Am. J. Hum. Genet. 48:959–970, 1991

Evidence for the Segregation of a Major Gene in Human Susceptibility/Resistance to Infection by *Schistosoma mansoni*

Laurent Abel,*'† Florence Demenais,† Aluizio Prata,‡ Antonio E. Souza,‡ and Alain Dessein§

*INSERM U.155, Laboratoire d'Anthropologie Biologique, Université Paris VII, Paris; †Division of Epidemiology and Biostatistics, Howard University Cancer Center, Washington, DC,‡ Faculdade de Medicina do Triangulo Mineiro, Uberaba, Brazil; and Centro de Saude de Caatinga do Moura, Jacobina, Brazil; and §Centre d'Immunologie de Marseille-Luminy, Marseille

Summary

Severe clinical disease caused by the major human parasite *Schistosoma mansoni* is the consequence of high and prolonged infections. Epidemiological studies indicate that, for individuals having frequent contacts with cercaria-infested waters, both infection intensities and reinfection after treatment depend, in large part, on their intrinsic susceptibility/resistance to infection, suggesting the role of genetic factors in human resistance to *S. mansoni*. To investigate whether a major gene controls human susceptibility/resistance to infection by *S. mansoni*, segregation analysis of infection intensities, adjusted for the factors relevant in schistosomiasis (water contact, age, sex), was performed on 20 Brazilian pedigrees (269 individuals), using both the unified mixed model and the regressive model of analysis. The results are consistent with the hypothesis that there is a codominant major gene controlling human susceptibility/resistance to infection by *S. mansoni*. Parameter estimates indicate a frequency of .20–.25 for the deleterious allele; thus, about 5% of the population is predisposed to high infections, 60% is resistant, and 35% has an intermediate, although fairly good, level of resistance. These findings provide a genetic basis for earlier observations on the lower resistance and the predisposition to reinfection of certain individuals. In addition to the detection of a major gene effect, the data suggest that immunity to *S. mansoni* develops progressively during childhood to reach a maximum around the age of puberty. The implications of these results for the strategy to be used in endemic areas to reduce morbidity and to control parasite transmission are discussed.

Introduction

The profound influence of the genetic makeup of the host on resistance to viral and bacterial infections has been established in numerous animal studies. These results provided important information for the analysis of genetic factors involved in human resistance to these infections (see Skamene [1985] for a review). Similar studies on animals infected with parasites led to the identification of several genes involved in susceptibility to most parasites (see Skamene [1985] for a review), including schistosomes (James and Sher 1983; Sher et al. 1984; Correa-Oliveira et al. 1986, 1988; Wright et al. 1988). However, few genetic studies have been carried out on human resistance to parasites, and, to our knowledge, malaria is the only parasitic disease in which the role of genetic factors in resistance to infection has been clearly established (Mims 1982).

Schistosomiasis affects more than 200 million people throughout the world, according to estimations of the World Health Organization (UNDF/World Bank/WHO 1985), and remains a major problem of public health in many developing countries. Infection by *Schistosoma mansoni* is initiated by the penetration of infective larvae through the skin of subjects in contact with infested waters. Within a few days, the larvae migrate into the portal and mesenteric veins, where they mature into either male or egg-laying female adult worms. Though most eggs are laid in the mesenteric

Received July 5, 1990; final revision received December 28, 1990.

Address for correspondence and reprints: Dr. Laurent Abel, INSERM U.155, Laboratoire d'Anthropologie Biologique, Université Paris VII, Tour 16, 2 Place Jussieu, 75005 Paris, France.

0002-9297/91/4805-0015$02.00

veins and pass into intestinal lumen, a number of them are taken by the portal blood flow to the liver, where they stop in small vessels. Hepatosplenomegaly and portal hypertension, which are the major manifestations of disease due to *S. mansoni,* are the consequence of the inflammatory reactions triggered by eggs trapped in liver small vessels (Warren et al. 1967; Warren 1976) and occur more frequently in subjects with high infections. Earlier observations on subjects living in areas where *S. mansoni* are endemic have shown that infection intensities, as determined by fecal egg counts, in individuals having frequent contact with infested waters depend, in a large part, on their intrinsic susceptibility/resistance to infection (Katz et al. 1978; Butterworth et al. 1985; Dessein et al. 1988), and that some of them are predisposed to rapid and severe reinfection after parasitological cure with oxamniquine (Dessein et al. 1988; Tingley et al. 1988). Preliminary analysis of epidemiological data also indicated that the highest infection levels were aggregated within certain families rather than randomly distributed (A. J. Dessein, unpublished data). These observations suggested that human susceptibility/resistance to infection by *S. mansoni* might be genetically determined, but this possibility has not been evaluated so far.

The aim of the present study is to investigate the presence of a major gene determining human susceptibility to infection by *S. mansoni* and to do so by using complex segregation analysis. Segregation analysis is the first step to determine, from family data, the mode of inheritance of a given trait, with the main goal of identifying the role of a major gene. This study was performed on 20 Brazilian pedigrees randomly collected and using two different models of analysis, the unified mixed model (Lalouel et al. 1983) and the regressive model (Bonney 1984).

Material and Methods

Family Data and Measures

A family study was undertaken as part of a survey started in 1982 in Caatinga do Moura, a village of Bahia, a northeastern state of Brazil; details of this survey have been described elsewhere (Dessein et al. 1988). The village is divided into two banks by a river, which is the principal source of water for domestic uses. The ascertainment scheme was complete selection of the persons living within a certain geographic area, and all the families located on the left bank (about one-third of the whole population) were enrolled for this study. These families comprised twenty pedigrees consisting of 269 subjects living within 200 m of the infested river and were homogeneous with respect to socioeconomic factors. None of the collected subjects had received any schistosomicidal drug for the past 5 years.

Levels of water contact (W) were assessed as described elsewhere (Dessein et al. 1988), over a period of 18–24 mo for the majority of the study persons, and over a shorter 6-mo period for a few individuals, mostly parents, who entered later in the study. Two kinds of water-related activities were considered to assess W: (1) domestic and agricultural activities, quantified in hours per day or per week, such as washing dishes, clothes, hair, and working in irrigation channels or in flooded fields; (2) baths in the river, which imply total immersion and major risk of infection, quantified in number of baths per week. The subjects were then divided into four groups of W and coded as follows: 1, less than 1 h/wk in the water and no bath (N = 48); 2, 1–3 h once or twice a week and no bath (N = 70); 3, 1–3 h/d and less than three baths per week in the river (N = 88); and 4, same as group 3 with three or more baths per week (N = 63).

