Current Biology 19, 312–318, February 24, 2009 ©2009 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2008.12.049

Report

Origins and Genetic Diversity of Pygmy Hunter-Gatherers from Western Central Africa

Paul Verdu,[1,*] Frederic Austerlitz,[2,3,4] Arnaud Estoup,[5] Renaud Vitalis,[1] Myriam Georges,[1] Sylvain Théry,[1] Alain Froment,[1] Sylvie Le Bomin,[1] Antoine Gessain,[6] Jean-Marie Hombert,[7] Lolke Van der Veen,[7] Lluis Quintana-Murci,[8] Serge Bahuchet,[1] and Evelyne Heyer[1]

[1]Ecoanthropology and Ethnobiology
UMR 5145
CNRS-MNHN-Université Paris 7
Musée de l'Homme
75016 Paris
France

[2]Laboratoire Ecologie Systématique et Evolution
CNRS UMR 8079
F-91405 Orsay
France

[3]Université Paris-Sud
F-91405 Orsay
France

[4]AgroParisTech
F-75231 Paris
France

[5]Institut National de la Recherche Agronomique (INRA)
UMR CBGP (INRA/IRD/Cirad/Montpellier SupAgro)
Campus International de Baillarguet, CS 30016
F-34988 Montferrier-sur-Lez cedex
France

[6]Unité d'Epidémiologie et Physiopathologie des Virus Oncogènes
Institut Pasteur
75015 Paris
France

[7]UMR 5596 Institut des Sciences de l'Homme
69363 Lyon
France

[8]Human Evolutionary Genetics Unit
CNRS URA3012
Institut Pasteur
75015 Paris
France

Summary

Central Africa is currently peopled by numerous sedentary agriculturalist populations neighboring the largest group of mobile hunter-gatherers, the Pygmies [1–3]. Although archeological remains attest to *Homo sapiens'* presence in the Congo Basin for at least 30,000 years, the demographic history of these groups, including divergence and admixture, remains widely unknown [4–6]. Moreover, it is still debated whether common history or convergent adaptation to a forest environment resulted in the short stature characterizing the pygmies [2, 7]. We genotyped 604 individuals at 28 autosomal tetranucleotide microsatellite loci in 12 nonpygmy and 9 neighboring pygmy populations. We found a high level of genetic heterogeneity among Western Central African pygmies, as well as evidence of heterogeneous levels of asymmetrical gene flow from nonpygmies to pygmies, consistent with the variable sociocultural barriers against intermarriages. Using approximate Bayesian computation (ABC) methods [8], we compared several historical scenarios. The most likely points toward a unique ancestral pygmy population that diversified ~2800 years ago, contemporarily with the Neolithic expansion of nonpygmy agriculturalists [9, 10]. Our results show that recent isolation, genetic drift, and heterogeneous admixture enabled a rapid and substantial genetic differentiation among Western Central African pygmies. Such an admixture pattern is consistent with the various sociocultural behaviors related to intermariages between pygmies and nonpygmies.

Results

Genetic Variation within and between Populations

We first compared genetic variation with and among pygmy and nonpygmy populations in Western Central Africa (see Figure 1 and Table S1, available online). Expected heterozygosity in pygmies (H_e = 73.6%, SD = 1.2%; see Tables S3 and S4) was not significantly different from that in nonpygmies (H_e = 74.1%, SD = 0.7%; Wilcoxon rank sum test: p = 0.30). However, the number of alleles per population was significantly lower (p = 0.015) in pygmies (average of 6.40) than in nonpygmies (average of 6.62). A significant proportion of the total genetic variance was found between nonpygmy and pygmy populations (hierarchical AMOVA; p < 0.001). Overall, pygmy populations were considerably more differentiated (F_{ST} = 0.019, p value < 0.001) than nonpygmy populations (F_{ST} = 0.004, p value < 0.001), with all but one significant pairwise F_{ST} estimates between pygmy populations and a majority (68.2%) of nonsignificant pairwise F_{ST} estimates between nonpygmy populations (Table S5). Consistently, the 12 nonpygmy populations were tightly clustered in the principal component analysis (PCA), whereas the various pygmy populations were more scattered (Figure 2). As a noticeable exception, the two Bongo pygmy subsamples clustered together with the nonpygmies. Importantly, none of the pygmy populations clustered tightly with its immediate nonpygmy neighbor.

Asymmetrical Admixture from Nonpygmies to Pygmies

We performed individual multilocus genotype clustering analyses with STRUCTURE [11] and varied the number, K, of putative clusters. For K = 2, one cluster (in blue) mainly included individuals from nonpygmy populations (Figure 3), whose individuals had, on average, 84.3% (SD = 9.5%) of their genotype membership in the "Blue" cluster, thus showing little evidence of coancestry with individuals clustering in the alternative "Red" cluster, which mainly included pygmy individuals.

