Am. J. Hum. Genet. 50:1308–1317, 1992

Segregation Analysis Detects a Major Gene Controlling Blood Infection Levels in Human Malaria

Laurent Abel,* Michel Cot,†,‡ Lambertus Mulder,§ Pierre Carnevale,‡ and Josué Feingold†

*INSERM Unité 194, Département de Biomathématiques, Hôpital Pitié-Salpêtrière; and †INSERM Unité 155, Unité de Recherches en Gènètique Epidemiologique, Paris; ‡ORSTOM, OCEAC, Yaoundé, Cameroon; and §Department of Parasitology, Catholic University of Nijmegen, and University of Wageningen, Nijmegen, The Netherlands

Summary

The profound influence that the genetic makeup of the host has on resistance to malaria infection has been established in numerous animal studies. This genetic heterogeneity is one of the main causes of the difficulties in developing an effective malaria vaccine. Segregation analysis is the first step in identifying the nature of genetic factors involved in the expression of human complex diseases, as infectious diseases. To assess the role of host genes in human malaria, we performed segregation analysis of blood parasite densities in 42 Cameroonian families by using both the unified mixed model and the class D regressive model of analysis. The results provide clear evidence for the presence of a recessive major gene controlling the degree of infection in human malaria. Parameter estimates show a frequency of .44–.48 for the deleterious allele, indicating that about 21% of the population is predisposed to high levels of infection.

Introduction

Host and parasite genetic factors have been demonstrated to influence the outcome of malaria infections in experimental animals (Stevenson et al. 1982; Stevenson and Skamene 1985; Sayles and Wassom 1988; Wunderlich et al. 1988). Certain genetic disorders of the red cell (abnormal hemoglobins, glucose-6-phosphate dehydrogenase deficiency, and absence of Duffy blood group antigen) are known to influence the resistance against malaria infection in humans (Weatherall 1987; Nagel and Roth 1989). However, even though these genetic red cell variants have achieved polymorphic frequencies in many populations, their role in the individual degree of protection may be small, especially among adults (Weatherall 1987). Some associations between HLA genes and malaria infection have also been reported at the population level (Osoba et al. 1979), and a recent case-control study showed that common west African HLA antigens were associated with protection from severe malaria (Hill et al. 1991). However, thus far, few familial genetic studies on human resistance to malaria infection have been carried out.

Malaria is undergoing a worldwide resurgence, mainly because of the spread of drug-resistant parasite strains, and malaria vaccines are being developed to complement traditional control measures (Miller et al. 1986). However, major difficulties arise from the genetic restriction of the immune response to subunit vaccine candidates (Grau et al. 1987; Good et al. 1988; Weiss et al. 1989; Londono et al. 1990), and one might expect that a better knowledge about host genes regulating the response to malaria infection would improve the vaccine research strategy.

The aim of the present study is to investigate the role of a major gene determining human susceptibility/resistance to malaria infection and to do so by using complex segregation analysis. Segregation analysis is the first step in the effort to determine, from family data, the mode of inheritance of a complex trait, and we have followed in this work a strategy similar to the one used to study the genetic control of susceptibility/

Received November 12, 1991; revision received January 24, 1992.

Address for correspondence and reprints: Dr. Laurent Abel, INSERM Unité 194, Département de Biomathématiques, Hôpital de la Pitié-Salpêtrière, 91 Boulevard de l'Hopital, 75013 Paris, France.

0002-9297/92/5006-0020$02.00

resistance to infection by *Schistosoma mansoni* (Abel et al. 1991). This study was conducted on 42 nuclear families from Cameroon by using two different models of analysis—the unified mixed model and a regressive model.

Methods

Family Data and Measures

A family study on malaria was carried out from January 1988 to July 1990 in the Bilalang district of Edea, an industrial town in southwest Cameroon. This district was constructed by the aluminum factory ALUCAM for its employees and was chosen for this study for two main reasons. First, nuclear families originating mainly from south Cameroon were settled in Bilalang district by the factory more than 10 years ago and were comparable in terms of socioeconomic levels. Second, the ALUCAM medical service is well developed and efficient, so that the follow up of these families is comparable to that observed in industrialized countries. Therefore, one might expect the reliability in these subjects' responses concerning antimalarial prophylaxis habits to be greater than that in other African populations. The ascertainment scheme was complete selection of the 42 nuclear families living in three subdistricts of Bilalang and totaling 285 persons. The families were questioned about their habits concerning chemoprophylaxis intake and were classified into three groups: (*a*) no intake (15 families), (*b*) irregular intake (14 families), and (*c*) regular intake (13 families). However, the dose of chloroquine taken by the members of groups *b* and *c* was generally lower than the one recommended by the WHO (300 mg/wk for adults, 5 mg/kg/wk for children).

During the study, each family was visited 10 times, and blood samples were taken from family members present (table 1). Each person was not present at each of the 10 visits, and the mean number of samples per subject was 5.2 (range 2–10). Determination of parasite density (PD) was based on the parasite/leukocyte ratio (Trape 1985), by counting 500 leukocytes on a Giemsa-stained thick smear; more than 95% of identified parasites were *Plasmodium falciparum*. The proportion of persons who have been infected, i.e., who had at least one positive PD, was 81%. For further analyses, a logarithmic transformation, based on log(PD + 1), was applied to PDs, to allow for zero counts; the log-transformed PDs will be denoted "LPDs."

