Genomic Differentiation Among Natural Populations Of Orang-Utan *(Pongo pygmaeus)*

Lu Zhi[1], William B. Karesh[2], Dianne N. Janczewski[3], Harmony Frazier-Taylor[4], Dondin Sajuthi[5], Francis Gombek[6], Mahedi Andau[7], Janice S. Martenson[8] and Stephen J. O'Brien [8]

[1] Laboratory of Genomic Diversity, National Cancer Institute, Frederick, Maryland 21702-1201, USA, Department of Environmental Biology and Ecology, College of Life Sciences, Peking University, Beijing 100871, China; [2] New York Zoological Society, Bronx, New York 10460, USA; [3] Laboratory of Genomic Diversity, National Cancer Institute, Frederick, Maryland 21702-1201, USA, Consolidated Safety Services, 4031 University Drive, Fairfax, Virginia 22030, USA; [4] Woodland Park Zoological Gardens, Seattle, Washington 98103, USA;[5] Institut Pertanian Bogor, Bogor, Indonesia; [6] National Parks and Wildlife Division, Kuching, Sarawak, Malaysia; [7] Department of Wildlife, Kota Kinabalu, Sabah, Malaysia; [8] Laboratory of Genomic Diversity, National Cancer Institute, Frederick, Maryland 21702-1201, USA

Received 10 April 1996; Revised 6 June 1996; accepted 8 July 1996; Available online 7 August 2004

Abstract

Background

Orang-utans exist today in small isolated populations on the islands of Borneo (subspecies *Pongo pygmaeus pygmaeus*) and Sumatra (subspecies *P. p. abelii*). Although, on the basis of their morphological, behavioral and cytogenetical characteristics, the Bornean and Sumatran orangutan populations are generally considered as two separate subspecies, there is no universal agreement as to whether their genetic differentiation is sufficient to consider and manage them as species, subspecies or population level taxonomic units. A more precise phylogenetic description would affect many conservation management decisions about captive and free-ranging orang-utans.

Results

We analyzed the amount and patterns of molecular genetic variation in orang-utan populations using cellular DNA from orang-utans from two locations in Sumatra and nine locations – representing four isolated populations – in Borneo. Genetic and phylogenetic analyses of mitochondrial DNA restriction fragment length polymorphisms, nuclear minisatellite (or variable number tandem repeat) loci and mitochondrial 16S ribosomal RNA sequences led to three major findings. First, the genetic distance and phylogenetic differentiation between Sumatran and Bornean orang-utans is large, greater than that between the common chimpanzee, *Pan troglodytes*, and the pygmy chimpanzee or bonobo, *Pan paniscus*. The genetic distance suggests that the two island subspecies diverged ~1.5–1.7 million years ago, well before the two islands separated and long enough for species-level differentiation. Second, there is considerable endemic genetic diversity within the Bornean and Sumatran orang-utan populations, suggesting that they have not experienced recent bottlenecks or founder effects. And third, there is little

genetic differentiation among four geographically isolated populations of Bornean orang-utans, consistent with gene flow having occurred between them until recently.

Conclusions

Our results are consistent with the view that the genetic differentiation between Sumatran and Bornean orang-utans has reached the level of distinct species. Furthermore, our findings indicate that there is not a genetic imperative for the separate management of geographically isolated Bornean populations.

Article Outline

Background:

The orang-utan, *Pongo pygmaeus*, is the only great ape species that exists outside of Africa. Its present range includes the dense rainforest habitats of Borneo and north Sumatra, where the species has become severely threatened because of poaching and habitat destruction. The two island populations are traditionally designated as separate subspecies — *P. pygmaeus pygmaeus* (Bornean) and *P. p. abelii* (Sumatran) — on the basis primarily of their distinctive morphological and behavioral characteristics [1, 2, 3, 4, 5 and 6]. A pericentric inversion of chromosome 2 [7] is a cytogenetic difference that has been used as an indicator of subspecies for the management of captive and confiscated orang-utans [8].

A series of molecular genetic analyses — of blood proteins and isozymes [9, 10, 11 and 12], fibroblast proteins resolved by two dimensional gel electrophoresis [12], DNA–DNA hybridization [13], mitochondrial DNA restriction fragment length polymorphisms (mtDNA–RFLPs) [8, 14 and 15], and the sequence of the mitochondrial *COII* gene [16] — has revealed that the genetic divergence between the subspecies is large, and in some cases approaches species-level genetic distance. In addition, Groves *et al*. [5] concluded from cranial measurements that orang-utan skulls from southwestern Borneo are as distinct from other

Bornean specimens as are the skulls of Bornean and Sumatran orang-utans; this level of difference is consistent with near species-level divergence. But recent reports of mitochondrial DNA sequences have contradicted these inferences of large distinctiveness and again raised the controversy about the most accurate designation of orangutan populations [17 and 18].

The taxonomic distinctions of orang-utan subspecies and populations are relevant not only to systematic issues but also to species conservation, because species and subspecies are the units of protection and captive propagation [19, 20 and 21]. Furthermore, phylogeographic descriptions of individual populations can reveal the presence or absence of a recognizable genetic substructure, which is useful for identifying the population/subspecies origin of confiscated illegal pets. In addition, the existence (or not) of a population-specific genetic subdivision is important in deciding whether reintroduction or relocation programs need to consider the genetic distinctiveness of isolated populations. In this study, we address these issues explicitly by assessing the extent and character of endemic genetic variation among five geographically isolated populations of Bornean and Sumatran orang-utans, using three distinct genetic measures of variation: mtDNA–RFLPs, nuclear DNA minisatellite (variable number tandem repeat) loci and mitochondrial 16S ribosomal RNA (rRNA) sequences.

Results

Analysis of mtDNA–RFLPs

We examined cellular DNA from six unrelated Sumatran and 33 unrelated Bornean orang-utans using 30 restriction enzymes. Samples were obtained from two Sumatran locales (n = 6) considered to contain a single population, and nine Bornean locales (n = 33) representing four geographically isolated populations (Figure 1). We scored a total of 149 restriction sites representing 720 nucleotides (4.4 % of the 16 500 base pair (bp) mitochondrial genome [22]).


