Studies on African Pygmies. IV. A Comparative Study of the HL-A Polymorphism in the Babinga Pygmies and Other African and Caucasian Populations

JULIA G. BODMER[1] AND WALTER F. BODMER[1]

INTRODUCTION

Population studies of complex polymorphisms such as the Rhesus and Gm system have contributed substantially to an understanding of their genetics. They have also helped in delineating the relationships between the various human races, and in studying microevolutionary differentiation in human populations. The high degree of variability and the complexity of the human HL-A tissue antigen polymorphism make this an especially interesting subject for population studies. In this paper we describe the results of a study of the distribution of the HL-A antigens in the Babinga Pygmies of West Africa and compare this with the distribution in Bantu (generic name used for non-Pygmy Africans in the area we studied), American black, and Caucasian populations.

During the last few years the identification and definition of human leucocyte antigens has progressed by leaps and bounds. There are now thirteen antigens in the HL-A system defined and given an official WHO designation (WHO 1968, 1970 [in press]), and many laboratories have reported or are working on the definition of further antigens (see e.g., Terasaki 1970).

The first genetic polymorphism for white cell antigens, called group 4, was described by van Rood (1962) as a simple two-allele system. Subsequently (see e.g., Balner et al. 1965; Curtoni et al. 1967) this turned out to be a complex system including many other associated antigens. A second apparently independent polymorphism called LA was described by Payne et al. (1964). The independence was based mainly on a lack of population associations between the then known LA and 4 antigens and also on incorrect family data (see W. Bodmer et al. 1966). Subsequently Dausset, Ceppellini, and others (see Curtoni et al. 1967) presented both family and population association data which clearly showed that the genetic determinants for all the antigens of the LA and 4 (or 7) series were very closely linked, thus belonging to one complex polymorphic system called HL-A. However, as pointed out by Bodmer (Curtoni et al. 1967, pp. 185–186; see also W. Bodmer et al. 1969), the general lack

Received January 9, 1970; revised March 18, 1970.

This work was supported in part by research grants GM 14650 and 10452 and by a career deve'opment award to W. F. B. (GM 35002) from the National Institutes of Health, U.S. Public Health Service.

[1] Department of Genetics, Stanford School of Medicine, Stanford University, Stanford, California 94305.

of population associations between the LA and 4 antigens, and the different characteristics of the corresponding sera still supported the concept that these two series of antigens were in some sense differentiated within the HL-A system. The LA series was developed from two "segregant" (see Mi and Morton 1966) antigens HL-A1 and HL-A2 to a multiple allelic series including HL-A3 (W. Bodmer et al. 1966) and HL-A9 (formerly LA4, Bodmer in Curtoni et al. 1967, pp. 185-186, W. Bodmer et al. 1969; Lc11, Walford et al. 1968; TO12, Ceppellini et al. 1967; 15, Dausset et al. 1968*a;* Te4, Mickey et al. 1969) by searching for new specificities which showed a mutually exclusive segregant relationship to HL-A1, HL-A2, and subsequently HL-A3. A similar relationship between some of the antigens associated with the 4 (or 7) series had been emphasized by van Rood et al. (1965), and also had been pointed out by Dausset et al. (1965) and Ceppellini et al. (1967). It has now been recognized that the 4 series, like the LA series, can be defined in terms of a mutually exclusive segregant set of specificities (Dausset et al. 1968*b;* Kissmeyer-Nielsen et al. 1968; Mickey et al. 1969). The original division into antigens 4a and 4b, and the subsequent confusing complexity of this series of antigens is probably explained by the extensive occurrence of cross reactions among the antigens grouped within each of these two categories.

In summary, it now appears that the majority of the HL-A antigens can be separated into two major series, LA and 4 (or 7), each of which behaves as if it were controlled by a set of alleles. All haplotypes, therefore, consist of one genetic determinant (or a blank) from each of the LA and 4 antigen series.

The present study on the HL-A types of the Pygmies was carried out in January of 1968 in collaboration with Cavalli-Sforza, and is based on blood samples taken from Pygmies from three areas belonging to a large group called Babingas, living near the boundaries of the Central African Republic, Cameroon, and Congo Brazza. In addition, blood samples were taken from 24 Bantu patients in a hospital and five Bantu from a neighboring village to use as a basis of comparison. In all, blood samples from 266 Pygmies and 29 Bantu were collected and typed for leucocyte antigens.

African Pygmies form one of the few human populations still living a life of hunting and food gathering, in the manner of all human beings during the Pleistocene. These very primitive people, who probably once occupied the whole of the tropical forest belt of Africa, are now only to be found in scattered groups in the forest areas. Because they are one of the few remaining primitive populations, and because of their interesting position in African anthropology and the uncertainty of their origin, Cavalli-Sforza and coworkers (1969) have been carrying out an intensive combined population genetic, demographic, and anthropological study of the Pygmies, before the influence of neighboring African cultures submerges their identity as a genetically and culturally distinct group. At the present time, although the Pygmies living on the periphery of the forest near other African groups are to some extent influenced by their culture, it is probable that there is relatively little gene flow between the Pygmies and their immediate neighbors. The reason is that almost the only intermarriage which takes place is between Pygmy women and Bantu men, resulting in the Pygmy women leaving the Pygmy group and joining the Bantu village. Relationships between Bantu women and Pygmy men are apparently very rare or non-

existent, so that there is little or no admixture with the Pygmy groups who maintain their traditional way of life in the forest. (See Turnbull [1965] for a description of a closely similar Pygmy group.)

