

Supplemental Information.

Material and Methods

Source of samples. The microbial community diversity present in 47 fecal specimens from 35 SIVcpz negative chimpanzees (*Pan troglodytes schweinfurthii*) at the Gombe National Park (Tanzania) was assayed by both 16S rDNA pyrotag and *iTag* sequencing procedures (1–4). After collection, fecal specimens were mixed into an equal volume of RNAlater (Qiagen) and transported to the laboratory, where they were stored at -80°C (5). All samples were typed with 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 (6). With few exceptions (Tubi_2008, Darbee_2008), we analyzed a sample collected between October and December 2008 from each chimpanzee. Note that some of the selected hosts reside in a different geographic community than their parents, grandparents, and/or siblings, and that some have many different levels of kinship to other hosts whose microbiotae were assayed (Fig. 1). For seven chimpanzee hosts, we selected samples from additional years collected during the same three-month interval. For one sample (Gremlin_2000, GM7), we performed multiple technical replicates to evaluate the reproducibility of our experimental methods (Table S1).

Sample preparation and pyrotag sequencing. Total DNA was extracted from fecal samples using the QIAmp DNA stool kit according the manufacturer's instructions (Qiagen). DNA was quantified with the PicoGreen DNA assay (Invitrogen) on a Victor X3 plate reader (Perkin Elmer). Total fecal DNA was used as template to amplify fragments corresponding to the V6–V9 region of 16S ribosomal RNA using the universal primers 926F and 1492R (sequences in lower case below). To multiplex samples onto a single sequencing run, 50 forward primers

were designed as fusions of the 454 FLX Titanium Lib-L 'A' sequencing primer, a 2-nucleotide spacer, an 8- or 10-nucleotide barcode (designated by **N**'s) and a conserved 16S rDNA sequence (primer **TAXX-926F**: 5'-CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG-**NNNNNNNN**-CT-aaa ctY aaa Kga att gac gg-3'). All reactions used the same reverse primer whose sequence included a conserved 16S sequence separated by a two nucleotide spacer to the 454 FLX Titanium Lib-L 'B' sequencing primer (primer **TB-1492R**: 5'-CCT ATC CCC TGT GTG CCT TGG CAG TCT CAG-TC-tac ggY tac ctt gtt acg act t-3').

Each DNA sample was amplified in triplicate 30 µl reactions using 0.4 U of Phusion DNA polymerase (NEB), 1X final concentration of HF buffer, 0.5 µM each primer, 200 µM dNTPs and 3–10 ng of template DNA. Reactions were denatured for 1 minute at 98°C then cycled 20 times at 98°C for 10 seconds, 55°C for 10 seconds, and 72°C for 15 seconds, followed by a final elongation step at 72°C for 5 minutes. Individual reactions from each sample were pooled, verified on 1.0% agarose gels, purified with the MinElute PCR purification kit (Qiagen), and quantified with the PicoGreen DNA assay. Equimolar amounts of amplicons were mixed and submitted for pyrosequencing using primer 'A' of the Lib-L kit. Technical replicates were performed on the DNA isolated from host GM7 using primers having the same or different barcode sequences to assess sample reproducibility (Table S1).

Pyrotag processing. Raw pyrosequencing reads were first quality trimmed with **Lucy** using an error rate of 0.5% (equivalent to **Phred** quality score of 23) set for the maximum average error and errors at the ends. Following this initial quality trimming, any reads less than 260 nucleotides in length, or that still contained undefined bases (**N**'s) or lacked perfect barcode and primer sequences, were removed. Remaining reads were trimmed to 260 nucleotides from

the 3' end of the forward primer. Taxonomic assignments of the quality-trimmed reads (aka "pyrotags") were performed with the **RDP classifier** (7), and any sequence read matching plant chloroplast sequences was removed.

The trimmed, taxonomically assigned sequences were de-replicated to yield the set of unique pyrotags (100% OTUs), which were then queried against the **Silva** database using **BlastN** to identify eukaryotic and chimeric sequences. To be deemed as eukaryotic in origin, sequences were required to match have $\geq 80\%$ nucleotide identity and an E-value $< 10^{-40}$ to an eukaryotic ssuRNA, and pyrotags with $\geq 90\%$ identity to a restricted section ($< 90\%$) of a known sequence were considered chimeras. Chimeric sequences were also identified with **UCHIME** using the *de novo* detection algorithm with default parameters (8). Read filtering removed 17,424 high-quality, eukaryotic pyrotags, the majority derived from unicellular eukaryotes (16,399/17,424) and dominated by Trichomonads and *Blastocystis* spp. Although these pyrotags represent non-specific PCR amplification and relative abundances are unreliable, 33 of the 35 nominally healthy chimpanzees were positive for one or both of these eukaryotic parasites. Previous surveys have directly enumerated the persistence and prevalence of gastrointestinal parasites within these chimpanzee communities (9–11). Sample TA39 harbored a highly atypical community, with $> 60\%$ of the reads assigned to the genus *Tetragenococcus*, which was otherwise very rare or lacking (constituting only 47/645,769 reads) in the remaining samples. Therefore, this sample was removed and all analyses were performed on the microbial communities of 34 chimpanzees.

All of the remaining unique tags were aligned with the **RDPaligner** and clustered at a range of similarity values using the **RDP** website (90 – 99%). Using custom Perl scripts, unique

reads were inflated to add back all constituent tags in order to reflect actual read abundances and clusters at each similarity cutoff were partitioned by sample. After this processing step, clusters represented by tags derived from a single chimpanzee fecal sample were removed, as they are not phylogenetically informative clusters and may represent amplification artifacts or contamination. Using the **RDP** web resources, Shannon diversity, Evenness, Sørensen distances and rarefaction curves were estimated. Additional diversity measures (e.g., Morsita-Horn) were calculated with **EstimateS** (12). In addition, raw 454 flowgrams were treated as above and then subjected to denoising procedures, as implemented in QIIME (13) using default parameters for 454-Titanium reads (`denoise_wrapper.py`), in order to remove any spurious phylotypes that might be generated during the clustering procedures.

Statistical analyses. Based on the occurrences and abundances of 99% OTUs (i.e., clusters of pyrotag sequences that differ by 1% or less) that were present in two or more samples, we generated abundance tables (formatted for **PC-ORD** v4.25; 14) relating the abundance of each OTU present in the samples collected in 2008 from 35 hosts and for the longitudinal samples (from 2000, 2001, and 2008) available for seven hosts. Outlying samples (TA39) and technical replicates were not included (TA05, TA52). Separate sample tables describing the sex, community affiliation, age, and sampling date of individuals were also produced. Initial nonmetric multidimensional scaling (NMS) analyses were performed using the “autopilot” mode to determine an appropriate number of axes. Autopilot first calculates between sample diversity (Sørensen distances) and then performs multiple ordinations using 1–4 axes determining the “stress” index of real and randomized data for each dimensionality. The optimal number of axes is defined as the number of dimensions (n) that reduce the final stress

by ≥ 5 relative to $n-1$ axes and has a final stress $< 95\%$ of the randomized data. Using a random starting seed, both data sets were re-run with the optimal number of axes, a stability criterion of 0.00001 (standard deviation in stress for preceding 10 iterations), and p -values were determined through Monte Carlo simulation.

Using **PC-ORD**, we performed multi-response permutation procedures (MRPP) to further characterize patterns observed in the NMS results. Sørensen distances were used to calculate chance-corrected within-group agreement (A) of samples grouped by individual characteristics, such as sex, community affiliation, age, and sampling date. To identify those 99% OTUs significantly associated with a particular sample characteristic, we applied an Indicator species analysis (15), with the significance of observed maximum indicator values (IV) tested by Monte Carlo simulations. These statistical procedures were also performed on randomly resampled data ($n = 8,250$ tags / sample) to assess the variation that may arise due to differences in the total number of tags recovered per sample. Departures from the patterns observed with the real data (if any) are recorded in the results.

Community phylogenetic reconstructions. Absolute abundances were converted to log abundances and formatted for **PAUP** 4b10 (16) as in (1). Phylogenies inferring the relatedness of samples based on their microbial community compositions and the bootstrap support for branching orders in these phylogenies were generated using default Parsimony and Neighbor-joining parameters.

Integrating data of Gombe chimpanzees with previous analyses of great ape microbiotae. The original pyrotag sequences obtained for fecal samples from five great ape species (1) were retrieved, quality trimmed and processed as described above. The filtered pyrotags were

combined with filtered tags recovered from the samples from the Gombe chimpanzees (this study), resulting in a large number (>200,000) of unique OTUs. OTUs present in a single sample were removed, and the remaining 100% OTUs aligned and clustered in the **RDP** pipeline (7). Nonmetric multidimensional scaling and phylogenetic analyses were carried out as described.

Illumina tag (iTag) sequencing. To examine changes in gut microbial communities at a very high level of resolution, we implemented an *Illumina*-based 16S-tag procedure that is analogous in principal to the pyrotag procedure described above, but differs by interrogating a 100-nt region from the 5'-end of amplicons spanning the V4 region of 16S rDNA (2). Because the short read-lengths produced by this method both hampers the phylogenetic assignment of large proportions of the sequencing reads and precludes direct comparisons to available pyrotags that targeted other regions of the 16S rDNA molecule, we applied these procedures to (i) monitor phylotype diversity within individual hosts over time, and (ii) examine the effects of long-term sample storage on changes in phylotype diversity.

