Population genetics of abnormal haemoglobins in Burkina Faso, West Africa

R. Devoucoux†, C. Hurpin‡, D. Baudon†, J. F. Molez‡, J. F. Roux†,
M. Guilloud-Bataille§, P. Carnevale‡ and J. Feingold§

† Centre Muraz, Bobo-Dioulasso. Burkina Faso
‡ Antenne Orstom—Centre Muraz, Bobo-Dioulasso
§ Unité d'Epidémiologie Génétique, INSERM, Paris, France

Summary. The gene frequencies of haemoglobin A (HbA), HbS and HbC were studied in Burkina-Faso (BF) and in a neighbouring region of Niger, Ayorou. The frequency of HbS was higher in the Sahel region, (northern part of BF and Ayorou) than in the Savanna region. The reverse was true for the HbC gene. The major findings of this study are: (a) confirmation of a peak of HbC gene frequency in the central region of BF (Mossi plateau); (b) a possible negative correlation between the frequencies of HbS and HbC–Cavalli–Sforza and Bodmer have observed that this correlation is at a significantly different level from that expected because of the allelic relationship between HbS and HbC; (c) comparison with the data collected by Livingstone shows a modification in the fitness of the different genotypes in the last thirty years: AS individuals have a lower and AA and SS a higher fitness. Our data favour a partial selection relaxation in this region.

1. Introduction

Burkina-Faso (Upper Volta) is a landlocked country in West Africa surrounded by the Niger river. It has an area of about $270,000 \, km^2$ and a population of about $6\frac{1}{2}$ million. The country is made up of different ethnic groups, and their geographic distribution is linked to the history of Burkina Faso (BF). Various climatic regions exist, related to rainfall. The southern part of the country is savanna and the north is semi-desert sahel. *Falciparum* malaria is widespread throughout the country.

The object of this study is to establish the phenotype and gene frequencies of various haemoglobinopathies in Burkina Faso and in a neighbouring region of Niger (Ayorou). The prevalence of genes for abnormal haemoglobins in areas endemic for *falciparum* malaria has been suggested to be due to balanced polymorphism (for reviews see Allison 1964, Livingstone 1971, Motulsky 1975, Luzzatto 1979).

2. Methods

Blood samples collected by finger-prick into heparinized tubes were obtained in three regions of BF and one town (Ayorou) of the Republic of Niger located near the northeastern border of BF. Table 1 gives the characteristics of the villages and towns studied, including regional location, climatic area and major ethnic groups of the inhabitants. The samples were obtained during an epidemiological survey on rural malaria; hospitalized patients and those seen in the outpatient clinic were excluded from the study. Oral informed consent was obtained from the parents of the children and from the adults.

2.1 Haemoglobin studies

All samples were electrophoresed in Helena cellulose acetate plates at pH 8·6 in Centre Muraz (Bobo-Dioulasso). With this method some abnormal haemoglobins could not be detected; β-thalassaemia was not studied.

Table 1. Genotype and gene frequency of β chain mutations in villages and small towns of Burkina Faso and in Ayorou (Niger).