However, many of these pygmy individuals showed a substantial signal of admixture with the "Blue" cluster, which suggests asymmetrical gene flow from nonpygmies into pygmy populations. The proportion of genotypes belonging to the nonpygmy cluster varied among the pygmy groups: 44.2% (SD = 14.8%), on average, among the Baka individuals

*Correspondence: verdu@mnhn.fr

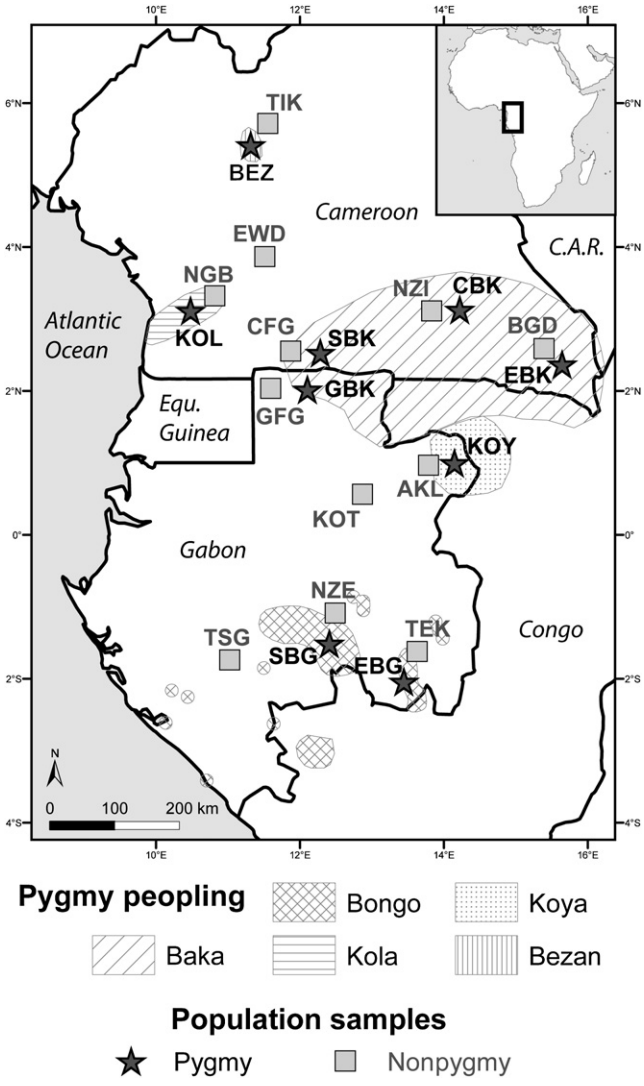

Figure 1. Population Samples

Geographical distribution of the 9 hunter-gatherer pygmy (stars) and the 12 neighboring agriculturist nonpygmy sample populations (gray squares). See Table S1 for population sample details. AKL, Akele (N = 13); BEZ, Bezan (N = 29); BGD, Bangando (N = 30); CBK, Central Baka (N = 29); CFG, Camerounese Fang (N = 30); EBG, Eastern Bongo (N = 30); EBK, Eastern Baka (N = 29); EWD, Ewondo (N = 24); GBK, Gabonese Baka (N = 30); GFG, Gabonese Fang (N = 30); KOL, Kola (N = 31); KOT, Kota (N = 30); KOY, Koya (N = 29); NGB, Ngumba (N = 30); NZE, Nzebi (N = 30); NZI, Nzime (N = 31); SBG, Southern Bongo (N = 30); SBK, Southern Baka (N = 29); TEK, Teke (N = 30); TIK, Tikar (N = 30); TSG, Tsogho (N = 30). Pygmy peopling areas were inferred from our ethnological fieldwork. C.A.R. stands for Central African Republic. Equ. Guinea stands for Equatorial Guinea.

across the four populations; 52.0% (SD = 23.3%) among the Bezan individuals; 56.9% (SD = 19.5%) among the Koya individuals; and 63.1% (SD = 19.4%) among the Kola individuals. Interestingly, the Eastern and Southern Bongo pygmies showed an extreme level of admixture, clustering only slightly less in the "Blue" cluster than nonpygmies: 79.9% (SD = 10.7%) membership, on average.

For K = 3, a large fraction of the Bezan individuals clustered in a third, distinct cluster (in yellow). They had, on average, 44.4% (SD = 34.7%) of genotype membership in this third cluster. Further increasing the number of assumed clusters did not change the overall interpretation of our clustering results (see Supplemental Experimental Procedures and Figures S4–S6).

Do All Pygmy Groups in Western Central Africa Have a Common Ancestor?

We used an ABC approach to discriminate among putative evolutionary scenarios [8]. Because of their high level of genetic similarity with nonpygmy populations (see above), we first excluded the two Bongo pygmy populations and their immediate nonpygmy neighbors. Then, we performed additional specific ABC treatments including these two Bongo populations. Note that, in both treatments, power analyses based on test data sets simulated under various scenarios indicated that we achieved sufficient power to discriminate between the competing scenarios (see Supplemental Experimental Procedures for details on type I and type II error rates).

Treatments Excluding Bongo Pygmies

We compared eight competing evolutionary scenarios (Figure 4; Figure S1). Scenarios 1a, 1b, 1c, and 1d assumed a common origin of pygmy populations, whereas scenarios 2a, 2b, 2c, and 2d assumed successive and independent divergence of pygmy groups from the nonpygmy lineage. Scenarios 1a, 1c, 2a, and 2c assumed two potential introgression events from the nonpygmy gene pool into each one of the pygmy lineages, whereas, in scenarios 1b, 1d, 2b, and 2d, the introgression rates were set to 0. Finally, scenarios 1a, 1b, 2a, and 2b further considered a change of effective size in the nonpygmy lineage.

The relative posterior probabilities computed for each scenario provided strong statistical support for scenario 1a (Prob. = 0.96; see Supplemental Experimental Procedures and Table S6). This scenario assumed a common ancestral pygmy population, which derived itself from a population ancestral to both the pygmy and the nonpygmy lineages, the latter having undergone a demographic expansion. This scenario also included two introgression events from nonpygmies into pygmy populations: one in the ancestral pygmy population and one in each one of the four pygmy lineages that diverged from the ancestral pygmy populations. Scenarios 2a, 2b, 2c, and 2d, which assumed independent origins of the pygmy populations, were statistically poorly supported by the data (Table S6).