DNA was extracted from faecal samples by a bead-beating method (17), cleaned using a QIAGEN QIAquick PCR purification kit and diluted to 25 ng/μl. PCR reaction were carried out in triplicate 25-μl reactions using universal forward primer F515 (5'- AAT GAT ACG GCG ACC ACC GAG ACG TAC GTA CGG Tgt gcc agc mgc cgc ggt aa) and universal reverse primer R806 (CAA GCA GAA GAC GGC ATA CGA GAT-**NNNNNNNNNNNN**-AGT CAG TCA GCC-gga cta chv ggg twt cta at), each fused to the *Illumina* adaptor-sequences. Note that reverse primers were indexed with a unique 12-nucleotide barcode sequences (designated by **N**'s) situated between *Illumina* adaptor sequence and the R808 conserved sequence (lower case), thereby allowing us to multiplex all samples into a single *Illumina* flowcell. Triplicate reactions were pooled and

purified using AMPure XP beads (Beckman Coulter Genomics), quantified using Quant-iT PicoGreen (Invitrogen), combined in equimolar amounts, and run over a QIAquick column (QIAGEN) for final purification. Samples were sequenced at the Yale Center for Genome Analysis using *Illumina* HiSeq 2000 technology.

iTags were generated for a total of 39 samples, which represent annual samples spanning 10 consecutive years for seven chimpanzees that were also surveyed at multiple time-points by the pyrotag procedure (Table S4). Initial quality filtering was performed in **Qiime v1.3**, using default parameters (12), which was then applied to determine OTUs using a closed reference picking protocol against precomputed 97% and 99% **Greengenes** OTUs with the **UCLUST** algorithm. OTUs were further culled by removing those identified in only a single sample and/or found at abundance <0.01% of the total number of reads, noting that such aggressive quality-trimming is essential when using *Illumina*-based 16S tag approaches (16, 18).

Literature Cited

1. Ochman H, et al. (2010) Evolutionary relationships of wild hominids recapitulated by gut microbial communities. *PLoS Biol* 8:e1000546.
2. Caporaso JG, et al. (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA* 108:4516–4522.
3. Hamady M, Walker JJ, Harris JK, Gold NJ, Knight R (2008) Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat Methods* 5:235–237.
4. Degnan PH, Ochman H (2012) *Illumina*-based analysis of microbial community diversity. *ISME J*, 6:183–194.

5. Keele BF, et al. (2009) Increased mortality and AIDS-like immunopathology in wild chimpanzees infected with SIVcpz. *Nature* 460:515–519.
6. Constable JL, Ashley MV, Goodall J, Pusey AE (2001) Noninvasive paternity assignment in Gombe chimpanzees. *Mol Ecol* 10:1279–1300.
7. Cole JR, et al. (2009) The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* 37:D141–D145.
8. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27:2194-2200.
9. File SK, McGrew WC, Tutin CE (1976) The intestinal parasites of a community of feral chimpanzees, *Pan troglodytes schweinfurthii*. *J Parasitol* 62:259–261.
10. Bakuza, JS, Nkwenguilila G. (2009) Variation over time in parasite prevalence among free-ranging chimpanzees in Gombe National Park, Tanzania. *Int J Primatol* 30:40–53.
11. Gillespie TR, et al. (2010) Demographic and ecological effects of patterns of parasitism in eastern chimpanzees (*Pan troglodytes schweinfurthii*) in Gombe National Park, Tanzania. *Amer J Phys Anthropol* 143:534–544.
12. Colwell RK (2009) User's Guide to EstimateS Vers. 8.2: Statistical estimation of species richness and shared species from samples. Online User's Guide. University of Connecticut. 22 pp.
13. Caporaso JG, et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7:335–336.
14. McCune B, Medford MJ (1999) PC-ORD: multivariate analysis of ecological data. Version 4. MJM, Glenden Beach, OR.

15. Dufrêne M, Legendre P (1997) Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecol Monogr* 67:345–366.
16. Swofford DL (2004) PAUP*: Phylogenetic Analysis Using Parsimony (and Other Methods). Sinauer Associates, Inc., Sunderland, MA.
17. Goodman AL, et al. (2011). Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proc Natl Acad Sci USA* 108:6252–6257.
18. Degnan PH, Ochman H (2012) *Illumina*-based analysis of microbial community diversity. *ISME J*, 6:183–194.

Fig. S1. Relative taxon abundances of Gombe chimp 16S rRNA pyrotags. Stacked Bar graphs illustrate the abundances of **(A)** Phyla **(B)** Class assignments for quality trimmed and clustered 99% OTUs as determined by the RDP Classifier for the 50 pyrotag chimp samples.

Fig. S2. Comparison of alpha diversity (Shannon's H) of gut microbial communities and age of individual. Alpha diversity indices were calculated for each of the 34 samples collected in 2008 and plotted by the age of each individual. (Infants, 0-5 yrs.; Juveniles, 6 – 8 yrs.; Adolescents, 9 – 15 yrs., Adults, 16 – 33 yrs.; Elderly, > 33 yrs.) Nonparametric tests revealed that infants and juveniles had slight but significantly greater microbial community diversity than did adults. The negative slope estimated from a linear regression also suggests this subtle trend.

Fig. S3. Proportions of shared OTUs among longitudinally sampled chimps. Venn diagrams are used to represent the numbers of recurring 99% OTUs identified in six chimpanzees sampled over an 8 year period **(A–F)**. Labels adjacent to each colored circle (2000 - yellow, 2001 - orange, and 2008 - green) indicate the sample number, year and total number of 99% pyrotag reads recovered. Numbers in each sector indicate the corresponding number of OTUs (**bold**) and average percentage of reads (roman).

Fig. S4. Nonmetric multidimensional scaling ordination of denser longitudinal sampling of chimps. Using iTags, the 20 original temporal samples from 2000, 2001 and 2008 (**bold**) and 19 additional fecal samples from intervening years (roman) were sequenced, clustered into 99% OTUs and analyzed using NMS ordination. Similar to the pyrotag results presented in Figure 5,

samples from 2000 (blue) and 2001 (red) more or less overlap in the ordination space and samples from 2008 (yellow) are largely distinct. The key denotes the year and individual of the samples and ellipses encompasses all of the samples from 2000 (blue), 2001 (red) and 2008 (yellow). Related chimpanzees are indicated by the superscripts ¹Sparrow is the mother of Sheldon and Sandi; ²Darbee and Tubi are presumed to be half-sibs.

Fig. S5 Gombe chimps exhibit stable core community through time. Shared OTUs from each temporal sample are represented as heatmaps in which the height of each row within columns (representing a samples from the designated year) is proportional to the number of OTUs (labeled at left). Cells are colored-coded according to the \log_{10} transformed read abundance of those shared OTUs from the sample indicated.

Fig. S6. Alpha diversity values estimated from iTag data within chimps and through time.

Microbial diversity estimates for fecal samples collected over a nine-year period were calculated to examine changes in diversity that might result from DNA degradation during sample storage. Although there are differences observed among sample estimates of Shannon's H within chimpanzees (A) and between years (B), there is no systemic bias leading to inflated diversity estimates in all older samples.