In order to obtain a unique variable accounting for the degree of malaria infection, a mean PD (MPD1)

Table 1

Parasite Densities from January 1988 to July 1990 and Standardized MPD1 and MPD3 Values for Two Families Living in the Same Subdistrict

Family and sex of Individual (age[a] in years)	Elementary Parasite Densities Measured at Each Visit[b]											
	88/1	88/3	88/5	88/7	88/10	89/1	89/6	89/11	90/3	90/7	MPD1[c]	MPD3[d]
1:												
M (35)	. . .	. . .	1,600	. . .	. . .	100	. . .	0	. . .	. . .	.81	1.46
F (32)	. . .	0	0	9,200	2,000	100	. . .	0	. . .	750	.88	1.39
M (11)	. . .	3,200	4,500	14,400	10,000	6,000	. . .	65	1,600	5,000	2.75	2.71
F (7)	. . .	0	20,000	68,000	1,500	3,000	. . .	0	850	1,000	1.92	1.82
F (5)	. . .	1,400	0	4,800	1,000	16,000	. . .	10	750	600	1.77	1.73
M (4)	. . .	0	300	50,000	1,500	2,000	. . .	500	3,500	8,000	2.20	2.25
5:												
M (37)	0	0	1,200	0	0	. . .	. . .	. . .	. . .	0	−.55	.02
F (26)	0	0	0	0	0	. . .	. . .	0	. . .	0	−1.09	−1.13
F (11)	0	0	0	0	0	0	. . .	. . .	0	0	−1.10	−1.62
M (9)	0	0	0	100	300	0	. . .	. . .	120	0	−.20	−.60
M (8)	0	0	0	0	0	0	. . .	. . .	100	0	−.83	−1.29
F (7)	100	0	0	100	0	100	. . .	. . .	0	30	−.07	−.41
F (3)	0	0	1,500	40,000	1,000	100	. . .	0	0	0	.47	.40

[a] By April 1989.

[b] An ellipsis (. . .) indicates that the person was not present at the time of the visit.

[c] Determination as described in Methods.

[d] Final value used for segregation analysis by the unified mixed model, corrected for both location and age effects.

was determined. One-way analysis of variance showed that the mean LPDs varied across the visits ($P < .0001$). As shown in figure 1, the lowest infection levels were observed during the dry season (February–June), when mosquito transmission was low. Therefore, the LPDs were first corrected for the visit effect by substracting from each individual LPD the mean LPD of the corresponding visit. The MPD1 was then computed for each subject, as the mean of his visit-adjusted LPDs (table 1). No effect of the number of measurements per subject on the MPD1 values was detected ($P > .70$).

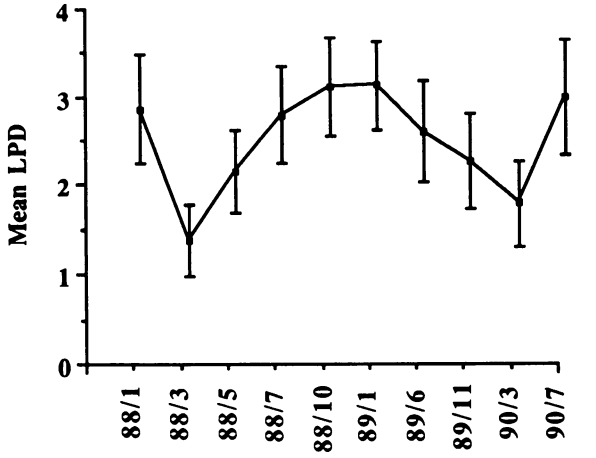

Figure 1 Influence that the visit had on the LPDs. Results are presented as the mean, with 95% confidence interval, of the LPD values observed during the 10 visits from January 1988 to July 1990.

Data Adjustment

Prior to segregation analysis we studied the effect that sex, prophylaxis intake (three groups), area of residence (with three subdistricts coded 1 [$n = 137$], 2 [$n = 82$], and 3 [$n = 66$]), and age categorized into nine classes (0–4, 5–8, 9–12, 13–16, 17–20, 21–30, 31–40, 41–50, and 50–60 years had on the MPD1 values, using analysis of variance. Correction for the place of residence was performed by substracting from each subject MPD1 value the mean MPD1 observed in the subdistrict where he or she lives. Two different strategies were then employed, according to the model of analysis. Under the regressive model, the studied phenotype was MPD1 adjusted for the location effect (denoted "MPD2"), and age was considered as a covariate influencing MPD2 during segregation analysis. Under the unified mixed model, MPD2 values were corrected for the age effect prior to segregation analysis, by means of forward stepwise polynomial regression using the SAS software. The MPD2 values adjusted on age (denoted "MPD3") were standardized and used for commingling and segregation analysis. The different steps of the adjustment for the location and for the age effect are illustrated with the data of four families presented in figure 2.