Full-size image (107K)

Figure 1. Sampling locations (black dots) of orangutans on Borneo and Sumatra. Two areas were sampled in Sumatra (n = 6); nine areas were sampled in Borneo (n = 33) from four geographically isolated populations: Sabah (n = 16), in northeastern Borneo (Malaysia); Sarawak (n = 12), in northwestern Borneo (Malaysia), including the northwestern Kalimantan (Indonesia); Kutai (n = 3), in east Borneo (Indonesia); and Gunung Palung (n = 2), in southwest Borneo (Indonesia). The pie charts show the distribution of thirteen mtDNA haplotypes (Table 1) in each population.

A total of 18 restriction enzymes produced polymorphic patterns of digestion at 18 restriction sites, and produced 13 distinct haplotypes, designated A–N (Table 1). Four haplotypes were unique to Sumatra and nine were specific to Borneo; there was no overlap between the haplotype distributions of the two subspecies (Figure 1 and Table 1). Within Borneo, orang-utans from Sabah and Sarawak shared two haplotypes, A and D, whereas single unique haplotypes, I and G, were found in the two south Bornean populations. However, these latter two populations were represented by rather small samples (two and three individuals, respectively; Figure 1).

Table 1. . mtDNA haplotypes and nucleomorphs for polymorphic restriction enzymes*.

Haplotype	*Acc*I	*Ava*I	*Ava*II	*Bam*HI	*Bcl*I	*Bgl*I	*Bst*UI	*Dra*I	*Eco*RV	*Hinc*II	*Hind*III	*Hpa*I	*Kpn*
A	A	A	A	A	A	A	A	A	A	A	A	A	A
B	A	A	B	A	A	A	A	A	A	A	A	A	A
C	A	A	A	A	A	A	A	A	A	A	A	A	A
D	A	B	C	A	A	A	A	B	A	A	B	A	A
E	A	A	A	A	A	A	A	C	A	A	A	A	A
F	A	A	A	A	A	A	A	C	A	A	A	A	A
G	A	A	A	A	A	A	A	C	B	A	A	A	A
H	A	B	C	A	A	A	A	B	A	A	B	A	A
I	B	A	A	A	A	A	A	C	A	A	A	A	A
K	C	C	D	B	B	B	B	D	A	B	C	B	B
L	C	D	E	B	B	B	B	D	A	B	C	B	B
M	D	D	F	B	A	C	C	D	C	C	C	C	A
N	E	D	G	B	B	C	C	E	D	C	C	C	A

Full-size table (21K)

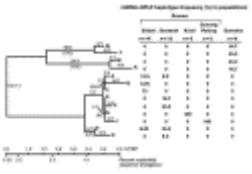
* Nucleomorphs represent the following patterns of restriction enzyme fragment sizes (kb): *Acc*I-A: 6.80, 4.37, 2.41, 1.05, 0.87, 0.35; -B: 4.98, 4.37, 3.55, 1.67, 1.05, 0.87, 0.35; -C: 6.80, 4.37, 2.67, 1.67, 1.05, 0.87, 0.48, 0.35; -D: 10.86, 2.41, 1.05, 0.87; -E: 5.03, 4.37, 2.01, 1.67, 1.05, 0.87, 0.35. *Ava*I-A: 10.71, 4.61, 3.02; -B: 10.71, 3.24, 3.02; -C: 8.68, 6.55, 3.02; -D: 11.91, 3.02. *Ava*II-A: 5.87, 3.36, 0.74, 0.67, 0.57, 0.24; -B: 4.31, 3.36, 0.74, 0.67, 0.52; -C: 5.87, 3.36, 0.67, 0.57, 0.43, 0.24, 0.16; -D: 5.35, 4.17, 1.25, 1.14, 0.61; E: 7.29, 1.86, 0.81, 0.61, 0.38, 0.24; -F: 5.35, 4.17, 1.25, 0.81, 0.65, 0.24; -F: 7.29, 3.57, 0.81, 0.61, 0.38, 0.24. *Bam*HI-A: 10.11, 8.14; -B: 12.19. *Bcl*I-A: 11.48, 3.01; -B: 9.75, 4.54, 3.01. *Bgl*I-A: 12.82, 2.49; -B: 8.25, 6.61, 2.49, 0.56; -C: 8.25, 7.04, 2.49, 0.56. *Bst*UI-A: 10.98, 3.27, 0.65, 0.30, 0.27; -B: 9.48, 1.76, 0.65, 0.63, 0.50, 0.30, 0.27; -C: 5.47, 4.55, 3.27, 1.63, 0.65, 0.30, 0.27. *Dra*I-A: 8.14, 1.93, 1.51, 0.75, 0.43, 0.37; -B: 8.14, 2.27, 1.51, 0.43, 0.37; -C: 8.14, 1.93, 1.51, 0.51, 0.43, 0.37; -D: 8.14, 2.27, 1.93, 1.51, 0.86; -E: 8.14, 1.93, 1.51, 1.36, 0.86. *Eco*RV-A: 13.43; -B: 11.25, 7.90; -C: 12.29, 3.69; -D: 10.67, 5.07, 2.50; -E: 10.67, 8.33. *Hinc*II-A: 5.62, 1.91, 0.63, 0.46, 0.33; -B: 3.23, 2.34, 2.02, 1.65, 1.00, 0.33; -C: 5.46, 2.34, 2.02, 1.65, 0.82, 0.33. *Hind*III-A: 4.13, 3.96, 3.75, 3.10, 1.61, 0.51. -B: 8.73, 3.96, 3.10, 1.61, 0.51; -C: 8.73, 6.83, 3.10, 0.51. *Hpal*-A: 11.73, 3.85, 0.51; -B: 9.79, 5.20, 3.31; -C: 11.73, 5.20. *Kpn*I-A: 8.66, 7.62, 2.90; -B: 10.85, 7.62. *Nco*I-A: 8.88, 5.54, 4.29; -B: 5.15, 4.29, 2.98, 0.62; -C: 8.88, 4.29, 0.62,; -D: 8.88, 0.62; -E: 9.91, 1.44, 0.90, 0.87; -F: 6.40, 5.95, 5.54, 0.90. *Nde*I-A: 11.01, 6.45; -B: 12.50; -C: 11.01, 4.88. *Pvu*II-A: 9.95, 9.35; -B: 13.59. *Stu*I-A: 3.64, 2.15, 1.95, 1.84, 0.90, 0.81, 0.37; -B: 3.64, 2.15, 1.84, 1.31, 0.90, 0.81, 0.40, 0.37; -C: 3.64, 2.15, 1.95, 0.90, 0.81, 0.64, 0.54; -D: 3.64, 2.15, 1.31, 0.90, 0.81, 0.64, 0.54, 0.40, 0.37. *Xba*I-A: 10.34, 8.76; -B: 9.06, 8.76, 1.47; -C: 13.03, 1.47. The following enzymes

produced monomorphic patterns in all orang-utans: *Cla*I, *Eco*RI, *Pst*I, *Sal*I, *Sst*I, *Sst*II, *Xho*I; the following enzymes did not cleave mtDNA: *Bgl*II, *Bst*EII, *Mlu*I, *Pvu*I, *Sma*I.