Figure S1

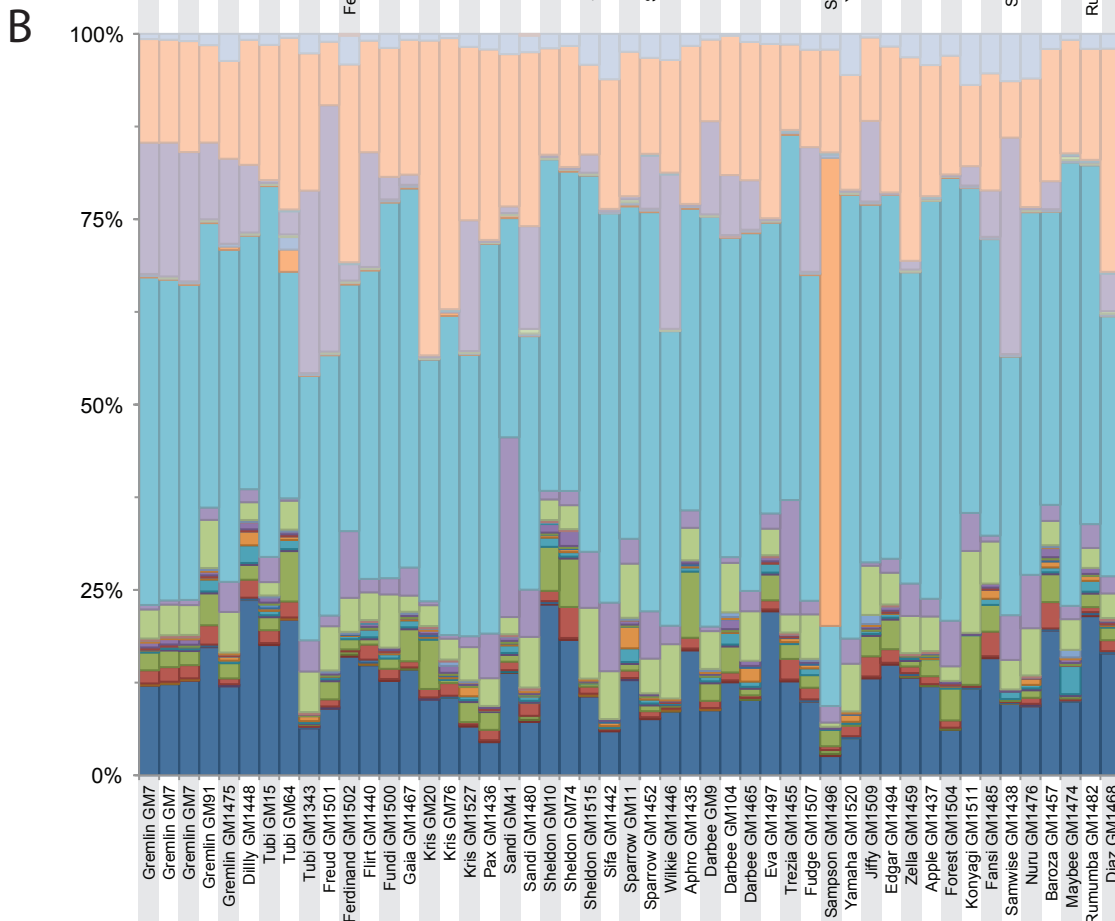
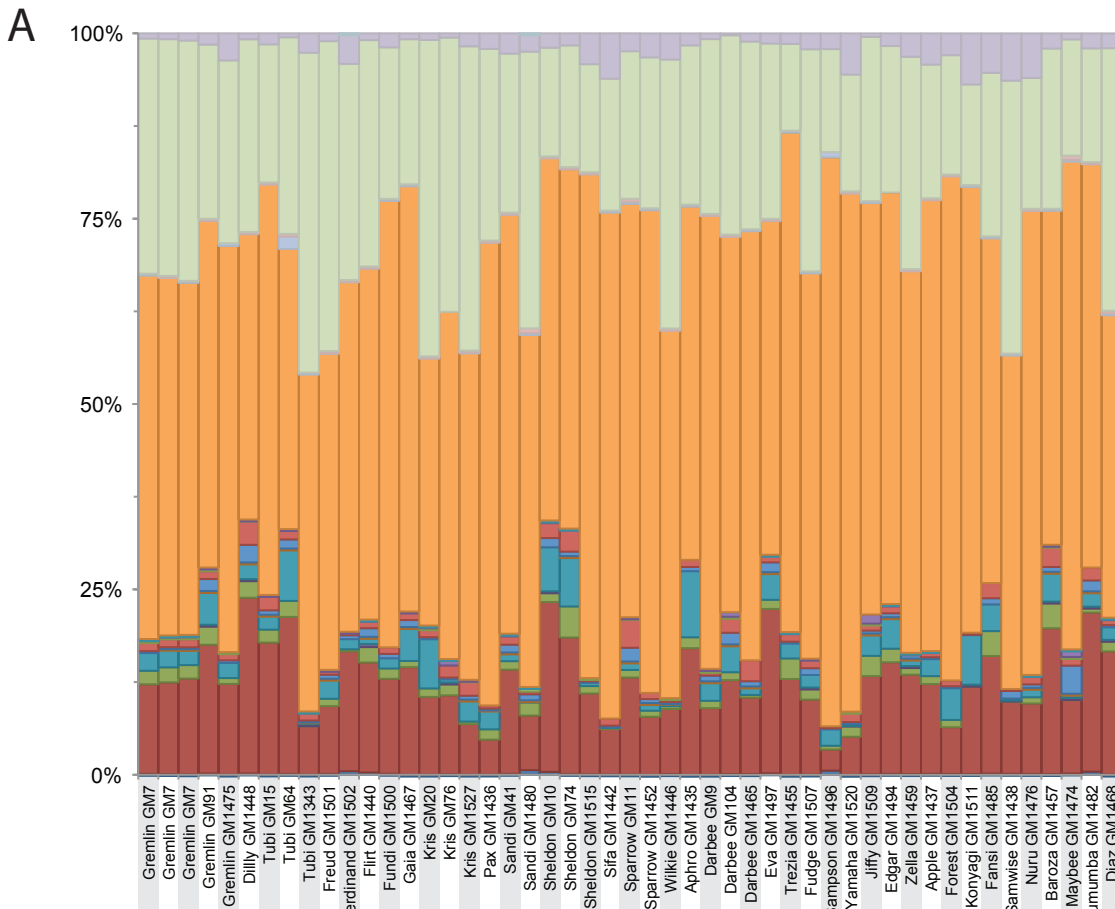