Given that scenario 1a was clearly favored, we inferred the posterior distributions of parameters for this model only. We found that the ancestral pygmy population diverged 3,587 generations ago (95% CI: 921–4,913) from the ancestral nonpygmy population, i.e., 89,675 years before present (YBP) (95% CI: 23,025–123,275) assuming a generation time of 25 years [12] (Table S2). The split time among pygmy populations was considerably smaller, i.e., $t_p = 105$ generations (95% CI: 29–1,371), which translates into 2,625 YBP (95% CI: 725–34,275).

Regarding introgression, we found that the estimated posterior distributions for the rate and time of the "ancient" introgression event from nonpygmies into the ancestral pygmy populations were relatively flat and hence noninformative ($r_a = 0.927$, 95% CI: 0.041–0.982; $tr_a = 771$, 95% CI: 212–3,749). The credibility intervals for the "recent" introgression rates were narrower: they ranged from $r_{r2} = 0.416$ (95% CI: 0.098–0.899) for the Bezan to $r_{r1} = 0.696$ (95% CI: 0.261 - 0.957) for the Baka, indicating heterogeneous levels of "recent" introgression from nonpygmies.

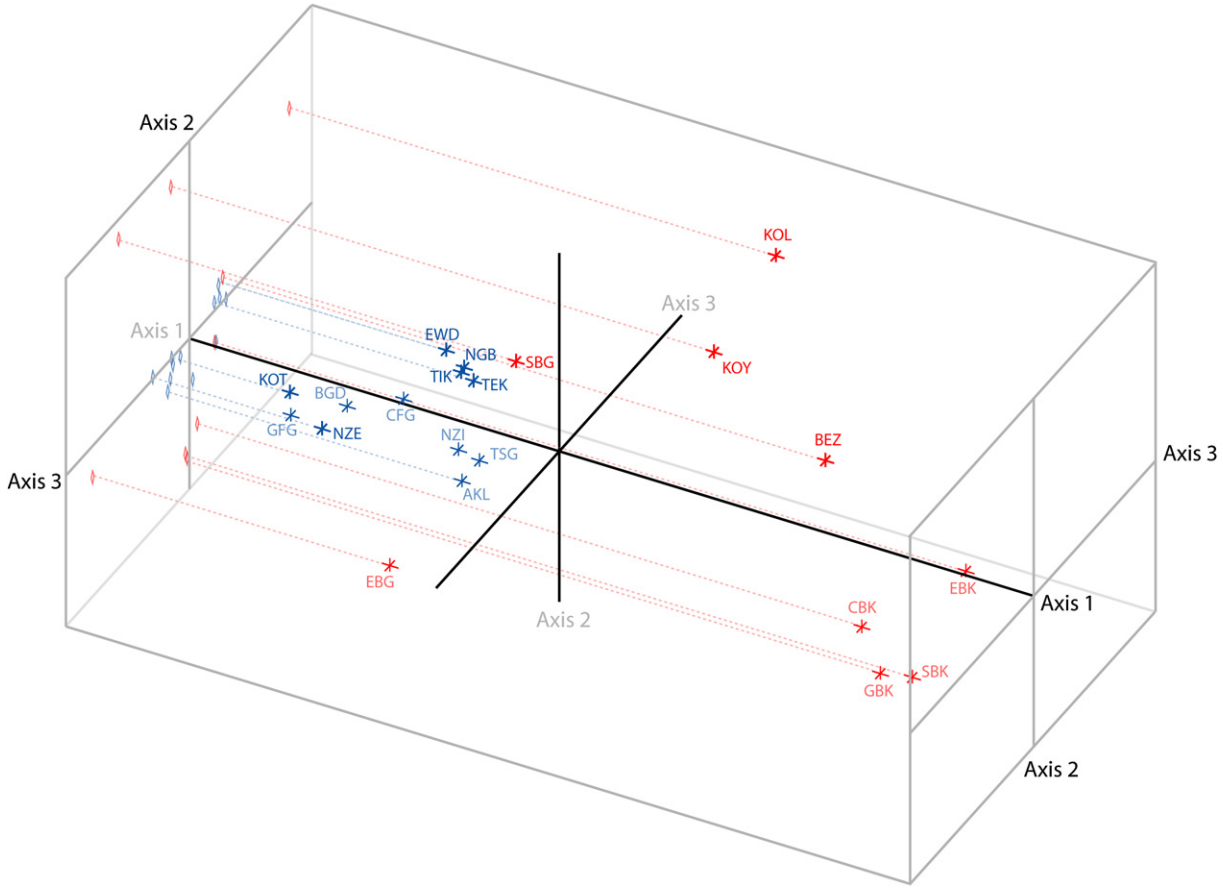

Figure 2. PCA Representation of Genetic Differentiation among Populations

We computed pairwise F_{ST} values among populations [37] by using SpaGeDi v1.2 [32] and performed a Principal Components Analysis (PCA) based on the matrix of pairwise F_{ST} values between populations by using GenAlex v6.0 [39]. The 9 pygmy and 12 nonpygmy populations are colored in red and blue, respectively. The first three principal components represented more than 70% of the total genetic variation among populations (39.7%, 18.3%, and 13.6%, respectively). The first axis reflected the genetic differentiation between nonpygmy and pygmy populations, more specifically the four subgroups of Baka pygmies. The second and third PCA axis reflected mainly the genetic differentiation found among pygmy populations. It is worth noting that the two Bongo pygmy samples cluster together with the nonpygmy populations. See the legend of Figure 1 and Table S1 for the population code.