The mitochondrial nucleotide diversity in orang-utans is large (π = 1.46 %) with appreciable diversity estimated in both Bornean (π = 0.33 %) and Sumatran (π = 1.75 %) population samples (Table 2). The Sumatran value is particularly high, and surpasses comparable estimates of mtDNA–RFLP diversity in humans (π = 0.32 %), gorillas (π = 0.55 %), common chimpanzees (π = 1.3 %), pygmy chimpanzees (π = 1.0 %), pumas (π = 0.35 %), leopards (π = 1.30 %), pocket gophers (π = 0.5 %) and humpback whales (π = 0.25 %) [14, 23, 24, 25, 26 and 27]. On Borneo, the diversity estimates were moderate within the populations from Sabah and Sarawak, but zero in Kutai and Gunung Palung (Table 2). Diversity estimates within the latter two populations may have been influenced by the small sample sizes.

The genetic divergence between the Bornean and Sumatran populations was estimated by computing d_{xy}, the average number of nucleotide differences between mitochondrial genomes from the two populations (Table 3). The average pairwise distance among the four Bornean populations was 0.44 (range 0.34–0.59), nearly ten times less than the average distance between Sumatran and Bornean mtDNA genomes (4.13, range 4.09–4.55).

To determine the extent of genetic differentiation between mtDNA haplotypes in distinct geographic locales, we carried out phylogenetic analyses of the restriction site data. An unrooted phylogeny, which treated RFLP site variation as discrete phylogenetic characters and was based upon the principle of maximum parsimony [28], is presented in Figure 2. Maximum likelihood and minimum evolution trees using the Fitch–Margoliash algorithm [29] were also applied to the mtDNA–RFLP data. Each of these analyses was concordant and showed deep phylogenetic distinction between Bornean and Sumatran haplotypes, two divergent mtDNA lineages (K and M *versus* L and N) within Sumatran orang-utans, and only slight phylogenetic structure (D plus H *versus* all others) among Bornean orang-utans. There was little indication of phylogeographic structuring between the Bornean populations, suggesting that their isolation was probably very recent (Figure 2).



Full-size image (19K)

Figure 2. Unrooted phylogenetic tree generated by maximum parsimony analysis of mtDNA haplotypes treated as unordered character states using the PAUP 3.1.1 computer program [28]. An identical tree of the shortest length (number of STEPS = 154; consistency INDEX = 0.838) was determined by a heuristic search (one of two trees retained) and a bootstrap resampling analysis based on 100 iterations. Numbers on branches represent the number of steps/number of

homoplasies. Numbers in parentheses are bootstrap values (out of 100 iterations) in support of adjacent nodes. A maximum likelihood tree generated by RESTML (PHYLIP version 3.5) produced an identical topology (ln LIKELIHOOD = −912.37) with significant node bifurcation in all cases, except among Bornean haplotypes: haplotypes D and H were resolved apart from the others, whereas the remaining Bornean haplotypes were not significantly resolved among themselves. A minimum evolution analysis of haplotype nucleotide divergence (d_{xy}) using the Fitch–Margoliash algorithm of FITCH and KITSCH (PHYLIP version 3.5) produced a similar result [29]. The scale represents percent nucleotide sequence divergence from the KITSCH analysis based upon the presumption of contemporaneous tips (constant rate molecular clock) using the calibrated mtDNA divergence rate of 2 % per million years in apes. MYBP, million years before the present [15].

The maximum parsimony tree (Figure 2) has a minimum length of 152 steps, one third of which (52 steps) separate Bornean *versus* Sumatran orang-utan haplotypes (Figure 2). A total of 38 of the changes between subspecies occur only once in the tree, consistent with the occurrance of derived (synapomorphic) characters. This level of deep phylogenetic divergence between Bornean and Sumatran haplotypes, as well as the large estimated nucleotide divergence (d_{xy}) between the island populations (Table 3), suggests that the two groups have been separated for a long period. Assuming a constant rate of mtDNA substitution of 2.0 % per million years for apes [15], it can be estimated that the period of separation occurred about 1.5 million years ago (Figure 2).

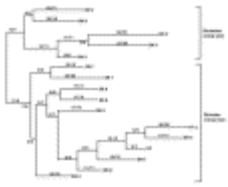
DNA fingerprint variation

The extent and patterns of nuclear minisatellite variation between orang-utans were determined using representatives with distinct mtDNA haplotypes from each population. Samples were analyzed for restriction fragment sharing following digestion with two restriction enzymes (*Hinf*I and *Hae*III) and hybridization with the human minisatellite probe 33.15. Genetic variation was assessed by computation of the mean average percent difference (MAPD) in band sharing between individuals and the estimated average heterozygosity (H) [30, 31 and 32].

The results showed that orang-utans have a considerable amount of minisatellite variation and that nearly every fragment is polymorphic. Estimates of minisatellite variation within populations and subspecies are presented in Table 2. In general, the minisatellite results affirmed the mtDNA estimates with appreciable variation in both subspecies and in three Bornean populations. A relatively low amount of variation was found in Kutai (MAPD = 25.9 %; H = 17.5 %); however, this population had a small sample size and we cannot exclude the possibility that the three Kutai individuals were related.

