

RESEARCH ARTICLE

# The mitochondrial ancestor of bonobos and the origin of their major haplogroups

Hiroyuki Takemoto<sup>1\*</sup>, Yoshi Kawamoto<sup>1</sup>, Shoko Higuchi<sup>1†</sup>, Emiko Makinose<sup>1</sup>, John A. Hart<sup>2</sup>, Tereese B. Hart<sup>2</sup>, Tetsuya Sakamaki<sup>1</sup>, Nahoko Tokuyama<sup>1</sup>, Gay E. Reinartz<sup>3</sup>, Patrick Guislain<sup>3</sup>, Jef Dupain<sup>4</sup>, Amy K. Cobden<sup>5</sup>, Mbangi N. Mulavwa<sup>6</sup>, Kumugo Yangozene<sup>6†</sup>, Serge Darroze<sup>7</sup>, Céline Devos<sup>8</sup>, Takeshi Furuichi<sup>1\*</sup>

**1** Primate Research Institute, Kyoto University, Inuyama, Japan, **2** Luku Foundation, Projet Tshuapa-Lomami-Lualaba (TL2), Kinshasa, Democratic Republic of Congo, **3** Bonobo and Congo Biodiversity Initiative, Zoological Society of Milwaukee, Milwaukee, Wisconsin, United States of America, **4** African Wildlife Foundation, Nairobi, Kenya, **5** Department of Anthropology, Emory University, Atlanta, Georgia, United States of America, **6** Research Center for Ecology and Forestry, Ministry of high Education and Scientific Research, Mabali, Democratic Republic of Congo, **7** Consultant Biodiversity, Sustainable Use of Natural Resources, Protected Areas Management and Adaptation to Climate Change, Bangkok, Thailand, **8** Department of Behavioral Biology, University of Liège, Liège, Belgium

☯ These authors contributed equally to this work.

† Deceased.

✉ Current address: Nagoya Head Office, HASHIMOTO UNYU Co., Ltd, Nagoya, Japan

\* [takemoto.hiroyuki.7e@kyoto-u.ac.jp](mailto:takemoto.hiroyuki.7e@kyoto-u.ac.jp) (HT); [furuichi.takeshi.7m@kyoto-u.ac.jp](mailto:furuichi.takeshi.7m@kyoto-u.ac.jp) (TF)



OPEN ACCESS

**Citation:** Takemoto H, Kawamoto Y, Higuchi S, Makinose E, Hart JA, Hart TB, et al. (2017) The mitochondrial ancestor of bonobos and the origin of their major haplogroups. PLoS ONE 12(5): e0174851. <https://doi.org/10.1371/journal.pone.0174851>

**Editor:** David Caramelli, University of Florence, ITALY

**Received:** November 18, 2016

**Accepted:** March 16, 2017

**Published:** May 3, 2017

**Copyright:** © 2017 Takemoto et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are available from the DNA Data Bank of Japan (DDBJ), accession numbers LC213801-LC213807.

**Funding:** This research was supported by the Ministry of the Environment, Japan (D-1007) (to TF); the Global Environment Research Fund (F-061) of the Ministry of the Environment, Japan (to Nishida); Japan Society for the Promotion of Science (JSPS) Grants-in-aid for Scientific Research (17255005 and 22255007, to TF); the JSPS Asia-Africa Science Platform Program

## Abstract

We report here where the most recent common ancestor (MRCA) of bonobos (*Pan paniscus*) ranged and how they dispersed throughout their current habitat. Mitochondrial DNA (mtDNA) molecular dating to analyze the time to MRCA (TMRCA) and the major mtDNA haplogroups of wild bonobos were performed using new estimations of divergence time of bonobos from other *Pan* species to investigate the dispersal routes of bonobos over the forest area of the Congo River's left bank. The TMRCA of bonobos was estimated to be 0.64 or 0.95 million years ago (Ma). Six major haplogroups had very old origins of 0.38 Ma or older. The reconstruction of the ancestral area revealed the mitochondrial ancestor of the bonobo populations ranged in the eastern area of the current bonobos' habitat. The haplogroups may have been formed from either the riparian forests along the Congo River or the center of the southern Congo Basin. Fragmentation of the forest refugia during the cooler periods may have greatly affected the formation of the genetic structure of bonobo populations.

## Introduction

Bonobos (*Pan paniscus*) range in the forest area of the southern Congo Basin, the left bank of the Congo River (Fig 1). Their divergence from chimpanzees (*Pan troglodytes*) has been estimated to be 0.8–2.1 million years ago (Ma) from genetic studies [1]. Previous genome studies on two *Pan* species suggested that bonobos have clearly been separated from chimpanzees [2,3]. However the recent genome analysis indicated the possibility that some gene flow occurred between two species during the late Pleistocene [4]. In any case, we need to examine

(2009–2011, 2012–2014, 2014–2017 to TF); the JSPS Grant-in-Aid for Challenging Exploratory Research (19657074 to HT); the JSPS Primate Origins of Human Evolution (HOPE) Project (#8 to Primate Research Institute, Kyoto University); US Fish and Wildlife (2007–2009, 2010–2011, to Lukuru Wildlife Research Foundation); Arcus Foundation (2007–2008, 2010, 2012, to Lukuru Wildlife Research Foundation); Wood Tiger Foundation (2010, 2012, to Lukuru Wildlife Research Foundation); Full Circle Foundation (2011, to Lukuru Wildlife Research Foundation); the US Fish and Wildlife Service Assistance Award (96200-0-G017, to African Wildlife Foundation); Arcus Foundation, the Columbus Zoo and Aquarium, and Barron Wall (to AWF); the L.L. Leakey Foundation (to AKC); the Congo Basin Forest Partnership/USAID-WWF (623-A-00-06-00072-00, to ZSM); Susan McConnell and Richard Scheller (Wildlife Conservation Network) (to ZSM); the AZA Ape TAG Conservation Initiative (to ZSM); WWF Belgium, WWF Netherlands, Belgian Cooperation funding for co-financing; and CARPE program (USAID) for program coordination, for WWF bonobo project in the Lac Tumba Landscape. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

when and how the internal genetic clades of bonobos had branched off from each other, in order to investigate the history of bonobos.

Six major mtDNA haplogroups have been found in wild bonobos to date. Eriksson et al. [5] reported two haplogroups (haplogroups A and B) in five wild bonobo populations and suggested that the Lomami River acted as a riverine barrier that affected the divergence of the two mtDNA clades. Furthermore, Zsurka et al. [6] reported a third haplogroup (haplogroup C) in captive bonobos. Next, Kawamoto et al. [7] distinguished six haplogroups (A1, A2, B1, B2, C, and D) in seven wild bonobo populations. Their analysis of the relationship between the genetic distance ( $F_{ST}$ ) and the geographical distance among the seven populations was similar to the analysis performed by Eriksson et al. [5] and indicated the linear geographical distances to show a significant correlation in the genetic distances among the populations. These results may suggest that the current riverine barrier, with the exception of the Lomami River, had a weak effect on the genetic structure of bonobos.

