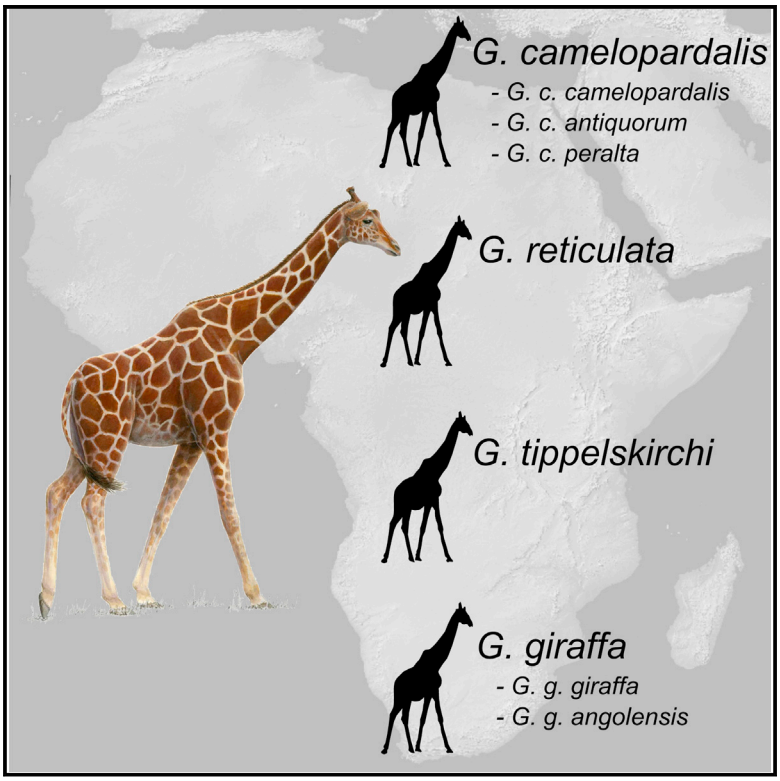
Report

Current Biology

Multi-locus Analyses Reveal Four Giraffe Species Instead of One

Graphical Abstract



Authors

Julian Fennessy, Tobias Bidon, Friederike Reuss, ..., Melita Vamberger, Uwe Fritz, Axel Janke

Correspondence

axel.janke@senckenberg.de

In Brief

Fennessy et al. show that multi-locus population genetic analyses indicate four distinct giraffe species instead of one. As the most inclusive genetic analysis of giraffe relationships to date, the findings highlight the need for targeted conservation efforts of the world's tallest megafauna.

Highlights

- Four genetically distinct giraffe clusters suggest separation into four species
- This is the first study using nuclear sequences and analyzing the Nubian giraffe
- Rothschild's giraffe should be subsumed into the nominate Nubian giraffe
- A giraffe survey genome produces valuable markers for phylogenomic analyses

Accession Numbers

LT596685
LT598170

Fennessy et al., 2016, Current Biology *26*, 2543–2549
September 26, 2016
http://dx.doi.org/10.1016/j.cub.2016.07.036

CellPress

Current Biology
Report

Multi-locus Analyses Reveal Four Giraffe Species Instead of One

Julian Fennessy,[1] Tobias Bidon,[2] Friederike Reuss,[2] Vikas Kumar,[2] Paul Elkan,[3] Maria A. Nilsson,[2] Melita Vamberger,[4] Uwe Fritz,[4] and Axel Janke[2,5,6,*]
[1]Giraffe Conservation Foundation, PO Box 86099, Windhoek, Namibia
[2]Senckenberg Biodiversity and Climate Research Centre, Senckenberganlage 25, 60325 Frankfurt am Main, Germany
[3]Wildlife Conservation Society, Bronx Zoo, 2300 Southern Blvd., Bronx, NY 10460, USA
[4]Museum of Zoology, Senckenberg Dresden, Königsbrücker Landstrasse 159, 01109 Dresden, Germany
[5]Institute for Ecology, Evolution & Diversity, Goethe University, Max-von-Laue-Strasse 13, 60439 Frankfurt am Main, Germany
[6]Lead Contact
*Correspondence: axel.janke@senckenberg.de
http://dx.doi.org/10.1016/j.cub.2016.07.036

SUMMARY

Traditionally, one giraffe species and up to eleven subspecies have been recognized [1]; however, nine subspecies are commonly accepted [2]. Even after a century of research, the distinctness of each giraffe subspecies remains unclear, and the genetic variation across their distribution range has been incompletely explored. Recent genetic studies on mtDNA have shown reciprocal monophyly of the matrilines among seven of the nine assumed subspecies [3, 4]. Moreover, until now, genetic analyses have not been applied to biparentally inherited sequence data and did not include data from all nine giraffe subspecies. We sampled natural giraffe populations from across their range in Africa, and for the first time individuals from the nominate subspecies, the Nubian giraffe, *Giraffa camelopardalis camelopardalis* Linnaeus 1758 [5], were included in a genetic analysis. Coalescence-based multi-locus and population genetic analyses identify at least four separate and monophyletic clades, which should be recognized as four distinct giraffe species under the genetic isolation criterion. Analyses of 190 individuals from maternal and biparental markers support these findings and further suggest subsuming Rothschild's giraffe into the Nubian giraffe, as well as Thornicroft's giraffe into the Masai giraffe [6]. A giraffe survey genome produced valuable data from microsatellites, mobile genetic elements, and accurate divergence time estimates. Our findings provide the most inclusive analysis of giraffe relationships to date and show that their genetic complexity has been underestimated, highlighting the need for greater conservation efforts for the world's tallest mammal.

RESULTS

Our nuclear and mitochondrial gene analyses clearly show that giraffes are not a homogeneous taxon but are deeply structured into distinct genetic groups. The degree of population genetic differentiation of seven nuclear markers from 105 individuals and concordance to mitochondrial (mt) data of 190 individuals suggests that some of the currently recognized subspecies are distinct species. Our nuclear dataset includes all currently recognized giraffe subspecies (Table 1; Figure 1) and, importantly, the elusive Nubian giraffe (*Giraffa c. camelopardalis*). A coalescent multi-locus (ML) tree analysis based on individual ML trees distinguishes four clusters: (1) a southern cluster comprising the South African and Angolan giraffe, (2) a Masai cluster corresponding to the Masai giraffe (including Thornicroft's giraffe [4, 6]), (3) the reticulated giraffe, and (4) a northern cluster including West African, Kordofan, and Nubian giraffe (Figure 2A). The monophyly of these groups is supported by $p > 0.95$ and is consistent with ML analysis of concatenated sequences.

Parsimony haplotype networks from the nuclear intron sequences revealed in general a similar pattern, reflecting the differentiation into four major clusters (Figure S1). Haplotype sharing between the Angolan and South African giraffe does not allow distinguishing between them; however, southern giraffe subspecies haplotypes are somewhat distinct. Two of the analyzed loci have exclusive haplotypes for the reticulated giraffe and the Masai giraffe (including Thornicroft's) that separate these from other giraffe subspecies. Intron 52 is characterized by a 33-bp-long insertion, a remnant of a DNA transposon, exclusive for the Masai giraffe.

The phylogenetic analysis of mtDNA from all nine giraffe subspecies (Figure 2B) produced a tree that conforms to previous analyses [3, 4], including the reciprocal monophyly of the seven distinct subspecies clades ($p > 0.95$). The lack of sequence variation shows that the Masai and Thornicroft's giraffe are not genetically distinct, supporting their grouping under the Masai giraffe [6]. Notably, mtDNA analyses also do not place the Nubian giraffe into a separate, monophyletic clade, but instead, two Nubian giraffe individuals group with the Kordofan giraffe, and three Nubian giraffe individuals group with Rothschild's giraffe (Figure 2B). Placing all Nubian giraffe individuals as a monophyletic sister group to either subspecies can be significantly rejected (approximately unbiased < 0.05). Specifically, the Nubian giraffe from Gambella National Park, Ethiopia, and two individuals from Bandingilo National Park, South Sudan, east of the Nile River, are placed in Rothschild's giraffe. In contrast, the Nubian giraffe individuals from northwest of

Table 1. Giraffe Subspecies, Distributions, and Population Sizes

Previous Scientific Name	Common Name	Principal Distribution	Population Size
G. c. peralta Thomas, 1898	West African giraffe	Niger	400
G. c. antiquorum (Jardine, 1835)	Kordofan giraffe	Central and Eastern Africa	2,000
G. c. camelopardalis (Linnaeus, 1758)	Nubian giraffe	South Sudan, Ethiopia	650
G. c. reticulata de Winton, 1899	reticulated giraffe	Kenya, Ethiopia, Somalia	8,660
G. c. rothschildi Lydekker, 1903[a]	Rothschild's giraffe	Uganda, Kenya	1,500
G. c. tippelskirchi Matschie, 1898	Masai giraffe	Kenya, Tanzania	32,000
G. c. thornicrofti Lydekker, 1911[a]	Thornicroft's giraffe	Zambia	550
G. c. angolensis Lydekker, 1903	Angolan giraffe	Botswana, Namibia	13,000
G. c. giraffa (von Schreber, 1784)	South African giraffe	Botswana, South Africa, Zambia, Zimbabwe	31,500

Giraffe (*Giraffa camelopardalis*) subspecies [2], common names, principal occurrences, and population sizes. Subspecies are listed in order of their approximate appearance from north to south. According to 2015 estimates by the Giraffe Conservation Foundation, the total giraffe population is ~ 90,000 individuals. See also Table S1. Data are from Giraffe Conservation Foundation, 2016; http://giraffeconservation.org.

