Local genetic adaptation to habitatin wild chimpanzees

Harrison J. Ostridge, Claudia Fontsere, Esther Lizano, Daniela C. Soto, Joshua M. Schmidt, Vrishti Saxena, Marina Alvarez-Estape, Christopher D. Barrat [et al.]

DOI:
https://doi.org/10.1126/science.adn7954

RESEARCH ARTICLE SUMMARY

EVOLUTION

Local genetic adaptation to habitat in wild chimpanzees

Harrison J. Ostridge *et al.*

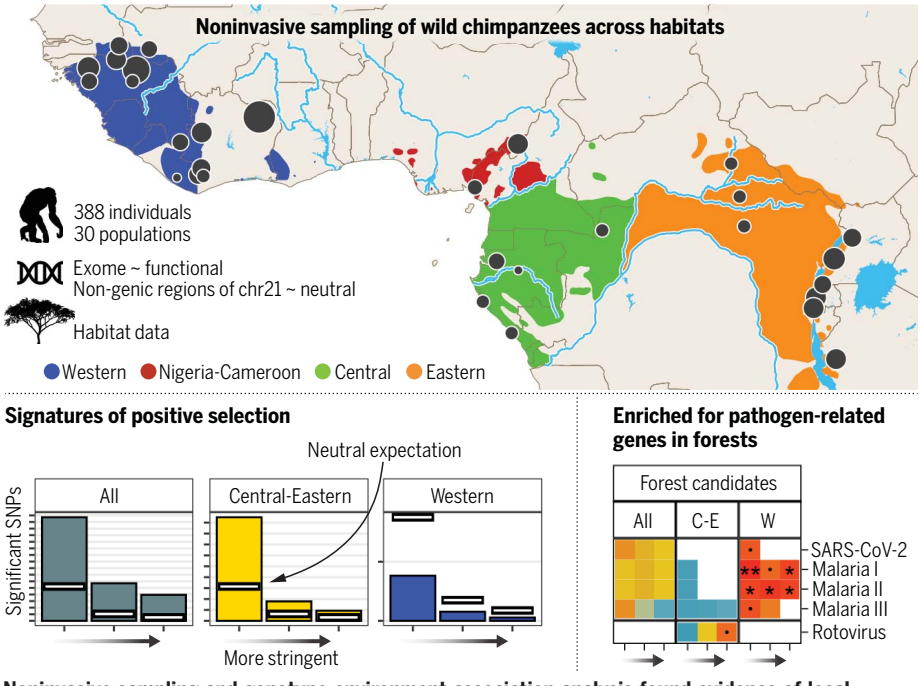
INTRODUCTION: How populations adapt to their environment is a fundamental question in biology. Yet, we know surprisingly little about this process, especially for endangered species, such as the nonhuman great apes. Chimpanzees, our closest living relatives, are particularly notable because they inhabit a diversity of habitats, from rainforest to woodland-savannah. Forests have closed canopies with high availability of food and water throughout the year, support high population densities, and harbor a diversity of pathogens and disease vectors. Conversely, savannahs are on the edge of the chimpanzee distribution in East and West Africa and are characterized by open canopies, higher temperatures, lower annual rainfall, and higher rainfall seasonality. Whether genetic adaptation facilitates chimpanzees' habitat diversity remains unknown, despite having wide implications for evolutionary biology and conservation.

RATIONALE: Investigating signatures of local adaptation requires genomic samples from wild individuals with known geographic origins. Noninvasive sampling is the only option for many protected species, including nonhuman great apes; however, recent technical and analytical advancements are beginning to enable population genomic analyses on such samples. With fecal samples collected as part of the Pan African Programme, we sequenced the full exome (i.e., protein-coding regions of the genome) from hundreds of wild chimpanzees across their geographic and environmental range. Putatively neutral regions in previously published chromosome 21 (chr21) sequences from the same samples were used to generate null expectations.

RESULTS: Integrating genetic and environmental data provides evidence of fine-scale local genetic adaptation in the form of an excess of single-nucleotide polymorphisms (SNPs) associated with a measure of habitat. This includes genetic adaptations to both forest and woodland-savannah habitats. These results suggest that although tool use and thermoregulatory behaviors are important in mitigating environmental stressors, selective pressures still drive genetic adaptation in chimpanzees. Thus, both behavioral flexibility and genetic adaptation may explain how chimpanzees inhabit such a range of habitats.

SNPs inferred to be under positive selection in forests are enriched for pathogen-related genes, consistent with the higher infectious disease burden in these habitats. This highlights the potential importance of genetic adaptation in shaping infectious disease mortality and, therefore, the dangers of displacement and environmental change. Most notably, forest candidate SNPs in the western subspecies are strongly enriched for malaria-related genes. A range of malaria parasites infect chimpanzees, including three species closely related to *Plasmodium falciparum*, which is responsible for 90% of global malaria mortality in humans. However, the fitness effects of malaria in wild chimpanzees are poorly understood. Our results indicate that this disease may have driven local adaptation and could have fitness effects in present-day wild populations. Genes with signatures of positive selection in chimpanzees underlie resistance and adaptation to malaria in humans. This is notable from an evolutionary point of view and demonstrates how understanding chimpanzee evolution can inform human evolution and medicine.

CONCLUSION: We found evidence for the presence of locally adaptive genetic differences among populations of wild chimpanzees, even at a fine geographic scale. Just as previous studies highlighted the importance of conserving behavioral diversity, we emphasize the need to consider local genetic adaptation in conservation efforts to ensure that individuals are adapted to their local environment and retain adaptive potential. This is particularly relevant, as direct anthropogenic destruction, climate change, and disease transmission are rapidly changing the environments experienced by chimpanzees. Our study also demonstrates the value and promise of noninvasive sampling to investigate genetic adaptation in wild populations of endangered species. ■



Noninvasive sampling and genotype-environment association analysis found evidence of local genetic adaptation in chimpanzees. Exomes and whole chr21 were sequenced from hundreds of samples spanning the chimpanzee geographical and environmental range. Population locations are shown as circles on a map of West, Central, and East Africa, with sizes proportional to the number of samples. Genotype-environment association analyses found an excess of SNPs strongly associated with habitat in the exome compared with nongenic regions. Candidate targets of positive selection in forests are enriched for pathogen-related genes, particularly malaria-related genes. All, *All* subspecies dataset; C-E, *Central-Eastern* subspecies dataset; W, *Western* subspecies dataset. •False discovery rate (FDR) < 0.1, *FDR < 0.05, **FDR < 0.01.

The list of authors and their affiliations is available in the full article online.
*Corresponding authors: Harrison J. Ostridge (harrison.ostridge@btinternet.com); Aida M. Andrés (a.andres@ucl.ac.uk)
Cite this article as H. J. Ostridge *et al.*, *Science* **387**, eadn7954 (2025). DOI: 10.1126/science.adn7954

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adn7954

RESEARCH ARTICLE

EVOLUTION

Local genetic adaptation to habitat in wild chimpanzees

Harrison J. Ostridge[1]*, Claudia Fontsere[2,3], Esther Lizano[3], Daniela C. Soto[4], Joshua M. Schmidt[5], Vrishti Saxena[1], Marina Alvarez-Estape[3], Christopher D. Barratt[6,7], Paolo Gratton[8], Gaëlle Bocksberger[9], Jack D. Lester[10], Paula Dieguez[11], Anthony Agbor[10], Samuel Angedakin[10], Alfred Kwabena Assumang[12], Emma Bailey[10], Donatienne Barubiyo[10], Mattia Bessone[10,13], Gregory Brazzola[10], Rebecca Chancellor[14], Heather Cohen[7,10,15], Charlotte Coupland[10], Emmanuel Danquah[12], Tobias Deschner[10,16], Laia Dotras[17,18], Jef Dupain[19], Villard Ebot Egbe[10], Anne-Céline Granjon[7,10,20], Josephine Head[21], Daniela Hedwig[22], Veerle Hermans[23], R. Adriana Hernandez-Aguilar[17,18], Kathryn J. Jeffery[24,25], Sorrel Jones[10], Jessica Junker[11,26], Parag Kadam[27], Michael Kaiser[10], Ammie K. Kalan[28], Mbangi Kambere[10], Ivonne Kienast[29,30], Deo Kujirakwinja[31], Kevin E. Langergraber[32,33], Juan Lapuente[10], Bradley Larson[34], Anne Laudisoit[35,36], Kevin C. Lee[10,32,33], Manuel Llana[17], Giovanna Maretti[10], Rumen Martín[10], Amelia C. Meier[10,37], David Morgan[38], Emily Neil[10], Sonia Nicholl[10], Stuart Nixon[39], Emmanuelle Normand[40], Christopher Orbell[24,41,42], Lucy Jayne Ormsby[10], Robinson Orume[43], Liliana Pacheco[44], Jodie Preece[10], Sebastien Regnaut[40], Martha M. Robbins[10], Aaron Rundus[14], Crickette Sanz[45,46], Lilah Sciaky[10], Volker Sommer[47,48], Fiona A. Stewart[10,47], Nikki Tagg[23,49], Luc Roscelin Tédonzong[23], Joost van Schijndel[10], Elleni Vendras[50], Erin G. Wessling[51,52], Jacob Willie[23,53], Roman M. Wittig[54,55], Yisa Ginath Yuh[10], Kyle Yurkiw[10], Linda Vigilant[10], Alex K. Piel[47], Christophe Boesch[40], Hjalmar S. Kühl[11,56], Megan Y. Dennis[4], Tomas Marques-Bonet[3,57,58,59], Mimi Arandjelovic[7,10], Aida M. Andrés[1]*

How populations adapt to their environment is a fundamental question in biology. Yet, we know surprisingly little about this process, especially for endangered species, such as nonhuman great apes. Chimpanzees, our closest living relatives, are particularly notable because they inhabit diverse habitats, from rainforest to woodland-savannah. Whether genetic adaptation facilitates such habitat diversity remains unknown, despite it having wide implications for evolutionary biology and conservation. By using newly sequenced exomes from 828 wild chimpanzees (388 postfiltering), we found evidence of fine-scale genetic adaptation to habitat, with signatures of positive selection in forest chimpanzees in the same genes underlying adaptation to malaria in humans. This work demonstrates the power of noninvasive samples to reveal genetic adaptations in endangered populations and highlights the importance of adaptive genetic diversity for chimpanzees.

Understanding how primates are adapted to their environments provides insights into our own evolution as well as information for conservation efforts. This is particularly relevant for our closest living relatives, nonhuman great apes, all of which are either endangered or critically endangered. Chimpanzees (*Pan troglodytes*) have the largest geographic and ecological range of any nonhuman ape [2.6 million km^2 (*1*)], spanning a variety of environments across Equatorial Africa, from dense tropical rainforest to open woodland-savannah mosaics (hereafter referred to as "savannah") (*2*). Aside from humans, they are the only great apes that inhabit savannah habitats (*2*). Yet, each of the four subspecies of chimpanzee [central (*P. t. troglodytes*), eastern (*P. t. schweinfurthii*), Nigeria-Cameroon (*P. t. ellioti*), and western (*P. t. verus*) (*3*–*5*)] are endangered (westerns, critically so), with numbers continuing to decline owing to hunting, habitat destruction, and infectious diseases (*1*, *6*–*8*). This decline has widespread negative impacts, as chimpanzees are an important conservation flagship species for biodiversity protection and are crucial ecosystem engineers (*9*–*11*).

Between the forest and savannah extremes, chimpanzees occupy a gradient of habitats, known as forest-savannah mosaics (*12*). Forests, which are likely closest to chimpanzee ancestral habitats (*3*, *13*), have closed canopies with high availability of food and water throughout the year and therefore tend to support high population densities (*2*). Forests also harbor a great diversity of pathogens and disease vectors (*14*). Conversely, savannahs are on the edge of the chimpanzee distribution in East and West Africa and are characterized by open canopies, higher temperatures, lower annual rainfall, and higher rainfall seasonality (*2*, *15*).

The occupation of such a range of habitats is facilitated by chimpanzee behavioral diversity (*16*). Behavioral adaptations include tool use in a range of contexts, such as foraging (*17*–*19*), water extraction (*20*–*22*), and communication (*23*). Savannah chimpanzees exhibit distinct thermoregulatory behaviors (*24*, *25*) and, on average, tend toward greater behavioral diversity than forest chimpanzees (*16*), a potential adaptation to higher environmental variability. Nevertheless, behavior does not fully compensate for stressors, as shown by physiological stress in response to pathogens (*26*–*29*) and environmental pressures (*15*, *30*).

Another mechanism that can facilitate the occupation of diverse habitats is genetic adaptation, just as local adaptation has contributed to genetic population differentiation in humans (*31*) despite great behavioral flexibility (*32*, *33*). In fact, humans have evolved local genetic adaptations to environmental pressures that differ between forest and savannah habitats, including pathogens (*34*–*36*), such as malaria (*37*, *38*); diet (*39*–*41*); solar exposure (*42*); and climatic variables, such as temperature and water availability (*43*, *44*). Culture can also promote genetic adaptations, similar to human adaptations to diet and zoonotic diseases associated with animal domestication (*41*, *45*).

Establishing whether genetic differences underlie local adaptation in chimpanzees is important to understanding primate evolution and critical for chimpanzee conservation. If adaptive genetic differences exist among populations, then this genetic diversity should be conserved to maintain existing adaptations and adaptive potential (*46*, *47*). Additionally, recent genetic adaptations highlight key selective pressures that likely shape fitness in the wild today and can help establish which populations may be more vulnerable to environmental change (*46*). This is particularly relevant in the face of anthropogenic climate change, which is increasing temperatures and precipitation seasonality within the chimpanzees' range (*2*). Furthermore, chimpanzees are excellent models for understanding our own evolution, particularly in savannah regions, which resemble early hominin habitats (*2*, *48*–*52*). Lastly, the close genetic similarity between humans and other great apes (*53*) has resulted in zoonotic disease transmissions (*54*, *55*), such as HIV-1 (*56*) and malaria (*57*). Understanding how chimpanzees have evolved to reduce the pathogenicity of microorganisms can thus reveal potential targets for treatments and vaccines (*58*–*61*).

