Am. J. Hum. Genet. 58:1303–1308, 1996

β-Globin Haplotype Analysis Suggests That a Major Source of Malagasy Ancestry Is Derived from Bantu-Speaking Negroids

Rachel Hewitt,[1] Amanda Krause,[1] Andrea Goldman,[1] Gwyn Campbell,[2] and Trefor Jenkins[1]

[1]Department of Human Genetics, School of Pathology, The South African Institute for Medical Research and University of the Witwatersrand, and [2]Department of Economic History, University of the Witwatersrand, Johannesburg

Summary

The origins of the inhabitants of Madagascar have not been fully resolved. Anthropological studies and preliminary genetic data point to two main sources of ancestry of the Malagasy, namely, Indonesian and African, with additional contributions from India and Arabia. The sickle-cell (β^S) mutation is found in populations of African and Indian origin. The frequency of the β^S-globin gene, derived from 1,425 Malagasy individuals, varies from 0 in some highland populations to .25 in some coastal populations. The β^S mutation is thought to have arisen at least five times, on the basis of the presence of five distinct β^S-associated haplotypes, each found in a separate geographic area. Twenty-five of the 35 Malagasy β^S haplotypes were of the typical "Bantu" type, 1 "Senegal" haplotype was found, and 2 rare or atypical haplotypes were observed; the remaining 7 haplotypes were consistent with the Bantu haplotype. The Bantu β^S mutation is thought to have been introduced into Madagascar by Bantu-speaking immigrants (colonists or slaves) from central or east Africa. The Senegal β^S mutation may have been introduced to the island via Portuguese naval explorers. This study provides the first definitive biological evidence that a major component of Malagasy ancestry is derived from African populations, in particular, Bantu-speaking Negroids. β^A haplotypes are also consistent with the claim for a significant African contribution to Malagasy ancestry but are also suggestive of Asian/Oceanic and Caucasoid admixture within the Malagasy population.

Introduction

Madagascar is the fourth-largest island in the world. A single term, "Malagasy," is used to refer to the island's people as well as its language. There is, however, as much cultural variation in Madagascar as there is diversity of climate, vegetation, and wildlife. The Malagasy are divided into 22 ethnic groups, distributed on the island as follows (fig. 1): east-coastal Madagascar: Betsimisaraka, Tanala, Antaimoro, Anataifasy, Antambahoaka, Antaisaka, Zafisoro, and Antanosy; southwest Madagascar: Antandroy, Bara, Mahafaly, Vezo, and Masikoro; west Madagascar: Sakalava and Makoa; north Madagascar: Antankarana, Tsimihety, and Sihanaka; and central highland Madagascar: Merina, Betsileo, Bezanozano, and Vakinankaratra. The Merina, Betsileo, Bezanozano, and Vakinankaratra (who are considered to be of Merina and Betsileo admixture) are usually referred to as the "highland" groups, and all others as the "lowland" groups.

The problem of Malagasy origins has been controversial and has still not been completely resolved. From the vast amount of anthropological, historical, and ethnographic data, and the limited genetic data, it appears that the base stock of the Malagasy people is a mixture of African and Indonesian immigrants. All Malagasy speak an essentially Indonesian language and possess a culture that contains many elements from both areas of origin. Several other likely sources of origin include the coast of Mozambique; the Comoros islands; the "Swahili Coast" of Tanzania and Kenya; the Persian Gulf; "Malay sailors" of the 13th or 14th centuries; Europeans; and additional associations with Great Zimbabwe, the Rift Valley of east Africa, and south India.

The hemoglobinopathies have reached high frequencies in certain geographical regions because of the selective advantage conferred by the carrier state against the malarial parasite *Plasmodium falciparum*. The sickle-cell mutation (β^S), the focus of this study, is found in many world populations, the highest frequencies occurring in Africans and their descendants in the Americas. The β^S gene is not found, however, in Southeast Asia (reviewed by Nagel and Fleming 1992). The β^S gene has been found in linkage disequilibrium with five major β-globin gene cluster haplotypes: four of these occur homogeneously and virtually exclusively in four segregated regions of Africa (Pagnier et al. 1984; Lapoumeroulie et al. 1992), and the fifth β^S haplotype is almost exclusively present in individuals from eastern Saudi

Received November 21, 1995; accepted for publication March 19, 1996.