Region and ecological area	Town or village	Population	Sample (n)	Genotype frequency						Gene frequency		
				AA	AS	AC	SC	SS	CC	A	S	C
Southwest region of BF; humid savanna	Soumousso	Bobo	102	0·696	0·059	0·225	—	—	0·020	0·838	0·030	0·132
	Bare	Tyefo	284	0·810	0·060	0·109	0·014	—	0·003	0·894	0·041	0·065
	Dandé	Bobo	533	0·679	0·124	0·171	—	0·003	0·018	0·826	0·066	0·108
	Kongodjan	Mossi	330	0·748	0·057	0·179	0·006	—	0·006	0·867	0·035	0·098
	Tago	Samo / Mossi	138	0·609	0·130	0·188	0·051	0·003	0·022	0·768	0·091	0·141
	Karankasso	Bobo / Sambla	298	0·762	0·070	0·154	0·007	—	0·007	0·874	0·039	0·087
	VK4	Mossi / Bobo	443	0·700	0·111	0·178	0·004	—	0·007	0·844	0·058	0·098
Centre and south region of BF; humid savanna	Koudougou	Mossi	197	0·716	0·036	0·213	0·015	—	0·020	0·840	0·025	0·135
	PO	Gourounsi	102	0·706	0·059	0·206	0·009	—	0·020	0·838	0·034	0·128
	Donse (*)	Gourounsi / Mossi	407	0·700	0·042	0·226	0·012	—	0·020	0·834	0·027	0·139
Northern region of BF; sahel	Dori	Peuls	275	0·826	0·069	0·091	0·003	—	0·011	0·906	0·036	0·058
	Koumbri	Mossi / Kurumba	185	0·730	0·070	0·178	0·011	—	0·011	0·854	0·041	0·105
	Oursi (*)	Songhrai / Mallebe	397	0·668	0·217	0·103	0·002	0·002	0·008	0·827	0·112	0·060
	Boulel (*)	Rimaibe	122	0·762	0·123	0·098	0·008	—	0·008	0·873	0·066	0·061
	Timbolo- (*) Gargassa	Bellah	133	0·835	0·158	—	0·007	—	—	0·913	0·083	0·004
Niger; sahel	Ayorou	Songhrai	573	0·764	0·132	0·089	0·011	0·004	—	0·875	0·075	0·050

* Data collected by Labie et al.

Prevalence and density of malaria parasites were determined by the thick-smear and thin-layer techniques. At least 200 fields were examined for malaria parasites before a slide was considered negative.

2.2 Genetic and statistical analysis

Gene frequencies were estimated by the counting method. The χ^2 test was used for frequency comparisons.

3. Results and discussion

Table 1 shows the prevalence of genotypes AA, AS, AC, SC, SS and SC and the frequencies of the β^A, β^s and β^c genes. The same table also gives the data recently reported by Labie et al. (1984) concerning Burkina Faso.

Our data confirm that there is a frequency peak of the β^c gene in the central region of BF (plateau Mossi) as noted by many authors (Sansarricq, Marill, Portier and Cabannes 1959, Livingstone 1967). From there the frequency decreased in all directions. The frequency of the β^c gene is lower in many populations of the northern region of BF, in Ayorou (Niger and in two population samples we studied in Mauritania and in Lome (Benin). The frequency of β^c in these two latter populations is respectively 0·21 and 0·054, and the frequency of the β^s gene is 0·030 and 0·079 (unpublished data).

Two-by-two statistical comparisons of the gene frequencies of the population reported in table 1 are given in table 2. Keeping in mind the great number of comparisons, we suggest that differences are statistically significant if $p < 0.01$. There are some discrepancies between our data and those of Labie et al., but the ethnic groups studied are not always the same. The three populations of the central region of BF do not differ. The populations of Boulel and Ayorou are very similar, and differ from all the other populations. Most of the differences observed can be explained by the ethnic origin of the populations, but genetic drift and/or sampling errors can also account for some statistical differences.

Considering the gene frequencies reported in table 1, we found a negative but not significant correlation ($r = -0.45$; $p < 0.10$) between β^s and β^c. This negative correlation was previously reported by Allison (1956) and by Cavalli-Sforza and Bodmer (1971). For these authors the correlation observed ($r = -0.20$) is different from that expected because of the allelic relationship between S and C. The figures in table 3 are obtained by computing the Hardy–Weinberg expectation for each population and summing over all the populations. Despite the discrepancy between observed SS individuals and the expected result, the difference is not significant (the χ^2 for the total data is 10·63 with 5 degrees of freedom; $p < 0.10$). The deficit of SS individuals can be partly explained by the exclusion of patients from the study. But if we consider only the SS phenotype the difference is significant ($p < 0.01$). We also note a slight but not significant excess of AS heterozygotes. By contrast, Cavalli-Sforza and Bodmer (1971), analysing the data from West Africa collected by Livingstone (1967), found a significant deviation from the Hardy–Weinberg equilibrium. They suggest the following order of fitness values: AS>AA, AC>SC>SS, with the fitness of CC homozygotes being anywhere in the range above that of SS and SC. Our data and those of Labie et al. suggest the following order: AS>(?) AA, AC, CC, SC>SS; the fitness of CC homozygotes is, as in the previous study, in the range above that of SS.