MATERIALS AND METHODS

The method of collecting the blood samples in the field and storing them for shipment to Stanford was described in detail by W. Bodmer and Gerbrandt (1968). In brief, 10–15 ml of venous blood was drawn into a sterile Vacutainer containing ten 2 mm-diameter glass beads, and defibrinated by vigorous shaking for six minutes. Red cells were sedimented in the Vacutainers for about 30 minutes, after the addition of a half volume of sterile 3% solution of Dextran in phosphate-buffered saline. The leucocyte supernatant was then removed and resuspended in 1 ml of McCoy's 5a medium with 30% fetal calf serum. White cells stored sterile in this way survive two weeks or more, and so could be shipped back to Stanford by air and frozen for later testing.

The fluorochromatic cytotoxicity assay, as described by W. Bodmer et al. (1967) was used for the determination of HL-A types. Sources of sera were, generally, as described by W. Bodmer et al. (1967, 1969), J. Bodmer et al. (1970), and Miggiano et al. (1970).

For the WHO-designated antigens, the principal sera used were as follows: HL-A1: 1014 (Gillespie), 4036 (Morrison); HL-A2: 1087, 4025 (Pinquette); HL-A3: 1045 (Malone), 4022 (Storm), 4040 (Thornton); HL-A7: 1047, 4007 (Cutten); HL-A8: 4010 (Willett); HL-A9: 3004 (Sykes), 4005 (Hunt), 4017 (MacDonald); and HL-A12: 1050 and 1089.

The sera 3001 (Tuckerman: HL-A1, HL-A9, and HL-A12), 3002 (Caminiti HL-A2, Ba*, LA-W [see below]), and 3003 (Cardona: HL-A9, HL-A12) were also used to help identify the relevant antigens. The sera 1014, 1087, 1045, 1047, 1050, and 1089 are Stanford pregnancy sera; 4036, 4025, 4022, 4040, 4007, 4010, and 4005 are sera made by Walford by planned immunization; 4017 is a serum from Amos; 3001, 3002, 3003, and 3004 are sera made by planned immunization, in collaboration with Dr. Payne and Hyland Laboratories (see Payne et al. 1970). All sera, except the pregnancy sera, are in the NIH Serum Bank. The antigen HL-A12, which is a 4a-associated antigen (formerly 4a11 or TO11), was defined mainly by the joint reaction of sera 1050 and 1089.

In addition to the WHO-designed HL-A1, 2, 3, 7, 8, 9, and 12, sera identifying at least two further antigens were used in our studies. The antigen 4c as originally defined by agglutination (W. Bodmer and Payne 1965) was most probably the same as HL-A5 (see W. Bodmer et al. 1966). Its definition is now based on the cytotoxic activity of a pregnancy serum (Rafter or 1027) which is now known to include the original 4c and HL-A5 (see Curtoni et al. 1967, Appendix) and, in addition, at least two further specificities (see Terasaki 1970). Absorption studies indicate that this serum contains only one antibody, which therefore presumably reacts with both HL-A5 and the other specificities (W. Bodmer et al. 1967; see also Thorsby 1969). In order to distinguish the reactions of 1027 (Rafter) from 4c, we refer to the corresponding set of specificities as 4c*. The antigens 4c*, HL-A7, 8, and 12 behave as

members of the 4 series. Serum 3002Q contains, in addition to anti-HL-A2, antibodies directed against Ba* (Kissmeyer-Nielsen et al. 1968) and a new antigen of the LA series, provisionally called LA-W (J. Bodmer et al. 1970).

Since sera tested in one population may, for a variety of reasons, behave somewhat differently in other populations (see the discussion below), we required, in general, agreement between two or more sera to be reasonably sure of an antigen assignment. Serum 1014, which has been shown to be monospecific using Caucasian populations, was negative with all the Pygmies. The sample of 4010 which was used has also been shown to be monospecific and behaved consistently in all populations. In general, the reactions of the other sera showed similar patterns of associations in non-Caucasian as in Caucasian populations, indicating the presence of the same common antibody specificities, and so reinforcing the basis for antigen assignment. The typing of the non Caucasian populations, nevertheless, will clearly be somewhat more subject to error because of the differences in the behavior of individual sera.

The white cells from the Pygmies and Bantu were tested on four occasions with different sets of leucocyte-typing sera. An initial test was done in Africa on about fifty of the cells, using the basic antigen-defining sera to test the viability of the cells. When the cells arrived in Stanford, an aliquot was taken from all the donors (255) before the cells were frozen and these were tested with thirty antisera to give a fairly comprehensive typing. After a few months, this was followed by a repeat test to confirm the reactions to these sera and to test their reactions to other unidentified leucocyte antibodies. In addition, sera from about a hundred multiparous Pygmy women were checked for antibody. Later the antigens HL-A9 and HL-A12 were more clearly defined, and the antigens Ba* and LA-W were identified in serum 3002Q, which also contains HL-A2, as described above. We therefore retyped about 180 of the cells (selecting those cells in good condition) to determine their reactions to these antibodies and to reconfirm previous results. In all, most of the cells were tested on about 100 antisera, defining nine known antigens and other so-far-unidentified specificities, in addition to some sera from multiparous Pygmy women.

RESULTS

Antigen Phenotype Frequencies

The population frequencies of the HL-A antigens in the Pygmies, Bantu, and two American population samples taken in and around Stanford, California, one black and one white, are shown in table 1. The heretogeneity χ^2 for testing the significance of the differences in antigen frequency between pairs of these populations are shown in table 2. The Bantu and American blacks do not differ significantly, which is hardly surprising in view of the small number of Bantu tested. Nevertheless, we have felt it worthwhile, because of the a priori expectation that they should differ, to treat them as a separate group. We have recently been able to obtain more Bantu samples from the Central African Republic and will report on these in a future publication. All the other pairs of populations show very significant overall differences. The interpretation of the extent of these differences will be discussed later in terms of genetic distances.