Kawamoto et al [7] also indicated that the east cohort (The area between the Lomami River and the Lualaba River in TL2) showed genetically closer proximity to the west cohort (Malebo and Lac Tumba) than to the central cohort (Lomako, Salonga, Wamba and Iyondji). Some haplogroups were found only within a certain cohort. For example, the A1 and C haplogroups were found in the central cohort, whereas the D haplogroup was found in the east cohort. On the other hand, the B1 haplogroup was found in two distant cohorts in the east and west, whereas the A2 and B2 haplogroups were found to be part of the two neighboring west and central cohorts. These genetic relationships may mean that the bonobo populations were not formed simply due to vicariance or dispersion. Therefore, paleoenvironment information in the Congo Basin must be reviewed, especially historical changes in the environment and the location of forest refugia during dry periods of the Pleistocene, for better understanding of the genetic structure of bonobos.

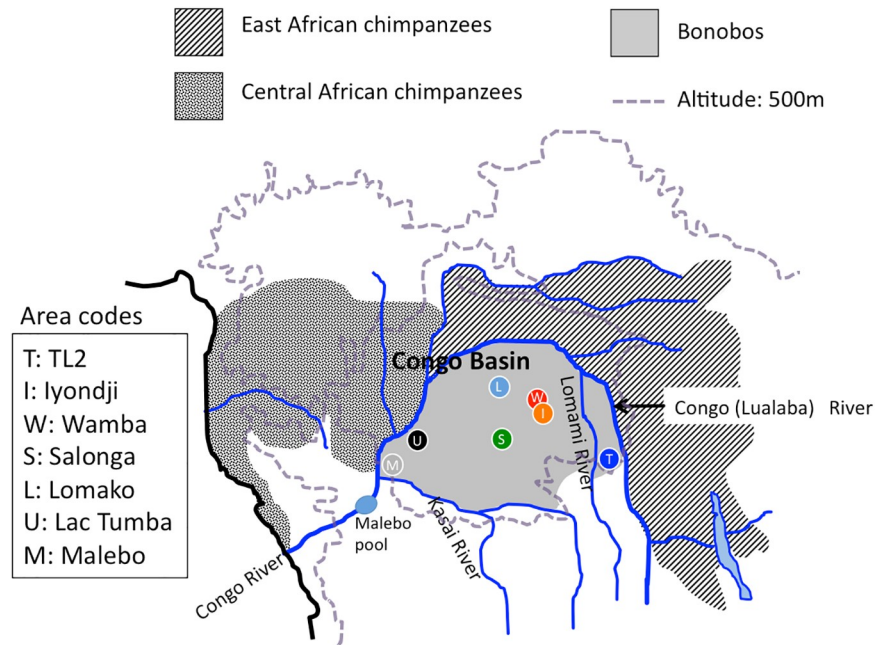
We conducted the historical analysis on the distribution of mtDNA haplogroups in bonobos. Questions to be solved were reasons for: 1) closer genetic proximity between the west and east cohort, 2) distant distribution of the B1 haplogroup, and 3) a weak effect of the current rivers on the genetic structure of bonobos. We first explored the ages and regions of the most recent common ancestor (MRCA) of bonobos and the MRCA of each haplogroup, and estimated the location of the forest refugia to be the left bank of Congo River in the Congo Basin. Then, we hypothesized the dispersal routes of bonobos after they entered the left bank of the Congo River [1].

## Methods

### DNA samples and sequencing

DNA sampling and methods for sequence determinations were previously described by Kawamoto et al. [7]. DNA samples were extracted from fecal samples, which were collected non-invasively from seven wild populations, i.e. TL2 (by HT, JH and TH), Iyondji (by TS), Wamba (by HT and NT), Salonga (by GN and PG), Lomako (by HT, JD and AC), Lac Tumba (by HT, MM and KY) and Malebo (by HT, SD and CD), that covered most of the eastern limit to the western limit of the current range of bonobos (Fig 1). We collected fecal samples from beneath of the nests or trees from which bonobos had already left, thus we never interacted or interfered with bonobos. Eighteen new fecal samples from Salonga were added to the data set of Kawamoto et al. [7] in this study.

The Scientific Authority for CITES and National Scientific Committee for the Worldwide Heritage of UNESCO in DRC has confirmed that we can publish the results obtained from the fecal samples of bonobos carried out from DRC as far as we have research permission that



**Fig 1. Distribution of *Pan* species in central Africa and sampling areas for mtDNA of bonobos.** Sampling populations are expressed as small circles with area codes.

<https://doi.org/10.1371/journal.pone.0174851.g001>

includes the permission to use those samples. Research Permissions during this study were issued by following authorities: 024/ICCN/BP-MA/2010 (for TL2), 051/ICCN/DG/ADG/KV/2011 (for Lomako), 1577/ICCN/ADG/ANG/DG/2008 (for Salonga) were given by the Institut Congolais pour la Conservation de la Nature (ICCN). MIN.RS/SG/002/2010 (for Iyondji), MIN.RS/SG/003/2010 (for Wamba), 008//MINRS/CREF/MAB/DG/01MNIK/2011 (for Lac Tumba) were given by Ministère de la Recherche Scientifique (MIN). 001/CREF/2012 (for Malebo) was given by the Centre de Recherche en Ecologie et Foresterie (CREF).

Obtained sequence data were deposited in DDBJ/EMBL/GenBank databases (Accession Numbers LC213801-LC213807). Sorting via gBlocks of 1117 nt sites from 154 fecal samples were identified. Haplotypes were identified by multiple alignments with ClustalX vr. 2.1 [8] of the sorted sequences.

Molecular phylogenetic relations for haplotypes were inferred using the neighbor joining (NJ) method. The mtDNA diversity within populations was estimated in terms of haplotype (gene) diversity, mean number of pairwise difference, and nucleotide diversity using the program Arlequin (v 3.5) [9]. Genetic differentiation between populations was quantified from calculations of intra- and interpopulation distances with pairwise  $F_{ST}$  distance [10] and average pairwise difference [11].

### Dating nodes of the tree

In order to estimate the time to MRCA (TMRCA) for each haplogroup, we added 44 mtDNA haplotypes of chimpanzees (*Pan troglodytes schweinfurthii*: 18, *P. t. troglodytes*: 8, *P. t. eliotti*: 7, *P. t. verus*: 11) and those of humans (9) from GenBank to the bonobos' haplotypes (61 haplotypes). These sequence data were then sorted using gBlocks, resulting in 1073 bp aligned sequence. Each node was dated in the phylogenetic tree using the Bayesian MCMC method BEAST (2.2.0) with the HKY and Relaxed Clock Log Normal model [12,13]. The root prior was set at 6–8 Ma for Human/*Pan* divergence, and 0.95–1.05 Ma (estimation 1) and 1.7–1.8

Ma (estimation 2) for bonobo chimpanzee divergence as uniform distributions, according to the estimations from various evidences [1]. The MCMC process was performed 50,000,000 times. The phylogenetic trees generated via TreeAnnotator (v2.2) yielded the same topology for the main haplogroups that was shown by the NJ methods. These results were detected by FigTree (v. 1.4.2) (<http://tree.bio.ed.ac.uk/software/figtree/>).