The comparison of the two sets of data is in favour of a partial selection relaxation in this region of Africa. In order to test this hypothesis we studied:

Table 2. Two-by-two comparisons of the gene frequencies (χ^2 test).

	Bare	Dande	Kongodjan	Tago	Karankasso	VK4	Koudougou	PO	Donse	Dori	Koumbri	Oursi	Boulel	Timbolo	Ayorou
Soumousso	**	NS	NS	*	NS	NS	NS	NS	NS	**	NS	***	*	***	***
Bare		**	NS	***	NS	*	***	*	***	NS	NS	NS	NS	***	*
Dandé			*	NS	*	NS	**	NS	***	***	NS	***	NS	***	***
Kongodjan				***	NS	NS	NS	NS	*	*	NS	***	*	***	***
Tago					***	*	***	*	***	***	**	***	**	***	***
Karankasso						NS	*	NS	**	NS	NS	***	NS	***	***
VK4							**	NS	***	**	NS	***	NS	***	***
Koudougou								NS	NS	***	NS	***	**	***	***
PO									*	**	NS	***	*	***	***
Donse										***	NS	***	***	***	***
Dori											*	***	NS	***	***
Koumbri												***	NS	***	***
Oursi													NS	***	***
Boulel														***	NS
Timbolo															**

Statistical significance: NS = not significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Statistical significance: NS = not significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3. Observations of Haemoglobin A, S, C genotypes and Hardy–Weinberg expectations (see text). Computation of fitness (observed) expected frequencies and standardized fitness considering AS fitness equal to 1.

	Genotype	AA	SS	CC	AS	AC	SC
Data collected by Livingstone and analyzed by Cavalli-Sforza and Bodmer	Observed	25,374	67	108	5482	1737	130
	Expected	25,615·5	306·87	74·69	4967·2	1768·6	165·01
	Fitness	0·991	0·218	1·466	1·104	0·982	0·788
	Standardized fitness	0·89	0·20	1·31	1	0·89	0·70
Data given in table 1	Observed	3298	5	46	456	672	42
	Expected	3303·5	17·2	41·0	433·2	682·3	41·6
	Fitness	0·998	0·291	1·122	1·053	0·985	1·010
	Standardized fitness	0·95	0·28	1·07	1	0·94	0·96

1. the *Plasmodium falciparum* parasitaemia in AA, AS and AC individuals, malaria being responsible for the AS heterozygote advantage;
2. the prevalence of these three genotypes according to age—rise in the prevalence with age of AS and AC individuals is expected in case of heterozygote advantage.

Cot, (1989) found no significant difference in the density of *P. falciparum* densities among AA, AS and AC individuals in Ayorou. The same is true among the inhabitants of Dandé. In this town the prevalence of parasite densities higher than 1000/mm³ in children below two years of age was 42% in the 106 AA children and 39% in the 18 AS children. If we consider the children under 6 years, these proportions are respectively 38% ($n = 175$) and 31% ($n = 31$). These difference are not statistically significant. Labie *et al.* made the same observation in Donse and Oursi (BF).

In tables 4 and 5 the frequencies of AA, AS and AC individuals in the population of Dandé and Ayorou are given according to age. Like Labie *et al.* in the same region, we observed no significant variation in the frequency of AS individuals with age. For the frequency of AC individuals we noted no variation in Dandé but we observed a low frequency of AC children in the 0–2-year age group in Ayorou. All these data are in favour of a partial selection relaxation.

Table 4. Frequency of AA, AS and AC individuals according to age in Dandé (BF).