Address for correspondence and reprints: Dr. Trefor Jenkins, Department of Human Genetics, South African Institute for Medical Research, P. O. Box 1038, Johannesburg, South Africa, 2000. E-mail: 058TJ@chiron.wits.ac.za

© 1996 by The American Society of Human Genetics. All rights reserved.
0002-9297/96/5806-0024$02.00

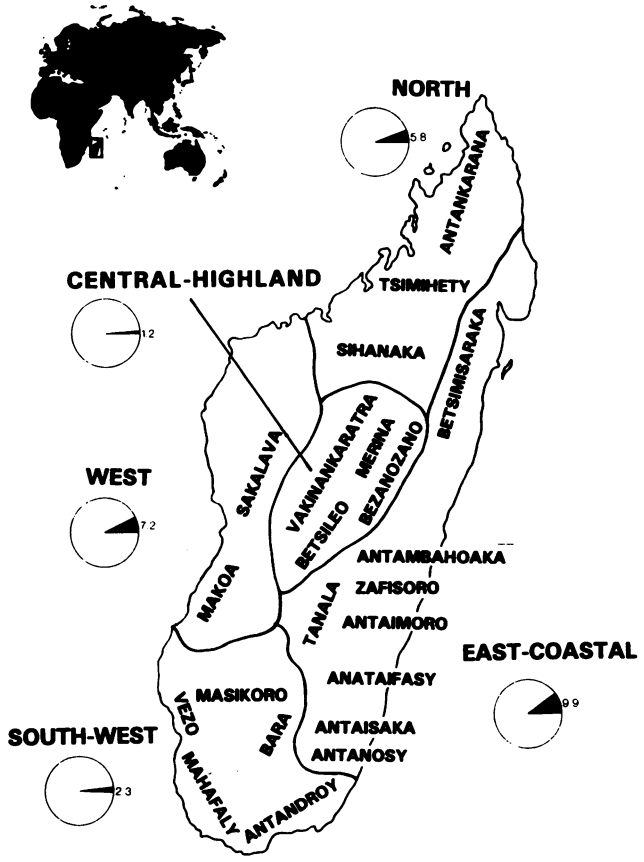

Figure 1 Map of Madagascar, showing the approximate distribution of the ethnic groups and the β^S frequencies in the different regions.

Arabia and India (Kulozik et al. 1986). Thus, knowledge of the β^S haplotype that occurs in a population may be used to gauge that population's origin(s). Because sickle-cell anemia is known to occur in Madagascar, and, because of the unresolved debate concerning the ancestry of the Malagasy, this study was undertaken to investigate the biological diversity of the Malagasy. This was done by determining the frequency and distribution of the sickle mutation in Madagascar, and constructing β^A and β^S haplotypes of Malagasy chromosomes.

Subjects and Methods

Subjects

Blood was obtained from 1,425 unrelated Malagasy individuals who had the same patrilineal and matrilineal ethnic groups and 22 families, comprising 72 individuals, during four field trips to Madagascar, undertaken during the periods May/June 1992, October 1992, June 1993, and October 1993. Appropriate informed consent was obtained from all subjects. The data on the sickle-cell status of 1,425 random individuals were used to calculate the β^S-gene frequencies of the different ethnic groups. The families were used for the construction of β^A and β^S haplotypes.

Methods

Hematological analysis.—Hematological analysis was carried out on all samples. This included full blood count, cellulose acetate electrophoresis, and citrate agar electrophoresis, to confirm the presence of HbS.

β^S *determination.*—The sickle-cell status of each individual was initially ascertained from the family members themselves and later confirmed by hematological and DNA analysis. DNA was extracted from buffy coats according to the salting-out method of Miller et al. (1988). The sickle-cell mutation was detected by PCR amplification of the 5′ region of the β-globin gene, followed by restriction-enzyme digestion with *Dde*I (Old and Ludlam 1991).