Gene Frequencies

A clearer picture of the comparative antigen patterns in different populations is obtained by looking at the gene rather than the phenotype frequencies. Gene frequencies were fitted separately by maximum likelihood for the LA and 4 series of antigens, assuming genetic control by multiple allelic series (including a blank allele)

TABLE 1

PHENOTYPE FREQUENCIES OF HL-A ANTIGENS IN FOUR POPULATIONS

HL-A ANTIGEN	POPULATION							
	American Whites		American Blacks		Bantu		Pygmies	
	Total	%	Total	%	Total	%	Total	%
LA Series:								
HL-A1	277	25	75	9	29	7	184	0
HL-A2	276	51	75	37	27	56	176	24
HL-A3	276	38	75	19	28	18	178	21
HL-A9	260	24	74	35	29	34	166	20
3002Q (HL-A2, Ba*, LA-W)	98	68	47	70	27	81	110	68
4 Series:								
HL-A12	275	31	72	28	29	48	183	19
4c*	275	42	75	33	29	34	185	17
HL-A7	276	33	75	19	29	28	182	17
HL-A8	272	17	62	13	27	7	186	7

TABLE 2

POPULATION HETEROGENEITY OF HL-A ANTIGEN FREQUENCIES

ANTIGENS	POPULATION COMPARISONS			
	Pygmy and Bantu	Bantu and American Black	Combined American Black and Bantu and White	Pygmy and American White
LA:				
HL-A1	12.3	0.05	7.2	52.0
HL-A2	11.0	3.0	5.3	29.3
HL-A3	0.02	0.1	8.7	17.7
HL-A9	6.4	0.03	3.2	1.2
3002Q	1.9	0.6	0.03	0.04
4:				
HL-A12	12.0	3.8	0.5	6.9
4c*	5.1	0.03	1.1	31.7
HL-A7	1.3	0.06	4.9	34.3
HL-A8	0.01	0.2	0.3	8.8
Total	50.03	7.87	31.23	181.94
df	9	9	9	8
P in %	$<<0.1$	>10.0	<0.1	$<<<0.1$

NOTE.—The values in the body of the table are heterogeneity χ^2 between populations (1 df) calculated separately for each antigen.

and random mating. The frequency for the combination of Ba* and LA-W was estimated by subtracting the frequency of HL-A2, and fitted, ignoring 3002Q from that for 3002Q when used in place of HL-A2. The gene frequencies for the same four populations already discussed are given together with their standard errors in table 3. The fit of observed-to-expected phenotype frequencies based on these data was good, giving only one out of eight χ^2 significant at the 5% level. This supports the validity of the genetic model for the control of the HL-A antigens for all these populations.

There are several significant differences in the individual gene frequencies between the Caucasian and African populations. For example, HL-A1 is absent in the Pygmies and occurs with a much lower frequency in American blacks and the Bantu than in the American whites. The frequencies of all the remaining alleles, except for HL-A9

TABLE 3

FREQUENCIES OF THE LA AND 4 ANTIGEN SERIES ALLELES OF THE HL-A SYSTEM

ANTIGENS	PYGMY		BANTU		AMERICAN BLACK		AMERICAN WHITE	
	Freq.	SE	Freq.	SE	Freq.	SE	Freq.	SE
LA:								
HL-A1	.....	.....	.023	.023	.051	.019	.134	.020
HL-A2	.117	.022	.294	.075	.202	.036	.281	.027
HL-A3	.094	.019	.099	.045	.090	.025	.183	.023
HL-A9	.093	.019	.201	.064	.190	.035	.134	.020
LA-W+Ba*	.319	.046	.271	.117	.280	.063	.201	.045
Blank	.377	.052	.112	.123	.187	.068	.067	.046
4:								
HL-A12	.120	.022	.258	.071	.151	.031	.176	.023
4c*	.083	.018	.182	.061	.187	.034	.221	.025
HL-A7	.052	.015	.093	.045	.105	.026	.139	.020
HL-A8	.034	.012	.023	.023	.065	.021	.093	.017
Blank	.711	.031	.444	.083	.492	.045	.371	.030

NOTE.—Gene frequencies are fitted separately, by maximum likelihood estimation assuming random mating, for each series of antigens. The frequency of LA-W + Ba* is estimated by subtracting the HL-A2 frequency from the allele frequency obtained when 3002Q is substituted for HL-A2. The second set of figures in each population column are the standard errors of the gene-frequency estimates.

and LA-W + Ba*, are also higher in Caucasian than in the other populations, with antigens HL-A3, HL-A7, and HL-A8 showing the biggest difference. The combination LA-W + Ba* is the only one which has a consistently lower frequency in the Caucasians. Perhaps the most striking difference between Pygmies and Caucasians is the very much higher frequency of the blank alleles in Pygmies for both the LA and the 4 series. This suggests the existence of as yet unidentified antigens which have much higher frequencies in Pygmies than in Caucasians. Apart from HL-A3, HL-A9, and HL-A12, the Bantu and American black frequencies tend to lie between those of the Pygmies and the Caucasians.

