Am J Hum Genet 39:96–111, 1986

Genetic Origins of the Dogon Population in the Arrondissement of Boni (Mali)

M. H. CAZES[1]

SUMMARY

A study of probabilities of origin of genes was carried out on a Dogon population in Mali, spread over four massifs separated from each other by about 20 kilometers. Within each village, the founder contributions are very disparate and show that each village has a very specific origin. Therefore, the exchange of wives between massifs has not resulted in a homogenization of the population, which has remained strongly structured into four relatively independent isolates.

INTRODUCTION

In studying population genetics, the traditional approach, in particular where kinship study is involved, consists of measuring the mean coefficient of inbreeding. This depends on one hand, on the genetic drift, which grows in importance as the size of the population decreases, and on the other hand, on the possibilities of spouse choice based on a kinship criterion.

This coefficient, which translates by a single figure all the complexity of the kinship chains linking the different members of a group, considerably weakens the information contained in the genealogies. But even more importantly, it constitutes a measure of the information that exists on the population, rather than the measure of a real characteristic of this population; its numerical value is, in fact, higher when the number of common ancestors taken into account increases, that is, when the genealogy goes further back into the past [1].

Another possible approach for analyzing genealogical information consists of studying the "probabilities of origin of genes." This approach permits a description of the population in both time and space. As they are cumulative, these probabilities can be analyzed using any criterion for classifying the found-

Received May 14, 1985; revised February 20, 1986.

[1] Département de Génétique des Populations, Institut National d'Etudes Démographiques, Paris, France.

ers. This paper presents the findings of a study of the latter type, conducted on a Dogon isolate in Mali.

DATA: PRESENTATION OF THE POPULATION

The Dogons of the Arrondissement of Boni, between Douentza and Hombori, have settled on the sandstone massifs that, rising vertically 600 m above the plain, overhang the region. Wishing originally to protect themselves from looting by neighboring ethnic groups, they settled either at the summit of the plateau or halfway up. According to local legend, certain villages were settled in the region five or six centuries ago. In some villages, informers have actually furnished a list of heads of families that dates back about 20 generations; this corroborates local legend. But most detailed genealogies, with collateral branches, go back only two or three centuries. In several villages, they seem to go back as far as the first settlement. A certain number of these villages have since resettled at the foot of the massifs.

In all, there are 15 Dogon villages located on four massifs, with three or four villages per massif, and a distance of 30 km at most between the massifs. The village community is organized into families corresponding to patrilineal lineages issued from the one ancestor. These sedentary farmers, isolated in the South Sahel semidesert, only subsist on one millet crop a year.

The population of 4,400 inhabitants presents a very distinctive geographical structure (fig. 1, table 1), which is reflected in the marriage patterns: highly endogamous, the Dogons have contracted few marriages with the neighboring ethnic groups (Peul, Sonraï, Rimaïbe). Preference is given to marriage between individuals living in the same massif (84% on average).

The data collected in this population relies entirely on oral tradition and on the memory of the living individuals. In each village, the chosen informant, a village notable of about 40 years of age, was spontaneously surrounded and helped by the main family heads in answering the questions. In this way, the genealogies over 9–10 generations were reconstituted. The biological and social kinship are identical, births out of wedlock being extremely rare events. A double data collection and several cross-checks are strong points in favor of the internal coherence of the data. For the first generations, omissions concern the ancestors' wives, whose villages of origin and even massifs have been forgotten.

Each individual is identified by a number and is located within the group by: father's number, mother's number, sex, dates of birth and death when these could be estimated, massif and village of birth, patrilineal lineage (called "family"), number of marriages, and the numbers of his or her successive spouses.

Each individual was given a generation number, moving down the genealogies from the "founders" (that is, the first-known ancestors) and following the matrilineal filiations that are shorter in timespan than the paternal ones [2].

The genealogies for the whole of the population thus cover 14 generations. In all, 510 "founders" are at the origin of the different families in each massif, with a total of 9,128 persons in these reconstituted family histories. A large proportion of these are unidentified wives.

METHOD: THE PROBABILITIES OF ORIGIN OF GENES

If a random selection is made of one of an individual's genes, the probability that it comes from each parent is 1/2, each grandparent, 1/4, and so on. If we go back as far as the first-known ancestors, called "founders." totalizing the coefficients corresponding to the different paths leading to these founders, we obtain a series of parameters that represent the probability W_{ij} that a gene selected in an individual i comes from the founder j.