Figure S2

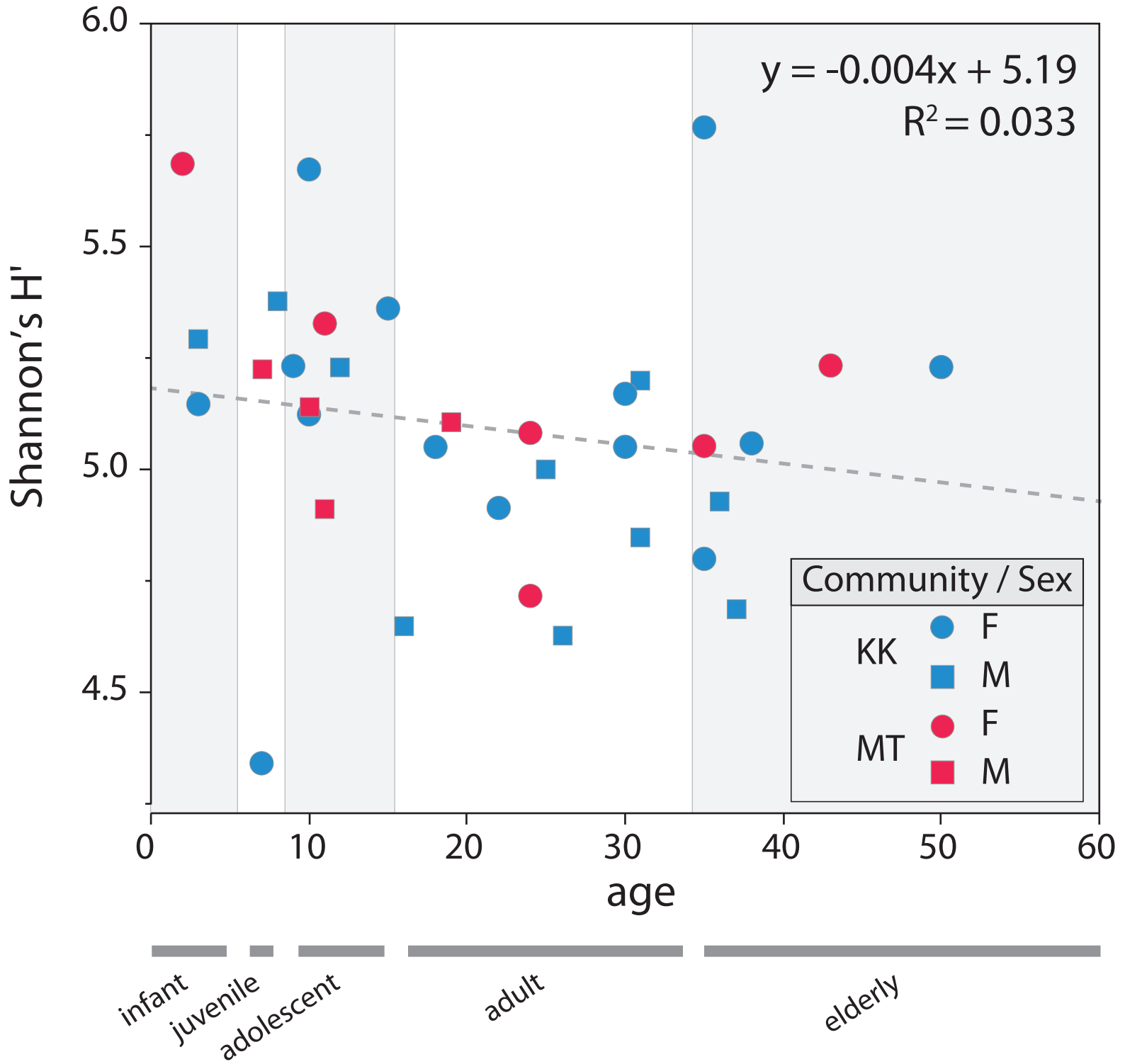


Figure S3

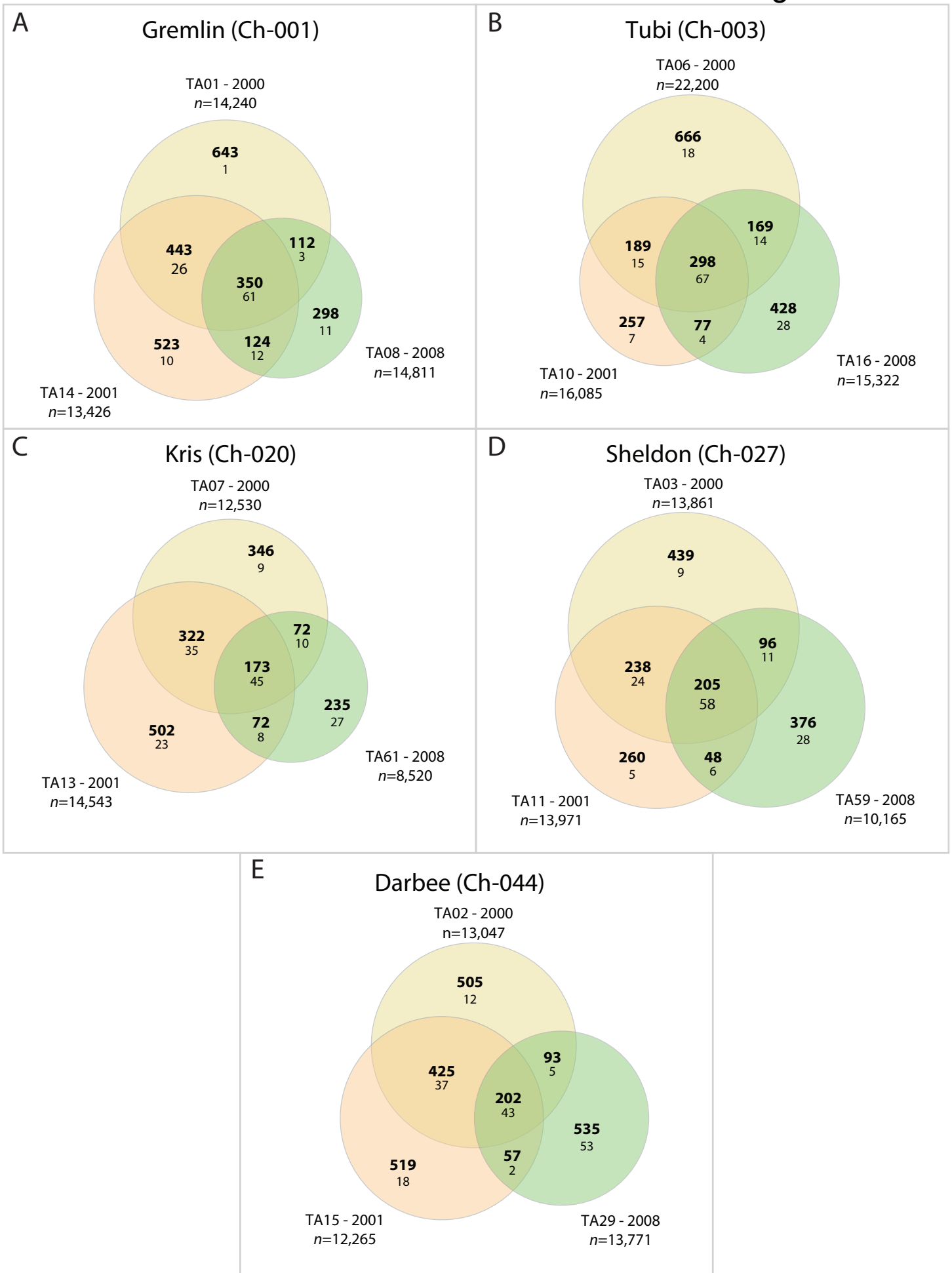
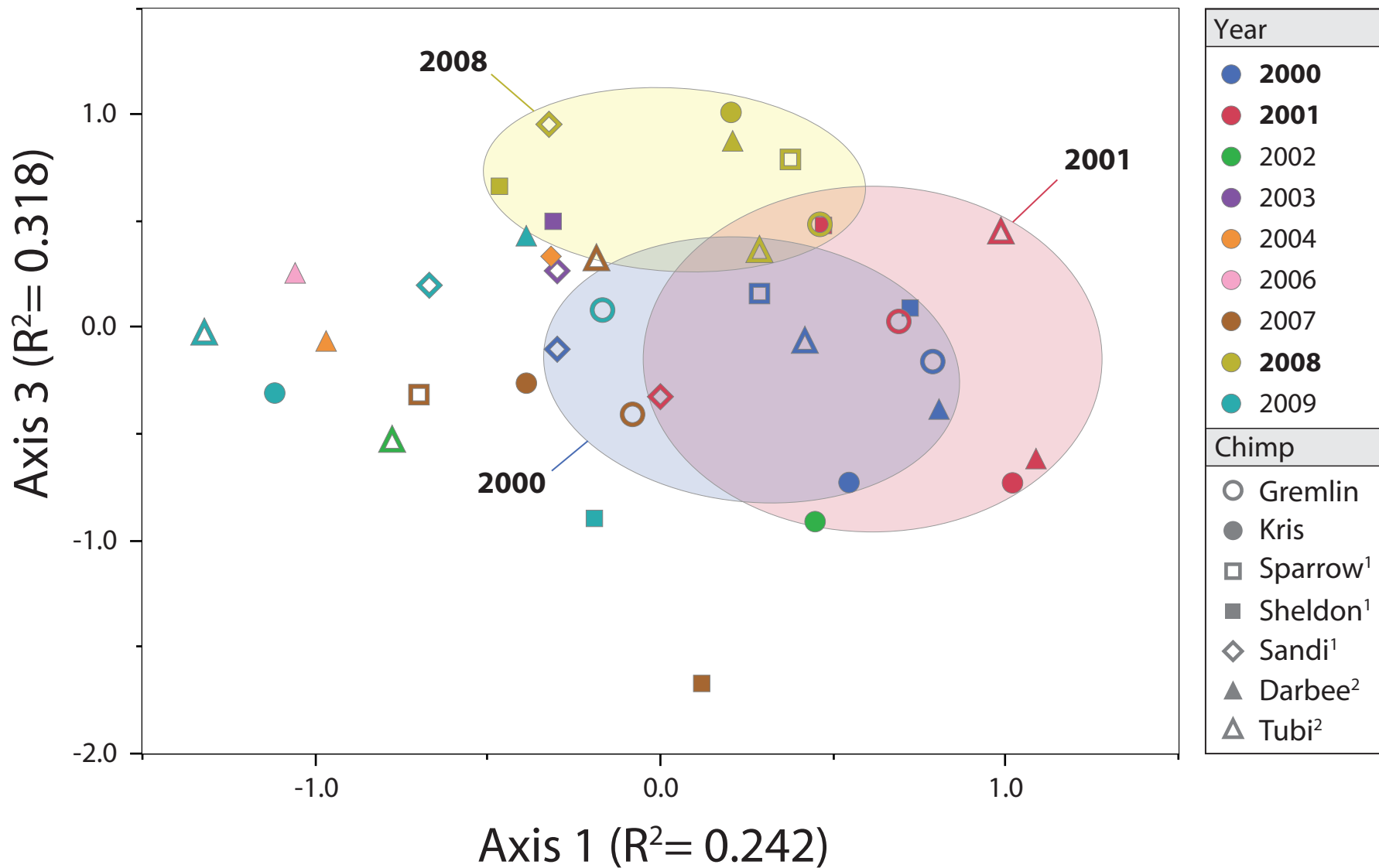
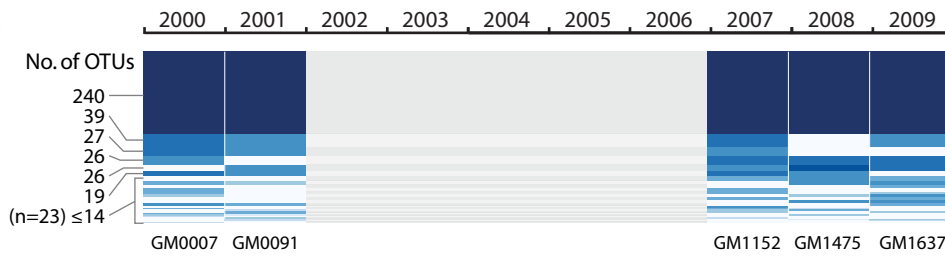