We have a growing understanding of chimpanzee demographic history thanks to population genomics studies (*3*, *5*, *62*), which have identified genetic differentiation among populations within each subspecies. However, our knowledge of genetic adaptation lags behind, largely owing to sample limitations. Because existing genomic datasets include only dozens of captive chimpanzees of unknown geographic origin (*5*, *62*–*64*), previous studies investigated adaptation only at the subspecies level (*63*, *65*–*73*), revealing notable subspecies-level adaptations, for example, to pathogens such as simian immunodeficiency virus (SIV) (*65*, *66*). However, habitats vary greatly within chimpanzee subspecies ranges (*12*); therefore, subspecies

comparisons are uninformative on adaptations to many potential selective pressures. Investigation of fine-scale local adaptation is essential for understanding adaptation in chimpanzees but requires large numbers of DNA samples from wild individuals of known geographic origin coupled with detailed environmental data.

Noninvasive sampling is the only ethical and feasible option for studying wild populations of many protected species (*74*), including nonhuman apes; now, recent technical and analytical advancements are beginning to enable population genomic analyses on such samples (*3*, *75*–*80*). By using fecal samples from wild individuals collected as part of the Pan African Programme: The Cultured Chimpanzee (PanAf) (*3*, *81*), we have generated full exome (protein-coding regions of the genome) sequences from hundreds of chimpanzees across their geographic and environmental range. We demonstrate that genomic data from noninvasive samples can be used to reveal the fine-scale adaptive history of endangered primates. When integrated with environmental data, the exomes revealed evidence of local genetic adaptation to habitat conditions in chimpanzees. In forests, we found signatures of positive selection in pathogen-related genes, including those mediating adaptation to malaria in humans.

Samples and sequences

Fecal samples from 828 distinct individuals were collected from 52 sampling sites across the geographic range of all four chimpanzee subspecies as part of PanAf (*3*, *81*). This represents a 10-fold increase in sample size and a massive increase in geographic coverage over existing genome-wide datasets of any nonhuman ape (*5*, *62*). The scale and resolution of the dataset are only comparable to the previously published chromosome 21 (chr21) sequences from the same samples (*3*).

Noninvasive samples typically contain low levels of endogenous DNA. Thus, we target-captured and sequenced full exomes [akin to chr21 (*3*)] because they are informative for the vast majority of functional sites in the genome, including both sequenced protein-coding and linked regulatory regions (e.g., promoters). Samples were strictly filtered to omit those with first-order relatives, contamination, or low read depth [supplementary text, section 3 (*82*)]. To mitigate the potential effects of the moderate read depth and to take advantage of the large sample size, we used genotype likelihoods and allele frequency–based methods, which minimize the effects of individual sequencing errors.

As expected, by using either exomes or chr21 (*3*), population structure analyses separated samples into four subspecies (figs. S13, S15, and S16), and within-subspecies population structure inferred with the exomes closely matched results from chr21 (*3*) (figs. S13 to S17). Each sample site was considered a genetic unit, which we refer to as a "population," except for four populations formed by combining very closely related sample sites [details in materials and methods and supplementary text, section 4 (*82*)]. After removing populations with fewer than eight samples, the final dataset contained 388 exomes (385 chr21) from 30 populations: 5 central, 9 eastern, 2 Nigeria-Cameroon, and 14 western. The resulting exomes have a median read depth per sample of 5.30-fold (0.51- to 52.27-fold) in the exome target space (60 mega–base pairs). We investigated the signatures of local adaptation within and across subspecies in four "subspecies datasets" containing populations from all subspecies (*All*), central and eastern together (*Central-Eastern*), Nigeria-Cameroon (*Nigeria-Cameroon*), and western (*Western*). Central and eastern have very low genetic differentiation, with a fixation index (F_{ST}) of 0.10 (Fig. 1) (*81*), and were combined to increase the sample size and environmental diversity, and, thus, the power of the analysis. *All* contains 521,015 single-nucleotide polymorphisms (SNPs) (covering 15,600 genes); *Central-Eastern*, 314,934 SNPs (15,518 genes); *Nigeria-Cameroon*, 108,382 SNPs (14,585 genes); and *Western*, which had a large sample size, 175,266 SNPs (15,278 genes) (tables S2 and S3). Although many of these SNPs are present in the existing small–sample size, high-coverage dataset (*5*, *62*) we also discovered thousands of high-quality SNPs (figs. S22 to S26) and generated a dataset of polymorphisms across wild chimpanzee populations [supplementary text, section 5.3 (*82*)]. The unfolded site frequency spectra (SFS) showed no abnormalities that would indicate biases or errors in the allele frequency estimations. The exome SFS has relatively fewer midfrequency alleles than the nongenic regions of chr21, as expected under stronger purifying selection at functional sites. We also replicated previous evidence for positive selection among chimpanzee subspecies (*65*) (fig. S26).

Allele frequency population differentiation

Local adaptation increases the frequency of alleles only where they are beneficial, generating large allele frequency differences among populations. We first investigated local positive selection by analyzing population differentiation with a genetics-only, hypothesis-free analysis using the BayPass (*84*) core model. BayPass

[1]UCL Genetics Institute, Department of Genetics, Evolution and Environment, University College London, London, UK. [2]Center for Evolutionary Hologenomics, The Globe Institute, University of Copenhagen, Copenhagen, Denmark. [3]Institute of Evolutionary Biology (UPF-CSIC), PRBB, Dr. Aiguader 88, Barcelona, Spain. [4]Department of Biochemistry & Molecular Medicine, Genome Center, MIND Institute, University of California, Davis, Davis, CA, USA. [5]Flinders Health and Medical Research Institute (FHMRI), Department of Ophthalmology, Flinders University Sturt Rd, Bedford Park, South Australia, Australia. [6]Naturalis Biodiversity Center, Darwinweg 2, Leiden, the Netherlands. [7]German Centre for Integrative Biodiversity Research (iDiv) Halle-Leipzig-Jena, Puschstrasse 4, Leipzig, Germany. [8]University of Rome "Tor Vergata," Department of Biology, Via Cracovia, 1, Roma, Italia. [9]Senckenberg Biodiversity and Climate Research Centre (SBiK-F), Senckenberganlage 25, Frankfurt am Main, Germany. [10]Max Planck Institute for Evolutionary Anthropology (MPI EVAN), Deutscher Platz 6, Leipzig, Germany. [11]Senckenberg Museum for Natural History Görlitz, Senckenberg – Member of the Leibniz Association, Am Museum 1, Görlitz, Germany. [12]Department of Wildlife and Range Management, Faculty of Renewable Natural Resources, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana. [13]University of Konstanz, Centre for the Advanced Study of Collective Behaviour, Universitätsstraße 10, Konstanz, Germany. [14]West Chester University, Department of Psychology, West Chester, PA, USA. [15]Institute for Biology, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany. [16]Institute of Cognitive Science, University of Osnabrück, Artilleriestrasse 34, Osnabrück, Germany. [17]Jane Goodall Institute Spain in Senegal, Dindefelo Biological Station, Dindefelo, Kedougou, Senegal. [18]Department of Social Psychology and Quantitative Psychology, Serra Hunter Programme, University of Barcelona, Barcelona, Spain. [19]Antwerp Zoo Foundation, RZSA, Kon.Astridplein 26, Antwerp, Belgium. [20]Leipzig University, Ritterstraße 26, Leipzig, Germany. [21]Arcus Foundation, 95 Regent Street, Cambridge, UK. [22]Elephant Listening Project, K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University, 159 Sapsucker Woods Road, Ithaca, NY, USA. [23]KMDA, Centre for Research and Conservation, Royal Zoological Society of Antwerp, Koningin Astridplein 20-26, Antwerp, Belgium. [24]School of Natural Sciences, University of Stirling, UK. [25]Agence National des Parcs Nationaux (ANPN) Batterie 4, Libreville, Gabon. [26]Re:wild, 500 N Capital of Texas Hwy Building 1, Suite 200, Austin, TX, USA. [27]Greater Mahale Ecosystem Research and Conservation Project P.O. Box 120, Mpanda, Tanzania. [28]Department of Anthropology, University of Victoria, 3800 Finnerty Rd, Victoria, Canada. [29]Department of Natural Resources and the Environment, Cornell University, Ithaca, NY, USA. [30]K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University, Ithaca, NY, USA. [31]Wildlife Conservation Society (WCS), 2300 Southern Boulevard. Bronx, NY, USA. [32]School of Human Evolution and Social Change, Arizona State University, 900 Cady Mall, Tempe, AZ Arizona State University, PO Box 872402, Tempe, AZ, USA. [33]Institute of Human Origins, Arizona State University, 777 East University Drive, Tempe, AZ Arizona State University, PO Box 872402, Tempe, AZ, USA. [34]FeralEarth Endeavors, Golden, CO, USA. [35]EcoHealth Alliance, New York, NY, USA. [36]Evolutionary Ecology Group (EVECO), University of Antwerp, Antwerp, Belgium. [37]Hawai'i Insititute of Marine Biology, University of Hawai'i at Manoa, 46-007 Lilipuna Place, Kaneohe, HI, USA. [38]Lester E. Fisher Center for the Study and Conservation of Apes, Lincoln Park Zoo, 2001 North Clark Street, Chicago, IL, USA. [39]North of England Zoological Society, Chester Zoo, Upton by Chester, CH2 1LH, UK. [40]Wild Chimpanzee Foundation (WCF), Abidjan, Côte d'Ivoire. [41]Panthera, 8 W 40TH ST, New York, NY, USA. [42]Institut de Recherche en Ecologie Tropicale, CENAREST, Libreville, Gabon. [43]Korup Rainforest Conservation Society, c/o Korup National Park, P.O. Box 36 Mundemba, Southwest Region, Cameroon. [44]Primate Project Inc., Miami, FL, USA. [45]Washington University in Saint Louis, Department of Anthropology, One Brookings Drive, St. Louis, MO, USA. [46]Congo Program, Wildlife Conservation Society, 151 Avenue Charles de Gaulle, Brazzaville, Republic of Congo. [47]University College London, Department of Anthropology, 14 Taviton Street, London, UK. [48]Gashaka Primate Project, Serti, Taraba State, Nigeria. [49]Born Free Foundation, Floor 2 Frazer House, 14 Carfax, Horsham, UK. [50]Frankfurt Zoological Society, Bernhard-Grzimek-Allee 1, Frankfurt, Germany. [51]German Primate Center, Leibniz Institute for Primate Research, Göttingen, Germany. [52]Johann-Friedrich-Blumenbach Institute for Zoology and Anthropology, Georg-August-University Göttingen, Göttingen, Germany. [53]Terrestrial Ecology Unit (TEREC), Department of Biology, Ghent University (UGent), K.L. Ledeganckstraat 35, Ghent, Belgium. [54]Ape Social Mind Lab, Institute for Cognitive Sciences Marc Jeannerod, CNRS UMR 5229 CNRS, 67 bd Pinel, Bron CEDEX, France. [55]Taï Chimpanzee Project, Centre Suisse de Recherches Scientifiques, Abidjan, Côte d'Ivoire. [56]International Institute Zittau, Technische Universität Dresden, Markt 23, Zittau, Germany. [57]Catalan Institution of Research and Advanced Studies (ICREA), Passeig de Lluís Companys, 23, Barcelona, Spain. [58]CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Baldiri i Reixac 4, Barcelona, Spain. [59]Institut Català de Paleontologia Miquel Crusafont, Universitat Autònoma de Barcelona, Edifici ICTA-ICP, c/ Columnes s/n, Cerdanyola del Vallès, Barcelona, Spain.
***Corresponding author. Email: harrison.ostridge@btinternet.com (H.J.O.); a.andres@ucl.ac.uk (A.M.A.)**

