subspecies designations in A and in ref. 10 with points representing previously analyzed samples labeled. The cluster of points corresponding to the gut microbial communities of Gombe chimpanzees is outlined with a dashed line.
occurs at the very tips of the tree: As such, they must be interpreted in the context of a monophyletic group of samples of recent origin and of high similarity relative to that observed in the species at large.

**Microbial Community Profiles.** We generated a total of 1,206,387 pyrosequencing reads for 48 samples: For each host (Fig. 1 and SI Appendix, Table S1), we examined the microbial diversity within a sample collected in October to December 2008; and for seven of the hosts, we examined additional samples collected during the same period in 2000 and/or in 2001. More than one-half of the sequencing reads (687,499) passed the initial quality and length filters and, of these, 31,837 were either chimeric, or of chloroplast or eukaryotic origins, and removed. This filtering yielded an average of 13,113 reads per sample (8,323–22,563), which were clustered into phylotypes applying 0.01% divergence cutoff [99% operational taxonomic units (OTUs)]. After removal of OTUs present in only one sample, the 5,788 phylogenetically informative phylotypes were assigned to taxonomic groups, and a total of 14 phyla were represented (SI Appendix, Fig. S1). The majority of phylotypes were classified as either Firmicutes (59.8%) or Bacteroidetes (17.9%), as observed in other mammals (7, 9, 10), and representatives of six other phyla were also present in all samples, but at substantially lower frequencies. Approximately 13% could not be assigned to a particular phylum; however, the majority of these unassigned phylotypes (476/784) have ≥95% BLAST similarity to 16S rRNA sequences retrieved from human gut samples.

**Factors Influencing Microbial Communities.** Estimates of diversity within (alpha) and between (beta) gut microbial communities were assessed for 34 chimpanzees that were each sampled in October to December 2008 (SI Appendix, Table S3). Individual samples contained 518–1,389 phylotypes, with high overall alpha diversity (Shannon’s H’, x = 5.09 ± 0.30) and sampling evenness E, x = 0.76 ± 0.031. There are no significant differences in Shannon diversity indices, sample evenness, or Chao1 species richness estimators with respect to sex, geographic community affiliation, or matriline. However, in contrast to what has been reported in humans (5), infant (0–5 y) exhibited significantly greater within sample diversity (H’) than did adults (16–33 y) (nonparametric Wilcoxon test; P = 0.044 (infant-adult)). In addition, adolescents have significantly greater within sample diversity (P = 0.032) and evenness (P = 0.011); infants harbor significantly greater phylotype diversity than adults (P = 0.031) (SI Appendix, Fig. S2). These statistical differences persist after applying a resampling procedure that accounts for variation in sequence-read numbers among samples.

We performed an ordination analysis by using nonmetric multidimensional scaling (NMS) with a Sørensen distance matrix generated for the 34 chimpanzees to assess beta diversity among samples. The two major axes, which together represent the majority of the observed variation (cumulative R² = 0.765), separate samples into two broad groups that roughly correspond to geographic community affiliation (Fig. 3). Results of multiresperror permutation procedures (MRPP) corroborated the NMS analyses: There is a significant chance-corrected within-group agreement (A) values when the sample is classified by community (A = 0.182; P = 0.0003), but not by any of the other parameters tested (e.g., sex, P = 0.38; age class, P = 0.10). A total of 108 phylotypes were significantly associated with separating hosts according to community affiliation (P < 0.05, indicator value ≥50%), and it is noteworthy that several of the borderline cases represent samples from hosts that switched community affiliation (Fig. 3).

The distributions of several beta diversity measures were compared within and between (i) geographic communities, (ii) hosts of different sexes, and (iii) individuals with different degrees of relatedness. Sørensen distances corroborated the observed distinction between the gut microbiota of hosts from the Kasekela (KK) and Mitumba (MT) chimpanzee communities. Individuals are labeled according to the key, with year and direction of immigration included when applicable.