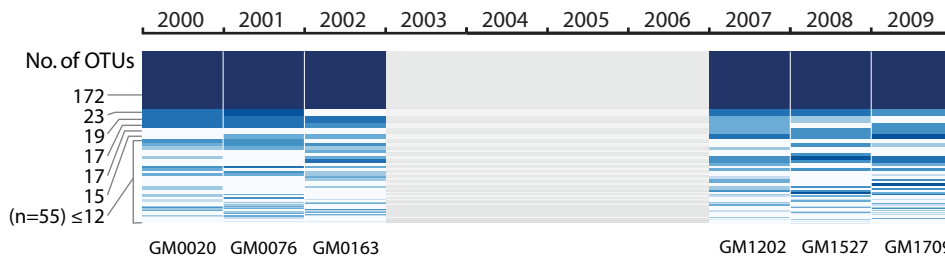
Figure S4



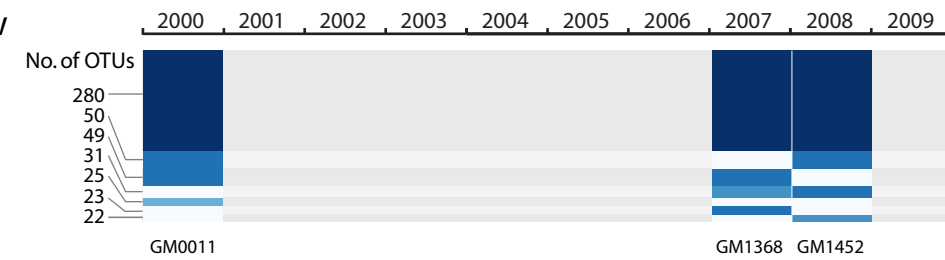
Gremlin



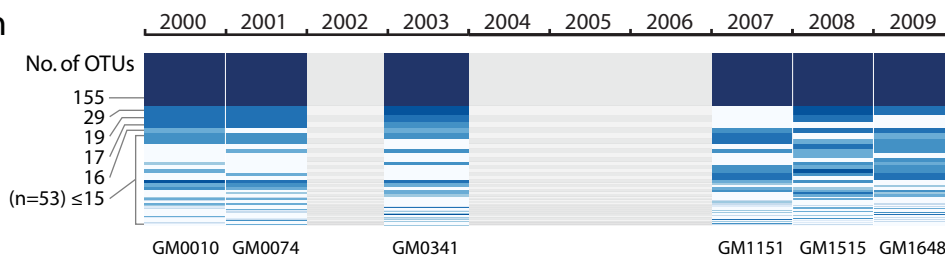
Kris



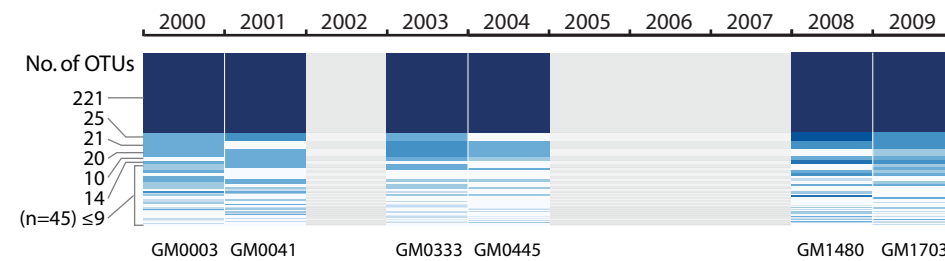
Sparrow



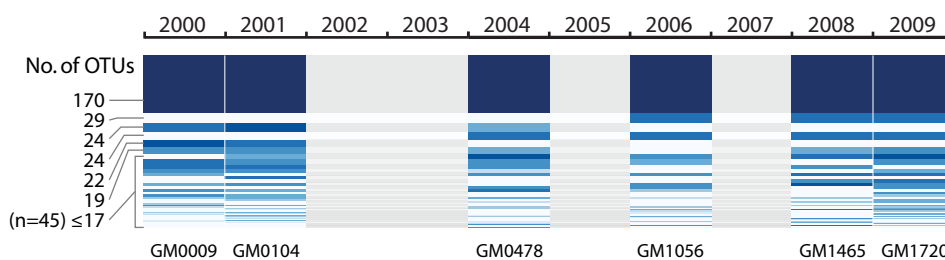
Sheldon



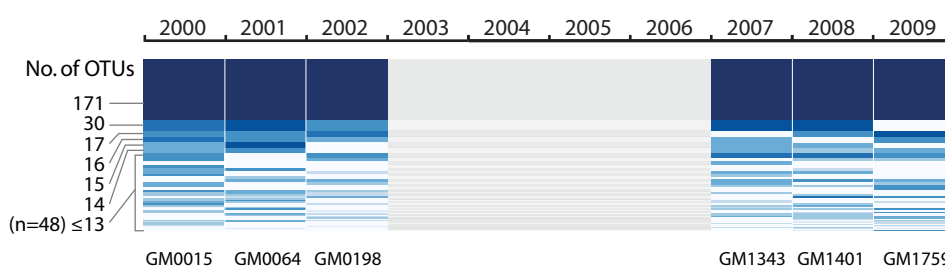
Sandi



Darbee



Tubi



Color Key

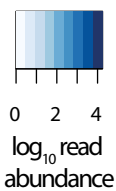


Figure S6

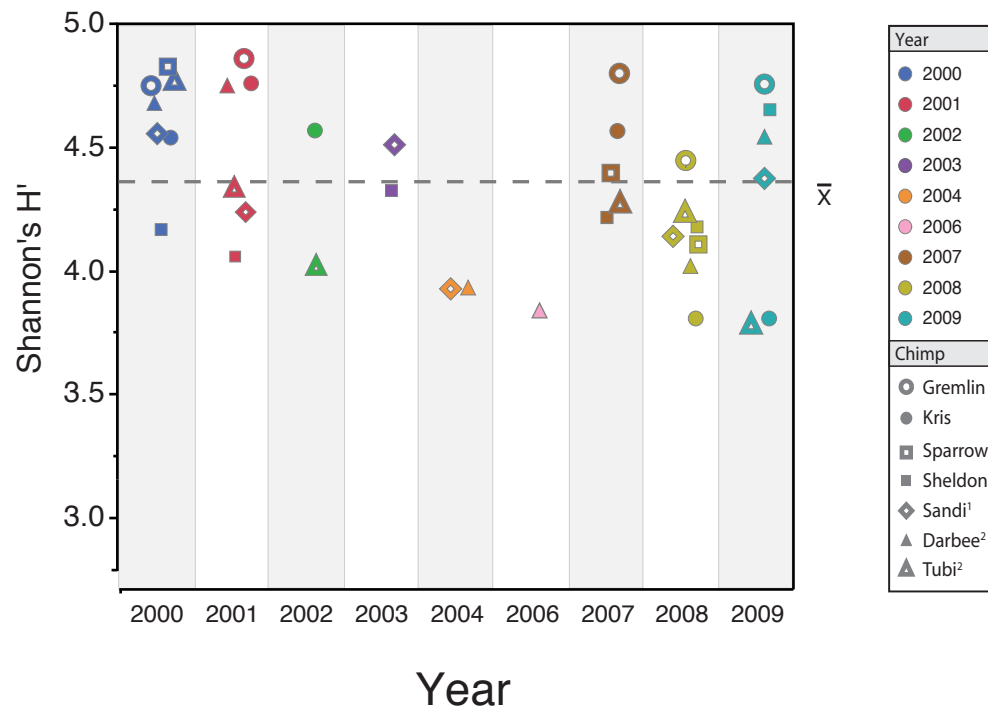
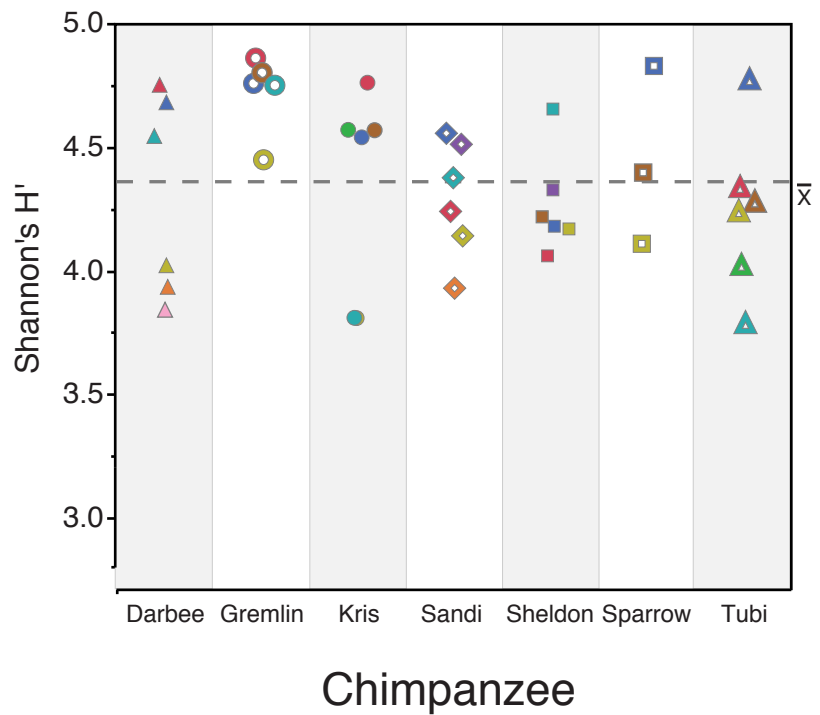