## Evolutionary analysis

Inferring the ancestral areas for the MRCA of bonobos and evolutionary events were performed using RASP (v3.0.2) [14]. We used two different models for RASP. S-DIVA is the statistical version of DIVA (Divergence-Vicariance Analysis) [15], which is based on the parsimony of the cost-event such as dispersion or extinction. However, there is some controversy in the use of DIVA analysis for inferring population histories [16]. We limited the maximum areas for one haplotype to three areas according to the assertion of Ronquist and Sanmartín [17], because two haplotypes (PPCR31 and PPCR32) were found from three sites in the central region of A2 clade (Iyondji, Wamba, and Lomako) (Fig 2). The 50,001 MCMC trees by BEAST were used as inputs in RASP, and 50 discard trees and 500 random trees for S-DIVA were set. The Bayesian Binary MCMC (BBM) method was also conducted in a similar way. The maximum area was kept to 3, with 10 chains, and 50,000 MCMC generations and JC models.

## Investigating the history of bonobos in the southern Congo Basin

First, we reviewed the reports on the paleoenvironment during the middle-to-late Pleistocene in the Congo Basin. The results of the molecular dating, the ancestral area reconstruction, and the paleoenvironment on the southern Congo Basin were combined to construct the forest areas of the southern Congo Basin during the cooler periods. From these results, we inferred how the genetic structure of bonobos was formed; namely, the history of bonobos after divergence from other *Pan* populations.

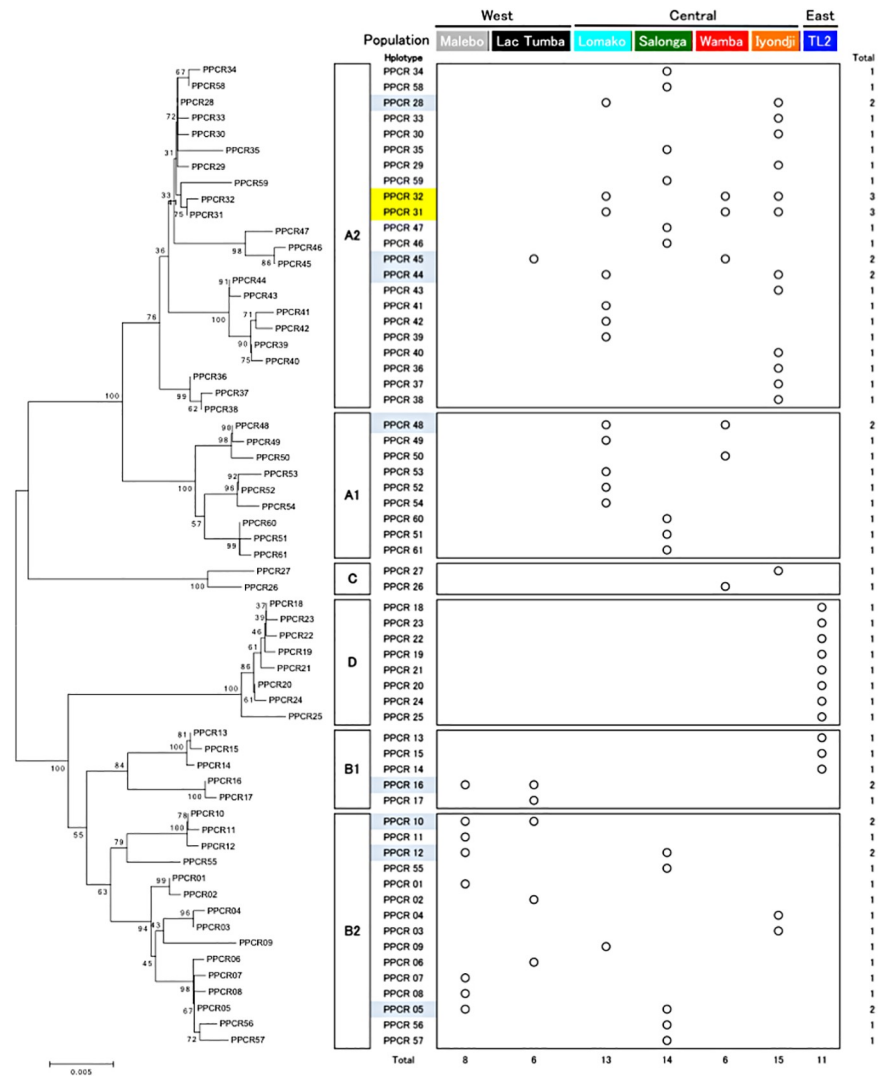
## Results

### Phylogenetic tree and haplogroup confirmation

We detected six haplogroups (A1, A2, B1, B2, C, and D) consisting of 61 mtDNA haplotypes, including 7 new haplotypes from Salonga, in the bonobo populations (Fig 2). Three cohorts (east: TL2; central: Iyondji, Wamba, Salonga, Lomako; west: Lac Tumba, Malebo) were revealed via clustering analyses and UPGMA method using  $F_{ST}$  distances (Fig 3). The  $F_{ST}$  distance among three cohorts indicated that the east cohort was genetically nearer to the west cohort than to the central cohort, which was incongruent with the geographical distances. Lac Tumba, Lomako, Salonga and Iyondji populations showed higher nucleotide diversity than the other populations (Tables 1 and 2).

### The TMRCA of bonobos and major haplogroups in bonobo populations

The TMRCA of bonobos was calculated at 0.64 Ma (0.50–0.79 with 95% HPD) by estimation 1 [0.95 Ma (0.70–1.22 Ma) by estimation 2; hereafter, a number in the square parenthesis means the age by estimation 2] (Table 3). All haplogroups diverged 0.38 Ma [0.53 Ma] or older. TMRCA of B1 haplogroup, which ranged distinctly between the west and east regions, was estimated at around 0.26 Ma [0.36 Ma]. Although a higher mutation rate of the branch was observed after divergence from chimpanzees to the MRCA of bonobos in estimation 2, there were no remarkable differences in the mutation rate among each haplogroup.

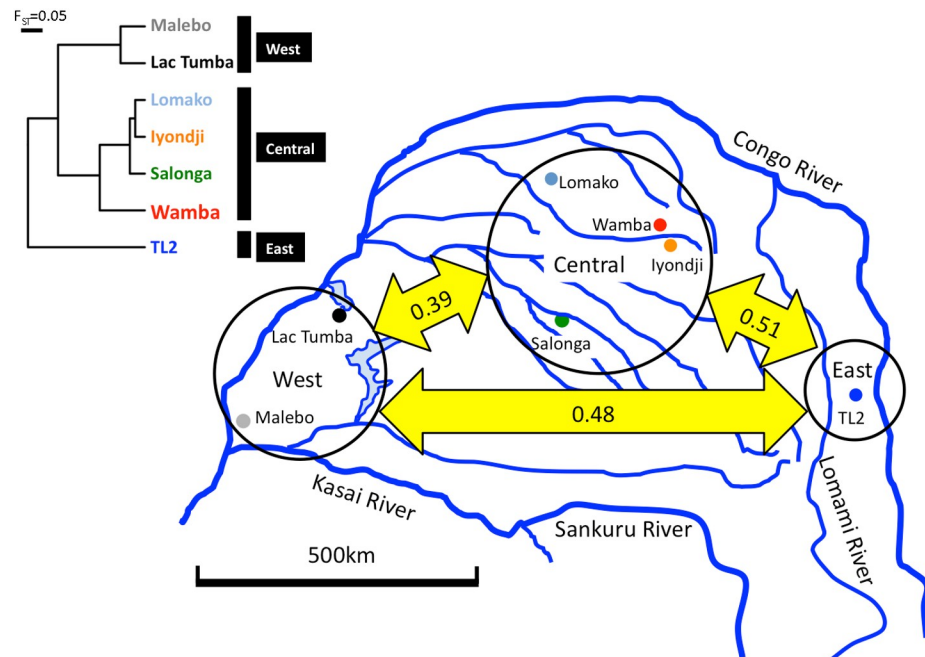


**Fig 2. Molecular phylogeny of haplotypes and their distribution in seven populations.** On the left is the phylogenetic tree constructed using the NJ method. The right side indicates the population where each haplotype was found. Blue shaded (two populations) and yellow shaded (three populations) haplotypes were found in multiple populations.