To explore the pattern of phylogeographic partition, we constructed minimum-length parsimony networks based on minisatellite DNA fragments for individuals from each subspecies (Figure 3). Analysis of the *Hae*III and *Hinf*I DNA fingerprints showed that orang-utans from each island subspecies were clustered together, indicating that sufficient divergence had occurred at these

minisatellite families to recapitulate the geographic separation of the island subspecies in a phylogenetic analysis. The consistency index (CI) for the topology is low (CI = 0.37), indicating that a high degree of homoplasy or parallel changes — due to allelic segregation within and between subspecies — are required to produce minimum length trees. There was no phylogenetic distinction among the four Bornean populations, consistent with recent gene flow between them.



Full-size image (16K)

Figure 3. Phylogenetic analyses of minisatellite data. A strict consensus tree was generated by PAUP 3.1.1 for the presence or absence of *Hae*III and *Hinf*I fingerprint fragments, using the human minisatellite probe 33.15. Two equally parsimonious trees were found (tree LENGTH = 411; consistency INDEX = 0.37). Percentages in parentheses refer to the number of bootstrap reiterations (out of 100) that support (> 70 %) a respective group. Numbers on limbs are number of steps/number of homoplasies. SM, Sumatra; SB, Sabah; KT, Kutai; SW, Sarawak; GP, Gunung Palung.

Mitochondrial 16S rRNA gene sequences

Homologous sequences of the mitochondrial 16S rRNA gene were obtained by polymerase chain reaction (PCR) amplification and direct sequence determination from six Sumatran and 13 Bornean orang-utans, three common chimpanzees, and two pygmy chimpanzees. These sequences were aligned with those from humans and gorillas, as shown in Figure 4.

Full-size image (174K)

Figure 4. Nucleotide sequence of mitochondrial 16S rRNA gene region from Sumatran and Bornean orang-utans, chimpanzee (*Pan troglodytes*, Ptr), pygmy chimpanzee (*Pan paniscus*, Ppa), human (*Homo sapiens*, Hum) and gorilla (*Gorilla gorilla*, Ggo). Orang-utan population and species code abbreviations are in the legend to Figure 3. The reference sequence is the

Sumatran orang-utan SM-3; dashes and letters represent identical and different nucleotides, respectively; asterisks indicate deletions.

Considerable sequence variation was apparent in the 16S rRNA genes, with five and 12 distinct genotypes found among Sumatran and Bornean orang-utans, respectively (Figure 4). Within the Sumatran samples, the average sequence mismatch between each pair of genotypes was 1.0 % (range 0.0–1.6 %); within Bornean samples it was 2.5 % (range 0.0–3.4 %). As with mtDNA–RFLPs (Table 3), the average pairwise divergence of the 16S rRNA sequence between Sumatran and Bornean orang-utans (4.8 %; range 3.4–5.7 %) was much larger than the sequence differences within the subspecies. Remarkably, the average pairwise difference between the 16S rRNA genes from the widely accepted species of pygmy chimpanzee and common chimpanzee [33 and 34] (1.6 %; range 1.3–1.8 %) was three-fold less than the average sequence difference between Bornean and Sumatran orang-utans.

The 16S rRNA gene sequences were analyzed using three phylogenetic methods: the distance-matrix-based Neighbor Joining (NJ) algorithm (PHYLIP 3.5) using Kimura distance estimates; maximum parsimony using PAUP 3.1.1; and maximum likelihood using DNAML (PHYLIP 3.5).

The results of each of these analyses (Figure 5) converged on three main conclusions. First, there was a highly significant bifurcation that separated Bornean from Sumatran 16S rRNA genotypes. Bootstrap resampling support for the separation was strong for Bornean lineage (100 % NJ and 96 % maximum parsimony) and for Sumatran lineage (75 % NJ and 90 % maximum parsimony). The maximum likelihood analyses also implied that the subspecies bifurcations were highly significant. Second, the phylogenetic divergence nodes between the two subspecies are very deep — comparable to, or greater than, those seen between the common and pygmy chimpanzees. Third, the Bornean genotypes do not display significant partitions among themselves, as indicated by low bootstrap values in the NJ and maximum parsimony analyses, and by non-significant nodes between Bornean genotypes in the maximum likelihood analysis. Furthermore, there is no apparent phylogeographic concordance among Bornean genotypes, again indicating little population substructure and, therefore, very recent isolation of the Bornean populations.


Full-size image (21K)

Figure 5. Phylogenetic tree derived from the mitochondrial 16S rRNA sequences in Figure 4 by maximum parsimony analysis using PAUP 3.1.1, designating gibbon (*Hylobates lar*) 16S rRNA as an outgroup (GenBank accession number HLU 39004). The tree was structured by a heuristic

search (one of two most parsimonious trees retained; number of STEPS = 237) and a bootstrap resampling analysis based on 100 iterations. (The second tree differed only by the disposition of sequences within the Bornean orang-utan group.) Limb lengths are assigned number of steps/number of homoplasies. The transition:transversion ratio was set at 4, the observed ratio of sequences from Figure 4. A minimum evolution tree estimated by the Neighbor Joining algorithm, based upon a matrix of pairwise nucleotide divergence among sequences, was also constructed using Kimura distance and a transition:transversion ratio of 4.0, the actual measured ratio in the sequence data. The major conclusions for both the parsimony and distance matrix method were the same. The bootstrap values out of 100 iterations for the maximum parsimony/minimum evolution analyses, respectively, are listed on each limb in parentheses. The orang-utan 16S rRNA sequences have a slightly increased similarity to human sequences compared with sequences from the African ape species (Table 4), which, in the presence of excessive orang-utan sequences, consistently led to difficulty in resolving the human–chimpanzee–gorilla trichotomy, regardless of the presence of the *Hylobates* outgroup sequence. When a subset consisting of two Bornean and two Sumatran 16s rRNA sequences was analyzed, the expected chimpanzee–human association became apparent. The failure to resolve the trichotomy with all sequences is the reason for presenting this great ape node as an unresolved polytomy. Finally, a maximum likelihood analysis [55] of the sequence data, performed using empirically derived nucleotide frequencies, produced topologies similar to the PAUP tree presented here (3101 trees examined; in likelihood, – 1148).