If, for instance, we totalize these parameters for all the living individuals of a village r, we obtain the probabilities of origin "of the village," that is, the part of the genome of

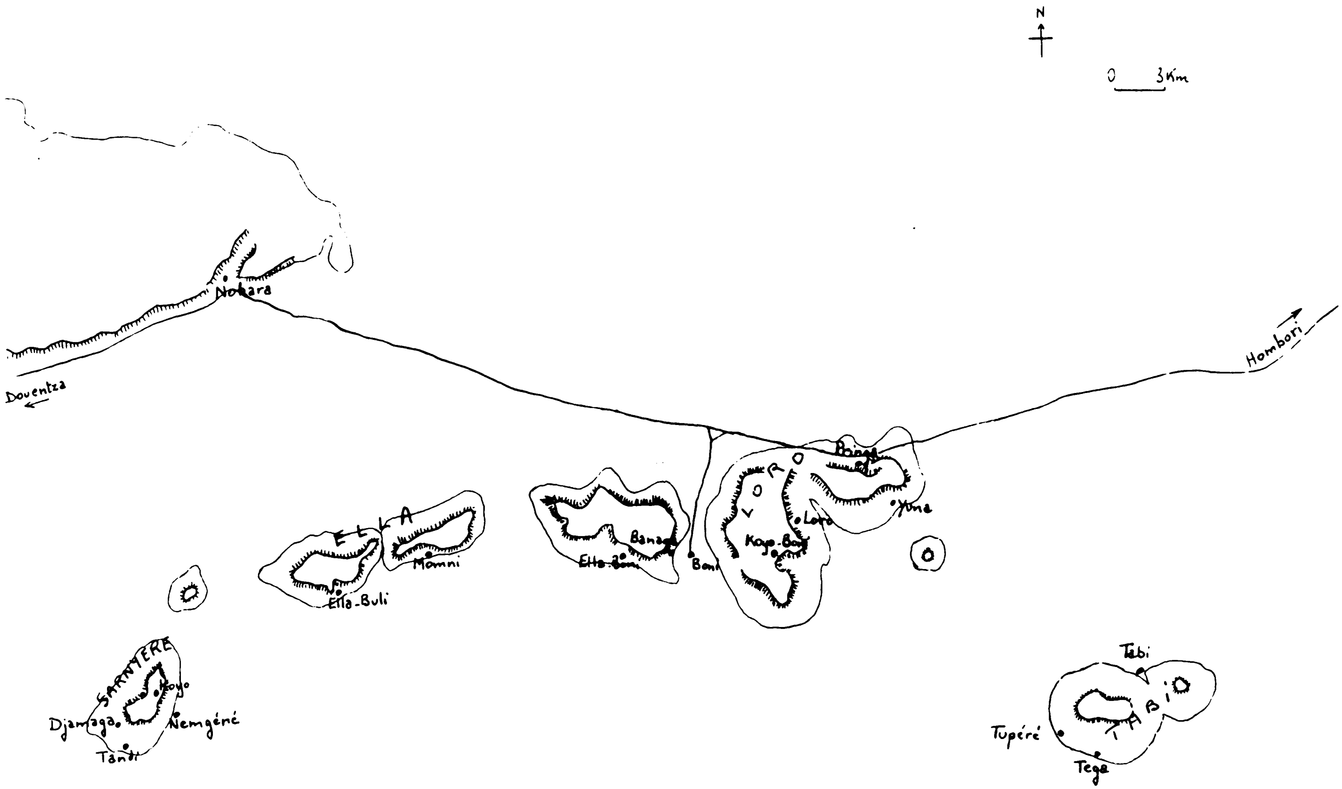

FIG. 1.—The massifs in the Boni Arrondissement and their 15 Dogon villages

TABLE 1

THE DOGON POPULATION OF BONI ARRONDISSEMENT, BY VILLAGES, IN JANUARY, 1980

Massifs	Villages	Males	Females	Total
Sarnyéré	Nemgéné	240	298	538
	Djamaga	154	181	335
	Tandi	148	133	281
	Koyo	73	86	159
Tabi	Tabi	512	551	1,063
	Tupéré	182	192	374
	Tega	197	224	421
Ella	Ella-Buli	44	49	93
	Ella-Boni	52	59	111
	Momni	60	74	134
	Banaga	52	63	115
Loro	Loro	141	160	301
	Yuna	25	35	60
	Koyo-Boni	105	85	190
	Pringa	129	114	243
	Total	2,114	2,304	4,418

the village *r* that comes from the founder *j*. Similarly, if we totalize all the founders at the origin of a village *s*, we obtain the part of the genome of the village *r* that comes from the founders of the village *s*. It is thus possible to measure the extent to which the inhabitants of a village issue from the founders of this village or from elsewhere.

To obtain an overall view of the origins of the 15 Dogon villages, we tabulated the probabilities of origin of genes for all the living population aged 0–30 years and still childless. This "population of youths," classified by village, is characterized by the genetic contribution of the founders from whom they issue (table 2); the founders are grouped according to the village in which they "founded."