Infection intensities were measured by the individual fecal egg counts, expressed in eggs per gram, that were the arithmetic mean of at least five determinations performed by Kato's method (Martin and Beaver 1968) on stool samples collected on different days.

Data Adjustment

Prior to segregation analysis, the raw phenotypic values were adjusted for the relevant factors known to influence the intensity of infection by *Schistosoma mansoni* (water contact, age, and sex). Since a significant sex effect was observed, polynomial regressions of eggs-per-gram values on W and age in years (Y) were performed separately for males and females, with only the terms significant at the 5% level being retained. The standardized residuals were then used for further analyses.

Test for a Mixture of Distributions

Evidence for a mixture of normal distributions accounting for the adjusted data 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 up to three normal distributions was tested, while correcting for residual skewness by means of a classical power transformation suggested

by MacLean et al. (1976). Maximum-likelihood estimates of the relevant parameters and tests of various hypotheses using the likelihood-ratio criterion (Morton et al. 1983) were carried out by means of the computer program SKUMIX (MacLean et al. 1976; Morton et al. 1983).

Segregation Analysis

In the method described above, all individuals are considered independent, and evidence for a mixture of distributions is not sufficient to demonstrate the presence of a major gene. Detection of a major gene effect requires that the familial dependence of the phenotypes be accounted for. This is performed by segregation analysis that tests whether the observed familial distributions of a trait fit Mendelian expectations computed using mathematical models of transmission. Two different models were used in this study, the unified mixed model (Lalouel et al. 1983) and the regressive model (Bonney 1984).

The unified mixed model assumes that the observed phenotype (i.e., the intensity of infection) results from the additive and independent contributions of a major transmissible effect, a multifactorial transmissible component and a random, nontransmitted, environmental effect. Under the hypothesis that there is a major gene, the major effect results from segregation of two alleles (A,a) at a single locus. Thus, the parameters of the major gene effect are the frequency, q, of the susceptibility allele A, and the three means of the phenotype, u_{AA}, u_{Aa}, and u_{aa}, corresponding to the three genotypes AA, Aa, and aa. The distribution of the phenotype given the genotype is assumed to be normal with equal variances for all three phenotypic distributions. Multifactorial transmission, accounting for both polygenic and environmental effects common to the family, is parameterized in terms of H and HZ, the multifactorial heritability in children and adults respectively. Under the mixed model (including the major effect with H and HZ), evidence for a major effect is obtained by rejecting the hypothesis that there is no major effect. If no major effect is demonstrated, family resemblance is imputed to the multifactorial transmissible component. If there is evidence for a major effect, tests concerning its parent-offspring transmission are needed before concluding that the major effect is likely a major gene. Transmission at the major locus is parameterized in terms of $\tau_{A/AA}$, $\tau_{A/Aa}$, and $\tau_{A/aa}$, which denote the probabilities of transmitting allele A for genotypes AA, Aa, and aa, respectively. Mendelian transmission corresponds to $\tau_{A/AA} = 1$, $\tau_{A/Aa} = .05$, and $\tau_{A/aa} = 0$; no parent-offspring transmission of the major effect is represented by $\tau_{A/AA} = \tau_{A/Aa} = \tau_{A/aa}$. Both the nonrejection of the Mendelian hypothesis and the rejection of the hypothesis that the major effect is not transmitted are required to demonstrate that a major gene is present (Demenais et al. 1986). The computer program POINTER (Lalouel and Morton 1981; Morton et al. 1983) was used to obtain maximum-likelihood parameter estimates, and required that the 20 pedigrees be broken into their 44 constitutive nuclear families.

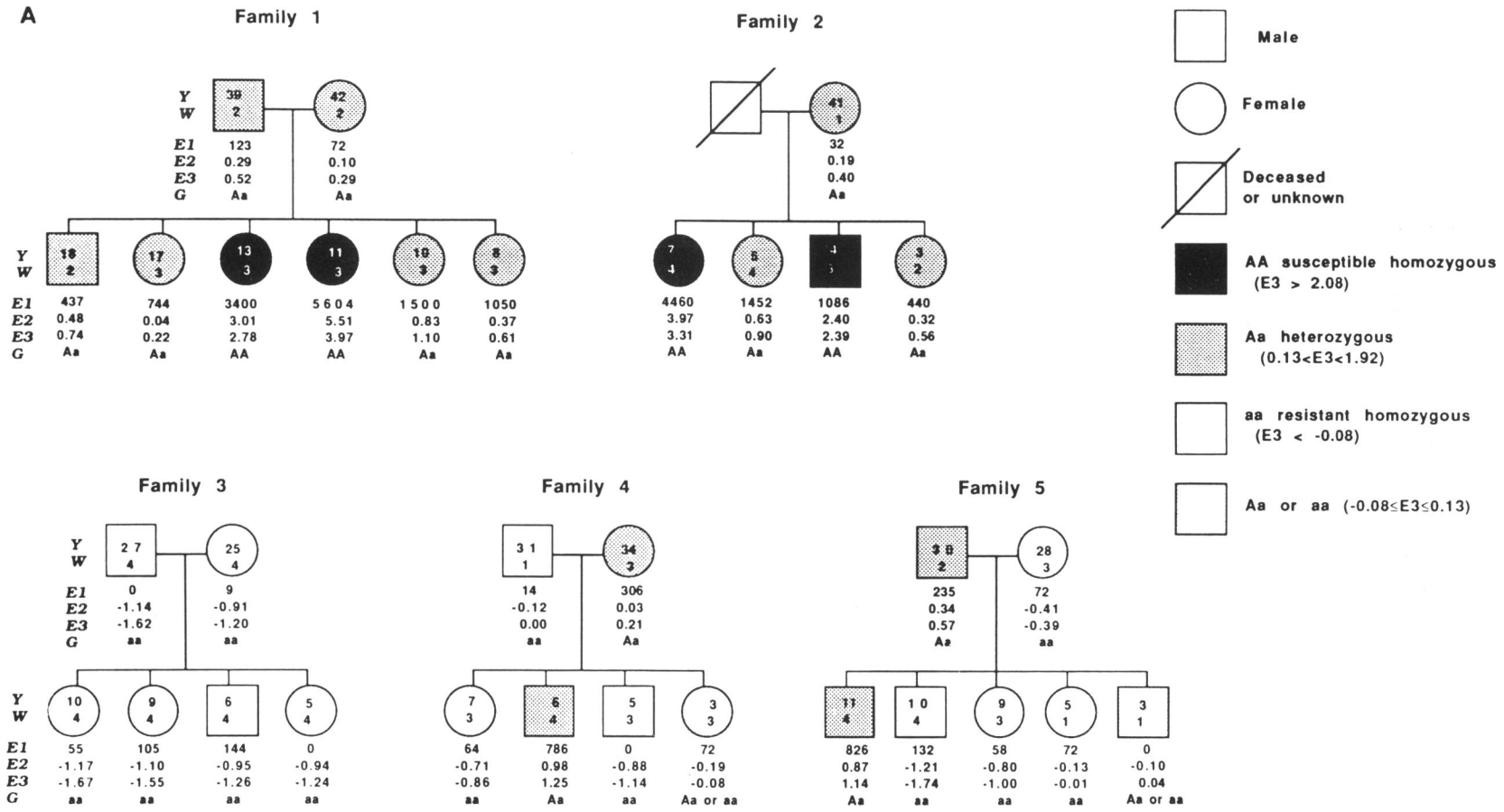
Regressive models are constructed by specifying a regression relationship between the phenotype of an individual and (1) a major gene effect, (2) the phenotypes of his preceding relatives, and (3) other explanatory variables or covariates (Bonney 1984). The major gene effect is specified in the same way as in the unified mixed model (with identical parameters). 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 A model (Bonney 1984; Bonney 1986), used in this analysis, these correlations are parametrized in terms of ρ_{FM}, the father-mother (or spouse) correlation, and ρ_{PO}, the parent-offspring correlation. When ρ_{FM} is set to zero, the equivalence between the regressive model and the unified mixed model has been established (Demenais and Bonney 1989). Since segregation analysis was performed on the adjusted data, no other covariate was introduced in the analysis model. Computations were carried out using the REGC program of the software package SAGE (Elston et al. 1986), which allows the pedigrees to be considered as a whole.