**Temporal Changes in Microbial Communities as Revealed by Pyrotags.** To examine the long-term stability of gut microbial communities, we first used the pyrotag approach to interrogate the microbial communities in first samples from 2000, 2001, and 2008 for multiple chimpanzee hosts. Among the chimpanzees sampled longitudinally, between 170 and 350 of the phylotypes detected in the initial year (2000) were present in both of the later samples (2001 and 2008) from an individual host. The conserved phylotypes were usually the most abundant microbes in a host’s gut microbial communities (constituting 40–70% of the total sequence reads). Within each of the longitudinally sampled hosts, there were a few (5–30) unique phylotypes maintained at low frequencies over the entire 8-y sampling period, suggesting that certain host-specific strains persist rather than originate from repeated reinfection (Fig. 5 and SI Appendix, Fig. S3). The taxonomic distribution of phylotypes was virtually identical in the samples from each year, but there is long-term turnover in the microbial community constituent at the level of bacterial phylotype. Samples collected in 2000 and 2001 were, for most chimpanzees, two to four times more similar to one another than either was to the 2008 sample. The greater similarity of the samples from 2000 and 2001 is due to the maintenance of particular phylotypes, which are subsequently replaced by new phylotypes from the same taxonomic division.
The nature of the temporal changes in phylotype diversity becomes apparent when the gut microbial communities were subjected to ordination methods. Considering the two major axes (cumulative $R^2 = 0.775$), collection year is the main factor driving the stratification of samples ($A = 0.0203$, $P = 0.0038$), with the 2008 samples from all seven hosts converging into a single cluster (Fig. 5). The similarity among samples from 2008 is due primarily to the increased abundance of a group of phylotypes classified as Prevotellaceae, Erysipelotrichaceae, and Coriobacteriaceae, which are taxonomically distinct from the indicator phylotypes from previous years.

**Microbial Community Continuity as Revealed by iTags.** To examine long-term changes in the gut microbiota at a finer level of resolution, we implemented an Illumina-based 16S-rDNA profiling on multiple samples collected from individual hosts during the interval spanning 2000–2008 (SI Appendix, Table S4). After quality filtering and the removal of reads unique to a single sample, the 39 samples each averaged nearly 50,000 iTag reads, which were clustered into 99% OTUs.

To determine whether the 99% OTUs based on iTags, which assayed a different region of the 16S rDNA gene, recapitulate the patterns observed for pyrotags, we applied analogous ordination methods to the iTag dataset. Again, we observed that the samples collected in 2008 tended to cluster together and apart from those from the same chimpanzees sampled in 2000 and 2001, even when we include samples from intervening and later years (SI Appendix, Fig. S4).

It is possible that long-term sample storage could either increase sequence diversity (because of DNA damage) or alternatively reduce diversity (through cellular degradation and decomposition), but there were no clear trends among samples. To determine how variation in bacterial OTUs within an individual changed through time, and whether there were systematic increases or decreases in bacterial diversity that might be attributable to sample storage, we examine the continuity of these 99% OTUs across multiple time-points for each chimpanzee host. On average, 40% of the OTUs identified from an individual are shared across every sample collected from the 9-y period, with the proportion of the shared OTUs representing as much as 94% to as few as 37% of the reads from a single sample (SI Appendix, Fig. S5). Furthermore, the amount of diversity was not associated with the age of a sample, and large proportions of very low frequency OTUs were maintained within hosts for close to a decade (SI Appendix, Fig. S6). For example, more than 25% of the OTUs that were present at frequencies of <0.1% (corresponding to an average of only 22 reads per samples) were detected within hosts over the entire sampling period. Although the number of OTUs (SI Appendix, Table S4) and the phylogenetic distribution of OTUs remained remarkably consistent over the sampling period, we examined the patterns of change in the frequencies of each OTU detected in all longitudinal samples from a given host. As expected, if the long-term storage of samples has no affect on the recovery and recognition of OTUs, there was no consistent change in the frequency of persistent OTUs among hosts: In four hosts, a majority of the persistent OTUs decreased through time, in two hosts, a majority increased, and in one host, there was no appreciable difference between the number of OTUs that increased or decreased in frequency.