The most remarkable finding is the extent of the intra-massif endogamy: in each village, the genetic pool of the youths comes, in 42% of cases on average, from their own massif, the phenomenon being much less pronounced in the massifs of Ella and Loro. A village endogamy is also observed, increasing with the size of the village to reach a maximum of 45% for Tabi.

However, before interpreting table 2, we must look into the nature of the "unknown founders," who are all unidentified wives. There are two types: (1) certain are known to be Dogon, but their village and massif of origin are unknown (they are classified in column "Dogon of unknown massif"); (2) for others, no information is available, we do not know whether or not they are Dogon or where they come from.

To simplify the analysis, we condensed these probabilities of origin by regrouping the founders into four categories (table 3): from the same massif, from another massif, from another ethnic group, and unknown. The column "unknown" reveals the necessity of further investigation into the probable origins of these wives.

In the earlier generations, it is the men and their descent who are remembered the most, for in a patrilineal society like that of the Dogons, they are the ones at the origin of the group. The wives, on the other hand, when they are born outside the village, are soon forgotten, and it is generally impossible to specify their origin: they are probably Dogon but there is no way of verifying this.

If we take as reference the same marriage behavior as that observed for the couples

TABLE 2

Origin of Genes of "the Youths" in Each of the 15 Dogon Villages

Founders origin	Sarnyéré: Nemg	Dja	Tan	Koyo	Unk	Σ	Ella: E.Buli	E.Boni	Mom	Ban	Unk	Σ	Loro: Loro	Yuna	K.Boni	Prin	Σ	Tabi: Tabi	Tupéré	Tega	Σ	Dogon of unknown Massif	Total: Dogon	Non-Dogon	Unknown
Nos.	230	195	220	157	1		144	106	125	109	3		121	108	145	152		160	112	100					
Villages:																									
Nemgéné	0.23	0.09	0.10	0.03	0.0071	0.45						0.02					0.0009				0.02	0.11	0.60	0.002	0.40
Djamaga	0.09	0.26	0.17	0.02	0.0005	0.54						0.04					...				0.003	0.07	0.65	0.04	0.30
Tandi	0.10	0.11	0.23	0.03	0.0017	0.47						0.05					0.0001				0.02	0.07	0.61	0.01	0.37
Koyo	0.11	0.14	0.12	0.07	0.0039	0.44						0.02					...				...	0.12	0.58	0.0007	0.42
Ella-Buli						0.06	0.18	0.04	0.06	0.01	0.0024	0.29					0.04				0.002	0.01	0.40	0.08	0.52
Ella-Boni						0.04	0.04	0.14	0.06	0.03	0.04	0.31					0.02				...	0.0005	0.37	0.15	0.48
Momni						0.03	0.03	0.05	0.14	0.06	...	0.28					0.03				0.0004	0.0035	0.34	0.11	0.55
Banaga						0.01	0.01	0.05	0.03	0.19	...	0.28					0.04				...	...	0.33	0.05	0.62
Loro						0.01						0.005	0.14	0.02	0.11	0.01	0.28				0.005	...	0.30	0.03	0.67
Yuna						...						0.01	0.19	0.05	0.02	0.01	0.27				0.03	...	0.31	0.01	0.68
Koyo-Boni						0.08						0.01	0.13	0.01	0.23	...	0.37				0.01	...	0.47	0.02	0.51
Pringa						0.06						0.01	0.006	0.004	...	0.14	0.15				0.04	0.06	0.32	0.01	0.67
Tabi						0.02						0.03					0.02	0.45	0.04	0.05	0.54	0.0004	0.61	0.04	0.34
Tupéré						0.03						0.13					0.01	0.10	0.12	0.10	0.32	...	0.49	0.08	0.42
Tega						0.06						0.04					0.002	9.18	0.03	0.22	0.43	...	0.53	0.05	0.42

NOTE: In Sarnyéré and Ella, column Unk: unknown consists of founders of these massifs of village unknown.

TABLE 3

ORIGIN OF GENES OF "THE YOUTHS" IN EACH OF THE FOUR DOGON MASSIFS (%)

"YOUTHS" MASSIF	No.	FOUNDERS' ORIGIN M Same massif	A Other massif	Ɖ Non-Dogon	U Unknown
Sarnyéré	720	0.48	0.14	0.01	0.37
Ella	252	0.29	0.07	0.10	0.54
Loro	427	0.26	0.09	0.02	0.63
Tabi	1,053	0.47	0.10	0.05	0.37
Weighted average		0.42	0.11	0.04	0.43

identified since the foundation of the group, the majority of these unknown wives were Dogon, and many of them came from a neighboring village. However, the limited size of the first generations may have led to a temporary opening of the population. In order to include these "unknown" wives in our data, we proceeded with the following assumption: the probability that a wife's origin is forgotten is higher if she is an alien than if she is a Dogon. Similarly, this probability of being forgotten is higher for a Dogon wife from another massif than for one from the same massif.