Associations between LA and 4 Series Alleles

The gametic association Δ (sometimes called the linkage disequilibrium parameter) between two alleles determining a corresponding pair of antigens can be estimated directly from the 2×2 table of phenotype frequencies for the two re-

spective antigens. Thus, if a, b, c, and d are the respective frequencies of + +, + −, − +, and − − phenotypes, then

$$\Delta = \sqrt{d/n} - \sqrt{\left(\frac{b+d}{n}\right)\left(\frac{c+d}{n}\right)}, \quad (1)$$

where $n = a + b + c + d$ (Ceppellini 1967; Ceppellini et al. 1967; W. Bodmer et al. 1969). The significance of the departure of a Δ estimate from zero is given simply by the 2 × 2 χ^2 calculated from the phenotype frequencies. A formula for the variance of the estimate of Δ given by equation (1) is derived in the Appendix. Estimates of Δ

TABLE 4

ASSOCIATIONS (Δ VALUES) BETWEEN LA AND 4 ALLELES FOR WHITE (W), BLACK (B), AND PYGMY (P) POPULATIONS

LA	4			
	HL-A12	4c*	HL-A7	HL-A8
HL-A1:				
W	−.022†			.042†††
B	−.008			.009
P				
HL-A2:				
W				
B				
P				
HL-A3:				
W		.014	.024††	
B		−.025	.011	
P		.028†††	.007	
HL-A9:				
W	.033††			
B	.030			
P	.025††			

NOTE.—Only those cases in which at least one population showed a significant association are given. All standard errors were approximately .01 (see Appendix).

† $P \sim .05$.

†† $P \sim .01$.

††† $P = .001$ or less.

for pairs of alleles, one from each of the LA and 4 series, are given for white, black, and Pygmy populations in table 4. Values are only given in those cases for which at least one of these three populations showed a significant association at the 5% level. Most pairs of alleles show no significant association, as pointed out before (see e.g., W. Bodmer et al. 1969; also Ceppellini et al. 1967). This, in part, justifies considering the LA and 4 series as if they were independent loci, from the point of view of estimating gene frequencies and assessing their differences between populations. The well-known association between HL-A1 and HL-A8 shows up in the Caucasians but not in the other populations, and is accompanied by a negative association between HL-A1 and HL-A12. Caucasians also show a significant association between HL-A3 and HL-A7. The only association consistent in all three populations is that between HL-A9 and HL-A12. This, however, is probably accounted for by some mistyping

for HL-A12, since several of our sera containing anti-HL-A12 also contain anti-HL-A9 (see W. Bodmer et al. 1969). The relationship between 4c* and HL-A3 is significant only in the Pygmies, where it is negative, though the Δ for this pair of alleles is positive in the American blacks. This is the only example of an association significant in the Pygmies but not in the Caucasians.

Differences between Pygmy Groups

As mentioned in the introduction, the Pygmy blood samples were obtained from three different areas within a radius of 50–100 miles. Each of these groups had their own language and so were presumably relatively isolated from each other. The separate frequencies of the HL-A antigens in these three groups and the χ^2's for testing the significance of the differences are shown in table 5. The overall heterogeneity

TABLE 5

HL-A ANTIGEN FREQUENCIES IN THREE DIFFERENT PYGMY GROUPS

ANTIGEN	BAGANDOU		LOKO		MONGOUMBA		HETEROGENEITY χ^2 (2 df)
	Total	%	Total	%	Total	%	
LA:							
HL-A1	59	0.0	85	0.0	32	0.0	.....
HL-A2	63	30.2	79	25.3	33	12.1	3.8
HL-A3	61	16.4	78	21.8	30	16.7	0.8
HL-A9	42	21.4	75	20.0	18	5.6	2.4
3002Q	39	61.5	60	70.0	11	81.8	3.0
4:							
HL-A12	63	19.0	87	23.0	33	9.1	5.5
4c*	63	20.6	89	19.1	33	3.0	6.3
HL-A7	63	3.2	90	15.6	63	9.5	5.1
HL-A8	63	12.7	90	3.3	33	6.1	1.8
Total (16 df)	.....	.....	.....	.....	.....	.....	28.7†

† $P = .025$.

χ^2 testing the difference is significant at the 2.5% level, indicating that these Pygmy populations are sufficiently isolated from each other to allow for the accumulation of genetic variation, most probably as a result of random genetic drift rather than differential selection.

Several measures of genetic distance based on differences in gene frequencies have been suggested. The most satisfactory so far seems to be that given by Cavalli-Sforza (1969; see also Cavalli-Sforza and Bodmer 1970), based on an analogy with inbreeding coefficients between populations measured in terms of Wahlund's variance in gene frequency. The distance f between two populations based on a locus with k alleles is given by

$$f = 4(1 - \cos\theta)/(k - 1)\,, \tag{2a}$$

where

$$\cos\theta = \sum_{i=1}^{k} \sqrt{p_{i1}p_{i2}}\,, \tag{2b}$$

and p_{i1}, p_{i2} are the respective allele frequencies in the two populations. The f values between the three Pygmy groups, calculated in each case as the average of the two values obtained separately for the LA and 4 series, are shown in table 6. Geographically, Loko lies between Bagandou and Mongoumba but is closer to the former. The genetic distances, measured in terms of f, correspond (at least roughly) to the geographical distances, in particular with respect to the relative similarity between the Bagandou and Loko populations as compared to the difference between each of these and Mongoumba. The level of variation is comparable to that found by Cavalli-Sforza (1969; see also Cavalli-Sforza and Bodmer 1970) between a group of villages

TABLE 6

GENETIC DISTANCES (f) BETWEEN THE THREE PYGMY GROUPS, BASED ON HL-A FREQUENCIES

Pygmy Groups	f
Bagandou-Loko	.014
Bagandou-Mongoumba	.045
Loko-Mongoumba	.031

NOTE.—Genetic distance (f) is the average of the two values calculated separately following Cavalli-Sforza (1969) for the LA and 4 series, assuming they are independent. The formula for a locus with k alleles is $f = 4(1 - \cos\theta)/(k - 1)$, where

$$\cos\theta = \sum_{i=1}^{k} \sqrt{p_{i1}p_{i2}}$$

and p_{i1}, p_{i2} are the respective allele frequencies in the two populations. Thus f is, approximately, the "Wahlund" variance of gene frequencies between the populations.

in the upper Parma valley in northern Italy, and is certainly compatible with a limited amount of migration between the groups.