All hypotheses were tested by means of the likelihood-ratio criterion. As the families were randomly collected, there was no need for ascertainment correction. The different steps of the analysis (adjustment, power transformation, segregation analysis) are illustrated with the data of six families, presented in figure 1.

Results

Presentation of the Raw Data and Influence of W, Y, and Sex on Infection Intensities

The geometric mean, based on log(egg count + 1) to allow for zero counts, over the whole sample was 95.5 eggs/g. The influence of Y on infection levels is shown in figure 2*A*, and the heaviest infections were observed during adolescence for both sexes, as described elsewhere (Dessein et al. 1988; Tingley et al. 1988). A


A
Family 1
Family 2
Family 3
Family 4
Family 5
Male
Female
Deceased or unknown
AA susceptible homozygous (E3 > 2.08)
Aa heterozygous (0.13<E3<1.92)
aa resistant homozygous (E3 < -0.08)
Aa or aa (-0.08≤E3≤0.13)
Y
W
E1
E2
E3
G

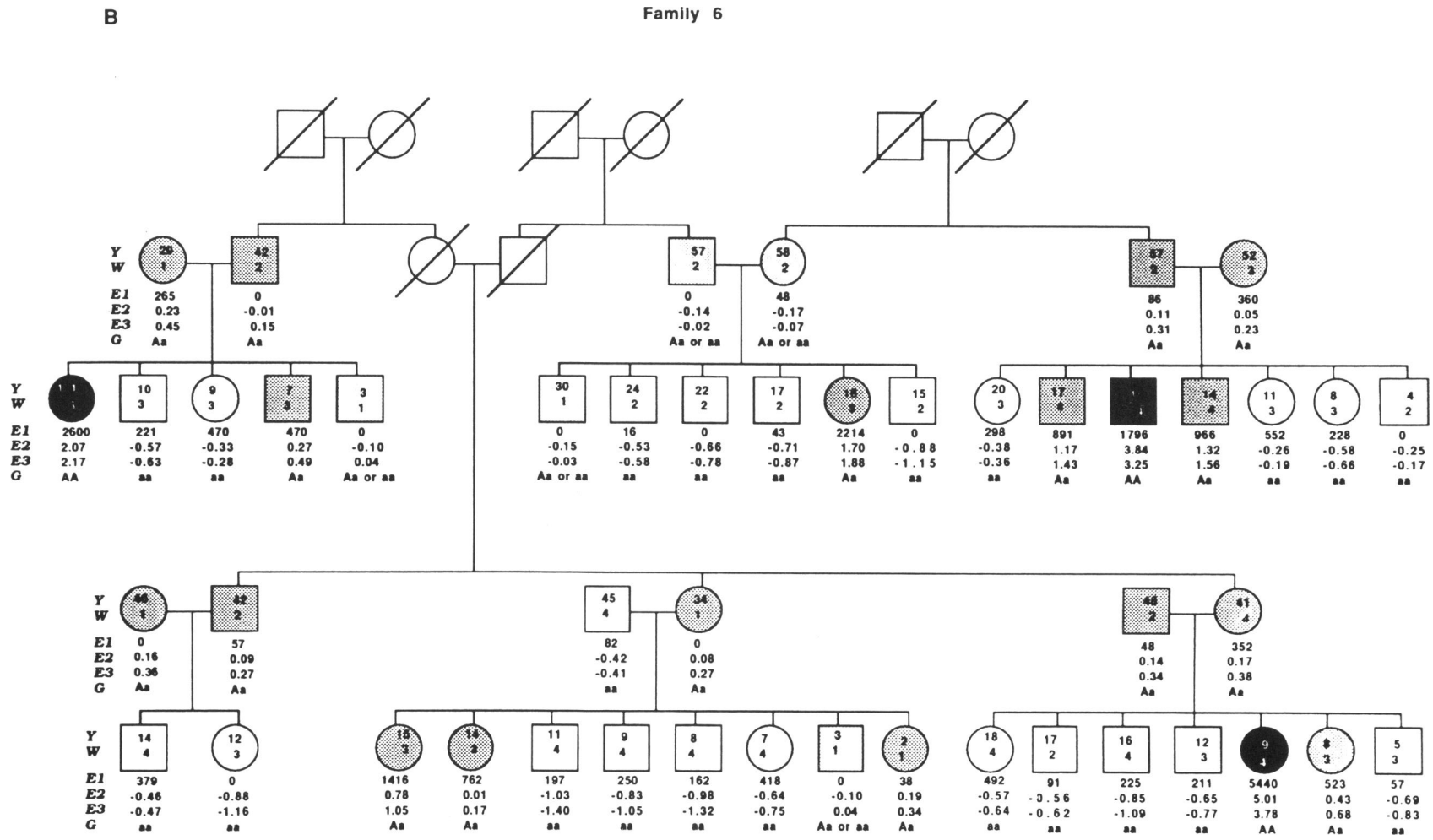


Figure 1 Five nuclear families (panel *A*) and one pedigree of the data (panel *B*) indicating for each person *Y*, *W*, raw eggs-per-gram values (E1), E2, E3, and most likely genotype (*G*) according to the E3 value under the hypothesis that there is a codominant major gene (A representing the deleterious allele). When the probability of a genotype is less than twice the probability of another genotype, both these genotypes are mentioned.

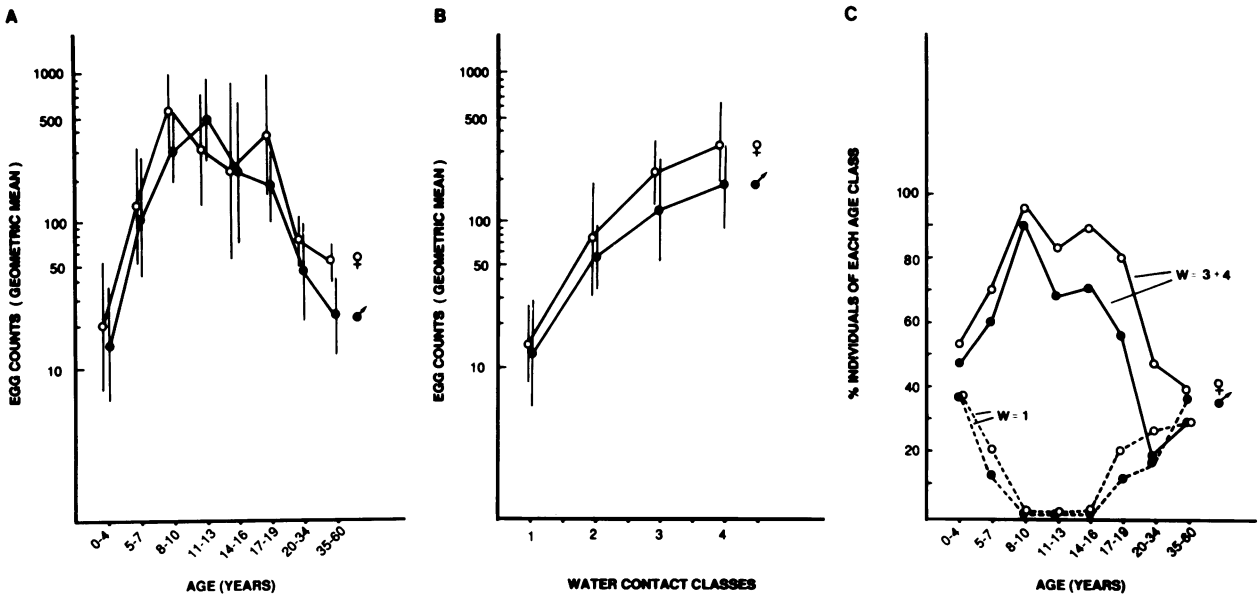


Figure 2 Presentation of raw data: influence on infection intensities of *Y* and *W*, according to sex. Results are represented as the geometric mean with 95% confidence interval of individual egg counts within (*A*) each *Y* group and (*B*) each *W* group. *C*, Modifications in water-related activities with age, according to sex. Results are presented as the proportion of individuals from *W* group 1, and both *W* groups 3 and 4, within each *Y* group.