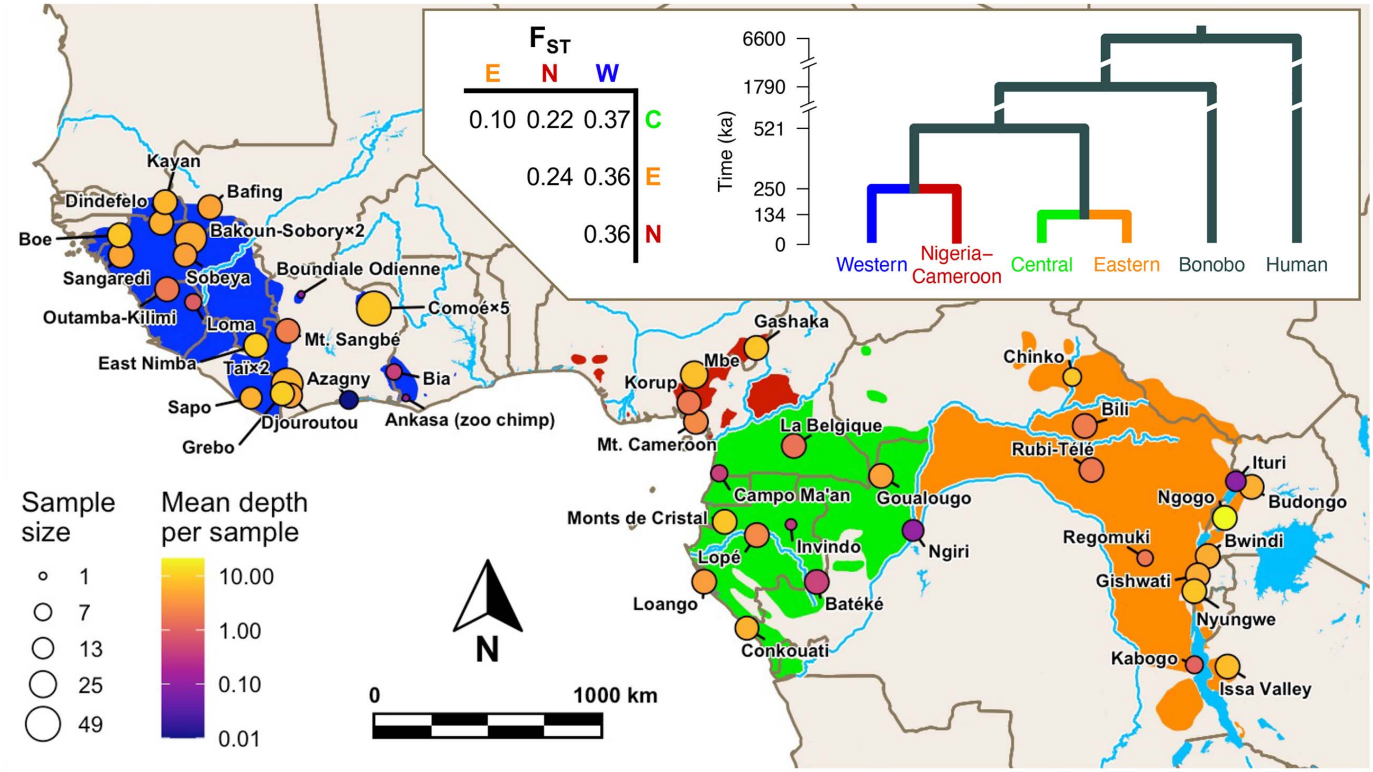


Fig. 1. Chimpanzee exome dataset distribution, sample size, and coverage. (Top inset) Hominini phylogenetic tree highlighting chimpanzee subspecies with estimated evolutionary split times (*62*, *83*) (ka, thousands of years ago) and pairwise F_{ST} between subspecies [supplementary text, section 4.1.1 (*82*)]. (Main) Map of West, Central, and East Africa indicating the location of sample sites (table S1), sample sizes, and mean exome sequencing read depth per sample. Each point represents a sample site, except for five geographically close sites sampled at Comoé, two at Taï, and two at Bakoun-Sobory. The geographic distribution of each subspecies is shown (blue, western; red, Nigeria-Cameroon; green, central; orange, eastern) (*1*), with major rivers and lakes indicated in light blue. The sample sizes and populations in the final filtered dataset used for selection analyses are shown in Fig. 3.

estimates the genome-wide population allele frequency covariance matrix, which is used to standardize allele frequencies for each SNP with respect to population structure (BayPass effectively accounts for population structure here; fig. S32). The variance across populations of these standardized frequencies is summarized in the test statistic X^tX^* (*85*). SNPs under local adaptation are expected to have exceptionally large population allele frequency differentiation and, therefore, the highest X^tX^* values in the genome. Approaches that rely on outliers to identify candidate targets of positive selection (hereafter referred to as "candidate SNPs") are sensitive to overinterpretation (*86*). Instead, we selected candidate SNPs by comparing exonic SNPs with expectations under neutrality generated using the nongenic regions of chr21 (nongenic-chr21) from these samples (*3*). Generating an empirical null frees the analysis from demographic assumptions and accounts for many potential confounding factors because nongenic-chr21 has an almost identical demographic history, sample size, and read depth to the exome and has been processed in the same way [supplementary text, section 6.2 (*82*)]. Because low coverage can lead to noisy allele frequency estimates, we matched SNPs in the exome and nongenic-chr21 by coverage to avoid the potential for read depth differences to generate false positives (*87*). This increased the stringency of our method (fig. S34) and effectively controlled for coverage differences (fig. S35). Candidate SNPs were defined as exome SNPs with higher X^tX^* than the values corresponding to estimated false positive rates (FPRs) of 0.5, 0.1, and 0.05% by using the nongenic-chr21 X^tX^* distribution [details in materials and methods and supplementary text, section 6.3.2 (*82*)]. We note that, although candidate SNPs show signatures of positive selection, they should not be assumed to be true positives, as some false positives likely exist.

If local adaptation drives population differentiation, then we expect exomes to show an excess of highly differentiated SNPs compared with neutral expectations. Contrary to this expectation, there are fewer SNPs with very large X^tX^* values in the exome compared with null expectations (Fig. 2 and fig. S33). This deficit could reflect the effects of purifying selection in the exome. Our method is conservative if there is any local adaptation in nongenic-chr21 loci. Moreover, even in the absence of genome-wide evidence, strong selective forces may have resulted in positive selection at a small number of key loci in the exome; however, an investigation into the putative functions of candidate genes did not find evidence of this [fig. S61; supplementary text, section 7.1 (*82*)].

Genetic adaptation to habitat

Integrating genetic and environmental data increases the power to detect signatures of local adaptation (*88*, *89*) and allows us to directly test the hypothesis that chimpanzees have adapted to selection pressures that vary between habitats. We thus performed a genotype-environment association (GEA) test by integrating an environmental covariable into the analysis using the BayPass AUX model (*84*). BayPass calculates a Bayes factor (BF) for each SNP that indicates the strength of evidence for a linear correlation between population allele frequencies and the environmental covariable while accounting for population structure [supplementary text, section 6.4.6 (*82*)]. SNPs evolving under local adaptation are expected to be highly correlated with the relevant environmental covariable and therefore have the highest BF in the genome.

Environmental metrics based on temperature, precipitation, or land cover do not correspond

well with researcher-defined forest and savannah regions (*12*). Therefore, we used a floristic measure informed by a large-scale biogeographic analysis that identified very different tree species compositions between forest and savannah regions and has been shown to produce more accurate maps of habitat distributions across Africa (*90*). Specifically, we used the percentage of trees identified as "forest specialists" (*90*) among all the classified trees recorded at each sample site (hereafter referred to as "forest-tree-percentage") (Fig. 3 and fig. S1) [see materials and methods and supplementary text, section 2 (*82*)]. This variable is ideal because the data was collected within the known ranges of the sampled populations, and the same field protocol can be applied to new sample sites, making our data and results comparable to those of future studies incorporating additional sample sites. Forest-tree-percentage was not used here to test for adaptation to tree species compositions, per se; rather, it was used to describe the chimpanzee habitat gradient, which summarizes many potential selective pressures (fig. S2). Using this single environmental variable focused our study on adaptation to the forest-savannah habitat gradient. This approach means that we cannot detect natural selection associated with variables that do not correlate closely with this gradient; however, it also aids the interpretation of our results and reduces the risk of false positives due to multiple testing. The GEA analysis was run with this covariable in each subspecies dataset, except *Nigeria-Cameroon*, as it has only two populations. Candidate SNPs were selected as in the genetics-only test [details in supplementary text, section 6.4.1 (*82*)].

In contrast to the genetics-only results, the GEA showed a substantial excess of SNPs strongly associated with forest-tree-percentage in the exome when compared with neutral expectations in *All* and *Central-Eastern* (Fig. 4 and fig. S42). This excess is precisely what we expect under local adaptation associated with habitat. The excess was only absent in *Western*, a subspecies dataset with considerable habitat diversity but with a demographic history that can explain this observation. Small long-term effective population size (N_e) (*3*, *5*, *62*) and potential high connectivity among populations (*3*) may have limited local adaptation in this subspecies. Further, low N_e increases the effects of random genetic drift, accelerating neutral allele frequency change and reducing our power to detect evidence of local adaptation at the genome-scale. We note that this result does not exclude the possibility that strong selective forces may have driven local adaptation in a few key genes in the western subspecies, and the SNPs with the highest BFs in the exome are the best candidate targets of positive selection.

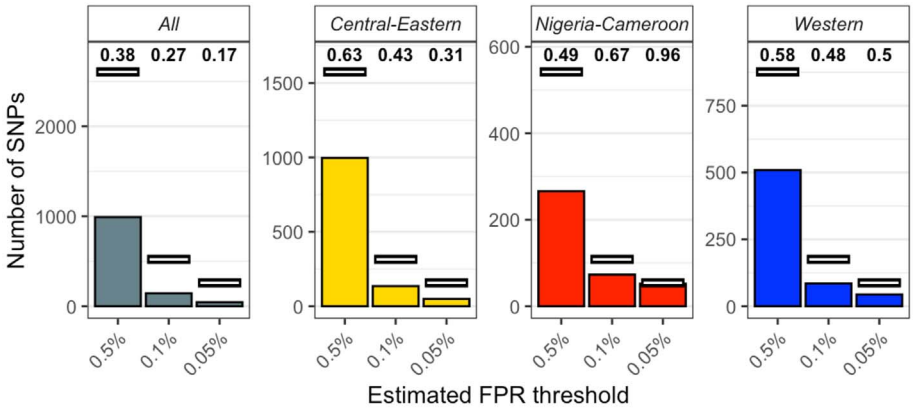


Fig. 2. Number of genetics-only candidate SNPs. The number of candidate SNPs from the genetics-only test (bars) compared with the null expectation (white lines) at X^tX^* thresholds corresponding to estimated FPRs of 0.5, 0.1, and 0.05% for each subspecies dataset. The exome to null expectation ratio is indicated at the top.

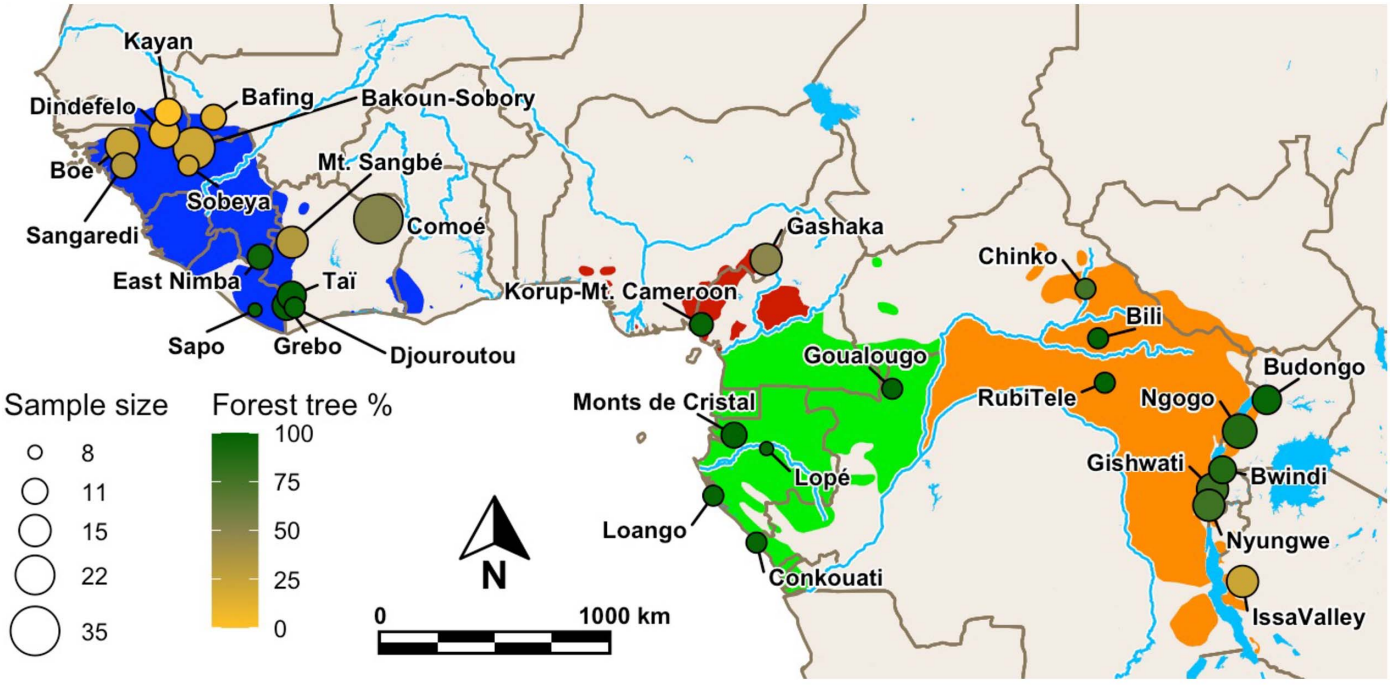


Fig. 3. BayPass analysis dataset. Map of West, Central, and East Africa showing the location, sample size (after filtering), and forest-tree-percentage for each population in the BayPass analyses. The ranges of the four subspecies are shown (blue, western; red, Nigeria-Cameroon; green, central; orange, eastern) (*1*), with major rivers and lakes indicated in light blue.

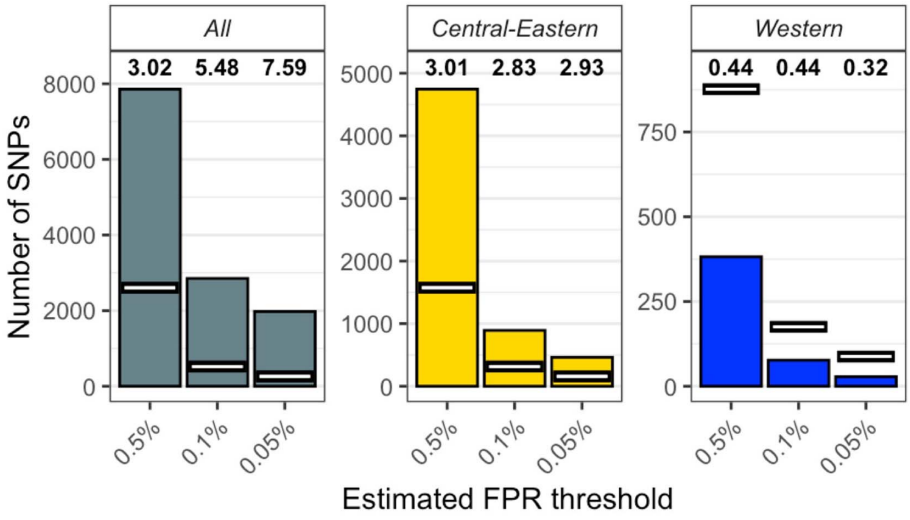


Fig. 4. Number of GEA candidate SNPs. The number of candidate SNPs from the GEA (bars) compared with the null expectation (white lines) at BF thresholds corresponding to estimated FPRs of 0.5, 0.1, and 0.05% for each subspecies dataset tested. The exome to null expectation ratio is indicated at the top.