Table 2. Estimations of mtDNA and DNA fingerprint variation in orangutan subspecies and populations.

	mtDNA			DNA fingerprint		
Population origin	n	Number of haplotypes	π (%)	n	MAPD	H (%)
Sabah	16	4	0.11	8	41.1	44.8
Sarawak	12	5	0.35	3	36.2	28.0
Kutai	3	1	0.0	3	25.9	17.5
Gunung Palung	2	1	0.0	2	40.2	20.0
Borneo (all)	33	9	0.33	11	51.8	61.3
Sumatra	6	4	1.75	5	55.2	52.9
Borneo + Sumatra	39	13	1.46	21	62.0	73.4

Full-size table (7K)

π (%), mean nucleotide diversity in a population; MAPD, mean average percent difference in minisatellite band-sharing using two restriction enzymes, *Hae*III and *Hinf*I; H (%), estimated average heterozygosity in minisatellite locus variation (see Materials and methods for details).

Discussion

The molecular genetic distinctiveness between Bornean and Sumatran samples is considerably greater than the amount of variation detectable within Bornean or Sumatran populations. For

example, the average sequence divergence of 16S rRNA sequences between orang-utans from different islands is 4.8 % (range 3.4–5.7 %), whereas within Sumatra the average is 1.0 % (range 0.0–1.6 %). For mtDNA–RFLP, the divergence between Bornean and Sumatran samples is 10-fold greater than the nucleotide divergence among animals from the same island subspecies (Table 3). The large Bornean *versus* Sumatran distinctiveness is also apparent by the occurrence of deep phylogenetic nodes between the two subspecies with high statistical (bootstrap) support with different gene families and several phylogenetic algorithms (Figure 2, Figure 3, Figure 5).

Table 3. Estimates of molecular genetic distances between orang-utan subspecies and populations.

		Bornean				
Population/subspecies	Bornean	Sabah	Sarawak	Kutai	Gunung Palung	Sumatran
Bornean	–	–	–	–	–	4.13 %
Sabah	–	–	0.38 %	0.34 %	0.56 %	4.09 %
Sarawak	–	39.3 %	–	0.37 %	0.59 %	4.09 %
Kutai	–	36.4 %	30.3 %	–	0.41 %	4.17 %
Gunung Palung	–	40.6 %	34.3 %	32.0 %	–	4.55 %
Sumatran	61.5 %	62.8 %	63.7 %	62.0 %	64.2 %	–

Full-size table (7K)

Above diagonal: average value of nucleotide differences of mtDNA–RFLP between each population (d_{xy}); below diagonal: MAPD of microsatellite band sharing between each population using two restriction enzymes (*Hae*III and *Hinf*I).

The estimated genetic distances between orang-utan subspecies vary with different gene families, but in most cases they approach or exceed the level of genetic distance measured between common and pygmy chimpanzees (Table 4). Because common and pygmy chimpanzees are widely recognized as distinct great ape species, because of their clear morphological, ecological and behavioral differences, [33, 34 and 35], they provide a precise measure of species-level distance among great apes. The distances between the orang-utan subspecies are also somewhat larger than the recently reported genetic divergence between recognized subspecies of gorillas and common chimpanzees [16 and 19], which the authors suggested may justify species designation for these taxa.

Table 4. Estimated divergence times between orang-utans and chimpanzee species based on molecular genetic distances for different gene families.

		Pygmy versus common chimpanzee		
	Human–chimp–gorilla AGD (I)	AGD (II)	II:I Ratio	Divergence date*
Gene family				
Blood protein	36.7 %	10.3 %	0.28	1.3 MY
Allozymes	20.6 %	7.5 %	0.36	1.7 MY
2DE fibroblast proteins	8.9 %	1.7 %	0.19	0.9 MY
mtDNA *COII* sequence	11.2 %	2.7 %	0.24	1.1 MY
mtDNA 16S rRNA sequence	6.7 %	1.6 %	0.24	1.1 MY
mtDNA–RFLP	ND	2.55 %	–	–
				Average 1.0

Full-size table (14K)

*Divergence date estimated as the ratio of chimpanzee species divergence (II) to average great ape species divergence (I) times 4.7 million years, the date of great ape species divergence [36]; see text.

†Divergence date estimated as the ratio of orang-utan subspecies divergence (III) to average great ape species divergence (I) times 4.7 million years [40], the date of great ape species divergence [36]; see text. AGD: average genetic distance; ND: not determined.

The Bornean–Sumatran genetic distances are estimated as 12–72 % of the average divergence between human, chimpanzee and gorilla species (Table 4), an evolutionary period estimated at 4.7–8 million years before present [15, 36 and 37]. Taking the most conservative date for human–chimpanzee–gorilla divergence of 4.7 million years [36], the estimate of the Bornean–Sumatran orangutan divergence date ranges from 0.6–3.4 million years before present, with an average of 1.7 million years ago (Table 4). This value supports our estimate, based on mtDNA–RFLP variation, that Bornean and Sumatran populations have been separated for about 1.5 million years or longer (Figure 2), a period approaching that required for new species development [38]. Older human–chimpanzee–gorilla calibration dates [15, 37 and 39] would yield proportionately older dates for the split between Bornean and Sumatran orang-utans. Considered together, a large body of molecular phylogenetic and morphological data (Table 4; [5, 8, 9, 12, 13, 14, 16 and 17]) are consistent with the conclusion that the genetic differentiation between the two island populations of orang-utan has reached the level of distinct species [40 and 41].

An important criterion for recognizing species is the achievement of reproductive isolation in nature [42 and 43]. Because the two orang-utan populations are isolated on two islands, it is not possible to discern whether effective isolation mechanisms have evolved to reinforce reproductive isolation *in situ*. Bornean and Sumatran orang-utans do form fertile F1 hybrids in captivity, and have been interbred for up to four generations. Inter-subspecies hybridization was suspended by the Species Survival Plan of the American Zoological Association ten years ago, and birth control has been applied to living hybrids. Preliminary observations have not produced evidence for reduced fertility [44]; however, these negative results should be interpreted

cautiously because they do not assess potential fitness reduction, which might reflect ecological or ethological adaptedness [42].