The estimated effective population sizes for the pygmy populations (from N_2 [Bezan] = 2,785 to N_1 [Baka] = 8,137) were roughly one order of magnitude lower than for nonpygmies (N_{np} = 77,157). The effective size of the ancestral pygmy population (N_{ap} = 8,007) was in the upper range of population sizes for modern pygmy populations. However, our results support the occurrence of a strong expansion in the nonpygmy population (from N_A = 1,071 to N_{np} = 77,157) that occurred at time t_A = 2,802 generations, i.e., 70,500 YBP (95% CI: 16,575–235,475). Posterior estimations of all parameters for scenario 1a were very consistent when an alternative set of priors was used (Priors set 2, see text in Supplemental Experimental Procedures and Table S7).

The Case of the Bongo Pygmies

Capitalizing on the above-described results, we then compared four additional scenarios (B1–B4, see text in Supplemental Experimental Procedures for details), including the two Bongo pygmy populations. Estimated posterior probabilities were again in favor of a historical scenario B1, in which all Western Central African pygmy populations derived recently from a single common ancestral population, with both ancient and more recent introgression events from nonpygmies (Table S8).

Under this most likely scenario, we found that the ancestral pygmy population diverged 53,975 YBP (95% CI: 20,625–121,475) from the ancestral population (Table S9). The split time among all seven Western Central African pygmy populations was again considerably shorter, i.e., t_p = 2,900 YBP (95% CI: 850–30,050). Interestingly enough, the two Bongo samples showed the highest levels of recent introgression from nonpygmies (r_{r5} = 0.679 and r_{r6} = 0.694 for the East and South Bongo populations, respectively; Table S9). Again, posterior estimations of these parameters were very consistent when alternative set of priors were used (see Supplemental Experimental Procedures and Table S10).

Discussion

Strong Autosomal Genetic Structure in Western Central Africa

Like other studies [13–15], we found very low levels of genetic differentiation among nonpygmy populations, which might be due to their recent demographic expansion in Central Africa

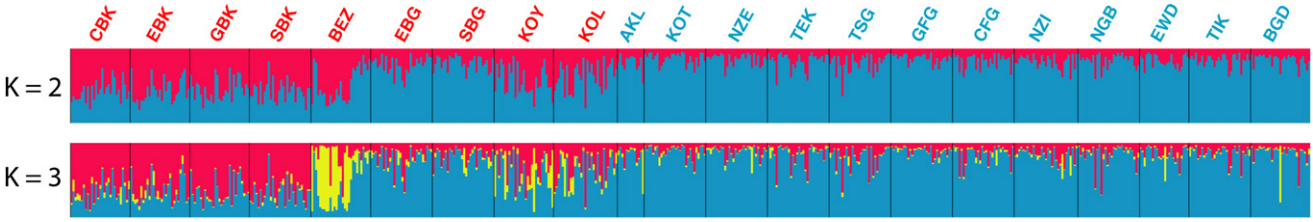

Figure 3. Genetic Structure of Western Central African Populations

The number of clusters assumed in STRUCTURE [11] is given by K. Individuals' genotype membership proportions in each cluster are represented by a single vertical line divided in K colors for each individual. Black lines separate individuals from different predefined populations. See Table S1 for the population code. We averaged, for each K value, results across runs belonging to the same mode as obtained with CLUMPP [38] and built summary barplots by using the program DISTRUCT v.1.1 [40]. We averaged results across runs showing the highest overall probability of observing the data and belonging to the same mode, for K = 2 and K = 3 (see Supplemental Experimental Procedures for detailed results). For K = 2, we found that all 50 independent runs gave very similar individual clustering results, and we therefore present here the averaged clustering solution among the 50 independent runs. For K = 3, we identified two modes (see Figure S4). One mode (displayed here) contained only 3 runs out of 50, but these runs systematically showed the highest overall probability of observing the data and were highly informative, with the third (yellow) cluster including a great proportion of Bezan pygmy individuals. Therefore, we present here the individual clustering results averaged across these three best runs (see Supplemental Experimental Procedures for more detailed results and discussion).

(2000–5000 YBP) [16]. The overall genetic variation in Western Central Africa is mainly structured by a substantial differentiation between pygmy populations (excluding the Bongo) and their nonpygmy neighbors, and also among all pygmy populations. Thus, the generic term "Pygmy," under which culturally and linguistically very diverse groups are gathered [3, 17, 18], hides a large genetic heterogeneity: the F_{ST} value (0.019) among pygmy populations sampled at this small geographical scale is in the upper range of F_{ST} values found at continental scales (e.g., 0.015 in Eurasia) [13]. In this context, the third cluster identified specifically in the Bezan group by using STRUCTURE for K = 3 could result from strong genetic drift due to the very small census size (at most 400 individuals nowadays), echoing the smallest effective population size found through our ABC treatments for this population. Indeed drift in a small isolated population might quickly lead to a marked divergence in allelic frequencies, thus resulting in a specific clustering with STRUCTURE.

Origin and Diversification of Hunter-Gatherer Pygmies in Western Central Africa

Despite the substantial genetic differentiation found among pygmy populations, our ABC analyses strongly support a common origin of Western Central African pygmies (Baka, Bezan, Koya, Kola, and Bongo). This scenario is consistent with the common maternal ancestry found in a similar pygmy population set when using mitochondrial DNA [14]. The initial divergence between ancestral pygmies and nonpygmies appears to be ancient: 53,975 YBP (95% CI: 20,625–121,475) when including the Bongo, and 89,675 YBP (95% CI: 23,025–123,275) without them, consistent with previous estimates [14, 19, 20]. These estimations of divergence times give insights into the time frame needed for human populations to acquire deep morphological differences. The wide credibility intervals found for this ancient divergence makes it nevertheless difficult to speculate on the potential causes for the original split and the subsequent morphological diversification of these human groups. This leaves room for future studies based on autosomal sequence data that could provide a more precise estimation of this original ancient divergence among Western Central African human populations.