These SNPs have strong evidence of positive selection that is not driven by potential confounding factors, such as read depth (fig. S45), allele frequency [fig. S48; supplementary text, section 6.4.8 (*82*)], linkage disequilibrium [fig. S49; supplementary text, section 6.4.9 (*82*)], or population substructure (fig. S54). Moreover, these signatures are not expected to be driven by background selection because its effects should not be associated with habitat, and X^tX^* shows no excess of highly differentiated exonic SNPs indicative of effects of background selection. For all thresholds and subspecies datasets, the minimum BF is very high: >14.7 for FPR < 0.5%, >18.3 for FPR < 0.1%, and >19.5 for FPR < 0.05% (fig. S44). Jeffrey's rule (*91*) defines 15 < BF < 20 as "very strong evidence" and BF > 20 as "decisive evidence," demonstrating that a vast majority of candidate SNPs have very strong evidence of being associated with habitat, with almost all SNPs in the 0.1% tail having decisive evidence (fig. S43). The candidate allele frequencies correlate strongly with forest-tree-percentage, as expected (figs. S55 and S56) and, as a set, show evidence of positive selection associated with forest-tree-percentage with Phylogenetic Generalized Least Squares (PGLS), a separate method that also accounts for population substructure [supplementary text, section 6.4.5 (*82*), Fig. S59].

These results provide strong evidence for local genetic adaptation to habitat in chimpanzees, revealing the presence of notable genetic differences among wild populations, even within subspecies, that likely shape fitness in an environment-dependent way.

These SNPs have strong, significant signatures of positive selection and are prime candidates to have mediated fine-scale local adaptation in chimpanzees, although we caution that, naturally, the set of candidate SNPs likely contains some false positives. Under the reasonable assumption that new adaptations are more commonly mediated by the new, derived allele than the ancestral one, we assigned SNPs as likely associated with forest or savannah adaptations according to the sign of their correlation coefficient. There is an excess of SNPs with high BFs in the exome for both savannah and forest candidates (fig. S52) in *All* and *Central-Eastern*, suggesting that adaptation in either direction contributes to the overall excess. To interpret these loci biologically, we investigated the genes that the candidate SNPs fall within (hereafter referred to as "candidate genes") by testing for an overrepresentation of functional categories in hypothesis-free gene set enrichment analyses. Given the relevance of pathogens as selective pressures (*65*, *66*), we also performed a hypothesis-driven enrichment analysis of pathogen-related genes [details in supplementary text, section 7 (*82*)]. These analyses point to potential different adaptations in savannah and forest chimpanzees.

Adaptations to savannah

Savannah candidate genes belong to many categories associated with physiological traits when compared with forest candidate genes (Fig. 5A). There are more than six times more "general" gene categories with a false discovery rate (FDR) < 0.5 in savannah candidates than in forest candidates (260 versus 42; Fig. 5A), although only two of these categories are significantly enriched (FDR < 0.05; Fig. 5B): negative regulation of nitrogen compound metabolic processes and negative regulation of cellular macromolecule biosynthetic processes. The large number of marginally enriched and few significantly enriched gene sets would be compatible with a large degree of polygenicity in the genetic adaptation of chimpanzees to the environmental extremes of savannahs. The diversity of categories and their overlap in genes makes it difficult to infer selection pressures that may drive this signal.

Restricted availability of water in savannahs during the dry season is a potential selection pressure (*24*, *30*) that could partially explain the enrichment of physiological categories in our candidates. However, there is no significant enrichment in the two dehydration response gene categories that we analyzed (*92*, *93*) [supplementary text, section 7 (*82*)], neither in the GEA (fig. S63) nor the genetics-only candidates (fig. S62). Chimpanzees may therefore have adapted to dehydration stress through genes not included in these categories or through behavioral adaptations [e.g., well digging (*20*, *94*)]. Alternatively, dehydration stress may be present but independent of habitat (*15*).

There is limited evidence of adaptation to pathogens in savannah populations (Fig. 5A, bottom left). Savannah candidates are significantly enriched (FDR < 0.05) for only one pathogen-related gene set: genes associated with AIDS progression in genome-wide association studies (GWASs) at the *Central-Eastern* 0.05% tail (Fig. 5B), which contains only two candidate genes (table S4). Viruses similar to SIV are not known to be associated with savannahs; instead, this result may be explained by adaptation in Issa Valley, which has a high prevalence of SIV (*95*, *96*) and a particularly low forest-tree-percentage in *Central-Eastern* (Fig. 3). Being extreme in forest-tree-percentage means that Issa Valley weighs heavily on the *Central-Eastern* savannah candidates, but without fully driving them [supplementary text, section 6.4.7 (*82*)]. The importance of adaptation to the eastern savannah habitat is highlighted by the fact that when central and eastern are analyzed separately, the evidence of local adaptation remains in the *Eastern* but not the *Central* datasets (figs. S46 and S47). Analyses of additional populations will help establish to what extent the evidence of adaptation is general across central and eastern savannah populations and identify the specific adaptive mechanisms and selective factors. In any case, the excess of exonic savannah candidate SNPs in *All* and *Central-Eastern* suggests that some chimpanzee populations do harbor genetic adaptations to savannah habitats.

Adaptations to forest

Although forest candidates show weak enrichment in general physiological categories, they show a pattern of stronger enrichment in pathogen-related genes, as shown in stronger enrichment for general "immunity genes" (*97*) and "innate immunity genes" (*98*), than savannah

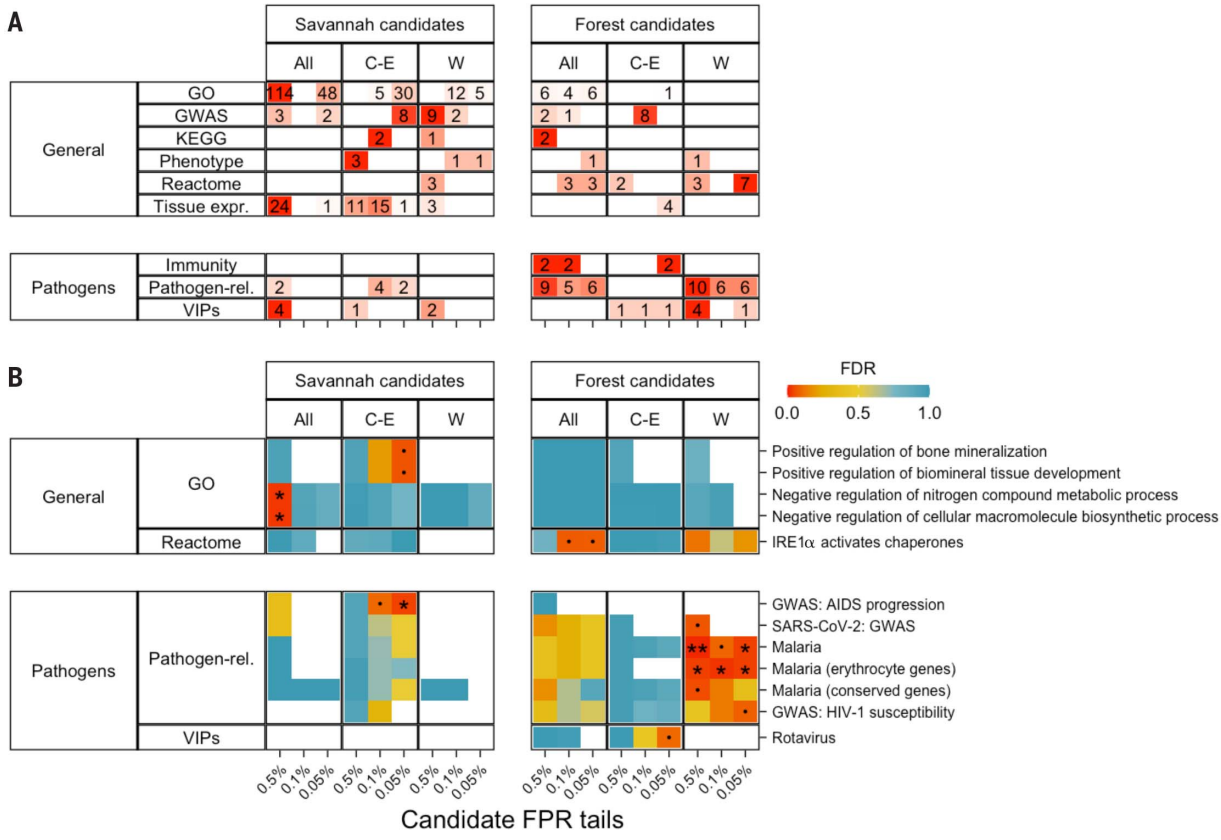


Fig. 5. GEA candidate gene set enrichment results. Results for 0.5, 0.1, and 0.05% FPR tails for savannah and forest candidate SNPs are shown. Vertical panels indicate results from each subspecies dataset. Horizontal panels show the broad categories to which the gene sets belong. Multiple testing correction was done within each gene set enrichment analysis run (i.e., each tail and gene set database, such as "Pathogen-related," "GWAS," and "Phenotype"). (**A**) The number of gene sets with FDR < 0.5; cells are colored in a gradient from white (zero) to red (the largest value per row). (**B**) The FDR values for the most-enriched gene sets with FDR < 0.1 for at least one candidate tail in at least one subspecies dataset. •FDR < 0.1, *FDR < 0.05, **FDR < 0.01. Pathogen-rel., pathogen related; Tissue expr., tissue expression; VIPs, viral interacting proteins; GO, Gene Ontology; All, "All" subspecies dataset; W, *Western* subspecies dataset; C-E, *Central-Eastern* subspecies dataset.

candidates in *All* and *Central-Eastern* (Fig. 5A). This pattern is also evident for individual pathogen categories in *All* and especially in *Western*, although not in *Central-Eastern* (Fig. 5). This is consistent with the higher population densities (*2*) and increased pathogen exposure (*14*) in forests resulting in a greater infectious disease burden. In humans, local adaptation has likely also been driven by high pathogen diversity (*99*), particularly in tropical forests (*34*–*36*). *Central-Eastern* does not show this pattern, likely owing to the presence of some eastern populations from montane forests, which are considerably cooler than lowland forests and therefore have lower levels of vector-borne diseases, such as malaria (*14*) [supplementary text, section 7.2.1 (*82*)]. Enrichment of pathogen-related categories in the *Western* forest candidates suggests that, although we do not see evidence of positive selection on the genome scale, strong selection at a limited number of pathogen-related genes is likely driving local adaptation in this subspecies.

Focusing on individual pathogens, the strongest and clearest signal is enrichment for malaria-related categories in *Western* forest candidates (Fig. 5B and table S4). They are significantly enriched (FDR < 0.05) in "Malaria-related genes" (*100*) at the 0.5 and 0.05% tails and, in "Erythrocyte genes related to malaria" (*101*), at all three tails; enrichment in *Plasmodium*-interacting proteins that are conserved across mammals (*102*) (thus excluding hemoglobin and glycophorin genes; see below) very narrowly exceeds the significance threshold (FDR = 0.050) in the 0.5% tail. *Western* lacks genome-wide evidence of positive selection, showing how strong selection in a few key genes can leave genomic signatures even in low-N_e populations and motivating further analyses to confirm the evidence of positive selection in these loci. Malaria infection probability in chimpanzees is closely correlated with tree cover (*14*), which is itself highly correlated with forest-tree-percentage in our dataset (Pearson correlation coefficient $r = 0.92$, $P = 9.044 \times 10^{-13}$) (*103*). Malaria is a major selection pressure and has driven some of the clearest examples of local adaptation in humans (*37*, *38*). Only five genetic variants have been significantly associated with severe malaria in human GWAS (*104*, *105*). Notably, two of these loci, which encode for hemoglobin (*HB*) and glycophorin (*GYP*) genes (Fig. 6A), contain chimpanzee forest candidate SNPs; both loci also underlie adaptations to malaria in humans (*106*–*109*).

For HB, candidate SNP chr11:5254366 (*Western* 0.5% tail, PGLS $P = 9.33 \times 10^{-3}$; *All* 0.05% tail, PGLS $P = 1.97 \times 10^{-4}$) lies within an intron of hemoglobin subunit delta (*HBD*), less than 5 kb upstream of the adjacent paralogue, hemoglobin subunit beta (*HBB*). Although mutations in *HBD* have little effect on malaria resistance owing to low expression in adults (*110*), the sickle hemoglobin (HbS) mutation in *HBB* is a classic example of balancing selection in humans, as heterozygotes are protected against severe malaria (*106*, *107*). Therefore, the signatures that we observe may reflect selection on a linked variant within *HBB*, the regulation of *HBB*, or *HBD* itself. In any case, it is notable that this locus shows evidence of local adaptation in both chimpanzees and humans.