β-*globin haplotype analysis.*—Nine RFLP sites within the β-globin gene cluster were analyzed, six (*Hinc*II/ε, *Hin*dIII/^{G}γ, *Hinc*II/5′ψβ, *Hinc*II/3′ψβ, *Ava*II/β, and *Hin*fI/β) with PCR amplification followed by restriction-enzyme digestion and three (*Xmn*I/^{G}γ, *Hin*dIII/^{A}γ, and *Hpa*I/β sites) by Southern blotting. The pγ probe was used for the *Xmn*I/^{G}γ and *Hin*dIII/^{A}γ sites (Tuan et al. 1979; Gilman and Huisman 1985), and the βIVS2 probe (Burns et al. 1981) was used to detect the *Hpa*I polymorphism. Probes were labeled with ^{32}P-dCTP labeled to a specific activity of 10^6 dpm/μg by using the random hexamer priming method and extension with the Klenow fragment of DNA polymerase I (Amersham) (Feinberg and Vogelstein 1983). Southern hybridization was performed by using Hybond-N nylon membrane. All PCR reactions were carried out in either a Perkin Elmer Cetus DNA thermal cycler or a Hybaid thermal reactor.

Each individual was scored for the presence (+) or absence (−) of each of the nine RFLP sites. β-globin gene cluster haplotypes were constructed by studying segregation in families to establish the linkage phase. In all families, both β^A and β^S haplotypes could be determined.

Results

Frequency of β^S *Gene in the Malagasy*

The overall frequency of the β^S gene was found to be .0156 in the highland populations and .0744 in the lowland populations. The highland populations as a whole were in Hardy-Weinberg equilibrium ($\chi^2 = .113$; $P = .737$), but the lowland population showed disturbance of Hardy-Weinberg equilibrium ($\chi^2 = 11.71$; $P < 10^{-3}$). The β^S-gene frequency for each ethnic group, as well as for each "regional group" (i.e., ethnic groups

Table 1

β^S Frequencies of Malagasy Ethnic and Regional Groups

Region and Ethnic Group	No. of Individuals			β^S Frequency (±SE)
		Heterozygotes		
	Total	No.	%	
East-Coastal:				
Betsimisaraka	142	9	6.34	.032(.010)
Tanala	11	3	27.3	.136(.073)
Antaimoro	91	29	31.9	.159(.027)
Antaifasy	45	7	15.6	.078(.028)
Antambahoaka	33	9	27.3	.136(.042)
Antaisaka	199	44	22.1	.111(.016)
Zafisoro	14	7	50.0	.250(.082)
Antanosy	40	6	15.0	.075(.029)
Overall	575	114	19.8	.099(.009)
Southwest:				
Antandroy	41	2	4.88	.024(0)
Bara	23	2	8.70	.044(.030)
Vezo	38	3	7.89	.039(.022)
Mahafaly	41	0	0.00	0(0)
Masikoro	11	0	0.00	0(0)
Overall	154	7	4.55	.023(.009)
West:				
Sakalava	76	11	14.5	.072(.021)
Makoa	8	0	0.00	0(0)
Overall	84	11	13.1	.065(.019)
North:				
Antankarana	14	3	21.4	.107(.058)
Tsimihety	126	12	9.52	.048(.013)
Sihanaka	6	2	3.33	.167(.108)
Overall	146	17	11.6	.058(.014)
Central:				
Merina	171	2	1.17	.006(.004)
Betsileo	154	1	0.65	.003(.003)
Bezanozano	61	6	9.84	.049(.020)
Vakinankaratra	6	0	0.00	0(0)
Overall	392	9	2.30	.012(.004)

classified according to their geographic location on the island), is shown in table 1. The frequencies in the regional groups are also shown in figure 1. The highest frequency (.099) occurs among the east-coastal groups, while the lowest frequency (.012), occurs among the central groups.

β^S *Haplotypes*

Complete β^S haplotypes for 28 unrelated chromosomes from nine ethnic groups, including both highland and lowland groups, were obtained. These were classified according to the common β^S haplotypes and are shown in table 2. "Rare (1)" has previously been reported in the Central African Republic (3.4% of β^S haplotypes), in Senegal (1.8%) (Pagnier et al. 1984), in India (1.6%), and in South Arabia (9.7%) (Kulozik et al. 1986) and accounts for 1.4% of all β^S haplotypes in the world. However, "rare (2)" has not been previously reported from any geographic region (reviewed by Flint et al. 1993*a*).