Interestingly, we found strong evidence for a recent divergence among pygmy populations: 2,625 YBP (95% CI: 725–34,275) and 2,900 YBP (95% CI: 850–30,050) without or with the Bongo pygmies, respectively. This dates back to the transition period from stone to metal techniques and the expansion of nonpygmy populations in this area (2,000–5,000 YBP) [6]. Our results hence support the hypothesis that the expansion of neighboring nonpygmy agriculturalists fundamentally affected the existing relationships within the ancestral pygmy population [2, 21]. This expansion probably introduced new constraints in the pygmy mobility and intermarriages, increasing isolation and thus genetic differentiation among pygmy populations. These findings are consistent with the small and constant effective population sizes found in these hunter-gatherer populations (see Table S2), generalizing previous results based on only two pygmy populations [16, 20]. The enhanced genetic drift may explain the significantly lower number of alleles in pygmies versus nonpygmies. Expected heterozygosities are not significantly different between the two groups, but such discrepancy could result from the demographic expansion of nonpygmies; indeed, the number of alleles is expected to increase faster than heterozygosity during demographic expansions [22, 23].

Heterogeneous Asymmetrical Genetic Introgression from Nonpygmies into Pygmies in Western Central Africa

Our second main result is the recent asymmetrical introgression from nonpygmy into pygmy populations found with both ABC and STRUCTURE treatments. Such asymmetrical admixture may stem from the intermarriage practices between pygmies and nonpygmies. Ethnologists often report a strong discrimination against pygmies, preventing marriages between a nonpygmy woman and a pygmy man. Conversely, marriages between pygmy women and nonpygmy men are less socially constrained [24–26]. These populations being strictly patrilocal, these marriages should result in a pygmy to nonpygmy gene flow, since married pygmy women live with their families-in-law. However, because of social pressures and discriminations, many of these intermarriages are broken and result in the return of the pygmy woman with her children to her original community, thereby increasing male gene flow from nonpygmy to pygmy populations. Moreover, this admixture pattern could be reinforced by the presence in the pygmy communities of illegitimate children from nonpygmy men and pygmy women [27].

These results are consistent with previous investigations based on Y chromosome and mitochondrial DNA data [15,

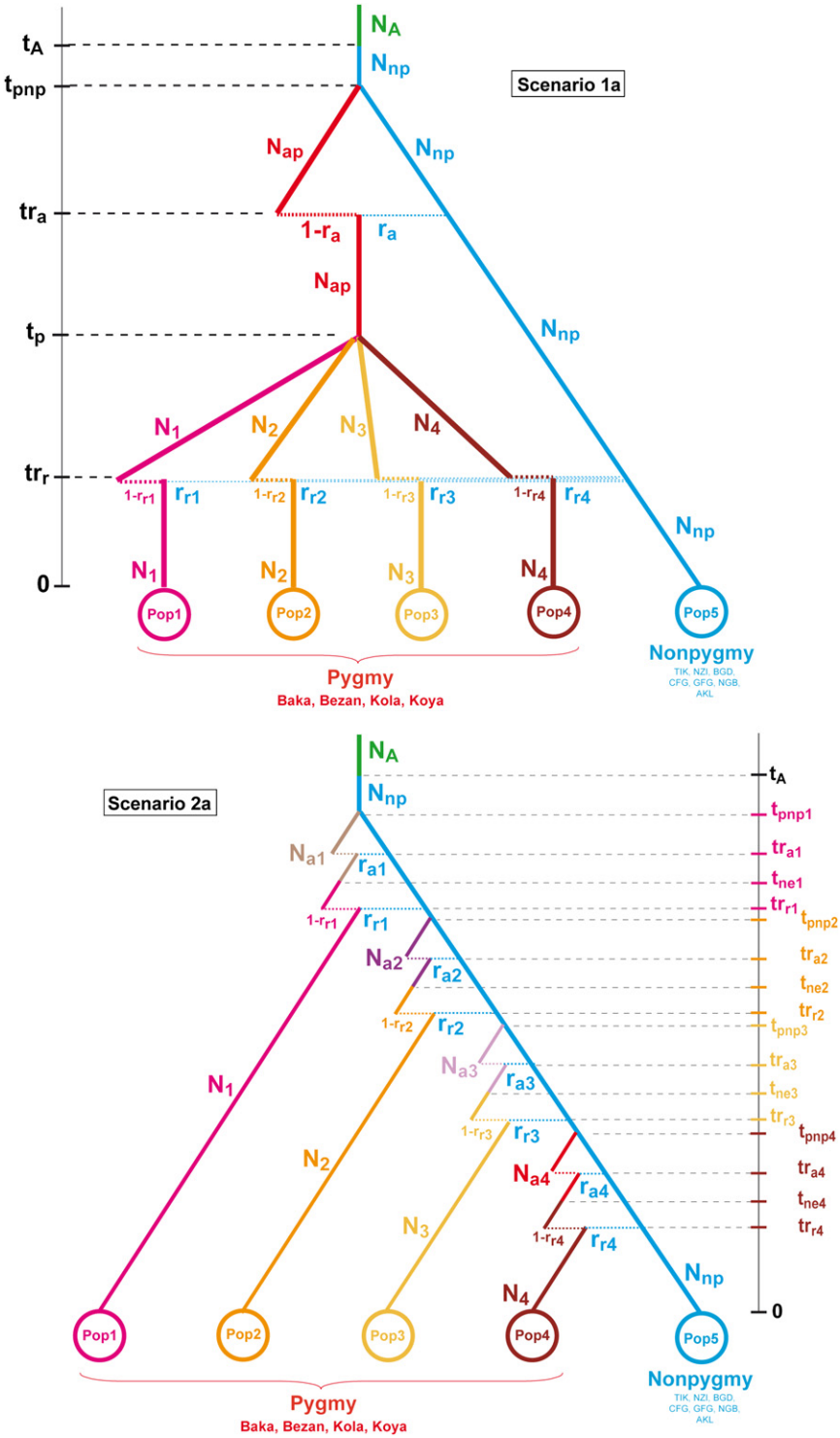

Figure 4. Two of the Eight Scenarios Compared through Approximate Bayesian Computations for the Origin and Diversification of Western Central African Populations