Age group (years)	AA		AS		AC		Total
	%	(n)	%	(n)	%	(n)	
<2	70·2	(106)	12	(18)	17·8	(27)	151
2–5	68·3	(69)	13·8	(14)	17·9	(18)	101
5–9	72·8	(59)	10	(8)	17·2	(14)	81
>9	69·5	(128)	14·1	(26)	16·4	(30)	184
Total	70	(362)	12·8	(66)	17·2	(89)	517

Table 5. Frequency of AA, AS and AC individuals according to age in Ayorou (Niger).

Age group (years)	AA		AS		AC		Total
	%	(n)	%	(n)	%	(n)	
<2	85·5	(89)	11·6	(12)	2·9	(3)	104
2–5	73·0	(120)	15·8	(26)	11·0	(18)	164
5–9	78·6	(151)	12·5	(24)	8·9	(17)	192
>9	74·3	(78)	13·3	(14)	12·4	(13)	105
Total	77·5	(438)	13·4	(76)	9·1	(51)	565

These results should be compared to recent studies in Africa because many changes have occurred in this region in the 20 years that have elapsed since the first studies of Allison (1954, 1956, 1964). In 1979 in Nigeria, Fleming, Storey, Molineaux, Iroko and Attai observed an increase in the frequency of AS individuals with age. *P. falciparum* density was less in these individuals up to age 3 years in the dry season only. Carnevale, Bosseno, Lallemant, Feingold, Lissouba, Molinier and Mouchet (1981) in the Republic of Congo, and Saurin (1984) in Senegal did not observe an increase in AS individuals with age. Bernstein, Bowman and Kaptue Noche (1980) in Cameroon, and Carnevale *et al.* (1981) found no significant difference in mean positive parasite counts in AA and AS children.

The discrepancies among these recent studies are difficult to explain, but some differences are related to:

1. the genetic background of each population (frequency of the different haemoglobinopathies, of α-thalassaemia, of the β region haplotypes associated with βs mutation);
2. the improvement in medical care (many individuals are now treated by anti-malarial drugs). This fact could mask potential genetic factors in favour of resistance to malaria.

Conclusion

Our data confirm the high frequency of the βc gene in the central region of Burkina Faso and the possible negative correlation between βc and βs in this area.

The comparisons of genotype distribution are in favour of a partial selection relaxation in this part of Africa.

In the future the complex relationships between *P. falciparum* malaria and the sickle-cell gene should be reappraised according to the epidemiology of the disease, the genetic background of the population, the level of medical care and drug resistance of *P. falciparum*.

References

ALLISON, A. C., 1954, Protection by the sickle-cell trait against subtertian malarial infection. *British Medical Journal*, **1**, 290–294.

ALLISON, A. C., 1956, Population genetics of abnormal human haemoglobins *Acta Genetica*, **6**, 430–434.

ALLISON, A. C., 1964, Polymorphism and natural selection in human populations. (*Cold Spring Harbor Symposium*) *Quantitative Biology*, **24**, 137–149.

BERNSTEIN, S. C., BOWMAN, J. E., and KAPTUE NOCHE, L., 1980, Population studies in Cameroon: haemoglobins, glucose-6-phosphate dehydrogenase deficiency and *falciparum* malaria. *Human Heredity*, **30**, 251–258.

CARNEVALE, P., BOSSENO, M. F., LALLEMANT, M., FEINGOLD, J., LISSOUBA, P., MOLINIER, N., and MOUCHET, J., 1981, Le paludisme à plasmodium falciparum et le gene de la drepanocytose en République populaire du Congo. *Annales de Génétique* **24**, 100–104.

CAVALLI-SFORZA, L. L., and BODMER, W. F., 1971, *The Genetics of Human Populations.* (San Francisco: Freeman).

COT, M., BOUSSINESQ, M., GAZIN, P., BAUDON, D., and CARNEVALE, P., 1988, Acquisition de l'immunité en zone d'endemie palustre (savane soudano-sahelienne). *Annales de la Société Belgique de Médécin Tropicale*, **68**, 25–35.