For GYP, two candidate SNPs, chr4:145040845 and chr4:145039806 [*Western* 0.5 and 0.05% tails,

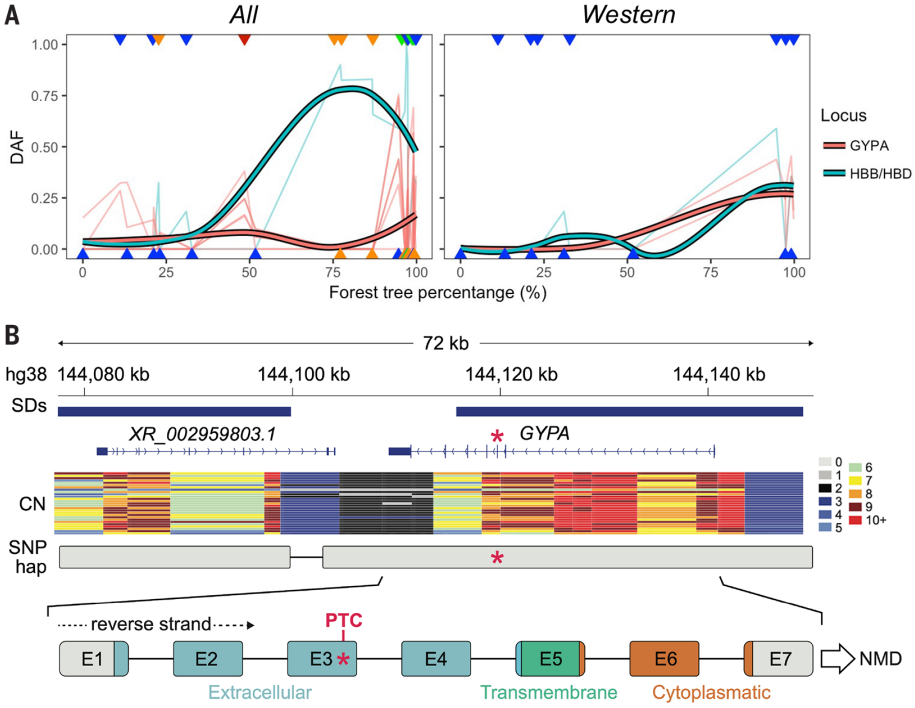


Fig. 6. Key malaria-related forest candidate genes. (**A**) Derived allele frequencies (DAF) of candidate SNPs at the *HBB/HBD* (green) and *GYPA* (red) loci plotted against forest-tree-percentage, with population values indicated with triangles colored according to subspecies (green, central; orange, eastern; red, Nigeria-Cameroon; blue, western) arbitrarily assigned to the top or bottom of the graph to reduce overlap. (Left) Candidate SNPs and populations from *All*. (Right) Candidate SNPs and populations from *Western*. Thin lines represent the estimated population allele frequencies for each candidate SNP, and thick lines show the smoothed pattern of all candidate SNPs per locus using LOESS. (**B**) Diagram of the *GYPA* locus in hg38 coordinates, including segmental duplications (SDs), copy numbers (CNs) across captive chimpanzees, representative long-read sequencing haplotype-containing candidate SNP C to A at chr4:145040845 in hg19 coordinates (red asterisk), and schematic representation of the candidate SNP location within *GYPA* exons (E1 to E7). PTC, premature termination codon; NMD, nonsense-mediated mRNA decay. The PTC SNP is 210 exonic base pairs upstream of the last exon-exon junction (between exons 6 and 7) and, therefore, likely to cause NMD according to the 50- to 55-nt rule (*121*).

respectively; PGLS $P = 6.92 \times 10^{-3}$ and 1.84×10^{-4}; supplementary text, section 6.4.5 (*82*)], lie within glycophorin A (*GYPA*) [Ensemble hg19 also places them within an intron of glycophorin B (*GYPB*), likely owing to an annotation error; see supplementary text, section 8.2 (*82*)]. The evidence for selection at this locus is very strong, with chr4:145039806 (FPR ≤ 2.91×10^{-4}) having the 23rd-highest BF in the *Western* exome data while accounting for read depth (and 16th-highest BF in nongenic-chr21). In *All*, there are six forest candidate SNPs in *GYPA*, including chr4:145040845 (0.05% tail) and chr4:145039806 (0.5% tail), and another *GYPA* SNP is a candidate in the genetics-only 0.5% tail (fig. S64).

GYPA and *GYPB* encode glycophorins used by *Plasmodium falciparum* to enter erythrocytes (*111*). In humans, structural variants associated with glycophorins mediate adaptation to malaria (*108*, *109*, *112*–*116*); therefore, we investigated structural variation at this locus in chimpanzees. Read depth in the PanAf exomes is not unusual at this locus, but low-coverage target capture data are not ideal for investigating structural variation. Copy number (CN) estimates from high-coverage short-read (n = 60 individuals) (*5*, *62*) and long-read data (n = 2) (*117*, *118*) from captive chimpanzees confirm that, in addition to the full-length and likely ancestral *GYPA*, chimpanzees also carry two to nine copies of truncated paralogues lacking the last two exons of *GYPA*, which encode for the cytoplasmic domain (*119*) [supplementary text, section 8.1 (*82*)] (Fig. 6B and figs. S65D and S68). Thus, like in humans, structural variants contribute to the complexity of the locus in chimpanzees. We note that PanAf exome SNPs in this region are in Hardy-Weinberg equilibrium [supplementary text, section 5.2 (*82*)], have single-copy coverage, and show no evidence of an association between forest-tree-percentage and read depth [supplementary text, section 8.1 (*82*)] (fig. S66). Further, the *GYPA* candidate SNPs are also present in both the long-read (*117*, *118*) and high-coverage short-read (*5*, *62*) data (fig. S67), confirming them to be true polymorphisms.

The long-read data (*117*, *118*, *120*) show the *GYPA* candidate SNPs residing in a single haplotype spanning the full-length gene (Fig. 6B and fig. S65B). The candidate allele at chr4:145040845 introduces a premature stop gain in exon 3 of *GYPA* (E76X), which is predicted to result in degradation of the mRNA by nonsense-mediated decay (*121*). Even if the truncated protein was translated, it would encode only a partial extracellular domain and be missing the remaining extracellular and entire transmembrane and cytoplasmic domains (*119*), resulting in nonfunctional GYPA. Thus, as suggested for *GYPB* deletions in humans (*113*), this *GYPA* SNP may have been selected for because it prevents the expression of a key receptor protein used by the malaria parasite to enter erythrocytes (*111*). Functional studies will be required to experimentally verify this hypothesis.

Conclusions

We present the largest population genomic study of natural selection in a nonhuman ape to date, capturing and sequencing the exome from noninvasive samples of hundreds of wild chimpanzees and integrating these exomes with previously published full chr21 sequences from the same samples (*3*). Even in the face of limitations from noninvasive sampling [supplementary text, section 1 (*82*)], this work demonstrates that population genomics can reveal the presence of local genetic adaptation in an endangered species.

The genotype-environment association analysis provides strong, genome-scale evidence of local adaptation to habitat in *All* and *Central-Eastern*, although not in *Western*, likely because of their small long-term N_e (*5*, *62*). This demonstrates the power of GEA analyses to identify positive selection by revealing signatures of local adaptation in the form of subtle allele frequency changes correlating with a relevant covariable [supplementary text, section 6.4.4 (*82*)]. Indeed, the GEA candidate SNPs differ consistently in allele frequency with respect to habitat but generally do not have large frequency differences between populations (fig. S55A). This is consistent with local adaptation in chimpanzees being mostly polygenic and driven by soft sweeps, as observed in humans (*32*, *122*, *123*), and suggests the presence of complex genetic adaptations even in the absence of fixed differences among populations.

Our findings suggest that although behaviors such as tool use (*16*, *124*) and thermoregulatory behaviors (*24*) are important in mitigating environmental stressors, selective pressures associated with habitat still appear to drive genetic adaptation in chimpanzees. Thus, both behavioral flexibility and genetic adaptation may explain how chimpanzees inhabit such a range of habitats. Far from replacing genetic adaptation, behavioral adaptations may drive genetic changes through gene-culture coevolution (*125*),

as seen with human diets (*41*), whereby behavioral flexibility facilitates exposure to new selection pressures that later drive genetic adaptations.

The evidence of genetic adaptation in forests demonstrates the importance of new adaptations even in habitats with high availability of resources that support high population densities. This is perhaps because the struggle against the high pathogen load of lowland forests shapes the evolution of these populations. This is not surprising, as pathogens have been important selective pressures for chimpanzees over longer timescales (*65*, *66*). Today, infectious diseases are a major cause of chimpanzee population decline (*1*), and recent increased exposure to humans has led to an increase in deadly outbreaks caused by cross-species transmission (*126*). Our findings highlight the importance of genetic adaptation in shaping infectious disease mortality in chimpanzees and suggest that individuals are adapted to the pathogens present in their local habitat, emphasizing the dangers of displacement and environmental change.

Signatures of positive selection in malaria-related genes in forests are particularly notable. A range of malaria parasites infect wild chimpanzees, including three *Laverania* species closely related to *P. falciparum*, which originated in gorillas (*57*, *127*) and is now responsible for 90% of global malaria mortality in humans (*128*). However, the fitness effects of malaria in wild chimpanzees are poorly understood (*129*). Its high prevalence in wild populations (*127*) and the few studies of captive chimpanzees suggest that severe effects are rare (*130*–*133*). However, health impacts in the wild may be more severe than in captivity, as demonstrated by SIV infections that are largely asymptomatic in captivity (*58*, *134*, *135*) but that have fitness effects in the wild (*29*, *136*). Young chimpanzees and pregnant mothers are particularly susceptible to malaria infection (*137*, *138*), which may lead to higher morbidity and mortality, as observed in humans (*139*). Our findings indicate that malaria may have been an important selection pressure in the recent past and may have fitness effects in present-day wild populations. Although the use of noninvasive samples limits our ability to verify these signatures with additional tests, the fact that signatures of selection are found at the same few genes in chimpanzees and humans provides additional evidence that these are likely to be true targets of natural selection. Further, it demonstrates how understanding chimpanzee evolution can provide insights into human evolution and medicine.

Chimpanzees also appear to have adapted genetically to savannah habitats, although identifying key selective pressures and adaptive traits is more challenging. Genomic and environmental data from additional savannah populations would help address this question. This would provide insight into how our ancestors may have adapted to similar habitats and have important implications for the conservation of wild chimpanzees as their habitats become hotter and more seasonal under climate change (*2*).

Just as previous studies highlighted the importance of conserving behavioral diversity (*16*, *140*, *141*), we emphasize the importance of conserving adaptive genetic diversity across chimpanzees' ecological range to maintain their adaptive potential and ensure long-term survival in the wild (*46*, *74*). This is notable because direct anthropogenic destruction (*1*, *142*, *143*), climate change (*144*), and disease transmission (*126*, *145*) are rapidly changing the environments experienced by chimpanzees. We emphasize the need to consider local genetic adaptations when planning conservation efforts to ensure that individuals are adapted to the local environment; although not the focus of our study, this information may help identify populations that may be especially vulnerable to forecasted environmental change (*146*). More generally, this study demonstrates the value and promise of noninvasive sampling to investigate genetic adaptation in endangered species.

Methods summary

Fecal samples from wild chimpanzees were collected across the geographic range of all four subspecies as part of PanAf (*3*, *81*). Full exomes were capture-sequenced for 828 individuals. Samples were then filtered extensively to exclude those with exceptionally low coverage or evidence of contamination and to remove first-order relatives, leaving 388 samples. We integrated these exomes with whole chr21 sequences from the same samples, generated in the same way (*3*). Population allele frequencies were estimated from genotype likelihoods using ANGSD v0.933 (*147*) at high-confidence SNPs. A dataset of nongenic-chr21 SNPs (>1 kb from a gene) was generated to be used as a proxy for expectations under neutrality; we then investigated potential signatures of positive selection in the exomes.

We first tested for evidence of local genetic adaptation in the form of exceptionally large allele frequency differentiation using a genetics-only test performed using BayPass v2.2 (*84*) under the core model. BayPass was run separately on the exome and nongenic-chr21, the latter was used to generate a null distribution to select candidate targets of positive selection in the exome. We also tested for evidence of natural selection in the form of allele frequencies correlating with a measure of habitat in a GEA performed using the BayPass AUX model. As above, non-genic-chr21 was used to generate a null distribution. To verify the signatures of local adaptation at candidate SNPs, we also performed a PGLS using the pgls() function in the R package "caper" (*148*)

To biologically interpret the signatures positive selection, we performed a gene set enrichment analysis using Gowinda (*149*). We tested for enrichment of general gene sets, such as Gene Ontology (GO) categories (*150*), KEGG pathways (*151*) and human GWAS traits (*152*), and performed hypothesis-driven tests using pathogen- and dehydration-related gene sets. We analyzed in detail a particularly interesting malaria-related candidate gene, *GYPA*. We confirmed the validity of the SNPs, estimated copy-number variants and investigated the structure of haplotypes in the *GYPA* locus using the data above, previously published high-coverage short-read data from 60 captive chimpanzees (*5*, *62*) and previously published long-read data from two captive chimpanzees (*117*, *118*).

