Genetic Affinities of the Andaman Islanders, a Vanishing Human Population

Kumarasamy Thangaraj,[1] Lalji Singh,[1] Alla G. Reddy,[1] V. Raghavendra Rao,[2] Subhash C. Sehgal,[3] Peter A. Underhill,[4] Melanie Pierson,[5] Ian G. Frame,[6] and Erika Hagelberg[6,7]

[1]Centre for Cellular and Molecular Biology
Uppal Road
Hyderabad 500 007
[2]Anthropological Survey of India
C.G.O. Complex
Seminary Hills
Nagpur 440 006
[3]Regional Medical Research Centre
Post Bag No. 13
Port Blair 744 101
Andaman and Nicobar Islands
India
[4]Department of Genetics
Stanford University
300 Pasteur Drive
Stanford, California 94305-5120
[5]Biochemistry Department
University of Otago
P.O. Box 56
Dunedin
New Zealand
[6]Department of Biology
University of Oslo
P.O. Box 1050 Blindern
N-0316 Oslo
Norway

[7]Correspondence: erika.hagelberg@bio.uio.no

Summary

Background: The Andaman Islands in the Bay of Bengal are inhabited by hunter-gatherers of unknown origin, now on the verge of extinction. The Andamanese and other Asian small-statured peoples, traditionally known as "Negritos," resemble African pygmies. However, it is generally believed that they descend from the early Australo-Melanesian settlers of Southeast Asia and that their resemblance to some Africans is due to adaptation to a similar environment, rather than shared origins.
Results: We analyzed mitochondrial DNA (mtDNA) sequences and RFLP polymorphisms, and Y chromosome biallelic markers and microsatellites, in present-day Andamanese of the Onge, Jarawa, and Great Andamanese tribes, and of inhabitants of the neighboring Nicobar Islands. We also analyzed mtDNA sequences from Andamanese hair samples collected by an ethnographer during 1906–1908. Living Andamanese exhibit low genetic variability that is consistent with their small population size and reproductive isolation.
Conclusions: Our data indicate that the Andamanese have closer affinities to Asian than to African populations and suggest that they are the descendants of the early Palaeolithic colonizers of Southeast Asia. In contrast, the Nicobarese have genetic affinities to groups widely distributed throughout Asia today, presumably descended from Neolithic agriculturalists.

Introduction

The native inhabitants of the Andaman Islands of the Indian Ocean (Figure 1) are one of several isolated groups of small-statured hunter-gatherers, sometimes known as "Negritos," who survive in isolated parts of Asia, like Malaysia and the Philippines [1]. The Andamanese earned a reputation for ferocity due to their violent resistance to foreign intrusions. They remained comparatively isolated from the outside world until the establishment of a British penal settlement in the islands, after the Indian Mutiny of 1857. The British befriended one of the tribes, the Great Andamanese, and employed them as bush police to recapture escaped convicts. The Great Andamanese suffered the brunt of colonialism, and, in the 19th and 20th centuries, their number collapsed from several hundred to a few dozen individuals. The Andaman and Nicobar Islands are now part of the territory of the Republic of India, with local administrative headquarters at Port Blair in the Andamans.

The origin of the Andamanese has been the subject of speculation for centuries. As recently as 1973, an author wrote, "I cannot explain the Andamanese, and nothing is known about their past" [2], and little more is known today. The Andamanese are near extinction following the population reduction caused by "pacification" and disease. Most of the

Immediate Early Publication

Published online November 26, 2002
DOI: 10.1016/S0960982202013362

present-day inhabitants of the islands stem from the Indian mainland. Of the dozen linguistically distinct tribes who populated the islands in the early 20th century, only four survive, the Sentinelese, Jarawa, Great Andamanese, and Onge, with a combined population of 400–500 individuals. The first two groups still maintain a traditional forest-dwelling lifestyle, while the Great Andamanese and Onge are settled and receive food and other aid from the Indian government [3].

The Andamanese and other Asian "Negrito" peoples share physical features, including short stature, dark skin, peppercorn hair, scant body hair, and sometimes steatopygia, with African pygmies. But, whereas the Asian pygmoid people from Southeast Asia to New Guinea resemble Melanesians in their cranial morphology, Andamanese crania resemble more closely those of Africans [4]. However, a 1950s study on blood groups and proteins suggested that the Andamanese were more closely related to Oceanic peoples than to Africans [5]. Genetic studies on Philippine Negritos, based on polymorphic blood enzymes and antigens, showed that they were similar to surrounding Asian populations and rejected the notion that they belonged to an ancient stratum of *Homo sapiens* in Asia [6]. The most favored current explanation of the origin of the Andamanese and other Negritos is that they are short-statured representatives of the early Australo-Melanesian settlers of Southeast Asia and Oceania, and not closely related to the African pygmoid peoples [7]. It has been suggested that the small stature of some nomadic hunting and gathering peoples in Asia and Africa might be a local adaptation to a tropical rainforest environment, rather than the result of a shared ancestry [8].

Little is known also about the Nicobarese peoples, who inhabit a small archipelago south of the Andamans as well as scattered locations in the Andaman Islands. They are thought to be related to the Malays and Burmese, and they speak dialects related to Mon-Khmer and the languages spoken in Vietnam, Malaysia, and parts of northeast India [9]. This view is supported by the result of mtDNA sequence analysis on 33 Nicobarese Islanders, which indicated affinities to populations of mainland Southeast Asia [10]

We present DNA sequence data on Y chromosome markers, and maternally transmitted mtDNA coding and non-coding region data, for several groups of Andaman Islanders and a sample of Nicobarese. This is the first molecular genetic evidence on the affinities of the Andaman Islanders, arguably the most enigmatic people on our planet.

Results

Mitochondrial DNA Analysis

The present-day Andamanese, bar the four Jarawa, were tested for the presence of the intergenic 9-base pair (9-bp) deletion common in Asian populations [11] and for their allelic status at position 10400, a C to T transition that defines the major "Asian" haplogroup M [12]. None of the Andamanese exhibited the 9-bp deletion. All but two of the Andamanese had the 10400T substitution, indicating that they belonged to haplogroup M.

Seven of the Andamanese (four Onge and three Great Andamanese) were analyzed further by DNA sequencing of selected mtDNA coding sites to investigate their affinity to previously identified subclades of haplogroup M, characterized by substitutions at the following sites: positions 6446, 6680, 12403, and 14110 [13], which define the M1 haplogroup in Ethiopia; position 11946, which defines the M_1 and M_2 haplogroups in India and the Philippines [13]; sites 153 and 3394 (M9), 4715 and 15487 (M8), 4833 (G), 5178 (D), 9824 (M7), and 10646 (M10), which define various M subgroups (shown in parentheses) observed in Han Chinese [14]; sites 7598, specific to haplogroup E of Borneo and Malaysia, and 16144, which defines haplogroup Q, a subclade of M common in Papua New Guinea. The seven Andamanese were identical to the mtDNA reference [15] at these sites and lacked the substitutions known to define previously observed M subclades.