Scenario 1a corresponds to a common origin of pygmy populations that diversified from a single ancestral pygmy population at time t_p. Ancestral pygmies diverged from the nonpygmy lineage at time t_{pnp}. Scenario 2a corresponds to an independent origin of pygmy groups that diverged from the nonpygmy lineage at time t_i. In order to account for asymmetrical gene flows from nonpygmies into pygmy lineages [21, 28], we simulated two events of introgression from the nonpygmy lineage into each pygmy lineage independently. Because human populations underwent fluctuating demographic regimes throughout history [4, 20], we added a potential stepwise change in effective population size that occurred in the ancestral population or in the nonpygmy lineage at time t_A. Finally, we compared scenarios in which each pygmy lineage encompassed the same number of historical events with parameters drawn from the same prior distributions, across all competing scenarios. Therefore, scenarios 2a–2d included an event of potential change in effective population size separating the two introgression events. In scenarios 1b/2b, all introgression rates were set to zero, whereas cenarios 1c and 2c assumed constant effective size in the nonpygmy lineage. Finally, scenarios 1d and 2d assumed neither introgression events nor change in effective size (details in Supplemental Experimental Procedures and in Figure S1). We considered five Western African samples: the Baka, Bezan, Kola, and Koya pygmy samples, and the Nonpygmy sample (see Table S1 for population groupings). For all scenarios, N_i indicates the effective population size of population i. Note that for scenarios 2a–2d, split times were drawn independently in corresponding prior distributions for each pygmy lineage, and thus the order in which these lineages split is not predefined. Finally, the time scale represented here is arbitrary, as these times will be estimated through our ABC treatments.

21, 28] that focused on a limited number of African pygmy samples. By sampling many different pygmy groups, we show the heterogeneity of the admixture levels among pygmy populations. The sociocultural patterns, such as the differences in marriage practices with nonpygmies, seem to explain the observed differences in the levels of genetic admixture. For instance, we found that the Bongo samples are the most admixed with both STRUCTURE and DIYABC treatments. This result is consistent with the substantial amount of intermarriages with nonpygmies among the Bongo, who are highly socially integrated as compared to all other pygmy populations (P.V., S.L.B., and S.B., unpublished data). Interestingly enough, the Bongo pygmies are also the tallest among all African pygmy groups ([7]; P.V., unpublished data). Intermediate introgression rates found with both methods for the Bezan, Koya, and Kola pygmies are also consistent with the more intermediate levels of intermarriages found nowadays between these groups and their nonpygmy neighbors (A.F., S.B., and S.L.B., unpublished data).

The Baka pygmies are known to be strongly constrained by intermarriages with nonpygmies ([29]; S.B., unpublished data). As expected, they appear as the least admixed individuals in the STRUCTURE analysis. However, our ABC treatments point toward a slightly higher level of introgression for them (but note

the large 95% CI). Note, however, that although our ABC treatments were found to achieve sufficient power to discriminate among scenarios and provide robust split time estimations, our set of independent loci lacked information on introgression parameters (results not shown based on simulated test data sets). Furthermore, we also found substantial levels of interindividual variability in the admixture levels in our STRUCTURE analyses for K = 2, with overlapping CIs across pygmy individuals, excluding Bongo pygmies. Finally, it is worth stressing here that the assumption of a limited number of punctual introgression events is by no means equivalent to assuming continuous gene flow among populations. This leaves room for future inferential treatments that would allow for treating such complex historical scenarios in which many populations exchange migrants in a continuous way.

Conclusions

To our knowledge, this study provides the first precise investigation of neutral autosomal genetic variation in a high-density sample of the largest modern hunter-gatherer groups, and their immediate agriculturalist neighbors. We provide insights into the time frame and demographic mechanisms needed to establish the deep morphological differences observed between very different human groups: sedentary agriculturalists and mobile hunter-gatherer pygmies.

Despite the substantial level of genetic differentiation found among pygmy populations, we identified a recent (about 2,800 YBP) common origin of all Western Central African pygmy populations, together with a more ancient (~54,000 or 90,000 YBP) divergence between the ancestral pygmy and nonpygmy populations. Finally, we found strong evidences of a recent asymmetrical and heterogeneous genetic introgression from nonpygmy into pygmy populations.

Our results hence converge toward a historical scenario in which the expansion of nonpygmy agriculturalist populations during the Neolithic revolution (2000–5000 YBP) in this area introduced new social constraints upon the ancestral pygmy population. This led to the rapid genetic diversification of the various Western Central African pygmy populations, through isolation, subsequent enhanced genetic drift, and heterogeneous asymmetrical introgression from nonpygmies into pygmies. We propose a summarized historical model of Western Central African peopling in Figure S3.

Experimental Procedures

Population Sampling and Marker Set

We genotyped 604 unrelated individuals from 9 pygmy and 12 nonpygmy populations from Cameroon and Gabon (see Figure 1 and Table S1). Oral and video-recorded informed consent was obtained for each donor. We used 28 autosomal microsatellites (Table S4) located on 18 chromosomes, from the data set provided by the Marshfield Foundation Mammalian Genotyping Service Screening Set 10 available at http://research.marshfieldclinic.org/genetics/GeneticResearch/screeningsets.asp. Levels of familial relationships were indirectly estimated from genetic data among individuals within each population by using RelPair v2.0 [30].