<https://doi.org/10.1371/journal.pone.0174851.g002>

### Ancestral areas and dispersal routes

Ancestral area reconstructions using the two estimations portrayed similar results (Figs 4 and 5 and Table 4). In the S-DIVA model, the ancestral range for the MRCA of the collected mtDNA samples was indicated over the TL2 and Iyondji areas, and it was separated into those two areas by a vicariance event. The A1, A2, and C haplogroups originated in the Iyondji area. The A1 and A2 haplogroups differentiated in the central region such as Iyondji or Lomako. On the other hand, the ancestral area for the B and D haplogroups was TL2. The D haplogroup was formed within TL2. The B haplogroups enlarged their range to the Malebo or Lac Tumba areas. The B1 haplogroup ranged over the TL2 and western region, and separated into east and west regions later. The ancestral range of the B2 haplogroup was estimated to be Malebo. In the BBM model, four of five major nodes indicated the plural possibility of ancestral area in



**Fig 3. Genetic distances among three cohorts.** Three circles mean east, central and west cohort respectively. Numeric values in double-headed arrows indicate  $F_{ST}$  distances among three cohorts calculated by UPGMA (the left tree).

<https://doi.org/10.1371/journal.pone.0174851.g003>

both estimations (Table 4). However, the area indicated by S-DIVA showed the highest probability in most cases (Figs 4 and 5).

Some haplotypes of B2 ranged the central region, but the central region was not indicated as a main dispersal route for the B2 haplogroup except for BBM in estimation 2. Similarly, although one haplotype (PPRC45) of A2 ranged in Lac Tumba (western region), this haplotype was concluded to be the result of recent gene flow from Salonga in the central region.

## Discussion

### Inferring the history of bonobos in the southern Congo Basin

Although previous genome studies suggested no gene flow between bonobos and any of the chimpanzee groups [2,3], some gene flow from chimpanzees to bonobos might have occurred after 0.65 Ma [4]. However mtDNA haplotypes of bonobos analyzed in this study coalesced to form one root (Figs 2, 4 and 5) and clearly separated from the haplotypes of chimpanzees. Therefore mtDNA haplotypes in this study can be thought to have originated within bonobo populations without gene flow from other *Pan* populations.

Molecular dating and reconstructing the ancestral areas for each node suggested that the MRCA of bonobos ranged the eastern areas, including TL2 and/or Iyondji, at around 0.64 Ma [0.95 Ma]. The A1, A2, and C haplogroups diverged in the central region, and the common

**Table 1. Genetic diversity for each cohort.**

Three cohorts	West	Central	East
N	23	115	16
Nucleotide diversity	0.01540 ± 0.007931	0.020855 ± 0.010232	0.012706 ± 0.006738

<https://doi.org/10.1371/journal.pone.0174851.t001>

**Table 2. Genetic diversity for each population.**

Seven populations	Malebo	Lac Tumba	Lomako	Salonga	Wamba	Iyondji	TL2
N	16	7	35	25	37	18	16
Nucleotide diversity	0.011588 ± 0.006173	0.022056 ± 0.012673	0.019060 ± 0.009576	0.022964 ± 0.011617	0.014601 ± 0.007396	0.018871 ± 0.009773	0.012706 ± 0.006738

<https://doi.org/10.1371/journal.pone.0174851.t002>

ancestor of the B and D haplogroups originated in the eastern region TL2. The B clades appeared to disperse from TL2 to the western region but not through the central region. Therefore, we need to consider how the B clades dispersed from the eastern to western regions directly.

**Forest refugia in the southern Congo Basin.** Palynological records since 1.35 Ma from the Ocean Drilling Program (ODP: 1985–2002) 1075 site in the deep sea fan sediments of the Congo River have shown that the most prominent change in pollen assemblages occurred at 1.05 Ma [18]. An increase in *Podocarpus* pollen, coupled with decreases in tropical, woodland, and mangrove tree pollen, suggests a shift to a cooler phase in the Congo Basin. The swampy open land probably dominated in the basin according to the retrogression of the tropical forest area in the basin during LGM [19,20,21]. Therefore, the wetland represented by *Cyperaceae*, *Xyris*, *Nymphaea lotus*, and *Laurembergia tetrandra*, as well as the savanna, might have also expanded to other cooler stages before LGM.

Two different areas have been proposed as the forest refugia for LGM. Most of them were figured as riparian forests along the main stream of the Congo River or the water-rich area around Lac Tumaba and Mai-Ndombe [22, 23, 24, 25, 26, 27]. The others were presumed as fragment forests in the center of the southern basin and Lomami (TL2) areas [28,29]. Each refugium possibly separated into more small forest fragments during severe dry phases or connected each other during relatively warmer dry phases, but main forest areas during the cooler periods would have remained situated along the Congo River and the center of the southern Congo Basin.

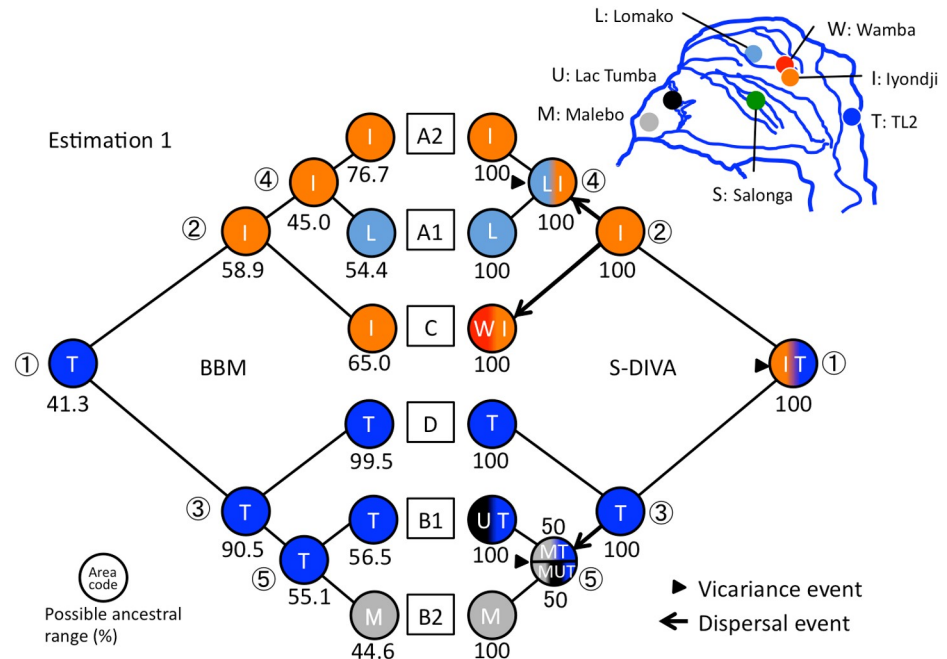
If both riparian forests and forests in the center of the southern Congo Basin have existed stably throughout the middle to late Pleistocene to a certain extent, the divergence through those forest refugia may explain the ancestral areas and direction of dispersion of the mtDNA haplotypes (Fig 6). The population inhabited around TL2 and Iyondji seems to have survived since 0.64 [0.95] Ma. The MRCA of haplogroups B and D seemed to have been formed around TL2 or over TL2 and the western region 0.47 Ma [0.67 Ma]. The MRCA of haplogroups A and C was born around 0.57 [0.84] Ma in the central region. Mutual exchanges of bonobo populations between these two forest refugia seem to have little possibility during the cooler stages. Montane forests, represented by *Podocarpus*, might approach from the Western Rift to the east to south regions of the bonobos' habitat during the cooler stages in the middle Pleistocene [30].