Sequences of approximately 400 bp of the first hypervariable region (HVR1) were generated for 87 present-day Andamanese. Only ten different mtDNA haplotypes were observed (Table 1). The Andamanese sequences within haplogroup M were of two HVR1 types. The first type, seen in all three Andamanese tribes, was represented by a single haplotype, And.hap1, and was characterized by substitutions at 16223 and 16311. The second type, represented by seven haplotypes, had substitutions at positions 16213, 16223, 16311, 16319, 16344, and 16357. Two further haplotypes were observed in two Great Andamanese, who lacked the 10400T substitution, and had control region substitutions at 16169, 16349, 16265, 16288, 16304, and 16311, suggesting that they were affiliated with Asian haplogroup F [14]. Nine of the haplotypes seemed unique to the

Immediate Early Publication **Published online November 26, 2002**

DOI: 10.1016/S0960982202013362

Andamans, while haplotype And.hap1 matched two previously identified individuals from the Indian mainland, known to belong to haplogroup M [16]. There was insufficient DNA in the four Jarawa samples to determine their status at position 10400 and the 9-bp deletion, but they had the same HVR1 sequences as other Andamanese (And.hap1 and And.hap6).

mtDNA data were generated for 47 Andaman hair samples collected at the beginning of the 20th century by the Cambridge ethnographer Alfred Radcliffe-Brown [17]. None exhibited the 9-bp deletion. Amplification of a 239-bp HVR1 fragment was possible for 42 of the hair extracts. A total of 27 different sequences of 190 bp were observed. The sequences of 13 of the hair samples matched those of living Andamanese in the section sequenced; namely, two matched And.hap1, three matched And.hap6, and eight matched And.hap7.

Seven different mtDNA haplotypes were detected in 13 Nicobarese. Four individuals had the 9-bp deletion and 10400C. They belong to the subset of lineages with the 9-bp deletion present throughout Asia. Seven Nicobarese also had 10400C, without the 9-bp deletion. These individuals had HVR1 sequences typical of Asian haplogroup F.

Y Chromosome Analysis

DNA of 48 present-day Andaman and Nicobar males was analyzed by using 23 binary markers. Six Y chromosome binary haplotypes, falling into four previously defined major global haplogroups, IV, VII, VIII, and X [18], also known as D, O, L/K/M, and P/Q, respectively [19], were detected. All Onge and Jarawa had the same binary haplotype D, previously observed in Asia [20]. All Nicobarese males exhibited haplotype O, previously observed in Southeast Asia [21–23]. The Great Andaman males had five different binary haplotypes, found previously in Southeast Asia, the Indian subcontinent, and Melanesia [18] (Figure 2).

Five Y chromosome microsatellite loci were analyzed in 37 males. Table 2 shows the haplotypes, the haplogroup affiliations, and the mean variance of the number of repeats across the five loci. Three of the microsatellite loci were monomorphic in the Onge, and one was monomorphic in the Nicobarese. All loci were polymorphic in the Great Andamanese, consistent with their higher haplogroup diversity.

Discussion

Our analysis of mtDNA and Y chromosome polymorphisms provides new insights into the history of the Andamanese. All the Onge, and all but two of the Great Andamanese, belonged to the previously described mtDNA haplogroup M, found in East Asia [14] and South Asia [16]. It has been suggested that haplogroup M is a genetic indicator of the migration of modern *Homo sapiens* from eastern Africa toward Southeast Asia, Australia, and Oceania [12]. Analysis of mtDNA coding sites indicated that these Andamanese fall into a subgroup of M not previously identified in human populations in Africa and Asia; this finding suggests an early split from these populations.

The mtDNA control region sequences recovered from present-day Andamanese people and the archival hair samples had a suite of substitutions, at positions 16223, 16311, 16319, 16344, and 16357, specific to the Andamanese and no other population. One Andamanese haplotype, And.hap1, shared substitutions at positions 16223 and 16311 with some Indians belonging to haplogroup M. It is worth noting that the transition at 16357 is only present at high frequencies in the Andaman Islands and in one of the major founding lineages of New Guinea and island Melanesia [24]. Interestingly, the linguist Joseph Greenberg proposed that the Andaman languages belonged to the same language phylum, the Indo-Pacific Phylum, as the languages of Tasmania, parts of Papua New Guinea, and island Melanesia [25]; this finding indicates a possible common origin of the peoples of these regions.

The mtDNA data on hair samples collected by the ethnographer Alfred Radcliffe-Brown at the beginning of the 20th century are of unique interest and are a valuable frame of reference for future studies. We tested 47 hair extracts for the 9-bp deletion and generated sequence information from 42 of the 70 hair extracts. The remaining extracts failed to yield a PCR product after repeated attempts, suggesting the original DNA was too degraded. Only short fragments of mtDNA could be retrieved from the almost century-old hair samples, but the mtDNA control region sequences of the hair extracts had the same pattern of nucleotide substitutions as that of living Andamanese and matched the latter in several instances. This suggests that the hair sequences are authentic. However, there were some individual mutations, which in some cases might be genuine but in others might have been caused by polymerase infidelity in the first cycles of

Immediate Early Publication

Published online November 26, 2002
DOI: 10.1016/S0960982202013362

amplification due to low template concentration. There was no opportunity for cross contamination with modern Andaman DNA, as the Andamanese hair and blood samples were extracted and analyzed independently in different countries, by different personnel, and at different times. No evidence of laboratory contamination was detected in the hair extracts. We present the sequences for their intrinsic historical interest, for comparison with the data on the extant Andamanese, and to demonstrate the potential of studies on archival materials.

The position of the Andamanese in the global Y chromosome phylogeny [18] is shown in Figure 2. Haplogroups A and B are present in Africans only. Clades C–Q are part of three major branches within the world phylogeny, with the M168 mutation at their root [26]. Haplogroup C is absent in Africa but present in Oceania, Asia, and North America [27], while haplogroups F–Q are rare in Africa but widespread elsewhere [18, 28]. The Onge and Jarawa males belong to the Asian-specific Y chromosome clade IV, also known as D [19].

Haplogroups D and E are characterized by the YAP+ insertion mutation. This mutation probably originated in Africa [20, 29] and is virtually fixed in sub-Saharan Africans, although it is also present in Asia [30]. African YAP+ lineages have the SRY4064 transition that characterizes haplogroup E [31], whereas those in Asia have M174, which defines haplogroup D [20]. Two Asian YAP+ lineages have been observed in Asia. One has the M15 mutation and is common in Tibet [21, 32, 33]. The second, with the M55 mutation, is common in Japan [18]. The Onge and Jarawa men had neither mutation, and they represent a subtype of haplogroup D not described previously.

In contrast to the single Y haplogroup observed in the Onge and Jarawa males, the Great Andamanese had five different binary mutation haplotypes, falling into haplogroups O, L, K, and P, among ten men. This suggests admixture with Indian (L-M11 and P-M45) and East Asian (O-M122 and O-M95) male lineages. European Y chromosome lineages, like those related to R-M173 [34], were not observed. These findings reflect the history of the Great Andamanese in colonial times. The tribe was pacified forcibly by the British in the 19th century and coexisted with thousands of male laborers and convicts from the Indian mainland. The few surviving members of the Great Andamanese tribe are now settled in Strait Island in the Andamans and bear little physical resemblance to other Andamanese.

The mtDNA diversity of the Great Andamanese was very low. The mtDNA haplotype diversity, mean pairwise differences between sequences, and nucleotide diversity in the Great Andamanese was the lowest of any population we examined (Table 3). Of the 20 individuals typed, 15 had the same haplotype, And.hap1, also present in the Onge and Jarawa. Two of the Great Andamanese had different mtDNA types than other Andaman types and lacked the 10400T mutation. It is possible that these sequences derive from non-Andaman women.

The differences between the mtDNA and Y chromosome haplotype substructure in the Andamanese indicate that the maternal and paternal lineages suffered different fates. It is likely that a proportion of the original maternal lineages have persisted in the Onge, Jarawa, and Great Andamanese. The Y chromosome haplogroup of the Onge and Jarawa probably represents an original paternal lineage, whereas it is likely that the Y chromosomes of the Great Andamanese are the result of admixture with non-Andaman males.