Taking the coefficients of proportionality α, β, γ, of the three events P(unk/M) = αa, P(unk/A) = βa, P(unk/Ɖ = γa*, where $\alpha < \beta < \gamma$, and using the observed probabilities P(M/known) = .74, P(A/known) = .19, P(Ɖ/known) = .07, P(known) = .57, we can find the value of parameter a by solving the algebraic system formed by these equalities and thus infer the conditional probabilities P(M/unk), P(A/unk), P(Ɖ/unk) as a function of the arbitrary coefficients α, β, γ (see APPENDIX).

If we choose $\alpha = 1$, $\beta = 1.7$, $\gamma = 2$, we thus obtain a new table (table 4) with only three origins. On average, 63% of the genetic pool of the youths comes from the same massif, 25% from another massif, and 12% is of some alien origin. The Ella and Loro massifs, although less closed than the other two, remain endogamous, with over 50% of genetic pool coming from the same massif.

This table is, of course, based on an a priori assumption where the values $\alpha = 1$, $\beta = 1.7$, $\gamma = 2$ are chosen arbitrarily, taking into account our knowledge of what the population tends to remember and forget. To test how sensitive the results are to this arbitrary choice, we recomputed with the values $\alpha = 1$, $\beta = 2.55$, $\gamma = 3$. (The probability of being forgotten is, in this case, three times higher for an "alien" than for a "same massif" woman.) We now obtain an average of .55 for the origin "same massif," .29 for "another massif," and .15 for an "alien" origin. These results, although based on a very extreme assumption, again reveal a preponderance of intra-massif origins.

A more detailed analysis, conducted village by village and generation by generation, has confirmed this endogamy. In each village, a set of common characteristics are observed. We have chosen to describe them for only one village, Tabi, in order to keep this paper as succinct as possible.

FINDINGS—EVOLUTION BY GENERATION: THE VILLAGE OF TABI

The largest Dogon village in the arrondissement, Tabi, has given its name to the massif. The study, based on a register updated in January 1980, concerns

* M = same massif; A = another massif; Ɖ = non-Dogon, that is, alien; unk = unknown.

TABLE 4

ORIGIN OF GENES OF "THE YOUTHS" IN EACH OF THE FOUR DOGON MASSIFS (%) INCLUDING UNKNOWN WIVES, WITH THE INITIAL ASSUMPTION $\alpha = 1$, $\beta = 1.7$, $\gamma = 2$

		FOUNDERS' ORIGIN		
"YOUTHS" MASSIF	No.	M Same massif	A Other massif	D non-Dogon
Sarnyéré	720	0.66	0.26	0.08
Ella	252	0.55	0.25	0.20
Loro	427	0.57	0.30	0.14
Tabi	1,053	0.65	0.22	0.12
Weighted average		0.63	0.25	0.12

2,118 individuals and covers 10 generations. The first founders appear in generation 4 (generations 1–3 only concern the other massifs); the children born on average since 1973 belong to generation 13. This last generation is incomplete, as yet accounting for only a small number of individuals, which explains why we have studied, at the same time, the group called "youths" comprising all the under-30s who are still childless. This group is roughly equivalent in size to generation 12 and represents the current genetic potential of the population.

(1) The 2,118 inhabitants of the village issue from only 160 founders out of the 510 listed in the general population register. This characteristic is repeated for all the Dogon villages, each one issuing from about 100 or, at most, 200 founders (see table 2): this reveals the specific nature of the origins of each village.

(2) Table 5, which represents the probabilities of origin of genes by generation, confirms this. The distribution of individuals by generation obviously retains a certain arbitrariness, and the decrease in numbers observed for the last 2 generations is only a question of attribution. The founders are classified in five groups, in decreasing order of contribution, this contribution differing greatly from one founder to another. The four founder-couples of generation 4, whose children are the seven persons in generation 5, contribute 33.6% to the current genome of the "youths." Of these first four, couple 3087 and 2619 of the Goni family—the first family to arrive in the village—alone contributes almost 12% to the genetic pool of the youths.

Over half the genome of these youths comes from only 15 founders. The second group of founders—19 in all, each with a contribution of around 1%–2%—accounts for 28% of the genetic pool of the youths. The remaining 126 have an overall contribution of under 21%. These figures confirm what the collection of genealogies had already revealed: that each village is an isolate.

(3) A third observation is the stabilization, as of the very first generations, of the contributions of the "weightiest" founders. For Tabi, the "last" 50 founders, who appear in generation 7, barely modify the genetic pool of the population, which appears to be stable as of generation 9.