REFERENCES AND NOTES

1. T. Humle, F. Maisels, J. F. Oates, A. Plumtre, E. A. Williamson, *Pan troglodytes* (IUCN, 2016), e.T15933A102326672.
2. S. Lindshield *et al*., Chimpanzees (Pan troglodytes) in savanna landscapes. *Evol. Anthropol.* **30**, 399–420 (2021). doi: 10.1002/evan.21924; pmid: 34542218
3. C. Fontsere *et al*., Population dynamics and genetic connectivity in recent chimpanzee history. *Cell Genomics* **2**, (2022). doi: 10.1016/j.xgen.2022.100133; pmid: 35711737
4. M. K. Gonder *et al*., A new west African chimpanzee subspecies? *Nature* **388**, 337 (1997). doi: 10.1038/41005; pmid: 9237749
5. J. Prado-Martinez *et al*., Great ape genetic diversity and population history. *Nature* **499**, 471–475 (2013). doi: 10.1038/nature12228; pmid: 23823723
6. A. Plumptre, J. A. Hart, T. C. Hicks, S. Nixon, A. K. Piel, *Eastern Chimpanzee (Pan troglodytes schweinfurthii): Status Survey and Conservation Action Plan 2010-2020* (IUCN, 2010).
7. T. Humle *et al*., *Pan troglodytes ssp. Verus* (IUCN, 2016); https://doi.org/10.2305/iucn.uk.2016-2.rlts.t15935a17989872.en.
8. J. F. Oates *et al*., *Pan troglodytes ssp. Ellioti* (IUCN, 2015); https://doi.org/10.2305/iucn.uk.2016-2.rlts.t40014a17990330.en.
9. N. D. Gross-Camp, M. Masozera, B. A. Kaplin, Chimpanzee seed dispersal quantity in a tropical montane forest of Rwanda. *Am. J. Primatol.* **71**, 901–911 (2009). doi: 10.1002/ajp.20727; pmid: 19603417
10. R. W. Wrangham, C. A. Chapman, L. J. Chapman, Seed dispersal by forest chimpanzees in Uganda. *J. Trop. Ecol.* **10**, 355–368 (1994). doi: 10.1017/S0266467400008026
11. W. D. Aguado, H. S. Rogers, J. D. Pruetz, Chimpanzees as ecosystem service providers: Seed dispersal of an economically important plant resource. *Biotropica* **54**, 656–669 (2022). doi: 10.1111/btp.13080
12. K. L. van Leeuwen, R. A. Hill, A. H. Korstjens, Classifying Chimpanzee (Pan troglodytes) Landscapes Across Large-Scale Environmental Gradients in Africa. *Int. J. Primatol.* **41**, 800–821 (2020). doi: 10.1007/s10764-020-00164-5
13. C. D. Barratt *et al*., Quantitative estimates of glacial refugia for chimpanzees (Pan troglodytes) since the Last Interglacial (120,000 BP). *Am. J. Primatol.* **83**, e23320 (2021). doi: 10.1002/ajp.23320; pmid: 34402081
14. E. J. Scully *et al*., The ecology and epidemiology of malaria parasitism in wild chimpanzee reservoirs. *Commun. Biol.* **5**, 1020 (2022). doi: 10.1038/s42003-022-03962-0; pmid: 36167977
15. E. G. Wessling *et al*., Seasonal Variation in Physiology Challenges the Notion of Chimpanzees (Pan troglodytes verus) as a Forest-Adapted Species. *Front. Ecol. Evol.* **6**, 60 (2018). doi: 10.3389/fevo.2018.00060
16. A. K. Kalan *et al*., Environmental variability supports chimpanzee behavioural diversity. *Nat. Commun.* **11**, 4451 (2020). doi: 10.1038/s41467-020-18176-3; pmid: 32934202
17. J. D. Pruetz *et al*., New evidence on the tool-assisted hunting exhibited by chimpanzees (Pan troglodytes verus) in a savannah habitat at Fongoli, Sénégal. *R. Soc. Open Sci.* **2**, 140507 (2015). doi: 10.1098/rsos.140507; pmid: 26064638

18. C. Boesch, C. Boesch, H. Boesch-Achermann, *The Chimpanzees of the Taï Forest: Behavioural Ecology and Evolution* (Oxford Univ. Press, 2000).
19. L. V. Luncz, R. Mundry, C. Boesch, Evidence for cultural differences between neighboring chimpanzee communities. *Curr. Biol.* **22**, 922–926 (2012). doi: 10.1016/j.cub.2012.03.031; pmid: 22578420
20. H. Péter, K. Zuberbühler, C. Hobaiter, Well-digging in a community of forest-living wild East African chimpanzees (Pan troglodytes schweinfurthii). *Primates* **63**, 355–364 (2022). doi: 10.1007/s10329-022-00992-4; pmid: 35662388
21. K. J. Hockings *et al*., Tools to tipple: Ethanol ingestion by wild chimpanzees using leaf-sponges. *R. Soc. Open Sci.* **2**, 150150 (2015). doi: 10.1098/rsos.150150; pmid: 26543588
22. A. Whiten *et al*., CHARTING CULTURAL VARIATION IN CHIMPANZEES. *Behaviour* **138**, 1481–1516 (2001). doi: 10.1163/156853901317367717
23. H. S. Kühl *et al*., Chimpanzee accumulative stone throwing. *Sci. Rep.* **6**, 22219 (2016). doi: 10.1038/srep22219; pmid: 26923684
24. J. Pruetz, P. Bertolani, Chimpanzee (Pan troglodytes verus) Behavioral Responses to Stresses Associated with Living in a Savannah-Mosaic Environment: Implications for Hominin Adaptations to Open Habitats. *Paleoanthropology* **2009**, 252–262 (2009). doi: 10.4207/PA.2009.ART33
25. N. Tagg *et al*., Nocturnal activity in wild chimpanzees (Pan troglodytes): Evidence for flexible sleeping patterns and insights into human evolution. *Am. J. Phys. Anthropol.* **166**, 510–529 (2018). doi: 10.1002/ajpa.23478; pmid: 29989158
26. T. Kaur *et al*., Descriptive epidemiology of fatal respiratory outbreaks and detection of a human-related metapneumovirus in wild chimpanzees (Pan troglodytes) at Mahale Mountains National Park, Western Tanzania. *Am. J. Primatol.* **70**, 755–765 (2008). doi: 10.1002/ajp.20565; pmid: 18548512
27. T. Matsuzawa, Jokro: The death of a wild infant chimpanzee from respiratory disease. *Primates* **61**, 339–346 (2020). doi: 10.1007/s10329-020-00819-0; pmid: 32337666
28. H. J. Barbian *et al*., Destabilization of the gut microbiome marks the end-stage of simian immunodeficiency virus infection in wild chimpanzees. *Am. J. Primatol.* **80**, e22515 (2018). doi: 10.1002/ajp.22515; pmid: 26676710
29. B. F. Keele *et al*., Increased mortality and AIDS-like immunopathology in wild chimpanzees infected with SIVcpz. *Nature* **460**, 515–519 (2009). doi: 10.1038/nature08200; pmid: 19626114
30. E. G. Wessling, H. S. Kühl, R. Mundry, T. Deschner, J. D. Pruetz, The costs of living at the edge: Seasonal stress in wild savanna-dwelling chimpanzees. *J. Hum. Evol.* **121**, 1–11 (2018). doi: 10.1016/j.jhevol.2018.03.001; pmid: 29685749
31. F. M. Key, Q. Fu, F. Romagné, M. Lachmann, A. M. Andrés, Human adaptation and population differentiation in the light of ancient genomes. *Nat. Commun.* **7**, 10775 (2016). doi: 10.1038/ncomms10775; pmid: 26988143
32. J. S. Rees, S. Castellano, A. M. Andrés, The Genomics of Human Local Adaptation. *Trends Genet.* **36**, 415–428 (2020). doi: 10.1016/j.tig.2020.03.006; pmid: 32396835
33. S. Fan, M. E. B. Hansen, Y. Lo, S. A. Tishkoff, Going global by adapting local: A review of recent human adaptation. *Science* **354**, 54–59 (2016). doi: 10.1126/science.aaf5098; pmid: 27846491
34. M. Lopez *et al*., Genomic Evidence for Local Adaptation of Hunter-Gatherers to the African Rainforest. *Curr. Biol.* **29**, 2926–2935.e4 (2019). doi: 10.1016/j.cub.2019.07.013; pmid: 31402299
35. C. M. Couto-Silva *et al*., Indigenous people from Amazon show genetic signatures of pathogen-driven selection. *Sci. Adv.* **9**, eabo0234 (2023). doi: 10.1126/sciadv.abo0234; pmid: 36888716
36. G. F. Harrison *et al*., Natural selection contributed to immunological differences between hunter-gatherers and agriculturalists. *Nat. Ecol. Evol.* **3**, 1253–1264 (2019). doi: 10.1038/s41559-019-0947-6; pmid: 31358949
37. I. Hamid, K. L. Korunes, S. Beleza, A. Goldberg, Rapid adaptation to malaria facilitated by admixture in the human population of Cabo Verde. *eLife* **10**, e63177 (2021). doi: 10.7554/eLife.63177; pmid: 33393457
38. S. N. Kariuki, T. N. Williams, Human genetics and malaria resistance. *Hum. Genet.* **139**, 801–811 (2020). doi: 10.1007/s00439-020-02142-6; pmid: 32130487
39. G. H. Perry *et al*., Diet and the evolution of human amylase gene copy number variation. *Nat. Genet.* **39**, 1256–1260 (2007). doi: 10.1038/ng2123; pmid: 17828263
40. M. Fumagalli *et al*., Greenlandic Inuit show genetic signatures of diet and climate adaptation. *Science* **349**, 1343–1347 (2015). doi: 10.1126/science.aab2319; pmid: 26383953
41. R. P. Evershed *et al*., Dairying, diseases and the evolution of lactase persistence in Europe. *Nature* **608**, 336–345 (2022). doi: 10.1038/s41586-022-05010-7; pmid: 35896751
42. N. G. Jablonski, G. Chaplin, Colloquium paper: Human skin pigmentation as an adaptation to UV radiation. *Proc. Natl. Acad. Sci. U.S.A.* **107** (suppl. 2), 8962–8968 (2010). doi: 10.1073/pnas.0914628107; pmid: 20445093
43. A.-S. Malaspinas *et al*., A genomic history of Aboriginal Australia. *Nature* **538**, 207–214 (2016). doi: 10.1038/nature18299; pmid: 27654914
44. X. Qi, W. L. Chan, R. J. Read, A. Zhou, R. W. Carrell, Temperature-responsive release of thyroxine and its environmental adaptation in Australians. *Proc. Biol. Sci.* **281**, 20132747 (2014). doi: 10.1098/rspb.2013.2747; pmid: 24478298
45. N. D. Wolfe, C. P. Dunavan, J. Diamond, Origins of major human infectious diseases. *Nature* **447**, 279–283 (2007). doi: 10.1038/nature05775; pmid: 17507975
46. J. C. Teixeira, C. D. Huber, The inflated significance of neutral genetic diversity in conservation genetics. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2015096118 (2021). doi: 10.1073/pnas.2015096118; pmid: 33608481
47. S. P. Flanagan, B. R. Forester, E. K. Latch, S. N. Aitken, S. Hoban, Guidelines for planning genomic assessment and monitoring of locally adaptive variation to inform species conservation. *Evol. Appl.* **11**, 1035–1052 (2017). doi: 10.1111/eva.12569; pmid: 30026796
48. J. Moore, in *Great Ape Societies*, W. C. McGrew, L. F. Marchant, T. Nishida, Eds. (Cambridge Univ. Press, 1996), chap. 20.
49. L. F. Marchant, E. G. Wessling, S. M. Lindshield, Introduction to the Special Issue on Savanna Chimpanzees. *Int. J. Primatol.* **41**, 767–774 (2020). doi: 10.1007/s10764-020-00188-x
50. R. Potts, Environmental hypotheses of hominin evolution. *Am. J. Phys. Anthropol.* **27** (suppl. 27), 93–136 (1998). doi: 10.1002/(SICI)1096-8644(1998)107:27+%3C93::AID-AJPA5%3E3.0.CO;2-X; pmid: 9881524
51. M. Domínguez-Rodrigo, Is the "savanna hypothesis" a dead concept for explaining the emergence of the earliest hominins? *Curr. Anthropol.* **55**, 59–81 (2014). doi: 10.1086/674530
52. R. C. Drummond-Clarke *et al*., Wild chimpanzee behavior suggests that a savanna-mosaic habitat did not support the emergence of hominin terrestrial bipedalism. *Sci. Adv.* **8**, eadd9752 (2022). doi: 10.1126/sciadv.add9752; pmid: 36516260
53. Chimpanzee Sequencing and Analysis Consortium, Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**, 69–87 (2005). doi: 10.1038/nature04072; pmid: 16136131
54. S. Calvignac-Spencer, S. A. J. Leendertz, T. R. Gillespie, F. H. Leendertz, Wild great apes as sentinels and sources of infectious disease. *Clin. Microbiol. Infect.* **18**, 521–527 (2012). doi: 10.1111/j.1469-0691.2012.03816.x; pmid: 22448813
55. D. S. Azevedo *et al*., One Health Perspectives on New Emerging Viral Diseases in African Wild Great Apes. *Pathogens* **10**, 1283 (2021). doi: 10.3390/pathogens10101283; pmid: 34684232
56. B. F. Keele *et al*., Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science* **313**, 523–526 (2006). doi: 10.1126/science.1126531; pmid: 16728595
57. P. M. Sharp, L. J. Plenderleith, B. H. Hahn, Ape Origins of Human Malaria. *Annu. Rev. Microbiol.* **74**, 39–63 (2020). doi: 10.1146/annurev-micro-020518-115628; pmid: 32905751
58. P. M. Sharp, G. M. Shaw, B. H. Hahn, Simian immunodeficiency virus infection of chimpanzees. *J. Virol.* **79**, 3891–3902 (2005). doi: 10.1128/JVI.79.7.3891-3902.2005; pmid: 15767392
59. G. Silvestri, M. Paiardini, I. Pandrea, M. M. Lederman, D. L. Sodora, Understanding the benign nature of SIV infection in natural hosts. *J. Clin. Invest.* **117**, 3148–3154 (2007). doi: 10.1172/JCI33034; pmid: 17975656
60. D. L. Sodora *et al*., Toward an AIDS vaccine: Lessons from natural simian immunodeficiency virus infections of African nonhuman primate hosts. *Nat. Med.* **15**, 861–865 (2009). doi: 10.1038/nm.2013; pmid: 19661993
61. M. Paiardini, I. Pandrea, C. Apetrei, G. Silvestri, Lessons learned from the natural hosts of HIV-related viruses. *Annu. Rev. Med.* **60**, 485–495 (2009). doi: 10.1146/annurev.med.60.041807.123753; pmid: 19630581
62. M. de Manuel *et al*., Chimpanzee genomic diversity reveals ancient admixture with bonobos. *Science* **354**, 477–481 (2016). doi: 10.1126/science.aag2602; pmid: 27789843
63. T. Bataillon *et al*., Inference of purifying and positive selection in three subspecies of chimpanzees (Pan troglodytes) from exome sequencing. *Genome Biol. Evol.* **7**, 1122–1132 (2015). doi: 10.1093/gbe/evv058; pmid: 25829516
64. C. Hvilsom *et al*., Extensive X-linked adaptive evolution in central chimpanzees. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 2054–2059 (2012). doi: 10.1073/pnas.1106877109; pmid: 22308321
65. J. M. Schmidt, M. de Manuel, T. Marques-Bonet, S. Castellano, A. M. Andrés, The impact of genetic adaptation on chimpanzee subspecies differentiation. *PLOS Genet.* **15**, e1008485 (2019). doi: 10.1371/journal.pgen.1008485; pmid: 31765391
66. H. Pawar, H. J. Ostridge, J. M. Schmidt, A. M. Andrés, Genetic adaptations to SIV across chimpanzee populations. *PLOS Genet.* **18**, e1010337 (2022). doi: 10.1371/journal.pgen.1010337; pmid: 36007015
67. A. Cagan *et al*., Natural Selection in the Great Apes. *Mol. Biol. Evol.* **33**, 3268–3283 (2016). doi: 10.1093/molbev/msw215; pmid: 27795229
68. J. Nye *et al*., Selection in the Introgressed Regions of the Chimpanzee Genome. *Genome Biol. Evol.* **10**, 1132–1138 (2018). doi: 10.1093/gbe/evy077; pmid: 29635458
69. J. Nye, M. Mondal, J. Bertranpetit, H. Laayouni, A fully integrated machine learning scan of selection in the chimpanzee genome. *NAR Genom. Bioinform.* **2**, lqaa061 (2020). doi: 10.1093/nargab/lqaa061; pmid: 33575612
70. S. Kovalaskas, J. K. Rilling, J. Lindo, Comparative analyses of the Pan lineage reveal selection on gene pathways associated with diet and sociality in bonobos. *Genes Brain Behav.* **20**, e12715 (2021). doi: 10.1111/gbb.12715; pmid: 33200560
71. S. Han, A. M. Andrés, T. Marques-Bonet, M. Kuhlwilm, Genetic Variation in Pan Species Is Shaped by Demographic History and Harbors Lineage-Specific Functions. *Genome Biol. Evol.* **11**, 1178–1191 (2019). doi: 10.1093/gbe/evz047; pmid: 30847478
72. K. Nam *et al*., Evidence that the rate of strong selective sweeps increases with population size in the great apes. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 1613–1618 (2017). doi: 10.1073/pnas.1605660114; pmid: 28137852
73. K. Nam *et al*., Extreme selective sweeps independently targeted the X chromosomes of the great apes. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6413–6418 (2015). doi: 10.1073/pnas.1419306112; pmid: 25941379
74. P. A. Hohenlohe, W. C. Funk, O. P. Rajora, Population genomics for wildlife conservation and management. *Mol. Ecol.* **30**, 62–82 (2021). doi: 10.1111/mec.15720; pmid: 33145846
75. L. C. White *et al*., A roadmap for high-throughput sequencing studies of wild animal populations using noninvasive samples and hybridization capture. *Mol. Ecol. Resour.* **19**, 609–622 (2019). doi: 10.1111/1755-0998.12993; pmid: 30637963
76. P. Frandsen *et al*., Targeted conservation genetics of the endangered chimpanzee. *Heredity* **125**, 15–27 (2020). doi: 10.1038/s41437-020-0313-0; pmid: 32346130
77. C. Fontsere *et al*., Maximizing the acquisition of unique reads in noninvasive capture sequencing experiments. *Mol. Ecol. Resour.* **21**, 745–761 (2021). doi: 10.1111/1755-0998.13300; pmid: 33217149
78. A. T. Ozga *et al*., Urine as a high-quality source of host genomic DNA from wild populations. *Mol. Ecol. Resour.* **21**, 170–182 (2021). doi: 10.1111/1755-0998.13260; pmid: 32985084
79. G. H. Perry, J. C. Marioni, P. Melsted, Y. Gilad, Genomic-scale capture and sequencing of endogenous DNA from feces. *Mol. Ecol.* **19**, 5332–5344 (2010). doi: 10.1111/j.1365-294X.2010.04888.x; pmid: 21054605
80. J. Hernandez-Rodriguez *et al*., The impact of endogenous content, replicates and pooling on genome capture from faecal samples. *Mol. Ecol. Resour.* **18**, 319–333 (2018). doi: 10.1111/1755-0998.12728; pmid: 29058768
81. J. D. Lester *et al*., Recent genetic connectivity and clinal variation in chimpanzees. *Commun. Biol.* **4**, 283 (2021). doi: 10.1038/s42003-021-01806-x; pmid: 33674780
82. Supplementary materials.
83. S. Besenbacher, C. Hvilsom, T. Marques-Bonet, T. Mailund, M. H. Schierup, Direct estimation of mutations in great apes reconciles phylogenetic dating. *Nat. Ecol. Evol.* **3**, 286–292 (2019). doi: 10.1038/s41559-018-0778-x; pmid: 30664699
84. M. Gautier, Genome-Wide Scan for Adaptive Divergence and Association with Population-Specific Covariates. *Genetics*