All the male Nicobarese in our sample had the same Y chromosome binary haplotype, O-M95. This is not unexpected, in view of what is known about the subsistence and languages of these people. Haplogroup O is very widely distributed in present-day Asian populations, which is probably the result of the demographic spread of Neolithic peoples in recent millennia [20]. The Nicobarese are sedentary agriculturalists thought to be of Neolithic descent. They resemble the Burmese in physical appearance and speak languages of the Autroasiatic family of mainland Southeast Asia [9].

Archaeological data for the Andaman Islands are relatively scarce, and the oldest radiocarbon dates are only about 2000 years before present [35]. However, the distinctive culture, appearance, and languages of the Andamanese argues for a longer separation from the surrounding mainland Asian populations. The meager archaeological record enhances the value of molecular genetic studies, although the lack of relevant independent archaeological data also prevents the verification of the conclusions drawn from molecular data. Nevertheless, the presence of a hitherto unidentified subset of the mtDNA Asian haplogroup M, and of the Asian-specific Y chromosome group D, is consistent with the view that the Andamanese are the descendants of Palaeolithic peoples who might have been widely dispersed in Asia in the past. Geographical isolation probably aided the

Immediate Early Publication **Published online November 26, 2002**

DOI: 10.1016/S0960982202013362

survival of ancient human lineages in the Andamans and other locations, whereas, in most of Asia, they were later largely assimilated or supplanted by Neolithic agriculturalists.

The phylogeography of Y chromosome binary haplotypes suggests that early modern human colonizers of East Asia and Southeast Asia belonged to both the C and D lineages [20]. It is possible that these people spread during episodes of low sea levels in the Pleistocene. D lineages are unevenly distributed on mainland Asia and have only been detected at high frequencies in Japan and the Himalayas [18, 22, 32, 33]. In Japan, such lineages are thought to be the genetic trace of a pre-Jomon culture [36]. C lineages are found in India, including in tribal populations [37], and reach their greatest frequency in Oceania [23, 38].

In conclusion, molecular genetic data support the view that the Andamanese belong to an ancient human substratum in Asia, later replaced in most of that continent by Neolithic agriculturalists. Our data agree with linguistic studies that suggest that the Andamanese languages might be the last representatives of pre-Neolithic Southeast Asia, perhaps going back to the initial settlement by modern humans, and might be largely unaffected by the spread of large language families, like the Austronesian language family, which occurred in Southeast Asia during the Neolithic [39]. Intriguingly, Andamanese languages have grammatical and lexical affinities to the language spoken by the Kusunda, an isolated Nepalese ethnic group with the only language of the Asian mainland thought to belong to the Indo-Pacific Phylum (M. Ruhlen, personal communication).

The low diversity in mtDNA, and in both Y chromosome binary markers and STRs, is genetic evidence of the population collapse suffered by the Andamanese. Although it is unlikely that the Andamanese genes will vanish completely, the irretrievable loss of the unique culture and languages of the Andamanese peoples is a tragedy.

Experimental Procedures

Modern DNA Samples

Blood samples were collected from 68 Onge, 20 Great Andamanese, and 14 Nicobarese (of whom 23, 10, and 11, respectively, were male). Buccal swabs were collected from four Jarawa males. Samples were collected after consultation with the Directorate of Tribal Welfare, the local administration, medical officer, and the donors. DNA was isolated from the samples by using proteinase K digestion and phenol-chloroform extraction and was analyzed for polymorphisms in mtDNA and the Y chromosome. Genetic analysis of the modern DNA extracts was performed at the Centre for Cellular and Molecular Biology, Hyderabad, India.

Archival Hair DNA Samples

DNA was extracted from samples of 70 hair locks collected by A. Radcliffe-Brown during an expedition to the Andaman Islands between 1906 and 1908; this collection was later part of the Duckworth Collection, University of Cambridge, United Kingdom. Radcliffe-Brown's account suggests that the samples were from Great Andamanese peoples. Hair DNA was extracted with strict precautions to avoid contamination by modern DNA [40] by using a modification of a previously described method [41, 42]. Extraction and analysis of the hair DNA were performed at the former ancient DNA laboratory at the University of Cambridge.

Analysis of mtDNA Loci

DNA extracts were analyzed by PCR amplification of selected regions of the mtDNA locus. Firstly, they were assayed for the presence of the intergenic 9-bp deletion by amplification of a 120-bp fragment, as described previously [11], and subsequent electrophoresis on 3% NuSieve: 1% agarose gels. mtDNA sequencing of the modern DNA extracts was carried out as described previously, with minor modifications [43].

Hair DNA was amplified with primers (shown in a 5′ to 3′ direction) L16210: TCCCCATGCTTACAAGCAAGTA and H16405: CGGGATATTGATTTCACGGAGGAT, which specify a 239-bp fragment of the mtDNA control region. A subset of ten extracts was amplified with the internal primer H16314: GTGCTATGTACGGTAAATGGCTTTAT, which yielded a 148-bp fragment. PCR products were used as templates for direct sequencing on an Applied Biosystems ABI 377A DNA sequencer.

Present-day individuals were assayed for their allelic status at mtDNA positions 10398 and 10400, diagnostic for the Asian haplogroup M, with the previously published primers 15F and 15R [44]. Seven individuals were also analyzed at selected mtDNA coding sites to

Immediate Early Publication

Published online November 26, 2002
DOI: 10.1016/S0960982202013362

determine their affinity to previously characterized African and Asian haplogroups.

Relevant population statistics were calculated with the ARLEQUIN software package [45] and included the haplotype diversity (h), the mean number of pairwise differences between sequences (π), and nucleotide diversity.

Analysis of Y Chromosome Binary Polymorphisms

The following 23 biallelic markers were analyzed: YAP, RPS4Y, SRY4064, sY81 = M2, M5, M9, M11, M15, M45, M55, M70, M95, M119, M120, M122, M145, M173, M174, M175, M176, M203 [20], 92R7 [46], and SRY1532 [47]. Nested polymerase chain reaction (PCR) was used for some markers in the Jarawa samples, since the quantity of DNA was limited. Sequencing products were purified by ethanol precipitation, suspended in formamide, and sequenced with an ABI 3700 DNA sequencer. Sequence traces were compared with a control consensus sequence by using the AutoAssembler program to identify the allelic status at the polymorphic site.

Analysis of Y Chromosome Microsatellites

Five Y chromosome-specific short tandem repeat (STR) loci were analyzed by multiplex PCR amplification and fragment analysis. The forward primers were labeled at the 5′ end with a fluorescent dye. The primers used were DYS389, DYS390, DYS391 [48], and DYS19 [49]. PCR products were analyzed on an ABI 377 instrument with 5% Long Ranger gels (FMC). Allele sizes were obtained by using the GeneScan and Genotyper analysis applications.

Acknowledgments

We are grateful to the Andaman and Nicobar people for providing DNA samples and to the Indian Council of Medical Research (ICMR), Government of India for permission to collect samples. We thank D. Mohanti for help in collecting blood, K. Price for assistance, R. D'Andrade and T. Kivisild for helpful discussions, and R. Pitchappan for sharing unpublished Indian data. This study was supported by grants from the Indian Government Council of Scientific and Industrial Research (CSIR) to L.S., by grants from the ICMR to V.R.R., L.S., and S.C.S., by grants from the Natural Environment Research Council, the Kon-Tiki Museum, and Otago University to E.H., and National Institutes of Health grants GM28428 and GM 55273 to L.L. Cavalli-Sforza. The Andaman hair samples were obtained from the Duckworth Collection, University of Cambridge. The analysis of hair DNA was carried out by C. Lalueza-Fox.