Table S1. Summary of pyrosequenced Gombe chimp samples

Sample tag	Barcode	Name	Abbr.	Chimp	Sex	Community ¹	YOB ²	Age when collected	Sample number	Total No. of processed reads ³
TA01	ACAGCACT	Gremlin	GM	Ch-001	F	KK	1970	30	GM7	14,752
TA02	ACAGCTCA	Darbee	DB	Ch-044	F	MT	1984	17	GM9	13,662
TA03	ACGAGTAG	Sheldon	SL	Ch-027	M	KK	1983	17	GM10	14,014
TA04	ACGTACGT	Sparrow	SW	Ch-031	F	KK	1958	42	GM11	15,416
TA05	ACTCACAG	Gremlin	GM	Ch-001	F	KK	1970	30	GM7	18,515
TA06	ACTCGTGA	Tubi	TB	Ch-003	M	KK	1977	23	GM15	22,563
TA07	ACTGCAGT	Kris	KS	Ch-020	M	KK	1982	19	GM20	12,747
TA08	AGACGACA	Gremlin	GM	Ch-001	F	KK	1970	38	GM1475	14,896
TA09	AGACTCTG	Sandi	SA	Ch-025	F	KK	1973	28	GM41	16,617
TA10	AGAGCAGT	Tubi	TB	Ch-003	M	KK	1977	24	GM64	16,192
TA11	AGCAGTAG	Sheldon	SL	Ch-027	M	KK	1983	19	GM74	14,036
TA13	AGTCGTCA	Kris	KS	Ch-020	M	KK	1982	20	GM76	14,939
TA14	AGTGACAG	Gremlin	GM	Ch-001	F	KK	1970	31	GM91	13,898
TA15	AGTGCTCA	Darbee	DB	Ch-044	F	MT	1984	18	GM104	12,616
TA16	ATCGCGTA	Tubi	TB	Ch-003	M	KK	1977	31	GM1343	15,566
TA17	ATGCGCTA	Aphro	AP	Ch-042	F	MT	1973	35	GM1435	14,092
TA18	CACTCAGA	Pax	PX	Ch-023	M	KK	1977	31	GM1436	16,131
TA19	CACTGAGT	Apple	APL	Ch-063	M	MT	1998	11	GM1437	15,972
TA20	CAGACTGT	Samwise	SAM	Ch-073	F	KK	2001	7	GM1438	17,732
TA21	CAGATCTG	Flirt	FLI	Ch-012	F	KK	1998	10	GM1440	12,804
TA22	CAGTCTGA	Sifa	SIF	Ch-029	F	KK	1978	30	GM1442	14,580
TA23	CATCGATG	Wilkie	WL	Ch-035	M	KK	1972	36	GM1446	14,136
TA24	CATGACGT	Dilly	DL	Ch-002	F	KK	1986	22	GM1448	13,234
TA25	CATGTCGA	Sparrow	SW	Ch-031	F	KK	1958	50	GM1452	15,817
TA26	CGATGCAT	Trezia	TZ	Ch-050	F	KK	1978	30	GM1455	12,567
TA27	CGTAGCTA	Baroza	BRZ	Ch-090	M	KK	2005	3	GM1457	15,139

TA28	CTACGTAG	Zella	ZEL	Ch-061	F	KK	1999	9	GM1459	13,647
TA29	CTACTGCT	Darbee	DB	Ch-044	F	MT	1984	25	GM1465	14,077
TA30	CTAGTCGA	Gaia	GA	Ch-015	F	KK	1993	16	GM1467	15,085
TA31	CTCACTGT	Diaz	DIA	Ch-097	F	KK	2005	3	GM1468	13,306
TA32	CTCAGACA	Maybee	MAY	Ch-094	F	MT	2006	3	GM1474	12,237
TA34	CTGACAGT	Nuru	NUR	Ch-079	F	KK	1990	18	GM1476	11,402
TA35	CTGTCAGA	Sandi	SA	Ch-025	F	KK	1973	35	GM1480	12,464
TA36	TACGCGAT	Rumumba	RUM	Ch-096	F	MT	1997	11	GM1482	12,164
TA37	TAGCTACG	Fansi	FAN	Ch-068	M	MT	2001	7	GM1485	11,037
TA38	TCACGACA	Edgar	EDG	Ch-059	M	MT	1989	19	GM1494	11,410
TA39	TCACTGAG	Sampson	SN	Ch-054	M	KK	1996	12	GM1496	9,893
TA40	TCAGTCAG	Eva	EVA	Ch-049	F	MT	1965	43	GM1497	9,141
TA51	TCTCGTCA	Fundi	FND	Ch-014	M	KK	2000	8	GM1500	9,736
TA52	ACACGATGCT	Gremlin	GM	Ch-001	F	KK	1970	30	GM7	9,159
TA53	CACATGCTAG	Freud	FD	Ch-005	M	KK	1971	37	GM1501	9,774
TA54	CACTCTACTG	Ferdinand	FE	Ch-010	M	KK	1992	16	GM1502	8,323
TA55	CGAGCTAGCT	Forest	FOR	Ch-065	M	MT	1997	11	GM1504	9,143
TA56	CTAGCATGCT	Fudge	FU	Ch-051	M	KK	1996	12	GM1507	9,738
TA57	CTATCGTACG	Jiffy	JF	Ch-057	F	KK	1973	35	GM1509	10,422
TA58	CTCGTATACG	Konyagi	KON	Ch-066	F	MT	1984	24	GM1511	8,721
TA59	TGATATCGCG	Sheldon	SL	Ch-027	M	KK	1983	25	GM1515	10,225
TA60	TGCGATATCG	Yamaha	YAM	Ch-055	F	KK	1998	10	GM1520	9,616
TA61	TGCGATCGTA	Kris	KS	Ch-020	M	KK	1982	27	GM1527	8,545

¹Community affiliation at time of sampling KK= Kasekela, MT= Mitumba

²Year of birth

³Includes phylogenetically uninformative reads

Table S2. Summary of 99% phylotypes used in the analysis of all great ape samples

Sample tag	Name ¹	Ape	Gender	Popu- lation	Sample number	Collection date	Common name	Location	Abbreviation	No. of Reads	No. of OTUs
TA01	Gremlin	Ch-001	F	KK	GM7	8-Nov-00	Chimp	Gombe	Pts	12,740	1,150
TA05	Gremlin	Ch-001	F	KK	GM7	8-Nov-00	Chimp	Gombe	Pts	16,263	1,254
TA52	Gremlin	Ch-001	F	KK	GM7	8-Nov-00	Chimp	Gombe	Pts	8,120	952
TA14	Gremlin	Ch-001	F	KK	GM91	31-Dec-01	Chimp	Gombe	Pts	12,181	1,091
TA08	Gremlin	Ch-001	F	KK	GM1475	21-Oct-08	Chimp	Gombe	Pts	13,870	713
TA24	Dilly	Ch-002	F	KK	GM1448	27-Oct-08	Chimp	Gombe	Pts	12,465	644
TA06	Tubi	Ch-003	M	KK	GM15	20-Nov-00	Chimp	Gombe	Pts	20,380	1,002
TA10	Tubi	Ch-003	M	KK	GM64	11-Dec-01	Chimp	Gombe	Pts	15,098	650
TA16	Tubi	Ch-003	M	KK	GM1343	16-Apr-08	Chimp	Gombe	Pts	14,369	751
TA53	Freud	Ch-005	M	KK	GM1501	15-Oct-08	Chimp	Gombe	Pts	9,053	624
TA54	Ferdinand	Ch-010	M	KK	GM1502	16-Nov-08	Chimp	Gombe	Pts	7,751	435
TA21	Flirt	Ch-012	F	KK	GM1440	21-Oct-08	Chimp	Gombe	Pts	11,283	1,021
TA51	Fundi	Ch-014	M	KK	GM1500	21-Oct-08	Chimp	Gombe	Pts	9,078	669
TA30	Gaia	Ch-015	F	KK	GM1467	15-Oct-08	Chimp	Gombe	Pts	13,997	884
TA07	Kris	Ch-020	M	KK	GM20	10-Nov-00	Chimp	Gombe	Pts	11,433	763
TA13	Kris	Ch-020	M	KK	GM76	15-Dec-01	Chimp	Gombe	Pts	13,206	897
TA61	Kris	Ch-020	M	KK	GM1527	9-Oct-08	Chimp	Gombe	Pts	8,072	448
TA18	Pax	Ch-023	M	KK	GM1436	21-Oct-08	Chimp	Gombe	Pts	14,896	722
TA09	Sandi	Ch-025	F	KK	GM41	24-Dec-01	Chimp	Gombe	Pts	15,471	669
TA35	Sandi	Ch-025	F	KK	GM1480	21-Oct-08	Chimp	Gombe	Pts	11,651	590
TA03	Sheldon	Ch-027	M	KK	GM10	10-Nov-00	Chimp	Gombe	Pts	12,676	735
TA11	Sheldon	Ch-027	M	KK	GM74	24-Dec-01	Chimp	Gombe	Pts	13,123	583
TA59	Sheldon	Ch-027	M	KK	GM1515	17-Oct-08	Chimp	Gombe	Pts	9,421	607
TA22	Sifa	Ch-029	F	KK	GM1442	2-Nov-08	Chimp	Gombe	Pts	13,547	586
TA04	Sparrow	Ch-031	F	KK	GM11	6-Nov-00	Chimp	Gombe	Pts	14,085	699
TA25	Sparrow	Ch-031	F	KK	GM1452	15-Oct-08	Chimp	Gombe	Pts	14,808	743
TA23	Wilkie	Ch-035	M	KK	GM1446	17-Oct-08	Chimp	Gombe	Pts	13,011	616
TA17	Aphro	Ch-042	F	MT	GM1435	14-Oct-08	Chimp	Gombe	Pts	13,126	669
TA02	Darbee	Ch-044	F	MT	GM9	26-Oct-00	Chimp	Gombe	Pts	11,484	897
TA15	Darbee	Ch-044	F	MT	GM104	30-Dec-01	Chimp	Gombe	Pts	10,952	996
TA29	Darbee	Ch-044	F	MT	GM1465	30-Sep-08	Chimp	Gombe	Pts	12,879	622
TA40	Eva	Ch-049	F	MT	GM1497	11-Oct-08	Chimp	Gombe	Pts	8,513	620
TA26	Trezia	Ch-050	F	KK	GM1455	7-Nov-08	Chimp	Gombe	Pts	11,569	713