The estimated period of separation between the two orang-utan island populations (*circa* 1.5 million years) does not, however, seem to be consistent with geological evidence. The islands of Borneo and Sumatra were physically connected until the late Pleistocene period, approximately 10 000–20 000 years ago [45]. Fossil records suggest that ancestors of the modern orang-utan first appeared in South China or Indo-China in early Pleistocene [2, 3, 46 and 47]. During the Pleistocene glacial period the tropical fauna moved south *via* land bridges to Sunda Land, an extensive continent stretching from the South China Sea to Java. The Sunda Land was exposed, submerged and re-exposed during the Pleistocene as many as three times [48 and 49]. There is geological evidence for an ancient river system separating modern Borneo and Sumatra when orang-utans arrived ~30 000–40 000 years ago [47], but it is unlikely that this would have posed an effective migration barrier. The geographical barriers may have been reinforced by behavioral or physiological reproductive barriers that had evolved between the two orang-utan lineages before their arrival in Sunda Land. Although we cannot be certain of this explanation, the consistent evidence of substantial genetic differentiation implies that effective or persistent hybridization has been rare or absent in nature for on the order of 1–2 million years.

Sumatran and Bornean orang-utan populations have an appreciable level of overall genetic diversity, based upon analysis of mtDNA–RFLP, minisatellites (Table 2), mitochondrial 16S rRNA genes and allozymes [12]. We conclude that neither island group has encountered a significant population bottleneck or founder effect since the Pleistocene period. This suggests that Pleistocene migration to southeast Asia was a mass movement. Furthermore, little molecular genetic differentiation exists between four geographically separate Bornean populations. Population-specific genotypes were observed for each gene family (Figure 1, Figure 4); however, the populations are closely related phylogenetically (Figure 2, Figure 3, Figure 5). Molecular phylogenetic analyses did not affirm the marked morphological distinctions reported among Bornean specimens from Gunung Palung *versus* other locales [5].

The lack of substantial molecular genetic differentiation among Bornean populations indicates that there is not a genetic imperative for relocating Bornean orang-utans to their natal populations in Borneo. Consequently, more important concerns for release of confiscated orang-utans include issues such as habitat carrying capacity, social behavior (if an animal can be socially accepted in an occupied area and reproductively integrate) and infectious disease. A further understanding of these and other factors affecting species survival is important for developing efficacious management programs for endangered species.

Materials and methods

Biological specimens of free-ranging orang-utans

Using remote biopsy darts [50], skin samples of wild and presumed unrelated orang-utans were collected from two locations in Sumatra ($n = 6$) and nine locations in Borneo ($n = 33$) (Figure 1). Fibroblast cell lines were established by tissue culture (as described previously) [12 and 51].

Genomic DNA from nuclei and mitochondria was extracted from the cell lines using the phenol–chloroform method.

Analysis of mtDNA–RFLPs

Genomic DNA (1 μg) from each animal was digested separately with the 30 restriction enzymes listed in Table 1, separated by electrophoresis in 1 % agarose gels and transferred to nylon filters (UV Duralon; Stratagene) by Southern blotting. DNA fragments on the membrane were hybridized with a [^{32}P]dCTP-labeled molecular clone of cat mitochondrial DNA [52], and visualized by autoradiography. Intrapopulation variation was estimated using π, an index of nucleotide diversity which measures the probability that two randomly selected sequences from two individuals within a population will have different nucleotides at a given position [53]. Average nucleotide diversity between populations (d_{xy}) [53] estimates the probability that two randomly selected sequences from two populations will not share the same sites. The net nucleotide diversity between two populations, d_a, which discounts the intrapopulation variations, is calculated as follows: $d_a = d_{xy} - (\pi_x + \pi_y)/2$. Values for π, d_{xy} and d_a were calculated by the computer program MAXLIKE [54].

DNA fingerprint (minisatellite) analysis

DNA (6 μg) from individual samples was digested with *Hinf*I and *Hae*III, separated by electrophoresis in 1 % agarose gels, transferred to nylon filters, and hybridized to the [^{32}P]dCTP-labeled human minisatellite probe 33.15 [30 and 31]. Population variation was estimated by computation of the average percent difference (APD), a measure of band-sharing between individuals, the mean value of APD from different enzymes (MAPD), and average heterozygosity (H) [32].

Mitochondrial 16S rRNA sequence analysis

A 387 bp sequence was obtained by PCR amplification and direct sequencing (ABI Sequencer) using oligonucleotide primers: 5′–GTGCAAAGGTAGCATAATCA–3′ and 5′–TGTCCTGATCCAACATCGAG–3′ (A.R. Hoelzel, personal communication). Six Sumatran samples and 13 Bornean samples from each mtDNA–RFLP haplotype and from each location were used for this analysis. DNAs from three common chimpanzees and three pygmy chimpanzees were also sequenced. Chimpanzee samples were from unrelated animals located at the San Diego Zoo and the Institute for Medical Research. Human and gorilla sequences were obtained from GenBank. Pairwise distance was obtained by computing the percent base pair divergence between two individuals with gaps given a weight of one residue. All derived sequences have been submitted to GenBank; see below for details.

Accession numbers

The following GenBank accession numbers have been assigned to the mitochondrial 16S rRNA sequences: U63486–U63510.

Phylogenetic analysis of data

Phenograms describing the associations among individuals and populations were constructed from the distance matrix for each gene family using the Neighbor Joining and the Fitch–Margoliash algorithm (FITCH and KITSCH; PHYLIP version 3.5) [29]. Character data for mtDNA restriction sites and minisatellite fragments were generated for each individual and population by transforming allelic frequencies into discrete character states – that is, each polymorphic site was coded as a discrete character and scored for its presence or absence in each individual, subspecies, or population. Character data including nucleotide sequences were analyzed by maximum parsimony using the program PAUP version 3.1.1 [35], and by a maximum likelihood algorithm available in RESTML (for mtDNA–RFLP) and in DNAML (for nucleotide sequences) of PHYLIP version 3.5 [29].