Received: July 1, 2002

Revised: October 30, 2002

Accepted: October 30, 2002

Published online: November 26, 2002

Published online November 26, 2002
DOI: 10.1016/S0960982202013362

References

1. Huxley, T.H. (1870). On the geographical distribution of the chief modifications of Mankind. J. Ethnolog. Soc. Lond. *2*, 404–412.

2. Howells, W.W. (1973). The Pacific Islanders (Sydney: Reed).

3. Venkateswar, S. (1999). The Andaman Islanders. Sci. Am. *280*, 72–78.

4. Howells, W.W. (1973). Cranial Variation in Man: A Study by Multivariate Analysis of Patterns of Difference among Recent Human Populations (Cambridge, MA: Peabody Museum of Archaeology and Ethnology, Harvard University).

5. Lehmann, H., and Ikin, E.W. (1954). Study of Andamanese Negritos. Trans. R. Soc. Trop. Med. Hyg. *48,* 12–15.

6. Omoto, K. (1985). The Negritos: genetic history and microevolution. In Out of Asia: Peopling the Americas and the Pacific, R. Kirk, and E. Szathmary, eds. (Canberra, Australia: The Journal of Pacific History), pp. 123–131.

7. Bellwood, P. (1978). Man's Conquest of the Pacific (Auckland: William Collins).

8. Cavalli-Sforza, L.L. (1986). African pygmies: an evaluation of the state of research. In African Pygmies, L.L. Cavalli-Sforza, ed. (Orlando: Academic Press), pp. 361–426.

9. Bellwood, P. (1997). Prehistory of the Indo-Malaysian Archipelago, Revised Edition (Honolulu: University of Hawai'i Press).

10. Prasad, B.V., Ricker, C.E., Watkins, W.S., Dixon, M.E., Rao, B.B., Naidu, J.M., Jorde, L.B., and Bamshad, M. (2001). Mitochondrial DNA variation in Nicobarese Islanders. Hum. Biol. *73*, 715–725.

11. Wrischnik, L.A., Higuchi, R.G., Stoneking, M., Erlich, H.A., Arnheim, N., and Wilson, A.C. (1987). Length mutations in human mitochondrial DNA: direct sequencing of enzymatically amplified DNA. Nucleic Acids Res. *15*, 529–542.

12. Quintana-Murci, L., Semino, O., Bandelt, H.-J., Passarino, G., McElreavey, K., and Santachiara-Benerecetti, A.S. (1999). Genetic evidence of an early exit of *Homo sapiens sapiens* from Africa through eastern Africa. Nat. Genet. *23*, 437–441.

13. Maca-Meyer, N., Gonzalez, A.M., Larruga, J.M., Flores, C., and Cabrera, V.M. (2001). Major genomic mitochondrial lineages delineate early human expansions. BMC Genet. *2*, 13.

14. Yao, Y.G., Kong, Q.P., Bandelt, H.J., Kivisild, T., and Zhang, Y.P. (2002). Phylogeographic differentiation of mitochondrial DNA in Han Chinese. Am. J. Hum. Genet. *70*, 635–651.

15. Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., et al. (1981). Sequence and organization of the human mitochondrial genome. Nature *290*, 457–465.

16. Kivisild, T., Bamshad, M.J., Kaldma, K., Metspalu, M., Metspalu, E., Reidla, M., Laos, S., Parik, J., Watkins, W.S., Dixon, M.E., et al. (1999). Deep common ancestry of Indian and western-Eurasian mitochondrial DNA lineages. Curr. Biol. *9*, 1331–1334.

17. Radcliffe-Brown, A.R. (1922). The Andaman Islanders (Cambridge: Cambridge University Press).

18. Underhill, P.A., Shen, P., Lin, A.A., Jin, L., Passarino, G., Yang, W.H., Kauffman, E., Bonne-Tamir, B., Bertranpetit, J., Francalacci, P., et al. (2000). Y chromosome sequence variation and the history of human populations. Nat. Genet. *26*, 358–361.

19. The Y Chromosome Consortium. (2002). A nomenclature system for the tree of human Y-chromosomal binary haplogroups. Genome Res. *12*, 339–348.

20. Underhill, P.A., Passarino, G., Lin, A.A., Shen, P., Mirazon Lahr, M., Foley, R.A., Oefner, P.J., and Cavalli-Sforza, L.L. (2001). The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. Ann. Hum. Genet. *65*, 43–62.

21. Su, B., Xiao, J., Underhill, P., Deka, R., Zhang, W., Akey, J., Huang, W., Shen, D., Lu, D., Luo, J., et al. (1999). Y-chromosome evidence for a northward migration of modern humans into Eastern Asia during the last Ice Age. Am. J. Hum. Genet. *65*, 1718–1724.

22. Su, B., Xiao, C., Deka, R., Seielstad, M.T., Kangwanpong, D., Xiao, J., Lu, D., Underhill, P., Cavalli-Sforza, L., Chakraborty, R., et al. (2000). Y chromosome haplotypes reveal prehistorical migrations to the Himalayas. Hum. Genet. *107*, 582–590.

Immediate Early Publication **Published online November 26, 2002**

DOI: 10.1016/S0960982202013362

23. Capelli, C., Wilson, J.F., Richards, M., Stumpf, M.P., Gratrix, F., Oppenheimer, S., Underhill, P., Pascali, V.L., Ko, T.M., and Goldstein, D.B. (2001). A predominantly indigenous paternal heritage for the Austronesian-speaking peoples of insular Southeast Asia and Oceania. Am. J. Hum. Genet. *68*, 432–443.

24. Hagelberg, E., Goldman, N., Liò, P., Whelan, S., Schiefenhövel, W., Clegg, J.B., and Bowden, D.K. (1999). Evidence for mitochondrial DNA recombination in a human population of island Melanesia. Proc. R. Soc. Lond. B Biol. Sci. *266*, 485–492.

25. Ruhlen, M. (1991). A Guide to the World's Languages (London: Edward Arnold).

26. Ke, Y., Su, B., Song, X., Lu, D., Chen, L., Li, H., Qi, C., Marzuki, S., Deka, R., Underhill, P., et al. (2001). African origin of modern humans in East Asia: a tale of 12,000 Y chromosomes. Science *292*, 1151–1153.

27. Karafet, T.M., Zegura, S.L., Posukh, O., Osipova, L., Bergen, A., Long, J., Goldman, D., Klitz, W., Harihara, S., de Knijff, P., et al. (1999). Ancestral Asian source(s) of new world Y-chromosome founder haplotypes. Am. J. Hum. Genet. *64*, 817–831.

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. Underhill, P.A., and Roseman, C.C. (2001). The case for an African rather than an Asian origin of the human Y-chromosome YAP insertion. In Genetic, Linguistic and Archaeological Perspectives on Human Diversity in Southeast Asia, L. Jin, M. Seielstad, and C. Xiao, eds. (River Edge, New Jersey: World Scientific Publishing), pp. 43–56.

30. Hammer, M.F. (1994). A recent insertion of an Alu element on the Y chromosome is a useful marker for human population studies. Mol. Biol. Evol. *11*, 749–761.

31. Hammer, M.F., Karafet, T., Rasanayagam, A., Wood, E.T., Altheide, T.K., Jenkins, T., Griffiths, R.C., Templeton, A.R., and Zegura, S.L. (1998). Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. Mol. Biol. Evol. *15*, 427–441.