clear increase of infection intensities with *W* was observed for males and females (fig. 2*B*). Using two-way analysis of variance, the effects of both *W* and sex on values of log(egg count + 1) were significant ($P < 10^{-4}$ and $P < .007$, respectively), and no interaction between these two factors was observed. Figure 2*C* shows clearly that major changes in water-related activities occurred with *Y*, the children from age 5 to 16 years being the most exposed to high *W*. However, these modifications in water-related activities cannot entirely explain the influence of *Y* on infection intensities (fig. 2*A*), since egg counts adjusted for *W* present persistent variations with *Y* (fig. 3*A*). Furthermore, the evidence for a significant residual effect of *Y* on infection levels while accounting for *W* is shown with the multivariate analysis performed for the final data adjustment and described in the following paragraph. It was also noted in figure 3*A* that, for each sex, the peak of infection is reached around puberty and, in subjects older than 8 years, the overall geometric mean of egg counts adjusted for *W* remained significantly higher in females than in males ($P < .03$). The global effect of *Y* in adjusted data on both *W* and sex is presented in figure 3*B*.

Data Adjustment

For the reasons mentioned above, polynomial regression of eggs-per-gram values on *Y* and *W* were carried out separately for males ($N = 121$) and females ($N = 148$). The final regression equation used for data adjustment was a function of W^2, $Y^{1/2}$, *Y*, and Y^2 for both sexes. For each sex, the residuals were divided by their respective SDs. The adjusted stan-

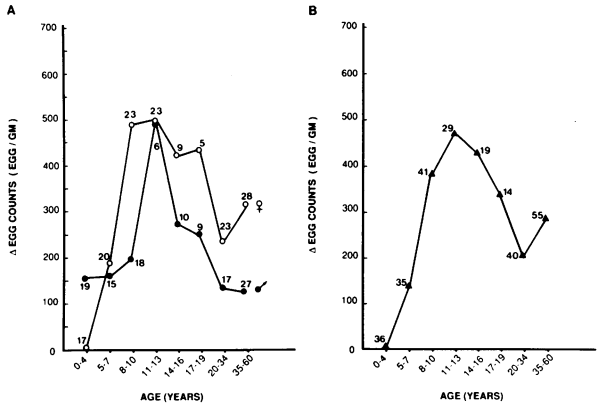


Figure 3 Influence of *Y* on infection intensities after adjustment for *W* and for both *W* and sex. Results are presented as the differences between groups in geometric means of individual egg counts adjusted for (*A*) *W* and (*B*) both *W* and sex, with the corresponding number of subjects in each *Y* group. The zero reference is taken as the value observed in 0–4 years in the (A) female group and (B) male-plus-female group.