References

References

1. 1. ML Jones, The geographic races of the orang utan. In: HO Hofer, Editor, *Recent Advances in Primatology*, Karger, Basel (1969), pp. 217–223.
2. 2. J MacKinnon, The behaviour and ecology of wild orang utans. *Anim Behavior* **22** (1974), pp. 3–74. Abstract | PDF (17826 K) | View Record in Scopus | Cited By in Scopus (114)
3. 3. J MacKinnon. In: *In Search of the Red Ape*, Holt, Reinehart and Winston, New York (1974).
4. 4. JJC Mallinson, Cocktail orang-utan and the need to preserve purebred stock. *Dodo* **15** (1978), pp. 69–77.
5. 5. CP Groves, CB Westwood and BT Shea, Unfinished business: Mahalanobis and a clockwork orang. *J Hum Evol* **22** (1992), pp. 327–340. Abstract | **Article** | PDF (944 K) | View Record in Scopus | Cited By in Scopus (31)
6. 6. BT Shea and HJ Coolidge, Craniometric differentiation and systematics in the genus *Pan*. *J Hum Evol* **17** (1988), pp. 671–685. Abstract | **Article** | PDF (1127 K) | View Record in Scopus | Cited By in Scopus (16)
7. 7. HN Seuanez, HJ Evans, DE Martin and J Fletcher, An inversion of chromosome 2 that distinguishes between Bornean and Sumatran orangutans. *Cytogenet Cell Genet* **23** (1979), pp. 137–140. **Full Text** via CrossRef | View Record in Scopus | Cited By in Scopus (18)
8. 8. OA Ryder and LG Chemnick, Chromosomal and mitochondrial DNA variation in orang utans. *J Hered* **84** (1993), pp. 405–409 94014292 . View Record in Scopus | Cited By in Scopus (18)
9. 9. E Bruce and F Ayala, Phylogenetic relationships between man and the apes: electrophoretic evidence. *Evolution* **33** (1979), pp. 1040–1056. **Full Text** via CrossRef
10. 10. LEM De Boer and PM Khan, Haemoglobin polymorphisms in Bornean and Sumatran orangutans. In: LEM de Boer, Editor, *The Orang Utan: Its Biology and Conservation*, W Junk, The Hague (1982), pp. 125–134.
11. 11. JT Wijinen, H Rijksen, LEM De Boer and PM Khan, Glucose-6-phosphate-dehydrogenase (G6PD) variation in the orang utan. In: LEM de Boer, Editor, *The Orang Utan: Its Biology and Conservation*, W Junk, The Hague (1982), pp. 109–118.

12. 12. DN Janczewski, D Goldman and SJ OBrien, Molecular genetic diversity of orang-utan subspecies based on isozyme and two dimensional gel electrophoresis. *J Hered* **81** (1990), pp. 375–387 91036893 . View Record in Scopus | Cited By in Scopus (20)
13. 13. A Caccone and JR Powell, DNA divergence among hominoids. *Evolution* **45** (1989), pp. 925–942. **Full Text** via CrossRef
14. 14. SD Ferris, WM Brown, WS Davidson and AC Wilson, Extensive polymorphism in the mitochondrial DNA of apes. *Proc Natl Acad Sci USA* **78** (1981), pp. 6319–6323 82082449 . **Full Text** via CrossRef | View Record in Scopus | Cited By in Scopus (47)
15. 15. AC Wilson, RL Cann, SM Carr, M George, U Gyllensten and KM Helm-Bychowski, et al., Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol J Linn Soc* **26** (1985), pp. 375–400. **Full Text** via CrossRef | View Record in Scopus | Cited By in Scopus (393)
16. 16. M Ruvolo, D Pan, S Zehr, T Goldberg, TR Disotell and M von Dornum, Gene trees and hominoid phylogeny. *Proc Natl Acad Sci USA* **91** (1994), pp. 8900–8904 94377460 . **Full Text** via CrossRef | View Record in Scopus | Cited By in Scopus (97)
17. 17. J Raloff, Caste-off orangs: controversy surrounds implications of a hybrid label. *Sci News 1995* **147** (1995), pp. 184–189. **Full Text** via CrossRef
18. 18. A Uchida, What we dont know about great ape variation. *Trends Ecol Evol* **11** (1996), pp. 163–168. **Article** | PDF (949 K) | View Record in Scopus | Cited By in Scopus (15)
19. 19. PA Morin, JJ Moore, R Chakraborty, L Jin, J Goodall and DS Woodruff, Kin selection, social structure, gene flow, and the evolution of chimpanzees. *Science* **265** (1994), pp. 1193–1201 94345392 . View Record in Scopus | Cited By in Scopus (300)
20. 20. V Geist, Endangered species and the law. *Nature* **357** (1992), pp. 274–276 92269950 . **Full Text** via CrossRef | View Record in Scopus | Cited By in Scopus (16)
21. 21. SJ OBrien, Genetic and phylogenetic analyses of endangered species. *Annu Rev Genet* **28** (1994), pp. 467–489 95200265 . View Record in Scopus | Cited By in Scopus (49)
22. 22. S Anderson, AT Bankier, BG Barrell, MH de Bruijn, AR Coulson and J Drouin, et al., Sequence and organization of the human mitochondrial genome. *Nature* **290** (1981), pp. 456–457 81173052 .
23. 23. R Cann, M Stoneking and A Wilson, Mitochondrial DNA in human evolution. *Nature* **325** (1987), p. 31. **Full Text** via CrossRef | View Record in Scopus | Cited By in Scopus (801)
24. 24. SJ OBrien, ME Roelke, N Yuhki, KW Richards, WE Johnson and WL Franklin, et al., Genetic introgression within the Florida panther (*Felis concolor coryi*). *Natl Geo Res* **6** (1990), pp. 485–494. View Record in Scopus | Cited By in Scopus (39)
25. 25. S Miththapala, J Seidensticker and SJ OBrien, Molecular genetic diversity within and between subspecies of leopards (*Panthera pardus*). *Cons Biol* **10** (1996), pp. 1115–1132. View Record in Scopus | Cited By in Scopus (26)
26. 26. JC Avise, C Giblin-Davidson, J Larem, JC Patton and RA Lansman, Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher (Geomys pinetis). *Proc Natl Acad Sci USA* **76** (1979), pp. 6694–6698. **Full Text** via CrossRef | View Record in Scopus | Cited By in Scopus (62)
27. 27. CS Baker, SR Palumbi, RH Lambertson, MT Weinrich, J Caalmbokidis and SJ OBrien, Influence of seasonal migration on geographic distribution of mitochondrial