32. Qian, Y., Qian, B., Su, B., Yu, J., Ke, Y., Chu, Z., Shi, L., Lu, D., Chu, J., and Jin, L. (2000). Multiple origins of Tibetan Y chromosomes. Hum. Genet. *106*, 453–454.

33. Karafet, T., Xu, L., Du, R., Wang, W., Feng, S., Wells, R.S., Redd, A.J., Zegura, S.L., and Hammer, M.F. (2001). Paternal population history of East Asia: sources, patterns, and microevolutionary processes. Am. J. Hum. Genet. *69*, 615–628.

34. Semino, O., Passarino, G., Oefner, P.J., Lin, A.A., Arbuzova, S., Beckman, L.E., De Benedictis, G., Francalacci, P., Kouvatsi, A., Limborska, S., et al. (2000). The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. Science *290*, 1155–1159.

35. Cooper, Z. (1993). The origins of the Andaman Islanders: local myth and archaeological evidence. Antiquity *67*, 394–399.

36. Hammer, M.F., and Horai, S. (1995). Y chromosomal DNA variation and the peopling of Japan. Am. J. Hum. Genet. *56*, 951–962.

37. Redd, A.J., Roberts-Thomson, J., Karafet, T., Bamshad, M., Jorde, L.B., Naidu, J.M., Walsh, B., and Hammer, M.F. (2002). Gene flow from the Indian subcontinent to Australia: evidence from the Y chromosome. Curr. Biol. *12*, 673–677.

38. Kayser, M., Brauer, S., Weiss, G., Schiefenhovel, W., Underhill, P.A., and Stoneking, M. (2001). Independent histories of human Y chromosomes from Melanesia and Australia. Am. J. Hum. Genet. *68*, 173–190.

39. Burenhult, N. (1996). Deep linguistic prehistory with particular reference to Andamanese. Lund University Department of Linguistics. Working Papers *45*, 5–24.

40. Hagelberg, E., and Clegg, J.B. (1991). Isolation and characterization of DNA from archaeological bone. Proc. R. Soc. Lond. B Biol. Sci. *244*, 45–50.

41. Higuchi, R., von Beroldingen, C.H., Sensabaugh, G.F., and Erlich, H.A. (1988). DNA typing from single hairs. Nature *332*, 543–546.

42. Boom, R., Sol, C.J.A., Salimans, M.M.M., Jansen, C.L., Wertheim-van Dillen, P.M.E., and van der Noordaa, J. (1990). Rapid and simple method for purification of nucleic acids. J. Clin. Microbiol. *28*, 495–503.

Immediate Early Publication **Published online November 26, 2002**

DOI: 10.1016/S0960982202013362

43. Thangaraj, K., Ramana, G.V., and Singh, L. (1999). Y-chromosome and mitochondrial DNA polymorphisms in Indian populations. Electrophoresis *20*, 1743–1747.

44. Rieder, M.J., Taylor, S.L., Tobe, V.O., and Nickerson, D.A. (1998). Automating the identification of DNA variations using quality-based fluorescence re-sequencing: analysis of the human mitochondrial genome. Nucleic Acids Res. *26*, 967–973.

45. Schneider, S., Roessli, D., and Excoffier, L. (2000). Arlequin: A Software for Population Genetics Data Analysis (Geneva: University of Geneva, Genetics and Biometry Laboratory).

46. Kwok, C., Tyler-Smith, C., Mendonca, B.B., Hughes, I., Berkovitz, G.D., Goodfellow, P.N., and Hawkins, J.R. (1996). Mutation analysis of the 2 kb 5′ to SRY in XY females and XY intersex subjects. J. Med. Genet. *33*, 465–468.

47. Hurles, M.E., Veitia, R., Arroyo, E., Armenteros, M., Bertranpetit, J., Perez-Lezaun, A., Bosch, E., Shlumukova, M., Cambon-Thomsen, A., McElreavey, K., et al. (1999). Recent male-mediated gene flow over a linguistic barrier in Iberia, suggested by analysis of a Y-chromosomal DNA polymorphism. Am. J. Hum. Genet. *65*, 1437–1448.

48. Kayser, M., Caglia, A., Corach, D., Fretwell, N., Gehrig, C., Graziosi, G., Heidorn, F., Herrmann, S., Herzog, B., Hidding, M., et al. (1997). Evaluation of Y-chromosomal STRs: a multicenter study. Int. J. Legal Med. *110*, 125–133.

49. Roewer, L., and Epplen, J.T. (1992). Rapid and sensitive typing of forensic stains by PCR amplification of polymorphic simple repeat sequences in case work. Forensic Sci. Int. *53*, 163–171.

Accession Numbers

Please see the final printed manuscript for the accession numbers for the mtDNA sequences.