[a]*G. c. rothschildi* and *G. c. thornicrofti* are now subsumed under *G. c. tippelskirchi* [6] and *G. c. camelopardalis* (this study), respectively.

Shambe National Park, west of the Nile River, South Sudan, group with the Kordofan giraffe (Figure 2B), despite varying pelage patterns (Figures 2C–2E).

Bayesian multi-locus clustering analysis shows that nuclear loci support four distinct groupings (Figure 3A). The highest ΔK [8] is observed for K = 4 clusters, with one cluster each corresponding to (1) southern giraffe (Angolan and South African), (2) the Masai giraffe (including Thornicroft's giraffe), (3) the reticulated giraffe, and (4) the northern giraffe (West African, Kordofan, Nubian, and Rothschild's giraffe). At K = 3, the northern cluster and reticulated giraffe are merged, while the other clusters remain distinct. Using K = 5 or higher values does not reveal additional clusters but rather shows increasing admixture. Analyzing the northern giraffe cluster separately shows that the West African giraffe is somewhat distinct but shares haplotypes with the Nubian giraffe. The southern giraffe cluster does not show further structuring (Figures S3A–S3C). Accordingly, principal component analyses (PCAs) of giraffe haplotypes find significant support for only four giraffe groups (Figure 3B). PCAs do not find support for additional groups according to mtDNA or traditional subspecies, or for a separate West African cluster (Figures S3D and S3E). The distinctness of these four clusters is in addition supported by significant fixation index (Fst) values and by Bayesian posterior probability (BPP) analyses of nuclear data receiving significant support for four clusters. BPP analyses that allow for additional clusters (e.g., West African giraffe being separate), clustering according to the mtDNA data, or six clusters [3] lack significant support ($p = 0.65$, $p = 0.32$, $p = 0.47$, respectively). Thus, population genetic, phylogenetic, and network analyses of nuclear sequences demonstrate that the giraffe is genetically well structured into four distinct species. This is consistent with divergence times of 1.25 to 2 million years ago (mya) among the four clusters (Figure 3C).

The survey genome assembly of a Kordofan giraffe produced 5,042 scaffolds > 10,000 bp. Repeat identification identified similar occurrences and relative numbers of short interspersed elements, endogenous retroviruses, and DNA transposons as in other Ruminantia [9], which may be suitable markers for conservation genetics. The genome assembly identified 2,239 protein-coding genes, of which 588 are orthologous to other mammals. After rigorous filtering, ≈540,000 bp remained for phylogenetic analyses, which places giraffe as sister group to cattle, antelope, and sheep and allows estimating the emergence of Giraffidae at 28.2 mya with high accuracy (Figure 3D), slightly longer ago than previously suggested [10]. Extracting microsatellites with >21 repeats identified 54 putatively informative loci.

DISCUSSION

The giraffe was first described in 1758 in Linnaeus' *Systema Naturae* [5]. As later revealed, Linnaeus based his description on the Nubian giraffe [11], corresponding to the nominotypical subspecies, *Giraffa camelopardalis camelopardalis*. Linnaeus had never seen a living giraffe and referred to 200-year-old descriptions [2]. Further descriptions of additional giraffe subspecies were later based on variable and taxonomically unreliable morphological traits, such as coat markings, ossicones, and geographic distribution [1, 2]. As an example, Thornicroft's giraffe from eastern Zambia was described as a distinct subspecies [12] but is morphologically similar to the Masai giraffe that occurs some 500 km to the north. However, genetic studies could not differentiate between the two subspecies, and Thornicroft's giraffe was therefore synonymized with the Masai giraffe [4, 6], as previously described [13]. Although the known geographical distribution of some giraffe subspecies remains uncertain, we genetically assigned individuals from the Sioma Ngwezi and Mosi-oa-Tunya National Parks in Zambia to the South African giraffe, in contrast to the previous assumption that they were Angolan giraffe.

Further north, the identification and classification of the nominotypical Nubian giraffe was even less certain [7]. The giraffe samples from west of the Nile River, which were assumed to be Nubian giraffe based on geography and morphology, turned out to share haplotypes with the Kordofan giraffe, whereas samples from east of the Nile River were genetically similar to Rothschild's giraffe. This is in agreement with the previous vague suggestion that South Sudan's giraffe populations from west of the Nile River (Figure 1) could be either Nubian, Rothschild's, or Kordofan giraffe [2, 7]. Thus, the putative Nubian giraffe samples that group with the Kordofan giraffe provide the first evidence that giraffes west of the Nile River actually belong to

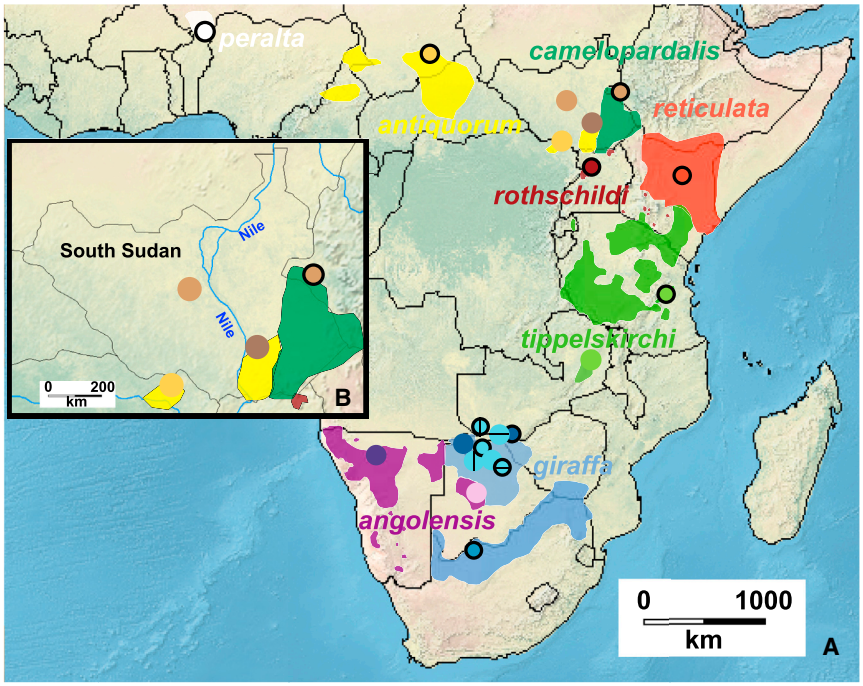


Figure 1. Distribution and Sampling Locations of Different Giraffe Subspecies in Africa

(A) Distribution ranges (colored shading) provided by the Giraffe Conservation Foundation [7], plotted on a map of Africa (http://www.naturalearthdata.com/). Circles represent sampling locations; for coding, see Figure 2.

(B) Enlarged view of the South Sudan region. Note that the samples of the putative Nubian giraffe were taken west and east of the Nile River.

See also Table S1.

Kordofan giraffe. Since the type locality of the Nubian giraffe had been previously restricted to "Sudan, Sennar" [11], east of the Nile River, it is clear that this name refers only to giraffe populations in this region. Yet, Rothschild's giraffe was also described from east, and further south, of the Nile River [1]; however, the Nubian and Rothschild's giraffe are genetically indistinguishable. For nomenclatorial priority reasons [14], Article 23, Rothschild's giraffe, *G. c. rothschildi* Lydekker, 1903 therefore needs to be synonymized with the earlier described Nubian giraffe, *G. c. camelopardalis* [5].

Numerous efforts have been made to define species, but a clear-cut consensus has not yet been reached [15, 16]. Common to many species concepts is that "species" represent distinct evolutionary units with limited gene flow to other, similar units, and concordance among different character sets has been suggested to support species distinctness [17]. In giraffe, multi-locus nuclear gene analyses, morphological data [2], mtDNA sequences [3, 4, 6, 18], and microsatellites [3] concordantly suggest genetically distinct groupings within giraffe. Concordance between maternally inherited mitochondrial and biparentally inherited nuclear markers indicates reproductive isolation for at least four giraffe groups. This lack of gene flow is unexpected, because wild giraffes are highly mobile [19] and can interbreed in captivity [20]. However, the genetic differentiation between the four giraffe groups is strong despite their similar appearance. Their divergence times are consistent with previous estimates [3, 4] and are on the order of divergence times of other mammals [21].