DNA haplotypes in humpback whales. *Nature* **344** (1990), pp. 238–240 90190837 . **Full Text** via CrossRef | View Record in Scopus | Cited By in Scopus (67)
28. 28. DL Swofford. In: *Phylogenetic Analysis Using Parsimony (PAUP), Version 3.0*, Natural History Survey, Illinois (1991).
29. 29. J Felsenstein. In: *PHYLIP: Phylogenetic Inference Package, 1993, Version 3.5*, University of Washington, Seattle (1993).
30. 30. AJ Jeffreys, V Wilson and SL Thein, Individual specific fingerprints of human DNA. *Nature* **316** (1985), pp. 76–79 85240582 . **Full Text** via CrossRef | View Record in Scopus | Cited By in Scopus (556)
31. 31. DA Gilbert, C Packer, AE Pusey, JC Stephens and SJ OBrien, Analytical DNA fingerprinting in lions: parentage, genetic diversity, and kinship. *J Hered* **82** (1991), pp. 378–386 92043591 . View Record in Scopus | Cited By in Scopus (78)
32. 32. JC Stephens, DA Gilbert, N Yuhki and SJ OBrien, Estimation of heterozygosity for single-probe, multilocus DNA fingerprints. *Mol Biol Evol* **9** (1992), pp. 729–743 92334143 . View Record in Scopus | Cited By in Scopus (89)
33. 33. AD Horn, The taxonomic status of the bonobo chimpanzee. *Am J Phys Anthrop* **51** (1979), pp. 273–282.
34. 34. RL Susman. In: *The Pygmy Chimpanzee: Evolutionary Biology and Behavior*, Plenum, New York (1984).
35. 35. DE Wilson and DM Reeder. In: *Mammal Species of the World, a Taxonomic and Geographic Reference,* (2 ed.), Smithsonian Institute Press, Washington, DC (1993).
36. 36. S Horai, Y Satta, K Hayasaka, R Kondo, T Inoue and T Ishida, et al., Mans place in hominoidea revealed by mitochondrial DNA genealogy. *J Mol Evol* **35** (1992), pp. 32–43. **Full Text** via CrossRef | View Record in Scopus | Cited By in Scopus (90)
37. 37. CG Sibley and JE Ahlquist, DNA hybridization evidence of hominoid phylogeny: results from an expanded data set. *J Mol Evol* **26** (1987), pp. 99–121 88141307 . **Full Text** via CrossRef | View Record in Scopus | Cited By in Scopus (97)
38. 38. SM Stanley. In: *Extinction*, W.H. Freeman and Company, New York (1987).
39. 39. Y Coppens, East side story: the origin of humankind. *Sci Am* **270** (1994), pp. 88–95 94255755 . **Full Text** via CrossRef | View Record in Scopus | Cited By in Scopus (33)
40. 40. CP Groves, Systematics of the great ape. In: DR Swindler and J Erwin, Editors, *Comparative Primate Biology: Systematics, Evolution, and Anatomy Vol 1*, Liss, New York (1986), pp. 187–217.
41. 41. P Andrews. *Primate Eye* **23** (1984), pp. 34–37.
42. 42. . E Mayr, Editor, *Animal Species and Evolution*, Harvard University Press, Cambridge (London) (1963).
43. 43. . D Otte and JA Endler, Editors, *Speciation and Its Consequences*, Sinauer, Sunderland, Massachusetts (1989).
44. 44. RJ Markham, Captive orang utans: present status and future prospects. *Bull Zoo Manag* **23** (1985), pp. 31–36.
45. 45. JHF Umbgrove. In: *Structural History of the East Indies*, Cambridge University Press, Cambridge (1949).
46. 46. GHR von Koeningswald, The relationship between the fossil mammalian faunae of Java and China, with special reference to early man. *Peking Natl Hist Bull* **13** (1938–1939), pp. 293–298.

47. 47. RJ Smith and DR Pilbeam, Evolution of the orang-utan. *Nature* **284** (1980), pp. 447–448. **Full Text** via CrossRef | View Record in Scopus | Cited By in Scopus (15)
48. 48. E Banks, The distribution of mammals and birds in the south China sea and west Sumatran islands. *Bull Natl Hist Mus Singapore* **30** (1961), pp. 92–96.
49. 49. D Jongsma, Eustatic sea level changes in the Arafura sea. *Nature* **228** (1970), pp. 150–151. **Full Text** via CrossRef | View Record in Scopus | Cited By in Scopus (9)
50. 50. WB Karesh, F Smith and H Frazier-Taylor, A remote method for obtaining skin biopsy samples. *Cons Biol* **3** (1987), pp. 261–262. **Full Text** via CrossRef
51. 51. WS Modi, WG Nash, AC Ferrari and SJ OBrien, Cytogenetic methodologies for gene mapping and comparative analyses in mammalian cell culture systems. *Gene Anal Tech* **4** (1987), pp. 75–85 89172772 . Abstract | PDF (1859 K) | View Record in Scopus | Cited By in Scopus (24)
52. 52. JV Lopez, N Yuhki, R Masuda, WS Modi and SJ OBrien, *Numt*, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. *J Mol Evol* **39** (1994), pp. 174–190 95018295 . View Record in Scopus | Cited By in Scopus (173)
53. 53. M Nei and W-H Li, Mathematical model for studying genetic variation in the terms of restriction endonucleases. *Proc Natl Acad Sci USA* **76** (1979), pp. 5269–5273. **Full Text** via CrossRef | View Record in Scopus | Cited By in Scopus (3614)
54. 54. M Nei and F Tajima, Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. *Genetics* **105** (1983), pp. 207–217 84005840 . View Record in Scopus | Cited By in Scopus (153)
55. 55. J Felsenstein, Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* **17** (1981), pp. 368–376 82033219 . **Full Text** via CrossRef | View Record in Scopus | Cited By in Scopus (3332)

Corresponding author.

Vitae

- L Zhi, DN Janczewski, JS Martenson and SJ O'Brien (corresponding author), Laboratory of Genomic Diversity, National Cancer Institute, Frederick, Maryland 21702-1201, USA.
- WB Karesh, New York Zoological Society, Bronx, New York 10460, USA.
- H Frazier-Taylor, Woodland Park Zoological Gardens, Seattle, Washington 98103, USA.
- D Sajuthi, Institut Pertanian Bogor, Bogor, Indonesia.
- F Gombek, National Parks and Wildlife Division, Kuching, Sarawak, East Malaysia.
- M Andau, Department of Wildlife, Kota Kinabalu, Sabah, East Malaysia.

Present address for L Zhi, Department of Environmental Biology and Ecology, College of Life Sciences, Peking University, Beijing 100871, China.
Present address for DN Janczewski, Consolidated Safety Services, 4031 University Drive, Fairfax, Virginia 22030, USA.