TA56	Fudge	Ch-051	M	KK	GM1507	21-Oct-08	Chimp	Gombe	Pts	8,854	728
TA39	Sampson	Ch-054	M	KK	GM1496	21-Oct-08	Chimp	Gombe	Pts	9,332	270
TA60	Yamaha	Ch-055	F	KK	GM1520	23-Oct-08	Chimp	Gombe	Pts	8,952	554
TA57	Jiffy	Ch-057	F	KK	GM1509	24-Nov-08	Chimp	Gombe	Pts	8,998	959
TA38	Edgar	Ch-059	M	MT	GM1494	10-Nov-08	Chimp	Gombe	Pts	10,677	609
TA28	Zella	Ch-061	F	KK	GM1459	17-Oct-08	Chimp	Gombe	Pts	12,355	712
TA19	Apple	Ch-063	M	MT	GM1437	19-Nov-08	Chimp	Gombe	Pts	14,767	631
TA55	Forest	Ch-065	M	MT	GM1504	13-Nov-08	Chimp	Gombe	Pts	8,630	509
TA58	Konyagi	Ch-066	F	MT	GM1511	21-Oct-08	Chimp	Gombe	Pts	8,199	491
TA37	Fansi	Ch-068	M	MT	GM1485	23-Oct-08	Chimp	Gombe	Pts	9,891	659
TA20	Samwise	Ch-073	F	KK	GM1438	17-Oct-08	Chimp	Gombe	Pts	16,700	513
TA34	Nuru	Ch-079	F	KK	GM1476	21-Oct-08	Chimp	Gombe	Pts	10,551	618
TA27	Baroza	Ch-090	M	KK	GM1457	4-Dec-08	Chimp	Gombe	Pts	13,925	834
TA32	Maybee	Ch-094	F	MT	GM1474	24-Nov-08	Chimp	Gombe	Pts	10,692	814
TA36	Rumumba	Ch-096	F	MT	GM1482	18-Oct-08	Chimp	Gombe	Pts	11,219	679
TA31	Diaz	Ch-097	F	KK	GM1468	21-Nov-08	Chimp	Gombe	Pts	11,677	710
LK645	n.a.	B_KSG0645	n.a.	n.a.	KSG0645	24-Mar-06	Bonobo	Lui-kotal, DRC	Pp	13,299	728
LK647	n.a.	B_KSG0647	n.a.	n.a.	KSG0647	28-Mar-06	Bonobo	Lui-kotal, DRC	Pp	11,757	701
LK652	n.a.	B_KSG0652	n.a.	n.a.	KSG0652	6-Apr-06	Bonobo	Lui-kotal, DRC	Pp	5,120	385
LK660	n.a.	B_KSG0660	n.a.	n.a.	KSG0660	5-May-06	Bonobo	Lui-kotal, DRC	Pp	16,493	775
LK673	n.a.	B_KSG0673	n.a.	n.a.	KSG0673	17-May-06	Bonobo	Lui-kotal, DRC	Pp	16,251	880
GM173	Vincent	Cs_GM173	M	MT	GM173	8-May-02	Chimp	Gombe	Pts	37,130	1,081
BI242	n.a.	Cs_KSG242	M	n.a.	KSG242	12-May-05	Chimp	Babingi	Pts	7,273	305
BI246	n.a.	Cs_KSG246	F	n.a.	KSG246	19-May-05	Chimp	Babingi	Pts	2,886	208
WK12	n.a.	Cs_WLK12	M	n.a.	WLK12	10-May-04	Chimp	Walikale, DRC	Pts	4,790	368
WK03	n.a.	Cs_WLK3	F	n.a.	WLK3	11-May-04	Chimp	Walikale, DRC	Pts	1,920	205
MT334	n.a.	Ct_CR334	F	n.a.	CR334	1-Mar-04	Chimp	Minta	Ptt	11,094	749
MT348	n.a.	Ct_CR348	M	n.a.	CR348	1-Mar-04	Chimp	Minta	Ptt	18,792	996
WE464.1	n.a.	Ct_CR464	n.a.	n.a.	CR464	28-May-04	Chimp	Wassa Emtse	Ptt	10,090	562
CP469	n.a.	Ct_CR469	F	n.a.	CR469	12-May-04	Chimp	Campo Ma'an	Ptt	18,066	857
CP470	n.a.	Ct_CR470	M	n.a.	CR470	14-May-04	Chimp	Campo Ma'an	Ptt	54,381	971
BB089	n.a.	Ct_CR89	M	n.a.	CR89	7-Jun-03	Chimp	Boumba Bek	Ptt	32,984	1,309
BB095	n.a.	Ct_CR95	F	n.a.	CR95	7-Jun-03	Chimp	Boumba Bek	Ptt	31,437	1,160
WE464.2	n.a.	Ct_WCR464	n.a.	n.a.	WCR464	28-May-04	Chimp	Wassa Emtse	Ptt	22,813	665
WE457	n.a.	Cv_CR457	F	n.a.	CR457	28-May-04	Chimp	Wassa Emtse	Pte	44,128	1,168
WE458	n.a.	Cv_CR458	M	n.a.	CR458	28-May-04	Chimp	Wassa Emtse	Pte	108,342	1,686
ME2528	n.a.	G_CR2528	n.a.	n.a.	CR2528	20-Jul-07	Gorilla	Ndoki NP, CAP	Gg	12,385	515

ME2530	n.a.	G_CR2530	n.a.	n.a.	CR2530	20-Jul-07	Gorilla	Ndoki NP, CAP	Gg	12,544	476
KE1471	n.a.	G_KSG1471	n.a.	n.a.	KSG1471	26-Feb-07	Gorilla	Walikale, DRC	Gb	15,779	485
KE1507	n.a.	G_KSG1507	n.a.	n.a.	KSG1507	26-Feb-07	Gorilla	Walikale, DRC	Gb	8,157	306
KS477	n.a.	H_KSG477	M	n.a.	KSG477	21-Dec-06	Human	Congo	Hs	1,190	52
US1AZ	n.a.	H_US1AZ	M	n.a.	US1AZ	3-Jun-08	Human	Arizona	Hs	831	100