FLEMING, A. F., STOREY, J., MOLINEAUX, L., IROKO, E. A., and ATTAI, E. Q. E., 1979, Abnormal haemoglobins in the Sudan savanna of Nigeria. I Prevalence of haemoglobins and relationship between sickel-cell trait, malaria and survival. *Annals of Tropical Medicine and Parasitology*, **73**, 161–172.

LABIE, D., RICHIN, C., PAGNIER, J., GENTILLINI, M., and NAGEL, R. L., 1984, Haemoglobins S and C in Upper Volta. *Human Genetics*, **65**, 300–302.

LIVINGSTONE, F. B., 1967, *Abnormal Haemoglobins in Human Populations.* (Chicago: Aldine).

LIVINGSTONE, F. B., 1971, Malaria and human polymorphisms. *Annual Review of Genetics*, **33**, 33–64.

LUZZATTO, L., 1974, Genetic factors in malaria. *Bulletin of the World Health Organization*, **50**, 192–202.

LUZZATTO, L., 1979, Genetics of red cells and susceptibility to malaria. *Blood*, **54**, 961–976.

MOTULSKY, A. G., 1975, Glucose-6-phosphate dehydrogenase and abnormal haemoglobin polymorphisms. Evidence regarding malaria selection. In *The role of natural selection in human evolution*, edited by F. M. Salzano. (Amsterdam: North Holland).

SANSARRICQ, H., MARILL, G., PORTIER, A., and CABANNES, R., 1959, Les haemoglobinopathies en Haute-Volta. *Le Sang*, **30**, 503–511.

SAURIN, W., 1984, Etude de la selection naturelle dans les populations de faible effectif: le cas de la drepanocytose chez les Bedik et les Malinké de la région de Kedougou (Senegal Oriental). *Bulletin et Mémoires de la Société d'Anthropologie de Paris*, **14**, 37–56.

Résumé Les fréquences géniques des hémoglobines A (HbA), HbS et HbC, ont été étudiées au Burkina-Faso (BF) ainsi que dans une région avoisinante due Niger: Ayorou. La Fréquence d'HbS est plus élevée dans le sahel (nord du BF et Ayorou) que dans la zone de savane. L'inverse est trouvé pour HbC. Les résultats essentiels de ce travail sont: (a) la confirmation d'un pic de la fréquence génique de HbC dans le centre du BF (plateau Mossi); (b) une corrélation négative entre les fréquences de HbS et HbC. Cavalli-Sforza et Bodmer ont observé que cette corrélation atteint une intensité significativement différente de celle qui est attendue, par suite de la relation allélique entre HbS et HbC; (c) la comparaison de ces résultats avec les données rassemblées par Livingstone montre une modification de la valeur sélective des différents génotypes au cours des dernières années. Les individus AS ont une valeur sélective plus basse et les AA et SS en ont une plus haute. Nos données sont favorables à l'hypothèse du relachement partiel de la sélection dans cette région.

Zusammenfassung. In Burkina-Faso (BF) und einer benachbarten Region in Niger, Ayorou, wurden die Genfrequenzen von Hämoglobin A (HbA), HbS und HbC bestimmt. Die Frequenz von HbS war in der Sahelregion (nördlicher Teil von BF und Ayorou) höher als in der Savannenregion während die Frequenz des HbC Gens eine gegenläufige Tendenz aufwies. Die wichtigsten Ergebnisse dieser Studie sind: (a) die Bestätigung eines möglichen negativen Korrelation zwischen den Frequenzen von HbS und HbC – Cavalli-Sforza und Bodmer haben beobachtet, daß diese Korrelation auf einem signifikant anderen Level existiert als auf dem nach den Beziehungen zwischen HbS und HbC zu erwartenden; (c) ein Vergleich mit den von Livingstone gesammelten Daten zeigt eine Modifikation der Fitness der verschiedenen Genotypen in den letzten 30 Jahren: AS-Individuen haben eine geringere, AA- und SS-Individuen eine höhere Fitness. Die vorliegenden Daten sprechen für die Annahme eines partiellen Nachlassens des Selektionsdrucks in dieser Region.