DISCUSSION

In order to test for new and different antigens in Pygmies, either belonging to the HL-A or to some other system, we tested the sera from over a hundred multiparous Pygmy women compared to both their own relatives and random American blacks and whites. The term "Pygmy family" is not used, since hardly any of the putative families turned out to be legitimate families in the conventional sense, as shown by absence of expected patterns of inheritance for white and red cell types and for isozyme variants (Cavalli-Sforza, personal communication). In any case, only two out of 100 proved to contain leucocyte antibodies. One reacted with 2% of the Pygmies, 3% Bantu, no black Americans, and 68% whites. The other had frequencies of 23%, 35%, 25%, and 89% in these respective groups. The antibodies were not positively identified and their reactions were relatively weak, but it is clear from the frequencies that they did not define a predominantly "Pygmy" or even "black" antigen. The question arose why so few of the sera contained antibody, when it has been found that about 20% of American women produce antibody after two or more children

(Payne and Rolfs 1958; Terasaki and McClelland 1964). Further studies are needed on this point. It may be that the presence of filariae and other parasites in the blood inhibit the production of antibody. Filariae are seen microscopically in nearly every sample of leucocytes tested from the Pygmies. These and other parasites are presumably responsible for the increased number of eosinophils found in our samples. The eosinophils were not killed off by prolonged storage and freezing as were the granulocytes. They also fluoresced brightly in many of the preparations and were not killed by the leucocyte antibodies (W. Bodmer and Gerbrandt 1968). They therefore complicated the reading of the assay but did not significantly interfere with the results, since their much larger size made them easily distinguishable.

Another problem which arose is one often encountered in population studies of red cell blood groups and already mentioned above. That is that sera tested and found to contain certain specificities in one population sometimes show different specificities in another population. One possible explanation for this is the presence in the serum of an extra antibody defining an antigen which is comparatively rare in the first population but more frequent in the second. A second reason might be that the overall distribution of antigens is sufficiently different in two populations for different patterns of cross reactivity to be expressed (see also Ivanyi and Dausset 1966). The existence of and complicating effects of cross reactivity are becoming increasingly evident (see, e.g., Svejgaard and Kissmeyer-Nielsen 1968; W. Bodmer et al. 1969; Thorsby and Kissmeyer-Nielsen 1969), and are further suggested by the possible production in a serum of a specificity which may not present in the immunizing cell donor (Ceppellini, personal communication; J. Bodmer et al. 1970). As the behavior of these cross reactions has not yet been fitted into a predictable pattern, this is still a matter for speculation. It is clear, however, that apparent demonstration of monospecificity by absorption or matching of donor-recipient pairs for serum production in one population is not necessarily going to have validity in another population. Therefore, as already mentioned, our criteria for assigning antigens became much stricter when faced with apparently discrepant serum reactions in the Pygmies. We did not in general assume that any single serum was monospecific and capable of defining an antigen by itself, and required agreement of two or more sera in order to be reasonably sure of the assignment. This is also partly the reason for the repeated testing of the Pygmy cells.

There have been a few previous studies on the distribution of the HL-A antigens in other than Caucasian populations (Dausset et al. 1965, 1967; Rubinstein et al. 1967; Gabb 1968; Singal et al. 1969). Dausset et al. (1965) compared the distribution among 29 Japanese and 27 Malian blacks with those of the French. Their main findings were: (1) a lower frequency of HL-A2 in Malians than in the French; (2) a lower frequency of their antigen 4 (now HL-A12) in Japanese and Malians; and (3) the absence of HL-A8 and a low frequency of HL-A7 in Japanese. These studies were extended by Dausset et al. (1967) to include a sample of American blacks and Orientals from New York, with essentially similar results. Recently, Singal et al. (1969) have published more comprehensive data on a sample of 94 Japanese, 89 blacks, and 233 whites from the Los Angeles area. Their data confirm the low frequency of HL-A7 and HL-A8 in Japanese, and the lower frequency of HL-A2 in

American blacks. They also showed that HL-A1 was absent from Japanese and present in a relatively low frequency in American blacks.

Rubinstein et al. (1967) tested predominantly for the 4 (or 7) series antigens. They found relatively low frequencies of 4a and 4b, indicating a higher blank allele frequency for the 4 series in their Indian population than in Caucasians. They also report a remarkably high frequency for HL-A8 and changed associations between other 4 series antigens (now known not to be "pure"), probably accounted for by differences in the frequencies of their constituent components.

Gabb (1968) found that HL-A1 was absent from 67 Australian aborigines tested and present at a low frequency in New Zealand Maori (gene frequency .057 based on 128 individuals); the low frequency could be explained by Caucasian admixture. He also found that HL-A9 gene frequencies were higher in Australian aborigines and New Zealand Maori than in Australian whites (respective gene frequencies .27, .34,

TABLE 7

GENETIC DISTANCES (f VALUES) BETWEEN AFRICAN, ASIAN, AND CAUCASIAN POPULATIONS BASED ON HL-A FREQUENCIES

Data	f
This paper:	
Pygmy-American black	.038
Pygmy-American white	.102
American black-American white	.022
Singal et al. (1969):*	
American black-American white	.023
Japanese-American white	.099
Japanese-American black	.080

* The HL-A gene frequencies were calculated by maximum-likelihood fitting from their data. See the legend to table 6 for the method of calculating f values.

and .1). The respective HL-A2 gene frequencies in these three populations were .28, .17, and .31. The high HL-A9 frequencies in aborigines and Maori are comparable to those found by Singal et al. (1969) in the Japanese.