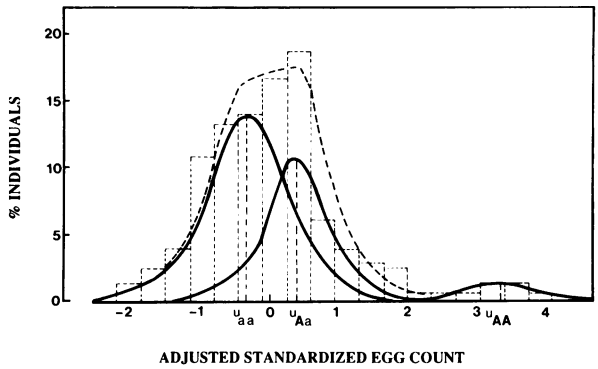


Figure 4 Histogram of the E3's in 269 persons. The curves represent the predicted distribution of the trait under a codominant major gene model (model 4 in table 2), assuming a normal distribution for each monogenotype. The estimated variance within each component is .47, and the proportions of aa, Aa, and AA individuals, according to the estimated gene frequency, are .60, .35, and .05, respectively.

dardized eggs-per-gram values of males and females (E2; fig. 1) were then combined for further analysis.

Test for a Mixture of Distributions

The skewness computed as the expectation of x^3, where x is E2, was high: 2.75 (this expectation is zero for a normal distribution). The likelihood-ratio tests showed that the most parsimonious hypothesis for the adjusted data was a mixture of two distributions in the presence of residual skewness. This result means that the high skewness observed in the adjusted data comes from (1) a mixture of two distributions and (2) the fact that those two distributions are skewed, needing a power transformation to restore normality. The adjusted data were then transformed, using the maximum-likelihood estimate of the power-transformation parameter p under the two-distribution hypothesis ($p = -0.78$) to provide the final phenotypes used for segregation analysis (the adjusted power-transformed standardized eggs-per-gram values are denoted by E3 in fig. 1). The residual skewness calculated after power transformation, and due to the mixture of the two normal distributions, was 1.70. The distribution of the E3's is shown in the histogram of figure 4. Nine subjects had values greater than 2.0; in two nuclear families (1 and 2 in fig. 1), two of these subjects were sibs.

Segregation Analysis

Results of segregation analysis obtained with the unified mixed model are presented in table 1. Under the mixed model (major gene with H and HZ), the presence of a codominant major gene (model V vs. model III) was highly significant ($\chi^2_3 = 45.2, P < 10^{-9}$). Addition of a residual multifactorial component to the codominant major gene (IV_1 vs. III) was not necessary ($\chi^2_2 = 4.2, P > .10$). Both the recessive (IV_2 vs. IV_1) and the dominant (IV_3 vs. IV_1) hypotheses for the major gene were rejected with $\chi^2_1 = 11.5$ ($P < .001$) and $\chi^2_1 = 32.7$ ($P < 10^{-6}$), respectively. The Mendelian

Table 1

Segregation Analysis of 44 Nuclear Families Using Unified Mixed Model

Model	Parameters										
	q	u_{aa}	u_{Aa}	u_{AA}	$\tau_{A/AA}$	$\tau_{A/Aa}$	$\tau_{A/aa}$	H	HZ	σ^2 [a]	$-2\ln L + c$ [b]
I. General transmission of codominant major effect (free τ's), H = HZ = 0	.18	−.40	.56	3.29	1.00[c]	.42	.05	(0)	(0)	.45	.00
II. No transmission of codominant major effect (equal τ's), H = HZ = 0	.19	−.12	−.12	3.10	.24	.24	.24	(0)	(0)	.64	11.7
III. Mixed Mendelian codominant	.20	−.31	.25	3.05	(1)	(.5)	(0)	.16	.04	.55	1.4
IV. Mendelian, H = HZ = 0:											
1. Codominant	.20	−.38	.40	3.11	(1)	(.5)	(0)	(0)	(0)	.49	5.6
2. Recessive	.20	−.13	−.13	2.97	(1)	(.5)	(0)	(0)	(0)	.63	17.1
3. Dominant	.05	−.19	1.58	1.58	(1)	(.5)	(0)	(0)	(0)	.70	38.3
V. Multifactorial (equal u's, q = 0)	(0)	(0)[d]	(0)	(0)	. . .	. . .	. . .	.33	.27	(1)[d]	46.6

NOTE.—Parameters in parentheses are fixed at the values shown.

[a] σ^2 represents the residual variance from the major effect.

[b] The values $-2\ln L + c$ are scaled by the transformed likelihood of the general model I. L = likelihood; c = 2 ln$L$ of the general model I.

[c] Parameter is set to a bound.

[d] For all models, the overall mean and variance of the phenotype were fixed at 0 and 1, respectively.

Table 2

Segregation Analysis of 20 Pedigrees Using Regressive Model

Model	Parameters										
	q	u_{aa}	u_{Aa}	u_{AA}	$\tau_{A/AA}$	$\tau_{A/Aa}$	$\tau_{A/aa}$	ρ_{FM}	ρ_{PO}	$\sigma^{2\,a}$	$-2\ \ln L + c^b$
I. General transmission of codominant major effect with ρ_{FM} and ρ_{PO}	.27	−.57	.30	2.45	1.00[c]	.43	.00[c]	.94	.16	.46	.00
II. No transmission of codominant major effect with ρ_{FM} and ρ_{PO}	.24	−.49	.55	3.16	.17	.17	.17	.88	.45	.47	10.6
III. Mendelian with ρ_{FM} and ρ_{PO}:											
1. Codominant	.25	−.59	.27	2.41	(1)	(.5)	(0)	.93	.17	.47	1.4
2. Recessive	.22	−.19	−.19	2.72	(1)	(.5)	(0)	.70	.19	.66	21.1
3. Dominant	.06	−.15	1.61	1.61	(1)	(.5)	(0)	.72	.20	.75	40.8
IV. Mendelian, $\rho_{FM} = \rho_{PO} = 0$ codominant	.22	−.43	.37	3.08	(1)	(.5)	(0)	(0)	(0)	.47	19.7
V. Familial correlations (equal u's, $q = 0$)	(0)	−.05	−.05	−.05	—	—	—	.80	.29	1.11	53.5

NOTE.—Parameters in parentheses are fixed at the values shown.