The existence of riparian forest refugia might have been easily imagined because the area could be regarded as the water rich area. However, what maintained the forest refugia in the center of the southern Congo Basin during cooler periods remains unknown. One possibility is the higher level of groundwater level at that location. The area where the top of the basement

**Table 3. Divergence time (Ma) for major clades calculated by BEAST (with 95% HPD).**

Root prior (chimpanzee/bonobo)	Node 1 (TMRCA)	Node 2 (C/A)	Node 3 (B/D)	Node 4 (A1/A2)	Node 5 (B1/B2)
0.95–1.05 (Estimation 1)	0.64 (0.50–0.79)	0.57 (0.42–0.74)	0.47 (0.33–0.64)	0.44 (0.28–0.60)	0.38 (0.26–0.53)
1.7–1.8 (Estimation 2)	0.95 (0.70–1.22)	0.84 (0.62–1.19)	0.67 (0.45–0.92)	0.60 (0.38–0.85)	0.53 (0.34–0.75)

<https://doi.org/10.1371/journal.pone.0174851.t003>



**Fig 4. Ancestral area reconstructions by RASP based on estimation 1.** Each pie graph shows the possible ancestral range (%) in each node. The same color indicates the same area. If the ancestral area of a node is indicated by more than one area, the gradient colors for these areas are used. Only the most likely state (i.e. the highest probability area) is shown in BBM. The triangles or arrows at each node show vicariance or dispersal events indicated by S-DIVA. Numbers in circles mean major nodes in the phylogenetic tree (Table 3).

<https://doi.org/10.1371/journal.pone.0174851.g004>

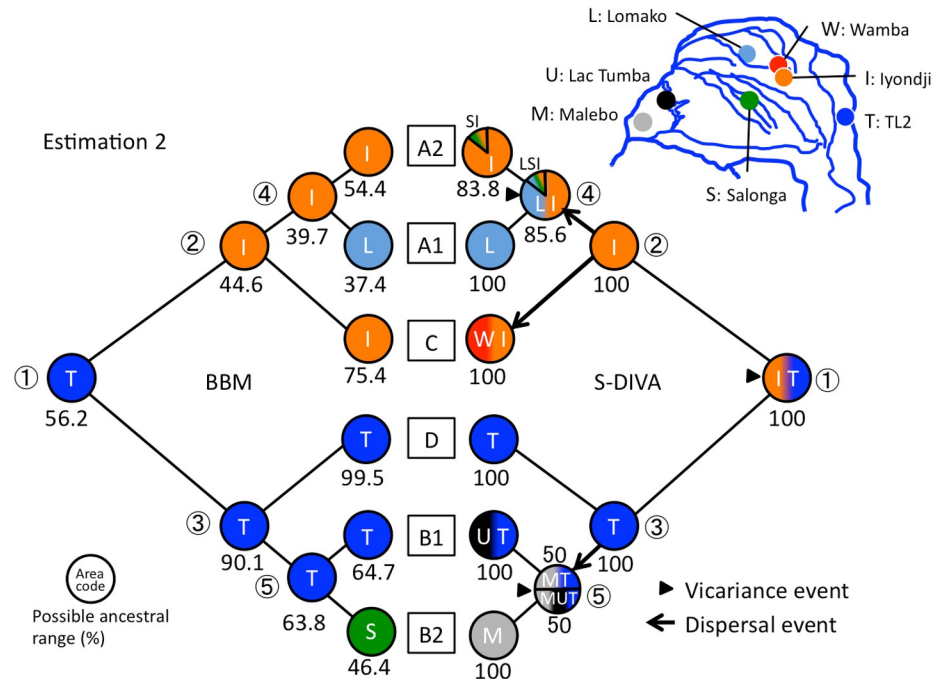
is less deep than the other areas lies in the region that is south-west to the center of the southern Congo Basin [31,32]. This subsurface structure, called Lokonia (or Lonkononia) High, might have supported the existence of the forests during the dry phases with higher level of groundwater.

**The dispersal routes of the bonobo populations in the left bank of the Congo River.**

Takemoto et al. [1] hypothesized that the bonobos’ ancestor crossed the Congo River to its left bank ca. 1.0 Ma [1.7–1.8 Ma], and that the most probable crossing point was the upper stream of the Congo River in the eastern region. The present study showed that the mtDNA ancestral populations most likely inhabited the eastern region of the current bonobos’ habitat at around 0.64 Ma [0.95 Ma] (Fig 7), supporting the previous supposition. Either 0.64 or 0.95 Ma was one of the coldest ages from 1.3 Ma [1,18, 33,34] and the bonobos’ range might have been very limited along the river. Bonobos might have crossed the Lomami River from the Lomami area to the Iyondji area during this cold age, or they might have already ranged over both banks of the Lomami River before this age. During the warm period immediately following it, the water levels of the Lomami River might have risen and divided the bonobo population across both banks of the Lomami River (A’ and B’).

The A’ population might have expanded their range to the center of the southern basin according to the increase of forest area. They would have survived in the refugia of the center of the southern basin and differentiating into the C, A1, and A2 haplogroups later. The C haplogroup remained near to the east areas such as the Iyondji and Wamba. The A2 haplogroup expanded to a wide range including the Lac Tumba of the western region, whereas the range of A1 remained in the central region. A part of the B’ population might have migrated to the left bank of the Lomami River at another age when the water level of the Lomami River





**Fig 5. Ancestral area reconstructions by RASP based on estimation 2.** Each pie graph shows the possible ancestral range (%) in each node. The same color indicates the same area. If the ancestral area of a node is indicated by more than one area, the gradient colors for these areas are used. Only the most likely state (i.e. the highest probability area) is shown in BBM. The triangles or arrows at each node show vicariance or dispersal events indicated by S-DIVA. Numbers in circles mean major nodes in the phylogenetic tree (Table 3).