**Discussion**

We examined a large and diverse set of chimpanzees from Gombe National Park in Tanzania and found that their gut microbial communities represent a discrete, phylogenetic clade distinct from those present in other chimpanzees and great apes (Fig. 2). Previous analysis of the gut microbiota of several species...
of great apes (including humans) revealed that the branching order of phylogenies based on microbial community compositions paralleled the evolutionary relationships of great ape species (10). The phylogenetic concordance between gut communities and host species, which resembles a pattern of vertical transmission, is surprising given that gut microbes are acquired from external sources each generation. This observation suggests that genetically determined factors (such as the host immune system), which accumulate differences over evolutionary timescales, have played a major role in the diversification of the great ape gut microbial communities.

Despite the concordance of the species-level phylogenies, which indicate that each great ape species has signature gut microbiota that is most similar to that in its most closely related species, the majority of gut community variation occurs among individuals within each species. Such within-species variation has been observed in comparisons of distal gut microbial communities of human hosts (7, 12, 13, 18) and is evident from the long terminal branches in the microbial community phylogenies (Fig. 2). Because numerous factors can potentially produce the large within-species diversity, we assayed samples collected over the past decade from Gombe chimpanzees to determine the contribution of genetic, social, geographic, and temporal variables on gut microbial community composition and continuity.

Given the strong phylogenetic signal at the core of the relationships among the gut microbiota of great apes, one might expect that the degree of genetic relatedness between individuals would dictate the level of similarity in the compositions of their gut microbial communities. However, we detected no association between coefficients of relatedness between individuals and the overall similarity in their gut microbial communities (Fig. 4). This situation in Gombe chimpanzees differs from previous findings on humans in which the gut microbial communities of twins and of parent-offspring pairs were more similar than were those of unrelated individuals (6, 18). Within the Gombe chimpanzees, there are significant sex-specific differences in the beta diversity of their gut microbial communities. Such dissimilarities have been observed (8) but are not common, which suggests that they stem not from physiological differences but from differences in the feeding behaviors of male and female chimpanzees. Male chimpanzees tend to seek food over the entire range of their community, whereas females typically feed in small, confined areas (23, 24), a pattern that might contribute to disparity between the sexes. However, the difference between males and females is observed in infants/juveniles (who feed with their mothers) as well as in adults, suggesting that other sex-specific traits contribute to differences in the gut microbiota.

Several aspects of the variation in gut microbial communities of the Gombe chimpanzees support the view that the differentiation (or convergence) in their gut microbial communities is shaped by inputs from environmental sources (such as diet, provenance, and/or social contact). For example, gut microbial communities differentiate according to the geographic origin and community affiliation of the host (Fig. 3), with most of the borderline cases being females who switched communities or immigrated from outside territories. The differentiation of individuals from the two communities is caused by approximately 100 phylotypes, which were either unique to members of a particular community or exhibited up to a 100-fold difference in mean abundance between communities. Despite the similarities in the gut microbiota in the members of a chimpanzee community, which could be due to a common diet and/or increased social interactions, higher-order factors are also contributing to the contents and composition of the gut microbiota. By examining the fecal microbiota of individual chimpanzees sampled over 8 y, we detected both host-specific and large-scale components to the differentiation of gut microbial communities.

The longitudinally sampled chimpanzees harbored broadly overlapping sets of core phylotypes that were maintained at relatively high frequencies throughout the entire sampling period. However, samples taken 1 y apart were, on average, twice as similar as those separated by 7 or 8 y, reflecting the perpetuation of resident phylotypes and gradual turnover of transient members. This investigation represents the longest temporal sampling of an individual’s gut microbiota, although such long-term studies of the microbial communities on and within humans are beginning to accumulate. Another level of temporal structuring to these microbial communities was observed: The microbiota from all seven longitudinally sampled chimps converge in 2008, even those derived from chimpanzees from different communities. Moreover, after implementing a method that returned a higher resolution of 16S rDNA phylotypes, we found no evidence indicating either that these temporal patterns of variation among hosts or that the changes in microbial communities over time stem from the use of samples that have been stored frozen for more than a decade. The most likely explanation for the convergence in the gut microbiota of chimpanzees sampled 2008 comes from field observations that members from the Kasanka community were frequently ranging into the Mitumba community territory in 2007 and 2008, which would increase similarity in the dietary sources of microbes among members of the two communities.