We have recently found that HL-A1 is also absent from a population of American Indians in Guatemala (unpublished results from a cooperative study with Drs. Cann and Barnett of Stanford). The antigen HL-A1 may well be unique to the Caucasian populations. Our data on American blacks is generally in good agreement with the frequencies given by Dausset et al. (1965, 1967) and Singal et al. (1969). The main differences lie in the frequencies of the antigen HL-A12 which was not, perhaps, as well defined as the others. Singal et al. (1969) also report a relatively high frequency for HL-A7 in American blacks.

Overall differences in HL-A frequencies between the various populations can best be summarized in terms of genetic distances (f) as in our discussion of the differences between the three Pygmy groups. Some f values between Pygmies, American blacks, whites, and Japanese based on our data and those of Singal et al. (1969) are given in table 7. As expected, the Pygmies are much closer to the American blacks (and also,

of course, to the Bantu) than they are to Caucasians. In fact, as pointed out earlier, they show in most cases a more pronounced difference from the Caucasians than do the Bantus and American blacks, and in the same direction as these latter groups. This is consistent with the suggestion of Cavalli-Sforza et al. (1969) that these Pygmies may be a "proto-African" race. There is very good agreement between the f values for the American black-American white difference given by the two independent bodies of data.

An f of approximately .02 is relatively low for the difference between two disparate racial groups. Gene frequencies in American blacks and whites for a number of genetic polymorphisms were summarized by Workman et al. 1963. Using their data, and excluding polymorphisms known or suspected to be associated with resistance to malaria (G6PD and hemoglobin S), the f value between American blacks and whites lies between .13 and .16. Values for individual alleles, however, range from about .8 for $R_o(cDe)$ of the Rhesus system and .47 for the Duffy allele Fy^a, to .006 for $R_2(cDE)$ of the Rhesus system and M of the MN system, and .024 for A of the ABO system and Hp_1 (of haptoglobins). The HL-A value of .02 is clearly on the low side of the spectrum.

All the f values obtained with the HL-A system are substantially lower than the average obtained for other polymorphisms. Thus Cavalli-Sforza (1969) gives the following f values based on averages over 16 loci: African-Caucasoid, .352; African-Eastern, .416; Caucasoid-Eastern, .242. Values of f, averaged over a large representative sample of human populations calculated directly as a Wahlund's variance, were also given separately for a number of polymorphisms by Cavalli-Sforza (1966). They range from .029 for the Kell blood groups to .382 for the $R_o(cDe)$ allele of the Rh system, with the ABO, MNSs, Diego, and haptoglobin systems lying in the range .07 to .1. These averages are, of course, expected to be lower than the f values between pairs of major racial groups. Relatively low values of f for a polymorphism imply a more restricted range of variation and so stabilizing or convergent (as opposed to differential) selection with respect to the polymorphism. Our data on HL-A frequencies, especially the relatively small difference between American blacks and whites, suggest therefore, relatively uniform stabilizing selection with respect to the HL-A polymorphism. More data on the distribution of the HL-A polymorphism in various populations, and especially in African populations to allow comparison among Africans, American blacks, and Caucasians, should help to show whether this important and highly variable polymorphism is, in fact, subject to relatively strong selection pressures.

SUMMARY

The history of the HL-A polymorphism and its present description in terms of two major multiple allelic series, LA and 4, is briefly reviewed. Data on the antigen and gene frequencies of HL-A antigens HL-A1, 2, 3, 7, 8, 9, and 4c*, HL-A12, and LA-W + Ba* for a Pygmy population from West Africa are compared with data on Bantu, American black, and American white populations. The American black and Bantu populations do not differ significantly, while all other pairs of populations show highly significant overall differences in gene frequencies. There are relatively

few significant gametic associations between alleles of the LA and 4 series. Genetic differences between the three groups of Pygmies typed correspond roughly to their geographical relationships. An analysis of the genetic distances between Pygmy, Bantu, American black, and Caucasian populations, based on the HL-A data, confirms the relationships between Pygmies and Bantu and is consistent with the Pygmies being a proto-African race. Genetic differences between the major African, Caucasian, and Oriental racial groups measured in terms of HL-A frequencies seem to be relatively small and indicate the possible existence of convergent, as opposed to differential, selection with respect to the HL-A polymorphism.

ACKNOWLEDGMENTS

We are very grateful to Professor L. L. Cavalli-Sforza of the Istituto di Genètica, Università di Pavia, for allowing us to accompany him on his expedition to study the Babinga Pygmies, which was under the auspices of the International Biological Program. The expedition was supported, in part, by the Consiglio Nazionale della Ricerche, Italy, the Wenner-Gren Foundation, and the World Health Organization. We should also like to thank Dr. Roy Walford and the NIH Serum Bank for their gifts of sera. We are grateful to Juliana Hwang for her help with the computer programs, and to Michelle Morris for her help with the data reduction. We are also grateful to Millie Tripp and Grace Gerbrandt for their technical assistance.

APPENDIX

CALCULATION OF THE STANDARD ERROR OF Δ

Given observations in a 2 × 2 table for a pair of antigens in the form

		1st antigen	
		+	−
2nd antigen	+	a	b
	−	c	d

the gametic association Δ between their respective genetic determinants is estimated by

$$\Delta = \sqrt{\frac{d}{n}} - \sqrt{\left(\frac{b+d}{n}\right)\left(\frac{c+d}{n}\right)}, \tag{A1}$$

where $n = a + b + c + d$ (see Ceppellini 1967; W. Bodmer et al. 1969). Fisher (see Bailey 1961, Appendix 2) derived a general formula for the approximate variance of a function of observed frequencies, such as that given by equation (A1). His formula can be expressed in the form

$$V(T) \approx \sum_i \left(\frac{\partial T}{\partial a_i}\right)^2 a_i - n\left(\frac{\partial T}{\partial n}\right)^2, \tag{A2}$$

where T is the function, a_i the observations, n the total, and where the partial differential coefficients are evaluated at the observed values a_i. Applying formula (A2) to (A1) gives

$$V(\Delta) \approx \frac{1}{4n} + \frac{1}{4n^2}(b + c + 3d) - \frac{d^{1/2}}{2n^{3/2}}(x^{1/2} + x^{-1/2}) - \frac{y}{n^3} + \frac{(yd)^{1/2}}{n^{5/2}}, \tag{A3}$$

where $x = (c + d)/(b + d)$ and $y = (b + d)(c + d)$.