Test for a Mixture of Distributions

Evidence for a mixture of normal distributions accounting for the MPD3 values is consistent with a major-gene hypothesis but can be confounded by skewness in the sample distribution (MacLean et al. 1976; Demenais et al. 1986). Therefore, the presence of a mixture of as many as three normal distributions can be tested, while correction for residual skewness can be done by means of a classical power transformation (MacLean et al. 1976). Maximum-likelihood estimates of the relevant parameters were determined using the computer program SKUMIX (MacLean et al. 1976; Morton et al. 1983), and the likelihood-ratio criterion (Morton et al. 1983) was used to test hypotheses.

Segregation Analysis

In the above method, all individuals are considered independent, and tests for a major-gene effect require that the familial dependence of the phenotypes be accounted for; these tests are performed by complex segregation analysis. Two different models of segregation analysis were used in this study—the unified mixed model (Lalouel et al. 1983) and a regressive model (Bonney 1984).

The unified mixed model assumes that the phenotype adjusted for measured covariates (i.e., MPD3) results from the additive and independent contributions of a major transmissible effect, a multifactorial transmissible component, and a random nontransmitted environmental effect. Under a genetic hypothesis, the major effect results from the segregation of two alleles (A and a) at a single locus; the frequency of allele A is denoted q. Within each genotype g (g is aa, Aa, or AA), the distribution of the phenotype is assumed to be normal, with mean μ_g and variance σ^2. Transmission at the major locus is parametrized in terms of τ_{AAA}, τ_{AaA}, and τ_{aaA}, which denote the probability of transmitting allele A for genotypes AA, Aa, and aa, respectively. Mendelian transmission corresponds to $\tau_{AAA} = 1$, $\tau_{AaA} = .5$, and $\tau_{aaA} = 0$; no parent-offspring transmission of the major effect is represented by $\tau_{AAA} = \tau_{AaA} = \tau_{aaA}$. Both the nonrejection