201, 1555–1579 (2015). doi: 10.1534/genetics.115.181453; pmid: 26482796
85. L. Olazcuaga *et al*., A Whole-Genome Scan for Association with Invasion Success in the Fruit Fly Drosophila suzukii Using Contrasts of Allele Frequencies Corrected for Population Structure. *Mol. Biol. Evol.* **37**, 2369–2385 (2020). doi: 10.1093/molbev/msaa098; pmid: 32302396
86. P. Pavlidis, J. D. Jensen, W. Stephan, A. Stamatakis, A critical assessment of storytelling: Gene ontology categories and the importance of validating genomic scans. *Mol. Biol. Evol.* **29**, 3237–3248 (2012). doi: 10.1093/molbev/mss136; pmid: 22617950
87. A. R. Barton *et al*., Insufficient evidence for natural selection associated with the Black Death. bioRxiv 2023.03.14.532615 [Preprint] (2023); https://doi.org/10.1101/2023.03.14.532615.
88. C. W. Ahrens *et al*., The search for loci under selection: Trends, biases and progress. *Mol. Ecol.* **27**, 1342–1356 (2018). doi: 10.1111/mec.14549; pmid: 29524276
89. S. De Mita *et al*., Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Mol. Ecol.* **22**, 1383–1399 (2013). doi: 10.1111/mec.12182; pmid: 23294205
90. J. C. Aleman *et al*., Floristic evidence for alternative biome states in tropical Africa. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 28183–28190 (2020). doi: 10.1073/pnas.2011515117; pmid: 33109722
91. H. Jeffreys, *Theory of Probability* (Oxford Univ. Press, ed. 3, 1961).
92. L. Kordonowy, M. MacManes, Characterizing the reproductive transcriptomic correlates of acute dehydration in males in the desert-adapted rodent, Peromyscus eremicus. *BMC Genomics* **18**, 473 (2017). doi: 10.1186/s12864-017-3840-1; pmid: 28645248
93. F. M. Giorello *et al*., An association between differential expression and genetic divergence in the Patagonian olive mouse (Abrothrix olivacea). *Mol. Ecol.* **27**, 3274–3286 (2018). doi: 10.1111/mec.14778; pmid: 29940092
94. C. Boesch, G. Hohmann, L. Marchant, *Behavioural Diversity in Chimpanzees and Bonobos* (Cambridge Univ. Press, 2002). doi: 10.1017/CBO9780511606397
95. R. S. Rudicell *et al*., High prevalence of simian immunodeficiency virus infection in a community of savanna chimpanzees. *J. Virol.* **85**, 9918–9928 (2011). doi: 10.1128/JVI.05475-11; pmid: 21775446
96. Y. Li *et al*., Eastern chimpanzees, but not bonobos, represent a simian immunodeficiency virus reservoir. *J. Virol.* **86**, 10776–10791 (2012). doi: 10.1128/JVI.01498-12; pmid: 22837215
97. J. Klunk *et al*., Evolution of immune genes is associated with the Black Death. *Nature* **611**, 312–319 (2022). doi: 10.1038/s41586-022-05349-x; pmid: 36261521
98. M. Deschamps *et al*., Genomic Signatures of Selective Pressures and Introgression from Archaic Hominins at Human Innate Immunity Genes. *Am. J. Hum. Genet.* **98**, 5–21 (2016). doi: 10.1016/j.ajhg.2015.11.014; pmid: 26748513
99. M. Fumagalli *et al*., Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLOS Genet.* **7**, e1002355 (2011). doi: 10.1371/journal.pgen.1002355; pmid: 22072984
100. J. T. Daub *et al*., Evidence for polygenic adaptation to pathogens in the human genome. *Mol. Biol. Evol.* **30**, 1544–1558 (2013). doi: 10.1093/molbev/mst080; pmid: 23625889
101. E. R. Ebel, F. A. Kuypers, C. Lin, D. A. Petrov, E. S. Egan, Common host variation drives malaria parasite fitness in healthy human red cells. *eLife* **10**, e69808 (2021). doi: 10.7554/eLife.69808; pmid: 34553687
102. E. R. Ebel, N. Telis, S. Venkataram, D. A. Petrov, D. Enard, High rate of adaptation of mammalian proteins that interact with Plasmodium and related parasites. *PLOS Genet.* **13**, e1007023 (2017). doi: 10.1371/journal.pgen.1007023; pmid: 28957326
103. M. C. Hansen *et al*., High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013). doi: 10.1126/science.1244693; pmid: 24233722
104. G. Band *et al*., Insights into malaria susceptibility using genome-wide data on 17,000 individuals from Africa, Asia and Oceania. *Nat. Commun.* **10**, 5732 (2019). doi: 10.1038/s41467-019-13480-z; pmid: 31844061
105. E. R. Ebel, L. H. Uricchio, D. A. Petrov, E. S. Egan, Revisiting the malaria hypothesis: Accounting for polygenicity and pleiotropy. *Trends Parasitol.* **38**, 290–301 (2022). doi: 10.1016/j.pt.2021.12.007; pmid: 35065882
106. A. C. Allison, Protection afforded by sickle-cell trait against subtertian malareal infection. *BMJ* **1**, 290–294 (1954). doi: 10.1136/bmj.1.4857.290; pmid: 13115700
107. L. Pauling, H. A. Itano, S. J. Singer, I. C. Wells, Sickle cell anemia a molecular disease. *Science* **110**, 543–548 (1949). doi: 10.1126/science.110.2865.543; pmid: 15395398
108. E. M. Leffler *et al*., Resistance to malaria through structural variation of red blood cell invasion receptors. *Science* **356**, eaam6393 (2017). doi: 10.1126/science.aam6393; pmid: 28522690
109. W. Algady *et al*., The Malaria-Protective Human Glycophorin Structural Variant DUP4 Shows Somatic Mosaicism and Association with Hemoglobin Levels. *Am. J. Hum. Genet.* **103**, 769–776 (2018). doi: 10.1016/j.ajhg.2018.10.008; pmid: 30388403
110. M. H. Steinberg, J. G. Adams III, Hemoglobin A2: Origin, evolution, and aftermath. *Blood* **78**, 2165–2177 (1991). pmid: 1932737
111. E. Jaskiewicz, M. Jodłowska, R. Kaczmarek, A. Zerka, Erythrocyte glycophorins as receptors for Plasmodium merozoites. *Parasit. Vectors* **12**, 317 (2019). doi: 10.1186/s13071-019-3575-8; pmid: 31234897
112. S. N. Kariuki *et al*., Red blood cell tension protects against severe malaria in the Dantu blood group. *Nature* **585**, 579–583 (2020). doi: 10.1038/s41586-020-2726-6; pmid: 32939086
113. C. Gassner *et al*., Two prevalent ~100-kb GYPB deletions causative of the GPB-deficient blood group MNS phenotype S-s-U- in black Africans. *Transfus. Med. Hemother.* **47**, 326–336 (2020). doi: 10.1159/000504946; pmid: 32884505
114. E. Tokunaga *et al*., Two apparently healthy Japanese individuals of type MkMk have erythrocytes which lack both the blood group MN and Ss-active sialoglycoproteins. *J. Immunogenet.* **6**, 383–390 (1979). doi: 10.1111/j.1744-313X.1979.tb00693.x; pmid: 521666
115. T. J. Hadley *et al*., Falciparum malaria parasites invade erythrocytes that lack glycophorin A and B (MkMk). Strain differences indicate receptor heterogeneity and two pathways for invasion. *J. Clin. Invest.* **80**, 1190–1193 (1987). doi: 10.1172/JCI113178; pmid: 3308959
116. E. J. Hollox, S. Louzada, Genetic variation of glycophorins and infectious disease. *Immunogenetics* **75**, 201–206 (2023). doi: 10.1007/s00251-022-01280-7; pmid: 36224278
117. D. C. Soto *et al*., Identification of structural variation in chimpanzees using optical mapping and nanopore sequencing. *Genes* **11**, 276 (2020). doi: 10.3390/genes11030276; pmid: 32143403
118. Z. N. Kronenberg *et al*., High-resolution comparative analysis of great ape genomes. *Science* **360**, eaar6343 (2018). doi: 10.1126/science.aar6343; pmid: 29880660
119. A. W. Bigham, K. Magnaye, D. M. Dunn, R. B. Weiss, M. Bamshad, Complex signatures of natural selection at GYPA. *Hum. Genet.* **137**, 151–160 (2018). doi: 10.1007/s00439-018-1866-3; pmid: 29362874
120. Y. Mao *et al*., Structurally divergent and recurrently mutated regions of primate genomes. *Cell* **187**, 1547–1562.e13 (2024). doi: 10.1016/j.cell.2024.01.052; pmid: 38428424
121. E. Nagy, L. E. Maquat, A rule for termination-codon position within intron-containing genes: When nonsense affects RNA abundance. *Trends Biochem. Sci.* **23**, 198–199 (1998). doi: 10.1016/S0968-0004(98)01208-0; pmid: 9644970
122. J. K. Pritchard, J. K. Pickrell, G. Coop, The genetics of human adaptation: Hard sweeps, soft sweeps, and polygenic adaptation. *Curr. Biol.* **20**, R208–R215 (2010). doi: 10.1016/j.cub.2009.11.055; pmid: 20178769
123. D. R. Schrider, A. D. Kern, Soft Sweeps Are the Dominant Mode of Adaptation in the Human Genome. *Mol. Biol. Evol.* **34**, 1863–1877 (2017). doi: 10.1093/molbev/msx154; pmid: 28482049
124. A. Whiten *et al*., Cultures in chimpanzees. *Nature* **399**, 682–685 (1999). doi: 10.1038/21415; pmid: 10385119
125. H. Whitehead, K. N. Laland, L. Rendell, R. Thorogood, A. Whiten, The reach of gene-culture coevolution in animals. *Nat. Commun.* **10**, 2405 (2019). doi: 10.1038/s41467-019-10293-y; pmid: 31160560
126. K. Gilardi *et al*., *Best practice guidelines for health monitoring and disease control in great ape populations* (IUCN/SSC Primate Specialist Group, 2015).
127. J. C. Rayner, W. Liu, M. Peeters, P. M. Sharp, B. H. Hahn, A plethora of Plasmodium species in wild apes: A source of human infection? *Trends Parasitol.* **27**, 222–229 (2011). doi: 10.1016/j.pt.2011.01.006; pmid: 21354860
128. R. W. Snow, Global malaria eradication and the importance of Plasmodium falciparum epidemiology in Africa. *BMC Med.* **13**, 23 (2015). doi: 10.1186/s12916-014-0254-7; pmid: 25644195
129. H. M. De Nys, T. Löhrich, D. Wu, S. Calvignac-Spencer, F. H. Leendertz, Wild African great apes as natural hosts of malaria parasites: Current knowledge and research perspectives. *Primate Biol.* **4**, 47–59 (2017). doi: 10.5194/pb-4-47-2017; pmid: 32110692
130. J. Rodhain, Contribution à l'étude des plasmodiums des singes africains. *Ann. Soc. Belg. Med. Trop.* **18**, 237–253 (1938).
131. J. Rodhain, Plasmadia of Central African Apes and their Relationship to Human Plasmodia. *Ann. Soc. Belg. Med. Trop.* **20**, 489-505 (1940).
132. T. Hayakawa *et al*., Identification of Plasmodium malariae, a human malaria parasite, in imported chimpanzees. *PLOS ONE* **4**, e7412 (2009). doi: 10.1371/journal.pone.0007412; pmid: 19823579
133. A. Herbert *et al*., Malaria-like symptoms associated with a natural Plasmodium reichenowi infection in a chimpanzee. *Malar. J.* **14**, 220 (2015). doi: 10.1186/s12936-015-0743-y; pmid: 26032157
134. N. G. de Groot, C. M. C. Heijmans, R. E. Bontrop, AIDS in chimpanzees: The role of MHC genes. *Immunogenetics* **69**, 499–509 (2017). doi: 10.1007/s00251-017-1006-6; pmid: 28695283
135. R. V. Gilden *et al*., HTLV-III antibody in a breeding chimpanzee not experimentally exposed to the virus. *Lancet* **1**, 678–679 (1986). doi: 10.1016/S0140-6736(86)91749-6; pmid: 2869366
136. R. S. Rudicell *et al*., Impact of simian immunodeficiency virus infection on chimpanzee population dynamics. *PLOS Pathog.* **6**, e1001116 (2010). doi: 10.1371/journal.ppat.1001116; pmid: 20886099
137. H. M. De Nys *et al*., Age-related effects on malaria parasite infection in wild chimpanzees. *Biol. Lett.* **9**, 20121160 (2013). doi: 10.1098/rsbl.2012.1160; pmid: 23720517
138. H. M. De Nys *et al*., Malaria parasite detection increases during pregnancy in wild chimpanzees. *Malar. J.* **13**, 413 (2014). doi: 10.1186/1475-2875-13-413; pmid: 25331753
139. World Health Organization, *World Malaria Report 2022* (WHO, 2022).
140. H. S. Kühl *et al*., Human impact erodes chimpanzee behavioral diversity. *Science* **363**, 1453–1455 (2019). doi: 10.1126/science.aau4532; pmid: 30846610
141. S. Carvalho *et al*., Using nonhuman culture in conservation requires careful and concerted action. *Conserv. Lett.* **15**, e12860 (2022). doi: 10.1111/conl.12860
142. N. T. Laporte, J. A. Stabach, R. Grosch, T. S. Lin, S. J. Goetz, Expansion of industrial logging in Central Africa. *Science* **316**, 1451 (2007). doi: 10.1126/science.1141057; pmid: 17556578
143. J. Junker *et al*., Recent decline in suitable environmental conditions for African great apes. *Divers. Distrib.* **18**, 1077–1091 (2012). doi: 10.1111/ddi.12005
144. T. L. Graham, H. D. Matthews, S. E. Turner, A global-scale evaluation of primate exposure and vulnerability to climate change. *Int. J. Primatol.* **37**, 158–174 (2016). doi: 10.1007/s10764-016-9890-4
145. S. Köndgen *et al*., Pandemic human viruses cause decline of endangered great apes. *Curr. Biol.* **18**, 260–264 (2008). doi: 10.1016/j.cub.2008.01.012; pmid: 18222690