A slightly simpler, more approximate result can be obtained as follows. It is well known (see e.g., W. Bodmer et al. 1969) that the 2×2 χ^2 and the binary correlation coefficient r, which is a measure of the association in a 2×2 table, are related by the equation

$$nr^2 = \chi^2 . \tag{A4}$$

Thus, since $r^2/V(r) \approx \chi^2$, we have

$$V(r) \approx \frac{1}{n} . \tag{A5}$$

Now, from the expression for r given by W. Bodmer and Payne (1965), it follows that

$$r = n\Delta[2(b+d)^{1/2}(c+d)^{1/2} + n\Delta]/\sqrt{(a+b)(a+c)(b+d)(c+d)} , \tag{A6a}$$

so that

$$n^2\Delta^2 + 2\Delta n\sqrt{(b+d)(c+d)} - r\sqrt{(a+b)(a+c)(b+d)(c+d)} = 0 . \tag{A6b}$$

Solving for Δ, assuming r is small, gives

$$\Delta \approx \frac{r}{2n}\sqrt{(a+b)(a+c)} . \tag{A7}$$

Therefore

$$V(\Delta) \approx \frac{(a+b)(a+c)}{4n^2} ; \qquad V(r) = \frac{1}{4n^3}(a+b)(a+c) , \tag{A8}$$

from equation (A5).

REFERENCES

BAILEY, N. T. J. 1961. *Mathematical theory of genetic linkage.* Clarendon Press, Oxford.

BALNER, H.; CLETON, F. J.; and EERNISSE, J. G., eds. *Histocompatibility testing, 1965.* Series Haemat. 11. Munksgaard, Copenhagen.

BODMER, W.; BODMER, J.; ADLER, S.; PAYNE, R.; and BIALEK, J. 1966. Genetics of 4 and LA human leukocyte groups. *Ann. N.Y. Acad. Sci.* **129**:473–489.

BODMER, W.; BODMER, J.; IHDE, D.; and ADLER, S. 1969. Genetic and serological association analysis of the HL-A leukocyte system. Pp. 117–127 in N. E. Morton (ed.), *Computer applications in genetics.* Honolulu, Univ. Hawaii Press.

BODMER, J.; COUKELL, A.; BODMER, W.; PAYNE, R.; and SHANBROM, E. 1970. A new allele for the LA series of HL-A antigens: the analysis of a complex serum. In P. Terasaki (ed.), *Histocompatibility testing, 1970.* Munksgaard, Copenhagen (in press).

BODMER, W. F., and GERBRANDT, G. 1968. Short term room temperature storage of human lymphocytes for white cell typing. *Vox Sang.* **14**:451–455.

BODMER, W. F., and PAYNE, R. 1965. Theoretical consideration of leukocyte grouping using multispecific sera. Pp. 141–149 in H. Balner, F. J. Cleton, and J. G. Eernisse (eds.), *Histocompatibility testing, 1965.* Series Haemat. 11. Munksgaard, Copenhagen.

BODMER, W. F.; TRIPP, M.; and BODMER, J. 1967. Application of a fluorochromatic cytotoxicity assay to human leukocyte typing. Pp. 341–350 in E. S. Curtoni, P. L. Mattiuz, and R. M. Tosi (eds.), *Histocompatibility testing, 1967.* Munksgaard, Copenhagen.

CAVALLI-SFORZA, L. L. 1966. Population structure and human evolution. *Proc. Roy. Soc. (London)*, B **164**:362–379.

CAVALLI-SFORZA, L. L. 1969. Human diversity. *Proc. 12th Int. Cong. Genet.* **3**:405–416.

CAVALLI-SFORZA, L. L., and BODMER, W. F. 1970. *The genetics of human populations.* Freeman, San Francisco (in press).

CAVALLI-SFORZA, L. L.; ZONTA, L. A.; NUZZO, F.; BERNINI, L.; DE JONG, W. W. W.; MEERA KHAN, P.; RAY, A. K.; WENT, L. N.; SINISCALCO, M.; NIJENHIUS, L. E.; VAN LOGHEM, E.; and MODIANO, G. 1969. Studies on African pygmies. I. A pilot investigation of Babinga pygmies in the Central African Republic (with an analysis of genetic distances). *Amer. J. Hum. Genet.* **21**:252–274.

CEPPELLINI, R. 1967. Genètica delle immunoglobuline. Atti 12th Riunione Ass. Genet. Italiana, Parma, October 28–30, 1966. *A.G.I.* **12**:3.

CEPPELLINI, R.; CURTONI, E. S.; MATTIUZ, P. L.; MIGGIANO, V.; SCUDELLER, G.; and SERRA, A. 1967. Genetics of leukocyte antigens. A family study of segregation and linkage. P. 149 in E. S. Curtoni, P. L. Mattiuz, and R. M. Tosi (eds.), *Histocompatibility testing, 1967.* Munksgaard, Copenhagen.

CURTONI, E. S.; MATTIUZ, P. L.; and TOSI, R. M., eds. 1967. *Histocompatibility testing, 1967.* Munksgaard, Copenhagen.

DAUSSET, J.; COLOMBANI, J.; COLOMBANI, M.; LEGRAND, L.; and FEINGOLD, N. 1968*a*. Un nouvel antigène du système HL-A (Hu-1): l'antigène 15, allele possible des antigènes 1, 11, 12. *Nouv. Rev. Franc. Hemat.* **8**:398–406.