<https://doi.org/10.1371/journal.pone.0174851.g005>

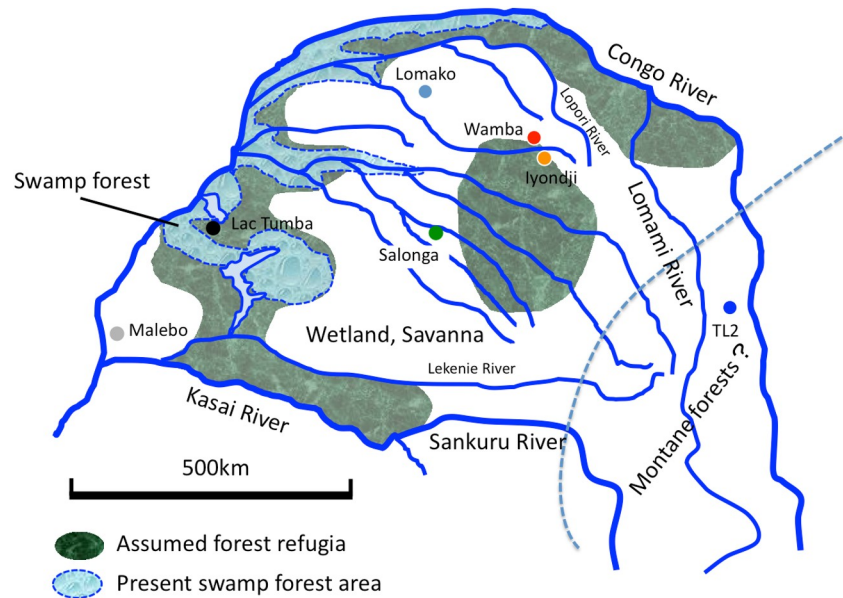
decreased. They seemed to have arrived at the Malebo through the forest refugium along the Congo River to become the common ancestor of the B2 and B1 haplogroups. The B2 haplogroup enlarged its range over the western and central regions. Furthermore, the B1 haplogroup was separated into the eastern and western region at around 0.26 Ma [0.36 Ma].

Mutual exchange of populations between the western and central regions might have occurred at plural times according to the forest expansion and contraction by glacial-inter glacial cycles. The eastern region, TL2, however did not seem to be connected to the central region directly, but through the western region due to the riverine barrier of the Lomami

**Table 4. Possible ancestral ranges (%) for each node in the phylogenetic tree indicated by RASP.**

S-DIVA	Node 1 (MRCA)	Node 2 (C/A)	Node 3 (B/D)	Node 4 (A1/A2)	Node 5 (B1/B2)
Estimation 1 0.95–1.05	TL2·lyondji (100.0)	lyondji (100.0)	TL2 (100.0)	lyondji·Lomako (100.0)	TL2·Malebo (50.0), TL2·Lac Tumba·Malebo (50.0)
Estimation 2 1.7–1.8	TL2·lyondji (100.0)	lyondji (100.0)	TL2 (100.0)	lyondji·Lomako (85.6), lyondji·Salonga·Lomako (14.4)	TL2·Malebo (50.0), TL2·Lac Tumba·Malebo (50.0)
BBM	Node 1 (MRCA)	Node 2 (C/A)	Node 3 (B/D)	Node 4 (A1/A2)	Node 5 (B1/B2)
Estimation 1 0.95–1.05	TL2 (41.3), lyondji (20.0), Wamba (12.2)	lyondji (58.9), Wamba (20.2), Lomako (7.2)	TL2 (90.5)	lyondji (45.0), Lomako (26.1), Wamba (10.8)	TL2 (55.1), Malebo (16.9), Lac Tumba (10.1), Salonga (5.7)
Estimation 2 1.7–1.8	TL2 (56.2), lyondji (13.3), Wamba (11.9)	lyondji (44.6), Wamba (31.3), TL2 (8.7)	TL2 (90.1)	lyondji (39.7), Wamba (27.8), Salonga (12.2), Lomako (7.2)	TL2 (63.8), Malebo (14.5), Lac Tumba (6.9)

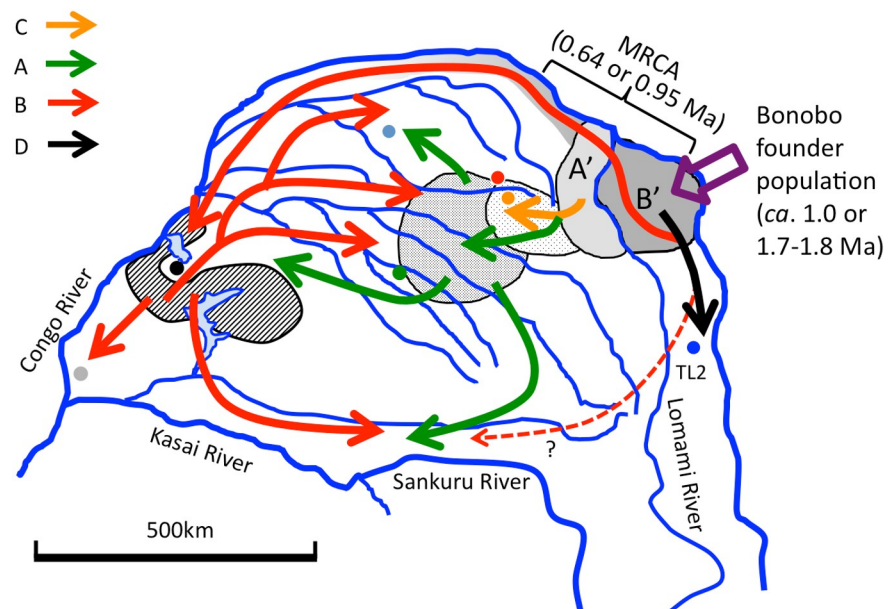
<https://doi.org/10.1371/journal.pone.0174851.t004>



**Fig 6. Assumed forest refugia of the bonobos' habitat during glacial stages.** Forest refugia might have existed stably along the main stream of the Congo River and the center of the southern Congo basin through cooler periods during the Pleistocene. Montane forests may have expanded to the basin in such periods (dotted line).

<https://doi.org/10.1371/journal.pone.0174851.g006>

River during the warm period and the distinct distribution of the forest refugia in the cooler periods. There is another possibility that the B clade dispersed to the western region through the southern route when the Montane forest approached near the basin. This dispersion route



**Fig 7. The hypothetical dispersal routs of bonobo populations since the mid-Pleistocene.** The areas of A' and B' represent the area of the MRCA of bonobos. Other hatched areas indicate the core refugia for the main haplogroups. Arrows in each color indicates the dispersal routs of each haplogroup. A dotted red arrow means another possible dispersion route for haplogroup B.