Data Analysis

Genetic Variation within and between Populations

We computed the expected heterozygosities (*He*, [31]) and mean allelic number across loci in each population by using SpaGeDi v.1.2 [32]. Differences between the 9 pygmy and the 12 nonpygmy populations were tested with Wilcoxon rank sum tests [33], as implemented in R [34]. We performed a hierarchical AMOVA [35] with Arlequin v.3.00 [36], grouping all pygmy populations on one side and all nonpygmy populations on the other side, and computed global indexes of multilocus differentiation (F_{ST}, [37]) among both groups.

Clustering Methods Based on Individual Genotypes

To characterize the admixture patterns among populations, we used the Bayesian method of individual clustering implemented in STRUCTURE v.2.1 [11]. We used the admixture model that assigns individual proportions of genotypes to each one of K clusters, K values ranging here from 1 to 5. We performed 50 independent runs for each K value, with separate values for the Dirichlet parameters α for each assumed clusters. Each run included 600,000 iterations following a burnin period of 200,000 iterations. We used the *Greedy* algorithm implemented in CLUMPP [38] in order to identify potential distinct modes among the results of the 50 STRUCTURE runs for each K value.

Approximate Bayesian Computations

In order to reconstruct the unknown history of divergence and migrations among Central African populations, we performed approximate Bayesian computations (ABC) with the computer package DIYABC, which allows for the handling of large microsatellite data sets for various complex demographic scenarios involving any combination of population divergences, admixtures, and stepwise population size changes [8].

See Figure 4, Supplemental Experimental Procedures, and Figure S1 for a precise description of the scenarios compared by using DIYABC [8]. The first treatments excluding Bongo pygmies comprise four groups of pygmy samples (i.e., Baka, Bezan, Kola, and Koya; see Table S1 for details on sample composition) and a single nonpygmy sample, merging all seven nonpygmy populations neighboring these pygmies. We pooled these seven nonpygmy populations into a single sample, as well as the four Baka subgroups into another sample, due to very low levels of genetic differentiation observed between these populations (see Results). Treatments regarding the two Bongo pygmy samples were conducted separately due to their high level of genetic similarity with the nonpygmies (see the Results section). For these treatments, we compared four specific scenarios, which are detailed in Supplemental Experimental Procedures and Figure S2.

The ABC treatment relies on simulated data sets produced by drawing the model parameters from a set of prior distributions. We evaluated the sensitivity of our ABC treatment to the set of priors by performing the same treatments with two different sets of priors (Table S11). Finally, we evaluated the power of our ABC methodology to discriminate between scenarios by analyzing simulated data sets with the same number of loci and individuals as in our real data set (see text in Supplemental Experimental Procedures for details).

Supplemental Data

Supplemental Data include Experimental Procedures, six figures, and eleven tables and can be found with this article online at http://www.current-biology.com/supplemental/S0960-9822(09)00542-9.

Acknowledgments

We gratefully thank the Biological Resource Center of the Foundation Jean Dausset-Centre d'Etude du Polymorphisme Humain for genotyping facilities, especially Howard Cann and Hélène Blanchet. We thank Kenneth K. Kidd, Barry Hewlett, Giovani Destro-Bisol, Raphaëlle Chaix, Bruno Toupance, Begoña Martinez-Cruz, as well as four anonymous reviewers for useful comments and suggestions. This work was supported by the Centre National de la Recherche Scientifique, the Museum National d'Histoire Naturelle, and a Action Concertée Incitative Prosodie research grant; the European Science Foundation (the The Origin of Man, Language and Languages Eurocores programme); and the French Agence Nationale de la Recherche (ANR) grants NT05-4-42230 and ANR-BLAN-0196-01.

Received: July 11, 2008
Revised: December 23, 2008
Accepted: December 24, 2008
Published online: February 5, 2009

References

1. Diamond, J., and Bellwood, P. (2003). Farmers and their languages: The first expansions. Science *300*, 597–603.
2. Cavalli-Sforza, L.L. (1986). African Pygmies: An Evaluation of the State of Research (Orlando: Academic Press).
3. Hewlett, B. (1996). Cultural diversity among African Pygmies. In Cultural Diversity among Twentieth-Century Foragers: An African Perspective, S. Kent, ed. (Cambridge: Cambridge University Press), pp. 215–244.