DAUSSET, J.; COLOMBANI, J.; LEGRAND, L.; and FEINGOLD, N. 1968*b*. Le deuzième sublocus du système HL-A, *Nouv. Rev. Franc. Hemat.*, **8**:841–846.

DAUSSET, J.; IVANYI, P.; COLOMBANI, J.; FEINGOLD, N.; and LEGRAND L. 1967. The Hu-1 system. P. 189 in E. S. Curtoni, P. L. Mattiuz, and R. M. Tosi (eds), *Histocompatibility testing, 1967.* Munksgaard, Copenhagen.

DAUSSET, J.; IVANYI, P.; and IVANYI, D. 1965. Tissue alloantigens in humans: identification of a complex system (Hu-1). P. 41 in H. Balner, F. J. Cleton, and J. G. Eernisse (eds.), *Histocompatibility testing, 1965.* Series Haemat. 11. Munksgaard, Copenhagen.

GABB, B. W. 1968. Complement fixing platelet antigens and antibodies in man. Unpublished Ph.D. thesis. Univ. Adelaide, Australia.

IVANYI, P., and DAUSSET, J. 1966. Allo-antigens and antigenic factors of human leukocytes. A hypothesis. *Vox Sang.* **11**:326–331.

KISSMEYER-NIELSEN, F.; SVEJGAARD, A.; and HAUGE, M. 1968. Genetics of the human HL-A transplantation system. *Nature* **29**:1116–1119.

MI, M. P., and MORTON, N. E. 1966. Blood factor association. *Vox Sang.* **11**:434–449.

MICKEY, M. R.; SINGAL, D. P.; and TERASAKI, P. I. 1969. Serotyping for homotransplantation XXV. Evidence for three HL-A subloci. *Transplantation Proc.* **1**:347–351.

MIGGIANO, V. C.; NABHOLZ, M.; and BODMER, W. F. 1970. Detection of HL-A and other antigens on fibroblast micro monolayers using a fluorochromatic cytotoxicity assay. In P. Terasaki (ed.), *Histocompatibility testing, 1970.* Munksgaard, Copenhagen (in press).

PAYNE, R.; BODMER, J.; BODMER, W.; and SHANBROM, E. 1970. Characterization of HL-A antisera produced by planned immunization. In P. Terasaki (ed.), *Histocompatibility testing, 1970.* Munksgaard, Copenhagen (in press).

PAYNE, R., and ROLFS, M. R. 1958. Fetomaternal leukocyte incompatibility. *J. Clin. Invest.* **37**:1756–1763.

PAYNE, R.; TRIPP, M.; WEIGLE, J.; BODMER, W.; and BODMER, J. 1964. A new leukocyte isoantigen system in man. *Cold Spring Harbor Sympos. Quant. Biol.* **29**:285.

ROOD, J. J. VAN. 1962. Leukocyte grouping. A method and its application. Unpublished doctoral thesis, Leiden.

ROOD, J. J. VAN; LEEUWEN, A. VAN; SCHIPPERS, A. M. J.; VOOYS, W. H.; FREDERIKS, E.; BALNER, H.; and EERNISSE, J. G. 1965. Leukocyte groups, the normal lymphocyte transfer test and homograft sensitivity. Pp. 37–50 in H. Balner, F. J. Cleton, and J. G. Eernisse (eds.), *Histocompatibility testing, 1965.* Munksgaard, Copenhagen.

RUBENSTEIN, P.; COSTA, R.; LEEUWEN, A. VAN; and ROOD, J. J. VAN. 1967. The leukocyte antigens of Mapuche Indians. Pp. 251–255 in E. S. Curtoni, P. L. Mattiuz, and R. M. Tosi (eds.), *Histocompatibility testing, 1967.* Munksgaard, Copenhagen.

SINGAL, D. P.; MICKEY, M. R.; and TERASAKI, P. I. 1969. Serotyping for homotransplantation. *Transplantation:* **7**:246–258.

SVEJGAARD, A., and KISSMEYER-NIELSEN, F. 1968. Cross-reactive human HL-A isoantibodies. *Nature* **219**:868–869.

TERASAKI, P., ed. 1970. *Histocompatibility testing, 1970.* Munksgaard, Copenhagen (in press).

TERASAKI, P. I., and MCCLELLAND, J. D. 1964. Microdroplet assay of human serum cytotoxins. *Nature* **204**:998–1000.

THORSBY, E. 1969. HL-A antigens and genes. II. Studies of the antigen MH (=HL-A5) of the 4 sub-locus. *Vox Sang.* **17**:93–101.

THORSBY, E., and KISSMEYER-NIELSEN, F. 1969. HL-A antigens and genes. III. Production of HL-A typing antisera of desired specificity. *Vox Sang.* **17**:102–111.

TURNBULL, C. M. 1965. *Wayward servants.* Natural History Press, Garden City, N.Y.

WALFORD, R. L.; WALLACE, O.; SHANBROM, E.; and TROUP, G. M. 1968. Lc-11 (Hunt B. Jones) as a mutually exclusive specificity to Lc-1, 2 and 3 in the main human leukocyte group. *Vox Sang.* **14**:338.

WORKMAN, P. L.; BLUMBERG, B. S.; and COOPER, A. J. 1963. Selection, gene migration and polymorphic stability in a U.S. white and Negro population. *Amer. J. Hum. Genet.* **15**:429–437.

WORLD HEALTH ORGANIZATION. 1968. Nomenclature for factors of the HL-A system. *Bull. WHO.* **39**:483–486.

WORLD HEALTH ORGANIZATION. 1970 (in press).