Immediate Early Publication **Published online November 26, 2002**

DOI: 10.1016/S0960982202013362

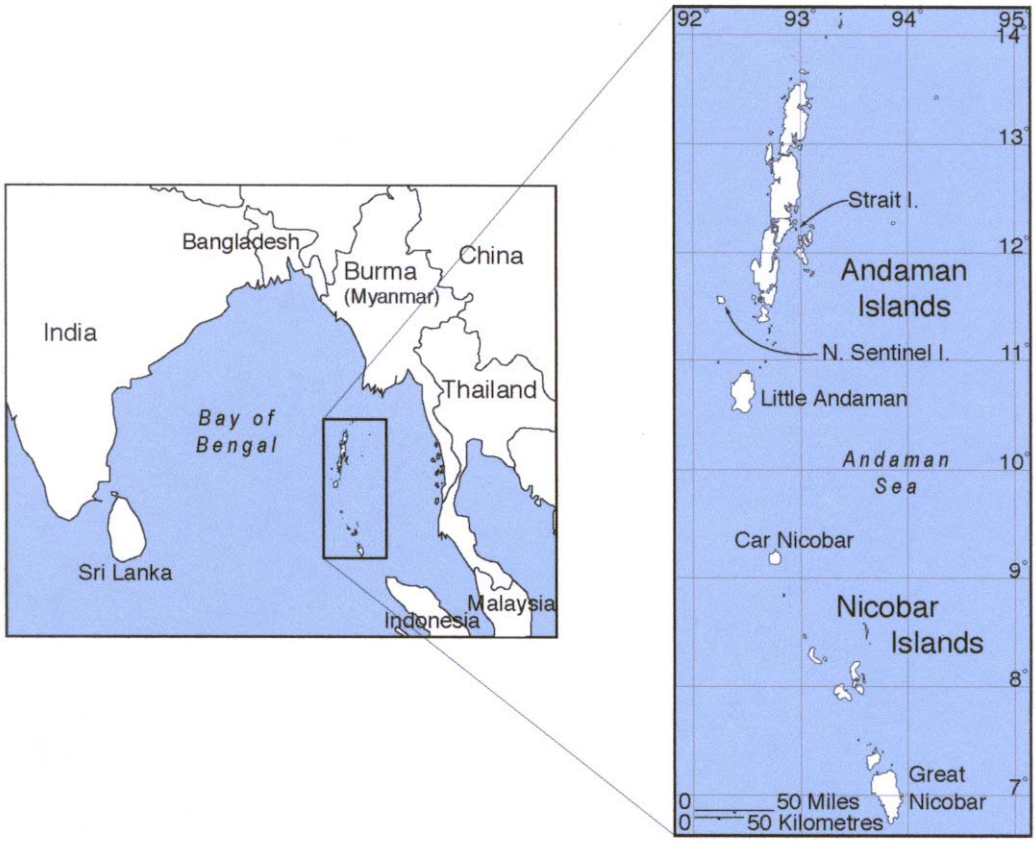
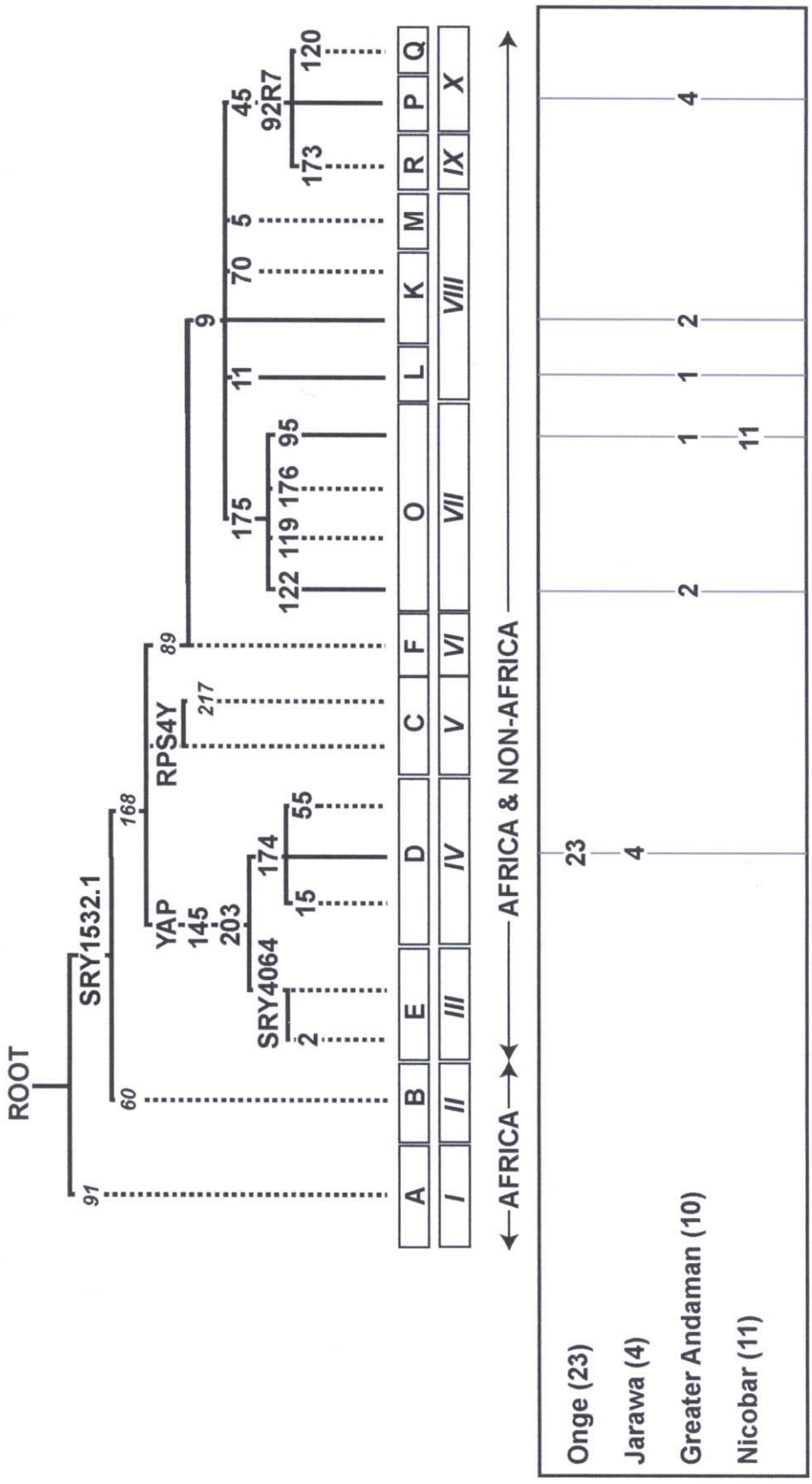


Figure 1. Map of Southeast Asia

This map shows the location of the Andaman and Nicobar Islands, Republic of India.

Published online November 26, 2002
DOI: 10.1016/S0960982202013362



Published online November 26, 2002
DOI: 10.1016/S0960982202013362

Figure 2. Maximum Parsimony Phylogenetic Tree of Worldwide Human Y Chromosome Haplogroups

The tree was constructed by using 28 binary markers that define 21 African and non-African haplogroups [18]. The bold vertical lines indicate the distribution of the six binary haplotypes observed in this study. The dashed lines indicate the remaining unobserved haplotypes. The 23 markers typed are indicated in large font. The remaining markers, shown in smaller font, provide phylogenetic context. The sample size and haplotypes observed are indicated.

Immediate Early Publication **Published online November 26, 2002**

DOI: 10.1016/S0960982202013362

Table 1. Mitochondrial DNA Substitutions in Andamanese and Nicobarese People

	9 bp del	78	93	108	129	140	162	169	172	179	183	189	213	223	234	239	242	248	249	261	265	266	270	287	288	294	296	301	304	311	319	342	344	355	357	362	364	373	10398	10400	Andamanese 1908[1] n=42	Onge n=63	Jarawa[2] n=4	Great Andamanese n=20	Nicobarese n=13
CRS	–	A	T	C	G	T	A	C	T	C	A	T	G	C	C	C	C	C	T	C	A	C	C	C	T	C	C	C	T	T	G	T	C	C	T	T	C	G	A	C					
And.hap1	–	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	G	T	2	24	2	15	
And.hap2	–	.	.	.	.	.	.	.	.	.	.	.	A	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	T	.	C	.	.	.	G	T		17			
And.hap3	–	.	.	.	.	.	.	.	.	.	.	.	A	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	C/T	.	C	.	.	.	G	T		1			
And.hap4	–	.	.	.	.	.	.	.	.	.	.	.	A	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	A	.	T	.	C	.	.	.	G	T		2			
And.hap5	–	.	.	.	.	.	.	.	.	.	.	.	A	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	T	.	C	.	.	.	G	T		3			
And.hap6	–	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	T	.	C	.	.	.	G	T	3	16	2	1	
And.hap7	–	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	C	.	.	.	G	T	8			1	
And.hap8	–	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	T	.	.	.	?	?	G	T				1	
And.hap9	–	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	C	.	C	.	.	.	C	.	.	.	C	C	.	.	.	.	.	.	.	.	.	.				1	
And.hap10	–	C	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	C	.	C	.	.	.	C	.	.	.	C	C	.	.	.	.	.	.	.	?	.	.				1	
And.hap11	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	?	?	1				
And.hap12	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	C	.	.	.	.	.	.	.	.	?	?	1				
And.hap13	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	?	?	1				
And.hap14	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	A	.	.	.	.	.	.	?	?	4				
And.hap15	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	C	.	.	.	?	?	1				
And.hap16	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	.	A	T	.	.	.	.	.	C	.	.	.	.	.	.	.	.	?	?	1				
And.hap17	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	C	.	.	.	?	?	1				
And.hap18	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	C	.	A	.	.	.	.	.	.	?	?	1				
And.hap19	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	A	.	.	C	.	.	.	?	?	2				
And.hap20	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	.	T	.	.	T	.	.	.	C	A	.	.	.	.	.	.	.	?	?	1				
And.hap21	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	A	.	.	.	C	.	.	.	?	?	1				
And.hap22	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	A	.	T	.	C	.	.	.	?	?	1				
And.hap23	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	A	.	.	.	C	.	.	A	?	?	1				
And.hap24	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	.	.	C	.	.	.	?	?	1				
And.hap25	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	A	.	T	.	.	.	.	.	?	?	1				
And.hap26	–	?	?	?	?	?	?	?	?	?	?	?	.	T	T	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	?	?	1				
And.hap27	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	?	?	1				
And.hap28	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	C	.	.	A	?	?	1				
And.hap29	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	T	.	.	.	.	.	.	.	.	T	T	.	.	.	.	.	.	T	.	.	.	.	?	?	1				
And.hap30	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	?	?	2				
And.hap31	–	?	?	?	?	?	?	?	?	?	?	?	.	T	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	1				
And.hap32	–	?	?	?	?	?	?	?	?	?	?	?	.	.	.	.	.	.	.	.	.	.	.	.	.	T	T	.	C	.	.	.	.	.	.	.	.	.	?	?	1				
And.hap33	–	?	?	?	?	?	?	?	?	?	?	?	.	.	.	.	.	.	.	.	.	.	.	.	.	T	T	.	C	.	.	.	.	.	.	.	T	.	?	?	1				
And.hap34	–	?	?	?	?	?	?	?	?	?	?	?	.	.	.	.	.	T	.	.	.	.	A	.	.	T	T	.	.	C	.	.	.	.	.	.	.	.	?	?	1				
Nic.hap1	–	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	C	.	.	.	C	.	.	.	.	.	.	.	.	.	G	.					1
Nic.hap2	–	.	.	T	A	.	G	.	C	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	C	.	.	.	.	.	.	.	.	.	.					1
Nic.hap3	–	.	.	T	A	.	G	.	C	.	.	C	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	C	.	.	.	.					1
Nic.hap4	–	.	.	T	A	.	G	.	C	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.					1
Nic.hap5	–	.	.	.	A	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.					4
Nic.hap6	+	.	.	.	.	C	.	.	.	.	C	C	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.					4
Nic.hap7	–	C	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	C	.	.	G	T					1
Haplotypes per population																																									27	6	2	6	7