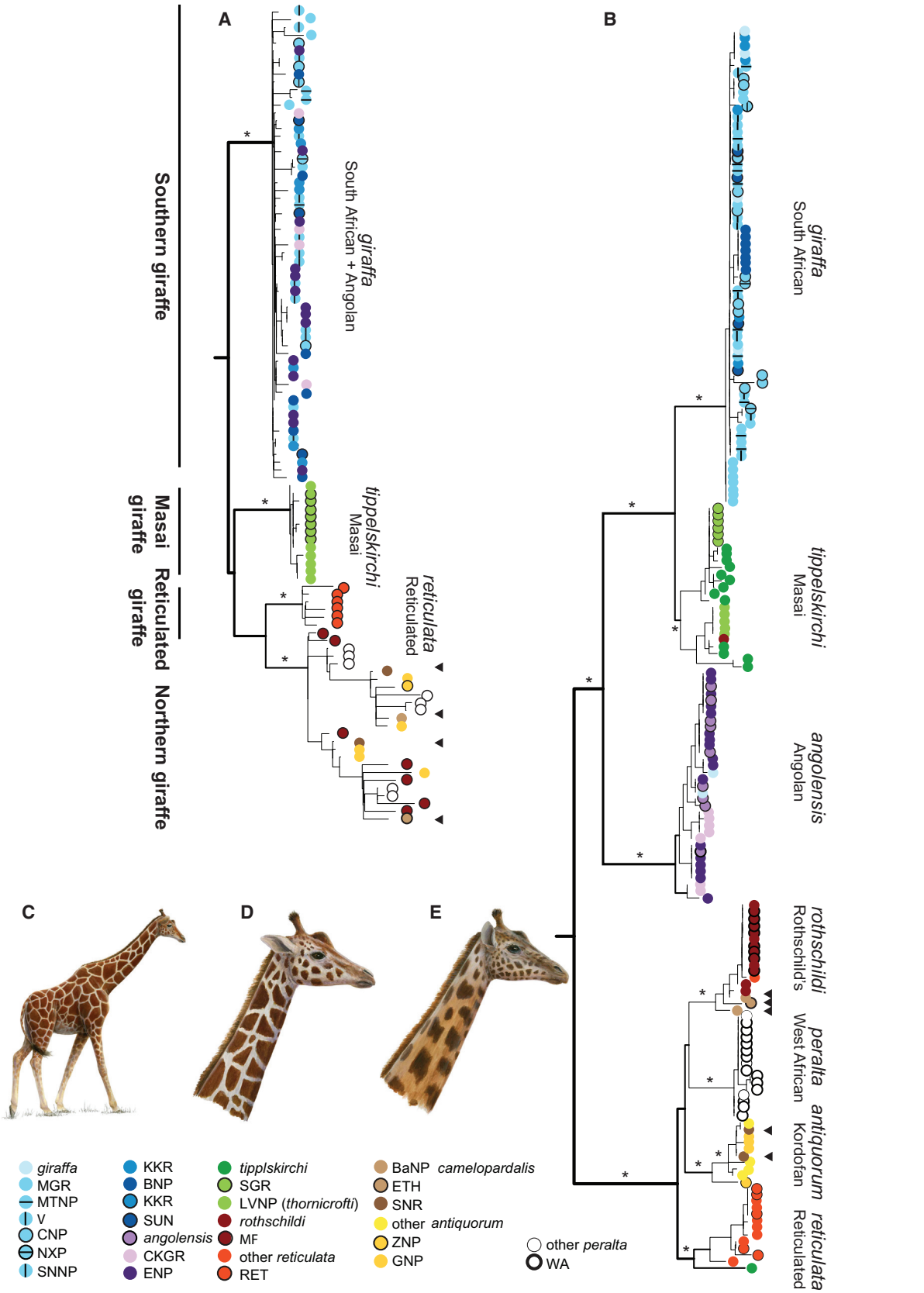
Although previous microsatellite analyses suggested the distinctness of six subspecies, with West African, Angolan, and South African giraffe being separate clusters [3], the statistical support is not clear. Our multi-locus coalescent-based analyses on sequence data allow for rigorous statistical testing and did not find support for such a grouping. Based on these data and analyses, and using the genealogical concordance method of phylogenetic species recognition [17] and fulfilling the requirements of the genetic species concept [22], we suggest recognizing four distinct giraffe species:

(1) southern giraffe (*G. giraffa*), comprising two distinct subspecies, Angolan giraffe (*G. g. angolensis*) and South African giraffe (*G. g. giraffa*);
(2) Masai giraffe (*G. tippelskirchi*), which includes the formerly recognized Thornicroft's giraffe;
(3) reticulated giraffe (*G. reticulata*); and
(4) northern giraffe (*G. camelopardalis*), which includes Nubian giraffe (*G. c. camelopardalis*) and its new synonym, Rothschild's giraffe (*G. c. rothschildi*), with Kordofan giraffe (*G. c. antiquorum*) and West African giraffe (*G. c. peralta*) as distinct subspecies. In the face of small population sizes, especially for the West African giraffe (Table 1), concerted conservation efforts are necessary for preserving these genetically differentiated subspecies.

Genome survey approaches have been successful in other species [9, 23, 24], including giraffe [25]. Our assembly and analysis of a Kordofan giraffe genome produced over 500,000 bp of protein coding sequence, numerous putatively informative microsatellites, and mobile element loci. Taken together with the recently published Masai giraffe genome [25], these markers are valuable for future analyses on conservation genomics. Thus, for two of the of the four giraffe species, there are now genomes available for conservation research.

Conclusions

For the first time, we have analyzed nuclear gene data from all formerly recognized giraffe subspecies, including the nominotypical Nubian giraffe, in a multi-locus analysis. Our findings demonstrate that most giraffe subspecies are composed of genetically divergent lineages. Two previously recognized subspecies (Thornicroft's and Rothschild's giraffe) turned out to be identical with the Masai giraffe and Nubian giraffe, respectively, and have been synonymized with these. The remaining former giraffe subspecies cluster genetically into four highly distinct groups, and we suggest that these should be recognized as discrete species. The conservation implications are obvious, as giraffe population numbers and habitats across Africa continue to dwindle due to human-induced threats.



(legend on next page)

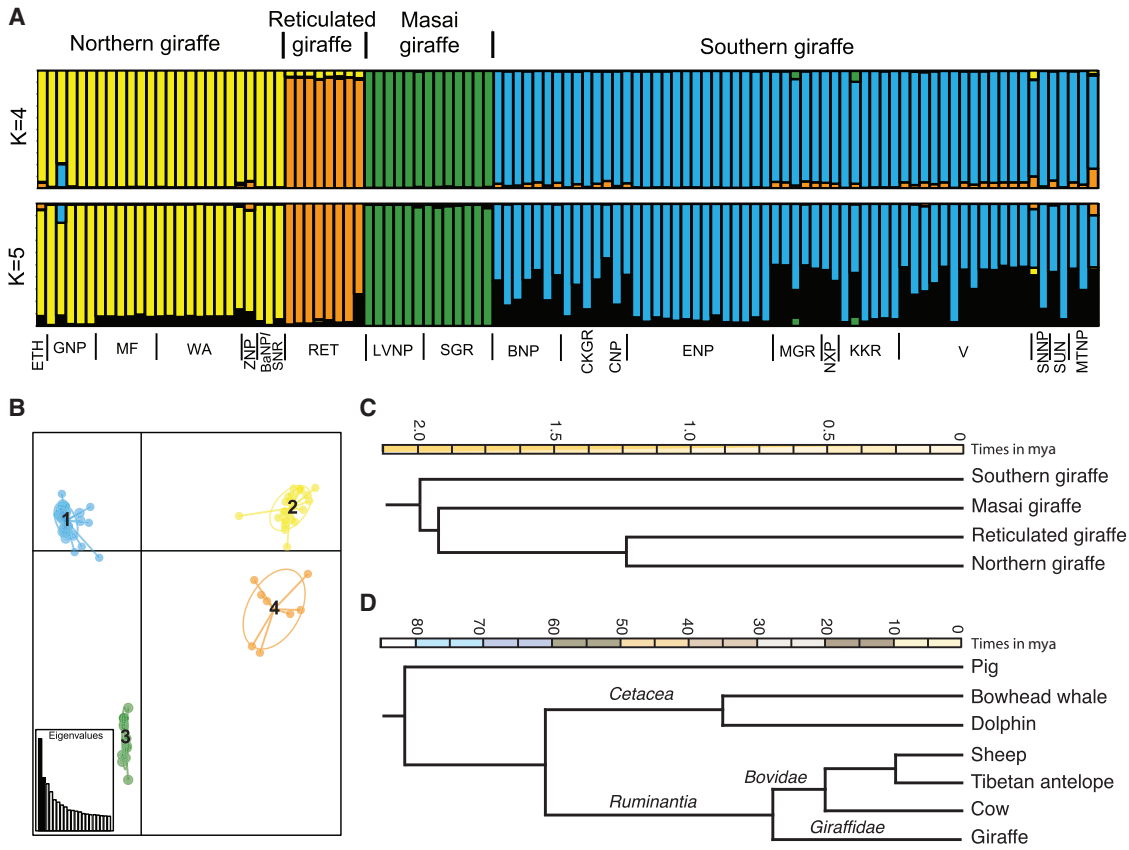


Figure 3. Population Structuring and Giraffe Divergence Times

(A) STRUCTURE analysis of seven nuclear loci for 105 individuals. Vertical bars show the membership in a cluster for each individual. Separate colors represent separate clusters. K = 4 has the highest credibility and shows well-resolved groups: blue: southern cluster (South African plus Angolan giraffe); green: Masai giraffe; orange: reticulated giraffe; yellow: northern cluster of the remaining subspecies. K = 5 or higher shows no further resolution.