β^A *Haplotypes*

At least 13 different β^A haplotypes (i.e., haplotypes comprising all of the nine polymorphic sites analyzed in the study) were found. Since there are little data available for full β^A haplotypes in other populations, the 5′ haplotypes, comprising the five polymorphic sites *Hinc*II/ε, *Hin*dIII/$^G\gamma$, *Hin*dIII/$^A\gamma$, *Hinc*II/5′ψβ, and *Hinc*II/3′ψβ, were used for comparison. Five such 5′ haplotypes occur in the Malagasy, and their frequencies, together with those in other world populations, are shown in table 3.

Discussion

β^S-*Gene Frequencies in Madagascar*

The β^S gene is present throughout Madagascar. The higher frequency of the β^S gene among the coastal populations than among the highlanders is, perhaps, not unexpected, in view of the known selective advantage conferred by the sickle-cell trait against malaria, which is hyperendemic in the coastal regions of the country. The β^S-gene frequency is highest among the groups in the southeast of the island, where it is wet and humid throughout the year and experiences a high average rainfall, almost double that at Antananarivo, situated on the central plateau (Campbell, in press). The β^S-gene frequency is lower among the populations along the western and northern coasts, which are drier and more desertlike. The lowest β^S-gene frequency occurs among the central groups inhabiting the cool, dry central plateau. The central plateau, traditionally malaria-free, has been visited by pandemics since the 1870s (Campbell 1992). Among the central group, the Bezanozano have a relatively high β^S frequency. They are arguably not highland people, since they inhabit the valleys on the eastern escarpment, but they may be descended from the Merina (Leroi-Gourhan and Poirier 1953). Their higher β^S frequency may be the result of selection.

Although similar β^S frequencies have been reported previously in the Malagasy, they were based on small samples that were not divided according to ethnic groups. Singer et al. (1957) reported 4.9% "sicklers" in a study of 1,546 Malagasy individuals. In a more extensive study of 3,684 individuals, Fourquet et al. (1974) found the average frequencies of heterozygotes to be 8.2% (average gene frequency of .04), with a much lower average frequency (4%) on the plateau compared to the coastal areas (12%).

Table 2

Malagasy β^S Haplotypes

Haplotype	RFLP Site[a]									No. of Chromosomes
	1	2	3	4	5	6	7	8	9	
Bantu	−	−	+	−	−	−	+	+	+	32[b]
Senegal	−	+	+	−	+	+	+	+	+	1
Benin	−	−	−	−	−	+	+	−	+	0
Cameroon	−	−	+	+	−	+	+	+	−	0
Arab-Indian	+	+	+	−	+	+	+	+	−	0
Rare (1)	+	−	−	−	−	−	+	+	+	1
Rare (2)	−	+	+	+	+	+	+	+	−	1

[a] The numbers of the sites correspond to *Hin*cII/ε; *Xmn*I/$\Delta^G\gamma$; *Hin*dIII/$^G\gamma$; *Hin*dIII$^A\gamma$; *Hin*cII/5′Ψβ; *Hin*cII/3′Ψβ; *Ava*II/β; *Hpa*I/β; and *Hin*fI/β.

[b] Twenty-five were fully determined Bantu β^S-haplotypes; 7 were consistent with the Bantu β^S-haplotype, but the phase could not be determined conclusively at two or more sites.

β^S *Haplotypes of the Malagasy*

Since the actual β^S mutation is identical in all population groups, the pattern of haplotypes is important in distinguishing one population from another. The data on the Malagasy β^S haplotypes offer the most convincing evidence for an African contribution to Malagasy ancestry. The β^S mutation is thought to have arisen at least five times (Pagnier et al. 1984; Kulozik et al. 1986; Lapoumeroulie et al. 1992) or to have had two origins (Africa and Asia), with five areas of expansion that correlate to the five common β^S haplotypes (Flint et al. 1993*b*). β^S expansion is believed to have occurred when malaria became hyperendemic in world populations. Four different haplotypes have been described that are in linkage disequilibrium with the β^S gene in four distinct areas of Africa: the "Senegal" haplotype in Atlantic west Africa; the "Benin" haplotype in central west Africa; the "Bantu" haplotype in Bantu-speaking Africa (Pagnier et al. 1984); and the "Cameroon" haplotype found among the Eton ethnic group of the Cameroon (Lapouméroulie et al. 1992). A fifth haplotype, the "Arab-Indian" haplotype, was described from east Saudi Arabia and India (Kulozik et al. 1986).