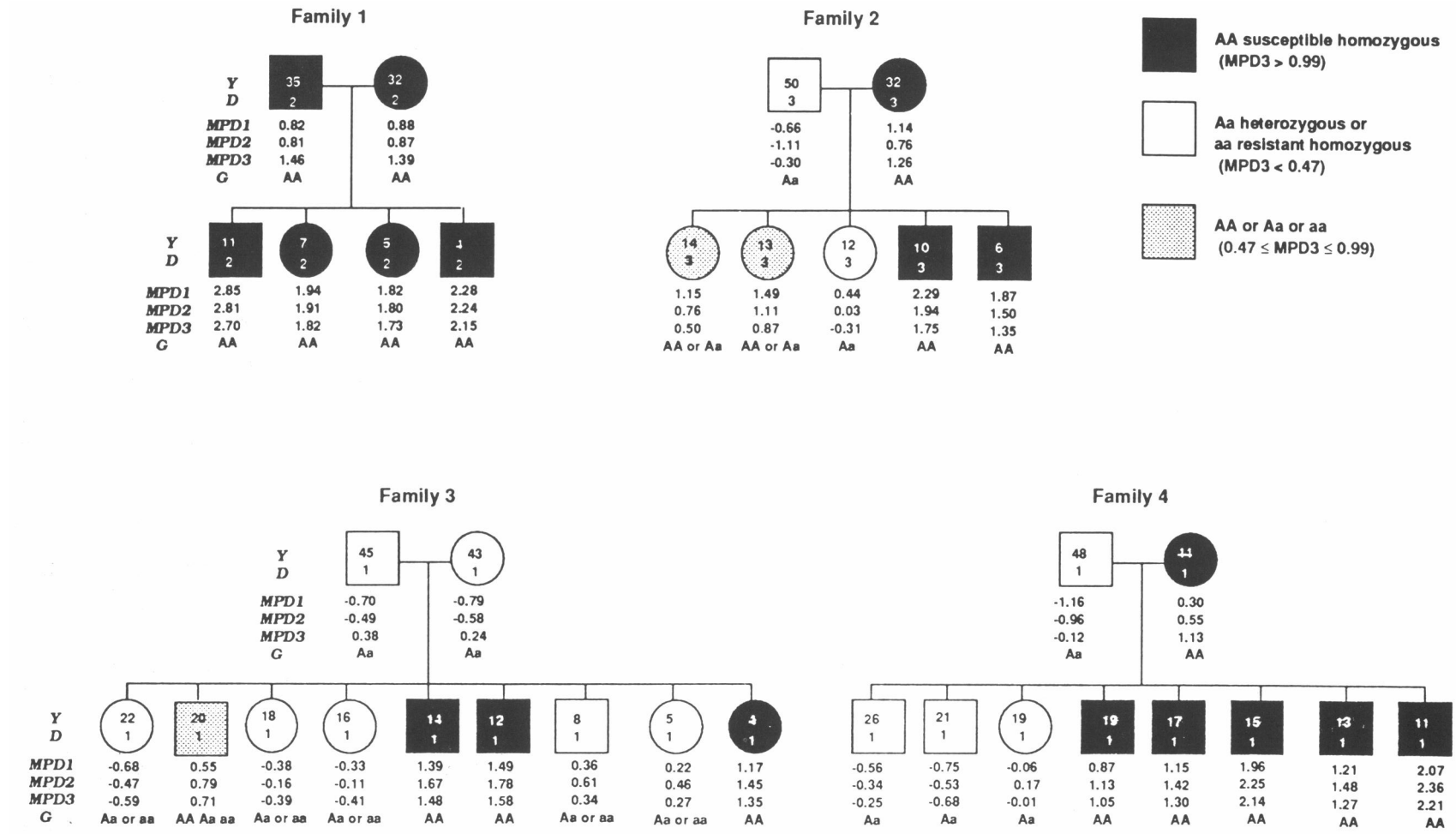

Figure 2 Four nuclear families, indicating, for each person, age in years (*Y*); district of residence (*D*); standardized MPD1, MPD2, and MPD3 values; and most likely genotype according to MPD3 value and Mendelian transmission laws, under the hypothesis of a recessive major gene (A = deleterious allele). The MPD3 value of .99 corresponds to an AA genotype probability greater than .75, and the bound of .47 corresponds to an AA probability less than .25.

of the Mendelian hypothesis and the rejection of the hypothesis of nontransmission of the major effect are required to conclude that the major effect is actually due to a major gene (Demenais et al. 1986). Multifactorial transmission, accounting for both polygenic and environmental effects common to the family, is parametrized in terms of H and HZ (multifactorial heritability in children and adults, respectively). Under each model defined by fixing the appropriate parameters, the joint likelihood of parents and offspring was maximized by means of the computer program POINTER (Lalouel and Morton 1981; Morton et al. 1983).

Regressive models specify a regression relationship between the phenotype of an individual (i.e., MPD2) and (1) a major gene effect, (2) the phenotype of his preceding relatives, and (3) other explanatory measured variables or covariates (Bonney 1984). The major-gene effect is specified in the same way as in the unified mixed model, with identical parameters.

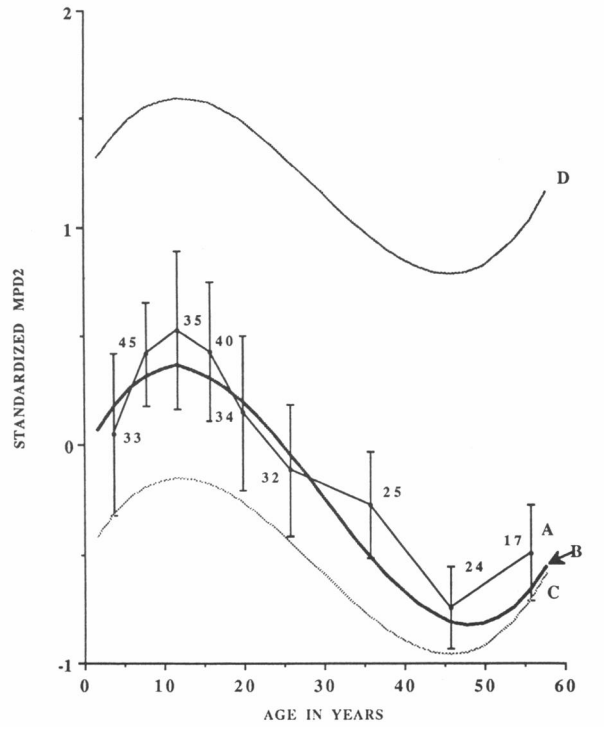

Figure 3 Influence that age has on MPD2s. Curve A represents the mean, with 95% confidence interval, of MPD2 values observed within nine age groups (i.e., 0–4, 5–8, 9–12, 13–16, 17–20, 21–30, 31–40, 41–50, and 51–60 years), with the corresponding number of subjects in each group. Curve B is the distribution of MPD2s that is predicted by the polynomial regression in age. Curves C and D represent the evolution with age, of the MPD2 mean for Aa- or aa-resistant (curve C) and AA-susceptible (curve D) persons, under recessive model 2b in table 3.

Different patterns of dependence between a person and preceding relatives can be considered and are simply expressed in terms of phenotypic correlations. In the class D model (Bonney 1984, 1986), used in this analysis, these correlations are parametrized in terms of ρ_{FM} (the father-mother [or spouse] correlation), ρ_{PO} (the parent-offspring correlation), and ρ_{SS} (the sib-sib correlation). In the absence of a covariate, the class D regressive model is equivalent to the unified mixed model in the particular case where $H = HZ$, $\rho_{FM} = 0$, and $\rho_{PO} = \rho_{SS}$, with the following correspondence: $\rho_{PO} = \rho_{SS} = H/2$ (Demenais and Bonney 1989). Age was introduced into this model as a covariate, so that the parameters of the polynomial regression in age were estimated simultaneously with the genetic and familial correlation parameters. Computations were performed using the REGC program of the software package SAGE (Elston et al. 1986).

All hypotheses were tested by means of the likelihood-ratio criterion (Morton et al. 1983). For example, under the mixed model including the major effect with the multifactorial component (model 3 in table 2), evidence for a major effect is obtained by rejecting the multifactorial model (model 5a); this test will be denoted "model 5a versus model 3," where 5a represents the reduced model and where 3 represents the complete model. As the 42 families were not chosen for any particular reason (i.e., there was random selection), there was no need for any ascertainment correction.