(B) PCA axes 1–2 for four distinct giraffe clusters (1: southern; 2; northern; 3: Masai; 4: reticulated giraffe) according to STRUCTURE clusters (K = 4). The x axis explains 12.5% and the y axis 7.15% of variation. The oval outlines represent 95% confidential intervals and are colored after STRUCTURE clusters. Non-overlapping frames denote significantly different clusters. Analyses along axes 1–3 (data not shown) produced nearly identical results.

(C) Divergence times among giraffe species estimated by BEAST to 1.99, 1.89, and 1.25 mya, respectively.

(D) Time-calibrated phylogenomic analysis based on 540,000 bp of protein coding sequences. The divergence time of Giraffidae was estimated at 28.7 mya.

See also Figure S3 and Table S2.

EXPERIMENTAL PROCEDURES

Samples and DNA Extractions

Tissue samples from all currently recognized subspecies were collected by or through the Giraffe Conservation Foundation (GCF) using remote biopsy darts from 141 wild individual giraffes with country-specific research permits and permission between 2009 and 2015. The geographical distribution and sampling locations are shown in Figure 1. For individuals' IDs and geographic origin, see Table S1 and Figure S2. DNA was extracted using a Machery-Nagel NucleoSpin Tissue Kit. Genomic DNA for low-coverage genome sequencing was prepared from a Kordofan giraffe (ZNP01) using a standard phenol-chloroform extraction method. The genome was sequenced to 10× coverage for paired-end libraries at the Beijing Genome Institute using Illumina technology.

Amplification and Sequencing of Intron and Mitochondrial Markers

Mitochondrial cytochrome *b* and control region sequences were PCR amplified and sequenced as described previously [4]. Nucleotide substitutions and insertion/deletions conformed to previous observations [4]; therefore,

Figure 2. Evolutionary Relationships among Giraffe

(A) A coalescent multi-locus tree from seven nuclear loci (4,294 bp) from okapi and 105 giraffe individuals identifies four monophyletic clades with significant support, p > 0.95: southern giraffe (*G. giraffa*, *G. angolensis*), Masai giraffe (*G. tippelskirchi*), reticulated giraffe (*G. reticulata*), and northern giraffe (*G. antiquorum*, *G. camelopardalis*, *G. peralta*, *G. rothschildi*). This grouping is consistent with STRUCTURE, PCA, and BPP analyses (Figures 3A and 3B). Asterisks indicate statistically significant support (p > 0.95). Arrowheads indicate Nubian giraffe individuals. Individuals without geographic ID are from databases.

(B) mtDNA BEAST tree for 190 individuals. Except for the Nubian giraffe (*G. c. camelopardalis*), all seven subspecies are well separated and form monophyletic groups. The Masai giraffe and Thornicroft's giraffe are subsumed into one subspecies, *G. c. tippelskirchi*. Misplaced individuals of the reticulated, Rothschild's, and South African giraffe are from databases and are likely to represent misidentifications [3]. The MTNP (Mosi-oa-Tunya National Park) and SNNP (Sioma Ngwezi National Park) giraffes are likely South African giraffes.

(C–E) Drawings showing distinctive coat markings of the reticulated giraffe (C), Kordofan giraffe (D), and Nubian giraffe (E).

See also Table S1 and Figure S2.

nuclear mitochondrial sequences are not suspected. Intron markers were developed from an alignment of sheep (*Ovis aries)* and cattle (*Bos taurus*) genomes [26, 27]. Intron markers were selected for analyses if they were ≈800 bp long, contained three or more variable sites, had invariable adjacent exons, and could be PCR amplified in giraffe and okapi samples. PCRs were performed using 10–15 ng genomic DNA with primers placed in adjacent exons (Supplemental Experimental Procedures). Each PCR setup contained negative controls and was inspected using agarose gel electrophoresis. The sequences were analyzed on an Applied Biosystems ABI 3730 instrument.

Tree Analyses

The mitochondrial sequences and published sequences were aligned and trimmed [4] and analyzed with BEAST v1.7.5 [28] under HKY+G+I as suggested by jModelTest v2.1.1 [29] with default priors, 2×10^9 generations, and sampling every 20,000th iteration. Convergence was analyzed in Tracer [30]. TreeAnnotator was used to construct a maximum clade credibility tree after discarding 10% as burn-in. ML trees were reconstructed by RAxML v8.2.4 [31] using the GTR+G+I model. Approximately unbiased (AU) statistics [32] using CONSEL v1.20 [33] were used to evaluate alternative phylogenies. Sequence data were generated from forward and reverse sequences and were edited and aligned using Geneious v5.6.4 (Biomatters). Heterozygous indels were resolved both by eye and with the help of Indelligent [34] and verified by allele-specific primers when necessary. All sequences were trimmed for gaps, missing information, and ambiguous sequences. Trees were generated under the GTR+G+I model of sequence evolution. From individual ML trees, a coalescent species tree was constructed in Astral [35] with default parameters, and branch lengths were calculated from sequences using RAxML. Divergence times were estimated using MCMCtree in PAML v4 [36] on four fossil-based references (Table S2). The okapi was the outgroup in all analyses.

Population Genetic Analysis

Haplotypes of nuclear sequences were deduced by PHASE in DnaSP v5.0 [37] using a threshold of 0.6 and allowing for recombination. Parsimony haplotype networks were inferred from TCS v1.21 [38], with a connection probability of 0.95. The Bayesian clustering algorithm implemented in STRUCTURE v2.3.4 [39] was used to infer admixture. The haplotype information deduced by PHASE was used to code individuals. We sampled 40,000 steps following a burn-in of 10,000 steps, for K = 1–10, with 20 replicates each. The results were averaged using CLUMPP [40]. Structure Harvester [41] was used to infer the most likely K. To assess the degree of similarity between a priori defined populations, PCAs were performed with the R package adegenet [42] for R v3.2.3. Bayesian phylogenetic and phylogeographic analyses for species delineation were performed in BPP v3.2 [43] using algorithm A11, 1,000,000 generations, and a burn-in of 10,000 with gamma priors of $\theta \sim 2$, 2000 and $\tau \sim 2$, 2000. Convergence was checked by repeated analyses and using different guide trees (four giraffe species, West African or Angolan giraffe separate [3], or eight mtDNA clusters). Fixation index (Fst) values were calculated on nuclear haplotypes using Arlequin v3.5.2.1 [44] (see Supplemental Experimental Procedures).

De Novo Genome Survey Assembly

Genomic DNA sequences were generated as 125-bp paired-end reads on one Illumina lane from a 250-bp insert size library. Raw reads were trimmed with Trimmomatic [45] using a minimum base quality of 20 bp and a minimum read length of 75 bp. SoapDenovo [46] assembled the reads on odd *k*-mers from 45 to 69 with default parameters. The assembly with *k*-mer 51 produced the largest number of scaffolds with a minimum sequence length of 10 kb and expected assembly statistics. Annotation, assembly statistics, and phylogenetic analysis are described in Supplemental Experimental Procedures.

Repeat and Microsatellite Sequences in the Giraffe Genome

We selected assembled scaffolds longer than 5,000 bp. Removal of duplicate sequence clustering was done using CD-HIT-454 v4.5.4 [47], followed by repeat masking [48] using mobile genetic elements from Cetartiodactyla. The percentage of repeat types was calculated with a custom script. Putatively informative microsatellites were identified with SciRoKo v3.4 [49], and Primer 3-designed primers [50] to optimally amplify microsatellite loci (see Supplemental Experimental Procedures).

ACCESSION NUMBERS

The accession numbers for the sequences reported in this study are European Nucleotide Archive: LT596685–LT598170, and the study accession number is European Nucleotide Archive: PRJEB12634.

SUPPLEMENTAL INFORMATION

Supplemental Information includes three figures, two tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2016.07.036.

AUTHOR CONTRIBUTIONS

A.J. and J.F. conceived, designed, and financed the project. J.F. and P.E. collected samples. T.B., F.R., V.K., M.V., and M.A.N. produced and analyzed the data. U.F. aided the systematic interpretation. A.J., J.F., and U.F. wrote the manuscript with input from all authors.

ACKNOWLEDGMENTS

We thank Kathinka Schulze, Clara Heumann-Kiesler, Sven Winter, and Gabriele Krenzer-Misch for support with analyses and editing. This study was supported by the State of Hesse's funding program LOEWE, the Leibniz Association, the Giraffe Conservation Foundation, the Leiden Conservation Foundation, the Auckland Zoo, and various African government partners and international supporters.