[a] σ^2 represents the residual variance from the major effect.

[b] The values $-2\ \ln L + c$ are scaled by the transformed likelihood of the general model I. L = likelihood; $c = 2\ \ln L$ of the general model I.

[c] Parameter is set to a bound.

transmission of the codominant major effect (IV_1 vs. I) was compatible with the data ($\chi^2_3 = 5.6$, $P > .10$), and its nontransmission (II vs. I) was rejected ($\chi^2_2 = 11.7$, $P < .005$). In conclusion, a codominant major gene alone accounted for the familial distribution of the phenotypes. Hypotheses were then tested using the second model of analysis.

Table 2 shows the results of segregation analysis using the regressive model. Without the major effect (model V), ρ_{FM} (.80) and ρ_{PO} = (.29) were both significant. The presence of the codominant major gene (V vs. III_1) was again highly significant ($\chi^2_3 = 52.1$, $P < 10^{-9}$). In the presence of the major gene, residual ρ_{FM} and ρ_{PO} were significant ($\chi^2_2 = 18.3$, $P < .001$). These residual correlations were mainly due to the ρ_{FM}, whereas the ρ_{PO} parameter was of borderline significance. In the presence of these familial correlations, both the recessive (III_2 vs. III_1, $\chi^2_1 = 19.7$, $P < 10^{-4}$) and the dominant (III_3 vs. III_1, $\chi^2_1 = 39.4$, $P < 10^{-9}$) hypotheses were rejected. In the presence of ρ_{FM} and ρ_{PO}, the Mendelian transmission of the codominant major gene (III_1 vs. I) was compatible with the data ($\chi^2_3 = 1.4$, $P > .50$), and the nontransmission hypothesis (II vs. I) was rejected ($\chi^2_2 = 10.6$, $P < .01$).

Parameter estimates of the major gene obtained using the two models were similar. The frequency of A was estimated between .20 and .25. The curves of figure 4 show the predicted distribution of the phenotype under the codominant major gene model, and figure 1 provides the most likely genotypes for the subjects of the presented families. The resistance level of heterozygotes is closer to that of resistant homozygotes than to that of susceptible homozygotes, implying that heterozygotes have an intermediate, although fairly good, level of resistance.

To our knowledge, no test of goodness of fit has been widely accepted for segregation analysis. However, to evaluate the fit of the codominant major gene model to our data, we computed the expected proportions of individuals in the five genotype classes defined in figure 1, that is aa (E3<−0.08), aa or Aa (−0.08<E3<0.13), Aa (0.13<E3<1.92), Aa or AA (1.92<E3<2.08), and AA (E3>2.08). The expected proportion of individuals having E3 values between bounds y and z is given by

$$P(y<E3<z) = \sum_g P(g)P(y<E3<z|g)\ ,$$

where $P(y<E3<z|g)$ is computed using cumulative normal distributions. As presented in table 3, the expected proportions are close to the observed ones. However, the parameters of the codominant major gene model used to compute the expected proportions were estimated from the same data as the ones used to determine the observed proportions; therefore, there is no simple procedure or test to compare the expected with the observed proportions.

Table 3

Expected and Observed Proportions in Five Genotype Classes under Hypothesis that There Is a Codominant Major Gene, and Observed Proportions among Parents and Offspring

	aa (E<.08)	aa or Aa (−.08<E3<0.13)	AA (.13<E3<1.92)	Aa or AA (1.92<E3<2.08)	AA (E3>2.08)
Expected.....	.496	.097	.349	.005	.051
Observed:					
Total ($N=269$).....	.468	.100	.390	.004	.034
Parents ($N=75$).....	.173	.227	.600	. . .	. . .
Offspring ($N=194$)...	.582	.051	.309	.005	.046

NOTE.—The genotype class of an individual is determined by his E3 as described in fig. 1.

Discussion

Several epidemiological observations made on residents of areas where *Schistosoma mansoni* are endemic suggested the role of genetic factors in human resistance to this parasite; first, reinfection intensities observed after parasitological cure with oxamniquine in adolescents with the same *W* still presented a large variability, indicating the role of some patient-intrinsic factor in resisting reinfection (Butterworth et al. 1985; Dessein et al. 1988); second, predisposition to rapid and severe reinfection had been observed in certain individuals (Dessein et al. 1988; Tingley et al. 1988); third, preliminary analysis of epidemiological data showed that high infection intensities were aggregated within certain families rather than randomly distributed in the population (A. J. Dessein, unpublished data). The results presented here confirm this view and are consistent with the hypothesis that there is a codominant major gene controlling human susceptibility/resistance to *S. mansoni*.

It is not entirely clear whether a residual familial correlation is needed in the model explaining the familial distributions of infection intensities. Using the unified mixed model, the residual multifactorial contribution (H, HZ) is marginally not significant, but we cannot exclude that it could become significant with a larger number of families. According to this latter hypothesis, both the presence of a major gene and adjustment for factors such as *W*, *Y*, and sex could not entirely account for the observed familial correlation of infection levels, and additional sources of resemblance, either genetic (other gene[s]) or environmental (unmeasured factors), would have to be considered. Using the regressive model, a strong residual ρ_{FM} was observed, and the interpretation of the ρ_{FM} parameter in this pedigree analysis needs some comments. A simple correlation coefficient computed between father and mother phenotype values, without taking into account the pedigree structure, was not significant, indicating that there is probably no actual dependence between the phenotypes of the spouses. The unexpected high spouse correlation obtained from analysis of pedigree data may in this case suggest some generation-specific effect. As presented in table 3, the observed distribution of genotypes in parents under the codominant major gene model is quite different from the one observed in the offspring generation. In particular, the absence of AA susceptible individuals among parents, probably due to mortality selection, may be responsible for a greater homogeneity of the trait in parents and contribute to the observed spouse correlation.