The identification of core groups of resident taxa in the Gombe chimpanzee is reminiscent of the recent description of human enterotypes, which correspond to groups of individuals clustered in characteristic gut microbiota profiles (25). We find that multiple factors—including geographic, temporal, sex-, and age-specific—are associated with the long-term composition and diversity of the gut microbial communities harbored by Gombe chimpanzees, but that despite the variation over space and time, their gut communities remain distinct from those of other great apes, including other subspecies of chimpanzees.

Materials and Methods

The microbial community diversity present in 47 fecal specimens from 35 SIVcpz negative chimpanzees (P.t. schweinfurthii) at the Gombe National Park (Tanzania) was assayed by 16S rDNA pyrotag and iTag sequencing procedures (10, 26, 27). The individual source of all samples was determined by genetic markers and could be unequivocally assigned to hosts, such that the sex, age at time of sampling, community affiliation, and genealogical relationships are known. For each chimpanzee, we analyzed a sample collected between October and December 2008 with few exceptions (Tubi_2008, Darbee_2008). For seven chimpanzee hosts, we selected additional samples from several additional years collected during the same 3-mo interval. Total DNA was extracted and used as a template to amplify fragments corresponding to the V6–V9 region of 16S ribosomal RNA by using the universal primers 926F and 1492R, which were barcoded for multiplexing two samples into a single pyrosequencing run. Raw pyrosequencing reads were quality-trimmed as in ref. 10 by using an error rate of 0.5% (equivalent to Phred quality score of 23), and reads <260 nucleotides in length, or that contained undefined bases or lacked perfect barcode and primer sequences, were removed. Taxonomic assignments of the quality-trimmed reads (also known as “pyrotags”) were performed with the RDP classifier (28). Trimmed, taxonomically assigned sequences were dereplicated, chimera-checked, aligned, and clustered at a range of similarity values (90–99%). Identical sequences were performed on 454 flowgrams that were denoised in QIIME (29). Pyrotag analyses of microbial communities were based on occurrences and abundances of 99% OTUs (i.e., clusters of pyrotag sequences that differ by 1% or less) that were phylogenetically informative (i.e., present in two or more samples). To examine changes in gut microbial communities at higher levels of resolution, we implemented an Illumina-based 16S-tag (iTag) approach that was analogous to the pyrotag procedure but differs by interrogating a 100-nt region from the 5′-end of amplicons spanning the V4 region of 16S rDNA (27). iTags were generated for a total of 39 samples, which represent annual samples spanning 10 consecutive years for the seven chimpanzees that were surveyed at multiple time-points by the pyrotag procedure.
ACKNOWLEDGMENTS. PCR and sequencing primers for iTag analysis were generously provided by Rob Knight (University of Colorado) and Integrated DNA Technologies. We thank the Jane Goodall Institute and the National Science Foundation through Grant BSC-0648481 (to M.L.W.) for supporting collection of behavioral data and fecal samples from chimpanzees at the Gombe Stream Research Centre, and the Tanzania Commission for Science and Technology, the Tanzania Wildlife Research Institute, and the Tanzania National Parks for permission to conduct research in Gombe. This work was supported in part by National Institutes of Health Grants R01 AI50529, R01 AI58715 (to B.H.H.), and R01 GM74738 and R01 GM101209 (to H.O.); University of Alabama at Birmingham Center for AIDS Research Grant P30 AI 27767, and the National Science Foundation (A.E.P.). R.S.R. was funded by a Howard Hughes Medical Institute Med-into-Grad Fellowship.