Received: May 6, 2016
Revised: June 14, 2016
Accepted: July 14, 2016
Published: September 8, 2016

REFERENCES

1. Lydekker, R. (1904). On the subspecies of *Giraffa camelopardalis*. Proc. Zool. Soc. Lond. *74*, 202–229.
2. Dagg, A.I., and Foster, J.B. (1976). The Giraffe: Its Biology, Behaviour and Ecology (Van Nostrand Reinhold).
3. Brown, D.M., Brenneman, R.A., Koepfli, K.-P., Pollinger, J.P., Milá, B., Georgiadis, N.J., Louis, E.E., Jr., Grether, G.F., Jacobs, D.K., and Wayne, R.K. (2007). Extensive population genetic structure in the giraffe. BMC Biol. *5*, 57.
4. Bock, F., Fennessy, J., Bidon, T., Tutchings, A., Marais, A., Deacon, F., and Janke, A. (2014). Mitochondrial sequences reveal a clear separation between Angolan and South African giraffe along a cryptic rift valley. BMC Evol. Biol. *14*, 219.
5. Linnaeus, C. (1758). Systema Naturae per Regna Tria Naturae, Secundum Classes, Ordines, Genera, Species, cum Characteribus, Differentiis, Synonymis, Locis (Tomus I, Editio Decima) (Laurentius Salvius).
6. Fennessy, J., Bock, F., Tutchings, A., Brenneman, R., and Janke, A. (2013). Mitochondrial DNA analyses show that Zambia's South Luangwa Valley giraffe (*Giraffa camelopardalis thornicrofti*) are genetically isolated. Afr. J. Ecol. *51*, 635–640.
7. Krumbiegel, I. (1971). Die Giraffe (A. Ziemsen Verlag).
8. Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. *14*, 2611–2620.
9. Gallus, S., Kumar, V., Bertelsen, M.F., Janke, A., and Nilsson, M.A. (2015). A genome survey sequencing of the Java mouse deer (*Tragulus javanicus*) adds new aspects to the evolution of lineage specific retrotransposons in Ruminantia (Cetartiodactyla). Gene *571*, 271–278.
10. Hassanin, A., Delsuc, F., Ropiquet, A., Hammer, C., Jansen van Vuuren, B., Matthee, C., Ruiz-Garcia, M., Catzeflis, F., Areskoug, V., Nguyen, T.T., and Couloux, A. (2012). Pattern and timing of diversification of

Cetartiodactyla (Mammalia, Laurasiatheria), as revealed by a comprehensive analysis of mitochondrial genomes. C. R. Biol. *335*, 32–50.

11. Harper, F. (1940). The nomenclature and type localities of certain Old World mammals. J. Mammal. *21*, 322–332.
12. Lydekker, R. (1911). Two undescribed giraffes. Nature *87*, 484.
13. Matschie, P. (1898). Einige anscheinend noch nicht beschriebene Säugetiere aus Afrika. Sitzungsber. Ges. Naturforsch. Freunde Berlin *1898*, 75–81.
14. ICZN (1999). International Commission on Zoological Nomenclature: International Code of Zoological Nomenclature (International Trust for Zoological Nomenclature).
15. Coyne, J.A., and Orr, H.A. (2004). Speciation (Sinauer).
16. De Queiroz, K. (2007). Species concepts and species delimitation. Syst. Biol. *56*, 879–886.
17. Avise, J.C., and Ball, R.M., Jr. (1990). Principles of genealogical concordance in species concepts and biological taxonomy. In Oxford Surveys in Evolutionary Biology, *Volume 7*, D. Futuyama, and J. Antonovics, eds. (Oxford University Press), pp. 45–67.
18. Hassanin, A., Ropiquet, A., Gourmand, A.-L., Chardonnet, B., and Rigoulet, J. (2007). Mitochondrial DNA variability in *Giraffa camelopardalis*: consequences for taxonomy, phylogeography and conservation of giraffes in West and central Africa. C. R. Biol. *330*, 265–274.
19. Flanagan, S.E., Brown, M.B., Fennessy, J., and Bolger, D.T. (2016). Use of home range behaviour to assess establishment in translocated giraffes. Afr. J. Ecol. Published online April 21, 2016. http://dx.doi.org/10.1111/aje.12299.
20. Lackey, L. (2009). Giraffe Studbook *Giraffa camelopardalis* North American Regional/Global (International Species Information System).
21. Hailer, F., Kutschera, V.E., Hallström, B.M., Klassert, D., Fain, S.R., Leonard, J.A., Arnason, U., and Janke, A. (2012). Nuclear genomic sequences reveal that polar bears are an old and distinct bear lineage. Science *336*, 344–347.
22. Baker, R.J., and Bradley, R.D. (2006). Speciation in mammals and the genetic species concept. J. Mammal. *87*, 643–662.
23. Wernersson, R., Schierup, M.H., Jørgensen, F.G., Gorodkin, J., Panitz, F., Staerfeldt, H.-H., Christensen, O.F., Mailund, T., Hornshøj, H., Klein, A., et al. (2005). Pigs in sequence space: a 0.66X coverage pig genome survey based on shotgun sequencing. BMC Genomics *6*, 70.
24. Kuramoto, T., Nishihara, H., Watanabe, M., and Okada, N. (2015). Determining the Position of Storks on the Phylogenetic Tree of Waterbirds by Retroposon Insertion Analysis. Genome Biol. Evol. *7*, 3180–3189.
25. Agaba, M., Ishengoma, E., Miller, W.C., McGrath, B.C., Hudson, C.N., Bedoya Reina, O.C., Ratan, A., Burhans, R., Chikhi, R., Medvedev, P., et al. (2016). Giraffe genome sequence reveals clues to its unique morphology and physiology. Nat. Commun. *7*, 11519.
26. Elsik, C.G., Tellam, R.L., Worley, K.C., Gibbs, R.A., Muzny, D.M., Weinstock, G.M., Adelson, D.L., Eichler, E.E., Elnitski, L., Guigó, R., et al.; Bovine Genome Sequencing and Analysis Consortium (2009). The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science *324*, 522–528.
27. Archibald, A.L., Cockett, N.E., Dalrymple, B.P., Faraut, T., Kijas, J.W., Maddox, J.F., McEwan, J.C., Hutton Oddy, V., Raadsma, H.W., Wade, C., et al.; International Sheep Genomics Consortium (2010). The sheep genome reference sequence: a work in progress. Anim. Genet. *41*, 449–453.
28. Drummond, A.J., Suchard, M.A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. *29*, 1969–1973.
29. Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods *9*, 772.
30. Rambaut, A., Suchard, M., Xie, D., and Drummond, A. (2014). Tracer v1.6. http://beast.bio.ed.ac.uk/Tracer.
31. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313.
32. Shimodaira, H. (2002). An approximately unbiased test of phylogenetic tree selection. Syst. Biol. *51*, 492–508.
33. Shimodaira, H., and Hasegawa, M. (2001). CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics *17*, 1246–1247.
34. Dmitriev, D.A., and Rakitov, R.A. (2008). Decoding of superimposed traces produced by direct sequencing of heterozygous indels. PLoS Comput. Biol. *4*, e1000113.
35. Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S., and Warnow, T. (2014). ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics *30*, i541–i548.
36. Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. *24*, 1586–1591.
37. Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics *25*, 1451–1452.
38. Clement, M., Posada, D., and Crandall, K.A. (2000). TCS: a computer program to estimate gene genealogies. Mol. Ecol. *9*, 1657–1659.
39. Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics *155*, 945–959.
40. Jakobsson, M., and Rosenberg, N.A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics *23*, 1801–1806.
41. Earl, D.A., and vonHoldt, B.M. (2011). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. *4*, 359–361.
42. Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics *24*, 1403–1405.
43. Yang, Z. (2015). The BPP program for species tree estimation and species delimitation. Curr. Zool. *61*, 5.
44. Excoffier, L., and Lischer, H.E.L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. *10*, 564–567.
45. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.
46. Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., et al. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascience *1*, 18.
47. Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics *22*, 1658–1659.
48. Smit, A., Hubley, R., and Green, P. (2015). RepeatMasker Open-4.0, 2013-2015. http://www.repeatmasker.org.
49. Kofler, R., Schlötterer, C., and Lelley, T. (2007). SciRoKo: a new tool for whole genome microsatellite search and investigation. Bioinformatics *23*, 1683–1685.
50. Rozen, S., and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. *132*, 365–386.