¹n.a. = Information not available

Table S3. Alpha diversity indices measured for Gombe chimpanzee samples

Sample tag	Name	Sex	Community ¹	YOB ²	Sample number	Stage	Collection year	Matriline ³	Total no. of						
									reads ⁴	99% OTUs	Chao1	Shannon's H'	Evenness (E)	OTU F:B ratio	Read F:B ratio
TA01	Gremlin	F	KK	1970	GM7	adult	2000	G	14,240	1,548	2,447	5.80	0.790	3.68	1.55
TA02	Darbee	F	MT	1984	GM9	adult	2000	B	13,047	1,225	1,906	5.56	0.782	5.00	2.60
TA03	Sheldon	M	KK	1983	GM10	adult	2000	S	13,860	977	1,306	5.30	0.770	4.18	3.35
TA04	Sparrow	F	KK	1958	GM11	old	2000	S	15,297	868	1,214	5.25	0.776	3.72	2.80
TA05	Gremlin	F	KK	1970	GM7	adult	2000	G	17,968	1,683	2,365	5.81	0.783	4.05	1.51
TA06	Tubi	M	KK	1977	GM15	adult	2000	B	22,200	1,322	1,993	5.62	0.781	4.67	2.98
TA07	Kris	M	KK	1982	GM20	adult	2000	U	12,530	913	1,299	5.00	0.733	3.75	0.85
TA08	Gremlin	F	KK	1970	GM1475	old	2008	G	14,811	884	1,215	5.06	0.746	4.35	2.22
TA09	Sandi	F	KK	1973	GM41	adult	2001	S	16,500	799	1,027	4.79	0.716	3.14	2.64
TA10	Tubi	M	KK	1977	GM64	adult	2001	B	16,085	821	1,199	4.94	0.736	2.51	1.42
TA11	Sheldon	M	KK	1983	GM74	adult	2001	S	13,971	751	1,062	5.03	0.760	3.22	2.95
TA13	Kris	M	KK	1982	GM76	adult	2001	U	14,543	1,069	1,451	5.60	0.803	3.72	1.27
TA14	Gremlin	F	KK	1970	GM91	adult	2001	G	13,426	1,440	2,178	5.75	0.791	3.94	1.99
TA15	Darbee	F	MT	1984	GM104	adult	2001	B	12,265	1,203	1,637	5.80	0.818	4.85	1.89
TA16	Tubi	M	KK	1977	GM1343	adult	2008	B	15,322	972	1,619	4.85	0.705	2.56	1.06
TA17	Aphro	F	MT	1973	GM1435	old	2008	A	14,003	829	1,108	5.05	0.752	3.68	2.22
TA18	Pax	M	KK	1977	GM1436	adult	2008	U	15,974	901	1,288	5.20	0.764	4.01	2.43
TA19	Apple	M	MT	1998	GM1437	adolescent	2008	A	15,805	779	1,070	5.14	0.772	3.70	3.39
TA20	Samwise	F	KK	2001	GM1438	juvenile	2008	S	17,635	660	940	4.34	0.669	2.28	1.23
TA21	Flirt	F	KK	1998	GM1440	adolescent	2008	F	12,450	1,389	2,285	5.67	0.784	4.06	1.55
TA22	Sifa	F	KK	1978	GM1442	adult	2008	U	14,437	737	1,063	5.05	0.765	5.27	3.85
TA23	Wilkie	M	KK	1972	GM1446	old	2008	U	13,966	818	1,271	4.93	0.735	2.57	1.37
TA24	Dilly	F	KK	1986	GM1448	adult	2008	D	13,196	802	1,201	4.91	0.735	2.46	1.49
TA25	Sparrow	F	KK	1958	GM1452	old	2008	S	15,734	930	1,388	5.23	0.765	4.29	3.20
TA26	Trezia	F	KK	1978	GM1455	adult	2008	T	12,481	895	1,262	5.17	0.761	4.93	5.76
TA27	Baroza	M	KK	2005	GM1457	infant	2008	U	14,917	1,089	1,627	5.29	0.757	3.08	2.09
TA28	Zella	F	KK	1999	GM1459	adolescent	2008	T	13,343	977	1,550	5.23	0.760	2.82	1.80
TA29	Darbee	F	MT	1984	GM1465	adult	2008	B	13,771	887	1,465	5.08	0.749	3.19	2.29
TA30	Gaia	F	KK	1993	GM1467	adult	2008	G	14,935	1,063	1,424	5.36	0.769	5.06	2.94
TA31	Diaz	F	KK	2005	GM1468	infant	2008	D	12,868	961	1,652	5.15	0.749	3.10	1.16
TA32	Maybee	F	MT	2006	GM1474	infant	2008	B	11,947	995	1,297	5.69	0.824	8.48	4.21
TA34	Nuru	F	KK	1990	GM1476	adult	2008	U	11,317	799	1,282	5.05	0.756	3.19	3.56
TA35	Sandi	F	KK	1973	GM1480	old	2008	S	12,387	720	1,058	4.80	0.729	2.37	1.28
TA36	Rumumba	F	MT	1997	GM1482	adolescent	2008	U	12,087	865	1,267	5.33	0.788	3.87	3.56
TA37	Fansi	M	MT	2001	GM1485	juvenile	2008	F	10,735	949	1,488	5.22	0.762	2.93	2.12
TA38	Edgar	M	MT	1989	GM1494	adult	2008	E	11,340	754	1,138	5.11	0.771	3.29	2.81
TA39	Sampson	M	KK	1996	GM1496	adolescent	2008	S	9,823	326	493	2.79	0.482	2.05	5.53
TA40	Eva	F	MT	1965	GM1497	old	2008	E	9,102	778	1,160	5.23	0.786	3.24	1.91
TA51	Fundi	M	KK	2000	GM1500	juvenile	2008	F	9,668	820	1,193	5.38	0.801	3.75	2.95
TA52	Gremlin	F	KK	1970	GM7	adult	2000	G	8,937	1,198	1,870	5.71	0.806	4.15	1.47
TA53	Freud	M	KK	1971	GM1501	old	2008	F	9,714	785	1,170	4.69	0.703	3.36	1.02
TA54	Ferdinand	M	KK	1992	GM1502	adult	2008	F	8,265	518	737	4.65	0.744	2.04	1.62

TA55	Forest	M	MT	1997	GM1504	adolescent	2008	F	9,047	619	880	4.91	0.764	5.24	4.24
TA56	Fudge	M	KK	1996	GM1507	adolescent	2008	F	9,588	955	1,596	5.23	0.762	3.27	1.74
TA57	Jiffy	F	KK	1973	GM1509	old	2008	U	10,052	1,120	1,505	5.77	0.821	6.85	2.51
TA58	Konyagi	F	MT	1984	GM1511	adult	2008	U	8,696	582	858	4.72	0.741	3.82	4.44
TA59	Sheldon	M	KK	1983	GM1515	adult	2008	S	10,165	725	1,014	5.00	0.759	4.32	4.68
TA60	Yamaha	F	KK	1998	GM1520	adolescent	2008	U	9,592	690	1,016	5.12	0.784	4.20	4.45
TA61	Kris	M	KK	1982	GM1527	adult	2008	U	8,520	552	794	4.63	0.733	2.49	1.08

¹Community affiliation at time of sampling KK= Kasekela, MT= Mitumba

²Year of birth

³G=Gremlin, B=Little Bee, A=Athena, D=Dilly, S=Sparrow, F=Fifi, T=Trusha,U=unknown

⁴Number of phylogenetically informative reads in 99% OTUs

Table S4. Alpha diversity indices calculated for longitudinally sampled chimps using iTag resequencing.

Chimpanzee	Sex	Community ¹	YOB ²	Collection date	Age when collected	Total No. of reads	No. of reads		Chao1	Shannon's H'	Evenness (E)
							No. of 99% OTUs	in 99% OTUs			
Gremlin	F	KK	1970	8-Nov-00	30	58,682	418	24,889	463.2	4.75	0.79
				31-Dec-01	31	27,907	399	10,497	434.3	4.86	0.81
				31-Jan-07	37	41,370	444	21,369	464.8	4.80	0.79
				21-Oct-08	38	40,832	374	22,225	400.6	4.45	0.75
				11-Oct-09	39	37,527	422	20,962	442.0	4.76	0.79
Tubi	M	KK	1977	20-Nov-00	23	44,216	403	23,574	420.5	4.78	0.80
				11-Dec-01	24	64,882	357	25,800	366.4	4.34	0.74
				24-Apr-02	25	32,242	337	18,179	394.2	4.02	0.69
				16-Apr-08	31	24,901	361	13,213	396.8	4.24	0.72
				17-Sep-07	30	40,430	369	22,729	390.1	4.28	0.72
				9-Oct-09	32	76,627	325	25,442	368.2	3.79	0.66
Kris	M	KK	1982	10-Nov-00	18	34,402	354	17,537	393.1	4.54	0.77
				15-Dec-01	19	58,157	355	25,577	382.0	4.76	0.81
				8-May-02	20	29,549	353	13,507	390.0	4.57	0.78
				7-Apr-07	25	46,064	387	14,156	411.0	4.57	0.77
				9-Oct-08	26	43,281	351	23,546	372.6	3.81	0.65
				9-Oct-09	27	48,409	365	24,602	418.3	3.81	0.65
Sandi	F	KK	1973	14-Nov-00	27	32,793	404	17,509	425.1	4.56	0.76
				24-Dec-01	28	51,009	370	25,180	408.3	4.24	0.72
				5-Aug-03	30	76,721	390	30,199	422.8	4.52	0.76
				30-Jun-04	31	43,517	328	24,893	340.8	3.93	0.68
				21-Oct-08	35	42,371	371	23,143	414.4	4.15	0.70
				27-Sep-09	36	44,540	393	20,428	414.6	4.38	0.73

Sheldon	M	KK	1983	10-Nov-00	17	45,318	361	23,973	395.7	4.17	0.71
				24-Dec-01	18	50,240	349	25,580	379.9	4.06	0.69
				27-Aug-03	20	37,438	349	21,855	366.6	4.33	0.74
				30-Jan-07	24	51,887	328	23,613	389.2	4.22	0.73
				17-Oct-08	25	41,672	384	22,429	426.5	4.18	0.70
				8-Sep-09	26	31,884	373	15,863	407.9	4.65	0.79
Sparrow	F	KK	1958	6-Nov-00	42	50,347	404	27,137	435.1	4.83	0.80
				16-Dec-01	43	36,502	371	19,587	385.5	4.80	0.81
				26-Jan-07	49	38,037	383	20,658	426.4	4.40	0.74
				15-Oct-08	50	60,380	383	35,091	407.8	4.11	0.69
Darbee	F	MT	1984	26-Oct-00	16	39,375	347	19,392	361.6	4.68	0.80047
				30-Dec-01	17	51,668	375	22,635	396.7	4.75	0.80198
				10-Jun-04	20	54,386	358	29,825	391.5	3.94	0.66967
				6-May-06	22	48,400	333	25,729	352.8	3.85	0.66246
				30-Sep-08	24	54,683	343	28,143	370.4	4.02	0.68934
				28-Oct-09	25	52,066	392	28,063	405.6	4.55	0.76123

¹Community affiliation at time of sampling KK= Kasekela, MT= Mitumba

²Year of birth