

## Reconstructing the Origin of Andaman Islanders

Kumarasamy Thangaraj,<sup>1</sup> Gyaneshwer Chaubey,<sup>1</sup>  
 Toomas Kivisild,<sup>2</sup> Alla G. Reddy,<sup>1</sup> Vijay Kumar Singh,<sup>1</sup>  
 Avinash A. Rasalkar,<sup>1</sup> Lalji Singh<sup>1\*</sup>

The Andaman and Nicobar Islands are inhabited by six indigenous tribal populations. Our previous work (1) suggested that Andamanese “Negritos” have closer affinities with Asian than with African populations and that the Nicobarese have close genetic affinities to Southeast Asians. In a parallel study by Endicott *et al.* (2), which also suggested Asian affinities for the mitochondrial DNA (mtDNA) lineages of the Andamanese, analyses of museum specimens related the two major lineage groups on the islands with haplogroups M2 and M4, which are found commonly throughout India (3–5). Analysis of a hypervariable sequence (HVS-I) motif indicated a link with Papuan populations (1). These links, however, relied on nucleotide positions in the control region of mtDNA that are known to be hypervariable. To clarify the phylogenetic position of the mtDNA lineages of the Andaman Islanders, we analyzed the complete mtDNA sequence of five Onge, five Great Andamanese, and five Nicobarese individuals.

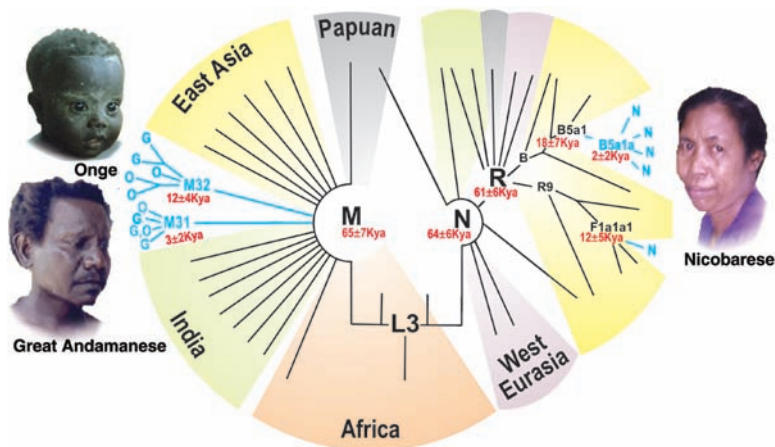
Analysis of the Onge and Great Andamanese complete mtDNA sequences revealed the existence of two previously uncharacterized clades, which we named M31 and M32 (Fig. 1). Analysis of the complete mtDNA sequences shows that none of the coding region mutations defining these two haplogroups overlap with the known Indian or East Asian mtDNA haplogroups (1–5). In our survey of ~6500 mtDNA sequences from mainland India, none of the M lineages carried the coding region mutations specific to M31 and M32 (6). Furthermore, none of the haplogroup M complete sequences reported so far share any of the mutations that define M31 and M32, suggesting that these two haplogroups are likely to have evolved in situ on these islands.

Our previous studies have shown that all Eurasian and Oceanian founder haplogroups—mitochondrial M, N, and R and Y-chromosomal C, D, and F—coexist in South Asia, suggesting their comigration along the southern coastal

route in one wave after the exit of modern humans from Africa (1, 3–5). Because the Andaman Negrito populations carry only one mitochondrial founder haplogroup (M) and only one Y-chromosomal haplogroup (D), it is tempting to relate this phylogeographic pattern with a one-haplogroup–one-migration scenario. However, the shallow phylogenetic time depths of M31 and M32 in Onge and Great Andamanese populations (Fig. 1) more likely reflect the effect of genetic drift in these extremely small populations. Therefore, it is not surprising that the founding mtDNA

Nicobarese F sequence clustered it into haplogroup F1a1a1. This subclade of haplogroup F has been observed in China (8), Malaysia, and Thailand (9), confirming a Southeast Asian origin of this lineage in Nicobarese. Relatively older coalescence times for the most recent common ancestors of Nicobarese and Southeast Asian B5a1 and F1a1a1 lineages (Fig. 1) are based on only a few samples from Southeast Asia. Therefore, these coalescence times reflect the time of the origin of these lineages in mainland Southeast Asia rather than the time of their migration to the Nicobar Islands.

Our data indicate that two ancient maternal lineages, M31 and M32 in the Onge and the Great Andamanese, have evolved in the Andaman Islands independently from other South and Southeast Asian populations. These lineages have likely been isolated since the initial penetration of the northern coastal areas of the Indian Ocean by anatomically modern humans, in their out-of-Africa migration ~50 to 70 thousand years ago. In contrast, the Nicobarese show a close genetic relation with populations in Southeast Asia, suggesting their recent arrival from the east during the past 18 thousand years.



**Fig. 1.** A tree showing the phylogenetic position of the complete mtDNA sequences of Andaman and Nicobar islanders. O, G, and N represent Onges, Great Andamanese, and Nicobarese lineages, respectively; kya, thousands of years ago. Coalescent times of haplogroups M, N, and R were obtained from Mishmar *et al.* (10); a mutation rate of  $1.26 \times 10^{-8}$  was used for other haplogroups estimated from the present data (7, 10).

haplogroups N\* and R\* and Y-chromosomal haplogroups C and F may have completely disappeared from the extant Andaman and Nicobar islanders.

Most Nicobarese mtDNA lineages belong to either of the two common haplogroups B and F (1), which are specific to East Asia. All four Nicobarese B5a sequences clustered together in a branch defined by three coding region substitutions (at nucleotide positions 11881, 13145, and 13395) (fig. S1) (7). The low variation observed both in control and coding region sequences implies that these lineages among Nicobarese coalesce to their most recent common ancestor within the past couple of thousand years. The phylogenetic analysis of a

### References and Notes

1. K. Thangaraj *et al.*, *Curr. Biol.* **13**, 86 (2003).
2. P. Endicott *et al.*, *Am. J. Hum. Genet.* **72**, 178 (2003).
3. T. Kivisild *et al.*, *Am. J. Hum. Genet.* **72**, 313 (2003).
4. P. Endicott, V. Macaulay, T. Kivisild, C. Stringer, A. Cooper, *Am. J. Hum. Genet.* **72**, 1590 (2003).
5. M. Metspalu *et al.*, *BMC Genet.* **5**, 26 (2004).
6. K. Thangaraj, unpublished data.
7. Materials and methods are available as supporting material on Science Online.
8. Y. G. Yao, Q. P. Kong, H.-J. Bandelt, T. Kivisild, Y. P. Zhang, *Am. J. Hum. Genet.* **70**, 635 (2002).
9. M. Ingman, H. Kaessmann, S. Pääbo, U. Gyllensten, *Nature* **408**, 708 (2000).
10. D. Mishmar *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 171 (2003).
11. We thank C. Tyler-Smith and P. M. Bhargava for their comments.

### Supporting Online Material

www.sciencemag.org/cgi/content/full/308/5724/996/DC1  
 Materials and Methods  
 SOM Text  
 Fig. S1  
 References and Notes  
 19 January 2005; accepted 10 March 2005  
 10.1126/science.1109987

<sup>1</sup>Centre for Cellular and Molecular Biology, Hyderabad-500 007, India. <sup>2</sup>Estonian Biocenter, Riia 23, Tartu-51010, Estonia.

\*To whom correspondence should be addressed. E-mail: lalji@cmb.res.in