Current Biology, Volume *26*

Supplemental Information

Multi-locus Analyses Reveal Four Giraffe Species Instead of One

Julian Fennessy, Tobias Bidon, Friederike Reuss, Vikas Kumar, Paul Elkan, Maria A. Nilsson, Melita Vamberger, Uwe Fritz, and Axel Janke

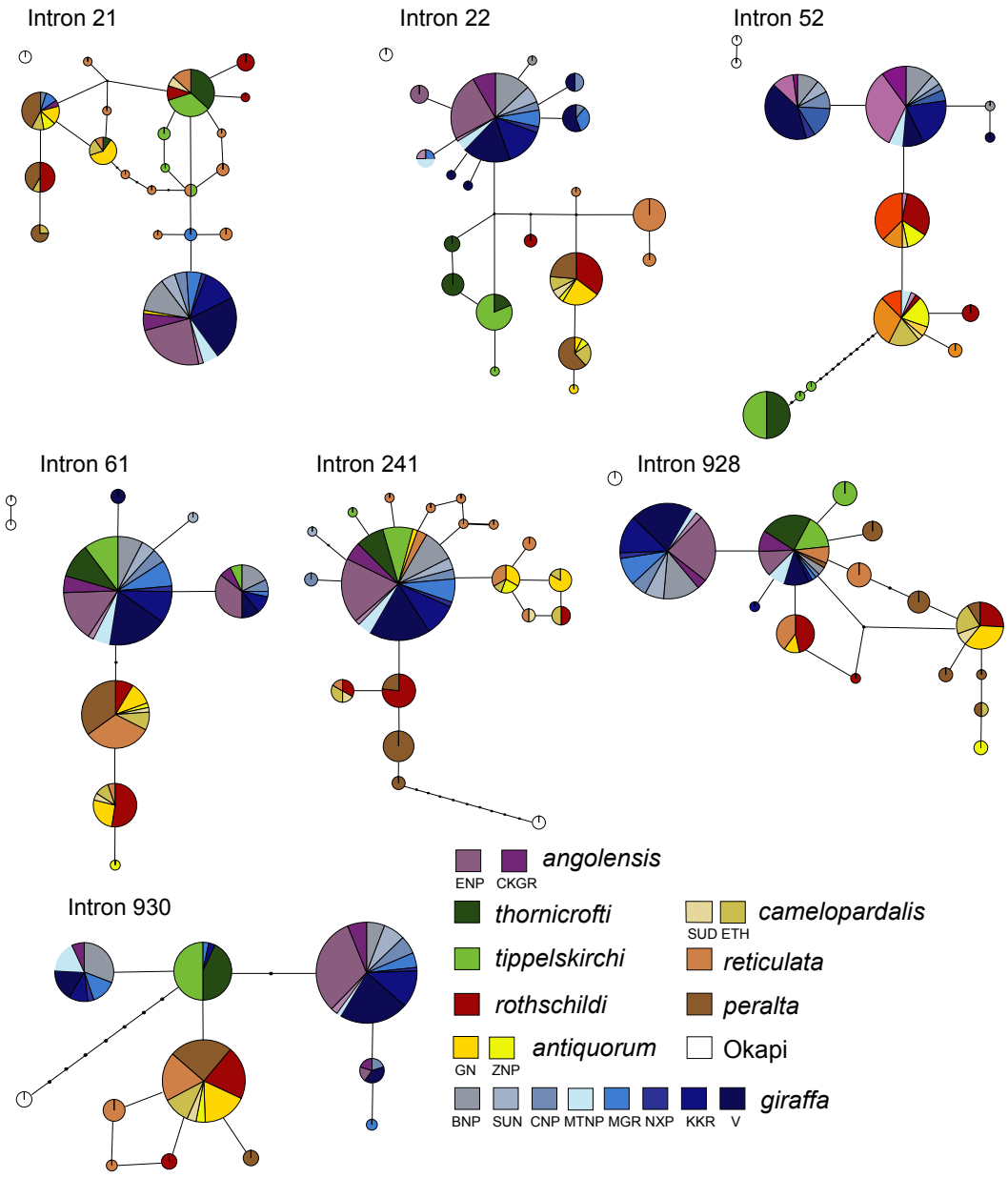


Figure S1. Haplotype networks of seven intron sequences. Related to Figure 2 and 3. The networks for 105 giraffe show that numerous single alleles are shared and that for most loci subspecies cannot be easily distinguished. Notable exceptions are intron 52 and 930 that are exclusive for Masai (*tippelskirchi*, including the formerly recognized Thornicroft's giraffe) and 241 that is nearly exclusive for the West African giraffe (*peralta*). Furthermore, a southern clade (*angolensis* plus *giraffa)* and northern clade (*antiquorum*, *peralta*, *rothschildi*) are prominent for most loci. The lack of further resolution is possibly a consequence of insufficient data and gene flow remains uncertain.

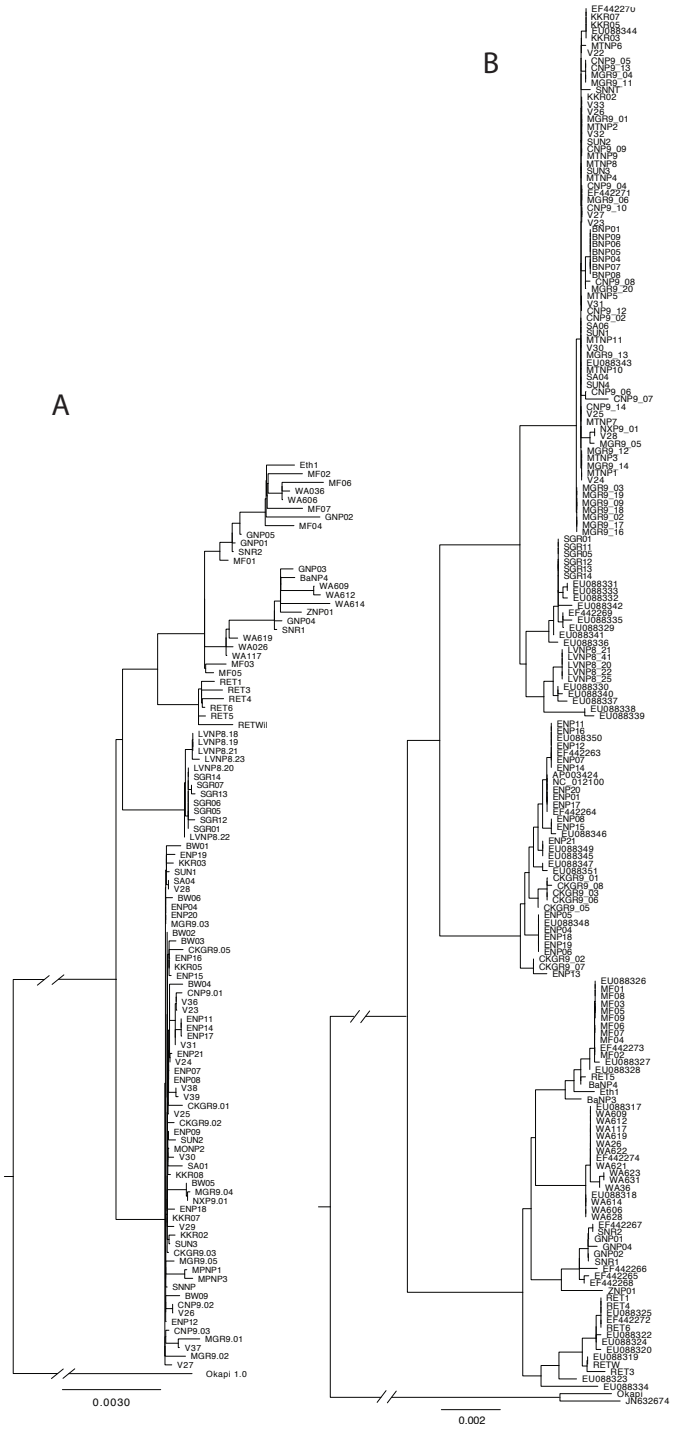


Figure S2. Evolutionary trees with details on the individual IDs. Related to Figure 2. A) ASTRAL tree from individual nuclear loci with ML branch lengths. While analysis of concatenated sequences is problematic [S1] a ML tree based on concatenated sequences, which shows the West African Giraffe separate can be found in doi:10.5061/dryad.h3tc2. B) BEAST mtDNA tree with details for accession numbers and individual IDs and their location. Note – the okapi branch (root) is not to scale in both figures to allow for better resolution among giraffe branches. Genbank accession numbers for published data are shown and are detailed in [S2] and individual ID and sample location can be found in doi:10.5061/dryad.h3tc2.