4. Cavalli-Sforza, L.L., Menozzi, P., and Piazza, A. (1994). The History and Geography of Human Genes (Princeton, NJ: Princeton University Press).
5. Mercader, J., ed. (2003). Under the Canopy. The Archaeology of Tropical Rain Forests (New Brunswick: Rutgers University Press).
6. Phillipson, D.W. (2005). African Archaeology (Cambridge: Cambridge University Press).
7. Froment, A. (1993). Adaptation biologique et variation dans l'espèce humaine: Le cas des Pygmées d'Afrique. Bull. Mem. Soc. Anthropol. Paris *5*, 417–448.
8. Cornuet, J.M., Santos, F., Beaumont, M.A., Robert, C.P., Marin, J.M., Balding, D.J., Guillemaud, T., and Estoup, A. (2008). Inferring population history with DIY ABC: A user-friendly approach to Approximate Bayesian Computation. Bioinformatics *24*, 2713–2719.
9. Vansina, J. (1995). New linguistic evidence and 'the Bantu expansion'. J. Afr. Hist. *36*, 173–195.
10. Rexová, K., Bastin, Y., and Frynta, D. (2006). Cladistic analysis of Bantu languages: A new tree based on combined lexical and grammatical data. Naturwissenschaften *93*, 189–194.
11. Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics *155*, 945–959.
12. Reich, D.E., and Goldstein, D.B. (1998). Genetic evidence for a Paleolithic human population expansion in Africa. Proc. Natl. Acad. Sci. USA *95*, 8119–8123.
13. Rosenberg, N.A., Pritchard, J.K., Weber, J.L., Cann, H.M., Kidd, K.K., Zhivotovsky, L.A., and Feldman, M.W. (2002). Genetic structure of human populations. Science *298*, 2381–2385.
14. Quintana-Murci, L., Quach, H., Harmant, C., Luca, F., Massonnet, B., Patin, E., Sica, L., Mouquiama-Daouda, P., Comas, D., Tzur, S., et al. (2008). Maternal traces of deep common ancestry and asymmetric gene flow between Pygmy hunter-gatherers and Bantu-speaking farmers. Proc. Natl. Acad. Sci. USA *105*, 1596–1601.
15. Hammer, M.F., Karafet, T.M., Redd, A.J., Jarjanazi, H., Santachiara-Benerecetti, S., Soodyall, H., and Zegura, S.L. (2001). Hierarchical patterns of global human Y-chromosome diversity. Mol. Biol. Evol. *18*, 1189–1203.
16. Excoffier, L., and Schneider, S. (1999). Why hunter-gatherer populations do not show signs of pleistocene demographic expansions. Proc. Natl. Acad. Sci. USA *96*, 10597–10602.
17. Bahuchet, S. (1993). L'Invention des Pygmées. Cah. Etud. Afr. *129*, 153–181.
18. Bahuchet, S. (1993). History of the inhabitants of the central African rain forest: Perspectives from comparative linguistics. In Tropical Forests, People and Food. Biocultural Interactions and Applications to Development, H. Cea, A. Hladik, O.F. Linares, H. Pagezy, A. Semple, et al., eds. (Paris/Lancs: UNESCO/Parthenon), pp. 37–54.
19. Batini, C., Coia, V., Battaggia, C., Rocha, J., Pilkington, M.M., et al. (2007). Phylogeography of the human mitochondrial L1c haplogroup: Genetic signatures of the prehistory of Central Africa. Mol. Phylogenet. Evol. *43*, 635–644.
20. Zhivotovsky, L.A., Rosenberg, N.A., and Feldman, M.W. (2003). Features of evolution and expansion of modern humans, inferred from genomewide microsatellite markers. Am. J. Hum. Genet. *72*, 1171–1186.
21. Destro-Bisol, G., Donati, F., Coia, V., Boschi, I., Verginelli, F., Cagliá, A., Tofanelli, S., Spedini, G., and Capelli, C. (2004). Variation of female and male lineages in sub-Saharan populations: The importance of sociocultural factors. Mol. Biol. Evol. *21*, 1673–1682.
22. Cornuet, J.M., and Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics *144*, 2001–2014.
23. Nei, M., Maruyama, T., and Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. Evol. Int. J. Org. Evol. *29*, 1–10.
24. Bahuchet, S., and Guillaume, H. (1982). Aka-farmer relations in the Northwest Congo Basin. In Politics and History of Band Societies, E. Leacock and R.B. Lee, eds. (Cambridge, Paris: Cambridge University Press/MSH), pp. 189–211.
25. Seitz, S. (1993). Pygmées d'Afrique Centrale (Paris: Peeters/SELAF).
26. Turnbull, C. (1965). Wayward Servants, the Two Worlds of the African Pygmies (New York: The Natural History Press).
27. Seitz, S. (1977). Die Zentralafrikanischen Wildbeuterkulturen, Eighth Edition (Wiesbaden, Germany: Steiner).
28. Cruciani, F., Santolamazza, P., Shen, P., Macaulay, V., Moral, P., Olckers, A., Modiano, D., Holmes, S., Destro-Bisol, G., Coia, V., et al. (2002). A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. Am. J. Hum. Genet. *70*, 1197–1214.
29. Joiris, D.V. (2003). The framework of Central African hunter-gatherers and neighbouring societies. Afr. Study Monogr. (*Suppl 28*), 57–79.
30. Epstein, M.P., Duren, W.L., and Boehnke, M. (2000). Improved inference of relationship for pairs of individuals. Am. J. Hum. Genet. *67*, 1219–1231.
31. Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics *89*, 583–590.
32. Hardy, O.J., and Vekemans, X. (2002). SPAGeDi: A versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol. Ecol. Notes *2*, 618–620.
33. Sokal, R.R., and Rohlf, F.J. (2001). Biometry: The Principles and Practice of Statistics in Biological Research, Eight Edition (San Francisco: W.H. Freeman).
34. R Development Core Team. (2007). R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing).
35. Excoffier, L., Smouse, P.E., and Quattro, J.M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics *131*, 479–491.
36. Excoffier, L., Laval, L., and Schneider, S.E. (2005). Arlequin (version 3.0): An integrated software package for population genetics data analysis. Bioinformatics *1*, 47–50.
37. Weir, B.S., and Cockerham, C.C. (1984). Estimating F-statistics for the analysis of population-structure. Evol. Int. J. Org. Evol. *38*, 1358–1370.
38. Jakobsson, M., and Rosenberg, N.A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics *23*, 1801–1806.
39. Peakall, R., and Smouse, P.E. (2006). GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes *6*, 288–295.
40. Rosenberg, N.A. (2004). DISTRUCT: A program for the graphical display of population structure. Mol. Ecol. Notes *4*, 137–138.