Results

Influence of Sex, Drug Intake, Area of Residence, and Age on MPD1 Values

There was clearly no evidence for a difference in the MPD1 values by sex ($P > .80$) and chloroquine intake ($P > .87$). The low dose of chloroquine taken by the persons from families (b) and (c) and the presence of chloroquine resistance in Edea (Gazin et al. 1990) can explain this last result; nevertheless, tests of homo- geneity of the sample, according to prophylaxis habits, were performed later in segregation analysis. The effects that both the area of residence and age (categorized into nine classes) had on MPD1 values were significant ($P < .0001$ for both), and no interaction between these two factors was observed ($P > .21$). The effect of the area of residence, accounting for 6% of the MPD1 variability, was easily explained by geographical differences in transmission conditions. The highest infection

levels were observed in subdistrict 3, which is located near a swamp area infested by mosquitoes and which had an MPD1 mean of .41 (95% confidence interval .16 to .66), whereas this mean was .03 (− .19 to .25) and − .22 (− .38 to − .06) in subdistricts 2 and 1, respectively. The MPD1s were corrected for this factor, as described in Methods, to constitute the MPD2 values. The standardized MPD2 values were the phenotypes used for segregation analysis by means of the regressive model. Figure 3 shows the mean MPD2 values observed in the nine age classes. The infection levels were the highest among children (9–12 years) and then decreased with age, consistent with the development of an acquired immunity. Polynomial regression in age was performed before segregation analysis by the unified mixed model. The final regression equation was a function of age ($P < .02$), age^2 ($P < .002$), and age^3 ($P < .002$), with the coefficients MPD2 = − .03 + .072 age − .0041age^2 + .000047age^3 and explained 15% of the variance of the MPD2s. The MPD2 values predicted by this regression equation are shown in figure 3. The distribution of the standardized residuals, i.e., the MPD3's, is presented in figure 4. Figure 2 shows the standardized MPD1, MPD2, and MPD3 values for four nuclear families.

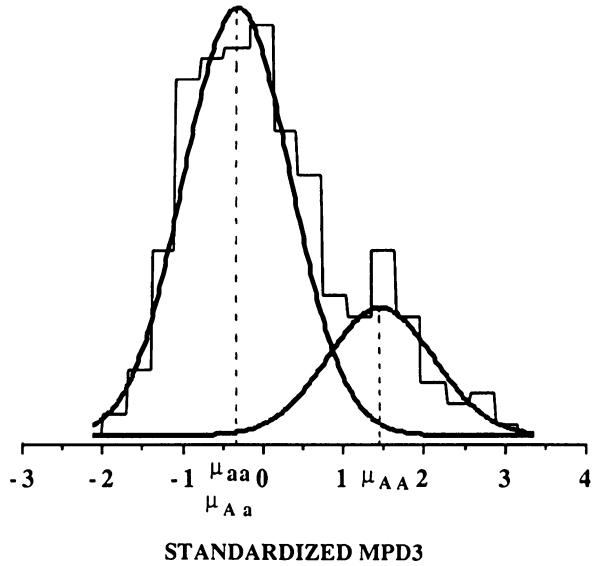

Figure 4 Histogram of the standardized MPD3s in 285 persons. The curves represent the predicted distribution of the MPD3, under the recessive major gene model (model 4b in table 2) and under the assumption of a normal distribution for each genotype (for a recessive model, aa and Aa individuals have the same distribution). The estimated variance within each component is .43, and the proportion of AA individuals, according to the estimated gene frequency, is .23.

Evidence for a Mixture of Distributions

The likelihood-ratio tests showed that the most parsimonious hypothesis for the MPD3 data was a mixture of two normal distributions with no residual skewness. This conclusion for a bimodal distribution is compatible with but is not evidence of a major-gene effect for which familial dependence of the phenotypes, as performed by segregation analysis, has to be taken into account.

Segregation Analysis

Results of segregation analysis obtained with the unified model are presented in table 2. The presence of a recessive major gene (model 5a vs. model 3) was highly significant ($\chi^2_{2\,dfs} = 28.9$, $P < .001$), and addition of a residual multifactorial component to this major gene (model 4b vs. model 3) was not needed ($\chi^2_2 = .1$, $P > .9$). With respect to the degree of dominance of the major gene, the recessive hypothesis was not rejected when compared with the codominant one (model 4b vs. model 4a); however, the test was borderline significant ($\chi^2_1 = 3.7$, $P = .07$). A dominant major gene did not fit the data (model 4c vs. model 4a). The Mendelian transmission of the recessive major effect was compatible with the data whether H and HZ were fixed at 0 (model 4b vs. model 1b; $\chi^2_3 = 4.6$, $P > .19$) or estimated (model 3 vs. model 1a; $\chi^2_3 = 4.7$, $P > .18$). Furthermore, the hypothesis of nontransmission was excluded whether H or HZ were fixed at 0 (model 2b vs. model 1b; $\chi^2_3 = 19.4$, $P < .001$) or estimated (model 2a vs. model 1a; $\chi^2_3 = 15.3$, $P < .01$). Although not shown in table 2, the same results were observed with a codominant major gene. The curves of figure 4 show the predicted distribution of the MPD3s under the recessive model (model 4b); and in figure 2 are indicated the most likely genotypes of family members, according to their MPD3 values.