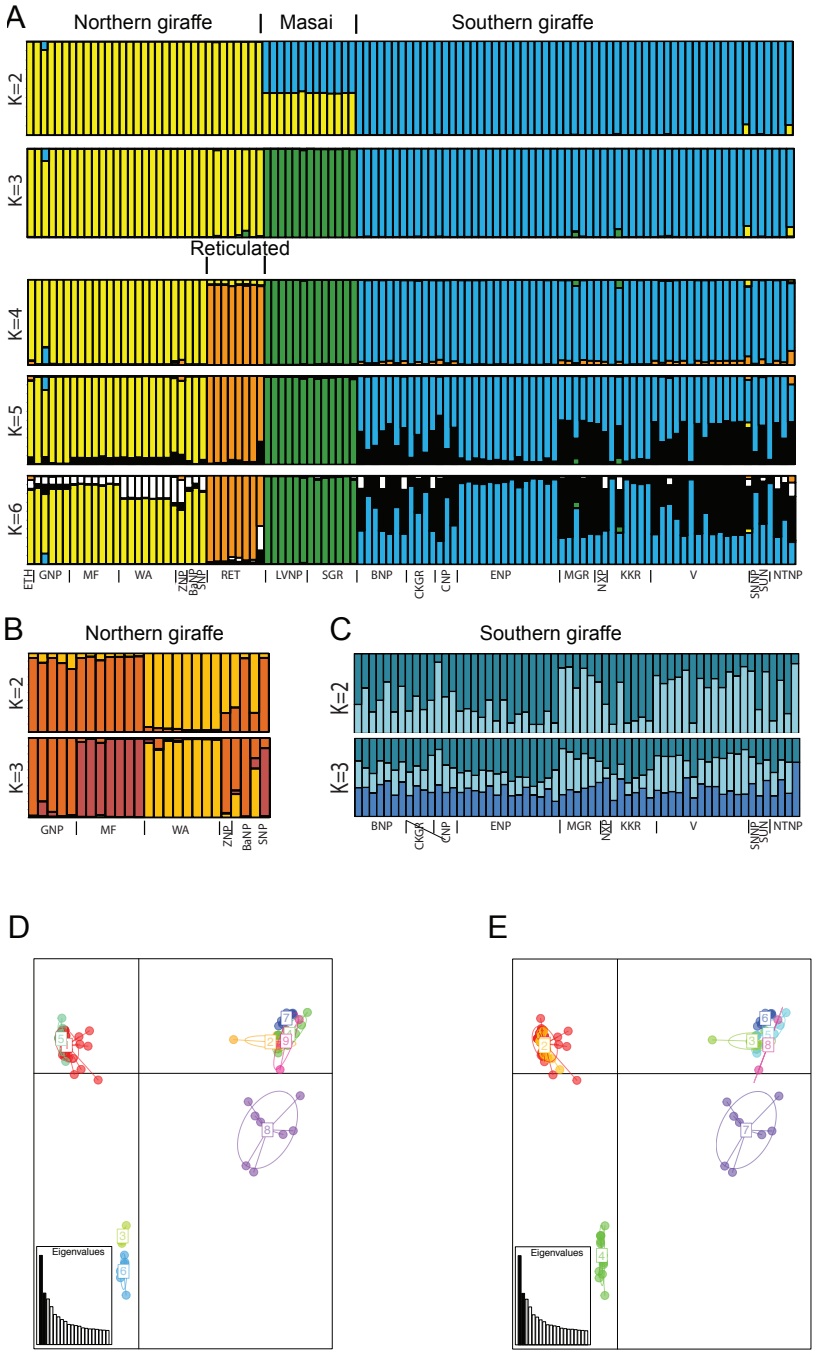


Figure S3. Additional Structure and PCA analyses. Related to Figure 3.
A) Structure analysis for all subspecies for K=2 to K=6. K=4 has the highest delta K and from K=5 increasing admixtures is evident.
B) A separate Structure analysis for the subspecies of the northern giraffe reveals evidence for additional cluster of West African (WA) and Nubian (former Rothschild's, MF) giraffe. However, in this data set haplotypes sharing with other subspecies is evident and these are not distinct in other analyses.
C) Southern giraffe do not show additional clustering when analyzed separately. This is in contrast the the clear separation of the subspecies by mtDNA sequences.
Abbreviations for the geographic origin are explained in Table S1.
D) PCAs of giraffe haplotypes with grouping according to traditional nine subspecies classification 1 - *G. c. angolensis,* 2 - *G. c. antiquorum,* 3 – *G. c. thornicrofti,* 4 - *G. c.*

rothschildi, 5 - *G. c. giraffa,* 6 - *G. c. tippelskirchi*, 7 - *G. c. peralta*, 8 - *G. c. reticulata*, 9 - *G. c. camelopardalis)*.
PCAs along Axis 1-3 (not shown) produces nearly identical results. In this analyses four distinct clades are evident and these correspond to the four giraffe species suggested by other analyses, other subspecies are overlapping.
E) PCAs of giraffe haplotypes with grouping according to mtDNA differentiation 1- *G. c. giraffa,* 2 - *G. c. angolensis*, 3 - *G. c. antiquorum*, 4 - *G. c. tippelskirchi*, 5 - *G. c. rothschildi*, 6 - *G. c. peralta*, 7 - *G. c. reticulata*, 8 - *G. c. camelopardalis*). In this analyses four distinct clades are evident and these correspond to the four giraffe species suggested by other analyses, other subspecies are overlapping.

Table S1. Origin, abbreviation, number of individuals (n) and traditional subspecies designation of analyzed giraffe sequences. Related to Table 1, Figure 1 and Figure 2.

Geographical origin	**Abbreviation**	**n**	**Previous subspecies designation**	**MtDNA subspecies (this study)**
Badingilo National Park, South Sudan	BaNP	2	*camelopardalis*	*camelopardalis*
Bwabwata National Park, Namibia	BNP	7	*angolensis*	*giraffa*
Chobe National Park, Botswana	CNP	11	*angolensis*	*giraffa*
Gambella National, Ethiopia	ETH	1	*camelopardalis*	*camelopardalis*
Etosha National Park, Namibia	ENP	17	*angolensis*	*angolensis*
Garamba National Park, DRC	GNP	3	*antiquorum*	*antiquorum*
Khamab Kalahari Reserve, South Africa	KKR	6	*giraffa*	*giraffa*
Luangwa Valley National Park, Zambia	LVNP	5	*thornicrofti*	*tippelskirchi*
Moremi Game Reserve, Botswana	MGR	16	*angolensis*	*giraffa*
Mosi-oa-Tunya National Park, Zambia	MTNP	11	*angolensis*	*giraffa*
Murchison Falls National Park, Uganda	MF	9	*rothschildi*	*camelopardalis*
Nxai Pans, Botswana	NXP	1	*angolensis*	*giraffa*
Nuernberg/Stuttart Zoo	RET	6	*reticulata*	*reticulata*
Selous Game Reserve, Tanzania	SGR	6	*tippelskirchi*	*tippelskirchi*
Shambe National Park, South Sudan	SNR	2	*camelopardalis*	*antiquorum*
Sioma Ngwezi NP, Zambia	SNNP	1	*angolensis*	*giraffa*
Sun hotel, Livingstone, Zambia	SUN	4	*giraffa*	*giraffa*
Koure, Niger	WA	13	*peralta*	*peralta*
Vumbura Concession, Botswana	V	11	*angolensis*	*giraffa*
Zakouma National Park, Chad	ZNP	1	*antiquorum*	*antiquorum*

Table S2. Detailed divergences time estimates. Related to Figure 3.
A: Estimated divergence times and confidence intervals.

Divergence	Estimated divergence time (Ma)
Southern giraffe – (Masai giraffe, (reticulated giraffe, northern giraffe)	2.00 (1.23-3.12)
Masai giraffe – (reticulated giraffe, northern giraffe)	1.87 (1.16-2.90)
Reticulated giraffe – northern giraffe	1.25 (0.74-1.97)

Note – divergence times were estimated based on a divergence of okapi to 11.5 million years ago (Ma) using BEAST and nuclear loci. For more information see: doi:10.5061/dryad.h3tc2.

B: Estimated divergence times and confidence intervals.

Species pair	Estimated divergence time
Whale-Dolphin	34.85 (34.0-36.0)
Sheep- Antelope	10.85 (7.1-14.6)
(Sheep,Antelope)-Cow	20.0 (18.0-22.0)
(Bovidae)-Giraffe	28.2 (22.8-35.2)
(Ruminantia,(Cetacea))-Pig	61.91 (54.2-66.1)
Cetartiodactyla-Dog	83.04 (67.8-98.1)
(Cetartiodactyla,Dog)-Human	90.85 (74.7-103.1)

Note – on four fossil based, independent divergences times included: whale-dolphin divergence at 34-36 Ma [S3], cattle-antelope 18-22 Ma [S4], divergence time of pig – remaining Artiodactyla at 52.4-65.8 Ma [S5], and Carnivora (dog)-Artiodactyla 62.5-131.0 [S5]. The analysis was for a sample size of 200,000, burn-in of 20,000 and sampling trees from every second iteration.

Supplemental Experimental Procedures

Coding sequences were extracted from scaffolds >10,00 bp. For annotation the scaffolds were repeat-masked using Repeat masker [S6]. From these scaffolds, genes were predicted by AUGUSTUS [S7] and coding sequences were extracted using the perl script provided with AUGUSTUS. For phylogenetic analysis orthology searches were made using using the recursive Blast method [S8] from nine genomes: human (*Homo sapiens)*, dog (*Canis familiaris*), pig (*Sus scrofa*), cattle (*Bos taurus*), Tibetan antelope (*Pantholops hodgsonii*), sheep (*Ovis aries*), bottlenose dolphin (*Tursiops truncates*), and bowhead whale (*Balaena mysticetus*) available from the Ensembl database (http://www.ensembl.org/info/data/ftp/index.html) and from the bow head whale (*Balaena mysticetus*) genome (http://alfred.liv.ac.uk/downloads/bowhead_whale/bowhead_whale_coding_sequences.zip) and Tibetan antelope (*Pantholops hodgsonii*) from ftp://climb.genomics.cn/pub/10.5524/100001_101000/100027/. Only the orthologous groups with data from seven species, including giraffe, were selected for further phylogenetic analysis. The selected orthologous sequences were translated using custom perl script and aligned using MAFFT [9]. PAL2NAL [S10] generated a nucleotide alignment from the aligned amino acids sequences and alignment-gaps were removed by the program GBLOCKS [S11]. Alignments with more than 25% observed nucleotide distance among any species were removed to avoid artifacts from multiple substitutions and unrecognized alignment artifacts. Only coding sequences larger than 180 bp were

selected for further phylogenetic analyses. The best evolutionary model was predicted by JMODELTEST. RAxML version 8.2.4 [S12] was used to reconstruct a maximum likelihood (ML) tree using the GTR+G+I model.