ACKNOWLEDGMENTS

This manuscript is dedicated to the enduring memory of Christophe Boesch, whose holistic approach to studying chimpanzee behavior, evolution and conservation was a driving force behind the PanAf project. We thank H. Burbano, M. Reuter, and all the members of the Andrés and Burbano groups for useful comments and suggestions. We also thank J. Wang for his support and supervision during H.J.O.'s PhD, D. Lavelle for assistance in rebase calling Oxford Nanopore Technology (ONT) data using guppy 5.0.11, and Agilent for their collaboration in the project. We thank L. Llovera for assistance in the laboratory. We thank T. Aebischer, F. Aubert, E. Ayuk Ayimisin, A. Bamba, M. Bonnet, C. Cipoletta, K. Corogenes, B. Curran, L. D'Auvergne, J. C. Dengui, T. Desarmeaux, K. Dierks, E. Dilambaka, D. Dowd, A. Dunn, M. Eno-Nku, T. Freeman, A. Goedmakers, J. Hart, M. Ter Heegde, I. Imong, M. Kambi, L. Kehoe, V. Lapeyre, V. Leinert, J. M. Linder, S. Marrocoli, M. Masozera, R. McElreath, V. Mihindou, G. Muhanguzi, F. Mulindahabi, M. Murai, P. Niyigaba, N. W. Njomen, N. Ntare, B. Perodeau, A. M. Siaka, A. Tickle, E. Ton, R. Tshombe, H. Vanleeuwe, P. Álvarez Varona, V. Vergnes, M. Kambale Vyalengerera, and K. Zuberbuehler for assistance in field site coordination and sample collection. We thank the field assistants and volunteers from the

Jane Goodall Institute Spain in Senegal for their help with sample collection in Dindefelo and the Greater Mahale Ecosystem Research and Conservation (GMERC) field assistants who provide crucial support and data collection at the Issa Valley field site in Tanzania. We thank the following government agencies for their support in conducting field research in their countries: Ministère de la Recherche Scientifique et de l'Innovation, Cameroon; Ministère des Forêts et de la Faune, Cameroon; Ministère des Eaux et Forêts, Cote d'Ivoire; Ministère de l'Enseignement Supérieur et de la Recherche Scientifique in Côte d'Ivoire; Institut Congolais pour la Conservation de la Nature, DR-Congo; Ministère de la Recherche Scientifique, Democratic Republic of the Congo; Agence Nationale des Parcs Nationaux, Gabon; Center National de la Recherche Scientifique (CENAREST), Gabon; Société Equatoriale d'Exploitation Forestière (SEEF), Gabon; Department of Wildlife and Range Management, Ghana; Forestry Commission, Ghana; Ministère de l'Agriculture de l'Elevage et des Eaux et Forets, Guinea; Instituto da Biodiversidade e das Áreas Protegidas (IBAP), Guinea-Bissau; Ministro da Agricultura e Desenvolvimento Rural, Guinea-Bissau; Forestry Development Authority, Liberia; Eaux et Forets, Mali; Ministre de l'Environnement et de l'Assainissement et du Developpement Durable du Mali; Conservation Society of Mbe Mountains (CAMM), Nigeria; National Park Service, Nigeria; Ministère de l'Economie Forestière, R-Congo; Ministère de le Recherche Scientifique et Technologique, R-Congo; Ministry of Education, Rwanda; Rwanda Development Board, Rwanda; Direction des Eaux, Forêts, et Chasses, Senegal; Reserve Naturelle Communautaire de Dindefelo, Senegal; Ministry of Agriculture, Forestry, and Food Security, Sierra Leone; National Protected Area Authority, Sierra Leone; Tanzania Commission for Science and Technology, Tanzania; Tanzania Wildlife Research Institute, Tanzania; Makerere University Biological Field Station (MUBFS), Uganda; Uganda National Council for Science and Technology (UNCST), Uganda; Uganda Wildlife Authority, Uganda; National Forestry Authority, Uganda; Agence Congolaise de la Faune et des Aires Protégées; and The Wild Chimpanzee Foundation (WCF). **Funding:** H.J.O. is supported by the Natural Environment Research Council through the London NERC Doctoral Training Programme scholarship, grant no. NE/L002485/1. A.M.A. is supported by UCL's Wellcome Trust ISSF3 award no. 204841/Z/16/Z, the BBSRC grant no. BB/W007703/1, and the Leverhulme trust grant no. RPG-2021-414. The Pan African Program: The Cultured Chimpanzee (PanAf) is generously funded by the Max Planck Society, the Max Planck Society Innovation Fund, and the Heinz L. Krekeler Foundation. T.M.-B. is supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 864203), grant PID2021-126004NB-100 funded by MICIU/AEI/ 10.13039/501100011033 and ERDF/EU (MICIIN/FEDER, UE), and Secretaria d'Universitats i Recerca and CERCA Programme del Departament d'Economia i Coneixement de la Generalitat de Catalunya (GRC 2021 SGR 00177). We also acknowledge funding from the "la Caixa" Foundation doctoral fellowship program LCF/BQ/DE15/ 10360006 (to C.F.), US National Institutes of Health (NIH) grants from the Office of the Director and National Institute of Mental Health (DP2MH119424 and R01MH132818 to MYD), Formación de Personal Investigador (FPI) PRE2018- 083966 from Ministerio de Ciencia, Universidades e Investigacion (to M.A.-E.), and PID2020-116908GB-I00 (MCIN) (to E.L.). F.A.S. and A.K.P. thank the UCSD/Salk Center for Academic Research and Training in Anthropogeny (CARTA) and Department of Human Origins, Max Planck Institute for Evolutionary Anthropology for support of GMERC, the Issa Valley field site, and ongoing sample collection in Tanzania. The research carried out in the Ituri province, Democratic Republic of the Congo, was supported by AL's National Geographic Explorer's grant (NGS-61178C-19). **Author contributions:** Conceptualization: A.M.A., T.M.-B., M.A., C.B.; Formal analysis: H.J.O., C.F., D.C.S., A.M.A., M.Y.D., T.M.-B.; Methodology: H.J.O., A.M.A., C.F., D.C.S., M.Y.D., J.M.S., T.M.-B.; Project design and planning: J.M.S., L.V., C.B., C.F., E.L., T.M.-B., M.A., H.S.K., A.M.A.; Bioinformatic processing: H.J.O., C.F., D.C.S.; Interpretation of results: H.J.O., A.M.A., C.F., J.M.S., C.D.B., P.G., A.P., H.S.K., T.M.-B., M.A., M.Y.D., D.C.S.; Supervision or execution of data collection: C.F., E.L., M.A.-E., P.D., A.A., S.A., A.K.A., E.B., D.B., M.B., G.B., R.C., H.C., C.C., E.D., A.-C.G., J.H., D.H., V.H., R.A.H-.A., K.J.J., S.J., J.J., P.K., Mi.K., A.K.K., Mb.K., I.K., D.K., K.E.L., J.L., B.L., A.L., K.C.L., M.L., G.M., R.M., A.C.M., D.M., Emi.N., So.N., St.N., Emm.N., C.O., L.J.O., R.O., L.P., J.P., S.R., M.M.R., A.R., C.S., L.S., V.S., F.A.S., N.T., L.R.T., J.v.S., E.V., E.C.W., J.W., R.M.W., Y.G.Y., K.K.Y., L.V., A.K.P., C.B., H.S.K., M.A.; Visualization: H.J.O., D.C.S.; Funding acquisition: T.M.-B., M.A., A.M.A., M.Y.D., H.S.K., C.B.; Project administration: A.M.A., M.A., H.S.K., T.M-.B., C.B. Supervision: A.M.A., M.A., T.M-.B.; Writing – original draft: H.J.O., A.M.A.; Writing – review and editing: all authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All genomic data generated will be available at a public repository (ENA) under the accession code ENA: PRJEB76176. All other relevant data will be available at the UCL Research Data Repository (*153*). All code is available at https://github.com/HarrJO/PanAf-local-adaptation.git and permanently archived at Zenodo (*154*, *155*); the latter contains the code for the *GYPA* structural analyses.

SUPPLEMENTARY MATERIALS

Submitted 6 January 2024; accepted 22 October 2024
10.1126/science.adn7954