TABLE 5

PROBABILITIES OF ORIGIN OF GENES (IN %) OF TABI VILLAGE INHABITANTS IN TERMS OF THE 160 FOUNDERS

Generations	5	6	7	8	9	10	11	12	13	Total	"Youths"
Nos.	7	23	41	118	185	333	707	603	101	2118	594
Founders:											
3087 and 2619	125	196	165	140	132	132	117	116	129	124	117
3193 and 2623	375	109	189	115	116	100	100	99	98	105	99
3138	...	...	73	40	47	45	46	41	45	44	44
4654 and 2639	250	44	67	58	69	67	67	69	70	68	67
2629	...	22	18	22	30	27	27	30	26	28	28
3711 and 2628	250	109	73	58	61	53	52	56	54	56	53
4243	...	43	12	15	23	28	24	25	28	24	25
4098	...	...	...	25	11	18	26	29	19	23	23
5000	...	43	64	28	27	22	22	24	24	25	23
2977	...	22	52	23	25	20	20	21	19	21	19
2640	...	22	30	16	20	20	21	18	20	19	19
The first 15	1000	610	743	540	561	532	522	528	532	537	517
The following 19	...	330	192	310	265	285	271	279	270	274	277
The following 29	...	20	35	100	115	131	135	137	158	129	147
The following 47	...	40	30	40	54	49	64	47	36	53	52
The last 50	...	...	...	10	5	3	8	9	4	7	7

If we examine the number of descendants per founder, we observe, for each generation, a parallel between the founders with the largest contributions and those with the highest number of descendants. However, these two concepts do not cover the same reality. They are strongly correlated in this case, but this is not always so. A similar study conducted on the Samaritans [3] revealed a contrast: the founders with the highest number of descendants did not necessarily have a large contribution in terms of genetic pool. In both studies, on the other hand, we observe the same stability of contributions over the different generations. This is established when the genes of the main founders are present in almost all the population.

These observations are explainable if we take into account the fact that the differential fertility of the couples, at this level, no longer has any influence. The difference in the overall probabilities of origin of genes between one generation and the next, is related to the covariance between the effective number of children and the overall probabilities of origin [1]:

For a given founder, i, we have: $Wi' - Wi = (1/m)\mathrm{cov}\,(x, Wix)$, where m is the mean number of children per family and Wix is, for the individuals with x children, the mean probability that one of their genes comes from the founder i. The probabilities of origin are therefore stabilized when this covariance becomes null, which occurs in the case of Tabi with generation 9.

Outside contributions may, however, radically change a population's genetic pool, if they occur early enough. An example of this was given by the immigrants who settled among the Jicaque Indians in the Honduras [4, 5]: sparsely introduced in the first generations of the group, they accounted for 29% of the genes at generation 6. Obviously, the same immigrants, introduced in the population 3 or 4 generations later in the same proportions, would not have had the same impact on the genetic evolution of the population.

(4) By grouping the founders according to the family they belong to, we can measure the weight of each lineage in the biological constitution of the village. From the outset, four lineages out of the nine existing in Tabi could be distinguished for their importance in the village's history: the Goni family, founder of the village, and holder of the authority; the Dungoba and Yerba families, which both dispute the chieftainship, the former much smaller in size but larger in traditional prestige, the latter numerically more important; and the Zoriba family, the largest in size.

When dictating their genealogies, the Dogons had themselves respected this order for the first four families. The other five families, on the contrary, had been collected according to the availability of the head of the family. It would seem that the Dogons' perception of the preponderance of these families corresponds to a biological reality, since these first four families all appear in table 6 in the exact order of their hierarchy. The true weight of each family in the current genetic constitution of the population is thus revealed.

Out of the 66% of identified founders, 45% are from the village itself and 54% from the same massif. The contributions from the other three massifs are very small compared with those from the same massif.

TABLE 6

ORIGIN OF THE GENES CLASSIFIED BY FAMILIES, VILLAGES, OR MASSIFS FOR TABI VILLAGE YOUTHS (594 YOUTHS, 160 FOUNDERS)

	Weight of founders
Families:	
Goni	0.0767
Dungoba	0.0669
Yerba	0.0515
Zoriba	0.0495
Warguri	0.0368
Yukaraba	0.0332
Ganemé	0.0301
Seguiwa	0.0181
Morba	0.0019
Unknown	0.0819
Tabi village total	0.4466
Yukaraba	0.0154
Ankoleba	0.0009
Unknown	0.0344
Tega village total	0.0507
Koreba	0.0201
Yamni	0.0032
Yerba	0.0011
Unknown	0.0179
Tupéré village total	0.0423
Tabi massif total	0.5396
Villages:	
Ella-Boni	0.0298
Others	0.0032
Ella massif total	0.033
Nemgéné	0.0206
Tandi	0.0029
Sarnyéré massif total	0.0235
Loro	0.0184
Others	0.0033
Loro massif total	0.0217
Dogon, of unknown massif	0.0004
Dogon total	0.6182
Non-Dogon	0.0430
Unknown	0.3388