The giraffe divergence time was estimated using MCMC tree in PAML version 4 [S13]. The molecular clock was calibrated on four fossil based, independent divergences times: whale-dolphin divergence at 34-36 Mya [S3], cattle-antelope 18-22 Mya [S4], divergence time of pig – remaining Artiodactyla at 52.4-65.8 Mya [S5], and Carnivora (dog)-Artiodactyla 62.5-131.0 [S5]. The analysis was for a sample size of 200,000, burn-in of 20,000 and sampling trees from every second iteration.

Analysis on 18 giraffe individuals covering all the major subspecies of giraffe were used to estimate the divergence time using BEAST on seven nuclear loci. Corresponding Okapi nuclear loci were Sanger sequenced and ortholog nuclear loci of cow (*Bos taurus*) genome were fetched from UCSC https://genome.ucsc.edu/ database. All the sequences were aligned using MAFFT [9]. Later BEAST was run with the settings of 100 million generations, HKY+I+G model, log normal relaxed clock and tree prior of Yule process. We used the molecular calibration point of 11.5 Ma [14] with standard deviation of 0.5 for the okapi and giraffe split with cow as the outgroup. Convergence was checked and confirmed with Tracer [S15].

Additional analysis using the Bayesian program, BPP version 3.2 was done to test the delimitation of different species in giraffe [S16 –18]. First method (algorithm) A00 was run to estimate the alpha and beta parameters. Slight deviations from alpha=2 and beta=2000 for θ and τ turned out to be non-crucial to the analyses. Other parameters were left to default. We used the method (algorithm) A11 to search various species delimitation models and different species phylogeny. The sequences from seven nucleotide loci were clustered into four groups of southern, Masai giraffe, Reticulate and northern giraffe according to the multi-locus coalescent tree and clades suggested by Structure and PCA analyses. In addition, the probability of five species (*G. c. peralta* being separate) the classic and mtDNA grouping was calculated. Each analysis was run with 1,000,000 generations and a burn-in of 10,000 with gamma (G) priors of $\theta \approx (2, 2000)$ and $\tau \approx (2, 2000)$. Convergence was checked by repeated analysis and with different guide trees.

The search for the microsatellites using SCIROKO version 3.4 [S19] identified useful microsatellites for future studies for Giraffe is shown in doi:10.5061/dryad.h3tc2. In addition, the repeat masking of the selected Scaffolds >5kb after removal of duplicates, identified different types of repeat elements similar to other ruminants such as mouse deer, sheep, Tibetan antelope and cow shown in doi:10.5061/dryad.h3tc2.

***De novo* assembly statistics for the 10X coverage giraffe genome assembly from 125bp paired-end Illumina reads.**

Total number of raw reads	549,679,536
Quality trimmed reads	464,882,128
Total size include N after assembly	2,432,441,945
Total size without N after assembly	2,397,009,050
Total no. of scaffold	30,24,215
Mean Size	804
Median Size	244
Longest Seq	35,995
Shortest Seq	100
Scaffolds>=10kb	5,042
N50	2,201

Note – additional information is provided in doi:10.5061/dryad.h3tc2.

List of Primer sequences and PCR conditions for amplification of nuclear introns and mtDNA in giraffe.

Name & locus	Sequence 5'-3'	PCR conditions
Intron 21 RASSF4, 9th intron	for: CAGTGTCCATCACACAAC rev: GCACCGGCATTTCAAACTTA	TD-PCR (T_a=65-55°C; 10 cycles), standard PCR (T_a=55°C; 30 cycles)
Intron 22 ACP5, 6th intron	for: CAGCAGCCAAGGAGGACTAC rev: ATCTCCTTGGGGCTGATCTC	TD-PCR (T_a=67-57°C); 10 cycles), standard PCR (T_a=57°C; 30 cycles)
Intron 52 UBN2, 4th intron	for: ACTGGCACTCTCCAGTTTCG rev: CTTCCTCTTTCCGCTTCCTC	INT_{UBN2} Bock et al. 2014b
Intron 52_140825 UBN2, 4th intron	for: GACAACCAAAAGCACAAACC rev: CACTTACCCCAGTTGTTTGG	TD-PCR (T_a=69-62°C; 14 cycles), standard PCR (T_a=62°C; 26cycles)
Intron 61 CWF19L1, 9th intron	for: GCTGGGAGGAAGGTAGCAATG rev: AATGTTGACCACCAAATGC	TD-PCR (T_a=66-59°C); 14 cycles), standard PCR (T_a= 59°C; 26 cycles)
Intron 241 NUP155, 23rd intron	for: GCTGCTGTTGATGGCATTAG rev: GGTCCACCTGATTGCTGATT	See Intron 61
Intron 928 OTOF, 12th intron	for: GCAGAGCACCAGTTCCA rev: GCTCGGTAGATCTTCACGTAG	INT_{OTOF} Bock et al. 2014b
Intron 930 SOS1, 11th intron	for: CAAAGTCCAAAGCACCCTG rev: CATGTTACTTCCTCCTTGCTTG	TD-PCR (T_a=67-60°C; 14 cycles), standard PCR (T_a=60°C; 26 cycles)
Control Region, mtDNA	F: TACACTGGTCTTGTAAGC R: TCGCTTTGGTGTTTAAGC	Bock et al. 2014b
Cytochrome b, mtDNA	F: GAAAAACCATCGTTGTCG R: TGGGAGTATATTAATAGC	Bock et al. 2014b

Note – for: forward primer. rev: reverse primer. TD-PCR: touchdown PCR. T_a: primer annealing temperature. The locus is the gene name of the human ortholog and the respective intron.

List of Fst values for seven nuclear loci of four giraffe species.

	Southern	Northern	Masai	Reticulated
Southern	0			
Northern	0.559**	0		
Masai	0.591**	0.522**	0	
Reticulated	0.608**	0.273**	0.595**	0

Note – Double asterisks indicate all Fst values are significant at $p<0.05$. Southern giraffe (*G. giraffa*) comprises the historic Angolan (*G. c. angolensis*) and South African giraffe (*G. c. giraffa*). Northern giraffe (*G. camelopardalis*) includes the historic Nubian giraffe (*G. c. camelopardalis*), Rothschild's giraffe (*G. c. rothschildi)*, Kordofan giraffe (*G. c. antiquorum*) and West African giraffe (*G. c. peralta*). Masai giraffe (*G. tippelskirchi*) includes historic Thornicroft's giraffe (*G. c. thornicrofti*). Reticulated giraffe (*G. reticulata*) includes only itself.

References

S1. Edwards SV, Xi Z, Janke A, Faircloth BC, McCormack JE, Glenn TC, Zhong B, Wu S, Lemmon EM, Lemmon AR, et al. (2016). Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. Mol. Phylogenet. and Evol.. *231*:447-62.

S2. Bock, F., Fennessy, J., Bidon, T., Tutchings, A., Marais, A., Deacon, F., and Janke, A. (2014). Mitochondrial sequences reveal a clear separation between Angolan and South

African giraffe along a cryptic rift valley. BMC Evol. Biol. *14*. 219
S3. de Muizon, C. (2009). L'origine et l'histoire évolutive des Cétacés. Comptes Rendus Palevol 8, 295–309.
S4. Vrba, E., and Schaller, G. (2000). Antelopes, deer, and relatives. fossil record, behavioral ecology, systematics and conservation (Yale University Press, New Haven, London).
S5. Benton, M., Benton, P., and Asher, R. (2009). Calibrating and constraining molecular clocks. In The Timetree of Life, Hedges SB, Kumar S. ed. (Oxford University Press).
S6. Smit, A., Hubley, R., and Green, P. (2015). RepeatMakser Open-4.0.2013-2015. Available from http://www.repeatmasker.org.
S7. Stanke, M., Schöffmann, O., Morgenstern, B., and Waack, S. (2006). Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. BMC Bioinformatics *7*, 62.
S8. Hallström, B.M., and Janke, A. (2008). Resolution among major placental mammal interordinal relationships with genome data imply that speciation influenced their earliest radiations. BMC Evol. Biol. *8*, 162.
S9. Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. *30*, 3059–3066.
S10. Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. *34*, W609-612.
S11. Talavera, G., and Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst. Biol. *56*, 564–577.
S12. Stamatakis, A. (2014). RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. Bioinformatics *30*,1312-1313.
S13. Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. *24*, 1586–1591.
S14. Agaba M, Ishengoma E, Miller WC, McGrath BC, et al. (2016). Giraffe genome sequence reveals clues to its unique morphology and physiology. Nat. Commun. 7, 11519.
S15. Rambaut, A., Suchard, M., Xie, D., and Drummond, A. (2014). Tracer v1.6, Available from http://beast.bio.ed.ac.uk/Tracer.
S16. Rannala B, Yang Z. (2003). Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. Genetics *164*, 1645–1656.
S17. Yang Z, Rannala B. (2010). Bayesian species delimitation using multilocus sequence data. Proc. Natl. Acad. Sci. U. S. A. *107*, 9264–9269.
S18. Yang Z. (2015). The BPP program for species tree estimation and species delimitation. Curr. Zool. *61*, 854-865.
S19. Kofler, R., Schlötterer, C., and Lelley, T. (2007). SciRoKo: a new tool for whole genome microsatellite search and investigation. Bioinformatics *23*, 1683–1685.