Table 3 shows the results of segregation analysis using the class D regressive model. For all models, likelihood-ratio tests showed that the best fit was obtained with a cubic polynomial regression in age needing three more parameters, denoted "β_{age}," "β_{age^2}," and "β_{age^3}" in table 3. Model 3b ($\rho_{FM} = 0$, $\rho_{PO} = \rho_{SS}$) is equivalent to model 5b of table 2, for the familial correlations, and the values of ρ_{PO} or ρ_{SS} (.159) estimated simultaneously with the age effect were very close to $H/2$ (.161). The presence of a recessive major gene (model 3a vs. model 1b) was again highly significant ($\chi^2_2 = 42.3$, $P < .001$), and addition of residual familial correlations to this major gene (model 2b

Table 2

Segregation Analysis of MPD3s by Using the Unified Mixed Model

Model	Parameter[a]										
	q	μ_{aa}	μ_{Aa}	μ_{AA}	τ_{AAA}	τ_{AaA}	τ_{aaA}	H	HZ	σ^{2b}	$-2\ \ln L + c$
1. General transmission of recessive major effect (free τ's):											
a. Mixed	.47	−.40	−.40	1.42	.92	.59	.12	.02	.02	.42	752.7
b. $H = HZ = 0$	.47	−.40	−.40	1.43	.93	.59	.11	(0)	(0)	.42	752.9
2. No transmission of recessive major effect (equal τ's):											
a. Mixed	.49	−.40	−.40	1.29	.49	.49	.49	.14	.14	.48	768.0
b. $H = HZ = 0$	.47	−.41	−.41	1.40	.47	.47	.47	(0)	(0)	.43	772.3
3. Mixed Mendelian recessive	.48	−.41	−.41	1.38	(1)	(.5)	(0)	.01	.01	.44	757.5
4. Mendelian, $H = HZ = 0$:											
a. Codominant	.46	−.67	−.23	1.46	(1)	(.5)	(0)	(0)	(0)	.40	754.0
b. Recessive	.48	−.41	−.41	1.38	(1)	(.5)	(0)	(0)	(0)	.43	757.6
c. Dominant	.14	−.39	1.07	1.07	(1)	(.5)	(0)	(0)	(0)	.58	773.2
5. Multifactorial (equal μ's, $q = 0$):											
a. H and HZ	(0)	(0)[c]	(0)	(0)	. . .	. . .	. . .	.32	.33	(1)[c]	786.4
b. $H = HZ$	(0)	(0)	(0)	(0)	. . .	. . .	. . .	.32	.32	(1)	786.4

[a] Values in parentheses are fixed.
[b] Variance residual from the major effect.
[c] For all models, the overall mean and variance of the MPD3s were fixed at 0 and 1, respectively.

vs. model 1b) was not necessary ($\chi^2_3 = 1.9$, $P > .5$). The recessive hypothesis was not rejected when compared with the codominant one, whether the residual familial correlations were estimated (model 1b vs. model 1a; $\chi^2_1 = 2.4$, $P > .12$) or not (model 2b vs. model 2a; $\chi^2_1 = .5$, $P > .45$). However, problems of convergence occurred when the transmission probabilities and the β parameters of the polynomial regression in age were estimated simultaneously, and tests of transmission of the major effect could not be performed. It can be noted that the estimates of the β parameters under recessive model 2b are close to the ones obtained in the previous paragraph under the hypothesis of independence between individuals, and figure 3 shows the evolution of the μ_g's with age, as predicted by model 2b.

The class D regressive model was also used to perform segregation analysis on the MPD3 values. In this case no additional covariate was needed in the analysis model. The conclusion of the analysis was the same as that of the analysis conducted with the unified mixed model—i.e., the presence of a recessive major gene with no residual familial correlations. The parameter estimates were identical to those of model 4b in table 2, since in the absence of both familial correlations and covariates, the regressive and unified mixed models are equivalent.

Test of Homogeneity of the Sample, According to Prophylaxis Habits

To assess the influence of prophylaxis habits on our conclusions, we performed segregation analysis, with the unified mixed model, separately in two subsets of families: families (a), who did not take any prophylaxis (15 families), and families (b + c), who had irregular or regular prophylaxis intake (27 families). Under the recessive-major-gene model, a test of homogeneity done by calculating twice the difference between the likelihood of the overall data and the summed likelihoods of the two subsamples clearly showed no heterogeneity ($\chi^2_2 = 0.1$, $P > .9$). Furthermore, in both subsamples, separate segregation analyses showed evidence for the presence of a Mendelian recessive major gene with parameter estimates close to the ones obtained in the overall data, indicating that prophylaxis habits do not influence the conclusions of the analysis.