DISCUSSION

The information given in table 2, presented above and which is a summary of the history of the 15 Dogon villages, is actually decreased because of the high proportion of unknown female spouses in the first generations. If consanguineous marriages were as frequent then as nowadays—which, in our view, was probably the case—some of these spouses must descend from some of the ancestors. The omission of these kinship links considerably decreases the calculated endogamy. Here, we wish once again to emphasize that an inbreeding measure is a numerical value of an information (that given by the known genealogies) and not a reality.

The possible omissions of the male ancestors, in these first generations, must be of small importance: the memory of a male is linked to the descendance he fathered. If some male ancestors have been forgotten, either they had no children or their descendance died out rather rapidly. These omissions are therefore of little importance for our study. On the other hand, it may happen that a founding ancestor was a Dogon from another massif who came and married—a rare exception to the rule of patrilocality that imposes on the wife to live with her husband's family. Because of that emigration, the identity of the parents may be lost. At Tupéré, the link with the village of origin could be found once, but for two other families in the massif, the parents of the ancestor were unknown and he was registered in his offspring's village. Omitting these kinship links with a neighboring massif decreases, in terms of origin, the column "other massif" in table 3 and thus increases the column "same massif." But, in terms of group endogamy, these omissions have no effects since, once the ancestor has emigrated, his descent always marries within the massif of arrival. Table 4, presented above in order to include the unknown wives, seems to us to offer a fairly good evaluation of the true origins of the Dogon villages.

The characteristics observed for Tabi are repeated for all the other villages studied: we find everywhere the same disparity in the contributions of the different founders, proving how specific the origins of each village are. The majority of the contributions come from roughly 10 founders per village. This is a characteristic of small endogamous populations, which are necessary conditions for a founder effect. It is to be found in several other isolate studies, whether religious in the United States [6], geographic [7], or cultural [8].

It is in the Sarnyéré massif that this disparity is the least pronounced: on one hand, the number of founders observed for each village is higher—some 200 founders, except for Koyo (see table 2); on the other hand, the highest contributions per founder are around 3% (as against 6% in the other massifs). This may be explained by the fact that the population of Sarnyéré covers a larger number of generations, the oldest ancestors belonging to generation 1, whereas they appear only at generation 4 in Tabi and many other villages. The probabilities of origin obtained for the current generation are therefore smaller in the Sarnyéré massif.

The Ella massif is characterized by smaller villages—100 inhabitants at most, which necessitates more exchange between villages within the massif. These small village populations may explain the opening of the massif to other ethnic groups: on average, 10% of its genetic pool comes from "alien" wives.

In the Loro massif, almost all the villages (three out of four) are coupled by name with a Peul village in the plain, at the foot of the mountain. However, intermarriages are rare between these "coupled villages." The identified alien wives (Peul, but above all Sonraï or Rimaïbe) only contribute 1%–3% to the population's genetic pool. "Unknown" origin, on the other hand, is 10% higher than in Ella, which could correspond to the presence of alien wives in the first generations.

The second possible aspect of this study is inbreeding. The averages of the inbreeding coefficients for each generation for all massifs or villages were measured from the genealogical network [9] (table 7).

If the thirteenth generation is excluded, since most individuals have yet to reach the reproductive age, it can be seen that the Tabi massif is the most inbred. This is a result of the village size—particularly, that of Tabi, which is large enough to be a self-sufficient matrimonial market. The population is large enough so that there is no need to marry a woman alien to the village or the massif. This increases the complexity of close kinship network. The Sarnyéré, whose profile of origin of genes is very similar to that of Tabi is, nevertheless, the least inbred. In their case, it is the endogamy at the massif level that is predominant. The relations between the villages are very close, and this creates a genealogical network at the massif level rather than at the village level. Finally, the massifs of Ella and Loro, which presented a wider origin of genes as the result of the presence of a certain percentage of "alien" women in the first generations, nevertheless attain values of inbreeding comparable to those of the other massifs. Although the colonization of Ella and Loro is as ancient, if no more, as that of Sarnyéré or Tabi, the inbreeding there appears later but also more rapidly, as soon as the population size is sufficient to allow an endogamous behavior. Thus, the results calculated from the probability of origin of genes are in agreement with the relatively intense inbreeding found in all four massifs.