The positions are compared to the human mtDNA reference sequence [15], minus 16000. A dot (.) denotes identity with the reference sequence. The allelic status at the COII/tRNA^{lys} region (9-bp), and at positions 10398 and 10400, is shown. The Andaman hair sequences spanned only from 16211 to 16400, but matches with living Andamanese over this span are indicated. "CRS," Cambridge Reference Sequence.

[1]Not assayed for the allelic status at sites 10398 and 10400.

[2]Not assayed for the allelic status at sites 10398 and 10400, or the 9-bp deletion.

Immediate Early Publication **Published online November 26, 2002**

DOI: 10.1016/S0960982202013362

Table 2. Y Chromosome Microsatellite Alleles in Present-Day Andamanese and Nicobarese

	Defining Mutation	n	DYS389-I	DYS389-II	DYS19	DYS390	DYS391	Mean Variance
Onge	M174	7	11	18	14	22	10	
	M174	3	12	18	14	22	10	
	M174	5	11	17	14	22	10	
	M174	3	11	18	14	22	10	
	M174	1	11	20	14	22	10	
			0.14	0.47	0	0	0	0.12
Great Andamanese	M45	4	10	18	14	23	10	
	M9	2	10	17	14	22	10	
	M11	1	9	17	15	22	11	
	M122	1	10	17	14	23	10	
	M122	1	9	17	15	23	10	
	M95	1	10	17	14	22	10	
			0.18	0.27	0.18	0.27	0.1	0.20
Nicobarese	M95	2	11	15	16	25	11	
	M95	2	11	16	15	25	11	
	M95	1	11	16	16	25	11	
	M95	1	10	15	14	25	11	
	M95	1	12	15	16	25	11	
	M95	1	10	16	15	25	10	
			0.41	0.29	0.55	0	0.13	0.28

Y chromosome microsatellite alleles at five loci in present-day Andamanese and Nicobarese, with the defining binary mutation. The alleles are given as the number of tandem repeats, and the mean variance across the five loci is shown for each population sample.

Immediate Early Publication **Published online November 26, 2002**

DOI: 10.1016/S0960982202013362

Table 3. Mitochondrial DNA Molecular Diversity of the Andamanese and Nicobarese Population Samples

	n	Number of Haplotypes	Haplotype Diversity, h	SD	π^a	SD	Nucleotide Diversity	SD
Andamanese 1908	42	27	0.95	0.02	4.19	2.13	0.022	0.012
Great Andamanese	20	6	0.45	0.14	2.15	1.25	0.007	0.005
Onge	63	5	0.72	0.03	2.49	1.36	0.008	0.005
Jarawa	4	2	0.67	0.20	2.67	1.78	0.008	0.007
Nicobarese	13	7	0.85	0.08	6.18	3.14	0.020	0.011

Mitochondrial DNA molecular diversity data of the Andamanese and Nicobarese. The calculations were based on the segment between positions 16063 and 16376 in the present-day populations and positions 16211 and 16400 in the hair DNA sequences.

[a]π is the mean number of pairwise differences between sequences.

Immediate Early Publication **Published online November 26, 2002**

DOI: 10.1016/S0960982202013362

Supplementary Material

Supplementary Experimental Procedures

Extraction of DNA from Archival Hair Samples

Hair DNA was extracted with strict precautions to avoid contamination by modern DNA [S1] by using a modification of a previously described method [S2, S3]. Strands were washed with sterile water and 95% ethanol, dried, and extracted overnight at 41°C in 250 μl 10 mM Tris-HCl (pH 8), 10 mM EDTA (pH 8), 0.1 M NaCl, 50 μg/ml proteinase K, 39 mM DTT, and 2% Triton X-100. After the addition of 50 μl of a buffer consisting of 10 M guanidinium thiocyanate (GuSCN), 0.1 M Tris-HCl (pH 7.4), and 0.2 M EDTA, the lysate was incubated for 25 min at 60°C. After a further centrifugation step, 200 μl of the supernatant was mixed with 1 ml 8.2 M NaI and 25 μl silica suspension and was incubated at room temperature for 6 min. After centrifugation, the silica pellet was washed three times with 70% ethanol. DNA was eluted with 15 μl TE buffer. Control extractions with no hair were carried out in parallel to check for contamination. Four hair samples were extracted in duplicate to check reproducibility.

Analysis of mtDNA Loci

mtDNA sequencing of the modern DNA extracts was carried out as described previously [S4], with the following modifications: control region amplicons for 20 Onge were purified with exonuclease I and shrimp alkaline phosphatase. Sequencing was done by cycle sequencing dye terminator reactions with an ABI 377 automated sequencer. Reducing the concentration of amplification primers to 5 pmol, and the concentration of dNTPs to 50 μM, eliminated the need for post-PCR template purification of the remaining samples, which were sequenced on an ABI 3700 sequencer.

Hair DNA was amplified with primers (shown in a 5′ to 3′ direction) L16210: TCCCCATGCTTACAAGCAAGTA and H16405: CGGGATATTGATTTCACGGAGGAT, which specify a 239-bp fragment of the mtDNA control region. Aliquots of 1 μl of hair DNA extract were amplified in 25 μl reaction volumes containing 1 U of *Taq* start antibody and *Taq* DNA polymerase (1:2) (HT Biotechnology), 10× reaction buffer (100 mM Tris-HCl [pH 9], 15 mM $MgCl_2$, 500 mM KCl, 1% [w/v] gelatin), 2.5 mM $MgCl_2$, 0.2 mM dNTPs, and 20 pmoles of each primer. Amplification was for performed for 40 cycles at 92°C for 40 s, 55°C for 40 s, and 72°C for 1 min, with an initial denaturation step of 1.5 min at 94°C and a final step of 5 min at 75°C. In some cases, a second PCR was carried out with primers L16210 and H16401. A subset of ten extracts was amplified with the internal primer H16314: GTGCTATGTACGGTAAATGGCTTTAT, which yielded a 148-bp fragment. PCR products were electrophoresed on 2% low-melting agarose gels and were visualized under UV light after ethidium bromide staining. PCR products were excised from the gels, silica purified, and used as templates for direct sequencing on an Applied Biosystems ABI 377A DNA sequencer.