Almost all of the Malagasy β^S haplotypes are of the typical Bantu type, a result that provides the first definitive biological evidence for a Bantu-speaking Negroid contribution to the ancestry of the Malagasy. Since the β^S gene is under positive selective pressure in malarial environs, it is not possible to determine the magnitude of the contribution of Bantu speakers to the Malagasy. Selectively neutral markers will be required to address these issues.

The β^S mutation in Bantu populations of equatorial, eastern, and southern Africa occurs almost exclusively on a single haplotype, the Bantu type, and therefore is thought to have occurred before the Bantu expansion,

Table 3

5′ β^A-Haplotype Frequencies in World Populations, Including the Malagasy

Haplotype	Europe[a]	Africa[b]	Asia[c]	Oceania[d]	Madagascar
−++−+	13.1	7.6	7.8	1.1	12.0
−−−−+	.3	33.8	0	.3	32.0
+−−−−	61.0	19.9	68.3	84.7	40.0
−+−++	21.2	16.9	17.4	6.2	12.0
+−−−+	.0	1.3	0	.3	4.0
−+−−−	.3	1.7	.2	.5	0
−+−−+	1.0	15.2	2.4	1.5	0
No. of Chromosomes	679	302	1005	659	25

[a] Italy, Germany, Greece, Britain, Sardinia, and Cyprus.

[b] West, north, and south Africa, and U.S. Blacks.

[c] China, Thailand, Cambodia, Japan, India, and Indonesia.

[d] Melanesia, Micronesia, and Polynesia (reviewed by Flint et al. 1993*b*).

~2,000–3,000 years ago, most likely in the region of the Benue River or in northern Cameroon. The southern Bantu populations do not have appreciable frequencies of β^S, despite having other genetic markers in common with northern Bantu speakers, although malaria is endemic among them. The β^S gene is found at very low frequencies south of the Zambezi River and is virtually absent in populations living south of the Limpopo (Ramsay and Jenkins 1985, 1987). It seems that gene flow has carried the β^S gene southward, but only after the population expansion began. Thus, the Bantu β^S mutation was most likely to have been introduced to Madagascar by people from central or east Afica but not Africa south of the Zambezi.

The single Senegal β^S haplotype found in this study is of particular interest, since this mutation has not been reported from east Africa. It may have originated from European explorers to the island, particularly Portuguese during the sixteenth century. The three main African haplotypes (Bantu, Benin, and Senegal) have been shown to be associated with the β^S mutation in Portugal (Montiero et al. 1989). There were no lasting Portuguese settlements in Madagascar, but there were shipwrecks. It is not clear whether the ships carried slaves, but there are records of criminals being transported. The rare atypical haplotypes found in our sample could be the product of mutation in the RFLP loci or the product of recombination or gene conversion.

β^A *Haplotypes of the Malagasy*

The data on the 5′ β^A haplotypes obtained in this study (table 3) are consistent with a major African contribution to the ancestry of the Malagasy, since the common African haplotype occurs at relatively high frequencies among the Malagasy. The 5′ β^A haplotype data also suggest an Asian/Oceanic influence, as well as possibly some Caucasoid (including Indian and European) ancestry. Since the sample size of the present study is very small, only the most general statistical treatment has been accorded these data. mtDNA studies of the Malagasy are also consistent with these data, because the 9-bp deletion was found on both Asian and African haplotypes (Soodyall et al., in press).

Since sickle-cell anemia does not occur in Southeast Asia, it has not been possible to determine the likely Indonesian contribution to Malagasy ancestry in this study. However, a large Arab-Indian contribution to Malagasy ancestry is unlikely, because no "Arab-Indian" β^S haplotype was detected among the Malagasy individuals included in this study. Further studies will involve characterization of other nuclear markers, including further globin variants, in an attempt to clarify the other sources of Malagasy ancestry and the genetic contributions of each to the gene pool.