<https://doi.org/10.1371/journal.pone.0174851.g007>

**Table 5. Molecular dating for TMRCA of bonobos.**

	Bonobos/Chimpanzees Divergence time	TMRCA of bonobos	Calibrations and materials	Citation
Previous studies	1.8 ± 0.3 Ma	0.5 ± 0.27 Ma	Human/ <i>Pan</i> : 5–7 Ma, Y chromosome	Stone et al. 2002
	-	0.54 (0.43–0.66) Ma	Human/ <i>Pan</i> : 6–8 Ma, mtDNA, captive bonobos	Zsurka et al. 2010
	-	0.41–0.46 Ma	mtDNA, HV1	Eriksson et al. 2006
	-	0.04–0.045 Ma	Y chromosome	Eriksson et al. 2006
	1.74 (1.54–1.96) Ma	0.40 (0.35–0.45) Ma	Whole genome	Manuel et al, 2016
This study	(1.03 Ma)	0.64 (0.50–0.79) Ma	Human/ <i>Pan</i> : 6–8 Ma, Bonobo/Chimpanzee: 0.95–1.05 Ma, mtDNA	Estimation 1
	(1.76 Ma)	0.95 (0.70–1.22) Ma	Human/ <i>Pan</i> : 6–8 Ma, Bonobo/Chimpanzee: 1.7–1.8 Ma, mtDNA	Estimation 2

<https://doi.org/10.1371/journal.pone.0174851.t005>

would have yielded the same results with the riparian forest refugia, allowing the TL2 population to migrate to the central region via the western region (Fig 7).

The eastern region was indicated as the first colonization areas in the left bank of the Congo River not only for bonobos but also for other forest living animals [1,35]. Furthermore, the morphological study of cranial measurements indicated the TL2 population was divergent slightly more from the Lukenie population, ranging between the Lukenie and Sankuru River, than from the central region [36]. This is consistent with our hypothesis because the Lukenie area is the farthest area from TL2 when the forest refugium is followed along the mainstream. The hypothesis presented here will be verified by genetic studies for the populations along the mainstream of the Congo River. This area was assumed to represent migration routes for bonobo populations. If the B1 or D haplogroup, which is restricted to the west or east cohort, are found in these areas, our hypothesis may become more plausible.

### Differences in genetic diversity among populations

The BBM analysis of estimation 2 suggested the possibility that the ancestral area of B2 was the Salonga in the central region (Fig 5). This might be caused by a geographical proximity between the Salonga and western core refugia (Fig 7). The western core refugia was probably situated between Lac Tumba and Salonga, which is now occupied by swamp forests (Fig 6). The Salonga and Lac Tumba populations showed the highest and second highest nucleotide diversities (Table 1). That may be because they are the nearest populations belonging to different cohorts that maintained the different haplogroups, A and B. In other words, the sympatric existence of A and B haplogroups in a population appears to be a causal factor for the higher nucleotide diversity in the Salonga, Iyondji, Lomako, and Lac Tumba populations.

The genetic diversity of populations usually declines from the center of a geographical range to the periphery, and the differences in the genetic diversity among the study populations matched this pattern [7]. However, the mechanism that generates this pattern is not clear [37]. The present analysis showed the possibility that the central area has a higher genetic diversity involving various genetic clades because chances of exchanges between different genetic clades with neighboring areas might be greater in the central area than in the peripheral area. It may be helpful to investigate the genetic clades in populations from the viewpoint

of the historical biogeography in order to clarify the geographical pattern of the genetic diversity in these populations.

## Molecular dating for TMRCA of bonobo populations

Two studies [4, 38] estimated that the TMRCA of bonobos was 0.4–0.5 Ma and the divergence time of bonobos from chimpanzees to be around 1.8 Ma. Other four studies also suggested that TMRCA of bonobos by mtDNA was at around 0.5 Ma [4,5,6] (Table 5). If these estimations are true, three quarters or two-thirds of the bonobos' genetic history remains unknown. The current study estimated that the TMRCA of bonobos was 0.95 (0.70–1.22) Ma, using a divergence time of 1.7–1.8 Ma for bonobos/chimpanzees and 6–8 Ma for human/*Pan*. Similarly, when using a divergence time of 0.95–1.05 Ma, TMRCA of bonobos was estimated as 0.64 Ma. Our estimations suggest that the current population of bonobos had an older origin than previously reported after the branching off from chimpanzees. Our analysis shortened the gap between the divergence time of chimpanzees/bonobos and TMRCA of bonobos.

We delineated the history of bonobos in the southern Congo Basin from the limited DNA samples and information of the basin's paleoenvironment. Therefore, the hypothetical history of the bonobos presented here should be renewed in the near future. At this stage, however, it is difficult to obtain DNA samples from other sites, as well as additional detailed paleoenvironmental data on the early to middle Pleistocene in the basin. Addition of mtDNA samples from new sites to our analysis or analyzing nuclear DNA may elucidate the bonobos' population history in more detail.

## Acknowledgments

We thank the Centre de Recherche en Ecologie et Foresterie (CREF), Ministère de la Recherche Scientifique (MIN), the African Wildlife Foundation (AWF), the World Wide Fund for Nature (WWF), the DRC staff of the Zoological Society of Milwaukee's Bonobo and Congo Biodiversity Initiative (ZSM), the Institut Congolais pour la Conservation de la Nature (ICCN), ICCN Salonga National Park guards, the Tshuapa-Lomami-Lualaba (TL2) Project and Mbou-Mon-Tour Project for field research assistance, Takekazu Yumoto and Itaru Koizumi for useful comments, Michael A. Huffman for comments on the English and Isabel Behncke, Andrew Fowler, Laure Déruiti, Menard Mbende, Yukimaru Sugiyama and Tom and Nancy Gallagher for field collaboration. KY passed away before the submission of the final version of this manuscript. HT and TF accept responsibility for the integrity and validity of the data collected and analyzed.