Analysis of Y Chromosome Binary Polymorphisms

The following 23 biallelic markers were analyzed: YAP, RPS4Y, SRY4064, sY81 = M2, M5, M9, M11, M15, M45, M55, M70, M95, M119, M120, M122, M145, M173, M174, M175, M176, M203 [S5], 92R7 [S6], and SRY1532 [S7]. Nested polymerase chain reaction (PCR) was used for some markers in the Jarawa samples, since the quantity of DNA was limited, as follows: for each marker tested, an initial PCR reaction was performed by using outer primers YAP-OF: TCACATAATTTCATTTTCCC, YAP-OR: CAAGTTAGCTGTCCATACTG, M130-OF: GAGTGGGAGGGACTGTGAGA, M130-OR: CCACAGAGATGGTGTGGGTA, M9-OF: CGGCGTCTTTGATCTCTCAATCC, M9-OR: CAGCTTCTTTCCCAATTATGCAAG, M173-OF: AAGAAATGTTGAACTGAAAGTTGAT, M173-OR: AGGTGTATCTGGCATCCGTTA, M175-OF: TTGAGCAAGAAAAATAGTACCCA, and M175-OR: CTCCATTCTTAACTATCTCAGGGA. Amplified material from the first PCR reactions was diluted 200-fold and reamplified with the published primers in the second,

Immediate Early Publication

Published online November 26, 2002
DOI: 10.1016/S0960982202013362

nested PCR reaction. For the Onge, Great Andamanese, and Nicobarese samples, a single three-temperature PCR was performed, with 50 ng of genomic DNA and marker-specific annealing temperatures and times. After an initial cycle of 94°C for 5 min, 30 PCR cycles were carried out at 94°C for 1 min and 72°C for 1 min, followed by a final cycle of 72°C for 5 min. Specific temperature/time annealing conditions were as follows: YAP-Outer: 52°C/20 s, SRY1532: 62°C/20 s, 92R7: 62°C/10 s, sY81: 60°C/20 s, M70: 50°C/30 s. The same annealing conditions of 55°C/30 s were used for YAP-I, M5, M9, M11, M15, M45, M70, M95, M119, M120, M122, M145, RPS4Y, M173, M174, M175, M176, and M203. Appropriate precautions were followed to control against amplicon contamination. Suitable allelic controls were also evaluated. Products from YAP amplification were fractionated on 1.5% agarose gels and visualized under UV light after staining with ethidium bromide. The PCR products of 92R7 were digested with HindIII and were separated on 12% acrylamide gels, followed by ethidium bromide staining. All other PCR amplicons were checked on 1.5% agarose gels prior to sequencing after purification with exonuclease I and shrimp alkaline phosphatase. Sequencing reactions for both forward and reverse strands were performed for each marker and used approximately 75 ng PCR product, 8 μl Big Dye Terminator (Perkin-Elmer cycle sequencing kit), and 4 pmol of the PCR primer as the sequencing primer. Sequencing products were purified by ethanol precipitation, suspended in formamide, and sequenced with an ABI 3700 DNA sequencer. Sequence traces were compared with a control sequence by using the AutoAssembler program to identify the allelic status at the polymorphic site.

Analysis of Y Chromosome Microsatellites

Five Y chromosome-specific short tandem repeat (STR) loci were analyzed by PCR amplification and fragment analysis. The forward primers were labeled at the 5′ end with a fluorescent dye. The primers used were DYS389, DYS390 and DYS391 [S8], and DYS19 [S9]. DYS389 yields two polymorphic loci. As the DYS389-11 amplicon contains the DYS389-1 fragment, the latter was subtracted [S10]. Multiplex reactions were performed with 5 ng DNA, 10 pM each primer, 200 mM dNTPs, 1× buffer with 1.5 mM $MgCl_2$, and 1 U AmpliTaq Gold. Amplifications were carried out in a Perkin-Elmer GeneAmp 9600 thermal cycler, as follows: denaturation at 95°C for 10 min, followed by 30 cycles at 94°C for 1 min, 55°C for 30 s, and 72°C for 2 min. PCR products were denatured with GS-ROX500 in the presence of formamide:blue dextrone (5:1) and were analyzed on an ABI 377 instrument by using 5% Long Ranger gels (FMC). Allele sizes were obtained by using the GeneScan and Genotyper analysis applications.

Immediate Early Publication **Published online November 26, 2002**

DOI: 10.1016/S0960982202013362

Supplementary References

S1. Hagelberg, E., and Clegg, J.B. (1991). Isolation and characterization of DNA from archaeological bone. Proc. R. Soc. Lond. B Biol. Sci. *244*, 45–50.

S2. Higuchi, R., von Beroldingen, C.H., Sensabaugh, G.F., and Erlich, H.A. (1988). DNA typing from single hairs. Nature *332*, 543–546.

S3. Boom, R., Sol, C.J.A., Salimans, M.M.M., Jansen, C.L., Wertheim-van Dillen, P.M.E., and van der Noordaa, J. (1990). Rapid and simple method for purification of nucleic acids. J. Clin. Microbiol. *28*, 495–503.

S4. Thangaraj, K., Ramana, G.V., and Singh, L. (1999). Y-chromosome and mitochondrial DNA polymorphisms in Indian populations. Electrophoresis *20*, 1743–1747.

S5. Underhill, P.A., Passarino, G., Lin, A.A., Shen, P., Mirazon Lahr, M., Foley, R.A., Oefner, P.J., and Cavalli-Sforza, L.L. (2001). The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. Ann. Hum. Genet. *65*, 43–62.

S6. Kwok, C., Tyler-Smith, C., Mendonca, B.B., Hughes, I., Berkovitz, G.D., Goodfellow, P.N., and Hawkins, J.R. (1996). Mutation analysis of the 2 kb 5′ to SRY in XY females and XY intersex subjects. J. Med. Genet. *33*, 465–468.

S7. Hurles, M.E., Veitia, R., Arroyo, E., Armenteros, M., Bertranpetit, J., Perez-Lezaun, A., Bosch, E., Shlumukova, M., Cambon-Thomsen, A., McElreavey, K., et al. (1999). Recent male-mediated gene flow over a linguistic barrier in Iberia, suggested by analysis of a Y-chromosomal DNA polymorphism. Am. J. Hum. Genet. *65*, 1437–1448.

S8. Kayser, M., Caglia, A., Corach, D., Fretwell, N., Gehrig, C., Graziosi, G., Heidorn, F., Herrmann, S., Herzog, B., Hidding, M., et al. (1997). Evaluation of Y-chromosomal STRs: a multicenter study. Int. J. Legal Med. *110*, 125–133.

S9. Roewer, L., and Epplen, J.T. (1992). Rapid and sensitive typing of forensic stains by PCR amplification of polymorphic simple repeat sequences in case work. Forensic Sci. Int. *53*, 163–171.

S10. Cooper, G., Amos, W., Hoffman, D., and Rubinsztein, D.C. (1996). Network analysis of human Y microsatellite haplotypes. Hum. Mol. Genet. *5*, 1759–1766.

Immediate Early Publication **Published online November 26, 2002**

DOI: 10.1016/S0960982202013362