Another possible explanation for a generation effect is the *Y* adjustment, which has likely not taken entirely into account the complex *Y* effect observed in the data (fig. 2*A*). In this case, a significant ρ_{FM} may reflect the persistent differences in the trait due to *Y* between parents and children, as observed in simulated data (L. Abel and G. E. Bonney, unpublished data). The results presented in figure 1 also suggest that the adjustment did not accommodate very well the persons with low W ($W = 1$) and zero egg counts; these subjects are at the two extremes of the *Y* distribution, either children younger than 4 years who have not yet been in the river or adults older than 30 years who have no more contacts with water. The adjusted egg counts of these persons vary between −0.04 and 0.45, suggesting they are likely heterozygous, whereas they have no detectable eggs in the stool. To assess the influence of this problem on our results, we again performed a segregation analysis using the unified mixed

model and considering all the phenotypes of the 26 persons with $W = 1$ and zero egg counts as unknown. The conclusions are quite similar to the previous analysis, with strong evidence for a codominant major gene, and the estimates of q, u_{aa}, u_{Aa}, and u_{AA} were 0.20, −0.42, 0.48, and 2.99 respectively. This result indicates that the subjects with low W and zero egg counts provide little information for the analysis and, even if they are slightly misclassified, do not affect the conclusions of the study.

The main result of the study remains the fit of the data for a major gene model, and we are confident of the robustness of this conclusion for several reasons. First, the proportion of false conclusions that there is a major gene effect is minimum for the unified mixed model when, as observed in our study, the residual skewness is high (>0.4), the Mendelian hypothesis is not rejected, and the hypothesis that the major effect is not transmitted is rejected (Demenais et al. 1986); this observation also applies to the class A regressive model in which the proportion of spurious detections of a major gene has been shown to be not greater than the nominal 5% level when tests of transmission are performed properly (Demenais et al. 1990). Second, two models different in the way of accounting for familial dependence and handling families (the unified mixed model used nuclear families; the regressive model considered pedigrees as a whole) lead to similar conclusions (detection of a codominant major gene) and close estimates of the major gene parameters (allelic frequency and means of the phenotype). Third, our conclusions are consistent with previous findings in mice showing the role of a single gene(s) involved in the induction of protective immunity to *S. mansoni* (Correa-Oliveira et al. 1986, 1988; Wright et al. 1988), and studies of leprosy indicate that it might be possible to draw analogies between mice and humans regarding the genetic control of susceptibility to infectious diseases (Abel and Demenais 1988; Schurr et al. 1989). This work should be completed by linkage analysis to confirm and locate the detected major gene and by immunological studies to assess the effect of this gene on human immunity. Earlier work has associated increased susceptibility to infection with an impairment of the antibody response against a major larval antigen (Dessein et al. 1988), and ongoing studies will evaluate whether such a defect is related to the major gene detected in this study.

Though conflicting data exists on an association between HLA loci and hepatosplenomegaly caused by *S. mansoni* (Salam et al. 1979; Wakelin 1985), it is important to determine whether or not susceptibility to infection and susceptibility to hepatosplenomegaly are controlled by the same gene. Several observations, however, argue against this possibility. First, high infections are not necessarily associated with hepatosplenomegaly (Prata and Schroeder 1967). Second, H-2-linked genes were reported to regulate the immunopathology elicited by schistosomes and their eggs without influencing susceptibility to a primary infection (Claas and Deelder 1979); furthermore, mouse propensity to develop severe disease, organomegaly, and portal hypertension is not inherited in a simple Mendelian pattern (Fanning et al. 1981) and appears to depend on several genes linked to the major histocompatibility complex and outside that locus (Fanning et al. 1981; Jones et al. 1983). Third, hepatosplenomegaly is thought to result from a hyperimmune reactivity to adult worm and egg antigens (Warren et al. 1967; Warren 1976; Byram and Lichtenberg 1977; Chensue et al. 1980; Dessein et al. 1984), whereas high infections are associated with an impaired immune reactivity to the parasite (Dessein et al. 1988; Butterworth et al. 1988; Goudot-Crozel et al. 1989). This question could be solved by a linkage analysis between the gene detected in this work and the HLA system.

Finally, the conclusion that a fraction of the population living in areas where *S. mansoni* are endemic might be less resistant as a consequence of a major gene effect has several implications for the control of this endemic. It emphasizes the necessity of evaluating the response of less resistant individuals to candidate vaccines. Control methods should primarily target these children and adolescents, since they have a high risk of developing severe clinical disease and represent a major reservoir of parasites in the endemic area. The possibility should be evaluated that chemotherapy targeting this relatively small group of subjects could markedly reduce morbidity and parasite transmission. Such a control program would require easy and cheap methods to identify persons homozygous for the deleterious allele. It would have, however, several advantages over mass chemotherapy: it would be cheaper, would require fewer well-trained personnel, and would be less likely to select resistant strains of parasites.

In addition to the genetic results, this study shows that variations of infection intensities with Y cannot be fully accounted for by changes in water-related activities. Infection levels adjusted for W increased markedly between the time of first contact (around age

2 to 3 years) and puberty, and then declined during adolescence. This finding is consistent with the view that immunity to *S. mansoni* infection increases gradually during childhood and reaches maximal levels around puberty. This conclusion is in agreement with both immunological studies suggesting the role of blocking antibodies in limiting the expression of immunity in young children (Butterworth et al. 1987, 1988) and theoretical mathematical studies, based on the assumption of a gradual acquisition of immunity during childhood, predicting such a convex pattern of age-intensity curves in human helminth infections (Crombie and Anderson 1985; Jose 1989). Thus, an implication of this work is that the age-dependent development of immunity is a major constraint on infection intensity patterns in addition to the effects of ecological (i.e., water contact) and genetic factors.

Acknowledgments

We thank N. Feingold and J. Feingold for helpful discussion, G. E. Bonney for constructive comments, and L. Leserman for reading the manuscript. This work received financial assistance from the Rockefeller Foundation, INSERM, CNRS, and UNDP/World Bank/WHO Special Program for Research and Training in Tropical Diseases.