## Author Contributions

**Conceptualization:** HT YK TF.

**Data curation:** HT YK TF.

**Formal analysis:** HT YK.

**Funding acquisition:** TF.

**Investigation:** HT YK SH EM.

**Methodology:** HT YK.

**Project administration:** TF.

**Resources:** JH TH TS NT GR PG JD AC MM KY SD CD.

**Supervision:** TF.

**Validation:** HT YK.

**Visualization:** HT.

**Writing – original draft:** HT YK TF.

**Writing – review & editing:** JH TH TS NT GR PG JD AC MM KY SD CD.

## References

1. Takemoto H, Kawamoto Y, Furuichi T. 2015. How did bonobos come to range south of the Congo River? Reconsideration of the divergence of *Pan paniscus* from other *Pan* populations. *Evolutionary Anthropology*. 34: 170–184.
2. Fischer A, Prüfer K, Good JM, Halbwax M, Wiebe V, Andre C, et al. 2011. Bonobos Fall within the Genomic Variation of Chimpanzees. *PLoS ONE* 6(6): e21605 <https://doi.org/10.1371/journal.pone.0021605> PMID: 21747915
3. Prüfer K, Munch K, Hellmann I, Akagi K, Miller JR, Walenz B, et al. 2012. The bonobo genome compared with the chimpanzee and human genomes. *Nature* 486:527–531. <https://doi.org/10.1038/nature11128> PMID: 22722832
4. de Manuel M, Kuhlwilm M, Frandsen P, Sousa VC, Desai T, Prado-Martinez J, et al. (2016). Chimpanzee genomic diversity reveals ancient admixture with bonobos. *Science*, 354(6311), 477–481. <https://doi.org/10.1126/science.aag2602> PMID: 27789843
5. Eriksson J, Hohmann G, Boesch C, Vigilant L (2004) Rivers influence the population genetic structure of bonobos (*Pan paniscus*). *Mol Ecol* 13:3425–3435. <https://doi.org/10.1111/j.1365-294X.2004.02332.x> PMID: 15488001
6. Zsurka G, Kudina T, Peeva V, Hallmann K, Elger CE, Khrapko K, et al. (2010) Distinct patterns of mitochondrial genome diversity in bonobos (*Pan paniscus*) and humans. *BMC Evol. Biol.*, 10, 270. <https://doi.org/10.1186/1471-2148-10-270> PMID: 20813043
7. Kawamoto Y, Takemoto H, Higuchi S, Sakamaki T, Hart JA, Hart TB, et al. 2013. Genetic Structure of Wild Bonobo Populations: Diversity of Mitochondrial DNA and Geographical Distribution. *PLoS ONE* 8: e59660. <https://doi.org/10.1371/journal.pone.0059660> PMID: 23544084
8. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam F, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948. <https://doi.org/10.1093/bioinformatics/btm404> PMID: 17846036
9. Schneider S, Roessli D, Excoffier L. 2000. Arlequin ver. 2.000: a software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
10. Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.
11. Nei M. 1987. *Molecular evolutionary genetics*. New York: Columbia University Press. 448p.
12. BEAST 2 development team. 2014. BEAST v2.1.2.
13. Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, et al. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Comput Biol* 10(4): e1003537. <https://doi.org/10.1371/journal.pcbi.1003537> PMID: 24722319
14. Yu Y, Harris AJ, He X. 2010. S-DIVA (Statistical Dispersal-Vicariance Analysis): a tool for inferring biogeographic histories. *Molecular Phylogenetics and Evolution*, 56(2), 848–850. <https://doi.org/10.1016/j.ympev.2010.04.011> PMID: 20399277
15. Ronquist F. 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Syst. Biol.* 46:195–203
16. Kodandaramaiah U. 2010. Use of dispersal-vicariance analysis in biogeography—a critique. *Journal of Biogeography*, 37(1), 3–11.
17. Ronquist F, Sanmartín I. 2011. Phylogenetic methods in biogeography. *Annual Review of Ecology, Evolution, and Systematics*, 42(1), 441.
18. Dupont LM, Donner B, Schneider R, Wefer G. 2001. Mid-Pleistocene environmental change in tropical Africa began as early as 1.05 Ma. *Geology* 29:195–198.
19. Elenga H, Schwartz D, Vincens A. 1994. Pollen evidence of late Quaternary vegetation and inferred climate changes in Congo. *Palaeogeography, Palaeoclimatology, Palaeoecology*. 109:345–356.

20. Dupont LM, Behling H, Jahns S, Marret F, Kim JH. 2007. Variability in glacial and Holocene marine pollen records offshore from west southern Africa. *Vegetation History and Archaeobotany*, 16(2–3), 87–100.
21. Jahns S. 1996. Vegetation history and climate changes in West Equatorial Africa during the Late Pleistocene and Holocene, based on a marine pollen diagram from the Congo fan. *Vegetation History and Archaeobotany*, 5(3), 207–213.
22. Kingdon JS. 1980. The role of visual signals and face patterns in African forest monkeys (guenons) of the genus *Cercopithecus*. *The Transactions of the Zoological Society of London*, 35(4), 425–475.
23. Hamilton AC, Taylor D. 1991. History of climate and forests in tropical Africa during the last 8 million years. *Climatic Change*, 19: 65–78.
24. Maley J (2001) The impact of Arid phases on the African rain forest through geological history. *In African Rain Forest Ecology and Conservation* (Weber W., White L.J.T., Vedder A. and Naughton-Treves L. eds.), pp 68–87. Yale University Press, New Haven.
25. Ray N, Adams J. 2001. A GIS-based vegetation map of the world at the last glacial maximum (25,000–15,000 BP). *Internet Archaeology*, 11.
26. Rommerskirchen F, Eglinton G, Dupont L, Rullkötter J. 2006. Glacial/interglacial changes in southern Africa: Compound-specific  $\delta^{13}C$  land plant biomarker and pollen records from southeast Atlantic continental margin sediments. *Geochemistry, Geophysics, Geosystems*, 7(8).
27. Holtvoeth J, Wagner T, Schubert CJ. 2003. Organic matter in river-influenced continental margin sediments: The land-ocean and climate linkage at the Late Quaternary Congo fan (ODP Site 1075). *Geochemistry, Geophysics, Geosystems*, 4(12).
28. Anhuf D, Ledru MP, Behling H, Da Cruz FW Jr, Cordeiro RC, Van der Hammen T, et al. 2006. Paleoenvironmental change in Amazonian and African rainforest during the LGM. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 239(3), 510–527.
29. Colyn M, Gautier-Hion A, Verheyen W. 1991. A re-appraisal of palaeoenvironmental history in central Africa: evidence for a major fluvial refuge in the Zaïre Basin. *J Biogeogr* 18: 403–407.
30. Adie H, Lawes MJ. 2011. Podocarps in Africa: temperate zone relicts or rainforest survivors. *In Ecology of the Podocarpaceae in tropical forests* (Turner BL, Lucas C. eds). pp 79–100. Smithsonian Contributions to Botany, No. 95. Smithsonian Institution Scholarly Press, Washington, D.C.
31. Kadima E, Delvaux D, Sebagenzi SN, Tack L, Kabeyaz SM. 2011. Structure and geological history of the Congo Basin: an integrated interpretation of gravity, magnetic and reflection seismic data. *Basin Research* 23: 499–527.
32. Daly MC, Lawrence SR, Diemu-Tshiband K, Matouana B. 1992. Tectonic evolution of the Cuvette Centrale, Zaïre. *Journal of the Geological Society*, 149(4), 539–546.
33. Schefuß E, Schouten S, Jansen JF, Damsté JSS. 2003. African vegetation controlled by tropical sea surface temperatures in the mid-Pleistocene period. *Nature*, 422(6930), 418–421. <https://doi.org/10.1038/nature01500> PMID: 12660780
34. Schefuß E, Sinninghe Damsté JS, Jansen JH. 2004. Forcing of tropical Atlantic sea surface temperatures during the mid-Pleistocene transition. *Paleoceanography*, 19(4).
35. Stanton DWG, Hart J, Galbusera P, Helsen P, Shephard J, Kumpel NF, et al. 2014. Distinct and Diverse: Range-Wide Phylogeography Reveals Ancient Lineages and High Genetic Variation in the Endangered Okapi (*Okapia johnstoni*). *PLoS ONE* 9(7): e101081. <https://doi.org/10.1371/journal.pone.0101081> PMID: 25007188
36. Pilbrow V, Groves C. 2013. Evidence for Divergence in Populations of Bonobos (*Pan paniscus*) in the Lomami-Lualaba and Kasai-Sankuru Regions Based on Preliminary Analysis of Craniodental Variation. *International Journal of Primatology*, 34(6), 1244–1260.
37. Eckert CG, Samis KE, Loughheed SC (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Mol Ecol* 17: 1170–1188. <https://doi.org/10.1111/j.1365-294X.2007.03659.x> PMID: 18302683
38. Stone AC, Griffiths RC, Zegura SL, Hammer MF. 2002. High levels of Y chromosome nucleotide diversity in the genus *Pan*. *Proc Natl Acad Sci USA* 99: 43–48. <https://doi.org/10.1073/pnas.012364999> PMID: 11756656