Discussion

The results provide clear evidence for the segregation of a recessive major gene controlling the levels of infection in human malaria. Under the recessive-major-gene model (model 4b in table 2), about 23% of the population is predisposed to high infections and 77% is resistant. With respect to the degree of domi-

Table 3

Segregation Analysis of MPD2s by Using the Regressive Model

Model	Parameter[a]														
	q	μ_{aa}	μ_{Aa}	μ_{AA}	τ_{AAA}	τ_{AaA}	τ_{aaA}	ρ_{FM}	ρ_{PO}	ρ_{SS}	σ^{2}[b]	β_{age}	β_{age^2}	β_{age^3}	$-2\ \ln L + c$
1. Mendelian with ρ_{FM}, ρ_{PO}, and ρ_{SS}:															
a. Codominant	.45	−.67	−.37	1.15	(1)	(.5)	(0)	.05	−.13	.00[c]	.30	.066	−.0036	.000042	694.9
b. Recessive	.45	−.52	−.52	1.16	(1)	(.5)	(0)	−.06	−.10	.00[c]	.32	.066	−.0036	.000043	697.3
2. Mendelian, ρ_{FM} = ρ_{PO} = ρ_{SS} = 0:															
a. Codominant	.44	−.60	−.44	1.24	(1)	(.5)	(0)	(0)	(0)	(0)	.32	.065	−.0036	.000042	698.7
b. Recessive	.44	−.51	−.51	1.23	(1)	(.5)	(0)	(0)	(0)	(0)	.32	.064	−.0035	.000042	699.2
c. Dominant	.13	−.43	.91	.91	(1)	(.5)	(0)	(0)	(0)	(0)	.49	.073	−.0040	.000046	726.0
3. Familial correlations (equal μ's, q = 0):															
a. ρ_{FM}, ρ_{PO}, and ρ_{SS}	(0)	−.09	−.09	−.09	. . .	. . .	. . .	.64	.22	.17	.89	.079	−.0044	.000050	734.9
b. ρ_{FM} = 0, ρ_{PO} = ρ_{SS} ...	(0)	−.06	−.06	−.06	. . .	. . .	. . .	(0)	.16	.16	.83	.074	−.0042	.000048	739.6

[a] Values in parentheses are fixed.

[b] Variance residual from both the age and the major-gene effects.

[c] Parameter is set to a bound.

nance of the major gene, the results obtained using the unified mixed model cannot exclude a codominant major gene. Under this hypothesis (model 4a in table 2), the susceptibility level of heterozygotes (50% of the population) remains much closer to that of resistant homozygotes (29%) than to that of susceptible homozygotes (21%).

Although segregation analysis of the MPD2s by means of the class D regressive model could not be entirely completed, it is important to note that two different strategies of age correction yielded close results. Furthermore, the shape of the regression function between the MPD2s and age was similar, whether a major gene was included in the analysis (models 2 and 3 in table 3) or not (model 1 in table 3), indicating that there is probably no interaction between the major gene and the age effect. To strengthen our findings we repeated the analysis with the unified mixed model, using an alternative mode of age correction. As different patterns of commingling in parents and children could lead to conflicting results (Rice et al. 1990), both the adjustment on age and the commingling analysis were carried out separately in parents and in offspring. There was evidence for a bimodal distribution in parents and in children, and the conclusions of segregation analysis were the same as previously, i.e., the presence of a recessive major gene, with close parameter estimates ($q = .53$, $\mu_{aa} = \mu_{Aa} = -.45$, and $\mu_{AA} = 1.15$).

The presence of a major gene controlling infection intensities in human malaria is consistent with experimental findings in mice, which show the role of a non–H2-linked major gene controlling resistance to *Plasmodium chabaudi* (Stevenson et al. 1982; Stevenson and Skamene 1985), and studies of leprosy indicate that parallels concerning the genetic control of susceptibility to infectious diseases between mice and humans could be drawn (Abel and Demenais 1988; Schurr et al. 1989). However, other H-2 and non–H-2-linked genetic factors have been shown to be involved in the susceptibility/resistance of mice to various species of *Plasmodium* (Sayles and Wassom 1988; Wunderlich et al. 1988), and one might expect a similar complexity in human genetic control. In this case, the isolation of the effect of a single gene is of major interest and provides an impetus for locating this gene by linkage analysis. Further linkage studies will determine the relationships between the red cell genetic defects known to influence the resistance against malaria infection (Weatherall 1987; Nagel and Roth 1989) and the HLA system with the major gene that we have detected.

Protective immunity as induced by injection of irradiated *P. yoelii* sporozoites have also been shown to be genetically controlled by both non-H-2 and H-2 genes (Weiss et al. 1989). This genetic restriction of the host response to immunization is one of the major obstacles to developing an effective malaria vaccine (Grau et al. 1987; Good et al. 1988; Weiss et al. 1989; Londono et al. 1990). Our results showing the role of a major gene in the determination of blood infection levels suggest that protective immunity in human malaria could be genetically controlled also. This conclusion is the first step toward identifying the genetic mechanisms of the human immune responsiveness to malaria infection and may have important implications for the control of this disease.