The genetic origins matrix cross-classifying the 510 Dogon founders with the group of youths divided into villages provides a general, statistical representation of the origins of this population. Each village is represented by a vector of 510 elements, over 3/5 of which are null.

TABLE 7

AVERAGE INBREEDING COEFFICIENTS PER GENERATION AND PER MASSIF, IN DOGON POPULATION

	Massifs				
Generation	Sarnyéré	Tabi	Ella	Loro	Total
1–6	0.0	0.0	0.0	0.0	...
6	0.01	0.0	0.0	0.0	*
7	*	0.19	0.0	0.0	0.06
8	0.18	0.23	0.0	0.0	0.17
9	0.24	1.28	0.0	0.0	0.64
10	0.63	1.77	0.12	0.16	1.02
11	1.11	2.29	0.76	0.90	1.54
12	1.29	2.67	1.82	1.46	1.98
13	2.11	3.02	2.53	3.08	2.65
Total	0.93	2.11	1.01	1.01	1.45

* Not zero, but negligible.

Finally, we considered a useful conclusion to this study would be to compute the distances between villages (using the distance of χ^2, which is well suited to the probabilities of origin of genes), and make a spatial representation of them through a classical factorial analysis. Each of the 15 villages is defined in the 510-dimensional space of the founders. Figure 2 shows the representation obtained by projecting this space into an optimum three-dimensional space (46% of the variance). The analogy between this representation and the geographical structure of the group is evident. The grouping of the villages into four massifs holds no surprise: it reflects the intra-massif endogamy that has been clearly demonstrated.

The Dogons of the Arrondissement of Boni form a well-defined population, genetically distinct from the other populations—Peul, Sonraï, Rimaïbe—with which it is in contact; but it is not homogeneous for all that and by no means represents a "panmictic population" according to the classical model used by geneticians. It is, in fact, very strongly structured, with internal barriers that, although they may be crossed, have significant consequences on the population's genetic pool.

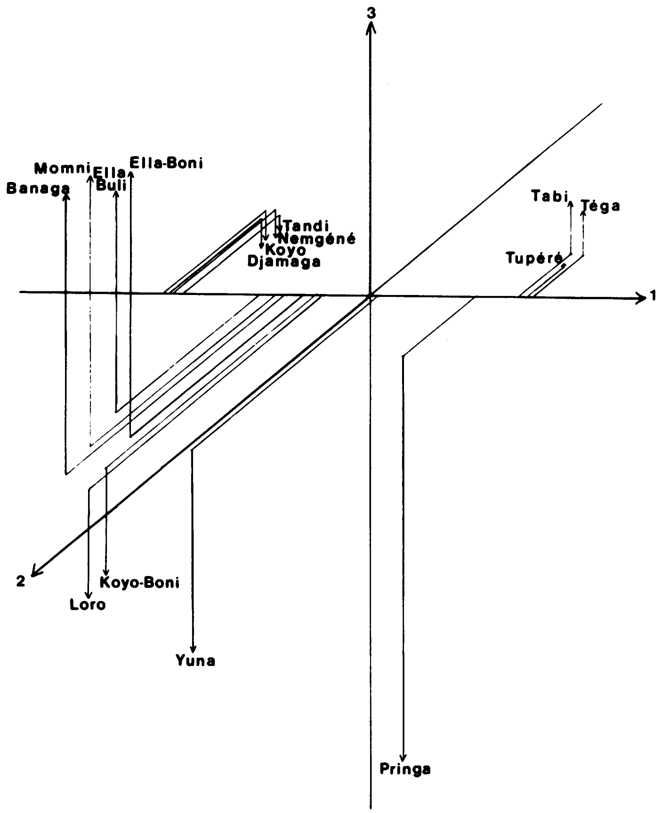

FIG. 2.—Factorial analysis of the genetic origins matrix; representation of the 15 Dogon villages in a three-dimensional space.

Analyses of the current composition of this pool (blood and immunological systems, various types of marker), as opposed to the sole analysis of its origins, must, if their findings are to be validly interpreted, take into account this internal structure. For instance, the frequency observed for the *A* allele of the ABO system is 7% in Sarnyéré and 29% in Tabi: the mean frequency of this allele among the Dogons has practically no significance. Once more, attention must be drawn to the danger of a rapid interpretation of data obtained from insufficiently complete observations.

There is a tendency to measure differences between groups or even between "races." However, most of the genetic variations observed in mankind are, as this study confirms, not between the means of different groups but between individuals within a same group.