Acknowledgments

We are grateful to Dr. T. de Ravel, Dr. P. Willem, Prof. J. Roux, Dr. A. Rakouth, and Dr. S. Quanbeck, for their assistance in the collection of blood samples in Madagascar. We also thank all the individuals who provided us with blood samples. We acknowledge the hard work of the Department of Haematology of the South African Institute for Medical Research for doing all the hematological analyses for this project. The Medical Research Council, the National Geographic Society (grant 5113-93), the South African Institute for Medical Research, and the University of the Witwatersrand provided financial support.

References

Burns A, Spence S, Kosche K, Ramirez F, Mears J, Schreiner H, Miller C, et al (1981) Isolation and characterisation of cloned DNA: the δ- and β-globin genes in homozygous β-thalassaemia. Blood 57:140–146

Campbell G (1992) The state and precolonial demographic history: the case of nineteenth-century Madagascar. J Afr Hist 32:427–433

———. An economic history of Madagascar, 1790–1895. Lit Verlag, Hamburg (in press)

Feinberg A, Vogelstein B (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Flint J, Harding R, Boyce A, Clegg J (1993*a*) The population genetics of the haemoglobinopathies. Ballieres Clin Haematol 6:215–262

Flint J, Harding R, Clegg J, Boyce A (1993*b*) Why are some genetic disease so common? Distinguishing selection from other processes by molecular analysis of globin gene variants. Hum Genet 91:91–117

Fourquet R, Sarthou J, Roux J, Aori K (1974) Hemoglobine S et origines du peuplement de Madagascar: Nouvelle hypothese sur son introduction en Afrique. Arch Inst Pasteur Madagascar 43:185–220

Gilman JG, Huisman THJ (1985) DNA sequence variation associated with elevated fetal $^G\gamma$ globin production. Blood 66:783–787

Kulozik AE, Wainscoat JS, Serjeant GR, Kar BC, Al-Awamy B, Essan GJF, Falusi AG, et al (1986) Geographical survey of β^S-globin gene haplotypes: evidence for an independent Asian origin of the sickle-cell mutation. Am J Hum Genet 39:239–244

Lapouméroulie C, Dunda O, Ducrocq R, Trabuchet G, Mony-Lobe M, Bodo JM, Carnevale P, et al (1992) A novel sickle cell mutation of yet another origin in Africa: the Cameroon type. Hum Genet 89:333–337

Leroi-Gourhan A, Poirier J (1953) Ethnologie de l'Union Francaise. II Asie-Oceanie-Amerique. Presses Universitaires de France, Paris

Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16:1215

Monteiro C, Rueff J, Falcao A, Portugal S, Weatherall D, Kulozik A (1989) The frequency and origin of the sickle cell mutation in the district of Coruche/Portugal. Hum Genet 82:255–258

Nagel R, Fleming A (1992) Genetic epidemiology of the β^S gene. Ballieres Clin Haematol 5:331–365

Nagel R, Ranney H (1990) Genetic epidemiology of structural mutations of the β-globin gene. Semin Haematol 27:342–359

Old J, Ludlam C (1991) Antenatal diagnosis. Ballieres Clin Haematol 4:391–428

Pagnier J, Mears J, Dunda-Belkhodja O, Schaefer-Rego K, Beldjord C, Nagel R, Labie D (1984) Evidence for the multicentric origin of the sickle cell haemoglobin gene in Africa. Proc Natl Acad Sci USA 81:1771–1773

Ramsay M, Jenkins T (1985) The *Hpa*I β-globin restriction fragment length polymorphism in southern Africa. S Afr J Sci 81:98–100

——— (1987) Globin gene–associated restriction-fragment-length polymorphisms in southern African peoples. Am J Hum Genet 41:1132–1144

Singer R, Budtz-Olsen OE, Brown P, Saugrain J (1957) Physical features, sickling and serology of the Malagasy of Madagascar. Am J Phys Anthropol 15:91–124

Soodyall H, Jenkins T, Hewitt R, Krause A, Stoneking M. The peopling of Madagascar. In: Boyce AJ, Mascie-Taylor CGN (eds) Molecular biology and human diversity. Cambridge University Press, Cambridge (in press)

Tuan D, Biro PA, deRiel JK, Lazarus H, Forget BG (1979) Restriction endonuclease mapping of the human γ globin gene loci. Nucleic Acids Res 6:2519–2544