Acknowledgments

We thank V. Robert, G. Le Goff, and E. Mooh, Department of Medical Entomology (OCEAC), for assistance, and we thank A. Mallet, INSERM U194, and A. Dessein, INSERM U136, for helpful comments on the manuscript.

References

Abel L, Demenais F (1988) Detection of major genes for susceptibility to leprosy and its subtypes in a Caribbean island: Desirade Island. Am J Hum Genet 42:256–266

Abel L, Demenais F, Prata A, Souza AE, Dessein A (1991) Evidence for the segregation of a major gene in human susceptibility/resistance to infection by *Schistosoma mansoni*. Am J Hum Genet 48:959–970

Bonney GE (1984) On the statistical determination of major gene mechanisms in continuous human traits: regressive models. Am J Med Genet 18:731–749

——— (1986) Regressive logistic models for familial diseases and other binary traits. Biometrics 42:611–625

Demenais F, Bonney GE (1989) Equivalence of the mixed and regressive models for genetic analysis. I. Continuous traits. Genet Epidemiol 6:597–617

Demenais F, Lathrop M, Lalouel JM (1986) Robustness and power of the unified model in the analysis of quantitative measurements. Am J Hum Genet 38:228–234

Elston RC, Bailey-Wilson JE, Bonney GE, Keats BJ, Wilson AF (1986) S.A.G.E.: a package of computer programs to perform statistical analysis in genetic epidemiology. Paper presented at the Seventh International Congress of Human Genetics, Berlin, September 22–26

Gazin P, Louis JP, Mulder L, Eberle F, Jambou R, Moyroud

R, Hengy C (1990) *In vivo* monitoring of chloroquine sensitivity of *Plasmodium falciparum* in Edea, Sanaga maritime district, south-west Cameroon. Med Trop 50: 27–31

Good MF, Miller LH, Kumar S, Quakyi IA, Keister D, Adams JH, Moss B, et al (1988) Limited immunological recognition of critical malaria vaccine candidate antigens. Science 242:574–577

Grau GE, Del Giudice G, Lambert PH (1987) Host immune response and pathological expression in malaria: possible implications for malaria vaccines. Parasitol. 94:8123–8137

Hill AVS, Allsopp CEM, Kwiatkowski D, Anstey NM, Twumasi P, Rowe PA, Bennett S, et al (1991) Common west African HLA antigens are associated with protection from severe malaria. Nature 352:595–600

Lalouel JM, Morton NE (1981) Complex segregation analysis with pointers. Hum Hered 31:312–321

Lalouel JM, Rao DC, Morton NE, Elston RC (1983) A unified model for complex segregation analysis. Am J Hum Genet 35:816–826

Londono JA, Gras-Masse H, Dubeaux C, Tartar A, Druilhe P (1990) Secondary structure and immunogenicity of hybrid synthetic peptides derived from two *Plasmodium falciparum* pre-erythrocytic antigens. J Immunol 145:1557–1563

MacLean CJ, Morton NE, Elston RC, Yee S (1976) Skewness in commingled distributions. Biometrics 32:694–699

Miller LH. Howard RJ, Carter R, Good MF, Nussenzweig V, Nussenzweig RS (1986) Research toward malaria vaccines. Science 234:1349–1356

Morton NE, Rao DC, Lalouel JM (1983) Methods in genetic epidemiology. Karger, Basel

Nagel RL, Roth EF (1989) Malaria and red cell genetic defects. Blood 74:1213–1221

Osoba D, Dick HM, Voller A, Gooson TJ, Gooson T, Draper CC, de The G (1979) Role of the HLA complex in the antibody response to malaria under natural conditions. Immunogenetics 8:323–338

Rice T, Bouchard C, Borecki IB, Rao DC (1990) Commingling and segregation analysis of blood pressure in a French-Canadian population. Am J Hum Genet 46:37–44

Sayles PC, Wassom DL (1988) Immunoregulation in murine malaria. J Immunol 141:241–248

Schurr E, Skamene E, Forget A, Gros P (1989) Linkage analysis of the Bcg gene on mouse chromosome 1: identification of a tightly linked marker. J Immunol 142:4507–4513

Stevenson MM, Lyanga JJ, Skamene E (1982) Murine malaria: genetic control of resistance to *Plasmodium chabaudi*. Infect Immun 38:80–88

Stevenson MM, Skamene E (1985) Murine malaria: resistance of AXB/BXA recombinant inbred mice to *Plasmodium chabaudi*. Infect Immun 47:452–456

Trape JF (1985) Rapid evaluation of malaria parasite density and standardization of thick smear examination for epidemiological investigations. Trans R Soc Trop Med Hyg 79: 181–184

Weatherall DJ (1987) Common genetic disorders of the red cell and the 'malaria hypothesis.' Ann Trop Med Parasitol 81:539–548

Weiss WR, Good MF, Hollingdale MR, Miller LH. Berzofsky JA (1989) Genetic control of immunity to *Plasmodium yoelii* sporozoites. J Immunol 143:4263–4266

Wunderlich F, Mossman H, Helwig M, Schillinger G (1988) Resistance to *Plasmodium chabaudi* in B10 mice: influence of the H-2 complex and testosterone. Infect Immun 56:2400–2406