APPENDIX

ESTIMATION OF PROBABILITY OF GEOGRAPHIC ORIGIN OF UNKNOWN SPOUSES

Given:

$$\text{P(same massif/known)} = \text{P(M/k)} = .74$$

$$\text{P(other massif/known)} = \text{P(A/k)} = .19$$

$$\text{P(non-Dogon/known)} = \text{P}(\not{D}\text{/k}) = .07$$

$$\text{P(known)} = \text{P(k)} = .565 \,,$$

we take:

$$\text{P(unknown/same massif)} = \text{P(unk/M)} = \alpha a$$

$$\text{P(unknown/other massif)} = \text{P(unk/A)} = \beta a$$

$$\text{P(unknown/non-Dogon)} = \text{P}(\text{unk}/\not{D}) = \gamma a \,.$$

According to the Bayes theorem, we infer

$$\text{P(M/k)} = \frac{\text{P(M)} \cdot \text{P(k/M)}}{\text{P(k)}}$$

$$= \frac{\text{P(M)}(1 - \alpha a)}{.565} = .74 \,,$$

and whence

$$\text{P(M)} = \frac{\text{P(M/k)} \cdot \text{P(k)}}{\text{P(k/M)}} = \frac{.74 \times .565}{1 - \alpha a} \,.$$

Similarly, we obtain:

$$P(A) = \frac{P(A/k) \cdot P(k)}{P(k/A)} = \frac{.19 \times .565}{1 - \beta a}$$

and P(Đ) = 1 − P(M) − P(A). If we start with

$$P(\text{Đ}/k) = \frac{P(\text{Đ}) \cdot P(k/\text{Đ})}{P(k)}$$

$$= \frac{[1 - P(M) - P(A)] \cdot P(k/\text{Đ})}{P(k)} ,$$

we have the following equality:

$$\frac{[1 - P(M) - P(A)] \cdot (1 - \gamma a)}{.565} = .07$$

or $1 - \gamma a - (1 - \gamma a) \cdot P(M) - (1 - \gamma a) \cdot P(A) = .04$, which leads to a third-degree equation in a: $a\beta\gamma\, a^3 - (.96\alpha\beta + .89\alpha\gamma + .58\, \beta\gamma)a^2 + (.85\alpha + .54\beta + .47\gamma)a - .435 = 0$.

The root a of the equation must satisfy the condition: $0 \leqslant \gamma a \leqslant 1$ as $0 \leqslant a \leqslant 1/\gamma$.

By taking, for example, $\alpha = 1$, $\beta = 1.7$, $\gamma = 2$, we obtain the following equation: $3.4a^3 - 5.396\, a^2 + 2.723\, a - 0.435 = 0$.

The suitable root must satisfy $0 \leqslant a \leqslant 0.5$. Once we have found $a = .3359$, we have:

$$P(unk/M) = .3359$$

$$P(unk/A) = .5711$$

$$P(unk/\text{Đ}) = .6719 ,$$

and we want to calculate

$$P(M/unk) = \frac{P(M) \cdot P(unk/M)}{P(unk)} .$$

We can break down P(M), the probability of belonging to the same massif:

$$P(M) = P(M \cap k) + P(M \cap unk)$$

$$= P(k) \cdot P(M/k) + P(unk) \cdot P(M/unk) .$$

Whence

$$P(M/unk) = \frac{P(k) \cdot P(M/k) \cdot P(unk/M)}{P(unk) \cdot P(k/M)} .$$

P(A/unk) and P(𝔻/unk) are calculated in the same way, which allows us to distribute the unknown origins in table 2 according to the proportions obtained.

In our example, this gives:

$$P(M/unk) = .4868$$

$$P(A/unk) = .329$$

$$P(\mathbb{D}/unk) = .1865 \ .$$

REFERENCES

1. Jacquard A: *Concepts en Génétique des Populations*. Paris, Ed. Masson, 1977
2. Cazes MH: "Chacun appartient évidemment à une génération. Mais laquelle?" *Population*. In press, 1986
3. Cazes MH, Bonne-Tamir B: Genetic evolution of the Samaritans. *J Biosoc Sci* 16:177–187, 1984
4. Chapman A, Jacquard A: Un isolat d'Amérique Centrale: les Indiens Jicaques du Honduras, in *Cah Travmat et Doc* no. 60, INED, PUF, 1971
5. Pison G, Vu Tien J: Les Indiens Jicaques du Honduras. Mise à jour des généalogies. *Cah Anthropol Ecol Hum* II(2):131–141, 1974
6. Eaton J, Mayer A: Man's capacity to reproduce. *Hum Biol* 25:1–58, 1954
7. Roberts DF, Bear JC: Measures of genetic change in an evolving population. *Hum Biol* 52:773–786, 1980
8. Chaventre A: Evolution anthropo-biologique d'une population Touarègue. *Cah Travmat et Doc* no. 103, INED, PUF, 1983
9. Serre JL, Babron MC, Cazes MH: The growth of inbreeding in a hierarchical population. Interest in anthropological studies. Submitted for publication