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

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 disease and other binary traits. Biometrics 42:611–625

Butterworth AE Bensted-Smith R, Capron A, Capron M, Dalton PR, Dunne DW, Grzych JM, et al (1987) Immunity in human schistosomiasis mansoni: prevention by blocking antibodies of the expression of immunity in young children. Parasitology 94:281–300

Butterworth AE, Capron M, Cordingley JS, Dalton PR, Dunne DW, Kariuki HC, Koech D, et al (1985) Immunity after treatment of human schistosomiasis mansoni. II. Identification of resistant individauls, and analysis of their immune responses. Trans R Soc Trop Med Hyg 79:393–408

Butterworth AE, Dunne DW, Fulford A, Capron M, Khalife J, Capron A, Koech D, et al (1988) Immunity in human schistosomiasis mansoni: cross reactive IgM and IgG2 anti-carbohydrate antibodies block the expression of immunity. Biochimie 70:1053–1059

Byram JE, Lichtenberg F (1977) Altered schistosome granuloma formation in nude mice. Am J Trop Med Hyg 26: 944–956

Chensue SW, Boros DL, David CS (1980) Regulation of granulomatous inflamation in murine schistosomiasis: in vitro characterization of T-lymphocyte subsets involved in the production and suppression of migration inhibition factor. J Exp Med 151:1398–1412

Claas FHS, Deelder AM (1979) H-2 linked immune response to murine experimental *Schistosoma mansoni* infections. J Immunogenet 6:167–175

Correa-Oliveira R, James SL, McCall D, Sher A (1986) Identification of a genetic locus, Rsm-1, controlling protective immunity against *Schistosoma mansoni.* J Immunol 137: 2014–2019

Correa-Oliveira R, James SL, Sher A (1988) Genetic complementation of defects in vaccine-induced immunity against *Schistosoma mansoni* in P- and A-strain inbred mice. Infect Immun 56:708–710

Crombie JA, Anderson RM (1985) Population dynamics of *Schistosoma mansoni* in mice repeatedly exposed to infection. Nature 315:491–493

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

Demenais F, Murigande C, Bonney GE (1990) A search for faster methods of fitting the regressive models to quantitative traits. Genet Epidemiol 7:319–334

Dessein AJ, Begley M, Demeure C, Caillol D, Fueri J, Galvao dos Reis M, Andrade ZA, et al (1988) Human resistance to *Schistosoma mansoni* is associated with IgG reactivity to a 37-kDa larval surface antigen. J Immunol 140:2727–2736

Dessein AJ, Lenzi HL, Bina JC, Carvalho EM, Weiser WY, Andrade ZA, David JR (1984) Modulation of eosinophil cytotoxicity by blood mononuclear cells from healthy subjects and patients with chronic schistosomiasis mansoni. Cell Immunol 85:100–113

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

Fanning MM, Peters PA, Davis RS, Kazura JW, Mahmoud AAF (1981) Immunopathology of murine infection with *Schistosoma mansoni:* relationship of genetic background to hepatosplenic disease and modulation. J Infect Dis 144: 148–153

Goudot-Crouzel V, Caillol D, Djabazi M, Dessein AJ (1989) The major parasite surface antigen associated with human resistance to schistosomiasis is a 37-kD glyceraldehyde-3P-deshydrogenase. J Exp Med 170:2065–2080

James SR, Sher A (1983) Mechanisms of protective immunity against *Schistosoma mansoni* infection in mice vaccinated with irradiated cercariae. III. Identification of a mouse strain, P/N, that fails to respond to vaccination. Parasite Immunol 5:567–575

Jones JT, McCaffery DM, Kusel JR (1983) The influence of the H-2 complex on responses to infection by *Schistosoma mansoni* in mice. Parasitology 86:19–30

Jose MV (1989) On the solution of mathematical models of herd immunity in human helminth infections. J Math Biol 27:707–715

Katz N, Zicker F, Rocha RS, Oliveira VB (1978) Reinfection of patients in schistosomiasis mansoni endemic areas after specific treatment. Rev Inst Med Trop Sao Paulo 20:273–278

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

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

Martin LK, Beaver PC (1968) Evaluation of Kato thick smear technique for quantitative diagnosis of helminth infection. Am J Trop Med Hyg 17:382–387

Mims CA (1982) Innate immunity to parasitic infections. In: Cohen S, Warren K (eds) Immunology of parasitic infections. Blackwell, Palo Alto, pp 3–27

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

Prata A, Schroeder S (1967) A comparison of whites and negroes infected with *Schistosoma mansoni* in a hyperendemic area. Gazeta Med Bahia 67:93–101

Salam EA, Ishaac S, Mahmoud AAF (1979) Histocompatibility-linked susceptibility for hepatosplenomegaly in human schistosomiasis mansoni. J Immunol 123:1829–1832

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

Sher A, Hieny S, James S (1984) Mechanisms of protective immunity against *Schistosoma mansoni* infection in mice vaccinated with irradiated cercariae. VI. Influence of the major histocompatibility complex. Parasite Immunol 6: 319–328

Skamene E (ed) Genetic control of host resistance to infection and malignancy. Liss, New York

Tingley GA, Butterworth AE, Anderson RM, Kariuki HC, Koech D, Mugambi M, Ouma JM, et al (1988) Predisposition of humans to infection with *Schistosoma mansoni:* evidence from the reinfection of individuals following chemotherapy. Trans R Soc Trop Med Hyg 82:448–452

UNDF/World Bank/WHO (1985) Tropical disease research: seventh program report. Barthelemy, Geneva

Wakelin D (1985) Genetic control of immunity to helminth infections. Parasitol Today 1:17–23

Warren KS (1976) Schistosomiasis: a multiplicity of immunopathology. J Invest Dermatol 47:464–469

Warren KS, Domingo EO, Cowan RBT (1967) Granuloma formation around schistosome eggs as a manifestation of delayed hypersensitivity. Am J Pathol 51:735–741

Wright MD, Tiu WU, Wood SM, Walker JC, Garcia EG, MitchelL GF (1988) *Schistosoma mansoni* and *S. japonicum* worm numbers in 129/J mice of two types and dominance of susceptibility in F1 hybrids. J Parasitol 74:618–622